New roles for arginine methylation in RNA metabolism and cancer

Isabelle Goulet

This thesis is submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the

Doctor of Philosophy degree
in Cellular and Molecular Medicine

September 2011

Department of Cellular and Molecular Medicine
Faculty of Medicine
University of Ottawa

AUTHORIZATIONS

Manuscript 1 (Chapter 2)
Alternative splicing yields protein arginine methyltransferase 1 isoforms with distinct activity, substrate specificity, and subcellular localization.
Journal of Biological Chemistry. 2007; 282: 33009-21.

Excerpt from the American Society for Biochemistry and Molecular Biology Website
< http://www.jbc.org/site/misc/Copyright_Permission.xhtml >:

Copyright Permission Policy

These guidelines apply to the reuse of articles, figures, charts and photos in the Journal of Biological Chemistry, Molecular & Cellular Proteomics and the Journal of Lipid Research.

For authors reusing their own material:

Authors need NOT contact the journal to obtain rights to reuse their own material. They are automatically granted permission to do the following:

- Reuse the article in print collections of their own writing.
- Present a work orally in its entirety.
- Use an article in a thesis and/or dissertation.
- Reproduce an article for use in the author's courses. (If the author is employed by an academic institution, that institution also may reproduce the article for teaching purposes.)
- Reuse a figure, photo and/or table in future commercial and noncommercial works.
- Post a copy of the paper in PDF that you submitted via BenchPress.

  Only authors who published their papers under the "Author's Choice" option may post the final edited PDFs created by the publisher to their own/departmental/university Web sites.

  All authors may link to the journal site containing the final edited PDFs created by the publisher.

Please note that authors must include the following citation when using material that appeared in an ASBMB journal:

"This research was originally published in Journal Name. Author(s). Title. Journal Name. Year; Vol:pp-pp. © the American Society for Biochemistry and Molecular Biology."

Manuscript 2 (Chapter 4)
TDRD3, a novel Tudor domain-containing protein, localizes to cytoplasmic stress granules. *Human Molecular Genetics* (United-Kingdom). 2008; 17: 3055-74.

Excerpt from the Oxford Journals Website
< http://www.oxfordjournals.org/access_purchase/publication_rights.htmll >:

Rights retained by ALL Oxford Journal Authors

- The right, after publication by Oxford Journals, to use all or part of the Article and abstract, for their own personal use, including their own classroom teaching purposes;
- The right, after publication by Oxford Journals, to use all or part of the Article and abstract, in the preparation of derivative works, extension of the article into book-length or in other works, provided that a full acknowledgement is made to the original publication in the journal;
- The right to include the article in full or in part in a thesis or dissertation, provided that this not published commercially.

For the uses specified here, please note that there is no need for you to apply for written permission from Oxford University Press in advance. Please go ahead with the use ensuring that a full acknowledgment is made to the original source of the material including the journal name, volume, issue, page numbers, year of publication, title of article and to Oxford University Press and/or the learned society.

Manuscript 3 (Appendix I)
Protein arginine methylation facilitates co-transcriptional recruitment of pre-mRNA splicing factors. *Molecular and Cellular Biology*. 2010; 30:5245-56.

AMERICAN SOCIETY FOR MICROBIOLOGY LICENSE
TERMS AND CONDITIONS

Apr 19, 2011

This is a License Agreement between Isabelle Goulet ("You") and American Society for Microbiology ("American Society for Microbiology") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by American Society for Microbiology, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	2652721199899
License date	Apr 19, 2011
Licensed content publisher	American Society for Microbiology
Licensed content publication	Molecular and Cellular Biology
Licensed content title	Protein Arginine Methylation Facilitates Cotranscriptional Recruitment of Pre-mRNA Splicing Factors
Licensed content author	Yin-Chu Chen, Michael C. Yu
Licensed content date	Nov 1, 2010
Volume	30
Issue	21
Start page	5245
End page	5256
Type of Use	Dissertation/Thesis
Format	Print and electronic
Portion	Full article

Order reference number	
Title of your thesis / dissertation	New roles for arginine methylation in RNA metabolism and cancer
Expected completion date	Apr 2011
Estimated size(pages)	200
Billing Type	Invoice
Customer reference info	
Total	0.00 USD

Terms and Conditions

Publisher Terms and Conditions

The publisher for this copyrighted material is the American Society for Microbiology (ASM). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

1. ASM hereby grants to you a non-exclusive license to use this material. Licenses are for one-time use only with a maximum distribution equal to the number that you identified in the licensing process; any form of republication must be completed within 1 year from the date hereof (although copies prepared before then may be distributed thereafter). The copyright of all material specified remains with ASM, and permission for reproduction is limited to the formats and products indicated in your license. The text may not be altered in any way without the express permission of the copyright owners.

2. The licenses may be exercised anywhere in the world.

3. You must include the copyright and permission notice in connection with any reproduction of the licensed material, i.e. Journal name, year, volume, page numbers, DOI and reproduced/amended with permission from American Society for Microbiology.

4. The following conditions apply to photocopies:
a. The copies must be of high quality and match the standard of the original article.
b. The copies must be a true reproduction word for word.
c. No proprietary/trade names may be substituted.
d. No additional text, tables or figures may be added to the original text.

e. The integrity of the article should be preserved, i.e., no advertisements will be printed on the article.
f. The above permission does NOT include rights for any online or other electronic reproduction.

5. The following conditions apply to translations:
a. The translation must be of high quality and match the standard of the original article.
b. The translation must be a true reproduction word for word.
c. All drug names must be generic; no proprietary/trade names may be substituted.
d. No additional text, tables or figures may be added to the translated text.
e. The integrity of the article should be preserved, i.e., no advertisements will be printed on the article.
f. The translated version of ASM material must also carry a disclaimer in English and in the language of the translation. The two versions (English and other language) of the disclaimer MUST appear on the inside front cover or at the beginning of the translated material as follows:

The American Society for Microbiology takes no responsibility for the accuracy of the translation from the published English original and is not liable for any errors which may occur. No responsibility is assumed, and responsibility is hereby disclaimed, by the American Society for Microbiology for any injury and/or damage to persons or property as a matter of product liability, negligence or otherwise, or from any use or operation of methods, products, instructions or ideas presented in the Journal. Independent verification of diagnosis and drug dosages should be made. Discussions, views, and recommendations as to medical procedures, choice of drugs and drug dosages are the responsibility of the authors.

g. This license does NOT apply to translations made of manuscripts published ahead of print as "[ASM Journal] Accepts" papers. Translation permission is granted only for the final published version of the ASM article. Furthermore, articles translated in their entirety must honor the ASM embargo period, and thus may not appear in print or online until 6 months after the official publication date in the original ASM journal.

6. While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by ASM or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. In addition, permission granted is contingent upon author permission, which you MUST obtain, and appropriate credit (see item number 3 for details). If you fail to comply with any material provision of this license, ASM shall be entitled to revoke this license immediately and retain fees paid for the grant of the license. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described

in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and ASM reserves the right to take any and all action to protect its copyright in the materials.

7. ASM reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. ASM makes no representations or warranties with respect to the licensed material and adopts on its own behalf the limitations and disclaimers established by CCC on its behalf in its Billing and Payment terms and conditions for this licensing transaction.

9. You hereby indemnify and agree to hold harmless ASM and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

10. This license is personal to you, but may be assigned or transferred by you to a business associate (or to your employer) if you give prompt written notice of the assignment or transfer to the publisher. No such assignment or transfer shall relieve you of the obligation to pay the designated license fee on a timely basis (although payment by the identified assignee can fulfill your obligation).

11. This license may not be amended except in a writing signed by both parties (or, in the case of ASM, by CCC on ASM 's behalf).

12. Objection to Contrary terms: ASM hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and ASM (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

13. The following terms and conditions apply to Commercial Photocopy and Commercial Reprint requests and should be considered by requestors to be additional terms. All other ASM terms and conditions indicating how the content may and may not be used also apply.

Limitations of Use:

The Materials you have requested permission to reuse in a commercial reprint or commercial photocopy are only for the use that you have indicated in your request, and

ABSTRACT

Because it can expand the range of a protein's interactions or modulate its activity, post-translational methylation of arginine residues in proteins must be duly coordinated and 'decoded' to ensure appropriate cellular interpretation of this biological cue. This can be achieved through modulation of the enzymatic activity/specificity of the protein arginine methyltransferases (PRMTs) and proper recognition of the methylation 'mark' by a subset of proteins containing 'methyl-sensing' protein modules known as 'Tudor' domains. In order to gain a better understanding of these regulatory mechanisms, we undertook a detailed biochemical characterization of the predominant member of the PRMT family, PRMT1, and of the novel Tudor domain-containing protein 3 (TDRD3). First, we found that PRMT1 function can be modulated by 1) the expression of up to seven PRMT1 isoforms (v1-7), each with a unique N-terminal region that confers distinct substrate specificity, and by 2) differential subcellular localization, as revealed by the presence of a nuclear export sequence unique to PRMT1v2. Second, our findings suggest that TDRD3 is recruited to cytoplasmic stress granules (SGs) in response to envrironmental stress potentially by engaging in methyl-dependent protein-protein interactions with proteins involved in the control of gene expression. We also found that arginine methylation may serve as a general regulator of overall SG dynamics. Finally, we uncovered that alteration of PRMT1, TDRD3, and global arginine methylation levels in breast cancer cells may be closely associated with disease progression and poor prognosis. Therefore, further studies into the pathophysiological consequences ensuing from misregulation of arginine methylation will likely lead to the development of novel strategies for the prevention and treatment of breast cancer.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 2

Chapter 4

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

LIST OF ABBREVIATIONS

3D	tridimensional
3’ or 5’ ss	3’ or 5’ splice site(s)
53BP1	tumor protein p53 binding protein 1
ActD	actinomycin D
aDMA	asymmetrically dimethylated arginine(s)
AdoMet	*S*-adenosyl-*L*-methionine
AdOx	adenosine-2’,3’-dialdehyde
Akt	v-akt murine thymona viral oncogene homolog (serine/threonine-protein kinase)
AMV	avian myeloblastosis virus
ANOVA	analysis of variance
Ars	sodium arsenite
ASF/SF2	alternative splicing factor/splicing factor 2
ASYM	asymmetric dimethylarginine
ATCC	American Type Culture Collection
AZFa	azospermia factor a
BLAP75	Bloom’s Syndrome-associated polypeptide 75
BLAT	BLAST (Basic Local Alignment Search Tool)-Like Alignment Tool
bp	base pair
BRCA1	breast cancer type 1 susceptibility protein
BSA	bovine serum albumin
BTG1/2	B-cell translocation gene 1/2
CA150	transcription elongation regulator 1 (TCERG1)
CARM1	coactivator-associated arginine methyltransferase 1 (a.k.a PRMT4)
CBs	Cajal bodies
cDNA	complementary deoxyribonucleic acid
CF l(m) 68	68-kDa subunit of cleavage factor l(m)
Chromo	chromatin organization modifier domain
CHX	cycloheximide
CIRP	cold-inducible RNA-binding protein

CREB	cAMP response element binding protein
CRM1	chromosome region maintenance protein 1 (exportin 1)
CTD	C-terminal domain (of RNA polymerase II)
DAPI	4',6-diamidino-2-phenylindole
DBTSS	database of human transcription start sites
DC	detergent compatible
DDX3	DEAD/H box-3 protein
DMA	dimethylated arginine
DMEM	Dubelccos's Modified Eagle's Medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DSBs	DNA double-strand breaks
DUF	domain of unknown function
E2F1	E2F transcription factor 1
EBM	exon junction complex-binding motif
ECFP	enhanced cyan fluorescent protein
EEF1A	eukaryotic translation elongation factor 1 alpha 1
EGFP	enhanced green fluorescent protein
eIF3, 4E, 4G, etc.	eukaryotic translation initiation factor 3, 4E, 4G, etc.
EJC	exon junction complex
ERα	estrogen receptor alpha
ES (cells)	embryonic stem (cells)
EST (clone)	expressed sequence tag (clone)
EWS	Ewing sarcoma breakpoint region 1
FAK	focal adhesion kinase
FBX010-11	F-box only protein 10 and 11
FMRP	Fragile X mental retardation protein
Fop	friend of PRMT1
FRAP	fluorescence recovery after photobleaching
FUS	fusion (involved in t(12;16) in malignant liposarcoma)
FXR1p	Fragile X mental retardation-related protein 1
G3BP	GTPase activating protein (SH3 domain) binding protein 1

GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GAR (motif)	glycine- and arginine-rich (motif)
GFP	green fluorescent protein
GST	glutathione S transferase
GW182	glycine-tryptophan protein of 182 kDa (now known as trinucleotide repeat containing 6A)
hCAF1	CCR4-associated factor 1
HDAC3	histone deacetylase 3
His	histidine
HIV-1 Rev	human immunodeficiency virus 1 regulator of virion protein
HMM	hidden Markov model
Hmt1/Rmt1	heterogeneous nuclear ribonucleoprotein methyltransferase 1
hnRNP A1, K, etc.	heterogeneous nuclear ribonucleoprotein A1, K, etc.
Hs	heat shock
Hsp 72	heat shock protein 72
HuR	Hu antigen R (now ELAV-like 1)
IB	immunoblotting
IF	immunofluorescence
IPTG	isopropyl-D-thiogalactopyranoside
IVTT	*in vitro* coupled transcription/translation (reactions)
JMJD2A	jumonji domain containing protein 2A
KH	K homology domain
KLH	keyhole limpet hemocyanin
KSRP	KH-type splicing regulatory protein
LeptB	leptomycin B
LMB	leptomycin B
Magoh	mago-nashi homolog, proliferation-associated protein
MAPK	mitogen-activated protein kinase
MBT	malignant brain tumor
MEFs	mouse embryonic fibroblasts
MEM	Modified Eagle's Medium
miRNA	micro RNA

MLL	mixed lineage leukaemia
MMA	monomethylated arginine
MRE11	meiotic recombination 11 homolog
mRNA	messenger ribonucleic acid
mRNP	messenger ribonucleoprotein
MTA	5'-deoxy-5'-methylthioadenosine
NCBI	National Center for Biotechnology Information
NES	nuclear export sequence
NF-κB	nuclear factor of kappa light polypeptide gene enhancer in B-cells
NFAR	nuclear factor associated with double-stranded RNA
NMD	nonsense-mediated mRNA decay (pathway)
NN	neural network
nt	nucleotide(s)
OB-fold	oligonucleotide/oligosaccharide binding-fold
Osm	osmotic
p53	tumor protein p53
p54/nrb	non-POU domain containing, octamer-binding (NONO)
PABP-1	poly(A) binding protein, cytoplasmic 1
PBS	phosphate buffered saline
PBST	phosphate buffered saline with Tween
PCR	Polymerase Chain Reaction
Pfam	protein family (database)
PGC-1α	peroxisome proliferator-associated receptor gamma coactivator 1 alpha
PI3K	phosphoinositide 3-kinase
piRNA	piwi-interacting RNA
PIWIL1/Miwi	piwi-like 1
PML	promyelocytic leukemia protein
PNCs	perinucleolar compartments
PPARγ	peroxisome proliferator-activated receptor gamma
pre-mRNA	precursor messenger ribonucleic acid
PRMT1-11	protein arginine methyltransferase 1 to 11

PSF	polypyrimidine track binding protein -associated splicing factor
PSP1, -2	paraspeckle protein 1, 2
PTM	posttranslational modification
PVDF	polyvinylidene fluoride
PWWP	‘proline-tryptophan-tryptophan-proline’ domain
RACE	rapid amplification of cDNA ends
Rev (HIV-1)	human immunodeficiency virus 1 regulator of virion protein
RG-rich (motif)	arginine- and glycine-rich (motif)
RIP140	receptor interacting protein of 140 kDa
RIPA (buffer)	radioimmunoprecipitation assay (buffer)
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNP	ribonucleoprotein particle complex
RPMI	Roswell Park Memorial Institute media
RRL	rabbit reticulocyte lysate
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
RUNX1	runt-related transcription factor 1
Sam68	Src-associated in mitosis 68K protein
SAP49	splicing factor 3b, subunit 4, 49kDa (SF3B4)
SCO	Sertoli-cell-only (SCO) syndrome
sDMA	symmetrically dimethylated arginine(s)
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SERBP1	serpine1 mRNA binding protein 1
SG	stress granule
SH3	SRC homology domain 3
shRNA	small hairpin ribonucleic acid
siRNA	small interfering ribonucleic acid
Sm	spliceosomal proteins
SMA	splinal muscular atrophy
SMN	survival of motor neuron protein
SND1	staphylococcal nuclease and tudor domain containing 1
snRNA	small nuclear ribonucleic acid

snRNP	small nuclear ribonucleoprotein
snoRNP	small nucleolar ribonucleoprotein
SPF30	splicing factor 30 (now known as survival motor neuron domain containing 1)
SR	serine- and arginine-rich protein
Src	sarcoma-inducible gene
SRSF9	serine/arginine-rich splicing factor 9 (a.k.a. SRp30c)
SUMO	small ubiquitin-like modifier
Syk	spleen tyrosine kinase
SYM	symmetric dimethylarginine
TCL	total cell lysate
Tdr	Tudor domain
TDRD1, 2, 3, etc.	Tudor domain-containing protein 1, 2, 3, etc.
TIA-1	T-cell-restricted intracellular antigen-1 (cytotoxic granule-associated RNA binding protein)
Tob	transducer of ErbB2
TOP3β	DNA topoisomerase III beta
TR3 (NR4A1)	nuclear receptor subfamily 4 group A member 1 (a.k.a. orphan nuclear receptor TR3)
TSS	transcriptional start site
U1-70K	small nuclear ribonucleoprotein 70 kDa (SNRNP70)
U1C	small nuclear ribonucleoprotein polypeptide C (SNRPC)
UBA	ubiquitin-binding domain A
UCSC	University of California, Santa Cruz
UTRs	untranslated region(s)
v1-7	variant 1 to 7
VEGF	vascular endothelial growth factor
VHL	tumor suppressor Von Hippel-Lindau protein
W	watts
WW (domain)	protein domain with two highly conserved tryptophans that binds proline-rich peptide motifs
Y14	RNA binding protein
YY1	YY1 transcription factor

ACKNOWLEDGMENTS

This thesis would not have been possible without the support and guidance of a number of individuals. First and foremost, I would like to thank my supervisor Dr. Jocelyn Côté for giving me this opportunity and above all, for his unwavering support, guidance, understanding, and patience during all these years.

Second, I would like to express my sincere gratitude to my friends and all members of Côté's Laboratory, both former and current. I would especially like to thank Dr. Hanan Abramovici to whom I owe the greatest debt, as the submission of this thesis would not have been at all possible without him. I would also like to thank Dr. Lise Boily, Sophie Boisvenue, Dr. Louise Chrétien, Josée Davies, Gabrielle Gauvin, Dr. Erik Harvey-Girard, Benoît Paquette, Helina Tadesse, and Dr. Hector Valderrama to have generously shared their expertise and wisdom, to have been a great source of titillating and enlightening conversations. On a more personal note, I would like to thank my father, Jocelyn Goulet, and his life partner, Lynda Dumont, for their support and encouragement; Gilles and Ghislaine Bourdeau for their cordial welcome and support during the first years of this journey; and last but not least, my life partner, Ray Pilon, whose love, confidence, and support have been crucial in the successful completion of this lifelong goal.

Third, I would like to thank the members of my advisory committee – Dr. John Bell, Dr. Martin Pelchat, and Dr. Barbara Vanderhyden – for their insightful comments and suggestions.

Finally, I would like to acknowledge the funding agencies which have supported the research presented in this thesis: *Canadian Breast Cancer Foundation* (I. Goulet), *Fonds de la recherche en santé du Québec* (I. Goulet), *University of Ottawa* (J. Côté and I. Goulet), *Canadian Institutes of Health Research* (J. Côté), and *Cancer Research Society* (J. Côté).

Music listened to while writing this thesis:

Barenaked Ladies, *Gordon*
Beatles (The), *Best of*
Black Tiger Sex Machine, *2000+*, *La Panthère Magique*, *The Galaxy Express*, *Marvin Gaye – Too Busy*, *Sapce Bordello*, *Keep On Makin' Me High*
Joe Pass Trio (The), *Eximious*
Lhasa, *La LIorona*
Norah Jones, *Come away with me*, *Feels like home, The Fall*
Pat Metheny, *One quiet night*
Rolling Stones (The), *Stripped*
Stanley Turrentine, *Ballads*
Susie Arioli Band, *It's wonderful*, *Pennies from heaven, That's for me*
Sweet and Lowdown, *movie soundtrack*
Tracy Chapman, *New beginning*
Zero 7, *Simple Things*

“Triumph without peril brings no glory.”
Pierre Corneille

"*I hate being a DNA molecule... there's so much to remember!*"
Frank and Ernest, by Bob Thaves

In loving memory of my mother
Monik Vadeboncoeur

Preface

In accordance with the Faculty of Graduate and Postdoctoral Studies' (FGPS) thesis regulations (University of Ottawa), the experimental findings presented in this thesis are comprised of one unpublished and two published journal articles, reprinted here with permission from respective publishers (see Authorizations). Published collaborative work done during my doctoral studies is also presented in appendix, reprinted with permission from publisher as well (see Authorizations). Each of the articles included in this thesis is prefaced with a short description, followed by a statement of contributions of collaborators and co-authors, as prescribed by the FGPS.

The first chapter of this thesis (Chapter 1) contains a general overview of our current knowledge on arginine methylation, the protein arginine methyltransferase 1, and the methyl-binding Tudor domains. Chapter 2 is a manuscript describing the influence PRMT1 isoforms' unique N-terminal sequences have on its localization, methyltransferase activity, and substrate specificity. Groundwork done on the biochemical characterization of the novel TDRD3 protein forms the basis of Chapter 3. The manuscript describing TDRD3 recruitment to cytoplasmic stress granules is presented in Chapter 4. The final chapter (Chapter 5) is dedicated to a general discussion of these results and their relevance to the field of arginine methylation and (breast) cancer biology.

CHAPTER 1

General Introduction

A cell's survival depends on its capacity to maintain homeostasis, i.e. its ability to continuously adapt to fluctuating environmental conditions in order to maintain viability. Such adaptation is made possible by the integrated actions of thousands of intracellular molecules contrived to quickly sense, process, and respond to changes that threaten to upset this delicate balance.

As the 'workhorses' of the cell, proteins often are the first molecules to respond to such changes, coordinating and carrying out thousands of biochemical reactions at any time with the ultimate goal of regulating various biological functions (e.g. growth, differentiation, death). Because of the complexity of such biochemical reaction networks, specific mechanisms have evolved to ensure that cellular proteins respond appropriately and accurately to extra- and intracellular cues. Such mechanisms include regulation of protein turnover, modulation of protein activity, protein-protein interaction, and protein localization. The cell uses three major means to achieve such fine-tuning of protein function and levels: *de novo* protein synthesis, protein degradation, and addition of post-translational modifications (PTMs) to existing proteins. Deregulation of any of these mechanisms can ultimately lead to the emergence of a number of pathophysiological processes culminating in disease.

In eukaryotic cells, protein synthesis and degradation are relatively slow processes, require a multitude of players and represent a significant energy investment for the cell (1, 2). On the other hand, reversible addition of post-translational modifications (PTMs) to existing proteins offers an advantageous alternative to modulate the activity of the target proteins: it proceeds on a time scale of seconds to minutes, requires relatively few proteins, and therefore, places fewer demands on the cell's energy stores (3-5). Virtually all of the 20 amino acid residues used as building blocks for proteins can be post-translationally modified by covalent addition of different functional groups (or other polypeptides) to extend the range of protein function (6). The ensuing biological effects of these modifications can differ greatly, depending on the types of modifications and their targets, but PTMs usually cause changes in protein activity, localization, dynamic interactions with other proteins, or stability (6). Thus, PTMs represent a cheap, energy-efficient, and powerful way to regulate protein function.

While kinase-dependent phosphorylation of protein serine, threonine, or tyrosine residues is indisputably the most studied and understood of these PTMs, in recent years, the development of new detection techniques has revealed methylation to be another major PTM involved in controlling protein function. In contrast to phosphorylation, protein methylation occurs on lysine and arginine residues (7), offering a distinct range of potential substrates in the cell.

1.1. Protein arginine methylation: chemistry and general biological significance

Arginine methylation is one of the most common protein methylation reactions in mammalian cells (8, 9) and a large array of different proteins contain methylated arginine(s) (10): transcription and splicing factors, nucleic acid binding proteins, signal transducers, and histones are but a few examples of the ever expanding repertoire of arginine-methylated proteins.

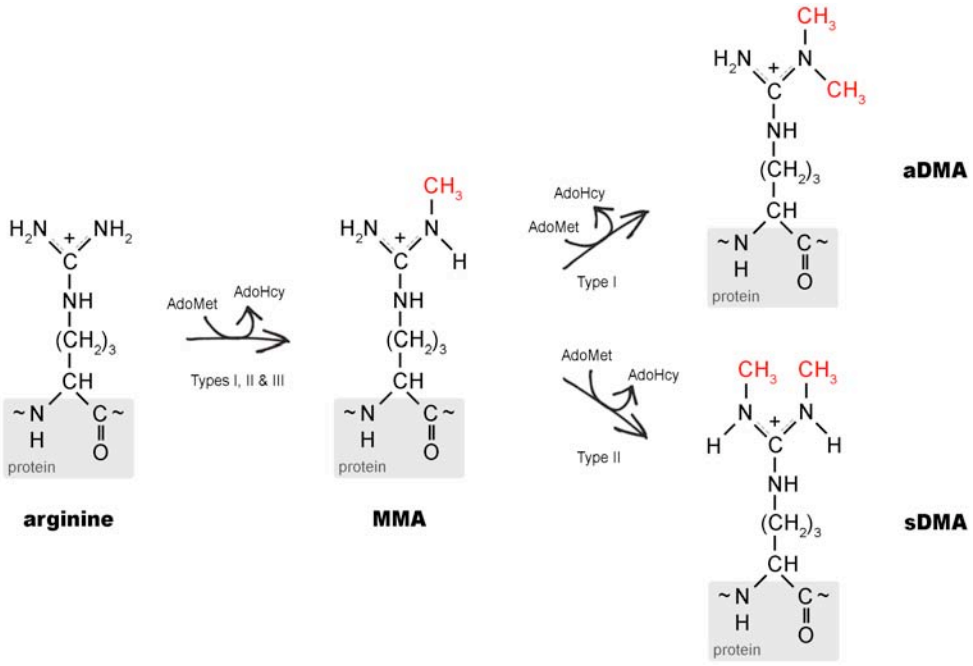


FIGURE 1. Types of methylated arginines generated by mammalian protein arginine methyltransferases. Methylation of arginine proceeds by the addition of a methyl group (CH_3), donated from S-adenosyl-methionine (AdoMet), to one of the guanidino amino groups of arginine, generating ω-N^G-monomethylated arginine (MMA). Type I PRMT enzymes catalyze the formation of asymmetrically ω-N^G,N^G-dimethylated arginine (aDMA) through the addition of a second methyl group to the same nitrogen. Symmetrically ω-N^G,N'^G-dimethylated arginine (sDMA) is generated by Type II PRMTs, which catalyze the addition of a methyl group to the other terminal nitrogen of the arginine. The use of AdoMet as a source of methyl group by PRMTs produces S-adenosyl-homocysteine (AdoHcy), a by-product that is recycled back to AdoMet within the cell.

Arginine methylation results from the transfer of methyl group(s) from S-adenosyl-L-methionine (AdoMet) to the side chain guanidino group of arginines in protein substrates

by proteins known as methyltransferases (11) (Fig. 1). Although methylation does not alter the positive charge of the arginine, it increases its bulkiness and hydrophobicity, and consequently, modifies its binding properties (12). Thus, as a means of post-translational regulation, arginine methylation may provide otherwise identical proteins with distinct binding properties which can affect their activity and ultimately their function. This simple and rapid modification of protein chemical properties in turn enables cells to respond quickly to both external and internal stimuli, without having to rely on the rather slow and energy-consuming process of synthesizing new proteins and degrading old ones.

1.2. Protein arginine methyltransferases

The mammalian protein arginine methyltransferases (PRMTs), the enzymes responsible for arginine methylation, compose a family of nine related members that have been classified into three types: I, II, and III (Fig. 2). All three types of enzymes catalyze the formation of ω-N^G-monomethylated arginine (MMA), where a single methyl group is placed on a terminal nitrogen atom of the guanidino group (13) (Fig. 1). Type I PRMTs [PRMT1 (14), PRMT2 (15), PRMT3 (16), PRMT4 (a.k.a. CARM1) (17), PRMT6 (18), and PRMT8 (19, 20)] are responsible for the subsequent generation of the most prevalent type of methylated arginine derivative: asymmetrically ω-N^G,N^G-dimethylated arginine (aDMA) (8). In this case, two methyl groups are placed on one of the terminal nitrogen atoms of the guanidino group (13) (Fig. 1). PRMT5 (21-23) and PRMT7 (13) are the only known type II enzymes and promote the formation of symmetrically ω-N^G,N'^G-dimethylated arginine (sDMA), where one methyl group is placed on each of the terminal

guanidino nitrogens (13) (Fig. 1). PRMT7 also exhibits type III enzymatic activity, which is restricted to the generation of monomethylarginines (13). However, the data regarding PRMT7 must be interpreted with caution as the types of methylation reaction catalyzed by this protein are still under investigation and remain to be clearly demonstrated. To date, no methyltransferase activity has been observed for PRMT9 (24) and no enzyme has been found that forms both aDMA and sDMA derivatives (13).

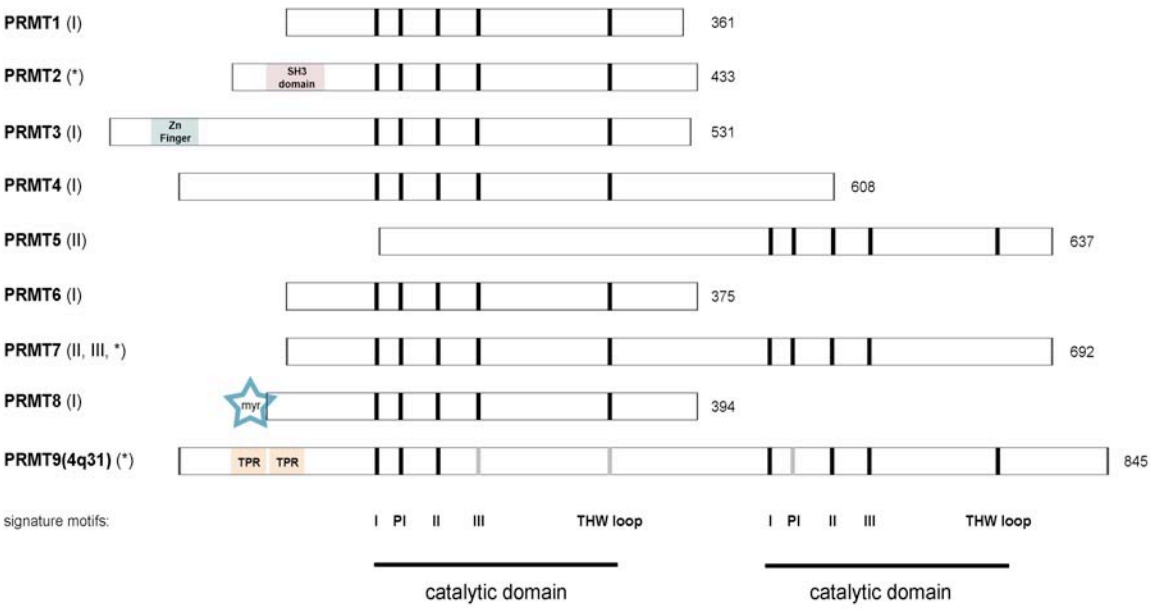


FIGURE 2. The mammalian protein arginine methyltransferase family. All PRMTs contain a highly conserved methyltransferase domain of approximately 310 amino acids harboring four signature motifs (I, post-I, II, and III) as well as conserved "THW" loop involved in AdoMet binding and subsequent transfer of methyl groups to arginine residues (25-29). Each PRMT exhibits a unique N-terminal region that varies considerably in length, but which usually contains two α helices as well as other domains possibly involved in substrate recognition (13, 26). Types of arginine methylation catalyzed by each enzyme are indicated in parentheses (type I, II, III), whereas the types of arginine methylation catalyzed by the enzymes marked with an asterisk are still under investigation.

Two related members of the F-box only family of proteins, FBXO10 and FBXO11, have also been suggested to be protein arginine methyltransferases (24, 30). Although neither protein seems to share the characteristic seven β-strand methyltransferase motifs

common to PRMT1-9, comparative sequence analysis revealed distant amino acid sequence similarities between the two F-box only proteins and the PRMT family (13, 24, 28). While biochemical characterization of FBXO10 remains to be performed (30), the attribution of type II methyltransferase activity to FBXO11 (24) may have actually been the result of contamination of the affinity-purified enzyme with PRMT5 (31) as an independent study reported no activity for the human or worm FBXO11 proteins (13, 32). Thus, further research will be required to determine whether F-box only proteins can be protein arginine methyltransferases (13).

As the role of PRMTs continues to be better understood, the number of arginine-methylated proteins and the cellular processes they govern continues to expand. Therefore, for the sake of simplicity and concentrating on the main findings relevant to the present thesis, this chapter will focus on the current knowledge regarding PRMT1, the functional implications of its methyltransferase activity, and its putative link to specific pathophysiological processes such as carcinogenesis. Separately, this chapter will also summarize information regarding a subset of proteins containing so-called 'Tudor' domains and their roles in mediating methyl-dependent protein-protein interactions in various cellular processes regulated by protein methylation.

1.3. PRMT1: the major arginine methyltransferase

Of the nine mammalian PRMTs, PRMT1 performs an estimated 85% of all the arginine methylation reactions in human cells (33). PRMT1 is found in the nucleus and the cytoplasm of virtually all cell types and is expressed in all tissues of the human body (34,

35). The abundance of PRMT1 and its cellular distribution emphasize its important physiological role as the dominant PRMT in the cell.

1.3.1. Substrate specificity and enzyme activity

Although much research has been performed on PRMT1 over the last fifteen years, the mechanism defining its substrate specificity remains poorly understood. Most methylated arginine residues in PRMT1 protein substrates reside in glycine ("RGG")- and arginine ("RXR")-rich (GAR) motifs (11), but there are exceptions to this rule (36). According to Wooderchak (37), the fact that PRMT1 substrates may not be limited to proteins bearing such GAR sequences strongly suggests that its substrate proteome may be more diverse than previously anticipated. Therefore, understanding how this enzyme selects certain substrates over others and performs catalysis is of crucial importance to get the full spectrum of biological functions of PRMT1 (37).

A number of recent studies have begun to provide some clues as to how PRMT1 might recognize its substrates. While negatively charged residues in the surface grooves of PRMT1 have been shown to be important for substrate binding through pull-down experiments (36, 38), surface residues distal to the active site might also contribute to substrate specificity. For example, PRMT1 exhibits increased activity towards full-length proteins in comparison to smaller peptides with similar recognition sites (39), suggesting that PRMT1 methyltransferase activity might be enhanced by additional enzyme-substrate contacts outside of the catalytic core (40). Complementing this finding, residues distal to the site of methylation in its protein substrates were also shown to be important

for PRMT1 substrate binding and catalysis (36, 41). Taken together, these observations suggest that PRMT1 activity and specificity towards its substrates may also involve other residues than those constituting the catalytic core (motifs I, post-I, II, and III in Fig. 2).

Aside from the highly conserved 310-amino acid methyltransferase catalytic core shared by all PRMTs, each PRMT displays unique N- and C-terminal extensions. The function of these PRMT-specific 'non-core' sequences is still not well understood, but results from independent studies suggest putative contributions of the N-terminal sequences (varying in length from approximately 30 to 200 amino acids (26); Fig. 2) to PRMT activity and specificity. For example, deletion of the N-terminus of PRMT3, which contains a zinc-finger motif (Fig. 2), decreases its enzymatic activity and affects its protein substrate specificity (16, 42). Similarly, the presence of the PRMT4/CARM1 'pre-core' N-terminal domain is required for methylation of the arginines at position 2 and 26 in histone 3 (H3R2 and H3R26) (43). The unique N-terminal region of PRMT1, in contrast, is only ~30 amino acids in length, indicating that its catalytic core may carry out most PRMT1 functions.

Complicating this picture are studies demonstrating that alternative splicing of the primary human PRMT1 transcript can generate at least three protein isoforms (v1-3) differing only in their first 6 to 30 N-terminal amino acids (34, 35) (Fig. 3). According to existing PRMT1 structural information (25), the varying sequences are N-terminal to the cofactor-binding α helices (αX and αY in Fig. 3), which suggests that these three N-terminal variants should still be able to interact with the AdoMet cofactor to a similar

extent and therefore, should all still be enzymatically active. Therefore, these short extreme N-terminal sequences are more likely to govern isoform-specific interactions with distinct sets of substrates during catalysis. This hypothesis, however, remains to be explored. Indeed, determining whether PRMT1 variant's unique N-terminal sequences can influence its substrate specificity is of crucial importance, as this could in turn provide a molecular mechanism for differences in substrate specificities and enzymatic activities observed between various PRMTs.

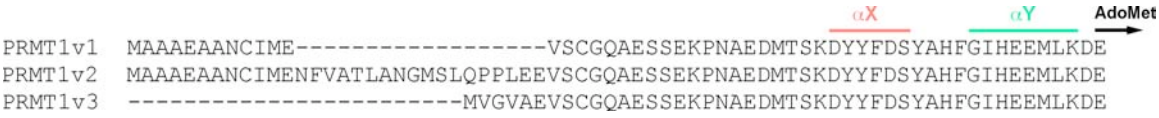


FIGURE 3. Multiple sequence alignment showing the N-terminal unique domains of PRMT1v1-3. The position of αX and αY helices as well as the beginning of the AdoMet binding domain is indicated above the sequences.

1.3.2. Alternative mechanisms regulating PRMT1 activity and substrate specificity

In addition to the intrinsic structural and substrate binding properties of PRMT1 and the production of a set of distinct protein isoforms, the activity and specificity of PRMT1 may be regulated in a number of different ways. One way is through the regulation of subcellular localization since PRMT1 is highly mobile between the cytoplasm and the nucleus (44). The relative distribution of cytoplasmic versus nuclear PRMT1 also varies across different cell types (45), suggesting the existence of cell- or tissue-specific mechanisms regulating PRMT1 localization, although little is known regarding the mechanisms responsible for such differences in localization.

The activity of PRMT1 could also be influenced by post-translational modifications, although the evidence to date for such a mechanism is tenuous. Nevertheless, some

evidence suggests that PRMT1 possesses automethylation activity (40) (as it has also been shown for PRMT4/CARM1 (46), PRMT6 (18), and PRMT8 (20)). The mechanisms involved and the functional consequences of automethylation, however, remain unknown. The activity of PRMT1 could also hypothetically be modified in a rapid manner by phosphorylation as a conserved serine residue (S217) found amongst the various PRMTs can be phosphorylated in PRMT4/CARM1, resulting in complete abolition of methyltransferase activity by preventing methyl donor AdoMet binding (47). Higashimoto and colleagues also reported that phosphorylation at another conserved serine residue (S229) could inhibit PRMT4/CARM1 methyltransferase activity, possibly by interfering with PRMT4/CARM1 homodimerization (48). The presence of post-translational modifications adjacent to arginine methylation motifs located in target proteins could also block PRMT accessibility to these sites. For example, various PTMs (especially methylation and acetylation) on N-terminal histone tails, which dictate chromatin-DNA interactions, have been shown to block methylation of adjacent arginines at positions 2 and 8 (49-51) and of adjacent lysines at position 4 and 9 on the Histone 3 (52, 53). Thus, post-translational modifications on both the enzyme and the substrate represent another potential mechanism by which the activity or substrate specificity of PRMT1 may be modulated, but further research is required to determine the prevalence and importance of this regulatory mechanism.

Although PRMT1 is an active methyltransferase in the absence of other polypeptide species *in vitro*, significant evidence exists indicating that its activity and/or specificity can be regulated through binding to additional partners *in vivo* (13). For example,

PRMT1 was first isolated through its interaction with the antiproliferative proteins BTG1 and BTG2, and both proteins can stimulate PRMT1's activity towards selected substrates (14, 54-61). Interestingly, depletion of BTG2 leads to a decrease in methylated arginines in the nucleus, but not in the cytoplasm, suggesting that modulators of PRMT1 activity are present in different cellular compartments (62). PRMT1 activity is also modulated in a substrate-dependent manner by hCAF1, a known interactor of BTG1, which inhibits, among others, the methylation of Sam68 and Histone 4 arginine 3 (H4R3) (63, 64). Binding of the nuclear orphan receptor TR3 (65), a transcription factor that plays key roles in cell proliferation and apoptosis, or the binding of protein phosphatase 2A (66), involved in the control of many cellular functions, to PRMT1 also result in repression of PRMT1 activity, consequently decreasing the methylation of several downstream substrates. Taken together, these findings suggest that interaction with other proteins is a major mechanism by which the methyltransferase activity of PRMT1 can be modulated *in vivo*.

1.3.3. Functional implications of PRMT1 expression

While loss of the expression of the PRMT1 homolog (*Hmt1/Rmt1*) in *Saccharomyces cerevisiae* does not affect cell viability, it results in mislocalization of several cellular proteins as well as defects in maintaining silent chromatin (67, 68). In mice, reduction of PRMT1 expression to approximately 5% of its normal level leads to embryonic lethality, but yet does not compromise the viability of embryonic stem (ES) cells generated from these mice, even though they exhibit a significant decrease in protein arginine methylation (69). In contrast, complete loss of PRMT1 expression in mouse embryonic

fibroblasts (MEFs) results in spontaneous DNA damage, cell cycle progression delay, checkpoint defects, aneuploidy, and polyploidy, which suggest that PRMT1 is required for genome integrity and cell proliferation (70).

1.3.3.1. Roles in DNA repair

In response to DNA double-strand breaks (DSBs), generated by genotoxic agents or by cellular endonucleases as intermediates of several important physiological processes, cell cycle checkpoints are activated to delay cell cycle progression. Such actions give the cell time to initiate DNA repair pathways, in which two of the protein players are currently known substrates of PRMT1: MRE11 and 53BP1. Arginine methylation of MRE11 is required for its localization to DNA damage foci and for proper regulation of its exonuclease activity on double-stranded DNA (71). On the other hand, 53BP1's methylatable GAR motif regulates its ability to oligomerize (72) and to associate with DNA (73), although independently of PRMT1 activity. In addition, PRMT1-deficient MEFs are defective for the recruitment of RAD51 to DNA damage foci (70), indicating that PRMT1 is also required for RAD51 recruitment to sites of DNA DSB repair. Furthermore, loss of PRMT1 in MEFs leads to genomic instability, a condition generally observed in cells that fail to repair DNA DSBs. The phenotypic abnormalities observed in PRMT1-deficient MEFs (e.g. spontaneous DNA damage, sustained delay in cell cycle progression, genetic instability, etc.) indeed suggest that PRMT1 function is essential in the cellular response to DNA DSBs. However, the exact molecular mechanisms by which PRMT1 functions in DNA repair pathways remains largely unknown and will require further investigation and research.

1.3.3.2. Roles in the regulation of gene expression

Histones represent major targets for PRMTs, where methylation is part of the histone post-translational modification code regulating chromatin structure and function (74). For example, PRMT1, functioning synergistically with PRMT4/CARM1, is recruited to specific genomic regions, in part by interactions with various transcription factors (including p53 (75), YY1 (76), NF-κB (77), PPARγ (78), RUNX1 (79), CREB (80) and E2F1 (81)), leading to alterations in histone methylation patterns that dictate changes in gene expression (82). Addition of aDMA to histones by PRMT1 generally serves as an activation signal for gene expression (83, 84). For example, PRMT1 specifically dimethylates the R3 residue of the histone H4 tail, which in turn promotes histone H3 and H4 acetylation (83). In general, increased acetylation levels of H3 and H4 are characteristic of an open chromatin state and active gene transcription (85-87). PRMT1 methyltransferase activity, however, can be also involved in repression of gene expression (84). For example, the yeast homolog of PRMT1, *Hmt1*/*Rmt1*, is involved in the maintenance of silent chromatin in yeast cells (68)), suggesting that the effect of PRMT1 function may be context-specific. Thus, arginine methylation of histones provides PRMT1 with a means to impact on the regulation of gene expression.

1.3.3.3. Roles in signal transduction and transcriptional co-regulation pathways

Analogously, a number of studies have suggested a function for PRMT1 in the regulation of hormone-dependent transcription, a process which plays a part in controlling the expression of cell-cycle genes that influence cell proliferation (88-93). One hyperproliferative disorder characterized by defects in hormone-dependent signaling

pathways impinging on transcription is breast cancer and PRMT1 may have a role to play here. Thus, the following section focuses on PRMT1's putative roles in the regulation of estrogen signaling and estrogen-dependent transcription.

PRMT1 methylates the estrogen receptor α (ERα) rapidly following estrogen (E2) induction (89). In breast epithelial cells, this methylated form of ERα can be found exclusively in the cytoplasm (89). PRMT1-mediated methylation is required for the extra-nuclear (nongenomic) function of the receptor in mediating the assembly of a cytoplasmic signaling complex consisting of ERα, the tyrosine kinase Src, the p85 subunit of the lipid kinase PI3K (Phosphoinositide 3-kinase), and the focal adhesion kinase (FAK), a substrate of Src involved in cell migration (89). The formation of this complex then triggers the activation of signaling kinases such as Akt and MAPK, whose downstream signaling cascades orchestrate cell survival and cell proliferation (94).

Methylation of ERα in response to estrogen is normally transitory in the cytoplasm of healthy breast epithelial cells, suggesting the existence of yet unknown mechanism(s) that can rapidly reverse the methylation of ERα (89, 95). The methylation of ERα, however, is significantly increased in a subset of breast cancer tumors (89, 95). Whether high levels of methylated ERα in breast tumors are the result of deregulated demethylation mechanism(s) and/or hyperactivated PRMT1 methyltransferase activity remains to be elucidated. Regardless, such persistent methylation may lead to sustained kinase activation, notably of Akt, which in turn activates estrogen signaling pathways, even in the absence of estrogen: a potentially selective survival advantage for primary tumor

cells. Moreover, the methyalted arginine (R260) in ERα lies within its DNA-binding domain and within one of its nulcear localization signals, which might interfere with its nuclear import (see below for the genomic function of ERα) (89, 95). These findings therefore suggest a role for PRMT1 and methylated ERα in signal transduction pathways whose deregulation may lead to mammary tumorigenesis (95).

In the nucleus, PRMT1 (synergistically with CARM1) forms part of the ER coactivating complex, in which it executes several methylation reactions, such as at the R3 residue of the histone H4 tail that activates gene transcription (17, 96, 97). PRMT1-dependent methylation also increases PGC-1α's (Peroxisome proliferator-associated receptor gamma coactivator 1 alpha) coactivator functionality for several nuclear receptors, including ERα, triggering target gene expression (98). The association of Fop (Friend of PRMT1) with PRMT1 and its methylation by PRMT1 are also critical events for ER-regulated activation of gene transcription (99). PRMT1-induced methylation of RIP140, a hormone-dependent ER-binding corepressor, reduces its repressive activity by two mechanisms. First, methylation of the R240 impairs RIP140 interaction with histone deacetylase 3 (HDAC3), thereby preventing chromatin condensation and subsequent gene silencing (100). Second, methylation of R240, R650, and R948 increases RIP140 interaction with the nuclear exporter CRM1 (101). This results in RIP140 sequestration to the cytoplasm, thereby diminishing its nuclear repressive activity.

Besides its roles in genomic and non-genomic estrogen pathways, PRMT1 is also an essential component of a novel Mixed Lineage Leukemia (MLL) oncogenic

transcriptional complex that modifies histones in a similar fashion to the ER coactivating complex (102). Methylation by PRMT1 also influences the capacity of other transcription factors to bind to specific promoters (75, 76, 79), such as BRCA1 (Breast cancer type 1 susceptibility protein), which may ultimately affect its tumor suppressor activity (103). Thus, arginine methylation of histones, transcriptional factors, and co-regulators provides PRMT1 with various ways to impact on the epigenetic regulation of gene expression, and for which deregulation has been shown to play important roles in tumor initiation and progression (104).

1.3.3.4. Roles in pre-mRNA splicing

Several studies have also suggested functions for arginine methylation in the regulation of co- and post-transcriptional processes, including pre-mRNA splicing (105-107), mRNA export (67, 108-111), and translation (112, 113).

Direct evidence for a role of arginine methylation in pre-mRNA splicing first came from the observation that splicing efficiency is greatly reduced in nuclear extracts prepared from cells grown in the presence of AdoMet-dependent methylation inhibitors, indicating that normal levels of methylarginine-containing proteins are required for pre-mRNA splicing *in vitro* (105). In addition, preincubation of nuclear extracts with an increasing amount of DMA-specific antibodies resulted in complete inhibition of splicing, demonstrating that DMA-containing proteins are part of the active spliceosome (105). Consistent with these findings, several heterogeneous nuclear ribonucleoproteins (hnRNPs), small nuclear ribonucleoproteins (snRNPs), serine- and arginine-rich (SR)

proteins, and other splicing factors involved in pre-mRNA splicing were identified in a large proteomic survey of arginine-methylated protein complexes (114). Methylation of a number of these splicing factors by PRMT1 (such as SRSF9/SRp30c (115), U1-70k (116), and ASF/SF2 (117-119)) and other PRMTs (106, 120-122) has since been confirmed.

In the nucleus, CARM1 and PRMT6, which catalyze asymmetric dimethylation of target arginines like PRMT1, are transcriptional coactivators, and also the only type I mammalian PRMTs that have been clearly identified as regulators of alternative splicing (106, 107, 123). The enzymatic activity of PRMT6 increases skipping of alternative exons in VEGF and Syk genes in a steroid hormone-independent manner (107). PRMT6 most likely regulates alternative splicing by methylating other splicing factors although this remains to be confirmed. Similarly, CARM1 methylates the splicing factors CA150, SAP49, SmB, and U1C, and promotes exon skipping in an enzyme-dependent manner (106). In yeast, the mammalian PRMT1 homolog *Hmt1*/*Rmt1*, which also generates aDMA, has been shown to promote the dynamic interactions between RNA-binding proteins and pre-mRNAs required for proper assembly of mRNPs during transcription (116, 124-127). Recent findings from our collaborators further demonstrate that *Hmt1*-mediated methylation regulates the specificity of the splicing machinery by recruiting splicing factors co-transcriptionally to particular pre-mRNAs (116) (and Appendix II). Since PRMT1 and CARM1 function synergistically as general transcriptional coactivators in mammals, this raises the possibility that mammalian PRMT1 could also function as a co-transcriptional regulator of pre-mRNA splicing, like *Hmt1*, in

partnership with CARM1. Together, these observations strongly suggest that arginine methylation is essential for accurate and efficient transcription-coupled processing of a number of pre-mRNAs. However, the precise molecular mechanisms involved in the regulation of these processing events by arginine methylation remain largely unknown and are research topics our laboratory is actively pursuing.

One potential mechanism involves the formation of a 'docking motif' created by arginine methylation of RNA binding proteins and splicing factors to recruit methyl-binding effector proteins that could help orchestrate the proper localization and function of transcription-coupled mRNA processing events. Supporting this hypothesis, a pilot series of titration experiments, in which HeLa nuclear extracts were incubated with increasing amounts of DMA-containing peptides, resulted in gradual inhibition of splicing (Côté, J., unpublished data). This observation provided us with functional evidence that constitutive methyl-dependent interactions with 'DMA-sensing' proteins are essential for splicing reactions to occur.

1.4. Tudor domains: 'DMA-sensing' protein modules

1.4.1. Types of Tudor domains

In most studies reported to date, arginine methylation acts as a negative regulator of protein-protein interactions (79, 101, 128-132). Since arginines are especially important in hydrogen bonding, methylation of arginine residues occurring at the interaction interface is likely to disrupt intra- and intermolecular interactions. However, increasing experimental evidence suggests arginine methylation may not always be inhibitory. In

some cases, methylation of arginine residues actually promotes protein-protein interaction, such as is the case with a subset of proteins containing so-called 'Tudor' domains.

The Tudor domain was first discovered through a protein sequence comparison analysis, performed to identify conserved patterns in *Drosophila* TUDOR (133, 134), a protein involved in germ cell formation that harbors 11 repeated Tudor domains. Since then, over 200 Tudor-containing proteins have been identified in a wide variety of eukaryotes, from fungi to animals, but not in prokaryotes (135).

Tudor domains consist of approximately 60 amino acids that form a hydrophobic barrel-like structure composed of β-sheets, surrounded by negatively charged and aromatic residues, which together likely constitute a protein-protein interaction surface (136, 137). This structural fold shares similarities with the Chromo domain, a motif known to interact with methylated lysines in histones (138). This resemblance led to the proposition that methyl-substrate binding might be a common feature of the related Tudor, Chromo, MBT, PWWP, and Agenet-like domains, which together form the so-called 'Royal Family' of domains that engage in protein-protein interactions (138).

Four types of Tudor domains can be distinguished based on their flanking sequences and/or linkage to other conserved motifs, providing different Tudor domain-containing proteins with distinct properties and biological functions (139) (Fig. 4). The original germline type Tudor domain proteins (TDRD1, 2, 4, 5, 6, 7, 8, and 9) form one group

(Fig. 4). This group is characterized by multiple Tudor repeats, some of which are flanked on their amino-terminal side by an α helix and β strands, and which bind to proteins with dimethylated arginine residues (135, 139). Proteins containing this type of Tudor domain accumulate in germline-specific ribonucleoprotein (RNP) complexes called nuage/polar/germinal granules, which in turn are essential for the recruitment of methylated Piwi-type proteins involved in retrotransposon silencing by piwi-interacting RNAs (piRNAs) during gametogenesis (135, 140-144).

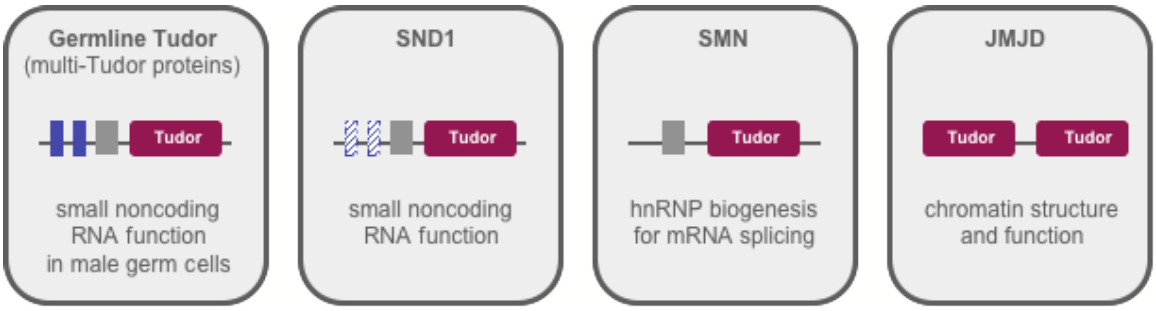


FIGURE 4. Classes of Tudor domain types. Four types of Tudor domains (germline Tudor, SND1, SMN, and JMJD) as well as their known or predicted functions are indicated. The structural organization of each Tudor domain class is also indicated with the colored boxes: Tudor domains are marked in dark red, the N-terminal α helices in gray, and the known or predicted β-strand extensions in blue. Adapted from (131).

The Tudor domain of SND1 (a.k.a. p100, Tudor-SN) characterizes a second group of Tudor domains (Fig. 4). SND1 has five tandem staphylococcal nuclease (SN)-like domains with its Tudor domain inserted into the fifth one, thereby creating a composite SN-Tudor structure (145). Like other germline Tudor proteins, SND1 associates with arginine methylated PIWIL1/Miwi in germ cells (146). SND1 is also a main component of the RNA-induced silencing complex (RISC) involved in gene regulation by small interfering RNAs (siRNAs) or micro RNAs (miRNAs) (147).

A third group is found in pre-mRNA splicing factors (Fig. 4), in which the Tudor domains of the 'survival of motor neuron' (SMN) protein and of the splicing factor 30 kDa (SPF30) have been demonstrated to be required for the assembly of spliceosomal small nuclear ribonucleoproteins (snRNPs) (148-151) and of the spliceosome (152, 153) through methyl-dependent interactions with a number of arginine methylated RNA binding proteins during pre-mRNA splicing (105, 106, 148, 154-158).

The last group of Tudor domains, dubbed "tandem Tudor", recognizes methylated lysines on histone tails (158, 159). These include the Tudor domains of 53BP1 (160, 161), involved in DNA repair, and of JMJD2A (162), a transcriptional repressor (Fig. 4). Collectively, these observations suggest that Tudor domain-containing proteins can recognize 'methyl marks' associated with the regulation of a number of transcriptional and RNA processing events.

1.4.2. TDRD3: a putative methyl-binding effector protein in mRNA processing events?

A proteomic analysis of the complete repertoire of the mammalian splicing machinery has identified another Tudor domain-containing protein in pre-spliceosomal complexes: the novel Tudor domain-containing protein 3 or TDRD3 (163). TDRD3 is a modular protein, known to be expressed based on the presence of several ESTs, but remains poorly characterized. Its Tudor domain is highly similar to that of SMN and SPF30, and it has the ability to bind dimethylated arginines in polypeptides (148, 158). In addition to its Tudor domain, TDRD3 harbors a DUF/OB-fold motif (a putative sugar, nucleic acid, or protein recognition motif) (164), a UBA (ubiquitin-binding) domain (165), and an

exon junction complex (EJC)-binding motif (EBM) (166) (Fig. 5). Although the multidomain structure of TDRD3 is well characterized, little is known regarding its putative biological functions. The presence of putative nucleic acid-, post-translational modification (PTM)-, and EJC-binding intrinsic capacities could however be useful in TDRD3's putative role as a methyl-binding effector protein in co-transcriptional mRNA processing events regulated by arginine methylation.

FIGURE 5. Schematic representation of the TDRD3 protein. The encoded TDRD3 protein consists of 744 amino acids arranged in a modular fashion. Data obtained from the NCBI Conserved Domains and the Protein Family (Pfam version 22.0) databases predict the presence of three structural domains in TDRD3: a DUF1767/OB-fold (a.a. 13-169), a UBA (a.a. 286-328), and a Tudor (a.a. 651-710) domain. The last 20 amino acids form an EJC-binding motif (see text for details).

1.5. Research questions and objectives

Alternative splicing of the human PRMT1 primary transcript can generate at least three protein isoforms differing in their first 6 to 30 N-terminal amino acids (34, 35). According to current PRMT1 structural information (25), these three N-terminal variants should all interact with the AdoMet cofactor to a similar extent and therefore, should all be enzymatically active. Therefore, based on results from previous studies (section 1.3.1), we hypothesized that production of various isoforms with unique short N-terminal sequences could have a significant impact on PRMT1's functions by providing each variant with distinct intrinsic activities and/or substrate specificities (hypothesis #1). Moreover, observations by Scorilas and colleagues suggest that the balance between these three PRMT1 alternatively spliced isoforms is altered in breast cancer cells (34).

Thus, altered expression of PRMT1 variants, each having a potential and distinct set of interactions and substrates, may result in an overall shift among PRMT1 targets in breast cancer cells, which in turn may contribute to mammary tumorigenesis.

The results presented in Chapter 2 will address hypothesis #1 by addressing the following objectives:

i. Clarifying the gene structure of PRMT1;

ii. Confirming differential splicing of PRMT1 in breast cancer;

iii. Characterizing the roles/properties of alternatively spliced PRMT1 isoforms.

Signaling of alterations in the cellular microenvironment through post-translational modification (PTM) of proteins is a fundamental mechanism of cellular homeostasis. Because a large number of post-translational modifications are possible, they must be properly 'decoded' to ensure that cells respond appropriately to biological cues (3). Modular protein domains that can recognize and 'read' specific modified amino acids in proteins are therefore crucial in relaying such PTM signals to the appropriate downstream pathways (3). While advances in the field of arginine methylation have clearly identified critical roles for PRMTs and methylated arginines in the regulation of virtually every step that control gene expression pathways, methyl-binding effector proteins that monitor these 'methyl marks' to synchronize transcriptional with post-transcriptional RNA processing events are yet to be identified (13).

TDRD3 is a yet uncharacterized methyl-binding protein (148, 158) that was first identified as a component of pre-spliceosomal complexes (163). Since TDRD3 has putative nucleic acid-, PTM-, and EJC-binding intrinsic capacities, we hypothesized that TDRD3 could play a central role in (co-transcriptional) mRNA processing events regulated by arginine methylation (hypothesis #2). This may be of significant importance since TDRD3's cellular function may also be linked to mammary tumorigenesis, as it has been identified among genes whose overexpression has a strong predictive value for poor prognosis of estrogen receptor-negative breast cancers (167).

The results presented in Chapter 3 and 4 will address hypothesis #2 by addressing the following objectives:

i. Assessing if TDRD3 plays a role as a methyl-binding protein in pre-mRNA processing;
ii. Identifying methylated splicing factors involved in these interactions.

CHAPTER 2

Alternative Splicing Yields Protein Arginine Methyltransferase 1 Isoforms with Distinct Activity, Substrate Specificity, and Subcellular Localization.

Isabelle Goulet, Gabrielle Gauvin, Sophie Boisvenue, and Jocelyn Côté.

From the Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario K1H 8M5, Canada

To whom correspondence should be addressed: Jocelyn Côté, University of Ottawa, Faculty of Medicine, Department of Cellular and Molecular Medicine, 451 Smyth Road, Room 3111, Ottawa, Ontario K1H 8M5, Canada, Tel: 613-562-5800 (ext. 8660); Fax: 613-562-5636; E-mail: jcote@uottawa.ca

Received for publication, May 29, 2007, and in revised form, August 27, 2007. Published, JBC Papers in Press, September 11, 2007.

This research was originally published in *The Journal of Biological Chemistry*. 2007; 282: 33009-21. © the American Society for Biochemistry and Molecular Biology.

Description and statement of contributions of collaborators and co-authors

This manuscript describes the influence PRMT1 isoforms' unique N-terminal sequences have on both its methyltransferase activity and its substrate specificity. It also uncovers and characterizes a functional CRM1-dependent nuclear export sequence unique to PRMT1 v2 that regulates its subcellular localization. Finally, it suggests that the relative balance of PRMT1 isoforms is altered in breast cancer.

I. Goulet and G. Gauvin (under close supervision from I. Goulet) performed all the experiments presented in this manuscript, with occasional assistance from S. Boisvenue and guidance from Dr. J. Côté. Reagents were provided as outlined in the text. I. Goulet and Dr. J. Côté (fig. 1A, F; fig. 2C, D; fig. 6C) created all figures. Dr. J. Côté and I. Goulet contributed equally to the manuscript writing, editing, and revisions prior to publication.

Abstract

PRMT1 is the predominant member of a family of protein arginine methyltransferases (PRMTs) that have been implicated in various cellular processes, including transcription, RNA processing and signal transduction. It was previously reported that the human PRMT1 pre-mRNA was alternatively spliced to yield three isoforms with distinct N-terminal sequences. Close inspection of the genomic organization in the 5'-end of the PRMT1 gene revealed that it can produce up to seven protein isoforms, all varying in their N-terminal domain. A detailed biochemical characterization of these variants revealed that unique N-terminal sequences can influence catalytic activity as well as substrate specificity. In addition, our results uncovered the presence of a functional nuclear export sequence in PRMT1v2. Finally, we find that the relative balance of PRMT1 isoforms is altered in breast cancer.

Introduction

Protein structural and functional diversity is reliant on the covalent posttranslational modification of its amino acid residues. One of these modifications, termed arginine methylation, results from the transfer of methyl groups from *S*-adenosyl-*L*-methionine to the guanidino nitrogen atom of arginine residues in protein substrates (11). The protein arginine methyltransferases (PRMTs) responsible for this process consist of a family of nine enzymes that have been classified as either type I (PRMT1-4, 6 and 8) or type II (PRMT5, 7 and 9), according to whether they promote the formation of asymmetric ω-N^G,N^G-dimethylarginines (aDMA) or symmetric ω-N^G,N'^G-dimethylarginines (sDMA), respectively. An increasing number of PRMTs' targets have recently been identified, that are involved in a broad range of cellular processes, including DNA repair, transcriptional regulation, RNA processing and signal transduction (96).

PRMT1 is the most abundant PRMT and accounts for more than 85% of the asymmetrically dimethylated arginines generated in mammalian cells (168). PRMT1 activity has mainly been linked with the regulation of protein-protein interactions in signal transduction pathways and cell growth. For example, arginine methylation of Sam68 is required for its proper localization (169) and differentially regulates its interaction with SH3 and WW domains (128). Recent observations have shown that arginine methylation of the proline-rich domains of the heterogeneous nuclear ribonucleoprotein K (hnRNP K) also regulates its interaction with the SH3 domain of the tyrosine kinase *c*-Src (132). The hnRNPs represent a major protein family that contains

aDMA *in vivo* (170) and arginine methylation has also been shown to regulate the nucleo-cytoplasmic localization of some of its members (171).

Surprisingly, very little is known about how arginine methylation is regulated. PRMT1 was first isolated through its interaction with BTG1 and BTG2, two members of the BTG/Tob family of proteins involved in negative regulation of cell growth (14, 54, 56, 58-61). BTG2 is a direct substrate of PRMT1 (55) and both BTG1 and 2 can stimulate its activity (14, 57). Recent observations have also suggested that PRMT1 activity could be regulated by its dynamic subcellular localization (44). As an example, the protein hCAF1 (CCR4-associated factor 1), a known interactor of BTG1, has been shown to form a complex with PRMT1 in nuclear speckles and to regulate its activity in a substrate-dependent manner (63).

Although all PRMTs contain a highly conserved methyltransferase domain of approximatively 310 amino acids, each PRMT has distinct protein substrate specificities. Beyond their 'core' region, each PRMT exhibits a unique N-terminal region that varies considerably in length. The function of this PRMT-specific sequence is still not well understood, but results from deletion studies suggest that it can contribute to PRMT enzymatic activity and/or protein substrate recognition. For example, deletion of the N-terminus of PRMT3, which contains a zinc-finger motif, decreases its enzymatic activity and affects its protein substrate specificity (16, 42). Similarly, deletion of the N-terminal part of Hmt1, the yeast PRMT1 homologue, significantly impairs its oligomerization and reduces its methyltransferase activity both *in vivo* and *in vitro* (172).

It was previously reported that alternative splicing of the human PRMT1 primary transcript could generate at least three protein isoforms, themselves differing at their N-terminus (34, 35). We describe here the complex genomic organization in the 5'-end of the PRMT1 gene that can produce up to seven protein isoforms expressed in a tissue-specific manner. Biochemical characterization of these seven isoforms (designated as v1 to v7) revealed that they are all active, except for v7, and that their unique N-terminal sequences can confer distinct substrate specificity. Moreover, we demonstrate that the amino acid sequence unique to v2 (encoded by exon 2) contains a CRM1-dependent nuclear export sequence (NES) that regulates its subcellular localization. Finally, we find that the relative balance of these PRMT1 splicing variants is altered in breast cancer cell lines.

Experimental procedures

Cells, reagents and antibodies

The human HeLa cervical carcinoma cell line was purchased from ATCC (Manassas, Virginia) and grown as a monolayer in DMEM medium supplemented with 1 mM sodium pyruvate, 50 IU/ml penicillin, 50 mg/ml streptomycin, and 10% foetal calf serum (Wisent). The human normal breast cell line Hs 578 Bst and the human breast cancer cell lines Hs 578T, BT-20, BT-549, MCF7, MDA-MB-231, and T-47D were procured from ATCC and grown as monolayers in complete DMEM medium supplemented with MEM non-essential amino acids and 10% foetal bovine serum (Wisent). Mouse wild type and PRMT1$^{-/-}$ ES cells were kindly provided by Dr. David Lohnes (University of Ottawa, ON, Canada) and Dr. H. Earl Ruley (Vanderbilt University Medical Center, Nashville, TN), respectively, and were maintained as previously described (69). Transfections were performed using the Lipofectamine Plus transfection reagent (Invitrogen) according to manufacturer's instructions.

Polyclonal anti-GFP antibodies were purchased from Sigma. Rabbit polyclonal antibodies against human PRMT1, Sam68, asymmetric dimethylarginine (ASYM24-25) were described previously (114, 169, 173) and were a kind gift from Dr. Stéphane Richard (McGill University, Montréal, Qc, Canada). Monoclonal antibodies against β-actin and GAPDH were purchased from Sigma and Covance, respectively.

DNA constructs

Total RNA was extracted from HeLa cells using Trizol reagent (Invitrogen) according to the manufacturer's instructions. RNA concentration was measured with a spectrophotometer, and RNA quality was determined by agarose gel electrophoresis. First strand cDNA synthesis was performed using 5 μg total RNA and the AMV Reverse Transcriptase (Promega) with an oligo-dT primer. cDNAs were amplified using PRMT1 isoform-specific oligonucleotides (v1-7) designed to introduce the amplified fragments in frame with a hexahistidine tag in the EcoR I and Xho I sites of the pET-20b expression vector (Novagen). PRMT1v4 was amplified from an EST clone (IMAGE: 387210) using v4-specific primers and PRMT1v7 was isolated using the v1,2-specific primers on HeLa cDNA. The resulting plasmids were digested with BamH I and Xho I and the released full length inserts were subsequently cloned in frame with EGFP into the Bgl II and Xho I sites of the pEGFP-N1 vector (Clontech). Using a similar strategy, full length PRMT1 isoforms were excised from the pET-20bPRMT1v1-7 constructs with EcoR I and Xho I and subcloned into the pcDNA3.1/myc-His vector (Invitrogen) using the same restriction sites. The PRMT1v2-EGFP nuclear export signal (NES) mutant, in which Val-15 and Leu-18 were replaced by alanines, was generated by using the Stratagene Quick-Change mutagenesis kit. The pGFP-Rev-NES vector was kindly provided by Dr. Stephen Lee (University of Ottawa, ON, Canada).

The expression vectors encoding GST-fused RG motif of coilin (V385 to V423) and Sam68 (G285 to P308) proteins were kindly provided by Dr. Stephane Richard (McGill University, Qc, Canada). The expression vectors encoding GST-fused RG motifs of

MRE11 (F554 to R680), SmB (P206 to L231) and SmB' (P206 to P240) have been described (71, 174). Full length fibrillarin and hnRNP A1 were amplified from HeLa cDNA and introduced in frame with a GST tag in the BamHI/Xho I restriction sites of the pGEX-4T2 vector (GE Healthcare). The DNA sequences of all constructs were verified by sequencing (StemCore Laboratories, Ottawa, ON, Canada).

RT-PCR analysis

Total human tissue cDNAs were obtained from BioChain (Hayward, California). PCR were performed in 25 μl reaction mixture using Taq DNA polymerase (Qiagen). cDNAs were amplified using the PF / PR primers (34) specific for PRMT1variant 1 to 3 or using either a forward PRMT1v4-specific primer (5'-AAATCTTCCAGCGGGGTCGCG-3'), PRMT1v5-specific primer (5'-ACTGGAGAGATGGTGTCCTGTGG-3'), PRMT1v6-specific primer (5'-AAGCTGACCAGACAAAGAGAGG-3') or a PRMT1v7-specific primer (5'-TGCATCATGGAGGAGATGCTGAAGG-3') with the reverse PR primer. Primers specific for actin (34) were used to show that an equal amount of total cDNA was used for each reaction. An initial incubation at 98°C for 2 min was followed by 35 cycles consisting of a 95°C denaturation step (30 s), a 65°C annealing step (30 s), and a 72°C extension step (30 s). A final extension step at 72°C was included for 10 min. For PRMT1 expression analysis in breast cancer cells, total RNA was extracted from cells using Trizol reagent (Invitrogen) according to the manufacturer's instructions and as described above. First strand cDNA synthesis was performed using 5 μg total RNA and the AMV Reverse Transcriptase (Promega) with an oligo-dT primer. PCR were performed as described above and GAPDH was used as a positive control. All PCR

products were electrophoresed on 2.5% agarose gels and visualized by ethidium bromide staining.

Protein purification

6xHis-fusion proteins were overexpressed in *E. coli* BL-21 cells (Stratagene) by induction with a final concentration of 1 mM isopropyl-b-D-thiogalactopyranoside (IPTG). Following induction, cells were spun down, resuspended in 10 ml of Lysis buffer (50 mM Na_2HPO_4, 300 mM NaCl, 10 mM imidazole, 20 mM 2-mercaptoethanol) and subsequently broken down by sonication (5 pulses of 15 sec at 12 W). Cell debris were discarded through centrifugation for 15 min at 10,000 x g. 6xHis-tagged proteins were then purified using Ni-NTA matrix (Qiagen) according to manufacturer's instructions and eluted from the beads using 250 mM imidazole. GST-fusion protein substrates (GST-coilin, -fibrillarin, -hnRNPA1, -MRE11, -Sam68, -SmB and -SmB') were expressed in *E. coli* BL-21 cells by induction with a final concentration of 0.1 mM IPTG. Cells were lysed by sonication as described above in 10 ml 1X PBS supplemented with a Complete protease inhibitor cocktail (Roche). The GST-fusion proteins were purified using glutathione-agarose beads (Sigma) and after extensive washes with PBS wash buffer (1X PBS, 1% Triton X-100, Complete Protease inhibitor), were eluted from the beads with 60 mM glutathione in PBS adjusted to pH 7.5. The purified tagged proteins were dialyzed against 1 liter of phosphate-buffered saline (1X PBS: 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4, pH 7.4) at 4°C, overnight. The dialysates were then concentrated using Centricon centrifugal devices (Millipore). Protein concentration was determined by using the *DC* Protein Assay reagent (Bio-Rad).

Methylation assays

Endogenous protein substrates from ES PRMT1$^{-/-}$ cells mock-treated or treated with 20 μM of the methyltransferase inhibitor adenosine-2',3'-dialdehyde (AdOx) (Sigma) for 48 h were prepared as previously described (175). Briefly, cells were harvested in 50 mM sodium phophaste buffer pH 7.5 and lysed by sonication. Cell debris were discarded through centrifugation for 10 min at 18,000 x g and cell extracts were then heat-inactivated at 70°C for 10 min. Adox-treated samples were desalted by overnight dialysis against 50 mM sodium phosphate buffer pH 7.5 at 4°C. Alternatively, ES PRMT1$^{-/-}$ cells were fractionated using the QProteome nuclear protein kit (Qiagen) following manufacturer's instructions. All obtained fractions were then dialysed against 50 mM sodium phosphate buffer pH 7.5 at 4°C, for 16 h. Protein concentration was measured by using the *DC* Protein Assay reagent (Bio-Rad). In vitro methylation of the GST-fusion proteins (5 μg) was performed at a final concentration of 25 mM Tris·HCl pH 7.4. Methylation of the endogenous protein substrates from ES PRMT1$^{-/-}$ cells (25 μg of total proteins or 10 μg of each fraction) was performed in 50 mM sodium phosphate buffer pH 7.5. Equal amounts of each purified 6xHis-tagged PRMT1 protein isoform v1 to v7 (1 µg) was added to each reaction. Methylation reactions were initiated by the addition of 0.8 μM *S*-adenosyl-L-[methyl-^{3}H]methionine in a final volume of 40 μl and were allowed to proceed at 37 °C for 2 h. All methylation reactions were resolved by SDS-PAGE. Following electrophoresis, gels were stained in Coomassie Brilliant Blue R-250 for 20-30 min, destained in a 10% methanol (v/v), 5% acetic acid (v/v) destaining solution to visualize protein bands, and then soaked in En^{3}Hance (PerkinElmer Life Sciences)

according to the manufacturer's instructions. Gels were dried *in vaccuo*, and ^{3}H-labeled proteins visualized by fluorography.

Results

Genomic structure of the human Prmt1 gene

Upon initial identification of the Prmt1 genomic locus, it was observed that three transcripts could be produced through alternative splicing of exons 2 and 3 (named v1-v3; Fig. 1*A*) and expression of these variants at the RNA level was later confirmed by an independent study (34, 35). Zhang and collaborators then predicted the existence of three additional isoforms (named v4-v6; Fig. 1*A*) by comparing available ESTs with human PRMT1 genomic sequence, but did not investigate further how these mRNAs were produced (25). Mapping of the unique sequences present at the 5'end of v4 and v6 mRNAs onto the human PRMT1 genomic locus (human Mar. 2006 hg18 assembly; NCBI Build 36.1), using the UCSC Genome Bioinformatics (http://genome.ucsc.edu/) BLAT alignment tool (176), suggested the existence of at least two additional exons upstream of the previously identified first exon. Further examination of splicing junctions in the genomic sequence using UCSC Genome Browser (177) revealed the presence of four alternative 5'-exons in the Prmt1 gene, that we have renamed e1a-d, e1d being the exon previously labelled exon 1 (Fig. 1*A*). E1c and e1d do not seem to have functional 3' splice sites (3'ss) and hence are never joined together and are utilized as alternate first exons. In contrast, exon 1b has two alternative 3'ss that can be paired with the 5'ss of e1a. In addition, the intron between these exons is sometimes retained in the mRNA (Fig. 1*A*). These splicing events would take place ~1 kb away from the previously identified exon 1 and support the presence of a transcriptional start site (TSS1) upstream of e1a.

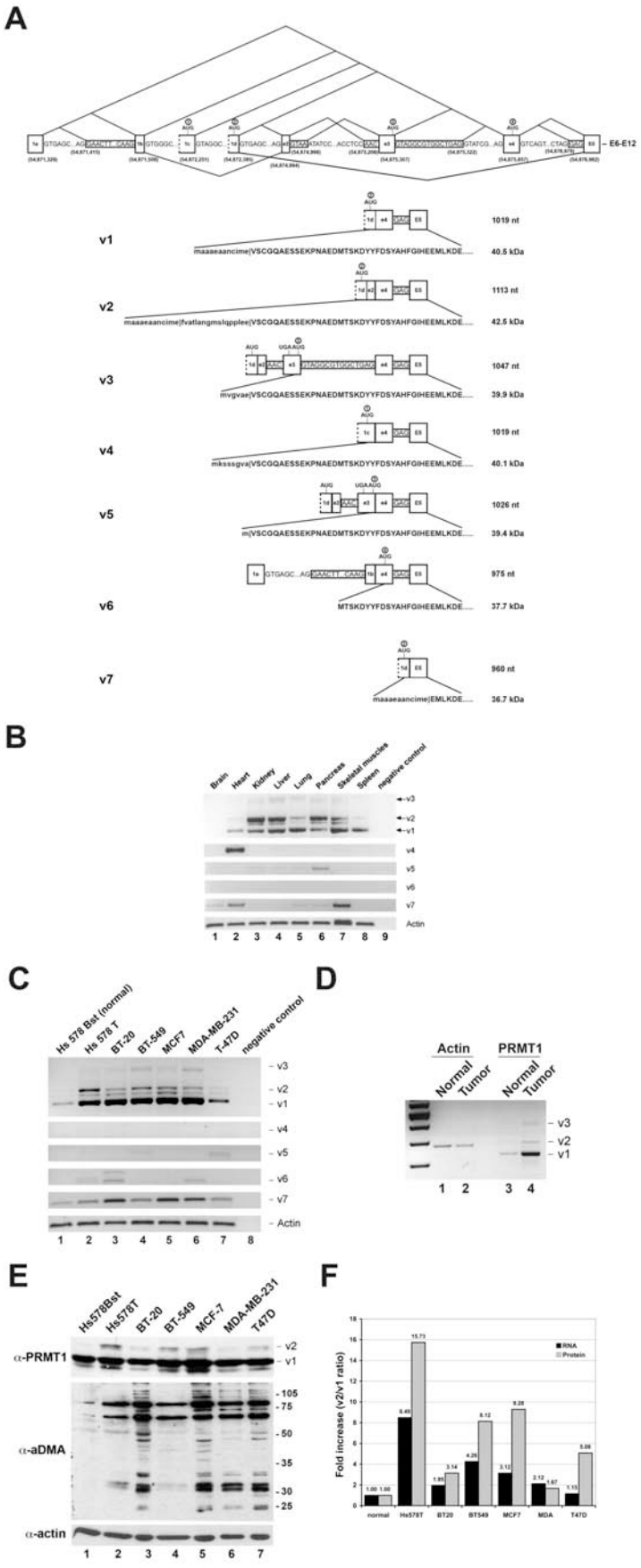


FIGURE 1. **Genomic organiza-tion and expression profile of PRMT1 isoforms.** *A,* The 5' region of the gene undergoing extensive alternative splicing has been blown-up to show details of splicing junctions. Exons are shown as boxes and sequence of intron boundaries is represented. Constitutive exons are indicated by "E" and alternative exons by "e". Genomic coordinates from UCSC Genome Browser are shown below the diagram and represent exon boundaries. Representative splicing patterns associated with each PRMT1 N-terminal isoform are shown below the diagram. *B*, PRMT1 variant 1 to 7 expression levels in normal human tissues. Total cDNA from eight human tissues were amplified by PCR using isoform-specific oligonucleotides. Actin was used as a control for equal amount of cDNA in each reaction. *C*, PRMT1 variant 1 to 7 expression levels in normal and breast cancer cell lines. Total RNA was extracted from breast cancer cells and analysed by semi-quantitative RT-PCR. *D*, Expression of PRMT1 v1 to 3 mRNA isoforms in normal and tumorigenic breast tissues, analysed as above using semi-quantitatitve RT-PCR. *E*, Total protein extracts from human normal and cancerous breast cell lines were analysed by immunoblotting with anti-PRMT1 and asymmetric dimethylarginine-specific antibodies. Immunoblotting with anti-actin antibodies is shown to confirm equal loading. *F,* v1 and v2 mRNA and protein signals were quantified by densitometry, normalized to actin levels, and plotted as fold increases in the v2/v1 ratios, relative to normal breast cells.

Indeed, a promoter (genomic coordinates 54,870,654-54,870,904) and a putative TSS (54,871,225) are predicted in this region using the "Promoter Scan" and "DBTSS" algorithms (178, 179). However, the highest score obtained for a TSS, using the Neural Network Promoter Prediction algorithm (180), maps to a site just upstream of e1d (54,872,277), which is compatible with this exon being the most represented first exon among ESTs. Preliminary results from a 5'-RACE analysis of PRMT1 mRNAs confirm the existence of this TSS2 (data not shown and Fig. 1*A*). Thus, these observations are consistent with a model where alternative promoters would dictate at least in part the choice of 5'-exon usage. The splicing patterns of alternative exons 2, 3, 4 and 5 are also more complex than initially anticipated from previous reports (34, 35), and involve the use of multiple alternative 5' and 3'ss. For example, the intron between e2 and e3 includes minor spliceosome AT-AC splice sites (181-184), in addition to a non-canonical GT-cc pair. In this case, the 5'ss used determines the reading frame, irrespective of the 3'ss to which it is paired, and thus will dictate AUG start codon choice: GT results in the use of a start codon in e3 and AT in e4, yielding protein isoforms v3 and v6, respectively. Finally, protein isoform v5 results from pairing of e1d-e2-e3 (GT-cc) but with usage of an alternative 5'ss downstream of e3 (Fig. 1*A*).

The predicted protein-coding regions (excluding 5' and 3' UTRs) of the various Prmt1 transcripts are comprised of 1,059 nt (v1), 1,113 nt (v2), 1,041 nt (v3), 1,047 nt (v4), 1,026 nt (v5), 975 nt (v6) or 960 nt (v7), and would encode 7 deduced polypeptides with a predicted molecular weight of 40.5, 42.5, 39.9, 40.1, 39.4, 37.7, and 36.7 kDa, respectively, excluding any posttranslational modifications (Fig. 1*A*).

Expression profile of PRMT1 isoforms in normal human tissues

Experiments investigating the expression pattern of PRMT1v1, v2 and v3 in different human tissues have demonstrated that these variants are ubiquitously expressed (34, 35). In order to address the expression profile of the newly isolated isoforms, we analysed the relative expression levels of the seven PRMT1 isoforms in human normal tissues by semi-quantitative RT-PCR using the expression level of the actin gene as an internal control (Fig. 1*B*). PRMT1v1 was mostly expressed in the kidney, liver, lung, skeletal muscle and spleen, whereas PRMT1v2 was predominantly detectable in the kidney, liver and pancreas. The PRMT1v3 isoform was present at comparable levels in all tested tissues. In contrast, PRMT1v4 to 7 showed a higher degree of tissue-specificity in their expression pattern. For example, v4 was detectable only in the heart, whereas v5 was mostly expressed in the pancreas, and v7 was predominantly present in the heart and skeletal muscles (Fig. 1*B*). Finally, the predicted PRMT1v6 was not detectable in the normal tissues tested under the experimental conditions used here. These results show that the expression level of the PRMT1 variants varies significantly between human tissues.

PRMT1 expression in breast cancer cells

It has been previously reported that the relative prevalence of PRMT1 isoforms 1 to 3 is different between normal and cancerous breast tissues (34). In addition, a large proportion of ESTs for PRMT1v6 were obtained from cancer cells libraries, suggesting that the balance of specific alternative isoforms might be altered in transformed cells. To assess whether the expression of PRMT1 N-terminal variants is distinct in breast cancer

cells, we examined the expression of the *Prmt1* gene in normal and cancerous breast cell lines using RT-PCR. Actin mRNA was used as a control for the amount of cDNA used (Fig. 1*C* (*bottom panel*) and Supplementary Fig. 1). We find that PRMT1 expression level is on average 14 fold higher among the breast cancer cell lines tested (Fig. 1*C*). Furthermore, the relative expression ratio of the isoforms also varied between normal and breast cancer cell lines. Specifically, the ratio of v2 over v1 mRNAs was increased on average 3.5 fold in breast cancer cell lines (Fig. 1*F*), suggesting that the v2 isoform is selectively increased relative to v1. Among the other isoforms, PRMT1v5 and v6 were detected uniquely in specific breast cancer cell lines, but not in normal breast cells (Fig. 1*C*, compare *lanes 3*, *6*, and *7*, with *lane 1*). The newly identified v7 isoform expression level was also on average 3 fold higher in the breast cancer lines tested (Fig. 1*C*, compare *lane 1* with *lanes 2-7*). Finally, an overall 9.5 fold increase in PRMT1 expression was detected in a human breast tumor sample, as compared with the adjacent normal breast tissue (Fig. 1*D*), confirming the observations made in breast cancer cell lines.

In order to evaluate whether the observed PRMT1 isoforms mRNA profiles correlated with protein levels, we immunoblotted normal and cancerous breast cell protein extracts with our anti-PRMT1 antibody (169). Since this antibody is directed against the common C-terminal domain, it is predicted that it should recognize all PRMT1 isoforms. Three major bands can be detected/resolved using this antibody in HeLa cell extracts (169). Based on its predicted molecular weight, the upper band should correspond to PRMT1v2 (42.5 kDa), and this was confirmed using a v2-specific siRNA duplex (see Supplementary Fig. 2). PRMT1v1 (40.5 kDa) migrates with the stronger

middle band (Fig. 1*E*, *top panel*, *lane 2*), and because of their very small difference in predicted sizes, it is likely that isoforms v3-5 cannot be resolved from each other and also migrate as part of this band. The identity of the lower band is more difficult to predict, but it could correspond to PRMT1v6 (37.7 kDa) and/or v7 (36.7 kDa). Using this antibody, the major middle band (labeled v1 here for simplicity), as well as a weak faster-migrating species, were observed in normal breast fibroblasts (Fig. 1*E*, *top panel*, *lane 1*). Unexpectedly, this band does not change significantly in breast cancer cells, although we can not rule out the possibility that an increase in PRMT1v1 could be masked by co-migration with other isoforms. In contrast, PRMT1v2 was clearly present at higher levels in the Hs578T, BT-549 and MCF-7 breast cancer cells (Fig. 1*E*, l*anes 2, 4* and *5*, respectively), which is consistent with the profile observed for mRNA levels (Fig. 1*F*). The intensity of the faster migrating protein species also varies slightly between the various samples, but these changes did not follow the profile observed for PRMT1v6 or v7 mRNAs. Hence, the identity of that band remains unclear.

PRMT1 is the predominant Type I PRMT in cells, and we have found that the expression profile of at least some of its splicing variant is altered in breast cancer. Thus, we hypothesized that the overall state of Type I arginine methylation should also be affected in these cancer cells. To test this hypothesis, total protein extracts were immunoblotted with asymmetric dimethylarginine-specific antibodies ASYM24 and ASYM25 (114). Strikingly, an overall increase of arginine methylated protein levels was observed in the breast cancer cell lines, as compared to the normal breast cell line (Fig. 1*E*; *middle panel*). Moreover, the asymmetric dimethylarginine-specific antibodies

immunoreacted with similar but distinct proteins between normal and cancerous breast cells, as well as between the different breast cancer cell types. Taken together, our results show that an altered PRMT1 isoforms expression profile correlates with a differential pattern of arginine methylation in breast cancer cell lines. This, in turn, suggest that misregulation of arginine methylation could contribute to the onset of breast cancer.

Unique N-terminal sequences can influence PRMT1 enzymatic activity and substrate specificity

In order to assess the contribution of the alternatively spliced N-terminus in PRMT1 function, the methyltransferase activity of each recombinant PRMT1 protein isoform (v1 to 7) was tested *in vitro*. The PRMT1 variants were expressed as C-terminally 6xHis-tagged proteins, purified, and incubated with a selection of known arginine methylation substrates (Fig. 2*A*). Following incubation in the presence of [^{3}H]-*S*-adenosyl-*L*-methionine (AdoMet) as a methyl donor, the reactions were resolved by SDS-PAGE and the methylated proteins were visualized by fluorography. In these experiments, PRMT1 variants 1 to 6 were active towards all substrates tested, with v3 and v4 being the less active isoforms overall (Fig. 2*B-E*). Remarkably, each active isoform preferentially methylated a distinct subset of protein substrates. Sam68 and SmB, for example, were better substrates for PRMT1v1 and v2, whereas hnRNP A1 was more effectively methylated by v5 and v6 (see the respective bar graph in Fig. 2*C*). Among the substrates tested, Coilin is a proposed substrate of PRMT5 and was shown to harbor sDMA *in vivo* (105, 157). Similarly, Sm proteins B and B' are targets of PRMT5 and CARM1, and accordingly were shown to contain both sDMA and aDMA (106, 121, 149, 185, 186).

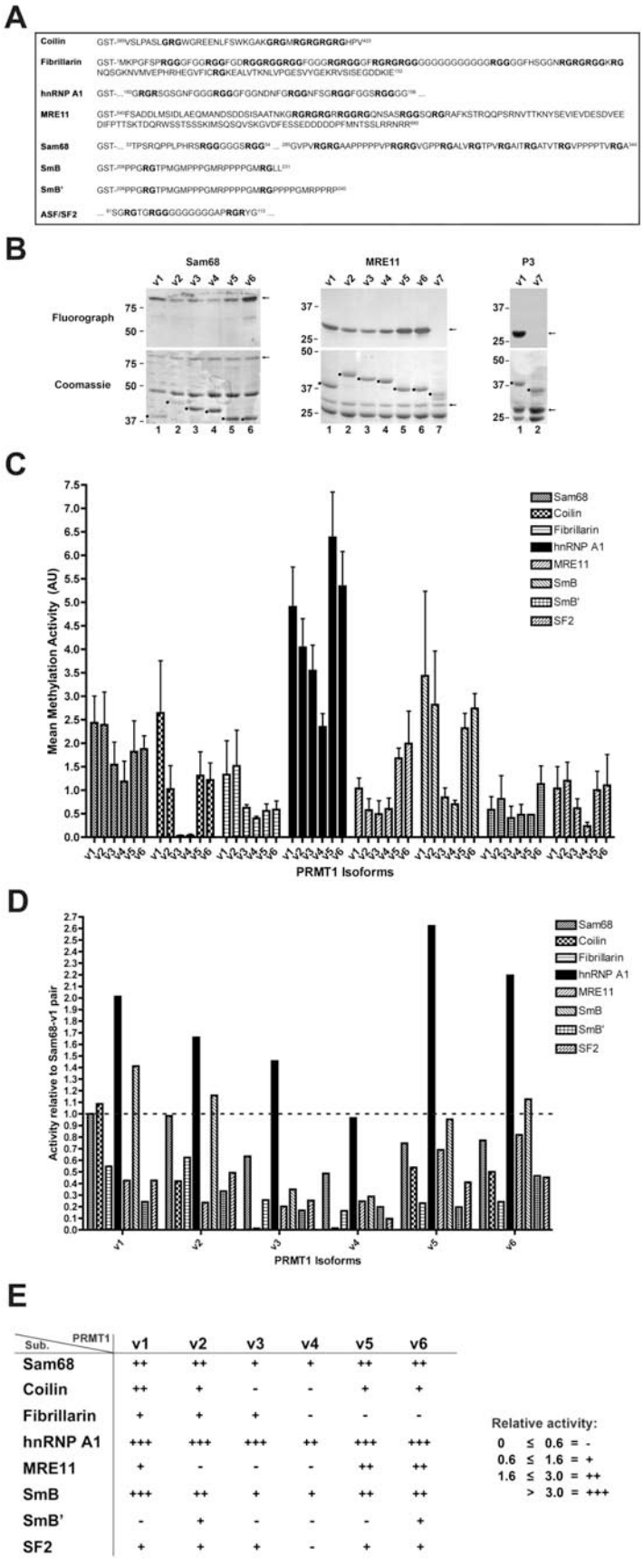


FIGURE 2. **PRMT1 isoforms enzymatic activity and substrate specificity *in vitro*.** *A,* Amino acid sequences of the known RG-rich containing methylation substrates used in our *in vitro* methylation assays. *B, In vitro* methylation assays were performed by incubating the recombinant substrates with equal amounts of purified 6xHis-tagged PRMT1 protein isoforms v1 to v7, in the presence of [^{3}H]AdoMet as a methyl donor. Methylation reactions were resolved by SDS-PAGE, gels were stained with Coomassie Brilliant Blue to visualize protein bands, and ^{3}H-labeled proteins were visualized by fluorography. All fluorographs were exposed at -80°C for 1 week. *In vitro* methylation of GST-Sam68, and MRE11 are shown as representative examples (*left and middle panels*). Methylation reactions were performed with purified PRMT1v1 and v7, using the P3 (G285 to P308) fragment of Sam68 (*right panel*). *C*, Mean methylation activity ± SEM of each isoform is expressed as an arbitrary unit derived from the band intensity normalized to the amount of isoform and substrate seen on the Coomassie stain, as determined by densitometry. These values were then plotted, grouped by substrates. Each methylation assay has been reproduced at least 3 times and one-way ANOVA statistical analysis indicated that differences in means were significant ($p<0.0001$; 95% confidence intervals). *D,* The same data is represented normalized relative to the mean methylation activity of v1 towards Sam68 (set as 1; dotted line), and grouped by PRMT1 isoforms. *E,* Table summarizing the relative *in vitro* methyltransferase activity of each PRMT1 variant towards the RG-rich substrates. + and – signs are representative of the relative activity of each isoform for each substrate, following the scale shown on the right.

Surprisingly, we have detected methyltransferase activity of specific PRMT1 isoforms toward these substrates (Fig. 2*C*-*E*). Finally, ASF/SF2 is a well-known general and alternative splicing factor, and was previously identified in two distinct proteomic screens for arginine methylated proteins (114, 118), although the specific methylase responsible for these modifications remained unknown. We report here for the first time that ASF/SF2 is a substrate of PRMT1 *in vitro* (Fig. 2*C*).

In contrast, PRMT1v7, which lacks exons 2 to 4, was completely inactive when incubated with the same selection of substrates (Fig. 2*B* and data not shown). This observation corroborated previous results suggesting important roles for helix αX (encoded by exon 4) both in cofactor binding and catalysis (25). Since PRMT1v7 is the only inactive variant and that PRMT1 dimerization is essential for its activity (25), we speculated that PRMT1v7 might regulate the activity of other PRMT1 isoforms, by acting as a "natural" dominant negative. To address this possibility, a competitive methylation assay between PRMT1 variant 1 and 7 was performed (Fig. 3*A*). Specifically, methylation assays were performed as described above with constant amounts of GST-Sam68 substrate and PRMT1v1, but in the presence of increasing concentration of PRMT1v7. Again, PRMT1v7 resulted in no detectable methylation activity and PRMT1v1 efficiently methylated GST-Sam68, when each were used alone (Fig. 3*A*, *lane 1* and *6*, respectively). However, addition of increasing amounts of PRMT1v7 did not influence the PRMT1v1 methyltransferase activity toward Sam68 in these assays (Fig. 3*A*, *lanes 2-5*). The same result was obtained when the competition assay was performed using v7 with v6, or any of the other PRMT1 isoforms (Fig. 3*B* and

data not shown, respectively). Hence, the functional significance of the v7 variant remains unknown.

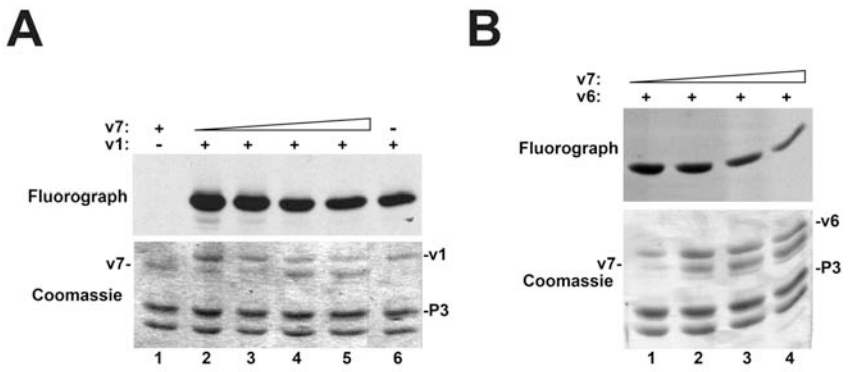


FIGURE 3. **PRMT1v7 does not act as a dominant negative for the acitvity of other isoforms.** *A*, Methylation assays were performed as described above with constant amounts of the GST-P3 substrate and PRMT1v1, but in the presence of increasing concentration of PRMT1v7. Coomassie staining is shown below the fluorograph to confirm the relative aboundance of v1 and v7. *B*, The same experiment was performed with the PRMT1v6 and v7 pair.

In order to determine if these observations can be generalized to additional endogenous substrates, hypomethylated cellular extracts prepared from PRMT1$^{-/-}$ ES cells were labeled *in vitro* using the 6xHis-tagged PRMT1 isoforms in the presence of [^{3}H]-AdoMet. Consistent with our observations on purified substrates, PRMT1v1 to v6 efficiently methylated cellular proteins, whereas PRMT1v7 showed no activity (Fig. 4*A*). As was observed with purified substrates, PRMT1v3 and v4 exhibited lower enzymatic activity towards endogenous proteins (Fig. 4*A*, *lanes 4* and *5*). However, all PRMT1 isozymes methylated an apparently similar set of cellular proteins and differences in substrate specificity were observed mainly between PRMT1v1 and v2 (marked with a "•" in Fig. 4*A*, compare *lane 1* and *2*). The absence of more drastic differences in the methylation pattern was not due to the activity of endogenous methyltransferases in the

lysates, since PRMT1$^{-/-}$ cell lysates were heat-inactivated prior to the methylation assays and lacked detectable methyltransferase activity (data not shown).

We reasoned that reducing the complexity of the protein lysates used as substrates would result in an increased PRMT/substrates ratio, which in turn might help reveal differences in specificity between the isoforms. PRMT1$^{-/-}$ ES cells were fractionated into cytoplasmic, nuclear and insoluble fractions (using the QProteome nuclear protein kit from Qiagen). Quality of cellular fractionation was assessed by western blotting using antibodies against the nuclear protein Sam68 and the cytoplasmic protein GAPDH (Fig. 4*B*). All obtained fractions were dialysed against sodium phosphate buffer and used for *in vitro* methylation assays as above, with purified PRMT1v1 and v6 enzymes. Again, no major differences in the profile of methylated bands were observed in the total (T) and cytoplasmic (C) cell extract fractions for each isoform (Fig. 4*C*, compare *lanes 1-2* with *lanes 5-6*, respectively). Strikingly, drastic differences were found in the methylation profile of PRMT1v6 when incubated with the nuclear and insoluble fractions (Fig. 4*C*, compare *lanes 3-4* with *lanes 7-8*). Taken together, these results show that the unique sequences at the N-terminus of PRMT1 isoforms can influence substrate specificity.

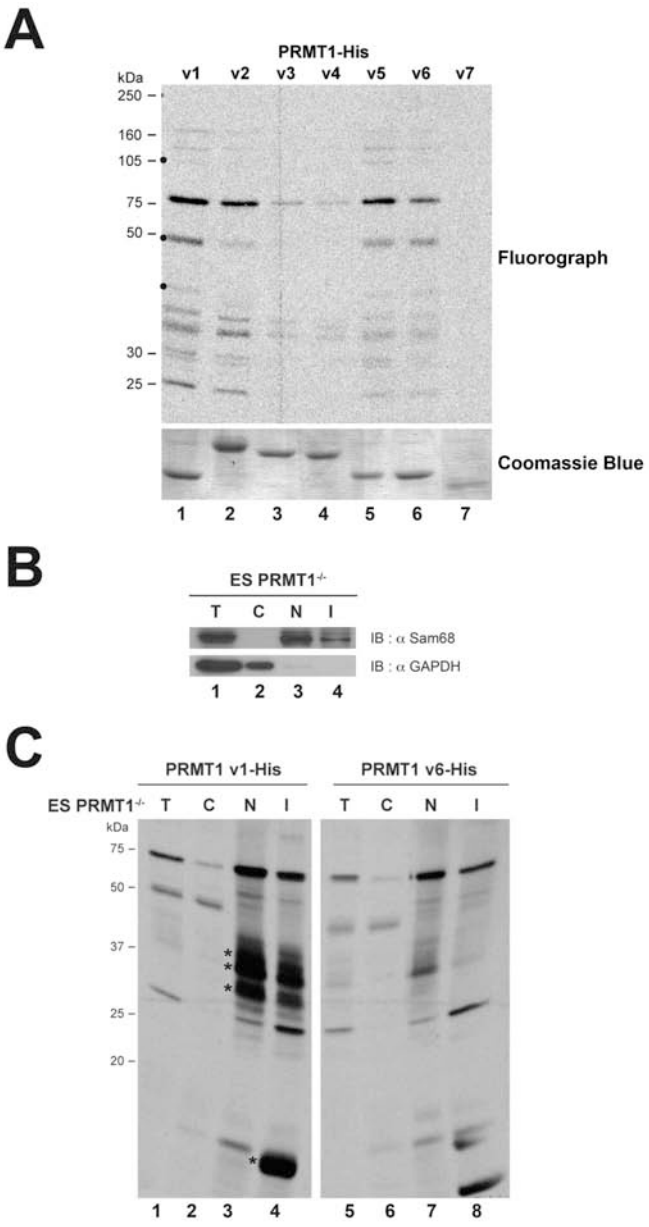


FIGURE 4. **PRMT1 isoforms have distinct substrate specificity towards endogenous substrates.** *A*, 15 μg of endogenous protein substrates from mouse PRMT1$^{-/-}$ ES cells were incubated with equal amounts of purified 6xHis-tagged PRMT1 protein isoform v1 to v7, in the presence of [^{3}H]AdoMet. All methylation reactions were resolved by SDS-PAGE, gels were stained in Coomassie to visualize protein bands, and ^{3}H-labeled proteins were revealed by fluorography. Fluorographs were exposed at -80°C for 1 week. Bands that show differences between isoforms are indicated with a "•". *B*, Cellular fractionation was assessed by western blot with antibodies against the nuclear protein Sam68 and the cytoplasmic protein GAPDH. *C*, Mouse PRMT1$^{-/-}$ ES cell extracts were fractionated using the QProteome nuclear protein kit (Qiagen). All obtained fractions were then dialysed against 50 mM sodium phosphate buffer pH 7.5. *In vitro* methylation assays were performed as described in *A* with 10 μg of total proteins. Fluorographs for Total (T), cytoplasmic (C), nuclear (N) and nuclear insoluble (I) fractions are shown. Bands that show differences between isoforms are indicated with a "*".

N-terminal sequences of PRMT1 can influence subcellular localization

Since cellular fractionation has uncovered differences in substrate specificity between certain PRMT1 isoforms, we next wanted to investigate the role of unique N-terminal sequences on PRMT1 intracellular distribution. Plasmid constructs expressing the different PRMT1 isoforms fused to a C-terminal EGFP tag were generated and used to transiently transfect HeLa cells grown on glass coverslips. Intracellular distribution of each PRMT1-GFP isoform was assessed 36 h post-transfection using fluorescence microscopy (Fig. 5*A*). Equal expression of each PRMT1-GFP fusion was confirmed by anti-GFP immunoblotting (Fig. 5*B*). Cells expressing PRMT1v3, v4, v5 and v6 showed an even distribution of the GFP fusion proteins between the nucleus and the cytoplasm (Fig. 5*A*, *respective panels*). Cells expressing PRMT1v1 and v7, however, presented a more intense nuclear staining (Fig. 5*A*, *respective panels*). Strikingly, PRMT1v2 expression resulted in a predominantly cytoplasmic staining, concentrated around the nuclear compartment (Fig. 5*A*). These observations were quantified by plotting the fluorescence intensity distribution of representative PRMT1-GFP fusions signals, relative to the exclusively nuclear signal distribution of DAPI (Fig. 5*A*, *graphs next to each series of panels*). Intracellular localization was also confirmed using confocal microscopy (data not shown).

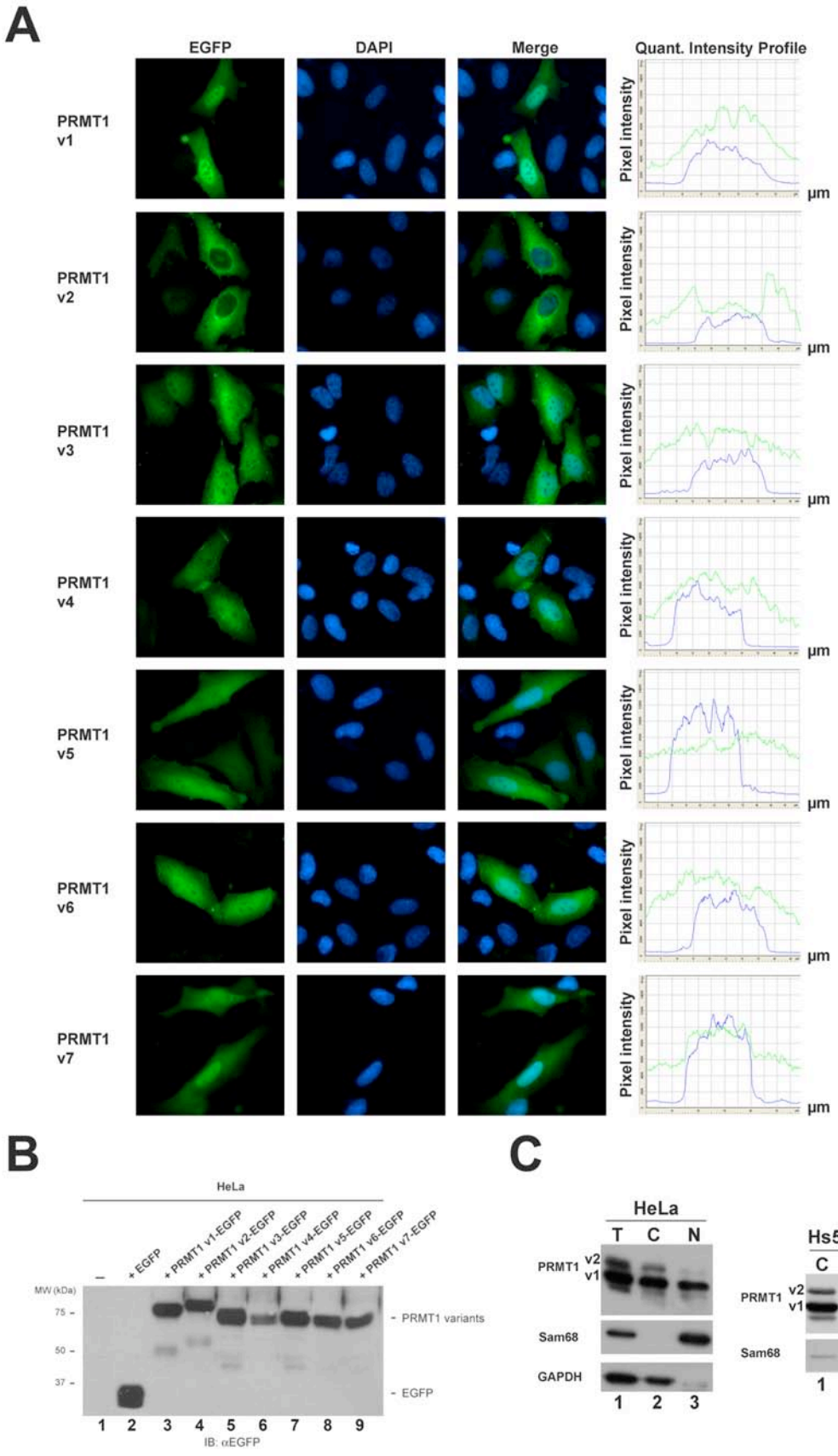

A
EGFP
DAPI
Merge
Quant. Intensity Profile
PRMT1 v1
PRMT1 v2
PRMT1 v3
PRMT1 v4
PRMT1 v5
PRMT1 v6
PRMT1 v7
Pixel intensity
µm
B
HeLa
– + EGFP + PRMT1 v1-EGFP + PRMT1 v2-EGFP + PRMT1 v3-EGFP + PRMT1 v4-EGFP + PRMT1 v5-EGFP + PRMT1 v6-EGFP + PRMT1 v7-EGFP
MW (kDa)
75
50
37
- PRMT1 variants
- EGFP
1 2 3 4 5 6 7 8 9
IB: αEGFP
C
HeLa
T C N
PRMT1 v2 v1
Sam68
GAPDH
1 2 3
Hs578T
C N
PRMT1 v2 v1
Sam68
1 2

FIGURE 5. **PRMT1 N-terminal unique sequences affect intracellular localization.** HeLa cells were transiently transfected with C-terminally EGFP-tagged PRMT1variant 1 to 7 for 36 h. *A*, Cells were fixed with 4% paraformaldehyde, stained with DAPI and viewed by fluorescence microscopy. Fluorescence intensity of a representative cell for each isoform was analysed using the AxioVision digital imaging software (Zeiss). Green lines represent the pixel intensity distribution of the GFP signal and blue depicts the profile of the exclusively nuclear DAPI staining. *B*, Total cell lysates of HeLa cells transiently expressing the EGFP-tagged PRMT1 were resolved by SDS-PAGE, transferred on a PVDF membrane and immunoblotted with anti-GFP antibodies. *C*, HeLa (*left panel*) and Hs578T cells (*right panel*) were fractionated using the QProteome Nuclear Proteins Kit (Qiagen). Subcellular distribution of endogenous PRMT1v1 and v2 between the cytoplasmic (C) and nuclear (N) fractions was analysed by Western blotting using anti-PRMT1 antibodies. Immunoblots with antibodies directed against the nuclear protein Sam68 and/or against the cytoplasmic protein GAPDH are shown to confirm cellular fractionation.

In order to confirm that the observed differential subcellular localization of PRMT1 isoforms was not due to overexpression and/or the GFP tag, we used cellular fractionation to assess the distribution of the endogenous PRMT1 variants 1 and 2. HeLa cells were fractionated as above, and an equal proportion of each fraction was resolved by SDS-PAGE, transferred to a nitrocellulose membrane and immunoblotted using the polyclonal anti-PRMT1 antibodies described above. Consistent with the distribution observed for transfected PRMT1v2-GFP, endogenous PRMT1v2 is detected at higher level in the cytoplasmic fractions of HeLa and Hs578T cells (Fig. 5*C*, *left and right panel*, respectively). Taken together, these results show that the alternatively spliced N-terminus of PRMT1 can direct its intracellular localization.

The alternative PRMT1 exon 2 encodes a leucine-rich nuclear export signal

Because PRMT1v2 is the only isoform predominantly located in the cytoplasm, we sought to identify the amino acid sequence responsible for this effect. We focused our attention on the sequences encoded by exon 2 since PRMT1v2 is the only isoform harboring this exon (Fig. 6*A*). The amino acid sequence encoded by exon 2 was submitted to the NetNES 1.1 prediction server (187), which is a web interface for an

algorithm designed to predict the presence of leucine-rich nuclear export signals (NESs) in peptide sequences. Specifically, this algorithm combines the predictive power of neural networks and hidden Markov models (NN and HMM in Fig. 6*B*, respectively), to derive a NES score, which is deemed significant if the output value is above a cut-off threshold of 0.5. Using this predictor, a putative leucine-rich NES (^{15}VATLANGMSL^{24}) was identified within the 18 amino acid sequence encoded by exon 2 (Fig. 6*B*). Leucine-rich NESs are short sequence motifs, first identified in the HIV-1 Rev protein, that are recognized by the nuclear export receptor CRM1 (188-190). Alignment of the exon 2 sequence with various known NES sequences demonstrated a very good match to the consensus ($Fx_{2-3}Fx_{2-3}FxF$; where F denotes conserved hydrophobic residues) (Fig. 6*C* and (190)). To confirm that the cytoplasmic localization of PRMT1v2 was mediated by this exon 2-encoded NES sequence, a mutated EGFP-tagged PRMT1v2, in which the first two conserved hydrophobic residues were replaced by alanines (V15A/L18A), was expressed as before in HeLa cells and analyzed by fluorescence microscopy (Fig. 6*D*). In this experiment, wild-type PRMT1v2 localized to the cytoplasm as expected (Fig. 6*D, panels a-c*). In contrast, the V15A/L18A NES mutant was distributed equally between the cytoplasm and nucleus (*panels d-f*), consistent with some of the protein now being retained in the nucleus. Leptomycin B (LMB) is a drug that inhibits CRM1-dependent nuclear export by forming a covalent bond with CRM1, hence preventing its interaction with other proteins to be exported. Hela cells transiently expressing PRMT1v2 were treated with 10 μM LMB for 1 h and analyzed using fluorescence microscopy. Following this treatment, PRMT1v2-EGFP was partly retained in the nucleus (Fig. 6*D, panels g-i*). Nuclear accumulation of a control NES sequence, GFP-Rev-NES, was also observed in

the presence of LMB (Fig. 6*D, m-o*). Hence, our results have uncovered the presence of a functional NES sequence in the alternative exon 2, targeting the PRMT1v2 isoform to the cytoplasm.

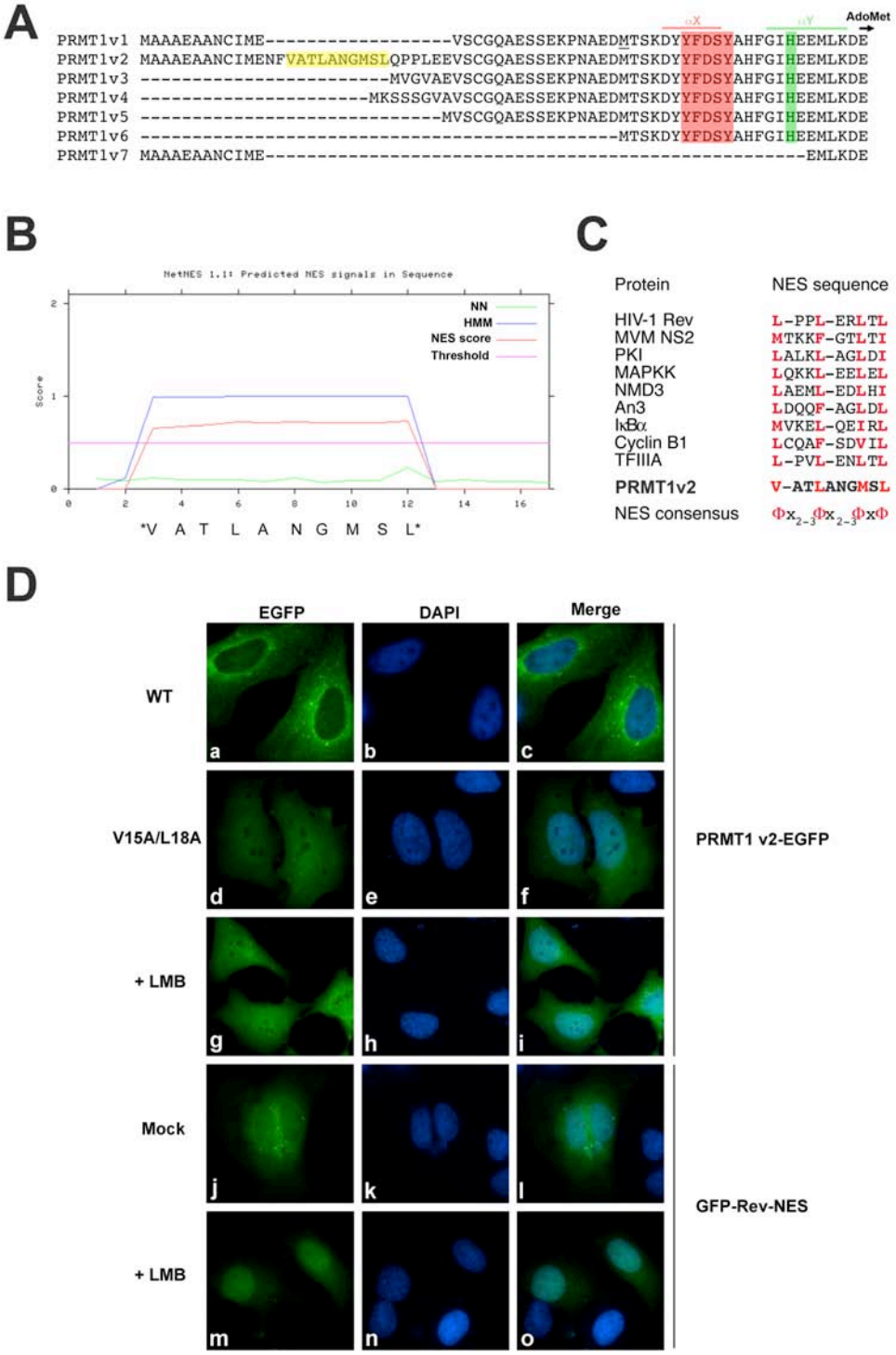


FIGURE 6. **Identification of a CRM1-dependent nuclear export signal in PRMT1v2.** *A*, Multiple sequence alignment showing the N-terminal unique domains of PRMT1v1-7. The position of αX and αY alpha helices, as well as the beginning of the AdoMet binding domain, is indicated above the sequences. Conserved YFxxY motif and H45, both predicted to make contacts with AdoMet, are highlighted in red and green, respectively. The NES sequence identified in PRMT1 v2 is highlighted in yellow. *B*, The NetNES algorithm predicts a putative leucine-rich nuclear export signal (NES) encoded by exon 2 in PRMT1v2. Scores obtained from the neural network (NN) and hidden Markov model (HMM) calculations are plotted in green and blue, respectively. The combined score obtain from the NetNES algorithm is plotted in orange and cut-off treshold is shown as a pink line. Only the portion of the exon 2-encoded sequence giving scores above treshold is shown. *C*, The putative PRMT1v2 NES sequence shows a good degree of conservation when aligned with other known CRM1-dependent NESs (modified from (190)). The conserved hydrophobic residues are shown in red. *D*, Essential hydrophobic residues Val-15 and Leu-18, were replaced by alanines using the Stratagene QuickChange mutagenesis method. HeLa cells were transiently transfected with PRMT1v2-EGFP (*a-c*) or PRMT1v2(V15A/L18A)-EGFP (*d-f*) for 36 h, then fixed with 4% paraformaldehyde, stained with DAPI and viewed by fluorescence microscopy. HeLa cells transiently expressing a vector encoding for GFP-Rev-NES (*j-o*) as a control or PRMT1v2-EGFP (*a-c, g-i*) were left untreated (*a-c, j-l*) or treated with 10 μM Leptomycin B for 1 h (*g-i, m-o*). Fluorescence microscopy was performed as described above.

Supplemental Figures

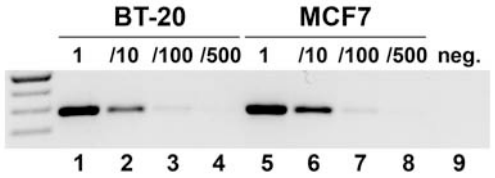


FIGURE S1. **PCR amplification conditions are in the linear range.** PCR reactions were performed with serial dilutions of the cDNA samples from BT-20 and MCF-7 cell lines, using primers specific for actin.

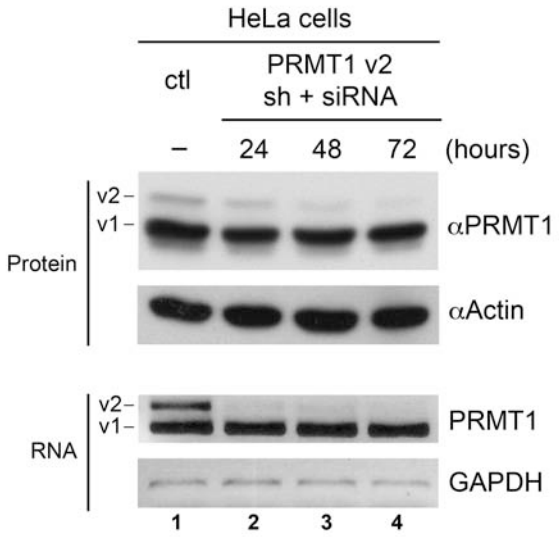


FIGURE S2. **Identification of the band corresponding to PRMT1v2 in PRMT1 immunoblots.** A PRMT1v2-specific siRNA duplex and shRNA expression plasmid were cotransfected in HeLa cells. 24, 48, and 72-h posttransfection, cells were harvested in either RIPA buffer (protein) and Trizol reagent (RNA). 50 μg of proteins were resolved by SDS-PAGE, and immunoblotted with polyclonal antibodies against the C-terminal domain of PRMT1. Immunoblotting with anti-actin antibodies was used to confirm equal loading (*upper panels*). PRMT1v2 mRNA silencing efficiency was determined by RT-PCR as described above. GAPDH was amplified to confirm that an equal amount of cDNA was used for each reaction (*lower panels*).

Discussion

In this study, we describe the complex genomic organization in the 5'-end of the PRMT1 gene that can produce up to seven protein isoforms expressed in a tissue-specific manner. Biochemical characterization of these seven isoforms revealed that they are all active, except for v7, and that their unique N-terminal sequences can confer distinct substrate specificity. Moreover, we demonstrate that the amino acid sequence unique to v2 contains a CRM1-dependent NES that regulates its subcellular localization. Finally, we report the differential expression of these PRMT1 splicing variants in breast cancer cell lines.

Complex genomic structure at the 5' end of the PRMT1 gene leads to the expression of multiple PRMT1 isoforms

The human PRMT1 gene locus was originally named *Hrmt1l2* because of its high sequence similarity with yeast *Saccharomyces cerevisiae Hmt1* (35). This early study had identified three different mRNAs (termed v1-v3) varying at their 5' end, likely arising from the same gene through alternative splicing. The existence of v1-v3 at the RNA level was further confirmed in a later study where the genomic locus was more precisely mapped and sequenced (34). Close inspection of available ESTs and comparison with complete genomic sequences strongly suggested the existence of previously unidentified exons upstream from the exon harboring the AUG start codon for v1-v3 (25). We reported here the detailed genomic organization of ~ 1 kb of sequence upstream from that previously identified AUG that can yield, through a complex pattern of alternative splicing, at least 7 different protein isoforms (Fig. 1*A*). Among these novel isoforms, v4

is the only mRNA where translation initiation actually starts at a new AUG located in the upstream 5' exon. For isoforms v3 and v5, inclusion of alternative exons e2 and e3 in the message introduces an in-frame STOP codon (UGA) and results in translation initiating at the next AUG in exon 3. The occurrence of in-frame premature STOP codons can sometimes result in mRNAs getting targeted and degraded through the so-called "nonsense-mediated mRNA decay" or NMD pathway (191). This pathway is thought to be triggered when such in-frame STOP codon is located ≥ 50-55 nt upstream of an exon-exon junction (191). Hence, following this rule, the mRNAs for v3 and v5 should not trigger the NMD pathway since the in-frame STOP codon is located at 48 and 33 nt, upstream of the closest exon-exon junction, respectively. However, recent developments suggest that this rule does not necessarily apply to every mRNA and that whether or not an mRNA is subjected to NMD needs to be determined experimentally (192-194).

According to EST databases, a number of different mRNAs can lead to the production of the v5 and v6 proteins, from various possible pairing combinations of exons e1a, e1b, e2 and e3 (Fig. 1*A*). The presence of these alternative 5' exons in mRNAs, as detected by RT-PCR (Fig. 1*B*, 1*C* and data not shown) strongly argues for the existence of at least one additional TSS upstream of exon 1a. However, we can not rule out the existence of TSSs associated with each individual alternative 5' exon and further studies will be required to determine precisely the contribution of alternative promoters and/or alternative splicing of 5' exons in the regulation of PRMT1 gene expression. Interestingly, these newly identified alternative splicing patterns would lead to the production of mRNAs with distinct 5' UTRs (e.g. v3, v5 and v6). These unique 5' UTR

sequences might contribute to regulate the expression of these mRNAs post-transcriptionally, by affecting their stability and/or translation efficiency, a phenomenon that looks to be more common than previously appreciated (195).

PRMT1v1-v3 represent the vast majority of human ESTs, with respectively, 251, 31 and 113 representative clones listed in the most recent Aceview database. As was reported previously (34), these 3 major isoforms are present in every tissue we have tested, although the relative amount of each isoform does vary between tissues (Fig. 1*B*). In our hands, the expression of all PRMT1 isoforms was very low in brain, although we have been able to amplify the major isoforms by increasing the PCR cycle numbers or the amount of input cDNA (data not shown). This observation is consistent with the original Northern blot results of Scott *et al.*, that showed strong expression of PRMT1 in fetal brain but weak expression in the adult tissue (35). This profile has also recently been confirmed at the protein level in mice, where PRMT1 expression was found to be high during mouse embryonic development, but then decrease rapidly following birth (196).

Expression of the minor isoforms is more restricted to specific tissues, with PRMT1v4 being found only in heart and v5 only in pancreas, while v7 was expressed mainly in heart and skeletal muscles (Fig. 1*B*). Interestingly, Scott *et al.* noted the presence of smaller bands in adult heart and skeletal muscles on Northern blots (35); these likely correspond to v4 and/or v7. Finally, we did not detect the presence of PRMT1v6 in any normal tissues tested, even at higher cycle numbers, but did detect expression in specific breast cancer cell lines (Fig. 1*C* and see below).

Unique N-terminal sequences affect enzymatic activity and substrate specificity

PRMT1 alternative isoforms v1 and v2 are conserved in mouse and it was previously shown that they can methylate different, but overlapping sets of proteins in PRMT1$^{-/-}$ ES cell extracts (175). Intriguingly, addition of an N-terminal His tag to these variant enzymes abolished the differences in substrate specificity (175). Using a similar assay, we now report that differences in substrate specificity can also be observed between human PRMT1v1 and v2, and in addition, PRMT1v3 and v4 display reduced enzymatic activity (Fig. 4*A*). However, the differences were not as drastic as those observed with the mouse isoforms (175), which could be a consequence of using mouse ES cell extracts as a source of substrates in combination with human PRMT1 isozymes. In addition, Pawlak and colleages were producing the PRMT1 isoforms in insect cells, which could reflect the need for post-translational modifications for the full spectrum of differences to be observed. Nevertheless, significant differences were revealed between PRMT1v1 and v6 upon fractionation of protein extracts used for the methylation assays (Fig. 4*C*). For one, this could be explained by a lower complexity of the protein extracts, which likely reflects more accurately the enzyme:substrate ratio found *in vivo*. Alternatively, it is possible that the fractionation protocol permits the extraction and solubilization of specific substrates that were not present in the sonicated extracts used in Fig. 4*A*. This would be consistent with the fact that PRMT1 is known to methylate many proteins associated with the nuclear matrix (e.g. hnRNPs) and chromatin.

Alignment of known mammalian PRMTs revealed that the N-terminal domains are the more divergent between various PRMTs. Although the function of these PRMT-

specific sequences remains unclear in most cases, it was suggested that they may contribute to enzymatic activity and/or substrate specificity (e.g. see (16, 42)). According to the PRMT1 3D crystallographic structure (25), the N-terminal sequences would be juxtaposed to an already fairly exposed alpha helix (αY), located in space between the AdoMet and substrate binding pockets (see Fig. 6*A* and (25)). In a structural analysis of the PRMT3 enzyme, an additional alpha helix (termed αX) was observed N-terminal to αY and shown to be important for cofactor binding and catalysis (26). Furthermore, deletion of the N-terminal part of Hmt1, the yeast PRMT1 homologue, significantly impairs its oligomerization and reduces its methyltransferase activity (172). This is consistent with PRMT1v7 showing no detectable activity in our hands (Figs. 2-3), since this isoform does not harbor the αY helix and the invariant YFxxY motif (25, 26). As our *in vitro* competition assays suggest that PRMT1v7 does not act as a dominant negative (Fig. 3 and data not shown), the functional significance of having a methylase-dead version of PRMT1 in cells remains unclear. Nevertheless, we can not rule out the possibility that v7 might behave differently in cells or could be highly specific for a substrate(s) not present in mouse ES cell extracts.

In contrast, the αX and αY helices, as well as the actual AdoMet-binding domain, are all present in PRMT1v1-v6, which suggest that, at least according to existing structural information, these isoforms should all interact with the AdoMet cofactor to a similar extent. Interestingly, no X-ray diffraction data was obtained for the extreme N-terminal domain of the protein crystallized by Zhang and colleagues, which is usually an indication that this portion of the structure is highly mobile in solution (25). Hence, we

envisage a model where the extreme N-terminal sequences would fold back and make crucial contacts with the substrate during catalysis. This could in turn provide a molecular mechanism for the differences in overall activity and substrate specificity observed between the various PRMT1 N-terminal variants. Since PRMT1 is known to interact relatively stably with its substrates, this model would predict that PRMT1 N-terminal variants should interact with different subsets of proteins in cells. We have performed a large scale protein-protein interaction screen with purified His-tagged PRMT1v1 and v2 using Invitrogen's ProtoArray® Human Protein Microarray v3.0. In support of the above hypothesis, a number of proteins interacted with both isoforms, but a subset showed reproducible preference for either v1 or v2 (J. Côté, unpublished data). Hence, this approach, that looks at thousands of interactions in one experiment, supports our *in vitro* methylation assay data and strongly suggest that N-terminal unique sequences can influence PRMT1 protein-protein interactions.

Identification of a functional NES in PRMT1 alternative isoform v2

The current study uncovered a functional CRM1-dependent NES in PRMT1 isoform v2 (Figs. 5-6). This means that by regulating the alternative splicing of PRMT1 exon 2, cells can modulate the nuclear versus cytoplasmic pools of PRMT1 isoforms that possess distinct protein-protein interaction profiles and substrrate specificity. We have mapped this NES to amino acids 15-24 of PRMT1v2 (^{15}VATLANGMSL^{24}), which matches the consensus derived from the sequence of multiple, experimentally tested, functional NESs ((190) and Fig. 6*C*). The use of LMB or the expression of NES mutants (V15A/L18A) resulted in v2 being localized in both the nucleus and cytoplasm (Fig 6*D*), suggesting that

PRMT1v2 normally enters the nucleus and is then efficiently exported to the cytoplasm through its CRM1-dependent NES. This is in agreement with a recent study showing that PRMT1 is a highly mobile protein in both the nucleus and cytoplasm, and can relocalize between compartments upon general inhibition of methylation (44). Our results might also provide an explanation for some of the discrepancy observed in the literature regarding the intracellular localization of PRMT1 (18, 44, 63, 69, 168, 169). For instance, depending on the antibodies used to detect the endogenous PRMT1, only specific isoform(s) may be recognized. Alternatively, the use of constructs with N-terminal tags or with truncated N-terminal domains might also result in differences in localization. Most strikingly, it was recently reported that PRMT1 localized to nuclear splicing speckles, a sub-nuclear compartment enriched in splicing factors (63). However, we have not observed the localization of any of the PRMT1 isoforms to nuclear speckles and we have found the localization of PRMT1 isoforms to be actinomycin D-independent (data not shown). Arginine methylation by PRMT1 was shown to regulate the intracellular localization of many substrates (e.g. see (101, 169, 197) and references therein), and, at least for RIP140, this effect seems to be mediated by PRMT1 directly modulating interaction with the nuclear export receptor CRM1 (101). This suggest the possible existence of a regulatory loop involving PRMT1 and it will be interesting to investigate how the various isoforms (and in particular v2) respond to changes in arginine methylation in cells.

Arginine methylation and cancer?

A number of studies have provided evidence linking arginine methylation with antiproliferative signaling or with a protective role against cancer (57, 59, 198-204). However, these observations may not be applicable to all PRMTs and all cancer types. For example, the fact that PRMT1 and CARM1 can act as coactivators of nuclear receptors makes them likely candidates to be overexpressed in hormone-dependent cancers, including breast cancer (96). Indeed, it has been found that CARM1 overexpression correlates with androgen independence in human prostate carcinomas (205, 206). It has also been observed that increased methylation of a PRMT6 substrate correlated with breast tumors of higher metastatic potential (207, 208). Consistent with these predictions, we observed that overall PRMT1 mRNA levels are increased in a panel of breast cancer cell lines and in a human breast tumor (Fig. 1*C* and 1*D*, respectively). Moreover, the relative balance of alternatively spliced PRMT1 isoforms is altered in these cell lines when compared to immortalized normal breast cells. Among the major isoforms, the PRMT1v2 mRNA, which encodes for an enzyme that is mostly localized to the cytoplasm, is selectively increased relative to v1 (Fig. 1*F*). Most strikingly, among the minor variants, PRMT1v7 was overexpressed in all of the breast cancer lines tested and PRMT1v6 was detected in BT-20 and MDA-MB-231 (Fig. 1*C*). A second RT-PCR product of slightly higher size was also reproducibly detected in BT-20 cells using v6-specific primers (i.e. a forward primer specific for exon 1a). The size of that band would be compatible with an EST-predicted mRNA containing alternative exons e2 and e3, in addition to 1a and 1b (see Fig. 1*A*), but further experiments will be required to confirm the identity of this product.

At the protein level, immunoblotting with our polyclonal antibody against PRMT1 did not reveal major differences in the level of v1 between normal and breast cancer cell line extracts. This suggests that translational and/or post-translational mechanisms (like e.g. protein stability) may be involved in regulating PRMT1v1 protein level in cancer cells. However, we can not rule out the possibility that differences in PRMT1v1 level could be masked by the fact that our polyclonal antibody also recognizes other isoforms that likely co-migrates with v1 because of their very similar molecular weights. In contrast, an increase in the relative abundance of PRMT1v2 is clearly observed in breast cancer cell lines (Fig. 1*E* and 1*F*). Although seen in every breast cancer cell lines tested, this increase in the v2 protein relative to v1 is most prominent for the Hs578T, BT-549 and MCF-7 cell lines (see Fig. 1*F*). This correlates well with the relative mRNA expression profiles observed for these cell lines, as well as for the BT-20 cell line, which shows a smaller increase both at the level of mRNA and protein. In contrast, the levels of the v2 mRNA and protein do not correlate well in the MDA-MB-231 and T-47D cell lines, suggesting the contribution from additional regulatory mechanisms, at the translational or post-translational level, in these breast cancer cell lines.

Finally, overall levels of aDMA-containing proteins were increased in all the breast cancer cell lines tested (Fig. 1*F*). Although this profile does not correlate with the cell line-specific overexpression of PRMT1v2, other type I PRMTs could also contribute to this increase in aDMA arginine methylation. Indeed, as mentioned above, CARM1 and PRMT6 were found to be overexpressed in certain cancers, and we have also found these PRMTs to be expressed at higher levels in the panel of breast cancer cell lines used in

this study (unpublished data, *see* Supplemental Data). Further work will be required to address the functional significance of PRMT overexpression in various cancers, as well as the general role(s) that arginine methylation pathways might play in cellular transformation.

In addition to PRMT1, alternative splicing was also reported for CARM1 and PRMT7 pre-mRNAs (123, 209), although extensive biochemical characterization of the respective PRMT isoforms generated is still lacking. As we have shown in the current study, alternative splicing can lead to the production of PRMT variants with a wide array of different biochemical and regulatory properties. It will be important to explore this additional level of complexity in the expression of other PRMTs in future studies.

Acknowledgments

This work was supported by a research grant from The Cancer Research Society Inc. (CRS) to J.C.. I.G. is supported by a scholarship from Fonds de la recherche en santé du Québec (FRSQ). J.C. is a recipient of a Canada Research Chair (Tier 2) in RNA metabolism. We wish to acknowledge the contribution of Josée Deschâtelets, Helina Tadesse and Huda Cathum in the early phases of this work, and Benoit Paquette for his help with microscopy. We would like to thank Helina Tadesse and Dr. Julie Deschenes-Furry for critically reading the manuscript.

Supplemental Data

Experimental procedures

Cells and antibodies

The various human (non-breast) cancer cell lines (Table I) were kindly provided by Dr. John Bell (University of Ottawa, Ottawa, ON, Canada) and grown as monolayers in complete DMEM medium for adherent cells or complete RPMI medium for non-adherent cells, supplemented with MEM non-essential amino acids and 10% foetal bovine serum (Wisent). The human breast cancer cell lines were procured from ATCC and grown as described elsewhere (117).

Rabbit polyclonal antibodies against human PRMT1, human PRMT5, asymmetric dimethylarginine (aDMA), and symmetric dimethylarginine (sDMA) were described previously (105, 114) and were a kind gift from Dr. Stéphane Richard (McGill University, Montréal, Qc, Canada). Monoclonal antibodies against β-actin were purchased from Covance. Rabbit polyclonal anti-CARM1 and -PRMT7 were purchased from Upstate Biotechnology (Upstate, NY). Anti-PRMT6 antibodies were purchased from Imgenex (San Diego, CA).

Table SDI. Cancer cell lines used in this study.

Cell line	Tissue of origin	Cancer type
K562	Bone marrow	Chronic myelogenous leukemia
Hs 578 Bst	Breast	Normal fibroblasts
Hs 578 T	Breast	Ductal carcinoma
BT-20	Breast	NOS carcinoma
BT-549	Breast	Invasive ductal carcinoma
MCF-7	Breast	NOS adenocarcinoma, pleural effusion
MDA-MB-231	Breast	NOS adenocarcinoma, pleural effusion
T-47D	Breast	Infiltrating ductal carcinoma
SF-539	Brain, Central Nervous System	Glioblastoma
SNB-19	Brain, Central Nervous System	Glioblastoma
U87-MG	Brain, Central Nervous System	Glioblastoma
HCT-116	Colon	Colorectal carcinoma
786-O	Kidney	Renal cell adenocarcinoma
EKVX	Lung	Adenocarcinoma
HOP-62	Lung	Adenocarcinoma
NCI-H226	Lung	Squamous cell carcinoma; mesothelioma; pleural effusion
Molt-4	Lymphocyte, T	Acute lymphoblastic leukemia
DU-145	Prostate	Carcinoma, derived from brain metastatic site
PC-3	Prostate	Adenocarcinoma, derived from bone metastatic site

Results

Overall levels of aDMA-containing proteins were increased in all cancer cell lines tested (Fig. SD1 and SD3). Although this profile does not always correlate with cell line-specific overexpression of PRMT1v2, other type I PRMTs could also contribute to this increase in aDMA arginine methylation. Indeed, CARM1 and PRMT6 were found to be expressed at higher levels in the panel of breast cancer cell lines used in this study (Fig. SD1), and we have also found these PRMTs to be overexpressed in other cancer types (Fig. SD3). Overall levels of sDMA-containing proteins were also increased in all cancer cell lines tested (Fig. SD1 and SD3), a profile that correlates well with cell line-specific overexpression of both type II PRMTs (PRMT5 and -7). Interestingly, the relative balance of alternatively spliced PRMT1 isoforms 1-3 mRNA is not altered in cancer cell lines when compared to immortalized normal breast cells (Fig. SD2). However, immunoblotting with our polyclonal antibody against PRMT1 revealed significant differences in v2 protein levels between normal breast and cancer cell line extracts. This suggests that translational and/or post-translational mechanisms (like e.g. protein stability) may be involved in regulating PRMT1v2 protein level in cancer cells. Further work will be required to address the functional significance of PRMT overexpression in various cancers, as well as the general role(s) that arginine methylation pathways might play in cellular transformation.

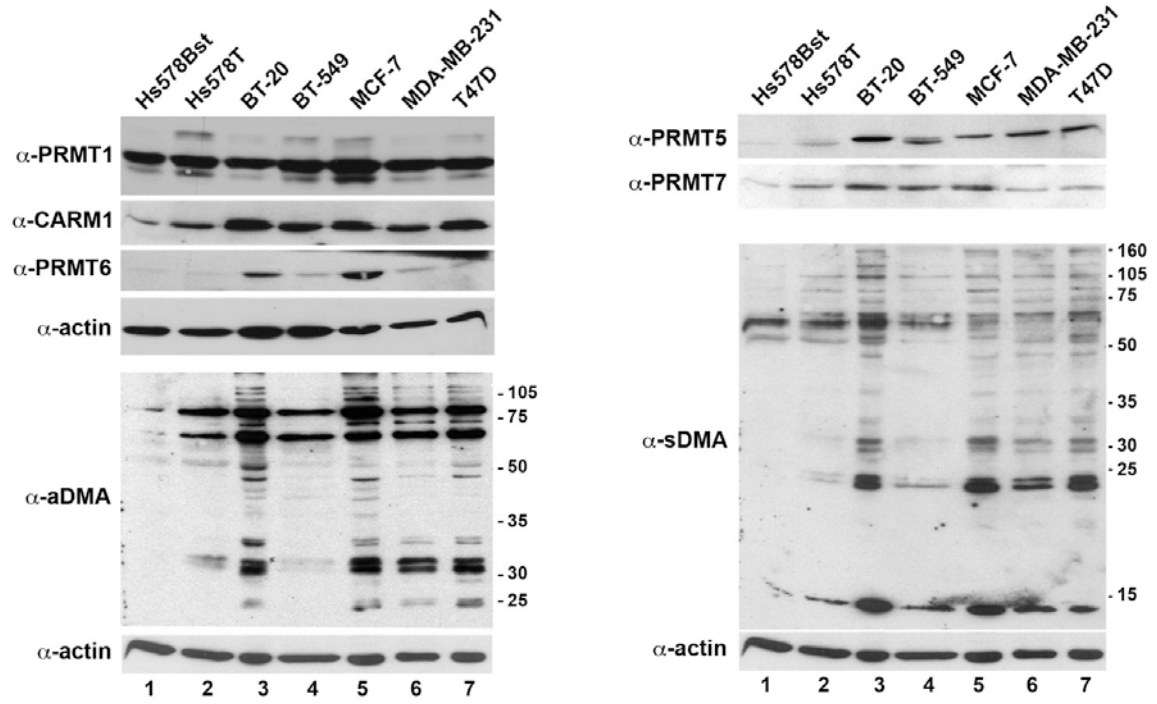


FIGURE SD1. PRMTs protein levels in breast cancer cell lines. Human normal breast and cancerous breast cells were lysed in RIPA buffer and protein concentration was measured spectrophotometrically using the *DC* Protein Assay reagent (Bio-Rad). 25 µg of proteins were resolved by SDS-PAGE, transferred to nitrocellulose, and immunoblotted with the indicated methyltransferase- and methylarginine-specific antibodies. Immunoblotting with anti-β-actin antibodies was used as a loading control.

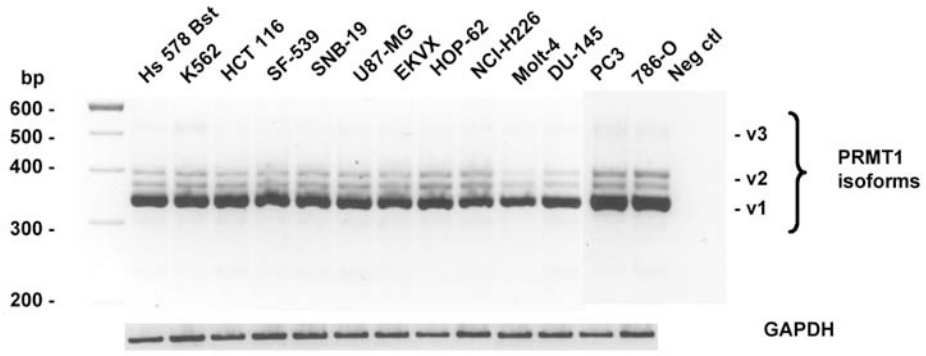


FIGURE SD2. Expression of PRMT1 splicing variants in normal vs cancer cell lines. Total RNA was extracted from breast cancer cells (see Table SDI) using Trizol reagent. RNA concentration was measured with a spectrophotometer, and RNA quality was determined by agarose gel electrophoresis. First strand cDNA synthesis was performed using 5 µg total RNA and the AMV Reverse Transcriptase (Promega) with an oligo-dT primer. 1/40 of total cDNAs were then amplified by PCR using the PF / PR primers (34) with Taq DNA polymerase (Qiagen). PCR products were resolved on a 2.5% agarose gel and visualized by ethidium bromide staining.

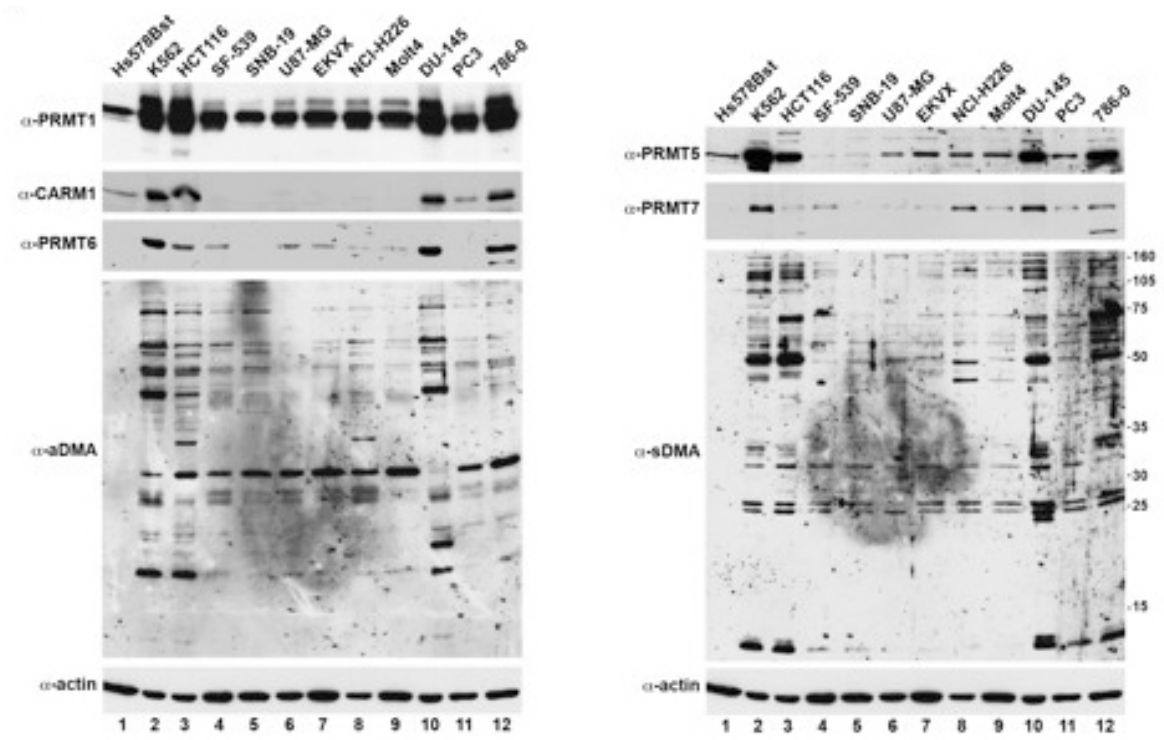


FIGURE SD3. PRMTs' expression profile and arginine methylation pattern in various cancer cell lines. Human normal and cancerous cells (see Table SDI) were lysed in RIPA buffer and protein concentration was measured spectrophotometrically using the *DC* Protein Assay reagent (Bio-Rad). 25 μg of proteins were resolved by SDS-PAGE, transferred to nitrocellulose, and immunoblotted with the indicated methyltransferase- and methylarginine-specific antibodies. Immunoblotting with anti-β-actin antibodies was used as a loading control.

CHAPTER 3

TDRD3: a Novel Tudor-Containing Protein in Search of a Function.

Isabelle Goulet and Jocelyn Côté.

Department of Cellular and Molecular Medicine and Center for Neuromuscular Disease, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada K1H 8M5

Description and statement of contributions of collaborators and co-authors

This unpublished manuscript uncovers TDRD3 as a novel component of Promyelocytic leukemia (PML) nuclear bodies possibly associated with sites of up-regulated mRNA transcription.

I. Goulet, performed all the experiments presented in this manuscript with the sole exception of the sequence alignment described in Fig. 3-5*B*, which was performed with the assistance of Dr. M. Khacho and Dr. K. Mekhail. Reagents were provided as outlined in the text. I. Goulet created all figures. I. Goulet performed writing of this manuscript with subsequent editing by Dr. J. Côté prior to thesis submission.

Abstract

Previous observations have suggested that constitutive methyl-dependent interactions with 'methyl-sensing' proteins (e.g. Tudor domain-containing proteins) are essential for pre-mRNA splicing to occur. The Tudor domains of the 'survival of motor neuron' (SMN) protein and of the splicing factor 30 kDa (SPF30) have been demonstrated to participate in pre-mRNA splicing through methyl-dependent interactions with a number of arginine-methylated RNA binding proteins. Yet, another Tudor domain-containing protein has been identified in a proteomic analysis of pre-spliceosomal complexes: the uncharacterized Tudor domain-containing protein 3 (TDRD3). In addition to its Tudor domain, TDRD3 harbors a putative nucleic acid recognition motif (OB-fold), a ubiquitin-associated (UBA) domain, and an exon junction complex (EJC)-binding motif (EBM). As suggested by the presence of putative nucleic acid-, EJC-binding, and methyl-binding properties, we speculated that TDRD3 could be central in pre-mRNA processing and/or its regulatory pathways through interactions with arginine-methylated proteins. We report here that TDRD3 localizes diffusely in both the cytoplasm and the nucleus, but not in nuclear domains typically associated with splicing factors. In ~20% of the cell population, however, TDRD3 is retained in nuclei in 20-40 large foci that are completely coincident with PML bodies. Expression of TDRD3 deletion mutants in HeLa cells revealed that both its OB-fold and its Tudor domain are sufficient for its tethering to nuclear foci. We also report that its localization to distinct nuclear foci is dependent on active RNA transcription, but have yet to establish a functional role for TDRD3 in pre-mRNA splicing. Taken together, our results suggest that TDRD3 likely participates in

some aspects of mRNA metabolism, by exploiting its putative nucleic acid-, EJC-binding, and/or methyl-binding properties.

Introduction

During eukaryotic transcription, the nascent pre-mRNA transcript assembles with a number of RNA-binding proteins, allowing a series of RNA processing events to occur in a transcription-dependent manner (210). These processing steps include addition of the cap structure, pre-mRNA splicing, and polyadenylation. Many factors involved in these steps are recruited to the carboxy-terminal domain of the largest subunit of the RNA polymerase II (211). A number of extensively studied RNA-binding proteins (e.g. SR proteins and hnRNPs) function in pre-mRNA splicing and the intracellular transport of mature and stable mRNAs (212, 213). Once transcription is complete, the mature message is coated with a unique complement of RNA-binding proteins, forming a messenger ribonucleoprotein particle (mRNP) that is then exported from the nucleus to the cytoplasm, where it is translated into proteins. Given the coupling between particular steps of transcription and specific RNA processing events, regulatory mechanisms must be in place to coordinate mRNP biogenesis in a dynamic manner[1].

Several studies have suggested roles for arginine methylation in the regulation of these co- and post-transcriptional processes, including pre-mRNA splicing (105, 106), mRNA export (215), and translation (112, 113). Direct evidence for a role of arginine methylation in pre-mRNA splicing first came from the observation that splicing efficiency is greatly reduced in hypomethylated nuclear extracts, indicating that normal

[1] First paragraph reproduced/amended with permission from 214. Yu, M.C. (2011) The role of protein arginine methylation in mRNP dynamics. *Molecular Biology International,* **2011,** 10. copyright under the Creative Commons Attribution License.

levels of methylarginine-containing proteins are required for pre-mRNA splicing *in vitro* (105). In addition, preincubation of nuclear extracts with an increasing amount of DMA-specific antibodies resulted in complete inhibition of splicing, demonstrating that DMA-containing proteins are part of the active spliceosome (105). Consistent with these findings, several hnRNPs, snRNP proteins, SR proteins, and other splicing factors were identified in a large proteomic survey of arginine-methylated protein complexes (114) and methylation of a number of these RNA binding proteins has since been confirmed (106, 115-122). However, the precise molecular mechanisms involved in the regulation of mRNA processing events by arginine methylation remain largely unknown.

One potential mechanism involves the formation of a 'docking motif' created by arginine methylation of RNA binding proteins and splicing factors to recruit methyl-binding effector proteins that could help orchestrate the proper localization and function of mRNA processing events. Supporting this hypothesis, a pilot series of titration experiments, in which HeLa nuclear extracts were incubated with increasing amounts of DMA-containing peptides, resulted in gradual inhibition of splicing (Côté, J., unpublished data). This observation provided us with functional evidence that constitutive methyl-dependent interactions with 'DMA-sensing' proteins are essential for splicing reactions to occur.

In eukaryotes, methylated arginines are specifically recognized by Tudor domains and several of the first Tudor domain-containing proteins that were identified also possessed RNA-binding capacity, which led to the suggestion that Tudor domains may

indeed play a role in RNA metabolism (134). Further evidence supporting this hypothesis has now been provided and common themes are starting to emerge. First, structural analysis of the N-terminal domain of FMRP, a well-characterized KH-containing RNA-binding protein, revealed the presence of a Tudor domain (216). Second, the germline type Tudor domain proteins (TDRD1, 2, 4, 5, 6, 7, 8, and 9) recruit to specific ribonucleoprotein (RNP) complexes methylated Piwi-type proteins involved in silencing transposons and retrotransposons by piwi-interacting RNAs during gametogenesis (135, 140-144). Third, the SND1 protein (a.k.a. p100, Tudor-SN) was shown to participate in both RNA interference and RNA editing pathways (145-147, 217). Fourth, the Tudor domains of the 'survival of motor neuron' (SMN) protein and of the splicing factor 30 kDa (SPF30) have been demonstrated to be required for the assembly of spliceosomal snRNPs (148-151) through known and predicted methyl-dependent interactions with a number of arginine methylated RNA binding proteins during pre-mRNA splicing (105, 106, 148, 154-158, 218). Collectively, these observations suggest that these Tudor domain-containing proteins could serve as integrators of arginine methylation signals in RNA processing reactions.

A proteomic analysis of the complete repertoire of the mammalian splicing machinery has identified another Tudor domain-containing protein in pre-spliceosomal complexes: the novel Tudor domain-containing protein 3 or TDRD3 (163). TDRD3 is a modular protein of unknown function. Its Tudor domain is highly similar to that of SMN and SPF30, and it has been demonstrated to have the ability to bind arginine-methylated polypeptides (148, 158). In addition to its Tudor domain, TDRD3 harbors a DUF/OB-

fold motif (putative nucleic acid recognition motif) (164) and a UBA (ubiquitin-binding) domain (165). More recently, an EJC (exon junction complex)-binding motif (EBM) has been found in the extreme C-terminus of TDRD3 that mediate interaction with two of the four core EJC proteins: Y14 and Magoh, and the peripheral interacting factor PABP1 (166). As suggested by the presence of nucleic acid-, methyl-, and EJC-binding properties, we speculated that TDRD3 could be central in pre-mRNA processing and/or its regulatory pathways through interactions with arginine-methylated proteins.

In order to gain insight into its cellular function, we generated antibodies against TDRD3 and undertook its characterization using a combination of biochemical and cellular biology approaches. Immunofluorescence microscopy studies showed that TDRD3 localizes diffusely in both the cytoplasm and the nucleus, but not in nuclear domains typically associated with splicing factors. We also report that TDRD3 localization to distinct nuclear domains is dependent on active RNA transcription and requires the expression of either its OB-fold or its Tudor domain.

Material and methods

Cell culture

The human HeLa cervical carcinoma cell line was purchased from the American Type Culture Collection (ATCC, Manassas, VA) and the human uterine leiomyosarcoma SKN cell line was a generous gift from Dr. Stéphane Richard (McGill University, Montréal, QC, Canada). Both cell lines were grown as a monolayer in DMEM medium supplemented with 1 mM sodium pyruvate, 50 IU/ml penicillin, 50 mg/ml streptomycin, and 10% fetal calf serum (Wisent, St-Bruno, QC, Canada). The human breast cancer cell line MCF7 was procured from ATCC and grown as a monolayer in complete DMEM medium supplemented with MEM non-essential amino acids, 0.01 mg/ml insulin (Invitrogen, Burlington, ON, Canada), and 10% foetal bovine serum (Wisent). Cells were transfected with DNA plasmids using Lipofectamine 2000 (Invitrogen) according to manufacturer's instructions.

DNA constructs

The pECFP-ER, -Golgi, and –Mito vectors were purchased from Clontech Laboratories Inc. (Mountain View, CA). Cloning of TDRD3's isolated domains (OB-fold, UBA, Central linker, and Tudor) into pCMV-Myc vector, as well as cloning of SMN, SPF30, and TDRD3 Tudor domains into pGEX-4T2 vector have been described elsewhere (219).

Antibodies

A peptide corresponding to the last 22 C-terminal amino acids of TDRD3, ^{723}DGQPRRSTRPTQQFYQPPRARN^{744}, was synthesized at W.M. Keck Foundation Biotechnology Resource Laboratory (Yale University, New Heaven, CT). Polyclonal antibodies were generated by Pocono Rabbit Farm and Laboratory Inc. (Canadensis, PA) using rabbits injected with the synthetic peptide coupled to keyhole limpet hemocyanin (KLH). The polyclonal antibodies were affinity-purified over the antigenic peptide coupled to Affi-Gel 10 beads (Bio-Rad, Hercules, CA) following manufacturer's instructions, eluted in 100 mM Glycine pH 2.5, buffered with 1 M Tris-HCl pH 8.0, dialyzed against 1X phosphate-buffered saline (1X PBS: 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4, pH 7.4) at 4°C overnight, and concentrated using Centricon centrifugal devices (Millipore, Bedford, MA). The integrity and concentration of the affinity-purified antibodies were determined by SDS-PAGE followed by Coomassie blue staining to visualize Ig bands. The affinity-purified anti-TDRD3 was used at a 1:150 dilution for immunofluorescence (IF) and at a 1:1000 dilution for western blot (IB) analysis.

Monoclonal antibodies against ASF/SF2 were purchased from Zymed Laboratories Inc. (San Francisco, CA). Rabbit anti-PSP1 polyclonal antibody was a generous gift from Dr. Archa H. Fox (University of Western Australia, Crawley WA Australia) (220). The 72B9 anti-fibrillarin mouse monoclonal antibody was a gift from Drs. Michael P. and Rebecca M. Terns (University of Georgia, Athens, GA) and was used at a 1:1000 dilution (IF). Monoclonal anti-p80-coilin antibodies were obtained from Dr. Gregory A. Matera

(Case Western Reserve University, Cleveland, OH). Anti-PML monoclonal antibody 5E10 was kindly provided by Dr. Roel Van Driel (State University of Utrecht, The Netherlands) and used at a 1:10 dilution (IF) (221). The 9E10 anti-myc monoclonal antibody (mAb) was procured from ATCC. All antibodies were used at a 1:200 dilution in IFs, unless otherwise specified. Mouse monoclonal anti-Elongation Factor 1 Alpha (EEF1A) antibody was purchased from Upstate Biotechnology (Upstate, NY) and used at a 1:3000 dilution (IB). Rabbit polyclonal anti-GST antibodies were described elsewhere (148). Anti-SMN antibody was purchased from BD Transduction Laboratories (Mississauga, ON, Canada). Monoclonal anti-snRNPs, antibody purified from Y12 hybridoma (mAb), was generously provided by Dr. Robin Reed (Havard University, Boston, MA).

Cell treatments and immunofluorescence

For indirect IF microscopy, cells were grown and transfected (when indicated) directly onto glass cover slips. RNA polymerase II transcription blockade was induced by treatment with 10 μg/ml actinomycin D (Sigma Aldrich, Saint-Louis, MO), while translation blockade was induced by treatment with 20 μg/ml cycloheximide (Sigma), both for 3 hours. Leptomycin B (Cedarlane, Burlington, ON, Canada) was used at a 10 μM concentration to inhibit CRM1-dependent nuclear export pathway in 6 h-treated cells. Cells were fixed with 4% paraformaldehyde for 10 min, permeabilized with 0.5% Triton X-100 for 5 min and blocked with 0.1% BSA in 1X PBS for 30 min prior to immunostaining. Cells were incubated with primary antibodies diluted in 0.1% BSA in 1X PBS for 1 h. Cells were then washed once with 0.1% Triton X-100 in 1X , twice with

0.1% BSA in 1X PBS and incubated with the appropriate fluorophore-annexed secondary antibodies in 0.1% BSA in 1X PBS, in the dark, for 1 h. All incubations were performed at room temperature. The cells were washed again as above, counterstained with DAPI and mounted onto glass slides. Fluorescence was visualized with a Z.1 AxioImager upright microscope (Carl Zeiss Canada Ltd, Toronto, ON, Canada) and images were captured with an AXIOCAM HRM R 2.0 CCD digital camera.

Protein purification, pulldowns, and mass spectrometry

GST-fusion proteins were overexpressed in *E. coli* BL-21 cells (Stratagene, La Jolla, CA) by induction with a final concentration of 0.1 mM isopropyl-D-thiogalactopyranoside (IPTG). Following 4-hour induction, cells were spun down, resuspended in 10 ml of 1X PBS containing Complete™ protease inhibitor cocktail (Roche Applied Science) and sonicated (5 pulses of 15 sec each, at a power output of 12 watts). Cell debris were discarded through centrifugation for 20 min at 10,000 X *g*. GST fusion proteins were then purified using glutathione-agarose (Sigma) and were kept on glutathione-agarose as a 50% slurry in 1X PBS for GST pull-down experiments. Pull-downs and mass spectrometry analysis have been extensively described elsewhere (219).

In vitro splicing reactions

Pre-mRNA substrates were synthesized *in vitro* using T3 RNA polymerase (Promega, Madison, WI) from *Sca*I-linearized plasmid kindly provided by Dr. Benoit Chabot (Université de Sherbrooke, QC, Canada). RNA was gel-purified as previously described (222) and incubated for 2 h under standard splicing conditions, or as specified, using a

HeLa nuclear extract as described elsewhere (223). The spliced RNAs were precipitated and reverse transcribed to cDNA using the AMV Reverse Transcriptase (Promega) with a primer specific to the adenovirus L2 exon (224). First strand cDNAs were then amplified by PCR with the Taq DNA polymerase (Qiagen, Mississauga, ON, Canada) using exon 7- and L2 exon-specific primers (224). PCR products were resolved on a 2% agarose gel and visualized by ethidium bromide staining.

Results

TDRD3 intracellular distribution

To begin investigating the cellular functions of TDRD3, we first generated a rabbit polyclonal antibody against a synthetic peptide coupled to keyhole limpet hemocyanine (KLH) corresponding to the last 22 C-terminal amino acids of TDRD3 (^{723}DGQPRRSTRPTQQFYQPPRARN^{744}). The affinity-purified antibody recognized a single band of ~83 kDa on immunoblots of SKN, MCF7, and HeLa total cell extracts (Fig. 3-1*A*). This band corresponds to the predicted molecular weight of TDRD3.

We then used the polyclonal antibody in indirect IF microscopy in HeLa cells to examine the subcellular distribution of endogenous TDRD3. In most cells, TDRD3 staining appeared predominantly cytoplasmic along with weak, diffuse nuclear staining. In the cytoplasm, the fluorescent signal was generally more intense in the perinuclear region, while gradually decreasing towards the edge of the cytoplasm (Fig. 3-1*B*, 'a' panels), a staining pattern reminiscent of proteins associated with polyribosomes and/or the endoplasmic reticulum. In the nucleus, we observed a strong nuclear accumulation of the fluorescent signal into 20-40 foci per nucleus (that did not resemble splicing factor-containing nuclear speckles) in ~ 20% of the cell population (Fig. 3-1*B*, 'b' and 'c' panels). This was frequently associated with a diffuse pattern of nucleolar staining (Fig. 3-1*B*, 'c' panel). Nucleoli are visible in the DAPI panels as dark regions in the nucleus. Preincubation of the affinity-purified antibody with an increasing amount of the immunogenic peptide resulted in a gradual decrease in the intensity of the IF signal, both

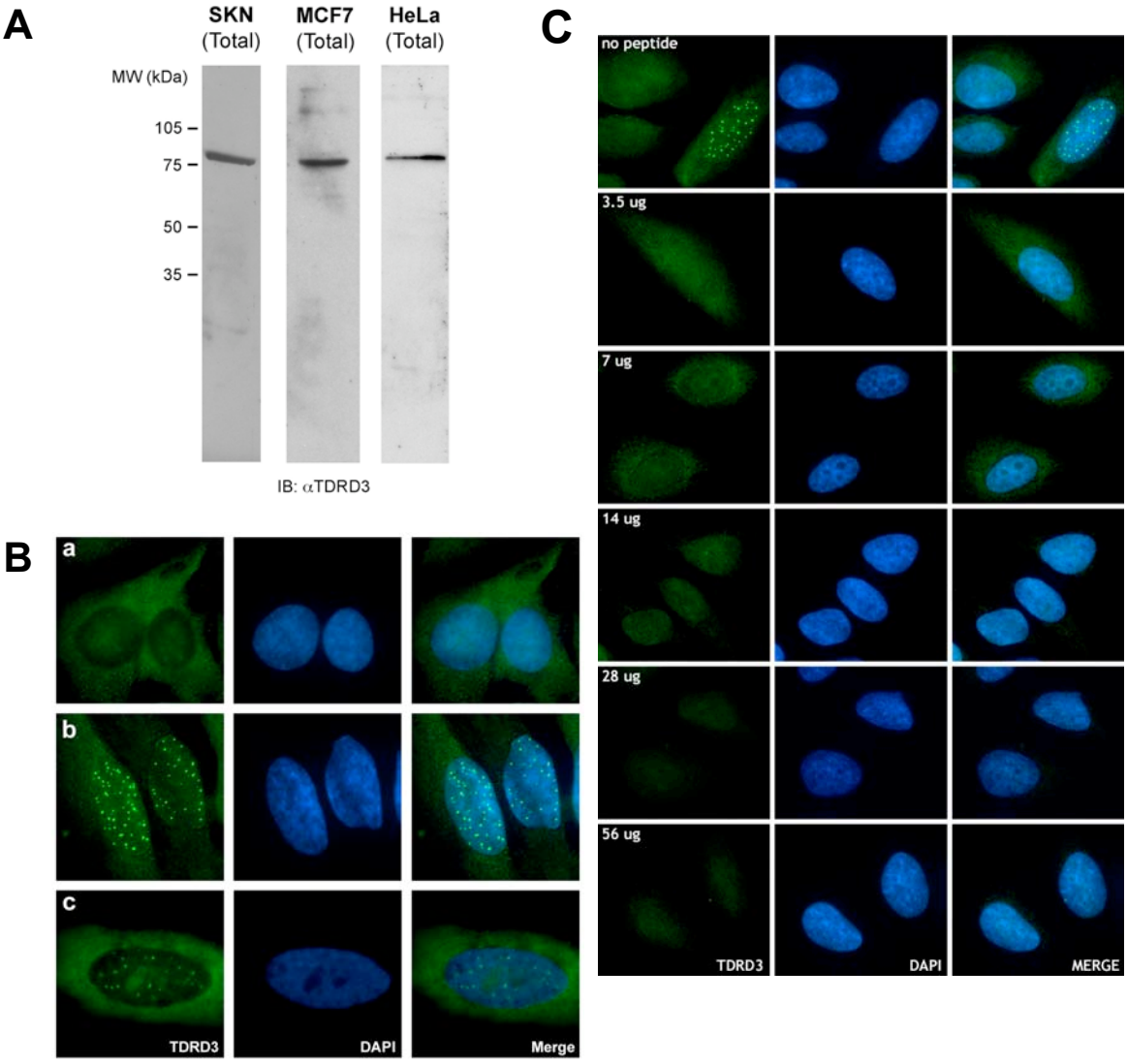


FIGURE 3-1. Characterization of a polyclonal antibody against the novel Tudor domain-containing protein 3 (TDRD3). *A*) The affinity-purified polyclonal anti-TDRD3 antibody recognizes a single band corresponding to the predicted molecular weight of TDRD3 in mammalian total cell extracts. *B*) Immunofluorescence microscopy of actively growing HeLa cells stained with the TDRD3-specific antibody (green). In most cells, TDRD3 is predominantly cytoplasmic (panel a). In contrast, TDRD3 accumulates into 20-40 foci per nucleus in approximately 20% of the cell population (panel b). This is frequently associated with a diffuse nucleolar staining (panel c). Nucleoli are visible in the DAPI panels as dark regions in the nucleus. *C*) Anti-TDRD3 antibodies (green) were pre-incubated with an increasing amount (3.5 to 56 µg) of the immunogenic peptide before immunofluorescence staining. Note the gradual decrease in TDRD3 staining (green), both in the cytoplasm and in the nucleus, as the amount of competing peptide increases. Merge panels show superimposition of green (TDRD3) and DAPI (nucleic acids) signals.

in the cytoplasm and in the nucleus, demonstrating the specificity of the observed staining patterns (Fig. 3-1*C*). Sub-cellular localization of endogenous TDRD3 was also

confirmed by sub-cellular fractionation (see Chapter 4, Fig. 2*D*). TDRD3 was found predominantly in the cytosolic fraction (63%), which correlated well with the sub-cellular distribution observed by IF microscopy and it was also detectable in the nuclear and nuclear insoluble fractions, albeit less so (37%) (219). Taken together, these observations suggest that TDRD3 generally predominates in the cytoplasmic compartment, but also has the capacity to accumulate in distinct nuclear structures.

Co-localization studies

As a step towards identifying the nuclear structures in which TDRD3 accumulates, we next performed double IF experiments in HeLa cells using TDRD3 antibodies in combination with antibodies directed against specific markers of various nuclear domains or 'bodies'. In order to verify our hypothesis that TDRD3 is a *bona fide* pre-mRNA splicing factor, we first assessed whether it co-localizes with other pre-mRNA splicing factors in nuclear 'speckles': nuclear domains located in the interchromatin regions of the nucleoplasm of mammalian cells. When observed by IF microscopy, they usually appear as 20-50 irregularly shaped nuclear structures that vary in size. TDRD3 nuclear aggregates did not co-localize with ASF/SF2, a splicing factor that typically resides in these structures (Fig. 3-2, panels a-d). Thus, TDRD3 does not accumulate in nuclear speckles, but in nuclear foci often found in close proximity to speckles (Fig. 3-2, panel d).

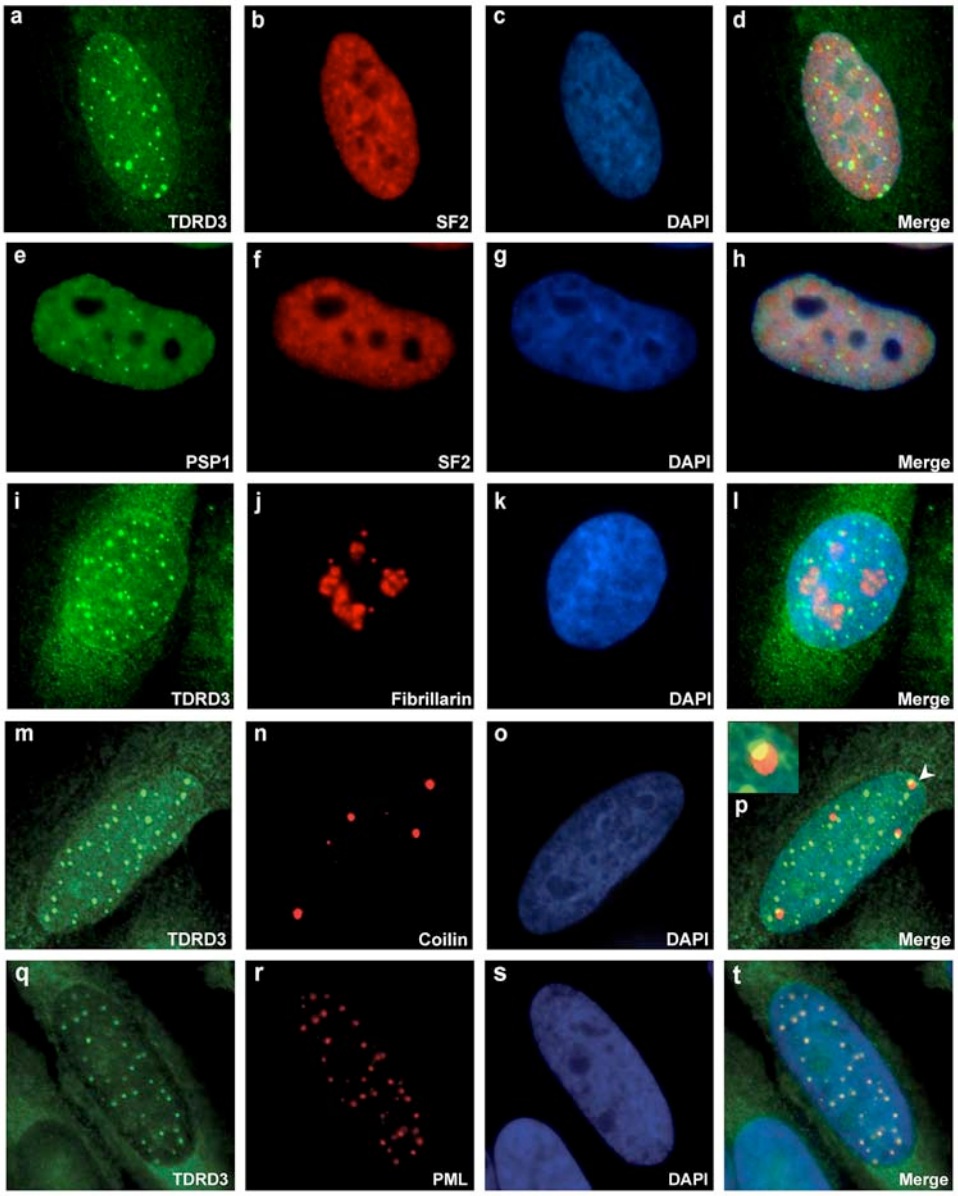


FIGURE 3-2. TDRD3 accumulates in PML nuclear bodies in a small subset of a cultured HeLa cell population. Co-analysis of TDRD3 (green) and SF2 (red) localization in HeLa cells by immunofluorescence microscopy (panels a-d), PSP1 (green) and SF2 (red; panels e-h), TDRD3 (green) and fibrillarin (red; panels i-l), TDRD3 (green) and coilin (red; panels m-p) with inset showing enlargement of close association between distinct coilin and TDRD3 foci (indicated by an arrowhead in panel p), and TDRD3 (green) and PML (red; panels q-t). Merge panels show superimposition of green, red and DAPI (nucleic acids) signals.

Paraspeckles are discrete bodies (10-30/nucleus) found in the interchromatin nucleoplasmic space that are always located adjacent to speckles in actively transcribing cells (220). They contain at least five RNA binding proteins (PSP1 (220), PSP2 (220),

p54/nrb (225), PSF (225), and CFl(m)68 (226)]) that also interact dynamically to form a crescent-shaped perinucleolar cap in the nucleolus in the absence of RNA polymerase II transcription (220). The function of paraspeckles is not fully understood, but recent findings have led to the speculation that they may represent storage sites to retain in the nucleus highly edited Alu-containing RNAs critical in the control of gene expression (227). Because both our TDRD3 and PSP1 antibodies were raised in the same species, we couldn't perform double IF co-localization studies. However, the expected punctate nuclear staining observed with the anti-PSP1 antibody (Fig. 3-2, panel e), shown in relation to ASF/SF2 staining in speckles (panel f), clearly did not resemble the punctate pattern observed for TDRD3 (panels a, i, m, and q). In addition, a diffuse nucleolar staining (Fig. 3-1*B*, 'c' panels and Fig. 3-2, panel q) or a nucleolar staining reminiscent of perinucleolar compartments (PNCs) were often observed by indirect immunofluorescence staining with our affinity-purified anti-TDRD3. Although its diffuse nucleolar staining co-localized with fibrillarin (a nucleolar marker), TDRD3's nuclear foci were typically not associated with nucleoli (Fig. 3-2, panels i-l). Thus, we conclude that TDRD3 does not localize to either paraspeckles or perinucleolar caps.

Cajal bodies (CBs) are dynamic structures within the nucleus involved in RNA-related metabolic processes such as the RNP biogenesis-maturation-trafficking-recycling of spliceosomal snRNPs, snoRNPs, telomerase, and U7 snRNP; as well as histone mRNA 3'-end processing (228-235). Human nuclei usually contain between 0-5 CBs, readily detected with antibodies against the marker protein p80-coilin (236-238). Double IF experiments in HeLa cells with the anti-TDRD3 and anti-coilin antibodies showed an

accumulation of TDRD3 in foci immediately adjacent to CBs (Fig. 3-2, panels m-p). Fibrillarin, a predominantly nucleolar protein, has been shown to localize to CBs as well (single dots in Fig. 3-2, panel j) and partial co-localization of TDRD3 with fibrillarin in these nuclear domains was also observed (panel l). CBs and PML (promyelocytic leukemia protein) nuclear bodies can associate (239, 240) and localize at sites of U snRNA expression (241). This prompted us to ask whether the nuclear TDRD3-containing foci could in fact be PML bodies. Moreover, the average mammalian cell contains 10-30 PML nuclear bodies, which makes them good candidates for the nuclear structures observed with the anti-TDRD3 antibody. Perfect co-localization was observed between PML protein staining in PML nuclear bodies and the foci detected using our affinity-purified anti-TDRD3 antibody (Fig. 3-2, panels q-t). Because PML nuclear bodies have been associated with a large number of nuclear functions, such as transcription, DNA repair, viral defense, stress, cell cycle regulation, proteolysis, senescence, and apoptosis (242-244), recruitment to these nuclear structures alone does not provide us with a precise hint about the putative functions of TDRD3.

Finally, in an attempt to identify in which cellular structures TDRD3 specifically localizes to in the cytoplasm, we transfected HeLa cells with commercially available pECFP vectors encoding the enhanced cyan fluorescent protein fused to an endoplasmic reticulum (pECFP-ER), a Golgi apparatus (pECFP-Golgi), or a mitochondrial targeting sequence (pECFP-Mito). Cells were fixed 24 h post-transfection and indirect IF staining was performed using our TDRD3 antibodies. In contrast to its usual diffuse distribution, endogenous TDRD3 was found concentrated in cytoplasmic granule-like foci in most

transfected cells. This observation led us to speculate that these cytoplasmic foci might be the result of transfection-induced stress, a hypothesis that was later confirmed (219). The biochemical characterization of TDRD3's recruitment to cytoplasmic stress granules (SGs) is described in detail in Chapter 4.

TDRD3 OB-fold and Tudor domains are recruited to nuclear foci

To identify which part of TDRD3 was important for its localization to nuclear foci as well as to cytoplasmic stress granules, a series of different myc-tagged constructs each containing individual TDRD3 domains were engineered (see Chapter 4, Fig. 5*A*) (219). In this early work, however, the myc-TDRD3 OB-fold (a.a. 98-174) construct did not include the putative DUF motif (a.a. 1-98), since it had been generated based on the 'truncated' TDRD3 amino acid sequence available at the time in databases. The fusion proteins were expressed in HeLa cells and immunofluorescence microscopy was performed using myc antibodies. The predicted molecular weight of these myc-tagged TDRD3 isolated domains ranges from 4 to 36 kDa, making it likely they will diffuse freely between the cytoplasm and the nucleus. As expected, expression of TDRD3's 'central linker' region resulted in diffuse cytoplasmic and nuclear staining (Fig. 3-3, 'myc-central linker' panels). In contrast, expression of TDRD3's OB-fold motif and its UBA or Tudor domains resulted in a predominantly nuclear staining pattern (Fig. 3-3, 'myc-OB-fold, -UBA, and –Tudor' panels). Remarkably, TDRD3's OB-fold motif and Tudor domain were both reproducibly and efficiently recruited to nuclear foci (Fig. 3-3, 'myc-OB-fold and –Tudor' panels), although the identity of these nuclear foci could not be determined due to incompatibilities between the antibodies used in this study, as they

were recognized by the same secondary antibody. Taken together, these results suggest that the UBA domain, and especially the OB-fold motif and the Tudor domain (which includes the exon junction binding motif) of TDRD3 may contribute to its recruitment to specific nuclear domains.

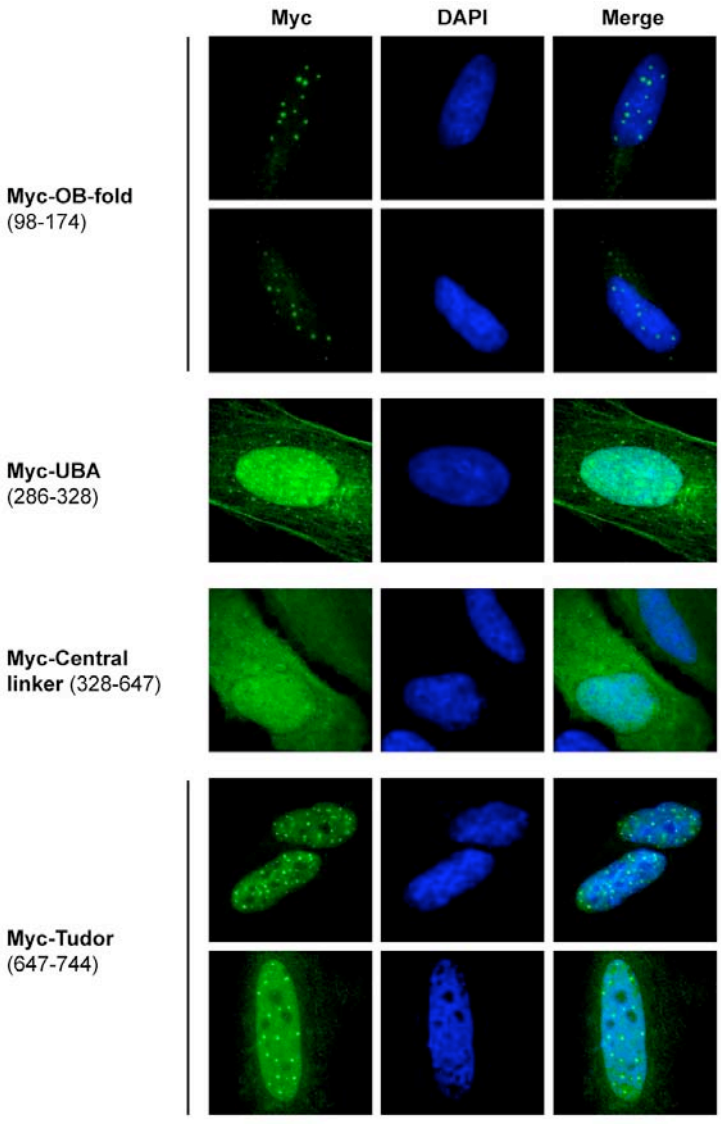


FIGURE 3-3. TDRD3 OB-fold motif and Tudor domain are recruited to nuclear domains. HeLa cells were transiently transfected with constructs expressing individual, myc-tagged TDRD3 domains (amino acid numbers are indicated in parentheses). Forty-eight hours post-transfection, cells were labeled with anti-myc antibodies and visualized by IF microscopy. Myc-tagged proteins are in green. Merge panels show superimposition of green and DAPI (nucleic acids) signals.

TDRD3 nuclear localization is transcription-dependent

Several RNA binding proteins involved in post-transcriptional regulation of RNA shuttle in a transcription-dependent manner between the nucleus and the cytoplasm (245, 246). The sub-cellular distribution of TDRD3, along with its putative link to RNA processing events, suggests that its nuclear localization to PML bodies may be regulated in such a fashion. In order to verify this hypothesis, HeLa cells were examined by immunofluorescence microscopy before and after treatment with actinomycin D (ActD), an inhibitor of RNA polymerase transcription. Untreated cells exhibited a normal TDRD3 distribution pattern (Fig. 3-4*A*, "untreated" panels). A three-hour treatment with ActD completely abolished TDRD3 recruitment to both PML nuclear bodies and nucleoli (Fig. 3-4*A*, "+ActD" panels). In contrast, inhibition of protein synthesis with cycloheximide had no effect on the distribution of TDRD3 (Fig. 3-4*A*, "+CHX" panels) and ongoing protein synthesis was not required for the redistribution effect observed under transcriptional inhibition by ActD (Fig. 3-4*A*, "+ActD/CHX" panels). Taken together, these results suggest that TDRD3 localization to PML nuclear bodies is transcription-dependent and is the result of inhibition of new RNA synthesis, not of the failure to synthesize a product from transcribed RNAs.

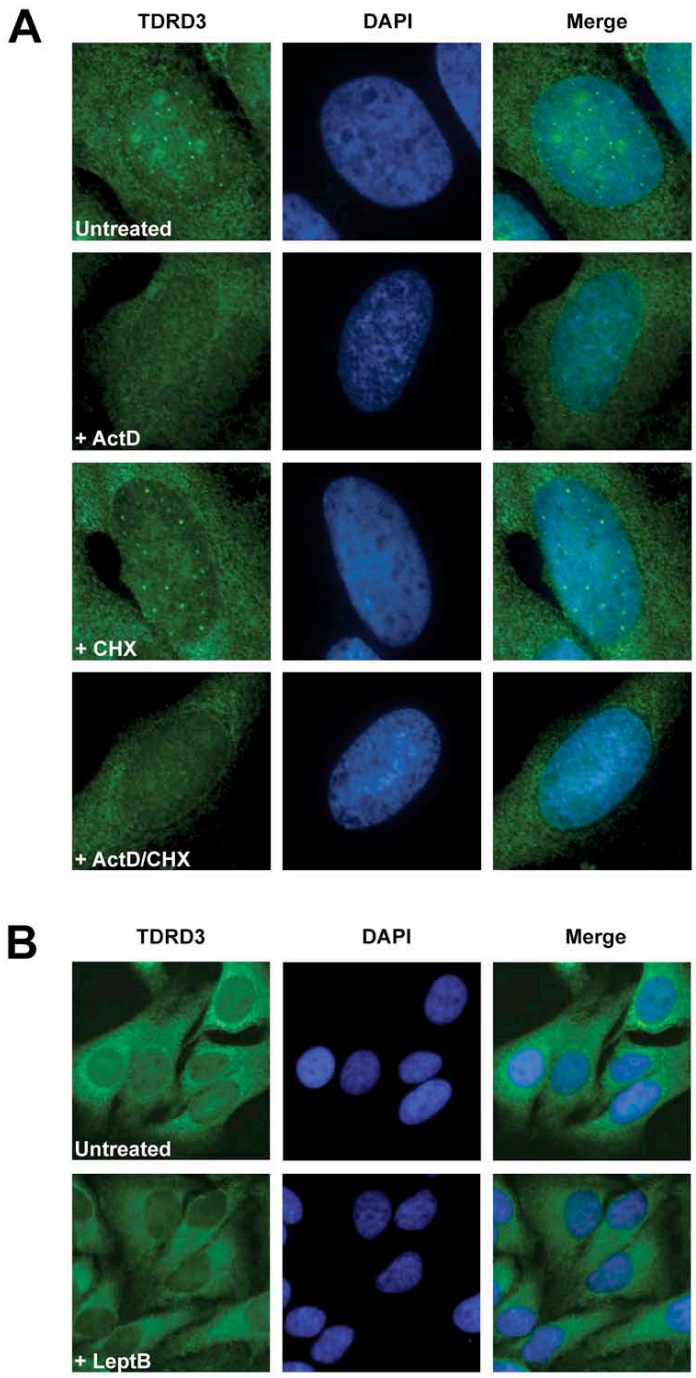


FIGURE 3-4. TDRD3 nuclear localization to PML bodies is transcription-dependent. *A*) HeLa cells were grown on coverslips in normal medium (untreated) or in medium containing 10 μg/ml actinomycin D (+ ActD), 20 μg/ml cycloheximide (+ CHX), or both agents (+ ActD/CHX) for 3 h. Cells were then fixed and immunostained with anti-TDRD3 antibodies (green). *B*) Alternatively, HeLa cells were grown in medium containing 10 μM Leptomycin B (+LeptB) for 6 h prior to anti-TDRD3 (green) immunofluorescence microscopy analysis. Merge panels show superimposition of TDRD3 (green) and DAPI (nucleic acids) signals.

Inhibition of RNA polymerase-mediated transcription did not result in noticeable redistribution of TDRD3 from the nucleus to the cytoplasm, as often observed with proteins that shuttle in a transcription-dependent manner between these two cellular compartments (245). One major nuclear export pathway relies on the nuclear export receptor CRM1. CRM1 mediates the nuclear export of cellular proteins bearing a leucine-rich nuclear export signal (NES) through the nuclear pore complex embedded in the nuclear membrane (188-190). Leptomycin B is a drug that inhibits CRM1-dependent nuclear export by forming a covalent bond with CRM1, hence preventing its interaction with proteins destined for nuclear export. If TDRD3 shuttles between the nucleus and the cytoplasm through CRM1-mediated nuclear export, we would expect a nuclear accumulation of TDRD3 after leptomycin B treatment, because it cannot be exported back to the cytoplasm. To verify this hypothesis, HeLa cells were treated with 10 μM leptomycin B for 6 h and analyzed by immunofluorescence microscopy with anti-TDRD3 antibody. This treatment did not affect TDRD3 localization (Fig. 3-4*B*, compare "untreated" with "+LeptB" panels), suggesting that its nuclear export (if any) is independent of the CRM1 nuclear export pathway. However, meaningful interpretation of this negative result would require appropriate positive controls since TDRD3 has recently been observed to accumulate in the nucleus of cells treated with leptomycin B (247). Thus, further studies will be needed to ascertain that TDRD3 has dynamic nucleocytoplasmic trafficking properties and whether they are dependent on the CRM1 nuclear export pathway or not.

TDRD3's Tudor domain interacts specifically with EEF1A

In order to begin identifying potentially arginine-methylated proteins interacting specifically with the Tudor domain of TDRD3, a glutathione-S-transferase (GST) pull-down assay was performed on HeLa cell extracts using a recombinant GST fusion protein consisting of the C-terminal region of TDRD3 (a.a. 647-744) encompassing the Tudor domain and the EBM (GST-TDRD3-Tdr) or a control GST protein coupled to glutathione agarose (see Chapter 4) (219). The retained proteins were resolved by SDS-PAGE and stained with Coomassie Blue. Five protein bands that were bound by GST-TDRD3-Tdr, and not by GST alone, were excised from the gel and subjected to mass spectrometry identification. The top matches obtained for each band are EWS, FUS, DDX3, SERBP1, and EEF1A (see Chapter 4, Table I for details) (219). Remarkably, all these proteins have known roles in RNA metabolism, which reinforces our hypothesis that TDRD3 likely functions in co- and/or post-transcriptional RNA processes.

We next assessed if all the proteins identified in our screen (see Chapter 4) were also able to bind the Tudor domains of SMN and SPF30. These experiments revealed that interaction with EEF1A was specific to TDRD3 (Fig. 3-5*A*). Interestingly, Khacho and colleagues recently identified EEF1A as a novel component of the nuclear protein export machinery, as the nuclear export of the tumor suppressor Von Hippel-Lindau (VHL) protein is transcription-dependent and is mediated by protein-protein interaction with EEF1A (248, 249). This group also mapped the region of VHL mediating the interaction with EEF1A to the motif shown in Fig. 3-5*B* (250). Since TDRD3 nuclear localization is also transcription-dependent, we compared the amino acid sequence of its Tudor domain

with the VHL motif and found a similar motif (Fig. 3-5*B*). In addition, the spinal muscular atrophy (SMA)-causing E134K mutation, located in the Tudor domain of SMN and conserved in the Tudor domain of TDRD3 (E691), is an amino acid required for this specific interaction (see Chapter 4, Fig. 6*D*) (219). Thus, TDRD3 nuclear export (if any) may be mediated by protein-protein interaction of its Tudor domain with EEF1A, which may explain why inhibition of the CRM1-dependent nuclear export pathway had no effect on TDRD3 sub-cellular localization (Fig. 3-4*B*). Future studies in the laboratory will aim to understand the importance of this interaction in TDRD3's function(s).

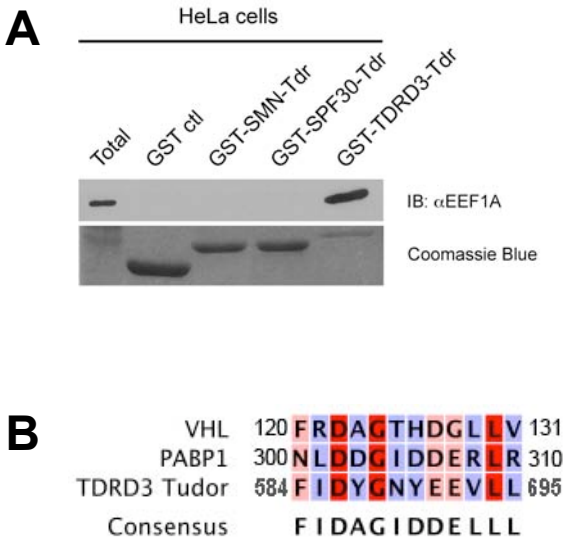


FIGURE 3-5. The Tudor domain of TDRD3 contains a binding motif to EEF1A and interacts specifically with it. *A*) HeLa cell extracts were prepared as described in 'Material and methods'. Cell lysates were then incubated with purified recombinant GST (as a control), GST-SMN-Tdr, GST-SPF30-Tdr, or GST-TDRD3-Tdr proteins coupled to glutathione-agarose for 16 h at 4°C. The retained proteins were resolved by SDS-PAGE. Bottom part of the gel was stained with Coomassie Blue to show GST-fusion proteins. Upper part of the gel was transferred to nitrocellulose, and immunoblotted with anti-EEF1A antibody. *B*) Sequence alignment of the EEF1A binding motif of VHL, PABP1, and TDRD3 (adapted from (250)). The color coding is red for conserved amino acids, pink for less conserved, and blue for amino acids that are not conserved.

Is TDRD3 involved in pre-mRNA splicing?

In order to assess whether TDRD3 is essential for splicing reactions to occur, as suggested by its putative association with pre-spliceosomal complexes, we next performed *in vitro* splicing reactions using a synthetic hybrid RNA containing the 5' splice site of hnRNPA1 exons 7 and 7B in competition for the 3' splice site of the adenovirus major late L2 exon as pre-mRNA substrate (224). In normal HeLa nuclear extracts, the proximal 5' splice site is preferred, which lead to the accumulation of a spliced fragment of approximately 300 bp after the splicing reaction (Fig. 3-6, lane 2). Pre-incubation of *in vitro* splicing reactions with antibodies directed against SMN or snRNPs (Y12) resulted in an inhibition of splicing, because the antibodies interfere with the normal functions of these proteins in pre-mRNA splicing (Fig. 3-6, lanes 6 and 7). Using the same approach, an inhibitory effect on splicing was observed when *in vitro* splicing reactions were performed with an increasing amount of polyclonal anti-TDRD3 antibody (Fig. 3-6, lanes 3-5). However, the same inhibitory effect was observed in our negative control using polyclonal anti-GST antibody (Fig. 3-6, lane 8). Thus, we cannot yet stipulate upon a putative role for TDRD3 in pre-mRNA splicing and further experiments will be needed to determine if TDRD3 is present in spliceosomal complexes during splicing.

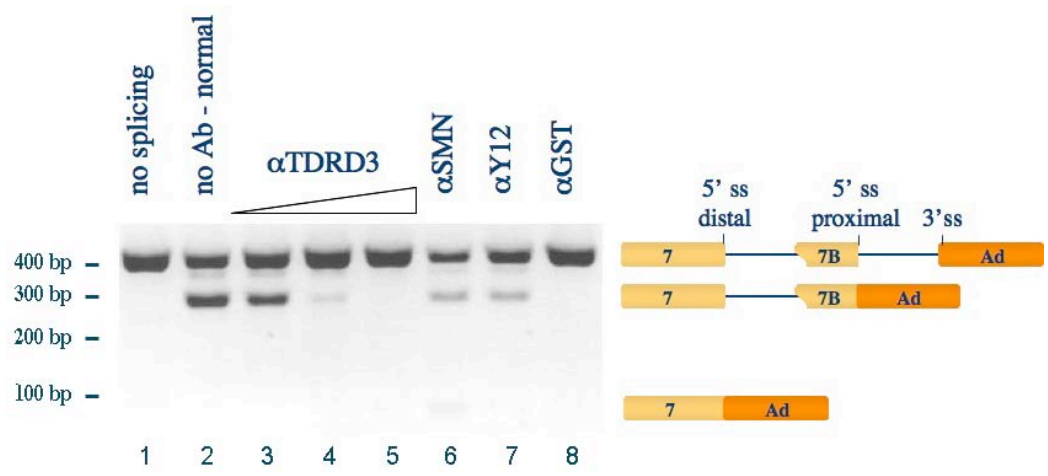


FIGURE 3-6. Effect of anti-TDRD3 antibodies on splicing reactions. Pre-mRNA substrates were incubated in a HeLa nuclear extract in the absence (lane 2) or in the presence of increasing amounts of anti-TDRD3 antibody (lanes 3-5), or in the presence of positive control antibodies against SMN (lane 6) and Y12 (lane 7), or negative control anti-GST antibody (lane 8), under splicing conditions for 2 h. The spliced RNAs were precipitated and reverse transcribed to cDNA using the AMV Reverse Transcriptase with a primer specific to the adenovirus L2 exon. First strand cDNAs were then amplified by PCR with the Taq DNA polymerase using exon 7- and L2 exon-specific primers. PCR products were resolved on a 2% agarose gel and visualized by ethidium bromide staining. The size of the different splicing products is depicted (right panels).

Discussion

We report here the first characterization of TDRD3, a novel Tudor domain-containing protein. Our findings identify TDRD3 as a novel component of both PML nuclear bodies and cytoplasmic stress granules. All three main domains (OB-fold, UBA, and Tudor/EBM) of TDRD3 potentially contribute to its nuclear localization, but only its OB-fold and its Tudor domain/EBM are efficiently recruited to distinct nuclear domains. We also show that TDRD3 recruitment to these nuclear domains is transcription-dependent and that EEF1A, a novel component of the nuclear protein export machinery, interacts specifically with its Tudor domain. Based on our preliminary results, however, we cannot yet stipulate upon a putative role for TDRD3 in pre-mRNA splicing.

TDRD3 is a novel protein recruited to PML nuclear bodies

We have found that the steady-state localization of the novel Tudor-containing protein TDRD3 is predominantly cytoplasmic, but also nuclear diffuse (Fig. 3-1). In the nucleus, it does not co-localize with splicing factors in nuclear speckles, but is rather transiently recruited to Promyelocytic leukemia (PML) protein-containing nuclear bodies under yet unknown cellular condition(s) (Fig. 3-2). This result does not exclude a putative role for TDRD3 in pre-mRNA splicing, since splicing factors can also be found within the nucleus both in a diffuse form and in other "non-speckle" nuclear domains (251). Co-localization of TDRD3 and PML in only a small subset (~20%) of the cell population suggests that TDRD3 recruitment to PML bodies may be regulated. This could be

achieved either in a cell cycle-dependent manner or very transiently, through another mechanism (see below).

Current models envision PML as a 'master recruiter' of a number of seemingly unrelated protein partners together with many protein-modifying enzymes, such as acetylases, kinases, as well as ubiquitin and SUMO ligases (244). Protein arginine methyltransferase 1 (PRMT1) is also present in PML bodies, where it methylates the DNA repair MRE11 protein prior to its translocation to sites of DNA damage (252). Through facilitation of partner protein post-translational modifications and subsequent partner activation, sequestration or degradation, PML bodies are proposed to anchor and regulate many nuclear functions, including DNA replication, transcription, and epigenetic silencing (244, 253).

In agreement with a putative function for PML nuclear bodies as integrators of post-translational protein modifications, the novel transient partner TDRD3 contains one putative and two confirmed PTM-specific binding domains: an OB-fold, a UBA and a Tudor domain, respectively. Although its function has not yet been characterized in the context of the TDRD3 protein, OB-fold domains are compact structural motifs usually involved in oligosaccharide and nucleic acid binding as well as in protein-protein interactions (164). The UBA domain of TDRD3 has been shown to specifically bind to Lys48-linked polyubiquitin chains, which typically mark proteins for degradation by the proteasome (254); whereas its Tudor domain specifically recognizes dimethylated arginines in protein substrates (148, 158), although some RNA-binding properties have

been suggested for this domain as well (134) (Leroux, G. and Côté, J. *et al.*, unpublished data). TDRD3 deletion mapping experiments clearly showed expression of all three domains in the nucleus where TDRD3 OB-fold and Tudor domains accumulated in distinct nuclear foci (Fig. 3-3). Although the identity of these nuclear bodies has not been determined, these observations suggest that TDRD3 OB-fold and Tudor domains are sufficient for its recruitment to distinct nuclear domains. Because of their binding properties, this could possibly be achieved through interaction with either nucleic acid-rich or PTM-rich domains in the nucleus.

TDRD3 recruitment to distinct nuclear domains is transcription-dependent

Intriguing questions remain about how TDRD3 nuclear accumulation in PML bodies is regulated. Our results strongly suggest that TDRD3 localizes to PML bodies in a transcription-dependent manner as inhibition of DNA-dependent RNA polymerase-mediated transcription by actinomycin D completely abolishes TDRD3 recruitment to distinct nuclear domains (Fig. 3-4). Yet, if we assume that approximately 75% of an unsynchronized cultured HeLa cell population is actively transcribing housekeeping genes as cells quiesce (255), why would TDRD3 protein immunostaining in PML bodies be strongly detected in only 10-20% of the cell population? This percentage may be actually more representative of cells in S phase, during which DNA replication occurs. Actinomycin D is a polypeptide antibiotic, which binds DNA at the transcription initiation complex and therefore, prevents elongation by RNA polymerases (256). However, it can also bind DNA duplexes, and it has been shown to interfere with DNA replication as well (257). Thus, TDRD3 accumulation in PML nuclear bodies could be

DNA replication-dependent rather than transcription-dependent or dependent on both processes (DNA replication-dependent transcription), a hypothesis that could be addressed in future experiments by using cell cycle-specific inhibitors.

Multiple lines of evidence favor the latter model, i.e. that TDRD3 is recruited to PML bodies during DNA replication-dependent transcription (S phase of the cell cycle). First, TDRD3 accumulates in PML bodies in cells with elongated nuclei (Fig. 3-1*B* and *C*, and Fig. 3-2), a nuclear shape that is representative of cells in S and G2 phases. Second, TDRD3 co-localizes with PML in cells containing higher number of PML bodies (Fig. 3-2). PML bodies fluctuate in number and size throughout the cell cycle, with the highest levels reached during both G1 and S phases (258, 259). Third, results from two independent studies suggest that a large fraction of PML bodies might be closely associated with active DNA replication-dependent transcription sites during S phase, since the majority (~70%) of PML bodies are not only associated with DNA replication domains (240), but also with active transcription sites during both G1 and S phases (260). Thus, TDRD3 may be transiently recruited to PML bodies associated with nuclear sites of up-regulated mRNA transcription tightly coupled to DNA synthesis. CBs are thought to participate in the preassembly of transcriptosomes involved in transcription and processing of RNA (261), and they have been implicated in replication-dependent histone gene transcription (262, 263). In the nuclei of unsynchronized mammalian cells, at least one PML body is usually spatially associated with a Cajal body (239, 240). In contrast, we observe a spatial association of each of the nuclear CBs with a TDRD3-containing PML body in cells potentially ongoing DNA synthesis (Fig. 3-2, panels i-p). The

functional significance of such close proximity between these two nuclear domains during DNA replication is unknown. Since PML bodies are composed primarily of proteins and do not generally contain RNA or DNA (264), one could speculate that they are recruited in the vicinity of CBs (or other transcription sites) to serve as a scaffold for (post-translationally modified) factors involved in the expression of histone (or other specific) genes.

In agreement with a putative role in replication and/or transcriptional processes, TDRD3 has been reported to directly interact with the DNA topoisomerase III beta (TOP3β) protein, both *in vitro* and *in vivo* (265). DNA topoisomerase IIIβ belongs to the type IA subfamily of topoisomerases, which catalyze the removal of negative DNA supercoils by cleaving one strand of the DNA to resolve DNA's double helix inherent topological problems during replication and transcription. TOP3β interaction with TDRD3 has been mapped to its N-terminal (a.a. 1-375) region, possibly encompassing both its DUF/OB-fold and its UBA domains (265). In addition, the Tudor domain of TDRD3 interacts specifically with EEF1A (Fig. 3-5), a protein proposed to be involved in transcription-dependent nuclear export of a number of proteins (248). Remarkably, expression of the TDRD3 Tudor domain/EBM alone is sufficient for both its accumulation in distinct nuclear domains and its relocalization to cytoplasmic stress granules upon stress (see Chapter 4, Fig. 5*C*). Specific interaction with EEF1A could therefore provide a possible molecular mechanism for TDRD3 translocation from transcription sites to cytoplasmic SGs under stressful conditions.

Is TDRD3 involved in pre-mRNA splicing?

We next sought to determine whether or not TDRD3 could be involved in pre-mRNA splicing. Such a role for TDRD3 has been proposed since its identification as a component of pre-spliceosomal complexes (163), but has never been confirmed. In the present study, we show that TDRD3 does not co-localize with splicing factors in nuclear speckles, but with PML bodies that do not contain either snRNPs or splicing factors (266), and consequently, are unfavorable sites to host mRNA splicing reactions. However, PML bodies are recruited to nuclear sites of active transcription and a number of studies have demonstrated that transcription and mRNA processing are intimately connected, one directly influencing the other (267). Thus, PML bodies may be indirectly involved in pre-mRNA splicing, by serving as a 'scaffold' not only for factors involved in transcription, but also for factors involved in transcription-coupled mRNA processing events. Another possibility is that TDRD3 has other functions outside PML bodies as well.

Pre-incubation of *in vitro* splicing reactions with an increasing amount of anti-TDRD3 antibody results in gradual inhibition of splicing (Fig. 3-6), in a manner similar to that observed with antibodies against the 'survival of motor neuron' protein SMN (150, 268) or snRNPs (Y12) involved in pre-mRNA splicing. This suggests that TDRD3 may be involved in mRNA processing, because the antibodies specifically interfere with its normal function in splicing reactions. However, pre-incubation of splicing reactions with high concentrations of antibodies directed against GST, a protein that has not been shown to be involved in pre-mRNA splicing whatsoever, results in similar inhibition of

splicing. Therefore, unless GST plays an unsuspected role in mRNA processing, the observed inhibitory effects may result from unspecific antibody interference in splicing reactions or may be the inevitable result of too high antibody concentration. In future experiments, it would be important to assess the inhibitory curve of increasing amount of GST antibodies on splicing reactions. This could provide us with a linear 'subinhibitory' range of GST antibody and corresponding TDRD3 antibody concentrations to use in our *in vitro* splicing reactions. Alternatively, classical immuno-depletion from nuclear extracts/add back experiments or RNA interference approaches could be used to help us elucidate TDRD3's implication in splicing. In the meantime, however, we cannot conclude, based on our preliminary results, that TDRD3 is involved in pre-mRNA splicing.

In a recent study, Kashima and colleagues identified an exon junction complex binding motif (EBM) in TDRD3 (166). The exon junction complexes (EJCs) assemble onto exon-exon junctions of mRNAs during splicing to form mature messenger ribonucleoproteins (mRNPs) (124, 269). The EJCs represent a 'molecular memory' of the splicing processes and remain stably bound to the mRNPs as they are exported out of the nucleus and into the cytoplasm (124). The EJCs and their positions are thought to help determining whether the transcripts are defective or not and therefore, are thought to play a major role in post-transcriptional regulation of mRNAs (270, 271). TDRD3's EBM motif mediates interactions between TDRD3 and the core EJC proteins Y14 and Magoh, as well as the Poly(A)-binding protein (PABP), which suggests a role for this motif in the recruitment of TDRD3 to EJCs and mRNAs (166) and a role for TDRD3 in post-

transcriptional regulation of mRNAs. Until very recently, the EJC had been linked exclusively to post-splicing events. But new findings suggest that the EJC is required for proper splicing of long intron-containing genes such as mapk (MAP kinase) transcripts by a mechanism that could possibly control exon definition (272). Thus, initial observations suggesting an association between TDRD3 and spliceosomal complexes may not have been fortuitous, although this association is most likely indirect.

In conclusion, recruitment to PML bodies possibly associated with up-regulated sites of mRNA transcription suggests a role for TDRD3 in this process. We cannot, however, yet stipulate upon a putative role for TDRD3 in pre-mRNA splicing. Thus, a putative role for TDRD3 in the synchronization of transcriptional with post-transcriptional processes remains uncertain. It will be important in future studies to determine whether TDRD3 can function as a transcriptional and/or splicing regulator and to identify (arginine-methylated) proteins with which TDRD3 interacts in such contexts. This could in turn provide significant insights into how arginine methylation and proper 'reading' of 'methyl marks' by TDRD3 regulates gene expression pathways.

Important note on TDRD3 antibody availability

Before we could complete this study, we ran out of our anti-TDRD3 serum. Therefore, we generated a new rabbit polyclonal antibody against TDRD3 using the same immunogenic peptide (encompassing the exon junction complex binding motif) coupled to KLH. As expected, TDRD3 immunofluorescence staining, using the new affinity-purified antibody, was detected as being predominantly cytoplasmic and weakly nuclear diffuse, as well as in cytoplasmic SGs upon stress. However, absolutely no TDRD3 accumulation in nuclear foci could be detected. Such discrepancy may result from cross-reactivity of our anti-peptide antibodies with cognate proteins. Another possibility may be epitope variability, as both antibodies may not recognize the exact same amino acid sequence on the immunogenic peptide (and consequently, on the protein). Thus, the epitope recognized by the new antibody may not be accessible for attachment when TDRD3 is associated with nuclear complexes such as the exon junction complex or PML bodies. Therefore, until another 'nuclear foci-detecting' anti-TDRD3 antibody becomes available, I decided to focus my efforts on the characterization of TDRD3's recruitment to cytoplasmic stress granules, which is described in details in the next chapter (see Chapter 4).

CHAPTER 4

TDRD3, a Novel Tudor Domain-Containing Protein, Localizes to Cytoplasmic Stress Granules

Isabelle Goulet[1], Sophie Boisvenue[1], Sophie Mokas[2], Rachid Mazroui[2], and Jocelyn Côté[1,].*

[1]Department of Cellular and Molecular Medicine and Center for Neuromuscular Disease, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada K1H 8M5
[2]Unité de recherche en génétique humaine et moléculaire, Centre de recherche de l'Hôpital Saint-François d'Assise, Université Laval, Québec, Québec, Canada G1L 3L5

*To whom correspondence should be addressed. Tel: +1 6135625800 ext. 8660; Fax: +1 6135625636; Email: jcote@uottawa.ca

Received May 21, 2008; revised and accepted July 11, 2008.

This research was originally published in *Human Molecular Genetics.* 2008;

17: 3055-74. © Oxford University Press.

Description and statement of contributions of collaborators and co-authors

This manuscript uncovers TDRD3 as a novel component of cytoplasmic stress granules, which recruitment to is dependent on the methyl-binding surface in its Tudor domain. Five novel TDRD3 Tudor interacting partners are also identified, most of which are potentially methylated RNA binding proteins as well as known or novel constituents of cytoplasmic SGs. Finally, this manuscript presents evidence supporting a role for arginine methylation in the regulation of SG dynamics.

I. Goulet performed all the experiments presented in this manuscript, with occasional assistance from S. Boisvenue, with the sole exception of the experiments described in Fig. 4, which were performed by S. Mokas. Reagents were provided as outlined in the text. I. Goulet created all figures. Dr. J. Côté and I. Goulet contributed equally to the manuscript writing, editing, and revisions prior to publication.

Abstract

Our previous work has demonstrated that the Tudor domain of the "survival of motor neuron" protein and the Tudor domain-containing protein 3 (TDRD3) are highly similar and that they both have the ability to interact with arginine-methylated polypeptides. TDRD3 has been identified amongst genes whose overexpression has a strong predictive value for poor prognosis of estrogen receptor-negative breast cancers, although its precise function remains unknown. TDRD3 is a modular protein, and in addition to its Tudor domain, it harbors a putative nucleic acid recognition motif and a ubiquitin-associated domain. We report here that TDRD3 localizes predominantly to the cytoplasm, where it co-sediments with the Fragile X Mental Retardation Protein on actively translating polyribosomes. We also demonstrate that TDRD3 accumulates into Stress Granules (SGs) in response to various cellular stresses. Strikingly, the Tudor domain of TDRD3 was found to be both required and sufficient for its recruitment to SGs, and the methyl-binding surface in the Tudor domain is important for this process. Pull down experiments identified five novel TDRD3 interacting partners, most of which are potentially methylated RNA binding proteins. Our findings revealed that two of these proteins, SERPINE1 mRNA binding protein 1 and DEAD/H box-3 (a gene often deleted in Sertoli-cell-only syndrome), are also novel constituents of cytoplasmic SGs. Taken together, we report the first characterization of TDRD3 and its functional interaction with at least two proteins implicated in human genetic diseases and present evidence supporting a role for arginine methylation in the regulation of SG dynamics.

Introduction

Cell signaling pathways rely heavily on modular proteins containing protein-protein interaction domains to sense, transmit, and process signals that regulate cellular functions (273). The domains involved in these interactions recognize specific peptide motifs in their binding partner, and this recognition can be further influenced by the presence of posttranslational modifications, such as phosphorylation, ubiquitination, or methylation. These modifications can be either added or removed to create dynamic changes in protein binding properties, thus enabling rapid cellular responses to both external and internal stimuli.

Arginine methylation is a common post-translational modification that has been shown to play central roles in signal transduction pathways regulating many cellular processes, including cell growth, transcription, and DNA repair (7, 96, 169). Although a number of proteomic screens have now significantly expanded the repertoire of arginine-methylated proteins (12, 114, 118), RNA binding proteins remain a predominant class of the proteins harboring this modification. Consistent with this observation, several studies have also suggested roles for arginine methylation in the regulation of various post-transcriptional processes, including pre-mRNA splicing (105, 106), mRNA export (215), and translation (112, 113). However, the precise molecular mechanisms involved in the regulation of these processes remain largely unknown. Several lines of evidence support the idea that arginine methylation may serve as an important regulatory signal in these cellular pathways. First, methylated arginines have been shown to be mainly involved in

the regulation of protein-protein interactions (106, 128, 129, 131). Second, these modified arginines are specifically recognized by a protein module termed "Tudor domain" (148, 158). Third, the recent discovery of enzymes that can remove the methylation mark (274-276) confirms that this modification is dynamic.

The Tudor domain is a 60-amino acid motif that was first discovered through a protein sequence comparison analysis, performed to identify conserved patterns in *Drosophila* TUDOR (133, 134), a protein involved in germ cell formation that harbors 11 repeated Tudor domains. Tudor domains are highly conserved throughout evolution: they can be found in virtually all organisms, from bacteria to mammals. In the latter, ~15 "Tudor Domain containing proteins" (TDRD1, 2, etc.) can be found in sequence databases; the functions of most remaining largely unknown. The Tudor domain of the "survival of motor neuron" (SMN) protein, the causative gene for spinal muscular atrophy, is by far the best characterized. Three-dimensional structure determination of the SMN Tudor fold revealed a barrel-like structure composed of β-sheets forming a hydrophobic pocket, surrounded by negatively charged and aromatic residues, which together likely constitute the protein-protein interaction surface (136, 137). This structural fold shares similarities with the Chromo domain, a motif known to interact with methylated lysines in histones (138). This resemblance led to the proposition that methyl-substrate binding might be a general feature of these related domains. Several studies have now indeed demonstrated methyl-dependent interactions between the Tudor domain of SMN and a number of arginine methylated proteins (105, 148, 154-157). Other studies have also suggested the existence of another subclass of Tudor domains, dubbed "tandem

Tudor", which recognize methylated lysines instead of arginines (158, 159). These include the Tudor domains of 53BP1 (160, 161) and JMJD2A (162). Taken together, these findings support the notion that the Tudor domain is a *bona fide* "protein module".

Interestingly, several of the first Tudor domain-containing proteins that were identified also possessed RNA-binding capacity, which led to the suggestion that Tudor domains may play a role in RNA metabolism (134). Further evidence supporting this hypothesis has now been provided, as common themes are starting to emerge. First, structural analysis of the N-terminal domain of FMRP, a well-characterized KH-containing RNA-binding protein, revealed the presence of a Tudor domain (216). This Tudor-KH configuration is also found in at least one other protein (277). This suggests that these modular proteins could serve as scaffolding proteins involved in the integration of arginine methylation signals and RNA processing. Second, the Tudor-SN protein was shown to participate in both RNA interference and RNA editing pathways (217). Third, SMN promotes the transfer of core Sm proteins to snRNAs in the cytoplasmic assembly of spliceosomal snRNPs (148-151), and it is thought to serve as a "master assembler" of many additional ribonucleoprotein particle (RNP) complexes (278-280). Finally, SMN (106, 150, 268) and two other Tudor domain-containing proteins, SPF30 (153, 281) and p100 (145, 282), have all been shown to play a role in pre-mRNA splicing. These observations suggest that Tudor domains may serve as molecular adaptors in post-transcriptional processes regulated by arginine methylation.

The Tudor domain-containing protein 3 (TDRD3) is a novel modular protein of unknown function. It was identified amongst genes whose overexpression has a strong predictive value for poor prognosis of estrogen receptor-negative breast cancers (167). This suggests that TDRD3 may contribute to the progression of breast cancer. Its Tudor domain is highly similar to that of SMN and we have recently demonstrated that it also has the ability to bind arginine-methylated polypeptides (148). In addition to its Tudor domain, TDRD3 harbors a DUF/OB-fold motif (putative nucleic acid recognition motif), as well as a UBA (ubiquitin-binding) domain (165). As suggested by the presence of both nucleic acid- and methyl-binding properties, we speculated that TDRD3 could be central in RNA processing regulatory pathways involving arginine methylation.

In order to gain insight into its cellular function, we generated antibodies against TDRD3 and undertook its biochemical characterization. Immunofluorescence microscopy studies showed that TDRD3 localizes diffusely in both the cytoplasm and the nucleus. Interestingly, we observed that TDRD3 colocalizes with TIA-1, G3BP, and FMRP in cytoplasmic Stress Granules (SGs) in response to various cellular stresses. SGs become visible at the microscopic level upon various environmental stresses, but recent studies suggest they may represent natural sites of cytoplasmic mRNA sorting/triage that play a major role in the posttranscriptional regulation of gene expression by regulating the localization, stability, and translation of many mRNAs (283-285). Known constituents of SGs can be catalogued within the following categories: (i) translation factors, (ii) mRNAs, (iii) mRNA-binding proteins with roles in the regulation of translation and/or mRNA stability, (iv) proteins linked with mRNA metabolism and (v)

signaling proteins with no known links with RNA metabolism (286). Expression of TDRD3 deletion mutants in HeLa cells revealed that its Tudor domain is both required and sufficient for its localization to stress granules. Pull down experiments using its isolated Tudor domain led to the identification of five interacting proteins with known functions in various aspects of RNA handling. Strikingly, we show that two of them, DEAD/H box-3 (an RNA helicase also known as DDX3X/Y, DBX/Y, HLP2, and DDX14) and SERPINE1 mRNA binding protein 1 (SERBP1), also localize to SGs following sodium arsenite treatment. Taken together, our results place TDRD3 in complexes with at least two proteins implicated in human genetic diseases (FMRP and DDX3) and suggest a role for arginine methylation in the regulation of SG dynamics.

Results

TDRD3 is a novel modular protein

Data obtained from the UCSC Genome Browser (human Mar. 2006 hg18 assembly; NCBI Build 36.1) revealed that the human *TDRD3* gene consists of 14 exons spanning a region of approximately 176,5 Kb of genomic DNA on chromosome 13q21.2 (Fig. 1*A*). All splicing junctions conform to the GU-AG general consensus and the resulting mRNA is approximately 2.9 Kb in length (Fig. 1*B*). The *TDRD3* mRNA sequence found in databases (GenBank accession no. NM_030794) contains 10 coding exons (exons 4-13 in Fig. 1*A*) with the ATG initiation codon (here designated as M94) located within exon 4. The *TDRD3* transcript derived from our analysis of the human genomic sequence, with reference to reported cDNAs and ESTs, unveiled an additional upstream in-frame initiation codon within the first exon. The sequence preceding this new ATG (M1) is a better match to the Kozak consensus sequence (287) than the previously proposed initiation codon (Fig. 1*B*). This extended *TDRD3* coding sequence therefore contains three additional coding exons (1-3) upstream of the previously described exons 4-13. The STOP codon (TAA) falls within exon 13 and the polyadenylation signal (AATAAA) is located in the non-coding exon 14 (Fig. 1*A* and *B*). The presence of these upstream exons in the mRNA coding sequence was confirmed by RT-PCR and subsequent sequencing using total RNA from HeLa cells.

The predicted TDRD3 protein consists of 744 amino acids and has a calculated molecular weight of 82.7 kDa. Amino acid sequence similarity analysis provided by the

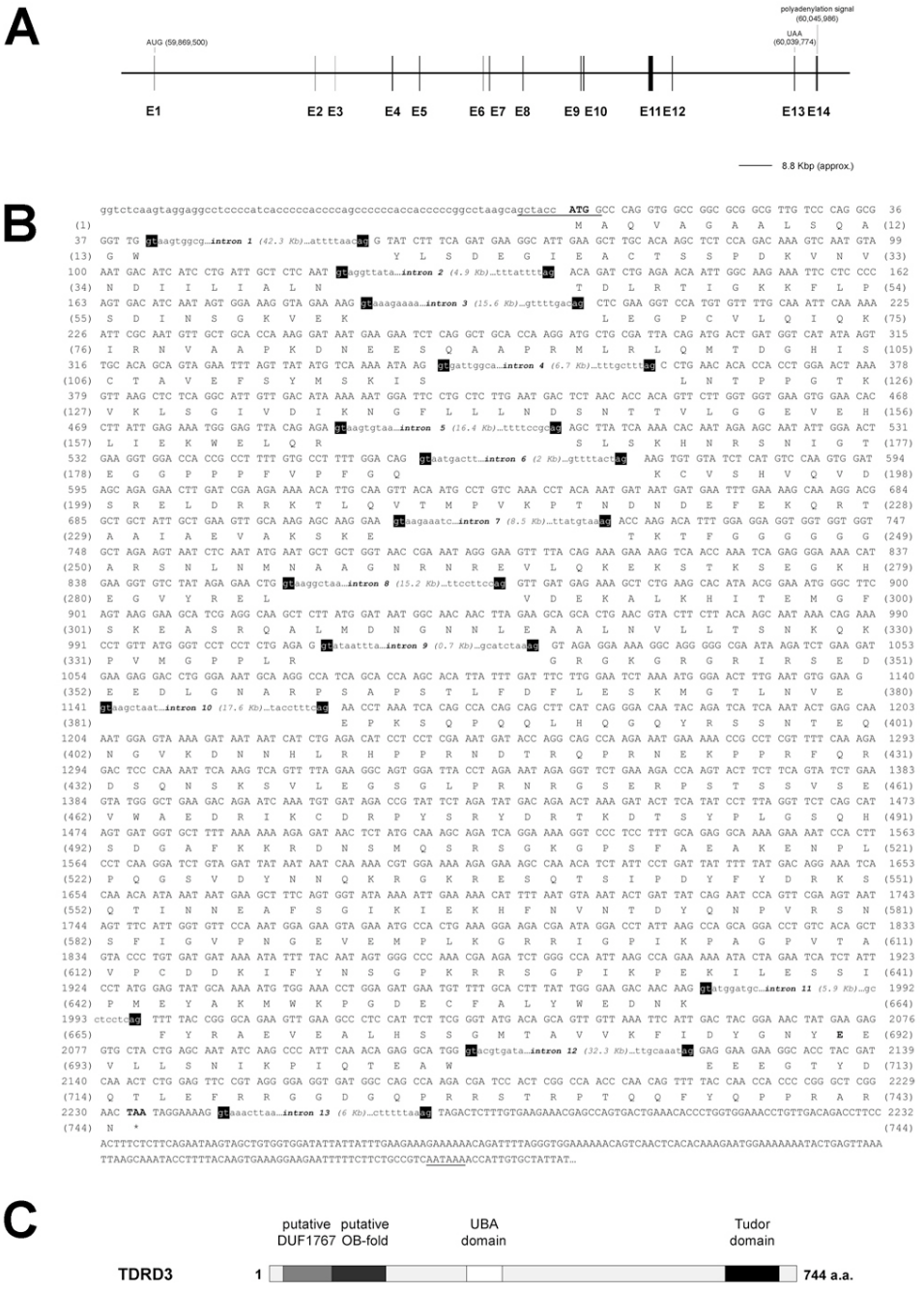


FIGURE 1. **Genomic organization of the human TDRD3 locus and schematic representation of the TDRD3 protein.** Analysis of the human TDRD3 gene locus revealed 14 exons spanning over ~176.5 Kb of genomic DNA, encoding a 2.9-Kb mRNA transcript (**A**). The sequence of exons is displayed with capital letters above the amino acid sequences they encode. Only short sequences around the intron boundaries are shown with small case letters. All of the exon-intron junctions (shown in black boxes) conform to the GU-AG general consensus. The start and stop codons are in bold, whereas the Kozak sequence and the polyadenylation signal are underlined. Both nucleotides and amino acids are numbered starting from the ATG codon (**B**). The encoded TDRD3 protein consists of 744 amino acids arranged in a modular fashion. Data obtained from the NCBI Conserved Domains and the Protein Family (Pfam version 22.0) databases predict the presence of three structural domains in TDRD3: a DUF1767/OB-fold (a.a. 13-169), a UBA (a.a. 286-328), and a Tudor (a.a. 651-710) domain (see text for details) (**C**).

NCBI Conserved Domains and the Protein Family (Pfam version 22.0) databases suggests that TDRD3 is a modular protein (Fig. 1*C*). In addition to its Tudor domain (a.a. 647-710), which selectively recognizes arginine-methylated motifs in proteins (148), TDRD3 also contains an ubiquitin-associated (UBA) domain (a.a. 286-328), a motif found in several proteins having connections to ubiquitin and the ubiquitination pathway (165). Moreover, while searching for proteins sharing homologies with the Bloom's Syndrome-associated polypeptide 75 (BLAP75) protein, Yin and colleagues identified a putative OB-fold in the N-terminal region of TDRD3 (a.a. 94-169) (288). OB-folds are compact structural motifs usually formed of five β-sheets that are often used for nucleic acid recognition, although they have been observed at protein-protein interfaces as well (164). This search, however, was limited to the "truncated" TDRD3 sequence available in databases. A new search within the Pfam database using full length TDRD3 amino acid sequence revealed that its OB-fold is associated with another putative motif, the DUF1767 domain (a.a. 13-95). This eukaryotic domain has no known function, but is often found at the N-terminus of nucleic acid binding motif (Pfam number PF08585). The modular domain structure of TDRD3 suggests that it may serve as an adaptor molecule for various signaling pathways.

TDRD3 intracellular distribution

In order to characterize the TDRD3 protein, a rabbit polyclonal antibody was raised against a synthetic peptide (coupled to KLH) corresponding to the last 22 C-terminal amino acids of TDRD3 (^{723}DGQPRRSTRPTQQFYQPPRARN^{744}). The affinity-purified antibody recognized a single band of approximately 83 kDa on immunoblots performed

using HeLa cell extracts (Fig. 2*A*, lane 1), which corresponds to the predicted molecular weight of TDRD3. Immunoreactivity to endogenous TDRD3 in Western blots was completely abolished by preincubation of TDRD3 antibodies with the antigenic peptide, but not with an unrelated peptide, demonstrating the specificity of our antibody (Fig. 2*A*, compare lanes 2 with 3). To further confirm that our antibodies specifically recognize the TDRD3 protein, *in vitro* coupled transcription/translation (IVTT) reactions were programmed with a construct designed to produce a hexahistidine-tagged truncated version of TDRD3 (ΔN328 in Fig. 5*A*), which still contains the reactive epitope. This TDRD3 C-terminal domain was effectively recognized by our TDRD3 antiserum in Western blot analysis of the programmed IVTT reactions but not in control reactions (Fig. 2*B*, compare lanes 2 with 1), confirming that our antibody recognizes the TDRD3 protein sequence. Interestingly, the antibody also specifically recognized a single band corresponding to TDRD3 predicted molecular weight in the Rabbit Reticulocyte Lysates (RRL), suggesting that TDRD3 is found endogenously in these lysates (Fig. 2*B*, lanes 1 and 2).

We next performed indirect immunofluorescence microscopy with our polyclonal antibody in HeLa cells to examine the subcellular distribution of endogenous TDRD3. In most cells, TDRD3 was detected in a predominant cytoplasmic staining combined with a weaker nuclear diffused staining. In the cytoplasm, the fluorescent signal was generally more intense in the perinuclear region while gradually decreasing towards the edge of the cytoplasm (Fig. 2*C*a), a staining pattern reminiscent of proteins associated with polyribosomes and/or the endoplasmic reticulum. Again, pre-incubation of the affinity-

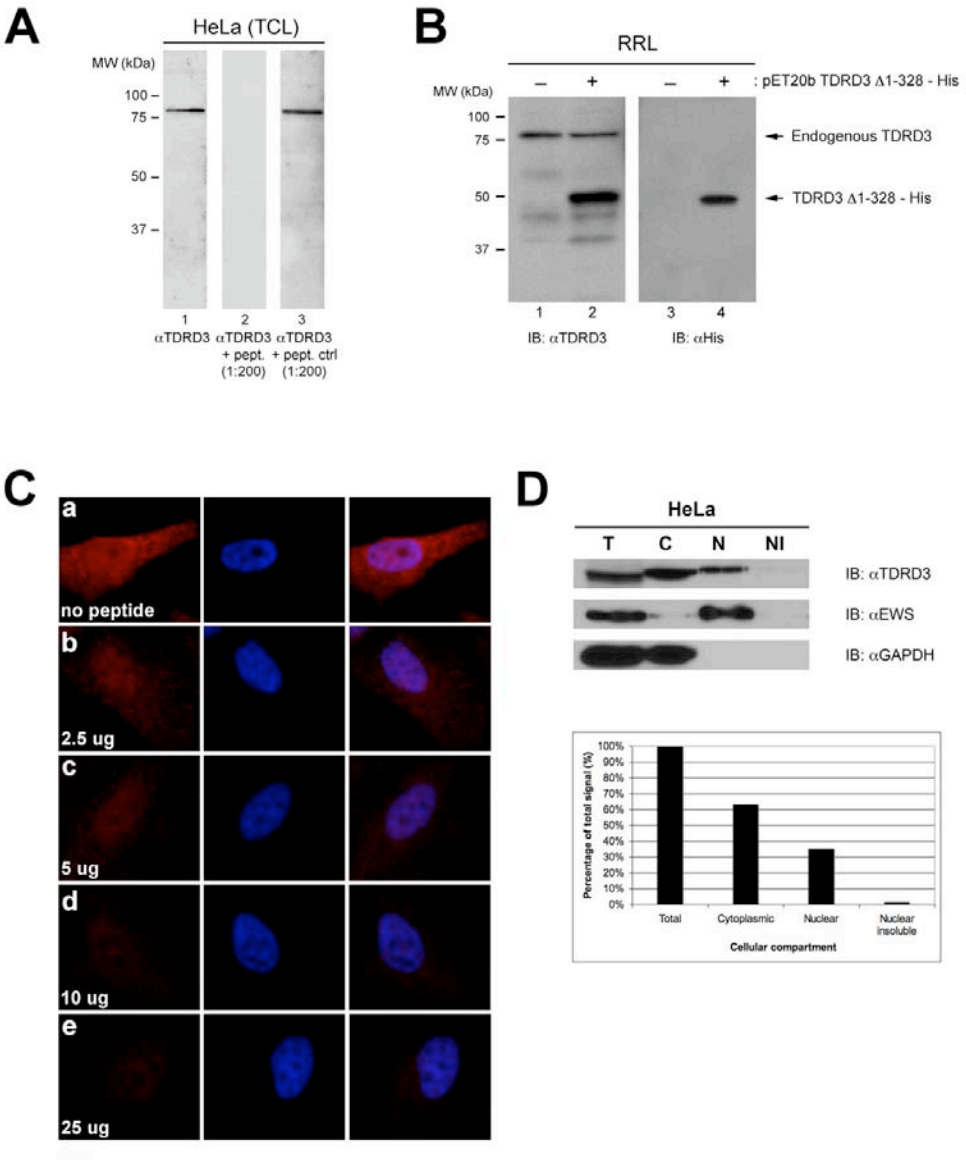


FIGURE 2. **A polyclonal antibody against the novel TDRD3 protein.** A HeLa total cell extract was immnublotted using our affinity-purified anti-TDRD3 polyclonal antibody (lane 1). Alternatively, immunoblots were performed using the TDRD3 antibody pre-incubated with a 200-fold molar excess of either the antigenic peptide (lane 2) or an unrelated peptide (ctrl; lane 3) (**A**). *In vitro* transcription and translation reactions were programmed to express a hexahistidine-tagged truncated version of TDRD3 (ΔN328). Control (-) and programmed (+) reactions were resolved by SDS-PAGE, transferred to a PVDF membrane, and immunoblotted with either TDRD3 (lanes 1 and 2) or His (lanes 3 and 4) antibodies (**B**). Actively growing HeLa cells were labeled for immunofluorescence with our TDRD3 antibody (panel a). Alternatively, TDRD3 antibodies were preincubated with an increasing amount (2.5 to 25 ⌠g) of the immunogenic peptide for 1 h on ice prior to immunofluorescence staining (panels b to e) (**C**). HeLa cells were fractionated using the Qproteome Nuclear Protein Kit (Qiagen). 2.5% of each fraction was resolved by SDS-PAGE, transferred to PVDF, and immunoblotted with affinity-purified anti-TDRD3. Cellular fractionation efficiency was assessed using antibodies against the nuclear protein EWS and the cytoplasmic protein GAPDH. Total (T), cytoplasmic (C), nuclear (N), and insoluble nuclear protein (NI) fractions are shown (top panel). Quantification of the relative abundance of TDRD3 in each fraction is presented in a bar graph (bottom panel) (**D**).

purified antibody with an increasing amount of the immunogenic peptide resulted in a gradual decrease in the intensity of the immunofluorescence signal, demonstrating the specificity of the observed staining pattern (Fig. 2*C*b-e).

In order to confirm the observed TDRD3 subcellular distribution, we fractionated HeLa cells using the Qproteome Nuclear Protein Kit (Qiagen). Total proteins from each fraction were resolved by SDS-PAGE and immunoblotted with our TDRD3 antibodies, as well as with nuclear (EWS) and cytoplasmic (GAPDH) markers antibodies to assess the efficiency of our fractionation (Fig. 2*D*, top). Relative abundance of TDRD3 in each fraction was determined by quantification of band intensity using the *ImageJ* version 1.37 software and expressed as a percentage of the total signal obtained from the addition of all fractions' intensities. Although detectable in the nuclear and nuclear insoluble fractions (37%), TDRD3 was found predominantly associated with the cytosolic fraction (63%) (Fig. 2*D*, bottom), which correlated well with the subcellular distribution observed by immunofluorescence. Taken together, these observations suggest that TDRD3 is found predominantly in the cytoplasmic compartment.

TDRD3 is a novel component of cytoplasmic stress granules

In an attempt to identify in which cellular structures TDRD3 specifically localizes in the cytoplasm, we transfected HeLa cells with commercially available pECFP vectors encoding the enhanced cyan fluorescent protein fused to either an endoplasmic reticulum (pECFP-ER), a Golgi apparatus (pECFP-Golgi) or a mitochondrial targeting sequence (pECFP-Mito). The cells were fixed and indirect immunofluorescence staining was

performed using TDRD3 antibodies. Surprisingly, TDRD3 was found concentrated in cytoplasmic granule-like foci in most transfected cells (data not shown), which led us to hypothesize that these cellular foci might be resulting from transfection-induced stress.

In order to confirm the identity of these TDRD3-containing foci, HeLa cells were exposed to different stress stimuli and double immunofluorescence experiments were performed using TDRD3 antibodies in combination with an antibody directed against the RNA-binding protein TIA-1, a well-accepted marker of cytoplasmic stress granules (SGs). Upon oxidative, heat, and osmotic shocks, endogenous TDRD3 was redistributed to cytoplasmic SGs, where it co-localized with TIA-1 (Fig. 3*A*). For simplicity, sodium arsenite was used in all subsequent experiments to induce cellular stress response. The Fragile X Mental Retardation protein (FMRP) is another protein that is well-known to relocalize to SGs (289, 290), and it was recently found to interact with TDRD3 (254). Accordingly, a colocalization of TDRD3 and FMRP into SGs following stress was observed (Fig. 3*B*). Relocalization of TDRD3 in SGs was further confirmed by its co-localization with a GFP fusion of another component of these granules, the RNA-binding protein G3BP (Fig. 3*C*, bottom, Ars). G3BP overexpression is known to induce the formation of SGs, even in the absence of any other stress stimuli (291), and as anticipated, TDRD3 was also observed in G3BP-induced SGs (Fig. 3*C*, top, C). Importantly, TDRD3 was detectable in SGs as early as 10 min following sodium arsenite treatment (Fig. 3*D*), which is comparable to the kinetics of TIA-1 relocalization to SGs. Since TIA-1 is thought to serve as a nucleator for SG assembly (292), these results

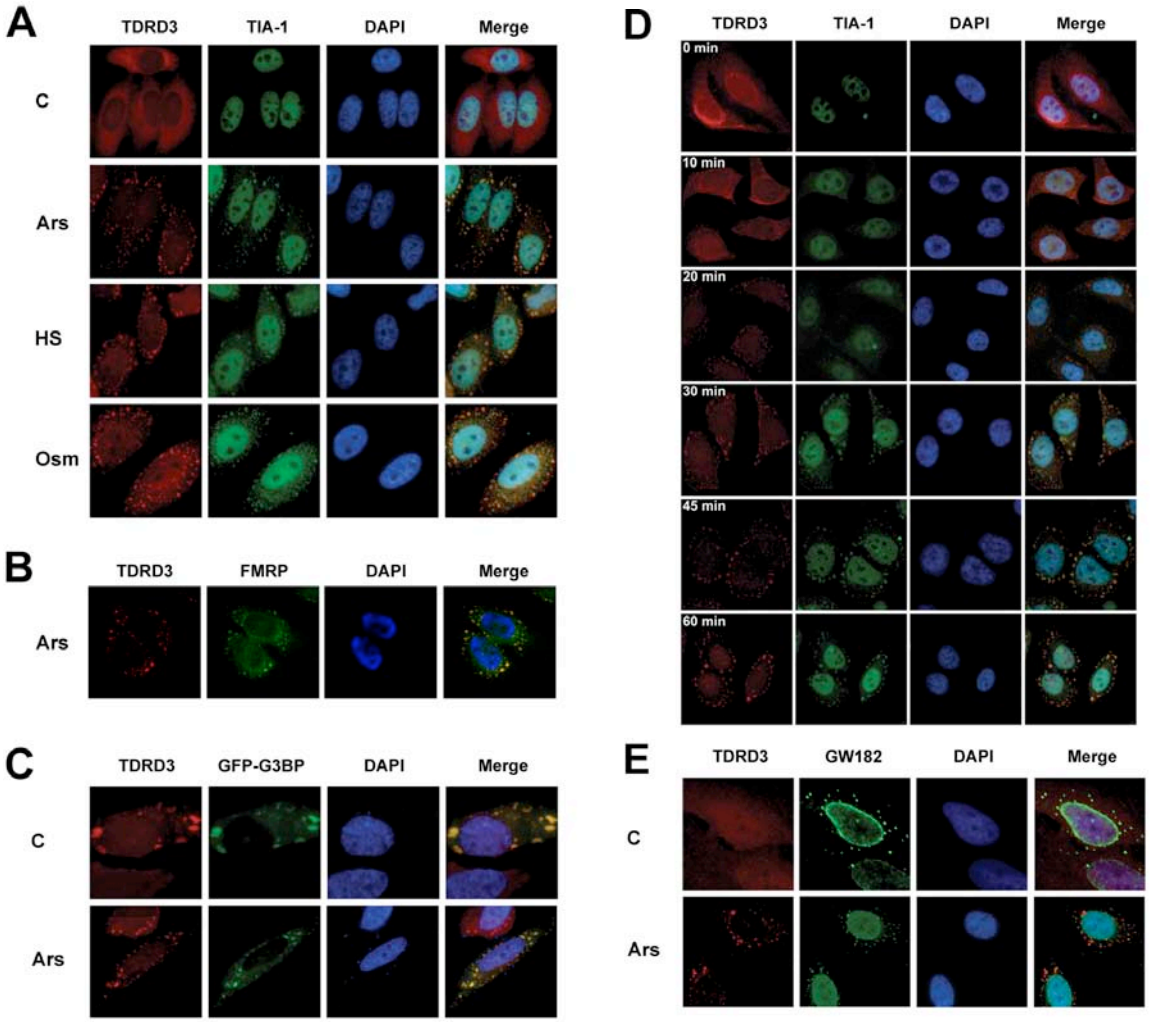


FIGURE 3. **TDRD3 localizes to cytoplasmic stress granules during stress response.** HeLa cells cultured on glass coverslips were left untreated (C) or exposed to oxidative stress (Ars; 0.5 mM sodium arsenite for 30 min), heat shock (HS; 43ºC for 30 min), or high-osmolarity medium (Osm; 1 M sorbitol in DMEM for 1 h followed by a 30-min recovery in normal DMEM). The cells were fixed and immunostained with TDRD3 and TIA-1 antibodies to detect the endogenous proteins (**A**). Cells were stressed with 0.5 mM sodium arsenite (Ars) for 30 min, followed by immunostaining of both TDRD3 and FMRP endogenous proteins (**B**). To visualize G3BP, a GFP fusion construct was transfected into cells 24 h prior to immunofluorescence analysis. Cells were left untreated (C) or exposed to oxidative stress (Ars) for 30 min, followed by immunostaining for endogenous TDRD3 (**C**). Cells were stressed with 0.5 mM sodium arsenite for the indicated time length before their fixation and double immunofluorescence with TDRD3 and TIA-1 antibodies (**D**). HeLa cells were either left untreated (C) or stressed for 30 min with 0.5 mM sodium arsenite (Ars). After fixation, endogenous TDRD3 and GW182 (a marker of P-bodies) were detected by fluorescence microscopy (**E**).

suggest that TDRD3 may also participate in this process. However, further experiments will be required to confirm this hypothesis. Finally, TDRD3 did not localize to other cytoplasmic mRNA foci, such as P-bodies, as shown by the absence of co-localization with GW182 (Fig. 3*E*), an RNA-binding protein resident of these structures (293, 294). Taken together, our results indicate that TDRD3 is a novel component of cytoplasmic SGs.

TDRD3 is associated with translating polyribosomes

SGs are considered to be stalled aggregates of 48S preinitiation complexes and to contain a number of translation factors, including eIF3, eIF4E, eIF4G, PABP-1, and FMRP (283). In contrast to TIA-1, which is not normally associated with the translational machinery (295), FMRP can be found on polyribosomes (296, 297). To determine if TDRD3 is associated with translational complexes, post-nuclear supernatants were subjected to velocity sedimentation through sucrose gradients and each collected fraction was analyzed by Western blot for the presence of TDRD3 (Fig. 4, control panels). Using this protocol, it is possible to resolve free ribosomal subunits (40S and 60S) and monomers (80S) from actively translating polysomes, which sediment towards to bottom of the gradient, as confirmed by the distribution of the ribosomal S6 protein in each fraction (Fig. 4, S6 panel). Strikingly, TDRD3 was found mainly distributed in the heavy sedimenting fractions of the sucrose gradient, following a pattern similar to what is observed for FMRP (Fig. 4, respective panels). Hsp72 was used as a negative control in these experiments as it should not be associated with ribosomes. These results indicate that TDRD3 is associated with actively translating polyribosomes in cycling HeLa cells.

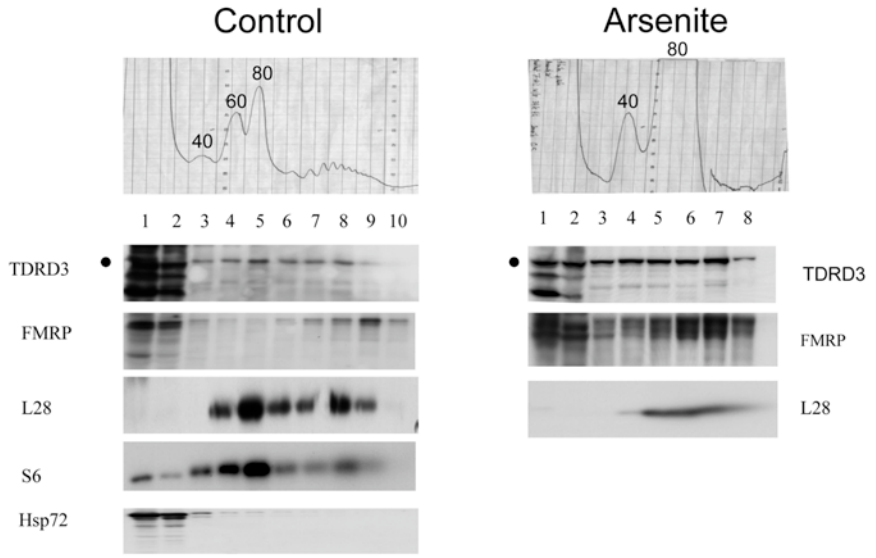


FIGURE 4. **TDRD3 is associated with polyribosomes in HeLa cells.** Cytoplasmic extracts were prepared from HeLa cells grown in normal conditions and centrifuged on a 10-60% w/w linear sucrose gradient. Fractions were collected and analyzed by Western blot with antibodies against the ribosomal S6 protein and FMRP (as positive controls), TDRD3, and Hsp72 (as a negative control) (Control panel). Cytoplasmic extracts from HeLa cells treated with 0.5 mM sodium arsenite for 30 min were analyzed in parallel and the collected fractions were analyzed by Western blot using the ribosomal L28 protein as positive control (Arsenite panel). Fractions from the top to the bottom of the gradient are shown from left to right. The positions of free small (40S) and large (60S) ribosomal subunits, monosomes (80S), and polysomes are indicated in each profile. The band corresponding to TDRD3 is indicated by a 'dot' on the side of the respective panels. Additional bands detected on the immunoblots represent non specific reactivity with our polyclonal antibody.

Under stress conditions such as heat shock or arsenite, FMRP shifts away from polyribosomes, which are largely disassembled due to inhibition of translation initiation (283, 289, 290, 298). Thus, we next determined if TDRD3 would follow the same pattern in ribosome profiles following stress induced by arsenite treatment (Fig. 4, Arsenite panels). HeLa cells were either left untreated (Control) or incubated with sodium arsenite as above (Arsenite), before being subjected to velocity sedimentation through sucrose gradients. Each collected fraction,was then analyzed by Western blot for the presence of TDRD3. As expected, polysome peaks were almost undetectable following arsenite treatment, and this was associated with an increase of the 80S monomer fraction (Fig. 4,

Arsenite UV profile). Under these stress conditions, the bulk of TDRD3 was now found associated with monomeric ribosomes and lighter fractions, again following the pattern previously reported for FMRP (Fig. 4, respective panels). Taken together, these results suggest that TDRD3's function may be at the level of translational regulation.

The Tudor domain is both essential and sufficient for TDRD3 translocation to cytoplasmic stress granules

To identify which part of TDRD3 was important for its localization to SGs, a series of myc epitope-tagged deletion constructs was generated (Fig. 5*A*). The fusion proteins were expressed in HeLa cells and immunofluorescence microscopy was performed using myc antibodies (Fig. 5*B* and *E*). As its endogenous counterpart, full length myc-tagged TDRD3 localized predominantly to the cytoplasm with a weaker diffuse nuclear staining (Fig. 5*B*a). Overall, the intracellular distribution patterns of deletion mutants were similar to that of full length TDRD3 (Fig. 5*B*b-e), although a certain degree of variation was observed depending on expression levels. A second series of constructs was engineered to express specific conserved domains of the TDRD3 protein (Fig. 5*A*). The predicted molecular weight of these myc-tagged TDRD3 isolated domains ranges from 4 to 36 kDa, which means that they are likely to diffuse freely between the cytoplasm and the nucleus. As expected, expression of TDRD3's isolated domains in HeLa cells resulted in both cytoplasmic and nuclear diffused staining (Fig. 5*B*f-i). Finally, in contrast with G3BP, overexpression of TDRD3 (or any deletion mutant) did not induce the formation of SGs in the absence of external stress stimuli. Taken together, these experiments did not permit the clear identification of a domain influencing normal TDRD3 subcellular localization.

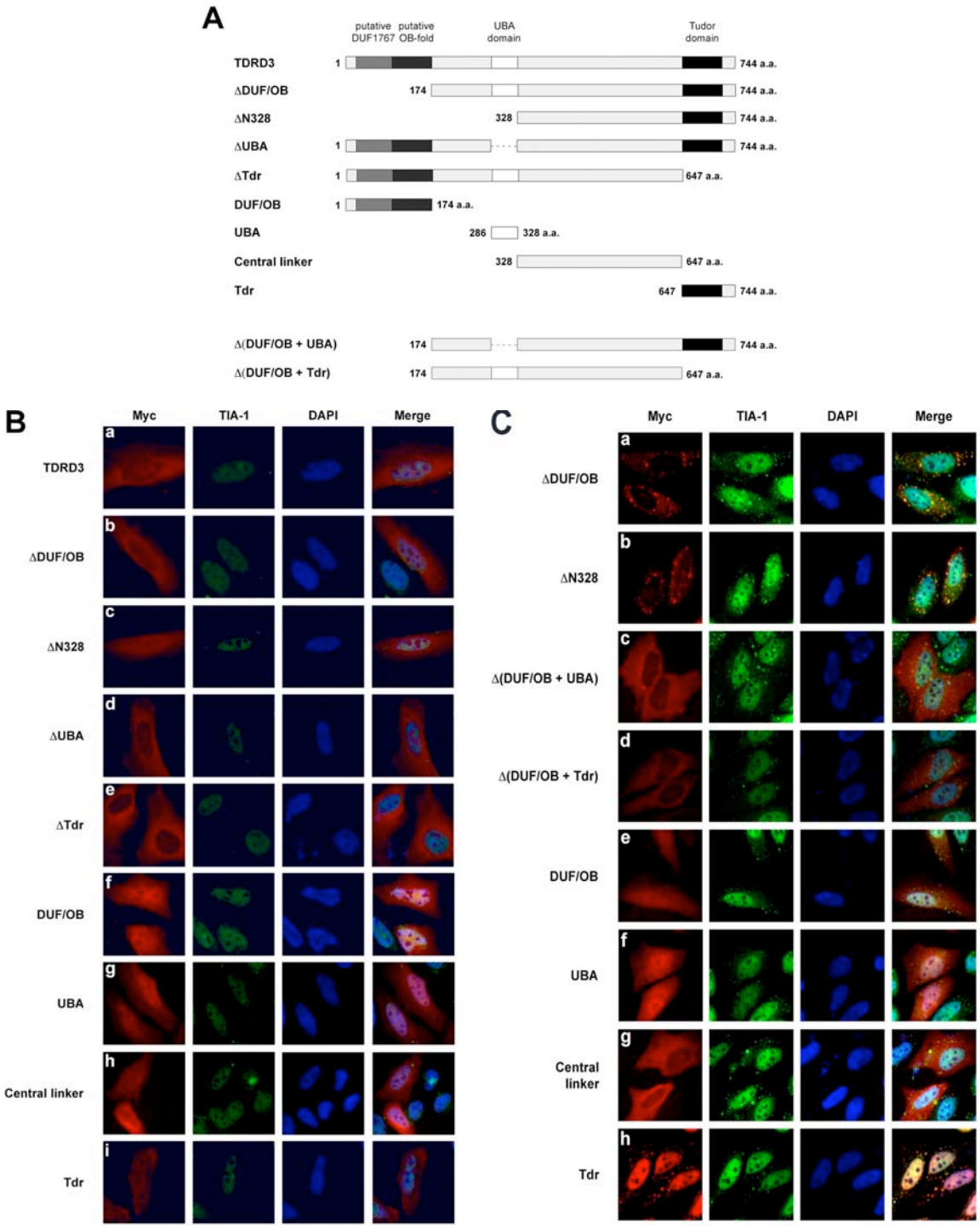

A
putative DUF1767
putative OB-fold
UBA domain
Tudor domain
TDRD3
ΔDUF/OB
ΔN328
ΔUBA
ΔTdr
DUF/OB
UBA
Central linker
Tdr
Δ(DUF/OB + UBA)
Δ(DUF/OB + Tdr)
1
174
328
286
647
744 a.a.
174 a.a.
328 a.a.
647 a.a.
B
C
Myc
TIA-1
DAPI
Merge
a
b
c
d
e
f
g
h
i

FIGURE 5.

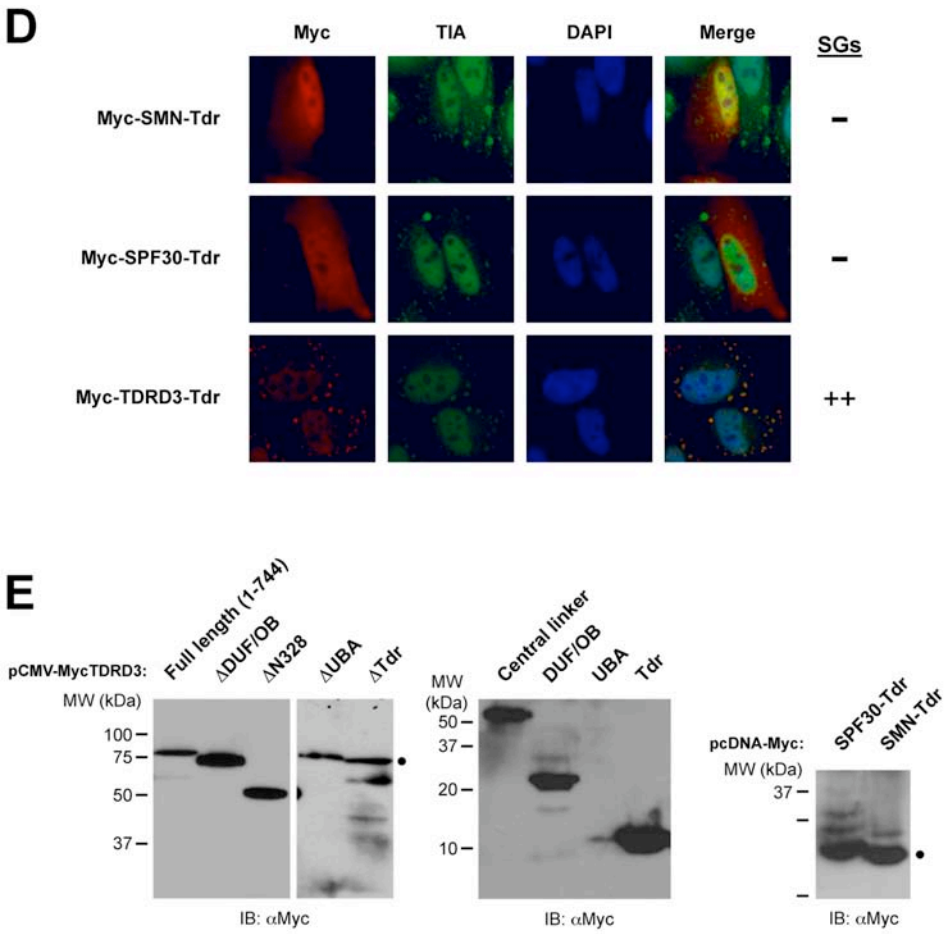


FIGURE 5 (continued). **The Tudor domain of TDRD3 is both required and sufficient for its recruitment to stress granules.** Diagram showing the various myc epitope-tagged TDRD3 deletion mutants used in this study (**A**). HeLa cells were transiently transfected with each deletion mutant. 24 h post-transfection, cells were left untreated and labeled for immunofluorescence with TIA-1 and myc antibodies to detect endogenous TIA-1 protein and recombinant myc-tagged proteins (**B**). Alternatively, transfected cells were treated with 0.5 mM sodium arsenite for 30 min. Indirect immunofluorescence staining was performed as described above (**C**). HeLa cells were transiently transfected with constructs expressing recombinant myc-tagged Tudor domain of SMN, SPF30, or TDRD3. 24 h post-transfection, cells were treated with 0.5 mM sodium arsenite for 30 min. Indirect immunofluorescence staining was performed using myc and TIA-1 antibodies (**D**). Total cell lysates of HeLa cells transiently expressing myc-tagged TDRD3 deletion mutants, as well as SMN and SPF30 Tudor domains, were immunoblotted with myc antibodies to confirm expression (**E**).

We next assessed whether each TDRD3 deletion mutant had retained the property of being relocalized to cytoplasmic SGs following sodium arsenite treatment (Fig. 5*C*). Intriguingly, over-expression of any myc-tagged TDRD3 proteins harboring the DUF1767/OB-fold motifs was toxic for cells when they were put under stress conditions.

In order to circumvent this problem, we generated additional deletion mutants of TDRD3, lacking the N-terminal DUF1767/OB-fold domains [Fig. 5*A*; TDRD3 Δ(DUF/OB + UBA) and TDRD3 Δ(DUF/OB + Tdr)]. These mutants displayed an intracellular distribution that was identical to their respective N-terminal-containing counterparts in the absence of stress (data not shown). Following arsenite treatment, ΔDUF/OB, ΔN328, and Δ(DUF/OB + UBA) all still relocalized to SGs, albeit to varying degrees (Fig. 5*C*a-c). In contrast, Δ(DUF/OB + Tdr) was not detectable in cytoplasmic SGs following stress stimuli (Fig. 5*C*d), which suggest a requirement for the Tudor domain in this process. Accordingly, when TDRD3's isolated domains were tested, only the Tudor domain (Tdr) was found to be sufficient for relocalization to SGs upon arsenite treatment (Fig. 5*C*e-h). Interestingly, the central linker fragment (amino acids 328-647) was redistributed almost entirely to the cytoplasm following arsenite treatment (compare Fig. 5*B*h with 5*C*g), suggesting that this region may contain an element favoring its export from the nucleus to the cytoplasm upon stress. Taken together, these results show that the Tudor domain of TDRD3 is both required and sufficient for its recruitment to cytoplasmic SGs during cellular stress response, although the central linker region may also contribute to this process by promoting nuclear-cytoplasmic translocation.

SMN, which also harbors a Tudor domain, was previously reported to relocalize to SGs upon stress (299). Hence, we next wanted to determine whether this feature was unique to the Tudor domain of TDRD3 or common to other known Tudor domains. To investigate this possibility, HeLa cells were transiently transfected with the isolated Tudor domain of SMN, SPF30, or TDRD3 fused to a myc epitope tag (Fig. 5*D* and *E*).

Twenty-four post-transfection, cells were treated with sodium arsenite and immunostained with myc antibodies. The Tudor domain of TDRD3 was effectively recruited to stress granules upon stress (Fig. 5*D*). However, neither SMN nor SPF30's Tudor domain localized to SGs under the same conditions (Fig. 5*D*). These observations therefore suggest that the mechanism by which TDRD3 relocalizes to SGs likely involves interactions that are mediated by and that are unique to its Tudor domain.

Identification of TDRD3 Tudor domain interactors

In order to begin the identification of proteins interacting specifically with the Tudor domain of TDRD3, a GST pull-down was performed using HeLa total cell extracts. A recombinant fusion protein was generated, consisting of glutathione-S-transferase (GST) fused to the C-terminal region of TDRD3 (a.a. 647-744) encompassing the Tudor domain (GST-TDRD3-Tdr). GST-TDRD3-Tdr and GST alone were then coupled to glutathione-agarose and used as affinity columns to pull down interacting proteins from HeLa cell lysates. The retained proteins were resolved by SDS-PAGE and stained with Coomassie Blue. GST-TDRD3-Tdr bound several polypeptides that were not bound by GST alone (data not shown). Five of these specific protein bands were excised from the gel and subjected to mass spectrometry identification (300). The top matches obtained for each band are EWS, FUS, DDX3, SERBP1 and EEF1A (see Table I for details). Remarkably, all these proteins have known roles in RNA handling, which reinforced our hypothesis that TDRD3 functions in post-transcriptional processes.

TABLE I. **Potentially methylated proteins involved in RNA processing interact with the Tudor domain of TDRD3.**

MW SDS-PAGE	Protein name (GenPept acc. no.)	Score	RG-motifs	Known or putative function(s)[a]	Calc. mass (kDa)
p80	Ewing sarcoma breakpoint region 1 (NP_005234)	70.26	19 x RGG ^{309}RGRGRGG^{316} ^{314}RGGRGGGRGG^{323}	RNA binding protein (transcription and pre-mRNA splicing machineries)	68.592
p75	DEAD/H box-3 (NP_001347)	40.17	10 x RG ^{121}RGG^{123}	RNA splicing and transport	73.244
p60	Fus (NP_004951)	40.15	2 x RGG ^{216}RGGRGRG^{222} ^{245}RGRGGGRGGRGG^{256} ^{389}RGGGNGRGGRG RGGPMGRGG^{399} ^{475}RRGGRGGYDRGGYRGRGG DRGGFRGGRGGGDRGG^{508}	RNA splicing and nuclear/cytoplasmic shuttling	53.377
p50	Eukaryotic translation elongation factor 1 alpha 1 (NP_001393)	150.19	2 x RG	Translation anchoring	50.141
p35	SERPINE1 mRNA binding protein 1 (NP_001018077)	60.11	1 x RG, 1 x RGG ^{163}RGRGGLGRGRG GRGRGMGRG^{182} ^{346}RGGRGGRGGRGRGGR^{360}	Regulation of mRNA stability	44.966

HeLa cells were lysed and incubated with either purified recombinant GST (as a control) or GST-TDRD3-Tdr proteins coupled to glutathione-agarose (see Materials & Methods). The polypeptides specifically pulled down by TDRD3-Tdr, that were not pulled down by GST alone, were excised from the gel and sent for mass spectrometry analysis and identification. Sequencing data were analyzed using the Sequest software and only the top match obtained for each band is listed (highest score). Potentially methylated RG-rich motifs found in the amino acid sequence (GenPept database) of each protein are also depicted, as well as their known or putative function (see text for details and references).
[a]see text for reference.

Interestingly, EWS was previously shown to interact with the Tudor domain of SMN (301), although arginine methylation was not a requirement for binding in this case. A *C. elegans* orthologue of SERBP1 was also found to interact with SMN (302). Finally, FUS is present in a complex with SMN and the NFAR proteins (303), although a direct interaction between SMN and FUS was not reported. These observations prompted us to verify if all the proteins identified in our screen were also able to bind the Tudor domain of SMN. These experiments revealed that EWS and FUS were also interacting with SMN-Tdr, but the interactions with DDX3 and EEF1A were specific to TDRD3 (Goulet and Côté, unpublished data). SERBP1 was not tested in these experiments. The functional relevance of each of these interactions will need to be addressed in future

studies.

Direct contact with methylated arginines contribute to TDRD3 relocalization to SGs

All the identified proteins also contain RG-rich motifs, which represent common target sites for protein arginine methyltransferases (Table I). In fact, the presence of methylated arginines in EWS and FUS has been documented (114, 118, 304), and we have also detected modified arginines in our proteomic analysis (data not shown). Since the Tudor domain of TDRD3 can specifically recognize methylated arginine motifs in proteins (148, 158), we speculated that TDRD3 might be recruited to SGs through methyl-dependent interaction with one or more of its binding partners. Previous studies have demonstrated that an E134K amino acid substitution in the Tudor domain of SMN abolished its interaction with arginine-methylated polypeptides (148). This amino acid is conserved in the Tudor domain of TDRD3 and corresponds to residue E691. In order to determine if the equivalent E691K substitution in TDRD3 Tudor domain would also abolished its capacity to interact with methylated arginines, an *in vitro* binding assay was used, with synthetic $(RG)_4$ biotinylated peptides containing either no modified amino acids or symmetrically dimethylated arginines (sDMA). Purified recombinant GST-TDRD3-Tdr fusion proteins were incubated with the biotinylated peptides coupled to streptavidin-agarose. As previously reported (148), purified GST-TDRD3-Tdr bound to sDMA peptide columns in a dose-dependent manner, but did not bind the unmethylated peptide (Fig. 6*A*, compare lanes 5-7 with lanes 2-4). Remarkably, this binding was completely abolished with the introduction of the E691K mutation in the TDRD3 Tudor domain (Fig. 6*A*, GST-TDRD3tdr E691K). This result confirms that the Tudor domain of

TDRD3 behaves as a methyl-binding protein module similar to that of SMN, in which the E691 residue is crucial for direct contact with dimethylated arginine motifs. This TDRD3 mutant was next used to determine if this methylarginine binding surface was important for recruitment to SGs. HeLa cells were transfected with a myc-tagged TDRD3 Tudor fragment containing the E691K point mutation. 24 h post-transfection, cells were stressed with sodium arsenite and labeled for immunofluorescence with myc antibodies. Strikingly, the introduction of the E691K mutation in the Tudor domain of TDRD3 reduced its recruitment to stress granules upon stress by at least 50% (Fig. 6*B*). This observation suggests that TDRD3's recruitment to stress granules is likely mediated through specific protein-protein interactions involving methylated arginines.

In order to fulfill this role, an interacting protein would have to satisfy at least two conditions. First, the protein has to be relocalized to SGs upon stress. Second, its interaction with TDRD3 must be affected by the E691K mutation. To verify the first criteria, immunofluorescence microscopy was performed using commercially available antibodies for EWS, DDX3, and FUS, or epitope-tagged versions of EEF1A (GFP) and SERBP1 (myc) (Fig. 6*C*). Strikingly, only DDX3 and SERBP1 relocalized significantly to SGs following arsenite treatment, while the other proteins showed little to no response to the stress stimuli (Fig. 6*C*, respective panels). To verify the second condition, a GST-TDRD3-Tdr pull down was performed as above using total HeLa cells lysates as input, and the retained proteins were analyzed by Western blot with commercially available antibodies specific to EWS, DDX3, FUS and EEF1A (Fig. 6*D*, lane 3). The E691K

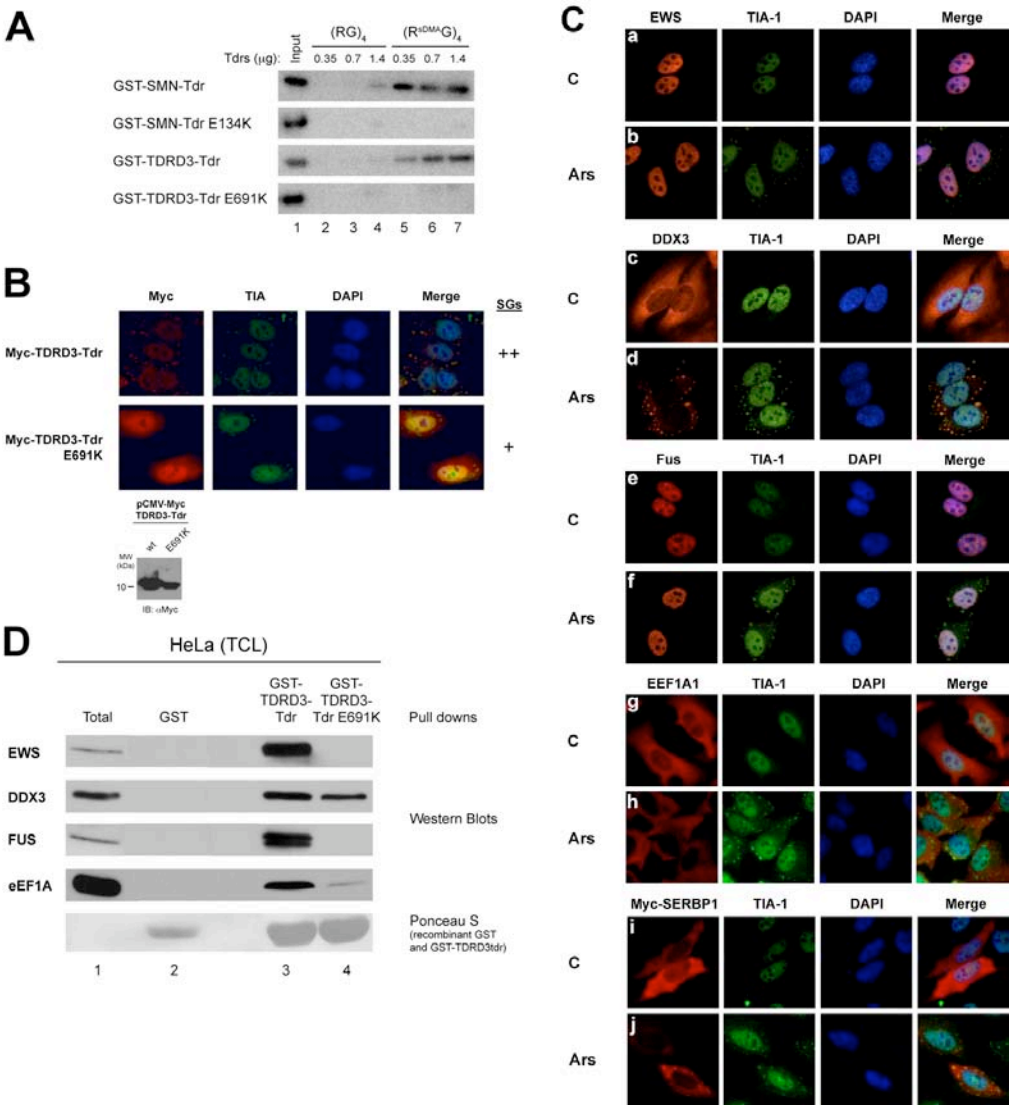


FIGURE 6. **Direct contact with methylated arginines contributes to TDRD3 relocalization to SGs.** Biotinylated (RG_4) peptides containing either unmethylated arginines or symmetrically dimethylated arginines (sDMA) were bound to streptavidin-agarose and used as affinity columns to measure the binding of purified GST fusion SMN and SMN E134K (as controls) and/or TDRD3 and TDRD3 E691K Tudor domains. The bound GST-Tdr fusion proteins were resolved by SDS-PAGE, transferred to a PVDF membrane, and detected by immunoblotting using GST antibodies (**A**). HeLa cells were transiently transfected with constructs expressing recombinant myc-tagged wild type or mutated (E691K) TDRD3 Tudor domain. 24 h post-transfection, cells were treated with 0.5 mM sodium arsenite for 30 min. Indirect immunofluorescence staining was performed using myc and TIA-1 antibodies (**B**). HeLa cells were either left untreated (C) or treated with 0.5 mM sodium arsenite for 30 min (Ars). Cells were then labeled for immunofluorescence with EWS (panels a and b), DDX3 (panels c and d), FUS (panels e and f), or EEF1A1 (panels g and h) in combination with TIA-1 antibodies to detect endogenous proteins. HeLa cells transiently transfected to express myc-tagged SERBP1 (panels i and j) were either left untreated (C) or stressed as described above (Ars), before being immunostained with TDRD3 and myc (SERBP1) antibodies (**C**). HeLa cells from 2 x 150 mm plates were lysed and incubated with purified recombinant GST (as a control), GST-TDRD3-Tdr, or GST-TDRD3-Tdr E691K proteins coupled to glutathione-agarose. The retained proteins were resolved by SDS-PAGE, transferred to PVDF, and immunoblotted with the specified antibodies to confirm mass spectrometry identifications. The membrane was stained with Ponceau Red prior to immunoblotting, in order to show the GST fusion proteins (**D**).

mutation completely prevented the interaction with EWS and FUS, while reducing binding to DDX3 and EEF1A by 50% and 90%, respectively (Fig. 6*D*, lane 4). The E691K mutation completely prevented the interaction with EWS and FUS, while reducing binding to DDX3 and EEF1A by 50% and 90%, respectively (Fig. 6*D*, lane 4). We have not yet assessed if the interaction of SERBP1 with TDRD3 was also affected by the presence of this mutation due to antibody availability. Taken together, these results identify DDX3 as a good candidate to mediate the recruitment of TDRD3 to SGs (or vice versa). Moreover, these experiments show that the methyl-binding surface of TDRD3's Tudor domain is important for this process.

Arginine methylation: a general regulator of SG assembly/disassembly?

A recent study uncovered a new subset of smaller cytoplasmic foci that are clearly distinct from SGs, which contain FMRP (305). FMRP is an RG-rich protein that is an *in vitro* substrate for protein arginine methyltransferases (306). FMRP relocalizes to *bona fide* SGs following treatment with various stimuli, and it was suggested that arginine methylation may regulate its translocation from smaller cytoplasmic foci into SGs by promoting its heterodimerization with FXR1p (305). Since this mechanism is reminiscent of what we propose for the recruitment of TDRD3 along with DDX3 and/or SERBP1 to SGs, we wanted to determine if arginine methylation may play a general role in the assembly of SGs. HeLa cells were grown for 20 h in the presence of 1 mM 5'-deoxy-5'-methylthioadenosine (MTA), a well-characterized general inhibitor of methylation (105). Following this pre-treatment, which reduced steady state arginine methylation levels by ~

50% (Fig. 7*C*), cells were incubated with sodium arsenite and immunofluorescence microscopy was performed with antibodies against TDRD3 and TIA-1 to monitor the formation of SGs. First, it was confirmed that the MTA treatment alone did not cause the formation of SGs (data not shown). As observed in non-treated HeLa cells (see time-course in Fig. 3*D*), TDRD3 was detectable in nascent SGs as early as 10-15 min after the addition of arsenite in Mock-treated cells, and TIA-1 after ~ 20 min (data not shown). Stress granules containing both TDRD3 and TIA-1 were present in 100% of cells after 30 min. Similarly, cells pre-treated with the methylation inhibitor MTA all contained TDRD3- and TIA-1-positive SGs after 30 min of stress stimulus, although small differences were observed at shorter time-points. Specifically, recruitment of TDRD3 to SGs was delayed to some extent, while more cells with TIA-1-containing SGs could be clearly discerned after 15 min (data not shown). Thus, reducing arginine methylation levels by ~ 50% did not prevent assembly of SGs, although the precise composition of these granules seemed to be altered. Another series of stressed cells (pre-treated or not with MTA) was immunostained for TDRD3 at various time points after being returned to normal growth media. As expected from previous studies (307), SGs first decreased in number, but increased in size, and then gradually disappeared following stress withdrawal, so that they were no longer detectable after 120-150 min in mock-treated cells (Fig. 7*A*, DMSO panels). In contrast, the number of SG-positive cells had increased by ~ 6 fold in cells pre-treated with the methylation inhibitor MTA after 150 min of recovery (Fig. 7*A*, MTA panels, and 7*B*). These observations were also confirmed using TIA-1 as a SG marker (data not shown). Taken together, our experiments support a role for methylation in regulating dynamics of the SG assembly and disassembly process,

potentially by regulating protein-protein interactions and impacting on the composition of SGs.

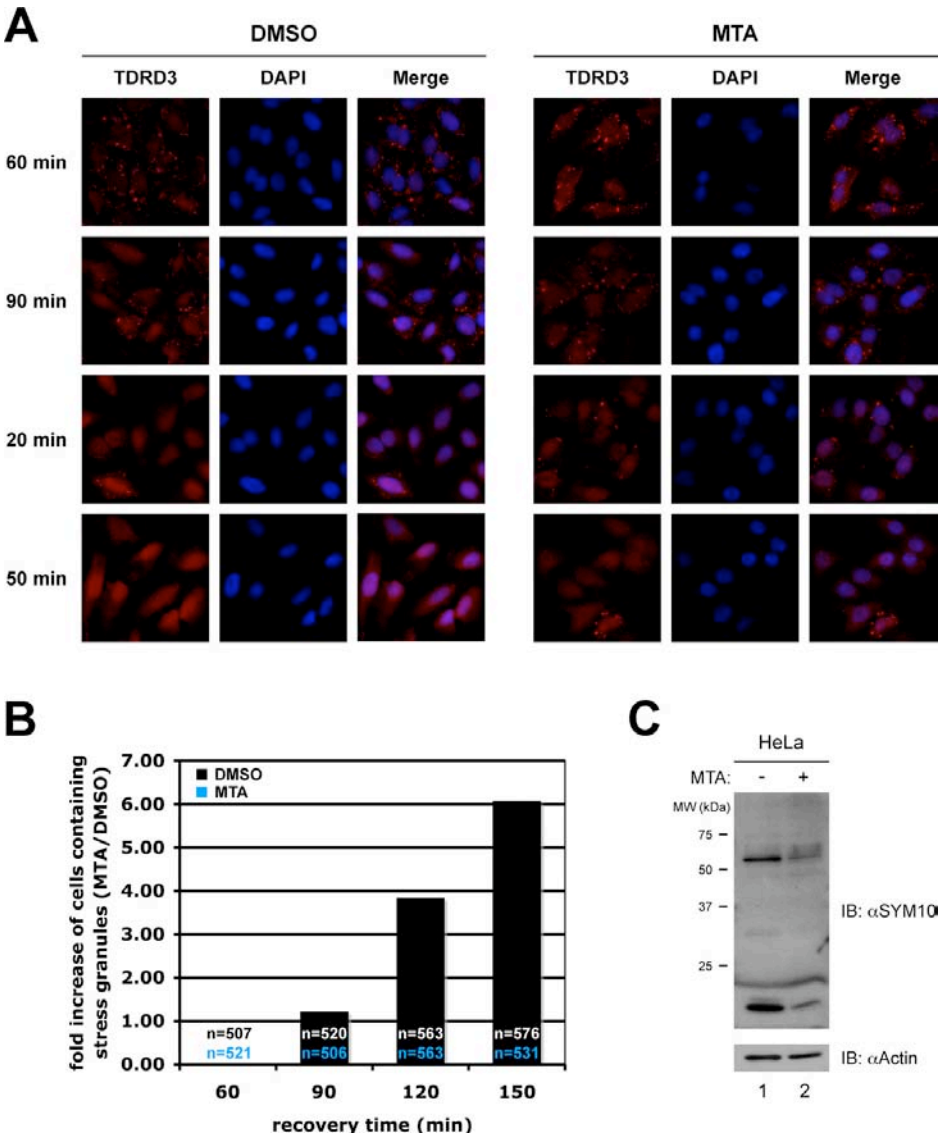


FIGURE 7. **Arginine methylation and SG dynamics**. HeLa cells grown on coverslips were stressed with 0.5 mM sodium arsenite for 30 min. Cells were then returned to normal growth media and prepared for TDRD3/TIA-1 immunostaining after the indicated incubation periods (**A**). The experiment was repeated twice using different batches of cells. At least 500 cells were counted for each time point and condition. Fold increase of SG-positive cells in the MTA-treated cells as compared to DMSO-treated cells is depicted in a bar graph (**B**). Cell extracts prepared from HeLa cells grown in the presence (+) or the absence (-) of the general methylation inhibitor MTA were immunoblotted with the SYM10 antibody to confirm the reduction in steady state arginine methylation levels. The same extracts were immunoblotted with actin antibodies to control for equal loading (**C**).

Discussion

We report here the first characterization of TDRD3, a novel Tudor-containing protein associated with poor prognosis of estrogen receptor-negative breast cancers (167). Our findings identify TDRD3 as a novel component of cytoplasmic SGs, where it relocalizes following various cellular stresses, including oxidative, heat and osmotic shocks. We also show that TDRD3 follows FMRP in polysome profiles under normal and stress conditions. A proteomic screen led to the identification of five novel TDRD3 interacting partners with known or putative roles in various aspects of RNA metabolism. Moreover, all these TDRD3 interacting proteins harbor RG-rich motifs that represent consensus sites for protein arginine methyltransferases. Our experiments revealed that two of these proteins, SERBP1 and DDX3 (a gene associated with Sertoli-cell-only syndrome), are also novel constituents of cytoplasmic SGs. The Tudor domain of TDRD3 was both required and sufficient for its relocalization to SGs. Specifically, the use of a conserved mutation helped us demonstrate that the methyl-binding surface in the Tudor domain of TDRD3 is involved in this process. Finally, we found that arginine methylation may be a general regulator of SG dynamics and protein composition.

TDRD3 is a novel component of cytoplasmic Stress Granules

We report here that the novel Tudor-containing protein TDRD3 relocalizes with TIA-1, a well-known SG marker upon arsenite treatment, heat shock and osmotic shock (Fig. 3*A*). TDRD3 also colocalized in SGs with FMRP following arsenite treatment, and it can be recruited to transfection-induced SGs upon GFP-G3BP overexpression (Fig. 3*B* and 3*C*,

respectively). In contrast, TDRD3 did not colocalize with processing body (also known as GW bodies) markers under any conditions tested here (Fig. 3*E* and data not shown). TIA-1 is considered to be a crucial mediator of SG assembly, as it is thought to promote aggregation through its prion-like glutamine-rich C-terminal domain (292). FMRP is also thought to participate, along with TIA-1, in the "primary aggregation/nucleation" phase of SG assembly (283). Since we found that TDRD3 can be detected in SGs concomitantly or even prior to TIA-1 (Fig. 3*C*), it may participate in the early phases of SG assembly. This is also consistent with the association of TDRD3 with polyribosomes (Fig. 4), since SG assembly involves stalled translation initiation complexes, polysome disassembly and mRNP aggregation (283). However, TDRD3 did not significantly induce SG formation on its own in the absence of additional stress stimuli, in contrast to TIA-1, G3BP and FMRP, which are all thought to physically nucleate SG assembly (283). In fact, our experiments suggest that TDRD3 may participate in the SG disassembly process by engaging in protein-protein interactions regulated by arginine methylation (see below), although further experiments will be required to fully confirm this hypothesis.

TDRD3 interacts with proteins involved in RNA handling and/or metabolism

We have uncovered, using a large-scale Tudor domain pull down approach, five novel protein-binding partners for TDRD3 (Table I). Strikingly, we found that two of these proteins, SERBP1 and DDX3, also relocalize to SGs following sodium arsenite treatment, although we cannot rule out that some of the other TDRD3 binding partners may still transiently associate with SGs (Fig. 6*C*). In addition, each of the newly

identified TDRD3 interactors has previously been linked with post-transcriptional regulatory processes. EEF1A1 is a translation initiation factor that has been shown to participate in the anchoring of specific mRNAs to the cytoskeleton (308), while DDX3 has been implicated in pre-mRNA splicing, nuclear export, mRNA transport and translation initiation (309-317). TDRD3 also interacts with FMRP and associates with translating polyribosomes. Taken together, these observations strongly suggest that TDRD3 may function as a translational regulator.

In contrast, SERBP1 (a.k.a. PAI-RBP1) has been suggested to play a role in mRNA stability (318), while EWS and FUS have both been implicated in alternative splicing regulation (319-325). Intriguingly, TDRD3 has been identified in a proteomic screen of proteins associated with spliceosome complexes (152). However, our experiments did not provide any indications to support a role for TDRD3 in splicing (Goulet and Côté, unpublished data). Further experimentation will be needed to establish its function in this precise mechanism.

The Tudor domain of TDRD3 is both required and sufficient for its localization to SGs

All the proteins identified in our GST-TDRD3-Tdr pull down are either known or are very likely to contain methylated arginines (Table I). This is consistent with our previous studies, which demonstrated that the Tudor domain of TDRD3 can specifically recognize methylated arginine motifs in proteins (148). Since DDX3 and SERBP1 are the only proteins that could be detected in SGs, these proteins are good candidates to mediate the methyl-dependent recruitment of TDRD3 to these structures. In this study, we have

presented several evidences supporting this model. First, thorough deletion mapping clearly showed that the Tudor domain of TDRD3 is both sufficient and required for its localization to SGs, although the central linker region may also contribute to this process by promoting nuclear-cytoplasmic translocation (Fig. 5). Second, a single E->K substitution (E691K) in the methylarginine-binding interface of TDRD3's Tudor domain is sufficient to greatly reduce its recruitment to SGs (Fig. 6*B*). Third, this same E691K mutation also reduces the interaction with DDX3 (Fig. 6*D*). Finally, general methylation inhibitors perturbed the dynamics of SG assembly/disassembly as well as their composition (Fig. 7 and see below).

Arginine methylation and Stress Granule dynamics

A simple prediction from our model would have been for methylation inhibitors to prevent or at least reduce the recruitment of TDRD3 to SGs, potentially even preventing their formation. In contrast, our results show that MTA treatment did not prevent SG assembly, as 100% of cells contained a normal number of SGs 30 min after arsenite treatment (data not shown). However, we cannot eliminate the possibility that a 50% reduction in overall arginine methylation may not be sufficient to observe a drastic effect. Moreover, our estimate of the efficiency of the MTA treatment does not take into account the different turnover rate of all the proteins that could be involved in SG assembly. Nevertheless, we did observe a tendency for TDRD3 and TIA-1 to no longer follow the same kinetics of recruitment as seen at early time-points of SG formation. SGs are highly dynamic structures with components rapidly and continuously shuttling in and out (307). Hence, more in-depth studies using Fluorescence Recovery After Photobleaching (FRAP)

experiments or similar approaches would be required to confirm and quantify these effects. Following the same idea, a longer time window may explain why we were able, under the same conditions of MTA treatment, to observe more profound differences in the recovery phase after the return to normal growth conditions (Fig. 7). Taken together, our results can be reconciled by proposing a role for arginine methylation in the regulation of many protein-protein interactions (including but not limited to the TDRD3/DDX3 interaction) important for SG dynamics. This is consistent with, and expands, the model proposed by Denman and Colleagues for the recruitment of FMRP to SGs (305). Finally, this interpretation is also reminiscent of the role that arginine methylation plays in other cellular processes; for instance, the assembly of spliceosomal snRNPs, where it serves to promote the efficient and specific assembly of large ribonucleoprotein macromolecular complexes.

Tudor domains and RNA granules

Several recent studies have provided links between Tudor domain-containing proteins and various types of "RNA granules". We have recently uncovered a novel interaction between the Tudor domain of SMN and KSRP, a KH-type RNA-binding protein that colocalizes with SMN in neuronal RNA granules (122). Moreover, each of the novel TDRD3 interactors (including FMRP) have also been identified as components of neuronal RNA granules (285, 311, 313, 326-328). Neuronal granules are considered to be highly related, both in composition and function, to cytoplasmic SGs (326), since they mediate the transport of specific mRNAs, in a translationally silent state, along dendrites and axons of neuronal cells (329). Whether TDRD3 plays a role in regulating translation

within SGs and/or neuronal RNA granules will require further experimentation. In fact, SMN has been shown to relocalize to SGs per se, although a functional requirement for its Tudor domain was not established in these studies (299). In our hands, relocalization to SGs was a property unique to the Tudor domain of TDRD3, as the isolated Tudor domains of SMN and SPF30 were not found in SGs under any of the conditions tested here (Fig. 5*D* and data not shown). Nevertheless, it is possible that the association of SMN with SGs may be cell type-specific, more transient, or may require other domains of the protein (e.g. its C-ter dimerization domain).

Studies in both *Drosophila* and mouse have uncovered a crucial role for Tudor domain-containing proteins in germline development and polar granules architecture (330-333). Polar or germ granules are specialized structures containing many mRNAs and RNA-binding proteins needed for germ cell formation (334), and it was suggested that several parallels could be drawn between these structures and SGs (326). Sm proteins and PRMT5 were also found to localize to polar granules and to be essential for *Drosophila* germ cell specification and maintenance (335, 336). Most importantly, Gonsalvez *et al*. showed that absence of PRMT5 in flies results in the loss of arginine methylation on Sm as well as other proteins, producing a phenotype resembling that of TUDOR mutants, which suggests that methyl-dependent interactions are crucial for this phenomenon. Lastly, DDX3 is thought to play a role in the early phase of spermatogenesis (see below), thus it will be interesting in future studies to look for a role of TDRD3 in germ cell biology.

TDRD3, stress granules, and human disease?

TDRD3 can be linked with human genetic diseases. Firstly, DDX3 has a Y-linked homologue (>94% homology) that is also known as DBY. The peptides identified in our proteomic analysis do not allow us to determine if the protein we have isolated was DDX3X or DDX3Y (data not shown). This gene is located in the azospermia factor a (AZFa) region on the human Y chomosome (Yq11.21). Deletion of this Y interval is known to be a major cause for the occurrence of a severe testicular pathology associated with a complete germ cell loss, the Sertoli-cell-only (SCO) syndrome, and specific loss of DBY expression has been documented and correlated with this human pathology (337-339). Secondly, TDRD3 was found to interact with FMRP, the gene associated with Fragile X Mental Retardation syndrome (Utz Fischer, personal communication). Thirdly, TDRD3 was identified as one of the top hits in a screen for genes strongly correlated with poor post-operative prognosis of estrogen receptor-negative breast cancers (167). Interestingly, several of the proteins identified in our interaction screen have links with cancer. For example, EWS and FUS are both involved in recurrent chromosome translocations that result in novel fusion oncoproteins associated with many forms of sarcomas (340). These resulting fusion proteins have aberrant transcriptional function compared to their wild-type counterparts and thereby result in a variety of altered cellular properties that contribute to the tumorigenic process. A role for TDRD3 in transcription was never postulated, but a significant proportion (~ 30%) of total TDRD3 can be detected in the nucleus (Fig. 2). Thus, it would be interesting to investigate the functional relevance of TDRD3 interaction with common sarcoma-associated EWS and FUS oncoprotein fusions. EEF1A1 was found to be overexpressed in rat mammary

adenocarcinomas and its expression level correlates with metastasis ((341) and reference therein). Most strikingly, an oncogenic role for DDX3 was recently shown to be important for breast cancer biogenesis (342). Other components of SGs, including HuR, have also been linked with cancer in many ways (343). Finally, the Tudor domain of TDRD3 is highly conserved with the Tudor domain of SMN, the causative gene of spinal muscular atrophy (148). Whether TDRD3 participates in these pathways, and whether it does so in the context of SGs, remains unclear and will be the focus of future studies.

Materials and Methods

Cell culture and transfection

The human HeLa cervical carcinoma cell line was purchased from ATCC (Manassas, VA) and grown as a monolayer in DMEM medium supplemented with 1 mM sodium pyruvate, 50 IU/ml penicillin, 50 mg/ml streptomycin, and 10% fetal calf serum (Wisent, St-Bruno, QC, Canada). Cells were transfected with DNA plasmids using the Lipofectamine Plus reagent (Invitrogen, Burlington, ON, Canada) according to manufacturer's instructions.

DNA constructs

Total RNA was extracted from HeLa cells using Trizol reagent (Invitrogen) according to manufacturer's instructions. RNA concentration was measured with a spectrophotometer, and RNA quality was assessed by agarose gel electrophoresis. First strand cDNA synthesis was performed using 5 μg of total RNA and the avian myeloblastosis virus reverse transcriptase (Promega, Madison, WI) with an oligo-dT primer. cDNAs were amplified using oligonucleotides designed to introduce either full length or deletion mutants of TDRD3 (ΔDUF/OB, ΔN328, ΔUBA, ΔTdr, Δ(DUF/OB + UBA), Δ(DUF/OB + Tdr), DUF/OB, UBA, Central linker, and Tdr) in frame with a myc tag into the EcoRI and XhoI restriction sites of the pCMV-Myc expression vector (Clontech Laboratories Inc., Mountain View, CA). The TDRD3 Tudor (Tdr; a.a. 647-744) fragment was also inserted into the EcoRI/XhoI sites of the pGEX-4T2 vector (GE Healthcare, Piscataway, NJ). The TDRD3 Tudor E691K mutation was introduced using overlap extension

mutagenesis (344) with the following mutagenic oligonucleotides: 5'-ggaaactat**a**aagaggtgctactg-3' and 5'-cagtagcacctctt**t**atagtttcc-3', while using TDRD3-specific primers as flanking oligonucleotides. The mutated fragment was then inserted into the EcoRI and XhoI restriction sites of both the pCMV-Myc and the pGEX-4T2 vectors. The human SMN Tudor domain was amplified using the following oligonucleotides: 5'-ccggaattcgcagtggaaagttggggacaaa-3' and 5'-ccggaattcttaattagctacttcacagattgg-3', and inserted into the EcoRI restriction site of a modified pcDNA$_{3.1}$ vector (Invitrogen) containing the myc epitope tag sequence (345). Cloning of the wild type and E134K SMN Tudor domains into the pGEX-4T2 vector have been previously described (148). The Tudor domain of SPF30 was amplified from the SPF30 expression vector (148) by using PCR with the 5'-ttttgaattcttctactcaacctactcattcatgg-3' and 5'-ttttctcgagttactcctttgccttccttccttctt-3' oligonucleotides, and the amplified DNA fragment was inserted into the EcoRI and XhoI sites of the pcDNA$_{3.1}$-Myc expression vector. Human SERBP1 was amplified from a full length cDNA clone obtained from Open Biosystems (Huntsville, AL) by using PCR with the 5'-tttttggaattctatgcctgggcacttacaggaagg-3' and 5'-tttttctcgagattaagccagagctgggaatgcctctgg-3' oligonucleotides. The amplified DNA fragment was then cloned in frame with a myc tag into the pcDNA$_{3.1}$-Myc expression vector using the same restrictions sites as above. The expression vector encoding the GFP-fused endoribonuclease G3BP protein (291) was a generous gift from Dr. Nancy Kedersha (Harvard Medical School, Boston, MA). The pECFP-ER, -Golgi, and –Mito vectors were purchased from Clontech Laboratories Inc. (Mountain View, CA). All DNA

construct sequences were confirmed by automated DNA sequencing (StemCore Laboratories, Ottawa, ON, Canada).

Antibodies

A peptide corresponding to the last 22 C-terminal amino acids of TDRD3, ^{723}DGQPRRSTRPTQQFYQPPRARN^{744}, was synthesized at W.M. Keck Foundation Biotechnology Resource Laboratory (Yale University, New Heaven, CT). Polyclonal antibodies were generated by Cedarlane Laboratories (Burlington, ON, Canada) using rabbits injected with the synthetic peptide coupled to keyhole limpet hemocyanine (KLH). The polyclonal antibodies were affinity-purified over the antigenic peptide coupled to Affi-Gel 10 beads (Bio-Rad, Hercules, CA) following manufacturer's instructions, eluted in 100 mM Glycine pH 2.5, buffered with 1 M Tris-HCl pH 8.0, dialyzed against 1X phosphate-buffered saline (1X PBS: 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4, pH 7.4) at 4°C overnight, and concentrated using Centricon centrifugal devices (Millipore, Bedford, MA). The affinity-purified anti-TDRD3 antibody (0.25 mg/ml) was used at a 1:100 dilution for immunofluorescence (IF) and at a 1:1000 dilution for western blot (IB) analysis.

Monoclonal antibodies against ®-actin (Sigma, St-Louis, MO) and glyceraldehyde-3-phosphate dehydrogenase (Covance Canada Inc., Dorval, Qc) were both used at a 1:3000 dilution (IB). Goat polyclonal anti-TIA-1 antibody (C-20) and rabbit anti-ribosomal protein L28 (FL-137) antibody were purchased from Santa Cruz Biotechnology (Santa Cruz, CA) and used at 1:50 (IF) and 1:2500 (IB) dilutions respectively (IF). Mouse anti-

FMRP, clone 1C3 (ascites) was purchased from Millipore and used at 1:50 (IF) or 1:2500 (IB) dilutions. Human anti-GW182 (1:600 dilution, IF) and fluorophore-coupled anti-human (1:900 dilution, IF) antibodies were generous gifts from Dr. Ken Dimock (University of Ottawa). Rabbit anti-Hsp70 (Hsp72) polyclonal antibody was obtained from Stressgen (Ann Arbor, MI) and used at a 1:10,000 dilution (IB). Rabbit S6 ribosomal protein (5G10) antibody was purchased from Cell Signaling Technology (Danvers, MA) and used at a 1:5000 dilution (IB). The supernatant of a murine hybridoma producing an anti-myc monoclonal antibody (9E10), purchased from ATCC, was used at a 1:50 dilution (IF). Rabbit polyclonal anti-GST antibodies were described elsewhere (148) and used at a 1:5000 dilution (IB). Mouse monoclonal anti-Elongation Factor 1 alpha antibody was purchased from Upstate Biotechnology (Upstate, NY) and used at a 1:3000 dilution (IB). Rabbit polyclonal antibodies directed against the Ewing Sarcoma breakpoint region 1 protein (EWS; 1:40,000 dilution, IB; 1:900 dilution, IF), DEAD box polypeptide 3 (DDX3; 1:1000 dilution, IB; 1:300 dilution, IF), and FUS/TLS (1:40,000 dilution, IB; 1:900 dilution, IF) were all purchased from Bethyl Laboratories Inc. (Montgomery, TX). SYM10 and SYM11 antibodies were described elsewhere (105, 114) and were used at a 1:750 and 1:1000 dilution (IB), respectively. Fluorophore-coupled anti-goat, -mouse, and –rabbit secondary antibodies were obtained from Jackson ImmunoResearch (West Grove, PA) and used at a 1:100-200 dilution (IF). Horseradish peroxidase-conjugated goat anti-mouse and goat anti-rabbit secondary antibodies (MP Biomedicals, Solon, OH) were used at a 1:4000 dilution.

Peptide competition assay and immunoblotting

50 μg of HeLa total protein extracts were resolved on a 10% SDS-PAGE in triplicate, transferred to a PVDF membrane and prepared as three identical western blot strips. PVDF strips were blocked with 5% dry milk in PBST (1X PBS, 0.05% Tween 20) for 1 h at room temperature, and probed with either TDRD3 antibodies (0.25 μg/ml) or an antibody:peptide solution, for 2 h at room temperature. Preadsorbed antibody:peptide solutions were prepared by preincubation of TDRD3 antibodies with a 200-fold molar excess of the antigenic peptide (0.186 μg/ml) or an unrelated synthetic peptide (0.113 μg/ml) in immunoblotting solution (1% dry milk in PBST), for 16 h at 4°C. Incubation with primary antibodies was followed by three washes with PBST and probing with secondary antibody under the same conditions. Proteins were detected by chemiluminescence (Millipore) after three final washes in PBST.

Cell treatments and immunofluorescence

When prepared for indirect immunofluorescence microscopy, cells were grown directly onto glass coverslips. Transfected HeLa cells were treated 24 to 48 h post-transfection. Cellular stress response was induced either by treatment with 0.5 mM sodium arsenite for 30 min (oxidative stress; Ars), by incubation at 43°C for 30 min (Heat shock; Hs) or by treatment with 1 M sorbitol for 1 h followed by a 30-min recovery in isotonic culture media (osmotic shock; Osm) (295). To inhibit methylases, HeLa cells were treated for 20 h with 1 mM of the methyltransferase inhibitor 5'-deoxy-5'-methylthioadenosine (MTA; Sigma) prepared in dimethyl sulfoxide (DMSO). Control cells were mock-treated with equal volume of DMSO. Untreated and treated cells were fixed with 4%

paraformaldehyde for 10 min, permeabilized with 0.5% Triton X-100 in PBS for 5 min, and blocked with 0.1% BSA in PBS for 30 min prior to immunostaining. Cells were incubated with primary antibodies diluted in 0.1% BSA in PBS for 1 h. Cells were then washed once with 0.1% Triton X-100 in PBS, twice with 0.1% BSA in PBS, and incubated with the appropriate fluorophore-annexed secondary antibodies in 0.1% BSA in PBS, in the dark, for 1 h. All incubations were performed at room temperature. The cells were washed again as above, counterstained with DAPI, and mounted onto glass slides. Fluorescence was visualized with a Z.1 AxioImager upright microscope (Carl Zeiss Canada Ltd., Toronto, ON, Canada) and images were captured with an AXIOCAM HRM R 2.0 CCD digital camera.

Protein purification

GST-fusion proteins were overexpressed in *E. coli* BL-21 cells (Stratagene, La Jolla, CA) by induction with a final concentration of 0.1 mM isopropyl-D-thiogalactopyranoside (IPTG). Following induction, cells were spun down, resuspended in 10 ml of 1X PBS containing Complete™ protease inhibitor cocktail (Roche Applied Science) and subsequently broken down by sonication (5 pulses of 15 sec at 12 watts). Cell debris were discarded through centrifugation for 20 min at 10,000 X *g*. GST fusion proteins were then purified using glutathione-agarose (Sigma). For GST pull-down experiments, fusion proteins were kept on glutathione-agarose as a 50% slurry in 1X PBS. Alternatively, GST fusion proteins were eluted from the beads using 60 mM glutathione in 1X PBS adjusted to pH 7.5. The purified tagged proteins were dialyzed against 1 liter of 1X PBS at 4°C overnight. The dialysates were then concentrated using Amicon centrifugal devices

(Millipore). Protein concentration was measured by using the *DC* protein assay reagent (Bio-Rad).

Peptide binding assay

Peptide binding assays were performed as previously described (148). Briefly, $(RG)_4$ and $(R^{sDMA}G)_4$ peptides were biotinylated and pre-bound to streptavidin-agarose (Sigma) to generate peptide affinity columns. 0.35-1.4 ∫g of eluted GST-fusion Tudor domain proteins were diluted in 300 μl of lysis buffer (1% Triton X-100, 20 mM Tris pH 7.4, 150 mM NaCl, and Complete™ protease inhibitor cocktail) and mixed with 20 μl of a 50% slurry of the respective peptide affinity column. The mixtures were incubated at 4°C for 30 min with constant end-over-end mixing, and the beads were washed twice with lysis buffer and once with 1X PBS. Retained proteins were then resolved by SDS-PAGE and detected using a rabbit anti-GST serum.

Pull downs and mass spectrometry

HeLa cells from 5 x 150 mm plates were lysed in lysis buffer for 6 h at 4°C and then centrifuged to remove cellular debris. Cell lysates were incubated with either purified recombinant GST (as a control), GST-TDRD3-Tdr or GST-TDRD3-Tdr E691K proteins coupled to glutathione-agarose for 16 h at 4°C. The beads were then washed twice with lysis buffer and once with 1X PBS. The retained proteins were resolved by SDS-PAGE and the gel, stained with Coomassie Blue. The polypeptides specifically pulled-down by TDRD3-Tdr, that were not pulled-down by GST alone, were excised from the gel and sent out for mass spectrometry analysis and identification (WEMB Biochem Inc.,

Toronto, ON, Canada). Briefly, gel slices were incubated with trypsin and the resulting digested peptides ran through a matrix-assisted laser desorption ionization-quadrupole-time of flight (MALDI-Qq-TOF) mass spectrometer coupled to a LCQ DECA XP ion trap. The database search for peptide mass fingerprint and MS/MS sequencing data from MALDI-Qq-TOF was done using an in-house database and the Mascot software. Sequencing data generated from the ion trap was analyzed using the Sequest software and a human database subset created from NCBI nonredundant data (300).

Polysomal profile analysis and protein extraction from sucrose gradient fractions

5 x 10^6 HeLa cells were grown in 100-mm tissue culture dishes, left untreated or incubated with 0.5 mM sodium arsenite for 30 min, harvested, and proceeded immediately. Prior to analysis, the cells were resuspended in 1 ml of polysomal buffer (20 mM Tris, pH 7.5, 150 mM NaCl, 1.25 mM $MgCl_2$, 5 U/ml of RNAsine (Amersham Bioscience, Piscataway, NJ), EDTA-free protease inhibitor cocktail (Complete™, Roche), and 1 mM dithiothreitol). Nonidet P-40 was then added to a final concentration of 1% and cells were lysed for 15 min on ice. Extracts were clarified by centrifugation at 12,000 X *g* for 20 min at 4°C. Twenty optical density units (260 nm) of cytoplasmic extracts were loaded on each 10-60% linear sucrose gradient and further analyzed as described elsewhere (297). Briefly, following ultracentrifugation (40,000 X *g* for 4 h) of the sucrose gradients, 0.5-ml fractions were collected using the gradient density fractionator system (Teledyne Isco). mRNPs of each collected fraction were precipitated overnight at –20°C by addition of 1 ml of cold ethanol and centrifuged for 30 min at 14,000 X *g*. Proteins were then resuspended in 150 ml Laemmli sample buffer.

Funding

I. G. is supported by a scholarship from Fonds de la recherche en santé du Québec. J.C. is the recipient of a Canada Research Chair (tier 2) in RNA metabolism and is supported by operating grants # MOP-68943/86746 from CIHR and a grant from Families of Spinal Muscular Atrophy. Funding to pay the Open Access charge was provided by operating grant # MOP-86746 to J.C. from the Canadian Institutes of Health Research.

Acknowledgements

We are grateful to Dr. Edouard W. Khandjian and Dr. Utz Fischer for their helpful comments in the early phases of this work. We also wish to thank Dr. Nancy Kedersha for providing us with useful advice for successful immunofluorescence stainings of stress granules. We would like to thank researchers acknowledged in Material & Methods for providing reagents, as well as Helina Tadesse for critically reading the manuscript. Last but not least, we would like to thank Moaweya Zayed for his help in cloning the Δ(DUF/OB + UBA) and the Δ(DUF/OB + Tdr) constructs.

CHAPTER 5

General discussion

5.1. The PRMT1 alternative isoform v2 possesses distinct substrate specificity, intracellular localization, and is predominantly expressed in breast cancer

5.1.1. Unique N-terminal sequences affect PRMT1 activity and substrate specificity

The ever-expanding repertoire of arginine-methylated proteins (114, 118) indicates that arginine methylation is an important post-translational modification involved in the regulation of several cellular processes, including DNA repair, transcription, signal transduction, protein localization, and RNA processing (12, 13, 96, 214). PRMT1, the first identified PRMT, is responsible for most of the arginine methylation reactions in human cells (33). Because PRMT1 has a multitude of substrates and plays significant roles in a number of distinct cellular processes, its enzymatic activity must be tightly regulated. PRMT1 activity can be modulated in multiple ways: by the formation of homo-oligomeric complexes (25), by post-translational modifications (40), and by protein-protein interactions (14, 54-61, 63-65). The findings presented in Chapter 2 add a fourth mechanism to this list, suggesting that PRMT1 activity can also be modulated at the co-transcriptional level. This finding is supported by the evidence of tissue-specific expression of up to seven PRMT1 isoforms (v1 to v7) resulting from alternative mRNA splicing of the primary PRMT1 transcript in human cells. Furthermore, we have shown that each PRMT1 splicing variant exhibits a unique N-terminal hydrophobic region that confers distinct substrate specificity (117). Thus, N-terminal unique sequences can

influence PRMT1 protein-protein interactions, a finding that could be extrapolated to other PRMTs, as N-terminal domains are generally the most divergent between PRMTs (16, 42). This finding not only adds additional complexity to the way in which PRMT1 function is regulated, but could also have important implications for PRMT1's downstream targets (see 5.1.3.).

5.1.2. The PRMT1 alternative isoform v2 contains a nuclear export signal

Endogenous PRMT1 was first reported to reside predominantly in the nucleus (16), even though it had been identified through purification from a cytoplasmic macromolecular complex (14). As an N-terminal GFP fusion protein, PRMT1 also localized predominantly to the nucleus (18), whereas N-terminal myc-tagged PRMT1 localized mostly in the cytoplasm, with some cells showing both cytoplasmic and nuclear localization (169). The explanation for the variation in the subcellular localization of PRMT1 became obvious when we uncovered a functional CRM1-dependent nuclear export signal encoded by exon 2 of the primary PRMT1 transcript, an exon only retained in PRMT1 isoform v2 (Chapter 2). As a result, PRMT1v2 mainly localizes to the cytoplasm, whereas the other PRMT1 isoforms predominantly localize to the nucleus of the cell. Thus, the subcellular localization of PRMT1 varies depending on which splicing variant is used and/or detected. This finding suggests that by regulating the alternative splicing of PRMT1 exon 2, cells can modulate the nuclear *versus* cytoplasmic pools of PRMT1 isoforms that possess distinct protein-protein interaction profiles and substrate specificity.

More recently, Herrmann and colleagues found that despite the presence of the nuclear export signal, PRMT1v2 could be predominantly nuclear in certain cell types (45), suggesting a cell type-specific regulation of the nucleo-cytoplasmic shuttling of PRMT1v2. This group also showed that both enzymatic activity and appropriate methylation status of substrate proteins are required for the intracellular shuttling of PRMT1 (44, 346). Additional work will be necessary to dissect the precise mechanisms by which PRMT1v2 shuttles in and out of the nucleus, as this might shed further light on the regulation of PRMT1 substrate methylation and the pathophysiological consequences ensuing from its misregulation (45) (see below).

5.1.3. Arginine methylation and cancer

A number of studies have provided evidence of functional links between misregulation of PRMTs and cancer (49, 102, 347-350). The fact that PRMT1 is a known co-activator of the estrogen receptor makes it an attractive candidate to be overexpressed in breast cancer. We observed that overall PRMT1 mRNA levels are on average 14-fold higher in a panel of breast cancer cell lines and in a human breast tumor (Chapter 2). Moreover, the relative balance of alternatively spliced PRMT1 isoforms is altered in these cell lines when compared with immortalized normal breast cells. Among the major isoforms, PRMT1v2 mRNA, which encodes for an enzyme that mostly localizes to the cytoplasm, is selectively increased by 3.5 fold relative to v1 (Chapter 2). An increase in the relative abundance of PRMT1v2 (but not of PRMT1v1) is also clearly observed at the protein level in the breast cancer cell lines (Chapter 2). The functional significance of PRMT1 (and particularly PRMT1v2) overexpression in breast cancer is still unknown, but these

observations suggest a possible relation between altered expression pattern of PRMT1 and breast cancer cell progression.

For example, overexpressed cytoplasmic PRMT1 isoform could be responsible for the prolonged methylation of the cytoplasmic estrogen receptor α observed in a subset of breast cancers (Chapter 1), which might in turn lead to sustained activation of estrogen signaling pathways and ultimately, to mammary tumorigenesis (89, 95). On the other hand, PRMT1v2-dependent methylation of the ligand-dependent corepressor of nuclear receptors RIP140 protein may facilitate RIP140 nuclear export by mediating its interaction with the nuclear exporter CRM1 (101) (Chapter 1). Cytoplasmic sequestration of RIP140, away from the nucleus where it functions as a tumor suppressor (101, 351), could provide PRMT1 with another means to contribute to breast cancer pathogenesis. Future work in our laboratory will be aimed at identifying the breast cancer-specific molecular pathways in which PRMT1 plays a role.

Aberrant alternative splicing patterns are common in cancers and cancer-specific transcript variants may represent promising biomarkers and targets for diagnostic, prognostic, and treatment purposes (352). PRMT1v2 expression has recently been shown to be associated with nodal status and tumor grade in colon cancer patients, and therefore, could potentially be used as a marker of unfavorable prognosis (353, 354). In future studies, it will be important to determine whether or not PRMT1v2 could be a promising biomarker of unfavorable breast cancer prognosis as well.

Finally, we have demonstrated that globally the levels of arginine-methylated proteins are increased in all cancer cell lines tested (Chapter 2). Although this profile does not necessarily correlate with cell line-specific overexpression of PRMT1v2, other type I PRMTs could also contribute to this increase in arginine methylation. Indeed, CARM1 and PRMT6 have been shown to be overexpressed in certain cancers (205, 355, 356), and we found these PRMTs to be also expressed at higher levels in our panel of cancer cell lines (Chapter 2). A recent study confirmed significantly higher expression of PRMT1 and PRMT6, with a corresponding increase in aDMA levels, in cancer cells of various tissues compared to non-neoplastic cells (356). Therefore, an altered PRMT1 (and other PRMTs) expression profile appears to correlate with a differential pattern of arginine methylation in cancer cell lines, suggesting that misregulation of arginine methylation could contribute to carcinogenesis.

5.2. TDRD3: a novel Tudor-containing protein in search of a function

5.2.1. Transcriptional regulation

Our characterization of TDRD3 showed that it localizes to both nuclear (40%) and cytoplasmic (60%) compartments of the cells (Chapter 4). In the nucleus, we found that TDRD3 accumulates in Promyelocytic leukemia (PML) nuclear bodies possibly associated with sites of up-regulated mRNA transcription (Chapter 3). This observation is in agreement with recent findings showing that TDRD3 localizes to transcriptional start sites inside promoter regions of transcriptionally active genes that are enriched for PRMT1 and PRMT4/CARM1 (247). These two PRMTs function synergistically at specific promoter regions to generate ‘methyl marks’ on the core histone tails that are

usually associated with transcriptional activation of gene expression (83, 84). PRMT1 and PRMT4/CARM1 are also co-regulators of a large number of transcription factors (75-81). Moreover, it has recently been demonstrated that the already extensively post-translationally modified C-terminal domain (CTD) of the RNA polymerase II is also methylated at a single arginine residue (R1810) by PRMT4/CARM1, which results in facilitated expression of select small nuclear and nucleolar RNAs (357). Such 'hypermethylation' in the neighborhood of active gene promoters may in turn function as a 'docking' motif to recruit methylarginine-binding (e.g. Tudor domain-containing) proteins that facilitate transcription. The Tudor domain of TDRD3, for example, binds strongly to PRMT1- and CARM1-generated 'methyl marks' associated with transcriptional activation (247, 357). This may in turn favor selective recruitment of TDRD3 to distinct genes where it may function as a transcriptional co-activator (247, 357). In addition, TDRD3 recruitment to sites of transcriptional activation could be further stabilized by the interaction of its UBA domain with ubiquitinated proteins (e.g. histones, pol II CTD, transcription factors), some of which may congregate in PML bodies (247, 358).

5.2.2. *Post-transcriptional regulation*

Post-transcriptional functions have also been suggested for TDRD3 because an exon junction complex (EJC)-binding motif (EBM) in the extreme C-terminus of TDRD3 has been found to interact with two of the four core EJC proteins: Y14 and Magoh, and the peripheral interacting factor PABP1, suggesting a role for this motif in the recruitment of TDRD3 to EJCs and mRNAs (166). The EJC is deposited onto mRNA during splicing,

on which it remains firmly bound during nuclear export into the cytoplasm, where it is thought to play a major role in post-transcriptional regulation of the spliced mRNA (e.g. localization, translation, and surveillance) (270, 271). The functional significance of the interaction between TDRD3 and the EJC, however, remains to be investigated.

Our studies along with work by others demonstrate that in the cytoplasm, TDRD3 interacts and co-sediments with the translational regulator Fragile X Mental Retardation Protein (FMRP) on actively translating polyribosomes (219, 254) (Chapter 4). When exposed to life-threatening stresses (e.g. oxidative stress, heat and osmotic shocks, viral infection, hypoxia, etc.), cells reconfigure their proteome to concentrate on molecular damage repair activities. Under such conditions, we and others have shown that TDRD3 (219, 254) and FMRP (290) move into non-translating cytoplasmic stress granules (SGs) where they may function as translational repressors of key transcripts (254). SGs represent sites of mRNA sorting/triage that play a major role in the post-transcriptional regulation of gene expression by regulating the localization, stability, and translation of many mRNAs (283, 286).

Another important finding from our work is that the Tudor domain of TDRD3 is both necessary and sufficient for its recruitment to SGs (Chapter 4). In addition, the introduction of a mutation that disrupts the recognition of methylated substrates demonstrated that the 'methyl-binding' surface in the Tudor domain of TDRD3 is important for this process. Our findings therefore suggest that TDRD3 may be recruited to SGs by engaging in protein-protein interactions regulated by arginine methylation. The

fact that methylation status of both FMRP (305) and the cold-inducible RNA-binding protein (CIRP) (359) has been shown to impact positively on SG assembly (360) and that general methylation inhibitors impinge upon SG dynamics (Chapter 4) suggest that protein methylation may play a more important role in SG function than previously expected. Future studies in our laboratory will focus on elucidating the mechanisms regulating the recruitment of TDRD3 to SGs as well as its contribution to SG formation/dissociation and dynamics.

In agreement with a putative role for arginine methylation in regulated recruitment of TDRD3 to SGs, our pull-down experiments identified several proteins interacting with the Tudor domain of TDRD3, many of which are potentially methylated RNA binding proteins (Chapter 4). Furthermore, amongst these proteins, we and others have confirmed the relocalization of DDX3 (361), EWS (362), FUS (362), and SERBP1 (363) to SGs upon various stress stimuli (Chapter 4). These proteins are therefore good candidates to promote the recruitment of TDRD3 to SGs (or vice-versa) and confirmation of the observed interactions in the context of the full-length protein will be crucial to future studies addressing their biological significance.

5.2.3. TDRD3 as a scaffold protein in the mRNP lifecycle?

Gene expression processes, from transcription to translation and degradation, depend on the ordered and precise assembly of messenger ribonucleoprotein (mRNP) particles (214, 267). The mRNP lifecycle begins during transcription as the nascent pre-mRNA is coated with a selection of RNA-binding proteins (267, 326). Its protein composition then

fluctuates as it mediates post-transcriptional RNA processing events (capping, splicing, polyadenylation, surveillance) to generate a mature, nuclear export-, and translational-competent mRNP (214). Because of the diverse range of cellular processes they govern, the organizational structure of mRNPs requires strict spatial and temporal coordination in order for this 'gene expression system' (267) to convey the proper message. This is achieved mainly through regulation of protein-nucleic acid and protein-protein interactions in which post-translational modifications, such as arginine methylation, have been shown to modulate various facets (214, 360).

Proteins that contain multiple interaction domains are thus particularly important in the scaffolding of these complexes by providing a platform for bringing together various protein partners that control these cellular processes (364). Growing evidence suggest that TDRD3 may serve as a scaffold molecule in mRNA metabolism. First, TDRD3 is composed of four distinct protein domains or motifs: a DUF/OB-fold, a UBA domain, a Tudor domain, and an EBM motif (Chapter 1) providing known and predicted binding surfaces to assemble nucleic acids, ubiquitinated and arginine-methylated proteins, as well as RNA-binding proteins (166, 219, 247, 254). Second, our studies along with work by others suggest that TDRD3 may be an integral component of mRNPs, from mRNA biogenesis (Chapter 3 and (247)) to translation (Chapter 4 and (254)) and degradation (through interaction with the exon junction complex) (166). Collectively, these findings support our hypothesis that TDRD3 may play a central role in synchronizing transcriptional with post-transcriptional processes in part regulated by arginine methylation.

5.2.4. TDRD3 and cancer

All the RNA binding proteins that we identified as putative interactors of the Tudor domain of TDRD3 (Chapter 4) have been linked to cancer. For example, EWS and FUS are both involved in recurrent chromosome translocations that result in novel fusion oncoproteins associated with many forms of sarcomas (340). SERBP1 is overexpressed in tumor epithelial cells from ovarian cancer, which correlates with disease progression (365). EEF1A1 is overexpressed in rat mammary adenocarcinomas and its expression level correlates with metastasis (341). Finally, DDX3 activation has been shown to promote growth, proliferation, and neoplastic transformation of breast epithelial cells (342). These findings strongly suggest a role for these proteins in carcinogenesis and it is intriguing that all of them (except EEF1A) relocalize to SGs upon various stress stimuli. Even more interestingly, TDRD3 has also been identified amongst genes whose overexpression has a strong predictive value for poor prognosis of ER-negative breast cancers (167). This suggests that TDRD3 may have an important role to play in progression of cancer, possibly by regulating the localization/activity of the above-mentioned oncoproteins.

Recent findings have demonstrated the physiological importance of mammalian SG formation in cancer cells resistance to chemotherapeutic drugs (366), a phenomenon often associated with late-stage cancers (367). Upon exposure to oxidative stress (arsenite) (368), radiation (369), hypoxia (370), and chemotherapeutic agents (e.g. proteasome inhibitors (371) and bortezomib (366)), formation of SGs has been shown to promote tumor cells resistance to apoptosis, which in turn contributes to cancer cells and

solid tumors resistance to therapy. In some cases, this effect is attributed to the sequestration of signaling proteins into SGs to promote cell survival in response to stress (370, 372). In other cases, SGs are involved in the regulation of the expression of key anti-apoptotic factors (371). Finally, clinical relevance of SG contribution to aggressive tumor growth and poor tumor response to conventional therapies is underscored by the observation of SGs in hypoxic cells, the most difficult tumor cells to kill, located in the core of breast tumors (369, 373). Whether TDRD3 participates in these resistance mechanisms, and whether it does so in the context of SGs, remains unclear and will be the focus of future studies.

5.3. Conclusion and outlook

Despite significant advances in the field of arginine methylation in the last two decades, much of the molecular mechanisms underlying how arginine methylation is regulated and how it influences different steps in the control of gene expression remains to be clarified. The work presented in this thesis, and that of numerous others, has provided a better understanding of the mechanisms involved in the modulation of PRMT1 enzymatic activity and specificity towards its substrates. It has also provided significant insight into how arginine methylation influences protein-protein interactions with the novel methyl-binding effector protein TDRD3 and their putative impacts on mRNP dynamics in transcriptional and post-transcriptional RNA processing events.

Finally, continued efforts aimed at understanding the etiology and improving breast cancer diagnosis are essential in answering the need for improved prevention and

treatment. Thus, continuing to study the significance of PRMT1 and TDRD3 upregulation in breast cancer and how this correlates with disease progression and poor prognosis is of crucial importance. Future studies on PRMT1 and TDRD3 will likely uncover additional novel pathways and molecules involved in breast cancer pathogenesis, which in turn may represent new targets for the formulation of novel detection and intervention strategies.

REFERENCES

1. Sherwood, L. (2008) The energy stored within ATP is used for synthesis, transport, and mechanical work. In *Human physiology: from cells to systems.* 7th ed. Cengage Learning, p. 928.

2. Newsholme, E. and Leech, A. (2009) Amino acid and protein metabolism - Protein turnover. In *Functional Biochemistry in Health and Disease*. John Wiley & Sons, p. 560.

3. Deribe, Y.L., Pawson, T. and Dikic, I. (2010) Post-translational modifications in signal integration. *Nat Struct Mol Biol,* 17, 666-72.

4. Segre, C.V. and Chiocca, S. (2011) Regulating the regulators: the post-translational code of class I HDAC1 and HDAC2. *Journal of biomedicine & biotechnology,* 2011, 690848.

5. Eyers, C.E. and Gaskell, S.J. (2008) Mass spectrometry to identify posttranslational modifications. In *Wiley Encyclopedia of Chemical Biology*. John Wiley and Sons.

6. Seo, J. and Lee, K.J. (2004) Post-translational modifications and their biological functions: proteomic analysis and systematic approaches. *Journal of biochemistry and molecular biology,* 37, 35-44.

7. Lee, D.Y., Teyssier, C., Strahl, B.D. and Stallcup, M.R. (2005) Role of protein methylation in regulation of transcription. *Endocr Rev,* 26, 147-70.

8. Paik, W.K. and Kim, S. (1980) Natural occurence of various methylated amino acid derivatives. *In Protein Methylation, A. Meister, ed. (New York: John Wiley & Sons)*, 8-25.

9. Najbauer, J., Johnson, B.A., Young, A.L. and Aswad, D.W. (1993) Peptides with sequences similar to glycine, arginine-rich motifs in proteins interacting with RNA are efficiently recognized by methyltransferase(s) modifying arginine in numerous proteins. *J Biol Chem,* 268, 10501-9.

10. Pahlich, S., Zakaryan, R.P. and Gehring, H. (2006) Protein arginine methylation: Cellular functions and methods of analysis. *Biochim Biophys Acta*.

11. Gary, J.D. and Clarke, S. (1998) RNA and protein interactions modulated by protein arginine methylation. *Prog Nucleic Acid Res Mol Biol,* 61, 65-131.

12. Boisvert, F.M., Chenard, C.A. and Richard, S. (2005) Protein interfaces in signaling regulated by arginine methylation. *Science STKE,* 2005, re2.

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

14. Lin, W.J., Gary, J.D., Yang, M.C., Clarke, S. and Herschman, H.R. (1996) The mammalian immediate-early TIS21 protein and the leukemia-associated BTG1 protein interact with a protein-arginine N-methyltransferase. *J Biol Chem,* 271, 15034-44.

15. Lakowski, T.M. and Frankel, A. (2009) Kinetic analysis of human protein arginine N-methyltransferase 2: formation of monomethyl- and asymmetric dimethyl-arginine residues on histone H4. *Biochem J,* 421, 253-61.

16. Tang, J., Gary, J.D., Clarke, S. and Herschman, H.R. (1998) PRMT 3, a type I protein arginine N-methyltransferase that differs from PRMT1 in its oligomerization, subcellular localization, substrate specificity, and regulation. *J Biol Chem,* 273, 16935-45.

17. Chen, D., Ma, H., Hong, H., Koh, S.S., Huang, S.M., Schurter, B.T., Aswad, D.W. and Stallcup, M.R. (1999) Regulation of transcription by a protein methyltransferase. *Science,* 284, 2174-7.

18. Frankel, A., Yadav, N., Lee, J., Branscombe, T.L., Clarke, S. and Bedford, M.T. (2002) The novel human protein arginine N-methyltransferase PRMT6 is a nuclear enzyme displaying unique substrate specificity. *J Biol Chem,* 277, 3537-43.

19. Lee, J., Sayegh, J., Daniel, J., Clarke, S. and Bedford, M.T. (2005) PRMT8, a new membrane-bound tissue-specific member of the protein arginine methyltransferase family. *J Biol Chem*.

20. Sayegh, J., Webb, K., Cheng, D., Bedford, M.T. and Clarke, S.G. (2007) Regulation of protein arginine methyltransferase 8 (PRMT8) activity by its N-terminal domain. *J Biol Chem,* 282, 36444-53.

21. Pollack, B.P., Kotenko, S.V., He, W., Izotova, L.S., Barnoski, B.L. and Pestka, S. (1999) The human homologue of the yeast proteins Skb1 and Hsl7p interacts with Jak kinases and contains protein methyltransferase activity. *J Biol Chem,* 274, 31531-42.

22. Rho, J., Choi, S., Seong, Y.R., Cho, W.K., Kim, S.H. and Im, D.S. (2001) Prmt5, which forms distinct homo-oligomers, is a member of the protein-arginine methyltransferase family. *J Biol Chem,* 276, 11393-401.

23. 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, 32971-6.

24. Cook, J.R., Lee, J.H., Yang, Z.H., Krause, C.D., Herth, N., Hoffmann, R. and Pestka, S. (2006) FBXO11/PRMT9, a new protein arginine methyltransferase, symmetrically dimethylates arginine residues. *Biochem Biophys Res Commun,* 342, 472-81.

25. Zhang, X. and Cheng, X. (2003) Structure of the predominant protein arginine methyltransferase PRMT1 and analysis of its binding to substrate peptides. *Structure,* 11, 509-20.

26. Zhang, X., Zhou, L. and Cheng, X. (2000) Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *The EMBO journal,* 19, 3509-19.

27. Kagan, R.M. and Clarke, S. (1994) Widespread occurrence of three sequence motifs in diverse S-adenosylmethionine-dependent methyltransferases suggests a common structure for these enzymes. *Archives of biochemistry and biophysics,* 310, 417-27.

28. Katz, J.E., Dlakic, M. and Clarke, S. (2003) Automated identification of putative methyltransferases from genomic open reading frames. *Mol Cell Proteomics,* 2, 525-40.

29. Cheng, X., Collins, R.E. and Zhang, X. (2005) Structural and sequence motifs of protein (histone) methylation enzymes. *Annu Rev Biophys Biomol Struct,* 34, 267-94.

30. Krause, C.D., Yang, Z.H., Kim, Y.S., Lee, J.H., Cook, J.R. and Pestka, S. (2007) Protein arginine methyltransferases: Evolution and assessment of their pharmacological and therapeutic potential. *Pharmacology & therapeutics,* 113, 50-87.

31. Nishioka, K. and Reinberg, D. (2003) Methods and tips for the purification of human histone methyltransferases. *Methods (San Diego, Calif,* 31, 49-58.

32. Fielenbach, N., Guardavaccaro, D., Neubert, K., Chan, T., Li, D., Feng, Q., Hutter, H., Pagano, M. and Antebi, A. (2007) DRE-1: an evolutionarily conserved F box protein that regulates C. elegans developmental age. *Developmental cell,* 12, 443-55.

33. 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, 7723-30.

34. Scorilas, A., Black, M.H., Talieri, M. and Diamandis, E.P. (2000) Genomic organization, physical mapping, and expression analysis of the human protein arginine methyltransferase 1 gene. *Biochem Biophys Res Commun,* 278, 349-59.

35. Scott, H.S., Antonarakis, S.E., Lalioti, M.D., Rossier, C., Silver, P.A. and Henry, M.F. (1998) Identification and characterization of two putative human arginine methyltransferases (HRMT1L1 and HRMT1L2). *Genomics,* 48, 330-40.

36. Wooderchak, W.L., Zang, T., Zhou, Z.S., Acuna, M., Tahara, S.M. and Hevel, J.M. (2008) Substrate profiling of PRMT1 reveals amino acid sequences that extend beyond the "RGG" paradigm. *Biochemistry,* 47, 9456-66.

37. Wooderchak, W.L. (2009) Characterization of the substrate specificity and mechanism of protein arginine methyltransferase 1. *Biochemistry.* Utah State University, Logan, Vol. Doctor of Philosphy, p. 218.

38. Lee, D.Y., Ianculescu, I., Purcell, D., Zhang, X., Cheng, X. and Stallcup, M.R. (2007) Surface-scanning mutational analysis of protein arginine methyltransferase 1: roles of specific amino acids in methyltransferase substrate specificity, oligomerization, and coactivator function. *Molecular endocrinology (Baltimore, Md,* 21, 1381-93.

39. Osborne, T., Roska, R.L., Rajski, S.R. and Thompson, P.R. (2008) In situ generation of a bisubstrate analogue for protein arginine methyltransferase 1. *Journal of the American Chemical Society,* 130, 4574-5.

40. Kuhn, P. and Xu, W. (2009) Chapter 9 Protein Arginine Methyltransferases Nuclear Receptor Coregulators and Beyond. *Progress in molecular biology and translational science,* 87C, 299-342.

41. Osborne, T.C., Obianyo, O., Zhang, X., Cheng, X. and Thompson, P.R. (2007) Protein arginine methyltransferase 1: positively charged residues in substrate peptides distal to the site of methylation are important for substrate binding and catalysis. *Biochemistry,* 46, 13370-81.

42. Frankel, A. and Clarke, S. (2000) PRMT3 is a distinct member of the protein arginine N-methyltransferase family. Conferral of substrate specificity by a zinc-finger domain. *J Biol Chem,* 275, 32974-82.

43. Yue, W.W., Hassler, M., Roe, S.M., Thompson-Vale, V. and Pearl, L.H. (2007) Insights into histone code syntax from structural and biochemical studies of CARM1 methyltransferase. *The EMBO journal,* 26, 4402-12.

44. Herrmann, F., Lee, J., Bedford, M.T. and Fackelmayer, F.O. (2005) Dynamics of human protein arginine methyltransferase 1(PRMT1) in vivo. *J Biol Chem,* 280, 38005-10.

45. Herrmann, F., Pably, P., Eckerich, C., Bedford, M.T. and Fackelmayer, F.O. (2009) Human protein arginine methyltransferases in vivo--distinct properties of eight canonical members of the PRMT family. *J Cell Sci,* 122, 667-77.

46. Xu, W., Chen, H., Du, K., Asahara, H., Tini, M., Emerson, B.M., Montminy, M. and Evans, R.M. (2001) A transcriptional switch mediated by cofactor methylation. *Science,* 294, 2507-11.

47. Feng, Q., He, B., Jung, S.Y., Song, Y., Qin, J., Tsai, S.Y., Tsai, M.J. and O'Malley, B.W. (2009) Biochemical control of CARM1 enzymatic activity by phosphorylation. *J Biol Chem,* 284, 36167-74.

48. Higashimoto, K., Kuhn, P., Desai, D., Cheng, X. and Xu, W. (2007) Phosphorylation-mediated inactivation of coactivator-associated arginine methyltransferase 1. *Proc Natl Acad Sci U S A,* 104, 12318-23.

49. Pal, S., Vishwanath, S.N., Erdjument-Bromage, H., Tempst, P. and Sif, S. (2004) Human SWI/SNF-associated PRMT5 methylates histone H3 arginine 8 and negatively regulates expression of ST7 and NM23 tumor suppressor genes. *Mol Cell Biol,* 24, 9630-45.

50. Guccione, E., Bassi, C., Casadio, F., Martinato, F., Cesaroni, M., Schuchlautz, H., Luscher, B. and Amati, B. (2007) Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. *Nature,* 449, 933-7.

51. Iberg, A.N., Espejo, A., Cheng, D., Kim, D., Michaud-Levesque, J., Richard, S. and Bedford, M.T. (2008) Arginine methylation of the histone H3 tail impedes effector binding. *J Biol Chem,* 283, 3006-10.

52. Hyllus, D., Stein, C., Schnabel, K., Schiltz, E., Imhof, A., Dou, Y., Hsieh, J. and Bauer, U.M. (2007) PRMT6-mediated methylation of R2 in histone H3 antagonizes H3 K4 trimethylation. *Genes Dev,* 21, 3369-80.

53. Rathert, P., Dhayalan, A., Murakami, M., Zhang, X., Tamas, R., Jurkowska, R., Komatsu, Y., Shinkai, Y., Cheng, X. and Jeltsch, A. (2008) Protein lysine methyltransferase G9a acts on non-histone targets. *Nature chemical biology,* 4, 344-6.

54. Rouault, J.P., Falette, N., Guehenneux, F., Guillot, C., Rimokh, R., Wang, Q., Berthet, C., Moyret-Lalle, C., Savatier, P., Pain, B. *et al.* (1996) Identification of BTG2, an antiproliferative p53-dependent component of the DNA damage cellular response pathway. *Nature genetics,* 14, 482-6.

55. Lim, I.K., Park, T.J., Kim, S., Lee, H.W. and Paik, W.K. (1998) Enzymatic methylation of recombinant TIS21 protein-arginine residues. *Biochem Mol Biol Int,* 45, 871-8.

56. Berthet, C., Guehenneux, F., Revol, V., Samarut, C., Lukaszewicz, A., Dehay, C., Dumontet, C., Magaud, J.P. and Rouault, J.P. (2002) Interaction of PRMT1 with BTG/TOB proteins in cell signalling: molecular analysis and functional aspects. *Genes Cells,* 7, 29-39.

57. Bakker, W.J., Blazquez-Domingo, M., Kolbus, A., Besooyen, J., Steinlein, P., Beug, H., Coffer, P.J., Lowenberg, B., von Lindern, M. and van Dijk, T.B. (2004) FoxO3a regulates erythroid differentiation and induces BTG1, an activator of protein arginine methyl transferase 1. *J Cell Biol,* 164, 175-84.

58. Duriez, C., Moyret-Lalle, C., Falette, N., El-Ghissassi, F. and Puisieux, A. (2004) BTG2, its family and its tutor. *Bull Cancer,* 91, E242-53.

59. Tirone, F. (2001) The gene PC3(TIS21/BTG2), prototype member of the PC3/BTG/TOB family: regulator in control of cell growth, differentiation, and DNA repair? *J Cell Physiol,* 187, 155-65.

60. Boiko, A.D., Porteous, S., Razorenova, O.V., Krivokrysenko, V.I., Williams, B.R. and Gudkov, A.V. (2006) A systematic search for downstream mediators of tumor suppressor function of p53 reveals a major role of BTG2 in suppression of Ras-induced transformation. *Genes Dev,* 20, 236-52.

61. Lim, I.K. (2006) TIS21 (/BTG2/PC3) as a link between ageing and cancer: cell cycle regulator and endogenous cell death molecule. *Journal of cancer research and clinical oncology,* 132, 417-26.

62. Miyata, S., Mori, Y. and Tohyama, M. (2008) PRMT1 and Btg2 regulates neurite outgrowth of Neuro2a cells. *Neuroscience letters,* 445, 162-5.

63. Robin-Lespinasse, Y., Sentis, S., Kolytcheff, C., Rostan, M.C., Corbo, L. and Le Romancer, M. (2007) hCAF1, a new regulator of PRMT1-dependent arginine methylation. *J Cell Sci*.

64. Rouault, J.P., Prevot, D., Berthet, C., Birot, A.M., Billaud, M., Magaud, J.P. and Corbo, L. (1998) Interaction of BTG1 and p53-regulated BTG2 gene products with mCaf1, the murine homolog of a component of the yeast CCR4 transcriptional regulatory complex. *J Biol Chem,* 273, 22563-9.

65. Lei, N.Z., Zhang, X.Y., Chen, H.Z., Wang, Y., Zhan, Y.Y., Zheng, Z.H., Shen, Y.M. and Wu, Q. (2009) A feedback regulatory loop between methyltransferase PRMT1 and orphan receptor TR3. *Nucleic acids research,* 37, 832-48.

66. Duong, F.H., Christen, V., Berke, J.M., Penna, S.H., Moradpour, D. and Heim, M.H. (2005) Upregulation of protein phosphatase 2Ac by hepatitis C virus modulates NS3 helicase activity through inhibition of protein arginine methyltransferase 1. *J Virol,* 79, 15342-50.

67. Shen, E.C., Henry, M.F., Weiss, V.H., Valentini, S.R., Silver, P.A. and Lee, M.S. (1998) Arginine methylation facilitates the nuclear export of hnRNP proteins. *Genes Dev,* 12, 679-91.

68. Yu, M.C., Lamming, D.W., Eskin, J.A., Sinclair, D.A. and Silver, P.A. (2006) The role of protein arginine methylation in the formation of silent chromatin. *Genes Dev,* 20, 3249-54.

69. 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, 4859-69.

70. Yu, Z., Chen, T., Hebert, J., Li, E. and Richard, S. (2009) A mouse PRMT1 null allele defines an essential role for arginine methylation in genome maintenance and cell proliferation. *Mol Cell Biol,* 29, 2982-96.

71. Boisvert, F.M., Dery, U., Masson, J.Y. and Richard, S. (2005) Arginine methylation of MRE11 by PRMT1 is required for DNA damage checkpoint control. *Genes Dev,* 19, 671-6.

72. Adams, M.M., Wang, B., Xia, Z., Morales, J.C., Lu, X., Donehower, L.A., Bochar, D.A., Elledge, S.J. and Carpenter, P.B. (2005) 53BP1 oligomerization is independent of its methylation by PRMT1. *Cell cycle (Georgetown, Tex,* 4, 1854-61.

73. 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 (Georgetown, Tex,* 4, 1834-41.

74. Jenuwein, T. and Allis, C.D. (2001) Translating the histone code. *Science,* 293, 1074-80.

75. An, W., Kim, J. and Roeder, R.G. (2004) Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53. *Cell,* 117, 735-48.

76. Rezai-Zadeh, N., Zhang, X., Namour, F., Fejer, G., Wen, Y.D., Yao, Y.L., Gyory, I., Wright, K. and Seto, E. (2003) Targeted recruitment of a histone H4-specific methyltransferase by the transcription factor YY1. *Genes Dev,* 17, 1019-29.

77. Covic, M., Hassa, P.O., Saccani, S., Buerki, C., Meier, N.I., Lombardi, C., Imhof, R., Bedford, M.T., Natoli, G. and Hottiger, M.O. (2005) Arginine methyltransferase CARM1 is a promoter-specific regulator of NF-kappaB-dependent gene expression. *The EMBO journal,* 24, 85-96.

78. Yadav, N., Cheng, D., Richard, S., Morel, M., Iyer, V.R., Aldaz, C.M. and Bedford, M.T. (2008) CARM1 promotes adipocyte differentiation by coactivating PPARgamma. *EMBO Rep,* 9, 193-8.

79. Zhao, X., Jankovic, V., Gural, A., Huang, G., Pardanani, A., Menendez, S., Zhang, J., Dunne, R., Xiao, A., Erdjument-Bromage, H. *et al.* (2008) Methylation of RUNX1 by PRMT1 abrogates SIN3A binding and potentiates its transcriptional activity. *Genes Dev,* 22, 640-53.

80. Krones-Herzig, A., Mesaros, A., Metzger, D., Ziegler, A., Lemke, U., Bruning, J.C. and Herzig, S. (2006) Signal-dependent control of gluconeogenic key enzyme genes through coactivator-associated arginine methyltransferase 1. *J Biol Chem,* 281, 3025-9.

81. 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, 13351-6.

82. Nicholson, T.B., Chen, T. and Richard, S. (2009) The physiological and pathophysiological role of PRMT1-mediated protein arginine methylation. *Pharmacol Res,* 60, 466-74.

83. 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, 1885-93.

84. Kleinschmidt, M.A., Streubel, G., Samans, B., Krause, M. and Bauer, U.M. (2008) The protein arginine methyltransferases CARM1 and PRMT1 cooperate in gene regulation. *Nucleic acids research,* 36, 3202-13.

85. Litt, M., Qiu, Y. and Huang, S. (2009) Histone arginine methylations: their roles in chromatin dynamics and transcriptional regulation. *Bioscience reports,* 29, 131-41.

86. Pal, S. and Sif, S. (2007) Interplay between chromatin remodelers and protein arginine methyltransferases. *J Cell Physiol,* 213, 306-15.

87. Wysocka, J., Allis, C.D. and Coonrod, S. (2006) Histone arginine methylation and its dynamic regulation. *Front Biosci,* 11, 344-55.

88. Klinge, C.M., Jernigan, S.C., Mattingly, K.A., Risinger, K.E. and Zhang, J. (2004) Estrogen response element-dependent regulation of transcriptional activation of estrogen receptors alpha and beta by coactivators and corepressors. *Journal of molecular endocrinology,* 33, 387-410.

89. 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. *Molecular cell,* 31, 212-21.

90. Matsuda, H., Paul, B.D., Choi, C.Y., Hasebe, T. and Shi, Y.B. (2009) Novel functions of protein arginine methyltransferase 1 in thyroid hormone receptor-mediated transcription and in the regulation of metamorphic rate in Xenopus laevis. *Mol Cell Biol,* 29, 745-57.

91. Strahl, B.D., Briggs, S.D., Brame, C.J., Caldwell, J.A., Koh, S.S., Ma, H., Cook, R.G., Shabanowitz, J., Hunt, D.F., Stallcup, M.R. *et al.* (2001) Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Curr Biol,* 11, 996-1000.

92. Teyssier, C., Le Romancer, M., Sentis, S., Jalaguier, S., Corbo, L. and Cavailles, V. (2010) Protein arginine methylation in estrogen signaling and estrogen-related cancers. *Trends in endocrinology and metabolism: TEM,* 21, 181-9.

93. 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, 853-7.

94. Dreesen, O. and Brivanlou, A.H. (2007) Signaling pathways in cancer and embryonic stem cells. *Stem cell reviews,* 3, 7-17.

95. Le Romancer, M., Treilleux, I., Bouchekioua-Bouzaghou, K., Sentis, S. and Corbo, L. (2010) Methylation, a key step for nongenomic estrogen signaling in breast tumors. *Steroids,* 75, 560-4.

96. Bedford, M.T. and Richard, S. (2005) Arginine methylation an emerging regulator of protein function. *Molecular cell,* 18, 263-72.

97. Koh, S.S., Chen, D., Lee, Y.H. and Stallcup, M.R. (2001) Synergistic enhancement of nuclear receptor function by p160 coactivators and two coactivators with protein methyltransferase activities. *J Biol Chem,* 276, 1089-98.

98. Teyssier, C., Ma, H., Emter, R., Kralli, A. and Stallcup, M.R. (2005) Activation of nuclear receptor coactivator PGC-1alpha by arginine methylation. *Genes Dev,* 19, 1466-73.

99. van Dijk, T.B., Gillemans, N., Stein, C., Fanis, P., Demmers, J., van de Corput, M., Essers, J., Grosveld, F., Bauer, U.M. and Philipsen, S. (2010) Friend of Prmt1, a novel chromatin target of protein arginine methyltransferases. *Mol Cell Biol,* 30, 260-72.

100. Augereau, P., Badia, E., Carascossa, S., Castet, A., Fritsch, S., Harmand, P.O., Jalaguier, S. and Cavailles, V. (2006) The nuclear receptor transcriptional coregulator RIP140. *Nuclear receptor signaling,* 4, e024.

101. Mostaqul Huq, M.D., Gupta, P., Tsai, N.P., White, R., Parker, M.G. and Wei, L.N. (2006) Suppression of receptor interacting protein 140 repressive activity by protein arginine methylation. *The EMBO journal,* 25, 5094-104.

102. Cheung, N., Chan, L.C., Thompson, A., Cleary, M.L. and So, C.W. (2007) Protein arginine-methyltransferase-dependent oncogenesis. *Nature cell biology,* 9, 1208-15.

103. Guendel, I., Carpio, L., Pedati, C., Schwartz, A., Teal, C., Kashanchi, F. and Kehn-Hall, K. (2010) Methylation of the tumor suppressor protein, BRCA1, influences its transcriptional cofactor function. *PloS one,* 5, e11379.

104. Lund, A.H. and van Lohuizen, M. (2004) Epigenetics and cancer. *Genes Dev,* 18, 2315-35.

105. 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, 957-69.

106. Cheng, D., Cote, J., Shaaban, S. and Bedford, M.T. (2007) The arginine methyltransferase CARM1 regulates the coupling of transcription and mRNA processing. *Molecular cell,* 25, 71-83.

107. Harrison, M.J., Tang, Y.H. and Dowhan, D.H. (2010) Protein arginine methyltransferase 6 regulates multiple aspects of gene expression. *Nucleic acids research,* 38, 2201-16.

108. Green, D.M., Marfatia, K.A., Crafton, E.B., Zhang, X., Cheng, X. and Corbett, A.H. (2002) Nab2p is required for poly(A) RNA export in Saccharomyces cerevisiae and is regulated by arginine methylation via Hmt1p. *J Biol Chem,* 277, 7752-60.

109. Xu, C. and Henry, M.F. (2004) Nuclear export of hnRNP Hrp1p and nuclear export of hnRNP Npl3p are linked and influenced by the methylation state of Npl3p. *Mol Cell Biol,* 24, 10742-56.

110. McBride, A.E., Cook, J.T., Stemmler, E.A., Rutledge, K.L., McGrath, K.A. and Rubens, J.A. (2005) Arginine methylation of yeast mRNA-binding protein Npl3 directly affects its function, nuclear export, and intranuclear protein interactions. *J Biol Chem,* 280, 30888-98.

111. Hung, M.L., Hautbergue, G.M., Snijders, A.P., Dickman, M.J. and Wilson, S.A. (2010) Arginine methylation of REF/ALY promotes efficient handover of mRNA to TAP/NXF1. *Nucleic acids research,* 38, 3351-61.

112. Bachand, F. and Silver, P.A. (2004) PRMT3 is a ribosomal protein methyltransferase that affects the cellular levels of ribosomal subunits. *The EMBO journal,* 23, 2641-50.

113. Swiercz, R., Cheng, D., Kim, D. and Bedford, M.T. (2007) Ribosomal protein rpS2 is hypomethylated in PRMT3-deficient mice. *J Biol Chem,* 282, 16917-23.

114. Boisvert, F.M., Cote, J., Boulanger, M.C. and Richard, S. (2003) A Proteomic Analysis of Arginine-methylated Protein Complexes. *Mol Cell Proteomics,* 2, 1319-30.

115. Bressan, G.C., Moraes, E.C., Manfiolli, A.O., Kuniyoshi, T.M., Passos, D.O., Gomes, M.D. and Kobarg, J. (2009) Arginine methylation analysis of the splicing-associated SR protein SFRS9/SRP30C. *Cell Mol Biol Lett,* 14, 657-69.

116. Chen, Y.C., Milliman, E.J., Goulet, I., Cote, J., Jackson, C.A., Vollbracht, J.A. and Yu, M.C. (2010) Protein Arginine Methylation Facilitates Co-transcriptional Recruitment of pre-mRNA Splicing Factors. *Mol Cell Biol*.

117. 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, 33009-21.

118. Ong, S.E., Mittler, G. and Mann, M. (2004) Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. *Nature Methods,* 1, 119-26.

119. 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, 2762-74.

120. Gonsalvez, G.B., Tian, L., Ospina, J.K., Boisvert, F.M., Lamond, A.I. and Matera, A.G. (2007) Two distinct arginine methyltransferases are required for biogenesis of Sm-class ribonucleoproteins. *J Cell Biol,* 178, 733-40.

121. 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. *Current Biology,* 11, 1990-4.

122. Tadesse, H., Deschenes-Furry, J., Boisvenue, S. and Cote, J. (2008) KH-type splicing regulatory protein interacts with survival motor neuron protein and is misregulated in spinal muscular atrophy. *Human molecular genetics,* 17, 506-24.

123. Ohkura, N., Takahashi, M., Yaguchi, H., Nagamura, Y. and Tsukada, T. (2005) Coactivator-associated arginine methyltransferase 1, CARM1, affects pre-mRNA splicing in an isoform-specific manner. *J Biol Chem,* 280, 28927-35.

124. Dreyfuss, G., Kim, V.N. and Kataoka, N. (2002) Messenger-RNA-binding proteins and the messages they carry. *Nature reviews,* 3, 195-205.

125. Kress, T.L., Krogan, N.J. and Guthrie, C. (2008) A single SR-like protein, Npl3, promotes pre-mRNA splicing in budding yeast. *Molecular cell,* 32, 727-34.

126. Piruat, J.I. and Aguilera, A. (1998) A novel yeast gene, THO2, is involved in RNA pol II transcription and provides new evidence for transcriptional elongation-associated recombination. *The EMBO journal,* 17, 4859-72.

127. Yu, M.C., Bachand, F., McBride, A.E., Komili, S., Casolari, J.M. and Silver, P.A. (2004) Arginine methyltransferase affects interactions and recruitment of mRNA processing and export factors. *Genes Dev,* 18, 2024-35.

128. Bedford, M.T., Frankel, A., Yaffe, M.B., Clarke, S., Leder, P. and Richard, S. (2000) Arginine methylation inhibits the binding of proline-rich ligands to Src homology 3, but not WW, domains. *J Biol Chem,* 275, 16030-6.

129. Kwak, Y.T., Guo, J., Prajapati, S., Park, K.J., Surabhi, R.M., Miller, B., Gehrig, P. and Gaynor, R.B. (2003) Methylation of SPT5 regulates its interaction with RNA polymerase II and transcriptional elongation properties. *Molecular cell,* 11, 1055-66.

130. Meissner, T., Krause, E., Lodige, I. and Vinkemeier, U. (2004) Arginine Methylation of STAT1; A Reassessment. *Cell,* 119, 587-9.

131. Mowen, K.A., Tang, J., Zhu, W., Schurter, B.T., Shuai, K., Herschman, H.R. and David, M. (2001) Arginine methylation of STAT1 modulates IFNalpha/beta-induced transcription. *Cell,* 104, 731-41.

132. Ostareck-Lederer, A., Ostareck, D.H., Rucknagel, K.P., Schierhorn, A., Moritz, B., Huttelmaier, S., Flach, N., Handoko, L. and Wahle, E. (2006) Asymmetric arginine dimethylation of heterogeneous nuclear ribonucleoprotein K by protein-arginine methyltransferase 1 inhibits its interaction with c-Src. *J Biol Chem,* 281, 11115-25.

133. Callebaut, I. and Mornon, J.P. (1997) The human EBNA-2 coactivator p100: multidomain organization and relationship to the staphylococcal nuclease fold and to the tudor protein involved in Drosophila melanogaster development. *Biochem J,* 321 (Pt 1), 125-32.

134. Ponting, C.P. (1997) Tudor domains in proteins that interact with RNA. *Trends Biochem Sci,* 22, 51-2.

135. Lasko, P. (2010) Tudor domain. *Curr Biol,* 20, R666-7.

136. 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.

137. Sprangers, R., Selenko, P., Sattler, M., Sinning, I. and Groves, M.R. (2003) Definition of domain boundaries and crystallization of the SMN Tudor domain. *Acta Crystallographica Section D Biological Crystallography,* 59, 366-8.

138. Maurer-Stroh, S., Dickens, N.J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F. and Ponting, C.P. (2003) The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends Biochem Sci,* 28, 69-74.

139. Jin, J., Xie, X., Chen, C., Park, J.G., Stark, C., James, D.A., Olhovsky, M., Linding, R., Mao, Y. and Pawson, T. (2009) Eukaryotic protein domains as functional units of cellular evolution. *Science signaling,* 2, ra76.

140. Chen, C., Jin, J., James, D.A., Adams-Cioaba, M.A., Park, J.G., Guo, Y., Tenaglia, E., Xu, C., Gish, G., Min, J. *et al.* (2009) Mouse Piwi interactome identifies binding mechanism of Tdrkh Tudor domain to arginine methylated Miwi. *Proc Natl Acad Sci U S A,* 106, 20336-41.

141. 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. *Nature cell biology,* 11, 652-8.

142. Siomi, M.C., Mannen, T. and Siomi, H. (2010) How does the royal family of Tudor rule the PIWI-interacting RNA pathway? *Genes Dev,* 24, 636-46.

143. 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, 1749-62.

144. Wang, J., Saxe, J.P., Tanaka, T., Chuma, S. and Lin, H. (2009) Mili interacts with tudor domain-containing protein 1 in regulating spermatogenesis. *Curr Biol,* 19, 640-4.

145. Shaw, N., Zhao, M., Cheng, C., Xu, H., Saarikettu, J., Li, Y., Da, Y., Yao, Z., Silvennoinen, O., Yang, J. *et al.* (2007) The multifunctional human p100 protein 'hooks' methylated ligands. *Nature Structural Molecular Biology,* 14, 779-84.

146. Liu, K., Chen, C., Guo, Y., Lam, R., Bian, C., Xu, C., Zhao, D.Y., Jin, J., MacKenzie, F., Pawson, T. *et al.* (2010) Structural basis for recognition of arginine methylated Piwi proteins by the extended Tudor domain. *Proc Natl Acad Sci U S A,* 107, 18398-403.

147. Caudy, A.A., Ketting, R.F., Hammond, S.M., Denli, A.M., Bathoorn, A.M., Tops, B.B., Silva, J.M., Myers, M.M., Hannon, G.J. and Plasterk, R.H. (2003) A micrococcal nuclease homologue in RNAi effector complexes. *Nature,* 425, 411-4.

148. Cote, J. and Richard, S. (2005) Tudor domains bind symmetrical dimethylated arginines. *J Biol Chem,* 280, 28476-83.

149. Friesen, W.J., Paushkin, S., Wyce, A., Massenet, S., Pesiridis, G.S., Van Duyne, G., Rappsilber, J., Mann, M. and Dreyfuss, G. (2001) The methylosome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. *Mol Cell Biol,* 21, 8289-300.

150. Meister, G., Buhler, D., Laggerbauer, B., Zobawa, M., Lottspeich, F. and Fischer, U. (2000) Characterization of a nuclear 20S complex containing the survival of motor neurons (SMN) protein and a specific subset of spliceosomal Sm proteins. *Human molecular genetics,* 9, 1977-86.

151. Pu, W.T., Krapivinsky, G.B., Krapivinsky, L. and Clapham, D.E. (1999) pICln inhibits snRNP biogenesis by binding core spliceosomal proteins. *Mol Cell Biol,* 19, 4113-20.

152. Neubauer, G., King, A., Rappsilber, J., Calvio, C., Watson, M., Ajuh, P., Sleeman, J., Lamond, A. and Mann, M. (1998) Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nature genetics,* 20, 46-50.

153. Rappsilber, J., Ajuh, P., Lamond, A.I. and Mann, M. (2001) SPF30 is an essential human splicing factor required for assembly of the U4/U5/U6 tri-small nuclear ribonucleoprotein into the spliceosome. *J Biol Chem,* 276, 31142-50.

154. Barth, S., Liss, M., Voss, M.D., Dobner, T., Fischer, U., Meister, G. and Grasser, F.A. (2003) Epstein-Barr virus nuclear antigen 2 binds via its methylated arginine-glycine repeat to the survival motor neuron protein. *J Virol,* 77, 5008-13.

155. 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, 1531-42.

156. Friesen, W.J., Massenet, S., Paushkin, S., Wyce, A. and Dreyfuss, G. (2001) SMN, the product of the spinal muscular atrophy gene, binds preferentially to dimethylarginine-containing protein targets. *Molecular cell,* 7, 1111-7.

157. Hebert, M.D., Shpargel, K.B., Ospina, J.K., Tucker, K.E. and Matera, A.G. (2002) Coilin methylation regulates nuclear body formation. *Development Cell,* 3, 329-37.

158. Kim, J., Daniel, J., Espejo, A., Lake, A., Krishna, M., Xia, L., Zhang, Y. and Bedford, M.T. (2006) Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep,* 7, 397-403.

159. Charier, G., Couprie, J., Alpha-Bazin, B., Meyer, V., Quemeneur, E., Guerois, R., Callebaut, I., Gilquin, B. and Zinn-Justin, S. (2004) The Tudor tandem of 53BP1: a new structural motif involved in DNA and RG-rich peptide binding. *Structure (Camb),* 12, 1551-62.

160. Botuyan, M.V., Lee, J., Ward, I.M., Kim, J.E., Thompson, J.R., Chen, J. and Mer, G. (2006) Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair. *Cell,* 127, 1361-73.

161. Corsini, L. and Sattler, M. (2007) Tudor hooks up with DNA repair. *Nat Struct Mol Biol,* 14, 98-9.

162. Klose, R.J. and Zhang, Y. (2007) Regulation of histone methylation by demethylimination and demethylation. *Nature reviews,* 8, 307-18.

163. Hartmuth, K., Urlaub, H., Vornlocher, H.P., Will, C.L., Gentzel, M., Wilm, M. and Luhrmann, R. (2002) Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. *Proc Natl Acad Sci U S A,* 99, 16719-24.

164. Theobald, D.L., Mitton-Fry, R.M. and Wuttke, D.S. (2003) Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct,* 32, 115-33.

165. Hurley, J.H., Lee, S. and Prag, G. (2006) Ubiquitin-binding domains. *Biochem J,* 399, 361-72.

166. Kashima, I., Jonas, S., Jayachandran, U., Buchwald, G., Conti, E., Lupas, A.N. and Izaurralde, E. (2010) SMG6 interacts with the exon junction complex via two conserved EJC-binding motifs (EBMs) required for nonsense-mediated mRNA decay. *Genes Dev,* 24, 2440-50.

167. Nagahata, T., Onda, M., Emi, M., Nagai, H., Tsumagari, K., Fujimoto, T., Hirano, A., Sato, T., Nishikawa, K., Akiyama, F. *et al.* (2004) Expression profiling to predict postoperative prognosis for estrogen receptor-negative breast cancers by analysis of 25,344 genes on a cDNA microarray. *Cancer Sci,* 95, 218-25.

168. 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, 19866-76.

169. Cote, J., Boisvert, F.M., Boulanger, M.C., Bedford, M.T. and Richard, S. (2003) Sam68 RNA binding protein is an in vivo substrate for protein arginine N-methyltransferase 1. *Mol Biol Cell,* 14, 274-87.

170. Liu, Q. and Dreyfuss, G. (1995) In vivo and in vitro arginine methylation of RNA-binding proteins. *Mol Cell Biol,* 15, 2800-8.

171. Nichols, R.C., Wang, X.W., Tang, J., Hamilton, B.J., High, F.A., Herschman, H.R. and Rigby, W.F. (2000) The RGG domain in hnRNP A2 affects subcellular localization. *Exp Cell Res,* 256, 522-32.

172. Weiss, V.H., McBride, A.E., Soriano, M.A., Filman, D.J., Silver, P.A. and Hogle, J.M. (2000) The structure and oligomerization of the yeast arginine methyltransferase, Hmt1. *Nat Struct Biol,* 7, 1165-71.

173. Chen, T., Boisvert, F.M., Bazett-Jones, D.P. and Richard, S. (1999) A role for the GSG domain in localizing Sam68 to novel nuclear structures in cancer cell lines. *Mol Biol Cell,* 10, 3015-33.

174. Bedford, M.T., Reed, R. and Leder, P. (1998) WW domain-mediated interactions reveal a spliceosome-associated protein that binds a third class of proline-rich motif: the proline glycine and methionine-rich motif. *Proc Natl Acad Sci U S A,* 95, 10602-7.

175. Pawlak, M.R., Banik-Maiti, S., Pietenpol, J.A. and Ruley, H.E. (2002) Protein arginine methyltransferase I: substrate specificity and role in hnRNP assembly. *J Cell Biochem,* 87, 394-407.

176. Kent, W.J. (2002) BLAT--the BLAST-like alignment tool. *Genome research,* 12, 656-64.

177. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M. and Haussler, D. (2002) The human genome browser at UCSC. *Genome research,* 12, 996-1006.

178. Knudsen, S. (1999) Promoter2.0: for the recognition of PolII promoter sequences. *Bioinformatics (Oxford, England),* 15, 356-61.

179. Suzuki, Y., Yamashita, R., Nakai, K. and Sugano, S. (2002) DBTSS: DataBase of human Transcriptional Start Sites and full-length cDNAs. *Nucleic acids research,* 30, 328-31.

180. Reese, M.G. (2001) Application of a time-delay neural network to promoter annotation in the Drosophila melanogaster genome. *Computers & chemistry,* 26, 51-6.

181. Tarn, W.Y. and Steitz, J.A. (1997) Pre-mRNA splicing: the discovery of a new spliceosome doubles the challenge. *Trends Biochem Sci,* 22, 132-7.

182. Tarn, W.Y. and Steitz, J.A. (1996) A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vitro. *Cell,* 84, 801-11.

183. Hall, S.L. and Padgett, R.A. (1996) Requirement of U12 snRNA for in vivo splicing of a minor class of eukaryotic nuclear pre-mRNA introns. *Science,* 271, 1716-8.

184. Hall, S.L. and Padgett, R.A. (1994) Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice sites. *J Mol Biol,* 239, 357-65.

185. 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, 17122-9.

186. Miranda, T.B., Khusial, P., Cook, J.R., Lee, J.H., Gunderson, S.I., Pestka, S., Zieve, G.W. and Clarke, S. (2004) Spliceosome Sm proteins D1, D3, and B/B' are asymmetrically dimethylated at arginine residues in the nucleus. *Biochem Biophys Res Commun,* 323, 382-7.

187. la Cour, T., Kiemer, L., Molgaard, A., Gupta, R., Skriver, K. and Brunak, S. (2004) Analysis and prediction of leucine-rich nuclear export signals. *Protein Eng Des Sel,* 17, 527-36.

188. Fried, H. and Kutay, U. (2003) Nucleocytoplasmic transport: taking an inventory. *Cell Mol Life Sci,* 60, 1659-88.

189. Gorlich, D. and Kutay, U. (1999) Transport between the cell nucleus and the cytoplasm. *Annual review of cell and developmental biology,* 15, 607-60.

190. Kutay, U. and Guttinger, S. (2005) Leucine-rich nuclear-export signals: born to be weak. *Trends in cell biology,* 15, 121-4.

191. Maquat, L.E. (2005) Nonsense-mediated mRNA decay in mammals. *J Cell Sci,* 118, 1773-6.

192. Chang, Y.F., Chan, W.K., Imam, J.S. and Wilkinson, M.F. (2007) Alternatively spliced T-cell receptor transcripts are up-regulated in response to disruption of either splicing elements or reading frame. *J Biol Chem,* 282, 29738-47.

193. Isken, O. and Maquat, L.E. (2007) Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. *Genes Dev,* 21, 1833-56.

194. Anczukow, O., Ware, M.D., Buisson, M., Zetoune, A.B., Stoppa-Lyonnet, D., Sinilnikova, O.M. and Mazoyer, S. (2008) Does the nonsense-mediated mRNA decay mechanism prevent the synthesis of truncated BRCA1, CHK2, and p53 proteins? *Human mutation,* 29, 65-73.

195. Hughes, T.A. (2006) Regulation of gene expression by alternative untranslated regions. *Trends Genet,* 22, 119-22.

196. Ikenaka, K., Miyata, S., Mori, Y., Koyama, Y., Taneda, T., Okuda, H., Kousaka, A. and Tohyama, M. (2006) Immunohistochemical and western analyses of protein arginine N-methyltransferase 3 in the mouse brain. *Neuroscience,* 141, 1971-82.

197. Smith, W.A., Schurter, B.T., Wong-Staal, F. and David, M. (2004) Arginine methylation of RNA helicase a determines its subcellular localization. *J Biol Chem,* 279, 22795-8.

198. Cimato, T.R., Tang, J., Xu, Y., Guarnaccia, C., Herschman, H.R., Pongor, S. and Aletta, J.M. (2002) Nerve growth factor-mediated increases in protein methylation occur predominantly at type I arginine methylation sites and involve protein arginine methyltransferase 1. *J Neurosci Res,* 67, 435-42.

199. el-Ghissassi, F., Valsesia-Wittmann, S., Falette, N., Duriez, C., Walden, P.D. and Puisieux, A. (2002) BTG2(TIS21/PC3) induces neuronal differentiation and prevents apoptosis of terminally differentiated PC12 cells. *Oncogene,* 21, 6772-78.

200. Abramovich, C., Yakobson, B., Chebath, J. and Revel, M. (1997) A protein-arginine methyltransferase binds to the intracytoplasmic domain of the IFNAR1 chain in the type I interferon receptor. *The EMBO journal,* 16, 260-6.

201. Altschuler, L., Wook, J.O., Gurari, D., Chebath, J. and Revel, M. (1999) Involvement of receptor-bound protein methyltransferase PRMT1 in antiviral and antiproliferative effects of type I interferons. *J Interferon Cytokine Res,* 19, 189-95.

202. Singh, V., Miranda, T.B., Jiang, W., Frankel, A., Roemer, M.E., Robb, V.A., Gutmann, D.H., Herschman, H.R., Clarke, S. and Newsham, I.F. (2004) DAL-1/4.1B tumor suppressor interacts with protein arginine N-methyltransferase 3 (PRMT3) and inhibits its ability to methylate substrates in vitro and in vivo. *Oncogene,* 23, 7761-71.

203. Jiang, W., Roemer, M.E. and Newsham, I.F. (2005) The tumor suppressor DAL-1/4.1B modulates protein arginine N-methyltransferase 5 activity in a substrate-specific manner. *Biochem Biophys Res Commun,* 329, 522-30.

204. Fackelmayer, F.O. (2005) Protein arginine methyltransferases: guardians of the Arg? *Trends Biochem Sci,* 30, 666-71.

205. Hong, H., Kao, C., Jeng, M.H., Eble, J.N., Koch, M.O., Gardner, T.A., Zhang, S., Li, L., Pan, C.X., Hu, Z. *et al.* (2004) Aberrant expression of CARM1, a transcriptional coactivator of androgen receptor, in the development of prostate carcinoma and androgen-independent status. *Cancer,* 101, 83-9.

206. Majumder, S., Liu, Y., Ford, O.H., 3rd, Mohler, J.L. and Whang, Y.E. (2006) Involvement of arginine methyltransferase CARM1 in androgen receptor function and prostate cancer cell viability. *Prostate,* 66, 1292-301.

207. Edberg, D.D., Bruce, J.E., Siems, W.F. and Reeves, R. (2004) In vivo posttranslational modifications of the high mobility group A1a proteins in breast cancer cells of differing metastatic potential. *Biochemistry,* 43, 11500-15.

208. Sgarra, R., Lee, J., Tessari, M.A., Altamura, S., Spolaore, B., Giancotti, V., Bedford, M.T. and Manfioletti, G. (2006) The AT-hook of the chromatin architectural transcription factor high mobility group A1a is arginine-methylated by protein arginine methyltransferase 6. *J Biol Chem,* 281, 3764-72.

209. Gros, L., Renodon-Corniere, A., de Saint Vincent, B.R., Feder, M., Bujnicki, J.M. and Jacquemin-Sablon, A. (2006) Characterization of prmt7alpha and beta isozymes from Chinese hamster cells sensitive and resistant to topoisomerase II inhibitors. *Biochim Biophys Acta*.

210. Moore, M.J. and Proudfoot, N.J. (2009) Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell,* 136, 688-700.

211. Perales, R. and Bentley, D. (2009) "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions. *Molecular cell,* 36, 178-91.

212. Han, S.P., Tang, Y.H. and Smith, R. (2010) Functional diversity of the hnRNPs: past, present and perspectives. *Biochem J,* 430, 379-92.

213. Zhong, X.Y., Wang, P., Han, J., Rosenfeld, M.G. and Fu, X.D. (2009) SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. *Molecular cell,* 35, 1-10.

214. Yu, M.C. (2011) The role of protein arginine methylation in mRNP dynamics. *Molecular Biology International,* 2011, 10.

215. Lukong, K.E. and Richard, S. (2004) Arginine methylation signals mRNA export. *Nature Structural Molecular Biology,* 11, 914-5.

216. Ramos, A., Hollingworth, D., Adinolfi, S., Castets, M., Kelly, G., Frenkiel, T.A., Bardoni, B. and Pastore, A. (2006) The structure of the N-terminal domain of the fragile X mental retardation protein: a platform for protein-protein interaction. *Structure,* 14, 21-31.

217. O'Connell, M.A. and Keegan, L.P. (2006) Drosha versus ADAR: wrangling over pri-miRNA. *Nature Structural Molecular Biology,* 13, 3-4.

218. Little, J.T. and Jurica, M.S. (2008) Splicing factor SPF30 bridges an interaction between the prespliceosome protein U2AF35 and tri-small nuclear ribonucleoprotein protein hPrp3. *J Biol Chem,* 283, 8145-52.

219. Goulet, I., Boisvenue, S., Mokas, S., Mazroui, R. and Cote, J. (2008) TDRD3, a novel Tudor domain-containing protein, localizes to cytoplasmic stress granules. *Human molecular genetics,* 17, 3055-74.

220. Fox, A.H., Lam, Y.W., Leung, A.K., Lyon, C.E., Andersen, J., Mann, M. and Lamond, A.I. (2002) Paraspeckles: a novel nuclear domain. *Curr Biol,* 12, 13-25.

221. Stuurman, N., de Graaf, A., Floore, A., Josso, A., Humbel, B., de Jong, L. and van Driel, R. (1992) A monoclonal antibody recognizing nuclear matrix-associated nuclear bodies. *J Cell Sci,* 101 (Pt 4), 773-84.

222. Chabot, B. (1994) *Synthesis and purification of RNA substrates.* Oxford University Press, Oxford.

223. Chabot, B., Blanchette, M., Lapierre, I. and La Branche, H. (1997) An intron element modulating 5' splice site selection in the hnRNP A1 pre-mRNA interacts with hnRNP A1. *Mol Cell Biol,* 17, 1776-86.

224. Nasim, F.U., Hutchison, S., Cordeau, M. and Chabot, B. (2002) High-affinity hnRNP A1 binding sites and duplex-forming inverted repeats have similar effects on 5' splice site selection in support of a common looping out and repression mechanism. *Rna,* 8, 1078-89.

225. Fox, A.H., Bond, C.S. and Lamond, A.I. (2005) P54nrb forms a heterodimer with PSP1 that localizes to paraspeckles in an RNA-dependent manner. *Mol Biol Cell,* 16, 5304-15.

226. Dettwiler, S., Aringhieri, C., Cardinale, S., Keller, W. and Barabino, S.M. (2004) Distinct sequence motifs within the 68-kDa subunit of cleavage factor Im mediate RNA binding, protein-protein interactions, and subcellular localization. *J Biol Chem,* 279, 35788-97.

227. Fox, A.H. and Lamond, A.I. (2010) Paraspeckles. *Cold Spring Harbor perspectives in biology,* 2, a000687.

228. Carmo-Fonseca, M. (2002) New clues to the function of the Cajal body. *EMBO Rep,* 3, 726-7.

229. Cioce, M. and Lamond, A.I. (2005) Cajal bodies: a long history of discovery. *Annual review of cell and developmental biology,* 21, 105-31.

230. Gall, J.G. (2000) Cajal bodies: the first 100 years. *Annual review of cell and developmental biology,* 16, 273-300.

231. Jady, B.E., Richard, P., Bertrand, E. and Kiss, T. (2006) Cell cycle-dependent recruitment of telomerase RNA and Cajal bodies to human telomeres. *Mol Biol Cell,* 17, 944-54.

232. Kiss, A.M., Jady, B.E., Darzacq, X., Verheggen, C., Bertrand, E. and Kiss, T. (2002) A Cajal body-specific pseudouridylation guide RNA is composed of two box H/ACA snoRNA-like domains. *Nucleic acids research,* 30, 4643-9.

233. Kiss, T. (2004) Biogenesis of small nuclear RNPs. *J Cell Sci,* 117, 5949-51.

234. Ogg, S.C. and Lamond, A.I. (2002) Cajal bodies and coilin--moving towards function. *J Cell Biol,* 159, 17-21.

235. Schumperli, D. and Pillai, R.S. (2004) The special Sm core structure of the U7 snRNP: far-reaching significance of a small nuclear ribonucleoprotein. *Cell Mol Life Sci,* 61, 2560-70.

236. Andrade, L.E., Chan, E.K., Raska, I., Peebles, C.L., Roos, G. and Tan, E.M. (1991) Human autoantibody to a novel protein of the nuclear coiled body: immunological characterization and cDNA cloning of p80-coilin. *The Journal of experimental medicine,* 173, 1407-19.

237. Raska, I., Andrade, L.E., Ochs, R.L., Chan, E.K., Chang, C.M., Roos, G. and Tan, E.M. (1991) Immunological and ultrastructural studies of the nuclear coiled body with autoimmune antibodies. *Exp Cell Res,* 195, 27-37.

238. Tuma, R.S., Stolk, J.A. and Roth, M.B. (1993) Identification and characterization of a sphere organelle protein. *J Cell Biol,* 122, 767-73.

239. Dundr, M. and Misteli, T. (2001) Functional architecture in the cell nucleus. *Biochem J,* 356, 297-310.

240. Grande, M.A., van der Kraan, I., van Steensel, B., Schul, W., de The, H., van der Voort, H.T., de Jong, L. and van Driel, R. (1996) PML-containing nuclear bodies: their spatial distribution in relation to other nuclear components. *J Cell Biochem,* 63, 280-91.

241. Sun, J., Xu, H., Subramony, S.H. and Hebert, M.D. (2005) Interactions between coilin and PIASy partially link Cajal bodies to PML bodies. *J Cell Sci,* 118, 4995-5003.

242. Bernardi, R. and Pandolfi, P.P. (2007) Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. *Nature reviews,* 8, 1006-16.

243. Bernardi, R., Papa, A. and Pandolfi, P.P. (2008) Regulation of apoptosis by PML and the PML-NBs. *Oncogene,* 27, 6299-312.

244. Lallemand-Breitenbach, V. and de The, H. (2010) PML nuclear bodies. *Cold Spring Harbor perspectives in biology,* 2, a000661.

245. Pinol-Roma, S. and Dreyfuss, G. (1991) Transcription-dependent and transcription-independent nuclear transport of hnRNP proteins. *Science,* 253, 312-4.

246. Pinol-Roma, S. and Dreyfuss, G. (1992) Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. *Nature,* 355, 730-2.

247. 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. *Molecular cell,* 40, 1016-23.

248. Khacho, M., Mekhail, K., Pilon-Larose, K., Pause, A., Cote, J. and Lee, S. (2008) eEF1A is a novel component of the mammalian nuclear protein export machinery. *Mol Biol Cell,* 19, 5296-308.

249. Lee, S., Neumann, M., Stearman, R., Stauber, R., Pause, A., Pavlakis, G.N. and Klausner, R.D. (1999) Transcription-dependent nuclear-cytoplasmic trafficking is required for the function of the von Hippel-Lindau tumor suppressor protein. *Mol Cell Biol,* 19, 1486-97.

250. Khacho, M., Mekhail, K., Pilon-Larose, K., Payette, J. and Lee, S. (2008) Cancer-causing mutations in a novel transcription-dependent nuclear export motif of VHL abrogate oxygen-dependent degradation of hypoxia-inducible factor. *Mol Cell Biol,* 28, 302-14.

251. Spector, D.L. (1993) Macromolecular domains within the cell nucleus. *Annual review of cell biology,* 9, 265-315.

252. Boisvert, F.M., Hendzel, M.J., Masson, J.Y. and Richard, S. (2005) Methylation of MRE11 regulates its nuclear compartmentalization. *Cell cycle (Georgetown, Tex,* 4, 981-9.

253. Stuurman, N., Meijne, A.M., van der Pol, A.J., de Jong, L., van Driel, R. and van Renswoude, J. (1990) The nuclear matrix from cells of different origin. Evidence for a common set of matrix proteins. *J Biol Chem,* 265, 5460-5.

254. Linder, B., Plottner, 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, 3236-46.

255. Rao, P.N. and Engelberg, J. (1966) Mitotic non-disjunction of sister chromatids and anomalous mitosis induced by low temperatures in HeLa cells. *Exp Cell Res,* 43, 332-42.

256. Sobell, H.M. (1985) Actinomycin and DNA transcription. *Proc Natl Acad Sci U S A,* 82, 5328-31.

257. Guy, A.L. and Taylor, J.H. (1978) Actinomycin D inhibits initiation of DNA replication in mammalian cells. *Proc Natl Acad Sci U S A,* 75, 6088-92.

258. Dellaire, G., Ching, R.W., Dehghani, H., Ren, Y. and Bazett-Jones, D.P. (2006) The number of PML nuclear bodies increases in early S phase by a fission mechanism. *J Cell Sci,* 119, 1026-33.

259. Terris, B., Baldin, V., Dubois, S., Degott, C., Flejou, J.F., Henin, D. and Dejean, A. (1995) PML nuclear bodies are general targets for inflammation and cell proliferation. *Cancer Res,* 55, 1590-7.

260. Kiesslich, A., von Mikecz, A. and Hemmerich, P. (2002) Cell cycle-dependent association of PML bodies with sites of active transcription in nuclei of mammalian cells. *Journal of structural biology,* 140, 167-79.

261. Gall, J.G., Bellini, M., Wu, Z. and Murphy, C. (1999) Assembly of the nuclear transcription and processing machinery: Cajal bodies (coiled bodies) and transcriptosomes. *Mol Biol Cell,* 10, 4385-402.

262. Ma, T., Van Tine, B.A., Wei, Y., Garrett, M.D., Nelson, D., Adams, P.D., Wang, J., Qin, J., Chow, L.T. and Harper, J.W. (2000) Cell cycle-regulated phosphorylation of p220(NPAT) by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. *Genes Dev,* 14, 2298-313.

263. Zhao, J., Kennedy, B.K., Lawrence, B.D., Barbie, D.A., Matera, A.G., Fletcher, J.A. and Harlow, E. (2000) NPAT links cyclin E-Cdk2 to the regulation of replication-dependent histone gene transcription. *Genes Dev,* 14, 2283-97.

264. Boisvert, F.M., Hendzel, M.J. and Bazett-Jones, D.P. (2000) Promyelocytic leukemia (PML) nuclear bodies are protein structures that do not accumulate RNA. *J Cell Biol,* 148, 283-92.

265. Plottner, O. (2002) SMNrp und TUBA: isolierung und charakterisierung zweier neuer gene mit homologie zu SMN, dem krankheitsgen der spinalen muskelatrophie. *Mathematisch-Naturwissenschaftlichen Fakultat.* Rheinischen Friedrich-Wilhelms-Universitat Bonn, Bonn, Vol. Doctorate, pp. 1-157.

266. Lamond, A.I. and Earnshaw, W.C. (1998) Structure and function in the nucleus. *Science,* 280, 547-53.

267. Hieronymus, H. and Silver, P.A. (2004) A systems view of mRNP biology. *Genes Dev,* 18, 2845-60.

268. Pellizzoni, L., Kataoka, N., Charroux, B. and Dreyfuss, G. (1998) A novel function for SMN, the spinal muscular atrophy disease gene product, in pre-mRNA splicing. *Cell,* 95, 615-24.

269. Aguilera, A. (2005) Cotranscriptional mRNP assembly: from the DNA to the nuclear pore. *Curr Opin Cell Biol,* 17, 242-50.

270. Chang, Y.F., Imam, J.S. and Wilkinson, M.F. (2007) The nonsense-mediated decay RNA surveillance pathway. *Annual review of biochemistry,* 76, 51-74.

271. Le Hir, H. and Andersen, G.R. (2008) Structural insights into the exon junction complex. *Current opinion in structural biology,* 18, 112-9.

272. Ashton-Beaucage, D., Udell, C.M., Lavoie, H., Baril, C., Lefrancois, M., Chagnon, P., Gendron, P., Caron-Lizotte, O., Bonneil, E., Thibault, P. *et al.* (2010) The exon junction complex controls the splicing of MAPK and other long intron-containing transcripts in Drosophila. *Cell,* 143, 251-62.

273. Pawson, T. (2004) Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems. *Cell,* 116, 191-203.

274. Chang, B., Chen, Y., Zhao, Y. and Bruick, R.K. (2007) JMJD6 is a histone arginine demethylase. *Science,* 318, 444-7.

275. Cuthbert, G.L., Daujat, S., Snowden, A.W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P.D., Tempst, P., Bannister, A.J. *et al.* (2004) Histone deimination antagonizes arginine methylation. *Cell,* 118, 545-53.

276. 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, 279-83.

277. Roiha, H., Shuster, E.O., Brow, D.A. and Guthrie, C. (1989) Small nuclear RNAs from budding yeasts: phylogenetic comparisons reveal extensive size variation. *Gene,* 82, 137-44.

278. Pellizzoni, L. (2007) Chaperoning ribonucleoprotein biogenesis in health and disease. *EMBO Rep,* 8, 340-5.

279. Terns, M.P. and Terns, R.M. (2001) Macromolecular complexes: SMN--the master assembler. *Curr Biol,* 11, R862-4.

280. Yong, J., Wan, L. and Dreyfuss, G. (2004) Why do cells need an assembly machine for RNA-protein complexes? *Trends in cell biology,* 14, 226-32.

281. Meister, G., Hannus, S., Plottner, O., Baars, T., Hartmann, E., Fakan, S., Laggerbauer, B. and Fischer, U. (2001) SMNrp is an essential pre-mRNA splicing factor required for the formation of the mature spliceosome. *The EMBO journal,* 20, 2304-14.

282. Yang, J., Valineva, T., Hong, J., Bu, T., Yao, Z., Jensen, O.N., Frilander, M.J. and Silvennoinen, O. (2007) Transcriptional co-activator protein p100 interacts with snRNP proteins and facilitates the assembly of the spliceosome. *Nucleic acids research,* 35, 4485-94.

283. Anderson, P. and Kedersha, N. (2008) Stress granules: the Tao of RNA triage. *Trends Biochem Sci,* 33, 141-50.

284. Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fritzler, M.J., Scheuner, D., Kaufman, R.J., Golan, D.E. and Anderson, P. (2005) Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J Cell Biol,* 169, 871-84.

285. Kiebler, M.A. and Bassell, G.J. (2006) Neuronal RNA granules: movers and makers. *Neuron,* 51, 685-90.

286. Kedersha, N. and Anderson, P. (2007) Mammalian stress granules and processing bodies. *Methods Enzymol,* 431, 61-81.

287. Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell,* 44, 283-92.

288. Yin, J., Sobeck, A., Xu, C., Meetei, A.R., Hoatlin, M., Li, L. and Wang, W. (2005) BLAP75, an essential component of Bloom's syndrome protein complexes that maintain genome integrity. *The EMBO journal,* 24, 1465-76.

289. Dolzhanskaya, N., Merz, G. and Denman, R.B. (2006) Oxidative stress reveals heterogeneity of FMRP granules in PC12 cell neurites. *Brain Res,* 1112, 56-64.

290. Mazroui, R., Huot, M.E., Tremblay, S., Filion, C., Labelle, Y. and Khandjian, E.W. (2002) Trapping of messenger RNA by Fragile X Mental Retardation protein into cytoplasmic granules induces translation repression. *Human molecular genetics,* 11, 3007-17.

291. Tourriere, 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, 823-31.

292. Gilks, N., Kedersha, N., Ayodele, M., Shen, L., Stoecklin, G., Dember, L.M. and Anderson, P. (2004) Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol Biol Cell,* 15, 5383-98.

293. Eystathioy, T., Chan, E.K., Tenenbaum, S.A., Keene, J.D., Griffith, K. and Fritzler, M.J. (2002) A phosphorylated cytoplasmic autoantigen, GW182, associates with a unique population of human mRNAs within novel cytoplasmic speckles. *Mol Biol Cell,* 13, 1338-51.

294. Eystathioy, T., Jakymiw, A., Chan, E.K., Seraphin, B., Cougot, N. and Fritzler, M.J. (2003) The GW182 protein colocalizes with mRNA degradation associated proteins hDcp1 and hLSm4 in cytoplasmic GW bodies. *RNA,* 9, 1171-3.

295. Kedersha, N.L., Gupta, M., Li, W., Miller, I. and Anderson, P. (1999) RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules. *J Cell Biol,* 147, 1431-42.

296. Corbin, F., Bouillon, M., Fortin, A., Morin, S., Rousseau, F. and Khandjian, E.W. (1997) The fragile X mental retardation protein is associated with poly(A)+ mRNA in actively translating polyribosomes. *Human molecular genetics,* 6, 1465-72.

297. Khandjian, E.W., Corbin, F., Woerly, S. and Rousseau, F. (1996) The fragile X mental retardation protein is associated with ribosomes. *Nature genetics,* 12, 91-3.

298. Kim, S.H., Dong, W.K., Weiler, I.J. and Greenough, W.T. (2006) Fragile X mental retardation protein shifts between polyribosomes and stress granules after neuronal injury by arsenite stress or in vivo hippocampal electrode insertion. *J Neurosci,* 26, 2413-8.

299. Hua, Y. and Zhou, J. (2004) Survival motor neuron protein facilitates assembly of stress granules. *FEBS Lett,* 572, 69-74.

300. Zhang, R., Barker, L., Pinchev, D., Marshall, J., Rasamoelisolo, M., Smith, C., Kupchak, P., Kireeva, I., Ingratta, L. and Jackowski, G. (2004) Mining biomarkers in human sera using proteomic tools. *Proteomics,* 4, 244-56.

301. Young, P.J., Francis, J.W., Lince, D., Coon, K., Androphy, E.J. and Lorson, C.L. (2003) The Ewing's sarcoma protein interacts with the Tudor domain of the survival motor neuron protein. *Brain Res Mol Brain Res,* 119, 37-49.

302. Burt, E.C., Towers, P.R. and Sattelle, D.B. (2006) Caenorhabditis elegans in the study of SMN-interacting proteins: a role for SMI-1, an orthologue of human Gemin2 and the identification of novel components of the SMN complex. *Invertebrate Neuroscience,* 6, 145-59.

303. Saunders, L.R., Perkins, D.J., Balachandran, S., Michaels, R., Ford, R., Mayeda, A. and Barber, G.N. (2001) Characterization of two evolutionarily conserved, alternatively spliced nuclear phosphoproteins, NFAR-1 and -2, that function in mRNA processing and interact with the double-stranded RNA-dependent protein kinase, PKR. *J Biol Chem,* 276, 32300-12.

304. Belyanskaya, L.L., Gehrig, P.M. and Gehring, H. (2001) Exposure on cell surface and extensive arginine methylation of ewing sarcoma (EWS) protein. *J Biol Chem,* 276, 18681-7.

305. Dolzhanskaya, N., Merz, G., Aletta, J.M. and Denman, R.B. (2006) Methylation regulates the intracellular protein-protein and protein-RNA interactions of FMRP. *J Cell Sci,* 119, 1933-46.

306. Denman, R.B. (2002) Methylation of the arginine-glycine-rich region in the fragile X mental retardation protein FMRP differentially affects RNA binding. *Cell Mol Biol Lett,* 7, 877-83.

307. Kedersha, N., Cho, M.R., Li, W., Yacono, P.W., Chen, S., Gilks, N., Golan, D.E. and Anderson, P. (2000) Dynamic shuttling of TIA-1 accompanies the recruitment of mRNA to mammalian stress granules. *J Cell Biol,* 151, 1257-68.

308. Liu, G., Grant, W.M., Persky, D., Latham, V.M., Jr., Singer, R.H. and Condeelis, J. (2002) Interactions of elongation factor 1alpha with F-actin and beta-actin mRNA: implications for anchoring mRNA in cell protrusions. *Mol Biol Cell,* 13, 579-92.

309. Yedavalli, V.S., Neuveut, C., Chi, Y.H., Kleiman, L. and Jeang, K.T. (2004) Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function. *Cell,* 119, 381-92.

310. Askjaer, P., Rosendahl, R. and Kjems, J. (2000) Nuclear export of the DEAD box An3 protein by CRM1 is coupled to An3 helicase activity. *J Biol Chem,* 275, 11561-8.

311. Elvira, G., Wasiak, S., Blandford, V., Tong, X.K., Serrano, A., Fan, X., del Rayo Sanchez-Carbente, M., Servant, F., Bell, A.W., Boismenu, D. *et al.* (2006) Characterization of an RNA granule from developing brain. *Mol Cell Proteomics,* 5, 635-51.

312. Zhou, Z., Licklider, L.J., Gygi, S.P. and Reed, R. (2002) Comprehensive proteomic analysis of the human spliceosome. *Nature,* 419, 182-5.

313. Kanai, Y., Dohmae, N. and Hirokawa, N. (2004) Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. *Neuron,* 43, 513-25.

314. Merz, C., Urlaub, H., Will, C.L. and Luhrmann, R. (2007) Protein composition of human mRNPs spliced in vitro and differential requirements for mRNP protein recruitment. *RNA,* 13, 116-28.

315. Deckert, J., Hartmuth, K., Boehringer, D., Behzadnia, N., Will, C.L., Kastner, B., Stark, H., Urlaub, H. and Luhrmann, R. (2006) Protein composition and electron microscopy structure of affinity-purified human spliceosomal B complexes isolated under physiological conditions. *Mol Cell Biol,* 26, 5528-43.

316. You, L.R., Chen, C.M., Yeh, T.S., Tsai, T.Y., Mai, R.T., Lin, C.H. and Lee, Y.H. (1999) Hepatitis C virus core protein interacts with cellular putative RNA helicase. *J Virol,* 73, 2841-53.

317. Shih, J.W., Tsai, T.Y., Chao, C.H. and Wu Lee, Y.H. (2008) Candidate tumor suppressor DDX3 RNA helicase specifically represses cap-dependent translation by acting as an eIF4E inhibitory protein. *Oncogene,* 27, 700-714.

318. Heaton, J.H., Dlakic, W.M., Dlakic, M. and Gelehrter, T.D. (2001) Identification and cDNA cloning of a novel RNA-binding protein that interacts with the cyclic nucleotide-responsive sequence in the Type-1 plasminogen activator inhibitor mRNA. *J Biol Chem,* 276, 3341-7.

319. Yang, L., Embree, L.J., Tsai, S. and Hickstein, D.D. (1998) Oncoprotein TLS interacts with serine-arginine proteins involved in RNA splicing. *J Biol Chem,* 273, 27761-4.

320. Yang, L., Embree, L.J. and Hickstein, D.D. (2000) TLS-ERG leukemia fusion protein inhibits RNA splicing mediated by serine-arginine proteins. *Mol Cell Biol,* 20, 3345-54.

321. Yang, L., Chansky, H.A. and Hickstein, D.D. (2000) EWS.Fli-1 fusion protein interacts with hyperphosphorylated RNA polymerase II and interferes with serine-arginine protein-mediated RNA splicing. *J Biol Chem,* 275, 37612-8.

322. Ohkura, N., Yaguchi, H., Tsukada, T. and Yamaguchi, K. (2002) The EWS/NOR1 fusion gene product gains a novel activity affecting pre-mRNA splicing. *J Biol Chem,* 277, 535-43.

323. Chansky, H.A., Hu, M., Hickstein, D.D. and Yang, L. (2001) Oncogenic TLS/ERG and EWS/Fli-1 fusion proteins inhibit RNA splicing mediated by YB-1 protein. *Cancer Res,* 61, 3586-90.

324. Knoop, L.L. and Baker, S.J. (2001) EWS/FLI alters 5'-splice site selection. *J Biol Chem,* 276, 22317-22.

325. Meissner, M., Lopato, S., Gotzmann, J., Sauermann, G. and Barta, A. (2003) Proto-oncoprotein TLS/FUS is associated to the nuclear matrix and complexed with splicing factors PTB, SRm160, and SR proteins. *Exp Cell Res,* 283, 184-95.

326. Anderson, P. and Kedersha, N. (2006) RNA granules. *J Cell Biol,* 172, 803-8.

327. Brendel, C., Rehbein, M., Kreienkamp, H.J., Buck, F., Richter, D. and Kindler, S. (2004) Characterization of Staufen 1 ribonucleoprotein complexes. *Biochem J,* 384, 239-46.

328. Villace, P., Marion, R.M. and Ortin, J. (2004) The composition of Staufen-containing RNA granules from human cells indicates their role in the regulated transport and translation of messenger RNAs. *Nucleic acids research,* 32, 2411-20.

329. Sossin, W.S. and DesGroseillers, L. (2006) Intracellular trafficking of RNA in neurons. *Traffic (Copenhagen, Denmark),* 7, 1581-9.

330. Arkov, A.L., Wang, J.Y., Ramos, A. and Lehmann, R. (2006) The role of Tudor domains in germline development and polar granule architecture. *Development,* 133, 4053-62.

331. Chuma, S., Hosokawa, M., Kitamura, K., Kasai, S., Fujioka, M., Hiyoshi, M., Takamune, K., Noce, T. and Nakatsuji, N. (2006) Tdrd1/Mtr-1, a tudor-related gene, is essential for male germ-cell differentiation and nuage/germinal granule formation in mice. *Proc Natl Acad Sci U S A,* 103, 15894-9.

332. Hosokawa, M., Shoji, M., Kitamura, K., Tanaka, T., Noce, T., Chuma, S. and Nakatsuji, N. (2007) Tudor-related proteins TDRD1/MTR-1, TDRD6 and TDRD7/TRAP: domain composition, intracellular localization, and function in male germ cells in mice. *Dev Biol,* 301, 38-52.

333. Smith, J.M., Bowles, J., Wilson, M., Teasdale, R.D. and Koopman, P. (2004) Expression of the tudor-related gene Tdrd5 during development of the male germline in mice. *Gene Expr Patterns,* 4, 701-5.

334. Thomson, T. and Lasko, P. (2005) Tudor and its domains: germ cell formation from a Tudor perspective. *Cell Res,* 15, 281-91.

335. Anne, J., Ollo, R., Ephrussi, A. and Mechler, B.M. (2007) Arginine methyltransferase Capsuleen is essential for methylation of spliceosomal Sm proteins and germ cell formation in Drosophila. *Development,* 134, 137-46.

336. Gonsalvez, G.B., Rajendra, T.K., Tian, L. and Matera, A.G. (2006) The Sm-protein methyltransferase, dart5, is essential for germ-cell specification and maintenance. *Curr Biol,* 16, 1077-89.

337. Foresta, C., Ferlin, A. and Moro, E. (2000) Deletion and expression analysis of AZFa genes on the human Y chromosome revealed a major role for DBY in male infertility. *Human molecular genetics,* 9, 1161-9.

338. Kamp, C., Huellen, K., Fernandes, S., Sousa, M., Schlegel, P.N., Mielnik, A., Kleiman, S., Yavetz, H., Krause, W., Kupker, W. *et al.* (2001) High deletion frequency of the complete AZFa sequence in men with Sertoli-cell-only syndrome. *Molecular human reproduction,* 7, 987-94.

339. Foresta, C., Ferlin, A., Garolla, A., Moro, E., Pistorello, M., Barbaux, S. and Rossato, M. (1998) High frequency of well-defined Y-chromosome deletions in idiopathic Sertoli cell-only syndrome. *Human reproduction (Oxford, England),* 13, 302-7.

340. Xia, S.J. and Barr, F.G. (2005) Chromosome translocations in sarcomas and the emergence of oncogenic transcription factors. *Eur J Cancer,* 41, 2513-27.

341. Edmonds, B.T., Wyckoff, J., Yeung, Y.G., Wang, Y., Stanley, E.R., Jones, J., Segall, J. and Condeelis, J. (1996) Elongation factor-1 alpha is an overexpressed actin binding protein in metastatic rat mammary adenocarcinoma. *J Cell Sci,* 109 (Pt 11), 2705-14.

342. Botlagunta, M., Vesuna, F., Mironchik, Y., Raman, A., Lisok, A., Winnard, P., Jr., Mukadam, S., Van Diest, P., Chen, J.H., Farabaugh, P. *et al.* (2008) Oncogenic role of DDX3 in breast cancer biogenesis. *Oncogene*.

343. Lopez de Silanes, I., Lal, A. and Gorospe, M. (2005) HuR: post-transcriptional paths to malignancy. *RNA Biol,* 2, 11-3.

344. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. and Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene,* 77, 51-9.

345. Chen, T., Damaj, B.B., Herrera, C., Lasko, P. and Richard, S. (1997) Self-association of the single-KH-domain family members Sam68, GRP33, GLD-1, and Qk1: role of the KH domain. *Mol Cell Biol,* 17, 5707-18.

346. Herrmann, F. and Fackelmayer, F.O. (2009) Nucleo-cytoplasmic shuttling of protein arginine methyltransferase 1 (PRMT1) requires enzymatic activity. *Genes Cells,* 14, 309-17.

347. Aggarwal, P., Vaites, L.P., Kim, J.K., Mellert, H., Gurung, B., Nakagawa, H., Herlyn, M., Hua, X., Rustgi, A.K., McMahon, S.B. *et al.* (2010) Nuclear cyclin D1/CDK4 kinase regulates CUL4 expression and triggers neoplastic growth via activation of the PRMT5 methyltransferase. *Cancer cell,* 18, 329-40.

348. Jansson, M., Durant, S.T., Cho, E.C., Sheahan, S., Edelmann, M., Kessler, B. and La Thangue, N.B. (2008) Arginine methylation regulates the p53 response. *Nature cell biology,* 10, 1431-9.

349. Pal, S., Baiocchi, R.A., Byrd, J.C., Grever, M.R., Jacob, S.T. and Sif, S. (2007) Low levels of miR-92b/96 induce PRMT5 translation and H3R8/H4R3 methylation in mantle cell lymphoma. *The EMBO journal,* 26, 3558-69.

350. Scoumanne, A., Zhang, J. and Chen, X. (2009) PRMT5 is required for cell-cycle progression and p53 tumor suppressor function. *Nucleic acids research,* 37, 4965-76.

351. Dobrzycka, K.M., Townson, S.M., Jiang, S. and Oesterreich, S. (2003) Estrogen receptor corepressors -- a role in human breast cancer? *Endocrine-related cancer,* 10, 517-36.

352. Skotheim, R.I. and Nees, M. (2007) Alternative splicing in cancer: noise, functional, or systematic? *The international journal of biochemistry & cell biology,* 39, 1432-49.

353. Mathioudaki, K., Papadokostopoulou, A., Scorilas, A., Xynopoulos, D., Agnanti, N. and Talieri, M. (2008) The PRMT1 gene expression pattern in colon cancer. *British journal of cancer,* 99, 2094-9.

354. Papadokostopoulou, A., Mathioudaki, K., Scorilas, A., Xynopoulos, D., Ardavanis, A., Kouroumalis, E. and Talieri, M. (2009) Colon cancer and protein arginine methyltransferase 1 gene expression. *Anticancer research,* 29, 1361-6.

355. Kim, Y.R., Lee, B.K., Park, R.Y., Nguyen, N.T., Bae, J.A., Kwon, D.D. and Jung, C. (2010) Differential CARM1 expression in prostate and colorectal cancers. *BMC cancer,* 10, 197.

356. Yoshimatsu, M., Toyokawa, G., Hayami, S., Unoki, M., Tsunoda, T., Field, H.I., Kelly, J.D., Neal, D.E., Maehara, Y., Ponder, B.A. *et al.* (2011) Dysregulation of PRMT1 and PRMT6, Type I arginine methyltransferases, is involved in various types of human cancers. *International journal of cancer,* 128, 562-73.

357. Sims, R.J., 3rd, Rojas, L.A., Beck, D., 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.

358. Muratani, M. and Tansey, W.P. (2003) How the ubiquitin-proteasome system controls transcription. *Nature reviews,* 4, 192-201.

359. De Leeuw, F., Zhang, T., Wauquier, C., Huez, G., Kruys, V. and Gueydan, C. (2007) The cold-inducible RNA-binding protein migrates from the nucleus to cytoplasmic stress granules by a methylation-dependent mechanism and acts as a translational repressor. *Exp Cell Res,* 313, 4130-44.

360. Takbum, O. and Anderson, P. (2010) The role of posttranslational modification in the assembly of stress granules. *Wiley Interdisciplinary Reviews: RNA,* 1, 486-93.

361. Lai, M.C., Lee, Y.H. and Tarn, W.Y. (2008) The DEAD-box RNA helicase DDX3 associates with export messenger ribonucleoproteins as well as tip-associated protein and participates in translational control. *Mol Biol Cell,* 19, 3847-58.

362. Andersson, M.K., Stahlberg, A., Arvidsson, Y., Olofsson, A., Semb, H., Stenman, G., Nilsson, O. and Aman, P. (2008) The multifunctional FUS, EWS and TAF15 proto-oncoproteins show cell type-specific expression patterns and involvement in cell spreading and stress response. *BMC cell biology,* 9, 37.

363. Goodier, J.L., Zhang, L., Vetter, M.R. and Kazazian, H.H., Jr. (2007) LINE-1 ORF1 protein localizes in stress granules with other RNA-binding proteins, including components of RNA interference RNA-induced silencing complex. *Mol Cell Biol,* 27, 6469-83.

364. Bhattacharyya, R.P., Remenyi, A., Yeh, B.J. and Lim, W.A. (2006) Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. *Annual review of biochemistry,* 75, 655-80.

365. Koensgen, D., Mustea, A., Klaman, I., Sun, P., Zafrakas, M., Lichtenegger, W., Denkert, C., Dahl, E. and Sehouli, J. (2007) Expression analysis and RNA localization of PAI-RBP1 (SERBP1) in epithelial ovarian cancer: association with tumor progression. *Gynecologic oncology,* 107, 266-73.

366. Fournier, M.J., Gareau, C. and Mazroui, R. (2010) The chemotherapeutic agent bortezomib induces the formation of stress granules. *Cancer cell international,* 10, 12.

367. Raguz, S. and Yague, E. (2008) Resistance to chemotherapy: new treatments and novel insights into an old problem. *British journal of cancer,* 99, 387-91.

368. Kwon, S., Zhang, Y. and Matthias, P. (2007) The deacetylase HDAC6 is a novel critical component of stress granules involved in the stress response. *Genes Dev,* 21, 3381-94.

369. Moeller, B.J., Cao, Y., Li, C.Y. and Dewhirst, M.W. (2004) Radiation activates HIF-1 to regulate vascular radiosensitivity in tumors: role of reoxygenation, free radicals, and stress granules. *Cancer cell,* 5, 429-41.

370. Arimoto, K., Fukuda, H., Imajoh-Ohmi, S., Saito, H. and Takekawa, M. (2008) Formation of stress granules inhibits apoptosis by suppressing stress-responsive MAPK pathways. *Nature cell biology,* 10, 1324-32.

371. Mazroui, R., Di Marco, S., Kaufman, R.J. and Gallouzi, I.E. (2007) Inhibition of the ubiquitin-proteasome system induces stress granule formation. *Mol Biol Cell,* 18, 2603-18.

372. Eisinger-Mathason, T.S., Andrade, J., Groehler, A.L., Clark, D.E., Muratore-Schroeder, T.L., Pasic, L., Smith, J.A., Shabanowitz, J., Hunt, D.F., Macara, I.G. *et al.* (2008) Codependent functions of RSK2 and the apoptosis-promoting factor TIA-1 in stress granule assembly and cell survival. *Molecular cell,* 31, 722-36.

373. Baguet, A., Degot, S., Cougot, N., Bertrand, E., Chenard, M.P., Wendling, C., Kessler, P., Le Hir, H., Rio, M.C. and Tomasetto, C. (2007) The exon-junction-complex-component metastatic lymph node 51 functions in stress-granule assembly. *J Cell Sci,* 120, 2774-84.

APPENDIX 1

Protein Arginine Methylation Facilitates Cotranscriptional Recruitment of pre-mRNA Splicing Factors

*Yin-Chu Chen[1], Eric J. Milliman[1], Isabelle Goulet[2], Jocelyn Côté[2], Christopher A. Jackson[1], Jennifer A. Vollbracht[1], and Micheal C. Yu[1]**

[1]Department of Biological Sciences, State University of New York at Buffalo, Buffalo, New York.
[2]Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario, Canada

*Corresponding author. Mailing address: Department of Biological Sciences, State University of New York at Buffalo, 109 Cooke Hall, Buffalo, NY 14260. Phone: (716) 645-4931. Fax: (716) 645-2975. E-mail: mcyu@buffalo.edu

Received 26 March 2010/ Returned for modification 3 May 2010/ Accepted 17 August 2010.

This research was originally published in *Molecular and Cellular Biology.* 2010;

30: 5245-56. © 2010, American Society for Microbiology.

Description and statement of contributions of collaborators and co-authors

This manuscript describes the regulation by arginine methylation of the co-transcriptional recruitment of splicing factors to promote target specificity in pre-mRNA splicing in *Saccharomyces cerevisiae*.

I. Goulet contributed to this manuscript by performing experiments linked to the findings presented in Figs. 1 and S1.

MOLECULAR AND CELLULAR BIOLOGY, Nov. 2010, p. 5245–5256 Vol. 30, No. 21
0270-7306/10/$12.00 doi:10.1128/MCB.00359-10

Protein Arginine Methylation Facilitates Cotranscriptional Recruitment of Pre-mRNA Splicing Factors▽#

Yin-Chu Chen,[1] Eric J. Milliman,[1] Isabelle Goulet,[2] Jocelyn Côté,[2] Christopher A. Jackson,[1] Jennifer A. Vollbracht,[1] and Michael C. Yu[1]*

Department of Biological Sciences, State University of New York at Buffalo, Buffalo, New York,[1] and Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario, Canada[2]

Received 26 March 2010/Returned for modification 3 May 2010/Accepted 17 August 2010

Cotranscriptional recruitment of pre-mRNA splicing factors to their genomic targets facilitates efficient and ordered assembly of a mature messenger ribonucleoprotein particle (mRNP). However, how the cotranscriptional recruitment of splicing factors is regulated remains largely unknown. Here, we demonstrate that protein arginine methylation plays a novel role in regulating this process in *Saccharomyces cerevisiae*. Our data show that Hmt1, the major type I arginine methyltransferase, methylates Snp1, a U1 small nuclear RNP (snRNP)-specific protein, and that the mammalian Snp1 homolog, U1-70K, is likewise arginine methylated. Genome-wide localization analysis reveals that the deletion of the *HMT1* gene deregulates the recruitment of U1 snRNP and its associated components to intron-containing genes (ICGs). In the same context, splicing factors acting downstream of U1 snRNP addition bind to a reduced number of ICGs. Quantitative measurement of the abundance of spliced target transcripts shows that these changes in recruitment result in an increase in the splicing efficiency of developmentally regulated mRNAs. We also show that in the absence of either Hmt1 or of its catalytic activity, an association between Snp1 and the SR-like protein Npl3 is substantially increased. Together, these data support a model whereby arginine methylation modulates dynamic associations between SR-like protein and pre-mRNA splicing factor to promote target specificity in splicing.

In eukaryotic cells, pre-mRNA is processed and packaged into a mature messenger ribonucleoprotein particle (mRNP) prior to its export from the nucleus (reviewed in references 12, 25, and 44). The correct formation of an mRNP requires a web of physical interactions among RNA processing factors during transcription. An important step in the processing of eukaryotic RNA is pre-mRNA splicing, in which noncoding introns are removed to generate mature, translatable mRNAs. The splicing reaction is catalyzed by the spliceosome, which is composed of five small nuclear ribonucleoprotein particles (snRNPs) and many associated proteins (reviewed in references 62, 67, and 68). Like many other RNA processing factors that have been studied thus far, the components of the spliceosome are recruited cotranscriptionally (19, 36, 47). Chromatin immunoprecipitation (ChIP) experiments have shown that, *in vivo*, spliceosome components assemble on intron-containing genes (ICGs) in a stepwise manner, consistent with findings from *in vitro* studies of splicing complexes (19, 32, 47). Specifically, the U1 snRNP is recruited to the 5′ splice site (ss), and the branchpoint binding protein (BPP) and Mud2 (human U2AF65) are recruited to the intronic branch site and nearby sequences, respectively. Together, these factors define basic intron/exon consensus features and "commit" a pre-mRNA substrate to splicing. Subsequent assembly involves ordered recruitment of the U2 snRNP, the U5/U4/U6 tri-snRNP, and spliceosome activation factors such as the "nineteen complex" (NTC) (9). Posttranscriptional splicing can occur, both *in vivo* and *in vitro* (64), but the coupling of splicing to transcription is thought to maximize the fidelity and efficiency of the process (13, 26). Thus, differential cotranscriptional recruitment of splicing factors represents a mechanism by which splicing can be regulated.

Protein arginine methylation is a posttranslational modification that is common to many RNA-binding proteins (RBPs) (reviewed in references 3 and 4). The enzymes that catalyze this process are termed protein arginine methyltransferases (PRMTs). In both yeast and mammalian cells, heterogeneous nuclear ribonucleoproteins (hnRNPs)—which like the snRNPs are associated with pre-mRNAs and involved in mRNA biogenesis—are major substrates of the PRMTs. Methylated hnRNPs possess at least one N-terminal RNA recognition motif (RRM)-type RNA-binding motif in addition to C-terminal arginine-glycine-glycine (RGG)-rich repeats, where arginine methylation is often found (42).

Many studies have demonstrated that arginine methylation plays an important role in modulating protein-protein interactions. For example, arginine methylation of the mammalian transcriptional elongation factor Spt5 regulates its interaction with RNA polymerase II (Pol II), thereby affecting transcription at a global level (35). The loss of arginine methylation on the mammalian STAT1 protein prohibits its association with PIAS, the inhibitor of STAT1, resulting in a decreased STAT1-mediated interferon response (49). In some cases, arginine methylation of a specific factor can modulate subsequent posttranslational modification events. For example, arginine methylation of mammalian FOXO transcription factors inhibits their phosphorylation by Akt (70).

* Corresponding author. Mailing address: Department of Biological Sciences, State University of New York at Buffalo, 109 Cooke Hall, Buffalo, NY 14260. Phone: (716) 645-4931. Fax: (716) 645-2975. E-mail: mcyu@buffalo.edu.

Supplemental material for this article may be found at http://mcb.asm.org/.

▽ Published ahead of print on 7 September 2010.

In *Saccharomyces cerevisiae*, the arginine methyltransferase that catalyzes the formation of most asymmetric dimethylarginines is called Hmt1 (also known as Rmt1) (17, 24). Hmt1 is important for the nuclear transport of several of the hnRNPs that participate in the mRNP biogenesis pathway (58) and for the biochemical association between Npl3 and Tho2 (71). Npl3, an Hmt1 substrate, is an RNA-binding protein that has characteristics of both SR family (15) and hnRNP-like family proteins (8, 57); it is important for mRNA export (29) and also plays a role in pre-mRNA splicing (33). Tho2 is a component of the evolutionarily conserved transcription/export) (TREX) complex and plays a role during the elongation phase of transcription (52). RNA *in situ* hybridization analysis has demonstrated that the loss of Hmt1 activity results in slowed release of the *HSP104* mRNA from the site of transcription (71). Genome-wide localization analysis (also known as ChIP-chip) revealed that Hmt1 coordinates cotranscriptional assembly of the hnRNPs Nab2, Hrp1, and Yra1 (71). Together, these data demonstrated that Hmt1 promotes the dynamic interactions between RNA-binding proteins and pre-mRNAs that facilitate proper assembly of an mRNP during transcription. However, the previous studies left open the question of whether the catalytic activity of Hmt1 or some other feature of this protein is required for the regulation of the cotranscriptional recruitment of pre-mRNA splicing factors.

In this study, we demonstrate that Hmt1-catalyzed arginine methylation, specifically, facilitates the cotranscriptional recruitment of pre-mRNA splicing factors. Also, we identify Snp1, a component of the U1 snRNP, as a novel Hmt1 substrate. We further show that arginine methylation of Snp1 is evolutionarily conserved in higher eukaryotes. Using ChIP-chip, we determined the impact of protein arginine methylation on the cotranscriptional recruitment of splicing factors to ICGs; specifically, we show that in mutants lacking Hmt1 or its catalytic activity, the recruitment of splicing factors to genes that undergo regulated splicing is aberrant, resulting in the enhanced splicing of these transcripts. Finally, we show that arginine methylation regulates the association of Snp1 with the SR-like protein Npl3. Together, these results demonstrate that protein arginine methylation regulates the specificity of the splicing machinery targeting to particular pre-mRNAs.

MATERIALS AND METHODS

Yeast strains used in this study. All yeast strains used are listed in Table S1 in the supplemental material. Cells were grown at 30°C on YEPD medium (1% yeast extract, 2% Bacto peptone, 2% [wt/vol] D-glucose) unless otherwise stated. Gene deletions and C-terminal epitope tags were generated as described previously (30, 43).

***In vitro* methylation of Snp1.** The *in vitro* methylation was performed as described previously (46) using recombinant Hmt1 and tandem affinity purification (TAP)-tagged Snp1 expressed in Δ*hmt1* cells. The TAP was carried out as described previously (55), with the exception of binding with IgG-Sepharose beads (GE Amersham) in 600 mM NaCl during the first affinity purification step to minimize the capturing of interactors. The entire *in vitro* methylation reaction was resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using a NuPAGE 4 to 12% Bis-Tris gel (Invitrogen), followed by fluorography as previously described (46). In parallel, the purified TAP-tagged Snp1 was subjected to immunoblotting using an anti-CBP (Open Biosystems) antibody.

Identification of *in vivo* methylation sites within U1-70K. Monoclonal antibodies against U1-70K were used to immunoprecipitate endogenous U1-70K from HeLa cells grown in medium specifically designed for SILAC experiments (66). The immunoprecipitates were resolved using a NuPAGE 12% Bis-Tris gel (Invitrogen) and Coomassie blue stained. The major band slice at 70 kDa was excised and subjected to preparation for mass spectrometry analysis as previously described (63, 66). The liquid chromatography/tandem mass spectrometry (LC/MS-MS) analysis was carried out as previously described (66).

Snp1 localization studies. Wild-type (WT) or *hmt1*-null yeast strains containing green fluorescent protein (GFP)-tagged Snp1 from the GFP-tagged yeast library (27) were grown in synthetic complete (SC) medium until mid-log phase, at which time DAPI (4′,6-diamidino-2-phenylindole) stain was added to the medium for a final concentration of 1.2 μg/ml. Cells were then incubated for an additional 30 min. The differential interference contrast (DIC) and fluorescence microscopies were performed using an Axioplan 2 fluorescent microscope (Zeiss). Digital images were captured using the AxioCam MRm camera (Zeiss). Image acquisition and analysis were carried out using AxioVision 4.4 software (Zeiss).

Npl3-Snp1 interaction studies. Yeast strains containing TAP-tagged Snp1 were grown in 500 ml of YEPD medium to mid-log phase and lysed with glass beads and a FastPrep machine (MP Biomedical) with a GigaPrep adapter five times at a setting of 6.5 for 30 s at 4°C. The cells were lysed in PBSMT buffer (phosphate-buffered saline with 0.5% Triton X-100, 3 mM KCl, and 2.5 mM $MgCl_2$) supplemented with protease inhibitors (1 mM phenylmethylsulfonyl fluoride, 1.3 mM benzamide, and 2.5 μg/ml each leupeptin, chymostatin, antipain, pepstatin A, and aprotinin). Lysates were normalized for total protein, and TAP-tagged Snp1 fusion protein was purified by 2 h of incubation at 4°C with a 25-μl packed volume of IgG-Sepharose beads that had been prepared as previously described (71). Beads were washed three times with 1 ml of lysis buffer and twice with 1× TEV cleavage buffer (25 mM Tris-HCl [pH 8], 150 mM NaCl, 0.1% NP-40, 0.5 mM EDTA, 0.1 mM dithiothreitol [DTT]) prior to TEV cleavage with 1 μl of ProTEV protease (Promega) for 1 h at 16°C. The cleaved samples were eluted from the beads by a MolBicol column and precipitated with 12.5% trichloroacetic acid. Protein samples were dried, resuspended in protein sample buffer, resolved by SDS-PAGE using a 4 to 12% NuPAGE Bis-Tris gel (Invitrogen), and immunoblotted using anti-Npl3 antibody (gift from Pam Silver) (8). The immunoblot was then stripped and reprobed with an anti-CBP antibody (Open Biosystems). The immunoblot signals were digitally captured by a Bio-Rad ChemiDoc system.

Genome-wide localization (ChIP-chip) studies of splicing factors. ChIP-chip studies were performed essentially as described previously (71). All of the experiments were performed with biological triplicates. The data obtained were analyzed using the Rosetta Resolver bioinformatics platform. Only genes with P values of <0.02 were considered for the final data analysis. Genes were considered bound if the immunoprecipitated fraction (IP)/background (whole-cell extract [WCE]) ratio was greater than 1. All of the raw output data from Rosetta Resolver and corrected data are available in Files S1 to S5 in the supplemental material.

Calculation of the intron distribution index (IDI). For each ChIP-chip data set, a ranked list for total genes bound by a splicing factor was generated based on the IP/WCE ratio as described previously (71). Any ICG present within this ranked list is identified, and its distribution across the entire list is graphically plotted (see Fig. S2A to C in the supplemental material). Based on the total number of ICGs identified, a model distribution was generated to reflect a scenario in which the same numbers of ICGs are evenly distributed across a profile, yielding a straight line at a 45° angle (see Fig. S2A to C in the supplemental material). By plotting the actual experimental ICG ranking data as the *x*-axis coordinates with the model ICG ranking data as the *y*-axis coordinates, a nonlinear binding relationship was observed in which the line was skewed away from the assumed evenly distributed binding event. We then quantified the distance between this plot and the assumed model plot to measure the changes in the overall intron distribution between the experimental data and the model distribution data using the least-mean square distance (D_{rand}) calculated by the following equation:

$$D_{\text{rand}} = \sqrt{\frac{\sum_i \frac{(X_i - Y_i)^2}{2}}{N}}$$

In this equation, X_i represents the ranking of each ICG in the data set being tested and Y_i represents the corresponding ranking within the model reference. N is the number of ICGs shared between the two profiles compared (such as the WT versus Δ*hmt1* or the WT versus *hmt1-G68R*).

Directed ChIP experiments. Directed ChIP experiments were performed as described previously (41, 71). For each immunoprecipitation, monoclonal anti-Myc (9E11; Fisher) antibody was precoupled to protein A-Sepharose beads. The

oligonucleotides used in the ChIP experiments are listed in Table S2 in the supplemental material.

RNA isolation, cDNA preparation, and qPCR. Total RNA was isolated by using the hot phenol method as described previously (71) followed by a DNase I (Invitrogen) treatment as per the manufacturer's instruction. The DNase I-treated RNA was then cleaned using an RNeasy kit (Qiagen). First-strand cDNA was synthesized from poly(A)-RNA using Superscript III reverse transcriptase (RT) (Invitrogen), oligo(dT), and 10 μg of RNA as described previously (47). cDNA dilution series were analyzed by PCR (24 to 26 cycles) to establish a quantitative, linear range of cDNA input for each set of quantitative PCR (qPCR) performed. The procedure for qPCR and the primer sets used were described previously (47), except [α-^{33}P]dCTPs were used to radiolabel qPCR products. The radiolabeled PCR products were then resolved using an 8% Tris-borate-EDTA (TBE) gel, dried, and exposed to a PhosphorImager screen. The signals from spliced and unspliced PCR products were quantified using a PhosphorImager (Molecular Dynamics). The average ratios were calculated from triplicate sets of biological duplicates for *HOP2* and triplicate sets of biological triplicates for *REC107*. Each RNA sample that had been treated by DNase I but was lacking the RT step was tested first by PCR using *ADH3* primers to determine any potential genomic DNA contamination.

RESULTS

The U1 snRNP components Snp1 and U1-70K are arginine methylated. Recent mammalian studies have implicated protein arginine methylation in pre-mRNA splicing (6, 11, 50). Snp1, a component of the *S. cerevisiae* U1 snRNP, has previously been shown to interact with the arginine methyltransferase Hmt1 in a comprehensive two-hybrid analysis (28). Analysis of the amino acid sequence of Snp1 revealed a number of arginine-glycine (RG) motifs that could potentially serve as sites for arginine methylation (see Fig. S1B, Snp1 sequence, in the supplemental material). We next set out to determine whether Snp1 is a substrate for Hmt1. We used a modified tandem affinity purification method to purify TAP-tagged Snp1 expressed from Δ*hmt1* cells under native, high-salt conditions. We then used this substrate, which has no endogenous methyl marks, in an *in vitro* methylation assay using recombinant Hmt1 as described previously (46).

The fluorograph from the *in vitro* methylation assay revealed a strong, positive signal corresponding to a band migrating at approximately 41 kDa (Fig. 1, left, arrow). Immunoblotting of the purified Snp1-CBP using an anti-CBP antibody (Fig. 1, right, arrow) confirmed that the 41-kDa band observed in the fluorograph is Snp1-CBP. Recombinant glutathione *S*-transferase (GST)-tagged Rps2 was used as a positive control (Fig. 1, asterisk), as described previously (51).

Hmt1 has been implicated in controlling nuclear localization of hnRNP-like proteins (20, 58). To determine if Hmt1 loss perturbs the cellular localization of Snp1, we chromosomally expressed a C-terminal GFP-tagged Snp1 in both wild-type and Δ*hmt1* cells and compared its *in vivo* cellular localization in these two cell types (see Fig. S1A in the supplemental material). In wild-type cells, Snp1-GFP was enriched in the nucleus (see Fig. S1A, WT panel, in the supplemental material), consistent with a previous report from a large-scale protein localization study in yeast (27). The same was true in the Δ*hmt1* cells (see Fig. S1A, Δ*hmt1* panel, in the supplemental material). Thus, Hmt1 loss has no effect on the cellular localization of Snp1.

The mammalian homolog of Snp1 is U1-70K (60), which was previously identified in a proteomic analysis of arginine-methylated protein complexes in HeLa cells (7). Although U1-70K

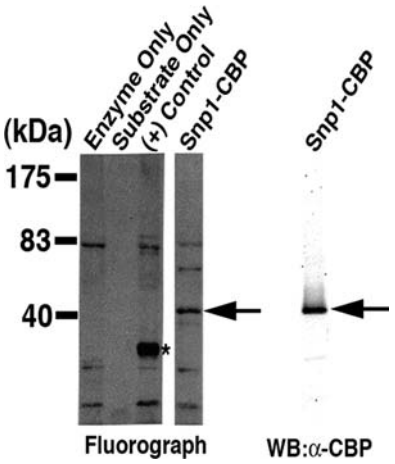


FIG. 1. Arginine methylation is a feature of the yeast splicing factor Snp1. TAP-tagged Snp1 purified from the Δ*hmt1* strain was subjected to an *in vitro* methylation assay using recombinant Hmt1 and [*methyl*-^{3}H]SAM. (Left) The full protein complement in each reaction was resolved on a 4 to 12% gel by SDS-PAGE and visualized by fluorography; (right) the substrate used in the *in vitro* methylation reaction was resolved in parallel, also on a 4 to 12% gel, and immunoblotted with an anti-CBP antibody. A band corresponding to *methyl*-^{3}H-labeled Snp1-CBP was observed only in the presence of Hmt1 (Snp1-CBP lane, arrow [middle], compared to substrate only lane [left]). Recombinant GST-tagged Rps2 served as a control [left, (+) control lane, asterisk]. Immunoblotting carried out with an anti-CBP antibody (right) confirmed that the methylated band is Snp1-CBP (middle, arrow).

was isolated using antibodies specific for symmetrical dimethylation (SYM10 and SYM11) (7), these antibodies strongly recognize core Sm proteins. Thus, the presence of U1-70K in those immunoprecipitates could have resulted from coisolation with native U1 snRNP particles. The presence of conserved RG motifs in the human U1-70K and the consensus sequences (see Fig. S1B, U1-70K and consensus sequence, in the supplemental material) suggests that these proteins may harbor modified arginine residues, which could account for their recognition by methylation-specific antibodies. We tested this possibility using a mass spectrometry approach and discovered that methylated arginines are present on U1-70K immunoprecipitated from HeLa cells (using a monoclonal anti-U1-70K antibody). Our mass spectrometry results indicated that 56% of the U1-70K protein was covered in the analysis, and the analysis of b and y ion series spectra indicated that 10 arginine residues (see Fig. S1B, asterisks, in the supplemental material) in the area covered were either mono- or dimethylated. Among the identified methylated arginines, many were found to match the RG and RXR motifs typical of most PRMT preferred sites (see Fig. S1B in the supplemental material). In summary, our data show that Snp1 is an *in vitro* substrate of Hmt1 and that its mammalian homolog, U1-70K, is likewise arginine methylated.

Hmt1 modulates the cotranscriptional recruitment of Snp1. Since Hmt1 has been shown to signal the cotranscriptional recruitment of mRNP components (71) and U1 snRNP proteins are cotranscriptionally recruited (32, 36), we wanted to examine whether Hmt1 is necessary for cotranscriptional recruitment of Snp1 to its targets throughout the genome. To this end, we performed genome-wide occupancy analysis (ChIP-

chip) of Snp1 in wild-type and Δ*hmt1* cells. A list of the relative rankings of all ICGs bound by Snp1 and other splicing factors examined in this study, in both wild-type and Δ*hmt1* cells, is presented in File S4 in the supplemental material. To probe for methylation-specific changes within the Snp1 ChIP-chip profiles, we used two complementary, quantitative analytical methods. First, we calculated the percentage of total Snp1-bound genes that were intron-containing genes in order to determine whether the Hmt1 loss leads to an increase or decrease in the percentage of ICGs bound within a profile. However, because a factor may bind to the same number and complement of ICGs yet the relative rankings of these genes may differ depending on the status of Hmt1, we also devised an intron distribution index (IDI) that assesses cumulative changes in the distribution of binding by Snp1 in a global manner (see Materials and Methods and Fig. S2 in the supplemental material). We calculated the IDI for each binding profile to distinguish changes in the relative rank of ICGs compared to that of other bound genes as a consequence of changes to Hmt1. A lower IDI in the Hmt1 mutant than in wild-type cells would reflect an overall decrease in the relative occupancy of the bound ICGs in the Hmt1 mutants. For example, although the same population of ICGs would be bound by a given factor, the factor's relative affinity for ICGs would be lower in Hmt1 mutants. Thus, using both approaches to analyze our Snp1 ChIP-chip data provides us with two separate but complementary methods to determine how methylation affects the cotranscriptional recruitment of Snp1.

Our Snp1 ChIP-chip data revealed that this protein binds a higher percentage of ICGs in Δ*hmt1* cells than in wild-type cells (Fig. 2A, left). However, this is mostly based on the observation that Snp1 binding to non-ICGs was generally reduced in the Δ*hmt1* cells (Fig. 2A, middle); the total number of ICGs bound increased only slightly in the Δ*hmt1* cells (Fig. 2A, right). In sum, Snp1 displays less affinity for non-ICGs in the absence of Hmt1.

In the absence of Hmt1, the recruitment of U1 snRNP-associated proteins to ICGs is increased. To determine if Hmt1 affects the recruitment of other factors that associate with U1 snRNP, we performed ChIP-chip on the U1 snRNP-associated protein Prp40 and the pre-mRNA–U1 snRNP complex component Mud2. Prp40 interacts with the transcriptional machinery (48) and is considered one of the earliest splicing factors to be cotranscriptionally recruited. Thus, Prp40 has been proposed to sample transcripts for introns and to be specifically retained at ICGs. Similar to the trend observed for Snp1, Prp40 bound to a slightly higher percentage of ICGs in Δ*hmt1* cells than in wild-type cells (Fig. 2B, left) and to a slightly lower number of both non-ICGs and ICGs in total (Fig. 2B, middle).

Mud2 associates with the 3′ end of an intron (2) and interacts with the U2 snRNP by recognizing the branchpoint sequence (1). Our ChIP-chip data show that Mud2 is bound to a higher percentage of ICGs in Δ*hmt1* cells than in wild-type cells (Fig. 2C, left). As is the case for Snp1, this is based on the fact that a significantly lower overall number of non-ICGs bound (Fig. 2C, middle), despite only a slight decrease in the total number of ICGs bound (Fig. 2C, right). This overall decrease in the number of non-ICGs bound suggests that

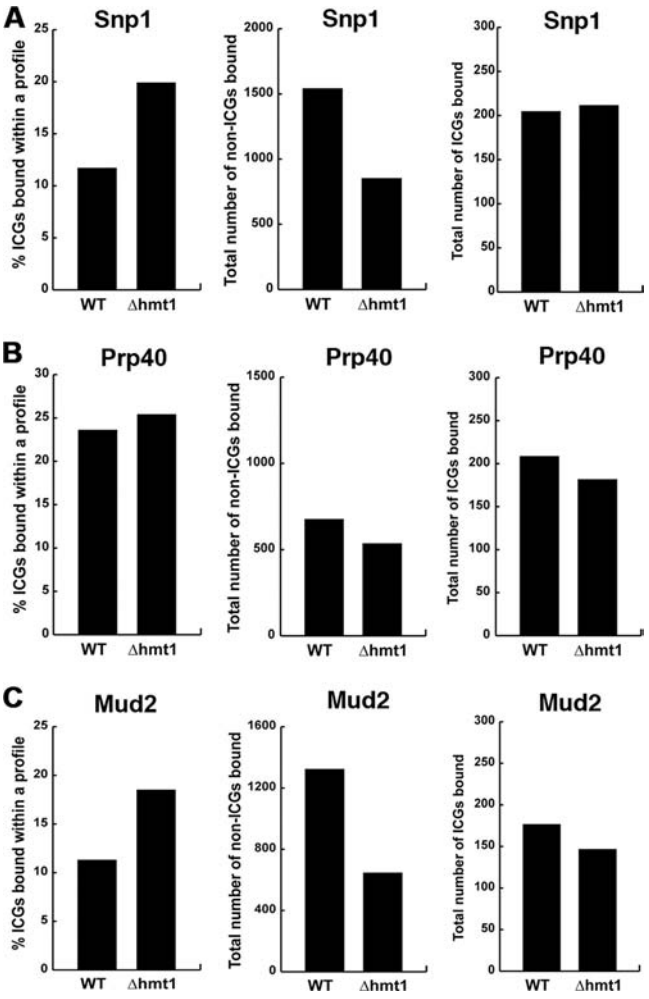


FIG. 2. Hmt1 influences the cotranscriptional recruitment of U1 snRNP and its associated proteins. ChIP-chip was used to assess the effects of Hmt1 on the cotranscriptional recruitment of U1 snRNP and its associated proteins, which are Snp1 (A), Prp40 (B), and Mud2 (C). The impact of Hmt1 was assessed as the change in the percentage of intron-containing genes (ICGs) within the total number of genes bound by a splicing factor, as denoted by the bar graphs on the left. The total numbers of non-ICGs bound by each factor are shown by the middle bar graphs, and the total numbers of ICGs bound by each splicing factor are shown by the bar graphs on the right. The wild-type ChIP-chip data for Prp40 and Mud2 were obtained from Moore et al. (47).

Mud2 has a significantly lower relative affinity for non-ICGs in the absence of Hmt1.

Together, the Prp40 and Mud2 data suggest that Hmt1 prevents Mud2 from being recruited to additional spurious non-ICGs and that this is also the case for Prp40, but to a lesser degree.

In the absence of Hmt1 activity, the cotranscriptional recruitment of splicing factors after U1 snRNP addition is decreased. Our data show that Hmt1 affects the cotranscriptional recruitment of U1 snRNP and its associated components, and we wanted to determine if Hmt1 also affects splicing factors involved in other major steps of the canonical spliceosome assembly pathway. Therefore, we extended this approach to investigation of the splicing factors Prp11, Prp5, Prp2, and

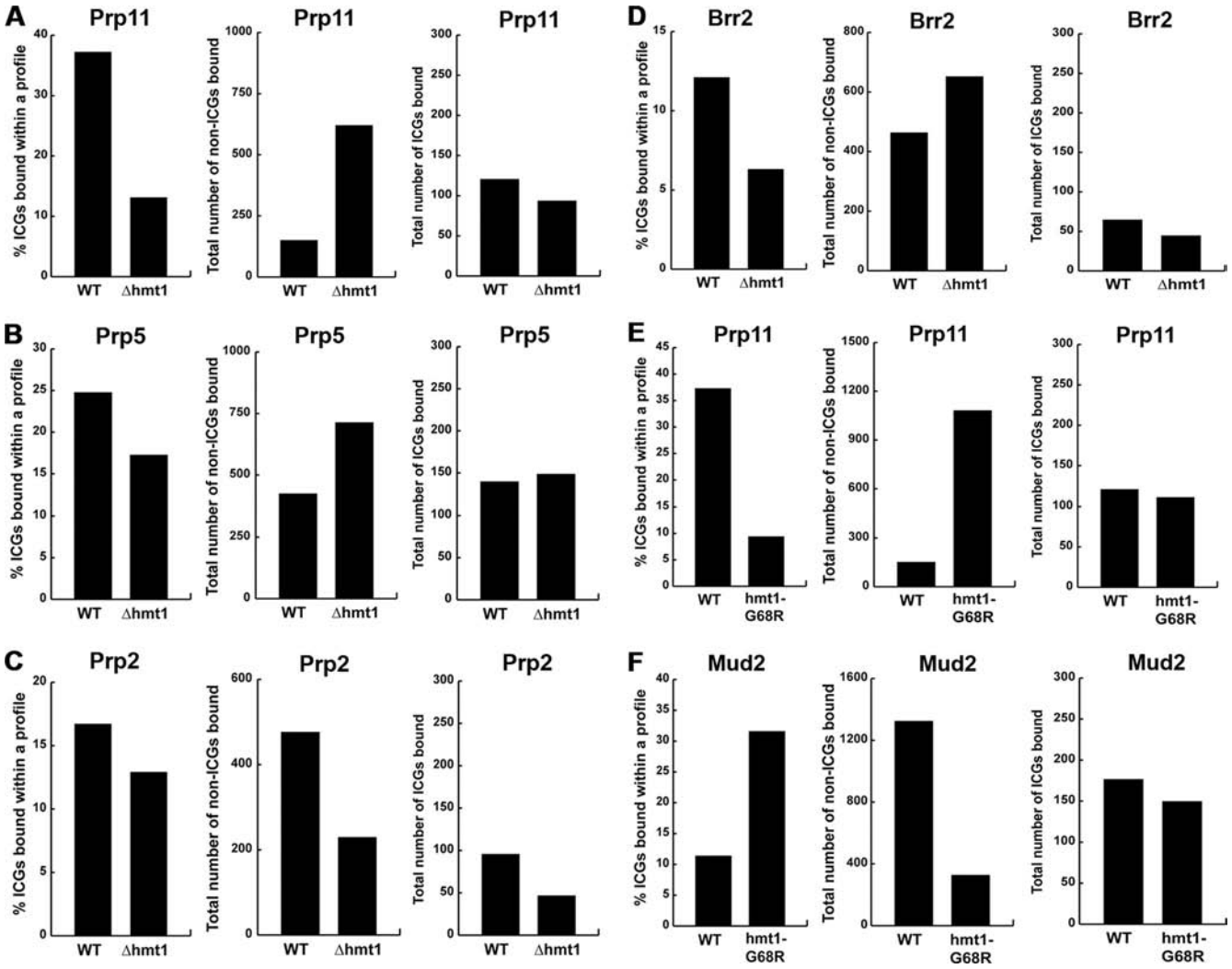


FIG. 3. Arginine methylation negatively affects the recruitment of U2 and U4/U5/U6 tri-snRNP components and their associated proteins. The genome-wide cotranscriptional recruitment of Prp11 (A), Prp5 (B), Prp2 (C), and Brr2 (D) in the wild-type and Δ*hmt1* strains is mapped as described in the legend to Fig. 2. The requirement for Hmt1 activity (versus Hmt1 presence) in cotranscriptional recruitment was tested for Prp11 (E) and Mud2 (F). The wild-type ChIP-chip data for Prp11, Prp5, Prp2, and Brr2 were obtained from Moore et al. (47).

Brr2. Prp11 is a component of the U2 snRNP and is required for binding of the U2 snRNP to pre-mRNA during spliceosome assembly (56). Prp5 bridges the U1 and U2 snRNPs to allow U2 snRNP to stably associate with introns (31). Prp2 is an RNA-dependent ATPase of the DEAH-box family and is required for the first transesterification reaction of splicing (59). Brr2 is an RNA helicase; it associates with the U5 snRNP (23) and is required for activating the spliceosome for catalysis by disrupting U4/U6 base pairing in native snRNPs (39, 53).

In comparison to the U1 snRNP and the associated components that have been examined, all four of the U2 and U4/U5/U6 components and their associated factors bound to a lower percentage of ICGs in the absence of Hmt1 (Fig. 3, left panels). In the case of Prp11, this result was particularly striking, with ICG binding reduced to only 13% in the Δ*hmt1* mutant (Fig. 3A, left). Moreover, for each of these proteins other than Prp2, the total number of non-ICGs bound was increased (Fig. 3, middle panels), whereas the total number of ICGs bound was generally decreased, with the only exception being Prp5 (Fig. 3, right panels). Nevertheless, we observed reduced IDIs for all four components in the context of Δ*hmt1* cells, with the reduction being particularly drastic in the cases of both Prp11 and Brr2 (see Fig. S2D, panels Prp11, Prp2, Prp5, and Brr2, in the supplemental material). These findings further support an overall loss of relative affinity for ICGs, an overall increase of relative affinity for non-ICGs, or both. Together, these data reveal that the absence of Hmt1 significantly increases the relative affinity of post-U1 snRNP addition splicing factors for non-ICGs and, to a lesser degree, decreases their relative affinities for ICGs.

To establish the role of the catalytic activity of Hmt1 in the cotranscriptional recruitment of splicing factors, we took advantage of a previously published Hmt1 catalytic mutant, *hmt1-G68R* (46). We focused on Mud2 and Prp11, since these proteins play key roles in the formation of the commitment complex yet display opposite responses in recruitment to ICGs

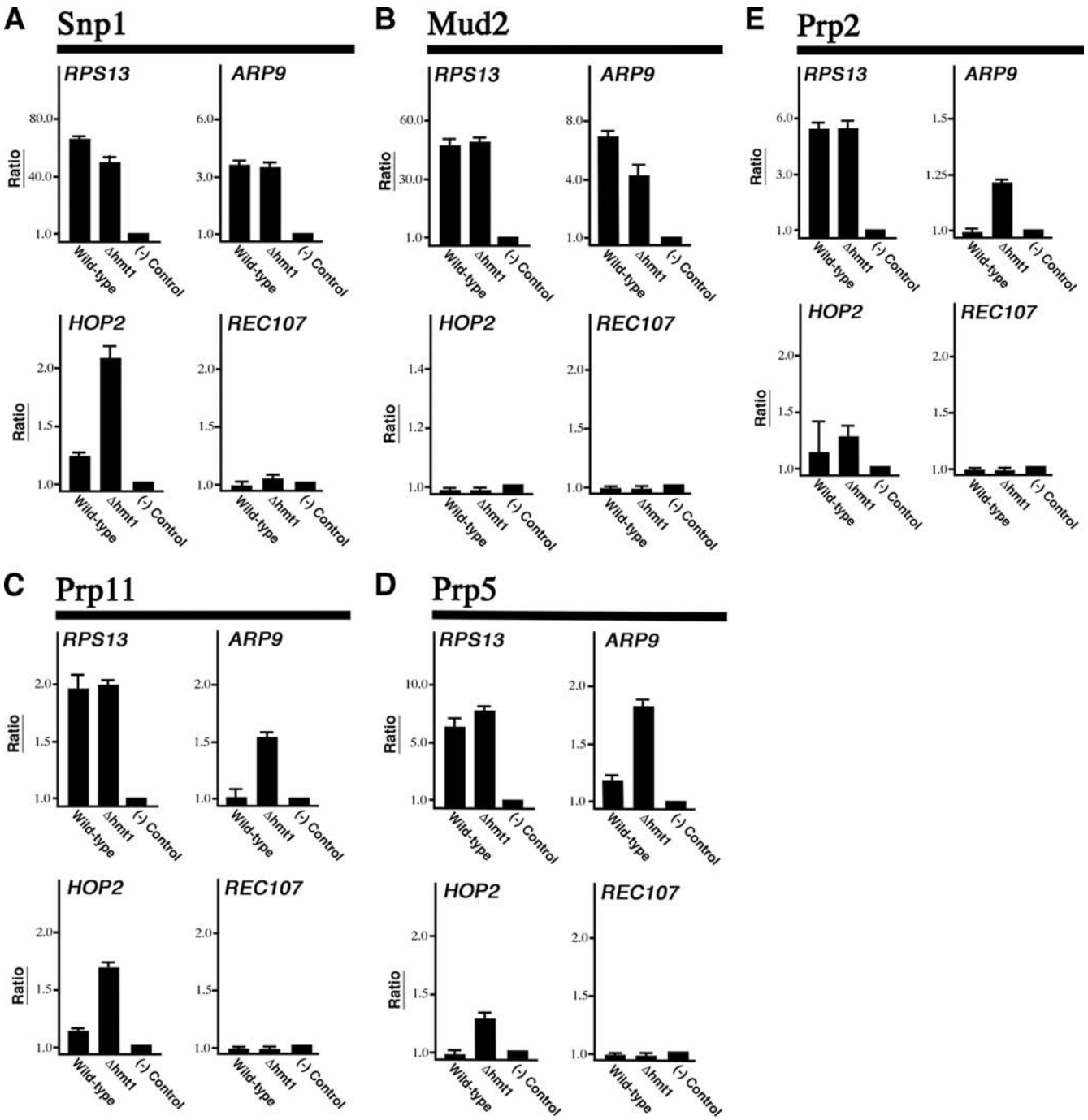


FIG. 4. Validation of genome-wide occupancy results. Directed ChIPs were performed on either wild-type or *hmt1*-null cells using anti-Myc. The splicing factors examined are Snp1 (A), Mud2 (B), Prp11 (C), Prp5 (D), and Prp2 (E). The choice of genes tested is based on the ChIP-chip data obtained. The bar graph depicts the normalized ratio of experimental signal to the intergenic region.

in the context of *HMT1* deletion. Our data with the *hmt1-G68R* mutants reveal that arginine methylation affects the cotranscriptional recruitment of both Mud2 and Prp11 in a manner that is consistent with the loss of the entire gene (Fig. 3E and F). Thus, we attribute the changes observed in our ChIP-chip analysis to the catalytic activity of Hmt1.

Directed ChIP validation of genome-wide occupancy analysis. To validate the results of our genome-wide binding studies, we performed directed ChIP on genes bound by Snp1, Mud2, Prp11, Prp2, and Prp5 in both wild-type and Δ*hmt1* cells (Fig. 4). In each case, we chose genes that the ChIP-chip data suggest have either different or similar binding based on the ranked list created with each profile. We also chose genes with different expression levels, including a highly expressed one (*RPS13*), a moderately expressed one (*ARP9*), and ones expressed at low levels (meiotic ICGs *REC107* and *HOP2*). In the

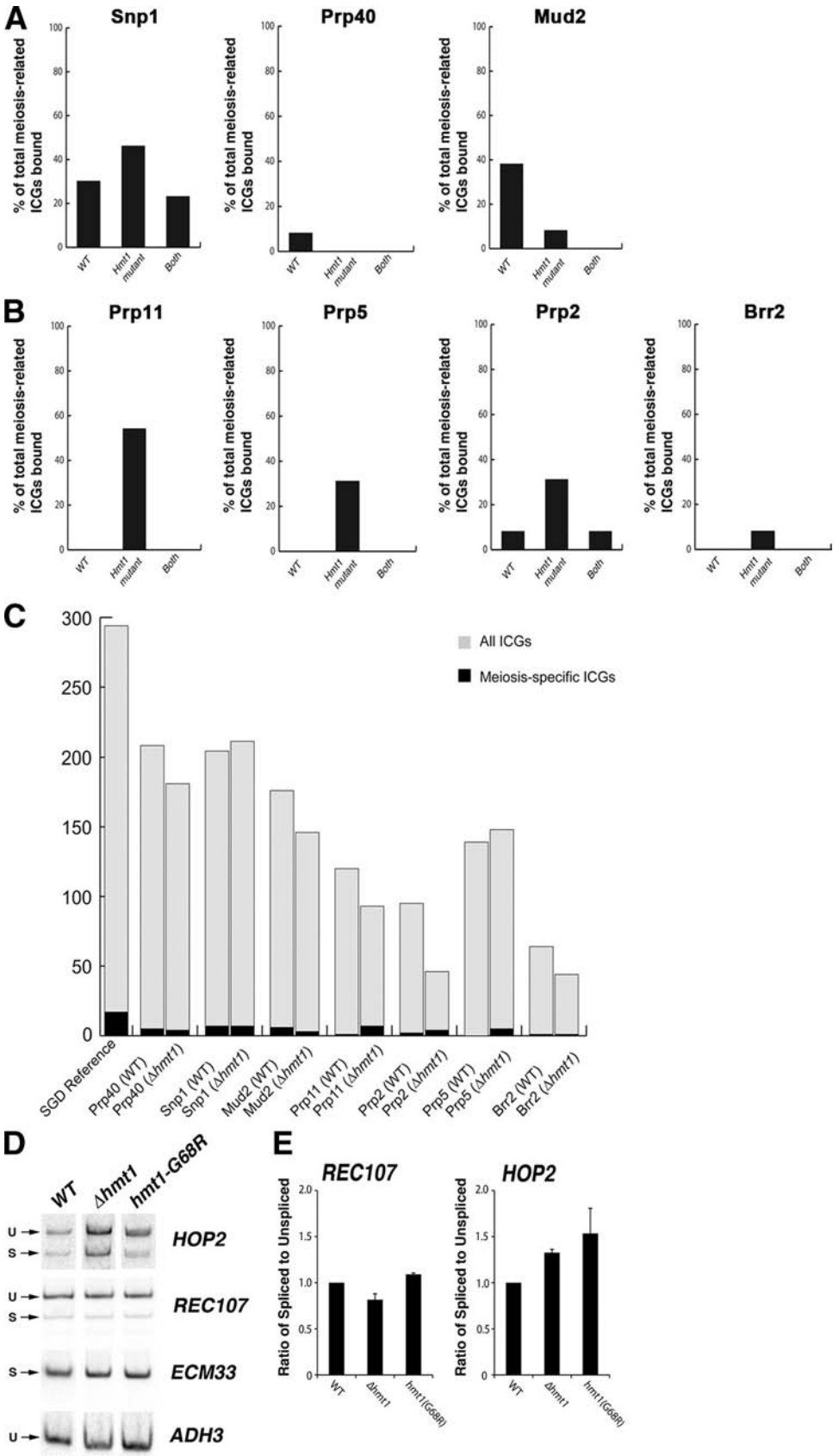
case of *RPS13*, a relatively high level of binding was detected for all five splicing factors in both wild-type and Δ*hmt1* cells, consistent with the data from genome-binding studies (Fig. 4A to E, *RPS13* graphs). In the case of *ARP9*, we did not detect any difference in binding by Snp1 and observed only a slight change in Mud2 binding in Δ*hmt1* cells compared to that in wild-type cells (Fig. 4A and B, *ARP9* graphs). Nevertheless, we did observe a significant enrichment for Prp11, Prp2, and Prp5 binding to *ARP9* in Δ*hmt1* cells (Fig. 4C to E, *ARP9* graphs). Overall, these trends support the ChIP-chip data obtained for both *RPS13* and *ARP9*.

In the case of both *REC107* and *HOP2*, our ChIP results were also generally consistent with our ChIP-chip data (Fig. 4A to E, *REC107* and *HOP2* graphs). The one exception had to do with Prp11 recruitment to *HOP2* in Δ*hmt1* cells. In this instance, a positive ChIP result was obtained (Fig. 4C, *HOP2* graph). Upon reanalysis of the original Prp11 ChIP-chip data for *HOP2* in Δ*hmt1* cells, we discovered that binding had been observed, but its *P* value was only 0.16 and thus did not meet the stringent *P* value cutoff of 0.02 (see File S5 in the supplemental material). As such, these data points were excluded from the data set that was further evaluated, despite the fact that it had an IP/input ratio that reflected a bound state (i.e., greater than 1). Notably, our ChIP-chip binding data for Prp11 in *hmt1*-G68R cells indicate that Prp11 is bound to *HOP2* (see File S6 in the supplemental material). Thus, it is likely that Prp11 is also bound to *HOP2* in the Δ*hmt1* cells and that this interaction was reflected in our directed ChIP data. As a whole, our directed ChIP results are in agreement with the ChIP-chip data.

Hmt1-dependent arginine methylation is required for regulated splicing of meiosis-related genes. We previously demonstrated that differential spliceosome recruitment predicts regulated splicing and found that this was most striking in the case of meiosis-related genes (47). The genes that encode these regulated transcripts are transcribed in the vegetative growth state but are not spliced until meiosis (14). Since our ChIP-chip data indicated that the recruitment and distribution of splicing factors to ICGs change in Hmt1 mutants, we examined whether these changes affect regulated splicing in Hmt1 mutants. To this end, we monitored the changes in pre-mRNA splicing for these meiosis-related transcripts. We analyzed our ChIP-chip data and evaluated changes in recruitment of the seven splicing factors assayed to meiosis-related genes in the Hmt1-null mutants (see File S6 in the supplemental material). We found that both Snp1 and Mud2 bound a significant fraction of meiosis-related ICGs in wild-type cells and that the loss of Hmt1 had some effect on the recruitment of Snp1, but not Mud2, to this population of ICGs (Fig. 5A, compare WT and Hmt1 mutant). All other splicing factors examined, with the exception of Prp40, showed increased association with meiosis-related ICGs in Δ*hmt1* cells, with little or no binding to these genes in wild-type cells (Fig. 5B, compare WT and Hmt1 mutant). These data indicated that Hmt1 contributes to the recruitment of specific splicing factors to meiosis-related ICGs. To further test this possibility, we performed a gene ontology (GO) test by using the GO Slim Mapper feature of the *Saccharomyces* Genome Database (SGD) to determine the effect of Hmt1 on the occupancy of meiotic ICGs for each of the splicing factors examined in this study (Fig. 5C). As a reference, the genome frequency for "meiosis" as a GO slim term is 155 genes out of 6,310 genes. Of the 294 annotated ICGs in the SGD, 17 map to the "meiosis" GO slim term, as indicated in the bar graph (Fig. 5C, SGD reference bar). Using the GO test, we found that in the cases of U1 snRNP and its associated factors Prp40, Snp1, and Mud2, the number of meiotic ICGs bound in Δ*hmt1* cells either did not change or decreased with respect to the number bound in wild-type cells (Fig. 5C, Prp40, Snp1, and Mud2 bars). In the case of splicing factors that are added after U1 snRNP is added, however, the number of meiotic ICGs bound was consistently higher in Δ*hmt1* cells than in their wild-type counterparts (Fig. 5C, Prp11, Prp2, and Prp5 bars). This trend also holds true if one calculates the "cluster frequency" by dividing the number of meiotic ICGs bound by the number of total ICGs bound for each splicing factor. Thus, Hmt1 loss leads to an increase in the impact of occupancy of meiotic ICGs by Prp11, Prp2, Prp5, and Brr2.

The enhanced recruitment of spliceosome components to meiosis-related genes in Hmt1 mutants suggests that splicing of these transcripts may be enhanced in Hmt1 mutants. To test this possibility, we examined splicing of the meiosis-related gene *HOP2*, which encodes a meiosis-specific protein that prevents synapsis between nonhomologous chromosomes while ensuring synapsis between homologs (10). In Δ*hmt1* cells, five of the seven splicing factors examined bound the *HOP2* locus, whereas only two of them bound this locus in wild-type cells (see File S6 in the supplemental material). For comparison, we examined *REC107*, another meiosis-related gene, which is known to undergo Mer1-dependent meiosis-specific splicing (61). Our data indicate that *REC107* is not bound by any splicing factors in Δ*hmt1* cells and that it is bound by only Prp40 in the wild-type cells (see File S6 in the supplemental material).

To examine how arginine methylation affects the splicing of *REC107* and *HOP2*, we performed quantitative PCR (qPCR) on cDNA produced from wild-type, Δ*hmt1*, and *hmt1-G68R* cells (Fig. 5D). The signal corresponding to unspliced and predicted spliced transcripts in each background was quantified to calculate an internal ratio for comparison (Fig. 5E). Mutations in Hmt1 had little effect on the ratio of spliced to unspliced transcripts for *REC107* (Fig. 5E, *REC107* graph), consistent with the observation that the loss of Hmt1 did not affect the recruitment of splicing factors to this locus. In contrast, the level of splicing increased 30% and 50% for the *HOP2* transcript in the cases of the Δ*hmt1* and *hmt1-G68R* mutants, respectively (Fig. 5E, *HOP2* graph), supporting our prediction that increased recruitment of splicing factors to the *HOP2* gene enhances the splicing of *HOP2* transcripts in vegetative cells. As a control for our qPCRs and isolation of total RNA, we examined the level of a constitutively spliced transcript encoded by *ECM33* and the level of non-intron-containing transcript encoded by *ADH3*. Levels of these RNAs were similar in the wild-type and mutant strains (Fig. 5D, *ECM33* and *ADH3*). A control in which a DNase I-treated, non-reverse-transcribed total RNA template was used produced no products, indicating that our RNA sample was not contaminated with genomic DNA (data not shown). Thus, our data demonstrate that arginine methylation can affect cotranscriptional recruitment of pre-mRNA splicing factors, leading to increased pre-mRNA splicing of the affected target.


A
Snp1
Prp40
Mud2
% of total meiosis-related ICGs bound
WT
Hmt1 mutant
Both
B
Prp11
Prp5
Prp2
Brr2
C
All ICGs
Meiosis-specific ICGs
SGD Reference
Prp40 (WT)
Prp40 (Δhmt1)
Snp1 (WT)
Snp1 (Δhmt1)
Mud2 (WT)
Mud2 (Δhmt1)
Prp11 (WT)
Prp11 (Δhmt1)
Prp2 (WT)
Prp2 (Δhmt1)
Prp5 (WT)
Prp5 (Δhmt1)
Brr2 (WT)
Brr2 (Δhmt1)
D
WT
Δhmt1
hmt1-G68R
HOP2
REC107
ECM33
ADH3
U
S
E
REC107
HOP2
Ratio of Spliced to Unspliced
WT
Δhmt1
hmt1(G68R)

Hmt1-dependent arginine methylation modulates the biochemical interaction between Npl3 and Snp1. Protein arginine methylation was previously shown to be critical in modulating protein complexes early in mRNP formation (71). Therefore, arginine methylation of Snp1 may regulate its interactions with nascent mRNPs. Npl3, a yeast SR-like protein (15) with known roles in mRNA export (29), is recruited during the early stages of transcription (40). Recently, Npl3 has been demonstrated to promote pre-mRNA splicing for a large subset of ICGs and to facilitate cotranscriptional binding of early splicing factors to their genomic targets (33). Since Npl3 and Snp1 have been shown to interact with each other as part of the cap-binding protein (Cbc2) complex (18) and since both proteins are substrates of Hmt1, we hypothesized that arginine methylation may modulate their biochemical interaction. To test this hypothesis, we expressed a TAP-tagged Snp1 fusion protein (Snp1-TAP) in wild-type, Δ*hmt1*, and *hmt1-G68R* backgrounds and then copurified Snp1 and its associated proteins. Immunoblotting using an anti-Npl3 antibody was then performed to probe for changes in the amount of Npl3 that copurified with Snp1 (Fig. 6, left). Immunoblotting revealed an intense band migrating at approximately 55 kDa in both the Δ*hmt1* and *hmt1-G68R* lanes (Fig. 6, left); this molecular mass corresponds to the size of Npl3 in its unmethylated form (46). In wild-type cells, a fainter band of slightly higher molecular mass corresponding to that expected for methylated Npl3 was also observed (Fig. 6, left, WT lane). The difference in the migration between the methylated and unmethylated forms of Npl3 is more readily visible in a control immunoblot, using a fraction of the input lysate (Fig. 6, input panel). The drastic contrast in the amount of Npl3 associated with Snp1 in the Hmt1 mutants indicates that the biochemical interaction between Npl3 and Snp1 was increased in the absence of Hmt1 or of its catalytic activity. The immunoblot was stripped and reprobed with an anti-CBP antibody to show that similar amounts of Snp1 were purified and loaded onto the gel (Fig. 6, right). A negative control was generated by purifying proteins from a yeast strain that does not express any TAP-tagged proteins [Fig. 6, (-) control lane].

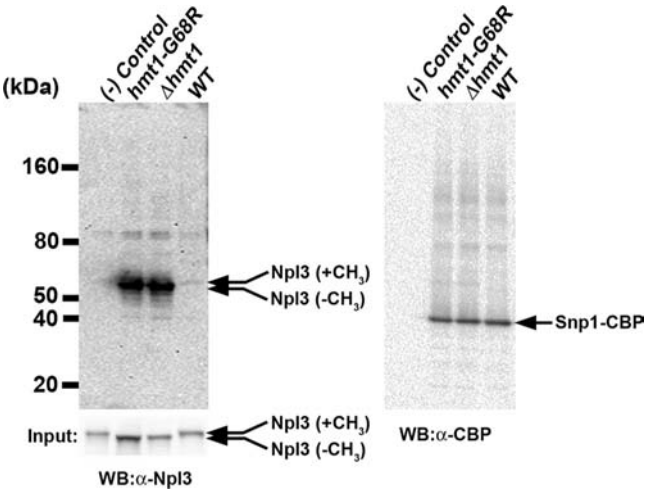


FIG. 6. Arginine methylation modulates the biochemical association between Npl3 and Snp1. TAP-tagged Snp1 expressed in wild-type, Δ*hmt1*, and *hmt1-G68R* cells was purified and resolved on a 4 to 12% gel by SDS-PAGE. (Left) The amount of Npl3 that copurified with Snp1 was determined using an anti-Npl3 antibody. The arrow (Npl3-CH_3) denotes a band corresponding to Npl3 lacking methylated arginines (lanes Δ*hmt1* and *hmt1-G68R*). A faint, slightly higher-molecular-mass band for the wild-type cells corresponds to arginine-methylated Npl3 (WT lane, arrow with Npl3 plus CH_3). (Input panel) The degree of change is readily visualized from the input of each lysate, probed with anti-Npl3 antibody (arrows). (Right) The same immunoblot was stripped and reprobed with anti-CBP antibody to detect the level of Snp1-CBP (arrow denotes Snp1-CBP). A control using a strain that does not express any TAP-tagged protein was run in parallel [(-) control lane].

DISCUSSION

During mRNP biogenesis, a network of cross-stimulatory connections and physical interdependencies is crucial for the formation of an export-competent mRNP (reviewed in references 12 and 44). Cotranscriptional recruitment of splicing factors constitutes an integral part of mRNP formation. In yeast, the ordered association of spliceosomal components during transcription has been shown to facilitate efficient cotranscriptional mRNA processing (19, 47). In the current study, we define a novel role for protein arginine methylation in the cotranscriptional recruitment of pre-mRNA splicing factors. We have identified Snp1, a component of the canonical U1 snRNP (which participates in the early steps of spliceosomal complex assembly during transcription [60]), as a novel substrate of Hmt1 *in vitro*. Using ChIP-chip, we uncovered a

FIG. 5. Changes in arginine methylation-dependent cotranscriptional recruitment influences regulated splicing. (A) The percentages of meiosis-related genes bound by the U1 snRNP components and associated factors Prp40, Snp1, and Mud2 in wild-type and Δ*hmt1* cells are shown. Where a specific meiosis-related gene is bound by a particular factor in both wild-type and Δ*hmt1* cells, the percentage of these bound genes within the total number of meiosis-related genes is calculated and denoted as "Both" on the bar graph. (B) The percentages of meiosis-related genes bound by U2 and the U4/U5/U6 tri-snRNP components and associated proteins Prp11, Prp5, Prp2, and Brr2 in wild-type and Hmt1 mutants are shown (representation is as described for panel A). (C) The effects of Hmt1 on the occupancy of meiotic ICGs relative to all ICGs bound by a splicing factor. A gene ontology (GO) test using the "Yeast GO-Slim: Process" feature of the SGD was performed, testing the total number of ICGs bound by each splicing factor in either wild-type or Δ*hmt1* cells. Of the entire 294 ICGs annotated in the yeast genome (SGD reference, gray bar), 17 map to the "meiosis" GO slim term (SGD reference, black bar). The results for each splicing factor are plotted as bars. (D) The splicing of meiotically regulated transcripts *REC107* and *HOP2* in wild-type, Δ*hmt1*, and *hmt1-G68R* cells was measured by quantitative PCR using [α-^{33}P]dCTP as described in Materials and Methods. The products were resolved on an 8% TBE gel, dried, exposed to a phosphorimager screen, and quantified by a PhosphorImager. The unspliced (U) and predicted spliced (S) products are indicated on the left. *ECM33*, a constitutively spliced transcript, and *ADH3*, an unspliced transcript, were used as controls in the quantitative PCR to demonstrate similarity in the levels of input cDNA used in the quantitative PCR. (E) The ratios of spliced to unspliced transcript levels from *REC107* and *HOP2* in wild-type, Δ*hmt1*, and *hmt1-G68R* cells are shown.

role for protein arginine methylation in modulating the recruitment of splicing factors to ICGs. A number of these data were validated by directed ChIP experiments. Changes in the cotranscriptional recruitment of splicing factors due to the loss of arginine methylation affects regulated splicing. We show that the absence of arginine methylation leads to an aberrant interaction between Snp1 and the SR-like protein Npl3. Overall, our data demonstrate that protein arginine methylation plays a critical role in defining proper cotranscriptional recruitment of pre-mRNA splicing factors.

To date, studies characterizing the recruitment of pre-mRNA splicing factors on a genome-wide scale have focused on measuring changes in the absolute number or percentage of ICGs bound by a splicing factor (32, 47, 64). This kind of measurement does not reveal how the distribution of bound ICGs changes in the context of a given condition or mutation. It is possible, for example, that a splicing factor is bound to the identical number and complement of ICGs under different conditions but that the ICGs bound in one case occupy the highly bound portion of the total bound population and in the other occupy the least-bound portion. Thus, the IDI serves as a novel method to assay the relative changes in ICG association when comparing two ChIP-chip profiles. In the study here, it reveals that the loss of Hmt1 or of its activity perturbs the relative degree of occupancy of ICGs for all U snRNPs examined, with the exception of U1 snRNP and its associated proteins.

Since the dynamics of U1 and U2 snRNP recruitment are intimately linked (19, 37, 65), it is likely that a change in the recruitment dynamics of the U1 snRNP would produce a downstream effect on the recruitment of U2 and other U snRNPs. This could be achieved by perturbing the formation of the commitment complex. Our ChIP-chip data are consistent with such an effect, as they demonstrate that recruitment of the U2 snRNP and the U4/U5/U6 snRNP, as well as of their associated proteins, is reduced in Hmt1 mutants relative to that in wild-type cells. Since the maximal U1 snRNP recruitment occurs near the 5′ ss (64), it is likely that the loss of arginine methylation in Hmt1 mutants results in increased retention of the U1 snRNP and its associated proteins at ICGs. This may be reflected by the observed decrease in binding of U1 snRNP and its associated proteins at non-ICGs, since fewer U1 snRNPs would be available to "scan" the non-ICGs as they are effectively "titrated away" as a consequence of longer ICG retention. Disruption of this *in vivo* recruitment dynamic could theoretically prevent recognition of the other components of the spliceosomal complex (such as proteins from the U2 snRNP and the U4/U5/U6 tri-snRNP) and thus their efficient recruitment to genomic targets. This could give rise to a scenario in which formation of the commitment complex is compromised. Indeed, our ChIP-chip data—according to which the U2 snRNP, the U4/U5/U6 tri-snRNP, and their associated proteins all exhibited lower relative affinities for ICGs in Hmt1 mutants—support such a scenario.

Our genome-wide assays led to the identification of a meiosis-related gene whose transcripts display enhanced splicing during vegetative growth in Hmt1 mutants. The observation that the loss of arginine methylation leads to aberrant recruitment of splicing factors and misregulated splicing of these targets supports our previous study, in which we showed that differential spliceosome recruitment predicts regulated splicing in precisely the same population of genes, i.e., meiosis-related ICGs (47). The relatively small increases in the pre-mRNA splicing of *HOP2* may be due to the nuclear mRNA surveillance machinery monitoring sites of regulated splicing, since increased levels of unspliced meiotic mRNAs have been observed in Δ*rrp6* cells, which lack a subunit of the nuclear exosome (47). Testing a mutant in which both *HMT1* and *RRP6* are deleted may reveal more dramatic changes. Notably, Δ*hmt1* cells display decreased sporulation frequency (16). Whether this abnormality results from altered splicing of meiosis-related transcripts remains to be investigated. Nevertheless, our present study has established arginine methylation as a requirement for regulated recruitment of the spliceosome.

Arginine methylation has been demonstrated to affect the function of protein substrates by modulating their interactions with other proteins (5). For example, the interaction of mammalian PRMT1 and transcription factor Ying Yang 1 (YY1) is necessary for the recruitment of histone H4-specific methyltransferase activity (54). Hmt1 has been demonstrated to control the biochemical association between mRNA export factor Npl3 and transcriptional elongation factor Tho2, as well as the self-association of Npl3 (71). While the predominant role of Npl3 is in the export of mRNAs (29), it was recently revealed that Npl3 also promotes pre-mRNA splicing, as it is required for the cotranscriptional recruitment of early splicing factors such as the U1 snRNP protein Prp40 and the U2 snRNP protein Lea1 (33). Furthermore, Npl3 is a yeast SR-like protein (15), and SR proteins in mammalian cells are known to help stabilize the U1 snRNP during the spliceosome assembly (69). These facts suggested that the loss of Hmt1 or its activity would impact how Npl3 interacts with the early splicing factors and that this interaction might account for the decreased recruitment of U2 and U4/U5/U6 tri-snRNP proteins in our ChIP-chip experiment. In support of this hypothesis, we observed an aberrant biochemical association between Npl3 and Snp1 in both the Δ*hmt1* and *hmt1-G68R* mutants (Fig. 5). Since Hmt1 also methylates Npl3, it is possible that arginine methylation regulates the cotranscriptional recruitment of splicing factors by altering Npl3 recruitment. However, our previous analysis shows that the loss of Hmt1 does not alter the recruitment of Npl3 to either ICGs or non-ICGs (see Fig. S4 in the supplemental material). Rather, the dynamics of biochemical association between Npl3 and Snp1 are most likely to control how the rest of the splicing factors are recruited in the context of mRNP biogenesis. New studies, using methylation-specific mutants of both Npl3 (45) and Snp1, are under way to address the influence of methylation on the ability of each factor to promote this association.

A recent study demonstrated that the chromatin-modifying activity of the histone acetyltransferase Gcn5 is functionally linked to the cotranscriptional recruitment of pre-mRNA splicing factors (21). Specifically, Gcn5 was found to acetylate the ICG-bound histone H3 (21). Notably, a screen used to identify protein-protein interactions that are triggered by post-translational modifications identified Hmt1 as a binding partner for acetylated histones (22). Hmt1 binds both acetylated histones H3 and H4 but methylates only H4 (34). Given that Hmt1 methylates histone H4 at position 3 (H4R3) in a chromatin-specific context (72) and that Hmt1 loss does not abolish

bulk changes in H4R3 methylation (38), it would be interesting to determine whether Hmt1 plays a role in the status of H4R3 methylation within ICG-bound histones. Such a study would determine whether the chromatin-modifying activity of Hmt1 is linked to its role in optimizing the cotranscriptional recruitment of splicing factors.

Overall, our data support a model in which Hmt1-catalyzed arginine methylation controls the cotranscriptional recruitment of splicing factors by promoting proper Npl3-Snp1 interaction. Npl3 is cotranscriptionally recruited during the early stages of transcription as part of mRNP biogenesis. As a yeast SR-like protein, it may stabilize the U1 snRNP, much as SR family proteins do in mammalian cells. During the early phase of the transcription process, the Snp1-containing U1 snRNP samples transcribed genes to identify any introns. The detection of an intron leads to the formation of a commitment complex, followed by subsequent recruitment of the rest of the spliceosome. Given that mRNP assembly involves a multitude of associations and dissociations of its components, aberrant interactions between Npl3 and Snp1 in the Hmt1 mutants may cause the U1 snRNP proteins to be retained at their genomic targets for longer times than optimal. This disruption in the dynamics of U1 snRNP recruitment, in turn, would likely affect the formation of the commitment complex, judging by the ChIP-chip data for Prp11 and Mud2. Alternatively, it is possible that a change in the dynamics of U1 snRNP recruitment due to Hmt1 activity results in a rearrangement within the spliceosome that subsequently prevents proper recruitment of the U2 and U4/U5/U6 tri-snRNP components. In summary, our study establishes a novel and key regulatory role for protein arginine methylation in controlling the dynamics of cotranscriptional recruitment of pre-mRNA splicing factors, an important facet in the assembly of an mRNP. The fact that the arginine methylation of the Snp1 homolog U1-70K is conserved in higher eukaryotes suggests that the role of arginine methylation in regulating the spliceosome may also be conserved.

ACKNOWLEDGMENTS

We thank M. Ares, S. Komili, E. P. Lei, A. McBride, M. Moore, and S. Roberts for critical reading of the manuscript, the Cullen lab members for microscopy assistance, and members of the Yu laboratory for helpful discussions. We also thank Pam Silver for antibodies and yeast strains and Manny Ares for research reagents and technical advice.

This work was supported by a Scientist Development grant (0830279N) from the American Heart Association to M.C.Y.

REFERENCES

1. **Abovich, N., X. C. Liao, and M. Rosbash.** 1994. The yeast MUD2 protein: an interaction with PRP11 defines a bridge between commitment complexes and U2 snRNP addition. Genes Dev. **8:**843–854.
2. **Abovich, N., and M. Rosbash.** 1997. Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. Cell **89:**403–412.
3. **Bachand, F.** 2007. Protein arginine methyltransferases: from unicellular eukaryotes to humans. Eukaryot. Cell **6:**889–898.
4. **Bedford, M. T., and S. G. Clarke.** 2009. Protein arginine methylation in mammals: who, what, and why. Mol. Cell **33:**1–13.
5. **Bedford, M. T., A. Frankel, M. B. Yaffe, S. Clarke, P. Leder, and S. Richard.** 2000. Arginine methylation inhibits the binding of proline-rich ligands to Src homology 3, but not WW, domains. J. Biol. Chem. **275:**16030–16036.
6. **Boisvert, F. M., J. Cote, M. C. Boulanger, P. Cleroux, F. Bachand, C. Autexier, and S. Richard.** 2002. Symmetrical dimethylarginine methylation is required for the localization of SMN in Cajal bodies and pre-mRNA splicing. J. Cell Biol. **159:**957–969.
7. **Boisvert, F. M., J. Cote, M. C. Boulanger, and S. Richard.** 2003. A proteomic analysis of arginine-methylated protein complexes. Mol. Cell. Proteomics **2:**1319–1330.
8. **Bossie, M. A., C. DeHoratius, G. Barcelo, and P. Silver.** 1992. A mutant nuclear protein with similarity to RNA binding proteins interferes with nuclear import in yeast. Mol. Biol. Cell **3:**875–893.
9. **Chan, S. P., D. I. Kao, W. Y. Tsai, and S. C. Cheng.** 2003. The Prp19p-associated complex in spliceosome activation. Science **302:**279–282.
10. **Chen, Y. K., C. H. Leng, H. Olivares, M. H. Lee, Y. C. Chang, W. M. Kung, S. C. Ti, Y. H. Lo, A. H. Wang, C. S. Chang, D. K. Bishop, Y. P. Hsueh, and T. F. Wang.** 2004. Heterodimeric complexes of Hop2 and Mnd1 function with Dmc1 to promote meiotic homolog juxtaposition and strand assimilation. Proc. Natl. Acad. Sci. U. S. A. **101:**10572–10577.
11. **Cheng, D., J. Cote, S. Shaaban, and M. T. Bedford.** 2007. The arginine methyltransferase CARM1 regulates the coupling of transcription and mRNA processing. Mol. Cell **25:**71–83.
12. **Cole, C. N., and J. J. Scarcelli.** 2006. Transport of messenger RNA from the nucleus to the cytoplasm. Curr. Opin. Cell Biol. **18:**299–306.
13. **Das, R., K. Dufu, B. Romney, M. Feldt, M. Elenko, and R. Reed.** 2006. Functional coupling of RNAP II transcription to spliceosome assembly. Genes Dev. **20:**1100–1109.
14. **Davis, C. A., L. Grate, M. Spingola, and M. Ares, Jr.** 2000. Test of intron predictions reveals novel splice sites, alternatively spliced mRNAs and new introns in meiotically regulated genes of yeast. Nucleic Acids Res. **28:**1700–1706.
15. **Dreyfuss, G., V. N. Kim, and N. Kataoka.** 2002. Messenger RNA-binding proteins and the messages they carry. Nat. Rev. Mol. Cell Biol. **3:**195–205.
16. **Enyenihi, A. H., and W. S. Saunders.** 2003. Large-scale functional genomic analysis of sporulation and meiosis in Saccharomyces cerevisiae. Genetics **163:**47–54.
17. **Gary, J. D., W. J. Lin, M. C. Yang, H. R. Herschman, and S. Clarke.** 1996. The predominant protein-arginine methyltransferase from Saccharomyces cerevisiae. J. Biol. Chem. **271:**12585–12594.
18. **Gavin, A. C., M. Bosche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J. M. Rick, A. M. Michon, C. M. Cruciat, M. Remor, C. Hofert, M. Schelder, M. Brajenovic, H. Ruffner, A. Merino, K. Klein, M. Hudak, D. Dickson, T. Rudi, V. Gnau, A. Bauch, S. Bastuck, B. Huhse, C. Leutwein, M. A. Heurtier, R. R. Copley, A. Edelmann, E. Querfurth, V. Rybin, G. Drewes, M. Raida, T. Bouwmeester, P. Bork, B. Seraphin, B. Kuster, G. Neubauer, and G. Superti-Furga.** 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature **415:**141–147.
19. **Gornemann, J., K. M. Kotovic, K. Hujer, and K. M. Neugebauer.** 2005. Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. Mol. Cell **19:**53–63.
20. **Green, D. M., K. A. Marfatia, E. B. Crafton, X. Zhang, X. Cheng, and A. H. Corbett.** 2002. Nab2p is required for poly(A) RNA export in Saccharomyces cerevisiae and is regulated by arginine methylation via Hmt1p. J. Biol. Chem. **277:**7752–7760.
21. **Gunderson, F. Q., and T. L. Johnson.** 2009. Acetylation by the transcriptional coactivator Gcn5 plays a novel role in co-transcriptional spliceosome assembly. PLoS Genet. **5:**e1000682.
22. **Guo, D., T. R. Hazbun, X. J. Xu, S. L. Ng, S. Fields, and M. H. Kuo.** 2004. A tethered catalysis, two-hybrid system to identify protein-protein interactions requiring post-translational modifications. Nat. Biotechnol. **22:**888–892.
23. **Hacker, I., B. Sander, M. M. Golas, E. Wolf, E. Karagoz, B. Kastner, H. Stark, P. Fabrizio, and R. Luhrmann.** 2008. Localization of Prp8, Brr2, Snu114 and U4/U6 proteins in the yeast tri-snRNP by electron microscopy. Nat. Struct. Mol. Biol. **15:**1206–1212.
24. **Henry, M. F., and P. A. Silver.** 1996. A novel methyltransferase (Hmt1p) modifies poly(A)$^+$-RNA-binding proteins. Mol. Cell. Biol. **16:**3668–3678.
25. **Hieronymus, H., and P. A. Silver.** 2004. A systems view of mRNP biology. Genes Dev. **18:**2845–2860.
26. **Howe, K. J., C. M. Kane, and M. Ares, Jr.** 2003. Perturbation of transcription elongation influences the fidelity of internal exon inclusion in Saccharomyces cerevisiae. RNA **9:**993–1006.
27. **Huh, W. K., J. V. Falvo, L. C. Gerke, A. S. Carroll, R. W. Howson, J. S. Weissman, and E. K. O'Shea.** 2003. Global analysis of protein localization in budding yeast. Nature **425:**686–691.
28. **Ito, T., T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, and Y. Sakaki.** 2001. A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc. Natl. Acad. Sci. U. S. A. **98:**4569–4574.
29. **Kadowaki, T., S. Chen, M. Hitomi, E. Jacobs, C. Kumagai, S. Liang, R. Schneiter, D. Singleton, J. Wisniewska, and A. M. Tartakoff.** 1994. Isolation and characterization of Saccharomyces cerevisiae mRNA transport-defective (mtr) mutants. J. Cell Biol. **126:**649–659.
30. **Knop, M., K. Siegers, G. Pereira, W. Zachariae, B. Winsor, K. Nasmyth, and E. Schiebel.** 1999. Epitope tagging of yeast genes using a PCR-based strategy: more tags and improved practical routines. Yeast **15:**963–972.
31. **Kosowski, T. R., H. R. Keys, T. K. Quan, and S. W. Ruby.** 2009. DExD/H-box Prp5 protein is in the spliceosome during most of the splicing cycle. RNA **15:**1345–1362.
32. **Kotovic, K. M., D. Lockshon, L. Boric, and K. M. Neugebauer.** 2003. Co-

transcriptional recruitment of the U1 snRNP to intron-containing genes in yeast. Mol. Cell. Biol. **23:**5768–5779.
33. **Kress, T. L., N. J. Krogan, and C. Guthrie.** 2008. A single SR-like protein, Npl3, promotes pre-mRNA splicing in budding yeast. Mol. Cell **32:**727–734.
34. **Kuo, M. H., X. J. Xu, H. A. Bolck, and D. Guo.** 2009. Functional connection between histone acetyltransferase Gcn5p and methyltransferase Hmt1p. Biochim. Biophys. Acta **1789:**395–402.
35. **Kwak, Y. T., J. Guo, S. Prajapati, K. J. Park, R. M. Surabhi, B. Miller, P. Gehrig, and R. B. Gaynor.** 2003. Methylation of SPT5 regulates its interaction with RNA polymerase II and transcriptional elongation properties. Mol. Cell **11:**1055–1066.
36. **Lacadie, S. A., and M. Rosbash.** 2005. Cotranscriptional spliceosome assembly dynamics and the role of U1 snRNA:5′ss base pairing in yeast. Mol. Cell **19:**65–75.
37. **Lacadie, S. A., D. F. Tardiff, S. Kadener, and M. Rosbash.** 2006. In vivo commitment to yeast cotranscriptional splicing is sensitive to transcription elongation mutants. Genes Dev. **20:**2055–2066.
38. **Lacoste, N., R. T. Utley, J. M. Hunter, G. G. Poirier, and J. Cote.** 2002. Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase. J. Biol. Chem. **277:**30421–30424.
39. **Laggerbauer, B., T. Achsel, and R. Luhrmann.** 1998. The human U5-200kD DEXH-box protein unwinds U4/U6 RNA duplices in vitro. Proc. Natl. Acad. Sci. U. S. A. **95:**4188–4192.
40. **Lei, E. P., H. Krebber, and P. A. Silver.** 2001. Messenger RNAs are recruited for nuclear export during transcription. Genes Dev. **15:**1771–1782.
41. **Lei, E. P., and P. A. Silver.** 2002. Intron status and 3′-end formation control cotranscriptional export of mRNA. Genes Dev. **16:**2761–2766.
42. **Liu, Q., and G. Dreyfuss.** 1995. In vivo and in vitro arginine methylation of RNA-binding proteins. Mol. Cell. Biol. **15:**2800–2808.
43. **Longtine, M. S., A. McKenzie III, D. J. Demarini, N. G. Shah, A. Wach, A. Brachat, P. Philippsen, and J. R. Pringle.** 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast **14:**953–961.
44. **Luna, R., H. Gaillard, C. Gonzalez-Aguilera, and A. Aguilera.** 2008. Biogenesis of mRNPs: integrating different processes in the eukaryotic nucleus. Chromosoma **117:**319–331.
45. **McBride, A. E., J. T. Cook, E. A. Stemmler, K. L. Rutledge, K. A. McGrath, and J. A. Rubens.** 2005. Arginine methylation of yeast mRNA-binding protein Npl3 directly affects its function, nuclear export, and intranuclear protein interactions. J. Biol. Chem. **280:**30888–30898.
46. **McBride, A. E., V. H. Weiss, H. K. Kim, J. M. Hogle, and P. A. Silver.** 2000. Analysis of the yeast arginine methyltransferase Hmt1p/Rmt1p and its in vivo function. Cofactor binding and substrate interactions. J. Biol. Chem. **275:**3128–3136.
47. **Moore, M. J., E. M. Schwartzfarb, P. A. Silver, and M. C. Yu.** 2006. Differential recruitment of the splicing machinery during transcription predicts genome-wide patterns of mRNA splicing. Mol. Cell **24:**903–915.
48. **Morris, D. P., and A. L. Greenleaf.** 2000. The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. J. Biol. Chem. **275:**39935–39943.
49. **Mowen, K. A., J. Tang, W. Zhu, B. T. Schurter, K. Shuai, H. R. Herschman, and M. David.** 2001. Arginine methylation of STAT1 modulates IFNalpha/beta-induced transcription. Cell **104:**731–741.
50. **Ohkura, N., M. Takahashi, H. Yaguchi, Y. Nagamura, and T. Tsukada.** 2005. Coactivator-associated arginine methyltransferase 1, CARM1, affects pre-mRNA splicing in an isoform-specific manner. J. Biol. Chem. **280:**28927–28935.
51. **Perreault, A., C. Lemieux, and F. Bachand.** 2007. Regulation of the nuclear poly(A)-binding protein by arginine methylation in fission yeast. J. Biol. Chem. **282:**7552–7562.
52. **Piruat, J. I., and A. Aguilera.** 1998. A novel yeast gene, THO2, is involved in RNA Pol II transcription and provides new evidence for transcriptional elongation-associated recombination. EMBO J. **17:**4859–4872.
53. **Raghunathan, P. L., and C. Guthrie.** 1998. RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. Curr. Biol. **8:**847–855.
54. **Rezai-Zadeh, N., X. Zhang, F. Namour, G. Fejer, Y. D. Wen, Y. L. Yao, I. Gyory, K. Wright, and E. Seto.** 2003. Targeted recruitment of a histone H4-specific methyltransferase by the transcription factor YY1. Genes Dev. **17:**1019–1029.
55. **Rigaut, G., A. Shevchenko, B. Rutz, M. Wilm, M. Mann, and B. Seraphin.** 1999. A generic protein purification method for protein complex characterization and proteome exploration. Nat. Biotechnol. **17:**1030–1032.
56. **Ruby, S. W., T. H. Chang, and J. Abelson.** 1993. Four yeast spliceosomal proteins (PRP5, PRP9, PRP11, and PRP21) interact to promote U2 snRNP binding to pre-mRNA. Genes Dev. **7:**1909–1925.
57. **Russell, I. D., and D. Tollervey.** 1992. NOP3 is an essential yeast protein which is required for pre-rRNA processing. J. Cell Biol. **119:**737–747.
58. **Shen, E. C., M. F. Henry, V. H. Weiss, S. R. Valentini, P. A. Silver, and M. S. Lee.** 1998. Arginine methylation facilitates the nuclear export of hnRNP proteins. Genes Dev. **12:**679–691.
59. **Silverman, E. J., A. Maeda, J. Wei, P. Smith, J. D. Beggs, and R. J. Lin.** 2004. Interaction between a G-patch protein and a spliceosomal DEXD/H-box ATPase that is critical for splicing. Mol. Cell. Biol. **24:**10101–10110.
60. **Smith, V., and B. G. Barrell.** 1991. Cloning of a yeast U1 snRNP 70K protein homologue: functional conservation of an RNA-binding domain between humans and yeast. EMBO J. **10:**2627–2634.
61. **Spingola, M., and M. Ares, Jr.** 2000. A yeast intronic splicing enhancer and Nam8p are required for Mer1p-activated splicing. Mol. Cell **6:**329–338.
62. **Staley, J. P., and J. L. Woolford, Jr.** 2009. Assembly of ribosomes and spliceosomes: complex ribonucleoprotein machines. Curr. Opin. Cell Biol. **21:**109–118.
63. **Tadesse, H., J. Deschenes-Furry, S. Boisvenue, and J. Cote.** 2008. KH-type splicing regulatory protein interacts with survival motor neuron protein and is misregulated in spinal muscular atrophy. Hum. Mol. Genet. **17:**506–524.
64. **Tardiff, D. F., S. A. Lacadie, and M. Rosbash.** 2006. A genome-wide analysis indicates that yeast pre-mRNA splicing is predominantly posttranscriptional. Mol. Cell **24:**917–929.
65. **Tardiff, D. F., and M. Rosbash.** 2006. Arrested yeast splicing complexes indicate stepwise snRNP recruitment during in vivo spliceosome assembly. RNA **12:**968–979.
66. **Trinkle-Mulcahy, L., S. Boulon, Y. W. Lam, R. Urcia, F. M. Boisvert, F. Vandermoere, N. A. Morrice, S. Swift, U. Rothbauer, H. Leonhardt, and A. Lamond.** 2008. Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. J. Cell Biol. **183:**223–239.
67. **Wachtel, C., and J. L. Manley.** 2009. Splicing of mRNA precursors: the role of RNAs and proteins in catalysis. Mol. Biosyst. **5:**311–316.
68. **Wahl, M. C., C. L. Will, and R. Luhrmann.** 2009. The spliceosome: design principles of a dynamic RNP machine. Cell **136:**701–718.
69. **Wu, J. Y., and T. Maniatis.** 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. Cell **75:**1061–1070.
70. **Yamagata, K., H. Daitoku, Y. Takahashi, K. Namiki, K. Hisatake, K. Kako, H. Mukai, Y. Kasuya, and A. Fukamizu.** 2008. Arginine methylation of FOXO transcription factors inhibits their phosphorylation by Akt. Mol. Cell **32:**221–231.
71. **Yu, M. C., F. Bachand, A. E. McBride, S. Komili, J. M. Casolari, and P. A. Silver.** 2004. Arginine methyltransferase affects interactions and recruitment of mRNA processing and export factors. Genes Dev. **18:**2024–2035.
72. **Yu, M. C., D. W. Lamming, J. A. Eskin, D. A. Sinclair, and P. A. Silver.** 2006. The role of protein arginine methylation in the formation of silent chromatin. Genes Dev. **20:**3249–3254.

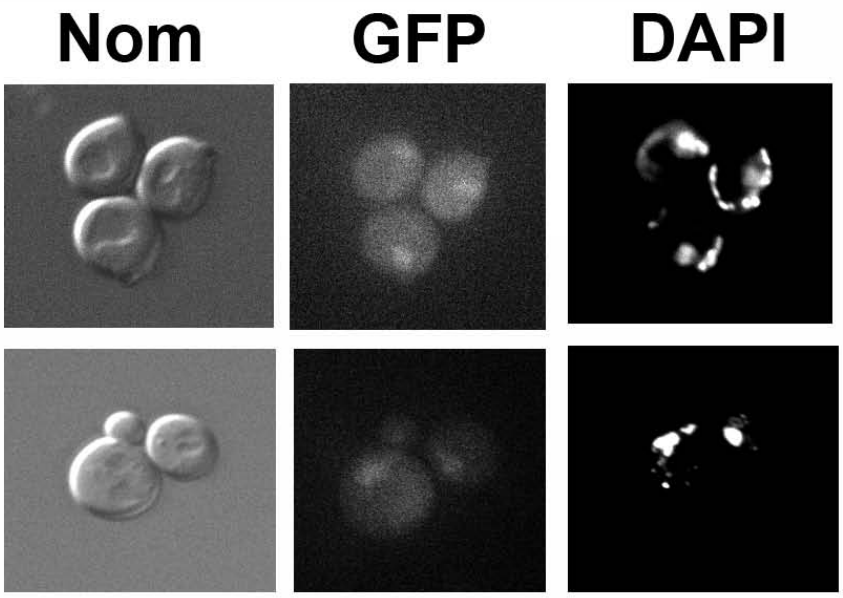
Supplemental Figure 1: Arginine methylation of Snp1's mammalian homolog U1-70K and its effect on the cellular localization of Snp1.

A) The cellular localization of GFP-tagged Snp1 is visualized using DIC (panel labeled as "Nom") and fluorescent microscopy (panels labeled as "GFP" and "DAPI").

B) Amino acid alignment of Snp1 with its human homolog U1-70K. The consensus sequence is listed below. Blue represents amino acids that are conserved between Snp1 and U1-70K. Methylated arginines (denoted by an asterisk) within U1-70K were identified by mass spectrometry, as described in the Materials and Methods section.

Chen et al.
Supplmental Figure 1

A.



B.

```
                                  20                                  40                                  60
human U1 snRNP 70k  MTQF-LP--P NLLA-LFAPR DPIPYLPPLE KLPHEKHH-N QP-YCGIAPY I----REF-E 49
SNP1                MN-YNLSKYP DDVSRLFKPR PPLSYKRPTD -YPYAKRQTN -PNITGVANL LSTSLKHYME 57
Consensus           MXQXNLXKYP XXXXRLFXPR XPXXYXXPXX KXPXXKXXTN QPNXXGXAXX XSTSLXXXME

                                 80                                 100                                 120
                                 * *
human U1 snRNP 70k  D-PRDAPPPT RAETREERME RKRREKIERR QQEVETELKM WDPHN-DPNA QG-DAFKTLF 106
SNP1                EFP-EGSPNN HLQ-RYEDI--KL-SKI-KN AQLLDRRLQN WNP-NVDPHI KDTDPYRTIF 110
Consensus           XFPRXXXPXX XXXTRXEXXE RKXRXKIEXX XQXXXXXLXX WXPHNVDPXX XXTDXXXTXF

                                 140                                 160                                 180
human U1 snRNP 70k  VARVNYDTTE SKLRREFEV- YGPIKRIHMV YSKRSGKPRG YAFIEYEHER D-MHS--AYK 162
SNP1                IGRLPYDLDE IELQKYF-VK FGEIEKIRIV KDKITQKSKG YAFIVF---K DPISSKMAFK 166
Consensus           XXRXXYDXXE XXLXXXFEVK XGXIXXIXXV XXKXXXKXXG YAFIXXEHEX DPXXSKMAXK

                                 200                                 220                                 240
                        *                   *                   *
human U1 snRNP 70k  ----HADGKK I-DGRRV-LV DVERGRTVKG WRPRRLGGGL GGTRRGGADV NIRHSGRDDT 216
SNP1                EIGVH-RGIQ IKD--RICIV DIERGRTVKY FKPRRLGGGL GG--RG---- ---YSNRD-- 212
Consensus           EIGVHAXGXX IKDGRRXCXV DXERGRTVKX XXPRRLGGGL GGTRRGGADV NIRXSXRDDT

                                 260                                 280                                 300
human U1 snRNP 70k  SRYDERPGPS PLPHRDRDRD RERERRERSR ERDKERERRR SRSRDRRRRS RSRDKEERRR 276
SNP1                SRL---PG-- ----R-FA-S ASTSN-P-A- ERNYA-PRL- PR-RE---TS -S-------- 243
Consensus           SRXDERPGPS PLPHRDXXRX XXXXXRXRXR ERXXXRXRXR XRSRXRRRXS RSRDKEERRR

                                 320                                 340                                 360
human U1 snRNP 70k  SRERSKDKDR DRKRRSSRSR ERARRER-ER KEELRGGGGD MAEPSEAGDA PPDDGPPGEL 335
SNP1                S-AYSA--DR ---YGSS-TL D-ARY-RGNR --PLLSA--- -ATPT-A--A ---------V 276
Consensus           SRXXSXDKDR DRKXXSSRXX XRARXERGXR KEXLXXXGGD MAXPXEAGDA PPDDGPPGEX

                                 380                                 400                                 420
                                                     * *                 *                   *       * *
human U1 snRNP 70k  GPDGPDGPEE KGRDRDRERR RSHRSERERR RDRDRDRDRD REHKRGERGS ERGRDEARGG 395
SNP1                ---------- ---------- -T--S----- --V------- --YK-----S ---RN-SR-- 286
Consensus           GPDGPDGPEE KGRDRDRERR RXHRSERERR RDXDRDRDRD REXKRGERGS ERGRXEXRGG

                                 440                                 460
human U1 snRNP 70k  GGGQDNGLEG LGNDSRDMYM ESEGGDGYLA PENGYLMEAA PE- 437
SNP1                ---------- ----TR---- ESQ--P---A PK-----EA- PDY 300
Consensus           GGGQDNGLEG LGNDXRDMYM ESXGGXGYLA PXNGYLMEAA PXY
```

Supplemental Figure 2: Graphical representation used in the calculation of IDI and NIDI. The actual ChIP-chip data of Prp11 is used to illustrate how the IDI is graphically represented.

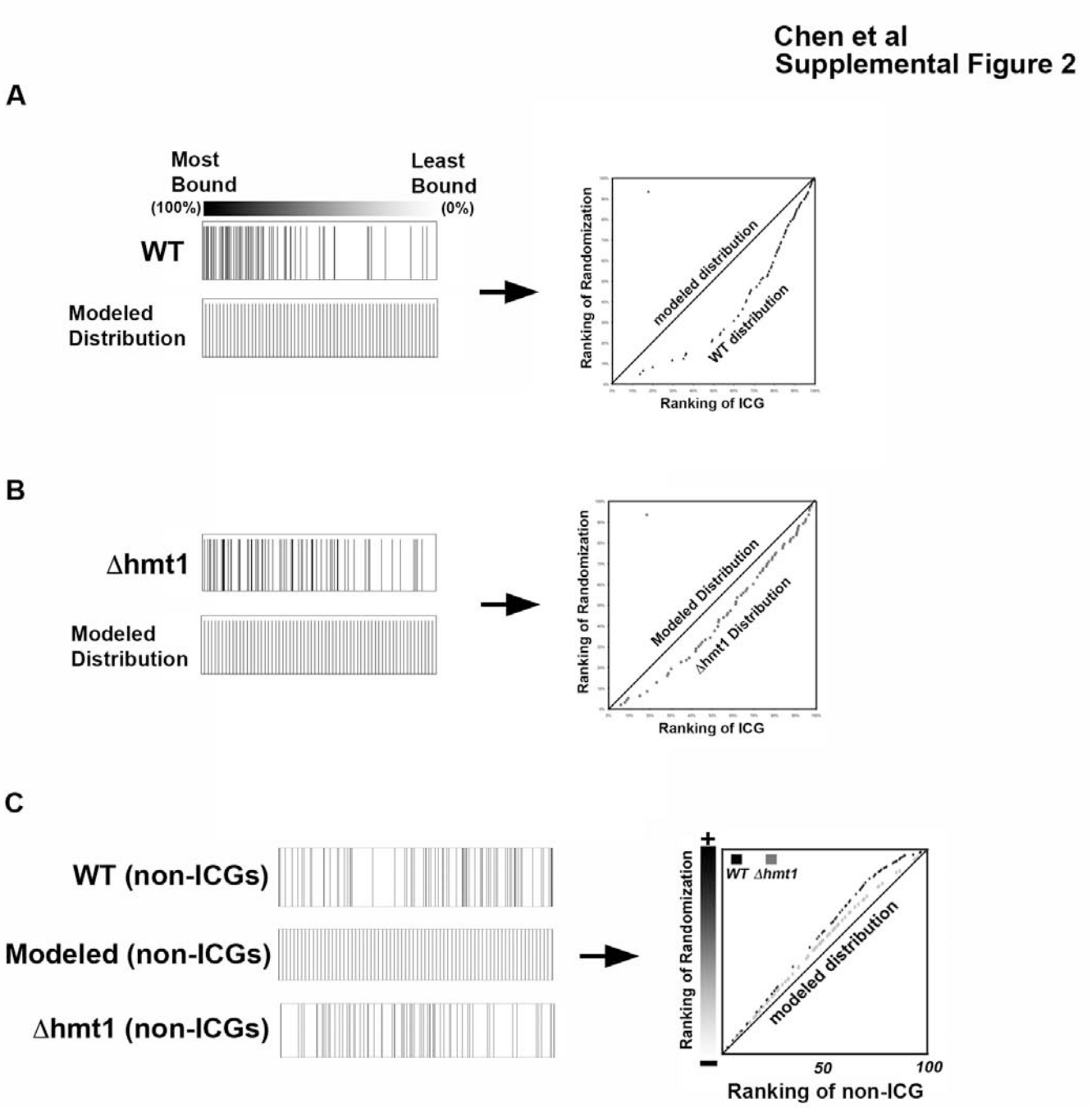
A) The top horizontal bar (WT) represents a complete ChIP-chip profile, ranking each gene bound by the splicing factor of interest. As previously shown, most of the ICGs that are bound by a splicing factor are enriched in the highly-bound proportion of a profile (47). Each vertical bar within the profile denotes a bound ICG. The bottom horizontal profile (Modeled Distribution) represents a modeled scenario in which the same ICGs bound are evenly distributed across a profile. When these two profiles are plotted together, the X coordinate represents the actual ICG ranking (Ranking of ICG) and the Y coordinate represents the modeled ranking (Ranking of Randomization). Since splicing factors are known to preferentially associate with ICGs (47), the ICGs bound by the splicing factors tend to display higher ordered ranking relative to non-ICGs. The 45° straight line denotes a modeled scenario in which the binding of ICGs is evenly distributed across a profile, which would indicate a lack of preferential associations. The IDI (or D_{rand}) is calculated by taking the average of the sum of the distance between the actual rankings of ICGs and the modeled rankings of ICGs (see Materials and Methods).

B) An analysis similar to that described above was carried out for Prp11 recruitment in *Δhmt1*. In both A and B, the Y coordinates (Modeled Distribution) for the 'WT' and 'Δhmt1' profiles are identical in a pair-wise comparison, which reflects the fact that a common set of bound ICGs was examined.

C) Actual ChIP-chip data for Prp5 were used to illustrate how the non-ICG distribution index (NIDI) is represented graphically. In this case, each vertical bar within the rectangle denotes a bound non-ICG, and the same equation with different variables (to reflect non-ICGs) is used to calculate the NIDI. Pair-wise comparison and graphical output are as described for IDI. In contrast to the ICGs, non-ICGs are enriched mostly in the lower ranked portion of the total profile, and thus, are generally inversely correlated with the ICG distribution.

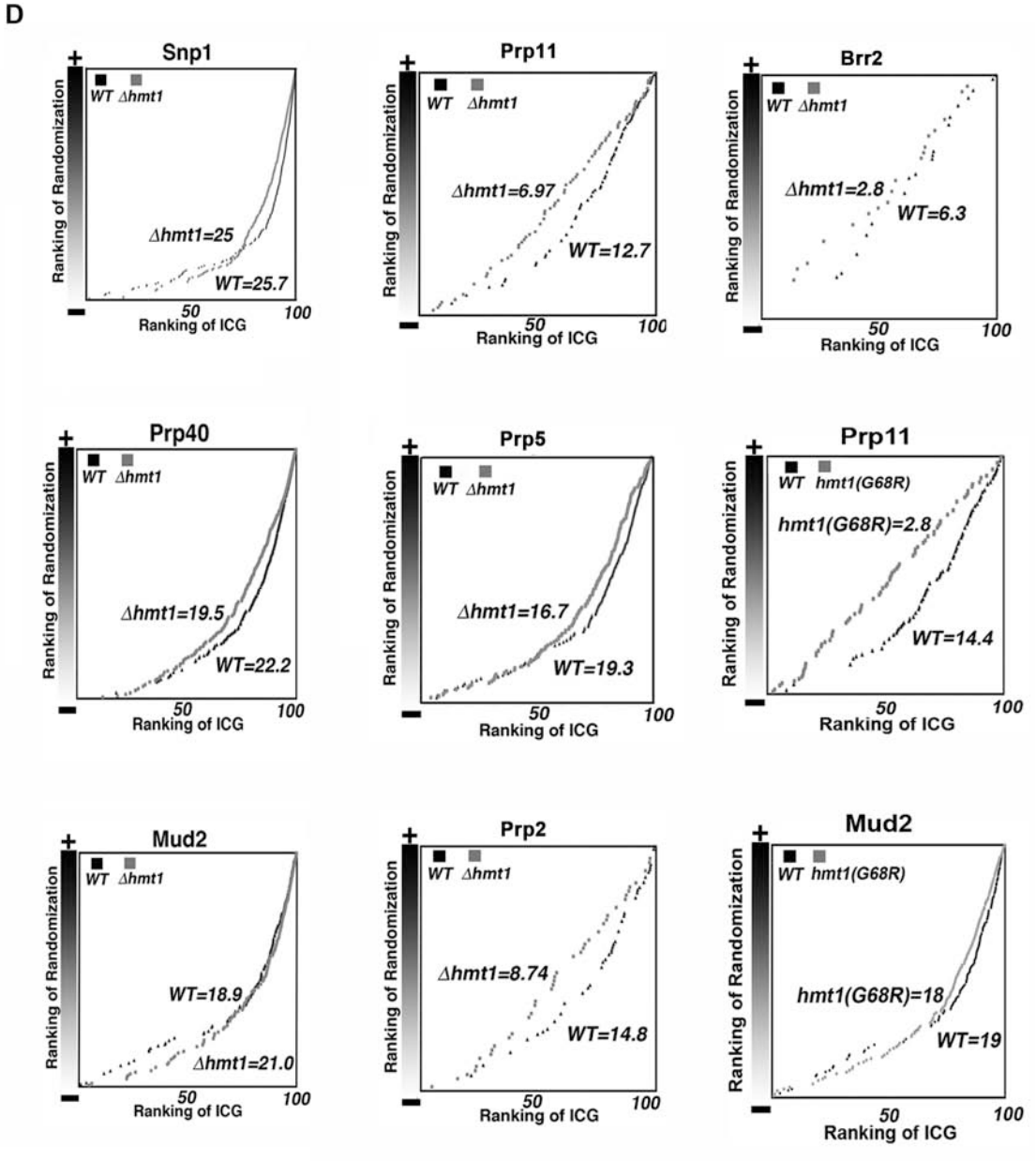
D) The change in the intron distribution index (IDI) for the two profiles compared was plotted as described in Supplemental Fig. 2, parts A-C. For each ICG plotted, the X-axis coordinate is the ranking within a binding profile (in either the wild-type or *Δhmt1* strain) and the Y-axis coordinate, which is identical for both profiles, represents the ranking of randomization based on a modeled distribution.

Chen et al
Supplemental Figure 2

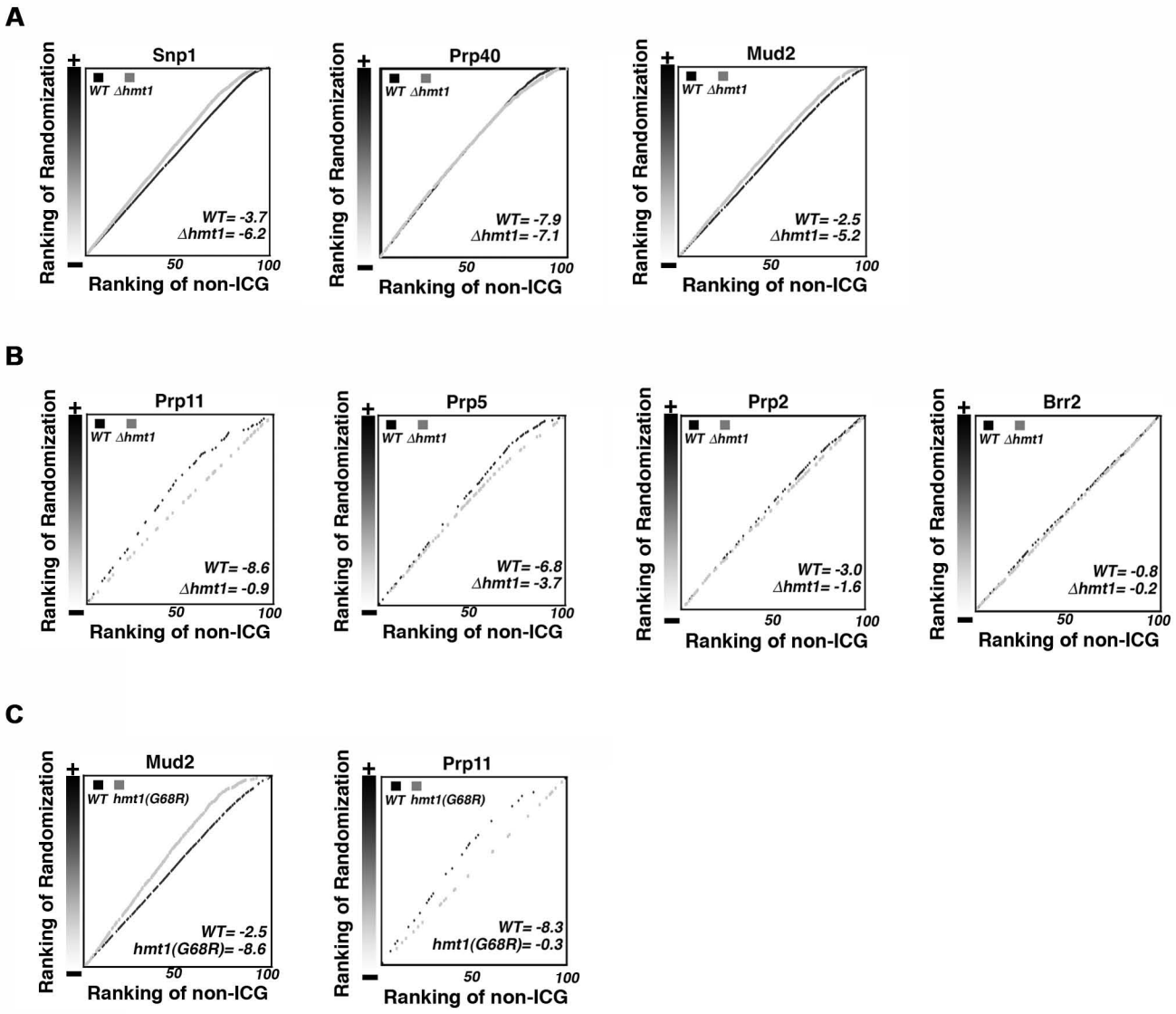


Chen et al
Supplemental Figure 2

D



Supplemental Figure 3: The non-ICG distribution (NIDI) for the splicing factors studied. ChIP-chip data obtained in the wild-type and Hmt1 mutants were used to determine how Hmt1 affects the recruitment of splicing factors to non-ICGs. The graphical output of each pair-wise comparison is shown: Snp1 (panel A), Prp40 (panel B), Mud2 (panel C), Prp11 (panel D), Prp5 (panel E), Prp2 (panel F), and Brr2 (panel G). The calculation for the non-ICG distribution index (NIDI) is the same as for the IDI except that the rankings of non-ICGs within a profile, rather than rankings of ICGs (see Materials and Methods and Supplemental Fig. 2), are used. The wild-type ChIP-chip data for Prp40, Mud2, Prp11, Prp5, Prp2, and Brr2 were obtained from Moore et al 2006.


A
Snp1
WT Δhmt1
WT= -3.7
Δhmt1= -6.2
Prp40
WT= -7.9
Δhmt1= -7.1
Mud2
WT= -2.5
Δhmt1= -5.2
B
Prp11
WT= -8.6
Δhmt1= -0.9
Prp5
WT= -6.8
Δhmt1= -3.7
Prp2
WT= -3.0
Δhmt1= -1.6
Brr2
WT= -0.8
Δhmt1= -0.2
C
Mud2
WT hmt1(G68R)
WT= -2.5
hmt1(G68R)= -8.6
Prp11
WT= -8.3
hmt1(G68R)= -0.3
Ranking of Randomization
Ranking of non-ICG
50
100
+
−

Supplemental Figure 4: Loss of Hmt1 does not affect recruitment of Npl3 to ICGs.

A) The percentages of Npl3-bound ICGs in wild-type and Hmt1 mutants.

B) The total number of Npl3-bound genes in wild-type and Hmt1 mutant.

C) Comparison of ChIP-chip data using the IDI method showed no change in the distribution of Npl3-bound ICGs in wild-type and Hmt1 mutant cells.

D) The same holds true for non-ICGs as well. The ChIP-chip data used in all calculations were obtained from Yu et al, 2004.

Chen et al.
Supplemental Figure 4