INVESTIGATIONS ON PROTEIN ARGININE METHYLTRANSFERASES IN DIFFERENTIATION AND CANCER

ΟΔΥΣΣΕΑΣ ΠΑΤΟΥΝΑΣ
ΒΙΟΛΟΓΟΣ

Πανεπιστημιο Ιωαννινων
Σχολη Επιστημων Υγειας
Τμημα Βιολογικων εφαρμογων
& Τεχνολογιων

Ιδρυμα Τεχνολογιας & Ερευνας
Ινστιτουτο Μοριακης Βιολογιας
& Βιοτεχνολογιας
Τμημα Βιοϊατρικων Ερευνων

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΙΩΑΝΝΙΝΑ 2017

Στην οικογένειά μου και στους καλούς φίλους που είναι πάντα στο πλευρό μου...

Acknowledgements

The experiments for the present thesis were carried out in the group of Dr. Frank O. Fackelmayer at the Institute of Molecular Biology and Biotechnology in Ioannina (FORTH-IMMB).

I would like to thank Dr. Frank O. Fackelmayer for giving me the opportunity to work in his lab, and for always being next to me in every problem that I was facing during my work.

I would also like to thank Dr. Carmen Eckerich for her help in every technical problem, for helping me with my experiments and for the cooperation we had in the lab.

I would also like to say thanks to the students who carried out their diploma thesis in our lab during these years; Mr. Christoforos Pappas, Mr. Ioannis Manolaras, Mrs. Ioanna Papacharalampous, Mrs. Athena Tsiali, Mrs. Elina Athanailidou and Mrs. Anieza Gkiokai. It was a great pleasure to work with and supervise them.

A big thanks also goes to Dr. Diana Scholz from the University of Konstanz, Germany, for her help in every theoretical question and technical issue I had about LUHMES cells; Prof. Anastasia Politou and Dr. Panagiota Georgoulia for our collaboration on LBR; Prof. Evangelos Kolettas and Dr. Georgios Markopoulos for supplying us with the lung cancer cell lines and their help with the FACS experiments.

I would also like to thank the members of the committee who accepted to evaluate the present thesis; Prof. Savvas Christoforidis, Dr. Frank O. Fackelmayer, Prof. Efstathios Hatziloukas, Prof. Evangelos Kolettas, Prof. Georgios Mavrothalassitis, Prof. Theologos Michailidis and Prof. Georgios Thyfronitis.

Finally I would like to say thanks to my family and my friends who supported me during my PhD studies.

This work was partially supported by Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: THALES "IDIPRO" (MIS 379440), KRIPIS-BIOSYS and a PhD scholarship from the Institute of Molecular Biology and Biotechnology.

Published paper (attached in the end of the thesis):

Patounas, O., Papacharalampous, I., Eckerich, C., Markopoulos, G. S., Kolettas, E. and Fackelmayer, F. O., "A novel splicing isoform of Protein Arginine Methyltransferase 1 (PRMT1) that lacks the dimerization arm and correlates with cellular malignancy." *J. Cell. Biochem.* 2017

Posters:

Patounas, O., Papacharalampous, I., Eckerich, C. and Fackelmayer, F. O., "A new splicing isoform of Protein Arginine Methyltransferase 1 (PRMT1) that lacks the multimerization arm of the enzyme." *67th Annual Conference of the Hellenic Society for Biochemistry and Molecular Biology, Ioannina Greece*, 2017

Georgoulia, P., Soupsana, K., Louka, A., **Patounas, O.**, Fackelmayer, F., Georgatos, S., Caflisch, A., Politou, A., S., "Structure and Dynamics of Lamin B Receptor: Order through Disorder", *Patra Greece*, 2015

Table of Contents

ABBREVIATIONS **10**

ABSTRACT **13**

ΠΕΡΙΛΗΨΗ **15**

1. INTRODUCTION **19**

1.1 Regulation of gene expression 19

1.2 Epigenetics in cancer and differentiation 20

1.3 Epigenetic regulation of gene expression by arginine methylation 21

1.3.1 Arginine methylation *21*

1.3.2 Protein aRginine MethylTransferases (PRMTs) *23*

1.3.3 Transcriptional regulation by PRMTs *26*

1.4 Members of PRMT family 27

1.4.1 Type I PRMTs *27*

1.4.2 Type II PRMTs *29*

1.4.3 Type III PRMTs *30*

1.5 PRMT1 31

1.5.1 Structure of the protein *32*

1.5.2 Oligomerization and mechanism of action of PRMT1 ..... *33*

1.5.3 Alternative splicing isoforms of PRMT1 ..... *34*

1.5.4 PRMT1, alternative splicing and cancer *37*

1.6 PRMT8 40

1.7 LUHMES – an appropriate model for studies on PRMT8 48

1.8 Aim of the thesis 51

2. MATERIALS AND METHODS **53**

2.1 Cell culture methods 53

2.1.1 Human cancer cell lines and normal human lung fibroblasts *53*

2.1.2 LUHMES *53*

2.2 Transfection and infection of cells 54

2.2.1 Transfection of cells with Polyethylenimine (PEI) and Lipofectamine 2000 ..... *54*

2.2.2 Generation of lentiviruses and infection of LUHMES *54*

2.3 Microscopy methods 55

2.3.1 Immunofluorescence microscopy *55*

2.3.2 Fluorescence Recovery After Photobleaching (FRAP) *56*

2.4 Biochemical and Molecular Genetics Methods 56

2.4.1 Quantitation of total protein *56*

2.4.2 Protein extraction/precipitation according to Wessel&Flügge (1984) *56*

2.4.3 Gel electrophoresis of proteins and Western blot *57*

2.4.4 Immunoprecipitation *57*

2.4.5 In vitro methylation assay *58*

2.4.6 RNA analysis - Preparation of total RNA and cDNA synthesis *59*

2.4.7 Polymerase chain reaction (PCR) *59*

2.5 Work with bacteria 60

2.5.1 Transformation of bacteria by electroporation *60*

2.5.2 Small-scale preparation of plasmid DNA *60*

2.5.3 Medium-scale preparation of plasmid DNA *60*

2.6 Other methods 61

2.6.1 Fluorescence-activated cell sorting (FACS) *61*

2.6.2 Identification of molecules in proximity with the BioID method *61*

2.6.3 Studying the distribution of PRMT8M16 between the cytoplasm and the nucleus *62*

2.6.4 Antibody production in rabbits ... *62*

3. RESULTS ... **67**

3.1 A NEW PRMT1 SPLICING VARIANT ... 67

3.1.1 Bioinformatic identification of the new variant ... *67*

3.1.2 Expression of Δarm in different cells ... *68*

3.1.3 Deletion of the dimerization arm and amino-terminal variation of PRMT1 ... *69*

3.1.4 Comparison of the subcellular localization of Δarm with the other amino-terminal variants of PRMT1 ... *70*

3.1.5 Expression levels of Δarm ... *71*

3.1.7 Dimerization ability of PRMT1v1Δarm ... *75*

3.1.8 Enzymatic activity of PRMT1v1Δarm ... *76*

3.1.9 Molecular dynamics of Δarm ... *77*

3.1.10 Involvement of Δarm expression in lung cancer ... *78*

3.1.11 Epithelial to mesenchymal transition (EMT) increases Δarm expression levels ... *81*

3.1.12 Δarm expression during differentiation ... *83*

3.2 NOVEL PRMT1 SUBSTRATES ... 85

3.3 TOOLS FOR THE STUDY OF PRMT8 ... 88

3.3.1 PRMT8 is endogenously expressed in differentiated LUHMES - Morphological and molecular characterization of the system ... *89*

3.3.2 PRMT8 mRNA is highly up-regulated during LUHMES differentiation ... *94*

3.3.3 PRMT8 antibodies ... *95*

3.3.4 PRMT8 hypothetical amino terminus ... *97*

3.3.5 Subcellular localization of PRMT8 ... *98*

3.3.6 PRMT8 nuclear amounts are regulated ... *100*

3.3.7 Knocking-down PRMT8 ... *103*

3.3.8 Tools for the identification of interactions partners and/or substrates of PRMT8 ... *104*

4. DISCUSSION ... **107**

4.1 PRMT1: IDENTIFICATION AND CHARACTERIZATION OF A NOVEL SPLICE VARIANT AND ITS IMPACT ON CANCER ... 107

4.2 PRMT8: ROLE IN NEURONAL DIFFERENTIATION USING A HUMAN CELLULAR MODEL *IN VITRO* ... 113

5. REFERENCES ... **123**

APPENDIX ... **133**

Abbreviations

ADMA	Asymmetric dimethylarginine
AdoHcy	S-adenosylhomocysteine
AdoMet	S-adenosylmethionine
BTG	B-cell translocation gene
CARM1	Coativator-associated arginine methyltransferase 1
cDNA	complementary DNA
CNS	Central nervous system
DRD2	Dopamine receptor D2
EMT	Epithelial to mesenchymal transition
EST	Expressed sequence tag
FACS	Fluorescence-activated cell sorting
FGF	Fibroblast growth factor
FRAP	Fluorescence recovery after photobleaching
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GAR	Glycine arginine rich
GDNF	Glial cell-derived neurotrophic factor
GFP	Green fluorescent protein
GMNN	Geminin
h/mESC	human/mouse embryonic stem cell
H3R3me2a	Histone H3, arginine 3 dimethylation
HA	Hemagglutinin
HEK293	Human embryonic kidney cell line
HeLa	Henrietta Lacks cervix carcinoma
Hmt	HnRNP MethylTransferase
hpf	hours post fertilization
HRP	Horseradish peroxidase
iRC	induced regenerative competent
LBR	Lamin B receptor
LIF	Leukemia inhibitor factor

LUHMES	Lund Human Mesencephalon
MMA	Monomethylated arginine
mRNA	messenger RNA
NES	Nuclear export signal
NF165	Neurofilament 165
NGF	Nerve growth factor
NLGN	Neuroligin
ORF	Open reading frame
PAX3	Paired box 3
PCR	Polymerase chain reaction
PEI	Polyethylenimine
PGM-rich	Proline glycine methionine rich
PLD	Phospholipase D
PRMT	Protein arginine methyltransferase
qPCR	quantitative PCR
RGG-rich	Arginine glycine rich
RS domain	Arginine serine domain
RT-PCR	Real Time PCR
SAF-A	Scaffold attachment factor A
SAM	S-adenosylmethionine
SCID	Severe combined immunodeficiency
SDMA	Symmetric dimethylarginine
shRNA	short hairpin RNA
Sox2	Sex determining region Y-box 2
SV2	Synaptic vesicle protein 2
SYP	Synaptophysin
TH	Tyrosine hydrolase
THW loop	Threonine, histidine, tryptophane loop
XIST	X-inactive specific transcript
YFP	Yellow fluorescent protein

Abstract

"Investigations on Protein Arginine Methyltransferases (PRMTs) in differentiation and cancer"

Methylation of arginine residues by the family of *Protein Arginine Methyltransferase* (PRMT) enzymes is an important modulator of protein function that is involved in epigenetic regulation of gene expression, DNA damage response, RNA maturation and cell signaling. The pre-mRNA of the predominant enzyme of the family, PRMT1, is alternatively spliced in the 5'- end and produces seven different isoforms. PRMT1 isoforms vary in their aminoterminal region, are expressed at different levels in different tissues, and have distinct substrate specificity and intracellular localization. Here, we characterize a novel splicing isoform of PRMT1, which lacks introns 8 and 9. These exons encode the dimerization arm of the enzyme that is essential for PRMT1 enzymatic activity. Consequently, the isoform does not form catalytically active oligomers with the other endogenous PRMT1 isoforms. Photobleaching experiments reveal an immobile fraction of the enzyme in the nucleus, in accordance with earlier results from our laboratory that had shown a tight association of inhibited or inactivated PRMT1 with chromatin and the nuclear scaffold. This isoform is detected in a variety of cell lines, but it is expressed at higher levels in cancer cells and it is further induced by the EMT-inducing transcription factor Snail. Thus, the novel isoform could act as a modulator of PRMT1 activity in cancer cells by acting as a competitive inhibitor that shields substrates from access to active PRMT1 oligomers.

Moreover, we provide evidence that lamin B receptor (LBR) is a novel PRMT1 substrate, which might contribute to the epigenetic regulation of gene expression through non-histone substrates.

The second part of the thesis focuses on PRMT8, a protein that shares a high degree of homology with PRMT1, but is mainly expressed in the central nervous system. Until now, PRMT8 was poorly studied due to the lack of appropriate tools. However, in the present thesis, we generated several tools and performed a wide variety of experiments in order to determine the physiological role of PRMT8 in neuronal differentiation and maintenance. We find that in a neuronal cell system, LUHMES, PRMT8 expression is induced during differentiation. A recombinant PRMT8:GFP construct accumulates in the

nucleus while LUHMES cells are differentiating. Furthermore, we provide evidence that the amount of the protein in the nucleus cannot exceed certain levels and reach a plateau when the total amount of the protein increases in the cell. Moreover, we generated and validated lentiviral vectors to succesfully achieve the knock-down of endogenous PRMT8, and to identify interaction partners and substrates through the use of the BioID method in future experiments.

Collectively, the work described in the present thesis shows that PRMT1 and PRMT8 have important roles in cancer and neuronal differentiation, respectively. The new knowledge will be instrumental to further investigations, which will elucidate the potential use of PRMT1 and PRMT8 as diagnostic markers or drug targets in cancer and neurodegenerative diseases.

Περίληψη

«Μελέτη του ρόλου των μεθυλοτρανσφερασών της αργινίνης (PRMTs) στην κυτταρική διαφοροποίηση και στον καρκίνο»

Η μεθυλίωση των καταλοίπων αργινίνης από την οικογένεια των *μεθυλοτρανσφερασών της αργινίνης* (Protein Arginine Methyltransferases, PRMTs) είναι ένας σημαντικός ρυθμιστής της λειτουργίας των πρωτεϊνών και εμπλέκεται σε διαδικασίες όπως: η επιγενετική ρύθμιση της γονιδιακής έκφρασης, οι αποκρίσεις σε βλάβες του γενετικού υλικού, η ωρίμανση του RNA και η κυτταρική σηματοδότηση. Ο κυριότερος εκπρόσωπος της οικογένειας των ενζύμων αυτών είναι η PRMT1, από το γονίδιο της οποίας παράγονται εφτά διαφορετικές ισομορφές μετά από εναλλακτικό μάτισμα του αρχικού μεταγράφου στο 5’ άκρο. Οι ισομορφές αυτές εκφράζονται σε διαφορετικά επίπεδα ανάλογα με τον κυτταρικό τύπο, επιδεικνύουν διακριτή ειδικότητα για συγκεκριμένα υποστρώματα καθώς και διαφορετικό υποκυττάριο εντοπισμό. Στην παρούσα διδακτορική διατριβή, ανακαλύψαμε μια νέα ισομορφή της PRMT1 που δε σχετίζεται με το αμινοτελικό άκρο της πρωτεΐνης, όπως οι εφτά γνωστές έως τώρα ισομορφές. Η νέα αυτή ισομορφή δεν εμπεριέχει τα εξόνια 8 και 9 τα οποία είναι υπεύθυνα για την κωδικοποίηση του βραχίονα διμερισμού του ενζύμου.

Εξαιτίας αυτού, το ένζυμο δεν μπορεί να δημιουργήσει καταλυτικά ενεργά ολιγομερή με τις υπόλοιπες ισομορφές του. Πειράματα FRAP (Fluorescent Recovery After Photobleaching) έδειξαν την ύπαρξη ενός ακινητοποιημένου κλάσματος της πρωτεΐνης στον πυρήνα. Τα αποτελέσματα αυτά είναι σύμφωνα με παλιότερα ευρήματα του εργαστηρίου μας, που έδειξαν ότι υπάρχει ισχυρή πρόσδεση της ανενεργής ή ανεσταλμένης PRMT1 στη χρωματίνη και στο πυρηνικό ικρίωμα (nuclear scaffold). Η νέα ισομορφή μπορεί να προσδεθεί στα ίδια υποστρώματα με την ενζυματικά ενεργή PRMT1. Η έκφραση της ανιχνεύθηκε σε διαφορετικές κυτταρικές σειρές και ήταν αυξημένη σε αυτές με τα πιο έντονα καρκινικά χαρακτηριστικά ή μετά από την υπερέκραφαση του μεταγραφικού παράγοντα Snail που επάγει τη μετάπτωση από επιθήλιο σε μεσέγχυμα (EMT). Θεωρούμε ότι η νέα ισομορφή μπορεί να λειτουργήσει ως ρυθμιστής της δράσης της PRMT1 στα καρκινικά κύτταρα, δρώντας ως ανταγωνιστικός αναστολέας που δεν επιτρέπει την πρόσβαση των ενεργών ολιγομερών της PRMT1 στα υποστρώματά τους.

Επίσης, παρουσιάζουμε νέα δεδομένα που καταδεικνύουν ότι ο υποδοχέας της λαμίνης Β (Lamin B Receptor, LBR) αποτελεί υπόστρωμα της PRMT1, κάτι που δεν ήταν γνωστό μέχρι σήμερα. Η μεθυλίωση του LBR από την PRMT1, πιθανώς να αποτελεί μέρος της επιγενετικής ρύθμισης των γονιδίων μέσω υποστρωμάτων που δεν ανήκουν στην οικογένεια των ιστονών.

Στο δεύτερο τμήμα της διδακτορικής διατριβής, μελετήσαμε την PRMT8, το όγδοο μέλος της οικογένειας των PRMTs. Η αλληλουχία της PRMT8 είναι κατά ένα μεγάλο ποσοστό ομόλογη με αυτήν της PRMT1, αλλά η έκφραση της είναι περιορισμένη στο κεντρικό νευρικό σύστημα. Μέχρι σήμερα, η PRMT8 δεν είχε μελετηθεί ενδελεχώς, εξαιτίας της έλλειψης κατάλληλων εργαλείων και κυτταρικών συστημάτων.

Στα πλαίσια της παρούσας διατριβής αναπτύξαμε νέα εργαλεία και διεξαγάγαμε μία σειρά πειραμάτων με σκοπό να διερευνήσουμε τον φυσιολογικό ρόλο της PRMT8 στη διαφοροποίηση και τη διατήρηση του νευρικού ιστού. Από τα πειράματα μας, προκύπτει ότι η PRMT8 εκφράζεται ενδογενώς στο κυτταρικό μοντέλο νευρικής διαφοροποίησης, LUHMES (LUnd Human MESencephalon). Η υπερέκφραση της PRMT8, συντηγμένης με την πράσινη φθορίζουσα πρωτεΐνη (Green Fluorescent Protein, GFP), κατέδειξε ότι η PRMT8 συσσωρεύεται στον πυρήνα καθώς τα LUHMES διαφοροποιούνται. Βάσει των αποτελεσμάτων μας φαίνεται ότι τα επίπεδα της πρωτεΐνης στον πυρήνα δεν μπορούν να ξεπεράσουν ένα όριο ανεξάρτητα από τη συνολική ποσότητά της στο κύτταρο. Τέλος, κατασκευάσαμε και αξιολογήσαμε νέους φορείς λέντι-ιών για την αποσιώπηση της έκφρασης της PRMT8 και την ανίχνευση μορίων με τα οποία αλληλεπιδρά μέσω της μεθόδου που ονομάζεται BioID.

Συνοπτικά, τα αποτελέσματα της διδακτορικής διατριβής μου αποδεικνύουν ότι η PRMT1 και η PRMT8 επιτελούν σημαντικό ρόλο στον καρκίνο και στη νευρική διαφοροποίηση, αντίστοιχα. Τα αποτελέσματα αυτά θα έχουν μεγάλη χρησιμότητα σε επόμενες μελέτες όπου θα διερευνηθεί η πιθανότητα χρήσης τους ως διαγνωστικά εργαλεία ή φαρμακευτικοί στόχοι στον καρκίνο και σε νευροεκφυλιστικές ασθένειες.

1. Introduction

1.1 Regulation of gene expression

Pathways affecting gene expression can be categorized into “short-term” mechanisms that allow cells to rapidly and transiently respond to extracellular and intracellular signals, and “long-term” mechanisms that govern the functional identity of a particular cell type. Short-term mechanisms usually involve pathways that transmit signals inside the cell, e.g. from the plasma membrane to the nucleus, and lead to the transcriptional activation or inactivation of target genes. In contrast, long-term mechanisms involve relatively stable changes in the general status of genes, and establish whether genes are silenced or potentially available for short-term regulation. Thus, short- and long-term mechanisms are interconnected and act together to assure homeostasis of the organism.

Long-term regulation of gene expression is usually referred to as “epigenetics”, because it is based on an additional layer of information on top of the DNA sequences in the genome. In other words, epigenetic mechanisms put “marks” onto chromatin that can be stably inherited to daughter cells over mitosis, and transmit information about the expression status of a gene without changing the underlying DNA sequence itself. These marks affect gene expression by allowing or prohibiting the access of particular transcription factors to gene promoters and enhancers. Epigenetic regulation occurs through two basic principles: the methylation of cytosine residues in the DNA sequence, and the post-translational modification of histones (Fig. 1.1). Both mechanisms cooperate to set up and faithfully replicate epigenetic marks on chromatin, and are indispensable for all developmental, differentiation and maintenance processes in an organism. The first mechanism, DNA methylation, is catalyzed by a family of enzymes called *DNA methyltransferases*. It usually occurs in CpG (cytosine-phosphate-guanine) islands in regulatory regions such as promoters, and lead to silencing of the associated gene. DNA methylation is critically involved in the formation of heterochromatin, and plays an important role in genomic imprinting, dosage compensation, and telomere stability.

In the present thesis, I focus on histone modifications, and more specifically on a family of enzymes involved in the methylation of arginine residues in histones, in healthy (differentiation of neural precursor cells to mature neurons) as well as in pathological conditions (cancer). These enzymes could become potential drug targets to alter specific

pathways and gene expression, and thus to treat a broad range of diseases including cancer and neurodegenerative diseases. Other mechanisms such as RNA interference are also studied as epigenetic mechanisms.

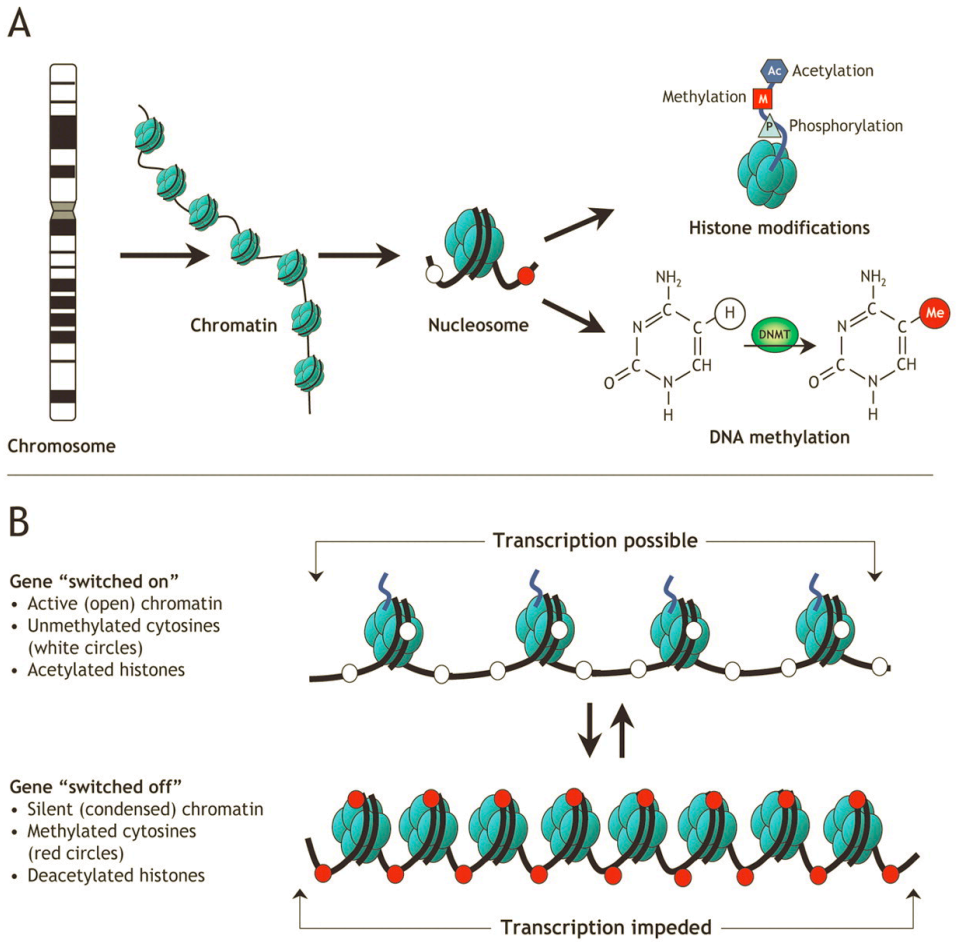


Fig. 1.1: Epigenetic regulation of gene expression. (A) Epigenetic marks occur through the methylation of DNA cytosine residues and through modifications on chromatin histones. These histone modifications include acetylation, methylation, and phosphorylation as depicted, but also others that fulfill specific functions. (B) Depending on the structure of the chromatin, gene expression is either activated (gene "switched on") or inactivated (gene "switched off") (image obtained from http://cnx.org/content/m26565/latest/).

1.2 Epigenetics in cancer and differentiation

The development of a complete eukaryotic organism from one single cell, the fertilized oocyte, is one of the most fascinating processes in biology. Extensive research in the last decades has elucidated many mechanisms that contribute to this process, although we are still far from seeing and appreciating the whole picture. Clearly, gaining a deeper

understanding of how pluripotent cells differentiate to post-mitotic cells with unique traits, how tissues and organs are formed and later maintained in healthy homeostasis, is crucial for developing new ways to rationally manipulate the involved pathways for medical purposes.

Cancer, the leading cause of death worldwide, is a consequence of defects in these homeostatic mechanisms. The biology of human tumors has been characterized as having six hallmarks: sustained proliferative capacity, evasion of growth suppressors, resisting cell death (apoptosis), enabling replicative immortality, inducing angiogenesis and activating invasion and metastasis (Hanahan and Weinberg, 2011). These features cooperate to promote tumor development, growth and aggressiveness. Understanding how the implicated molecular mechanisms are regulated, will facilitate the development of better therapeutic strategies, improve the life quality and increase the survival of the patients.

Although the development of an organism is a physiological procedure and cancer is a pathological condition, both share a common characteristic. Development and cancer are governed by epigenetic mechanisms that direct cells to differentiate into specific cell types or de-differentiate to become more aggressive, respectively. Primarily, development from the fertilized oocyte to an embryo and later to a mature organism is achieved through changes in the expression of genes. Even though all cells of the body have the same genome, only around 20% of the genes are expressed in any given cell type, and determine the phenotype and physiological properties of the cell. Disruption of epigenetic mechanisms can lead to gene silencing of tumor suppressors as well as to the activation of oncogenes, resulting in malignant cellular transformation. Thus, cancer is characterized by global changes in the epigenome of the cancer cell.

1.3 Epigenetic regulation of gene expression by arginine methylation

1.3.1 Arginine methylation

Arginine methylation is a post-translational modification of proteins that, in addition to other functions, is involved in epigenetic gene regulation through the histone code. Arginine is a positively charged amino acid due to the presence of the strongly basic (pKa around 12.0) guanidine group that is protonated at all physiologically relevant pH values.

The positive charge, together with multiple hydrogen bond donating capacity and a high degree of polarity, ensures surface localization of arginine side chains, except in rare instances of formation of buried guanidinium-carboxylate salt bridges. Arginine residues function in proteins generally as positively charged groups, contributing to the binding of negatively charged ligands such as phosphate and phosphate esters, including nucleic acids. The three nitrogens in the guanidino group of its side chain can be methylated. The possible products of arginine methylation are monomethylated arginine (omega-N_G-monomethylarginine, MMA) with a single methyl group on the terminal nitrogen atom, asymmetric dimethylarginine (omega-N_G,N_G-dimethylarginine, ADMA) where two methyl groups are placed on one of the terminal nitrogen atoms of the guanidine group, and symmetric dimethylarginine (omega-N_G,N'_G-dimethylarginine, SDMA) where one methyl group is placed on each of the terminal guanidine nitrogens. In yeast, arginine can also be methylated at the internal δ-guanidino nitrogen atom, producing delta-monomethylated arginine (Fig. 1.2).

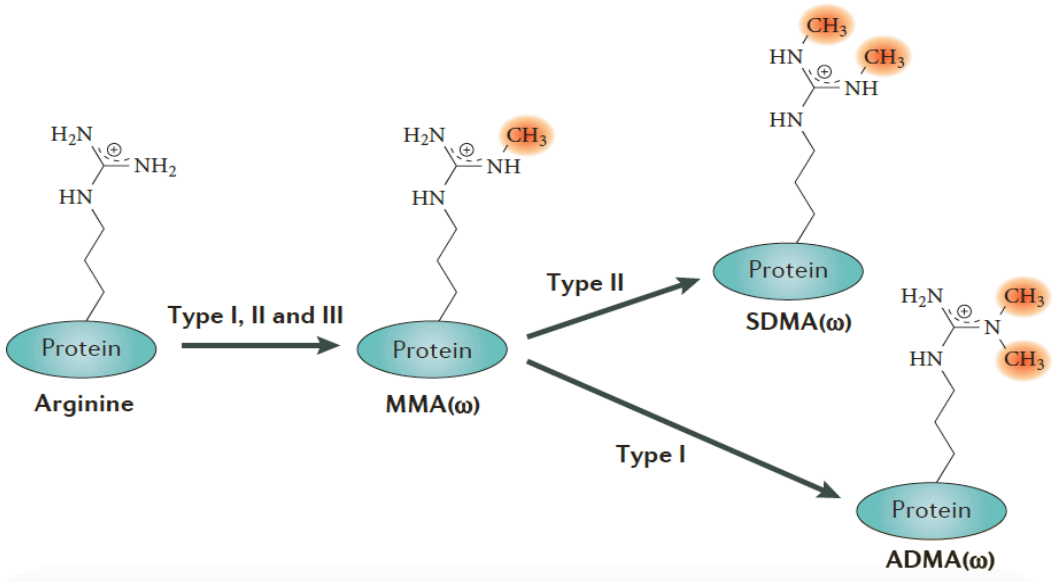


Fig. 1.2: Types of methylation on arginine residues. Arginine can be monomethylated (MMA(ω)) and subsequently either symmetrically (SDMA(ω)) or asymmetrically dimethylated (ADMA(ω)). Type I, II and III indicate the type of the enzyme that carries out each reaction (see next chapter)(Yang and Bedford, 2013).

The addition of each methyl group changes the residue's shape and removes a potential hydrogen bond donor. Consequently, the post-translational methylation of arginine residues has been shown to modify protein function by regulating protein-

protein interactions, both negatively and positively.

1.3.2 Protein aRginine MethylTransferases (PRMTs)

In mammals, arginine methylation is a procedure as common as phosphorylation, acetylation and ubiquitination (Larsen et al., 2016). However, in contrast to e.g. phosphorylation, arginine methylation of proteins is catalyzed by only a small family of highly related enzymes, the so-called *Protein aRginine MethylTransferases* (PRMTs) (Bedford and Clarke, 2009) (Fig. 1.3) and maybe a few unrelated enzymes such as the rat NDUFAF7 protein (Zurita Rendón et al., 2014).

PRMT (locus)	Domain structures of human enzymes*	Function	Family	Primary substrates‡
PRMT1 (19q13.3)	1 316 a b c d e	Transcription activation, signal transduction, RNA splicing and DNA repair	Type I	H4R3, MRE11, 53BP1 and SAM68
PRMT2 (21q22.3)	1 SH3 domain 433 a b c d e	Transcription regulation	Type I	H3R8
PRMT3 (11p15.1)	1 Zn finger 531 a b c d e	Ribosomal homeostasis	Type I	RPS2 and p53
CARM1 (19p13.2)	1 608 a b c d e	Transcription activation, RNA splicing, cell cycle progression and DNA repair	Type I	H3R17, AIB1, p300, CBP and RNA Pol II CTD
PRMT5 (14q11.2)	1 637 a b c d e	Transcription repression, signal transduction and piRNA pathway	Type II	H3R8, H4R3, E2F1, p53, EGFR and CRAF
PRMT6 (1p13.3)	1 375 a b c d e	Transcription regulation	Type I	H3R2 and H2AR29
PRMT7 (16q22.1)	1 692 a b c d e a b c d	Male germline gene imprinting	Type II and type III	H4R3, H2AR3 and H3R2
PRMT8 (12p13.3)	myr 1 394 a b c d e	Brain-specific function	Type I	Unknown
PRMT9 (4q31.23)	1 TPR 843 a b c d e a b c d e	Unknown	Not classified	Unknown

Fig. 1.3: The Protein aRginine MethylTransferase (PRMT) family. The mammalian PRMT family consists of 9 enzymes that all harbor five conserved domains: I, post-I, II, III and a THW loop. The vertical dark blue lines represent PRMT domains with good sequence similarity to the indicated PRMT motif. In contrast, red vertical lines indicate poor sequence similarity to the indicated PRMT motif. a: domain I, b: domain post-I, c: domain II, d: domain III, e: THW loop and TPR: tetratricopeptide repeat. Some examples of primary substrates are presented for each molecule, as well as their type and function (modified from Yang & Bedford 2013).

In humans, the PRMT family consists of nine members (for a detailed description of the individual members, see chapter 1.4 below). The relation between these proteins was revealed by sequence homology. All PRMTs contain five conserved domains called I, post-

I, II, III (common also to other seven-beta strand methyltransferases) and a THW (threonine, histidine, tryptophan) loop. PRMTs catalyze the transfer of a methyl group from S-adenosyl-methionine (SAM) to the guanidino nitrogen atoms of arginine. This reaction results in the formation of methylarginine and S-adenosylhomocysteine (AdoHcy). Based on the end product of their reaction, PRMTs are classified into four different types. Type I, II, and III methylate the terminal (or ω) guanidino nitrogen atoms. While type III produces only monomethylarginine, type I and II form asymmetrically or symmetrically dimethylated arginine, respectively, via a monomethylated intermediate. Type IV is found in yeast, and is responsible for the methylation of the internal (δ) guanidino nitrogen atom. Interestingly, yeast harbors only one PRMT of each type I and type II, whereas in humans type I PRMTs are present as six distinct enzymes (PRMT1, PRMT2, PRMT3, PRMT4, PRMT6 and PRMT8), and type II as two enzymes (PRMT5 and PRMT9). Type III enzymes only catalyze the formation of monomethylargine, but do not process further to dimethylarginine (Yang et al., 2015). In humans, PRMT7 is the only type III methyltransferase, and histones are its only substrates known today (Feng et al., 2013).

PRMTs (except for PRMT4) methylate arginines preferentially in arginine-glycine (RGG) rich motifs that often function in both nucleic acid binding and mediating protein-protein interactions (Thandapani et al., 2013). Surrounding glycine residues increase the flexibility of the area and gives access to the active site of PRMTs. Exception in this rule is PRMT4, which prefers arginine neighboring proline, glycine and methionine (PGM) rich motifs (Yang and Bedford, 2013) and PRMT7, which prefers RxR motifs in a lysine-rich environment (Feng et al., 2013).

The methylated arginines interact mainly with Tudor domains of proteins (Gayatri and Bedford, 2014). The Tudor domains of SMN (Survival of motor neuron), SPF30 (Splicing factor 30), and TDRD1/2/3/6/9/11 (Tudor domain-containing protein) are some examples of interaction partners of methylarginine (Tripsianes et al., 2011). There are approximately 36 human proteins with a Tudor domain. On the other hand, there are over 2000 proteins containing RGG motifs. This suggests that other methylarginine-interacting domains are likely to be discovered.

Members of the PRMT family are involved in a multitude of cellular processes. They can act as transcriptional co-activators or co-repressors by methylating transcription

factors, other co-activators and histones, thereby modulating their interaction with other proteins and DNA, or their subcellular localization. PRMTs also methylate splicing factors, which can regulate the formation of alternative splicing products by promoting exon skipping in an enzyme-dependent manner. Moreover, PRMTs have been implicated in DNA repair due to the methylation of Mre11 that participates in the repair of DNA double strand breaks by homologous recombination, and also modulates the p53 response. Finally evidence has been provided that PRMTs are involved in signal transduction through protein-protein interactions, mRNA translation and protein stability (Auclair and Richard, 2013; Yang and Bedford, 2013) (Fig. 1.4).

The existence of arginine demethylases is still under debate. Recently it was shown that some lysine demethylases (KDM3A, KDM4E, KDM5C) also possess arginine demethylation activity *in vitro* (Walport et al., 2016). These new observations need to be further investigated in order to clarify the dual role of these enzymes.

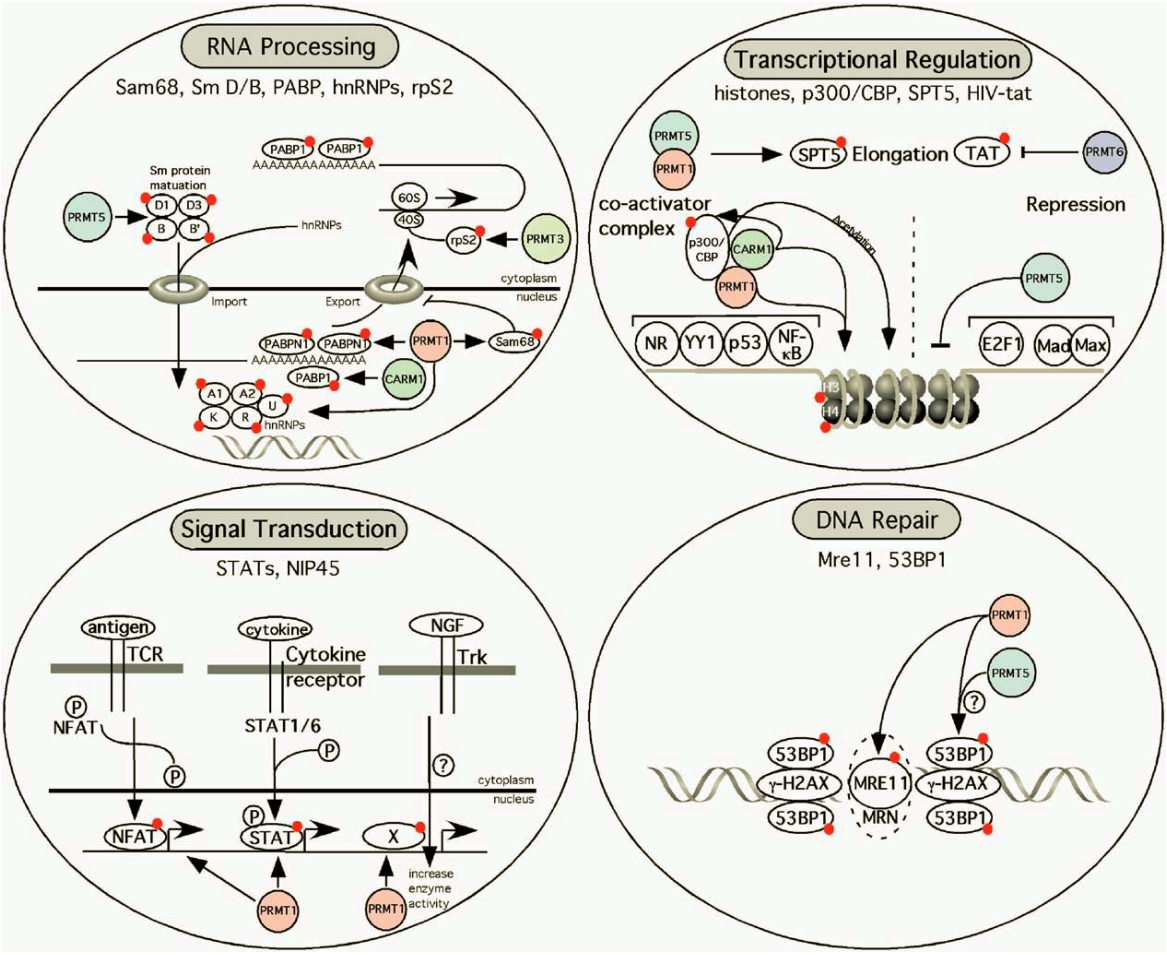


Fig. 1.4: An overview of some cellular processes regulated by arginine methylation. (Bedford and Richard, 2005).

1.3.3 Transcriptional regulation by PRMTs

The activity of PRMTs is regulated by post-translational modifications, either on them or on their substrates, binding to regulatory proteins, microRNAs, as well as through their subcellular localization. In a regulated manner, PRMTs act as activators or repressors of gene transcription by modifying histones and transcription factors. Some examples of transcriptional regulation of gene expression by arginine methylation are depicted in figure 1.5. PRMT1 and PRMT4 mainly represent co-activating PRMTs. The methylation of H4R3me2a by PRMT1, for example, leads to recruitment of p300/CBP-associated factor complex that results in acetylation of histone H3 at lysines 9 and 14 and thus activating gene expression by binding to transcription factors (Bedford and Clarke, 2009). This crosstalk between H4R3me2a and H3K9/14Ac was reported in various biological processes (Li et al., 2015). In a similar way, PRMT4 (CARM1) methylates H3R17, H3R26, and H3R42, which subsequently also activate the affected genes (Yang and Bedford, 2013).

In contrast to PRMT1 and PRMT4, PRMT5 and PRMT6 act as co-repressors of gene activity. PRMT5 methylates H4R3me2s and H3R8me2s, while PRMT6 methylates H3R2me2a. H4R3me2s methylated by PRMT5 is recognized by DNA methyltransferase DNMT3a, which is recruited and methylates neighboring DNA to further repress the expression of genes (Zhao et al., 2009). PRMT6 can repress gene expression by methylation at H2R29me2a (Waldmann et al., 2011). However it can also activate transcription in a similar way to PRMT4/CARM1 through methylation of H3R42 (Casadio et al., 2013).

PRMTs also regulate transcription via methylation of non-histone substrates. For example, PRMT1 regulates TNFα/NF-κB signaling via methylation of p65/RelA R30, which prevents its binding to DNA (Reintjes et al., 2016). Moreover, in the present doctoral thesis we provide evidence that Lamin B Receptor (LBR) is a PRMT1-substrate *in vitro*. LBR is a protein of the inner membrane of the nuclear envelope, which is known to anchor heterochromatin to the nuclear periphery. It is speculated that it mediates the interaction between chromatin and lamin B. It also interacts with the long non-coding XIST RNA in mouse cells, and potentially assists in spreading the XIST across X chromosome and achieve X inactivation. For yet unknown reasons and with unknown consequences,

PRMT1 methylates LBR in a phosphorylation-dependent manner.

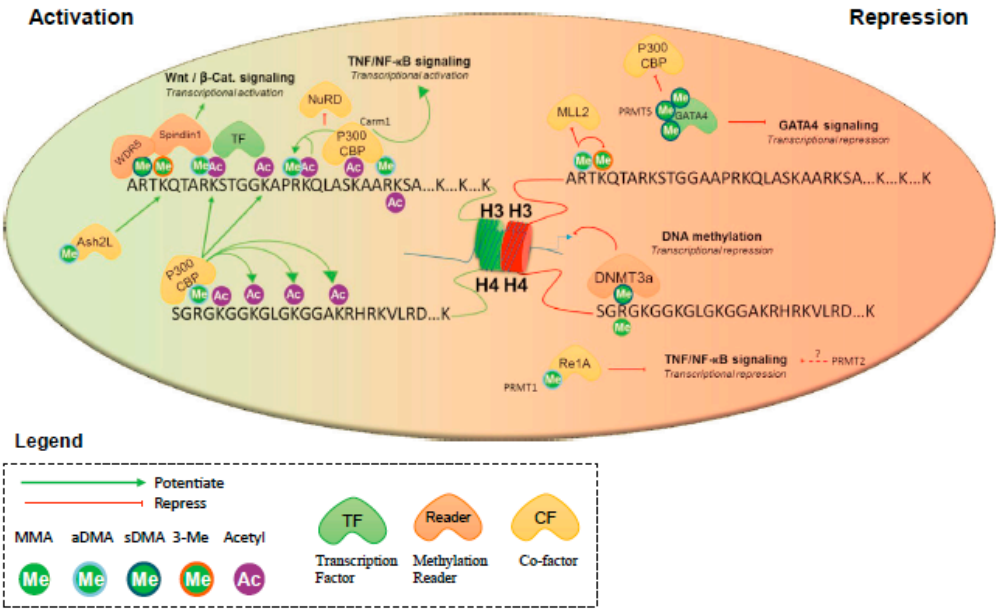


Fig. 1.5: Arginine methylation and transcriptional regulation. Arginine methylation activates (left) or represses (right) gene expression by potentiating or inhibiting the interaction with other factors, which further enhance the effect by following post-translational modifications. Some examples of modifications in histones H3 (upper part) and H4 (lower part) as well as in non-histone substrates, involved in various signaling pathways, are presented(Blanc and Richard, 2017).

1.4 Members of PRMT family

1.4.1 Type I PRMTs

The six Protein Arginine Methyltransferases of type I are the most abundant PRMTs in humans, and are briefly compared in the following paragraphs.

PRMT1 was discovered by its sequence similarity to the yeast arginine methyltransferase Hmt/Rmt1. It is the predominant mammalian type I enzyme, being responsible for more than 85% of all protein arginine methylation (Lin et al., 1996), and is an essential protein because PRMT1-knockout mice die shortly after implantation (Pawlak et al., 2000). Currently, seven splicing isoforms of PRMT1 are known, some of which are cell-type specific, whereas others are ubiquitously found in all cell types. This enzyme will be further described in one of the next chapters, as it is one of the molecules studied in this thesis.

PRMT2 was identified as the first relative to PRMT1 (Katsanis et al., 1997). PRMT2 was long considered enzymatically inactive, but our group has shown that it has a weak

activity (Herrmann and Fackelmayer, 2009). This finding was later independently confirmed. PRMT2 knockout mice are viable and have no detectable defects (Yoshimoto et al., 2006). PRMT2 is the only member of the PRMT family that contains an SH3 domain (Scott et al., 1998); these domains are usually involved in protein-protein interactions because they recognize and bind proline-rich motifs on other proteins. Functionally, it was shown to act as an androgen and estrogen receptor co-activator (Meyer et al., 2007), and can also promote apoptosis by blocking the export of IκB-α from the nucleus (Ganesh et al., 2006). Unpublished results from our group show that PRMT2 co-localizes with known markers of splicing speckles in the nucleus, and a yeast two-hybrid screen for PRMT2-interacting proteins has identified several proteins involved in RNA metabolism.

PRMT3 is predominantly localized in the cytoplasm and contains a zinc finger in its N-terminus, which is unique among PRMTs and is responsible for its substrate specificity (Frankel and Clarke, 2000; Tang et al., 1998). This motif can be involved in substrate recognition or DNA binding of the nuclear fraction of the enzyme (Frankel and Clarke, 2000). One of its known substrates is the ribosomal protein S2, which is part of the small subunit of ribosomes. In this case, the Zn-finger has been shown to be necessary for binding to S2 protein, recruiting PRMT3 to the ribosomes. Mouse embryos with disrupted PRMT3 are small, but they can survive after birth and achieve normal size in adulthood (Swiercz et al., 2005).

PRMT4 is also called CARM1 (Co-activator Associated aRginine Methyltransferase 1) because of its transcriptional co-activator activity. It methylates substrates very distinct from those of PRMT1 and other type I PRMTs, because it does not recognize the glycine-arginine motifs that are typical targets of these PRMTs (Lee and Bedford, 2002). PRMT4 is known to methylate steroid receptor co-activators and other transcriptional co-activators, thereby regulating gene expression (Chen et al., 1999). It also methylates histone H3 and contributes to the epigenetic histone code, being involved in epigenetic reprogramming of early embryos (Bedford and Richard, 2005). PRMT4 knockout mice die perinatally because the lungs cannot inflate after birth due to defects in apoptotic pathways required for organogenesis (Yadav et al., 2003). PRMT4 has been shown to be overexpressed in certain forms of cancer, similar to some isoforms of PRMT1 (El Messaoudi et al., 2006; Hong et al., 2004). Our group has shown that PRMT4 shuttles

between the cytoplasm and the nucleus during the cell cycle, and that it accumulates in the nucleus after DNA damage (Herrmann et al., 2009). This is compatible with other reports that suggested a role of PRMT4 in the p53 response (Yadav et al., 2008; Zhao et al., 2008). CARM1 is involved in regulating pre-mRNA splicing (Cheng et al., 2007; Kuhn et al., 2011) and its expression is regulated by various microRNAs including miR-181c (Z. Xu et al., 2013), miR-223 (Vu et al., 2013), and miR-15 (Liu et al., 2014).

PRMT6 is the only PRMT family member that contains a nuclear localization signal and is exclusively found in the cell nucleus. It was the first PRMT where the ability of automethylation was described (Frankel et al., 2002); we have later shown that PRMT4 has the same ability, and others have found the same for PRMT8. It has been reported that automethylation increases the stability of the molecule (Singhroy et al., 2013). The substrate pattern of PRMT6 is very similar to that of PRMT1, and it methylates the same GAR motifs, e.g. in histones H3 and H4 (Lee et al., 2004). The activity of the enzyme is mainly associated with transcriptional repression since it methylates arginine 2 of histone H3 (H3R2me2a) (Neault et al., 2012). It also methylates and stimulates Polymerase β and thereby regulates the DNA base excision repair (El-Andaloussi et al., 2006). Finally, PRMT6 has been shown to methylate a number of HIV proteins, which motivates research in the direction of medical applications (Invernizzi et al., 2007).

PRMT8 was identified bioinformatically through its high sequence homology to PRMT1. The only major difference between the two proteins is that PRMT8 possesses a unique N-terminus, which contains a myristoylation motif on glycine 2. Myristoylated PRMT8 is tethered to the plasma membrane, which makes it the only membrane-associated PRMT (J. Lee et al., 2005). However, it was suggested that different start methionines could be used *in vivo*, resulting in a non-myristoylated form of the enzyme that is localized in the nucleus (Kousaka et al., 2009). One important aspect of PRMT8 is that it is exclusively expressed in neurons of the central nervous system (Kousaka et al., 2009; J. Lee et al., 2005; Taneda et al., 2007). A more detailed description of PRMT8 is provided below, since it is one of the molecules studied in this thesis.

1.4.2 Type II PRMTs

PRMT5 is the major type II mammalian enzyme member of the family responsible for

the majority of the symmetrically dimethylated arginine in the cell. PRMT5 forms a heterooctameric complex with MEP50 (methylosome protein 50) *in vitro*, which is required for activity. MEP50 is responsible for the substrate specificity and interaction partners of PRMT5 (Antonysamy et al., 2012). However, the necessity of MEP50 *in vivo* remains unclear (Tee et al., 2010). PRMT5 methylates its substrates in a distributive manner by releasing the monomethylated product before the second methylation step (Wang et al., 2014). It can be localized in different complexes both in cytoplasm and the nucleus. In the nucleus, it binds to COPR5 (cooperator of PRMT5) and methylates histones H3 and H4. However, it prefers H4R3 over H3R8 (Lacroix et al., 2008). COPR5 is also responsible for PRMT5 transcriptional co-repressing properties. PRMT5 binds to the hSWI/SNF ATP-dependent chromatin remodeling proteins and functions as a transcriptional co-activator (Dacwag et al., 2007). It also associates with regulators of transcriptional elongation (Liu et al., 2007).

PRMT9 (4q31) is a product of a gene at human chromosome 4q31 and was discovered due to its homology with the other family members (Cook et al., 2006). This protein is designated with its human chromosomal location because the term PRMT9 had previously been used for the product of the FBXO11 gene, which was later shown not to belong to the PRMT family (Cook et al., 2006). It has recently been shown to be a type II PRMT that is not redundant with PRMT5 (Hadjikyriacou et al., 2015; Yang et al., 2015). It is similar in structure to PRMT7, which is an “intramolecular dimer” containing two AdoMet-binding motifs and active centers with conserved sequences in the double E-loop important for substrate specificity and activity (Hadjikyriacou et al., 2015; Jain et al., 2016). It contains two tetratricopeptide repeats (TPR) motifs, which are often found to be involved in protein-protein interactions (Bedford, 2007). It has been shown that mutations in the catalytic domain can change PRMT9 activity from type II to type III enzyme (Jain et al., 2016).

1.4.3 Type III PRMTs

PRMT7 was first identified in a genetic screen for susceptibility to chemotherapeutic cytotoxicity (Gros et al., 2003). It is the one of two PRMTs that contain two AdoMet-binding motifs (Miranda et al., 2004). Different laboratories have classified PRMT7 in two

different enzymatic classes. PRMT7 was classified as type III enzyme because it was shown to catalyze the formation of MMA, but not DMA (Miranda et al., 2004). In contrast, others showed that PRMT7 could catalyze the formation of SDMA in particular substrates. Hence, it was classified as a type II enzyme (Lee et al., 2005). Today, PRMT7 is considered to be a type III enzyme, which produces monomethylated arginine. Another reason for classifying PRMT7 as a type III enzyme is that it preferentially methylates RxR motifs, a property which differentiates it from the main enzymes of type I and II (PRMT1 and PRMT5, respectively). The connection between type III and type I and II PRMTs is not yet understood. The enzyme acquires a double E-loop necessary for its activity and substrate specificity. Mutations in two particular residues in the E-loops switch PRMT7 to type I or II class of enzymes (Debler et al., 2016; Jain et al., 2016). In contrast to type I and II enzymes, which create dimers, and due to its structure, PRMT7 acts as a monomer that acquires a homodimer-like structure with two catalytic domains both essential for its activity. PRMT7 activity is associated with either resistance or sensitivity to DNA-damaging agents (Gros et al., 2006; Verbiest et al., 2008) and sensitivity of the kidney to damage caused by certain antibiotics (Zheng et al., 2005). Depletion of PRMT7 leads to up-regulation of genes related to DNA repair machinery (POLD1, POLD2, ALKBH5, and APEX2) (Karkhanis et al., 2012). PRMT7 also interacts with CTCFL protein to “write” histone methylation marks onto chromatin in Imprinting Control Regions (ICR), which then recruits DNA methyltransferases and methylates DNA in the regulatory regions of imprinted genes (Jelinic et al., 2006).

In the present thesis we characterized a newly identified isoform of PRMT1 and showed that it is implicated in lung cancer. In addition, we provide evidence that Lamin B Receptor (LBR) is a novel substrate of PRMT1 *in vitro*. Moreover, we generated new tools for studying the role of PRMT8 in neuronal development and maintenance. For these reasons an extended description of both enzymes is given in the next chapters.

1.5 PRMT1

Protein Arginine Methyltransferase 1 (PRMT1) is the predominant member of the family of protein arginine methyltransferases, which is involved in a large variety of cellular processes. PRMT1 methylates histones and regulates gene expression by acting as

a transcriptional co-activator. Moreover, it methylates RNA binding proteins thus affecting RNA metabolism and DNA damage proteins which are involved in the maintenance of genome stability (Auclair and Richard, 2013). PRMT1 is also involved in signal transduction. It was identified by sequence similarity to the yeast arginine methyltransferase Hmt1/Rmt (Lin et al., 1996) and was the first mammalian protein arginine methyltransferase identified as a single gene product. PRMT1 mRNA is detected in all embryonic and adult tissues analyzed, indicating the importance of the enzyme in cells (Pawlak et al., 2000). PRMT1 accounts for approximately 85% of the asymmetrically dimethylated arginines generated in mammalian cells (Tang et al., 2000). It has a complex pattern of substrate proteins in both the nucleus and the cytoplasm (Herrmann et al., 2005), and preferentially (but not without several exceptions) methylates the arginine in RGG motifs (Thandapani et al., 2013).

PRMT1 knock-out mice die shortly after implantation at day E6.5 indicating the crucial role of PRMT1 in early development (Pawlak et al., 2000). However, embryonic stem cells derived from these mice are viable with decreased levels in protein arginine methylation, suggesting that lack of PRMT1 is not prohibiting for cell viability. *Saccharomyces cerevisiae* cells lacking the PRMT1 homolog Hmt1/Rmt1 are viable (Gary et al., 1996; Henry and Silver, 1996), but they have multiple defects, including the mislocalization of proteins within the cell (Shen et al., 1998) and chromatin silencing defects (Yu et al., 2006).

1.5.1 Structure of the protein

The structure of PRMT1 is divided into four parts: N-terminal, AdoMet-binding (S-adenosylmethionine-binding), β-barrel and dimerization arm (Fig. 1.6). The AdoMet-binding domain has the consensus fold, conserved in other AdoMet-dependent methyltransferases (Cheng and Roberts, 2001; Fauman et al., 1999), but the β barrel is unique to the PRMT family (Zhang et al., 2000). The active site is located between the AdoMet-binding site and the β-barrel. Two negatively charged glutamate residues within these domains are responsible for interacting with the positively charged arginine and are important for the catalysis of the reaction (Zhang and Cheng, 2003).

1.5.2 Oligomerization and mechanism of action of PRMT1

PRMT1 has the ability to form homodimers through its dimerization arm and the outer surface of the AdoMet-binding domain that provides a hydrophobic dimer interface. The homodimers can further oligomerize and form a disulfide bond between two not conserved cysteines (C254) residues in the crystal lattice. However, this higher order multimerization does not occur in the absence of dimerization, and the formation of the disulfide bond is not required for oligomerization (Zhang and Cheng, 2003) since the C254 is not conserved among PRMTs (Fig. 1.7). The di- or oligomerization is a common characteristic among the members of the PRMT family that is crucial for the AdoMet binding and the activity of the enzyme. As it has been shown by other groups and by experiments presented in this thesis, PRMT1 becomes inactive if it lacks the dimerization arm.

The suggested mechanism of PRMT1 action is that the methylation of the substrate is carried out in a processive manner. The ring-like dimer could allow monomethylarginine, the product of the first methylation reaction, to enter the active site of the second molecule of the dimer without releasing the substrate from the ring or replenishing the methyl donor. Two methionine residues, M48 and M155, seem to play an important role in the catalytic activity of the enzyme since if they are mutated, they alter substrate specificity and abolish catalytic activity (Gui et al., 2014).

PRMT1 has a wide range of substrates thus being implicated in many cellular functions. One of the important substrates of PRMT1 is histone H4 that is methylated by the enzyme at arginine 3 (H4R3). This modification is considered to be a marker for transcriptional activation. However, PRMT1 can also be involved in gene repression (Kleinschmidt et al., 2008), indicating that epigenetic regulation is depended on different factors, such as surrounding post-translational modifications. The fine-tuning of epigenetic regulation by PRMT1 also depends on the interaction with other transcription factors and regulatory proteins, which recruit the enzyme to specific genomic loci in order to carry out the methylation of histones. The transcriptional role of PRMT1 is also obvious on non-histone substrates including co-activators, transcriptional elongation factors and splicing factors. Deregulation of such mechanisms can lead to serious diseases including cancer (see below for examples).

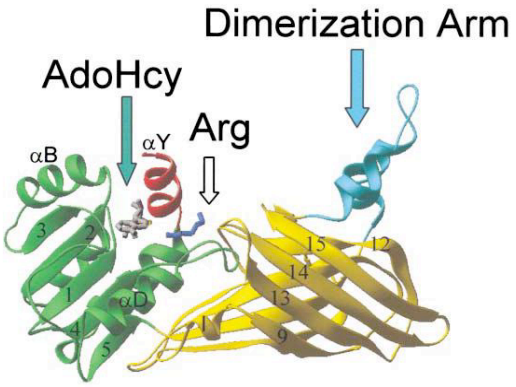


Fig. 1.6: 3D structure of PRMT1 monomer. The four domains of PRMT1 are depicted in different colors. The N-terminal helix is shown in red, the AdoMet-binding domain in green, the β barrel in yellow and the dimerization arm in light blue. The bound arginine is shown in dark blue in the active site, which is located between the AdoMet-binding domain and the β barrel. AdoHcy is shown in grey (modified fromZhang & Cheng 2003).

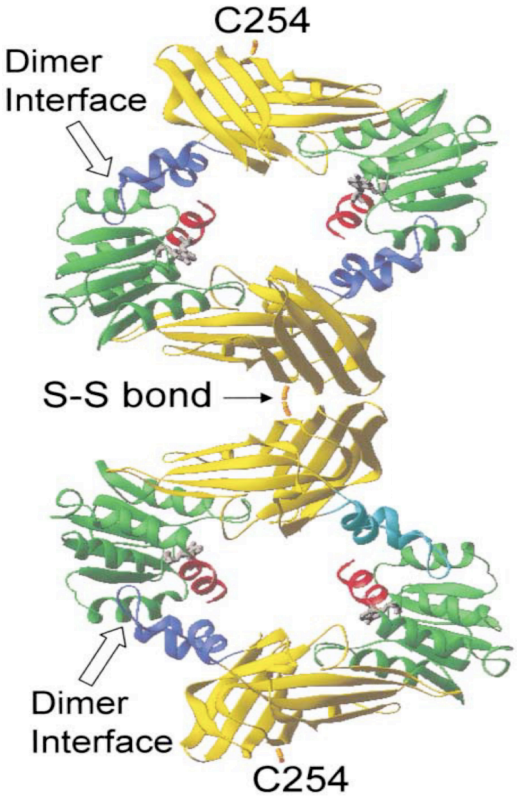


Fig. 1.7: Dimerization and oligomerization of PRMT1. Two ring-like PRMT1 dimers connected via a disulfide bond on cysteine 254. The dimer is formed through the dimerization arm and the outer surface of AdoMet binding domain and they are connected via cysteine 254 creating larger homooligomers (Zhang and Cheng, 2003).

1.5.3 Alternative splicing isoforms of PRMT1

The Prmt1 gene is composed of 12 coding exons. Initial studies on the Prmt1 genomic locus revealed three different transcripts which were products of alternative splicing of exons 2 and 3 (Scott et al., 1998) and the expression of these variants was later confirmed on the mRNA level by another study (Scorilas et al., 2000). Later, the group of Jocelyn Cote, after inspection of upstream sequences and EST comparison, identified additional exons in the Prmt1 gene that could produce at least seven protein isoforms. These differ in the amino-terminal region that is involved in substrate recognition and subcellular localization (Goulet et al., 2007); the rest of the sequence remains identical among variants. Among these variants, variant 4 starts from a new AUG codon placed in a novel

upstream exon (1c). Variants 3 and 5 include alternative exons e2 and e3, which introduce an in-frame stop codon (UGA) and consequently the translation initiates at the next start codon (AUG) in exon 3. The predicted coding regions of the seven prmt1 transcripts excluding the 5' and 3' untranslated regions are composed of 1059 nt (v1), 1113 nt (v2), 1041 nt (v3), 1047 nt (v4), 1026 nt (v5), 975 nt (v6), or 960 nt (v7) and it was predicted that they encode seven polypeptides with molecular mass of 40.5, 42.5, 39.9, 40.1, 39.4, 37.7, and 36.7 kDa, respectively. The combinations of exons, which participate in the formation of each variant, as well as the predicted transcript and protein sizes, are depicted in Figure 1.8

All seven isoforms, designated variants 1-7 (v1-v7), were found to be expressed in a tissue-specific manner. PRMT1v1 is mainly expressed in kidney, liver, lung, skeletal muscle, and spleen, whereas PRMT1v2 is detected in kidney, liver and pancreas. The PRMT1v3 isoform is present in all tested tissues, although at low levels. PRMT1v4-v7 show a higher degree of tissue specificity in their expression patterns. For example, v4 is detectable only in heart, v5 in pancreas, and v7 is predominantly present in the heart and skeletal tissue. Figure 1.9 shows the expression pattern of each variant in different tissues after PCR analysis. The only exception is variant 6, which is not detectable in any of the tissues tested. However, in this thesis is we have shown its expression in LUHMES cells (see results).

Isoforms 1-6 are active, but variants 3 and 4 appear to have lower activity. The only exception is variant 7, which is catalytically inactive; the functional role of this variant remains unknown. The unique amino-terminus of each variant offers the possibility of discrimination between different substrates. For example, Sam68 and SmB are better substrates for v1, whereas hnRNP A1 is better methylated by v5 and v6.

The subcellular localization also differs among variants, and seems to be dependent on the cell type. For example, v3-v6 are distributed evenly between the cytoplasm and the nucleus, v1 and v7 present a more intense nuclear staining, and variant 2 is restricted to the cytoplasm in HeLa cells. Variant 2 acquires a CRM1-dependent leukine-rich nuclear export signal (NES) that regulates its subcellular localization (Fig. 1.10).

In the present thesis we present data showing that the localization of the seven PRMT1 isoforms slightly differs in HEK293 compared to the corresponding published

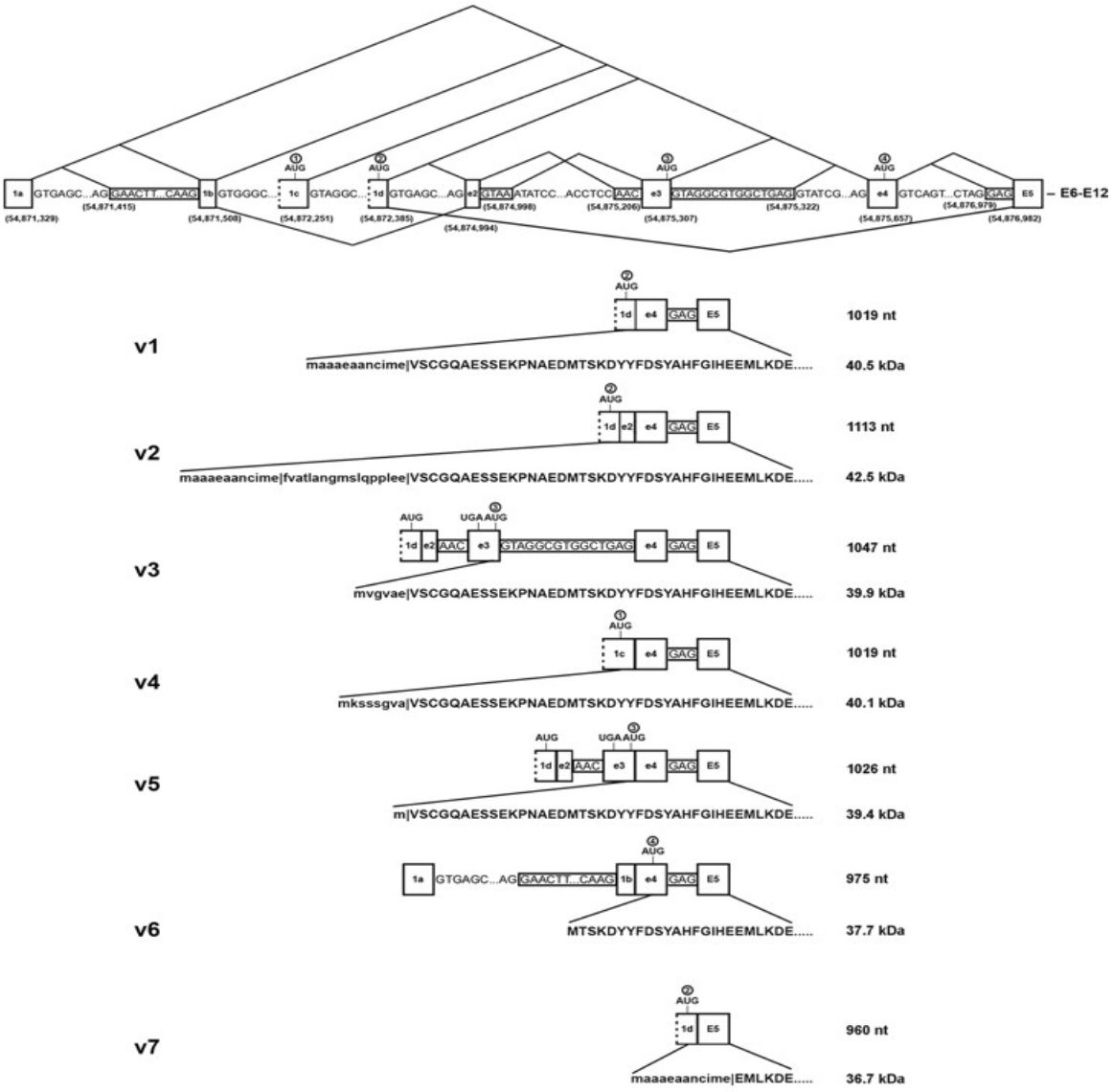


Fig. 1.8: Alternative splicing isoforms of PRMT1. The genomic locus of Prmt1 gene is presented on the top of the figure. Below, different exons are spliced in order to create the seven amino-terminal variants of PRMT1. Exons are drawn as boxes and the sequence of intron boundaries is also shown. "E" indicates constitutive exons and "e" alternative exons. The combinations of the exons as well as the transcription start codons are shown. Predicted sizes of the mRNA and the protein of each variant are given on the right side (modified from Goulet et al. 2007).

localization in HeLa cells.

Another important observation is that the molecular dynamics of PRMT1 variants differ and their mobility is, in some cases, depended on their enzymatic activity. One striking example, published by our laboratory in the past, is the mobility behavior of PRMT1v2. This variant is predominantly localized in the cytoplasm due to the nuclear export signal mentioned above, but it shuttles in and out of the nucleus depending on the methylation status of its substrates (Herrmann et al., 2005). A mutated form of variant 2,

which destroys the enzymatic activity, leads to its accumulation in the cell nucleus. There, it forms a granular pattern and becomes partially immobilized (as shown by fluorescence recovery after photobleaching experiments, FRAP) due to the stable binding on substrates (histone H3, SAF-A) (Herrmann and Fackelmayer, 2009).

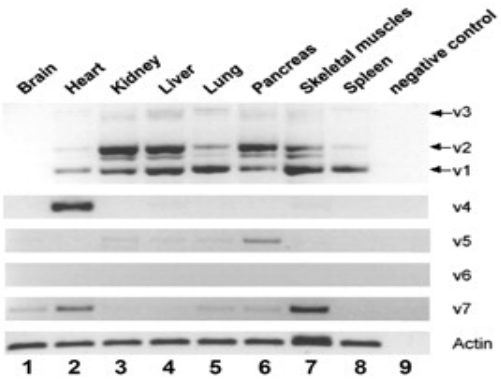


Fig. 1.9: Tissue-specific expression of PRMT1 isoforms. RT-PCR indicated that variants 1-3 are more widely expressed, whereas v4-v7 show higher tissue specificity (Goulet et al., 2007).

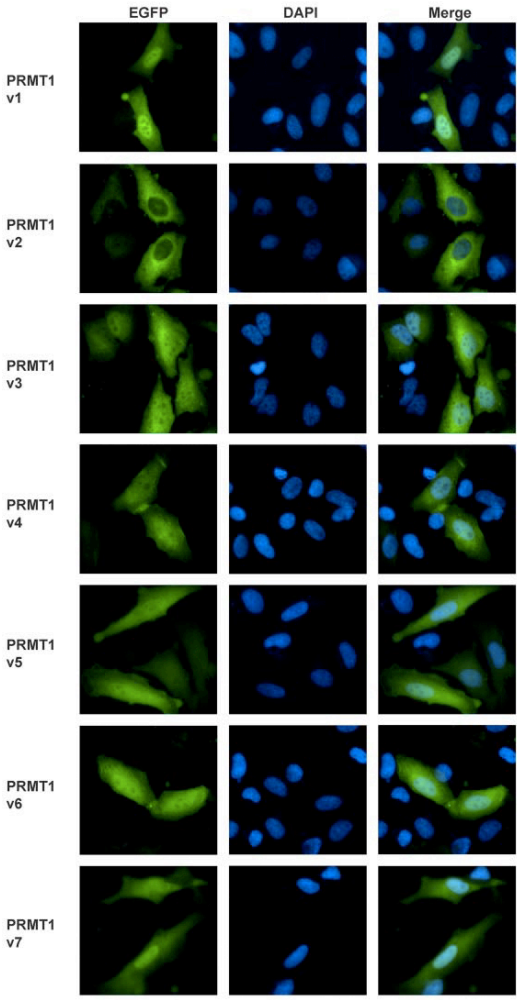


Fig. 1.10: Subcellular localization of PRMT1 variants in HeLa cells (right). HeLa cells were transfected with carboxy-terminal GFP tagged variants of PRMT1 and were fixed after 36 hours. DAPI was used as a nuclear staining. v3-v6 are distributed evenly between the cytoplasm and the nucleus, v1 and v7 present a more intense nuclear staining and variant 2 is restricted in the cytoplasm of the cells due to its CRM1-dependent leukine-rich nuclear export sequence (modified from Goulet et al. 2007)

1.5.4 PRMT1, alternative splicing and cancer

The seven variants of PRMT1 were found to be differently expressed in breast cancer tumors and cell lines; v5 and v6 were detected in specific cancer cell lines but not in normal breast cells. Overall PRMT1 expression was increased by 9,5 fold in human tumor sample when compared with the adjacent normal breast tissue and 14 fold higher among breast cancer cell lines. Expression of v2 was increased 3,5 times in comparison to v1. Moreover, v7 is increased approximately 3 fold in breast cancer cell lines (Goulet et al., 2007). This fact has an impact in the overall arginine-methylated protein levels and

pattern between normal and cancerous breast cells and thus makes PRMT1 variants molecules of high medical interest.

These findings are not the only example of the implication of arginine methylation in cancer. Additional publications support the idea that alterations in arginine methylation status and in the expression levels of PRMTs and their splicing isoforms may represent a driving force leading to the development, progression and aggressiveness of several types of cancer. PRMT1v2 knock-down leads to significant reduction of cell viability and growth in breast cancer cells as a result of apoptosis (Baldwin et al., 2012). On the other hand, overexpression of the same variant increased the growth rate of the same cells, an effect that does not take place after variant 1 overexpression in the same cell line. It is proposed that PRMT1v2 has a key role in the survival of the cells. Moreover, high levels of PRMT1v1 have been correlated with poor patient prognosis and reduced disease-free survival (Mathioudaki et al., 2011). In addition to cell cycle and apoptosis regulation, deregulation of PRMT1 activity contributes to cancer progression by increasing the invasiveness of the cells. For example, PRMT1v2 methylates Axin, increases its stability and thereby decreases β-catenin levels. Depletion of variant 2 leads to decreased invasiveness of MCF-7 cells in a Matrigel barrier. Overexpression of the same variant in the same cell line leads to increased motility of the cells (Baldwin et al., 2012). These are only few examples that highlight the importance of the different variants in cancer development and progression, and make a strong point why we should study them as independent molecules.

Under normal conditions, alternative splicing provides the cell with the possibility to create variable proteins for different purposes from a single gene. In that way, the structural transcript variation and proteome diversity are increased. In cancer, there are indications that splicing is deregulated resulting in both functional and non-functional products, and this might be one of the reasons of cancer progression. The alterations in splicing in cancer cells can be associated with changes either *in cis*-acting splicing regulatory sequences or *in trans*-acting splicing factors. Genes which are related to cell cycle regulation, apoptosis, growth suppression and genome stability have been found to be alternatively spliced in different cancer types similarly to PRMT1. It is being discussed whether these splicing variants can be used as diagnostic markers or drug-target molecules for therapeutic approaches.

Other studies, which did not discriminate among the variants, have shown an involvement of PRMT1 in cancer in different ways. For example, PRMT1 is implicated in cell proliferation and viability as well as in cell cycle progression. Depletion of PRMT1 leads to G0/G1 cell cycle phase arrest, and thus in decreased cell proliferation in osteosarcoma, breast, bladder and lung cancer cell lines (Le Romancer et al., 2008; Yoshimatsu et al., 2011; Yu et al., 2009). In addition, PRMT1 affects the activation of estrogen receptor (ERα) by methylation at arginine 260, resulting in cytoplasmic retention of ERα and its interaction with Src focal adhesion kinase (FAK) and the regulatory subunit of PI-3 kinase (p85). These three proteins are involved in cancer cell signaling which promotes cell survival and invasiveness (Hernandez-Aya and Gonzalez-Angulo, 2011; Weber et al., 2011; Zheng et al., 2011). In a similar way, PRMT1 affects the activity of other receptors by methylating them directly or in an indirect way by methylating receptor binding proteins. An example is TGFβ signaling, where PRMT1 methylates Smad6 (Xu et al., 2013). Except of the large numbers of PRMT1 substrates, which participate in cell cycle regulation, there are others, which belong to apoptotic pathways, such as apoptosis signal regulating kinase 1 (ASK1). Methylation of ASK1 inhibits its activity and thus promotes its interaction with the negative regulator thioredoxin, making breast cancer cells resistant to treatment with paclitaxel (Cho et al., 2012). Other substrates of PRMT1 that are involved in transcription (forkhead box protein 1, FOXO1), telomere stability (telomeric repeat binding factor2, TRF2) and DNA repair (MRE11 and p53 binding protein 1, 53BP1) as well as the tumor suppressor gene BRCA1 might lead to cancerous conditions and tumor formation in cases of deregulated PRMT1 activity.

PRMT1 has been shown to correlate with epithelial to mesenchymal transition (EMT) in lung cancer cells. EMT is an essential process for cancer progression and metastasis. Overexpression of PRMT1 in a non-transformed bronchial epithelial cell line results in the induction of EMT, by decreasing E-cadherin, and increasing N-cadherin expression levels at the same time. Knock-out of PRMT1 gives the opposite effect, thus reducing migration and invasion of the cells. The suggested mechanism describes that PRMT1 overexpression suppress E-cadherin expression through an intermediate EMT-inducing transcription factor, Twist1, which is responsible for E-cadherin repression only when is methylated at arginine 34 (R34). It is speculated that methylation of Twist1 at R34 drives the protein in

the nucleus, and thereby allows it to transcriptionally suppress E-cadherin (Avasarala et al., 2015).

As described before, PRMT1 is up-regulated in some types of cancer. It also epigenetically modifies histone H4 at arginine 3 (H4R3me2a), which is associated with transcription activation. It has been shown that the methylation status of H4R3 correlates with increasing tumor grade in a positive way and could be used as a prediction marker of prostate cancer reappearance (Seligson et al., 2005).

In conclusion, these and additional studies point to the crucial role of PRMT1 in cancer through various pathways and epigenetic regulation. The existence of different isoforms adds an extra level of complexity and creates challenging opportunities for the design of new anticancerous drugs targeting PRMT1.

In the present PhD thesis, I describe the natural existence of a new splicing isoform of PRMT1 that lacks exons 8 and 9. These exons code for the dimerization arm of the enzyme. This is the first time where variation in the catalytic core of PRMT1 is observed. The expression levels of the new variant (designated Δarm) are correlated with the aggressiveness of lung cancer cell lines. Moreover, I present data indicating that Lamin B Receptor (LBR) is a novel PRMT1 substrate and its methylation is depended on the phosphorylation status of the surrounding serine residues in the RS domain of the protein (see results).

PRMT	Cancer type	Role(s) in cancer
PRMT1	Breast cancer, Lung cancer, Colon cancer, Bladder cancer, Acute myeloid leukemia, Mixed lineage leukemia	Cell proliferation and survival, Transformation, Resistance to DNA damaging agents, Invasion
PRMT2	Breast cancer	Cell proliferation and invasion
PRMT3	Breast cancer	Cell survival
CARM1/PRMT4	Breast cancer, Prostate cancer, Colorectal cancer	Cell proliferation
PRMT5	Lung cancer, Leukemia, Lymphoma, Melanoma, Gastric cancer, Colorectal cancer	Cell proliferation, Transformation, Invasion, Resistance to DNA damaging agents
PRMT6	Lung cancer, Bladder cancer	Cell proliferation
PRMT7	Breast cancer	Resistance to DNA damaging agents
PRMT8	ND	ND
PRMT9	ND	ND

Fig. 1.11: An overview of the involvement of PRMT family members in different cancers. ND: not determined (modified from Baldwin et al. 2014).

1.6 PRMT8

PRMT8 (originally also called Hrmt1l4 for “hnRNP methyltransferase-like 4” or Hrmtl3) was identified when the human genome was scanned for novel open reading frames (ORFs) that harbor conserved methyltransferase domains. The reading frame identified in

this way is closely related to PRMT1, with more than 80% identity between the two full-length sequences. However, the difference is attributed to a unique amino-terminus of around 60 amino acids, which is not found in any of the known splicing isoforms of PRMT1 – the remainder of the two enzymes are more than 95% identical (Fig. 1.12). Bioinformatic analysis of the genes encoding these proteins has shown that they are localized on different chromosomes. In the human genome, PRMT1 is located on chromosome 19, while PRMT8 is located in the subtelomeric region of chromosome 12. In mouse, PRMT1 maps to chromosome 7, and PRMT8 to chromosome 6. Interestingly, the intron/exon boundaries of the genes encoding PRMT1 and PRMT8 are identical in both the organisms, suggesting a recent duplication of an ancestral locus. Like PRMT1, PRMT8 is an active arginine methyltransferase which displays type I activity, forming N_G-monomethylarginine and ADMA. The high homology between PRMT1 and PRMT8 explains why both proteins can form heterodimers, and catalyze the methylation of the same substrates *in vitro*.

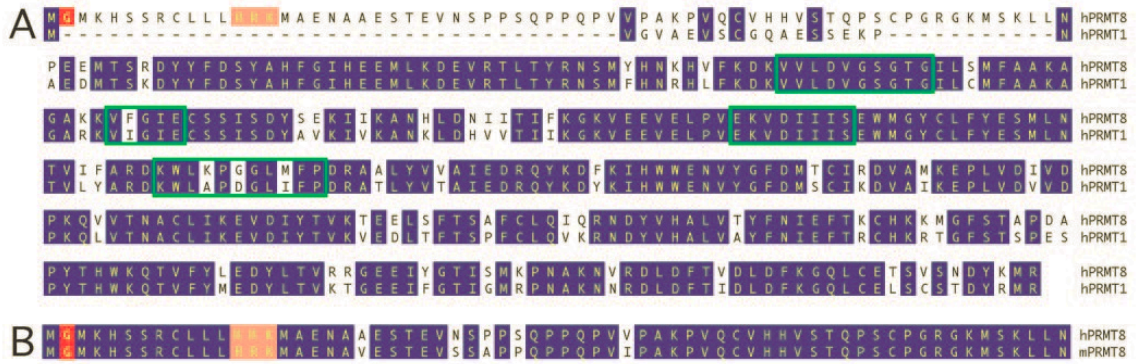


Fig. 1.12: Alignment of hPRMT1 and hPRMT8 protein sequences. (A) Alignment of human PRMT1 and PRMT8 protein sequence. The two proteins are 80% homologous but PRMT8 obtains a unique amino-terminus of 60 amino acids. Matching amino acids are highlighted in blue. Green boxes represent contain the three typical domains of a PRMT (I, post-I, II, III). (B) Alignment of the human and mouse PRMT8 amino-terminus. The glycine 2, which is myristoylated, is highlighted in red. Thegroup of basic amino acids close to the myristoylation motif is marked with pink (from J. Lee et al. 2005).

One fundamental difference between PRMT1 and PRMT8 is the tissue specific expression pattern. PRMT8 is largely expressed in brain and spinal cord (central nervous system, CNS), but is not detected in other tissues. PRMT1, on the other hand, is ubiquitously expressed in all cell types and tissues. The CNS-specific expression of PRMT8 was confirmed by three independent methods. First, Northern blots with RNA from a large variety of human tissues gave positive signals only with RNA from brain. Second,

ESTs analysis identified PRMT8 sequences almost exclusively in brain cDNA libraries; a low number of transcripts was also detected in mouse retina and in human lung, testes, nasopharynx and kidney. Third, when neural precursor cells were isolated by FACS from Sox1-GFP knock-in mouse embryos, microarray analysis revealed PRMT8 as one of the transcripts that are selectively expressed in neural progenitor cells from mid-gestation mouse embryos (Lee et al., 2005). Later, *in situ* hybridization demonstrated that PRMT8 is expressed in the central nervous system (CNS), and more specifically in the somatosensory and limbic system as well as in certain subtypes of motoneurons (Fig. 1.13). Positive signals were exclusively found in neurons, with no signals being observed in glial cells (Taneda et al., 2007).

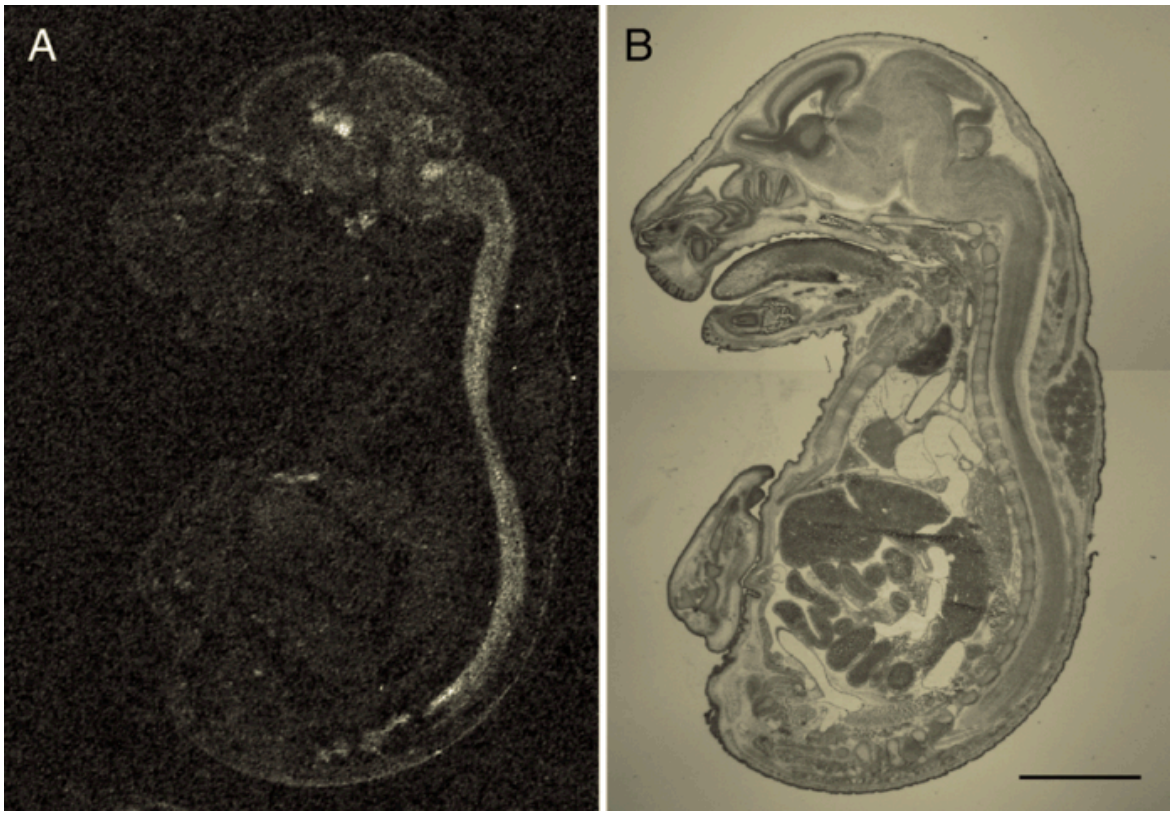


Fig. 1.13: Expression of PRMT8 mRNA in a mouse embryo. (A) Sagittal sections obtained from a d16.5 mouse embryo were hybridized with antisense probe specific for PRMT8. (B) Brightfield image of adjacent section illustrated in (A) stained with hematoxylin/eosin. Note that PRMT8 is exclusively expressed in the brain and spinal cord. Scale bar=3 mm (from Taneda et al. 2007).

As mentioned above, PRMT8 harbors a unique amino-terminal region that contains a myristoylation motif on glycine 2, which is myristoylated after the post-translational removal of the initiating methionine. This myristoylation tethers PRMT8 to the plasma membrane (Fig. 1.14), and mutation of glycine 2 to alanine leads to a cytoplasmic and

nuclear localization of the enzyme. Interestingly, the myristoylation consensus motif (MGXXX(S/T)) is conserved between mouse, human and puffer fish. This myristoylation motif is surrounded by basic amino acids, which electrostatically enhance the interactions between the myristoyl group and membrane lipids (J. Lee et al., 2005).

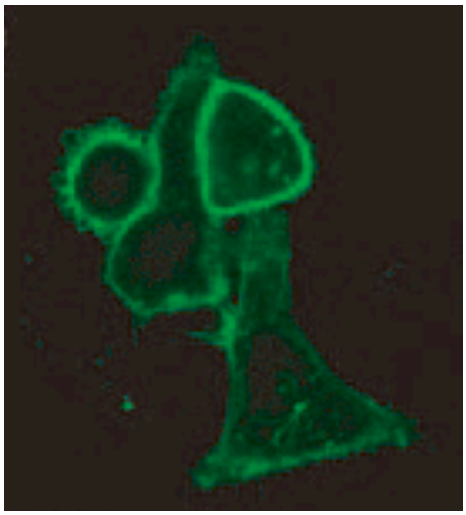


Fig. 1.14: Subcellular localization of PRMT8. PRMT8:GFP localizes to the plasma membrane of transfected HeLa cells, through its myristoylated amino-terminus and the surrounding basic amino acids. (fromJ. Lee et al. 2005)

Lee et al. proposed a full coding sequence of PRMT8 based on bioinformatic data, using the 5'-most in-frame AUG codon as the translation initiation point. Thus the predicted protein contains the myristoylation site and is tethered to the plasma membrane as described above. However, the findings of Kousaka et al. (2009) based on immunohistochemistry experiments with mouse brain, suggested a different localization of the protein. PRMT8 was found mainly in the nucleus of neurons of the central nervous system, rather then on the plasma membrane. In order to explain this observation, they suggested that PRMT8 may make use of alternative start codons. To investigate this point in more detail, they compared the size of the endogenous PRMT8 protein with the size of three recombinant proteins with different in-frame start codons, the one proposed by Lee et al., and two other methionines downstream. Their results demonstrated that the major form of the endogenous PRMT8 was comparable in size to the protein expressed from the third AUG codon, methionine 16. Based on these data, they concluded that the resulting protein does not contain the N-myristoylation site from the 5'-most start codon, and would therefore not be associated with the plasma membrane. Also this methionine is conserved among the investigated species (mouse, rat, and human) (Kousaka et al., 2009).

One surprising result of the first characterization of PRMT8 was the low enzymatic activity. As PRMT8 is almost identical to PRMT1 in the domains responsible for catalytic activity, a similar specific activity would have been expected. Based on this observation, Sayegh and co-workers studied if and how the unique amino-terminal domain of PRMT8 could affect its enzymatic activity. They hypothesized that the amino-terminal domain of PRMT8 might sterically inhibit the methyltransferase activity. Indeed, they demonstrated that the proteolytic removal of the first 60 amino acids of PRMT8 caused a 10-fold increase of the enzymatic activity, reaching enzymatic activity levels comparable to PRMT1. On the basis of these results, the authors suggest that the amino-terminus may fold back onto the enzyme to autoinhibit its activity, and may be displaced by interaction with one or more physiological inducers (Sayegh et al., 2007).

PRMT8 is automethylated near its amino-terminal domain *in vitro* and *in vivo*, with an asymmetrically dimethylated arginine residue at position 73 and a monomethylated arginine at position 58 (Sayegh et al., 2007). Automethylation appears to be incomplete, as methylated and unmethylated residues are found at the same time in different protein molecules. The ability of PRMT8 to methylate its own amino-terminus points towards a possible mechanism for regulation of the enzyme. However, the physiological function of automethylation of PRMT8 has not yet been investigated in detail. It is conceivable that automethylation affects interaction with other proteins, such as other members of the PRMT family or unrelated proteins. In fact, PRMT8 harbors two proline-rich domains that can interact with SH3 domains of other proteins. Specifically, our laboratory has shown that PRMT8 interacts with PRMT2, the only member of the PRMT family that contains an SH3 domain (Herrmann and Fackelmayer, 2009). In addition, as PRMT8 can homodimerize and heterodimerize with PRMT1, automethylation might be involved in regulating the formation of enzymatically active PRMT complexes (J. Lee et al., 2005). These mixed oligomers would most probably differ from homo-oligomers with respect to subcellular localization and/or substrate specificity, and would only occur in cells of the CNS where both enzymes are expressed. Crystal structure analysis of hPRMT8 indicated that it forms an octamer in solution. This octameric structure is necessary for proper localization to the plasma membrane and efficient methyltransferase activity (Toma-Fukai et al., 2016).

Recently, new evidence for PRMT8 expression in other cell types than brain came to light. In 2016, Sarah Hernandez demonstrated expression of PRMT8 in induced regeneration competent cells (iRC) and in human embryonic stem cells (hESCs). iRC cells are derived by exogenous addition of human fibroblast growth factor FGF2 to fibroblasts and culture in reduced oxygen concentration (2%) for seven days. These cells show increased proliferative lifespan and increased time to cellular senescence, while lacking the tendency to form tumors when injected into SCID mice, a capability that is characteristic of immortalized and pluripotent cells. The induction of normal fibroblasts to iRC cells leads to up-regulation of PRMT8 (Hernandez and Dominko, 2016). However, whether the detection of PRMT8 in iRC cells is an outcome of the lower oxygen levels that the cells were supplied with (closer to physiological) or a cell specific trait, is not clear yet. It is being discussed whether the lower oxygen levels induce the expression of PRMT8 or if the high levels of oxygen, which are usually used in cell culture conditions, repress its expression. Moreover, there are other publications, which have shown that other PRMTs (1-7) are regulated by hypoxic conditions (Yildirim et al., 2006).

These observations together with the findings from Li et al.(2013) are the first evidence for PRMT8 having a role in development before becoming localized specifically to mature brain tissue. Li and colleagues, as the other groups, detected PRMT8 in adult zebrafish brain but not in other adult tissues. However, Prmt8 mRNA was also present in the zebrafish embryo from 0,25 hours post fertilization (hpf) throughout 72 hpf. By whole mount in situ hybridization, ubiquitous prmt8 expression was detected at early developmental stages from one cell to 12 hpf. Expression of prmt8 in the eyes and brain, especially at the mid-hindbrain boundary and the hindbrain was clearly detected at 24 hpf and the expression in the head regions continued to 60 hpf. Even distribution in the somites was also detected during these periods. The signals were restricted to the somites at 72 hpf, which later faded and were present mainly in the midbrain at 96 hpf. Thus, embryonic Prmt8 expression was shown for first time in zebrafish. Knock down of PRMT8 led to severe phenotypes during various morphogenetic events of the zebrafish embryo (Lin et al., 2013). PRMT8 has also been detected in mouse embryonic stem cells, and increased during their neuronal differentiation. The promoter of the gene is activated by Oct4, Sox2 and Nanog, and PRMT8 expression is down-regulated when Sox2 mRNA is

knocked down by short hairpin RNA (Solari et al., 2016). These findings indicate that methyltransferase activity of PRMT8 might have a critical role in early embryonic stages, in addition to its role in the nervous system.

In a different study, PRMT8 has been found to act together with PRMT1 as a rheostat to integrate retinoid signaling into neuronal specific gene expression governed by retinoids (Simandi et al., 2015). PRMT8 was induced upon retinoic acid treatment of mouse embryonic stem cells and was present only in differentiated neurons with nuclear localization. On the other hand, PRMT1, the protein that is highly homologous to PRMT8, was strongly expressed and its expression remained constant during neuronal differentiation showing a slight increase upon retinoic acid treatment. The model suggested by this group is that only PRMT1 is expressed in early developmental stages and it acts as a selective repressor of retinoic acid induced genes. After retinoic acid induction, PRMT8 is expressed and collaborates with PRMT1 or as a homo-dimer with itself acting as a co-activator that potentiates retinoid response. The same group highlighted the importance of PRMT1 and PRMT8 for the proper function of ion channels after knock down of the two proteins, a result that has to be further investigated (Simandi et al., 2015).

The group of Tanja Dominko demonstrated the expression of a new PRMT8 variant (designated variant 2) in human fibroblasts, indicating human PRMT8 expression outside of the central nervous system (Hernandez and Dominko, 2016). Variant 2 is mostly expressed in normal fibroblasts as well as in iRC cells and is critical for proliferation of human fibroblasts. On the other hand, variant 1 was found mostly in hESCs. The protein product of PRMT8 variant 2 lacks the glycine residue near the N-terminus, which is myristoylated, resulting in the localization of PRMT8 variant 2 in the nucleus (Hernandez et al, unpublished data).

Recently, it was reported that PRMT8 has a dual enzymatic activity. The first one is the known methyltransferase activity. In addition, a conserved catalytic aminoacid sequence which is typical among phospholipases D (PLDs) was identified in PRMT8 sequence. This motif, HxK(x)4D(x)6GG/S (where x represents any amino acid) is referred to as the HKD motif. PRMT8 was found to act as a phospholipase by hydrolyzing phosphatidylcholine to choline and phosphatidic acid (Kim et al., 2015).

Recently PRMT8 was related to cancer. PRMT8 was found to be down-regulated in glioma tissues from patients. Whether this is because PRMT8 is a specific neural marker, which is not expressed in astrocyte-derived tumors or whether the loss of PRMT8 positively affects astrocyte differentiation, resulting in a shift in cell fate commitment, remains to be elucidated. There are findings, which support that loss of PRMT8 leads to marker increase (Cxcr4, Dhfrand, Efemp1) or decrease (Gfra2) related to astrocyte-derived glioma (Simandi et al., 2015). On the other hand, other groups support that PRMT8 is critical for the proliferation of grade IV glioblastomas, since knock-down of this molecule kills glioblastoma line U87MG. The same happened after the knock-down of PRMT8 variant 2 in human fibroblasts leading to the conclusion that PRMT8 knock-down provokes loss of cell proliferation in both non-tumorigenic and tumorigenic cell types (Hernandez and Dominko, 2016).

High levels of PRMT8 expression were also detected in various types of cancer, which normally do not express it, such as breast, glandular, cervical, head and neck, prostate and thyroid cancer. Other types of cancer were found to express moderate-high levels of PRMT8 (colorectal, endometrial, brain, lung, ovarian, pancreatic, skin, testicular and urothelial cancer) and others were found with low-high levels (renal and stomach). An interesting observation was that in some cases PRMT8 expression levels were correlated with patient survival. For example, in patients with breast and ovarian cancer, high PRMT8 expression was correlated with increased patient survival, whereas in gastric cancer, high PRMT8 expression was correlated with decreased patient survival. However, in patients with non-small-cell lung cancer, no significant correlation was identified between patient survival and PRMT8 expression levels. The possibility that PRMT8 expression levels can be used as a prognostic biomarker for cancer or as a therapeutic target is being discussed. However, more experiments must be carried out in to this direction (Hernandez et al., 2017).

The interest of our group lies on the role of PRMT8 in neuronal differentiation and maintenance. We speculate that PRMT8 has an epigenetic role neuronal differentiation and maintenance, by modifying histones and thus activating genes related to neuronal lineage while suppressing others unrelated to neurons. One of the aims of this thesis was to set up new tools for this kind of studies, which they were not available before (see

results) and start a preliminary characterization of the behavior of the molecule.

1.7 LUHMES – an appropriate model for studies on PRMT8

One of our interests was to study the functional role of PRMT8 in neuronal differentiation and maintenance. For this purpose several tools as well as a reliable cellular model needed to be set up in the laboratory. We thus searched for a cell model that can be used to obtain a homogeneous population of human post-mitotic neurons. Rapid and large-scale differentiation models are necessary to investigate the cellular events that drive neuronal differentiation, and allow for studies on the cause of neurodevelopmental disorders. Already existing cell lines such as PC12 (cell line derived from a pheochromocytoma of the ratadrenal medulla), Neuro2a (mouse, albino, neuroblastoma), or embryonic stem cells (ES cells) are widely used in these studies, but all have several disadvantages. For example, PC12 and Neuro2a cells are derived from rat and mouse, respectively, so they are not suitable for studies of human diseases because rodent cells often react differently to growth factors or differentiation stimuli than human cells. The same applies to mouse embryonic stem cells. Human embryonic stem cells (hESCs), on the other hand, can differentiate into many different cell types including neurons, but differentiation to mature neurons is neither synchronous nor quantitative with presently available differentiation protocols. These protocols are time-consuming, not very efficient, and –due to the need for special growth factors and media– prohibitively expensive.

For the reasons described above, we decided to work with committed neural precursors immortalized with LINX-v-myc vector, where v-myc expression is tightly regulated (Hoshimaru et al., 1996) (Fig. 1.15). In particular, we used LUHMES (Lund human mesencephalic) cells. These cells are a subclone of the tetracycline-controlled, v-myc-overexpressing human mesencephalic-derived cell line MESC2.10, characterized at and originating from Lund University (Lund, Sweden) (Lotharius et al., 2002). The cells were derived from an 8-week-old human ventral mesencephalic tissue. The karyotype of these cells revealed a normal set of chromosomes and female phenotype (Paul et al., 2007). However, it was demonstrated that these cells were unstable and heterogeneous regarding to tyrosine hydroxylase (TH) expression, and they were not suitable for

replacement of dopaminergic neurons upon transplantation (Fountaine et al., 2008; Paul et al., 2007). Due to these reasons, a subclone of MESC2.10 was created, designated LUHMES, which was initially used to study dopamine related cell death mechanisms (Lotharius et al., 2005; Schildknecht et al., 2009).

LUHMES can be induced to differentiate into morphologically and biochemically mature dopaminergic neurons. The differentiation of LUHMES is triggered by shutting down the expression of v-myc after the addition of tetracycline to the differentiation medium, together with glial cell line-derived neurotrophic factor (GDNF), dibutyryl-cAMP (cAMP) and N2 medium supplement. The procedure contains a two-step optimized protocol and results in large-scale post-mitotic neurons after six days (see methods). The differentiated cells exhibit the same dopaminergic and neuronal characteristics as MESC2.10 cells (e.g., intense β-III-tubulin immunoreactivity, extensive neuritic processes, time-dependent induction of tyrosine hydroxylase, and extracellular dopamine release). This is accompanied by the down-regulation of some precursor markers and the up-regulation of neuronal markers.

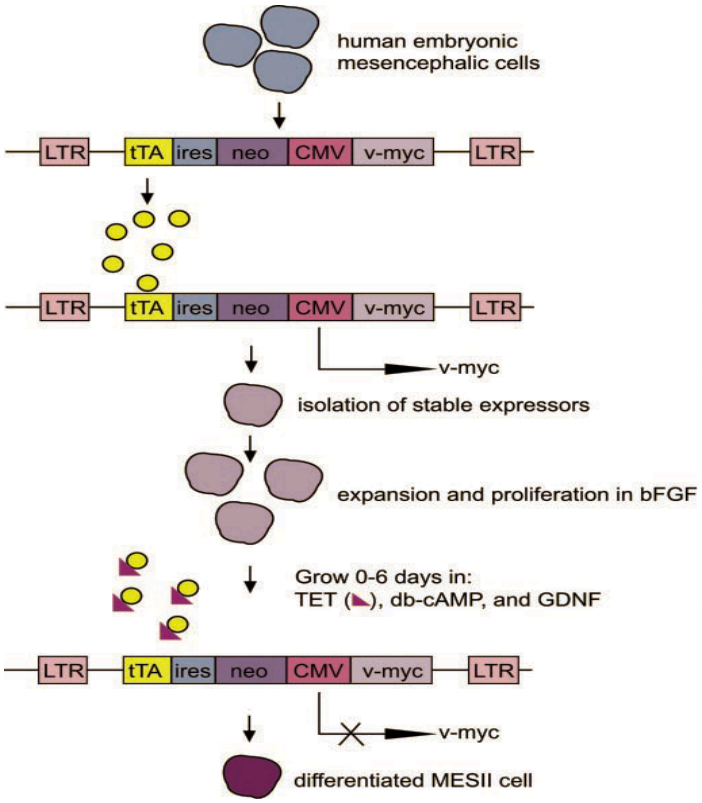


Fig. 1.15: Construction of committed immortalized neuronal precursors. Human embryonic mesencephalic cells are retrovirally infected with a LINX v-myc vector. The vector contains a neomycin resistance gene for selection. A tetracycline-controlled transactivator (tTA, yellow circles) strongly activates transcription from a minimal CMV promoter in the absence of tetracycline. Constitutive expression of v-myc in the absence of tetracycline allows cells to proliferate continuously. Induction with tetracycline (purple triangle) abolishes transcription activation by tTA and blocks the expression of v-myc, leading to differentiation to dopaminergic neurons(Lotharius et al., 2002)

Proliferating LUHMES grow in colonies with absent neurites but small structures, called primary cilia, are frequently observed by electron microscopy. Two days after the

induction of differentiation, LUHMES homogeneously extend structures with characteristics of lamellipodia and filopodia. On day 3, neurites of 50-150 µm with large membrane protrusions are observed. The proximal part of these protrusions is stained for the cytoskeletal bIII-tubulin and the distal part is positive for F-actin but negative for β-III-tubulin. This morphology as well as β-III-tubulin and F-actin distribution are typical for neurites and growth cones in primary neuronal cultures. Monitoring of the growth behavior during LUHMES differentiation revealed morphological characteristics of dopamine neurons in the substantia nigra from histochemical stainings (Arsenault et al., 1988; Jaeger et al., 1989). A percentage of the cells display one neurite and another percentage two neurites at opposite poles. One of the two is elongated constantly with a velocity of 20 µm/h. It often happens that the extending neurite extends and then retracts into different directions and sometimes it forms a second extension followed by the collapse of the original growth cone and formation of a new one at the new edge. The elongation is slowed down on day 5,and a high percentage of cells lacks the growth cone after this time. The average length of the axon on day 5 is 500-1000 µm (fig. 1.16).

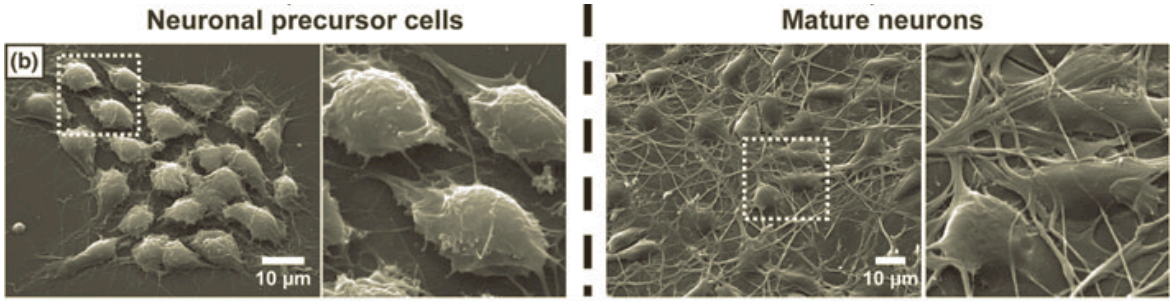


Fig. 1.16: Morphological characteristics of undifferentiated and differentiated LUHMES cells. Inspection of LUHMES under the electron microscope revealed that undifferentiated cells (left) grow in colonies with absent neurites. However, primary cilia are visible. After five days of differentiation (right) cells grow neurites, creating a homogeneous network. Marked squares contain cells, which are shown in higher magnification (modified from Scholz et al. 2011).

Under "maintenance" cell culture conditions in the presence of basic FGF, these conditionally immortalized cells are simultaneously expressing some neuroblast and stem cell markers. Undifferentiated cells express the neuronal cytoskeletal protein β-III-tubulin. The mRNA of this gene is up-regulated during differentiation, as well as the mRNA of Fox-3/NeuN (marker of post-mitotic neurons). On the other hand, CDK1 (cell cycle regulator) is down-regulated during differentiation. Undifferentiated cells are also positive for the proliferation marker Ki-67, which is not detectable in differentiated, post-mitotic LUHMES

cells. Patch-clamp analysis revealed that neuronal electrical features increase from day 5 to day 12 of differentiation, while the post-mitotic state was already reached from day 5.

Analysis of typical differentiation markers revealed a group of genes being up-regulated during differentiation. These genes were related to neuronal function or they were related to neurites or to synapsis formation (neuregulin 1, pentraxin 1, dopamine receptor D2). Synaptic markers are being up-regulated very fast, within two days of differentiation in contrast to post-synaptic markers (GRIN1, DLG4, neuroligin 1) and synaptic vesicle 2a which were up-regulated later. Other neuronal genes are detected in undifferentiated cells and their expression remains stable throughout differentiation (effector of neurogenesis MEF2C, neurite growth-promoting factor 2 MDK, tyrosine activation protein YWHAH) indicating the existence of neuronal features already in undifferentiated LUHMES. Markers such as Sox2 and the marker for migrating neuroblasts Pax3 are down-regulated and completely absent on day 5 of differentiation. Immunostaining revealed the homogeneous expression of the markers among all cells that remain the same after day 5 regarding localization and strength.

LUHMES cells can also differentiate to post-mitotic neurons with an alternative differentiation protocol where GDNF and cAMP are omitted (-/-). In this case, the differentiated cells look morphologically and immunocytologically identical to those obtained from the +GDNF/+cAMP protocol (+/+). They also express similar levels of marker proteins. The only difference is that the expression levels of tyrosine hydroxylase (TH) and dopamine receptor D2 (DRD2) are significantly reduced. This result shows that LUHMES cell fate is predetermined via a robust, endogenous program and only few features (mainly TH) need external signals to be regulated.

LUHMES cells are therefore an ideal model for our study. Their human origin and the optimized 2-step protocol allow for a rapid and synchronous homogenous population of post-mitotic neurons with characterized properties which are reproducible. Most importantly differentiated LUHMES express PRMT8 in high levels (see results).

1.8 Aim of the thesis

Given the important role of arginine methylation in cancer and differentiation, the aim of the present thesis was to find new PRMT1 variants, characterize them and study their

role in cancer. In parallel, we looked for novel nuclear substrates of PRMT1. Finally, we intended to set up new tools for studies on the role of PRMT8 in neuronal differentiation and maintenance including a cellular model, shRNAs and expression vectors as well as antibodies and proceed with a preliminary characterization of the molecule with regard to expression levels, localization etc.

2. Materials and Methods

2.1 Cell culture methods

2.1.1 Human cancer cell lines and normal human lung fibroblasts

Cells were cultivated as a monolayer in tissue-culture grade plastic dishes (Sarstedt) at 37°C under humidified atmosphere containing 5% CO_2. The growth medium consisted of DMEM (PAA) for HEK293 (immortalized human embryonic kidney), HeLa (human cervix carcinoma), HFL-1 and MRC5 (normal human lung fibroblasts), and the human lung cancer cell lines H1299 and A549, or RPMI-1640 (Gibco) for the human lung cancer cell lines H1437 and H1792, according to instructions provided by ATCC. All media contained 10% fetal calf serum (FCS; Gibco) and 1x Penicillin/Streptomycin (PAA). Confluent cells were split every 2-3 days. After rinsing with PBS (136.9 mM NaCl, 2.68 mM KCl, 8.09 mM Na_2HPO_4, 1.47 mM KH_2PO_4) the cells were detached using a 0.05% trypsin/0.022% EDTA solution in PBS, and transferred to new dishes with fresh medium in a ratio of 1:3 to 1:5.

2.1.2 LUHMES

LUHMES cells were cultured at 37^{o}C in a humidified atmosphere containing 5% CO_2. Proliferating cells were washed with PBS and enzymatically dissociated with 0.025% trypsin/0.01% EDTA solution in PBS and seeded in T75 Nunclon^{TM} (Nunc) flasks coated with 40 µg/ml poly-L-ornithine and 1 µg/ml Fibronectin (Sigma-Aldrich). The proliferation medium consisted of Advanced DMEM/F12 (Gibco) supplied with 2 mM L-Glutamine (PAA), 1x N-2 supplement (PAA) and 0.04 µg/ml recombinant basic fibroblast growth factor (FGF; Peprotech). $2x10^6$ or $1x10^6$ cells were seeded every two or three days, respectively.

Differentiation of LUHMES cells was performed in two steps. As a predifferentiation step, $2.5x10^6$ cells were seeded on T75 flasks in proliferation medium. After 24 hours (day 0, d0), proliferation medium was substituted by differentiation medium. Differentiation medium consisted of Advanced DMEM/F12, 2 mM L-Glutamine, 1x N-2 supplement, 1 µg/ml tetracycline, 2 ng/mL recombinant human glial cell-derived neurotrophic factor (GDNF) and 1 mM dibutyryl-cAMP (Sigma-Aldrich). In many experiments, GDNF and dibutyryl-cAMP were omitted. The cells were incubated with differentiation medium for 48 hours in order to stop their proliferation. On day 2 they were detached with

trypsin/EDTA solution and, depending on the experiment, were seeded on new poly-L-ornithin/Fibronectin flasks, dishes or glass coverslips (Marienfeld) in a density of 100.000 cells/cm^2.

2.2 Transfection and infection of cells

2.2.1 Transfection of cells with Polyethylenimine (PEI) and Lipofectamine 2000

One day before transfection, cells were split at a ratio of 1:5 to obtain around 50% confluency at the time of transfection. Prior to transfection, the medium on the dishes (100 mm) was reduced from 10 ml to 3 ml. A transfection "working solution" of PEI (Sigma-Aldrich) was prepared by freshly mixing 9 µl of PEI stock solution (25% PEI in sterile distilled water) with 0.5 ml of sterile distilled water. For each transfection, 4.5 µl of the PEI working solution was further diluted into 300 µl of serum free medium. In a separate tube, 10 µg of plasmid were diluted in 50 µl of serum free medium. PEI-DNA complexes were formed by thoroughly mixing the plasmid dilution with the readily prepared PEI/medium dilution and incubation for 30 min at room temperature. The resulting suspension was dropwise added to the cells, mixed well and incubated from 3 to 5 hours, before the volume of the medium was adjusted to 10 ml again. Cells were analyzed at least 24 h after transfection.

Transfection of cells with Lipofectamine 2000 (Invitrogen) was performed according to manufacturer's instructions. The cells were analyzed at least 24 hours post-transfection.

Depending on the resistance gene that each vector was carrying, HEK293 cells were selected either with 1 µg/ml puromycin (Clontech) for one week or with 400 µg/ml G418 antibiotic (Biosera) for 4-6 weeks in order to obtain stable cell lines.

2.2.2 Generation of lentiviruses and infection of LUHMES

Lentiviruses were produced by PEI transfection of HEK293T cells, as described above, with a three-plasmids system consisting of pMD2.G (envelope vector), pBR8.91 (packaging vector) and a transfer vector, which was different according to the experiment. For protein overexpression, LeGO-iPuro2 was used, while for protein knock-down shRNAs were cloned in a LeGO-G/Puro vector. The ratio between the vectors was 2.3:1.6:1 (transfer vector:pBR8.91:pMD2.G). One day after transfection the medium, of

the cells was exchanged with LUHMES proliferation medium (Advanced DMEM/F12 with 2 mM L-Glutamine) for experiments in LUHMES cells, or DMEM for experiments in other cells. Two and three days after transfection the supernatants were harvested, filtered through 0.45 µm filters (Roth) and incubated overnight with Lenti-X Concentrator (Clontech) or "'homemade" precipitation solution (8% Polyethylene glycol-6000, 300 mM NaCl) at 4°C overnight. Then, the mixtures were centrifuged at 1500 g at 4°C for 45 min, the pellets were resuspended in 100 µl PBS, aliquoted and stored at -80°C. Different volumes of viral stocks were tested for their ability to infect LUHMES cells efficiently. The infected cells were selected with 0.1 µg/ml puromycin for one week in order to obtain stable cell lines.

2.3 Microscopy methods

Microscopic observation of live or fixed cells was carried out in a Leica TCS SP5 Confocal microscope and data acquirement and analysis were achieved by the accompanying software, LAS AF.

2.3.1 Immunofluorescence microscopy

LUHMES cells grown on poly-L-ornithin/Fibronectin-coated coverslips, or HEK293 cells grown on Alcian Blue-coated coverslips (0.1% w/v Alcian Blue in water), were washed briefly in PBS, and fixed for 30 min with 3.5% w/v paraformaldehyde solution in PBS. The coverslips were washed again in PBS and permeabilized with 0.3% v/v Triton for 10 min. The cells were then washed in PBS and they were blocked with 3.5% w/v bovine serum albumin (BSA) in PBS for 1 hour. Cells were incubated with the primary antibody (diluted appropriately in blocking solution) for 1 hour. The coverslips were washed several times with PBS to completely remove unbound antibodies. Then, the samples were incubated with the secondary antibody for 30 min followed again by several washing steps in PBS. For nuclear staining, samples were incubated either with 1:1000 DRAQ5™ (Thermo/Molecular Probes) or 1:300 TO-PRO™-3 Iodide (Invitrogen) or 1:20000 SYTOX™ Green (Invitrogen) diluted in PBS. To remove the PBS, the sample was washed once with water and mounted on a glass slide with mounting medium (100 mM Tris pH 8, 90% Glycerol, 2.5% DABCO).

2.3.2 Fluorescence Recovery After Photobleaching (FRAP)

Fluorescence Recovery After Photobleaching (FRAP) experiments were performed with whole populations of HEK293 cells stably expressing YFP fusion constructs of full length PRMT1v1 and PRMT1v1Δarm. Cells were seeded on 35 mm glass-bottom dishes so as to be 50-70% confluent at the time of the experiment. Dishes were placed on the stage of the Leica SP5 microscope and rectangular regions of interest (ROIs, height of bleached area=2.5 µm) from at least 25 cells containing a part of the cytoplasm and the nucleus were photobleached following the steps of the FRAP wizard of LAS AF software. The ROI was bleached with the 488 nm laser at 85% laser power. The pre-bleaching, bleaching and post-bleaching (recovery) steps were monitored by the software. 250 images of each cell were captured within 50 sec after bleaching using 5% laser intensity to minimize further bleaching. The recovery time in each cellular compartment (cytoplasm and nucleus) in the ROI was measured separately, as well as a region outside of the bleached area. The data were imported to Microsoft Excel for analysis. The pixel intensities on different time points were normalized to the total cell brightness and to the prebleaching brightness for cytoplasm and nucleus separately.

2.4 Biochemical and Molecular Genetics Methods

2.4.1 Quantitation of total protein

Colorimetric quantitation of total protein in cell lysates was done using the Pierce BCA Protein Assay Kit (Thermo Scientific), according to the manual.

2.4.2 Protein extraction/precipitation according to Wessel&Flügge (1984)

Cells were lysed in 400 µl of 1% SDS solution. The lysate was sonified to shear genomic DNA and thereby reduce viscosity. An equal volume of methanol was added to the sample and mixed well. Then, 100 µl of chloroform were added and mixed to obtain a phase separation, and precipitated protein was recovered by centrifugation for 5 min at full speed. The precipitated protein could be observed in the interphase between the methanol solution and the chloroform. The upper phase (methanol, water, salts and nucleic acids) was removed and 300 µl of methanol were added to the interphase and chloroform phase. Precipitated protein was pelleted by centrifugation for 2 min, the

supernatant was removed and the pellet was allowed to dry. The dry pellet could be stored at -20°C, or redissolved in an appropriate volume of SDS PAGE loading buffer, and boiled at 95°C for 5 min.

2.4.3 Gel electrophoresis of proteins and Western blot

Protein extracts were adjusted to 1x SDS PAGE loading buffer (5% mercaptoethanol, 10 mM Tris pH 7.4, 2.5% SDS) boiled at 95°C for 5 min and loaded onto an 8-12% acrylamide gel. The gel run was performed at 18 mA. Proteins were transferred to a PVDF membrane for 70 min at 350 mA by tank blotting in a Biorad chamber. The membrane was blocked with 1x Rotiblock solution (Roth) for 1 hour, rinsed for 5 min with TNT buffer (150 mM NaCl, 10 mM Tris pH 8, 0.005% v/v Tween 20), and then incubated with the primary antibody diluted in TNT for 1 hour to over night. The membrane was thoroughly washed with TNT for 30 min, with at least 5 changes of washing buffer. After washing, the membrane was incubated for 30 min with the secondary antibody diluted in TNT plus 1% dried milk. The membrane was again washed 3 to 4 times in TNT and quickly rinsed once in water before chemiluminescent detection with SuperSignal Femto reagent (Pierce), diluted 1:6 in water was carried out. A film was exposed to the membrane in an X-ray cassette and was developed in an X-ray film developer machine (Amersham).

2.4.4 Immunoprecipitation

Cultured cells were washed twice with cold PBS on the dish, then the PBS was completely removed and 1 ml of lysis buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 1% v/v NP-40) was added. The cells were scraped off the dish and transferred to a tube. The sample was sonified to shear DNA, and centrifuged for 5 min at 10000 rpm at 4°C for preclearing the lysate. The supernatant was transferred into a new tube and the pellet (preclear pellet), containing insoluble proteins complexes was stored at -20°C in case it was needed for analysis. A small amount of the supernatant was also kept as an input control at -20°C. The supernatant was mixed with the antibody (3 µg) and incubated for 2 hours at 4°C in an overhead shaker/mixer. Then, 30 µl of 50% Protein G Sepharose suspension (Pierce) were added, and the sample was incubated for another hour under the same conditions. Immunoprecipitated proteins bound to the beads were collected by

centrifugation (10 sec, 3000 g), and the supernatant was stored for controls. The beads were washed six times by centrifuging and resuspending in 1ml of cold lysis buffer. Precipitated proteins were either eluted from the beads by adding 2x Laemmli buffer (10% mercaptoethanol, 20 mM Tris pH 7.4, 5% SDS) and boiling at 95°C, or used immediately for enzymatic activity assays.

2.4.5 *In vitro* methylation assay

For *in vitro* methylation assays, HEK293 cells growing in 100 mm dishes were transfected with GFP- or YFP-tagged PRMTs by PEI transfection. One day after transfection, the enzymes were immunoprecipitated as described above, using 4 µg of anti-GFP antibodies (Roche). The beads with the precipitated PRMTs were mixed with either 30 µl of hypomethylated total HEK293 extract or with bacterially expressed and purified peptides (kindly provided from Prof. Anastasia's Politou lab, Medical School, University of Ioannina). In the reaction, 5 µl of radioactive methyl donor, S-Adenosyl-L-[methyl-^{3}H]Methionin (SAM[^{3}H]) (spec. activity 2.96 TBq/mmol, Amersham), were added, and the volume was filled up to 60 µl. The reactions were incubated at 37°C for 4 hours with constant shaking before they were terminated with the addition of 7 µl of 10x SDS PAGE loading buffer and boiling at 95°C for 5 min. For the analysis of the results, 1/3 of each reaction was run on a 10% acrylamide gel. After the run, the gel was stained with Coomassie staining solution (25% v/v Isopropanol, 10% v/v Acetic acid, 0.05% w/v Coomassie Brilliant Blue R250) for the visualization of the protein bands. The gel was then incubated in EN^{3}HANCEsolution (NEN Life Sciences) in a glass bowl for 45 min at room temperature and subsequently washed in water 3 times for 10 min each. After the washing steps, the gel was put onto a gel dryer on a wet Whatman filter paper, covered with Saran wrap, and dried for about 1-2 hours at a maximum of 55°C. The dried gel, after the removal of the Saran wrap, was exposed to an autoradigraphy film from overnight to 2 weeks at -80°C and developed in the developer machine.

Hypomethylated extracts were obtained by treatment of HEK293 cells with 15 µM of the S-adenosylhomocysteine hydrolase inhibitor, adenosine-2′, 3′-dialdehyde (Adox) for 48 hours. Cells were washed twice in PBS before they were rinsed briefly twice with dd$H_2$0. Then, they were harvested in a small volume of dd$H_2$0 and heated at 70°C for 10

min for the inactivation of the endogenous PRMTs. The samples were let to cool down and centrifuged at full speed for 15 min. The supernatant was kept and adjusted to 1x PBS by the addition of the appropriate of volume of 10x PBS, and the sample was centrifuged again. The supernatant was aliquoted and stored at -20°C until use.

2.4.6 RNA analysis - Preparation of total RNA and cDNA synthesis

Total cellular RNA was extracted and purified using the NucleoSpin RNA Mini kit (Macherey-Nagel) according to the manufacturer's manual. After quality control by agarose gel electrophoresis (1% w/v agarose in TAE buffer: 40 mM Tris pH 8, 1 mM EDTA), RNA was reverse-transcribed with the PrimeScript RT reagent kit (Takara #RR037A) according to protocol V of the manufacturer. The synthesized cDNA was stored at -20°C until further use. All procedures were done under RNase-free conditions.

2.4.7 Polymerase chain reaction (PCR)

For RT-PCR, 1 µl of cDNA was amplified by PCR in a volume of 20 µl, containing 10 pmole of sense- and antisense primers, 0.4 µl dNTP mix (containing 10 mM per nucleotide), 2 µl 10x concentrated reaction buffer (final concentrations: 50 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris pH 9, 0.5% Triton X100) and 2 units of homemade Taq polymerase. The final volume was adjusted to 20 µl with water. Amplification programs were carefully optimized in pilot experiments, according to melting temperatures of the primers, as given in table 2.3. Typically, DNA was initially denatured at 95°C for 5 min, before primers were annealed for 30 sec. DNA was elongated at 72°C for 30 sec, before denaturing again at 95°C for 30 sec. Denaturation, annealing, and elongation were repeated for 20 to 35 cycles, depending on the gene of interest. A final elongation step at 72°C for 10 min was performed, before the samples were cooled to 4°C. Reaction products were either analyzed by agarose gel electrophoresis or stored at -20°C.

To quantify transcripts, real-time quantitative PCR (qPCR) was performed in capillaries on a Roche LightCycler 2, using cyber green (Kapa #KK4602) in a total volume of 20 µl. All qPCR analyses were done on at least two biological replicates, in triplicate each. For normalization, we used primer sets for GAPDH.

2.5 Work with bacteria

2.5.1 Transformation of bacteria by electroporation

Aliquots (40 µl) of frozen, electrocompetent E.coli XL1 bacteria were thawed and mixed with 1 µl plasmid diluted in water to 1 ng/µl. An electric pulse was applied with a BioRad GenePulser II, using 1800 V, 200 Ohm, 25 µFa in a 1 mm electroporation cuvette. After the pulse, cells were immediately diluted to 1 ml with LB medium (1% w/v Bacto-Trypton, 0.5 w/v Yeast extract, 1% w/v NaCl), transferred to a glass test tube, and allowed to recover for 30 min at 37°C. After recovery they were plated on LB agar plates (LB medium plus 0.15% w/v agar) containing the appropriate antibiotic and incubated over night.

2.5.2 Small-scale preparation of plasmid DNA

Single colonies were picked from the Petri dishes, and cultured for 8 hours to over night at 37°C in 1 ml of LB with the appropriate antibiotic. Bacteria were harvested by centrifugation for 5 min at 5000 rpm in an Eppendorf centrifuge and resuspended in 100 µl of GET resuspension buffer (50 mM Glucose, 25 mM Tris pH 8, 10 mM EDTA). Bacteria were lysed in 200 µl of lysis buffer (0.2 N NaOH, 1% w/v SDS) by gently inverting the tube several times. Bacterial protein and genomic DNA were precipitated with 150 µl of cold salt solution (3 M Potassium acetate, 1.8 M Acetic acid). After gentle but thorough mixing, precipitated material was removed by centrifugation (5 min, 14000 rpm). 400 µl of the supernatant was mixed with 900 µl of cold ethanol, and precipitated DNA was pelleted by centrifugation (15 min, 14000 rpm). The supernatant was completely removed, and the pellet was allowed to air-dry. Finally, DNA was redissolved in 50 to 100 µl of TE (10 mM Tris-HCl pH 8, 1 mM EDTA) containing 100 ng/ml DNase-free RNase (Roche).

2.5.3 Medium-scale preparation of plasmid DNA

Larger amounts of plasmid were purified from 100 ml of bacterial culture, using the NucleoBond Xtra Midi kit (Macherey-Nagel) according to the manual.

2.6 Other methods

2.6.1 Fluorescence-activated cell sorting (FACS)

For fluorescence-activated cell sorting cells were washed with PBS and detached from the dishes with trypsin-EDTA solution. Then they were harvested and centrifuged at 300 g for 5 min. The cell pellet was incubated with hypotonic solution (0.1% w/v sodium citrate, 0.1% v/v Triton X-100, 50 mg/l propidium iodide). The stained cells were analyzed on a FACSCalibur (Becton-Dickinson) flow cytometer, equipped with 2 lasers (488 nm, 635 nm). Data analysis was carried with CellQuest software (Becton-Dickinson), for at least 10000 cells/sample. Untransfected and YFP-transfected HEK293 cells were used as a control for the "normal" cell cycle and the YFP fluorescence respectively.

2.6.2 Identification of molecules in proximity with the BioID method

BioID is a method being used for the identification of proteins, which are in proximity to a protein of interest. The principle of the method is based on the idea that a biotin ligase is fused to the protein of interest and then it is transfected/infected to the cells, where it will biotinylate neighboring proteins after the addition of biotin to the culture medium. The biotinylated proteins can then be isolated by streptavidin beads and identified by mass spectrometry. The principle of the BioID method is presented in figure 2.1. For our experiments we used a mutated form of the bacterial biotin ligase BirA, called BirA*(R118G). This mutant cannot make dimers and cannot bind to DNA (Kwon et al., 2000).It has been shown that BirA* promiscuously biotinylates the proteins in proximity *in vivo* after the supplementation of biotin in the culture medium (Roux et al., 2012).

PRMT8 fused with BirA* in its carboxy-terminus was cloned in a LeGO-iPuro2 vector and infected in LUHMES cells before selection with 0.1 μg/ml puromycin for one week. For control experiments a LeGo-iPuro2 vector encoding BirA* was used for creating a LUHMES stable cell line under the same conditions. The infected cells were then qualified for the ability to express and properly localize the proteins by immunofluorescence and Western blot with an anti-HA antibody (Roche) as describe before. For labeling proteins in proximity with biotin undifferentiated cells were supplied with 50 μM of biotin in the culture medium. 24 hours later the cells were lysed in 1% SDS. The proteins were precipated by the Wessel&Flügge protocol, analyzed by SDS-PAGE and transferred to a

PVDF membrane. The blot membrane was incubated with 1:25000 diluted streptavidin conjugated to horseradish peroxidase (HRP) and exposed to an autoradiography film after incubation with ECL.

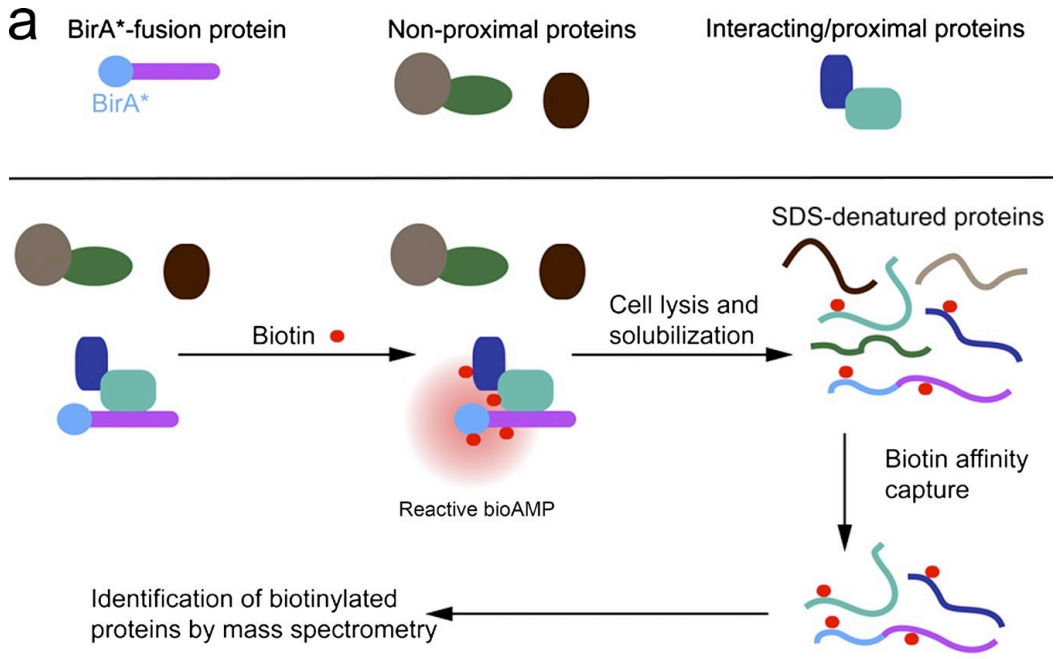


Fig. 2.1: Principle of the BioID method. A protein of interest is fused to a bacterial biotin ligase (light blue, purple) and transfected to cells. Proximal proteins (dark blue, petrol) are labeled after the addition of biotin (red dots) in the culture medium. Labeled proteins can then be purified by streptavidin magnetic beads and analyzed by mass spectrometry (from Roux et al. 2012).

2.6.3 Studying the distribution of PRMT8M16 between the cytoplasm and the nucleus

HEK293 cells stably expressing PRMT8M16:EYFP were seeded in to 35 mm glass-bottom dishes so as to be 50-70% confluent at the time of the experiment. Random fields of the cells were imaged at the confocal microscope keeping the same parameters from one image to the other. The acquired images were analyzed by the ImageJ software. There, a region of interest (ROI) of 15x15 pixels was used to measure the mean brightness of cellular and nuclear PRMT8M16:EYFP in more than one hundred cells. The numbers were statistically analyzed by Microsoft Excel.

2.6.4 Antibody production in rabbits

Two rabbits were injected with an epitope corresponding to a region in the amino-terminus of PRMT8 (NH_2-CPGRGKMSKLLNPEEMT-COOH) mixed with complete Freund's

adjuvant. Two weeks later, the injection was repeated and samples from their blood were tested for their antibody titer before the final bleeding.

Table 2.1: Antibodies used in the present thesis.

Antibody	Company	Order number
Alpha-tubulin	Hybridoma Bank	12G10
Geminin	Hybridoma Bank	CPTC-GMNN-1
Neurofilament 165	Hybridoma Bank	2H3
Neurofilament 200	Hybridoma Bank	RT97
Synaptic vesicle marker	Hybridoma Bank	SV2
Islet1	Hybridoma Bank	40.2D6
HA tag	Roche	11 583 816 001
GFP	Roche	11 814 460 001
Flag tag	Sigma-Aldrich	F3165
Secondary antibodies		
Alexa568 goat a-mouse IgG	Molecular Probes	A-11004
Alexa568 goat a-rabbit IgG	Molecular Probes	A-11036
a-mouse IgG (Western blot)	Sigma	A0545
a-rabbit IgG (Western blot)	Sigma	A9309

Table 2.2: Plasmids used in the present thesis.

Recombinant gene	Vector backbone	Origin
PRMT1:GFP	pGFP-C	Kindly provided by Prof. Mark Bedford
PRMT4:GFP		
PRMT6:GFP		
PRMT8:GFP	pGFP-N	
PRMT8M16:EYFP	pEYFP-N	This work
Seven aminoterminal variants of PRMT1 tagged with GFP	pGFP-N	Kindly provided by Prof. Jocelyn Coté
PRMT1v1:YFP	pEYFP-N	This work
PRMT1v1Δarm:YFP		
PRMT1v2Δarm:YFP		
Fibrilarin:FusionRed	pFusionRed-N	Kindly provided by Prof. Dmitriy Chudakov
Snail:FlagHA 6SA	pCMV-Tag 2B	Addgene #16221
shRNAs for PRMT8 knock-down (KD1-KD4, scrambled)	LeGo-G/Puro	http://www.lentigo-vectors.de/
PRMT8:GFP	LeGO-iPuro2	http://www.lentigo-vectors.de/
PRMT8:BirA*HA		This work
pcDNA3.1MCS-BirA(R118G)-HA	pcDNA3.1MCS	Addgene #36047

Table 2.3: PCR primers used in the present thesis.

Gene name	Sequence 5'→3'	Product size (base pairs)	Ta (°C)
PRMT8	CCAAAGCAAGTGGTGACCAATG	332	55
	TCTACTGTGAAATCGAGGTCTCG		
SOX2	GGCAATAGCATGGCGAGCG	324	55
	CACCGAACCCATGGAGCCAA		
PAX3	CCTTTCCGTTTCGCCTTCAC	402	62
	CTTCTCCACGTCAGGCGTT		
GMNN-1	GGAGGAGTCACCCAGGAGTCA	387	55
	TGGCTTTGCATCCGTAGAGGAA		
NF165	TGGGAAATGGCTCGTCATTT	294	55
	GGCCTCTTCCATTTCTGACT		
NLGN	TGAGGCCACTCAAGTTCCAA	384	58
	GTGCCATGGTCCCACATTAG		
SYP	CGAGTTCGAGTACCCCTTCA	377	55
	CTCTCAGCTCCTTGCATGTG		
DRD2	AAGCCACTCAGATGCTCGCC	412	58
	CACTGCCCTGGCAGAGTGAG		
TH	CACCTTCGCGCAGTTCTC	387	58
	CTGTCCAGCACGTCGATG		
GAPDH	AGCCACATCGCTCAGAACAC	472	53
	GAGGCATTGCTGATGATCTTG		
ACTIN BETA	GGATCTTCATGAGGTAGTCAGTC	626	53
	CCTCGCCTTTGCCGATCC		
PRMT1Δarm	CGGGACAAGTGGCTGGAGGTG	429	59
	CGAAACCGCCTAGGAACGCT		
P8VX	AGAAACGGTGCGCCGAGATGG		55
P8VYFW1	TCCCCTCCCTCTTGAAAATGC		
P8VYFW2	CATGTCTCAGGTCCCCTCC		
P8V1FW1	ATGGGCATGAAACACTCCTCC	471	
P8V1FW2	TGCTCCTGAGGAGGAAAATGG	442	
P8V2FW1	ATGGAGTCTCTGGCTTCAG	444	
P8V2FW2	GGCTTCAGATGGATTCAAGC	433	
P8REV	GTGGTTGGCCTTAATGATCTTC		
PRMT1V1	GAGGCCGCGAACTGCATCAT	331	59
PRMT1V2		385	
PRMT1V3		504	
PRMT1V4	TCCATTGCCAATGAAATCTTCCA	470	
PRMT 1V5	ACTGGAGAGATGGTGTCCTGTGG	318	
PRMT 1V6	AGAAGCTGACCAGACAAAGAGA	329	
PRMT 1V7	TGCATCATGGAGGAGATGCTGAAGG	219	

Common reverse primer for PRMT1 variants	TGGCTTTGACGATCTTCACC		
PRMT1Δarm_qPCR	CGGGACAAGTGGCTGGAGGTG	108	60
	CAGGGCGTGCACGTAGTCAT		
PRMT1_qPCR	ATGACTACGTGCACGCCCTG	113	
	CGTCTGCTTCCAGTGCGTGTA		
GAPDH_qPCR	TGCACCACCAACTGCTTAGC	87	
	GGCATGGACTGTGGTCATGAG		

Table 2.4: shRNAs sequences used in the present thesis.

Gene name	Sequence 5'→3'
PRMT8 KD1	AATGTGCGAGACCTCGATTTCACAGTAGA
PRMT8 KD2	AAGCAAGTGGTGACCAATGCCTGTTTGAT
PRMT8 KD3	GTGTGAAACATCTGTATCTAATGACTACA
PRMT8 KD4	GTGTTCAAGGACAAAGTGGTACTGGATGT
Scrambled control	GCACTACCAGAGCTAACTCAGATAGTACT

3. Results

3.1 A new PRMT1 splicing variant

Protein Arginine Methyltransferase 1 (PRMT1) is the predominant member of the family of protein arginine methyltransferases, which are involved in a large variety of cellular processes, including epigenetic regulation of transcription, RNA processing, DNA repair, and signal transduction. PRMT1 accounts for approximately 85% of the asymmetrically dimethylated arginines generated in mammalian cells (Tang et al., 2000), and has a complex pattern of substrate proteins in both the nucleus and the cytoplasm (Herrmann et al., 2005). It is known that PRMT1 obtain its enzymatic activity through the formation of oligomers, which start being formed from dimers (Zhang and Cheng, 2003). Until now, it was known that PRMT1 produces seven distinct splicing isoforms, which differ in their amino-terminal first exons, which are involved in substrate recognition and subcellular localization (Goulet et al., 2007). No variation has yet been described in the catalytic core of the protein. For the first time, we have discovered and characterized a novel ubiquitous splicing isoform of PRMT1, which lacks exons 8 and 9. These exons encode the multimerization arm of the enzyme, which is required for the formation of catalytically active enzyme oligomers.

3.1.1 Bioinformatic identification of the new variant

As published earlier by other laboratories (Goulet et al., 2007; Scorilas et al., 2000), PRMT1 can be spliced to produce seven different isoforms, which vary in the 5' region of the mRNA and consequently the amino-terminal end of the protein molecule (Fig. 3.1A). Inspired by this work, we decided to search for other PRMT1 variants that might have an interesting role in physiological or pathological conditions of the cell. In database analyses, we identified an Expressed Sequence Tag (EST) from neuroblastoma cells (GenBank accession number BX352789.2). This EST sequence was identical with the PRMT1v1 mRNA, but with the striking difference that exons 8 and 9 were absent. These two exons encode for the dimerization arm of PRMT1, which is responsible for the dimerization and hence for the activity of the enzyme (Fig. 3.1B, C). Therefore, the newly identified variant was named PRMT1Δarm (or shortly Δarm).

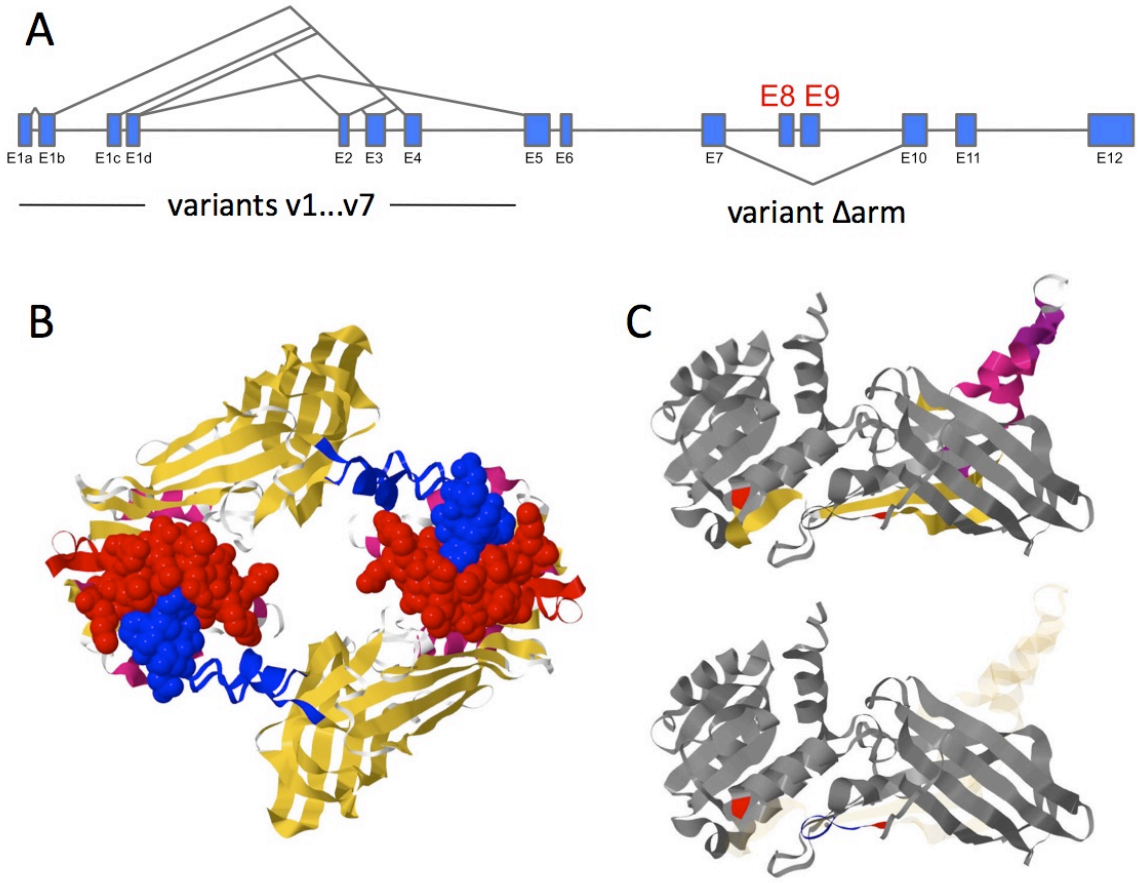


Fig. 3.1: Exon organization of the Prmt1 gene, showing the known splicing variants v1-v7, and the novel variant PRMT1Δarm (A) and structural context of the new variant (B, C). (B) Structure of the dimer of canonical PRMT1, showing in space fill mode the tip of the dimerization arm (blue) and the dimer interface on the AdoMet binding domain of the bound second monomer (red). (C) Structure of a PRMT1 monomer, with the dimerization arm highlighted in color (upper panel), and the PRMT1Δarm deletion variant with the missing part in transparent color (lower panel). Note that the two amino acids, amino-terminal (end of exon 7) and carboxy-terminal (start of exon 10) of the dimerization arm, are located closely together in the full-length protein, and can easily be joined without causing major structural changes in the protein.

3.1.2 Expression of Δarm in different cells

For the characterization of the Δarm isoform, we initially investigated whether the *in silico* identified isoform was expressed endogenously in different cell types. For that reason, we designed a primer pair, which would specifically amplify the Δarm isoform. The forward primer (5'-CGGGACAAGTGGCTGGAGGTG-3') was designed in a manner that it anneals to a sequence across the exon-exon junction of exons 7 and 10 of PRMT1 mRNA and is therefore specific for Δarm that lacks exons 8 and 9. The reverse primer (5'-CGAAACCGCCTAGGAACGCT-3') hybridizes to a sequence in exon 12 that is common to all known PRMT1 isoforms. PCR with this primer pair and the appropriate cDNA as a template is predicted to produce a specific amplicon of 429 bp, if Δarm is expressed. For

the experiment, total RNA was purified and reverse transcribed to cDNA from HEK293, HeLa and undifferentiated LUHMES cells (for extended description and characterization of these cells see introduction and below). The resulting cDNAs were analyzed by PCR with the described primer set, followed by agarose gel electrophoresis. GAPDH was used as a quality and loading control. Figure 3.2 shows that both cancer cell lines (HEK293 and HeLa) as well as the immortalized neuronal precursors (LUHMES) are positive for Δarm; thus, Δarm is being endogenously expressed in these cell lines of different origin, at least on the level of RNA (also see additional experiments with other cell lines below). Currently, the expression of Δarm on protein level cannot be measured due to the lack of an appropriate antibody, which would recognize specifically this variant, but not the remaining seven isoforms of PRMT1.

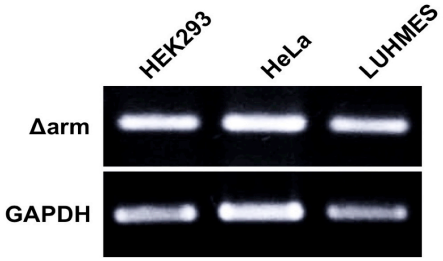


Fig. 3.2: PRMT1Δarm mRNA detection in different cell lines. Total RNA was purified from HEK293, HeLa and undifferentiated LUHMES, and subjected to RT-PCR with a primer pair specific for Δarm. Δarm was present in all three cell lines. GAPDH was used as a quality control.

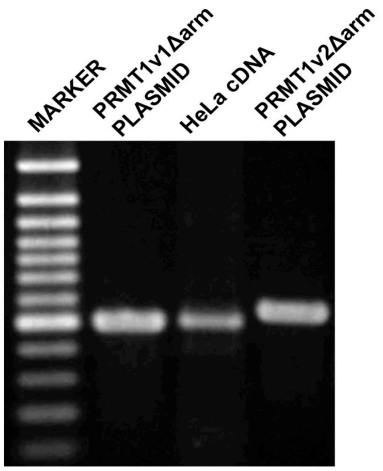


Fig. 3.3: The arm deletion preferentially occurs together with variant 1 of the amino-terminus. Expression of Δarm in HeLa cells was investigated by RT-PCR with a forward primer that anneals to exon E1d common to variants 1, 2, 3, 5 and 7, and a reverse primer specific for Δarm. This would give different amplicon sizes for different combinations of v1-v3 together with Δarm. For comparison, an identical PCR was performed with two plasmids encoding PRMT1v1Δarm:EYFP and PRMT1v2Δarm:EYFP respectively. Note that the amplicon size from HeLa cDNA is identical to the one from the PRMT1v1Δarm plasmid.

3.1.3 Deletion of the dimerization arm and amino-terminal variation of PRMT1

The originally identified EST included a part of exon 1d followed by exons 4-12, excluding of course exons 8 and 9. This is a strong indication that the arm deletion co-

occurs with variant 1 of the amino-terminus (see Fig. 3.1), but does not rule out that it may also be combined with the amino-terminus of other isoforms. To investigate this point, we performed total RNA isolation from HeLa cells, reverse transcribed it to cDNA and amplified it by PCR with a forward primer (5'-GAGGCCGCGAACTGCATCAT-3') that anneals to exon E1d common to variants 1, 2, 3, 5 and 7, and a reverse primer specific for Δarm (5'-GGTATAGATGTCCACCTCCAGCC-3'). This would give different amplicon sizes for different combinations of v1, v2, v3, v5 and v7 with Δarm (predicted product sizes, 507 bp, 561 bp, 680 bp, 665 bp and 405 bp, respectively). In parallel, an identical PCR was performed with cloned constructs encoding PRMT1v1Δarm:EYFP and PRMT1v2Δarm:EYFP for comparison. Figure 3.3 shows that HeLa cDNA produced a single band that has identical size with the one from the PRMT1v1Δarm:EYFP plasmid (507 bp). The conclusion is that the arm deletion occurs together with variant one of the amino-terminus as it was expected from the sequence of the EST. For this reason, we decided to focus the following experiments on the full-length and the Δarm form of variant 1.

3.1.4 Comparison of the subcellular localization of Δarm with the other amino-terminal variants of PRMT1

It is known from the literature that the amino-terminal variants of PRMT1 localize differently among the cell compartments in HeLa cells (Goulet et al., 2007), but also due to different conditions like the accumulation of the cytoplasmic variant 2 into the nucleus after enzymatic inactivation by site directed mutagenesis in HEK293 cells (Herrmann and Fackelmayer, 2009). This is speculated to be one of the reasons of their substrate specificity. Next, we thus compared the localization of Δarm to the seven amino-terminal variants of PRMT1, to clarify whether the lack of the dimerization arm can affect the localization of PRMT1v1. HEK293 cells were transfected with expression vectors coding the seven isoforms tagged with GFP in their carboxy-terminuses (a kind gift from Jocelyn Cote's lab at the Department of Cellular and Molecular Medicine, University of Ottawa), and stable cell lines were created after selection with G418 for five weeks. Live cells were observed under the microscope and confocal images were obtained. In parallel, PRMT1v1Δarm:EYFP vector was used and the same procedure was followed. As shown in Figure 3.4, the lack of the arm does not significantly affect the localization of the protein.

In general, all variants, including Δarm, were distributed between cytoplasm and nucleus, with the cytoplasmic levels being always higher. In contrast to HeLa cells where, variants 3, 4, 5, 6 showed an even distribution between cytoplasm and nucleus, and variants 1 and 7 showed a stronger nuclear signal (Goulet et al., 2007), HEK293 cells presented stronger cytoplasmic levels. This indicates that the localization of the variants is dependent on the cell type. The only exception was variant 2, which was completely excluded from the nucleus and it was detected exclusively in the cytoplasm. This result was expected since variant 2 is the only variant, where alternative exon 2 is translated and which codes for a nuclear export signal. We conclude that PRMT1v1 localization is not affected when the dimerization arm is missing. Moreover, the distribution of amino-terminal variants of PRMT1 between cytoplasm and nucleus depends on the cell type in which they are being expressed.

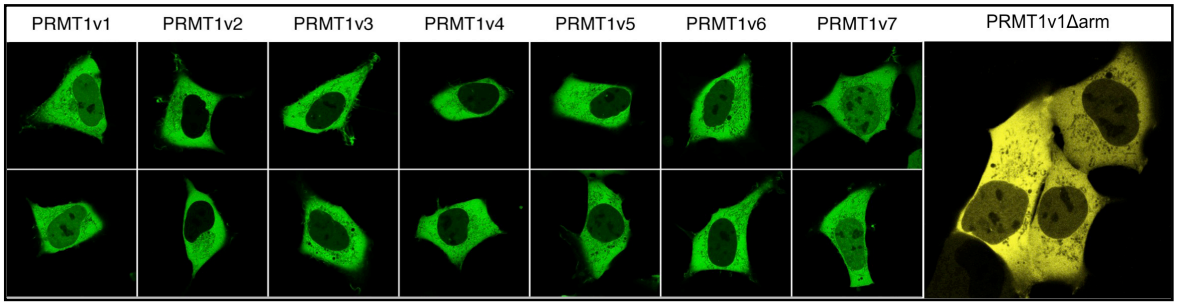


Fig. 3.4: Localization of PRMT1 isoforms in live cells. The seven known variants of PRMT1 were expressed as EGFP fusion proteins in HEK293 cells, and visualized by confocal microscopy. The Δarm of variant 1 was expressed as a fusion protein with EYFP. Two or three representative cells are shown in each case. Localization of variant 1 is not affected upon removal of the dimerization arm. Additionally, the seven variants show differences in their distribution between cytoplasm and nucleus compared to the one that has been show in HeLa cells.

3.1.5 Expression levels of Δarm

Microscopic observation of HEK293 cells transiently or stably transfected with PRMT1v1:EYFP and PRMT1v1Δarm:EYFP, repeatedly indicated much lower expression levels of the Δarm variant. Moreover, increased cell death was observed in the Δarm-transfected cells two days after transfection. In the stable cell lines, the cells which manage to survive were those with the lowest expression levels of Δarm. The same result was observed when whole cell extracts from these cells were analyzed by Western blot with antibodies against GFP (Roche) which recognize the EYFP tag of both proteins. In detail, HEK293 cells were transfected with PRMT1v1:EYFP and PRMT1v1Δarm:EYFP

coding vectors. After 24 hours, total cell extracts were obtained and analyzed by SDS-PAGE and immunoblotting. The acrylamide gel was stained with Coomassie for

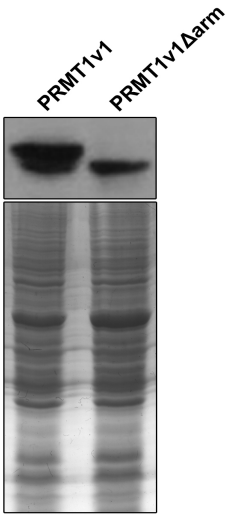


Fig. 3.5: Δarm is expressed in lower levels compared to those of the full-length PRMT1. HEK293 cells were transiently transfected with expression vectors encoding PRMT1v1:EYFP and PRMT1v1Δarm:EYFP. The total cell extracts were analyzed by SDS-PAGE and Western blot after 24 hours. Coomassie staining of the gel (lower part) indicated slightly lower total cell extract amounts loaded for the full-length PRMT1 expressing cells. Even in this case, Δarm is expressed in much lower levels than the full-length PRMT1 as the Western blot with a-GFP antibodies

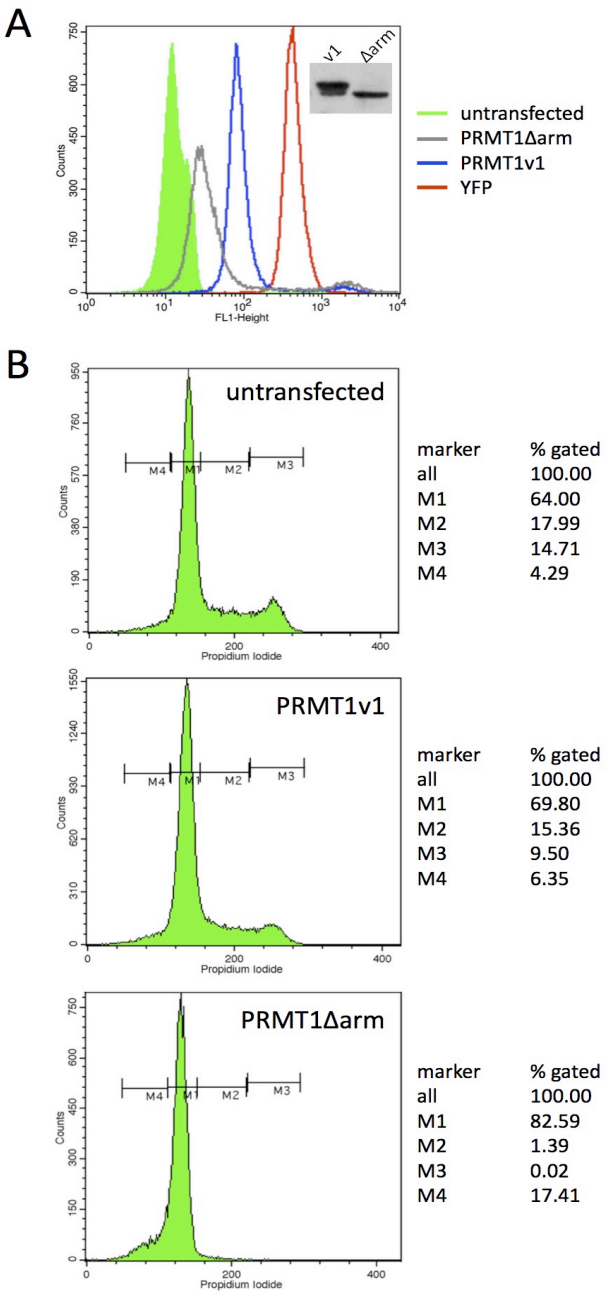


Fig. 3.6: FACS analysis of PRMT1v1:EYFP and PRMT1v1Δarm:EYFP transfected HEK293 cells. HEK293 cells were transiently transfected with constructs encoding PRMT1v1:YFP or PRMT1v1Δarm:YFP, respectively. 24 hours after transfection, cells were harvested and analyzed by fluorescence-activated cell sorting (FACS). (A) Cells were investigated for expression levels of the two constructs. Untransfected cells and cells transfected with empty eYFP vector were used as references. Insert: western blot with extracts from cells expressing the two variants, using anti-GFP antibodies. (B) Cells were analyzed for cell cycle by staining with propidium iodide. Note that cells expressing PRMT1v1 have a cell cycle profile identical with untransfected control cells, whereas PRMT1v1Δarm lead to an arrest in G1 and increased number of cells with sub-G1 amount of DNA. (M1= G1 phase, M2= S, M3= G2/M, M4=sub-G1)

normalization and the blot membrane was stained with antibodies against GFP. The results are presented in Figure 3.5, demonstrating that the level of the expressed full-length PRMT1 is around four times higher than that of Δarm. This result is compatible with our previous microscopic observations.

The expression levels of Δarm were also compared to those of the full-length PRMT1 by fluorescent activated cell sorting (FACS). HEK293 cells were transfected with the same plasmids and 24 hours later they were detached from the dishes and incubated with propidium iodide (PI) hypotonic solution (0.1% w/v sodium citrate, 0.1% v/v Triton X-100, 50 mg/l PI in distilled water) for 15 minutes. PI was used for cell cycle analysis in parallel with the expression levels comparison for the two proteins. The analysis of the results revealed that the Δarm expression levels were about four times less than the full-length protein (Fig. 3.6A). Due to this reason and the fact that a cell sorter for sorting cells with similar expression levels of the two proteins was not available, the experiments which required normalized expression levels was carried by the usage of 4 times more of the Δarm cell extract. It is important to mention here that the two proteins were cloned in the same vector backbone and were under the control of the same promoter (CMV). They were also transfected in exactly the same way to the cells.

PI incorporated levels from the DNA of each cell revealed that the cell cycle was blocked in the G1 phase in cells expressing Δarm. As shown in Figure 3.6B, expression of PRMT1v1Δarm:EYFP, but not full-length PRMT1v1:EYFP, results in a block in the G1 phase of the cell cycle, and in an increase of cell death as shown by an increase of the sub-G1 fraction of cells from 6.4% to 17.4%. We conclude that a considerable expression of PRMT1Δarm is not compatible with progression through the cell cycle, and consequently lead to cell death by apoptosis. Interestingly, already in short times after transfection, the expression levels of PRMT1v1Δarm and full length PRMT1v1 were different by a factor of around 4, even though they were expressed from the same vector backbone with the identical promoter. Δarm is not tolerated by the cells and leads to cell death, so that only cells with very low expression levels can survive and create stable cell lines. Whether the above is due to the blocking of the cell cycle or whether other cellular functions are affected and consequently block the cell cycle and lead to cell death remains to be elucidated.

3.1.6 Potential proteasomal degradation of Δarm

The next question we addressed was whether the low level of Δarm protein was an outcome of degradation through the proteasomal pathway in a higher degree than the full-length protein. To this end, HEK293 cells were transfected either with the full-length PRMT1 or with the Δarm form and incubated with 1 µM of proteasomal inhibitor MG132 (Sigma) for 8 and 16 hours. The total cell extracts were analyzed by Western blot with a-

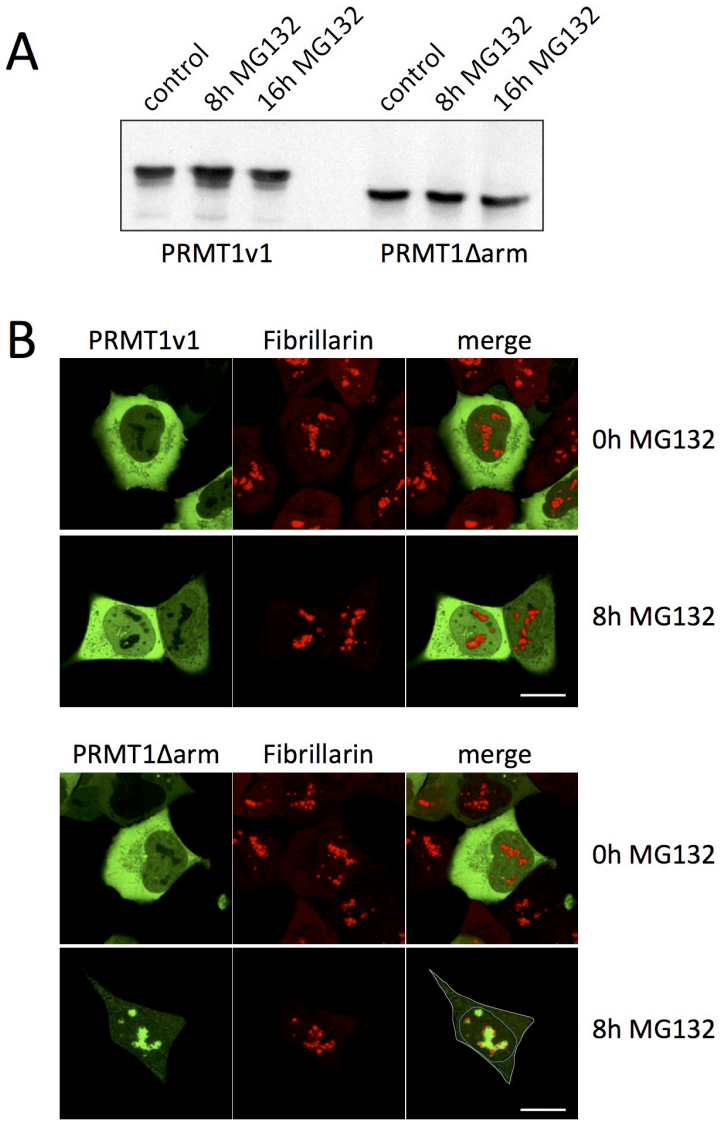


Fig. 3.7: Localization of PRMT1v1 and PRMT1v1Δarm are differently affected by inhibition of the proteasome with MG132. (A) Cells transiently expressing the two constructs were treated with MG132 for 8 or 16 hours, respectively, before total cell extract was prepared and analyzed by western blotting with an anti-GFP antibody. Even after 16 hours, the expression levels of both proteins remain constant. Full-length PRMT1 showed slightly increased levels after 8 hours of treatment. (B) HEK293 cells stably expressing Fibrillarin-FusionRed as a marker for nucleoli were transiently transfected with the two PRMT1 constructs, and then treated with MG132 as in (A). Analysis by confocal microscopy reveals that PRMT1v1Δarm, but not PRMT1v1, re-localize from predominantly cytoplasmic localization to the periphery of nucleoli upon MG132 treatment. Scale bar: 20 µm.

GFP antibodies. As shown in Figure 3.7A, the inhibition of the proteasomal machinery did not affect the expression levels of the proteins since the levels of both remain stable independently of the shorter or longer incubation times with MG132. The only exception is that after 8 hours slightly more PRMT1 is observed. Apparently, the increased proteasomal degradation is not the reason for the lower Δarm levels. Combining the data from this experiment and the FACS analysis, we conclude that elevated levels of Δarm possibly lead to growth disadvantage resulting only in cells with very low expression levels, or Δarm variant levels were suppressed with an unknown mechanism and for a yet unknown reason by the cell.

While inhibition of the proteasome did not result in changes at the levels of the two proteins, we microscopically observed a striking re-localization of Δarm after 8 hours of MG132 treatment in HEK293 cells, which did not occur for full-length PRMT1 (Fig. 3.7B). As also shown above in Figure 3.4, PRMT1v1:EYFP and PRMT1v1Δarm:EYFP do not differ in their subcellular localization under normal cell culture conditions. However, after eight hours of inhibition of the proteasome with MG132, Δarm is almost completely absent from the cytoplasm and has accumulated in the nucleolus, where it localizes in close vicinity to fibrillarin visualized by the nucleolar marker protein Fibrillarin:FusionRed (a kind gift of the laboratory of Dmitriy Chudakov, Moscow). Under the same conditions, full length PRMT1v1 does not significantly change its localization (Fig. 3.7B, upper panel). Thus, PRMT1v1Δarm accumulates in nucleoli when the proteasome is inhibited, while full- length PRMT1v1 is not affected. Similar re-localization after proteasome inhibition has been shown before for several other proteins, but the reason for this phenomenon is not yet clear (see discussion).

3.1.7 Dimerization ability of PRMT1v1Δarm

It has been shown earlier that PRMT1 becomes enzymatically active only after oligomerization (Zhang and Cheng, 2003). These authors had also used an artificial construct lacking the dimerization arm, similar to our naturally occurring Δarm isoform. As their recombinant construct eluted as a monomer from gel filtration columns, and lacked enzymatic activity, we hypothesized that the lack of exons 8 and 9 in the Δarm isoform, which code for the dimerization arm of PRMT1, would not allow the interaction of Δarm

with itself or with the other PRMT1 isoforms. In order to examine this possibility, we designed and carried out the following experimental procedure. HEK293 cells were transiently transfected with expression vectors encoding PRMT1v1:EYFP and PRMT1v1Δarm:EYFP. Cells were lyzed and the whole cell extracts were subjected to immunoprecipitation with anti-GFP antibodies. Four times more starting material was used from the cells which were expressing Δarm to account for the lower expression level of the protein (see above). The precipitated proteins were analyzed by Western blot. The blot membrane was stained with anti-PRMT1 antibodies (Cell Signalling), revealing the overexpressed fusion proteins and the endogenous PRMT1 isoforms and oligomers, which co-immunoprecipitated with the full-length PRMT1v1 protein, but not with the Δarm form of variant 1 (fig. 3.8A). PRMT1v1Δarm:EYFP appears in a lower molecular weight than the full-length PRMT1v1:EYFP due to the lack of the arm. Endogenous PRMT1 monomers, which appear in the first lane together with some high molecular weight bands (most probably representing remaining non-denatured PRMT1 oligomers) indicate that full-length PRMT1 is able to interact with endogenous PRMT1, while on the other hand, the Δarm isoform does not co-precipitate endogenous PRMT1 (lane 2). This result verifies that the lack of the dimerization arm leads to dimerization and further oligomerization incapability of Δarm variant with any of the endogenous PRMT1 variants.

3.1.8 Enzymatic activity of PRMT1v1Δarm

The lack of the dimerization arm and the consequent dimerization incapability of Δarm suggest that the Δarm isoform cannot be enzymatically active. To investigate this point, we carried out methylation reactions with hypomethylated protein extracts using the full-length variant 1 of PRMT1 compared to the Δarm form of the same variant. In particular, HEK293 cells were transfected with PRMT1v1:EYFP and PRMT1v1Δarm:EYFP plasmids. 24 hours post-transfection, cells were lysed and the lysates subjected to immunoprecipitation with an anti-GFP antibody. The precipitated enzymes were incubated with hypomethylated HEK293 total cell extract and with a radioactive methyl donor, S-Adenosyl-L-[methyl-^{3}H]Methionin (SAM[^{3}H]) at 37^{o}C for 4 hours. Due to the lower Δarm expression levels, four times more precipitate was used in comparison to the full-length protein. The hypomethylated cell extract was obtained from HEK293 cells

treated with 15 µM of the S-adenosylhomocysteine hydrolase inhibitor, adenosine-2′,3′-dialdehyde (Adox). The methylation reaction products were analyzed by SDS-PAGE and the dried gel was exposed to an autoradiography film. In contrast to full length PRMT1, the Δarm form of variant 1 is incapable of methylating the hypomethylated cell extract since the activity of the Δarm form is not higher than the background which might appear due to the traces of active endogenous PRMT1 in the hypomethylated cell extract (Fig. 3.8B).

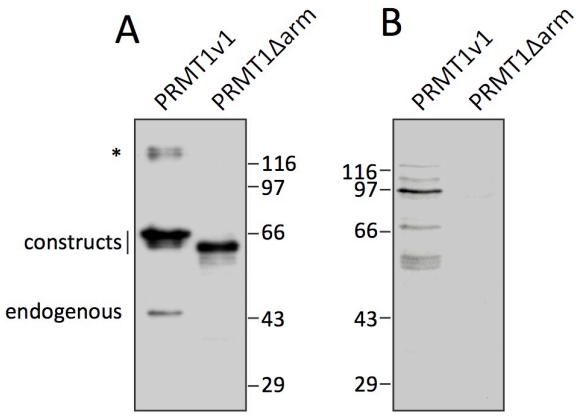


Fig. 3.8: The Δarm isoform does not dimerize with endogenous PRMT1 and is enzymatically inactive. (A) HEK293 cells were transfected with expression vectors encoding either PRMT1v1:EYFP or PRMT1v1Δarm:EYFP. 24 hours after transfection, whole cell extracts were prepared and subjected to immunoprecipitation with anti-GFP antibodies. Western blot with anti-PRMT1 antibody revealed the expressed fusion proteins and the endogenous PRMT1 monomer and oligomers, which co-immunoprecipitated with the full-length PRMT1v1 protein, but not with the Δarm isoform. (B) PRMT1 constructs were immunoprecipitated as in (A) and incubated with radioactive SAM[^{3}H] and hypomethylated whole cell extract from HEK293 cells. The products of the reactions were analyzed by SDS-PAGE and the gel was exposed to an autoradiography film. Δarm is incapable of methylating the hypomethylated substrates, apparently due to the absence of oligomerization with the other PRMT1 molecules.

3.1.9 Molecular dynamics of Δarm

Mobility of proteins is an important parameter because it can give information about if and how fast they move in the cell or if they are bound to cellular structures, which totally or partially immobilize them.

It has been earlier shown from our lab that an inactive form of PRMT1v2, obtained by site directed mutagenesis in the SAM binding site, accumulates in the nucleus and gains an immobile fraction there, but not in the cytoplasm. The immobile fraction is caused by stable interactions with nuclear substrates such as SAF-A and histone H3 (Herrmann and Fackelmayer, 2009). Considering the fact that Δarm is an inactive form of PRMT1, we have already observed that the localization of PRMT1v1 is not affected when the dimerization arm is removed from the protein. On the other hand, mobility might be

affected. In order to investigate the dynamics of Δarm and compare it to the full-length in live cells, we performed Fluorescence Recovery After Photobleaching (FRAP) experiments with HEK293 stably expressing the full-length and the Δarm form of PRMT1v1 tagged with EYFP. A rectangular bleaching area that included both cytoplasmic and nuclear parts in at least 20 cells was photobleached and the fluorescence recovery of each area (nucleus and cytoplasm individually) was monitored through 250 post-bleach images. The analysis was done by the FRAP wizard which is supplied with the Leica imaging software and later with Microsoft Excel. The percentage of fluorescence recovery during time is presented in Figure 3.9. The analysis revealed that both full-length PRMT1v1 and Δarm are fully mobile in the cytoplasm, with recovery half-time less than 3 seconds, behaving like soluble proteins. In contrast, in the nucleus, the two proteins recover to only 75% of the initial

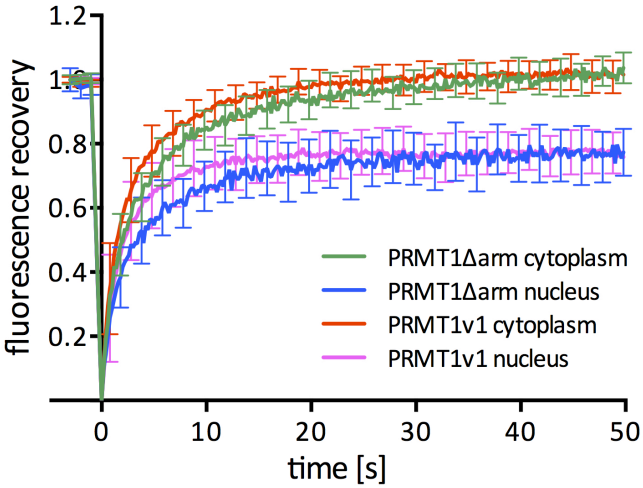


Fig. 3.9: Molecular dynamics of full-length PRMT1v1 and PRMT1v1Δarm. HEK293 cells stably expressing PRMT1v1:EYFP and PRMT1v1Δarm:EYFP were used for the analysis of the mobility of the two proteins. Rectangular regions including parts of the cytoplasm and the nucleus of at least 20 cells for each case were photobleached and analyzed for their fluorescence recovery, normalized to their initial fluorescence. The means of the fluorescence from all cells together with the SEM for each time point are presented here. In cytoplasm, both PRMT1 forms are fully mobile with a recovery half-time of 3 seconds, but both proteins have a 25% immobile fraction in the nucleus.

fluorescence. This result resembles earlier observations of the laboratory, in which Herrmann et al. (2005) had shown an immobilization of PRMT1v2 upon enzyme inhibition. The result for full length PRMT1v1 was surprising, as it differs significantly from PRMT1v2; this will require further research (see discussion). Interestingly, Δarm behaves very similar to full length PRMT1v1 with regard to its immobile fraction, but recovers more slowly, presumably indicating longer interaction with substrate proteins.

3.1.10 Involvement of Δarm expression in lung cancer

It has been previously reported that the expression levels of PRMT1 variants are altered between normal and cancerous breast tissues as well as between normal and

breast cancer cell lines (Baldwin et al., 2014; Goulet et al., 2007; Scorilas et al., 2000). Based on these observations, we decided to investigate whether something similar happens with Δarm between normal and lung cancer cell lines. For this purpose, two normal human lung fibroblasts cell strains (HFL-1 and MRC5) and four different human lung cancer cell lines (A549, H1299, H1437, H1792) were used for comparative assays (table 3.1). The cells lines were kindly provided by Prof. Evangelos Kolettas (Laboratory of Biology, School of Medicine, University of Ioannina, and Biomedical Research Division, Institute of Molecular Biology and Biotechnology-FORTH, Ioannina). Total RNA from all cell lines was isolated and reverse transcribed to cDNA, which was then amplified by PCR with primers corresponding to Δarm (forward primer 5'-CGGGACAAGTGGCTGGAGGTG-3', reverse primer 5'-CGAAACCGCCTAGGAACGCT-3'). GAPDH was used for normalization.

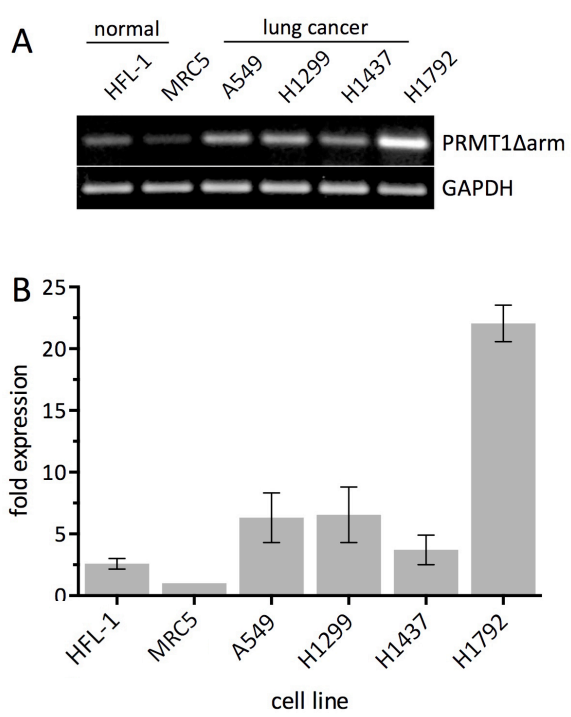


Fig. 3.10: PRMT1Δarm expression in two normal diploid lung fibroblast lines, HFL-1 and MRC5, and a panel of four lung cancer lines with different status of p53 and K-Ras (A549, H1299, H1437, H1792). (A) Total RNA was isolated from lung fibroblasts (HFL-1 and MRC5) and lung cancer cell lines (A549, H1299, H1437, H1792) and reverse transcribed to cDNA. PCR with these cDNAs and Δarm primers revealed unequal expression levels of Δarm among the cell lines. The lowest levels were detected in the normal lung fibroblasts. Cancer lines were different from this aspect with H1437 having the lowest and H1792 the highest levels. A549 and H1299 expressed intermediate levels. GAPDH was used as a loading control. (B) Quantitative RT-PCR was performed on RNA from the same cells. The means of PRMT1Δarm expression relative to MRC5 cells are given, together with their SEM. The data was normalized to GAPDH and analyzed by the ΔΔCt method.

The gel with the PCR products stained with ethidium bromide is presented in Figure 3.10A, demonstrating that the amounts of Δarm vary among the different types of cells. Normal human lung fibroblasts express the lowest levels of Δarm expression. In contrast, human lung cancer cells lines express increased Δarm levels with H1437 expressing the lowest and H1792 the highest levels. A549 and H1299 express intermediate and similar levels of Δarm.

In order to quantify these differences, we performed qPCR with the same cDNAs and the same primer pairs according to the KAPA SYBR® FAST qPCR Kit. The results were reproduced after the qPCR analysis with the ΔΔCt method. The relative expression levels of the four human lung cancer cell lines compared to each one of the normal human lung fibroblast strains are presented schematically in Figure 3.10B. According to our measurements, HFL1 fibroblasts express 2,6 times more than MRC-5. All cancer cell lines express higher levels of Δarm compared to fibroblasts. A549 and H1299 express 2,3 and 6,5 times more Δarm compared to HFL1 and to MRC-5 respectively. H1437 express 1,3 fold compared to HFL1 and 3,8 fold compared to MRC-5. H1792 expresses 9 and 22 times higher levels of Δarm compared to HFL1 and to MRC-5 respectively (Table 3.2). The fact that the more aggressive the cell line, the higher the levels of Δarm was a striking observation of potentially high importance. For example, H1792 cells which originate from a stage 4 cancer and is the only cell line with mutated K-Ras and almost null p53 had the highest levels of Δarm. This line is also a mixture of epithelial and mesenchymal cells, which means it might have gone through epithelial to mesenchymal transition (EMT), a change which occurs very often in cancer cells and has an important role in cancer progression, invasion and metastasis. Indeed, mesenchymal cancer cells are known for their invasiveness, aggressiveness and their metastatic abilities. On the other hand, p53 mutated H1437 cells, which have a wild type K-Ras express the lowest levels among the cancer cell lines.

Table 3.1: Characteristics of the cells used to compare Δarm expression levels between normal lung lung fibroblasts and human lung cancer cells.

Cells	Characteristics	Morphology	K-Ras	p53
HFL-1	Normal male human lung diploid fibroblasts	Spindle-shaped	wt	wt
MRC5	Normal male human lung diploid fibroblasts	Spindle-shaped	wt	wt
A549	58 years male, Caucasian, metastatic	Mostly epithelial but many mesenchymal	$K\text{-}Ras^{G12S}$	wt
H1299	43 years male, Caucasian	Epithelial	wt	null
H1437	60 years, male, Caucasian	Epithelial	wt	$p53^{R267P}$
H1792	50 years, male, Caucasian	Epithelial-Mesenchymal	$K\text{-}Ras^{G12S}$	mut (almost null)

Table 3.2: Δarm relative expression in lung cancer cell lines and normal lung fibroblasts.

Cells	Relative expression to HFL-1		Cells	Relative expression to MRC-5
HFL-1	1		**MRC5**	1
MRC5	0,42		**HFL-1**	2,60
A549	2,25		**A549**	6,45
H1299	2,30		**H1299**	6,70
H1437	1,32		**H1437**	3,79
H1792	8,96		**H1792**	22,13

3.1.11 Epithelial to mesenchymal transition (EMT) increases Δarm expression levels

The previous results suggested that the most aggressive lung cancer cell line expresses the highest Δarm levels. We therefore decided to test whether we can experimentally affect the level of Δarm by forcing cells to undergo Epithelial to Mesenchymal Transition (EMT). EMT is a normal morphogenetic process, which happens during the formation of the embryo. In addition, EMT has been observed to occur during cancer development, and it contributes to cancer progression. In this process, epithelial cells undergo molecular and morphological changes including loss of epithelial polarity and adhesive properties, and gain a fibroblastoid phenotype with increased cell motility and invasiveness. This procedure is governed by epigenetic mechanisms through DNA methylation, histone modifications and miRNAs. Snail is a transcriptional repressor, which plays an important role in EMT by suppressing genes supporting an epithelial phenotype such as *CDH1* which encodes E-cadherin (Tiwari et al., 2012; Wu et al., 2012).

For our experiments, we used a well-characterized plasmid, which expresses a mutated Snail variant (Snail 6SA; Addgene plasmid #16221) and has successfully be used in the literature to change the epithelial character of the cells to mesenchymal (Zhou et al., 2004). HEK293 cells were transfected with two different methods (Lipofectamine 2000, LF and Polyethylenimine, PEI) and with pCMV-Tag 2B vector encoding Flag:Snail 6SA. Transfected cells were selected with G418 for 5 weeks. Then, they were fixed and incubated with the monoclonal M2-anti-Flag antibody (Sigma) and SYTOX Green. A representative confocal image of a nucleus of a cell is presented in Figure 3.11B, where the expression and the nuclear localization of Snail are verified. Snail is co-localized with

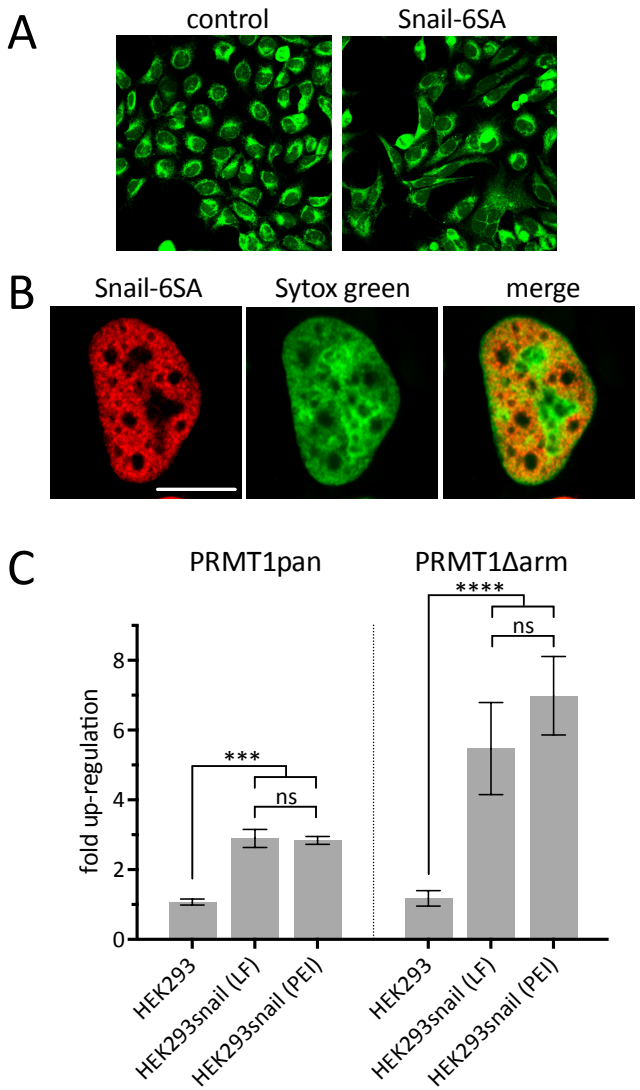


Fig. 3.11: PRMT1 and Δarm up-regulation upon stable Snail expression in HEK293 cells. (A, B) Cells' shape became more spindle-like after transfection with Snail. The expression and the correct localization of Snail were verified after fixation and immunofluorescence of the stable cells with the anti-Flag antibody (red) and SYTOX Green as a DNA counterstaining (green). A representative nucleus is shown here (scale bar: 10 μm). (C) HEK293 cells were stably transfected with Flag:Snail 6 expression vector either with Lipofectamine 2000 (LF) or Polyethylenimine (PEI). The expression of PRMT1 and Δarm was measured with qPCR with primer pairs corresponding to all seven amino-terminal variants of PRMT1 and Δarm respectively. The mean of individual experiments is presented graphically together with SEM for each case. PRMT1 was up-regulated 2,9 fold in both cases and Δarm 5,47 and 7 fold in LF and PEI transfected cells respectively. The stars indicate statistical significance (three stars: $p<0.01$, four stars: $p<0.005$).

SYTOX Green in the nucleus of the stable cell line. Moreover, transfected and untransfected cells were stained with the green-fluorescent dye, DiOC6 (Thermo Fisher Scientific), in order to visualize membranes and thus the shape of the cells. The population, which was expressing Snail, had an increased number of spindle-shaped cells in comparison to the untransfected population. This morphology is typical for mesenchymal cells and verifies that Snail triggered the desirable changes in the cells. After selection, total RNA from normal and stably transfected cells was purified and transcribed to cDNA. The cDNAs were analyzed by qPCR with a primer pair annealing to Δarm and another one annealing to conserved regions among the seven amino-terminal variants of PRMT1. GAPDH was used for normalization. The data was analyzed by the ΔΔCt method and depicted graphically in Figure 3.11C. Interestingly, total PRMT1 expression (all expressed splicing isoforms) was increased about 2,9 fold in both cases

(Lipofectamine and Polyethylenimine transfected cells). On the other hand, Δarm expression was increased by 5,5 and 7 fold in Lipofectamine and in Polyethylenimine transfected cells respectively.

These results demonstrate that overexpression of Snail can increase the expression levels of Δarm and PRMT1 in HEK293 cells. We conclude that the higher levels of Δarm are associated with the mesenchymal cancer cell phenotype, and its impact in cancer will be investigated in future experiments.

Table 3.3: Up-regulation factors of PRMT1 and Δarm upon Snail transfection.

Cells	PRMT1 fold	Δarm fold
Untransfected HEK293	1	1
LF transfected HEK293 with Flag:Snail 6SA	2,9	5,47
PEI transfected HEK293 with Flag:Snail 6SA	2,84	7

3.1.12 Δarm expression during differentiation

As shown above, the levels of Δarm are altered in lung cancer cell lines. Cancer, on the one hand, is a pathological condition where many changes occur in the cell in all aspects. Differentiation, on the other hand, represents also a condition where the cell changes, but this process is physiological. It is known that PRMT1 acts together with its paralogue PRMT8 to co-ordinate neuronal differentiation through retinoic acid signaling. PRMT1 is highly expressed in undifferentiated embryonic stem cells, acting mainly as a selective gene repressor of a large set of retinoic acid-induced genes including Hoxa1, Hoxb1, Pmp22 and Spsb1. Its expression remains constant during neural differentiation showing a slight increase upon RA-treatment (Simandi et al., 2015). For these reasons, we decided to examine if and how Δarm is regulated during differentiation of LUHMES cells (compare Fig. 3.2 for the expression of Δarm in undifferentiated LUHMES) in order to have a comparison of a physiological and a pathological condition in a cell culture context. To this end, LUHMES cells were differentiated and total RNA was prepared from undifferentiated cells (day 0) and cells which were induced to differentiate as well as terminally differentiated cells (day 1-8, day 10, day 12 and day 14). The isolated RNAs

were reverse transcribed to cDNA, followed by PCR with the Δarm primers. The PCR products were analyzed by agarose gel electrophoresis and the results are presented in figure 3.12A. Δarm mRNA was detected in the undifferentiated cells and in all stages of differentiation, but clearly decreased after day 7, when the neurons are mature.

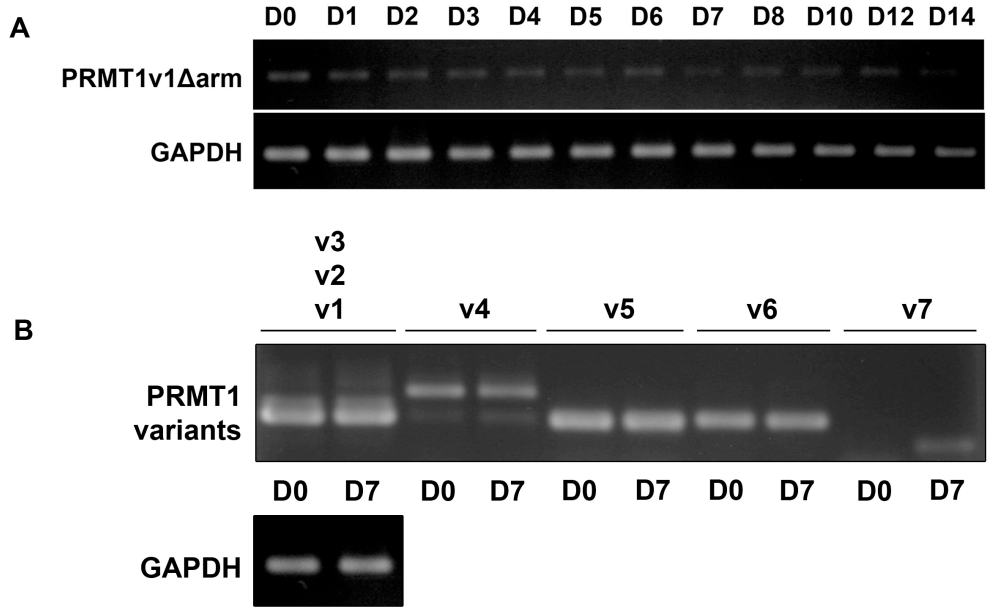


Fig. 3.12: Δarm and PRMT1 amino-terminal variants expression during neuronal differentiation of LUHMES cells. (A) LUHMES cells were differentiated and total RNA was isolated from undifferentiated (day 0, D0) and differentiating/differentiated cells of day 1–8, 10, 12, 14 (D1-D8, D10, D12, D14). Δarm expression was analyzed by RT-PCR, and the products were analyzed on an agarose gel. (B) cDNAs from undifferentiated (D0) and differentiated (D7) LUHMES were analyzed by PCR with primer pairs specific to the seven different amino-terminal variants of PRMT1. Variants 1-3 were amplified by the same primer pair and the three products can be discriminated by their size differences. The remaining variants (v4-v7) were amplified by a different forward primer and a reverse primer common for all variants. GAPDH was used for normalization in both experiments.

For comparison, we tested the expression of the seven amino-terminal variants of PRMT1 between undifferentiated and differentiated cells. cDNA from undifferentiated (day 0) and differentiated (day 7) LUHMES was used as a template with primer pairs specific for each one of the seven PRMT1 amino-terminal variants. In Figure 3.11B the agarose gel with the PCR products for each variant are presented in pairs of undifferentiated and differentiated cells. For variants 1, 2 and 3 a common primer pair was used and the three different products are discriminated by their size (see methods, Table 2.3). There are not striking differences between the two states for variants 1-6, but variant 7 is only detectable after cell differentiation. The absence or presence of variant 7 in undifferentiated and differentiated cells, respectively, creates intriguing questions

since variant 7 is enzymatically inactive and its role should be elucidated in future experiments.

3.2 Novel PRMT1 substrates

Our interest about PRMTs includes the identification of new substrates, which are being methylated by this family of enzymes. In collaboration with Dr. Panagiota Georgoulia from Prof. Anastasia Politou's group (Laboratory of Biological Chemistry, School of Medicine, University of Ioannina), whose focus lies on Intrinsically Disordered Proteins (IDPs), we investigated whether Lamin B Receptor (LBR) and Scaffold Attachment Factor B (SAF-B) are methylated by PRMTs.

IDPs are proteins or domains in proteins, without defined secondary or 3D structure. They are usually rich in hydrophilic amino acids, which makes them ideal substrates for post-translational modifications. The combination of different modifications gives plasticity to these molecules through the alteration of hydrogen bonds or salt bridges. Lamin B Receptor (LBR) is localized in the nuclear envelope and participates in a variety of functions such as the tethering of nuclear lamina to the inner nuclear membrane and chromatin remodeling. The nucleoplasm-facing amino-terminal part of the protein has been shown to mediate most of LBR's interactions. It harbors a 60-amino acid Tudor domain followed by a 40-residue region rich in charged amino acids (Arginines and Serines repeats, RS region), and a third 110-amino acid segment (SG domain) with no apparent homologies. RS region also exists in other IDPs and includes multiple phosphorylation sites. SAF-B participates in chromatin organization, regulation of gene expression and RNA maturation. This protein contains regions rich in Arginine-Glutamic acid dipeptides (p43 domain), and it also belongs to the IDP group. Hence, SAF-B is also a good candidate for post-translational modifications.

Our aim was to investigate the possibility of post-translational arginine methylation in LBR and/or SAF-B molecule by members of the PRMT family. Regarding LBR, our studies were focused in the amino-terminus of the protein and in the domains that constitute it (i.e. TD, RS and SG).

First, we tested whether the amino-terminal region of LBR could be methylated by one of the PRMTs. The enzymes with the highest possibility of methylating LBR, because of

their localization in the cell nucleus and their abundance and ubiquity, were PRMT1, PRMT4 or PRMT6. HEK293 cells were transiently transfected with three different vectors that express-under the control of a CMV promoter - PRMT1, PRMT4 and PRMT6 tagged with GFP in their amino-terminus. Twenty-four hours post-transfection, and after the expression of the enzymes was verified by observing the cells under the fluorescence microscope, the cells were lyzed in lysis buffer. The lysates were subjected to immunoprecipitation with anti-GFP antibodies. The precipitated enzymes were incubated with either the amino-terminal region of LBR or hypomethylated HEK293 total cell extract (as a control) and with a radioactive methyl donor, S-Adenosyl-L-[methyl-^{3}H]Methionin (SAM[^{3}H]) at 37^{o}C for 4 hours. A bacterially expressed LBR peptide was purified by Dr. Panagiota Georgoulia from Prof. Anastasia Politou's lab (School of Medicine, University of Ioannina). The hypomethylated cell extract was obtained from HEK293 cells treated with 15 μM of the S-adenosylhomocysteine hydrolase inhibitor, Adenosine-2′, 3′-dialdehyde (Adox). The methylation reaction products were analyzed by SDS-PAGE and the dried gel was exposed to an autoradiography film. According to the controls, PRMT1 strongly methylates the hypomethylated extract. PRMT4 and PRMT6 have lower enzymatic activity than PRMT1, with PRMT6 being very weak. This was expected because these three enzymes possess different enzymatic activity (Herrmann et al., 2009). In the assays with the LBR amino-terminus, only the incubation with PRMT1 gave a strong signal of methylated protein; the assays with PRMT4 and PRMT6 were negative (Fig. 3.13A). Control experiments without PRMT1 were also negative (Fig. 3.13B). Thus, LBR is a novel substrate of PRMT1. The next question we addressed was to clarify which domain of the LBR amino-terminus is being methylated by PRMT1. The domains that were used was TD, RS (two different preparations 1 & 2) and the combination of both, SG and SG combined with RS. These experiments show that methylation is restricted in the RS domain of LBR, and is independent of the presence or absence of Tudor domain. On the other hand, when RS was combined with SG domain the methylation was much weaker (Fig. 3.13C+D).

Finally, we carried out methylation reactions using the bacterially expressed RS peptide, an unmodified synthetic part of RS (NH_2-SSPSRRRGSRSRSRSRS-COOH) and three synthetic RS peptides with phosphorylations in different Serines (RS 1 - pS71, RS 2 – pS86,

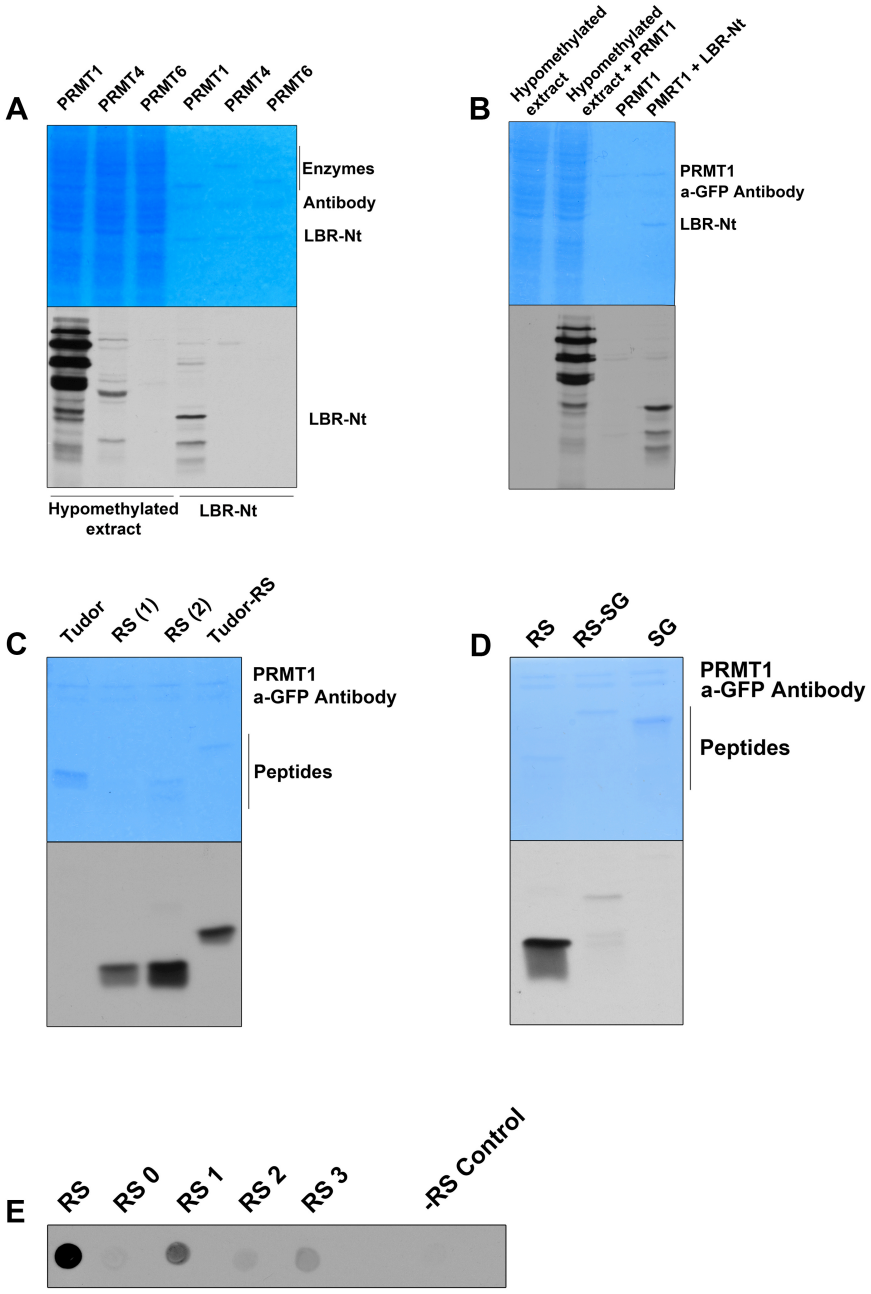


Fig. 3.13: LBR methylation by PRMT1. (A) HEK293 cells were transiently transfected with GFP tagged PRMT1, PRMT4 and PRMT6. 24 hours after transfection the cells were lysed in lysis buffer. The lysates were subjected to immunoprecipitation with antibodies against GFP. The precipitated enzymes were used for methylation reactions with hypomethylated extracts from HEK293 cells or with the bacterially expressed amino-terminal region of LBR as substrate and SAM[^{3}H] as a methyl donor. PRMT1, PRMT4 and PRMT6 methylate the hypomethylated extract with different intensity. PRMT1 is the only enzyme that methylates LBR amino-terminal region. (B) In the absence of PRMT1 neither the hypomethylated extract nor LBR give signal for methylation. (C), (D) Different domains and combinations of them from the LBR amino-terminal region (Tudor, RS, Tudor-RS, RS-SG and SG) were used as substrates in order to identify which of them are being methylated. The region of LBR that is being methylated by PRMT1 is the RS domain and the methylation is independent of the presence of Tudor domain. However, the presence of SG domain inhibits the methylation of RS. (E) Phosphorylated Serines in different positions of synthetic RS peptides affect the methylation by PRMT1 positively or negatively. Phosphorylation of Serine 71 (RS 1) has the strongest methylation signal. Unmodified RS (RS 0) and RS with phosphorylated Serine 86 (RS 2) have the weakest methylation.

RS 3 – pS78 & pS80). A minus RS control reaction was used as a reference point of the background signal. Due to the small size of the synthetic peptides, the reaction products were dot-blotted on a PVDF membrane, which was exposed on an autoradiography film. We find that the bacterially expressed RS peptide was indeed strongly methylated; this was further confirmed with the synthetic peptide RS1 (Fig. 3.13E). Peptide RS3 (phosphorylated on Serine 78) was also methylated, although weaker. On the other hand, we found almost no methylation for the unmodified (RS 0) and the S86 phosphorylated (RS 2) synthetic peptides.

Following these experiments with LBR as a substrate, we attempted to express in bacteria and purify p43 domain of SAF-B (rich in Arginine-Glutamic acid dipeptides) in order to use it in similar experiments. Several conditions of expression and purification were tested, but we did not manage to get reasonable amounts of soluble protein, which could be used for *in vitro* methylation assays.

3.3 Tools for the study of PRMT8

PRMT8 is the eighth member of the protein arginine methyltransferases family. It is a paralogue of PRMT1, but contains a unique amino-terminus with a myristoylation anchor, which targets the protein on the plasma membrane. In contrast to the other ubiquitous PRMT family members, the expression of PRMT8 is restricted in the central nervous system. This well-defined expression profile, in combination with the very specific localization, creates challenging questions about its role in neuronal differentiation and/or maintenance. However, due to the lack of available tools such as antibodies and appropriate model systems, and the fact that this protein has a strong homology with highly abundant PRMT1 which makes creating specific antibodies very challenging, PRMT8 has been poorly studied. To date, very few studies about PRMT8 have been published, and often with controversial results. Most of these studies have been done in models, which do not endogenously express PRMT8 (e.g. HeLa cells), combined with the overexpression of alternative hypothetical forms of the overexpressed protein. Recently, some publications studied endogenous PRMT8 in human embryonic stem cells and induced regeneration competent cells (Hernandez and Dominko, 2016), mouse stem cells (Solari et al., 2016) and neurons obtained from mouse stem cells upon retinoic acid

treatment (Simandi et al., 2015). Our aim was to set up a solid human neuronal differentiation model system, which expresses PRMT8 endogenously, as well as to create and validate tools for studying the role of this molecule in detail, in order to elucidate its physiological function in neurons and its potential role in diseases.

3.3.1 PRMT8 is endogenously expressed in differentiated LUHMES - Morphological and molecular characterization of the system

One of the aims of the present thesis was to set up new tools for studying the role of PRMT8 in neuronal differentiation and/or maintenance. Different cellular models were analyzed for their ability to differentiate to mature neurons, which express endogenous PRMT8. For example, PRMT8 mRNA was detected in low levels in mouse embryonic stem cells (mESCs) and in spontaneously differentiating mESCs after four days upon removal of leukemia inhibitor factor (LIF) (data not shown). The expression of PRMT8 in mESCs is compatible with recent findings (Solari et al., 2016). However, our interest was focused on the role of PRMT8 in differentiating or mature human neurons and not in a mixed mouse cell population. mESCs on the one hand, is a mouse model system and, on the other hand, it is challenging to set up protocols which reproducibly give high percentage of neurons after their differentiation. Our purpose was therefore to set up a human cellular system to study neuronal differentiation, which could differentiate efficiently and reproducibly in mature neurons with known molecular and morphological characteristics and with higher expression levels than the spontaneously differentiating mouse stem cells. As described in detail in the introduction, we chose LUHMES cells as a model for neuronal differentiation. This mesencephalic-derived human cell line can be induced to homogenously differentiate to post-mitotic mature dopaminergic neurons in almost 100% within 5 days. LUHMES cells are a subclone of the originally generated MESC2.10 cell line, which was obtained from human embryonic mesencephalic tissue that was conditionally immortalized by infection with a LINX v-myc retroviral vector system. In this system, the expression of *v-myc* oncogene is under the control of tetracycline. In the absence of tetracycline, LUHMES cells proliferate continuously. After the addition of tetracycline, *v-myc* expression is inhibited. Consequently, the cells stop proliferating and start differentiating to post-mitotic neurons (Lotharius et al., 2005). LUHMES cells were

kindly offered by Prof. Marcel Leist (Chair of *in vitro* Toxicology and Biomedicine, University of Konstanz, Germany), after my visit to his lab in order to be trained on the manipulation of the cells.

Although LUHMES cells are well-characterized (Scholz et al., 2011) to undergo profound morphological and molecular changes after several days of differentiation, we decided to ensure that they also behave as expected in our hands. To this end, we performed maintenance culture and several differentiations of the cells in the way is described in methods, and observed their morphology on daily basis. Moreover, we examined the cells for up-regulation or down-regulation of marker genes, either by RT-PCR or immunofluroescence with specific primers and antibodies respectively.

In Figure 3.14 the morphological changes after 6 days of differentiation are presented. While being kept in maintenance medium, cells grow in colonies and no neurites are observed (day 0). After shifting to differentiation medium and applying the two-step differentiation protocol, the cells gradually and synchronously create a network of neurons with long neurites (day 6). For a better visualization of the colonies and the neuronal network, in the undifferentiated and differentiated cells, respectively, LUHMES were stained with an antibody against the cytoskeletal protein α-tubulin. The protein was present in both states, distributed in the undifferentiated cells and along the somata and the axons of the differentiated cells.

Following the morphological observation of LUHMES differentiation, we proceeded to the molecular investigation of the procedure. Total mRNA was isolated from undifferentiated (day 0) and differentiated (day 7) LUHMES. RNA was reverse transcribed to cDNA and subjected to PCR with primer pairs specific for different marker genes. GAPDH and Actin beta were used for normalization (Fig. 3.15). The list of the primer pairs which were used in these experiments is given in methods. We find that markers such as SOX2 (sex determining region Y-box 2), PAX3 (paired box 3) and GMNN (geminin) are down-regulated. SOX2 is a transcription factor, essential for maintaining self-renewal and pluripotency of undifferentiated embryonic stem cells. It is also known for playing an important role in the maintenance of neural stem cells. The expression of SOX2 is down-regulated during LUHMES differentiation. The same kinetic behavior is observed for PAX3

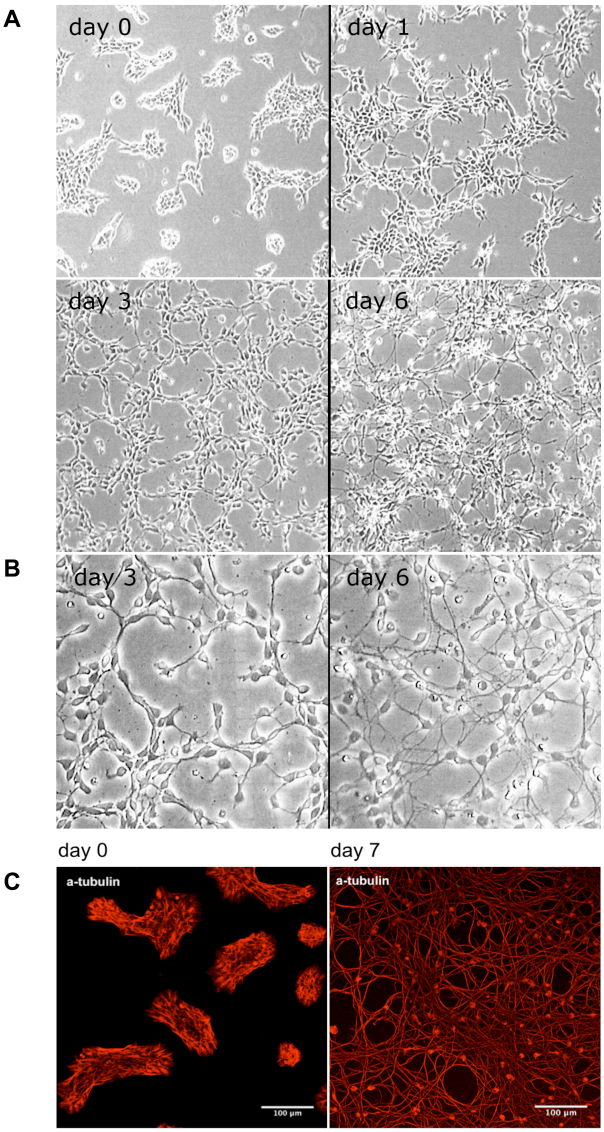


Fig. 3.14: Morphological changes during LUHMES differentiation. LUHMES cells display morphological changes during their differentiation from immortalized, undifferentiated cells to post-mitotic neurons. Before induction (day 0), the cells are growing in clusters and have no apparent protrusions. After induction with tetracyclin and replating the cells, the population creates a homogeneous network of functional neurons. (A) day 0 shows undifferentiated cells, day 1 and day 3 shows the changes during differentiation and day 6 shows mature post-mitotic neurons, 10x magnification lens. (B) day 3 (intermediate) and day 6 (mature neurons) of differentiation, 20x magnification lens. (C) For better visualization of the morphological changes, day 0 and day 7 LUHMES were cultivated on glass coverslips, fixed and stained with a primary anti-α-tubulin antibody and a secondary anti-mouse conjugated with Alexa 568(red).

and GMNN, a migrating neuroblast marker and a cell cycle control marker controlling G1/S phase transition, respectively. The traces of GMNN, which are detected on day 7 are explained through the immunofluorescence results in the next paragraph. On the other hand, NLGN (Neuroligin), a post-synaptic protein, which participates in the formation of the synapse and SYP (Synaptophysin), the major synaptic vesicle protein, are up-regulated. The same kinetic behavior is observed for DRD2 (dopamine receptor D2) and TH (tyrosine hydroxylase). NF165 (Neurofilament 165), a neuronal cytoskeleton protein, remains more or less the same. Finally, and most importantly PRMT8, the protein of our interest, appears only after differentiation of LUHMES. This result indicate, first, that PRMT8 is endogenously expressed in those cells and second, that PRMT8 expression is absent in undifferentiated cells and expressed only in differentiated LUHMES cells. This

suggests that it may be involved either in the differentiation process or in the maintenance of the mature neurons, or in both of them. Moreover, it verifies that LUHMES, apart from all the advantages they have as a cell line, are suitable for fulfilling our aim – the study of the role of PRMT8 in neurons.

In parallel with the transcriptional analysis of differentiation markers, LUHMES cells were cultured on glass coverslips and fixed with paraformaldehyde on day 0 and on day 7

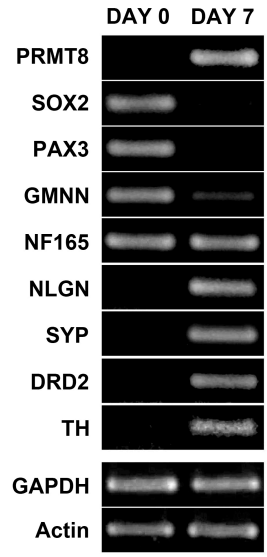


Fig. 3.15: Expression of markers in undifferentiated and differentiated LUHMES. Total RNA was isolated and reverse transcribed to cDNA from undifferentiated (Day 0) and differentiated (Day 7) LUHMES. The cDNAs were analyzed by PCR for the expression of markers related to neuronal differentiation. SOX2, PAX3 and GMNN are down-regulated. NF165 expression levels remain almost stable between the two states. On the other hand, NLGN, SYP, DRD2 and TH are up-regulated after differentiation. PRMT8 is also up-regulated on day 7. GAPDH and Actin were used for normalization.

for immunofluorescence experiments. Fixed cells were blocked with BSA and incubated with primary antibodies against Geminin (GMNN), Neurofilament 165 (NF165), Neurofilament 200 (NF200), synaptic vesicle protein 2 (SV2) and insulin gene enhancer protein Islet-1 (ISL1) (all antibodies were from Developmental Studies Hybridoma Bank; DSHB, Iowa), and secondary anti-mouse antibody conjugated with Alexa 568 (red). DRAQ5 was used as a DNA counterstain. The cells were observed under the microscope and confocal images were obtained.

Geminin is localized in the nucleus of undifferentiated cells, indicating that they are actively proliferating. After seven days of differentiation, Geminin has disappeared from the cells, which are post-mitotic. Very few cells “escape” the differentiation, and remain proliferating, creating small colonies between the neuronal network of the differentiated cells. Those cells are positive for Geminin like the undifferentiated cells. This also explains

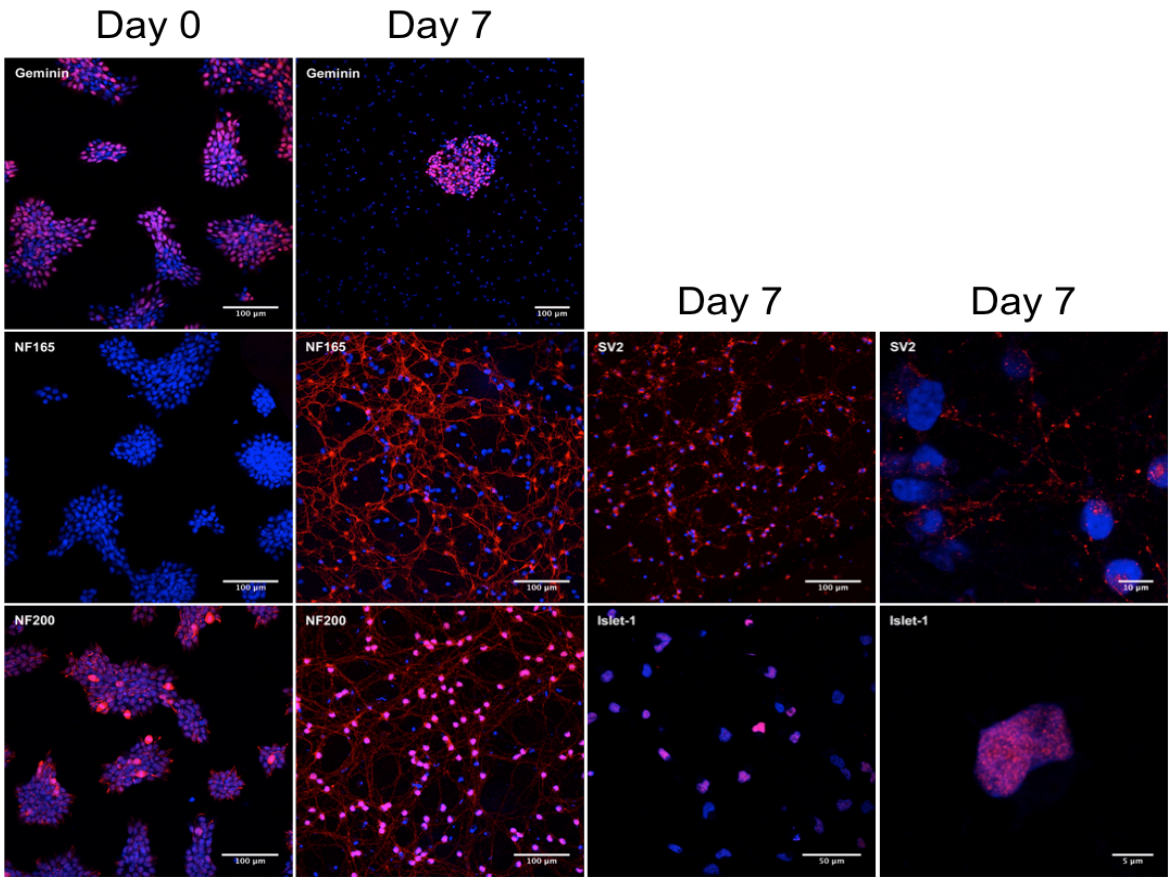


Fig. 3.16: LUHMES differentiation and expression of marker proteins. Undifferentiated and differentiated LUHMES cells were fixed with paraformaldehyde on day 0 and day 7 and stained with primary antibodies against Geminin (GMNN), Neurofilament 165 (NF165) and Neurofilament 200 (NF200). Differentiated cells of day 7 were also stained with antibodies against synaptic vesicle protein 2 (SV2) and insulin gene enhancer protein Islet-1 (ISL1). Primary antibodies were detected with a secondary anti-mouse antibody (red). DRAQ5 was used for staining DNA (blue). Proliferating LUHMES of day 0 are positive for the cell cycle marker Geminin but they become negative on day 7. Few cells, which "escape" the differentiation process and continue growing in clusters, remain positive on day 7. Undifferentiated LUHMES are negative for NF165 but positive for NF200. Both neuron specific cytoskeletal proteins are present in differentiated LUHMES. Day 7 cells are also positive for SV2 and Islet-1. Scale bars are indicated individually on every image.

the traces of Geminin mRNA detected by RT-PCR in the differentiated cells in the previous experiment.

Undifferentiated cells were negative for the neuronal cytoskeleton specific protein Neurofilament 165, but they were positive for Neurofilament 200. Differentiated LUHMES of day 7 were positive for both Neurofilament 165 and Neurofilament 200. Differentiated cells were also positive for the synaptic vesicle protein 2 (SV2), an abundant and conserved component of synaptic vesicles in vertebrates and Islet-1, a member of the family of homeodomain containing transcription factors. Islet-1 is expressed in all islet

cells in the pancreas and is considered to be an early marker for motor neuron differentiation (Fig. 3.16).

3.3.2 PRMT8 mRNA is highly up-regulated during LUHMES differentiation

As described above, LUHMES cells can differentiate in a similar way in the presence or absence of dibutyryl-cAMP and/or glial cell derived neurotrophic factor (GDNF). The differentiated cells follow the same gene expression pattern and the dopamine uptake and release is also identical. However, tyrosine hydroxylase (TH) requires the continuous presence of cAMP (Scholz et al., 2011). We have already shown that PRMT8 is up-regulated in differentiated LUHMES using the +cAMP/+GDNF protocol. Next, we aimed to analyze whether PRMT8 is being expressed in LUHMES after omitting cAMP and GDNF from the differentiation medium. The experiment was performed by collecting samples in different time points of differentiation. Total RNA was purified from the samples, and reverse transcribed to cDNA, followed by qPCR or semi-quantitative PCR. For the qPCR, cDNAs from day 0, 2, 4 and 6 were used in order to quantify the relative expression levels of PRMT8 to those that the undifferentiated cells (day 0) express. The results were normalized to GAPDH and analyzed by the ΔΔCt method. In parallel, the relative expression of PRMT1 and PRMT6 were measured for comparison. We find that PRMT8 was strongly up-regulated on day 6 (257 fold); on the other hand PRMT1 was slightly down-regulated, and PRMT6 slightly up-regulated (fig. 3.17A and table 3.4). Thus, PRMT8 up-regulation does not require cAMP and GDNF.

In order to monitor PRMT8 expression levels over a longer time course, we used cDNAs from day 0-8, 10, 12 and 14 as PCR templates with the same specific primers for PRMT8 (forward primer 5'-CCAAAGCAAGTGGTGACCAATG-3', reverse primer 5'-TCTACTGTGAAATCGAGGTCTCG-3') and GAPDH for normalization. The semi-quantitative PCR products were analyzed by agarose gel electrophoresis, stained with ethidium bromide. PRMT8 starts appearing on day 2 and increases up to day 7. From day 8 until day 14, expression of PRMT8 decreases gradually (Fig. 3.17B).

The expression levels of PRMT1 variants during LUHMES differentiation are presented in chapter 3.1.11. In comparison to PRMT8, PRMT1 variants remain more or less stable after differentiation, with variant 7 being the only exception.

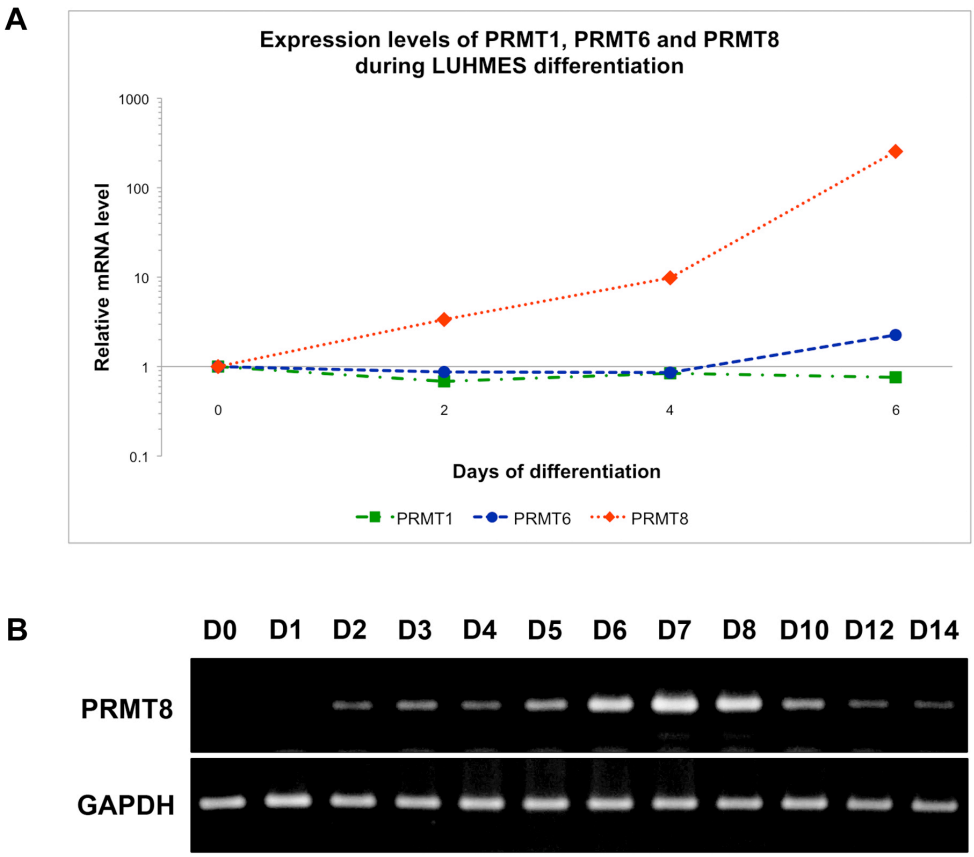


Fig. 3.17: PRMT8 up-regulation during LUHMES cell differentiation. Total RNA was isolated from LUHMES cells on day 0 to 8, day 10, 12 and 14. The RNAs were reverse transcribed to cDNA. (A) The expression levels of PRMT1, PRMT6 and PRMT8 were measured by qPCR on day 0, 2, 4 and 6. In contrast to PRMT1 and PRMT6, PRMT8 was up-regulated by 257 fold on day 6 of differentiation. (B) The expression levels of PRMT8 were monitored on day 0-8, day 10, 12 and 14 (D1-D8, D10, D12, D14). PRMT8 increases during differentiation, reaching a peak on day 7, and then decreases until day 14. GAPDH was used for normalization.

Table 3.4. Expression levels of PRMT1, PRMT6 and PRMT8 during LUHMES cell differentiation. PRMT1 and PRMT6 are not significantly modulated during differentiation. On the other hand, PRMT8 is highly up-regulated by 257 fold.

	PRMT1	PRMT6	PRMT8
Day 0	1,00	1,00	1,06
Day 2	0,69	0,93	3,44
Day 4	0,89	0,91	10,28
Day 6	0,75	2,26	257,29

3.3.3 PRMT8 antibodies

A significant obstacle in the studies of PRMT8 is the lack of a good commercially available antibody that recognizes endogenous PRMT8, but not the highly related and much more abundant PRMT1. The fact that many companies create anti-PRMT8

antibodies and later qualify them in cell lines where PRMT8 should not be present (e.g. HEK293 cells) or use the recombinant protein to show that their antibody react with PRMT8 in Western blots, creates suspicions about their quality and reliability. We have therefore decided to create new antibodies to PRMT8. To this end, we started collaborating with the American company Cell Signaling that is commercializing high quality antibodies to a large variety of proteins of medical interest. Different regions of PRMT8, mainly parts of its unique amino-terminus as well as the full protein sequence, were used as epitopes from immunization of rabbits. Over twenty different crude sera and purified antibodies from the immunized rabbits were sent to our laboratory for validation. As the first step of testing, the sera and the antibodies were checked for their ability to recognize the recombinant PRMT8. To this end, HEK293 cells were transiently transfected with expression vectors for GFP-tagged PRMT8. One day after transfection, whole cell protein extracts were prepared by lysing the cells with SDS, and precipitated according to Wessel&Flügge. The extracts were analyzed by SDS-PAGE and the results were visualized by immunoblotting. The different rabbit sera and the antibodies were used in a concentration of 1:1000. All tested sera and antibodies reacted with PRMT8, with different affinity, and a band appeared in the expected size (a representative Western blot with some of the tested antibodies is shown in Fig. 3.18). The next and more important step was to test whether the sera and the antibodies can be used for the immunodetection of endogenous PRMT8. Therefore LUHMES were differentiated for seven days and the procedure described above for HEK293 cells was followed. We knew already that on day 7, LUHMES express the highest levels of PRMT8 mRNA, and for this reason we chose to take protein extracts on that day. Extracts from undifferentiated LUHMES were used in parallel for comparison with those from the differentiated cells. To our surprise, none of the sera or of the antibodies detected endogenous PRMT8 in differentiated cells, even in higher concentrations (up to 1:100) and in low stringency buffers (TNT). There was either no signal at all or non-specific bands, but there was no protein band in the expected molecular weight size of 42 kDa, which does not exist in the undifferentiated cells and appears upon differentiation.

We have also tried to generate ourselves antibodies using an epitope from a sequence (NH_2-CPGRGKMSKLLNPEEMT-COOH), corresponding to a region in the amino-terminus of

PRMT8. The same sequence has been used earlier and it is described earlier (Kousaka et al., 2009). Two rabbits were immunized twice, and samples from their blood were received for testing before the final bleeding. Unfortunately, one of the rabbits died during the procedure. The final sera and the purified antibody (by affinity chromatography with the peptide we used for immunization) were tested in the same way we did with the Cell Signaling antibodies. Also our own antibody successfully detects recombinant PRMT8 expressed in HEK293 cells, but fails to detect the endogenous PRMT8 in LUHMES cell extracts (data not shown). For the reasons described above, we had to continue with experiments that did not rely on an antibody to PRMT8.

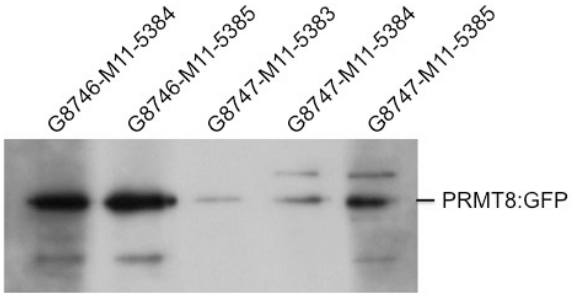


Fig. 3.18: PRMT8 antibodies test. Extracts from HEK293 cells overexpressing PRMT8:GFP were analyzed by Western blotting. The membrane was incubated with five different antibodies against PRMT8, supplied by Cell Signaling Company. The five antibodies react with the recombinant PRMT8 with different affinity each.

3.3.4 PRMT8 hypothetical amino terminus

Until today, there are apparently conflicting data in the literature about the full sequence of PRMT8. There are two different opinions, which are described in detail in the introduction. Briefly, PRMT8 is believed to acquire a unique amino-terminus with a myristoylation motif, which targets the protein on the plasma membrane (Lee et al., 2005). Earlier work suggested that PRMT8 does not include the first 15 amino-terminal amino acids, thus the myristoylation motif is absent and the protein is not associated with the plasma membrane. This was explained by the usage of an alternative starting codon 16 amino acids downstream of the original one (Kousaka et al., 2009). Moreover, alternative variants are given in GenBank.

In order to test which PRMT8 isoform is being expressed in LUHMES we designed different forward primers and combined them with a common reverse primer (P8rev; 5'-GTGGTTGGCCTTAATGATCTTC-3'), which anneals on a sequence of exon 3 of PRMT8. The designed forward primers anneal in regions 5' of the known exon 1, and which, according to their sequence and homology with a human sequences (GenBank accession:

NM_019854.4, designated "variant 1" of PRMT8, NM_001256536.1, designated “variant 2” of PRMT8) or a rat EST (GenBank accession #XP_002726479.1, designated "PRMT8-like"), respectively, could represent alternative 5' exons. Thus, they could encode different amino termini of human PRMT8. We used as a template RNA from differentiated (day 7) LUHMES cells. The PCR products were analyzed by agarose gel electrophoresis and visualized by ethidium bromide. The only transcript that was detected was variant 1 with two independent alternative forward primers. The other PCR reactions did not give any products (Fig 3.19). From these experiments we concluded that the sequences we chose for alternative forward primers are not present in naturally occuring variants of PRMT8.

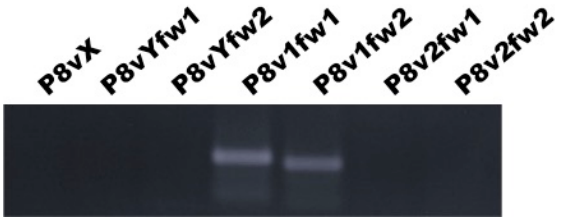


Fig. 3.19: PRMT8 hypothetical amino terminus. Differentiated LUHMES cDNA was analyzed by RT-PCR for the expression of alternative isoforms of PRMT8. Two alternative variants (variant 1 and variant 2) and two hypothetical variants (variant X and variant Y) from a rat EST were checked with two alternative primers each (except for variant X that was analyzed with one forward primer). Only variant 1 (GenBank accession number NM_019854.4) was found to be expressed in differentiated LUHMES cells of day 7. The legends above the image represent the name of the forward primer, which was used in each PCR. P8vX: PRMT8 hypothetical variant X; P8vYfw1/2: PRMT8 hypothetical variant Y-alternative forward primer 1/2; P8v1fw1/2: PRMT8 variant 1-alternative forward primer 1/2; P8v2fw1/2: PRMT8 variant 2-alternative forward primer 1/2. The reverse primer was common in all PCRs.

3.3.5 Subcellular localization of PRMT8

As it has been stated earlier, the first published form PRMT8 has a unique amino-terminus, which contains a myristoylation on glycine 2, and is thus anchored on the plasma membrane. This hypothesis was proved by overexpression of GFP fusion constructs of PRMT8 in HeLa cells (Herrmann et al., 2009; Lee et al., 2005). Later, a different group suggested that the myristoylation motif was not included in PRMT8 protein sequence due to the fact that the translation of the protein was starting from a start codon 16 amino acids downstream of the original one. This leads to a distribution of PRMT8 in the cell with a dominant nuclear pattern, instead of being bound to the plasma membrane (Kousaka et al., 2009). As the localization of a protein is crucial for its function, we decided to clarify these controversial opinions. To this end, we performed localization

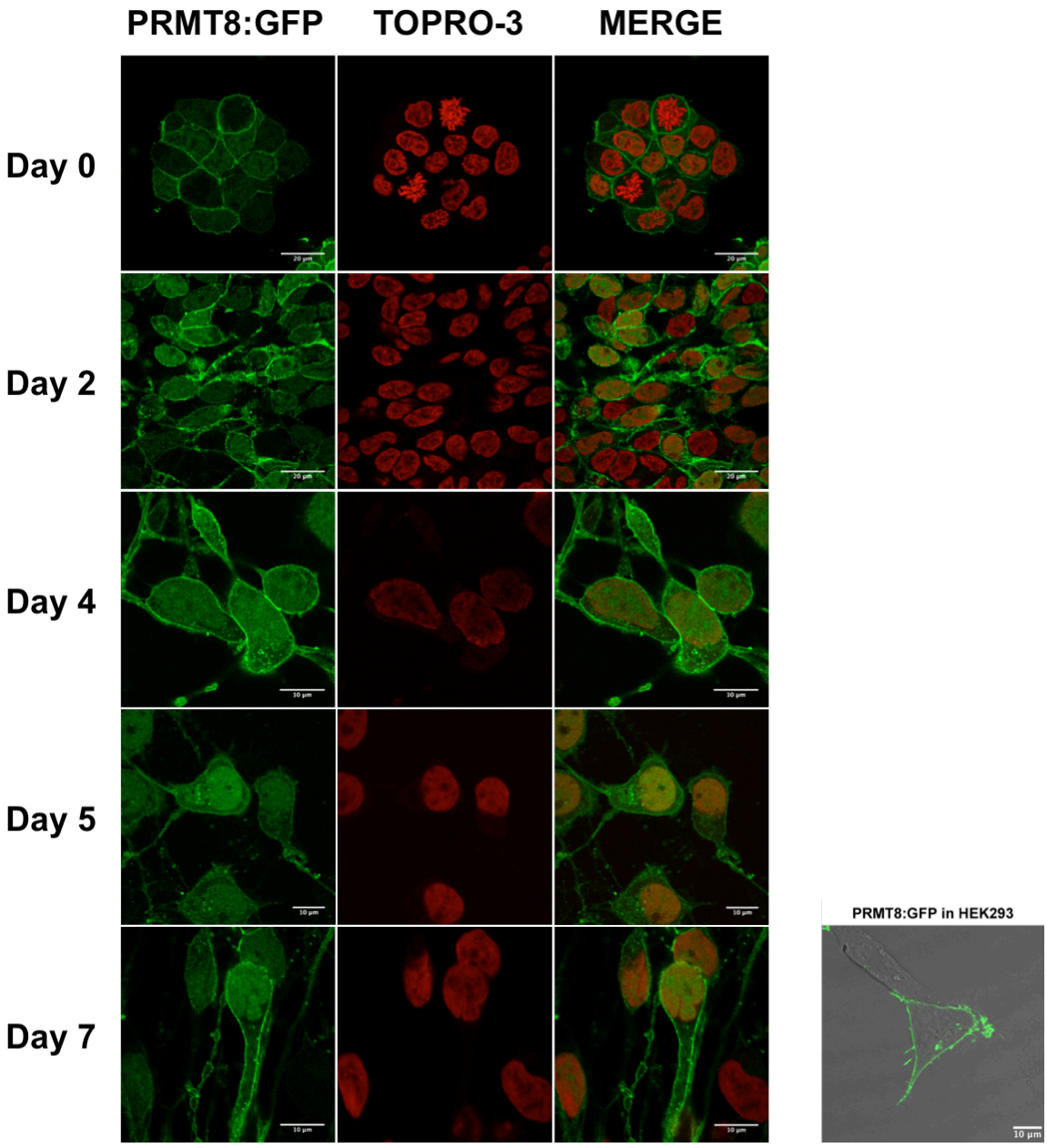


Fig. 3.20: PRMT8 localization during LUHMES differentiation. Undifferentiated LUHMES were infected with PRMT8:GFP LeGO-iPuro2 vector. After selection with Puromycin for one week, they were induced to differentiate. Cells were fixed with paraformaldehyde every day of differentiation and stained with TO-PRO-3 as a DNA staining. The samples were observed under the microscope and confocal images were obtained. PRMT8:GFP (green) is mainly localized in the cellular membrane of undifferentiated LUHMES. While differentiation is in process, more and more protein is localized in the cytoplasm and in the nucleus of the neurons. A fraction of it remains in the plasma membrane even on day 7. For comparison an image of HEK293 cells infected with the same construct is presented. PRMT8 is restricted in the plasma membrane of HEK293 cells.

experiments in cells where PRMT8 is endogenously expressed (i.e. LUHMES). As it was mentioned earlier, there were no available antibodies that recognize the endogenous protein. We therefore proceeded with recombinant GFP-fusion proteins. We hypothesized that the localization of recombinant PRMT8 in LUHMES should be

representative because it is an endogenous protein of the system and potential interaction partners, which could determine its localization, should be present in the system. We chose to work with the original "full length" (the one with the myristoylation anchor) because we had found that this is expressed in LUHMES on the mRNA level, and we do not have indications that the one which is shorter by 15 amino acids is the natural form of the protein. Undifferentiated LUHMES cells were infected with PRMT8:GFP LeGO-iPuro2 vector and selected for one week in 0.1 µg/ml of puromycin. The stable PRMT8:GFP-expressing LUHMES were induced to differentiate by adding tetracycline on glass coverslips to switch off *v-myc* oncogene expression. Images of fixed cells were obtained with a confocal microscope every day of differentiation. TO-PRO-3 was used as an infrared DNA counterstaining. In undifferentiated cells, the protein is mainly localized in the plasma membrane, but a small fraction of it is distributed equally between the cytoplasm and the nucleus. As the differentiation is progressing, the nuclear signal of the protein becomes stronger (Fig. 3.20). The same behavior was observed in live LUHMES cells. This means that – for yet unknown reasons – full-length PRMT8 accumulates in the nucleus. The reasons and the mechanism would be studied in future experiments.

3.3.6 PRMT8 nuclear amounts are regulated

Following the observation that the full-length PRMT8 translocates to the nucleus during differentiation, and in combination with the previous knowledge (Kousaka et al., 2009) that PRMT8 is enriched in the nucleus during differentiation of PC12 cells, we hypothesized that the full-length protein may be first processed and then moves into the nucleus. For example, PRMT8 amino-terminal region could be proteolytically cleaved or folded inside the molecule after specific signals and then dissociate from the membrane. Next, it could receive other signals or be accompanied by other proteins into the nucleus. The phenomenon is specifically observed in LUHMES but not in HEK293 cells. In order to mimic the shifting of PRMT8 from the plasma membrane and its distribution between the cytoplasm and the nucleus and quantify its amounts in each cellular compartment, we created an artificial system. We cloned PRMT8 cDNA in a pEYFP vector by removing a part of its amino-terminus, which encodes the first 15 amino acids (45 nucleotides). The new protein was designated PRMT8M16:EYFP, because the first amino acid is methionine

which corresponds to position 16 of the full-length protein. This creates a construct that corresponds to the shorter form of PRMT8 as described by Kousaka et al. 2009, which lacks the part that is responsible for the membrane anchoring of PRMT8. For this experiment, HEK293 cells were used due to the fact that is a simpler cellular system, and because these cells have a morphology which makes them very suitable for such experiments. HEK293 cells were transfected with PRMT8M16:EYFP and observed under the microscope. Indeed, PRMT8M16:EYFP was localized in the cytoplasm and in the nucleus of the cells. The localization of the protein was completely different to PRMT8:GFP in HEK293 cells, which was restricted to the plasma membrane (Fig. 3.21). In fact, it very much resembled the localization of PRMT1 or PRMT1Δarm (compare Fig. 3.4).

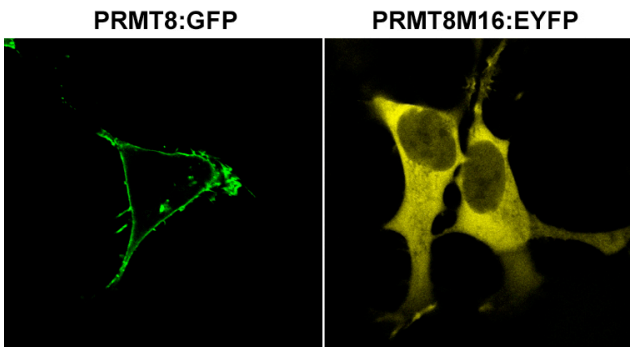


Fig. 3.21: Comparison of the subcellular localization of PRMT8:GFP and PRMT8M16:EYFP. HEK293 cells were transfected with encoding vectors either for PRMT8:GFP or for PRMT8M16:EYFP and confocal images were acquired. As expected, full-length PRMT8 was localized on the plasma membrane. In contrast, the form which lacks the fifteen first amino acids of its amino-terminus (PRMT8M16:EYFP) is released from the membrane and is distributed between the cytoplasm and the nucleus of the cell.

During these experiments, we also observed that, independently of the expression level of PRMT8M16:EYFP in each individual cell, the intensity of fluorescent signal seems to be similar in each of the nuclei. It appeared as if the signal in the nucleus of PRMT8M16:EYFP transfected cells could not exceed a certain intensity. Inspired by that observation, we quantified the relation between the total amount of protein per cell and the amounts in the nucleus or in the cytoplasm. Random fields of stable PRMT8M16:EYFP HEK293 cells were imaged under the same parameters. The images were analyzed with ImageJ software by measuring the mean fluorescent intensity in a region of interest (ROI) of 15x15 pixels in both the cytoplasm and the nucleus of more than one hundred cells. The numbers were statistically analyzed by Microsoft Excel, and are presented graphically in Figure 3.22A. The cells are arranged by increasing total brightness from left to right (green bars). It is obvious that cytoplasmic brightness (blue bars) increases together with the total brightness. On the other hand, nuclear brightness (red bars) cannot exceed 35

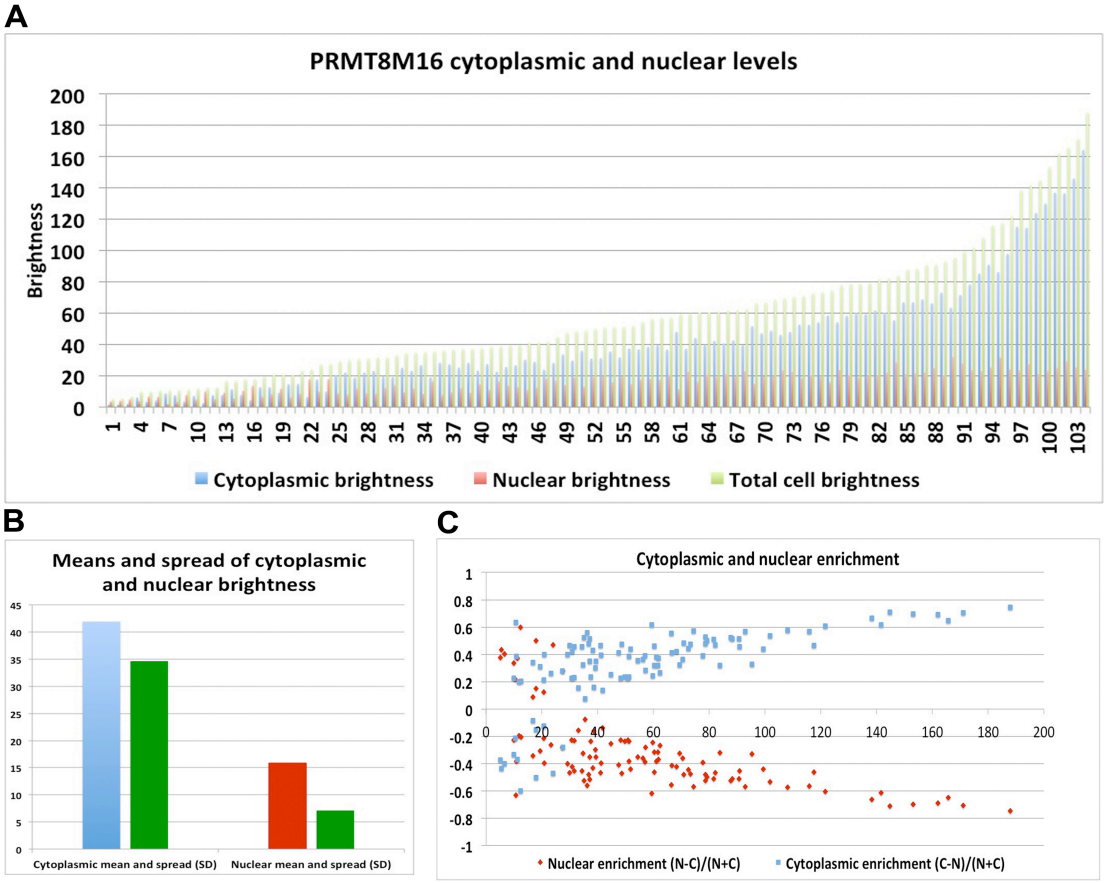


Fig. 3.22: PRMT8M16 distribution between the cytoplasm and the nucleus. Confocal images of random fields of stably transfected HEK293 cells expressing PRMT8M16:EYFP were acquired. The images were analyzed by ImageJ software. A 15x15 pixels ROI was used to measure the mean brightness of a cytoplasmic and a nuclear region of over one hundred cells. (A) The numbers were analyzed by Microsoft Excel and arranged by increasing total (cytoplasmic plus nuclear) brightness (green bars) from left to right. Blue bars represent the cytoplasmic brightness and red bars the nuclear brightness of each individual cell. Cytoplasmic brightness increases along with the total brightness. In contrast, nuclear brightness cannot exceed 35 units, independently of the total cell brightness. (B) The means of the cytoplasmic (light blue) and the nuclear (red) brightness of all cells are presented. The nuclear mean is lower with a small spread (nuclear mean=15,88, SD=7,06). On the other hand, the cytoplasmic mean is higher and the measurements are much more spread (cytoplasmic mean=41,95, SD=34,71). Green bars represent the standard deviation, which in this case is showing the spread of the measured intensities in cytoplasm and nucleus. (C) Cytoplasmic (blue) and nuclear (red) enrichment of PRMT8. When the total cell brightness (cytoplasmic plus nuclear, X axis) increases, PRMT8 is enriched in the cytoplasm. This is shown from the positive numbers that the cytoplasmic enrichment has in the cells with higher total amounts of PRMT8. In those cells the nuclear enrichment is negative.

arbitrary units, independently of the total cell brightness. The means of cytoplasmic and nuclear brightness of all cells were calculated, showing that the cytoplasmic mean (light blue bar) was higher and with a very high spread (cytoplasmic mean=41,95, SD=34,71), in stark contrast to the nuclear brightness (red bar), where the mean is lower with a much lower spread (nuclear mean=15,88, SD=7,06) (Fig. 3.22B). Thus, we could state that nuclear amounts of PRMT8 range around the same levels and the cytoplasmic amounts

differ according to the total amount of protein per cell. The cytoplasmic (C) and the nuclear (N) enrichment of every individual cell are presented. The formulas, which were used for those calculations are the difference of the cytoplasmic from the nuclear brightness of each cell and vice versa divided by the total cytoplasmic and nuclear brightness ((C-N)/(C+N) and (N-C)/(C+N)). For the first case, positive numbers indicate cytoplasmic enrichment and negative numbers nuclear enrichment of PRMT8. The same logic is applied in the second case where, positive numbers indicate nuclear enrichment and negative numbers cytoplasmic enrichment of PRMT8. It is clear that by increasing the total amount of protein per cell there is cytoplasmic enrichment of the protein (Fig. 3.22C).

3.3.7 Knocking-down PRMT8

The up-regulation of PRMT8 during the neuronal differentiation of LUHMES cells indicates that it may play a role in this process and/or in the maintenance of the mature neurons, resulting from LUHMES differentiation. An interesting issue to be answered in the future is the effect of PRMT8 knock-down in these cells. For that purpose, four different short hairpin sequences targeting PRMT8 mRNA:

- KD1: 5'-AATGTGCGAGACCTCGATTTCACAGTAGA-3',
- KD2: 5'-AAGCAAGTGGTGACCAATGCCTGTTTGAT-3',
- KD3: 5'-GTGTGAAACATCTGTATCTAATGACTACA-3',
- KD4: 5'-GTGTTCAAGGACAAAGTGGTACTGGATGT-3'

plus a scramble control sequence was cloned in LeGO-G/Puro vector. The four constructs (KD1 to KD4) and the scramble control (SCR) were qualified for their ability to reduce PRMT8 on the mRNA level by RT-PCR. LUHMES were infected with lentiviruses carrying those vectors and stable cells lines were created after a week of selection in 0.1 µg/ml puromycin. Successful infection was verified by microscopic observation of the cells, based on the green fluorescent protein encoded by the LeGO-G/Puro vector. All five-cell lines were induced to differentiate for 7 days. Total RNA was purified and reverse transcribed to cDNA. cDNAs were used as a template for a PCR reaction with PRMT8 and GAPDH specific primers. The PCR products were analyzed by agarose gel electrophoresis and visualized by ethidium bromide. The best knock-down effect was observed from KD2

construct, followed by KD1 and KD3. KD4 cells were expressing higher amounts than the scramble control, and will thus be omitted from the next experiments (Fig. 3.23). In future experiments, when specific antibodies for PRMT8 become available, it will also be necessary to verify the knockdown efficiency on protein level. Differentiated cells expressing shRNAs did not present any morphological changes compared to the scramble control.

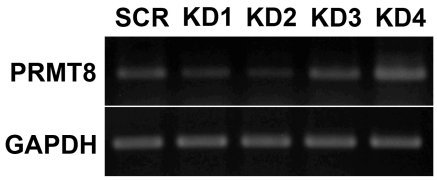


Fig. 3.23: Knock-down effect of the newly designed constructs on PRMT8 mRNA. LUHMES stable cell lines were created with vectors encoding four different shRNAs corresponding to PRMT8 mRNA (KD1, KD2, KD3, KD4), or a scrambled (SCR) control. The cells were differentiated for seven days before PRMT8 mRNA levels were measured by semi-quantitative PCR. KD2 had the best knock-down effect on PRMT8, followed by KD1 and KD3. KD4 showed, for unclear reasons, increased PRMT8 levels in comparison to the scrambled control.

3.3.8 Tools for the identification of interactions partners and/or substrates of PRMT8

Other important experiments, which will help us elucidate the cellular functions of PRMT8 will be the identification of its interaction partners or the substrates that methylates. Conventional methods such as co-immunoprecipitation (co-IP) or yeast two-hybrid system (Y2H) are not very suitable in the case of PRMT8. First, co-IP can currently not be applied due to the lack of a reliable antibody against endogenous PRMT8. Two-hybrid systems, on the other hand, must be performed in yeast or other specialized cell systems, which rules out identifying partners in the natural environment of the PRMT8 protein, cells from the central nervous system. A promising new alternative method appears to be the BioID system based on BirA. BioID can provide a list of proteins, which are in very close proximity of the protein of interest. It is based on a mutated form of the bacterial biotin ligase BirA (Bifunctional ligase/repressor BirA), designated BirA*, as described in the methods.

In our pilot experiments, we designed and validated the tools for BioID of PRMT8, to be used in LUHMES cells. To this end, BirA*HA and PRMT8:BirA*HA were cloned in the LeGO-iPuro2 vector. Next, undifferentiated LUHMES were infected with lentiviruses encoding the two constructs. The cells were selected in 0.1 µg/ml puromycin for one week in order to obtain stable cell lines. After selection, the two cell lines were tested for the expression and the correct localization of the recombinant proteins. Undifferentiated

and differentiated cells from each line were lyzed and the protein extracts were analyzed by SDS-PAGE and Western blot with the a-HA antibody. As expected, we found BirA*HA and PRMT8:BirA*HA both in undifferentiated (D0) and differentiated (D7) cells (Fig. 3.24B).

The correct localization of the two proteins is essential for this experiment because we take advantage of the labeling of neighboring proteins in order to find interaction partners or substrates of PRMT8. Cells from the two different cell lines were seeded on coverslips, fixed and stained with the anti-HA primary antibody and an anti-rat secondary antibody conjugated with Alexa568. DRAQ5 was used as a DNA counterstaining. Samples were observed under the microscope and confocal images were acquired. We observed that BirA*HA was distributed homogeneously in the cell. On the other hand, PRMT8:BirA*HA was localized mainly on the plasma membrane and a small fraction of it was detected in the cytoplasm and in the nucleus. Thus, the distribution of PRMT8:BirA*HA was exactly like that of PRMT8:GFP (Fig. 3.24A), verifying that the protein encoded by the construct is properly localized.

In the next step, undifferentiated cells of both cell lines were supplied with 50 μM of biotin (Sigma) for 24 hours. Afterwards, the cells were lyzed and the protein extracts were analyzed by SDS-PAGE and Western blot. Equal amounts of protein were loaded on a separate gel and later stained with Coomassie blue. The blot membrane was incubated with streptavidin conjugated to horseradish peroxidase (HRP) and exposed to an autoradiography film after incubation with ECL. We observed that our control, free BirA*, promiscuously biotinylates LUHMES proteins. On the other hand, PRMT8:BirA* presents a much more restricted biotinylation profile (Fig. 3.24C). As a control, the same experiment was repeated by omitting the biotin from the culture medium and no signal was detected on the film (data not shown). These results suggest that the newly obtained constructs, LeGO-iPuro2-Bira*HA and LeGO-iPuro-PRMT8:BirA*HA are expressed and localized properly in LUHMES cells, and that they are enzymatically active to biotinylate proximal proteins. In future experiments, biotinylated proteins will be precipitated by streptavidin magnetic beads and analyzed by mass spectrometry in order to obtain a list of proteins in proximity of PRMT8 and study the most interesting of them.

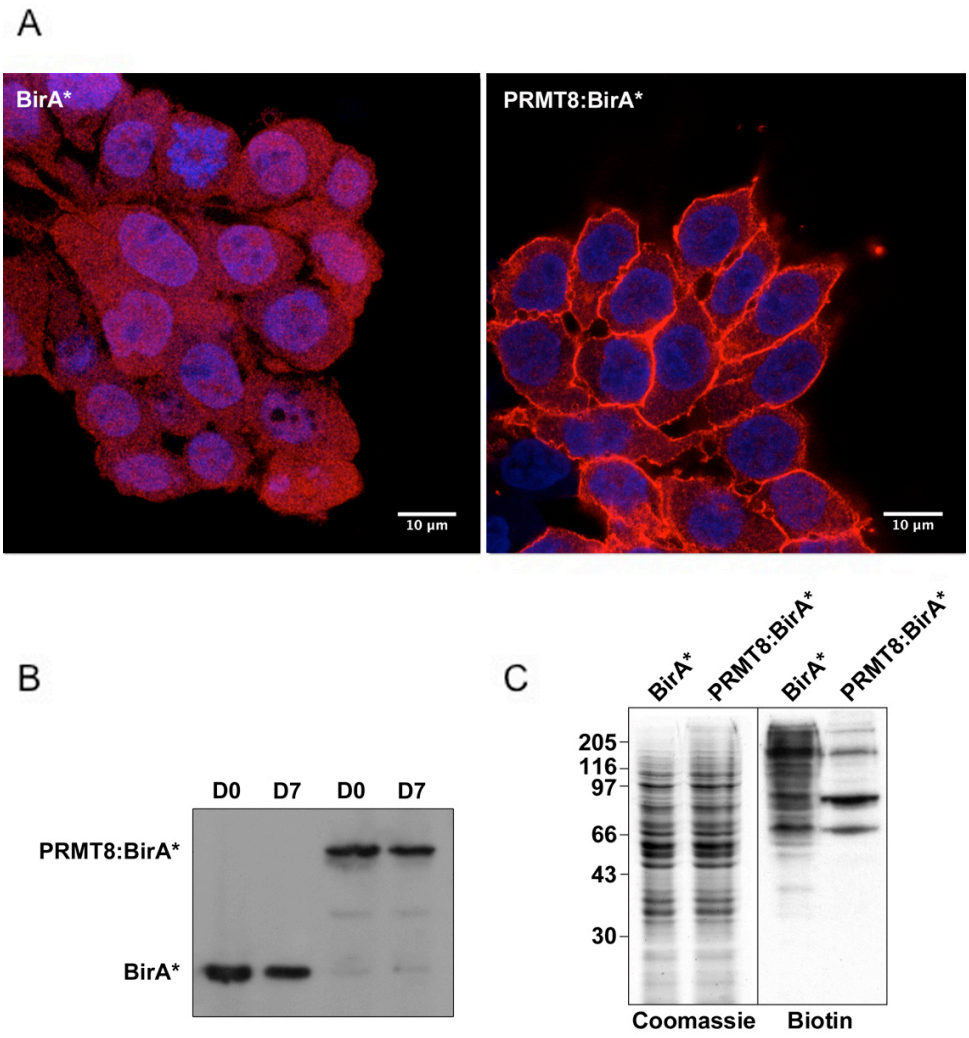


Fig. 3.24: Application of BioID method for recognition of proteins in proximity with PRMT8. (A) LUHMES cells were infected with lentiviruses carrying BirA*HA and PRMT8:BirA*HA in LeGO-iPuro2 vector and stable cell lines were created after one week of selection with puromycin. Undifferentiated cells grown on coverslips were stained with a primary anti-HA antibody and a secondary anti-rat conjugated with Alexa568 (red). DRAQ5 was used as a DNA staining (blue). The localization of the two constructs was observed under the microscope and the confocal images are presented here. (B) The same cell lines were differentiated for seven days and protein extracts from undifferentiated (D0) and differentiated (D7) cells were isolated. The extracts were analyzed by Western blotting with anti-HA primary antibody and an anti-rat secondary antibody, verifying the expression of the two constructs in both cellular states and in the correct molecular mass size. (C) Undifferentiated cells of both BirA*HA and PRMT8:BirA*HA cell lines were supplied with 50 µM of biotin and total cell extracts were obtained. The extracts were separated by SDS-PAGE, transferred to a membrane, and incubated with streptavidin conjugated with HRP. After incubation of the membrane with ECL and exposition to an autoradiography film the biotinylated proteins appeared in each one of the two cell lines.

4. Discussion

4.1 PRMT1: Identification and characterization of a novel splice variant and its impact on cancer

PRMT1 is the predominant member of the human family of protein arginine methyltransferases, and has previously been described to occur in a variety of splicing variants that differ in their amino terminus (Goulet et al., 2007). Here, we characterized a novel splicing variant that lacks the dimerization arm of the protein, and is catalytically inactive but able to bind to natural substrates. The novel variant, which we call PRMT1Δarm, correlates with the malignant phenotype in lung cancer lines, and is increased by expression of the transcription factor Snail.

Several previous studies had provided evidence for a role of protein arginine methylation and members of the PRMT family in different types of cancer (reviewed by Yang & Bedford 2013). PRMT1 and PRMT5 are enzymes that asymmetrically or symmetrically methylate identical arginine residues, respectively, for example arginine 3 of Histone H4. Symmetric R3 methylation of H4 by PRMT5 results in silencing of the underlying gene (Zhao et al., 2009), whereas asymmetric arginine methylation of H4 by PRMT1 results in transcriptional activation (Wang et al., 2001) and is associated with increased tumor grade in prostate cancer (Seligson et al., 2005). Both enzymes have recently been correlated with tumor aggressiveness, and their knock-down in both *in vitro* and *in vivo* studies had shown that it results in cell cycle arrest and reduced tumor volume (Baldwin et al. 2012; Han et al. 2014; Wang et al. 2012). This is also in line with a recent clinical study, which demonstrated that high PRMT1 mRNA expression in breast cancer tissues correlates with poor patient prognosis and a reduced disease-free survival (Mathioudaki et al., 2011). However, the contribution of PRMT1 to malignancy is still not well understood on a molecular level, nor do we know enough about how the activity of PRMT1 is regulated on specific substrates. This is of particular interest because PRMT1 is a highly abundant enzyme that fully methylates most of its substrates (see, e.g. Herrmann et al. 2004), but at the same time only a small fraction of histone tails are modified by the enzyme. One way of regulation is the interaction with other proteins such as B-cell translocation gene 1 (BTG1) and BTG2, which promote its multimerization and result in increased PRMT1 activity and co-activator function (Lin et al., 1996). Earlier results

suggested that the activity of PRMT1 can also be regulated by the balance of expressed splicing variants (Goulet et al., 2007), which are all able to form oligomers with each other, resulting in oligomeric holoenzymes with different activity and distinct substrate specificities in different cell types. Along the same line, our group had previously demonstrated that catalytically inactive PRMT1v2 stably binds to substrates; it might thus contribute to the regulation of PRMT1 activity by shielding selected substrates from being methylated by active monomers in the same oligomeric PRMT1 complex, thereby reducing its specific activity (Herrmann et al., 2005).

In contrast to all other known splicing variants of PRMT1 that differ in their amino terminus, the novel isoform described in this thesis lacks the dimerization arm of the enzyme, and thus represents the first example of a PRMT1 isoform with variation in the catalytic core of the protein. PRMT1Δarm cannot multimerize with other variants of endogenous PRMT1 (Fig. 3.8A) and is catalytically dead (Fig. 3.8B), but is still able to stably associate with substrates in the nucleus as measured by photobleaching analysis (Fig. 3.9). Surprisingly, we found here that full length PRMT1v1 also possesses an immobile fraction in the nucleus (Fig. 3.9), in contrast to PRMT1v2, which our group had investigated earlier (Herrmann et al., 2005). This is an interesting finding, because PRMT1v2 is the only splicing variant that contains a nuclear export signal (Goulet et al., 2007), and thus should be excluded from the nucleus. However, our group has shown that the localization of PRMT1v2 is highly dependent on the cell type, ranging from almost exclusively cytoplasmic to almost exclusively nuclear (Herrmann et al., 2009), most probably due to the localization of the majority of its substrates in the respective cell line. Also, its nucleo-cytoplasmic shuttling requires enzymatic activity, with inactive enzyme preferentially displaying nuclear localization. On the other hand, PRMT1v1 is present in the nucleus in much higher amount, as it lacks an export signal. We hypothesize that PRMT1v1 may stably associate with substrates (or regulators) in the nucleus, resulting in an immobilized fraction of the enzyme. We will investigate this point in further experiments, but for the present study it is mainly important to note that the immobile fractions of PRMT1Δarm and PRMT1v1 are identical, demonstrating that binding of PRMT1Δarm to the factors that also immobilize PRMT1v1 does not require multimerization, in contrast to enzymatic activity, which does. This suggests that

PRMT1Δarm associates with endogenous PRMT1 substrates without being a component of an active oligomeric holoenzyme itself, and can act as a competitive inhibitor. In fact, we find that experimental overexpression of PRMT1Δarm leads to a block in the G1 phase of the cell cycle, and to morphological changes indicative of apoptosis as well as an increase of cells with sub-G1 content of DNA (fig. 3.6). This is in line with observations from Wang and colleagues, who have demonstrated that knock-down of PRMT1 results in an arrest in the G1-S phase of the cell cycle, proliferation inhibition and apoptosis induction in glioma cell lines (Wang et al., 2012).

We find that the protein level in transient expression of the isoform is typically around one fourth of active PRMT1v1 (Fig. 3.5 & 3.6), and even less in stably expressing cells after several weeks of selection. Apparently, the cell counteracts the forced expression of elevated amounts of PRMT1Δarm, even though it is expressed from the same CMV promoter on the expression construct as full length PRMT1v1, suggesting that high levels of PRMT1Δarm may not be compatible with cellular growth. This is reminiscent of earlier results from our laboratory which had demonstrated that a point-mutated, inactive version of PRMT1 can likewise only be expressed at low levels (Herrmann and Fackelmayer, 2009). We thus investigated whether PRMT1Δarm (and possibly other inactive versions of the enzyme) is degraded by the proteasome pathway, in an attempt of the cell to lower the toxicity of PRMT1Δarm. Upon first inspection, the amount of the isoform does not change detectably after an 8 or 16 hour incubation with the proteasome inhibitor MG132. Unexpectedly, though, we microscopically found that only PRMT1v1Δarm, but not full length PRMT1v1, accumulates in nucleoli, next to but not exactly co-localizing with Fibrillarin (Fig. 3.7B). This re-localization is seen only after MG132 treatment, while in untreated cells the localization of PRMT1Δarm and full length PRMT1v1 are indistinguishable. The structures to which PRMT1Δarm re-localizes are reminiscent of “nucleolar aggresomes”, which have been found before in many cell types after treatment with proteasome inhibitors, and appear to contain dozens of proteins that accumulate there during proteotoxic stress, together with polyadenylated RNA (Latonen et al., 2011). Among the proteins found in these structures are components of the ubiquitin-proteasome system, including PML body proteins and ubiquitin itself, as well as targets including cancer-related transcription factors and cell cycle regulators (e.g.

p53, MDM2, cyclins) and proteins involved in neurodegenerative diseases (e.g. ataxin-1, Malin). Interestingly, we find that PRMT1Δarm is almost entirely cleared from the cytoplasm when it re-locates to nucleolar aggresomes. This is not a consequence of cytoplasmic PRMT1Δarm being degraded, because the total amount of the protein remains the same during MG132 treatment (Fig. 3.7A). Most probably, the nucleolar aggresome proteins are in the process of being degraded, and their normal half-lives are bound to have an effect on the kinetics of their appearance in nucleolar aggresomes (Latonen et al., 2011). We therefore hypothesize that the half-life of PRMT1Δarm is somewhat decreased in comparison to full length PRMT1v1, and thus the lower steady-state amount of the PRMT1Δarm isoform most likely results from faster degradation. The re-localization of PRMT1Δarm to nucleolar aggresomes is also remarkable from a different point of view. As the protein is predominantly cytoplasmic in untreated cells, one would expect it to localize to "classical" cytoplasmic aggresomes (Johnston et al., 1998; Rodriguez-Gonzalez et al., 2008) rather than nucleolar aggresomes, which are usually only used by nuclear proteins (Latonen et al., 2011). Taken together with the results from photobleaching, these indicate that the main place of action of PRMT1 may be the nucleus, irrespective of the predominant localization in the cytoplasm. This is also in line with re-localization of inhibited or genetically inactivated PRMT1v2 to the nucleus, where it stably but reversibly binds to substrates in chromatin and the nuclear scaffold, and the cell-type specificity that results in predominantly nuclear PRMT1v2 in some cells, such as MCF-7 breast cancer cells (Herrmann et al., 2009). Thus, PRMT1 activity appears to be at least partially regulated by its cytoplasmic-to-nuclear translocation.

We had originally found this new variant in database searches as an expressed sequence tag from a neuroblastoma (GenBank accession number BX352789.2); in follow-up experiments, we detected its expression in variable amounts in normal fibroblasts and different cancer cell lines (Fig. 3.10A). When investigated by semi-quantitative and quantitative RT-PCR, we found that the amount of PRMT1Δarm is significantly higher in non-small cell lung cancer cell lines than in normal diploid lung fibroblasts, and correlates with their reported aggressiveness (Fig. 3.10): it was higher in H1792 that originate from a stage 4 lung adenocarcinoma and expressing mutant K-Ras and mutant p53, and lower in H1437 from a stage 1 lung adenocarcinoma, expressing wild type K-Ras and a dominant

negative p53. The two other cancer cell lines, expressing mutant K-Ras and wild type p53 (A549) or wild type K-Ras and mutant p53 (H1299), displayed intermediate but almost identical levels of up-regulation.

In these experiments, we also observed a correlation with the mesenchymal phenotype, with higher expression of PRMT1Δarm in cell lines with a higher percentage of mesenchymal cancer cells. As HEK293 cells can be driven towards mesenchymal cells by expressing Snail (Kashyap et al., 2013), one of the master EMT-inducing transcription factors (see, e.g. Zhou et al. 2004), we generated a HEK293 cell line that stably expresses Snail, and indeed find that they have a significantly elevated level of PRMT1Δarm in comparison to the parental cells (Fig. 3.11). Thus, PRMT1Δarm is clearly up-regulated in cells undergoing EMT. Together, these results point to a link between PRMT1Δarm expression and cancer cell phenotype in lung cells. The specific way in which PRMT1Δarm is linked to cancer remains currently unknown. PRMT1Δarm could promote the transformation of normal to cancer cells in a similar way to splice variant 2 (Baldwin et al., 2012), and thus being expressed in higher levels in cancer cells. On the other hand, and maybe more likely in light of our results, PRMT1Δarm might not promote cancer, but rather be acting as a defence mechanism against neoplastic transformation. In this scenario, PRMT1Δarm could protect the cells from cancer progression by negatively regulating PRMT1-mediated methylation, which has been linked to cancer promotion in several studies. This could explain our observation that cells forced to express higher amounts of PRMT1Δarm have increased rate of apoptotic cell death, as they initiate apoptosis in order to avoid their transition to cancer cells through EMT. In other, words, PRMT1Δarm may be a cancer barrier. This is particularly plausible, as PRMT1 has recently been found to act as an important regulator of EMT through PRMT1-mediated methylation and activation of the E-cadherin repressor, Twist1, in non-small cell lung cancer (Avasarala et al., 2015) as well as in breast cancer (Gao et al., 2016). Through promoting migration and invasion, overexpression of PRMT1 may trigger cancer cells to metastasize (Gao et al., 2016). Overall PRMT1 was also found to be increased (2.9 fold) after induction of EMT but in a smaller factor than the Δarm (6.5 fold). The elevated levels of PRMT1 in the Snail transfected cells might be an outcome of the increase of particular variants and not of all seven. We cannot exclude that some of the seven variants might be

down-regulated. The increase of PRMT1Δarm in cells that underwent EMT may therefore represent an attempt to reduce PRMT1 activity, and in this way counteract PRMT1-induced activation of the EMT program in affected cells.

The expression of Δarm was detected in every cell line analyzed in the present thesis. It was present in normal lung fibroblasts and lung cancer cell lines as well as in HEK293 and HeLa cells. Moreover, it was present in LUHMES cells and its expression remained constant during differentiation (Fig. 3.2, 3.10 & 3.12). This points to a functional role of this variant in the cell, and also excludes the formal possibility that it is only detected in cancer cells because of general splicing defects in transformed cells. The observation that its expression levels remain stable during differentiation might indicate that Δarm is not specifically associated with differentiation of LUHMES cells. However, it might have a role in the survival of the cells, since it is detected in every stage of LUHMES differentiation as well as in every cell line we have analyzed.

A parallel experiment revealed that, although Δarm expression as well as the other six amino-terminal variants of PRMT1 were found to be unchanged during LUHMES differentiation, variant 7 was not detectable in undifferentiated LUHMES, but was up-regulated in differentiated LUHMES (Fig. 3.12). This indicates that variant 7 might have a role in differentiated cells or in cells that are in the process of differentiation. These results are very interesting because variants 4-7 were thought to be tissue specific (Goulet et al., 2007). According to Goulet, variant 4 was detectable only in the heart, whereas variant 5 was mostly expressed in the pancreas, and variant 7 was predominantly present in the heart and skeletal muscles. In contrast, we observed strong expression of variants 4, 5 and 6 in both undifferentiated and differentiated LUHMES which origin from mesencephalon. Moreover, variant 7 is detected in the differentiated cells which means that its expression is not restricted in the heart and skeletal muscles.

Future research in our laboratory will now extend these results, aiming at better defining the properties of PRMT1Δarm and its functional role in regulating PRMT1 activity, particularly in relation to cancer. An obstacle, which has to be overcome, is the production of suitable antibodies against Δarm. The epitope, which can be used for antibodies creation, is restricted because the biggest part of Δarm sequence is homologous to the rest PRMT1 variants. In principle, the only possibility of an appropriate

epitope would be the amino acid sequence at the point where exon 7 and 10 are fused. In addition, shRNAs with sequences derived from the same region of the molecule could be used for knock-down studies. By better understanding the function of Δarm function we could also investigate the possibility of its use as a diagnostic marker or a drug target in different types of cancer.

In a side project of this thesis, I investigated Lamin B receptor as a novel substrate of PRMT1, in collaboration with the laboratory of Prof. Anastasia Politou. Prof. Politou's laboratory is interested in understanding the physiological role of intrinsically disordered proteins or protein domains, of which LBR is an example. LBR is an important factor of the nuclear periphery, and is involved in gene regulation. The nucleoplasm-facing amino-terminal part of the protein has been shown to mediate most of the protein's interactions. It harbors a 60-residue Tudor domain followed by a 40-residue region rich in charged amino acids (Arginines and Serines repeats, RS region), and a third 110-amino acid segment (SG domain) with no apparent homologies. Sequence comparisons with PRMT consensus sequences suggested that the RS region might be a novel substrate of PRMT1, and that its methylation may be one factor in regulating the protein's functional interactions, adding LBR to one of the non-histone proteins that are post-translationally modified. As shown in Fig 3.13A, we indeed find a strong methylation of the RS domain of LBR by PRMT1, but not by PRMT4 or PRMT6 which were used as controls. Moreover, as Fig. 3.13C & D show, methylation of the RS domain is independent of the presence or absence of Tudor domain, but is negatively affected by the presence of the SG domain. Future work will focus on the functional consequences of the methylation of LBR.

4.2 PRMT8: Role in neuronal differentiation using a human cellular model *in vitro*

PRMT8 was discovered due to its high sequence homology with PRMT1. Initially, PRMT8 was believed to be restricted to the central nervous system (Kousaka et al., 2009; J. Lee et al., 2005; Taneda et al., 2007) and associated with the plasma membrane through the myristoylation anchor on glycine 2 (J. Lee et al., 2005; Taneda et al., 2007). Recent studies have demonstrated the expression of PRMT8 in human and mouse embryonic stem cells, in induced regenerative competent cells (iRC), in zebrafish embryo and in different types of cancer (Hernandez and Dominko, 2016; Hernandez et al., 2017;

Lin et al., 2013; Solari et al., 2016). Work from the group of Fukamizu suggests that PRMT8 has dual enzymatic activity, the known methyltransferase plus phospholipase activity (Kim et al., 2015). The expression of a second variant, which does not contain the myristoylation motif, was recently discovered in human fibroblasts and in iRC cells by the laboratory of Dominko (Hernandez and Dominko, 2016).

In the present thesis, we set up an appropriate human cellular model system for studying the role of PRMT8 in neuronal differentiation and maintenance. An expression profile analysis as well as the localization of the protein during differentiation are also presented. I also generated and validated vectors for PRMT8 knock-down and for the identification of its interaction partners. Motivated by our work on Δarm, we also looked for the existence of hypothetical alternative splicing variants.

For our experiments on PRMT8, we worked with the mesencephalic-derived cell line, LUHMES. This cell line has several advantages compared to other common cell lines being used in neurobiology, as described in the introduction. In the present thesis, this elegant model system of neuronal differentiation was set up and characterized in our lab. We propagated and differentiated the cells according to the described protocols and analyzed them for the expression of typical markers by RT-PCR and immunofluoresence. In parallel, we monitored the morphological changes of the cells during differentiation. Markers such as SOX2, PAX3 and Geminin were down-regulated. Although differentiated cells are post-mitotic, we detected traces of Geminin by RT-PCR (Fig. 3.15). This result was explained later through immunofluorescence experiments, where we detected small colonies of cells, among the network of differentiated neurons, which had managed to “escape” the differentiation process and continue proliferating as well as being positive for Geminin (Fig. 3.16). The number of the colonies was very small in comparison to the whole population of cells and might occur from few cells in which the tetracycline-controlled transactivator does not work properly. Other markers such as Neuroligin, Synaptophysin, DRD2, tyrosine hydroxylase and Neurofilament 165 were up-regulated (Fig. 3.15). Unexpectedly, Neurofilament 165 was detected in undifferentiated cells by RT-PCR but not by immunofluorescence, maybe due to post-translational modifications that render the protein invisible to the monoclonal antibody 2H3, and that differ between differentiated and non-differentiated cells.

The behavior of LUHMES cells in our hands (morphology and marker expression) is in line with the description by Dr. Diana Scholz (Scholz et al., 2011) where a detailed characterization of the system has been carried out. After verifying that LUHMES cells behave in the appropriate way, we measured the expression levels of PRMT8 in cells during differentiation compared to undifferentiated cells. In line with earlier publications (Kousaka et al., 2009; Lee et al., 2005; Taneda et al., 2007), PRMT8 expression increased during differentiation. In LUHMES cells, this up-regulation appears to happen in a highly synchronous way. PRMT8 expression reaches a peak on day 7, at a time point when the cells are fully differentiated, and then slightly decreases and remains stable until day 14 (Fig. 3.17B). As PRMT8 can readily be detected in brain, which mostly contains post-mitotic but not developing neurons, we speculate that the decrease of PRMT8 RNA in later days of LUHMES culture is a result of cell culture conditions that may not support maintenance of mature neurons for extended periods of time.

In line with the semi-quantitative analysis, quantification by qPCR revealed that PRMT8 is up-regulated by 257 fold on day 6 of differentiation. The relative up-regulation is very high compared to the changes of the levels of other members of the PRMT family; PRMT1 is down-regulated to 75% and PRMT6 is up-regulated by only 2.26 fold compared to undifferentiated cells (Fig. 3.17A). Individual PRMT1 variants remain also more or less stable except for variant 7, which is up-regulated in differentiated cells and might have a regulatory role in differentiation or in post-mitotic neurons (Fig. 3.12). PRMT1 has recently been shown to remain stable or slightly up-regulated during neuronal differentiation of mouse embryonic stem cells upon induction with retinoic acid, which is reminiscent of the results of our differentiation model (Simandi et al., 2015). Moreover it has been shown that all PRMT1 isoforms were detected in very low levels in adult brain in contrast to fetal brain where the expression was stronger (Goulet et al., 2007; Scott et al., 1998). It is thus possible that PRMT8 could replace PRMT1 in adult brain tissue.

In summary, LUHMES is a conditionally immortalized human cell line in which PRMT8 is inducibly expressed at high levels, and in a regulated manner, following the induced neuronal differentiatin of the cells. This makes LUHMES an excellent choice for the study of PRMT8. Moreover, the primers and the PCR conditions for analysing the expression levels of PRMT8, in normal conditions or after different treatments, have been

established and validated in this thesis. We have not detected PRMT8 expression in other cell lines (e.g. HEK293). However, recently it has been stated that PRMT8 is being expressed in some types of cancer (Hernandez et al., 2017). In my diploma work I have detected PRMT8 expression in mouse embryonic stem cells (unpublished data), which is in line with recent findings (Solari et al., 2016).

Unfortunately, we could not test for PRMT8 protein levels because of the lack of a reliable antibody. There are several antibodies on the market, but discussions with leading scientists in PRMT research have the consensus that these antibodies are not reliable, and mainly cross-react with the highly homologous PRMT1. For example, several commercial antibodies are presented in manufacturer's descriptions in western blots with extracts from HEK293 cells, a cell line that does not express the protein according to our and other laboratories' results. Second, many of the antibodies have been shown to recognize recombinant PRMT8, but not the endogenous protein. In our own attempts to create new antibodies against PRMT8, we collaborated with the American Cell Signaling company. Peptides from the amino-terminus of PRMT8, which is the only region of the protein that sufficiently differs from PRMT1, or the full-length protein were used for immunizing rabbits by the company. We received the crude sera and purified antibodies from the company, and we carried out a series of tests in order to check whether they recognize the recombinant and the endogenous protein, but not other members of the family. Indeed, this collaboration succeeded in creation of sera as well as purified antibodies, which strongly recognize PRMT8 in extracts from transfected cells (Fig. 3.18). The reactions were also highly specific, as the sera and antibodies did not detect PRMT1 or the other, less homologous, members of the PRMT family. However, all these antibodies failed to detect endogenous PRMT8 in differentiated LUHMES (day 7) extracts where we had previously detected PRMT8 transcripts by RT-PCR. Kousaka et al. (2009) had used a synthetic peptide from a region further downstream from the amino-terminus (positions 55-71) and created antibodies that recognized the endogenous PRMT8 in mouse brain extracts in Western blot and immunocytochemistry experiments. We have requested a small aliquot of this antibody, and did several tests, but did not manage to detect endogenous PRMT8 in LUHMES, and obtained somehow inconclusive results in mouse brain extracts. This antibody, like those produced by Cell Signaling, was also

working very well with over-expressed protein. To create a similar antibody for future work, we also used the same sequence (positions 55-71) as an epitope for injecting rabbits. Interestingly, our results were the same: the antibody recognized recombinant PRMT8, but did not see the endogenous protein. There are several potential reasons for this result. In contrast to the detection of protein by western blotting, PCR reactions from LUHMES might be positive, because of the high sensitivity of PCR amplification. This possibility would be even more pronounced if the expression level of PRMT8 protein was very low. Maybe PRMT8 is sufficient in very low levels in order to carry out its physiological function. However, as judged from the expression levels of other family members, this would be unusual for a PRMT. Also, the number of cycles required to obtain a positive PCR result suggested that the expression of PRMT8 is not extremely low. A more probable explanation could be the presence of post-translational modifications in the amino-terminus or in other domains of PRMT8, which could hinder antibody binding. Last, as there is still controversy about the exact amino-terminus of endogenous PRMT8 (see below), it is possible that the peptides sequences used to create antibodies are completely or partially missing on the endogenous protein. This could explain why the recombinant protein is strongly recognized by the antibodies (in which case the start codon is determined by cloning), but the endogenous protein escapes detection. Further work will be necessary to decide between these possibilities and draw final conclusions.

The data about full length sequence of endogenous PRMT8 are controversial. First, Dr. Lee published the PRMT8 sequence that is myristoylated in Glycine 2 and anchors to the plasma membrane (Lee et al., 2005). Later it was suggested that translation of the mRNA starts from an alternative start codon, 16 codons downstream from the original ATG (Kousaka et al., 2009) that Lee had published previously (Lee et al., 2005). Thus, the myristoylation anchor of the protein would be omitted, and the protein would not be attached on the plasma membrane. It was recently reported the expression of a novel splicing variant of PRMT8 (Gene Bank KR014345.1; Homo sapiens arginine methyltransferase 8 isoform 4) in human fibroblasts and in induced regenerative competent cells with an alternative exon 1 upstream of the one that was known until then (Hernandez and Dominko, 2016). In this variant, the myristoylation motif is also omitted (Hernandez and Dominko, 2016). In the present thesis I looked for the expression

of two hypothetical isoforms of the transcript (X, Y) and of variant 1 and variant 2 (GeneBank NM_019854.4 and NM_001256536.1 respectively) in differentiated LUHMES cells of day 7. The expression of variant 1 was verified with two alternative forward primers. On the other hand, no other variants or hypothetical variants could be detected. For this reason we decided to use variant 1 of PRMT8 for the following overexpression experiments.

In this thesis, localization experiments of PRMT8 during LUHMES differentiation revealed that the protein is mainly bound to the plasma membrane in undifferentiated cells and a small fraction is equally distributed between the cytoplasm and the nucleus. In differentiated cells, the signal in the nucleus becomes stronger. The reason for the accumulation of PRMT8 in the nucleus is still unknown and should be studied in future experiments. The nuclear accumulation points to a possible role of PRMT8 in epigenetic mechanisms regulated by exogenous signals. This is reminiscent of the accumulation of PRMT8M16 in the nucleus of PC12 cells upon NGF-induced differentiation (Kousaka et al., 2009). However, our results show that the amino-terminal region of PRMT8, which includes the myristoylation anchor cannot block the detachment of the protein from the plasma membrane and the translocation to other cellular compartments. The appropriate environment seems to play an important role in this procedure because when we infected HEK293 cells with exactly the same construct, the protein was strictly localized in the plasma membrane and it was completely excluded from the cytoplasm and the nucleus (Fig. 3.20). Therefore, in parallel with the creation of functional expression vectors for PRMT8:GFP and stable cell lines for the study of the localization of the molecule under different conditions, we found out that PRMT8 localization shifts during differentiation and accumulates in the nucleus in a cell-type depended manner. We believe that the localization of recombinant PRMT8 in LUHMES is representative of the molecule because it is an endogenous protein of the system, and the signals as well as the interaction partners, which determine its localization should be present there. We hypothesize that PRMT8 amino-terminal region could be proteolytically cleaved or folded inside the molecule after specific signals and then dissociate from the membrane. Next, it could receive other signals or be accompanied by other proteins into the nucleus. As our results show, this is specifically observed in LUHMES but not in HEK293 cells where

PRMT8 remains anchored on the plasma membrane.

The artificial system we created to mimic Kousaka's protein – HEK293 cells stably transfected with PRMT8M16 – revealed that the localization of this form of PRMT8 is distributed between the cytoplasm and the nucleus, and is reminiscent of the localization of full length PRMT1 or Δarm (Fig. 3.21 and Fig. 3.4). In this system we measured the levels of the protein in the cytoplasm and in the nucleus. Surprisingly, we found that PRMT8M16 levels could not exceed a particular level in the nucleus, independently of the total amount of the protein in the cell. In other words, the amounts of PRMT8M16 are increasing in the cytoplasm when the total amount of the protein is increasing in the cell. In contrast, the nuclear amount of PRTM8M16 reaches a plateau already after quite low expression levels (Fig. 3.22). This suggests that the concentration of PRMT8 in the nucleus is kept at a low level, either actively through regulated import into the nucleus, or passively through interaction with nuclear factors that bind and recruit PRMT8 to its subtrates. Possibly, this is one of the ways that the activity of the enzyme on its nuclear substrates is modulated.

Little is known about the effect of PRMT8 knock-down in cells. It has been shown that PRMT8 knock-down kills glioblastoma line U87MG (Hernandez and Dominko, 2016). Similar results were observed upon knock-down of PRMT8 variant 2 (isoform 4) in human fibroblasts (Hernandez and Dominko, 2016). Creation of tools for PRMT8 knock-down was considered of high importance by us. For this reason, we created and validated vectors carrying shRNAs targeting sequences derived from the PRMT8 sequence. Three out of four shRNAs knocked-down PRMT8 mRNA in differentiated LUHMES, although with different efficiency (Fig. 3.23). It was impossible to validate the efficiency of those constructs on the protein level because of the lack of the antibody at that stage. The effect of KD1, KD2, and KD3 is possible to be improved by infection of the cells with higher amounts of viruses, because in the particular experiment modest quantity of viruses was used to be able to better compare the knockdown efficiencies. Interestingly, LUHMES cells in which PRMT8 was knocked down did not show any morphological differences during differentiation compared to the scramble control. What remains is the analysis of PRMT8 knocked-down LUHMES cells on the molecular level. For example, the changes in the profile of methylated proteins could be studied in order to get hints about

PRMT8 possible substrates. Moreover, high throughput screening for neuronal markers deregulation could be carried out in order to clarify if PRMT8 is responsible for the regulation of neuronal specific proteins, and identify them.

One of the goals of our group is the identification of substrates and interaction partners of PRMT8. For this purpose, we created the tools in order to apply the novel BioID method in our system (Roux et al., 2012). BioID is a method for identifying proteins in proximity of a protein of interest, which might potentially be interaction partners or substrates in case that the protein of interest is an enzyme. BioID method has several advantages compared to other methods that are currently used for finding proximal proteins and interactions. For example, biotinylation is amenable to selective isolation because it is relatively uncommon in mammalian cells. Moreover, this procedure overcomes the limitations of the other methods like yeast two hybrid where the proteins might lack the appropriate post-translational modifications or the interacting partners are absent (while others, non-physiological exist and might create artifacts). It also has an advantage compared to co-immunoprecipitation, because it is taking place *in vivo* (cellular environment), it does not exclude the insoluble proteins and weak or transient interactions are also detected.

Of course, there are also limitations in this system. One of those is the problem, which we should consider in every experiment involving recombined proteins. BirA* is a protein, which might affect properties of the protein of interest such as the localization or the interaction with oligomers or multiprotein complexes. Another issue is the covalent bonding of biotin molecules mainly on the lysines of the proximal protein. The consequence is that the positive charge of those residues is neutralized and can affect other post-translational modifications. The result is that the protein (fusion or neighboring) behavior might change. The efficacy of the method is based on the availability of lysines in the neighboring proteins, so the amount of the biotinylated proteins should not indicate the strength of the interaction. In the opposite way, the absence of biotinylation does not mean absence of interaction or proximity. As a conclusion, this method provides a list of candidates, which might be interaction partners, but should be studied further to verify or not the interaction with the protein of interest.

In the present thesis we created and validated the tools for applying this method on PRMT8. We created lentiviral vectors encoding PRMT8 fused to a mutated bacterial biotin ligase, BirA*, and a control vector encoding for BirA* alone. We created stable cell lines and verified the expression and the proper localization of the constructs in LUHMES as well as their ability to biotinylate proteins upon supplementation of biotin in the culture medium. In future experiments, biotinylated proteins will be purified from LUHMES extracts by streptavidin magnetic beads and analyzed by mass spectrometry. This will reveal important information about the role of PRMT8 in neuronal differentiation and maintenance, and might help clarifying the role of PRMT8 in neurodegenerative diseases, where PRMT8 might be a novel drug target.

5. References

Antonysamy, S., Bonday, Z., Campbell, R.M., Doyle, B., Druzina, Z., Gheyi, T., Han, B., Jungheim, L.N., Qian, Y., Rauch, C., Russell, M., Sauder, J.M., Wasserman, S.R., Weichert, K., Willard, F.S., Zhang, A., Emtage, S., 2012. Crystal structure of the human PRMT5:MEP50 complex. Proc. Natl. Acad. Sci. 109, 17960–17965. doi:10.1073/pnas.1209814109

Arsenault, M.-Y., Parent, A., Seguela, P., Descarries, L., 1988. Distribution and Morphological Characteristics of Dopamine-Immunoreactive Neurons in the Midbrain of the Squirrel Monkey (*Saimiri sciureus*). J. Comp. Neurol. 267, 489–506. doi:10.1002/cne.902670404

Auclair, Y., Richard, S., 2013. The role of arginine methylation in the DNA damage response. DNA Repair (Amst). doi:10.1016/j.dnarep.2013.04.006

Avasarala, S., Van Scoyk, M., Rathinam, M.K.K., Zerayesus, S., Zhao, X., Zhang, W., Pergande, M.R., Borgia, J.A., DeGregori, J., Port, J.D., Winn, R.A., Bikkavilli, R.K., 2015. PRMT1 Is a novel regulator of epithelial-mesenchymal-transition in non-small cell lung cancer. J. Biol. Chem. 290, 13479–13489. doi:10.1074/jbc.M114.636050

Baldwin, R.M., Morettin, A., Côté, J., 2014. Role of PRMTs in cancer: Could minor isoforms be leaving a mark? World J. Biol. Chem. 5, 115–29. doi:10.4331/wjbc.v5.i2.115

Baldwin, R.M., Morettin, A., Paris, G., Goulet, I., Côté, J., 2012. Alternatively spliced protein arginine methyltransferase 1 isoform PRMT1v2 promotes the survival and invasiveness of breast cancer cells. Cell Cycle 11, 4597–612. doi:10.4161/cc.22871

Bedford, M.T., 2007. Arginine methylation at a glance. J. Cell Sci. 120, 4243–4246. doi:10.1242/jcs.019885

Bedford, M.T., Clarke, S.G., 2009. Protein Arginine Methylation in Mammals: Who, What, and Why. Mol. Cell. doi:10.1016/j.molcel.2008.12.013

Bedford, M.T., Richard, S., 2005. Arginine methylation: An emerging regulator of protein function. Mol. Cell. doi:10.1016/j.molcel.2005.04.003

Blanc, R.S., Richard, S., 2017. Arginine Methylation: The Coming of Age. Mol. Cell 65, 8–24. doi:10.1016/j.molcel.2016.11.003

Casadio, F., Lu, X., Pollock, S.B., LeRoy, G., Garcia, B. a, Muir, T.W., Roeder, R.G., Allis, C.D., 2013. H3R42me2a is a histone modification with positive transcriptional effects. Proc. Natl. Acad. Sci. U. S. A. 110, 14894–9. doi:10.1073/pnas.1312925110

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

Cheng, D., Côté, J., Shaaban, S., Bedford, M.T., 2007. The Arginine Methyltransferase CARM1 Regulates the Coupling of Transcription and mRNA Processing. Mol. Cell 25, 71–83. doi:10.1016/j.molcel.2006.11.019

Cheng, X., Roberts, R.J., 2001. AdoMet-dependent methylation, DNA methyltransferases and base flipping. Nucleic Acids Res. 29, 3784–95.

Cho, J.-H., Lee, M.-K., Yoon, K.W., Lee, J., Cho, S.-G., Choi, E.-J., 2012. Arginine methylation-dependent regulation of ASK1 signaling by PRMT1. Cell Death Differ. 19, 859–870. doi:10.1038/cdd.2011.168

Cook, J.R., Lee, J.-H., Yang, Z.-H., Krause, C.D., Herth, N., Hoffmann, R., Pestka, S., 2006. FBXO11/PRMT9, a new protein arginine methyltransferase, symmetrically

dimethylates arginine residues. Biochem. Biophys. Res. Commun. 342, 472–481. doi:10.1016/j.bbrc.2006.01.167

Dacwag, C.S., Ohkawa, Y., Pal, S., Sif, S., Imbalzano, A.N., 2007. The Protein Arginine Methyltransferase Prmt5 Is Required for Myogenesis because It Facilitates ATP-Dependent Chromatin Remodeling. Mol. Cell. Biol. 27, 384–394. doi:10.1128/MCB.01528-06

Debler, E.W., Jain, K., Warmack, R.A., Feng, Y., Clarke, S.G., Blobel, G., Stavropoulos, P., 2016. A glutamate/aspartate switch controls product specificity in a protein arginine methyltransferase. Proc. Natl. Acad. Sci. 113, 2068–2073. doi:10.1073/pnas.1525783113

El-Andaloussi, N., Valovka, T., Toueille, M., Steinacher, R., Focke, F., Gehrig, P., Covic, M., Hassa, P.O., Schär, P., Hübscher, U., Hottiger, M.O., 2006. Arginine Methylation Regulates DNA Polymerase β. Mol. Cell 22, 51–62. doi:10.1016/j.molcel.2006.02.013

El Messaoudi, S., Fabbrizio, E., Rodriguez, C., Chuchana, P., Fauquier, L., Cheng, D., Theillet, C., Vandel, L., Bedford, M.T., Sardet, C., 2006. Coactivator-associated arginine methyltransferase 1 (CARM1) is a positive regulator of the Cyclin E1 gene. Proc. Natl. Acad. Sci. 103, 13351–13356. doi:10.1073/pnas.0605692103

Fauman, E.B., Blumenthal, R.M., Cheng, X., 1999. Structure and evolution of adomet-dependent methyltransferases, in: S-Adenosylmethionine-Dependent Methyltransferases. World Scientific, pp. 1–38. doi:10.1142/9789812813077_0001

Feng, Y., Maity, R., Whitelegge, J.P., Hadjikyriacou, A., Li, Z., Zurita-Lopez, C., Al-Hadid, Q., Clark, A.T., Bedford, M.T., Masson, J.Y., Clarke, S.G., 2013. Mammalian protein arginine methyltransferase 7 (PRMT7) specifically targets RXR sites in lysine- and arginine-rich regions. J. Biol. Chem. 288, 37010–37025. doi:10.1074/jbc.M113.525345

Fountaine, T.M., Venda, L.L., Warrick, N., Christian, H.C., Brundin, P., Channon, K.M., Wade-Martins, R., 2008. The effect of α-synuclein knockdown on MPP+ toxicity in models of human neurons. Eur. J. Neurosci. 28, 2459–2473. doi:10.1111/j.1460-9568.2008.06527.x

Frankel, A., Clarke, S., 2000. PRMT3 Is a Distinct Member of the Protein Arginine *N* -Methyltransferase Family. J. Biol. Chem. 275, 32974–32982. doi:10.1074/jbc.M006445200

Frankel, A., Yadav, N., Lee, J., Branscombe, T.L., Clarke, S., Bedford, M.T., 2002. The Novel Human Protein Arginine *N* -Methyltransferase PRMT6 Is a Nuclear Enzyme Displaying Unique Substrate Specificity. J. Biol. Chem. 277, 3537–3543. doi:10.1074/jbc.M108786200

Ganesh, L., Yoshimoto, T., Moorthy, N.C., Akahata, W., Boehm, M., Nabel, E.G., Nabel, G.J., 2006. Protein Methyltransferase 2 Inhibits NF-κB Function and Promotes Apoptosis. Mol. Cell. Biol. 26, 3864–3874. doi:10.1128/MCB.26.10.3864-3874.2006

Gao, Y., Zhao, Y., Zhang, J., Lu, Y., Liu, X., Geng, P., Huang, B., Zhang, Y., Lu, J., 2016. The dual function of PRMT1 in modulating epithelial-mesenchymal transition and cellular senescence in breast cancer cells through regulation of ZEB1. Sci. Rep. 6, 19874. doi:10.1038/srep19874

Gary, J.D., Lin, W.J., Yang, M.C., Herschman, H.R., Clarke, S., 1996. The predominant protein-arginine methyltransferase from Saccharomyces cerevisiae. J. Biol. Chem. 271, 12585–94.

Gayatri, S., Bedford, M.T., 2014. Readers of histone methylarginine marks. Biochim. Biophys. Acta - Gene Regul. Mech. doi:10.1016/j.bbagrm.2014.02.015

Goulet, I., Gauvin, G., Boisvenue, S., Côté, J., 2007. Alternative splicing yields protein arginine methyltransferase 1 isoforms with distinct activity, substrate specificity, and subcellular localization. J. Biol. Chem. 282, 33009–33021. doi:10.1074/jbc.M704349200

Gros, L., Delaporte, C., Frey, S., Decesse, J., de Saint-Vincent, B.R., Cavarec, L., Dubart, A., Gudkov, A. V, Jacquemin-Sablon, A., 2003. Identification of new drug sensitivity genes using genetic suppressor elements: protein arginine N-methyltransferase mediates cell sensitivity to DNA-damaging agents. Cancer Res. 63, 164–71.

Gros, L., Renodon-Corniere, A., de Saint Vincent, B.R., Feder, M., Bujnicki, J.M., Jacquemin-Sablon, A., 2006. Characterization of prmt7alpha and beta isozymes from Chinese hamster cells sensitive and resistant to topoisomerase II inhibitors. Biochim. Biophys. Acta - Gen. Subj. 1760, 1646–1656. doi:10.1016/j.bbagen.2006.08.026

Gui, S., Gathiaka, S., Li, J., Qu, J., Acevedo, O., Hevel, J.M., 2014. A Remodeled Protein Arginine Methyltransferase 1 (PRMT1) Generates Symmetric Dimethylarginine. J. Biol. Chem. 289, 9320–9327. doi:10.1074/jbc.M113.535278

Hadjikyriacou, A., Yang, Y., Espejo, A., Bedford, M.T., Clarke, S.G., 2015. Unique Features of Human Protein Arginine Methyltransferase 9 (PRMT9) and Its Substrate RNA Splicing Factor SF3B2. J. Biol. Chem. 290, 16723–16743. doi:10.1074/jbc.M115.659433

Han, X., Li, R., Zhang, W., Yang, X., Wheeler, C.G., Friedman, G.K., Province, P., Ding, Q., You, Z., Fathallah-Shaykh, H.M., Gillespie, G.Y., Zhao, X., King, P.H., Nabors, L.B., 2014. Expression of PRMT5 correlates with malignant grade in gliomas and plays a pivotal role in tumor growth *in vitro*. J. Neurooncol. 118, 61–72. doi:10.1007/s11060-014-1419-0

Hanahan, D., Weinberg, R.A., 2011. Hallmarks of cancer: The next generation. Cell. doi:10.1016/j.cell.2011.02.013

Henry, M.F., Silver, P.A., 1996. A novel methyltransferase (Hmt1p) modifies poly(A)+-RNA-binding proteins. Mol. Cell. Biol. 16, 3668–78.

Hernandez-Aya, L.F., Gonzalez-Angulo, A.M., 2011. Targeting the phosphatidylinositol 3-kinase signaling pathway in breast cancer. Oncologist 16, 404–14. doi:10.1634/theoncologist.2010-0402

Hernandez, S., Dominko, T., 2016. Novel Protein Arginine Methyltransferase 8 Isoform Is Essential for Cell Proliferation. J. Cell. Biochem. 11, 2056–2066. doi:10.1002/jcb.25508

Hernandez, S.J., Dolivo, D.M., Dominko, T., 2017. PRMT8 demonstrates variant-specific expression in cancer cells and correlates with patient survival in breast, ovarian and gastric cancer. Oncol. Lett. 13, 1983–1989. doi:10.3892/ol.2017.5671

Herrmann, F., Bossert, M., Schwander, A., Akgün, E., Fackelmayer, F.O., 2004. Arginine methylation of scaffold attachment factor A by heterogeneous nuclear ribonucleoprotein particle-associated PRMT1. J. Biol. Chem. 279, 48774–48779. doi:10.1074/jbc.M407332200

Herrmann, F., Fackelmayer, F.O., 2009. Nucleo-cytoplasmic shuttling of protein arginine methyltransferase 1 (PRMT1) requires enzymatic activity. Genes to Cells 14, 309–317. doi:10.1111/j.1365-2443.2008.01266.x

Herrmann, F., Lee, J., Bedford, M.T., Fackelmayer, F.O., 2005. Dynamics of human protein arginine methyltransferase 1 (PRMT1) *in vivo*. J. Biol. Chem. 280, 38005–38010. doi:10.1074/jbc.M502458200

Herrmann, F., Pably, P., Eckerich, C., Bedford, M.T., Fackelmayer, F.O., 2009. Human protein arginine methyltransferases *in vivo* - distinct properties of eight canonical members of the PRMT family. J. Cell Sci. 122, 667–677. doi:10.1242/jcs.039933

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., MacLennan, G.T., Cheng, L., 2004. Aberrant expression of CARM1, a transcriptional coactivator of androgen receptor, in the development of prostate carcinoma and androgen-independent status. Cancer 101, 83–89. doi:10.1002/cncr.20327

Hoshimaru, M., Ray, J., Sah, D.W., Gage, F.H., 1996. Differentiation of the immortalized adult neuronal progenitor cell line HC2S2 into neurons by regulatable suppression of the v-myc oncogene. Proc. Natl. Acad. Sci. U. S. A. 93, 1518–23.

Invernizzi, C.F., Xie, B., Frankel, F.A., Feldhammer, M., Roy, B.B., Richard, S., Wainberg, M.A., 2007. Arginine methylation of the HIV-1 nucleocapsid protein results in its diminished function. AIDS 21, 795–805. doi:10.1097/QAD.0b013e32803277ae

Jaeger, C., Gonzalo Ruiz, A., Llinás, R., 1989. Organotypic slice cultures of dopaminergic neurons of substantia nigra. Brain Res. Bull. 22, 981–91.

Jain, K., Warmack, R.A., Debler, E.W., Hadjikyriacou, A., Stavropoulos, P., Clarke, S.G., 2016. Protein Arginine Methyltransferase Product Specificity Is Mediated by Distinct Active-site Architectures. J. Biol. Chem. 291, 18299–308. doi:10.1074/jbc.M116.740399

Jelinic, P., Stehle, J.-C., Shaw, P., 2006. The Testis-Specific Factor CTCFL Cooperates with the Protein Methyltransferase PRMT7 in H19 Imprinting Control Region Methylation. PLoS Biol. 4, e355. doi:10.1371/journal.pbio.0040355

Johnston, J.A., Ward, C.L., Kopito, R.R., 1998. Aggresomes: a cellular response to misfolded proteins. J. Cell Biol. 143, 1883–98.

Karkhanis, V., Wang, L., Tae, S., Hu, Y.-J., Imbalzano, A.N., Sif, S., 2012. Protein Arginine Methyltransferase 7 Regulates Cellular Response to DNA Damage by Methylating Promoter Histones H2A and H4 of the Polymerase δ Catalytic Subunit Gene, *POLD1*. J. Biol. Chem. 287, 29801–29814. doi:10.1074/jbc.M112.378281

Kashyap, A., Zimmerman, T., Ergul, N., Bosserhoff, A., Hartman, U., Alla, V., Bataille, F., Galle, P.R., Strand, S., Strand, D., 2013. The human Lgl polarity gene, Hugl-2, induces MET and suppresses Snail tumorigenesis. Oncogene 32, 1396–1407. doi:10.1038/onc.2012.162

Katsanis, N., Yaspo, M.L., Fisher, E.M.C., 1997. Identification and mapping of a novel human gene, HRMT1L1, homologous to the rat protein arginine N-methyltransferase 1 (PRMT1) gene. Mamm. Genome 8, 526–529. doi:10.1007/s003359900491

Kim, J., Park, K., Ishida, J., Kako, K., Hamada, J., Kani, S., Takeuchi, M., Namiki, K., Fukui, H., Fukuhara, S., Hibi, M., Kobayashi, M., Kanaho, Y., 2015. PRMT8 as a phospholipase regulates Purkinje cell dendritic arborization and motor coordination. Neurscience 1–9. doi:10.1126/sciadv.1500615

Kleinschmidt, M.A., Streubel, G., Samans, B., Krause, M., Bauer, U.-M., 2008. The protein arginine methyltransferases CARM1 and PRMT1 cooperate in gene regulation. Nucleic Acids Res. 36, 3202–3213. doi:10.1093/nar/gkn166

Kousaka, A., Mori, Y., Koyama, Y., Taneda, T., Miyata, S., Tohyama, M., 2009. The distribution and characterization of endogenous protein arginine N-methyltransferase 8 in mouse CNS. Neuroscience 163, 1146–1157. doi:10.1016/j.neuroscience.2009.06.061

Kuhn, P., Chumanov, R., Wang, Y., Ge, Y., Burgess, R.R., Xu, W., 2011. Automethylation of CARM1 allows coupling of transcription and mRNA splicing. Nucleic Acids Res. 39, 2717–2726. doi:10.1093/nar/gkq1246

Kwon, K., Streaker, E.D., Ruparelia, S., Beckett, D., 2000. Multiple Disordered Loops Function in Corepressor-induced Dimerization of the Biotin Repressor. J. Mol. Biol. 304, 821–833. doi:10.1006/jmbi.2000.4249

Lacroix, M., Messaoudi, S. El, Rodier, G., Le Cam, A., Sardet, C., Fabbrizio, E., 2008. The histone-binding protein COPR5 is required for nuclear functions of the protein arginine methyltransferase PRMT5. EMBO Rep. 9, 452–458. doi:10.1038/embor.2008.45

Larsen, S.C., Sylvestersen, K.B., Mund, A., Lyon, D., Mullari, M., Madsen, M. V., Daniel, J.A., Jensen, L.J., Nielsen, M.L., 2016. Proteome-wide analysis of arginine monomethylation reveals widespread occurrence in human cells. Sci. Signal. 9, rs9-rs9. doi:10.1126/scisignal.aaf7329

Latonen, L., Moore, H.M., Bai, B., Jäämaa, S., Laiho, M., 2011. Proteasome inhibitors induce nucleolar aggregation of proteasome target proteins and polyadenylated RNA by altering ubiquitin availability. Oncogene 30, 790–805. doi:10.1038/onc.2010.469

Le Romancer, M., Treilleux, I., Leconte, N., Robin-Lespinasse, Y., Sentis, S., Bouchekioua-Bouzaghou, K., Goddard, S., Gobert-Gosse, S., Corbo, L., 2008. Regulation of Estrogen Rapid Signaling through Arginine Methylation by PRMT1. Mol. Cell 31, 212–221. doi:10.1016/j.molcel.2008.05.025

Lee, J.-H., Cook, J.R., Yang, Z.-H., Mirochnitchenko, O., Gunderson, S.I., Felix, A.M., Herth, N., Hoffmann, R., Pestka, S., 2005. PRMT7, a New Protein Arginine Methyltransferase That Synthesizes Symmetric Dimethylarginine. J. Biol. Chem. 280, 3656–3664. doi:10.1074/jbc.M405295200

Lee, J., Bedford, M.T., 2002. PABP1 identified as an arginine methyltransferase substrate using high-density protein arrays. EMBO Rep. 3, 268–273. doi:10.1093/embo-reports/kvf052

Lee, J., Cheng, D., Bedford, M.T., 2004. Techniques in Protein Methylation, in: Signal Transduction Protocols. Humana Press, New Jersey, pp. 195–208. doi:10.1385/1-59259-816-1:195

Lee, J., Sayegh, J., Daniel, J., Clarke, S., Bedford, M.T., 2005. PRMT8, a new membrane-bound tissue-specific member of the protein arginine methyltransferase family. J. Biol. Chem. 280, 32890–32896. doi:10.1074/jbc.M506944200

Li, Y., Zhu, R., Wang, W., Fu, D., Hou, J., Ji, S., Chen, B., Hu, Z., Shao, X., Yu, X., Zhao, Q., Zhang, B., Du, C., Bu, Q., Hu, C., Tang, Y., Zhong, L., Yang, S., Zhao, Y., Cen, X., 2015. Arginine methyltransferase 1 in the nucleus accumbens regulates behavioral effects of cocaine. J. Neurosci. 35, 12890–12902. doi:10.1523/JNEUROSCI.0246-15.2015

Lin, W.J., Gary, J.D., Yang, M.C., Clarke, S., 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, 15034–15044. doi:10.1074/jbc.271.25.15034

Lin, Y. ling, Tsai, Y.J., Liu, Y.F., Cheng, Y.C., Hung, C.M., Lee, Y.J., Pan, H., Li, C., 2013. The Critical Role of Protein Arginine Methyltransferase prmt8 in Zebrafish Embryonic and Neural Development Is Non-Redundant with Its Paralogue prmt1. PLoS One 8. doi:10.1371/journal.pone.0055221

Liu, X., Wang, L., Li, H., Lu, X., Hu, Y., Yang, X., Huang, C., Gu, D., 2014. Coactivator-associated arginine methyltransferase 1 targeted by miR-15a regulates inflammation in acute coronary syndrome. Atherosclerosis 233, 349–356. doi:10.1016/j.atherosclerosis.2014.01.039

Liu, Z., Zhou, Z., Chen, G., Bao, S., 2007. A putative transcriptional elongation factor hIws1 is essential for mammalian cell proliferation. Biochem. Biophys. Res. Commun. 353, 47–53. doi:10.1016/j.bbrc.2006.11.133

Lotharius, J., Barg, S., Wiekop, P., Lundberg, C., Raymon, H.K., Brundin, P., 2002. Effect of mutant alpha-synuclein on dopamine homeostasis in a new human mesencephalic cell line. J. Biol. Chem. 277, 38884–94. doi:10.1074/jbc.M205518200

Lotharius, J., Falsig, J., van Beek, J., Payne, S., Dringen, R., Brundin, P., Leist, M., 2005. Progressive degeneration of human mesencephalic neuron-derived cells triggered by dopamine-dependent oxidative stress is dependent on the mixed-lineage kinase pathway. J. Neurosci. 25, 6329–42. doi:10.1523/JNEUROSCI.1746-05.2005

Mathioudaki, K., Scorilas, A., Ardavanis, A., Lymberi, P., Tsiambas, E., Devetzi, M., Apostolaki, A., Talieri, M., 2011. Clinical evaluation of PRMT1 gene expression in breast cancer. Tumor Biol. 32, 575–582. doi:10.1007/s13277-010-0153-2

Meyer, R., Wolf, S.S., Obendorf, M., 2007. PRMT2, a member of the protein arginine methyltransferase family, is a coactivator of the androgen receptor. J. Steroid Biochem. Mol. Biol. 107, 1–14. doi:10.1016/j.jsbmb.2007.05.006

Miranda, T.B., Miranda, M., Frankel, A., Clarke, S., 2004. PRMT7 Is a Member of the Protein Arginine Methyltransferase Family with a Distinct Substrate Specificity. J. Biol. Chem. 279, 22902–22907. doi:10.1074/jbc.M312904200

Neault, M., Mallette, F.A., Vogel, G., Michaud-Levesque, J., Richard, S., 2012. Ablation of PRMT6 reveals a role as a negative transcriptional regulator of the p53 tumor suppressor. Nucleic Acids Res. 40, 9513–9521. doi:10.1093/nar/gks764

Paul, G., Christophersen, N.S., Raymon, H., Kiaer, C., Smith, R., Brundin, P., 2007. Tyrosine hydroxylase expression is unstable in a human immortalized mesencephalic cell line—Studies *in vitro* and after intracerebral grafting *in vivo*. Mol. Cell. Neurosci. 34, 390–399. doi:10.1016/j.mcn.2006.11.010

Pawlak, M.R., Scherer, C.A., Chen, J., Roshon, M.J., Ruley, H.E., 2000. Arginine N-methyltransferase 1 Is Required for Early Postimplantation Mouse Development, but Cells Deficient in the Enzyme Are Viable. Mol. Cell. Biol. 20, 4859–4869. doi:10.1128/MCB.20.13.4859-4869.2000

Reintjes, A., Fuchs, J.E., Kremser, L., Lindner, H.H., Liedl, K.R., Huber, L.A., Valovka, T., 2016. Asymmetric arginine dimethylation of RelA provides a repressive mark to modulate TNFα/NF-κB response. Proc. Natl. Acad. Sci. 113, 4326–4331. doi:10.1073/pnas.1522372113

Rodriguez-Gonzalez, A., Lin, T., Ikeda, A.K., Simms-Waldrip, T., Fu, C., Sakamoto, K.M., 2008. Role of the Aggresome Pathway in Cancer: Targeting Histone Deacetylase 6-Dependent Protein Degradation. Cancer Res. 68, 2557–2560. doi:10.1158/0008-5472.CAN-07-5989

Roux, K.J., Kim, D.I., Raida, M., Burke, B., 2012. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J. Cell Biol. 196, 801–810. doi:10.1083/jcb.201112098

Sayegh, J., Webb, K., Cheng, D., Bedford, M.T., Clarke, S.G., 2007. Regulation of protein arginine methyltransferase 8 (PRMT8) activity by its N-terminal domain. J. Biol. Chem. 282, 36444–36453. doi:10.1074/jbc.M704650200

Schildknecht, S., Poltl, D., Nagel, D.M., Matt, F., Scholz, D., Lotharius, J., Schmieg, N., Salvo-Vargas, A., Leist, M., 2009. Requirement of a dopaminergic neuronal phenotype for toxicity of low concentrations of 1-methyl-4-phenylpyridinium to human cells. Toxicol. Appl. Pharmacol. 241, 23–35. doi:10.1016/j.taap.2009.07.027

Scholz, D., Pöltl, D., Genewsky, A., Weng, M., Waldmann, T., Schildknecht, S., Leist, M., 2011. Rapid, complete and large-scale generation of post-mitotic neurons from the human LUHMES cell line. J. Neurochem. 119, 957–71. doi:10.1111/j.1471-4159.2011.07255.x

Scorilas, A., Black, M.H., Talieri, M., Diamandis, E.P., 2000. Genomic Organization, Physical Mapping, and Expression Analysis of the Human Protein Arginine Methyltransferase 1 Gene. Biochem. Biophys. Res. Commun. 278, 349–359. doi:10.1006/bbrc.2000.3807

Scott, H.S., Antonarakis, S.E., Lalioti, M.D., Rossier, C., Silver, P.A., Henry, M.F., 1998. Identification and characterization of two putative human arginine methyltransferases (HRMT1L1 and HRMT1L2). Genomics 48, 330–40. doi:10.1006/geno.1997.5190

Seligson, D.B., Horvath, S., Shi, T., Yu, H., Tze, S., Grunstein, M., Kurdistani, S.K., 2005. Global histone modification patterns predict risk of prostate cancer recurrence. Nature 435, 1262–1266. doi:10.1038/nature03672

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

Simandi, Z., Czipa, E., Horvath, A., Koszeghy, A., Bordas, C., Põliska, S., Juhász, I., Imre, L., Szabõ, G., Dezso, B., Barta, E., Sauer, S., Karolyi, K., Kovacs, I., Hutõczki, G., Bognár, L., Klekner, Á., Szucs, P., Bálint, B.L., Nagy, L., 2015. PRMT1 and PRMT8 regulate retinoic acid-dependent neuronal differentiation with implications to neuropathology. Stem Cells 33, 726–741. doi:10.1002/stem.1894

Singhroy, D.N., Mesplède, T., Sabbah, A., Quashie, P.K., Falgueyret, J.-P., Wainberg, M.A., 2013. Automethylation of protein arginine methyltransferase 6 (PRMT6) regulates its stability and its anti-HIV-1 activity. Retrovirology 10, 73. doi:10.1186/1742-4690-10-73

Solari, C., Echegaray, C.V., Luzzani, C., Cosentino, M.S., Waisman, A., Petrone, M.V., Francia, M., Sassone, A., Canizo, J., Sevlever, G., Barañao, L., Miriuka, S., Guberman, A., 2016. Protein arginine Methyltransferase 8 gene is expressed in pluripotent stem cells and its expression is modulated by the transcription factor Sox2. Biochem. Biophys. Res. Commun. 473, 194–199. doi:10.1016/j.bbrc.2016.03.077

Swiercz, R., PERSON, M.D., BEDFORD, M.T., 2005. Ribosomal protein S2 is a substrate for mammalian PRMT3 (protein arginine methyltransferase 3). Biochem. J. 386, 85–91. doi:10.1042/BJ20041466

Taneda, T., Miyata, S., Kousaka, A., Inoue, K., Koyama, Y., Mori, Y., Tohyama, M., 2007. Specific regional distribution of protein arginine methyltransferase 8 (PRMT8) in the

mouse brain. Brain Res. 1155, 1–9. doi:10.1016/j.brainres.2007.03.086

Tang, J., Frankel, A., Cook, R.J., Kim, S., Paik, W.K., Williams, K.R., Clarke, S., Herschman, H.R., 2000. PRMT1 is the predominant type I protein arginine methyltransferase in mammalian cells. J. Biol. Chem. 275, 7723–7730. doi:10.1074/jbc.275.11.7723

Tang, J., Gary, J.D., Clarke, S., 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.

Tee, W.-W., Pardo, M., Theunissen, T.W., Yu, L., Choudhary, J.S., Hajkova, P., Surani, M.A., 2010. Prmt5 is essential for early mouse development and acts in the cytoplasm to maintain ES cell pluripotency. Genes Dev. 24, 2772–2777. doi:10.1101/gad.606110

Thandapani, P., O'Connor, T.R., Bailey, T.L., Richard, S., 2013. Defining the RGG/RG Motif. Mol. Cell 50, 613–623. doi:10.1016/j.molcel.2013.05.021

Tiwari, N., Gheldof, A., Tatari, M., Christofori, G., 2012. EMT as the ultimate survival mechanism of cancer cells. Semin. Cancer Biol. 22, 194–207. doi:10.1016/j.semcancer.2012.02.013

Toma-Fukai, S., Kim, J.D., Park, K.E., Kuwabara, N., Shimizu, N., Krayukhina, E., Uchiyama, S., Fukamizu, A., Shimizu, T., 2016. Novel helical assembly in arginine methyltransferase 8. J. Mol. Biol. 428, 1197–1208. doi:10.1016/j.jmb.2016.02.007

Tripsianes, K., Madl, T., Machyna, M., Fessas, D., Englbrecht, C., Fischer, U., Neugebauer, K.M., Sattler, M., 2011. Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. Nat. Struct. Mol. Biol. 18, 1414–1420. doi:10.1038/nsmb.2185

Verbiest, V., Montaudon, D., Tautu, M.T., Moukarzel, J., Portail, J.-P., Markovits, J., Robert, J., Ichas, F., Pourquier, P., 2008. Protein arginine (*N*)-methyl transferase 7 (PRMT7) as a potential target for the sensitization of tumor cells to camptothecins. FEBS Lett. 582, 1483–1489. doi:10.1016/j.febslet.2008.03.031

Vu, L., Perna, F., Wang, L., Voza, F., Figueroa, M., Tempst, P., Erdjument-Bromage, H., Gao, R., Chen, S., Paietta, E., Deblasio, T., Melnick, A., Liu, Y., Zhao, X., Nimer, S., 2013. PRMT4 Blocks Myeloid Differentiation by Assembling a Methyl-RUNX1-Dependent Repressor Complex. Cell Rep. 5, 1625–1638. doi:10.1016/j.celrep.2013.11.025

Waldmann, T., Izzo, A., Kamieniarz, K., Richter, F., Vogler, C., Sarg, B., Lindner, H., Young, N.L., Mittler, G., Garcia, B. a, Schneider, R., 2011. Methylation of H2AR29 is a novel repressive PRMT6 target. Epigenetics Chromatin 4, 11. doi:10.1186/1756-8935-4-11

Walport, L.J., Hopkinson, R.J., Chowdhury, R., Schiller, R., Ge, W., Kawamura, A., Schofield, C.J., 2016. Arginine demethylation is catalysed by a subset of JmjC histone lysine demethylases. Nat. Commun. 7, 11974. doi:10.1038/ncomms11974

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., Zhang, Y., 2001. Methylation of Histone H4 at Arginine 3 Facilitating Transcriptional Activation by Nuclear Hormone Receptor. Science (80-.). 293, 853–857. doi:10.1126/science.1060781

Wang, M., Fuhrmann, J., Thompson, P.R., 2014. Protein Arginine Methyltransferase 5 Catalyzes Substrate Dimethylation in a Distributive Fashion. Biochemistry 53, 7884–7892. doi:10.1021/bi501279g

Wang, S., Tan, X., Yang, B., Yin, B., Yuan, J., Qiang, B., Peng, X., 2012. The role of protein arginine-methyltransferase 1 in gliomagenesis. BMB Rep. 45, 470–475.

doi:10.5483/BMBRep.2012.45.8.022

Weber, G.L., Parat, M.-O., Binder, Z.A., Gallia, G.L., Riggins, G.J., 2011. Abrogation of PIK3CA or PIK3R1 reduces proliferation, migration, and invasion in glioblastoma multiforme cells. Oncotarget 2, 833–49. doi:10.18632/oncotarget.346

Wu, C.-Y., Tsai, Y.-P., Wu, M.-Z., Teng, S.-C., Wu, K.-J., 2012. Epigenetic reprogramming and post-transcriptional regulation during the epithelial–mesenchymal transition. Trends Genet. 28, 454–463. doi:10.1016/j.tig.2012.05.005

Xu, J., Wang, A.H., Oses-Prieto, J., Makhijani, K., Katsuno, Y., Pei, M., Yan, L., Zheng, Y.G., Burlingame, A., Brückner, K., Derynck, R., 2013. Arginine Methylation Initiates BMP-Induced Smad Signaling. Mol. Cell 51, 5–19. doi:10.1016/j.molcel.2013.05.004

Xu, Z., Jiang, J., Xu, C., Wang, Y., Sun, L., Guo, X., Liu, H., 2013. MicroRNA-181 regulates CARM1 and histone arginine methylation to promote differentiation of human embryonic stem cells. PLoS One 8, e53146. doi:10.1371/journal.pone.0053146

Yadav, N., Cheng, D., Richard, S., Morel, M., Iyer, V.R., Aldaz, C.M., Bedford, M.T., 2008. CARM1 promotes adipocyte differentiation by coactivating PPARγ EMBO Rep. 9, 193–198. doi:10.1038/sj.embor.7401151

Yadav, N., Lee, J., Kim, J., Shen, J., Hu, M.C.-T., Aldaz, C.M., Bedford, M.T., 2003. Specific protein methylation defects and gene expression perturbations in coactivator-associated arginine methyltransferase 1-deficient mice. Proc. Natl. Acad. Sci. 100, 6464–6468. doi:10.1073/pnas.1232272100

Yang, Y., Bedford, M.T., 2013. Protein arginine methyltransferases and cancer. Nat. Rev. Cancer 13, 37–50. doi:10.1038/nrc3409

Yang, Y., Hadjikyriacou, A., Xia, Z., Gayatri, S., Kim, D., Zurita-Lopez, C., Kelly, R., Guo, A., Li, W., Clarke, S.G., Bedford, M.T., 2015. PRMT9 is a Type II methyltransferase that methylates the splicing factor SAP145. Nat. Commun. 6, 6428. doi:10.1038/ncomms7428

Yildirim, A.O., Bulau, P., Zakrzewicz, D., Kitowska, K.E., Weissmann, N., Grimminger, F., Morty, R.E., Eickelberg, O., 2006. Increased Protein Arginine Methylation in Chronic Hypoxia. Am. J. Respir. Cell Mol. Biol. 35, 436–443. doi:10.1165/rcmb.2006-0097OC

Yoshimatsu, M., Toyokawa, G., Hayami, S., Unoki, M., Tsunoda, T., Field, H.I., Kelly, J.D., Neal, D.E., Maehara, Y., Ponder, B.A.J., Nakamura, Y., Hamamoto, R., 2011. Dysregulation of PRMT1 and PRMT6, Type I arginine methyltransferases, is involved in various types of human cancers. Int. J. Cancer 128, 562–573. doi:10.1002/ijc.25366

Yoshimoto, T., Boehm, M., Olive, M., Crook, M.F., San, H., Langenickel, T., Nabel, E.G., 2006. The arginine methyltransferase PRMT2 binds RB and regulates E2F function. Exp. Cell Res. 312, 2040–2053. doi:10.1016/j.yexcr.2006.03.001

Yu, M.C., Lamming, D.W., Eskin, J.A., Sinclair, D.A., Silver, P.A., 2006. The role of protein arginine methylation in the formation of silent chromatin. Genes Dev. 20, 3249–54. doi:10.1101/gad.1495206

Yu, Z., Chen, T., Hebert, J., Li, E., Richard, S., 2009. A Mouse PRMT1 Null Allele Defines an Essential Role for Arginine Methylation in Genome Maintenance and Cell Proliferation. Mol. Cell. Biol. 29, 2982–2996. doi:10.1128/MCB.00042-09

Zhang, X., Cheng, X., 2003. Structure of the predominant protein arginine methyltransferase PRMT1 and analysis of its binding to substrate peptides. Structure 11, 509–520. doi:10.1016/S0969-2126(03)00071-6

Zhang, X., Zhou, L., Cheng, X., 2000. Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. EMBO J. 19, 3509–3519. doi:10.1093/emboj/19.14.3509

Zhao, Q., Rank, G., Tan, Y.T., Li, H., Moritz, R.L., Simpson, R.J., Cerruti, L., Curtis, D.J., Patel, D.J., Allis, C.D., Cunningham, J.M., Jane, S.M., 2009. PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing. Nat. Struct. Mol. Biol. 16, 304–311. doi:10.1038/nsmb.1568

Zhao, X., Jankovic, V., Gural, A., Huang, G., Pardanani, A., Menendez, S., Zhang, J., Dunne, R., Xiao, A., Erdjument-Bromage, H., Allis, C.D., Tempst, P., Nimer, S.D., 2008. Methylation of RUNX1 by PRMT1 abrogates SIN3A binding and potentiates its transcriptional activity. Genes Dev. 22, 640–653. doi:10.1101/gad.1632608

Zheng, S., Huang, J., Zhou, K., Zhang, C., Xiang, Q., Tan, Z., Wang, T., Fu, X., 2011. 17β-Estradiol Enhances Breast Cancer Cell Motility and Invasion via Extra-Nuclear Activation of Actin-Binding Protein Ezrin. PLoS One 6, e22439. doi:10.1371/journal.pone.0022439

Zheng, Z., Schmidt-Ott, K.M., Chua, S., Foster, K.A., Frankel, R.Z., Pavlidis, P., Barasch, J., D'Agati, V.D., Gharavi, A.G., 2005. A Mendelian locus on chromosome 16 determines susceptibility to doxorubicin nephropathy in the mouse. Proc. Natl. Acad. Sci. 102, 2502–2507. doi:10.1073/pnas.0409786102

Zhou, B.P., Deng, J., Xia, W., Xu, J., Li, Y.M., Gunduz, M., Hung, M.-C., 2004. Dual regulation of Snail by GSK-3β-mediated phosphorylation in control of epithelial–mesenchymal transition. Nat. Cell Biol. 6, 931–940. doi:10.1038/ncb1173

Zurita Rendón, O., Silva Neiva, L., Sasarman, F., Shoubridge, E.A., 2014. The arginine methyltransferase NDUFAF7 is essential for complex I assembly and early vertebrate embryogenesis. Hum. Mol. Genet. 23, 5159–5170. doi:10.1093/hmg/ddu239

Appendix

Article

A novel splicing isoform of Protein Arginine Methyltransferase 1 (PRMT1) that lacks the dimerization arm and correlates with cellular malignancy†

Running Head: A new PRMT1 variant without dimerization arm

Odysseas Patounas[#,1], Ioanna Papacharalampous[#,2], Carmen Eckerich[#,1], Georgios S. Markopoulos[3], Evangelos Kolettas[3], and Frank O. Fackelmayer[1,*]

[1] Institute of Molecular Biology and Biotechnology (IMBB-FORTH), Laboratory for Epigenetics and Chromosome Biology, University Campus Bldg. E5, 45110 Ioannina, Greece

[2] present address: Cardiothoracic Pharmacology, National Heart and Lung Institute, Imperial College, Dovehouse Street, London SW3 6LY, UK

[3] Laboratory of Biology, School of Medicine, University of Ioannina, 45110 Ioannina, Greece, and Dept. of Biomedical Research, Institute of Molecular Biology and Biotechnology (IMBB-FORTH), University Campus Bldg. E5, 45110 Ioannina, Greece

[#] these authors contributed equally to this work

*corresponding author
Institute of Molecular Biology and Biotechnology (IMBB-FORTH)
Laboratory for Epigenetics and Chromosome Biology
University Campus, Bldg. E5
45110 Ioannina, Greece
frank@fackelmayer.eu
+30 26510 07199

Grant information: This work was partially supported by Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: THALES "CancerTFs" (MIS 379435) and THALES "IDIPRO" (MIS 379440), and by the EU COST Action CM1406 "Epigenetic Chemical Biology (EPICHEM)"

Additional Supporting Information may be found in the online version of this article.

Received 12 June 2017; Revised 22 August 2017; Accepted 24 August 2017
Journal of Cellular Biochemistry

DOI 10.1002/jcb.26373

Abstract

Methylation of arginine residues is an important modulator of protein function that is involved in epigenetic gene regulation, DNA damage response and RNA maturation, as well as in cellular signaling. The enzymes that catalyze this post-translational modification are called protein arginine methyltransferases (PRMTs), of which PRMT1 is the predominant enzyme. Human PRMT1 has previously been shown to occur in seven splicing isoforms, which are differentially abundant in different tissues, and have distinct substrate specificity and intracellular localization. Here we characterize a novel splicing isoform which does not affect the amino-terminus of the protein like the seven known isoforms, but rather lacks exons 8 and 9 which encode the dimerization arm of the enzyme that is essential for enzymatic activity. Consequently, the isoform does not form catalytically active oligomers with the other endogenous PRMT1 isoforms. Photobleaching experiments reveal an immobile fraction of the enzyme in the nucleus, in accordance with earlier results from our laboratory that had shown a tight association of inhibited or inactivated PRMT1 with chromatin and the nuclear scaffold. Thus, it apparently is able to bind to the same substrates as catalytically active PRMT1. This isoform is found in a variety of cell lines, but is increased in those of cancer origin or after expression of the EMT-inducing transcriptional repressor Snail1. We discuss that the novel isoform could act as a modulator of PRMT1 activity in cancer cells by acting as a competitive inhibitor that shields substrates from access to active PRMT1 oligomers.

keywords: protein arginine methylation; alternative splicing; dimerization arm; malignancy; epithelial-mesenchymal transition

Introduction

The transition from a normal to a neoplastic phenotype is a complex process that involves a multitude of pathways and factors that are not yet understood in sufficient detail. There is increasing evidence that the methylation of arginine residues in proteins plays an important role through its effects on epigenetic gene regulation, signal transduction, pre-mRNA splicing, DNA repair and the intracellular localization of proteins (for review, see Blanc and Richard, 2017). This common post-translational modification is catalyzed by a small family of protein arginine methyltransferases (PRMT, see e.g. Herrmann et al., 2009), of which PRMT1 is the predominant enzyme that is responsible for greater than 85% of protein arginine methylation (Tang et al., 2000). It is a type I arginine methyltransferase, and as such asymmetrically dimethylates arginine residues on one of the ω-guanidino nitrogen atoms of the arginine sidechain, using AdoMet as the methyl donor. Methylation by PRMT1 occurs preferentially at arginine residues that are flanked by one or more glycine residues (Gary and Clarke, 1998; Lee and Bedford, 2002). PRMT1 has a wide range of substrates in both the cytoplasm and the nucleus, is expressed in all embryonic and adult tissues, and is essential for early embryonic development (Pawlak et al., 2000).

The overall structure of monomeric PRMT1 can be divided into four parts: amino-terminal tail, AdoMet-binding domain, β-barrel and dimerization arm (Zhang and Cheng, 2003). The AdoMet-binding domain has the consensus fold conserved in other AdoMet-dependent methyltransferases (Cheng and Roberts, 2001), whereas the β-barrel domain is unique to the PRMT family (Zhang et al., 2000). The active site is located between these two domains, and contains two conserved glutamate residues. These negatively charged residues interact with the positively charged arginine in the substrate and are critical for catalysis (Zhang and Cheng, 2003). PRMT1 has the ability to form dimers and larger oligomers through its dimerization arm, and the oligomerization is essential for AdoMet binding and enzymatic activity. The ring-like dimer allows the product of the first methylation reaction, monomethylarginine, to enter the active site of the second molecule of the dimer, without releasing the substrate from the ring or replenishing the methyl donor. Thus, the dimerization permits the production of the final product, asymmetric dimethylarginine, in a processive manner. As for the substrate binding region of the PRMT1 molecule, three different peptide-binding channels have been identified, suggesting that different PRMT1 substrates might approach the active site from different angles (Zhang and Cheng, 2003).

Although the PRMT1 protein is often treated in the literature as a single entity, the Prmt1 gene is composed of 12 coding exons and the resulting pre-mRNA is subjected to extensive alternative splicing. Upon initial identification of the PRMT1 genomic locus, it was observed that three transcripts (named v1-v3) could be produced through alternative splicing of exons 2 and 3 (Scorilas et al., 2000). Later, inspection of available ESTs and comparison with complete genomic sequences revealed the existence of previously unidentified exons upstream of the exon harboring the AUG start codon for v1-v3 (Goulet et al., 2007). These alternative exons were called e1a-e1d, with e1d being the exon previously labeled as exon 1. A complex pattern of

alternative splicing takes place in about 1 kilobase of sequence upstream from that originally identified exon 1, and – potentially together with the use of alternative transcription start sites – gives rise to seven known isoforms (named v1-v7, Fig. 1A, suppl. Fig. S1) with distinct amino-terminal sequences. Among these isoforms, v4 is the only mRNA where translation initiation starts at a new AUG located in an 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. According to EST analysis, a number of different mRNAs can lead to the production of the v5 and v6 proteins from various possible pairing combinations of the exons e1a, e1d, e2 and e3. The predicted protein-coding regions (excluding 5'and 3'-untranslated regions) of the various Prmt1 transcripts are composed of 1059 nt (v1), 1113 nt (v2), 1041 nt (v3), 1047 nt (v4), 1026 nt (v5), 975 nt (v6), or 960 nt (v7) and encode seven deduced polypeptides with predicted molecular masses of 40.5, 42.5, 39.9, 40.1, 39.4, 37.7, and 36.7 kDa, respectively.

The isoforms appear to serve different roles, as their unique amino-terminal sequences affect their substrate specificity and enzymatic activity. In addition, the relative expression of these isoforms may regulate the overall PRMT1 function in the living cell, because their expression is partially tissue-specific and they exhibit distinct subcellular localization (Goulet et al., 2007). For example, the PRMT1v2 isoform contains a leucine-rich, CRM1-dependent nuclear export signal, and its relative expression was found to be increased in breast cancer cell lines and tumors. More recently, the same laboratory demonstrated that specific depletion of PRMT1v2 by RNA interference causes a decrease in cancer cell survival due to induction of apoptosis, and significantly decreased cell invasion in an aggressive cancer cell line (Baldwin et al., 2012). In keeping with this, PRMT1v2 overexpression in a non-aggressive cancer cell line was sufficient to render them more invasive. This novel tumor promoting activity was specific to PRMT1v2, as overexpression of other isoforms did not enhance cell invasion. Thus, distinct splicing isoforms of PRMT1 have a different impact on the process of neoplastic transformation. This motivated us to search for additional splicing isoforms that may play a role in cancer development.
Here, we describe the discovery and characterization of a novel isoform of PRMT1, which differs from the known variants because it is not adding variability to the amino terminus of the protein, but rather lacks the dimerization arm that is essential for the formation of enzymatically active PRMT1 oligomers. We find that the relative abundance of this splice variant is correlated with cellular malignancy, and can experimentally be increased by expression of the epithelial-to-mesenchymal-transition (EMT) inducing transcriptional repressor Snail1.

Materials & Methods

Cell culture and transfection. Cells were cultured as a monolayer in tissue-culture grade plastic dishes at 37°C under humidified atmosphere containing 5% CO_2. The growth medium consisted of DMEM for HEK293 (immortalized human embryonic kidney), HeLa (human cervix carcinoma), the normal lung fibroblasts HFL-1 and MRC-5, and the lung cancer lines H1299 and A549, or RPMI1640 for lung cancer lines H1437 and H1792, according to instructions provided by ATCC. All media contained

10% FCS and 1x Penicillin/Streptomycin. The cells were passaged 1:3 or 1:4 every 2-3 days shortly before they became confluent.
Cells were transfected with Polyethylenimine as described previously (Helbig and Fackelmayer, 2003; Mearini and Fackelmayer, 2006) or Lipofectamine 2000 (Invitrogen) as per manufacturer's instructions, and analyzed at least 24 h post transfection. To generate cell lines that stably express PRMT1v1:YFP, PRMT1Δarm:YFP or fibrillarin:FusionRed, respectively, HEK293 cells were split 1:10 one day post transfection, and were kept in selective medium (DMEM plus 400 µg/ml G418) without passaging until resistant colonies had developed (4 to 6 weeks). Cells were then trypsinized and passaged 1:3 for three times, before they were frozen for cryopreservation. For some experiments, single cell clones were selected from this population by limited dilution.
Live cell morphology (Fig. 4A) was analyzed after total membrane staining with 1 µg/ml of the lipophilic dye 3,3'-Dihexyloxacarbocyanine Iodide ($DiOC_6$; Invitrogen) in cell culture medium for 15 min; before microscopic analysis, cells were washed three times with medium to remove unbound dye.

Photobleaching experiments and FACS. Fluorescence Recovery After Photobleaching (FRAP) experiments were performed with whole populations of HEK293 cells stably expressing YFP fusion constructs of full length PRMT1v1 and PRMT1Δarm, exactly as described previously (Fackelmayer, 2005; Herrmann et al., 2005). Fluorescence-activated cell sorting (FACS) was performed by detaching the cells from the dishes with trypsin and incubation with hypotonic solution (0.1 % w/v sodium citrate, 0.1 % v/v Triton X-100) containing 50 mg/l propidium iodide for 15 minutes, as described previously (Vartholomatos et al., 2015).

Cell extracts and immunoprecipitation. For SDS PAGE analysis, whole cell extracts were prepared by lysing cells from one 60 mm dish in 400 µl of 1 % SDS in water, followed by brief sonication to reduce viscosity. Then, protein was precipitated according to Wessel and Flügge (1984) by vigorously mixing with an equal volume of methanol and addition of 100 µl of chloroform. After mixing, precipitated protein was recovered by centrifugation for 5 min at full speed, air dried, then re-dissolved in SDS gel loading buffer and heated to 65°C for 5 min.
To prepare native extracts for immunoprecipitation and activity assays, cells were washed twice with cold PBS, and lyzed in 1 ml of lysis buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 1 % NP-40). The sample was sonified to shear DNA, and centrifuged for 5 min at 10000 rpm at 4°C to preclear the lysate. A small amount of the supernatant was kept as loading control. For immunoprecipitation, the supernatant was mixed with 4 µg of antibody and incubated for 1 h at 4°C in a rotating mixer. Then, 30 µl of 50 % Protein G Sepharose suspension were added, and the sample was incubated for an additional hour under the same conditions.
Immunoprecipitated proteins bound to the beads were collected by centrifugation (10 sec, 3000 g). The beads were washed six times in 1 ml of cold lysis buffer each. Precipitated proteins were either eluted from the beads by adding 2x SDS gel sample buffer and boiling at 95°C, or used immediately for enzymatic activity assays.

RNA analysis. All procedures were done under RNase-free conditions. Total cellular RNA was extracted and purified using the NucleoSpin RNA Mini kit (Macherey-Nagel) according to the manufacturer's manual. After quality control by agarose gel electrophoresis, RNA was reverse-transcribed with the PrimeScript RT reagent kit (Takara #RR037A) according to protocol V of the manufacturer. The synthesized cDNA was stored at -20°C until further use. For RT-PCR, cDNA was diluted 1:5 to 1:10, and 1 µl of the dilution was amplified by PCR in a volume of 20 µl, containing 10 pmole of sense- and antisense primers, 0.4 µl dNTP mix (containing 10 mM per nucleotide), 2 µl 10x concentrated reaction buffer (final concentrations: 50 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris pH 9, 0.5 % Triton X100), and 2 units of homemade Taq polymerase. The final volume was adjusted to 20 µl with water. Amplification programs were carefully optimized in pilot experiments and according to melting temperatures of the primers, as given in suppl. table 1. To quantify transcripts, real-time quantitative PCR (qPCR) was performed in capillaries on a Roche LightCycler 2, using cyber green (Kapa #KK4602) in a total volume of 20 µl. All qPCR analyses were done on at least two biological replicates, in triplicate each. For normalization, we used primer sets for GAPDH, 18S RNA, and Alu, respectively (suppl. table 1).

Other methods. Radioactive activity assay, SDS page analysis, western blot, silver staining and agarose gel electrophoresis were carried out as described previously (Herrmann et al., 2004; Herrmann et al., 2009). Colorimetric quantitation of total protein in cell lysates was done using the Pierce BCA Protein Assay Kit (Thermo Scientific), according to the manual.

Results

PRMT1Δarm is a naturally occurring new variant of PRMT1

The presence of different isoforms, which are also found to vary in abundance between healthy and cancer cells (Goulet et al., 2007), suggests that the activity of PRMT1 is finely regulated through the balance of expressed isoforms. We have thus screened GenBank for potential additional isoforms, and found an Expressed Sequence Tag (EST) from human neuroblastoma cells (accession# BX352789.2) that differs from the previously characterized isoforms. In the new form, exons 8 and 9 are skipped, directly fusing exon 7 and 10. This results in a deletion of amino acids 168 to 234 of the enzyme, which exactly encode the known dimerization arm of PRMT1 (Fig. 1). It is thus similar to the artificial construct PRMT1Δarm that was investigated by Zhang & Cheng (2003). We find, for the first time, that such an isoform also exists in nature, at least as an EST in the GenBank database.

To verify the presence of PRMT1Δarm in cell lines, we developed specific primer sets (suppl. table 1; also see suppl. Fig. S1 for the positioning of all used PCR primers) for use in RT-PCR. For semi-quantitative PCR, the forward primer (P1Δarmfwd; 5'-CGGGACAAGTGGCTG*GAGGTG*-3') hybridizes to a sequence across the border of exons 7 and 10 (in italics) of the prmt1 gene, and is therefore specific for PRMT1Δarm that lacks exons 8 and 9. The reverse primer (Pan12r; 5'-CGAAACCGCCTAGGAACGCT-3') hybridizes to a sequence in exon 12 that is common to all PRMT1 isoforms. For PRMT1Δarm, this PCR produces a specific amplicon of

429bp; no amplification was observed from sequences containing exon 8 and 9, using cloned PRMT1v1 as a PCR template (not shown). For quantitative PCR, the same forward primer was combined with a different reverse primer (Pan10rq; 5'-CAGGGCGTGCACGTAGTCAT-3'), to produce a specific amplicon of 108bp. When performed on RNA prepared from HEK293 and HeLa cells, expression of PRMT1Δarm was clearly detectable at comparable levels after 25 cycles of semi-quantitative PCR amplification (Fig. 2A, also compare suppl. Fig. S2). To examine which other alternative exons, upstream of exon 7, are present in the novel isoform PRMT1Δarm, we designed a PRMT1Δarm-specific reverse primer, which again hybridized across the border of exon 7 and 10, but in the opposite orientation. This primer (P1Δarmrev; 5'-GGTATAGATGTCCACCTC*CAGCC* -3'), was then combined with published or newly designed forward primers specific for the known amino-terminal splicing isoforms (Fig. 2B). PCR reactions with these primer pairs would yield specific products that correspond to the 5' end of the known seven isoforms. For a first analysis, we used the published forward primer PF (Scorilas et al., 2000) located in exon e1d. This primer, combined with P1Δarmrev, would amplify all known isoforms except v4 and v6, with characteristic product sizes (v1: 507bp; v2: 561bp; v3: 680bp; v5: 665bp; v7: 405bp). We obtained only one strong band, at the predicted size corresponding to variant v1.

For a more detailed analysis, we used published primers (Goulet et al., 2007) for variant 5 and 7, and newly designed primers for variant 3 (v3s), located in the longer form of exon 3 unique to variant 3, for variant 4 (v4s), overlapping the junction between exon e1c and E4, and for variant 6 (v6s), located in exon e1b. The isoform-specific primers gave only very weak products for isoform 3 and 7, and no signals for isoform 4, 5, and 6. In control PCR reactions to verify the presence of the known isoforms in HeLa cells, we used the same specific forward primers and combined them with the common reverse primer (Pan12r) in exon 12. We find products of the expected size for all isoforms, with the exception of variant 4 which appears to be lacking in HeLa cells (Fig. 2B, lower panel). To further confirm that PRMT1Δarm contains the upstream exons of variant 1, but not those of the other major isoform v2, we cloned a "Δarm version" of PRMT1v2, and used it to compare the amplification product from this plasmid with the products from the Δarm version of variant 1, and from HeLa cDNA. We find only one band amplified from HeLa cDNA, with a size corresponding to the 5'-end of variant 1, but not of variant 2.

Thus, together, these two experiments clearly show that the lack of the dimerization arm in exon 8 and 9 is almost exclusively combined with a 5' end identical to variant 1 in HeLa cells.

PRMT1Δarm expression is correlated with cellular malignancy

To expand our analysis to other cell lines, we next analyzed the level of PRMT1Δarm expression in two normal diploid lung fibroblast strains, HFL-1 and MRC-5, and a panel of four lung cancer lines with different status of p53 and K-Ras (suppl. table 2). We find that the expression level of PRMT1Δarm is consistently higher in the cancer cell lines when investigated by semi-quantitative PCR (Fig. 3B) and real-time quantitative RT-PCR (Fig. 3C, also compare suppl. Fig. S2). The results, which were readily reproducible in five independent experiments, show that H1792 has the highest level of PRMT1Δarm expression of all the tested cell lines (7 times more than

HFL-1 and 25 times more than MRC-5). The second highest PRMT1Δarm expression levels are found in H1299 cells (3 times more than HFL-1 and 11 times more than MRC-5). A549 cells express slightly less than H1299 (2.5 times more than HFL-1 and 10 times more than MRC-5). Finally, H1437 express the lowest amount of PRMT1Δarm of all tested cancer cell lines (2 times more than HFL-1 and 6 times more than MRC-5). This higher abundance of the PRMT1Δarm splicing isoform in cancer cells is in line with reports that the expression levels of PRMT1 variants are altered between normal and cancer cell lines, in that case from breast cancer (Baldwin et al., 2012; Goulet et al., 2007; Scorilas et al., 2000). Furthermore, the relative expression levels suggest a correlation to the presence of cells that display mesenchymal phenotype. For example, the H1792 cell line originates from a stage 4 cancer and is the only cell line in our panel with mutated K-Ras and mutated p53 with almost p53 null phenotype, had the highest levels of PRMT1Δarm. This line is a mixture of epithelial and mesenchymal-like cells, suggesting that it has at least partially gone through epithelial-to-mesenchymal transition (EMT). Transformed cells that have undergone EMT are known for their invasiveness, aggressiveness and their metastatic abilities. On the other hand, H1437 cells have wild type K-Ras and a point mutation in p53 and strictly epithelial phenotype, show the lowest levels of PRMT1Δarm among the investigated cancer cell lines.

To investigate whether PRMT1Δarm directly correlates with EMT, we transfected HEK293 cells with an expression vector for flag-tagged Snail-6SA, a stabilized form of the transcriptional repressor Snail1 that is critically involved in driving EMT (see, e.g. Zhou et al., 2004). In line with earlier data from other laboratories who had used HEK293 cells as an EMT model (see e.g. Kashyap et al., 2013; Wu et al., 2009), we find that expression of Snail leads to morphological changes towards more spindle-shaped cells (Fig. 4A, left panel), concomitant with a strong reduction of E-Cadherin expression (Fig. 4A, right panel), both of which are characteristic for cells undergoing EMT. Immunofluorescence with an anti-flag antibody confirmed the proper nuclear localization of Snail (Fig. 3B). We also transfected HEK293 with the Snail-6SA expression vector by two independent methods of transfection (Lipofectamine 2000 and Polyethylenimine, respectively), to establish stable cell lines. After selection, we prepared RNA from these cells and investigated them by quantitative RT-PCR for the presence of total PRMT1 (all isoforms, amplified with the "pan" PRMT1 primer set Panfwd and Pan12r) and PRMT1Δarm (amplified with primers P1Δarmfwd and Pan10rq). This analysis showed that the level of total PRMT1 is increased by about a factor of 3 in comparison to the non-transfected parental HEK293 cell line. Interestingly, however, the level of PRMT1Δarm is significantly further increased, revealing a six- to sevenfold up-regulation of PRMT1Δarm in comparison to controls, and independently of the transfection method (Fig. 4C). Thus, Snail enhances the splicing of the PRMT1 pre-mRNA in favor of PRMT1Δarm, in line with a correlation between PRMT1Δarm and EMT or malignancy.

PRMT1Δarm expression leads to G1 cell cycle block and cell death

In a further set of experiments, we characterized the PRMT1Δarm protein, and its potential function in live cells. To this end, we cloned the coding region of PRMT1Δarm into a pEYFP-N1 vector to express a yellow fluorescent fusion protein, and used full length PRMT1v1:YFP as a control in all following experiments.

Microscopic observation of HEK293 cells transiently or stably transfected with PRMT1v1:EYFP and PRMT1Δarm:EYFP, respectively, consistently indicated significantly lower expression levels of the PRMT1Δarm variant. Moreover, increased cell death was observed for the PRMT1Δarm-transfected cells two days after transfection. In stable cell populations after six weeks of selection, the cells which managed to survive were those with the lowest expression levels of PRMT1Δarm, while a large range of expression levels was found for full length PRMT1v1:EYFP. To quantify the difference in average expression between PRMT1v1:EYFP and PRMT1Δarm:EYFP, we performed western blotting with total cell extracts prepared 24h post transfection with the two expression constructs. We find that the level of the expressed full-length PRMT1 is about four times higher than that of PRMT1Δarm (Fig. 5A), in agreement with the results from microscopic observation. Identical results were also obtained by fluorescence-activated cell sorting (FACS). In these experiments, we used propidium iodide for cell cycle analysis in parallel with the expression levels comparison for the two proteins. As shown in Fig. 5B, expression of PRMT1Δarm:EYFP, but not full length PRMT1v1:EYFP, results in a arrest in the G1 phase of the cell cycle, and in increased cell death as shown by an increase of the sub-G1 fraction of cells from 6.4% to 17.4%. We conclude that a considerable expression of PRMT1Δarm is not compatible with cell cycle progression, and consequently leads to cell death by apoptosis.

PRMT1Δarm accumulates in nucleoli under proteasome inhibition

Interestingly, already shortly after transfection, the expression levels of PRMT1Δarm and full length PRMT1v1 were different by a factor of around 4, even though they were expressed from the same promoter and vector backbone. We thus hypothesized that the difference may be on protein level, perhaps due to increased proteosomal degradation of PRMT1Δarm. To address this point, we transfected HEK293 with the two constructs, and, starting 24 h after transfection, incubated them with 5 µM of the proteasomal inhibitor MG132 for 8 or 16 hours. We find that, when we analyzed total cell extracts by immunoblotting, the levels of both proteins remained stable independently of incubation time with MG132 (Fig. 6A). While inhibition of the proteasome did not result in changes in the levels of the two proteins, we microscopically observed a striking re-localization of PRMT1Δarm after 8 hours of MG132 treatment in HEK293 cells, which did not occur for full-length PRMT1 (Fig. 6B&C). Then, PRMT1Δarm is almost completely absent from the cytoplasm and has accumulated at the nucleolus as visualized by the nucleolar marker protein fibrillarin:FusionRed (a kind gift of the laboratory of Dmitriy Chudakov, Moscow). Similar re-localization after proteasome inhibition has been shown before for several other proteins, but the functional consequences of this phenomenon are not yet entirely clear (see discussion). Under the same conditions, full length PRMT1v1 does not significantly change its localization, and remains excluded from nucleoli (Fig. 6B&C).

PRMT1Δarm is catalytically inactive, but stably interacts with nuclear proteins

It has been shown earlier that PRMT1 obtains its enzymatic activity only after oligomerization (Zhang and Cheng, 2003); these authors had also used an artificial construct lacking the dimerization arm, similar to our naturally occurring

PRMT1Δarm isoform. As their recombinant construct had eluted as a monomer from gel filtration columns, and completely lacked enzymatic activity, we hypothesized that the lack of exons 8 and 9 in the PRMT1Δarm isoform, which code for the dimerization arm of PRMT1, would also not allow the interaction of PRMT1Δarm with itself or with the other PRMT1 isoforms. In order to investigate this possibility, we transiently transfected HEK293 cells with expression vectors encoding PRMT1v1:EYFP and PRMT1Δarm:EYFP, and performed immunoprecipitation with anti-GFP antibodies. Four times more starting material was used from the cells expressing PRMT1Δarm to account for the lower expression level of the protein (see above). The precipitated proteins were analyzed by Western blot with anti-PRMT1 antibodies, which recognize both the overexpressed fusion proteins and the endogenous PRMT1 isoforms. We find that endogenous PRMT1 readily co-immunoprecipitated with the expressed full-length PRMT1v1:YFP protein (Fig. 7A, lane 1), but not with PRMT1Δarm (Fig. 7A, lane 2). Interestingly, lane 1 also shows some high molecular weight bands that most probably represent remaining non-denatured PRMT1 oligomers that co-precipitate exclusively with full-length PRMT1v1:YFP. These results verify that the lack of the dimerization arm indeed results in the inability of dimerization (and further oligomerization) of the PRMT1Δarm variant with endogenous PRMT1 variants. Consequently, the PRMT1Δarm variant also lacks enzymatic activity (Fig. 7B) as determined by an assay with hypomethylated protein extracts as a substrate (compare Herrmann et al., 2004). When we performed silver staining of the same immunoprecipitation assays, however, we obtained a very similar pattern of co-precipitated proteins for PRMT1v1 and PRMT1Δarm (Fig. 7C). This demonstrates that PRMT1Δarm still can bind to many of the same factors that are also bound by (active) PRMT1v1, although some differences were also obvious (Fig. 7C, asterisks).

We had earlier demonstrated that a chemically inhibited or mutated, catalytically inactive PRMT1v2 accumulates in the nucleus (Herrmann et al., 2005; Herrmann and Fackelmayer, 2009). Interestingly, we find a similar enrichment in the nucleus also for PRMT1Δarm (Fig. 7D). As PRMT1Δarm is a naturally occurring inactive form of PRMT1, its molecular dynamics could also resemble those of inhibited or genetically inactivated enzyme. In a last set of experiments, we therefore measured the dynamics of PRMT1Δarm and compare it to the full-length variant 1 in live cells, by performing Fluorescence Recovery After Photobleaching (FRAP) experiments. To this end, we used HEK293 stably expressing the full-length and the PRMT1Δarm form of PRMT1v1 tagged with EYFP. A rectangular bleaching area that included both cytoplasmic and nuclear parts in at least 20 cells was photobleached and the fluorescence recovery of each area (nucleus and cytoplasm individually) was monitored through 250 post-bleach images over a time period of 50 seconds (Fig 8). This analysis revealed that both full-length PRMT1v1 and PRMT1Δarm are fully mobile in the cytoplasm, with recovery half-time of around 3 seconds, and thus behave like soluble proteins. In the nucleus, on the other hand, both proteins recover to only 75% of the initial fluorescence, and have an immobile fraction of 25%. We also note that PRMT1Δarm recovers more slowly, presumably indicating longer interaction with substrate proteins. So, indeed, PRMT1Δarm – but also full length PRMT1v1 – display molecular dynamics identical with our earlier data on

substrate-bound inactive PRMT1 (Herrmann et al., 2005), which we have added for comparison (Fig. 8, grey dotted line). Further research will be necessary to understand the reason for the immobile fraction of full length PRMT1v1, as it differs significantly from the PRMT1v2 variant we had investigated previously (Herrmann et al., 2005).

Discussion

PRMT1 is the predominant member of the human family of protein arginine methyltransferases, and has previously been described to occur in a variety of splicing variants that differ in their amino terminus (Goulet et al., 2007). Here, we characterize a novel splicing variant that lacks the dimerization arm of the enzyme, and thus represents the first example of a PRMT1 isoform with variation in the catalytic core of the protein. The novel variant, which we call PRMT1Δarm, correlates with the malignant phenotype in lung cancer lines, and is increased by expression of the transcriptional repressor Snail.

Several previous studies have demonstrated a role of members of the PRMT family in cell homeostasis and in different types of cancer (reviewed by Yang and Bedford, 2013). For example, PRMT1 and PRMT5 have been correlated with tumor aggressiveness, and their knockdown in both in vitro and in vivo studies had shown cell cycle arrest and reduced tumor volume (e.g. Baldwin et al., 2012; Wang et al., 2012; Han et al., 2014; Yan et al., 2014); also, high PRMT1 expression in breast cancer tissues correlates with poor patient prognosis and a reduced disease-free survival (Mathioudaki et al., 2011). However, the contribution of PRMTs to malignancy is still not well understood on a molecular level. In the case of PRMT1, we in particular must understand how the activity of PRMT1 is regulated on specific substrates, as this highly abundant enzyme fully methylates most of its substrates (see, e.g. Herrmann et al., 2004), but at the same time methylates only a small fraction of histone tails. One way of regulation is the interaction with proteins such as B-cell translocation gene 1 (BTG1) and BTG2, which promote its multimerization and result in increased PRMT1 activity and co-activator function (Lin et al., 1996; Lee et al., 2007). Results from the Côté laboratory suggest that the activity of PRMT1 can also be regulated by the balance of expressed splicing variants (Goulet et al., 2007). These variants are all able to form oligomers with each other, and can result in oligomeric holoenzymes with different activity and distinct substrate specificities in different cell types. In fact, PRMT1v2 has been shown to promote the survival and invasiveness of human breast cancer cells (Baldwin et al., 2012). Along this line, we had demonstrated that catalytically inactive PRMT1v2 stably binds to nuclear substrates, and might thus contribute to the regulation of PRMT1 activity as a competitive inhibitor (Herrmann et al., 2005). In the present paper, we describe the same behavior for PRMT1Δarm: It is catalytically inactive (Fig. 7B), accumulates in the nucleus (Fig. 7D) through stable interaction with other proteins (Fig. 7C), and gains an immobile fraction (Fig. 8) very similar to a catalytically inactive mutant of PRMT1v2 (Herrmann et al., 2005; Herrmann and Fackelmayer, 2009). With regard to interacting proteins, we find that the general pattern of co-precipitated proteins is very similar for both PRMT1 constructs, but also show some distinct differences. For

example, we detect a protein of around 40 kD exclusively in the sample with full length PRMT1v1, while we find two proteins of around 48 and 50 kD exclusively in the sample with PRMT1Δarm (labeled with asterisks in Fig. 7C). Of note, the enrichment of the 50 kD protein in the PRMT1Δarm sample reproduces the result of inhibited PRMT1v2 from our earlier publication Herrmann et al. (2005).

As expected from the lack of the dimerization arm, PRMT1Δarm cannot multimerize with other variants of endogenous PRMT1 (Fig. 7A). Nevertheless, PRMT1Δarm associates with endogenous interaction partners (substrates or regulators) of PRMT1, without being a component of an active oligomeric holoenzyme itself. Thus, the recognition and binding of interacting proteins is independent of multimerization of PRMT1, and of its enzymatic activity. Due to these interactions, PRMT1Δarm may act as a competitive inhibitor, shielding selected substrates from being methylated by an active PRMT1 oligomer. This may explain how the experimental overexpression of PRMT1Δarm leads to a block in the G1 phase of the cell cycle, and to morphological changes indicative of apoptosis as well as an increase of cells with sub-G1 content of DNA (Fig. 5B). This is in agreement with the previous observation that knock-down of PRMT1 results in an arrest in the G1-S phase of the cell cycle, proliferation inhibition and apoptosis induction in human glioma cell lines (Wang et al., 2012).

We find that in transiently transfected cells, the protein level of PRMT1Δarm is typically about one-fourth of active PRMT1v1 (Fig. 5A), and even less in stably expressing cells after several weeks of selection. This is reminiscent of earlier results from our laboratory which had demonstrated that a point-mutated, inactive version of PRMT1v2 can likewise only be expressed in low levels (Herrmann and Fackelmayer, 2009). Apparently, the cell counteracts the elevated amounts of PRMT1Δarm, resulting from its forced expression, on protein level (as the RNA is expressed from the same CMV promoter as full length PRMT1v1). Interestingly, when we investigated whether PRMT1Δarm is degraded by the proteasome, we found that only PRMT1Δarm, but not full length PRMT1v1, re-locates from predominantly cytoplasmic localization to nucleoli in almost all cells upon MG132 treatment (Fig. 6B). This is not a consequence of cytoplasmic PRMT1Δarm being degraded, because the total amount of the protein remains the same during MG132 treatment (Fig. 6A). More specifically, we found that PRMT1Δarm localizes close to, but not exactly co-localizing with fibrillarin (Fig. 6C), to structures that are reminiscent of “nucleolar aggresomes”. These structures have been found before in many cell types after treatment with proteasome inhibitors, and contain dozens of proteins, together with polyadenylated RNA (Latonen et al., 2011). Among the proteins found in these structures are components of the ubiquitin-proteasome system, including PML body proteins and ubiquitin itself, as well as targets including cancer-related transcription factors and cell cycle regulators (e.g. p53, MDM2, cyclins), and proteins involved in neurodegenerative diseases (e.g. ataxin-1, Malin). Most probably, the nucleolar aggresome proteins are in the process of being degraded, and their normal half lives are bound to have an effect on the kinetics of their appearance in nucleolar aggresomes (Latonen et al., 2011). We therefore hypothesize that the half-life of PRMT1Δarm is somewhat decreased in comparison

to full length PRMT1v1, and thus the lower steady-state amount of the PRMT1Δarm isoform most likely results from faster degradation.

We had originally found the PRMT1Δarm variant in database searches as an expressed sequence tag from a neuroblastoma (GenBank accession number BX352789.2). In follow-up experiments, we detected its expression in variable amounts in normal fibroblasts and different lines of transformed cells (Fig. 2A, Fig. 3). When investigated by semi-quantitative and quantitative RT-PCR, we find that the amount of PRMT1Δarm is significantly higher in human non-small cell lung cancer cell lines than in human normal diploid lung fibroblasts, and correlates with their reported aggressiveness (Fig. 3B): It was highest in H1792 that originate from a stage 4 lung adenocarcinoma and expressing mutant K-Ras and mutant p53, and lowest in H1437 from a stage 1 lung adenocarcinoma, expressing wild type K-Ras and a dominant negative p53. The two other cancer cell lines, expressing mutant K-Ras and wild type p53 (A549) or wild type K-Ras and mutant p53 (H1299), displayed intermediate but almost identical levels of up-regulation. In these experiments, we also observed a correlation with the mesenchymal phenotype, displaying higher expression of PRMT1Δarm in cell lines with a higher percentage of mesenchymal cells. To recapitulate the up-regulation of PRMT1Δarm in an experimental system, we expressed Snail1, one of the master regulators of EMT (see, e.g. Zhou et al., 2004), in HEK293 cells that can be driven towards a mesenchymal-like phenotype (Kashyap et al., 2013; Wu et al., 2009). Indeed, concomitant with a change in cellular morphology and a down-regulation of E-Cadherin indicative of EMT, we find a significant up-regulation of PRMT1Δarm (Fig. 4). Together, these results point to a link between PRMT1Δarm expression and cancer cell phenotype in lung cells.

The particular way in which PRMT1Δarm is linked to cancer remains presently unknown. PRMT1Δarm could promote the transformation of normal cells to cancer cells in a similar way to variant v2 (Baldwin et al., 2012), and thus being expressed in higher levels in cancer cells. However, in the light of our results, it is more likely that PRMT1Δarm is involved in a defense mechanism against neoplastic transformation. In this scenario, PRMT1Δarm could protect the cells from cancer progression by negatively regulating PRMT1-mediated methylation, which has been linked to cancer promotion in several studies. This could explain our observation that cells forced to express higher amounts of PRMT1Δarm have an increased rate of cell death, as they initiate apoptosis in order to avoid their transition to cancer cells, perhaps through EMT. This is particularly plausible, as PRMT1 has recently been found to act as an important regulator of EMT through PRMT1-mediated methylation and activation of the E-cadherin repressor, Twist1, in non-small cell lung cancer (Avasarala et al., 2015) as well as in breast cancer (Gao et al., 2016). Through promoting cell migration and invasion, overexpression of PRMT1 may promote development of metastases (Gao et al., 2016). The Snail1-induced increase of PRMT1Δarm may therefore represent an attempt to reduce PRMT1 activity, and in this way counteract PRMT1-induced activation of the EMT program in affected cells.
Future research in our laboratory will now extend these experiments, with the aim to better understand the properties of PRMT1Δarm and its role in regulating PRMT1 activity, particularly in cancer development and progression.

Acknowledgments

We thank Jocelyn Côté, Ottawa, for the generous gift of the vector expressing PRMT1v1:GFP, and Dmitriy Chudakov, Moscow, for the generous gift of the vector expressing fibrillarin:FusionRed; the vector for flag-Snail 6SA was created by Mien-Chie Hung and supplied through Addgene (Addgene plasmid # 16221). This work was partially supported by Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: THALES "CancerTFs" (MIS 379435) and THALES "IDIPRO" (MIS 379440), and by the EU COST Action CM1406 "Epigenetic Chemical Biology (EPICHEM)".

References

Avasarala, S., Van Scoyk, M., Karuppusamy Rathinam, M.K., Zerayesus, S., Zhao, X., Zhang, W., Pergande, M.R., Borgia, J.A., DeGregori, J., Port, J.D.*, et al.* (2015). PRMT1 Is a Novel Regulator of Epithelial-Mesenchymal-Transition in Non-small Cell Lung Cancer. J Biol Chem *290*, 13479-13489.

Baldwin, R.M., Morettin, A., Paris, G., Goulet, I., and Cote, J. (2012). Alternatively spliced protein arginine methyltransferase 1 isoform PRMT1v2 promotes the survival and invasiveness of breast cancer cells. Cell Cycle *11*, 4597-4612.

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

Cheng, X., and Roberts, R.J. (2001). AdoMet-dependent methylation, DNA methyltransferases and base flipping. Nucleic Acids Res *29*, 3784-3795.

Fackelmayer, F.O. (2005). A stable proteinaceous structure in the territory of inactive X chromosomes. J Biol Chem *280*, 1720-1723.

Gao, Y., Zhao, Y., Zhang, J., Lu, Y., Liu, X., Geng, P., Huang, B., Zhang, Y., and Lu, J. (2016). The dual function of PRMT1 in modulating epithelial-mesenchymal transition and cellular senescence in breast cancer cells through regulation of ZEB1. Sci Rep *6*, 19874.

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.

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

Han, X., Li, R., Zhang, W., Yang, X., Wheeler, C.G., Friedman, G.K., Province, P., Ding, Q., You, Z., Fathallah-Shaykh, H.M.*, et al.* (2014). Expression of PRMT5 correlates with malignant grade in gliomas and plays a pivotal role in tumor growth in vitro. J Neurooncol *118*, 61-72.

Helbig, R., and Fackelmayer, F.O. (2003). Scaffold attachment factor A (SAF-A) is concentrated in inactive X chromosome territories through its RGG domain. Chromosoma *112*, 173-182.

Herrmann, F., Bossert, M., Schwander, A., Akgun, E., and Fackelmayer, F.O. (2004). Arginine methylation of scaffold attachment factor A by heterogeneous nuclear ribonucleoprotein particle-associated PRMT1. J Biol Chem *279*, 48774-48779.

Herrmann, F., and Fackelmayer, F.O. (2009). Nucleo-cytoplasmic shuttling of protein arginine methyltransferase 1 (PRMT1) requires enzymatic activity. Genes Cells *14*, 309-317.

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*, 38005-38010.

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. Journal of Cell Science *122*, 667-677.

Kashyap, A., Zimmerman, T., Ergul, N., Bosserhoff, A., Hartman, U., Alla, V., Bataille, F., Galle, P.R., Strand, S., and Strand, D. (2013). The human Lgl polarity gene, Hugl-2, induces MET and suppresses Snail tumorigenesis. Oncogene *32*, 1396-1407.

Latonen, L., Moore, H.M., Bai, B., Jaamaa, S., and Laiho, M. (2011). Proteasome inhibitors induce nucleolar aggregation of proteasome target proteins and polyadenylated RNA by altering ubiquitin availability. Oncogene *30*, 790-805.
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. Mol Endocrinol *21*, 1381-1393.
Lee, J., and Bedford, M.T. (2002). PABP1 identified as an arginine methyltransferase substrate using high-density protein arrays. EMBO Rep *3*, 268-273.
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*, 15034-15044.
Mathioudaki, K., Scorilas, A., Ardavanis, A., Lymberi, P., Tsiambas, E., Devetzi, M., Apostolaki, A., and Talieri, M. (2011). Clinical evaluation of PRMT1 gene expression in breast cancer. Tumour Biol *32*, 575-582.
Mearini, G., and Fackelmayer, F.O. (2006). Local chromatin mobility is independent of transcriptional activity. Cell Cycle *5*, 1989-1995.
Pawlak, M.R., Scherer, C.A., Chen, J., Roshon, M.J., and Ruley, H.E. (2000). Arginine N-methyltransferase 1 is required for early postimplantation mouse development, but cells deficient in the enzyme are viable. Mol Cell Biol *20*, 4859-4869.
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*, 349-359.
Tang, J., Kao, P.N., and Herschman, H.R. (2000). Protein-arginine methyltransferase I, the predominant protein-arginine methyltransferase in cells, interacts with and is regulated by interleukin enhancer-binding factor 3. J Biol Chem *275*, 19866-19876.
Vartholomatos, G., Alexiou, G.A., Batistatou, A., Lykoudis, E., Voulgaris, S., and Kyritsis, A.P. (2015). Rapid cell cycle analysis for intraoperative diagnosis of brain tumors. Brain Tumor Pathol *32*, 151-152.
Wang, S., Tan, X., Yang, B., Yin, B., Yuan, J., Qiang, B., and Peng, X. (2012). The role of protein arginine-methyltransferase 1 in gliomagenesis. BMB Rep *45*, 470-475.
Wessel, D., and Flügge, U.I. (1984). A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. Anal Biochem *138*, 141-143.
Wu, Y., Deng, J., Rychahou, P.G., Qiu, S., Evers, B.M., and Zhou, B.P. (2009). Stabilization of snail by NF-kappaB is required for inflammation-induced cell migration and invasion. Cancer Cell *15*, 416-428.
Yan, F., Alinari, L., Lustberg, M.E., Martin, L.K., Cordero-Nieves, H.M., Banasavadi-Siddegowda, Y., Virk, S., Barnholtz-Sloan, J., Bell, E.H., Wojton, J*., et al.* (2014). Genetic validation of the protein arginine methyltransferase PRMT5 as a candidate therapeutic target in glioblastoma. Cancer Res *74*, 1752-1765.
Yang, Y., and Bedford, M.T. (2013). Protein arginine methyltransferases and cancer. Nat Rev Cancer *13*, 37-50.
Zhang, X., and Cheng, X. (2003). Structure of the predominant protein arginine methyltransferase PRMT1 and analysis of its binding to substrate peptides. Structure *11*, 509-520.

Zhang, X., Zhou, L., and Cheng, X. (2000). Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. Embo J *19*, 3509-3519.
Zhou, B.P., Deng, J., Xia, W., Xu, J., Li, Y.M., Gunduz, M., and Hung, M.C. (2004). Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition. Nat Cell Biol *6*, 931-940.

Legends and Figures

Fig. 1. (A) Exon organization of the Prmt1 gene, showing the known splicing variants v1-v7, and the novel variant PRMT1Δarm. (B, C) Structural context of the new variant. (B) Structure of the dimer of canonical PRMT1, showing in space fill mode the tip of the dimerization arm (blue) and the dimer interface on the AdoMet binding domain of the bound second monomer (red). (C) Structure of a PRMT1 monomer, with the dimerization arm highlighted in colors (upper panel), and the PRMT1Δarm deletion variant with the missing part in transparent color (lower panel). Note that the flexible loop VDIYT (blue arrow) at the beginning of exon 10 is of suitable size to fold over and connect the two amino acids (red) amino-terminal (L167, end of exon 7) and carboxy-terminal (E236, start of exon 10) of the deleted arm, without causing major structural changes in the protein.

Fig. 2. (A) Expression of PRMT1Δarm on RNA level in HEK293 and HeLa cells. Total RNA was prepared from untreated HEK293 and HeLa cells and reverse-transcribed, followed by 20, 25 and 30 cycles of PCR with a primer set specific for PRMT1Δarm. (B) RT-PCR with isoform-specific forward primers and the PRMT1Δarm specific reverse primer (P1Δarmrev) was used to identify the exons present upstream of exon 7 (upper panel), or with isoform-specific forward primer and a general PRMT1pan reverse primer (lower panel). Note that a strong amplification product is only obtained with the forward primer PF ("v1…v3"). (C) Dimerization-arm deletion constructs were cloned for variant 1 and variant 2, and used for PCR with the primer pair PF and P1Δarmrev. The products were compared with the product from RT-PCR from HeLa cDNA. Note that the RT-PCR product is identical with that of the v1 construct.

Fig. 3. (A) PRMT1Δarm expression in two normal diploid lung fibroblast strains, HFL-1 and MRC-5, and a panel of four lung cancer lines (A549, H1299, H1437, H1792) with different status of p53 and K-Ras. Semi-quantitative PCR for 30 cycles was performed, and the products were analyzed by agarose gel electrophoresis, together with products from RT-PCR with the PRMT1pan primer pair (26 cycles) that amplifies all PRMT1 isoforms. GAPDH was used as a loading control. (B) Quantitative RT-PCR was performed on RNA from the same cells. The data was normalized to GAPDH and analyzed by the ΔΔCt method. The means of PRMT1Δarm expression relative to MRC-5 cells are given, together with their SEM. (one star: $p<0.05$; three stars: $p<0.001$, two-tailed t-tests)

Fig. 4. Expression of PRMT1Δarm is increased in cells expressing the EMT inducer Snail1. (A) HEK293 cells transfected with a Flag:Snail-6SA expression vector were analyzed for cell morphology by staining of live cells with $DiOC_6$, and for the expression of Snail and E-Cadherin by RT-PCR, using GAPDH as a normalization control. (B) Expression and correct nuclear localization of Snail were verified by immunofluorescence with anti-Flag antibody (red) and Sytox Green as a DNA counterstain (green). A representative nucleus is shown; scale bar: 10µm. (C) Total PRMT1 and PRMT1Δarm are up-regulated upon Snail expression in HEK293 cells stably transfected with Flag:Snail-6SA expression vector either with Lipofectamine

2000 (LF) or Polyethylenimine (PEI). The expression of PRMT1 and PRMT1Δarm was measured with qPCR with primer pairs that amplify all seven amino-terminal variants of PRMT1 or PRMT1Δarm, respectively. The data was normalized to GAPDH and analyzed by the ΔΔCt method. The mean of individual experiments is presented graphically together with SEM for each case. The stars indicate statistical significance (three stars: $p<0.01$, four stars: $p<0.005$, two-tailed t-tests).

Fig. 5. Expression of PRMT1Δarm leads to a G1 cell cycle arrest. HEK293 cells were transiently transfected with constructs encoding PRMT1v1:YFP or PRMT1Δarm:YFP, respectively. 24h after transfection, cells were harvested and analyzed by fluorescence-activated cell sorting (FACS). (A) Expression levels of the two constructs. Untransfected cells and cells transfected with empty YFP vector were used as references. Insert: western blot with extracts from cells expressing the two variants, using anti-GFP antibodies (upper panel), or a part of an identical gel stained with Coomassie blue as loading control (lower panel). (B) Cells cycle analysis of transfected cells by staining with propidium iodide. Note that cells expressing PRMT1v1 have a cell cycle profile identical with untransfected control cells, whereas PRMT1Δarm lead to an arrest in G1 and increased number of cells with sub-G1 amount of DNA. M1= G1 phase, M2= S, M3= G2/M, M4=sub-G1.

Fig. 6. Localization of PRMT1v1 and PRMT1Δarm are differently affected by inhibition of the proteasome with MG132. (A) Cells transiently expressing PRMT1v1 or PRMT1Δarm, respectively, were treated with 5 μM MG132 for 8 or 16 hours, respectively, before total cell extract was prepared and analyzed by western blotting with anti-GFP antibody (upper panel), or an identical gel stained with Coomassie blue as loading control (lower panel). (B) Cells transfected as in (A) and treated with MG132 for 8 h were investigated by confocal microscopy. Note how PRMT1Δarm, but not PRMT1v1, relocalizes from predominantly cytoplasmic localization to nucleoli in 97% of treated cells. Upper panel: typical microscopic fields with the expressed constructs in green, and a nuclear counterstain with To-Pro 3 in red; lower panel: statistical analysis of the four conditions. Each black dot represents one percent of cells with nucleolar localization; grey dots represent cells with non-nucleolar localization. $p<0.0001$, two-tailed t-test. (C) PRMT1Δarm accumulates in nucleolar aggresomes. Live cells stably expressing fibrillarin:FusionRed as a marker for nucleoli were transiently transfected with the two PRMT1 constructs, and then treated with MG132 as in (B).

Fig. 7. PRMT1Δarm does not multimerize with endogenous PRMT1, is catalytically inactive, and accumulates in the nucleus through stable interaction with nuclear factors. (A) Extracts from cells expressing PRMT1v1:YPF or PRMT1Δarm:YFP were subjected to immunoprecipitation with anti-GFP antibodies, and the precipitated proteins were analyzed by western blot with antibodies to PRMT1. (B) PRMT1 constructs were immunoprecipitated as in (A), and the immunoprecipitated native proteins were incubated with radioactive AdoMet and hypomethylated total cell extract as a substrate. Reaction products were separated by SDS PAGE and the radioactively labeled products were visualized by fluorography. (C) Identification of proteins that are preferentially associated with PRMT1v1 or PRMT1Δarm,

respectively. Total cell extracts were used for immunoprecipitation with monoclonal antibodies against GFP. After extensive washing, proteins were resolved by SDS-PAGE and visualized by silver staining. Note that some proteins differentially interact with PRMT1 and PRMT1Δarm, denoted by asterisks beside the lanes.
(D) Typical microscopic fields of cells expressing PRMT1v1 or PRMT1Δarm, respectively, under steady state conditions in live cells (left panel). For quantification, mean pixel intensities in several regions of interest in the cytoplasm and the nucleus were measured, and expressed as a ratio of nuclear versus cytoplasmic fluorescence; n≈110 cells per construct (right panel). Note that catalytically inactive PRMT1 accumulates in the nucleus approximately threefold in comparison to wild-type PRMT1; shown is the median with 25 and 75 percentile (box) and 5-95 percentile (whisker). $p<0.0001$, two-tailed t-test.

Fig. 8. HEK293 cells stably expressing PRMT1v1:YFP or PRMT1Δarm:YFP, respectively, were analyzed by fluorescence recovery after photobleaching (FRAP). Rectangular regions of interest spanning both the cytoplasm and the nucleus of cells were photobleached, and the recovery of fluorescence was recorded for 50sec. Cytoplasmic and nuclear regions were then analyzed separately for 25 cells per construct. Shown are the means (and SEM) of the measurements. Dotted grey line: FRAP recovery curve of a catalytically dead mutant of PRMT1v2, shown for comparison.

Supplementary Fig. S1. Schematic representation of the genomic structure of the prmt1 gene, and the individual exon composition of the mature mRNAs of all eight isoforms. The positioning of all primers used for PCR analysis is shown by arrows.

Supplementary Fig. S2. Comparison of PRMT1Δarm expression between lung fibroblasts, lung cancer lines, and HEK293 cells, as determined by semi-quantitative RT-PCR. GAPDH was used for normalization

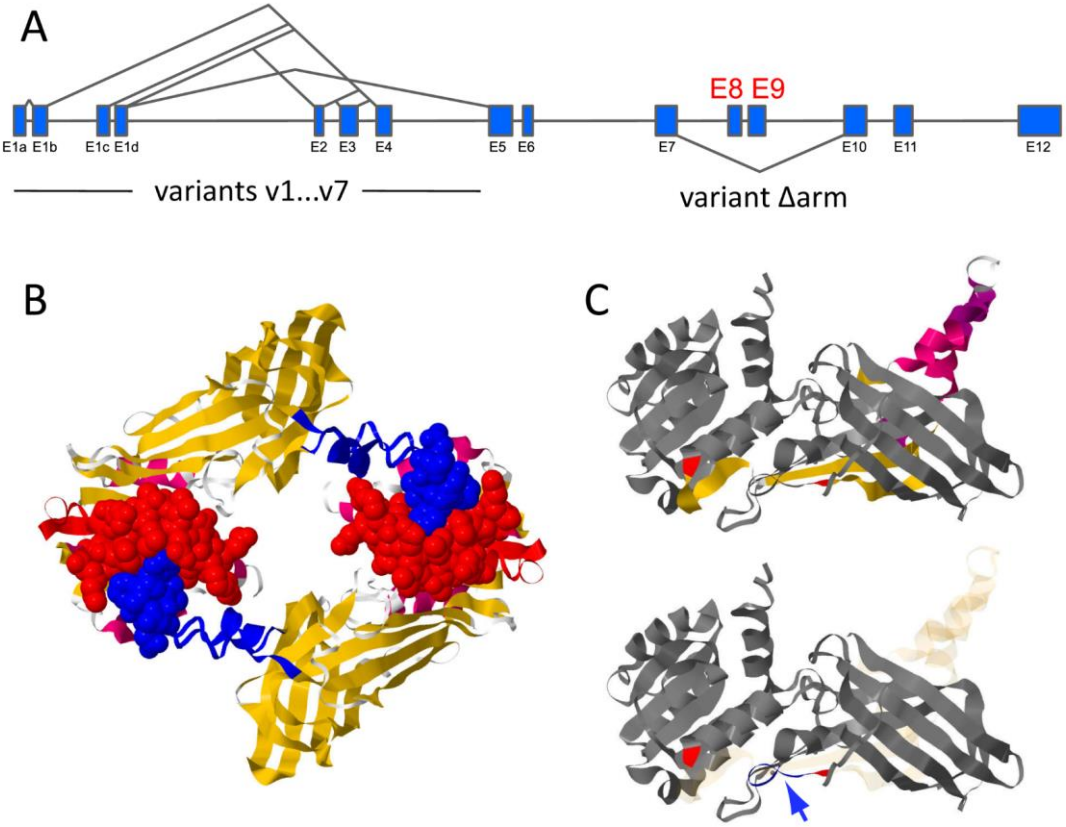


Figure 1

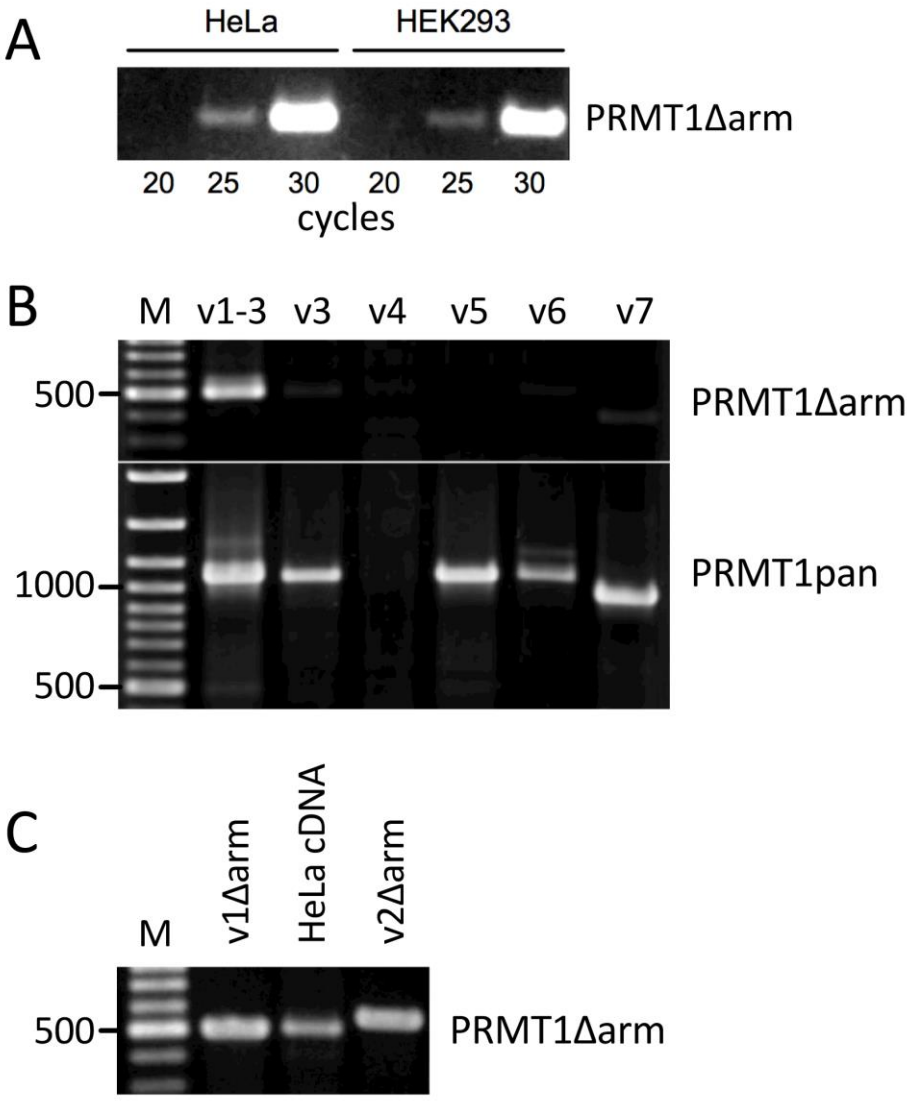


Figure 2

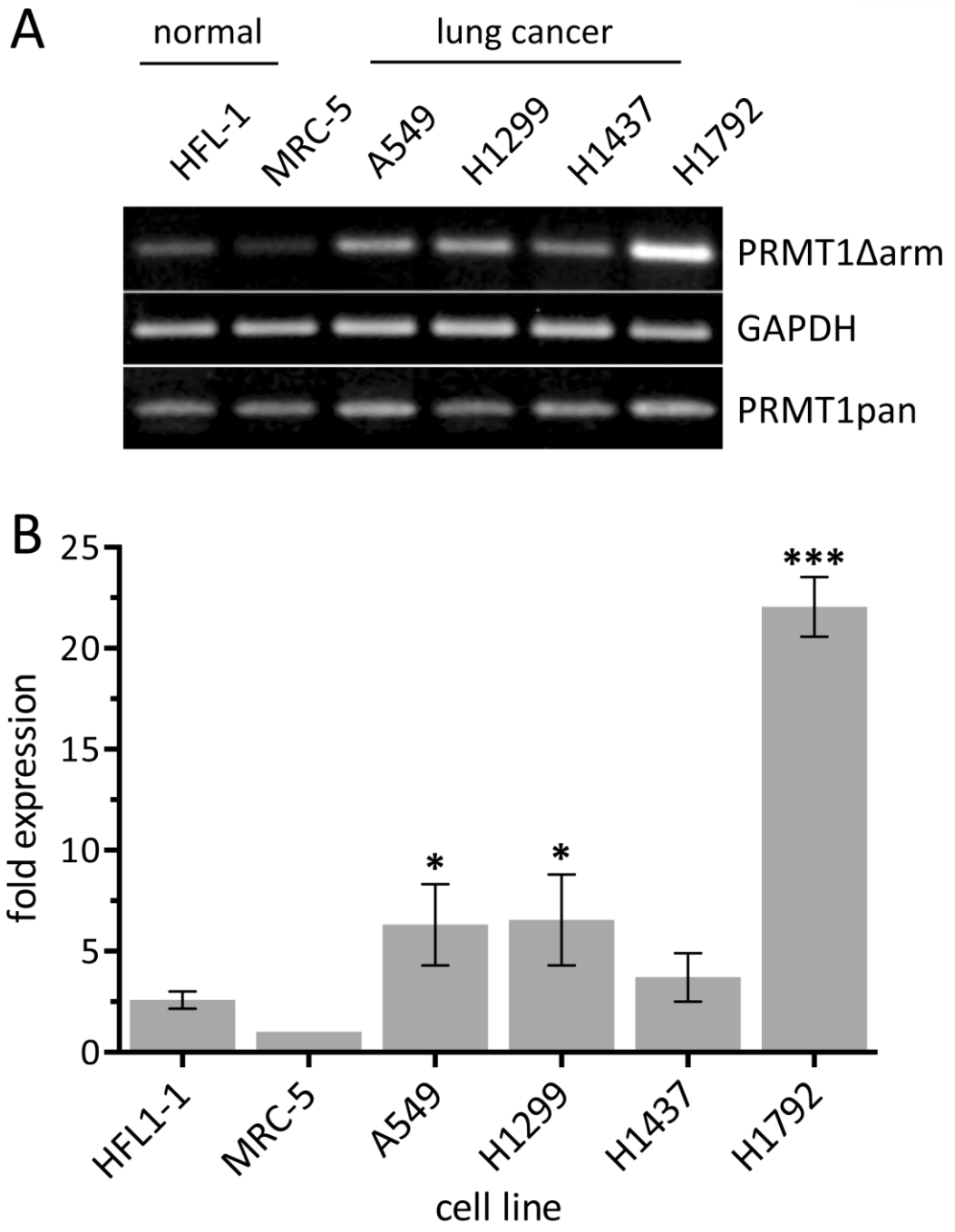


Figure 3

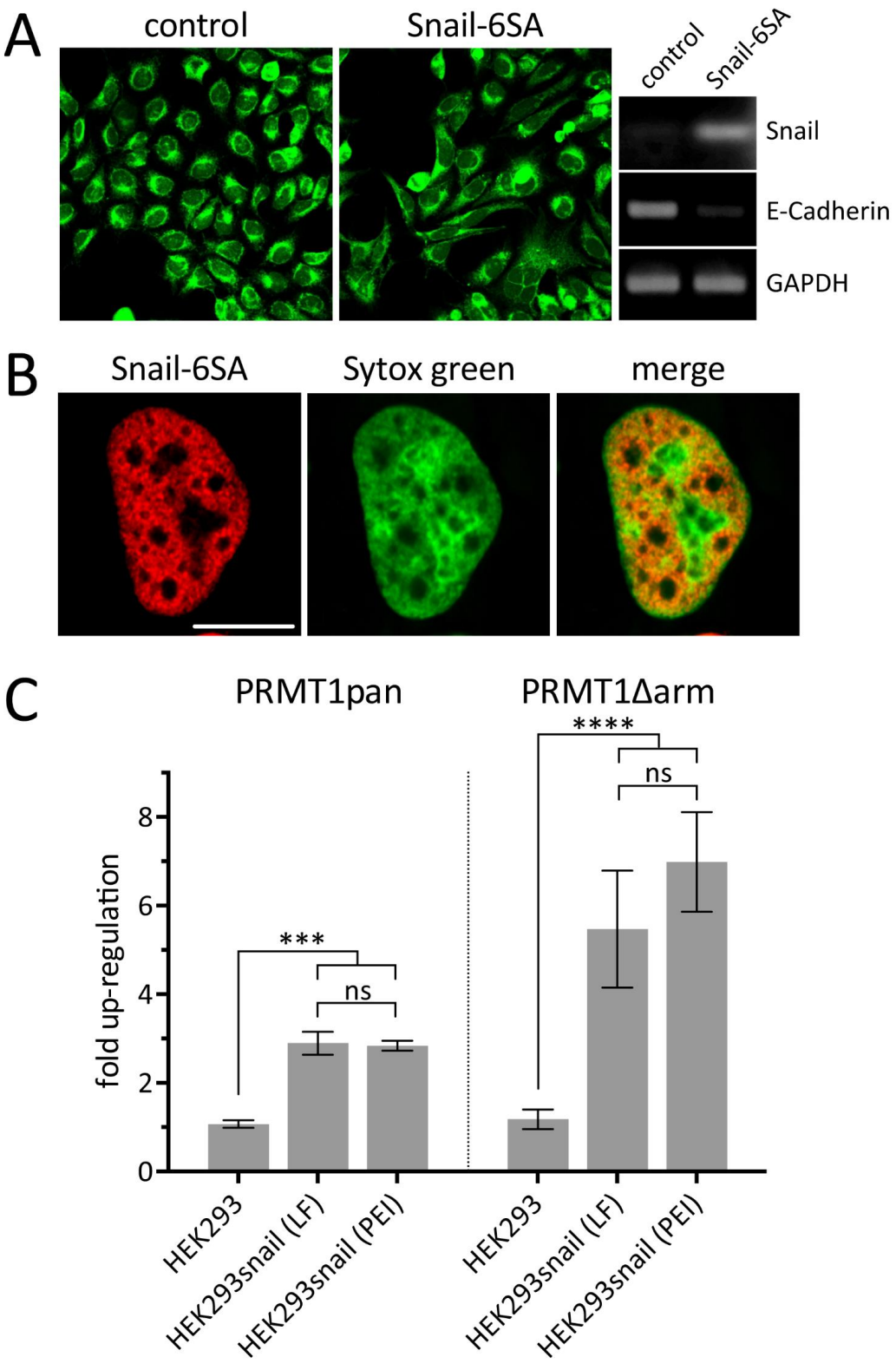


Figure 4

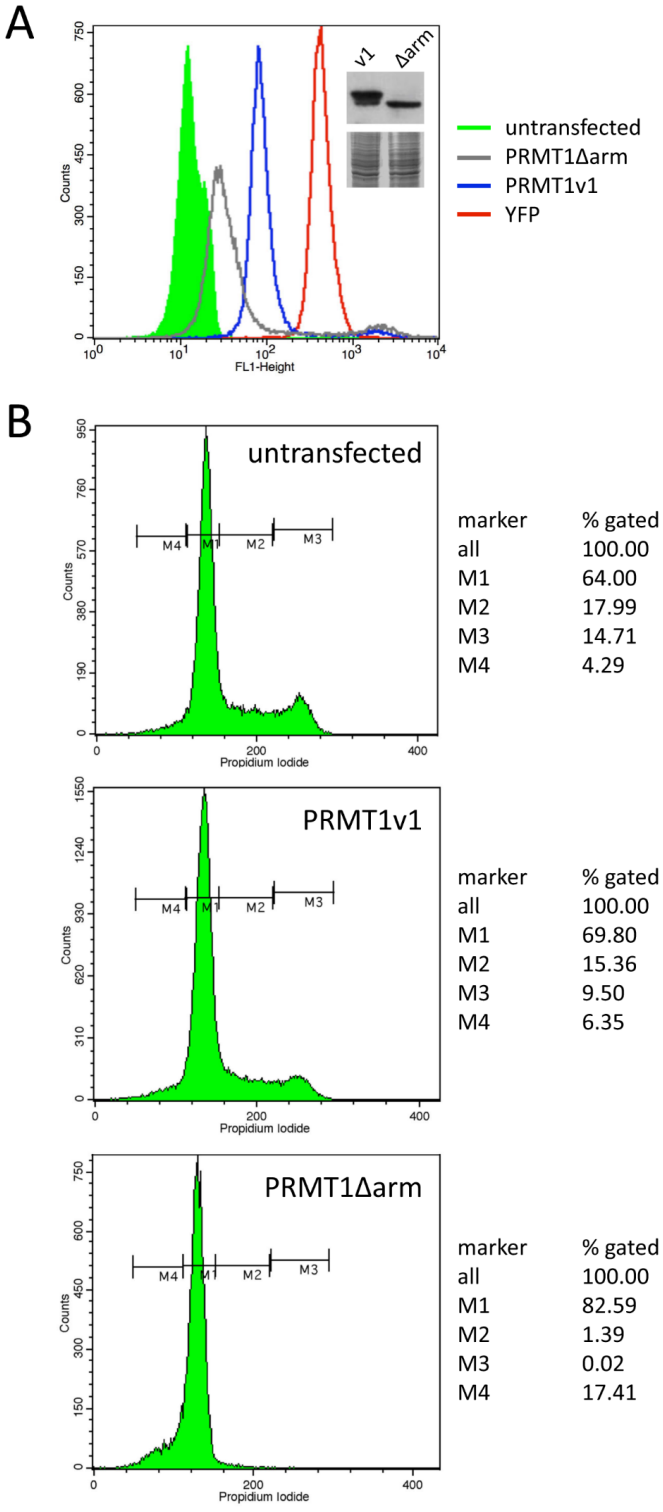


Figure 5

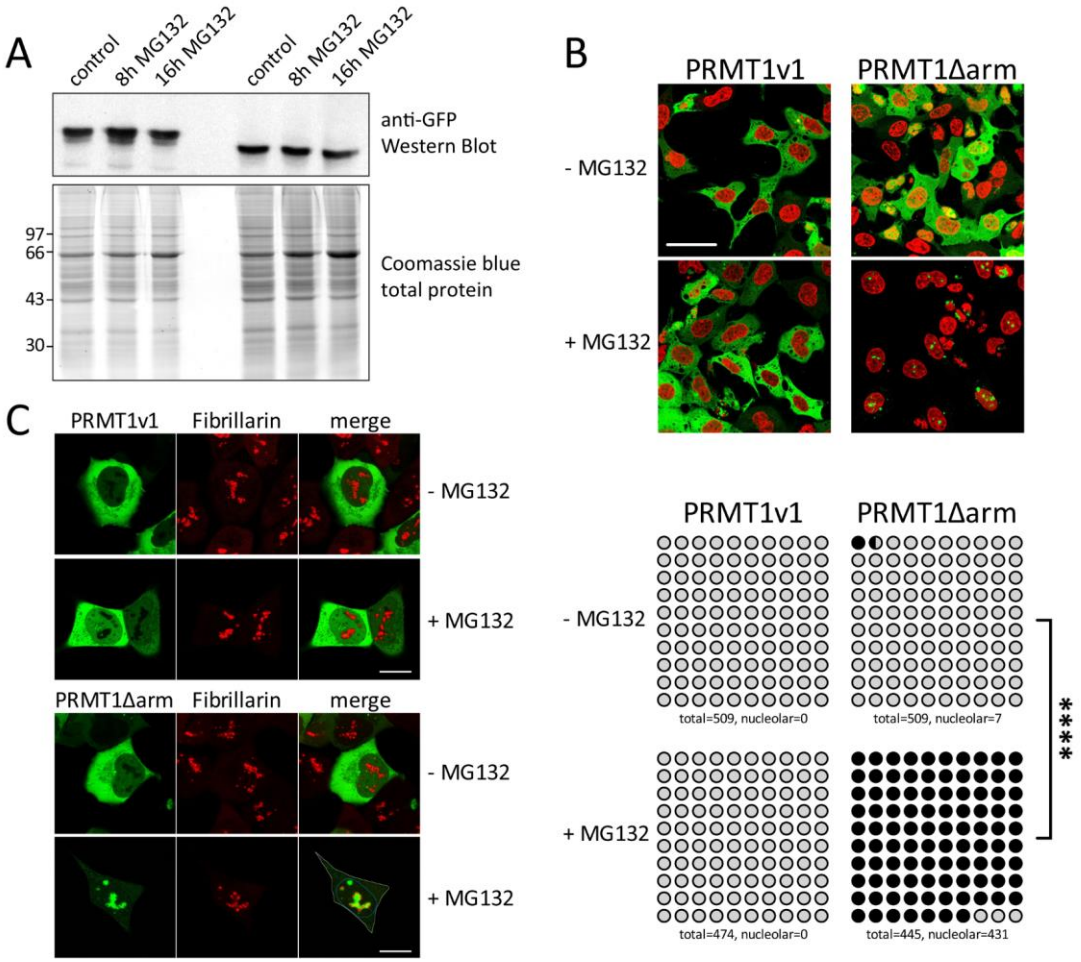


Figure 6

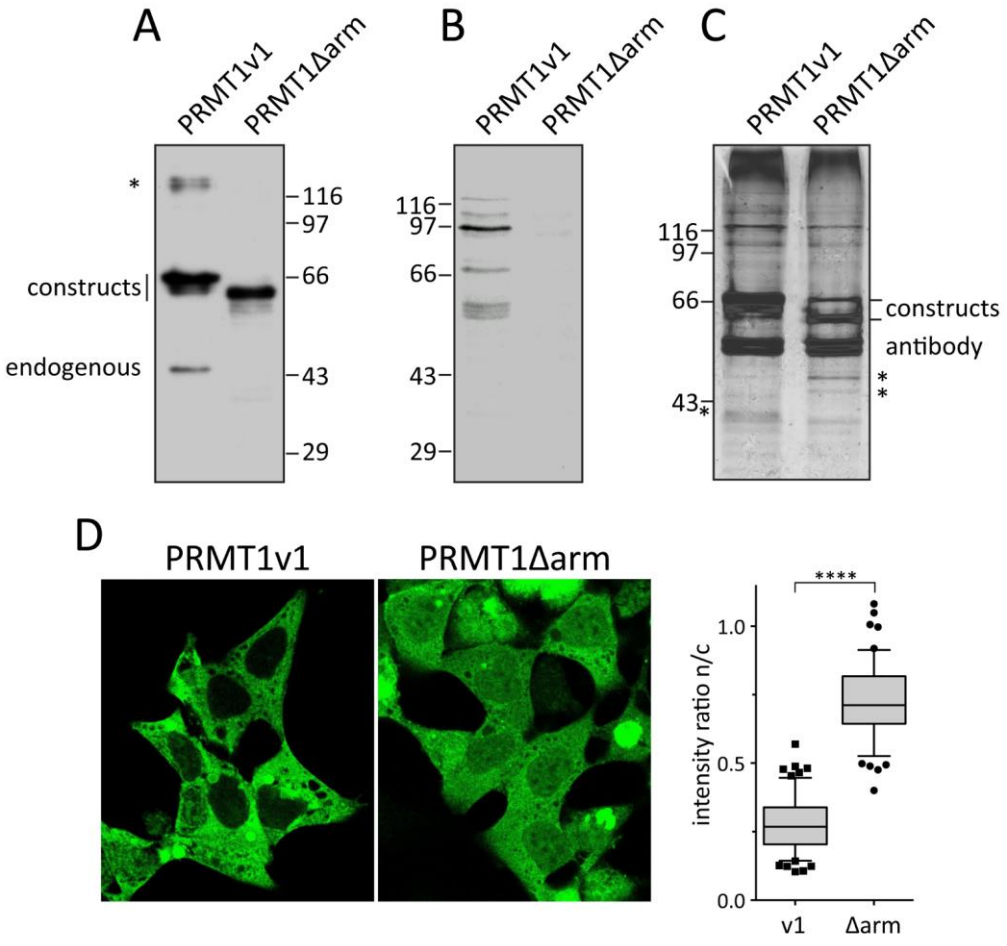


Figure 7

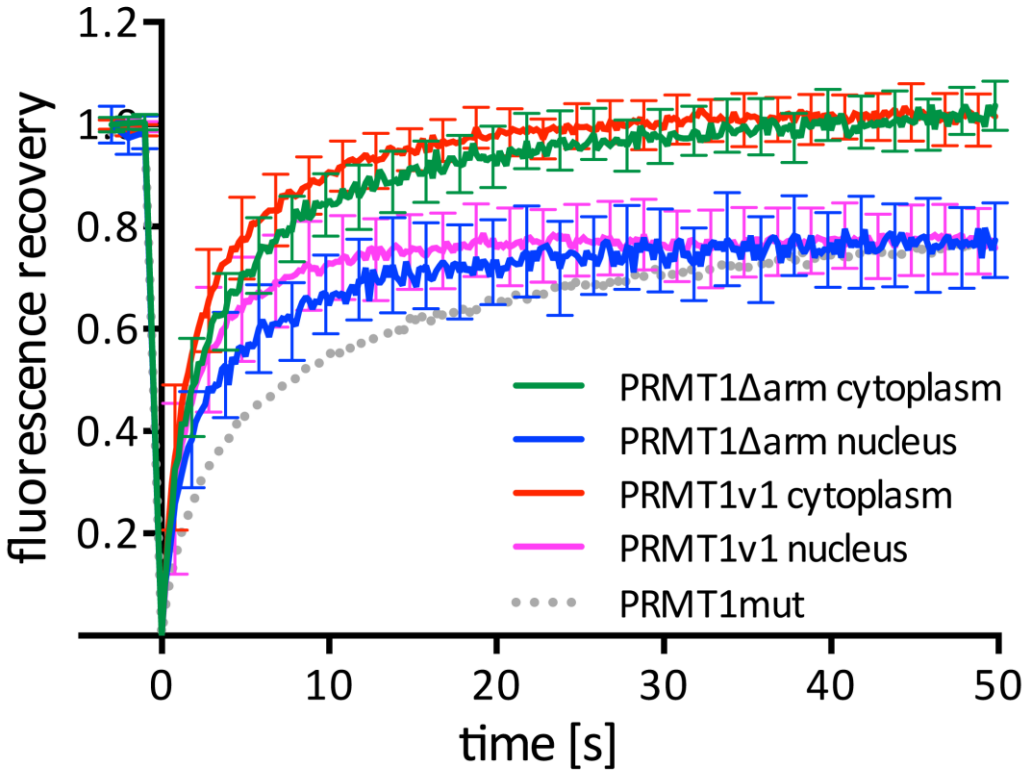
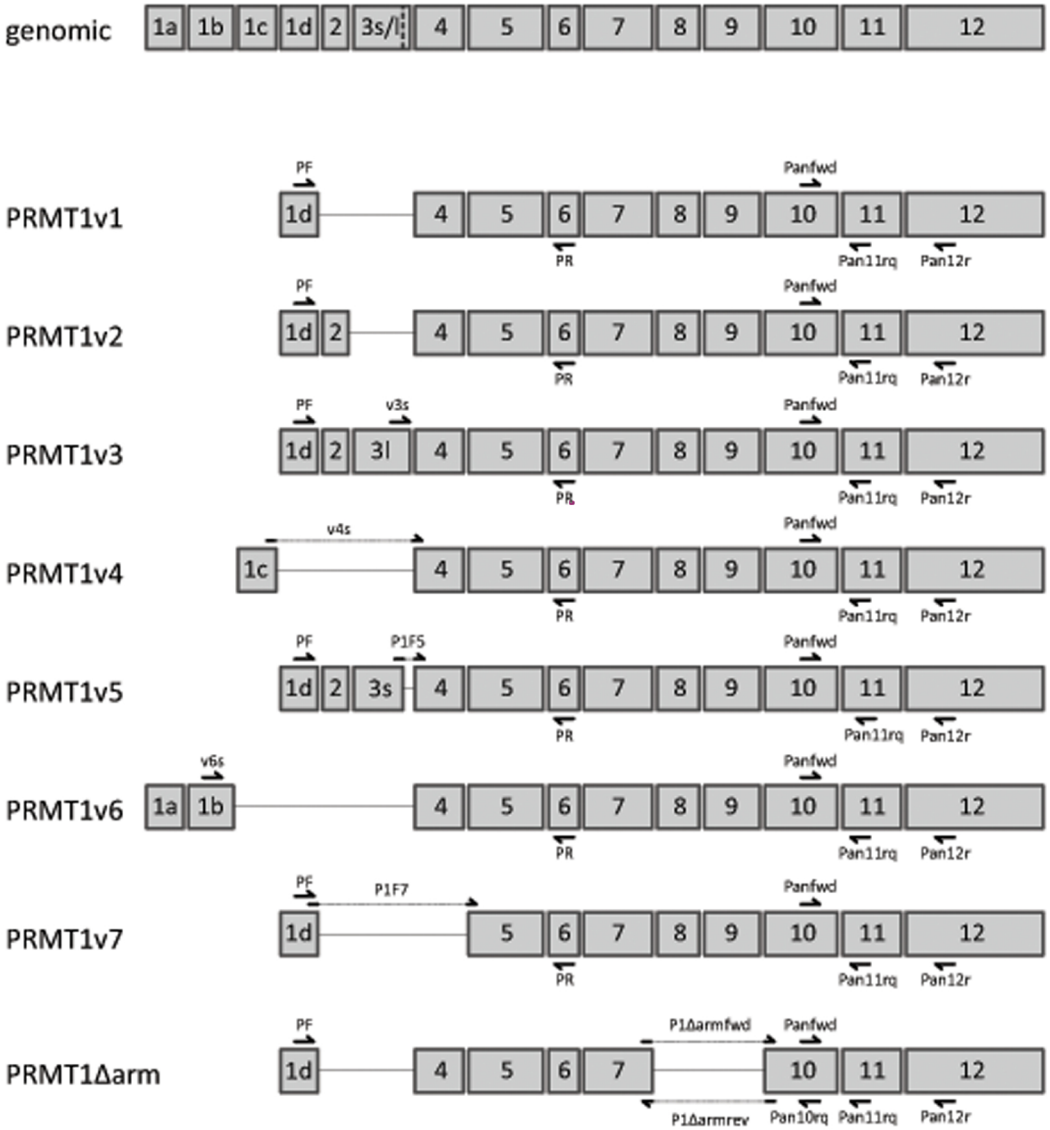
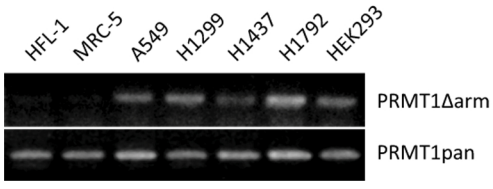


Figure 8



Supplementary Fig. S1



Supplementary Fig. S2

Tables

Supplementary TABLE I. DNA primer sequences and conditions for RT-PCR

primer name	sequence (5'->3')	use*	annealing	elongation
P1Δarmrev	GGTATAGATGTCCACCTCCAGCC	semi	65°C	30sec, 72°C
P1Δarmfwd	CGGGACAAGTGGCTGGAGGTG	semi&q	67°C	20sec, 67°C
Pan10rq	CAGGGCGTGCACGTAGTCAT	q	60°C	60sec, 60°C
Panfwd	ATGACTACGTGCACGCCCTG	semi&q	60°C	30sec, 72°C
Pan11rq	CGTCTGCTTCCAGTGCGTGTA	q	60°C	60sec, 60°C
Pan12r	CGAAACCGCCTAGGAACGCT	semi	60°C	30sec, 72°C
v3s	AGATGGTAGGCGTGGCTGA	semi	65°C	30sec, 72°C
v4s	GCGGGGTCGCGGTGTCCT	semi	65°C	30sec, 72°C
v6s	AGAAGCTGACCAGACAAAGAGA	semi	65°C	30sec, 72°C
PF (1)	GAGGCCGCGAACTGCATCAT	semi	65°C	30sec, 72°C
PR (1)	TGGCTTTGACGATCTTCACC	semi	65°C	30sec, 72°C
P1F5 (2)	ACTGGAGAGATGGTGTCCTGTGG	semi	65°C	30sec, 72°C
P1F7 (2)	TGCATCATGGAGGAGATGCTGAAGG	semi	65°C	30sec, 72°C
Snail_fwd (3)	ACCACTATGCCGCGCTCTT	semi&q	60°C	60sec, 60°C
Snail_rev (3)	GGTCGTAGGGCTGCTGGAA	semi&q	60°C	60sec, 60°C
ecadherin_fwd (3)	GTCAGTTCAGACTCCAGCCC	semi&q	60°C	60sec, 60°C
ecadherin_rev (3)	AAATTCACTCTGCCCAGGACG	semi&q	60°C	60sec, 60°C
GAPDH_fwd (4)	AGCCACATCGCTCAGAACAC	semi	57°C	45sec, 72°C
GAPDH_rev (4)	GAGGCATTGCTGATGATCTTG	semi	57°C	45sec, 72°C
GAPDH_RT_fwd (5)	TGCACCACCAACTGCTTAGC	q	60°C	60sec, 60°C
GAPDH_RT_rev (5)	GGCATGGACTGTGGTCATGAG	q	60°C	60sec, 60°C
18S_RT_fwd (6)	CTCAACACGGGAAACCTCAC	semi&q	60°C	60sec, 60°C
18S_RT_rev (6)	CGCTCCACCAACTAAGAACG	semi&q	60°C	60sec, 60°C
Alu_fwd (7)	CATGGTGAAACCCCGTCTCTA	semi&q	60°C	60sec, 60°C
Alu_rev (7)	GCCTCAGCCTCCCGAGTAG	semi&q	60°C	60sec, 60°C

* semi= semi-quantitative end point PCR; q= real-time qPCR

(1) Scorilas et al. (2000). Genomic organization, physical mapping, and expression analysis of the human protein arginine methyltransferase 1 gene. Biochem Biophys Res Commun 278, 349-359.
(2) Goulet et al. (2007). Alternative splicing yields protein arginine methyltransferase 1 isoforms with distinct activity, substrate specificity, and subcellular localization. J Biol Chem 282, 33009-33021.
(3) Medici et al. (2008). Snail and Slug Promote Epithelial-Mesenchymal Transition through beta-Catenin–T-Cell Factor-4-dependent Expression of Transforming Growth Factor-beta3. Mol. Biol. Cell 19: 4875–4887
(4) Lehmann et al. (2002). Pseudogene-free amplification of human GAPDH cDNA. Biotechniques 33(4):766, 769-70.
(5) Mavrakis et al. (2007). Arkadia enhances Nodal/TGF-b signaling by coupling phospho-Smad2/3 activity and turnover. PLoS Biol 5(3): e67
(6) Pop et al. (2016). Normalization of gene expression measurement of tissue samples obtained by transurethral resection of bladder tumors. Onco Targets Ther. 9: 3369–3380
(7) Rihani et al. (2013). Effective Alu Repeat Based RT-Qpcr Normalization in Cancer Cell Perturbation Experiments. PloS One 8 (8): e71776

Supplementary TABLE II. Cell lines used in this study

line	characteristics	morphology	K-Ras	p53
HFL-1	Normal diploid human lung fibroblasts, male	fibroblast	wt	wt
MRC-5	Normal diploid human lung fibroblasts, male	fibroblast	wt	wt
A549	58 years male, Caucasian, stage 2, NSCLC, primary tumor	mostly epithelial, few mesenchymal	G12S	wt
H1299	43 years male, Caucasian, NSCLC, metastatic site	epithelial	wt	null
H1437	60 years male, Caucasian, stage 1 adenocarcinoma; NSCLC	epithelial	wt	R267>P
H1792	50 years male, Caucasian, stage 4 NSCLC metastatic site	epithelial-mesenchymal	G12C	c.672+1G>A (almost null)