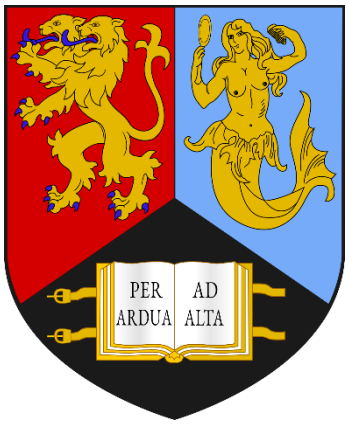


Detecting and Characterising the Protein Arginine Methyltransferase 1 Substrate Profile in Breast Cancer

By James Jarrold

A thesis presented to the College of Medical and Dental Sciences, University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

Institute of Cancer and Genomic Sciences

College of Medical and Dental Sciences

University of Birmingham

December 2020

Abstract

PRMT1 is a known contributor to breast cancer through methylation of histone and non-histone substrates, but PRMT1-mediated transcriptional coactivation by deposition of H4R3me2a remains understudied. As a result, potential contributions to malignant gene expression via H4R3me2a deposition remain largely unknown. To confront this issue, we made efforts to optimise bio-orthogonal profiling by utilising a PRMT1-Y29F/M38G-Pob-SAM pairing to label histone H4 with a surrogate alkynyl moiety. Although further optimisation is required, data gathered so far suggest that this technology may prove a viable option for genome-wide analysis of H4R3me2a in breast cancer, enabling global comparisons of this epigenetic mark with non-transformed mammary epithelial cells.

In addition, the biochemical consequences of PRMT1 interaction with histone code reader SPIN1 were characterised, after SPIN1 was identified as a breast cancer-enriched interactor of PRMT1 using SILAC quantitative proteomics. SPIN1 was found to positively regulate *PRMT1* transcription, making it challenging to validate PRMT1-dependent methylation of SPIN1 *in vitro*. In contrast, recombinant PRMT1 methylated immunoprecipitated SPIN1, possibly at R117 within its P loop, suggesting a role for R117 in modulating phosphate binding. Supporting this, PRMT1-mediated SPIN1 methylation in a cell-free context was enhanced in phosphate-free conditions. Finally, expression of SPIN1-R117K promoted S-phase cell cycle accumulation, implying that R117 methylation is biologically significant. Ultimately, these findings may expose a SPIN1-PRMT1 axis that can be therapeutically targeted in breast cancer.

Dedication

In loving memory of Fiona Patricia Jarrold.

All that I have become that may be considered good, I owe to you.

We will never stop missing you.

Acknowledgements

I would like to thank my supervisor Dr. Clare Davies for the opportunity to prove myself as a scientist and reach this milestone in my lifetime and career. I have learned such an enormous amount during my time as one of her students, and have grown so much under her supervision. I would also like to thank all members of the Davies group, both past and present. In particular Dr. Agnieszka Zielinska for her tutelage and patience, and Nicole McFadzean and Matthew Gillespie for their morale-boosting friendship during my stressful post-lockdown race to gather the final pieces of data for this thesis!

Additional thanks go to Dr. Krystian Ubych and Dr. Sati Jujh for their help, friendship and sense of humour, and all of the many peers with whom I have shared this experience over these last four years.

Finally, I would like to thank my whole family, particularly my siblings David, Max and Joanna, for their constant love and support and provision of second home(s) when I needed a day or two away.

Contents

List of Figures

List of Tables

Abbreviations

ADMA	Asymmetric dimethyl-arginine
AMKL	Acute megakaryocytic leukaemia
AML	Acute myeloid leukaemia
B-catenin	Beta-Catenin
BSA	Bovine serum albumin
cccDNA	covalently closed circular DNA
ChIP	Chromatin Immunoprecipitation
CliEn-seq	clickable chromatin enrichment with parallel DNA sequencing
co-IP	co-Immunoprecipitation
CuAAC	copper-catalysed azide–alkyne cycloaddition
DMA	Dimethyl-arginine
ESC	Embryonic stem cell
ESCC	oesophageal squamous cell carcinoma
H2AR3me2a	Histone H2A Aymmetric dimethyl-arginine 3
H3K4me3	Histone H3 trimethyl-lysine 4
H3R8me2a	Histone H3 Aymmetric dimethyl-arginine 8
H4R3me2a	Histone H4 Aymmetric dimethyl-arginine 3
HAT	Histone acetyltransferase
HR	Homologous repair
ID-2	Inhibitor Of DNA Binding 2
IP	Immunoprecipitation
K_{cat}	Enzymatic turnover
K_D	Dissociation constant
K_i	Binding affinity of a ligand to a molecule
K_m	Michaelis constant
MAT	methionine adenosyl transferase
MMA	Monomethyl-arginine
MS	Mass spectrometry
MSC	Muscle stem cell
MWCO	Molecular weight cutoff
NANOG	Homeobox Transcription Factor Nanog
OCT4	octamer-binding transcription factor 4
PMT	Protein Methyltransferase
Pob-met	S-4-propargyloxy-but-2-enyl homocysteine
Pob-SAM	4-propargyloxy-but-2-enyl-sam
PRMT	Protein Arginine Methyltransferase
PRMT1-XGX	PRMT1-S69A/G70A/T71A (Catalytic dead)
PTM	Post-translational modification
RG	Arginine-glycine motif
rpm	Revolutions per minute
SAM	S-adenosyl methionine
SDMA	Symmetric dimethyl-arginine
SILAC	Stable Isotope Labelling by Amino Acids in Cell culture
SOX2	Sex determining region Y-box 2

TAMRA	Tetramethylrhodamine-azide
TIAM1	T-Cell Lymphoma Invasion And Metastasis 1
Wnt	Wingless-Type MMTV Integration Site Family

Chapter 1: Introduction

1.1: Introduction to breast cancer

1.1.1: Origin and physiology of breast cancer

Breast cancer is an umbrella term describing a range of similar diseases that emerge as a result of uncontrolled proliferation of breast tissue. The origin of the disease is thought to rest with neoplastic transformation of a single cell into a breast cancer stem cell (BCSC). BCSCs comprise a small subset of cells within the tumour and are characterised by their propensity to not only initiate tumours, but also self-renew upon successive xenotransplantations[1].

Classically, the physiology of breast cancer follows a general pattern of progression in which the affected breast tissue transforms through increasingly malignant morphological increments, a model originally proposed by Wellings almost fifty years ago. Starting within the terminal duct lobular unit (TDLU), initial neoplastic development results in hyperplastic enlarged lobular units (HELUs), that are characterised by epithelial hyperplasia that increases their growth. HELU is subsequently capable of progressing to atypical ductal hyperplasia (ADH), in which features such as cell adhesion and polarity are altered, and acini (milk secretory structures) start to become distended. Up to this point, the cells within these aberrant growths are generally differentiated and show a histology more representative of their tissue of origin. However, progression to the next stage, ductal carcinoma *in situ* (DCIS) crosses this rubicon by greatly increasing histological diversity within the structure, as well as increasing the size of the tumour and allowing it to spread to other parts of the breast. Finally, the DCIS may further progress to invasive breast cancer (IBC), the most malignant form of

neoplastic growth in the breast, in which the tumour acquires the capability to invade distal tissues as a consequence of tumour cells undergoing epithelial to mesenchymal transition (EMT)[2].

1.1.2: Clinical subtypes of breast cancer

Although more sophisticated sub-categorisation of breast cancers has emerged in recent years (for example, pertaining to distinct transcriptomic signatures[3]), the 'classical' molecular subtyping system remains widely used. These molecular markers can range from presence of hormone receptors such as the estrogen and/or progesterone receptors, to amplification or overexpression of certain genes such as *HER2.* HER2-enriched breast cancers are characterised by an increased expression level of growth factor HER2, also known as Neu, a well-established oncogene. In the case of estrogen receptor-positive (ER^+) breast cancers, a further subdivision into Luminal A (high expression of ER and ER-responsive genes) and Luminal B (lower expression of ER and ER-influenced genes) is possible, with Luminal A breast cancers holding a generally more positive prognosis for patients[4].

The ability to treat both ER^+ and *HER2*-amplified breast cancers is attributable to the molecular drivers that they are named after, which provide targets for therapeutic intervention. The principal approach for targeting ER^+ breast cancer is with the application of endocrine therapeutic Tamoxifen, which competes with estrogen for binding to the estrogen receptor, inhibiting downstream signalling that drives the cancer[5]. HER2-positive cancers, which constitute approximately

twenty percent of breast cancers, were historically associated with a poorer prognosis. However, the advent of HER2-directed monoclonal antibodies such as Trastuzumab has provided an effective treatment strategy for targeting HER2-enriched breast cancers[6].

As a defining principle, the most clinically aggressive form of breast cancer lacks both HER2 overexpression and ER positive status. Termed ‘triple negative breast cancer’ (TNBC), this subtype also lacks expression of the progesterone receptor, and accounts for approximately 12-17% of breast cancer cases. This form of breast cancer can be further sub-categorised into either basal-like or claudin-low, the latter exhibiting a deficiency in cell-cell adhesion claudin family proteins, and a gene expression profile favouring epithelial to mesenchymal transition and stemness. An absence of the primary therapeutic targets makes this form of cancer particularly difficult to treat, with chemotherapy still providing the primary answer to this form of the disease[7].

Regardless of subtype, breast cancer remains the second highest source of cancer-related mortality in women (Cancer Research UK), and resistance to the previously-described treatments often emerges, rendering them ineffective. Thus, a more intricate understanding of the molecular underpinnings of breast cancer remains imperative if we are to reduce breast cancer to a largely treatable and even curable disease.

1.2: Arginine methylation

Arginine methylation is a post-translational modification (PTM) applied to the terminal guanidino nitrogen atoms on residues of the basic amino acid arginine. Despite increasing bulk and hydrophobicity, it does not actually alter the cationic charge of the residue **(Figure 1.1),** but does decrease the number of hydrogen bond donors provided by this group with the addition of each methyl group[8].

Despite its discovery in 1971[9,10], the ensuing fifty years of biological research would leave arginine methylation relatively understudied compared to other PTMs such as lysine methylation. Much of this distinction comes from difficulty in generating reliable anti-methylarginine antibodies, possibly due to the aforementioned electrophysical subtlety of the modification. However, since the cloning of PRMT1 in the 1990s[11] interest in arginine methylation has increased, and it has become clear that this highly abundant post-translational modification (occupying ~2% arginine residues in rat liver nuclei[12]) plays an important role in numerous vital cellular processes. Greater still perhaps is the growing appreciation that aberrant arginine methylation contributes to numerous oncogenic events, including both initiation and progression of cancers.

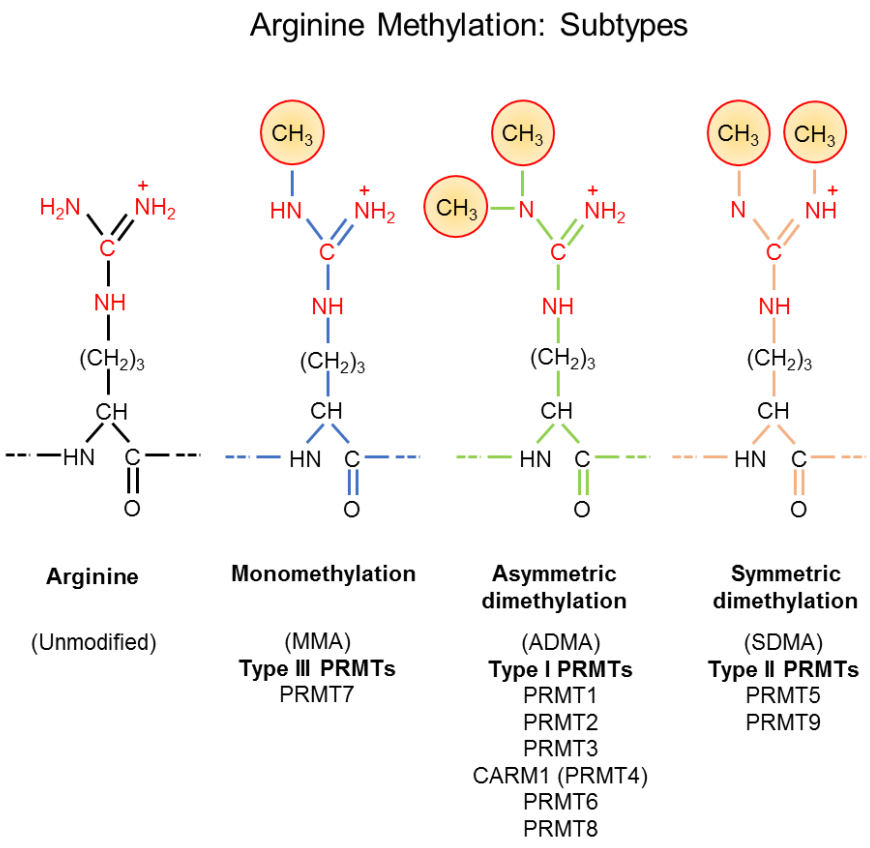


Figure 1.1: Different modes of arginine methylation

Depiction of unmodified arginine side chain that features a terminal guanidino group (red atoms), and potential arginine methylation events. PRMTs that catalyse each event are listed below. For simplicity, only PRMT7 is listed as capable of MMA, although all Type I and Type II PRMTs catalyse MMA as a precursor to DMA.

1.3: PRMTs: The writers of arginine methylation

1.3.1: Background

Post-translational modifications confer additional 'fine tuning' to the behaviour of the protein species to which they are applied. As a generic concept, this event is modulated by three factors: writers, readers and erasers. Writers are enzymatic applicators of these modifications, and in the context of arginine methylation this role is contributed by the protein arginine methyltransferases (PRMTs). This family of enzymes currently contains nine members, subdivided into three subtypes: I, II and III. Type I PRMTs include PRMT1, PRMT2, PRMT3, PRMT4 (CARM1), PRMT6 and PRMT8. PRMT1 and PRMT3 are generally observed to deliver their characteristic asymmetric dimethyl mark to target residues in a distributive rather than processive manner, where enzymatic dissociation occurs after a monomethyl mark is applied to the target arginine, requiring re-association of the enzyme in order to generate a dimethyl mark[13]. However in the case of PRMT1, certain amino acid sequences within the substrate can give rise to increasing levels of processivity, where the enzyme repeatedly modifies the same substrate without dissociation[14]. Type II PRMTs (PRMT5 and PRMT9) are defined by their ability to symmetrically dimethylate targets[15,16], whereas type III PRMTs (a group currently limited to PRMT7) are only capable of delivering a monomethyl mark **(Figure 1.1)**[17].

Despite this variability in target modification, there is a general consensus to the structural blueprint of PRMTs. Central to their activity, all feature a conserved catalytic core consisting of approximately 300 amino acids. This can be further

subdivided into an N-terminal SAM (S-adenosyl-L-methionine, the universal methyl donor) binding site (also known as the Rossman fold) and a β-barrel that serves as a substrate binding site. In the case of PRMT1, The N-terminal end of the β-barrel features a dimerisation arm that contacts the N-terminus of the Rossman fold of the opposing monomer during dimerisation[18]. Within these domains are six highly conserved motifs containing specific properties: motif I (VLD/VGxGxG) contains three critical glycine residues that contribute the core of the catalytic domain. Post motif I (V/I-X-G/A–X-D/E) is responsible for forming hydrogen bond contacts with the ribose-hydroxyl moiety of SAM. Motif II (E/K/VDII) stabilises motif I through β-sheet formation, and the double-E motif (SExMGxxLxxExM) contains two crucial glutamic acid residues that stabilise contacts with arginine substrate, and motif III (LK/xxGxxxP) forms a parallel β-sheet with motif II. Finally, the THW loop plays a role in stabilisation of the N-terminal α-helix as well as substrate binding. All of these motifs are found within the Rossman fold, with the exception of the THW loop, which is located in the β-barrell **(Figure 1.2)**[18].

Higher-order structuring occurs in all PRMTs with the exception of PRMT7 and PRMT9 as their methyltransferase domains are ancestrally duplicated and therefore present as pseudo-dimers in their monomeric form[19,20]. All other PRMTs are observed as homodimers at a minimum, a conformation that is essential to their catalytic activity. Some members of the family are capable of higher order oligomerisation, with PRMT5 forming a tetramer and PRMT1 and PRMT8 both

forming yet higher order structures, respectively (the former in a concentration-dependent manner)[18].

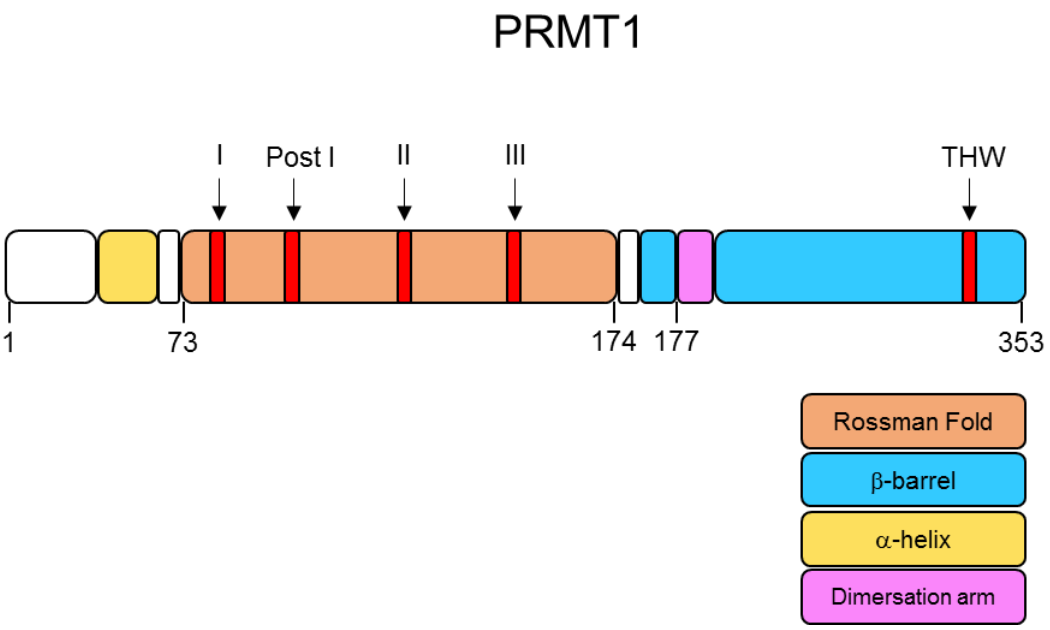


Figure 1.2: Linear representation of the PRMT1 functional domains

Functional motifs are marked by red regions. Motif I contributes to the core of the catalytic domain. Post motif I provides contacts involved in SAM binding. Motif II stabilises motif I via β-sheet formation, which is further supported by parallel β-sheet formation of motif III. "THW" refers to the THW loop, responsible for stabilisation of the N-terminal α-helix and substrate binding[18]. Amino acid numbers are listed below the structure. Adapted from Fulton et al., 2019[21].

1.3.2: Regulation of PRMTs: Target motif specificity

A fundamental mechanism by which PRMTs discriminate between target arginines in a non-random manner is through recognition of specific amino acid sequence motifs. The most well characterised of these is the glycine-arginine rich (GAR, also known as RG/RGG) motif, hereafter referred to as the RG motif. Such motifs often occur within intrinsically disordered domains and are generally targeted for methylation by PRMTs[22]. This trend is not without exceptions, and arginine methylation has been observed in the absence of RG motifs, such as in the case of PRMT1-mediated methylation of FOXO1 at arginines 248 and 250[23]. Further, CARM1 and PRMT5 are capable of targeting entirely different motifs: CARM1 is thought to specifically target arginine residues flanked by proline residues (known as the PGM motif) whereas PRMT5 can target both RG and PGM motifs[24,25]. More recently, PRMT5 RG motif preference has been further refined to 'GRG', with emphasis on the recipient arginine being flanked by glycines on both sides for it to be more effectively targeted[26].

1.3.3: Regulation of PRMTs: Post-translational modification

Several modes of regulation exist in order to fine-tune PRMT activity. One example of this is post-translational modification of the PRMTs themselves. Auto-methylation of CARM1 at R551 is necessary for its activity in transcription and alternative splicing[27], whereas auto-methylation of PRMT6 at R35 increases protein stability. In contrast, PRMT8 auto-methylation at R58 and R73 decreases its methylation activity[28,29].

Phosphorylation of certain PRMTs is another documented phenomenon. CARM1 is phosphorylated at S217 and S229, decreasing enzymatic activity by inhibiting SAM binding and disrupting dimerisation, respectively[30,31]. Methyltransferase activity of PRMT5 is negatively regulated by phosphorylation; Janus kinase JAK2-V617F-mediated phosphorylation of PRMT5 prevents binding of coactivator MEP50, and phosphorylation of T139 and T144 by liver kinase B1 (LKB1) prevents methyltransferase activity by impeding binding of cofactors MEP50, RIOK1 and PICln[32,33]. PRMT1 activity can also be affected via phosphorylation, with phosphorylation by CSNK1a1 altering genomic targeting of PRMT1, resulting in changes in gene transcription[34].

Ubiquitination is also applied to several PRMTs, and PRMT1 and PRMT5 are both known targets of E3 ubiquitin ligases. This modification (administered by E4B and CHIP, respectively) results in their targeted degradation by the proteasome. In the case of PRMT5, this activity is postulated to be tumour suppressive, although the relevant biological consequences of PRMT1 ubiquitin-mediated degradation have yet to be divulged[35,36].

1.3.4: Regulation of PRMTs: micro-RNAs

Several PRMTs are negatively regulated by micro-RNAs. miR-503 targets the 3'UTR of *PRMT1* mRNA, reducing PRMT1-mediated invasiveness in hepatocellular carcinoma through suppression of epithelial to mesenchymal transition (EMT)[37]. Similarly, miR-415 targets *PRMT5* mRNA in glioma, leading to apoptosis and arrest of cell growth. Effects of miR-415 are counteracted by its

sequestration by long non-coding RNA *SNHG16* which is highly expressed in clinical samples and correlates with poorer patient prognoses in glioma[38]. Conversely, miR-543 appears to produce a pro-oncogenic effect in osteosarcoma by targeting the 3'UTR of PRMT9 mRNA, as evidenced by higher miR-543 inversely correlating with PRMT9 levels in patient samples and osteosarcoma cell lines[39].

1.3.5: Regulation of PRMTs: Cofactor Binding

Interaction with other proteins can significantly alter PRMT behaviour. Although homodimerisation is the *de facto* configuration for PRMT activity, heterodimerisation between PRMT1 and PRMT8 has been documented *in vitro*. This may occur in order to recruit PRMT1 activity to the plasma membrane (PRMT8 is the only PRMT to be N-terminally myristoylated, allowing binding to the plasma membrane), and facilitated by the otherwise high structural similarity between both enzymes[40]. PRMT1 also binds hCAF1 and BTG1, transcription factors that interact in a synergistic manner as part of the mammalian CCR4-NOT complex. This complex is heavily involved in gene expression from transcription through to translation via a multitude of mechanisms[41], however as singular subunits hCAF1 and BTG apparently have opposing influences on PRMT1 enzymatic behaviour. Whereas binding of hCAF1 reduces PRMT1-mediated methylation of histone H4 and SAM68[42], BTG1 binds to PRMT1 to promote methylation of substrates[11]. PRMT1 also binds Hepatitis B protein that negatively regulates its methyltransferase activity[43].

As previously mentioned, PRMT5 is bound by its obligate cofactor MEP50, activating its methyltransferase activity by directing substrates to its active site. However, it can also be bound by serine-threonine protein kinase RIOK1 and PICln in a mutually exclusive manner, which alter substrate specifcity[44].

1.4: Readers of arginine methylation

A major strategy by which post-translational modifications produce a downstream effect is through interaction with proteins generically referred to as ‘readers’. The best-characterised family of methyl-arginine readers are Tudor domain proteins, which form part of a wider ‘Royal Family’ of proteins that also includes PWWP, chromodomain and MBT repeat domain proteins[45]. By establishing interactions through their aromatic cages, Tudor domain proteins are capable of interacting with methyl-lysine and methyl-arginine. However, variation in shape between different Tudor domain-containing proteins allows discrimination between target residues, with readers of methyl-arginine containing narrower aromatic cages than readers of methyl-lysine[46]. Specificity of methyl-arginine binding can be further subdivided into binders of symmetric versus asymmetric dimethylation: for example, SMN (survival of motor neurons) preferentially interacts with symmetric dimethyl-arginine, whereas TDRD3 (Tudor domain containing protein 3) preferentially recognises asymmetric dimethylation[46].

More recently, a Tudor-like SPIN-SSTY domain has been discovered in the SPIN protein family, which also relies on an aromatic cage for target interaction. This domain is structurally variable even within the same protein monomer, with SPIN1

containing three distinct SPIN-SSTY domains. Domain I interacts with asymmetrically dimethylated R8 on histone H3 (H3R8me2a) whilst domain II simultaneously contacts the trimethyl-lysine modification on H3K4 (H3K4me3)[47]. Currently, domain III is thought to be non-functional[48].

The Tudor domain and its structural relatives are not the only known readers of methyl-arginine however, and the WD40 β-propeller domain of transcriptional coactivator WDR5 has been shown to bind symmetric dimethylation of H3R2, supporting the maintenance of a euchromatic state[49]. In addition, PHD (plant homeodomain) domains of UHFR1 and RAG2, and the UBR (Ubiquitin E3 ligase n-recognin) domain of UBR2 are identified as putative arginine binding domains[50].

1.5: Erasers of arginine methylation

1.5.1: Arginine demethylases

There are currently no known enzymes that demethylate arginine residues unambiguously, such as in the case of lysine demethylases or deacetylases. Hence, the turnover of methylated arginines is a contentious topic within the field. JmjC histone lysine demethylases such as KDM4E and KDM5C have demonstrated weak arginine demethylase activity in a cell-free context, however evidence for this phenomenon *in vitro* has yet to be demonstrated, as has an ability to discriminate between symmetric and asymmetric dimethylation[51].

Perhaps most controversial among the arginine demethylase candidates is the 2OG hydroxylase JMJD6. Despite demonstrating weak reduction of asymmetric dimethylation in a cell-free reaction and expression levels correlating inversely with asymmetric dimethylation of histone H4 *in vitro*[52,53], inconsistency within the literature has prevented a robust consensus from forming over the true nature of the enzyme[54].

Finally, the conversion of methyl-arginine residues to citrulline by PADI (peptidylarginine deiminase) enzymes such as PADI4 has been proposed as a means to remove methylation from arginine. However, despite antagonizing CARM1-dependent methylation of histone H3, PADI4-dependent deimination of arginine residues occurs only in unmethylated or monomethylated arginine residues, and not those that are asymmetrically dimethylated. This suggests that citrullination of arginine may be a modification that impedes, rather than removes arginine methylation[55].

1.6: PRMT1: Normal functions and oncogenicity

1.6.1: Background

PRMT1 is the predominant protein arginine methyltransferase in mammalian cells, contributing approximately 85% of arginine methylation in mouse liver extract[56]. The *PRMT1* gene encodes seven translated isoforms **(Figure 1.3)** which differ in their N-termini and are all catalytically active, with the exception of isoform 7. These isoforms show tissue specific expression: variant 1 (PRMT1v1) is mostly expressed in the kidney, liver, lung, skeletal muscle, and spleen,

whereas PRMT1v2 is mostly found in the pancreas, kidney and liver. Variant 3 is found at low levels ubiquitously, however the remaining isoforms are significantly more restricted in their distribution. PRMT1v4 is limited to the heart, PRMT1v5 is predominantly pancreatic and PRMT1v7 is mostly expressed in skeletal muscle and the heart. So far, no tissues have been identified that express PRMT1v6.

The varying N-termini of these isoforms also confer changes in subcellular localisation. For example, PRMT1v1 is expressed throughout the cell, although it is predominantly nuclear. However, PRMT1v2 is almost entirely cytoplasmic due to the presence of an N-terminal nuclear export signal. Alternative splicing of PRMT1 also confers substrate specificity under cell-free circumstances. For example PRMT1v1 and PRMT1v2 showed superior ability to methylate SAM68 and SMB relative to other isoforms[57]. More recently, an additional isoform known as PRMT1Δarm was identified, lacking the dimerisation arm and rendering it catalytically inactive **(Figure 1.3)**. However, it has been observed that PRMT1Δarm interacts stably with PRMT1 substrates, perhaps providing a dominant-negative protective effect against PRMT1-mediated neoplastic progression. This is supported by a model in which HEK-293T cells were made to undergo EMT by ectopic expression of SNAIL1, causing an increase in expression of the PRMT1Δarm variant, resulting in apoptosis [58].

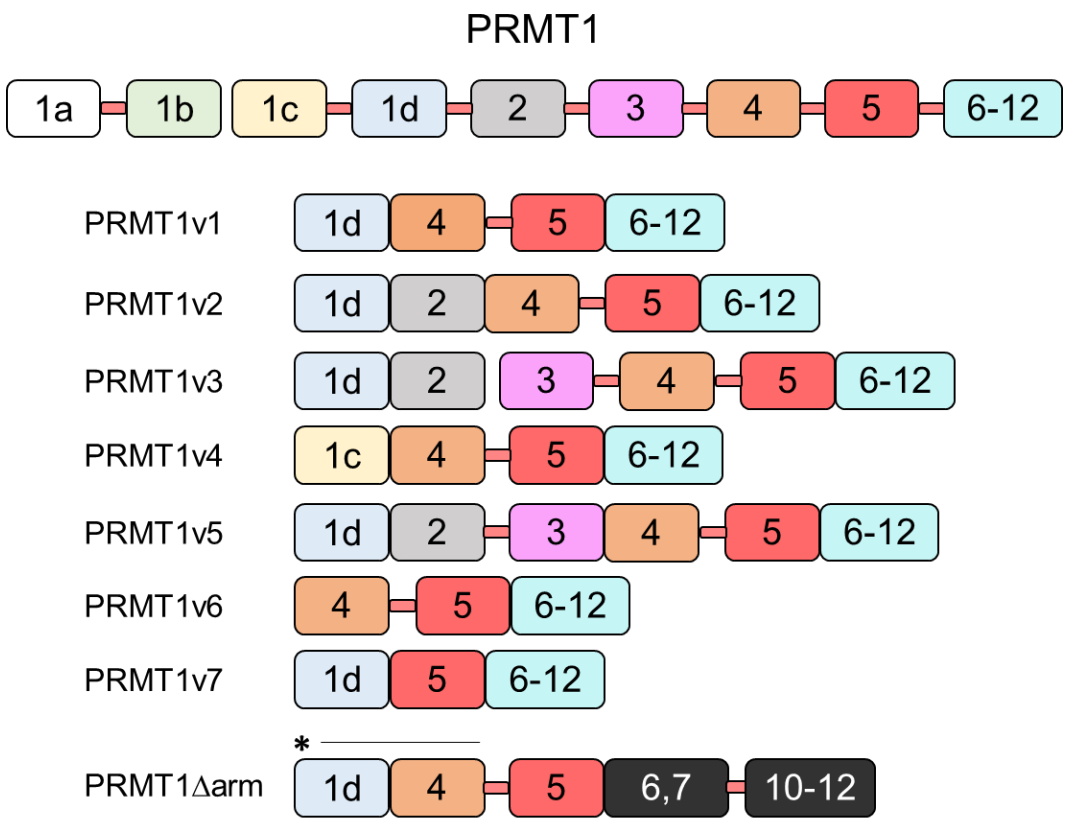


Figure 1.3: Exon inclusion schematic for all known PRMT1 isoforms

Top image shows all possible exons that can be included in a PRMT1 splice variant. Images below depict exons included in each known PRMT1 isoform. Red connectors imply introns from which alternative splicing is applied. The asterisk and line above the 5' end of Δarm (bottom) indicates the exons thought to be included in this region[58]. Adapted from Baldwin et al., 2014[59].

1.6.2: PRMT1 and cancer

PRMT1 influences a growing and diverse number of important cellular processes, however much research interest has accumulated around this enzyme due to its properties as an oncogene. This is starkly demonstrated by its contribution to neoplastic inception in the mammary gland *in vivo.* Female mice overexpressing PRMT1 under control of the Keratin-5 promoter developed hyperplasia of the mammary gland epithelium by 4 months of age, and were significantly more prone to spontaneous tumour development once aged. These findings suggest that PRMT1 contributes a 'priming' capacity in oncogenesis, by causing cells to develop into a pre-malignant state, with further mutations that occur over time providing a 'tipping point' for malignant growth to occur. Further, co-overexpression of Neu (Her2) and PRMT1 showed a significantly enhanced mammary tumour occurrence than in mice singularly overexpressing Neu (despite the latter genotype being sufficient for tumourigenesis). RNA seq and subsequent pathway analysis revealed that overexpression of PRMT1 in these mice caused an increase in pro-oncogenic PI3K/AKT signalling. Thus, PRMT1 was also implicated as a powerful enhancer of tumourigenesis, working synergistically with Neu-driven transformation[60].

Clinical data support this role from a correlative perspective, and the malignant potential of PRMT1 is documented across numerous cancer types: in prostate cancer, increased PRMT1 levels correlate positively with increasing tumour grade[61], and higher expression levels correlate with increased aggressiveness in colorectal cancer[62]. Perhaps most prominently explored is the contribution of PRMT1 to breast cancer, the focal point of this thesis. This disease is still the

second most common cause of cancer-related mortality in women in the UK (2017, Cancer Research UK) and high PRMT1 expression correlates with poorer patient prognoses[63] **(Figure 1.4).** As a result of its prolific role in oncogenesis and cancer progression, a type I PRMT inhibitor is in phase 1 clinical trials: GSK3368715 (trial number NCT03666988) for the treatment of refractory solid tumours (pancreatic, bladder, non-small cell lung cancer) and diffuse large B cell lymphoma (DLBCL).

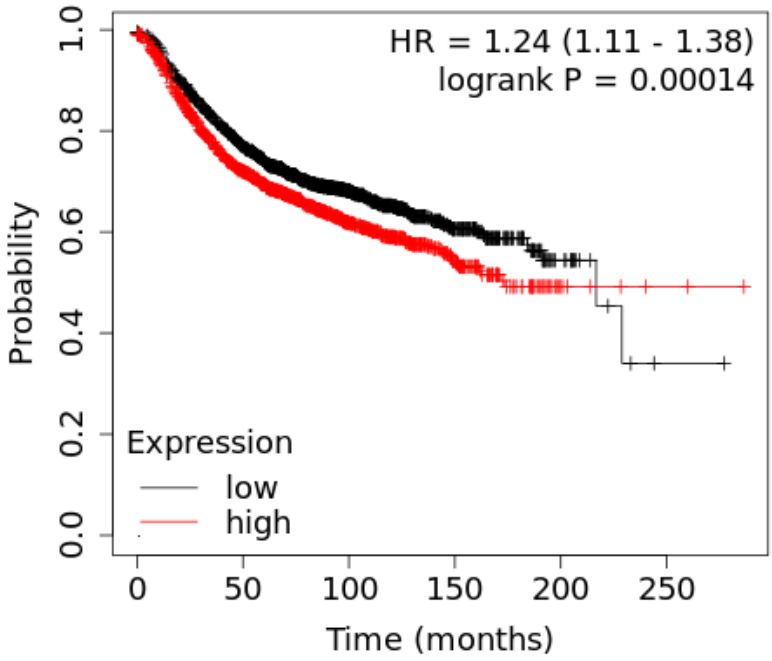


Figure 1.4: High *PRMT1* expression in breast cancer

Kaplan-Meier plot showing difference in breast cancer patient survival (all subtypes) in individuals with either high or low *PRMT1* expression (www.kmplot.com, Affymetrix probe ID 206445_s_at). Data were derived from the following datasets: E-MTAB-365, E-TABM-43, and GSE datasets 11121, 12093, 12276, 1456, 16391, 16446, 16716, 17705, 17907, 18728, 19615, 20194, 20271, 2034, 20685, 20711, 21653, 22093, 25066, 2603, 26971, 29044, 2990, 31448, 31519, 32646, 3494, 36771, 37946, 41998, 42568, 43358, 43365, 45255, 4611, 46184, 48390, 50948, 5327, 58812, 61304, 65194, 6532, 69031, 7390, 76275, 78958, 9195.

1.6.3: Epigenetics: introduction and links to cancer

First coined by Conrad Waddington in 1942[64], the classical definition of epigenetics refers to changes to the phenotype of an organism that occur without changes to its genotype. In the age of molecular biology, these concepts generally relate to chemical modifications applied to chromatin, (such as DNA methylation and post-translational modification of histones) although gene regulation by non-coding RNAs and chromatin remodelling also fall under the umbrella of epigenetic regulation[65]. Thus, epigenetics has emerged as one of the fundamental avenues through which cells modulate use of their genetic information.

As a result, it is perhaps unsurprising that epigenetic contributions to malignancy are increasingly documented. In response, this has also triggered attempts to design drugs against epigenetic readers and writers in order to curb the activity of oncogenes they regulate. Targeting of EZH2, the catalytic subunit of the Polycomb repressor complex is a notable example. Gain of function mutations in EZH2 lead to repression of genes that cause differentiation in B cells by aberrant application of the repressive H3K27me3 histone mark. This in turn maintains an uncommitted and highly proliferative state, providing a strong driving contribution to the leukaemic phenotype[65]. To counter this phenomenon, EZH2 inhibitors have been developed by GSK (GSK2816126) and Epizyme (Tamezostat) and are currently undergoing phase II clinical trials for B-cell lymphoma[66].

EZH2-mediated H3K27me3 has also been revealed alongside PRMT5-dependent arginine methylation as a crucial Achilles heel in epigenetic promotion of breast cancer. Treatment of MCF7 cells with Resveratrol has been shown to lower expression levels of PRMT5 and EZH2, reducing repressive H4R3me2s and H3K27me3 at the promoters of crucial tumour suppressive genes such as *p53*, *p21* and *BRCA1*[67]*.* Thus, focussing on epigenetic modulators for drug targeting holds significant promise, even within the context of arginine methylation and breast cancer.

1.6.4: PRMT1 in epigenetic regulation of gene expression

One of the best-characterised roles for PRMT1 is in its regulation of gene expression, most notably in its function as a transcriptional coactivator. By asymmetrically dimethylating histone H4 at arginine 3 (H4R3me2a), PRMT1 plays a role in the formation and maintenance of transcriptionally active regions of chromatin.

The fundamentals of this phenomenon were originally discovered through methylation assays involving exposure of PRMT1 derived from HeLa cell extracts to histone H4. Subsequent experimentation showed that pre-methylated H4 was more amenable to acetylation by histone acetyltransferase p300 under cell-free conditions, at K8 and K12. This finding was then reinforced *in vitro* in *Xenopus* oocytes, where expression from the CBP/p300 coactivator-dependent androgen receptor (AR) locus was enhanced by PRMT1 overexpression, but not overexpression of a catalytically inactive mutant variant[68].

Further characterisation of this event came through work on the *β-globin* locus in erythroid cells derived from chicken and mouse. In the first instance, presence of H4R3me2a potentiated the locus for an increase in H3K9 and H3K14 acetylation, as well as H4K5 and H4K12 acetylation at the upstream Folate receptor (*Fr*) HSA and 5'HS4 elements, all modifications associated with a euchromatic ('open' and transcriptionally amenable) state[69]. In the mouse, PRMT1 is recruited by transcription factor USF1 to the LCR (locus control region) enhancer region and β^{maj} promoter at the *β-globin* locus, increasing H4R3me2a in those regions, and maintaining H3K9 and H3K14 acetylation.

Despite containing a bromodomain for recognition of acetylated lysine[70], p300/CBP associated factor (PCAF) can bind H4R3me2a, and both PCAF and p300 show increased H3K9/14-targeted HAT (histone acetyltransferase) activity towards pre-methylated H4R3me2a nucleosomes in a cell-free context. This phenomenon is observed even when PRMT1 is removed from the reaction mixture prior to addition of either HAT, suggesting that the presence of the H4R3me2a mark rather than recruitment by PRMT1 is responsible for subsequent HAT activity. Overall, these modifications to the histone code lead to an increase in long-range interaction between the HS2 enhancer and β^{maj} promoter and increased transcription of the *β-globin* locus, promoting erythroid differentiation[70].

H4R3me2a is also recognised by the Tudor domain of TDRD3, which recruits the TDRD3 protein complex containing a mixture of RNA binding proteins and

transcription factors to facilitate transcriptional progression[71]. Among these interactions is the N-terminus of Topoisomerase III β (TOP3B), that stabilises TOP3B and recruits it to sites of active transcription. Here, it uses its topoisomerase activity to resolve transcription-stalling R-loops formed when nascent RNA transcripts re-anneal to DNA, and by negative supercoiling of these regions caused by the mechanical progression of RNA polymerase II[72].

Despite the pro-transcriptional nature of H4R3me2a being common knowledge in the field for some time, genome-wide ChIP data for H4R3me2a are scarce, possibly due to an electrostatic discreteness of the modification that results in poor antibody enrichment[73]. However, PRMT1 occupancy across the genome has been investigated using an HA-tagged version of the enzyme with anti-HA antibodies in human keratinocytes. In accordance with the pro-transcriptional nature of its histone mark, 98% of the peaks obtained from the experiment were associated with open or transcriptionally competent chromatin when cross-referenced with ENCODE DNase I hypersensitivity sites. Additionally, 37.7% of these peaks localised to gene promoters and intergenic distal regulatory regions. In the context of the study, phosphorylation of PRMT1 by CSNK1a1 was found to direct PRMT1 to genes that promote proliferation (and promote their transcription) as well as to the promoter of epidermal pro-differentiation gene *GRHL3*, repressing its expression. Why PRMT1 occupancy of *GRHL3* had an inverse and uncharacteristic effect was not elucidated, however the sum consequence of the aforementioned activity was crucial to maintaining progenitor status in epidermal cells[34].

1.6.5: Contributions to splicing

PRMT1 possesses numerous substrates that contribute to splicing events. PRMT1 methylates heterogeneous nuclear ribonucleoprotein A2 (hnRNP A2, an important agent in pre-mRNA processing), causing it to localise to the nucleus in NIH-3T3 and HEK-293T cells, facilitating splicing events[74]. Similarly, nuclear localisation of RNA-binding protein Sam68 and subsequent nuclear export of unspliced HIV mRNAs is dependent on PRMT1-mediated methylation of Sam68[75].

Another hnRNP targeted for methylation by PRMT1 is hnRNPA1. Asymmetric dimethylation of the C-terminal RG region by PRMT1 leads to decreased mRNA binding and ITAF (IRES *trans*-acting factor) activity of the protein (generation of transcripts from an IRES)[76]. PRMT1-mediated control of splicing also plays an influential role in cardiac homeostasis *in vivo*. Although the molecular mechanisms responsible for this have yet to be identified, Cre-dependent knockout of *PRMT1* in cardiac cells of the mouse results in aberrant alternative splicing of a large variety of genes, and ultimately, lethality via dilated cardiac myopathy (DCM)[77]. In addition to developmental defects, aberrant RNA splicing can also have oncogenic consequences. PRMT1 plays an indirect role in splicing in acute megakaryocytic leukaemia by methylating RNA-binding splicing factor RBM15 at R578. This mark serves as a methyl degron, leading to ubiquitination of RBM15 by E3 ubiquitin ligase CNOT4. Thus, by lowering RBM15 protein levels, PRMT1 alters splicing of transcripts such as *GATA1, RUNX1, TAL1* and *c-MPL* that are important for megakaryocyte maturation by reducing interaction of RBM15 with splicing factor SF3B1[78].

PRMT1 has been implicated as an important regulator of alternative splicing in AML, where its influence was identified through a drug screen to test vulnerability of splicing factor SRSF2 mutant cells to pharmacological intervention. Treatment with PRMT1 inhibitor MS023 produced a stronger response in K562 SRSF2^{P95H} cells than in their wild type counterparts, an effect synergised by co-inhibition of PRMT5 with GSK591. Although the substrate components of the splicing machinery that effected this change were not revealed, this event was thought to be influenced by restoration of wild-type EZH2 expression, rather than an alternatively-spliced EZH2 variant containing a poison exon that results in nonsense-mediated decay that is more highly expressed in SRSF2 and SF3B1 mutant cells[79].

1.6.6: Translation

Deregulation of translation can have pro-oncogenic properties, which can occur through multiple mechanisms such as down-regulation of tumour suppressive miRNA or hyper-activation of translation through increased ribosome biogenesis. Coincidentally, both scenarios are documented in cancers where *TP53* is deleted or mutated[80,81]. In the context of arginine methylation, this provides a therapeutic opportunity as osteosarcoma cells with deletions or mutations in *TP53* show a high degree of sensitivity to PRMT1 depletion. This was determined to be a result of PRMT1-mediated methylation of eIF4G1 (a key member of the translation initiation complex) at R689, promoting stability of the protein. Additionally, PRMT1 was found to methylate other members of the translation initiation complex eIF4A and eIF4E. Overall, these events appeared to conserve the

integrity of the translation initiation complex, as PRMT1 knockdown reduced association of eIF4G1 with eIF4A and Pabp1[82].

1.6.7: Cell Fate

Several different cell types rely on PRMT1-dependent arginine methylation for lineage determination. PRMT1 methylates the cytoplasmic IGα subunit of the membrane-bound B-cell receptor (BCR) in B cells, via a conserved RG motif at position 198 in close proximity to the immunoreceptor tyrosine-based activation motif (ITAM). This proximity allows the arginine methylation event to antagonise phosphorylation of the ITAM and subsequent binding by SYK phosphotyrosine kinase, preventing SYK and PI3K-dependent degradation of FOXO family transcription factors that promote differentiation. Thus, PRMT1 drives maturation of B cells from the pre-B cell state to the BCR^+ state[83].

Muscle stem cell (MSC) differentiation is another form of lineage commitment involving PRMT1, in which it methylates the eye-absent catalytic domain of phosphotyrosine phosphatase Eya1. Eya1 interacts with transcription factor Six1 and coactivates expression of genes involved in the myogenesis cascade, such as *Pax3, MyoD, Myf5,* and *Myf6*. Six1 activity at gene promoters is thought to be repressive in the absence of Eya1 binding, and reduction in PRMT1 levels decreases association of Eya1 and Six1, causing Six1 to singularly bind to the promoter of *MyoD,* an important factor in MSC differentiation. This suggests that PRMT-mediated methylation of Eya1 influences binding to Six1. Consequently, loss of this interaction results in expansion of the MSC population but a loss of

differentiation, which is backed up *in vivo,* where MSCs isolated from mice that were *PRMT1-/-* failed to align for fusion and form mature, multinucleated fibres. Notably, this results in a phenotype in which muscle regeneration is severely impaired after injury[84].

PRMT1 also regulates expression of retinoic acid (RA) responsive genes during neuronal cell differentiation. The enhancers of PRMT1-sensitive genes were characterised by presence of H4R3me2a, as well as occupancy by Oct3/4, Sox2, Klf4, Nanog and P300 binding. Although the precise role in the process is not elucidated, it is thought to help coordinate RA-dependent gene expression in a lineage stage-specific manner[85].

PRMT1 plays a role in early-stage differentiation of embryonic stem cells (ESCs) in the mouse through methylation of pluripotency factor Klf4 at R396. This modification has proven necessary for Klf4-dependent recruitment of the repressive mSin3a/HDAC complex to genes that influence primitive endoderm (PrE) commitment, thus reducing chromatin accessibility of pro-PrE genes such as *Gata4* and *Gata6*. This is further evidenced *in vivo*, where mouse embryos treated with PRMT1 inhibitor furamidine dihydrochloride induce expression of Gata6, biasing ESC stage cells towards PrE commitment[86].

An example of neoplastic involvement of PRMT1 in cell fate is seen in acute megakaryocytic leukaemia (AMKL), where methylation of RBM15 leads to

alternative splicing of genes that promote retention of a less differentiated state. For example, PRMT1 inhibition with DB75 in MEG-01, CMK and CMY AMKL cell lines increases the ratio of GATA1fl:GATA1s transcripts. Although both transcripts support differentiation, only the latter promotes subsequent proliferation. Thus, a decrease in GATA1s reduces the expansion of more differentiated cells. Additional consequences of these splicing differences include decreased RUNX1a relative to RUNX1b,c, (the former isoform acting as a dominant negative that prevents differentiation), and pro-differentiation c-MPL-exon9+ transcript variant also decreasing relative to c-MPL-exon9- under PRMT1 inhibition. As a result, PRMT1 was identified as a regulator that maintains the megakaryocyte-erythrocyte stage of development, and inhibits the emergence of mature polyploid $CD41^+/CD61^+CD42^+$ megakaryocytes – a phenomenon that could potentially contribute to the development of megakaryocytic leukaemia[87].

1.6.8: PRMT1 in breast cancer

PRMT1 is noted in contributing to breast cancer via deposition of pro-transcriptional H4R3me2a. In ER^+ and triple negative breast cancer cells (as well as normal mammary epithelial cells), H4R3me2a is found at the *ZEB1* promoter, a transcription factor that promotes EMT, cancer cell invasiveness and stem-like qualities such as increased self-renewal. H4R3me2a has also been implicated in the expansion of tumour-initiating cells in oesophageal squamous cell carcinoma[88], and in MLL-GAS7/MOZ-TIF2 acute myeloid leukaemias (AMLs) where it promotes downstream transcription of pro-leukaemic targets *Hoxa9* and *Meis1*[89]. Hence, histone H4R3me2a, catalysed by PRMT1, contributes to breast cancer as well as other forms of the disease.

PRMT1 is also heavily implicated in breast cancer through cell signalling activity. With notable relevance to the cell models used in this thesis, PRMT1 is extensively characterised in estrogen signalling-mediated cell growth. The estrogen receptor α (ERα) binds the steroid hormone estrogen changing the conformation of ERα and causing it to dimerise. Dimeric ERα then translocates to the nucleus allowing it to either directly interact with DNA via the estrogen response enhancer, or indirectly via binding to DNA-bound transcription factors, thus functioning as a co-activator.

Estrogen signalling also contributes to a wide array of other functions that are non-genomic in nature, such as MAPK and AKT signalling. One organ that is highly dependent on estrogen signalling for normal development is the mammary gland, where it plays a central role in differentiation, proliferation and development of mammary epithelial cells[90]. However, approximately 75% of breast cancers are also ER^{+}[91], which contributes to the cancer phenotype through strong mitogenic signalling. PRMT1 transiently methylates ERα at R260 after estrogen stimulation in the cytoplasm, facilitating assembly of a complex containing the p85 subunit of PI3K and tyrosine kinase Src (activity of both kinases is also necessary for this association)[92]. Alternatively, IGF1-R also associates with methyl-ERα in a PRMT1-dependent manner, activating ERα independently of estrogen binding[93]. Downstream of this event, signalling from serine/threonine kinase AKT is activated, promoting proliferative and anti-apoptotic signalling programs such as up-regulation of Cyclin D1. Importantly, this mechanism was elucidated in MCF7 breast cancer cells and methylated

ERα was detected in over half of clinical breast cancer samples tested, demonstrating the importance of PRMT1 in estrogen-dependent breast cancers[92].

PRMT1 also methylates the progesterone receptor (PR) at R637 within the nuclei of T47D breast cancer cells, promoting stability of PR. Knockdown of PRMT1 upon PR stimulation has revealed that PRMT1 influences expression of approximately 30% of genes downstream of PR activation. Included in the downregulated cohort are genes such as *EGFR, EGR1, SGK1*, and *CD44* which are crucial in mammary epithelial differentiation, but also contribute to invasiveness in the context of breast cancer. This is suggestive of PRMT1-mediated methylation of PR in breast cancer having a pro-oncogenic role, a conclusion supported by clinical data in which high PR/low PRMT1 patients experience a significantly longer relapse-free survival than high PR/high PRMT1 patients[94].

More recently, PRMT1 has been characterised as a key factor in resistance to radiotherapy in ER^+ and triple negative breast cancer models (MCF7 and MDA-MB-231, respectively). In response to ionizing radiation, cells were found to increase SAM production, resulting in an increase of PRMT1-mediated methylation of BRCA1. This event was critical for BRCA1-BARD1 association, which shuttles BRCA1 from the cytoplasm to the nucleus, ultimately promoting DNA repair by the error-free homologous repair (HR) mechanism. Another effect of this activity is that depletion of cytosolic BRCA1 through nuclear re-localisation

has an anti-apoptotic effect, by reducing interaction with anti-apoptotic protein Bcl2. This prevents Bcl2 being targeted for degradation via BRCA1-mediated ubiquitination, promoting cell survival. Thus, PRMT1 may be an important factor not just in oncogenic inception and progression, but also in resistance to conventional breast cancer treatment methods[95].

1.7: Introduction to Bio-orthogonal Profiling

1.7.1: Background

Key to understanding the significance of a methylation event, and how it is perturbed in disease, is the ability to understand the substrate repertoire of a specific enzyme. Until relatively recently, profiling of arginine-methylated proteins had been accomplished primarily through mass spectrometry. This method has been combined with techniques such as hmSILAC (heavy methyl Stable Amino acid Labelling in Cell Culture, which produces MS-detectable methyl-arginine by modifying arginines with heavy methyl groups) as well as antibody-mediated enrichment. These approaches have identified many novel arginine modifications within the proteome, and examination of nuclear and cytosolic fractions provides additional (albeit crude) cellular context[96].

However, these approaches have caveats. For example, antibodies only enrich for methyl-arginines within the context of the epitope(s) they were raised against, such as the RG motif. Since certain PRMTs tend to preferentially methylate different motifs (CARM1 favours PGM motifs, and PRMT7 prefers RxR motifs), this can generate a bias in the discovery of methylated proteins by neglecting certain target consensus sites[25,97]. More importantly still, within the context of studying epigenetic marks a mass spectrometry-based approach is unable to deliver the most crucial piece of information: the locations of the marks relative to the genome. Protein reader domains have been used for discovery of interactors bearing their target modification - for example the SH2 domain has been used to profile phosphotyrosine proteins[98]. This would in theory open up the prospect of

using TDRD3 (the reader of H4R3me2a) to profile H4R3me2a across the epigenome, however TDRD3 also binds a number of other PRMT1-methylated substrates[99], which could add 'noise' to sequencing data by co-eluting DNA that is not marked by H4R3me2a, but associated with other asymmetrically dimethylated proteins.

Consequently, whilst several genes are known to be transcriptionally up-regulated by PRMT1-mediated deposition of H4R3me2a, little analysis of this phenomenon across the genome has been undertaken. This is primarily attributable to a historical lack of available ChIP (chromatin immunoprecipitation) grade antibodies, as the epitope presented by H4R3me2a is relatively discreet, altering only the mass and not the charge of the terminal arginine guanidino groups[100]. Further, the N-terminal tail of histone H4 is rich in lysine residues that are formaldehyde-reactive. Thus, traditional fixation methodologies could potentially alter the structure (and thus, the epitope recognised by the anti H4R3me2a antibody) of the H4 tail upon formaldehyde fixation. Hence, knowledge of the global contributions of PRMT1 to the histone code remain largely obscured, and to date there exists no relevant data in ENCODE. This obstacle is pertinent to cancer research, as such a 'blind spot' in our knowledge of PRMT1-mediated transcription could be obscuring vital contributions to the cancer phenotype.

1.7.2: Bio-orthogonal profiling: Initial development

Bio-orthogonal profiling is one technology with the potential to mitigate the limitations of commercially available antibodies. By utilising engineered enzyme-cofactor pairs to deposit synthetic post-translational modifications onto substrates of the enzyme of interest, an enzyme's substrate profile can be studied. This is facilitated by modifying recipient proteins with chemical groups that are more amenable to detection than some more chemically 'subtle' modifications that occur endogenously, such as methylation. Interest in matching protein methyltransferases (PMTs) with specific synthetic cofactors is documented as early as 2001, where an engineered SAM analogue was used to inhibit activity of Rmt1 (the yeast homologue of PRMT1) mutant E117G but not the wild-type form of the enzyme, and another N^6-benzyl SAM analogue was demonstrated by the E117G mutant to label substrates in a cell-free reaction[101]. This innovation was based on the 'bump and hole' strategy developed by the Shokat group to screen the function of protein kinases. By mutating key residues that interact with the endogenous cofactor to smaller variants, an artificial 'hole' or binding pocket is created within the enzyme binding site. Thus, a specifically engineered cofactor could interact only with this mutant form of the enzyme, due to additional bulk of the cofactor generated by inclusion of an extra chemical group constituting the 'bump'[102].

Despite these vital preliminary advances, several key obstacles remained. Firstly, transfer of these principles onto human PRMTs was not possible at the time since the crystal structure of PRMT family members had yet to be solved, making structural predictions for generation of the 'hole' element challenging. Secondly,

the activity of the engineered Rmt1-E117G towards its N^6-benzyl SAM analogue produced a k_{cat}/K_m that was 500-fold lower than activity of native Rmt1 towards native SAM. Such low activity between an engineered enzyme-cofactor pair poses a considerable obstacle to eventual modification of target substrates *in vitro,* as this activity will be largely out-competed by endogenous analogues. Further, the technology was marred by a lack of complete substrate specificity between the enzyme and cofactor, with Rmt-E117G still exhibiting activity (albeit reduced to 25% of that of wild type Rmt) towards native SAM, preventing this enzyme-cofactor pair from being truly 'bio-orthogonal'. Finally, use of these synthetic cofactors *in vitro* was restricted by their poor membrane permeability[101].

The next wave of advances on the way to bio-orthogonal PMT-cofactor pairing came in the mid 2000s, through engineering SAM cofactors that could be utilised by native DNA and RNA methyltransferases. These cofactors were termed methyl-surrogate tags (m-tag), due to the methyl moiety of SAM having been replaced by numerous alkynyl variants. As such, target residues could be chemically modified with various alkynyl moieties, offering an application as chemical reporters[103]. However, it was not until the end of the decade that application of these sulfonium-β-sp1/sp2-alkyl SAM analogues was applied to protein methyltransferases, when cofactor EnYn-SAM was demonstrated to have activity towards Dim-5 (a yeast lysine methyltransferase), as well as the human lysine methyltransferases MLL and MLL4. Importantly, these modified enzymes were able to akynylate histone H3, thereby allowing addition of a biotin-azide group via CuAAC (Cu^I-catalyzed azide-alkyne 1,3-dipolar cycloaddition) 'click

chemistry'[104] **(Figure 1.5 A)**. Thus, early developments in surrogate modification of enzymatic substrates provided a foundation for future development of profiling epigenetic marks.

1.7.3: Bio-orthogonal profiling of the PRMT1 epigenome

In 2011, the Luo group applied bio-orthogonal technology to the protein arginine methyltransferase family. Using a high-throughput fluorogenic assay, wildtype PRMT1, PRMT3 and CARM1, and the lysine methyltransferases SUV39H2, SET7/9, SET8, G9a and GLP1, were all shown to be inert towards the bulky alkyl/azide SAM cofactors EnYn-SAM, Hey-SAM, and Pob-SAM[105]. In contrast, a G9a enzyme mutated at F1152A was able to utilise Hey-SAM, and another mutation (Y1154A) permitted activity towards both Hey-SAM and EnYn-SAM. These proof of principle studies demonstrated the potential of bio-orthogonal pairing and substrate labelling for protein methyltransferases[105].

Subsequently, a bio-orthogonal coupling between PRMT1-Y39F/M48G and 4-propargyloxy-but-2-enyl-SAM (Pob-SAM) was demonstrated (PRMT1 nomenclature is in reference to PRMT1v3 that had been N-terminally truncated to resemble PRMT1v1). From hereon, the PRMT1 bio-orthogonal mutant will be referred to as PRMT1-Y29F/M38G, according to the nomenclature of PRMT1v1. In order to find this pairing a panel of synthetic SAM cofactors were generated, substituting the methyl group with various iterations of the alkynyl moiety. These cofactors were tested in cell-free reactions against a panel of PRMT1 mutants in which key residues within the SAM binding pocket that interact with the SAM

sulfonium moiety, had been replaced with smaller substitutes to reduce steric incompatibility with bulkier SAM analogues. Alkynylation of recombinant histone H4 was detected by addition of a TAMRA-azide probe to the alkynylated histone via CuAAC and subsequent in-gel fluorescence. This principle was transferred to the hypomethylated protein lysate of HEK-293T cells for semi cell-free alkynylation (using substrate generated from mammalian cells for cell-free modification by recombinant enzyme), again demonstrating efficacy of recombinant PRMT1-Y29F/M38G via in-gel fluorescence. Importantly, this mutant-cofactor pairing was deemed to be truly bio-orthogonal, as wild-type PRMT1 demonstrated minimal activity towards Pob-SAM when labelling H4 peptide, and PRMT1 Y29F/M38G could not utilise native SAM, as determined my mass spectrometry[106]. Further, fluorogenic assay analysis determined Pob-SAM to be inert towards a panel of 8 protein methyltransferases (PRMT1, PRMT3, CARM1, SUV39H2, SET7/9, SET8, G9a and GLP1), suggesting that Pob-SAM would only be usable by enzymes with purpose-engineered binding pockets[105].

Although the labelling of whole-cell lysate with engineered recombinant PRMT1 and Pob-SAM has been used to discover novel PRMT1 interactors, application of alkynyl modifications to proteins in a semi cell-free system removes these interactions from their cellular contexts[78,107]. For example, PRMT1v1 is predominantly nuclear in its subcellular localisation, suggesting that any novel interactions with this isoform are likely occurring in the nucleus[59]. As such, alkynylation after cellular lysis is incompatible with ChIP-seq, as it prevents representative modification of histones within the normal functioning context of chromatin.

To overcome this limitation, the Luo group moved this technology into an *in vitro* context. Pob-SAM exhibits poor membrane-permeability and so cannot be directly supplemented to cells within the culture medium[108]. As a solution, methionine (the chemical precursor to SAM) was modified to produce the synthetic alkynyl variant S-4-propargyloxy-but-2-enyl homocysteine, called Pob-methionine (Pob-met). Within the cell, methionine is converted to SAM by an ATP-dependent process catalysed by the methionine adenosyl transferases (MAT) enzymes. The Luo group identified key conserved residues that constitute the methionine binding pocket in MAT2A by alignment with *E.coli* enzyme MAT, and using the analogous bump and hole principle, generated a library of MAT2A binding pocket mutants. Of these, MAT2A-I117A emerged as the most efficient at converting Pob-methionine to Pob-SAM.

Thus, in principle, MAT2A-I117A is able to convert Pob-met to Pob-SAM within cells, which is then utilised by PRMT1-Y29F/M38G in order to catalyse alkynylation of substrates, including Histone H4R3 **(Figure 1.5).** Consequently**,** alkynylation of Histone H4 provides a surrogate representation of the transcription-activating H4R3me2a mark. Isolation of alkynylated Histone H4R3 is made possible through CuAAC-catalysed addition of a biotin-azide group **(Figure 1.6 A)** enabling streptavidin affinity enrichment **(Figure 1.6 B)**. Importantly, the inclusion of native lysis conditions, rather than formaldehyde fixation, will enrich for histone alkynylation events. Non-histone proteins that are also targeted for alkynylation by PRMT1-Y29F/M38G will not be isolated due to their lower affinity association with chromatin. Ultimately, sequencing of the associated DNA is considered a representative of histone H4R3me2a

occupancy. The only theoretical caveat this system presents is potential co-enrichment of regions associated with alkynylated H2AR3. Although uncharacterised *in vitro,* PRMT1 methylates R3 of Histone H2A (which shares a 5 amino acid stretch at its N-terminus that is identical to H4) in a cell-free reaction[109]*.* As such, PRMT1-mediated alkynylation of histones may potentially profile two histone marks simultaneously.

Bio-orthogonal profiling of histone marks (termed “CliEn-seq”, or Clickable Chromatin Enrichment with parallel DNA sequencing) has so far been demonstrated through pairing of lysine methyltransferase mutants G9a-Y1154A and GLP1-Y1211A with MAT2A-I117A and Hey-SAM[108]. Thus, published demonstration of this technology’s effectiveness holds promise for interrogating the PRMT1 epigenome in breast cancer.

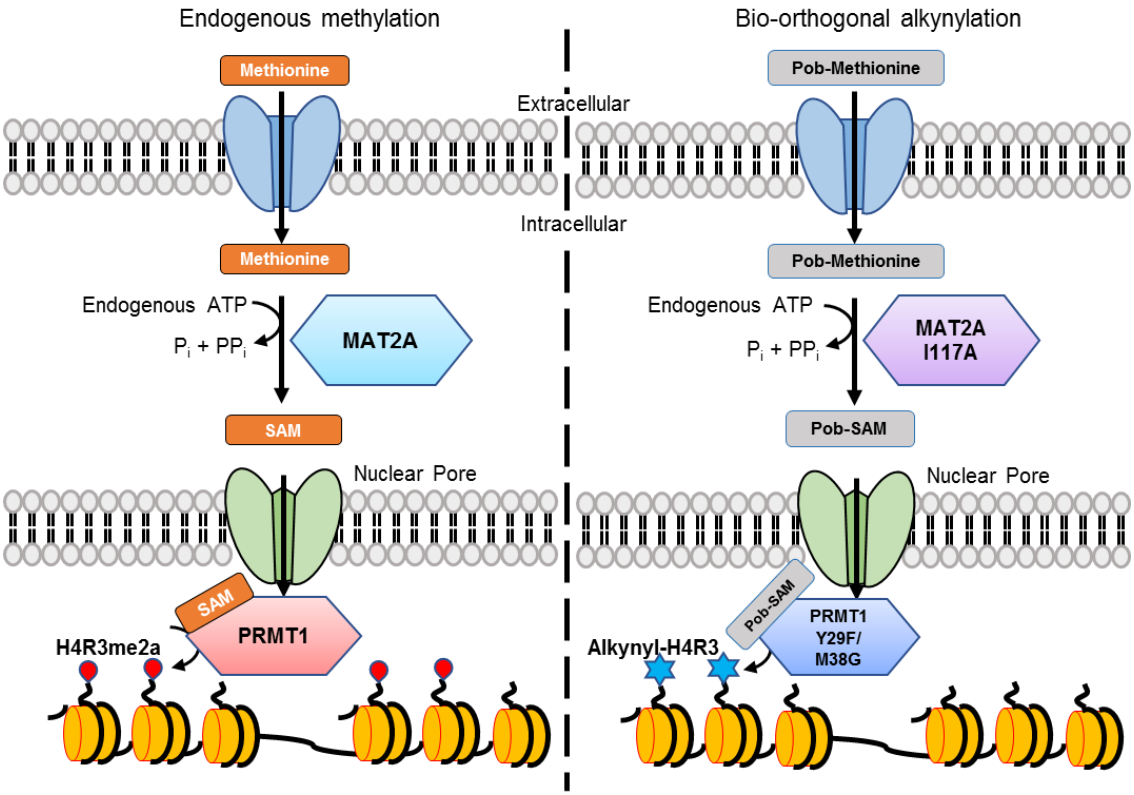


Figure 1.5: Schematic of PRMT1-depdenent alkynylation of histone H4 *in vitro*

Left hand panel depicts the endogenous cellular methylation pathway. Methionine is imported across the cell membrane and converted to the universal methyl donor S-adensyl methionine (SAM) in the cytoplasm by methionine adenosyl transferase (MAT) enzymes with the use of ATP. SAM can then enter the nucleus via nuclear pores, where it is utilised by PRMT1 to asymmetrically dimethylate arginine 3 of histone H4 (H4R3me2a). Right hand panel depicts the synthetic bio-orthogonal alkynylation pathway. A synthetic analogue of methionine, Pob-methionine is also membrane-permeable. It is converted in the cytoplasm by the bio-orthogonal MAT mutant to alkynyl donor Pob-SAM with the use of ATP. Pob-SAM then passes through the nuclear pore where PRMT1-Y29F/M38G can use it to apply the surrogate alkynyl moiety to arginine 3 of histone H4 (alkynyl-H4R3).

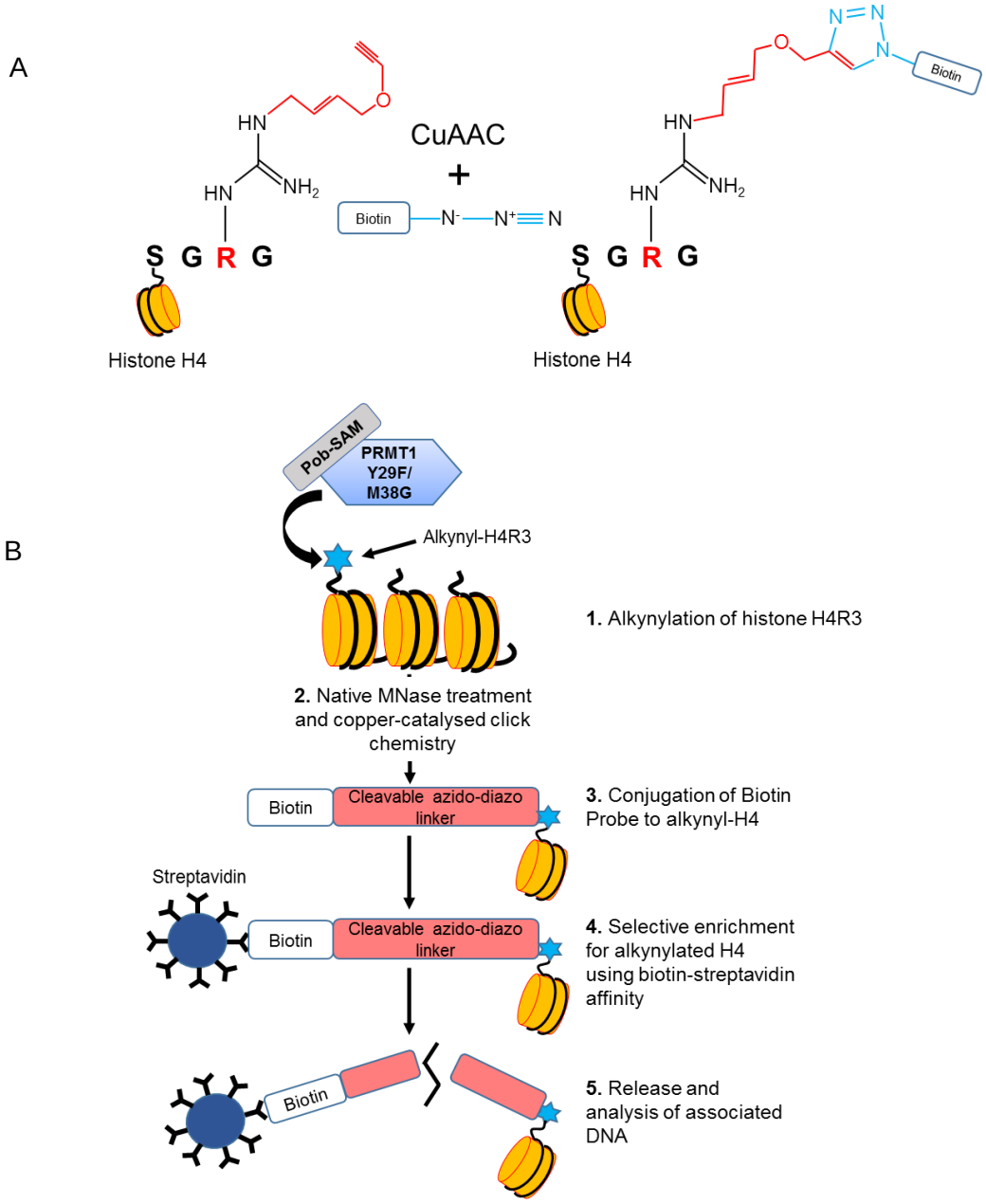


Figure 1.6: Copper-catalysed biotinylation of alkynylated histone H4 and isolation of DNA for CliEn-seq

(A) Chemical rendition of Histone H4 modifications during alkynylation (red atoms) of arginine 3 and subsequent addition of biotin-azide (blue atoms) to the alkynyl moiety via copper-catalysed click chemistry (CuAAC). Adapted from Wang et al., 2011[106]**. (B)** Schematic of bio-orthogonal native ChIP-seq (CliEn-seq). After alkynylation of histone targets, alkynyl H4R3 can be biotinylated via copper-catalysted click chemistry. Alkynyl-H4R3 enriched regions can then be isolated via biotin-streptavidin affinity and associated genomic regions eluted and subsequently sequenced.

*The following section details current functional knowledge on the gene *SPIN1* and its protein product SPIN1, which was identified as a breast cancer-enriched interactor of PRMT1 by SILAC quantitative proteomics. SPIN1 was selected from a cohort of thirty breast cancer-enriched interactors of PRMT1 for further characterisation of its relationship with PRMT1 in breast cancer **(please see results section 4.1 onwards).**

1.8: Introduction to SPIN1

SPINDLIN1, or *SPIN1*, is a gene originally described as a highly abundant maternal transcript in the mouse embryo (contributing 0.35% of all transcripts at the two cell stage), and named for its association with the meiotic spindle[110]. In humans, *SPIN1* encodes a globular 29 kDa protein, consisting primarily of 3 tandem Tudor-like SPIN/SSTY domains (these domains define the *spin/ssty* family of proteins, **Figure 1.7 A)** that exists as a homodimer *in vitro* **(Figure 1.7 B**)[111]. In addition to the 3 SPIN/SSTY domains (from here on referred to generically as 'Tudor domains'), SPIN1 also features an unstructured N-terminal tail consisting of amino acids 1-49. It is from this region that SPIN1 derives its subcellular localisation: a basic amino acid patch (28-44) containing a nucleolar localisation signal (NoLS) is bound by nucleolar transporter B23 (NPM1). In humans, nucleolar SPIN1 has been observed in HeLa cells, SW480 and HEK-293T cells, suggesting that this phenomenon is consistent across multiple tissue types[112]. Indeed, SPIN1 is detected across a plethora of tissue types in humans **(Figure 1.8)**.

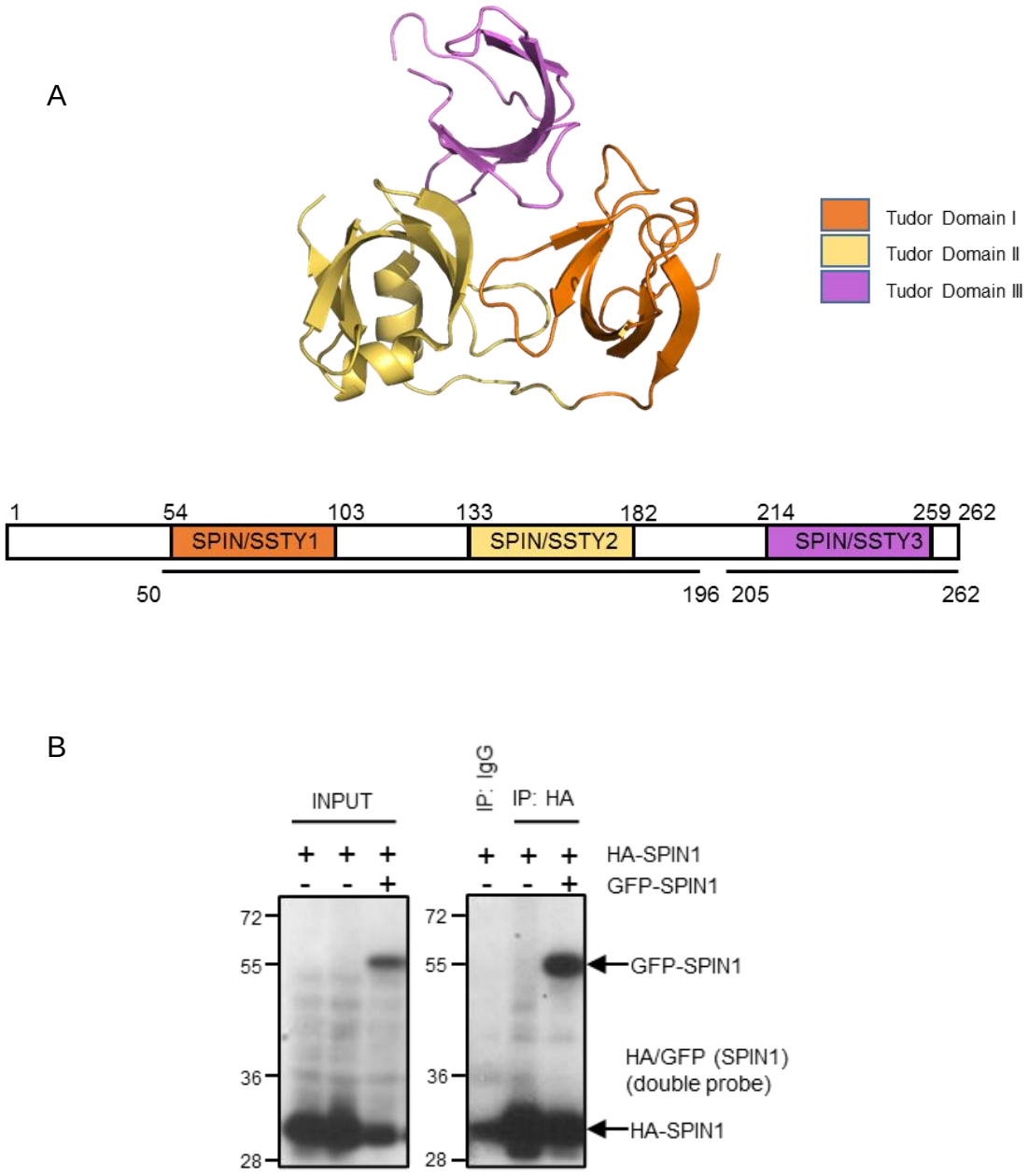


Figure 1.7: SPIN1 monomeric crystal structure and homodimersation *in vitro*

(A) Top panel: Crystal structure depicting monomeric human SPIN1 derived from 2NS2 (PDB DOI: 10.2210/pdb2NS2/pdb). Lower panel: Linear representation of SPIN1 primary structure. Numbers depict amino acid sequence flanking each key segment of the structure. Broken lines and numbers underneath represent how much of the amino acid structure is solved in 2NS2**. (B)** Co-immunoprecipitation of GFP-tagged SPIN1 and HA-tagged SPIN1 24 hours after transfection into HEK-293T cells corroborates *in vitro* dimerization of SPIN1.

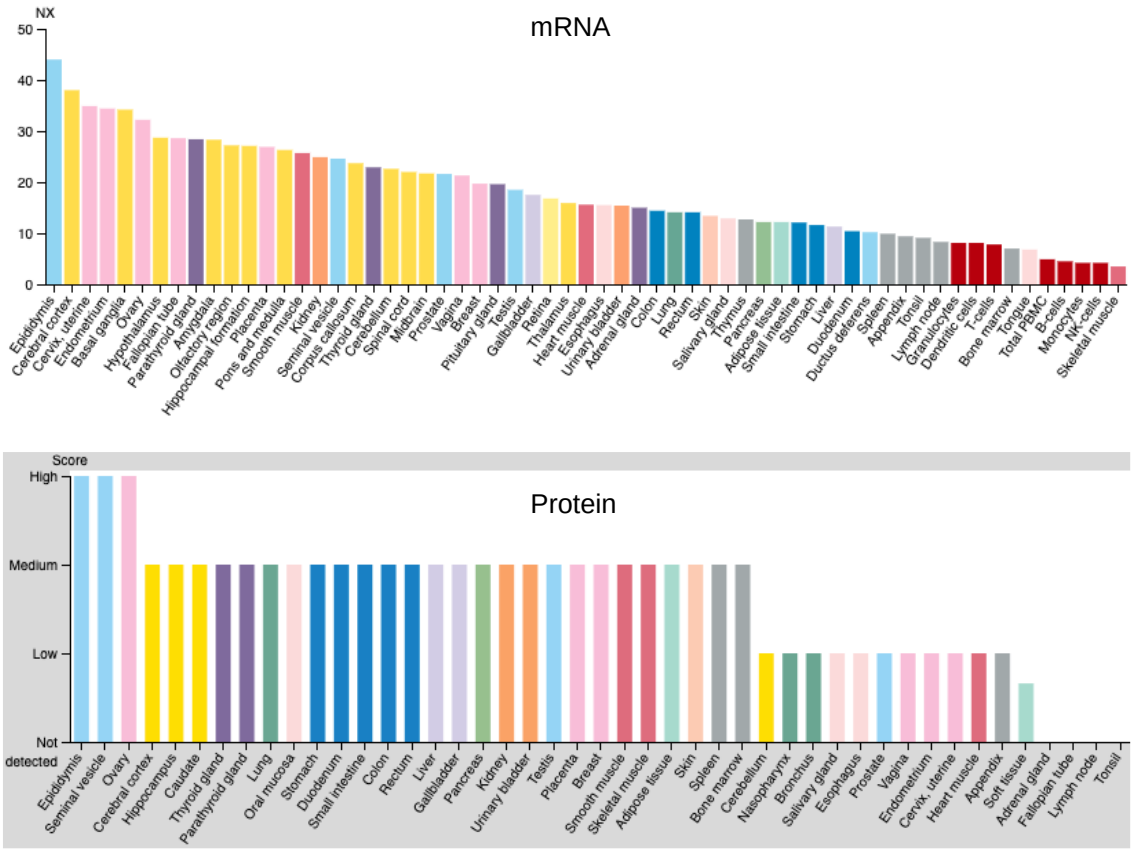


Figure 1.8: SPIN1 is expressed across a wide range of tissue types

Tissue specific expression of human SPIN1 in both mRNA (top panel) and protein (lower panel) ranked from high to low (derived from The Human Protein Atlas). mRNA data are a consensus of HPA, GTEx and FANTOM5 datasets. NX = normalized expression. Colours indicate tissue type.

1.9: Cellular functions of SPIN1

1.9.1: The cell cycle

Though its range of functions remains relatively uncharacterised, SPIN1 is implicated in several vital cellular processes. In keeping with its discovery in gametes, SPIN1 positively regulates meiotic resumption in prophase I mouse oocytes. Mechanistically, it is thought to accomplish this through interaction with RNA binding protein SERBP1 (mediated predominantly through Tudor domain II), which in turn maintains maternal transcript stability of cAMP-degrading enzyme PDE3A. Thus, maintenance of PDE3A expression lowers cAMP levels, facilitating meiotic resumption[113]. SPIN1 also plays an important role in metaphase I to anaphase transition in porcine oocytes, possibly by regulation of *BUB3* mRNA and protein stability[114]. In addition, SPIN1 influences the later-stages of porcine oocyte meiosis, where maintenance of ERK1/2 expression and positive regulation of $p34^{cdc2}$ (the latter being the catalytic subunit of MPF, which in turn maintains CDC2 and cyclin B1 levels) maintain metaphase II arrest[115].

1.9.2: Development

SPIN1 is essential for viability in the mouse, with one-day postnatal lethality apparent in SPIN1-knockout mice. When the deletion was instead specified to myoblast precursors, aberrant foetal myogenesis was reported. This was attributed to deregulation of basic helix-loop-helix (bHLH) transcription factors during early foetal myogenesis, and ultimately de-regulation of genes involved in muscle fibre development, glycogen metabolism and neuromuscular junction/excitation-contraction coupling[116].

1.9.3: Gene regulation

So far, the best-characterised function of SPIN1 is in epigenetic modulation of transcription, an activity mediated by binding of its Tudor domains I and II to histone marks H3R8me2a and H3K4me3, respectively (the latter interaction possessing the higher affinity of the two)[47]. Despite some structural similarity, Tudor domains I and II exhibit specificity in target binding as a result of different contact residues within their aromatic cages that electrostatically and sterically discriminate between targets **(Figure 1.9 A)**. It is presumably for this reason that a function has yet to be ascribed to domain III, as it lacks an intact aromatic cage[47,48]. The histone code reading ability of Tudor-domains I and II have been suggested to operate in a simultaneous (bivalent) manner, forming a consistent groove along the tertiary structure of the SPIN1 protein that is capable of accommodating a histone tail peptide modified with both a trimethylation of H3K4 and asymmetric dimethylation of H3R8 **(Figure 1.9 B)**. Further, this bivalent interaction has been shown to positively regulate Wnt target gene expression **(see section 1.10.2)**[47].

In HeLa cells, SPIN1 co-localises with H3K4me3 at promoters of actively transcribed ribosomal DNA (rDNA) loci, promoting gene expression. Partial abrogation via mutations to the aromatic cage of Tudor domain II suggests that this is mediated at least in part by its histone code reading ability[117]. More recently, histone marks H4K20me3 and H4R23me2a were discovered as cell-free binding targets of Tudor domains II and I respectively, albeit at a significantly lower affinity than the aforementioned H3-related marks. Further, these interactions do not appear to be a bivalent event. Although binding of SPIN1 to

H4K20me3/R23me2a occurs with five-fold higher affinity compared to singular binding of Tudor domain I to H4R23me2a, bidentate interaction (H4K20me3/R23me2a) has an eight-fold lower affinity than singular binding of domain II to H4K20me3[118].

SPIN1-mediated transcriptional modulation has also been implicated in antiviral defence. SPIN1 localises to episomal viral cccDNA (covalently closed circular DNA) of the DNA viruses Herpes Simplex Type I and Hepatitis B viruses, and represses transcription of viral DNA within the nucleus. Although this recruitment is dependent on Tudor domain II of SPIN1, the transcriptionally repressive consequence (and inverse correlation between SPIN1 and H3K4me3 occupancy) suggests that Tudor domain II is fulfilling this function via interaction with another protein that is also recruited to cccDNA.

In the case of HBV, this repressive activity is counteracted by binding of viral protein HBx, which binds to SPIN1 and, through interaction with HBx binding partner ubiquitin ligase Cul4A/DDB1, may alter SPIN1 localisation or target a critical SPIN1 binding partner for degradation[119]. Of particular note, HBx regulates HBV viral transcription via inhibition of other chromatin modifiers including PRMT1. Similarly to SPIN1, PRMT1 also inhibits HBV transcription, a phenomenon that is dependent on its catalytic activity (although the precise application of this activity is not yet elucidated) and is also counter-acted by direct binding of HBx[43].

Attenuation of SPIN1-mediated transcriptional modulation has also been observed in HEK-293T cells through a high-affinity interaction with SPIN1 docking protein (SPIN-DOC), which negatively regulates downstream target genes of the Wnt signalling pathway[120].

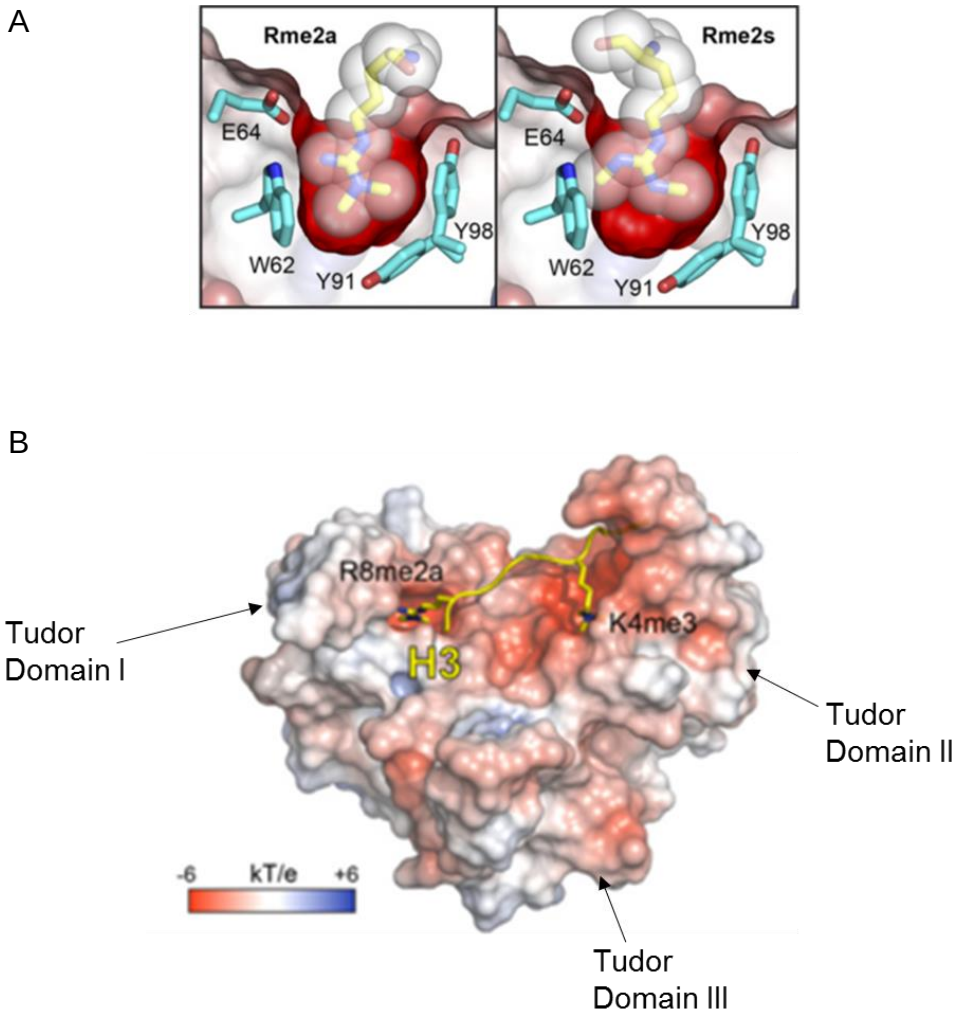


Figure 1.9: Differential structure of SPIN1 Tudor domains dictates substrate binding specificity

(A) Aromatic residues of Tudor domain I are sterically compatible with asymmetric but not symmetric dimethyl-arginine. **(B)** Space fill model of SPIN1 crystal monomer (2NS2) bound to bidentate histone H3 K4me3/R8me2a peptide. Colour coding (KT/e) represents electrostatic potential. Taken from Su et al., 2014[47].

1.10: SPIN1 in Cancer

1.10.1: Background

Considering the relative infancy in its characterisation, documentation of SPIN1-mediated oncogenicity is rapidly increasing. Initially identified as a highly expressed gene in ovarian carcinoma, oncogenic driving capacity of SPIN1 was subsequently demonstrated by its ability to transform NIH-3T3 cells and cause tumour formation in nude mice when overexpressed[121,122]. Malignant inception by SPIN1 dysregulation may be rooted at least partially in its capacity to cause genomic instability, resulting in chromosome mis-segregation, chromosome breakages and multinucleation, all of which are known oncogenic events[123,124].

1.10.2: Wnt Signalling

At the level of molecular oncology, SPIN1 is perhaps most associated with activation of the Wnt signalling pathway whose downstream effectors promote proliferation and invasiveness. SPIN1 has been shown to directly bind to Wnt pathway transcription factor TCF-4 in HeLa cells, an interaction partially mediated by S109 and S124 of SPIN1. Both serine residues are targets of serine/threonine kinase Aurora-A in a cell-free kinase assay, a protein that also co-localises with SPIN1 *in vitro*[125]. In addition, SPIN1 co-localises with H3K4me3 and H3R8me2a (deposited by MLLc and PRMT2, respectively) at the promoters of Wnt target genes *CYCLIN D1, C-MYC* and *AXIN2* and positively regulates the transcription of these genes in HCT116 human colon cancer cell lines[47].

1.10.3: PI3/AKT Signalling

In addition to Wnt signalling, an increasing number of reports suggest that SPIN1 promotes expression of a number of genes linked to PI3/AKT signalling. SPIN1 has been documented as a regulator of RET signalling (a known stimulator of PI3/AKT signalling[126]) in liposarcoma by binding to transcription factor MAZ (Myc-associated zinc finger protein) and co-localizing to the promoter of *GDNF* to positively regulate its expression. GDNF interacts with RET, causing it to autophosphorylate and recruit proteins that activate effectors in RAS-MAPK and PI3/AKT pathways, subsequently increasing cell proliferation and promoting survival[127].

SPIN1 also positively regulates PI3/AKT signalling in glioma, where reduction of SPIN1 levels (in response to miR-489 overexpression, see **Figure 1.10)** reduces levels of phospho-AKT, a marker of active AKT signalling. This in turn reduces transcription of downstream effectors in the pathway such as *CYCD1* and *BCL-XL*, both of which have an anti-apoptotic effects[128]. A similar event has also been observed in seminoma, where SPIN1 antagonises apoptosis through up-regulation of *CYCD1*[129]*.* Further, SPIN1-mediated PI3/AKT signalling has been implicated in enhancing growth and invasiveness in melanoma *in vivo* as well as enhancing chemoresistance in breast cancer, although the exact mechanisms by which SPIN1 accomplishes this have yet to be elucidated. Both of the aforementioned phenomena are modulated by non-coding RNAs that antagonise SPIN1 expression (see **Figure 1.10**)[130,131].

1.10.4: Potential as a Therapeutic Target

The rapidly-mounting evidence that aberrant SPIN1 behaviour drives malignancy has made it an attractive therapeutic target. Endogenous inhibition of SPIN1 expression by various micro-RNAs (see **Figure 1.10** for full list of non-coding RNA regulators of SPIN1) sensitises ER^+ and triple-negative breast cancers to a range of chemotherapeutics, and oesophageal squamous cell carcinoma cells (ESCCs) to radiotherapy[131–133].

Targeting SPIN1 activity may be triaged in future therapies according to patient tumour heterogeneity. In HCT116 human colon cancer cells for example, knockdown of SPIN1 was significantly more detrimental to viability (*in vitro* and *in vivo*) of *TP53* wild-type cells than in *TP53* null (*TP53*$^{-/-}$) cells. However, as SPIN1 knockdown also reduces viability of the *TP53*$^{-/-}$ cells, it is likely that an uncharacterised p53-independent mechanism also exists. With pertinence to p53 dependence, SPIN1 facilitates apoptotic resistance through MDM2-mediated degradation of p53. Here, SPIN1 sequesters uL18 (universal Large ribosomal subunit protein 18) to the nucleolus by binding it directly, preventing it from binding cooperatively with uL5 to MDM2 in the nucleoplasm. Thus, MDM2 is left to exert its E3 ubiquitin ligase activity upon p53, targeting it for degradation and reducing p53 levels when SPIN1 levels are high.

Due to interaction between uL18 and SPIN1 requiring an intact SPIN1 Tudor domain II, and numerous aforementioned oncogenes requiring domains I and II for SPIN1-mediated promotion of their transcription, targeting these domains with

small-molecule inhibitors has become an effort of several groups. Consequently, the Bedford and Oppermann groups have generated two bidentate compounds, designated EML631 and VinSpinIn, respectively[134,135]. Both compounds have been validated *in vitro,* where treatment of cells with either compound or SPIN1 siRNA yielded overlapping RNA-seq gene cohorts between compound and siRNA. However, VinSpinIn treatment was noted to only have a minor effect on cell viability[135].

Despite their demonstrated efficacy, these compounds still have room for improvement. For instance, VinSpinIn has been tested against a more rigorous panel of methyl reader and writer proteins than EML631 and features a potent IC_{50} of 33nM under cell-free conditions towards SPIN1, but still lacks specificity within the SPIN protein family, most notably showing a greater affinity for SPIN3 than SPIN1[134,135]. The relatively large size of bidentate inhibitors may hinder their progression in drug development, as lower molecular weight inhibitors allow greater margin for error when undergoing structural optimisation to navigate obstacles such as low potency, oral administrative suitability and membrane permeability.

To confront this issue, the Jin group have developed MS31, a small molecule inhibitor approximately half the size of the bidentate inhibitors. This small size is attainable by only targeting SPIN1 Tudor domain II, and this compound has shown minimal activity (IC_{50} >50μM) towards arginine and lysine methyltransferases and acetyltransferases, as well as improved selectivity

against other SPIN family members (presenting K_D values of 611nM, 1.66μM and 171nM for SPIN2A, SPIN3 and SPIN4, respectively, versus 91nM for SPIN1[136]). Taken together, the recent but rapid development in small molecule tools to probe SPIN1 function coincides with increasing implication of SPIN1 in cancer. As such, these molecules may play an invaluable role in future characterisation of SPIN1 function, as well as providing a foundation for development of cancer therapeutic drugs that target SPIN1.

RNA Designation	Cancer(s) Documented	Mechanism of Action	Functional Consequence(s)	Reference
LINC00473	Oespophageal Squamous Cell Carcinoma	Interacts with miR-374a-5p to prevent repression of SPIN1 expression.	Promotes SPIN1 expression, which correlates with radioresistance.	133
LncRNA MHENCR	Melanoma (A375, Sk-MEL-2)	Interacts with miR-489 to prevent repression of SPIN1 expression.	Promotes SPIN1 expression and subsequently PI3/AKT signalling, increasing cell migration and invasiveness.	130
miR-29b-1-5p	Triple Negative Breast Cancer (MDA-MB-231, BT-20)	Negatively regulates SPIN1 expression (binds 3' UTR of SPIN1 transcripts).	Reduction of SPIN1 levels reduces pro-invasive/anti-apoptotic behaviour. Also down-regulates genes associated with stemness *OCT4, SOX2* and *NANOG* through Wnt/AKT pathways, decreasing self-renewal. Increases sensitivity to Paclitaxel by reducing SPIN1.	137
miR-148/152	ER^+ Breast (MCF7)	Negatively regulates SPIN1 expression (binds 3' UTR of SPIN1 transcripts).	Reduces resistance to chemotherapeutic Adriamycin (doxorubicin) by suppressing SPIN1-mediated expression of drug-metabolizing enzymes CYP2C8, UGT2B4 and UGT2B17 and drug efflux transporter ABCB4.	132
miR-374a-5p	Oespophageal Squamous Cell Carcinoma	Negatively regulates SPIN1 expression (binds 3' UTR of SPIN1 transcripts).	Represses SPIN1 expression, which correlates with radiosensitivity.	133
miR-489	Glioma (U251, U87) ER^+ Breast (MCF7)	Negatively regulates SPIN1 expression (binds 3' UTR of SPIN1 transcripts).	Reduction of SPIN1 levels reduces pro-invasive/anti-apoptotic PI3/AKT signalling. Reduction of SPIN1 levels decreases adriamycin resistance, invasiveness and apoptosis resistance by downregulating PI3/AKT genes *PIK3CA, AKT1/2/3, CREB1* and *BCL2*.	128
miR-1271	ER^+ (MCF7), Triple Negative Breast Cancer (MDA-MB-453)	Negatively regulates SPIN1 expression (binds 3' UTR of SPIN1 transcripts).	Reduces SPIN1 and B-Catenin levels, reducing motility and proliferation.	138

Figure 1.10: Non-coding RNAs that target SPIN1 and their downstream effects

1.11: Thesis Objectives

PRMT1 is increasingly characterised as an oncogene across many cancer types, including breast cancer. However, the highly pleiotropic nature of the enzyme suggests many unexplored avenues through which it imposes its oncogenic influence. In appreciation of this, the objectives of this thesis are divided into two distinct areas of focus. They are as follows:

i) PRMT1 is important for breast cancer and is known to modulate gene expression in numerous cell types through methylation of the histone H4 tail. However, the genome-wide pattern of this modification and downstream consequences are poorly understood. This phenomenon is being investigated within the context of breast cancer, under the hypothesis that the distribution of H4R3me2a differs between normal mammary epithelial cells and breast cancer cells – and that this difference drives the breast cancer phenotype. In order to address this, we will employ bio-orthogonal profiling as a substitute to conventional antibody-directed chromatin immunoprecipitation. Ultimately, this technology will be used to identify the distribution of PRMT1-directed epigenetic marks in non-transformed versus breast cancer cells on a genome-wide basis.

ii) PRMT1 interacts with and methylates a plethora of non-histone targets, which is predicted to contribute to downstream pro-oncogenic activity. By using SILAC quantitative proteomics to identify novel breast cancer-enriched substrates of PRMT1, SPIN1 was identified as one such interactor, and chosen for further study. Thus, the objective derived from this work was to characterise the

relationship between PRMT1 and a SPIN1, with a predominant focus on the biochemical consequences of SPIN1-PRMT1 interaction.

Chapter 2: Materials and Methods

2.1: Tissue Culture Techniques

2.1.1: Cell lines

All cell lines used in this study are summarised in **Table 1.**

2.1.2: Tissue Culture Medium

MCF7, HCT116 and HEK-293T cells were cultured in Dulbecco's Modified Eagle's Medium (Sigma) containing 1% v/v Penicillin Streptomycin (Sigma), 1% v/v L-Glutamine (Sigma) and 10% v/v foetal bovine serum (Gibco). MCF10A cells were cultured in DMEM F12 1:1 (Gibco) with 5% v/v horse serum (Gibco), 20ng/μl EGF, 0.5μg/ml (Miltenyi) hydrocortisone (Sigma), 100ng/ml cholera toxin (Sigma) 10μg/ml Insulin (Sigma) and 1% v/v Penicillin Streptomycin (Sigma).

2.1.3: Culture and passage of cells

Cells were incubated at 37^0C with 5% CO_2. Prior to forming a confluent monolayer (approximately 70-90% confluence), medium was aspirated and cells were washed with sterile PBS (Sigma) and subjected to trypsin-mediated detachment at 37^0C. Once detached, the effects of trypsin were quenched with an equal volume of cell medium and cells pelleted via centrifugation at 1000*g* for 3 minutes. Supernatant was then aspirated and cells re-suspended in fresh medium. An appropriate number of cells (depending on imminence of requirement) were then plated onto a fresh cell culture dish and returned to the incubator.

2.1.4: SILAC cell culture

For SILAC, MCF7 cells were cultured in DMEM without L-arginine and lysine (DC Biosciences) and supplemented with 10% v/v dialysed FBS (MWCO 10kDa), 1% v/v L-glutamine, 1% v/v Penicillin Streptomycin and 0.4μM heavy arginine, and 0.76μM heavy lysine (Cambridge Isotope Laboratories, 18L-0063, 17E-093 respectively). MCF10A cells were grown in DMEM F12 1:1 for SILAC (Thermo Scientific) supplemented with dialysed horse serum (Dundee Cell Products Ltd) and 0.4μM light arginine, and 0.76μM light lysine (Sigma, A8094, L8662, respectively). Additional constituents are detailed in **2.1.2.** Cells were cultured in SILAC medium for >7 passages prior to harvest to ensure that they had incorporated either light or heavy amino acids into endogenous proteins. Complete incorporation of heavy isotope was validated by Mass Spectrometry (data not shown, conducted by Dr. Mark Skehel, LMB, Cambridge).

2.1.5: Cryopreservation and recovery of cells

Cells were trypsinised, pelleted and supernatant removed. Cells were re-suspended in a solution containing 80% v/v complete cell medium, 10% v/v Dimethyl Sulfoxide (Sigma) and 10% v/v FBS (Gibco) and cryovials placed at -80^{0}C in an isopropanol cooling box. Cryovials were transferred 48 hours later for long-term liquid nitrogen storage. Cells were collected for recovery on dry ice and thawed for 30 seconds in a 37^{0}C water bath. Once thawed, cells were quickly transferred into a 30ml tube containing 10ml fresh medium and centrifuged for 3 minutes at 1000*g*. Supernatant was then aspirated, and cells re-suspended in the appropriate volume of medium for the relevant culture dish size. Cells were then plated and incubated at 37^{0}C.

2.2: Cell Biology techniques

2.2.1: Transient cell transfection with DNA constructs

All transfection procedures took place approximately 24 hours after plating cells. Cells were plated to a density so that they were approximately 60% confluent upon transfection. For transfection of HEK-293T cells, 5μg or 10μg DNA was added to 500μl or 750μl Opti-Mem (Invitrogen) for 6 or 10 cm dishes, respectively. The solution was then briefly mixed by vortex before the addition of PEI (SIGMA) (diluted 1:1 ratio in Opti-Mem to 0.5mg/ml) at 3 times the volume of DNA added. This solution was vortexed for 10 seconds to mix and incubated at room temperature for 30 minutes before dropwise addition to the relevant cell dishes. For transfection of MCF7 cells in a 10cm culture dish, JetPei™ transfection reagent (PolyPlus) was used, with the same DNA quantities detailed above. DNA was added to 150mM NaCl to a final volume of 250μl, and vortexed briefly. In a separate tube, JetPei™ solution (10 or 20μl for 6cm or 10cm dishes, respectively) was diluted in 150mM NaCl to a final volume of 250μl and also vortexed. The JetPei™-NaCl mix was subsequently added to the DNA-NaCl mix, vortexed for 10 seconds and briefly centrifuged. The resulting mixture was incubated at room temperature for 30 minutes to allow formation of DNA-Jet Pei complexes. After 30 minutes, the entirety of each 500μl JetPei™/NaCl/DNA mix was added to the relevant dish dropwise. Cells were incubated with this mixture at 37°C for 3 hours, after which their media was aspirated and replaced with complete medium. Cells were then incubated under standard conditions detailed above for at least 24 hours before harvesting for experimentation.

2.2.2: Lentiviral infection for generation of stable cell lines

In order to generate pTIPZ-FLAG-SPIN1 MCF7 and MCF10A cell lines, HEK-293T cells were plated at a density of ~3 x 10^6 on 10cm culture dishes and transfected 24 hours later with PEI transfection reagent according to **2.2.1.** 9μg viral DNA, 4.5μg pMD2.G (for generation of viral envelope component VSV-G) and 6.75μg psPAX2 (for production of packaging proteins gag and pol) were transfected simultaneously within the same cocktail. Cell medium was changed six hours later and replaced with 5ml complete medium. 24 hours later, 6-well plates were seeded with either MCF7 or MCF10A cells at a density of ~0.2 x 10^6 cells per well. The following day, viral medium was passed through a 0.45μm filter and polybrene (Sigma) added to viral supernatant at a final concentration of 8μg/ml. Approximately 1.5ml of this mixture was added to each well of the 6-well plate, which was sealed with cling film and centrifuged at 12,000*g* for 90 minutes at room temperature. After centrifugation, viral medium was removed and replaced with complete medium. The process was repeated the following day and infected MCF7/MCF10A cells passaged onto a larger culture dish. 48 hours later, cells were subjected to puromycin selection at 1μg/ml for at least a week before experimental use.

2.2.3: Preparation of cells for Cell Cycle Analysis

One 10cm dish per treatment was set up with a seeding density of 1.5x 10^6 cells. All cells used were pTIPZ-FLAG-SPIN1 (or various mutants thereof) and were kept under puromycin selection (1μg/ml). Doxycycline at a concentration of 1μg/ml was added to induce expression of ectopic FLAG-SPIN1. After 48 hours, medium was aspirated from cells and placed in a 30ml universal tube to save detached mitotic cells. Dishes were washed with PBS which was also added to the respective tube for each treatment. Cells

were then detached via trypsinisation and added to the same tube as their saved medium/PBS wash, and pelleted via centrifugation at 6,000*g* for 5 minutes. The number of cells was standardised between samples by counting with a haemocytometer. Samples were then re-suspended in 400μl cold PBS and transferred to FACS tubes and placed on a slow vortex while 800μl cold ethanol was added dropwise to fix cells. Tubes were then covered with parafilm and stored at 4^0C for up to a week before FACS analysis. Prior to staining for DNA content, cells were washed twice in PBS, and re-suspended in 300μl PBS containing 100μg/ml RNase A and 25μg/ml propidium iodide. FACS was carried out by Miss Elizabeth Dufficy of the Davies Laboratory, using the Texas Red channel to visualise DNA content. Statistical analysis of cell cycle proportions was carried out using FlowJo (BD) and Microsoft Excel.

2.2.4: Drug treatments

For experiments where cells were treated with AdOx (Adenosine periodate, Sigma), the chemical was added to culture medium to a final concentration of 100μM for HEK-293T cells and 20μM for MCF7 pTIPZ-FLAG-SPIN1 cells 24 hours prior to harvest of cells. Type I PRMT inhibitor MS023 was added to cells at a concentration of 1μM for 48 hours prior to harvest of experiments.

2.3: Molecular Biology Techniques

2.3.1: PCR primers

All primers used are detailed in **Table 2**.

2.3.2: PCR amplification of genes of interest for cloning

Generally, constructs or cDNA containing the gene of interest were amplified in a 100μl reaction containing 0.5μM forward and reverse primers, 0.2mM dNTPs, 10μl 10x Accuzyme™ buffer, 2μl Accuzyme™ DNA polymerase (Bioline) and 5μl DMSO. Typical reaction conditions consisted of the following: 95^0C initial denaturation for 3 minutes, 95^0C denaturation for 30 seconds, 60^0C annealing for 30 seconds per kilobase, 72^0C extension for 60 seconds and final 72^0C extension for 10 minutes. Denaturation, annealing and extension were repeated for 35 cycles.

2.3.3: Agarose gel electrophoresis

PCR product amplification, site-directed mutagenesis and restriction digests were visualised on 1% w/v agarose gels diluted in 1x TAE buffer (40mM Tris, 20mM acetic acid 1mM EDTA pH8.0) and stained with 1/10,000 SYBR™ safe DNA gel stain (Thermo). Gel electrophoresis proceeded at 100 volts for approximately 1 hour and was visualised using a GeneFlash UV imager (Syngene).

2.3.4: Restriction digests

All enzymes used were supplied by New England Biolabs. Generally, restriction digests occurred in a reaction volume of 20μl and included 10u of each restriction enzyme, and CutSmart restriction enzyme buffer (New England Biolabs) diluted to 1x. Digests were carried out at 37^0C for one hour-overnight.

2.3.5: Ligation of cDNA into selected vectors

Digested vectors were treated with 1μl calf intestinal phosphatase (CIP, New England Biolabs) for 25 minutes at 37^0C in order to prevent the linearised plasmid

DNA from re-annealing. Digested cDNA and vectors were subjected to agarose gel electrophoresis and then purified using an E.Z.N.A. gel extraction kit (Omega Bio-tek). For ligation, 50ng purified vector DNA was used and cDNA added to either a 1:3 or 1:6 vector:insert quantity. 1μl T4 DNA ligase (New England Biolabs) was added to this mixture, along with T4 DNA ligase buffer, diluted to 1x concentration (New England Biolabs). The ligation mixture was left at room temperature for at least two hours prior to transformation.

2.3.6: Transformation of *E. coli*

NEB 5-alpha *E. coli* (New England Biolabs) were used for bacterial transformations. For recombinant protein generation, BL21 bacteria (New England Biolabs) were used. 50μl cells were thawed on ice and the entire ligation mixture added to the cell suspension. This mixture was incubated on ice for 30 minutes and then heat shocked at 42^0C for 30 seconds, before returning to ice for 2 minutes. An additional 450μl SOC outgrowth medium (New England Biolabs) was added and cells were recovered at 37^0C for at least one hour in a shaking incubator. The entire transformation suspension was then added to LB agar plates containing either kanamycin or ampicillin at concentrations of 50 and 100μg/ml, respectively. Plates were incubated at 37^0C overnight.

2.3.7: Site-directed mutagenesis

All site-directed mutagenesis was carried out with primers **(see Table 2)** generated from the NEBasechanger online tool (https://nebasechanger.neb.com/). The protocol itself (PCR through to bacterial

transformation) was carried out using the Q5 Site-Directed Mutagenesis Kit and associated protocol (New England Biolabs).

2.3.8: RNA extraction from human cells

Cells were harvested via trypsinisation and either stored at -20^{0}C or subjected immediately to RNA extraction using a GenElute mammalian total RNA miniprep kit (Sigma). RNA was then stored long term at -80^{0}C.

2.3.9: cDNA generation from extracted RNA

cDNA was generated from RNA using a SensiFAST^{TM} cDNA synthesis kit (Bioline). Reverse transcription was carried out in a Bio-Rad T100 Thermocycler according to the attached protocol: primer annealing at 25^{0}C for 10 minutes, reverse transcription at 42^{0}C for 15 minutes and inactivation at 85^{0}C for 5 minutes. cDNA was stored long term at -20^{0}C.

2.3.10: Quantitative real-time polymerase chain reaction (qPCR)

cDNA was diluted 1/25 and 2μl per well added in technical triplicate to a 96 well plate. To this, SYBR^{TM} Green 2x PCR master mix (Thermo) and 2μM primer mix (all primers used in qPCR are detailed in **Table 3**) were also added. 96 well plates were loaded into a Stratagene Mx3005P (Agilent Technologies) and quantitative PCR run according to the 'normal 2-step' (1x 10min at 95^{0}C, 40x 30s at 95^{0}C, 30s at 60^{0}C, 1x 60s at 95^{0}C, 30s at 55^{0}C, 30s at 95^{0}C) reaction protocol in MxPro (Agilent Technologies). Primer efficiencies were determined via standard curve

generation for each primer set before use (see **Table 3**). After running each experiment, dissociation curves were analysed to check for single product amplification. Gene expression levels were calculated in Microsoft Excel via the ΔΔCT method[139], normalising to actin levels.

2.4: Protein Biochemistry

2.4.1: Protein lysis and protein quantification

Protein concentration determination was carried out via addition of Bradford reagent (Bio-Rad). Protein was lysed in either 0.2% v/v NP40 (150mM NaCl, 20mM Tris pH7.5, 0.5mM Ethylenediaminetetraacetic acid (EDTA), 1mM Na_3VO_4, 50mM NaF, 1mM β-glycerol phosphate, 100μM phenylmethylsulfonyl fluoride (PMSF), 1μg/ml Leupeptin, 1μg/ml Aprotinin) or RIPA buffer (50mM Tris pH7.4, 150mM NaCl, 1mM EDTA, 1% NP40, 0.5% sodium deoxycholate, 0.1% w/v SDS, 10% v/v glycerol, 1mM PMSF, 50mM NaF, 10mM Na_3VO_4, 1μg/ml Leupeptin, 1μg/ml Aprotinin). Once resuspended, samples were sonicated using a small probe at 25% amplitude for 5 seconds (twice and left to rest on ice for 30 seconds between pulses) and left on ice for 10 minutes. Lysate was then centrifuged at 17,000*g* for 10 minutes and supernatant transferred to new tubes. 2μl protein lysate was added to 2ml cuvettes and diluted in 1ml of 20% v/v Bradford solution. Samples were measured in a spectrophotometer at a wavelength of 595nm. Samples were measured in duplicate and concentrations generated from a pre-determined BSA standard curve in Microsoft Excel.

2.4.2: SDS Poly-Acrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE gels were prepared on their day of use. Acrylamide content ranged from 10-16% v/v depending on the protein(s) under investigation. Constituents of a 10ml gel were 375mM Tris-HCl (pH8.8), 0.1% w/v SDS, 0.1% w/v Ammonium persulfate and H_2O, ProtoGel ultra-pure 30% acrylamide (national diagnostics) and TEMED adjusted in volume depending on the desired percentage. Isopropanol was added during resolving gel casting to remove bubbles and was subsequently removed with distilled H_2O once the gel had set for the addition of the stacking gel (125mM Tris-HCl (pH6.8), 5% v/v acrylamide, 0.1% w/v SDS, 0.1% w/v ammonium persulfate and 0.1% v/v TEMED). Gels were set up in the mini-gel format according to the equipment manufacturer's instructions (Bio-Rad) and gel tank apparatus filled with SDS-PAGE running buffer diluted to 1x (10x stock solution consisted of 0.25M Tris-HCl pH8.3, 1.92M glycine and 1% w/v SDS). Protein samples were typically loaded with a total content of 50μg, (with the exception of SPIN1, which usually used 100μg when detecting with the anti-SPIN1 antibody) and denatured in the appropriate volume of 4x Laemmli buffer (187.5mM Tris-HCl pH6.8, 6% w/v SDS, 30% v/v glycerol, 0.03% w/v bromophenol blue, 150mM DTT, 4% v/v β-mecaptoethanol) by boiling at 95^0C for 5 minutes. For molecular weight reference, PageRuler^{TM} Prestained protein ladder (Thermo) was run alongside samples. Gels were typically run for 90 minutes at 125V (constant voltage).

2.4.3: Immunoprecipitation

Cell quantities for immunoprecipitation when using HEK-293T cells was approximately 8 x 10^6 (or one confluent 10cm culture dish). Immunoprecipitations were carried out by lysing in 500μl lysis buffer (0.2% v/v NP40 for co-immunoprecipitations and RIPA for determination of methylation status). Cells were sonicated as in section **2.4.1**. Samples were then pre-cleared for 2 hours by rotating at 4^0C with 20μl Protein G Sepharose 4 Fast Flow resin (GE Healthcare Life Sciences) that had been pre-washed and diluted 1:1 with the lysis buffer of choice. After pre-clearing, samples were centrifuged at 2,400*g* for 3 minutes at 4^0C and supernatant transferred to new tubes. At this point, 50μl of each sample was taken for input (10% of total sample) and transferred to a new tube and boiled at 95^0C in 10μl Laemmli buffer for 5 minutes. The remaining sample was increased in volume by addition of another 250μl lysis buffer and 20μl 1:1 beads with an affinity for the protein to be immunoprecipitated (see **Table 5**) were added. Samples were then rotated at 4^0C overnight and washed 4 times with 500μl of the appropriate lysis buffer (including protease inhibitors). After the final wash, residual lysis buffer was removed with an insulin needle and beads were boiled at 95^0C for 5 minutes in 40μl Laemmli buffer in order to denature and dissociate immunoprecipitated proteins. For semi cell-free and *in vitro* methylation assays, (sections **2.4.10** and **2.4.11**) SPIN1 was subsequently isolated by FLAG peptide elution. Prior to elution, samples were washed twice in the appropriate lysis buffer, once in a high-salt version (350mM NaCl) of that same lysis buffer and twice in PBS. This was followed by 2x 20 minute elutions on a 4^0C thermoshaker shaking at 1200rpm. Each elution consisted of PBS,

50mM NaCl and 300μg/ml FLAG peptide (Sigma, F3290) and both elutions were combined to give a total volume of 40μl.

2.4.4: Western Blotting

After SDS-PAGE, stacking gel was removed and resolving gel laid onto a sheet of PVDF (GE Healthcare) pre-soaked in methanol and then rinsed in chilled transfer buffer (70% v/v distilled H_2O, 20% v/v methanol, 10% v/v 10x Tris-glycine electroblotting buffer, National Diagnostics) and sandwiched between two outer layers consisting of 3 sheets of Whatman paper (GE Healthcare) soaked in transfer buffer. The sandwich was then loaded into a Mini Trans-Blot Cell (Bio-Rad) and proteins were transferred at 4^0C ambient temperature for 90 minutes at 85 volts. After transfer, membranes were blocked for approximately 1 hour in 5% w/v milk (Marvel) in Tris-buffered saline solution (138mM NaCl, 20mM tris, pH7.6) with 0.1% v/v Tween (TBST) with agitation. When detecting SPIN1 with an anti-SPIN1 antibody, membranes were blocked in 5% w/v bovine serum albumin (Fisher Scientific). After this, membranes were rinsed 3 times with TBST to remove excess milk/BSA and placed in a 50ml tube (Corning) containing the requisite antibody diluted in 5% w/v BSA-TBST (primary antibodies and their working dilutions can be found in **Table 4**). Membranes were then incubated at 4^0C overnight on a roller. The following day, membranes were washed 3 times for 10 minutes in TBST and incubated in horseradish peroxidase-conjugated secondary antibody: either goat anti-rabbit (Dako) or anti-mouse (CST) at 1/5000 concentration or donkey anti-sheep (Sigma) at 1/10,000 in 5% w/v milk-TBST solution for 1 hour at room temperature. After secondary antibody incubation,

membranes were washed another 3 times in TBST and then incubated for 5 minutes in WesternBright ECL Spray (Advansta) or SuperSignal™ West Femto Maximum Sensitivity Substrate (Thermo). Membranes were then exposed to X-ray film to detect protein.

2.4.5: Preparation of SILAC samples for mass spectrometry

Fifteen 15 cm dishes of MCF7 and MCF10A cells stably overexpressing the pHIV-Zsgreen-FLAG-PRMT1v1-XGX construct were grown to near confluency for each experimental repeat. Samples were harvested via trypsinisation and pelleted before lysis and immunoprecipitated as described in **2.4.3,** with the exception that 30μl rather than 20μl 1:1 affinity resin:lysis buffer was used for pulldown. A total of 4 experimental repeats were produced (2 of these were contributed by Dr. Agnieszka Zielinska of the Davies Laboratory). Two experimental repeats were boiled at 95^0C for 5 minutes in 25μl GSB and fractionated via SDS-PAGE using pre-cast 4-12% bis-tris gels (Invitrogen) and stained with SYPRO™ Ruby Protein Gel Stain (Lonza) according to the manufacturer's protocol. Gels were then imaged for future reference using a ChemiDoc mp Imaging System (Bio-Rad) and bands cut out with a scalpel and placed in 300μl Pierce Water (LC-MS grade, Thermo 51140) and stored at 4^0C overnight. For a further two experimental repeats, residual lysis buffer was removed after washing beads and replaced with 20mM TRIS pH7.4 and stored at 4^0C overnight to enable on bead digest (performed by collaborators). Samples were then sent for analysis by Dr. J. Mark Skehel and Dr. Sarah Maslen at the MRC Laboratory of Molecular Biology (Cambridge UK).

2.4.6: Preparation of SPIN1 samples for arginine methylation and interaction analysis by mass spectrometry

SPIN1 was immunoprecipitated from pTIPZ-FLAG-SPIN1 MCF10A and MCF7 cells, as well as HEK-293T cells transiently transfected with pHIV-dTomato-HA-SPIN1. Cells were lysed in 500μl RIPA buffer and SPIN1 immunoprecipitated (see **Tables 4 and 5**) and fractionated via SDS-PAGE in the same manner detailed in section **2.4.5.** The only exception was that SYPRO™ Ruby was replaced with InstantBlue Coomassie Protein Stain (Abcam) due to the need for omission of a methanol-based fixation step with the former stain, which has been attributed to spontaneous methylation of substrates[140]. Bands were then cut out, imaged and sent to Dr. J. Mark Skehel and Dr. Sarah Maslen for tryptic digest and analysis at the MRC Laboratory of Molecular Biology (Cambridge UK).

2.4.7: *In silico* analysis of mass spectrometry data

In silico analysis used Scaffold (Proteome Software) for both SILAC quantitative analysis and arginine methylation analysis of SPIN1. For SILAC-based discovery of breast cancer-enriched interactors of PRMT1, quantitative interactome ratios were provided by Dr. Mark Skehel and Dr. Sarah Maslen (MRC Laboratory of Molecular Biology, Cambridge UK). In order to triage interaction phenomena that were considered ‘true’, several criteria were used. Firstly, the log fold change in PRMT1-substrate interaction had to equal or exceed 1.5 in MCF7 cells over MCF10A cells. Additionally, this phenomenon had to be observed in at least 2 of the four experimental repeats. Finally, Scaffold’s application of the T-test provided statistical significance to these fold-change events. For analysis of

SPIN1 arginine methylation, analysis of post-translational modifications was provided by Dr. J. Mark Skehel and Dr Sarah Maslen. Spectra featuring post-translational modifications (including arginine methylation events) could then be observed through Scaffold. When looking at SPIN1 interactors, band-by-band analysis was employed non-quantitatively to look for interactors of interest that fell primarily within the 55 and 72kDa range.

2.4.8: Generation of recombinant proteins

For all recombinant protein preparations (unless otherwise stated), 2x 800ml flasks of LB broth (Sigma) containing kanamycin or ampicillin at concentrations of 50 and 100μg/ml (respectively) were inoculated with 10ml starter cultures (incubated at 37^0C with shaking overnight). Bacterial growth proceeded at 37^0C with shaking and was monitored at wavelength of 600nm using a Jenway 7310 spectrophotometer. Generation of recombinant protein was stimulated by adding 400μM IPTG once OD reached log growth phase (0.4-0.6). Bacterial cultures were then grown at 20^0C with shaking overnight and pelleted the next morning by centrifugation at 3,000*g* for 10 minutes. Pellets were then either stored at -80^0C or lysed immediately on ice. Pellets were re-suspended in 10ml lysis buffer per 800ml culture. For His-tagged proteins, lysis buffer consisted of 50mM NaH_2PO4, 50mM Na_2HPO_4, 300mM NaCl, 5% v/v glycerol and 10mM β-mercaptoethanol. For GST-tagged proteins, lysis buffer consisted of 50mM Tris (pH8) 100mM NaCl, and 1mM EDTA. Both buffers were supplemented with 1mM DTT, 1mM PMSF and 10μg/ml aprotinin and leupeptin. Once re-suspended, lysozyme (Sigma] was added to a final concentration of 1mg/ml and cells were

sonicated at 65% using a medium size probe four times for 60 seconds (10 seconds on, 10 seconds off). Lysate was then rotated at 4^0C for 15 minutes with 0.5% v/v Triton X100 before centrifugation by a pre-cooled JA-25.5 rotor centrifuge at 48,000*g*, 4^0C for 30 minutes. Supernatant was transferred to new 50 ml tubes and 300μl 1:1 affinity beads: **(see Table 5)** that had been pre-washed in the appropriate lysis buffer were added. Tubes were then rotated at 4^0C overnight. The following day, tubes were centrifuged to pellet beads (500*g* at 4^0C for 5 minutes) and supernatant removed by aspiration. Beads were washed 4 times with 30ml lysis buffer (without DTT or protease inhibitors) for GST-tagged proteins and 30ml lysis buffer with 50mM imidazole for His-tagged proteins. To elute GST-tagged protein, beads were re-suspended in 500μl 30mM Glutathione (pH8) dissolved in water and transferred to a 1.5ml eppendorf tube. Tubes were then rotated at room temperature for 20 minutes and centrifuged at 900*g* for 2 minutes at 4^0C to pellet beads. Eluate was then transferred to a fresh eppendorf tube and the elution process repeated twice more. To elute His-tagged protein, beads were transferred to a 1.5ml eppendorf tube in 400μl ice cold lysis buffer with 300mM imidazole (Sigma). Beads were then vortexed for 5 seconds and incubated on ice for 60 seconds before centrifugation (900*g* for 2 minutes at 4^0C) to pellet beads. This was repeated once more, and eluate transferred to a fresh tube with each elution. For both purification types, eluates were pooled and added to a pre-equilibrated 8kDa Maxi GeBaFlex-tube dialysis chamber (Generon) and left to dialyze at 4^0C in 4L PBS for at least 24 hours. Proteins were then concentrated by centrifugation in Vivaspin2 protein concentration columns (Generon) and 50% v/v glycerol added to make 20% v/v glycerol of total protein volume. Purified protein was stored at -80^0C. Protein concentration was

determined by running a BSA standard curve via SDS-PAGE and using ImageJ (NIH) pixel-counting.

2.4.9: Cell-free methylation of SPIN1 with 3-[H]-S-adenosyl methionine (^{3}[H]-SAM)

Reactions took place in a total volume of 60μl. Three different buffer types were tested: PBS diluted to 1x working concentration, Tris-HCl (20mM, pH7.4, found to be the most effective, (see section **4.3.3**) and 1M HEPES pH7.0-7.6 (Sigma). A full reaction consisted of 1.5μg His-SPIN1, 2μg GST-PRMT1 and 1μCi 3[H]-SAM (PerkinElmer: specific activity 55–85Ci/mmol). 2μg recombinant histone H4 (New England Biolabs, M2504S) was used as a positive control reaction. Aluminium hydroxide (1μM, Acros organics) was also added to certain reactions. Reactions were allowed to proceed overnight (~12 hours) on a benchtop thermoshaker at 37^{0}C, shaking at 800rpm, then boiled for 5 minutes at 95^{0}C with the addition of 20μl Laemmli buffer. Reactions were then resolved via SDS-PAGE using 12 or 16% gels and stained in InstantBlue Coomassie Protein Stain (Abcam) to visualize protein content. Gels were then washed in distilled water for 60 minutes with agitation (changing water every 20 minutes) and incubated in EN3HANCE autoradiography enhancer solution (PerkinElmer) for 1 hour. Gels were then washed another 3x20 minutes in distilled water. Once washed, gels were dried by immersing in Gel-Dry Drying Solution (Life Technologies) and rocking gently for 20 minutes at room temperature. Gels were then prepared and dried using a Novex DryEase Mini-Gel Dryer System (Life Technologies)

according to the manufacturer's instructions. Once dried, gels were exposed to X-ray film and left to develop at -80^{0}C for at least 24 hours before development.

2.4.10: Semi cell-free methylation of SPIN1 with ^{3}[H] S-adenosyl methionine (^{3}[H]-SAM)

In order to extract transfected SPIN1 expressed in human cells, HEK-293T cells were transfected with pCMV-2B-FLAG-SPIN1 and treated 24 hours later with 100μM oxidized Adenosine periodate (AdOx, Sigma) in order to prevent methylation of ectopically expressed SPIN1. When using MCF7 pTIPZ-FLAG-SPIN1 cells, FLAG-SPIN1 expression was induced 48 hours prior to harvest with 1μg/ml doxycycline. Cells were treated with 20μM oxidized Adenosine periodate (AdOx, Sigma) 24 hours prior to harvest. Lysis, immunoprecipitation and peptide elution are detailed in **2.4.3.** The methylation reaction was then set up with 2μg GST-PRMT1 and 1μCi ^{3}[H]-SAM (PerkinElmer) and volume made up to 50μl with PBS. Reactions, staining, SDS-PAGE resolution and enhancer treatment proceeded similarly to **2.4.9,** with the addition that 5μl of each reaction was analysed by western blotting to check the relative levels of SPIN1 between different reactions.

2.4.11: *In vitro* methylation of SPIN1 with L-methyl-^{3}[H]-methionine

pTIPZ-FLAG-SPIN1 MCF7 cells were seeded onto 3x10cm dishes, at a density of 3x10^6 cells. SPIN1 overexpression was induced for 24 hours by the addition of 1μg/ml doxycycline for MCF7 cells. The next day, cells were washed twice with Dulbecco's modified Eagle's medium without L-methionine, L-cystine and L-

glutamine (Sigma), supplemented with 10% v/v FBS, 1% v/v L-glutamine and 1% v/v Penicillin Streptomycin and then cultured in the same media supplemented with cycloheximide (100μg/ml) and chloramphenicol 40μg/ml) for 1 hour at 37^{o}C. Then, L-[methyl-^{3}H]-methionine was added (10μCi/ml media, specific activity 70-85Ci/mmol) and returned to 37^{0}C incubation for 3 hours. After incubation, cells were lysed in a chemical fume hood with 500μl RIPA without EDTA. Cells were scraped on ice and transferred to a pre-chilled eppendorf tube. 45 units of benzonase (Millipore, 2894174) were added to dissociate chromatin-bound proteins and tubes left on ice for 1 hour. Lysate was then centrifuged at 17,000*g* in a 4^{0}C benchtop centrifuge and supernatant transferred to new pre-chilled tubes. 50μl of each sample was taken to a new tube as an input and boiled at 95^{0}C for 5 minutes with the addition of 10μl Laemmli buffer. A further 250μl RIPA buffer without EDTA with protease inhibitors was added to the remaining lysate and 20μl 1:1 pre-washed FLAG M2 affinity gel (Sigma) added to each tube. Immunoprecipitation of FLAG-SPIN1 proceeded overnight by 4^{0}C rotation, followed by washing of beads and peptide elution **(see section 2.4.3)**, SDS-PAGE resolution, gel enhancing treatment and autoradiography are detailed in section **2.4.9.**

2.4.12: Cell Fractionation

Per treatment, 1x15cm dish was seeded with 2.5x10^{6} MCF7 pTIPZ-FLAG-SPIN1 or FLAG-SPIN1-F141A cells. Cells were kept under puromycin selection throughout the experiment (1μg/ml) and ectopic FLAG-SPIN1 expression induced at the time of plating with 1μg/ml doxycycline. Cells were harvested at

48 hours and pellets washed with PBS. Pellets were then re-suspended in 200μl 'Buffer A' (10mM HEPES pH7.9, 10mM KCl, 1.5mM $MgCl_2$, 340mM sucrose, 10% v/v glycerol, 1mM DTT – all buffers except for UTB included the same protease inhibitor cocktail used for NP40 lysis buffer, see **2.4.1**). Triton X-100 was then added to a final concentration of 0.1% v/v and samples incubated on ice for 8 minutes. Samples were centrifuged at 1,300*g* for 5 minutes in a 4^0C benchtop centrifuge. Supernatant was removed and kept as the cytoplasmic fraction and pellets were washed in the same volume of 'Buffer A', before incubation in 50μl 'Buffer B' (3mM EDTA, 200μM EGTA, 1mM DTT and protease inhibitors) on ice for 10 minutes. Samples were then centrifuged for 5 minutes at 1,700*g* at 4^0C. Supernatant was removed and kept as the soluble nuclear fraction. Pellets were washed in the same volume of 'Buffer B' and the incubated for 10 minutes at 4^0C in 50μl 'Buffer C' (3mM EDTA, 200μM EGTA, 200mM NaCl, 1mM DTT and protease inhibitors) with 0.1% v/v Triton X-100. Samples were pelleted by centrifugation (5 minutes, 1,700*g* at 4^0C) and washed with 50μl 'Buffer B' before centrifugation at 10,000*g* for 1 minute at 4^0C. Pellets were then re-suspended in 100μl UTB buffer (8M Urea, 50mM Tris,150mM b-mercaptoethanol) and sonicated twice for 10 seconds at 30% amplitude. All fractions were then clarified by centrifugation at 17,000*g* for 5 minutes at 4^0C. Prior to western blotting analysis, all samples were normalized for protein content via Bradford assay, using 'Buffer A', 'Buffer B' and UTB buffer as blank measurements for cytoplasmic, nuclear soluble and chromatin bound fractions respectively.

2.5: Bio-orthogonal Profiling Techniques

2.5.1: Cell-free alkynylation of recombinant histone H4 and detection via TAMRA conjugation

Alkynylation reactions consisted of 5μg (18mM) recombinant histone H4 (New England Biolabs, M2504S), 2μM recombinant His-PRMT1 Y29F/M38G (supplied by the Luo Laboratory, Memorial Sloan Kettering Cancer Center, NYC), BSA (1% w/v) and 100μM 4-Propargyloxy-but-2-enyl-*S*-adenosyl-*L*-methionine (Pob-SAM; synthesised by Dr. Krystian Ubych, Neely Laboratory, School of Chemistry, University of Birmingham). The reaction was buffered by 50mM HEPES (pH8.0) containing 0.005% v/v Tween-20, 0.0005% w/v BSA and 1mM TCEP and reactions made up to a final volume of 25μl. Reactions were left overnight at ambient temperature. The next day, 1μl 1% w/v BSA was added to aid precipitation: protein content was precipitated with 600μl methanol, 200μl chloroform and 400μl H_2O, with vortexing between each addition. Tubes were then centrifuged at 17,000*g* at 4^0C in a bench top centrifuge for 15 minutes. After centrifugation, the top layer was removed and 1ml ice-cold methanol added. Tubes were quickly vortexed and centrifuged again. This process was repeated one more time, resulting in aggregation of protein at the bottom of the tube. All supernatant was removed by pipette and left to air dry for approximately 30 minutes at room temperature. Pellets were then re-suspended in 100μl 50mM triethanolamine (pH7.4, Sigma) containing 4% w/v SDS and EDTA-free protease inhibitor cocktail (Roche). 25μl of this re-suspended protein was transferred to a new tube and used for the subsequent TAMRA click reaction, supplemented with 60μl 50mM triethanolamine (pH7.4, without SDS). To this, a 12.5μl reaction

cocktail consisting of 10μl CuSO4:BTTP (20mM:40mM, pre-complexed at room temperature for at least 30 minutes) and 2.5μl Sodium Ascorbate (from 100mM stock, dissolved in H_2O less than 30 minutes prior to use) was added. Finally, 2.5μl Tetramethylrhodamine 5-Carboxamido-(6-Azidohexanyl) 5-isomer (TAMRA-azide, Thermo T10182) was added to the reaction (from 10mM stock, dissolved in DMSO). Final concentrations of reagents were as follows: CuSO4 1mM, BTTP 2mM, Sodium Ascorbate 2.5mM, TAMRA-azide 250μM. Reactions were placed on a bench top thermoshaker at ambient temperature and shaken at 600rpm for 2 hours, and samples were covered with foil to protect them from light. After click incubation, protein was extracted and excess click reagents removed via methanol-chloroform precipitation (as described previously), with the exception that no additional BSA was added. After the final methanol wash, pellets were air dried under foil for 30 minutes and re-suspended in 40μl Laemmli buffer (without bromophenol blue, as this can interfere with TAMRA detection). Samples were then resolved via SDS-PAGE using a pre-cast NuPAGE™ 4-12% Bis-tris gel using NuPAGE™ MOPS SDS running buffer (Thermo) at 125V for at least 90 minutes. Gel tanks were covered during running to protect TAMRA from light-mediated degradation. TAMRA signal was then visualised using the 'Rhodamine' channel on a Bio-Rad ChemiDoc mp Imaging System. Once TAMRA had been detected, gels were stained with InstantBlue Coomassie Protein Stain (Abcam) in order to control for total protein content and imaged using the same system.

2.5.2: Semi cell-free alkynylation of whole-cell lysate and detection via TAMRA conjugation

Eight 10cm dishes of HEK-293T cells were seeded per treatment at a density of 3x 10^6 cells per dish. 24 hours later, cells were PEI transfected (see section **2.2.1**) with pHIV-Zsgreen (either an empty vector control or containing FLAG-PRMT1 or PRMT1-Y29F/M38G). AdOx (10μM) was added to culture medium immediately after the transfection cocktail to prevent *in vitro* methylation of endogenous substrates, thus providing more opportunity for their alkynylation during Pob-SAM incubation, which occurs during the cell-free component of the procedure. Cells were harvested after 48 hours and lysed in 500μl 50mM HEPES buffer pH8.0 with 0.005% v/v Tween20, 1mM TCEP (made fresh and dissolved in H_2O) and EDTA-free protease inhibitor cocktail (Roche). Cells were left on ice for 20 minutes before sonicating each sample with a small probe sonicator, using 40% amplitude for 5 seconds. Samples were subjected to 5 rounds of sonication in total, resting on ice for 30 seconds between each pulse. After sonication, samples were centrifuged for 20 minutes at 4^0C and 17,000*g* using a benchtop microcentrifuge. Supernatant was then transferred to new pre-chilled eppendorf and protein determination carried out using a pre-diluted Protein Assay standards BSA set (Thermo, 23208) according to the manufacturer's instruction. Samples were transferred to 15ml falcon tubes and 50mM HEPES pH8.0 lysis buffer added to reduce concentration of protein to 2mg/ml. Pob-SAM (final concentration 200μM) was added and samples vortexed briefly before incubating at room temperature for 12 hours. The following morning, proteins were precipitated via methanol-chloroform precipitation (see section **2.5.1**) with the omission of BSA and samples air dried for 30 minutes on ice before re-suspension in 100μl 50mM

Triethanolamine pH7.4 with 4% w/v SDS and EDTA-free protease inhibitor cocktail. Sonication with a small probe sonicator (6 x 5 second pulses, 35% amplitude) was necessary to re-suspend the large quantity of protein in a relatively small volume. Copper-catalysed click addition of TAMRA-azide to alkynylated protein, subsequent methanol-chloroform clean-up, SDS-PAGE resolution and TAMRA imaging proceeded in the same manner as in section **2.5.1**.

2.5.3: *In vitro* alkynylation of cell proteome and detection via TAMRA conjugation

Two 10 cm dishes for each treatment were seeded with $3x10^6$ HEK-293T cells. 24 hours later, dishes were PEI transfected with 10μg pcDNA3.1 construct (either empty vector control or containing various HA-MAT/FLAG-PRMT1 combinations, **see section 3.2.2**) and treated with AdOx (20μM) to hypomethylate the proteome. 24 hours later cells were washed twice with 2ml methionine-free DMEM and medium replaced with 4ml methionine-free DMEM without replenishing AdOx. Pob-methionine (synthesized by Dr. Krystian Ubych, Neely Laboratory, School of Chemistry, University of Birmingham) was then added to a final concentration of 2mM and cells incubated at 37^0C for 4 hours. At the end of incubation, cells were pelleted and a small amount of material (1/20th total pellet volume) kept for western blotting analysis. The remaining pellets were stored at -80^0C prior to resuming the experiment. Pellets were thawed on ice and lysed in 500μl 50mM HEPES buffer pH8.0 with 0.005% v/v Tween20, and EDTA-free protease inhibitor cocktail (Roche). Sonication was carried out in four rounds

using a small probe (3x 30%, 1x 35% amplitude) for 10 seconds each, returning each sample to ice for 30 second intervals. Samples were then centrifuged at 17,000*g*, 4^0C for 20 minutes in a benchtop centrifuge. Supernatant was transferred to fresh eppendorf tubes and further clarified by centrifugation under the same conditions for 5 minutes. Supernatant was then transferred to 15ml falcon tubes and protein isolated via methanol-chloroform purification (see section **2.5.1**). After last methanol wash, pellets were air-dried for 30 minutes and remaining methanol removed via syringe or pipette. Pellets were then re-suspended in 500μl 50mM triethanolamine (pH 7.4) containing 4% w/v SDS and EDTA-free protease inhibitor cocktail (Roche) via sonication: tubes were mounted in ice and sonicated at 30% amplitude with a small probe (10 seconds on, 10 seconds off) for 5 minutes in total. 25μl of this resuspension was used for addition of TAMRA-azide to alkynylated protein. The constituents of this click reaction were set up as in section **2.5.1.** Reactions were protected from light and shaken at ambient temperature for 90 minutes at 800rpm and protein subsequently methanol-chloroform purified and dried as previously described (section **2.5.1**). Samples were then re-suspended by pipetting in 40μl Laemmli buffer without bromophenol blue. Light-protected SDS-PAGE and imaging were carried out as previously described in **2.5.1.**

Table 1: Cell lines used in this thesis.

Cell Line	Source
MCF10A	ATCC
MCF10A pTIPZ-FLAG-SPIN1	Generated in-house
MCF7	Dr. Gillain Farnie, SGC, Oxford
MCF7 pTIPZ-FLAG-SPIN1	Generated in-house
MCF7 pTIPZ-FLAG-SPIN1-R14K	Generated in-house
MCF7 pTIPZ-FLAG-SPIN1-R117K	Generated in-house
MCF7 pTIPZ-FLAG-SPIN1-F141A	Generated in-house
MCF7 pTIPZ-FLAG-SPIN1-R152K	Generated in-house
HEK-293T	Davies Laboratory

Table 2: PCR primers used in this thesis. All primers listed below are for site-directed mutagenesis. Lower case base pairs indicate targeted mutations.

PCR primers	Sequences
SPIN1-R14K	**F** 5'-CCAGCGGTCCaaaGCTGATGCAG-3' **R** 5'-CCAGGTGTCTTTCCGAATGGGG-3'
SPIN1-Y98H	**F** 5'-TGACTGTGTTcatGGACTAGAAC-3' **R** 5'-AATCCATCGTATTTTATAAGATAC-3'
SPIN1-Y98R	**F** 5'-TGACTGTGTTaggGGACTAGAACTTAATAAAG-3' **R** 5'-AATCCATCGTATTTTATAAGATAC-3'
SPIN1-R117K	**F** 5'-CCTCCCTGATaaaGTTGCGACAT-3' **R** 5'-ACTTCAAGCGCAGAAACTC-3'
SPIN1-F141A	**F** 5'-GGAACATATGgcaGAGACAGAGGATG-3' **R** 5'-ACTGCTTTGCCAATCATTG-3'
SPIN1-R152K	**F** 5'-AGATGAGTGGaaaGGAATGGTCTTAG-3' **R** 5'-TTAGAACCATCCTCTGTC-3'
SPIN1-D184H	**F** 5'-ACTCTTAGATcatTACAAAGAAGG-3' **R** 5'-TGGTACATGTACAAGACAG-3'
SPIN1-F251H	**F** 5'-TGATGATGATcatCATATTTATGTCTACGATTTG-3' **R** 5'-AACTTGATGAAATAGACGG-3'
SPIN1-F251R	**F** 5'-TGATGATGATaggCATATTTATGTCTACGATTTG-3' **R** 5'-AACTTGATGAAATAGACGG-3'
PRMT1-Y29F	**F** 5'-CTTTGACTCCttcGCACACTTTG-3' **R** 5'-TAGTAATCTTTGGATGTCATG-3'
PRMT1-M38G	**F** 5'-CCACGAGGAGgggCTGAAGGACG-3' **R** 5'-ATGCCAAAGTGTGCGTAG-3'
MAT2A-I117A	**F** 5'-GTCACCAGATgctGCTCAAGGTGTTC-3' **R** 5'-TGTTGCTCCAAGGCTACC-3'

Table 3: qPCR primers used in this thesis.

qPCR primers	Sequences	Source	Efficiency
ACTIN	**F** 5'- CTCTTCCAGCCTTCCTTCCT-3' **R** 5'- GGATGTCCACGTCACACTTC-3'	**Ensembl/ BLAST**	2.01
PRMT1v1	**F** 5'-GACTCCTACGCACACTTTGG-3' **R** 5'-TGAAGAGGTGCCGGTTATGA-3'	**Ensembl/ BLAST**	1.2
AXIN2	**F** 5'-AGTGTGAGGTCCACGGAAAC-3' **R** 5'-CTTCACACTGCGATGCATTT-3'	[47]	1.93
CYCLIN D1	**F** 5'-CCGTCCATGCGGAAGATC-3' **R** 5'-ATGGCCAGCGGGAAGAC-3'	[47]	1.92
ID-2	**F** 5'-TCAGCCTGCATCACCAGAGA-3' **R** 5'-CTGCAAGGACAGGATGCTGAT-3'	[47]	1.85
TIAM1	**F** 5'-AAGACGTACTCAGGCCATGTCC-3' **R** 5'-GACCCAAATGTCGCAGTCAG-3'	[47]	2.2
OCT4	**F** 5'-TACAGCATGTCCTACTCGCA-3' **R** 5'-TGGAGTGGGAGGAAGAGGT-3'	**Ensembl/ BLAST**	1.62
SOX2	**F** 5'-CAACARGCTCAGGGAACAGG-3' **R** 5'-GTTAGCATGAGTTGGCACCC-3'	**Ensembl/ BLAST**	1.96
NANOG	**F** 5'-CGTGAAGCTGGAGAAGGAG-3' **R** 5'-TGCTCGAGTTCTTTCTGCAG-3'	**Ensembl/ BLAST**	1.66

Table 4: antibodies used in this thesis.

Antibody	Company	Raised in	Catalogue ID	Dilution Factor/Function
SPIN1	Abcam	Rabbit	Ab196938	1/1000 Western blot
PRMT1	CST	Rabbit	2449	1/1000 Western blot
FLAG	Sigma	Rabbit	F7425	1/1000 Western blot
HA.11	Covance	Mouse	MMS-101R	4μl (IP) 1/1000 Western blot
C-MYC (9E10)	Sigma	Mouse	M4439	1/1000 Western blot
GFP	Abcam	Rabbit	Ab6556	1/1000 Western blot

Table 5: affinity beads used in this thesis.

Name	Company	Catalogue ID	Function
Anti-FLAG M2 Affinity Gel	Sigma	A2220	FLAG IP
GFP-Trap agarose	Chromotek	gta-10	GFP IP
IgG Sepharose 6 Fast Flow	GE Healthcare	17-0969-01	IgG control IP
Protein G Sepharose 4 Fast Flow	GE Healthcare	17-0618-01	IP pre-clear//HA IP with HA.11
Ni-NTA Agarose	Invitrogen	R901-01	His-tag recombinant protein purification
Glutathione Sepharose 4B	GE Healthcare	17-0756-01	GST-tag recombinant protein purification

Chapter 3: Optimisation of Bio-Orthogonal Profiling

3.1: Cell-free validation of bio-orthogonal PRMT1

PRMT1-mediated transcriptional coactivation via deposition of H4R3me2a remains largely obscured on the epigenome-wide scale. Traditional X-ChIP using an antibody raised against H4R3me2a has produced relatively few publications[89,141], possibly due to the challenging nature of enriching for a relatively discrete epitope[73]. Bio-orthogonal profiling followed by native ChIP could remedy this issue through surrogate modification (alkynylation) of H4R3 by a matched and specifically engineered PRMT1-cofactor pair[142]. Ultimately, this would allow comparison of the PRMT1-mediated epigenetic signature across the epigenome between normal and breast cancer cells – providing crucial insight into PRMT1-mediated malignancy in breast cancer.

3.1.1: Cell-free validation histone H4 alkynylation: background

The first step in demonstrating the viability of our chosen PRMT1 mutant for bio-orthogonal profiling involved cell-free alkynylation of recombinant histone H4 with recombinant PRMT1. This particular PRMT1 contains a double Y29F/M38G mutation in its SAM binding pocket to accommodate the larger Pob-SAM alkynyl donor **(Figure 3.1).** Detection of this modification is made possible by subsequent addition of a TAMRA-azide probe by CuAAC (Cu^{I}-catalyzed azide-alkyne 1,3-dipolar cycloaddition) allowing alkynylated histones to be visualised by in-gel fluorescence.

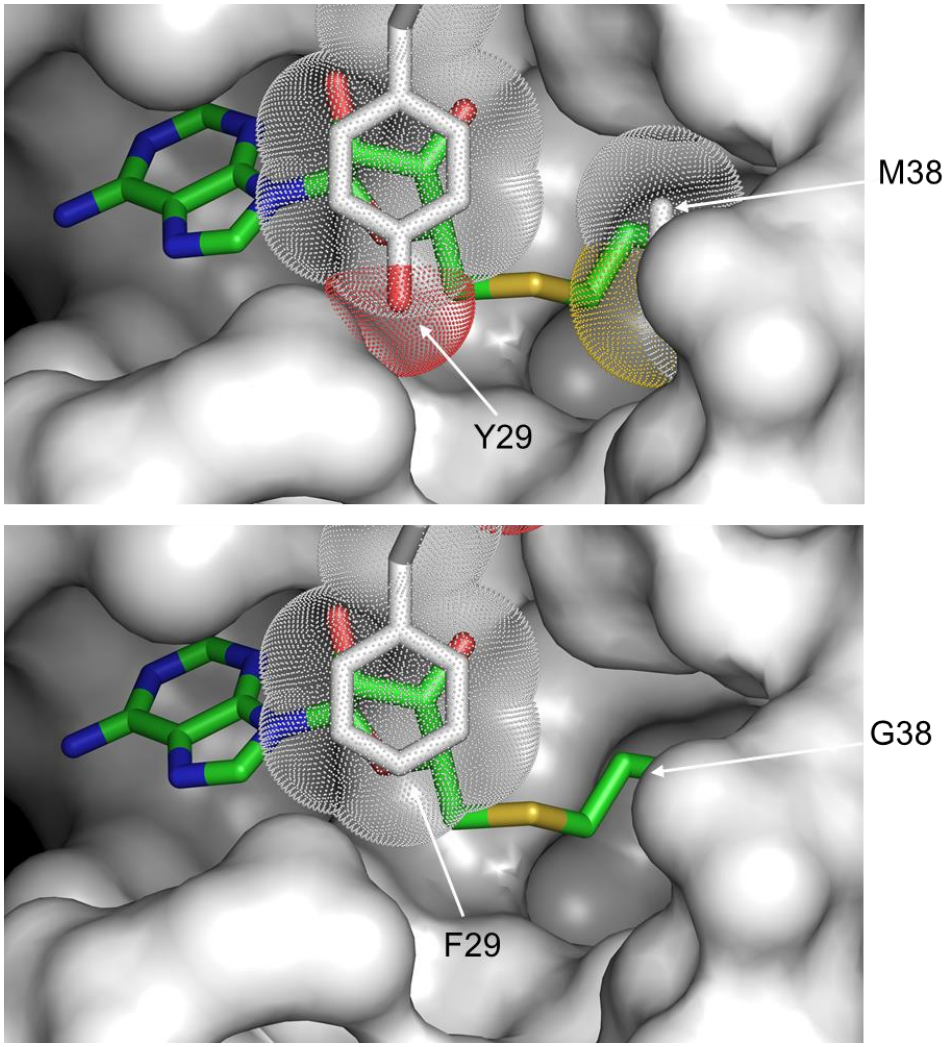


Figure 3.1: PRMT1 mutations to sterically accommodate alkynylating cofactor Pob-SAM

Top panel: Inside view of the wild-type PRMT1 SAM binding pocket, occupied by S-adenosylhomocysteine (SAH, the by-product of SAM after methyl donation). Arrows mark key residues selected for mutation to expand the binding pocket. The yellow segment of SAH denotes the sulfonium moiety purported to be contacted by these key residues during binding and catalysis. Dotted surfaces denote van der Waals forces. Red and yellow surfaces (Y29 and M38 respectively) imply forces that could sterically hinder Pob-SAM. Lower panel: steric proximity of bio-orthogonal mutant residues with SAH. Images generated by Professor Stephen Smerdon, University of Birmingham. PDB ID: 6NT2.

3.1.2 Cell-free alkynylation of histone H4: initial attempts

Although generation of recombinant His-PRMT1 and His-PRMT1-Y29F/M38G was a success **(Figure 3.2 A)**, this was accomplished after initial attempts to generate GST-PRMT1-Y29F/M38G failed due to apparent instability of the protein (data not shown). His-PRMT1 and His-PRMT1-Y29F/M38G were placed into cell-free reactions with Pob-SAM as an alkynyl donor and recombinant histone H4 as a substrate. After incubation, TAMRA-azide was conjugated to alkynylated H4 via CuAAC. This experiment failed to re-constitute the reported bio-orthogonal nature of His-PRMT1-Y29F/M38G[106] **(Figure 3.2 B).** Not only did His-PRMT1-Y29F/M38G fail to out-perform the wild-type control, but labelling of H4 was observable in the no-enzyme control lane **(Figure 3.2 B, lane 5)** at a similar intensity to lanes containing either wild-type His-PRMT1 or His-PRMT1-Y29F/M38G (**Figure 3.2 B, lanes 1–4**), suggesting that the signal observed at the size of histone H4 was largely enzyme-independent.

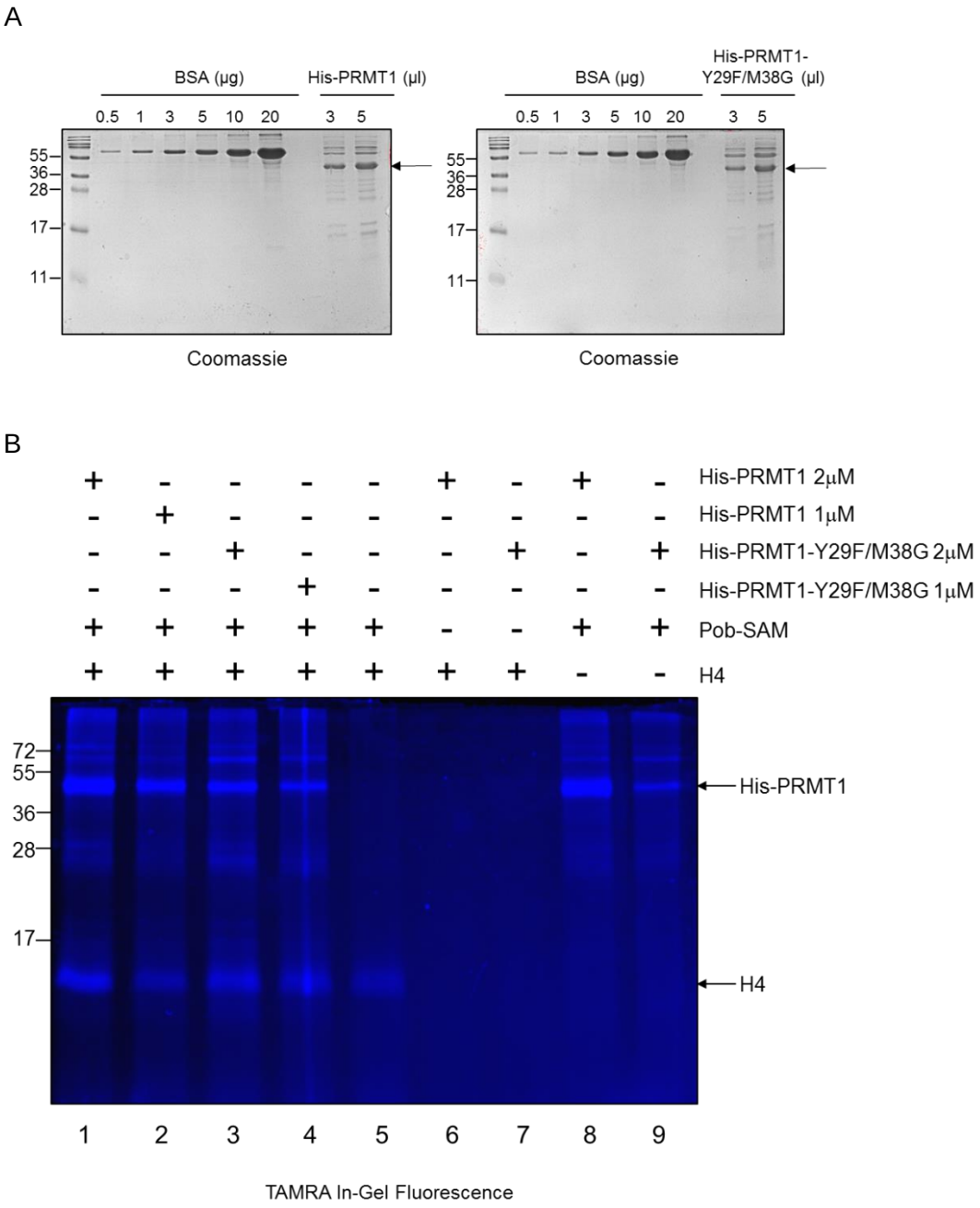


Figure 3.2: PRMT1 Y29F/M38G did not specifically utilise Pob-SAM in a cell-free reaction

(A) Coomassie staining of BSA gradient showing successful generation of His-PRMT1 and His-PRMT1-Y29F/M38G. Arrows denote relevant protein species. **(B)** TAMRA in-gel fluorescence readout of attempted alkynylation of recombinant histone H4 by recombinant His-PRMT1-Y29F/M38G compared to a wild type His-PRMT1 control. n=3.

3.1.3: Cell-free alkynylation of histone H4: troubleshooting

Efforts to optimise this experiment in-house proved unsuccessful in spite of multiple batch generations of recombinant enzyme and alterations to incubation time/temperature of the alkynylation reaction itself (data not shown). As such, additional training was undertaken at the Luo Laboratory (Memorial Sloan Kettering Cancer Center, New York, USA). First, we compared our UoB-generated His-PRMT1 with recombinant PRMT1 protein generated by the Luo lab via cell-free methylation using recombinant histone H4 or a peptide representing the first 21 amino acids of histone H4 as substrate. Surprisingly, it was found that for both substrate types, the Luo laboratory-generated PRMT1 was substantially more efficient at methylation than our recombinant enzyme, despite both enzymes being generated through His-tag purification **(Figure 3.3 A)**. This was deemed an issue related to the purity of our enzyme, which had only been batch purified, and therefore potentially contained protein contaminants that may have interfered with catalysis. This idea was supported by absence of any Coomassie stain signal after initial attempts to optimise the assay (such as in **Figure 3.2**) suggesting that protein contents of these reactions had been degraded, possibly by proteases derived from the bacterial origin of the His-tagged proteins (data not shown).

Consequently, it was rationalised that sub-optimal catalysis of recombinant His-PRMT1-Y29F/M38G was responsible for our inability to detect alkynylation of histone H4 using our recombinant His-PRMT1-Y29F/M38G **(Figure 3.2 B).** Hence, the Luo in-house generated His-PRMT1-Y29F/M38G, rather than our own enzyme, was used in the next attempt at cell-free alkynylation. The reaction

protocol was also altered to include BSA (bovine serum albumin) to aid in protein precipitation during purification stages of the protocol. This experiment produced a more robust increase in fluorescence of a protein corresponding to the molecular weight of histone H4 when His-PRMT1-Y29F/M38G was present within the reaction, suggesting active modification of histone H4. Additional bands suggested alkynylation of other proteins by His-PRMT1-Y29F/M38G, including a band corresponding to PRMT1 itself. This implies that His-PRMT1-Y29F/M38G can auto-alkynylate in a cell-free context, similar to its wild-type counterpart which can automethylate under such circumstances[143]. This phenomenon appeared to be specific and not a consequence of background alkynylation and/or TAMRA conjugation, as BSA was not alkynylated **(Figure 3.3 B)**.

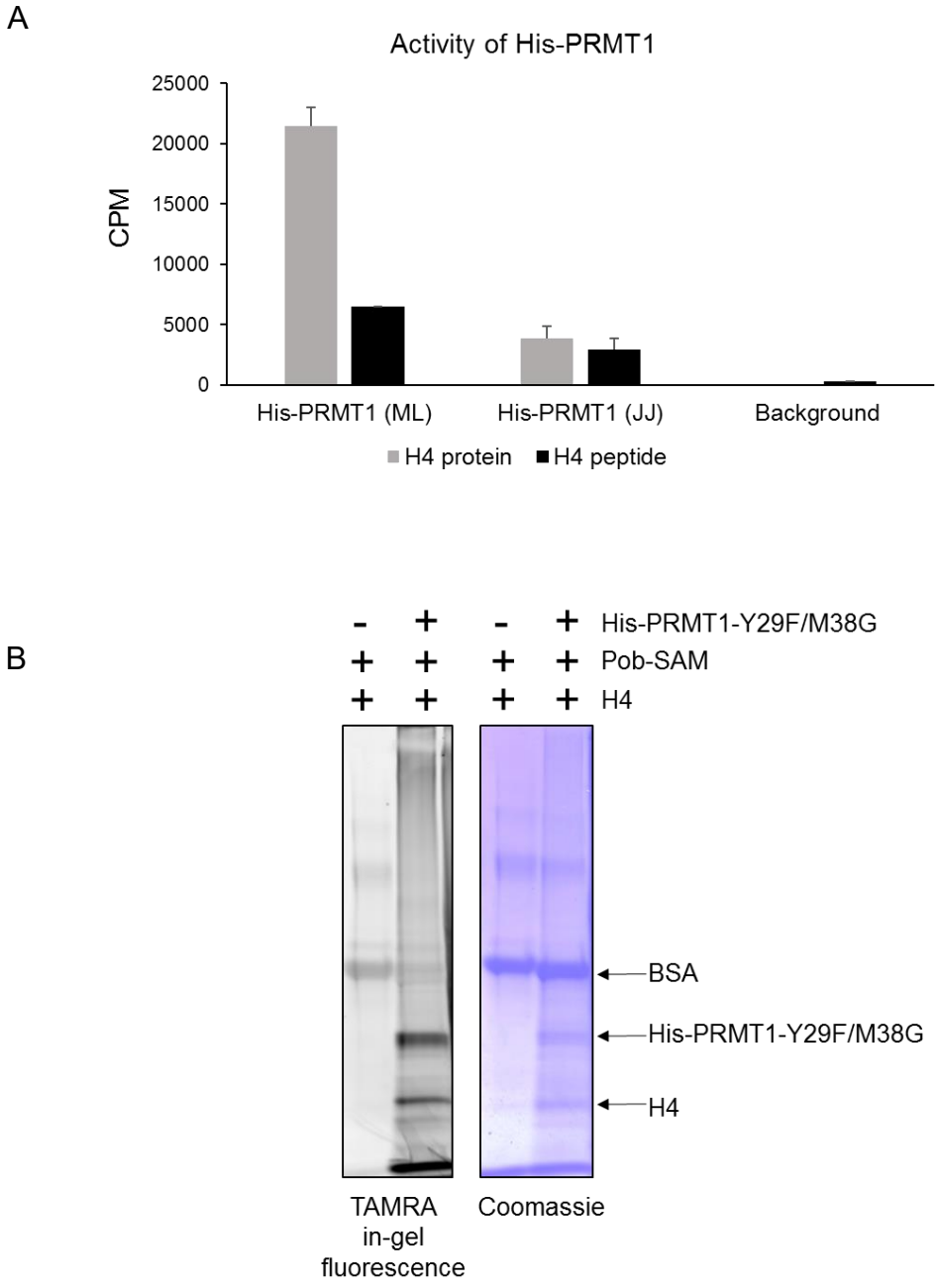


Figure 3.3: Successful cell-free alkynylation of histone H4

(A) Scintillation count readout of cell-free methylation assays comparing wild-type His-PRMT1 as generated in the Luo and Davies Laboratories (His-PRMT1 ML and His-PRMT1 JJ, respectively). **(B)** Left panel: cell-free alkynylation of histone H4 and in-gel visualisation with clickable TAMRA-azide probe using His-PRMT1-Y29F/M38G generated by the Luo Laboratory. Right hand panel shows Coomassie stain of gel shown in left panel to show overall protein content. n=1.

3.2 Semi cell-free validation of Bio-orthogonal PRMT1

3.2.1 PRMT1-Y29F/M38G utilisation of native SAM

As shown by Wang et al.[106], the bio-orthogonal PRMT1 mutant has been proposed to harbour negligible activity towards native SAM. Since the ultimate aim of this methodology is to express ectopic PRMT1-Y29F/M38G in mammalian cells, an inability of PRMT1-Y29F/M38G to bind and utilise native SAM is important as it will prevent competition with Pob-SAM for access to the active site, increasing alkynylation efficiency and thus, signal to noise ratio.

To validate this, we conducted semi cell-free methylation assays. HEK-293T cells were transiently transfected with FLAG-tagged PRMT1, a catalytically inactive control (FLAG-PRMT1-XGX[144]) and FLAG-PRMT1-Y29F/M38G. Ectopic PRMT1 was then isolated by immunoprecipitation and incubated in an cell-free methylation assay using recombinant histone H4 as a substrate and ^{3}[H]-SAM as a cofactor **(Figure 3.4 A).**

This experiment reproduced the findings of Wang et al., with wild-type FLAG-PRMT1 robustly methylating histone H4 in contrast to the FLAG-PRMT1-XGX negative control. Importantly, FLAG-PRMT1-Y29F/M38G proved incapable of utilising native SAM to methylate histone H4, suggesting that native SAM would not provide competition for Pob-SAM for PRMT1-Y29F/M38G occupancy *in vitro* **(Figure 3.4 B)**.

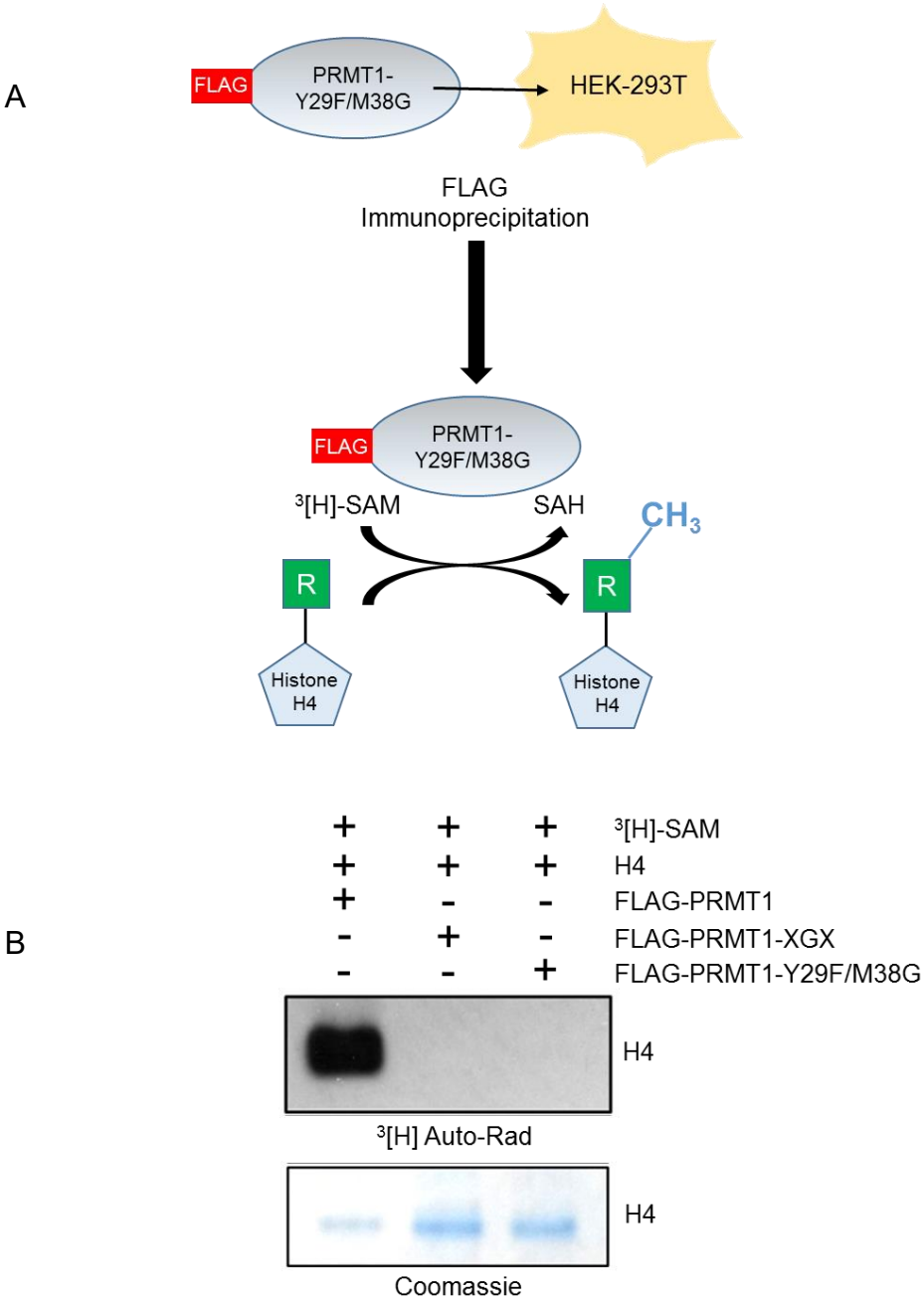


Figure 3.4: PRMT1-Y29F/M38G is inactive towards native SAM

(A) Schematic showing process of semi cell-free methylation assay using recombinant histone H4 as a substrate and ^{3}[H]-SAM as a cofactor. PRMT1-Y29F/M38G is transiently overexpressed in HEK-293T cells and immunoprecipitated 24 hours after transfection. Next, its ability to utilise native SAM to methylate histone H4 is tested in a cell-free context. (**B)** Upper panel shows auto-radiographic readout of H4 methylation, lower panel shows total H4 present in each reaction**.** n=3**.**

3.2.2: Alkynylation of whole-cell lysate

The finding that transfected FLAG-PRMT1 could be catalytically tested in a cell-free reaction suggested that FLAG-PRMT1-Y29F/M38G could be used similarly to test its efficacy with Pob-SAM. Since a direct comparison between His-PRMT1 and His-PRMT1-Y29F/M38G had not been managed in a cell-free context (the His-tagged PRMT1-Y29F/M38G protein was not prioritised for re-generation due to time constraints) semi cell-free alkynylation of whole-cell lysate using transfected FLAG-PRMT1 was considered a viable solution **(Figure 3.5 A)**. Using whole-cell lysate containing overexpressed FLAG-PRMT1-Y29F/M38G as a source of enzyme in reactions containing Pob-SAM and recombinant H4, rather than immunoprecipitating and enriching for FLAG-PRMT1-Y29F/M38G, holds significant advantages. The enzyme does not need to be isolated, and the much greater protein content acquired from whole cell lysate makes the multiple precipitation steps throughout the enzymatic/CuAAC segments of the protocol easier and less susceptible to loss of material.

HEK-293T cells were transfected with either FLAG-PRMT1-Y29F/M38G or a wild-type control and pre-treated with universal methylation inhibitor AdOx to prevent endogenous methylation of the proteome, thus leaving more unmodified substrate to be alkynylated after cell lysis. Samples were then lysed and incubated with Pob-SAM, and alkynylation of the whole-cell proteome observed by TAMRA in-gel fluorescence. This experiment was originally carried out in the Luo Laboratory (with material pre-generated in the Davies Laboratory), demonstrating a clearer pattern of modification when cells were transfected with FLAG-PRMT1-Y29F/M38G than those transfected with FLAG-PRMT1 or an

empty vector control **(Figure 3.5 B)**. Repetition of this protocol at the Davies Laboratory showed concordance, with a more striking banding pattern in the presence of FLAG-PRMT1-Y29F/M38G, thus suggesting an enhanced ability for PRMT1-Y29F/M38G to utilise Pob-SAM to alkynylate protein substrate **(Figure 3.6)**.

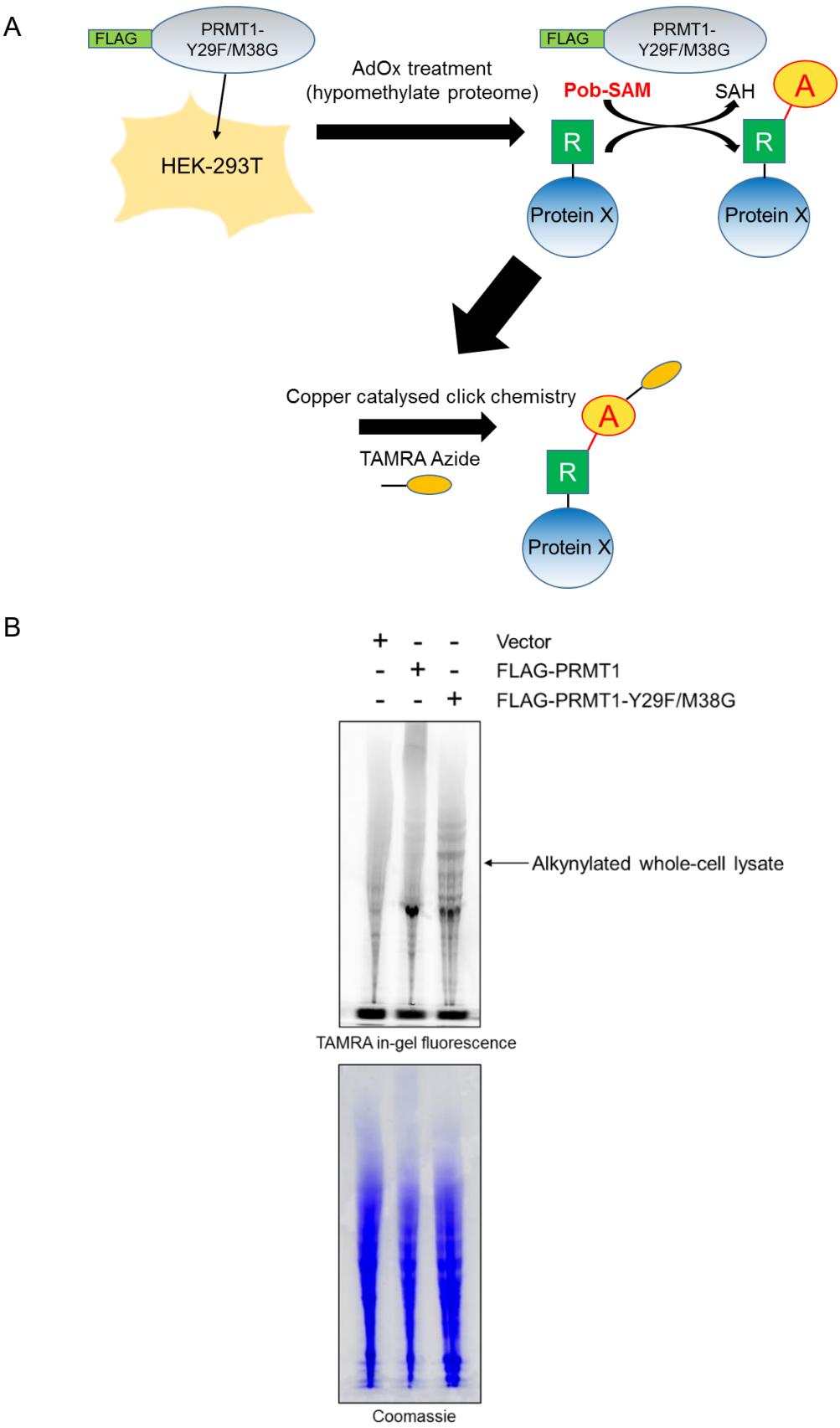


Figure 3.5: PRMT1-Y29F/M38G alkynylates whole-cell lysate by utilising Pob-SAM

(A) PRMT1-Y29F/M38G-mediated alkynylation of whole cell lysate. HEK-293T cells were transiently transfected with constructs for overexpression of either FLAG-PRMT1 or FLAG-PRMT1-Y29F/M38G and cells treated with 100μM AdOx. Cells were lysed 24 hours later and lysate incubated with 200μM Pob-SAM for 12 hours. Alkynylation of the proteome was viewed by in-gel fluorescence after addition of a TAMRA-azide probe to alkynylated protein by CuAAC. **(B)** Upper panel: TAMRA in-gel fluorescence readout of the experiment detailed in **A**, as carried out at the Luo Laboratory. Lower panel: Coomassie stain of the gel shown in the upper panel, showing total protein content in the gel. n=1.

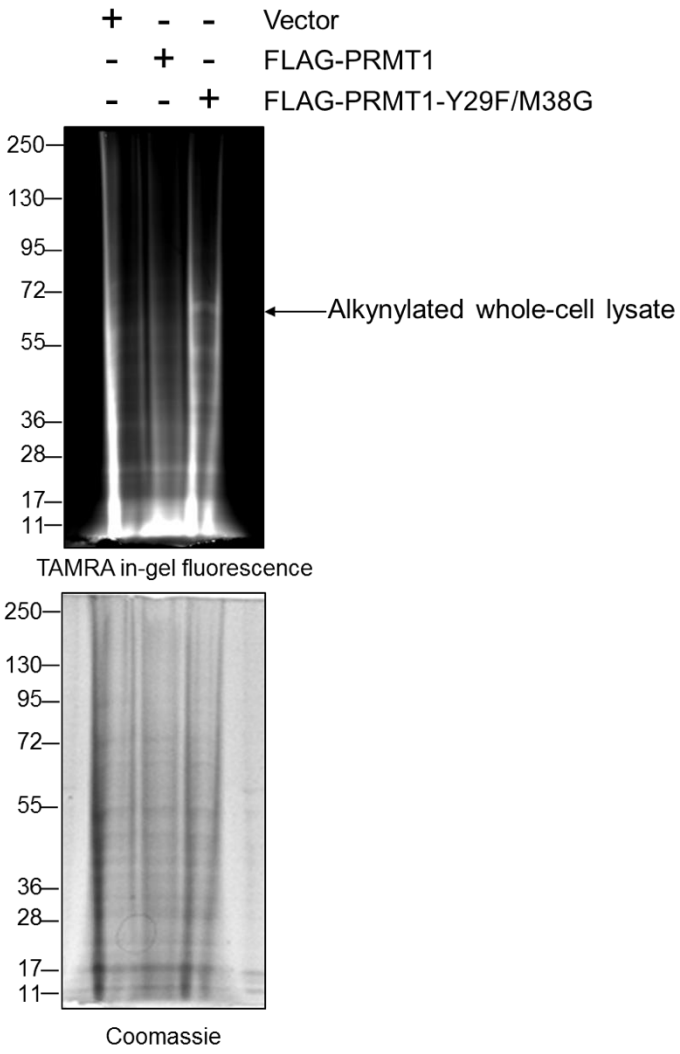


Figure 3.6: PRMT1-Y29F/M38G alkynylates whole-cell lysate by utilising Pob-SAM (repeat)

Repeat experiment of **3.5 B**, carried out in the Davies Laboratory. Upper panel: in-gel fluorescence representing the extent of alkynylation of the proteome. Lower panel: Coomassie stain showing total protein content in the gel shown in the upper panel. n=1.

3.3: *In vitro* validation of bio-orthogonal PRMT1

3.3.1: Design and generation of an *in vitro* system

Although data from semi cell-free experimentation suggested that PRMT1-Y29F/M38G had an increased capacity to alkynylate endogenous proteins with Pob-SAM **(Figures 3.5 and 3.6)**, this did not sufficiently represent the entire bio-orthogonal system *in vitro.* Due to the poor membrane permeability of Pob-SAM[142], supplementation with the precursor Pob-methionine would be necessary for labelling histones in living cells. This requires conversion of Pob-methionine to Pob-SAM via catalytic activity of MAT2A-I117A. Similar in rationale to the mutations in PRMT1-Y29F/M38G, this mutant variant of MAT2A has been reported by Wang et al. to be capable of converting Pob-methionine to Pob-SAM via expansion of its SAM binding pocket[108]. As with the semi cell-free experiment, this procedure used CuAAC to add TAMRA-azide to modified protein so that alkynylation could be determined by in-gel fluorescence. Previous attempts to validate the *in vitro* bio-orthogonal pathway using stable MCF10A cell lines overexpressing MAT2A-I117A and PRMT1-Y29F/M38G were unsuccessful with a very low TAMRA signal acquired from in-gel fluorescence (data not shown).

Although the bio-orthogonal pathway published by the Luo group contains MAT2A-I117A as a converter of Pob-methionine to Pob-SAM[108], no elaboration on the choice of this particular member MAT family was provided. Upon further research of the MAT enzyme family, MAT1A was selected as a potential replacement. The decision to test MAT1A was made for numerous reasons. First, MAT2A heterodimerises with MAT2B in order to be catalytically active (forming

the catalytic and regulatory subunits, respectively)[145]. If MAT2A-I117A had been the limiting factor in efficiency of the *in vitro* system, then perhaps this had been an issue of stoichiometry, since MAT2B was not simultaneously overexpressed, resulting in a large but mostly inactive pool of MAT2A-I117A. MAT1A could potentially circumvent this issue, since it is active as a homodimer or homotetramer[145]. Another advantage of MAT2B independence is that MAT1A is not subjected to the auto-inhibitory feedback that MAT2B places upon MAT2A **(Figure 3.7 A)**[145]. Super-imposition of MAT1A crystal structure onto MAT2A shows high structural similarity, and importantly, key residues in the SAM binding pocket of MAT2A that are almost entirely conserved within MAT1A **(Figure 3.7 B)**. Hence, it appeared likely that mutation of I117 to alanine in MAT1A would function in a similar manner to MAT2A-I117A, allowing for Pob-Methionine to Pob-SAM conversion.

To examine this possibility, a range of constructs containing FLAG-tagged PRMT1 (either wild-type or Y29F/M38G) and HA-MAT1A-I117A or HA-MAT2A-I117A were cloned. To reduce the amount of cellular engineering required for an effective bio-orthogonal system, these sequences were combined into a single construct, separated by the viral self-cleaving P2A sequence. As a result, both PRMT1 and MAT enzymes are translated as a single polypeptide that undergoes P2A peptide self-cleavage, producing the two proteins. Importantly, this means that each cell expressing the viral construct will express both mutated enzymes required to reconstitute *in vitro* the bio-orthogonal system, alleviating concerns that certain cells within the population may express these enzymes with mutual exclusivity.

These sequences were cloned into two vector backbones: pcDNA 3.1 for transient overexpression, and pTIPZ, a modified variant of the commercial pTRIPZ vector obtained from the Coleman Laboratory, University of Birmingham, for the production of stable cell lines with inducible expression. This vector was originally generated by Dharmacon for doxycycline-inducible knockdown of a gene of interest via shRNA induction. Both the choice shRNA and tRFP reporter under control of a tet-on CMV promoter were removed and replaced with the sequence(s) for the gene(s) of interest for doxycycline-induced overexpression. An inducible system was necessary for the eventual generation of normal mammary epithelial (MCF10A) and breast cancer (MCF7) cell lines, since constitutive overexpression of MAT2A-I117A and PRMT-Y29F/M38G in MCF10A cells resulted in slower growth, a higher rate of cell death and a more spindle-like morphology, suggesting that these enzymes had a cytotoxic effect (data not shown). This construct also holds the potential to include an shPRMT1 hairpin, in order to knock down endogenous PRMT1, thus reducing competition for substrate between the endogenous methylation pathway and ectopic alkynylation pathway. The range of constructs and cloning strategy are outlined in **Figure 3.8.**

A

MAT	MAT1A	MAT2A	MAT2B
Protein Product	α1	α2	V1 (same as β), V2
MAT isoenzyme	MATI (homotetramer) MATIII (homodimer)	Catalytic subunit of MATII	Regulatory subunit of MATII
Regulatory subunit	None	Both V1 and V2 lower MATII K_m for methionine, V1 also lowers K_i for SAMe	N/A
K_m for methionine	MATI: 23μM-1mM MATIII: 210μM-7mM	4-10μM	N/A
K_i for SAMe	MATI: 400μM MATIII: none	60μM	N/A

B

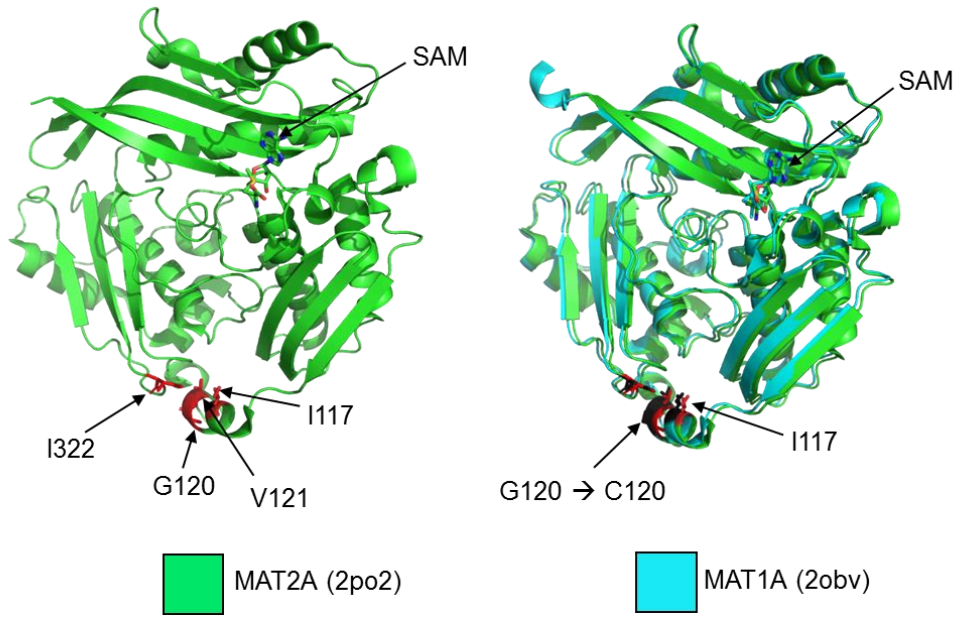


Figure 3.7: Comparison of MAT1A and MAT2A for Pob-methionine conversion *in vitro*

(A) Table outlining mechanisms of action and activity towards methionine for MAT1A and MAT2A, adapted from Maldonado et al., 2018[145]. **(B)** Left image: crystal structure of monomeric MAT2A, including SAM cofactor. Residues in red constitute the highly-conserved methylio binding pocket. Right image: Superimposition of MAT1A and MAT2A monomers. Residues in red are those highlighted in the left image. Residues depicted in black are the equivalent residues found in MAT1A. Structures 2po2 and 2obv were obtained from PDB (www.rcsb.org).

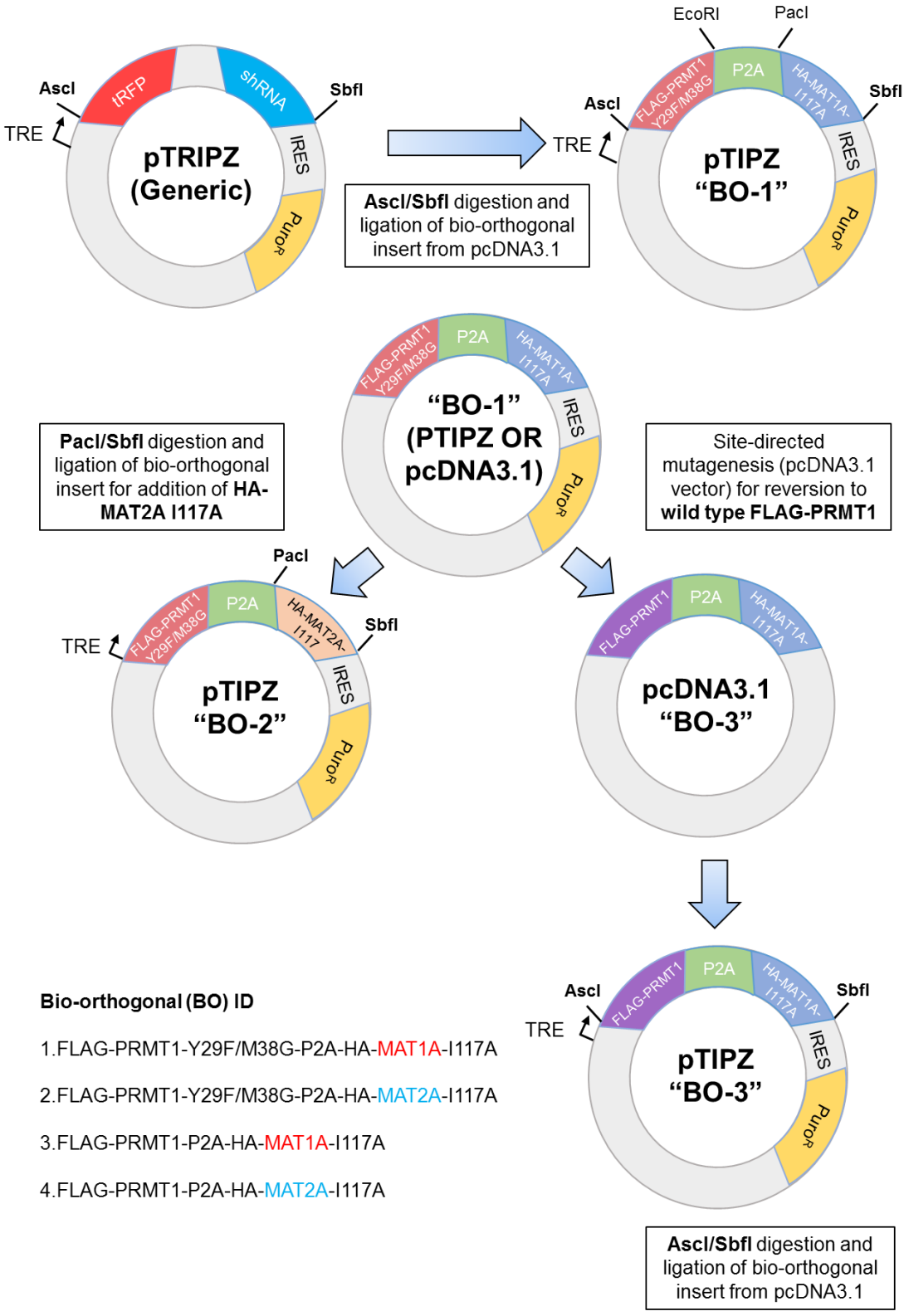


Figure 3.8: Design and generation of an *in vitro* bio-orthogonal expression system

Cloning strategy to generate range of bio-orthogonal constructs. The commercial pTRIPZ vector is modified to house a sequence containing a FLAG-tagged version of PRMT1 and an HA-tagged

MAT tagged MAT enzyme sequence, separated by the P2A self-cleaving peptide sequence. Restriction enzymes in bold denotes their use to attain the next iteration of the construct. Through sub-cloning and site-directed mutagenesis, a library of four constructs for attempting to recapitulate a full *in vitro* bio-orthogonal pathway were generated, listed below “Bio-orthogonal (BO) ID”.

3.3.2: Validation of *in vitro* constructs

Once the constructs detailed in **section 3.3.1** were successfully generated, their performance was validated by transient overexpression in HEK-293T cells. Initial transfections were carried out to test expression level and P2A-mediated cleavage efficiency of all constructs **(Figure 3.9 B)**. All enzyme combinations tested were initially in the pcDNA3.1 vector backbone, however, a single pTIPZ construct was tested to demonstrate doxycycline inducibility of enzyme expression **(Figure 3.10)**. All constructs expressed equally, and cleavage was efficient for all, with dominant bands corresponding to the molecular weight of cleaved species. For subsequent analysis, pcDNA3.1 constructs were chosen for testing Pob-methionine utilisation due to incomplete availability of the entire range of PRMT1/MAT combinations within the pTIPZ vector.

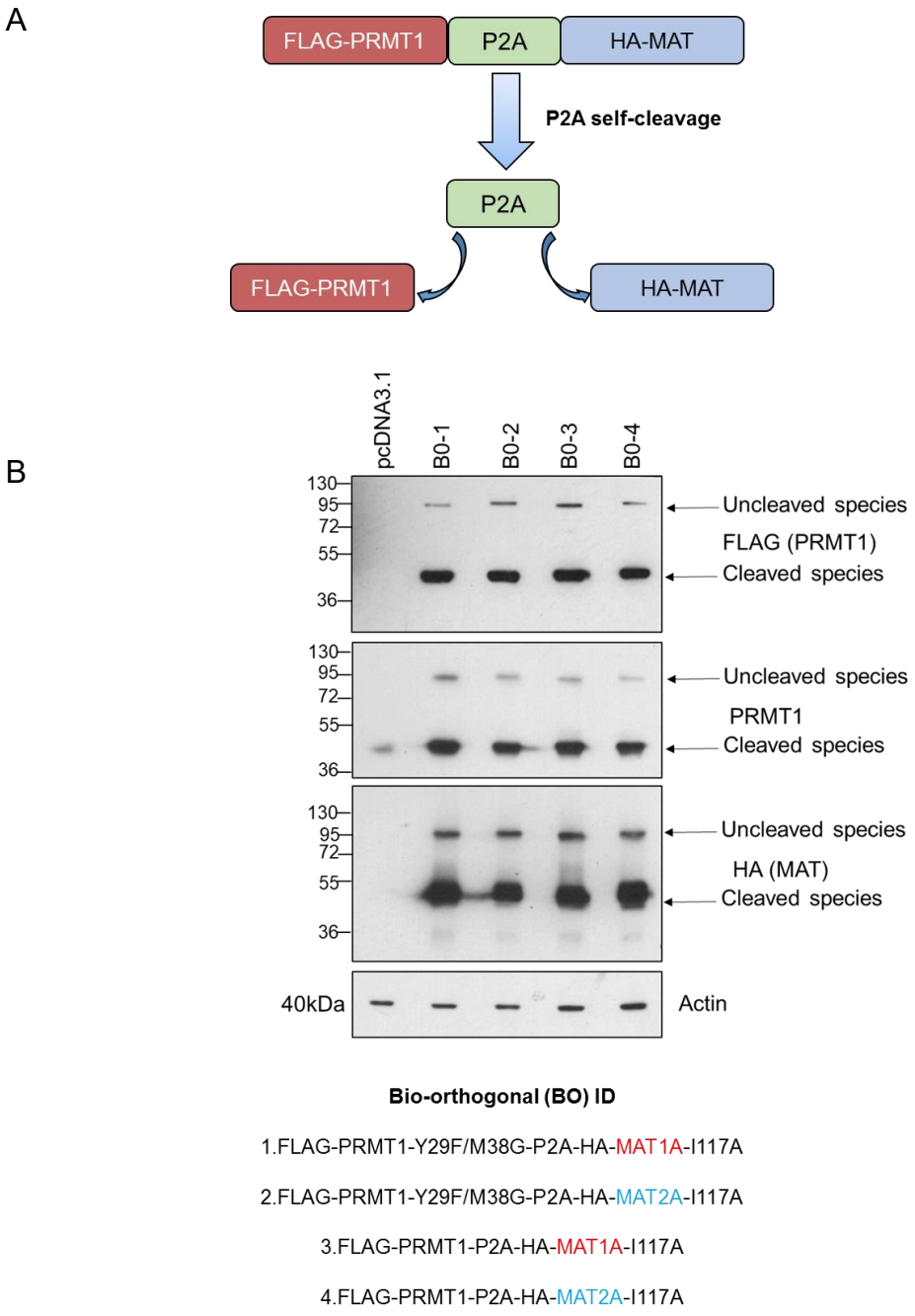


Figure 3.9: *In vitro* validation of constitutively expressed bio-orthogonal constructs

(A) Diagram showing generic bio-orthogonal insert and post-translational P2A-dependent cleavage. **(B)** Western blot of various bio-orthogonal enzyme combinations within the pcDNA3.1 backbone, transiently transfected into HEK-293T cells, harvested 24 hours after transfection. n=1.

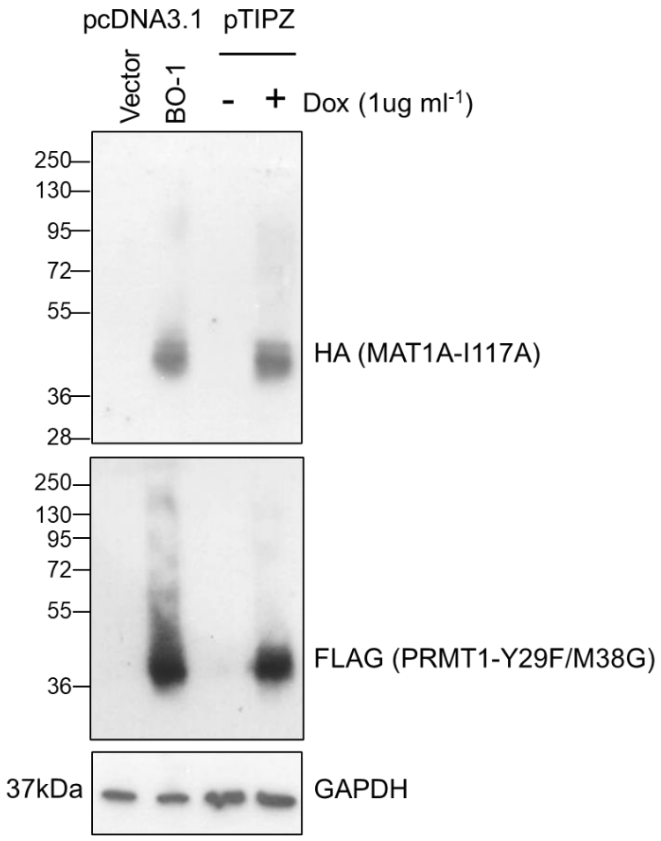


Figure 3.10: *In vitro* validation of doxycycline-inducible bio-orthogonal constructs

'BO-1' sequence was overexpressed by transient transfection of pcDNA3.1 and pTIPZ vectors into HEK-293T cells for 24 hours. The pTIPZ variant of the contstruct was induced with 1μg/ml doxycycline for twenty-four hours prior to harvest. n=1.

3.3.3: Validation of *in vitro* bio-orthogonal pathway

HEK-293T cells were transiently transfected with the range of four MAT/PRMT1 combinations, treated with AdOx to prevent endogenous methylation and cultured in media supplemented with Pob-methionine. Incorporation of alkynylated arginine was then determined by CuAAC-mediated TAMRA conjugation and in gel fluorescence. Surprisingly, the most striking difference in protein modification was observed between constructs containing HA-MAT1A-I117A and HA-MAT2A-I117A, with alkynylation occurring at higher efficiency in samples containing MAT2A-I117A. In contrast, both constructs containing MAT1A-I117A only generated a TAMRA signal comparable to the empty vector control **(Figure 3.11 A)**.

When comparing the efficiency of FLAG-PRMT1 with FLAG-PRMT1-Y29F/M38G for utilisation of intracellular Pob-SAM, the alkynylation pattern of the wild type suggested at least equal if not slightly greater efficiency compared to the bio-orthogonal mutant **(Figure 3.11 A, lanes 3 and 4)**. This result came as a surprise, due to the reported bio-orthogonal interaction between Pob-SAM and PRMT1-Y29F/M38G, which contrasted with an almost total inertness of Pob-SAM towards wild-type PRMT1[106].

TAMRA signal strength in the molecular weight region of approximately 25-40kDa appeared to slightly favour wild-type FLAG PRMT1 **(Figure 3.11 A, lane 3 marked by a green box)**, although this difference was minor. Both enzymes generated a dominant band at approximately 40kDa suggestive of auto-

alkynylation **(Figure 3.11 A, marked by an arrow);** however, this was much stronger in wild-type FLAG-PRMT1, despite expressing at a slightly lower level than FLAG-PRMT1-Y29F/M38G **(Figure 3.11 B).** Overall, validation of the full *in vitro* bio-orthogonal pathway suggested clear conversion of Pob-methionine to Pob-SAM by HA-MAT2A-I117A, but a relatively indistinct alkynylation signature when comparing FLAG-PRMT1-Y29F/M38G to a wild-type control.

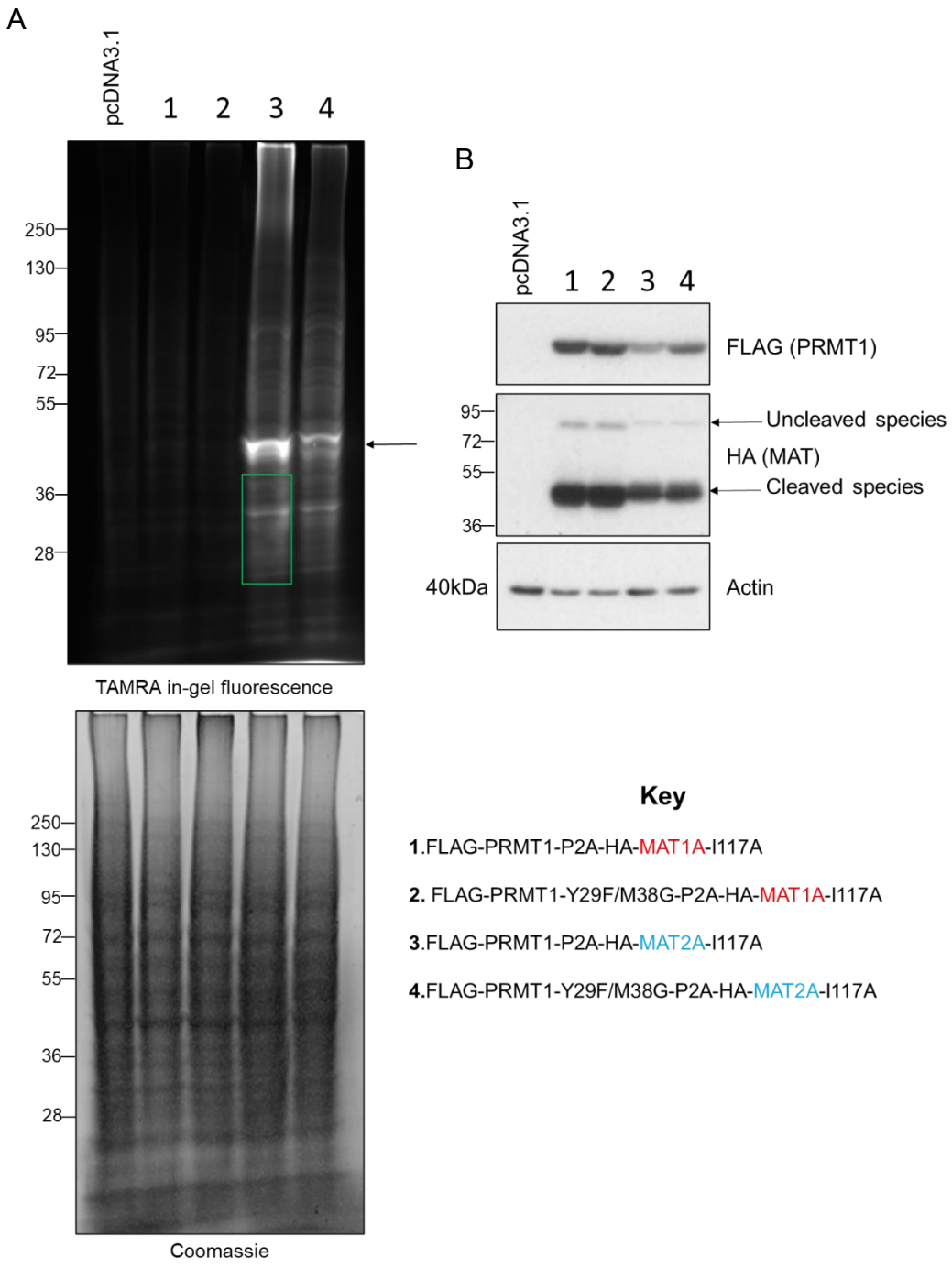


Figure 3.11: Comparison of *in vitro* efficacy between bio-orthogonal enzyme combinations

(A) Top panel: TAMRA in-gel fluorescence readout of *in vitro* alkynylation of whole cell proteome. Lower panel: Coomassie stain of the same gel shown in the top panel showing total protein

content of the gel. HEK-293T cells were transfected with constructs outlined in the figure key and treated with 10μM AdOx. 24 hours later, 2mM (final concentration) Pob-methionine was added to cell medium for 4 hours. Cells were then lysed and alkynylated protein visualised by CuAAc-mediated addition of a TAMRA-azide probe. **(B)** Western blot showing expression of MAT and PRMT1 variants in each sample. n=1.

3.4: Discussion

The data presented in this chapter show clear progress towards validation of a bio-orthogonal system that could be used for the genome-wide profiling of H4R3me2a, providing a much-needed comparison between normal and breast cancer cells. However, it is clear that more optimisation is required before the system can be utilised for epigenetic profiling of PRMT1 activity.

At the level of cell-free validation, recombinant His-PRMT1-Y29F/M38G demonstrates the ability to utilise Pob-SAM to alkynylate histone H4. Although no data have been acquired comparing performance of His-PRMT1-Y29F/M38G to a wild-type His-PRMT1 control, comparison of the wild-type and bio-orthogonal variants of the enzyme was possible in an semi cell-free context using endogenously-generated FLAG-tagged versions of both enzymes. Under these circumstances, FLAG-PRMT1-Y29F/M38G consistently alkynylated whole-cell protein extract more effectively than the wild-type counterpart, suggesting a certain degree of enzyme-cofactor bio-orthogonal specificity between PRMT1-Y29F/M38G and Pob-SAM.

Validation of the *in vitro* bio-orthogonal pathway requires further optimisation. One encouraging feature of the experiment was the clear distinction in TAMRA (alkynylation) signal when comparing samples containing HA-MAT1A-I117A and HA-MAT2A-I117A (with HA-MAT2A-I117A producing a greatly enhanced signal relative to HA-MAT1A-I117A). This disproves suspicion from previous attempts that poor signal was derived from inefficient conversion of Pob-methionine to

Pob-SAM (data not shown), a conclusion prompted by semi cell-free experiments showing that the ‘second half’ of the labelling pathway works in the presence of active cofactor.

Inclusion of HA-MAT1A-I117A under the impression that it may provide superior conversion of Pob-methionine to Pob-SAM was unsuccessful, although inability of HA-MAT1A-I117A to convert Pob-methionine to Pob-SAM did inadvertently provide a robust negative control for the conversion step of the pathway. Such incapability may arise for a number of reasons. Despite the conserved position of the isoleucine residue at position 117 within the methylio binding pocket, the MAT1A binding pocket is not identical to the MAT2A binding pocket, and a change in one of the key residues (G120 is a cysteine in MAT1A) may somehow alter the interaction of the bound cofactor with the enzyme binding pocket. Cysteine is a slightly larger amino acid than glycine due to an amine group side-chain. Although speculative, this may be sufficient to sterically hinder binding of Pob-methionine, or alternatively, alter the orientation of enzyme-cofactor interaction in such a way that mutation of I117 is unnecessary or even detrimental.

As such, further engineering of MAT1A to find the correct residue with which to make the ‘hole’ would be necessary before it would be usable for a bio-orthogonal pathway. One potentially useful modification would be generation of an I117A/C120G double mutant to render the conserved binding residues identical

to those in the MAT2A-I117A mutant. If functional, the higher V^{max} of the MAT1A homotetramer could result in an even more effective cofactor converter[146].

True enzyme-cofactor exclusivity appears to be the primary issue in demonstrating a viable *in vitro* bio-orthogonal pathway. Currently, a clear and distinct enhancement in Pob-SAM utilisation by PRMT1-Y29F/M38G relative to wild-type PRMT1 has not been demonstrated within a fully *in vitro* context. This may simply be a consequence of technical difficulty within the experiment, and numerous issues can confound the final readout. Chiefly among these are expression level of the enzymes in question and overall protein content that is involved in the CuAAC reaction and subsequently added to the gel. Coomassie staining in **Figure 3.11 A** suggests that slightly more protein was present in the lane containing the MAT2A-I117A/PRMT1 pathway **(Figure 3.11 A Lane 3)** than in the lane containing the MAT2A-I117A-/PRMT1-Y29F/M38G pathway **(Figure 3.11 A Lane 4).** This could account for the greater signal strength in certain regions of lane 3 (such as the region generally spanning the 25-40kDa range **Figure 3.11 A, lane 3 green box),** especially since expression of wild-type FLAG-PRMT1 was slightly lower than FLAG-PRMT1 Y29F/M38G, (which would otherwise infer enzymatic 'superiority' of wild-type PRMT1 over the double mutant in this scenario).

The intense doublet denoted by an arrow could correspond to auto-alkynylated PRMT1 that is subsequently modified by TAMRA. Auto-methylation of PRMT1 has been reported[143] and is a frequent occurrence in our cell-free methylation

assays using recombinant GST-PRMT1 **(see Figure 4.21**). Another explanation is TAMRA conjugation to Pob-SAM whilst it is enzyme-bound. This would explain two bright bands at a similar size, since both PRMT1 and MAT2A resolve within the 40-45 kDa range. Thus, a scenario in which HA-MAT2A-I117A appears to be modified due to TAMRA addition to *de novo*- generated Pob-SAM that is retained within the enzyme, and FLAG-PRMT1 appears to be modified due to TAMRA addition to bound but unused Pob-SAM is plausible. Further, Pob-SAM retention by PRMT1 and auto-alkynylation are not mutually exclusive scenarios.

Despite relatively equal band definition and intensity generated by wild-type and bio-orthogonal PRMT1 **(Figure 3.11 A lanes 4 and 5)** from ~50kDa upwards, this is not sufficient to infer 'bio-orthogonal' status for the PRMT1 double mutant *in vitro.* In addition, despite an advantage in favour of wild-type PRMT1 due to a discrepancy in protein content, the signal strength generated relative to PRMT1-Y29F/M38G conflicts with the supposedly negligible activity of wild-type PRMT1 towards Pob-SAM reported by Wang et al[106].

Thus, in order to demonstrate viability of this pathway *in vitro*, certain adjustments will have to be made to the experiment. First, a protein standardisation step prior to TAMRA conjugation could help mitigate imbalances in loading. Incubation time of cells in Pob-methionine could potentially be reduced, as lack of total specificity between Pob-SAM and PRMT1-Y29F/M38G means that other endogenous methyltransferases may be utilising free Pob-SAM, thus generating background signal. In the vein of this principle, a shorter time of incubation (<4hours) may

expose a crucial window where PRMT1-Y29F/M38G-dependent alkynylation still occurs sufficiently for bio-orthogonal profiling, but wild-type-dependent alkynylation is comparatively minimal. If such a window of stringency exists, then wild-type PRMT1 could be used as a ‘base line’ control for normalisation of background noise.

Finally, there is a possibility that the enzyme-cofactor pairing is not optimal, despite optimisation of experimental conditions. Although the M38G mutation appears to reduce steric hindrance placed upon the variable side chain of Pob-SAM, the Y39 residue does not appear to have sufficient proximity to the cofactor to warrant mutation to phenylalanine – a substitution that could potentially de-stabilise the structure of the binding pocket (personal communication, Professor Stephen Smerdon, University of Birmingham). As such, use of an M38G single mutant may be more effective, despite Wang et al reporting a 60-80% modification rate with this mutant, compared to the supposed 80-100% rate observed with the double mutant[106]. In the event that published bio-orthogonal mutants do not enhance Pob-SAM utilisation relative to wild-type PRMT1, a ‘saturation approach’ to CliEn-seq using wild-type PRMT1 could be employed. By including a MAT2A-I117A only control, background enrichment due to non-specific utilisation of the cofactor by endogenous enzymes could be accounted for. In addition, experimentation with another cofactor may be beneficial. When presenting proof-of-principle data for the *in vitro* pathway, Wang et al used engineered variants of lysine methyltransferases G9a and GLP, combined with a similar cofactor precursor S-(E)-hex-2-en-5-ynyl homocysteine (Hey-methionine). MAT2A-I117A was used to catalyse conversion of Hey-met to Hey-

SAM^{108} which could be subsequently utilised by PRMT1-Y29F/M38G (or a single M38G variant) with a perhaps greater degree of exclusivity.

Chapter 4: SPIN1 as a PRMT1 substrate in breast cancer

4.1: Identification of novel PRMT1 interactor SPIN1 using SILAC quantitative proteomics

4.1.1: Quantitative comparison of PRMT1v1 interactome in normal and breast cancer cells

MCF10A and MCF7 cells represent normal mammary epithelial and ER^+ breast cancer models respectively. Thus, quantitative comparison of their PRMT1 interactomes would allow identification of substrates that interact with PRMT1 at a greater frequency in breast cancer. In order to identify novel breast cancer-enriched substrates of PRMT1, SILAC (Stable Integration of Amino acids in Cell culture) quantitative proteomics was employed. First used in 2002, this technology offers a robust and reproducible method for quantitative proteomics without the need for prior modification of peptides, as is necessary for similar methods such as iTRAQ[147].

For the purpose of this investigation, this methodology involved culturing MCF10A and MCF7 cells in medium containing either light or heavy arginines and lysines, respectively. The mass increase in the heavy arginines and lysines is contributed by the presence of ^{13}C and ^{15}N, which cause a detectable mass difference when peptides derived from both cell lines are mixed for quantitative comparison via mass spectrometry **(Figure 4.1 A)**. Both MCF7 and MCF10A cell lines stably overexpressed FLAG-PRMT1 variant 1 (the predominantly nuclear isoform[57]) in equal levels (as determined by western blotting, **Figure 4.1 B**) to ensure that PRMT1 interaction with substrates was normalised between cell

lines. This is important, as unequal contribution of sample derived from either cell line would ultimately skew the quantitative data, rendering PRMT1:substrate enrichment ratios non-representative. As shown in **Figure 4.1 C**, FLAG-PRMT1 was effectively enriched in immunopreciptated complexes and a number of associated proteins co-immunoprecipitated. Lanes from two repeats consisting of mixed heavy/light immunoprecipitation were cut into bands and sent to Dr. Mark Skehel (LMB, Cambridge) for analysis and protein identification. A further two repeats were sent without SDS-PAGE resolution, and were instead subjected to low pH glycine elution (conducted by the Skehel group, LMB, Cambridge) prior to mass spectrometry analysis.

The ectopic FLAG-PRMT1 was also a catalytically inactive variant often referred to as 'PRMT1 XGX' in the literature, where residues S69, G70, and T71 were all mutated to alanines (**Figure 4.1 A)**[144]. The rationale behind this modification was to prevent the completion of FLAG-PRMT1-mediated catalysis, thus reducing dissociation of FLAG-PRMT1 from its substrates, increasing the co-immunoprecipitation (co-IP) efficiency of substrates with FLAG-PRMT1.

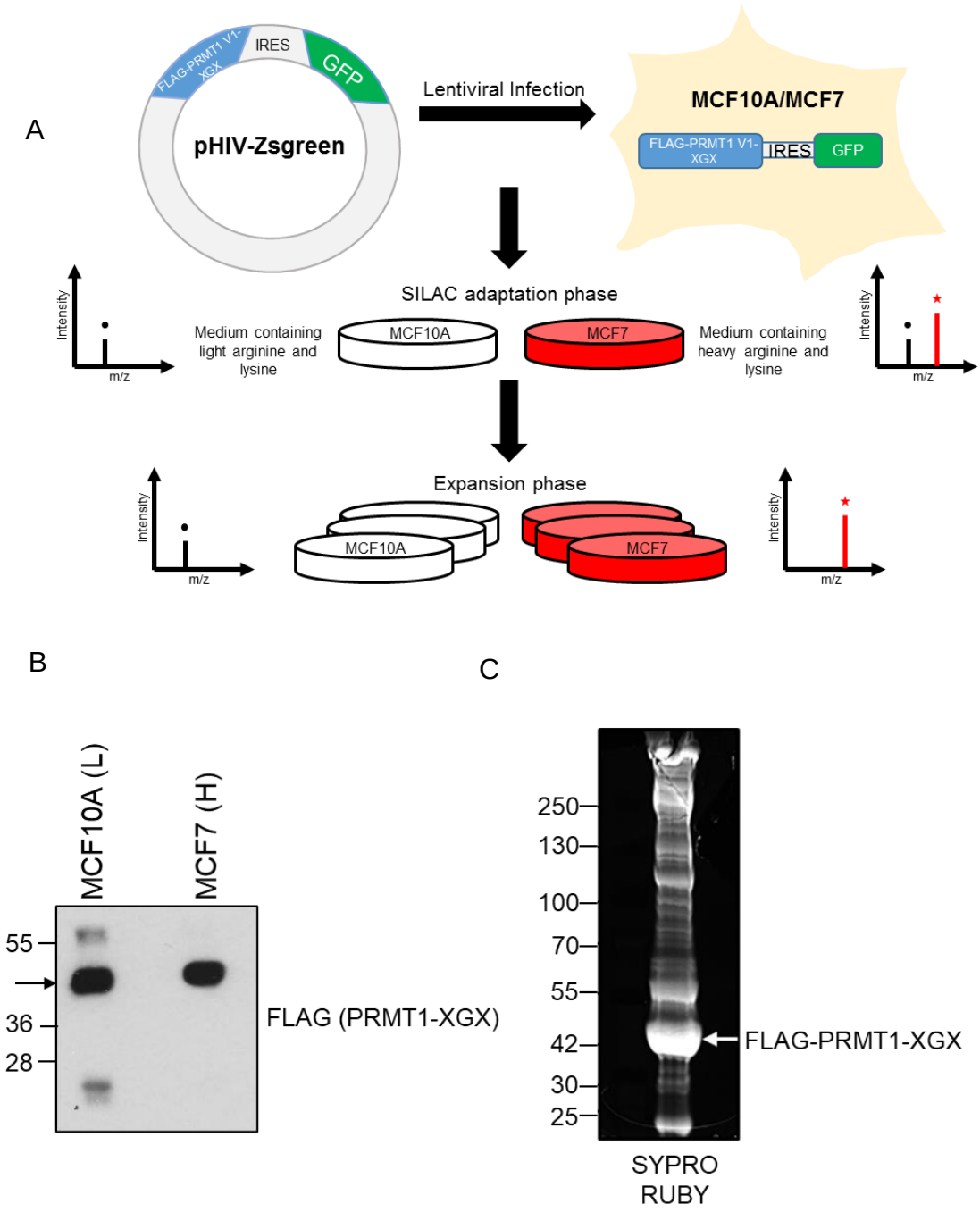


Figure 4.1: SILAC workflow for breast cancer-enriched PRMT1 interactor discovery

(A) Schematic showing generation and culture of SILAC cell lines in preparation for quantitative proteomics. Both non-transformed MCF10A and ER$^+$ breast cancer MCF7 cells stably overexpress FLAG-PRMT1-XGX. Culture of cell lines in medium with either light or heavy amino acid content causes a distinction in mass signature of a protein (as determined by mass spectrometry) depending on the amino acid supplementation, and therefore cell line of origin. **(B)**

Representative Western blot showing equal loading of FLAG-PRMT1-XGX substrate bait (marked by arrow) for both MCF10A and MCF7 cells respectively during SDS-PAGE resolution of substrates. '(L)' and '(H)' refer to light and heavy amino acid-containing media, respectively. **(C)** SDS-PAGE resolution of mixed co-immunoprecipitation of endogenous proteins with ectopic FLAG-PRMT1-XGX derived from MCF10A and MCF7 SILAC samples. n=4.

4.1.2: Triaging of PRMT1v1-interacting proteins

Mass spectrometry data were triaged by Dr. Agnieszka Zielinska of the Davies Laboratory. This was accomplished by setting two criteria: the FLAG-PRMT1-XGX:substrate interaction had to be enriched in MCF7 cells over MCF10A with a log fold change threshold of 1.5 or greater, and this event had to be present in two or more of the four repeats analysed. The resulting 30 candidate proteins are listed in **Figure 4.2.**

Out of this gene cohort SPIN1 was chosen for further study, despite several other genes appearing as enriched interactors in MCF7 cells more frequently and/or producing log fold change scores that were higher than those acquired for SPIN1 **(Figure 4.2)**. All such proteins (NT5E, LGALS1 OASL, QPCTL) were predominantly characterised in immune signalling (even within the context of cancers), eliminating them from further study due to the immune response falling outside our research topics. In addition, TRIM21 (which ranked mid-table on average enrichment and appeared to be enriched in all repeats) could have a case to be made for further study. However TRIM21 is a known antibody-binding protein[148], suggesting that its enrichment could be an artefact of the immunoprecipitation process, rather than a biological interaction with FLAG-PRMT1-XGX. Publicly-available mass spectrometry data (www.phosphosite.org) revealed that many of these putative interactors of PRMT1 are predicted to be arginine methylated. Indeed, TAF15, TRIM21 and LMNA have been characterised as substrates of PRMT1, PRMT5 and CARM1, respectively[149–151]. Further, cross-referencing with publicly available interaction data from the Harmonizome[152] revealed that many of these proteins are already identified as

PRMT1 interactors, further increasing confidence in the legitimacy of these interaction phenomena. These interaction data included SPIN1, which ranked as the second-highest in breast cancer enrichment of this sub-cohort. **(Figure 4.2).**

In addition to its high average enrichment, SPIN1 was chosen for its capacity as a histone code reader[48], in keeping with the epigenetically focussed approach to arginine methylation biology in this thesis. Thus, it was hypothesised that interaction of SPIN1 with PRMT1 may influence transcriptional coactivator activity of SPIN1, perhaps as a consequence of PRMT1-mediated methylation of SPIN1. The proclivity of both proteins to localise to the promoters of actively transcribed genes (PRMT1 as an epigenetic writer, SPIN1 as a reader) also supports this concept. In addition, SPIN1 interacts with asymmetric dimethylarginine (H3R8me2a, which has been demonstrated as oncogenic[47]), linking it yet more intimately to the role of arginine methylation in cancer.

Protein Name	No. repeats enriched	Average Log fold change (enrichment MCF7:MCF10A)	*P*-values	Documented Arginine Methylation
NT5E	2	4.95	<0.0001, -	No
LGALS1	2	4.53	0.071, <0.0001	No
OASL	2	4.41	-, -	No
QPCTL	2	4.10	0.0019, -	No
SPIN1	2	3.93	-, NS	No
DPYSL3	2	3.93	0.0038, -	Yes
DOCK4	3	3.86	<0.0001, <0.0001, NS	No
WDR1	2	3.76	NS, <0.0001	Yes
RPL24	2	3.74	0.031, -	Yes
HSPA5	2	3.71	<0.0001, <0.0001	Yes
ELMO2	2	3.64	0.064, <0.0001	No
TGM2	3	3.60	0.0018, 0.00066, NS	No
ARHGAP22	2	3.46	<0.0001, -	No
TAF15	2	3.18	0.0017, 0.0097	Yes (PRMT1)
TRIM21	4	2.86	0.00093, <0.0001, <0.0001, 0.00093	Yes (PRMT5)
FXR1	3	2.85	<0.0001, 0.0059, NS	Yes
THRAP3	2	2.75	0.0023, <0.0001	Yes
CTSB	2	2.67	0.013, 0.027	No
ERH	2	2.44	NS, 0.00028	Yes
PML	4	2.42	<0.0001, 0.0022, 0.0032, 0.00023	Yes
RRAS2	2	2.39	0.012, NS	No
CCT2	2	2.34	NS, NS	Yes
JAK1	2	2.28	0.00069, NS	No
HKDC1	2	2.21	0.0048, 0.041	No
STK38	2	2.12	0.026, 0.0001	Yes
SPTAN1	2	1.99	<0.0001, 0.028	Yes
VDAC2	2	1.99	0.0005, NS	No
RPLP2	2	1.89	NS, NS	No
AP2M1	2	1.73	0.033, NS	No
LMNA	2	1.72	0.0022, NS	Yes (CARM1)

Figure 4.2: Breast cancer enriched PRMT1-interacting proteins

SILAC quantitative proteomics identifies 30 proteins whose interaction with PRMT1 is significantly enriched in MCF7 compared to MCF10A cells. Genes are listed in descending order according to their average Log fold change enrichment in MCF7 over MCF10A cells. *P*-values correspond to individual repeats in which a particular PRMT1 interactor was breast cancer-enriched. NS = not significant, '-' = *P*-value unknown. Predicted arginine methylation targets (www.phosphosite.org) are highlighted in red in the right-hand column, and green where a PRMT has been ascribed to such an event. Arrows denote PRMT1 interactors that are already

documented in the Harmonizome database, using the Pathway Commons Protein-Protein Interactions dataset[152]. SPIN1, which was chosen for further investigation, is highlighted in yellow.

4.1.3: SPIN1 Expression in Cancer

Once chosen for further investigation as a putative breast cancer-enriched interactor of PRMT1, the expression levels of SPIN1 in cancer were investigated using a combination of publicly available data and western blotting in relevant cell lines (MCF10A and MCF7). Patient sample data showing SPIN1 expression at the mRNA level suggests that high SPIN1 expression actually correlates with an improved prognosis across all breast cancer subtypes, and in the specific case of ER^+ breast cancers **(Figure 4.3 A)**. Interestingly, co-expression of SPIN1 with PRMT1 using patient data across all breast cancer subtypes shows a negative correlation in expression at the transcript level, whilst expression of both genes correlate positively at the protein level **(Figure 4.3 B),** suggesting that protein-level monitoring may provide more representative insight into SPIN1 expression levels and its potential role in cancer.

Expression of SPIN1 and PRMT1 in MCF7 breast cancer and MCF10A mammary epithelial cells was assessed by western blotting throughout various stages of cell confluence, to account for potential fluctuations in expression of either gene as a result of differential cell cycling. This comparison suggests that PRMT1 levels remain consistently higher in MCF10A cells throughout proliferation, lowering once cells reach confluence. Conversely, MCF7 cells maintain a consistent expression level of PRMT1 that does not appear to be perturbed by confluence, thus surpassing PRMT1 levels observed in confluent MCF10As. By contrast, SPIN1 levels remain consistently higher past the 40% confluence stage in MCF7 cells compared to MCF10A cells, implying a generally higher expression level in ER^+ breast cancer **(Figure 4.4).**

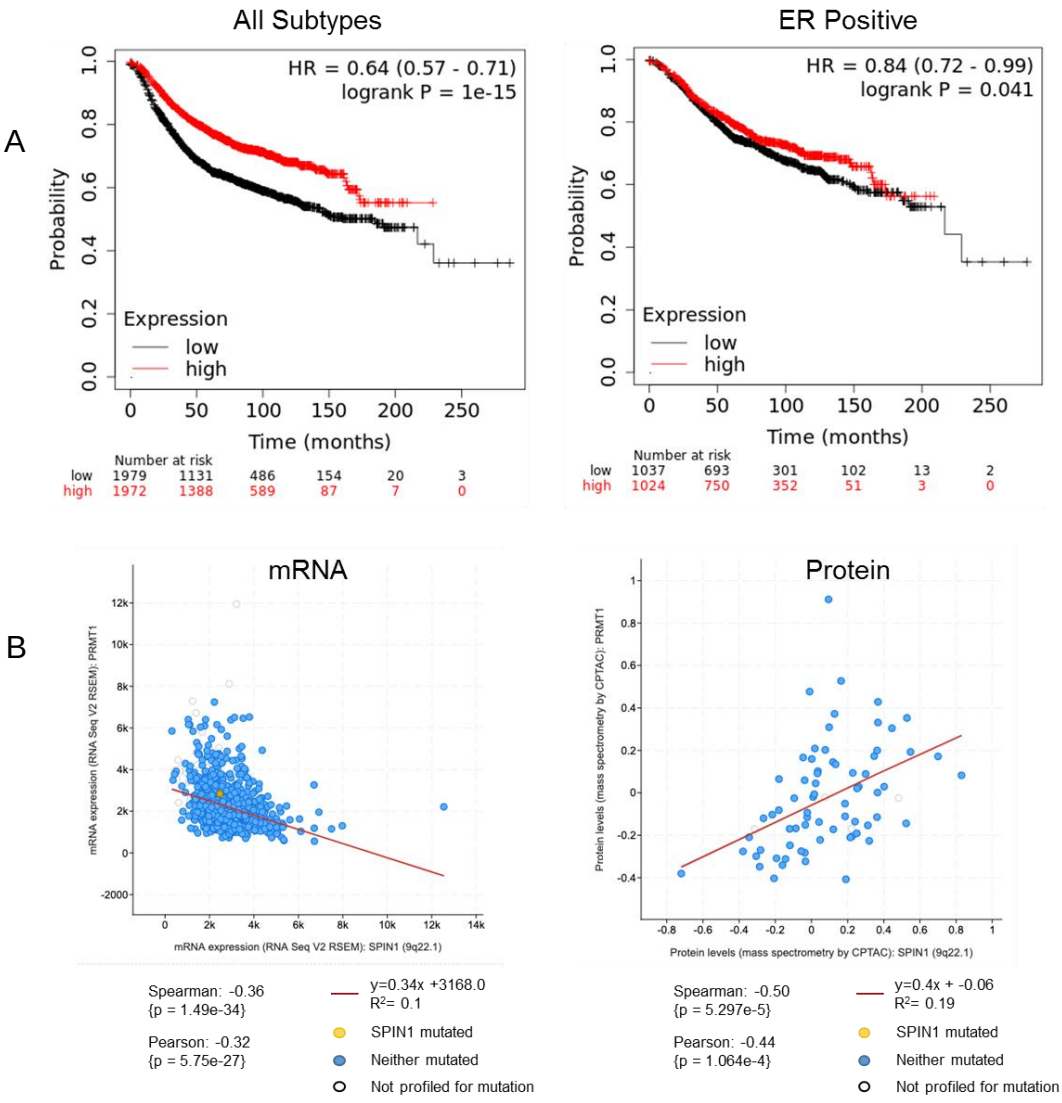


Figure 4.3: SPIN1 Expression in breast cancer patients

(A) Kaplan-Meier plots derived from SPIN1 mRNA expression across all breast cancer subtypes (left panel) and ER-positive breast cancer (right panel). Data gathered from www.kmplot.com, probe ID: 211005_at. HR denotes hazard ratio for high versus low SPIN1 expression. Data were derived from the following datasets: E-MTAB-365, E-TABM-43, and GSE datasets 11121, 12093, 12276, 1456, 16391, 16446, 16716, 17705, 17907, 18728, 19615, 20194, 20271, 2034, 20685, 20711, 21653, 22093, 25066, 2603, 26971, 29044, 2990, 31448, 31519, 32646, 3494, 36771, 37946, 41998, 42568, 43358, 43365, 45255, 4611, 46184, 48390, 50948, 5327, 58812, 61304,

65194, 6532, 69031, 7390, 76275, 78958, 9195. **(B)** Graphs depicting relationship between SPIN1 and PRMT1 expression level in breast cancer using mRNA (left panel) and protein level (right panel). Generated using www.cbioportal.org.

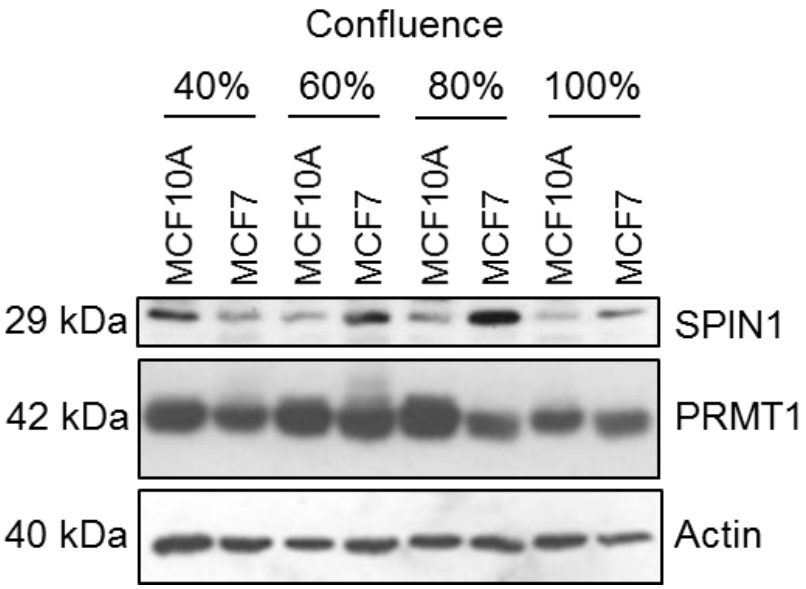


Figure 4.4: SPIN1 Expression in normal and malignant breast cells

Western blot showing endogenous expression of SPIN1 and PRMT1 in MCF10A and MCF7 cells. This was observed throughout various stages of confluence, in order to more accurately assess expression of both genes, since cell growth and participation in the cell cycle (both processes influenced by confluence) could potentially affect expression of either protein. n=1.

4.1.4: Validation of SPIN1-PRMT1 interaction *in vitro*

Prior to investigating SPIN1-PRMT1 contributions to the breast cancer phenotype, it was necessary to validate the interaction observed in the SILAC data between PRMT1 and SPIN1 *in vitro.* Preliminary validation involved transient co-transfection of Myc-tagged PRMT1-XGX with GFP-fused SPIN1 into HEK-293T cells, and immunoprecipitation of GFP-SPIN1. Western blotting showed robust and reproducible co-immunoprecipitation of Myc-PRMT1-XGX with GFP-SPIN1 demonstrating interaction of SPIN1 and PRMT1 *in vitro* **(Figure 4.5 A)**.

Despite proving *in vitro* interaction of SPIN1 and PRMT1, further investigation was necessary to prove this phenomenon unambiguously, using the relevant cell models. Initial attempts to immunoprecipitate endogenous SPIN1 in MCF10A and MCF7 cells using an anti-SPIN1 antibody were unsuccessful (data not shown), causing a resort to IP using affinity-tagged ectopic SPIN1. In order to overexpress SPIN1 in MCF10A and MCF7 cells (for the purposes of interaction validation, but also future experiments such as *in vitro* methylation assays, see section **4.2.6**), a stable overexpression system was opted for, since MCF7 and MCF10A cells cannot be transiently transfected efficiently. However, constitutive stable overexpression of FLAG-SPIN1 in MCF10A and MCF7 cells resulted in drastic morphological transformation, with cells forming bloated, amorphous masses suggestive of multinucleation/incomplete mitotic separation. At approximately two weeks after generation, populations crashed due to mass cell death (data not shown). Due to this observation in behaviour of both cell types, and considering

that SPIN1 overexpression has produced similar results in HeLa cells[124], this was considered a direct consequence of FLAG-SPIN1 overexpression.

To circumvent this issue, FLAG-SPIN1 was cloned into the pTIPZ vector **(see section 3.3.1)** in order to generate cell lines that could overexpress FLAG-SPIN1 after treatment with doxycycline. These MCF10A and MCF7 pTIPZ-FLAG-SPIN1 cells were subsequently used in an attempt to co-immunoprecipitate endogenous PRMT1. Although PRMT1 co-immunoprecipitated with FLAG-SPIN1 in both cell lines, the signal derived from MCF7 pTIPZ-FLAG-SPIN1 cells was noticeably stronger than the analogous signal obtained from MCF10A, despite the latter overexpressing FLAG-SPIN1 to a greater degree **(Figure 4.5 B)**. Thus, this observation implied an enhanced rate of interaction between SPIN1 and PRMT1 in MCF7 cells over MCF10A, validating the enriched interaction in MCF7 cells obtained from SILAC.

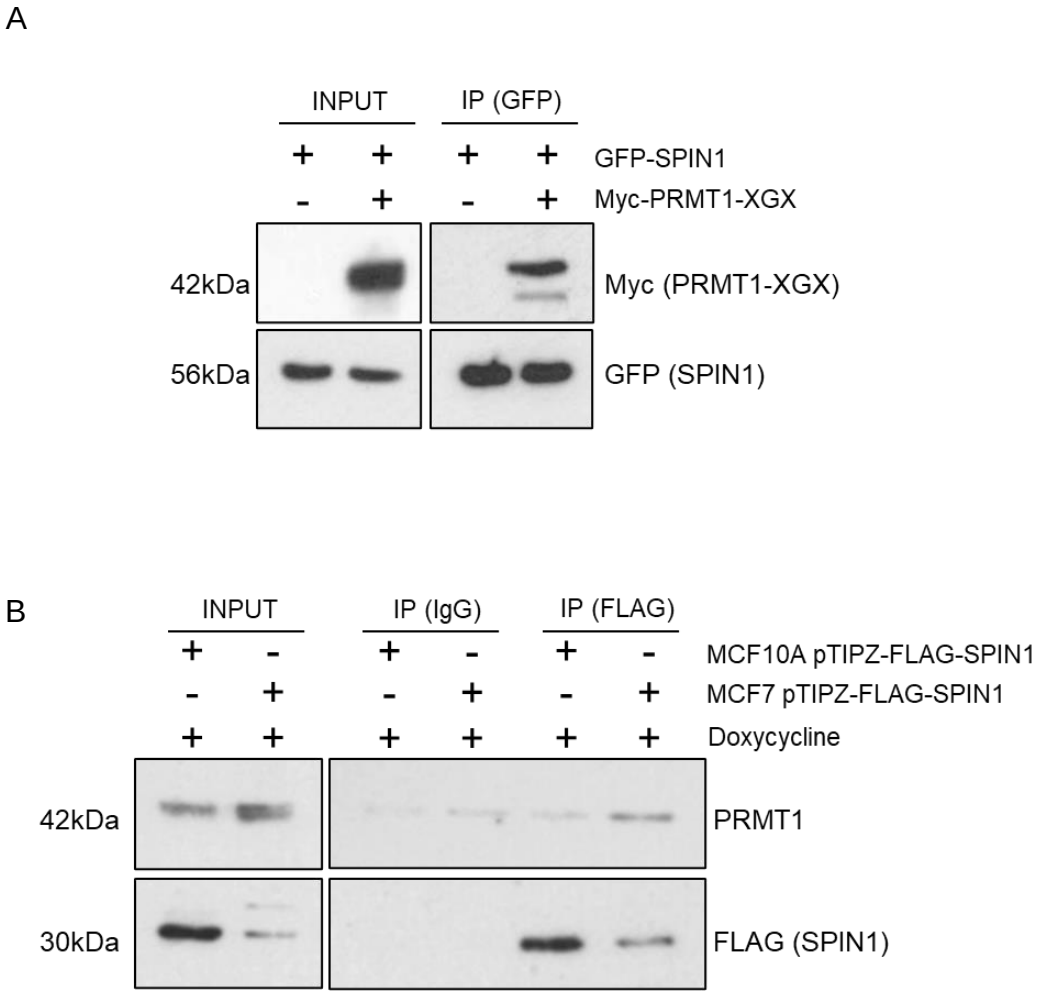


Figure 4.5: Validation of SPIN1-PRMT1 interaction *in vitro*

(A) Western blot showing co-immunoprecipitation transiently-overexpressed myc-PRMT1-XGX and GFP-SPIN1 in HEK-293T cells, 24 hours after co-transfection into HEK-293T cells. n=3. **(B)** Western blot showing co-immunoprecipitation of FLAG-SPIN1 with endogenous PRMT1 in doxycycline inducible pTIPZ-FLAG-SPIN1 MCF10A and MCF7 cells, 48 hours after induction of FLAG-SPIN1 expression with 1μg/ml doxycycline. n=1.

4.2: Methylation of SPIN1 by PRMT1

4.2.1: SPIN1 as a potential substrate of PRMT1

Once interaction between SPIN1 and PRMT1 had been confirmed *in vitro,* the implication that SPIN1 may be methylated by PRMT1 was investigated. Although no arginine methylation events have been identifed so far on SPIN1, including global mass spectrometry analysis using pan-methyl-arginine antibodies for enrichment **(Figure 4.6 A),** the protein contains 15 arginine residues, many of which are highly conserved in vertebrates **(Figure 4.6 B).** Further, one of these residues (R152) forms an RG motif (the preferred target motif of PRMT1[22]) and is placed in an exposed position on top of Tudor domain II, making it a particularly attractive prospect for methylation **(Figure 4.7)**. Taken together, SPIN1 appeared a likely target for methylation of arginine residues by PRMT1.

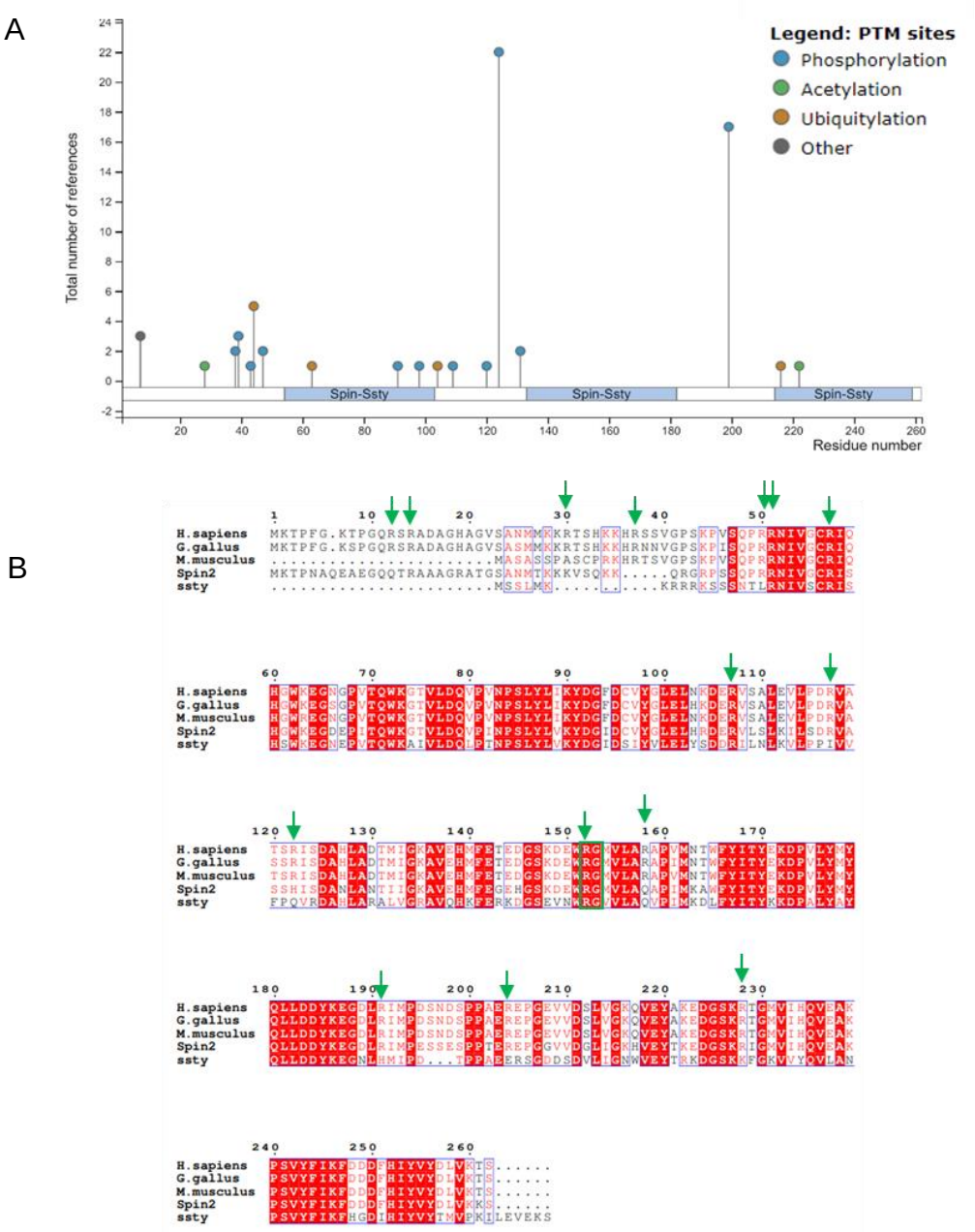


Figure 4.6: Known and potential post-translational modification sites of SPIN1

(A) Schematic of known post-translational modifications of SPIN1, and their relative position in the amino acid sequence (data obtained from www.phosphosite.org). **(B)** Amino acid sequence alignment of human SPIN1 (NCBI accession AAG38112), chicken *Gallus gallus* (NCBI accession NP_989964) and mouse *Mus musculus* (NCBI accession NP_035592). Also included are the human orthologue SPIN2 (NCBI accession NP_001006684) and the paralogue ssty, derived from mouse (NCBI accession NP_033246). Green arrows indicate position of arginine residues. Green box indicates position of R152 within an RG motif. Adapted from Zhao et al., 2007 [111].

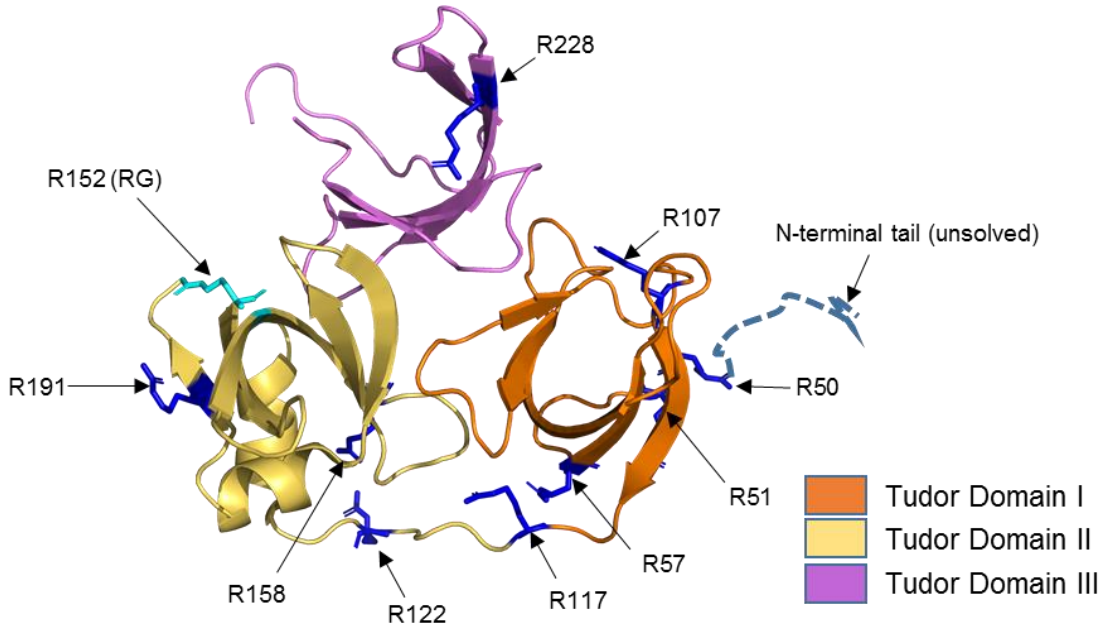


Figure 4.7: Arginine residue placement in the SPIN1 crystal structure

Crystal structure of monomeric SPIN1 with all solved arginine residues (R12,R14,R30,R37 and R204 are missing) shown in blue, with the exception of R152/RG, which is shown in cyan. The blue dotted line represents the theoretical position of the unstructured and unsolved N-terminal tail. Residues 197-204 are also missing from the structure. PDB ID: 2NS2.

4.2.2: Cell-free methylation of SPIN1 by PRMT1: Initial attempts

The initial approach to investigate PRMT1-dependent methylation of SPIN1 involved generation of recombinant His-SPIN1 **(Figure 4.8 A)**, which was placed into cell-free methylation assays with GST-PRMT1 and ^{3}[H]-SAM as a cofactor. This experiment initially proved successful, generating a robust autoradiographic signal corresponding to the molecular weight of His-SPIN1 **(Figure 4.8 B).** However, it was noticed that the pET-28a^{+} vector that was used to express recombinant SPIN1 contains two translated RG motifs between the N-terminal His-tag and the N-terminus of the SPIN1 protein. These motifs were found in the sequences that contribute to the BamHI restriction site and Thrombin cleavage sequence **(Figure 4.8 C)**. As such, these residues could potentially be methylated by GST-PRMT1, producing a false-positive result. In order to address this, the arginine residues in these sites were mutated to lysines producing a version of pET-28a+ that is devoid of the RG motif (pET-28a+ RG-KG) **(Figure 4.9 A).** In addition, recombinant GST-SPIN1 was also generated since GST is not a PRMT1 substrate in cell-free methylation assays **(Figure 4.9 B)**. Re-constitution of this assay demonstrated a loss of SPIN1 methylation when the RG motif within the thrombin cleavage site was mutated to KG, and the methylation signal was lost entirely when both RG motifs within the thrombin cleavage and BamHI restriction sites were mutated **(Figure 4.9 B, lanes 3 and 5, respectively).** Further, no methylation of GST-SPIN1 was detected **(lane 6)** whilst histone H4 was clearly methylated, validating the activity of the recombinant PRMT1 enzyme in this context. Taken together, these results suggested that recombinant SPIN1 was not methylated by recombinant PRMT1.

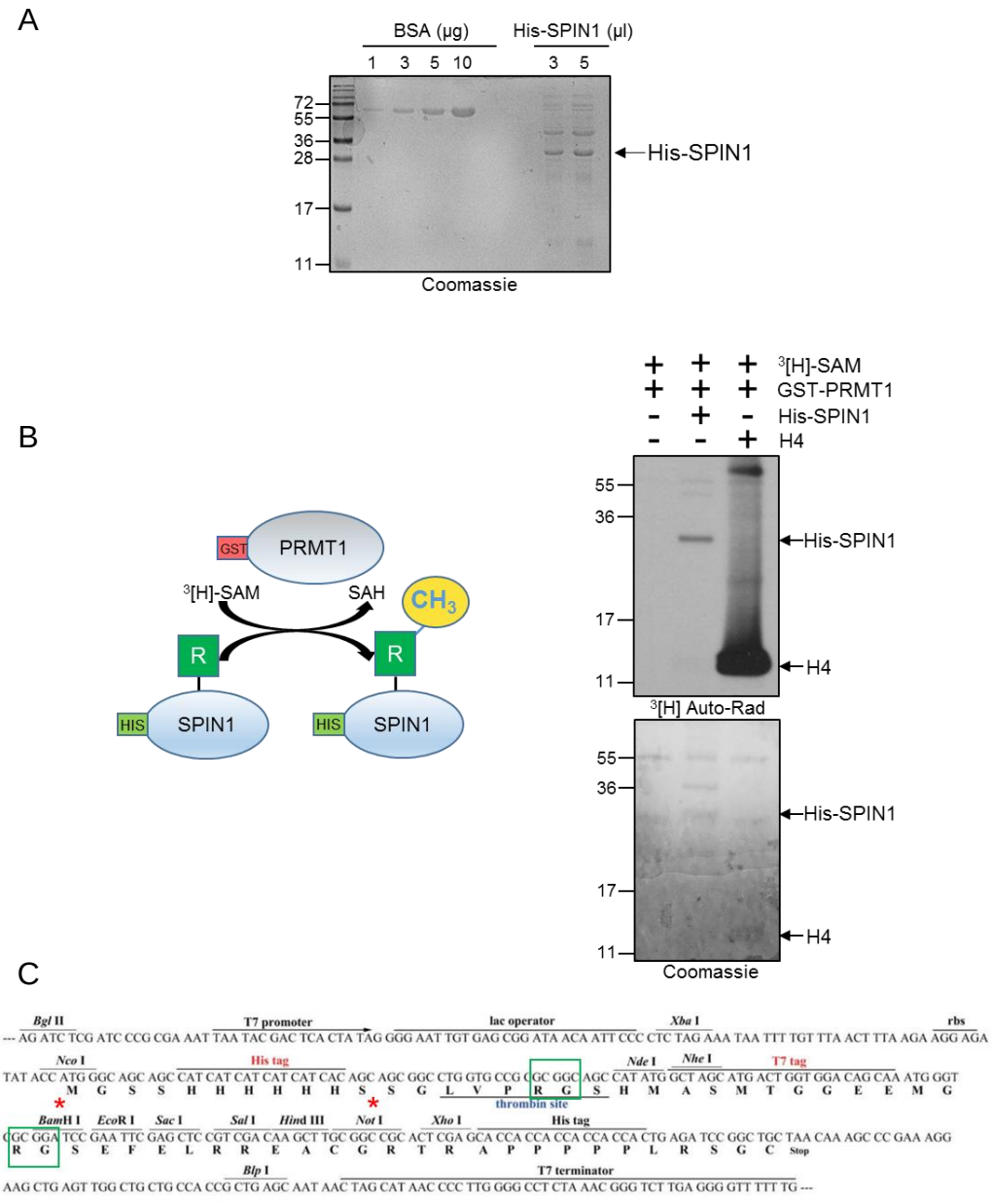


Figure 4.8: Cell-free methylation of SPIN1 generated from a non-mutated vector backbone

(A) Coomassie stain showing generation of recombinant His-SPIN1 expressed from the pET-28a+ vector. **(B)** Left panel: Diagram depicting constituents of cell-free methylation assay. Upper right panel: Autoradiographic readout of cell-free methylation of His-SPIN1. Lower right panel: Coomassie stain of autoradiograph in upper right panel, showing total protein content within the gel. n=2. **(C)** Schematic of pET-28a^{+} vector multiple cloning site situated between N and C-

terminal His-tags. RG motifs are boxed in green. Red asterisks mark restriction sites used to clone in the SPIN1 sequence. Adapted from https://www.genscript.com/gsfiles/vector-map/bacteria/pET-28a

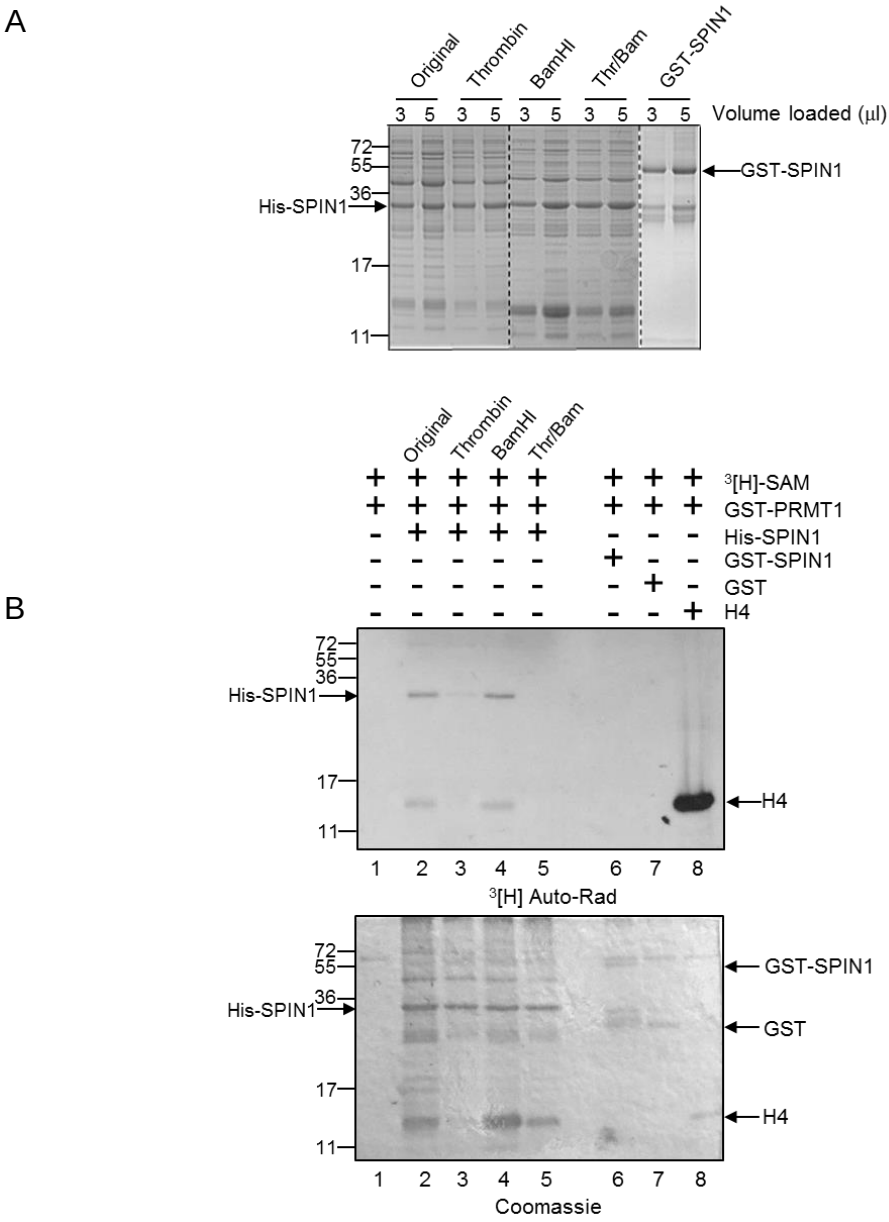


Figure 4.9: Cell-free methylation of SPIN1 generated from pET-28a+ (RG-KG)

(A) Coomassie stain showing generation of His-SPIN1 RG-KG vector mutants in Thrombin and BamHI sites, as well as GST-SPIN1. **(B)** Upper panel: autoradiographic readout of cell-free methylation of His-SPIN1 vector mutants and GST-SPIN1. Lower panel: Coomassie stain showing total protein content in autoradiograph in upper panel. n=2.

4.2.3 Semi cell-free methylation of SPIN1 by PRMT1

Although recombinant His-SPIN1 did not demonstrate evidence of methylation by recombinant GST-PRMT1, we have previously observed that RUVBL1, a key substrate of PRMT5 in the DDR, is not methylated in cell-free methylation assays (personal communication, Davies Laboratory), despite being a bona-fide substrate of PRMT5 *in vitro*[153]. Such discrepancies may result from improper folding of mammalian proteins in a bacterial expression system, or from an absence of other post-translational modifications that are required for methylation that may occur within the context of a human cell.

As such, a semi cell-free approach was utilised whereby HEK-293T cells were transfected with FLAG-SPIN1 that was subsequently immunoprecipitated and used as substrate in a cell-free reaction with recombinant PRMT1. Critically, after transfection, cells were treated with universal methylation inhibitor adenosine dialdehyde (AdOx) in order to prevent endogenous methylation of ectopically expressed FLAG-SPIN1, thus keeping the putative methyl-acceptor sites of FLAG-SPIN1 available for methylation by GST-PRMT1 **(Figure 4.10 A)**. This reaction used radioactive SAM as a cofactor, so that only cell-free methylation and not endogenous *a priori* methylation would be detected. The experiment reproducibly generated a robust methylation signal corresponding to the molecular weight of FLAG-SPIN1, suggesting that endogenously generated FLAG-SPIN1 could be methylated by recombinant GST-PRMT1. Importantly, a control sample in which FLAG-SPIN1-transfected HEK-293T cells were not treated with AdOx was included. These samples consistently generated a

negligible signal, implying that they had been methylated *a priori in vitro* **(Figure 4.10 B)**.

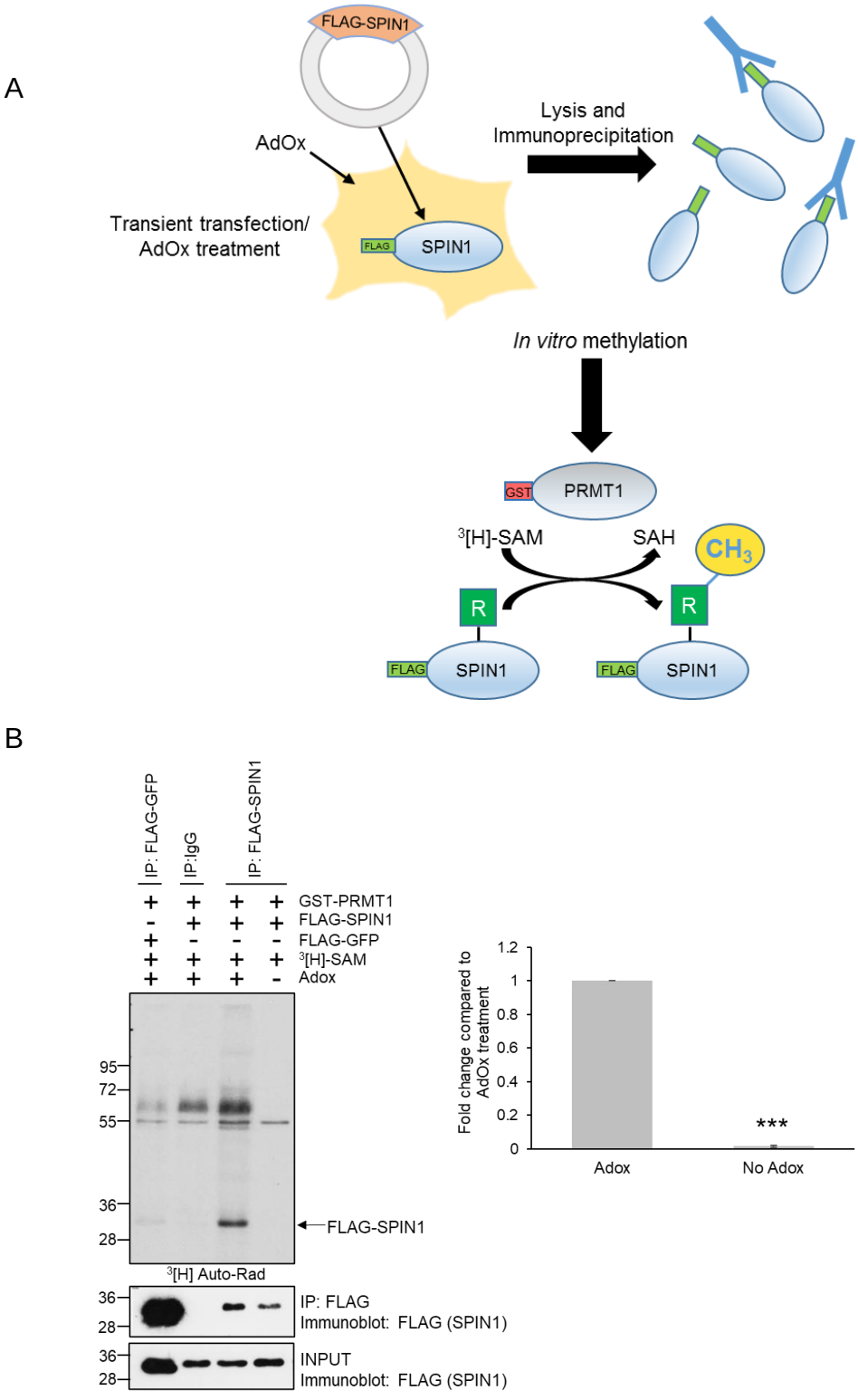


Figure 4.10: Semi cell-free methylation of SPIN1 by PRMT1

(A) Schematic of semi cell-free methylation of FLAG-SPIN1. HEK-293T cells are transiently transfected to overexpress FLAG-SPIN1. Cells were treated with 100μM AdOx for 24 hours to prevent *a priori* methylation of FLAG-SPIN1. FLAG-SPIN1 was immunoprecipitated 24 hours after

transfection and placed into a cell-free methylation reaction with recombinant GST-PRMT1. **(B)** Left upper panel: Autoradiograph showing semi cell-free methylation of FLAG-SPIN1 by GST-PRMT1. Middle and lower left panels: Western blot showing amount of FLAG-SPIN1 present in the immunoprecipitate and input, respectively. Right panel: quantification of auto-radiography signal generated from FLAG-SPIN1 derived from either AdOx-treated or untreated cells and normalized to the respective IP western blot. Error bars represent +/- SD. Students T-test *** represents *P*=<0.0001. n=3.

4.2.4 Antibody-based detection of SPIN1 methylation *in vitro*

The effects of AdOx treatment on SPIN1 methylation under semi cell-free conditions implied that SPIN1 is methylated *in vitro* **(Figure 4.10 B)**, which warranted further investigation. In addition to detecting methylation with the use of radioisotopes, antibody-based approaches presented another viable option. Their greater stringency with the type of methylation they detect (by delineating between arginine and lysine methylation; as well as mono or symmetric/asymmetric dimethylation of arginine) also affords basic triaging of the enzyme(s) responsible.

Applying this principle, HEK-293T cells were transfected with FLAG-SPIN1 that was subsequently immunoprecipitated, and probed for methylation with an anti-asymmetric dimethylarginine (ADMA) antibody **(Figure 4.11 A).** Although weak relative to the FXR1 positive control (confirmed as a PRMT1 substrate by Dr. Agnieszka Zielinska, Davies Laboratory, unpublished data), a signal corresponding to the size of SPIN1 could be detected with the anti-ADMA antibody supporting the notion that SPIN1 is methylated in cells. More dominant bands were detected at 55 and 72kDa, although whether these represented asymmetrically dimethylated SPIN1 binding partners or higher order oligomerisation was not clear **(Figure 4.11 B, see also section 4.9.1)**.

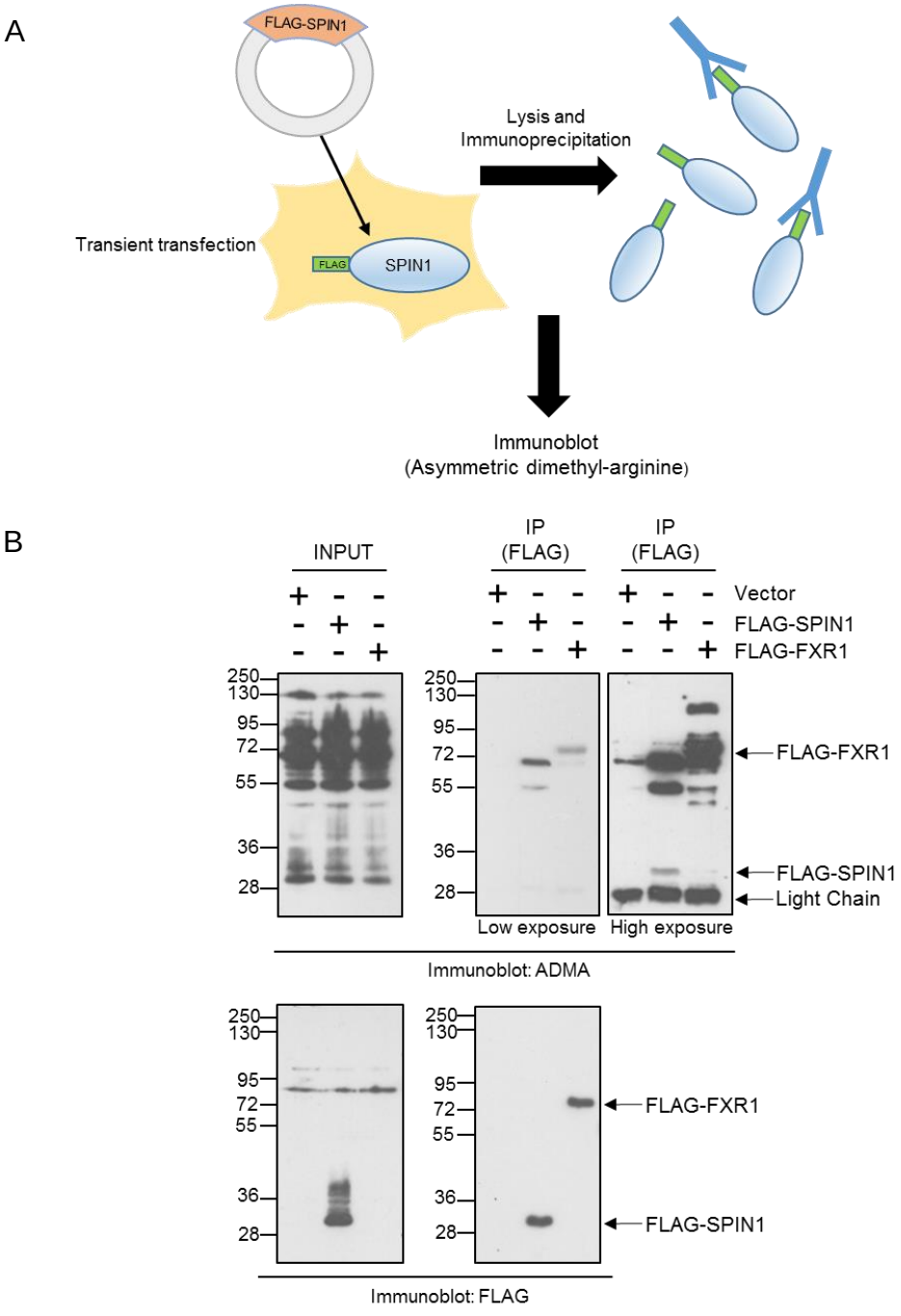


Figure 4.11: Antibody-based detection of SPIN1 methylation *in vitro*

(A) Schematic showing workflow of antibody-based detection of SPIN1 methylation *in vitro*. 24 hours after transient transfection of HEK-293T cells, ectopic FLAG-SPIN1 was immunoprecipitated and subjected to western blotting with an antibody raised against asymmetric dimethylarginine (ADMA). **(B)** Immunoblot showing asymmetric dimethylation of SPIN1 *in vitro*

in HEK-293T cells, derived from experimental plan outlined in **A.** Upper left panel: input showing overall state of asymmetric dimethylation in non-immunoprecipitated sample. Upper middle and upper right panels show species detected by the ADMA antibody after immunoprecipitation of FLAG-SPIN1 or FLAG-FXR1. Upper right panel is a higher exposure of upper middle panel, revealing an ADMA signal corresponding to the molecular weight of FLAG-SPIN1 (middle lane). Lower left and lower right panels show presence of FLAG-tagged species (SPIN1 and FXR1 positive control) in input and immunoprecipitate samples, respectively. n=2.

4.2.5 Optimising PRMT1 knockdown in the presence of SPIN1 overexpression

In order to demonstrate PRMT1-dependent methylation of SPIN1 *in vitro,* knockdown of PRMT1 is necessary. Normally, this would be accomplished through siRNA transfection. However, it was observed in numerous experiments that SPIN1 overexpression actually increased PRMT1 protein expression **(Figure 4.12 A).** This up-regulation of PRMT1 prevented sufficient knockdown of PRMT1 by siRNA, rendering *in vitro* methylation experiments challenging to conduct (data not shown).

Attempts to solve this issue using MCF7 shPRMT1 constitutive knockdown cells and transient transfection of SPIN1 proved unsuccessful due to poor SPIN1 transfection efficiency **(Figure 4.12 B, lane 2 compared to lanes 3 and 4/ lane 5 compared to lanes 6 and 7)**. This was important as preliminary autoradiographic detection of SPIN1 methylation (in MCF7 pTIPZ-FLAG-SPIN1 cells and transient transfection into HEK-293Ts) generated a weak signal over a relatively long period of time (data not shown), making a high transfection efficiency of SPIN1 was deemed critical. Indeed, despite robust stable knockdown of PRMT1 in MCF7 cells using two independent shRNA sequences (shPRMT1-1 and shPRMT1-2), even a modest transfection of FLAG-SPIN1 into these cells started to counteract the activity of the expressed shPRMT1 hairpins, increasing PRMT1 levels. Indeed, due to SPIN1's documented activity as a transcriptional coactivator of numerous genes through its histone code reader function[47], it was hypothesised that this activity was the cause of PRMT1 upregulation. Supporting this, qPCR data showed that induction of SPIN1 in

MCF7 pTIPZ-FLAG-SPIN1 cells also upregulated PRMT1 at the transcript level **(Figure 4.12 C)**. Taken together, this suggested that overexpressing FLAG-SPIN1 to usable levels may restore PRMT1 levels completely, rendering this methodology ineffective for investigating PRMT1-dependent methylation of SPIN1 *in vitro*.

Another interesting observation made during these experiments was the reduced expression of SPIN1 in both MCF7 shPRMT1 cell lines relative to the shC control **(Figure 4.12 B, lanes 2 and 5 compared to lane 1, respectively)**. This suggested that PRMT1 also positively regulates SPIN1 levels, implying a relationship of mutual positive regulation. Whether PRMT1 regulates SPIN1 expression at the transcriptional level or by promoting protein stability remains to be elucidated.

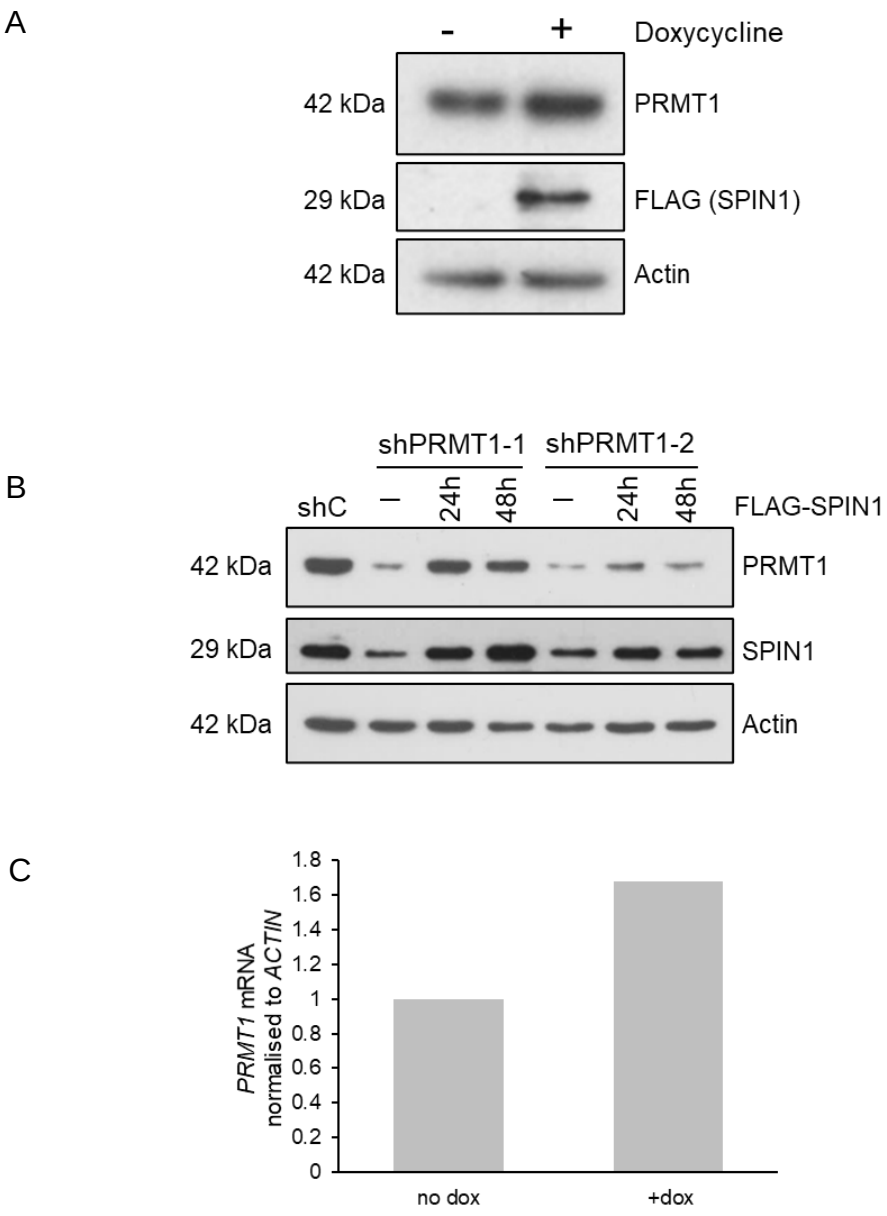


Figure 4.12: SPIN1 overexpression upregulates PRMT1 and suppresses PRMT1 knockdown

(A) Western blot showing PRMT1 up-regulation upon 48 hour induction of FLAG-SPIN1 overexpression in MCF7 pTIPZ-FLAG-SPIN1 cells using 1μg/ml doxycycline (n=2). **(B)** Western blot showing PRMT1 up-regulation upon transient transfection of FLAG-SPIN1 into two MCF7 cell lines stably expressing distinct shRNA against PRMT1 (n=1)**. (C)** qPCR data showing induction of PRMT1 after 48 hours of 1μg/ml doxycycline-induced FLAG-SPIN1 overexpression in MCF7 pTIPZ FLAG-SPIN1 cells (n=1).

Several publications have identified key residues within the aromatic cages of SPIN1 Tudor domains that are responsible for its contact with methyl-arginine (H3R8me2a) and methyl-lysine (H3K4me3) in the context of transcriptional activation, and a plethora of Tudor domain mutants have been characterised[47]. Of these candidates, two key residues F141 and D184 were selected due to their reported transcription coactivator incompetence. Both residues form part of Tudor domain II, with F141 interacting directly with trimethylated K4 on histone H3 tail peptide (as part of the aromatic cage), and D184 forming a stabilising salt bridge with R2 of the histone H3 tail **(Figure 4.13 A)** [47].

Both F141A and D184R mutants have been characterised in their suppression of SPIN1-mediated ribosomal RNA expression, with D184R proving the most potent D184 mutant due to a reversal in the amino acid charge[48,117]. However, since this mutant was to be used for detecting arginine methylation, providing additional arginine residues (albeit buried within the SPIN1 structure and thus unlikely to be reached by PRMTs) was deemed potentially confounding. As a result, arginine was substituted for the only other basic amino acid histidine, resulting in the D184H mutant.

In contrast to FLAG-SPIN1, transient overexpression of FLAG-SPIN1-F141A and FLAG-SPIN1-D184H in HEK-293T cells failed to upregulate transcription of *PRMT1* **(Figure 4.13 B).** As such, the F141A mutant was selected due to its greater efficacy over D184H for preventing *PRMT1* mRNA upregulation, and more widespread use in the literature. Through the generation of MCF7 cells

stably expressing doxycycline-inducible SPIN1-F141A (MCF7 pTIPZ-FLAG-SPIN1-F141A, **section 4.2.6),** studies investigating the role of PRMT1 for SPIN1 methylation could now be conducted.

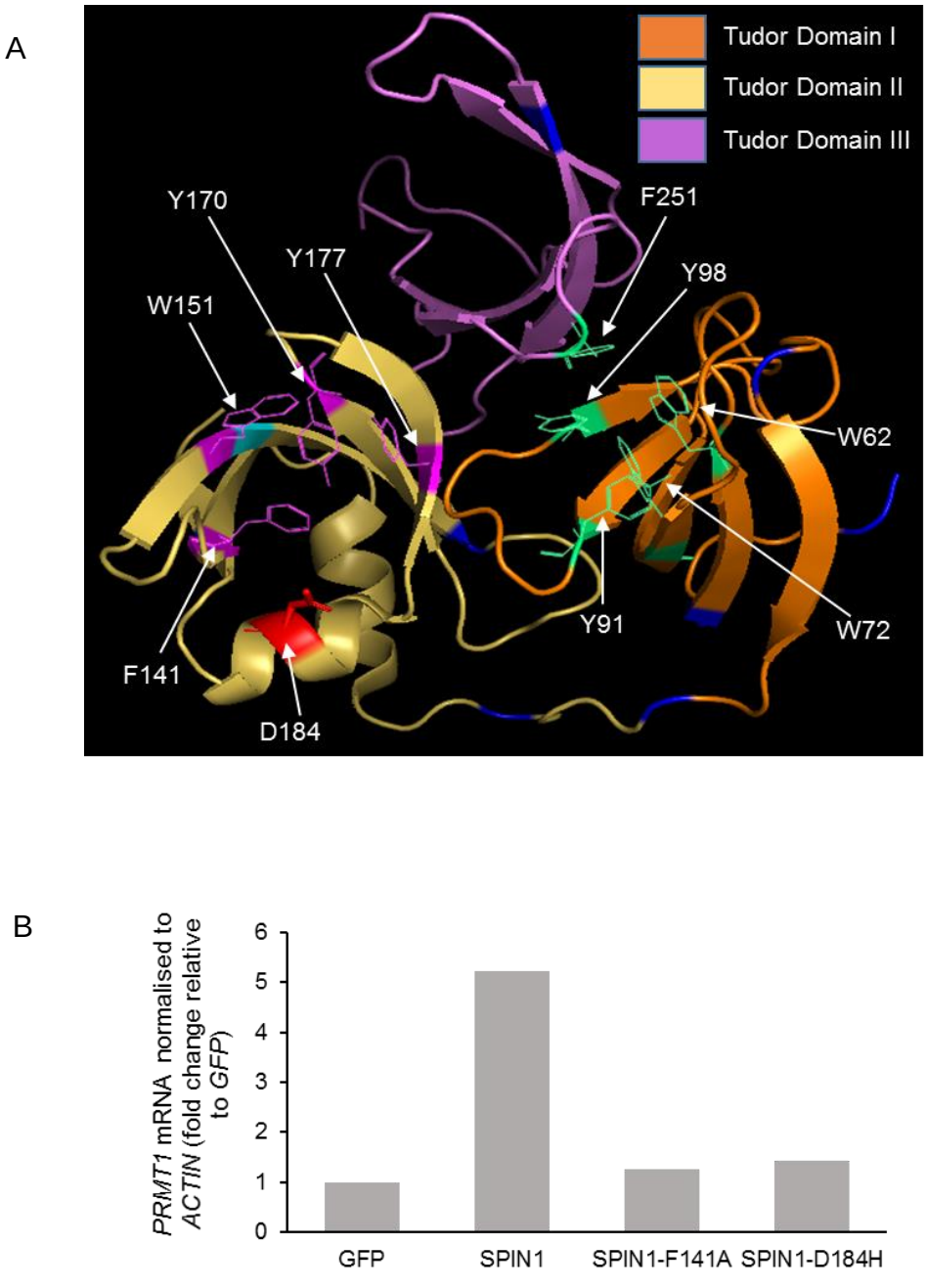


Figure 4.13: SPIN1 Tudor domain mutants suppress PRMT1 upregulation

(A) SPIN1 crystal structure PDB: 2NS2. Arginine residues are in dark blue (and cyan for R152). Aromatic cage residues (involved in histone code reader/transcriptional coactivator function) for Tudor domain I are shown in green, and in magenta for Tudor domain II. D184 is shown in red. **(B)** qPCR showing *PRMT1* transcript levels in HEK-293T cells transfected with either wild-type FLAG-SPIN or transcription incompetent mutants, 24h after transfection (n =1).

4.2.6: *In vitro* methylation of SPIN1 with PRMT1 knockdown

By using MCF7 pTIPZ-FLAG-SPIN1-F141A cells as a viable means of overexpressing SPIN1 and simultaneously knocking down endogenous PRMT1, it was possible to probe the contributions of PRMT1 to SPIN1 methylation *in vitro.* This was initially tested within the context of antibody-mediated detection of ADMA after immunoprecipitation of FLAG-SPIN1. Unexpectedly, knockdown of PRMT1 did not reduce asymmetric dimethylation of FLAG-SPIN1-F141A when detected with an anti–ADMA antibody **(Figure 4.14 A and B)**, suggesting that the methylation event(s) detected by the ADMA antibody were not PRMT1 dependent.

This result did not discourage further investigation since the ADMA antibody is specifically raised against methylated RG motifs, suggesting that it may be detecting methylation of R152, the only arginine within the SPIN1 structure belonging to an RG motif. Although PRMT1 appears to have a preference for methylated RG motifs, there are known exceptions[154] meaning that other arginine residues may be targeted by PRMT1 that are not detected by the ADMA antibody. To investigate this, we used *in vitro* methylation assays in which MCF7 pTIPZ-FLAG-SPIN1-F141A cells were supplemented with radioactive ^{3}[H]-methionine, in order to give a pan-methylation signature of SPIN1 through autoradiography. This experiment also included knockdown of PRMT1 by siRNA to interrogate dependency on PRMT1 for FLAG-SPIN1-F141A methylation. However, the same result was obtained, with no observable difference in FLAG-SPIN1-F141A methylation after PRMT1 depletion **(Figure 4.14 C and D)**. Inclusion of a transiently-transfected GFP control suggested that equal autoradiograph signals

were not a consequence of methionine incorporation into the polypeptide chain, and rather, a true comparison of post-translational methylation of FLAG-SPIN1-F141A. However, incomplete knockdown of PRMT1 for this experiment meant that PRMT1-dependent methylation cannot be completely ruled out, as a low expression threshold of PRMT1 protein may only be necessary to maintain SPIN1 methylation *in vitro.*

Due to the high efficiency with which FLAG-SPIN1 was methylated by GST-PRMT1 under semi cell-free conditions, the possibility that the F141A mutation somehow reduced PRMT1-dependent methylation *in vitro* by altering interactions with PRMT1 was investigated. Co-transfection of Myc-PRMT1-XGX and FLAG-SPIN1-F141A into HEK-293T cells demonstrated a modest reduction in PRMT1 interaction relative to wild-type FLAG-SPIN1. In contrast, the FLAG-SPIN1-D184H mutant, that was included as a potential replacement for the F141A mutant, showed no change in PRMT1 binding **(Figure 4.15 A)**. Next, cell fractionation was also carried out comparing wild-type FLAG-SPIN1 localisation to that of the F141A mutant. Western blotting and subsequent band quantification showed a shift from the chromatin-bound fraction to the nucleoplasm and cytoplasm in the F141A mutant relative to the wild-type FLAG-SPIN1, suggesting that this mutation may considerably alter the subcellular localisation of SPIN1 **(Figure 4.15 B)**. However, since the cytoplasmic marker tubulin is also present in the nuclear fraction and (to a lesser extent) the chromatin-bound fraction, this experiment remains inconclusive. A potential shift in subcellular localisation was not completely unexpected, as the F141 residue contributes to the Tudor domain II aromatic cage of SPIN1, which is vital in binding H3K4me3[47]. Thus, ability to

form this contact may have been one of the primary factors in maintaining close proximity with chromatin for FLAG-SPIN1.

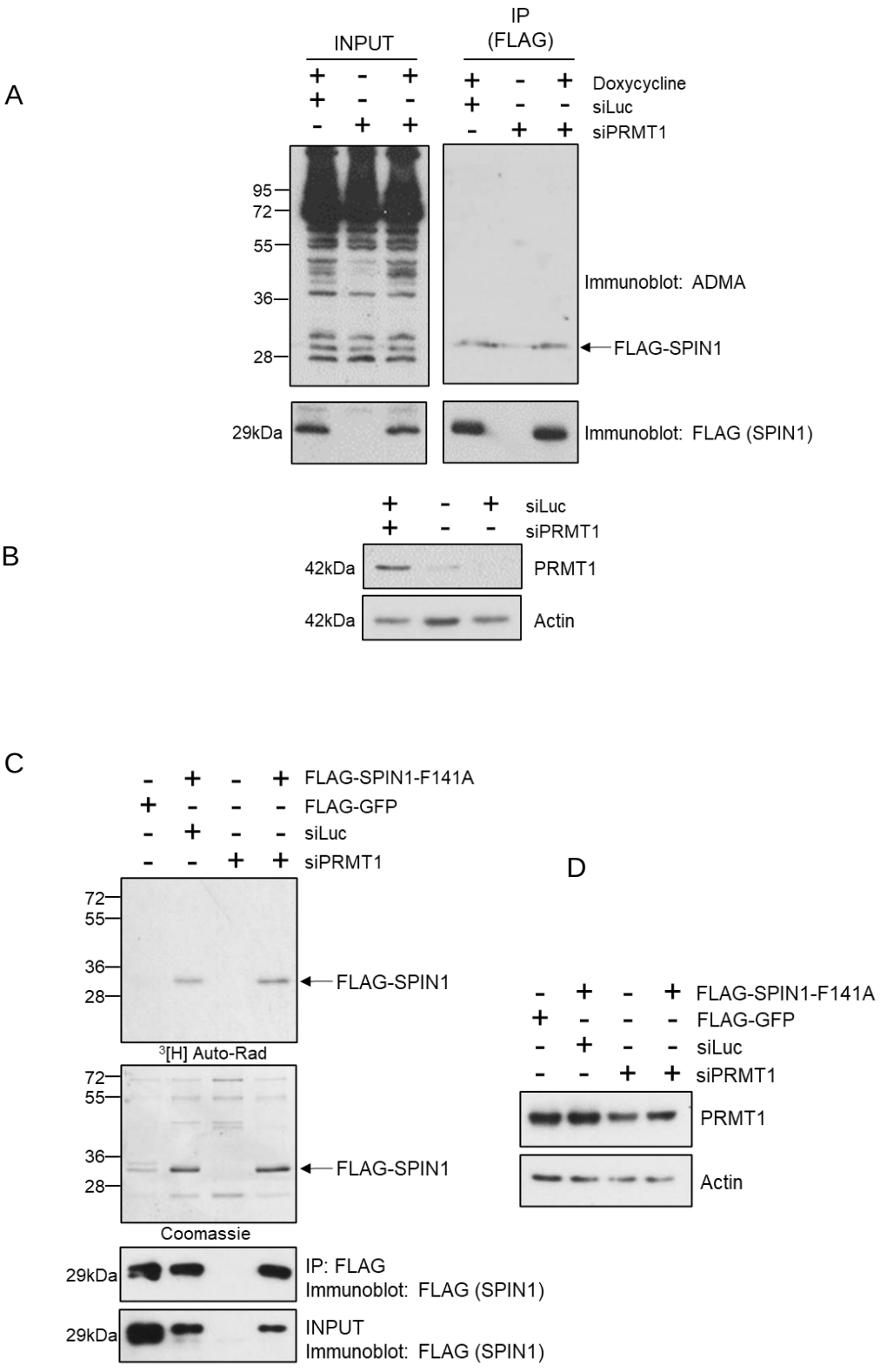


Figure 4.14: Effects of PRMT1 knockdown on SPIN1-F141A methylation status *in vitro*

(A) *In vitro* SPIN1 methylation assay using MCF7 pTIPZ-FLAG-SPIN1-F141A cells in the presence of 72h siRNA-mediated PRMT1 knockdown (using ADMA antibody to detect FLAG-

SPIN1-F141A methylation). Ectopic FLAG-SPIN1-F141A overexpression was induced for 48 hours prior to immunoprecipitation with 1μg/ml doxycycline. Upper left panel shows all asymmetrically dimethylated species in input (non-immunoprecipitated) sample. Upper right panel shows asymmetrically dimethylated species after immunoprecipitation of FLAG-SPIN1-F141A. Lower left and right panels show presence of FLAG-SPIN1-F141A in input and immunoprecipitated, respectively. **(B)** Western blot showing knockdown of PRMT1 in experiment depicted in **A.** n=2. **(C)** Equivalent experiment to **A** using autoradiography for detection of FLAG-SPIN1-F141A methylation. Upper panel: autoradiograph showing methylation of FLAG-SPIN1-F141A. Middle panel: Coomassie stain showing total protein content in gel that produced autoradiograph in upper panel. Lower panels: western blots showing presence of FLAG-tagged species (GFP and SPIN1-F141A) in immunoprecipitate and input samples, representative of middle and upper panels. **(D)** Western blot showing siRNA-mediated knockdown of PRMT1 after 72 hours, representative of data in **C.** n=3.

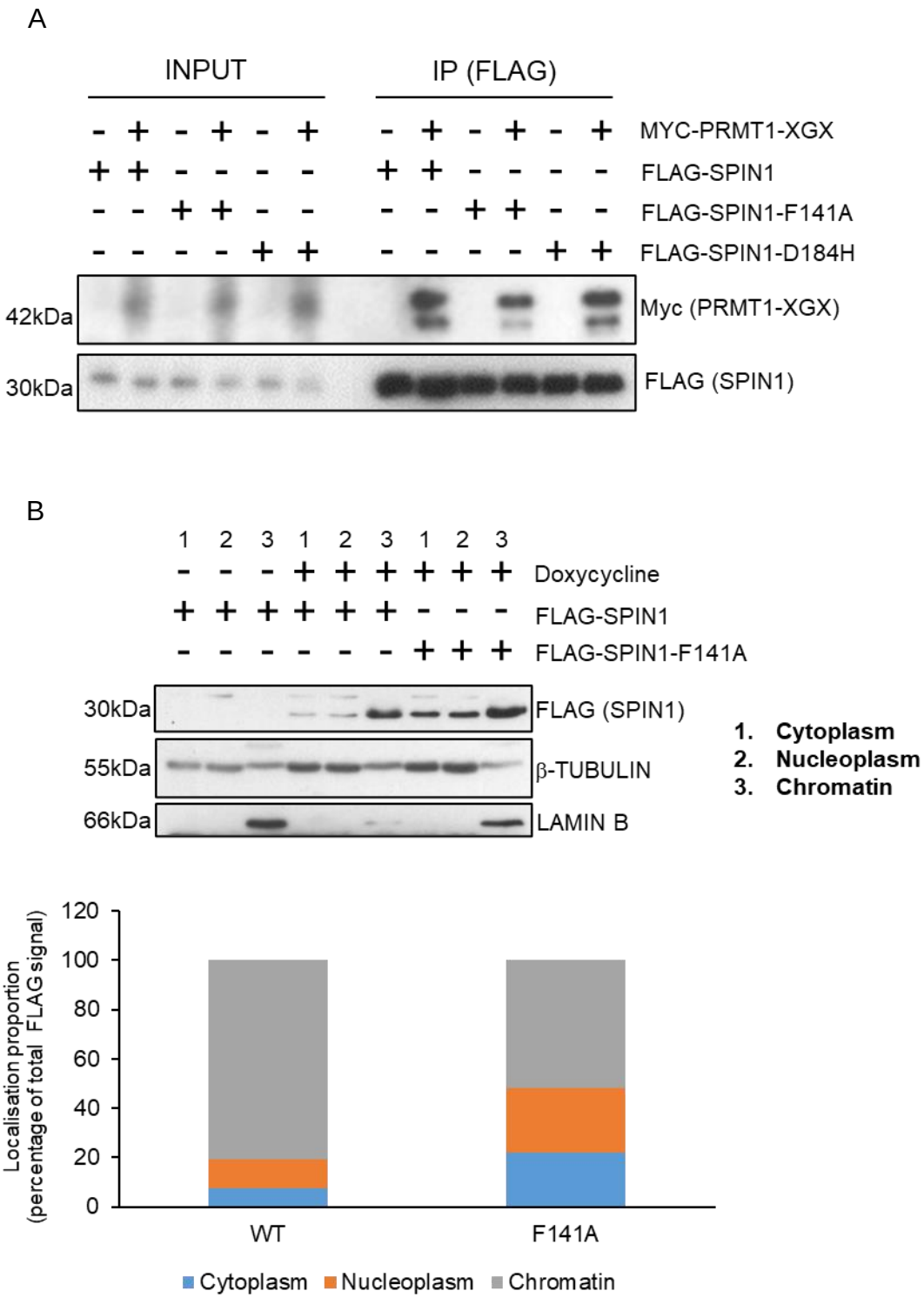


Figure 4.15: Effects of the F141A mutation on SPIN1 localisation and PRMT1 interaction

(A) Co-immunoprecipitation assay (24h transient transfection in HEK-293T cells) comparing *in vitro* interaction of wild-type FLAG-SPIN1, FLAG-SPIN1-F141A and FLAG-SPIN1-D184H with Myc-PRMT1-XGX, by immunoprecipitation of FLAG-tagged species. n=2. **(B)** Upper panel: cell

fractionation of MCF7 pTIPZ-FLAG-SPIN1 cells comparing subcellular localisation between wild-type FLAG-SPIN1 and the F141A mutant, 48 hours after 1μg/ml doxycycline induction. Lower panel: graphical representation of data presented in the upper panel using ImageJ band quantification. Localisation is presented as a percentage of the total signal from all 3 lanes (cytoplasm, nucleoplasm and chromatin) for each treatment. n=1.

4.2.7: Inhibition of PRMT1-mediated SPIN1 methylation *in vitro*

Given that FLAG-SPIN1-F141A exhibits altered subcellular localisation and reduced PRMT1 interaction relative to wild-type FLAG-SPIN1 **(Figure 4.15 A** and **B),** the F141A mutant was not considered biologically representative of SPIN1 methylation status *in vitro.* As such, reversion to studying wild-type SPIN1 in relation to PRMT1 activity was necessary for *in vitro* experimentation. Due to the aforementioned inability to knock down PRMT1 in the presence of SPIN1 overexpression, pan-type I PRMT inhibitor MS023 that has a greater potency for PRMT1 than other Type 1 enzymes[155], was employed to test PRMT1-dependent methylation. HEK-293T cells were transiently transfected with FLAG-SPIN1, protein synthesis inhibited, and incubated with ^{3}[H]-methionine to detect methylation of FLAG-SPIN1 via autoradiography. Treatment of cells with MS023 abrogated the methylation signal in comparison to the DMSO control, which generated a detectable signal corresponding to the size of FLAG-SPIN1 **(Figure 4.16)**. This experiment therefore suggested that in proper cellular context, SPIN1 may be methylated by PRMT1 *in vitro*.

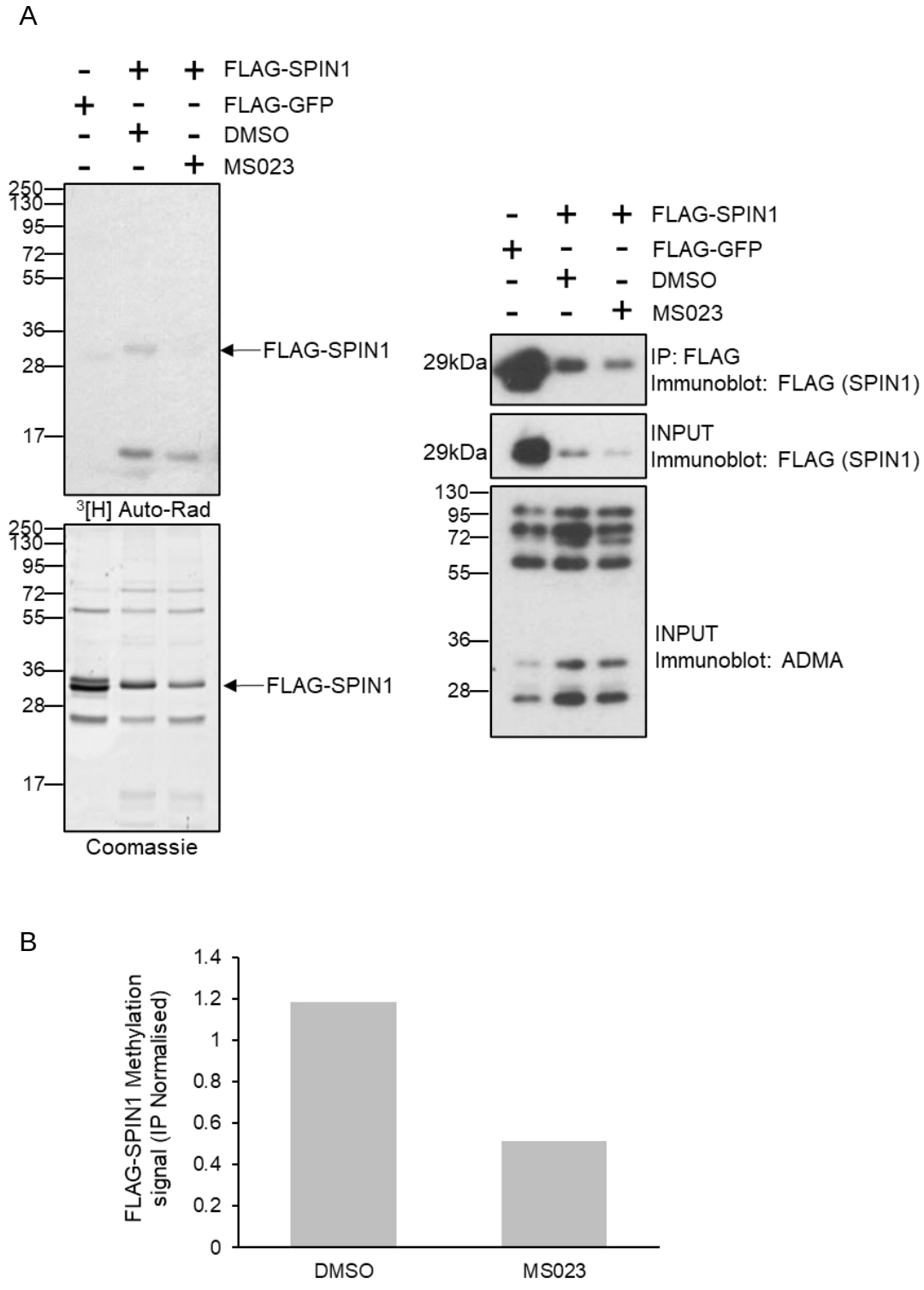


Figure 4.16. Type I PRMT inhibition reduces SPIN1 methylation *in vitro*

(A) *In vitro* methylation status of FLAG-SPIN1 after 48 hours of type I PRMT inhibitor MS023 treatment (48h transient transfection in HEK-293T cells). Upper left panel shows autoradiographic readout of methylation assay. Lower left panel: Coomassie stain showing all protein species

present in gel that produced autoradiograph in upper left panel. Upper and middle right-hand panels show expression of FLAG-tagged species (FLAG-SPIN1 and FLAG-GFP) in immunoprecipitate and input samples, respectively. Lower right-hand panel shows all asymmetrically dimethylated species detected in input samples. **(B)** Depicts band quantification of autoradiography signal (upper left-hand panel in **A**), normalised to IP western blot in upper-right panel of **A** (n=1).

4.3: Identification of SPIN1 arginine methylation sites

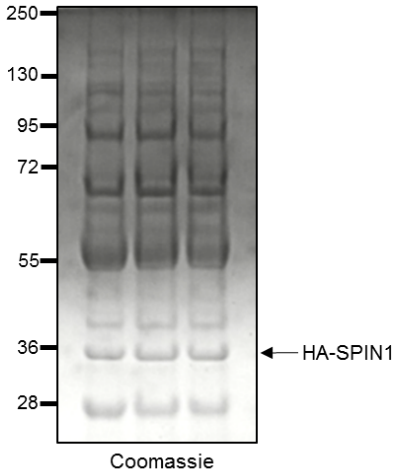
4.3.1: Identification of novel SPIN1 arginine methylation sites by mass spectrometry

Once SPIN1 was confirmed as a semi cell-free and *in vitro* PRMT1 substrate, attention was turned to identification of the specific target residue(s). In order to identify arginine methylation sites without systematically mutating all arginine residues within the SPIN1 structure, a mass spectrometry approach was applied. After transient overexpression in HEK-293T cells, HA-SPIN1 was enriched by immunoprecipitation and resolved by SDS-PAGE **(Figure 4.17 A)**. Samples were then sent for analysis by mass spectrometry after tryptic digestion. Among several phosphorylation events (some of which are novel), monomethylation at R14 and R117 was detected **(Figure 4.17 B)**. Although most arginine residues were covered across the three independent repeats, (albeit to varying extents), R158 was not covered in 75% repeats, and R30 was not represented at all.

This experiment was repeated with MCF7 and MCF10A pTIPZ-FLAG-SPIN1 cells, using FLAG immunoprecipitated SPIN1. Here, monomethylation at arginine 117, but not R14, was detected in MCF10A but not MCF7 cells. Lack of R14 methylation in MCF10A pTIPZ-FLAG-SPIN1 cells could be explained by their generation of fewer FLAG-SPIN1 spectra compared to HEK-293T cells (data not shown) leading to a reduced probability of detecting methylation events that occur at a low stoichiometry. This trend continues to a greater extreme in MCF7 pTIPZ-FLAG-SPIN1 cells, which produced even fewer spectra than their MCF10A counterparts, perhaps explaining their failure to reveal methylation of either

residue. However, reproduction of the R117 methylation event across HEK-293T and MCF10A cells reinforced the likelihood that it was a biologically relevant event. Similarly, although R14 methylation was only found in HEK-293T cells, peptides constituting the N-terminal tail that contain this residue were represented far more scarcely than most other segments of the polypeptide. Thus, methylation of R14 may be more abundant when considered in ratio with the relevant peptides rather than an absolute number of events. As such, both R14 and R117 methylation were selected for further investigation.

A



B

PTM	Documented?
S13 Phosphorylation	No
R14 Methylation (mono)	No
S38 Phosphorylation	Yes
S39 Phosphorylation	Yes
Y91 Phosphorylation	Yes
Y98 Phosphorylation	Yes
R117 Methylation (mono)	No
S124 Phosphorylation	Yes
T143 Phosphorylation	No
S147 Phosphorylation	No
T164 Phosphorylation	No
Y170 Phosphorylation	No
Y177 Phosphorylation	No
S196 Phosphorylation	No
S199 Phosphorylation	Yes

Figure 4.17: Identification of novel SPIN1 arginine methylation sites by mass spectrometry

(A) Coomassie gel showing 3 identical repeats of HA-SPIN1 immunoprecipitation after 24h transient overexpression in HEK-293T cells. **(B)** List of all post translational modifications pooled together from HEK-293T and MCF7/MCF10A pTIPZ-FLAG-SPIN1 immunoprecipitate samples where FLAG/HA-SPIN1 was transiently overexpressed (n=6 for each cell line), cross-referenced with known PTMs of SPIN1 (www.phosphosite.org).

4.3.2: Semi cell-free validation of prospective SPIN1 methylation sites

Since immunoprecipitated FLAG-SPIN1 can be methylated by GST-PRMT1 in a cell-free reaction **(Figure 4.10)** we chose to use this assay to assess a panel of FLAG-SPIN1 mutants whereby methyl-acceptor arginine residues were mutated to lysine. HEK-293T cells were transiently transfected with these FLAG-SPIN1 variants followed by FLAG immunoprecipitation and cell-free methylation with GST-PRMT1 and ^{3}[H]-SAM. Despite mass spectrometry analysis failing to identify methylation at R152, its status as the sole contributor of an RG motif combined with its placement on Tudor domain II (the best-characterised functional region of SPIN1), warranted its inclusion as a potential arginine methylation site. Further, previous detection of FLAG-SPIN1 methylation *in vitro* using an antibody raised against asymmetrically dimethylated RG motifs suggested that R152 may be methylated *in vitro* **(Figure 4.11)**. Therefore, we generated FLAG-SPIN1-R14K, FLAG-SPIN1-R117K and FLAG-SPIN1-R152K expression constructs.

Transient overexpression of FLAG-SPIN1 mutants in HEK-293T cells and subsequent semi cell-free methylation generated multiple bands via autoradiography **(Figure 4.18 A)**. Bands corresponding to FLAG-SPIN1 methylation were detected at the correct molecular weight and super-imposed perfectly with coomassie bands corresponding to FLAG-SPIN1 of the same gel. Another set of methylated proteins running slightly higher (36kDa) produced a slightly stronger signal, yet were more faintly visible in the coomassie stain. These may represent a smaller pool of FLAG-SPIN1 that has been post-translationally

modified (by phosphorylation for example) that are also more heavily methylated. However, due to the speculative nature of this assertion, these bands were not regarded as a means to judge FLAG-SPIN1 methylation. Finally, the most dominant were a set methylated proteins ranging between ~80kDa and 100kDa. Considering that these proteins were barely visible via coomassie, this suggested that they were heavily methylated.

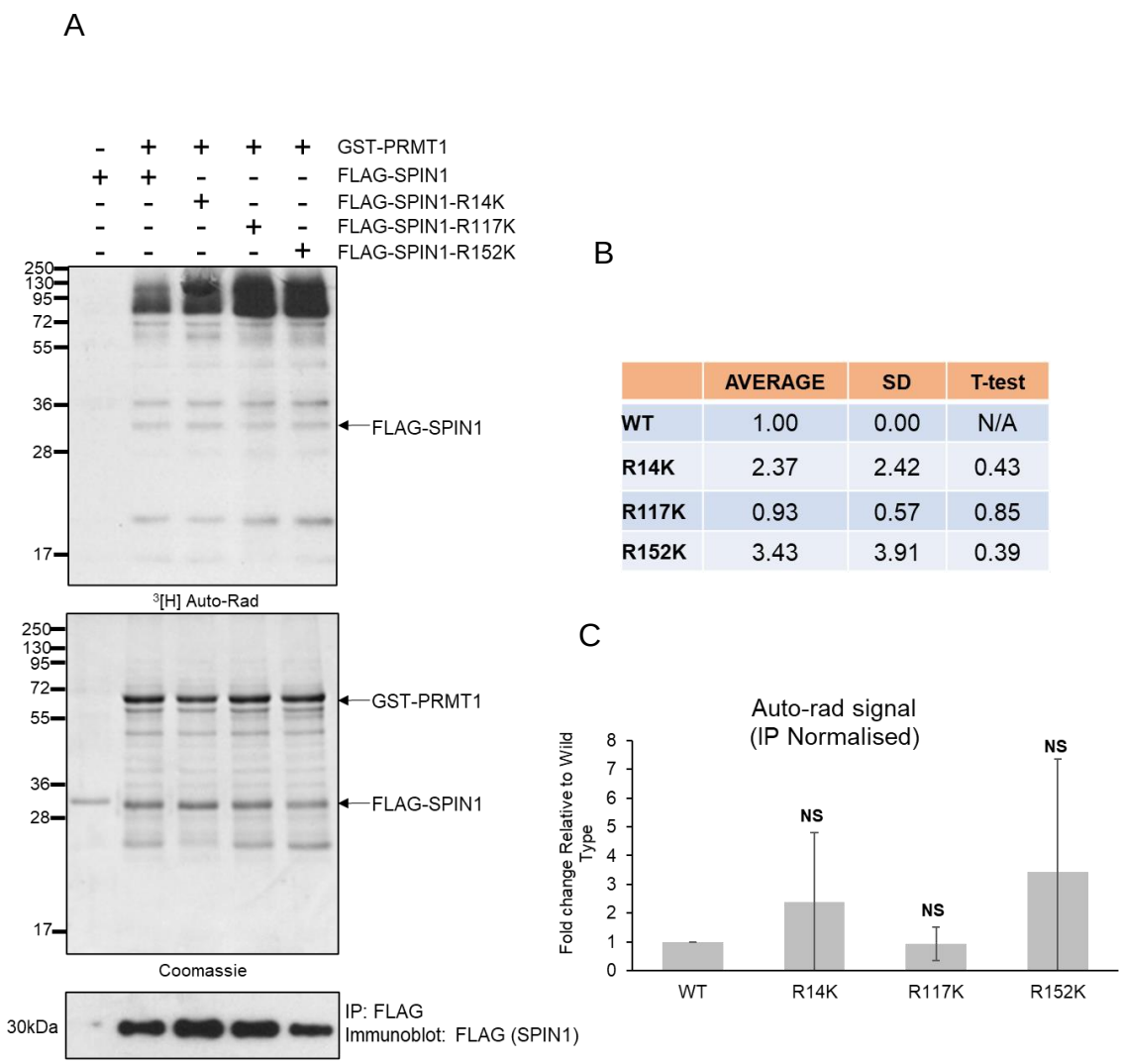
The methylation signals generated by FLAG-SPIN1 and its mutants did not demonstrate an obvious difference in autoradiograph signal by visual interpretation. To acquire a quantitative understanding of the data, the autoradiograph signal was normalised to the respective IP western blot using ImageJ software and represented as a fold-change in methylation relative to wild-type FLAG-SPIN1. This methodology showed minor reduction in methylation in the R117K mutant, in contrast to the R14K and R152K mutants, which generated generally much higher methylation signals than the wild type control. However, due to high variability in quantitative analysis of all mutants tested, none of the aforementioned phenomena were statistically significant **(Figure 4.18 B and C).**

Although these experiments failed to validate any of the candidate arginine methylation sites, one of the three experimental repeats for the R117K mutant produced a high outlying value, altering the methylation value such that it did not differ from wild-type FLAG-SPIN1 on average. Without this data point, (using only n=2) the R117K mutation abrogated 40% of the methylation signal (data not shown).

Because of inconsistencies in data generated from transient transfection into HEK-293T cells, the same experiments were carried out using MCF7 pTIPZ-FLAG-SPIN1 cells **(Figure 4.19 A)**. For this analysis, the autoradiograph signal was again normalised to the amount of total FLAG-SPIN1 immunoprecipitated in each sample. With this approach, a statistically significant reduction (33%) in FLAG-SPIN1-R117K methylation was observed relative to the wild type control **(Figure 4.19 B and C)**. FLAG-SPIN1-R14K demonstrated an even higher average signal (a ~3 fold compared to ~2.4 fold increase) than in **Figure 4.18,** however this difference was also not statistically significant, again as a result of variability within the data. Contrasting the findings in **Figure 4.18**, the R152K mutant showed a 25% reduction in methylation on average compared to the wild type, although this was also statistically non-significant **(Figure 4.19 B and C)**.

Interestingly**,** both experiments repeatedly showed that inclusion of GST-PRMT1 appeared to stabilise Flag-SPIN1 expression, (**Figures 4.18 and 4.19 A, compare lane 1 with lanes 2-5**). However, whether this stabilisation was incumbent on FLAG-SPIN1 methylation, or simply interaction with GST-PRMT1 was indeterminable. For the stable MCF7 cell lines, expression of the R117K and R152K mutants was far greater than that of wild-type FLAG-SPIN1 and the R14K mutant (which were approximately equal, **Figure 4.19 A**). One possible reason for this difference is that cell lines were generated with un-titred lentivirus and on different occasions. Thus, the amount of viral particles produced, and hence infection rate, may be variable. As such, the apparent increased expression of

FLAG-SPIN1-R117K and R152K could simply be due to increased viral particle integration.

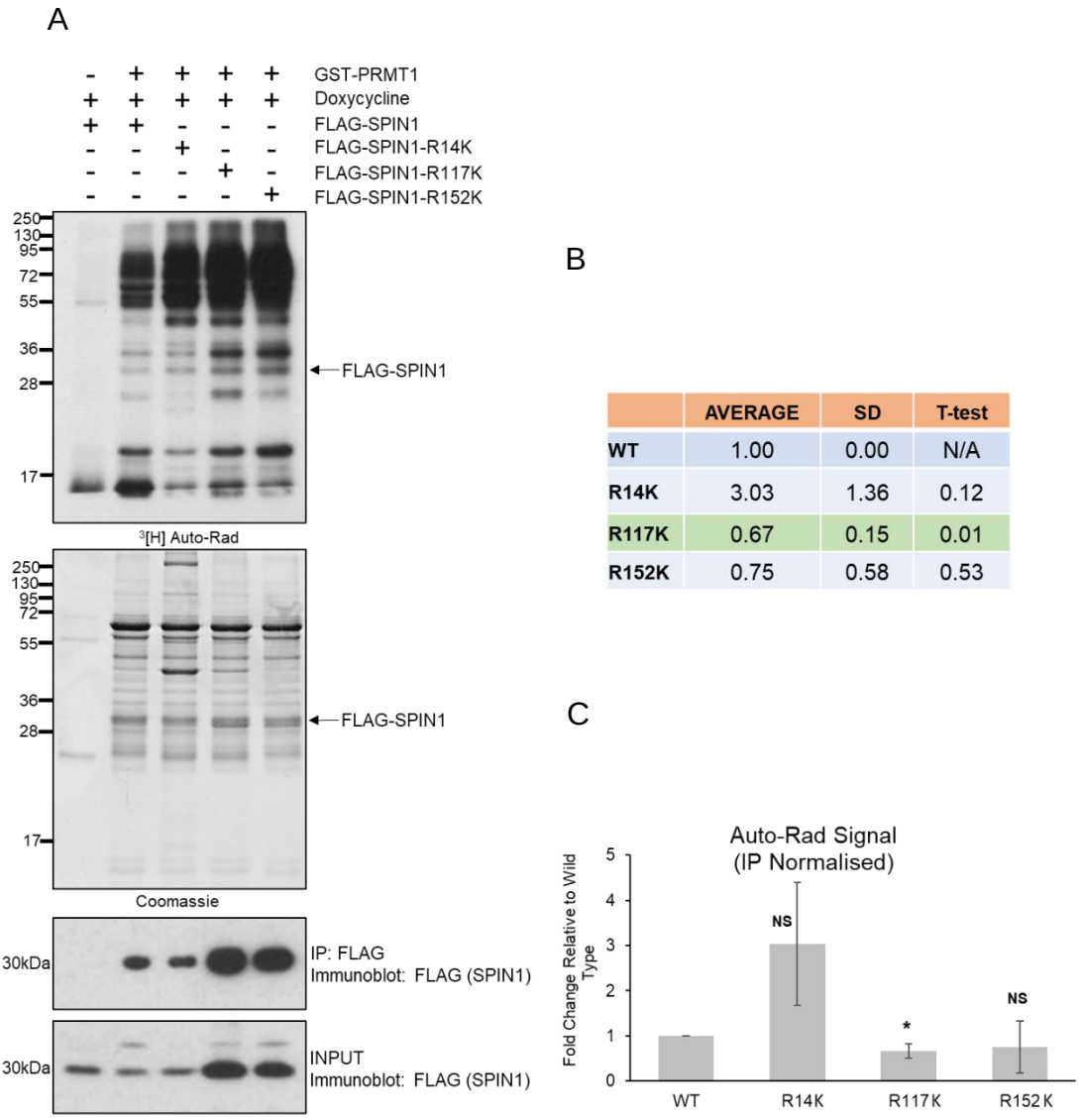


	AVERAGE	SD	T-test
WT	1.00	0.00	N/A
R14K	2.37	2.42	0.43
R117K	0.93	0.57	0.85
R152K	3.43	3.91	0.39

Figure 4.18: Semi cell-free methylation of HEK-293T-derived SPIN1 arginine mutants

(A) Semi cell-free methylation assay of FLAG-SPIN1 mutants extracted from HEK-293T cells after 24h transient overexpression. Upper panel: autoradiograph showing methylation state of FLAG-SPIN1 and mutants for candidate PRMT1-targeted residues (marked by arrow). Middle panel: Coomassie stain showing protein content of the autoradiograph gel shown in upper panel. Lower panel: western blot showing FLAG-SPIN1 content in immunoprecipitates. **(B)** Pixel counting data derived from autoradiograph band quantification, normalised to respective IP western blots. **(C)**

Bar graph representation of data depicted in **(B)**. Statistical analysis carried out using Student's one-tailed T-test. NS (not significant) n=3.



	AVERAGE	SD	T-test
WT	1.00	0.00	N/A
R14K	3.03	1.36	0.12
R117K	0.67	0.15	0.01
R152K	0.75	0.58	0.53

Figure 4.19: Semi cell-free methylation of MCF7 pTIPZ-FLAG-SPIN1-derived SPIN1 arginine mutants

(A) Semi cell-free methylation assay of FLAG-SPIN1 and mutants for prospective methylation targeted residues extracted from MCF7 pTIPZ FLAG-SPIN1 cells after 48h induction with 1μg/ml doxycycline. Upper panel: autoradiograph showing methylation state of different FLAG-SPIN1 variants (marked by arrow). Middle panel: Coomassie stain showing total protein content of gel used for autoradiograph in upper panel. Lower two panels: western blots showing presence of

FLAG-SPIN1 species in immunoprecipitate and input samples, respectively. **(B)** Data derived from autoradiograph band quantification using ImageJ pixel counting, normalised to pixel counting of IP western blots. Data for the R117K mutant, which showed a statistically significant reduction in methylation, are highlighted in green. **(C)** Bar graph representation of data depicted in **B**. Statistical analysis carried out using Student's one-tailed T-test. NS (not significant) * (*P*=<0.05) n=3.

4.3.3: Optimising cell-free SPIN1 methylation

Upon acquiring semi cell-free methylation data suggesting that R117 may be a PRMT1-methylated site, the SPIN1 crystal structure was used as a means to divulge potential functions of this modification. R117 is located within the phosphate–binding loop (P loop) that connects Tudor domains I and II **(Figure 4.20).** Residues within this loop chelate two phosphate ions, and R117 contributes to the affinity of a phosphate ion designated C301, that is preserved in the crystal structure, by provision of a hydrogen bond[111]. Due to high steric proximity between phosphate C301 and R117, it was hypothesised that methylation of this residue may be a mutually exclusive event with chelation of phosphate C301. This is reasoned because of spatial competition, but also as a result of methylation reducing the hydrogen bond-donating capacity of arginine by one with the addition of each methyl group[156].

In order to test the influence of phosphate presence on SPIN1 methylation, cell-free methylation reactions were set up using a range of buffers set to neutral pH (7.4): PBS (which had been used in previously unsuccessful cell-free His-SPIN1 methylation reactions, see section **4.2.2**), 20mM Tris-HCl and 1M HEPES. In addition, aluminium hydroxide (AlOH) was added to a duplicate reaction for each buffer type. This chemical is a strong phosphate binding agent that is versatile over a wide pH range, and has been used historically in patients with chronic kidney disease, for whom hyperphosphataemia is a risk[157]. Thus, phosphate chelation could be abrogated with the use of phosphate-free buffers (Tris-HCl, HEPES) or by direct sequestration from His-SPIN1 by aluminium hydroxide.

Addition of aluminium hydroxide did not appear to alter the methylation state of His-SPIN1, however PRMT1-catalysed methylation of SPIN1 was greatly enhanced by the use of Tris-HCl buffer compared to PBS and HEPES buffer **(Figure 4.21 A)**. Indeed, the potential for Tris-HCl as an efficient buffer for His-SPIN1 methylation was robustly reproducible **(Figure 4.21 B)**. Taken together, the data imply that cell-free methylation of His-SPIN1 by GST-PRMT1 that was not detected previously **(Figure 4.8)** could be due to methylation reactions containing PBS as the buffering agent. It is therefore likely that phosphate chelation adjacent to R117 competes with methylation of this residue, providing additional circumstantial evidence that R117 is a PRMT1-dependent methylation site.

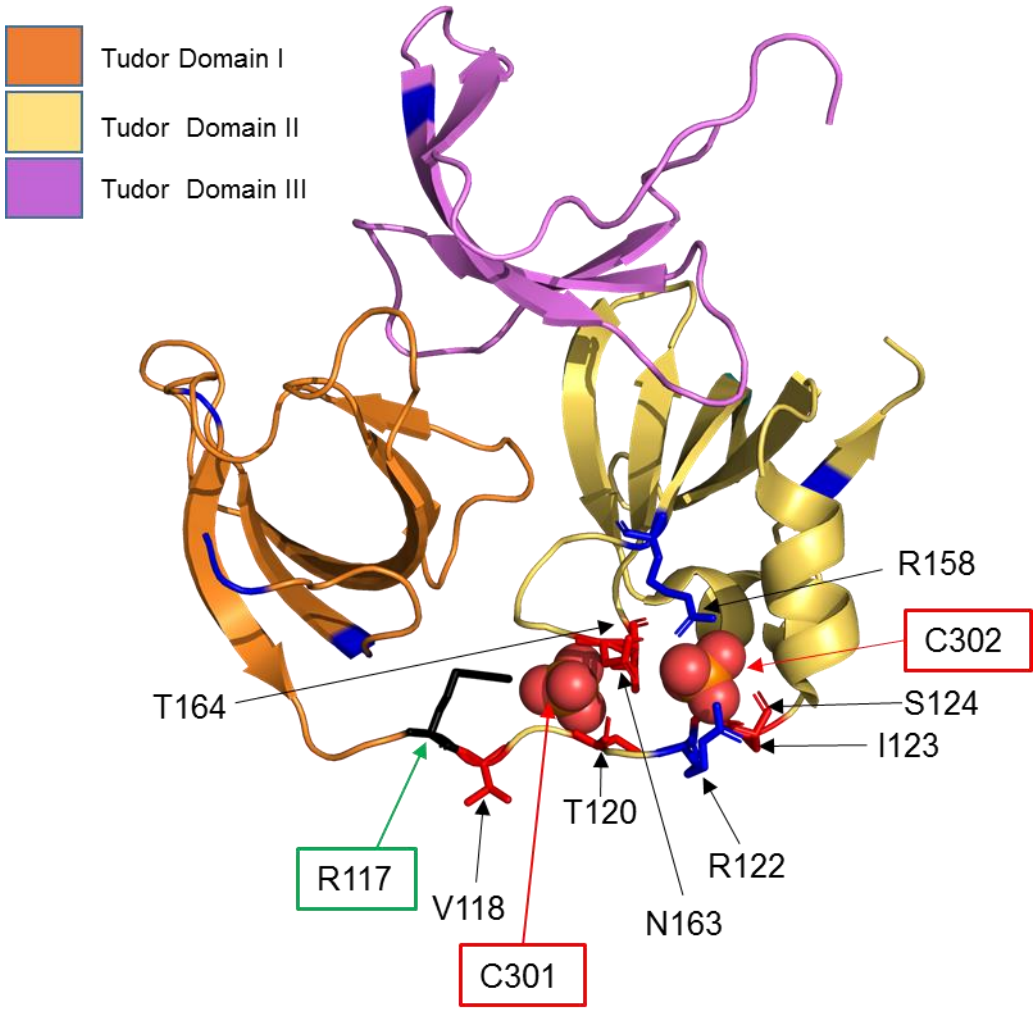


Figure 4.20: SPIN1 crystal structure suggests steric competition between R117 methylation and phosphate chelation

SPIN1 crystal structure showing residues in the phosphate-binding loop (represented as sticks) that chelate C301 and C302 phosphates according to Zhao et al., 2007[111]. Arginine residues are presented in blue, except for R117, which is presented in black. All other non-arginine residues involved in this process are presented in red. Note that residues are labelled according to the full length human SPIN1 amino acid sequence, not the nomenclature derived from the 2NS2 crystal structure[111].

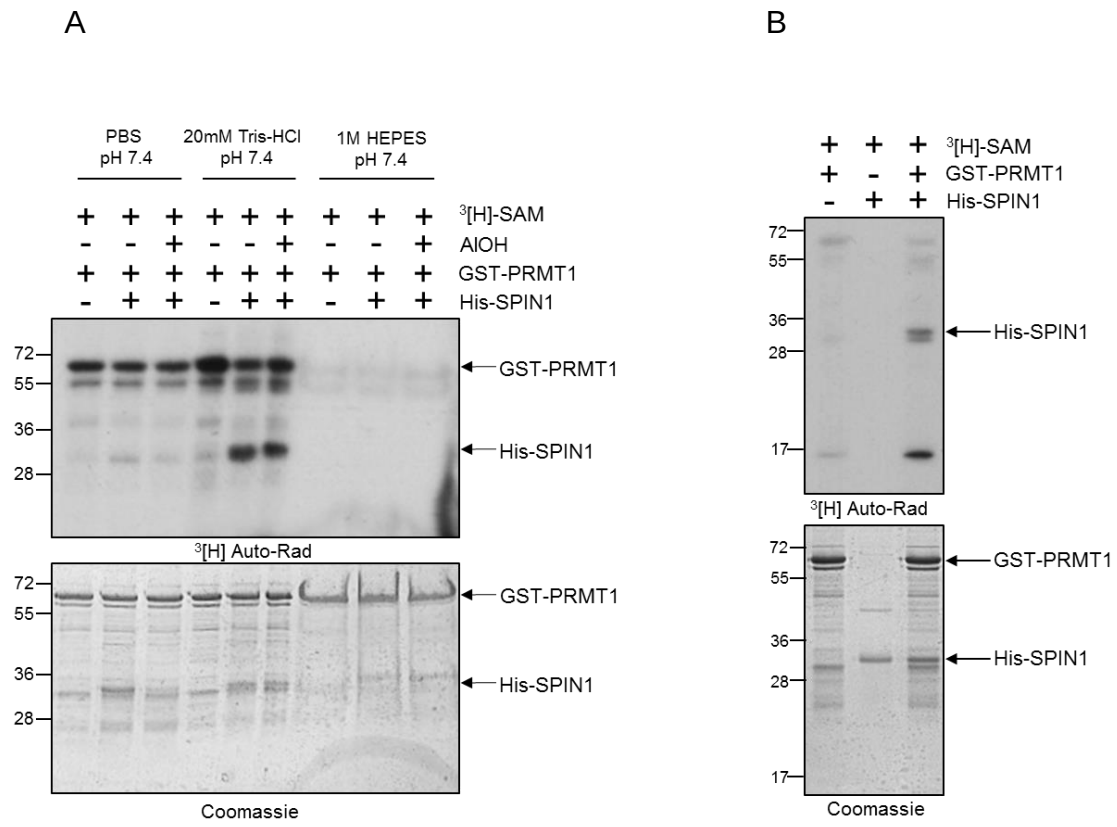


Figure 4.21: SPIN1 is methylated by PRMT1 under cell-free conditions when incubated in a phosphate-free buffer

(A) Cell-free assay testing the effect of buffer and aluminum hydroxide on GST-PRMT1-dependent methylation of His-SPIN1. Upper panel: autoradiograph showing methylation state of His-SPIN1 under various reaction conditions. Lower panel: Coomassie stain showing total protein content of gel that produced autoradiograph in upper panel. n=2**. (B)** Upper panel: Autoradiograph readout of cell-free methylation assay using optimized conditions with Tris-HCl.

Lower panel: Coomassie stain showing total protein content of gel that produced autoradiograph in upper panel. n=3.

4.4: Functional characterisation of SPIN1 arginine methylation

Having demonstrated that PRMT1 methylates SPIN1 in cell-free and semi cell-free contexts, and possibly *in vitro,* the functional consequences of this event were investigated. To date, SPIN1 is primarily implicated in gene regulation through its transcriptional coactivator function, and influencing the cell cycle in both meiosis and mitosis, although roles in the latter processes remain poorly characterised. Using these cellular processes as a starting point, the role of SPIN1-PRMT1 interaction in breast cancer was investigated.

4.4.1: SPIN1-PRMT1 regulation of transcription in breast cancer

SPIN1-mediated transcription activation is documented in genes that contribute to drug resistance such as drug metabolising enzymes CYP2C8, UGT2B4 and UGT2B17 in MCF7 breast cancer cells. However, this effect is observed most optimally in subsets of these cells that have been pre-selected to acquire adriamycin resistance[132]. To investigate potential cooperation between PRMT1 and SPIN1 in promoting oncogenic transcription in MCF7 cells that have not been cultured to acquire drug resistance, SPIN1-dependent gene sets discovered in liposarcoma[127], colorectal cancer[47] and triple-negative breast cancer[137] were brought into focus.

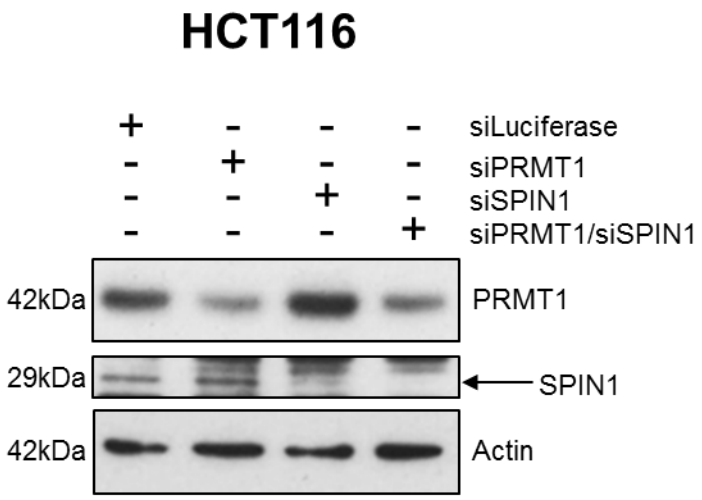
As a result of cell-type specific gene expression, not all genes identified as SPIN1-dependent could be analysed in MCF7 cells. For example, SPIN1 drives pro-proliferative RET signalling in liposarcoma through up-regulation of RET

activator GDNF[127], however amplification of GDNF could not be detected in MCF7 cells, even when SPIN1 was transiently overexpressed (data not shown).

Another well-characterised oncogenic signalling pathway involving SPIN1 is Wnt/β-catenin signalling. Here, SPIN1 promotes proliferation and invasiveness[125], and transcriptionally upregulates expression of downstream effectors such as *CYCLIN D1, AXIN2, ID-2* and *TIAM1* in HCT116 colorectal cancer cells[47].

First, we attempted to verify the findings of Su et al[47]., by depleting SPIN1 in the colorectal cancer cell line, HCT116. To determine if PRMT1 also regulates the same set of genes, and if this occurred in an epistatic or cooperative manner, cells were also depleted for PRMT1, and combination PRMT1/SPIN1, respectively **(Figure 4.22 A)**. Surprisingly, depletion of SPIN1, PRMT1, or combination depletion did not show any significant decrease in *CYCLIN D1* or *TIAM1* transcripts (ANOVA *P*= >0.05). In contrast, depletion of either SPIN1 or PRMT1 decreased *ID-2* expression, whilst combination knockdown of PRMT1 and SPIN1 lead to an additive reduction in *ID-2* transcripts (*P* =<0.05), suggesting a cooperative mechanism of gene expression **(Figure 4.22 B)**. *AXIN2* levels were only significantly reduced by PRMT1 depletion **(Figure 4.22 B)**.

A



B

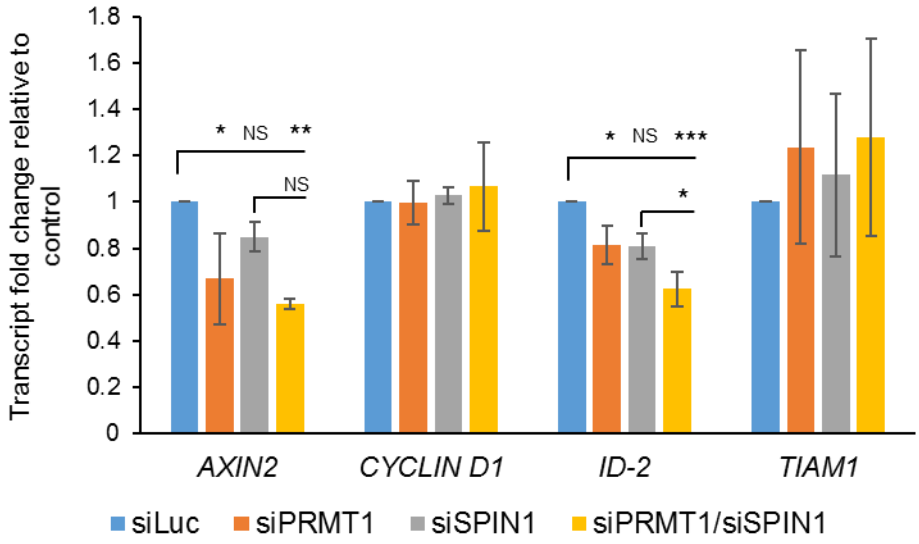


Figure 4.22: Influence of SPIN1 and PRMT1 on Wnt target gene transcription in HCT116 colorectal cells

(A) Western blot showing expression of endogenous SPIN1 and PRMT1 after 72h siRNA-mediated knockdown. **(B)** qPCR of Wnt/β-catenin target genes in the presence of siRNA directed against SPIN1 and PRMT1. Error bars represent +/- SD. Statistics are one-way ANOVA with Tukey's HSD post-hoc analysis. * signifies *P*= <0.05, ** *P*= <0.01, NS = not significant. n=3.

Despite our inability to fully validate the published SPIN1-dependent transcription profile in HCT116 cells, attention was turned to the MCF7 model to investigate whether *AXIN2, ID-2, TIAM1* and *CYCLIN D1* could be detected at the transcript level in this cell type. Again, changes in expression were interrogated after siRNA-mediated knockdown of either PRMT1, SPIN1 or both genes simultaneously **(Figure 4.23 A)**. Although statistical analysis was not possible for this experiment due to variability in knockdown efficiency of PRMT1 and SPIN1 (data not shown), a minor reduction in *ID-2* transcripts was attained after SPIN1 knockdown. However, combination knockdown of PRMT1 and SPIN1 did not show any additive effect over knockdown of SPIN1 alone **(Figure 4.23 B).** No reduction in transcript level was detected for *CYCLIN D1, AXIN2* or *TIAM1* after SPIN1 knockdown, nor was any change observed with PRMT1 knockdown, or combined PRMT1/SPIN1 depletion **(Figure 4.23 B)**. Interestingly, knockdown of SPIN1 appeared to have a greater effect on reducing PRMT1 protein levels than knockdown of PRMT1 itself **(Figure 4.23 A)**. However this only occurred in one of the two experimental repeats. Knockdown of SPIN1 has been observed to reduce PRMT1 protein levels in a small number of previous experiments (data not shown); however, inconsistency in these findings suggests that this phenomenon depends on some unknown condition. Further, *PRMT1* appeared to increase slightly at the transcript level when SPIN1 was singularly knocked down **(Figure 4.23 B)**, suggesting that reduction of SPIN1 somehow decreases PRMT1 protein stability.

Next, it was decided that expression of these genes would be analysed using our doxycycline-inducible MCF7 pTIPZ-FLAG-SPIN1 overexpression cell line

(**Figure 4.23 C).** qPCR analysis showed a statistically-insignificant decrease in *AXIN2* levels, as well as a minor increase in *CYCLIN D1* and *TIAM1* levels that were also statistically insignificant (*P*= >0.05). In agreement with the SPIN1 knockdown experiments (**Figure 4.23 A and B**) *ID*-2 showed a significant increase in expression after SPIN1 induction (*P*= <0.05) **(4.23 C and D).**

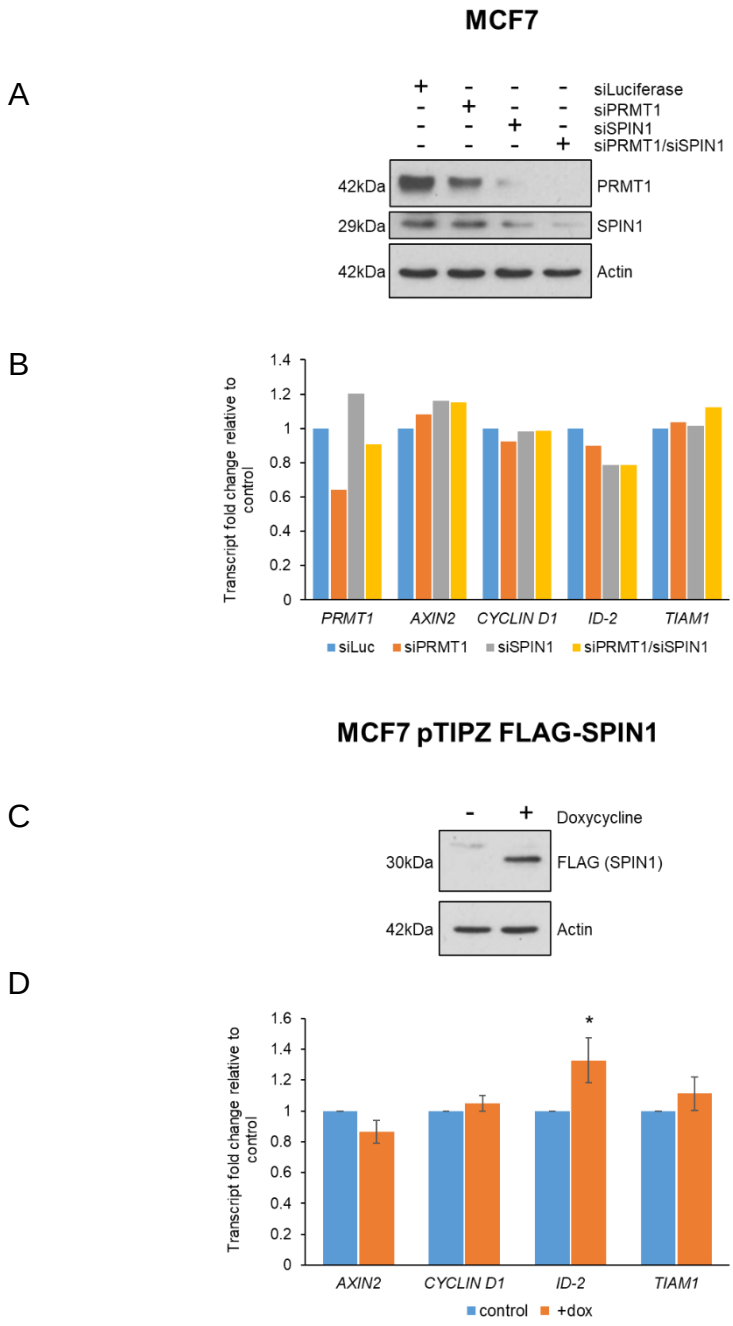


Figure 4.23: SPIN1-PRMT1 regulation of Wnt gene transcription in MCF7 cells

(A) Western blot showing expression of endogenous SPIN1 and PRMT1 after 72h siRNA-mediated knockdown. **(B)** qPCR of Wnt/β-catenin target genes in the presence of siRNA directed against SPIN1 and PRMT1 (n=2). **(C)** Western blot showing doxycycline-induced (48h, 2μg/ml) overexpression of FLAG-SPIN1 in MCF7 pTIPZ FLAG-SPIN1 cells. **D)** qPCR showing transcript levels of Wnt/b-catenin target genes in the presence of FLAG-SPIN1 overexpression (n=3). Error bars represent +/- SD. Student's T-test, * signifies *P*=<0.05.

Another group of genes that are positively regulated by SPIN1 and associated with aggressiveness and resistance to treatment in cancer are the classical stemness regulators *OCT4, SOX2* and *NANOG*[158]*.* This has been observed in triple-negative breast cancer cell lines such as MDA-MB-231 that are typically more invasive than ER^+ breast cancers[137]. Since basic tissue homology suggested a greater likelihood that these genes were also dependent on SPIN1 in MCF7 cells, the same cDNA derived from MCF7 PRMT1 and SPIN1 knockdown experiments **(Figure 4.23 A and B)** was subjected to qPCR with primers against *OCT4, SOX2* and *NANOG*.

Although preliminary, results obtained from knockdown of SPIN1 in MCF7 cells contrasted with reported behaviour in MDA-MB-231 cells, appearing to produce increases in *OCT4* and *SOX2*. Conversely, depletion of PRMT1 only slightly increased *SOX2* expression. If these data are reproducible, it again suggests that SPIN1 regulates gene expression in a cell type specific manner, even in cells derived from the same tissue. Simultaneous depletion of PRMT1 and SPIN1 appeared to increase transcript levels of *OCT4* and *SOX2*, although this was most pronounced in *OCT4*. However, combination depletion of PRMT1 and SPIN1 did not increase levels of *OCT4 or SOX2* to a greater extent than depletion of SPIN1 alone. Indeed, the only decrease across all three stemness genes came from singular knockdown of SPIN1, which reduced *NANOG* levels (**Figure 4.24 B)**.

To acquire a more comprehensive understanding of these patterns, the same cDNA from FLAG-SPIN1 overexpression in MCF7 cells that was used to interrogate Wnt target genes **(Figure 4.23 C and D)** was subjected to qPCR in the presence of stem gene primers. Overexpression of FLAG-SPIN1 stimulated an increase in all three genes, however this was only significant (P= <0.05) in *NANOG*, suggesting that SPIN1 influences at least one of the classical stemness genes in MCF7 breast cancer cells, in accordance with the observed decrease in *NANOG* upon SPIN1 knockdown **(Figure 4.24 D and B, respectively)**.

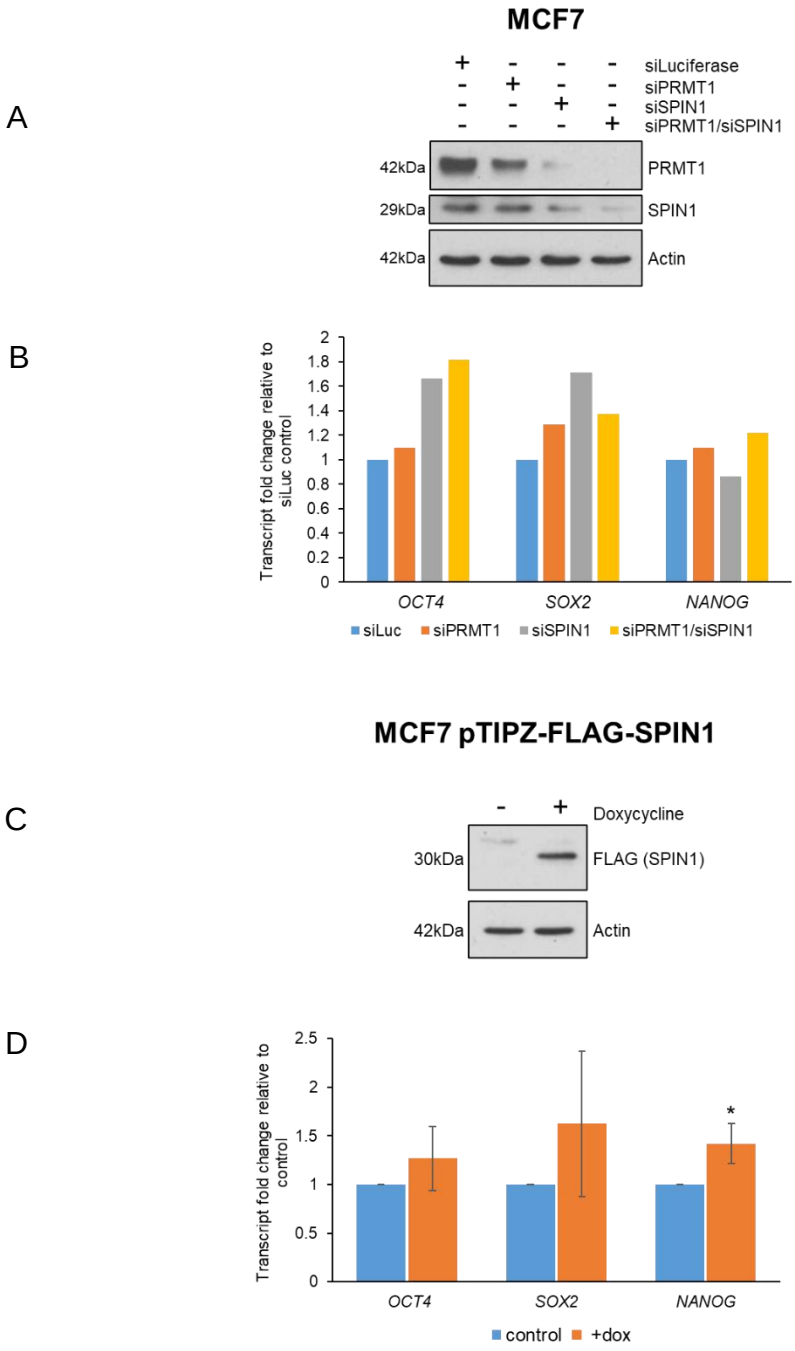


Figure 4.24: SPIN1-PRMT1 regulation of stem gene transcription in MCF7 cells

(A) Western blot showing expression of endogenous SPIN1 and PRMT1 after 72h siRNA-mediated knockdown. **(B)** qPCR of stem genes in the presence of siRNA directed against SPIN1 and PRMT1 (n=2). **(C)** Western blot showing doxycycline-induced overexpression of FLAG-SPIN1 in MCF7 pTIPZ-FLAG-SPIN1 cells. (**D)** qPCR showing transcript levels of stem genes in the presence of FLAG-SPIN1 overexpression (n=3). Error bars represent +/- SD. Student's T-test, * signifies *P*=<0.05.

Following the finding that SPIN1 overexpression up-regulated *NANOG,* the question as to whether PRMT1 activity played a role in this phenomenon was addressed. Given that depletion of PRMT1 in MCF7 cells was marginal, it was decided that the role of PRMT1 would be investigated through use of the Type I PRMT inhibitor MS023. Upon treatment with the compound, *NANOG* levels were significantly reduced, suggesting that PRMT1 activity may be required for SPIN1-mediated upregulation of *NANOG* expression **(Figure 4.25)**.

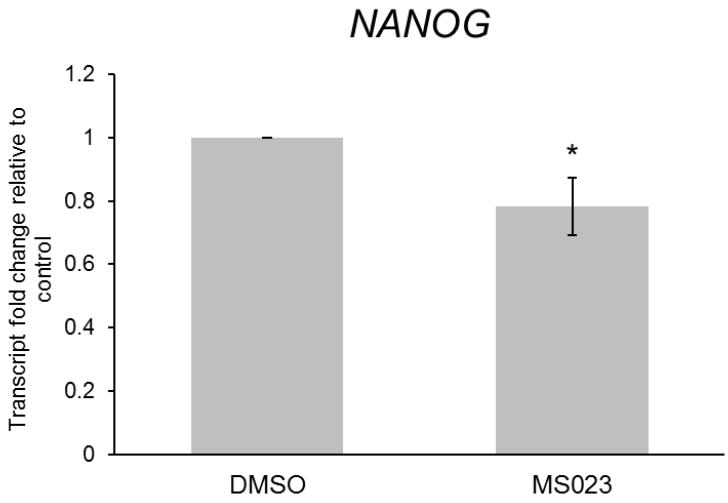


Figure 4.25: Type I arginine methyltransferase inhibition decreases *NANOG* expression in MCF7 cells.

qPCR of *NANOG* levels in MCF7 pTIPZ-FLAG-SPIN1 cells with induced overexpression of FLAG-SPIN1. Cells were either treated with DMSO vehicle control or 1μM type I PRMT inhibitor MS023 for 48 hours. (n=3). Error bars represent +/- SD. Student's T-test, * signifies *P*=<0.05.

4.4.2: SPIN1 methylation in cell cycle regulation

SPIN1 promotes cell cycle progression in glioma cells[128] and increases the proportion of G_2/M phase when overexpressed in HeLa cells[111]. PRMT1 also plays a role in cell cycle control, for example by methylating transcription factor C/EBPα, preventing its interaction with transcriptional co-repressor HDAC3. This event promotes expression of cyclin D1, increasing cell growth rate and pathogenesis in breast cancer[159]. In order to gauge whether SPIN1 and PRMT1 cooperatively influence cell cycle regulation in MCF7 breast cancer cells, cell cycle was analysed in our panel of MCF7 pTIPZ-FLAG-SPIN1-arginine mutant cells (R14K, R117K and R152K (see section **4.3**). The pTIPZ-FLAG-SPIN1-F141A mutant cell line was also included to investigate whether any observed phenotypes in SPIN1 R-K mutant cells phenocopied overexpression of the transcription-incompetent SPIN1-F141A Tudor domain II mutant. Such an event could provide crucial insight into cooperation of PRMT1 and SPIN1 in cell cycle control, since both proteins act as transcriptional coactivators. Thus, SPIN1 could potentially influence expression of cell cycle regulators via interaction with H3R8me2a and H3K4me3 at their promoters, an event mediated by SPIN1 Tudor domains I and II[47].

FLAG-SPIN1 was induced for 48 hours prior to harvest, and percentages of cells within each population occupying either G1, S or G2/M phases analysed. Interestingly, transient overexpression of wild-type FLAG-SPIN1 did not adversely affect cell cycle progression (with all stages of the cell cycle appearing similar between the induced ‘WT’ sample and ‘No Dox’ control), implying that the mitotic defects observed after SPIN1 overexpression[111,124] either manifest over

several cell cycles or are protein threshold dependent. One-way ANOVA analysis between cell lines expressing the various SPIN1 R-K mutants showed significant variation in G1 and S phases; ANOVA (*P*=0.0028 and *P*=9.65 E-07, respectively) but not for samples in G2/M phase (*P*=0.05), despite R14K and R152K cell lines showing accumulations in this phase **(Figure 4.26 A** and **B)**. Application of Student's T-test to G1 phase samples identified a significant reduction of R117K cells in this phase (*P*= 0.03). However, all other cell lines presented an insignificantly different G1 profile compared to the uninduced control, despite visible reductions in the number of R14K and R152K cells in this phase.

Application of the same statistical analysis to the S phase populations validated the most striking phenotype, where R117K mutant-expressing cells produced a ~6-fold accumulation (*P*= <0.0025) relative to the wild-type uninduced control **(Figure 4.25, A, B and C)**. Another notable feature of the S phase data was the comparatively normal profile of all other cell lines, which failed to produce any significant increases or decreases in S-phase accumulation. The unique phenotype presented by the R117K mutant was suggestive of a dominant-negative effect imposed specifically by substitution of the R117 residue. Further, inability of the F141A mutant to phenocopy R117K was suggestive of a cell-cycle control mechanism that is independent of the ability of SPIN1 to bind to H3K4me3 via its Tudor domain II.

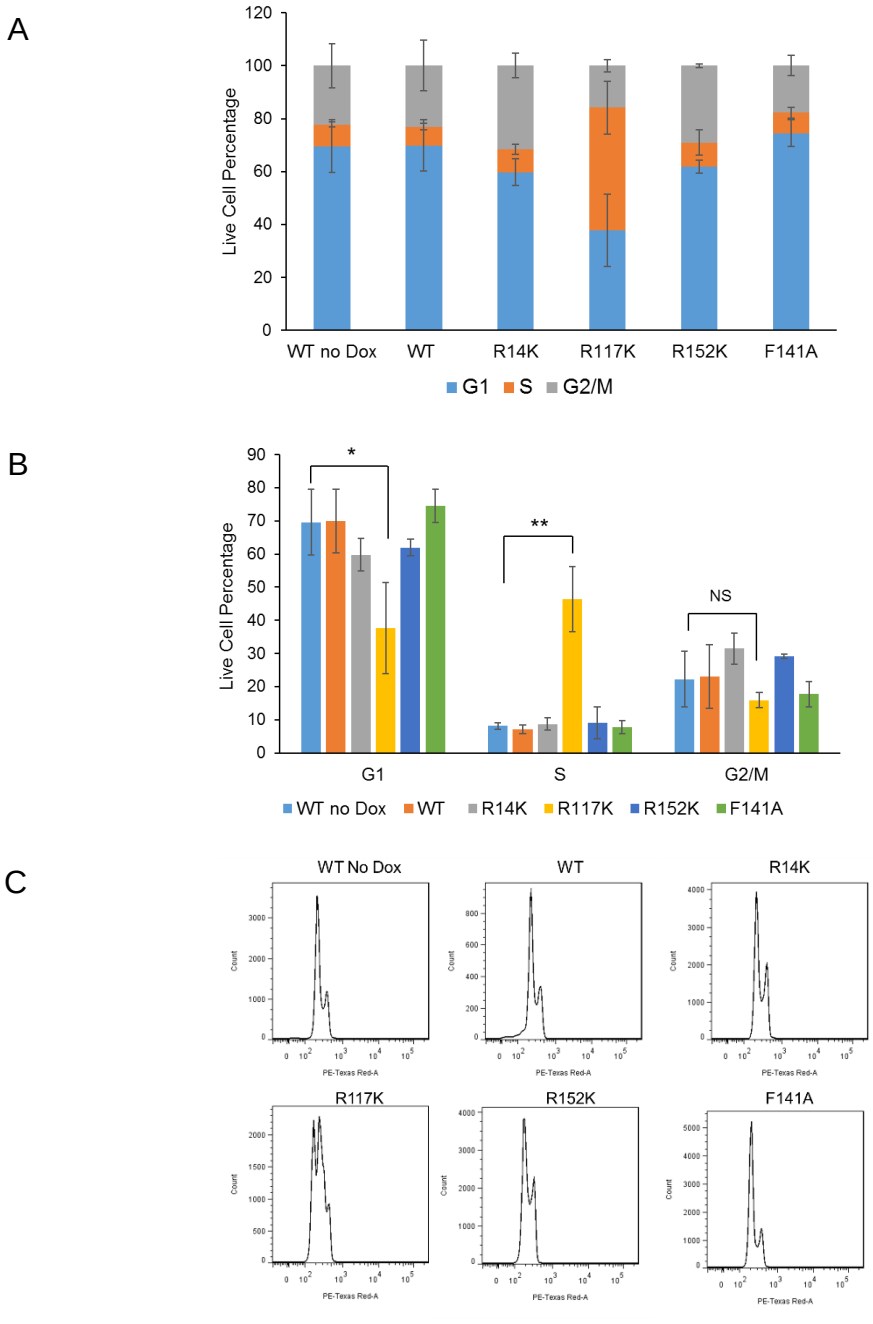


Figure 4.26: Overexpression of SPIN1 R117K mutant causes G1 loss and S phase accumulation in MCF7 cells

(A) Stacked bar graph depicting stages of the cell cycle in various MCF7 pTIPZ-FLAG-SPIN1 cell lines. FLAG-SPIN1 expression was induced with 1μg/ml doxycycline for 48 hours prior to cell cycle analysis. (**B)** Comparison of the same data in **A** grouped according to cell cycle stage (n=3).

Error bars represent +/- SD. Student's T-test (* P=<0.05 ** P=<0.01). (**C)** Flow cytometry profiles illustrating representative distributions within the cell cycle for each cell line analysed.

4.9: Identification of SPIN1 Binding Partners

SPIN1 binds asymmetrically dimethylated R8 on histone H3 via Tudor domain I, facilitating its transcriptional coactivator activity in conjunction with binding H3K4me3 via Tudor domain II [47]. To date, the methyl-histone H3 tail is one of the few known Tudor domain-directed binding partners of SPIN1. However, type I PRMTs methylate a plethora of non-histone targets, suggesting that the Tudor domain I of SPIN1 may have a functional role in binding non-histone proteins. Moreover, given the cooperativity between Tudor domains I and II in binding the histone H3 tail, it is interesting to speculate that SPIN1 interactions of non-histone proteins may also require the same methyl-lysine/methyl-arginine consensus sequence. Hence, SPIN1 may therefore be a novel reader of tandem lysine and arginine methylation.

4.9.1: SPIN1 oligomerisation and methylation status

Previous SPIN1 methylation experiments involving antibody-based detection of asymmetric dimethyl-arginine consistently detected proteins at 55kDa and 72kDa **(Figure 4.27 A)**. Initially, the 29kDa molecular weight of SPIN1 aroused suspicion that these methylated species represented dimeric (55kDa) and trimeric (72kDa) forms of SPIN1, and that the oligomerisation of SPIN1 potentially facilitated methylation or was more highly methylated compared to monomeric SPIN1.

Although higher order structuring should not survive denaturing SDS-PAGE, SPIN1 is known to dimerise *in vitro*, and chemical cross-linking experiments have suggested that it may form yet higher-order structures[111]. To investigate this,

HEK-293T cells were transfected with HA-SPIN1, which was immunoprecipitated and samples subsequently resolved via SDS-PAGE. Bands across a wide range of molecular weights were then cut from the gel and sent for mass spectrometry analysis to examine their HA-SPIN1 peptide content. Due to the simplistic nature of the investigation, an IgG control was not included, although a list of potential SPIN1 binding partners was extracted from the data, pending validation (data not shown). The sample corresponding to the 36kDa band as seen in the coomassie stain **(Figure 4.27 B)** consistently generated the most HA-SPIN1 spectra, suggesting that at least post denaturation, monomeric SPIN1 was by far the most abundant form **(Figure 4.27 C)**. Thus, the higher molecular weight proteins that are detected as methylated species in various SPIN1 methylation assays were deemed unlikely to be dimeric or trimeric SPIN1.

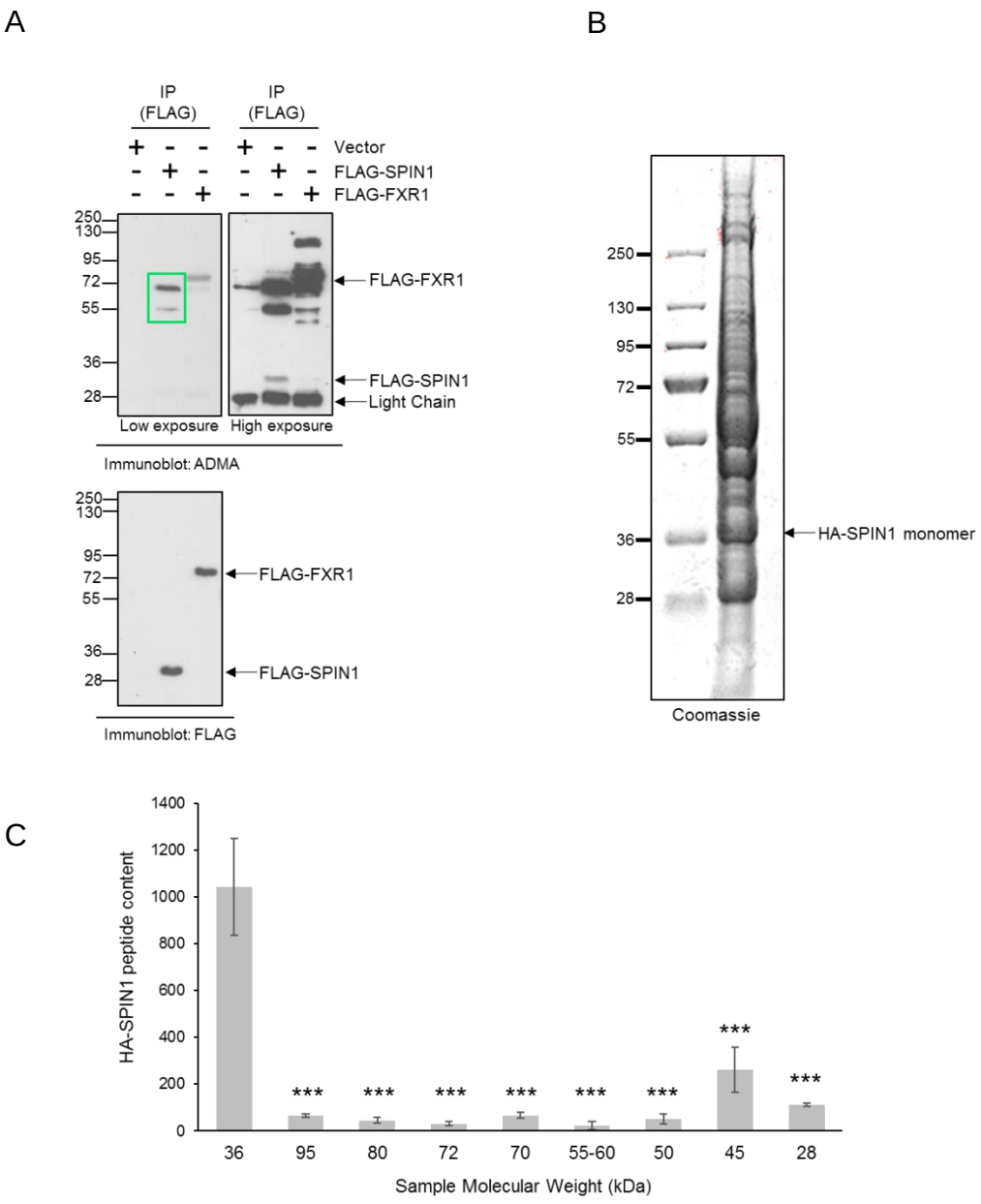


Figure 4.27: Detection of denatured SPIN1 spectra across a range of molecular weights

(A) Detection of SPIN1 methylation *in vitro* by western blotting for asymmetric dimethyl-arginine detects dominant bands at 55kDa and 72 kDa (green box). Data are acquired from immunoprecipitation of FLAG-tagged species after 24h transient transfection into HEK-293T cells. n=3. (**B)** Representative coomassie stain of HA-SPIN1 immunoprecipitation resolved via SDS-PAGE for mass spectrometry analysis. HA-SPIN1 was transiently overexpressed in HEK-

293T cells for 24h. **(C)** Average abundance of HA-SPIN1 peptides discovered in the assigned molecular weight samples, compared to the dominant HA-SPIN1-containing band at 36kDa (n=3). Error bars represent +/- SD *** *P*=<0.001.

4.9.2: SPIN1 binds asymmetrically dimethylated non-histone proteins

Since the higher molecular weight bands that arose in SPIN1 methylation assays *in vitro* **(Figure 4.27 A)** could not be attributed to SPIN1 oligomerisation **(Figure 4.27 C)**, this suggested SPIN1 could bind asymmetrically dimethylated non-histone proteins *in vitro*, possibly via Tudor domain I. To investigate this, FLAG-SPIN1 was immunoprecipitated from HEK-293T cells pre-treated with the type I PRMT inhibitor MS023, and western blotting used to detect the asymmetric dimethylation status of FLAG-SPIN1 and co-immunoprecipitated protein. Although a reduction in ADMA signal corresponding to FLAG-SPIN1 molecular weight was observed in one out of three repeats (data not shown), the methylation status of the 55kDa and 72kDa bands was reproducibly decreased **(Figure 4.28 A).** Reduction of these bands could have occurred through impairment of FLAG-SPIN1 binding to the proteins they represent (due to the reduced methylation of these unknown species), or simply as a consequence of their reduced methylation, despite still co-precipitating with FLAG-SPIN1.

To further elucidate this, a panel of FLAG-SPIN1 Tudor domain mutants were transiently overexpressed in HEK-293T cells and isolated via immunoprecipitation. Co-immunoprecipitated arginine-methylated proteins were then detected by anti-ADMA antibody. With the exception of D184H, all Tudor domain I and domain II mutants abrogated the 55 and 72kDa signals, suggesting that co-immunoprecipitation of these proteins with FLAG-SPIN1 depends upon both arginine and lysine methylation events that enable their interactions with SPIN1 *in vitro* **(Figure 4.28 B)**.

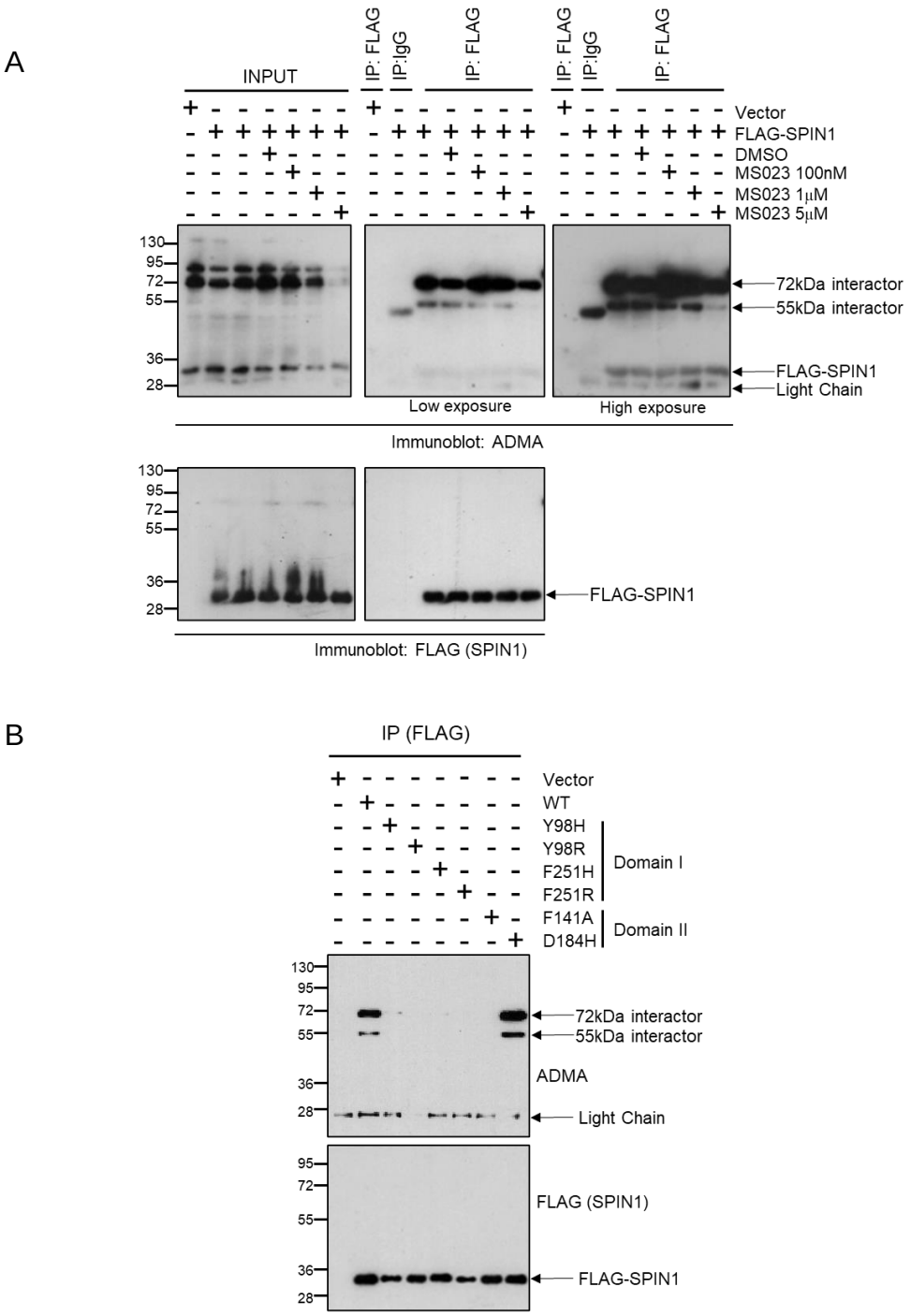


Figure 4.28: SPIN1 interacts with non-histone proteins via its Tudor domains I and II

(A) Effect of Type I PRMT inhibitor MS023 on methylation state of FLAG-SPIN1 and co-immunoprecipitated binding partners. HEK-293T cells were transiently transfected to overexpress FLAG-SPIN1 and treated with various concentrations of MS023 for 48h prior to immunoprecipitation. Upper panels: Western blot for asymmetric dimethyl-arginine content within

input (left panel) and immunoprecipitate (middle panel shows lower exposure, right-hand panel shows higher exposure). Lower panels: western blot to assess FLAG-SPIN1 content of different treatments in input (left) and immunoprecipitate (right) samples. n=3. **(B)** Western blot comparing presence of asymmetrically dimethylated SPIN1 binding partners when co-immunoprecipitated with wild-type FLAG-SPIN1 and various FLAG-SPIN1 Tudor domain mutants 24h after transfection into HEK-293T cells. n=2.

4.9.3 SPIN1 as a probe for Type I PRMT substrates

Identification of the proteins responsible for the asymmetric dimethylation signals at 55 and 72kDa requires further investigation. However, the ability of SPIN1 to discriminate between binding partners *in vitro* in a manner that likely depends upon their arginine methylation status opens up the opportunity to use SPIN1 as a novel tool for type I PRMT substrate discovery. Using reader domains that recognise post-translational modifications to profile the proteome for recipients of the respective modification is not a new idea. For example, the SH2 (Src homology 2) domain has been used as a tool to identify proteins that are tyrosine phosphorylated[98], showing successful precedent of this methodology. In addition, the survival motor neuron protein (SMN, a Tudor-domain protein) has been optimised for profiling proteins that are modified with symmetric dimethyl-arginine marks (Davies Lab, unpublished data). Accordingly, transient overexpression of affinity-tagged SPIN1 and co-immunoprecipitation of arginine-methylated substrates could allow their identification via mass spectrometry. This approach could even be applied quantitatively using SILAC in order to study conditional methylation events. For example, searching for breast-cancer enriched substrates of type I PRMTs would be possible using MCF7 and MCF10A pTIPZ-FLAG-SPIN1 cell lines, which have already been generated. This could be performed in parallel with Tudor domain I mutant cell lines (such as Y98R) that would enable discrimination between true Tudor domain interactions and non-specific binders.

To acquire qualitative indication that this methodology holds potential, HEK-293T cells were transiently transfected with FLAG-SPIN1 and mutants in Tudor

domains I (Y98R) and II (F141A), as well as a double mutant for both domains (Y98R/F141A). FLAG-SPIN1 variants were then immunoprecipitated and the contents of each sample resolved by SDS-PAGE. The respective binding profiles were subsequently visualised by SYPRO ruby staining. This method showed that the interaction of some SPIN1 binding proteins was reduced in the Tudor Domain I mutant compared to wild-type SPIN1 **(Figure 4.29,** proteins marked by green arrows**)**, suggesting that SPIN1 binds an array of asymmetrically dimethylated non-histone substrates *in vitro.* In addition, the F141A mutant demonstrated a unique binding profile, with clear reduction of a band at 72kDa **(Figure 4.29,** red arrow**)** that was not mirrored by the Y98R mutant. This suggests that certain proteins bound by SPIN1 may not depend on both Tudor domains, as in the case of SPIN1-histone H3 interaction. Consequently, the F141A mutant could also be used as a negative control in profiling for lysine methylated proteins.

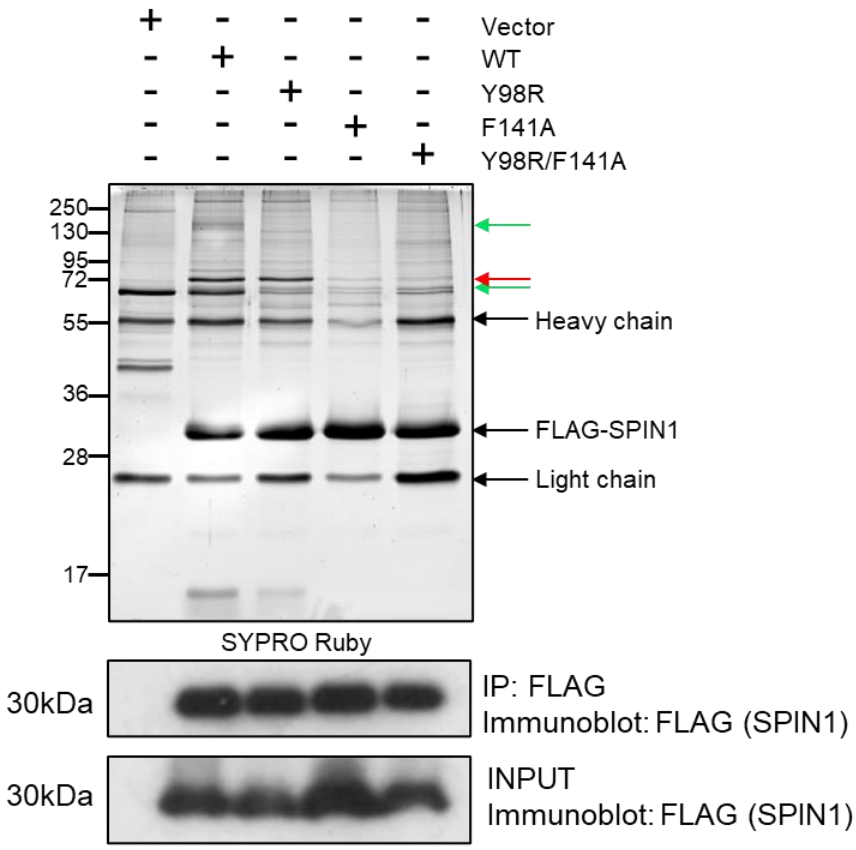


Figure 4.29: SPIN1 as a tool for profiling Type I PRMT substrates

SYPRO ruby stain comparing co-immunoprecipitation profiles of FLAG-SPIN1 with various Tudor domain mutants. FLAG-SPIN1 variants were transiently overexpressed for 24h prior to immunoprecipitation. Green arrows indicate regions in which wild-type FLAG-SPIN1 immunoprecipitated more protein than the Y98R Tudor domain I mutant. Red arrow indicates protein immunoprecipitated by wild-type FLAG-SPIN1 and the Y98R mutant, but not the F141A mutant in methyl-lysine binding Tudor domain II. n=1.

4.10: Discussion

4.10.1: PRMT1 and SPIN1: characterising the interaction

Quantitative mass spectrometry using SILAC identified a number of novel breast cancer enriched interactors of PRMT1. Among this cohort was histone code reader SPIN1, which is implicated in numerous oncogenic activities such as activation of Wnt/β–catenin signalling, mediated through its pro-transcriptional chromatin interactions[47]. Although mRNA data suggest high SPIN1 expression is of prognostic benefit in ER^+ and pan-breast cancer datasets **(Figure 4.3 A)**, SPIN1 and PRMT1 expression do correlate positively in breast cancer at the protein level **(Figure 4.3 B)**, and SPIN1 is expressed at higher levels in the MCF7 ER^+ breast cancer cell model than the MCF10A non-transformed control **(Figure 4.4).**

The finding that GFP-SPIN1 and myc-PRMT1-XGX interact *in vitro* when transiently overexpressed in HEK-293T cells **(Figure 4.5 A)** suggested that the SILAC interaction data were valid. However, as both proteins were overexpressed, this approach could provide false-positive data, since the interaction may be derived from mismanaged subcellular localisation and/or general saturation that forces proximity between both proteins that may not occur endogenously. Nevertheless, co-immunoprecipitation of endogenous PRMT1 did result in detectable SPIN1 interaction in MCF10A and MCF7 cells stably expressing inducible SPIN1, supporting the conclusion that SPIN1 is a *bona fide* PRMT1 interacting protein **(Figure 4.5 B)**. Interestingly, and in agreement with

the quantitative proteomics, interaction of endogenous PRMT1 with FLAG-SPIN1 was more robust in pTIPZ-MCF7-FLAG-SPIN1 cells despite higher expression levels of FLAG-SPIN1 in MCF10A cells, suggesting a genuinely heightened rate of interaction in MCF7 cells. Ultimately, co-IP of both endogenous SPIN1 and PRMT1 would provide an unambiguous answer to their interaction ratios between both cell lines, however attempts to IP endogenous SPIN1 were unsuccessful (data not shown). Moreover, immunoprecipitating PRMT1 followed by SPIN1 immunoblotting may not be successful as the SPIN1 antibody lacks sensitivity even within the context of a typical western blot of whole-cell lysate (data not shown), Hence, if the proportion of PRMT1 complexes that contain SPIN1 are small, this may not be detectable with commercially available SPIN1 antibodies.

The finding that recombinant SPIN1 was not methylated by recombinant PRMT1 under cell-free conditions **(Figure 4.9)** was unexpected given the numerous surface exposed arginine residues of SPIN1, and that one such residue forms part of an RG motif (R152). However, subsequent investigation using mammalian cells to generate FLAG-SPIN1 for use as a substrate of GST-PRMT1 demonstrated robust methylation of FLAG-SPIN1 **(Figure 4.10)**. This difference in outcome could be attributed to proper folding of FLAG-SPIN1 in human cells that is not replicated in bacterially-generated His-SPIN1, either exposing the requisite residue(s) for modification or providing the appropriate binding surface for GST-PRMT1 in the tertiary structure of SPIN1. Likewise, recombinant SPIN1 may form higher order structures that are not physiologically relevant but preclude cell-free methylation by PRMT1 (a phenomenon that could be determined by gel filtration analysis of recombinant SPIN1). Another possibility is

that additional post-translational modifications of SPIN1 are necessary to permit methylation by PRMT1, which can be recapitulated in mammalian but not prokaryotic cells. Numerous phosphorylation events are documented for SPIN1 **(Figure 4.6 A)**, and although arginine methylation is often characterised as an event upstream of protein phosphorylation (by either permitting or preventing the event[160]), the inverse has also been observed. For example, phosphorylation of C/EBPβ by MAPK abrogates its methylation by preventing binding of CARM1[161].

Methylation of SPIN1 could be tested as a phosphorylation-dependent event in a semi cell-free context. By removing phosphatase inhibitors sodium orthovanadate, β-glycerol phosphate and sodium fluoride from the lysis buffer and/or treating extracts with Lambda phosphatase to dephosphorylate immunoprecipitated FLAG-SPIN1, SPIN1 phosphorylation could be largely reduced for subsequent semi cell-free methylation. Comparison to a 'phospho-preserved' control that was extracted in lysis buffer with phosphatase inhibitors and/or not treated with Lambda phosphatase could then reveal whether phosphorylation of SPIN1 is permissive or repressive of subsequent methylation by GST-PRMT1. Importantly, such information could be integral in determining the mechanism and function of SPIN1 methylation *in vitro.*

Transiently overexpressed FLAG-SPIN1 is methylated *in vitro* in HEK-293T cells using a pan-ADMA antibody for detection. Although this does not confirm PRMT1 dependency, the detection of asymmetric dimethylation with an antibody that is raised against dimethylated RG motifs narrows down the range of potential

enzymes responsible to Type I PRMTs, most notably PRMT1 and PRMT6. In contrast, *in vitro* methylation assays using ^{3}[H]-methionine as a source of ^{3}[H]-SAM cofactor will detect all methylation events including lysine methylation. Detection of methyl-FLAG-SPIN1 by a pan-ADMA antibody was apparent, but significantly weaker than that of FLAG-FXR1, a protein known to be methylated by PRMT1 (Davies Laboratory, unpublished data). Interestingly, SPIN1 methylation did not represent the dominant methyl-species detected in Flag-SPIN1 immunoprecipitates **(Figure 4.11)**. This suggests that SPIN1 is not as heavily methylated *in vitro* as FXR1, or that methylation is conditional, for example, incumbent on cell cycle stage or some form of cellular stress/signalling. Alternatively, SPIN1 may also be methylated at other arginine residues outside of the RG motif, and thus the ADMA antibody is under representing the extent of *in vitro* SPIN1 methylation.

To confirm that PRMT1 was the Type I arginine methyltransferase responsible for methylating SPIN1, PRMT1 knockdown was initially attempted via siRNA in MCF7-pTIPZ-FLAG-SPIN1 cells (data not shown). However, this analysis proved challenging because overexpressed SPIN1 positively regulated PRMT1 gene expression, resulting in inefficient siRNA and shRNA-mediated knockdown of PRMT1 **(Figure 4.12 A-C)**. To overcome this, a SPIN1 construct containing a mutation in Tudor domain II (F141A) was generated, which nullified the transcriptional coactivator capacity of FLAG-SPIN1 **(Figure 4.1 B)**. Expression of Flag-SPIN1-F141A allowed for effective siRNA-mediated depletion of PRMT1, confirming *PRMT1* as a gene transcriptionally regulated by SPIN1. However, knockdown of PRMT1 did not reduce FLAG-SPIN1-F141A methylation *in vitro,*

as detected by the anti-ADMA antibody **(Figure 4.14 A and B)**. This was considered a potential consequence of the ADMA antibody detecting a non-PRMT1-mediated form of FLAG-SPIN1-F141A methylation, and the experiment was repeated using autoradiography to detect all forms of FLAG-SPIN1–F141A methylation. To our surprise, this experiment also failed to show a reduction of FLAG-SPIN1-F141A methylation **(Figure 4.14 C and D),** suggesting that SPIN1 was not methylated by PRMT1 *in vitro.* One notable caveat to this experiment was the relatively inefficient knockdown of PRMT1, which may not have been sufficient to abrogate PRMT1-dependent methylation of FLAG-SPIN1-F141A, leaving this result somewhat ambiguous.

Due to consistent demonstration that PRMT1 and SPIN1 interact *in vitro* and robust methylation of SPIN1 by recombinant GST-PRMT1 in a semi cell-free context, the F141A mutation was investigated further as a potential antagonist in proving *in vitro* PRMT1-mediated methylation of SPIN1. It emerged that this mutation reduced the interaction of SPIN1 and PRMT1 in HEK-293T cells, and potentially altered subcellular localisation of SPIN1 in MCF7 cells **(Figures 4.15 A** and **B).** As such, the methylation of FLAG-SPIN1-F141A that was detected may have been contributed by other endogenous arginine and/or lysine methyltransferases whose interaction with SPIN1 is less dependent on the integrity of the Tudor domain. Reduced interaction of SPIN1 and PRMT1 may result from altered subcellular localisation as both proteins localise to chromatin, which would provide an opportunity for the hypothetical PRMT1-dependent methylation event. Although the subcellular localisation experiments require optimisation (due to the cytoplasmic marker tubulin contaminating nuclear and

chromatin bound fractions) wild-type FLAG-SPIN1 was overwhelmingly chromatin localised (80% of total FLAG-SPIN1), whereas FLAG-SPIN1-F141A was only 50% chromatin bound and showed a 3-fold increase in the cytoplasmic fraction over the wild-type. This may suggest that preservation of Tudor domain II is important for retention in the nucleus at large (possibly by mediating interaction with an unknown chaperone), as well as chromatin association. Whilst in close proximity to chromatin, additional protein factors may contribute to PRMT1-mediated methylation of SPIN1, either by addition of other methylation-permissive modifications to SPIN1, or by providing a physical support for both proteins to interact.

Although interaction of PRMT1 and SPIN1 may occur within close proximity to chromatin, engagement of SPIN1 with the histone epigenetic marks (H3K4me3 and H3R8me2a) might not be necessary to facilitate this interaction. This is suggested by the D184H mutant showing no reduction in PRMT1 interaction **(Figure 4.15 A)**, despite similar incompetence to the F141A mutant in activating transcription **(Figure 4.13 B).** Although subcellular localisation of the D184H mutant has not been studied, comparison to wild-type SPIN1 and the F141A mutant by cellular fractionation could prove useful as the D184H mutant may display unaltered localisation, due to its potentially histone-specific mutation **(see section 4.2.5)**. Further, the regions of SPIN1 that are responsible for PRMT1 interaction are not known and the possibility that an intact lysine-interacting Tudor domain II may be necessary for more efficient interaction of SPIN1 with PRMT1 cannot be ruled out. Although PRMT1 is not known to harbour methyl-lysine modifications (www.phosphosite.org), it could be feasible that a cofactor-like

PRMT1 binding partner is lysine methylated and that this promotes SPIN1 methylation by physically mediating SPIN1-PRMT1 interaction. If true, then the D184H mutant may allow simultaneous overexpression of SPIN1 and knockdown of PRMT1 in MCF7 cells without removing cellular context that would permit PRMT1-dependent methylation of SPIN1, thus enabling PRMT1-dependent methylation of SPIN1 to be tested in MCF7 cells.

Type I PRMT-dependent methylation of FLAG-SPIN1 was demonstrated in HEK-293T cells by using the MS023 inhibitor **(Figure 4.16)**. MS023 has also shown activity towards PRMT3 and PRMT6, however it has a higher specificity towards PRMT1, with a reported 12-fold and 6-fold greater efficiency than its inhibition of PRMT3 and PRMT6, respectively[155]. The preliminary observation (n=1) that MS023 decreased SPIN1 methylation in cells, and that PRMT1 robustly methylates SPIN1 in semi cell-free circumstances, does imply that PRMT1 can methylate SPIN1 *in vitro.* In the future, it would be interesting to demonstrate this in breast cancer cells, and determine if methylation is enhanced compared to non-transformed mammary epithelial cells. With the systems now in place, this important question will be able to be addressed.

4.10.2: Putative sites of SPIN1 methylation

Considering the small molecular weight of SPIN1 (29kDa), the protein is subjected to a high volume of post-translational modifications. Already documented among these are lysine methylation, acetylation and ubiquitination and several serine and threonine phosphorylation events

(www.phosphosite.org). Mass spectrometry analysis of SPIN1 transiently overexpressed in HEK-293T, MCF10A and MCF7 cells has identified several novel serine, threonine and tyrosine phosphorylation events **(Figure 4.17B)**. Of the known PTMs that were identified through MS (phosphorylation of S38, S39, Y91, Y98, S124 and S199) only S124 is detailed in the literature. This residue is targeted by Aurora A kinase, however the functional significance of this event is not known[125]. In addition, monomethylation of SPIN1 at R14 and R117 was also detected. Although PRMT1 is typically associated with asymmetric dimethylation of arginine residues, it is also capable of monomethylation, which is generally regarded as an intermediate in the formation of asymmetric dimethylation[13].

Validation of PRMT1 and SPIN1 interaction *in vitro* and confirmation that recombinant GST-PRMT1 can methylate FLAG-SPIN1 prompted further interrogation of these methylation sites using semi cell-free methylation assays. In addition to R14 and R117, R152 was included due to its placement with the sole RG motif, as well as the finding that immunoprecipitated FLAG-SPIN1 cross reacted with an ADMA antibody raised against methyl-RG motifs. Semi cell-free methylation assays were initially undertaken with FLAG-SPIN1 variants derived from HEK-293T cells. These experiments did not yield apparent differences in methylation between wild-type FLAG-SPIN1 and any of the mutants, and normalisation to their IP western blot failed to demonstrate a significant decrease in methylation **(Figure 4.18)**. However, a 40% average decrease in methylation for the R117K mutant in two out of three experimental repeats promoted further validation in another system. Indeed, analysis of FLAG-SPIN1-R117K methylation when derived from MCF7 cells inducibly overexpressing the protein

demonstrated a statistically significant decrease in methylation compared to wild-type FLAG-SPIN1 **(Figure 4.19)**. This observation further reinforced the likelihood that R117 is a target for methylation *in vitro*, possibly by PRMT1.

By contrast, mutation of R14 to a lysine appeared to increase methylation of FLAG-SPIN1-R14K relative to the wild-type control when proteins were extracted from from either HEK-293T and MCF7 cells, although this was not significant. Although methylation was not reduced by this mutation, R14 is suggested by mass spectrometry to be methylated. However, the MS data do not indicate which PRMT is responsible for this modification, meaning that SPIN1 may not be methylated by PRMT1 at R14. Another factor that suggests PRMT1 may not methylate SPIN1 at R14 is the location of the residue within an RxR motif (specifically RSR at positions 12, 13 and 14 respectively), which is the preferred target motif of PRMT7[97]. As the only type III PRMT, PRMT7 is also the only member of this family to be exclusively capable of delivering a monomethyl mark to arginine residues (MMA)[97], and methylated R14 was consistently found in the monomethyl form by mass spectrometry.

Although the N-terminal tail in which R14 is found remains unsolved and poorly-characterised, several post-translational modifications have been discovered in this region, including methylation at K7, acetylation at K28 and numerous serine phosphorylation events towards the C-terminal region of the tail, suggesting that this region is heavily modified and likely important in fine-tuning various modes of SPIN1 function. This is typical of structurally disordered regions, which often

rely on post-translational modifications to bring structure and function to their otherwise amorphous nature[162]. Indeed, arginine methylation has been counted among these modifications such as in the case of histone tails, altering their potential for further functional modification[50]. Mass spectrometry also identified the novel phosphorylation event at S13 directly adjacent to R14. Although S13 phosphorylation requires confirmation, this close proximity suggests that cross talk between PTMs occurring on S13 and R14 could possibly occur. Phenomena of this nature are increasingly documented, for example in the case of transcription factor FOXO1, which is methylated by PRMT1 at R248 and R250, preventing phosphorylation of FOXO1 by AKT[23].

Functionally, the N-terminal tail (amino acids 1-49) has been shown to modulate subcellular localisation of SPIN1 as it contains a nucleolar localisation sequence (NoLS), a patch of basic amino acids consisting of amino acids 28-44 that is recognised by nucleolar chaperone B23 (Nucleophosmin)[112]. Whether methylation of R14 plays a role in subcellular localisation has yet to be determined. Another feature to consider with the N-terminal tail is the relatively high preponderance of serine and arginine residues, some of which follow one another. For example, R14 sits within an RSR sequence, and a similar sequence, RSS encompasses residues 37-39 (another sequence, RTS constitutes amino acids 30-32). Regions that are rich in these motifs are known as RS domains, and are typically found in proteins belonging to the SR family, which also feature an RNA binding domain that contains an RNA-recognition motif (RRM). SR proteins are often associated with pre-mRNA splicing, and the RS domain is key in determining localisation of the protein to perichromatin fibrils, which are co-

transcriptional splice sites, from inter-chromatin granule clusters (IGCs, storage sites for mRNA splicing factors). Further, this activity is reliant on phosphorylation of serine residues within the RS domain[163] and all serine residues within uninterrupted RS sequences in the SPIN1 N-terminal tail (S13, S38 and S39) were confirmed phosphorylation sites by mass spectrometry analysis **(Figure 4.17).**

Another important function of the RS domain is contact and recruitment of other protein factors involved in splicing, as well as forming a contact with RNA in the prespliceosome branchpoint[164]. It is therefore possible that the N-terminal tail of SPIN1 is involved in formation of the spliceosome by influencing subcellular localisation of SPIN1, as well as physically recruiting other SR family proteins and perhaps directly binding mRNA. DAVID gene ontology analysis (https://david.ncifcrf.gov/) of SPIN1 co-immunoprecipitated proteins supported this, showing high enrichment of proteins involved in mRNA splicing **(Appendix B)**. Thus, localisation of SPIN1 to perichromatin fibrils could be mediated through methylation at R14, by either promoting or antagonising serine phosphorylation of adjacent residues. It should be noted that all of the proposed biology surrounding R14 methylation is speculative, and this residue has yet to be confirmed experimentally as a target of arginine methylation. Further, the hypothetical PRMT responsible for this event has also not been identified.

The finding that R117 appeared to be methylated by GST-PRMT1 under semi cell-free conditions suggests that at least one of the methylated sites discovered

by mass spectrometry is a PRMT1-dependent event. Although cell-free methylation of recombinant His-SPIN1-R117K has yet to be compared to wild type His-SPIN1, optimisation of this assay using Tris-HCl buffer as a replacement for PBS may further legitimise R117 as a target of PRMT1. The 2NS2 crystal structure of SPIN1 illustrates the phosphate binding capacity of the protein via a phosphate-binding loop (P loop) that connects Tudor domains I and II, and chelates two phosphate ions (C301 and C302). Among the residues responsible for contributing hydrogen bonds to this interaction is R117 **(Figure 4.20)**, which chelates C301[111]. Due to steric exclusivity and methylation removing hydrogen bond donor sites from arginine residues, it was hypothesised that R117 may be methylated by PRMT1, in a manner that is mutually exclusive with chelation of phosphate C301. This was investigated in a cell-free scenario by attempting to methylate His-SPIN1 with GST-PRMT1 in the presence (PBS) and absence (Tris-HCl, HEPES) of phosphate within the buffer. In addition, aluminium hydroxide was added to a duplicate of each reaction in an attempt to sequester phosphate from His-SPIN1. Use of aluminium hydroxide for this purpose was unsuccessful, although certain caveats, such as poor solubility and lack of positive control for this condition make it difficult to infer whether this was simply a technical issue.

More encouraging was the striking increase in His-SPIN1 methylation when 20mM Tris-HCl was used to buffer the reaction in place of PBS. Part of this increase could be attributed to generic differences in buffer action, as Tris-HCl also appeared to enhance auto-methylation of GST-PRMT1, suggesting that this buffer may increase the general efficiency of GST-PRMT1 catalysis in cell-free methylation reactions. However, the difference in GST-PRMT1 auto-methylation

between PBS and Tris-HCl reactions is far less dramatic than the difference between His-SPIN1 methylation from the respective reactions, suggesting that removal of phosphate from the reaction may specifically facilitate His-SPIN1 methylation **(Figure 4.21 A)**. Surprisingly, a HEPES-containing buffer was less able to support the activity of GST-PRMT1, as evidenced by its weak auto-methylation despite recombinant His-PRMT1-Y29F/M38G being active in HEPES buffer when utilising Pob-SAM to alkynylate recombinant histone H4 in cell-free assays **(Figure 3.3).** However, the alkynylation assays utilised HEPES at a 20-fold lower concentration (50mM), implying that inability of GST-PRMT1 to methylate His-SPIN1 using SAM as a cofactor may be attributed to excessive buffer concentration. Thus, re-attempting cell-free methylation of His-SPIN1 with 50mM HEPES pH 7.4 may prove successful, further supporting the hypothesis that absence of phosphate promotes methylation of His-SPIN1 in a cell-free context.

Although semi cell-free data suggest that R117 is a PRMT1 methylation site, mutation to lysine failed to completely supress SPIN1 methylation. Whilst this could be caused by heterodimersation of FLAG-SPIN1-R117K with endogenous SPIN1 (with the endogenous/wild-type half of the dimer partially compensating for lost methylation), it also suggests that additional sites may be targeted for methylation by GST-PRMT1. Identification of additional methylation sites could be approached by truncation of FLAG-SPIN1 via progressive deletion of Tudor domains (Δdomain I, Δdomain II) and subjection to semi cell-free methylation. However, based on current findings, a more targeted approach could be applied

first. If binding of phosphate ions is modulated by methylation of R117, this principle could be applied to two more arginine residues within the phosphate-binding pocket: R122 and R158. Both of these residues are exposed and therefore accessible, and their role in binding phosphate C302 could mean that their methylation could further prevent phosphate binding **(Figure 4.20)**. Although speculative, if true this could mean that PRMT1-mediated methylation of SPIN1 *in vitro* functions as a switch to modulate phosphate binding of SPIN1.

One confounding difference between the semi cell-free methylation assays and the finding that phosphate chelation inhibits SPIN1 methylation is that semi cell-free methylation assays were carried out in PBS. Why (if phosphate presence is an influence) FLAG-SPIN1 can be methylated in PBS under these circumstances is unclear. This may arise from additional post-translational modifications that occur *a priori* within the cell to a sub-population of the FLAG-SPIN1 pool that discourage phosphate chelation during the methylation reaction (such as phosphorylation of other phosphate-chelating residues in the P loop), allowing R117 to be methylated. Reliance on *a priori* P loop phosphorylation for semi cell-free SPIN1 methylation makes sense, as FLAG-SPIN1 generates a relatively weak signal by autoradiography, despite coomassie staining suggesting that it is the most abundant species (with the exception of GST-PRMT1, **Figures 4.18 and 4.19**). If other PTMs are promoting semi cell-free FLAG-SPIN1 methylation in spite of phosphate presence, the relative inefficiency of this methylation could be solved by carrying out the reaction(s) in Tris-HCl, potentially maximising the difference between methylation of the R117K mutant and wild-type FLAG-SPIN1.

Another means of detecting SPIN1 methylation by GST-PRMT1 would be through cell-free and/or semi cell-free methylation using non-radioactive SAM, coupled with mass spectrometry. This approach had not been applied previously, as attempts to detect SPIN1 methylation by western blotting after semi cell-free methylation with GST-PRMT1 using an anti-ADMA antibody had proven unsuccessful (data not shown). Given subsequent experiments using MS023 have demonstrated that SPIN1 methylation *in vitro* is PRMT1-independent when detected by an anti-ADMA antibody **(Figure 4.28)**, it is clear that the ADMA antibody is not a suitable tool to detect PRMT1-dependent events. Approaches such as mass spectrometry would allow unbiased analysis of cell-free PRMT1-mediated methylation of SPIN1, and could yield important findings.

Analysis of His or FLAG-tagged SPIN1 species by mass spectrometry also opens up the opportunity to investigate other methylated bands that appear through autoradiography. Of particular interest is a band that appears in semi cell-free methylation assays at 36kDa **(Figures 4.18 and 4.19)**, due to its closeness in size with SPIN1, suggesting that it may be a more heavily modified sub-population of SPIN1.

Finally, despite appearing an attractive target for PRMT1-mediated methylation, semi cell-free methylation of FLAG-SPIN1-R152K failed to produce a significant reduction in methylation signal relative to wild type FLAG-SPIN1. This, in

combination with the previous finding that knockdown of PRMT1 *in vitro* did not reduce methylation of FLAG-SPIN1-F141A when detected by an anti ADMA antibody suggests that R152 may be methylated, but not by PRMT1. Of the other Type I PRMTs such as PRMT2, PRMT3 or PRMT6, which also target RG motifs for asymmetric dimethylation[22], PRMT2 may be the logical choice for initial investigation, as it provides the H3R8me2a mark that is bound by Tudor domain I of SPIN1[47], placing the two proteins in close proximity. Given the placement of R152 on top of Tudor domain II, methylation of this residue may alter the binding capacity of SPIN1 to H3K4me3 by altering the conformation of this domain.

4.10.3: PRMT1 and SPIN1: transcriptional regulation of oncogenes

SPIN1 and PRMT1 are both known modulators of transcriptional activation. In addition to promoting expression of ribosomal RNA[112], SPIN1 regulates transcription of numerous genes involved in oncogenic pathways, such as Wnt/β-catenin effectors[47] and stemness genes[137].

In order to gain an understanding of the possible contribution of SPIN1 and PRMT1 to ER^+ breast cancer, both genes were knocked down singularly and in combination in MCF7 cells via siRNA treatment. Transcript levels of Wnt/β-catenin effector genes *AXIN2, CYCLIN D1, TIAM1* and *ID-2* were then analysed via qPCR to understand the contribution of PRMT1 and SPIN1 for expression of these oncogenic transcripts. Data acquired from these experiments did not suggest a strong influence of PRMT1 or SPIN1 on expression of these genes **(Figure 4.23 A and B)**. A more obvious effect was detected in HCT-116

colorectal cells in which SPIN1 dependency was originally described in the context of Wnt gene expression[47]. Although knockdown of SPIN1 reduced *AXIN2* levels, this was not significant – however a significant reduction was observed in the case of PRMT1 knockdown. SPIN1 reduction did significantly reduce *ID-2* levels however, resulting in an average *ID-2* transcript level that was highly similar to the effect of PRMT1 knockdown. Simultaneous knockdown of SPIN1 and PRMT1 further reduced *AXIN2* and *ID-2* levels, although when compared to singlular knockdown of SPIN1, this was only significant in the context of *ID-2* expression. This observation suggested that SPIN1 and PRMT1 do not regulate *ID-2* expression in an epistatic manner, and therefore influence its expression via different pathways **(Figure 4.22)**.

These differences between MCF7 and HCT-116 cells suggest that SPIN1-mediated control of genes in this pathway may be cell-type specific. However, technical issues may also have impacted upon these experiments, because in contrast to Su et al[47]., we could not detect any changes in *TIAM1* and *CYCLIN D1* levels after SPIN1 depletion in HCT116 cells. This may be due to less efficient knockdown of SPIN1 in our HCT116 experiments, and suggests a critical threshold at which SPIN1 is still able to maintain a normal expression profile of certain genes. Similarly, inefficient knockdown of PRMT1 may have obscured its role in modulating expression of these genes, as PRMT1 has been previously implicated as a positive transcriptional regulator of *CYCLIN D1* levels in triple negative breast cancer[159], as well as in MCF7 cells through methylation of the estrogen receptor[92].

This knockdown ‘threshold’ issue may also explain the apparent lack of dependency on SPIN1 and PRMT1 for expression of *OCT4, SOX2* and *NANOG* in MCF7 cells **(Figure 4.24 A and B).** This insufficient knockdown hypothesis is supported by overexpression of FLAG-SPIN1 upregulating all three genes (although this event was only statistically significant in the case of *NANOG,* **Figure 4.24 C and D).** Reduction of *NANOG* levels in response to MS023 treatment during FLAG-SPIN1 overexpression **(Figure 4.25)** suggested co-operation between SPIN1 and a type I PRMT (potentially PRMT1) in regulating *NANOG* expression at the transcriptional level. However, whether this results from mutual occupation of the *NANOG* promoter has yet to be determined. These observations support other findings that SPIN1 and PRMT1 levels positively correlated with transcriptional up-regulation of all three stemness genes in triple-negative breast cancer and oesophageal squamous cell carcinoma, respectively[88,137].

A potentially confounding factor when interrogating genes that regulate stemness is that changes in their expression can be masked when testing a whole cell population, since cells with cancer stem cell (CSC) properties comprise only a small proportion of the whole cell population. Indeed, Drago-Ferrante et al, who determined the positive regulatory role of SPIN1 on stemness genes in triple-negative breast cancer did so with cancer-stem cell enriched mammospheres[137]. As such, future investigation into the effects of SPIN1 and PRMT1 on cancer stem properties in MCF7 cells may require enrichment of the cancer stem cell population to enable further study of their relationship to PRMT1 and SPIN1. It has previously been shown that PRMT1 knockdown reduces the number of stem

cells within the MCF7 whole cell population, despite PRMT1 not being necessary for self-renewal[165]. If chromatin immunoprecipitation experiments prove simultaneous occupation of the *NANOG* promoter, then inhibition of PRMT1 and SPIN1 could prove devastating to the MCF7 cancer stem cell niche by attacking two key transcriptional agonists simultaneously.

4.10.4: SPIN1: cell cycle regulation by putative arginine methylation sites

Due to basic characterisation of SPIN1 as an modulator of the cell cycle[111], flow cytometry was employed to investigate the potential role of SPIN1 methylation in cell cycle control. From a panel of MCF7 pTIPZ-FLAG-SPIN1 cells overexpressing various mutants for potential methylation sites, the FLAG-SPIN1-R117K variant showed a dramatic accumulation in S-phase and a concurrent decrease in the G1 population. This phenotype is most likely unrelated to the transcriptional coactivator properties of SPIN1, as the cell cycle of FLAG-SPIN1-F141A expressing cells did not differ from the wild-type control **(Figure 4.26).**

The high accumulation of cells in S phase may be indicative of S phase checkpoint activation, a prominent cause of which is defective chromatin assembly during DNA replication. Other causes of S phase arrest such as exposure to DNA damaging agents[166] and dNTP metabolite deprivation[167] are documented. However, the literature, as well as experiments in this thesis suggest that SPIN1 is highly chromatin associated **(Figure 4.15).** This supports the idea that the phenotype may occur through defective interaction with chromatin, suggesting that FLAG-SPIN1-R117K overexpression may affect DNA

repair during replication, or chromatin assembly. Defective chromatin assembly has been demonstrated via a mutation in the p150 subunit of CAF1, a protein complex integral in chromatin assembly during DNA replication. The dominant-negative effect of this subunit reduced DNA replication and increased spontaneous DNA damage, resulting in S-phase arrest[168]. Similarly, overexpression of FLAG-SPIN1-R117K without knockdown of endogenous SPIN1 suggests a dominant-negative effect of the mutation. Although the role of SPIN1 is unclear in this phenotype, we have identified CAF1B (the p60 subunit of the CAF complex) as a putative interactor of SPIN1 from mass spectrometry analysis (data not shown). Knockdown of CAF1B has been reported to cause a ~10-fold reduction in nucleosome assembly, inhibiting DNA replication and causing DNA damage[169].

The finding that R117 contributes phosphate-chelating hydrogen-bonds[111] suggests that its methylation would hinder interaction with phosphate groups by removal of hydrogen bond donors. Thus, SPIN1 may either bind the phosphate backbone of DNA directly during replication, or bind to phosphorylated proteins involved in coordinating DNA replication and chromatin assembly during S-phase. Methylation of SPIN1 at R117 would likely therefore prevent such interactions or stimulate disengagement of SPIN1 from replicating DNA or the phosphorylated protein to which it is bound. SPIN1 binds DNA in a sequence-independent manner as determined by gel-shift analysis, in which it exhibits a preference for superhelical double stranded DNA over more open or single-stranded DNA[111]. Supercoiled DNA is a well-documented by-product of *de novo* replication, and is crucially relieved by the activity of Topoisomerases to prevent

replication stalling and subsequent replication stress[170]. Since SPIN1 exhibits a preference for binding DNA in a superhelical conformation, it may therefore aid in the resolution of superhelices by dissociating from replicating DNA in a methylation-dependent manner, allowing topoisomerase activity to commence. Further, many proteins involved in replication stress such as RPA1 and MCM7 are phosphorylated during S-phase; (the former by Chk1 and ATR in response to replication stress, and the latter by $p56^{Lyn}$, ILK and AKT, all of which promote chromatin loading)[171], and both examples have been identified during this thesis as putative SPIN1 interactors (data not shown). Thus, an abundance of phosphorylation events could facilitate interaction of factors crucial to the replication stress response and/or chromatin loading with the phosphate-binding domain of SPIN1.

Whether SPIN1 influences replication, and if aberrant function causes replication stress requires further study. Immunostaining for RPA foci or performing DNA fibre assays to observe replication fork stalling would provide a visual indication of SPIN1-R117K induced replication stress. Testing for activation of the ATR pathway (by western blotting for increases in phosphorylated Chk1 and Chk2) in the presence of SPIN1-R117K overexpression is also another means to corroborate this phenomenon. To interrogate improper assembly of chromatin, digestion by micrococcal nuclease (Mnase) could be applied, as improperly loaded nucleosomes would increase sensitivity to the enzyme.

Another potential cause of cell cycle arrest is nucleolar/ribosomal stress caused by aberrant ribosome biogenesis. SPIN1 negatively modulates ribosomal stress signalling via suppression of the Rp-MDM2-p53 pathway. By binding the large ribosomal subunit uL18 (RPL5) and sequestering it to the nucleolus, SPIN1 prevents it from binding ubiquitin ligase MDM2. Without this interaction, MDM2 is free to ubiquitinate p53, targeting it for degradation with anti-apoptotic consequences. However, reduction of SPIN1 levels activates p53 by disrupting ribosome biogenesis, resulting in a surplus of free ribosomal proteins that can subsequently bind MDM2 and prevent degradation of p53[172].

A major caveat to the insinuation that R117 methylation mediates Rp-MDM2-p53 is that this pathway causes cell cycle arrest in G1 when activated[173], in contrast to the S phase accumulation observed during FLAG-SPIN1-R117K overexpression. SPIN1 may therefore cause cell cycle arrest through aberrant ribosomal modulation in a manner that is not yet characterised. For example, the phosphate binding loop of SPIN1 could help co-ordinate ribosomal DNA (rDNA) in the formation of nucleolar organising regions (NORs), areas in which the chromosomal regions containing ribosomal DNA are transcribed, and upon which the structure of the nucleolus depends[174]. Repression of rDNA transcription causes nucleolar stress, and importantly, this is a phenomenon that has been observed to cause S phase arrest in MCF7 cells. This was found to be a result of the Rp-MDM2-p53 response being circumvented by MCF7 cells overexpressing MDM2, thus reducing p53 and preventing G1 arrest. MCF7 cells are therefore allowed to enter S phase, where activation of ATR signalling in response to replication stress activates the intra-S phase checkpoint[175]. Hence, the p53

deficient quality of MCF7 cells allows a connection between both hypotheses presented (replication stress or ribosomal/nucleolar stress as a cause of S phase arrest), suggesting that they may not mutually exclude one another in the context of FLAG-SPIN1-R117K-mediated S phase accumulation.

Verification that SPIN1-R117K induces ribosomal stress could be acquired through a number of means. Firstly, co-localisation of SPIN-R117K to the nucleolus would need to be established by immunostaining SPIN1 and nucleolar/ribosomal markers such as Fibrillarin. Second, up-regulation of p53 and downregulation of nucleolar marker nucleophosmin (signs of the nucleolar stress response) could be verified by western blotting and immunostaining, respectively. Finally, changes in the level of ribosomal genes (by qPCR for ribosomal RNA, and western blotting for ribosomal proteins) could be compared to control cells in which SPIN1-R117K is not overexpressed.

In the event that SPIN1 is suspected to influence organisation of ribosomal DNA, 3D approaches such as 3C could be used to interpret contact alterations within the same gene. Although contacts between different chromosomes using Hi-C is not possible due to the highly repetitive nature of the ribosomal genes[176], high resolution imaging after staining nucleoli for UBF (Upstream Binding Transcription factor, a marker of active transcription) has been recently used to acquire visual understanding on the organisation of ribosomal DNA[177].

4.10.5: Therapeutic targeting of a theoretical PRMT1-SPIN1 axis in cancer

The preliminary findings that SPIN1 increases transcription of stemness gene *NANOG* and that prevention of SPIN1 methylation at R117 may cause cell cycle arrest (potentially through replication stress and/or ribosomal stress) presents several opportunities for therapeutic intervention. PRMT1 is a gene already identified as a drug target, with compound GSK3368715 being tested for efficacy on solid and haematological tumours (trial ID NCT03666988). Anti-SPIN1 drug development has not yet commenced, however several small molecule inhibitors have been developed for probing SPIN1 function[134,135,178]. This combined with ever-mounting evidence that SPIN1 influences many facets of malignancy across an array of cancers places SPIN1 as a potential future candidate for cancer therapy.

Before choosing the approach for therapy however, a more detailed mechanistic understanding of the aforementioned phenomena is required. For example, transcription of *NANOG* could be targeted through inhibition of SPIN1 Tudor domains to prevent its transcriptional coactivator function, in tandem with inhibition of PRMT1, which would prevent deposition of transcription-promoting H4R3me2a mark. This could prove highly effective in targeting the cancer stem cell niche in breast cancer **(**and potentially other cancers, **Figure 4.30 A)**, however ChIP data showing co-localisation of these factors to the promoter of *NANOG* are required to support this hypothesis.

With regard to the role of the P loop and general phosphate-binding capacity of SPIN1 in cell cycle control, the question as to whether R117 is methylated by PRMT1 in cell-free and *in vitro* contexts needs to be answered. Following this, testing whether inhibition of PRMT1 causes a similar phenotype to overexpression of the SPIN1-R117 mutant remains. Focussing on the P loop itself, phosphate-binding residues T120, S124 and T164 are modified by phosphorylation events. Mutation of these residues could be used to investigate whether their phosphorylation is permissive of repressive of R117 (and potentially R122 and R158) methylation. Mutation of T120 to an alanine is necessary to prevent S and G2 phase accumulation when SPIN1 is overexpressed in HeLa cells[111], suggesting that phosphorylation of these residues may have functional overlap with methylation of R117. This is further supported by HeLa cells responding to nucleolar stress in an analogous manner to MCF7 cells, circumventing G1 arrest caused by Rp-MDM2-p53 signalling and instead aggregating in S phase[175].

If any of these phosphorylation events encourage R117/122/158 methylation (or synergise functionally), then identification of their respective kinases would be of great use for therapeutic targeting in combination with PRMT1 inhibition **(Figure 4.30 B).** Phosphorylation of S124 (as well as 109) is administered by Aurora A kinase[125], although kinases that target the other phosphorylated residues in the P loop have yet to be identified. Additionally, if SPIN1 is found to interact with factors involved in the ribosomal stress response/replication stress via phosphorylation of these factors, then the kinase(s) responsible for this interaction could also be targeted alongside PRMT1.

Inhibition of PRMT1 and SPIN1, or PRMT1 and a kinase that facilitates P loop mediated activity could have pleiotropic anti-cancer effects. For example, targeting SPIN1 activity in a manner that potentially causes ribosomal stress could inhibit stemness without directly interfering with transcription of stem genes. This is implied by the finding that epithelial to mesenchymal transition (EMT, which is responsible for metastasis and de-differentiation) is positively regulated by ribosome biogenesis in MCF7 cells. EMT induction in these cells via hypoxia-induced Notch activation formed highly invasive basal-like tumours in mouse models. However, inhibition of ribosome biogenesis caused re-differentiation to a more benign state, coinciding with an increase in ERα expression[179]. Thus, confrontation of the CSC niche by forcing re-differentiation may be possible through ribosomal stress induction. In the PRMT1-SPIN1 axis scenario, this could be accomplished by inhibiting PRMT1-SPIN1 mediated transcriptional activation of rRNA genes (since SPIN1 transcriptionally activates rRNA genes[112]), or by interfering with the speculated P loop-mediated ribosome homeostasis **(Figure 4.30, A and B).**

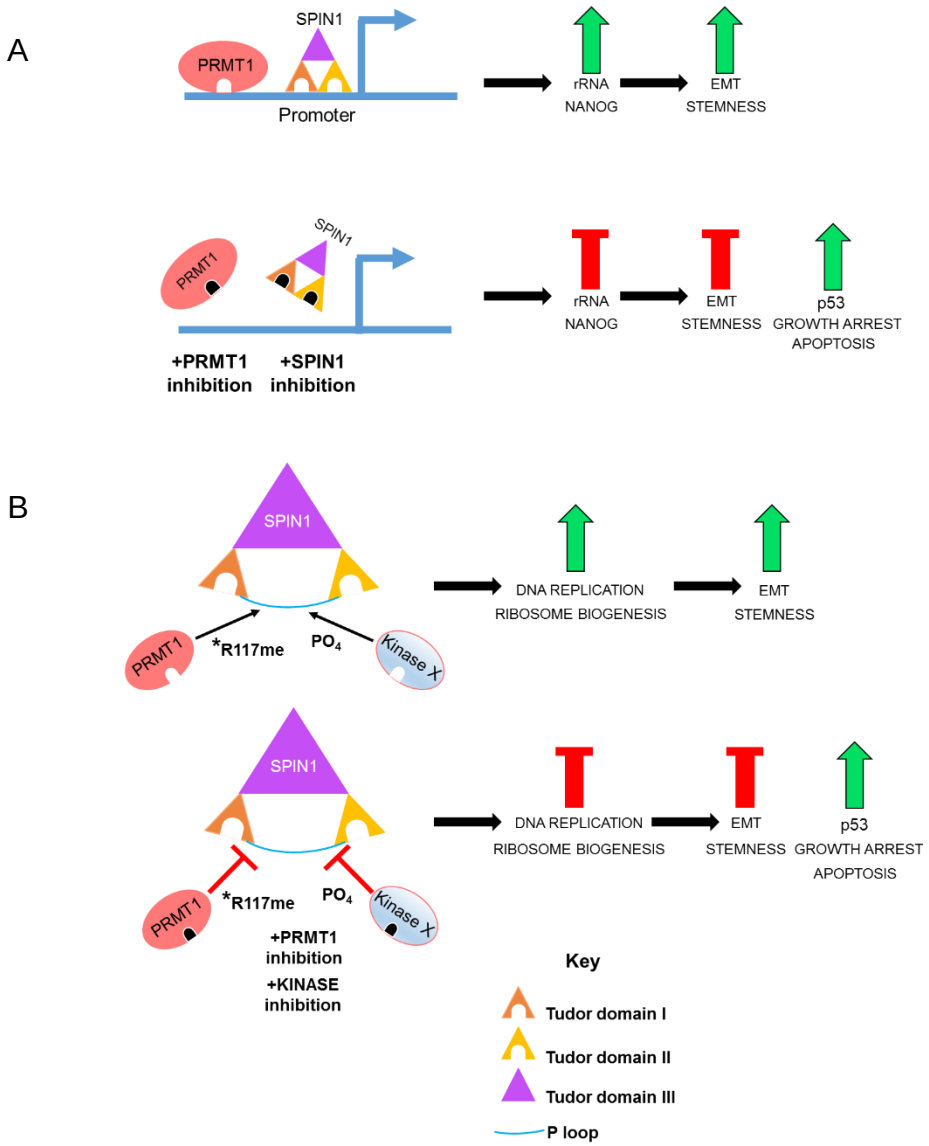


Figure 4.30: Speculative models for PRMT1 and SPIN1 involvement in the cancer phenotype, and therapeutic interventions

(A) Top panel: SPIN1 and PRMT1 co-localise to the promoter region of *NANOG* and rRNA genes providing transcriptional coactivator function. Lower panel: Inhibition of SPIN1 Tudor domains and PRMT1 methylation prevents SPIN1 histone code reader function and PRMT1-mediated deposition of H4R3me2a, downregulating transcription of ribosomal RNA and *NANOG*. **(B)** Top panel: PRMT1 and unknown kinase(s) methylate and phosphorylate the phosphate binding domain of SPIN1, respectively promoting DNA replication and/or ribosome biogenesis. Lower panel: inhibition of PRMT1 and kinase X inhibits SPIN1-mediated promotion of DNA replication and/or ribosome biogenesis.

Finally, it should be noted that SPIN1 and PRMT1 inhibition could have a drastic effect on expression of both genes. *PRMT1* transcription is up-regulated by SPIN1 overexpression, and constitutive knockdown of PRMT1 greatly reduces SPIN1 protein levels in MCF7 cells **(Figure 4.12)**. Thus, both genes appear to positively modulate one another's expression **(Figure 4.31)**. Whether PRMT1 positively regulates *SPIN1* transcription, or if it somehow promotes stability of the SPIN1 protein is not known, although both genes correlate positively at the protein level rather than the transcript level in breast cancer patient data **(Figure 4.3 B**), suggesting that the latter scenario may be true. Importantly, higher expression of SPIN1 in MCF7 cells (and an even higher level in triple negative breast cancer[132]) relative to MCF10A **(Figure 4.4)** could imply a greater dependency of breast cancer cells on SPIN1 activity relative to their healthy counterparts.

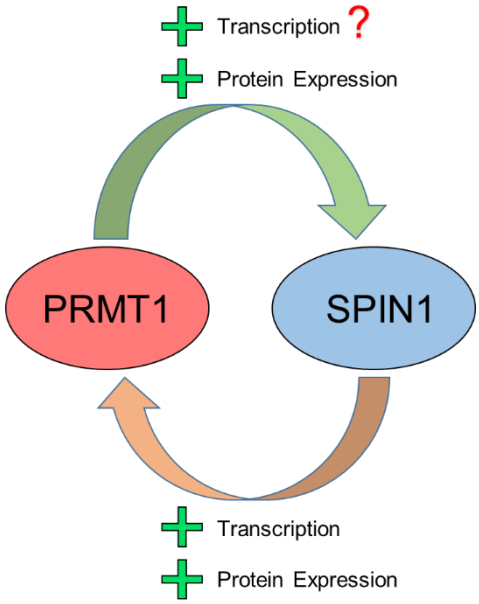


Figure 4.31: PRMT1 and SPIN1 positively regulate each other's expression in MCF7 cells

Diagram depicting mutual preservation of expression between SPIN1 and PRMT1. Transcriptional modulation of *SPIN1* by PRMT1 is marked with a question mark, pending further investigation.

4.10.6: SPIN1 as a profiling tool for proteomics

Given that SPIN1 dimerises *in vitro* **(Figure 1.7 B)**, the identification of ADMA proteins at 55kDa and 72kDa after FLAG-SPIN1 immunoprecipitation **(Figure 4.27 A)** were hypothesised to be dimeric and trimeric species, respectively. However, mass spectrometry analysis showed a negligible number of FLAG-SPIN1 peptides in all molecular weight categories outside of the dominant species that resolved at 36kDa **(Figure 4.27 C).** This implies that the ADMA bands corresponding to 55kDa and 72kDa were not SPIN1 oligomers. The observation that MS023 reduces the appearance of the 55 and 72kDa ADMA species **(Figure 4.28 A)** strongly suggests that SPIN1 bound additional non-histone but arginine-methylated proteins *in vitro.* This was further substantiated by abrogation of the 55kDa and 72kDa bands when Tudor domains I (Y98R) and II (F141A) were mutated **(Figure 4.28 B)** suggesting that asymmetric dimethyl-arginine and lysine trimethylation are required for binding of these species to SPIN1.

This observation may prove important in identifying the methylated proteins within the 55kDa and 72kDa bands, since it suggests their interaction with SPIN1 depends equally on the presence of methylated arginines and lysines. SPIN1 presumably interacts with these partners in the same bidentate configuration with which it binds H3K4me3/R8me2a. As such, candidate binding proteins likely contain closely-situated arginine and lysine residues within RG-dense regions, since the ADMA antibody is raised against methyl RG motifs. Overall, this may provide clues to the identity of these proteins from mass spectrometry data.

Of all the Tudor domain mutants tested, only the D184H mutant was capable of co-immunoprecipitating the 55 and 72kDa interacting proteins. This unique quality is likely a result of this residue forming a salt bridge contact with the R2 side chain of histone H3[47]. Thus, D184 does not contribute directly to the aromatic cage of SPIN1 Tudor domain II which directly contacts the trimethyl moiety of K4. As a result, this residue may be uniquely important in interaction with histone H3, and therefore ineffective at reducing binding of non-histone substrates. The slight increase in the methylation of the 55kDa and 72kDa signals could be testament to this, by reducing the amount of FLAG-SPIN1-D184H that would otherwise be committed to H3 binding; more FLAG-SPIN1-D184H could be available to bind non-histone substrate.

Using the information acquired from these experiments, the potential for SPIN1 to be used as a type I PRMT substrate discovery tool was briefly explored. Visual interpretation of the co-IP profile suggested obvious differences between the binding profile of wild-type SPIN1 and SPIN1-Y98R **(Figure 4.29)**. This result implied SPIN1 could be used as a substrate trap to affinity purify asymmetrically dimethylated proteins via its Tudor domain I, perhaps using the Y98R mutant as a negative control for non-specific interactors. In addition, SPIN1 could be used to discover targets of lysine methylation, given the even more striking change in banding pattern of co-immunoprecipitated proteins with FLAG-SPIN1-F141A compared to the wild type. Taken together, the data in section **4.9** suggest that the SPIN1 interactome far surpasses that of a histone H3 tail reader, potentially

widening the functional significance of this protein in normal and malignant cellular function.

Chapter 5: Summary and Final Conclusions

As the mechanisms of malignant growth continue to be revealed, we move ever closer to an age of anti-cancer medicine that is more effective, more specific and more personalised than ever before. Originally, transcription factors that act as highly influential nodes in oncogenic gene expression were considered ideal candidates for drug-based targets. However, hope for this methodology was turned on its head when it became clear that absence of a catalytic active site posed a significant challenge to development of small molecule inhibitors[180]. Thus, targeting oncogene expression at the transcriptional level in cancers returned to the realm of unfeasibility.

More recent advances in small molecule design have pioneered disruption of protein-protein interactions, or post-translational modifications that regulate transcription factor activity. For example, c-Myc, an often crucial oncogenic transcription factor could be targeted by interfering with deubiquitinating enzyme USP22 which maintains c-Myc levels[180].

During the interim of these developments, targeting epigenetic mediators of gene expression arose as a hopeful solution to the aforementioned problems. Since writers, readers and erasers of epigenetic information all have distinct binding pockets/catalytic sites that are more easily targetable; the prospect of targeting more direct players in transcription was once again opened. This has led to the development of small molecule inhibitors such as GSK2816126 and tazemezostat, which target activity of the lysine methyltransferase EZH2 and are currently undergoing Phase II clinical trials for the treatment of B cell lymphoma.

Closer to home for this thesis are clinical trials for epigenetic writers PRMT1 (GSK3368715) and PRMT5 (JNJ-64619178, PF-06939999) all of which are undergoing phase I clinical trials for a range of cancers.

Although the prospect of PRMT1 inhibition in cancer is already firmly in motion, more information on the PRMT1-mediated transcriptome would provide crucial insight into which genes would be affected by this treatment, and the mechanism of action. This could not only aid in eradication of targeted subsets of patient tumours, but also in the prediction of potentially harmful side-effects.

Knowledge of the PRMT1 signature H4R3me2a across the epigenome remains scarce, with ChIP-seq data still lacking. Despite requiring further optimisation, bio-orthogonal profiling detailed in this thesis could provide an answer to this absence where antibodies against H4R3me2a have yet to succeed.

Another gene that may hold great promise in epigenetic drug targeting is the histone code reader SPIN1. This gene has already been revealed as a key transcriptional orchestrator of numerous oncogenic programs in a plethora of cancers. The finding presented in this thesis that SPIN1 interacts with PRMT1 may infer their co-localisation to a subset of overlapping genes in breast cancer, perhaps placing them both in the same therapeutic crosshair.

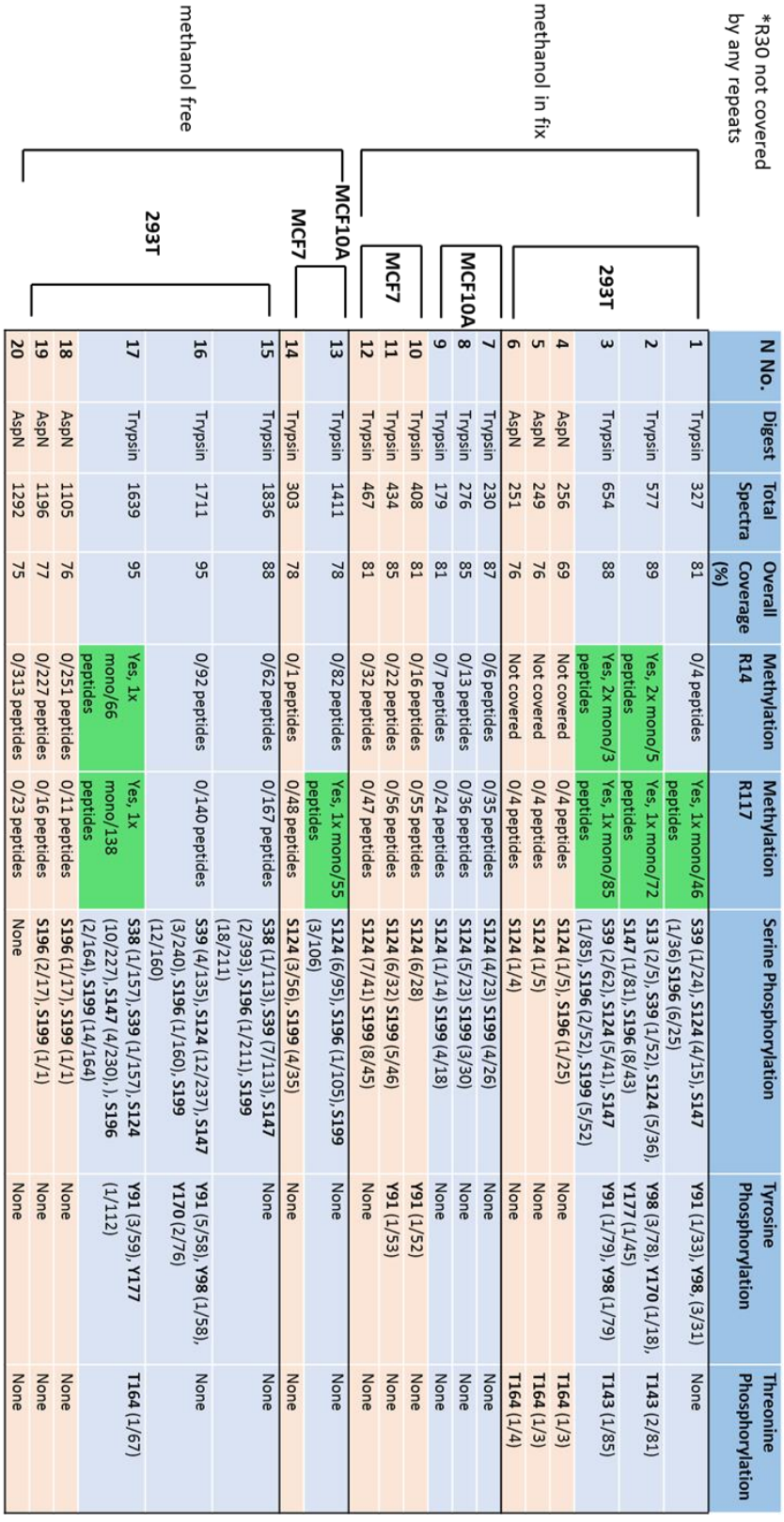
This relationship between PRMT1 and SPIN1 also appears to regulate critical cellular function outside of the realm of epigenetics, perhaps in DNA replication or ribosome biogenesis. If true, then targeting both genes simultaneously could interfere with several vital processes that promote malignancy, to devastating effect.

More work is required to prove these hypotheses, and thus, the utility of targeting both genes in breast cancer; however preliminary experimentation and the hard-found discoveries of others offer a glimmer of hope in pursuing PRMT1 and SPIN1 as bona-fide therapeutic targets in cancer.

Chapter 6: Appendices

6.1: Appendix A

Table summarising all SPIN1 PTMs discovered through mass spectrometry



*R30 not covered by any repeats

N No.		Digest	Total Spectra	Overall Coverage (%)	Methylation R14	Methylation R117	Serine Phosphorylation	Tyrosine Phosphorylation	Threonine Phosphorylation
methanol in fix	293T	1	Trypsin	327	81	0/4 peptides	Yes, 1x mono/46 peptides	S39 (1/24), S124 (4/15), S147 (1/36) S196 (6/25)	Y91 (1/33), Y98, (3/31)	None
		2	Trypsin	577	89	Yes, 2x mono/5 peptides	Yes, 1x mono/72 peptides	S13 (2/5), S39 (1/52), S124 (5/36), S147 (1/81), S196 (8/43)	Y98 (3/78), Y170 (1/18), Y177 (1/45)	T143 (2/81)
		3	Trypsin	654	88	Yes, 2x mono/3 peptides	Yes, 1x mono/85 peptides	S39 (2/62), S124 (5/41), S147 (1/85), S196 (2/52), S199 (5/52)	Y91 (1/79), Y98 (1/79)	T143 (1/85)
	MCF10A	4	AspN	256	69	Not covered	0/4 peptides	S124 (1/5), S196 (1/25)	None	T164 (1/3)
		5	AspN	249	76	Not covered	0/4 peptides	S124 (1/5)	None	T164 (1/3)
		6	AspN	251	76	Not covered	0/4 peptides	S124 (1/4)	None	T164 (1/4)
	MCF7	7	Trypsin	230	87	0/6 peptides	0/35 peptides	S124 (4/23) S199 (4/26)	None	None
		8	Trypsin	276	85	0/13 peptides	0/36 peptides	S124 (5/23) S199 (3/30)	None	None
		9	Trypsin	179	81	0/7 peptides	0/24 peptides	S124 (1/14) S199 (4/18)	None	None
	MCF10A	10	Trypsin	408	81	0/16 peptides	0/55 peptides	S124 (6/28)	Y91 (1/52)	None
		11	Trypsin	434	85	0/22 peptides	0/56 peptides	S124 (6/32) S199 (5/46)	Y91 (1/53)	None
		12	Trypsin	467	81	0/32 peptides	0/47 peptides	S124 (7/41) S199 (8/45)	None	None
methanol free	MCF7	13	Trypsin	1411	78	0/82 peptides	Yes, 1x mono/55 peptides	S124 (6/95), S196 (1/105), S199 (3/106)	None	None
		14	Trypsin	303	78	0/1 peptides	0/48 peptides	S124 (3/56), S199 (4/35)	None	None
		15	Trypsin	1836	88	0/62 peptides	0/167 peptides	S38 (1/113), S39 (7/113), S147 (2/393), S196 (1/211), S199 (18/211)	None	None
	293T	16	Trypsin	1711	95	0/92 peptides	0/140 peptides	S39 (4/135), S124 (12/237), S147 (3/240), S196 (1/160), S199 (12/160)	Y91 (5/58), Y98 (1/58), Y170 (2/76)	None
		17	Trypsin	1639	95	Yes, 1x mono/66 peptides	Yes, 1x mono/138 peptides	S38 (1/157), S39 (1/157), S124 (10/227), S147 (4/230),), S196 (2/164), S199 (14/164)	Y91 (3/59), Y177 (1/112)	T164 (1/67)
		18	AspN	1105	76	0/251 peptides	0/11 peptides	S196 (1/17), S199 (1/1)	None	None
		19	AspN	1196	77	0/227 peptides	0/16 peptides	S196 (2/17), S199 (1/1)	None	None
		20	AspN	1292	75	0/313 peptides	0/23 peptides	None	None	None

6.2: Appendix B

DAVID gene ontology analysis of co-immunoprecipitated proteins from HA-SPIN1 co-IP using HEK-293T cells.

~130kDa

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_BP_DIRECT	transcription, DNA-templated	RT		43	17.6	2.6E-3	1.2E-1
	GOTERM_BP_DIRECT	mRNA splicing, via spliceosome	RT		24	9.8	6.1E-14	3.6E-11
	GOTERM_BP_DIRECT	RNA processing	RT		23	9.4	6.5E-21	7.8E-18
	GOTERM_BP_DIRECT	cell-cell adhesion	RT		21	8.6	1.3E-9	5.2E-7
	GOTERM_BP_DIRECT	positive regulation of transcription from RNA polymerase II promoter	RT		21	8.6	5.5E-2	6.2E-1
	GOTERM_BP_DIRECT	mRNA processing	RT		15	6.1	2.1E-7	6.2E-5
	GOTERM_BP_DIRECT	viral process	RT		14	5.7	2.9E-4	2.4E-2
	GOTERM_BP_DIRECT	regulation of transcription from RNA polymerase II promoter	RT		14	5.7	8.8E-3	2.4E-1
	GOTERM_BP_DIRECT	cell division	RT		13	5.3	3.7E-3	1.4E-1
	GOTERM_BP_DIRECT	RNA splicing	RT		12	4.9	2.1E-5	3.1E-3

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_CC_DIRECT	nucleus	RT		126	49.8	7.4E-12	5.4E-10
	GOTERM_CC_DIRECT	cytoplasm	RT		122	48.2	1.8E-11	9.4E-10
	GOTERM_CC_DIRECT	nucleoplasm	RT		117	46.2	4.1E-32	1.5E-29
	GOTERM_CC_DIRECT	cytosol	RT		91	36.0	1.1E-11	6.7E-10
	GOTERM_CC_DIRECT	membrane	RT		84	33.2	5.8E-19	1.1E-16
	GOTERM_CC_DIRECT	extracellular exosome	RT		58	22.9	1.2E-3	1.8E-2
	GOTERM_CC_DIRECT	nucleolus	RT		34	13.4	7.6E-8	2.5E-6
	GOTERM_CC_DIRECT	intracellular ribonucleoprotein complex	RT		24	9.5	7.2E-19	8.8E-17
	GOTERM_CC_DIRECT	cell-cell adherens junction	RT		21	8.3	2.1E-8	7.7E-7
	GOTERM_CC_DIRECT	focal adhesion	RT		17	6.7	1.0E-4	2.2E-3

~ 100-120kDa

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_BP_DIRECT	transcription, DNA-templated	RT		54	14.2	4.2E-2	5.1E-1
	GOTERM_BP_DIRECT	mRNA splicing, via spliceosome	RT		36	9.5	1.4E-20	2.3E-17
	GOTERM_BP_DIRECT	positive regulation of transcription from RNA polymerase II promoter	RT		31	8.2	3.2E-2	4.4E-1
	GOTERM_BP_DIRECT	intracellular protein transport	RT		29	7.7	2.0E-13	1.1E-10
	GOTERM_BP_DIRECT	cell-cell adhesion	RT		28	7.4	3.4E-11	1.1E-8
	GOTERM_BP_DIRECT	viral process	RT		24	6.3	1.3E-7	2.7E-5
	GOTERM_BP_DIRECT	RNA processing	RT		21	5.5	1.5E-14	1.2E-11
	GOTERM_BP_DIRECT	mRNA processing	RT		19	5.0	6.0E-8	1.6E-5
	GOTERM_BP_DIRECT	cell division	RT		19	5.0	5.9E-4	2.8E-2
	GOTERM_BP_DIRECT	RNA splicing	RT		17	4.5	5.8E-7	9.5E-5

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_CC_DIRECT	nucleus	RT		192	50.7	4.2E-19	3.0E-17
	GOTERM_CC_DIRECT	cytoplasm	RT		178	47.0	3.6E-15	2.0E-13
	GOTERM_CC_DIRECT	nucleoplasm	RT		168	44.3	2.5E-43	1.1E-40
	GOTERM_CC_DIRECT	cytosol	RT		145	38.3	2.5E-21	2.1E-19
	GOTERM_CC_DIRECT	membrane	RT		142	37.5	1.8E-38	3.8E-36
	GOTERM_CC_DIRECT	extracellular exosome	RT		81	21.4	9.5E-4	1.3E-2
	GOTERM_CC_DIRECT	nucleolus	RT		68	17.9	7.1E-22	7.6E-20
	GOTERM_CC_DIRECT	mitochondrion	RT		41	10.8	8.6E-3	7.4E-2
	GOTERM_CC_DIRECT	intracellular ribonucleoprotein complex	RT		32	8.4	5.9E-24	8.5E-22
	GOTERM_CC_DIRECT	cell-cell adherens junction	RT		31	8.2	4.5E-12	2.2E-10

~ 90kDa

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_BP_DIRECT	transcription, DNA-templated	RT		38	14.5	7.4E-2	6.6E-1
	GOTERM_BP_DIRECT	cell-cell adhesion	RT		25	9.5	2.5E-12	1.6E-9
	GOTERM_BP_DIRECT	positive regulation of transcription from RNA polymerase II promoter	RT		24	9.2	2.1E-2	3.6E-1
	GOTERM_BP_DIRECT	mRNA splicing, via spliceosome	RT		22	8.4	1.8E-11	7.5E-9
	GOTERM_BP_DIRECT	viral process	RT		20	7.6	1.2E-7	1.9E-5
	GOTERM_BP_DIRECT	RNA processing	RT		17	6.5	9.8E-13	1.2E-9
	GOTERM_BP_DIRECT	ER to Golgi vesicle-mediated transport	RT		16	6.1	1.7E-8	3.6E-6
	GOTERM_BP_DIRECT	mRNA processing	RT		16	6.1	7.6E-8	1.4E-5
	GOTERM_BP_DIRECT	intracellular protein transport	RT		16	6.1	2.6E-6	3.0E-4
	GOTERM_BP_DIRECT	positive regulation of transcription, DNA-templated	RT		15	5.7	2.2E-2	3.7E-1

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_CC_DIRECT	cytoplasm	RT		135	51.5	9.6E-16	8.6E-14
	GOTERM_CC_DIRECT	nucleus	RT		129	49.2	6.0E-12	2.3E-10
	GOTERM_CC_DIRECT	membrane	RT		100	38.2	4.9E-28	1.7E-25
	GOTERM_CC_DIRECT	nucleoplasm	RT		98	37.4	3.8E-19	4.3E-17
	GOTERM_CC_DIRECT	cytosol	RT		98	37.4	6.4E-14	4.4E-12
	GOTERM_CC_DIRECT	extracellular exosome	RT		61	23.3	4.7E-4	6.7E-3
	GOTERM_CC_DIRECT	nucleolus	RT		45	17.2	6.7E-14	3.8E-12
	GOTERM_CC_DIRECT	mitochondrion	RT		29	11.1	2.1E-2	1.4E-1
	GOTERM_CC_DIRECT	cell-cell adherens junction	RT		27	10.3	7.4E-13	3.7E-11
	GOTERM_CC_DIRECT	intracellular ribonucleoprotein complex	RT		25	9.5	8.2E-20	1.4E-17

~ 65-70kDa

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_BP_DIRECT	mRNA splicing, via spliceosome	RT		32	12.3	1.1E-21	1.3E-18
	GOTERM_BP_DIRECT	cell-cell adhesion	RT		21	8.0	3.1E-9	7.4E-7
	GOTERM_BP_DIRECT	mRNA processing	RT		19	7.3	1.4E-10	5.4E-8
	GOTERM_BP_DIRECT	viral process	RT		17	6.5	8.6E-6	7.2E-4
	GOTERM_BP_DIRECT	RNA splicing	RT		16	6.1	2.1E-8	4.2E-6
	GOTERM_BP_DIRECT	protein folding	RT		15	5.7	4.1E-7	4.8E-5
	GOTERM_BP_DIRECT	negative regulation of apoptotic process	RT		15	5.7	7.1E-3	2.0E-1
	GOTERM_BP_DIRECT	rRNA processing	RT		14	5.4	1.6E-5	1.2E-3
	GOTERM_BP_DIRECT	gene expression	RT		13	5.0	3.1E-12	1.8E-9
	GOTERM_BP_DIRECT	intracellular protein transport	RT		13	5.0	1.9E-4	9.6E-3

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_CC_DIRECT	nucleus	RT		130	49.8	2.4E-12	1.2E-10
	GOTERM_CC_DIRECT	cytoplasm	RT		128	49.0	9.1E-13	5.3E-11
	GOTERM_CC_DIRECT	nucleoplasm	RT		109	41.8	1.5E-25	2.7E-23
	GOTERM_CC_DIRECT	cytosol	RT		91	34.9	6.1E-11	2.7E-9
	GOTERM_CC_DIRECT	membrane	RT		85	32.6	1.1E-18	1.2E-16
	GOTERM_CC_DIRECT	extracellular exosome	RT		57	21.8	3.7E-3	4.7E-2
	GOTERM_CC_DIRECT	nucleolus	RT		48	18.4	6.4E-16	4.7E-14
	GOTERM_CC_DIRECT	mitochondrion	RT		36	13.8	2.5E-4	4.5E-3
	GOTERM_CC_DIRECT	intracellular ribonucleoprotein complex	RT		32	12.3	6.5E-29	2.3E-26
	GOTERM_CC_DIRECT	myelin sheath	RT		24	9.2	1.7E-17	1.5E-15

~ 60kDa

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_BP_DIRECT	mRNA splicing, via spliceosome	RT		28	14.4	2.4E-20	2.3E-17
	GOTERM_BP_DIRECT	cell-cell adhesion	RT		16	8.2	4.7E-7	4.6E-5
	GOTERM_BP_DIRECT	mRNA processing	RT		15	7.7	1.6E-8	2.2E-6
	GOTERM_BP_DIRECT	protein folding	RT		15	7.7	1.7E-8	2.1E-6
	GOTERM_BP_DIRECT	negative regulation of apoptotic process	RT		12	6.2	1.5E-2	3.0E-1
	GOTERM_BP_DIRECT	RNA splicing	RT		11	5.6	2.0E-5	1.3E-3
	GOTERM_BP_DIRECT	viral process	RT		11	5.6	2.2E-3	6.7E-2
	GOTERM_BP_DIRECT	positive regulation of telomere maintenance via telomerase	RT		10	5.1	5.6E-11	1.1E-8
	GOTERM_BP_DIRECT	gene expression	RT		10	5.1	2.9E-9	4.6E-7
	GOTERM_BP_DIRECT	protein stabilization	RT		10	5.1	2.4E-5	1.5E-3

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_CC_DIRECT	cytoplasm	RT		113	57.9	2.8E-18	2.6E-16
	GOTERM_CC_DIRECT	nucleus	RT		106	54.4	2.4E-13	1.1E-11
	GOTERM_CC_DIRECT	nucleoplasm	RT		83	42.6	2.6E-20	3.6E-18
	GOTERM_CC_DIRECT	cytosol	RT		74	37.9	5.0E-11	1.2E-9
	GOTERM_CC_DIRECT	membrane	RT		70	35.9	4.3E-18	2.9E-16
	GOTERM_CC_DIRECT	extracellular exosome	RT		63	32.3	2.9E-9	6.0E-8
	GOTERM_CC_DIRECT	nucleolus	RT		35	17.9	1.8E-11	4.8E-10
	GOTERM_CC_DIRECT	intracellular ribonucleoprotein complex	RT		28	14.4	4.6E-27	1.2E-24
	GOTERM_CC_DIRECT	myelin sheath	RT		19	9.7	3.1E-14	1.7E-12
	GOTERM_CC_DIRECT	microtubule	RT		17	8.7	1.9E-7	3.2E-6

~ 55kDa

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_BP_DIRECT	transcription, DNA-templated	RT		30	16.4	3.4E-2	3.4E-1
	GOTERM_BP_DIRECT	mRNA splicing, via spliceosome	RT		26	14.2	6.6E-19	5.7E-16
	GOTERM_BP_DIRECT	mRNA processing	RT		15	8.2	6.0E-9	8.6E-7
	GOTERM_BP_DIRECT	cell-cell adhesion	RT		15	8.2	1.0E-6	8.1E-5
	GOTERM_BP_DIRECT	regulation of mRNA stability	RT		14	7.7	5.6E-11	1.6E-8
	GOTERM_BP_DIRECT	gene expression	RT		12	6.6	2.0E-12	8.4E-10
	GOTERM_BP_DIRECT	RNA processing	RT		12	6.6	5.6E-9	9.5E-7
	GOTERM_BP_DIRECT	RNA splicing	RT		11	6.0	1.0E-5	5.7E-4
	GOTERM_BP_DIRECT	protein folding	RT		10	5.5	1.2E-4	3.4E-3
	GOTERM_BP_DIRECT	protein polyubiquitination	RT		10	5.5	1.4E-4	3.9E-3

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_CC_DIRECT	nucleus	RT		118	64.5	7.3E-23	9.9E-21
	GOTERM_CC_DIRECT	cytoplasm	RT		105	57.4	2.9E-16	1.5E-14
	GOTERM_CC_DIRECT	nucleoplasm	RT		78	42.6	4.7E-19	3.2E-17
	GOTERM_CC_DIRECT	extracellular exosome	RT		76	41.5	1.4E-17	7.6E-16
	GOTERM_CC_DIRECT	cytosol	RT		72	39.3	1.6E-11	5.6E-10
	GOTERM_CC_DIRECT	membrane	RT		70	38.3	1.0E-19	9.3E-18
	GOTERM_CC_DIRECT	intracellular ribonucleoprotein complex	RT		27	14.8	2.5E-26	6.7E-24
	GOTERM_CC_DIRECT	nucleolus	RT		25	13.7	4.0E-6	7.2E-5
	GOTERM_CC_DIRECT	mitochondrion	RT		20	10.9	6.7E-2	3.1E-1
	GOTERM_CC_DIRECT	myelin sheath	RT		17	9.3	2.2E-12	8.6E-11

~ 40kDa

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_BP_DIRECT	mRNA splicing, via spliceosome	RT		39	16.5	8.6E-31	8.4E-28
	GOTERM_BP_DIRECT	cell-cell adhesion	RT		22	9.3	1.7E-10	3.4E-8
	GOTERM_BP_DIRECT	translation	RT		21	8.9	3.3E-10	5.4E-8
	GOTERM_BP_DIRECT	translational initiation	RT		19	8.1	5.3E-13	1.7E-10
	GOTERM_BP_DIRECT	gene expression	RT		17	7.2	1.5E-18	7.1E-16
	GOTERM_BP_DIRECT	RNA processing	RT		16	6.8	4.3E-12	1.1E-9
	GOTERM_BP_DIRECT	rRNA processing	RT		15	6.4	1.6E-6	1.1E-4
	GOTERM_BP_DIRECT	SRP-dependent cotranslational protein targeting to membrane	RT		14	5.9	5.1E-10	7.1E-8
	GOTERM_BP_DIRECT	viral transcription	RT		14	5.9	4.6E-9	5.6E-7
	GOTERM_BP_DIRECT	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	RT		14	5.9	9.7E-9	9.5E-7

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_CC_DIRECT	nucleus	RT		131	55.5	1.9E-17	1.2E-15
	GOTERM_CC_DIRECT	cytoplasm	RT		109	46.2	3.5E-9	7.7E-8
	GOTERM_CC_DIRECT	nucleoplasm	RT		104	44.1	9.7E-27	1.5E-24
	GOTERM_CC_DIRECT	cytosol	RT		97	41.1	5.9E-17	5.8E-15
	GOTERM_CC_DIRECT	extracellular exosome	RT		91	38.6	1.6E-18	1.7E-16
	GOTERM_CC_DIRECT	membrane	RT		67	28.4	9.4E-12	3.2E-10
	GOTERM_CC_DIRECT	intracellular ribonucleoprotein complex	RT		40	16.9	2.4E-42	7.4E-40
	GOTERM_CC_DIRECT	mitochondrion	RT		40	16.9	7.3E-7	1.3E-5
	GOTERM_CC_DIRECT	nucleolus	RT		38	16.1	5.7E-11	1.5E-9
	GOTERM_CC_DIRECT	focal adhesion	RT		26	11.0	3.3E-11	9.3E-10

~ 36kDa

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_BP_DIRECT	translation	RT		31	18.9	1.8E-24	1.3E-21
	GOTERM_BP_DIRECT	mRNA splicing, via spliceosome	RT		25	15.2	9.3E-19	3.5E-16
	GOTERM_BP_DIRECT	rRNA processing	RT		21	12.8	1.6E-14	1.7E-12
	GOTERM_BP_DIRECT	translational initiation	RT		19	11.6	8.0E-16	1.2E-13
	GOTERM_BP_DIRECT	SRP-dependent cotranslational protein targeting to membrane	RT		18	11.0	2.0E-17	5.0E-15
	GOTERM_BP_DIRECT	viral transcription	RT		18	11.0	4.4E-16	8.3E-14
	GOTERM_BP_DIRECT	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	RT		18	11.0	1.4E-15	1.7E-13
	GOTERM_BP_DIRECT	cell-cell adhesion	RT		16	9.8	4.9E-8	3.4E-6
	GOTERM_BP_DIRECT	gene expression	RT		11	6.7	2.2E-11	2.1E-9
	GOTERM_BP_DIRECT	mitochondrial translational elongation	RT		11	6.7	7.9E-9	6.5E-7

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_CC_DIRECT	nucleus	RT		98	59.8	1.3E-15	1.1E-13
	GOTERM_CC_DIRECT	cytoplasm	RT		87	53.0	5.0E-11	1.0E-9
	GOTERM_CC_DIRECT	extracellular exosome	RT		68	41.5	1.5E-15	8.8E-14
	GOTERM_CC_DIRECT	cytosol	RT		67	40.9	1.6E-11	3.5E-10
	GOTERM_CC_DIRECT	nucleoplasm	RT		66	40.2	1.4E-14	7.0E-13
	GOTERM_CC_DIRECT	membrane	RT		56	34.1	2.6E-13	8.9E-12
	GOTERM_CC_DIRECT	mitochondrion	RT		35	21.3	1.7E-8	2.6E-7
	GOTERM_CC_DIRECT	nucleolus	RT		30	18.3	4.7E-10	8.9E-9
	GOTERM_CC_DIRECT	ribosome	RT		26	15.9	8.0E-24	2.0E-21
	GOTERM_CC_DIRECT	mitochondrial inner membrane	RT		25	15.2	1.2E-12	3.2E-11

Chapter 7: References

1. Zhang, M., Lee, A. V & Rosen, J. M. The cellular origin and evolution of breast cancer. *Cold Spring Harb. Perspect. Med.* **7,** (2017).

2. Allred, D. C., Wu, Y., Mao, S., Nagtegaal, I. D., Lee, S., Perou, C. M., Mohsin, S. K., O'Connell, P., Tsimelzon, A. & Medina, D. Ductal carcinoma in situ and the emergence of diversity during breast cancer evolution. *Clin. Cancer Res.* **14,** 370–8 (2008).

3. Liu, Y. R., Jiang, Y. Z., Xu, X. E., Yu, K. Da, Jin, X., Hu, X., Zuo, W. J., Hao, S., Wu, J., Liu, G. Y., Di, G. H., Li, D. Q., He, X. H., Hu, W. G. & Shao, Z. M. Comprehensive transcriptome analysis identifies novel molecular subtypes and subtype-specific RNAs of triple-negative breast cancer. *Breast Cancer Res.* **18,** 33 (2016).

4. Peddi, P. F., Ellis, M. J. & Ma, C. Molecular basis of triple negative breast cancer and implications for therapy. *Int. J. Breast Cancer* **2012,** 217185 (2012).

5. Kubo, M. Adjuvant endocrine treatment for estrogen receptor (ER)-positive/HER2-negative breast cancer. *Chinese Clin. Oncol.* **9,** 33–33 (2020).

6. Jelovac, D. & Wolff, A. C. The adjuvant treatment of HER2-positive breast cancer. *Curr. Treat. Options Oncol.* **13,** 230–239 (2012).

7. Bergin, A. R. T. & Loi, S. Triple-negative breast cancer: Recent treatment advances. *F1000Research* **8,** (2019).

8. Tripsianes, K., Madl, T., MacHyna, M., Fessas, D., Englbrecht, C., Fischer, U., Neugebauer, K. M. & Sattler, M. Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. *Nat. Struct. Mol. Biol.* **18,** 1414–1420 (2011).

9. Brostoff, S. & Eylar, E. H. Localization of methylated arginine in the A1 protein from myelin. *Proc. Natl. Acad. Sci. U. S. A.* **68,** 765–769 (1971).

10. Baldwin, G. S. & Carnegie, P. R. Specific enzymic methylation of an arginine in the experimental allergic encephalomyelitis protein from human myelin. *Science (80-.).* **171,** 579–581 (1971).

11. Lin, W. J., Gary, J. D., Yang, M. C., Clarke, S. & Herschman, H. R. 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 (1996).

12. Boffa, L. C., Karn, J., Vidali, G. & Allfrey, V. G. Distribution of NG, NG-dimethylarginine in nuclear protein fractions. *Biochem. Biophys. Res. Commun.* **74,** 969–976 (1977).

13. Kölbel, K., Ihling, C., Bellmann-Sickert, K., Neundorf, I., Beck-Sickinger, A. G., Sinz, A., Kühn, U. & Wahle, E. Type I arginine methyltransferases PRMT1 and PRMT-3 act distributively. *J. Biol. Chem.* **284,** 8274–8282 (2009).

14. Gui, S., Wooderchak-Donahue, W. L., Zang, T., Chen, D., Daly, M. P.,

Zhou, Z. S. & Hevel, J. M. Substrate-induced control of product formation by protein arginine methyltransferase 1. *Biochemistry* **52,** 199–209 (2013).

15. Sun, L., Wang, M., Lv, Z., Yang, N., Liu, Y., Bao, S., Gong, W. & Xu, R. M. Structural insights into protein arginine symmetric dimethylation by PRMT5. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 20538–20543 (2011).

16. 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. PRMT9 is a Type II methyltransferase that methylates the splicing factor SAP145. *Nat. Commun.* **6,** 1–12 (2015).

17. Zurita-Lopez, C. I., Sandberg, T., Kelly, R. & Clarke, S. G. Human protein arginine methyltransferase 7 (PRMT7) is a type III enzyme forming ω-N G-monomethylated arginine residues. *J. Biol. Chem.* **287,** 7859–7870 (2012).

18. Tewary, S. K., Zheng, Y. G. & Ho, M. C. Protein arginine methyltransferases: insights into the enzyme structure and mechanism at the atomic level. *Cell. Mol. Life Sci.* **76,** 2917–2932 (2019).

19. Hasegawa, M., Toma-Fukai, S., Kim, J. D., Fukamizu, A. & Shimizu, T. Protein arginine methyltransferase 7 has a novel homodimer-like structure formed by tandem repeats. *FEBS Lett.* **588,** 1942–1948 (2014).

20. Hadjikyriacou, A., Yang, Y., Espejo, A., Bedford, M. T. & Clarke, S. G. Unique features of human protein arginine methyltransferase 9 (PRMT9) and its substrate RNA splicing factor SF3B2. *J. Biol. Chem.* **290,** 16723–16743 (2015).

21. Fulton, M. D., Brown, T. & George Zheng, Y. The biological axis of protein arginine methylation and asymmetric dimethylarginine. *Int. J. Mol. Sci.* **20,** 1–18 (2019).

22. Thandapani, P., O'Connor, T. R., Bailey, T. L. & Richard, S. Defining the RGG/RG Motif. *Mol. Cell* **50,** 613–623 (2013).

23. Yamagata, K., Daitoku, H., Takahashi, Y., Namiki, K., Hisatake, K., Kako, K., Mukai, H., Kasuya, Y. & Fukamizu, A. Arginine Methylation of FOXO Transcription Factors Inhibits Their Phosphorylation by Akt. *Mol. Cell* **32,** 221–231 (2008).

24. Lee, J. & Bedford, M. T. PABP1 identified as an arginine methyltransferase substrate using high-density protein arrays. *EMBO Rep.* **3,** 268–273 (2002).

25. Cheng, D., Côté, J., Shaaban, S. & Bedford, M. T. The Arginine Methyltransferase CARM1 Regulates the Coupling of Transcription and mRNA Processing. *Mol. Cell* **25,** 71–83 (2007).

26. Musiani, D., Bok, J., Massignani, E., Wu, L., Tabaglio, T., Ippolito, M. R., Cuomo, A., Ozbek, U., Zorgati, H., Ghoshdastider, U., Robinson, R. C., Guccione, E. & Bonaldi, T. Proteomics profiling of arginine methylation defines PRMT5 substrate specificity. *Sci. Signal.* **12,** 8388 (2019).

27. Kuhn, P., Chumanov, R., Wang, Y., Ge, Y., Burgess, R. R. & Xu, W. Automethylation of CARM1 allows coupling of transcription and mRNA splicing. *Nucleic Acids Res.* **39,** 2717–2726 (2011).

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

29. Dillon, M. B. C., Rust, H. L., Thompson, P. R. & Mowen, K. A. Automethylation of protein arginine methyltransferase 8 (PRMT8) regulates activity by impeding S-adenosylmethionine sensitivity. *J. Biol. Chem.* **288,** 27872–27880 (2013).

30. Feng, Q., He, B., Jung, S. Y., Song, Y., Qin, J., Tsai, S. Y., Tsai, M. J. & O'Malley, B. W. Biochemical control of CARM1 enzymatic activity by phosphorylation. *J. Biol. Chem.* **284,** 36167–36174 (2009).

31. Higashimoto, K., Kuhn, P., Desai, D., Cheng, X. & Xu, W. Phosphorylation-mediated inactivation of coactivator-associated arginine methyltransferase 1. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 12318–12323 (2007).

32. Liu, F., Zhao, X., Perna, F., Wang, L., Koppikar, P., Abdel-Wahab, O., Harr, M. W., Levine, R. L., Xu, H., Tefferi, A., Deblasio, A., Hatlen, M., Menendez, S. & Nimer, S. D. JAK2V617F-Mediated Phosphorylation of PRMT5 Downregulates Its Methyltransferase Activity and Promotes Myeloproliferation. *Cancer Cell* **19,** 283–294 (2011).

33. Lattouf, H., Kassem, L., Jacquemetton, J., Choucair, A., Poulard, C., Trédan, O., Corbo, L., Diab-Assaf, M., Hussein, N., Treilleux, I. & Le Romancer, M. LKB1 regulates PRMT5 activity in breast cancer. *Int. J. Cancer* **144,** 595–606 (2019).

34. Bao, X., Siprashvili, Z., Zarnegar, B. J., Shenoy, R. M., Rios, E. J., Nady, N., Qu, K., Mah, A., Webster, D. E., Rubin, A. J., Wozniak, G. G., Tao, S., Wysocka, J. & Khavari, P. A. CSNK1a1 Regulates PRMT1 to Maintain the Progenitor State in Self-Renewing Somatic Tissue. *Dev. Cell* **43,** 227-239.e5 (2017).

35. Zhang, H. T., Zeng, L. F., He, Q. Y., Tao, W. A., Zha, Z. G. & Hu, C. D. The E3 ubiquitin ligase CHIP mediates ubiquitination and proteasomal degradation of PRMT5. *Biochim. Biophys. Acta - Mol. Cell Res.* **1863,** 335–346 (2016).

36. Bhuripanyo, K., Wang, Y., Liu, X., Zhou, L., Liu, R., Duong, D., Zhao, B., Bi, Y., Zhou, H., Chen, G., Seyfried, N. T., Chazin, W. J., Kiyokawa, H. & Yin, J. Identifying the substrate proteins of U-box E3s E4B and CHIP by orthogonal ubiquitin transfer. *Sci. Adv.* **4,** (2018).

37. Li, B., Liu, L., Li, X. & Wu, L. MiR-503 suppresses metastasis of hepatocellular carcinoma cell by targeting PRMT1. *Biochem. Biophys. Res. Commun.* **464,** 982–987 (2015).

38. Lu, Y. F., Cai, X. L., Li, Z. Z., Lv, J., Xiang, Y. A., Chen, J. J., Chen, W. J., Sun, W. Y., Liu, X. M. & Chen, J. B. LncRNA SNHG16 Functions as an Oncogene by Sponging MiR-4518 and Up-Regulating PRMT5 Expression in Glioma. *Cell. Physiol. Biochem.* **45,** 1975–1985 (2018).

39. Zhang, H., Feng, X., Wang, T., Hu, Z., Que, X., Tian, Q., Zhu, T., Guo, G., Huang, W., Li, X. & Guo, X. MiRNA-543 promotes osteosarcoma cell proliferation and glycolysis by partially suppressing PRMT9 and stabilizing HIF-1α protein. *Oncotarget* **8,** 2342–2355 (2017).

40. Lee, J., Sayegh, J., Daniel, J., Clarke, S. & Bedford, M. T. PRMT8, a new membrane-bound tissue-specific member of the protein arginine methyltransferase family. *J. Biol. Chem.* **280,** 32890–32896 (2005).

41. Rodríguez-Gil, A., Ritter, O., Saul, V. V., Wilhelm, J., Yang, C. Y., Grosschedl, R., Imai, Y., Kuba, K., Kracht, M. & Lienhard Schmitz, M. The CCR4-NOT complex contributes to repression of Major Histocompatibility Complex class II transcription. *Sci. Rep.* **7,** 1–15 (2017).

42. Robin-Lespinasse, Y., Sentis, S., Kolytcheff, C., Rostan, M. C., Corbo, L. & Le Romancer, M. hCAF1, a new regulator of PRMT1-dependent arginine methylation. *J. Cell Sci.* **120,** 638–647 (2007).

43. Benhenda, S., Ducroux, A., Riviere, L., Sobhian, B., Ward, M. D., Dion, S., Hantz, O., Protzer, U., Michel, M.-L., Benkirane, M., Semmes, O. J., Buendia, M.-A. & Neuveut, C. Methyltransferase PRMT1 Is a Binding Partner of HBx and a Negative Regulator of Hepatitis B Virus Transcription. *J. Virol.* **87,** 4360–4371 (2013).

44. Guderian, G., Peter, C., Wiesner, J., Sickmann, A., Schulze-Osthoff, K., Fischer, U. & Grimmler, M. RioK1, a new interactor of protein arginine methyltransferase 5 (PRMT5), competes with pICln for binding and modulates PRMT5 complex composition and substrate specificity. *J. Biol. Chem.* **286,** 1976–1986 (2011).

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

46. Liu, K., Guo, Y., Liu, H., Bian, C., Lam, R., Liu, Y., Mackenzie, F., Rojas, L. A., Reinberg, D., Bedford, M. T., Xu, R.-M. & Min, J. Crystal structure of TDRD3 and methyl-arginine binding characterization of TDRD3, SMN and SPF30. *PLoS One* **7,** e30375 (2012).

47. Su, X., Zhu, G., Ding, X., Lee, S. Y., Dou, Y., Zhu, B., Wu, W. & Li, H. Molecular basis underlying histone H3 lysine-arginine methylation pattern readout by Spin/Ssty repeats of Spindlin1. *Genes Dev.* **28,** 622–636 (2014).

48. Yang, N., Wang, W., Wang, Y., Wang, M., Zhao, Q., Rao, Z., Zhu, B. & Xu, R.-M. Distinct mode of methylated lysine-4 of histone H3 recognition by tandem tudor-like domains of Spindlin1. *Proc. Natl. Acad. Sci.* **109,** 17954–17959 (2012).

49. Migliori, V., Müller, J., Phalke, S., Low, D., Bezzi, M., Mok, W. C., Sahu, S. K., Gunaratne, J., Capasso, P., Bassi, C., Cecatiello, V., De Marco, A., Blackstock, W., Kuznetsov, V., Amati, B., Mapelli, M. & Guccione, E. Symmetric dimethylation of H3R2 is a newly identified histone mark that supports euchromatin maintenance. *Nat. Struct. Mol. Biol.* **19,** 136–145 (2012).

50. Lorton, B. M. & Shechter, D. Cellular consequences of arginine methylation. *Cell. Mol. Life Sci.* **76,** (2019).

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

52. Poulard, C., Rambaud, J., Hussein, N., Corbo, L. & Le Romancer, M. JMJD6 regulates ERα methylation on arginine. *PLoS One* **9,** 87982 (2014).

53. Shen, C., Quan, Q., Yang, C., Wen, Y. & Li, H. Histone demethylase JMJD6 regulates cellular migration and proliferation in adipose-derived mesenchymal stem cells. *Stem Cell Res. Ther.* **9,** 212 (2018).

54. Bottger, A., Islam, S., Wolf, Rasheduzzaman Chowdhury, R., Schofield, C. J. & Wolf, A. The oxygenase Jmjd6–a case study in conflicting assignments. *Biochem. J.* **468,** 191–202 (2015).

55. Cuthbert, G. L., Daujat, S., Snowden, A. W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P. D., Tempst, P., Bannister, A. J. & Kouzarides, T. Histone deimination antagonizes arginine methylation. *Cell* **118,** 545–553 (2004).

56. Tang, J., Frankel, A., Cook, R. J., Kim, S., Paik, W. K., Williams, K. R., Clarke, S. & Herschman, H. R. PRMT1 is the predominant type I protein arginine methyltransferase in mammalian cells. *J. Biol. Chem.* **275,** 7723–30 (2000).

57. Goulet, I., Gauvin, G., Boisvenue, S. & Côté, J. Alternative splicing yields protein arginine methyltransferase 1 isoforms with distinct activity, substrate specificity, and subcellular localization. *J. Biol. Chem.* **282,** 33009–21 (2007).

58. Patounas, O., Papacharalampous, I., Eckerich, C., Markopoulos, G. S., Kolettas, E. & 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.* **119,** 2110–2123 (2018).

59. Baldwin, R. M., Morettin, A. & Côté, J. Role of PRMTs in cancer: Could minor isoforms be leaving a mark? *World J. Biol. Chem.* **5,** 115–29 (2014).

60. Bao, J., Di Lorenzo, A., Lin, K., Lu, Y., Zhong, Y., Sebastian, M. M., Muller, W. J., Yang, Y. & Bedford, M. T. Mouse models of overexpression reveal distinct oncogenic roles for different type I protein arginine

methyltransferases. *Cancer Res.* **79,** canres.1995.2018 (2018).

61. Seligson, D. B., Horvath, S., Shi, T., Yu, H., Tze, S., Grunstein, M. & Kurdistani, S. K. Global histone modification patterns predict risk of prostate cancer recurrence. *Nat. Lett.* **435,** 1262–1266 (2005).

62. Mathioudaki, K., Papadokostopoulou, A., Scorilas, A., Xynopoulos, D., Agnanti, N. & Talieri, M. The PRMT1 gene expression pattern in colon cancer. *Br. J. Cancer* **99,** 2094–2099 (2008).

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

64. Waddington, C. H. The epigenotype. 1942. *Int. J. Epidemiol.* **41,** 10–13 (2012).

65. Flavahan, W. A., Gaskell, E. & Bernstein, B. E. Epigenetic plasticity and the hallmarks of cancer. *Science (80-.).* **357,** (2017).

66. Gulati, N., Béguelin, W. & Giulino-Roth, L. Enhancer of zeste homolog 2 (EZH2) inhibitors. *Leuk. Lymphoma* **59,** 1574–1585 (2018).

67. Chatterjee, B., Ghosh, K. & Kanade, S. R. Resveratrol modulates epigenetic regulators of promoter histone methylation and acetylation that restores BRCA1, p53, p21CIP1 in human breast cancer cell lines. *BioFactors* **45,** 818–829 (2019).

68. 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. Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science (80-.).* **293,** 853–857 (2001).

69. Huang, S., Litt, M. & Felsenfeld, G. Methylation of histone H4 by arginine methyltransferase PRMT1 is essential in vivo for many subsequent histone modifications. *Genes Dev.* **19,** 1885–93 (2005).

70. Li, X., Hu, X., Patel, B., Zhou, Z., Liang, S., Ybarra, R., Qiu, Y., Felsenfeld, G., Bungert, J. & Huang, S. H4R3 methylation facilitates beta-globin transcription by regulating histone acetyltransferase binding and H3 acetylation. *Blood* **115,** 2028–37 (2010).

71. Yang, Y., Lu, Y., Espejo, A., Wu, J., Xu, W., Liang, S. & Bedford, M. T. TDRD3 Is an Effector Molecule for Arginine-Methylated Histone Marks. *Mol. Cell* **40,** 1016–1023 (2010).

72. Yang, Y., McBride, K. M., Hensley, S., Lu, Y., Chedin, F. & Bedford, M. T. Arginine Methylation Facilitates the Recruitment of TOP3B to Chromatin to Prevent R Loop Accumulation. *Mol. Cell* **53,** 484–497 (2014).

73. Rakow, S., Pullamsetti, S. S., Bauer, U. M. & Bouchard, C. Assaying epigenome functions of PRMTs and their substrates. *Methods* **175,** 53–65 (2020).

74. Nichols, R. C., Wang, X. W., Tang, J., Hamilton, B. J., High, F. A.,

Herschman, H. R. & Rigby, W. F. C. The RGG domain in hnRNP A2 affects subcellular localization. *Exp. Cell Res.* **256,** 522–532 (2000).

75. Côté, J., Boisvert, F. M., Boulanger, M. C., Bedford, M. T. & Richard, S. Sam68 RNA binding protein is an in vivo substrate for protein arginine N-methyltransferase 1. *Mol. Biol. Cell* **14,** 274–287 (2003).

76. Wall, M. L. & Lewis, S. M. Methylarginines within the RGG-Motif Region of hnRNP A1 Affect Its IRES Trans-Acting Factor Activity and Are Required for hnRNP A1 Stress Granule Localization and Formation. *J. Mol. Biol.* **429,** 295–307 (2017).

77. Murata, K., Lu, W., Hashimoto, M., Ono, N., Muratani, M., Nishikata, K., Kim, J. D., Ebihara, S., Ishida, J. & Fukamizu, A. PRMT1 Deficiency in Mouse Juvenile Heart Induces Dilated Cardiomyopathy and Reveals Cryptic Alternative Splicing Products. *iScience* **8,** 200–213 (2018).

78. Zhang, L., Tran, N. T., Su, H., Wang, R., Lu, Y., Tang, H., Aoyagi, S., Guo, A., Khodadadi-Jamayran, A., Zhou, D., Qian, K., Hricik, T., Côté, J., Han, X., Zhou, W., Laha, S., Abdel-Wahab, O., Levine, R. L., Raffel, G., *et al.* Cross-talk between PRMT1-mediated methylation and ubiquitylation on RBM15 controls RNA splicing. *Elife* **4,** (2015).

79. Fong, J. Y., Pignata, L., Goy, P. A., Kawabata, K. C., Lee, S. C. W., Koh, C. M., Musiani, D., Massignani, E., Kotini, A. G., Penson, A., Wun, C. M., Shen, Y., Schwarz, M., Low, D. H., Rialdi, A., Ki, M., Wollmann, H., Mzoughi, S., Gay, F., *et al.* Therapeutic Targeting of RNA Splicing Catalysis through Inhibition of Protein Arginine Methylation. *Cancer Cell* **36,** 194-209.e9 (2019).

80. Bommer, G. T., Gerin, I., Feng, Y., Kaczorowski, A. J., Kuick, R., Love, R. E., Zhai, Y., Giordano, T. J., Qin, Z. S., Moore, B. B., MacDougald, O. A., Cho, K. R. & Fearon, E. R. p53-Mediated Activation of miRNA34 Candidate Tumor-Suppressor Genes. *Curr. Biol.* **17,** 1298–1307 (2007).

81. Treré, D., Ceccarelli, C., Montanaro, L., Tosti, E. & Derenzini, M. Nucleolar size and activity are related to pRb and p53 status in human breast cancer. *J. Histochem. Cytochem.* **52,** 1601–1607 (2004).

82. Hsu, J. H. R., Hubbell-Engler, B., Adelmant, G., Huang, J., Joyce, C. E., Vazquez, F., Weir, B. A., Montgomery, P., Tsherniak, A., Giacomelli, A. O., Perry, J. A., Trowbridge, J., Fujiwara, Y., Cowley, G. S., Xie, H., Kim, W., Novina, C. D., Hahn, W. C., Marto, J. A., *et al.* PRMT1-mediated translation regulation is a crucial vulnerability of cancer. *Cancer Res.* **77,** 4613–4625 (2017).

83. Infantino, S., Benz, B., Waldmann, T., Jung, M., Schneider, R. & Reth, M. Arginine methylation of the B cell antigen receptor promotes differentiation. *J. Exp. Med.* **207,** 711–719 (2010).

84. Blanc, R. S., Vogel, G., Li, X., Yu, Z., Li, S. & Richard, S. Arginine Methylation by PRMT1 Regulates Muscle Stem Cell Fate. *Mol. Cell. Biol.* **37,** (2017).

85. 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., *et al.* PRMT1 and PRMT8 regulate retinoic acid-dependent neuronal differentiation with implications to neuropathology. *Stem Cells* **33,** 726–741 (2015).

86. Zuo, Z., Yang, G., Wang, H., Zhang, Y., Cai, Y., Chen, F., Xiao, Y., Cheng, M., Huang, Y. & Zhang, Y. Klf4 methylated by Prmt1 is required for lineage segregation of epiblast and primitive endoderm. *bioRxiv* 2020.04.24.059055 (2020). doi:10.1101/2020.04.24.059055

87. Zhang, L., Tran, N. T., Su, H., Wang, R., Lu, Y., Tang, H., Aoyagi, S., Guo, A., Khodadadi-Jamayran, A., Zhou, D., Qian, K., Hricik, T., Côté, J., Han, X., Zhou, W., Laha, S., Abdel-Wahab, O., Levine, R. L., Raffel, G., *et al.* Cross-talk between PRMT1-mediated methylation and ubiquitylation on RBM15 controls RNA splicing. *Elife* **4,** 7938 (2015).

88. Zhao, Y., Lu, Q., Li, C., Wang, X., Jiang, L., Huang, L., Wang, C. & Chen, H. PRMT1 regulates the tumour-initiating properties of esophageal squamous cell carcinoma through histone H4 arginine methylation coupled with transcriptional activation. *Cell Death Dis.* **10,** 1–17 (2019).

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

90. Nilsson, S., Mäkelä, S., Treuter, E., Tujague, M., Thomsen, J., Andersson, G., Enmark, E., Pettersson, K., Warner, M. & Gustafsson, J. Å. Mechanisms of estrogen action. *Physiol. Rev.* **81,** 1535–1565 (2001).

91. Yersal, O. & Barutca, S. Biological subtypes of breast cancer: Prognostic and therapeutic implications. *World J. Clin. Oncol.* **5,** 412–424 (2014).

92. Le Romancer, M., Treilleux, I., Leconte, N., Robin-Lespinasse, Y., Sentis, S., Bouchekioua-Bouzaghou, K., Goddard, S., Gobert-Gosse, S. & Corbo, L. Regulation of Estrogen Rapid Signaling through Arginine Methylation by PRMT1. *Mol. Cell* **31,** 212–221 (2008).

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

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

95. Montenegro, M. F., González-Guerrero, R., Sánchez-del-Campo, L., Piñero-Madrona, A., Cabezas-Herrera, J. & Rodríguez-López, J. N. PRMT1-dependent methylation of BRCA1 contributes to the epigenetic defense of breast cancer cells against ionizing radiation. *Sci. Rep.* **10,** 1–

14 (2020).

96. Bremang, M., Cuomo, A., Agresta, A. M., Stugiewicz, M., Spadotto, V. & Bonaldi, T. Mass spectrometry-based identification and characterisation of lysine and arginine methylation in the human proteome. *Mol. Biosyst.* **9,** 2231–2247 (2013).

97. 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. Mammalian protein arginine methyltransferase 7 (PRMT7) specifically targets RXR sites in lysine- and arginine-rich regions. *J. Biol. Chem.* **288,** 37010–37025 (2013).

98. Machida, K., Thompson, C. M., Dierck, K., Jablonowski, K., Kärkkäinen, S., Liu, B., Zhang, H., Nash, P. D., Newman, D. K., Nollau, P., Pawson, T., Renkema, G. H., Saksela, K., Schiller, M. R., Shin, D. G. & Mayer, B. J. High-Throughput Phosphotyrosine Profiling Using SH2 Domains. *Mol. Cell* **26,** 899–915 (2007).

99. Goulet, I., Boisvenue, S., Mokas, S., Mazroui, R. & Côté, J. TDRD3, a novel Tudor domain-containing protein, localizes to cytoplasmic stress granules. *Hum. Mol. Genet.* **17,** 3055–3074 (2008).

100. Tripsianes, K., Madl, T., MacHyna, M., Fessas, D., Englbrecht, C., Fischer, U., Neugebauer, K. M. & Sattler, M. Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. *Nat. Struct. Mol. Biol.* **18,** 1414–1420 (2011).

101. Lin, Q., Jiang, F., Schultz, P. G. & Gray, N. S. Design of allele-specific protein methyltransferase inhibitors. *J. Am. Chem. Soc.* **123,** 11608–11613 (2001).

102. Bishop, A. C., Buzko, O. & Shokat, K. M. Magic bullets for protein kinases. *Trends Cell Biol.* **11,** 167–172 (2001).

103. Lukinavičius, G., Lapiene, V., Staševskij, Z., Dalhoff, C., Weinhold, E. & Klimašauskas, S. Targeted labeling of DNA by methyltransferase-directed transfer of activated groups (mTAG). *J. Am. Chem. Soc.* **129,** 2758–2759 (2007).

104. Peters, W., Willnow, S., Duisken, M., Kleine, H., Macherey, T., Duncan, K. E., Litchfield, D. W., Lüscher, B. & Weinhold, E. Enzymatic site-specific functionalization of protein methyltransferase substrates with alkynes for click labeling. *Angew. Chemie - Int. Ed.* **49,** 5170–5173 (2010).

105. Wang, R., Ibáñez, G., Islam, K., Zheng, W., Blum, G., Sengelaub, C. & Luo, M. Formulating a fluorogenic assay to evaluate S-adenosyl-L-methionine analogues as protein methyltransferase cofactors. *Mol. Biosyst.* **7,** 2970–2981 (2011).

106. Wang, R., Zheng, W., Yu, H., Deng, H. & Luo, M. Labeling substrates of protein arginine methyltransferase with engineered enzymes and matched S-adenosyl-L-methionine analogues. *J. Am. Chem. Soc.* **133,** 7648–51 (2011).

107. Guo, H., Wang, R., Zheng, W., Chen, Y., Blum, G., Deng, H. & Luo, M. Profiling substrates of protein arginine N-methyltransferase 3 with S-adenosyl-L-methionine analogues. *ACS Chem. Biol.* **9,** 476–84 (2014).

108. Wang, R., Islam, K., Liu, Y., Zheng, W., Tang, H., Lailler, N., Blum, G., Deng, H. & Luo, M. Profiling genome-wide chromatin methylation with engineered posttranslation apparatus within living cells. *J. Am. Chem. Soc.* **135,** 1048–56 (2013).

109. Waldmann, T., Izzo, A., Kamieniarz, K., Richter, F., Vogler, C., Sarg, B., Lindner, H., Young, N. L., Mittler, G., Garcia, B. A. & Schneider, R. Methylation of H2AR29 is a novel repressive PRMT6 target. *Epigenetics and Chromatin* **4,** 11 (2011).

110. Oh, B., Hwang, S. Y., Solter, D. & Knowles, B. B. Spindlin, a major maternal transcript expressed in the mouse during the transition from oocyte to embryo. *Development* **124,** 493–503 (1997).

111. Zhao, Q., Qin, L., Jiang, F., Wu, B., Yue, W., Xu, F., Rong, Z., Yuan, H., Xie, X., Gao, Y., Bai, C., Bartlam, M., Pei, X. & Rao, Z. Structure of human spindlin1: Tandem tudor-like domains for cell cycle regulation. *J. Biol. Chem.* **282,** 647–656 (2007).

112. Zhang, X., Zhu, G., Su, X., Li, H. & Wu, W. Nucleolar localization signal and histone methylation reader function is required for SPIN1 to promote rRNA gene expression. *Biochem. Biophys. Res. Commun.* **505,** 325–332 (2018).

113. Chew, T. G., Peaston, A., Lim, A. K., Lorthongpanich, C., Knowles, B. B. & Solter, D. A Tudor Domain Protein SPINDLIN1 Interacts with the mRNA-Binding Protein SERBP1 and Is Involved in Mouse Oocyte Meiotic Resumption. *PLoS One* **8,** e69764 (2013).

114. Choi, J. W., Zhou, W., Nie, Z. W., Niu, Y. J., Shin, K. T. & Cui, X. S. Spindlin1 alters the metaphase to anaphase transition in meiosis I through regulation of BUB3 expression in porcine oocytes. *J. Cell. Physiol.* **234,** 8963–8974 (2019).

115. Choi, J.-W., Zhao, M.-H., Liang, S., Guo, J., Lin, Z.-L., Li, Y.-H., Jo, Y.-J., Kim, N.-H. & Cui, X.-S. Spindlin 1 is essential for metaphase II stage maintenance and chromosomal stability in porcine oocytes. *Mol. Hum. Reprod.* **23,** 166–176 (2017).

116. Greschik, H., Duteil, D., Messaddeq, N., Willmann, D., Arrigoni, L., Sum, M., Jung, M., Metzger, D., Manke, T., Günther, T. & Schüle, R. The histone code reader Spin1 controls skeletal muscle development. *Cell Death Dis.* **8,** e3173 (2017).

117. Wang, W., Chen, Z., Mao, Z., Zhang, H., Ding, X., Chen, S., Zhang, X., Xu, R. & Zhu, B. Nucleolar protein Spindlin1 recognizes H3K4 methylation and stimulates the expression of rRNA genes. *EMBO Rep.* **12,** 1160–6 (2011).

118. Wang, C., Zhan, L., Wu, M., Ma, R., Yao, J., Xiong, Y., Pan, Y., Guan, S.,

Zhang, X. & Zang, J. Spindlin-1 recognizes methylations of K20 and R23 of histone H4 tail. *FEBS Lett.* **592,** 4098–4110 (2018).

119. Ducroux, A., Benhenda, S., Rivière, L., Semmes, O. J., Benkirane, M. & Neuveut, C. The Tudor Domain Protein Spindlin1 Is Involved in Intrinsic Antiviral Defense against Incoming Hepatitis B Virus and Herpes Simplex Virus Type 1. *PLoS Pathog.* **10,** (2014).

120. Bae, N., Gao, M., Li, X., Premkumar, T., Sbardella, G., Chen, J. & Bedford, X. M. T. A transcriptional coregulator, SPIN-DOC, attenuates the coactivator activity of Spindlin1. *J. Biol. Chem.* **292,** 20808–20817 (2017).

121. Yue, W., Sun, L.-Y., Li, C.-H., Zhang, L.-X. & Pei, X.-T. Screening and identification of ovarian carcinomas related genes. *Aizheng Chinese J. cancer* **23,** 141–5 (2004).

122. Gao, Y., Yue, W., Zhang, P., Li, L., Xie, X., Yuan, H., Chen, L., Liu, D., Yan, F. & Pei, X. Spindlin1, a novel nuclear protein with a role in the transformation of NIH3T3 cells. *Biochem. Biophys. Res. Commun.* **335,** 343–350 (2005).

123. Zhang, P., Cong, B., Yuan, H., Chen, L., Lv, Y., Bai, C., Nan, X., Shi, S., Yue, W. & Pei, X. Overexpression of spindlin1 induces metaphase arrest and chromosomal instability. *J. Cell. Physiol.* **217,** 400–408 (2008).

124. Yuan, H., Zhang, P., Qin, L., Chen, L., Shi, S., Lu, Y., Yan, F., Bai, C., Nan, X., Liu, D., Li, Y., Yue, W. & Pei, X. Overexpression of SPINDLIN1 induces cellular senescence, multinucleation and apoptosis. *Gene* **410,** 67–74 (2008).

125. Wang, J.-X., Zeng, Q., Chen, L., Du, J.-C., Yan, X.-L., Yuan, H.-F., Zhai, C., Zhou, J.-N., Jia, Y.-L., Yue, W. & Pei, X.-T. SPINDLIN1 Promotes Cancer Cell Proliferation through Activation of WNT/TCF-4 Signaling. *Mol. Cancer Res.* **10,** 326–335 (2012).

126. Pitt, S. C. & Chen, H. The phosphatidylinositol 3-kinase/akt signaling pathway in medullary thyroid cancer. *Surgery* **144file:*//*,** 721–724 (2008).

127. Franz, H., Greschik, H., Willmann, D., Ozretić, L., Annette Jilg, C., Wardelmann, E., Jung, M., Buettner, R. & Schüle, R. The histone code reader SPIN1 controls RET signaling in liposarcoma. *Oncotarget* **6,** 4773–4789 (2015).

128. Li, Y., Ma, X., Wang, Y. & Li, G. miR-489 inhibits proliferation, cell cycle progression and induces apoptosis of glioma cells via targeting SPIN1-mediated PI3K/AKT pathway. *Biomed. Pharmacother.* **93,** 435–443 (2017).

129. Janecki, D. M., Sajek, M., Smialek, M. J., Kotecki, M., Ginter-Matuszewska, B., Kuczynska, B., Spik, A., Kolanowski, T., Kitazawa, R., Kurpisz, M. & Jaruzelska, J. SPIN1 is a proto-oncogene and SPIN3 is a tumor suppressor in human seminoma. *Oncotarget* **9,** 32466–32477 (2018).

130. Chen, X., Dong, H., Liu, S., Yu, L., Yan, D., Yao, X., Sun, W., Han, D. &

Gao, G. *Long noncoding RNA MHENCR promotes melanoma progression via regulating miR-425/489-mediated PI3K-Akt pathway*. *Am. J. Transl. Res.* **9,** (2017).

131. Chen, X., Wang, Y. W., Xing, A. Y., Xiang, S., Shi, D. B., Liu, L., Li, Y. X. & Gao, P. Suppression of SPIN1-mediated PI3K–Akt pathway by miR-489 increases chemosensitivity in breast cancer. *J. Pathol.* **239,** 459–472 (2016).

132. Chen, X., Wang, Y. W. & Gao, P. SPIN1, negatively regulated by miR-148/152, enhances Adriamycin resistance via upregulating drug metabolizing enzymes and transporter in breast cancer. *J. Exp. Clin. Cancer Res.* **37,** (2018).

133. Chen, W., Zhang, Y., Wang, H., Pan, T., Zhang, Y. & Li, C. LINC00473/miR-374a-5p regulates esophageal squamous cell carcinoma via targeting SPIN1 to weaken the effect of radiotherapy. *J. Cell. Biochem.* **120,** 14562–14572 (2019).

134. Bae, N., Viviano, M., Su, X., Lv, J., Cheng, D., Sagum, C., Castellano, S., Bai, X., Johnson, C., Khalil, M. I., Shen, J., Chen, K., Li, H., Sbardella, G. & Bedford, M. T. Developing Spindlin1 small-molecule inhibitors by using protein microarrays. *Nat. Chem. Biol.* **13,** 750–756 (2017).

135. Fagan, V., Johansson, C., Gileadi, C., Monteiro, O., Dunford, J. E., Nibhani, R., Philpott, M., Malzahn, J., Wells, G., Faram, R., Cribbs, A. P., Halidi, N., Li, F., Chau, I., Greschik, H., Velupillai, S., Allali-Hassani, A., Bennett, J., Christott, T., *et al.* A Chemical Probe for Tudor Domain Protein Spindlin1 to Investigate Chromatin Function. *J. Med. Chem.* **62,** 9008–9025 (2019).

136. Xiong, Y., Greschik, H., Johansson, C., Seifert, L., Bacher, J., Park, K. S., Babault, N., Martini, M., Fagan, V., Li, F., Chau, I., Christott, T., Dilworth, D., Barsyte-Lovejoy, D., Vedadi, M., Arrowsmith, C. H., Brennan, P., Fedorov, O., Jung, M., *et al.* Discovery of a Potent and Selective Fragment-like Inhibitor of Methyllysine Reader Protein Spindlin 1 (SPIN1). *J. Med. Chem.* **62,** 8996–9007 (2019).

137. Drago-Ferrante, R., Pentimalli, F., Carlisi, D., De Blasio, A., Saliba, C., Baldacchino, S., Degaetano, J., Debono, J., Caruana-Dingli, G., Grech, G., Scerri, C., Tesoriere, G., Giordano, A., Vento, R. & Di Fiore, R. Suppressive role exerted by microRNA-29b-1-5p in triple negative breast cancer through SPIN1 regulation. *Oncotarget* **8,** 28939–28958 (2017).

138. Du, H.-Y. & Liu, B. MiR-1271 as a tumor suppressor in breast cancer proliferation and progression via targeting SPIN1. *Eur. Rev. Med. Pharmacol. Sci.* **22,** 2697–2706 (2018).

139. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25,** 402–8 (2001).

140. Chen, G., Liu, H., Wang, X. & Li, Z. In vitro methylation by methanol: Proteomic screening and prevalence investigation. *Anal. Chim. Acta* **661,**

67–75 (2010).

141. Gao, Y., Zhao, Y., Zhang, J., Lu, Y., Liu, X., Geng, P., Huang, B., Zhang, Y., Lu, J., Fan, L., Hay, E. D., Kalluri, R., Weinberg, R. A., Bonnomet, A., Trimboli, A. J., Sarrio, D., Prat, A., Davis, F. M., Stewart, T. A., *et al.* 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 (2016).

142. Wang, R. & Luo, M. A journey toward Bioorthogonal Profiling of Protein Methylation inside living cells. *Curr. Opin. Chem. Biol.* **17,** 729–37 (2013).

143. Lakowski, T. M., Zurita-Lopez, C., Clarke, S. G. & Frankel, A. Approaches to measuring the activities of protein arginine N-methyltransferases. *Anal. Biochem.* **397,** 1–11 (2010).

144. Balint, B. L., Szanto, A., Madi, A., Bauer, U.-M., Gabor, P., Benko, S., Puskás, L. G., Davies, P. J. A. & Nagy, L. Arginine Methylation Provides Epigenetic Transcription Memory for Retinoid-Induced Differentiation in Myeloid Cells. *Mol. Cell. Biol.* **25,** 5648–5663 (2005).

145. Maldonado, L. Y., Arsene, D., Mato, J. M. & Lu, S. C. Methionine adenosyltransferases in cancers: Mechanisms of dysregulation and implications for therapy. *Exp. Biol. Med.* **243,** 107–117 (2018).

146. Pérez, C., Pérez-Zúñiga, F. J., Garrido, F., Reytor, E., Portillo, F. & Pajares, M. A. The oncogene PDRG1 is an interaction target of methionine adenosyltransferases. *PLoS One* **11,** 1–29 (2016).

147. Chen, X., Wei, S., Ji, Y., Guo, X. & Yang, F. Quantitative proteomics using SILAC: Principles, applications, and developments. *Proteomics* **15,** 3175–3192 (2015).

148. Foss, S., Watkinson, R., Sandlie, I., James, L. C. & Andersen, J. T. TRIM21: A cytosolic Fc receptor with broad antibody isotype specificity. *Immunol. Rev.* **268,** 328–339 (2015).

149. Jobert, L., Argentini, M. & Tora, L. PRMT1 mediated methylation of TAF15 is required for its positive gene regulatory function. *Exp. Cell Res.* **315,** 1273–1286 (2009).

150. Gullà, A., Hideshima, T., Bianchi, G., Fulciniti, M., Kemal Samur, M., Qi, J., Tai, Y. T., Harada, T., Morelli, E., Amodio, N., Carrasco, R., Tagliaferri, P., Munshi, N. C., Tassone, P. & Anderson, K. C. Protein arginine methyltransferase 5 has prognostic relevance and is a druggable target in multiple myeloma. *Leukemia* **32,** 996–1002 (2018).

151. Kim, S. J., Yoo, B. C., Uhm, C. S. & Lee, S. W. Posttranslational arginine methylation of lamin A/C during myoblast fusion. *Biochim. Biophys. Acta - Proteins Proteomics* **1814,** 308–317 (2011).

152. Rouillard, A. D., Gundersen, G. W., Fernandez, N. F., Wang, Z., Monteiro, C. D., McDermott, M. G. & Ma'ayan, A. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database (Oxford).* **2016,** 1–16 (2016).

153. Clarke, T. L., Sanchez-Bailon, M. P., Chiang, K., Reynolds, J. J., Herrero-Ruiz, J., Bandeiras, T. M., Matias, P. M., Maslen, S. L., Skehel, J. M., Stewart, G. S. & Davies, C. C. PRMT5-Dependent Methylation of the TIP60 Coactivator RUVBL1 Is a Key Regulator of Homologous Recombination. *Mol. Cell* **65,** 900-916.e7 (2017).

154. Wooderchak, W. L., Zang, T., Zhou, Z. S., Acuña, M., Tahara, S. M. & Hevel, J. M. Substrate Profiling of PRMT1 Reveals Amino Acid Sequences That Extend Beyond the RGG Paradigm. *Biochemistry* **47,** 9456–9466 (2008).

155. Eram, M. S., Shen, Y., Szewczyk, M. M., Wu, H., Senisterra, G., Li, F., Butler, K. V, Kaniskan, H. Ü., Speed, B. A., dela Seña, C., Dong, A., Zeng, H., Schapira, M., Brown, P. J., Arrowsmith, C. H., Barsyte-Lovejoy, D., Liu, J., Vedadi, M. & Jin, J. A Potent, Selective, and Cell-Active Inhibitor of Human Type I Protein Arginine Methyltransferases. *ACS Chem. Biol.* **11,** acschembio.5b00839 (2015).

156. Bedford, M. T. & Clarke, S. G. Protein Arginine Methylation in Mammals: Who, What, and Why. *Mol. Cell* **33,** 1–13 (2009).

157. Chan, S., Au, K., Francis, R. S., Mudge, D. W., Johnson, D. W. & Pillans, P. I. Phosphate binders in patients with chronic kidney disease. *Aust. Prescr.* **40,** 9–14 (2017).

158. Hepburn, A. C., Steele, R. E., Veeratterapillay, R., Wilson, L., Kounatidou, E. E., Barnard, A., Berry, P., Cassidy, J. R., Moad, M., El-Sherif, A., Gaughan, L., Mills, I. G., Robson, C. N. & Heer, R. The induction of core pluripotency master regulators in cancers defines poor clinical outcomes and treatment resistance. *Oncogene* **38,** 4412–4424 (2019).

159. Liu, L. M., Sun, W. Z., Fan, X. Z., Xu, Y. L., Cheng, M. Bin & Zhang, Y. Methylation of C/EBPa by PRMT1 inhibits its tumor-suppressive function in breast cancer. *Cancer Res.* **79,** 2865–2877 (2019).

160. Rust, H. L. & Thompson, P. R. Kinase consensus sequences: A breeding ground for crosstalk. *ACS Chem. Biol.* **6,** 881–892 (2011).

161. Kowenz-Leutz, E., Pless, O., Dittmar, G., Knoblich, M. & Leutz, A. Crosstalk between C/EBPB phosphorylation, arginine methylation, and SWI/SNF/Mediator implies an indexing transcription factor code. *EMBO J.* **29,** 1105–1115 (2010).

162. Uversky, V. N. Intrinsically disordered proteins and their ‘Mysterious’ (meta)physics. *Front. Phys.* **7,** 10 (2019).

163. Shepard, P. J. & Hertel, K. J. The SR protein family. *Genome Biol.* **10,** 242 (2009).

164. Shen, H., Kan, J. L. C. & Green, M. R. Arginine-Serine-Rich Domains Bound at Splicing Enhancers Contact the Branchpoint to Promote Prespliceosome Assembly. *Mol. Cell* **13,** 367–376 (2004).

165. Chiang, K., Zielinska, A. E., Shaaban, A. M., Sanchez-Bailon, M. P., Jarrold, J., Clarke, T. L., Zhang, J., Francis, A., Jones, L. J., Smith, S.,

Barbash, O., Guccione, E., Farnie, G., Smalley, M. J. & Davies, C. C. PRMT5 Is a Critical Regulator of Breast Cancer Stem Cell Function via Histone Methylation and FOXP1 Expression. *Cell Rep.* **21,** 3498–3513 (2017).

166. Barr, A. R., Cooper, S., Heldt, F. S., Butera, F., Stoy, H., Mansfeld, J., Novák, B. & Bakal, C. DNA damage during S-phase mediates the proliferation-quiescence decision in the subsequent G1 via p21 expression. *Nat. Commun.* **8,** 1–17 (2017).

167. Patel, D., Menon, D., Bernfeld, E., Mroz, V., Kalan, S., Loayza, D. & Foster, D. A. Aspartate rescues S-phase arrest caused by suppression of glutamine utilization in KRas-driven cancer cells. *J. Biol. Chem.* **291,** 9322–9329 (2016).

168. Ye, X., Franco, A. A., Santos, H., Nelson, D. M., Kaufman, P. D. & Adams, P. D. Defective S phase chromatin assembly causes DNA damage, activation of the S phase checkpoint, and S phase arrest. *Mol. Cell* **11,** 341–351 (2003).

169. Nabatiyan, A. & Krude, T. Silencing of Chromatin Assembly Factor 1 in Human Cells Leads to Cell Death and Loss of Chromatin Assembly during DNA Synthesis. *Mol. Cell. Biol.* **24,** 2853–2862 (2004).

170. Yang, Y., Mcbride, K. M., Hensley, S., Lu, Y., Chedin, F. & Bedford, M. T. Arginine methylation facilitates the recruitment of TOP3B to chromatin to prevent R-loop accumulation. *Mol. Cell* **6,** 484–497 (2014).

171. Fei, L. & Xu, H. Role of MCM2-7 protein phosphorylation in human cancer cells. *Cell Biosci.* **8,** 43 (2018).

172. Fang, Z., Cao, B., Liao, J. M., Deng, J., Plummer, K. D., Liao, P., Liu, T., Zhang, W., Zhang, K., Li, L., Margolin, D., Zeng, S. X., Xiong, J. & Lu, H. SPIN1 promotes tumorigenesis by blocking the uL18 (Universal large ribosomal subunit protein 18)-MDM2-p53 pathway in human cancer. *Elife* **7,** 1–21 (2018).

173. Liu, Y., Deisenroth, C. & Zhang, Y. RP-MDM2-p53 Pathway: Linking Ribosomal Biogenesis and Tumor Surveillance. *Trends in Cancer* **2,** 191–204 (2016).

174. Latonen, L. Phase-to-phase with nucleoli - Stress responses, protein aggregation and novel roles of RNA. *Front. Cell. Neurosci.* **13,** (2019).

175. Dong, Z., Zhu, C., Zhan, Q. & Jiang, W. The roles of RRP15 in nucleolar formation, ribosome biogenesis and checkpoint control in human cells. *Oncotarget* **8,** 13240–13252 (2017).

176. Potapova, T. A. & Gerton, J. L. Ribosomal DNA and the nucleolus in the context of genome organization. *Chromosom. Res.* **27,** 109–127 (2019).

177. Maiser, A., Dillinger, S., Längst, G., Schermelleh, L., Leonhardt, H. & Németh, A. Super-resolution in situ analysis of active ribosomal DNA chromatin organization in the nucleolus. *Sci. Rep.* **10,** 1–11 (2020).

178. Xiong, Y., Greschik, H., Johansson, C., Seifert, L., Bacher, J., Park, K. S., Babault, N., Martini, M., Fagan, V., Li, F., Chau, I., Christott, T., Dilworth, D., Barsyte-Lovejoy, D., Vedadi, M., Arrowsmith, C. H., Brennan, P., Fedorov, O., Jung, M., *et al.* Discovery of a Potent and Selective Fragment-like Inhibitor of Methyllysine Reader Protein Spindlin 1 (SPIN1). *J. Med. Chem.* **62,** 8996–9007 (2019).

179. Prakash, V., Carson, B. B., Feenstra, J. M., Dass, R. A., Sekyrova, P., Hoshino, A., Petersen, J., Guo, Y., Parks, M. M., Kurylo, C. M., Batchelder, J. E., Haller, K., Hashimoto, A., Rundqivst, H., Condeelis, J. S., Allis, C. D., Drygin, D., Nieto, M. A., Andäng, M., *et al.* Ribosome biogenesis during cell cycle arrest fuels EMT in development and disease. *Nat. Commun.* **10,** (2019).

180. Bushweller, J. H. Targeting transcription factors in cancer — from undruggable to reality. *Nat. Rev. Cancer* **19,** 611–624 (2019).