Arginine Methylation on E2F1



Yi-Chien Lu

Linacre College

A thesis submitted for the degree of
Doctor of Philosophy
Trinity 2014

Arginine methylation on E2F1

Yi-Chien Lu — Department of Oncology
Linacre College — University of Oxford
A thesis submitted for the degree of Doctor of Philosophy — *Trinity 2014*

Abstract

E2F1 is a transcription factor which paradoxically has major influence on both apoptosis and cell cycle progression. One of the most important questions in E2F1 biology therefore is the mechanism underlying regulation of these opposing physiological outcomes.

Post-translational modifications (PTM) provide proteins with an additional layer of complexity, potentially altering interactions with partner DNA and protein. The importance of arginine methylation has recently been implicated in modulating the activity of the tumour suppression pathway proteins, p53 and E2F1. Previous studies have established that the methyltransferase, PRMT5, is responsible for the symmetrical dimethylation of E2F1, which inhibits its pro-apoptotic activity.

In this thesis, E2F1 was found to be a substrate of PRMT1, which catalysed **a**symmetrical dimethylation of E2F1 at arginine 109. In addition, a positive correlation was found between the percentage of apoptotic cells and levels of PRMT1. Conversely, an increase in cancer cell colony formation was shown when the site of PRMT1 methylation on E2F1 was changed from arginine to lysine at position 109. These findings suggested a growth inhibition effect by PRMT1 methylation on E2F1.

At the transcriptional level, depletion of PRMT1 increased E2F1 binding to the promoter region of Cdc6, a cell cycle regulator, and decreased binding to the promoter region of Apaf1, which has a pro- apoptotic role. Genome-wide ChIP-sequencing technology was undertaken and results further clarified that the depletion of PRMT1 preferably enriched E2F1 binding to promoters of positive regulators of cell proliferation and promoters of the cell cycle.

Collectively, the findings of this thesis suggested that the opposing roles E2F1 demonstrated in promoting both cell proliferation and apoptosis was due to different types of arginine methylation which trigger E2F1 binding to different promoters.

Lastly, arginine methylation was shown to influence protein-protein interactions. PRMT5 induction resulted in the identification by mass spectrometry of β-catenin as an E2F1 interacting partner. As the Wnt/ β-catenin signalling pathway is broadly recognised as having pro- cell proliferation activity, this finding is consistent with previous reports that suggest the oncogenic role PRMT5 methylation has on E2F1.

Acknowledgement

It has been an amazing journey to be able to come abroad to live and to learn science in Oxford. None would be possible without my parents' full support. I am very fortunate to have a pair of wonderful parents and a very close family. I thank Minister of Education of Taiwanese government for the study abroad studentship and also Oxford Oncology Department for the stipend, both have provided me with the most substantial help.

The person I want to thank most for my DPhil study is Prof Nick La Thangue. He is a pioneer in E2F biology and a very accomplished scientist. It's an honour to be able to learn from him. I cherished and enjoyed all our meetings together for those are the time I learn to form insights quickly from a pile of data. He is also a very experienced and amazing supervisor. He is always there in the office/lab and is always eager to discuss results. For a very busy professor like he is, he has helped me at all stages in my DPhil including many corrections on the transfer report, a mock viva which really helped me. For the preparation of DPhil thesis, he has used his private free time to read the thesis and almost gave me immediate feedbacks chapter by chapter which really helped to make this happen. More importantly, he has helped me in sorting out more funding so that I can continue to live and learn in Oxford.

I have received a lot of helps from a lab of very capable scientists. In particular, I want to thank Dr Amanda Coutts, Dr Shonagh Munro and Dr Panos Zalmas. It is amazing to discuss and enjoy science together. I learnt a great deal from them and truly enjoyed spending time with them inside or outside of the lab. All La Thangue group previous and current members play important roles in supporting me in my study and in my days in Oxford.

For this study, I am very fortunate to have many collaboration opportunities for different expertise. I have learnt a great deal a lot from Prof Benedikt Kessler's proteomics group and special thanks to Dr Rebecca Konietzny and Dr Joanna McGouran. I really had fun learning and discussing in Kessler's lab and also made a lot of friends. The other collaboration is done with Wellcome Trust Genomics group and I thank the project manager, Dr Lorna Witty. On this account, I need to delicate my special thanks to Dr Robert Siddaway from Ludwig Institute for Cancer Research who I met in the bioinformatics workshop and ended up being an important help for the bioinformatics analysis. Both Robert and Panos are my teachers and I enjoy learning from them immensely.

Finally, I thank all my friends who have shared my ups and downs in all these years, including those who set up Oxbridge Biotech Roundtable together and those in the Taiwanese Student Society. Lastly, I have most sincere gratitude to Mr William Stokes and Dr Charles Sande, Dr Tania Garcia and Ms Kylie Hassall. They have been there from the start to the end as my UK family, my mentors and my dearest friends.

Table of Contents

Chapter 1 Introduction....11
Ch1.1 Cancer....12
Ch1.1-1 Social and Economic Impact of Cancer....12
Ch1.1-2 Carcinogenesis hypotheses....14
Ch1.2 E2F1 and E2F family....21
Ch1.3 Cell Cycle....24
Ch1.3-1 Cell cycle cyclins and CDKs control....24
Ch1.3-2 Cell cycle checkpoints controls and DNA damage....27
Ch1.4 E2F1 and pocket protein in cell cycle regulation....28
Ch1.5 E2F1 in tumour progression and suppression....29
Ch1.6 Post-Translational Modifications and E2F1 regulation....31
Ch1.7 PRMT family and arginine methylation....35
Ch1.9 Research Objective....40
Chapter 2 Materials and Methods....41
Ch2.1 Cell Lines and Tissue Culture....42
Ch2.1-1 Human cancer cell lines....42
Ch2.1-2 Generation of PRMT1 and PRMT5 inducible cell lines....42
Ch2.1-3 Generation of stably expressed E2F1 cells....43
Ch2.2 Site-directed mutagenesis....44
Ch2.3 DNA plasmids transfection....44
Ch2.4 small interfering RNA (siRNA)....47
Ch2.5 Recombinant protein purification....48
Ch2.6 *In vitro* ^{3}H-SAM Methylation Assay....49
Ch2.7 SDS-PAGE and Immunoblotting....50
Ch2.8 Antibodies....52
Ch2.9 Immunoprecipitation....52
Ch2.10 Mass Spectrum sample preparation and analysis....53
Ch2.10-1 in solution digestion....54
Ch2.10-2 silver staining and in gel digestion....55
Ch2.10-3 mass spectrum analysis....57
Ch2.11 Flow Cytometry....58
Ch2.12 Immunofluorescence staining....58
Ch2.13 Colony Formation....59

Ch2.14 Chromatin immunoprecipitation (ChIP), sample preparation for ChIP-sequencing and double ChIP....60
Ch2.14-1 ChIP....60
Ch2.14-2 ChIP-sequencing grade ChIP sample preparation....61
Ch2.14-3 Double ChIP....63
Ch2.15 Semi-quantitative Poly Chain Reaction (PCR) and quantitative PCR (qPCR)64
Ch2.16 ChIP-sequencing analysis and validation....66
Ch2.16-1 ChIP-sequencing bioinformatics analysis....66
Ch2.16-1 ChIP-sequencing validation....67
Ch2.17 Quantitation and statistical analysis....68
Chapter 3 E2F1 methylation by PRMT1....69
Ch3.1 Introduction....70
Ch3.2 E2F1 as a substrate of PRMT1 *in vitro*....74
Ch3.3 E2F1 as a substrate of PRMT1 *in vivo*....80
Ch3.4 PRMT1 methylates R109 on E2F1....83
Ch3.4-1 Samples preparation for mass spectrometry....84
Ch3.4-2 E2F1 R109 arginine dimethylation by PRMT1....87
Ch3.5 Summary....95
Chapter 4 Impact of arginine methylation on E2F1 by PRMT1....96
Ch4.1 Characterisation of inducible cell lines....97
Ch4.2 Growth inhibition of PRMT1-induced ADMA on E2F1....107
Ch4.3 Promoter specificity....117
Chapter 5 E2F1 ChIP-sequencing....127
Ch5.1 Introduction....128
Ch5.2 E2F1 ChIP-sequencing samples preparation....134
Ch5.3 E2F1 ChIP-sequencing analysis....141
Chapter 6 Arginine methylation interacting protein partner....158
Ch6.1 Introduction....159
Ch6.2 E2F1 interacting-partners network....160
Ch6.3 PRMT5 induced E2F1 and □□Catenin interaction....162
Ch6.4 Arginine methylation on E2F1 and □□Catenin interaction....165
Chapter 7 Summary & Discussion....167

Table of Figures

FIGURE1. 1. THE STRUCTURAL PROPERTIES OF THE E2F FAMILY.....22
FIGURE1. 2 THE CELL CYCLE.....25
FIGURE1. 3 E2F1 POST-TRANSLATIONAL MODIFICATIONS.....33
Figure1. 4 Arginine methylation by PRMTs.....36

FIGURE3. 1 PRMT5 SUBSTRATES' SEQUENCE MOTIF ALIGNMENT.....72
FIGURE3. 2 PRMT1 METHYLATES E2F1 *IN VITRO*.....77
FIGURE3. 3 PRMT1 METHYLATES E2F1 IN CELLS.....80
FIGURE3. 4 OPTIMISATION OF MASS SPECTROMETRY SAMPLES PREPARATION.....84
FIGURE3. 5 (A) PRMT1 METHYLATES E2F1 ON R109: E2F1 MASS SPECTROMETRY.....87
FIGURE3. 6 (B) PRMT1 METHYLATES E2F1 ON R109: R111/113K E2F1 MASS SPECTROMETRY.....90
FIGURE3. 7 (C) PRMT1 METHYLATES E2F1 ON R109: LOSS OF ADMA WITH E2F1 R109K.....92

FIGURE4. 1 CELLULAR LOCALISATION OF INDUCIBLE PRMT1.....98
FIGURE4. 2 ENZYME ACTIVITY OF INDUCIBLE PRMT1.....100
FIGURE4. 3 CELLULAR LOCALISATION OF INDUCIBLE PRMT5.....102
FIGURE4. 4 ENZYME ACTIVITY OF INDUCIBLE PRMT5.....104
FIGURE4. 5 CORRELATION BETWEEN PRMT1 AND SUB-G1 CELLS.....112
FIGURE4. 6 CHANGE OF APOPTOTIC MARKERS UPON PRMT DEPLETION.....114
FIGURE4. 7 ARGININE METHYLATION SITES AND CELL GROWTH.....116
FIGURE4. 8 (A) E2F1 PROMOTER SPECIFICITY UPON PRMT INDUCTION.....121
FIGURE4. 8 (B) E2F1 PROMOTER SPECIFICITY UPON PRMT DEPLETION.....124
FIGURE4. 8 (C) E2F1 PROMOTER SPECIFICITY UPON ASYMMETRIC OR SYMMETRIC DI-METHYLATION. 127

FIGURE5. 1 OVERVIEW OF CHIP-SEQUENCING WORKFLOW.....131
FIGURE5. 2 CHIP-SEQUENCING WORKFLOW BREAKDOWN.....132
FIGURE5. 3 CHIP-SEQUENCING DNA FRAGMENTATIONS.....135
FIGURE5. 4 CHIP-SEQUENCING SAMPLES QUALITY CONTROL.....139
FIGURE5. 5 CHIP-SEQ BIOINFORMATICS ANALYSIS: CHROMOSOMAL DISTRIBUTION OF CHIP REGIONS. 143
FIGURE5. 6 CHIP-SEQ BIOINFORMATICS ANALYSIS: CATEGORIES OF CHIP REGIONS.....144
FIGURE5. 7 THE CHIP-SEQUENCING STATISTICS.....148
FIGURE5. 8 CHIP-SEQUENCING BIOINFORMATICS ANALYSIS: DIFFERENTIAL BINDING.....151
FIGURE5. 9 FUNCTIONAL ANNOTATION ON DIFFERENTIAL BINDING ANALYSIS HITS OF PRMTS

DEPLETED E2F1 CHIP-SEQ.....154
FIGURE5. 10 BIOCHEMICAL VALIDATION OF DIFFERENTIALLY BOUND E2F1CHIP-SEQ GENES.....156

FIGURE6. 1 E2F1 INTERACTING PARTNER SEARCH BY MASS SPECTROMETRY.....160
FIGURE6. 2 E2F1 AND BETA-CATENIN INTERACTION UPON PRESENCE OF PRMT5.....163
Figure6. 3 Arginine methylation sites and E2F1-Beta-Catenin interaction.....165

Abbreviations

ADMA	asymmetric dimethyl arginine
Ac	acetylation
AdoMet	*S*-adenosyl-L-methionine
amu	atomic mass units
APC	adenomatosis polyposis coli
ATM	Ataxia Telangiectasia Mutated
ATR	ATM- and Rad3-related
BP	biological process
caspase-3	cysteine-aspartic acid protease-3
CDK	cyclin-dependent kinase
ChIP	chromatin immunoprecipitation
ChIP-seq	Chromatin Immunoprecipitation Sequencing
Chk1	Checkpoint kinase 1
Chk2	Checkpoint kinase 2
CK1 α	casein kinase1 α
CKI	CDK inhibitor
DAVID	Database for Annotation, Visualisation and Integrated Discovery
DBD	DNA binding domain
DHFR	dihydrofolate reductase
DOC	deoxycholic acid sodium salt
Dox	upon doxycycline
DP	DRTF- associated protein
DPM	disintegrations per minute
DRTF	differentiation regulated transcription factor
ES	Enrichment Score
EtBr	Ethidium bromide
G1	Gap 1
G2	Gap 2
GAR	glycine- and arginine- rich
GO	Gene Ontology
GSK3	glycogen synthase kinase 3
GST	glutathione *S*-transferase
HA	human influenza hemagglutinin
HAT	histone acetylase
HRP	horseradish peroxidase

inP1	PRMT1 inducible cells
inP5	PRMT5 inducible cells
IP	immunoprecipitation
M	Mitosis
MACS	Model-based analysis of ChIP-seq
Me	methylation
MMA	mono methyl arginine
MS	mass spectrometry
MTase	methyltransferase
NT	non-targeting
NUP155	Nucleoporin 155kDa
PARP	poly ADP-ribose polymerase
PGM	proline/glycine/methionine
PI	propidium iodide
PIKK	phosphatidyl-inositol-3-OH kinase-like kinase
PIR	Protein Information Resource
PP2A	protein phosphatase 2A
PRMT	protein arginine methyl transferase
RefSeq	NCBI Reference Sequence Database
S	Synthesis
SDMA	symmetric dimethyl arginine
siPRMT1	PRMT1 siRNA
siPRMT1	PRMT5 siRNA
siRNA	small interfering RNA
SP	Swiss-Prot
TA	transactivation domain
UTR	untranslated region
XRCC5	X-ray repair cross-complementing protein 5

Chapter 1 Introduction

Ch1.1 Cancer

Neoplasia (new-formation), the presence of formation of new, abnormal growth of tissue, has been classified into four main groups: benign neoplasms, in situ neoplasms, malignant neoplasms, and neoplasms of uncertain or unknown behaviour, in the International Statistical Classification of Diseases (ICD-10) (Who.int, 2010). Neoplasm does not always result in lump formation, but when it caused a lump, it is referred as tumour.

Ch1.1-1 Social and Economic Impact of Cancer

The malignant neoplasm is cancer, a collective term describing a group of diseases characterised by uncontrollable cell proliferation that potentially becomes metastatic (spread and invade to other parts of body from the formation of origin) at later stages and results in pain, post-operative/treatment disability, or death (Wigmore et al., 2013) (Xing et al., 2013). Cancer poses a threat not only to the patients' physical well-being, but also causes psychological distress resulting in fear, depression, anxiety and potentially, suicide, with adverse knock-on effects on families and society as a whole (Errico, 2014) (Badr et al., 2014).

Cancer has taken an enormous human toll claiming approximately 7.6 million deaths in 2008, and 8.2 million deaths in 2012 globally according to World Health Organisation (WHO) (W and P, 2014). Overall, the numbers of diagnosed cancer cases are expected to

rise from 14 million to 22 million within the next two decades (IARC, 2013). Following on from these estimations, the *World Cancer Report 2014* (W and P, 2014) reported that of the these 14 million diagnosed cases in 2012, more than 60% were from Africa, Asia and Central and South America, and overall over 70% of the 8.2 million deaths worldwide recorded in 2012 occurred in these regions. Overall, cancer has a negative impact on people individually, on families and the society at large.

A study of global economic cost of all causes of death helped to assess the scope of the problem by converting these challenges into monetary terms (American Cancer Society, 2010). This report combined World Health Organisation (WHO) methodology and data with World Bank's *World Development Report* of 2004. Overall, the economic impacts of cancer (premature death or disability) was $895 billion in 2008 which was accounted for 1.5% of the world's gross domestic product (GDP), and this figure is nearly 19% higher than heart disease, the second leading cause of economic loss ($895 billion and $753 billon respectively) in *the global economic cost of cancer* (American Cancer Society, 2010). Altogether, when considered the years lost in life and productivity, cancer remained the single largest drain on the global economy, benchmarked with all sorts of deaths globally, including infectious diseases and HIV/AIDS (American Cancer Society, 2010).

Ch1.1-2 Carcinogenesis hypotheses

A cell is a basic unit of life, comprised with a membrane-enclosed nucleus and membrane-enclosed organelles in the cytoplasm. In the nucleus, the chromosome carries genetic materials that can be passed on following an ordered sequence of event in the life of a cell (Reece and Campbell, 2014).

The human body is composed of about one hundred trillion cells (Lodish, 2007). Each cell serves a specific function and is classified into more than 200 types of cells (CRUK). All human cells, are derived from a zygote, and hence contain the same genetic material. These cells are clones of one another and therefore it is in the evolutionary interest to ensure the propagation of genes and to keep the growth in check by both chemical or physical signals from the surrounding cells (Raff, 1992).

Ch1.1-2-1 Genetic mutations of cancer

Cancer is a collective term referring to a set of diseases that are characterised by uncontrollable cell proliferation. Each time a human cell divides, its DNA replicates, transmitting up to three billion nucleotides to the resulting daughter cells in the process. Even with many corrective mechanisms in place, the average mutation rate is estimated to be about 2.5×10^{-8} mutations per nucleotide site or 175 mutations per diploid genome per generation (Nachman and Crowell, 2000). These estimates were later adjusted following

advances in sequencing techniques, and reported to be approximately 1.2×10^{-8} bp^{-1} $generation^{-1}$ (Scally and Durbin, 2012).

Many factors influence the mutation rate, for example, time dependent DNA damage (Stamatoyannopoulos et al., 2009) or reactive oxygen species (ROS) derived from normal metabolic processes, or simply challenges from viruses, mutagens or radiation can further introduce DNA damage which greatly increases mutation rates (Grivennikov et al., 2010).

A way to determine the impact of inheritance in tumour formation is to assess the likelihood for identical twins to develop the same type of cancer. A study showed that 75-year-old identical twins shared a particular type of cancer with a probability between 10 and 15%. This demonstrated that while inherited genes play a significant influence on cancer developing, greater impact comes from genes that are altered during one's lifetime (Karp and Patton, 2013).

In Knudson's multiple-hit hypothesis, it was proposed that cancer requires multiple genes to be deregulated over time through accumulated mutations, and thus the correlation between cancer and longevity (Knudson, 1971). The hypothesis was revised because of mutations found in *RB1* or *TP53* tumour suppressor genes. Patients with these mutations develop cancer at early age possess cancer at early age due to a small number of genetic aberrations (Rushlow et al., 2009).

Ch1.1-2-2 Tumour suppressor genes and oncogenes

The concept of ‘tumour suppressor genes’ was established following research in the 1960’s that showed that the fusion of normal and malignant rodent cells resulted in the loss of malignant characteristics. A normal cell is thus considered to possess factors that contribute to the suppression of uncontrollable malignant cell growth. Additionally, observations of consistent gene deletions in chromosomal regions of certain types of cancer identified genes, that when absent, correlated to the development of a tumour. Examples of these tumour suppressor genes and their proposed functions were summarised in Table 1.1.

Oncogenes, on the other hand, acted as accelerators of cell proliferation. Early studies suggested that RNA tumour viruses transformed normal cells to malignant cells and therefore acted as the sources of oncogenes. However, once the *src* oncogene (which previously was believed to be solely expressed by avian sarcoma virus) was identified to be expressed in uninfected cells, it became clear that oncogenes were not viral genes but cellular genes that incorporated into the viral genome during previous infection. These genes were referred to as proto-oncogenes, which possess the potential to drive a normal cell towards malignancy (Karp and Patton, 2013).

Gene	Primary Tumour	Proposed Function
APC	Colorectal	Binds β-catenin acting as transcription factor
ARF	Melanoma	p53 activation
BRCA1	Breast	DNA repair
MSH2, MLH1	Colorectal	Mismatch repair
E-Cadherin	Breast, colon, etc.	Cell adhesion molecule
INK4a	Melanoma, pancreatic	p16:CdK inhibitor $p14^{ARF}$: stabilises p53
NF1	Nerofibromas	Activates GTPase of Ras
NF2	Meningiomas	Links membrane to cytoskeleton
TP53	Sarcomas, lymphomas, etc.	Transcription factor (cell cycle and apoptosis)
PTEN	Breast, thyroid	RIP_3 phosphatase
RB	Retinal	Binds to E2F
VHL	Kidney	Protein ubiquitnation and degradation
WT1	Wilms tumor of kidney	Transcription factor

Table1.1 Tumour suppressor genes and their proposed function (Karp and Patton, 2013)

Ch1.1-2-3 Darwinian microevolution

Apart from the initial genetic material inherited through the germ line, mutations present in the somatic (body) cells are associated with for example neurodevelopmental diseases and cancer (Poduri et al., 2013). These changes in genomic sequence are known as acquired somatic mutation (Evan and Vousden, 2001). Due to the diploid nature of chromosome in somatic cells, these changes contribute to phenotypic defects if derived from a single allele copy would still become silent.

In molecular oncology, one concept adopted to understand carcinogenesis is Darwinian microevolution. As cells become cancerous from defects accumulated through pathway deregulation, the implication was that these cells once held normal cellular functions. For transformation to occur, two Darwinian evolution criteria must be fulfilled: selective advantages and inheritability. Both somatic mutations and epigenetic changes satisfy these conditions. Epigenetic changes such as chromatin remodelling, histone methylation or DNA methylation, instead of changing the underlying DNA sequences like somatic mutations, alter genetic expression and are heritable (Jablonka and Raz, 2009) (Franklin and Mansuy, 2010).

Ch1.1-2-4 Hallmarks of cancer

In 2011, Hanahan and Weinberg updated their report of 2000, which collected and organised the process a cell undergoes to become cancerous into a total of eight hallmarks in the hope of providing an organising principle in order to rationalise this complex

neoplastic disease. These eight characteristics are sustaining proliferative signalling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, activating invasion and metastasis, and two new additions, reprogramming of energy metabolism and evading immune destruction. Genome instability contributes to these hallmarks and the hypothesis proposes that cancer occurs in multiple developmental steps in which these hallmarks are acquired (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011).

Ch1.1-2-5 Tumour microenvironment

Many studies indicated that tumour growth depends not only the cancer cells but also on the tumour stroma, which is composed of the basement membrane, immune cells, capillaries, fibroblasts, and the extracellular matrix (ECM) that surrounding the cancer cells (Kalluri and Zeisberg, 2006). Evidence of the stroma providing oncogenic signals to facilitate tumourigenesis was demonstrated in chickens infected with Rous sarcoma virus and then physically wounded. Wounding led to the formation of invasive carcinomas within tissues that would otherwise have remained tumour-free (Sieweke et al., 1990) (Dolberg et al., 1985).

Carcinoma-associated fibroblasts (CAFs) were defined as phenotypically distinct fibroblasts that were irreversibly activated in the tumour microenvironment. CAFs are the most abundant cell type in the stroma of many different cancers including breast and

pancreatic (Cirri and Chiarugi, 2011). The transforming growth factor-β (TGF-β) pathway, which promotes wound healing, may mediate the transition of a normal fibroblast into a CAF as the phenotypic features of CAFs can be induced by TGF-β treatment in culture (Aoyagi et al., 2004). Carcinogenesis was interpreted as a by-product of the wound-healing process, in which healing is never achieved and cells are consumed by exhaustion (Meng et al., 2012).

Ch1.2 E2F1 and E2F family

E2F1 is the first member of E2F family to have been discovered. Originally, it was described as a cellular factor required for adenovirus E1A protein to bind and activate its E2 promoter, and hence the name E2 promoter-binding Factor (E2F) (Kovesdi et al., 1986). At the same time, independent research found a novel protein bearing functional resemblance to the adenovirus E1A protein, which during the differentiation of F9 embryonic cancer stem cells, reduced its DNA binding activity (La Thangue and Rigby, 1987). Later, this novel protein named differentiation regulated transcription factor (DRTF) was found to be functionally indistinguishable and therefore, these two factors recognised as being very similar (Trimarchi and Lees, 2002). An essential cofactor was found to dimerise with E2F1 by affinity purification and microsequencing and was named the DRTF-associated protein (DP) (Girling et al., 1993).

To date, there are eight members, which share similar functional domains and are grouped into the E2F family. E2F1 for instance, denotes the first member of the family. Over the last decade, these members have been found to be the key components of the regulatory machinery for cell cycle (Sahin and Sladek, 2010b). The transcription of cell cycle regulators, cyclin E, cyclin A, Cdc2, Cdc25A for examples, and the nucleotide biosynthesis enzymes such as dihydrofolate reductase, thymidylate synthetase and thymidine kinase, and also the DNA-replication machinery such as CDC6, ORC1 and Mcm

proteins, are all governed by E2F (Trimarchi and Lees, 2002). The competing functionalities between the E2F members are well-documented and are broadly subdivided into activators and repressors. While E2F1, E2F2, and a longer spliced E2F3 (E2F3α) are, activators and, promote cell cycle progression, the shorter spliced E2F3β and E2F4-8 are, repressors and, take part in gene repression. Cell proliferation can be completely abolished when E2F1, E2F2 and E2F3α are mutated altogether (Wu et al., 2001). From within the E2F family, the transcriptional activity of E2F1 was shown to be repressed directly by E2F7 and E2F8 following DNA damage (Zalmas et al., 2008).

Three members of DP family have previously identified, of which DP-1 (Helin et al., 1993) and DP-2 (Zhang and Chellappan, 1995) serve as the co-activator of E2F1, while DP-3 participates as a transcription co-repressor (Qiao et al., 2007).

Structurally, all E2F members (E2F1 - 8) contain the highly conserved DNA binding domain (DBD) (Figure1.1). E2F1 - 6 contain a DP domain and after forming a heterodimer, the E2F-DP complex binds DNA. In contrast, the DNA binding of E2F7 - 8 is independent of DP proteins. E2F1 - 5 contain a retinoblastoma (Rb) protein family binding domain (DeGregori and Johnson, 2006).

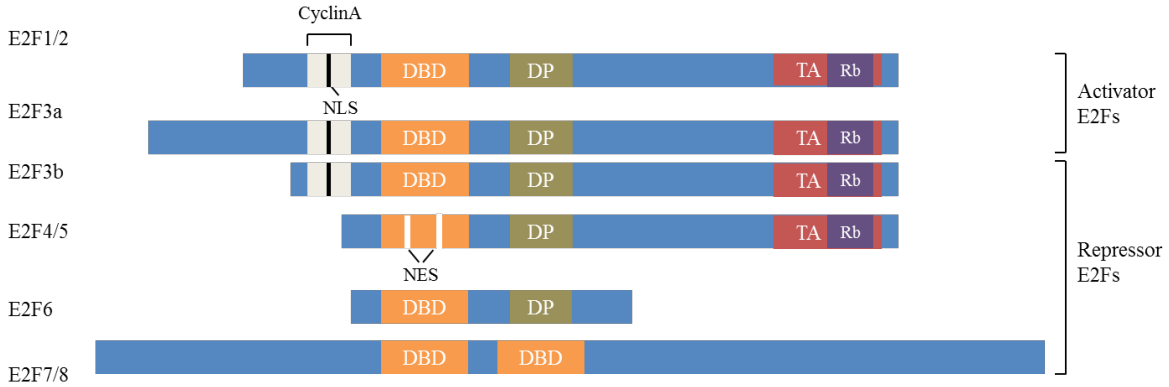


Figure1. 1. The structural properties of the E2F family

E2F family share structural and sequence similarity. They are grouped into activator E2Fs and repressor E2Fs. All members from E2F1 to E2F8 contain DNA binding domain (DBD), E2F1 to E2F6 contain DP dimerisation domain, and E2F1 to E2F5 contain the Rb family binding region in the transactivation domain (TA).Other regions indicated as cyclinA binding, nuclear localisation signal (NLS) and nuclear export signal (NES).

Ch1.3 Cell Cycle

In the life of a cell, an ordered sequence of events occurs to complete the cell division cycle. The synthesis (S) phase, in which DNA replicated in a cell in preparation for cell division and mitotic (M) phase, is sandwiched by two gap phases, the gap 1 (G1) and gap 2 (G2). The G1, S, G2 and M cycle is governed by a checkpoint system, which confirms the fulfilment of the prerequisites of each phase, for example, the completion of DNA replication, before allowing entry to the next phase.

Ch1.3-1 Cell cycle cyclins and CDKs control

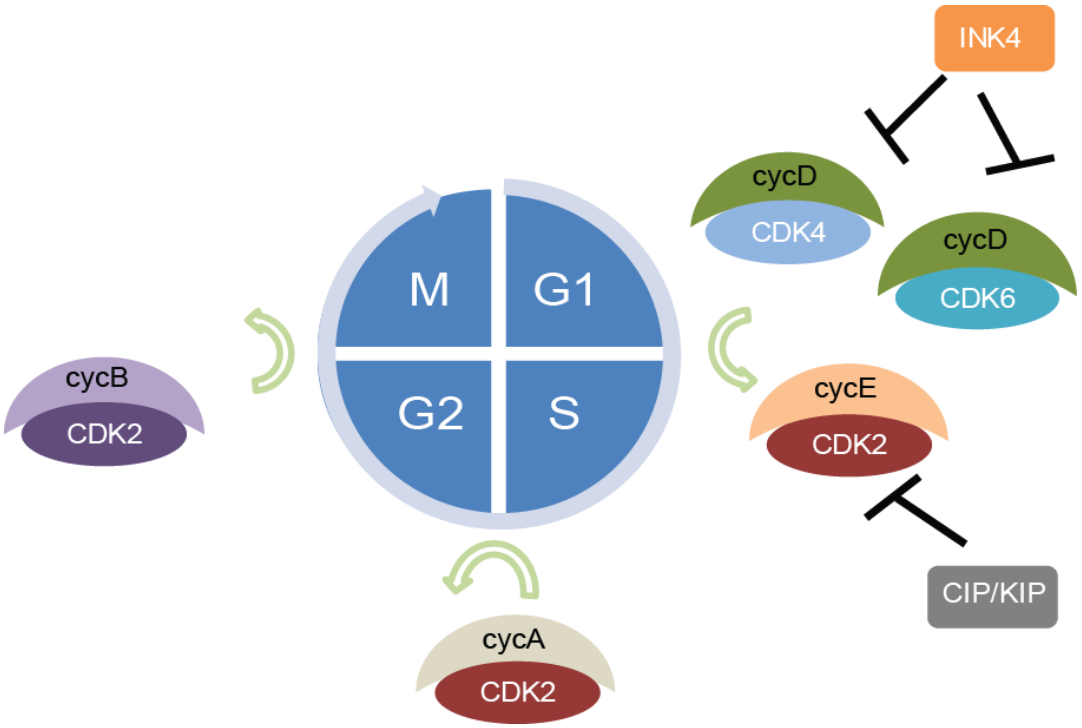
Starting with G1, cells are sensitive to growth stimulation. A critical point is the G1/S checkpoint, which once passed, commits a cell to complete the full cycle and is therefore named the restriction point (Foster et al., 2010) (Bertoli et al., 2013). The progression of a cell cycle is propelled by the dynamic fluctuations of different types of cyclins, a family of proteins controls cell cycle through binding to, and activating of different types of cyclin-dependent kinase (CDK) (Morgan, 2007). Through the binding of different CDK to different cyclins, a cell cycles through the sequenced stages to complete a division cycle. The CDK is activated through cyclin-CDK complex formation, and the cyclin-dependent kinase phosphorylates their substrates on serines and threonines. They are thus defined as serine-threonine kinases. As different cyclins are formed at different stages of cell cycle,

through the binding of different CDK to different cyclins, a cell cycled through the sequenced stages to complete a division cycle (Williams and Stoeber, 2012).

The break system employed is the CDK inhibitors (CKIs) when aberrations are detected in the process of the cycle, for example, DNA damage. The INK4 family that inhibits CDK4 and CDK6 or the CIP/KIP family members that targets to inhibit CDK2 activity.

Cyclin D are short-lived proteins whose assembly with CDK4/6 in the G1 phase is dependent on mitogenic signalling. As long as the mitogenic signalling continues, the CyclinD/CDK4 or CDK6 activity goes on. Following Cyclin E levels peaking at the G1/S transition, Cyclin A levels increase during S phase. Both Cyclin E and Cyclin A interact with CDK2 and later in S/G2 transition, Cyclin A can interact with CDK1 (Giacinti and Giordano, 2006) (Figure1.2).

A



B

Interphase	**Cyclin- CDK**
G1/S	Cyclin D- CDK4 Cyclin D- CDK6 Cyclin E- CDK2
S/G2	Cyclin A- CDK2
G2/M	Cyclin B- CDK1

Figure1. 2 The cell cycle

A cell cycle comprises with the G1, S, G2 and M phase. Cell cycle progression is mediated by different cyclin-CDK interaction at different stages. The INK or CIP/KIP family inhibits CDK activity. (B) The cell cycle transition interphase and the corresponding cyclin-CDK combination at each phase.

Ch1.3-2 Cell cycle checkpoints controls and DNA damage

To ensure the proliferation of only those cells fit for division, the cell cycle relies on three major checkpoints to progress through a cell division cycle, the G1/S, intra-S and G2/M checkpoints. In response to DNA damage, the “effectors” are activated and transmit DNA damage signals to initiate the checkpoint system and inhibit cell cycle progression. The important effectors include ATM (ataxia telangiectasia mutated) and ATR (ATM- and Rad3-related), and PI(3)K (phosphatidyl-inositol-3-OH kinase)-like kinase (PIKKs).

The DNA damage response depends on the cooperation of ATR, ATM and the transducer kinases Chk1 and Chk2 which unfold a cascade of phosphorylation events. The aim of the checkpoint system is to transiently delay the progression of a cell cycle, even beyond the restriction point, and to allow time for DNA repair.

A way to stop cells at the G1 phase is to phosphorylate the tumour suppressor gene, p53 and its suppressor MDM2, resulting in a more stable tumour suppressor p53 with an increased transcriptional activity. One of the genes targeted by p53 is *p21* which mediates the inhibition of cyclin E/CDK2 interaction at late G1/S and therefore halt cells at in G1/S phase (Ishikawa et al., 2006). Finally, to halt cells at G2 phase, the main target is Cyclin B/CDK1 as its activity is abolished upon phosphorylation mediated by ATM/ATR, Chk1/Chk2 or p38 kinase (Malumbres and Barbacid, 2007).

Ch1.4 E2F1 and pocket protein in cell cycle regulation

Pivotal to cell cycle regulation is the pocket protein family, the retinoblastoma (Rb), p107 and p130. When a cell does not divide and stays arrested at G0 phase due to the deprivation of growth factors, or the occurrence of nutrient insufficiency in high cell density environments, p130 expression predominates. In the mid G1 phase, p107 replaces p130 as the predominant protein at E2F responsive promoters. The Rb expression enriched towards the late G1 phase. All three pocket proteins are expressed differentially to prevent the transcription of E2F regulated genes and therefore control the progression of a cell. The cell cycle arrest achieved by negative regulation on two positive regulators of cell cycle entry, the inhibition of E2F transcription and the low level of cyclin-CDK activity (Henley and Dick, 2012).

One of the objectives of cyclin-CDK complex formation is the phosphorylation of a pocket protein which results in the release of E2F, a family of transcription factors, pivotal to the orchestration of the cell cycle. In the absence of mitogenic growth factors, the E2F-DP1 heterodimer commonly interacts with C terminal domain of Rb (RbC). A study of the overall structure of the E2F1-DP1-RbC complex revealed that RbC directly binds with E2F1-DP1 in two discontinuous regions. Moreover, E2F1-DP1 heterodimerisation was found to be required for RbC binding (Rubin et al., 2005). In the presence of mitogenic stimuli, a series of phosphorylation events take place, such as the phosphorylation of p130

and p107 pocket proteins by cyclin E/CDK2 and later by cyclin A/CDK2 (Giacinti and Giordano, 2006). Entry of the S phase of the cell cycle relies on hyperphosphorylation of Rb. Phosphorylation of S788 and S705 on Rb for example, weakens the direct association between Rb, E2F1 and its heterodimerisation partner DP (Burke et al., 2014). Moreover, the binding of phosphorylated RbC to the pocket and to E2F1-DP1 is mutually exclusive (Rubin et al., 2005). The phosphorylation of pocket proteins releases and activates E2F1 which in turn initiates a positive feedback as E2F1 further promotes cyclin E expression (Henley and Dick, 2012). Once committed to S phase, the cyclin E/CDK2 balances its expression by binding to unbound cyclin E for ubiquitination and hence reduces the cyclin E/CDK2 activity. In short, E2F1 binds to and is negatively regulated by cell cycle dependent phosphorylation of Rb by cyclin dependent kinases. The hyper-phosphorylated Rb releases E2F1 to commit cells to progress to cell cycle synthesis (S) phase irreversibly. Nearly one third of human cancers result from defects in Rb/E2F1 pathway (Nevins, 2001).

At the end of mitosis phase, protein phosphatase-1 targets Rb and protein phosphatase-2 targets all three pocket proteins for dephosphorylation and thereby supress E2F1 activity (Hirschi et al., 2010).

Ch1.5 E2F1 in tumour progression and suppression

As the best-characterised member of the E2F transcription factor family, E2F1, regulated more than 1000 genes that are involved in various cellular functions (Sahin and

Sladek, 2010b). Besides its well-documented roles in the cell cycle, especially those in favour of cell cycle progression, E2F1 has been reported to be able to introduce apoptosis (La Thangue, 2003) (Meng and Ghosh, 2014). Interestingly, although all activating E2Fs are capable of driving cells through S phase entry, only E2F1 is able to introduce apoptosis (DeGregori et al., 1997).

There is a wealth of literature reporting that E2F1 stimulates cell cycle progression. The cell cycle regulator genes such as Cyclin A, Cyclin E, Cdc2, Cdc6 and dihydrofolate reductase (DHFR) are all E2F1 responsive genes (Wolf, 2009). The over-expression of E2F1 is sufficient to stimulate quiescent cells to re-enter the cell cycle and also induce the transformation of primary cells (Helin et al., 1992, Lees et al., 1993, Johnson et al., 1994). In $Rb^{+/-}$ mice, frequency of tumours is reduced upon deactivation of E2F1 (Yamasaki et al., 1998), suggesting a tumour progression role for E2F1.

Paradoxically, in $E2F1^{-/-}$ mice, increased incidence of tumour formation is seen (Field et al., 1996) which suggests that there is tumour suppression ability in E2F1. In retinoblastoma tumour suppressor protein (Rb) deficient mice, mutated E2F1 suppresses the apoptosis and prolonged the survival of mouse embryos (Tsai et al., 1998). Several known apoptotic genes have been identified as E2F1 responsive genes, these include both p53 dependent genes, such as $p19^{ARF}$ and p53 independent genes, such as p73, pathways (Wu et al., 2009).

Ch1.6 Post-Translational Modifications and E2F1 regulation

The yin and yang properties of E2F1 highlighted the importance to understand E2F1 regulation. Post-translational modifications (PTM) provide a means to transmit regulatory signals to proteins, for example, phosphorylation, ubiquitination, NEDDylation, SUMOylation, acetylation, ribosylation, citrullination and methylation.

Initially, Allis and colleagues described PTMs on histone proteins and proposed the combination of these dynamic covalent histone modifications is like a “new language” that provides histone proteins with multi-layer complex regulations (Strahl and Allis, 2000). This histone code hypothesis referred to the phenomenon principally emphasised in two parts, firstly the “writing” and “erasing” of numbers of different modifications on histone proteins in a given cellular content or environment, and secondly the “reading” and the resultant effects of a given combination of PTMs. The two concepts together makes up of the “new language” which extend beyond what basic composition of histone proteins may offer to respond (Chi et al., 2010).

In the past decades, the understanding of post-translational modification has extended from histones to non-histone proteins (Singh et al., 2010). Many PTMs contribute to the E2F1 regulation and therefore impact on E2F1 activity, for examples, phosphorylation, ubiquitination, acetylation and methylation.

The phosphorylation events on E2F1 are important for cells to respond to DNA damage, these are the S31 phosphorylation mediated by ataxia telangiectasia mutated/ataxia telangiectasia and Rad3-related (ATM/ATR) kinases and S364 mediated by Chk2 enzyme which contribute to the stabilisation of E2F1 protein and apoptosis (Lin et al., 2001) (Stevens et al., 2003). Other examples include when phosphorylation occurred on S332 and S337 of E2F1 which help to prevent E2F1 binding to Rb (Sahin and Sladek, 2010a). This is crucial to keep E2F1 free from the negative regulator Rb in the cell cycle. The p38 MAP mediated phosphorylation at S403 and S433 of E2F1 assists E2F1 exportation from the nucleus for its degradation (Ivanova et al., 2009).

Ubiquitination is a way to monitor E2F1 protein level in the cells. The F-box-containing protein $p45^{SKP2}$ is responsible for E2F1 degradation during S/G2 transition in the cell cycle (Jackson and Eldridge, 2002). E2F1 was also found ubiquitinated by the APC/C complex and the resultant K11 linkage-specific ubiquitin chain which help target E2F1 for degradation (Budhavarapu et al., 2012)

The acetylation mediated by P/CAF, a protein associated with histone acetylase (HAT) p300/CBP complex, at K117, K120 and K125 is conserved between E2F1, E2F2 and E2F3 (Martinez-Balbas et al., 2000) . The acetylation event increases the stability of E2F1 with increased p73 gene binding specificity which suggests the enhancement of the apoptotic effect of E2F1 (Ozaki et al., 2009). The “erasing” of acetylation in this case is the

deacetylase HDAC1 which function is dependent on the E2F1-Rb interaction (Martinez-Balbas et al., 2000).

The ubiquitin-like molecule, NEDD8, traditionally catalyses NEDDylation on the cullins substrate. E2F1 is a non-cullins substrate for NEDDylation mediated by NEDD8, and interestingly E2F1 methylation at lysine (K)185 was required for this event. This is one of the examples of cross talk between post-translational modifications on the same substrate. The NEDDylation on E2F1 contributed to a slow cell growth, a decline in E2F1 stability and transcriptional activity (Loftus et al., 2012).

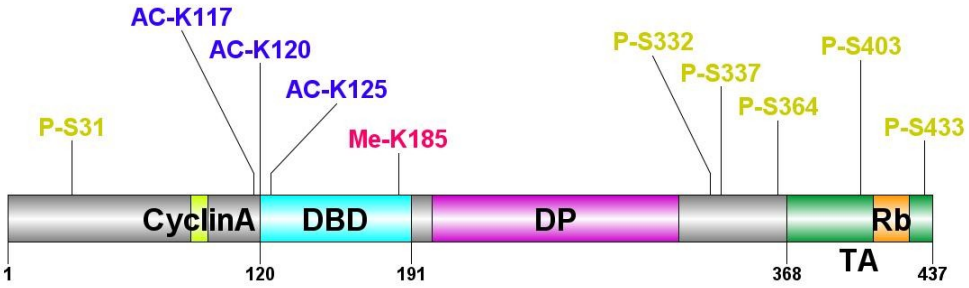


Figure1. 3 E2F1 post-translational modifications

The schematic diagram shows the structure of E2F1 (amino acid (aa)1-437) and the post-translational modifications. E2F1 protein domains: Cyclin A (aa.87-95), DNA binding domain (DBD) (aa.120-191), DP binding domain (aa.201-317), transactivation domain (TA) (aa. 368-437) and the Rb binding domain (aa.409-426). The post-translational modifications are colour coded and displayed as the "type of modification- site of modification," for example, the P-S31 means phosphorylation (P) occurs on serine at position 31 (S31), other phosphorylation sites include S332, S337, S364, S403 and S433. The acetylation (AC) sites are lysine (K) 117,120,125. The lysine methylation (Me) occurs on the lysine position 185.

Ch1.7 PRMT family and arginine methylation

Increasingly, arginine methylation has been identified as an important post-translational modification involved in a multitude cellular functions such as signal transduction, cellular proliferation, RNA processing, DNA repair, chromatin regulation and transcriptional control. The methylation occurs at the N- and C-termini and amino acid side chain residues, of which the methylation on arginine side chain guanidine groups reported to likely be quantitatively most extensive in mammalian cells (Clarke and Tamanoi, 2006) (Gary and Clarke, 1998).

As the arginine side chain guanidino groups contain five possible hydrogen-bond donors, which are in favour of interaction from hydrogen-bond acceptors. The addition of each methyl group compensates one such hydrogen bond donor and therefore methylation changes not only the structure but also acts as an inhibition mechanism to some of downstream binding partners (Luscombe et al., 2001).

Arginine methylation is catalysed by a family of enzymes known as protein arginine methyltransferases (PRMTs). To date, there are ten members discovered but the activity of the newest member, PRMT10, has not been characterised (Wei et al., 2014) and will therefore not be discussed further. The PRMTs catalyse the transfer of the methyl group from the donor molecule, *S*-adenosyl-L-methionine (AdoMet), to the guanidine nitrogen atom of the arginine residues in the target protein. In mammalian cells, the mono- and

dimethylation reactions on arginine result in the ω-N^G- monomethylarginine (MMA), symmetric ω-N^G, $N^{G'}$ -dimethylarginine (SDMA), or asymmetricω-N^G, N^G-dimethylarginine (ADMA). The ω symbol was given by IUPAC nomenclature and means that the methylation of the arginine happens on the terminal nitrogen atoms. The MMA and ADMA are catalysed by type I PRMTs (PRMT1, 2, 3, 4, 6, 8) whilst the formation of MMA and SDMA are catalysed by type II PRMTs (PRMT5, 7, 9). Amongst different types of arginine methylation, the most prevalent is ω-N^G, N^G- dimethylarginine, the asymmetric dimethylarginine (ADMA), in which two methyl groups are placed on one of the terminal nitrogen atoms of the guanidino group followed by the ω-N^G, $N^{G'}$ -dimethylarginine, the symmetric dimethylarginine (SDMA), where the two methyl groups are arranged to separately and placed as one methyl group to each of the terminal guanidino nitrogens with an occurrence frequency of 20% to that of ADMA. The third type of arginine methylation is the single methyl group on the terminal nitrogen atom, the monomethylarginine (MMA), which happens 50% of the levels of ADMA (Bedford and Richard, 2005) (Bedford and Clarke, 2009). The level each types of arginine methylation influences each other. The loss of PRMT1 activity for example corresponded to major increases in global level of MMA and SDMA (Dhar et al., 2013).

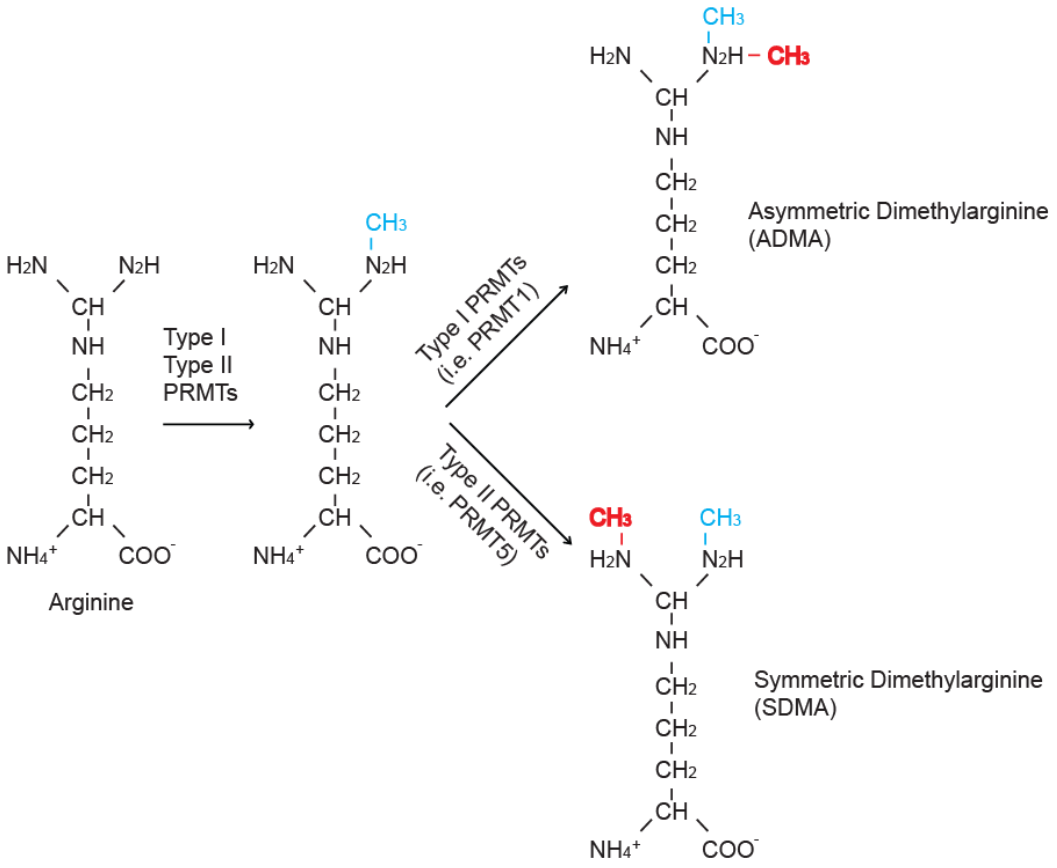


Figure1. 4 Arginine methylation by PRMTs.

Both type I (PRMT1) and type II (PRMT5) PRMTs can generate monomethyl-arginine by transferring a methyl group (indicated in blue) on the arginine. However, PRMT1 (Type I) transfers the second methyl group (indicated in red) asymmetrically onto the arginine to form asymmetric dimethyl-arginine (ADMA) whilst PRMT5 (Type II) adds the second methyl group symmetrically (indicated in red) to form symmetric dimethyl-arginine (SDMA).

The PRMT family although differing in sequence, length, and specificity for distinct substrates, are highly conserved evolutionally with a core methyltransferase (MTase) domain, a dimerization arm and a β-barrel (Bedford and Clarke, 2009, Bachand, 2007, Wolf, 2009). The enzyme activity is to utilise the ubiquitous cofactor *S*-adenosyl-L-methionine (AdoMet) as a donor of methyl groups. Mostly, PRMT targets a glycine and arginine (GAR) rich patch motif with the exception of PRMT4 that prefers proline/glycine/ methionine (PGM) rich motifs (Lee and Bedford, 2002) (Yang and Bedford, 2013).

Although common features are shared amongst the family, each member contains its distinct features. PRMT1 and PRMT5 are the major arginine methyltransferases that deposit asymmetric and symmetric dimethyl arginines on its target, in their respective subdivision, type I and type II PRMTs (Yu et al., 2009) (Tee et al., 2010).

The most dominant player in mammalian cells is PRMT1 which is responsible for more than 80% of all PRMT activity in a cell. It is also the smallest in size containing 361 amino acids and has 3 major transcription variants contributing to isoforms with an alternative transcription start site. This is the first member that was cloned, purified and characterised, in the PRMT family (Krause et al., 2007). With the exception of PRMT8 which locates primarily at the plasma membrane, and PRMT3 in the cytoplasm, the remaining type I PRMTs, PRMT1, PRMT2, PRMT4 and PRMT6 are all predominantly present in the nucleus (Wei et al., 2014). It is interesting to note that PRMT8 shares over

80% structural similarity to PRMT1 yet the unique N-terminal myristoylation motif which helps locate PRMT8 to plasma membrane also makes its methylation ability approximately 10 fold lower than it is to PRMT1 (Sayegh et al., 2007). Structurally, PRMT1 exists as a homodimer and the three peptide-binding channels in the catalytic domain allow its contact to a diverse set of substrates (Zhang and Cheng, 2003).

PRMT1 is implicated necessary in the early mouse development as the knockout mice die at an early stage. The aberrant isoforms overexpression is associated with breast (Mathioudaki et al., 2011), lung (Yoshimatsu et al., 2011), colon (Mathioudaki et al., 2008), prostate (Seligson et al., 2005) bladder cancer and leukaemia.

PRMT5 is the major enzyme deposit SDMA mark on the target protein. Like the rest of type II PRMTs with the exception of PRMT9, PRMT5 primarily locates in the cytoplasm (Wei et al., 2014). There is an accumulation of evidence that PRMT5 is associated with the repression of tumour suppressor pathways which render its oncogenic feature. These substrates include histone H2AR3, H3R8 and H4R3 and the elevated levels were detected in leukaemia and lymphoma cells (Pal et al., 2004).

As a DNA damage co-factor of the p53 complex, PRMT5 was found to methylate p53 which affects the target gene specificity. PRMT5 depletion triggers p53 dependent apoptosis (Jansson et al., 2008). More recently, studies have shown that PRMT5 influences another major tumour suppressor pathway. Studies have shown that PRMT5 catalyses

arginine methylation on E2F1 at R111 and R113. Functionally, depletion experiments of PRMT5 or the ectopic R111/113K E2F1 showed the stabilisation of E2F1 and an increase in sub-G1 cell population as well as the increase of activation transcriptionally. PRMT5 mediated arginine methylation inhibited apoptosis effect of E2F1 (Cho et al., 2012).

The PRMT1 and PRMT5 each has an extensive list of targets that is summarised at Table 3.1.

Ch1.9 Research Objective

The aim of the study is to describe a level of E2F1 control mediated by arginine methylation and proposes the arginine methylation as an underlying mechanism, which contributes to the opposing E2F1 impacts on cells. With the addition of symmetric or asymmetric dimethylation marks on E2F1, the cell cycle regulation or the apoptotic effects of E2F1 is then correspondingly signalled to unfold. Therefore, this project aims at testing the hypothesis of different type of arginine methylation serves as a switch which regulates E2F1 and offers a possible explanation to the unsolved enigma of E2F1 by assigning the conflicting resultant cell fate to different type of arginine methylation.

Chapter 2 Materials and Methods

Ch2.1 Cell Lines and Tissue Culture

There are three different human cancer cell line systems:

Ch2.1-1 Human cancer cell lines

U2OS (osteosarcoma), SAOS-2 (sarcoma osteogenic), HeLa (cervical), C33A (cervical), A2058 (skin), MCF-7 (breast adenocarcinoma), MDA-MB-231 (breast adenocarcinoma), HCT116 (colon carcinoma) cell lines were maintained in Dulbecco modified Eagle medium (DMEM) (Invitrogen) with 5% Foetal calf serum (FCS) (Sigma®) and 1% (v/v) Penicillin/Streptomycin (Pen-Strep) (Gibco). H1299 (lung carcinoma) cell line was maintained with RPMI with 10% Foetal calf serum (FCS) and 1% Pen-Strep. Cells were cultured at 37°C with 5% CO_2. Cells were regularly passaged at 80-90% confluency. The tissue culture process was that cells were rinsed with Phosphate Buffered Saline (PBS) (Oxoid) and incubated with Trypsin –EDTA (Lonza) for 5-10 minutes to detach the cells from the plates.

Ch2.1-2 Generation of PRMT1 and PRMT5 inducible cell lines

HA-PRMT1 and HA-PRMT5 plasmids were cloned into pTre2-hyg vectors and transfected into G418 resistant Tet-On® U2OS cells. After 48 hours transfection, cells were selected with selection medium (200 μg/ml G418,5% Tet-negative FCS, 1% P/S and 0.3% hygromycin) until when selection in control cells (pcDNA3.1 transfected, no G418 antibiotics resistance) completed (cells died due to the lack of antibiotics resistance). The

selected cells were then maintained in maintenance medium (DMEM supplement with 5% Tet-negative FCS, 100μg/ml G418 (Clontech), 1% Pen-Strep and 150 μg/ml hygromycin). 1 μg/ml doxycycline incubation for 24 hours condition was used to induce PRMT1 or PRMT5 expression. The cell culture passage followed routine 80% confluency and Trypsin-EDTA approached protocol.

Ch2.1-3 Generation of stably expressed E2F1 cells

Stably inducible cells, HA-WT-E2F-1, HA-R109K-E2F-1, HA-R111/113KK-E2F-1, and mammalian transfection control empty vector (pcDNA3.1) cell lines were generated by transfecting 1 μg respective DNA plasmids into HeLa cells for 48 hours before selecting with 800μg/ml G418 for 5-7 days. As the mammalian DNA plasmids transfected (pcDNA3.1 vector, HA-WT-E2F1, HA-R109K-E2F1 and HA-R111/113KK-E2F1) contained G418 resistance, the generation of stably induced E2F1 and derivatives cells utilised G418 containing medium to select plasmids containing cells. This is because G418 is an antibiotic structurally similar to gentamicin B1 which blocked polypeptide synthesis in cells.

Cells were incubated with 5% Foetal calf serum (FCS) (Sigma) and 1% (v/v) Penicillin/Streptomycin (Pen-Strep) (Gibco). The cell lines kept for experiments for 2 weeks before disposal due to the apoptotic nature of E2F1 plasmids. Reproduction required re-transfection and re-selection with no maintenance needed.

Ch2.2 Site-directed mutagenesis

Catalytic-enzyme dead PRMT1 (G80R) was generated by site mutagenesis using primers:

Primers	Sequence
Forward	5’-CTA GCG CGG CCG CTT ATC ACT AGC GCA TCC GGT AGT C-3’
Reverse	5’-CAT GGA TCC ATG TAC CCT TAT GAC GTG CCC GAT TAC GCT GCG GCA GCC GAG-3’

PCR was performed using Pfx50 DNA polymerase (Invitrogen) at an annealing temperature of 55°C and an extension temperature of 68°C. The PCR products were then ligased with appropriate vectors using T4 DNA ligase (NewEngland BioLabs) and transformed into competent XL1-Blue cells (Aglient Technologies). The purified plasmids were then sequenced to confirm the point mutation.

Ch2.3 DNA plasmids transfection

The DNA plasmids transfection was done with pre-incubating the transfection mix, GeneJuice® (Novagen) transfection reagent and OPTI-MEM® (Invitrogen), for 5 minutes, and then DNA plasmids incubated with transfection mix for future 15 minutes in the room temperature before drop-wisely transfected to the plates pre-seeded with cells. Cells were harvested 48 hours post plasmid transfection. The DNA plasmids used:

Plasmids	Tag	Expression	Description

pTRE-hyg vector	-	Mammalian	Vector template as a control
pTRE-hyg-HA-PRMT1 (WT)	HA	Mammalian	Induced by 1 μg/ml doxycycline
pTRE-hyg-HA-PRMT5 (WT)	HA	Mammalian	Induced by 1 μg/ml doxycycline
pcDNA3.1 vector	-	Mammalian	Vector template as a control
pcDNA3.1-HA-PRMT1 (WT)	HA	Mammalian	Expressed enzymatically active PRMT1
pcDNA3.1-HA-PRMT1 (G80R)	HA	Mammalian	Gly to Arg point mutation at residue 80 of catalytic domain
pcDNA3.1-HA-PRMT5 (WT)	HA	Mammalian	Expressed enzymatically active PRMT5
pcDNA3.1-HA-PRMT5 (Δ365-369)	HA	Mammalian	Deletion mutation of catalytic domain
pcDNA3.1-Flag-PRMT5 (WT)	Flag	Mammalian	Expressed enzymatically active PRMT5
pcDNA3.1-Flag-E2F1 (WT)	Flag	Mammalian	Wild-type E2F1
pcDNA3.1-Flag-E2F1 (1-194)	Flag	Mammalian	E2F1 5’ terminal from residue 1-194
pcDNA3.1-HA-E2F1 (WT)	HA	Mammalian	Wild-type E2F1
pcDNA3.1-HA-E2F1 (R109K)	HA	Mammalian	Arg to Lys point mutation at

(K)			residue 109
pcDNA3.1-HA-E2F1 (R111/113K) (KK)	HA	Mammalian	Arg to Lys point mutations at residue 111 and 113
pGEX vector	-	Bacteria	Vector template as a control
pGEX-PRMT1 (WT)	-	Bacteria	Expressed enzymatically active PRMT1
pGEX- PRMT1 (G80R)	-	Bacteria	Gly to Arg point mutation at residue 80 of catalytic domain
pGEX- PRMT5 (WT)	-	Bacteria	Expressed enzymatically active PRMT5
pGEX- PRMT5 (Δ365-369)	-	Bacteria	Deletion mutation of catalytic domain
pGEX-E2F1 (WT)	-	Bacteria	Wild-type E2F1
pGEX-E2F1 (R109K) (K)	-	Bacteria	Arg to Lys point mutation at residue 109
pGEX-E2F1 (R111/113K) (KK)	-	Bacteria	Arg to Lys point mutations at residue 111 and 113

Ch2.4 small interfering RNA (siRNA)

Oligofectamine reagent (Invitrogen) and siRNAs were incubated with OPTI-MEM ® Reduced Serum Media (Invitrogen) separately at room temperature for 5 minutes before the

all co-incubated for further 15 minutes. The Oligofectamine-siRNA-OPTI-MEME solution was then added drop-wisely onto pre-seeded plates with 30-40% cell confluency for 72 hours before harvest. Customised Dharmacon siRNAs listed as below:

siRNAs	**Concentration**	**Sequence**	**Source**
Non-targeting control (NT)#1	50nM	5'-UUC UCC GAA CGU GUC AGG-3'	Dharmacon
Non-targeting control (NT)#2	50nM	5'-UAA GGC UAU GAA GAG AUA CUU-3'	Dharmacon
PRMT1	50nM	5'-GGA CAU GAC AUC CAA AGA-3'	Dharmacon
PRMT5	50nM	5'-CCG CUA UUG CAC CUU GGA A-3'	Dharmacon
E2F1#1	50nM	5'-AAC UCC UCG CAG AUC GUC AUC-3'	Dharmacon
E2F1#2	50nM	5'-GCU UUA AUG GAG CGU UAU U-3'	Dharmacon

Ch2.5 Recombinant protein purification

Escherichia coli (E. coli) cells (BL21 strain) were transformed with pGEX bacteria expression vector backboned plasmids and plated onto agar dishes containing 100 μg/ml ampicillin overnight at 37°C. Selected one of the amplified colonies and cultured with 5ml

ampicillin (100 μg/ml) containing Luria Broth (Luria) for 8 hours at 37°C on the shaker (starter culture). Monitored the optical density (OD) of the culture medium once the starter culture transferred to 200ml pre-autoclaved 100 μg/ml ampicillin containing Luria Broth until it reached OD_{595}~0.6-0.8. IPTG (VWR International Ltd) was then added when growth reach the exponential log phase to induce protein expression overnight at 20°C. Bacteria cell pellets were sonicated at 23% amplitude in 1% Triton X-100. The lysates were centrifuged at 13,000 rpm for 30 minutes (Sorvall Evolution centrifuge, SLA3000 rotor) and the supernatant was transferred and mixed with glutathione sepharose beads (GE healthcare). The glutathione sepharose beads were then spun down at 1000 rpm (Beckman GS-6S) and washed three times. Recombinant proteins were eluted using 20mM reduced glutathione dissolved in 50mM Tris-HCl (pH 8.0). The elutions were repeated three times and pooled together. Eluates were then subject to buffer exchange by injection into Slide-A-Lyzer Dialysis Cassette (ThermoScientific) with 10kDa molecular weight cut-off to remove excess reduced glutathione.

Ch2.6 *In vitro* ^{3}H-SAM Methylation Assay

Mixed 0.5μg enzyme and 1μg substrate with 1 μCi *S*-adenosyl-[^{3}H-methyl]-methionin (^{3}H-SAM) with methylation buffer (50mM Tris-Hcl (pH 8.0), 20mM KCl, 5mM DTT, 1mM EDTA) and incubated at 30°C for 1 hour. The enzyme, substrate table listed below:

Enzyme	Tag	Source	Enzyme status
PRMT1	GST	Recombinant enzyme purified from bacteria	enzymatically active
PRMT1 (G80R)	GST	Recombinant enzyme purified from bacteria	enzymatically dead
PRMT1	HA	enzyme purified from mammalian cells	enzymatically active
PRMT5	GST	Recombinant enzyme purified from bacteria	enzymatically active
PRMT5 (Δ365-369)	GST	Recombinant enzyme purified from bacteria	
PRMT5	HA	enzyme purified from mammalian cells	enzymatically active
Set7/9	GST	Recombinant enzyme purified from bacteria	enzymatically active
Substrate	**Tag**	**Source**	
E2F1	GST	Recombinant protein purified from bacteria	
Histone	-	Calf thymus histone	

The reactions were spotted on to the P81 filters (Whatman®). The air-dried filters were washed with the washing buffer (200mM NaHCO, pH 9.2) three times. The filters were rinsed with acetone briefly after the final wash. Following the acetone wash, filters were air-dried and placed into scintillator vials containing Ready Protein ™ Liquid Scintillation cocktail (Beckman Coulter ®).The disintegrations per minute (DPM) were counted using a scintillation counter.

Ch2.7 SDS-PAGE and Immunoblotting

Cell were harvested and spun down at 1000 rpm (table centrifuge) at 4°C for three minutes for pelleting the cells. The cell pellets were then lysed with TNN buffer (50mM Tris-HCl pH7.4, 150mM NaCl, 5mM EDTA, 0.5% Igepal (NP-40), 50mM NaF, 0.2mM sodium orthovanadate and protease inhibitor cocktails (1M PMSF, leupeptin [1μg/ml], aprotinin [1μg/ml], and pepstatin A [1μg/ml])) for at least 20 minutes on ice. The cell lysate was then homogenised and spun down at 13,000rpm (table centrifuge) at 4°C for 10 minutes to separate the debris. Protein extracts were measured and normalised using Bradford assay (Bio-Rad). Normalised samples were mixed with SDS loading dye (125mM β-mercaptoethanol, 50mM Tris-HCl pH 6.8, 2% SDS, 7.5% glycerol, 0.1% bromophenol blue) and heated to denature at 95°C for 5 minutes. ColorPlus Prestained Protein ladder (New England Biolans) was run alongside samples for molecular mass estimation control.

Samples were separated by size using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Acrylamide gels were constituted with either 8% or 12% acrylamide, 375mM Tris-HCl Ph8.8, 0.1% SDS, 0.15% ammonium persulphate (APS), 0.15% tetramethylethylenediamine (TEMED); staking gel: 3.2% acrylamide,125mM Tris-HCl pH6.8, 0.1% SDS, 0.15% APS, 0.15% TEMED, and carried out in Laemmli running buffer (25Mm Tris, 192mM glycine, 0.1% SDS) at 100V for 50 minutes. Electrophoresed samples were transferred onto protein nitrocellulose or PVDF membranes (pre-rinsed in

methanol for 2 minutes) (Whatman). The Laemmli transfer buffer (25mM Tris, 192Mm glycine) was used at 400mA for 1 hour or 150mA overnight.

To validate the transfer of proteins, the membranes were stained with Ponceau S dye (Sigma®). The dye could be reversed with ddH_2O and the membranes were incubated with blocking solution (5% non-fat milk, 0.1% Tween-20, phosphate buffered saline (PBS)) for 5 minutes at room temperature before overnight 4°C incubation with primary antibodies (diluted in blocking solution with pre-determined concentration). Membranes were washed three times in 0.1% Tween-20/PBS for 10 minutes each time before 1 hour incubation at room temperature with horseradish peroxidase (HRP)-tagged secondary antibody in blocking solution. Signals of interested proteins were measured with SuperSignal West Pico Substrate (Pierce®) and exposing X-ray films (Kodak) to the membranes. The AGFA CP1000 automatic developer was used to develop the films.

Ch2.8 Antibodies

E2F1 (KH95 and C-20), DP-1(K20) (Santa Cruz); β-Actin (Sigma), Flag peptide monoclonal antibody M2 (Sigma), Anti-HA coupled agarose beads, Anti-Flag peptide monoclonal antibody M2 coupled agarose beads were from Sigma. PRMT1 (Abcam), PRMT5 (Millipore), HA11 (Covance), Cleaved-caspase 3 (Cell Signalling), Cleaved-PARP, β-catenin were from BD Pharmingen, mouse and rabbit secondary antibodies were from Dako. Anti-Symmetric Dimethyl-arginine (SDMA) antibody was purchased from Cell

Signalling and Anti Asymmetric Dimethyl-arginine (ADMA) antibody was purchased from Abcam.

Ch2.9 Immunoprecipitation

Samples underwent immunoprecipitation followed the Immunoblotting process until the protein lysates were measured and normalised by Bradford assay (Bio-Rad). The immunoprecipitation was carried out incubating cell lysates with the indicated antibodies overnight in TNN buffer on the rotating wheel at 4°C. Protein-A/G sepharose beads (GE Healthcare) were added for 1 hour at 4°C for binding to the incubated antibodies. The beads were spun at 1000rpm for 1 minute to collect proteins targeted by antibodies which were pelleted by binding to the beads. The beads were washed five times with TNN buffer before eluting the samples in SDS loading dye and heated to 95°C to detach and denature captured proteins. The rest followed Immunoblotting process (SDS-PAGE and immunoblotting) to detect immunoprecipitated signals.

Ch2.10 Mass Spectrum sample preparation and analysis

The preparation of mass spectrum samples comprised with two stages. The first step was to immunoprecipitate proteins of interest, and the second part is to digest the proteins.

Mass spectrum samples followed Immunoprecipitation process with overnight antibody incubation replaced with agarose-antibody conjugate (HA or Flag conjugated

agarose) incubation at 4°C overnight. The Flag or HA antibody-conjugated agarose (Sigma) were used to immunoprecipitate ectopically expressed Flag or HA tag. Samples were spun down at 3000rpm at 4°C for 1 minute. The agarose were pelleted and washed five times with TNN buffer before eluted with 0.5 mg/ml 3x-Flag peptide (Sigma) or HA peptide (Sigma) diluted in 0.25M Tris, pH7.4. The peptides completed epitopes binding with agarose conjugated HA or Flag hence the release of immunoprecipitated proteins from antibodies.

At this stage, there are two ways to digest immunoprecipitated proteins. One is to follow in solution digestion (Kessler Lab Proteomics protocols, 2010) and the other is to follow in gel digestion (Kessler Lab Proteomics protocols, 2010).

Ch2.10-1 in solution digestion

The protocol was optimised for a sample volume of 200μl for 10μg total protein. Proteins released by peptide competition in 0.25M Tris (pH7.4) was normalised to 200μl and 5μl of DTT reducing reagent (200mM DTT in 0.1M Tris buffer, Ph7.8) was added and homogenised by vortex followed by 30-60 minutes incubation at room temperature. The iodoacetamide alkylating reagent (200Mm iodoacetamide in 0.1M Tris buffer, Ph7.8) 20μl was added into mixture for another 30-60 minutes room temperature incubation. Samples were subjecting to Methanol/Chloroform Extraction which started with mixing 600μl methanol and 150μl chloroform with samples. After vortex, added 450μl autoclaved ddH_2O

into the mix and the solution became two-phases. Centrifuged the samples at room temperature for 3 minutes The upper aqueous phase was discarded without disrupting the precipitate in the interface (protein extracted) and replaced with 450μl methanol. After centrifuge at room temperature for 3 minutes, supernatant was removed and the protein pellets was suspended in 50μl 6M urea buffer (pH7.8) by sonication and vortex. In order to dilute the urea concentration to a final concentration less than 1M, HPLC grade MilliQ-H_2O was added with vortex. The enzyme Trypsin was added in a 1:50 ration regarding to the total protein content of the sample. The digestion was carried out at 37°C overnight.

Following the Sep-Pak C18 purification protocol, the Sep-Pak column was equilibrated by flushing 2.5ml of solution B (65% CH_3CN, 35% HPLC grade MilliQ-H_2O, 0.1% FA) and then 5ml of solution A (98% HPLC grade MilliQ-H_2O, 2% CH_3CN, 0.1% FA). The Trypsin digested samples were added into the solution A and flew through the Sep-Pak column slowly. Extracted proteins left in the column and washed with 5ml solution A. The elution was collected by adding 600μl solution B twice in separate tubes. The purified samples dried down completely by speed-vac. The samples re-suspended in solution A for mass spectrum analysis.

Ch2.10-2 silver staining and in gel digestion

The other protocol to digest proteins for mass spectrum analysis was in gel digestion protocol. The proteins that released from peptide (Flag or HA) competition after

immunoprecipitation with antibody (HA or Flag) conjugated agarose were then subjected to electrophoresis using NuPAGE® Novex® Gel system (Invitrogen). The NuPAGE Bis-Tris precast gels were used and samples were run 50 minutes at 200V in the 1X NuPAGE® MES SDS Running Buffer. The gel was further subjected to silver staining (Life Technologies Mass Spectrometry-compatible Silver Staining of Proteins in Polyacrylamide Gels, 2012) to visualise proteins of interest. The silver staining based on the chemical reduction of silver ions to metallic silver on a protein band (Rabilloud et al., 1994). The target protein was separated by size, which was indicated by protein ladder, and visualised by silver stain. The target protein (i.e. Flag-E2F1 (1-194)) was cut out into smaller peices from the gel to separate tubes and rinsed the gel pieces overnight in 200 μl washing buffer (50% methanol, 5% acetic acid) on the shaking platform at room temperature.

The gel pieces were washed repeatedly with 200μl washing buffer for 2-3 hours at room temperature and added 200μl acetonitrile to dehydrate for 5 minutes twice, each time removed the acetonitrile completely. To alkylate the gel pieces, 50mM iodoacetamide for 30 minutes. After removed the solution, 200μl acetonitrile added again to dehydrate the gel pieces for 5 minutes and removed acetonitrile completely. The gel pieces were rehydrated in 200μl of 100mM ammonium bicarbonate for 10 minutes and dehydrated with 200μl acetonitrile for 5 minutes. After acetonitrile removed, the gel pieces were dried in a vacuum centrifuge for 3 minutes. The enzyme prepared for digestion had to be kept on ice at all

times, the final concentration of 20ng/μg trypsin 30μl was added for no more than 10 minutes on ice. After centrifuge, the excess of trypsin solution was removed and 10μl of 50M ammonium bicarbonate was added to the samples for trypsin pre-absorbed gel pieces to carry out digestion at 37°C overnight. The final process was to top up 50 μl of 50M ammonium bicarbonate and incubated 10 minutes with gentle vortex. The solution was collected into a separate tube. The next two steps repeated incubating gel pieces (10 minutes x2) with 50μl extraction buffers (buffer 1: 50% acetonitrile, 5% formic acid; buffer2: 85% acetonitrile, 5% formic acid). The extractions were pooled with the initial ammonium bicarbonate extracted solution and dried down completely in a vacuum centrifuge (8 hours or overnight) before re-suspended in 20μl buffer A (98% HPLC grade MilliQ-H_2O, 2% acetonitrile, 0.1% formic acid) (Kessler Lab Proteomics protocols, 2010).

Ch2.10-3 mass spectrum analysis

The mass spectrum analysis was carried out by Dr. Joanna McBane, Dr. Rebecca Konietzny in Professor. Benedikt M Kessler Proteomics laboratory. The “in gel digestion samples,” for examples, Flag-E2F1 (1-194) and Flag-R111/113K E2F1 (KK) samples, were submitted and subjected to analysis by tandem mass spectrometry (LC-MS/MS) (Taylor and Goodlett, 2005) (Batycka et al., 2006) (Fischer et al., 2012). The MS/MS spectrum (mass-to-charge ratio, m/z) of the modified tryptic peptide (protein samples that digested by Trypsin and become peptides) was compared to m/z data at The Universal Protein Resource

(UniProt) (Uniprot.org). The prediction of post translational modification sites was defined principally by m/s differences on particular peptide. For example, the identification of R109 methylation was the mass-to-charge differences (the two methyl groups' difference) of peptide compared to that peptide of human E2F1 (UniProt: Q01094). The "in solution samples" followed similar principle to identify all tryptic peptides available in the solution submitted therefore predicted protein interacting partners from immunoprecipitated samples.

Ch2.11 Flow Cytometry

Cells were harvested and fixed with 50% Ethanol/PBS overnight at 4°C. Fixed cells were washed with PBS and RNase (Qiagen) was added. 0.1mg/ml propidium iodide (PI) (Sigma) was added to the cell and incubated in the dark on ice for 30 minutes before analysis by Flow Cytometer (BD C6 Accuri). With the amount of PI stain detected in each cell, cells in different stages of the cell cycle were identified. G1 phase contains 1X DNA, G2 phase 2X DNA and DNA synthesis phase falls in between. PI was an intercalating agent and as hypodiploidy is a feature of apoptotic cells, Sub-G1 phase indicated cells contain hypodiploidy therefore the implication of the amount of degrading DNA resulted from apoptosis.

Ch2.12 Immunofluorescence staining

Cells were seeded onto the 13 mm glass coverslips. After relevant experimental treatments, cells were washed with PBS and fix with 3.7% formaldehyde for 15 minutes. Cells were permeablised with 0.5% TritonX-100/PBS. The coverslips were then incubated with appropriate primary antibodies for 1 hour followed by washes with 0.025% Tween-PBS before the fluorochrome conjugated secondary antibodies supplied for 30 minutes at the room temperature. The coverslips were mounted with a DAPI containing reagent to permit visualisation of cell nuclei. The protein localisation was visualised by BX60 transmitted-reflective light fluorescence microscope (Olympus) and analysed with Openlab software (Improvision).

Ch2.13 Colony Formation

Cells were seeded in 6-well plate and transfected with intended DNA plasmids (pcDNA3.1 vector, HA-WT-E2F1, HA-R109K-E2F1, HA-R111/113KK-E2F1) for 48 hours before G418 (800μg/ml) containing medium (DMEM) select cells for 5-7 days to stably induced plasmids contacting cells due to their G418 resistance. Selected cells were trypnised and re-seeded for 1000 cells per well for 10 days. The culturing medium was gently aspirated and the plates were rinsed with cold PBS and cells were fixed with ice-cold methanol (stored at -20°C) for 10 minutes on ice. Carefully aspirated methanol and applied

3 ml Crystal violet stain to the cells and incubated at room temperature for 10 minutes. After collected stain from the cells, the cell plates were rinsed in ddH_2O and left to air-dry. The scanning was done with a Gelcount Colony Counter (Oxford Optronics).

Ch2.14 Chromatin immunoprecipitation (ChIP), sample preparation for ChIP-sequencing and double ChIP

Ch2.14-1 ChIP

After relevant experimental process, for example, 48 hours after DNA plasmids transfection, 72 hours after siRNAs treated cells, cells subjected to chromatin immunoprecipitation were cross-linked by 3.7% formaldehyde addition for 10 minutes with 0.125M glycine neutralisation in the end. Cells were washed with cold PBS before harvested to lyse in ChIP Lysis buffer 1 (5mM Tris-HCl pH8.0, 85mM KCl, 0.5% Igepal) containing protease inhibitors. The cell pellets were re-suspended in ChIP Lysis buffer 2 (10mM Tris-HCl pH7.4, 150 mM NaCl, 1 mM EDTA, 1% Igepal, 1% deoxycholic acid sodium salt (DOC) and 0.1% SDS) containing protease inhibitors for 5 minutes and the cell lysates were sonicated 10 times at 25% power for 10 seconds (Branson Digital Sonifier 450). The cell lysates were then pre-cleared 3 hours with pre-blocked protein A and G agarose beads (1mg/ml BSA and 400μg/ml salmon sperm DNA) containing protease inhibitors and non-specific antibody at 4°C. The step pre-cleared non-specific binding of

lysates to non-specific antibody or to agarose beads. The lysates were spun down at 1000rpm for 5 minutes and the agarose (and antibody attached) were discarded. The lysates were immunoprecipitated with 1-2μg either non-specific or specific (i.e. anti-E2F1 or anti-DP-1) antibody overnight at 4°C. Pre-blocked agarose A/G beads were added for 2 hours at 4°C before the beads were washed 4 times with washing buffer 1 (20mM Tris-HCl pH8.0, 150mM NaCl, 2mM EDTA, 0.1% SDS and 1% Trixon-100(Sigma)), 2 times with washing buffer 2 (10mM Tris-HCl pH8.0, 250mM LiCl, 1mM EDTA, 1% Igepal and 1% DOC) and final 1 wash with TE buffer (10mM Tris-HCl, 1mM EDTA). The Protein/DNA immunocomplexes were eluted in 1% SDS and 0.1$NaHCO_3$ elution buffer twice on the rotating wheel at room temperature and the elution were pooled. If no double ChIP (Ch2.14-3) intended, the cross-link was broken by adding 200mM NaCl, 10mM EDTA, 40mM Tris-HCl pH6.5, 40μg/ml proteinase K and 20μg/ml RNAse A, 3 hours at 55°C followed by overnight at 65°C. The immunoprecipitated chromatin was then purified following Qiaquick PCR kit (Qiagen). Input DNA was purified form the non-specific antibody control supernatant.

Ch2.14-2 ChIP-sequencing grade ChIP sample preparation

Two areas optimised from chromatin immunoprecipitation (ChIP) protocol were, firstly the ChIP process, and secondly the fragmentation of ChIP DNA.

I. ChIP process: the change of blocking buffer, magnetic beads:

Instead of the pre-blocked agarose protein A/G beads, ChIP-seq grade ChIP samples were immunoprecipitated by magnetic Dynabeads® Protein G beads (Invitrogen) as the alternative throughout the entire process. The magnetic Dynabeads were pre-blocked by 5mg/ml BSA without salmon sperm that standardised in the agarose beads to avoid DNA contamination from fish. At the pre-clear stage, while 1μg non-specific antibody and 5-10μl magnetic Dynabeads were incubated with lysates, the additive step was that a separate 20-25μl magnetic Dynabeads were in parallel incubated with non-specific (i.e. rabbit IgG) or intended specific antibody (rabbit polyclonal E2F-1 (C20) (Santa Cruz) in the 5mg/ml BSA solution to prepare the antibody-Dynabeads complex. After 3 hours of pre-cleared non-specific antibody and magnetic Dynabeads with the lysates, the pelleted Dynabeads were applied onto magnetic rack and supernatant were separated and collected for overnight immunoprecipitation at 4°C with the prepared antibody-Dynabeads complex.

II. ChIP fragmentation:

In order to optimal DNA fragmentation result to the size smaller than 1000 base pair (bp), the sonication process has been tested for time and concentration factors (demonstrated in Figure 5.3), 20 cell culture petri dishes (150mm) with HeLa cells grown to 70% confluency after 72hours siRNAs treatment were best optimised to divide to 5 batches of sonication for 10 times at 25% power for 10 seconds (Programme set up at

Branson Digital Sonifier 450). The rest procedure referred to standard chromatin immunoprecipitation process (Ch2.14-1 ChIP).

Ch2.14-3 Double ChIP

The double ChIP was a secondary ChIP sequentially immunoprecipitated from primary ChIP materials. After overnight immunoprecipitation and washes, the immunocomplexes were eluted by elution buffer (1% SDS and 0.1$NaHCO_3$) and instead of the reverse cross-link; the elutions were subjected to a second round of immunoprecipitation. For example, a double ChIP was carried out to investigate chromatin that bound to the asymmetrically dimethylated E2F1 and the symmetrically dimethylated E2F1, 15 full confluency U2OS cells seeded in 10mm cell culture dishes were harvested and the non-specific (IgG) and specific (E2F1) antibodies with agarose protein A/G beads were incubated for immunoprecipitation and washes, these two elutions (500μl/each sample) were divided into four portions for non-specific, ADMA, SDMA, and DP-1 antibodies to further immunoprecipitate from. The primary elution 125μl was diluted 1:8 by topped up 875μl of double ChIP lysis buffer, the Lysis buffer 2 without the SDS and DOC component (10mM Tris-HCl pH7.4, 150 mM NaCl, 1 mM EDTA, 1% Igepal) and incubated overnight at 4°C. The aimed of the step was to capture specific set of chromatins bound to E2F1 that was asymmetrically dimethylated (ADMA), symmetrically

dimethylated (SDMA) and DP-1 dimerised (positive control) (Cress and Nevins, 1994) with a secondary non-specific control (negative control) included.

Primary ChIP (500μl)	**Secondary ChIP** (1000μl= 125μl primary ChiP/875 lysis buffer)
IgG (125μl)	IgG
IgG (125μl)	ADMA
IgG (125μl)	SDMA
IgG (125μl)	DP-1
E2F1 (125μl)	IgG
E2F1 (125μl)	ADMA
E2F1 (125μl)	SDMA
E2F1 (125μl)	DP-1

The reverse cross-link step helped disassociate chromatin from immunoprecipitated proteins and the purification of chromatin by Qiaquick PCR kit (Qiagen) were identical with ChIP protocol (Ch2.14-1).

Ch2.15 Semi-quantitative Poly Chain Reaction (PCR) and quantitative PCR (qPCR)

The Qiaquick PCR kit (Qiagen) purified immunoprecipitated chromatin was then subjected to amplification by poly chain reaction (PCR), either semi-quantitative to

visualise by DNA gels or quantitative PCR (qPCR). Input DNA was purified form the non-specific antibody control supernatant.

The chromatin was analysed by PCR using 1.7U Taq DNA polymerase (Stratagene), 0.5mM dNTP (Promega) in PCR buffer (75mM Tri-HCl H8.8, 20mM $(NH_4)_2SO_4$, 0.01% Tween-20) following setting as follows: initial denaturation at 95°C for 5 minutes, 30 cycles 95°C for 3 seconds, 61°C annealing temperature for 1 min and 72°C for 1 min, and the final elongation 72°C for 5 minutes. PCR products were visualised on 2% agarose gel with 0.5μg/ml ethidium bromide. The primers used as follows:

Primers	**Sequence**
Albumin Forward	5'- TGGGGTTGACAGAAGAGAAAAGC-3'
Albumin Reverse	5'- TACATTGACAAGGYCTTGTGGAG-3'
E2F1 Forward	5'- AGGAACCGCCGCCGTTGTTCCCGT-3'
E2F1 Reverse	5'- GCTGCCTGCAAAGTCCCGGCCACT-3'
Cdc6 Forward	5'- GGCCTCACAGCGACTCTAAGA-3'
Cdc6 Reverse	5'- CTCGGACTCACCACAAGC-3'
Apaf1 Forward	5'- CTTGGCCAGGCTGGTCTTGAAT-3'
Apaf1 Reverse	5'- GCAGCCATTCAAAATTATGACACAT-3'

For real time quantitative PCR, the template DNA was amplified using the Brilliant III SYBR Green QRT-PCR master mix (Stratagene) according to the manufacturer's protocol

on a MX3005P QPCR system (Stratagene). The dissociation curve analysis was used to check for specific amplification and the data analysis was performed on the MXPro QPCR software. The relative quantities were plotted by CT value obtained (threshold value) against serial dilution of positive control sample (input). The data represented the mean values from three reactions and are shown as fold change, relative to the untreated sample (input), with errors bars indicating standard deviation.

Ch2.16 ChIP-sequencing analysis and validation

Ch2.16-1 ChIP-sequencing bioinformatics analysis

The ChIP-Sequencing bioinformatics analysis was carried out by Dr. Robert Siddaway (Ludwig Institute for Cancer Research, University of Oxford). The raw reads were aligned to human genome build hg19, using Bowtie 0.12.8, and duplicate reads discarded to remove PCR artefacts, using Picard Tools (version 1.96, http://picard.sourceforge.net). Aligned read files were randomly down-sized to the size of the smallest dataset to facilitate downstream analysis. For the "peak calling," peaks were identified using MACS 2.0.10 (Zhang et al., 2008) and the default settings were with a q-value cut-off of 10^{-5}. The q-value is an adjusted p-value based on the false-discovery rate (Storey, 2002)). For the "peak annotation," peaks were annotated as being linked with a gene if they lay either within the gene body of a RefSeq gene, or within 20 kb of the transcription start and end sites of the gene, using CisGenome 2.0 (Ji et al., 2008). Categorisation of peaks as falling into introns,

exons, 5’ and 3’ untranslated regions, core and distal promoters, or intergenic regions was performed with Cis-regulatory Element Annotation System (CEAS; version 1.0.2) (Shin et al., 2009). Differentially bound peaks were identified using the “bdgdiff” function of MACS 2.0.10, using default settings and the “bedgraph” (graph in .bed format) and peak files output from the initial peak detection.

Ch2.16-1 ChIP-sequencing validation

The validation of differential binding of ChIP-sequencing analysis was performed with semi-quantitative PCR (Ch2.15) on 2% agarose gel containing 0.5μg/ml ethidium bromide and visualised under UV. The primers used as follows:

Gene	**Forward**	**Reverse**
CAMTA1	5’-TGGGAATTTACGGTGCTGTG-3’	5’-ACCCTCTGAGAAAGCCGAC-3’
XRCC5	5’-GCGGTCGGGGAATAAGGTAT-3’	5’-GCCCCAAATCTCCAGCTTTT-3’
NUP155	5’-ATCTCAGCATATCCTCGCCG-3’	5’-GCGGCGGATCTAAGCTAACT-3’
GATA1	5’-CTGAGGTGATGGAGTGGGAG-3’	5’-TCCCAGGTCTGCACTTAACT-3
C9orf100	5’-CTCGGTCTCTAGCAGCTCC-3’	5’-CAACCTTGCCCTGCCCTG-3’
RGS3	5’-GTCCTGCTGGCTCCTTGG-3’	5’-GGTGGTGAGGTAGAAGGACA-3’
STMN1	5’-GAGAATTAGACGCAGCACCG-3’	5’-CACCCGCTGAACTGTGAAG-3’
BLM	5’-CGGACTCTGATTGGGCTTTG-3’	5’-CAGACTAGAGGACTCAGCGG-3’

Ch2.17 Quantitation and statistical analysis

The quantitation was done with image processing software Imgae J (National Institutes of Health, NIH) as it calculated image area and pixel value statistics. Differences between experimental groups were evaluated for statistical significance (* = $p<0.05$; ** = $p< 0.01$) using Student's *t-test* for unpaired two group comparison. P value less than 0.05 were considered to be statically significant.

Chapter 3 E2F1 methylation by PRMT1

Ch3.1 Introduction

Since the discovery of PRMT as histone arginine methyltransferases, many non-histone substrates have been identified and characterized. Arginine methylation on non-histone substrates is an emerging area of interest as many of these substrates play important roles in various cellular functions (Table3.1).

The transcriptional factor E2F1 has been recently identified as a novel PRMT5 substrate with symmetric di-methylarginines (SDMA) detected on R111 and R113 (Cho et al., 2012). Through motif sequence alignment, the RGRGR short sequence was found to be present in E2F1. This sequence is highly similar to the RGRER sequence found in p53, where PRMT5-induced arginine methylation occurs (SDMA on arginine 333, 335 and 337 of p53) (Jansson et al., 2008). Many other substrates also contain the PRMT motif (Bedford, 2007), namely the glycine- and arginine- rich patches (GAR motif) (Figure 3.1).

Since E2F1 possesses a PRMT GAR motif and is a substrate of PRMT5, it is possible for other members of PRMT family to target it as a substrate. One of such examples is substrate histone protein H4. Both PRMT1 and PRMT5 are capable of targeting histone protein H4 and generate asymmetric or symmetric di-methyl marks on the arginine 3 (H4R3) respectively (Strahl et al., 2001) (Zhao et al., 2009).

PRMT5-induced methylation (SDMA) on histone H4 arginine 3 and H3 arginine 8 has previously been shown to trigger the transcriptional *silencing* of genes (Karkhanis et al.,

2011). PRMT1 mediated methylation (ADMA) on the same substrate, H4R3, is important for transcriptional *activation* by allowing subsequent acetylation on histones (Huang et al., 2005) (Zhao et al., 2009) .

The studies outlined in this chapter, set out to investigate whether E2F1 is a substrate of PRMT1, a chief member of type I PRMT. The aim is to hypothesise the potential PRMT1 resultant ADMA, together with the previously reported PRMT5 resultant SDMA on E2F1 as the underlying regulation mechanism which contributes to the opposing effect of E2F1.

PRMTs member	Known histone substrate	Known non-histone protein substrates
PRMT1 (typeI)	H4R3	hnRNP1 A1(R194), BTG1, TIS2, IFNα/β, ILF3, SPT5(R681,696,698), SAF-A, 53BP1(GAR motif),MRE11 (GAR motif), FMRP (R544, 546), Sam68, SLM(GAR motif), ERα（R260), RUNX1 (R206, 210), TAF15 (RGG repeats), BCR (R198), CFIm59 (GAR motif), CFIm68, Ash2L (R296)
PRMT5 (typeII)	H4R3, H3R8, H3R2, H2AR3	MBP (R107), LSm4, SmD1(GAR motif), Sm D3 (GAR motif), EBNA-2 (R325-376), SPT5 (R698), EBNA-1 (R325-376), RBL2, p53 (R333,335,337), CBP-1 (R234), CF Im68 (GAR motif), Ash2L (R296), PDCDC4 (R110), HoxA (R140), NFkB (R30)

Table3. 1 Arginine methylation by PRMT1 and PRMT5 on histone and non-histone proteins (Wei et al., 2014)

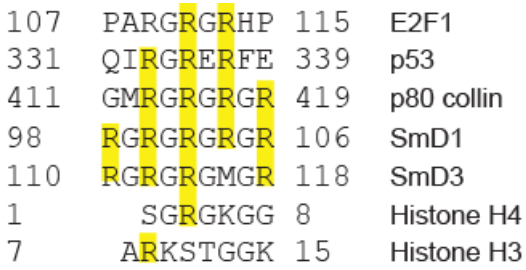


Figure3. 1 PRMT5 substrates' sequence motif alignment

Comparison of PRMT5 arginine methylation sites in E2F1 to other proteins. Symmetrically dimethylation sites were highlighted in yellow. The PRMT glycine (G)- and arginine (R)-rich motif demonstrated in many PRMT5 substrates.

Ch3.2 E2F1 as a substrate of PRMT1 *in vitro*

As E2F1 contains a PRMT binding motif and is also a substrate for PRMT5 (type II PRMT) (Figure3.2A), it is interesting to test whether PRMT1, the predominant member of type I PRMT, can also methylate E2F1.

To test this idea, the recombinant E2F1, wild-type (WT) and enzymatic dead PRMT1 proteins were purified from *E.coli* and co-incubated in the presence of radioactive tritium labelled SAM (H^3-SAM) using a filter system to measure *in vitro* methylation activity. The incorporation of H^3 was measured by a scintillation counter in disintegrations per minute (DPM). The DPM count indicates the occurrence of methylation.

The methyltransferase activity of PRMT1 was abrogated using site-directed mutagenesis by inserting a glycine (G) to arginine (R) mutation at position 80 of the enzyme's methyltransferase (MTase) region (Zhang and Cheng, 2003). The E2F1 derivatives were generated with the charge and structure neutral arginine (R) to lysine (K) mutation (Figure3.2B). These constructs were sub-cloned into the glutathione *S*-transferase (GST) containing pGEX-4T1 vectors, therefore the transformed and purified proteins were GST tagged (Figure3.2C). The sub-cloning was enabled by the designed flanking restriction digestion enzyme sites (BamHI and NotI). These restriction digestions sites were

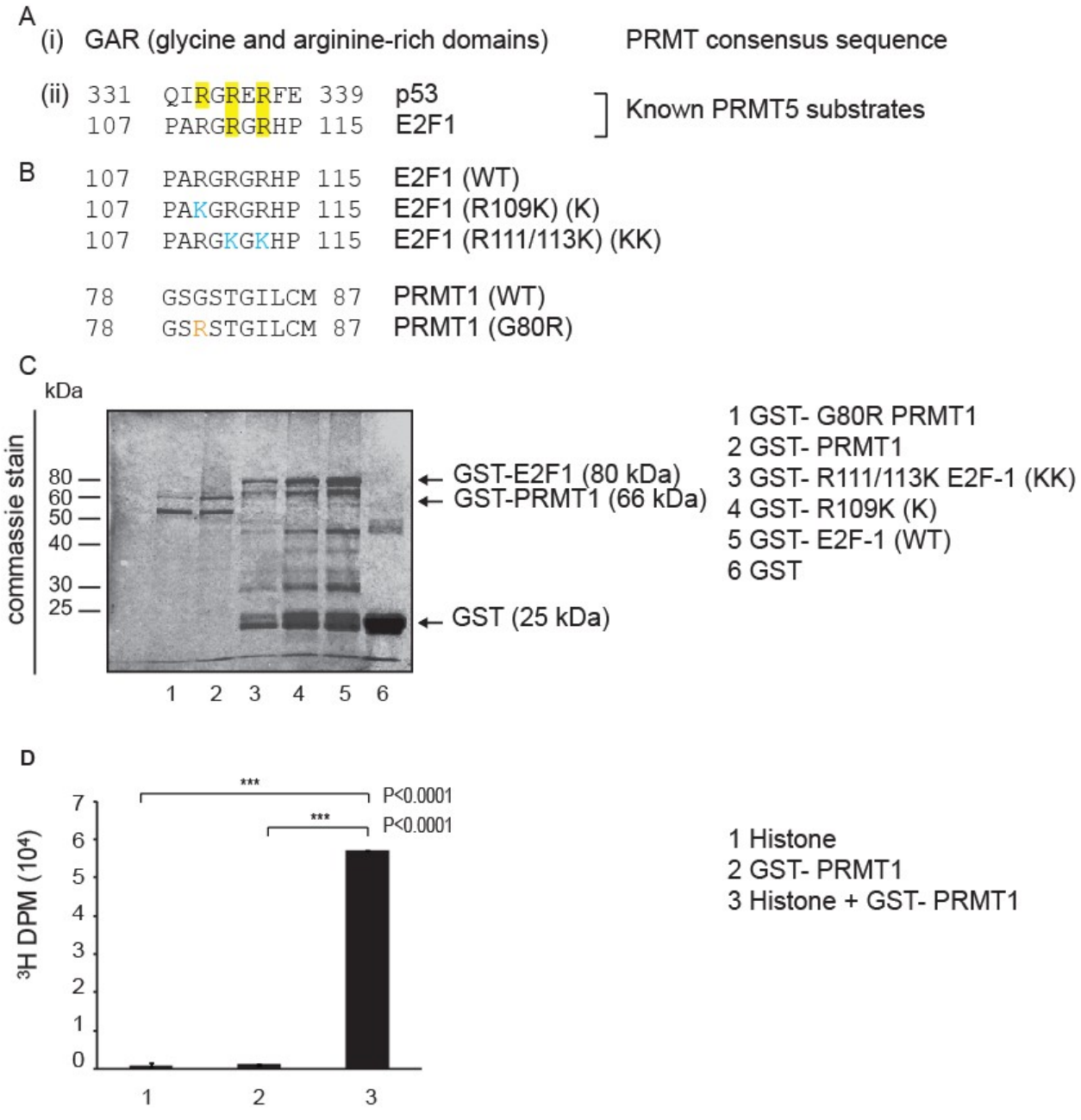
also present in the pcDNA3.1 and pTRE-Hyg vectors that were used for mammalian cell transfection and the Tet-on inducible cell systems respectively.

The enzyme activity of purified recombinant GST-PRMT1 was measured by co-incubation with a known substrate, histone protein, in the presence of radioactive H^3-SAM. The results of this experiment indicated that the purified PRMT1 is enzymatically active as the co-incubation between PRMT1 and histone accounted for a statistically significant H^3 DPM count (Figure3.2D lane3) compared to that of the substrate or enzyme alone controls (Figure3.2D lane1 and 2).

To test if E2F1 is a substrate of active PRMT1 *in vitro*, the GST tag alone was used as a substrate to serve as negative controls. When GST tag was incubated with GST tagged WT PRMT1 enzyme or with GST tagged G80R PRMT1, H^3 the DPM count remained constant at the background level (Figure3.2E lane1-3). The Set 7/9, a protein lysine methyltransferase, was used as a positive control as it methylated K185 on E2F1 (Kontaki and Talianidis, 2010) (Figure3.2E lane7). It should be noted that both the enzymes and the substrates were GST tagged. The possibility of dimerization of the GST tags of GST-PRMT1 and GST-E2F1 is addressed by using a HA tagged PRMT1 in the next set of experiments.

The results showed significantly higher H^3 DPM counts when PRMT1 was co-incubated with E2F1 compared to with GST tag (Figure3.2 lane 2 and 5). This suggests

E2F1 itself, not the tag, is a potential substrate of PRMT1. The results also showed significantly higher H^3 DPM counts measured when WT PRMT1 was co-incubated with E2F1 compared to G80R PRMT1, whose catalytic activity had been abrogated (Figure3.2E lane 5 and 6). This suggests the methyltransferase ability of PRMT1 is responsible for this methylation event. Additionally, when repeated this set of experiments with enzyme to substrate dilution (1:1, 1:2 and 1:4), the pattern remained similar and the methylation event measured remained statistically significant (Figure 3.2E lane 1-7 repeats). These attempts tapped into basic enzyme kinetics. The aim is to confirm the methylation detected is not due to the saturation of enzyme to substrate but bona fide interaction between enzyme and substrate proteins.


A
(i) GAR (glycine and arginine-rich domains)
PRMT consensus sequence
(ii) 331 QIRGRERFE 339 p53
107 PARGRGRHP 115 E2F1
Known PRMT5 substrates
B
107 PARGRGRHP 115 E2F1 (WT)
107 PAKGRGRHP 115 E2F1 (R109K) (K)
107 PARGKGKHP 115 E2F1 (R111/113K) (KK)
78 GSGSTGILCM 87 PRMT1 (WT)
78 GSRSTGILCM 87 PRMT1 (G80R)
C
kDa
commassie stain
80
60
50
40
30
25
GST-E2F1 (80 kDa)
GST-PRMT1 (66 kDa)
GST (25 kDa)
1 2 3 4 5 6
1 GST- G80R PRMT1
2 GST- PRMT1
3 GST- R111/113K E2F-1 (KK)
4 GST- R109K (K)
5 GST- E2F-1 (WT)
6 GST
D
3H DPM (10⁴)
7
6
5
4
3
2
1
0
***
P<0.0001
***
P<0.0001
1 2 3
1 Histone
2 GST- PRMT1
3 Histone + GST- PRMT1

E

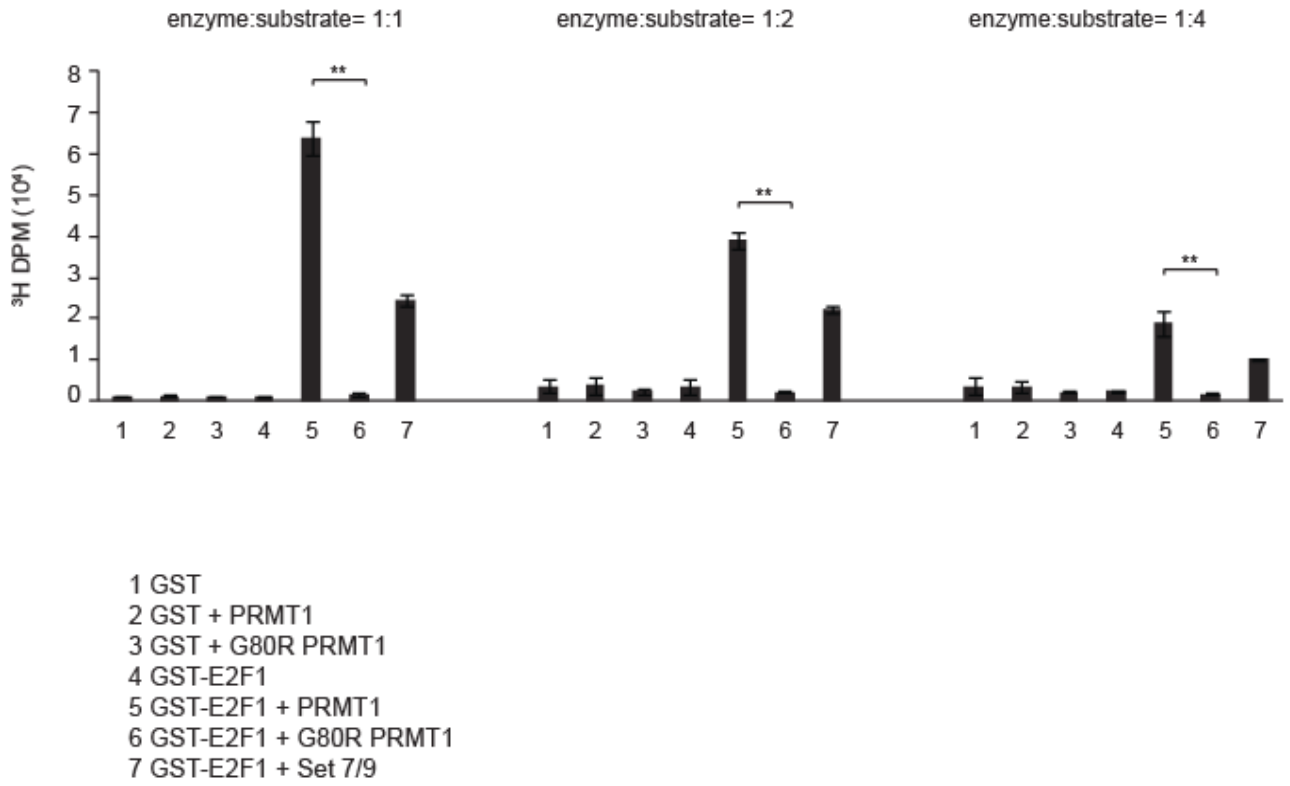


Figure3. 2 PRMT1 methylates E2F1 *in vitro*

(A) (i) PRMT consensus motif sequence (ii) arginine methylation sites (highlighted in yellow) of PRMT5 substrates. (B) Wild-type sequences of E2F1, PRMT1 and their derivatives by mutagenesis (highlighted in colour). The arginine (R) to lysine (K) mutation at E2F1 position 111 and 113 (R111/113K or KK) erased PRMT5 methylation sites on E2F1. The remaining arginine in the ARG patch was also mutated from R to K (R109K or K). PRMT1 catalytic MTase domain was abrogated by glycine (G) to arginine (R) site mutagenesis at position 80 of PRMT1. (C) GST tagged recombinant proteins were purified, separated by size (indicated with arrow) using gel electrophoresis and subjected to Commassie staining for visualisation. (D) 1μg bacteria purified recombinant PRMT1 and 1μg histone, known substrate of PRMT1, were tested alone (lane1,2) or together (lane3) in the *in vitro* methylation assay The incorporation of ^{3}H containing methyl group was measured by a scintillation counter in disintegrations per minute (DPM). Data shown is best representative of 3 independent experiments; $p<0.0001$ (E) 1μg bacteria purified recombinant GST-E2F1 were incubated with 1μg or 500ng, or 250ng of enzymes, such as GST tagged wild-type PRMT1, enzymatically dead G80R PRMT1 or positive control GST tagged Set7/9 in the *in vitro* methylation assay. The methyltransferases activity was measured in DPM. Data shown is most representative of 3 independent experiments.

Ch3.3 E2F1 as a substrate of PRMT1 *in vivo*

While E2F1 arginine methylation by PRMT1 is seen *in vitro*, it is crucial to test if this methylation event occurs *in vivo*. The human osteosarcoma U2OS cell lines were used. One of which is an engineered PRMT1 inducible Tet-On U2OS cell line (inP1), whose PRMT1, tagged with human influenza hemagglutinin (HA), is induced to express upon doxycycline (Dox) treatment (details in chapter 4 inducible cell lines generation). The inducible HA-PRMT1 U2OS cell line is a good tool to monitor changes in cells in the presence of PRMT1.

The idea is to investigate E2F1 asymmetric arginine dimethylation by PRMT1 i*n vivo*. The existence of asymmetrically dimethylated E2F1 was assessed by detection of ADMA mark on E2F1 which were immunoprecipitated from cells. To probe further into the relationship between PRMT1 and asymmetrically dimethylated E2F1, two conditions: when the PRMT1 was induced to express or was depleted, were tested.

In the presence of Dox, the level of HA labelled PRMT1 increased (Figure3.3A input). Both uninduced (Dox negative) and induced (Dox positive) cells were previously transfected with Flag tagged E2F1 and the E2F1 were immunoprecipitated for ADMA mark detection.

Upon induction of HA-PRMT1 by Dox treatment, the ADMA mark was detected on immunoprecipitated E2F1 (Figure3.3A lane 2 and 4) which indicates the existence of

asymmetrically dimethylated E2F1 in cells. This result also demonstrates that an inducible ADMA signal on E2F1 corresponds with the induction of PRMT1 (Figure3.3A).

In a similar but reverse fashion, when comparing PRMT1 and non-targeting siRNA transfected cells (siP1 vs. NT), the endogenous PRMT1 were depleted or left intact (Figure3.3B input). Under the siP1 treated condition, the ADMA mark detected on immunoprecipitated E2F1 was observed to decline (Figure3.3B lane 2 and 4). This result suggests the loss of the ADMA signal on E2F1 corresponds to the depletion of endogenous PRMT1. The doublet observed in the E2F1 blot might be addressed with phosphate for a possible phosphorylation modified form (Figure 3.3B).

These observations, the inducible ADMA upon PRMT1 induction and the loss of ADMA upon PRMT1 depletion, reveal a strong correlation relationship between E2F1 ADMA signal and PRMT1. This elucidates the PRMT1 dependant nature of ADMA mark of E2F1 which supports that PRMT1 is responsible for the asymmetric dimethyl arginines of E2F1 in cells.

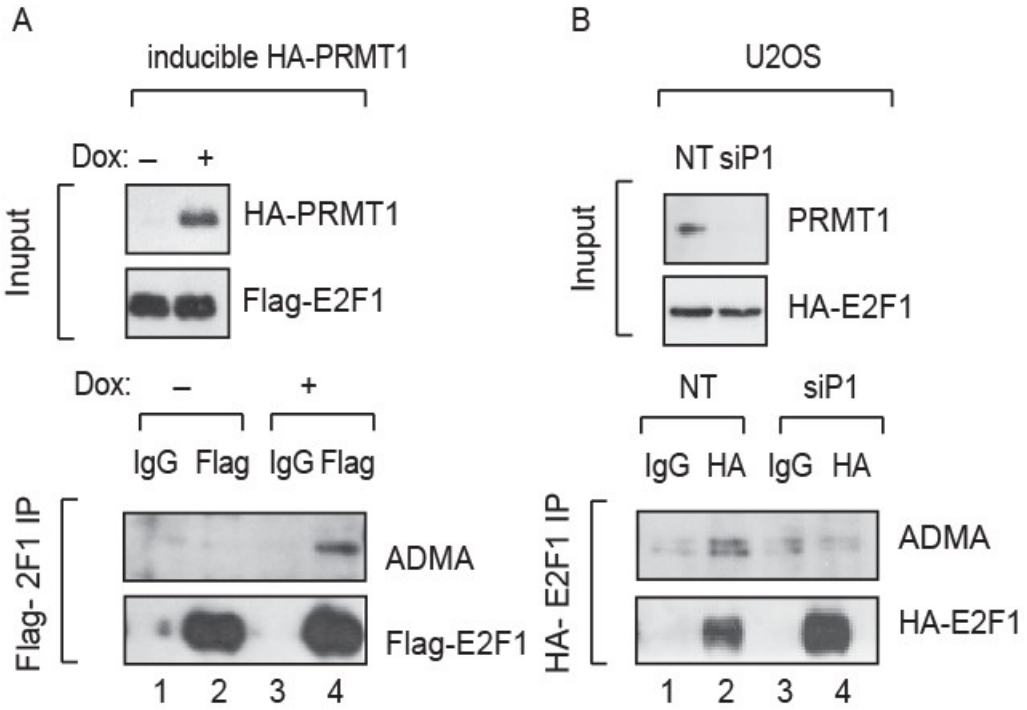


Figure3. 3 PRMT1 methylates E2F1 in cells.

(A) Inducible PRMT1 cells were treated with or without 1μg/ml doxycycline and transfected with 1μg Flag-E2F1 for 48 hours before harvest for Flag immunoprecipitation. The input were probed with Flag and HA antibodies for Flag-E2F1 transfection and HA-PRMT1 induction confirmation. The immunoprecipitaion(IP) was done with 1μg non-specific IgG and 1μg Flag antibodies and the IP materials were probed for Flag and ADMA antibodies to confirm the amount the ADMA signal on immunoprecipitated Flag-E2F1. (B) The human osteosarcoma U2OS cells were treated with 50nM non-targeting (NT) or PRMT1 siRNA (siP1) for 72 hours. The 1μg HA-E2F1 was transfected after first 24 hours siRNAs treatment thus remained 48 hours before harvest for immunoprecipitaion. The input was probed for PRMT1 and HA antibodies for endogenous PRMT1 level and HA-E2F1 transfection. The IP materials were probed with ADMA and HA antibodies to access the ADMA signal on immunoprecipitated HA-E2F1.

Ch3.4 PRMT1 methylates R109 on E2F1

As PRMT1 was shown to methylate E2F1 both *in vitro* and *in vivo*, it is important to map its actual methylation site(s) for further dissection of the effects of PRMT1-induced arginine methylation on E2F1. Traditionally, the mapping of modification site(s) has been done by site/region mutagenesis of the entire protein sequence, coupled with *in vitro* assays to test the potential modification sites which can then be validated in cells.

On the basis of previous work with PRMT5/E2F1 and preliminary *in vitro* data on PRMT1/E2F1, the approach adopted here is mass spectrometry (MS) directly using materials from the cells. The MS helps to screen modification of protein of interest, biochemically, in an amino acid by amino acid fashion. Mass spectrometry based proteomics analysis was done with the aim of investigating biochemically the location of arginine methylation modification by PRMT1 on E2F1 in cells.

Ch3.4-1 Samples preparation for mass spectrometry

The simple breakdown of the MS procedure is firstly ionizing the samples, this breaks samples into charged fragments, and secondly, producing their mass-to-charge ratio profile, this is done by accelerating fragmented samples and subjecting them to an electric or magnetic field where elements of the sample behave differently according to its chemical properties. This, collectively, produces an 'intensity versus m/z (mass-to-charge)' plot, the mass spectrum of a sample. By correlating to the known masses (or m/z), samples were identified and its chemical modifications addition can thereafter computed.

There are two obstacles to overcome in preparing samples for mapping methylation site in cells using this approach. There are namely, firstly, to supply enough amount of sample, and secondly, to ensure coverage of the target sequence by selected enzyme.

To investigate the methylation sites in cells, PRMT1 and E2F1 were ectopically expressed in cells. The E2F1 was therefore eluted in the presence of high levels of PRMT1 in a physiologically relevant environment. The optimisation of the MS experiments includes the use of short E2F1 (amino acid 1-194), instead of full length E2F1. This doubles the amount of sample obtained from the cells as a result of enhanced transfection and elution efficiency (Figure3.4A(i) and (ii)). The short E2F1 (aa.1-194) contains the glycine-and-arginine rich PRMT consensus motif (109 RGRGR 113) and the intact DNA binding domain of E2F1.

The other optimisation steps focused on the choice of enzyme to breakdown the eluted E2F1. Although different enzymes are predicated to be capable of digesting E2F1 into suitable fragmentation sizes (using PeptideCutter, ExPAsy) with their own cutting preferences that kept RG rich fragment intact. In order for the mass spectrometer to capture and analyse, four such attempts: endoproteinase Lys-C, Chymotrypsin, the combination of Lys-C and Chymotrypsin (Figure3.4B(i) and(ii)), or the non-specific enzyme Elastase (no cutting preference), were unsuccessful in fragmenting the target, eluted E2F1. Paradoxically, the use of Trypsin improved the fragmentation success albeit it is not the ideal enzyme for this application. The enzyme's lysine (K), arginine (R) targeting preference would be expected to fragment E2F1 too extensively at the sites of interest (RG rich PRMT motif) and for retention by the mass spectrometer (Figure3.4B(iii)).

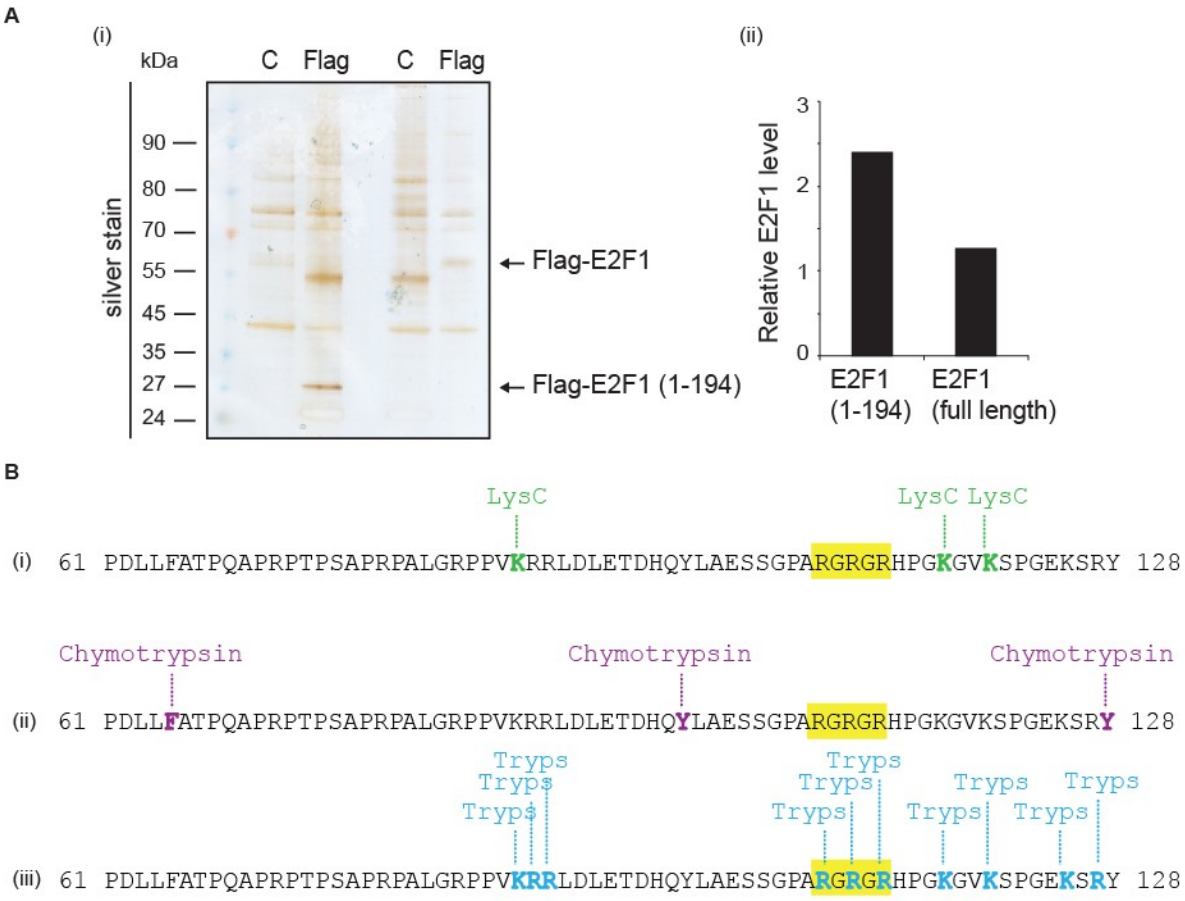


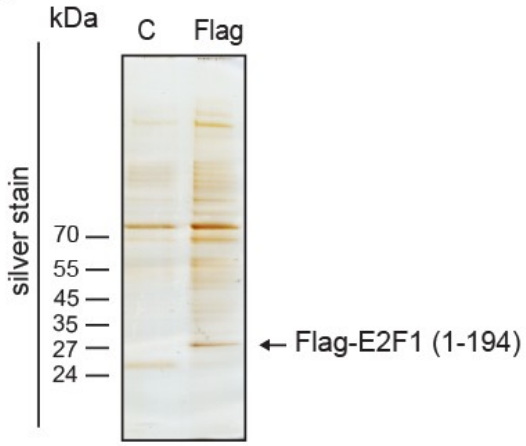
Figure3. 4 Optimisation of Mass Spectrometry samples preparation

(A) (i)Human U2OS cells were transfected with 1μg control empty vector (C) Flag-E2F1 or 1μg Flag-E2F1 (1-194) at the scale of 15 cell petri dishes (10mm) with 70-80% confluency for 48 hours. Cells were harvested, immunoprecipitated with Flag conjugated agarose and eluted with 30μl, 0.5 mg/ml 3x-Flag peptide in 0.25M Tris, pH7.4. The elutions were subjected to electrophoresis using NuPAGE® Novex® Gel system with NuPAGE Bis-Tris precast gels in the 1X NuPAGE® MES SDS Running Buffer for 50 minutes at 200V. The silver staining was used to visulaise proteins following manufacture's protocol. The arrow indicated Flag E2F1 (full length or 1-194) identifed by size and the comparsion to control. The visulaised E2F1 was cut out from the gel into smaller pieces from the gel. (A)(ii) The amout of short (1-194) or full length Flag-E2F1 were quantified from the silver statining gel by Image J. (B) Schematic E2F1 61-128 sequence showed PeptideCutter (ExPasy) prediction on enzyme cutting sites. Enzymes and relavant cutting sites were highlighted with correponsing colour (LysC in green, Chymotrypsin in purple, Trypsin (Tryps) in blue) and PRMT consensus motif RG rich patch on E2F1 sequence were highlighted in yellow.

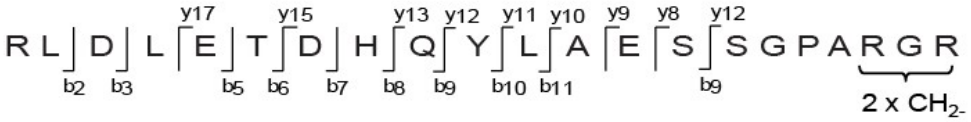
Ch3.4-2 E2F1 R109 arginine dimethylation by PRMT1

With PRMT1 and E2F1 (aa. 1-194) co-expressed in the cells, the eluted E2F1 was digested with Trypsin following the standard ‘in gel digestion protocol (chapter2 mass spec sample preparation).' The MS proteomic analysis identified eluted protein as E2F1 and also assigned two methyl groups to the region 105 SGPARGR 111 in peptide 91-111 according to its mass-to-charge ratio versus relative abundance plot (Figure3.5A(ii)). This means that compared to the reference m/z of the peptide 91-111, the digested E2F1 peptide 91-111 was two methyl groups (labelled as 2 x CH_3-) ‘heavier’ and ‘more charged.’ More specifically, the two methyl groups were assigned in the peptide fragment of 105-111 where only two arginines were found, arginine 109 or 111 as potential PRMT site(s) to catalyse methyl groups on. This suggests possible dimethylation on either R109 or R111 of E2F1 or monomethylation on R109 and R111.

A (i)


kDa
C
Flag
silver stain
70
55
45
35
27
24
← Flag-E2F1 (1-194)

(ii)


R L D L E T D H Q Y L A E S S G P A R G R
y17 y15 y13 y12 y11 y10 y9 y8 y12
b2 b3 b5 b6 b7 b8 b9 b10 b11 b9
2 x CH2-

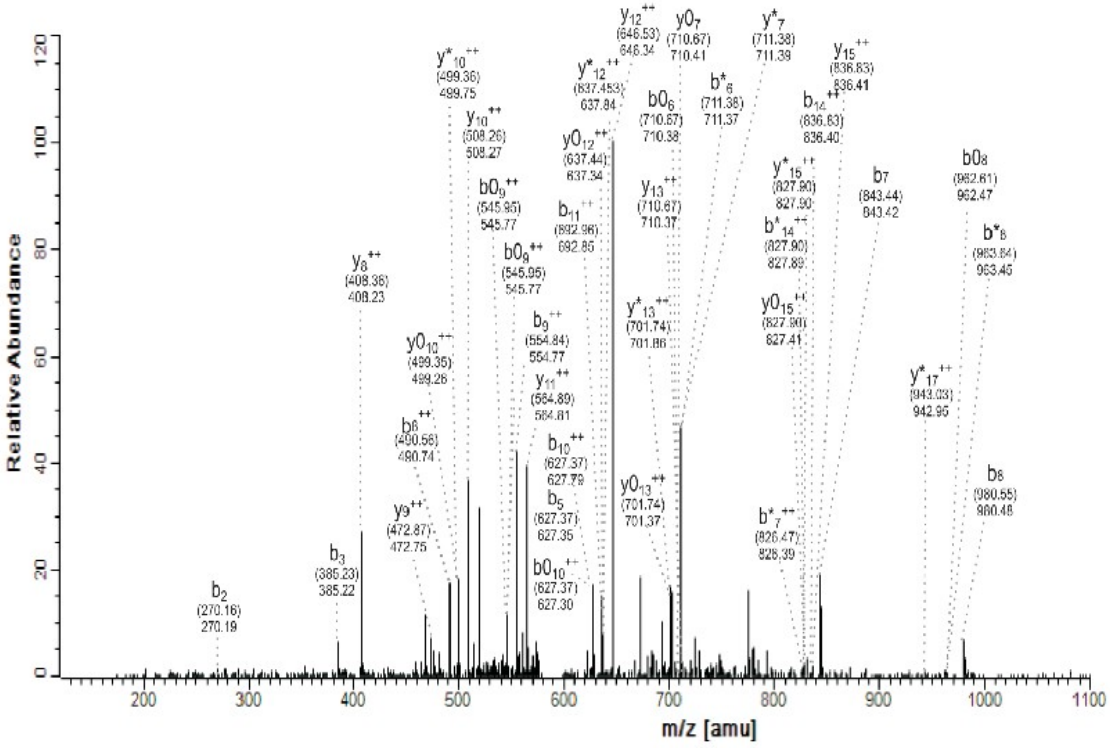
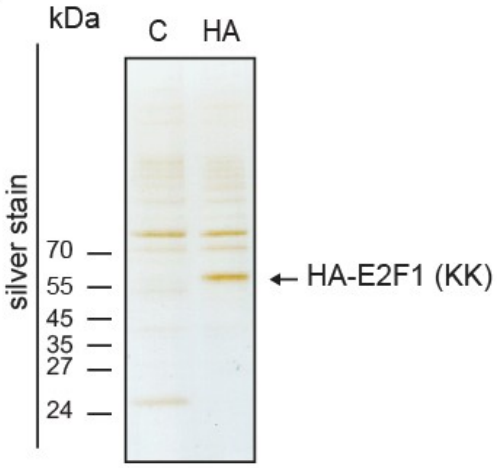

Relative Abundance
m/z [amu]
0
20
40
60
80
100
120
200
300
400
500
600
700
800
900
1000
1100

Figure3. 5 (A) PRMT1 methylates E2F1 on R109: E2F1 mass spectrometry.
(A)(i) Flag-E2F1 (1-194) and empty expression vector control were immunoprecipitated from U2OS co-expressed with 1μg HA-PRMT1 and 1μg Flag-E2F1 (1-194) per each 10mm cell petri dishes (x15 plates). The immunoprecipitated control and Flag-E2F1 elutions were subjected to gel electrophoresis using NuPAGE® Novex® Gel system with NuPAGE Bis-Tris precast gels in the 1X NuPAGE® MES SDS Running Buffer. The protein bands were visualised by silver staining and Flag-E2F1 was identifiied by size and comparison to expression control vector. The arrow indicated gel band was isolated from the gel and subjected to trypsin “in gel digestion (Ch2.10-2)” before analysed by liquid chromatography/tandem mass spectrometry (LC-MS/MS). (ii) LC-MS/MS analysis revealed the submitted sample matched to E2F1 peptide (UniProt: Q1094) and the detailed annotation mass spectrum (m/z, mass-to-charge ratio) of the modified tryptic peptide RLDLETDHQYLAESSGPARGR (position 91-111) was shown in the atomic mass units (amu) with fragment ions indicated as b and y ions for abundance. Each fragment peak represented the detected m/z with the theoretical m/z value indicated in the bracket. There were two methyl groups (2 x CH_{2}-) assigned to the two arginines containing peptide SGPARGR (position 105-111) fragment.

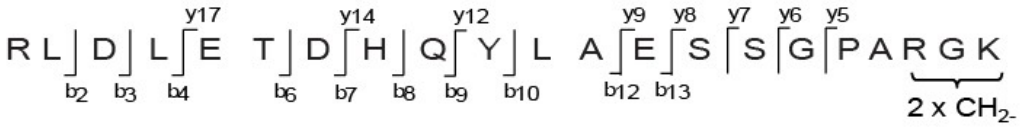
To further clarify the arginine methylation site(s) in cells by PRMT1, both PRMT1 and R111/113K E2F1 (KK E2F1) were transfected into cells and the KK E2F1 was immunoprecipitated and subjected to mass spectrometry analysis. In this setting, R111 and R113 were substituted by lysines (K) which kept the charge and structural change neutral. This approach provided a tool to keep R109 of E2F1 intact but eliminates R111 and R113 in the glycine- and arginine- rich PRMT consensus motif. Customised database containing E2F1 KK sequence has built in as a mass-to-charge reference. The results revealed the Trypsin digested peptide matched to the E2F1 KK sequence and the dimethylation event was assigned to the peptide region 107-111 (PARGK) (Figure3.5B). This demonstrates the two methyl groups detected in the peptide 107-111 contains one arginine at the position 109 of E2F1 as a potential site for the two methyl groups transferred by PRMT1.

To summarise, when examining E2F1 under high level PRMT1 influence, the first MS analysis indicates two methyl groups found in the region covers R109 and R111 of E2F1. With arginine 111 and 113 of PRMT motif mutated to lysines (R111/113K), there are still 2 methyl groups assigned to the fragmented E2F1 containing R109. Collectively, the results suggest that PRMT1 contributes to asymmetrical dimethylarginines at arginine 109 of E2F1.

B (i)


kDa
C
HA
silver stain
70
55
45
35
27
24
← HA-E2F1 (KK)

(ii)


R L D L E T D H Q Y L A E S S G P A R G K
y17 y14 y12 y9 y8 y7 y6 y5
b2 b3 b4 b6 b7 b8 b9 b10 b12 b13
2 x CH2-

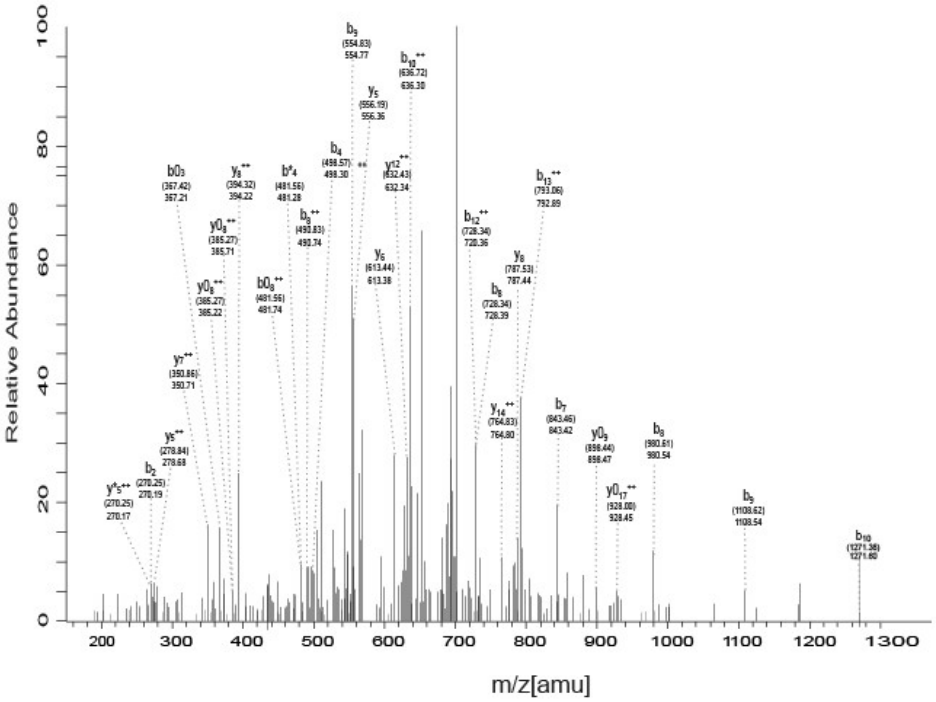

Relative Abundance
m/z[amu]
0
20
40
60
80
100
200
300
400
500
600
700
800
900
1000
1100
1200
1300

Figure3. 6 (B) PRMT1 methylates E2F1 on R109: R111/113K E2F1 mass spectrometry.
(B)(i) HA-E2F1 (KK), HA tagged, full length E2F1 with arginine to lysine substituted by site-directed mutagenesis at position 111 and 113. The HA-R111/113K E2F1 (KK) and expression vector control were immunoprecipitated from U2OS cells co-expressed 48 hours with 1μg HA-R111/113K E2F1 (KK) and 1μg PRMT1 per each 10mm cell petri dishes (x15 plates). The immunoprecipitated control and HA-E2F1(KK) elutions were separated by gel electrophoresis using NuPAGE® Novex® Gel system with NuPAGE Bis-Tris precast gels in the 1X NuPAGE® MES SDS Running Buffer. The protein bands were visualised by silver staining and HA-E2F1(KK) was identifiied by size and comparison to expression control vector (indicated by arrow). The arrow indicated gel band was isolated from the gel and subjected to trypsin “in gel digestion (Ch2.10-2)” before analysed by tandem mass spectrometry LC-MS/MS. (ii) The reference sequence (R111/113K E2F1) was built in for the LC-MS/MS analysis. The detailed annotation mass spectrum (m/z, mass-to-charge ratio) of the modified tryptic peptide 91-111 RLDLETDHQYLAESSGPARGR was shown in the atomic mass units (amu) with fragment ions indicated as b and y ions for abundance. Each fragment peak represented the detected m/z with the theoretical m/z value indicated in the bracket. There were two methyl groups (2 x CH_{2}-) assigned to the one arginine containing peptide 105-111 SGPARGK fragment.

In order to validate if R109 is PRMT1 methylation site, R109K point mutated E2F1 was introduced to compare to the wild-type E2F1 on methylation status. The mutation of arginine 109 to lysine (R109K) on E2F1 has coincided with the loss of ADMA signal on E2F1 (Figure 3.5C). This result further supports the hypothesis that E2F1 is methylated by PRMT1 at arginine 109 in cells.

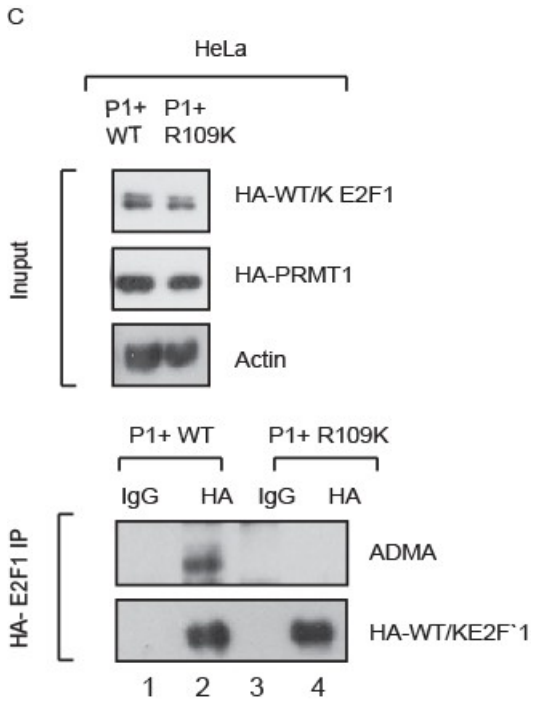


Figure3. 7 (C) PRMT1 methylates E2F1 on R109: Loss of ADMA with E2F1 R109K.
(C) HeLa cells were co-transfected with 1μg PRMT1 and 1μg HA-WT-E2F1 or 1μg HA-R109K-E2F1 (K) for 48 hours before immunoprecipitation with non-specific (IgG) or HA antibodies. The input proteins were separated using immunoblotting by size and probed with HA and Actin antibodies. The immunoprecipitation lysates were probed with ADMA and HA antibodies to detect ADMA signals on immunoprecipitated HA-WT or HA-R109K E2F1.

Ch3.5 Summary

This Chapter investigated E2F1 as a substrate of PRMT1.

Through sequence alignment analysis, it was revealed that E2F1 contains PRMT glycine- and arginine- rich patches (GAR) motif, 109 RGRGR113. Also, another PRMT family member, PRMT5 methylated within this region of E2F1 at arginine 111 and 113. It is possible that PRMT1 also methylates E2F1 and make it a common substrate.

In this Chapter, E2F1 was firstly tested as a substrate of PRMT1 in the *in vitro* methylation assay. In cells, the correlation between PRMT1 and the ADMA modification signal on E2F1 were demonstrated, such as the inducible ADMA mark when PRMT1 level elevated and the decrease of ADMA signal along with the depletion of PRMT1 by siRNA. Furthermore, with the help of mass spectrometry, the addition of two methyl groups on E2F1 was found biochemically located within the highlighted glycine- and arginine- rich region 109 RGRGR113. The employment of KK E2F1 transformed the region to 109RGKGK113 and the detection of two methyl groups remained. These collective efforts suggest E2F1 dimethylation at arginine 109 by PRMT1. Lastly, the ADMA modification signal was reduced when lysine substituted arginine at position 109 (R109K). These collectively show the importance of arginine 109 of E2F1 in PRMT1 methylation event.

Chapter 4 Impact of arginine methylation on E2F1 by PRMT1

Ch4.1 Characterisation of inducible cell lines

In the previous chapter, E2F1 was established to be a substrate of PRMT1. Next is to assess what this asymmetric dimethyl addition to E2F1 does. Moreover, it is interesting to compare what physiological effect the resultant ADMA contributes, to that of the SDMA by PRMT5 on the common substrate, E2F1.

The Tet-On inducible expression system, with its strength in easy regulation of gene expression, was therefore adopted. The system makes use of the feature of a type of transcriptional activation, which in response to activation, increased expression of the gene or genes downstream of its promoter. This activation is Doxycycline (Dox) dependent, which means the resultant gene expression, expressed or not, depends on the presence or absence of Dox. As the activation came from antibiotic tetracycline or one of its derivatives (i.e. doxycycline) therefore, the “Tet-On, expression on” relationship concludes the inducible feature of the system.

The generation of HA-tagged inducible cell lines (inducible PRMT1 (inP1) and inducible PRMT5 (inP5)) (Ch2.1-2), aimed to harnessing such Dox responsive control to regulate PRMT1 or PRMT5 protein level in cells. These cell lines thus are good tools to monitor changes in cells organically under influences of the ADMA catalyst, PRMT1 or the SDMA catalyst, PRMT5.

The PRMT construct was designed to carry a HA tag and flanked with BamHI and

NotI restriction digestion sites, which enable the subsequent sub-clone into the hygromycin resistance containing Tet-On pTRE2hyg vector. Following transfection into Tet-On parental osteosarcoma U2OS cells and selection by hygromycin, HA-tagged inducible PRMT1, PRMT5 and an empty vector cell lines were developed.

The characterisation of inducible cells began with the confirmation of protein production. In this section, the inP1 cells were harvested at different doxycycline incubation time points before subjected to immunoblotting for HA-PRMT1 detection. Immunofluorescence staining was used to visualise the localisation of the induced PRMT1.

In the absence of doxycycline, no HA-PRMT1 was detected. After incubated inP1 cells with doxycycline for 24 hours or 48 hours, HA-PRMT1 was detected, and there is an increased amount of HA-PRMT1 detected at 48 hours than it is at 24 hours (Figure4.1A). The result suggested the HA-PRMT1 Tet-On system is inducible by doxycycline as the expression of HA-PRMT1 detected in the presence of Dox, a confirmation of a Dox inducible response.

Further, the morphology of the inP1 cells remained similar to the normal U2OS osteosarcoma orthogonal shape (Khyrul et al., 2004), regardless the presence or the duration of Dox treatment (Figure4.1B ph3 or DAPI stained).

Finally, the induction of HA-PRMT1 was visualised by immunofluorescence staining to investigate the localisation of the induced PRMT1. At Dox treatment time 0 (untreated), the HA-tagged PRMT1 remained undetectable at background level and then became

detectable at hour 24 and at hour 48 (Figure4.1B HA 0h, 24h, 48h) which again confirmed the Dox responsive-inducible nature of the inP1 cells. The results suggested that the induced PRMT1 mainly localised in the DAPI-stained nucleus which is in agreement of previous report with regard of PRMT1 location in cells (Tang et al., 1998). Notably, also, there was an improved induction coverage rate at 48 hour compared to 24 hour.

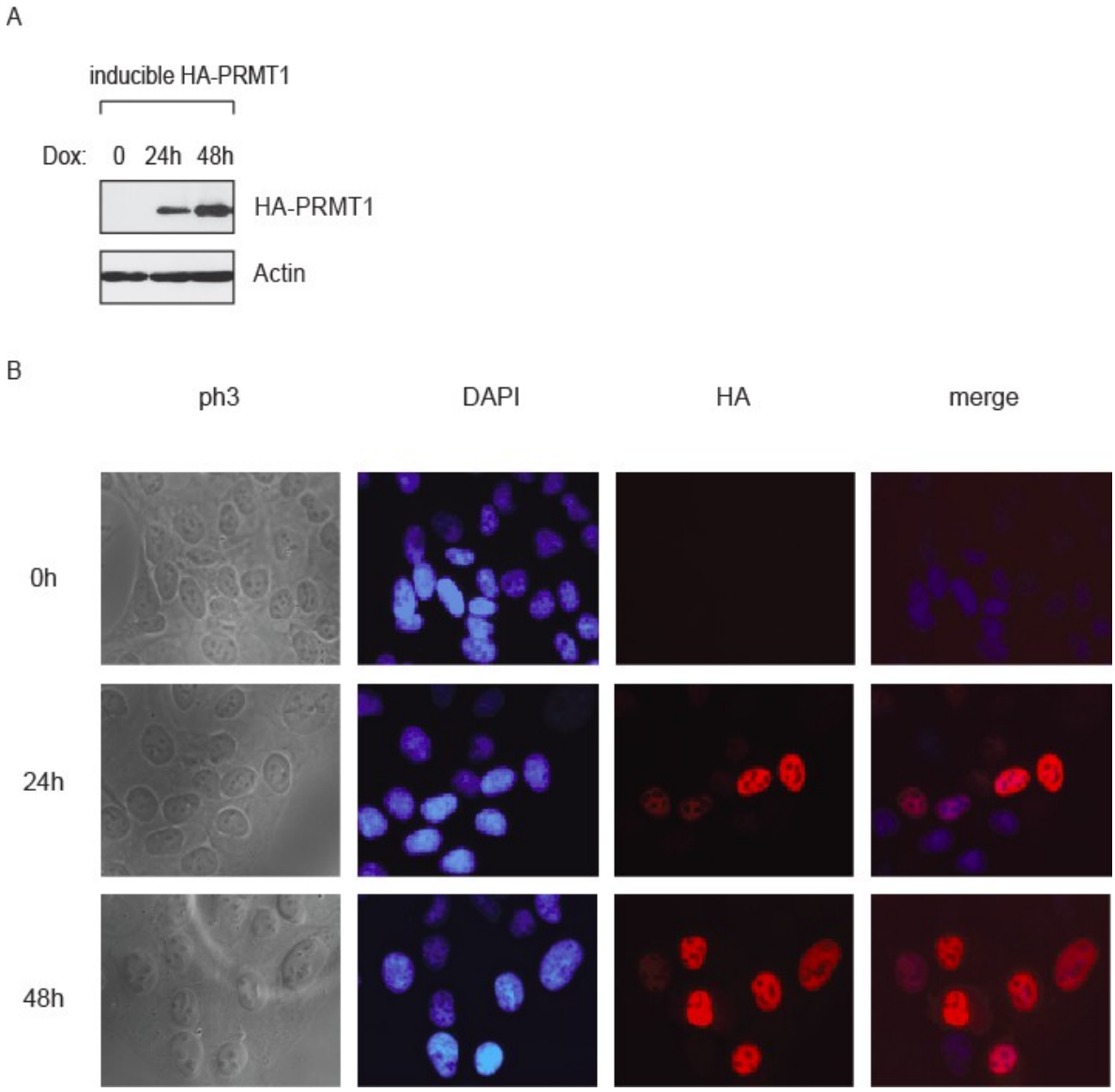


Figure4. 1 Cellular localisation of inducible PRMT1.

(A) Inducible HA-PRMT1 cells were harvested at time point 0 as untreated control or incubated with 1μg/ml doxycycline for 24 hours and 48 hours before harvest for immunoblotting. The induction of HA-PRMT1 was confirmed probing with HA antibody by immunoblotting. The β-actin, the highly conserved cytoskeletal protein, was measured as a loading control. (B) Immunofluorescence staining was applied to doxycycline time-course treated inducible PRMT1 cells on the pre-seeded coverslips in parallel to cells subjected to above immunoblotting. HA-PRMT1 proteins were visualised by receiving mercury lamp excited 1:200 1μg HA primary antibody in conjunction with 1:500 fluorochrome-conjugated secondary antibody stained cells. The DAPI staining was used to label nuclei in the cells (blue). The additional phase contrast (ph3) image demonstrate light contrast image taken without mercury lamp excited fluorochrome emission and focused on morphology of cells.

In brief summary, three aspects of inP1 cells were checked. Firstly, the HA-PRMT1 is inducible upon doxycycline treatment and, secondly, the morphology of its osteosarcoma backbone remained unaffected by the induction and, thirdly, the induced PRMT1 chiefly localised in the nucleus which is in consistent with the previous report.

It is important next to test if the induced PRMT1 is indeed enzymatically active. This was done by immunoprecipitating the PRMT1 by its designed tag (HA) from the cells be subjecting it to the *in vitro* methylation assay. The empty vector cell line (pTRE vector, V) was used as a negative control and the recombinant GST-E2F1 was used as a substrate. Instead of using both recombinant enzymes and substrates purified from *E.coli*, this semi-*in vitro* setting aims at measuring the methyltransferase ability of enzyme by co-incubating it with its potent substrate, E2F1, in the *in vitro* setting. Two findings were concluded from this experiment, firstly, when both vector and inP1 cells were treated with Dox, only the HA-PRMT1 were induced and detected (Figure4.2A). Secondly, the results suggested the Dox-induced PRMT1 was active as it accounted for a statistically higher ^{3}H DPM count compared to the negative controls (Figure4.2B). The results demonstrated the induced enzyme PRMT1 is active and functions as a methyltransferase enzyme.

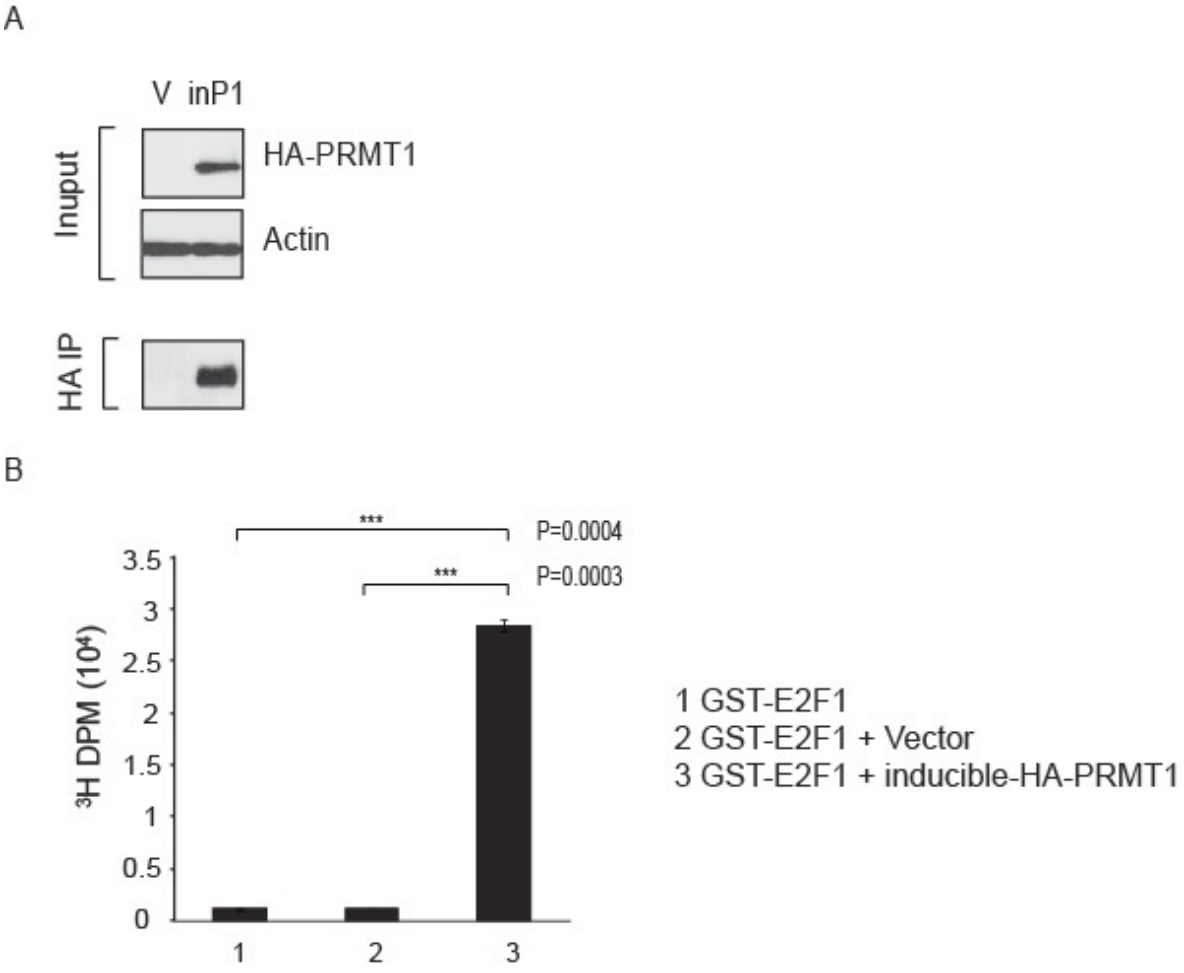


Figure4. 2 Enzyme activity of inducible PRMT1.

(A) The empty vector (pTRE-hyg) and inducible HA-PRMT1 cells were induced with doxycycline for 48 hours and both harvested for HA immunoprecipitation. Input and immunoprecipitated control or HA-PRMT1 were confirmed with HA antibody immunoblotting. Actin was served as loading control. (B) The immunoprecipitated control (vector) or PRMT1 was incubated with radioactive ^{3}H-SAM and bacteria purified 1μg GST-E2F1 in the *in vitro* methylation assay measured by disintegrations per minute DPM. The standard deviation derived from three independent experiments. P value< 0.001 (***). (N=3)

The same rules applied to the characterisation of PRMT5 inducible cells. The expression of HA-PRMT5 was confirmed when inP5 cells were treated with Dox and the level of HA-PRMT5 increased at 48 hours post Dox treatment compared to that of 24 hours (Figure4.3A). The same U2OS osteosarcoma orthogonal shape was remained in the inP5 cells with or without doxycycline exposure.

The immunofluorescence staining was used to visualise the localisation of active PRMT5. In the absence of doxycycline, no HA was detected. After incubation with doxycycline for 24 hours or 48 hours, HA-PRMT5 was detected (Figure4.3B). The results suggested that the majority of Dox-induced PRMT5 localised in the cytoplasm. This observation was in agreement with previous PRMT update in humans (Wolf, 2009). It is interesting to note here that it was previously reported that PRMT5 physically interacted with E2F1 by immunoprecipitation and also the PRMT5-induced symmetrically dimethylated E2F1 in different cancer cell lines including HeLA and U2OS cells (Cho et al., 2012).

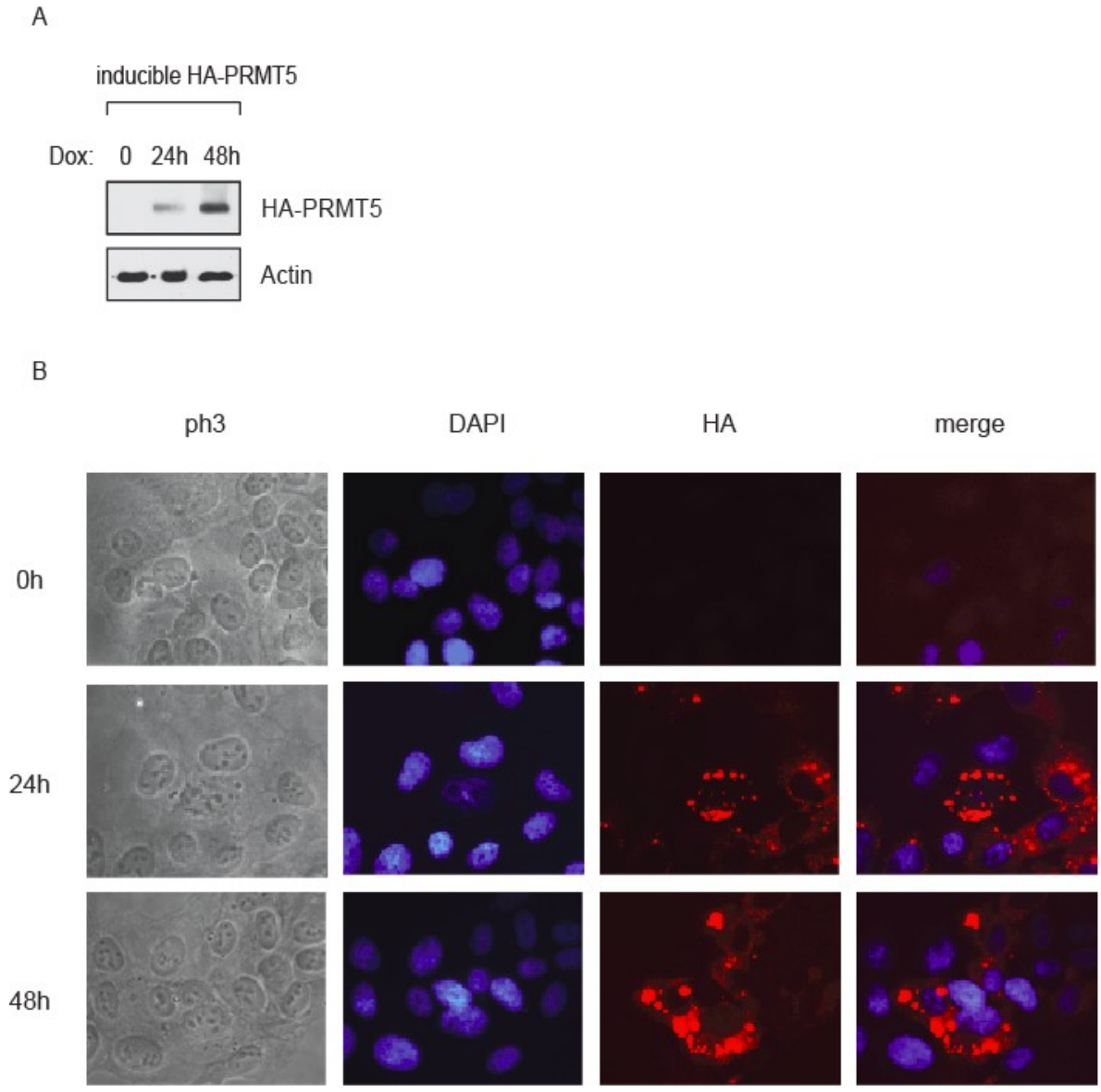


Figure4. 3 Cellular localisation of inducible PRMT5.

(A) Inducible HA-PRMT5 cells were harvested at time point 0 as untreated control or incubated with 1μg/ml doxycycline for 24 hours and 48 hours before harvest for immunoblotting. The induction expression of HA-PRMT5 was confirmed probing with HA antibody by immunoblotting. The β-actin served a loading control. (B) Immunofluorescence staining was applied to doxycycline time-course treated inducible PRMT5 cells on the pre-seeded coverslips in parallel to cells subjected to above immunoblotting. HA-PRMT1 proteins were visualised by receiving mercury lamp excited 1:200 1μg HA primary antibody in conjunction with 1:500 fluorochrome-conjugated secondary antibody stained cells. The DAPI staining was used to label nuclei in the cells (blue). The additional phase contrast (ph3) image demonstrate light contrast image taken without mercury lamp excited fluorochrome emission and focused on morphology of cells.

The methyltransferase activity of induced PRMT5 was measured by co-incubating cell-derived HA-PRMT5 and its known substrate, histone protein, in the *in vitro* methylation assay with the radioactive tritium. The results suggested that only when in the presence of Dox, the HA-PRMT5 expressed and detected (Figure4.4A). The immunoprecipitated PRMT5 was then incubated with histone protein. The enzyme alone and substrate alone, both negative controls, showed background level of tritium ^{3}H count (Figure4.4 lane 1 and 2) while the co-incubation of HA-PRMT5 and histone demonstrated a significant higher tritium ^{3}H count (Figure4.4 B lane3).

The results indicated the occurrence of the methyl groups transfer (radioactive tritium ^{3}H incorporated) when incubating Dox-induced HA-PRMT5 with its known substrate, and therefore concluded the Dox-induced, inP5 cell-derived PRMT5, functioned as an active methyltransferase enzyme towards its substrate.

In summary, the characterisation of the Tet-On inducible HA-PRMT5 cell line was confirmed to be inducible, during Dox induction, predominantly localised in the cytoplasm and the ectopic PRMT5 was enzymatically active.

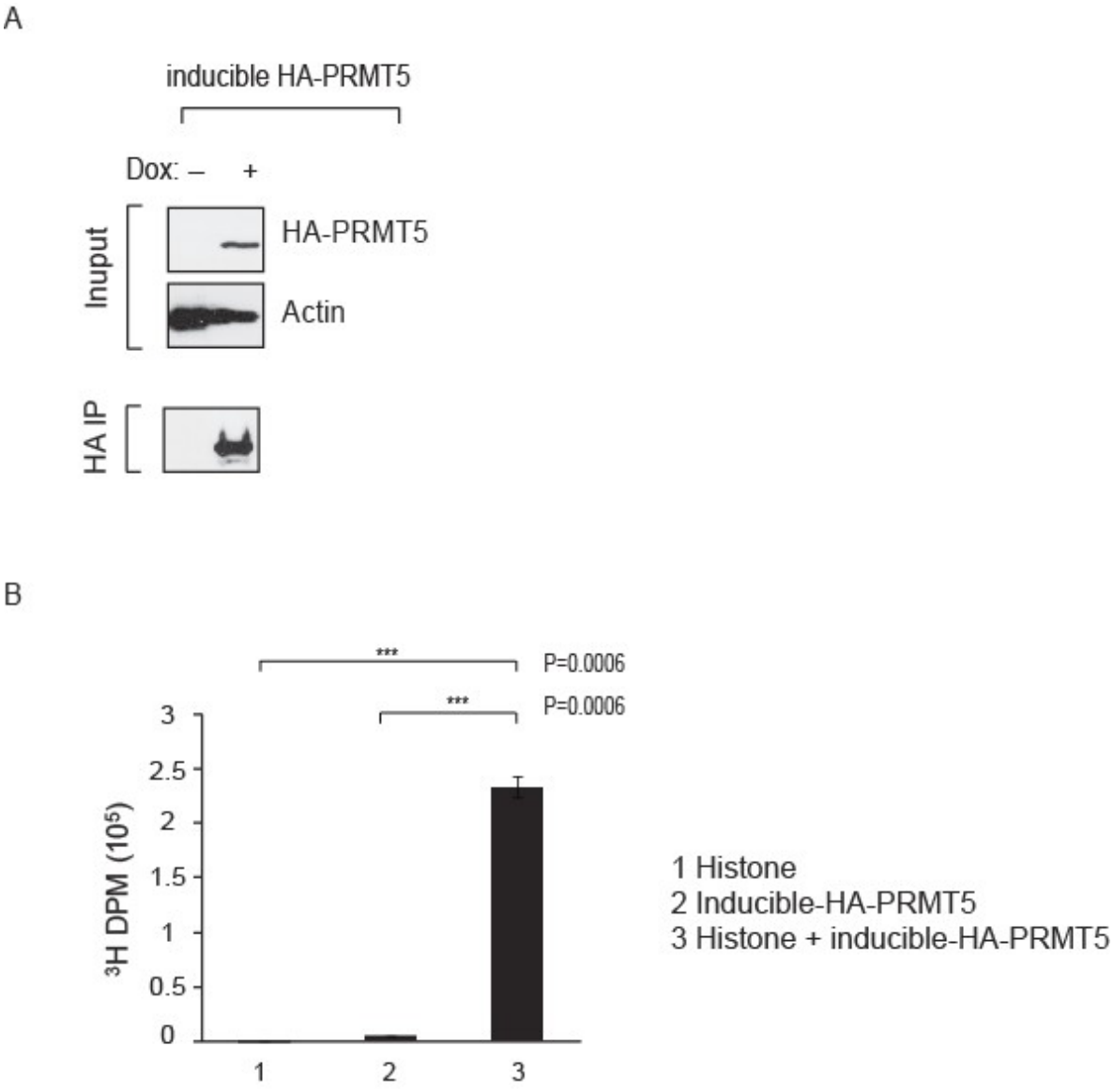


Figure4. 4 Enzyme activity of inducible PRMT5.

(A) The HA-PRMT5 cells were induced with or without doxycycline for 48 hours and both harvested for HA immunoprecipitation. Input and immunoprecipitated HA-PRMT5 were confirmed with HA antibody immunoblotting. Actin was served as loading control. (B) The immunoprecipitated PRMT5 was incubated alone or with 1μg histone in the radioactive ^{3}H-SAM enabled *in vitro* methylation assay, measured by disintegrations per minute DPM. The standard deviation derived from three independent experiments. P value< 0.001 (***).

Ch4.2 Growth inhibition of PRMT1-induced ADMA on E2F1

In the previous section, the PRMT1 induced ADMA and the PRMT5 induced SDMA on E2F1 were observed in the cells. Next, it was intriguing to compare effect of PRMT1 resultant ADMA on E2F1 with the previously reported SDMA mark on E2F1 by PRMT5. One of the key findings from previous work found that PRMT5 dependent symmetrically dimethylated E2F1 impacted on the E2F1-dependent growth control (Cho et al., 2012). More specifically, PRMT5 depletion coincided with a decreased growth rate and was associated with apoptosis. In other words, it implied a role in the inhibition of apoptosis.

A simple measure of apoptosis is the percentage of sub-G1 cells in a cell cycle profile. The mammalian cell division cycle can be segmented into four basic phases. The phase a cell grows to duplicate its DNA content, occurred in the DNA synthesis phase (S). The mitosis (M) phase is when the duplicated DNA content separated to become two individual cells. The two gaps in between S phase and M phase are called gap 1 (G1), prior to S phase; and gap 2 (G2), after S phase (Sherr, 1996).

The propidium iodide (PI) is an intercalating fluorescent agent used as a nucleic acids dye. As PI binds to nucleic acids stoichiometrically, its binding is therefore proportional to the DNA (Pollack and Ciancio, 1990). The sub-G1 analysis employs this quantitative DNA measurement and focused on the hypodiploid cells. Hypodiploidy is a feature of apoptotic cells during endonucleases-activated DNA fragmentation (Riccardi and Nicoletti, 2006). In

other words, these hypodiploid cells were the products by endonucleases degrading DNA in the cells during apoptosis.

The following experiments aimed at utilising the inducible HA-PRMT1 cells and PRMT1 siRNA treated osteosarcoma cells (the backbone cell-type of the inducible cells) as a pair of tools to observe changes in the percentage of sub-G1 cells, under influences of high and low level of the E2F1 ADMA catalyst, PRMT1. Another factor to consider was the E2F1 dependency, which was performed by comparing non-targeting and targeting (E2F1) siRNAs transfected into the inducible PRMT1 cells, in parallel to validate the E2F1 relevance in the sub-G1 experiments set.

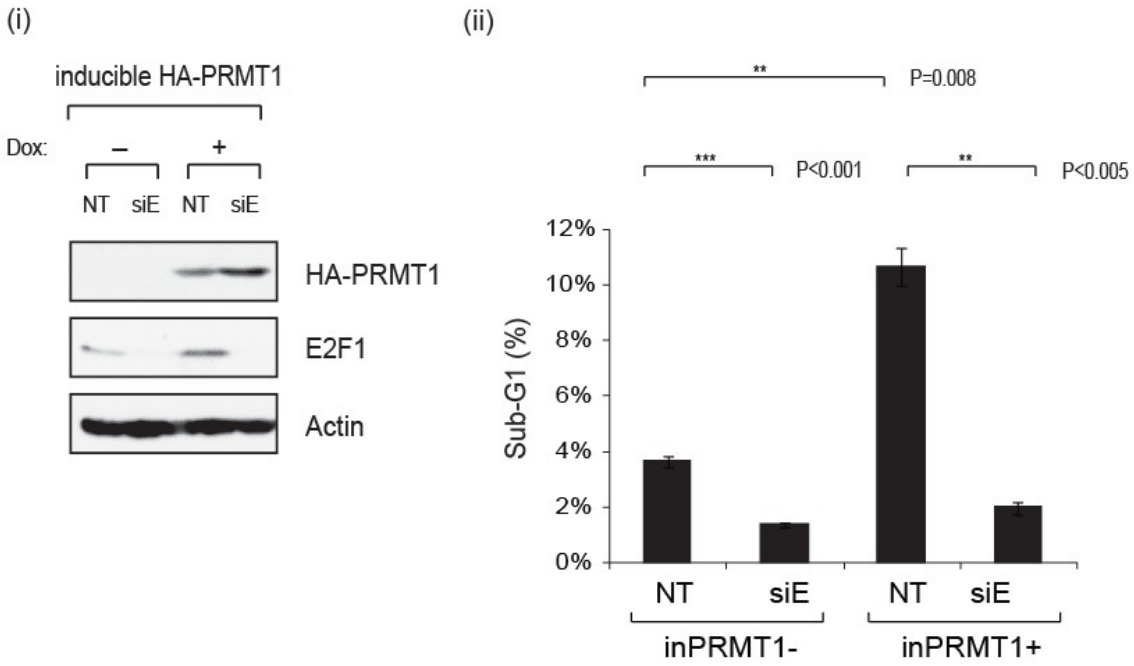
In the first set of experiments, the induction of PRMT1 detected when cells were treated with Dox (Figure4.5 A(i)). Both Dox-untreated and treated cells (or PRMT1-uninduced and induced cells) were transfected with a control non-targeting (NT) siRNA or the E2F1 siRNA. The E2F1 depletion was confirmed by immunoblotting (Figure4.5 A(i)). When focusing on both PRMT1-uninduced and induced cells, the percentage of sub-G1 cells tripled from 3.64% to 10.66% along with the induction of PRMT1 (Figure4.5 A(ii) inPRMT1- NT vs. inPRMT1+ NT). The results suggested a positive correlation between the presence of PRMT1 and the increased amount of hypodiploid cells. Moreover, the pattern seen in the NT group could not be reproduced when E2F1 were removed from the cells (Figure4.5 A(ii) NT PRMT1-/+ vs. siE PRMT1-/+). This means the existence of E2F1

is crucial to that PRMT1 responsive build-up of the sub-G1 cells. Overall, these observations suggested an E2F1 dependent, PRMT1 inducible accumulation of sub-G1 cells. As these hypodiploid sub-G1 cells were implied as apoptotic cells, the accumulation of sub-G1 cells revealed a PRMT1-induced pro-apoptosis effect which is E2F1 dependent.

Opposite to what induced PRMT1 cells provided, the depletion of endogenous PRMT1 was performed using PRMT1 siRNA transfection and confirmed with immunoblotting (Figure4.5 B(i)). This was to test further for the correlation between apoptotic hypodiploid cells and PRMT1. With more PRMT1, more sub-G1 cells observed, the loss of PRMT1 corresponded to a decrease of percentage of sub-G1 cells by roughly two fold from 5.47% to 2.66% of all cells measured (Figure4.5 B(ii)).

To summarise, the correlation between PRMT1 and apoptotic cells is strong as the increase or decrease amount of sub-G1 cells responded to the increase or decrease of PRMT1 level. What is more, this relationship is E2F1-dependent. Therefore it is interesting to note that while impact of E2F1 SDMA by PRMT5 demonstrated a role in the inhibition of apoptosis, the sub-G1 analysis probe revealed a promotion of apoptosis which is E2F1 dependent and inducible by PRMT1.

A


(i)
inducible HA-PRMT1
Dox:
–
+
NT
siE
NT
siE
HA-PRMT1
E2F1
Actin
(ii)
**
P=0.008
***
P<0.001
**
P<0.005
Sub-G1 (%)
12%
10%
8%
6%
4%
2%
0%
NT
siE
NT
siE
inPRMT1-
inPRMT1+

B

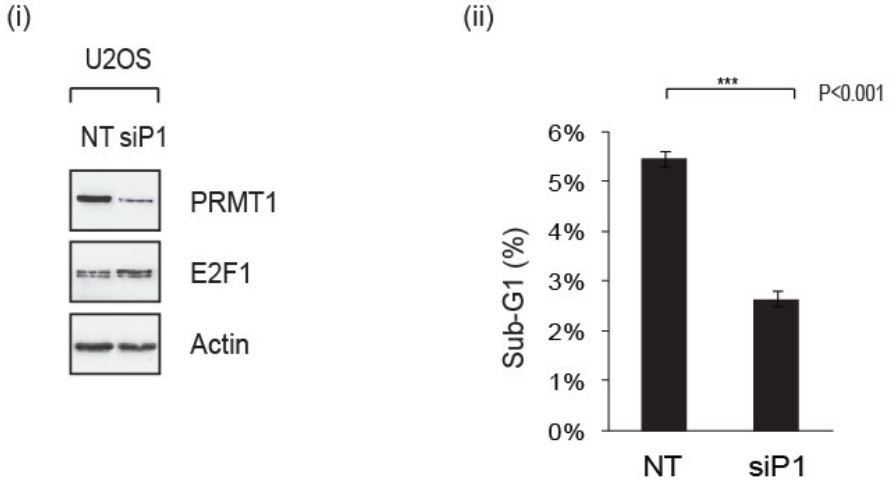

(i)
U2OS
NT siP1
PRMT1
E2F1
Actin
(ii)
***
P<0.001
Sub-G1 (%)
6%
5%
4%
3%
2%
1%
0%
NT
siP1

Figure4. 5 Correlation between PRMT1 and sub-G1 cells.
(A)(i) The inducible PRMT1 cells were transfected with 50nM non-targeting (NT) or E2F1 siRNAs 24 hours before treated with or without 1μg/ml doxycycline for further 48 hours. After 72 hours, cells were harvested for immunoblotting and sub-G1 analysis by flow cytometer. Cells harvested for immunoblotting used HA, E2F1 and Actin antibodies to confirm HA-PRMT1 induction and E2F1 siRNA transfection with loading control. A(ii) Cells processed to sub-G1 analysis were stained with PI, which was proportioned to DNA content. Flow cytometer recorded width and numbers of cells flew through and matched to the received PI emission by UV excitement. The sub-G1 percentage indicated percentage of cells with degraded DNA which served as indicator for apoptotic cells. The flow cytometer was set for capturing 20, 000 cells. The data shown was the mean and standard deviation of triplicate samples using student's *t-test*. P value< 0.01 ** ;P value< 0.001 *** for statistical significance. (B)(i) U2OS cells were treated with 50 nM NT or PRMT1 siRNA for 72 hours and subjected to immunoblotting with PRMT1, E2F1 and Actin antibodies. B(ii) The percentage of sub-G1 recorded by Flow cytometer capturing 20,000 cells. Data shown is the mean and standard deviation applied to triplicate samples. P value<0.001***.

In order to address further on this E2F1 dependent association between PRMT1 induction and apoptosis, another common human cancer cell line, the cervical cancer HeLa cells were used with PRMT1 and PRMT5 siRNAs transfection in parallel. As previously reported, PRMT5 depletion has demonstrated a role in inhibition of apoptosis, and therefore in these experiments, PRMT1 depletion was performed side by side with PRMT5 depletion for the detection of classic apoptosis markers.

When the effect of siRNAs transfection confirmed was a decreased level of PRMT1 and PRMT5 proteins by immunoblotting, the cleaved form of cysteine-aspartic acid protease-3 (caspase-3) and the cleaved form of poly ADP-ribose polymerase (PARP) increased along with the decline of PRMT5 protein level, not with PRMT1 (Figure4.6). The caspase cascade is the hallmark of the apoptotic pathway, in which caspase-3 plays a pivotal role, and the cleaved form of caspase-3 as a proof of an apoptosis process in action (McIlwain et al., 2013). The poly ADP-ribose polymerase (PARP) senses stress in cells and begins the repair process by adding poly ADP-ribose chain which acted as a platform for the repair complex formation. It is inactivated when cleaved hence the detection of cleaved form of PARP indicated the prevention of repair and the initiation of apoptosis (Chaitanya et al., 2010). Consequently, PRMT5 implicated as an apoptosis inhibitor while PRMT1 depletion demonstrated a similar level of expression of the apoptotic markers when compared with control.

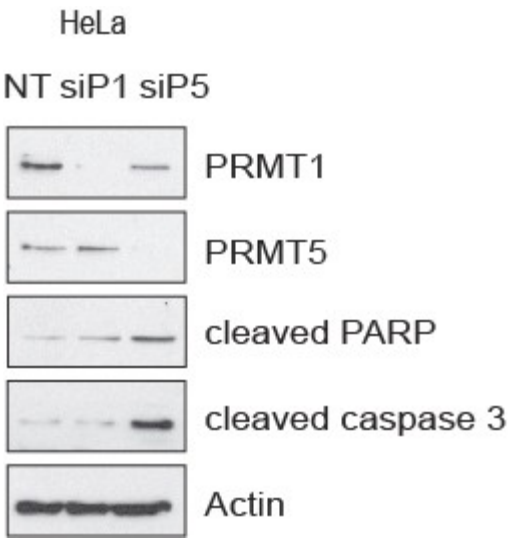


Figure4. 6 Change of apoptotic markers upon PRMT depletion

HeLa cells were treated with 50nM NT, PRMT1 or PRMT5 siRNAs for 72 hours. Immunoblotting used PRMT1 and PRMT5 antibodies to confirm the effect of siRNA transfections by comparison to control NT, Actin antibody used as loading control and cleaved PARP and cleaved caspase 3 antibodies were used to detect the level of these proteins. Data shown was most representative of three independent experiments (N=3).

The cell growth colony assay adopted here is a mean to visualise changes of cell growth simply by the numbers of cell colonies stained. The “stably induced E2F1 cells” (Ch2.1-3), the expression empty vector, wild-type E2F1, and E2F1 mutant derivatives, E2F1 R109K (K) and E2F1 R111/113K (KK) HeLa cells were employed in the colony assay. As the arginine to lysine point mutation at position 109 removed PRMT1 methylation site and position 111 and 113 removed PRMT5 methylation sites, the experiments aimed at addressing the relationship between arginine methylation and cell growth directly.

When comparing Vector and WT-E2F1, cells transfected with WT-E2F1 demonstrated the apoptotic trait of E2F1 as the density of colonies decreased (Figure4.8A Vector vs. WT-E2F1). Interestingly, R109K E2F1 transfection has seen HeLa cells responded with more colonies than there were with WT-E2F1, and when the R111/113K E2F1 transfected, on the other hand, there were fewer colonies seen in comparison with that of WT E2F1 (Figure4.7A). When arginine substituted to lysine at position 109 (R109K), more cell colonies counted, on the contrary, R111/113K (KK) accounted for a decrease in colonies formation (Figure4.8B R109K vs. R111/113K). These results indicated R109 site of E2F1 was crucial in cell growth resistance as once mutated to R109K, more cell colonies seen; on the contrary, R111 and R113 sites of E2F1 were important to sustain cell growth as once mutated to R111/113K, fewer colonies seen.

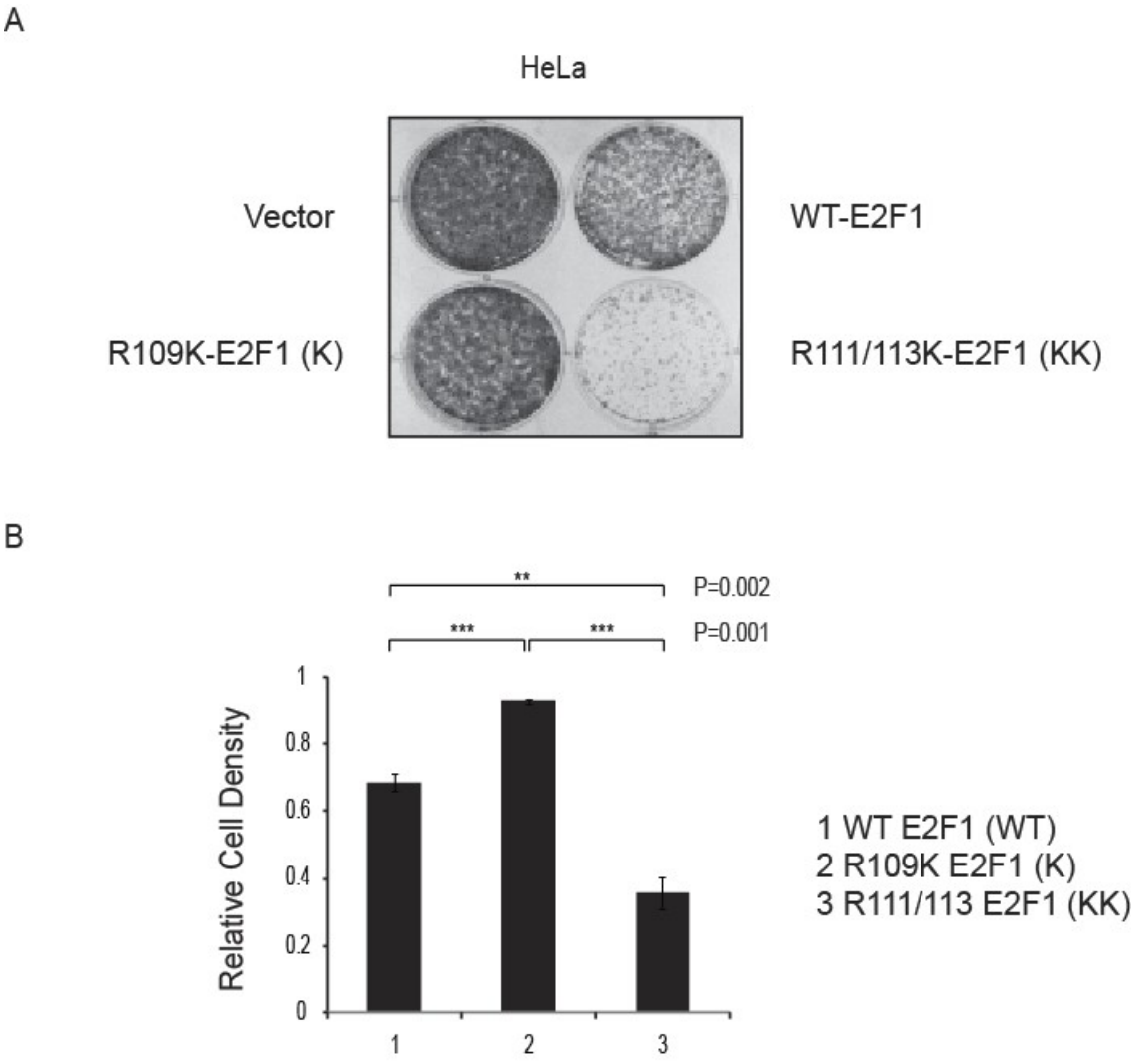


Figure4. 7 Arginine methylation sites and cell growth

(A) HeLa cells were transfected with 1μg DNA plasmids, the mammalian expression pcDNA3.1 vector (Vector), wild-type E2F1 (WT-E2F1), R109K-E2F1, R111/113K-E2F1, in the 60% confluent cell petri dish (10mm) for 48 hours and selected with 800μg/ml G418 containing DMED for 5-7 days when the control HeLa cells (no antibiotics resistance acquired by the pcDNA3.1 plasmid transfection) ceased to grow. Cells were trypnised and counted, 1000 cells were seeded and maintained in the selection medium (800μg/ml G418, 5% FCS, DMEM) for 10 days. The cell colonies were stained with crystal violet dye and the cell plate was scanned by the Gelcount Colony Counter. (B) The quantitation was done by Image J calculating image pixels which reflected the cell colonies density. The relative cell density was normalised against Vector control. Data shown was mean and standard deviation of three independent experiments. The student *t-test* was used and P value<0.1**;

P value< 0.001***. (N=3).

In summary, when the PRMT1 level increased, the apoptotic sub-G1 cells accumulated in an E2F1 dependent fashion; when PRMT1 level decreased, the apoptotic markers tested remained low at the background level, in contrast to the elevated level of apoptotic markers seen with PRMT5 depletion. Moreover, loss of PRMT1 dimethylation on R109 prevented growth resistance that wild-type E2F1 imposed. This was the reverse effect the loss of PRMT5 dimethylation on R111 and R113, where reduced growth occurred. The evidence suggests PRMT1 arginine methylation on E2F1 functioned in a fashion that favours apoptosis or growth resistance, hence the conclusion of growth inhibition by PRMT1-indcued ADMA on E2F1.

Ch4.3 Promoter specificity

It is interesting to observe the growth inhibitory effect that PRMT1 had on E2F1 which is opposite to the apoptosis inhibition demonstrated by PRMT5 on the same substrate. This phenomenon echoed some previous findings where opposing effects recorded executed by PRMT1 and PRMT5 on the histone protein H4 (Majumder et al., 2010) (Huang et al., 2005). The observations supported the hypothesis that different arginine methylation signals (ADMA or SDMA) as the underlying mechanism for the dual roles of E2F1 in cell cycle progression and apoptosis.

One possible explanation is based on the chief function of E2F1 as a transcription factor. It is possible that through promoter specificity, E2F1 responded to ADMA and SDMA signals differently which consequently resulted in the opposing effects on the cell fate. In another words, ADMA and SDMA signalled E2F1 to target specific promoter groups that eventually contributed to the corresponding apoptosis or cell cycle progression.

This hypothesis was first examined using the inducible cell lines. In the absence and presence of doxycycline, PRMTs were absent or induced correspondingly (Figure4.8A(i)) and E2F1 or control chromatin were immunoprecipitated from each conditions whether Dox is presence or not. The representative targets chosen were the classic E2F1 promoters, Cdc6 and Apaf1, which traditionally classified in the cell cycle regulation and the apoptosis

categories respectively. The result showed an intriguing opposite trend when compared inducible PRMT1 and PRMT5 cells, especially apparent in the Cdc6 promoter binding.

When PRMT1 was induced, which previously demonstrated a correlation with a high level ADMA on E2F1, the decrease of Cdc6 promoter binding was seen compared to the uninduced PRMT1 cells (Figure4.8A(ii) lane 2 and 4 on Cdc6) while interestingly, the opposite trend was seen in the Apaf1 promoter where induction of PRMT1 encouraged E2F1 binding (Figure4.9A(ii) lane 2 and 4 on Apaf1). This was opposite to the PRMT5 induction where Cdc6 promoter binding accumulated (Figure 4.8A(ii) lane 8 and 10) upon PRMT5 induction. The relative enrichment graphs demonstrated this opposing trend in the uninduced and induced PRMT1 in the Cdc6 and Apaf1 promoter binding (Figure 4.9 (iii) (a) and (c)) and the opposing trend of Cdc6 binding comparing PRMT1 and PRMT5 induction (Figure 4.8 (iii) (a) and (b)).

The results suggested a trend, when PRMT1 was introduced, there was a promoter specificity discovered towards the inhibition of cell growth (Cdc6 decreased) and a favourable accumulation on the apoptotic promoter (Apaf1 increased). The promoter specificity was demonstrated again when PRMT5 was induced and E2F1 chromatin binding was found increased on the cell cycle progression promoter rather than the apoptotic target (Figure4.9A). Collectively the results supported the promoter specificity hypothesis.

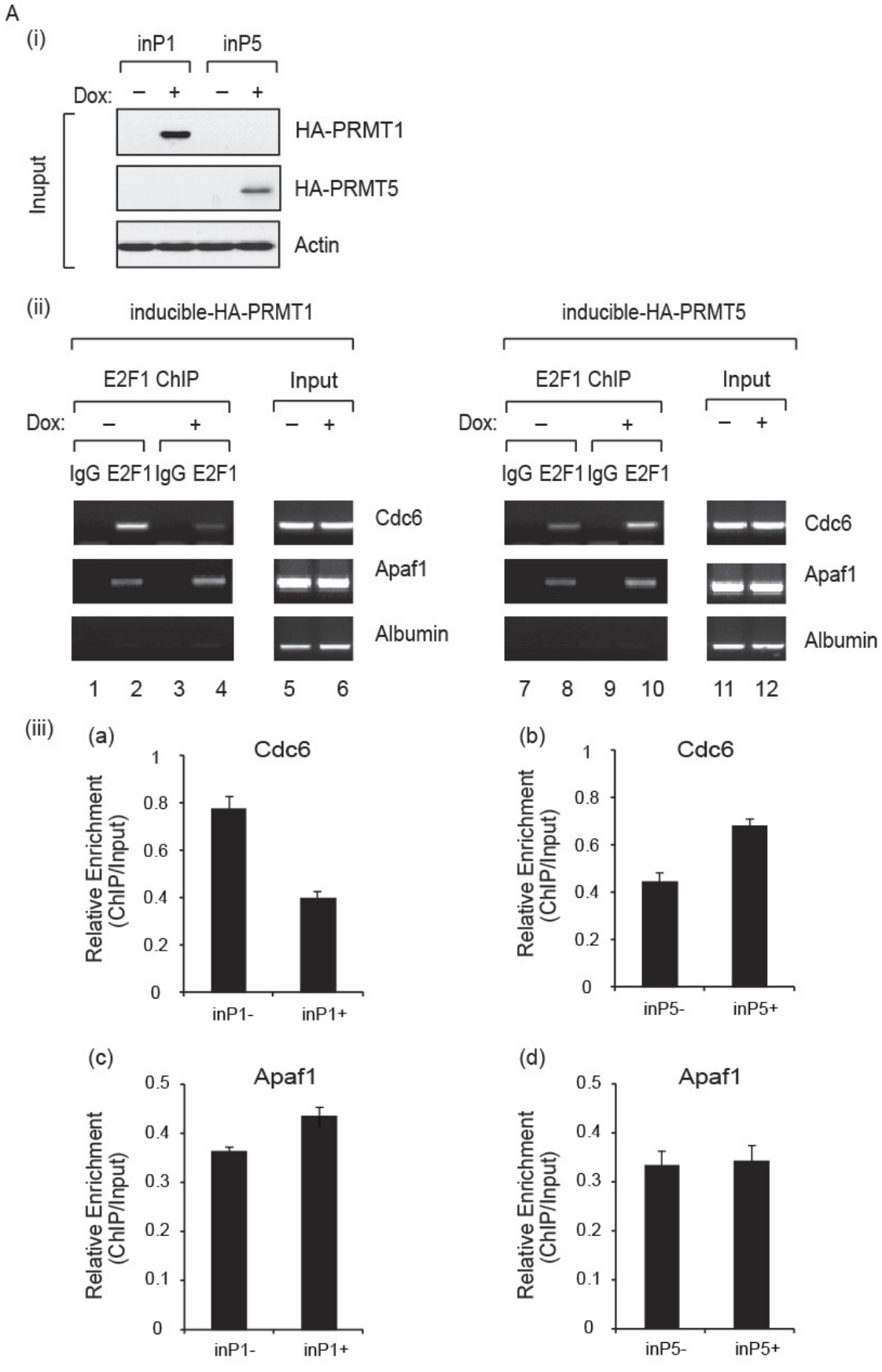

A
(i)
inP1
inP5
Dox: – + – +
Inuput
HA-PRMT1
HA-PRMT5
Actin
(ii)
inducible-HA-PRMT1
inducible-HA-PRMT5
E2F1 ChIP
Input
Dox: – + – +
IgG E2F1 IgG E2F1
Cdc6
Apaf1
Albumin
1 2 3 4 5 6
7 8 9 10 11 12
(iii)
(a)
Cdc6
Relative Enrichment (ChIP/Input)
inP1-
inP1+
(b)
Cdc6
inP5-
inP5+
(c)
Apaf1
(d)
Apaf1

Figure4. 8 (A) E2F1 promoter specificity upon PRMT induction

(A)(i) Inducible PRMT1 or PRMT5 cells were treated with or without 1μg/ml doxycycline (Dox) for 24 hours before fixed and cross-linked for E2F1 chromatin immunoprecipitation. Using immunoblotting, the input was confirmed with induction of PRMT with HA antibody. The Actin antibody was used for the loading control. (ii) 1μg of non-specific or E2F1 antibodies were used to immunoprecipitate chromatin from cross-linked and sonicated cells. Purified input and ChIP DNA were amplified by PCR with Cdc6, Apaf1 forward and reverse primer pairs. The housekeeping albumin gene was used as the chromatin loading control. The PCR products were visualised on 2% agarose gel with 0.5μg/ml ethidium bromide, excited by UV. Data shown was best representative of two independent experiments. (iii) Image J software (NIH) was used to quantitate the relative enrichment of ChIP over input on, (a)(b) E2F1 and Cdc6 binding, (c)(d) E2F1 and Apaf1 binding (N=3).

To further test the promoter specificity hypothesis, the cervical cancer HeLa cells were treated with PRMT1 and PRMT5 siRNAs to eliminate the endogenous PRMTs from the cells. The setting was to investigate where endogenous E2F1 bind at the chromatin level under influence of PRMT1 and PRMT5 absence.

When compared between PRMT1 and PRMT5 siRNAs transfections, the amount of E2F1 chromatin immunoprecipitated appeared accumulated on Cdc6 promoter and absent on Apaf1 promoter upon PRMT1 depletion (Figure 4.9B (ii) lane 4 and 6)(Figure 4.9B (iii)). These binding pattern was unique to the selected promoters as E2F1 was found of no specific binding pattern on the control house-keeping albumin promoter (Figure4.8B(ii)). The trend was reversed when PRMT5 was depletion.

These results demonstrated the relationship between PRMTs elimination and the promoters binding of the transcriptional factor E2F1. In short, in the absence of PRMT1, E2F1 bound to the cell cycle progression promoter Cdc6 and was off the apoptotic Apaf1 promoter. This indicated that PRMT1, when present, was influential by inhibiting E2F1 binding to the cell cycle regulation promoter Cdc6 and by the binding of the apoptotic promoter Apaf1. This finding was in tune with the promoter specificity assumption and was compatible with effects previously seen with E2F1 ADMA modification by PRMT1, the inhibition of cell growth (Figure4.8 cell growth colony assay) and pro-apoptosis (Figure4.6 sub-G1 analysis) observation.

B

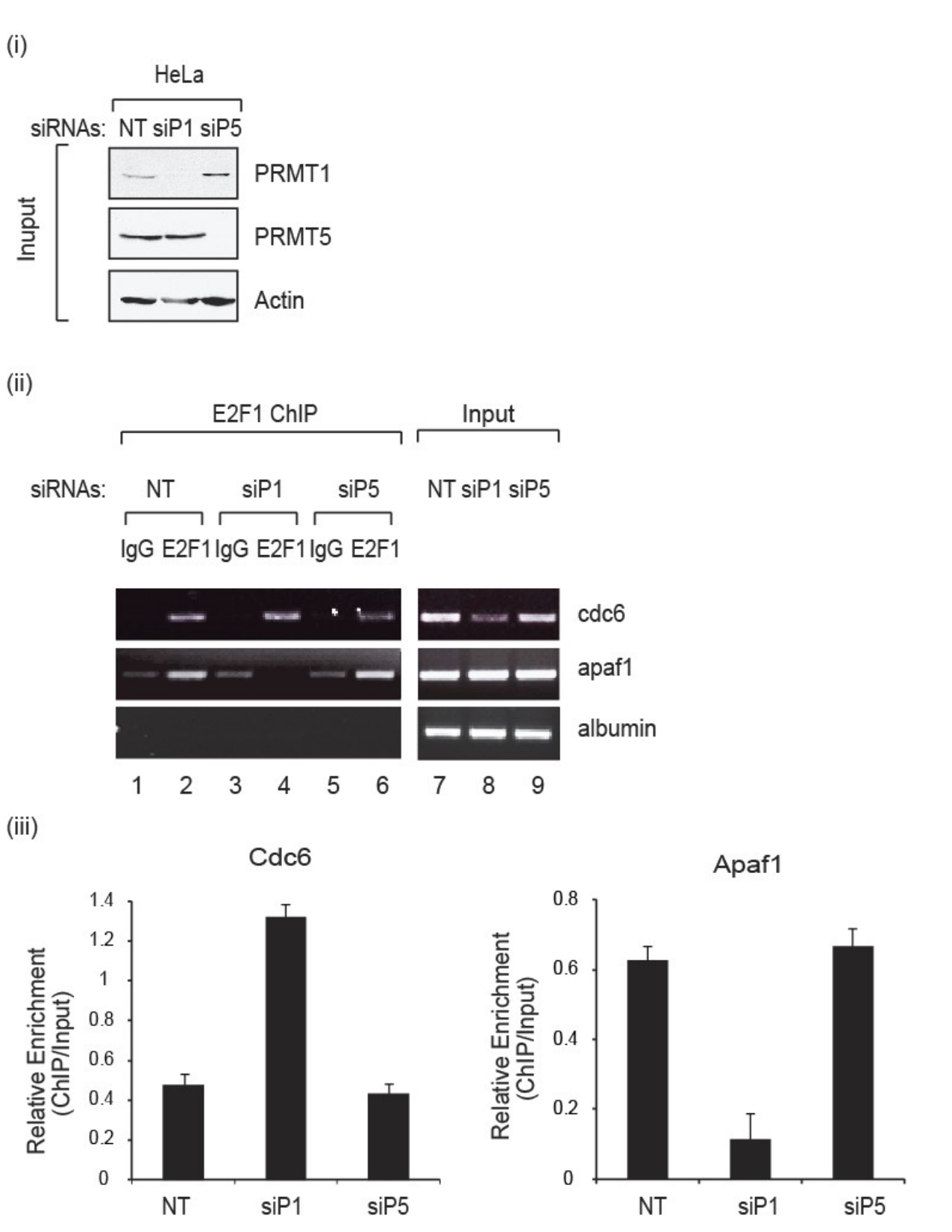

(i)
HeLa
siRNAs: NT siP1 siP5
Inuput
PRMT1
PRMT5
Actin
(ii)
E2F1 ChIP
Input
siRNAs: NT siP1 siP5 NT siP1 siP5
IgG E2F1 IgG E2F1 IgG E2F1
cdc6
apaf1
albumin
1 2 3 4 5 6 7 8 9
(iii)
Cdc6
Relative Enrichment (ChIP/Input)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
NT siP1 siP5
Apaf1
Relative Enrichment (ChIP/Input)
0.8
0.6
0.4
0.2
0
NT siP1 siP5

Figure4.8 (B) E2F1 promoter specificity upon PRMT depletion

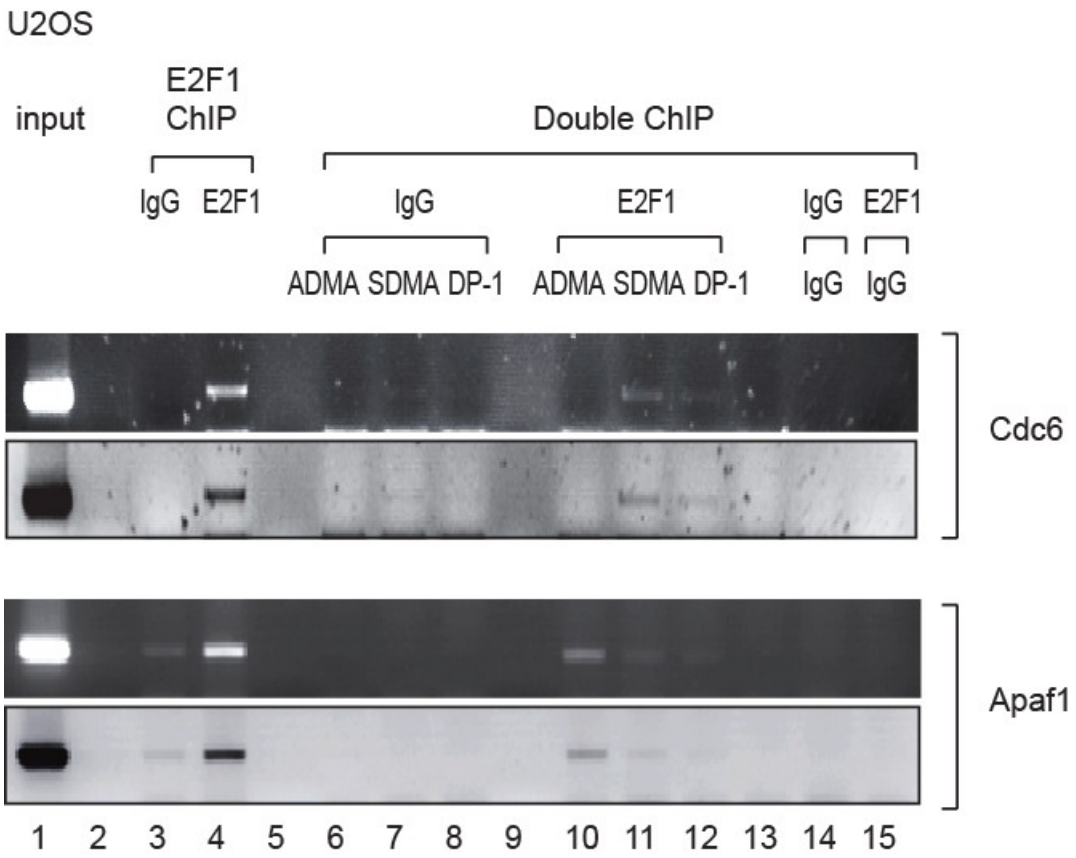
(B)(i) HeLa cells were transfected with 50nM NT, PRMT1, PRMT5 siRNA for 72 hours before harvesting for E2F1 chromatin immunoprecipitation. The transfection was confirmed with immunoblotting using PRMT1, PRMT5 and actin antibodies. (ii) 1μg of non-specific or E2F1 antibodies was used to immunoprecipitate chromatin from cross-linked and sonicated cells. Purified input and ChIP DNA were amplified with Cdc6, Apaf1 and albumin primers by PCR. The PCR products were visualised on 0.5μg/ml ethidium bromide containing 2% agarose gel. Data shown was the best representative of three different experiments. (iii) Quantitation of E2F1 binding on different promoters were performed by Image J.

To strengthen the arginine methylation with promoter specificity hypothesis, a double chromatin immunoprecipitation (double ChIP) was attempted. The idea was to investigate which promoter the asymmetric dimethylated E2F1 versus symmetric dimethylated E2F1 bind to on the transcriptional level. Representative E2F1 cell cycle promoter Cdc6 and, apoptotic promoter, Apaf1 were included in the methyl specific double ChIP experiments. The primary chromatin immunoprecipitation (ChIP) was done using E2F1 specific antibody while the secondary ChIP was done by using specific dimethylation antibodies (Asym or Sym) from the existing primary E2F1 ChIP materials. Known E2F1 dimerisation partner DP-1 was included as a positive control.

Revealed by the primary ChIP, E2F1 bound to both Cdc6 and Apaf1 as predicted as the E2F1 targeting promoters (Figure4.8C lane4). The non-specific control IgG binding remained at the background level on both promoters (Figure4.9C lane3). This suggested a successful chromatin immunoprecipitation as the specific ChIP (E2F1) was clear compared to the non-specific ChIP control (Figure4.8C lane4 and 3). The secondary ChIP was executed from the primary E2F1 ChIP materials hence the results allowed an investigation on top of the endogenous E2F1 chromatin binding of specific arginine dimethylation, both asymmetric and symmetric, on the transcriptional level. The negative controls for the secondary ChIP remained low at the background level. These are the secondary non-specific IgG ChIP (Figure4.8C lane 14 and 15) demonstrating their counterparts, ADMA

(asym), SDMA (sym) and DP-1 binding of Cdc6 or Apaf1 were specific by comparison. The DP-1 served as a positive control as it was known to bind to E2F1 and then the biding of promoters Cdc6 and Apaf1 (Figure4.8C lane 12).

These experiments focused on the differences between asymmetrically dimethylated E2F1 versus symmetrically dimethylated E2F1 on the promoter binding specificity using Cdc6 and Apaf1 as examples. The ADMA (asym) signal had E2F1 preferentially bind to the apoptotic promoter Apaf1 (Figure4.8C lane10) whilst the SDMA (sym) signal the cell cycle progression promoter Cdc6 predominantly(Figure4.8C lane11). It is interesting while endogenous E2F1 bound to both Cdc6 and Apaf1, different arginine methylation signals dictated promoter binding preferences. This was an example supporting how promoter specificity could serve as one of the mechanisms for the duel functions E2F1 possessed.



ADMA: Asymmetrically di-methylated arginine (lane 6, 10)

SDMA: Symmetrically di-methylated arginine (lane 7,11)

Figure4. 8 (C) E2F1 promoter specificity upon asymmetric or symmetric di-methylation.
(C) U2OS cells were subjected to double chromatin immunoprecipitation (Ch2.14-3). 2μg control IgG or E2F1 antibodies immunoprecipitated chromatin from the cross-linked and sonicated cells. The elutions from primary immunoprecipitation were collected and divided equally by volume (125μl) into the repeated immunoprecipitation using 1μg asymmetric di-methyl arginine (ADMA), 1μg symmetric di-methyl arginine (SDMA), and 1μg DP-1 antibodies diluted in 875 μl lysis buffer without SDS and DOC components (Ch2.14-3) respectively. The reverse cross-linked process released chromatin captured by immunoprecipitation and the purified chromatin was amplified by PCR with Cdc6 and Apaf1 primers. The PCR products were visualized on ethidium bromide containing 2% agarose gel. Data shown was the best representative agarose gel images of two independent experiments (the inverted figures shown were for the better visualisation purpose and therefore two figures per each promoter were shown). The primary non-specific IgG (lane2) and the secondary non-specific IgG (lane14,15) served as negative control for specific E2F1 chromatin immunoprecipitation. The E2F1 dimerisation partner, DP-1, was selected as positive control.

Chapter 5 E2F1 ChIP-sequencing

Ch5.1 Introduction

Given that the preliminary results were encouraging in search of an explanation for duel functions of E2F1 and were concluded from a small pilot panel of promoters, it was reasonable to follow up with the global search using Chromatin Immunoprecipitation Sequencing (ChIP-seq) technique. One of the strengths ChIP-seq provided was its ability in decoding the entire DNA-protein relationship network. This is done by integrating the chromatin immunoprecipitation (ChIP) with massively parallel DNA sequencing. The aim was to probe more into E2F1 as a master transcription factor and the influence of arginine methylation. The completion of ChIP-seq could be principally divided into three phases. The preparation of the ChIP-seq grade ChIP samples, the actual sequencing process and finally the bioinformatics analysis (Figure5.1).

Going forward, the better physiologically relevant siRNAs approach, as oppose to ectopic transfection, was adopted. The workflow breakdown could be simplified into six stages (Figure5.2). The process began with endogenous E2F1 immunoprecipitated from control non-targeting (NT), PRMT1 (siP1) siRNA or PRMT5 (siP5) siRNA treated human cancer cells (Figure5.2 stage1). The goal was to understand the impact of these two types of arginine methyltransferases on genome-wide chromatin, and more importantly, whether the previous promoter specificity pattern could be applied when compared E2F1 DNA interactions between endogenous PRMT1 and PRMT5 depletions. At the ChIP stage, the

inputs which chromatin were immunoprecipitated from, were kept for later use as background control to identify ChIP specific features by comparison.

Both input and the E2F1 ChIP then went through DNA library preparation stage. The goal of this stage was to coordinate DNA fragments into the final products: the thymine (T)-based adaptor sequence attached fragments, sized around 500bp. The adding of the T-based adapter sequence was done by utilising polymerase that adding poly adenine (A) tails to the ChIP materials (both input and ChIP DNA). Therefore by the adenine-thymine complementary principle, the T-based adaptor sequence could be edited on to the DNA. Considered the T-based adaptor sequence as a tag and was attached and edited onto the end of each input or E2F1 ChIP DNA, the later amplification and size selection (approx. 200-500bp) contributed to the formation of DNA library (Figure5.2 stage2). The design of siRNAs treated experiments ended up building six DNA libraries (NT input, NT E2F1 ChIP, siP1 input, siP1 E2F1 ChIP, siP5 input, and siP5 E2F1 ChIP).

Throughout the ChIP-sequencing stage (Figure5.2 stage3) the DNA base pair complementarity principle (thymine (T) - adenine (A); cytosine (C) – guanine (G) complementarity) was incorporated repeatedly. The ChIP-sequencing platform (Illumina Cluster Station®) contained sequence that is complimentary to the T-based adaptor sequence tag and therefore provided a mean to attach the DNA library materials (tagged input or E2F1ChIP) on to the system using the complementarity principle. The platform

attached DNA library was used as templates in the process of massive parallel sequencing. As nucleotides that used input or ChIP DNA as templates to sequence were florescent and once excited, emitted specific signals corresponding to the specific nucleotide, the raw reads of the entire DNA library sequence were therefore generated (Figure5.2 Stage3).

The following stages aimed to translate the recorded DNA sequence raw reads into meaningful biological contents by the bioinformatics analysis. The raw reads were aligned to human genome build hg19, which was released and updated by University of California Santa Cruz (UCSU) on 2009 from the initial sequencing and analysis of human genome (Lander et al., 2001), by the short read aligner, Bowtie (John Hopkins University) (Figure5.2 stage4).

The Model-based analysis of ChIP-seq (MACS) is a computational algorithm help identify genome-wide regions where aligned reads were enriched (Feng et al., 2012). The inputs (NT input, siP1 input, siP5 input) were used as controls to identify where E2F1 ChIP (NT ChIP, siP1 ChIP, siP5 ChIP) reads were enriched respectively. The normalised ChIP sequencing enrichments were defined as “peaks,” the “peak calling” by MACS therefore presented the genome-wide result of human genome regions where chromatin immunoprecipitated by target protein (E2F1), bound and enriched (Figure 5.2 Stage5).

As the chromatin regions E2F1 targeted revealed, these regions (peaks) were annotated as being linked with a gene if they lay either within the gene body of a RefSeq

gene (NCBI Reference Sequence Database), or within 20 kb of the transcription start and end sites of the gene, using Cis-regulatory Element Annotation System (CisGenome 2.0) (Ji et al., 2008) (Figure5.2 Stage6).

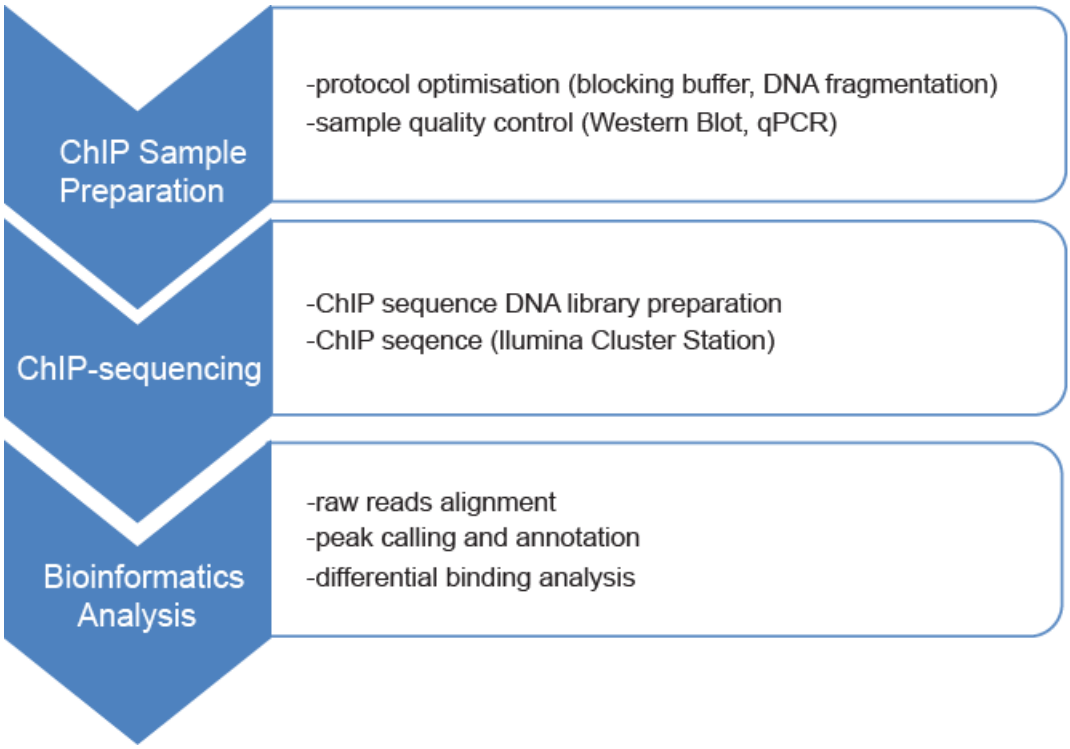


Figure5. 1 Overview of ChIP-sequencing workflow

The workflow of ChIP-sequencing principally divided into three phases. Firstly, the "ChIP-sequencing grade ChIP sample preparation" was achieved by standard chromatin immunoprecipitation optimisations and quality controls. Secondly, the massively parallel sequencing technique provided by Illumina Cluster Station applied to the DNA library, which was edited and amplified from ChIP samples submitted. Thirdly, the bioinformatics analysis was accomplished by many bioinformatics tools. The initial sequencing raw reads were aligned to human genome using Bowtie 0.12.8. The model-based Analysis of ChIP-seq (MACS) was the algorithm used to identify the genome-wide enrichment of chromatin, or the peaks. The peaks (chromatin enrichment) were annotated by CisGenome which adopted RefSeq definition of human genes information. Many primary analyses data directly derived from the algorithm (MACS2.0.10), such as the differential binding analysis.

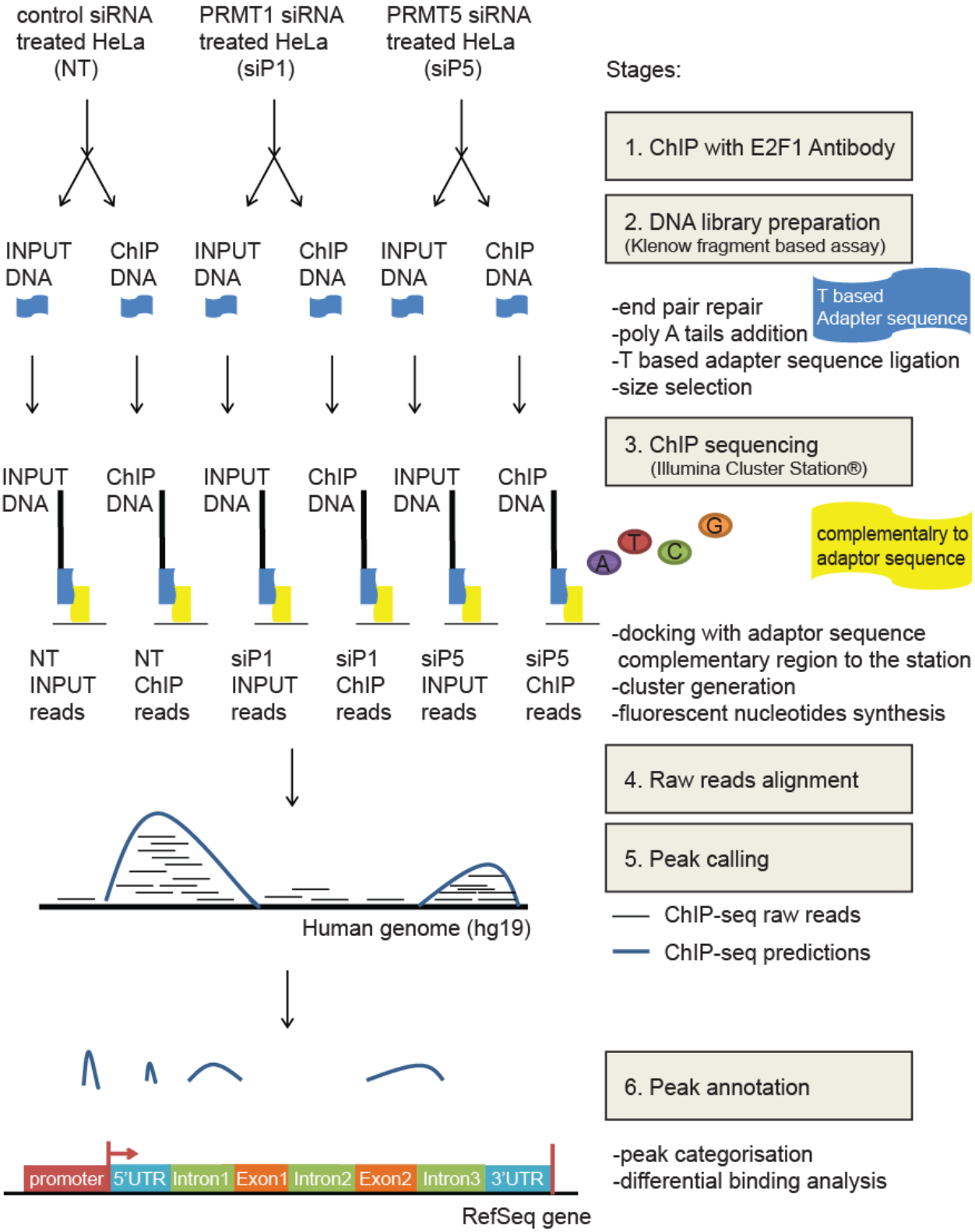


Figure5. 2 ChIP-sequencing workflow breakdown

Schematic diagram describes the stages breakdown of the ChIP-sequencing workflow.

Ch5.2 E2F1 ChIP-sequencing samples preparation

To maximise the success rate and the use of the technology, many steps have been taken to prepare a set of samples. Although the basic principles of ChIP were the same, the quality requirement and the scale for smooth completion of ChIP-seq needed further optimisation. These included some trial and error attempts in the ChIP process, for example, on the non-specific binding blocking agent choice: to avoid selecting of salmon sperm, used as a blocking agent in case of possible contamination of fish DNA in the human sample, or the altering the resin-based with the magnetic beads-based IP materials for an optimal output for both efficiency and cleanness.

One of the difficulties lay in the process of fragmenting DNA into a suitable size as it influenced the intake for the later DNA library preparation, which is crucial for the overall ChIP-seq outcome. The optimisation steps included firstly evaluating the time and volume factors, and secondly the concentration of samples. To understand the volume and time factors impacts on DNA fragmentation outcome, small scale siRNA-transfected samples were treated with the different conditions (Figure5.3A). Ethidium bromide (EtBr) is commonly used as an intercalating agent, which when excited by UV, reflected proportionally to the amount of DNA. As demonstrated, the dilution (every 2 dishes in lysis buffer volume from 1ml to 2ml) or prolonged duration of time (programme unit from 1 to 2) did not influence the fragmented the DNA. Not only the size of the DNA of all

conditions concentrated around 3000 bp, the amount of the DNA amount deteriorated when compared with the initial condition (Figure5.3A lane1).

The sonication time remained the same however, instead of diluting the lysed cells, the cell concentration became five-fold and two-fold (Figure5.3B). Unlike conventional ChIP, where the DNA isolated from protein of interest may vary and were acceptable beyond 3000 bp (Figure5.3B Set1), ChIP-seq required DNA fragment sizes much below 1000bp to better define protein-DNA interrelation when mapping binding sites globally. The results allowed optimised conditions to be established, where the majority of DNA across siRNAs treated samples was around and below 1000bp (Figure5.3B Set2).

A

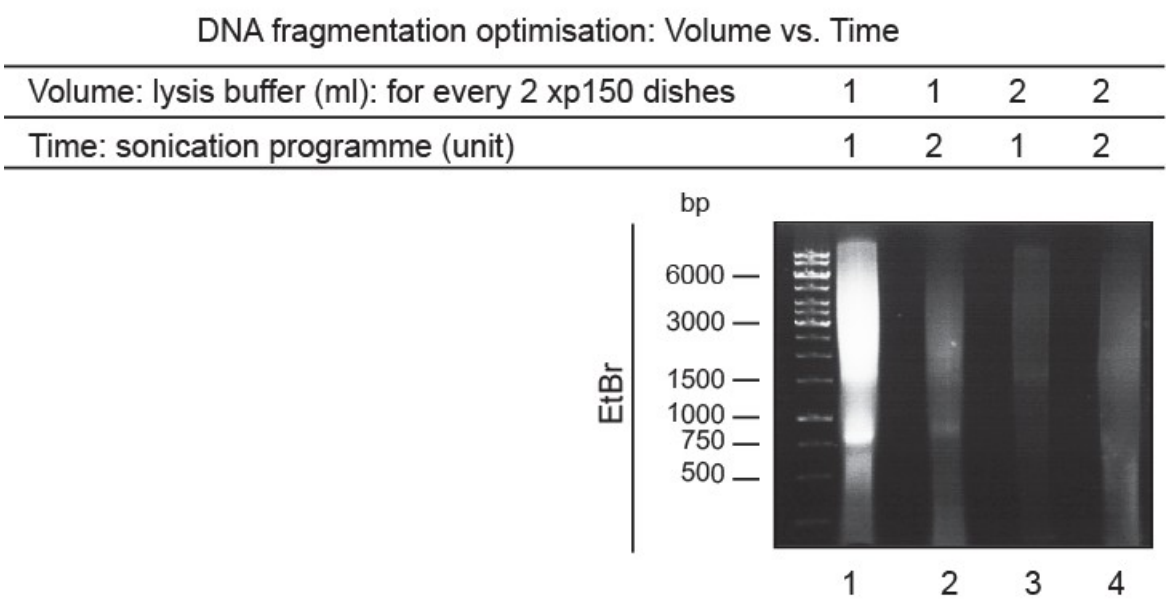


Figure5. 3 ChIP-sequencing DNA fragmentations

(A) HeLa cells with 60% confluency pre-seeded on two cell petri dishes (150mm) were lysed and sonicated with indicated combination before reverse cross-linked with 40μg/ml proteinase K and 20μg/ml RNAse A, for 3 hours at 55°C followed by 65°C overnight on the hot block. The products were loaded on 2% agarose gel containing 0.5μg/ml ethidium bromide (EtBr) and excited by UV for visualisation. One sonication programme unit was set as 25% power for 10 seconds repeated 10 times at (Programme set up at Branson Digital Sonifier 450).

B

DNA fragmentation optimisation: Concentration

Concentration: $X = \frac{2 \text{ x p150 dishes}}{1 \text{ ml lysis buffer}}$	5X	2X
Time: sonication programme (unit)	1	1

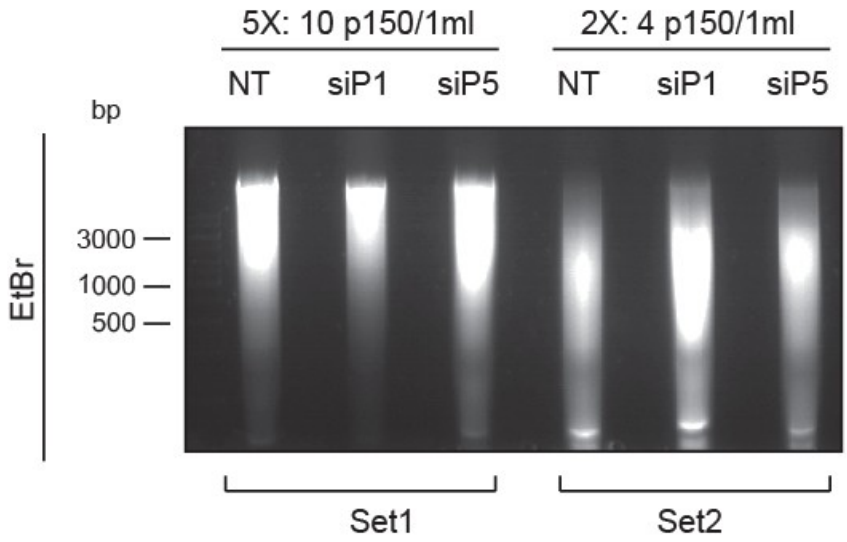


Figure 5.3 ChIP-sequencing DNA fragmentations

(B) HeLa cells were treated with 50nM non-targeting, PRMT1, PRMT5 siRNAs (NT, siP1, siP5) for 72 hours on the scale of 14 cell petri dishes (150mm) per siRNA treatment. Cells harvested and divided into two different concentrations, 10 cell petri dishes (150mm) lysed in 1 ml lysis buffer (Set 1) or 4 cell petri dishes (150mm) in 1 ml lysis buffer (Set 2) for previously optimised sonication power and time (10 repeats of 25% power for 10 seconds at Branson Digital Sonifier 450). The sonicated lysates were reverse cross-linked with 40μg/ml proteinase K and 20μg/ml RNAse A, 3 hours at 55°C followed by 65°C overnight on the hot block. The products were loaded on 2% agarose gel containing 0.5μg/ml ethidium bromide (EtBr) and excited by UV for visualisation.

The quality control of samples focused on the specificity of ChIP, the meaningful E2F1 ChIP defined as statistically significant binding on targeted promoters compared to non-specific Immunoglobulin G (IgG) immunoprecipitation. The large scale of PRMT1 or PRMT5 siRNAs transfection was confirmed with immunoblotting. The protein level of PRMT1 or PRMT5 decreased with respective siRNAs transfection while the housekeeping actin protein level remained constant (Figure5.4A).

In the semi-quantitative PCR approach, with chromatin immunoprecipitated E2F1 from non-targeting (NT) or PRMT depleted cells, specific binding was seen on E2F1 promoter compared to that of control IgG group (Figure5.4B lane 2,4,6 versus 1,3,5). By quantitative PCR (qPCR) in both Cdc6 and Aparf1 promoters, the fold-enrichment (ChIP/input) was statistically significant from all siRNAs treated groups (NT, siP1 or siP5) compared to their controls (Figure5.4C(i) and (ii)). This means the E2F1 ChIP was specific and ready for the ChIP-seq analysis. It is interesting to note that the promoter specificity pattern was reproducible in this initial ChIP preparation, for example, the depletion of PRMT1 caused E2F1 to preferably located on cell cycle regulation Cdc6 promoter rather than the apoptotic Aparf1 promoter ((Figure5.4C(i) versus (ii)).

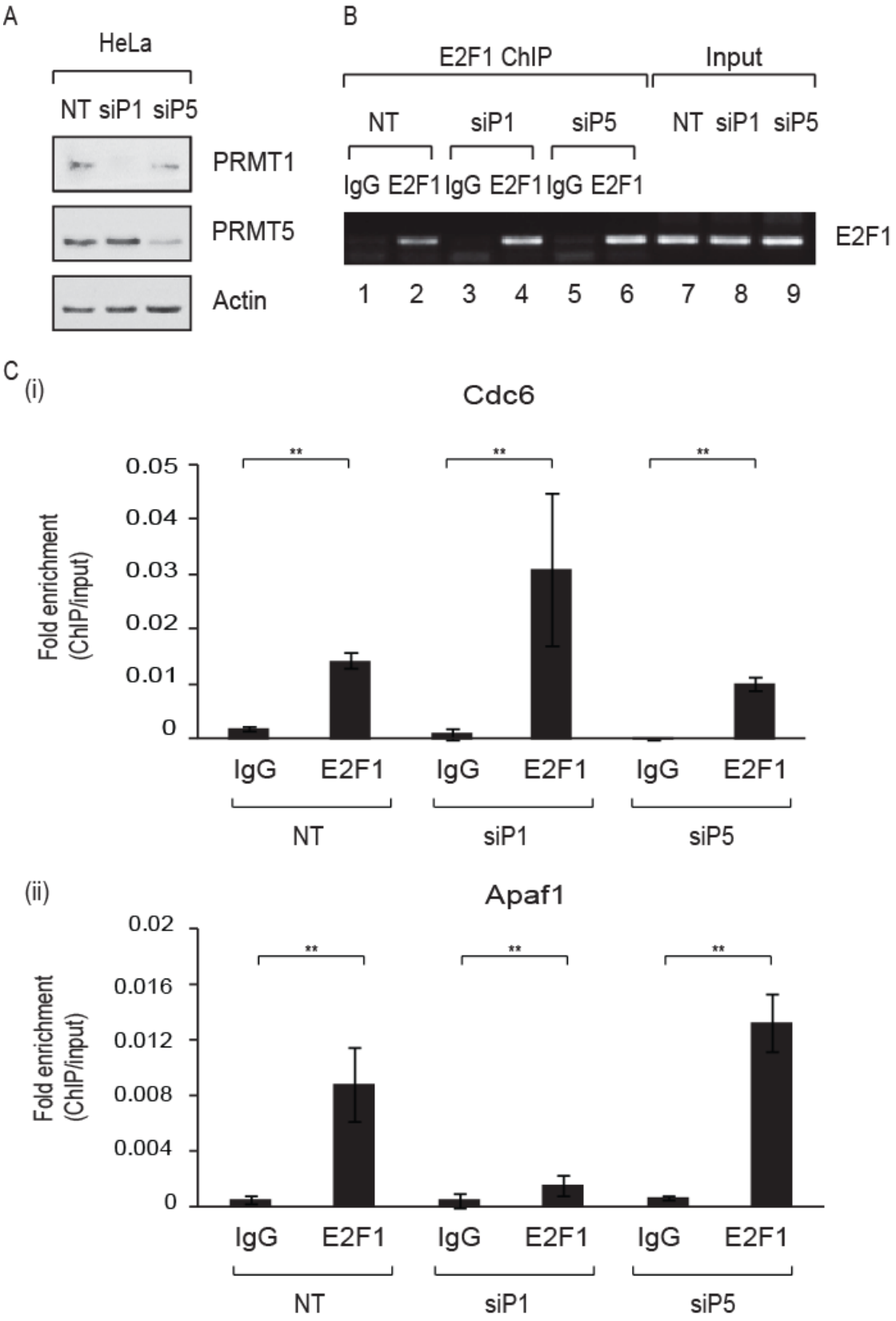

A
HeLa
NT siP1 siP5
PRMT1
PRMT5
Actin
B
E2F1 ChIP
Input
NT
siP1
siP5
NT siP1 siP5
IgG E2F1 IgG E2F1 IgG E2F1
E2F1
1 2 3 4 5 6 7 8 9
C
(i)
Cdc6
Fold enrichment (ChIP/input)
0.05
0.04
0.03
0.02
0.01
0
**
IgG E2F1 IgG E2F1 IgG E2F1
NT siP1 siP5
(ii)
Apaf1
Fold enrichment (ChIP/input)
0.02
0.016
0.012
0.008
0.004
0
**
IgG E2F1 IgG E2F1 IgG E2F1
NT siP1 siP5

Figure5. 4 ChIP-sequencing samples quality control

HeLa cells were pre-seeded on the scale of 21 of 150mm cell petri dishes per siRNA treatment. 50nM non-targeting, PRMT1 or PRMT5 siRNAs (NT, siP1, siP5) were transfected into the cells for 72 hours. Cells were then subjected for immunoblotting (1x 150mm cell dish) and chromatin immunoprecipitation (20x 150mm cell dishes). (A) Cells were harvested and lysed with TNN buffer, separated by size using gel electrophoresis. Actin antibody was used as loading control and PRMT1, PRMT5 antibodies were used to confirm the depletion PRMTs by immunoblotting. (B) Cells subjected for chromatin immunoprecipitation were cross-linked by fixing with 3.7% formaldehyde and neutralised by 0.125M glycine. The cross-linked cells were harvested and sonicated with pre-optimised time and concentration (10 repeats of 25% power for 10 seconds at Branson Digital Sonifier per every 4 150mm cell petri dishes amount of cells lysed in 1 ml ChIP lysis buffer). 1μg of non-specific IgG with 5-10 μl magnetic Dynabeads® were used for 3 hours pre-clear at 4°C while 20-25μl magnetic Dynabeads were mixed with 5mg/ml BSA and 2 μg non-specific IgG or 2μg E2F1 antibodies for 4 hours to prepare magnetic Dynabeads-Antibody complex mix. The pre-clear of lysates completed when IgG-Dynabeads were removed by magnetic rack. The pre-cleared lysate were divided into Dynabeads-IgG or Dynabeads-E2F1 for immunoprecipitation overnight incubation at 4°C. The input was kept before washes and elution, the reverse cross-linked was done with 3 hours at 55°C followed 65°C overnight on both input and the elution. The immunoprecipitated chromatin was then purified by Qiaquick PCR kit (Qiagen). The purified chromatin was subjected to semi-quantitative PCR or quantitative PCR (qPCR). The chromatin was amplified by semi-quantitative PCR using E2F1 primers, the PCR products were loaded on 2% agarose gel containing 0.5μg/ml ethidium bromide (EtBr) and excited by UV for visualisation. (C) The purified chromatin was subjected to quantitative PCR (qPCR) using the Brilliant III SYBR Green QRT-PCR master mix on the MX3005P QPCR system (Stratagene) and analysed by MXPro QPCR software. The relative quantities were plotted by CT value obtained (threshold value) against serial dilution of positive control sample (input). The data represented the mean values from three reactions and are shown as fold change, relative to the untreated sample (input), with errors bars indicating standard deviation. The Student t-test was used for statistical significance P value< 0.05**. (i) E2F1 ChIP versus input on Cdc6 promoter (ii) E2F1 ChIP versus input on Apaf1 promoters. Data shown was representative of three independent experiments. ChIP-sequencing samples were required of statistically significant qPCR specificity (IgG control versus specific (E2F1) being specifically ChIPed) before DNA library preparation and ChIP-sequencing.

Ch5.3 E2F1 ChIP-sequencing analysis

The bioinformatics analysis tools helped to align ChIP-seq raw reads to the human genome, where each ‘peak’ was determined by the enrichment of ChIP-seq raw reads. The human genome assembly used was hg19 released and updated on February 2009 by University of California Santa Cruz (UCSU) from the 2001 initial sequencing and analysis of human genome (Lander et al., 2001).

Following the peak-prediction were the annotations, which assign the peaks (enrichment of reads) to the nearby genes by the definition of NCBI Reference Sequence Database (RefSeqGene) (Pruitt et al., 2000) (Pruitt and Maglott, 2001).

The chromatin immunoprecipitated by E2F1 from the siRNAs treated cells had been sequenced and the results were aligned and annotated against the entire human genome. As the alignment and annotation lined up ChIP-seq results with biological content, much information was made available.

There were three dimensions of the initial inspections on the ChIP-seq results. Firstly, where the chromosomal location the immunoprecipitated chromatin was. Secondly, what were the regions the immunoprecipitated chromatin in relations to genes, such as promoters, untranslated region (UTR), coding region or other distal intergenic region. Thirdly, the basic statistical overview on how many genes targeted and what they are.

As a DNA binding transcription factor, the ChIP-seq results allowed a genome-wide overview on the E2F1 chromatin binding location (Figure5.5). When compared with the distribution of human genetic materials on chromosomes, the 'Genome,' E2F1 ChIP-seq (NT, siP1 and siP5 treated cells) revealed a similar chromosomal distribution pattern. Although variation existed on each chromosome, the distribution pattern overall was similar with less than 1.5% differences across all chromosomes when compared between PRMT1 and PRMT5 depletions (siP1 versus siP5), less than 2.5% differences when compared PRMTs depletions (both siP1 and siP5) to control non-targeting siRNA treated cells (siPRMTs versus NT). The implication was that the depletion of PRMTs, regardless of PRMT1 or PRMT5, had limited effect on chromosomal distribution of E2F1 targets.

When inspected, these E2F1 DNA targets in relation to the nearby genes, enrichment of these E2F1 targets fell into different categories (Figure5.6). The most distinctive feature demonstrated by comparing the Genome and the E2F1 ChIPs (NT, siP1 and siP5) was that the majority of E2F1 ChIPs DNA fragments fell into the region defined as promoters rather than distal intergenic (the region that was not promoter, downstream, UTR, coding or intron) which comprised of 50% in the Genome (Figure5.6 highest percentage region: Promoter (NT, siP1 or siP5) versus Distal Intergenic (Genome)). The enrichment for DNA located within 1000bp to the promoters was more than thirty-six fold in percentage (Genome: NT: siP1: siP5 =1.10%: 36.30%: 38.80: 41.20%), or a minimum thirty-eight fold

increase in the region within 2000bp to the promoters (Figure5.6 combined percentages within 1000bp and between 1000bp-2000bp: Genome: NT: siP1: siP5 =1.80%: 38.20%: 41.20: 43.10%). Also the 5' untranslated region (5'UTR) had seen a distribution of 0.4% in the human genome (Genome) increased to 16.80%, 15.30%, 18% in the NT, siP1, and siP5 treated E2F1 ChIPs, which was a minimum more than thirty eight-fold increase (Figure5.6 5'UTR: NT, siP1 or siP5 versus Genome).

The initial inspections revealed that although the genetic material of human genome shared a similar pattern in chromosomal distribution with ChIP-seq result, the regions where these chromatins identified was a complete different picture. The majority of E2F1 ChIP demonstrated a preferable binding in the promoter area, whereas human genome located mostly in distal intergenic region which is not promoter, exon, intron or untranslated region.

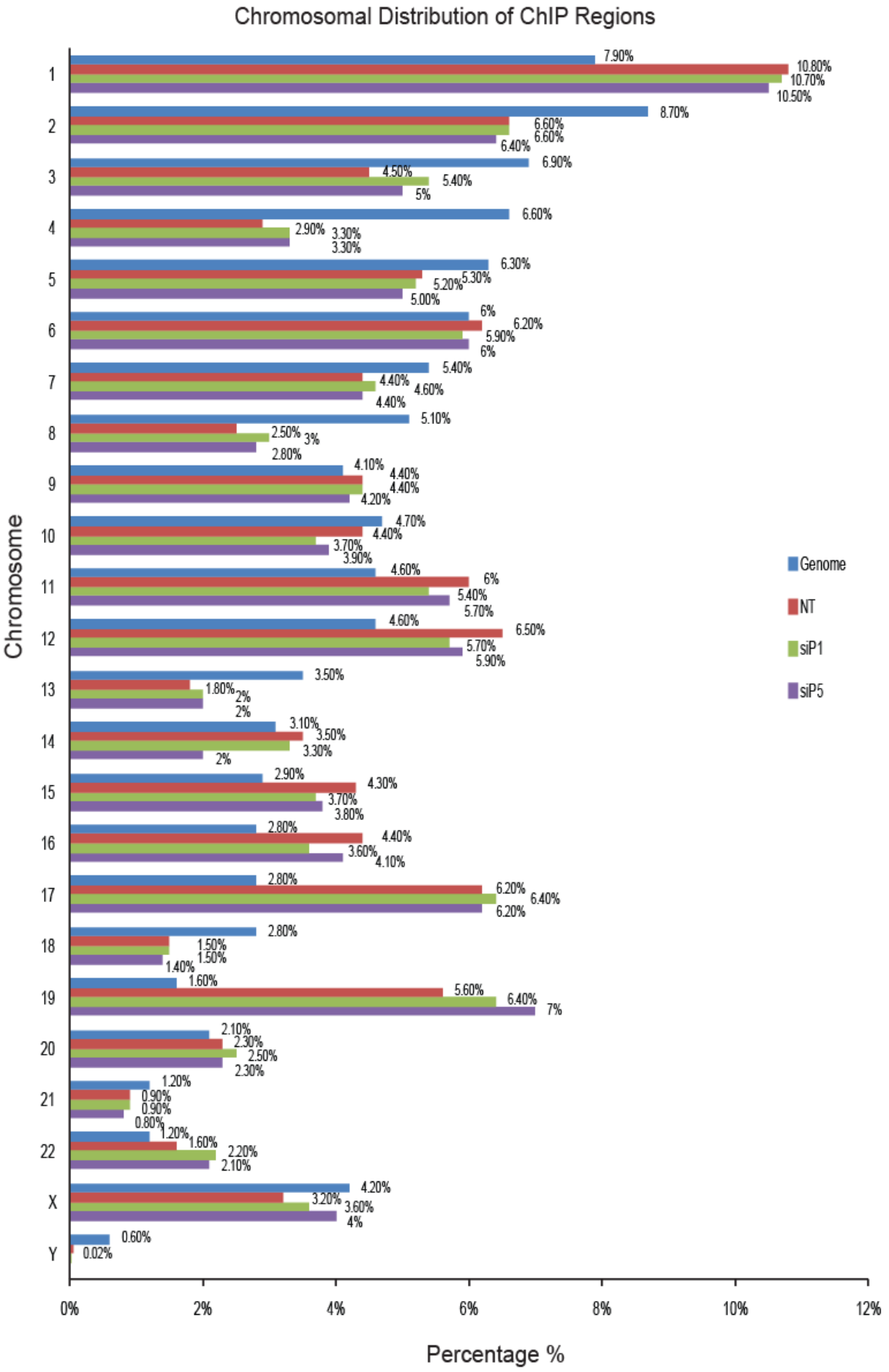

Chromosomal Distribution of ChIP Regions
Chromosome
Percentage %
Genome
NT
siP1
siP5
0%
2%
4%
6%
8%
10%
12%
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
7.90%
10.80%
10.70%
10.50%
8.70%
6.60%
6.60%
6.40%
6.90%
4.50%
5.40%
5%
6.60%
2.90%
3.30%
3.30%
6.30%
5.30%
5.20%
5.00%
6%
6.20%
5.90%
6%
5.40%
4.40%
4.60%
4.40%
5.10%
2.50%
3%
2.80%
4.10%
4.40%
4.40%
4.20%
4.70%
4.40%
3.70%
3.90%
4.60%
6%
5.40%
5.70%
4.60%
6.50%
5.70%
5.90%
3.50%
1.80%
2%
2%
3.10%
3.50%
3.30%
2%
2.90%
4.30%
3.70%
3.80%
2.80%
4.40%
3.60%
4.10%
2.80%
6.20%
6.40%
6.20%
2.80%
1.50%
1.50%
1.40%
1.60%
5.60%
6.40%
7%
2.10%
2.30%
2.50%
2.30%
1.20%
0.90%
0.90%
0.80%
1.20%
1.60%
2.20%
2.10%
4.20%
3.20%
3.60%
4%
0.60%
0.02%

Figure5. 5 ChIP-seq bioinformatics analysis: chromosomal distribution of ChIP regions.

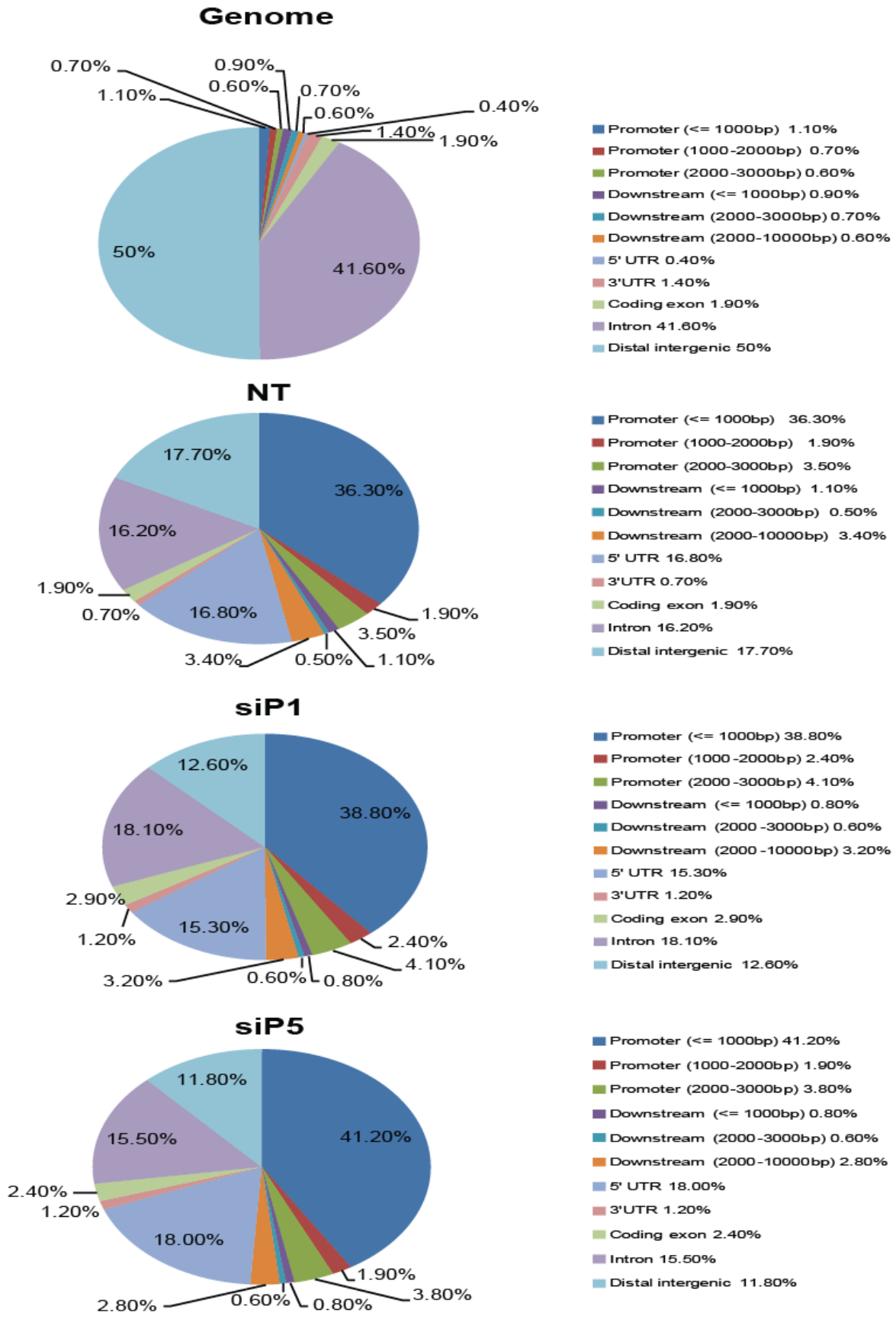


Figure5. 6 ChIP-seq bioinformatics analysis: categories of ChIP regions.

Figure 5.5 ChIP-seq bioinformatics analysis: chromosomal distribution of ChIP regions
Figure 5.6 ChIP-seq bioinformatics analysis: categories of ChIP regions.
The bioinformatics analysis of chromosomal distribution of ChIP regions and categories of ChIP regions were completed with the Cis-regulatory Element Annotation System (CEAS; version 1.0.2). With the RefSeq data incorporated with ChIP-seq raw data (BED file format) which contained the chromosomes, start and end locations of ChIP regions identified by peak caller (MACS 2.0.10), the algorithm provided integrated information on chromosomal distribution and categories (RefSeq definition) of ChIP regions.

After the analysis of ChIP-seq results on the chromosomal distribution and the type of categories that ChIPs have fallen into, a closer exam on E2F1 targets was performed (Figure5.7). The Model-based analysis of ChIP-seq (MACS; MACS 2.0.10) was employed as the algorithm to identify ChIP-seq peaks by aligning reads to human genome and by normalising ChIP reads against corresponding input reads. The peaks were therefore the enrichment of E2F1 ChIP. During the peak annotation process, when peaks were assigned to specific nearby genes (RefGene), some peaks located outside the regions related to any; sometimes multiple peaks were seen to be located in close proximity to the same gene, and thus the 'peak count' was always more than that of the 'Unique genes'.

In order to address the initial question, namely the overall impact of arginine methylation on E2F1 chromatin binding, how many genes (Unique genes) were broadly affected by the depletion of PRMT1 or PRMT5 was important to ascertain (Figure5.7). When compared the absolute numbers of genes affected, siP1 E2F1 ChIPs (4,876 genes) topped the table, followed by siP5 group (3,212 genes) with an approximate three-fold and two-fold more genes identified than the genes targeted by E2F1 in the non-targeting (NT) group (1,730 genes) (siP1 4,876/ NT 1,730=2.81; siP5 3,212/ NT 1,730=1.86) (Figure5.7A). These results indicated an overall greater influences by arginine methylated E2F1 on E2F1 chromatin binding. When compared the gene numbers affected between PRMT1 and PRMT5 depletion, one and half fold slightly more in siP1 than in siP5 (siP1:

4,876: siP5: 3,212). In another words, the impact of arginine methylation on E2F1 chromatin binding was marginally more by PRMT1 than PRMT5.

What is also interesting was the centre overlapping E2F1 targets shared in between control, siP1 and siP5 which was similar in size with the common targets between siP1 and siP5 but excluding the overlapped targets (Figure5.7B NT&siP1&siP5: 1466 genes versus siP1&siP5 excluding NT&siP1&siP5: 1499 genes). The centre overlapping genes comprised 84.7% of E2F1 targets genes in the NT control group (1466/1730=0.847). There were 2,965 shared E2F1 targets between siP1 and siP5 (1,466+1,499=2,965) which was 92.3% of E2F1 targets in the siP5 group (2965/3212=0.923). There were two conclusions from the statistical overview. Firstly, these two types of arginine methyltransferases covered the majority of traditional E2F1 chromatin binding targets (control group) and beyond, secondly, the majority of genes affected by PRMT5 were also targets of PRMT1.

A

Samples	Peaks count	Unique genes	Other genomic region
NT	2179	1730	343
PRMT1 siRNA	6343	4876	665
PRMT5 siRNA	3921	3212	391

Samples	Unique genes
NT & PRMT1 siRNA	1633
NT & PRMT5 siRNA	1487
PRMT1 & PRMT5 siRNA	1466
NT & PRMT1 & PRMT5 siRNA	1499

B

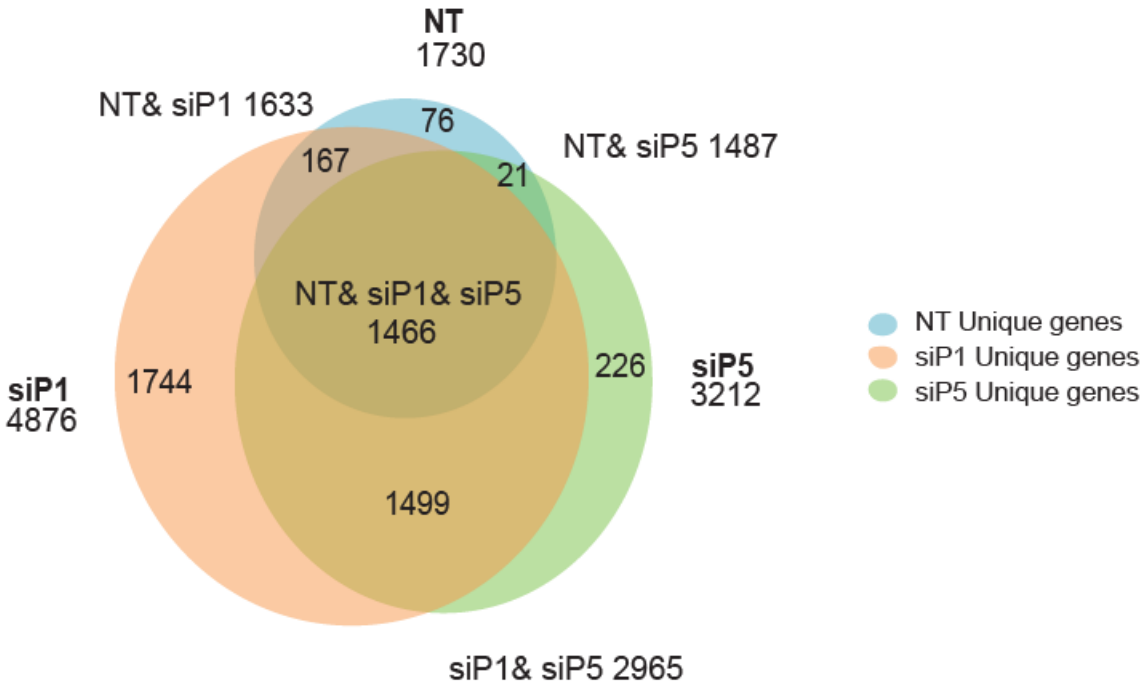


Figure5. 7 The ChIP-sequencing statistics

The Cis-regulatory Element Annotation System (CEAS; version 1.0.2) annotated ChIP-seq peaks to nearby genes according to RefSeqGene definition. The CEAS recorded the length, start and end of chromosome, and the related gene names of ChIP-seq. Data was secondary analysis by Microsoft Excel, using functions combination. “Filter” was established for individual column, and “custom sort” was used to highlight duplicated or singled genes and therefore enable the calculation of the numbers of unique or overlapped genes across ChIP-seq NT, siP1 and siP5 data.

Beyond the absolute numbers, a critical part of ChIP-seq was the quantitative information provided. Amongst information made available were the exact binding location and the quantitative fold change in binding enrichment on each respective location. To advance from the numbers of genes affected to the amount of enrichment, the differentially bound peaks were identified using the "bdgdiff" function of the algorithm MACS 2.0.10 (Zhang et al., 2008) (Figure5.8). This enabled a three way comparison to find out where in the genome has differential enrichment between two conditions. The prerequisite was that this region should be at least enriched in either condition. This means all genes identified in previous analysis (Figure5.7) qualified and were included in the analysis. A log10 likelihood ration cut-off was applied.

When compared to non-targeting control, just over half of E2F1 ChIP (56.06%) from PRMT1 depletion enhanced their binding whilst almost half of E2F1 ChIP region (43.86%) had a similar or insignificant change, in the amount binding (Figure5.8 siP1 vs NT). Reflected on the numbers of genes E2F1 bound to, depletion of PRMT1 had almost three-fold more genes bound by E2F1 compared to non-targeting siRNA treatment (Figure5.7). Amongst these different E2F1 bound genes by numbers, nearly half of the enrichment by amount was statistically insignificant and just over half of these genes favoured an enrichment in E2F1 binding when PRMT1 was depleted.

The depletion of PRMT5 versus non-targeting control identified nearly two-fold increase in gene numbers however the differential binding analysis showed one fifth (20.02%) of genes had an increase in E2F1 binding as opposed to nearly 80% genes remained similar in the amount enriched (Figure5.8 siP5 vs NT).

In the effort to identify PRMT influences on genome-wide E2F1 chromatin binding, when directly compared the change in E2F1 ChIP enrichment between PRMT1 and PRMT5 depletions, the differential binding analysis revealed the majority of genes (97.36%) remained similar in the amount enriched whilst 2.28% of genes reflected an enhanced binding if PRMT1 was removed (Figure5.8 siP1 vs siP5).

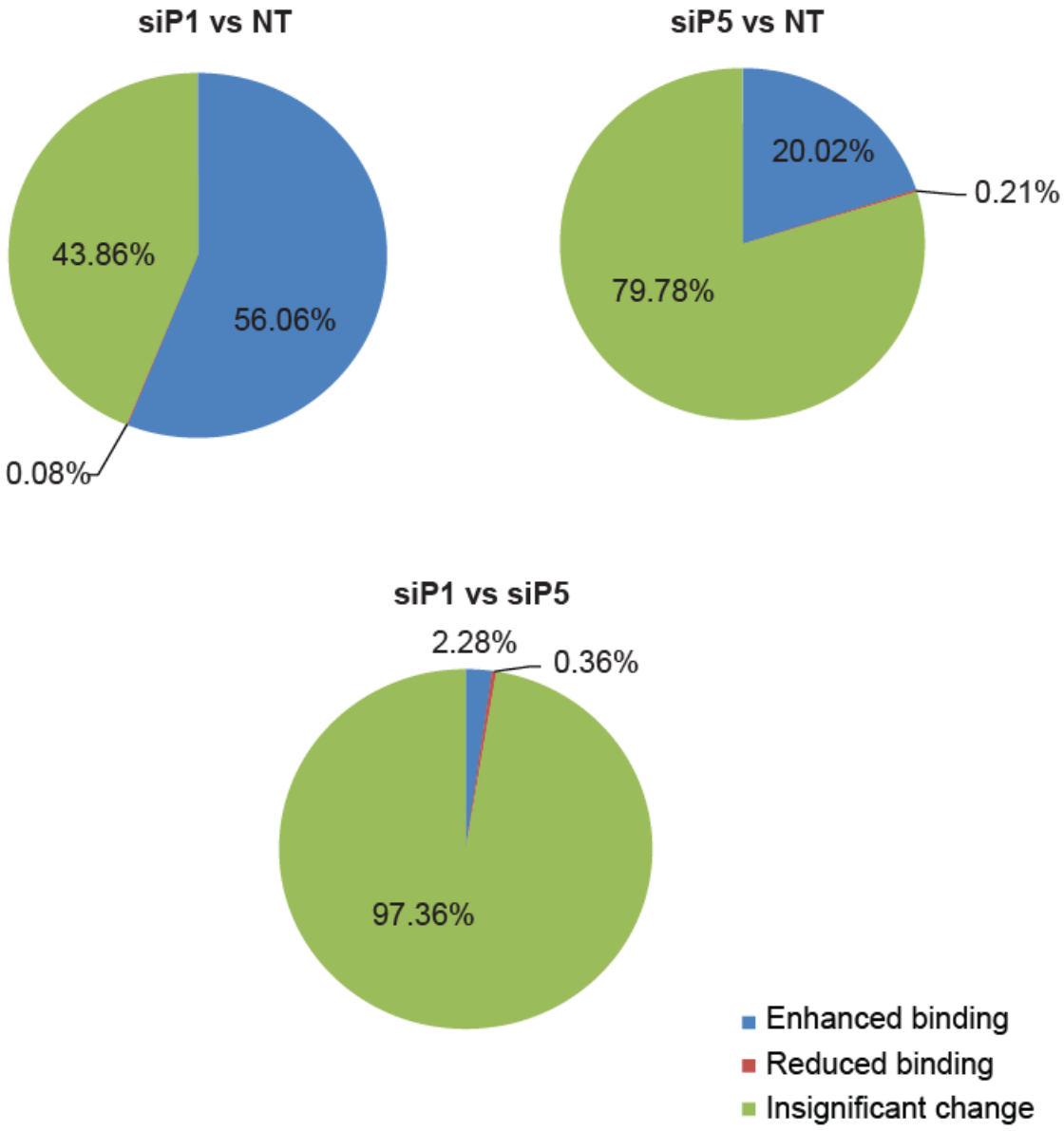


Figure5. 8 ChIP-sequencing bioinformatics analysis: differential binding.

The “bdgdiff” function of algorithm MACS2.0.10 was applied on ChIP-seq data with a log10 likelihood ratio cutoff point to define statistical significance in binding whether it is Enhanced (blue), Reduced (red) or Insignificant (green) changes. The report achieved by this three way comparison to find out where differential enrichment between two conditions has occurred. Data shown was all combination of three E2F1 ChIP-seq, siP1 vs NT, si P5 vs NT and siP1 vs siP5.

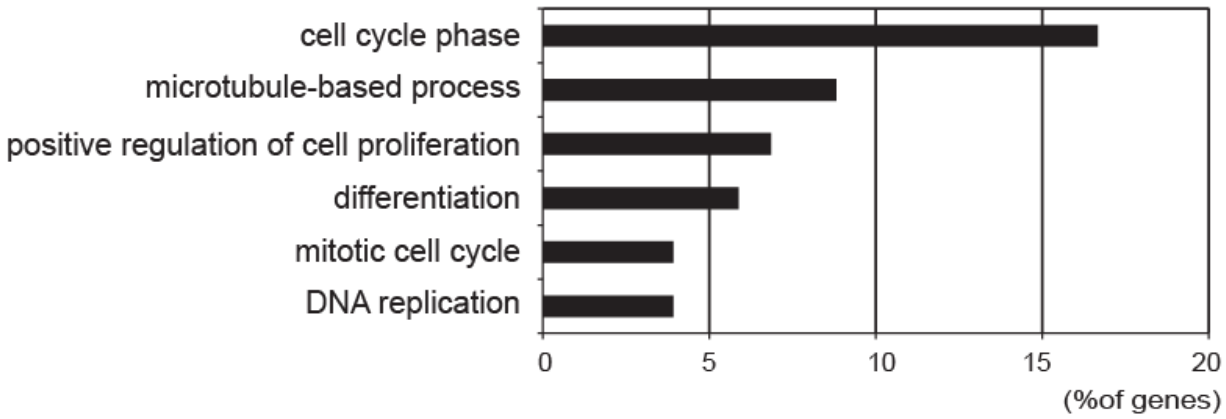
Although the comparison between depletions of two PRMTs on E2F1 chromatin binding showed a relatively small portion of genes (2.28% enhanced binding, or 0.36% reduced binding) demonstrated differences in the amount enriched (Figure5.8 siP1 vs siP5), it was intriguing even more on these hits that were highlighted in the differential binding analysis. To systematically examine these genes, the bioinformatics programme Database for Annotation, Visualisation and Integrated Discovery (DAVID v.6.7; sixth version updated in 2009) (Huang da et al., 2009) was employed to functionally annotate these hits identified by differential binding comparison between PRMTs depleted E2F1 ChIP-seq data sets (Figure5.9).

The programme DAVID integrated many bioinformatics database and tools such as Gene Ontology (GO) terms, a unified representative gene attributes (Gene Ontology, 2008), Swiss-Prot (SP) Protein Information Resource (PIR) (Wu et al., 2003), a collection of databases that combine protein sequence, classification or annotation integrated protein informatics, to help functionally annotate genes into different clusters.

When analysed genes highlighted as enhanced bound by differential binding analysis comparing the amount bound by E2F1 ChIP from PRMT1 depletion versus that from PRMT5 depletion cells, top functional clusters revealed to be highly cell cycle related (cell cycle phase, mitotic cell cycle , DNA replication), and when ranked by the percentage of

genes in each cluster, the third functional cluster was classified as positive regulation of cell proliferation which took up more than 5% of genes that enhanced bound by E2F1 when PRMT1 was depleted (Figure5.9A,B). The analysis indicated that when PRMT1 was removed, E2F1 found to be preferably bound to genes that associated with positive regulation of cell proliferation and other cell cycle related genes. This implication echoed the growth inhibition indication that PRMT1 was associated with (Ch4.2) as the removal of PRMT1 released E2F1 to be bound on genes that clustered into genes that related to positive cell proliferation.

A



B

Functional Clusters	ES	p-value	%Genes	Genes
cell cycle phase	5.85	8.94E-10	16.67	PDS5A, SAC3D1, NCAPG2, FBXO3
microtubule-based process	2.65	8.43E-05	8.82	KIF23, KIFC1, RANBP9, SAC3D1
positive regulation of cell proliferation	1	2.87E-02	6.86	BLM, GNAL2, ADM, RTKN2, ICMT
differentiation	1.43	8.31E-02	5.88	PEG10, SEMA6B, HES4, TLL1
mitotic cell cycle	1.33	2.02E-02	3.92	BLM, CDKN2C, POLE, RCC1
DNA replication	1.29	1.00E-02	3.92	RFC4, BLM, POLE, REV3L

Figure5. 9 Functional annotation on differential binding analysis hits of PRMTs depleted E2F1 ChIP-seq.

(A) The input data was the differential binding analysis hits which E2F1 enriched in binding, when compared E2F1 ChIP-seq from PRMT1 depletion to PRMT5 depletion. The functional annotation was performed using the programme Database for Annotation, Visualisation and Integrated Discovery (DAVID) 2008. The parameters set to apply were Gene Ontology (GO) terms biological process (BP) all and Swiss-Prot (SP) Protein Information Resource (PIR) Keywords. The y-axis was the result of top functional clusters that E2F1 hits, derived by PRMT1 siRNA (siP1) versus PRMT5 (siP5) differential binding analysis, categorised in. The x-axis was the percentage of hits in individual functional cluster. (B) The table summarising the top functional clusters using DAVID, indicating the Enrichment Score (ES) and p-value (EASE score) both reflecting the statistical confidence of formation of each cluster. Also shown were examples of top genes in each cluster.

What happened next was the biochemical validation on these genes (Figure5.10). Chromatin immunoprecipitation of ChIP-seq grade was repeated with HeLa cells (Ch2.14-2). The genes that highlighted by differential binding analysis were derived from comparison between PRMT1 depletion to PRMT5 depletion on the level of E2F1 enrichment. Genes selected from the enhanced binding group were GATA-binding factor 1, Bloom syndrome, RecQ helicase-like (BLM), and Regulator of G-protein signalling family 3, which demonstrated an enhanced E2F1 binding when PRMT1 siRNA was transfected (Figure5.10B). In the reduced binding group, E2F1 enrichment was expected to be reduced upon PRMT1 depletion when compared to that of PRMT5 depletion which were validated in X-ray repair cross-complementing protein 5 (XRCC5), Nucleoporin 155kDa (NUP155) genes (Figure 5.10).

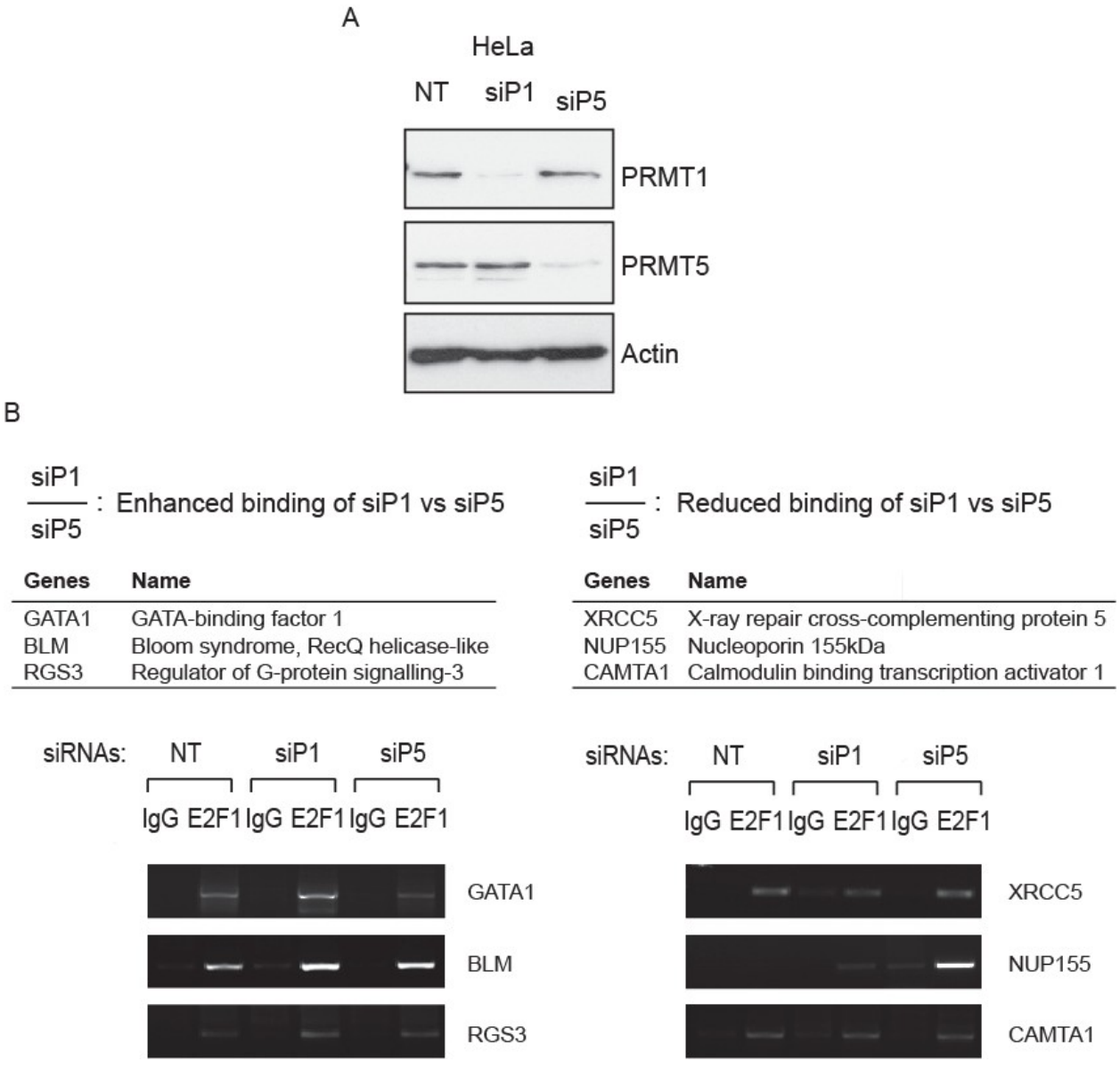


Figure5. 10 Biochemical validation of differentially bound E2F1ChIP-seq genes.
(A) Immunoblotting with PRMT1, PRMT5 and Actin antibodies on 72 hours 50nM non-targeting, PRMT1 and PRMT5 siRNAs treated HeLa cells. (B) Chromatin immunoprecipitated from NT, siP1, siP5 treated HeLa cells. The products were amplified by PCR with primers GATA1, BLM, RGS3 that identified as genes that E2F1 binding enhanced and primers XRCC5, NU155, CAMTA1 identified as genes that E2F1 binding reduced. The PCR products were loaded on 2% agarose gel containing 0.5μg/ml ethidium bromide (EtBr) and excited by UV for visualisation.

Chapter 6 Arginine methylation interacting protein partner

Ch6.1 Introduction

The methylation of arginine residues in proteins altered the protein-protein interaction landscape. The modification gave rise to the discovery of a group of methylarginine-binding proteins (best known examples, Tudor domain containing proteins) (Chen et al., 2011), which by recognising the methylarginine binding site, or by hindering other proteins sterically to neighbouring post-translationally modifications (PTM) sites, changed the protein-protein interaction and thus the physiological outcomes. The Writer-Reader (PRMTs ‘write’ the arginine methylation and the methylarginine binding proteins ‘read’) concept has therefore proposed in the understanding of arginine methylation or other PTM by large.

Ch6.2 E2F1 interacting-partners network

Inspired by the emerging concept, it is curious how E2F1 interacting-partners network responded to arginine methylation. The mass spectrometry technique has adopted for its power in recognising peptides, only the application shifted its focus from identifying methylation sites to the overall interacting protein-partners network by investigating E2F1 immunoprecipitation (IP) materials as a whole in the solution format for all its associated enzyme-digested peptides for this purpose. The control and inducible cells (inP1 and inP5) were previously exposed to doxycycline stimulation and therefore overexpressed accordingly (Figure6.1A). The transfections of mammalian vector and Flag tagged E2F1 plasmids have been confirmed by immunoblotting and then the electrophoresis in combination of silver staining confirmed the E2F1 IP (Figure6.1B). As demonstrated, the vector (C) served as controls to confirm its counterpart, the specific E2F1 IP (marked as arrow). The remaining 90% of the three E2F1 IP materials in solution, obtained from

Vector (pTRE inducible empty vector), induced PRMT1 and induced PRMT5 cells respectively, were then submitted to mass spectrometry.

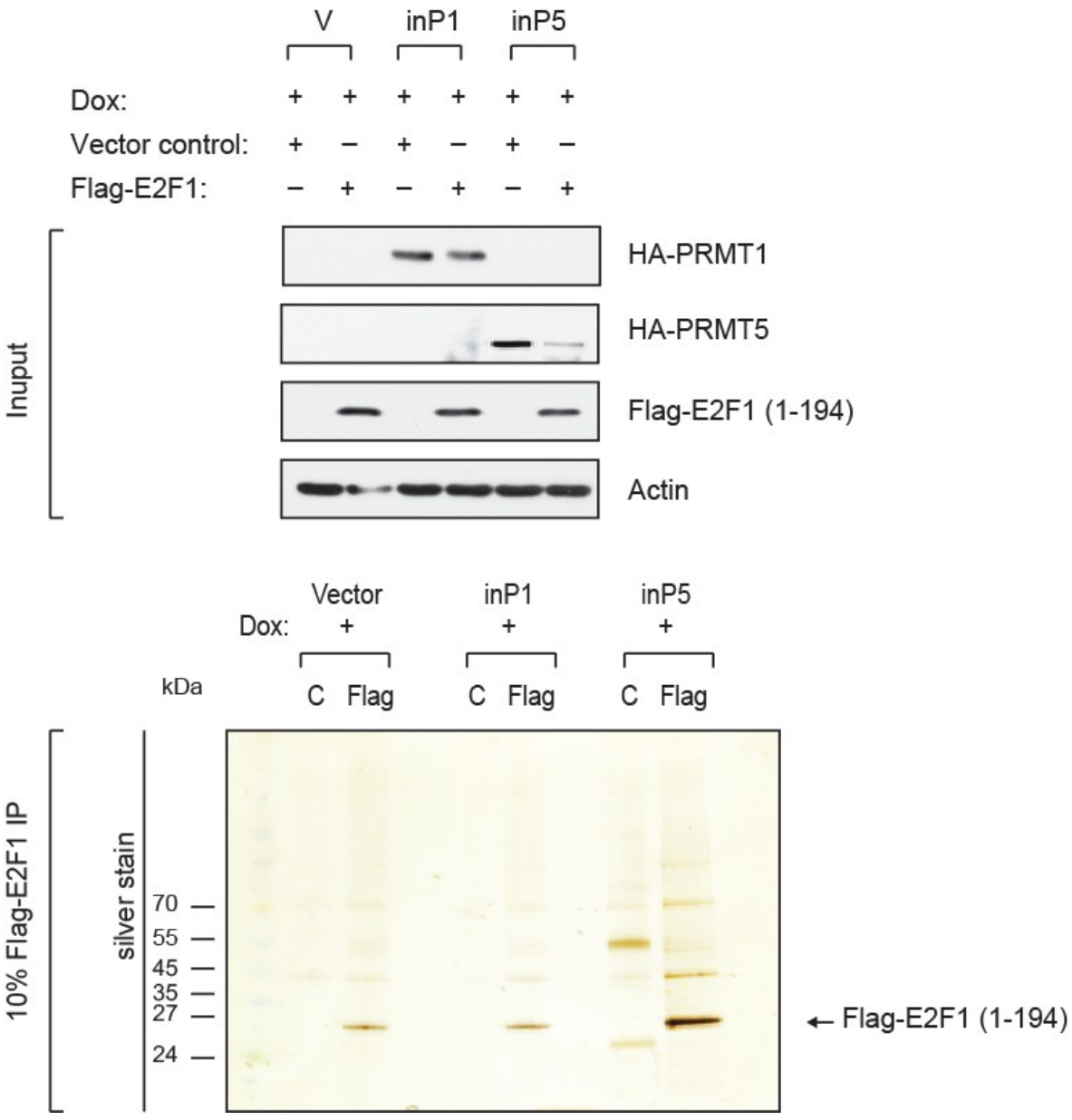


Figure6. 1 E2F1 interacting partner search by Mass Spectrometry.

Expression vector (pTre-hyg) control (V), inducible PRMT1 (inP1) and inducible PRMT5 (inP5) cells were transfected with 1μg mammalian vector (pcDNA3.1) control (C) or 1μg Flag E2F1 (1-194) 24 hours before a further 24 hours induction with1μg/ml doxycycline. The immunoblotting was used to confirm the induction of HA tagged PRMT1 and PRMT5 by HA antibody and Flag antibody for the Flag-E2F1 (1-194) transfection. Actin served as loading control. Flag conjugated agarose beads were used to immunoprecipitate Flag tagged E2F1 from inducible cells. The elutions were done with Flag peptide in 0.125M Tris solution competing binding to Flag agarose beads. 10% of elutions were loaded and separated by gel electrophoresis using NuPAGE® Novex® Gel system which later stained with silver staining. The remaining 90% elutions were subjected to digest in solution (in solution digestion) before submitted for mass spectrometry analysis.

Ch6.3 PRMT5 induced E2F1 and β-Catenin interaction

Following mass spectrometry analysis, one of the E2F1 interacting proteins appeared when PRMT5 was identified as β-catenin, a major player in the Wnt/ β-catenin signalling pathway. The activation of Wnt/ β-catenin pathway featured the accumulation of cytoplasmic β-catenin (Valenta et al., 2012), a consequential translocation from cytoplasm to nucleus where it modulated genes associated with cell growth, invasion or cell transformation, for examples c-MYC and cyclin D1(Clevers and Nusse, 2012) (Geyer et al., 2011).

The level of β-Catenin remained low by the destruction complex (comprised with Axin, adenomatosis polyposis coli (APC), protein phosphatase 2A (PP2A), glycogen synthase kinase 3 (GSK3) and casein kinase1α (CK1α)) targeting it for ubiquitination (Stamos and Weis, 2013). The mutation of component of the destruction complex, APC, led

to high-activity signalling both in genetic and non-genetic human diseases such as brain injury and colon cancer (Fancy et al., 2014) (Cancer Genome Atlas, 2012).

As effects of PRMT5 on E2F1 established to be favourable to cell growth (Cho et al., 2012), the involvement of β-Catenin, the center piece of a pathway known to be associated with cell proliferation and cell development in the fields of carcinogenesis and embryonic development (Nusse and Varmus, 2012) (Klaus and Birchmeier, 2008) was an interesting discovery.

To validate relation between PRMT5 presence and E2F1-β-Catenin interaction, HA-E2F1 were immunoprecipitated from cells with or without ectopically transfected Flag-PRMT5 for endogenous β-Catenin detection. These semi-endogenous E2F1-β-Catenin interaction experiments demonstrated a physical interaction between E2F1 and β-Catenin suggested by mass spectrometry analysis indeed existed in the presence of PRMT5 (Figure 6.2). Besides the confirmation of this PRMT5 driven interaction, the conditions not only reinforced SDMA signal with PRMT5 (Figure 6.2 PRMT5 versus SDMA)and correlated it in the setting but also demonstrated the competitive relationship between PRMT1 and PRMT5 on E2F1 as endogenous PRMT1 level decreased related to an increased level of PRMT5 (Figure 6.2 PRMT1 versus Flag-PRMT5).

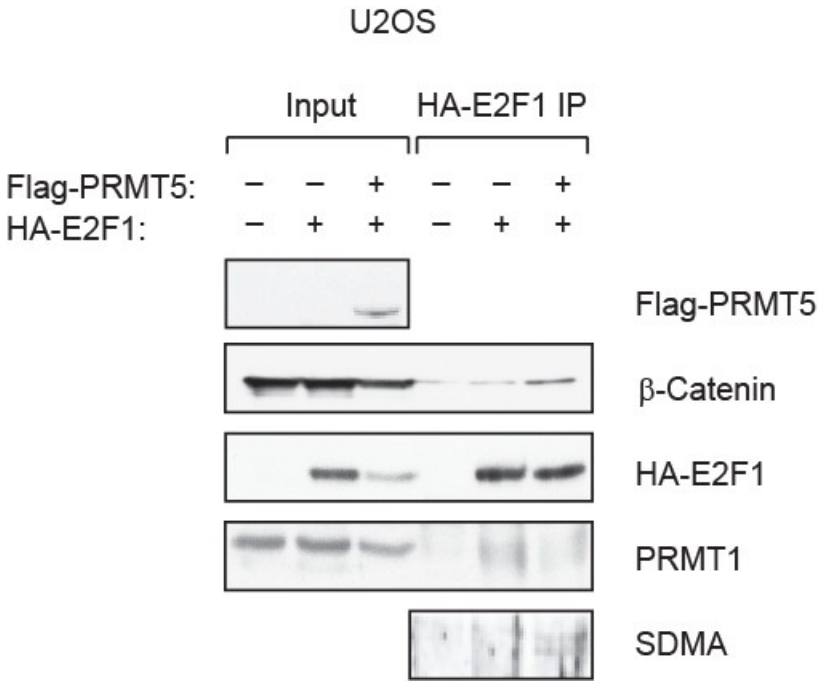


Figure6. 2 E2F1 and β-Catenin interaction upon presence of PRMT5.
U2OS cells were transfected with1μg HA tagged E2F1 with or without 1μg Flag tagged PRMT5. HA-E2F1 was immunoprecipitated and probed with β-Catenin, PRMT1 and SDMA. Data shown representative of three independent experiments.

Ch6.4 Arginine methylation on E2F1 and β-Catenin interaction

In order to further dissect details of arginine methylation and E2F1-β-Catenin interaction, E2F1 and its arginine methylation sites mutation derivatives were used in the experiments. The PRMT5 methylation sites mutation R111/113K (KK) and PRMT1 methylation site mutation R109K (K) were incorporated to compare with wild-type E2F1, the highlight of the experiments was that the E2F1-β-Catenin interaction level differed when different types of PRMT methylation sites mutated. Following removal of the PRMT1 methylation site on E2F1 by mutation (R109K), an increase level of the E2F1-β-Catenin interaction was detected compared to that of wild-type E2F1 (Figure 6.3 lane 3 versus lane 2). The substitution of lysine to arginine on position 111 and 113 of E2F1 deducted PRMT5 methylation sites from E2F1, and then the interaction with β-Catenin reduced compared to that of wild-type E2F1 (Figure 6.3 lane 4 versus lane 2)

The implications were encouraging for the arginine methylation impacts on E2F1 association with β-Catenin. Not only the PRMT5 sites mutation resulted in disassociation of E2F1 from β-Catenin reversely confirmed the concept of this PRMT5 dependent interaction, but also the mutation of PRMT1 methylation site, which by competitive nature between PRMT1 and PRMT5, strengthened the PRMT5 effect on E2F1 hence the elevated physical interaction between E2F1 and β-Catenin.

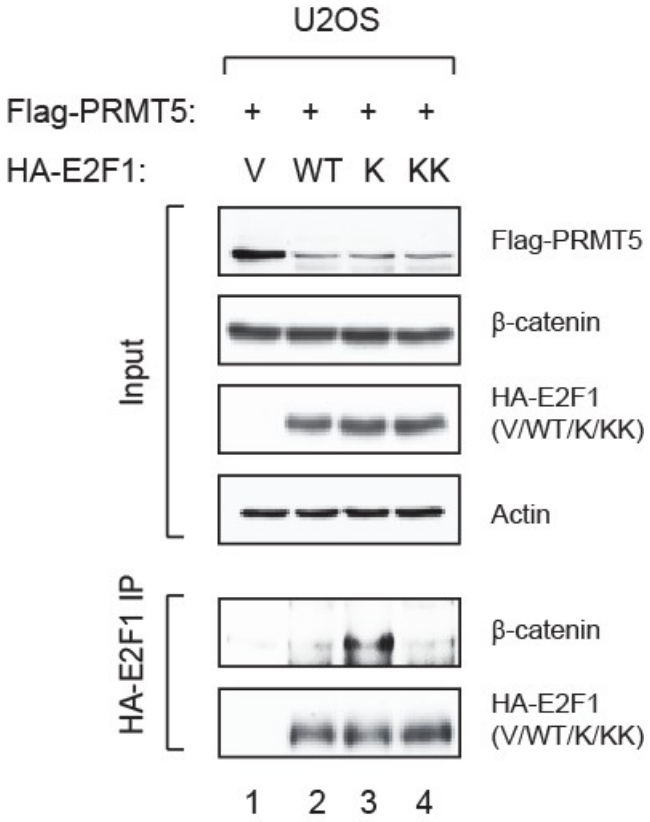


Figure6. 3 Arginine methylation sites and E2F1-β-Catenin interaction.
U2OS cells were transfected with 1μg Flag-PRMT5 and with either 1μg mammalian expression vector (V), or 1μg wild-type E2F1 (WT), or 1μg R109K E2F1 (K), or 1μg R111/113K E2F1(KK) for 48 hours. Flag, HA, β-Catenin, Action antibodies were used for immunoblotting on input proteins. HA-E2F1 was immunoprecipitated and probe with β-Catenin (N=3).

Chapter 7 Summary & Discussion

One of the most important questions in E2F1 biology is the mechanism underlying the opposing physiological outcomes that contribute to the E2F1 activity. As a master transcription factor that governs a spectrum of diverse genes that impact on a variety of cellular functions, many systems have found to influence E2F1 regulation.

Post-translational modification (PTM) is a way of providing proteins with an additional layer of complexity on top of their amino acids structure, resulting in extra ways of responding to external stimuli.

In this study, arginine methylation was proposed to be one, of a set of mechanisms dictating E2F1 activity between cell growth and cell death. It was hypothesised that the opposing choice of cell proliferation or apoptosis pathways was the result of E2F1 responding to different types of arginine methylation.

The rationale behind this argument was the formation of asymmetric or symmetric dimethyl arginines, catalysed by type I (i.e. PRMT1) or type II (i.e. PRMT5) PRMTs, has been identified in both histones and non-histone protein substrates. What is more interesting is these events have been shown to contribute to an opposing effect on common substrates.

PRMT5 induced methylation on histone H4 arginine 3 and H3 arginine 8 has previously been shown to trigger the transcriptional *silencing* of tumour suppressor genes (Pal et al., 2004) (Zhao et al., 2009). However, on the same substrate, the histone H4

arginine 3, PRMT1 mediated methylation has contributed to transcriptional *activation* by allowing by allowing subsequent acetylation on histones (Huang et al., 2005).

In non-histone protein substrates, PRMT5 mediated arginine methylation has previously been shown to regulate the p53 response and to affect the target gene specificity of p53. Deletion of PRMT5 triggers p53 dependent apoptosis (Jansson et al., 2008).

Unlike the resultant apoptosis inhibition effect on p53 via PRMT5, PRMT1 induced methylation on heterogeneous nuclear ribonucleoproteins K (hnRNP K) enhances its affinity with p53 in response to UV radiation. Suppression of methylation attenuates the recruitment of p53 to the p21 promoter and reduces p21 mRNA expression, which is traditionally associated with p53 dependent apoptosis (Chen et al., 2008).

The work presented in this thesis was conceived on the basis of previous confirmation of PRMT5 arginine methylation on E2F1 (Cho et al., 2012). It was found out that the symmetric dimethylation on arginine 111 and arginine 113 of E2F1 resulted in inhibition of the apoptotic function of E2F1, therefore implying an oncogenic feature of this event. It was proposed that the choice of E2F1's conflicting activities was dictated by different types of arginine methylation modifications.

As symmetric dimethylation on E2F1 demonstrated the apoptosis inhibitory effect, it was then interesting to investigate the potential effects of different types of arginine methylation on E2F1, such as the asymmetric dimethylation.

To confirm whether E2F1 can also be a substrate for type I methyltransferase, PRMT1, the enzyme and the substrate (PRMT1 and E2F1) were incubated together and the radioactive methyltransferase activity was detected *in vitro*, implying that E2F1 is a potential substrate for PRMT1. Later, two mass spectrometry analyses were employed to identify the methylation site using E2F1 immunoprecipitated from cells co-transfected with PRMT1 and E2F1. The two mass spectrometry analyses complimented on each other as the first analysis using wild type E2F1, narrowed down the methylation site at E2F1 peptide fragment containing arginine 109 and 111, and then the two methyl groups were still detected in the E2F1 peptide fragment containing arginine 109, only this time arginine 111 has been substituted by lysine using site directed mutagenesis at position 111. Together these proteomic approaches suggested PRMT1 was able to transfer dimethyl groups on to arginine 109 of E2F1. The third approach adopting R109K mutated E2F1 co-transfected with PRMT1 into human osteosarcoma cells and the detection of asymmetric dimethyl arginine mark (ADMA) is lost on the immunoprecipitated E2F1 comparing to the wild type. This suggested arginine 109 of E2F1 plays an important role in PRMT1 methylation on E2F1. These observations suggested the formation of asymmetric dimethylated arginines on E2F1 by PRMT1 which made E2F1 a common substrate of PRMT1 and PRMT5.

The investigation on physiological effect of PRMT1 methylation focused on the comparison between cell proliferation and cell death. The measurements included the apoptotic sub-G1 cell percentage analysis, apoptotic markers detection and cell colonies growth assay, all of which demonstrated a trend that PRMT1 was preferably associated with a growth inhibition effect or aided the apoptotic function of E2F1.

Following this, PRMT1 was tested for its ability to methylate E2F1 at arginine 109. The asymmetric methylation event demonstrated a growth inhibitory tendency that is opposing to what observed with PRMT5. What followed was to search for a possible explanation contributed to these observations.

When examined this transcription factor at the transcriptional level, the depletion of PRMT1 or PRMT5 demonstrated selectivity on the E2F1 promoter binding. Interestingly while PRMT1 was removed, level of the cell progression promoter Cdc6 bound by E2F1 increased and the level of apoptotic promoter Apaf1 bound by E2F1 decreased. This intriguing pattern was reversible by overexpression of PRMT1 and PRMT5 in the inducible cell system and validated again with double chromatin immunoprecipitation, which directly addressed the promoter specificity of asymmetric versus symmetric methylated E2F1 promoter binding.

The promoter specificity served as a possible explanation to convey post transcriptional modification signals to E2F1 and contributed to the corresponding effects

that originally seemed conflicting. A similar strategy was adopted by another transcription factor p53. The acetylation on the lysine at position 120 (K120) mediated by Tip60/hMOF was in the DNA binding domain where the majority of p53 mutation found in cancer. The acetylation at K120 is indispensable for the activation of target genes involved in apoptosis but not cell cycle arrest, which suggested a way to control targeted genes in a promoter specific fashion that influenced the cell fate (Sykes et al., 2006) (Tang et al., 2006).

In order to test the promoter specificity hypothesis further, next generation sequencing technology, and chromatin immunoprecipitation sequencing (ChIP-seq) was adopted with the aim of utilising massively parallel sequencing technique on all chromatin bound by E2F1, upon depletion of PRMT1 or PRMT5, on a genome-wide scale. The rationale is to observe whether the E2F1 sustained the promoter specificity at the genome-wide level.

The ChIP-seq results revealed E2F1 bound to 1,730 genes and the depletion of PRMT1 and PRMT5 recruited E2F1 on to three-fold and two- fold more genes.

Amongst the 20,000-25,000 protein coding genes sequenced internationally by the Human Genome Project (International Human Genome Sequencing, 2004), ChIP-seq revealed to be bound to 1,730 genes takes up 8.65% of 20,000 human genes. The figure was comparable with the prediction that E2F influenced 7% of human genes (Muller et al., 2001).

Also the ChIP-seq result was in line with evidence from the literature that shows that the preferences of these binding locations were at the core promoter regions or overlapped

with transcription start sites (Balciunaite et al., 2005). An average 40% of E2F1 ChIP were located proximally to the promoters which was comparable to a previous ChIP based study that 30% of the genes possessed an E2F1 binding site in the core promoters (Bieda et al., 2006).

The bioinformatics tools allowed a comparison between PRMT1 and PRMT5 depletion conditions and very intriguingly, one of the leading group of genes that was more affected by PRMT1 removal than PRMT5 removal was annotated to be the positive regulators for cell proliferation amongst other groups that were cell cycle related. The implication echoed a degree of promoter specificity yet slightly adjusted to the cell progression inhibition rather than the combination of growth inhibition and apoptosis promotion from PRMT1 methylation. In other words, PRMT1's impact on E2F1 was geared towards the repression of cell proliferation rather than induction of apoptosis according to E2F1 ChIP-seq data.

Numerous investigations could have followed on form the inaugural use of E2F1 ChIPs-seq data. For example, in the comparison between PRMT1 and PRMT5 depletion, although at the protein coding gene arena, ChIP regions identified influenced by PRMT1 was more evident, many E2F1 ChIP region identified upon PRMT5 depletion failed to be related to any genes by RefSeq definition. The implication was that the recruitment of E2F1 by PRMT5 had influences perhaps at the level not exclusively to protein coding genes. The

E2F1 regulations could be achieved by microRNAs for example, such as miR-449 and miR-34, which are induced by both E2F1 and p53 and shown to repress E2F1 at mRNA level that prevented excessive E2F1-induced proliferation (Yan et al., 2012).

Apart from the gene regulation by arginine methylation, the last part of the study began the investigation of the effects arginine methylation on the protein-protein interactions. By another proteomic mass spectrometry analysis, the main player of Wnt signalling pathway, the β-catenin was identified and confirmed to be interacting with E2F1 physically in the presence of PRMT5. The data suggested this interaction has demonstrated a degree of arginine methylation dependency as different methylation sites mutant prompted a difference in its binding to E2F1. There are many follow up experiments needed to be done to clarify the impact of this association, however the establishment of this association, the implied oncogenic PRMT5 signal with the Wnt/ β-catenin pathway, which traditionally linked as underlying mechanism for excessive cell growth behind many cancer was an interesting finding. The concept was comparable with a group of methylarginine-binding proteins (best known examples, Tudor domain containing proteins (Chen et al., 2011)) that influenced E2F1 regulation. A closer example is the p100-TSN tudor domain containing protein that reads symmetric arginine methylation mark on E2F1 and downregulated E2F1 apoptotic activity (Zheng et al., 2013).

The interaction of β-catenin with PRMT5 methylated E2F1 could be further investigated in two ways. These aspects may involve with the role of β-catenin in the Wnt signalling pathway. Firstly, β-catenin accumulates and then translocates from the cytoplasm to nucleus upon activation of the Wnt signalling pathway (Li et al., 2012). It might be interesting to understand whether the translocation ability of β-catenin plays a role in localising PRMT5 methylated E2F1 in the nucleus. There is evidence that the majority of PRMT5 is located in cytoplasm, E2F1 in the nucleus and PRMT5 methylated E2F1 in the nucleus (Cho et al., 2012). Secondly, it is also known that β-catenin enters the nucleus to engage with the DNA bound T cell factor/lymphoid enhancer factor (TCF/LEF) transcription factor family that activate Wnt target genes expression (Behrens et al., 1996) (Daniels and Weis, 2005). In colon cancer cells, the β-catenin/TCF-4 complex imposes a proliferative phenotype by upregulating the TCF target gene, c-Myc, which represses $p21^{CIP1/WAF1}$ transcription, mediating G1 arrest and differentiation (van de Wetering et al., 2002). Under hypoxic conditions, hypoxia inducible factor-1α (HIF-1α) competes directly with TCF-4 for the binding of β-catenin. Binding of HIF-1α to β-catenin enhances HIF-1 mediated transcription and promoting cell proliferation and adaptation to hypoxia (Kaidi et al., 2007). It would be intriguing to follow up and understand whether the PRMT5 induced β-catenin-E2F1interaction has any implications on the transcriptional level of the β-catenin associated transcriptional factors TCF, HIF1-α or other responsive Wnt genes.

These studies would further allow insights into the influence that arginine methylation imposes on protein-protein interaction networks and cross-talk between major transcription factor families in cancer biology.

References

AOYAGI, Y., ODA, T., KINOSHITA, T., NAKAHASHI, C., HASEBE, T., OHKOHCHI, N. & OCHIAI, A. 2004. Overexpression of TGF-beta by infiltrated granulocytes correlates with the expression of collagen mRNA in pancreatic cancer. *Br J Cancer,* 91**,** 1316-26.

BACHAND, F. 2007. Protein arginine methyltransferases: from unicellular eukaryotes to humans. *Eukaryot Cell*, 6**,** 889-98.

BADR, H., SMITH, C. B., GOLDSTEIN, N. E., GOMEZ, J. E. & REDD, W. H. 2014. Dyadic psychosocial intervention for advanced lung cancer patients and their family caregivers: Results of a randomized pilot trial. *Cancer*.

BALCIUNAITE, E., SPEKTOR, A., LENTS, N. H., CAM, H., TE RIELE, H., SCIME, A., RUDNICKI, M. A., YOUNG, R. & DYNLACHT, B. D. 2005. Pocket protein complexes are recruited to distinct targets in quiescent and proliferating cells. *Mol Cell Biol,* 25**,** 8166-78.

BATYCKA, M., INGLIS, N. F., COOK, K., ADAM, A., FRASER-PITT, D., SMITH, D. G., MAIN, L., LUBBEN, A. & KESSLER, B. M. 2006. Ultra-fast tandem mass spectrometry scanning combined with monolithic column liquid chromatography increases throughput in proteomic analysis. *Rapid Commun Mass Spectrom,* 20**,** 2074-80.

BEDFORD, M. T. 2007. Arginine methylation at a glance. *J Cell Sci*, 120**,** 4243-6.

BEDFORD, M. T. & CLARKE, S. G. 2009. Protein arginine methylation in mammals: who, what, and why. *Mol Cell*, 33**,** 1-13.

BEDFORD, M. T. & RICHARD, S. 2005. Arginine methylation an emerging regulator of protein function. *Mol Cell*, 18**,** 263-72.

BEHRENS, J., VON KRIES, J. P., KUHL, M., BRUHN, L., WEDLICH, D., GROSSCHEDL, R. & BIRCHMEIER, W. 1996. Functional interaction of beta-catenin with the transcription factor LEF-1. *Nature,* 382**,** 638-42.

BERTOLI, C., SKOTHEIM, J. M. & DE BRUIN, R. A. 2013. Control of cell cycle transcription during G1 and S phases. *Nat Rev Mol Cell Biol*, 14**,** 518-28.

BIEDA, M., XU, X., SINGER, M. A., GREEN, R. & FARNHAM, P. J. 2006. Unbiased location analysis of E2F1-binding sites suggests a widespread role for E2F1 in the human genome. *Genome Res*, 16**,** 595-605.

BUDHAVARAPU, V. N., WHITE, E. D., MAHANIC, C. S., CHEN, L., LIN, F. T. & LIN, W. C. 2012. Regulation of E2F1 by APC/C Cdh1 via K11 linkage-specific ubiquitin chain formation. *Cell Cycle,* 11**,** 2030-8.

BURKE, J. R., LIBAN, T. J., RESTREPO, T., LEE, H. W. & RUBIN, S. M. 2014. Multiple mechanisms for E2F binding inhibition by phosphorylation of the retinoblastoma protein C-terminal domain. *J Mol Biol*, 426**,** 245-55.

CANCER GENOME ATLAS, N. 2012. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, 487**,** 330-7.

CHAITANYA, G. V., STEVEN, A. J. & BABU, P. P. 2010. PARP-1 cleavage fragments: signatures

of cell-death proteases in neurodegeneration. *Cell Commun Signal*, 8**,** 31.

CHEN, C., NOTT, T. J., JIN, J. & PAWSON, T. 2011. Deciphering arginine methylation: Tudor tells the tale. *Nat Rev Mol Cell Biol*, 12**,** 629-42.

CHEN, Y., ZHOU, X., LIU, N., WANG, C., ZHANG, L., MO, W. & HU, G. 2008. Arginine methylation of hnRNP K enhances p53 transcriptional activity. *FEBS Lett*, 582**,** 1761-5.

CHI, P., ALLIS, C. D. & WANG, G. G. 2010. Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers. *Nat Rev Cancer*, 10**,** 457-69.

CHO, E. C., ZHENG, S., MUNRO, S., LIU, G., CARR, S. M., MOEHLENBRINK, J., LU, Y. C., STIMSON, L., KHAN, O., KONIETZNY, R., MCGOURAN, J., COUTTS, A. S., KESSLER, B., KERR, D. J. & THANGUE, N. B. 2012. Arginine methylation controls growth regulation by E2F-1. *EMBO J*, 31**,** 1785-97.

CIRRI, P. & CHIARUGI, P. 2011. Cancer associated fibroblasts: the dark side of the coin. *Am J Cancer Res*, 1**,** 482-97.

CLARKE, S. G. & TAMANOI, F. 2006. Protein methyltransferases. *The enzymes v. 24*. Amsterdam ; London: Academic Press.

CLEVERS, H. & NUSSE, R. 2012. Wnt/beta-catenin signaling and disease. *Cell*, 149**,** 1192-205.

CRESS, W. D. & NEVINS, J. R. 1994. Interacting domains of E2F1, DP1, and the adenovirus E4 protein. *J Virol*, 68**,** 4213-9.

DANIELS, D. L. & WEIS, W. I. 2005. Beta-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. *Nat Struct Mol Biol*, 12**,** 364-71.

DEGREGORI, J. & JOHNSON, D. G. 2006. Distinct and Overlapping Roles for E2F Family Members in Transcription, Proliferation and Apoptosis. *Curr Mol Med*, 6**,** 739-48.

DEGREGORI, J., LEONE, G., MIRON, A., JAKOI, L. & NEVINS, J. R. 1997. Distinct roles for E2F proteins in cell growth control and apoptosis. *Proc Natl Acad Sci U S A*, 94**,** 7245-50.

DHAR, S., VEMULAPALLI, V., PATANANAN, A. N., HUANG, G. L., DI LORENZO, A., RICHARD, S., COMB, M. J., GUO, A., CLARKE, S. G. & BEDFORD, M. T. 2013. Loss of the major Type I arginine methyltransferase PRMT1 causes substrate scavenging by other PRMTs. *Sci Rep*, 3**,** 1311.

DOLBERG, D. S., HOLLINGSWORTH, R., HERTLE, M. & BISSELL, M. J. 1985. Wounding and its role in RSV-mediated tumor formation. *Science*, 230**,** 676-8.

ERRICO, A. 2014. Cancer therapy: A SMaRT-approach to depression might improve cancer survival. *Nat Rev Clin Oncol*.

EVAN, G. I. & VOUSDEN, K. H. 2001. Proliferation, cell cycle and apoptosis in cancer. *Nature*, 411**,** 342-8.

FANCY, S. P., HARRINGTON, E. P., BARANZINI, S. E., SILBEREIS, J. C., SHIOW, L. R., YUEN, T. J.,

HUANG, E. J., LOMVARDAS, S. & ROWITCH, D. H. 2014. Parallel states of pathological Wnt signaling in neonatal brain injury and colon cancer. *Nat Neurosci*, 17, 506-12.

FENG, J., LIU, T., QIN, B., ZHANG, Y. & LIU, X. S. 2012. Identifying ChIP-seq enrichment using MACS. *Nat Protoc*, 7, 1728-40.

FIELD, S. J., TSAI, F. Y., KUO, F., ZUBIAGA, A. M., KAELIN, W. G., JR., LIVINGSTON, D. M., ORKIN, S. H. & GREENBERG, M. E. 1996. E2F-1 functions in mice to promote apoptosis and suppress proliferation. *Cell*, 85, 549-61.

FISCHER, R., TRUDGIAN, D. C., WRIGHT, C., THOMAS, G., BRADBURY, L. A., BROWN, M. A., BOWNESS, P. & KESSLER, B. M. 2012. Discovery of candidate serum proteomic and metabolomic biomarkers in ankylosing spondylitis. *Mol Cell Proteomics*, 11, M111 013904.

FOSTER, D. A., YELLEN, P., XU, L. & SAQCENA, M. 2010. Regulation of G1 Cell Cycle Progression: Distinguishing the Restriction Point from a Nutrient-Sensing Cell Growth Checkpoint(s). *Genes Cancer*, 1, 1124-31.

FRANKLIN, T. B. & MANSUY, I. M. 2010. Epigenetic inheritance in mammals: evidence for the impact of adverse environmental effects. *Neurobiol Dis*, 39, 61-5.

GARY, J. D. & CLARKE, S. 1998. RNA and protein interactions modulated by protein arginine methylation. *Prog Nucleic Acid Res Mol Biol*, 61, 65-131.

GENE ONTOLOGY, C. 2008. The Gene Ontology project in 2008. *Nucleic Acids Res*, 36, D440-4.

GEYER, F. C., LACROIX-TRIKI, M., SAVAGE, K., ARNEDOS, M., LAMBROS, M. B., MACKAY, A., NATRAJAN, R. & REIS-FILHO, J. S. 2011. beta-Catenin pathway activation in breast cancer is associated with triple-negative phenotype but not with CTNNB1 mutation. *Mod Pathol*, 24, 209-31.

GIACINTI, C. & GIORDANO, A. 2006. RB and cell cycle progression. *Oncogene*, 25, 5220-7.

GIRLING, R., PARTRIDGE, J. F., BANDARA, L. R., BURDEN, N., TOTTY, N. F., HSUAN, J. J. & LA THANGUE, N. B. 1993. A new component of the transcription factor DRTF1/E2F. *Nature*, 362, 83-7.

GRIVENNIKOV, S. I., GRETEN, F. R. & KARIN, M. 2010. Immunity, inflammation, and cancer. *Cell*, 140, 883-99.

HANAHAN, D. & WEINBERG, R. A. 2000. The hallmarks of cancer. *Cell*, 100, 57-70.

HANAHAN, D. & WEINBERG, R. A. 2011. Hallmarks of cancer: the next generation. *Cell*, 144, 646-74.

HELIN, K., LEES, J. A., VIDAL, M., DYSON, N., HARLOW, E. & FATTAEY, A. 1992. A cDNA encoding a pRB-binding protein with properties of the transcription factor E2F. *Cell*, 70, 337-50.

HELIN, K., WU, C. L., FATTAEY, A. R., LEES, J. A., DYNLACHT, B. D., NGWU, C. & HARLOW, E.

1993. Heterodimerization of the transcription factors E2F-1 and DP-1 leads to cooperative trans-activation. *Genes Dev*, 7**,** 1850-61.

HENLEY, S. A. & DICK, F. A. 2012. The retinoblastoma family of proteins and their regulatory functions in the mammalian cell division cycle. *Cell Div*, 7**,** 10.

HIRSCHI, A., CECCHINI, M., STEINHARDT, R. C., SCHAMBER, M. R., DICK, F. A. & RUBIN, S. M. 2010. An overlapping kinase and phosphatase docking site regulates activity of the retinoblastoma protein. *Nat Struct Mol Biol*, 17**,** 1051-7.

HUANG DA, W., SHERMAN, B. T. & LEMPICKI, R. A. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, 4**,** 44-57.

HUANG, S., LITT, M. & FELSENFELD, G. 2005. Methylation of histone H4 by arginine methyltransferase PRMT1 is essential in vivo for many subsequent histone modifications. *Genes Dev*, 19**,** 1885-93.

INTERNATIONAL HUMAN GENOME SEQUENCING, C. 2004. Finishing the euchromatic sequence of the human genome. *Nature*, 431**,** 931-45.

ISHIKAWA, K., ISHII, H. & SAITO, T. 2006. DNA damage-dependent cell cycle checkpoints and genomic stability. *DNA Cell Biol*, 25**,** 406-11.

IVANOVA, I. A., NAKRIEKO, K. A. & DAGNINO, L. 2009. Phosphorylation by p38 MAP kinase is required for E2F1 degradation and keratinocyte differentiation. *Oncogene*, 28**,** 52-62.

JABLONKA, E. & RAZ, G. 2009. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Q Rev Biol*, 84**,** 131-76.

JACKSON, P. K. & ELDRIDGE, A. G. 2002. The SCF ubiquitin ligase: an extended look. *Mol Cell*, 9**,** 923-5.

JANSSON, M., DURANT, S. T., CHO, E. C., SHEAHAN, S., EDELMANN, M., KESSLER, B. & LA THANGUE, N. B. 2008. Arginine methylation regulates the p53 response. *Nat Cell Biol*, 10**,** 1431-9.

JI, H., JIANG, H., MA, W., JOHNSON, D. S., MYERS, R. M. & WONG, W. H. 2008. An integrated software system for analyzing ChIP-chip and ChIP-seq data. *Nat Biotechnol*, 26**,** 1293-300.

JOHNSON, D. G., CRESS, W. D., JAKOI, L. & NEVINS, J. R. 1994. Oncogenic capacity of the E2F1 gene. *Proc Natl Acad Sci U S A*, 91**,** 12823-7.

KAIDI, A., WILLIAMS, A. C. & PARASKEVA, C. 2007. Interaction between beta-catenin and HIF-1 promotes cellular adaptation to hypoxia. *Nat Cell Biol*, 9**,** 210-7.

KALLURI, R. & ZEISBERG, M. 2006. Fibroblasts in cancer. *Nat Rev Cancer*, 6**,** 392-401.

KARKHANIS, V., HU, Y. J., BAIOCCHI, R. A., IMBALZANO, A. N. & SIF, S. 2011. Versatility of PRMT5-induced methylation in growth control and development. *Trends Biochem*

Sci, 36**,** 633-41.

KARP, G. & PATTON, J. G. 2013. *Cell and molecular biology : concepts and experiments*, Hoboken, NJ, John Wiley.

KHYRUL, W. A., LALONDE, D. P., BROWN, M. C., LEVINSON, H. & TURNER, C. E. 2004. The integrin-linked kinase regulates cell morphology and motility in a rho-associated kinase-dependent manner. *J Biol Chem*, 279**,** 54131-9.

KLAUS, A. & BIRCHMEIER, W. 2008. Wnt signalling and its impact on development and cancer. *Nat Rev Cancer*, 8**,** 387-98.

KNUDSON, A. G., JR. 1971. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A*, 68**,** 820-3.

KONTAKI, H. & TALIANIDIS, I. 2010. Lysine methylation regulates E2F1-induced cell death. *Mol Cell*, 39**,** 152-60.

KOVESDI, I., REICHEL, R. & NEVINS, J. R. 1986. Identification of a cellular transcription factor involved in E1A trans-activation. *Cell*, 45**,** 219-28.

KRAUSE, C. D., YANG, Z. H., KIM, Y. S., LEE, J. H., COOK, J. R. & PESTKA, S. 2007. Protein arginine methyltransferases: evolution and assessment of their pharmacological and therapeutic potential. *Pharmacol Ther*, 113**,** 50-87.

LA THANGUE, N. B. 2003. The yin and yang of E2F-1: balancing life and death. *Nat Cell Biol*, 5**,** 587-9.

LA THANGUE, N. B. & RIGBY, P. W. 1987. An adenovirus E1A-like transcription factor is regulated during the differentiation of murine embryonal carcinoma stem cells. *Cell*, 49**,** 507-13.

LANDER, E. S., LINTON, L. M., BIRREN, B., NUSBAUM, C., ZODY, M. C., BALDWIN, J., DEVON, K., DEWAR, K., DOYLE, M., FITZHUGH, W., FUNKE, R., GAGE, D., HARRIS, K., HEAFORD, A., HOWLAND, J., KANN, L., LEHOCZKY, J., LEVINE, R., MCEWAN, P., MCKERNAN, K., MELDRIM, J., MESIROV, J. P., MIRANDA, C., MORRIS, W., NAYLOR, J., RAYMOND, C., ROSETTI, M., SANTOS, R., SHERIDAN, A., SOUGNEZ, C., STANGE-THOMANN, N., STOJANOVIC, N., SUBRAMANIAN, A., WYMAN, D., ROGERS, J., SULSTON, J., AINSCOUGH, R., BECK, S., BENTLEY, D., BURTON, J., CLEE, C., CARTER, N., COULSON, A., DEADMAN, R., DELOUKAS, P., DUNHAM, A., DUNHAM, I., DURBIN, R., FRENCH, L., GRAFHAM, D., GREGORY, S., HUBBARD, T., HUMPHRAY, S., HUNT, A., JONES, M., LLOYD, C., MCMURRAY, A., MATTHEWS, L., MERCER, S., MILNE, S., MULLIKIN, J. C., MUNGALL, A., PLUMB, R., ROSS, M., SHOWNKEEN, R., SIMS, S., WATERSTON, R. H., WILSON, R. K., HILLIER, L. W., MCPHERSON, J. D., MARRA, M. A., MARDIS, E. R., FULTON, L. A., CHINWALLA, A. T., PEPIN, K. H., GISH, W. R., CHISSOE, S. L., WENDL, M. C., DELEHAUNTY, K. D., MINER, T. L., DELEHAUNTY, A., KRAMER, J. B., COOK, L. L., FULTON, R. S., JOHNSON, D. L., MINX, P. J., CLIFTON, S. W., HAWKINS, T., BRANSCOMB, E., PREDKI, P., RICHARDSON, P., WENNING, S.,

SLEZAK, T., DOGGETT, N., CHENG, J. F., OLSEN, A., LUCAS, S., ELKIN, C., UBERBACHER, E., FRAZIER, M., et al. 2001. Initial sequencing and analysis of the human genome. *Nature*, 409**,** 860-921.

LEE, J. & BEDFORD, M. T. 2002. PABP1 identified as an arginine methyltransferase substrate using high-density protein arrays. *EMBO Rep*, 3**,** 268-73.

LEES, J. A., SAITO, M., VIDAL, M., VALENTINE, M., LOOK, T., HARLOW, E., DYSON, N. & HELIN, K. 1993. The retinoblastoma protein binds to a family of E2F transcription factors. *Mol Cell Biol*, 13**,** 7813-25.

LI, V. S., NG, S. S., BOERSEMA, P. J., LOW, T. Y., KARTHAUS, W. R., GERLACH, J. P., MOHAMMED, S., HECK, A. J., MAURICE, M. M., MAHMOUDI, T. & CLEVERS, H. 2012. Wnt signaling through inhibition of beta-catenin degradation in an intact Axin1 complex. *Cell*, 149**,** 1245-56.

LIN, W. C., LIN, F. T. & NEVINS, J. R. 2001. Selective induction of E2F1 in response to DNA damage, mediated by ATM-dependent phosphorylation. *Genes Dev*, 15**,** 1833-44.

LODISH, H. F. 2007. *Molecular cell biology*, New York, Freeman.

LOFTUS, S. J., LIU, G., CARR, S. M., MUNRO, S. & LA THANGUE, N. B. 2012. NEDDylation regulates E2F-1-dependent transcription. *EMBO Rep*, 13**,** 811-8.

LUSCOMBE, N. M., LASKOWSKI, R. A. & THORNTON, J. M. 2001. Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res*, 29**,** 2860-74.

MAJUMDER, S., ALINARI, L., ROY, S., MILLER, T., DATTA, J., SIF, S., BAIOCCHI, R. & JACOB, S. T. 2010. Methylation of histone H3 and H4 by PRMT5 regulates ribosomal RNA gene transcription. *J Cell Biochem*, 109**,** 553-63.

MALUMBRES, M. & BARBACID, M. 2007. Cell cycle kinases in cancer. *Curr Opin Genet Dev*, 17**,** 60-5.

MARTINEZ-BALBAS, M. A., BAUER, U. M., NIELSEN, S. J., BREHM, A. & KOUZARIDES, T. 2000. Regulation of E2F1 activity by acetylation. *EMBO J*, 19**,** 662-71.

MATHIOUDAKI, K., PAPADOKOSTOPOULOU, A., SCORILAS, A., XYNOPOULOS, D., AGNANTI, N. & TALIERI, M. 2008. The PRMT1 gene expression pattern in colon cancer. *Br J Cancer*, 99**,** 2094-9.

MATHIOUDAKI, K., SCORILAS, A., ARDAVANIS, A., LYMBERI, P., TSIAMBAS, E., DEVETZI, M., APOSTOLAKI, A. & TALIERI, M. 2011. Clinical evaluation of PRMT1 gene expression in breast cancer. *Tumour Biol*, 32**,** 575-82.

MCILWAIN, D. R., BERGER, T. & MAK, T. W. 2013. Caspase functions in cell death and disease. *Cold Spring Harb Perspect Biol*, 5**,** a008656.

MENG, P. & GHOSH, R. 2014. Transcription addiction: can we garner the Yin and Yang functions of E2F1 for cancer therapy? *Cell Death Dis*, 5**,** e1360.

MENG, X., ZHONG, J., LIU, S., MURRAY, M. & GONZALEZ-ANGULO, A. M. 2012. A new

hypothesis for the cancer mechanism. *Cancer Metastasis Rev*, 31**,** 247-68.
MORGAN, D. O. 2007. *The cell cycle : principles of control*, London
Sunderland, MA, New Science Press ;
Sinauer Associates.
MULLER, H., BRACKEN, A. P., VERNELL, R., MORONI, M. C., CHRISTIANS, F., GRASSILLI, E., PROSPERINI, E., VIGO, E., OLINER, J. D. & HELIN, K. 2001. E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis. *Genes Dev*, 15**,** 267-85.
NACHMAN, M. W. & CROWELL, S. L. 2000. Estimate of the mutation rate per nucleotide in humans. *Genetics*, 156**,** 297-304.
NEVINS, J. R. 2001. The Rb/E2F pathway and cancer. *Hum Mol Genet*, 10**,** 699-703.
NUSSE, R. & VARMUS, H. 2012. Three decades of Wnts: a personal perspective on how a scientific field developed. *EMBO J*, 31**,** 2670-84.
OZAKI, T., OKOSHI, R., SANG, M., KUBO, N. & NAKAGAWARA, A. 2009. Acetylation status of E2F-1 has an important role in the regulation of E2F-1-mediated transactivation of tumor suppressor p73. *Biochem Biophys Res Commun*, 386**,** 207-11.
PAL, S., VISHWANATH, S. N., ERDJUMENT-BROMAGE, H., TEMPST, P. & SIF, S. 2004. Human SWI/SNF-associated PRMT5 methylates histone H3 arginine 8 and negatively regulates expression of ST7 and NM23 tumor suppressor genes. *Mol Cell Biol*, 24**,** 9630-45.
PODURI, A., EVRONY, G. D., CAI, X. & WALSH, C. A. 2013. Somatic mutation, genomic variation, and neurological disease. *Science*, 341**,** 1237758.
POLLACK, A. & CIANCIO, G. 1990. Cell cycle phase-specific analysis of cell viability using Hoechst 33342 and propidium iodide after ethanol preservation. *Methods Cell Biol*, 33**,** 19-24.
PRUITT, K. D., KATZ, K. S., SICOTTE, H. & MAGLOTT, D. R. 2000. Introducing RefSeq and LocusLink: curated human genome resources at the NCBI. *Trends Genet*, 16**,** 44-7.
PRUITT, K. D. & MAGLOTT, D. R. 2001. RefSeq and LocusLink: NCBI gene-centered resources. *Nucleic Acids Res*, 29**,** 137-40.
QIAO, H., DI STEFANO, L., TIAN, C., LI, Y. Y., YIN, Y. H., QIAN, X. P., PANG, X. W., LI, Y., MCNUTT, M. A., HELIN, K., ZHANG, Y. & CHEN, W. F. 2007. Human TFDP3, a novel DP protein, inhibits DNA binding and transactivation by E2F. *J Biol Chem*, 282**,** 454-66.
RABILLOUD, T., VUILLARD, L., GILLY, C. & LAWRENCE, J. J. 1994. Silver-staining of proteins in polyacrylamide gels: a general overview. *Cell Mol Biol (Noisy-le-grand)*, 40**,** 57-75.
RAFF, M. C. 1992. Social controls on cell survival and cell death. *Nature*, 356**,** 397-400.
REECE, J. B. & CAMPBELL, N. A. 2014. *Campbell biology / Jane B. Reece ... [and five others]*, Boston, Benjamin Cummings : imprint of Pearson.

RICCARDI, C. & NICOLETTI, I. 2006. Analysis of apoptosis by propidium iodide staining and flow cytometry. *Nat Protoc*, 1**,** 1458-61.

RUBIN, S. M., GALL, A. L., ZHENG, N. & PAVLETICH, N. P. 2005. Structure of the Rb C-terminal domain bound to E2F1-DP1: a mechanism for phosphorylation-induced E2F release. *Cell*, 123**,** 1093-106.

RUSHLOW, D., PIOVESAN, B., ZHANG, K., PRIGODA-LEE, N. L., MARCHONG, M. N., CLARK, R. D. & GALLIE, B. L. 2009. Detection of mosaic RB1 mutations in families with retinoblastoma. *Hum Mutat*, 30**,** 842-51.

SAHIN, F. & SLADEK, T. L. 2010a. E2F-1 binding affinity for pRb is not the only determinant of the E2F-1 activity. *Int J Biol Sci*, 6**,** 382-95.

SAHIN, F. & SLADEK, T. L. 2010b. E2F-1 has dual roles depending on the cell cycle. *Int J Biol Sci*, 6**,** 116-28.

SAYEGH, J., WEBB, K., CHENG, D., BEDFORD, M. T. & CLARKE, S. G. 2007. Regulation of protein arginine methyltransferase 8 (PRMT8) activity by its N-terminal domain. *J Biol Chem*, 282**,** 36444-53.

SCALLY, A. & DURBIN, R. 2012. Revising the human mutation rate: implications for understanding human evolution. *Nat Rev Genet*, 13**,** 745-53.

SELIGSON, D. B., HORVATH, S., SHI, T., YU, H., TZE, S., GRUNSTEIN, M. & KURDISTANI, S. K. 2005. Global histone modification patterns predict risk of prostate cancer recurrence. *Nature*, 435**,** 1262-6.

SHERR, C. J. 1996. Cancer cell cycles. *Science*, 274**,** 1672-7.

SHIN, H., LIU, T., MANRAI, A. K. & LIU, X. S. 2009. CEAS: cis-regulatory element annotation system. *Bioinformatics*, 25**,** 2605-6.

SIEWEKE, M. H., THOMPSON, N. L., SPORN, M. B. & BISSELL, M. J. 1990. Mediation of wound-related Rous sarcoma virus tumorigenesis by TGF-beta. *Science*, 248**,** 1656-60.

SINGH, B. N., ZHANG, G., HWA, Y. L., LI, J., DOWDY, S. C. & JIANG, S. W. 2010. Nonhistone protein acetylation as cancer therapy targets. *Expert Rev Anticancer Ther*, 10**,** 935-54.

STAMATOYANNOPOULOS, J. A., ADZHUBEI, I., THURMAN, R. E., KRYUKOV, G. V., MIRKIN, S. M. & SUNYAEV, S. R. 2009. Human mutation rate associated with DNA replication timing. *Nat Genet*, 41**,** 393-5.

STAMOS, J. L. & WEIS, W. I. 2013. The beta-catenin destruction complex. *Cold Spring Harb Perspect Biol*, 5**,** a007898.

STEVENS, C., SMITH, L. & LA THANGUE, N. B. 2003. Chk2 activates E2F-1 in response to DNA damage. *Nat Cell Biol*, 5**,** 401-9.

STRAHL, B. D. & ALLIS, C. D. 2000. The language of covalent histone modifications. *Nature*, 403**,** 41-5.

STRAHL, B. D., BRIGGS, S. D., BRAME, C. J., CALDWELL, J. A., KOH, S. S., MA, H., COOK, R. G., SHABANOWITZ, J., HUNT, D. F., STALLCUP, M. R. & ALLIS, C. D. 2001. Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Curr Biol*, 11**,** 996-1000.

SYKES, S. M., MELLERT, H. S., HOLBERT, M. A., LI, K., MARMORSTEIN, R., LANE, W. S. & MCMAHON, S. B. 2006. Acetylation of the p53 DNA-binding domain regulates apoptosis induction. *Mol Cell*, 24**,** 841-51.

TANG, J., GARY, J. D., CLARKE, S. & HERSCHMAN, H. R. 1998. PRMT 3, a type I protein arginine N-methyltransferase that differs from PRMT1 in its oligomerization, subcellular localization, substrate specificity, and regulation. *J Biol Chem*, 273**,** 16935-45.

TANG, Y., LUO, J., ZHANG, W. & GU, W. 2006. Tip60-dependent acetylation of p53 modulates the decision between cell-cycle arrest and apoptosis. *Mol Cell*, 24**,** 827-39.

TAYLOR, G. K. & GOODLETT, D. R. 2005. Rules governing protein identification by mass spectrometry. *Rapid Commun Mass Spectrom*, 19**,** 3420.

TEE, W. W., PARDO, M., THEUNISSEN, T. W., YU, L., CHOUDHARY, J. S., HAJKOVA, P. & SURANI, M. A. 2010. Prmt5 is essential for early mouse development and acts in the cytoplasm to maintain ES cell pluripotency. *Genes Dev*, 24**,** 2772-7.

TRIMARCHI, J. M. & LEES, J. A. 2002. Sibling rivalry in the E2F family. *Nat Rev Mol Cell Biol*, 3**,** 11-20.

TSAI, K. Y., HU, Y., MACLEOD, K. F., CROWLEY, D., YAMASAKI, L. & JACKS, T. 1998. Mutation of E2f-1 suppresses apoptosis and inappropriate S phase entry and extends survival of Rb-deficient mouse embryos. *Mol Cell*, 2**,** 293-304.

VALENTA, T., HAUSMANN, G. & BASLER, K. 2012. The many faces and functions of beta-catenin. *EMBO J*, 31**,** 2714-36.

VAN DE WETERING, M., SANCHO, E., VERWEIJ, C., DE LAU, W., OVING, I., HURLSTONE, A., VAN DER HORN, K., BATLLE, E., COUDREUSE, D., HARAMIS, A. P., TJON-PON-FONG, M., MOERER, P., VAN DEN BORN, M., SOETE, G., PALS, S., EILERS, M., MEDEMA, R. & CLEVERS, H. 2002. The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell*, 111**,** 241-50.

W, S. B. & P, W. C. 2014. World Cancer Report 2014. Geneva: World Health Organization.

WEI, H., MUNDADE, R., LANGE, K. C. & LU, T. 2014. Protein arginine methylation of non-histone proteins and its role in diseases. *Cell Cycle*, 13**,** 32-41.

WIGMORE, T. J., FARQUHAR-SMITH, P. & LAWSON, A. 2013. Intensive care for the cancer patient - unique clinical and ethical challenges and outcome prediction in the critically ill cancer patient. *Best Pract Res Clin Anaesthesiol*, 27**,** 527-43.

WILLIAMS, G. H. & STOEBER, K. 2012. The cell cycle and cancer. *J Pathol*, 226**,** 352-64.

WOLF, S. S. 2009. The protein arginine methyltransferase family: an update about function, new perspectives and the physiological role in humans. *Cell Mol Life Sci*, 66**,** 2109-21.

WU, C. H., YEH, L. S., HUANG, H., ARMINSKI, L., CASTRO-ALVEAR, J., CHEN, Y., HU, Z., KOURTESIS, P., LEDLEY, R. S., SUZEK, B. E., VINAYAKA, C. R., ZHANG, J. & BARKER, W. C. 2003. The Protein Information Resource. *Nucleic Acids Res*, 31**,** 345-7.

WU, L., TIMMERS, C., MAITI, B., SAAVEDRA, H. I., SANG, L., CHONG, G. T., NUCKOLLS, F., GIANGRANDE, P., WRIGHT, F. A., FIELD, S. J., GREENBERG, M. E., ORKIN, S., NEVINS, J. R., ROBINSON, M. L. & LEONE, G. 2001. The E2F1-3 transcription factors are essential for cellular proliferation. *Nature*, 414**,** 457-62.

WU, Z., ZHENG, S. & YU, Q. 2009. The E2F family and the role of E2F1 in apoptosis. *Int J Biochem Cell Biol*, 41**,** 2389-97.

XING, X. Z., GAO, Y., WANG, H. J., YANG, Q. H., HUANG, C. L., QU, S. N., ZHANG, H., WANG, H., XIAO, Q. L. & SUN, K. L. 2013. Risk factors and prognosis of critically ill cancer patients with postoperative acute respiratory insufficiency. *World J Emerg Med*, 4**,** 43-7.

YAMASAKI, L., BRONSON, R., WILLIAMS, B. O., DYSON, N. J., HARLOW, E. & JACKS, T. 1998. Loss of E2F-1 reduces tumorigenesis and extends the lifespan of Rb1(+/-)mice. *Nat Genet*, 18**,** 360-4.

YAN, F., LIU, H., HAO, J. & LIU, Z. 2012. Dynamical behaviors of Rb-E2F pathway including negative feedback loops involving miR449. *PLoS One*, 7**,** e43908.

YANG, Y. & BEDFORD, M. T. 2013. Protein arginine methyltransferases and cancer. *Nat Rev Cancer*, 13**,** 37-50.

YOSHIMATSU, M., TOYOKAWA, G., HAYAMI, S., UNOKI, M., TSUNODA, T., FIELD, H. I., KELLY, J. D., NEAL, D. E., MAEHARA, Y., PONDER, B. A., NAKAMURA, Y. & HAMAMOTO, R. 2011. Dysregulation of PRMT1 and PRMT6, Type I arginine methyltransferases, is involved in various types of human cancers. *Int J Cancer*, 128**,** 562-73.

YU, Z., CHEN, T., HEBERT, J., LI, E. & RICHARD, S. 2009. A mouse PRMT1 null allele defines an essential role for arginine methylation in genome maintenance and cell proliferation. *Mol Cell Biol*, 29**,** 2982-96.

ZALMAS, L. P., ZHAO, X., GRAHAM, A. L., FISHER, R., REILLY, C., COUTTS, A. S. & LA THANGUE, N. B. 2008. DNA-damage response control of E2F7 and E2F8. *EMBO Rep*, 9**,** 252-9.

ZHANG, X. & CHENG, X. 2003. Structure of the predominant protein arginine methyltransferase PRMT1 and analysis of its binding to substrate peptides. *Structure*, 11**,** 509-20.

ZHANG, Y. & CHELLAPPAN, S. P. 1995. Cloning and characterization of human DP2, a novel dimerization partner of E2F. *Oncogene*, 10**,** 2085-93.

ZHANG, Y., LIU, T., MEYER, C. A., EECKHOUTE, J., JOHNSON, D. S., BERNSTEIN, B. E., NUSBAUM, C., MYERS, R. M., BROWN, M., LI, W. & LIU, X. S. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*, 9**,** R137.

ZHAO, Q., RANK, G., TAN, Y. T., LI, H., MORITZ, R. L., SIMPSON, R. J., CERRUTI, L., CURTIS, D. J., PATEL, D. J., ALLIS, C. D., CUNNINGHAM, J. M. & JANE, S. M. 2009. PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing. *Nat Struct Mol Biol*, 16**,** 304-11.

ZHENG, S., MOEHLENBRINK, J., LU, Y. C., ZALMAS, L. P., SAGUM, C. A., CARR, S., MCGOURAN, J. F., ALEXANDER, L., FEDOROV, O., MUNRO, S., KESSLER, B., BEDFORD, M. T., YU, Q. & LA THANGUE, N. B. 2013. Arginine methylation-dependent reader-writer interplay governs growth control by E2F-1. *Mol Cell*, 52**,** 37-51.