Histone H3.3-G34W in Giant Cell Tumor of Bone drive DNA hypomethylation through interaction with PWWP of DNMT3A

A Thesis Submitted to
the Department of Cancer Biomedical Science
in Partial Fulfillment of the Requirements
for the Master's Degree of Science

DAO HONG NHUNG

July 2021

National Cancer Center
Graduate School of Cancer Science and Policy

Professor Anders Lindroth

We hereby approve the M.S. thesis of
Dao Hong Nhung.

July 2021

Eun Jung Park **sign**
Chairman of Thesis Committee

Byeong Il Lee **sign**
Thesis Committee Member

Anders Lindroth **sign**
Thesis Committee Member

NATIONAL CANCER CENTER
GRADUATE SCHOOL OF CANCER SCIENCE AND POLICY

ABSTRACT

Histone H3.3-G34W in Giant Cell Tumor of Bone drive DNA hypomethylation through interaction with PWWP of DNMT3A

As vital components of chromatin configuration, histones and their variants are involved in several epigenetic activities contributing to genome stability and gene expression. The histone H3.3 variants have been reported to be frequently mutated mainly in brain and bone cancers. These mutations occur in the N-terminal tail of histone H3.3 subject to epigenetic modifications and a binding platform for many proteins regulating chromatin and gene expression. One prominent example of H3.3 mutations leads to G34 substitution to tryptophan, denoted H3.3-G34W, which occur in more than 90% of cases of giant cell tumor of bone. It was recently observed that H3.3-G34W stromal cells undergo distinct loss of DNA methylation. In this thesis I have addressed the potential link between H3.3-G34W and hypomethylation.

With surface plasmon resonance methodology and protein-docking prediction experiments, I found that the binding properties of H3K36me3 and the PWWP-domain of the DNA methyltransferase is similar to the properties of the interaction between PWWP and H3.3-G34W, which is predicted to be hydrophobic in nature.

To test if the interaction to DNMT3A would affect the DNA methyltransferase (DNMT) activity, I found that isogenic cell lysates with H3.3-G34W reduce DNMT activity with 34%. Similar *in vitro* experiments were performed with purified recombinant DNMT3A enzyme mixed with peptides of H3K36me3, H3.3-G34W showed a distinct reduced activity in the latter compared to the control H3.3-WT.

We propose that H3.3-G34W reduces H3K36me3 at the same molecule and replaces binding for PWWP of DNMT3A. This leads to reduced methyltransferases activity, in keeping with the observed hypomethylation in giant cell tumor of bone, suggesting an active role in epigenetic changes.

Table of contents

List of Tables

List of Figures

1. Introduction

1.1 DNA methylation and histone methylation

DNA methylation is a well-described epigenetic modification associated with gene expression control, including gene silencing, with a critical function in mammalian embryogenesis (1). The mechanism is carried out by the DNA methyltransferase (DNMT) family of proteins. Among its members, DNA methyltransferase 3 alpha (DNMT3A) and DNA methyltransferase 3 beta (DNMT3B) are responsible for *de novo* DNA methylation in early embryo development, closely correlated with differentiation (2). They catalyze the transfer of a methyl group from S-adenosyl methionine (SAM) to the fifth carbon (C5) of cytosine in the CpG context, forming 5-methylcytosine. The mechanism by which DNMT3 proteins are recruited is not clear, but certain histone modification, in particular trimethylation of histone H3 at lysine 36 (H3K36me3) is closely correlated (3).

Histones undergo numerous post-translational modifications (PTMs) and are involved in many genome regulation processes (4). The most common modifications are methylation of arginine or lysine or the acetylation of lysine (5). Among four core histone proteins which form nucleosome with 147 base pairs of DNA (H2A, H2B, H3 and H4), histone H3 is the most extensively modified. The site as well as the degree (mono-, di- or tri-) of methylation contribute to significantly different gene expression (6, 7). For example, H3K4me3 and H3K36me2/3, are hallmarks of active transcription, while

H3K27me3 and H4K20me3 are silenced chromatin marks and involved in chromatin repression (7, 8). H3K36 trimethylation, which is induced by SETD2 (SET-domain-containing 2), is associated with active chromatin as well as transcriptional repression, alternative splicing and other DNA process in eukaryotes during development (9-12). Histone H3.3, a variant of histone H3, is encoded by the H3F3A and H3F3B genes and has been shown to participate in gene activation, silencing and chromosome segregation (4, 13). PTMs direct how reader proteins interact with nucleosomes (14). It is not surprising that epigenetic dysregulation caused by mutations in H3.3 might have a dramatic impact on gene expression and chromatin integrity. Gain-of-function mutations in histones lead to abnormal gene expression and change the chromatin towards cancer formation. The H3.3-K27M mutation drastically alter the epigenetic landscape by reduction of histone mark H3K27me3 and DNA methylation via blocking the function of Polysomy repressive complex 2 (15). Hence, both K27M and G34R/V affect PTMs and drive pediatric glioblastoma, yet mechanisms of the latter is less well understood (16). While the effect of H3.3-K27M occurs in the nucleoplasm and less in the chromatin, H3.3-G34R act in the context of the nucleosome and the chromatin, strongly influencing expression of genes where it appears in the nucleosome.

1.2 Mechanisms of DNA methylation involve histone modification binding

DNMT3A and DNMT3B consist of a C-terminal catalytic domain (methyltransferase – MTase) and a N-terminal region with regulatory function that include the ADD (found in ATRX, DNMT3, and DNMT3L) and the PWWP domain. While ADD interacts with DNA (17), PWWP domain reads the proper position in CpG island through trimethylated H3K36 (3). A loss of recognition of H3 tail containing K36me3 and a decrease in DNA methylation were found as a result of a point mutations in DNMT3B-PWWP in the ICF syndrome (18). PWWP (Proline-Tryptophan-Tryptophan-Proline), like other members of the Royal super family domains including chromodomain, MBT and Tudor, participates in chromatin targeting by identifying histone tail modification, mostly the methylation or acetylation at Lysine positions (19). Structurally, all PWWP family members share similar fold consisting of β−barrels and helical substructures. Typically, there are five N-terminal antiparallel β-strands of the PWWP domain, and more variable sequences in the C-terminal helical part of the domain with two to six α-helices (20). As shown in Figure 1, PWWP of DNMT3A/B is replaced with SWWP in the β2 barrel and has three conserved aromatic residues forming a putative methyl-lysine aromatic cage, essential in the interaction with H3K3me3 (21). The cage constitutes F303/206, the third residue of the PWWP motif W306/239 in the beginning of β2 and W330/263 in the end of β3 strand in DNMT3A and DNMT3B (21). Similar methylated–

histone binding features of PWWP with trimethylated H3 at Lysine 36 was also found in other PWWP proteins by multiple sequencing and crystallography studies such as BRPF1, ZMYND11, PSIP1, PWWP2B (22).

Figure 1. Sequence alignment with secondary structure between DNMT3A, DNMT3B and PWWP2B. Blue dots: aromatic cage residues. Both DNMT3A/B and PWWP2B have the same three highly conserved amino acids central to the aromatic cage of PWWP domain, suggesting shared ability for interaction.

1.3 DNA methylation is commonly altered in cancer

In normal cells, the two opposing processes of methylation and demethylation are tightly regulated in development. However, in cancer this balance is wildly disrupted and has been closely correlated with cancer development (23). Typically, stringent hypermethylation of promoter CpG islands is commonly observed. In contrast, less severe global hypomethylation is also frequent and may lead to chromosomal instability and loss of imprinting. This bimodal deregulation of the epigenetic landscape is found across virtually all human cancers (24). DNA hypomethylation occur in virtually all tumors, e.g. cervical (25) and breast cancers (26), while hypermethylation varies widely. In this study, I have investigated the hypomethylation in giant cell tumor of bone.

1.4 Giant Cell Tumor of Bone is driven by H3.3-G34W

Giant cell tumor of bone (GCTB) is a rare, aggressive benign tumor, characterized by the presence of multinucleated giant cells (osteoclast-like cells) and stromal cells. It is more common in young adults between ages 20 and 40 after skeletal maturity and mostly located in the long bones, in particular the distal femur area. It is not known why GCTB only affects younger individuals but RANKL signaling is strongly correlated (27). The most common symptom is intensive pain in the area of the tumor due bone overgrowth (requiring frequent curettage) but also osteoporosis leading to fractures (28).

Behjati, S. and colleagues found that glycine-to-tryptophan substitution at position 34 of the H3.3 gene H3F3A (29) accounted for the highest mutation rate (92%) in GCTB stromal cells, but also arginine (G34R) and valine (G34V) substitutions were observed (30). This result was consisted with a German and a South Korean cohort showing more than 90% of H3.3-G34W (31). With a complex gene expression pattern and higher proliferative and infiltrative activity in GCTB, H3.3-G34W has been proposed as the likely driver of tumorigenesis (31). A strong relationship between G34-alteration and its nearby modifications has also been observed. For instance, G34R/V mutations have been found to cause a significant loss of H3.3K36me3 and H3.3K27 methylation *in cis* (29), suggesting that these mutations may differentially influence gene transcription where incorporated.

1.5 G34W substitution of H3.3 affects epigenetic modifications

Located in close proximity to lysine 36, G34W has been shown to alter histone H3K36 methylation and affect the binding of readers specifically interacting with K36 methylation (29, 32-34). A comprehensive epigenetic analysis was recently published by Lindroth, Plass and colleagues in GCTB-derived stromal cells, with and without H3.3-G34W. In that study, mutant cells showed an increase in chromatin accessibility along with a 20% decrease in DNA methylation, as seen in Figure 2A (35). Subsequent genome-wide analysis found that subtelomeric areas, displayed in Figure 2B, were hypomethylated in GCTB (Lindroth et. al., personal communication). The underlying mechanism(s) behind DNA hypomethylation is widely debated. Loss of DNA methylation may be linked to H3K36me3 through interaction with the PWWP domain of DNMT3s. Due to its proximal location, H3.3-G34W may be involved in reducing DNA methylation via PWWP. Previous experiments in the Lindroth lab showed indications of binding between G34W and DNMT3A-PWWP, similar to binding of H3K36me3 by the PWWP domain (Lindroth et. al., personal communication). This suggests the possibility that loss of H3.3K36me3 may instead accept PWWP binding to H3.3-G34W, effectively contributing to the distinct DNA hypomethylation in GCTB. In this study, I have focused on this promising mechanism behind H3.3-G34W and DNA hypomethylation.

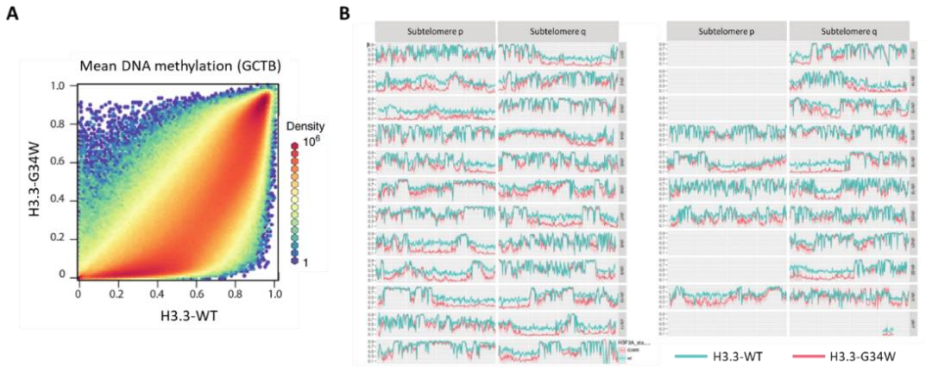


Figure 2. Loss of DNA methylation in GCTB.

A. A scatter plot with the distribution of β-values when comparing mean DNA methylation between H3.3-WT (x-axis) and H3.3-G34W (y-axis) of GCTB stromal cells (35). The color represents the density gradient from 1 (blue) to 10^6 (red)**.** The higher hexagon population with warmer colors closed to x-axis reveals the decrease in DNA methylation in the mutant cells.

B. Telomeric region with the illustration of the general DNA methylation in every chromosome (Lindroth et. al., personal communication). In general, the DNA methylation in H3.3-G34W (red line) is lower than it is in wildtype H3.3 (green line), indicating the reduced DNA methylation in the latter sample.

2. Materials and Methods

2.1 Cell lines and proliferation assays in isogenic cells

The human cervical cancer cell line HeLa and the osteosarcoma cell MG-63 were obtained from the Korean Cell Line Bank. The two cell lines were transfected with constructs carrying H3.3 mutations producing isogenic cell lines with H3.3 wild-type (isoH3.3-WT) and a change of amino acid residues glycine 34 to tryptophan (isoH3.3-G34W) targeting the endogenous H3F3A locus. The SETD2 gene, encoding histone H3 trimethyl transferase, was knocked out by means of CRISPR/Cas9 methodology using both isogenic cell lines to identify the impact on cell growth. Two independent clones of each isogenic cell line were isolated and used in the assay. The culture conditions were MEM media (Welgene) for HeLa or DMEM media (Welgene) for MG-63, 10% fetal bovine serum (FBS, Corning) and 5 μg/ml gentamicin (Gibco). The cells were grown and maintained under conditions of 80% humidity and 5% CO_2 at 37°C and passaged every 3 days when confluence reached 80–90%.

Proliferation assay was conducted by seeding 20,000 isoHeLa cells in each well of 12-well plates and manually counting the cells for five days continuously. Here, the Counting Chamber was utilized after the cells were stained with 0.4% trypan blue (Gibco) to determine dead cells in the counting. The cell number in each day was done in triplicates. All results and statistical analysis were performed using GraphPad Prism 6 with unpaired t-test and presented as mean ± SD.

2.2 Migration assays in isogenic MG-63 cells with H3.3 mutations

Infiltration assays estimating migration of cells through a gel matrix was according to the Boyden assay protocol (36). Isogenic MG-63 with H3.3-WT, -K27M, -G34R and -G34W developed in the Lindroth lab (31) was used. Each chamber cup (Transwell Permeable Supports, Corning) was prepacked with 100 µl extracellular matrix (ECM) diluted 1:8 in ice cold serum-free medium (SFM) and transferred to a well with 700 µl SFM. After overnight incubation in cell growth conditions, the cups were rinsed with SFM and added 35,000 cells/100 µl SFM. The packed cups were then placed in 700 µl of SFM with 10% FBS. Only cells able to degrade the gel layer could migrate through the microporous membrane into the lower compartment, where nutrients were present (without antibiotic). After 15 hours incubation, the cups were taken out and stained in Diff Quik kit (green/red/blue, Sysmex). The cups were cleaned with a cotton swab and water to remove cells which did not migrate across the membrane, after which the membrane was cut out and mounted on slide with VectaMount Permanent Mounting Medium (Vector Laboratories). The infiltration rate was estimated by manual counting the cells that migrated across the membrane.

2.3 Interaction analysis between the PWWP domain of DNMT3A and H3.3-G34W

2.3.1. H3.3 peptide production and purification

Peptides corresponding to 20 amino acids (residue 21-40) of H3.3 N-terminal tail were synthesized with 95% purity (GenScript, USA). The amino acid sequences are as follows:

H3.3K36me3: ATKAARKSAPSTGGV**K(me3)**KPHR

H3.3-WT: ATKAARKSAPSTG**G**VKKPHR

H3.3-G34W: ATKAARKSAPSTG**W**VKKPHR

All peptides were dissolved in water but reconstituted in PBS for the SPR analysis in 2.3.3.

2.3.2 Purification of PWWP protein from DNMT3A

The PWWP construct was obtained from Addgene (pGEX-GP-2-DNMT3A-PWWP, plasmid #59696) and transformed into chemically competent *E. coli* BL21-DE3 strain. The plasmid vector pGEX-6P-2 containing DNMT3A-PWWP generates a fusion protein with GST. BL21-DE3 was cultured in LB media with 100 μg/mL ampicillin at 37 °C and 150 rpm rotation. When OD_{600} reached 0.4 - 0.6, translation of pGEX-GP-2-DNMT3A-PWWP was induced by 0.2 mM IPTG at 20 °C overnight and 150 rpm rotation. The bacterial cells were pelleted by centrifugation. The pellet was resuspended in cold buffer containing 20 mM Tris pH 7.5, 500 mM NaCl, 1% NP-40, 5% deoxycholate, with protease inhibitor

cocktail (cOmplete Mini, EDTA-free, 1 tablet for 10 mL of solution, Roche) and 0.1 mM PMSF by using sonication in cold condition for 15 minutes. The GST-tagged DNMT3A-PWWP was purified by Glutathione Spin Column (ThermoFisher). The GST tag was then cleaved from the fusion protein by HRV 3C protease (Takara) overnight. Digestion of the fusion protein with HRV 3C protease (22 kDa) resulted in a 24 kDa GST peptide and a 16.2 kDa PWWP protein. Since the protease has a N-terminal 6xHN tag, which is a modified His-tag, it can be easily eliminated from the protease reaction solution through HisPur Ni-NTA Spin Column (ThermoFisher). Using size-exclusion chromatography (SEC) method, the domain was finally polished over a Hiload 16/600 Superdex 75 pg column in PBS buffer, followed by concentrated by concentrator (Amicon Ultra-15 Centrifugal Filter Unit – 3kDa cutoff). All the expression and purification stages were checked by SDS-PAGE and Coomassie assay.

2.3.3 Surface plasmon resonance analysis of H3.3-G34W and PWWP interaction

Surface Plasmon Resonance (SPR) is an optical biosensing technique which provides affinity and kinetic information of molecular interaction. The purified DNMT3A-PWWP protein described in 2.3.2 and the synthesized H3.3 peptides described in 2.3.1 were tested for interaction with SPR. The PWWP-protein as the ligand was immobilized on the surface of the sensor chip using well-defined

chemistry. This allowed an analyte, in this case the peptides, to flow over the ligand at various concentrations to measure the interactions between the ligand/analyte pair. The increase in mass associated with a binding event caused a proportional increase in the refractive index at the surface of the gold sensor chip. These changes were measured as changes in the resonance angle of refracted light when the analyte (peptide) flows through and binds to the immobilized ligand and increases in density at the sensor chip. Here, SPR was performed by Reichert SR7500DC system and Scrubber2 software with CMDH Chip. Binding assay was applied with PBS, pH 7.4 as flow buffer and NaOH 10 mM as regeneration buffer. Sample preparations prior to analysis are described in Table 1.

Table 1. Sample preparation for SPR experiment

Type	Sample	Stock	Solvent	M.W.
Ligand 1	DNMT3A-PWWP	1.332 mg/mL	PBS	16.2 kDa
Analyte 1	H3.3K36me3	3.885 mM	PBS	2316.65 Da
Analyte 2	H3.3-WT	3.956 mM	PBS	2274.66 Da
Analyte 3	H3.3-G34W	3.744 mM	PBS	2403.83 Da

The ligand immobilization on sample channel involved 3 steps: activation, immobilization, and deactivation. The ligand stock (PWWP domain) was diluted to 10 µg/mL with sodium acetate 10 mM, pH 5.0 and allowed to flow at 20 µL/min. For each of 3 analytes, testing concentration included 6.25, 12.5, 25, 50, 100 and 200 µM; the flow rate was 30 µL/min, and association time and

dissociation time were 3 and 4 minutes, respectively. The strength of two molecule interaction was characterized by the equilibrium dissociation constant K_D, which was based on the formula: K_D = [P] [L] / [PL], where [P] is concentration of analyte, [L] is concentration of ligand and [PL] is concentration of the complex. At equilibrium, K_D is calculated by the relationship between the association rate k_a (the rate of complex formation) and the dissociation rate k_d (the rate of breakdown), which is based on this formular: K_D [M] = k_d [s^{-1}] / k_a [$M^{-1}.s^{-1}$]. A strong affinity binding is shown as small K_D value, as well as large k_a value and small k_d value. Half-life is the time to dissociate to half the starting value and depends only on the dissociation constant: $t_{1/2}$ [s] = ln2 / k_d. The data was graphed out using Prism 6 (GraphPad Software).

2.4 DNA methyltransferase activity assays

2.4.1 Total DNA methyltransferase activity assay

DNA methyltransferase (DNMT) activity assays were designed to measure and compare the total DNMT activity between wildtype (WT) and mutated samples. The purpose was also to determine whether the interaction between G34W and PWWP affects the global DNMT activity in the tumor cells.

Isogenic HeLa cells with H3.3-WT or H3.3-G34W described in 2.1 were harvested with 0.25% trypsin-EDTA, pelleted by centrifugation and lysed by RIPA buffer (20 mM Tris pH 7.5, 500 mM NaCl, 1% NP-40, 5% deoxycholate)

with protease prohibitor (10x cOmplete protease inhibitor EDTA-free) on ice for 1.5 hour. After high-speed centrifugation, the supernatant was collected as whole cell lysate, followed by measuring the total protein content using the Pierce BCA protein assay (ThermoFisher). The lysate contained the DNMT enzymes present in the cells to be measured by the assay: DNMT3A, DNMT3B, DNMT1. The colorimetric EpiQuik DNMT Activity/Inhibition Assay Ultra Kit (EpiGentek) was applied for two individual clones of isoH3.3-WT and two clones of isoH3.3-G34W using 8 µg of total protein in each reaction. The kit is designed with the DNA substrate coated and crosslinked into the microplate wells. According to the kit instructions, each well was supplemented with Adomet and DNMT buffer. Following the incubation at 37 °C for 1.6 hours, the DNMT enzymes in the cell lysates catalyze the transfer of methyl groups from Adomet to cytosine of the DNA substrate fixed to the plate. A 5–methylcytosine antibody recognizing methylated DNA was used to estimate activity where the amount of signal from the antibody is proportional to enzyme activity. After antibody incubation, color development solution is added, and absorbance is measured by reading in the microplate spectrophotometer at a wavelength of 450 nm. DNMT activity then was calculated based on this formula:

$$\text{DNMT Activity (OD/h/mg)} = \frac{(Sample\ OD - Blank\ OD)}{(Protein\ amount\ (\mu g)\ x\ Time\ (hour))} \times 1000$$

In the described assay, total protein amount was 8 µg and incubation time 1.6 hours.

2.4.2 DNA methyltransferase 3A activity assay

The DNA methyltransferase 3A (DNMT3A) Direct Activity Assay (BPS BioScience) is an ELISA-based method with a DNA substrate coated in the bottom of a 96-well plate. This is an efficient template for methylation by the supplied DNMT3A/3L enzyme. Following the manufacturers protocol, recombinant DNMT3A provided in the kit was combined with various concentrations of the peptides described in section 2.3.1. Each well included one concentration of one H3.3 peptide and DNMT assay buffer well-mixed with S-adenosylmethionine beforehand. There was no peptide in control wells and no enzymes in blank wells. The reaction was initiated by adding DNMT3A/3L complex and then incubated at 37 ºC for 2 hours. An anti-5-methylcytosine antibody was then added to detect methylation, followed by detecting the primary antibody with horseradish peroxidase (HRP) secondary antibody. Every well was washed by TBST buffer then added Blocking buffer after each step. The DNMT3A activity is directly proportional to the chemiluminescent signal which we detected with an Infinite 200 PRO plate reader (TECAN). The values of chemiluminescence, after being subtracted the blank value, was compared with the control without peptide and arbitrarily set to maximum (100%) activity. Statistical analysis of the collected data was by nonlinear regression analyses (curve-fit): dose response variable slope (four parameters) in Prism 6 (GraphPad Software). Tabulated is the peptide concentrations used in the assay and the formula for calculating DNMT3A activity.

Concentration	Peptide: H3.3K36me3; H3.3-WT; H3.3-G34W							
[μM]	0.5	1	2	4	8	16	32	64

$$\text{DNMT3A Activity (\%)} = \frac{(Sample\ luminescence - Blank\ luminescence)}{(Control\ luminescennce - Blank\ luminescence)} \times 100$$

2.5. Docking predictions, PWWP domain family phylogeny and visualization

Numerous crystal structures of PWWP-proteins have been published and are freely available via PDBe, UniProt, InterPro and others. Information of individual PWWP proteins was obtained from HGNC (genenames.org) and the UniProt database (uniprot.org). They were then aligned by T-Coffee (Tree-based Consistency Objective Function for alignment Evaluation, tcoffee.crg.cat), a progressive alignment similar to ClustalW to identify similarities and differences between sequences. A web-based server ESPript 3.0 (espript.ibcp.fr/ESPript) was applied to visualize the alignment with the structural information formatting with the crystal structure of the PWWP domain of Human DNA (cytosine-5-)-methyltransferase 3 alpha (PDB code: 3LLR). This is a program showing sequence similarities and secondary structure information from aligned sequences. Phylogenetic trees were built from the multiple alignment of the PWWP sequences by ClustalW (genome.jp/tools-bin/clustalw), contributing to observation of the relationship between these proteins and their H3K36me3-binding possibility.

Protein-protein docking experiments were conducted with the online tool ClusPro 2.0 (cluspro.bu.edu) by fitting H3.3-G34W peptide, residues 31-40 (STG**W**VKKPHR), with the PWWP domain of DNMT3A and others. Using PyMOL (37), the software for molecular visualization, the peptides were modelled as antiparallel beta structure. Based on multiple alignment and phylogenetic tree, the PWWP proteins are collected from Protein Data Bank (PDB) listed in Table 2.

Table 2. PWWP crystal structures applied to docking analysis

PWWP proteins	PDB code	Chain	Name
DNMT3B	5CIU	A	Structural basis of the recognition of H3K36me3 by DNMT3B PWWP domain
DNMT3A	3LLR	A	Crystal structure of the PWWP domain of Human DNA (cytosine-5-)-methyltransferase 3 alpha
PWWP2B	4LD6	A	PWWP domain of human PWWP Domain-Containing Protein 2B
BRPF1	3MO8	A	PWWP Domain of Human Bromodomain and PHD finger-containing protein 1 In Complex with Trimethylated H3K36 Peptide
BRD1	3LYI	A	PWWP Domain of Human Bromodomain-Containing Protein 1
BRPF3	3PFS	A	PWWP Domain of Human Bromodomain and PHD finger-containing protein 3
ZMYND11	4N4I	A	Crystal structure of the Bromo-PWWP of the mouse zinc finger MYND-type containing 11 isoform alpha in complex with histone H3.3K36me3

MSH6	6OQM	A	Crystal structure of the MSH6 PWWP domain
PWWP3A	3PMI	B	PWWP Domain of Human Mutated Melanoma-Associated Antigen 1
WHSC1L1 (NSD3) N-terminal	6G3T	A	X-ray structure of NSD3-PWWP1
WHSC1L1 (NSD3) C-terminal	4RXJ	A	Crystal structure of WHSC1L1-PWWP2
HDGFRP2	3EAE	A	PWWP domain of human hepatoma-derived growth factor 2 (HDGF2)
HDGFRP3	6IIS	A	Complex structure of the HRP3 PWWP domain with both a 16-bp TA-rich DNA and an H3K36me3-containing histone peptide
PSIP1	4FU6	A	Crystal structure of the PSIP1 PWWP domain

The ClusPro database provides machine-learning algorithms that models protein-protein interactions as balanced, electrostatic, hydrophobic, van der Waals and electrostatics forces. It provides correlative model scores, which consist of size, energy of the cluster center, and the lowest energy of the cluster. As stated in the ClusPro manual, the size of a cluster is approximately proportional to the probability of near-native structure and lower energy that contribute to build a larger cluster of docked structure. In other words, the most likely conformation is selected based on the most populated cluster and the lowest energy landscape is considered as an indirect criterion (38-41). After docking between PWWP structures as the receptor and H3.3-G34W peptides as the ligand, the models were visualized by PyMOL to evaluate the predicted

binding between the conserved aromatic cage of the crystal structure of PWWP and the modified histone H3.3.

Another docking method is AutoDock Vina. The ligand was prepared by building the peptide H3.3-G34W residue 32-38 (TG**W**V**K**KP) by Maestro, which is a powerful, multi-platform graphical user interface to build structures. The exported peptide was then docked with PWWP domains of DNMT3A and B by AutoDock Vina, a program using energy-based scoring method. The criterion to select the best model from AutoDock Vina is based on binding energy, unlike ClusPro which provides structures with highly populated clusters. The models from AutoDock Vina were shown in order of final total binding energy, which is the affinity measured in kcal/mol.

3. Results

3.1 Proliferation assay of isoH3.3-G34W and SETD2 knockout indicate opposing growth defects

To determine if the mutated H3.3 influence growth, a proliferation assay was performed. The proliferation assay revealed that the isogenic HeLa cell line H3.3-G34W (denoted isoH3.3-G34W) proliferated with increased rate compared to isoH3.3-WT (Figure 3). As H3K36 methylation is relevant to G34W, the enzyme catalyzing trimethylation of H3K36, SETD2, is tested for the proliferation and differentiation. In this assay, two independent SETD2 knockout-clones both grew less frequently compared to control. Together this suggests that loss of H3K36me3 globally is negative for proliferation of normal and isoH3.3-G34W cells.

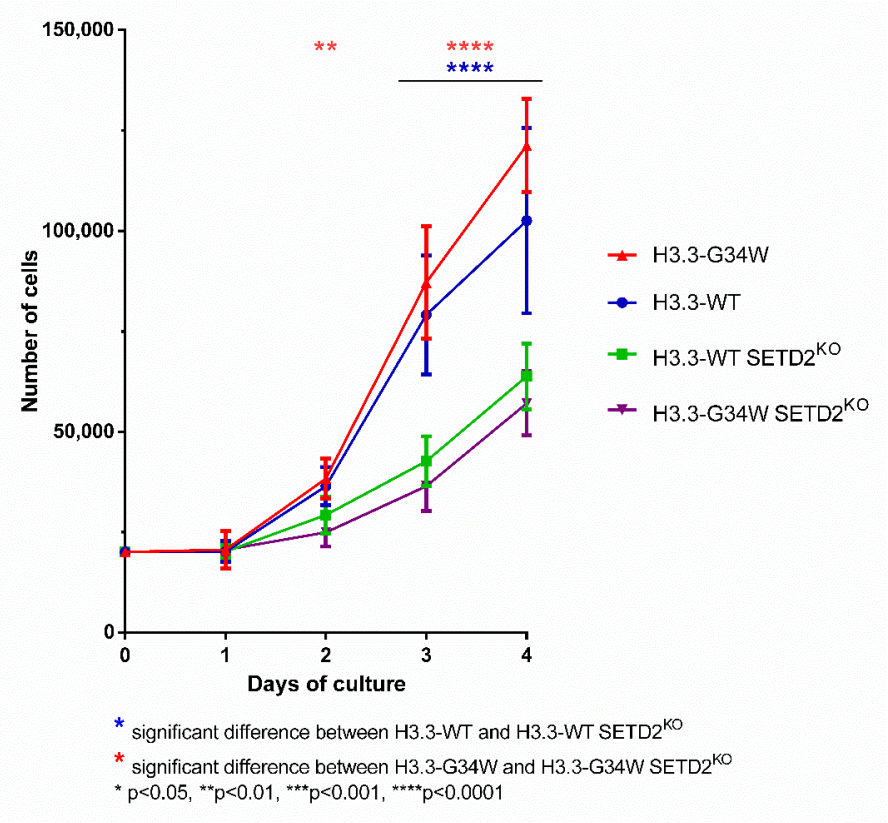


Figure 3. Isogenic HeLa cells with H3.3-G34W with slightly elevated proliferation status. Manual counting of two isogenic H3.3-WT and -G34W clones, for which cells were counted over four days, did not reveal a significant difference. Removal of H3K36me3 by knockout of SETD2 reduced cell growth equally much in both H3.3-WT and -G34W cells. Two technical replicates per time point and clone were tested. Student t-test is indicated with asterisk for significance: * p-value ≤ 0.05, ** p-value ≤ 0.01, *** p-value ≤ 0.001, **** p-value ≤ 0.0001.

3.2 The change of infiltrative capacity with G34W mutated histone

One of the key features of cancer is the ability to migrate. To determine if isoH3.3-mutants migrate more readily, the Boyden chamber assay was employed in isogenic MG-63 cell lines. In this assay, cells would have to degrade the extracellular matrix in order to migrate through a porous membrane and reach the opposite surface that is facing the serum-rich media to survive. Cells that fail to migrate through the membrane will be scraped off and will not be included in the assay (36). Thirty-five thousand cells were seeded in each chamber cup and counted for migration, after which t-test statistical analysis was applied to the data. While Figure 4 indicates that all three clones migrated in larger numbers compared to wildtype, the two G34-mutants exhibited significant greater migration rate in MG-63 cells compared to WT (*p*-value is 0.03 and 0.02 for G34R and G34W, respectively). In line with previously published observations on MG-63 from the Lindroth lab (31), G34W substitution shows more aggressive migratory features than normal cells.

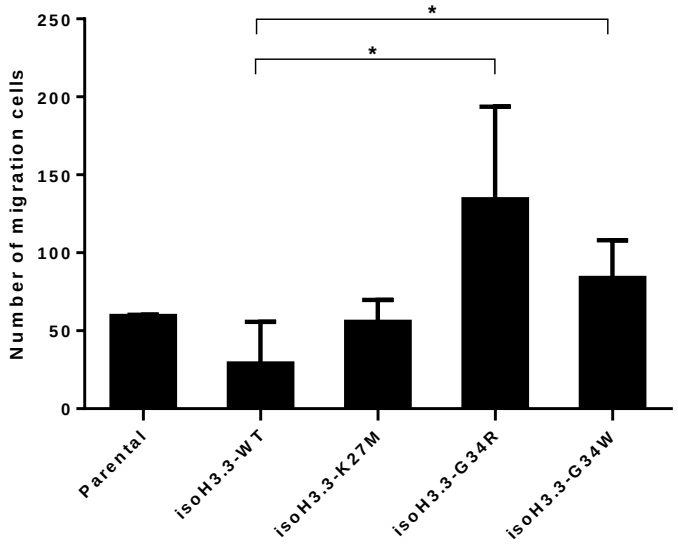


Figure 4. Isogenic MG-63 cells with G34 mutations migrate more readily WT. Migration of MG-63 cells was estimated after applying extracellular matrix to isogenic cells in a Boyden chamber assay. While isoH3.3-K27M did not reveal a migratory advantage, H3.3-G34R and G34W migrated about 5 and 3 folds more than WT, respectively. Student t-test (unpaired two-tail test) is indicated with asterisk for significant difference: * p-value ≤ 0.05.

3.3 Detailed analysis of binding properties between H3.3-G34W and the PWWP-domain of DNMT3A indicates unequivocal interaction

Experiments previously performed in the Lindroth lab indicated that H3.3-G34W interacts with the PWWP-domain of DNMT3A. Those experiments were

pull-down assays of purified DNMT3A and H3.3-peptides, as well as co-immunoprecipitation assays using transfected isoH3.3 HeLa cells with DNMT3A full-length constructs. In order to generate a binding constant for the interaction, surface plasmon resonance (SPR) experiment was performed. For DNMT3A-PWWP domain and H3.3K36me3, the interaction was tested in six distinct concentrations (from 6.25 to 200 µM), chosen based on published studies in that concentration range (42). The dose-response curve was used to estimate the association rate and the dissociation rate, and from that, the equilibrium dissociation constant $K_D = k_d / k_a$. A summary of the data is presented in Table 3 below, showing similar K_D for both H3.3K36me3 and H3.3-G34W interaction.

The K_D value of 630 µM determined by SPR for WT vis-à-vis PWWP interaction revealed weak to no affinity as it displayed extremely weak K_D-value. This was mainly caused by the increased dissociation rate k_d compared to histone H3K36me3 and H3.3-G34W. Importantly, there was no binding signal until the injected concentration of wildtype H3.3 peptide was higher than 25 µM. By contrast, the binding affinities of PWWP-K36me3 and PWWP-G34W interactions were much stronger and similar to each other where the K_D values for H3.3K36me3 and H3.3-G34W were 219 µM and 203 µM, respectively. The SPR response of each peptide to PWWP is depicted in Figure 5.

Table 3. The summarized kinetic parameters estimated by SPR methods.

Ligand	Analyte	Concentration (μM)	k_a (1/Ms)	k_d (1/s)	$t_{1/2}$ (s)	K_D (μM)
DNMT3A - PWWP	H3.3K36me3	6.25, 25, 50, 100, 200	3.8(1)e2	0.083(2)	8.351	219(6)
	H3.3-WT	~~6.25, 25~~, 50, 100, 200	3.4(5)e2	0.21(2)	3.301	630(90)
	H3.3-G34W	~~6.25~~, 25, 50, 100, 200	5.5(1)e2	0.111(2)	6.245	203(4)

Where: the number in the parentheses indicates the error value, and the unit follows the last digit.

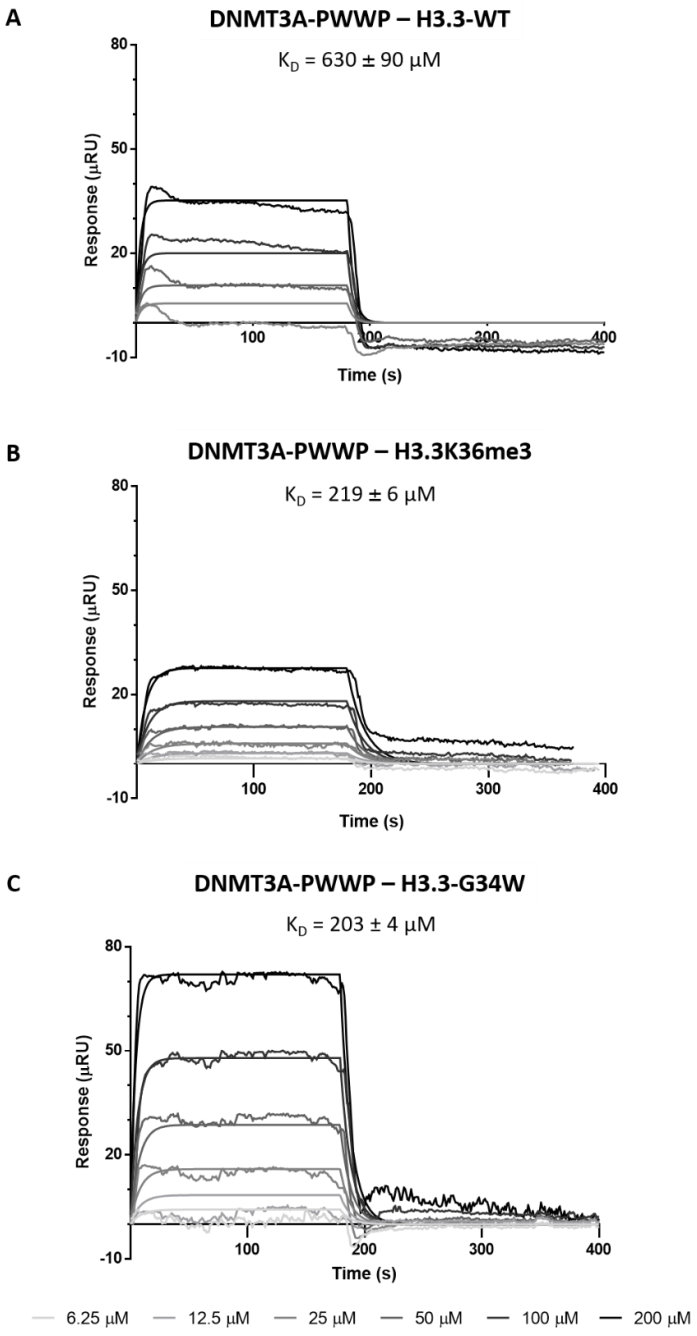


Figure 5. Binding between DNMT3A-PWWP and H3.3-G34W indicate similar affinity as to H3.3K36me3. Surface plasmon resonance binding assay

with PWWP of DNMT3A as ligand and H3 peptides as analytes suggest weak or absent binding to H3.3-WT peptide (K_D 630 μM) (A). In contrast, binding to H3.3K36me3 and H3.3-G34W peptides (B, C) were almost identical, with K_D equal to 219 and 203 μM, respectively.

3.4 DNA methyltransferase activity is diminished in the presence of H3.3-G34W

If binding of the PWWP-domain of DNMT3A to H3.3-G34W is to be contributing to the observed loss of DNA methylation in GCTB, it is reasonable to suspect a diminished catalytic activity of the methyltransferase. This was tested by measuring DNMT activity in isogenic cell lysates on DNA substrates immobilized in a microtiter plate-like ELISA setting (see details in 2.4.1). DNMT activity assays using isogenic HeLa cells with and without G34W was conducted to determine whether DNMT vis-à-vis H3.3-G34W interaction would affect the global DNMT activity in tumor cells.

Table 4. DNMT activity assay of isoH3.3-WT and -G34W cells

	OD_{450}	DNMT Activity	Average	GW decrease (%)
H3.3-WT	0.386155	28.533	29.221	**33.834**
	0.397755	29.500		
	0.395955	29.350		
	0.401555	29.817		
	0.380096	28.070		
	0.421896	31.336		
	0.390796	28.906		
	0.365696	26.945		
	0.411596	30.531		
H3.3-G34W	0.244955	16.767	19.334	
	0.259155	17.950		
	0.277255	19.458		
	0.294755	20.917		
	0.272355	19.050		
	0.257896	18.523		
	0.303996	22.125		
	0.247996	17.750		
	0.295596	21.469		

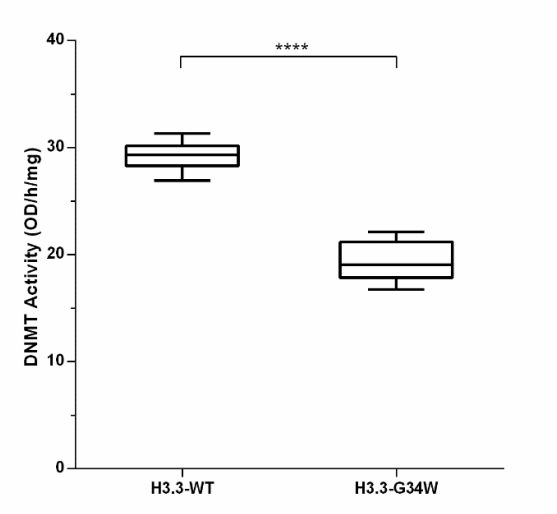


Figure 6. DNMT activity of isoH3.3 HeLa cells with H3.3-WT and -G34W cells indicate reduced activity in mutant cells. With the use of an ELISA-based DNMT assay, not distinguishing which enzymes provide methylation to the provided DNA substrate, isogenic cells with H3.3-G34W show a 33.8% decrease in methyltransferase activity. p value < 0.0001 determined using two-tailed unpaired t-test.

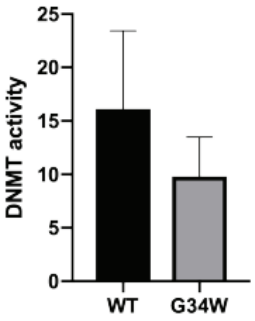


Figure 7. The reduced DNMT activity of GCTB stromal cells with H3.3-G34W cells compared to H3.3-WT (Lindroth et. al., personal communication). Using the same DNMT assay kit, the whole cell lysates of

GCTB were added to DNA substrate, showing reduced DNA methylation activity.

The assay is quantitative and was applied to whole cell lysates of unmutated isoH3.3-WT and isoH3.3-G34W. I noted a significant reduction of DNMT activity in H3.3-G34W clones compared to H3.3-WT clones in an unpaired two-tailed t-test with p-value < 0.0001 (Figure 6). This indicated that the interaction between PWWP of DNMT3A and H3.3-G34W negatively influenced DNA methyltransferase activity as catalytic activity decreased on average 33.8% compared to wildtype sample (Table 3). The result is consistent with the decrease in general DNMT activity calculated in H3.3-G34W GCTB stromal cells when compared to normal cells without the mutation (Figure 7).

3.5 Binding between DNMT3A-PWWP and H3.3-G34W reduce the methyltransferase activity of specific targets

The DNMT3A assay was conducted to determine if the interaction between H3.3-G34W and the PWWP-domain of DNMT3 was negatively influencing activity. Before measurement, chemiluminescence values were blanked by including all reagents minus substrate. A lower DNMT3A activity was observed when using H3.3-G34W peptide compared to H3.3K36me3 and H3.3-WT peptides, with the lowest inflection point at 4 nM. As indicated in Figure 8, the reactions H3.3K36me3 and H3.3-WT peptides did not show a reduced DNMT3A activity and did not generate an inflection point to report on.

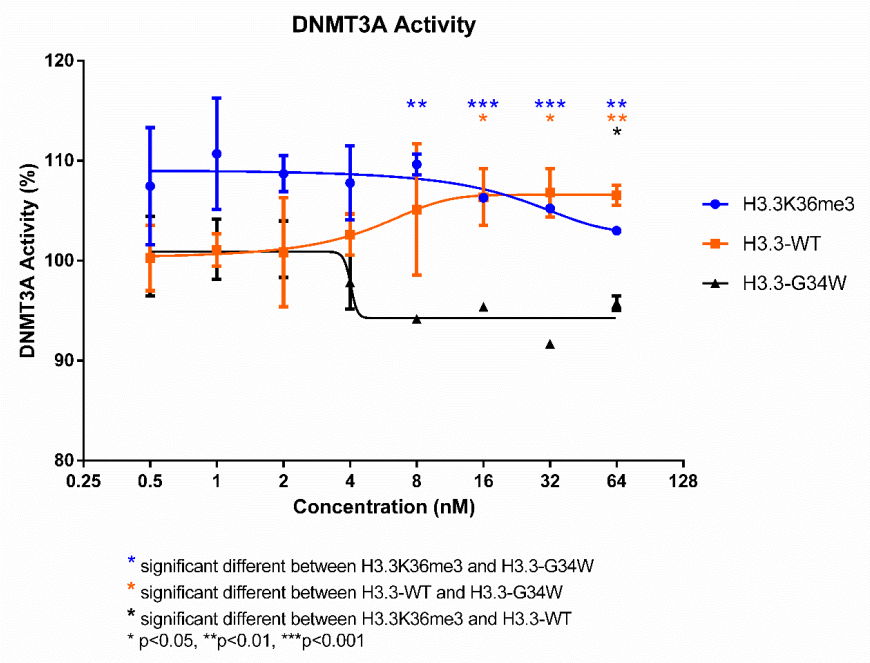


Figure 8. DNMT3A in the presence of H3.3-G34W deflates methyltransferase activity compared to WT and H3K36me3. Specifically testing DNMT3A in a methyltransferase assay in the presence of H3-peptides showed that WT and H3K36me3 did not substantially affected activity. In contrast, H3.3-G34W produced an inflection point of diminished activity at 4 nM, with an estimated decrease of activity nearly 10%.

3.6 Alignment and phylogenetic analysis of all known human PWWP domains

The observed analysis of H3.3-G34W being bound to PWWP of DNMT3A granted a detailed analysis of this and other PWWP-domain containing proteins.

The PWWP-domain exist in numerous proteins, sometimes occurring more than once, and is especially prominent in chromatin-related proteins (22). Alignment and phylogentic analysis provide details on conserved residues and may provide insight into specific functions which may explain how H3.3-G34W affect the proteins by which it interact, for example how DNMT3A display reduced activity upon interaction.

Using the multiple sequence alignment of all collected human PWWP domains by T-Coffee, ESPript was utilized to produce a synthesis of both sequence and structural information. PWWP-domains of DNMT3A and DNMT3B have been reported as DNA reader proteins by interactions with H3K36me3 at F303/206, W306/239 and W330/263 (21). The resulting alignment in Figure 9 revealed a PWWP domain, which in the literature have been well described with binding charecteristics to H3K36me3, containing a five-stranded beta-sheet packed with two alpha-helices, which was conserved among all PWWP family members. More interestingly, these three amino acids are completely maintained in PWWP2B, indicating promising binding potential as DNMTs.

Using the alignment of the PWWP domains from T-Coffee, which runs a similar progression strategy as ClustalW, a phylogenetic tree was built by ClustalW. With a neighbor-joining algorithm, the ClustalW calculates all pairwise alignments for an input set of 27 sequences, and constructs a phylogenetic tree for the input sequences based on their pairwise similarity

scores. In Figure 12, the phylogenetic tree of PWWP family revealed a close relationship between DNMT3A/B and PWWP2A/B proteins.

To visualize the binding properties of H3.3-G34W to PWWP, a computational docking model was generated by ClusPro (40). A peptide sequence representing residue 31-40 (STG**W**V**K**KPHR) of the H3.3 protein, where glycine 34 was replaced with tryptophan, were built by PyMOL in comparison to the WT sequence of the same length (Figure 10). The protein-protein docking tool ClusPro was used to predict binding properties based on a database of pre-existing proteins with known amino acid properties (39-41). ClusPro clusters the low-energy docked conformations based on the RMSD (pairwise Root Mean Square Deviation) and classifies the clusters according to their size. I applied the modelled H3.3-G34W peptide as the ligand with fourteen human PWWP crystal structures as the receptor (listed in Table 2) and docked them with the ClusPro server to find the binding motif when glycine 34 is substituted. This showed that G34W could fit the cavity of the aromatic cage normally occupied by H3K36me3 (Figure 11A, B and Supplementary 4). The same docking pattern was predicted in DNMT3A and PWWP2B (Figure 11C, D and Supplementary 4). For those PWWP proteins which did not indicate a spontaneous docking between G34W and the aromatic cage, an advanced biased option using receptor attraction was applied to point out the position of the three conserved residues in relation to G34W position. The docking prediction generated by ClusPro (Figure 11 and Supplementary 3) suggested that the H3.3-G34W occupies the aromatic cage of all PWWP domains much the same as PWWP docking to H3K36me3. While

published information on the intermolecular interactions between DNMT3A/3B and H3K36me3 was of the electrostatic type (21), the near-native structures of all PWWP-G34W conformations predicted by Cluspro favour hydrophobic interaction. This likely reflects the properties of tryptophan over glycine and the differences to methylated lysine 36. Docking results is indicated in the phylogenetic tree with an asterisk and numbers representing binding type (Figure 12). A similar binding fashion with only non-polar contacts observed was also found in the results of AutoDock Vina, showing the final binding energy between PWWP of DNMT3A and B with H3.3-G34W peptide is -6.3 and -7.1 kcal/mol, respectively.

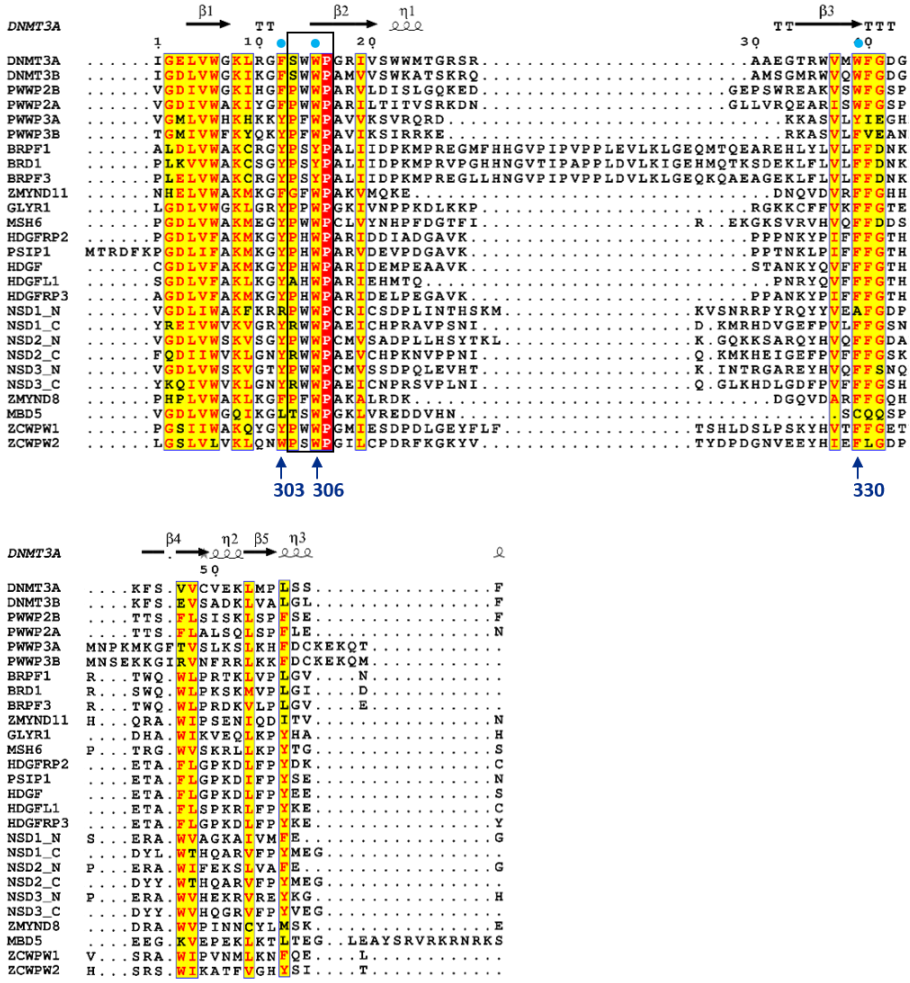


Figure 9. Alignment of all known human PWWP domains highlight strong evolutionary conservation. A total of 27 PWWP domains from 24 proteins were used for sequence comparison using the ClustalW algorithm. β-barrels β1-5 and alpha helices 1-3 are indicated on top. The residues forming the aromatic cage of DNMT3A that interacts with H3K36me are indicated with blue dots and numbers.

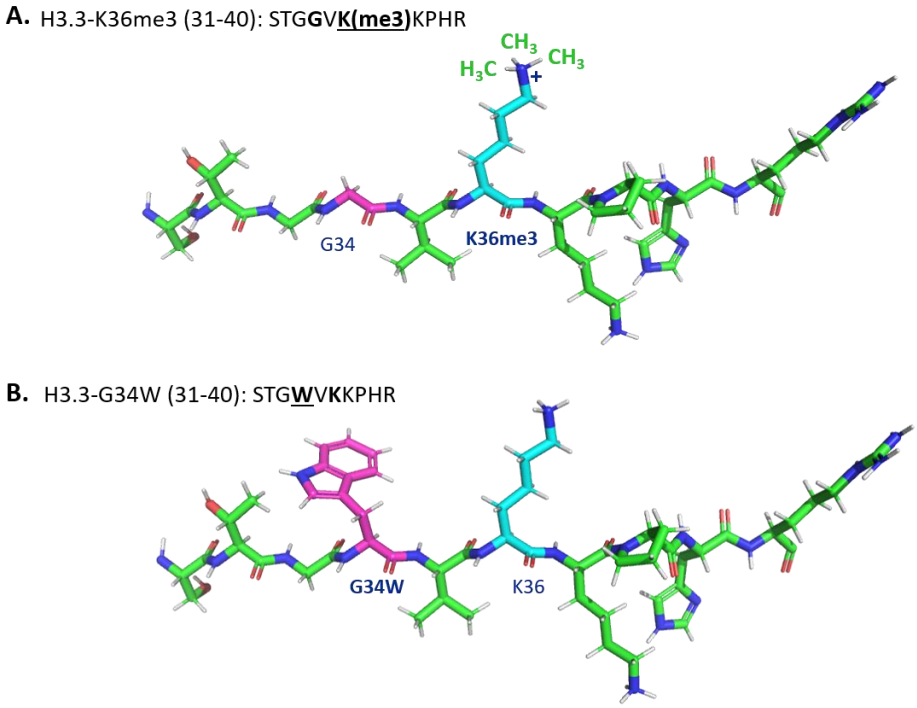
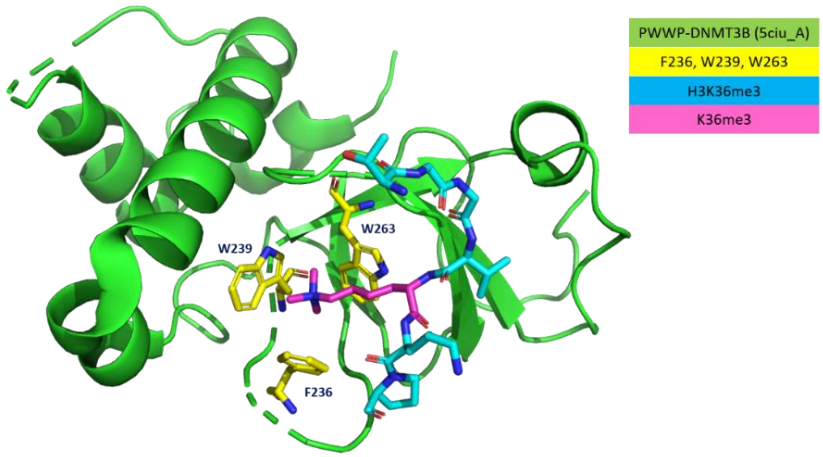
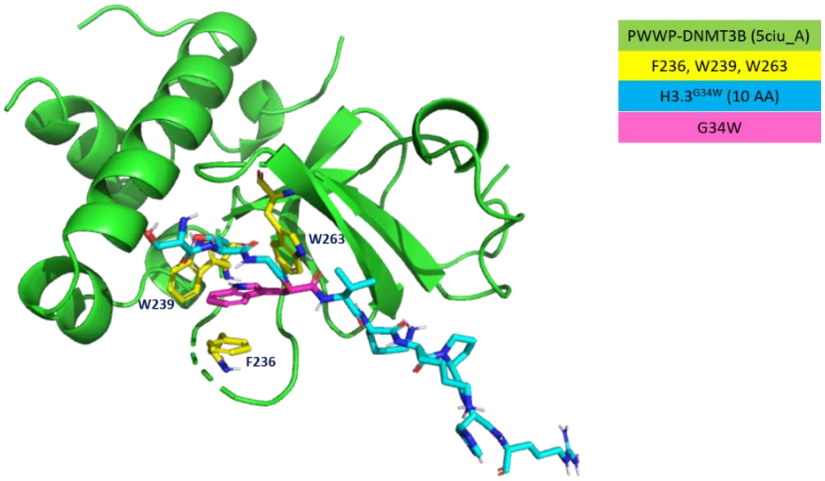


Figure 10. Stick-model illustration of H3.3K36me3 (A) and H3.3-G34W (B) peptides indicate close similarity. Ten amino acids centered around the modifications were used for each peptide and modeled in PyMOL. The lysine 36 (blue) in the former model involves three methyl groups, while G34 in the latter is replaced by tryptophan (pink) with similar size of K36me3. The H3.3-G34W model (B) then was used in ClusPro docking analysis as the ligand.

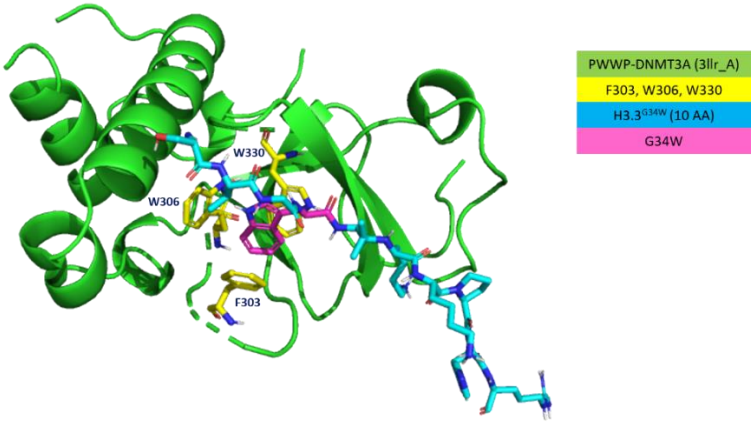
A


PWWP-DNMT3B (5ciu_A)
F236, W239, W263
H3K36me3
K36me3
W263
W239
F236

B


PWWP-DNMT3B (5ciu_A)
F236, W239, W263
H3.3G34W (10 AA)
G34W
W263
W239
F236

C


PWWP-DNMT3A (3llr_A)
F303, W306, W330
H3.3G34W (10 AA)
G34W
W330
W306
F303

D

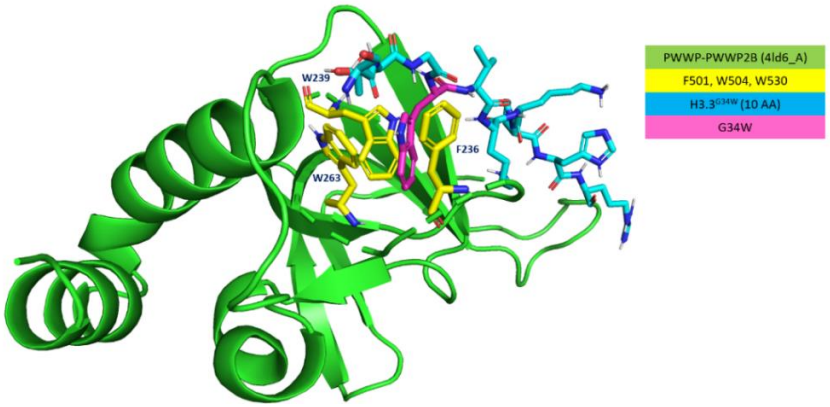


Figure 11. Protein-protein docking experiments of PWWP-domains and H3 peptides predict that H3.3-G34W occupy the same space as H3K36me3. Using the ClusPro database, H3K36me3 interaction guided docking of H3.3-G34W sequences.

A) Crystal structure of the recognition of H3K36me3 by DNMT3B PWWP domain (PDB: 5CIU) where K36me3 is highlighted in pink.

B) Docking model of S31-R40 segment of H3.3-G34W with PWWP domain of DNMT3B (5CIU_A). Histone peptide is colored as blue sticks. Tryptophan 34 (pink) is clustered to the aromatic residues (yellow) by hydrophobic link.

C) Hydrophobic-bonding structure between H3.3 peptide and G34W (pink) and DNMT3A (3LLR_A) in the cage of PWWP domain (yellow).

D) The hydrophobic structure cluster of the recognition of G34W sequence by PWWP2B-PWWP (4LD6_A) is depicted in stick mode with yellow for the three aromatic amino acids surrounding tryptophan 34 (pink).

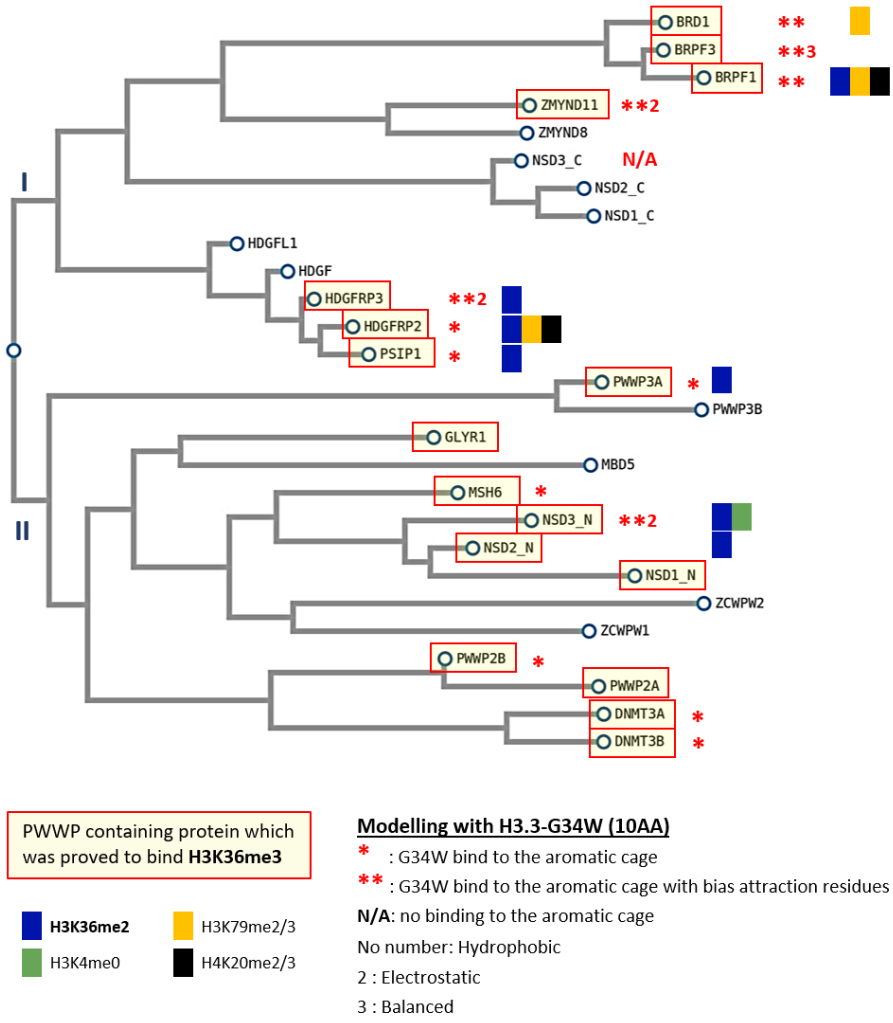


Figure 12. Phylogenetic tree based on multiple alignment of PWWP domains indicate monophyletic groups suggestive of functional distinction. Applying the T-Coffee alignment above, a phylogentic tree was constructed using ClustalW. The PWWP members with the H3K36me3 recognition ability are highlighted in red boxes. H3.3-G34W binding prediction to the aromatic cage of PWWP by ClusPro are indicated as asterisks and numbers correlated to chemical bond. Two monophyletic groups, denoted I and II indicated close to the

ancestral branch, both bind H3K36me3, yet more so in II than I. Group I contains transcription factors while II chromatin remodelers. The best interaction to H3.3-G34W is observed amoung group II members, especially DNMT3A and PWWP2B.

4. Discussion

Histones are crucial components of a functional chromatin structure and provide guidelines for gene expressions promoting cellular differentiation. In addition, post translational modification of the N-terminal tails of histones widely contribute to transcription regulation and outcomes. The last decade has uncovered frequent mutations of histones in cancer, especially childhood brain and bone cancers. Giant cell tumor of bone with its high frequency of glycine-to-tryptophan substitutions in H3.3 elevates proliferative activity, colony formation, migration capacity, and leading to abnormal RNA processing (31).

Published evidence suggest that G34 substitutions affect acquisition of H3K36 methylation, and this may in part explain the heterogenic nature of cancer with this alteration (29, 32). Other known G34 substitutions, including valine (V) and arginine (R) and asparagine (D), also prevents H3K36 from being methylated. One study suggests interaction with mismatch repair protein MutSα may lead to genome instability and tumorigenesis (33). Global decrease in H3K36me3 is demonstrated in H3.3-K36M which inhibit the methyltransferases MMSET/NSD2 (multiple myeloma SET domain/Nuclear SET Domain-Containing Protein 2) as well as SETD2. I observed decreased proliferation in SETD2 knock-out in both H3.3-G34W and WT HeLa cells, suggesting that H3K36me3 operate independent of H3.3-G34W and that the mutation is not driving tumor formation based on H3K36me3 status (Figure 3).

A recent study from the Lindroth lab uncovered genome-wide epigenetic alterations in GCTB, the main change was loss of DNA methylation (35). Yet, the connection between H3.3-G34W and DNA methylation remained unresolved. I have in this thesis studied the connection between PWWP-domains of the DNA methyltransferases DNMT3 and H3.3-G34W as they are both connected to H3K36me3. As stated before, both DNMT3A and DNMT3B are *de novo* methyltransferases essential for differentiation (2, 43), guiding DNA methylation via H3K36 gene body methylation by means of their PWWP domains (3). I used surface plasmon resonance assays to identify the properties by which PWWP recognizes H3.3K36me3 and H3.3-G34W. They were shown to interact with similar values of affinity parameters (Figure 5). To further labor these findings, docking and modelling experiments, PWWP domain in both DNMT3A and B, indicated the three conservation aromatic residues involved in the interaction with not only trimethylated H3 at K36 but also G34W. The DNMT3A and B share 53% of identity and 67% of similarity but the essential cage residues are perfectly conserved (21). The proposed cluster models provided similar binding between the two DNMT3 PWWPs and H3.3-G34W as with what was reported using the PWWP crystal structure onto H3K36me3 (Figure 11). I found that the interaction shifted from electrostatic in H3K36me3 to hydrophobic in H3.3-G34W, potentially explained that tryptophan is a hydrophobic residue similar to methylated lysine 36. Yet, the modellings suggest that G34W substitution fit the pocket of the conserved aromatic cage in the absence of H3K36me3.

It has been clearly shown in both tumor cells of GCTB and isogenic HeLa cells with H3.3-G34W that the mutation trigger the epigenetic changes, prominently leading to a 20% decrease in DNA methylation (35). How does this hypomethylation come about and can we explain the mechanism behind this epigenetic pattern? I performed DNMT assays that significantly pointed to a consistent decrease in DNA methylation in the presence of H3.3-G34W compared to unmutated or unmodified H3.3. *In vitro* assays with only DNMT3A and H3.3-G34W indicated a decrease sufficient to explain the diminished activity. Together, these data provide credence to the notion that H3.3-G34W is the sole contributor to reduced methyltransferase activity and DNA hypomethylation in GCTB.

The PWWP domain has been described in many chromatin proteins, and almost all of them have been shown to interact with H3K36 di- or tri-methylation. Multiple articles have shown that the molecular interactions are important for transcription, to maintain genome integrity and cell differentiation (10, 44-46). Inhibiting DNMT3s activity leading to loss of DNA methylation, H3.3-G34W may be sufficient to drive oncogenesis by substituting for loss of H3.3K36me3 in other PWWP-carrying proteins. Based on sequence alignment and phylogeny analysis of all members of the PWWP family, H3.3-G34W may generate aberrant interactions furthering the damage beyond DNA methylation. A detailed analysis of those interaction and activities that may have as a result of the interactions, will be an important task for the future. As chromatin proteins, they may have global effects on gene expression and cancer cell behavior.

I have in this thesis shown that the dominant-negative oncohistone H3.3-G34W interact with a domain, PWWP, that influence catalytic activity independent of H3K36me3 status. Only the two *de novo* methyltransferases have PWWP while the maintenance methyltransferase DNMT1 does not. This set apart the necessity to target a new compared to maintain DNA methylation as essential to the pattern of DNA methylation. It will be important to determine how *de novo* methylation drive tumorigenesis, and giant cell tumor with its unique sole contributor in H3.3-G34W may provide novel insight into the generality of this phenomenon.

5. Future direction

To fully uncover the scope and importance of the findings in this thesis, I would suggest venturing into making more sophisticated *in vivo* models. This would entail introducing oncohistones into mouse models. One important question would be to determine how strong tumorigenic capacity the H3.3-G34W have.

The docking modelling can be developed further by generating a crystal structure of H3.3-G34W peptide in complex with DNMT3A, DNMT3B and other PWWP candidates. This would solidify the predictions as something that indeed occur in the cell.

Another important aspect would be to identify treatments strategies to block the mutations or the activities. Due to the knowledge of the conservation of PWWP domain proteins where the aromatic cage recognizes histone H3.3K36me3, it is plausible to find a possible cancer treatment for G34W-related GCTB related to other PWWP-containing members. If methods suggested above are effective, they can be tested on the giant cell tumor of bone cancer cells before a Phase 2-3 study on patients.

There are also interesting possibilities based on these findings to understand how DNA methylation is regulated. Those findings can be applied to all kind of subjects related with DNA methylation regulation.

Bibliography

1. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev. 2002;16(1):6-21.

2. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999;99(3):247-57.

3. Dhayalan A, Rajavelu A, Rathert P, Tamas R, Jurkowska RZ, Ragozin S, et al. The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. J Biol Chem. 2010;285(34):26114-20.

4. Maze I, Noh KM, Soshnev AA, Allis CD. Every amino acid matters: essential contributions of histone variants to mammalian development and disease. Nat Rev Genet. 2014;15(4):259-71.

5. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res. 2011;21(3):381-95.

6. Volkel P, Angrand PO. The control of histone lysine methylation in epigenetic regulation. Biochimie. 2007;89(1):1-20.

7. Wozniak GG, Strahl BD. Hitting the 'mark': interpreting lysine methylation in the context of active transcription. Biochim Biophys Acta. 2014;1839(12):1353-61.

8. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet. 2007;39(3):311-8.

9. Edmunds JW, Mahadevan LC, Clayton AL. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. EMBO J. 2008;27(2):406-20.

10. Wen H, Li Y, Xi Y, Jiang S, Stratton S, Peng D, et al. ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression. Nature. 2014;508(7495):263-8.

11. Guo R, Zheng L, Park JW, Lv R, Chen H, Jiao F, et al. BS69/ZMYND11 reads and connects histone H3.3 lysine 36 trimethylation-decorated chromatin to regulated pre-mRNA processing. Mol Cell. 2014;56(2):298-310.

12. Baubec T, Colombo DF, Wirbelauer C, Schmidt J, Burger L, Krebs AR, et al. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature. 2015;520(7546):243-7.

13. Elsaesser SJ, Goldberg AD, Allis CD. New functions for an old variant: no substitute for histone H3.3. Curr Opin Genet Dev. 2010;20(2):110-7.

14. Bannister AJ, Schneider R, Kouzarides T. Histone methylation: dynamic or static? Cell. 2002;109(7):801-6.

15. Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, et al. The Polycomb group protein EZH2 directly controls DNA methylation. Nature. 2006;439(7078):871-4.

16. Schwartzentruber J, Korshunov A, Liu XY, Jones DT, Pfaff E, Jacob K, et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature. 2012;482(7384):226-31.

17. Zhang Y, Jurkowska R, Soeroes S, Rajavelu A, Dhayalan A, Bock I, et

al. Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. Nucleic Acids Res. 2010;38(13):4246-53.

18. Hansen RS, Wijmenga C, Luo P, Stanek AM, Canfield TK, Weemaes CM, et al. The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. Proc Natl Acad Sci U S A. 1999;96(25):14412-7.

19. Chen C, Nott TJ, Jin J, Pawson T. Deciphering arginine methylation: Tudor tells the tale. Nat Rev Mol Cell Biol. 2011;12(10):629-42.

20. Rona GB, Eleutherio ECA, Pinheiro AS. PWWP domains and their modes of sensing DNA and histone methylated lysines. Biophys Rev. 2016;8(1):63-74.

21. Rondelet G, Dal Maso T, Willems L, Wouters J. Structural basis for recognition of histone H3K36me3 nucleosome by human de novo DNA methyltransferases 3A and 3B. J Struct Biol. 2016;194(3):357-67.

22. Qin S, Min J. Structure and function of the nucleosome-binding PWWP domain. Trends Biochem Sci. 2014;39(11):536-47.

23. Bhattacharjee D, Shenoy S, Bairy KL. DNA Methylation and Chromatin Remodeling: The Blueprint of Cancer Epigenetics. Scientifica (Cairo). 2016;2016:6072357.

24. Ehrlich M. DNA methylation in cancer: too much, but also too little. Oncogene. 2002;21(35):5400-13.

25. Kim YI, Giuliano A, Hatch KD, Schneider A, Nour MA, Dallal GE, et al. Global DNA hypomethylation increases progressively in cervical dysplasia and

carcinoma. Cancer. 1994;74(3):893-9.

26. Soares J, Pinto AE, Cunha CV, Andre S, Barao I, Sousa JM, et al. Global DNA hypomethylation in breast carcinoma: correlation with prognostic factors and tumor progression. Cancer. 1999;85(1):112-8.

27. Werner M. Giant cell tumour of bone: morphological, biological and histogenetical aspects. Int Orthop. 2006;30(6):484-9.

28. Sobti A, Agrawal P, Agarwala S, Agarwal M. Giant Cell Tumor of Bone - An Overview. Arch Bone Jt Surg. 2016;4(1):2-9.

29. Shi L, Shi J, Shi X, Li W, Wen H. Histone H3.3 G34 Mutations Alter Histone H3K36 and H3K27 Methylation In Cis. J Mol Biol. 2018;430(11):1562-5.

30. Behjati S, Tarpey PS, Presneau N, Scheipl S, Pillay N, Van Loo P, et al. Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. Nat Genet. 2013;45(12):1479-82.

31. Lim J, Park JH, Baude A, Yoo Y, Lee YK, Schmidt CR, et al. The histone variant H3.3 G34W substitution in giant cell tumor of the bone link chromatin and RNA processing. Sci Rep. 2017;7(1):13459.

32. Fellenberg J, Sahr H, Mancarella D, Plass C, Lindroth AM, Westhauser F, et al. Knock-down of oncohistone H3F3A-G34W counteracts the neoplastic phenotype of giant cell tumor of bone derived stromal cells. Cancer Lett. 2019;448:61-9.

33. Fang J, Huang Y, Mao G, Yang S, Rennert G, Gu L, et al. Cancer-driving H3G34V/R/D mutations block H3K36 methylation and H3K36me3-MutSalpha

interaction. Proc Natl Acad Sci U S A. 2018;115(38):9598-603.

34. Jain SU, Khazaei S, Marchione DM, Lundgren SM, Wang X, Weinberg DN, et al. Histone H3.3 G34 mutations promote aberrant PRC2 activity and drive tumor progression. Proc Natl Acad Sci U S A. 2020;117(44):27354-64.

35. Lutsik P, Baude A, Mancarella D, Oz S, Kuhn A, Toth R, et al. Globally altered epigenetic landscape and delayed osteogenic differentiation in H3.3-G34W-mutant giant cell tumor of bone. Nat Commun. 2020;11(1):5414.

36. Chen H-C. Boyden chamber assay. Cell migration: Springer; 2005. p. 15-22.

37. Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. 2015.

38. Desta IT, Porter KA, Xia B, Kozakov D, Vajda S. Performance and Its Limits in Rigid Body Protein-Protein Docking. Structure. 2020;28(9):1071-81 e3.

39. Vajda S, Yueh C, Beglov D, Bohnuud T, Mottarella SE, Xia B, et al. New additions to the ClusPro server motivated by CAPRI. Proteins. 2017;85(3):435-44.

40. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, et al. The ClusPro web server for protein-protein docking. Nat Protoc. 2017;12(2):255-78.

41. Kozakov D, Beglov D, Bohnuud T, Mottarella SE, Xia B, Hall DR, et al. How good is automated protein docking? Proteins. 2013;81(12):2159-66.

42. Weinberg DN, Papillon-Cavanagh S, Chen H, Yue Y, Chen X, Rajagopalan KN, et al. The histone mark H3K36me2 recruits DNMT3A and

shapes the intergenic DNA methylation landscape. Nature. 2019;573(7773):281-6.

43. Okano M, Xie S, Li E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat Genet. 1998;19(3):219-20.

44. Huen MS, Huang J, Leung JW, Sy SM, Leung KM, Ching YP, et al. Regulation of chromatin architecture by the PWWP domain-containing DNA damage-responsive factor EXPAND1/MUM1. Mol Cell. 2010;37(6):854-64.

45. van Nuland R, van Schaik FM, Simonis M, van Heesch S, Cuppen E, Boelens R, et al. Nucleosomal DNA binding drives the recognition of H3K36-methylated nucleosomes by the PSIP1-PWWP domain. Epigenetics Chromatin. 2013;6(1):12.

46. Vezzoli A, Bonadies N, Allen MD, Freund SM, Santiveri CM, Kvinlaug BT, et al. Molecular basis of histone H3K36me3 recognition by the PWWP domain of Brpf1. Nat Struct Mol Biol. 2010;17(5):617-9.

47. Zhang T, Wei G, Millard CJ, Fischer R, Konietzny R, Kessler BM, et al. A variant NuRD complex containing PWWP2A/B excludes MBD2/3 to regulate transcription at active genes. Nat Commun. 2018;9(1):3798.

48. Wu H, Zeng H, Lam R, Tempel W, Amaya MF, Xu C, et al. Structural and histone binding ability characterizations of human PWWP domains. PLoS One. 2011;6(6):e18919.

49. Li F, Mao G, Tong D, Huang J, Gu L, Yang W, et al. The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with

MutSalpha. Cell. 2013;153(3):590-600.

50. Fei J, Ishii H, Hoeksema MA, Meitinger F, Kassavetis GA, Glass CK, et al. NDF, a nucleosome-destabilizing factor that facilitates transcription through nucleosomes. Genes Dev. 2018;32(9-10):682-94.

51. Qiao Q, Li Y, Chen Z, Wang M, Reinberg D, Xu RM. The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation. J Biol Chem. 2011;286(10):8361-8.

52. Tian W, Yan P, Xu N, Chakravorty A, Liefke R, Xi Q, et al. The HRP3 PWWP domain recognizes the minor groove of double-stranded DNA and recruits HRP3 to chromatin. Nucleic Acids Res. 2019;47(10):5436-48.

Supplementary 1: Summary of DNMT3A activity assay with three histone H3.3 peptides.

H3.3K36me3	Chemiluminescence		Activity (%)		
[nM]	replicate1	replicate2	replicate1	replicate2	Average
Blank	889008	881004			
0	3537285	3494955	100.804	99.196	100.000
0.5	3821773	3603513	111.617	103.322	107.469
1	3901514	3694407	114.648	106.776	110.712
2	3778889	3711785	109.987	107.437	108.712
4	3789934	3652745	110.407	105.193	107.800
8	3750869	3789346	108.922	110.384	109.653
16	3687934	3676786	106.530	106.106	106.318
32	3653755	3654762	105.231	105.269	105.250
64	3594329	3596547	102.972	103.057	103.015

H3.3-WT	Chemiluminescence		Activity (%)		
[nM]	replicate1	replicate2	replicate1	replicate2	Average
0	3537285	3494955	100.804	99.196	100.000
0.5	3462974	3585001	97.980	102.618	100.299
1	3514393	3574463	99.934	102.217	101.076
2	3640035	3437376	104.710	97.007	100.858
4	3623775	3546519	104.092	101.155	102.623
8	3773047	3529648	109.765	100.514	105.140
16	3630805	3736427	104.359	108.373	106.366
32	3650386	3740291	105.103	108.520	106.812
64	3707486	3669660	107.273	105.836	106.554

H3.3-G34W	Chemiluminescence		Activity (%)		
[nM]	replicate1	replicate2	replicate1	replicate2	Average
0	3537285	3494955	100.804	99.196	100.000
0.5	3602053	3453912	103.266	97.636	100.451
1	3602826	3490150	103.295	99.013	101.154
2	3599490	3494110	103.169	99.163	101.166
4	3510021	3409990	99.768	95.966	97.867
8	3347455	3360280	93.590	94.077	93.833
16	3391413	3398356	95.260	95.524	95.392
32	3302819	3292698	91.893	91.508	91.701
64	3418318	3391995	96.283	95.282	95.783

Supplementary 2: 17 proteins in Figure 12 carrying PWWP domain which can bind to histone H3.3K36me3 (with references)

No.	Protein	Name	Reference
1	DNMT3A	DNA (cytosine-5-)-methyltransferase 3 alpha	(21)
2	DNMT3B	DNA (cytosine-5)-methyltransferase 3 beta	(21)
3	PWWP2B	PWWP domain-containing protein 2B	(47)
4	PWWP2A	PWWP domain-containing protein 2A	(47)
5	BRPF1	Bromodomain and PHD finger-containing protein 1	(46)
6	BRD1	Bromodomain-containing protein 1	(48)
7	BRPF3	Bromodomain and PHD finger-containing protein 3	(48)
8	ZMYND11	Zinc finger MYND domain-containing protein 11	(10)
9	MSH6	DNA mismatch repair protein Msh6	(49)
10	PWWP3A	PWWP domain-containing DNA repair factor 3A	(48)
11	GLYR1	Putative oxidoreductase GLYR1 (Glyoxylate reductase 1 homolog)	(50)
12	NSD1_N	Nuclear receptor-binding SET domain-containing protein 1	(51)
13	NSD2_N	Nuclear receptor-binding SET domain-containing protein 2	(48)
14	NSD3_N	Nuclear receptor-binding SET domain-containing protein 3	(48)
15	HDGFRP2	Hepatoma-derived growth factor-related protein 2	(48)
16	HDGFRP3	Hepatoma-derived growth factor-related protein 3	(52)
17	PSIP1	PC4 and SFRS1-interacting protein (Lens epithelium-derived growth factor)	(45)

Supplementary 3: Summary of interaction possibility with the predicted chemical bond between the aromatic cage of PWWP domains and tryptophan 34 of H3.3-G34W peptide (10 amino acids) (ClusPro)

Protein	PDB code	Without Receptor Attraction	With Receptor Attraction
DNMT3B	5CIU_A	Hydrophobic	
DNMT3A	3LLR_A	Hydrophobic	
PWWP2B	4LD6_A	Hydrophobic	
BRPF1	3MO8_A	No binding	Hydrophobic
BRD1	3LYI_A	No binding	Hydrophobic
BRPF3	3PFS_A	No binding	Balanced
ZMYND11	4N4I_A	No binding	Electrostatic
MSH6	6OQM_A	Hydrophobic	
PWWP3A	3PMI_B	Hydrophobic	
NSD3_N	6G3T_A	No binding	Electrostatic
NSD3_C	4RXJ_A	No binding	No binding
HDGFRP2	3EAE_A	Hydrophobic	
HDGFRP3	6IIS_A	No binding	Electrostatic
PSIP1	4FU6_A	Hydrophobic	

Supplementary 4: ClusPro docking results of PWWP domains - H3.3-G34W interaction.

ClusPro suggests the most likely models of the complex should be selected as the centers of the large clusters with low energy structures rather than simply low energy structures. The outcomes shown in next three pages (58-60) list the clusters of docked structures in the order of cluster size (the number of docked structures), followed by the energy of the cluster center (the structure that has the highest number of neighbor structures in the cluster), and the energy of the lowest energy structure in the cluster. The highlighted clusters are the best models which all are hydrophobic interaction with both highest population and lowest binding energy.

DNMT3A-PWWP - H3.3-G34W

Balanced

Cluster	Members	Representative	Weighted Score
0	257	Center	-591.4
		Lowest Energy	-811.1
1	133	Center	-576.8
		Lowest Energy	-792.7
2	122	Center	-594.8
		Lowest Energy	-651.3
3	121	Center	-764.4
		Lowest Energy	-764.4
4	86	Center	-678.5
		Lowest Energy	-678.5
5	86	Center	-675.1
		Lowest Energy	-675.1
6	68	Center	-653.3
		Lowest Energy	-653.3
7	38	Center	-589
		Lowest Energy	-643
8	30	Center	-644.3
		Lowest Energy	-644.3
9	25	Center	-620.6
		Lowest Energy	-620.6
10	12	Center	-594.7
		Lowest Energy	-594.7

Electrostatic-favored

Cluster	Members	Representative	Weighted Score
0	153	Center	-575.4
		Lowest Energy	-746
1	130	Center	-597.9
		Lowest Energy	-780.5
2	96	Center	-574.9
		Lowest Energy	-730.1
3	95	Center	-592.1
		Lowest Energy	-660.2
4	73	Center	-673.9
		Lowest Energy	-673.9
5	66	Center	-593.5
		Lowest Energy	-627
6	61	Center	-649.9
		Lowest Energy	-649.9
7	52	Center	-685.9
		Lowest Energy	-685.9
8	45	Center	-582.2
		Lowest Energy	-648.8
9	45	Center	-582.2
		Lowest Energy	-620.8
10	41	Center	-593.6
		Lowest Energy	-648
11	31	Center	-578.4
		Lowest Energy	-615.8
12	19	Center	-649.5
		Lowest Energy	-649.5
13	19	Center	-616.5
		Lowest Energy	-616.5
14	18	Center	-595.9
		Lowest Energy	-604.9
15	14	Center	-575.7
		Lowest Energy	-683.5
16	13	Center	-576
		Lowest Energy	-607.8
17	4	Center	-583.4
		Lowest Energy	-583.4
18	2	Center	-577
		Lowest Energy	-578.4
19	1	Center	-575.3
		Lowest Energy	-575.3

Hydrophobic-favored

Cluster	Members	Representative	Weighted Score
0	**400**	**Center**	**-844.3**
		Lowest Energy	**-1069.7**
1	163	Center	-895.5
		Lowest Energy	-1077.7
2	137	Center	-1031.7
		Lowest Energy	-1031.7
3	63	Center	-772.9
		Lowest Energy	-881.5
4	60	Center	-868.7
		Lowest Energy	-868.7
5	54	Center	-871.5
		Lowest Energy	-871.5
6	41	Center	-850.7
		Lowest Energy	-850.7
7	36	Center	-828.3
		Lowest Energy	-828.3
8	18	Center	-840.6
		Lowest Energy	-840.6
9	13	Center	-812.2
		Lowest Energy	-812.2

Van der Waals + electrostatics

Cluster	Members	Representative	Weighted Score
0	260	Center	-128.7
		Lowest Energy	-149
1	228	Center	-122.1
		Lowest Energy	-149.7
2	82	Center	-124.5
		Lowest Energy	-144.8
3	73	Center	-122
		Lowest Energy	-137.7
4	65	Center	-131.5
		Lowest Energy	-140.6
5	52	Center	-124.4
		Lowest Energy	-154.3
6	51	Center	-124.8
		Lowest Energy	-140.4
7	41	Center	-122.1
		Lowest Energy	-141.3
8	39	Center	-123.7
		Lowest Energy	-136.3
9	27	Center	-123
		Lowest Energy	-131
10	27	Center	-125
		Lowest Energy	-135.7
11	26	Center	-124
		Lowest Energy	-135
12	18	Center	-121.4
		Lowest Energy	-128

DNMT3B-PWWP – H3.3-G34W

Balanced

Cluster	Members	Representative	Weighted Score
0	337	Center	-558.4
		Lowest Energy	-584.1
1	143	Center	-601.8
		Lowest Energy	-601.8
2	94	Center	-529.5
		Lowest Energy	-607.7
3	57	Center	-522.2
		Lowest Energy	-620.2
4	47	Center	-539.9
		Lowest Energy	-607.5
5	42	Center	-561
		Lowest Energy	-561
6	38	Center	-551.7
		Lowest Energy	-551.7
7	35	Center	-566.2
		Lowest Energy	-566.2
8	27	Center	-525.1
		Lowest Energy	-563.3
9	26	Center	-570.9
		Lowest Energy	-570.9
10	24	Center	-572.3
		Lowest Energy	-572.3
11	22	Center	-525.2
		Lowest Energy	-538.6
12	15	Center	-540.2
		Lowest Energy	-540.2
13	14	Center	-531.4
		Lowest Energy	-531.4
14	13	Center	-518.8
		Lowest Energy	-567
15	13	Center	-540.5
		Lowest Energy	-540.5
16	10	Center	-520.5
		Lowest Energy	-533.5
17	8	Center	-524.9
		Lowest Energy	-542.8
18	2	Center	-541.2
		Lowest Energy	-541.2

Electrostatic-favored

Cluster	Members	Representative	Weighted Score
0	192	Center	-550
		Lowest Energy	-571.3
1	145	Center	-596.8
		Lowest Energy	-596.8
2	98	Center	-569.2
		Lowest Energy	-569.2
3	91	Center	-534.7
		Lowest Energy	-609.7
4	78	Center	-511.3
		Lowest Energy	-586.1
5	77	Center	-632.9
		Lowest Energy	-632.9
6	54	Center	-520.5
		Lowest Energy	-559.7
7	52	Center	-560.5
		Lowest Energy	-569.3
8	36	Center	-517
		Lowest Energy	-597.1
9	30	Center	-546.8
		Lowest Energy	-546.8
10	26	Center	-513.8
		Lowest Energy	-564.4
11	24	Center	-541.1
		Lowest Energy	-541.1
12	23	Center	-523.8
		Lowest Energy	-568.8
13	18	Center	-527.9
		Lowest Energy	-540.1
14	12	Center	-509.4
		Lowest Energy	-534.3
15	11	Center	-533.7
		Lowest Energy	-533.7
16	6	Center	-524
		Lowest Energy	-529.2
17	1	Center	-515.6
		Lowest Energy	-515.6

Hydrophobic-favored

Cluster	Members	Representative	Weighted Score
0	**372**	**Center**	**-742.4**
		Lowest Energy	**-804.6**
1	351	Center	-838
		Lowest Energy	-849.6
2	190	Center	-768.3
		Lowest Energy	-768.3
3	27	Center	-763.7
		Lowest Energy	-763.7
4	22	Center	-703.5
		Lowest Energy	-741.1
5	14	Center	-731.1
		Lowest Energy	-731.1
6	12	Center	-740
		Lowest Energy	-740

Van der Waals + electrostatics

Cluster	Members	Representative	Weighted Score
0	282	Center	-95.5
		Lowest Energy	-139
1	97	Center	-96.1
		Lowest Energy	-120.8
2	81	Center	-99.6
		Lowest Energy	-115.6
3	76	Center	-97.1
		Lowest Energy	-108.7
4	75	Center	-98
		Lowest Energy	-107.1
5	67	Center	-98.6
		Lowest Energy	-107.8
6	52	Center	-94.6
		Lowest Energy	-117
7	41	Center	-94.4
		Lowest Energy	-107.5
8	36	Center	-98.2
		Lowest Energy	-110.7
9	35	Center	-95.2
		Lowest Energy	-105.7
10	31	Center	-93.8
		Lowest Energy	-103.4
11	28	Center	-94
		Lowest Energy	-102.5
12	25	Center	-95.4
		Lowest Energy	-109.2
13	23	Center	-101.2
		Lowest Energy	-102.3
14	22	Center	-94.5
		Lowest Energy	-107.6
15	10	Center	-100
		Lowest Energy	-109.3
16	4	Center	-95.3
		Lowest Energy	-98.2
17	3	Center	-93.7
		Lowest Energy	-98.9
18	2	Center	-97.2
		Lowest Energy	-97.2
19	1	Center	-96.5
		Lowest Energy	-96.5
20	1	Center	-100.6
		Lowest Energy	-100.6

PWWP2B-PWWP – H3.3-G34W

Balanced

Cluster	Members	Representative	Weighted Score
0	617	Center	-677.9
		Lowest Energy	-754.9
1	149	Center	-605.5
		Lowest Energy	-667.1
2	111	Center	-591.2
		Lowest Energy	-723.6
3	68	Center	-603.2
		Lowest Energy	-713.6
4	23	Center	-673.3
		Lowest Energy	-673.3
5	15	Center	-613.2
		Lowest Energy	-613.2
6	10	Center	-600.8
		Lowest Energy	-600.8

Electrostatic-favored

Cluster	Members	Representative	Weighted Score
0	854	Center	-668.7
		Lowest Energy	-826.8
1	81	Center	-616
		Lowest Energy	-703.4
2	46	Center	-627.9
		Lowest Energy	-663.8
3	15	Center	-618.8
		Lowest Energy	-645

Hydrophobic-favored

Cluster	Members	Representative	Weighted Score
0	**561**	**Center**	**-857**
		Lowest Energy	**-1008.7**
1	159	Center	-859
		Lowest Energy	-930.4
2	127	Center	-868.2
		Lowest Energy	-976.1
3	121	Center	-826.7
		Lowest Energy	-1025.6
4	31	Center	-866.2
		Lowest Energy	-866.2

Van der Waals + electrostatics

Cluster	Members	Representative	Weighted Score
0	175	Center	-141.8
		Lowest Energy	-165.3
1	172	Center	-136.7
		Lowest Energy	-172.7
2	118	Center	-136.5
		Lowest Energy	-167.6
3	99	Center	-138.7
		Lowest Energy	-156.4
4	82	Center	-135.1
		Lowest Energy	-152.4
5	55	Center	-135.7
		Lowest Energy	-154.2
6	50	Center	-138
		Lowest Energy	-159.3
7	46	Center	-144.4
		Lowest Energy	-155.7
8	36	Center	-144
		Lowest Energy	-144
9	35	Center	-138.5
		Lowest Energy	-153.7
10	28	Center	-139.3
		Lowest Energy	-159.8
11	21	Center	-140
		Lowest Energy	-149.6
12	19	Center	-142.7
		Lowest Energy	-148.9
13	17	Center	-137
		Lowest Energy	-146.8
14	16	Center	-136
		Lowest Energy	-146.3
15	9	Center	-137.1
		Lowest Energy	-140.8
16	8	Center	-135.2
		Lowest Energy	-139.9

ACKNOWLEDGEMENT

This thesis is the milestone of my study, and I wish to express my sincerest appreciation to all people helping me complete the wonderful journey at NCC-GCSP.

My deepest gratitude goes to my precious advisor, Professor Anders Lindroth, who expertly guided me through my 2-year graduate education and shared his passion for science with me. His patience and encouragement inspired my interest in cancer research and helped me do the right things even when the road got tough.

I would like to pay my special regards to my committee members, Professor Eun Jung Park and Professor Byeong Il Lee. Their insights and suggestions help me improve and accomplish this study.

I would like to recognize the invaluable support from my lab mates, Dr. Su Man Lee and Ji Hye Park, and my senior, Nguyen Viet Ha. My study would not have been possible without them, and I treasure every moment of working and staying with them. I also appreciate all NCC-GCSP Professors and NCC Staff, who provided me broad knowledge and assistance for my academic and daily life.

Special thanks go to my Vietnamese and foreign friends in NCC. I had good fortune to accompany my beloved brother and sister, and my roommate during

these two years. I am so grateful to meet you all in my life and I wish you all the best.

To my dear Family, thank you for supporting me to chase my dream and for your unconditional love.

Goyang, June 2021

Dao Hong Nhung