# Tagging endogenous H2A.Z1 and H2A.Z2 histone variant isoforms using CRISPR-Cas9 and identifying SUMOylated proteins in proximity to H2A.Z-containing nucleosome.

Myat Su Nwe Ma

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

**Master of Science**

GRADUATE PROGRAM IN BIOLOGY

YORK UNIVERSITY

TORONTO, ONTARIO

August 2018

## Abstract

H2A.Z is a variant of the H2A core histone, and is well known to be important for transcriptional regulation in eukaryotic cells. It was recently discovered that higher eukaryotes have two almost identical isoforms of H2A.Z that differ only by 3 amino acids. Although current biochemical studies are unable to distinguish between the native H2A.Z isoforms, several studies have shown potential non-redundant functions between the two. For my thesis work, CRISPR-Cas9 was used to insert a small epitope tag to the C-terminus region of either H2A.Z1 and H2A.Z2 in 293T cells to allow us to study the two isoforms individually in their native states. We successfully generated 3 H2A.Z1-Strep and 3 H2A.Z2-Strep clones, and confirmed that there were minimal off-target effects by Cas9 nuclease. However, the H2A.Z1-Strep clones were found to have much lower than normal H2A.Z1 expression whereas H2A.Z2-Strep clones have normal levels of H2A.Z expression. Strep-tag ChIP data showed H2A.Z1 being enriched on the proximal region of the transcription start site (TSS) of the MTA1 gene compared to H2A.Z2. For a second project, an Avi-tag BirA system was used to test for SUMOylation of H2A.Z in mammalian cells. This method was also used to determine whether there are proteins in proximity to H2A.Z that can be SUMOylated. Our assay did not detect any H2A.Z SUMOylation; however, there were a few proteins within close proximity to H2A.Z that appeared to be SUMOylated. These putative SUMOylated proteins (which are biotinylated by the fusion H2A.Z-BirA) were purified by Streptavidin-coupled beads and awaiting identification by mass spectrometry.

Acknowledgements

First and foremost, I would like to thank Dr. Peter Cheung for allowing me the chance to work in his lab and learn so much about everything, both academically and non-academically. I remember coming to you for an opportunity to research in your lab even without the most outstanding transcript nor much lab experience, and you still took me in anyways. Looking back, I realized how much you've played a role in allowing me to grow and improve as a person. I will never forget the skills and lessons I've learnt and apply them to everyday life. Thank you for being patient with me when I needed it and giving me the resources to act on my curiosity, albeit not all since grants don't come in easily. Your presence and support throughout my undergraduate and graduate years have really made some of the toughest times during my research bearable. Also, thank you for the caffeine, snacks, and the lab lunches!

Secondly, I would like to extend my gratitude and appreciation to Dr. Emanuel Rosonina, my advisor. Thank you for everything during this process, especially when we needed anything (antibodies, glycogen, agar plates etc.) and could come to you for it. Being able to just pop in your office to get your suggestions on failed experiments (i.e. heat shock) or talk to you about my Western blot results to get your feedback were extremely helpful.

To Dr. Mark Bayfield and Dr. Gerald Audette, thank you for sitting in my defense committee. I appreciate the time spent and hope my thesis would be a smooth read.

I would also like to thank all the members of Cheung lab. Firstly, to Dr. Kashif Khan, the post-doc in our lab, I appreciate all the times when I went to you for

everything; whining about negative results, bouncing off ideas about an experiment, talking about owning pets, asking about experiments, debating about the strike situations, fashion, life lessons about everything from lab experiences to personal life, and many more. To Sangyeon Han, and Faraz Elahi, thank you for welcoming me to the lab so warmly. I will never forget all the advice. Thank you Sang for showing me cat pictures when I needed it and Faraz for giving me tips on online shopping and socks. To Marlee Ng, thank you for sharing some of your knowledge on experiments. It really helped make some of my experiments go smoother. Jessica D'Angelo, spending my whole time at Cheung lab with you was such a joy and I'm glad I had you to rant about ice hockey. We should really come back and see how long our hockey themed clock in the lab would last. Alegria Indio and Shahir Morcos, thanks for making my time in the lab more upbeat and lively.

To my friends in other labs, the LSB squad, especially Veroni, John, Marjon, Gio, Fernand, and Daniel Miller, thank you for being around to talk about anything. Just having y'all around made my time in the lab more interesting and joyful.

Last but not least, I would like to thank me mom, dad, sister, Reema, Kim, and Steve for always being there, in both good times and bad. I can always count on you to listen to me without judgement and let it all out. Thank you for believing in me, supplying me food, and understanding on so many occasions.

Words cannot describe the amount of gratitude I feel towards everyone mentioned. Please know that I am extremely thankful for having you in my life and making my MSc journey one of the best experience ever.

## Table of Contents

## List of Tables

## List of Figures

# List of abbreviations

| | |
|---|---|
| ℃ | Degree Celsius |
| TSS | Transcription start site |
| SUMO | Small Ubiquitin-like Modifier |
| ChIP | Chromatin immunoprecipitation |
| C-terminus | Carboxyl-terminus |
| CRISPR-Cas9 | Clustered Regularly Interspaced Short Palindromic Region with Cas9 |
| HDR | Homology-directed repiar |
| IP | Immunoprecipitation |
| kDa | Kilo-dalton |
| N-terminus | Amino-terminus |
| NHEJ | Non-homologous end joining |
| PI | Protease inhibitors |
| RT | Room temperature |
| SB | Sample buffer |
| ssODNs | Single stranded oligodeoxynucleotides |
| Ub | Ubiquitin/Ubiquitylated |

# Chapter 1: General Introduction

## 1.1 DNA compaction in nucleus and its role in transcription regulation

The human genome is made up of approximately 6 billion base-pairs (bp) of DNA, which translates to roughly 2 meters in length. It is compacted into a nucleus that is typically 5-10 um in diameter (Nelson & Cox, 2008). In the eukaryotic nucleus, DNA is wrapped around histone proteins to form nucleosomes, which are the basic repeating units of chromatin. Each nucleosome comprises of about 147 bp of DNA wrapped around an octamer that contains 2 copies each of the 4 core histones, H2A, H2B, H3, and H4 (Richmond *et al*, 1984). This forms the first structural level of DNA organization. Nucleosomes are connected to one another by linker DNA, and some nucleosomes may also contain the linker histone, H1, that binds and locks in place the linker DNA regions that enter and exit the nucleosome structure (Olins and Olins, 1974; Luger *et al*, 1997; Carruthers *et al*, 1990). This forms a higher structural level of DNA organization. Despite the need to compact DNA in the nucleus, DNA also has to be accessible to DNA binding factors at the right time and regions. The eukaryotic genome can be organized into euchromatin, which are regions where the chromatin is less compact, or heterochromatin, which are regions containing more compact chromatin (Campos and Reinberg, 2009). Heterochromatin is further classed into constitutive heterochromatin, where it is always highly condensed, or facultative heterochromatin, where chromatin compaction can vary at different times during differentiation. Packing of the genome into different chromatin states allows for the control of DNA accessibility to nuclear proteins involved in different genome-mediated processes such as transcription. Therefore,

regulating chromatin compaction and nucleosomal positioning are important factors that contribute to the proper control of gene expression.

The importance of chromatin in transcription was elucidated when it was discovered that nucleosomes impeded transcription *in vitro* (Knezetic and Luse, 1986; Lorch *et al*, 1987) and that gene expression was affected in a specific manner when the histones or their basic tails were deleted (Han and Grunstein, 1988; Kayne *et al*, 1988). Additionally, the role of chromatin in transcription was further supported by the identifications of the first transcription-linked histone acetyltransferase (Brownell *et al*, 1996) and transcription-related chromatin-remodeling complexes (Cote *et al*, 1994). Accumulated evidence suggests that chromatin structure also impacts on other DNA-related metabolic processes besides transcription, such as recombination, DNA damage and repair, and replication. However, this thesis focuses on transcription, and more importantly, on epigenetic regulation of transcription. More specifically, my work focuses on post-transcription modifications of histones and the functions of the histone variant, H2A.Z.

## 1.2 Histone variants and its role in transcription regulation

### 1.2.1 Histones variants in chromatin structure and transcription

Chromatin structure can be modulated by ATP-dependent chromatin remodeling complexes as well as by replacement of core histones within the nucleosome context by histone variants of its own class (Cheung and Lau, 2005). Histone variants provide additional levels of chromatin organization since they can interact with variant-specific proteins or impart distinct structural and functional properties to the nucleosomes.

Deposition of histone variants are mediated by distinct chaperones and mechanisms, and represent additional levels of regulation by the cell (Szenker, Ray-Gallet, & Almouzni, 2011; Talbert & Henikoff, 2010). Histone variants are encoded by separate genes distinct from their core counterparts. Outside of the amino acid sequences that are conserved between the core and histone variants, each variant also has stretches of unique amino acid sequences that confer specific properties and functions. The core histones are replication dependent as they are mostly expressed in S phase and their incorporation into chromatin is connected to synthesis of DNA. This ensures that there are sufficient core histones for incorporation into nucleosomes at newly replicated DNA strands, and allow for efficient chromatin packaging. The 4 core histones are each encoded by multi-copy genes and their transcripts lack introns and poly A tails (Marzluff, Wagner, & Duronio, 2008). In contrast, transcription of histone variant genes is not restricted to S phase and is not dependent on replication of DNA (Wu and Bonner, 1981). Histone variants are encoded by single-copy genes and their transcripts contain both introns and poly A tails. When core histones are replaced by histone variants of its own class within the nucleosome, the biochemical properties of the nucleosome are modified. Therefore, histone variants add another layer of complexity as it can change the combinations of potential post-translational modifications (PTMs) of the histones within the nucleosomes, and can alter interactions of the nucleosome with other proteins or DNA, and/or change the structure of the chromatin (Talbert & Henikoff, 2010). Nucleosomes that contain two copies of the same histone are homotypic while nucleosomes that contain a core histone and a variant histone or two different variants

are called heterotypic. The existence of both homotypic and heterotypic nucleosome supports the 'nucleosome code' hypothesis (Bernstein & Hake, 2006), which explains increased flexibility in nucleosome formation, stability and structure.

### 1.2.2 H2A.Z histone variant and its positioning in the genome for gene expression regulation

In eukaryotes, there are different variants of H2A core histone, of which H2A.X, H2A.Bbd, macroH2A, and H2A.Z are the most widely studied. H2A.Z histone variant shares 60% homology in the amino acid sequences with its core counterpart H2A (Zlatanova and Thakar, 2008). It is also one of the more conserved histone variants, with approximately 80% similarity across different species (Lowell *et al*, 2005). Taken together, it implies that H2A.Z has non-redundant roles that other H2A variants are unable to compensate for. H2A.Z is deposited by the SWR-1 ATPase complex to replace the H2A within the nucleosome (Kobor *et al*, 2004). The structure and function of H2A.Z are evolutionarily conserved as phenotypes associated with the disruption of the H2A.Z-coding gene in yeast can be rescued by introduction of the H2A.Z gene from the ciliate protozoan *Tetrahymena* (Jackson and Gorozky, 2000). H2A.Z also has an essential role in eukaryotic cells since loss of H2A.Z in more complex organisms, such as *Tetrahymena thermophile, Drosophila melanogaster, Xenopus laevis*, and mice, all resulted in lethality (Liu *et al*, 1996; Clarkson *et al*, 1999; Faast *et al*, 2001). On the other hand, H2A.Z is not essential in less complex eukaryotes, such as *Saccharomyces cerevisiae*, since H2A.Z-null mutant yeast strains were viable but had defects in

transcriptional activation of inducible genes (Santisteban *et al*, 2000; Adam *et al*, 2001). Knock-out of H2A.Z in *Schizosaccharomyces pombe* also led to phenotypes indicating severe growth defects (Carr *et al*, 1994). Additionally, H2A.Z was found to be involved in recruitment of RNAPII in yeast and human cells (Adam *et al*, 2000; Hardy *et al*, 2009). These functional studies suggest that H2A.Z is involved in transcriptional initiation. H2A.Z- and H3.3-containing nucleosomes are less stable than homotypic nucleosomes, and are localized on nucleosome-free regions (NFR) of active promoters, enhancers and insulators (Chen, Wang, & Li, 2014; Jin et al., 2009). Specifically, histone variants H2A.Z and H3.3 are known to be involved in transcription regulation (Weber & Henikoff, 2014). Analyses of H2A.Z positioning within the yeast and mammalian genomes revealed H2A.Z-containing nucleosomes preferentially localized to promoter regions of genes and at nucleosomes flanking the nucleosome-free regions (NFR) at transcription start sites (Barski *et al*, 2007; John *et al*, 2008; Schones *et al*, 2008). This suggests that nucleosomes containing H2A.Z facilitates gene regulation by preventing stability of nucleosomes around regulatory regions that could be displaced easily by chromatin-binding complexes to both activate and repress gene expression (Hu et al., 2013). Chromatin with H2A.Z nucleosomes were also shown to result in higher susceptibility to DNA cleavage by nucleases compared to H2A nucleosomes (Tolstorukov *et al*, 2009). This indicated that more DNA was exposed when DNA was bound to nucleosomes containing H2A.Z than H2A histones, which may allow easier access by transcriptional activating proteins. Enhancers enriched with H2A.Z were linked to chromatin accessibility, epigenetic modifications and recruitment of RNA

polymerases (Brunelle *et al*, 2015). Enhancers are important in achieving transcriptional activation by recruitment of transcription factors and co-activator complexes to regulate target genes. Despite many studies showing a positive role of H2A.Z on transcription, it is important to note that there are also studies that showed H2A.Z can also localize to heterochromatin and centromeric regions, and are involved in negatively regulating gene transcription (Cheung and Draker, 2009). Moreover, there are also studies that showed contradictory results in the role of H2A.Z in nucleosome stability (Marque *et al*, 2010). H2A.Z was initially found at heterochromatin sites (Rangasamy *et al*, 2003), and proposed to promote higher-order chromatin structure formation. This is dependent on the 2 amino acid residues that resulted in an extended acidic patch of H2A.Z (compared to core H2A), which also interacts with HP1$\alpha$, a heterochromatin protein, to contribute or maintain constitutive heterochromatin domains (Fan *et al*, 2002). However, Fan *et al* (2002) also showed that H2A.Z inhibited the formation of higher order chromatin structure that stemmed from intermolecular association. This suggests that H2A.Z could play a role in constructing special chromatin regions poised for activation of transcription. Furthermore, H2A.Z could also localize to promoters of inducible genes to poise them for activation, such as those found at constitutive and hormone-induced glucocorticoid receptor binding sites (John *et al*, 2008). An enrichment of H2A.Z was also observed at the promoter of the p21 gene regulated by p53 in U2OS osteosarcoma cells, particularly at sites where p53 bind (Gevry*et al*, 2007). These studies suggested that the H2A.Z variant play significant roles in transcriptional activation at multiple levels of gene expression regulation. Despite some gaps in our full knowledge, the general

consensus is that H2A.Z plays highly dynamic roles in transcription regulation depending on its localization to different regions of the genome.

### 1.3 Role of H2A.Z in cancer

Epigenetic states can be regulated by histone variants by allowing or preventing stabilization of chromatin structure at the nucleosomal level (Bonisch & Hake, 2012). Cancer could hence be contributed by the deregulation of histone variants and their regulators at the transcription level. H2A.Z has been implicated in promoting cancer by increasing epigenetic plasticity in cancer cells, causing them to lose their ability to maintain genomic integrity, altering transcriptional regulation of target genes, and potentially affecting the cells' ability to inhibit senescence. The link between H2A.Z and cancer was first reported in sporadic colorectal tumors, where its overexpression was observed by genome-wide gene expression profiling (Dunican *et al*, 2002). H2A.Z was also found to be overexpressed inn progressive breast cancer, bladder and malignant melanoma (Zink & Hake, 2016). H2A.Z was also found to be involved in androgen receptor (AR) gene transactivation and prostate carcinoma (PCa) progression. Ubiquitin-specific protease (USP10) deubiquitylation of H2A.Z in K120, K121 and K125 resulted in AR-regulated PSA and KLK3 gene transcriptional activation (Draker *et al*, 2011). Additionally, PCa patients undergoing androgen-deprivation therapy generally expressed higher levels of H2A.Z over time, which could be indicative of androgen-independence in primary PCa progression (Dryhurst *et al*, 2012). These findings

suggest that H2A.Z could be exploited as markers of disease progression and as response to cancer therapy.

### 1.4 CRISPR-Cas9 genome editing tool

Clustered, regularly interspaced short palindromic repeats (CRISPR)-associated (Cas) system is an immune system of *Streptococcus pyogenes* that utilizes its endogenous nuclease guided by foreign RNA to cleave foreign genetic elements (Bhaya *et al*, 2011). It has been adapted and optimized as a powerful genome editing technique. The most characterized type of CRISPR system out of the three identified is the Type II (Gasiunas *et al*, 2012) system. It is composed of Cas9 protein and the CRISPR RNA (crRNA), whereby crRNA is matured by a trans-activating CRISPR RNA (tracrRNA) containing sequence complementary to repeat regions of pre-crRNA. Cas9 is an RNA-guided DNA endonuclease enzyme that induces DNA cleavage. Two domains in Cas9 endonuclease are responsible for its catalytic activity on the target DNA strand; 1. RuvC domains cuts the non-complementary strand; 2. HNH domain cuts the complementary strand of the DNA. The optimized CRISPR-Cas9 system for genome editing has developed a chimeric, single-guide RNA (sgRNA) that fused crRNA and tracrRNA together (Jinek *et al*, 2012) to guide the Cas9 nuclease to its target site on the DNA for cleavage. The sgRNA that guides Cas9 to its complementary target on the genome has to be immediately trailed by a 5'-NGG PAM sequence, where N is any of the four possible nucleotides. Double-stranded break induced by Cas9 endonuclease would occur 3 bp upstream of PAM (Jinek *et al*, 2012). After cleavage, there are two

distinct DNA repair pathways that the cell can undergo depending on the cell cycle stage: non-homologous end joining (NHEJ) or homology-directed repair (HDR) (Figure 1-1). NHEJ repairs DNA damage by directly ligating the broken ends without any templates, which often results in deletions or random nucleotide insertions. In HDR pathway, DNA templates that contain homology to the break ends are used to repair the gene with high fidelity. The nature of genome editing differs when different DNA repair pathways are taken. NHEJ often leads to gene knockouts whereas HDR would be useful for knock-in experiments.

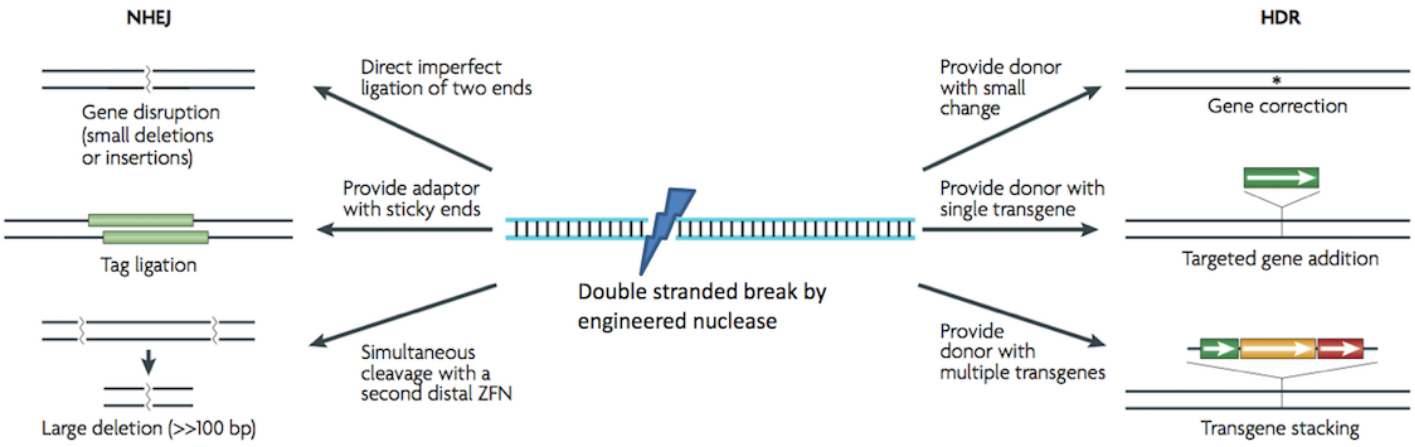



Figure 1-1. **Genome editing tools take advantage of endogenous DNA repair machineries to induce edits in the genome.** After double-stranded break by Cas9 to the target sequence upstream of PAM sequence, repair mechanisms follow to prevent permanent DNA damage. Depending on the cell cycle, non-homologous end joining (NHEJ) or homology-directed repair (HDR) can occur. In NHEJ, the blunt ends are directly ligated without any templates, resulting in deletions or random nucleotide insertions. In HDR, a template that contains homology to the break ends is provided to fix the DNA break to avoid any mutations in the gene. Through this system, gene knockout, via NHEJ, or knock-in, via HDR, lines can be created. (adapted from Urnov *et al*, 2010)

### 1.5 Avi-tag BirA system

The avi-tag BirA system is based on *in vivo* biotinylation that takes advantage of the *E. coli* biotin ligase BirA's ability to recognize and attach biotin to a short 15-amino acid peptide, Avi-tag (26) (Figure 1-2). This system is highly specific as there are no known endogenous proteins in mammalian cells that contain the 15-amino acid recognition sequence. Additionally, biotinylated avi-tagged proteins could be purified using Streptavidin beads in pull-down assays. Biotin binds to avidin/streptavidin beads with high affinity and so allows for highly stringent washes during purification of biotinylated proteins without fear of early elution or loss of the biotinylated proteins.

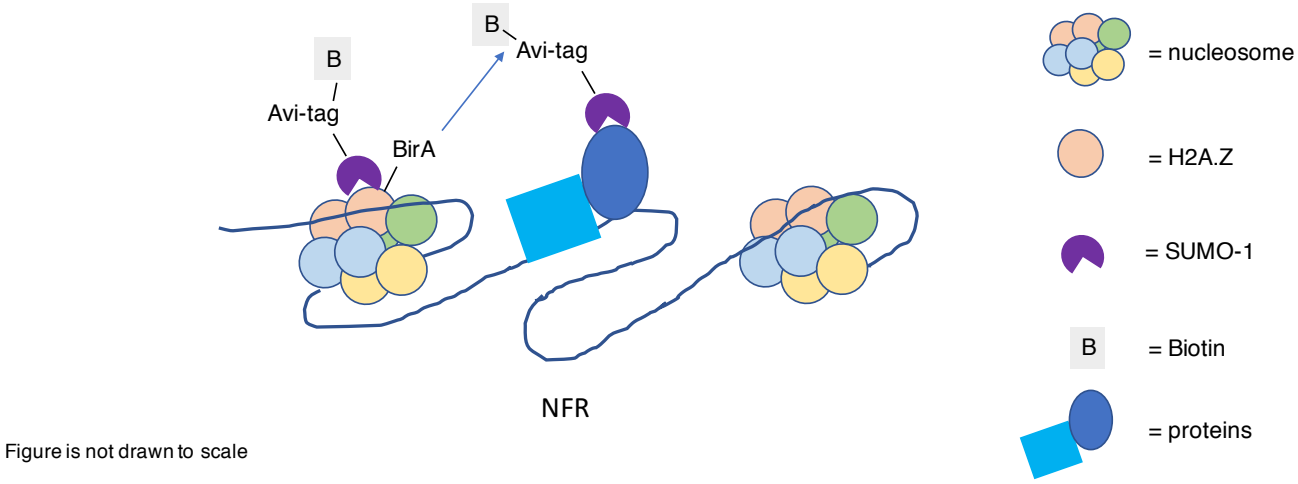



Figure 1-2. **General representation of the Avi-tag BirA biotinylation system used to target SUMOylated proteins on the nucleosome free region of the chromatin.** BirA ligase is fused to H2A.Z incorporated into nucleosomes, and Avi-tag is fused to SUMO-1. If proteins located at the transcription start site on the NFR are SUMOylated, a portion of SUMO and H2A.Z would be from the endogenous pool while others would contain the fusion constructs. When both constructs are localized in proximity to each other, BirA would recognize a unique sequence on the 15-amino acid Avi-tag and attach biotin to the lysine residue on Avi-tag.

# Chapter 2: Generation of endogenous H2A.Z1-Strep and H2A.Z2-Strep cell lines using CRISPR-Cas9

## 2.1 Introduction

### 2.1.1 H2A.Z histone variant isoforms, H2A.Z1 and H2A.Z2

Our current understanding of the H2A.Z histone variant is complicated by the recent discovery that there are two distinct isoforms, H2A.Z1 and H2A.Z2, that are encoded by two non-allelic genes H2AFZ and H2AFV respectively. Even though the DNA sequences of these two genes are not conserved, the protein amino acid sequences of the two isoforms are almost identical except for 3 amino acid differences at residues 15, 39, and 128 (Figure 2-1) (Dryhurst *et al*, 2009). Other highly conserved yet functionally distinct histone variants have been described before. For example, despite high similarity in the primary amino acid sequences between the core histone H3.1 and the histone variant H3.3 (they differ only by 4 amino acids), the latter's incorporation into nucleosomes significantly changes the stability of the nucleosome (Jin et al., 2009). Therefore, minor changes in amino acid sequences can impart distinct functional properties to highly conserved histone variants. The H2AFZ and H2AFV genes are driven by distinct promoters, raising the possibility that their expressions are regulated by distinct mechanisms and pathways. Despite their almost identical amino acid sequences, several pieces of evidence suggests H2A.Z1 and H2A.Z2 have non-redundant functions. Developing mice that have a deletion of the H2A.Z1 alleles were found to be embryonically lethal (Faast *et al*, 2004), suggesting that the presence of the intact H2A.Z2 gene is not sufficient to compensate for the loss of H2A.Z1 during

embryonic development. Knock-out of the genes coding for either H2A.Z1 or H2A.Z2 in chicken DT40 cells resulted in distinct phenotypic characteristics in terms of cell growth and expression of genes (Mastuda *et al*, 2010). For example, cells deficient in H2A.Z2 proliferated slower than cells lacking H2A.Z1. In melanoma cells, H2A.Z2 isoform was found to be overexpressed and required for the expression of genes targeted by the E2F transcription factors, whereas H2A.Z1 was not involved in the regulation of those target genes (Vardabasso *et al*, 2015). Knock-down of either H2A.Z1 or H2A.Z2 in the melanoma cells also resulted in differential gene transcription profile, supporting the differential functional roles of H2A.Z1 and H2A.Z2. Crystallography studies of H2A.Z1- and H2A.Z2-containing nucleosomes found distinct structural differences in the L1 loops of the respective isoforms, resulting in higher turnover rates of H2A.Z1 compared to H2A.Z2 in normal conditions (Horikoshi *et al*, 2013). In response to DNA damage resulting double stranded breaks, the opposite effect was observed where nucleosomal H2A.Z2 is more rapidly swapped than H2A.Z1 (Nishibuchi *et al*, 2014). All together, these studies illustrate the differential functions of H2A.Z1 and H2A.Z2 isoforms in transcriptional regulation and other nuclear processes.

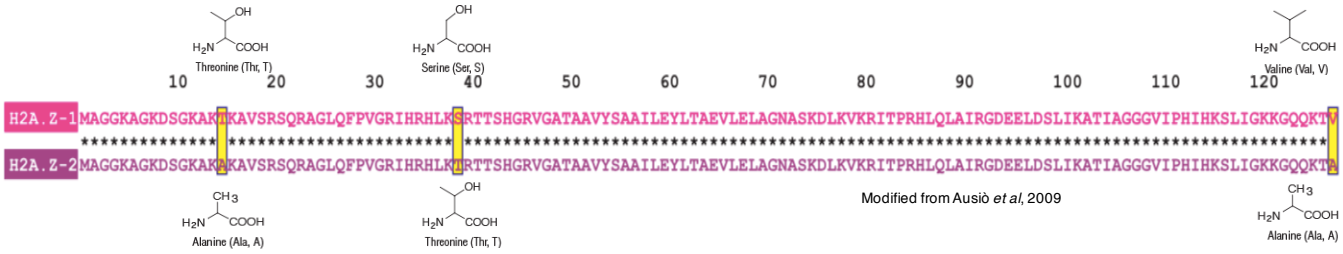



Figure 2-1. **Difference in amino acid sequence between H2A.Z1 and H2A.Z2.** While most of the amino acid sequences between the two isoforms are identical, there are three residues where they differ. At residues 15, 39, and 128, it is threonine, serine, and valine in H2A.Z1 while it is alanine, threonine, and alanine in H2A.Z2 respectively.

2.1.2 Research rationale and objective

To date, most biochemical studies of H2A.Z do not distinguish between H2A.Z1 and H2A.Z2 since there are no commercially available antibodies that can distinguish between the two isoforms. The three amino acids that differ between the isoforms are far apart and do not form a single distinguishable epitope together. Therefore, it is not easy to generate isoform-specific antibodies. Functional studies in mammalian cells often rely on ectopic expression of epitope-tagged H2A.Z1 and H2A.Z2 proteins to distinguish between the two isoforms. One of the main drawbacks of these ectopic expression studies is that the transgenes are not expressed from the endogenous promoters and therefore, not subjected to the native regulated expression of the endogenous genes. Also, differential expression of the H2A.Z isoforms can only be determined at the mRNA levels, but not at the protein levels. Therefore, the goal of this project is to add the DNA sequence of an epitope tag in frame to the endogenous H2A.Z1 or H2A.Z2 coding sequences in order to generate fusion protein that can be detected by antibodies against the tag (Figure 2-2). The tagged isoforms would still be driven by the respective endogenous promoters and, therefore, allow more physiologically relevant studies of the two isoforms. To achieve our goal, we utilized the genome editing system, CRISPR-Cas9, to insert a Strep-tag DNA sequence just before its stop codon at the 3' end of the endogenous H2A.Z1 and H2A.Z1 genes. Once expressed, H2A.Z1 or H2A.Z2 would have a Strep-tag present at its C-terminus. Strep-tag, which comprises 8 amino acids and approximately 1 kDa in size, was chosen for its small size and molecular weight addition. We hoped that the small tag would hence not

occlude or interfere with the functionality of the regions on H2A.Z1 or H2A.Z2. Additionally, Strep-tag binds with a strong affinity to Streptactin, which would be useful for affinity purification approaches. For example, immunoprecipitation of Strep-tagged H2A.Z1 and H2A.Z2 could be conducted and used to identify co-purifying interactors. Chromatin immunoprecipitation (ChIP) could also be performed and identify genomic regions that differentially associate with the Strep-tagged H2A.Z1 or H2A.Z2 isoforms. We decided to generate endogenous H2A.Z1-Strep and H2A.Z2-Strep cell lines in 293T cells due to the easy handling and high transfection efficiency of these cells.

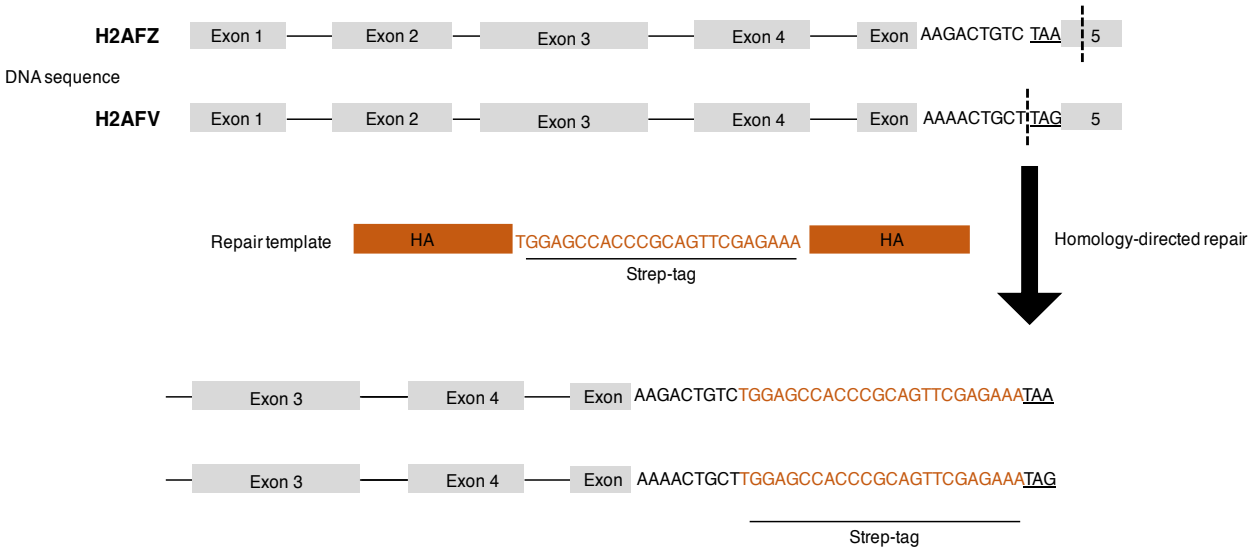



Figure 2-2. **Objective of using CRISPR-Cas9 genome editing system.** Strep tag would be inserted right before the Stop codon of H2AFZ and H2AFV without resulting in any mutations elsewhere. This would prevent any frameshift mutations as well. Repair templates would be present for the cell to undergo homology-directed repair. When the genes are expressed, the Strep-tag would be transcribed and translated as well. Stop codon is underlined.

2.1.3 Designing an optimized CRISPR-Cas9 genome editing process for insertion of Strep-tag to H2A.Z histone variant isoforms

In the CRISPR-Cas9 system, double strand breaks introduced by the Cas9 nuclease can either be repaired by the NHEJ or HDR pathways. For this study, given the goal is to insert the Strep-tag sequence in frame to the endogenous H2A.Z1 and H2A.Z2 coding sequence, the HDR mechanism is needed to perform this task (Figure 2-3). HDR generally occurs in the late S and G2 phase of the cell cycle, when sister chromatids are present to serve as repair templates after DNA replication to fix any errors that might have occurred (Heyer *et al*, 2010). In contrast, NHEJ takes place throughout the whole cell cycle and acts as a competitor to HDR for DNA reparation, even at S and G2 phase. It also actively inhibits HDR at M phase and early G1 (Orthwein *et al*, 2014). Therefore, the overall frequency of HDR occuring in normal cycling cells is generally quite low, which explains why inserting DNA sequences into the genome by the CRISPR-Cas9 system is known to be difficult and low in efficiency.

We have attempted to enhance HDR efficiency by using Nocodazole to arrest cell cycle at G2/M phase, where HDR is known to be most efficient (Gutschner *et al* 2016). Prior optimization experiment using FACS analysis showed that Nocodazole treatment at 200 ng/mL for 20 hours synchronized the highest amount of cell population at the G2/M phase using FACS analysis and higher amounts of Cas9 protein expressed. Additionally, i53 protein was also used to inhibit 53BP1, which is an important regulator of DSB repair pathway choice that normally promotes NHEJ by preventing end resection to inhibit HDR (Canny et al., 2017). By inhibiting 53BP1 that

promotes NHEJ, others have shown that addition of i53 to the CRISPR-Cas9 system can improve the success rate of generating genomic insertions of targeted DNA sequences.

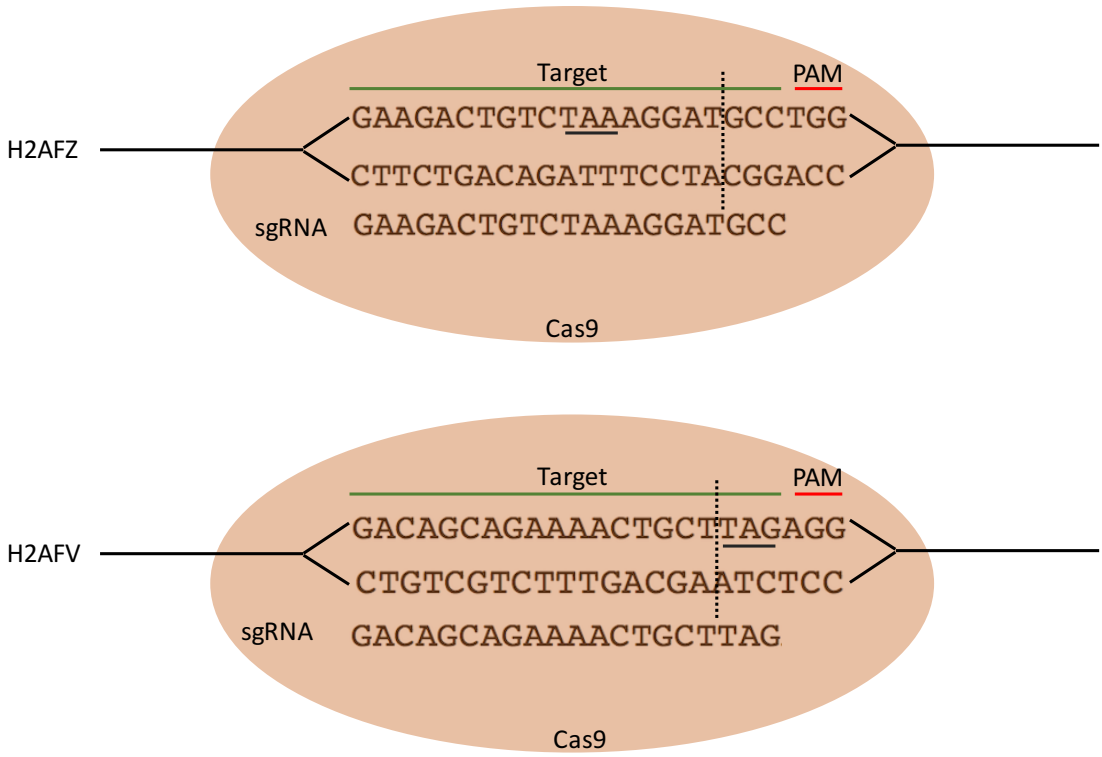



Figure 2-3. **Designs of sgRNA and repair templates to genes of H2A.Z1 and H2A.Z2; H2AFZ and H2AFV.** The respective sgRNA is targeted to the C-terminus of the 2 genes, and Cas9 will cleave 3 nucleotides upstream of the unique PAM sequence. For H2AFZ, the cut site is 5 nucleotides downstream of the site of insertion. For H2AFV, the cut site is exactly at the site of insertion, making it ideal. Repair templates for both H2AFZ and H2AFV contain the Strep-tag sequence and 3 linkers upstream, with both flanked by symmetric homology arms to their respective genes. The linker nucleotide sequence used is GGGAGGGAGGGATCC, which would translate to GGGS. The linker sequences were used to distance the H2A.Z1 and H2A.Z2 protein from the Strep-tag.

## 2.2 Materials and methods

### 2.2.1 CRISPR-Cas9 procedure, and plasmids

Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin were used to grow 293T cells. They were then used when their passage numbers were still low for transfection and generation of CRISPR-Cas9 clones with Strep-tag inserted at the C-terminus of H2A.Z1 and H2A.Z2. The CRISPR-Cas9 expression constructs were based on the PX459 plasmid (Addgene #62988, Ran *et al*, 2013) backbone containing the gene sequence for the SpCas9 protein and a cloning site for the single guide RNAs (sgRNAs). Previously, the unique sgRNA sequence targeting the C-terminus of H2A.Z1 and H2A.Z2, as well as a Scrambled sgRNA sequence that is known to not target any sites on the human genome have been successfully cloned into PX459 plasmid vector. PX459 plasmid containing sgRNA targeting C-terminus of H2A.Z1 and H2A.Z2 would be henceforth referred to as H2A.Z1 sgRNA and H2A.Z2 sgRNA respectively. PX459 plasmid containing the sgRNA for the Scrambled sequence would be referred to interchangeably as Scr-sgRNA. Repair template used had 35 base pairs serving as 5' and 3' homology arms to the C-terminus region of either H2A.Z1 or H2A.Z2 flanking a three-linker sequence (GGGSGGGSGGGSA) and strep-tag (WSHPQFEK). All co-transfections were conducted using PolyJet reagent with either H2A.Z1-sgRNA or H2A.Z2-sgRNA or Scr-sgRNA and the respective repair templates as well as the plasmid encoding i53. PolyJet-containing DMEM was replaced with media containing 200 ng/mL and treated for 20 hours. Fresh media was replaced and the cells were

released from cell synchronization for 24 hours before Puromycin selection was conducted at 1 ug/mL for 3 days.

2.2.2 DNA analysis: extraction, PCR, restriction enzyme analysis and gel analysis

DNA was extracted from clones via the QuickExtract™ DNA Extraction Protocol and quantified using the Thermo Scientific NanoDrop 2000/2000c Spectrophotometer. Equal concentrations of 500 ng DNA were used in the PCR analysis to determine if Strep-tag has been inserted. The primer pairs for H2A.Z1 and H2A.Z2 were both designed to amplify the last exon (exon 5) and some region at 3'UTR of H2A.Z1 or H2A.Z2, and inclusive of the Strep-tag integration if it had been integrated between the last exon and 3'UTR. The primer pair for Strep-tag was designed to target just the Strep-tag in the gene. PCRs (0.25 uL Taq polymerase, 5 uL ThermoPol Buffer, 1 uL each of forward and reverse primer, and 1 uL of 10mM dNTP per sample) for both H2A.Z1 and H2A.Z2 primer pairs were performed under the same conditions except for the annealing temperature. PCR products were then purified with the Wizard SV Genomic DNA Purification System from Promega. Restriction enzyme digestion was performed on H2A.Z1 and H2A.Z2 PCR products with BamHI-HF and 10X CutSmart buffer in 37℃ for 3 hours. Larger PCR products were run on a 1% agarose gel with RedSafe while smaller PCR products were run on 10% TBE polyacrylamide gel with 1X TBE buffer at 150 volts for 50 minutes before soaking it in RedSafe Dye for 20 minutes. The gel was washed in dd$H_2O$ 3 times for a minute each before exposure to UV light to observe bands.

2.2.3 Western blot analysis

For whole cell protein extraction, boiling hot (@95℃) 2x sample buffer (SB) was added to the cell pellets and immediately mixed. They were then sonicated at 35% duty cycle and 40% output control 13 pulses. The samples were then centrifuged at high speed and the supernatant was transferred to new tubes. Samples were then stored at -20℃ or used in subsequent Western blot analyses. To obtain proteins from nuclear extract, cell pellets were washed with Buffer A and lysed with Buffer A + 0.1% Triton-X100 for 5 minutes on ice. The samples were spun at 1300G for 5 minutes at 4℃ and the supernatant containing the cytoplasmic fraction was removed. Nuclear fraction was lysed with boiling hot 2x sample buffer (SB) and sonicated. The protein samples were either stored at -20℃ or used in subsequent Western blot analyses.

Protein samples were loaded in 15% polyacrylamide gels when blotting for histones or 10% polyacrylamide gels when blotting for higher molecular weight proteins. SDS gel-electrophoresis was conducted at 150V for 17 minutes and 200V for 58 minutes in SDS running buffer. The proteins separated on the SDS gels were transferred to PVDF membranes using semi-dry when detecting for low molecular weight proteins, or wet transfer apparatus when detecting for high molecular weight proteins. The PVDF membranes were then blocked in 5% milk in TBS-T for 1 hour at room temperature (RT), and blotted with incubated with appropriate primary antibodies overnight at 4℃ in 2% milk with TBS-T. Strep-tag II antibody was used to detect the Strep-tagged H2A.Z isoforms. In addition, three other antibodies were used to detect

the endogenous as well as tagged H2A.Z isoforms: Active motif H2A.Z and antibody #1379 (previously made in the lab) recognize the C-terminus of H2A.Z, whereas antibody #1380 (also previously made in the lab) is directed against the L1 loop of the H2A.Z histone variant. The endogenous H2A.Z isoforms should be detected as a 14 kDa band on the Western blot, whereas the Strep-tagged form should have an apparent molecular weight of 15 kDa. The ubiquitylated form of H2A.Z or H2A.Z-Strep should be detected as bands at 22 kDa or 23 kDa respectively. Instead of transferring the SDS gel with the separated proteins, some gels were stained in Coomasie B stain for an hour at RT for an hour on a shaker and de-stained overnight to use the four histone bands as a loading control.

2.2.4 Fluorescence-activated cell sorting (FACS) analysis

Cell pellet was obtained after trypsinization of the cells from the cell culture plate and was washed in 1 mL of Phosphate-Buffered Saline (PBS) before centrifuging it at 5000 rpm for 2-5 minutes. PBS was then removed and the pellet was re-suspended with 300 mL of PBS and adding 700 mL of -20℃ 100% high grade ethanol on ice for 30 minutes. The resulting samples were then stored in -20℃ until FACS analysis was conducted. On the day of FACS analysis, the samples were centrifuged at 1000 rpm for 5 minutes before removing the supernatant. The pellet for each sample was re-suspended in 10 mL of cold PBS and centrifuged at 1000 rpm for 5 minutes before removing the PBS. The pellet was then re-suspended in 10 mL of staining buffer and centrifuged at 1000 rpm for 5 minutes before aspirating the buffer. The resulting pellet

was again re-suspending in 1 mL staining buffer with RNaseA at final concentration of 50ug/ml and propidium iodide at final concentration of 50 ug/mL. The samples were then left to incubate in the dark at RT for an hour before conducting FACS analysis in the machine.

2.2.5 Reverse transcriptase Quantitative Polymerase Chain Reaction (RT-qPCR)

Scrambled, 3 H2A.Z1-Strep, and 3 H2A.Z2-Strep clones were seeded in 60 mm plates and RNA was harvested 24 hours after according to the QiaGEN RNeasy kit. Once RNA was extracted from the sample, their concentration was determined and 2 ug was treated with DNAse1 at 37℃ for 10 minutes. The reaction was terminated with 5 mM EDTA and heated at 75℃ for 10 minutes. First-strand cDNA was synthesized initially by addition of 50 uM oligo(dT) and 10 mM dNTP before incubation at 65℃ for 10 minutes. The samples were then snap-chilled on ice for 5 minutes. First-strand buffer, 0.1M DTT, and ribonuclease inhibitor were added to the samples and incubated at 50℃ for 2 minutes. SuperScript III Reverse Transcriptase was added and incubated at 50℃ for 50 minutes. To inactivate the reaction, the samples were heated at 70℃ for 15 minutes. At the end of the reaction, cDNA was stored at -20℃ until use.

qPCR was conducted in triplicate with PerfeCta SYBR Green SuperMix (Quanta Biosciences) and primers specific to the transcript. Gene expression was normalized with the housekeeping gene Cytochrome C.

2.2.6 Chromatin immunoprecipitation

Cells from two 15-cm cell culture dishes were collected by trypsinization from the CRISPR-Cas9 clones; Scrambled, H2A.Z1-Strep #2, H2A.Z2-Strep #5, H2A.Z2-Strep #13, and H2A.Z2-Strep #20. Cells were counted to ensure equal amounts were used for protein extraction. They were then washed with 10 mL PBS + protease inhibitors (PI) and fixed by incubating with 1% formaldehyde in PBS at RT for 30 minutes. Formaldehyde treatment was quenched with 125 mM glycine and incubated at RT for 5 minutes. The fixed cells were then washed 3x in cold PBS + PI and incubated in nuclei isolation buffer (50 mM HEPES pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 140 mM NaCl, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100) + PI for 10 minutes at 4℃ on a nutator. Nuclei were pelleted with centrifugation and washed with wash buffer (10 mM HEPES pH 8.0, 1 mM EDTA, 05 mM EGTA, 200 mM NaCl) at 4℃. Pelleted nuclei were resuspended in sonication buffer (50 mM HEPES pH 8, 1 mM EDTA, 140 mM NaCl, 1% Triton-X100, 0.1% SDS, 0.5% SDC) + PI and nutated at 4℃ for 10 minutes. Sonication was conducted at 90% duty cycle and 40% output with 8 pulses per cycle and for a total of 8 cycles. Samples were chilled on ice 1 minute between each cycle. The samples were then centrifuged at max speed at 4℃ for 10 minutes and the supernatants were transferred to clean microcentrifuge tubes. To ensure that the amount of input DNA were approximately equal for IP, a small aliquot of 50 uL was taken from each sample and the remaining was stored at -20℃. To the small aliquot, 250 uL of Elution Buffer (2% SDS, 10 mM DTT, 0.1M Sodium Bicarbonate) was used to top up the volume and 5M NaCl was added before incubation at 65℃ overnight to reverse cross-linking. They

were then treated with RNAse A and Protinase K, followed by DNA extraction with standard phenol-chloroform protocol and lyophilized in a speed-vac before re-dissolving the sample in 50 uL molecular biology-grade water. The DNA concentration from the samples were then determined, and 5 ug of chromatin from the remaining sample was used for IP. The remaining supernatant were then used pre-cleared with Protein G Dynabeads (Invitrogen) for 2 hours at room temperature on a nutator. The samples were then incubated with following antibodies overnight at 4℃: Strep-tag, AM H2A.Z, and H3. Chromatin-antibody complexes were then pulled down with new Protein G Dynabeads for 2 hours at room temperature on a nutator and washed sequentially in the following buffers containing PI: RIPA (150 mM NaCl, 50 mM Tris pH 8, 0.1% SDS, 0.5% SDC), High Salt (500 mM NaCl, 50 mM Tris pH 8, 0.1% SDS), LiCl buffer (250 mM LiCl, 50 mM Tris pH 8, 0.5% SDC), and 2 washes with TE buffer (10 mM Tris pH 8, 1 mM EDTA). Chromatin were then eluded from the beads with Elution buffer and DNA precipitated as described above. An aliquot of 5 uL for each sample was used in triplicates for qPCR analysis described previously with primer pairs targeting the promoters of genes known to be enriched by H2A.Z-containing nucleosomes. The gene used to normalize ChIP elute samples was the (-)2kb PSA gene that is not enriched by H2A.Z. The following primers were then used for ChIP analyses: MTA1 promoter region 250-350 bp (MTA1 250-350) *forward 5' cctcactgcggaaagagcac 3', reverse 5' agaggccacctccaaagacc 3'*; MTA promoter region 650-700 bp (MTA1 650-700) *forward 5' gaaatggcccttgtggaggat 3', reverse 5' caccagaacccccctggtagat 3'*; TRIM21 *forward 5' aaactcagtagcccgtggtc 3', reverse 5' agcgtctaggtgtggagtga 3'*. ChIP data are presented

as enrichment to account for any differences in nucleosome density between samples, which were calculated by dividing % input of the IPs by the % input of the total H3 IPs.

## 2.3 Results

### 2.3.1 Increased HDR efficiency with both i53 and Nocodazole

Cas9 endonuclease induces double-stranded breaks (DSBs), leading to DNA damage and inducing DNA repair mechanisms; NHEJ or HDR. For efficient insertion of Strep-tag into the genome, we require a high proportion of cells to undergo HDR. Nocodazole is a drug that arrests cells at $G_2$/M, a phase of the cell cycle that is known to have a higher HDR rate in response to DNA damage (Gutschner *et al* 2016). Additionally, i53 is a reagent developed to inhibit the 53BP1 protein that normally promotes NHEJ, and enhance HDR efficiencies instead. To examine the effects of i53 and Nocodazole on HDR efficiency, 293T cells were transfected with the PX459 plasmid, encoding both Flag-tagged Cas9 and the sgRNA targeting either H2A.Z1 or Scrambled sequences, and the repair template with the H2A.Z1-Strep sequence. The Scrambled sgRNA, which would not target Cas9 to the H2A.Z1 allele, was used as a negative control for Cas9-mediated modification of the H2A.Z1 gene. Genomic DNA was extracted and PCR was used to amplify the region around the site of Strep-tag insertion at H2A.Z1 gene. The amount of PCR products corresponding to H2A.Z1 sequence with the Strep-tag insertion was used as an approximate measurement of HDR efficiency (the relative portion of transfected cells having the Cas9-directed insertion of the Strep-tag sequence). If majority of the cells underwent the NHEJ pathway to repair the DNA damage, we would observe any large insertion into the H2A.Z alleles since most break ends are simply re-ligated together. PCR products that do not contain Strep-tag insertion would show a smaller band at 508 bp, whereas the

presence of Strep-tag insertion would produce a larger 571 bp PCR product. The intensity of the band roughly suggests the abundance of the PCR product; the higher the intensity of the band, the more abundant it is. Without drug treatment, we observed 2 bands at 508 bp and 571 bp in the Cas9-H2A.Z1-sgRNA + repair template sample (henceforth named as H2A.Z-sgRNA), indicating the presence of both Strep-tagged H2A.Z1 and unmodified H2A.Z1 alleles (Figure 2-4). The 571 bp was presumed to correspond to the Strep-tag-containing PCR product since it was absent in the samples transfected with the Cas9-Scrambled-sgRNA + H2A.Z1 repair template control (henceforth named as Scrambled-sgRNA). When the H2A.Z1-sgRNA transfected cells were treated with either Nocodazole or i53, there was a small increase in the relative abundance of the 571 bp PCR product compared to the no-treatment samples, suggesting addition of either reagent slightly improved the efficiency of insertion of the Strep-tag sequences. However, we also saw the presence of the 571 bp band in the Scrambled-sgRNA-transfected sample treated with Nocodazole or i53. This was unexpected since the Scrambled-sgRNA should not target Cas9 to modify the H2A.Z1 gene. We do not have an explanation for this observation but the presence of the 571 bp band in these samples could be due to a small amount of PCR-template contamination. In the conditions where both Nocodazole and i53 were added, the sample from cells transfected with the Scrambled-sgRNA control again did not show any of the 571 bp product, whereas in the H2A.Z1 sgRNA-transfected sample there was a clear reduction of the 508 bp product (corresponding to the unmodified H2A.Z1 allele) and a concomitant increase in the 571 bp product (corresponding to the modified allele).

Although these results were qualitative and not perfect, they do suggest that the combination of both reagents helped improve the efficiency of Cas9-mediated modification of the H2A.Z1 allele.

To further confirm the 571 bp PCR product observed corresponds to alleles that have insertion of the Strep-tag DNA sequences, we also performed BamHI restriction enzyme digestion of the PCR samples. The Strep-tag insertion should also add three BamHI digestion sites (Figure 2-4, red GS) in the linker region between the H2A.Z.1 and Strep-tag DNA sequences (lower panel of Figure 2-4). BamHI digestion of the PCR samples from Scrambled-sgRNA-transfected cells yielded only one intact band corresponding to the expected 508 bp products whereas BamHI digestion of the PCR samples from H2A.Z1-sgRNA-transfected cells yielded a main 508 bp band plus a smaller size band. BamH1 digestion of the insert-containing (571 bp) PCR product is expected to yield two main fragments at 290 bp and 257 bps plus 2 very small fragments corresponding to sequences between the tandem BamHI sites. Although the agarose gel shown did not clearly resolve two distinct smaller size DNA bands, the presence of the smaller size band specifically in the H2A.Z1-sgRNA- but not in the Scrambled-sgRNA-samples suggests that the smaller band detected does correspond to the BamH1 digested 571 bp PCR product. In this analysis, we observed slightly more of the smaller size BamH1-digested bands in the Nocodazole/i53-treated samples compared to the no-treatment sample; however, there was no significant difference in

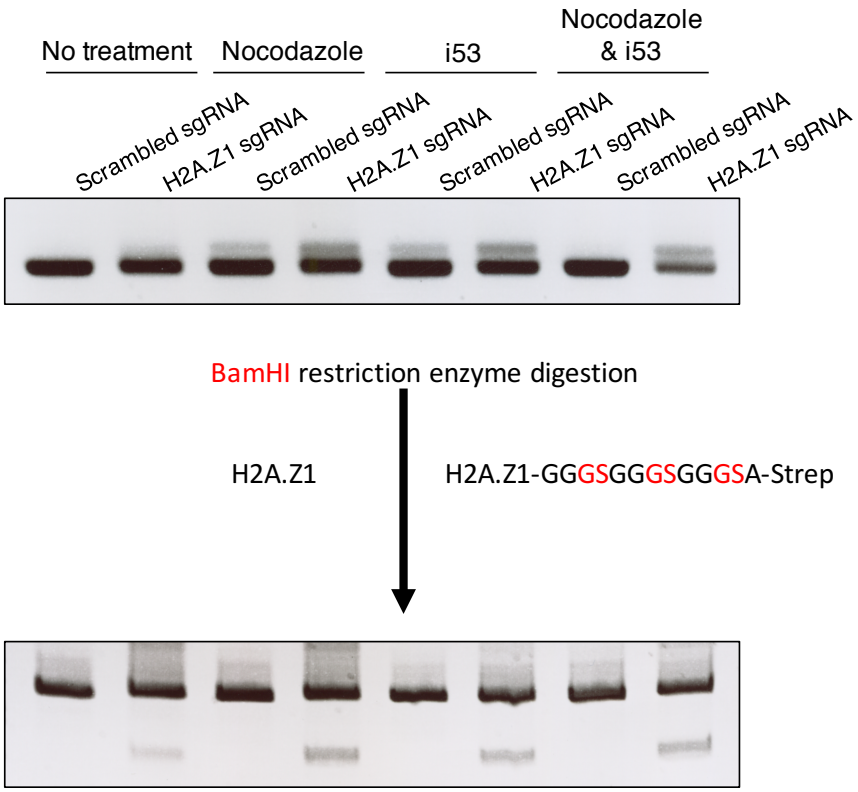



Figure 2-4. **Efficiency of Nocodazole and i53 in increasing HDR for CRISPR-Cas9 insertion of Strep-tag in C-terminus of endogenous H2A.Z1 using semi-quantitative PCR technique.** Scrambled clone was used as a negative control for Strep-tag insertion, which was used as a measurement for efficiency of HDR in the 293T cells. Nocodazole drug was administered to CRISPR-Cas9 transfected cells for 20 hours, while i53 was co-transfected along with the respective sgRNAs. Genomic DNA were extracted from the transfected cell population and PCR was conducted on the samples with a pair of primers that encompass part of the C-terminus end of endogenous H2A.Z1. Strep tag insertion at the C-terminus would increase the PCR product size, indicating by the bigger PCR product. Subsequently, restriction enzyme digestion was conducted with BamHI, whose recognition sites are only found in the PCR products including Strep-tag. There would be a total of 4 digestion products after cleavage, but two would be of tiny fragment size undetectable in the gel analysis.

the amount of the smaller product in the separate or combined Nocodazole/i53 treatments. Nevertheless, our results suggest that there may be some benefit using Nocodazole and i53, and therefore, we decided to include both treatments in our actual experimental attempts to generate cell lines containing Strep-tagged H2A.Z1/H2A.Z2 alleles.

### 2.3.2 Generation of Strep-tagged H2A.Z1 & H2A.Z2 clones in 293T cells

A modified CRISPR-Cas9 genome editing system was used to insert Strep-tag into the 3' end of the endogenous genes of H2A.Z1 or H2A.Z2 histone variant isoforms, H2AFZ and H2AFV (Figure 2-5). PX459 plasmid encoding the Flag-tagged Cas9 with sgRNA targeting Scrambled or the 3' end of either H2AFZ or H2AFV were transfected in low passaged 293T cells, along with the corresponding repair template and i53. The transfected cells were treated with Nocodazole for 20 hours and then grown in Puromycin for three days to select for cells that expressed the PX459 plasmid. Whole cell protein lysates were extracted initially from a portion of the cell population to check for expression of Cas9 in the cells after Puromycin selection (data not shown), and the remaining cells were used in clonal dilution and isolation of individual colonies. Single cell colonies were selected and expanded; a portion of each cell clone was cryopreserved whereas the rest of the cells were further tested to confirm Strep-tag insertion. Preliminary checks at the level of genomic DNA were done by PCR and at the protein level by Western blots to screen for colonies that indeed had Strep-tag insertion in H2A.Z1 or H2A.Z2. A few rounds of CRISPR-Cas9 transfections were conducted to

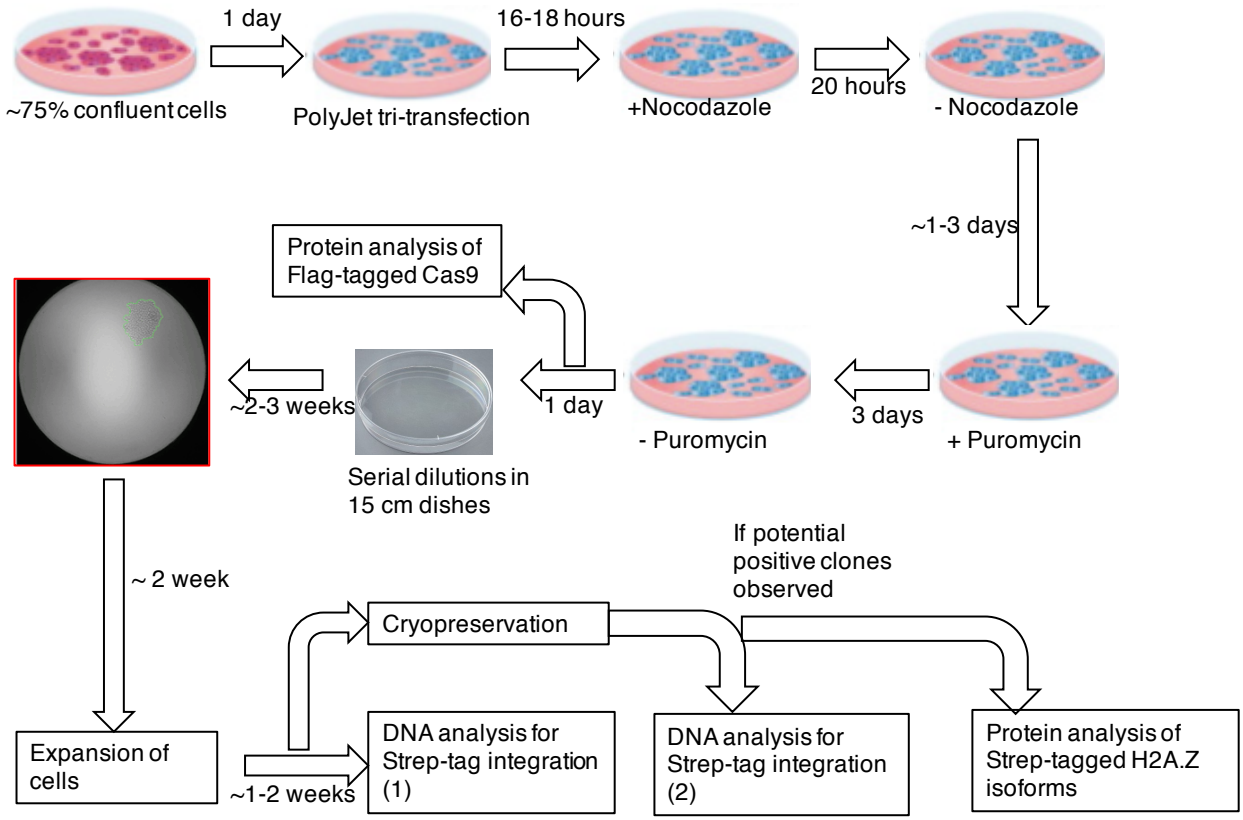



Figure 2-5. **Schematic of the CRISPR-Cas9 genome editing system to insert Strep-tag to 3' ends of endogenous H2AFZ AND H2AFV genes.** Cells were transfected with the plasmids expressing the SpCas9 endonuclease with the respective sgRNA and i53 protein, as well as the respective repair template in the form of single stranded oligonucleotides (ssODNs). Media was replaced 16-18-hours post-transfection and Nocodazole treatment was conducted on the cells for 20 hours. Fresh media was replaced to wash away Nocodazole and the cells were treated to puromycin selection for 3 days. To relieve cells of stress from the drug treatment, cells were incubated in fresh media for a day before conducting serial dilutions in 15 cm dishes. Proteins were also extracted to ensure expression of Cas9 endonuclease. Clonal isolation was conducted over 3 weeks and single cells were selected and transferred to 24-well plates and expanded over 1-2 weeks. Once there would be enough cells, these selected single clones would have their DNA extracted for confirmation of Strep-tag insertion at the correct site and the remaining cells would be cryopreserved. Once DNA analysis showed positive clones for Strep-tag insertion, these clones would be taken from cryopreservation and cultured for protein analysis of Strep-tagged H2A.Z isoforms by Strep-tag antibody.

obtain clones that contain insertion of Strep-tag sequences at the 3' end of the endogenous H2A.Z1 or H2A.Z2 genes. After a few failures in the puromycin selection step and the isolation of clones, we finally obtained viable clones that were expanded. A total of 54 clones were collected; 26 H2A.Z1-Strep clones, 23 H2A.Z2-Strep clones, and 5 Scrambled clones. Genomic DNA was extracted from these surviving clones that potentially have either H2A.Z1-Strep or H2A.Z2-Strep endogenous genes. PCR was conducted with primer pairs that amplified the 3' end of either H2A.Z1 or H2A.Z2 and restriction enzyme digestion was conducted with BamHI to cleave sites only found in the inserted DNA sequence (Figure 2-6A). Before cleavage, we can clearly observe the band at 508 bp for unmodified H2A.Z1 in the negative controls, Scrambled and 293T, as well as in H2A.Z1-Strep #2, and a slightly larger band at 571 bp was observed in all 3 H2A.Z1 clones (Figure 2-6B). H2A.Z2 clones all showed the slightly larger band at 493 bp indicating Strep-tag insertion, and the 430 bp band for unmodified H2A.Z2 was observed in both the negative controls and H2A.Z2-Strep #5 (Figure 2-6C). The presence of the two bands suggested heterogeneity of both unmodified and modified H2A.Z1 and H2A.Z2 alleles respectively, indicating these two clones might be heterozygous for Strep-tag insertion. After cleavage, we see that there were 3 positive clones for H2A.Z1-Strep, clone #2, #3, #5, that showed cleavage products at 290 and 257 bp (Figure 2-6D). In addition, there is also a band at 508 bp in H2A.Z1-Strep #2 clone aligned to the band observed with the Scrambled-sgRNA negative control, which suggested a heterozygous nature of the insertion. BamHI restriction digest in H2A.Z2-Strep clones showed cleaved PCR products at 260 bp and 209 bp in 4 of the clones;

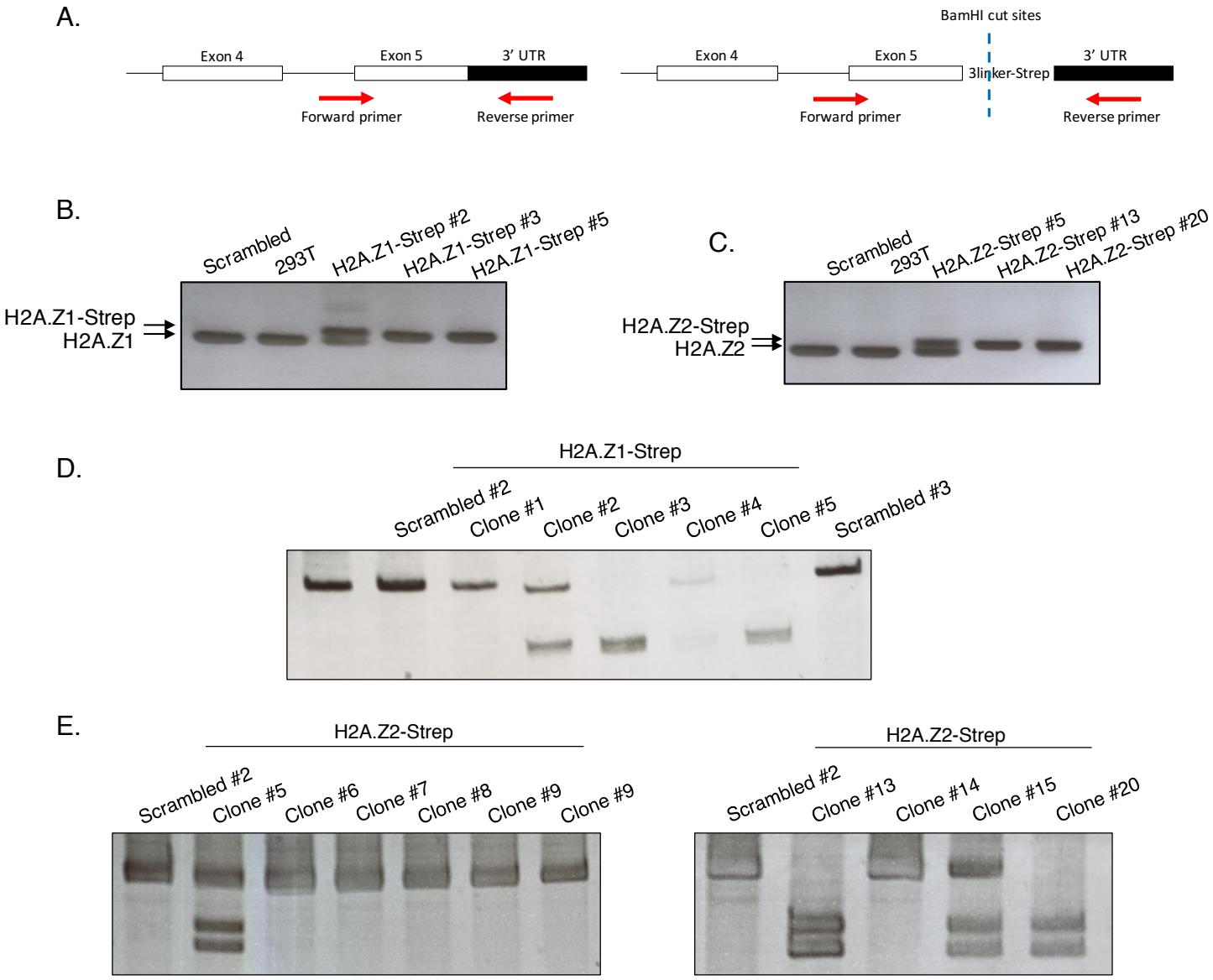
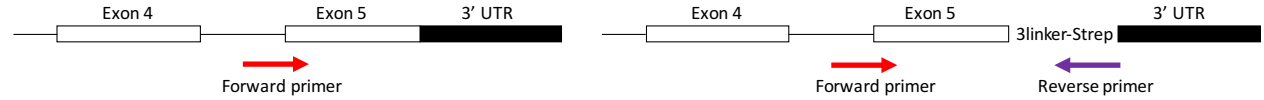


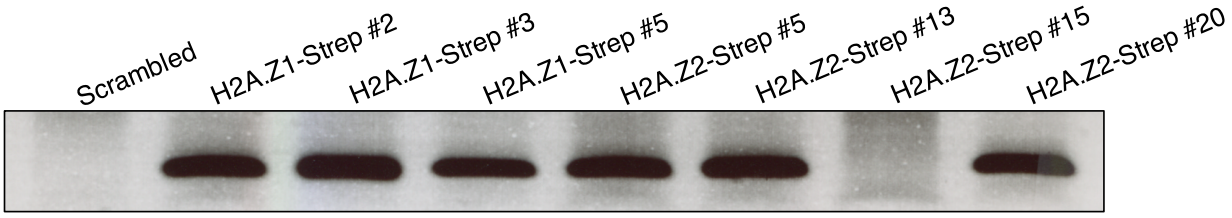

Figure 2-6. **Confirmation of Strep-tag insertion at the C-terminus of endogenous H2A.Z1 or H2A.Z2 using PCR and restriction enzyme digestion.** A) Genomic DNA were extracted from the obtained viable clones and PCR was conducted with primer pairs targeting similar regions in H2A.Z1 for potential H2A.Z1-Strep clones and in H2A.Z2 for potential H2A.Z2-Strep clones. Scrambled clones were used as negative controls for Strep-tag insertion as well as for unmodified H2A.Z isoforms. B) PCR was conducted with H2A.Z1-targeting primers. Strep-tag inserted clones showed a bigger band at 571 kDa in comparison to Scrambled and 293T at 508 kDa. C) Restriction enzyme digestion with BamHI showed 3 clones #2, #3, and #5, containing Strep-tag insertion as there were products of cleavage at 290 kDa and 257 kDa. H2A.Z1-Strep clone #2 contains an additional band at 508 kDa after cleavage. D) PCR was conducted with H2A.Z2-targeting primers. Strep-tag inserted clones showed a bigger band at 493 kDa in comparison to the negative controls, Scrambled and 293T, at 430 kDa. E) Restriction enzyme digestion with BamHI also showed 3 clones #5, #13, #15, and #20 had Strep-tag insertion to H2A.Z2 with cleavage products showing at 260 kDa and 209 kDa. H2A.Z2-Strep clone #5 and #15 contains an additional band at 508 kDa after cleavage.

#5, #13, #15, and #20, suggesting Strep-tag insertion (Figure 2-6E). The lack of the larger band representing unmodified H2A.Z1 or H2A.Z2 in both H2A.Z1-Strep #3 and #5 as well as in H2A.Z2-Strep #13 and #20 supported the homogeneity of modified H2A.Z1 or H2A.Z2. We next wanted to confirm Strep-tag integration again by replacing the reverse primer with one that recognizes a site directly on the Strep-tag (Figure 2-7A). The negative control, Scrambled, showed an absence of band as the reverse primer was unable to anneal for amplification (Figure 2-7B). All H2A.Z1-Strep clones and H2A.Z2-Strep clones showed a band at 314 bp and 266 bp respectively indicating Strep-tag presence, except for H2A.Z2-Strep #15 clone. Moving forward, we decided not to use H2A.Z2-Strep #15.

To confirm Strep-tag insertion at the level of protein expression, whole cell lysates were used for Western blot analysis using a commercial antibody against the Strep-tag. Lysates from cells transfected with a Strep-tagged H2A.Z-expression construct were included as a positive control for Western blot detection in the initial round of protein analysis (Figure 2-8, represented by the +). Bands corresponding to the expected sizes of H2A.Z-Strep (~ 14 kDa) and ubiquitylated H2A.Z-Strep (~ 22 kDa) proteins were detected by the Strep-tag antibody in all the clones deemed positive for Strep-tag insertion by DNA analysis for both H2A.Z1 (Figure 2-8A) and H2A.Z2 (Figure 2-8B). Additionally, the levels for H2A.Z1-Strep clone #2 and H2A.Z2-Strep clone #5 were clearly lower than other positive clones within their own class, supporting the previous interpretation that those 2 clones are heterozygous for Strep-tag insertion.

A.



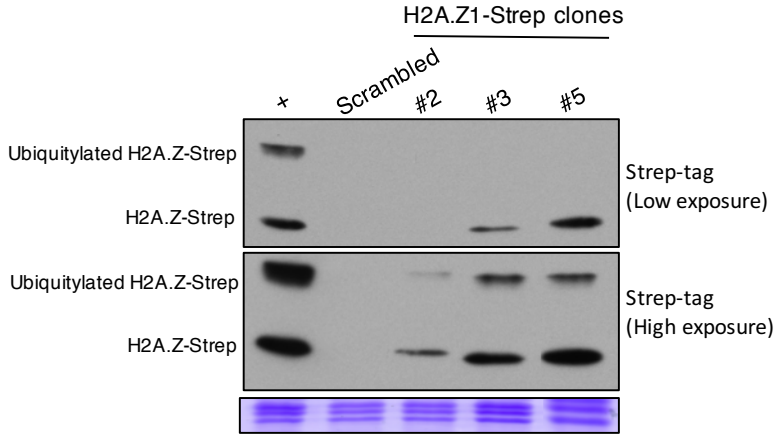

B.

Figure 2-7. **Confirmation of Strep-tag insertion at the C-terminus of endogenous H2A.Z1 or H2A.Z2 using PCR with reverse primer directly against the Strep-tag.** (A) Genomic DNA were extracted from chosen clones and PCR was conducted with either H2A.Z1 or H2A.Z2 forward primer and a reverse primer targeting the Strep-tag sequence directly. (B) Should Strep-tag not be inserted at the C-terminus of H2AFZ OR H2AFV in the respective clones, a PCR product would not be observed as the reverse primer would be unable to recognize any sequence and anneal to the genome for amplification. Additionally, Scrambled acts as the negative control as there should not be any Strep-tag inserted. A band was observed in all the H2A.Z1-Strep and H2A.Z2-Strep clones except in H2A.Z2-Strep #15. The specificity of these band was validated when there was no band observed in Scrambled.

Table 2-1. **Summary of the Strep-tagged H2A.Z1 and H2A.Z2 clones obtained**. A total of 3 clones each were obtained for both isoforms. Each isoform has 1 heterozygous Strep-tag insertion and 2 homozygous.

| | **H2A.Z1-Strep clones** | **H2A.Z2-Strep clones** |
|---|---|---|
| **Heterozygous** | #2 | #5 |
| **Homozygous** | #3<br>#5 | #13<br>#20 |

A.




B.

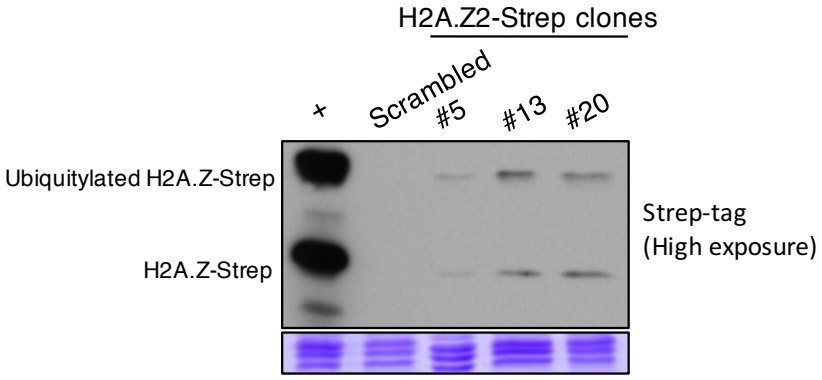



Figure 2-8. **Confirmation of Strep-tag insertion at the C-terminus of endogenous H2A.Z1 or H2A.Z2 using Western-blot analysis with Strep-tag antibody.** Western blot analyses of proteins from the whole-cell lysates from the identified Strep-tag positive clones with Strep-tag antibody. The positive control for H2A.Z-Strep was obtained from cells overexpressing the H2A.Z1-Strep protein. (A) H2A.Z1-Strep clones show presence of Strep-tag indicating H2A.Z1-Strep and ubiquitylated H2A.Z1-Strep. (B) H2A.Z2-Strep clones show presence of Strep-tag indicating H2A.Z2-Strep and ubiquitylated H2A.Z2-Strep, albeit at much lower levels.

Overall, we successfully obtained 6 Strep-tag inserted clones from the 48 clones selected. PCR and protein analyses suggested that H2A.Z1-Strep #2 and H2A.Z2-Strep #5 were clones with heterozygous insertion of the Strep-tag sequences where H2A.Z1-Strep #3, H2A.Z1-Strep #5, H2A.Z2-Strep #13, and H2A.Z2-Strep #20 were clones with homozygous insertions (Table 2-1).

Nuclear extract from all 6 positive clones as well as from the negative control (Scrambled-sgRNA-transfected) clones were obtained and analyzed again by Western blot with all the samples ran in parallel to compare their expression levels. H2A.Z2-Strep levels were detected by the Strep-tag antibody at much lower levels than H2A.Z1-Strep, which is consistent with the known lower RNA expression of H2A.Z2 compared to H2A.Z1 (Figure 2-9). Also as expected, the heterozygous clones for both H2A.Z1-Strep and H2A.Z2-Strep had approximately half the amount of the tagged histones (detected by the Western blots) compared to the homozygous clones. H2A.Z2-Strep clones had a much lower expression detected in the Strep-tag blots, indicating that H2A.Z2 seemed to be expressed at a much lower rate than H2A.Z1. Flag-tag antibody was also used to probe for the presence of Cas9 endonuclease, and H2A.Z1-Strep #3 was the only clone that still expressed Cas9. Since the transfected cells were selected with puromycin for only 3 days, it is possible that Cas9 expression was eventually lost in most of the clones after multiple cell passages without selection. The presence of Cas9 in H2A.Z1-Strep #3 might result in off-target cleaving and lead to potential genome mutations. Finally, when blotted with two different H2A.Z antibodies, expression of H2A.Z was found to greatly

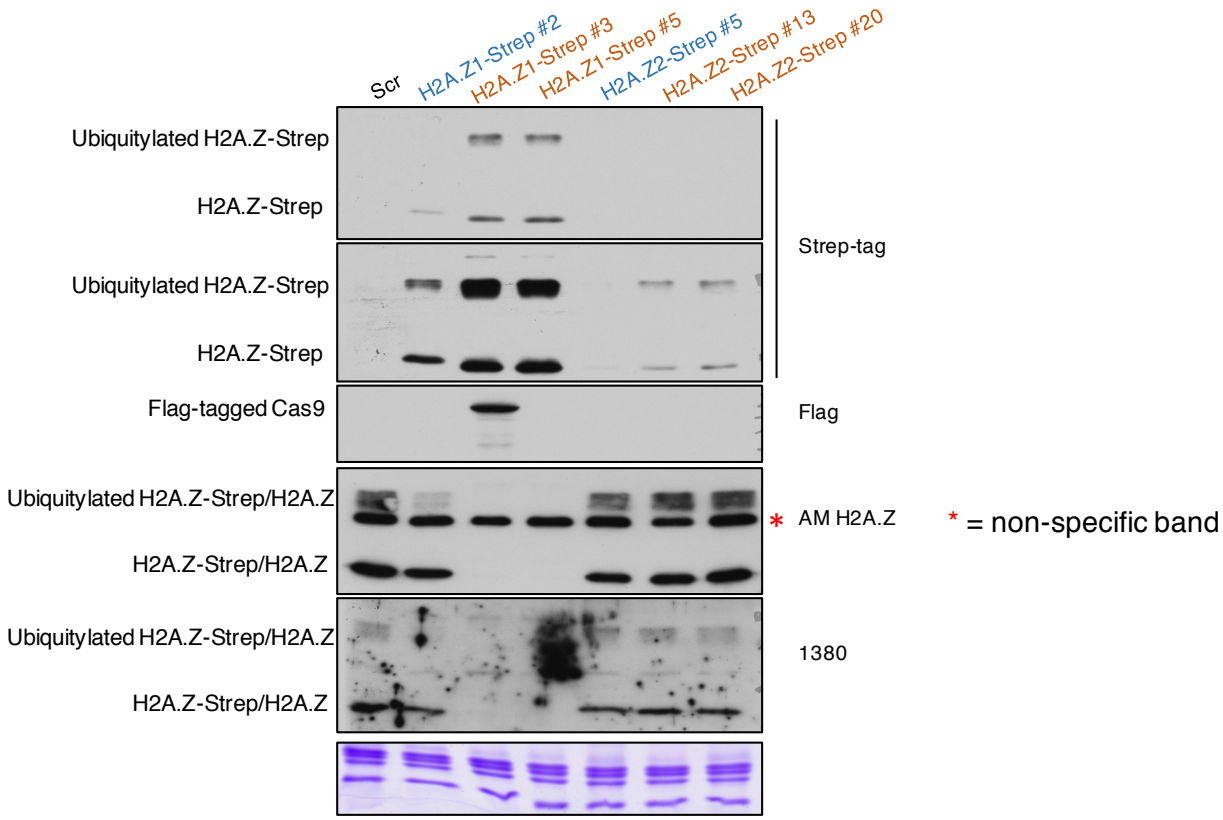



Figure 2-9. **Strep-tag is successfully inserted to endogenous genes of H2A.Z1 and H2A.Z2.** Western blot analysis of nuclear lysates from H2A.Z1-Strep and H2A.Z2-Strep clones, as well as the Scrambled clones. Coomasie was used as loading control and showed approximately equal loading of proteins. Strep-tag antibody was used to blot for the tag, while AM H2A.Z and 1380 antibodies were used to detect the H2A.Z1 proteins, both modified and unmodified. Flag antibody was used to detect the presence of Cas9 nuclease that is tagged with Flag. Heterozygous clones are color-coded in blue label while the homozygous clones are in orange.

reduced in the homozygous H2A.Z1-Strep clones compared to controls of H2A.Z2-Strep clones. Since the same observation was seen with two separate antibodies directed at different regions of the H2A.Z protein, this suggested the lack of detectable H2A.Z in the H2A.Z1-Strep clones was due to a knockdown of expression rather than an epitope occlusion (the inserted tag blocking recognition by the antibody) problem. Instead, it appears that the DNA sequence inserted right before the stop codon of the endogenous H2A.Z1 gene may have affected its expression at either the transcription, mRNA stability, or protein translation levels.

### 2.3.3. Lower endogenous H2A.Z2 protein levels in comparison to H2A.Z1 in 293T cells

The difference in the levels of H2A.Z1-Strep and H2A.Z2-Strep proteins as detected by Strep-tag antibody versus endogenous H2A.Z antibodies suggested that Strep-tag antibody is more sensitive at detecting the tagged proteins. However, endogenous H2A.Z antibodies are used for detecting the H2A.Z1 isoform. As H2A.Z1 isoform in H2A.Z2-Strep clones are unmodified, their levels were still detectable by the endogenous H2A.Z antibodies in the clones. Strep-tag antibody measures the levels of H2A.Z1 protein in H2A.Z1-Strep clones and H2A.Z2 protein in H2A.Z2-Strep clones. Data showed that H2A.Z2 seems to be expressed at a much lower level than H2A.Z1 as indicated by the intensity of the bands (Figure 2-9). We wanted to get an estimate difference in protein expression levels between the two histone variant isoforms in 293T cells. The heterozygous H2A.Z1-Strep #2 clone had comparable H2A.Z1 signal picked up by Strep-tag antibody that was not too intense when compared to H2A.Z2 levels in

H2A.Z2-Strep clones, and was hence used as a comparison. Furthermore, endogenous H2A.Z antibodies were still able to detect some levels of H2A.Z1 protein expression levels in the heterozygous H2A.Z1-Strep #2 clone. Nuclear extracts from H2A.Z1-Strep #2 and the H2A.Z2-Strep clones were obtained and analyzed with Western blot. Dilutions of H2A.Z1-Strep #2 were conducted by taking 1/3$^{th}$, 1/6$^{th}$. 1/9$^{th}$, and 1/12$^{th}$ of the sample and diluting it with 2x SB and protein loading dye to equalize the loading volume. We observed that H2A.Z1 levels at 1/9$^{th}$ or 1/12$^{th}$ showed similar band intensity to H2A.Z2 (Figure 2-10). This suggested that H2A.Z2 is expressed 1/18$^{th}$ or 1/24$^{th}$ lower than H2A.Z1 in 293T cells, as we are comparing heterozygous H2A.Z1-Strep levels to H2A.Z2-Strep levels and we have assumed that the homozygous clones have twice the number of modified alleles than heterozygous clones.

### 2.3.4 Characterization of H2A.Z1-Strep and H2A.Z2-Strep clones

Protein analysis showed a reduction in endogenous H2A.Z1-Strep levels. However, we do not know at which stage its gene expression was altered. To test whether insertion of the Strep-tag sequence affected the transcription rate of the H2A.Z1 and H2A.Z2 genes, we next used reverse transcriptase quantitative PCR (RT-qPCR) to measure the steady-state amount of H2A.Z1-Strep/H2A.Z2-Strep mRNA in the stable cell clones. Stable cells transfected with the Scrambled-sgRNA was also used as a negative control to test whether Cas9 expression affects H2A.Z1/2 transcription. Finally, normal 293T cells that haven't been through the CRISPR/Cas9 editing were used to determine the normal steady-state expression levels of the H2A.Z

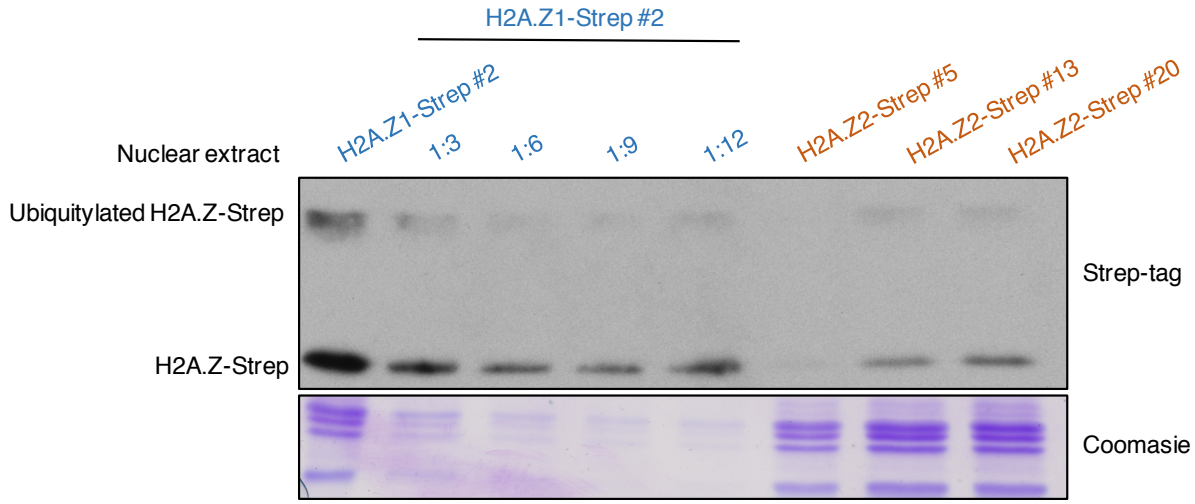
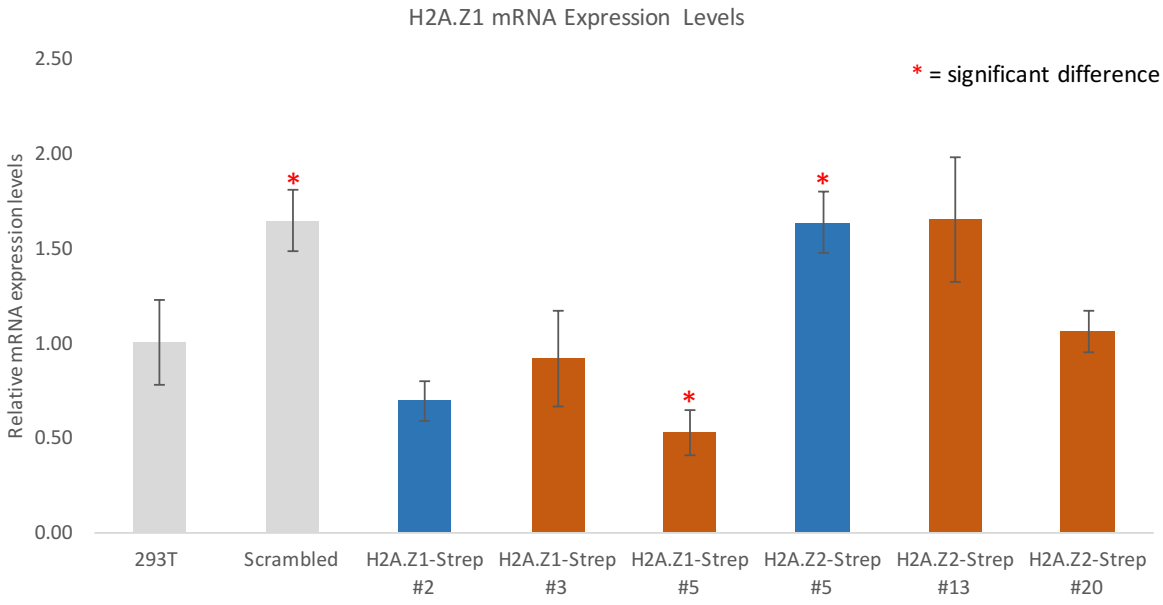



Figure 2-10. **Endogenous H2A.Z2 is expressed at a lower level compared to endogenous H2A.Z1.** Nuclear extracts from heterozygous H2A.Z1-Strep #2 clone and H2A.Z2-Strep clones were obtained and ran in a Western blot and Coomasie for loading control. Strep-tag antibody was used to detect Strep-tagged H2A.Z1 or H2A.Z2. Heterozygous H2A.Z1-Strep #2 was diluted accordingly and used as a comparison for H2A.Z1 protein levels versus H2A.Z2 protein levels. Heterozygous clones are color-coded in blue label while the homozygous clones are in orange.

variant genes. RNA was harvested from the various cell clones and H2A.Z1/2 mRNA was converted to cDNA for use in qPCR. H2A.Z1 expression levels in most of the Strep-tagged clones were comparable to 293T, with a few exceptions. Scrambled and H2A.Z2-Strep #5 seemed to have a relatively higher H2A.Z1 expression compared to 293T, while H2A.Z1-Strep #5 has a lower H2A.Z1 expression levels (Figure 2-11A). H2A.Z2 expression levels in all the Strep-tagged and Scrambled sg-RNA clones seemed to be similar to levels in 293T cells (Figure 2-11B). The detection of mRNA expression levels of H2A.Z1 and H2A.Z2 indicates that the genes were being transcribed; however, the data did show some reduction in homozygous H2A.Z1-Strep mRNA levels compared to H2A.Z1 expression in normal 293T cells. Additionally, we wanted to see if modifying the genome by insertion of Strep-tag would have disrupted the cell cycle progression of the clones, and FACS analysis was conducted. Most of the clones and Scrambled had the same trend as 293T, with most of the cell population at G1 phase, followed by the G2/M phase (Figure 2-12). This suggests that insertion of Strep-tag or modification of the genome by Cas9 did not affect cell cycle progression of the cells. Altogether, our data suggest that although the H2A.Z2-Strep clones seem to be comparable to the unmodified parent cells, the H2A.Z protein levels in the homozygous H2A.Z1-Strep clones were undetectable with endogenous H2A.Z antibodies and this effect could not be fully explained by transcription rate alone.

A.




B.

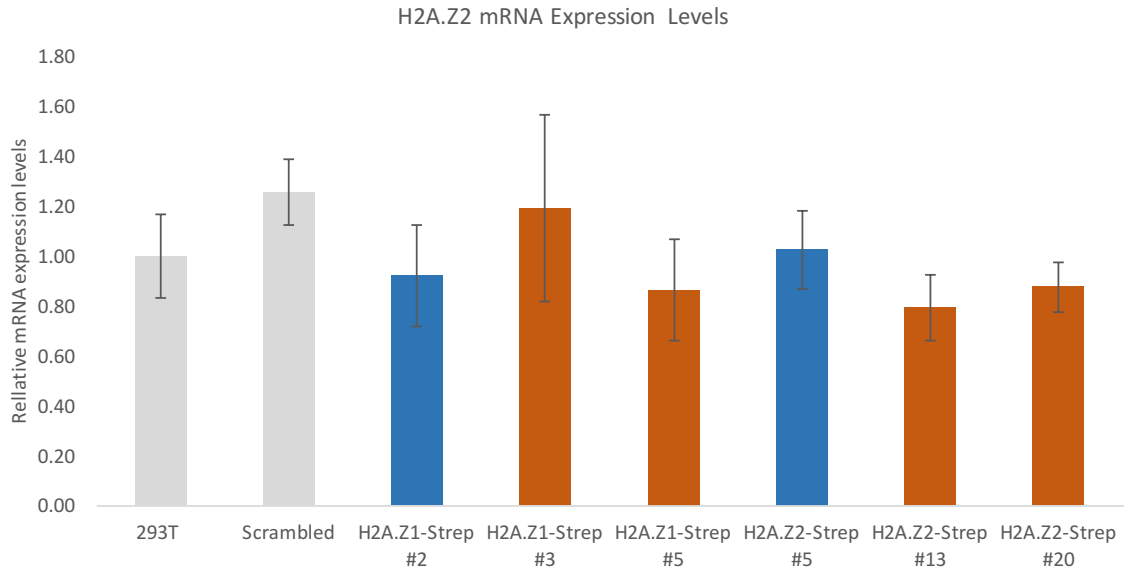



Figure 2-11. **Mostly normal levels of H2A.Z1 and H2A.Z2 mRNA expression in H2A.Z1-Strep and H2A.Z2-Strep clones.** 293T was both used as a control to compare expression of mRNA in Strep-tagged clones. RNA was extracted from cells and converted to cDNA with reverse transcriptase for use as template for quantitative PCR (qPCR). Each PCR reaction was performed in triplicates. Asterisks show significant difference between the sample and the control (293T). A) H2A.Z1 primer pair was used to detect relative H2A.Z1 expression levels. B) H2A.Z2 primer pair was used to detect relative H2A.Z2 expression levels. Heterozygous clones are color-coded in blue label while the homozygous clones are in orange.

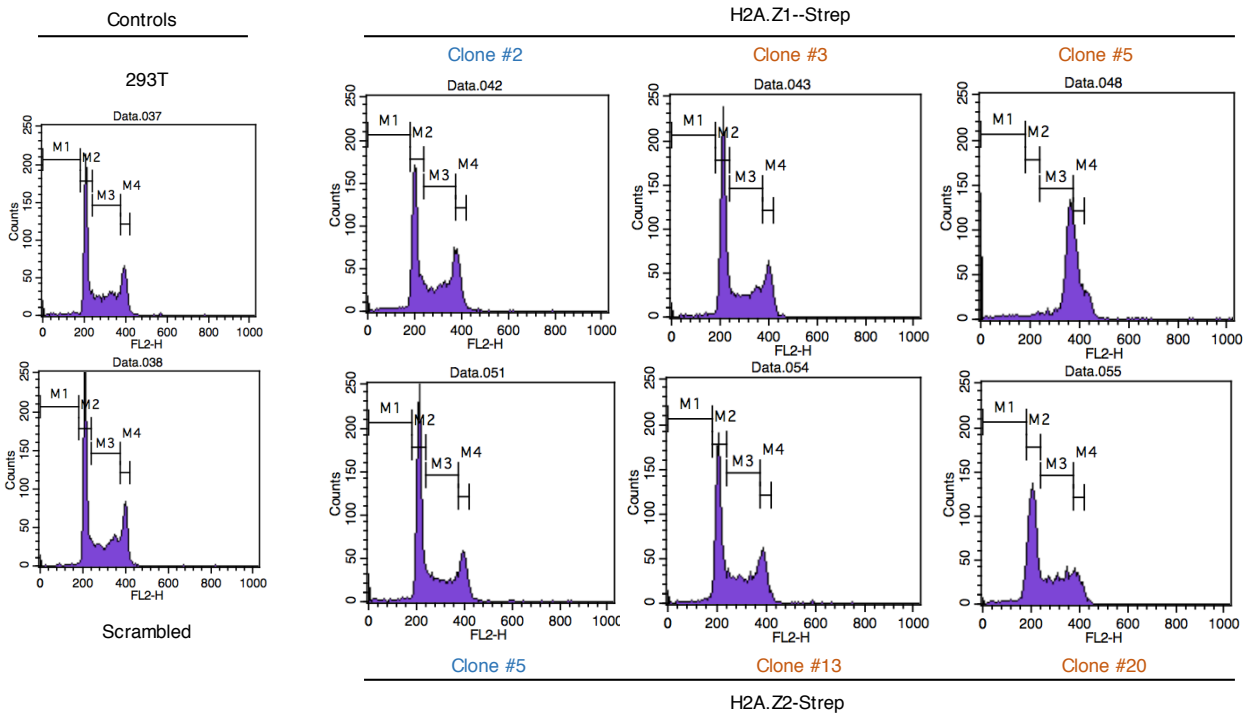



Figure 2-12. **Characterization of cell cycle phases of CRISPR-Cas9 modified H2A.Z1-Strep and H2A.Z2-Strep clones by FACs analysis.** H2A.Z1-Strep and H2A.Z2-Strep clones cells were fixed and stained for analysis in flow cytometry. 293T and Scrambled were used as comparison for normal cell cycle phases, with the latter as a control for Cas9-modified cells. M1 represents pre-G1 phase, M2 is the G1 phase, M3 is S phase, and M4 is the G2/M phase. Heterozygous clones are color-coded in blue label while the homozygous clones are in orange.

2.3.5 Minimal off-target effects using CRISPR-Cas9 for insertion of Strep-tag into C-terminus of H2A.Z histone variant isoforms in 293T cells

We wanted to ensure there weren't any single base mutation/deletion near the site of insertion in our clones, which might also explain the reduced levels of H2A.Z detected by the endogenous H2A.Z antibodies in homozygous H2A.Z1-Strep clones. Due to time constraints and a failed attempt to sequence the genomic DNA of the clones, we utilized an assay that would give us an overview of the presence of mutations. Surveyor mutagenesis assay was conducted in the 3 H2A.Z1-Strep clones and 3 H2A.Z2-Strep clones with Scrambled-sgRNA as a control (Figure 2-13). 293T cells was used as our reference sample as it is the parent cell line of our clones and should mimic unmodified H2A.Z1 or H2A.Z2 gene sequences. DNA was extracted from our reference sample and both the H2A.Z1-Strep clones and H2A.Z2-clones. PCR was conducted with the same primer pairs as described above in Figure 2-7A for both H2A.Z1 and H2A.Z2 gene sequences. The PCR products from H2A.Z1-Strep or H2A.Z2-Strep clones were then denatured and annealed with the H2A.Z1 or H2A.Z2 PCR product from 293T to form hetero/homo duplexes. Scrambled, used as a negative control, should not show any cleavage bands for both H2A.Z1 and H2A.Z2 when hybridized with the reference as the locus should not be targeted by Cas9 and therefore won't be modified (i.e. should have identical DNA sequence). Presence of Strep-tag would yield heteroduplexes containing a mismatch at the site of Strep-tag insertion to show 2 cleavage bands while absence of Strep-tag would produce homoduplexes and not yield any cleavage bands (Figure 2-13). Additionally, any mutations resulting from

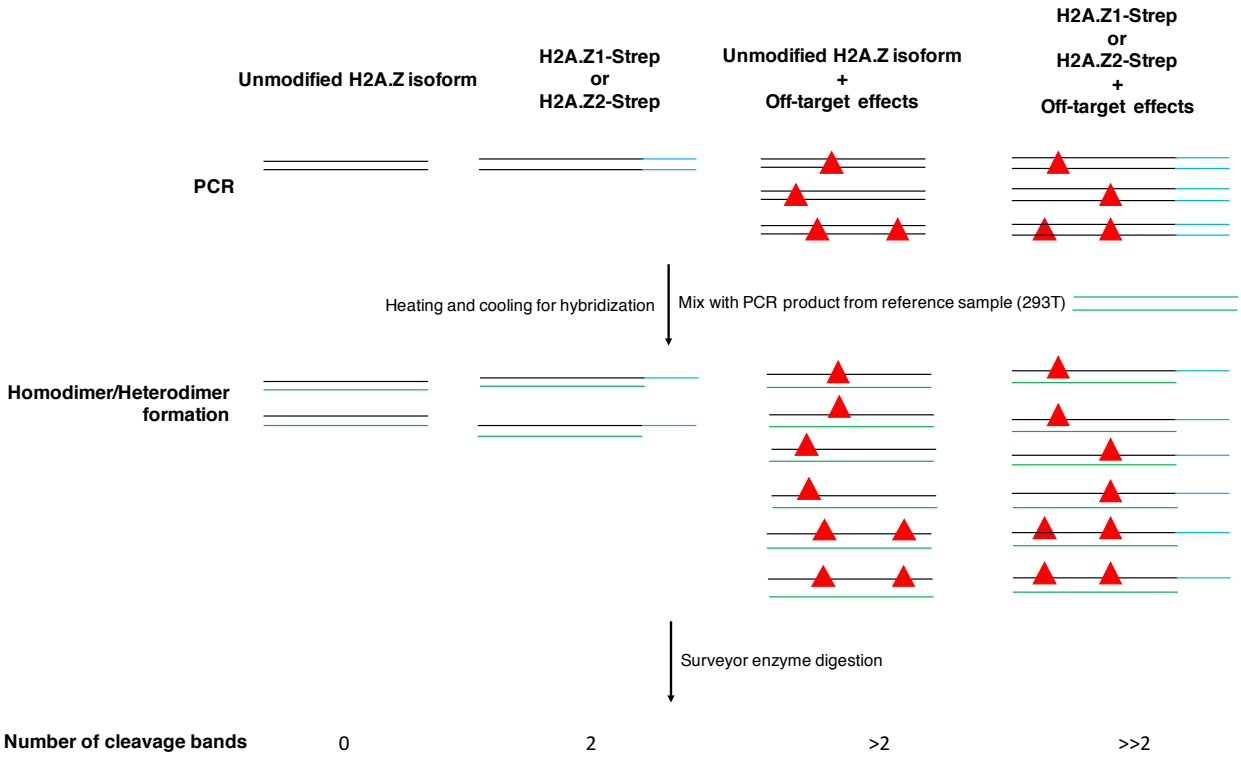
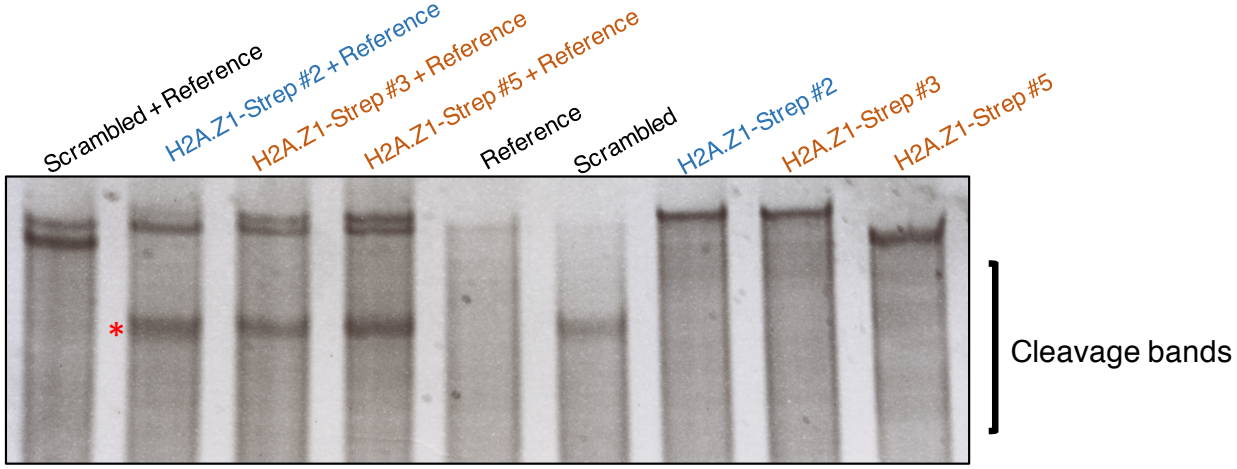



Figure 2-13. **Concept of Surveyor mutagenesis assay in determining off-target effects by CRISPR-Cas9 genome editing system when inserting Strep-tag into C-terminus of H2A.Z histone variant isoforms.** Genomic DNA was used in PCR to amplify the H2A.Z1 region for H2A.Z1-Strep clones and H2A.Z2 region for H2A.Z2-Strep clones. Scrambled and 293T cells were also used to amplify H2A.Z1 or H2A.Z2 region as negative control for Strep-tag insertion and as reference sample respectively. After amplification, the PCR products from the Strep-tag clones were incubated with the PCR product from the reference gene (H2A.Z1 or H2A.Z2 in 293T) for hybridization; H2A.Z1-Strep clones were incubated with H2A.Z1 gene while H2A.Z2-Strep clones were incubated with H2A.Z2 gene. After Surveyor nuclease cleavage, there would be 2 bands observed if Strep-tag was inserted, and more than 2 bands if there were any off-target effects resulting in mutations. If Strep-tag was not inserted and no off-target effects, there would not be a cleavage band.

Cas9 off-target effects within the PCR product would also result in heteroduplex formation with PCR product from the reference sample. However, this would yield different sized bands or more than 2 bands as mutations are random and would occur indiscriminately, potentially leading to more sites of mismatches for Surveyor nuclease to cut. Surveyor nuclease would then cleave the heteroduplexes, and the digested DNA products were analyzed with gel electrophoresis. For H2A.Z1-Strep clones, we observed a thick cleavage band at similar size across samples that could be 2 unresolved bands absent in the negative control (Figure 2-14A, red asterisk), suggesting Strep-tag insertion to H2A.Z1 alleles. The PCR products from the clones without the reference sample were also ran on a gel alongside the samples hybridized with the reference to ensure that the cleavage bands observed were specific to the formation of heteroduplexes. The absence of any cleavage bands in the H2A.Z1-Strep clones without the reference sample supported the specificity of the cleavage bands observed earlier with the reference sample. Two cleavage bands were observed in all three H2A.Z2-Strep clones after Surveyor mutagenesis assay that were absent in the negative control (Figure 2-14B, red asterisk). However, in the control assay without the hybridization with the gene from reference sample, the same two bands were observed in H2A.Z2-Strep #5 and H2A.Z2-Strep #20. This suggests that the two cleavage bands observed with hybridization might not have been very specific to heteroduplex formation from Strep-tag insertion, and there could be slight mutations. This assay was conducted only once and should be replicated again to ensure reliability of results.

A.




B.

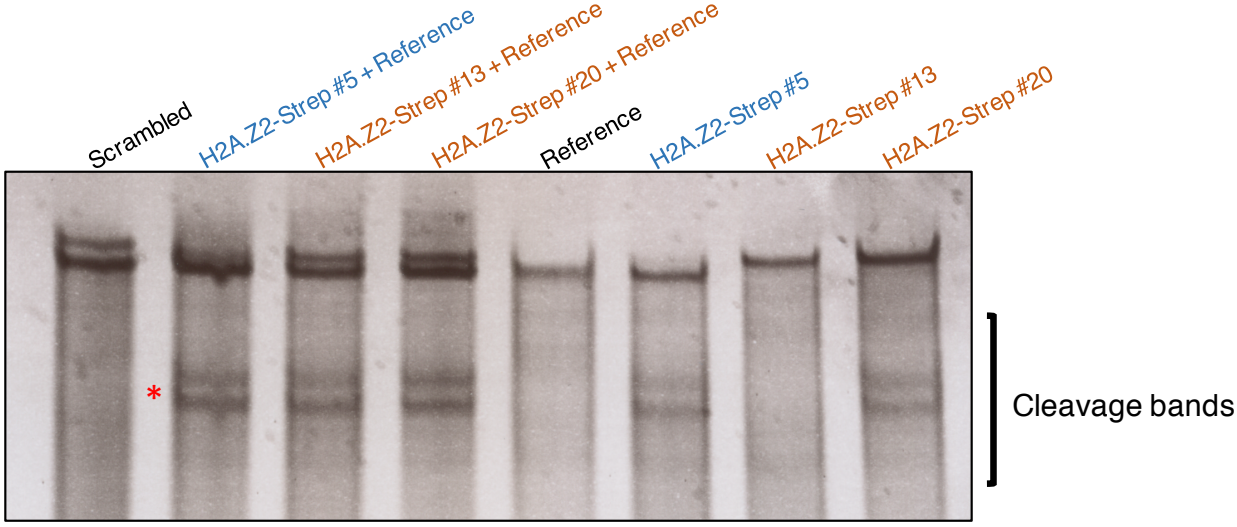



Figure 2-14. **No significant mutations observed in H2A.Z1-Strep and H2A.Z2-Strep clones**. Surveyor mutagenesis assay was carried out as described in the previous figure on the 3 H2A.Z1-Strep clones and 3 H2A.Z2-Strep clones, with Scrambled as a control to assure any digested bands observed were not random. 293T was used as the reference sample. Red asterisk was used to represent the cleavage bands not observed in Scrambled control. A) H2A.Z1 PCR product from the reference sample was hybridized with H2A.Z1 PCR products from the clones. Cleavage products by Surveyor enzyme was run on 10% polyacrylamide gel. B) H2A.Z2 PCR product from the reference sample was hybridized with H2A.Z2-PCR products from the clones. Cleavage products by Surveyor enzyme was run on 10% polyacrylamide gel. Heterozygous clones are color-coded in blue label while the homozygous clones are in orange.

2.3.6 Comparison of localization of H2A.Z1-Strep-containing vs H2A.Z2-Strep-contaiing nucleosomes on specific DNA sequences *in vivo*

We next wanted to confirm if the Strep-tagged H2A.Z1 and H2A.Z2 clones are incorporated into chromatin and at genomic locations/genes where H2A.Z is known to be present (e.g. based on previous ChIP-Seq data). The homozygous H2A.Z1-Strep clones #3 and #5 were not used as the levels of H2A.Z1-Strep in these seemed to be extremely low as endogenous H2A.Z antibodies were unable to detect it. ChIP was conducted using Scrambled, H2A.Z1-Strep #2, and all 3 H2A.Z2-Strep clones to examine the localization of the Strep-tagged H2A.Z1 or H2A.Z2 at specific promoters of genes. Chromatin was immunoprecipitated using the Strep-tag antibody to isolate for chromatin enriched by either H2A.Z1-Strep or H2A.Z2-Strep. They were then probed with primer pairs that were previously designed to amplify the MTA1 gene and the TRIM21 gene. For the MTA1 gene, two sets of PCR primers were used corresponding to two different regions at different distances from the TSS (Figure 2-15A). These genes were formerly identified with RNA-sequencing to have changes in its expression after H2A.Z1 knock-down (previously conducted in our lab) and reported to have enriched H2A.Z by ChIP experiments. Scrambled-sgRNA sample pulled-down with Strep-tag antibody was used to mimic a no antibody control for Strep-tag ChIP as there is no Strep-tag present for the antibody to pull-down. H2A.Z1-Strep enrichment was found on the MTA1 gene at a more proximal region to its TSS, but this enrichment was not observed with the H2A.Z2 clones (Figure 2-15B). There was not any significant enrichment of either H2A.Z1-Strep or H2A.Z2-Strep on the distal region of MTA1 gene

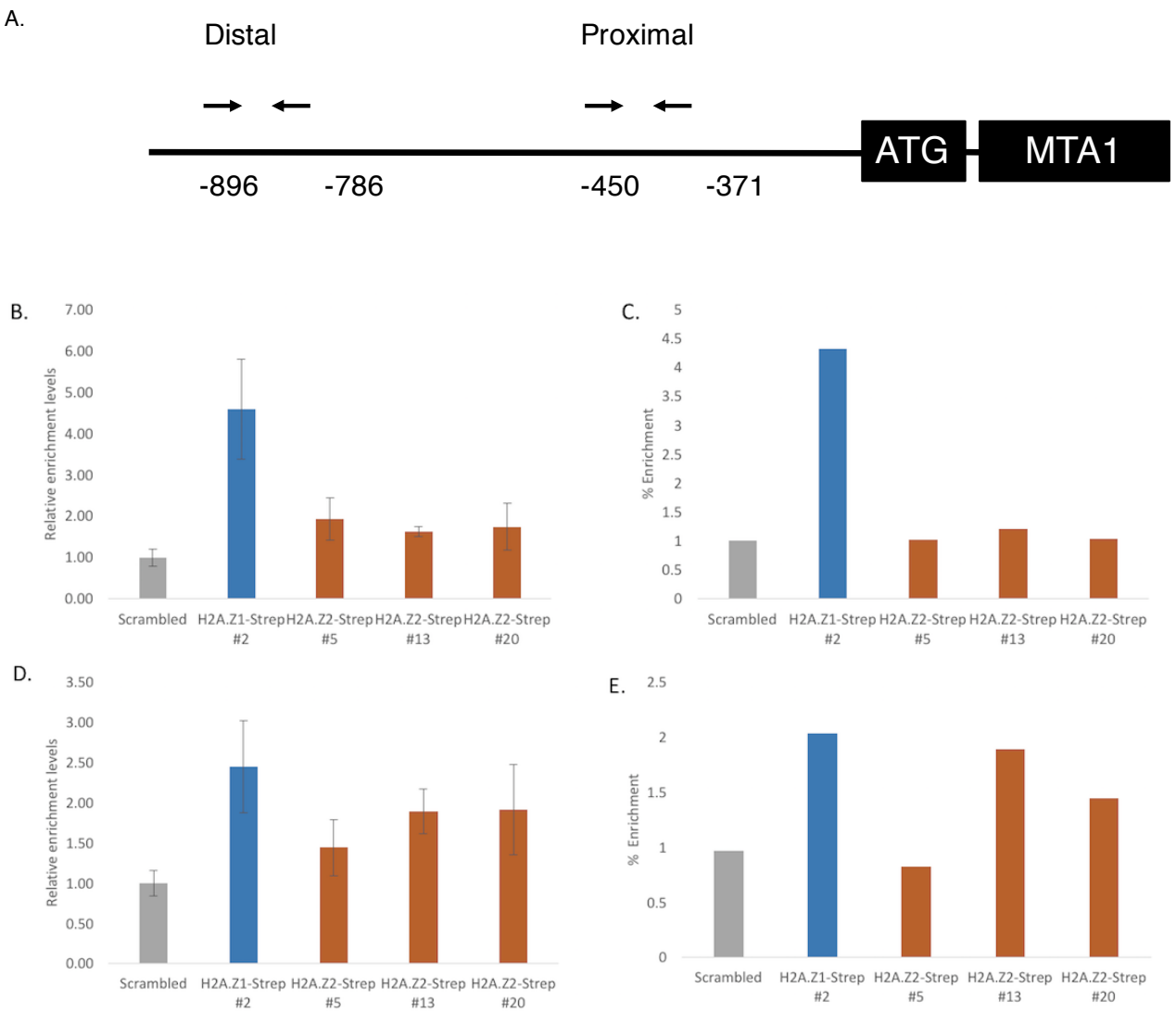



Figure 2-15. **H2A.Z1 is highly enriched on MTA1 at the proximal region to its TSS compared to H2A.Z2 and equal levels of enrichment between H2A.Z1 and H2A.Z2 on the distal region of MTA1 TSS.** Heterozygous H2A.Z1-Strep #2 and all three H2A.Z2-Strep clones were fixed in 1% formaldehyde and nuclear fraction material were extracted to use in chromatin immunoprecipitation (ChIP) with Strep-tag antibody. Scrambled is used to mimic a no-antibody control as there are no Strep-tag for the antibody to IP. A) The two primer pairs used in qPCR correspond to different regions of the MTA1 gene. B) ChIP analysis of H2A.Z1-Strep and H2A.Z2-Strep at the proximal region of MTA1 transcription start site (TSS). C) Data from A was divided by the data from H3 ChIP to account for any changes nucleosomal density across samples. D) ChIP analysis of H2A.Z1 and H2A.Z2 at the distal region of MTA1 TSS. E) Data from C was divided by the data from H3 ChIP to account for any changes nucleosomal density across samples. Heterozygous clones are color-coded in blue label while the homozygous clones are in orange.

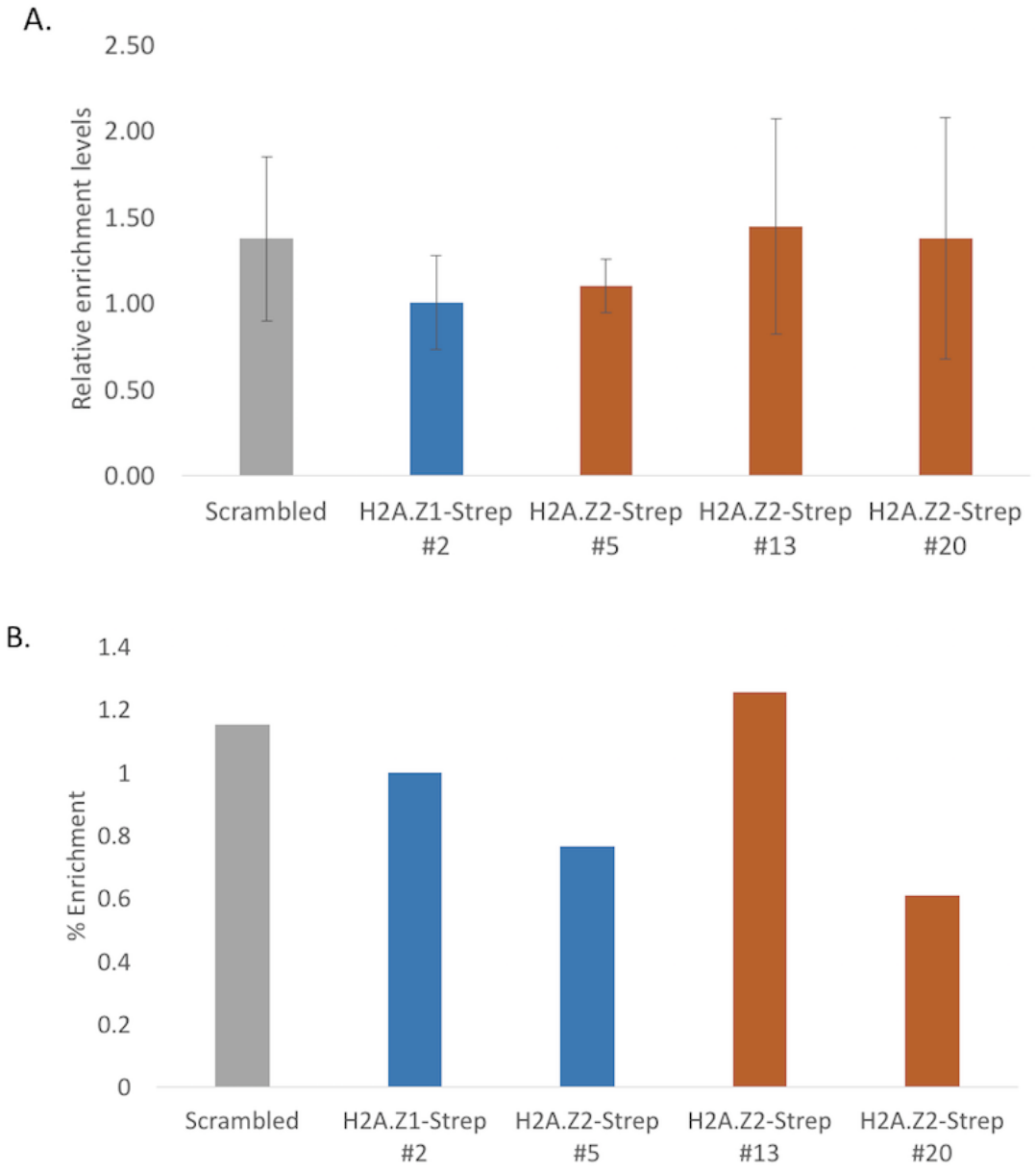



Figure 2-16. **No difference in localization of H2A.Z1 and H2A.Z2 at TRIM21 gene**. Heterozygous H2A.Z1-Strep #2 and all three H2A.Z2-Strep clones were fixed in 1% formaldehyde and nuclear fraction material were extracted to use in chromatin immunoprecipitation (ChIP) with Strep-tag antibody. Scrambled is used to mimic a no-antibody control as there are no Strep-tag for the antibody to IP. A) ChIP analysis of H2A.Z1 and H2A.Z2 at the TRIM21 gene. B) Data from C was divided by the data from H3 ChIP to account for any changes nucleosomal density across samples. Heterozygous clones are color-coded in blue label while the homozygous clones are in orange.

TSS (Figure 2-15D) nor at TRIM21 (Figure 2-16A). To ensure that the data we observed was not due to uneven nucleosome density between the samples, we've normalized the H2A.Z1-Strep/H2A.Z2-Strep enrichment over the H3 levels. After normalization of the data, H2A.Z1-Strep is still more enriched on the proximal region of the MTA1 gene TSS in comparison to H2A.Z2-Strep (Figure 2-15C). There weren't any significant changes in the localization between H2A.Z1-Strep and H2A.Z2-Strep on the distal site of the MTA1 gene TSS (Figure 2-15E) and on TRIM21 gene (Figure 2-16B). My data seemed to suggest that H2A.Z1-Strep and H2A.Z2-Strep may have different enrichment levels or patterns. However, we need to check/compare the H2A.Z1-Strep and H2A.Z2-Strep protein levels in the different clones to rule out differences of ChIP results was due to difference in H2A.Z1-Strep/H2A.Z2-Strep protein levels. Additionally, this was conducted only once and should be replicated again to ensure the data observed are consistent and accurate.

## 2.4 Discussion

The highly conserved H2A.Z variant of the core histone H2A has been implicated in transcription regulation. The two recently discovered isoforms, H2A.Z1 and H2A.Z2, were found to localize on promoters and regions flanking transcription start sites, consistent with their involvement in transcription activation. The goal of this project is to insert the Strep-tag DNA sequence in frame to the 3' terminus of endogenous H2A.Z1 or H2A.Z2 coding sequences using the CRISPR-Cas9 genome editing system. We aimed to address the problem of there being no commercial antibodies that can distinguish between the two histone variant isoforms. Fusion proteins that can be detected by antibodies against the tag can be expressed, and studies can be done to learn more about the distinct functions of the two isoforms. To increase the efficiency of the genome editing system, we wanted to increase efficiency of the HDR pathway in response to double-stranded breaks for better incorporation of Strep-tag in the repair templates. A drug and an inhibitor, Nocodazole and i53 respectively, were both used as it seemed to show a cleaner and more efficient Strep-tag insertion when used together rather than individually.

We generated 293T clones that contained Strep-tag in the 3' terminus of H2AFZ and H2AFV genes. For endogenous Strep-tagged H2A.Z1, H2A.Z1-Strep #2 is the heterozygous clone for Strep-tag insertion while H2A.Z1-Strep #3 and #5 are homozygous for the insertions. For endogenous Strep-tagged H2A.Z2, H2A.Z2-Strep #5 clone is heterozygous while H2A.Z2-Strep #13 and #20 are homozygous for Strep-tag

insertion. These positive H2A.Z1-Strep and H2A.Z2-Strep clones were confirmed at the DNA level with PCR and at the protein level with Western blot against the Strep-tag, which showed presence of the protein and its ubiquitylated form. H2A.Z2 is seemingly expressed at around 18 to 24 times lower compared to H2A.Z1 based on Strep-tag antibody detection. Endogenous H2A.Z antibodies that detect H2A.Z1 levels in the cells were able to detect normal levels in H2A.Z2-Strep clones, but was unable to pick up a signal in the homozygous H2A.Z1-Strep clones, and a weak signal in the heterozygous H2A.Z1-Strep #2 clone. This itself might suggest that expressing Strep-tag at the C-terminus of H2A.Z1 might have knocked out H2A.Z1-Strep in homozygous clones by interfering with its gene expression. However, we know that there is detectable amount of H2A.Z1 protein in the homozygous clones based on Strep-tag blots, which suggests that Strep-tag antibody is more sensitive than the H2A.Z antibodies. This indicates that H2A.Z1 expression in homozygous is not knocked out, but potentially severely diminished. Additionally, knockout of H2A.Z1 is unlikely as studies have shown that H2A.Z1 deficiency is either lethal (Faast *et al*, 2001) or results in distinct alterations in cells (Matsuda *et al*, 2010). However, the clones were still viable and surviving. Although H2A.Z1 levels in all H2A.Z2-Strep clones seem to be normal, we do not know for sure if inserting Strep-tag to the C-terminus of H2A.Z2 has affected its protein expression as the endogenous H2A.Z antibodies are for detecting H2A.Z1 and not H2A.Z2.

To observe steady state amounts of mRNA of H2A.Z1 and H2A.Z2 in H2A.Z1-Strep clones and H2A.Z2-Strep clones, RT-PCR was conducted to provide a general

sense of RNA expression levels. The relative mRNA expression levels of H2A.Z1 and H2A.Z2 in the respective clones showed close to normal expression levels with a few exceptions, whereby 293T was used as normalization. H2A.Z1-Strep #5 had higher than normal and H2A.Z2-Strep #5 had lower than normal levels of H2A.Z1 mRNA transcribed. Relative mRNA expression was still observed in all clones, which supports the fact that H2A.Z1 should not have been knocked out, but that there might be some problems after transcription of the tagged genes. H2A.Z2 mRNA levels in all the Strep-tagged clones seem to be relatively comparable to the levels in 293T. Cell cycle progression in the cell population also seemed to be normal for most clones, except for H2A.Z1-Strep #5. We checked for mutations that might have led to the H2A.Z antibodies being unable to recognize H2A.Z, as Cas9 endonuclease might still be present in the cells and cleaving the genome non-specifically. Ideally, sequencing of the genomic loci of the H2A.Z1/H2A.Z2 alleles would be conducted to be sure of the correct insertions and no addition mutations. Due a failed attempt to do so and time constraints, Surveyor mutagenesis assay was instead conducted with 293T as the reference sample, but did not seem to show significant variance/mutation in the clones.

One speculation on why the protein levels of H2A.Z1 decreased drastically when Strep-tag was inserted to the C-terminus of H2AFZ could be because Strep-tag might be occluding certain regions when folded. A study conducted in *Drosophila* showed that the essential regions for H2A.Z to function properly were in its C-terminus region (Clarkson *et al*, 1999). In lieu with this, the SWR-1 ATPase chromatin remodeling

complex was found to target the region at the C-terminus of H2A.Z (Wu et al., 2005) and the acidic patch crucial for unique interactions of H2A.Z was found in the C-terminus in *S. cerevisiae* (Jensen *et al,* 2011). These findings highlight the importance of the C-terminus in the function of H2A.Z1, which could be blocked by Strep-tag when H2AFZ-Strep is expressed and folded in its protein structure. The only other region where an epitope tag could be inserted is at the 5' terminus of the gene. However, the challenge in doing so is that there is a splice site located after the first methionine ATG codon in the endogenous H2A.Z1 gene. If the Strep-tag sequence was inserted right after the first ATG codon, this might alter the normal splicing recognition sites and affect its proper splicing. This could pose a bigger problem than the current ones we're facing, and so we decided to continue with our C-terminus Strep-tagged H2A.Z1 and H2A.Z2 clones instead.

Although H2A.Z1 protein levels in H2A.Z1 homozygous clones were not detected by Western blot analysis using several H2A.Z antibodies, there were still detectable H2A.Z levels in the heterozygous H2A.Z1-Strep #2 clone. We decided to use it with H2A.Z2-Strep clones to compare the two histone variant isoforms. ChIP was conducted to determine if Strep-tagged H2A.Z1 and H2A.Z2 were incorporated into nucleosomes and whether they had similar characteristics as their unmodified counterparts. Strep-tag ChIP was used to probe for DNA regions occupied by nucleosomes containing either H2A.Z1-Strep or H2A.Z2-Strep. The immunoprecipitated chromatin was used in quantitative PCR with primers corresponding to the promoters of genes known to be

enriched for H2A.Z1-containing nucleosomes, such as MTA1 and TRIM21. H2A.Z1-Strep was found to be highly enriched at the proximal region on the TSS of the MTA1 gene in comparison to H2A.Z2-Strep. This might suggest some functional differentiation between the two isoforms, but additional biological and technical repeats should be done. Additionally, we should also factor in the difference in H2A.Z1-Strep and H2A.Z2-Strep protein levels between the clones, and the reduced H2A.Z2 protein levels could contribute to the differential enrichment observed. On the distal region of the TSS of MTA1 gene and on TRIM21, we did not observe any significant differences in occupancy by H2A.Z1 and H2A.Z2. In the future, we could probe the immunoprecipitated chromatin with more primers targeting different genes that are already known to be occupied by H2A.Z1-containing nucleosomes and determine if there are other differences between H2A.Z1 and H2A.Z2. However, this experiment was only conducted once due to time constraints, and needs to be replicated to state anything conclusively.

### 2.4.1 Future directions and implications

Moving forward, we could repeat the Surveyor mutagenesis assay to ensure that there are indeed no mutations occurring due to off-target effects by Cas9 at the site of Strep-tag insertion. Biological and technical repeats of ChIP qPCR should be conducted to ensure the trend we've observed is consistently seen. Additionally, we could use a more recently optimized CRISPR-Cas9 system to increase the efficiency of positive Strep-tagged H2A.Z1 or H2A.Z2 clones in other cell lines that are more difficult to

transfect. Instead of using the wild type *Streptococcus pyogenes* Cas9 (SpCas9), we could utilize the genetically modified Cas9 with enhanced specificity (eSpCas9) that has been shown to have a higher accuracy at inducing double-stranded breaks with less off-target effects (Slaymaker *et al*, 2016). Plasmids encoding eSpCas9 with sgRNA targeting either H2A.Z1, H2A.Z2 or Scrambled has already been successfully generated. To further increase HDR efficiency for the Strep-tag insertion, we've made some changes to the single stranded oligodeoxynucleotides (ssODNs) in our repair templates for both H2A.Z1-Strep and H2A.Z2-Strep. Currently, the repair templates we've used have symmetric homology arms, with 35 bp homology on both the 5' and 3' regions. A recent study suggests that asymmetric single stranded donor DNA seemed to be favored for insertions (Liang *et al*, 2016). New sets of repair templates targeting H2A.Z1 or H2A.Z2 have been redesigned to be flanked by asymmetric homology arms; with 67 bp on the 5' end and 35 bp at the 3' region. Additionally, to lower the chances of off-target effects by Cas9, the last nucleotide of the PAM sequences has been replaced by a Wobble base pair to prevent re-recognition by the Cas9 endonuclease. Taken together, we could use this enhanced system to determine if we could generate homozygous Strep-tagged H2A.Z1 and H2A.Z2 clones that mimic unmodified H2A.Z1 and H2A.Z2 levels in the parent cells. Additionally, we would need assurance that any future results obtained about H2A.Z1 and H2A.Z2 from the H2A.Z1-Strep and H2A.Z2-Strep 293T clones are not cell-line specific but universal. After optimizing our technique, it can be used to help elucidate the differential roles of both isoforms in the native environment during transcription regulation on chromatin.

The significance of H2A.Z's role in chromatin structure and function was highlighted by the fact that it is essential for viability in higher eukaryotes. It is specifically localized at the promoters of many genes in human and yeast. However, the mechanisms for the requirement of H2A.Z has yet to be clearly understood. Furthermore, the presence of two non-redundant isoforms, H2A.Z1 and H2A.Z2, adds another layer to the functional complexity of the H2A core histone variant. Understanding the functional difference between these two isoforms might provide insight into the dynamic roles of H2A.Z that were observed before the discovery of the presence of the H2A.Z2 isoform. This can thus contribute towards learning more about its role in epigenetic modifications on gene regulation and therefore, develop targeted therapeutic means to combat diseases like cancer.

## 3. Characterization of SUMOylated proteins localized to the transcription start site via the Avi-tag BirA system

### 3.1 Introduction

The solving of the nucleosome structure by crystallography paved the way for a better understanding of the interactions holding the nucleosome core particles together and how these interactions contribute to the forming of higher order chromatin structures (Luger *et al*, 1997). As histones are highly basic, the histone octamer contacts with the negatively-charged phosphate backbone of the DNA. The N-terminus of the histones proteins, also known as the N-terminal tails, are extended out of the nucleosome particle and the DNA double helix. They are thought to play a role in the formation of higher order chromatin structures as they can interact with neighbouring nucleosomes. The absence of histone tails in archaea suggests that it was an important evolutionary trait that helped in compaction of DNA, as well as in regulation of DNA accessibility to nuclear proteins (Malike and Henikoff, 2003). It is now widely known that histone post-translational modifications (PTMs) have a role in regulating chromatin structure and function. Most of the PTM sites are mainly found on the N-terminal tails extending past the DNA double helix, although some modifications can be found on the C-terminal tails and within the globular domains. Unlike other histones where their C-terminus tails are buried in the core of the nucleosome complex, H2A variants have both unstructured N- and C-terminal tails extending out of the complex (Corujo & Buschbeck, 2018).

### 3.1.1 Post-translational modifications (PTMs) of histones by SUMOylation

Histones and proteins can be modified by small chemical groups or via covalent binding of small polypeptides. Modifications of core histones by covalent binding of small polypeptides could modify properties of the nucleosome unit or the ability of certain nuclear proteins to bind with it and; therefore, could play a role to regulate the proper functions of chromatin (Lawrence *et al*, 2016). The most common PTM by covalent polypeptide addition is ubiquitin. One of its major roles is to bind to proteins and target them for proteolytic degradation by the proteasome. Aside from ubiquitin, there are other forms of ubiquitin-like (UBL) proteins that can conjugate to and modify the activities of its substrate proteins. SUMOylation, a process where small ubiquitin-like modifiers (SUMOs) are covalently attached to or detached from its substrate proteins, is another form of PTM. Studies have observed that these 13 kDa SUMOs conjugate to a large variety of proteins that function in important cellular processes such as transcription regulation and nuclear localization (Flotho and Melchior, 2013). SUMO and ubiquitin are similar in their structure (Bayer *et al*, 1998), but the distribution of the charged residues on both polypeptides are different. These surface residues contribute to the binding specificities that differ between ubiquitin and SUMO (Gill, 2004). Vertebrates, such as humans, express 4 paralogs of the SUMO protein, SUMO-1, SUMO-2, SUMO-3, and SUMO-4 while invertebrates express a single SUMO (Kerscher, 2007 OR). SUMO-2 and SUMO-3 proteins were found to share approximately 97% sequence identity and are commonly referred to as SUMO-2/3 (Wang and Dasso, 2009). SUMO-2/3 also share approximately 86% with SUMO-4, but

much is unknown about the ability of SUMO-4 to conjugate to other proteins (Owerbach *et al*, 2005; Wei *et al*, 2008 OR). On the other hand, they only share about 50% identity with SUMO-1 (Gareau and Lima, 2010 OR). The 4 paralogs of SUMO share homology in certain regions and are different in few regions in its amino acid sequence (Figure 3-1). Hence, they appear to alter both common and different substrates. For example, some substrates can be modified by both SUMO-1 and SUMO-2/3, while other substrates can only be modified by either of the SUMO paralog. To date, not much is yet known about the SUMO-4 paralog.

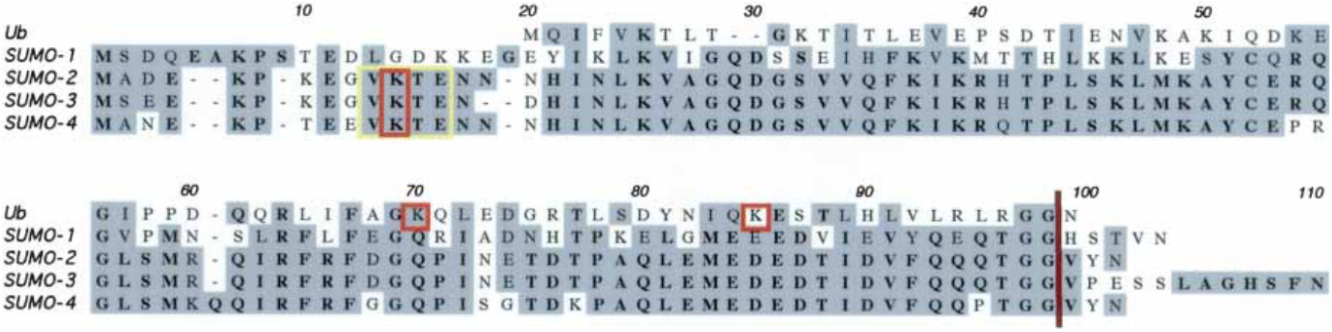



Figure 3-1. **Alignment of amino acid sequences of ubiquitin, SUMO1, SUMO2, SUMO3, and SUMO4.** Shaded areas indicate similarities, and bolded sequence represent identical sequences. The consensus motif for SUMOylation is outlined in yellow, and the SUMO acceptor, lysine, is outlined in red. Both were absent in SUMO1, suggesting the inability of poly-SUMOylation on SUMO-1. The bold red line after glycine-glycine at residue 98 for all 5 proteins indicated the cleavage site of the C-terminus tail to result in mature proteins. (Bohren *et al*, 2004)

SUMOs are conjugated to their target substrates via enzymatic machineries in a reversible process very similar to ubiquitylation (Figure 3-2). SUMO protein moiety is transcribed from its gene as an inactive precursor protein with a C-terminal extension (Wilson and Hochstrasser, 2016). A SUMO-specific protease, SENP, is then required to cleave the C-terminal extension of the precursor SUMO protein for its maturation, which

would expose two glycine residues required for conjugation (Li *et al*, 1999; Mukhopadhyay and Dasso, 2007). To initiate the SUMO conjugation process, E1 SUMO-activating enzyme, a heterodimer of SAE1 and SAE2 subunits (Desterro *et al*, 1999), transfers the AMP from ATP to adenylated SUMO (Hay, 2005). This AMP-SUMO bond is then broken and a thioester bond is formed with SAE2. Ubc9, a SUMO-specific conjugating enzyme, then transfers the thioester bond from the E1 enzyme to itself (Desterro *et al*, 1999). Unlike ubiquitin E2 conjugating enzymes, Ubc9 is able to directly distinguish substrate proteins for SUMOylation, and can catalyze formation of an isopeptide bond between SUMO and its target protein independent of an E3 ligase. The lysine residue that Ubc9 recognizes in the substrate protein is part of the SUMO consensus motif, ΨKxE, where Ψ indicates a large hydrophobic residue, x is any amino acid residue, and K lysine is the site of SUMO conjugation (Rodriguez *et al*, 2001). However, this consensus motif might not be found in every protein that can be modified by SUMO, suggesting the presence of some other sequence or a need for the E3 ligase.

More importantly, SUMO-addition consensus motifs can also be found in the N-terminus of SUMO-2/3 and SUMO-4, suggesting that these could be used for polymeric SUMO chains in these SUMO paralogs, but not in SUMO-1 (Tatham *et al*, 2001). Three SUMO paralogs, SUMO-1, SUMO-2 and SUMO-3, were later observed to form polymerization chains *in vitro* but its functional significance was not identified (Pichler *et al*, 2002). *In* vivo evidence of the SUMO polymerization was later observed by mass

spectrometry (Matic *et al*, 2008), and can occur via the lysine residue in the consensus motif *in vivo* and *in* vitro (Skilton *et al*, 2009), or via a lysine residue not in the consensus site *in vitro* (Wilson and Heaton, 2008).

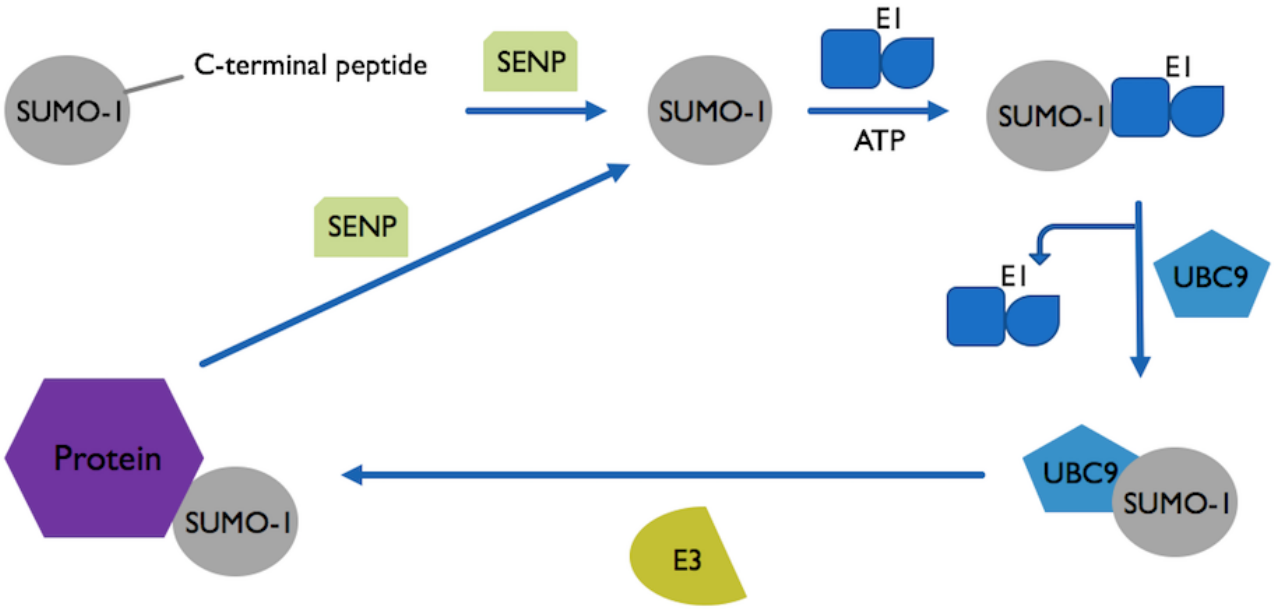



Figure 3-2. **SUMO conjugation pathway.** SUMO is transcribed as a precursor protein with an extended C-terminal peptide. SENP, a SUMO-specific protease, would cleave it to produce a mature SUMO protein. E1 ligase, a heterodimer, would then bind to SUMO for its transfer to the SUMO E2 conjugating enzyme, Ubc9. SUMO is then conjugated to its substrate from Ubc9 by the SUMO E3 ligase. For deconjugation from its substrate protein, SENP enzyme can cleave the bond between SUMO and its protein. The conjugation and deconjugation of SUMO is a rapid and dynamic process.

### 3.1.2 Role of SUMO in epigenetic transcription regulation

SUMO was first identified in mammalian cells to function in repression of transcription modifying H4 in an E3-independent manner (Shilo and Eisenman, 2003). Similarly, all 4 histone proteins in *S. cerevisiae* can also be SUMOylated *in vivo* and

function to repress transcriptional activity (Nathan *et al*, 2006). There are different ways SUMOylation of histones could function as a repressor for transcription. One way would be the recruitment of other proteins, such as histone deacetylase HDAC1 and heterochromatin protein 1 (HP1), to the DNA. This would repress transcription and maintain silenced regions in the chromatin structure (Shilo and Eisenman, 2003). Another way would be to compete with activating histone modifications, such as acetylation, for sites of modification on histones. This was observed in *S. cerevisiae* when there was an inverse correlation between histone H2B acetylation and H2B SUMOylation at the GAL1 locus (Nathan *et al*, 2006). In contrast to this finding, histone H3 SUMOylation on a nearby lysine residue could be stimulated by histone acetylation in mammalian cancer cell line, HeLa (Hendriks *et al*, 2014). Hence, mammalian histone H3 could simultaneously be modified by both acetyl and SUMO on nearby lysine residues. When these cells were treated with a HDAC inhibitor, increased levels of histone H3 SUMOylation was observed. However, these levels decreased when cells were treated with a histone acetyltransferase inhibitor. The function of this novel cross-talk between the 2 histone PTMs have yet to be determined, but suggests the complexity of SUMOylation in modifying histones for epigenetic regulation of transcription.

In addition to recruiting proteins or competing with other PTMs to regulate gene transcription levels, histone SUMOylation on H4 was found to modulate chromatin structure and long-range chromatin interactions (Dhall *et al*, 2014). The N-terminal tail of

H4 plays a role in chromatin condensation (Dorigo *et al*, 2003). By conjugating SUMO to its N-terminal tail, it could lead to other modifications in chromatin structure. In *S. cerevisiae*, all 4 core histone proteins and H2A-family variant, H2A.Z, have been detected in their SUMO-modified forms (Nathan *et al*, 2006; Kalocsay *et al*, 2009). In mammalian cells however, only H3 and H4 have been confirmed to be sumoylated (Hendriks *et al*, 2014). More importantly, histone SUMOylation have been identified in a myriad of organisms such as plants (Miller *et al*, 2010) and protozoan parasites (Issar *et al*, 2008). Its presence across species suggests histone SUMOylation as an evolutionarily conserved and significant chromatin PTM that adds complexity to gene regulation via epigenetics. Understanding the mechanisms and interplay of SUMOylation could allow us to potentially manipulate it as therapeutic targets in diseases resulting from un-regulated changes in gene transcription.

### 3.1.3 Dichotomous role of SUMOylation in transcription regulation

Instead of directly attaching SUMO to histones for induction of PTMs, there are other ways SUMO can impact transcription. Most of the SUMO target proteins are involved in nuclear function, especially in transcription regulation (Seeler and Dejean, 2003). SUMO can be conjugated to chromatin-modifying enzymes or proteins involved in transcription activation or repression to modify their functions and stability. SUMO-1 conjugation to HDAC4 was observed to promote its deacetylation activity and serve to repress gene transcription levels (Kirsh *et al*, 2002). Overexpression of SUMO or Ubc9 to hyper-SUMOylate specific gene promoters primarily repressed genes (Chupreta *et al*,

2005; Shiio and Eisenman, 2003) while knocking-down SUMOylation in cells resulted in gene expression activation (Spektor *et al*, 2011). One of the most common classes of proteins modified by SUMO includes transcription factors and co-regulators. Aside from mediating repression of transcription, SUMO can also enhance transcriptional activity by modifying certain transcription factors, such as Ikaros (Gómez-del Arco *et al*, 2005). In yeast studies, SUMO-modified proteins were found at promoters of constitutively active genes and Ubc9 and SUMO at promoters of inducible genes upon induction of gene activation (Rosonina *et al*, 2010).

Unlike the acetylation of histones that generally lead to increase in gene transcription by allowing easier access of the promoter region, SUMOylation of histones does not result in a direct black-and-white effect, and can lead to either activation or repression of transcription. Most of the current studies and research point to the fact that SUMOylation of a transcription factor results in recruitment of histone deacetylases, leading ultimately to the repression of gene transcription and expression (Yang and Sharrocks, 2004; Lindberg *et al*, 2010; Murata *et al*, 2010). However, there are some findings to suggest SUMOylation could enhance certain gene transcription, or even its activation (Guo and Sharrocks, 2009; Lyst *et al*, 2006). SUMO enrichment at active promoters was observed in *S. cerevisiae* and human fibroblasts (Neyret-Kahn *et al*, 2013; Rosonina *et al*, 2010; Liu *et al*, 2015;). The dichotomous role of SUMOylation on transcription, partly due to its extensive impacts on chromatin structure and modifications, was exemplified with studies on Sp3, a transcription factor modified by

SUMO-1. Impeding SUMOylation of Sp3 resulted in differential effects on Sp3-target promoters to modify the structure of chromatin on multiple levels such as reducing levels of DNA and histone methylation, histone methyltransferases, HP1, and two ATP-dependent chromatin remodelers (Ross *et al*, 2002; Stielow *et al*, 2010, 2008a, 2008b). These findings point to the fact that SUMOylation is not a simple switch to just silence or activate genes, but contributes to an important role in regulation gene transcription.

### 3.1.4 SUMO and diseases

SUMOylation has been found to influence certain proteins involved in different diseases and in cellular processes characteristic to cancer, such as cell proliferation and apoptosis. Interestingly, SUMO was recently found to contribute to Alzheimer's disease, which is a progressive neurodegenerative disease characterized by an accumulation of amyloid beta plaques and neurofibrillary tangles containing tau protein associated with microtubules. The amyloid precursor protein and the tau protein was found to be SUMOylated for its regulation, and findings suggest its involvement in the pathogenic pathway of Alzheimer's diseases (Hoppe *et al*, 2015). Hence, by understanding more of its contribution to the pathway, SUMO could be used as a neuroprotectant to potentially treat Alzheimer's disease. Additionally, SUMO is also implicated in modifying transcription factors essential for normal cardiac development, and deregulation of the baseline SUMOylation could result in cardiac diseases (Wang, 2012). These studies suggest a role of SUMOylation in disease development. By

understanding the exact mechanistic pathway of SUMO, it can potentially be used as a therapeutic target for diseases.

Cancer is a type of disease that can target most of the body's organs and impact people the most. In addition to being involved in meiosis and mitosis, genome-wide RNA interference screens scored several SUMO pathway components as genes required for cell proliferation (Schlabach *et al*, 2008). Loss of SUMO in human cancer cells resulted in decreased cell growth *in vitro*, with multinucleated cells going into senescence, and increased apoptosis (He *et al*, 2015). Furthermore, from the same study, *in vivo* xenograft tumor model showed significantly impaired tumor growth. This suggested the importance of SUMO pathway in cancer cell proliferation *in vitro* and *in vivo*, implying the use of SUMO pathways as a potential target for cancer therapies. However, more research is required to ensure the specificity of SUMO in cancer tumor cells as SUMO impacts pathways for regular cellular processes.

### 3.1.5 Research rationale and objective

SUMOylation on histones generally suggests a role in repressing gene expression, but data have shown SUMOylation of some transcription factors can also function in gene activation. Studies in yeast found that H2A.Z is SUMOylated at K126 and K133 of the C-terminal tail of this variant in response to DNA double-strand breaks and DNA repair (Kalocsay *et al*, 2009). Taking these studies into consideration, we want to determine if H2A.Z, which is also involved in both gene repression and activation,

could be SUMOylated in mammalian cells. Secondarily, we want to take advantage of the proximity of H2A.Z to the nucleosome-free region (NFR) to identify potential SUMOylated targets (non-histone targets) recruited to the TSS. Our lab previously successfully utilized and described a technique termed BICON, which stands for Biotinylation-assisted Isolation of CO-modified Nucleosomes. We wanted to determine if we could use a modified version of this technique to isolate for biotinylated proteins that were modified by SUMO in proximity of H2A.Z. To do so, we fused the BirA ligase to H2A.Z and added the BirA substrate target, Avi-tag, to SUMO. Flag-NLS-BirA was used as a negative control to ensure the biotinylated bands observed were specific for H2A.Z instead of by random chance of BirA being in the nucleus, and the Flag-tag in H2A.Z-Flag-BirA was used for detecting the expression of the protein by Flag antibody. Once we observe biotinylated SUMOylated proteins in context of H2A.Z proximity, we could conduct mass spectrometry analysis to identify putative transcription-related proteins that are SUMOylated. However, we cannot assume that all biotinylated proteins are at the TSS as H2A.Z is not exclusively localized to the TSS and could also be found in other regions of the genome.

## 3.2 Materials and Methods

### 3.2.1 Cell culture, transfection, and plasmids

293T and HeLa cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) High-Glucose with 10% fetal bovine serum (FBS) in 37℃ incubators with 5% $CO_2$. Cells were passaged every 3-4 days when the cells were at 80-90% confluency. Plasmids were then transfected using PolyJet by the manufacturer's protocol instructions. In co-transfection experiments, the ratio of H2A.Z-Flag-BirA/H2A.Z2-Flag-BirA/Flag-NLS-BirA/H2A.X-Flag-BirA to Avi-SUMO1/Avi-SUMO3/Avi-Ub were 1:1. Because H2A.Z2 histone variant isoform was only discovered relatively recently, H2A.Z1 has generally been referred to as H2A.Z. With that said, H2A.Z would be interchangeably used with H2A.Z1. Cells were then harvested 48-hours post-transfection and washed 2 times in Phosphate Buffered Saline (PBS) before centrifuging them down as cell pellets.

### 3.2.2 Western blot analysis

Whole cell protein extraction was first conducted by trypsinization to lift adhesive cells off the culturing plate and collected as cell pellets. The cell pellets were washed with PBS once and lysed with boiling hot 2x SB before sonication at 40% duty cycle and 30% output control for 13 pulses. The samples were then boiled for an additional 5 minutes and spun down before transferring the supernatant to a new microcentrifuge tube. The protein samples from whole cell fraction were stored at -20℃ for analyses in the future.

Proteins in the nuclear fraction was isolated by lysing the cell pellet with 0.2% Triton-X100 in Buffer A for 5 minutes on ice and centrifuged at 1300 G for 5 minutes at 4℃. The supernatant containing cytoplasmic fraction was removed and the pellet containing the nuclear fraction was lysed in boiling hot 2x Sample Buffer or RIPA buffer before sonication at 40% duty cycle and 30% output control for 13 pulses or 95% output control 8 cycles at 8 pulses each. The nuclear fraction protein samples were either stored at -20℃ for Western blots or used immediately in immunoprecipitation.

Equal concentrations of proteins per sample, along with DTT, bromophenol blue, and additional 2X SDS buffer were boiled for 5 minutes before samples were resolved on 10% or 15% tris-glycine gels by SDS-PAGE. Gels were stained in Coomassie Brilliant Blue dye for an hour at room temperature before being de-stained in destaining solution overnight at room temperature on a nutator. Gels containing the IP-ed samples not observed by Coomassie Brilliant Blue dye were then stained with Silver stain (Pierce kit #). Samples for Western blot analysis were wet transferred from the gels to PVDF membrane by wet transfer method at 100V for 1 hour.

PVDF membranes were blocked with 5% non-fat milk in 1X TBS-T (for anti-Flag mono-clonal or anti-SUMO1) or 4% BSA in 1X TBS-T (for Avi-HRP antibody) for an hour at room temperature on a nutator. The membranes were then washed 3 times for 5 minutes each with 1X TBS-T and probed overnight at 4℃ with primary antibodies in 2% milk in TBS-T at the respective dilutions; M2 Flag (1:15000), anti-SUMO1 (1:200), Avi-

HRP (1:80 000). M2 Flag antibody was used to detect for expression of the constructs/plasmids while Avi-HRP was used to detect biotinylated proteins. SUMO1 antibody was used to detect SUMOylated proteins. Membranes were then washed in TBS-T 3 times for 5 minutes and incubated in ThermoScientific Supersignal West Pico Chemiluminescent Substrate or the higher sensitivity Millipore Immobilon Western Chemiluminescent HRP Substrate for 5 minutes. They were then exposed to autoradiography film and developed in a dark room accordingly.

### 3.2.3 Heat shock and DNA damage treatments

Heat shock was conducted at 42℃ for 1 hour and DNA damage was induced with 200 ng/uL of methyl methanesulfonate (MMS) for 3 hours at 37℃. Both treatments were respectively given to 293T cells and HeLa cells 48 hours post-transfection, and nuclear fraction protein were then obtained with Buffer A + 0.2% Triton-X100 and 2x SB.

### 3.2.4 Transcription inhibition with Flavopiridol and Triptolide drugs

Transfected 293T cells were treated with 1 uM of Flavopiridol resuspended in water and 1 uM of Triptolide resuspended in DMSO for 24 hours and proteins in the nuclear fraction were obtained with Buffer A + 0.2% Triton-X100 and 2x SB.

### 3.2.5 Immunoprecipitation of biotinylated SUMOylated proteins with Streptavidin sepharose beads

Nuclear fraction proteins from 2 15-cm cell culture dishes were diluted in equal volume of Buffer D with 2x protease inhibitors for proteins lysed with 2x Sample Buffer or RIPA buffer with 2x protease inhibitors for proteins lysed with RIPA buffer. Streptavidin sepharose beads were washed 3x in either Buffer D or RIPA buffer, depending on the lysis buffer, and centrifuged at 7000 G for 30 seconds. Nuclear fraction proteins were then incubated with 50 uL Streptavidin sepharose beads on a nutator at 4℃ overnight, and washed 3x each in RIPA buffer with protease inhibitors and 50 mM ammonium bicarbonate with protease inhibitors. Biotinylated proteins bound to the streptavidin beads were then eluted with 2x sample buffer and boiled for 5 minutes. Samples were then stored at -20℃.

### 3.3.6 Sample preparation for mass spectrometry

Samples to be sent for mass spectrometry (MS) were all prepared in a laminar flow hood for most steps possible to prevent or minimize keratin contamination in the samples. Lysis buffers and wash buffers were also freshly prepared in the laminar flow hood before the start of the pulldown. There were 2 different ways the samples were prepared for MS; 1) on-bead trypsinization of samples; and 2) in-gel tryptic digest of samples. Nuclear proteins were extracted from transfected cells with Buffer A + 0.2% Triton-X100 and 2x SB or with RIPA and IP-ed with Streptavidin beads.

3.2.6.1 On-bead trypsinization of samples

Instead of eluting biotinylated proteins bound on the Streptavidin beads with 2x Sample Buffer and boiling it, the beads were incubated with 1 ug of Trypsin Gold overnight at 37℃ on a nutator and an additional 1 ug was added the next day and incubated further for 2 hours. Supernatant was later collected and transferred to a clean microcentrifuge tube, and beads were rinsed twice with Molecular Grade water. The rinses were combined with original supernatant and centrifuged at high speed for 10 minutes before transferring most of the supernatant to a clean microcentrifuge tube to prevent contamination of beads. The samples were then lyophilized in a speed-vac and resuspended in 5% formic acid. They were stored at -80℃ until ready for mass spectrometry analysis.

3.2.6.2 In-gel tryptic digest of samples

Biotinylated samples were incubated with Streptavidin sepharose beads overnight at 4℃ and eluted as normal in 2x SB. The elute samples were run on a pre-cast 10% tris-glycine gel and SDS-PAGE was conducted at 150V for 17 minutes and 200V for 60 minutes. Specific protein spots were excised depending on their molecular weight ranges with a sterile blade using a Silver stained gel as guide and transferred into clean microcentrifuge tubes. The gel pieces were generously covered with HPLC-grade acetonitrile and incubated for 10-20 minutes, when the gel would shrink. The gel pieces were then washed with HPLC-grade water for 10-20 minutes for it to re-swell. This shrinking and swelling was conducted 2 additional times before the gel pieces were

dried in a speed-vac. The dried gel pieces were then incubated in 10 mM DTT in 100 mM ammonium bicarbonate for 60 minutes at 37℃. The solution was removed and the gel pieces were re-suspended in enough acetonitrile for 20 minutes to allow shrinkage. The solution was removed and 55 mM IAA in 100 mM ammonium bicarbonate was added for 15 minutes before drying the gel pieces with a speed-vac. The dried gel pieces were then incubated with enough 100 mM ammonium bicarbonate containing 1 ug of Trypsin Gold and digested overnight at 37℃. The samples were centrifuged the next day for 5 minutes at 14000 RPM and supernatant was transferred to a new clean microcentrifuge tube to be set aside. Peptide from the gel pieces were then further extracted by incubating with enough 50% acetonitrile with 5% formic acid for 30 minutes and vortexed briefly every 10 minutes before centrifugation. The supernatant from this step was then combined with the initial supernatant. Extraction was repeated once more and supernatants were lyophilized with a speed-vac before storage in -80℃ until ready for mass spectrometry analysis.

## 3.3 Results

### 3.3.1 H2A.Z is not SUMOylated in 293T cells

H2A.Z is known to be ubiquitylated in mammalian cells. We used H2A.Z ubiquitylation to test if we could use the Avi-tag BirA system to detect the known PTM of H2A.Z before using it to test for the SUMOylation of H2A.Z. To do so, BirA fused to H2A.Z-Flag and Avi-tag fused to ubiquitin were co-expressed in 293T cells. In parallel, we also co-transfected either H2A.Z-Flag-BirA or Flag-NLS-BirA with either Avi-SUMO1 or Avi-SUMO3. SUMO2 homolog was not used as it has high amino acid sequence similarity to SUMO3, and so SUMO3 could be regarded as SUMO2/3. Whole cell extract proteins were then used in Western blot. Flag-antibody was used to detect for the expression levels of the transfected constructs and the potential SUMOylation of H2A.Z. H2A.Z-Flag-BirA runs at approximately 50 kDa on SDS gels, whereas its Avi-ubiquitylated form runs at 59 kDa and the Avi-SUMOylated form runs at 63 kDa. Endogenous ubiquitin modified H2A.Z-Flag-BirA corresponds to a band at 58 kDa. Flag-NLS-BirA runs at 35 kDa, and was used as a control to ensure that the bands observed with H2A.Z-Flag-BirA were specific for H2A.Z and not purely due to chance. Bands specific to H2A.Z would therefore not show up with Flag-NLS-BirA, and so, Flag-NLS-BirA would interchangeably be referred to as a negative control. Expected band sizes are represented in Table 3-1 for future reference.

Table 3-1. **Expected protein sizes of constructs transfected and modified H2A.Z-Flag-BirA.** Proteins under the "Unmodified" represents the expected molecular weight of the proteins that should be expressed from the transfected plasmids. "Modified H2A.Z-Flag-BirA" indicates the predicted molecular weight of H2A.Z-Flag-BirA modified by either Avi-ubiquitin, endogenous ubiquitin, or Avi-SUMO.

| | Protein | Expected molecular weight (kDa) |
|---|---|---|
| **Unmodified** | Flag-NLS-BirA | 35 |
| | H2A.Z-Flag-BirA | 50 |
| | Avi-Ubiquitin | 9 |
| | Avi-SUMO | 13 |
| **Modified H2A.Z-Flag-BirA** | Av-ubiquitylated H2A.Z-Flag-BirA | 59 |
| | Ubiquitylated H2A.Z-Flag-BirA | 58 |
| | Di-ubiquitylated H2A.Z-Flag-BirA | 66 |
| | Avi-SUMOylated H2A.Z-Flag-BirA | 63 |

H2A.Z Avi-Ubiquitylation was observed with a high intensity band at 59 kDa with the Avi-HRP blot (Figure 3-3, blue arrow). Although additional bands were observed after high exposure of the Avi-HRP blot, they were of extremely low intensity compared to the band at 59 kDa. The Flag blot also supported H2A.Z-Flag-BirA ubiquitylation when a similar band at 59 kDa was also picked up by the Flag-antibody (Figure 3-3, blue arrow). A band at 58 kDa was also observed with the Flag blot, corresponding to the endogenous ubiquitylated H2A.Z-Flag-BirA protein (Figure 3-3, red arrow). The detection of ubiquitylated H2A.Z by both Avi-HRP and Flag-antibody suggests the efficiency of the Avi-tag BirA and that we could reliably use the system to determine SUMOylation of H2A.Z. Doing so, a band at 63 kDa was detected with Avi-HRP blot and SUMO1 blot to indicate Avi-SUMOylated H2A.Z-Flag-BirA (Figure 3-3, red asterisk). We also observed multiple bands of equal intensity around 63 kDa and additional bands at other molecular weights in the Avi-HRP and SUMO-1 blots. Most importantly, compared to the intensity of the Avi-HRP-detected Avi-ubiquitylated H2A.Z-Flag-BirA band, it is clear that if there were any Avi-SUMOylation of H2A.Z-Flag-BirA, the levels were very low in comparison to the Avi-ubiquitylated form. Furthermore, a similar band at 63 kDa was observed with Flag-NLS-BirA control with Avi-HRP and SUMO1 antibody, indicating the non-dependency of that band on H2A.Z. To support the hypothesis that the 63 kDa band observed with H2A.Z-Flag-BirA is not the SUMOylated form, the Flag antibody was unable to pick up the band at 63 kDa. We hence conclude that although there is no clear SUMOylation of H2A.Z as no dominant band was observed at its expected size and the band that was detected was not specific to H2A.Z. However,

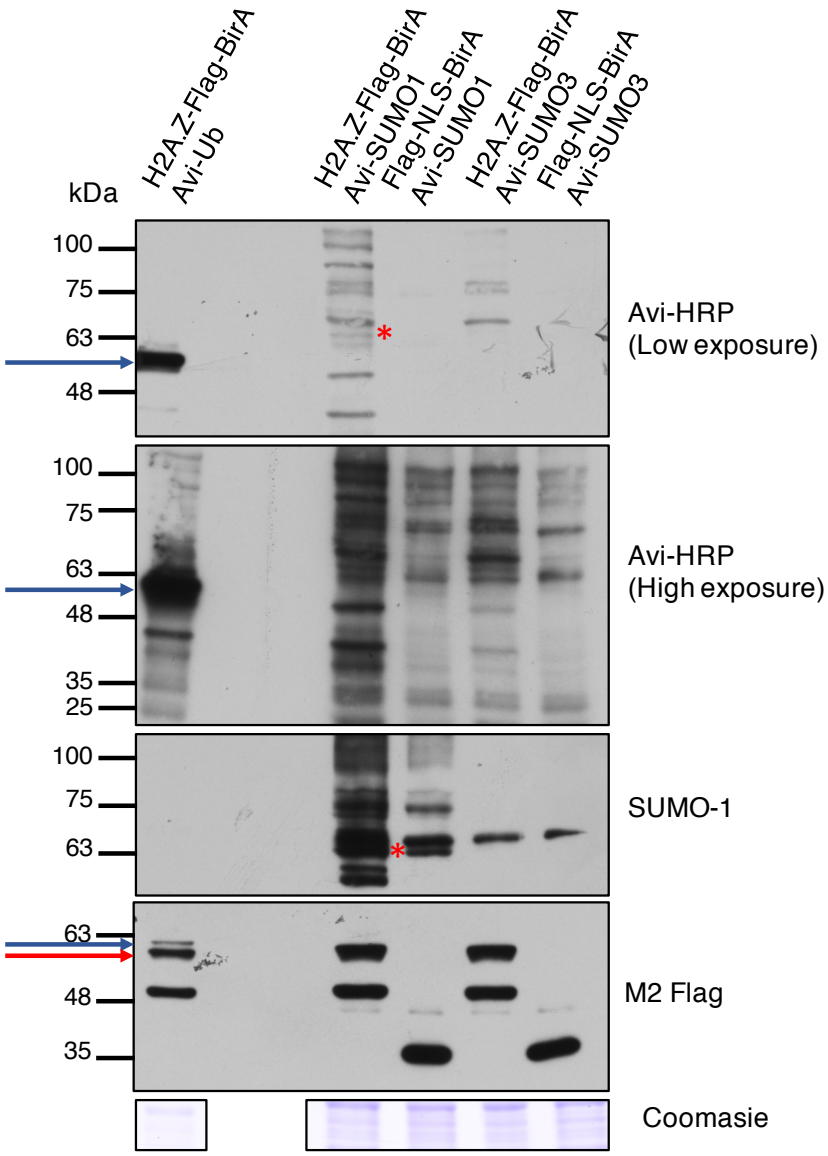



Figure 3-3. **Preliminary test for biotinylation of SUMOylated proteins in proximity of H2A.Z.** Flag blot ensured the expression of the plasmids transfected were roughly equal. Avi-Ubiquitin (Avi-Ub) was used as a positive control for biotinylation of H2A.Z and Flag-NLS-BirA was used as a control to ensure bands observed were specific to H2A.Z and not due to its presence in the nucleus. Blue arrows represent Avi-ubiquitylated H2A.Z-Flag-BirA, and the red arrow indicates endogenous ubiquitylated H2A.Z-Flag-BirA. Red asterisks show the 63 kDa band that could correspond to Avi-SUMOylated H2A.Z-Flag-BirA.

there might be low levels of H2A.Z SUMOylation and the presence of multiple bands with H2A.Z-Flag-BirA suggests potential H2A.Z-dependent Avi-SUMOylation of other target proteins. We can then use the Avi-tag BirA system with H2A.Z to also identify some target proteins of SUMO at the TSS. Moving forward, we decided to focus on SUMO1 as more biotinylated bands were observed in Avi-HRP blot compared to SUMO3.

H2A.Z-Flag-BirA or Flag-NLS-BirA was again co-expressed with Avi-SUMO1 in 293T cells. Whole cell extract and nuclear extract were prepared as an optimization step to determine which protein fractionation method would allow for clearer detection of biotinylated bands. The data showed that there were more distinct bands with the nuclear extracts when blotted with Avi-HRP and SUMO1 antibody (Figure 3-4A). Therefore, we consistently harvested nuclear extracts for our analyses in subsequent experiments. For higher molecular weight proteins, we decided to run the samples in lower percentage polyacrylamide gels for better resolution. We observed the 63 kDa band corresponding to Avi-SUMOylated-H2A.Z-Flag-BirA in Avi-HRP, SUMO1, and Flag blots (Figure 3-4B blue arrows). However, we cannot be certain of this since the same size band was also seen in the Flag-NLS-BirA control in Avi-HRP blot, and the intensity of the 63 kDa band in H2A.Z-Flag-BirA sample is very low compared to other bands detected in SUMO1 and Flag blots. This supports our statement that H2A.Z is either lowly SUMOylated, or not modified by SUMO at steady states in 293T cells.

Additional bands were again observed in H2A.Z-Flag-BirA samples that were absent with Flag-NLS-BirA, suggesting potential SUMO targets dependent on H2A.Z.

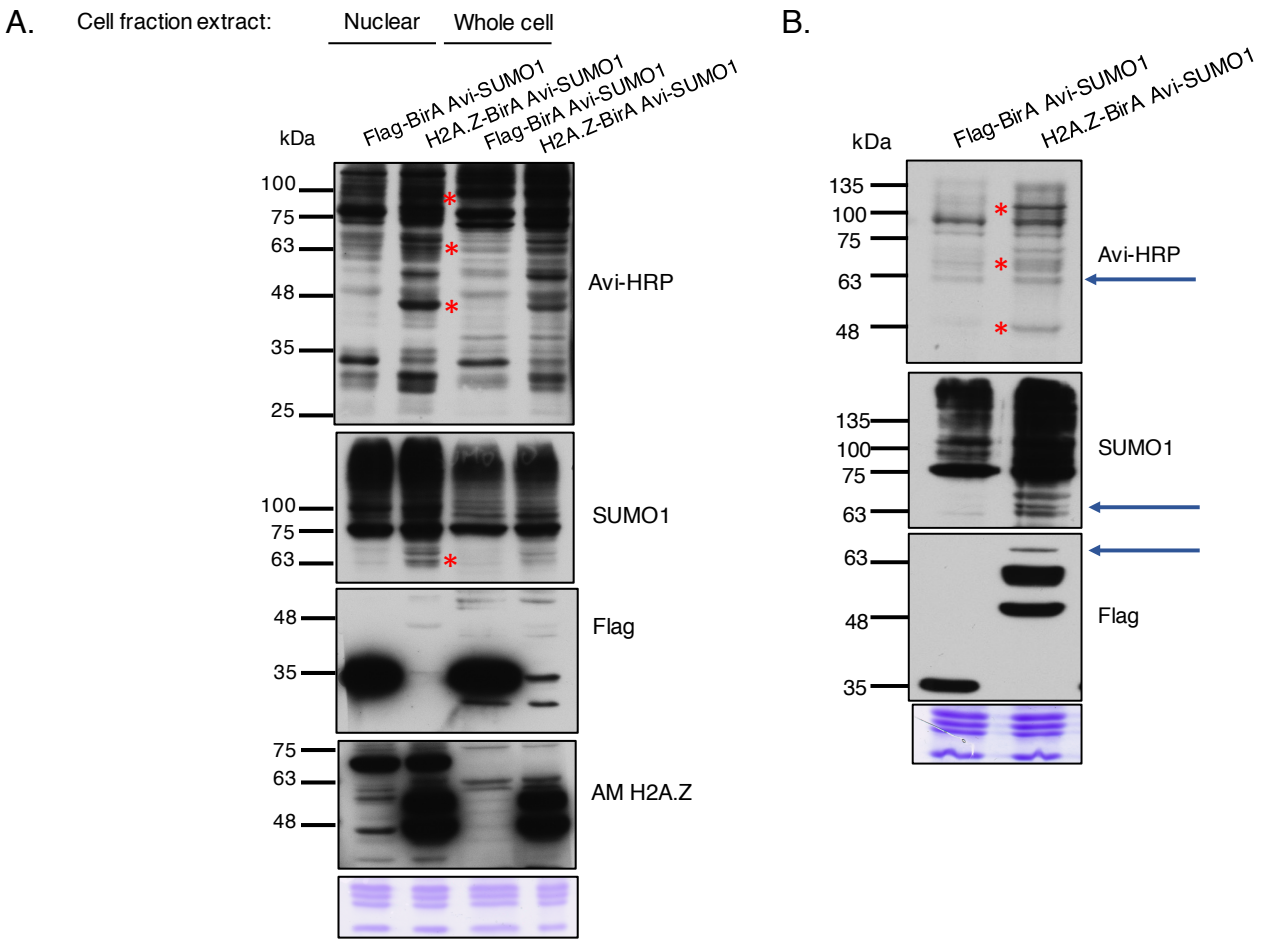



Figure 3-4. **Biotinylation of potential transcription start site proteins modified by Avi-SUMO1 near proximity of H2A.Z histone variant tagged with BirA ligase.** (A) Western blot analysis conducted on nuclear and whole cell extract from 293T cells transfected with Avi-SUMO1 and Flag-NLS-BirA or H2A.Z-Flag-BirA. (B) A biological repeat was conducted and the nuclear fraction proteins were then extracted and used in Western blot for analysis with Avi-HRP, SUMO1 antibody, Flag antibody and AM H2A.Z antibody. The 63 kDa band that could potentially be Avi-SUMOylated-H2A.Z-Flag-BirA is highlighted in blue arrows while H2A.Z-Flag-BirA-dependent-Avi-SUMOylated targets are shown in red asterisks.

A few bands at different molecular weights were observed in the H2A.Z-Flag-BirA-expressing sample and absent in the Flag-NLS-BirA control; a band at 45 kDa, bands above the 63 kDa mark, a band at 110 kDa (Figure 3-4B, red asterisks) This suggested that they might be SUMOylated proteins that were biotinylated due to its proximity to H2A.Z-BirA. These bands are consistently observed throughout the rest of the experiments described in this project, raising the confidence that they are putative H2A.Z-Flag-BirA-SUMOylated targets.

### 3.3.2 H2A.Z is not SUMOylated in response to heat shock or DNA damage

Previous studies have shown that heat shock treatment can induce transcription of some target genes and lead to SUMOylation of some transcription factors. We wanted to test if H2A.Z was SUMOylated in response to heat shock, and whether any of the biotinylated and Avi-SUMOylated proteins detected in our system were related to the heat shock response. If H2A.Z SUMOylation be involved in response to heat shock, a high intensity band at 63 kDa would be observed with Avi-HRP, SUMO1 antibody and Flag antibody after heat shock is conducted. If SUMO targets proteins that are involved in heat shock response, more biotinylated bands would be observed. Heat shock treatment was conducted on transfected 293T cells for one hour at 42℃. We did not observe the expected results, and no additional bands were observed after heat shock (Figure 3-5), suggesting that heat shock does not induce SUMOylation of H2A.Z or H2A.Z dependent target proteins.

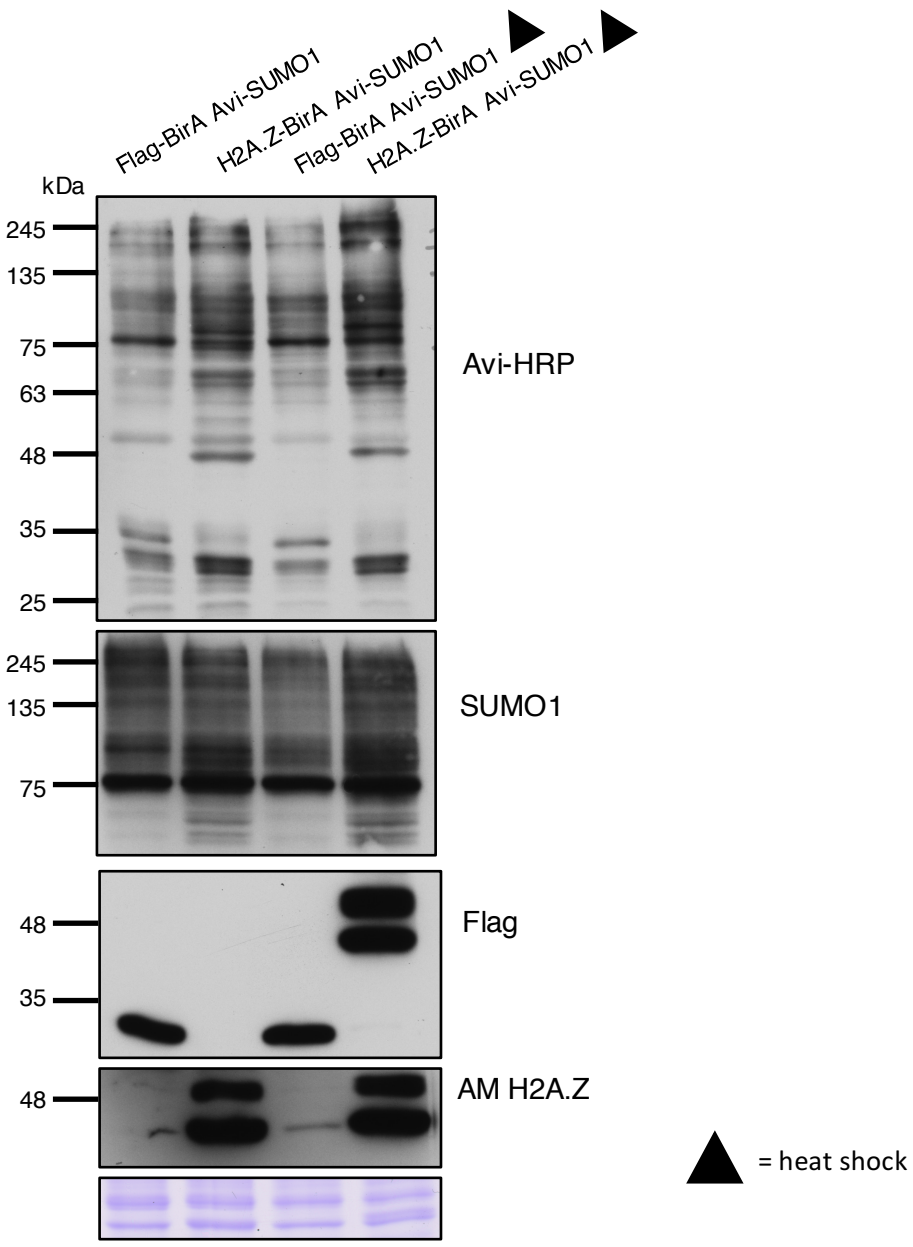



Figure 3-5. **No changes in state of H2A.Z SUMOylation in response to heat shock treatment**. 293T cells were transfected with Avi-SUMO1 and either H2A.Z-BirA or Flag-NLS-BirA in duplicate. One set was then incubated in 42℃ for an hour while the other was in normal 37℃ incubator. Nuclear fraction were then extracted and used in Western blot for analysis with Avi-HRP, SUMO1 antibody, Flag antibody and AM H2A.Z antibody. No additional biotinylated bands were observed with the heat shock treatment.

H2A.Z SUMOylation was originally detected in yeast under DNA damage. To test whether this pathway was conserved in mammalian cells, DNA damage was induced in the transfected 293T cells using MMS treatment for 3 hours. For comparison, we also performed the same experiment using HeLa cells as well since that cell line has been more commonly used for DNA-damage studies. In addition, another isoform of the histone variant, H2A.Z2-Flag-BirA, was also co-transfected along with Avi-SUMO1 since a recent paper (Fukuto *et al*, 2018) claimed that this specific isoform of H2A.Z was SUMOylated in a DNA damage-dependent manner. Gamma H2A.X ($\gamma$H2A.X) was used as an indicator of DNA damage. Higher $\gamma$H2A.X levels were observed in samples treated with MMS in comparison to samples without MMS treatment (Figure 3-6A). The increase in $\gamma$H2A.X levels were more evident in 293T cells than HeLa cells, indicating a more efficient DNA damage induction in 293T cells. After assuring induction of DNA damage in the cells, we went on to determine if doing so would result in SUMOylation of H2A.Z1 or H2A.Z2 in response. However, we did not see any a band of higher intensity at 63 kDa after DNA damage in both H2A.Z1 (Figure 3-6B) and H2A.Z2 (Figure 3-6C) in either of the cell lines. This suggests that neither H2A.Z1 nor H2A.Z2 SUMOylation is involved in DNA damage response.

### 3.3.3 Putative H2A.Z-Flag-BirA-SUMOylated targets are not dependent on on-going transcription

We wanted to further characterize the biotinylated (presumably SUMOylated) bands corresponding to potential putative H2A.Z-Flag-BirA-SUMOylated targets.

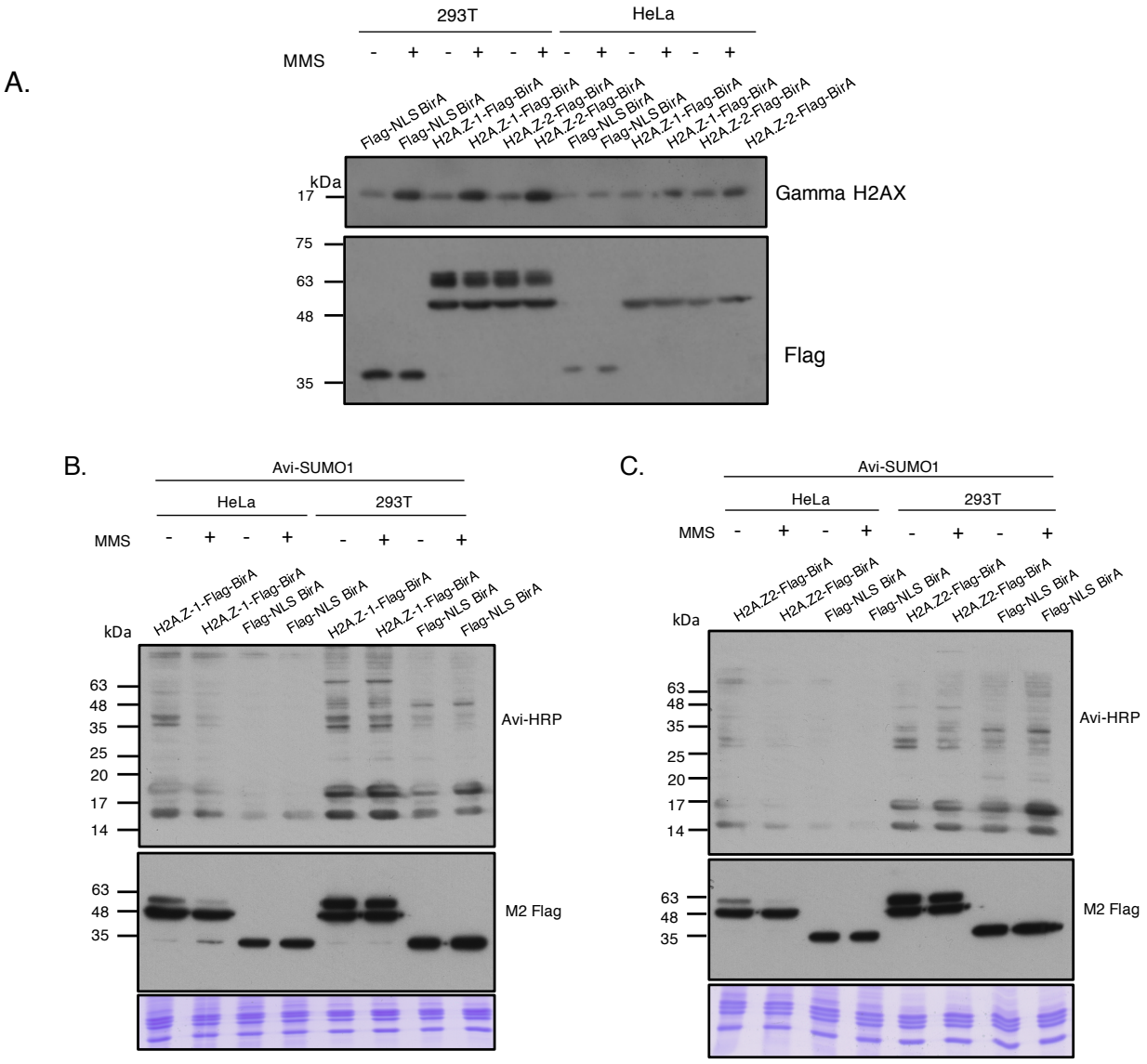



Figure 3-6. **SUMOylation is not involved in DNA damage response of H2A.Z1 or H2A.Z2 in 293T and HeLa cells.** DNA damage was induced to 293T cells 48-hours post-transfection by MMS for 3 hours at 37℃. (A) Gamma H2A.X detection was used to ensure DNA damage induction. In 293T cells, a more intense band was observed in the presence of MMS. This increase in intensity with MMS treatment was not very distinct in HeLa cells. Flag antibody was used as a control to ensure equal expression and loading of the plasmids transfected. (B) DNA damage was conducted as described to 293T and HeLa cells transfected with Avi-SUMO1 and H2A.Z1-BirA or Flag-NLS-BirA. Nuclear fraction proteins were analyzed with Western blot. C) DNA damage was conducted as described to 293T and HeLa cells transfected with Avi-SUMO1 and H2A.Z2-BirA or Flag-NLS-BirA. Nuclear fraction proteins were analyzed with Western blot.

SUMOylation seem to be related to activation of transcription, so we hypothesized that these SUMOylated proteins detected by H2A.Z are related to ongoing transcription. Transfected 293T cells were treated with two transcription inhibitors, Flavopiridol and Triptolife, for 24 hours before nuclear extract was prepared. There did not seem to be any distinct changes in the intensity nor the number of biotinylated SUMOylated bands with Flavopiridol (Figure 3-7A) as the same pattern of bands were observed after treatment. Similar observations were made with Triptolide treatment, as no changes in the pattern of biotinylated bands were observed (Figure 3-7B). This supports our conclusion that inhibition of transcription in 293T cells does not affect the H2A.Z-Flag-BirA-dependent biotinylated bands. Hence, our hypothesis that the biotinylated (presumably SUMOylated) proteins detected by H2A.Z are related to ongoing transcription.

### 3.3.4 Purification of the putative H2A.Z-Flag-BirA-SUMOylated targets

We've consistently observed bands representing biotinylated proteins dependent on H2A.Z-Flag-BirA in our nuclear extracts that were detected by the Avi-HRP blot. To confirm the modification by H2A.Z-BirA on Avi-SUMO1 targets, Streptavidin beads were used to pull down and purify the biotinylated proteins from the nuclear extracts of transfected cells. Streptavidin beads bound with the protein samples were washed with increasing stringency (0.2% Triton-X100 and 0.5% Triton-X100 in buffer A) to ensure proteins bound to the beads were specific. The biotinylated proteins were then eluded

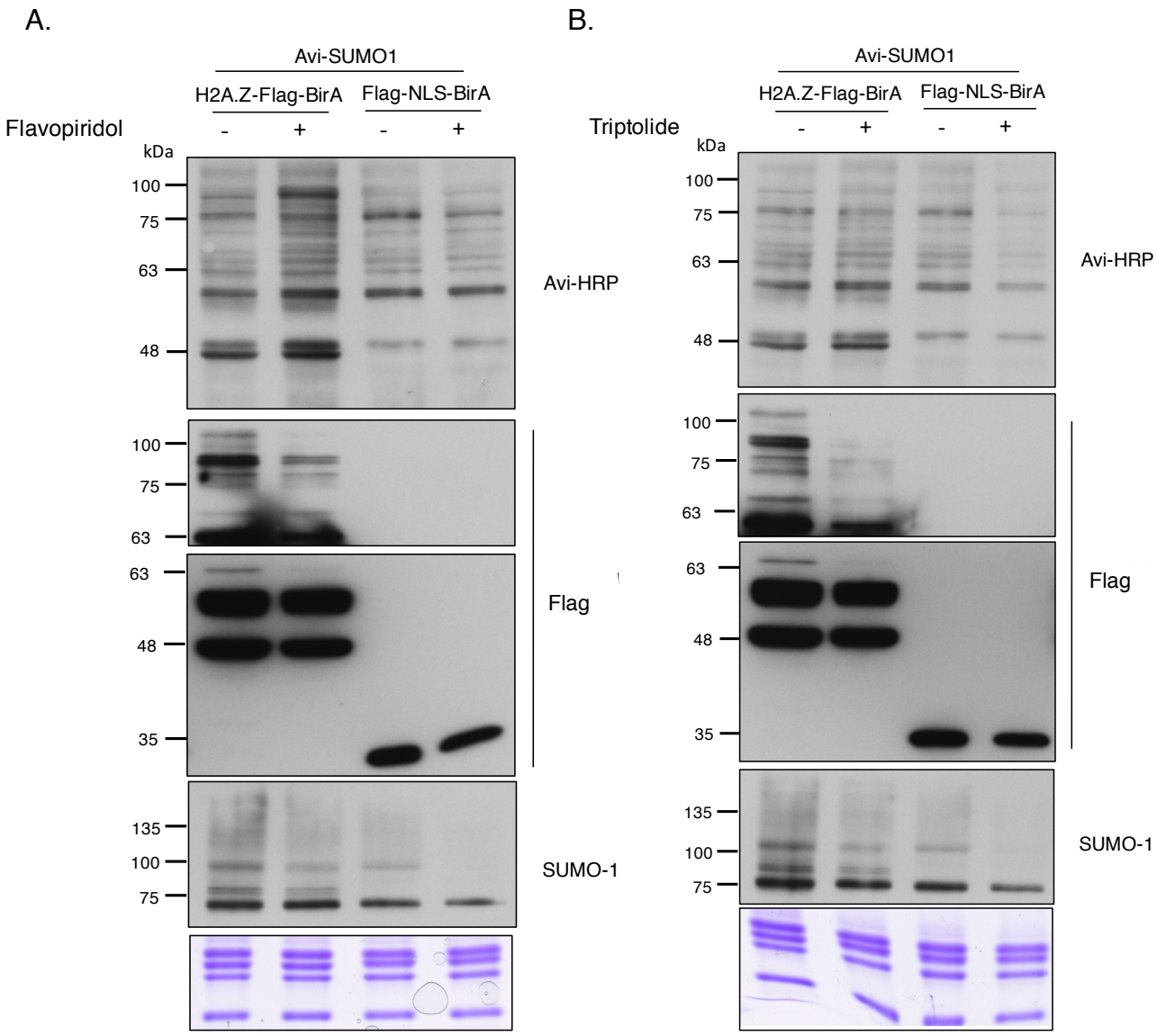



Figure 3-7. **Biotinylated proteins modified by Avi-SUMO1 potentially at the transcription start site are largely not dependent on transcription.** 293T cells were transfected with Avi-SUMO1 and H2A.Z-BirA or Flag-NLS-BirA, and nuclear fraction proteins were extracted for analysis with Western blot. Before protein extraction and 24-hours post-transfection, the cells were treated to either Flavopiridol (A) or Triptolide (B) for 24 hours to inhibit transcription. DMSO and water were used as the negative controls for the respective treatments. M2 Flag detected bands appropriately at 35 kDa, representing Flag-NLS-BirA, and at 50 kDa and 59 kDa, representing H2A.Z-Flag-BirA and ubiquitylated H2A.Z-Flag-BirA.

from the beads with 2x SB and analyzed with Western blot. The H2A.Z-Flag-BirA-dependent biotinylated bands at 45 kDa, around 63 kDa, and around 100 kDa were observed to be enriched in the eluate fraction (Figure 3-8, red asterisks). This signified that the proteins detected by Avi-HRP blotting of the total nuclear extracts were also captured by the Streptavidin beads, and that these bands do correspond with biotinylation.

Streptavidin immunoprecipitation was conducted again, but with increased washing stringencies to ensure that the proteins bound to the Streptavidin beads were indeed biotinylated. In addition, we also tried another method of extracting nuclear proteins by using RIPA and compared it to nuclear proteins extracted with 2x SB to ensure the reproducibility of the H2A.Z-Flag-BirA-dependent biotinylated bands. The input and eluate fractions were then analyzed in Western blot. H2A.Z-Flag-BirA-dependent biotinylated bands were still observed with the Avi-HRP blot and SUMO1 blot (Figure 3-9, red asterisks). However, SUMO1 antibody was consistently unable to detect the 45 kDa band, which might suggest that the corresponding protein might not be Avi-SUMOylated. Biotinylation cannot occur on proteins that do not contain the Avi-tag, and so SUMO1 antibody being unable to pick up the 45 kDa band might be due to Avi-tag or any modification on the protein itself occluding the recognition site of the antibody.

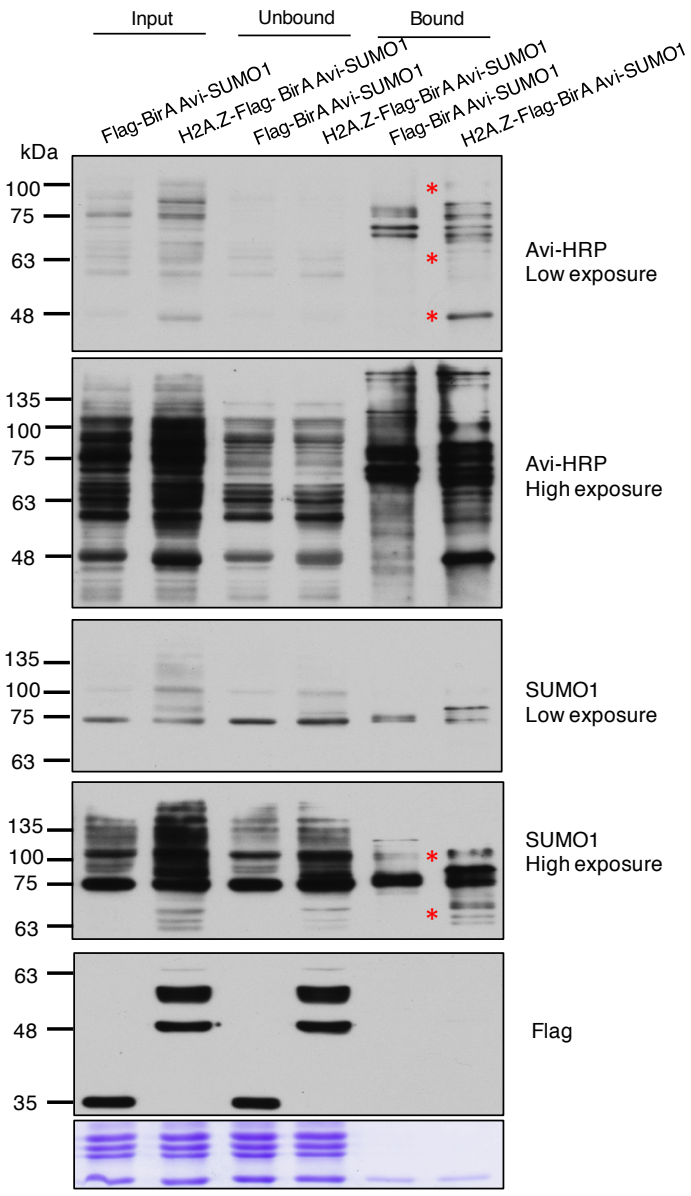



Figure 3-8. **Pull-down of biotinylated proteins using Streptavidin sepharose beads.** 293T cells were transfected with Avi-SUMO1 and either H2A.Z-BirA or Flag-NLS-BirA. Nuclear fraction was extracted with Buffer A + 0.2% Triton-X100 and 2x SB before diluting in Buffer D +PI for immunoprecipitation with Streptavidin sepharose beads. The unique bands observed with H2A.Z-BirA but absent in the negative control was enriched after IP but there were still bands observed in the negative control as well indicating non-specific binding of proteins to the beads. H2A.Z-Flag-BirA-dependent-Avi-SUMOylated targets are shown in red asterisks.

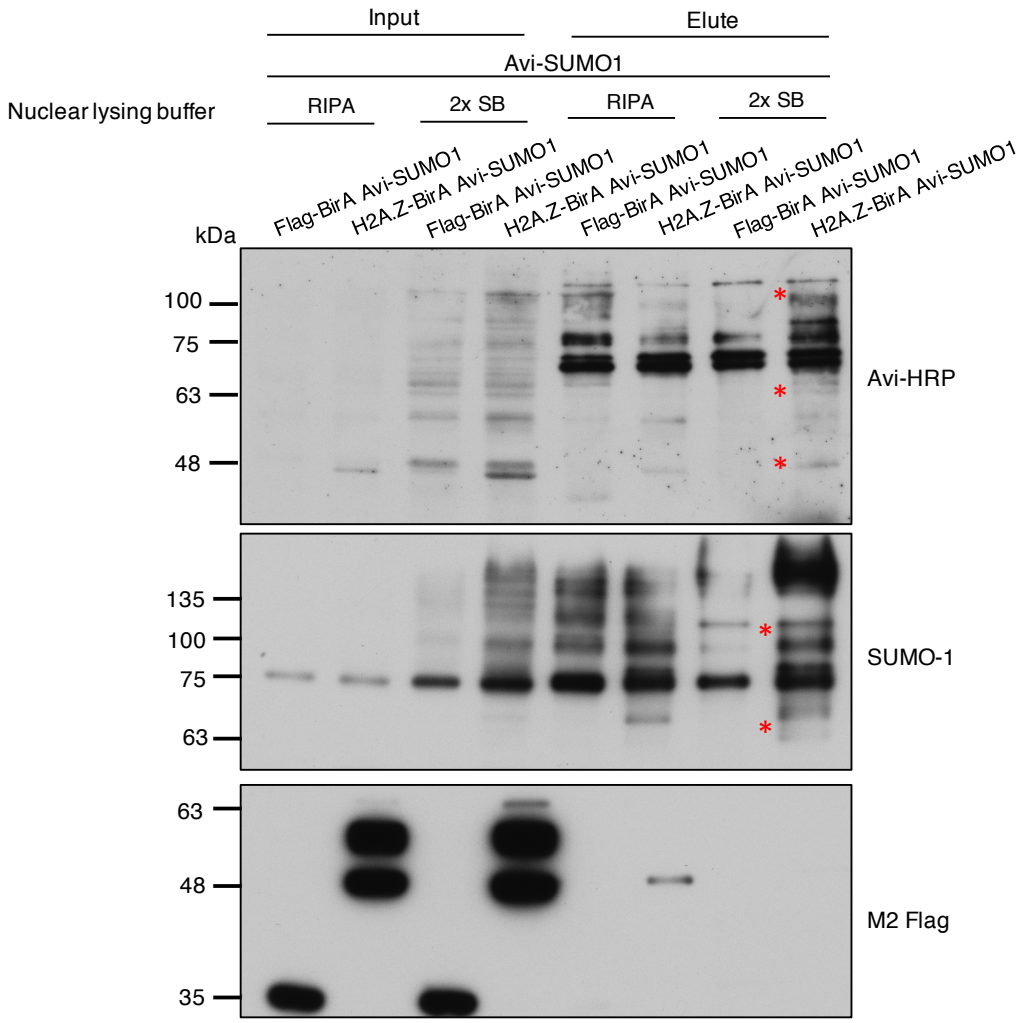



Figure 3-9. **Comparison of pull-down of biotinylated proteins using Streptavidin sepharose beads with different nuclear lysing buffer and use of more stringent bead washing steps.** 293T cells were transfected with H2A.Z-Flag or Flag-BirA with Avi-SUMO1 and nuclear fraction proteins were obtained 48-hours post-transfection with either RIPA buffer or 2x SB as indicated. Biotinylated proteins were then pulled-down with Streptavidin sepharose beads and washed more stringently with RIPA buffer +PI and 50 uM ammonium bicarbonate 3x each. Nuclear fraction with RIPA buffer seemed to be cleaner but the unique bands were more distinct when lysed with 2xSB. H2A.Z-Flag-BirA-dependent-Avi-SUMOylated targets are shown in red asterisks.

Altogether, our data support the H2A.Z-Flag-BirA-dependent biotinylated bands at 45 kDa and several other higher molecular weight proteins to be proteins that are potentially SUMOylated and biotinylated. Therefore, it would be of interest to identify those SUMOylation targets and whether they are normally recruited to transcription start sites and hence in proximity to H2A.Z.

### 3.3.5 Mass spectrometry identification of the putative H2A.Z-Flag-BirA-SUMOylated targets

The detection of the putative H2A.Z-Flag-BirA-SUMOylated proteins by Avi-HRP and also by Streptavidin beads pull down suggests that they are indeed biotinylated and SUMOylated. In order to prove that and determine the identity of the biotinylated bands corresponding to putative H2A.Z-Flag-BirA-SUMOylated targets, we hoped to use mass spectrometry (MS) as a tool for identification. The first round of samples sent for mass spectrometry (biochemical data not available) were from on-bead trypsinization of biotinylated proteins after nuclear extraction using 2x sample buffer. Analysis was conducted by Peter Liuni and are pending.

The second round of MS samples were more specific to the bands we wanted to identify, and in-gel trypsinization was used. RIPA buffer with optimized sonication was used instead to extract nuclear proteins and eluted with 2x SB before running on a pre-cast Novex gel. Silver stain was conducted with the samples beforehand to ensure proteins were present and to visualize bands to be excised. Bands that were cut in-gel

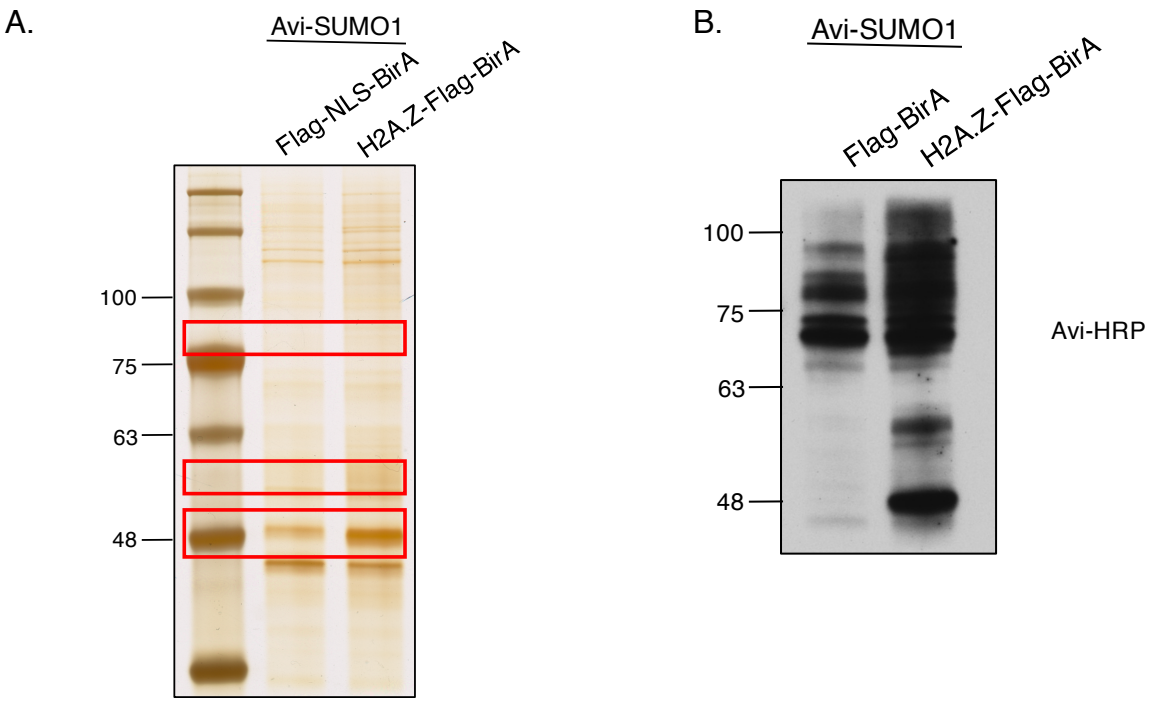



Figure 3-10. **Confirmation of pulled down biotinylated protein samples for mass spectrometry via silver stain and Western blots.** 293T cells were transfected as per usual with H2A.Z1-BirA or Flag-BirA and Avi-SUMO1 and nuclear lysates were collected. Unlike the prior round of mass spectrometry, nuclear fraction was lysed RIPA buffer and sonication this time. Proteins bound the Streptavidin beads were eluted with 2x SB and boiled for 5 minutes before running on a 10% gel. A small fraction of the elute samples were used for silver stain (A) to determine the band sizes to be excised for mass spectrometry and for Western blot analysis (B) to ensure the presence of the distinct band at approximately 45 kDa that was absent in the control (Flag-BirA) as well as a few other bands that seemed to be distinct in H2A.Z-BirA samples but not in the control. The red boxes shown in (A) represent the bands that were excised from the gel and sent for mass spectrometry analysis.

were highlighted in red on the silver stain (Figure 3-10A) and sent for MS analysis by Peter Liuni. Leftover samples were then used in Western blot analysis to confirm proper purification of biotinylated SUMOylated proteins, which was supported by presence of H2A.Z-Flag-BirA-dependent biotinylated proteins that run at 45 kDa, around 63 kDa, and around 100 kDa (Figure 3-10B).

### 3.4 Discussion

Studies have shown the importance of PTMs on influencing chromatin structure and function as well as regulating transcription. One such PTM involved in many nuclear functions is the covalent addition of the SUMO protein. Yeast studies have shown that H2A.Z is able to be modified by SUMO (Nathan *et al*, 2006; Kalocsay *et al*, 2009) while in mammalian cells, only H3 and H4 core histones were previously known to be SUMOylated (Hendriks *et al*, 2015). To test if we see the same phenomena for H2A core histone variant H2A.Z in mammalian cells, we utilized the Avi-tag BirA system to biotinylate Avi-SUMOylated targets and purify the biotinylated proteins. H2A.Z was not observed to be a significant target of SUMO modifications in 293T cells and HeLa cells. Although we did observe a band at 63 kDa potentially corresponding to biotinylated Avi-SUMOylated-H2A.Z-Flag-BirA, the intensity of the band was extremely low compared to other present bands. In addition, the same band was observed in the Flag-NLS-BirA control, indicating that the 63 kDa protein is not specific to H2A.Z-BirA. Altogether, our data showed that H2A.Z is not SUMOylated in mammalian cells. The presence of multiple bands corresponding to biotinylated H2A.Z-Flag-BirA-dependent Avi-SUMOylated proteins suggests that we could also use the Avi-tag BirA system to detect and identify SUMOylated proteins in proximity to H2A.Z. The presence of these additional bands dependent on H2A.Z-Flag-BirA suggests that the BirA ligase on H2A.Z was able to reach out to the Avi-tag on other SUMOylated proteins in proximity to H2A.Z and attach biotin to those SUMOylated proteins. We hence took advantage of the localization of H2A.Z-containing nucleosomes at the regions flanking the NFR or TSS to

detect other SUMO targets that could potentially be proteins involved in transcription. We noticed a clear band at 45 kDa that would always appear distinctly with H2A.Z-Flag-BirA but never in the negative control, suggesting the specificity of it to H2A.Z. In addition, there were additional bands above 63 kDa and around 100 kDa that were also present with H2A.Z-BirA and absent in the negative control. These multiple bands were consistently observed with Western blot, supporting the hypothesis that they are putative H2A.Z-Flag-BirA-dependent Avi-SUMOylated targets. Although H2A.Z mostly flanks the TSS, a caveat would be that all proteins biotinylated by H2A.Z-Flag-BirA cannot be assumed to be localized to the TSS or NFR as H2A.Z is not exclusively found at TSS.

Changes in SUMOylation states of specific proteins were observed in HeLa cells in response of heat shock (Golebiowski et al, 2009). We therefore conducted heat shock treatment to induce transcription of some target genes and determine if it would result in SUMOylation of H2A.Z or additional transcription factors. However, we were unable to detect any additional biotinylated-Avi-SUMOylated proteins after heat shock, suggesting that the H2A.Z was not modified by SUMO in response to heat shock. Putative H2A.Z-Flag-BirA-dependent Avi-SUMOylated targets were also not involved in heat shock response. Both SUMO and H2A.Z are individually known be involved in DNA damage repair responses (Hendriks & Vertegaal, 2015; Xu et al., 2012). Recently, a paper published their observations of H2A.Z2 histone variant isoform being SUMOylated in response to DNA damage in HeLa cells (Fukuto *et al*, 2018). To

determine if H2A.Z1 or H2A.Z2 is indeed SUMOylated in response to DNA damage in our hands, MMS treatment was conducted in both 293T and HeLa cell lines. We did not observe SUMOylation of H2A.Z2 in either cell lines. There were no modifications in the Western blot banding pattern before and after DNA damage for both H2A.Z1 and H2A.Z2 in both cell lines. This suggested that SUMOylation might not be involved in DNA damage in these cancer cell lines. However, the DNA damaging agent we used, MMS, is believed to stall replication forks to induce DNA damage (when it stalls replication forks, double-strand breaks can be formed) (Lundin *et al*, 2005), while SUMOylation of H2A.Z2 in response to DNA damage was observed through double-stranded breaks using ionizing radiation (Fukuto *et al*, 2018). This could potentially explain why we do not see the involvement of SUMO in facilitating H2A.Z2 exchange in nucleosomes in response to DNA damage.

We wanted to determine if these putative biotinylated H2A.Z-Flag-BirA-dependent Avi-SUMOylated proteins were dependent on on-going transcription, which would suggest their role in binding to the transcription start site and being SUMOylated for transcription. We used two transcription inhibitors, Flavopiridol and Triptolide, to ensure the results observed were consistent with each other. The distinct biotinylated SUMOylated bands observed in H2A.Z-Flag-BirA did not decrease with the inhibition of transcription. This suggests that the biotinylated SUMOylated proteins that were localized in proximity to H2A.Z might not contribute to on-going transcription. With that

said, these results are not conclusive as the successful occurrence of transcription inhibition by both drugs in 293T cells have yet to be validated due to time constraints.

To further ensure the distinct bands observed exclusively with H2A.Z-Flag-BirA were indeed biotinylated due to proteins modified by Avi-SUMO, Streptavidin beads were used to purify for biotinylated proteins as biotin has a strong affinity to Streptavidin. The detection of the bands at around 45 kDa, above 63 kDa, and 100 kDa with Avi-HRP after Streptavidin IP suggested they were indeed biotinylated and potentially SUMOylated, as BirA would only be able to biotinylate Avi-tag, which is fused to SUMO1. Furthermore, the faint band at 63 kDa corresponding to potential Avi-SUMOylated H2A.Z-Flag-BirA was no longer observed with Flag antibody after Streptavidin IP, indicating that H2A.Z is not SUMOylated in mammalian cells, or at least in 293T cells. This supported the postulation that the band at around 63 kDa observed with Flag antibody might have been the di-ubiquitylated form of H2A.Z-BirA. We then hoped to identify the putative H2A.Z-Flag-BirA-dependent Avi-SUMOylated targets we've consistently observed with Western blots and Streptavidin IPs with mass spectrometry (MS). We excised specific bands at the molecular weights previously mentioned to test whether MS could detect peptides corresponding to SUMO C-terminus joined to the target protein via isopeptide linkage. However, due to time constraints as well as equipment malfunction, we are unable to present any additional information about the identity of SUMOylated proteins in proximity to H2A.Z.

### 3.4.1 Future directions and implications

Going forward, we should also confirm with RT-qPCR that the transcription inhibition with Flavopiridol and Triptolide had indeed worked. This would provide support that the results we observed was actually due to the inhibition of transcription. The experiment can then be repeated to show that the data we observed is consistent and not due to chance. Additionally, we should also repeat the DNA damage experiment by inducing double-stranded breaks instead of stalling replication forks. One way we could achieve it is by using CRISPR-Cas9 with the guide RNA targeting Alu elements. As there are many Alu elements in the human chromosome, Cas9 would be able to induce double-stranded breaks at many sites, causing DNA damage. We could then mimic the conditions that Fukuto lab utilized to observe H2A.Z2 exchange facilitated by SUMO in response to DNA damage and determine if we could see the same effect in our hands. We could use more than one cell line, such as with 293T and HeLa cells, to ensure that SUMO facilitation of H2A.Z2 exchange in response to DNA damage is universal.

Once we've obtained data from the MS and analyzed them, we would be able to identify the SUMOylated proteins that were localized in proximity to H2A.Z. Additional biological and technical repeats for mass spectrometry could be conducted to ensure the results are reliable and accurate. Once the identities of these proteins are verified, we could learn more about the SUMO targets and the regulatory pathways they contribute to. In doing so, we would be able to understand the mechanistic role of SUMO in certain diseases and develop targeted therapeutics with less side effects.

## References


Adam, M., Robert, F., Larochelle, M., and Gaudreau, L. (2001). H2A.Z is required for global chromatin integrity and for recruitment of RNA polymerase II under specific conditions. Mol Cell Biol **21**(18)**,** 6270-9.

Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell **129**(4)**,** 823-37.

Bayer, P., Arndt, A., Metzger, S., Mahajan, R., Melchior, F., Jaenicke, R., and Becker, J. (1998) Structure determination of the small ubiquitin-related modifier SUMO-1. J Mol Biol. **280**, 275-286.

Bernstein, E., & Hake, S. B. (2006). The nucleosome: a little variation goes a long way. Biochemistry and Cell Biology, 84(4), 505–507. https://doi.org/10.1139/o06-085

Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Tetrahymena histone acetyltransfer- ase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell 84, 843–851.

Campos E.L. and Reinburg D. (2009). Histones: annotating chromatin. Annu Rev Genet. 43: 559-599.

Canny, M. D., Moatti, N., Wan, L. C. K., Fradet-Turcotte, A., Krasner, D., Mateos-Gomez, P. A., Zimmermann, M., Orthwein, A., Juang, Y.C., Zhang, W., Noordermeer, S.M., Seclen, E., Wilson, M.D., Vorobyov, A., Munro, M., Ernst, A., Ng, T.F., Cho, T., Cannon, P.M., Sidhu, S.S., Sicheri, F. and Durocher, D. (2017). Inhibition of 53BP1 favors homology-dependent DNA repair and increases CRISPR–Cas9 genome-editing efficiency. Nature Biotechnology, 36(1), 95–102. https://doi.org/10.1038/nbt.4021

Carr, A.M., Dorrington, S.M., Hindley, J., Phear, G.A., Aves, S.J. and Nurse, P. (1994). Analysis of a histone H2A variant from fission yeast: evidence for a role in chromosome stability. Mol Gen Genet. 245(5): 628-635.

Carruthers L.M., Bednar J., Woodcock C.L. and Hansen J.C. 1998. Linker histones stabilize the intrinsic salt-dependent folding of nucleosomal arrays: mechanistic ramifications for higher- order chromatin folding. Biochemistry. 37: 14776–14787.

Cheung, P. and Draker, R. (2009). Transcriptional and epigenetic functions of histone variant H2A.Z. Biochem Cell Biol. 87(1): 19-25.

Cheung, P. and P. Lau. (2009). Epigenetic regulation by histone methylation and histone variants. Mol Endocrinol. 19(3): 563-73.

Chupreta, S., Holmstrom, S., Subramanian, L. and Iniguez-Lluhi, JA. (2005). A small conserved surface in SUMO is the critical structural determinant of its transcriptional inhibitory properties. Mol Cell Biol. 25: 4272–4282.

Clarkson, M. J., Wells, J. R. E., Gibson, F., Saint, R., & Tremethick, D. J. (1999). Regions of variant histone His2AvD required for Drosophila development. Nature, 399(6737), 694–697. https://doi.org/10.1038/21436

Corujo, D. and Buschbeck, Marcus. (2018). Post-Translational Modifications of H2A Histone Variants and Their Role in Cancer. Cancers (Casel). 10(3): 59

Cote, J., Quinn, J., Workman, J.L., and Peterson, C.L. (1994). Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/ SNF complex. Science. 265, 53–60.

Desterro, J. M., Rodriguez, M. S., Kemp, G. D., and Hay, R. T. (1999) Identification of the enzyme required for activation of the small ubiquitin-like protein SUMO-1. J Biol Chem. **274**, 10618-10624

Dhall, A., Wei, S., Fierz, B., Woodcock, C. L., Lee, T. H., and Chatterjee, C. (2014) Sumoylated human histone H4 prevents chromatin compaction by inhibiting long-range internucleosomal interactions. J Biol Chem. **289**, 33827-33837

Dorigo, B., Schalch, T., Bystricky, K., and Richmond, T. J. (2003) Chromatin fiber folding: requirement for the histone H4 N-terminal tail. J Mol Biol. **327**, 85-96

Draker, R., Sarcinella, E. and Cheung, P. (2011) USP10 deubiquitylates the histone variant H2A.Z and both are required for androgen receptor-mediated gene activation. Nucleic Acids Res. 39(9):3529-3542.

Dryhurst, D., McMullen, B., Fazli, L., Rennie, P.S. and Ausio, J. (2012) Histone H2A.Z prepares the prostate specific antigen (PSA) gene for androgen receptor-mediated transcription and is upregulated in a model of prostate cancer progression. Cancer letters. 315(1):38-47.

Dunican, D.S., McWilliam, P., Tighe, O., Parle-McDermott, A. and Croke, D.T. (2002). Gene expression differences between the microsatellite instability (MIN) and chromosomal instability (CIN) phenotypes in colorectal cancer revealed by high-density cDNA array hybridization. Oncogene. 21(20):3253-3257.

Faast R., Thonglairoam V., Schulz T.C., Beall J., Wells J.R., Taylor H., Matthael K,

Rathjen P.D., Tremethick D.J. and Lyons I. 2001. Histone variant H2A.Z is required for early mammalian development. Curr. Biol. 11: 1183–1187.

Fan, J.Y., Gordon, F., Luger, K., Hansen, J.C. & Tremethick, D.J. (2002). The essential histone variant H2A.Z. regulates the equilibrium between different chromatin conformational states. Nat Struct Biol, 9: 172-176

Flotho, A. and Melchior, F. (2013) Sumoylation: a regulatory protein modification in health and disease. Annu Rev Biochem. **82**, 357-385

Fukuto, A., Ikura, M., Ikura T., Sun, J., Horikoshi, Y., Shima, H., Igarashi, K., Kusakabe, M., Harata, M., Horikoshi, N., Kurumizaka, H., Kiuchi, Y. and Tashiro, S. (2018). SUMO modification system facilitates the exchange of histone variant H2A.Z-2 at DNA damage sites. Nucleus. **9**(1): 87-94.

Gévry, N., Chan, H.M., Laflamme, L., Livingston, D.M. and Gaudreau, L. (2007). p21 transcription is regulated by differential localization of histone H2A.Z. Genes Dev. **21**(15): 1869-1881

Gill, G. (2004) SUMO and ubiquitin in the nucleus: different functions, similar mechanisms? Genes Dev. **18**, 2046-2059

Golebiowski, F., Matic, I., Tatham, M. H., Cole, C., Yin, Y., Nakamura, A., … Hay, R. T. (2009). System-Wide Changes to SUMO Modifications in Response to Heat Shock. Science Signaling. 2(72), ra24–ra24.

Gómez-del Arco, P., Koipally, J. and Georgopoulos, K. (2005). Ikaros SUMOylation: switching out of repression. Mol cell Biol. **25**(7): 2688-2697.

Guo, B., and Sharrocks, A. D. (2009) Extracellular signal-regulated kinase mitogen-activated protein kinase signaling initiates a dynamic interplay between sumoylation and ubiquitination to regulate the activity of the transcriptional activator PEA3. Mol Cell Biol. **29**, 3204-3218

Gutschner, T., Haemmerle, M., Genovese, G., Draetta, G. and Chin, L. (2016). Post-translational regulation of Cas9 during G1 enhances homology-directed repair. Cell Reports. 14: 1-12

Han, M., and Grunstein, M. (1988). Nucleosome loss activates yeast downstream promoters in vivo. Cell, 55, 1137–1145.

Hardy, S., Jacques, P.E., Gévry, N., Forest, A., Fortin, M. E., Laflamme, L., Gadreau L. and Robert F. (2009). The euchromatin and heterochromatin landscapes are shaped by antagonizing effects of transcription on H2A.Z deposition. PLoS Genet. **5**(10): e1000687

Hay, R. T. (2005) SUMO: a history of modification. Mol Cell. **18**, 1-12

He, X., Riceberg, J., Pulukuri, S. M., Grossman, S., Shinde, V., Shah, P., Brownell, J. E., Dick, L., Newcomb, J., and Bence, N. (2015) Characterization of the loss of SUMO pathway function on cancer cells and tumor proliferation. PLoS One. **10**, e0123882

Hendriks, I. A., D'Souza, R. C., Yang, B., Verlaan-de Vries, M., Mann, M., and Vertegaal, A. C. (2014) Uncovering global SUMOylation signaling networks in a site-specific manner. Nat Struct Mol Biol. **21**, 927-936

Hendriks, I. A., & Vertegaal, A. C. O. (2015). SUMO in the DNA damage response. Oncotarget, 6(18).

Heyer, W.D., Ehmsen, K.T. and Liu, J. (2010). Regulation of homologous recombination in eukaryotes. Annual Review of Genetics. 44:113–139.

Hoppe, J.B., Salbego, C.G. and Cimarosti, H. (2015). SUMOylation: Novel neuroprotective approach for Alzheimer's Disease? Aging Dis. **6**(5): 322-330.

Horikoshi, N., Sato, K., Keisuke, S., Arimura, Y., Osakabe, Akihisa., Hiroaki, T., Hayashi-Takanaka, Y., Iwasaki, W., Kagawa, W., Harata, M., Kimura, H. and kurumizaka, H. (2013). Structural polymorphism in the L1 loop regions of human H2A.Z1 and H2A.Z2. Acta Crystallogr D Biol Crystallogr. 69: 2431-2439.

Issar, N., Roux, E., Mattei, D., and Scherf, A. (2008) Identification of a novel post-translational modification in Plasmodium falciparum: protein sumoylation in different cellular compartments. Cell Microbiol. **10**, 1999-2011

Jackson J.D. and Gorozky M.A. 2000. Histone H2A.Z has a conserved function that is distinct from that of the major H2A sequence variants. Nucleic Acids Res. 28(19): 3811-3816.

Jensen, K., Santisteban, M. S., Urekar, C., & Smith, M. M. (2011). Histone H2A.Z acid patch residues required for deposition and function. Molecular Genetics and Genomics, 285(4), 287–296.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., & Felsenfeld, G. (2009). H3.3/H2A.Z double variant–containing nucleosomes mark “nucleosome-free regions” of active promoters and other regulatory regions. Nature Genetics, 41(8), 941–945. https://doi.org/10.1038/ng.409

John, S., Sabo, P. J., Johnson, T. A., Sung, M. H., Biddie, S. C., Lightman, S. L., Voss, T. C., Davis, S. R., Meltzer, P. S., Stamatoyannopoulos, J. A., and Hager, G. L. (2008). Interaction of the glucocorticoid receptor with the chromatin landscape.

Mol Cell **29**(5)**,** 611-24.

Kalocsay, M., Hiller, N. J., and Jentsch, S. (2009) Chromosome-wide Rad51 spreading and SUMO-H2A.Z-dependent chromosome fixation in response to a persistent DNA double-strand break. Mol Cell. **33**, 335-343

Kayne, P.S., Kim, U.J., Han, M., Mullen, J.R., Yoshizaki, F., and Grunstein, M. (1988). Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mat- ing loci in yeast. Cell. 55, 27–39.

Kirsh, O., Seeler, J. S., Pichler, A., Gast, A., Muller, S., Miska, E., Mathieu, M., Harel-Bellan, A., Kouzarides, T., Melchior, F., and Dejean, A. (2002) The SUMO E3 ligase RanBP2 promotes modification of the HDAC4 deacetylase. EMBO J. **21**, 2682-2691

Knezetic, J.A., and Luse, D.S. (1986). The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. Cell 45, 95–104.

Kobor M.S., Venkatasubrahmanyam S., Meneghini M.D., Gin J.W., Jennings J.L., Link A.J., Madhani H.D. and Rine J. 2004. A protein complex containing the conserved Swi1/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. PLoS Biol. 2:E131

Lawrence, M., Daujat, S. & Schneider, R. (2016). Lateral Thinking: How Histone Modifications Regulate Gene Expression. Trends Genet. 32:42–56. doi: 10.1016/j.tig.2015.10.007.

Li, S. J., and Hochstrasser, M. (1999) A new protease required for cell-cycle progression in yeast. Nature. **398**, 246-251.

Liang, Z., Potter, J., Kumar, S., Ravinder, N. and Chestnut, J.D. (2016). Enhanced CRISPR/Cas9-mediated precise genome editing by imrpoved design and delivery of gRNA, Cas9 nuclease, and donor DNA. J Biotech. 241: 136-146.

Lindberg, M. J., Popko-Scibor, A. E., Hansson, M. L., and Wallberg, A. E. (2010) SUMO modification regulates the transcriptional activity of MAML1. FASEB J. **24**, 2396-2404

Liu, H.W., Banerjee, T., Guan, X., Freitas, M.A. and Parvin, J.D. (2015). The chromatin scaffold protein SAFB1 localizes SUMO-1 to the promoters of ribosomal protein genes to facilitate transcription initiation and splicing. Nucleic Acids Res. **43**(7): 3605-3613.

Liu, X., Li, B., and Gorovsky, M.A. 1996. Essential and nonessen- tial histone H2A variants in Tetrahymena thermophila. Mol. Cell. Biol. 16: 4305–4311.

Lorch, Y., LaPointe, J.W., and Kornberg, R.D. (1987). Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. Cell 49, 203–210.

Lundin C, North M, Erixon K, Walters K, Jenssen D, Goldman AS, Helleday T (2005). "Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable in vivo DNA double-strand breaks". Nucleic Acids Research. **33** (12): 3799–3811.

Luger K., Mader A.W., Richmond R.K., Sargent D.F. and Richmond T.J. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 389:251–260.

Lyst, M. J., Nan, X., and Stancheva, I. (2006) Regulation of MBD1-mediated transcriptional repression by SUMO and PIAS proteins. EMBO J. **25**, 5317-5328

Malik, H. S., and Henikoff, S. (2003). Phylogenomics of the nucleosome. Nat Struct Biol **10**(11)**,** 882-91.

Marques, M., Laflamme, L., Gervais, A.L. and Gaudreau, L. (2010). Reconciling the positive and negative roles of histone H2A.Z in gene transcription. Epigenetics. **5**(4): 267-272

Matic, I., van Hagen, M., Schimmel, J., Macek, B., Ogg, S. C., Tatham, M. H., Hay, R. T., Lamond, A. I., Mann, M., and Vertegaal, A. C. (2008) In vivo identification of human small ubiquitin-like modifier polymerization sites by high accuracy mass spectrometry and an in vitro to in vivo strategy. Mol Cell Proteomics. **7**, 132-144

Matsuda, R., Hori, T., Kitamura, H., Takeuchi, K., Fukagawa, T., and Harata, M. (2010). Identification and characterization of the two isoforms of the vertebrate H2A.Z histone variant. Nucleic Acids Res. **38**(13)**.** 4263-73.

Miller, M. J., Barrett-Wilt, G. A., Hua, Z., and Vierstra, R. D. (2010) Proteomic analyses identify a diverse array of nuclear processes affected by small ubiquitin-like modifier conjugation in Arabidopsis. Proc Natl Acad Sci. **107**, 16512-16517

Monteiro, F. L., Baptista, T., Amado, F., Vitorino, R., Jerónimo, C., & Helguero, L. A. (2014). Expression and functionality of histone H2A variants in cancer. Oncotarget, 5(11). https://doi.org/10.18632/oncotarget.2007

Mukhopadhyay, D., and Dasso, M. (2007) Modification in reverse: the SUMO proteases. Trends Biochem Sci. **32**, 286-295

Murata, T., Hotta, N., Toyama, S., Nakayama, S., Chiba, S., Isomura, H., Ohshima, T., Kanda, T., and Tsurumi, T. (2010) Transcriptional repression by sumoylation of Epstein-Barr virus BZLF1 protein correlates with association of histone deacetylase. J Biol Chem. **285**, 23925-23935

Nathan, D., Ingvarsdottir, K., Sterner, D. E., Bylebyl, G. R., Dokmanovic, M., Dorsey, J. A., Whelan, K. A., Krsmanovic, M., Lane, W. S., Meluh, P. B., Johnson, E. S., and Berger, S. L. (2006) Histone sumoylation is a negative regulator in Saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone modifications. Genes Dev. **20**, 966-976

Neyret-Kahn, H., Benhamed, M., Ye, T., Le Gras, S., Cossec, J. C., Lapaquette, P., Bischof, O., Ouspenskaia, M., Dasso, M., Seeler, J., Davidson, I., and Dejean, A. (2013) Sumoylation at chromatin governs coordinated repression of a transcriptional program essential for cell growth and proliferation. Genome Res. **23**, 1563-1579

Nishibuchi, I., Suzuki, H., Kinomura, A., Sun, J., Liu, N.A., Horikoshi, Y., Shima, H., Kusakabe, M., Harata, M., Fukagawa, T., et al. (2014). Reorganization of damaged chromatin by the exchange of histone variant H2A.Z-2. International journal of radiation oncology. 89, 736-744.

Olins A.L. and Olins D.E. 1974. Spheroid chromatin units (v bodies). Science. 183: 330–332.

Orthwein A, Fradet-Turcotte A, Noordermeer SM, Canny MD, Brun CM, Strecker J, Escribano-Diaz C, Durocher D. (2014). Mitosis inhibits DNA double-strand break repair to guard against telomere fusions. Science. 344:189–193.

Pichler, A., Gast, A., Seeler, J. S., Dejean, A., and Melchior, F. (2002) The nucleoporin RanBP2 has SUMO1 E3 ligase activity. Cell. **108**, 109-120

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 8**(11)**: 2281-2308

Rodriguez, M. S., Dargemont, C., and Hay, R. T. (2001) SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. J Biol Chem. **276**, 12654-12659

Rosonina, E., Duncan, S. M., and Manley, J. L. (2010) SUMO functions in constitutive transcription and during activation of inducible genes in yeast. Genes Dev. **24**, 1242-1252

Santisteban M.S., Kalashnikova T., and Smith M.M. 2000. Histone H2A.Z regulats transcription and is partially redundant with nucleosome remodeling complexes. Cell. 103(3): p. 411-22.

Seeler, J. S., and Dejean, A. (2003) Nuclear and unclear functions of SUMO. Nat Rev Mol Cell Biol. **4**, 690-699

Schlabach, M. R., Luo, J., Solimini, N. L., Hu, G., Xu, Q., Li, M. Z., Zhao, Z., Smogorzewska, A., Sowa, M. E., Ang, X. L., Westbrook, T. F., Liang, A. C., Chang, K., Hackett, J. A., Harper, J. W., Hannon, G. J., and Elledge, S. J. (2008) Cancer proliferation gene discovery through functional genomics. Science. **319**, 620-624

Schones, D. E., Cui, K., Cuddapah, S., Roh, T. Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. Cell. **132**(5)**,** 887-98.

Shiio, Y., and Eisenman, R. N. (2003) Histone sumoylation is associated with transcriptional repression. Proc Natl Acad Sci. **100**, 13225-13230

Skilton, A., Ho, J. C., Mercer, B., Outwin, E., and Watts, F. Z. (2009) SUMO chain formation is required for response to replication arrest in S. pombe. PLoS One. **4**, e6750

Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X. and Zhang, F. (2016). Rationally engineered Cas9 nucleases with imrproved specificity. Science. 351**(6268)**: 84-88.

Spektor, T.M., Congdon, L.M., Veerappan, C.S. and Rice, J.C. (2011) The UBC9 E2 SUMO conjugating enzyme binds the PR-Set7 histone methyltransferase to facilitate target gene repression. PLoS ONE. 6:e22785

Tatham, M. H., Jaffray, E., Vaughan, O. A., Desterro, J. M., Botting, C. H., Naismith, J. H., and Hay, R. T. (2001) Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. J Biol Chem. 276: 35368-35374

Urnov FD, Rebar EJ, Holmes MC, Zhang HS, & Gregory PD. 2010. Genome editing with engineered zinc finger nucleases. Nat Rev Genet. 119: 636–646.

Vardabasso, C., Gaspar-Maia, A., Hasson, D., Pünzeler, S., Valle-Garcia, D., Straub, T., Keilhauer, E.C., Strub, T., Dong, J., and Panda, T. (2015). Histone Variant H2A.Z.2 Mediates Proliferation and Drug Sensitivity of Malignant Melanoma. Mol Cell. 59: 75-88.

Wang, J. (2012) Cardiac function and disease: emerging role of SUMO conjugation. Wiley Interdiscip Rev Syst Biol Med. **3**, 446-457

Wang, Y., and Dasso, M. (2009) SUMOylation and deSUMOylation at a glance. J Cell Sci. **122**, 4249-4252

Wilson, V. G., and Heaton, P. R. (2008) Ubiquitin proteolytic system: focus on SUMO. Expert Rev Proteomics. **5**, 121-135

Wilson, N. R., and Hochstrasser, M. (2016) The Regulation of Chromatin by Dynamic SUMO Modifications. Methods Mol Biol. **1475**, 23-38

Wu, W.-H., Alami, S., Luk, E., Wu, C.-H., Sen, S., Mizuguchi, G., … Wu, C. (2005). Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange. Nature Structural & Molecular Biology, 12(12), 1064–1071. https://doi.org/10.1038/nsmb1023

Wu, R.S., & Bonner, W.M. (1981). Separation of basal histone synthesis from S-phase histone synthesis in dividing cells. Cell. 27,321-330.

Xu, Y., Ayrapetov, M. K., Xu, C., Gursoy-Yuzugullu, O., Hu, Y., & Price, B. D. (2012). Histone H2A.Z Controls a Critical Chromatin Remodeling Step Required for DNA Double-Strand Break Repair. Molecular Cell, 48(5), 723–733.

Yang, S. H., and Sharrocks, A. D. (2004) SUMO promotes HDAC-mediated transcriptional repression. Mol Cell **13**, 611-617

Zlatanova J. and Thakar A. 2008. H2A.Z: view from the top. Structure. 16(2): p. 166-79.