Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Diane Mwelu Mwonga   April 2, 2021

Tumorigenic, oncogenic, and cell cycle effects of H2A/H3 oncohistone mutations explored through

structural modeling and analysis in *Saccharomyces cerevisiae*

by

Diane Mwelu Mwonga

Dr. Anita H. Corbett

Adviser

Department of Biology

Dr. Anita H. Corbett

Adviser

Dr. Jennifer Spangle

Co-Advisor/Committee Member

Dr. Eladio Abreu

Committee Member

Dr. José Soria

Committee Member

2021

Tumorigenic, oncogenic, and cell cycle effects of H2A/H3 oncohistone mutations explored through

structural modeling and analysis in *Saccharomyces cerevisiae*

By

Diane Mwonga

Dr. Anita H. Corbett

Adviser

Dr. Jennifer Spangle

Co-adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Department of Biology

2021

Abstract

Tumorigenic, oncogenic, and cell cycle effects of H2A/H3 oncohistone mutations explored through structural modeling and analysis in *Saccharomyces cerevisiae*
By Diane Mwelu Mwonga

Oncohistone mutations are primarily missense mutations that occur in the core histones genes that encode the building blocks of the nucleosome complex. The histone proteins that comprise nucleosomes are highly conserved through evolution and are critical for nearly all DNA transactions. An increased understanding of how these oncohistone mutations affect cellular properties such as growth, regulation, chromatin structure, and post-translational modifications (PTMs) of histones is critical to developing novel drug therapeutic agents to treat cancers. Further research in mutations in the histone H2A and H3 genes can facilitate the development of novel therapeutics for head, neck, and hematological cancers. In the present study, I employ two complementary approaches; 1) I analyze known oncohistones (H3) using a budding yeast model, and 2) explore potential novel oncohistones (H2A) modeled using PyMOL software to generate hypotheses about the consequences of these previously unstudied missense mutations. Through this study, we examine the consequences of histone H3 and H2A gene mutations, providing insight into oncohistone mutations in tumor and cancer pathways in these core histone proteins. H3 oncohistones were studied *in vivo* in a *Saccharomyces cerevisiae* model. H2A oncohistones were studied through remote and computational analyses of protein structure through nucleosome modeling. Taken together, these oncohistone studies could help to expand our understanding of how oncohistone mutations impact cell growth and support oncogenic transformation.

Tumorigenic, oncogenic, and cell cycle effects of H2A/H3 oncohistone mutations explored through

structural modeling and analysis in *Saccharomyces cerevisiae*

By

Diane Mwelu Mwonga

Dr. Anita H. Corbett

Adviser

Dr. Jennifer Spangle

Co-adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Department of Biology

2021

Acknowledgements

Table of Contents

Chapter 1: Introduction

The Nucleosome Complex and Histones

Cancer is one of the leading killers in the United States (Sidney, Go, & Rana, 2019). Although cancer is mainly characterized by uncontrolled cell growth, many mechanisms within the cell, including DNA processes such as transcription and regulatory functions, and post-translational modifications (PTMs) play a role in driving tumorigenesis and oncogenesis through mechanisms that are not yet well defined (Hanahan & Weinberg, 2011; Sarthy & Henikoff, 2019). The present study focuses on understanding how a class of cancer-causing mutations in histone genes, termed oncohistones (Nacev et al., 2019), may alter cellular biology to drive tumorigenesis.

Histone proteins comprise the nucleosome complex in DNA, consisting of core and linker histones that have DNA wrapped tightly around each nucleosome complex (McGinty & Tan, 2015). The nucleosome complex is central to the cell cycle because DNA is packaged as chromatin in order to fit into the cell nucleus (**See Figure 1**). This chromatin is tightly associated with histone proteins in the cell nucleus; hence, the nucleosome complex is the unit of chromatin folding. Furthermore, the nucleosome complex recruits enzymes that both modify chromatin structure and introduce post-translational modifications (PTMs) that modulate nucleosome stability and function (McGinty & Tan, 2015). As shown in Figure 2, nucleosomes are comprised of the four histone core proteins (H2A. H2B, H3, and H4), and each core protein is present in two copies, producing an octamer complex (McGinty & Tan, 2015). H2A/H2B as well as H3/H4 create a heterodimer protein complex (McGinty & Tan, 2015). The tails of these histones have an overall positive charge; thus, these histone proteins tightly associate with the negatively-charged phosphate backbone of DNA (McGinty & Tan, 2015). Consequently, the

nucleosome complex, in association with other chromatin modifying and regulating proteins dictates many aspects of cellular processes and can contribute to disruptions in various cell cycle mechanisms, and in turn, can affect oncogenesis through disruption of cellular pathways (Hanahan & Weinberg, 2011).

Oncohistones

Oncohistone mutations occur at high frequency and lead to the expression of mutant histones that exhibit oncogenic features (Nacev et al., 2019). Why these mutations in histone genes drive oncogenesis is not understood. In fact, in humans, histone genes are present in multi-gene arrays and mutation of a single one of these genes can cause cancer. For example, there are 15 copies of the histone H3 gene in humans (Shi, Wen, & Shi, 2017). This large number of copies of the histone genes in humans makes studying the functional consequences of these mutations challenging in humans.

While several known oncohistones (primarily H3 variants) have been identified and characterized, (Nacev et al., 2019), there could be additional novel oncohistone mutations that have not yet been discovered. Several known oncohistone mutations cause a surprising array of different types of cancers including pediatric glioma, head and neck cancer, and several different types of bone cancers (Lu et al., 2019). Specifically, H3K27M, H3G34R, and H3G34V known oncohistone missense mutations are implicated in gliomas, H3K36M in chondroblastoma, and H3G34W and H3G34L in bone tumors (Maze, Noh, Soshnev, & Allis, 2014; Qiu et al., 2018). It is not clear why these mutations cause cancers or why they cause the specific cancers they do. How the histone variants alter cellular function and drive oncogenesis has been challenging to tackle, at least in part because of the many histone genes present in humans.

Most known oncohistone mutations studied and characterized thus far primarily consist of mutations in the histone H3 gene; however, there is now evidence that mutations in the histone

H2B gene or the linker histone H1 can also drive cancer and these findings suggest that there could be an array of uncharacterized novel oncohistones that require further study (Bennett et al., 2019; Yusufova et al., 2021). Thus, in the present study, we utilized exome sequencing data from patients with hematological cancers to explore and understand the behaviors of possible novel oncohistone mutations within the H2A family. Experiments with *S. cerevisiae* provide a model to study the known H3 oncohistone mutations and understand how specific pathogenic amino acid substitutions could alter cellular function. While humans have 15 genes that encode histone H3 (Shi et al., 2017), budding yeast have only two genes that encode histone H3. These budding yeast genes are referred to as *HHT1* and *HHT2* (Johnson et al., 2015). Thus, budding yeast cells provide an ideal model system where histone variants can be analyzed as the sole copy of the essential histone proteins, as cells lacking either *HHT* gene are viable and thus engineering a single histone variant can produce cells that express only that histone variant.

For example, one well-studied histone missense mutation encodes H3K36M, which causes tumorigenesis in head and neck cancers (Nacev et al., 2019). The lysine at position 36 is methylated and thus changing this residue to methionine alters the post-translational modifications (PTMs) in this region. which is a key regulatory site in the histone tail (Cancer Genome Atlas, 2015; Sarthy & Henikoff, 2019). Thus, H3K36M can be utilized as a model known oncohistone mutation to understand the importance of H2A mutations and their impact on the chromatin structure, potential consequences for PTMs, and the overall protein structures both in the context of the isolated histone proteins and assembled nucleosomes. Fortunately, structures exist of DNA bound nucleosomes for both humans and budding yeast (Taguchi et al., 2017; White, Suto, & Luger, 2001), so we can readily compare how amino acid substitutions in specific positions are likely to alter the function of the protein or the complex (Sarthy & Henikoff, 2019).

<u>*Histone Post-translational Modifications (PTMs)*</u>

To better comprehend the potential consequences of amino acid substitutions that occur in oncohistones, it is necessary to understand the types of histone modifications that can occur in the four core histones (H2A, H2B, H3, and H4). A thorough understanding of histone modifications lends key insights that could help to explain how mutations drive tumorigenesis at the cellular level. Furthermore, understanding histone modifications allows for better predictions regarding the effects of novel oncohistones that have yet have to be studied in any model. Lysine and arginine amino acid residues on the N-terminal tails of histones can be modified through acetylation or methylation, (Bhaumik, Smith, & Shilatifard, 2007). Acetylation reduces the effects of positive charges by making N-terminal tails more negative, and phosphorylation, which can occur on serine, threonine, or tyrosine, has a similar effect (Bannister & Kouzarides, 2011; Nacev et al., 2019). Methylation does not affect charge behavior but impacts the steric interactions of N-terminal tails as it adds bulky methyl (CH_3) groups to these tail regions (Bannister & Kouzarides, 2011; Nacev et al., 2019). Taken together, these PTMs on histone tails regulate gene expression as they affect the nucleosome structure in close proximity or at histone tails (Nacev et al., 2019). Overall, PTMs of histones have substantial impact on gene expression, and thus can affect whether chromatin is transcriptionally active or repressed (Nacev et al., 2019). The question that remains is whether downstream biochemical and cellular process cause tumorigenesis and oncogenesis as a result of these oncohistone mutations, such as changes in RNA transcripts and protein products.

<u>*An S. cerevisiae model to study functional consequences of oncohistones*</u>

The CRISPR/Cas9 system was used to generate several histone H3 oncohistone models and explore the role of these oncohistones and functions *in vivo* in budding yeast cells. In humans, there are 15 copies of the H3 gene (Shi et al., 2017). We utilized a budding yeast model

as it only has two copies of the H3 gene, and the H3 gene for yeast has approximately 90%

identity to the human H3 protein (Johnson et al., 2015; McBurney et al., 2016). Overall, the goal

of the *in vivo* studies was to provide insight into how these histone H3 oncohistones alter cellular

function. Our principal objective was to determine the mechanisms of cellular growth and

regulatory events within the *S. cerevisiae* model. *S. cerevisiae* are well-equipped for homologous

recombination, allowing for the fragment containing mutant *HHT2* and the known oncohistone

mutation to be incorporated into *S. cerevisiae* as the sole copy of the essential histone H3 gene

(Rattray & Symington, 1995). An advantage of this system is that we can also employ a model

where the other histone H3 gene remains intact and test for dominant effects of oncohistone

mutations. Both models can be useful to study how missense mutations present in oncohistones

alter cellular properties, either as the sole copy of the histone H3 gene or in the presence of a

wild-type histone H3 gene, which closely aligns with the situation in patients. These functional

consequences can be analyzed by examining the growth of these oncohistone budding yeast

models on plates with various drugs that alter cell cycle progression, DNA repair pathways, or

various other cellular processes linked to cancer (Hyland et al., 2005).

Novel Oncohistones

Novel oncohistones refers to oncohistones that have yet to be modeled or studied. The

Oncohistone Project is a collaboration between the Corbett and Spangle Laboratories at Emory

University. This work builds on a collaboration with Foundation Medicine, where Dr. Spangle

obtained exome-sequencing data for a large cohort of patients with hematological cancers.

Hence, the data for modeling these H2A oncohistone mutations come from sequenced exons of

mRNA in cancer patients with different cancer types. These novel candidate oncohistone

mutations are the focus of the computational analysis described in this study. The initial step in

analyzing novel candidate oncohistones is to perform structural modeling to identify the specific

location of the amino acid substitution within the nucleosome complex. Through software analysis using PyMOL Version 2.0, these oncohistone missense mutations can initially be studied through structural modeling of the nucleosome to model the amino acid substitution. This structural modeling can be coupled with predictive software that provides information about whether this variant is likely to be damaging (Leslie, 2020). Following modeling and predictive analysis, the functional consequence of potential candidate novel oncohistone mutations in cell proliferation, cellular growth, and response to promotion or inhibition of the cell cycle can be studied *in vivo* through creating a *S. cerevisiae* model. Our group has developed an *in vivo* model of *S. cerevisiae* cells to evaluate the consequences of H3 oncogenic mutations, as was performed for known oncohistones. These functional assays include testing how oncohistone mutations alter histone function and the cell cycle through budding yeast cell experimentation. The evolutionary conservation of histone genes facilitates the ease of creating *in vivo* experimentation models using different model systems. In this study, we focus on two potential novel candidate oncohistones (H2A mutants) that show interesting properties from computational analyses and studied known H3 oncohistone phenotypes through cell culture systems to provide an introductory model of how candidate novel oncohistones can be studied *in vivo* after computational analysis.

In summary, I took a two pronged approach to explore why missense mutations in genes that encode histones cause cancer, meaning they function as oncohistones. First, I employed a budding yeast model to explore how several characterized oncohistone mutations in histone H3 alter cellular function. Second, I used modeling approaches to generate hypotheses about the potential consequences of novel candidate oncohistone mutations in the histone H2B gene. These combined approaches have the potential to provide insight into how single amino acid changes in

one out of 15 copies of a histone gene can drive tumorigenesis. Ultimately, we can combine

these approaches with studies in the Spangle laboratory that employ relevant cell culture systems

and mouse models, which could allow the pathways altered to be defined and targeted for novel

chemotherapeutics.

Chapter 2: Materials and Methods

Development of S. cerevisiae models to study oncohistones

All known oncohistone H3 variants were studied in *S. cerevisiae* budding yeast cells designed for H3 variants specifically. H3 mutants were generated including known oncohistone missense mutants H3K36R, H3K36M, H3G34W, H3G34L, H3G34R, and H3G34V. For mutagenesis for H3 variants, an *S. cerevisiae* strain deleted of the *HHT2* gene was employed. The *HHT2* gene was replaced with a Kanamycin resistant locus. Through a series of nested PCR reactions, a fragment with the KanMX cassette flanked by *HHT2* sequences was integrated into the genome to replace the endogenous *HHT2* gene. PCR confirmed the mutated *hht2* and known mutant oncohistone fragment were incorporated into budding yeast cells. The end strain of all known oncohistone mutations contained a functional wild type *HHT1* and were deleted for *HHT2*. Yeast mutagenesis was conducted following the budding yeast co-transformation protocol as described by (Duina & Turkal, 2017). This approach allowed us to introduce any missense mutation into the *HHT1* locus and this would be the only histone H3 gene present in these oncohistone models.

S. cerevisiae growth assays

Triplicates of serial dilution and spotting assays were performed for the *S. cerevisiae* growth assays. Each assay included wild-type control *S. cerevisiae* (two samples), and the following missense mutants: H3K36R, H3K36M, H3G34W, H3G34L, H3G34R, and H3G34V. All wild-type and known oncohistones mutations were inoculated overnight in liquid yeast extract peptone dextrose (YEPD) media at 30°C. Next, a 1:10 serial dilution was performed for the eight samples, all samples were vortexed before pipetting into a 96-well plate. Specifically, 200 µl of each sample was pipetted into Column 1. Then, 20 µl were extracted from Column 1,

and the 1:10 serial dilution was performed to Column 6. At the end of the serial dilution,

Columns 1-6 each had 180 µl of sample. Afterwards, well-plate samples were spotted onto agar

plates containing YEPD (control), yeast extract-peptone-glycerol (YPG), caffeine, formamide,

hydroxyurea, or phleomycin.

For the first spotting assay, samples were plated onto YEPD, caffeine, and formamide.

The YEPD plates were stored at three temperatures: 18°C (n=1), 30°C (n=2), and 37°C (n=1).

The remaining plates containing YPG, caffeine, formamide, hydroxyurea, and phleomycin were

stored at 30°C (n=5). All plates were incubated for approximately 3 to 5 days.

For the second and third spotting assays (n=8), The YEPD plates were stored at two

temperatures: 18°C (n=1) and 30°C (n=2). The remaining plates containing YPG, caffeine,

formamide, hydroxyurea, and phleomycin were stored at 30°C (n=5).

Novel oncohistone computational modeling

Novel oncohistones were compiled into a database for H2A, H2B, and H3 mutants. This

database contained exome-sequencing data for a large cohort of patients with hematological

cancers including patients' sample numbers, cancer type, altered gene within the various histones

(H2A, H2B, and H3 mutants), the type of alteration (point or truncation), and the description of

DNA and amino acid changes.

Through the NCBI Basic Local Alignment Search Tool (BLAST), UniProt FAST-All

(FASTA) sequences of the two-novel candidate oncohistone mutations H2AE57K and

and the 1:10 serial dilution was performed to Column 6. At the end of the serial dilution,

Columns 1-6 each had 180 µl of sample. Afterwards, well-plate samples were spotted onto agar

plates containing YEPD (control), yeast extract-peptone-glycerol (YPG), caffeine, formamide,

hydroxyurea, or phleomycin.

All plates with the aforementioned drugs were placed on paper that contained square

diagrams to allow for proper spotting alignment during pipetting. Each plate had a total of six

columns and eight rows for a total of 48 spots, and each spot contained 3.5 µl. Once all plates

were spotted, plates were air dried then stored in respective temperatures agar side up.

For the first spotting assay, samples were plated onto YEPD, caffeine, and formamide.

The YEPD plates were stored at three temperatures: 18°C (n=1), 30°C (n=2), and 37°C (n=1).

The remaining plates containing YPG, caffeine, formamide, hydroxyurea, and phleomycin were

stored at 30°C (n=5). All plates were incubated for approximately 3 to 5 days.

For the second and third spotting assays (n=8), The YEPD plates were stored at two

temperatures: 18°C (n=1) and 30°C (n=2). The remaining plates containing YPG, caffeine,

formamide, hydroxyurea, and phleomycin were stored at 30°C (n=5).

Novel oncohistone computational modeling

Novel oncohistones were compiled into a database for H2A, H2B, and H3 mutants. This

database contained exome-sequencing data for a large cohort of patients with hematological

cancers including patients' sample numbers, cancer type, altered gene within the various histones

(H2A, H2B, and H3 mutants), the type of alteration (point or truncation), and the description of

DNA and amino acid changes.

Through the NCBI Basic Local Alignment Search Tool (BLAST), UniProt FAST-All

(FASTA) sequences of the two-novel candidate oncohistone mutations H2AE57K and

H2AE57Q were compared against humans, flies, and mice sequences within the databank. These

oncohistone mutations are located in the altered gene of HIST1H2AM (*BLAST: Basic Local*

Alignment Search Tool n.d.; *UniProt*, n.d.). Sequence comparisons demonstrated that the residue

at position 57 is evolutionarily conserved, meaning that we could model both these changes and

then eventually model these changes in budding yeast.

Next, H2AE57K and H2AE57Q mutations were modeled through the computational

analysis software PyMOL Version 2.0. The Protein Data Bank (PDB) file PDB 5X7X (Taguchi

et al., 2017) was obtained from the Research Collaboratory for Structural Bioinformatics

(RCSB), as this PDB file best showed the protein of interest for these novel candidate H2A

mutants in humans (Bank, n.d.). This file was imported into PyMOL Version 2.0, and the

specific mutagenesis of H2AE57K and H2AE57Q were performed in two separate analyses.

Pictures of the various orientations of the mutations were taken. This enabled observations to be

made regarding the polarity, steric interactions, size, bulkiness, and flexibility of amino acid

residues consequent to oncohistone mutagenesis. Additionally, this allowed observation of

possible protein interactions taking place after mutagenesis such as acetylation, phosphorylation,

and methylation.

Concurrently with modeling, an mCSM-P1P2 predictive analysis was performed through

Biosig Lab to determine the changes in binding affinity, and various chemical interactions (clash,

hydrophobic, hydrogen bonding, and polar) between the wild-type and mutant H2A novel

candidate oncohistones studied (*MCSM-PPI2 | Home*, n.d.).

Chapter 3: Results

A Budding Yeast Model to Characterize Know Oncohistone Proteins

The first set of experiments I performed employed a budding yeast model to explore the cellular pathways impacted by a set of known oncohistone variants. The studies employ budding yeast as a model because with only two copies of each core histone gene (Johnson et al., 2015; McBurney et al., 2016), this system can be used both to explore how the missense mutations in oncohistones that drive caner alter the function of the histone when it is the sole cellular copy of that histone or to test for dominant effects when the other histone gene is present.

H3 oncohistone variants H3K36R, H3K36M, H3G34W, H3G34L, H3G34R, and H3G34V show growth defects on media containing YP glycerol, hydroxyurea, phleomycin, caffeine, and formamide

We employed a budding yeast model to assess how missense mutations in histone genes alter cellular function. The recombination approach described in Materials and Methods was used to create a series of H3 oncohistone variants in cells. For this model, the budding yeast cells contain one wild-type histone H3 gene *HHT1* and one mutant histone H3 gene *hht2*. This model allows us to test whether oncohistone mutations confer dominant phenotypes over a wild-type gene as is seen in patients where only one out of 15 histone H3 genes is mutated.

To assess what biochemical and cellular pathways were affected, *S. cerevisiae* models of the following oncohistone missense mutations: H3K36R, H3K36M, H3G34W, H3G34L, H3G34R, and H3G34V were spotted onto agar plates to examine growth on: YEPD (control), YP glycerol (YPG), hydroxyurea, phleomycin, caffeine, or formamide (**Figure 6**). YEPD control plates incubated at 18°C (n=1) and 30°C (n=2) showed normal growth as expected, where all *S. cerevisiae* models (Wild-type and H3 Mutants) showed growth at all dilutions examined, with

less growth across Columns 1-6 due to the 1:10 serial dilution (most cells in Column 1, 10-fold dilution in Column 2, etc…). In these studies, we also included a very well characterized oncohistone H3K36M variant as well as another control H3K36R, which is not an oncohistone, but a conservative change in this key residue.

The YPG plate incubated at 30°C (n=1) showed poor growth for all samples analyzed (Wild-type and H3 Mutants). The poor growth of all *S. cerevisiae* samples on YP Glycerol was expected as this media does not support growth of cells lacking mitochondrial DNA (e.g., petite or *pet* (p^-) mutant (Klein, Swinnen, Thevelein, & Nevoigt, 2017). Thus, *S. cerevisiae* strains on the YPG plate show on average smaller colonies across the plate in comparison to YEPD control plates, this is most evident in Columns 5 & 6. However, no significant differences were observed between the control wild-type and any of the oncohistone models, suggesting that oncohistone mutations do not significantly alter mitochondrial DNA in budding yeast. The hydroxyurea, phleomycin, caffeine, and formamide plates incubated at 30°C (n=4) reveal some interesting differences between the control wild-type cells and the oncohistone mutants. Hydroxyurea, phleomycin, caffeine, and formamide affect the following pathways: DNA replication pathway, cell proliferation, stress response, and RNA metabolism, respectively (Li et al., 2014). Interestingly, the caffeine plate incubated at 30°C (n=1) showed deficient growth almost similar to that seen with the YP Glycerol plate. These results could indicate that the stress response pathway plays a bigger role in growth deficiency compared to the other drugs of hydroxyurea, phleomycin, and formamide. Specifically, for the hydroxyurea treatment, which impacts the cell cycle (Kaushal & Vashishat, 1982), at 30°C (n=1), all *hht2*H3K36 & *hht2*H3G34 mutants in Columns 3-6 show very poor growth. In contrast, the *hht2* H3G34R mutant shows the most robust growth on plates containing hydroxyurea out of all *hht2*-H3G34 mutants. This observation

was unexpected, as this is a substitution at the same position as the other H3G34 mutants. Furthermore, H3G34R and H3G34V are both implicated in brain tumors (glioblastoma), and the H3G34V mutant did not show a comparable growth pattern to the H3G34R mutant (Qiu et al., 2018). Interestingly, the H3G34L mutant and the H3G34V oncohistone models show similar growth patterns, which likely reflects the change of glycine to a hydrophobic amino acid. The HG34R shows the strongest change in growth properties and this is the only substitution that changes this residue to a charged amino acid. Overall, the finding that different substitutions at the same position differentially alter yeast cell growth could help to explain why different mutations at this position drive different types of cancer.

For the phleomycin treatment, which induces DNA damage (Koy et al., 1995), at 30°C (n=1), all H3 mutant colonies except H3G34R show very decreased growth compared to the control. This suggests that cell proliferation is mildly impacted by H3 mutants with the exception of H3G34R. This finding also raises the question of how the H3G34R mutant operates differently from H3G34V mutant that also causes brain tumors and the H3G34W/L mutant. In the presence of the caffeine, which impacts the TOR growth regulation pathway (Kuranda, Leberre, Sokol, Palamarczyk, & François, 2006), at 30°C (n=1), we observe similar growth patterns as seen on the phleomycin plates. The H3G34R and H3G34V mutants show the least growth out of the slate of H3G34 mutants, but the worst growth profile is seen for the H3G34R mutant. Additionally, the H3K36R mutant shows a decrease in growth that is more notable than H3K36M. This result is interesting, because the caffeine plates suggest that introduction of an arginine residue alters growth in all conditions examined but particularly in caffeine.

Finally, in the presence of formamide, which alters RNA metabolism (Hoyos-Manchado et al., 2017), at 30°C (n=1), all H3K36 and H3G34 mutants show poor growth. This result

suggests that H3 mutants are highly impacted by defects in RNA metabolism and these H3

mutants could affect production of RNA transcripts (Qiu et al., 2018). Similar results were seen

for the other two experiments.

Novel Oncohistones

The second part of my analysis used structural modeling to make predictions about the possible functional consequences of potential novel oncohistones. These potential oncohistones were identified by exome sequencing of patients with hematological cancers in collaboration with Foundation Medicine. While candidate oncohistone variants were identified in all histone genes sequences, my studies focused on specific oncohistone candidates in Histone H2A. Ultimately, we would like to test these candidate oncohistones in the budding yeast model, so I focused my analysis on specific missense mutations that occur in a residue that is conserved between budding yeast and humans H2AE57.

Human Histone H2A shares 73% identity with S. cerevisiae Histone H2A

As described in Material and Methods, we utilized NCBI BLAST to assess amino acid identity between Human Histone H2A Type 1 and *S. cerevisiae* Histone H2A.1 (*BLAST: Basic Local Alignment Search Tool* n.d.; *UniProt*, n.d.). This BLAST search was repeated for other common model organisms including *D. melanogaste*r and *M. musculus*. The NCBI BLAST query search showed that there was a 73% identity match in sequence between humans and *S. cerevisiae* Histone H2A (**Figure 7**; *BLAST: Basic Local Alignment Search Tool* n.d.; *UniProt*, n.d.). Furthermore, this Histone H2A identity match in sequence between human and *D. melanogaster* decreased to 56% identity match and increased to 97% between humans and *M. musculus* (**Figure 7**). Thus, these results further validated *S. cerevisiae* use as a model to study oncohistones *in vivo*. Furthermore, these results show that the Histone H2A sequence is conserved across multiple organisms.

Histone H2A residue E57i's accessible at the surface of the Octamer Protein Complex

While I studied a number of candidate oncohistones, I focus this report on two missense mutations that change the glutamic acid at position 57 in Histone H2A to either a lysine (H2AE57K) or a glutamine (H2AE57Q). I chose this particular oncohistone for my focused analysis because we identified multiple changes at the same position, i.e., E57 changed to both lysine and glutamine, suggesting that changes at this position could be functionally important. In addition, the introduction of a lysine could create a site for novel post-translational modifications (**See Figure 5**) due to the presence of the amino group.

To examine the location of E57 in H2A and explore the possible consequences of amino acid substitutions at this position, we utilized the protein modeling software PyMOL Version 2.0. All analyses were based on the human nucleosome structure PDB 5X7X (Taguchi et al., 2017), and mutagenesis was performed on the altered HIST1H2AM gene within Chain C. Through PyMOL Version 2.0, we could visualize the overall nucleosome structure of WT H2AE57, H2AE57K, and H2AE57Q (**Figure 8**). Furthermore, we found that E57 is located on the surface of the histone octamer (**Figure 9**). This finding suggests that introduction of a lysine at this position on the octamer protein complex surface could provide a *de novo* site for PTMs such as acetylation and methylation. Furthermore, the change of the wild type of the central glutamic acid residue (negatively charged) to the lysine (positively charged) or glutamine (neutral) could change protein interactions between the central residue and other proteins within the overall chromatin. Thus, one hypothesis is that novel PTMs at least in H2AE57K or overall charge of the central residue could negatively affect downstream cellular pathways.

H2AE57K & H2AE57Q Mutant exhibit various steric clashes, van der Waals, hydrophobic, polar, and hydrogen bonding forces with surrounding residues

We utilized predictive mutational output from mCSM-P1P2 analysis to guide our PyMOL modeling. The WT H2AE57 exhibit minimal steric clashes with surrounding residues, the biggest steric clashes between the central glutamic acid residue and the surrounding residues of glutamine and phenylalanine (**Figure 10**). This result was as expected as steric classes were between the nitrogen/oxygen atoms in the glutamic acid and the surrounding residues nitrogen/oxygen atoms. A small amount of Van der Waals forces was seen as well (Figure 10), primarily between polar (nitrogen/oxygen atoms) and nonpolar (hydrogen on carbon atoms) regions of the central glutamic acid and the surrounding residues polar (nitrogen/oxygen atoms) and nonpolar (hydron on carbon atoms) regions. Furthermore, the vast majority of hydrophobic interactions were detected been the central glutamic acid residue nonpolar regions and the phenylalanine nonpolar regions (**Figure 10**). This result was expected due to the close association of glutamic acid and phenylalanine residues. Many polar interactions specifically between carbon and oxygen atoms were seen across all residue association (**Figure 10**). Lastly, significant hydrogen bonding was observed (**Figure 10**). This result was expected as many residues had highly electronegative nitrogen and oxygen atoms available to hydrogen bond with hydrogens of surrounding residues. In comparison to the WT H2AE57, the H2AE57K amino acid substitution showed slightly less steric clash, Van der Waals, and polar interactions with surrounding residues; similar hydrophobicity to WT H2AE57 was observed as well (**Figure 10**). Overall, in comparison to the WT H2AE57, the H2AE57Q mutant showed similar steric clash, Van der Waals, hydrophobic, and polar interactions with surrounding residues (**Figure 10**). Both the H2AE57K and H2AE57Q mutant showed decreased hydrogen bonding compared to WT H2AE57K (**Figure 11**). We hypothesize that this result allows for novel PTMs, including

acetylation and/or methylation of the lysine introduced due to the weakening of hydrogen bonding, which acts as a form of critical infrastructural bonds, as well as the substantial steric clash with the aromatic ring of phenylalanine. Thus, this conformational change in the glutamic acid on the surface of the histone octamer could alter local conformation and possibility introduce the opportunity for novel PTMs (H2AE57K), which could alter the surface of the nucleosome. Furthermore, charge changes for both mutants at the nucleosome surface could alter interactions with chromatin-interacting proteins.

Chapter 4: Discussion

This study used two complementary approaches to probe the functional changes that can occur in histone proteins to turn these essential chromatin organizing proteins into nefarious causes of tumors and oncogenesis. We employed a functional assay modeling previously characterized oncohistone variants of the histone H3 protein in budding yeast and discovered that specific changes alter different cellular pathways. We then employed structural modeling to examine two candidate oncohistone changes in the histone H2A protein. This latter approach allows us to screen through multiple candidate mutations and both develop hypotheses about how the changes could alter histone and/or nucleosome function and to rapidly screen a large series of candidate oncohistones. Taken together, these approaches provide a platform for discovering and analyzing oncohistones that could define cellular pathways altered to target in new therapeutic approaches.

A major challenge in exploring the functional consequences of the missense mutations present in oncohistones is the fact that human histone genes occur in many copies. For example, the human histone *H3* gene is present in 15 copies and other histone genes are present in similar copy numbers (Brown, 2001). These genes are quite similar to one another as they arose from ancient gene duplication events (Brown, 2001), but the H3.3 variant is the most common (Shi, Wen, & Shi, 2017). In fact the H3.3 variant is very similar to the the histone H3, typically exhibiting 4 to 5 amino acid changes (Shi et al., 2017). These H3.3 variants most likely are similar to the H3 histone due to genome duplication followed by divergence (Szenker, Ray-Gallet, & Almouzni, 2011). Similar trends are also seen across H2A variants. The question that remains is why H3.3 variants are implicated in various different cancer types such as H3K27M, H3G34R, and H3G34V in gliomas, H3K36M in chondroblastoma, and H3G34W and H3G34L

in bone tumors (Maze et al., 2014; Qiu et al., 2018). Currently, there is not much information about which cell types express any of the individual H3 variants, eliminating the simple explanation that some oncohistones could simply be expressed in specific cell types. Many of the best characterized oncohistones occur in the histone variant H3.3 (Nacev et al., 2019). As H3.3 variants are nearly identical to other histone H3 proteins, the hypothesis that mutations in different histone variants cause tumorigenesis and/or oncogenesis in specific cell types remains quite unlikely (Shi et al., 2017). Studies to assess the functional consequences of a missense mutation that drives cancer in a single copy of a gene that is present in 15 copies are both conceptually and technically challenging. For this reason, we can employ the yeast model, which contains only two copies of each histone gene to explore either dominant effects (when we have one wild-type and one mutant present) or recessive effects (when the mutant histone gene is supplied as the only functional copy of the histone gene). Thus, yeast studies such as the ones performed in this study, can provide insight into which specific cellular pathways are impacted by individual missense mutations that drive distinct forms of tumorigenesis or oncogenesis in patients.

One objective of this study was to exploit a budding yeast model to examine how known H3 oncohistone mutations alter cell growth and begin to define what cellular pathways might be altered by the cancer-causing missense mutations. This objective was achieved through analyzing the growth of known H3 oncohistones that alter residues K36 (H3K36M, H3K36R) and G34 (H3G34L, H3G34W, H3G34R, H3G34V) created in *S. cerevisiae*. We analyzed the growth of these budding yeast oncohistone models on rich growth media (YEPD) and also on plates containing a variety of additives or drugs to uncover cellular pathways that could be altered by the oncohistone mutations. The drugs would provide data on how missense mutations

in histone genes impact cellular function within the biological pathways of DNA replication (hydroxyurea), DNA damage response (phleomycin), stress response (caffeine), and RNA metabolism (formamide) (Li et al., 2014). Furthermore, because all *S. cerevisiae* known *hht2*-H3 mutants had a functional copy of the histone H3 gene, *HHT1*, we could assess any dominant growth defects of these H3 mutations in the presence of the wild type histone H3 gene, a situation that mimics patients where a single copy out of multiple copies of histone H3 is altered.

Results showed that the caffeine plate exhibited poor growth that parallel the YP Glycerol, with the *hht2*-H3K36M and *hht2*-H3G34R mutant having the worst growth profile out of all known K36 & G34 mutants, suggesting caffeine plays a significant role in growth deficiency of known H3 mutants compared to the other drugs of hydroxyurea, phleomycin, and formamide. Furthermore, regarding hydroxyurea treatment, which affects the cell cycle, overall, all known *hht2*-H3K36 and *hht2*-H3G34 mutant plates showed poor growth, with *hht2*-H3G34R mutants exhibiting the best growth out of all *hht2*-H3G34 mutants (Kaushal & Vashishat, 1982). In contrast, for phleomycin treatment, which stimulates DNA damage, all known *hht2*-H3K36 and *hht2*-H3G34 mutants showed decreased growth compared to the wild type control, with *hht2*-H3G34R mutants showing the worst growth out of all *hht2*-H3G34 mutants (Koy et al., 1995). Results of the phleomycin treatment closely parallel the caffeine treatment. Finally, for the formamide treatment, which changes RNA metabolism (Hoyos-Manchado et al., 2017), overall, all known *hht2*-H3K36 and *hht2*-H3G34 mutant plates showed poor growth. Thus, these studies suggest that the H3G43R mutant operates significantly differently from other mutants that have single amino acid substitutions at the same position. A very interesting finding from this analysis is that single amino acid substitutions at the same position G34 clearly have distinct effects on cellular growth, providing possible insight as to why mutations lead to different tumor

and cancer types. Lastly, through the caffeine and formamide treatments we concluded that all known H3 oncohistone mutants are highly impacted in terms of their stress response and RNA metabolism pathways, which are most likely critical in downstream cellular pathways of RNA processing and protein translation that could contribute to tumorigenesis and/or oncogenesis.

Of particular interest for the H3 G34 oncohistones is the observation that different amino acid substitutions at this same position cause distinct cancers. While H3G34R and H3G34V cause high grade glioblastomas, devastating brain cancers, H3G34L and H3G34W cause cancers of the bone, including chonodroblastomas and giant cell tumors of the bone (Mohammad & Helin, 2017). Why these different amino acid substitutions at a single position cause distinct cancers remains a mystery. In fact, the finding that HG34L and H3G34V cause distinct cancers is very surprising because valine (V) and leucine (L) are very similar amino acids. One might think that studies such as our modeling in yeast could help to uncover why these different changes cause distinct pathology.

Other studies have analyzed the role of the H3G34 mutant. H3G34 lies close to the K36 residue and could impair K36 methylation, and an amino acid substitution such as tryptophan or arginine can cause changes to the K36 residue that affect the accessibility of K36 (Qiu & Han, 2021). Additionally, in yeast fission experiments, H3G34R mutants on hydroxyurea showed a poor growth profile and was found to have issues with homologous recombination and genome stability in *S. pombe* (Lowe et al., 2021; Qiu & Han, 2021; Yadav et al., 2017). Hence this could explain why *hht2*-H3G34R showed different growth properties than the other mutants at this position in our studies. This could also explain why H3G34R exhibited the worst growth on phleomycin and caffeine as well as an overall low growth in formamide. Future experiments could be performed to assess whether H3G34R affects tumor and brain cancers through causing

altering PTMs (acetylation and methylation) of the N-terminal regions of H3K36. These future studies could provide important information for the next step in assessing downstream cellular pathways. In addition, knowledge about what specific pathways are affected could help to define new therapies to treat cancers caused by specific oncohistones.

In our present study we only studied *S. cerevisiae* known H3 oncohistone mutations. Future experiments should utilize other models such as *D. melanogaster*, *M. musculus*, or *S. pombe*. Furthermore, adding in additional drug treatments for the growth assay such as camptothecin, which blocks DNA topoisomerase activity could provide additional information (Qiu & Han, 2021; Yadav et al., 2017). Further studies of these known H3 oncohistone mutations in cell culture models or mice xenograft models could also help to define why specific oncohistones drive distinct types of cancer. Finally, for these studies, we employed a budding yeast model where a second histone H3 gene was present. We could extend our analysis to cells where the oncohistone variant analyzed is the sole copy of the histone H3 present. Such studies would define the functional consequences of the missense mutations present in these oncohistone models. Perhaps this approach would reveal more similarities or differences between the histone H3G34 models that cause distinct forms of cancer.

Few if any experiments have utilized PyMOL to model novel candidate H2A oncohistones as we did in this collaborative project between Foundation of Medicine and the Corbett/Spangle lab. However, some studies have employed PyMOL analysis with known H3 oncohistone mutations such in the H3K36 and H3G34 families. These studies have showed protein interactions such as the interaction of H3K36M and H3G34 with the SET domain, which exhibits methyltransferase activity, and affects chromatin structure (Alvarez-Venegas, 2014; Zhang et al., 2017). These past studies demonstrate that PyMOL provides a platform for further

analysis of potential oncohistone candidates in other histone proteins such as H2B, H3, and H4. A goal for the future is to not only model the amino acid changes as we did here but also to introduce potential novel PTMs. In future experiments, lysine residues such as the H2AE57K that we analyzed, which could be sites for de novo PTMs (acetylation and methylation) could be modeled in PyMOL with the PTM present. Furthermore, results from this study and previous studies analyzing known H3 oncohistones demonstrate the potential information that can be gleaned from creating a PyMOL model for novel and known oncohistones to accurately predict important protein interactions and PTMs that could affect tumorigenesis and/or oncogenesis. Overall, future experiments in yeast models, cell cultures, and xenograft mouse models *in vivo* should be performed with these novel H2A oncohistone candidates to observe their disruption of biological pathways that could cause tumorigenesis and/or oncogenesis.

Overall my studies which combine functional analysis of oncohistones in the laboratory and modeling of potential novel oncohistones provide some insight into cellular pathways that are affected by more well characterized missense mutations in histone H3 and provide the first analysis of potential missense mutations at one position in histone H2A. There are many more potential novel oncohistone mutations uncovered in the Foundation One dataset (nearly 2000), which could be analyzed by structural modeling and subsequent studies in budding yeast. These studies highlight the importance of considering potential oncohistone mutations beyond those that have already been characterized as potential drivers for cancer.

REFERENCES

Alvarez-Venegas, R. (2014). Bacterial SET domain proteins and their role in eukaryotic chromatin modification. *Front Genet, 5*, 65. doi:10.3389/fgene.2014.00065

Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res, 21*(3), 381-395. doi:10.1038/cr.2011.22

Bennett, R. L., Bele, A., Small, E. C., Will, C. M., Nabet, B., Oyer, J. A., . . . Licht, J. D. (2019). A Mutation in Histone H2B Represents a New Class of Oncogenic Driver. *Cancer Discov, 9*(10), 1438-1451. doi:10.1158/2159-8290.CD-19-0393

Bhaumik, S. R., Smith, E., & Shilatifard, A. (2007). Covalent modifications of histones during development and disease pathogenesis. *Nat Struct Mol Biol, 14*(11), 1008-1016. doi:10.1038/nsmb1337

Brown, D. T. (2001). Histone variants: are they functionally heterogeneous? *Genome Biol, 2*(7), Reviews0006. doi:10.1186/gb-2001-2-7-reviews0006

Cancer Genome Atlas, N. (2015). Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature, 517*(7536), 576-582. doi:10.1038/nature14129

Duina, A. A., & Turkal, C. E. (2017). Targeted in Situ Mutagenesis of Histone Genes in Budding Yeast. *J Vis Exp*(119). doi:10.3791/55263

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell, 144*(5), 646-674. doi:10.1016/j.cell.2011.02.013

Hoyos-Manchado, R., Reyes-Martín, F., Rallis, C., Gamero-Estévez, E., Rodríguez-Gómez, P., Quintero-Blanco, J., . . . Tallada, V. A. (2017). RNA metabolism is the primary target of formamide in vivo. *Sci Rep, 7*(1), 15895. doi:10.1038/s41598-017-16291-8

Hyland, E. M., Cosgrove, M. S., Molina, H., Wang, D., Pandey, A., Cottee, R. J., & Boeke, J. D. (2005). Insights into the role of histone H3 and histone H4 core modifiable residues in Saccharomyces cerevisiae. *Mol Cell Biol, 25*(22), 10060-10070. doi:10.1128/MCB.25.22.10060-10070.2005

Johnson, P., Mitchell, V., McClure, K., Kellems, M., Marshall, S., Allison, M. K., . . . Duina, A. A. (2015). A systematic mutational analysis of a histone H3 residue in budding yeast provides insights into chromatin dynamics. *G3 (Bethesda), 5*(5), 741-749. doi:10.1534/g3.115.017376

Kaushal, A., & Vashishat, R. K. (1982). Effect of hydroxyurea on UV-induced mitotic crossingover & gene conversion in Saccharomyces cerevisiae. *Indian J Exp Biol, 20*(3), 195-199.

Klein, M., Swinnen, S., Thevelein, J. M., & Nevoigt, E. (2017). Glycerol metabolism and transport in yeast and fungi: established knowledge and ambiguities. *Environ Microbiol, 19*(3), 878-893. doi:10.1111/1462-2920.13617

Koy, J. F., Pleninger, P., Wall, L., Pramanik, A., Martinez, M., & Moore, C. W. (1995). Genetic changes and bioassays in bleomycin- and phleomycin-treated cells, and their relationship to chromosomal breaks. *Mutat Res, 336*(1), 19-27. doi:10.1016/0921-8777(94)00040-d

Kuranda, K., Leberre, V., Sokol, S., Palamarczyk, G., & François, J. (2006). Investigating the caffeine effects in the yeast Saccharomyces cerevisiae brings new insights into the connection between TOR, PKC and Ras/cAMP signalling pathways. *Mol Microbiol, 61*(5), 1147-1166. doi:10.1111/j.1365-2958.2006.05300.x

Leslie, E. J. (2020). Embracing human genetics: a primer for developmental biologists. *Development, 147*(21). doi:10.1242/dev.191114

Li, F., Yang, J., Wu, X., Zhang, W., Wang, G., & Yue, X. (2014). [Effects of histone H3 K4L and K36L mutations on the cell growth and the transcription of GAL1, SSA3 and PHO5 in Saccharomyces cerevisiae]. *Yi Chuan, 36*(8), 827-834. doi:10.3724/sp.J.1005.2014.0827

Lowe, B. R., Yadav, R. K., Henry, R. A., Schreiner, P., Matsuda, A., Fernandez, A. G., . . . Partridge, J. F. (2021). Surprising phenotypic diversity of cancer-associated mutations of Gly 34 in the histone H3 tail. *Elife, 10*. doi:10.7554/eLife.65369

Lu, C., Ramirez, D., Hwang, S., Jungbluth, A., Frosina, D., Ntiamoah, P., . . . Hameed, M. (2019). Histone H3K36M mutation and trimethylation patterns in chondroblastoma. *Histopathology, 74*(2), 291-299. doi:10.1111/his.13725

Maze, I., Noh, K. M., Soshnev, A. A., & Allis, C. D. (2014). Every amino acid matters: essential contributions of histone variants to mammalian development and disease. *Nat Rev Genet, 15*(4), 259-271. doi:10.1038/nrg3673

McBurney, K. L., Leung, A., Choi, J. K., Martin, B. J., Irwin, N. A., Bartke, T., . . . Howe, L. J. (2016). Divergent Residues Within Histone H3 Dictate a Unique Chromatin Structure in Saccharomyces cerevisiae. *Genetics, 202*(1), 341-349. doi:10.1534/genetics.115.180810

McGinty, R. K., & Tan, S. (2015). Nucleosome structure and function. *Chem Rev, 115*(6), 2255-2273. doi:10.1021/cr500373h

Mohammad, F., & Helin, K. (2017). Oncohistones: drivers of pediatric cancers. *Genes Dev, 31*(23-24), 2313-2324. doi:10.1101/gad.309013.117

Nacev, B. A., Feng, L., Bagert, J. D., Lemiesz, A. E., Gao, J., Soshnev, A. A., . . . Allis, C. D. (2019). The expanding landscape of 'oncohistone' mutations in human cancers. *Nature, 567*(7749), 473-478. doi:10.1038/s41586-019-1038-1

Qiu, L., & Han, J. (2021). Histone H3G34 Mutation in Brain and Bone Tumors. *Adv Exp Med Biol, 1283*, 63-71. doi:10.1007/978-981-15-8104-5_5

Qiu, L., Hu, X., Jing, Q., Zeng, X., Chan, K. M., & Han, J. (2018). Mechanism of cancer: Oncohistones in action. *J Genet Genomics, 45*(5), 227-236. doi:10.1016/j.jgg.2018.04.004

Sarthy, J. F., & Henikoff, S. (2019). Bringing Oncohistones into the Fold. *Cancer Discov, 9*(10), 1346-1348. doi:10.1158/2159-8290.CD-19-0839

Shi, L., Wen, H., & Shi, X. (2017). The Histone Variant H3.3 in Transcriptional Regulation and Human Disease. *J Mol Biol, 429*(13), 1934-1945. doi:10.1016/j.jmb.2016.11.019

Sidney, S., Go, A. S., & Rana, J. S. (2019). Transition From Heart Disease to Cancer as the Leading Cause of Death in the United States. *Ann Intern Med, 171*(3), 225. doi:10.7326/l19-0202

Szenker, E., Ray-Gallet, D., & Almouzni, G. (2011). The double face of the histone variant H3.3. *Cell Res, 21*(3), 421-434. doi:10.1038/cr.2011.14

Taguchi, H., Xie, Y., Horikoshi, N., Maehara, K., Harada, A., Nogami, J., . . . Kurumizaka, H. (2017). Crystal Structure and Characterization of Novel Human Histone H3 Variants, H3.6, H3.7, and H3.8. *Biochemistry, 56*(16), 2184-2196. doi:10.1021/acs.biochem.6b01098

White, C. L., Suto, R. K., & Luger, K. (2001). Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *Embo j, 20*(18), 5207-5218. doi:10.1093/emboj/20.18.5207

Yadav, R. K., Jablonowski, C. M., Fernandez, A. G., Lowe, B. R., Henry, R. A., Finkelstein, D., . . . Partridge, J. F. (2017). Histone H3G34R mutation causes replication stress, homologous recombination defects and genomic instability in S. pombe. *Elife, 6*. doi:10.7554/eLife.27406

Yusufova, N., Kloetgen, A., Teater, M., Osunsade, A., Camarillo, J. M., Chin, C. R., . . . Melnick, A. M. (2021). Histone H1 loss drives lymphoma by disrupting 3D chromatin architecture. *Nature, 589*(7841), 299-305. doi:10.1038/s41586-020-3017-y

Zhang, Y., Shan, C. M., Wang, J., Bao, K., Tong, L., & Jia, S. (2017). Molecular basis for the role of oncogenic histone mutations in modulating H3K36 methylation. *Sci Rep, 7*, 43906. doi:10.1038/srep43906
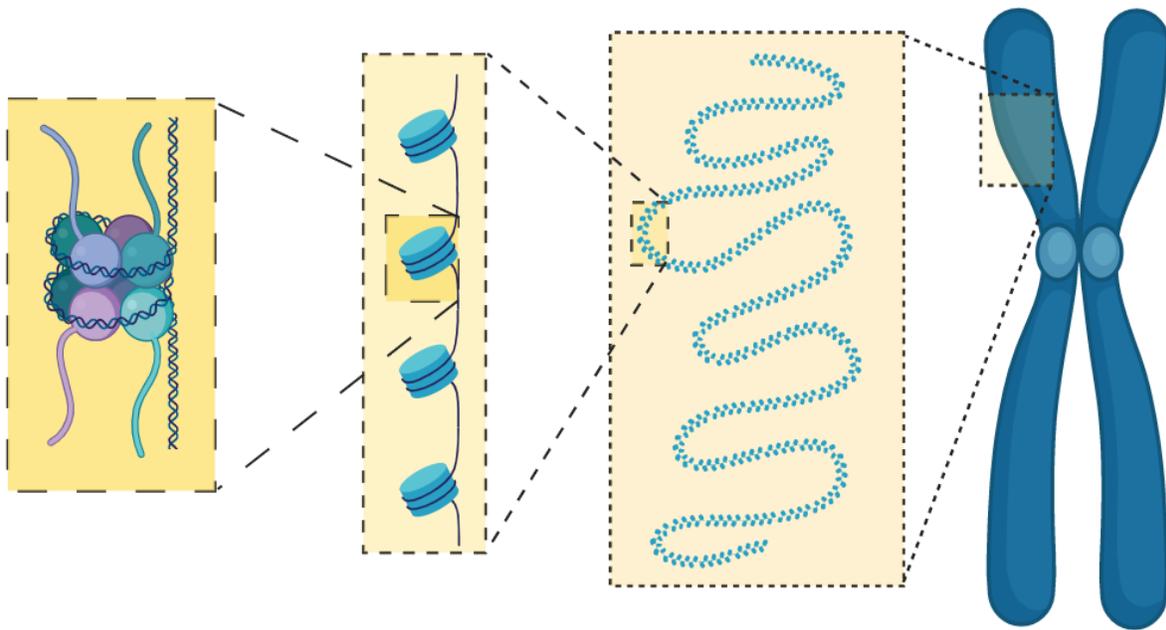
Figures

Figure 1. Molecular Structure of Nucleosomes on a Micro to Macro Level. A) Condensed version of a nucleosome containing the four core histone proteins (H2A, H2B, H3, and H4) in their octamer protein complex. Double-stranded DNA is wrapped around the octamer protein complex. Histone N-terminal tails are visible extending below and above the octamer protein complex. B) Decondensed version showing multiple nucleosomes. Histones are represented by blue cylinders, and DNA wraps these histones as depicted by the black string. C) Decondensed version of chromatin, which represents the basic DNA subunit of a chromosome. This illustration was created by Severin Lustenberger.
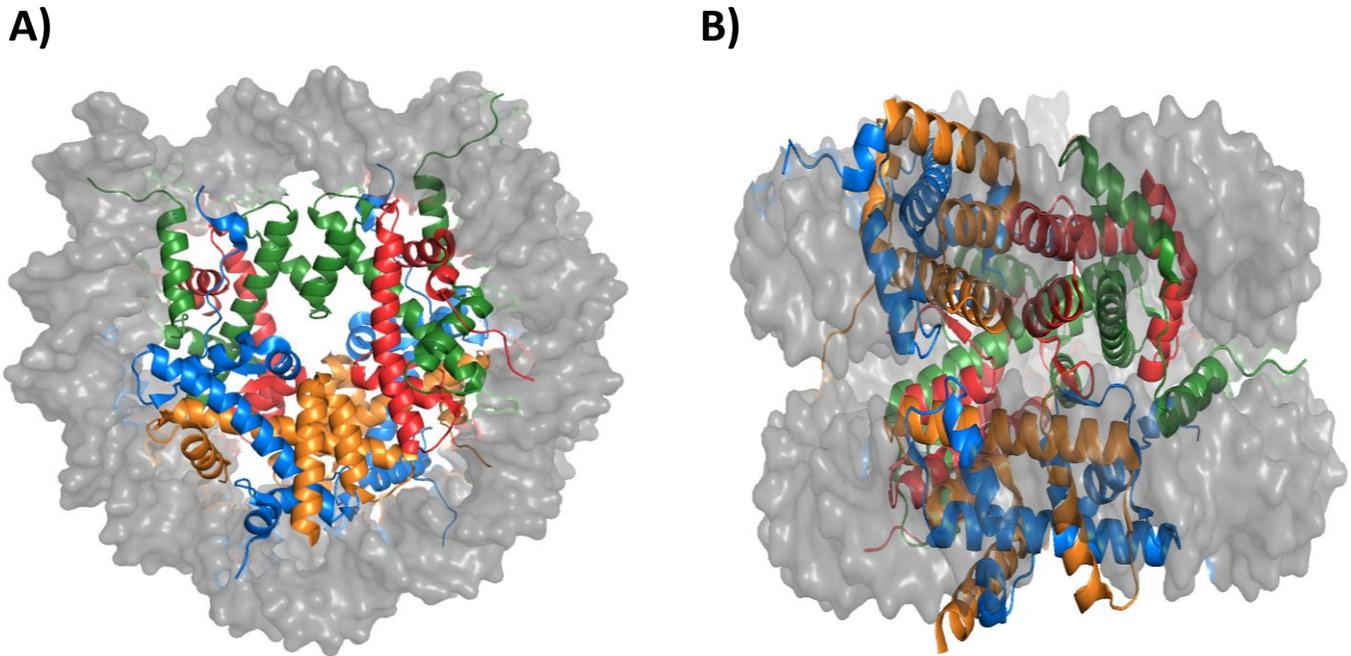
A)

B)

Figure 2. Nucleosomes consist of histone octamers wrapped by DNA. A) A front-view of a single nucleosome containing the four core histone proteins (H2A, H2B, H3, and H4) and DNA (gray) in the octamer protein complex. B) A side-view of a single nucleosome containing the four core histone proteins (H2A, H2B, H3, and H4) and DNA (gray). Histone H2A in blue, Histone H2B in orange, Histone H3 in green, and Histone H4 in red. This illustration was adapted from Severin Lustenberger. The structure was generated using PyMOL Version 2.0 based on PDB: 5X7X (Taguchi et al., 2017).
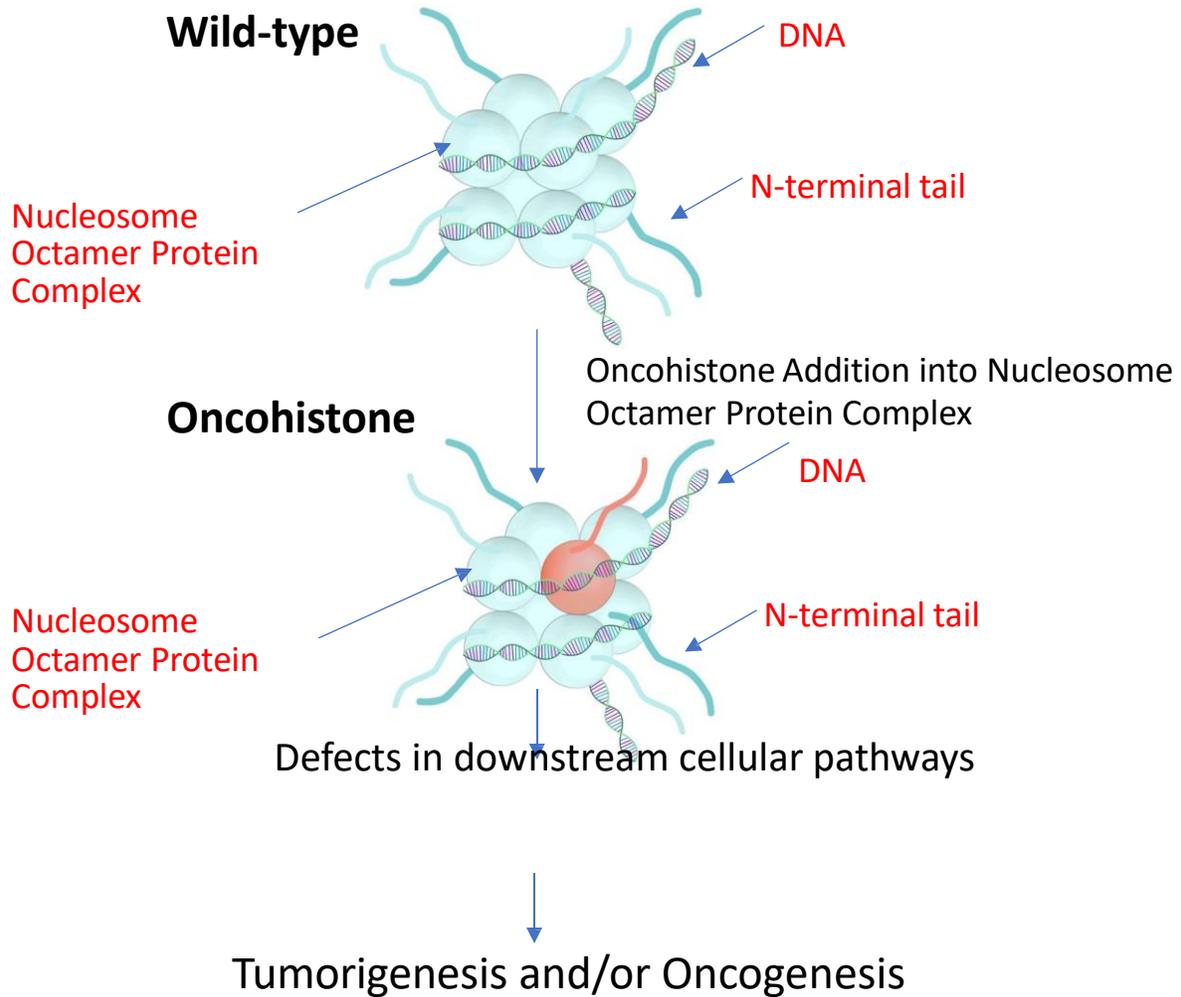
Figure 3. General Structure of an Oncohistone-containing Nucleosome. The defective oncohistone (red) is added to the nucleosome octamer protein complex. How the oncohistone alters the function of the nucleosome is largely unknown, but the hypothesis is that the oncohistone alters key cellular pathways that drive oncogenesis. This figure was generated via Canva and adapted from (Nacev et al., 2019).
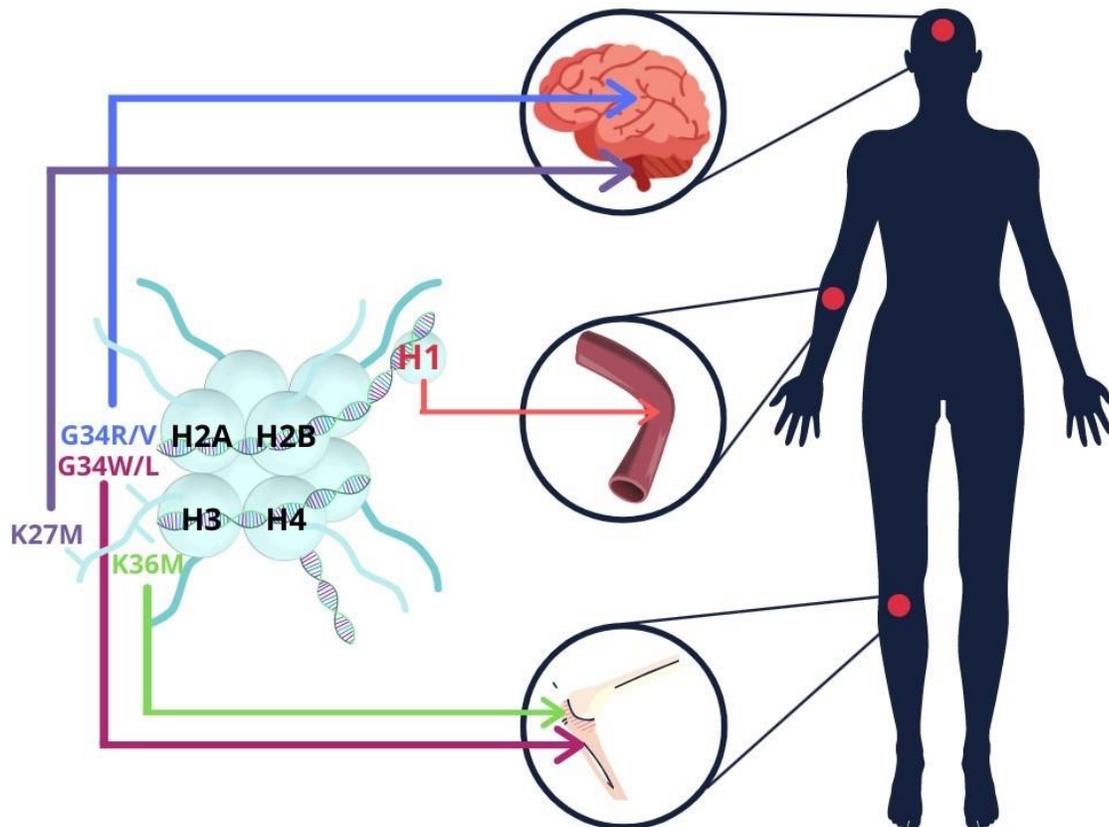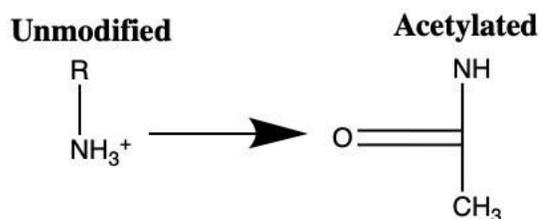
Figure 4. Oncohistones Drive Tumor Formation throughout the Body. Oncohistone H3K27M as well as G34 (H3G34R and H3G34V) mutations cause glioblastoma, a tumor of the brain. Histone H3K36M as well as G34 (H3G34W and H3G34L) mutations cause chondrosarcoma and giant cell tumors, cancers of the bone. This figure was created using Canva and adapted from (Qiu et al., 2018).

A) Acetylation

Unmodified → **Acetylated**

$$R-NH_3^+ \longrightarrow O=\overset{\overset{\displaystyle NH}{|}}{\underset{\underset{\displaystyle CH_3}{|}}{C}}$$

B) Methylation

Unmodified → **Monomethylated** → **Dimethylated** → **Trimethylated**

$$R-NH_3^+ \longrightarrow \underset{CH_3}{\overset{R}{NH_2^+}} \longrightarrow \underset{CH_3}{\overset{R}{NH^+}}-CH_3 \longrightarrow \underset{CH_3}{\overset{CH_3}{NH^+}}-CH_3$$
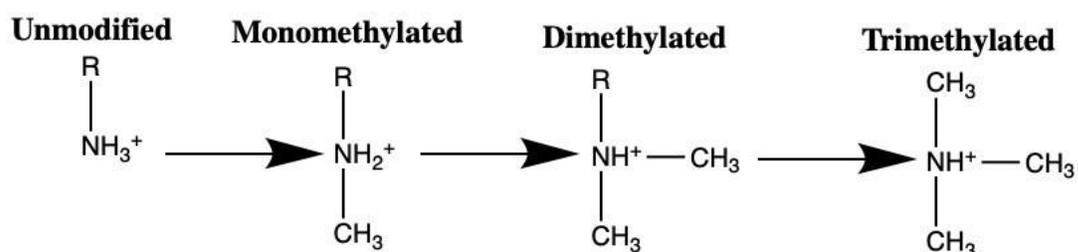
Figure 5. Acetylation and Methylation Post-translational modifications (PTMs) of the N-terminal tail or core domain regions of nucleosomes. A) Unmodified R group acetylated via the addition of the acetyl group (CH_3CO). B) Unmodified R methylated consecutively via the addition of the methyl group (CH_3). This illustration was adapted from Severin Lustenberger. The structure was generated using ChemDraw Version 20.0.0.38 (PerkinElmer Informatics).
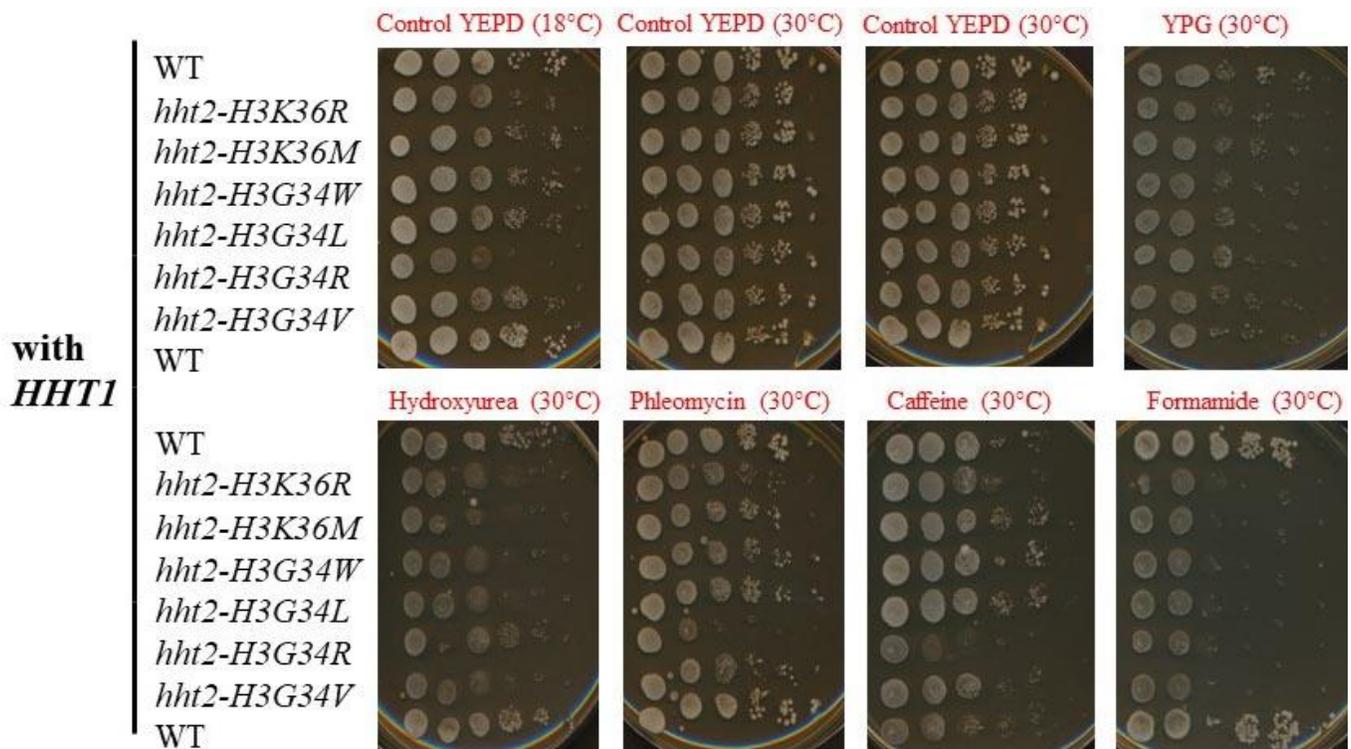
Figure 6. Histone H3 mutants show dominant growth defects when plated on YP Glycerol, hydroxyurea, phleomycin, caffeine, and formamide. Histone H3 mutants at residues K36 (H3K36R and H3K36M) and G34 (H3G34W, H3G34L, H3G34R, and H3G34V) grow poorly on media containing YP Glycerol (YPG), hydroxyurea, phleomycin, caffeine, and formamide, which affect stress DNA replication pathways, cell proliferation, stress response, and RNA metabolism, respectively. Results are more striking on plates containing formamide and hydroxyurea. The *hht2*-H3G34R also shows striking growth defects. The control wildtype (WT) cells grow well under all conditions employed. These changes in cell growth occur even in the presence of a wildtype copy of *HHT1*, revealing growth defects that are dominant to a wildtype histone H3 gene (*HHT1*). These results are representative of three independent experiments.

S. cerevisiae

NW Score	Identities	Positives	Gaps
464	97/132(73%)	111/132(84%)	1/132(0%)

```
Query    1  MSG-RGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKGNYAERVGAGAPVYLAAVLEYL   59
            MSG +G + G A    +++RS++AGL FPVGRVHRLLR+GNYA+R+G+GAPVYL AVLEYL
Sbjct    1  MSGGKGGKAGSAAKASQSRSAKAGLTFPVGRVHRLLRRGNYAQRIGSGAPVYLTAVLEYL   60

Query   60  TAEILELAGNAARDNKKTRIIPRHLQLAIRNDEELNKLLGKVTIAQGGVLPNIQAVLLPK   119
             AEILELAGNAARDNKKTRIIPRHLQLAIRND+ELNKLLG VTIAQGGVLPNI   LLPK
Sbjct   61  AAEILELAGNAARDNKKTRIIPRHLQLAIRNDDELNKLLGNVTIAQGGVLPNIHQNLLPK   120

Query  120  KTESHHKAKGK    130
            K+    KA  +
Sbjct  121  KSAKATKASQEL   132
```

M. musculus

NW Score	Identities	Positives	Gaps
618	126/130(97%)	128/130(98%)	2/130(1%)

```
Query    1  MSGRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKGNYAERVGAGAPVYLAAVLEYLT   60
            MSGRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKGNY+ERVGAGAPVYLAAVLEYLT
Sbjct    1  MSGRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKGNYSERVGAGAPVYLAAVLEYLT   60

Query   61  AEILELAGNAARDNKKTRIIPRHLQLAIRNDEELNKLLGKVTIAQGGVLPNIQAVLLPKK   120
            AEILELAGNAARDNKKTRIIPRHLQLAIRNDEELNKLLG+VTIAQGGVLPNIQAVLLPKK
Sbjct   61  AEILELAGNAARDNKKTRIIPRHLQLAIRNDEELNKLLGRVTIAQGGVLPNIQAVLLPKK   120

Query  121  TESHHKAKGK    130
            TESHHKAK
Sbjct  121  TESHHKAK      128
```

NW Score	Identities	Positives	Gaps
319	80/142(56%)	98/142(69%)	4/142(2%)

D. melanogaster

```
Query    1  MSG--RGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKGNYAE-RVGAGAPVYLAAVLE   57
            M+G   GK  GKA+AKA +RS+RAGLQFPVGR+HR L+     +  RVGA A VY AA+LE
Sbjct    1  MAGGKAGKDSGKAKAKAVSRSARAGLQFPVGRIHRHLKSRTTSHGRVGATAAVYSAAILE   60

Query   58  YLTAEILELAGNAARDNKKTRIIPRHLQLAIRNDEELNKLLGKVTIAQGGVLPNIQAVLL   117
            YLTAE+LELAGNA++D K  RI PRHLQLAIR DEEL+ L+ K TIA GGV+P+I   L+
Sbjct   61  YLTAEVLELAGNASKDLKVKRITPRHLQLAIRGDEELDSLI-KATIAGGGVIPHIHKSLI   119

Query  118  PKKTESHHKAKGK           130
            KK E+    + K
Sbjct  120  GKKEETVQDPQRKGNVILSQAY   141
```

Figure 7. Human Histone H2A Type 1 is evolutionarily conserved. Human Histone H2A type 1 protein shares 73% amino acid identity with *S. cerevisiae* Histone H2A.1, 97% amino acid identity with *M. musculus* Histone H2A Type 1-H, and 56% amino acid identity with *D. melanogaster* Histone H2A.v. The alignments are shown with the human H2A type 1 protein sequence on top (Query) and the other sequence on the bottom (Sbjct). All FASTA sequences were obtained from (*UniProt*). Data was inputted into NCBI BLAST and results were recorded (*BLAST: Basic Local Alignment Search Tool*).
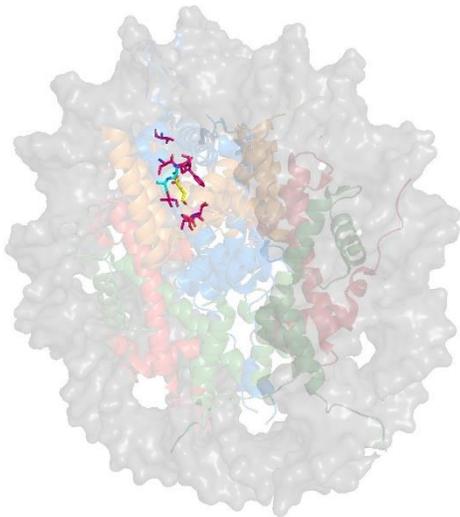
Key:
DNA: grey
Histone H2A Histone H2B
Histone H3 Histone H4
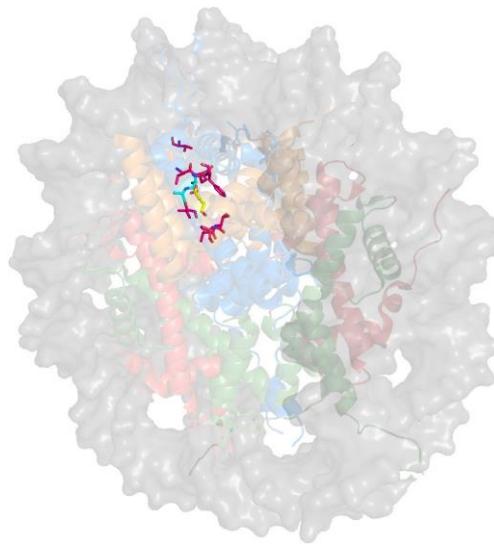WT residue of interest
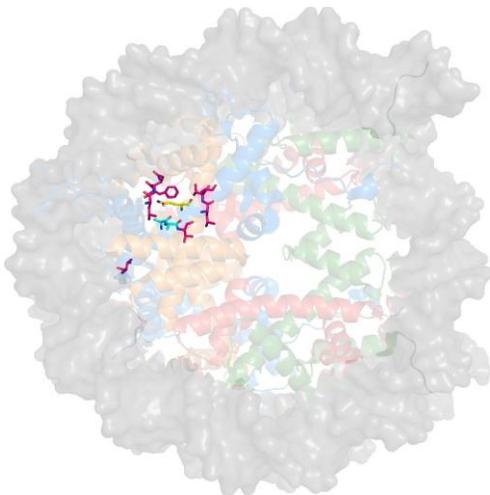Mutated residue of interest
Proximal Alanine
All other nearby Amino Acids

Wild-type H2AE57

H2AE57K

H2AE57Q

Figure 8. Nucleosome structure illustrating WT H2AE57, H2AE57K Mutant, and H2AE57Q Mutant. DNA and Histone domains are transparent to clearly see amino acid residues. The structure was generated using PyMOL Version 2.0 based on PDB: 5X7X (Taguchi et al., 2017).

Key:

DNA

Histone H2A

Histone H2B

Histone H3

Histone H4

WT or Mutant residue of interest
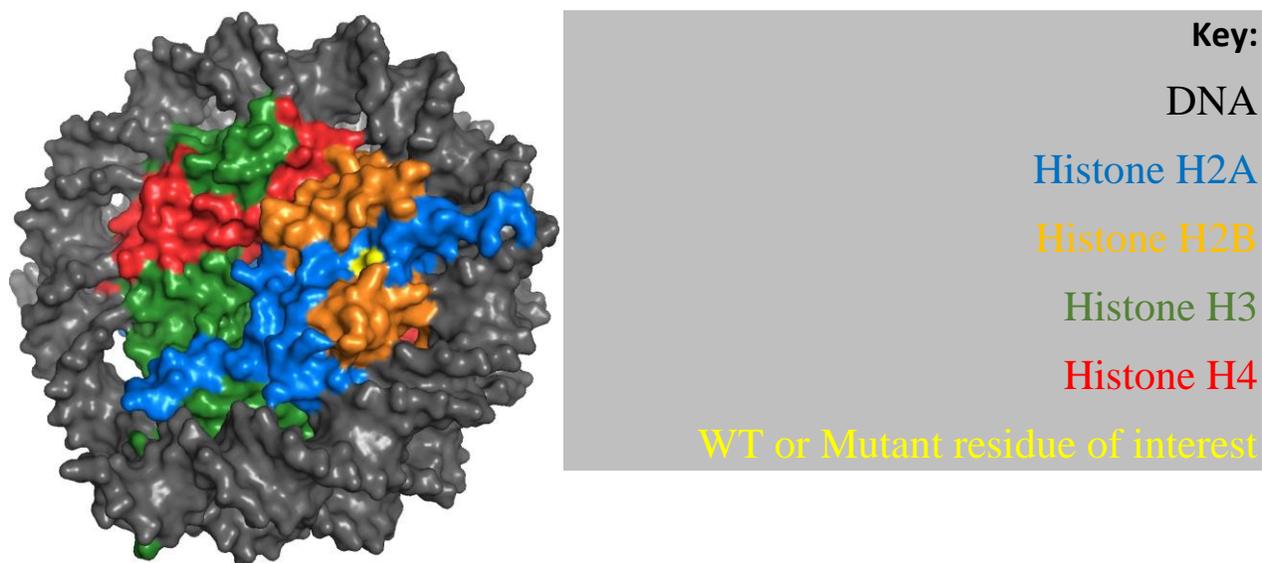
Wild-type H2AE57
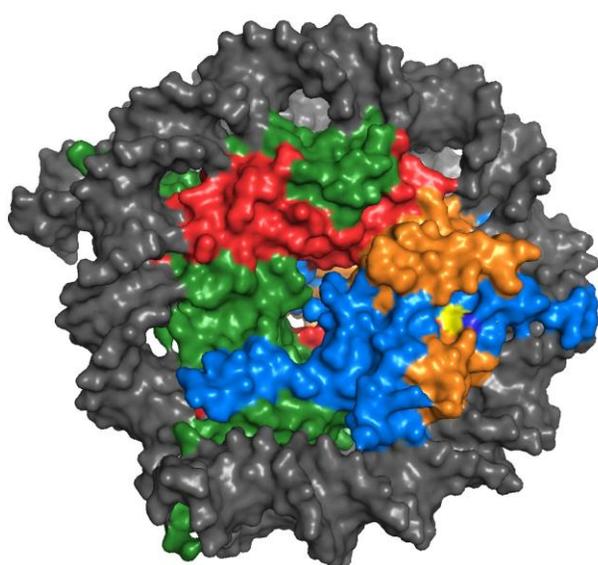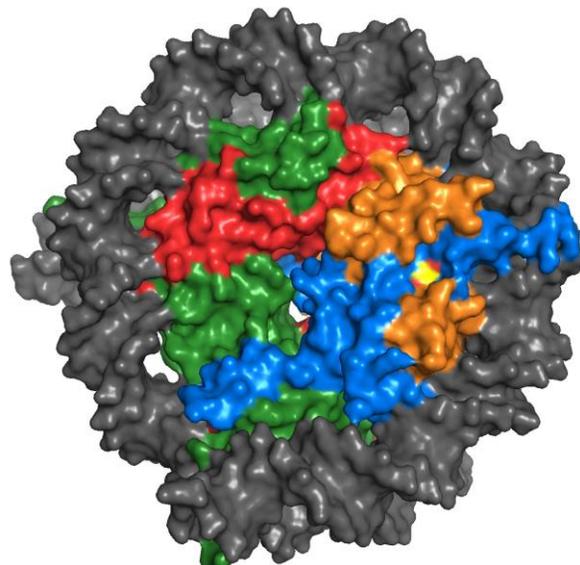
H2AE57K

H2AE57Q

Figure 9. WT H2AE57, H2AE57K Mutant, and H2AE57Q are located on the surface of the Octamer Protein Complex. Each of the histones is shown in a distinct color and the position of H2AE, E57K or E57Q is shown in yellow. The structure, which shows the surface representation, was generated using PyMOL Version 2.0 based on PDB: 5X7X (Taguchi et al., 2017).
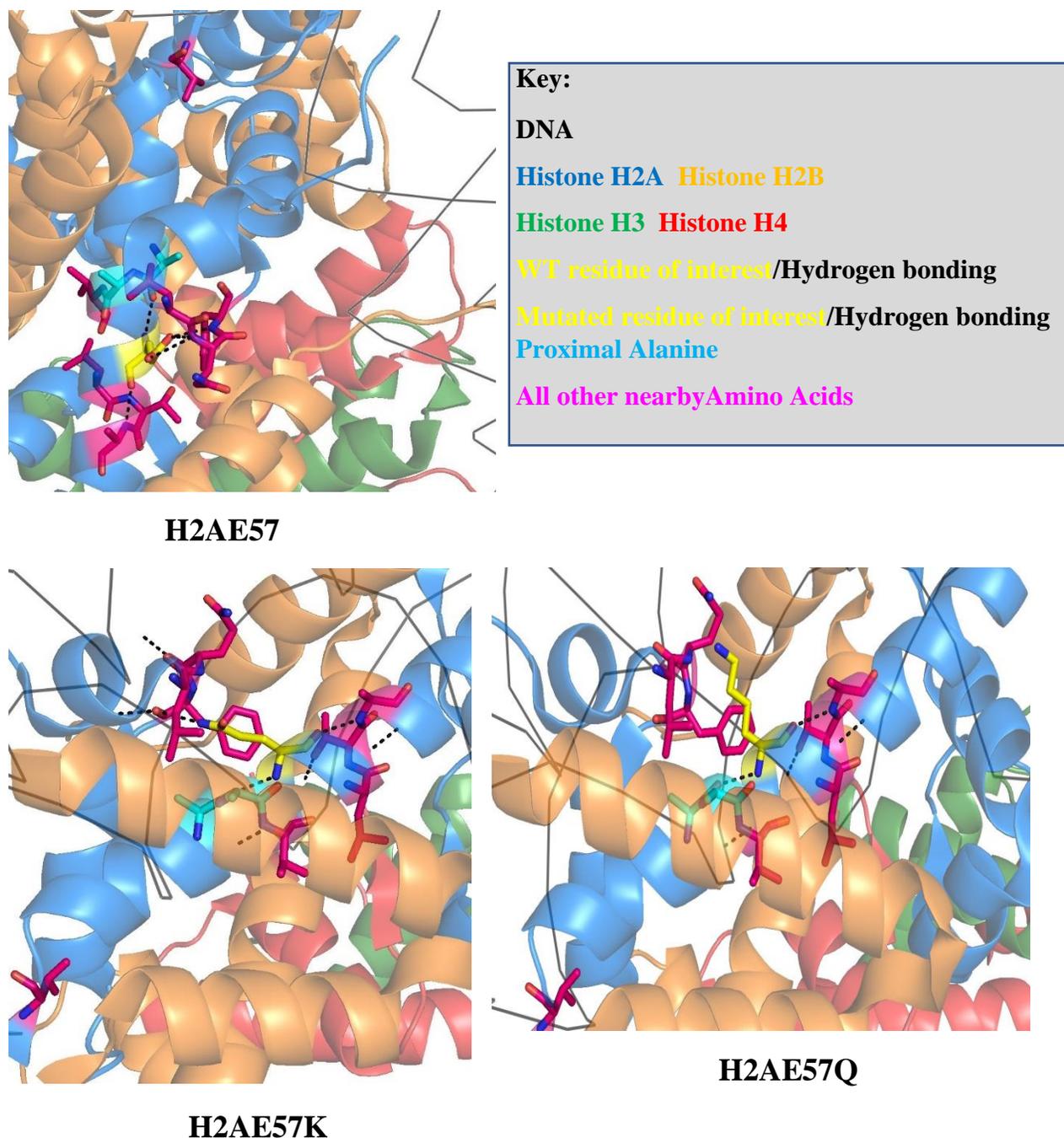
H2AE57

Key:

DNA

Histone H2A **Histone H2B**

Histone H3 **Histone H4**

WT residue of interest/**Hydrogen bonding**

Mutated residue of interest/**Hydrogen bonding**
Proximal Alanine

All other nearbyAmino Acids

H2AE57K

H2AE57Q

Figure 10. Characteristics and Chemical Interactions of WT H2AE57, H2AE57K, and H2A57Q. WT H2AE57K, H2AE57K, and H2AE57Q are positioned near the interface with H2B (green). There are various steric clashes with surrounding residues. Hydrophobic, polar, and hydrogen bonding forces are seen between the residue of interest and the surrounding amino acids. WT H2AE57 has minimal steric clashes with surrounding residues, allowing for normal nucleosome function. A Valine on the N-terminal tail of histone H2A associates closely with the surrounding DNA. Core hydrogen bonds of residues of interest are depicted with black dashed lines. Structure from PyMOL Version 2.0 based on PDB: 5X7X (Taguchi et al., 2017).
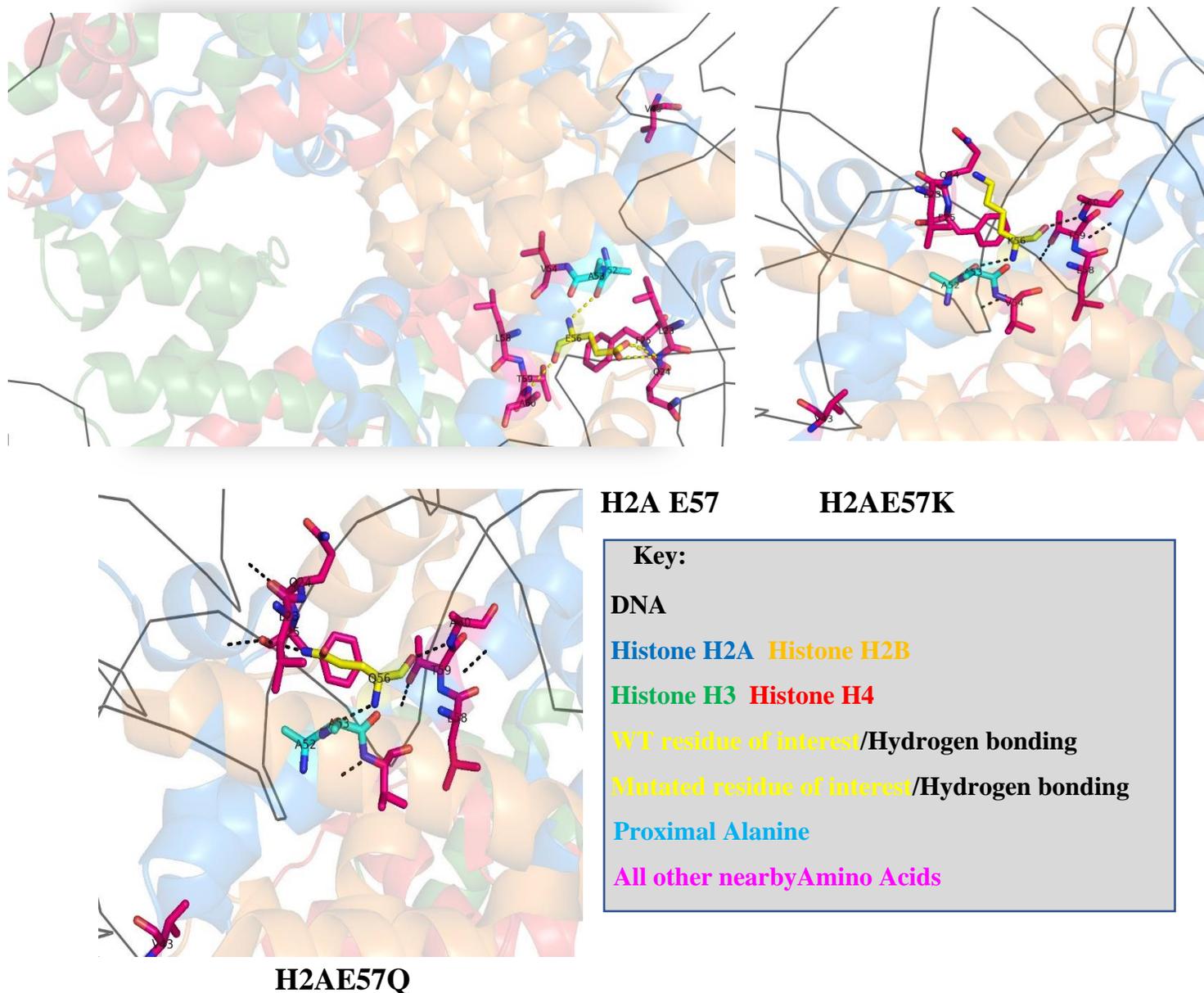
H2A E57   **H2AE57K**

Key:

DNA

Histone H2A Histone H2B

Histone H3 Histone H4

WT residue of interest/Hydrogen bonding

Mutated residue of interest/Hydrogen bonding

Proximal Alanine

All other nearbyAmino Acids

H2AE57Q

Figure 11. Histone H2AE57 is Critical for Proper Hydrogen Bonding. Changing H2AE57 to either K (H2AE57K) or Q (H2AE57Q) causes a loss of hydrogen boding with surrounding residues. We hypothesize that this lack of hydrogen bonding could facilitate introduction of de novo PTMs, methylation and acetylation for H2AE57K. This could be due to weaker bonding interactions with the surrounding residues and increased steric hinderance between the H2AE57K and the aromatic ring of Phenylalanine. Core hydrogen bonds of the residue of interest are depicted with either black or yellow dashed lines. The structure was generated using PyMOL Version 2.0 based on PDB: 5X7X (Taguchi et al., 2017).