

## **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.

Alicia Yin

April 10, 2023

Elucidating the role of CHD5 domains in the recruitment of CHD5 to DNA double-strand break sites

by

Alicia Yin

David S. Yu, MD, PhD  
Adviser

Biology

David S. Yu, MD, PhD  
Adviser

Heather Skye Comstra, PhD  
Committee Member

Edward Nam, PhD  
Committee Member

Alexandre Orthwein, PhD, MSc  
Committee Member

2023

Elucidating the role of CHD5 domains in the recruitment of CHD5 to DNA double-strand break sites

By

Alicia Yin

David S. Yu, MD, PhD

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

Biology

2023

## Abstract

Elucidating the role of CHD5 domains in the recruitment of CHD5 to DNA double-strand break sites

By Alicia Yin

DNA double-strand breaks (DSBs) are highly deleterious and pose a threat to genome integrity when left unrepaired, potentially driving the development or progression of cancer. Non-homologous end joining (NHEJ) and homologous recombination (HR) are the two primary repair pathways that eukaryotic cells rely on to repair DSBs. However, NHEJ is more mutagenic, so maintaining the balance between the two pathways within a cell is also essential for genomic stability. The chromatin remodeler protein CHD5 has been found to play a role in the DNA damage response (DDR), which may include mediating the balance between NHEJ and HR, as unpublished results from our lab have shown that loss of CHD5 shifts the DNA repair choice toward NHEJ. Therefore, we aim to further understand the mechanics behind CHD5's recruitment to DSBs. In this study, we investigate which domains of CHD5 are responsible for its recruitment. Using chromatin immunoprecipitation (ChIP) and real-time quantitative PCR (qPCR), we show that all domains except the domain of unknown function (DUF) and the CHDCT2 uncharacterized domain recruit independently to DSBs, suggesting that all of the domains except the DUF and CHDCT2 participate in facilitating recruitment. However, the lysine-rich domain at CHD5's N-terminal showed the greatest enrichment at DSBs, indicating that it may play the biggest role in facilitating recruitment. Finally, we use laser microirradiation and live-cell imaging to explore the recruitment and retention kinetics of the CHD5 protein, demonstrating that CHD5 recruits rapidly and retains for an extended period of time.

Elucidating the role of CHD5 domains in the recruitment of CHD5 to DNA double-strand break sites

By

Alicia Yin

David S. Yu, MD, PhD

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

Biology

2023

## Acknowledgements

I would like to express my sincere gratitude to Dr. David Yu for giving me the opportunity to be a part of his lab and pursue this research project. His tireless support and generosity have not only been instrumental in fostering my growth as a student, but have granted me an incredibly positive undergraduate research experience overall, and for that I am so thankful.

I am also deeply indebted to my mentor, Dr. Sandip Rath, for his unwavering kindness and patience. Thank you for your guidance throughout this entire project, for always taking the time to answer my endless questions, and for consistently going above and beyond your duties as a mentor.

Additionally, I wish to thank the members of my Honors Committee: Dr. Skye Comstra, Dr. Edward Nam, and Dr. Alexandre Orthwein. Thank you for taking the time out of your busy schedules to provide me with invaluable insight and feedback on my work.

To everyone in the Yu Lab, thank you for creating such an enjoyable and welcoming environment to work in, for being wonderful resources, and for helping me prepare for my defense.

Special thanks to my friends for their incredible love and support. Thank you for being there for me every step of the way and for always believing in me. I would not have been able to accomplish this without you all.

Finally, I would like to dedicate this thesis to my parents. Thank you for everything.

## Table of Contents

<b>Introduction</b>	<b>1</b>
DNA double-strand break repair	1
CHD5: A chromatin remodeler involved in cancer	1
Approaches for investigating protein recruitment to DSBs	6
Hypothesis	10
<b>Materials and Methods</b>	<b>11</b>
Cell Lines	11
Plasmid Extraction and Purification	11
Transfections	11
Laser Microirradiation	12
ImageJ Quantification	12
ChIP Assays	13
Real-Time Quantitative PCR	14
Protein Lysate Preparation and Western Blotting	15
Statistical Analysis	16
<b>Results</b>	<b>18</b>
Confirmation of plasmid expression	18
Quantifying recruitment kinetics of CHD5	20
ChIP reveals GFP-CHD5 1-337 construct has the highest enrichment at double-strand breaks	22
<b>Discussion</b>	<b>26</b>
<b>Supplementary Figures</b>	<b>29</b>
<b>References</b>	<b>31</b>

## Figures

<b>Figure 1.</b> Graphical illustration of the 4 main subfamilies of ATP-dependent chromatin remodelers	2
<b>Figure 2.</b> Graphical illustration of CHD subfamilies	3
<b>Figure 3.</b> DNA repair reporter assays	5
<b>Figure 4.</b> Schematic of CHD5 domains	7
<b>Figure 5.</b> Schematic of laser-induced DNA damage and protein recruitment and chromatin immunoprecipitation	8
<b>Figure 6.</b> Schematic showing Fok1 induction theme in U2OS cells	9
<b>Figure 7.</b> Real-time qPCR schematic	15
<b>Figure 8.</b> Western blot of GFP-CHD5 and its GFP-tagged truncated constructs	19
<b>Figure 9.</b> Full-length GFP-CHD5 recruits to laser-induced DNA DSB sites	21
<b>Figure 10.</b> GFP-CHD5 recruits to DNA DSBs within 1 minute following laser-induced damage and retains for 15 minutes	22
<b>Figure 11.</b> GFP-CHD5 1-337 shows greatest enrichment at DSB sites	24
<b>Figure S1.</b> Recruitment of GFP-CHD5 and its associated truncated constructs to laser-induced DNA DSB sites	29
<b>Figure S2.</b> U2OS cells transfected with GFP-CHD5 truncated constructs before and after laser microirradiation	30

## **Introduction**

### ***DNA double-strand break repair***

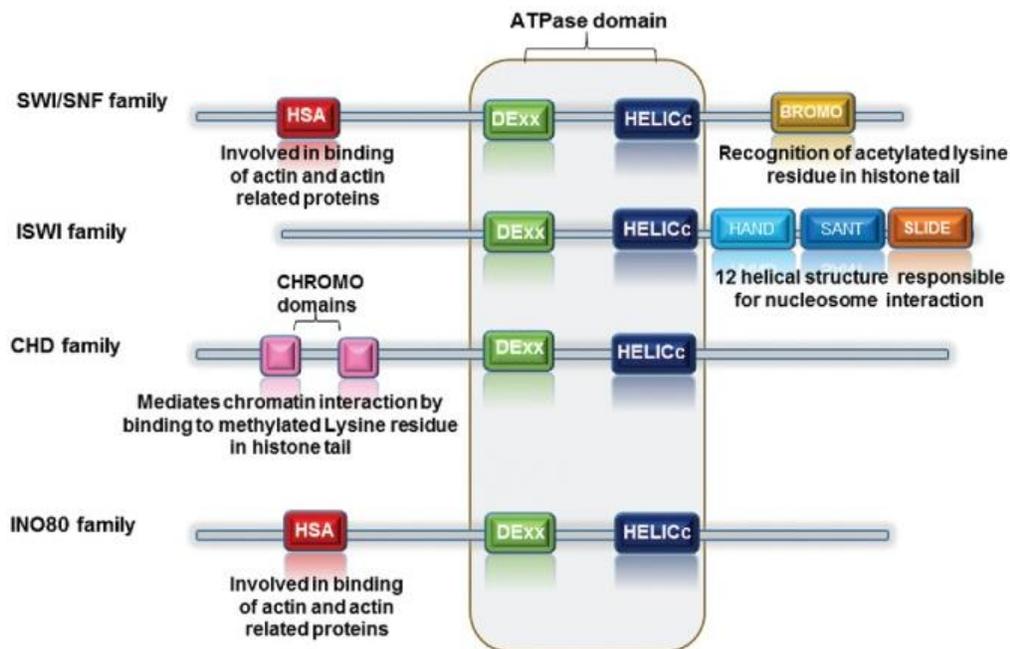
DNA double-strand breaks (DSBs) foster genomic instability that can lead to the development and or progression of cancer (Jeggo et al., 2016). To ensure genome maintenance, cells have evolved an intricate mechanism for DNA repair, known as the DNA damage response (DDR) (Jeggo et al., 2016) (Groelly et al., 2023). Deficiencies in the DDR promote genomic instability and can give rise to the proliferation of damaged cells, thereby increasing mutational load that eventually contributes to increased risk of cancer development (Jeggo et al., 2016). Consequently, ongoing research seeks to identify opportunities to target these repair pathways as approaches for potential cancer therapeutics (Groelly et al., 2023; Pilié et al., 2019).

There are two primary pathways by which cells can repair double-strand breaks: non-homologous end joining (NHEJ) and homologous recombination (HR) (Mao et al., 2008). NHEJ is the faster and more efficient pathway, but it is error-prone and mutagenic, as it is cell-cycle independent and does not rely on a homologous template (Mao et al., 2008). Instead, NHEJ directly ligates the two broken DNA ends together without accounting for nucleotides that were either lost or gained, therefore resulting in insertions or deletions at the repaired break site (Clouaire and Legube, 2015; Mao et al., 2008). In contrast, HR occurs only during the S and G2 phases of the cell cycle, when DNA replication has occurred, as it relies on the sister chromatid as a template for accurate repair (Clouaire and Legube, 2015; Mao et al., 2008).

### ***CHD5: A chromatin remodeler involved in cancer***

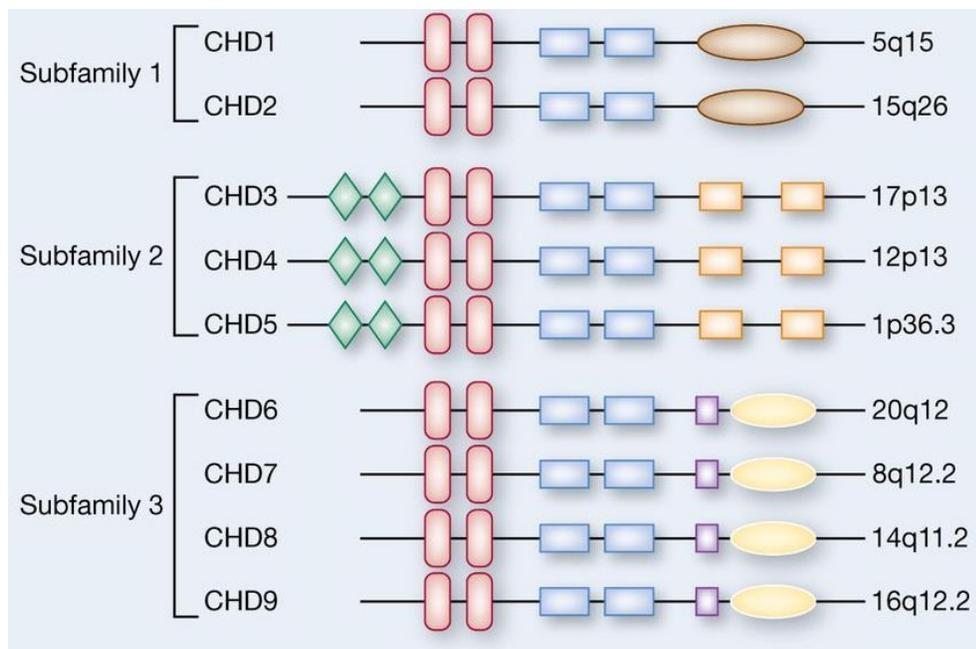
The structure and organization of chromatin in the regions surrounding double-strand breaks impacts the ability of DSB-repair machinery to detect, access, and repair DNA DSBs

(Clouaire and Legube, 2015). Chromatin remodeling factors alter the accessibility of DNA, hence playing an important role in orchestrating the DDR machinery (Clouaire and Legube, 2015). Chromatin remodeling occurs mostly via two ways: post-translational histone modifications, or by relying on energy from ATP hydrolysis to mediate contact between histones and DNA within the nucleosome (Price and D'Andrea, 2013; Tyagi et al., 2016). There are four main subfamilies of ATP-dependent chromatin remodelers: SWI/SNF, ISWI, INO80, and CHD (Price and D'Andrea, 2013; Wolffe, 2001).



**Figure 1. Graphical illustration of the 4 main subfamilies of ATP-dependent chromatin remodelers.** Representative image adapted from Tyagi et al. shows the characteristic domains of the 4 main subfamilies of ATP-dependent chromatin remodelers. All 4 families share an ATPase domain. The CHD family is characterized by tandem chromodomains found in the N-terminal region (Tyagi et al., 2016). Original image is Figure 1 from *Chromatin remodelers: We are the drivers!!* by Tyagi et al. published in *Nucleus* and © 2016 Taylor & Francis, reprinted by permission of Informa UK Limited, trading as Taylor & Taylor & Francis Group, <http://www.tandfonline.com>.

The chromodomain helicase DNA-binding (CHD) family of proteins is highly conserved in eukaryotes and has two signature sequence motifs: tandem chromodomains in the N-terminal region of the protein, and an SNF2-like ATPase domain in the central region (Marfella and Imbalzano, 2007). Within this family, there are three subfamilies that consist of nine proteins total—CHD1 and CHD2 form subfamily I; CHD3, CHD4, and CHD5 form subfamily II; and CHD6, CHD7, CHD8, and CHD9 form subfamily III (Marfella and Imbalzano, 2007; Price and D'Andrea, 2013; Tyagi et al., 2016) (Figure 2). The subfamilies are categorized based on shared motifs: subfamily I proteins contain a DNA-binding domain in the C-terminal region; subfamily II proteins have two plant homeodomain (PHD) zinc-finger-like motifs in the N-terminal region; and subfamily III proteins have additional functional domains, such as Brahma and Kismet (BRK) domains and a SANT domain (Marfella and Imbalzano, 2007).



**Figure 2. Graphical illustration of CHD subfamilies.** Representative image adapted from Kolla et al. shows the three subfamilies of chromodomain helicase DNA-binding chromatin remodelers. Green diamonds represent PHD finger domains, red vertical boxes represent chromodomains, blue rectangular boxes represent SNF2-like helicase/ATPase domains, brown ovals represent a DNA-

binding motif, orange boxes represent domain of unknown function (DUF1, DUF2), light blue boxes represent a SANT domain, and yellow ovals represent BRK domains (Kolla et al., 2014).

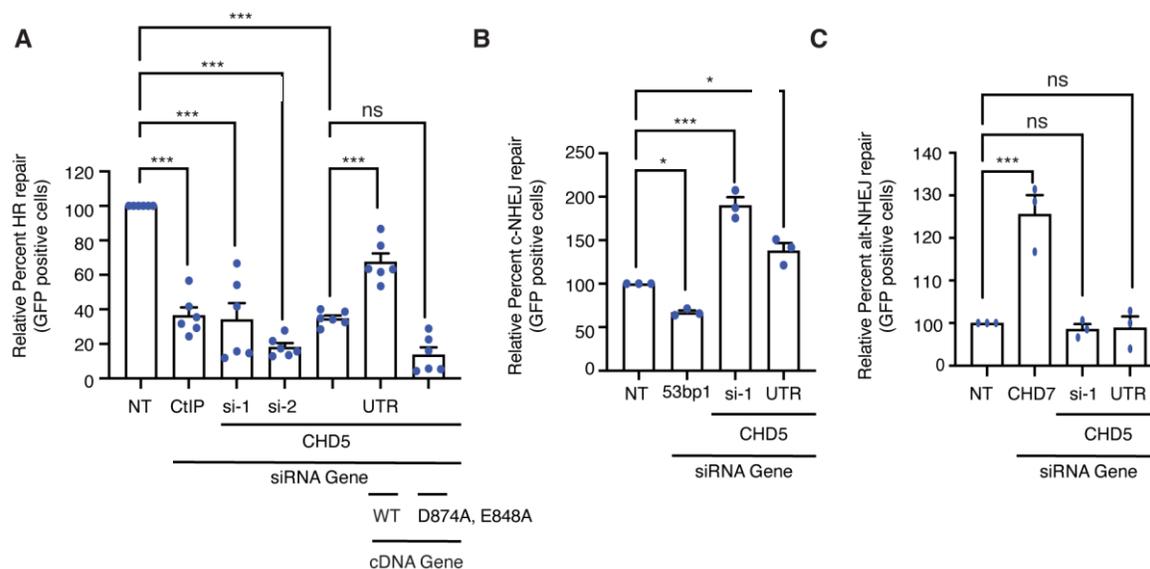
CHD5 was the first protein in the CHD family to be identified as playing a functional role in cancer (Bagchi et al., 2007). It is a potential tumor suppressor gene located at position 36.31 on the short arm of chromosome 1 (*1p36*), which is often deleted in several human cancers, including neuroblastoma, melanoma, and breast cancer, among many others (Bagchi et al., 2007; Fujita et al., 2008; Hall et al., 2014; Huang et al., 2020; Laut et al., 2022; Lv and Lin, 2022; Wu et al., 2012). CHD5 has also been found to be epigenetically silenced via DNA methylation in certain cancers, including lung, breast, and colon cancer, as well as renal cell carcinoma and glioma (Higashi et al., 2015; Huang et al., 2020; Laut et al., 2022; Lv and Lin, 2022; Wu et al., 2012).

Recent studies have shown that CHD5 positively regulates p53-mediated pathways and interacts with proteins involved in cell-cycle regulation, further suggesting its role in tumor suppression processes (Bagchi et al., 2007; Fujita et al., 2008). Quan et al. found that CHD5 represses the transcription of *WEE1*, a gene that plays a key part in the G2/M checkpoint of the cell cycle, but that has also been observed to act as an oncogene and is highly expressed in certain cancers (Quan et al., 2014). Another study found that CHD5 induces cell cycle G1 phase arrest and apoptosis to inhibit proliferation (Huang et al., 2020). Specifically, CHD5 was shown to activate pathways involving p53 and retinoblastoma protein (RB), both of which are known to help regulate cell cycle arrest and apoptosis (Huang et al., 2020).

Previous published work from our lab demonstrated that lower levels of CHD5 expression activate the DDR in human pancreatic cancer cells and was associated with decreased recurrence-free survival (RFS) and overall survival (OS) among patients (Hall et al., 2014). It was posited that increased activation of DDR may promote survival of pancreatic adenocarcinoma (PAC) via

heightened selection pressure and the development of resistance to DNA damage from chemotherapy and radiation therapy (Hall et al., 2014). However, the exact mechanism by which CHD5 contributes to DNA repair remains unknown.

Ongoing studies (unpublished results) from our lab have confirmed that CHD5 gets recruited to DNA double-strand break sites and that loss of CHD5 leads to decreased HR-repair efficacy, shifting repair pathways toward NHEJ, which is more mutagenic and therefore results in more genomic instability in the absence of HR (Figure 3). At DSBs, CHD5 appears to play a pivotal role in determining the specific pathway of repair, and because CHD5 has multiple functional domains, we aimed to explore which domains of the CHD5 protein are responsible for its recruitment to DNA DSBs. The results would shed light on the regulatory function of different domains in the recruitment of CHD5 to DSBs upon induction of DNA damage in cells.



**Figure 3. DNA repair reporter assays.** Unpublished results from the Yu Lab showing DNA repair efficacy after CHD5 knockdown using (A) DR-GFP, (B) EJ-5, and (C) EJ-2 assays. U2OS cells containing (A) an integrated DR-GFP homologous recombination repair (HR) reporter, (B) a classical non-homologous end joining (c-NEJ) reporter, and (C) an alternative non-homologous end joining (alt-NHEJ) reporter were silenced for CHD5 using multiple siRNAs, and transfected

with the I-SceI endonuclease and RFP (empty) or CHD5-RFP (WT, DHel) as shown. Live cells were assayed in flow cytometry. The RFP-positive cell population was sorted and percent of GFP-positive cells within the RFP-positive population was determined to analyze HR, c-NHEJ and alt-NHEJ repair efficacy. The mean  $\pm$  SEM from 3-5 biological experiments are shown. Data represented were analyzed using one-way ANOVA followed by Tukey test. (\*)  $p < 0.0332$  (\*\*)  $p < 0.0021$ , (\*\*\*)  $p < 0.0002$ .

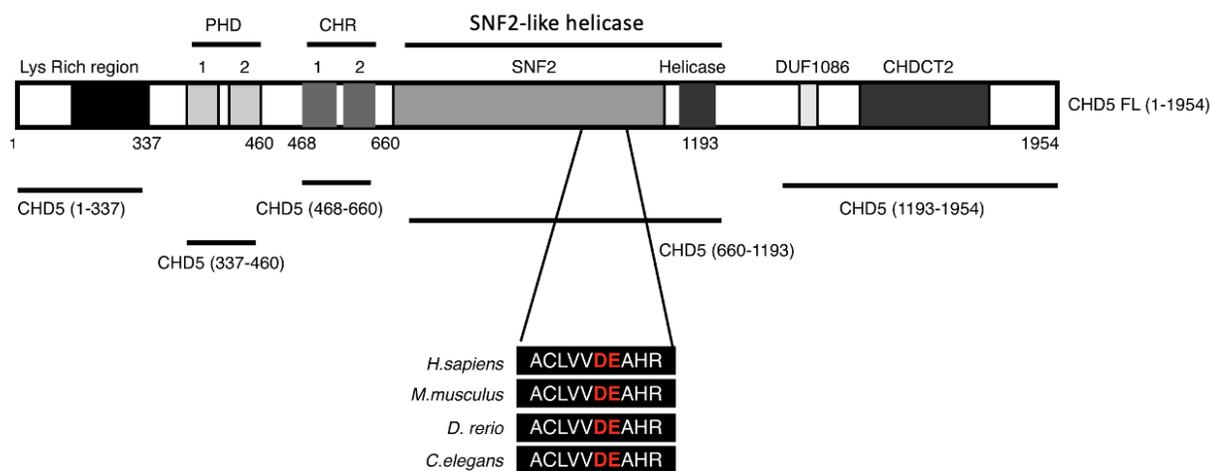
CHD5 has six different functional domains (Marfella and Imbalzano, 2007) (Figure 4). At the N-terminal, there is a lysine-rich region which can be helpful for binding DNA due to the positive charge of lysine and the negative charge of DNA (UniProt Q8TDI0). Adjacent to that are two plant homeodomains (PHDs), which are characteristic of CHD subfamily II proteins (Marfella and Imbalzano, 2007). CHD5 also contains two chromodomains, an SNF2/helicase domain, a domain of unknown function (DUF) and a CHDCT2 uncharacterized domain in the C-terminal region (Marfella and Imbalzano, 2007). It has been found that PHDs bind to the N-terminal residues of histone 3, and that this interaction is critical for CHD5's ability to prevent cellular proliferation, modulate other genes implicated in cancer, and suppress tumorigenesis (Paul et al., 2013). Chromodomains have also been found to bind histone 3, allowing for mediation of chromatin interactions (Paul et al., 2013). Finally, another study looking at a mutant helicase domain in CHD5 demonstrated that the helicase domain is similarly critical for proper associations with chromatin (Quan et al., 2014; Quan and Yusufzai, 2014).

### ***Approaches for investigating protein recruitment to DSBs***

We used two different approaches to study recruitment of CHD5 to DNA double-strand break sites. Both approaches involve the use of GFP-tagged CHD5 protein, as well as GFP-tagged truncated constructs of the CHD5 protein. We used five different truncated constructs, which each

contain a different domain type and are named based on the numbered positions of the amino acids they contain. The five truncated constructs we used are listed below with the domains they contain:

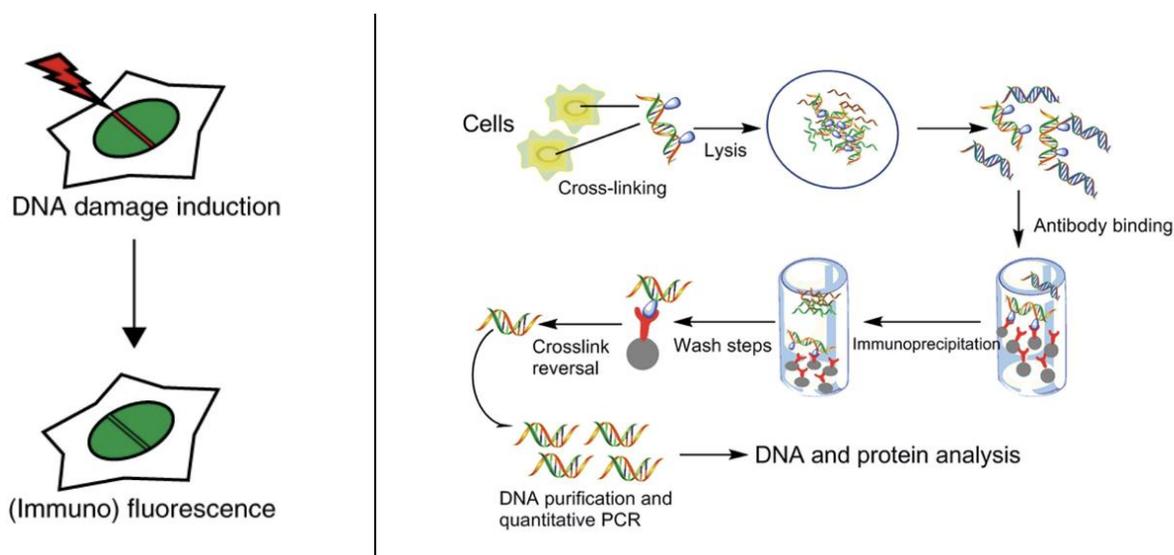
- GFP-CHD5 1-337 (lysine-rich region)
- GFP CHD5 337-460 (2 PHDs)
- GFP-CHD5 468-660 (2 chromodomains)
- GFP-CHD5 660-1193 (SNF2/helicase domain)
- GFP-CHD5 1193-1954 (domain of unknown function and CHDCT2)



**Figure 4. Schematic of CHD5 domains.** Illustration shows the domains of CHD5 and which domains are found in each of the truncated constructs. From N-terminal to C-terminal (left to right), the constructs and their corresponding domains are as follows: CHD5 1-337 (lysine-rich region), CHD5 337-460 (2 PHDs), CHD5 468-660 (2 CHRs), CHD5 660-1193 (SNF2-like helicase), and CHD5 1193-1954 (domain of unknown function and uncharacterized CHDCT2).

The first experimental approach was to use live microscopy and laser microirradiation, which has been utilized in previous studies to induce DNA DSBs (Holton et al., 2017) (Figure 5). These studies confirmed the successful induction of DNA DSBs by staining for  $\gamma$ H2AX, a marker of double-strand breaks. Notably, most of the DNA damage induced by laser microirradiation at a

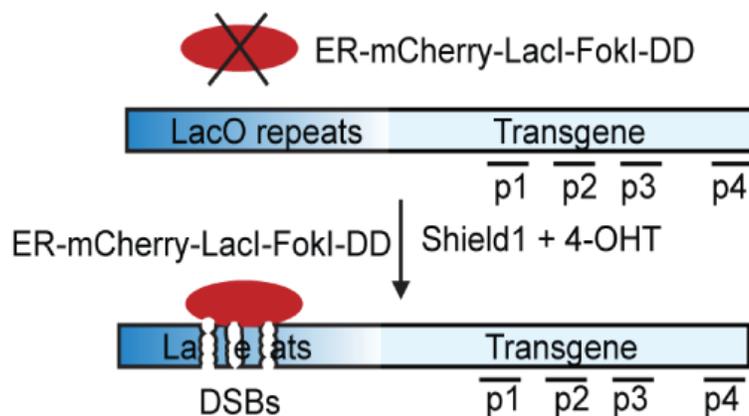
wavelength of 405 nm with moderate energy levels consists of double-strand breaks; thus, these were the wavelength and energy level used for this study (Muster et al., 2017). Following damage induction, images were taken of cells transfected with full-length GFP-CHD5 or one of its associated constructs at regular intervals. Finally, ImageJ was used to analyze recruitment and retention kinetics.



**Figure 5. Schematic of laser-induced DNA damage and protein recruitment (left) and chromatin immunoprecipitation (right).** The diagram on the left illustrates the process from laser-induced DNA damage to protein recruitment at the site of damage, which is visualized by immunofluorescence. The diagram on the right, adapted from Song et al., illustrates the primary steps that occur during the procedure for chromatin immunoprecipitation (ChIP). “Chromatin immunoprecipitation” from Choosing a suitable method for the identification of replication origins in microbial genomes by Song et al., 2015 is licensed [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/).

The second approach was to use chromatin immunoprecipitation (ChIP) (Figure 5) and real-time PCR (Figure 7) with an engineered cell line. Specifically, we used the U2OS-235 mCherry-LacI-FokI cell line (Shanbhag and Greenberg, 2013). These human osteosarcoma (U2OS) cells contain lac operator (LacO) repeats adjacent to a transgene, which are together integrated into a single location within the genome. Additionally, these cells include the fusion

protein ER-mCherry-LacI-FokI-DD, which is formed by fusing the FokI endonuclease protein to mCherryLacI, a modified estradiol receptor (ER), and a destabilization domain (DD) (Shanbhag and Greenberg, 2013) (Figure 6). Upon treatment with 4-Hydroxytamoxifen (4-OHT) and Shield1, the ER allows for nuclear translocation of the fusion protein and the DD allows for stabilization (Shanbhag and Greenberg, 2013). Within the nucleus, LacI helps with targeting and binding to the lac operator repeats, where the FokI nuclease creates DSBs (Shanbhag and Greenberg, 2013). Meanwhile, the red fluorescent mCherry protein allows for visualization and confirmation of localization at DSBs. This cell line was transfected with full-length GFP-CHD5 or one of its associated constructs, allowing for their overexpression. Following induction of double-strand breaks, ChIP and real-time PCR were used to quantify protein enrichment at DSBs. Because the exact break site cannot be sequenced, the p1 and p4 regions in the adjacent transgene were sequenced during real-time PCR to serve as a representation of the amount of CHD5 enrichment at the DSB site, as CHD5 also binds to the adjacent areas. Overall, while this approach does not allow us to observe the kinetics of recruitment, it still allows us to determine which domain of CHD5 is responsible for its recruitment to DSBs.



**Figure 6. Schematic showing FokI induction theme in U2OS cells.** The ER-mCherry-LacI-FokI-DD fusion protein binds to lac operator repeats, where FokI makes double-strand breaks.

The p1 and p4 regions in the adjacent transgene were sequenced during real-time PCR to quantify protein enrichment at the DSB site.

### ***Hypothesis***

CHD5 has multiple domains that have been characterized to bind unmodified or modified histones, and/or perform chromatin remodeling activity. We have found that CHD5 recruits to DNA DSBs, but of all the domains that CHD5 has, it remains unclear which of them actually helps CHD5 get recruited to DNA DSBs upon induction of DNA damage. In line with this, our present study aims to identify which domain of CHD5 helps in its recruitment to DNA DSBs. We also aim to understand which domain facilitates CHD5's retention at the break site to help it promote its DNA repair activity. We hypothesize that the lysine-rich region at the N-terminal, the zinc-finger-like plant homeodomains, and the chromodomains will be responsible for CHD5 recruitment to DSBs because it is widely known in literature that PHDs and chromodomains facilitate binding of proteins to DNA or histones (Paul et al., 2013). Additionally, the positive charge of lysine may be helpful for binding to negatively-charged DNA. Based on this hypothesis, we expect to observe GFP-CHD5 1-337, GFP-CHD5 337-460, and GFP-CHD5 468-660 as having the greatest enrichment at DSBs, while also observing negligible enrichment for the two other constructs (GFP-CHD5 660-1193 and GFP-CHD5 1193-1954). Interestingly, we find that GFP-CHD5 1-337 has the highest amount of enrichment, while GFP CHD5 337-460, GFP-CHD5 468-660, GFP-CHD5 660-1193 show enrichment to the same levels as full-length GFP-CHD5, suggesting that all domains except the domain of unknown function and CHDCT2 play a role in facilitating recruitment of CHD5 to DNA DSBs.

## **Materials and Methods**

### ***Cell Lines***

U2OS mammalian cell line was purchased from American Type Culture Collection (ATCC, Manassas, VA). U2OS-235 mCherry-LacI-FokI cell line was provided by Dr. Roger Greenberg. Both cell lines were cultured in 100 mm plates (Corning, #439293) and grown in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, #11995065) supplemented with 10% FBS. All GFP-tagged CHD5 plasmids were provided by the Custom Cloning Core Division of the Emory Integrated Genomics Core.

### ***Plasmid Extraction and Purification***

Plasmids expressing full-length GFP-CHD5 and its truncated constructs were transformed using DH-5 $\alpha$  competent cells. A single colony was isolated and grown in LB media with ampicillin or kanamycin and incubated overnight in an orbital shaker at 30°C and 220 rpm. The following day, the culture was centrifuged for 25 minutes at 4°C and 6000 rpm using a JA-10 rotor. The supernatant was discarded, and the pellet was kept. Plasmid extraction and purification were performed using QIAGEN Plasmid Maxi Kit (#12163) as per manufacturer's instructions. Plasmid concentrations were measured using Thermo Scientific NanoDrop 1000 Spectrophotometer.

### ***Transfections***

Plasmid transfections were performed to overexpress full-length GFP-CHD5 and its truncated constructs. Cells were seeded in 35 mm plates (MatTek) for laser microirradiation assays and 6-well plates for ChIP assays. 3 $\mu$ g of the desired plasmid was transfected using 2.5 $\mu$ g Lipofectamine 3000 Amplifier and 7.5 $\mu$ g Lipofectamine 3000 Reagent as per manufacturer's instructions

(ThermoFisher). Transfection was performed 1-hour post-plating for laser microirradiation experiments and 24 hours post-plating for ChIP assays.

### ***Laser Microirradiation***

Andor MicroPoint was used to induce double-stranded DNA breaks in transfected U2OS cells 24 hours after transfection. iQ3 Live Cell Imaging Software (Andor IQ.3.6.6) was used to control laser microirradiation. The following settings were used: 17 Hz repetition rate, 7 repeats, 75% energy output (arbitrary normalized fluorescence units in pixels). The laser wavelength was 405 nm. Cells were simultaneously imaged using ZEN microscopy software (ZEISS). For each cell, images were taken immediately following laser microirradiation. Images were taken every 30 seconds for 31 cycles, equaling 15 minutes total.

### ***ImageJ Quantification***

Fiji, a distribution of ImageJ, was used to quantify GFP fluorescence in cells following laser microirradiation. Images captured using ZEN microscopy software at the 0, 1-, 5-, 10-, and 15-minute time points were opened in Fiji and analyzed using the ROI Manager tool. Freehand selection was used to outline the cell nucleus and the region that experienced DNA damage via laser microirradiation. The outlined area of the cell nucleus was also used to capture background fluorescence. For each cell, fluorescent intensity at the DNA break site was calculated using the following equation:  $(\text{break site fluorescence} - \text{background fluorescence}) / (\text{cell nucleus fluorescence} - \text{background fluorescence})$ . To find relative intensity at each time point, the values calculated using this equation were normalized to the value calculated for the 0-minute time point.

### ***ChIP Assays***

To induce double-stranded breaks in DNA, (Sigma-Aldrich, #H7904) and Shield1 (Takara Bio, #632189) were diluted in DMEM (1: 20000 and 1: 500, respectively) and added to transfected U2OS-235 mCherry-LacI-Fok1 cells 48 hours post-transfection. After incubating for 4 hours, 1% formaldehyde (Sigma-Aldrich, #252549) (diluted in PBS) was added to induce crosslinking. Following 15 minutes incubation at room temperature, cells were washed twice with PBS and 0.125M glycine (Sigma-Aldrich, #G7126) (diluted was nuclease-free water) to stop excessive crosslinking due to formaldehyde. Cells were washed again with PBS and scraped to be collected and centrifuged for 15 minutes at 4°C and 13.2 rpm. The supernatant was removed, and the pellet was kept. Lysis buffer, prepared by combining 1x RIPA buffer (Millipore, #20-188) (diluted in nuclease-free water) with DTT and a protease and phosphatase inhibitors cocktail, was added to the samples, which were then incubated on ice for 30 minutes. The samples then underwent 25 cycles of sonication (one cycle = 30 seconds on and off) using Bioruptor® Pico sonication device (Diagenode) to break up the DNA into fragments of less than 500 base pairs. This was followed by another 10 minutes of centrifugation at 4°C and 13.2 rpm. DNA concentrations were then measured using Thermo Scientific NanoDrop 1000 Spectrophotometer. 1.5µg of each sample was added to individual wells in an 8-well assay strip from the EpigenTek *ChromaFlash™ One-Step ChIP Kit* (#P-2025). 1µL of anti-GFP antibody (Abcam, #ab290, produced in rabbit) was also added to each well. In one well, 0.8µL of non-immune IgG was added in place of anti-GFP antibody to serve as a negative control. Finally, CH2 ChIP buffer was added to each well such that the total volume in each well equaled 100µL. From here, ChIP was completed using EpigenTek *ChromaFlash™ One-Step ChIP Kit* (#P-2025) per manufacturer's instructions.

***Real-Time Quantitative PCR***

Following ChIP, real-time quantitative PCR (qPCR) was performed using Applied Biosystems™ 7500 Fast Real-Time PCR system (ThermoFisher). Samples were loaded into a 96-well plate, with 6 wells per sample and 2μL of the DNA sample per well. 10μL of PowerTrack™ SYBR Green Master Mix (#A4601, ThermoFisher), which contains antibody-mediated hot start DNA polymerase, was also added to each well. Finally, 1μL each of forward and reverse primer was added. The following primers were used:

**p1,**

forward: GGAAGATGTCCCTTGTATCACCAT

reverse: TGGTTGTCAACAGAGTAGAAAGTGAA

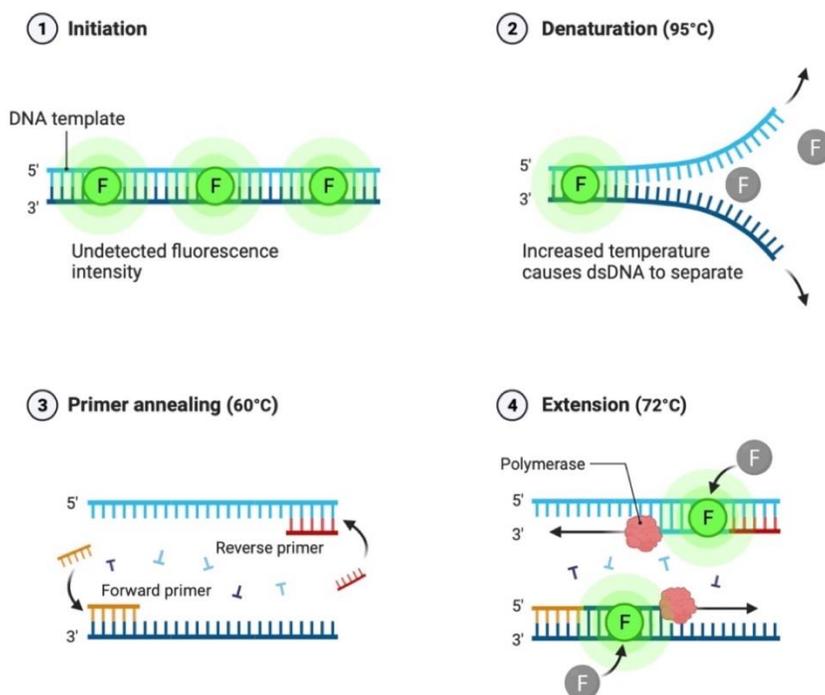
**p4,**

forward: CCACCTGACGTCTAAGAAACCAT

reverse: GATCCCTCGAGGACGAAAGG

The parameters used for the PCR system include a pre-warming stage of 50°C for 20 minutes, during which enzymes are activated; a denaturation period of 95°C for 10 minutes; and 40 cycles of melting at 95°C and annealing at 65°C. Upon completion of real-time qPCR, the Livak method was used with the resulting quantitative data to calculate relative fold enrichment of each of the protein constructs.

## Fluorescent Dye-Based Real Time PCR (qPCR)



**Figure 7. Real-time qPCR schematic.** Diagram illustrating the primary steps that occur during qPCR. The temperature is increased to 95°C to allow for denaturation of dsDNA, and then lowered to 60°C to allow for annealing of forward and reverse primers. Finally, DNA polymerase synthesizes new strands of DNA. Fluorescent dye (SYBR™ Green) binds to double-stranded DNA. The number of cycles required for the fluorescent signal to cross a detection threshold (detection beyond background signal) is represented by the Ct number, which is given in the resulting quantitative data. Figure courtesy of BioRender.

### *Protein Lysate Preparation and Western Blotting*

Transfected cells were harvested and resuspended in lysis buffer with 0.75% CHAPS, DTT, and a protease and phosphatase inhibitors cocktail. Protein concentrations were determined by adding samples to diluted Bradford reagent (Bio-Rad Protein Assay Dye Reagent Concentrate, #5000006) and measuring in BioMate™ 3S Spectrophotometer (Thermo Scientific, #14-386-503) at an absorbance wavelength of 595 nm. 50µg of protein was added to 5µL of SDS loading dye (5X)

and lysis buffer (without CHAPS) for a total volume of 30 $\mu$ L. The samples were heated at 100°C for 7 minutes and then loaded onto an 8% SDS-PAGE gel with 10 $\mu$ L of protein ladder (Precision Plus Protein Dual Color Standards, #1610374). The gel was run at 80V for 20 minutes, and then 120V for 1.5 hours. Afterwards, the gel was transferred onto a membrane overnight at 40V and 4°C.

The following day, the membrane was blocked in a solution of 5% w/v BSA in PBST (PBS with Tween 20) for 1 hour at room temperature on a shaker apparatus. The following primary antibodies were then added and incubated for 1 hour at room temperature on the shaker apparatus:

- Anti-GFP (Santa Cruz Tech, #SC9996, 1:1000, produced in mouse)
- Anti-GAPDH (Sigma-Aldrich, #G9545, 1:2000, produced in rabbit)

Anti-GFP antibody was used to view our proteins of interest (CHD5 and its associated truncated constructs), which were tagged with GFP. GAPDH was used as a loading control. After primary antibody incubation, the blot was washed three times in PBST (20 minutes each round) on the shaker apparatus at room temperature. The following secondary antibodies were then added and incubated for 1 hour at room temperature on the shaker:

- Donkey anti-mouse IR Dye 680RD (Licor Biosciences, # 926-68072, 1:5000)
- Donkey anti-rabbit IR Dye 800CW (Licor Biosciences, #926-32213, 1:5000)

The blot was washed twice in PBS (20 minutes each round) on the shaker at room temperature. It was then developed using the Li-Cor Odyssey system with ImageStudio 5.2 software.

### ***Statistical Analysis***

ChIP and real-time qPCR were performed in a biological triplicate (each biological replicate had three experimental replicates). All values plotted represent the mean. Error bars represent standard

error of the mean. Two-tailed Student's t-test was used for comparing the relative fold enrichment of full-length GFP-CHD5 and each truncated construct against the negative control. One-way ANOVA was used for comparing relative fold enrichment of each truncated construct against that of full-length GFP-CHD5.

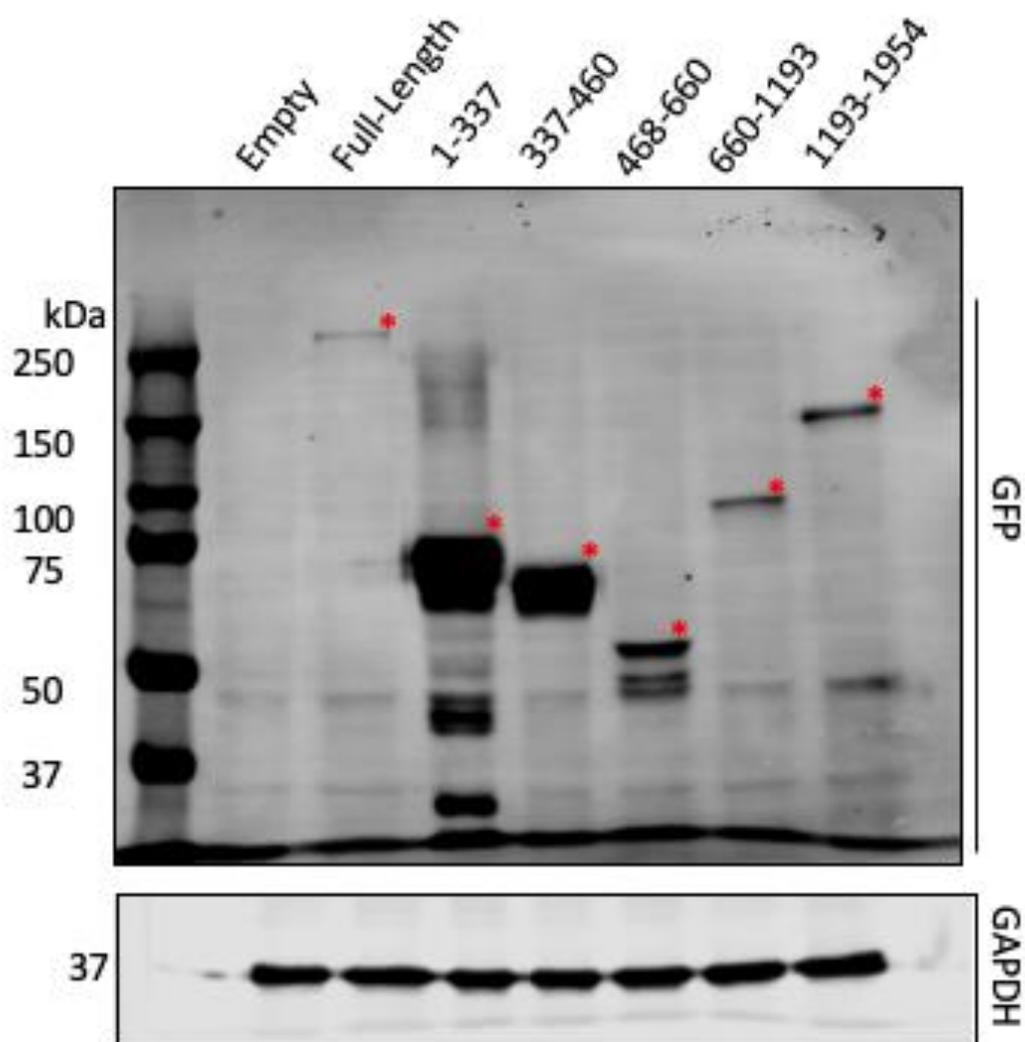
## Results

### *Confirmation of plasmid expression*

We used western blotting to confirm the expression of our plasmids encoding full-length GFP-CHD5 and the following CHD5 truncated constructs: GFP-CHD5 1-337, GFP CHD5 337-460, GFP-CHD5 468-660, GFP-CHD5 660-1193, and GFP-CHD5 1193-1954. We loaded one well with empty GFP vector to serve as our control. The molecular weights of full-length GFP-CHD5 and the truncated constructs are as follows:

- Full-Length GFP-CHD5 — 257 kDa
- GFP-CHD5 1-337 — 64.5 kDa
- GFP-CHD5 337-460 — 37.72 kDa
- GFP-CHD5 468-660 — 52.28 kDa
- GFP-CHD5 660-1193 — 88.3 kDa
- GFP-CHD5 1193-1954 — 113.32 kDa

Our western blot showed equal-sized bands for GAPDH across all samples, indicating that equal amounts of each sample were loaded. We also observed bands for each sample that were consistent with their respective molecular weights, confirming correct expression of each plasmid (Figure 8).



**Figure 8. Western blot of GFP-CHD5 and its GFP-tagged truncated constructs.** U2OS cells transfected with plasmids for GFP-CHD5 or one of the GFP-tagged truncated constructs were harvested and lysed. Western blot analysis was performed using anti-GFP antibody to confirm expression of plasmids. Red dots indicate that the observed molecular weight is consistent with the predicted molecular weight. GAPDH (MW = 36 kDa) was used as the loading control.

### *Quantifying recruitment kinetics of CHD5*

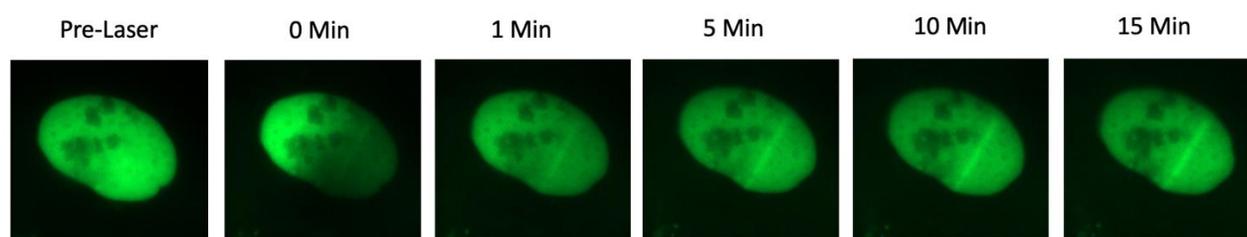
Previously completed unpublished work from our lab showed induction of DNA damage on U2OS cells transfected with either full-length GFP-CHD5 or one of the CHD5 truncated constructs via laser microirradiation under a previous version of the iQ3 Live Cell Imaging Software. These experiments were performed at the following settings: 15 Hz repetition rate, 5 repeats, and 70% energy output. Imaging results from these experiments visually displayed recruitment of full-length GFP-CHD5, GFP-CHD5 1-337, GFP CHD5 337-460, GFP-CHD5 468-660, and GFP-CHD5 660-1193 to the DNA break site by the one-minute mark post-damage induction (Figure S1). GFP-CHD5 1193-1954 did not show any recruitment within the 15 minutes following damage induction (Figure S1). Our work sought to replicate these results.

Our experiments were conducted under the newly updated version of the laser software (Andor IQ.3.6.6). Re-standardization of laser settings under this version of the software determined that the appropriate settings which allowed for sufficient induction of DNA damage without causing immediate cell death were as follows: 17 Hz repetition rate, 7 repeats, and 75% energy output. Under these settings, we induced DNA damage via laser microirradiation to U2OS cells transfected with either full-length GFP-CHD5 or one of the CHD5 truncated constructs.

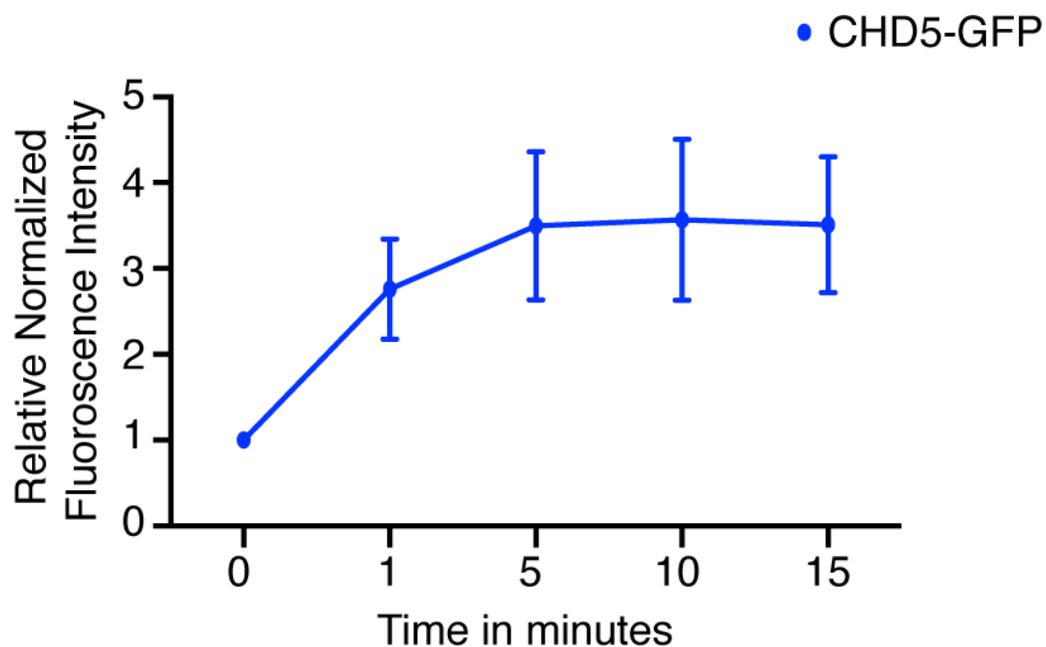
Recruitment kinetics quantified using ImageJ showed that full-length GFP-CHD5 recruited to the break site within one minute of damage induction (Figure 9, Figure 10). Average relative intensity (in pixels) peaked at the one minute time point and stayed consistent for the remaining 14 minutes. These results indicate that full-length GFP-CHD5 not only recruits quickly to the DNA break site following damage induction, but also retains at the break site for an extended period.

For the truncated constructs, laser microirradiation experiments were performed on cells transfected with GFP-CHD5 1-337, GFP CHD5 337-460, GFP-CHD5 468-660, GFP-CHD5 660-

1193, and GFP-CHD5 1193-1954. Over twenty experiments were performed for each construct (the specific number of experiments performed for each construct varied). However, no or minimal recruitment was observed for each of these truncated constructs, which was inconsistent with our lab's previously completed unpublished work (Figure S2). Because we were unable to replicate these prior results despite numerous repeated experiments, we believed that we had encountered a technical issue with the laser software, which will necessitate further inspection.



**Figure 9. Full-length GFP-CHD5 recruits to laser-induced DNA DSB sites.** Representative images of U2OS cells transfected and overexpressed with full-length GFP-CHD5 before and after subsection to laser microirradiation. Images of the cell nucleus were taken at the 0, 1, 5, 10, and 15-minute time points following laser microirradiation.



**Figure 10. GFP-CHD5 recruits to DNA DSBs within 1 minute following laser-induced damage and retains for 15 minutes.** Quantification of fluorescence intensity at the DSB site was found using fluorescence values acquired through ImageJ/Fiji and calculated using the equation (break site fluorescence – background fluorescence) / (cell nucleus fluorescence – background fluorescence). Relative intensity was calculated by normalizing all time point values to the value found at 0 min. The plotted points represent the average DSB fluorescence intensities for 17 different cells. Error bars represent standard error of the mean.

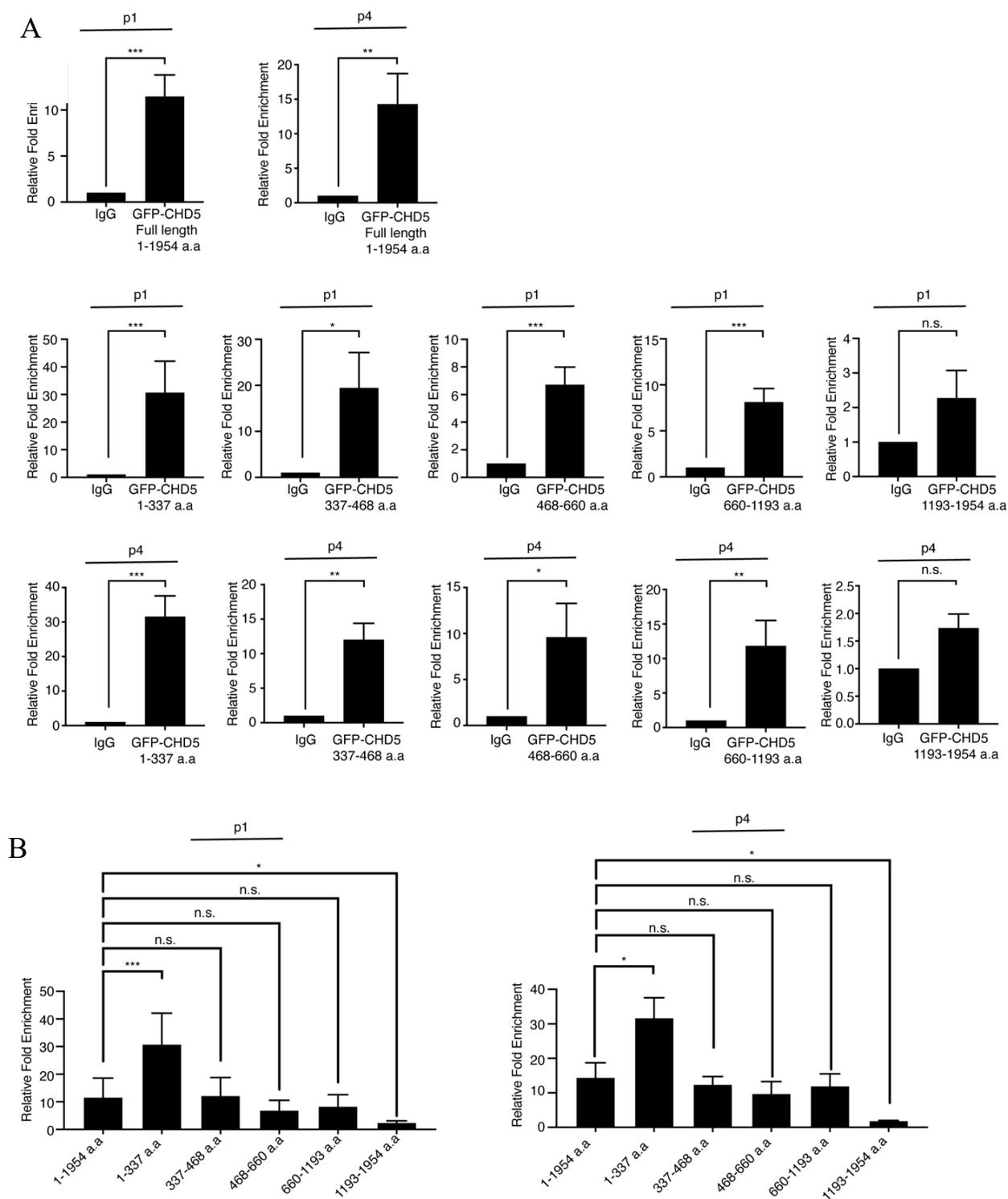
*ChIP reveals GFP-CHD5 1-337 construct has the highest enrichment at double-strand breaks*

Chromatin immunoprecipitation (ChIP) and real-time PCR were used to determine which domain of CHD5 is responsible for its recruitment to double-strand breaks. We began crosslinking 4 hours after adding 4-Hydroxytamoxifen (4-OHT) and Shield1 to induce double-strand breaks in U2OS-235 mCherry-LacI-FokI cells overexpressing full-length GFP-CHD5, GFP-CHD5 1-337, GFP-CHD5 337-460, GFP-CHD5 468-660, GFP-CHD5 660-1193, or GFP-CHD5 1193-1954. Results from real-time PCR showed that when compared to non-specific pulldown from IgG, which served as our negative control, full-length GFP-CHD5, GFP-CHD5 1-337, GFP-CHD5 337-460, GFP-CHD5 468-660, and GFP-CHD5 660-1193 all had significantly greater relative fold

enrichment at both the p1 and p4 regions of the transgene located adjacent to the LacO repeats, which was the site of our DNA double-strand breaks (Figure 11). Meanwhile, relative fold enrichment of GFP-CHD5 1193-1954 was not significantly different from that of the negative control (Figure 11). These results indicate that full-length GFP-CHD5 and all truncated constructs except for GFP-CHD5 1193-1954, which contains a domain of unknown function and a CHDCT2 uncharacterized domain, recruited to the double-strand break sites.

We also compared relative fold enrichment of each of the truncated constructs with that of full-length GFP-CHD5. GFP-CHD5 1-337 showed significantly greater enrichment, while GFP-CHD5 1193-1954 showed significantly less enrichment. The remaining three constructs (GFP-CHD5 337-460, GFP-CHD5 468-660, and GFP-CHD5 660-1193) did not show significantly different levels of enrichment compared to full-length GFP-CHD5. These results indicate that the GFP-CHD5 337-460, GFP-CHD5 468-660, and GFP-CHD5 660-1193 truncated constructs are able to independently recruit to double-strand break sites as effectively as full-length GFP-CHD5. Meanwhile, the GFP-CHD5 1-337 truncated construct can recruit to double-strand break sites more effectively than full-length GFP-CHD5.

Altogether, these results suggest that all domains except for the domain of unknown function and CHDCT2 play a role in facilitating recruitment of full-length GFP-CHD5 to double-strand break sites. However, the lysine-rich region at the N-terminal appears to be the domain which is most responsible for recruitment.



**Figure 11. GFP-CHD5 1-337 shows greatest enrichment at DSB sites.** ChIP was performed on U2OS-235 mCherry-LacI-Fok1 cells transfected and overexpressed with plasmids for GFP-CHD5 or one of the GFP-tagged truncated constructs 4 hours after addition of 4-OHT and Shield1. qPCR was performed to quantify protein enrichment at DSBs. The Livak method was used to calculate

relative fold enrichment. (A) Two-tailed Student's t-test was used to compare the relative fold enrichment of full-length GFP-CHD5 and each truncated construct against pulldown from non-immune IgG. (B) One-way ANOVA was used to compare relative fold enrichment of each truncated construct against that of full-length GFP-CHD5. (A-B) Error bars represent standard error of the mean. N = 3. (\*\*\*)  $p < 0.001$ , (\*\*)  $p < 0.01$ , (\*)  $p < 0.05$ , (n.s.)  $p \geq 0.05$ .

## Discussion

DNA double-strand breaks are extremely deleterious and present as a threat to genome stability if left unrepaired, potentially giving rise to disease states such as cancer (Groelly et al., 2023; Jeggo et al., 2016). Of the two main repair pathways for DNA DSBs, NHEJ is faster and more efficient, but it is also more mutagenic (Mao et al., 2008). Therefore, the choice of repair pathway and the balance between the two pathways within a cell is significant (Clouaire and Legube, 2015). Unpublished results from our lab demonstrate that loss of the chromatin remodeler protein CHD5 leads to decreased HR-repair efficacy, shifting the repair pathway balance toward NHEJ. Altogether, these results suggest that CHD5 likely plays a critical role in determining the pathway by which DNA DSBs are repaired.

Because CHD5's presence at DNA DSBs appears highly consequential, our study aimed to examine the domains of CHD5 to help us better understand the recruitment of CHD5 to DSB sites. Our findings demonstrate that all domains of CHD5 except for the domain of unknown function and the CHDCT2 uncharacterized domain can recruit independently to DSB sites, reaching equal or even greater enrichment levels compared to those of full-length CHD5, thereby suggesting that all domains except the domain of unknown function and CHDCT2 contribute to the facilitation of CHD5 recruitment to DSB sites. Notably, the truncated construct that contains the lysine-rich N-terminal region showed greater enrichment at DSB sites than full-length CHD5, suggesting that this domain, when on its own, recruits to DSB sites more effectively. We posit that this finding may be in part due to the truncated construct's smaller size as compared to full-length CHD5, allowing it to move to DSB sites more efficiently. Additionally, because the truncated construct lacks most of the residues that are present in full-length CHD5, the truncated construct will differ somewhat in folding and structure, which may in turn influence its binding capacity.

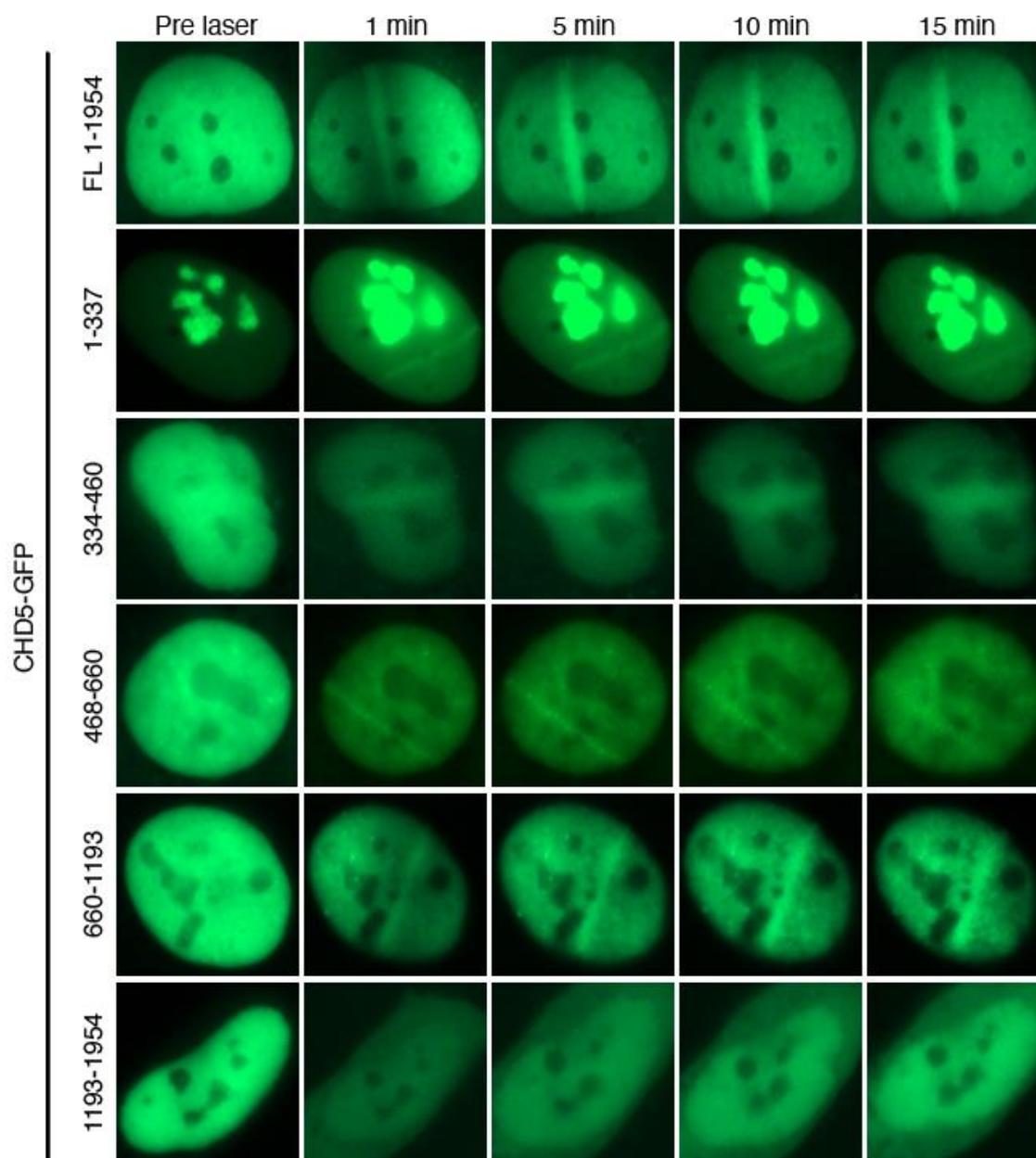
For instance, full-length CHD5 may contain additional interactions that interfere with its overall binding ability. To address this, future experiments could involve the use of full-length plasmids with mutated domains, rather than truncated constructs, to help further elucidate the roles of the different domains in facilitating recruitment to DSBs. Overall, however, due to the negative charge of lysine, which would be strongly attracted to the positive charge of DNA, it is unsurprising that the lysine-rich region of CHD5 demonstrated strong recruitment.

Using laser microirradiation, we also examined recruitment and retention kinetics of full-length GFP-CHD5. Our findings demonstrate that full-length CHD5 both recruits within 1 minute and retains at DSB sites for at least 15 minutes. As a result of technical challenges with our laser equipment, we were unable to determine recruitment and retention kinetics for the truncated constructs. Therefore, in future experiments, we hope to be able to use laser microirradiation to investigate retention kinetics and answer the question of which domain is responsible for the retention of CHD5 at DSB sites. An analysis of recruitment kinetics for all constructs would also provide further insight into the roles of the different domains in facilitating recruitment of CHD5 to DSBs. Future work may also include imaging for longer periods of time to determine the full duration for which CHD5 and its truncated constructs retain at DSB sites.

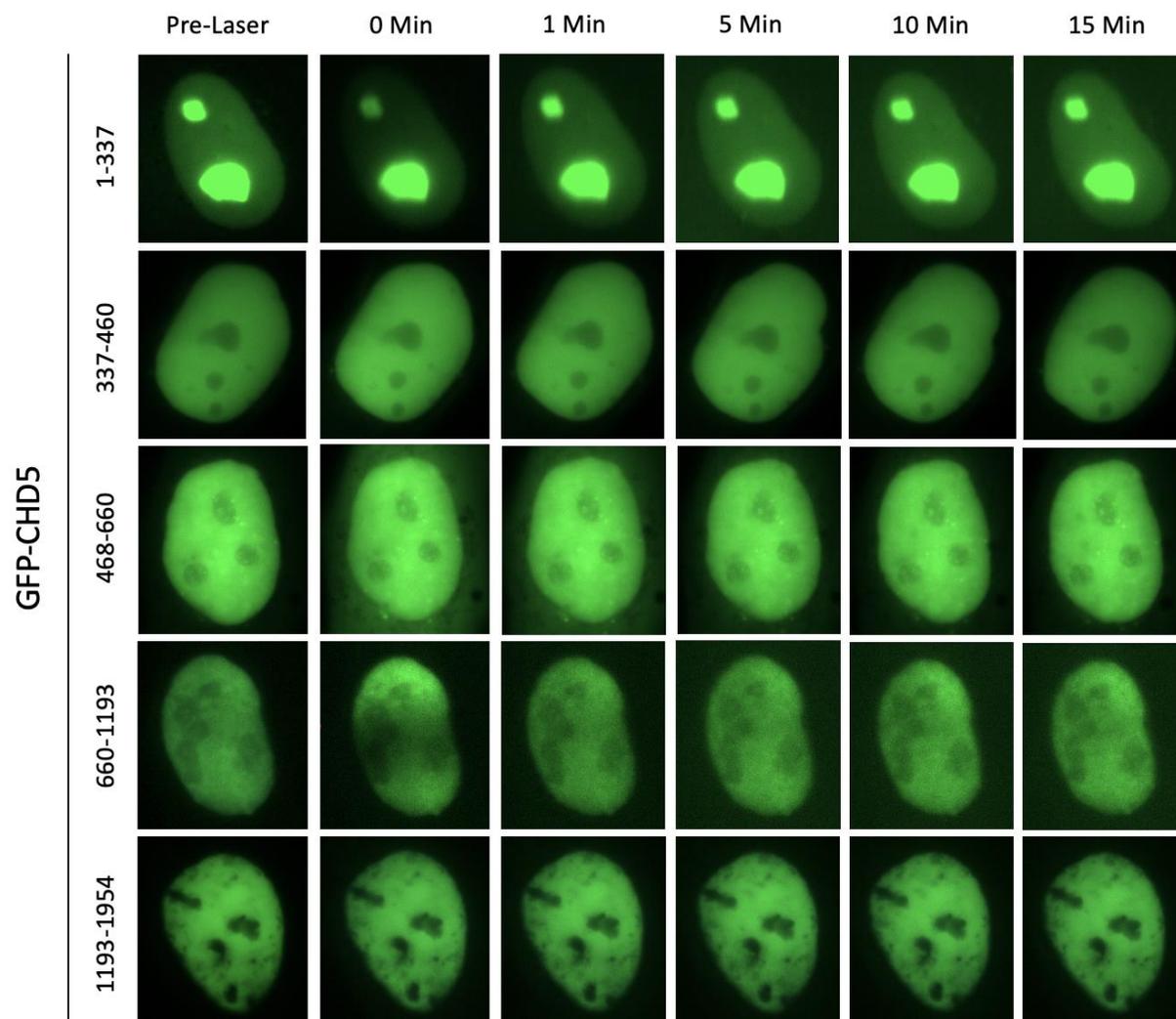
In summary, our study suggests that all domains except for the domain of unknown function and the CHDCT2 uncharacterized domain participate in recruitment of CHD5 to double-strand break sites, with the lysine-rich region at the N-terminal playing the biggest role. Additionally, our findings reveal that CHD5 both recruits quickly and retains at DSB sites for an extended period. Future work will aim to further clarify the specific roles of CHD5's domains in its recruitment and retention at DSBs via an analysis of recruitment and retention kinetics. Our hope is to be able to apply this knowledge toward our understanding of DSB repair pathway choice

and inform research exploring the targeting of these pathways as potential approaches for cancer therapeutics.

## Supplementary Figures



**Figure S1. Recruitment of GFP-CHD5 and its associated truncated constructs to laser-induced DNA DSB sites.** Unpublished results from the Yu Lab. Laser microirradiation was conducted under a previous version of the iQ3 Live Cell Imaging Software. Damage was induced in cells overexpressing full-length GFP-CHD5 (top row) or one of its truncated constructs. Images of the cell nucleus were taken at the 1, 5, 10, and 15-minute time points following laser microirradiation.



**Figure S2. U2OS cells transfected with GFP-CHD5 truncated constructs before and after laser microirradiation.** Laser microirradiation was performed on cells overexpressing GFP-CHD5 truncated constructs using the newly updated version of the iQ3 Live Cell Imaging Software. The settings used were 17 Hz repetition rate, 7 repeats, and 75% energy output. Images of the cell nucleus were taken at the 0, 1, 5, 10, and 15-minute time points following laser microirradiation. No protein recruitment was observed for all truncated constructs.

## References

- Bagchi, A., Papazoglu, C., Wu, Y., Capurso, D., Brodt, M., Francis, D., Bredel, M., Vogel, H., and Mills, A.A. (2007). CHD5 is a tumor suppressor at human 1p36. *Cell* *128*, 459-475.
- Clouaire, T., and Legube, G. (2015). DNA double strand break repair pathway choice: a chromatin based decision? *Nucleus* *6*, 107-113.
- Fujita, T., Igarashi, J., Okawa, E.R., Gotoh, T., Manne, J., Kolla, V., Kim, J., Zhao, H., Pawel, B.R., London, W.B., *et al.* (2008). CHD5 , a Tumor Suppressor Gene Deleted From 1p36.31 in Neuroblastomas. *JNCI: Journal of the National Cancer Institute* *100*, 940-949.
- Groelly, F.J., Fawkes, M., Dagg, R.A., Blackford, A.N., and Tarsounas, M. (2023). Targeting DNA damage response pathways in cancer. *Nature Reviews Cancer* *23*, 78-94.
- Hall, W.A., Petrova, A.V., Colbert, L.E., Hardy, C.W., Fisher, S.B., Saka, B., Shelton, J.W., Warren, M.D., Pantazides, B.G., Gandhi, K., *et al.* (2014). Low CHD5 expression activates the DNA damage response and predicts poor outcome in patients undergoing adjuvant therapy for resected pancreatic cancer. *Oncogene* *33*, 5450-5456.
- Higashi, M., Kolla, V., Iyer, R., Naraparaju, K., Zhuang, T., Kolla, S., and Brodeur, G.M. (2015). Retinoic acid-induced CHD5 upregulation and neuronal differentiation of neuroblastoma. *Mol Cancer* *14*, 150.
- Holton, N.W., Andrews, J.F., and Gassman, N.R. (2017). Application of Laser Micro-irradiation for Examination of Single and Double Strand Break Repair in Mammalian Cells. *J Vis Exp*.
- Huang, S., Yan, Q., Xiong, S., Peng, Y., Zhao, R., and Liu, C. (2020). Chromodomain Helicase DNA-Binding Protein 5 Inhibits Renal Cell Carcinoma Tumorigenesis by Activation of the p53 and RB Pathways. *Biomed Res Int* *2020*, 5425612.

- Jeggo, P.A., Pearl, L.H., and Carr, A.M. (2016). DNA repair, genome stability and cancer: a historical perspective. *Nature Reviews Cancer* *16*, 35-42.
- Kolla, V., Zhuang, T., Higashi, M., Naraparaju, K., and Brodeur, G.M. (2014). Role of CHD5 in human cancers: 10 years later. *Cancer Res* *74*, 652-658.
- Laut, A.K., Dorneburg, C., Fürstberger, A., Barth, T.F.E., Kestler, H.A., Debatin, K.M., and Beltinger, C. (2022). CHD5 inhibits metastasis of neuroblastoma. *Oncogene* *41*, 622-633.
- Lv, Y., and Lin, W. (2022). Comprehensive analysis of the expression, prognosis, and immune infiltrates for CHDs in human lung cancer. *Discov Oncol* *13*, 29.
- Mao, Z., Bozzella, M., Seluanov, A., and Gorbunova, V. (2008). Comparison of nonhomologous end joining and homologous recombination in human cells. *DNA Repair (Amst)* *7*, 1765-1771.
- Marfella, C.G.A., and Imbalzano, A.N. (2007). The Chd family of chromatin remodelers. *Mutat Res* *618*, 30-40.
- Muster, B., Rapp, A., and Cardoso, M.C. (2017). Systematic analysis of DNA damage induction and DNA repair pathway activation by continuous wave visible light laser micro-irradiation. *AIMS Genet* *4*, 47-68.
- Paul, S., Kuo, A., Schalch, T., Vogel, H., Joshua-Tor, L., McCombie, W.R., Gozani, O., Hammell, M., and Mills, A.A. (2013). Chd5 requires PHD-mediated histone 3 binding for tumor suppression. *Cell Rep* *3*, 92-102.
- Pilié, P.G., Tang, C., Mills, G.B., and Yap, T.A. (2019). State-of-the-art strategies for targeting the DNA damage response in cancer. *Nature Reviews Clinical Oncology* *16*, 81-104.
- Price, B.D., and D'Andrea, A.D. (2013). Chromatin remodeling at DNA double-strand breaks. *Cell* *152*, 1344-1354.

- Quan, J., Adelmant, G., Marto, J.A., Look, A.T., and Yusufzai, T. (2014). The chromatin remodeling factor CHD5 is a transcriptional repressor of WEE1. *PLoS One* 9, e108066.
- Quan, J., and Yusufzai, T. (2014). The tumor suppressor chromodomain helicase DNA-binding protein 5 (CHD5) remodels nucleosomes by unwrapping. *J Biol Chem* 289, 20717-20726.
- Shanbhag, N.M., and Greenberg, R.A. (2013). The dynamics of DNA damage repair and transcription. *Methods Mol Biol* 1042, 227-235.
- Song, C., Zhang, S., and Huang, H. (2015). Choosing a suitable method for the identification of replication origins in microbial genomes. *Front Microbiol* 6, 1049.
- Tyagi, M., Imam, N., Verma, K., and Patel, A.K. (2016). Chromatin remodelers: We are the drivers!! *Nucleus* 7, 388-404.
- Wolffe, A.P. (2001). Chromatin remodeling: why it is important in cancer. *Oncogene* 20, 2988-2990.
- Wu, X., Zhu, Z., Li, W., Fu, X., Su, D., Fu, L., Zhang, Z., Luo, A., Sun, X., Fu, L., *et al.* (2012). Chromodomain helicase DNA binding protein 5 plays a tumor suppressor role in human breast cancer. *Breast Cancer Res* 14, R73.