Designing Chromatin Substrates for Structural and Functional Analysis of Regulatory Complexes

Ryan Scott Stoner

Department of Biochemistry

University of Colorado Boulder

Defense Date

4 April, 2022

Thesis Advisor

Dr. Vignesh Kasinath, Department of Biochemistry

Honors Council Representative

Dr. Jeffrey Cameron, Department of Biochemistry

Defense Committee

Dr. Michael Stowell, Department of Molecular, Cellular & Developmental Biology

Abstract

The processes behind cellular differentiation and the maintenance of cellular identity are both thought to be regulated by chromatin modification. Furthering our understanding of these mechanisms could offer invaluable insights into both cancer and development. This study will focus on a specific group of proteins called polycomb repressive complexes, which help to control gene expression through the modifications they exert on chromatin. Polycomb group proteins consist of three families: PRC1 which deposits ubiquitin on H2AK119, PRC2 which mono-, di-, and tri-methylates H3K27, and PR-DUB which acts as a proofreading mechanism, removing misplaced ubiquitin on H2A. The exact mechanism by which PRC1, an E3 ubiquitin ligase, is recruited to and activated on chromatin is poorly understood. To better understand how PRC1 acts on chromatin, chromatin substrates were designed and assembled for experimental analysis. All four histones making up the histone octamer, as well as modified histones capable of crosslinking with ubiquitin, were cloned, expressed, and purified using a wide variety of biochemical techniques. This was used to form the nucleosome core particle which can undergo functional and structural assays in conjunction with PRC1-RYBP. This PRC1-RYBP was found to be active, capable of binding to CpG rich nucleosome sequences and modifying chromatin by ubiquitinating H2A. Preliminary tests that explored the interaction of PRC1-RYBP with the nucleosome found that PRC1-RYBP had a higher affinity for the nucleosome modified with ubiquitin on H2A than for the wild-type nucleosome. This affinity is higher than previously thought, binding much more tightly than other E3 ligases. More work involving PRC1, including additional activity and binding assays along with negative stain and cryo-EM structural analysis will continue to be carried out in the future, providing a deeper look into the critical role that PRCs play in controlling gene expression.

Acknowledgments

First and foremost, I would like to thank my thesis advisor and mentor, Vignesh Kasinath. Working in the Kasinath lab has proven to be an invaluable experience through which I have learned what working in a research lab entails. I have become a better thinker, reasoner, and researcher and I feel as though I will enter my future career prepared for what lies ahead.

I am especially grateful for my fellow undergraduates Tyler Hobbs, Angel Zhang, and Maggie Cornelius who has been by my side throughout this project. Their support, both on my project as well as outside of the lab has been invaluable in successfully completing the project.

Other members of the Kasinath lab including Akhil Gargey, Danielle Guillen, Mark Matyas, and Liqi Yao have also helped me as mentors and people I feel comfortable approaching should I have questions or need guidance. I especially would like to thank Akhil for assisting me by purifying PRC1 and working with me to design experiments.

I would also like to thank the Department of Biochemistry for providing me with intellectually challenging courses and wonderful professors who have further driven my passion for science. Specifically, I would like to thank Jeffery Cameron, my honor's council representative for taking the time to allow me to undertake an undergraduate thesis.

I also received support on this project from the Biological Sciences Initiative (BSI). The funding and professional development programs provided by BSI have helped enable me to carry out my research and complete an undergraduate thesis.

Table of Contents

Introduction

One of the fundamental questions in biology is why cells, which largely have near-identical DNA, can differentiate into such a large number of cell identities. Discovering the processes by which this occurs could have tremendous implications in our understanding of development as well as the onset of cancer. One way in which this cellular differentiation is thought to occur is through post-translational modifications (PTMs) of histones carried out by chromatin modifiers (Chen & Dent, 2013). Histones play an important role in the cell, serving as the protein units which make up the core nucleosome particle. The structure and sequence of histones are among the most conserved of all eukaryotic proteins, highlighting their importance in normal cell function (Marino-Ramirez et al., 2011). The nucleosome is the fundamental unit of chromatin, allowing for the compaction of DNA in the nucleus (Xu & Liu, 2021). Histones assemble into an octamer made up of two copies each of H2A, H2B, H3, and H4, around which DNA wraps, forming a left-handed supercoil (Luger et al., 1997). Protruding from this core particle are histone tails, disordered polypeptide chains stemming from the N-terminus of all histones, and the C-terminus of H2A (Iwasaki et al., 2013). These histone tails are capable of undergoing PTMs which have various impacts on transcriptional activation and silencing, as well as the assembly of chromatin (Marino-Ramirez et al., 2011). DNA compaction through histone PTMs not only allows DNA to be packed into higher-order chromatin structures but also impacts the ability of DNA-binding proteins to interact and initiate transcription (Michael & Thomä, 2021). Heterochromatin is characterized by an especially condensed region of chromatin in which gene expression is repressed. This formation of heterochromatin, carried out by chromatin

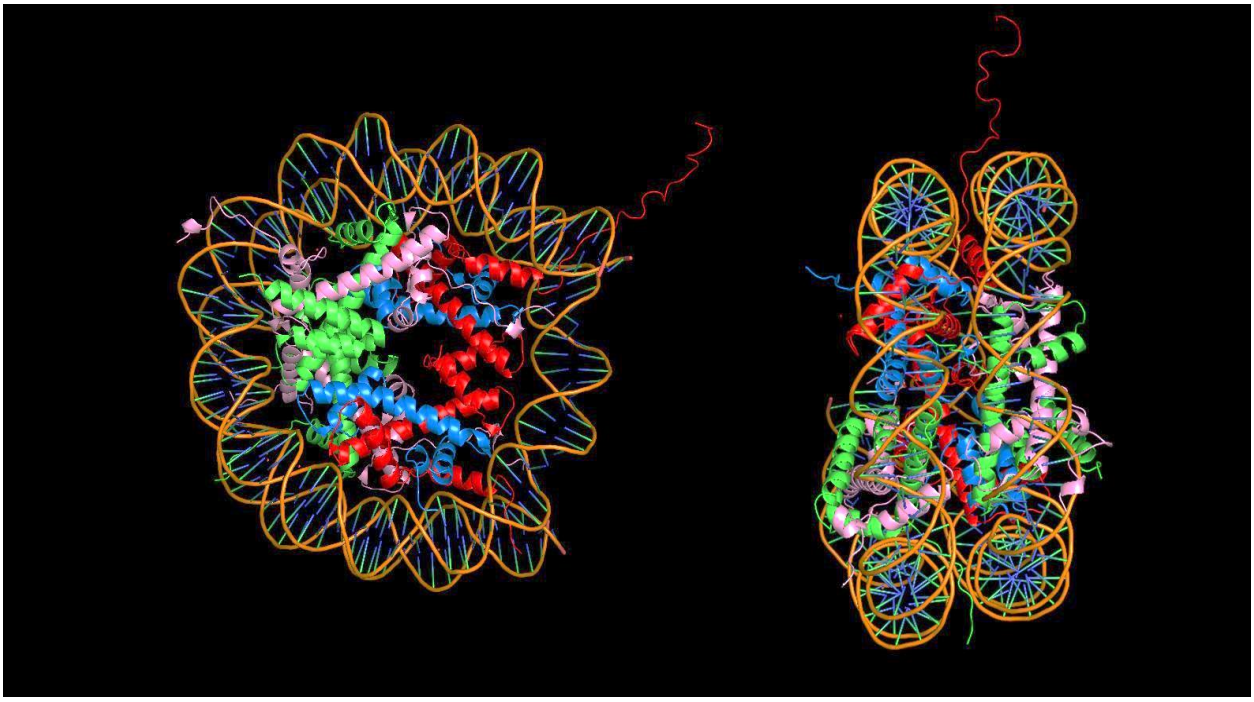

Figure 1: Nucleosome crystal structure looking down nucleosome superhelical axis (left), and perpendicular to the superhelical axis (right). Contains all four histone proteins (H3 = red, H4 = blue, H2A = Pink, H2B = Green), as well as a 146bp DNA segment wrapped around the histone octamer. Also notable are the histone tails protruding from the nucleosome core particle. Structure found on protein database, accession id 1AOI (Luger et al., 1997).

modifiers, is critical in the regulation of gene expression. Chromatin modifiers relevant to my research include polycomb repressive complexes (PRCs) as well as histone deacetylases (HDACs). PRCs are a group of proteins that remodel chromatin by covalently modifying histone tails, promoting gene silencing by heterochromatin formation (Luo et al., 2009). In this mode of chromatin modification, non-canonical PRC1, an E3 ubiquitin ligase, (ncPRC1) is first recruited to CpG islands, regions of the genome rich in cytosine and guanine (Lavarone et al., 2019). At these sites, ncPRC1 catalyzes the monoubiquitination of H2AK119 (H2AK119ub1). PRC2, a methyltransferase, then recognizes this mark and catalyzes the trimethylation of H3K27 (Lavarone et al., 2019). Lastly, canonical PRC1 (cPRC1) contains CBX proteins that recognize and bind to H3K27me3 and potentially H2AK119ub1, leading to heterochromatin formation and

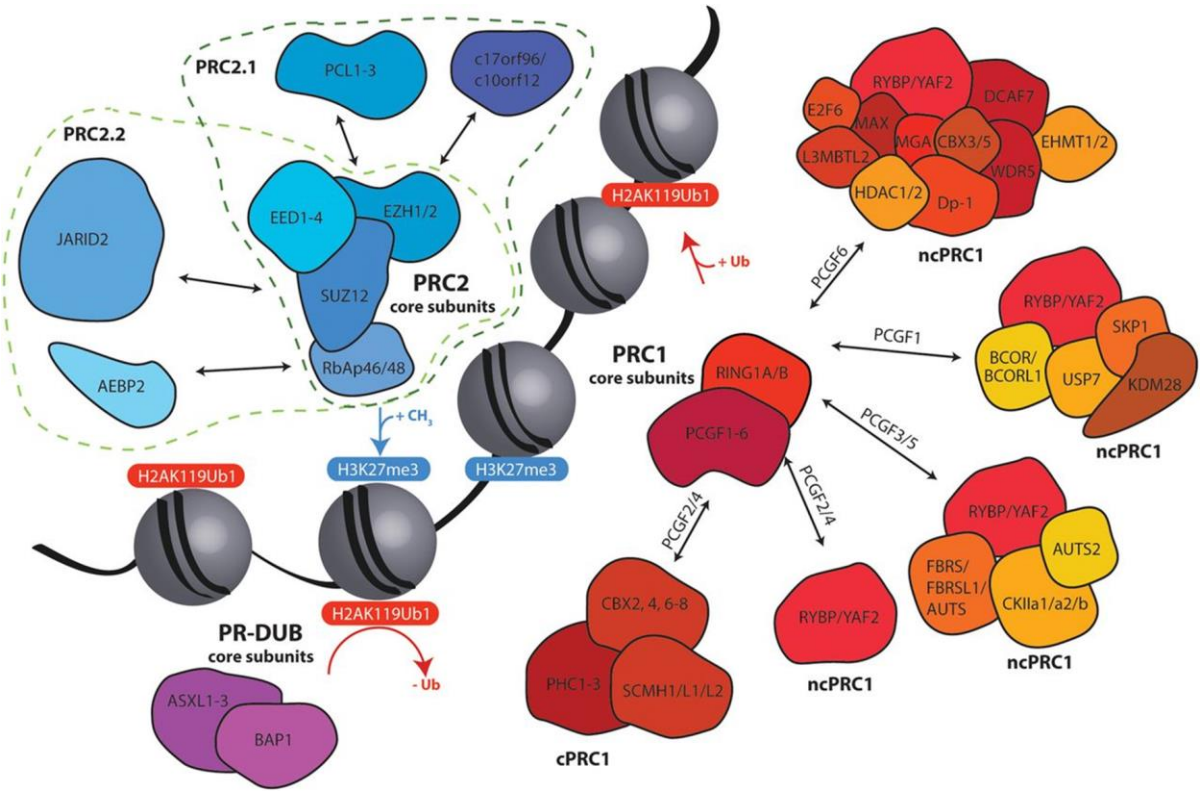

Figure 2: The complete family of polycomb group proteins, with the constituent proteins of each polycomb subtype listed. Both PRC1 and PRC2 are made up of a protein core present in all subtypes, and other subunits which make up the functional PRC1/2 complexes. PRC1 catalyzes the monoubiquitination of H2AK119 while PRC2 catalyzes the trimethylation of H3K27. PR-DUB is capable of removing the ubiquitin on H2AK119 (Chittock et al., 2017).

gene silencing (Lavarone et al., 2019). More specifically, this produces facultative heterochromatin, which is reversible and enables the regulation of gene expression through histone modification (Trojer and Reinberg, 2007). This also contributes to gene silencing during development and differentiation (Chantalat et al., 2011). Another group of chromatin modifiers called HDACs are involved in the deacetylation of H3 and H4 lysines, which is generally thought to be the first event leading to transcription repression (Seto & Yoshida, 2014). Besides histone-modifying enzymes, DNA-modifying enzymes such as DNMT1 also play a crucial role in gene regulation (Ren et al., 2020). DMNT1 is recruited to hemi-methylated DNA wrapped

around histones ubiquitinated at H3K18 and H3K23 (Vann & Kutateladze, 2017). Overall, all of these chromatin modifiers are involved in gene regulation and are the focus of much research.

The precise spatial and temporal regulation of gene transcription is critical not only for cell differentiation but also for establishing gene expression patterns and cell signaling. Mutations in PRCs or HDACs or DNMT1 are leading causes of cancer and neurodegenerative diseases as they result in a fundamental disruption of gene transcription (Glozak & Seto, 2007; Piunti & Shilatifard, 2021). Chromatin modifiers are among the most mutated genes in all human tumors, showcasing their importance in cancer development (Parreno et al., 2022). Mutations in Polycomb-group proteins or in the genes which regulate their activity have been shown to cause transcriptional deregulation, leading to tumor formation (Parreno et al., 2022). Because of PRC's importance in cellular differentiation, mutations in Polycomb-group proteins can lead to tumor suppressor inhibition and activation of proto-oncogenes (Wang et al., 2015). This could cause a loss in cellular identity and enable these cells to evade the various mechanisms by which the body prevents uncontrolled proliferation (Wang et al., 2015). Understanding the mechanism of how these proteins function is hence important for formulating appropriate therapeutic approaches for various cancers where these proteins are often deregulated. Furthermore, PRCs also play an important role in development. Both PRC1 and PRC2 are involved in transcriptionally repressing genes involved in development (Piunti & Shilatifard, 2021). As a result, over or underexpression of Polycomb-group proteins is commonly associated with deregulation of developmental genes, leading to developmental abnormalities (Piunti & Shilatifard, 2021). Loss of function studies of PRC1 and PRC2 subunits also highlights the critical role that Polycomb-group proteins play in development. In PRC2, all core subunits, as well as JARID2, are required for development, and in PRC1, both canonical and non-

canonical subunits have essential functions in development (Piunti & Shilatifard, 2021). Furthermore, PRCs play a critical role in ensuring proper dosage compensation. PRC2 is recruited to Xist RNA on the inactive chromosome in females, which then recruits PRC1, leading to H3K27me3 and H2AK119ub1 deposition (Brockdorff, 2017). This transcriptionally silences the inactive X chromosome, allowing proper levels of transcription (Brockdorff, 2017).

I will be building my research project around the current hypothesis that both DNA sequence and histone post-translational modifications play a role in the recruitment of chromatin modifiers to their appropriate site. In my research project, I will be working on reconstituting complex chromatin substrates which will be used in a variety of structural and activity studies to test this. I will be generating modified histones as well as the nucleosome DNA to ultimately reconstitute the nucleosome. This will enable me to then study how chromatin modifiers such as PRC1 interact with the nucleosome, hopefully elucidating the role they play in gene expression.

Materials and Methods

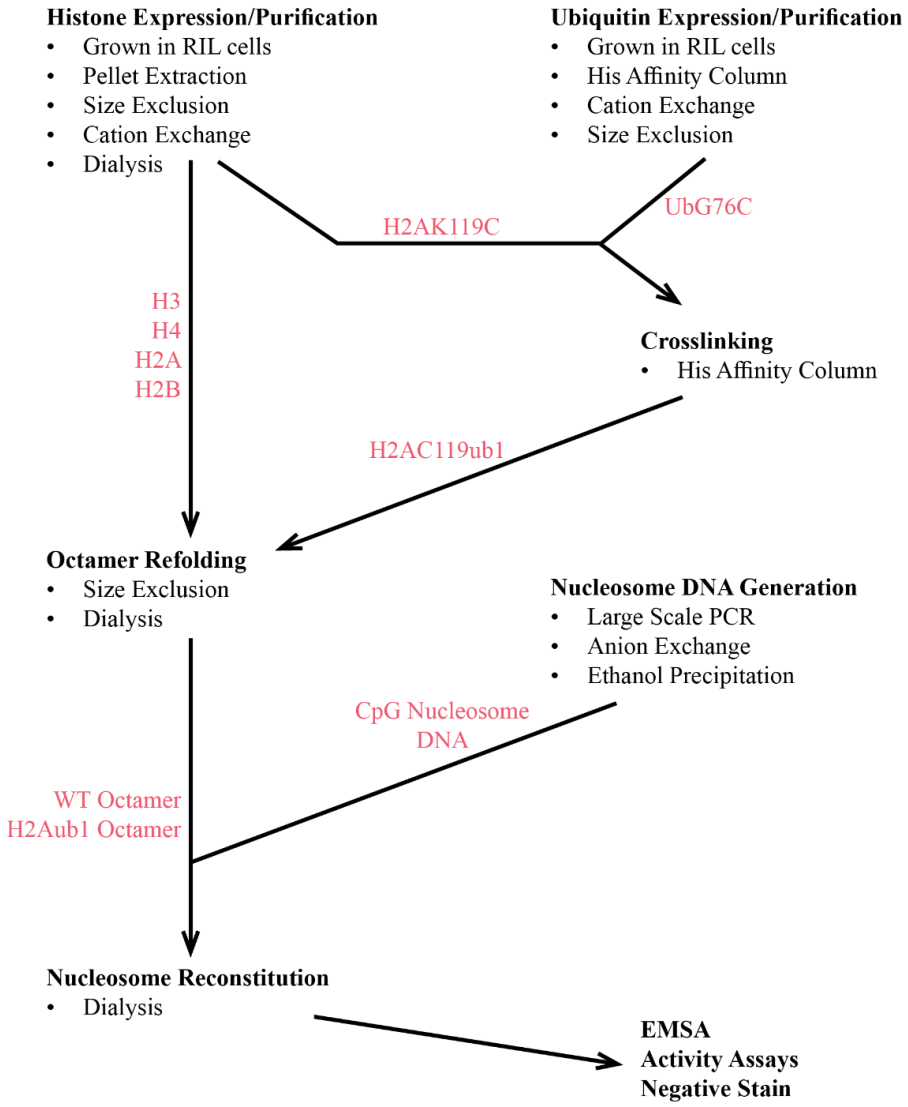

Figure 3: Flowchart describing the steps used throughout the project. Products generated in each step are highlighted in red on the arrows. First, all individual histones (H3, H4, H2A, H2B, H2AK119C, H2AK119C) and UbG76C were expressed and purified. H2AC119 and UbG76C were crosslinked to create H2AC119ub1 which, along with the other histones, was refolded into both a wild-type and H2Aub1 octamer. These octamers were reconstituted with nucleosome CpG DNA to form the nucleosome, which was used in activity assays, and EMSA binding assays with PRC1.

Modified Histone Sequence Generation

H3K18CK23C and H2AK119C were generated using site-directed mutagenesis of modified Xenopus histone sequences (Final DNA sequences used found in supplemental materials). All cysteine residues in the original sequence were first altered to alanine in order to prevent unwanted crosslinking. Following this, primers were designed to alter the H2AK119 to a cysteine, taking care to alter the fewest bases. This process was repeated for H3K18C then H3K23C. This mutagenesis was done using CloneAmp Hifi DNA polymerase under the following conditions: 10 ng template, 0.5 µM forward and reverse primers, 12.5 µL 2X CloneAmp HiFi PCR Premix, in a total reaction volume of 25 µL. Thermocycler conditions started with a 3-minute initial denaturation at 98°C, 10-second 98°C denaturation, 15-second 52.6°C annealing, 35-second 72°C extension, and 5-minute 72°C final extension with 30 cycles. Following each PCR-based mutation, the plasmid (pET3a) was transformed into Agilent XL1-Blue competent cells, then plated onto LB + agar plates with 0.1 µg/mL ampicillin. For transformation, 50 µL of Agilent XL1-Blue competent cells are thawed on ice, 100 ng of the plasmid is added, then allowed to sit on ice for 10 minutes. The cells are then heat shocked at 42°C for 45 seconds then placed back on ice for 5 minutes. 750 µL of LB is added to the cells and they are shaken at 600 rpm, 42°C for 1 hour, plating 200 µL onto the appropriate plate and allowing it to grow overnight at 37°C. An overnight culture was started using a colony from the plate in 5 mL of LB in a culture tube and the plasmid was isolated using a Macherey-Nagel miniprep kit.

Histone Expression

Individual histone plasmids were transformed into *Escherichia coli* (*E. coli*) Agilent BL21-CodonPlus RIL cells and plated onto LB + agar plates with 100 µg/mL ampicillin and 10 µg/mL

chloramphenicol. 50 mL overnight cultures were started in 2XYT media using 0.1% glucose and 100 µg/mL ampicillin and 10 µg/mL chloramphenicol. At an OD between 0.4 and 0.5 cultures were transferred into a large shaker flask containing 1.5 L of 2XYT, 0.1% glucose, and 100 µg/mL ampicillin and 10 µg/mL chloramphenicol. Cultures were grown to an OD of 0.4 and then induced with 0.2 mM IPTG for 3 hours. Cells were then spun down at 4000 rpm for 30 minutes at 4°C. The pellet was resuspended in histone wash buffer (20 mM Tris, pH 7.5, 200 mM NaCl, 1 mM EDTA, 5 mM βME) then frozen in liquid N2 for storage at -80°C.

Histone Purification

Cell pellets were thawed in a water bath, then refrozen in liquid N2. This process was repeated an additional two times to facilitate cell lysis. Following two freeze-thaw cycles, a Roche complete-protease inhibitor cocktail tablet, DNase, and 2.5 µL benzonase were added to the resuspended pellet. The cells were sonicated on ice for 1 minute at 100% power, 10 seconds on, 20 seconds off. The cells were then lysed using a C3 emulsiflux high-pressure homogenizer. The cell lysate was then spun down at 15,000 rpm for 20 minutes at 4°C. Individual over-expressed histones do not fold without their respective partners and are hence present in inclusion bodies i.e. the pellet fraction after lysis. Histones were extracted by resuspending the pellet in 20 mL histone wash buffer containing 1% triton, 2.5 µL benzonase, and a PIC tablet, then spun down at 10,000 rpm for 10 minutes at 4°C. Two additional resuspensions and spin downs were done using histone wash buffer, this time with no triton. After the final spin down, 300-500 uL of DMSO was added to the pellet, which was then minced using a spatula and allowed to sit at room temperature for 30 minutes. 10 mL of extraction buffer (7 M GuHCl, 20 mM Tris pH 7.5, 200 mM NaCl, 1 mM EDTA, 5 mM βME) was then added to the pellet/DMSO mixture, stirring for 1 hour. This was then spun down at 15,000 rpm for 10 minutes at 4°C, with the supernatant

collected and filtered through a gravity column. The supernatant was then loaded onto a HiPrep 26/60 Sephacryl S-200 size-exclusion column and ran overnight using 7 M Urea, 20 mM Na-Acetate, pH 5.2, 1 M NaCl, 1 mM EDTA, 5 mM βME as the buffer. Fractions containing protein (as determined by running an SDS-PAGE gel) were pooled and diluted using a no salt buffer (7 M Urea, 20 mM Na-Acetate, pH 5.2, 1 mM EDTA, 5 mM βME) to a final salt concentration of 200-250 mM. This was then loaded onto a cation exchange column (PL-SCX 4000A 30 µM resin), and eluted using a high salt buffer (7 M Urea, 20 mM Na-Acetate, 1 M NaCl, pH 5.2, 1 mM EDTA, 5 mM βME). The flowthrough from the cation exchange column was loaded back on and ran a second time to improve yield. Fractions containing histone were pooled and dialyzed in 4 L water with 2 mM βME using 3kDa cutoff dialysis tubing. Following dialysis, histones were concentrated to 1 mg/mL and lyophilized. (Dyer et al., 2003)

Nucleosome DNA Generation

Large-scale PCR was done in order to generate large amounts of nucleosome DNA containing the 147 bp widom sequence with 40 bp flanking sequences (Full sequence in supplemental materials). This was done using GoTaq Flexi DNA polymerase in 50 µL reactions on 96 well plates under the following conditions: 50 ng Template (pTwist 147 bp DNA), 0.5 µM Forward and Reverse primers, 10 µL 5X colorless GoTaq Flexi buffer, 2 mM MgCl2, 0.2 mM dNTPs, 3 µL DMSO, and 0.25 µL GoTaq Flexi enzyme in a 50 µL reaction. Thermocycler conditions started with a 3-minute initial denaturation at 95°C, 30-second 95°C denaturation, 30-second 58°C annealing, 15-second 72°C extension, and 2-minute 72°C final extension with 34 cycles. All PCR reactions were pooled and kept at 4°C for further purification. Pooled PCR reactions were diluted 1.5x in no salt buffer (20 mM Tris), then loaded onto a HiTrap Q 1 mL column. The DNA was eluted using high salt (2 M KCl, 20 mM Tris), and low salt (150 mM KCl, 20 mM

Tris) buffers, coming out in a single peak. 100 µL of 3 M sodium acetate was added to each 1 mL peak fraction, which was then split into three separate tubes. To each of these tubes, 1 mL of 100% ice-cold ethanol was added and the tubes were stored at -80°C for 3 hours or -20°C overnight. Each tube was then spun down at 11,000 rpm for 30 minutes at 4°C, and all liquid was decanted. DNA pellet was resuspended in 500 µL of ice-cold 70% ethanol and then spun down at 11,000 rpm for 15 minutes at 4°C. All liquid was removed, and the pellet was left to dry in a fume hood for 1 hour. All pellets were resuspended and combined into a single tube using 300 µL RB High buffer (10 mM Tris-HCl, pH 7.5, 2 M KCl, 1 mM EDTA, 1 mM DTT), where the DNA concentration was measured.

Ubiquitin Purification

The ubiquitin plasmid was transformed into *E. coli* Agilent BL21-CodonPlus RIL cells and plated onto LB + agar plates with 100 µg/mL ampicillin and 10 µg/mL chloramphenicol. 50 mL cultures were started in 2XYT media using 0.1% glucose and 100 µg/mL ampicillin and 10 µg/mL chloramphenicol. At an OD between 0.4 and 0.5 cultures were transferred into a large shaker flask containing 1.5 L of 2XYT, 0.1% glucose, and 100 µg/mL ampicillin and 10 µg/mL chloramphenicol. Cultures were grown to an OD of 0.4 and then induced with 0.2 mM IPTG for 3 hours. Cells were then spun down at 4000 rpm for 30 minutes at 4°C, and resuspended (150 mM Nacl, 25 mM HEPES, 10% glycerol). This was then spun down at 4900 rpm, discarding the supernatant, flash freezing the pellet in liquid N_2 for storage at -80°C. The pellet was resuspended using 25 mL buffer (0.5% CHAPS, 25 mM HEPES, 0.2 mM PMSF, 10 µM Leu peptin, 1 M NaCl, PIC Tablet, 1mM TCEP, 10 mM imidazole, DNase). The resuspension was then lysed using a C3 emulsiflux high-pressure homogenizer. The lysate was spun down at 15,000 rpm for 20 minutes at 4°C. The supernatant was filtered on a gravity column, then loaded

onto a 5 mL HisTrap HP column, eluting with a gradient from low imidazole (20 mM imidazole, 1 M NaCl, 25 mM HEPES, 0.2 mM PMSF, 10 µM leu peptin, 1mM TCEP) to high imidazole (500 mM imidazole, 1 M NaCl, 25 mM HEPES, 0.2 mM PMSF, 10 µM leu peptin, 1mM TCEP). Fractions containing ubiquitin were then diluted to 150 mM NaCl and loaded onto a cation exchange (PL-SCX 4000A 30 µM resin) column and eluted using a low salt (150mM NaCl, 25 mM HEPES) to high salt (2M NaCl, 25 mM HEPES) gradient. Lastly, fractions containing ubiquitin were loaded onto a size exclusion column (Superdex 200 Increase 10/300) for a final purification step (1M NaCl, 1 mM TCEP, 25 mM HEPES).

Histone Crosslinking

Lyophilized aliquots of H2AK119C and UbG76C were dissolved in 500 µL of urea buffer (50 mM Tris, 1 mM EDTA, 7 M Urea, pH 7.5) and sat at room temperature for 30 minutes. Protein concentration was measured at 280 nm, allowing H2AK119C and UbG76C to be combined at a 2:1 molar ratio. Borax buffer (50 mM Sodium Tetraborate, 5 mM TCEP, 6 M Urea, pH 8.5) was added to reach a final H2AK119C concentration of 100 µM. 1 M TCEP was added to reach a final concentration of 5 mM, then the mixture was allowed to sit at room temperature for 30 minutes. A 0.1 M solution of 1,3-dichloroacetone (DCA) dissolved in N,N′-dimethylformamide (DMF) was added to the histone-Ub mixture to reach a final DCA concentration twice the histone concentration. This was allowed to incubate on ice for one hour, after which the reaction was stopped using 5 mM βME. To purify the desired crosslinking products, a Ni-NTA HisTrap HP column was run to collect samples containing a 6x-His-tag on ubiquitin. The crosslink reaction was diluted 10x in denaturing binding buffer (50 mM NaPi, pH 8, 300 mM NaCl, 6 M urea, 10 mM imidazole, 5 mM βME), then incubated with Ni-NTA agarose at 4°C for 1 hour. The column was washed with binding buffer 3 times and then eluted using elution buffer (50

mM NaPi, pH 8, 300 mM NaCl, 6 M urea, 200 mM imidazole, 5 mM βME, 200 mM imidazole). The eluted fractions were then dialyzed into 4 L water and lyophilized. (Long et al., 2014)

Octamer Refolding

Equimolar amounts of histones were dissolved in unfolding buffer (6 M GuHCl, 20 mM Tris-HCl, pH 7.5, 5 mM DTT) and allowed to sit at room temperature for 1 hour. Each individual protein concentration was calculated and histones were mixed in equimolar amounts. Samples were then dialyzed in 10 kDa tubing against 2 L refolding buffer (10 mM Tris-HCl, pH 7.5, 2 M NaCl, 1 mM EDTA, 5 mM βME). Dialyzed samples were pooled and concentrated down to 500 µL, then loaded onto a Superdex 200 10/300 size exclusion column, and ran with refolding buffer. Fractions containing octamers are concentrated and stored at 4°C on ice (not frozen). (Dyer et al., 2003)

Nucleosome Reconstitution

Octamers and nucleosome DNA are combined together at a 1.1:1 octamer:DNA ratio, using 4.4 µM octamer and 4 µM DNA. DNA is added first, then octamer is added to the solution. This mixture is then dialyzed against a gradient of RB low (10 mM Tris-HCl, pH 7.5, 250 mM KCl, 1 mM EDTA, 1 mM DTT) into RB high (10 mM Tris-HCl, pH 7.5, 2 M KCl, 1 mM EDTA, 1 mM DTT), starting with 400 mL of RB-high buffer, and slowly pumping in 1600 mL of RB-low buffer while simultaneously pumping out the used dialysis buffer at the same rate. A final set of dialysis was done in which the nucleosome was dialyzed against 500 mL RB low then 500 mL TCS (20 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mM DTT). Nucleosomes are then stored at 4°C on ice (not frozen). (Dyer et al., 2003)

Activity Assay

E2-Ub and nucleosome were combined to a final concentration of 5 µM E2-Ub and 0.5 µM nucleosome. This was then allowed to incubate at room temperature for 15 minutes, after which 2 µM PRC1-RYPP was added. The final reaction volume was brought to 30 µL with buffer (150 mM NaCl, 25 mM HEPES). 10 µL was immediately taken and quenched with 4 µL of 4x Laemmli sample buffer containing no βME. 10 µL of the reaction was also taken at 5-minute and 45-minute time points, quenching with 4 µL of 4x Laemmli sample buffer containing no βME. Control samples were made including 5 µM E2, 5 µM E2-Ub, 2 µM PRC1-RYBP, 0.5 µM nucleosome, and 2 µM PRC1-RYBP + 5 µM E2 + 0.5 µM nucleosome. An SDS-PAGE gel was run at 200 V for 60 minutes, loading 8 µL of each sample to detect PRC1-RYBP activity.

EMSA

5% 59:1 Native gels were cast and preran in 0.2x TBE for 1 hour at 150 V. The entire gel apparatus was placed in an ice slushy in the cold room. 100 nM of wild-type nucleosome was mixed with different molar ratios of PRC1 (1:0, 1:1, 1:2.5, 1:5, 1:7.5, 1:10 nucleosome:PRC1), at a salt concentration of 100 mM NaCl in a reaction volume of 10 µL. The same reactions were made up using 100 nM of H2AK119ub1 nucleosome before loading on the gel. The 0.2x TBE from the gel prerun was replaced with fresh 0.2x TBE buffer before loading any samples. Each 10 µL sample was mixed with 2 µL of 20% sucrose, and 8 µL was loaded onto the gel, using one or two Bio-Rad EZ Load 100 bp ladders to differentiate the gels. The gels were run at 150 V for 135 minutes. The gel was stained in 30 mL of 0.2x TBE with 1x SYBR Gold for 10 minutes, then imaged on a GE Typhoon FLA 9000 Gel Imaging Scanner. (Rose et al., 2016)

Results & Discussion

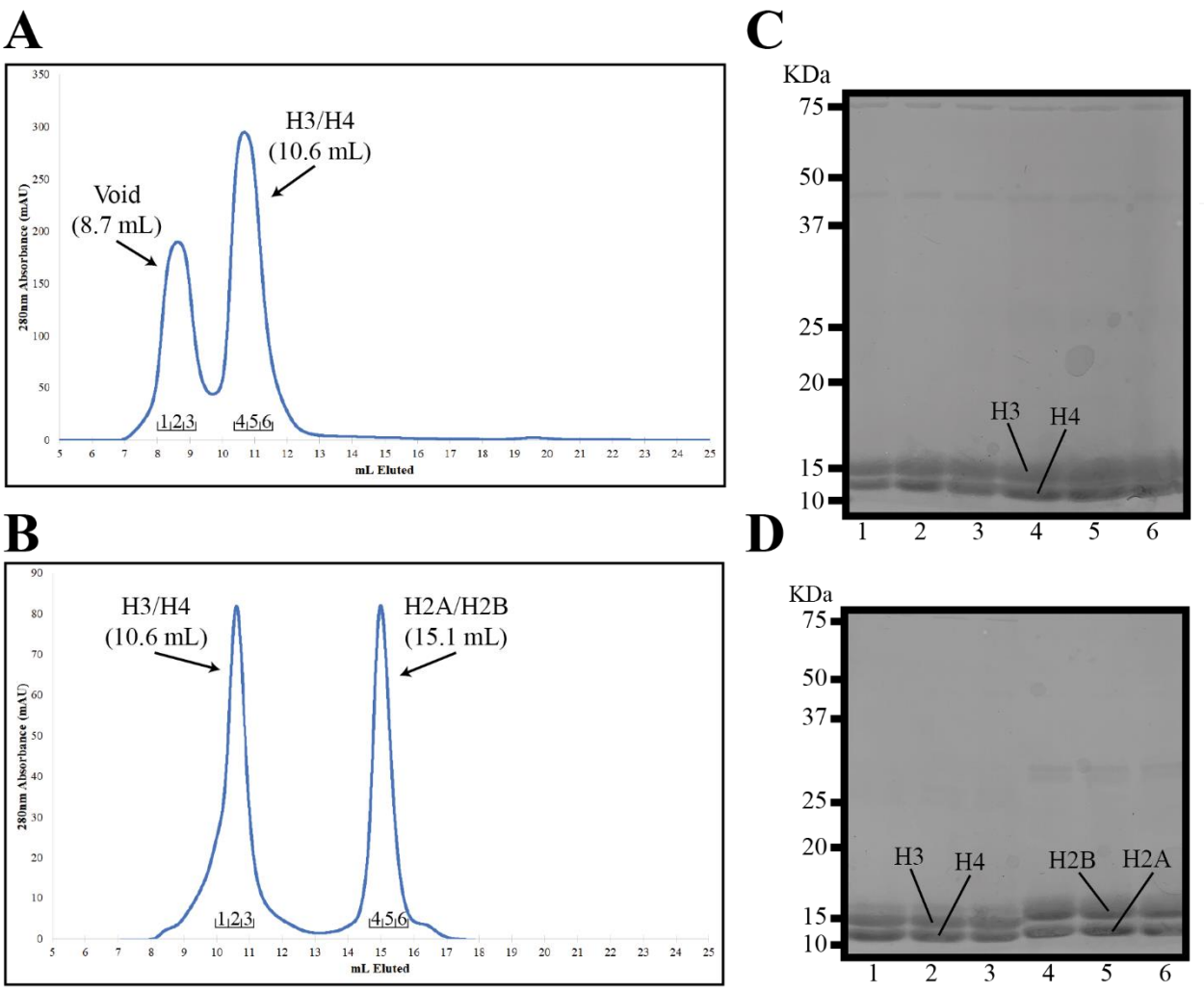

Figure 4: (A) FPLC size-exclusion (Superdex 200 Increase 10/300) trace of H3/H4, eluting from the column at a higher molecular weight than expected. **(B)** FPLC size-exclusion (Superdex 200 Increase 10/300) trace shows failed reconstitution of octamer. Peaks eluted at 10.6 mL and 15.1 mL, both having a peak absorbance of 80 mAU. **(C)** SDS-PAGE gel ran using H3/H4 size-exclusion fractions, with lanes 1-3 corresponding to the first peak (void) and lanes 4-6 corresponding to the second peak (H3/H4). **(D)** SDS-PAGE gel ran on octamer fractions. Lanes 1-3 correspond to the first peak (H3/H4) and lanes 4-6 correspond to the second peak (H2A/H2B). Fractions ran on the gels are outlined on the bottom of the trace.

The initial purification scheme for reconstituting the octamer involved growing H2A/H2B and H3/H4 in single plasmids in order to cut down on the number of required histone purifications. This ran into problems, however, when the purified H3/H4 appeared to form

higher-order structures that failed to combine with H2A/H2B. Instead of forming an H3/H4 tetramer with a mass of approximately 52.8 KDa, higher molecular weight complexes were being formed as seen in the FPLC size-exclusion traces. In **Figure 4A**, there are two peaks, one eluting at 8.7 mL corresponding to the void volume, and one eluting at 10.6 mL. These elution volumes correspond to molecular weights of 800 KDa and 440 KDa respectively on the Superdex 200 Increase 10/300, both far above the expected value (Cytiva, 2020). The SDS-PAGE gel ran on these peaks in **Figure 4C** showed that the eluted protein was indeed H3/H4, but denaturing conditions were required for the protein to run at the proper molecular weight. An octamer reconstitution was attempted using the H3/H4 fraction from the second peak, along with previously purified H2A/H2B, but the trace in **Figure 4B** revealed that the H2A/H2B and H3/H4 did not properly combine, leading to separate H3/H4 tetramer and H2A/H2B dimer peaks. The first peak on the trace corresponds to the H3/H4 tetramer, eluting at approximately the same molecular weight as the H3/H4 used in the reconstitution. The H2A/H2B dimer peak, on the other hand, elutes at 15.1 mL, corresponding to a protein size of approximately 35 KDa on the Superdex 200 Increase 10/300 (Cytiva, 2020). This is exactly where the H2A/H2B dimer should elute, having a size of 28 KDa. Furthermore, because equal molar amounts of both H3/H4 and H2A/H2B were used in the octamer refolding, it would be expected for both peaks to have the same absorbance if no octamer was formed, which is exactly what is seen. The SDS-PAGE gel ran on the fractions in **Figure 4D** confirms that no octamer reconstitution occurred. This is believed to be the result of overgrowing the *E. coli* to an OD far exceeding the optimal range of 0.4-0.5 before induction. Ideally, induction should be done during the log phase of bacterial growth in which the *E. coli* are undergoing exponential growth (Dyer et al., 2003). Failure to do so could result in the cells reaching the stationary phase where the bacteria fight for depleting

resources, causing a lag in bacterial growth. Instead of expressing abundant, properly folded protein, the bacteria are instead expending their available energy to improve their chances of survival in a resource-scarce environment. This could be the cause of the higher than expected H3/H4 molecular weight as the protein could have improperly folded, leading to the formation of higher-order structures.

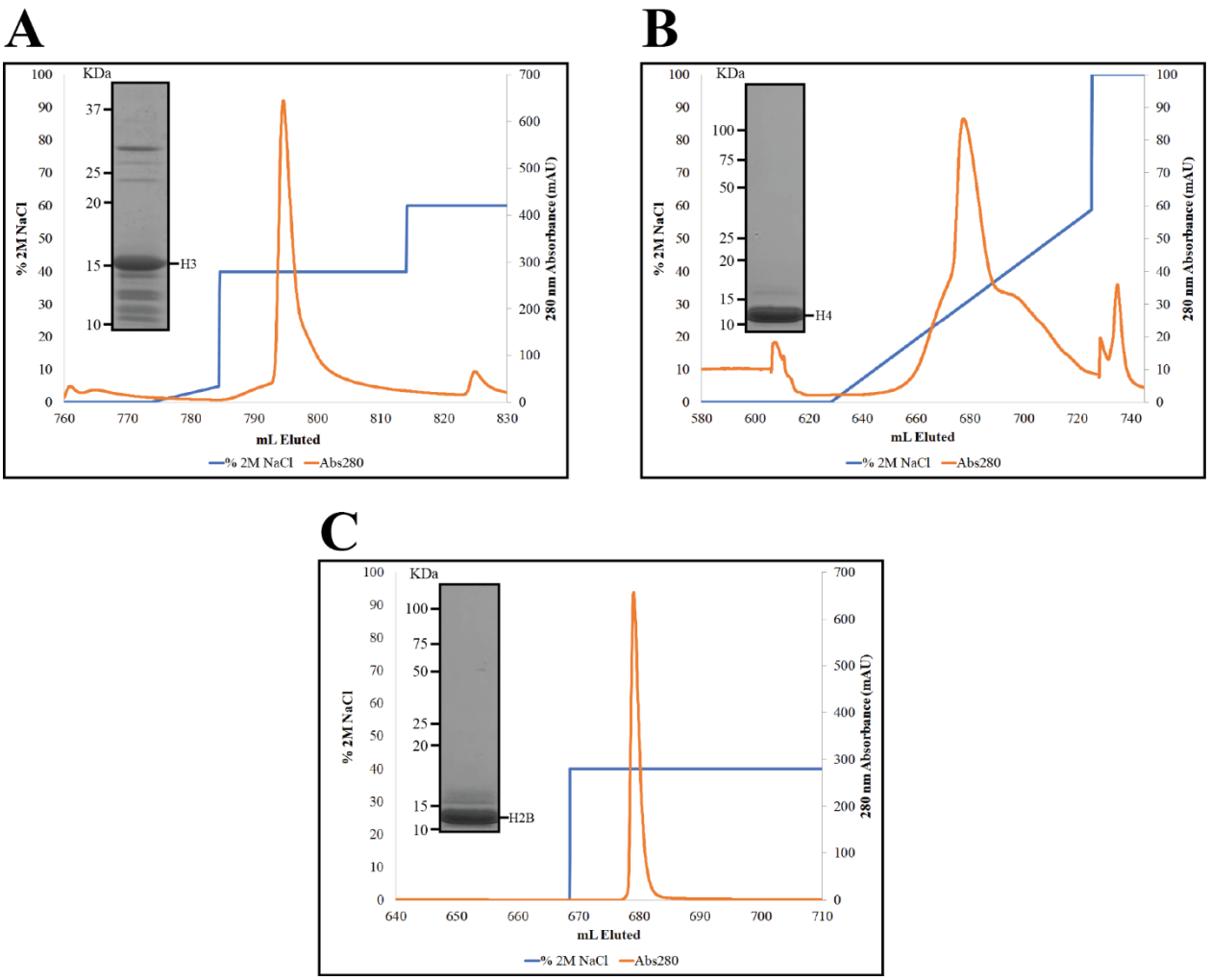

Figure 5: FPLC cation-exchange (PL-SCX 4000A 30 µM resin) traces of wild-type individual histones. **(A)** H3 eluted at 800 mM NaCl with a peak absorbance of 636 mAU. SDS-PAGE gel shows that the final product contains many impurities. **(B)** H4 eluted at approximately 600 mM NaCl with a peak absorbance of 86 mAU. SDS-PAGE gel shows relatively pure H4 sample. **(C)** H2B eluted at 800 mM NaCl with a peak absorbance of 655 mAU. SDS-PAGE gel confirms isolation of pure H2B sample.

To resolve these problems, a new purification scheme was developed in which each histone would be purified individually. Each individual histone was expressed in RIL cells and histones were extracted from exclusion bodies. The lysate was loaded onto a size exclusion column (HiPrep 26/60 Sephacryl S-200), with fractions containing protein then loaded onto a cation exchange resin (PL-SCX 4000A 30 µM resin). The individual histones were eluted using a 2M NaCl buffer, as the salt competes with the histones for binding to the resin, eluting the protein. **Figure 5B** shows that H2B eluted at a NaCl concentration of approximately 600 mM when using a linear salt gradient. Using this as a baseline, a step gradient of 800 mM NaCl was used to elute both H3 and H2B as seen in **Figures 5A & 5C**, yielding a sharp, singular histone peak. SDS-PAGE Gels ran on fractions from each cation exchange run revealed relatively pure fractions of both H4 and H2B, with only one minor impurity around 16 KDa. The H3 gel, however, revealed many impurities that neither size exclusion nor cation exchange could separate out. Although the sample was not pure, upon reconstitution of the octamer, only the H3 will combine with the other histone proteins, allowing separation of the much larger octamer from the other impurities.

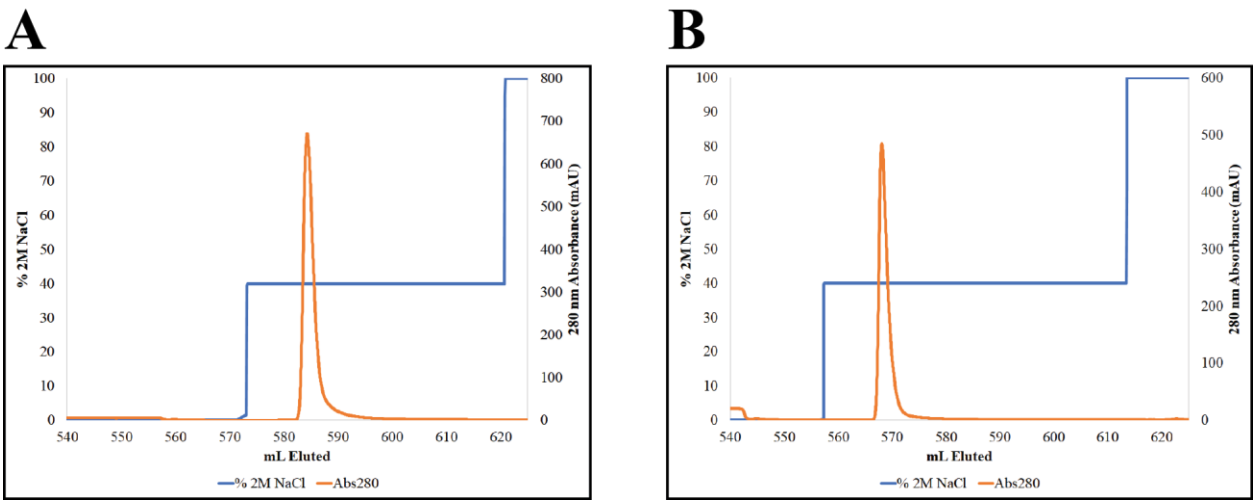

Figure 6: FPLC cation-exchange (PL-SCX 4000A 30 µM resin) traces of modified individual histones. **(A)** H2AK119C eluted at 800 mM NaCl with a peak absorbance of 666 mAU. **(B)** H3K18CK23C eluted at 800 mM NaCl with a peak absorbance of 482 mAU.

In order to generate ubiquitinated histones such as H2AC119ub1 and H3C18ub1C23ub1, lysine residues present in the histone tails must first be converted to cysteine residues. Doing this, along with the conversion of a ubiquitin glycine at position 76 to a cysteine residue allows for crosslinking of the two proteins using chemical methods. H2AK119C and H3K18CK23C were first generated using site-directed mutagenesis, then expressed in RIL cells, following the same procedure as done for the wild-type individual histones. **Figures 6A & 6B** show that each of the modified histones was eluted using 800 mM NaCl on a cation exchange column (PL-SCX 4000A 30 µM resin).

The wild-type octamer was generated by combining equimolar amounts of all four histones in an unfolding buffer (6 M GuHCl, 20 mM Tris-HCl, pH 7.5, 5 mM DTT) to completely unfold each protein before refolding the octamer into its native state by dialyzing in a refolding buffer (10 mM Tris-HCl, pH 7.5, 2 M NaCl, 1 mM EDTA, 5 mM βME). A size exclusion column

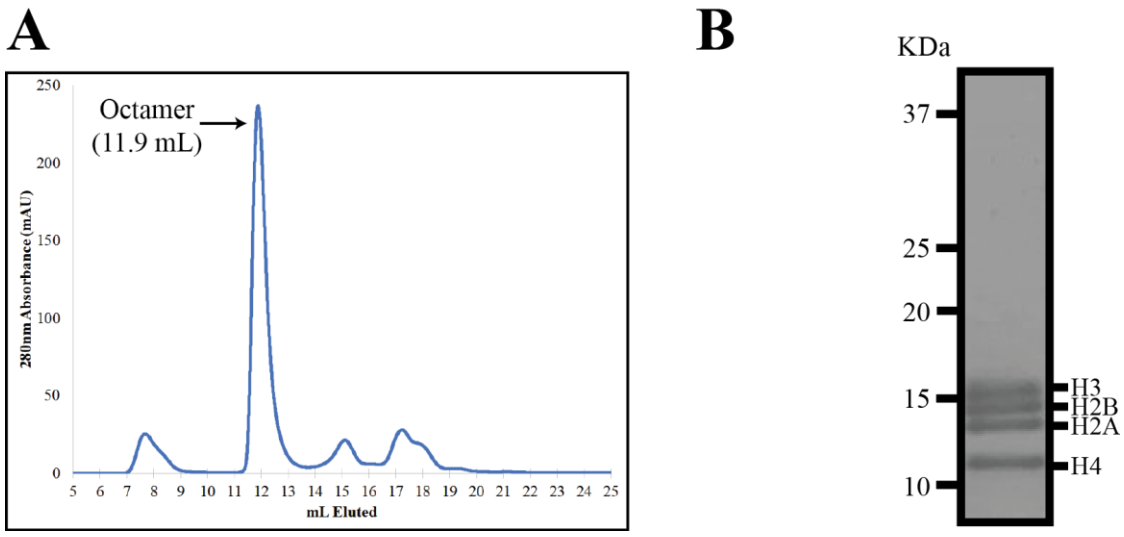

Figure 7: (A) FPLC size-exclusion (Superdex 200 Increase 10/300) trace of wild-type histone octamer, containing two copies each of H2A, H2B, H3, and H4. Peak eluted at 11.9 mL with a peak absorbance of 234 mAU. **(B)** SDS-PAGE gel ran on the peak fraction showing the presence of all four histones.

(Superdex 200 Increase 10/300) was then run to improve purity, with the trace seen in **Figure 7A**. The octamer peak eluted at 11.9 mL which corresponds to a molecular weight of approximately 158 KDa, which while higher than the expected molecular weight of 108 KDa, it is still within a reasonable range (Cytiva, 2020). An SDS-PAGE gel was run to confirm the presence of the octamer in the size-exclusion peak as seen in **Figure 7B**. All four histone bands can be clearly seen, and they are running at the appropriate height on the gel, confirming that the octamer was successfully isolated.

1,3-dichloroacetone (DCA) was used as a crosslinker to chemically conjugate UbG76C to H2AC119, generating H2AC119ub1. **Figure 8A** shows the difference between the H2A ubiquitination mechanism used natively in the cell and the chemical ubiquitination utilized in this project. Inside the cell, ubiquitin is first conjugated to E1, and the ubiquitin is then transferred to E2 (Streich & Lima, 2014). PRC1, acting as an E3 ligase, is capable of transferring the ubiquitin from E2 to the substrate, in this case, lysine 119 on H2A (Streich & Lima, 2014, Lavarone et al., 2019). The problem with this process for use in this project, however, is that the reaction requires a multitude of substrates, and it is hard to generate appreciable amounts of pure protein. Instead, a chemical conjugation using 1,3-dichloroacetate as a crosslinker was done to produce larger volumes of purer H2AC119ub1, a mimic of the in vivo cross-linking. Both the native and chemical conjugation bonds have approximately the same length, allowing for structural and functional assays to be run on the ubiquitinated nucleosome. **Figure 8B** shows an SDS-PAGE gel run following the chemical conjugation of H2AC119 to UbG76C. The brightest bands on the gel belong to H2A (MW = 13.96 KDa) and ubiquitin (MW = 8.6 KDa), indicating that a lot of the substrate did not react. This means that most of the product following the conjugation was likely unreacted H2A and ubiquitin. Despite this, there are three new bands that

appear following the zero minute conjugation time point which belong to the three cross-linked products. Because both H2A and ubiquitin contain cysteine residues, it is possible for H2A to crosslink with H2A, ubiquitin to crosslink with ubiquitin, and H2A to crosslink with ubiquitin. H2A*H2A would have a mass of 28 KDa, causing it to run at the top of the gel. H2AC119ub1, the desired product, would have a mass of 23.5 KDa and would run just below the H2A*H2A. Ub*Ub has a mass of just 17.2 KDa, running the lowest of the crosslinked products. These bands can all be observed in the gel and indicate that the crosslinking was successful. To separate the

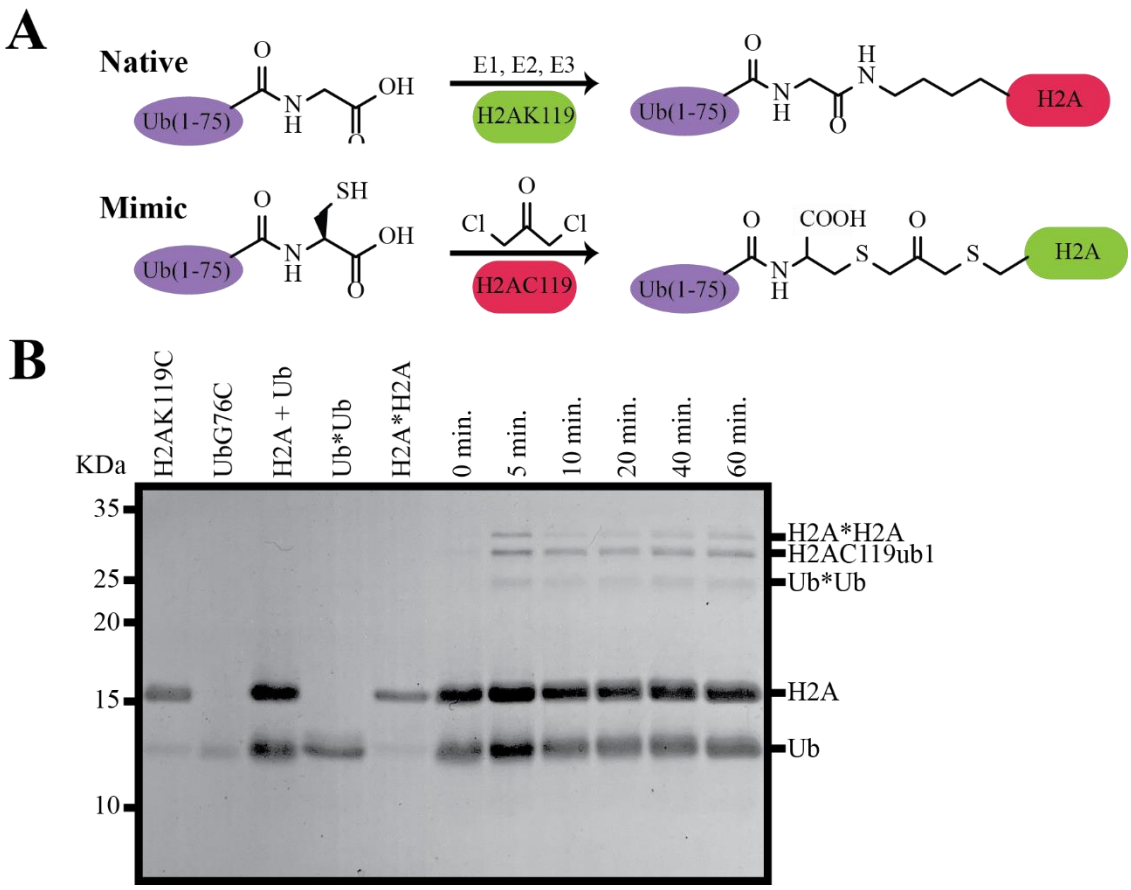

Figure 8: (A) Diagram showing the difference between the native and mimicked ubiquitination of H2A. In-vivo, cells utilize E1, E2, and E3 enzymes to transfer ubiquitin to H2AK119. The mimic utilizes 1,3-dichloroacetone as a crosslinker to attach UbG76C to H2AC119, yielding the same approximate structure as in-vivo ubiquitination. **(B)** SDS-PAGE gel displaying the successful monoubiquitination of H2AC119 along with the generation of the side products Ub, H2A, Ub*Ub, and H2A*H2A across various timepoints in the reaction.

H2AC119ub1 from all of the other products, a Ni-NTA HisTrap HP column was ran to isolate the products containing Ub, which has a 6x-His-tag. In lanes 4 and 5 on the gel in **Figure 8B**, H2A was attempted to be cross-linked with H2A and Ub was attempted to be cross-linked with Ub. Both of these control reactions failed as no higher molecular weight band can be seen on the gel. This is likely due to an inadequate 1,3 DCA concentration, preventing the crosslinking from taking place, leaving only the unreacted products visible on the gel.

H2AC119ub1 was used in conjunction with wild-type H3, H4, and H2B to make an octamer containing a ubiquitinated H2A. To do this, three steps were required to obtain the final octamer product. First, H3 and H4 were combined and purified, producing an H3/H4 tetramer. H2AC119ub1 was then combined with H2B to yield an H2Aub1/H2B dimer, which was mixed with the H3/H4 tetramer to form the ubiquitinated octamer. **Figure 9A** shows the FLPC size-exclusion (Superose 6 Increase 10/300) trace for the H3/H4 tetramer. The peak on the trace eluted at 16.9 mL, corresponding to an approximate molecular weight of 44 KDa, which is where the H3/H4 tetramer, having a mass of 52.8 KDa should elute (Cytiva, 2014). **Figure 9B** details the FLPC size-exclusion (Superose 6 Increase 10/300) trace for the H2Aub1/H2B dimer. This dimer will have an increased molecular weight because it contains H2A which is crosslinked to ubiquitin, as well as the H2B. Ubiquitin has a mass of 8.6 KDa, H2A has a mass of 14 KDa, and H2B is 13.8 KDa, giving a total mass of 36.4 KDa. The peak on the size-exclusion trace for the dimer eluted at 19 mL, which was higher than expected (Cytiva, 2014). An SDS-PAGE gel ran on the peak fractions, however, confirmed the presence of the dimer. Having both the H3/H4 tetramer and H2Aub1/H2B dimer purified, it was now possible to assemble the ubiquitinated octamer. It was important that the required histones were mixed in equimolar amounts, thus requiring twice the molar amount of H2Aub1/H2B dimer as H3/H4 tetramer. **Figure 9C** depicts

the size-exclusion (Superose 6 Increase 10/300) trace for the octamer, with the peak eluting at 16.5 mL. This elution volume correlates to a value of approximately 130 KDa, which is where the octamer containing two ubiquitin, totaling a mass of 125 KDa is expected to elute (Cytiva, 2014). This is further confirmed by the lowered elution volume compared to the H3/H4 and H2Aub1/H2B peaks, indicating an increase in mass. To ensure that the ubiquitinated octamer

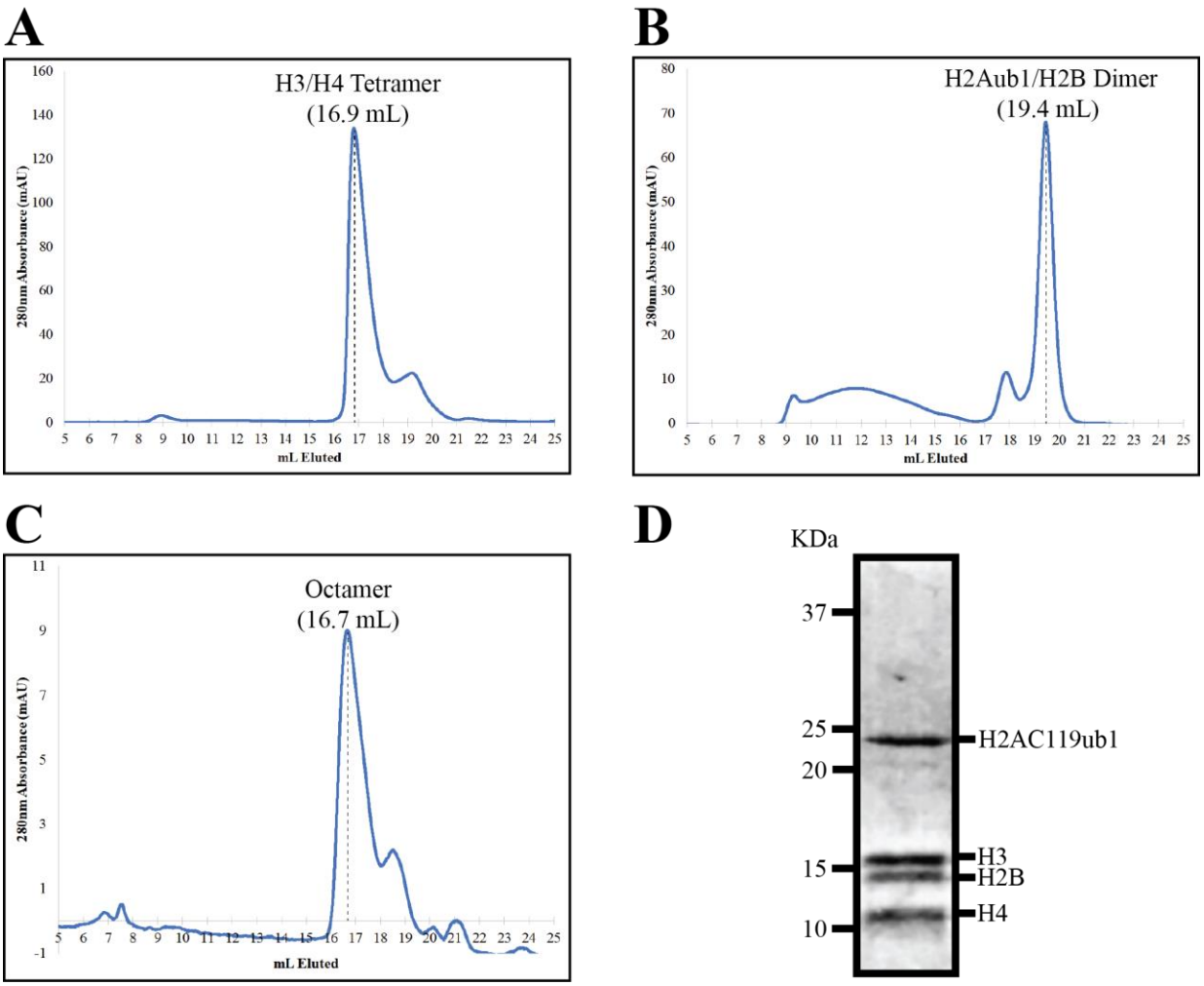

Figure 9: (A) FPLC size-exclusion (Superose 6 Increase 10/300) trace of the H3/H4 tetramer. The peak eluted at 16.9 mL with a max absorbance of 132 mAU. **(B)** FPLC size-exclusion (Superose 6 Increase 10/300) trace of the H2AC119ub1/H2B dimer. The peak eluted at 19.4 mL with a peak absorbance of 67 mAU. **(C)** FPLC size-exclusion (Superose 6 Increase 10/300) trace of the H2AC119ub1 octamer containing two copies each of H2AC119ub1, H2B, H3, and H4. The peak eluted at 16.7 mL with a max absorbance of 9 mAU. **(D)** SDS-PAGE gel of the octamer peak fraction showing the presence of all four histones with the H2AC119ub1 band shifted up.

was successfully purified, an SDS-PAGE gel was run as seen in **Figure 9D**. The three unmodified histones H2B, H3, and H4 all appear at the bottom of the gel, running at their appropriate molecular weights. The H2A band, on the other hand, is shifted upwards between 20 and 25 KDa on the gel, right where H2Aub1 should be at 22.8 KDa.

To test for activity in fractions of PRC1-RYBP isolated by Akhil, an activity assay was performed. This would test the ability of the PRC1-RYBP to transfer ubiquitin from E2-Ub to H2A on the wild-type CpG nucleosome. Before performing the assay, a test to determine the quality of the E2-Ub was done on an SDS-PAGE gel, seen in **Figure 10A**. E2 containing no ubiquitin was used as a control in lane 2 on the gel. An E2-Ub sample was run in

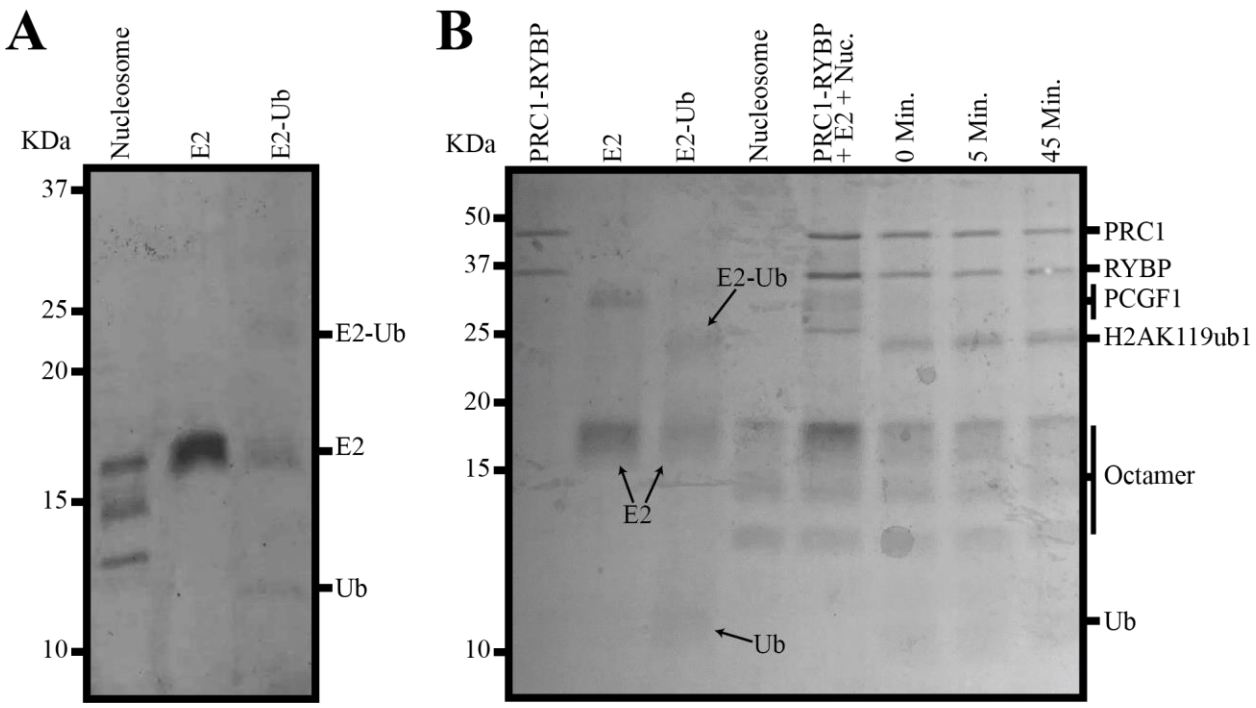

Figure 10: (A) SDS-PAGE gel of nucleosome, E2, and E2-Ub to ensure the presence of E2-Ub. **(B)** Activity assay ran utilizing PRC1-RYBP with E2-Ub and wild-type CpG DNA nucleosome to test for PRC1-RYBP activity. 2 µM PRC1-RYBP, 5 µM E2-Ub, and 0.5 µM nucleosome were incubated together at room temperature for 0, 5, and 45-minute time points, quenching the reaction with 4x Laemmli sample buffer containing no βME. Activity was determined by the shift of the H2A band, becoming H2AK119ub1.

lane 3, and three distinct bands are visible. At the top, there is a faint band that belongs to E2-Ub. Below this is the E2 band, right next to the control, and even further down on the gel is a band belonging to Ub. This gel indicated that despite using a sample buffer containing no βME, there was degradation in the E2-Ub, creating E2 and Ub. It is possible that this may be due to boiling the sample before loading, possibly disrupting the integrity of the H2A-Ub bond, however, the gel did confirm that E2-Ub was present. To run the activity assay, 2 µM PRC1-RYBP, 5 µM E2-Ub, and 0.5 µM nucleosome were incubated together at room temperature for 0, 5, and 45-minute time points. At these time points, the reaction was quenched with 4x Laemmli sample buffer containing no βME, and the resulting samples were run on a gel. A successful assay would be characterized by the appearance of a shifted H2A band towards the top of the gel, indicating that ubiquitin was added to the H2A by PRC1-RYBP. In the gel seen in **Figure 10B**, there is no H2AK119ub1 band seen in any of the control samples. Lane 5 contained a control with all the substrates used in the reaction, substituting E2-Ub for E2 as to prevent any H2AK119ub1 formation. In this lane, there is not any shifted H2A, which is to be expected due to the lack of Ub on E2. Only when E2-Ub is combined with PRC1-RYBP and the WT CpG nucleosome is activity seen. Starting immediately at the 0-minute time point, a new band can be seen, which is the shifted H2A band, representing the H2AK119ub1. This appearance of the band in the 0-minute time point suggests that the conjugation of ubiquitin to H2A occurs very quickly, occurring before the reaction could be quenched. There is a slight increase in band intensity in the 5 and 45-minute time points compared to the initial 0-minute timepoint, suggesting some lag in activity. Although the E2-Ub band and the H2AK119ub1 band run at approximately the same

height, the E2-Ub band is quite hazy and not very concise. The H2AK119ub1 band seen in the 0, 5, and 45-minute timepoints differs from this, appearing much more resolved.

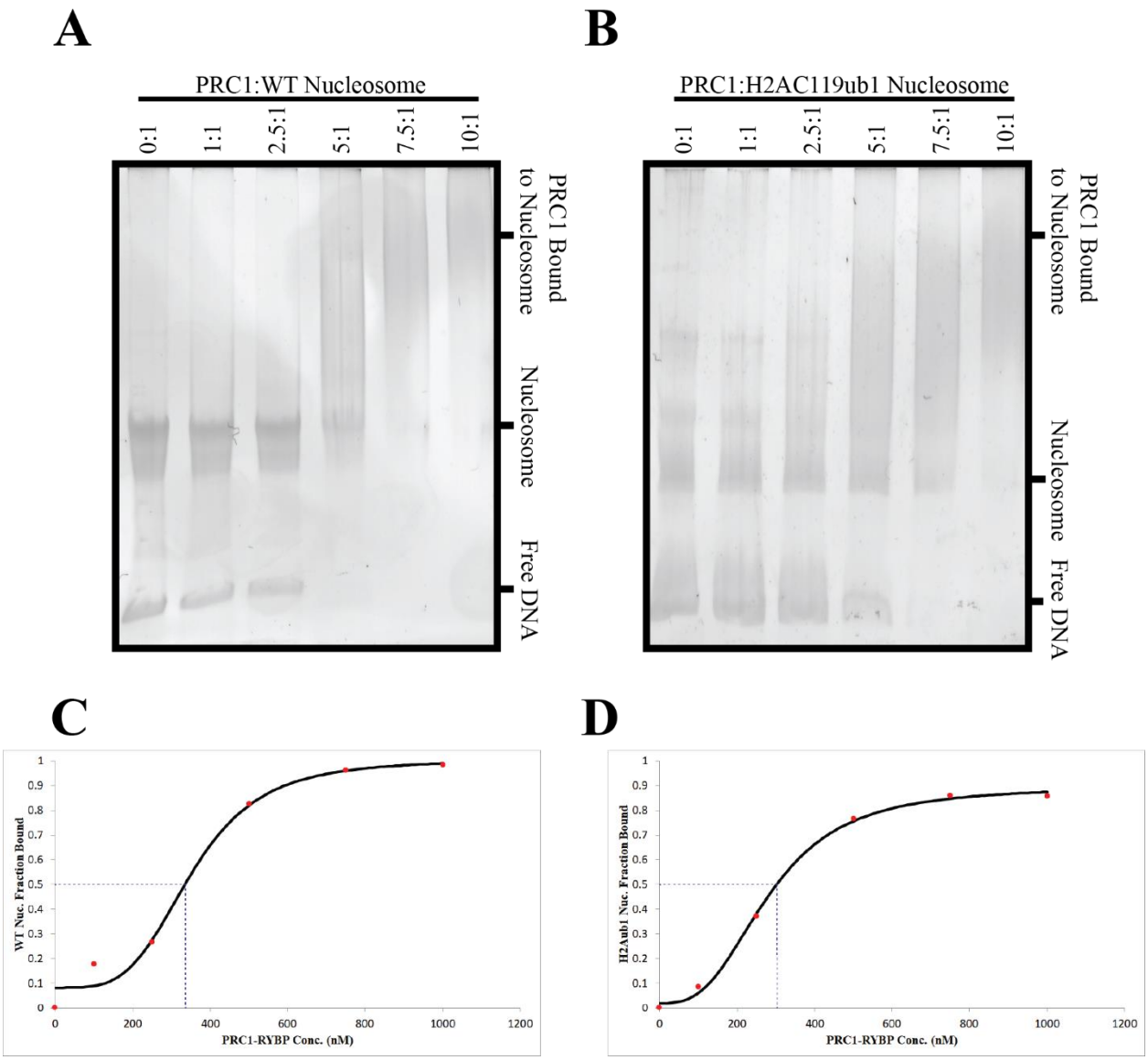

Figure 11: (A) Native Gel EMSA ran using varying ratios of PRC1-RYBP to the wild-type nucleosome. The WT nucleosome concentration of 100 nM remained constant while the PRC1-RYBP concentration ranged from 0 nM (0:1) to 1 µM (10:1). Higher molar ratios of PRC1-RYBP compared to nucleosome lead to increased binding of PRC1 to the nucleosome, causing an upward shift on the gel. **(B)** EMSA ran using varying ratios of PRC1-RYBP to the H2AC119ub1 nucleosome, ranging from 0 nM PRC1-RYBP (0:1) to 1 µM (10:1) PRC1-RYBP with a constant 100 nM of H2AC119ub1 nucleosome in each reaction. **(C)** Binding curve of PRC1-RYBP to the wild-type nucleosome, yielding a K_d of 334 nM. **(D)** Binding curve of PRC1-RYBP to the H2AC119ub1 nucleosome, yielding a K_d of approximately 301 nM.

In order to observe the binding interaction of PRC1-RYBP to the nucleosome, an electrophoresis mobility shift assay (EMSA) can be used. An EMSA uses a native, non-denaturing, gel to observe proteins in their folded state. This is especially useful to study protein interactions which depend on the structure and conformation of the proteins themselves. In this experiment, 100 nM of nucleosome was incubated with varying concentrations of PRC1-RYBP to observe the binding affinity of PRC1 to the nucleosome CpG DNA. In **Figure 11A**, the first lane contains only 100 nM nucleosome, setting a baseline for which the amount of bound protein can be determined. At the bottom of the gel is free DNA, and the nucleosome can be seen roughly at the center of the gel. As the concentration of PRC1-RYBP increases, a smear begins to form on the gel above the nucleosome band. This smear represents the protein-DNA complex being formed between the PRC1-RYBP and nucleosome, causing it to shift further up on the gel. As more PRC1-RYBP is added to the nucleosome, more of the protein will become bound to the DNA, leading to an even further shift up the gel. Also seen is a shift in the free DNA, with the band disappearing almost completely at a PRC1-RYBP concentration of 500 nM as the free DNA binds protein. **Figure 11B** displays another EMSA between PRC1-RYBP and the nucleosome, this time using the H2AC119ub1 CpG DNA nucleosome. In this gel, although the ratio of PRC1:nucleosome is the same, there is a less pronounced shift than what was seen with the WT nucleosome. The nucleosome band does not begin to disappear until a 1μM concentration of PRC1-RYBP is reached. The protein-DNA smear at the top of the gel is also much broader than the one found when using the wild-type nucleosome. **Figure 11C** shows a binding curve for PRC1-RYBP to the WT nucleosome, comparing the fraction bound to the concentration of PRC1-RYBP. This was done by comparing the band intensity of the nucleosome band, with decreasing intensity indicating increased binding. The equilibrium

dissociation constant (K_d) of the interaction can be found by determining the PRC1-RYBP concentration at which half of the H2Aub1 nucleosome core particle is bound. For the wild-type nucleosome, the K_d is 334 nM. **Figure 11D** displays the binding curve for the H2Aub nucleosome. The calculated K_d for the ubiquitinated nucleosome is 301 nM. Repetition of the experiment will likely be required to better fit the data to the ideal curve. This would also allow for the calculation of error and standard deviation when figuring out the K_d values. This collected data provides preliminary evidence that PRC1-RYBP has a higher affinity for the H2Aub1 nucleosome than it does for the WT nucleosome. The K_d of 301 nM for the H2Aub1 nucleosome is slightly lower than the K_d for the WT nucleosome of 334 nM. Because a lower K_d corresponds to tighter binding, the lower K_d for the H2Aub1 nucleosome to PRC-1-RYBP indicates that it has a higher affinity for it.

Conclusions

First of all, the isolated PRC-RYBP complex was found to be active. An activity assay ran using the PRC1-RYBP, CpG WT nucleosome, and E2-Ub showed that the PRC1-RYBP was capable of binding to CpG rich nucleosome sequences, attaching ubiquitin to H2AK119 on the nucleosome. This E3 ubiquitin ligase activity is critical for the modification of chromatin and plays an important role in gene regulation.

The PRC1-RYBP was also found to bind both the CpG WT nucleosome and the H2Aub1 nucleosome with very high affinity through the use of an EMSA. Higher PRC1-RYBP concentrations in relation to the nucleosome concentration correspond to a large increase in the fraction of PRC1-RYBP bound to the nucleosome. This was seen through a shift upwards in the gel of the nucleosome band, as the PRC1-RYBP forms a protein-DNA complex with the nucleosome. This affinity of PRC1-RYBP towards the nucleosome was also seen to be higher for the H2Aub1 nucleosome than for the WT nucleosome, as indicated by the observed K_d values. The K_d value of 301 nM for PRC1-RYBP to the H2Aub1 nucleosome is lower than the K_d value of 334 nM for the WT nucleosome. This difference in K_d suggests tighter binding of PRC1-RYBP to the H2Aub1 nucleosome.

Supplemental Information

Native H3

ATGGCCCGTACCAAGCAGACCGCCCNNNAANCCACCGGAGGGAAGGCTCCCCGCAA
GCAGCTGGCCACCAAGGCAGCCAGGAAGTCCGCTCCTGCTACCGGCGGAGTCAAGA
AACCTCACCGTTACCGGCCCGGCACAGTCGCTCTCCGCGAGATCCGCCGCTACCAGA
AATCCACCGAGCTGCTCATCCGCAAACTGCCTTTCCAGCGCCTGGTCCGGGAGATCG
CTCAGGACTTCAAGACCGACCTGCGCTTCCAGAGCTCGGCCGTTATGGCTCTGCAGG
AGGCCAGCGAGGCTTATCTGGTCGCTCTCTTTGAGGACACCAACCTGTGCGCCATCC
ACGCCAAGAGGGTCACCATCATGCCCAAGGACATCCAGCTGGCCCGCAGAATCCGA
GGCGAGAGGGCT

Native H4

ATGTCTGGTCGTGGTAAAGGTGGTAANGGTCTGGGTAAAGGTGGTGCTAAACGTCA
CCGTAAAGTTCTGCGTGACAACATCCAGGGTATCACCAAGCCGGCTATCCGTCGTCT
GGCTCGTCGTGGTGGTGTTAAACGTATCTCCGGTCTGATCTACGAAGAAACCCGCGG
TGTTCTGAAAGTTTTCCTGGAAAACGTTATCCGTGACGCTGTTACCTACACCGAACA
CGCTAAACGTAAAACCGTTACCGCTATGGACGTTGTTTACGCTCTGAAACGTCAGGG
TCGTACCCTGTACGGTTTCGGTGGT

Native H2A

ATGTCAGGAAGAGGCAAACAAGGCGGTAAAACCCGCGCTAAGGCCAAGACTCGCTC
ATCTCGGGCTGGGCTACAGTTCCCTGTTGGCCGTGTTCACCGGCTGTTAAGGAAAGG
CAATTATGCAGAGCGGGTGGGAGCTGGAGCTCCAGTCTATCTGGCTGCAGTGTTGGA
GTATCTGACCGCTGAGATTTTGGAATTGGCCGGGAATGCGGCCCGTGATAACAAGA
AGACTCGCATTATCCCCAGACACCTGCAGCTCGCTGTGCGCAACGATGAGGAACTG
AACAAACTGCTCGGAAGAGTCACTATCGCTCAGGGCGGGGTCCTGCCCAACATCCA
GTCCGTGCTGCTGCCCAAGAAAACCGAGAGTTCCAAGTCGGCCAAGAGCAAG

Native H2B

ATGGCCAAGTCCGCTCCAGCCCCGAAGAAAGGCTCCAAGAAAGCGGTGACCAAGAC
TCAGAAGAAAGACGGGAAAAAGCGCAGGAAGACAAGGAAGGAGAGTTATGCCATT
TACGTGTACAAGGTGCTGAAGCAGGTGCACCCCGATACCGGCATCTCGTCCAAGGC
CATGAGCATCATGAACTCCTTTGTCAACGATGTGTTTGAGCGCATCGCAGGGGAAGC
CTCCCGCCTGGCTCATTACAACAAGCGCTCCACCATCACCTCCCGGGAGATCCAGAC
CGCGGTCCGACTGCTGCTGCCTGGGGAGTTGGCCAAACACGCCGTGTCCGAGGGCA
CCAAGGCTGTCACCAAGTACACCAGCGCCAAG

H3K18CK23C

ATGGCCCGTACCAAGCAGACCGCCCGTAAATCCACCGGAGGGAAGGCTCCCCGCTG
CCAGCTGGCCACCTGCGCAGCCAGGAAGTCCGCTCCTGCTACCGGCGGAGTCAAGA
AACCTCACCGTTACCGGCCCGGCACAGTCGCTCTCCGCGAGATCCGCCGCTACCAGA
AATCCACCGAGCTGCTCATCCGCAAACTGCCTTTCCAGCGCCTGGTCCGGGAGATCG
CTCAGGACTTCAAGACCGACCTGCGCTTCCAGAGCTCGGCCGTTATGGCTCTGCAGG
AGGCCAGCGAGGCTTATCTGGTCGCTCTCTTTGAGGACACCAACCTGGCCGCCATCC
ACGCCAAGAGGGTCACCATCATGCCCAAGGACATCCAGCTGGCCCGCAGAATCCGA
GGCGAGAGGGCT

H2AK119C

ATGTCAGGAAGAGGCAAACAAGGCGGTAAAACCCGCGCTAAGGCCAAGACTCGCTC
ATCTCGGGCTGGGCTACAGTTCCCTGTTGGCCGTGTTCACCGGCTGTTAAGGAAAGG
CAATTATGCAGAGCGGGTGGGAGCTGGAGCTCCAGTCTATCTGGCTGCAGTGTTGGA
GTATCTGACCGCTGAGATTTTGGAATTGGCCGGGAATGCGGCCCGTGATAACAAGA
AGACTCGCATTATCCCCAGACACCTGCAGCTCGCTGTGCGCAACGATGAGGAACTG
AACAAACTGCTCGGAAGAGTCACTATCGCTCAGGGCGGGGTCCTGCCCAACATCCA
GTCCGTGCTGCTGCCCTGCAAAACCGAGAGTTCCAAGTCGGCCAAGAGCAAG

UbG76C

CAACAACCACGAAGTCTCAACACAAGATGAAGAGTAGACTCCTTTTGAATATTGTA
GTCAGACAAAGTACGTCCATCTTCCAGCTGCTTGCCAGCAAAGATCAGTCTCTGCTG
ATCAGGAGGAATTCCTTCCTTATCCTGGATCTTGGCCTTTACATTTTCTATCGTATCC
GAGGGTTCAACCTCGAGGGTGATGGTCTTCCCCGTAAGGGTTTTCACGAAAATCTGC
ATCA

CpG Nucleosome DNA (147bp widom sequence in bold)

CACGCGACTGTGTGCCCGTCAGACGCTGCGCTGCCGGCGG**CTGGAGAATCCCGGT
GCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGT
ACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAG
GCACGTGTCAGATATATACATCCTGT**ATGCATGCATATCATTCGATCGGAGCTCCC
GATCGATGC

References

Brockdorff, N. (2017). Polycomb complexes in X chromosome inactivation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1733), 20170021. https://doi.org/10.1098/rstb.2017.0021

Chantalat, S., Depaux, A., Héry, P., Barral, S., Thuret, J.-Y., Dimitrov, S., & Gérard, M. (2011). Histone H3 trimethylation at lysine 36 is associated with constitutive and facultative heterochromatin. *Genome Research*, *21*(9), 1426–1437. https://doi.org/10.1101/gr.118091.110

Chen, T., & Dent, S. Y. (2013). Chromatin modifiers and remodellers: Regulators of cellular differentiation. *Nature Reviews Genetics*, *15*(2), 93–106. https://doi.org/10.1038/nrg3607

Chittock, E. C., Latwiel, S., Miller, T. C. R., & Müller, C. W. (2017). Molecular architecture of Polycomb repressive complexes. *Biochemical Society Transactions*, *45*(1), 193–205. https://doi.org/10.1042/bst20160173

Cytiva. (2014). *Superose 6 Increase 10/300 GL - Instructions.* Retrieved from https://lab.rockefeller.edu/chen/assets/file/Superose_Tricorn_column_manual.ED.pdf

Cytiva. (2020). *Superdex 200 Increase 10/300 GL - Instructions for use.* Retrieved from https://cdn.cytivalifesciences.com/dmm3bwsv3/AssetStream.aspx?mediaformatid=10061&destinationid=10016&assetid=16443

Dyer, P. N., Edayathumangalam, R. S., White, C. L., Bao, Y., Chakravarthy, S., Muthurajan, U. M., & Luger, K. (2003). Reconstitution of nucleosome core particles from recombinant histones and DNA. *Chromatin and Chromatin Remodeling Enzymes, Part A*, 23–44. https://doi.org/10.1016/s0076-6879(03)75002-2

Glozak, M. A., & Seto, E. (2007). Histone deacetylases and cancer. *Oncogene*, *26*(37), 5420–5432. https://doi.org/10.1038/sj.onc.1210610

Iwasaki, W., Miya, Y., Horikoshi, N., Osakabe, A., Taguchi, H., Tachiwana, H., Shibata, T., Kagawa, W., & Kurumizaka, H. (2013). Contribution of histone N-terminal tails to the structure and stability of nucleosomes. *FEBS Open Bio*, *3*(1), 363–369. https://doi.org/10.1016/j.fob.2013.08.007

Lavarone, E., Barbieri, C. M., & Pasini, D. (2019). Dissecting the role of h3k27 acetylation and methylation in PRC2 mediated control of cellular identity. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-09624-w

Long, L., Furgason, M., & Yao, T. (2014). Generation of nonhydrolyzable ubiquitin–histone mimics. *Methods*, *70*(2–3), 134–138. https://doi.org/10.1016/j.ymeth.2014.07.006

Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., & Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, *389*(6648), 251–260. https://doi.org/10.1038/38444

Luo, M., Platten, D., Chaudhury, A., Peacock, W. J., & Dennis, E. S. (2009). Expression, imprinting, and evolution of rice homologs of the Polycomb group genes. *Molecular Plant*, *2*(4), 711–723. https://doi.org/10.1093/mp/ssp036

Marino-Ramirez, L., Levine, K. M., Morales, M., Zhang, S., Moreland, R. T., Baxevanis, A. D., & Landsman, D. (2011). The histone database: An integrated resource for histones and histone fold-containing proteins. *Database*, *2011*. https://doi.org/10.1093/database/bar048

Michael, A. K., & Thomä, N. H. (2021). Reading the chromatinized genome. *Cell*, *184*(14), 3599–3611. https://doi.org/10.1016/j.cell.2021.05.029

Parreno, V., Martinez, A.-M., & Cavalli, G. (2022). Mechanisms of Polycomb group protein function in cancer. *Cell Research*, *32*(3), 231–253. https://doi.org/10.1038/s41422-021-00606-6

Piunti, A., & Shilatifard, A. (2021). The roles of Polycomb repressive complexes in mammalian development and cancer. *Nature Reviews Molecular Cell Biology*, *22*(5), 326–345. https://doi.org/10.1038/s41580-021-00341-1

Ren, W., Fan, H., Grimm, S. A., Guo, Y., Kim, J. J., Yin, J., Li, L., Petell, C. J., Tan, X.-F., Zhang, Z.-M., Coan, J. P., Gao, L., Cai, L., Detrick, B., Çetin, B., Cui, Q., Strahl, B. D., Gozani, O., Wang, Y., … Song, J. (2020). Direct readout of Heterochromatic H3K9me3 regulates DNMT1-mediated maintenance DNA methylation. *Proceedings of the National Academy of Sciences*, *117*(31), 18439–18447. https://doi.org/10.1073/pnas.2009316117

Rose, N. R., King, H. W., Blackledge, N. P., Fursova, N. A., Ember, K. J., Fischer, R., Kessler, B. M., & Klose, R. J. (2016). RYBP stimulates PRC1 to shape chromatin-based communication between Polycomb repressive complexes. *eLife*, *5*. https://doi.org/10.7554/elife.18591

Seto, E., & Yoshida, M. (2014). Erasers of histone acetylation: The histone deacetylase enzymes. *Cold Spring Harbor Perspectives in Biology*, *6*(4). https://doi.org/10.1101/cshperspect.a018713

Streich, F. C., & Lima, C. D. (2014). Structural and functional insights to ubiquitin-like protein conjugation. *Annual Review of Biophysics*, *43*(1), 357–379. https://doi.org/10.1146/annurev-biophys-051013-022958

Trojer, P., & Reinberg, D. (2007). Facultative heterochromatin: Is there a distinctive molecular signature? *Molecular Cell*, *28*(1), 1–13. https://doi.org/10.1016/j.molcel.2007.09.011

Vann, K. R., & Kutateladze, T. G. (2017). Histone H3 dual ubiquitylation mediates maintenance DNA methylation. *Molecular Cell*, *68*(2), 261–262. https://doi.org/10.1016/j.molcel.2017.10.007

Wang, W., Qin, J.-J., Voruganti, S., Nag, S., Zhou, J., & Zhang, R. (2015). Polycomb Group (PCG) proteins and human cancers: Multifaceted functions and therapeutic implications. *Medicinal Research Reviews*, *35*(6), 1220–1267. https://doi.org/10.1002/med.21358

Xu, J., & Liu, Y. (2021). Probing chromatin compaction and its epigenetic states in situ with single-molecule localization-based super-resolution microscopy. *Frontiers in Cell and Developmental Biology*, *9*. https://doi.org/10.3389/fcell.2021.653077