Identifying Novel Interactions Between the CDK8 Kinase Module and the Deubiquitinating Enzyme Ubiquitin Specific Protease 7 (USP7)

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE

STUDIES IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

GRADUATE PROGRAM IN BIOLOGY

YORK UNIVERSIRY

TORONTO, ONTARIO

JULY 2022

ABSTRACT

Ubiquitin Specific Protease 7 (USP7) is a deubiquitinating enzyme that regulates and interacts

with various cellular substrates, playing regulatory roles in various pathways such as DNA

damage repair and epigenetics. This thesis explores the role of USP7 in regulating components

of gene expression regulators. The gene expression of target genes is fine-tuned by the

transcriptional co-activator, the Mediator complex. Our group has determined that Cyclin

Dependent Kinase 8 (CDK8) and Mediator Complex Subunit 12 (MED12), components of the

kinase module in the Mediator complex, contain consensus sequences for binding to the

N-terminus of USP7. Using various co-immunoprecipitation, pulldowns, tissue culture, and

western blotting related techniques, our findings show that USP7, using the DWGF motif at its

N-terminus, interacts with and stabilizes protein levels of MED12 and CDK8. This provides

insights onto the cellular role of USP7 in gene expression, giving such novel interactions

therapeutic potential for diseases such as cancer.

ACKNOWLEGMENTS

I would like to first and foremost, give thanks to my supervisor, Dr. Vivian Saridakis, for giving me the opportunity to be a part of her research team. Despite doing my master's during a pandemic, she was nevertheless always approachable, and provided me with unending support, guidance, and wise insights for my journey with this research project. Her support and belief in me have pushed and fueled my interest in exploring different avenues and techniques for the research that I have conducted. It was such an honour to work with her.

I would also like to extend my gratitude to my advisor, Dr. Dasantilla Golemi-Kotra, for finding the time to revise my work and provide exceptional suggestions and insights throughout my project. I would also like to thank Dr. Yi Sheng for helping me troubleshoot some of the experiments, as well as for letting me borrow some of her lab reagents and materials. Finally, I extend my gratitude to Dr. Emmanuel Rosonina for his valuable suggestions with respect to my project, as well as for letting me borrow his lab reagents and equipment such as the imager.

Thank you as well to Niha for her excellent training and expertise when I first started working in this lab. Her support and help extended beyond the lab hours. I would also like to thank my current lab mates Sukhdeep, Aneesah, and Nikola and former lab members, Anna and Helen, for their help, support, and encouragement throughout my research project.

Lastly, I would like to thank my family for all their love, patience, and support, from the beginning, making my research journey more enjoyable. Thank you as well to all my dear friends, for their amazing friendship, help, and motivation.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AML Acute Myeloid Leukemia

APC Adenomatous Polyposis Coli

BMI-1 B cell -specific Moloney murine leukemia virus integration site 1

BMP Bone Morphogenic Pathway

BMPs Bone Morphogenic Proteins

BRCA1 Breast Cancer Type 1 Susceptibility Protein

CDK7 Cyclin Dependent Kinase 7

CDK8 Cyclin Depeendet Kinase 8

co-IP Co-Immunoprecipitation

CRC Colorectal Cancer

CTD Carboxy Terminal Domain

DDR DNA Damage Response

DNMT1 DNA Methyl Transferase 1

DSB Double Stranded Breaks

DSIF DRB sensitivity-inducing factor

DUB Deubiquitinating Enzyme

EBNA1 Epstein-Barr Virus Nuclear Antigen 1

EMT Epithelial to Mesenchymal Transition

ER Estrogen Receptor

EV Empty Vector

EZH2 Enhancer of Zeste Homolog 2

E1 Ubiquitin Activating Enzyme

E2 Ubiquitin Conjugating Enzyme

E3 Ubiquitin Ligase

FOXO3 Forkhead box O proteins 3

FOXO4 Forkhead box O proteins 4

GREB1 Growth Regulating Estrogen Receptor Binding 1

HAUSP Herpes Virus Associated Protease

HCT116 Colorectal Carcinoma Cell Line

HCT116 USP7 -/- Colorectal Carcinoma Cell Line Usp7 Knock-Out

Hdm2 Human double minute 2 homolog

HECT Homologous to the E6Ap Carboxyl Terminus

HeLa Human Cervical Adenocarcinoma Cell Line

HSV Herpes Simplex Virus

H3K4 Histone 3 Lysine 4

H3K9 Histone 3 Lysine 9

H3K27 Histone H3 Lysine 27

ICD Intracellular Domain

ICP0 Infected Cell Protein 0

JAMM JAMM motif proteases

LSD1/KDM1 Lysine-specific demethylase 1

MATH Meprin and TRAF Homology

MCM Minichromosome maintenance protein complex

MCM-BP MCM binding protein

MDM2 Mouse double minute 2 homolog

MED12 Mediator Complex Subunit 12

MED14 Mediator Complex Subunit 14

MED23 Mediatory Complex Subunit 23

mH2A1 MacroH2A1

MJD Machado-Joseph Domain protease

MLL5 Methyltransferase Mixed lineage leukemia 5

ncRNA Noncoding RNA

NELF Negative Elongation Factor

NER Nucleotide-Excision Repair

OTU Ovarian Tumour related proteases

PPM1G Protein phosphatase 1G

PRC Polycomb Repressive Complexes

Pre-RC Pre-replicative complex

P-TEFb Positive Transcription Elongation Factor b

REST RE1 Silencing Transcription Factor

RBR Ring-Between-Ring

RING Really Interesting New Gene

RING1B Ring Finger Protein 1B

RNAPl II RNA Polymerase II

RNF168 Ring Finger Protein 168

RNF8 Ring Finger Protein 8

SEC P-TEFb-containing super elongation complex

SNAI1 & 2 Snail 1& 2

STAT1/3/5 Signal Transducer and Activator of Transcription 1/3/5

SUMO Small Ubiquitin-Like Modifier

TCF/LEF T-cell factor/lymphoid enhancer factor

TF Transcription Factors

TGFβ Transforming Growth Factor Beta

Tip60 Tat Interactive Protein 60

TRAF Tumor necrosis factor receptor-associated factor

UAS Upstream Activating Sequences

Ub Ubiquitin

UBA52 Ubiquitin A-52 Residue Ribosomal Protein Fusion Gene

UBA80 Ubiquitin Carboxyl Extension Protein 80 Gene

UBB Ubiquitin B Gene

UBC Ubiquitin C Gene

Ubl Ubiquitin-like

UCH Ubiquitin C-terminal Hydrolases

UHRF1 Ubiquitin-like with PHD And Ring Finger Domains

UPS Ubiquitin-Proteosome System

USP Ubiquitin-Specific Proteases

U2OS Human Osteosarcoma Cell Line

vIRF Viral Interferon Regulatory Factor

XLMR X-linked Mental Retardation

XPC Xeroderma pigmentosum complementation group C

CHAPTER 1: INTRODUCTION

1.1 THE REGULATION OF CELLULAR PROTIENS IN EUKARYOTIC CELLS

Protein synthesis is essential for cellular development and growth. Proteins possess cellular functions, respond to extracellular signals, and regulate cellular processes. However, regulatory mechanisms are needed to control the levels of proteins in the cell and prevent the accumulation of protein products when they are not required. The cell also requires signals to prevent protein turnover when a protein product is required for cellular functions. The balance between protein turnover and maintenance helps to achieve cellular homeostasis. The Ubiquitin-Proteasome (UPS) system governs the regulation of cellular protein levels in the cells.

1.1.1 UBIQUITIN

Ubiquitination is a tightly regulated post-translational modification in eukaryotes, whereby the C-terminal tail of a ubiquitin (Ub) molecule is covalently attached to a target protein substrate. This event alters the function and/or stability of the protein (Swatek and Komander, 2016). Ub is a highly conserved small protein molecule, 76 amino acids (8.5 kDa), that is ubiquitous across many tissues and organisms (Roos-Mattjus and Sistonen, 2004). In mammalian cells, Ub is encoded from 4 different genes, two of which, Ubiquitin B Gene (UBB) and Ubiquitin C Gene (UBC), produce polyubiquitin chains, while the Ubiquitin A-52 Residue Ribosomal Protein Fusion Gene (UBA52) and Ubiquitin Carboxyl Extension Protein 80 Gene (UBA80) encode for ubiquitin fused with ribosomal proteins. Subsequently, the resulting ubiquitin products are further processed to release free ubiquitin for protein ubiquitination (Kimura and Tanaka, 2010). Protein substrates can have various degrees of ubiquitination. The substrate can be monoubiquitinated, that is, contain one molecule of ubiquitin, have multiple

1

monoubiquitylations, or be polyubiquitinated, in which ubiquitin chains are formed on the target

substrate (Figure 1.1). Now, polyubiquitination can occur in linear or branched chains. Ubiquitin

itself contains seven lysine residues, Lys6, Lys11, Lys27, Lys29, Lys33, Lys48 and Lys63, that

can also form isopeptide-linked ubiquitin chains on its surface (Ikeda and Dikic, 2008). Thus, the

type of ubiquitin linkages can also vary. For instance, ubiquitination of the same lysine linkages

is known as homotypic, whereas ubiquitination containing various types of linkages is known as

heterotypic (Swatek and Komander, 2016). Ubiquitin molecules can also undergo

post-translational modifications which involves phosphorylation, acetylation and SUMOylation

(Figure 1.1). These modifications can result in ubiquitin molecules acting as second messengers

and participating in signal transduction in the cells. The modification of ubiquitin may also alter

its recognition by other enzymes. For instance, phosphorylation of the ubiquitin Ser65 residue

impacts the recognition of the molecule by some deubiquitinating enzymes (Swatek and

Komander, 2016).
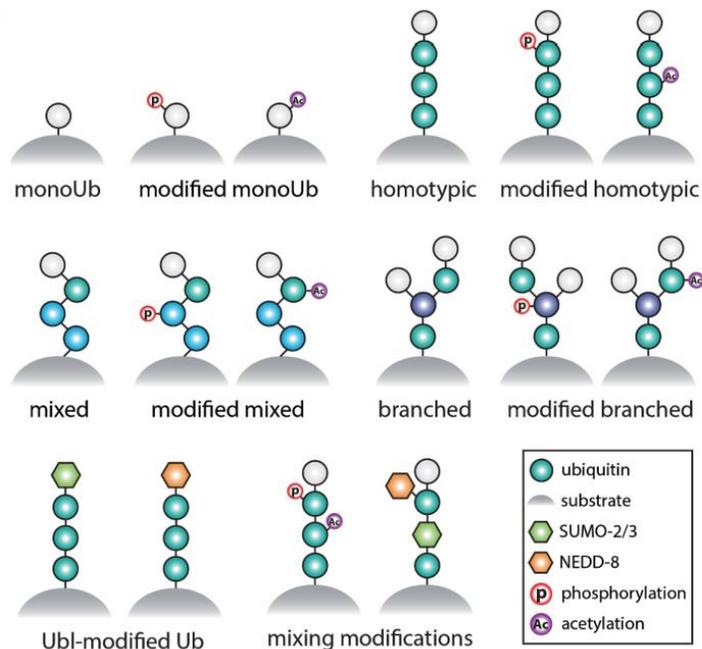
Figure 1.1. **Structural Organization of Ubiquitin**. Ubiquitin can be linked to a substrate either as a single molecule (monoubiquitylation) or as linear or branched chains (polyubiquitinations). The linkage types can be homotypic or heterotypic, with various ubiquitin modifications. *Retrieved from* Swatek and Komander, 2016.

1.1.2 PROTEIN UBIQUTINATION

The addition of ubiquitin to a substrate occurs via three enzymatic reactions. Firstly, ubiquitin is activated by the E1 ubiquitin-activating enzyme with the expenditure of energy, ATP, forming the thiol ester intermediate E1-Ub. The activated ubiquitin is then transferred to the active site cysteine residue of the E2 ubiquitin conjugating enzyme, forming a thiol ester bond (Pickart and Eddins, 2004). In the final step of the reaction, E2 enzyme binds to an E3 ubiquitin ligase and the E3 ligase catalyzes the addition of the ubiquitin molecule via the formation of an isopeptide bond between the carboxyl terminus of glycine residue in ubiquitin and an ε-amino group of a lysine residue on the target substrate (Figure 1.2) (Foot *et al.*, 2017). The human genome encodes for only two E1 enzymes, approximately 40 E2 enzymes, and approximately 600 E3 ligases (Chen and Sun, 2009). Due to their abundance, E3 ligases facilitate substrate specificity. That is, E3 ligases play a role in mediating the ubiquitination of different substrates. Based on the mechanism of ubiquitin transfer and the presence of specific domains, the E3 ligases can be divided into three main families: Really Interesting New Gene (RING), Homologous to E6-associated protein C-terminus (HECT), and RING-in-between-RING (RBR) ligases (Figure 1.2 B) (Foot *et al.,* 2017). The HECT E3 ligases ubiquitinate substrates in two steps. They first conjugate the ubiquitin molecules to their catalytic cysteine residue via a trans-thioesterification reaction, followed by the transfer of ubiquitin onto the substrate. Furthermore, RING E3 ligases are the largest group of E3 ligases and unlike the HECT E3 ligases, they act as a scaffold to directly transfer ubiquitin from the E2 enzyme to the target substrate (Foot *et al.,* 2017). Finally, the RBR E3 ligases also utilizes two-step ubiquitin transfer mechanisms as HECT E3 ligases.
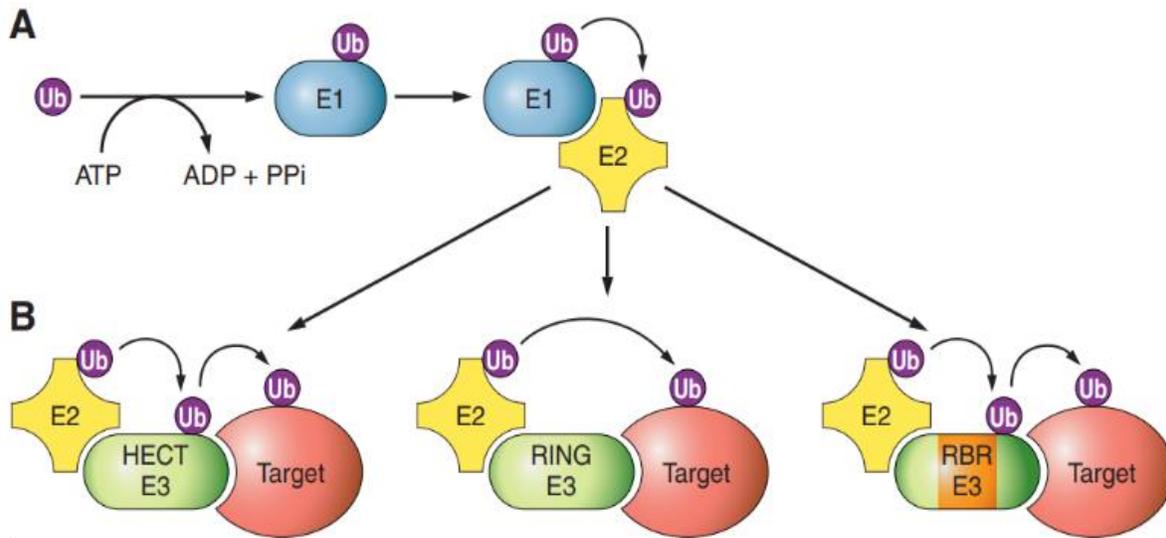
Figure 1.2. **The Ubiquitin-Substrate Conjugation.** A) Protein ubiquitination occurs by three main enzymatic reactions involving ubiquitin-activating (E1), ubiquitin-conjugating (E2) and ubiquitin-protein ligase (E3) enzymes. B) The mechanism of ubiquitin transfer depends on the family of E3 ligase present. *Retrieved from* Foot *et al.,* 2017.

Depending on the extent of ubiquitination and the linkage type formed on the substrate, several outcomes can result (Figure 1.3). For instance, monoubiquitylation of substrates have been shown to direct proteins in DNA repair and chromatin remodeling (Buetow and Huang, 2016). In contrast, polyubiquitination with different lysine linkages play other roles in the cell. More than 50% of protein ubiquitination involves Lys48 linked polyubiquitin chains, which serves to signal the substrate for protein degradation by the 26S proteasome complex (Swatek and Komander, 2016). Degradation of proteins via the UPS ensures that damaged, misfolded, or mutated proteins are removed, preventing their pathological accumulation. For this reason, any dysregulation in the UPS pathway can give rise to diseases such as cancer and neurodegenerative diseases (Reinstein and Ciechanover 2006).

Additionally, other types of ubiquitination linkages can also have non-degradative roles. For instance, polyubiquitinated proteins that involve mixed or branched Lys63 linkages promote NF-κB signaling, as the ubiquitinated protein acts as a scaffold at the immune receptors such as IL-1 receptors (Swatek and Komander, 2016). All in all, ubiquitin modifications regulate and govern various biological processes such as DNA repair and remodeling, endocytosis, endosomal trafficking, and gene transcription (Ikeda and Dikic, 2008).
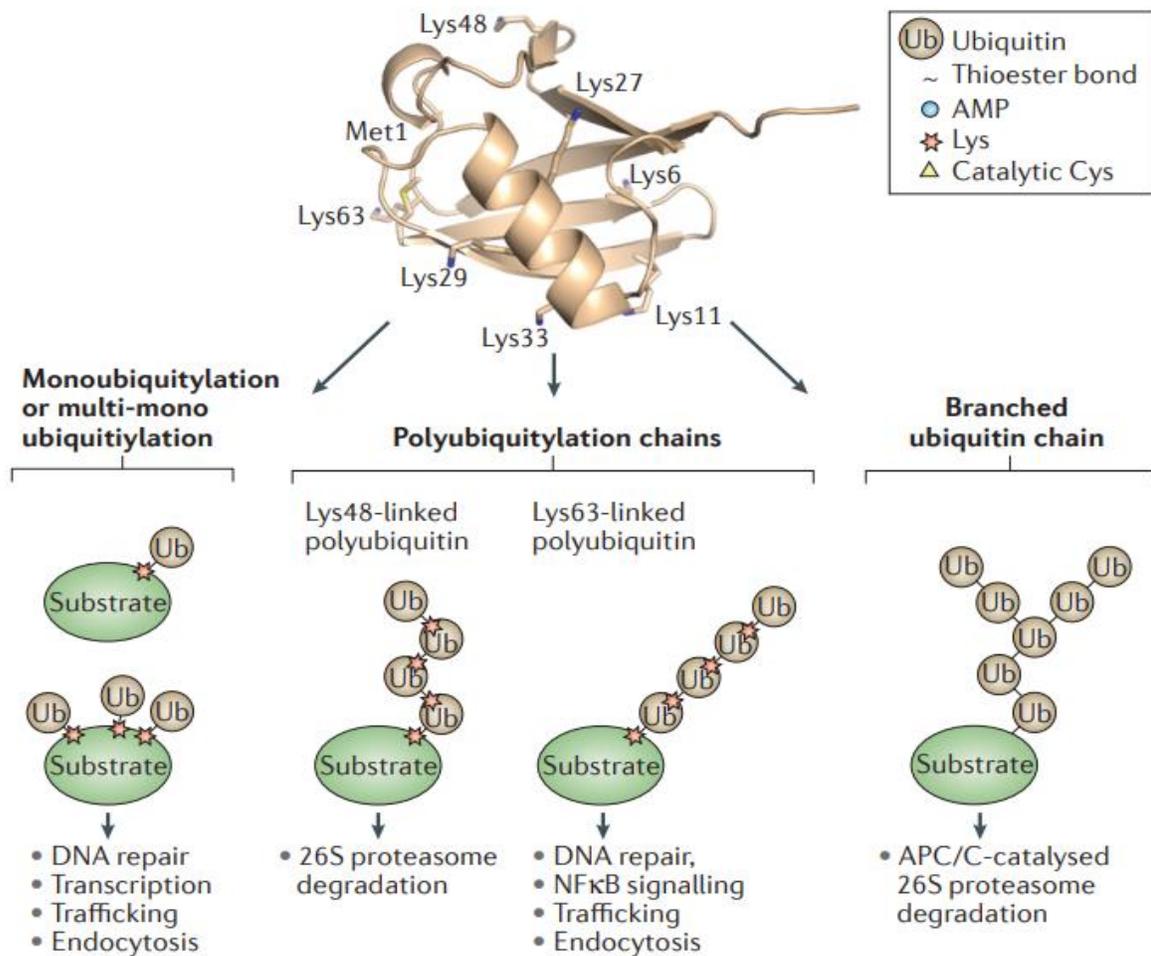
Figure 1.3. **Types and Functions of Ubiquitin Chain Linkages.** Ubiquitin contains 7 lysine residues itself that can undergo further ubiquitination and modification. Depending on the lysine linkage formed, it can target the substrate to various biological role. Lys48 linked polyubiquitination for instance marks substrate for protein degradation by the 26S proteosome complex. *Retrieved from* Buetow and Huang, 2016.

1.1.3 PROTEIN DEUBIQUTINATION

The process of ubiquitination is reversible and is known as protein deubiquitination. The removal of ubiquitin molecules from a target substrate is carried out by de-ubiquitinating enzymes, also known as DUBs. The human genome encodes for approximately 100 different DUBs. These enzymes remove ubiquitin chains from substrates to prevent them from being degraded by the 26S proteasome complex, or alter their function. Furthermore, DUBs help maintain the pool of free ubiquitin in the cells (Roos-Mattius and Sistonen, 2004). DUBS are also responsible for cleaving and processing newly synthesized ubiquitin fusion proteins or ubiquitin chains on a substrate. This results in the release of ubiquitin monomers and thus their recycling, maintaining a ubiquitin pool for future substrate conjugation (Snyder and Silva, 2021). The dynamic between protein ubiquitination and deubiquitination maintains the homeostasis of cellular proteins. Thus, any dysregulation or mutations in any of these two pathways can give rise to diseases such as cancer.

DUBs can be classified into two major classes: metalloproteases and cysteine proteases. DUBs in the family of metalloproteases catalyze the removal of ubiquitin isopeptide via a catalytic serine and utilize a zinc ion as the cofactor. Thus, metalloproteases are zinc dependent. In this class of DUBs, the JAB1/MPN/MOV34 (JAMM/ MPN+) motif is conserved among the metalloproteases.

On the other hand, cysteine proteases contain the catalytic cysteine, histidine, and either asparagine or aspartic acid, that facilitate the cleavage of ubiquitin molecules. Within this class, six different DUB families share such characteristic but also have their own conserved domains and organization. These include Ubiquitin-Specific Proteases (USP), Ubiquitin C-terminal Hydrolases (UCH), Machado-Joseph Domain Protease (MJD), Ovarian Tumour Related

Proteases (OTU), JAMM motif proteases (JAMM), K48 polyubiquitin specific MINDY domain families, and the newest discovered DUB family, the zinc finger with UFM1-specific peptidase domain (Snyder and Silva, 2021). The USP family is the largest group within the class of cysteine proteases. Members of this family conform a three-dimensional structure in the form of a hand that consists of a thumb, palm, and fingers subdomains. The catalytic active site, consisting of the catalytic triad Cys-Asp/Asn, and His, sit between the thumb and the palm subdomains, and the fingers subdomains provide stabilization for the interacting ubiquitin molecule. Figure 1.4 shows the three-dimensional structure of one of the most studied DUB from the family of USP, Ubiquitin-Specific Protease 7, also known as USP7. As seen in Figure 1.4, USP7 consists of the catalytic triad, Cys223, His464, Asp481, that mediates the cleavage of the isopeptide bond from the C-terminus of ubiquitin to release it from the substrate.

The catalytic mechanism of cysteine proteases involves the deprotonation of histidine by aspartic acid, which mediates the deprotonation of the cysteine inactive thiol group (Figure 1.5). Once deprotonation is performed, the thiol group is converted to an active thiolate and this mediates a nucleophilic attack on the acyl group of the isopeptide bond within the ubiquitin linkage. Then, the amide group of the isopeptide bond deprotonates the histidine residue, releasing the substrate from ubiquitin. In the final step, the ubiquitin intermediate is released from the nucleophilic cysteine residue via hydration reaction that reforms the ubiquitin monomer and the thiolate, thus releasing the substrate and ubiquitin (Snyder and Silva, 2021).
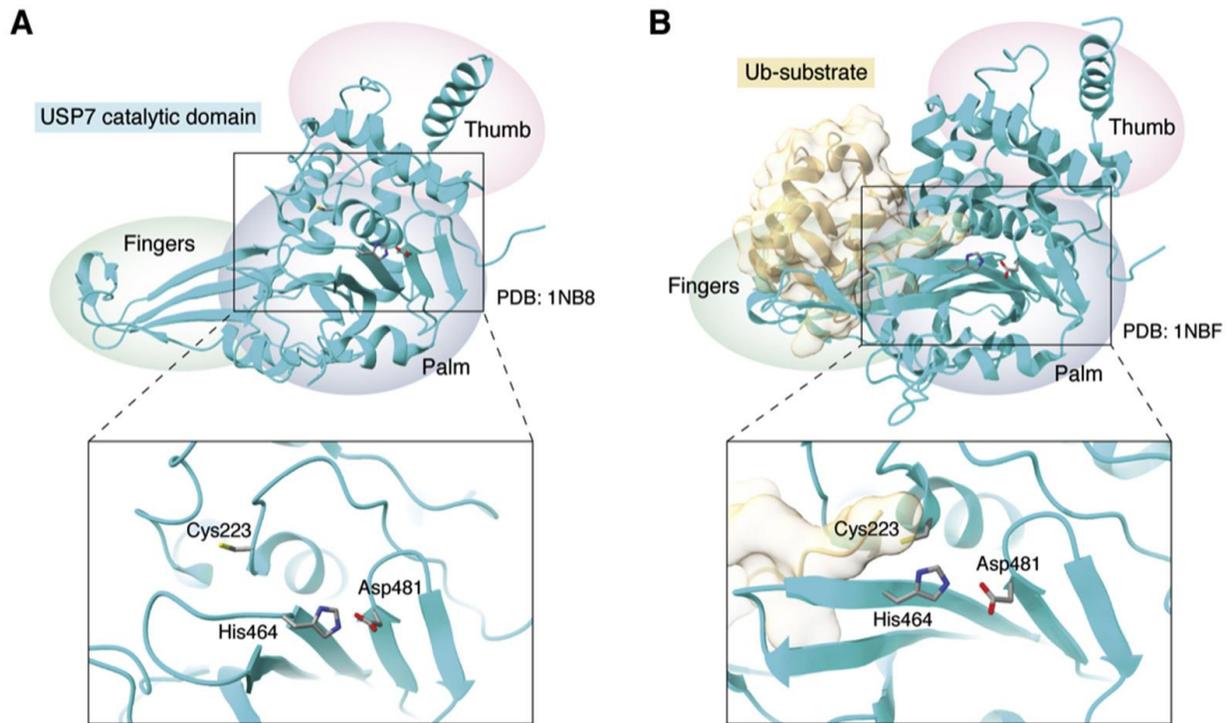
Figure 1.4. **Three-dimensional Structure of Ubiquitin and the Catalytic Site of Cysteine Proteases**. A) in the absence of ubiquitin, the catalytic cysteine of USP7 (cyan) is inactive and it is far from the other residues in the triad. However, upon ubiquitin binding (yellow)(B), a conformational change occurs, that brings the cysteine residue closer to histidine, which allows for its subsequent deprotonation and thus enable for the catalytic attack on the isopeptide bond of ubiquitin. *Adapted from* Snyder and Silva, 2021.
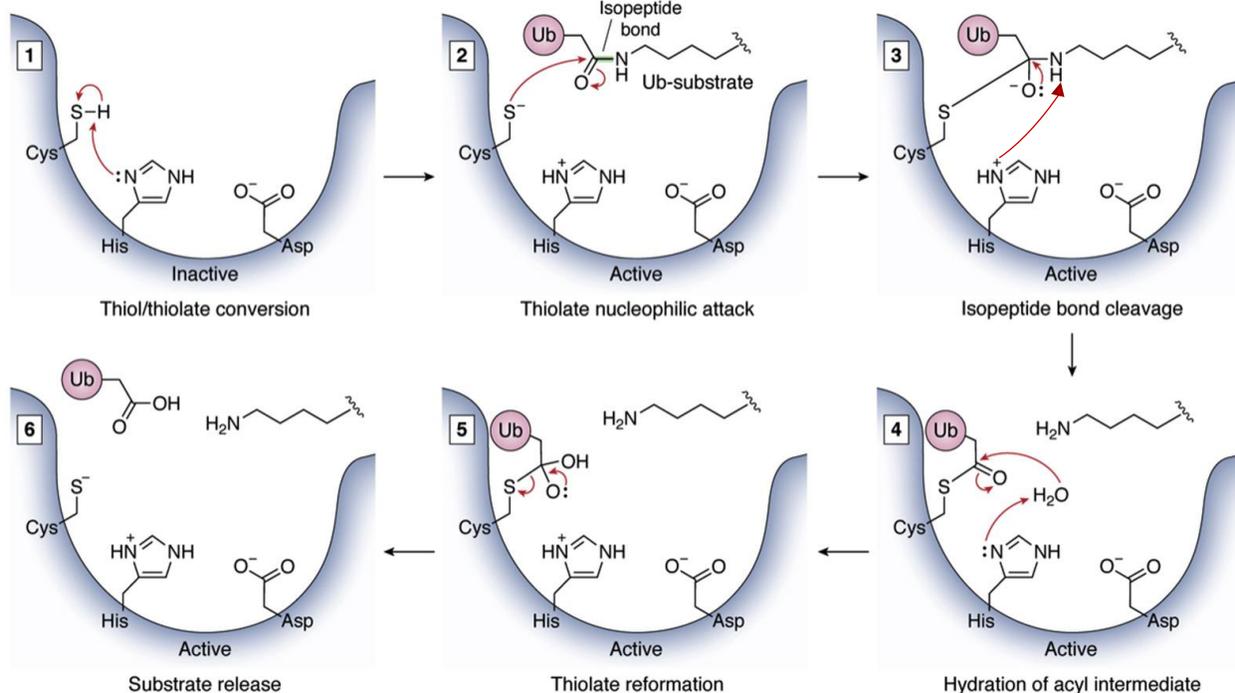
Figure 1.5. **The Mechanism of DUB Cysteine Proteases Catalytic Activity**. The catalytic mechanism of ubiquitin hydrolysis involves series of steps whereby the catalytic cysteine is primed and activated by histidine for a nucleophilic attack on the ubiquitin isopeptide bond. In subsequent reactions, ubiquitin-thiolate intermediate is formed, and the substrate is released from ubiquitin. In the final steps, hydration of the acyl intermediate takes place that enables the formation of ubiquitin monomer and thiolate group, releasing both the ubiquitin molecule and the substrate. *Retrieved from* Snyder and Silva, 2021.

1.2 USP7

USP7, also known as Herpes Virus Associated Protease (HAUSP) is a cysteine protease, belonging to the largest family of DUBs, the USP family. USP7 is located on chromosome 16p13.2 (Fountain *et al.,* 2019) and the protein product is made up of 1102 amino acids (135 kDa) (Pozhidaeva and Bezsonova, 2019); it is evolutionary conserved among eukaryotes. It has approximately 98.6% sequence homology to the mouse and rat proteins. Everett *et al.,* 1997, were the first to identify and characterize USP7 as a 135 kDa DUB. They discovered that a deubiquitinating enzyme was associated with a viral protein, Vmw110, also known as Infected

Cell Protein 0 (ICP0), from the Herpes Simplex Virus (HSV) type I, and this interaction enabled

the viral protein to simulate gene expression and viral lytic growth. Through various assays, they

showed that the protease was capable of successfully cleaving ubiquitin molecules from their

model substrate, enhancing its stability (Everett *et al.,* 1997).

USP7 functions to hydrolyze the ester, thioester, peptide and isopeptide bonds created by the

glycine residue, located in the carboxy-terminal of ubiquitin (Bhattacharya *et al.,* 2018). In

mouse models where USP7 was knocked out, embryonic lethality surfaced due to increased p53

levels, leading to cellular apoptosis and cell growth arrest. Subsequent deletion of p53 gene did

not extend the viability of the mice significantly, so this showed that USP7 had other protein

regulatory functions, outside of the p53-Mdm2 axis (Kon *et al.,* 2011). However, certain

pediatric patients have been shown to carry missense, nonsense and frameshift mutations

throughout the USP7 gene. Also, other individuals were presented with partial or complete

deletions of USP7. Such abnormalities resulted in a spectrum of neurodevelopmental disorders in

these children including speech impairments, autism spectrum disorder, and intellectual

disability (Fountain *et al.,* 2019). Thus, the expression and functional activity of USP7 are

essential for proper cellular functions, development, and regulation.

1.2.1 STRUCTURAL ORGANIZATION AND CONSERVED DOMAINS OF USP7

USP7 is organized into various domains that enable substrate interaction, and the catalytic

function of the protein. The N-terminus of the protein contains the Tumor necrosis factor

receptor-associated factor (TRAF), also known as the Meprin and TRAF Homology (MATH)

domain. This domain is one of the sites for protein-protein interaction. USP7 also consists of a

catalytic domain, which mediates ubiquitin interaction and cleavage. Finally, 5 ubiquitin-like

folds (Ubls) are present at the C-terminus of the protein and this region also contains a second site for protein-protein interaction (Figure 1.6 A).

The active site of the catalytic domain of USP7 consists of the conserved catalytic triad, Cysteine (Cys223), Histidine (His464), and Aspartate (Asp481). As with the USP family, the active site confers a conformation of a right hand, consisting of a thumb, palm and fingers subdomains. The thumb and the palm regions house the catalytic triad, whereas the fingers subdomain confers stability for ubiquitin binding (Pozhidaeva and Bezsonova, 2019). In the absence of ubiquitin, the catalytic cysteine is positioned far from the deprotonating histidine, rendering the catalytic site inactive. However, upon the binding of ubiquitin, a conformational change occurs in the active site, which positions the deprotonating histidine in close proximity to the catalytic cysteine (Kim and Sixma 2017). As a result, a nucleophilic attack can initiate the process of deubiquitination (Figure. 1.5). Furthermore, for efficient catalytic activity by the catalytic active site, the partial binding of CTD, more specifically Ubl45, to the catalytic region is needed. The Ubl45 binds the switching loop, W285, of the catalytic domain, close to the binding site of the C-terminus of ubiquitin. The binding of Ubl45 allows for the ubiquitin tail to get closer to the active site of the catalytic domain, stabilizing and enhancing the binding of ubiquitin to the active site by 100-fold (Faesen *et al.,* 2011). This ensures that USP7 functions at its full capacity for ubiquitin cleavage. Furthermore, our group has shown that the 5 Ubls are connected to the catalytic domain via a 26 amino acid alpha helical linker found between Ubl1 and the catalytic domain (Figure 1.6 B) (Pfoh *et al.,* 2015). In addition, the Ubls themselves are also connected via "hinges" that link Ubl12 to Ubl3, and Ubl3 to Ubl45. Thus, the overall arrangement of the CTD is organized in a 2-1-2 fashion, whereby Ubl 1 & 2 form a dimer, that connects to Ubl3 via hinges, which connects to the second dimer Ubl45 (Faesen *et al.,* 2011). The presence of hinges

and the alpha helical linker enable for the flexibility, movement, and folding back of the Ubls to result in an active conformation of USP7 for binding ubiquitin (Pfoh *et al.,* 2015; Kim and Sixma 2017).

For USP7 to deubiquitinate substrates, it must first recruit the substrates for interaction. One of the sites that USP7 uses to interact with the target substrate is through the TRAF-like/ MATH domain. The MATH domain, residues 63-205, is located at the N-terminus of the protein (Figure 1.6 A). It confers the structure of an eight-stranded, antiparallel β-sandwich folds (Pozhidaeva and Bezsonova, 2019). The MATH domain consists of a ^{164}DWGF^{167} motif, that recognizes and binds target substrate. The substrate consensus sequence that is recognized by the motif is P/A/ExxS. This consensus sequence is recognized by the shallow groove of the MATH domain, whereby the residues, D164 and W165 are critical for anchoring the binding motif of a substrate (Pozhidaeva and Bezsonova, 2019). The first protein to have been known to utilize this domain for binding was the viral protein Epstein Barr Nuclear antigen 1 (EBNA1) (Holowaty *et al.,* 2003). Later on, other cellular proteins have also been discovered to interact with USP7 utilizing the MATH domain. Such proteins include p53, and Mouse Double Minute 2 homolog (MDM2), as well as other viral proteins such as Viral Interferon Regulatory Factor 1 (vIRF1), that all interact at the same DWGF motif as determined by protein crystallography (Sheng *et al.,* 2006; Chavoshi *et al.,* 2016). Finally, mutations to the residues D164 and W165 have been shown to compromise substrate-binding. A mutational study conducted by Sheng *et al.,* 2006 showed that mutation of the residue W165 to Alanine completely abrogated binding to p53, EBNA1, and MDM2 GST fusion proteins. Furthermore, mutation of the residue D164 to Alanine decreased the binding to p53, MDM2, and EBNA1. Thus, the conservation of these critical residues is essential for mediating substrate interaction.

USP7 also contains a second site for substrate interaction. This is located at Ubl2 in the C-terminal domain of the protein. ICP0 was the first substrate identified to interact with the CTD. Several structural analyses of the interaction have shown that positively charged regions of ICP0, ^{618}PRKCARKT^{625} interacted with negatively charged regions found within Ubl2. This "negatively charged region" was determined to be the ^{758}DELMDGD^{764} motif, that recognizes substrates with the consensus sequence KxxxK or KxxxKxK, where K refers to lysine residues and x being any amino acid. The positively charged lysine residues found in the substrate consensus motif form direct salt bridges with the aspartic acid residues of the acidic MDGD motif found in Ubl2 (Pfoh *et al.,* 2015). Furthermore, other substrates, such as Ubiquitin-like with PHD And Ring Finger Domains 1 (UHRF1), DNA Methyl Transferase 1 (DNMT1), and Enhancer of Zeste Homolog 2 (EZH2), have all been shown to interact within that same region (Felle *et al.,* 2011; Cheng *et al.,* 2015; Gagarina *et al.,* 2020). Finally, mutating the aspartic residues in the MDGD motif to either alanine or arginine disrupted binding to ICP0 and UHRF1 peptides (Pfoh *et al.,* 2015). Thus, the aspartic residues are key for mediating substrate interaction with the CTD of USP7.

1.3 CELLULAR ROLES OF USP7

USP7 plays a very important role in the cell. It has been shown to regulate many cellular pathways in the cell because of its deubiquitinating activity on various target substrates. Figure 1.7 summarizes the pathways and processes that USP7 regulates. These include DNA damage response and repair, DNA replication, epigenetic regulations, and transcription factors, which will be discussed in greater detail in the next sections (Al-Eidan *et al.,* 2020).
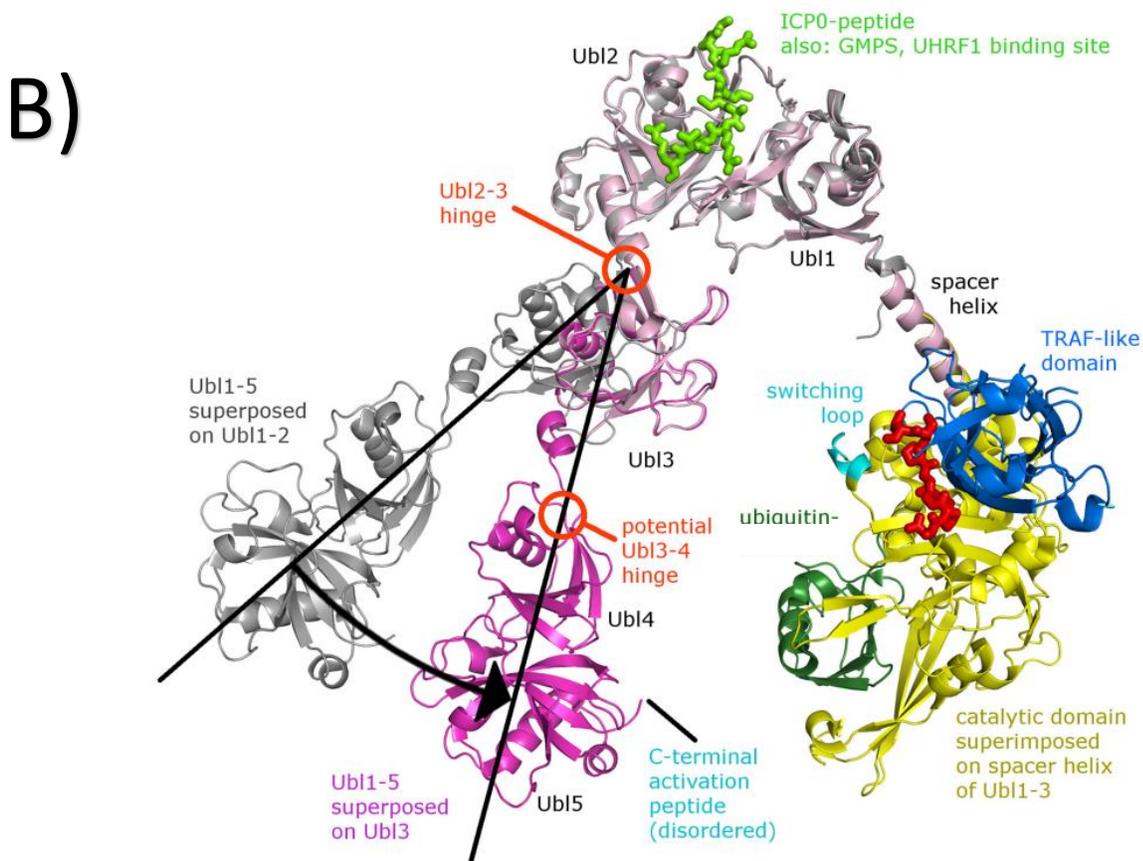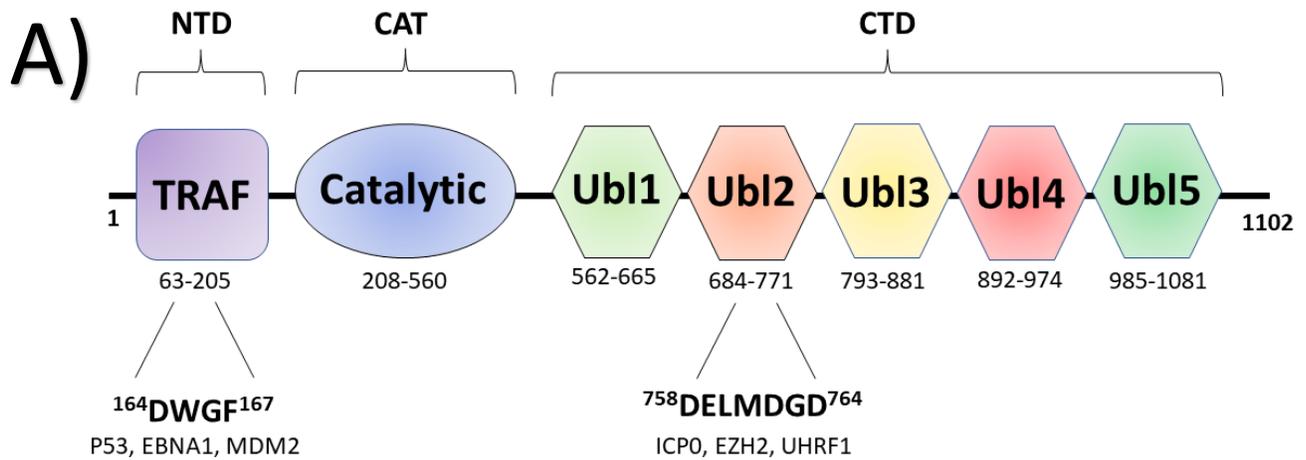
Figure 1.6. **Structural Organization of USP7**. A) USP7 is composed of a TRAF-like domain located at the N-terminus, a catalytic domain responsible for ubiquitin cleavage, and 5 ubiquitin-like folds (Ubls) located at the C-terminus. The TRAF-like domain contains the ^{164}DWGF^{167} motif for binding to substrates. The Ubl2 also contains a second site for substrate interaction with the motif ^{758}DELMDGD^{764}. B) Individual three-dimensional domains of USP7 such as NTD and catalytic domain (blue and yellow respectively, PBD 2F1Z) have been superposed to the spacer helix of Ubl123, and to the CTD (Magenta, PDB 2YLM) to generate a three-dimensional model of full length USP7. The flexible regions within the CTD,allow for the movement of Ubl45 for efficient USP7 activity. *Adapted from* Pfoh *et al.*, 2015.
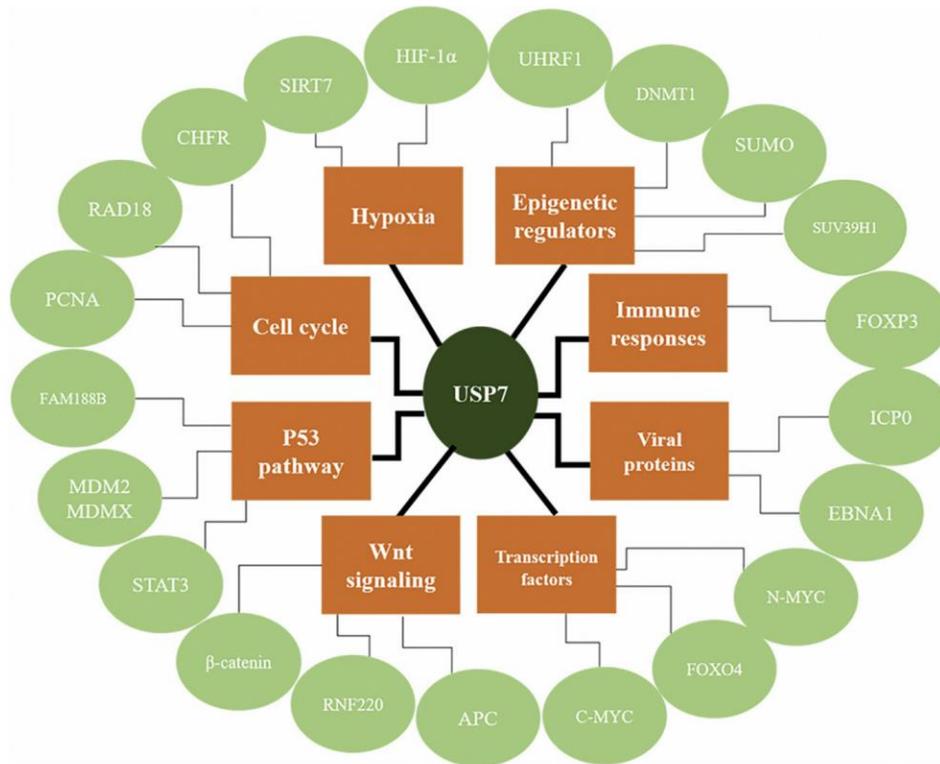
Figure 1.7. **Cellular Roles of USP7**. USP7 acts on and regulates the stability of various cellular and viral proteins. Such proteins are involved in different pathways such as immune response, DNA damage response, making USP7 a regulator of various important pathways. *Retrieved from* Al-Eidan *et al.,* 2020.

1.3.1 USP7 AND DNA DAMAGE RESPONSE

One of the well studied axes for DNA damage response is the USP7-Hdm2-p53 axis. USP7 has been shown to deubiquitinate both substrates, Hdm2 and p53 (Brooks *et al.,* 2007). Hdm2 is an E3 ligase that functions to add ubiquitin tags onto protein substrates. This signals the protein for degradation by the 26S proteasome complex. One of the substrates of Hdm2 is p53, a tumour suppressor protein, dubbed the guardian of the genome, that is critical in determining cell fate in the event of DNA damage (Pozhidaeva and Bezsonova, 2019). In the absence of DNA damage, USP7 is phosphorylated at Ser18 by the kinase CK2 (Khoronenkova *et al.,* 2012). This increases its affinity for Hdm2, causing it to preferentially interact with and deubiquitinate Hdm2. The

15

degradation of Hdm2 by the 26S proteosome complex is thus inhibited, stabilizing the levels of the protein in the cell. This allows it to carry out ubiquitination on p53, marking it for degradation by the 26S proteosome complex, and decreasing the levels of p53 in the cell (Brooks *et al.,* 2007). As a result, this prevents an unnecessary DNA damage response (DDR) in the absence of DNA damage. However, in the event of DNA damage, caused by ionizing radiation for instance, ATM-dependent dephosphorylation of Ser18 of USP7 takes place by the phosphatase Protein phosphate 1G (PPM1G) (Khoronenkova *et al.,* 2012). This dephosphorylation event decreases the affinity of USP7 for Hdm2. As a result, the levels of Hdm2 are not stabilized in the cell due to the decrease in the deubiquitination activity of USP7. Rather, USP7 deubiquitinates p53 and prevent its degradation by the 26S proteosome complex, increasing its levels in the cell and allowing it to mediate its function as a response to the DNA damage detected (Pozhidaeva and Bezsonova, 2019).

USP7 also regulates the pathway of Nucleotide-Excision Repair (NER) during DNA damage induced by UV irradiation. UV irradiation results in helix distorting DNA lesions. Xeroderma pigmentosum complementation group C (XPC) protein recognizes these DNA lesions and binds them to initiate NER. However, in the early stages of UV-light induced lesions, XPC is ubiquitinated, leading to its degradation. He *et al.,* 2014 have shown that USP7 interacts with and deubiquitinates XPC, preventing its degradation upon UV irradiation. Finally, USP7 also regulates vital proteins involved in the repair mechanism of Double Stranded Breaks (DSBs), such as RNF168. RNF168 is an E3 ligase that is recruited to the site of DNA damage, marked by phosphorylated histone H2AX (γH2AX) and H2A. RNF168, with the aid of RNF8, ubiquitinate γH2AX and H2A, allowing for the recruitment of important repair proteins such as Breast

Cancer Type 1 Susceptibility Protein (BRCA1) to the site of DNA damage to produce the proper

DDR (Zhu *et al.,* 2015).

1.3.2 USP7 AND DNA REPLICATION

Not only is USP7 involved in regulating DNA damage repair pathways, but it has also been

shown to play a role in DNA replication in mitosis. DNA replication requires the assembly of the

pre-replicative complex (pre-RC) at replication origin in the G1 phase of the cell cycle, loading

of replication factors at the origin and the activation of DNA synthesis at the start of S-phase

(Jagannathan *et al.,* 2014). The pre-RC consists of the Minichromosome Maintenance Complex

(MCM), a hexameric complex consisting of MCMs 2-7 (MCM2-7). This complex promotes the

progression of the replication forks by its DNA helicase activity. The MCM complex also

interacts with MCM Binding Protein (MCM-BP), a highly conserved protein, shown to unload

inactive MCM complex at replicative origins throughout S-phase. MCM-BP has been identified

to interact with USP7, and both proteins have been suggested to work together to efficiently

unload the MCM complex from the DNA at the late S/G2 progression (Jagannathan *et al.,* 2014).

Furthermore, USP7 plays a role in maintaining chromatin in a SUMOlyation rich, but

ubiquitination poor environment, around replisomes during DNA replication. USP7 is enriched

at replisome sites and acts as a SUMO deubiquitinase, whereby it deubiquitinates SUMO and

SUMOlayted proteins, to enable the progression of replication forks (Lecona *et al.,* 2017).

1.3.3 USP7 AND EPIGENETIC REGULATION

USP7 has been shown to regulate the proteins responsible for "reading" and "writing" chromatin

epigenetic markers, altering the expression of target genes (Pozhidaeva and Bezsonova, 2019).

For instance, USP7 has been shown to interact with replisome components, UHRF1 and

DNMT1, which produce tissue specific DNA methylation. UHRF1 is an E3 ligase and it ubiquitinates DNMT1, leading to its degradation. USP7 has been shown to stabilize the levels of DNMT1 by removing the ubiquitin markers and preventing protein degradation (Qin *et al.,* 2011), stimulating the maintenance and de novo DNA methylation by DNMT1 (Felle *et al.,* 2011). Furthermore, USP7 also deubiquitinates UHRF1 and regulate its stability (Felle *et al.,* 2011). Additionally, USP7 can directly deubiquitinate Histone H3. Histone H3 is ubiquitinated by UHRF1 and this ubiquitination event triggers the recruitment of DNMT1 to the site to result in DNA methylation. USP7 has been shown to form a complex with DNMT1 that allows for the recruitment and loading of USP7 at the ubiquitinated sites on histone H3, mediating deubiquitination of histone H3 (Yamaguchi *et al.,* 2017). Research conducted by Li *et al.,* 2020 showed that this deubiquitination event reduces the DNA methylation levels mediated by DNMT1. Thus, they suggested that USP7 acts as a negative regulator of DNA methylation by attenuating histone H3 ubiquitination-dependent recruitment of DNMT1. Lastly, other histone associated proteins are also stabilized and interact with USP7 such as histone 3 lysine 4 (H3K4) methyltransferase Mixed lineage leukemia 5 (MLL5) (Ding *et al.,* 2015), lysine-specific demethylase 1 (LSD1/KDM1) that remove mono and di-methyl marks from H3K4 (Yi *et al.,* 2016), and acetyltransferase TIP60, responsible for adding acetyl marks to histones and non-histone proteins such as p53 (Dar *et al.,* 2013).

Finally, components of the Polycomb Repressive Complexes (PRCs) were shown to be substrates of USP7. PRC is responsible for meditating gene silencing by ubiquitination of histone H2A (De Bie *et al.,* 2010). EZH2 from PRC2 methylates lysine 27 on histone H3 (H3K27) (Su *et al.,* 2021). This methylation recruits the PCR1 complex, which contains an E3 ligase, Ring Finger Protein 1B (RING1B), that monoubiquitinates histone H2A on lysine 119 (De Bie *et al.,*

2010), resulting in chromatin compaction and gene silencing (Su *et al.,* 2021). EZH2, RING1B, and other PRCs components such as Bmi1 are substrates of USP7 (Gagarina *et al.,* 2020; De Bie *et al.,* 2010; Maertens *et al.,* 2010). Therefore, USP7 maintains the stability of PRC components, in turn playing a role in regulating gene expression.

1.3.4 USP7 AND TRANSCRIPTION FACTORS

USP7 governs the integrity of various transcription factors to regulate their specific gene expression products. For instance, USP7, with the help of the bridging activity of RING finger E3 ligase RNF220, deubiquitinates β-catenin and stabilizes it. This increased protein stability of β-catenin enhances the canonical Wnt-signaling pathway as β-catenin can stimulate the expression of Wnt target genes such as c-Myc and cyclin D1 (Ma *et al.,* 2014). Additionally, the transcriptional activity of Forkhead box O proteins, FOXO3 and FOXO4 are also impacted by USP7 mediated deubiquitination (Van der Horst *et al.,* 2006). The transcription factors FOXO3 and FOXO4 are important regulators of cell cycle progression, metabolism, and cell death. During oxidative stress, FOXO3/44 are monoubiquitinated and this ubiquitination event induces nuclear localization, resulting in increased transcriptional activity. However, USP7 removes monoubiquitin tag on FOXO3 and FOXO4, resulting in their re-localization to the cytosol, inhibiting their transcriptional activities (Van der Horst *et al.,* 2006). Finally, other transcription factors such as N-Myc are also regulated by USP7. N-Myc is a proto-oncogene which is amplified in numerous end stage tumours and drives the tumorigenesis of diseases, such as neuroblastoma (Tavana *et al.,* 2016). The stability and deubiquitination of N-Myc is driven by USP7. Thus, Tavana *et al,* 2016 showed that increased deubiquitination of N-Myc prevented protein degradation, stabilized the protein levels, and increased its transcriptional activity for tumorigenesis.

Overall, previous studies have portrayed that USP7 plays an important role in regulating gene expression by acting on specific transcription factors to either enhance or supress the transcription of target genes. However, little is known about the role of USP7 when it comes to the Mediator complex, which is a known co-activator of gene expression.

1.4 GENE EXPRESSION AND THE MEDIATOR COMPLEX

Gene transcription is a crucial step for expressing a gene of interest. Through transcription, an mRNA product is formed, allowing for its subsequent translation to result in a protein product (Soutourina, 2018). The transcription of genes occurs via the binding of RNA Polymerase II (RNAPII), general transcription factors (GTF), and specific transcription factors (TF) to their designated regions on the DNA such as the promoter or regulatory regions like enhancer motifs. The activity of the TFs is further regulated by binding of a multiprotein complex known as the Mediator complex. Thus, the Mediator complex plays a functional role in transcription initiation, elongation, and chromatin architecture by facilitating enhancer–promoter gene looping (Allen and Taatjes, 2015).

The Mediator complex is an evolutionary conserved multi- protein complex that is composed of 25 protein subunits in yeast and up to 30 subunits in humans (Figure 1.8) (Soutourina, 2018). The Mediator complex is composed of two compartments: the core mediator, and the CDK8 kinase module. The core mediator consists of the head, middle, and tail modules. The Mediator complex subunit MED14, acts a scaffold that holds the three modules, head, middle, and tail, together. The subunits in the head and middle modules form contacts with the RNAP II, and GTFs, while the tail module forms contact with specific TF. The kinase module reversibly associates with the core mediator and can pose as positive or negative regulator of transcription (Jeronimo and Robert, 2017). Thus, the Mediator complex functions as a molecular bridge

between the DNA binding transcription factors and the RNAP II to regulate the transcription

levels of target genes (Soutourina, 2018).
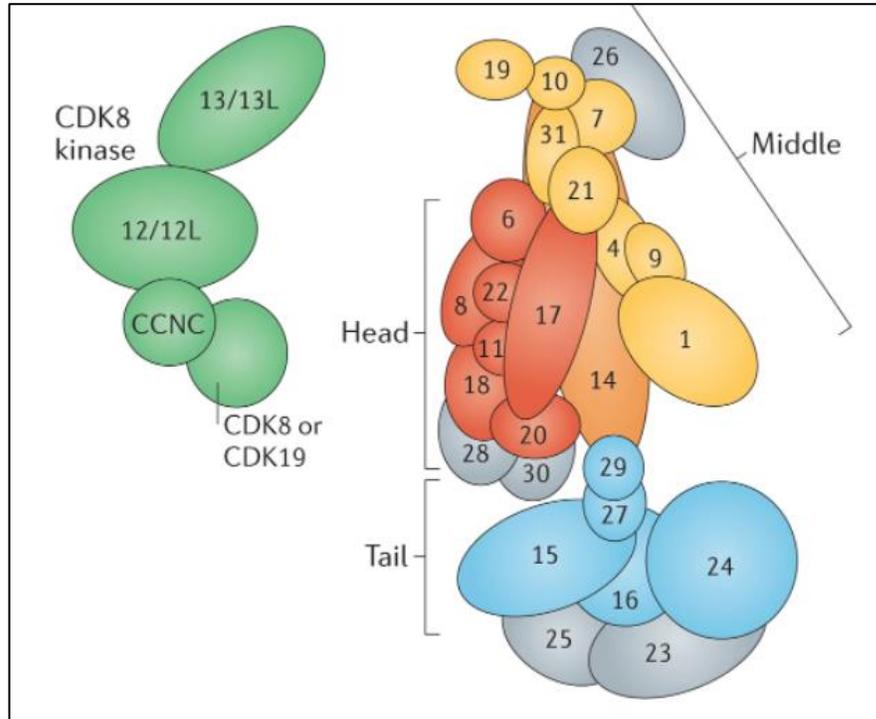
Figure 1.8. **Composition of the Mediator complex**. The Mediator complex is a multiprotein complex that consists of 30 proteins. It is composed of two modules: the core mediator, and the CDK8 kinase module. The core mediator interacts with DNA regulatory elements and transcription factors to facilitate the transcription of genes. The CDK8 kinase module functions to regulate the activity of the core mediator. *Adapted from* Soutourina, 2018.

1.4.1 TRANSCRIPTION INTIATION AND PIC ASSEMBLY

The Mediator is necessary for transcriptional regulation. It is recruited to enhancer regions or

upstream activating sequences (UAS) in yeast, by TFs forming contacts with subunits from the

tail module (Figure 1.9). Furthermore, recruitment of the Mediator to enhancer regions can also

be TF independent. For instance, a class of noncoding RNA (ncRNA), ncRNA-activating (nc-

RNA-a), has been shown to activate neighboring genes in cis. ncRNA can bind to Mediator and

recruit it to the ncRNA-a locus, allowing for the looping of chromatin towards its target gene

promotor (Jeronimo and Robert, 2017). When the Mediator is recruited to the enhancer elements, the kinase module is still attached to the core mediator. The CDK8 module can antagonize the recruitment and retention of mediator to enhancer regions by potential phosphorylation events of TFs responsible for Mediator recruitment or Mediator subunits themselves. Therefore, the kinase module provides a mechanism to fine tune gene expression. Furthermore, when the CDK8 kinase module is associated with the core mediator, it causes steric hinderance and blocks the core mediator from interacting with the RNAP II at the preinitiation complex (Jeronimo and Robert, 2017). Thus, gene transcription is turned off at that locus. However, when the kinase module is dissociated from the core mediator via phosphorylation events initiated by CDK8 (Osman *et al.,* 2021), this allows for the core to form and stabilize the preinitiation complexes (PIC) at core promoters. The TF do not directly bind to RNAP II, rather, they recruit RNAP II to the core promoter by recruiting the Mediator and having it bind to RNAP II. The Mediator complex stabilizes the formation of the PIC by contacting RNAPII at its CTD, and GTFs such as TFIIB, TFIIE, TFIIF, TFIIH, and TFIIS (Allen and Taatjes, 2015). Thus, the interaction of the Mediator with enhancer elements and core promoters allows for the enhancer-promoter looping to bring distal elements and TFs closer together for more efficient transcription, hence, acting as a molecular bridge (Figure 1.9 B) (Soutourina, 2018). Mutations in Mediator subunits such as MED10, altered the interaction between MED14 and TFIIB, impacting the assembly of TFIIB, TFIIE, TFIIH, and RNAP II at the promotors (Eychenne *et al.,* 2016). The interaction of Mediator with RNAP II causes a conformational change to the Mediator, placing the CTD domain of RNAP II in between the head, neck, and the middle knob of the complex. This positions the RNAP II near the kinase, CDK7 in the TFIIH complex, which stimulates the TFIIH kinase activity, resulting in the phosphorylation of Ser5on the CTD of RNAP II. This

phosphorylation event dissociates the Mediator-RNAP II complex, allowing for RNAP II promoter escape, to initiate elongation (Figure 1.9 C) (Jeronimo and Robert, 2017).

1.4.2 RNAP II PAUSING AND ELONGATION

Post initiation, RNAP II pauses 30-60 nucleotides into the process of elongation due to Negative Elongation Factor (NELF) and DRB sensitivity-inducing factor (DSIF). The release of such pausing occurs by phosphorylation of RNAP II CTD on Ser2, DSIF, and NELF, which is mediated by Positive Transcription Elongation Factor b (P-TEFb), which contains CDK9 in complex with cyclin T, or P-TEFb-containing super elongation complex (SEC) (Figure 1.9 D) (Chou *et al.,* 2020). The subunit MED26 of the Mediator is responsible for recruiting the SEC to the paused RNAP II to allow for the release of the polymerase. Furthermore, the pause release is also dependent on the kinase module, whereby CDK8 recruits other forms of P-TEFb (Figure 1.9 C). Thus, the recruitment of multiple forms of P-TEFb via different mechanisms cooperate to fully release the pausing of the RNAP II (Jeronimo and Robert, 2017). The escaped RNAP II can continue elongation by recruiting elongation factors.
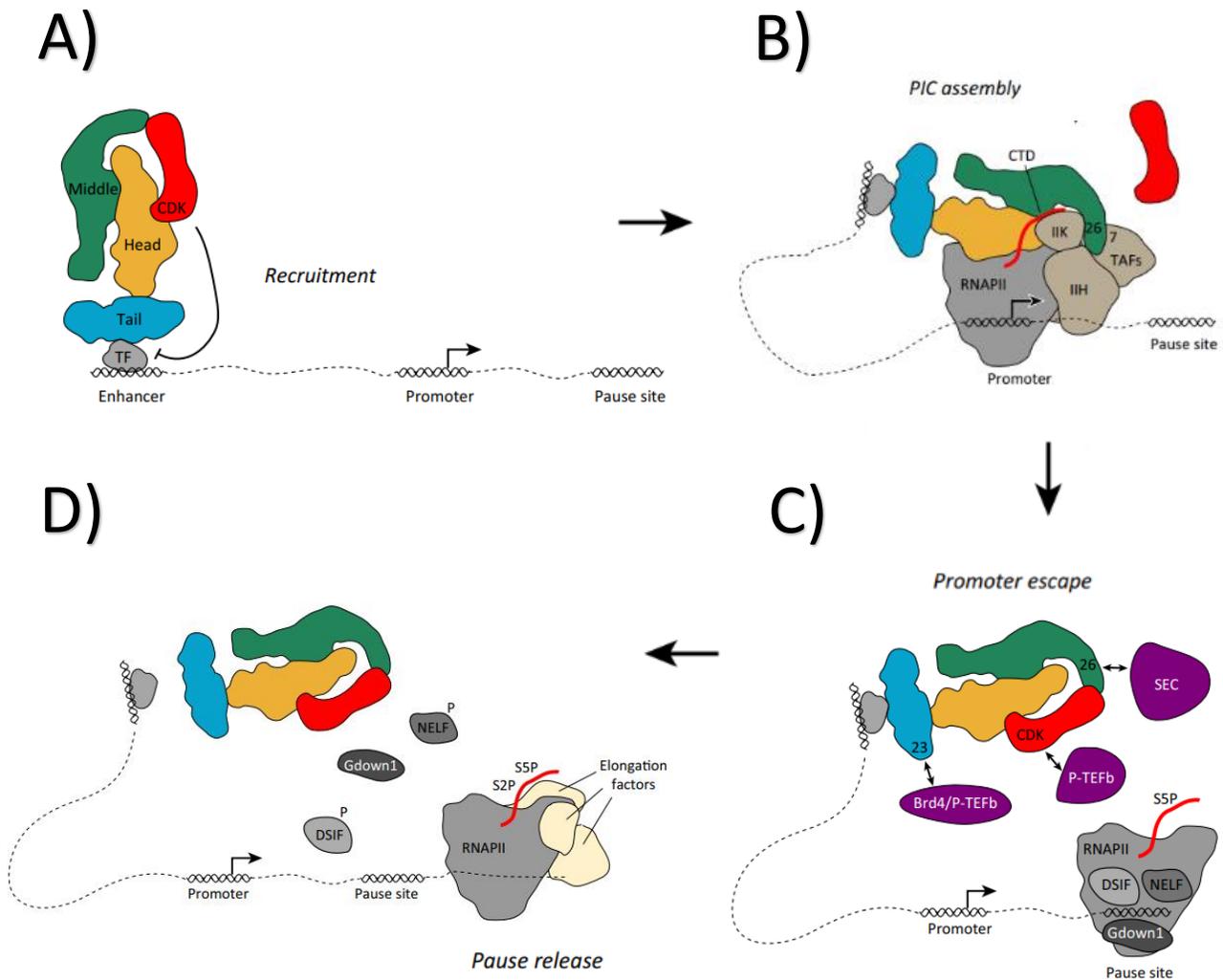
Figure 1.9. **The Functional Role of Mediator in Transcription Initiation.** A) The Mediator is recruited to enhancer elements by TF via binding to the tail module (blue). The kinase module (red) can negatively regulate the recruitment of the mediator by phosphorylating target TFs or Mediator subunits. B) The kinase module is ejected from the Mediator complex, allowing for the core mediator to loop the chromatin towards the core promotor, and binding to RNAP II and other TFs. C) Phosphorylation of the RNAPII CTD at Ser5 is mediated by TFIIH, in which the activity is stimulated by the core mediator. This phosphorylation event allows for promoter escape. D) The RNAP II pauses at pause site due to DSIF and NELF factors, but the pausing is released by the Mediator-dependent recruitment of P-TEFb and SEC factors that phosphorylate NELF and DSIF. This allows for subsequent progression of transcription elongation via the recruitment of elongation factors by RNAP II. *Adapted from* Jeronimo and Robert, 2017.

1.5 THE CDK8 KINASE MODULE

The CDK8 kinase module consists of Cyclin Dependent Kinase 8 (CDK8), cyclin C, MED12 and MED13 (Galbraith *et al.,* 2010). As previously mentioned, the kinase module is linked to the core mediator via an interaction between the MED13 subunit and a hook domain found in the middle module of the core mediator (Jeronimo and Robert, 2017). The reversible interaction of the kinase module with the core mediator fine tunes the activity of the core mediator and regulates the levels of transcription. Components of the kinase module such as CDK8 and MED12 contain paralogs, including CDK19 and MED12L respectively. These paralogs are expressed on different chromosomes from CDK8 and MED12 and their expression and function are tissue specific. Furthermore, their function is not redundant, but rather mutually exclusive in the kinase module. Cyclin C is needed for the function of CDK8, and together, they form a dimer conformation (Poss *et al.,* 2013). Finally, MED12 is responsible for bridging CDK8-cyclin C dimer with MED13 and the core mediator (Wu *et al.,* 2021). Of interest to my project are CDK8 and MED12 from the CDK8 kinase module.

1.5.1 CYCLIN DEPENDENT KINASE 8

CDK8 is a 53 kDa (464 amino acids) serine/threonine protein kinase, characterized as a transcriptional CDK rather than cell cycle affiliated CDK (Xi *et al.,* 2018; Poss *et al.,* 2013). CDK8 primarily consists of a conserved kinase domain (residues 21-335), responsible for its catalytic activity. The catalytic core consists of adenosine triphosphate (ATP)-binding pocket, a conserved PSTAIRE-like cyclin-binding domain for cyclin C binding, and an activating T-loop motif that regulates the active status of the kinase. When cyclin C is not bound to CDK8, the T-loop closes the catalytic cleft, rendering CDK8 inactive (Chou *et al.,* 2020). As mentioned previously, CDK8 is the predominant kinase in the kinase module and it fine tunes the activity of

the core mediator to regulate transcription. By forming a complex with the core mediator, the kinase module inhibits the core mediator's association with RNAP II, preventing transcription. Early yeast studies have suggested that the CDK8 homolog, SRB10, is a negative regulator of transcription due to its phosphorylation of RNAP II CTD, prior to the assembly of PIC (Hengartner *et al.,* 1998). Human CDK8 also inhibits transcription by phosphorylating cyclin H subunit in TFIIH (Akoulitchey *et al.,* 2000). However, other studies have shown that CDK8 also serves positive roles in transcription. CDK8 is critical in transcription elongation as it facilitates the release of RNAP II pausing by recruiting SEC factors for downstream phosphorylation events (Chou *et al.,* 2020). Thus, the function of CDK8 has been controversial and depending on the context of the disease or transcriptional situation, it can act as either a positive or negative regulator. Early studies in mice have suggested an essential role of CDK8 in early embryonic development. CDK8 knockout, CDK8 $^{-/-}$, in mice produced a phenotype of embryonic lethality on days 2.5 to 3 of development because of preimplantation defect. However, no deleterious effects were seen in adult mice with regards to development in CDK8 $^{-/-}$ (Westerling *et al.,* 2007). The embryonic lethality seen in CDK8$^{-/-}$ could be due to CDK8's role in regulating TFs essential for fetal development.

1.5.2 CDK8 AND REGULATION OF TRANSCRIPTION FACTORS

Depending on the cellular context and mechanism of action, CDK8 can act as a positive or negative regulator of transcription (Poss *et al.,* 2013). CDK8 has been shown to phosphorylate various substrates from the Mediator complex such as MED13 and CTD of RNAP II at Ser5 (Galbraith *et al.,* 2010). However, CDK8 also possesses functions independent from the Mediator complex. CDK8 can for instance regulate the expression of metabolic genes in response to changes in nutrients. Hirst *et al., (*1999) have shown that CDK8 phosphorylates the

transcription factor, GAL4, for efficient galactose-inducible transcription and metabolism in yeast. Furthermore, metabolic regulation of metazoans, such as in mouse and human cells, is also mediated by the CDK8 dependent phosphorylation of TFs to regulate their activity and subsequently impact their stability. For example, CDK8 activates the interferon or the notch signalling pathway by phosphorylating STAT1/3/5 (Bancerek *et al.,* 2013) or the notch receptor intracellular domain (ICD) (Fryer *et al.,* 2004). Likewise, CDK8 phosphorylates SMAD1/3 (Alarcon *et al.,* 2009) to allow for the activation of the bone morphogenic pathway (BMP) and transforming growth factor beta (TGFβ). In all three cases, the phosphorylation event activates the functions of the proteins, and subsequently alters their stability due to their ensuing degradation (Luyties and Taatjes, 2022). Furthermore, the CDK8 kinase function is critical for maintaining pluripotency and maintenance of cell state. CDK8 kinase inhibition studies in acute myeloid leukemia (AML) cells triggered cellular apoptosis due to the stimulation of differentiation programs as a result of disruption in STAT1 function (Nitulescu *et al.,* 2017). Human tumour cell pluripotency in mouse xenografts models was also shown to be dependent on CDK8. Knockdown of CDK8 via shRNA caused tumour cells differentiation (Adler *et al.,* 2012).

CDK8 can also have inhibitory effects on TF by repressing their activity. For example, in the Wnt/β-catenin pathway, E2F1 TF has various E2F1 DNA-binding sites across the genome and it plays a negative regulatory role in the Wnt/ β-catenin pathway (Zhao *et al.,* 2013). E2F1 acts as a β-catenin repressor by binding to β-catenin, interfering with its association with T-cell factor/lymphoid enhancer factor TCF/LEF, and inhibiting the formation of β-catenin/TCF mediated transcription. CDK8 phosphorylates E2F1 at Ser375 and represses its inhibitory effects on β-catenin, allowing for β-catenin to induce the expression of Wnt target genes by interacting

with the TCF/LEF transcription factors (Zhao *et al.,* 2013). β-catenin mediated gene expression is also positively regulated by the Mediator complex as β-catenin directly binds with MED12 in the kinase module and the Mediator can thus recruit RNAP II and initiate transcription of Wnt target genes (Zhao *et al.,* 2013).

1.5.3 CDK8 AND CANCER

CDK8 mis-regulation and overexpression has been implicated in various cancer types such as colorectal cancer, breast cancer, and pancreatic cancer. For instance, in 123 cases of colorectal cancer (CRC), CDK8 was amplified in 47% of the tumour samples (Firestein *et al.,* 2008). A negative correlation was also examined between the survival of CRC patients and the amplified expression of CDK8. Furthermore, at later stages of CRC, such as stages III and IV, CDK8 expression was further increased, suggesting a contribution of CDK8 in the progression from colorectal adenoma to carcinoma. In these tumour samples, an increased expression of β-catenin was present, and thus its transcriptional activity was also increased, positively corelating with the increased CDK8 expression (Seo *et al.,* 2010). This could be attributed to the direct and indirect regulation of β-catenin by CDK8 kinase module. As mentioned earlier, β-catenin interacts with MED12, allowing for the recruitment of Mediator complex to Wnt promoter to facilitate transcription of Wnt target genes. CDK8 also indirectly regulates β-catenin functioning by phosphorylating E2F1 and preventing its inhibitory grip on β-catenin. As a result, Wnt target genes such as c-Myc, and Axin2 are expressed, allowing for the increased proliferation and survival of the cancer cells (Wu *et al.,* 2021). However, *in vivo* mouse model of conditional CDK8 knockout showed CDK8 as a tumour suppressor. Young adult mice with CDK8 deletion in the background of APC Min mutations produced large intestinal tumours and increased rate of growth compared to mice with only the APC Min mutation. The life span of the mice was also

shortened in the CDK8 knockout mice (McCleland *et al.,* 2015). Thus, CDK8 can serve as both an oncogene or tumour suppressor depending on the disease and tissue type in question.

The expression of CDK8 and its homolog CDK19, are also increased in breast cancer. In breast cancer, the levels of MacroH2A1 (mH2A1) are reduced due to Skp2 mediated ubiquitination and degradation of mH2A1. Such reduced mH2A1 levels coincided with increased expression of CDK8 protein, promoting G2/M progression, polyploidy, and poor survival outcome for breast cancer patients (Xu *et al.,* 2015). Additionally, in Estrogen Receptor (ER) positive breast cancer, CDK8 acts as an oncogene that facilitates estrogen-induced transcription. Upon binding of estrogen to ER and signaling a cascade of events, CDK8 acts downstream of ER and binds to ER responsive Growth Regulating Estrogen Receptor Binding 1 (GREB1) gene promoter, enabling gene expression (McDermott *et al.,* 2017).

Overexpression of CDK8 in pancreatic cancer promotes the cancer's invasiveness and triggers the epithelial to mesenchymal transition (EMT) (Menzl *et al.,* 2019). Bone morphogenetic proteins (BMPs) are pleiotropic cytokines that regulates processes such as cellular proliferation and EMT in cancer. BMPs bind to their respective serine/threonine kinase receptor resulting in signaling cascade that mediates phosphorylation of C-terminus of SMAD1/5/8 proteins by CDK8, localizing it to the nucleus and mediating transcription of target genes. Thus, the expression of CDK8 and EMT markers such as Snail 1 and 2 (SNAI1 & SNAI2) showed a positive correlation and that the kinase activity of CDK8 was needed for BMP4-induced EMT in cancer tissues (Serrao *et al.,* 2018). In conclusion, CDK8 mis-regulation and overexpression have been implicated in various diseases, making it an interesting target for therapeutics.

1.5.4 MED12 AND CELLULAR FUNCTIONS

MED12 gene is located on chromosome Xq13.1 and it encodes the 240 kDa MED12 protein (2212 amino acids). MED12 is divided into four domains, organized from the N-terminus to C-terminus: the leucine-rich domain (L), the leucine/ serine-rich domain (LS), the proline/glutamine/leucine-rich domain (PQL) and the Opa domain (Wang *et al.,* 2013). MED12 is a component of the CDK8 kinase module and is essential for the assembly of the module (Rubinato *et al.,* 2020). As previously mentioned, MED12 interacts with the cyclin C-CDK8 dimer, which is essential for the efficient activity of CDK8. The N-terminus of MED12 contains the activation helix, which interacts with and stabilizes a disordered activation loop of CDK8. Oncogenic mutations in the N-terminus affecting the activation helix of MED12 showed reduced catalytic efficiency of CDK8 (Park *et al.,* 2018). Thus, the interaction between MED12 and CDK8 is necessary for enhancing CDK8 kinase activity. Furthermore. the C-terminus of MED12 binds MED13 subunit, which is responsible for binding and linking the core mediator to the kinase module.

MED12 regulates and possess various cellular activities. For instance, it promotes the expression of Wnt target genes by binding to β-catenin, which subsequently recruits the core mediator at the promotor for β-catenin-induced gene transcription. The binding of MED12 and β-catenin occurs via its PQL domain; mutations to this domain reduce the transactivation activity of β-catenin (Plassche *et al.,* 2021). Furthermore, MED12 is essential in neuronal gene regulation for effective specification of cell fate and development in the nervous system. Non-neuronal cells are inhibited from expressing neuronal genes via a repressive mechanism facilitated by the RE1 silencing transcription factor (REST) via methylation events. MED12 has been shown to interact with G9a, a histone methyltransferase that methylates histone 3 on lysine 9 residue (H3K9)

(Ding *et al.,* 2008). Thus, MED12 serves a role in recruiting G9a and linking it with REST to allow for G9a-dependent histone H3K9 di-methylation. This methylation event suppresses the expression of neuronal genes in non-neuronal cells (Ding *et al.,* 2008). It is thus clear that any mutations or disruptions to MED12 can have harmful effects on non-neuronal cells. For instance, missense mutations to MED12 have been the cause of X-Linked Mental Retardation (XLMR) disorders in humans such as FG syndrome and Lujan syndrome as a result of the disruption of its REST-specific corepressor function due to its inability to recruit G9a for neuronal genes methylation and repression in non-neuronal cells (Ding *et al.,* 2008).

1.6 PROJECT RATIONALE AND OBJECTIVES

USP7 regulates various cellular pathways, by altering protein activity or via deubiquitination and stabilization of protein, making it an essential regulator of epigenetics and transcription factors to modulate gene expression. However, its role in regulating gene expression via the Mediator complex has not been established. Furthermore, the mechanism by which the proteins, CDK8 and MED12, are regulated in the cell is not well understood. Thus, our group was interested in investigating whether USP7 regulates the protein levels of CDK8 and MED12. Understanding how their levels are maintained in cells is critical as such proteins are often overexpressed and mis-regulated, giving rise to human pathologies and cancer.

Our group has analyzed the sequences of both CDK8 and MED12 and we discovered that both proteins contain the P/A/ExxS motif which mediates protein interaction with USP7. Thus, we hypothesized that USP7 interacts with both, MED12 and CDK8, impacting their stability. Our first objective was to determine whether USP7 had any effects on the protein abundance of CDK8 and MED12. Once a regulatory relationship was established, our second objective was to investigate their interaction, both *in vitro* and *in vivo* in various cell lines. Finally, we wanted to

determine whether the interaction of USP7 with both proteins affected their downstream targets including β-catenin expression and phosphorylation of CTD of RNAP II.

CHAPTER 2: MATERIALS AND METHODS

2.1 PLASMIDS

The plasmids used for mammalian transfection were: pcDNA3 Myc empty vector, pCAN/Myc-USP7^{C223S}, pCAN/Myc-USP7^{MRGR}, pCAN/Myc-USP7$^{DW-MRGR}$, pCAN/Myc-USP7, pcDNA3.1/FLAG, and pcDNA3/FLAG-CDK8. The plasmids used for bacterial transformation are shown in table 2.1

Table 2.1. His-Tagged USP7 plasmid constructs.

Construct	Vector	Purification Tag	Residues
His-USP7-MATH (Wild-type)	pET15b	6x-His	54-205
His-USP7-MATH DW Mutant	pET15b	6x-His	54-205
His-USP7-CTD (Wild-type)	pET15b	6x-His	535-1102
His-USP7-CTD MRGR D762R/D764R Mutant	pET15b	6x-His	535-1102

2.2 MAMMLIAN CELL CULTURE, TRANSFECTION, AND PREPRATION

Human osteosarcoma (U2OS), Colorectal carcinoma HCT116 wild-type cells (John Hopkins University School of Medicine, Cell line #8 40-16), and HCT116 USP7 $^{-/-}$ (John Hopkins University School of Medicine, Cell line #89) were grown in Mcoy's media (Wisent Inc.) containing 10% Fetal Bovine Serum (Gibco) and 1mg/ml Penicillin-streptomycin (Wisent Inc.). Human cervical adenocarcinoma (HeLa) cells were grown in Dulbecco's Modified Eagle Medium (DMEM) (Wisent Inc.) media containing 10% FBS and 1mg/ml Penicillin-Streptomycin. The cells were incubated at 5% CO2 and 37°C in a Thermo Scientific Forma Steri-CycleTM CO2 Incubator for cellular growth and maintenance. DNA transfection was accomplished with PolyJet^{TM} In Vitro DNA transfection reagent (SignaGen Laboratories)

according to the manufacturer's instructions. For cellular lysates preparation, cells were

harvested and lysed with gentle rotation at 4 degrees for 30 minutes. The lysis buffer used was

100mM NaCl, 50mM Tris pH 7.5, 0.8% NP-40, 5% glycerol, with 1X Protease inhibitor cocktail

and 1X Protease inhibitor tablet (Roche cOmplete ULTRA Tablets). The lysates were cleared by

centrifugation at 17,000 x g for 15 minutes. The concentration of the lysates was quantified using

Pierce™ BCA Protein Assay Kit (ThermoFisher) according to the manufacturers protocol. 20 µg

of lysate samples were made using 5x SDS loading dye (250 mM Tris-Cl, pH 6.8, 2% (w/v)

SDS, 30% (v/v) glycerol, 0.5 M DTT, 0.05% (w/v) bromophenol blue, 5% β-mercaptoethanol).

The samples were then boiled for 5 minutes at 100 degrees and loaded onto an SDS-PAGE gel.

2.3 SDS-PAGE GEL & WESTREN TRANSFER

20 µg of lysate was resolved in 10% SDS polyacrylamide gel electrophoresis (SDS-PAGE) for 1

hour at 180 volts. For western blotting, a 0.45 µm PDVF membrane (Millipore) was used and

activated with 100% methanol for 45 seconds, rinsed with Milli-Q water, and calibrated with 1X

western transfer buffer. The western buffer used for the transfer contained (25 mM Tris [pH 8.3],

192 mM Glycine, and 20% methanol. The SDS gel was transferred on the membrane for 1 hour

and 30 minutes at 100 volts at 4 degrees. The transferred membrane was then blocked for either

1 hour at room temperature or overnight at 4 degrees with 5% skim milk in TBS-T (0.1%

Tween20) with rocking.

2.4 IMMUNOBLOTTING AND ANTIBODIES

The following antibodies were used for either membrane blotting for proteins of interest or for

immunoprecipitation studies: polyclonal rabbit anti-USP7 (Bethyl, A300-033A), monoclonal

mouse anti-USP7 (Millipore, 05- 1946), polyclonal rabbit anti-CDK8 (Cell-signaling, D6M3J),

polyclonal rabbit anti-MED12 (Cell-signaling, D9K5J), monoclonal mouse anti-Myc (Millipore, 05-724), monoclonal mouse anti-ECS (FLAG) (Sigma-Aldrich, F1804), polyclonal rabbit anti-ECS (FLAG) (Bethyl, A190-102A), monoclonal mouse anti-EZH2 (Cell-signaling, AC22), monoclonal mouse anti-GAPDH (Santa-Cruz, sc047723), polyclonal rabbit anti-Bmi1 (Cell-Signalling, 5856), polyclonal rabbit anti-β-catenin (Cell-signaling, 9582), normal rabbit IgG (Cell-Signaling 2729S), normal mouse IgG (Santa-Cruz, sc-2025), monoclonal mouse anti-His tag (Santa-Cruz, sc-8036), monoclonal mouse anti-Ser5p (Santa Cruz sc-44701), and monoclonal rat anti-Ser2p (Millipore, 04-1571). The transferred membrane was incubated, facing down, with 500 µl of diluted primary antibodies either at room temperature for 1.5 hours or overnight at 4 degrees. Post primary incubation, the membranes were washed with either 1X PBST or 1X TBST, 3 times in 5 minutes intervals with shaking. The washed membranes were then subjected to and incubated with secondary antibodies for 1 hour at room temperature. The secondary antibodies used were Trueblot anti-mouse (Rockland, 18-8817-33), HRP-conjugated goat polyclonal anti-mouse IgG (ThermoFisher, 31430) and HRP-conjugated goat polyclonal anti-rabbit IgG (Bethyl, A120-101P). Post secondary incubation, the membranes were washed 3 times (15 minutes intervals) with 1X PBST with shaking. For imaging the blots of interest, the membranes were incubated for 5 minutes with Amersham ECL Select Western Blotting Detection Reagent (Cytiva). The blots were imaged using Microchemi® (DNR-IS) system using different exposure timings.

2.5 USP7 RECONSTITUTION IN HCT116 USP7 [-/-] CELLS

HCT116 USP7 [-/-] cells were passaged the day before transfection and supplemented with fresh media so that on the day of transfection, cell confluency was around 75-80%. The cells were transfected with 5 µg of either Myc empty vector, pcDNA 3, or pCAN/Myc-USP7 plasmids

using PolyJet^{TM} In Vitro DNA transfection reagent to manufacturers protocol. After 5 hours post transfection, the media was replaced with fresh media to reduce cellular toxicity. The cells were then harvested after 24 hours post transfection, lysed using lysis buffer, and sample preparation was made as mentioned previously.

2.6 USP7 KNOCKDOWN VIA siRNA

Double transfection of siRNA was performed to knockdown the levels of USP7 in either U2OS or HeLa cells. Prior to siRNA transfection, the cells were passaged in 10 cm plates the day before to yield plate confluency of 75-80% on the day of transfection. The cells were transfected with 434 pmol of either siRNA or negative control (scrambled siRNA) using Lipofectamine™ 3000 Transfection Reagent according to the manufactures protocol (ThermoFisher). After 5 hours, the 5 ml media in the plates were topped to 5 ml of fresh media and the plates were placed back in the CO2 incubators for incubation. After 24 hours from initial transfection, the media was replaced with fresh media and cells were transfected again with 434 pmol of either siRNA against USP7 or scrambled siRNA using Lipofectamine™ 3000 Transfection Reagent according to the manufactures protocol. The cells were then harvested after 12-24 hours post the second transfection. Cells were lysed at 4 degrees for 30 minutes using gentle rotation. 20 µg of cellular lysate was resolved on 10% SDS-PAGE gel and western blotting was performed as mentioned before. The siRNA used in this experiment were:

siUSP7 5'- CCC AAA UUA UUC CGC GGA AAT T-3'(GenePharma) or

siNegative Control 5'- UUC UCC GAA CGU GUC ACG UTT -3'(GenePharma).

2.7 ENDOGENOUS CO-IMMUNOPRECIPITATION STUDIES

Wild-type U2OS, HCT116 and Hela cells were harvested and lysed in 50 mM Tris, 150 mM

NaCl, 0.8% NP-40 lysis buffer with 1X Protease inhibitor cocktail and 1X Protease inhibitor

tablet (Roche cOmplete ULTRA Tablets), with gentle agitation at 4 degrees for 30 minutes.

Lysates were centrifuged and lysate concentration was quantified with Pierce™ BCA Protein

Assay Kit (ThermoFisher) according to the manufacturers protocol. 1 mg of lysate was incubated

with 2 µg of either mouse or rabbit anti-USP7, and either 2 µg of normal mouse or rabbit IgG.

The samples were incubated overnight at 4 degrees with rotation. 50 µl of A/G plus beads were

precleared with 5% BSA in PBS for 1 hour at 4 degrees with rotation and the beads were washed

3 times with PBS and equilibrated with lysis buffer. The lysate containing the respective

antibody was added to the beads and incubated for 1 hour with rotation at 4 degrees. Post

incubation, the beads were washed 3 times with 100 mM NaCl, 50 mM Tris pH 7.4, and the

beads were eluted with 35 µl of 5x SDS and boiled for 7 minutes.

The reciprocal co-immunoprecipitation studies were done by incubating the lysates with 2 µg of

rabbit anti-CDK8 or 2 µg of normal rabbit IgG, and the procedure was followed as mentioned

above.

2.7.1 EXOGENOUS CO-IMMUNOPRECIPITATION STUDIES

U2OS cells were transfected with either: 5 µg of FLAG-CDK8, 5 µg of Myc-USP7, or

co-transfected with 5 µg of Myc-USP7 and 5 µg of FLAG-CDK8, for a total of 10 µg of DNA.

The harvested cells were then lysed, and the lysates were incubated with either: 2 µg of mouse

anti-FLAG or mouse normal IgG, 2 µg of mouse anti-Myc or mouse normal IgG, and 2 µg of

mouse anti-Myc or mouse normal IgG respectively. The reaction samples were then washed and prepared as outlined above.

2.7.2 FLAG CO-IMMUNOPRECIPITATION STUDIES

HCT116 USP7 wild-type cells were transfected with either 10 µg FLAG-CDK8, 5 µg Myc-USP7, 5 µg Myc-USP7^{MRGR}, and 5 µg Myc-USP7^{MRGRDW} using PolyJet^{TM} In Vitro DNA transfection reagent according to manufacturers protocol. Cells were harvested after 24 hours and lysed in FLAG IP buffer: 80mM NaCl, 50mM Tris pH 7.4, 1% Trixton, 1mM EDTA. For each reaction, 40 µl of ANTI-FLAG® M2 Affinity Gel (Sigma-Aldrich) resin was washed 3 times with 1X TBS-T and calibrated with the FLAG IP buffer. The lysates were added to the calibrated resin in a 1:2 ratio- 350 µg of FLAG-CDK8 with either 700 µg of Myc-USP7, Myc-USP7^{MRGR}, or Myc-USP7^{MRGRDW}. The negative control was 700 µg of Myc-USP7 lysate combined with 350 µg lysate of pcDNA 3.1 FLAG empty vector. Reaction tubes were incubated overnight at 4 degrees with rotation and the beads were washed 3 times with 100 mM NaCl, 50 mM Tris pH 7.4. Samples were eluted with 2x SDS (20 µl) and boiled for 3 minutes. The beads were spun down and supernatant was collected and resolved on 10% SDS-PAGE gel and ran for 120 volts for 1 hour and 30 minutes. The gel was then transferred to PDVF membrane at 100 volts for 1 hour and 45 minutes. The membrane was subjected to immunoblotting using the primary antibodies incubated with membrane overnight at 4 degrees: rabbit anti-FLAG (1:2000), mouse anti-Myc (1:1000). For secondary immunoblotting, anti-rabbit (1:5000) and Trueblot anti-mouse (1:1000) were used and incubated with the membrane for 1 hour at room temperature. The blots were washed and imaged as mentioned previously.

The reciprocal coimmunoprecipitation of mutant Myc-UPS7 constructs was also performed. HCT116 wild-type cells were transfected with either 5 µg of Myc-USP7, Myc-USP7^{MRGR}

mutant, or Myc-USP7$^{MRGR\ DW}$ double mutant. Cells were lysed and quantified as previously

mentioned. The Myc-USP7 constructs were immunoprecipitated using 2 µg of mouse anti-Myc.

Mouse IgG was also used as the negative control. The immunoprecipitation reaction was

incubated overnight at 4 degrees with rotation, subsequently washed with 1X PBS (3x), and

eluted using 5x SDS. The eluted fractions were resolved on 10% SDS-PAGE gel, transferred

onto PVDF membrane for 1 hour and 30 minutes at 100 volts. The proteins were then blotted

using rabbit anti-CDK8 (1:1000) and mouse anti-Myc (1:1000).

2.8 BACTERIAL TRANSFORMATION

1 µl of plasmid was added to 50 µl of DH5α or BL21 mgk *Escherichia coli* and incubated on ice

for 5 minutes. The cells were then heat shocked for 45 seconds at 42°C. Post heat shock, cells

were incubated on ice for 5 minutes and 450 µl of Luria Broth (LB) media was added to the

cells. The cells were then incubated at 37°C for 1 hour with shaking to facilitate growth. Post

incubation, the DH5α cells were added to 5 ml of LB containing 100 µg/ml Ampicillin, while the

BL21 mgk *E. coli* cells were plated on LB agar plates containing both 100 µg/ml Ampicillin and

50 µg/ml Kanamycin. The tubes were incubated overnight at 37 °C with shaking (200 rpm)

(Innova 43R, New Brunswick Scientific), while the plates were incubated overnight at 37°C

incubator. Post plates incubation, a single colony was selected and grown overnight at 37°C in

5 ml of LB containing 100 µg/ml Ampicillin and 50 µg/ml Kanamycin.

2.8.1 RECOMBINANT PROTEIN EXPRESSION IN *E. COLI* BL21 MGK CELLS

50 ml of LB media containing 100 µg/ml Ampicillin and 50 µg/ml Kanamycin were inoculated

with freshly transformed colonies containing plasmid of interest and incubated overnight at 37°C

with shaking at 200 rpm (Innova 43R, New Brunswick Scientific). The 50 ml culture was then

transferred to 1 litre of Terrific Broth (TB) containing 100 µg/ml Ampicillin and 50 µg/ml Kanamycin. The flask was then incubated at 37°C with shaking for several hours. Gene expression was induced using 0.4 mM Isopropyl β-D-1 thiogalactopyranoside (IPTG) (Bioshop) when the OD_{600nm} reached a value within the range 0.8-1. The induced culture was incubated at 16°C overnight with shaking (200 rpm). Post incubation, the cells were harvested by centrifugation at 6,000 x g for 30 minutes using a JLA 9.1000 Beckman rotor (Beckman Coulter Avanti J-E centrifuge).

2.8.2 PROTEIN PURIFICATION USING NICKEL AFFINITY CHROMATOGRAPHY

The pelleted cells were resuspended in 40 ml of Binding/Lysis buffer (500 mM NaCl, 20 mM Tris [pH 7.5], 5 mM Imidazole, 5% glycerol, and 1 X Protease inhibitor cocktail (0.001 M Benzamidine and 0.0005 M PMSF in ethanol) and 1X Protease inhibitor tablet (cOmplete EDTA-Free, Roche). The cells were sonicated on ice at 30% amplitude for 8 minutes, 10 seconds ON, 15 seconds OFF, (Branson Digital Sonifier D450). The sonicated fraction was pelleted using a JA-25.50 Beckman rotor at 18000 rpm for 30 minutes at 4°C (Beckman Coulter Avanti J E centrifuge). The supernatant was collected and added to a column containing 5 ml of washed and equilibrated Ni^{2+}-NTA-Agarose (Qiagen). The supernatant and the Nickel beads were incubated for 1 hour at 4°C with gentle rocking. Post incubation, the beads was washed with 100 ml of wash buffer (500 mM NaCl, 20 mM Tris [pH 7.5], 20 mM Imidazole, and 5% glycerol). The His-tagged proteins were then eluted from the beads twice, by incubating the resin with 5 ml of elution buffer (500 mM NaCl, 20 mM Tris [pH 7.5], 500 mM Imidazole, and 10% Glycerol) for 5 minutes at 4°C. The eluted proteins were dialyzed overnight in 150 mM NaCl, and 20 mM Tris [pH 7.5].

2.9 HIS-PULL DOWN ASSAY

The dialyzed purified His-protein constructs (Table 2.1) were concentrated using spin concentrators (Thermo Scientific™ Pierce™ Protein Concentrator PES) at 4°C. 300 µg of purified His-USP7 proteins were incubated with 100 µl of equilibrated Ni^{2+}-NTA-Agarose (Qiagen). The reaction was incubated for 1 hour at 4°C in binding buffer (500 mM NaCl, 20 mM Tris [pH 7.5], 5 mM Imidazole, 5% glycerol, and 1 X Protease inhibitor cocktail (0.001 M Benzamidine and 0.0005 M PMSF in ethanol) and 1X Protease inhibitor tablet (cOmplete EDTA-Free, Roche). This served as the bait for the pulldown assay. Post incubation, the beads were washed three times with the binding buffer. Then, 600 µg of mammalian lysates either from HCT116 or FLAG-CDK8 transfected HCT116 cells were added to the washed beads. The reaction was incubated overnight at 4°C with rotation. Post incubation, the reaction tubes were washed three times with 150 mM NaCl, and 20 mM Tris [pH 7.5] (1x TBS) and subsequently eluted using 100 µl of elution buffer (500 mM NaCl, 20 mM Tris [pH 7.5], 500 mM Imidazole, and 10% Glycerol). 5% input (containing His-USP7 proteins) and 35 µl of the eluted samples were made using 5x SDS loading dye and boiled for 6 minutes at 95 °C. The samples were then resolved on 12% SDS-polyacrylamide gel and transferred on PDVF membrane using western blotting as mentioned previously. The membrane was subjected to immunoblotting using anti-His tag (1:1000), anti-CDK8 (1:1000), anti-MED12 (1:1000), anti-Bmi1 (1:1000), anti-EZH2 (1:1000), and anti-FLAG (1:2000).

CHAPTER 3: RESULTS

3.1 PREDICTING INTERACTION SITES OF CDK8 & MED12

The MATH domain of USP7 contains the motif, ^{164}DWGF^{167}, that recognizes the consensus sequence P/A/ExxS on its target substrate, resulting in their interaction. Bioinformatic tools used by our lab indicate that both CDK8 and MED12 contain consensus sequence for binding to the MATH domain of USP7 at the N-terminus. The substrate interaction site of CDK8 and MED12 were predicated using a combination of bioinformatic tools such as Uniprot, Eukaryotic Linear Motif Resource (ELM), and BLAST sequence alignment. The protein sequence of both CDK8 and MED12 were obtained and entered in ELM to search for regions/motifs that could have potential for protein-protein interactions. Our lab has discovered that CDK8 has the sequence ^{436}PSTSQ^{440} as the potential site for interacting with the NTD of USP7. Furthermore, the potential site for MED12 interaction with USP7 has been predicted to be ^{337}PTSST^{341}. The interaction site of our proteins of interest were aligned with the interaction site of substrates that are known to interact with the DWGF motif of USP7. As shown in Figure 3.1, CDK8 and MED12 consensus sequence closely follows that of well-known substrates such as p53, CHFR, and MDM2, with the P and S amino acids residues being highly conserved. As a result, determining the potential interaction site allows us to investigate whether CDK8 and MED12 are indeed interactors of USP7 through various protein-protein interaction and protein abundance related assays.

3.2 THE PROTEIN LEVELS OF CDK8 AND MED12 ARE DEPENDENT ON USP7

To determine whether USP7 has any effect on the protein levels of CDK8 and MED12, the levels of CDK8 and MED12 were blotted for via western blotting in the two cell lines: HCT116

and HCT116 USP7$^{-/-}$. The HCT116 cell line possesses functional USP7 gene, and the HCT116 USP7$^{-/-}$ cell line has the USP7 gene knocked out. As seen in Figure 3.2, the protein bands for both CDK8 and MED12 are stronger in their intensity and more abundant in the HCT116 parental cell line containing functional USP7, in comparison with the levels found in the HCT116 USP7$^{-/-}$ cells. Likewise, EZH2, a positive control, also shows abundance in its protein levels in the presence of functional USP7 gene in comparison to the HCT116 USP7$^{-/-}$ cells.

The blots were further quantified using ImageJ software to determine whether such quantitative data possess any statistical differences. At least 3 biological replicates were used for the quantification. That is, at least 3 different plates of each cell line were harvested and probed for the proteins of interest. Thus, the quantification data presented in figure 3.2 B represent the mean fold change of protein signals normalized to GAPDH, the loading control. The fold change in protein signal is a ratio of protein signals in HCT116 USP7$^{-/-}$ to HCT116 parental cells. CDK8 showed a mean fold change of 0.38 ± 0.078 (SEM) and a paired T-test was performed, giving a p value of 0.00067, indicating statistically significant difference (Fig. 3.2 B). MED12 and EZH2 had a mean fold change of 0.27 ± 0.108 and 0.39 ± 0.15 respectively (Fig. 3.2 B). Additionally, the difference in the fold change was statistically significant as T-test produced a p value of 0.011 and 0.026 for MED12 and EZH2 respectively. Thus, this shows that the presence of USP7 impacts the protein levels of CDK8 and MED12 in the HCT116 cell line.

3.2.1 USP7 RECONSTITUTION IN HCT116 USP7 $^{-/-}$ CELLS RESCUES PROTEIN LEVELS OF CDK8 & MED12

Since the HCT116 USP7$^{-/-}$ cells lack a functional USP7 gene, we decided to investigate what would happen to the levels of both CDK8 and MED12 when USP7 is reintroduced back. Thus, HCT116 USP7$^{-/-}$ were transfected with either pCAN/Myc-USP7 or Myc empty vector plasmids

and the proteins levels of CDK8 and MED12 were blotted for and compared in both treatments. When USP7 is introduced back to the HCT116 USP7$^{-/-}$ cell lines, CDK8 (Fig 3.3 A) and MED12 (Fig 3.3 B) levels go back up and increase compared to the USP7 knockout cells transfected with only the empty vector. Likewise, the levels of EZH2 also increase when USP7 is reintroduced back to the knockout cells. At least two biological replicates were done for this experiment and the signals produced from the protein of interests were quantified using ImageJ, normalized to GAPDH, and presented as a mean fold change of signals from cells transfected with empty vector to Myc-USP7. CDK8 had a mean fold change of 0.68 ± 0.13 (Fig 3.3 C), with a p-value of 0.067 generated from a paired T-test. EZH2 also had a mean fold change of $0.60 +/- 0.15$ (Fig 3.3 C), with a p-value of 0.11. According to the T-test, no significant difference is seen for both CDK8 and EZH2. Finally, MED12 displayed a mean fold change of 0.46 ± 0.0040, with a p-value of 0.0024, indicating statistical significance (Fig 3.3 C). Despite the lack of statistical significance for both CDK8 and EZH2, quantitative blots for all three proteins show that the phenotype of protein stability is rescued when USP7 is re-introduced.

Figure 3.1. **Alignment of Substrate Interaction Site**. Bioinformatic tools such as Uniprot, protein BLAST and Eukaryotic Linear Motif Resource have predicted that CDK8 and MED12 can potentially interact with the DWGF motif of USP7. CDK8 and MED12 follow the substrate consensus sequence **P/A/ExxS** for binding to NTD of USP7, with the potential sequence of interaction as ^{436}**PST$\underline{S}$Q**^{440} and ^{337}**PTS$\underline{S}$T**^{341} respectively. The amino acid residues Proline (P) and Serine (S) are highly conserved among the aligned known substrates and as well as CDK8 and MED12.
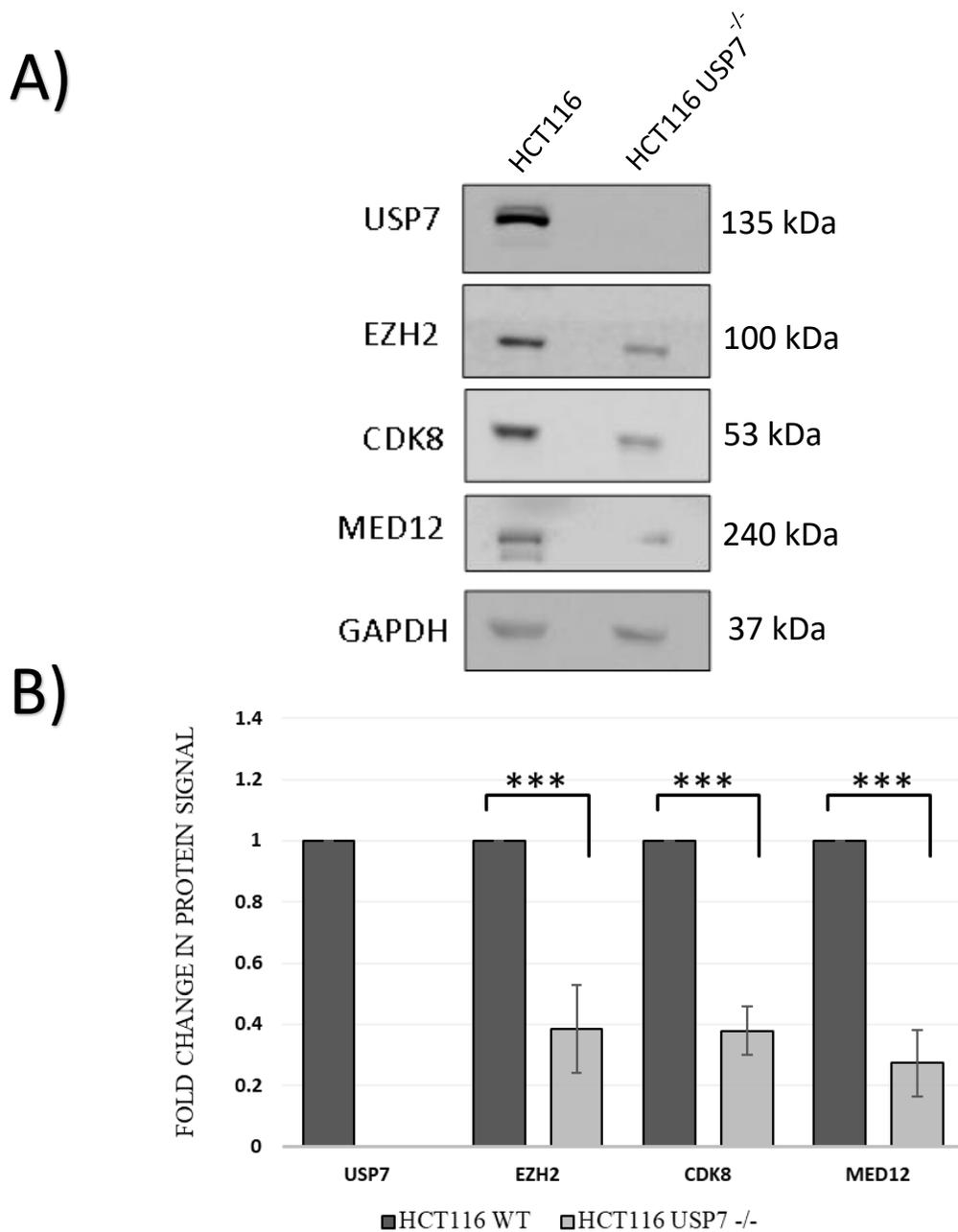
Figure 3.2. **The Levels of CDK8 and MED12 is Dependent on USP7**. A) HCT116 and HCT116 USP7[-/-] cells were lysed, and 20 µg of lysate was loaded on 10% SDS-PAGE gel, and transferred on PVDF membrane for 1 hour and 30 min at 100 volts. The indicated proteins were probed with rabbit anti-USP7, mouse anti-EZH2, rabbit anti-CDK8, rabbit anti-MED12, and mouse anti-GAPDH. GAPDH is used as the loading control and EZH2 is used as a positive control. B) The blots were quantified using ImageJ software and the fold change level of protein signals were determined as ratio of HCT116 USP7[-/-] to HCT116 signals. The ratios were then normalized to GAPDH. A paired T-test was conducted to evaluate statistical significance in the level changes of EZH2, CDK8 and MED12 ($p<0.05$= significant, indicated by ***). Error bars represent Standard Error of the Mean (SEM). EZH2 (3 biological replicates), CDK8 (5 biological replicates), MED12 (3 biological replicates).
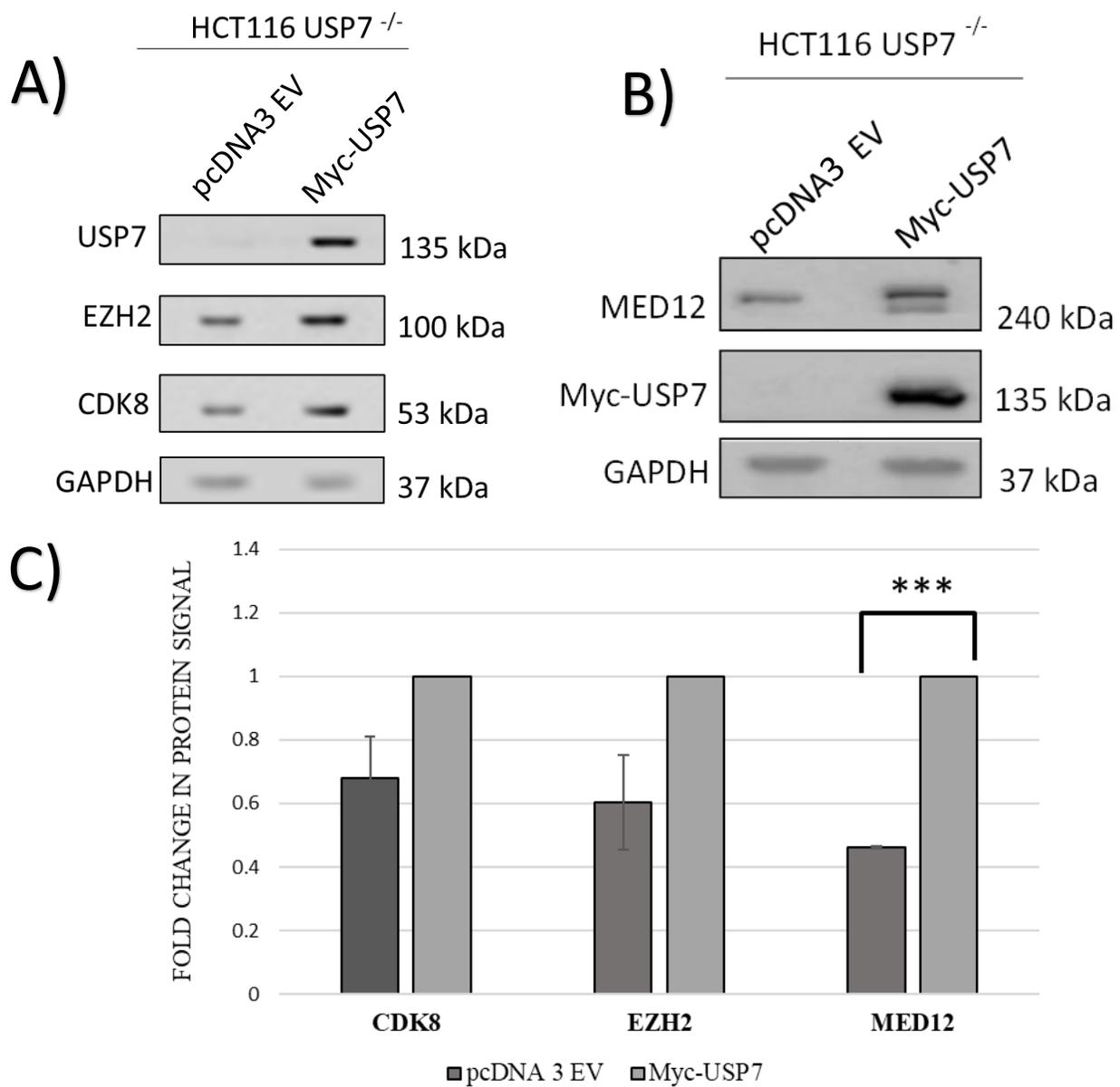
Figure 3.3. **Reconstitution of USP7 in HCT116 USP7 $^{-/-}$ Cells**. HCT116 USP7$^{-/-}$ were transfected with either 5 µg of pCAN/Myc-USP7 or pcDNA3 Myc empty vector (EV). Cells were harvested 24 hours post transfection and lysed. 20 µg of lysate was resolved on 10% SDS-PAGE gel and transferred on PVDF for 1 hour and 30 minutes at 100 volts. A) proteins were blotted for, using mouse anti-USP7, rabbit anti-CDK8, mouse anti-EZH2, and mouse anti-GAPDH. B) proteins were blotted using rabbit anti-MED12, mouse anti-Myc, and mouse anti-USP7. C) The blots were quantified using ImageJ software and the fold change level of protein signals were determined as ratio of cells transfected with pcDNA3 EV to Myc-USP7. The ratios were then normalized to GAPDH. A paired T-test was conducted to evaluate statistical significance in the level changes of EZH2, CDK8 and MED12 ($p<0.05$ indicated by ***). Error bars represent Standard Error of the Mean (SEM). EZH2 (2 biological replicates), CDK8 (3 biological replicates), and MED12 (2 biological replicates).

3.2.2 CDK8 & MED12 PROTEIN LEVELS ARE DEPENDENT ON THE CATALYTIC ACTIVITY OF USP7

Knowing that USP7 increases the levels of CDK8 and MED12 when it is introduced back to the USP7 knockout cell lines, we decided to investigate what would happen to the levels of the proteins if a catalytic mutant of USP7 was introduced. The USP7 catalytic mutant contains a missense substitution at amino residue 233, whereby serine is substituted for the catalytic Cysteine 233 (Sarkari, *et al.,* 2013). Such amino acid mutation is common in mutagenesis studies as it determines the effect of the absence of the formation of di-sulfide bonds on the protein's catalytic function. The USP7 C233S mutant makes USP7 catalytically inactive, compromising its deubiquitylation activities. To investigate whether the maintenance of MED12 and CDK8 protein abundance is dependent on the catalytic activity of USP7, we transfected the plasmids: pcDNA3 Myc empty vector, pCAN/Myc-USP7 WT or pCAN/Myc-USP7 Catalytic mutant (C233S) and probed for CDK8 and MED12 proteins in the various cell lines: HeLa, HCT116 USP7 $^{-/-}$, and parental HCT116. Based on the results obtained from Figure 3.4 A, B, & C), we can see that endogenous CDK8 is more abundant in all three cell lines, HCT116 USP7$^{-/-}$, HCT116, and HeLa, when the functional USP7 protein is present, compared to the cells transfected with only the empty vector and the catalytic mutant. Furthermore, MED12 levels are also more abundant as seen in Figure 3.4 A when USP7$^{-/-}$ cells are transfected with the WT Myc-USP7, in comparison to the cells transfected with either the empty vector or USP7 catalytic mutant.

Given that three and two biological replicates were done for HCT116 USP7$^{-/-}$ and HeLa cells respectively, quantification of the immunoblots were thus performed. The signals generated from the blots were quantified using ImageJ, normalized to GAPDH, and a mean signal fold change was plotted as a ratio of either pcDNA 3 EV to Myc-USP7 or Myc-USP7 C233S to Myc-USP7.

In the HCT116 USP7$^{-/-}$ cells, the protein levels of catalytic mutant and wild-type USP7 are similar, with a mean fold change of 0.99 ± 0.021 (mean +/- SEM) and 1 respectively. The mean fold change of CDK8 in HCT116 USP7$^{-/-}$ cells treated with only the empty vector was determined to be 0.069 ± 0.024. A paired T-test was also performed which produced a p-value of 0.0015, suggesting statistically significant difference (indicated by ***) between cells treated with EV versus Myc-USP7. Additionally, the mean fold change of CDK8 in HCT116 USP7$^{-/-}$ cells transfected with the USP7 catalytic mutant was determined to be 0.59 ± 0.069. Paired T-test was also conducted, producing a p-value of 0.013, indicating once again, statistically significant differences in the levels of CDK8 in the USP7 wild-type and catalytically mutant transfected cells. Finally, quantification of MED12 blots yielded a mean fold change of 0.68 ± 0.038 and 0.72 ± 0.069 in HCT116 USP7$^{-/-}$ cells transfected with empty vector and catalytically mutant USP7 respectively. However, statistically significant difference was only seen with cells transfected with EV. The p-value obtained from the T-test was 0.038. The p-value for catalytically mutant transfected cells was 0.076 (Fig 3.5 A).

Similar trends can be observed in the HeLa cell line. The protein levels of Myc-USP7 C233S are similar to the wild-type Myc-USP7, with a mean value of 1.0 ± 0.12. In HeLa cells transfected with the empty vector, the mean fold change of CDK8 band signals were quantified to be 0.49 ± 0.011. Finally, quantification of the immunoblots obtained for CDK8 from cells transfected with USP7 C223S mutant produced a mean fold change of 0.50 ± 0.067 (Fig 3.5 B). Paired T-test was conducted for both groups, which produced a p-value of 0.0069 for EV transfected cells and 0.043 for Myc-USP7 C233S transfected cells, indicating statistically significant differences between the two group and the HeLa cells transfected with wild-type Myc-USP7.
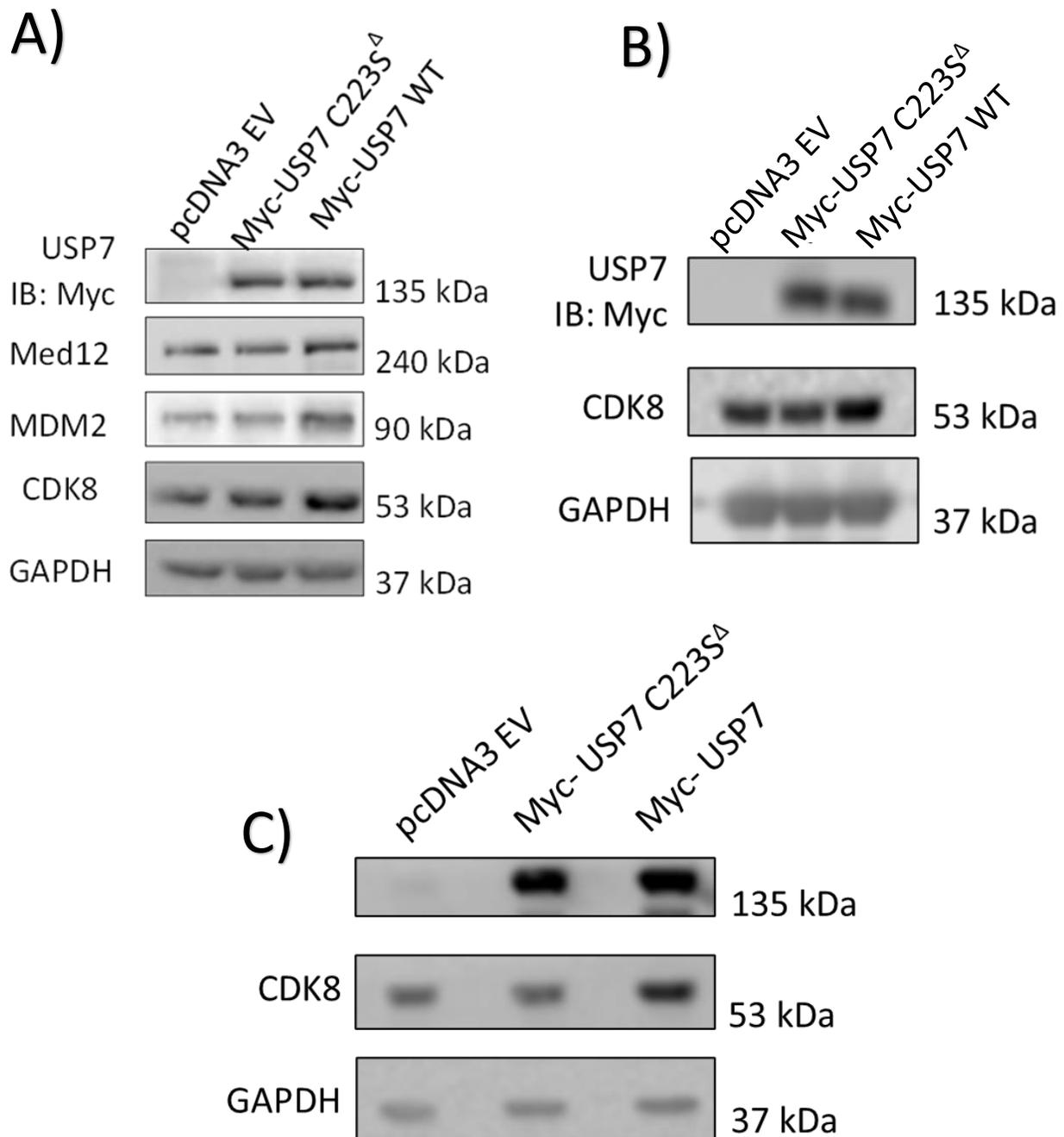
Figure 3.4. **CDK8 and MED12 Levels are Dependent on Catalytic Activity of USP7**.
pcDNA3 Myc EV, Myc-USP7 C223S mutant, and Myc-USP7 WT plasmids were transfected in
A) HCT116 USP7 $^{-/-}$ (N=3), B) HCT116 (N=1), and C) in HeLa (N=2). Cells were harvested 24
hours post transfection and 20 µg of cellular lysate was resolved onto 10% SDS-PAGE gel. The
gel was transferred onto PVDF membrane for 1 hour and 30 minutes at 100 volts and the
membrane was blotted with rabbit anti-MED12, rabbit anti-MDM2, rabbit anti-CDK8, mouse
anti-GAPDH, and mouse anti-Myc for Myc-USP7 constructs.
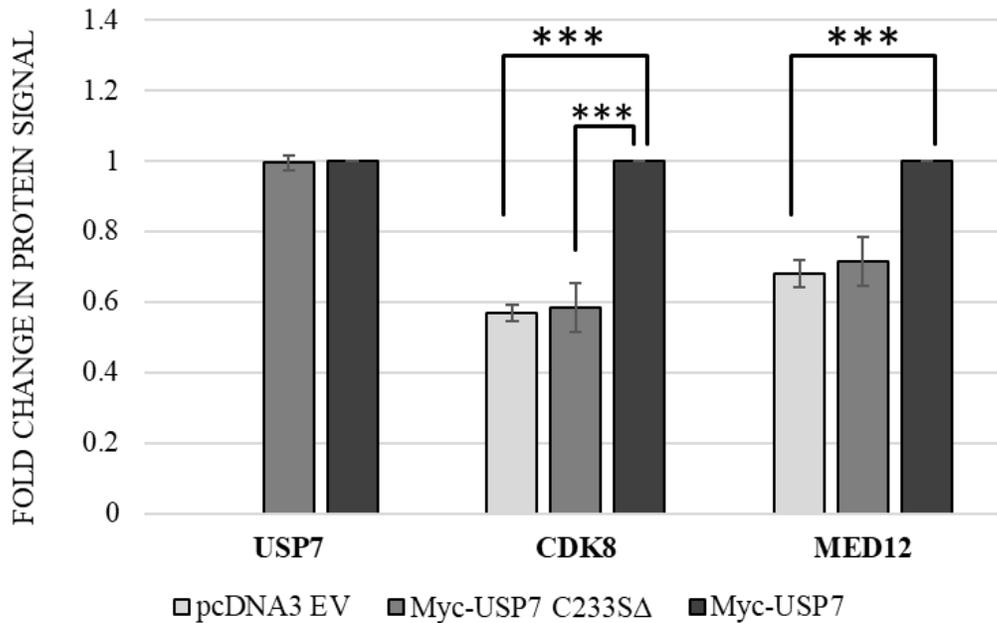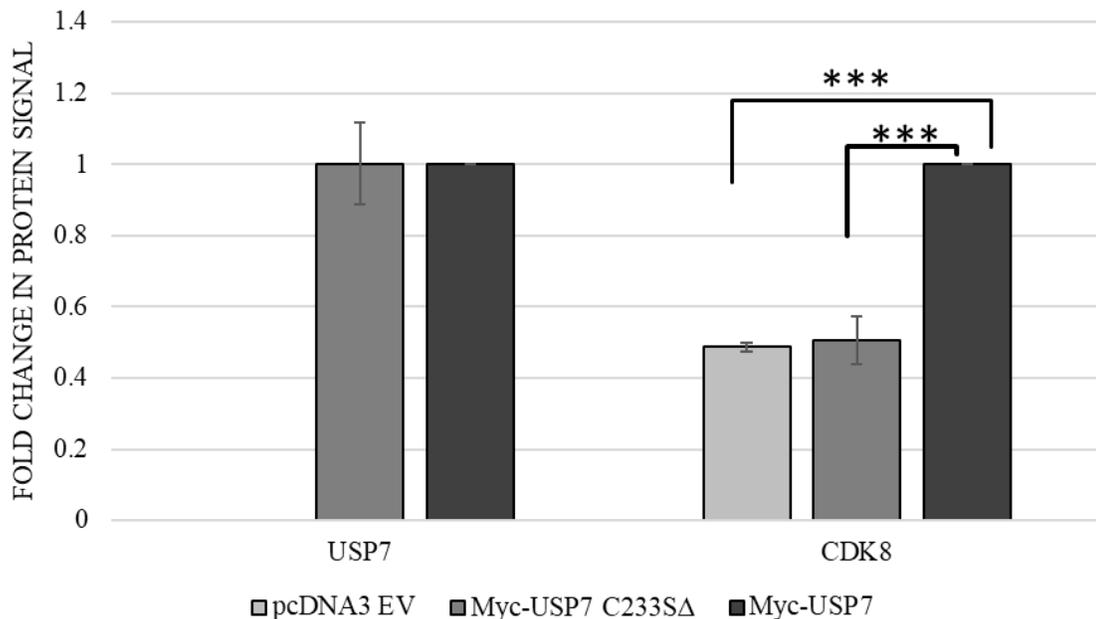
Figure 3.5. **Relative protein levels of CDK8 and MED12 in HCT116 USP7$^{-/-}$ and HeLa cells transfected with pcDNA3 EV, Myc-USP7 C223S$^\Delta$, and Myc-USP7.** The blots were quantified using ImageJ software and the fold change level of protein signals were determined as ratio of cells transfected with pcDNA3 EV to Myc-USP7 and Myc-USP7 C223S$^\Delta$ to Myc-USP7. The ratios were then normalized to GAPDH. A paired T-test was conducted to evaluate statistical significance in the level changes of CDK8 and MED12 (p<0.05, indicated by ***). Error bars represent Standard Error of the Mean (SEM). A) HCT116 USP7$^{-/-}$ cells: CDK8 (3 biological replicates) and MED12 (2 biological replicates). B) HeLa cells: CDK8 (2 biological replicates).

3.2.3 CDK8 & MED12 PROTEIN LEVELS ARE INCREASED UPON INCREASING USP7 CONCENTRATION

Finally, the effect of USP7 on CDK8 and MED12 will further be examined by increasing the concentration of USP7 and observe how that impacts the levels of our proteins. To do this, U2OS cells were transfected with varying amounts of pCAN/Myc-USP7 such as 0 µg, 1 µg, 5 µg, 10 µg and 15 µg of the plasmid. Then, endogenous levels of MED12 and CDK8, as well as β-catenin, were probed for via western blotting. β-catenin is a known substrate of USP7 and its expression is also impacted by the CDK8 kinase module. As shown in Figure 3.6 A, when the concentration of USP7 is increased, the levels of endogenous CDK8 and MED12 also increase as indicated by the increase in the thickness of the bands. Furthermore, there is more CDK8 and MED12 in the cells transfected with 15 µg of Myc-USP7 compared to the cells transfected with only the empty vector. Our positive control, β-catenin, also increases in the intensity of the protein bands, as the concentration of Myc-USP7 is increased from 0 µg to 15 µg. Since it seems like higher concentrations of USP7 do not show a drastic difference in the levels of CDK8, we then decided to further repeat this experiment but only focusing instead on transfecting 0 µg, 1 µg, and 5 µg of Myc-USP7 and blotting for the presence of endogenous CDK8. Consistent with the previous results, when the concentration of Myc-USP7 is increased to 5 µg, there is more CDK8 present as indicated by the thick dark protein band, compared to the cells transfected with the Myc empty vector. Also, when 1 µg vs 5 µg of Myc-USP7 is transfected, there is more CDK8 present in cells treated with 5 µg of Myc-USP7 than in cells with 1 µg of Myc-USP7 (Fig 3.6 B). Finally, there is slightly more CDK8 in cells transfected with 1 µg of USP7 than in cells transfected with only the empty vector (Fig 3.6 B). The blots for this experiment were quantified using ImageJ software and plotted by normalizing the band intensities to GAPDH. Figure 3.6 C shows that with increasing USP7 concentration, EZH2 and CDK8 also increase in response.
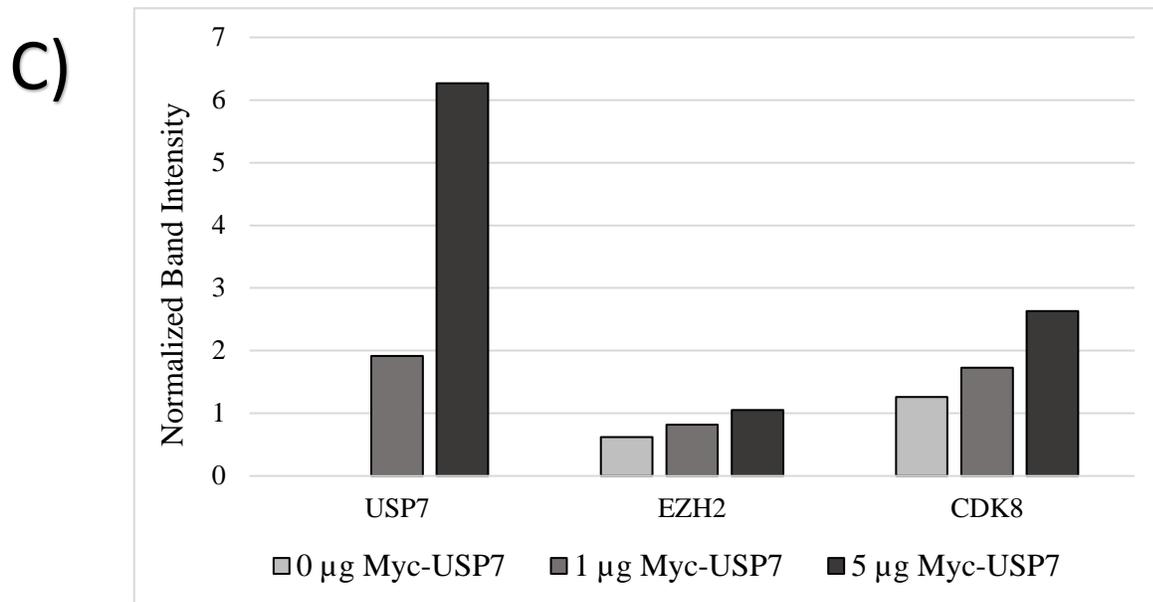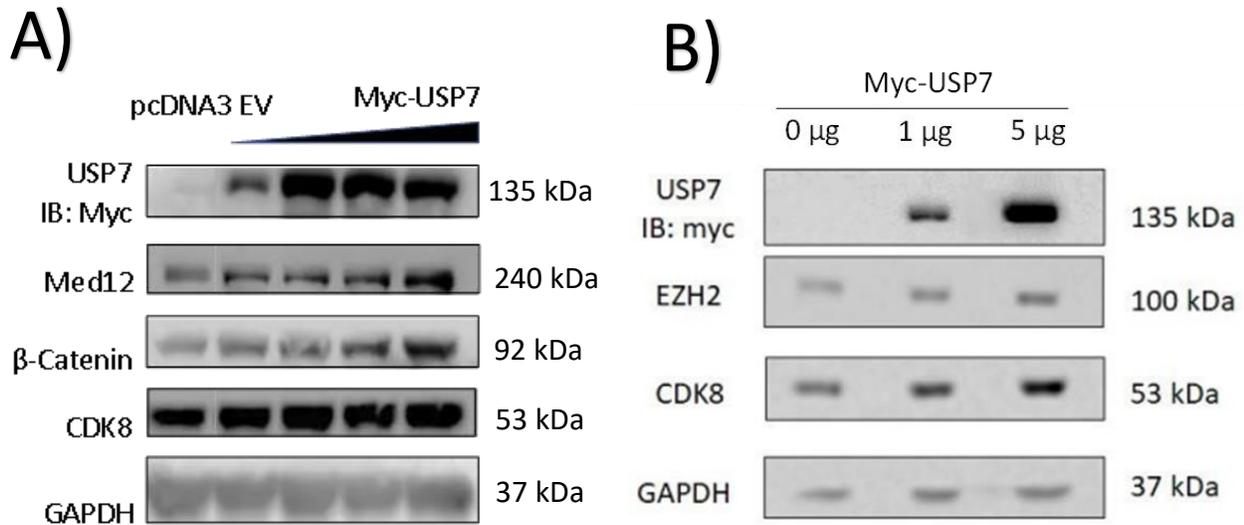
Figure 3.6. **Effect of Increasing Concentration of USP7 on CDK8 and MED12**. A) U2OS cells were transfected with either 5 µg pcDNA3 Myc EV (0 µg), or 1 µg, 5 µg, 10 µg, and 15 µg of Myc-USP7 (N=1). B) U2OS cells were transfected with only pcDNA3 EV, 1 µg or 5 µg of Myc-USP7 (N=1). Cells were harvested 24 hours post transfection. 20 µg of cellular lysate was loaded onto 10% SDS-PAGE gel and western blot was performed. The membrane was probed with mouse anti-Myc, rabbit anti-Med12, rabbit anti- β-catenin, rabbit anti-CDK8, and mouse anti-GAPDH. GAPDH is the loading control. C) Immunoblots from B were quantified using ImageJ software. The band intensities of each protein were normalized to GAPDH (1 biological replicate).

3.2.4 USP7 KNOCKDOWN DECREASES CDK8 & MED12 PROTEIN LEVELS

In the previous experiments, increase in the levels of CDK8 and MED12 were detected when Myc-USP7 concentration was increased in the cells. Thus, it would be interesting to see how the levels of CDK8 and MED12 change in response to decreasing the endogenous levels of USP7. This was investigated by doing a transient knocking down of USP7 using siRNA. Preliminary trials have resulted in successful knockdown of USP7 when cells were harvested 24, 48 and 72 hours post initial siRNA transfection. However, no changes were detected in the levels of our target proteins (data not shown). As a result, we have decided to transfect the siRNA, twice, once at t=0 hours and then again at t=24 hours and harvest the cells 12-24 hours post the second siRNA transfection. When U2OS cells are double transfected with 434 pmol of siRNA against the USP7 mRNA, we see a successful decrease in the protein levels of USP7, compared to that found in the cells transfected with the scrambled siRNA, which serves as the negative control (NC) (Fig. 3.6 A & B). Quantifying the USP7 blots obtained from the knockdown from three biological replicates produced a mean decrease to 0.37-fold ± 0.065 (Mean ± SEM), compared to the cells transfected with scrambled siRNA. Paired T-test was also conducted, giving a p-value of 0.00031, indicating statistical significance in the protein levels between the cells transfected with scrambled siRNA versus siRNA targeting USP7 (Fig 3.8 A). Interestingly, the levels of CDK8 are also reduced when USP7 is knocked down, as shown by the fainter protein band, in comparison to our negative control lane (Fig 3.7 A). Quantification of the CDK8 blots from three biological replicates showed an average decrease of 0.63-fold ± 0.038, with a p-value of 0.0049, also indicating statistical significance (Fig 3.8 A). Furthermore, the levels of MED12 In U2OS cells are also reduced (fainter protein band) when USP7 is knocked down, in comparison to our negative control lane (Fig 3.7 B). 4 biological replicates of MED12 blots were quantified using

ImageJ, producing an average decrease of MED12 levels to 0.40 folds ± 0.14. The p-value

obtained for this quantification was 0.012, also indicating statistically significant difference (Fig

3.8 A).

Finally, USP7 knockdown was also investigated in the HeLa cell line. HeLa cells were double

transfected with 434 pmol of siRNA against the USP7 mRNA or scrambled siRNA as the

negative control. Consistent with the results seen with U2OS cells, USP7 is successfully knocked

down when the siRNA against our target gene is present (Fig 3.7 C). Quantification of two

biological replicates for USP7 knockdown showed an average decrease of USP7 protein levels to

0.33 ± 0.0045 (Mean ± SEM). The p-value obtained for this knockdown from a paired T-test was

0.0021 (Fig 3.8 B). This denotes statistically significant difference in USP7 levels from

nonspecific verses USP7-specific knockdown. Additionally, we also see a reduction in the

protein band intensity of CDK8 when the cells are transfected with siRNA against USP7

compared to our negative control (Fig. 3.7 C). Further quantification of the CDK8 blots from two

biological replicates showed an average decrease of CDK8 levels to 0.66 ± 0.040. This data

showed statistical significance as the p-value determined for the data set was 0.038 (Fig 3.8 B)
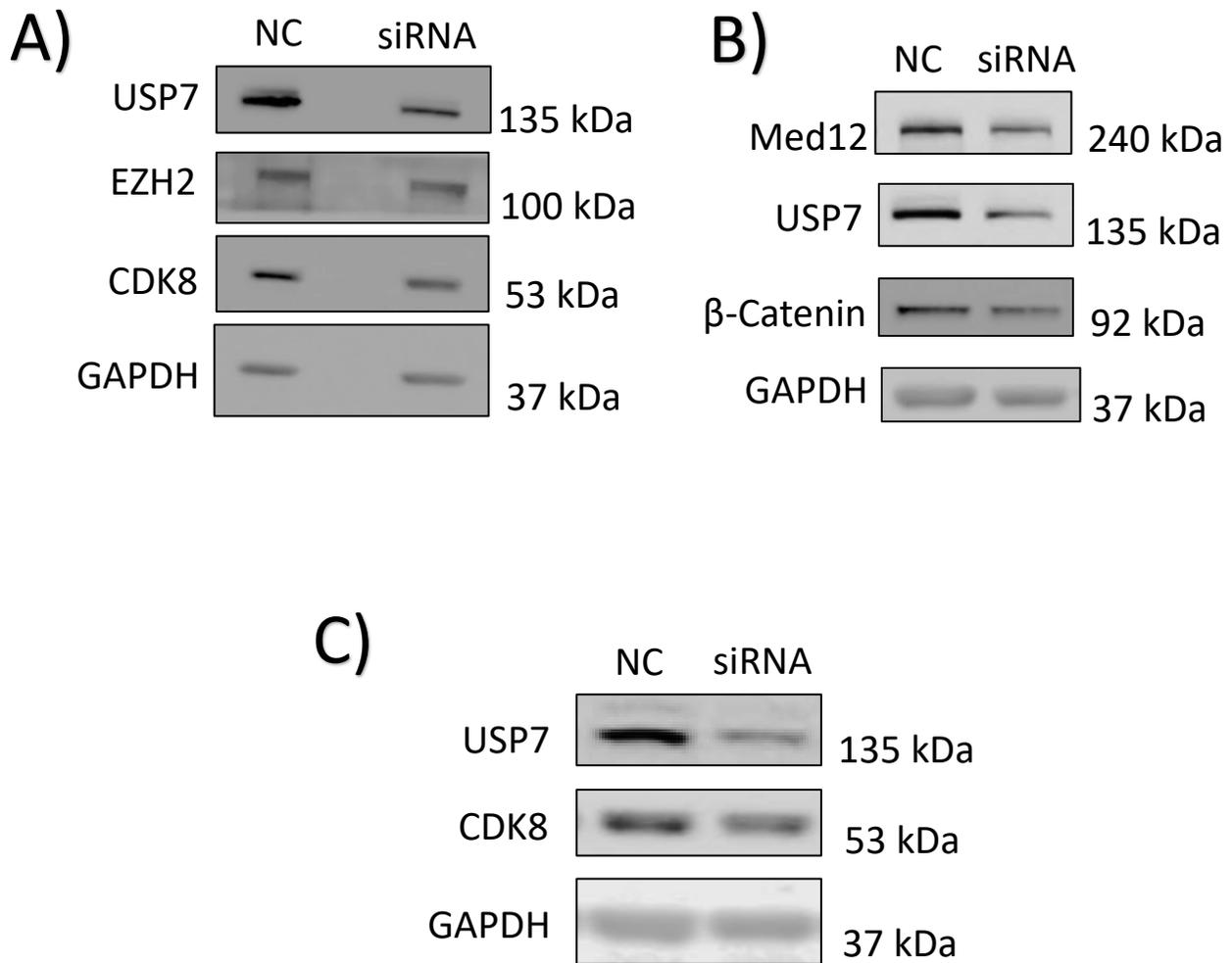
Figure 3.7**. USP7 knockdown decreases stability of CDK8 and MED12**. U2OS (A&B) and C) HeLa cells were double transfected with either 434 pmol of scrambled siRNA (negative control-NC) or siRNA against USP7. Cells were harvested 12-24 hours post second transfection. 20 µg of cellular lysates was resolved on 10% SDS-PAGE gel and transferred onto PVDF membrane via western blotting. The proteins were blotted for using rabbit anti-MED12, rabbit anti-CDK8, mouse anti-USP7, rabbit anti-β-catenin, mouse anti-EZH2, and mouse anti-GAPDH. A) N=3, B) N=4, C) N=2.
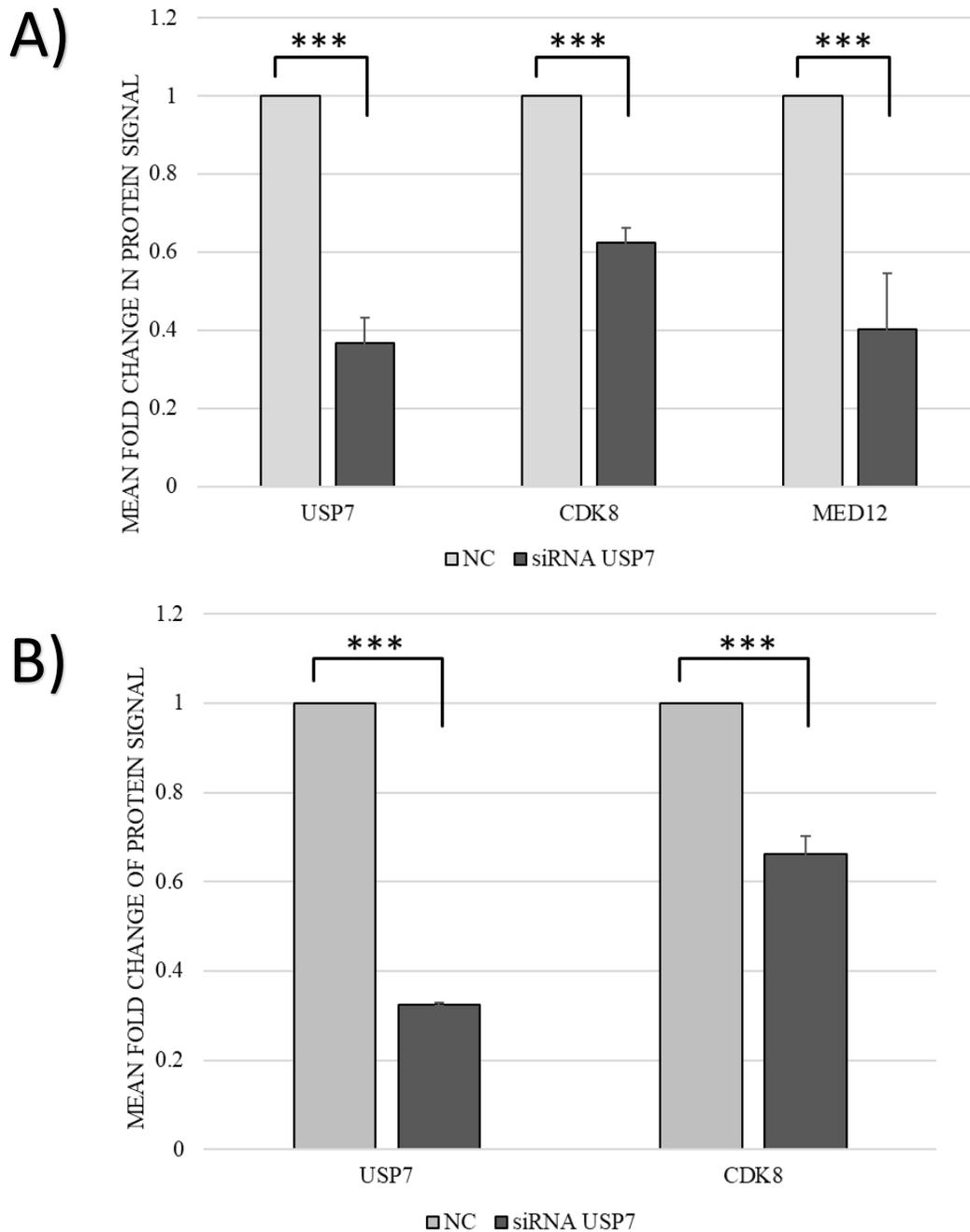
Figure 3.8. **Relative protein levels of CDK8 and MED12 as a response to USP7 siRNA.** A) the blots U2OS cells from Figure 3.7 A and B were quantified using ImageJ software and the mean fold change level of protein signals were determined as ratio of cells transfected with scrambled siRNA (NC) to siRNA targeting USP7. The ratios were then normalized to GAPDH (CDK8= 3 biological replicates, MED12= 4 biological replicates in U2OS). B) The levels of CDK8 and USP7 obtained from USP7 siRNA in HeLa cells (Fig 3.7 C) were quantified and plotted similarly as done in U2OS cells. CDK8=2 biological replicates. A paired T-test was conducted to evaluate statistical significance in the level changes of USP7, CDK8, and MED12 ($p < 0.05$, indicated by ***). Error bars represent Standard Error of the Mean (SEM).

3.2.5 CDK8 STABILIZATION IMPACTS THE PROTEIN'S DOWNSTREAM TARGETS

Previous research has shown that CDK8 can phosphorylate the CTD of RNAP II at serine 5 (Ser5P) (Galbraith *et al.,* 2010). Thus, for this experiment, we will examine whether the phosphorylation levels of Ser5 change when the abundance of CDK8 is compared in the cell lines containing a wild-type USP7 (HCT116) vs lacking a functional USP7 gene (HCT116 USP7 [-/-]). Furthermore, we will also examine phosphorylation of Ser2, another residue found in the CTD of RNAP II that is not predominantly phosphorylated by CDK8, but rather by CDK9 (Nekhai *et al.,* 2014); this will act as a negative control. Additionally, CDK8 overexpression has also been shown to correlate with increased expression of β-catenin. Thus, β-catenin was also blotted for to see if any changes can be detected in its level. However, we must note that β-catenin is also a substrate of USP7 and so any changes detected in the protein levels of β-catenin might be attributed to both factors. As seen in Figure 3.9, when USP7 is absent, as seen in HCT116 USP7[-/-] cell line, the levels of CDK8 are reduced; the protein band is less intense than that found in the USP7 wild-type lane. In turn, we also see a reduction in the phosphorylation of Ser5 residue on CTD of RNAP II; the detected protein band in the USP7 knockout lane is less than that found in the USP7 wild-type lane. However, we do not see any changes in phosphorylation level of Ser2 residue when CDK8 levels are reduced as seen in the HCT116 USP7[-/-] lane. Furthermore, the levels of EZH2, a positive control, is also reduced when USP7 is lacking in the cell, as shown in the HCT116 USP7[-/-] compared to the HCT116 USP7 wild-type lane. Finally, the β-catenin band intensity is also reduced in the absence of a functional USP7 protein. Now, as mentioned before, the changes in the protein levels of β-catenin could be a result of changes in the levels of both CDK8 and USP7 in the two cell types.
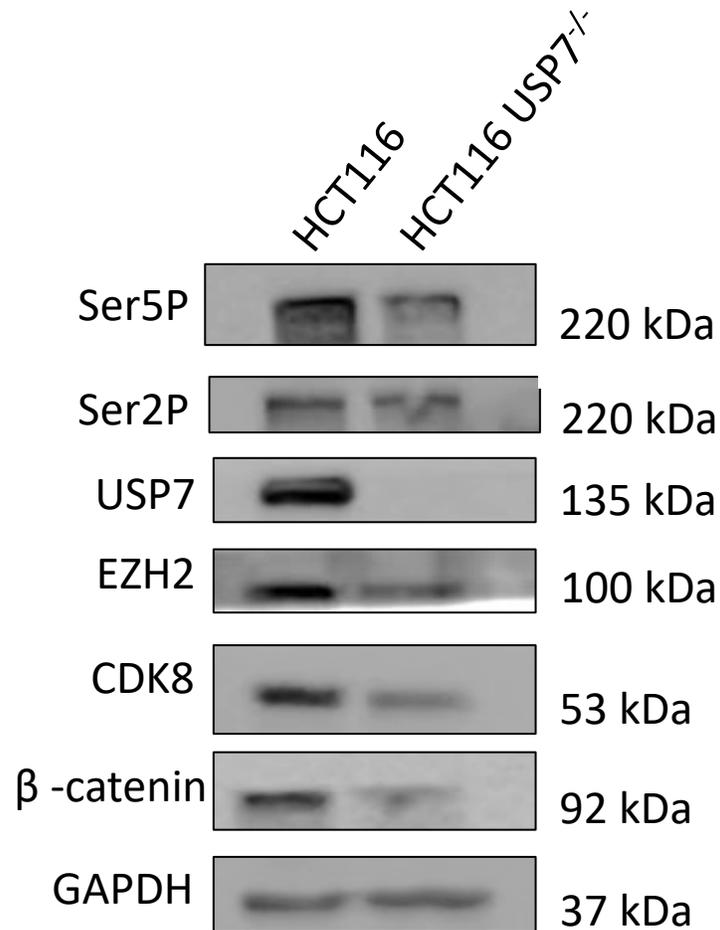
Figure 3.9. **USP7 Interplay on -CDK8 Levels and the Effect on CDK8 Downstream Targets**.
HCT116 and HCT116 USP7$^{-/-}$ cells were lysed, and 20 µg of lysate was resolved and blotted for
the indicated proteins via mouse anti-Ser5P, rat anti-Ser2P, mouse anti-USP7, mouse anti-EZH2,
rabbit anti-CDK8, rabbit anti- β-catenin, and mouse anti-GAPDH antibodies. Mouse anti-Ser5P
and rat anti-Ser2P detect phosphorylation marks on residues Ser5 and Ser2 on the CTD of RNAP
II.

3.3 CDK8 & MED12 INTERACT WITH NTD OF USP7 IN HIS-PULLDOWN ASSAYS

To determine whether USP7 interacts with CDK8, His-pulldown assay will be used to investigate the interaction in vitro. The His-USP7 constructs, His-USP7 NTD and His-USP7 NTD DW$^{\Delta}$ were grown and expressed in *E. coli* BL21 mgk Cells. The proteins were purified using His-affinity chromatography and 300 µg of purified His-USP7 constructs were used to perform the His-pulldown with CDK8. The His-USP7 NTD DW$^{\Delta}$ is a mutant construct that has a mutation at the DWGF motif of USP7's N-terminus.

When the wild-type NTD of USP7 is pulled down, endogenous CDK8 and MED12 are also pulled down with it. Similarly, the positive control, Bmi1 can also be seen co-eluting with the His-USP7 NTD construct (Fig 3.10 A) However, when the mutant His-USP7 construct is pulled down, there is less CDK8 and MED12 co-eluting with the mutant protein as indicated by the faintness of the protein bands respectively. Furthermore, Bmi1 is completely absent when His-USP7 NTD DW$^{\Delta}$ is pulled down. The empty beads lane (EB) serves as a negative control; the reaction tube had empty Nickel beads incubated with HCT116 lysate without any His-USP7 constructs.

In addition, we also tested this assay by overexpressing FLAG-CDK8 in the HCT116 cell line and the results are consistent. In the presence of the His-USP7 NTD DW$^{\Delta}$ mutant construct, FLAG-CDK8 is not co-eluted with the mutant. However, FLAG-CDK8 is seen when the wild-type His-USP7 NTD domain is present (Fig 3.10 B). EZH2 also follows a similar pattern, which is a known substrate of USP7. This experiment was also tried several times with CTD constructs of USP7 as bioinformatic tools show that both MED12 and CDK8 have potential consensus sequence for interacting with the CTD of USP7. However, in all attempts, CDK8 and MED12 were not pulled down with either the wild-type CTD of USP7 or the mutant construct of

CTD of USP7, His-USP7 MRGR$^{\Delta}$ (data not shown). Thus, this suggests that both CDK8 and MED12 might prefer to interact with USP7 using the USP7 MATH domain found in the N-terminus.
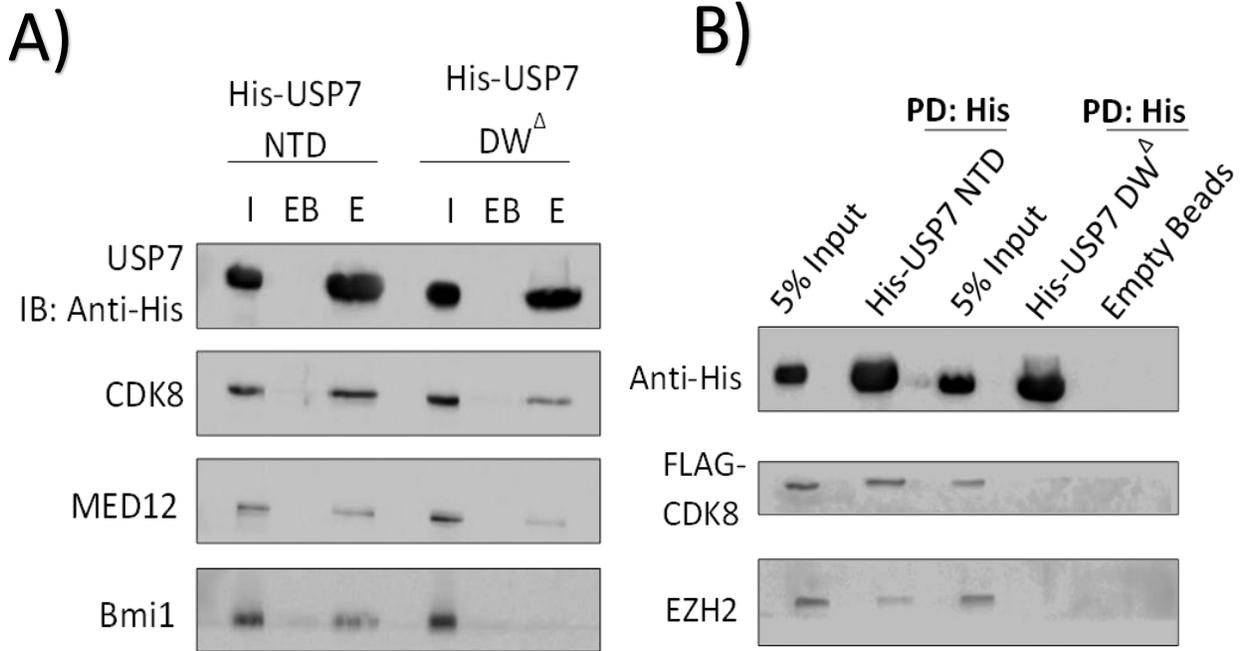
Figure 3.10. **USP7 His pulldown assay**. His-USP7 NTD constructs (WT and DW mutant) were grown, expressed, and purified from *E. coli* BL21 mgk Cells. A) 300 µg of purified His-USP7 constructs were incubated with 600 µg of wild-type HCT116 lysate overnight at 4°C. The reaction tubes were then washed and eluted with elution buffer. Eluted samples were resolved, and western blot was performed to blot for endogenous CDK8, MED12, and Bmi1 (positive control). B) HCT116 cells were transfected with 10 µg of FLAG-CDK8. 500 µg of lysate was incubated with 300 µg of purified USP7 NTD His construct overnight at 4°C. Reaction tubes were washed and then eluted using elution buffer. FLAG-CDK8 was blotted for using mouse anti-FLAG, mouse anti-His for His-USP7 constructs, and endogenous EZH2 (positive control) was blotted for using mouse anti-EZH2.

3.4 CO-IMMUNOPRECIPITATION OF USP7 WITH CDK8 & MED12

To determine whether USP7 physically interacts with CDK8 and MED12 *in vivo*, various co-immunoprecipitation studies were performed in various cell lines including HeLa, U2OS, and HCT116 USP7 wild-type cells.

FLAG-CDK8 and Myc-USP7 were co-expressed in U2OS cell line. The cells were lysed and Myc-USP7 was immunoprecipitated using mouse monoclonal anti-Myc tag; mouse immunoglobulin G (IgG) was used as a negative control for any non-specific interactions. The resulting immunoprecipitated reactions were analyzed by western blotting by blotting for FLAG-CDK8 using anti-FLAG and Myc-USP7 using anti-Myc. As seen in Figure 3.11 A, FLAG-CDK8 coimmunoprecipitates with Myc-USP7 as indicated by the faint grey protein band. Furthermore, a reciprocal coimmunoprecipitation was performed, whereby FLAG-CDK8 was immunoprecipitated using mouse monoclonal anti-FLAG and the endogenous USP7 was blotted for using rabbit polyclonal anti-USP7. Mouse IgG was used as a negative control. As shown in Figure 3.11 B endogenous USP7 coimmunoprecipitates with FLAG-CDK8.

Endogenous coimmunoprecipitation studies were also conducted in HeLa cell line. HeLa cells were harvested and lysed. Endogenous USP7 was immunoprecipitated using rabbit polyclonal anti-USP7 and rabbit IgG was used for negative control. Western blot was performed on the eluted fractions and the proteins of interest were blotted for using rabbit polyclonal anti-CDK8, and rabbit polyclonal anti-MED12. When USP7 is immunoprecipitated, both MED12 and CDK8 are also co-immunoprecipitated as shown in Figure 3.11 C There was not any significant nonspecific interaction as shown with the empty IgG lane. Thus, both CDK8 and MED12 interact with USP7 endogenously. In addition, reciprocal coimmunoprecipitation was performed in HeLa cells whereby endogenous MED12 was immunoprecipitated using rabbit polyclonal anti-MED12, and anti-CDK8 and anti-USP7 antibodies were used to elute for our proteins of interest. As shown in Figure 3.11 D, USP7 coimmunoprecipitates with MED12, like CDK8; CDK8 acts as a positive control for this experiment.

Finally, endogenous coimmunoprecipitation was also performed in HCT116 USP7 wild-type

cells and the results show consistency. When USP7 was immunoprecipitated using mouse

monoclonal anti-USP7, both CDK8 and MED12 were co-immunoprecipitated with USP7

(Fig 3.11 E). As a result, all the coimmunoprecipitation studies in the different cell lines show

that USP7 interacts with both CDK8 and MED12 and the interaction is reciprocal.

Further co-immunoprecipitation studies were also performed to pinpoint whether mutations to

the DWGF motif in the MATH domain of USP7 alter its interaction with CDK8 *in vivo*. The

following constructs were transfected into HCT116 cells: pcDNA 3.1 (FLAG EV),

FLAG-CDK8, pCAN/Myc-USP7, pCAN/Myc-USP7 MRGR (mutation in the CTD binding

motif), pCAN/Myc-USP7 MRGR DW (mutation in both the CTD and NTD substrate binding

motif).

FLAG-immunoprecipitation was conducted to immunoprecipitate FLAG-CDK8 and then the

USP7 constructs were blotted for using western blotting. In Figure 3.12, both the wild-type

Myc-USP7 and the Myc-USP7 MRGR mutant coimmunoprecipitated with FLAG-CDK8. This

shows that mutations to the MDGD motif in the CTD of USP7 does not compromise binding of

CDK8 to USP7. However, the double mutant, Myc-USP7 MRGR DW, did not

coimmunoprecipitate with FLAG-CDK8. Thus, the mutation in the DWGF motif diminished the

interaction between CDK8 and USP7.

The reciprocal coimmunoprecipitation was also performed. HCT116 UPS7 wild-type cells were

transfected with either Myc-USP7, Myc-USP7 MRGR mutant, or Myc-USP7 MRGR DW

double mutant. The Myc-USP7 constructs were immunoprecipitated using mouse monoclonal

anti-Myc and CDK8 was blotted for via western blotting. Consistent with the FLAG

coimmunoprecipitation. Endogenous CDK8 can be seen coimmunoprecipitating with both

Myc-USP7 and Myc-USP7 MRGR mutant. However, CDK8 did not coimmunoprecipitate with the Myc-USP7 MRGR DW mutant (Fig 3.13). As a result, the coimmunoprecipitation studies shows that the binding of CDK8 to USP7 is dependent on the presence of functional DWGF motif in the N-terminus of USP7. Furthermore, mutations in the MDGD motif in the C-terminus of USP7 does not impact CDK8 binding. This suggests that CDK8 mediates its interaction with USP7 mainly through the USP7 N-terminus.
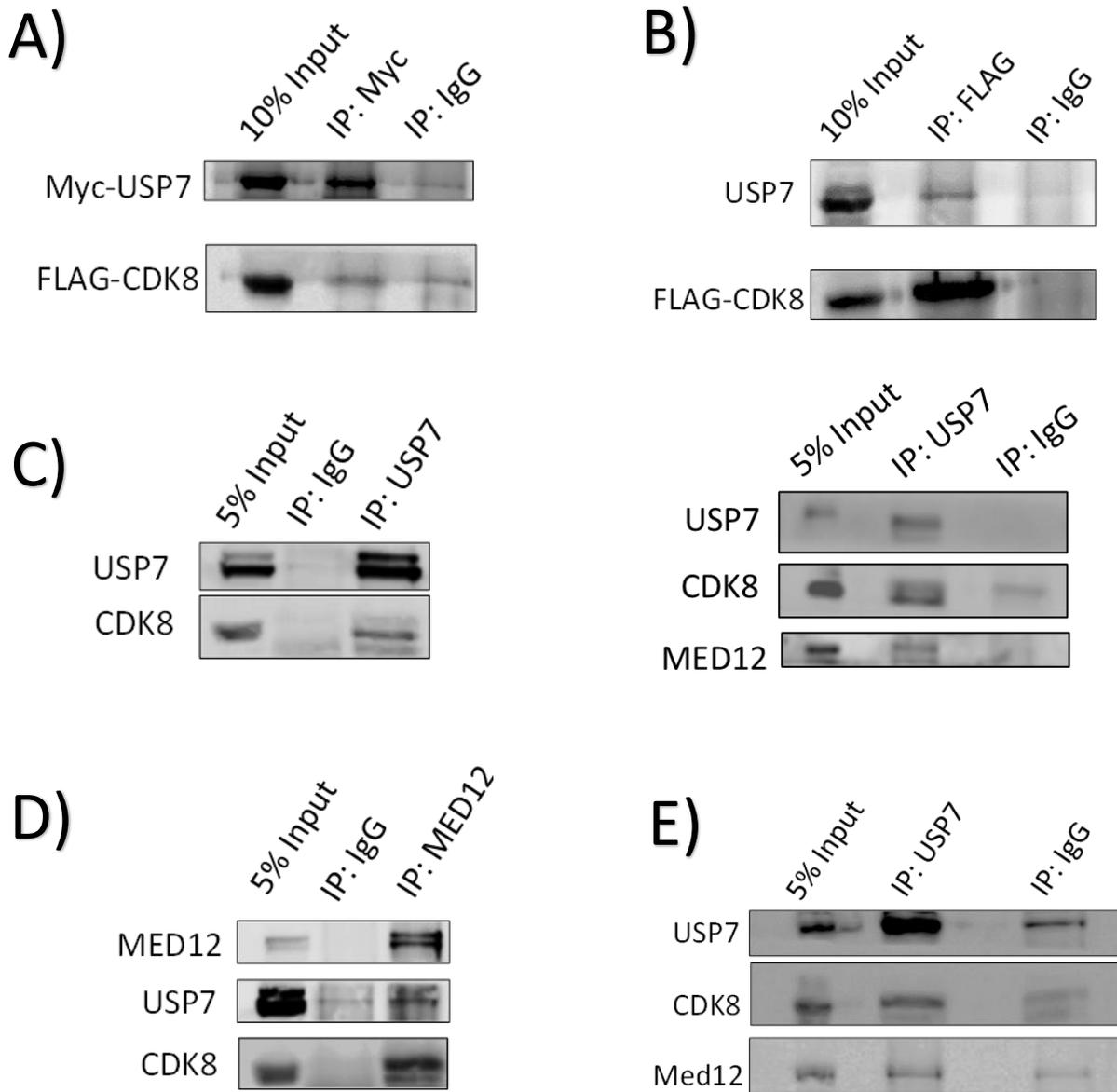
Figure 3.11. **Coimmunoprecipitation studies of USP7, CDK8, and MED12.** U2OS cells were transfected with Myc-USP7 and FLAG-CDK8. A) Myc-USP7 was immunoprecipitated using mouse anti-myc and FLAG-CDK8 was blotted for using rabbit anti-FLAG. B) FLAG-CDK8 was immunoprecipitated using mouse anti-FLAG and endogenous USP7 was blotted for using rabbit anti-USP7. C) Endogenous coIP was replicated in HeLa cells whereby USP7 was immunoprecipitated and endogenous CDK8 and MED12 were blotted for. D) Reciprocal coIP was performed in HeLa, whereby MED12 was immunoprecipitated while both endogenous CDK8 and USP7 were blotted for. CDK8 is a positive control. E) Endogenous coIP was also replicated in HCT116 cells by immunoprecipitating USP7 and blotting for endogenous CDK8 and MED12.
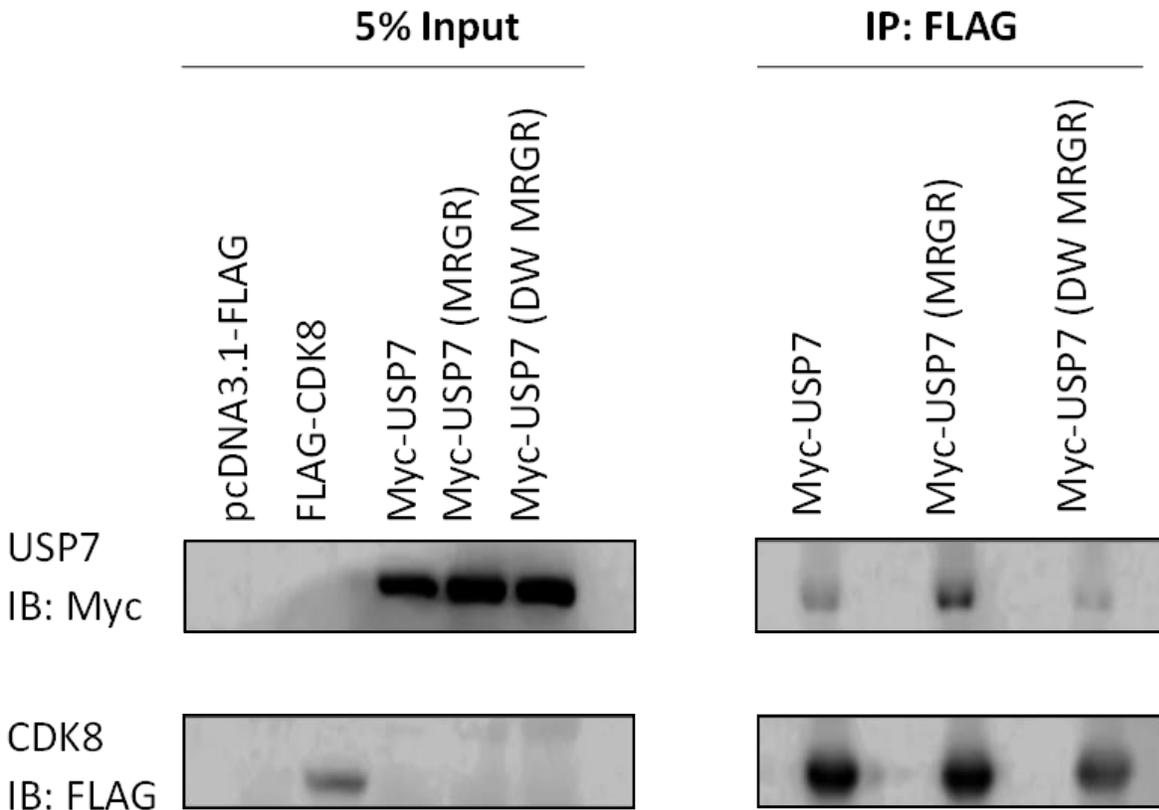
Figure 3.12. **FLAG co-immunoprecipitation of FLAG-CDK8 and USP7 constructs.** HCT116 cells were transfected with FLAG empty vector (pcDNA3.1), FLAG-CDK8, Myc-USP7 WT, Myc-USP7 MRGR, Myc-USP7 MRGR DW double mutant. 350 µg of FLAG-CDK8 lysate or EV was incubated with 750 µg of either Myc-USP7, Myc-USP7 MRGR, or Myc-USP7 MRGR DW. Reactions were incubated overnight at 4 degrees. Post incubation, samples were washed, eluted, resolved, and blotted for using rabbit anti-FLAG and mouse anti-Myc.
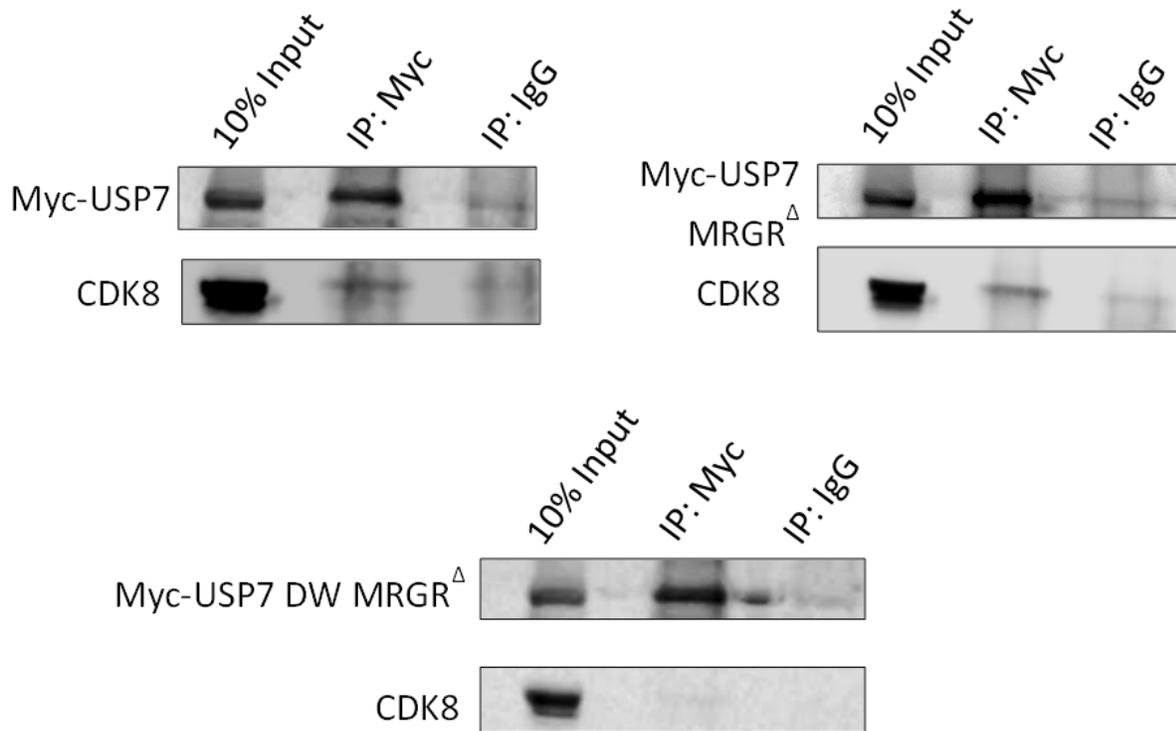
Figure 3.13. **Coimmunoprecipitation of CDK8 with Mutant USP7 Constructs**. HCT116 cells were transfected with either 5 µg of Myc-USP7, Myc-USP7 MRGR mutant, or Myc-USP7 MRGR DW double mutant. Cells were lysed and the USP7 constructs were immunoprecipitated using mouse anti-myc. Mouse IgG was also used as the negative control. The immunoprecipitation reaction was incubated overnight, subsequently washed with 1X PBS (3x), and eluted using 5x SDS. The eluted fractions were resolved on 10% SDS-PAGE gel, transferred onto PVDF membrane for 1 hour and 30 minutes at 100 volts. The proteins were then blotted using rabbit anti-CDK8 and mouse anti-Myc.

CHAPTER 4: DISCUSSION

The purpose of this research project was to investigate and study the hypothesized regulatory effect and interaction between USP7 and two components of the CDK8 module, CDK8 and MED12 proteins. Bioinformatic analysis suggested that both proteins contain a potential USP7 binding consensus sequence motif, P/A/ExxS. Our *in vitro* and *in vivo* studies provided strong evidence that both CDK8 and MED12 interact with the N-terminal domain of USP7 and that through this interaction, CDK8 and MED12 levels are maintained in the various cell lines examined: U2OS, HeLa, HCT116, and HCT116 USP7$^{-/-}$.

4.1 USP7 REGULATES THE PROTEIN LEVELS OF CDK8 AND MED12 *IN VIVO*

To determine if CDK8 and MED12 were novel substrates of USP7, we decided to first investigate the endogenous levels of the proteins in the absence of USP7. Immunoblotting for CDK8 and MED12 in HCT116 USP7 knockout cells showed a significant reduction in the levels of both proteins, when the blot signals were normalized to the parental cell line that had a functional USP7 protein (HCT116) (Fig. 3.2). Such preliminary result suggested that a regulatory relation might exist between USP7 and our proteins of interest. Further studies were conducted to examine the effects of reintroducing USP7 back in the HCT116 USP7 knockout cells. As anticipated, there was a significant increase in the protein levels of MED12 when USP7 was reintroduced. Now, the quantified protein signal of CDK8 did not pass the T-test for statistical significance as the reported p-value was 0.067. However, qualitative data (Fig 3.3 A, C) did show that CDK8 levels increased when USP7 was reconstituted in the HCT116 USP7 knockout cells. Furthermore, we also showed that the differences in the protein levels = mediated by USP7 is dependent on that catalytic activity of the protein. Transfecting a catalytically inactive USP7 protein harbouring a point mutation at the cysteine residue that changes it to

serine (C223S) decreased CDK8 and MED12 levels when compared to cells transfected with the

wild-type Myc-USP7. Furthermore, the signal bands of CDK8 and MED12 from cells

transfected with the catalytic mutant was similar to the cells transfected with the empty vector

only. This observation was shown across the cell lines: HCT116 USP7$^{-/-}$, HCT116 cells, and

HeLa (Fig 3.4, 3.5). Because USP7 is a deubiquitinating enzyme, these observations suggest

indirect evidence that USP7 stabilizes the levels of CDK8 and MED12 by deubiquitinating them,

to prevent their degradation by the 26S proteosome complex. Disrupting the catalytic core via

cysteine-serine mutation prevents the catalysis of ubiquitin from our target proteins, making

them more susceptible to proteasomal degradation. Thus, the catalytic activity of USP7 is

important for the protein abundance of CDK8 and MED12 in the cells.

Finally, our last sets of experiments further supported our hypothesis that a regulatory

relationship exists between USP7 and our proteins of interest. As seen in Figure 3.6, increasing

the concentration of USP7 from 0 µg to 5 µg or 0 µg to 15 µg, increased the protein levels of

both CDK8 and MED12. Conversely, decreasing the levels of USP7 via siRNA knockdown

significantly reduced the levels of CDK8 in both U2OS and HeLa cells, and MED12 in U2OS

USP7$^{-/-}$ cells (Fig 3.8). Therefore, based on the obtained results, the protein abundance of CDK8

and MED12 are dependent on the catalytic function of USP7.

4.2 USP7 INTERACTS WITH CDK8 & MED12 *IN VITRO* AND *IN VIVO*

Endogenous co immunoprecipitation studies revealed that USP7 can bind to and

immunoprecipitate CDK8 and MED12 from cellular lysate in various cell lines such as HCT116,

U2OS, and HeLa. This indicates that the interaction is not cell line specific but occurs across

different cell lines. Upon initial co-immunoprecipitation studies, we suspected that MED12 or

CDK8 might have been tagging along with one another when USP7 was immunoprecipitated. As

mentioned previously, MED12 interacts with cyclin C-CDK8 dimer and enhances CDK8 activity

(Park *et al.,* 2018), providing a potential reason for why MED12 or CDK8 might have also

showed up when USP7 was immunoprecipitated. However, through reciprocal

coimmunoprecipitation studies, whereby either CDK8 or MED12 were immunoprecipitated, we

were able to show that USP7 was also co-immunoprecipitated, suggesting a direct and reciprocal

interaction between the proteins (Fig 3.11 B-D). In some of the endogenous coIPs however, there

seemed to be a faint protein band of CDK8. This could be due to the possibility that since USP7

interacts with various other substrates in the cell (Al-Eidan *et al.,* 2020), such interactions can act

as a competitive inhibitor for the binding of CDK8. Thus, it was important to exogenously

express CDK8, USP7, or both, and repeat the co-immunoprecipitation studies. Similar to the

endogenous coIPs, exogenously expressed FLAG-CDK8 or Myc-USP7 showed reciprocal

interaction between the two proteins (Fig 3.11 A-B).

Once we have established that a reciprocal interaction occurs between MED12-USP7 and

CDK8-USP7, the next step was to determine which regions in USP7 facilitated substrate

binding. As mentioned earlier, USP7 possess two functional motifs that direct substrate binding:

the [164]DWGF[167] motif at the NTD, which recognizes substrates with the consensus sequence

P/A/ExxS (Pozhidaeva and Bezsonova, 2019), or the [758]DELMDGD[764] motif at the CTD, which

recognizes the consensus sequence KxxxK for substrate binding (Pfoh *et al.,* 2015). *In vitro*-pull

down assays were conducted to verify the sites of interaction. Purified His-USP7 NTD pulled

down endogenous CDK8 and MED12. Mutations to the [164]DWGF[167] as presented in the His-

USP7 DW$^\Delta$ construct reduced the interactions of both CDK8 and MED12 (Fig 3.10 A). It is

interesting to note however that the interaction is not 100% depleted as seen with Bmi1. One

possible explanation could be that other residues in USP7, CDK8 or MED12 might be playing a

role in somewhat stabilizing the interactions between the proteins, especially since CDK8 and

MED12 associate with one another. The importance of a functional ^{164}DWGF^{167} motif was also

examined *in vivo*. FLAG-coimmunoprecipitation was conducted to determine if mutations to

either the ^{164}DWGF^{167} motif or ^{758}DELMDGD^{764} abrogated substrate binding. HCT116 cells

were transfected with FLAG-CDK8, Myc-USP7 wild-type, Myc-USP7 MRGR mutant (Mutation

in the CTD), and a double mutant construct, Myc-USP7 MRGR DW, whereby both binding

motifs are mutated. Consistent with the in vitro pull-down results, FLAG-CDK8 was shown to

interact with both, the wild-type Myc-USP7 and as well as the MRGR mutant. However, the

interaction was diminished in the presence of the Myc-USP7 MRGR DW double mutant (Fig

3.12). A reciprocal coIP was also performed with the same constructs and similar results were

observed. Endogenous CDK8 was shown to interact with both the wild-type Myc-USP7 and the

Myc-USP7 MRGR mutant. Furthermore, interaction was completely diminished when the Myc-

USP7 MRGR DW double mutant was immunoprecipitated (Fig 3.13). This suggests that CDK8

utilizes the ^{164}DWGF^{167} for binding to USP7, rather than the ^{758}DELMDGD^{764} motif found at the

CTD. Mutations to the ^{758}DELMDGD^{764} motif did not alter the binding capacity of CDK8, but

mutations to the ^{164}DWGF^{167} motif as presented in the Myc-USP7 MRGR DW mutant prevented

the interaction of CDK8 with USP7. Similar observation was reported by Sheng *et al.,* 2006,

whereby mutations to D164 and W165 residue in the ^{164}DWGF^{167} abrogated the protein's ability

to interact with p53, MDM2 and EBNA1 peptide. All in all, both CDK8 and MED12 interact

with USP7 using the NTD and this interaction requires the presence of a functional ^{164}DWGF^{167}

motif.

4.3 USP7 AFFECTS CDK8 MEDIATED PHOSPHORYLATION ACTIVITY

After showing that USP7 interacts with and regulates the protein levels of CDK8 in the cells, we then wanted to determine if such interaction has any effects on the downstream targets of CDK8, more specifically, the phosphorylation of Ser5 found in the CTD of RNAP II. The CTD of RNAP II contains conserved heptapeptide repeat, Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7 (YSPTSPS) that can be differentially phosphorylated by various kinases (Rickert 1998). CDK8 has been shown to associate with the CTD and phosphorylate

Ser5 residue *in vitro* (Rickert 1998).

We predicted that if CDK8 is stabilized by USP7, then reduction in USP7 levels leads to reduction in CDK8 levels, which might potentially lead to reduction in the phosphorylation of Ser5p of the CTD of RNAP II. As seen in Figure 3.9, when USP7 is absent in the HCT116 USP7$^{-/-}$ cells, the levels of CDK8 are reduced, compared to the levels found in the parental cell line with functional USP7. Furthermore, we also see a reduction in the phosphorylation of Ser5p as detected by the antibody compared to the HCT116 lane.

However, it is important to note that CDK8 might not significantly impact the phosphorylation of Ser5p in a direct way. Another kinase that has been shown to phosphorylate serine 5 residue on CTD of RNAP II *in vivo* is Cyclin dependent kinase 7 (CDK7), the catalytic subunit of TFIIH complex (Egloff *et al.,* 2012). CDK7-Cyclin H is also associated with the transcriptional machinery and serves to prominently phosphorylate the CTD of RNAP II for promotor clearance.

Another limitation to this experiment is the inability to determine the transcriptional status of the cells by only examining the western blots. Phosphorylation of Ser5 residue on the CTD are

highest at the promotors, just before elongation takes place, as such phosphorylation event is needed for promotor clearance (Rickert 1998). Although there seems to be a significant difference in the intensity of Ser5 phosphorylation in between the two cell types, we cannot be 100% certain that this is mainly attributed to the reduced levels of CDK8 in the HCT116 USP7 $^{-/-}$. Thus, because USP7 acts on various substrates in the cells, the decrease in the global phosphorylation of Ser5 might be due to other factors that are currently not fully understood.

4.4 FUTURE WORKS AND DIRECTIONS

4.4.1 The gene expression levels of CDK8 Kinase module targets

Since many of the downstream effectors governed by CDK8 are also regulated by USP7 directly or by other kinases, it is challenging to determine whether the changes detected in the protein levels or their phosphorylation are mainly due to CDK8, USP7 or other factors. For this reason, we will look at the changes that take place at the mRNA levels of some of the downstream effectors, in which their expression is regulated by the CDK8 kinase module to determine whether increased/decreased stability of CDK8 and MED12 by USP7 changes the expression levels of some of the genes. More specifically, it would be interesting to investigate the gene expression profile induced by β-catenin as MED12 interacts with β-catenin and recruits the Mediator complex to facilitate transcription of Wnt-target genes. Furthermore, CDK8 also represses the inhibitory grip that E2F1 has on β-catenin, allowing it to go and stimulate the transcription of Wnt-target genes. Thus, we can investigate the expression of c-Myc, cyclin D1, and Axin using RT-qPCR in HCT116 cells vs HCT116 USP7$^{-/-}$ cells.

4.4.2 *In vitro* and *in vivo* deubiquitination assay

Although we have shown that USP7 interacts with and stabilizes both CDK8 and MED12, it would also be interesting to see whether USP7 can de-ubiquitinate CDK8 and MED12 both *in vivo* and *in vitro*. This will establish if CDK8 and MED12 are both novel substrates of USP7, who's stability is dependent on deubiquitination.

4.5 CONCLUSION

In conclusion, we have shown that CDK8 and MED12 are novel interactors of USP7. This research sheds light on how CDK8 and MED12 are regulated in the cell, and this is via their association with the deubiquitinating enzyme, USP7. USP7 interacts with both proteins using the DWGF motif found in the MATH domain of its N-terminus. This interaction allows for USP7 to maintain the levels of CDK8 and MED12 and this effect is dependent on its catalytic activity. Furthermore, any changes in the protein levels of USP7 result in direct changes in the levels of both MED12 and CDK8. When the concentration of USP7 is gradually increased, the protein levels of both MED12 and CDK8 also increase, which suggests that USP7 is meditating their stability within the cells. Furthermore, decreasing USP7 levels via siRNA knockdown also results in a decrease in both MED12 and CDK8 protein levels, which suggests that their stability is dependent on the presence of USP7. Such finding is essential because this highlights a role for USP7 in regulating proteins from the Mediator complex, which is important for transcriptional and non-transcriptional regulation.

Knowing that CDK8 is amplified in some cancer types such as breast cancer (Wu *et al.,* 2020), and colorectal cancer (Firestein *et al.,* 2008), and that its levels are dependent on USP7, one can take advantage of this interaction to perhaps come up with therapeutics that can hinder the

interaction between USP7 and CDK8, compromising the stability of the protein, and subsequently the action of the kinase module. This can hinder the module's ability from facilitating transcription of target oncogenes. This specifically becomes important in the Wnt signaling pathway as CDK8 and MED12 have been shown to stimulate β-catenin transcriptional activity. CDK8 phosphorylates and inhibits E2F1 from repressing β-catenin and MED12 recruits the Mediator complex to the β-catenin target locus, thus promoting tumorigenicity.

REFRENECES

1. Adler, A.S., McCleland, M.L., Truong, T., Lau, S., Modrusan, Z., Soukup, T.M., Roose-Girma, M., Blackwood, E.M. and Firestein, R., 2012. CDK8 maintains tumor dedifferentiation and embryonic stem cell pluripotency. *Cancer research*, *72*(8), pp.2129-2139.

2. Akoulitchev, S., Chuikov, S. and Reinberg, D., 2000. TFIIH is negatively regulated by cdk8-containing mediator complexes. *Nature*, *407*(6800), pp.102-106.

3. Al-Eidan, A., Wang, Y., Skipp, P. and Ewing, R.M., 2020. The USP7 protein interaction network and its roles in tumorigenesis. *Genes & Diseases*.

4. Allen, B.L. and Taatjes, D.J., 2015. The Mediator complex: a central integrator of transcription. *Nature reviews Molecular cell biology*, *16*(3), pp.155-166.

5. Bancerek, J., Poss, Z.C., Steinparzer, I., Sedlyarov, V., Pfaffenwimmer, T., Mikulic, I., Dölken, L., Strobl, B., Müller, M., Taatjes, D.J. and Kovarik, P., 2013. CDK8 kinase phosphorylates transcription factor STAT1 to selectively regulate the interferon response. *Immunity*, *38*(2), pp.250-262.

6. Bhattacharya, S., Chakraborty, D., Basu, M. and Ghosh, M.K., 2018. Emerging insights into HAUSP (USP7) in physiology, cancer and other diseases. *Signal transduction and targeted therapy*, *3*(1), pp.1-12.

7. Bojagora, A. and Saridakis, V., 2020. USP7 manipulation by viral proteins. *Virus Research*, p.198076.

8. Brooks, C.L., Li, M., Hu, M., Shi, Y. and Gu, W., 2007. The p53–Mdm2–HAUSP complex is involved in p53 stabilization by HAUSP. *Oncogene*, *26*(51), pp.7262-7266.

9. Buetow, L. and Huang, D.T., 2016. Structural insights into the catalysis and regulation of E3 ubiquitin ligases. *Nature reviews Molecular cell biology*, *17*(10), pp.626-642.

10. Chavoshi, S., Egorova, O., Lacdao, I.K., Farhadi, S., Sheng, Y. and Saridakis, V., 2016. Identification of Kaposi sarcoma herpesvirus (KSHV) vIRF1 protein as a novel interaction partner of human deubiquitinase USP7. *Journal of Biological Chemistry*, *291*(12), pp.6281-6291.

11. Chen, Z.J. and Sun, L.J., 2009. Nonproteolytic functions of ubiquitin in cell signaling. *Molecular cell*, *33*(3), pp.275-286.

12. Cheng, J., Yang, H., Fang, J., Ma, L., Gong, R., Wang, P., Li, Z. and Xu, Y., 2015. Molecular mechanism for USP7-mediated DNMT1 stabilization by acetylation. *Nature communications*, *6*(1), pp.1-11.

13. Chou, J., Quigley, D.A., Robinson, T.M., Feng, F.Y. and Ashworth, A., 2020. Transcription-associated cyclin-dependent kinases as targets and biomarkers for cancer therapy. *Cancer discovery*, *10*(3), pp.351-370.

14. Dar, A., Shibata, E. and Dutta, A., 2013. Deubiquitination of Tip60 by USP7 determines the activity of the p53-dependent apoptotic pathway. *Molecular and cellular biology*, *33*(16), pp.3309-3320.

15. De Bie, P., Zaaroor-Regev, D. and Ciechanover, A., 2010. Regulation of the Polycomb protein RING1B ubiquitination by USP7. *Biochemical and biophysical research communications*, *400*(3), pp.389-395.

16. Ding, N., Zhou, H., Esteve, P.O., Chin, H.G., Kim, S., Xu, X., Joseph, S.M., Friez, M.J., Schwartz, C.E., Pradhan, S. and Boyer, T.G., 2008. Mediator links epigenetic silencing of

neuronal gene expression with x-linked mental retardation. *Molecular cell*, *31*(3), pp.347-359.

17. Ding, X., Jiang, W., Zhou, P., Liu, L., Wan, X., Yuan, X., Wang, X., Chen, M., Chen, J., Yang, J. and Kong, C., 2015. Mixed lineage leukemia 5 (MLL5) protein stability is cooperatively regulated by O-GlcNac transferase (OGT) and ubiquitin specific protease 7 (USP7). *PloS one*, *10*(12), p.e0145023.

18. Egloff, S., Zaborowska, J., Laitem, C., Kiss, T. and Murphy, S., 2012. Ser7 phosphorylation of the CTD recruits the RPAP2 Ser5 phosphatase to snRNA genes. *Molecular cell*, *45*(1), pp.111-122.

19. Everett, R.D., Meredith, M., Orr, A., Cross, A., Kathoria, M. and Parkinson, J., 1997. A novel ubiquitin-specific protease is dynamically associated with the PML nuclear domain and binds to a herpesvirus regulatory protein. *The EMBO journal*, *16*(7), pp.1519-1530.

20. Eychenne, T., Novikova, E., Barrault, M.B., Alibert, O., Boschiero, C., Peixeiro, N., Cornu, D., Redeker, V., Kuras, L., Nicolas, P. and Werner, M., 2016. Functional interplay between Mediator and TFIIB in preinitiation complex assembly in relation to promoter architecture. *Genes & development*, *30*(18), pp.2119-2132.

21. Faesen, A.C., Dirac, A.M., Shanmugham, A., Ovaa, H., Perrakis, A. and Sixma, T.K., 2011. Mechanism of USP7/HAUSP activation by its C-terminal ubiquitin-like domain and allosteric regulation by GMP-synthetase. *Molecular cell*, *44*(1), pp.147-159.

22. Felle, M., Joppien, S., Nemeth, A., Diermeier, S., Thalhammer, V., Dobner, T., Kremmer, E., Kappler, R. and Längst, G., 2011. The USP7/Dnmt1 complex stimulates the DNA methylation activity of Dnmt1 and regulates the stability of UHRF1. *Nucleic acids research*, *39*(19), pp.8355-8365.

23. Firestein, R., Bass, A.J., Kim, S.Y., Dunn, I.F., Silver, S.J., Guney, I., Freed, E., Ligon, A.H., Vena, N., Ogino, S. and Chheda, M.G., 2008. CDK8 is a colorectal cancer oncogene that regulates β-catenin activity. *Nature*, *455*(7212), pp.547-551.

24. Foot, N., Henshall, T. and Kumar, S., 2017. Ubiquitination and the regulation of membrane proteins. *Physiological reviews*, *97*(1), pp.253-281.

25. Fountain, M.D., Oleson, D.S., Rech, M.E., Segebrecht, L., Hunter, J.V., McCarthy, J.M., Lupo, P.J., Holtgrewe, M., Moran, R., Rosenfeld, J.A. and Isidor, B., 2019. Pathogenic variants in USP7 cause a neurodevelopmental disorder with speech delays, altered behavior, and neurologic anomalies. *Genetics in medicine*, *21*(8), pp.1797-1807.

26. Fryer, C.J., White, J.B. and Jones, K.A., 2004. Mastermind recruits CycC: CDK8 to phosphorylate the Notch ICD and coordinate activation with turnover. *Molecular cell*, *16*(4), pp.509-520.

27. Gagarina, V., Bojagora, A., Lacdao, I.K., Luthra, N., Pfoh, R., Mohseni, S., Chaharlangi, D., Tan, N. and Saridakis, V., 2020. Structural Basis of the Interaction Between Ubiquitin Specific Protease 7 and Enhancer of Zeste Homolog 2. *Journal of Molecular Biology*, *432*(4), pp.897-912.

28. Galbraith, M.D., Donner, A.J. and Espinosa, J.M., 2010. CDK8: a positive regulator of transcription. *Transcription*, *1*(1), pp.4-12.

29. He, J., Zhu, Q., Wani, G., Sharma, N., Han, C., Qian, J., Pentz, K., Wang, Q.E. and Wani, A.A., 2014. Ubiquitin-specific protease 7 regulates nucleotide excision repair through deubiquitinating XPC protein and preventing XPC protein from undergoing ultraviolet light-induced and VCP/p97 protein-regulated proteolysis. *Journal of Biological Chemistry*, *289*(39), pp.27278-27289.

30. Hengartner, C.J., Myer, V.E., Liao, S.M., Wilson, C.J., Koh, S.S. and Young, R.A., 1998. Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases. *Molecular cell*, *2*(1), pp.43-53.

31. Hirst, M., Kobor, M.S., Kuriakose, N., Greenblatt, J. and Sadowski, I., 1999. GAL4 is regulated by the RNA polymerase II holoenzyme–associated cyclin-dependent protein kinase SRB10/CDK8. *Molecular cell*, *3*(5), pp.673-678.

32. Holowaty, M.N., Zeghouf, M., Wu, H., Tellam, J., Athanasopoulos, V., Greenblatt, J. and Frappier, L., 2003. Protein profiling with Epstein-Barr nuclear antigen-1 reveals an interaction with the herpesvirus-associated ubiquitin-specific protease HAUSP/USP7. *Journal of Biological Chemistry*, *278*(32), pp.29987-29994.

33. Jagannathan, M., Nguyen, T., Gallo, D., Luthra, N., Brown, G.W., Saridakis, V. and Frappier, L., 2014. A role for USP7 in DNA replication. *Molecular and cellular biology*, *34*(1), pp.132-145.

34. Jeronimo, C. and Robert, F., 2017. The mediator complex: at the nexus of RNA polymerase II transcription. *Trends in Cell Biology*, *27*(10), pp.765-783.

35. Khoronenkova, S.V., Dianova, I.I., Ternette, N., Kessler, B.M., Parsons, J.L. and Dianov, G.L., 2012. ATM-dependent downregulation of USP7/HAUSP by PPM1G activates p53 response to DNA damage. *Molecular cell*, *45*(6), pp.801-813.

36. Kim, R.Q. and Sixma, T.K., 2017. Regulation of USP7: a high incidence of E3 complexes. *Journal of molecular biology*, *429*(22), pp.3395-3408.

37. Kimura Y, Tanaka K. 2010. Regulatory mechanisms involved in the control of ubiquitin homeostasis. J Biochem. 147(6): 793-798.

38. Kon, N., Zhong, J., Kobayashi, Y., Li, M., Szabolcs, M., Ludwig, T., Canoll, P.D. and Gu, W., 2011. Roles of HAUSP-mediated p53 regulation in central nervous system development. *Cell Death & Differentiation*, *18*(8), pp.1366-1375.

39. Lecona, E., Rodriguez-Acebes, S., Specks, J., Lopez-Contreras, A.J., Ruppen, I., Murga, M., Munoz, J., Mendez, J. and Fernandez-Capetillo, O., 2016. USP7 is a SUMO deubiquitinase essential for DNA replication. *Nature structural & molecular biology*, *23*(4), pp.270-277.

40. Li, J., Wang, R., Jin, J., Han, M., Chen, Z., Gao, Y., Hu, X., Zhu, H., Gao, H., Lu, K. and Shao, Y., 2020. USP7 negatively controls global DNA methylation by attenuating ubiquitinated histone-dependent DNMT1 recruitment. *Cell discovery*, *6*(1), pp.1-19.

41. Luyties, O. and Taatjes, D.J., 2022. The Mediator kinase module: an interface between cell signaling and transcription. *Trends in biochemical sciences*.

42. Ma, P., Yang, X., Kong, Q., Li, C., Yang, S., Li, Y. and Mao, B., 2014. The ubiquitin ligase RNF220 enhances canonical Wnt signaling through USP7-mediated deubiquitination of β-catenin. *Molecular and cellular biology*, *34*(23), pp.4355-4366.

43. Maertens, G.N., El Messaoudi-Aubert, S., Elderkin, S., Hiom, K. and Peters, G., 2010. Ubiquitin-specific proteases 7 and 11 modulate Polycomb regulation of the INK4a tumour suppressor. *The EMBO journal*, *29*(15), pp.2553-2565.

44. McCleland, M.L., Soukup, T.M., Liu, S.D., Esensten, J.H., de Sousa e Melo, F., Yaylaoglu, M., Warming, S., Roose-Girma, M. and Firestein, R., 2015. Cdk8 deletion in the ApcMin murine tumour model represses EZH2 activity and accelerates tumourigenesis. *The Journal of pathology*, *237*(4), pp.508-519.

45. Menzl, I., Witalisz-Siepracka, A. and Sexl, V., 2019. CDK8-novel therapeutic opportunities. *Pharmaceuticals*, *12*(2), p.92.

46. Nakamura, A., Nakata, D., Kakoi, Y., Kunitomo, M., Murai, S., Ebara, S., Hata, A. and Hara, T., 2018. CDK8/19 inhibition induces premature G1/S transition and ATR-dependent cell death in prostate cancer cells. *Oncotarget*, *9*(17), p.13474.

47. Nekhai, S., Petukhov, M. and Breuer, D., 2014. Regulation of CDK9 activity by phosphorylation and dephosphorylation. *BioMed research international*, *2014*.

48. Osman, S., Mohammad, E., Lidschreiber, M., Stuetzer, A., Bazsó, F.L., Maier, K.C., Urlaub, H. and Cramer, P., 2021. The Cdk8 kinase module regulates interaction of the mediator complex with RNA polymerase II. *Journal of Biological Chemistry*, *296*.

49. Park, M.J., Shen, H., Spaeth, J.M., Tolvanen, J.H., Failor, C., Knudtson, J.F., McLaughlin, J., Halder, S.K., Yang, Q., Bulun, S.E. and Al-Hendy, A., 2018. Oncogenic exon 2 mutations in Mediator subunit MED12 disrupt allosteric activation of cyclin C-CDK8/19. *Journal of Biological Chemistry*, *293*(13), pp.4870-4882.

50. Pfoh, R., Lacdao, I.K., Georges, A.A., Capar, A., Zheng, H., Frappier, L. and Saridakis, V., 2015. Crystal structure of USP7 ubiquitin-like domains with an ICP0 peptide reveals a novel mechanism used by viral and cellular proteins to target USP7. *PLoS pathogens*, *11*(6), p.e1004950.

51. Plassche, S.V.D. and de Brouwer, A.P., 2021. MED12-related (neuro) developmental disorders: a question of causality. *Genes*, *12*(5), p.663.

52. Pozhidaeva, A. and Bezsonova, I., 2019. USP7: Structure, substrate specificity, and inhibition. *DNA repair*, *76*, pp.30-39.

53. Qin, W., Leonhardt, H. and Spada, F., 2011. Usp7 and Uhrf1 control ubiquitination and stability of the maintenance DNA methyltransferase Dnmt1. *Journal of cellular biochemistry*, *112*(2), pp.439-444.

54. Reinstein, E. and Ciechanover, A., 2006. Narrative review: protein degradation and human diseases: the ubiquitin connection. *Annals of internal medicine*, *145*(9), pp.676-684.

55. Rickert, P.K., 1998. *Functional and Biochemical Characterization of the cyclin C/CDK8 kinase*. Stanford University.

56. Roos-Mattjus, P. and Sistonen, L., 2004. The ubiquitin-proteasome pathway. *Annals of medicine*, *36*(4), pp.285-295.

57. Rubinato, E., Rondeau, S., Giuliano, F., Kossorotoff, M., Parodi, M., Gherbi, S., Steffan, J., Jonard, L. and Marlin, S., 2020. MED12 missense mutation in a three-generation family. Clinical characterization of MED12-related disorders and literature review. *European journal of medical genetics*, *63*(3), p.103768.

58. Sarkari, F., Wheaton, K., La Delfa, A., Mohamed, M., Shaikh, F., Khatun, R., Arrowsmith, C.H., Frappier, L., Saridakis, V. and Sheng, Y., 2013. Ubiquitin-specific protease 7 is a regulator of ubiquitin-conjugating enzyme UbE2E1. *Journal of Biological Chemistry*, *288*(23), pp.16975-16985.

59. Seo, J.O., Han, S.I. and Lim, S.C., 2010. Role of CDK8 and β-catenin in colorectal adenocarcinoma. *Oncology reports*, *24*(1), pp.285-291.

60. Serrao, A., Jenkins, L.M., Chumanevich, A.A., Horst, B., Liang, J., Gatza, M.L., Lee, N.Y., Roninson, I.B., Broude, E.V. and Mythreye, K., 2018. Mediator kinase

CDK8/CDK19 drives YAP1-dependent BMP4-induced EMT in cancer. *Oncogene*, *37*(35), pp.4792-4808.

61. Sheng, Y., Saridakis, V., Sarkari, F., Duan, S., Wu, T., Arrowsmith, C.H. and Frappier, L., 2006. Molecular recognition of p53 and MDM2 by USP7/HAUSP. *Nature structural & molecular biology*, *13*(3), pp.285-291.

62. Snyder, N.A. and Silva, G.M., 2021. Deubiquitinating enzymes (DUBs): Regulation, homeostasis, and oxidative stress response. *Journal of Biological Chemistry*, *297*(3).

63. Soutourina, J., 2018. Transcription regulation by the Mediator complex. *Nature reviews Molecular cell biology*, *19*(4), pp.262-274.

64. Su, D., Wang, W., Hou, Y., Wang, L., Yi, X., Cao, C., Wang, Y., Gao, H., Wang, Y., Yang, C. and Liu, B., 2021. Bimodal regulation of the PRC2 complex by USP7 underlies tumorigenesis. *Nucleic Acids Research*, *49*(8), pp.4421-4440.

65. Swatek, K.N. and Komander, D., 2016. Ubiquitin modifications. *Cell research*, *26*(4), pp.399-422.

66. Tavana, O., Li, D., Dai, C., Lopez, G., Banerjee, D., Kon, N., Chen, C., Califano, A., Yamashiro, D.J., Sun, H. and Gu, W., 2016. HAUSP deubiquitinates and stabilizes N-Myc in neuroblastoma. *Nature medicine*, *22*(10), pp.1180-1186.

67. Van der Horst, A., de Vries-Smits, A.M., Brenkman, A.B., van Triest, M.H., van den Broek, N., Colland, F., Maurice, M.M. and Burgering, B.M., 2006. FOXO4 transcriptional activity is regulated by monoubiquitination and USP7/HAUSP. *Nature cell biology*, *8*(10), pp.1064-1073.

68. Wang, H., Shen, Q., Ye, L.H. and Ye, J., 2013. MED12 mutations in human diseases. *Protein & cell*, *4*(9), pp.643-646.

69. Westerling, T., Kuuluvainen, E. and Mäkelä, T.P., 2007. Cdk8 is essential for preimplantation mouse development. *Molecular and cellular biology*, *27*(17), pp.6177-6182.

70. Wu, D., Jia, H., Zhang, Z. and Li, S., 2020. Capsaicin suppresses breast cancer cell viability by regulating the CDK8/PI3K/Akt/Wnt/β-catenin signaling pathway. *Molecular Medicine Reports*, *22*(6), pp.4868-4876.

71. Wu, D., Zhang, Z., Chen, X., Yan, Y. and Liu, X., 2021. Angel or Devil?-CDK8 as the new drug target. *European Journal of Medicinal Chemistry*, *213*, p.113043.

72. Xu, D., Li, C.F., Zhang, X., Gong, Z., Chan, C.H., Lee, S.W., Jin, G., Rezaeian, A.H., Han, F., Wang, J. and Yang, W.L., 2015. Skp2–MacroH2A1–CDK8 axis orchestrates G2/M transition and tumorigenesis. *Nature communications*, *6*(1), pp.1-14.

73. Yamaguchi, L., Nishiyama, A., Misaki, T., Johmura, Y., Ueda, J., Arita, K., Nagao, K., Obuse, C. and Nakanishi, M., 2017. Usp7-dependent histone H3 deubiquitylation regulates maintenance of DNA methylation. *Scientific reports*, *7*(1), pp.1-12.

74. Yi, L., Cui, Y., Xu, Q. and Jiang, Y., 2016. Stabilization of LSD1 by deubiquitinating enzyme USP7 promotes glioblastoma cell tumorigenesis and metastasis through suppression of the p53 signaling pathway. *Oncology reports*, *36*(5), pp.2935-2945.

75. Zhao, J., Ramos, R. and Demma, M., 2013. CDK8 regulates E2F1 transcriptional activity through S375 phosphorylation. *Oncogene*, *32*(30), pp.3520-3530.

76. Zhu, Q., Sharma, N., He, J., Wani, G. and Wani, A.A., 2015. USP7 deubiquitinase promotes ubiquitin-dependent DNA damage signaling by stabilizing RNF168. *Cell cycle*, *14*(9), pp.1413-1425.