Master of Science (by Research) in Oncology 2023/2024

Dissertation Thesis Title:

Characterization of ubiquitin-activated DNA-protein crosslink repair *in vitro*

Xiaosheng Zhu
Wolfson College
University of Oxford

Sep 2024

Characterization of ubiquitin-activated DNA-protein crosslink repair *in vitro*

Abstract

Introduction: DNA-Protein Crosslinks (DPCs) are a form of DNA damage caused by covalent attachment between proteins and DNA, which can lead to replication fork stalling, disruption of cell division, and ultimately cell death. DPCs repair is critical for maintaining genome stability, however, the knowledge of the DPC repair pathways is still limited. SPRTN, a DNA-dependent metalloprotease, has been identified as the major protease in repairing DPCs. The function of SPRTN in DPC repair is regulated by ubiquitination. The regulation of SPRTN activity, in particular the N-terminal SPRTN catalytic region (SprT) core, remains largely unknown.

Aim: The fluorescently labeled histone H1 is used as the model substrate for SPRTN proteolysis. This project is designed to study the effect of ubiquitin and ubiquitin-like proteins on the activation of SPRTN SprT core towards the model H1-DPC substrate. We are also aiming to explore the effect of SPRTN proteolysis on H1 with different post-translational modifications.

Results: SPRTN activity towards the model H1-DPC substrate is enhanced by the addition of ubiquitin chains on substrate, but not the ubiquitin-like protein ISG15. To better understand SPRTN activation by modified substrates, H1 constructs fused with mono-ubiquitin, linear tetra-ubiquitin or 2xISG15 were designed and purified. Surprisingly, the complete cleavage of tetra-ubiquitin-modified H1, but not the 2xISG15-H1 fusion, can be observed within 5 min by the SprT core, indicating the much more efficient activating effect of Ub chains on SPRTN proteolysis.

Conclusion: Our study has demonstrated that ubiquitination of the DPC substrates is a key pathway in regulating SPRTN activity. These findings will not only contribute to the understanding of the mechanisms by which SPRTN proteolysis in DPCs repair, but also provide a powerful *in vitro* approach for drug screening to develop new cancer treatment strategies.

Table of Contents

ABSTRACT 2
LIST OF FIGURES 5
LIST OF TABLES 6
ABBREVIATIONS 7
1 INTRODUCTION 9
1.1 DNA-Protein Crosslinks 9
1.1.1 Type of DPCs 10
1.1.2 DPC Repair Pathways 11
1.2 SPRTN 13
1.2.1 Molecular Structure of SPRTN 14
1.2.2 SPRTN as a DNA-dependent Protease 15
1.2.3 the role of SPRTN in DPC Repair Pathways 17
1.3 Histone1 18
1.3.1 Molecular Structure of Histone H1 19
1.3.2 Histone H1 Acts As a Model Substrate of SPRTN 20
1.4 Ubiquitin 21
1.4.1 Ubiquitin and Ubiquitin chains 21
1.4.2 Ubiquitination-dependent Degradation of DPCs 23
1.4.3 Ubiquitin and SPRTN 24
1.5 Objective of Thesis 26
2 METHODOLOGY AND MATERIALS 27
2.1 Plasmid Construction 27
2.1.1 Gene Cloning 27
2.2 SDS-PAGE 29
2.3 Protein Expression 30
2.4 Protein Purification 31
2.4.1 Sonication 31
2.4.2 Ni-NTA Affinity Chromatography and Tag Removal 32
2.4.3 Size-Exclusion Chromatography (SEC) 33
2.4.4 Protein Quantification 34
2.4.5 Proteins Purified in This Thesis 35
2.5 Cy5 Labeling 36
2.6 Making Model dsDNA-H1 DPC 37
2.7 SPRTN Assay 37
3 RESULTS 39
3.1 The Purification of Ubiquitin-fused H1 Substrates 39
3.1.1 Purification of ISG15 40
3.1.2 Purification of 2×ISG15-H1 42
3.1.3 Purification of Ub-H1 44
3.1.4 Purification of dsDNA-H1 DPC 46
3.2 Linear Ub_4, but not ISG15, Activates SprT Cleavage Towards H1-DPC 48

3.3 SprT Core Cleaves Ub_4-H1 within 5 minutes, but not H1, Ub-H1 or 2xISG15-H1 49
4 DISCUSSIONS 51
4.1 Protein Purification 52
4.2 Ubiquitin Chains Activate SPRTN Proteolysis of H1-DNA Conjugate 54
5 CONCLUSION AND OUTLOOK 56
BIBLIOGRAPHY 57
APPENDIX Ⅰ: INFORMATION OF PLASMIDS AND PROTEINS 61

List of Figures

Figure 1.1 Different types of DNA lesions.....9
Figure 1.2 Structure of SPRTN.....14
Figure 1.3 Protease activity of SPRTN and different types of DNA 15
Figure 1.4 Structural details of chromatin.....18
Figure 1.5 Secondary structure of histone H1.....19
Figure 1.6 Classification of ubiquitin chains.....21
Figure 1.7 Linear Ub_4, but not monoUb, Ub_2 or ISG15, activates SprT cleavage towards H1.....24
Figure 3.1 Schematics of His-ISG15, His-2×ISG15-H1 and His-Ub-H1.....39
Figure 3.2 Purification of His-tagged ISG15.....30
Figure 3.3 Purification of tag-free ISG15.....41
Figure 3.4 Purification of His-tagged 2×ISG15-H1.....42
Figure 3.5 Purification of His-tagged Ub-H1.....44
Figure 3.6 Purification of dsDNA-H1 DPC.....46
Figure 3.7 Linear Ub_4, but not ISG15, activates SprT core cleavage towards H1-DPC.....48
Figure 3.8 SprT core cleaves Ub_4-H1 in a faster manner than H1, Ub-H1 and 2xISG15-H1.....49
Figure 4.1 Secondary structure of M1-diUb and ISG15.....55

List of Tables

Table 1. List of molecular weight and purification method of the proteins.....35
Table 2. List of sequence of ssDNA.....37
Table 3. List of DNA, proteins and time points used in SPRTN assay.....38

Abbreviations

DDR	DNA Damage Response
SSBs	Single-strand Breaks
DSBs	Double-strand Breaks
DPCs	DNA-protein Crosslinks
IR	Ionizing Radiation
UV	Ultraviolet
FA	Formaldehyde
GCNA	Germ Cell Nuclear Antigen
RJALS	Ruijs-Aalfs Syndrome
HCC	Hepatocellular Carcinoma
RJALS	Ruijs-Aalfs Syndrome
ssDNA	Single-stranded DNA
dsDNA	Double-stranded DNA
PARP1	Poly(ADP-ribose) Polymerase 1
DNMT1	DNA Methyltransferase 1
TOP1	DNA Topoisomerase 1
TOP2	DNA Topoisomerase 2
TOP2cc	TOP2 cleavage complex
5-aza-dC	5-aza-2'-deoxycytidine
Wss1	Weak Suppressor of SMT3 Protein 1
ICLs	Interstrand Crosslinks
DDI1	DNA Damage Inducible 1
VCP	Valosin-containing Protein
BR	Basic DNA-binding Region
SHP	Ubiquitin Segregase p97-interacting Motif
PIP	Proliferating Cell Nuclear Antigen Interacting Motif
MPD	Metalloprotease Domain
UBZ	Ubiquitin-binding Zinc Finger
ZBD	Zn^{2+}-binding Sub-domain
PCNA	Proliferating Cell Nuclear Antigen
MIU	Motif Interacting with Ubiquitin
HR	Homologous Recombination
NTD	N-terminal Domain
GD	Globular Domain
CTD	C-terminal Domain
PTMs	Post-translational Modifications
SAR	Scaffold-associated Regions
CDK	Cyclin-dependent Kinases
ER	Endoplasmic Reticulum
DUBs	Deubiquitinases
E1	Ubiquitin-activating Enzyme

Lys	Lysine
Cys	Cysteine
E2	Ubiquitin-conjugating Enzymes
E3	Ubiquitin-ligase Enzymes
PARylation	Poly-ADP-ribosylation
PIAS4	Protein Inhibitor of Activated STAT 4
TLS	Translesion DNA Synthesis
USP7	Ubiquitin-specific Protease 7
VCPIP1	Valosin Containing Protein Interacting Protein 1
USP11	Ubiquitin-specific Protease 11
Ub_4	Tetra-ubiquitin Chains
Ub_2	Di-ubiquitin Chains
E. coli	*Escherichia Coli*
IPTG	Isopropyl β-D-1-thiogalactopyranoside
SDS-PAGE	Sodium Dodecyl Sulphate-polyacrylamide Gel Electrophoresis
His*	Affinity Chromatography by Ni-NTA Column
SEC	Size-exclusion Chromatography
IEX	Ion Exchange Chromatography
NCP	Nucleosome Core Particles
SMCC	Succinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate

1. Introduction

1.1 DNA-protein Crosslinks

DNA is the fundamental molecule that carries all the genetic information necessary for cell survival. However, the stability of DNA is constantly threatened by various internal and external factors. DNA lesions include DNA mutations, mismatches, single-strand breaks (SSBs), double-strand breaks (DSBs), interstrand crosslinks (ICLs) and DNA-protein crosslinks (DPCs) (1). To maintain genome stability, cell activates DNA damage response (DDR) to counteract these threats. If DNA damage is not repaired in time, it may accumulate mutations and initiate diseases such as cancers, neurodegenerative disorders, immune deficiencies, cardiovascular diseases and metabolic syndromes (2).

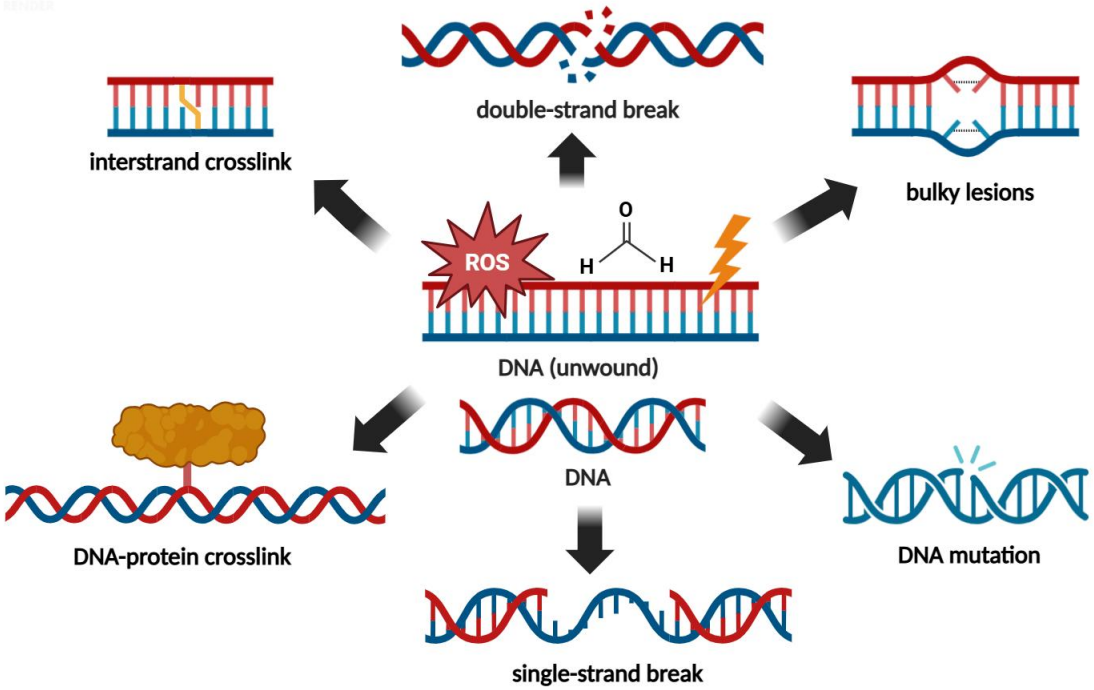


Figure 1.1 Different types of DNA lesions. *Under the stimulation of ROS, aldehydes, ultraviolet radiation, and other factors, DNA lesions are formed, including interstrand crosslinks, single-strand breaks, double-strand breaks, bulky lesions, DNA mutations, and DNA-protein crosslinks.*

DPCs are a type of DNA damage caused by the irreversible covalent attachment between proteins and DNA. Compared to other forms of DNA lesions, such as SSBs, DSBs, and ICLs, DPCs pose unique challenges to the repair mechanisms and can lead to more severe cellular outcomes, including replication fork stalling, which inhibits DNA replication, transcription and cell division, ultimately resulting in cell death (4). Therefore, DPCs are considered one of the most severe forms of DNA damage (3).

1.1.1 Type of DPCs

Based on the mechanisms of formation, DPCs can be categorized into enzymatic DPCs and non-enzymatic DPCs (5). Enzymatic DPCs are typically produced by enzymes involved in DNA metabolism, such as poly(ADP-ribose) polymerase 1 (PARP1), DNA methyltransferase 1 (DNMT1), DNA topoisomerase 1 (TOP1) and DNA topoisomerase 2 (TOP2) (6). Under endogenous conditions, enzymatic DPCs are often generated during normal cellular processes, such as DNA repair, replication, and transcription. For example, during DNA repair, PARP1 can attach poly(ADP-ribose) chains to DNA and proteins, forming a transient complex (6). However, under certain stress conditions or damaged by exposure to external stimuli, such as etoposide or camptothecin (CPT), making the enzymes fail to dissociate from DNA. Stable enzymatic DPCs are thus formed(7). For example, 5-aza-2'-deoxycytidine (5-aza-dC) can be incorporated into DNA and trap DNMT1 to form DPCs during replication (8).

Non-enzymatic DPCs, on the other hand, are caused by non-specific chemical interactions between DNA and proteins, potentially triggered by intracellular metabolic products or exogenous factors (9). Intracellular metabolic factors include some byproducts of cellular metabolism, such as aldehyde species. For example, formaldehyde can react with primary amines from the DNA and subsequently covalently bind to lysine or arginine residues from the proteins, to form DPCs. Other aldehydes, such as acetaldehyde, can also contribute to the formation of DPCs (9). Exogenous factors include ionizing radiation (IR), ultraviolet (UV) light and environmental carcinogens (such as transition metal ions), which can generate reactive oxygen species (ROS) leading to DNA-protein crosslinking (10).

1.1.2 DPC Repair Pathways

The deficiency in DNA repair can sensitise cells to DNA damaging agents, which could result in synthetic lethality. Most anticancer treatments work by inducing DNA damage in cancer cells. Combining these treatments with agents that inhibit DNA damage repair pathways might significantly enchance the effectiveness of anti-cancer therapies. Therefore, studying DNA damage repair pathways, particularly the repair of toxic DPCs, is highly beneficial for cancer treatment.

So far, several DPC repair pathways have been identified in cells. Currently known DPCs repair pathways include TDP1 and TDP2-mediated DPC repair pathway, the protease-dependent repair pathway, nuclease-dependent repair pathway, homologous recombination (HR) repair pathway, and Fanconi anaemia pathway (11).

For TDP1 and TDP2-mediated DPC repair pathway, tyrosyl-DNA phosphodiesterase (TDP) can cleave the DPCs by directly hydrolyzing the covalent bond between the topoisomerase's (TOP1 or TOP2) active site and DNA (12). TDP deficiency will lead to accumulation of TOP1 or TOP2 on DNA and prevent DNA replication (13). It is worth to mention that the SUMO (small ubiquitin-like modifier) ligase ZATT (ZNF451) binding to TOP2 cleavage complex (TOP2cc) activates the TDP2 hydrolase activity on stalled TOP2cc (14). Additionally, ZATT further promotes the interactions between TDP2 and SUMOylated TOP2, regulating the efficient DPC repair pathway (14).

Protease-dependent repair pathway is seen as the crucial mechanism to maintain genome stability. For example, weak suppressor of SMT3 protein 1 (Wss1) was first found from a synthetic lethality screen in yeast which disabled the survival of cells deficient for Tdp1 (15). Wss1 also works in repairing DPCs induced by formaldehyde, indicating that Wss1 can also resolve non-enzymatic DPCs (16). Spartan (SPRTN, also called DVC1) is identified as the functional homologue of Wss1 in human (17).

For nuclease-dependent repair pathway, the MRE11-RAD50-NBS1 (MRN) nuclease complex targets the DNA component of the DPCs, excising the damaged fragment (7). HR and Fanconi anaemia pathway provide resistance to some DPC inducing agents. Currently, it is known that HR repair pathway increases cellular tolerance to DPCs by repairing the damage via the MRN nuclease complex. The Fanconi anemia pathway is involved in the repair of interstrand crosslinks (ICLs) (18). Some Fanconi anemia proteins play a role in processing DPCs repair. For example, DNMT DPCs caused by 5-azadC cannot be repaired in FANCG-depleted CHO cells, leading to an increase in chromatid breaks (19). FANCG-depleted CHO cell line and FANCD2-depleted cells exhibited hypersensitivity toward acetaldehyde and formaldehyde respectively (20, 21). Additionally, recent study has found that cells are sensitive to induction that trap PARP on DNA if crucial Fanconi anaemia genes are lost (22). It shows that HR and Fanconi anaemia pathway have essential effect on DPC repair.

Although several DPC damage repair pathways have been identified, their detailed mechanisms still require further investigation. This study will focus on the protease-dependent repair pathway, in particular, SPRTN, in the repair of DPCs.

1.2 SPRTN

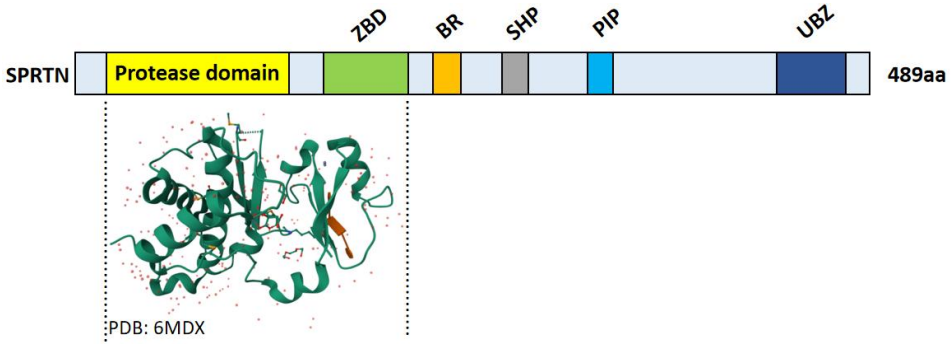
Proteases from the Wss1/SPRTN family cleave the protein of DPCs and are recognized as essential DPCs repair enzymes (17). Apart from SPRTN, there are other proteases related to the proteolysis of DPCs, such as Germ Cell Nuclear Antigen (GCNA) (33), FAM111A, FAM111B (23), DNA damage inducible 1 (DDI1) and DDI2 (24). However, SPRTN is the only DPC repair protease that is essential in cells. Biallelic mutations in human protease SPRTN result in Ruijs-Aalfs syndrome (RJALS), characterized by progeria and early onset hepatocellular carcinoma (HCC) (25). Similar pathological features were also observed in the hypomorphic SPRTN mouse model. Mice with homozygous SPRTN deletions died during the early stages of embryonic development (26). It proves that SPRTN is important for DPCs repair, cell survival and cell development.

1.2.1 Molecular Structure of SPRTN

Human SPRTN protease, at a molecular weight of 55 kDa, contains an N-terminal domain, a basic DNA-binding region (BR) and a C-terminal tail. The C-terminal of SPRTN contains the ubiquitin segregase p97 (or valosin-containing protein, VCP)-interacting motif (SHP), the proliferating cell nuclear antigen interacting box (PIP) and the ubiquitin binding domain (UBZ). SHP is required to interact and recruit ATPase p97 to the DNA damage sites. PIP box interacts with proliferating cell nuclear antigen (PCNA). UBZ domain can interact with ubiquitin (27). The C-terminal of SPRTN doesn't have DNA-binding ability, though *in vivo* activity of SPRTN-ΔC is significantly reduced, indicating the key regulatory roles from the C-terminal (31).

The N-terminal of SPRTN includes motif interacting with ubiquitin (MIU) domain, SprT domain and BR. SprT domain is responsible for the proteolytic cleavage of the substrate. The SprT domain consists of metalloprotease sub-domain (MPD) and a Zn^{2+}-binding sub-domain (ZBD). MPD domain contains three α helices and a three-stranded anti-parallel β sheet (28). ZBD and BR are DNA-binding domains (28).

(a)



(b)

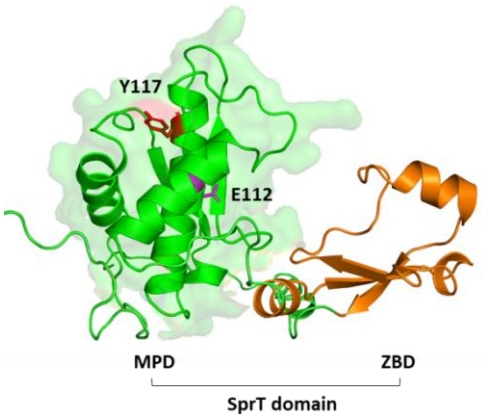


Figure 1.2 Structure of SPRTN. *(a) Domain structure of SPRTN. Color code: zinc-binding domain (ZBD) (green), the basic DNA-binding region (BR) (orange), the SHP box (grey), the PCNA-interacting motif (PIP) (light blue), and the ubiquitin-binding zinc finger (UBZ) (dark blue). (b) Structure of N-terminal domain of SPRTN (SPRTN 1-245, AlphaFold: Q9H404). Mutations at Y117 on the MPD domain and ΔC lead to Ruijs-Aalfs syndrome, whereas the E112 mutation does not.*

1.2.2 SPRTN as a DNA-dependent Protease

SPRTN is inactive without DNA, but its protease activity significantly increases in the presence of DNA with single- and double-stranded features such as gaps, dsDNA ends, ss/dsDNA junctions or DNA bubbles (see Fig. 1.3)(29). SPRTN's access to chromatin is regulated by DNA switch. DNA-binding induces conformational changes of SPRTN to better bind and degrade substrates (30). DNA serves as molecular scaffold to bring SPRTN and substrates in close proximity, which increases residence time between SPRTN and substrates (31). This allows the protein hydrolysis reaction to be fully carried out.

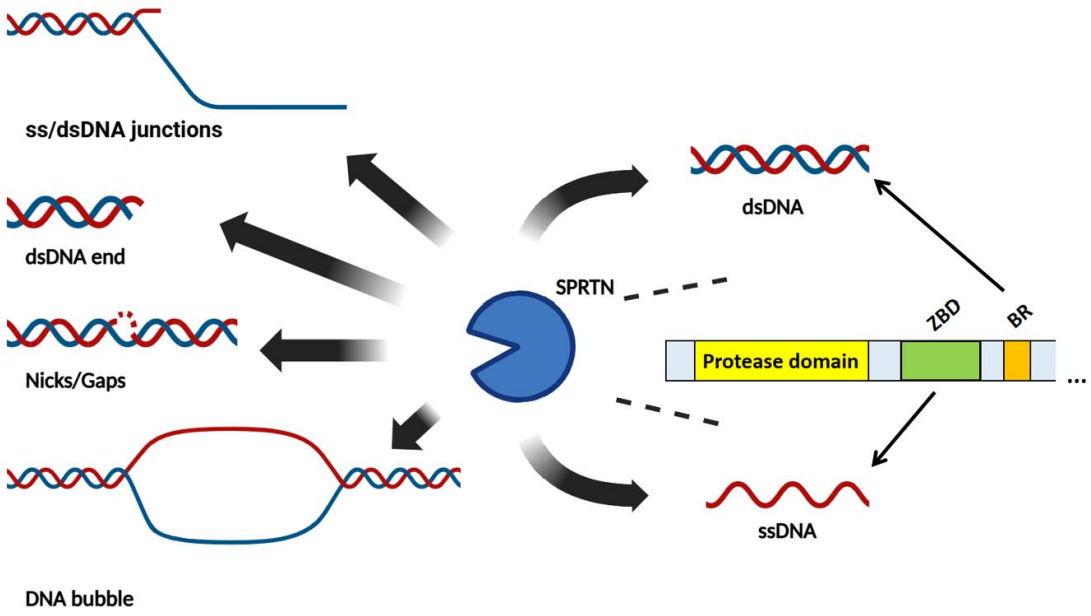


Figure 1.3 Protease activity of SPRTN and different types of DNA. *SPRTN can be activated by ss/dsDNA junctions, dsDNA end, nicks and DNA bubble. ZBD and BR binds to ssDNA and dsDNA, respectively (17).*

In vitro assay has demonstrated that N-terminal of SPRTN (SPRTN-ΔC) still retain the DNA-binding ability, whereas the isolated C-terminal of SPRTN can not (30), suggesting that the DNA binding site is located in the N-terminal region. Previous studies have localized this binding region to the N-terminal BR domain, which is identified as the primary DNA-binding domain (30). The ZBD domain, in particular, contains a binding site for ssDNA (Fig. 1.3) and shields active site located in the MPD domain (28). It has been reported that the ZBD domain specifically recognizes ssDNA at unwinding dsDNA ends, ss/dsDNA junctions, bubbles or at the ends of a DNA hairpin, while BR domain enhances binding through non-specific interactions with the double-stranded parts of these structures (29).

Different types of DNA binding to SPRTN can activate the DNA switch in different ways. The long circular double-stranded DNA (dsDNA) activates the auto-cleavage

activity of SPRTN, while short dsDNAs and long circular single-stranded DNA (ssDNA) can fully activate SPRTN to cleave substrates. Reinking et al. proposed that ZBD and BR domain in SPRTN are involved in recognizing different types of DNA. BR binds to negatively charged phosphate backbone of dsDNA , and ZBD constitutes a ssDNA binding domain (22). Mutations in the ZBD inhibit the protease activity of SPRTN. Moreover, the auto-cleavage of SPRTN achieves a balance between activating and deactivating its protease activity. SPRTN auto-cleavage removes its C-terminus, allowing SPRTN to detach from the DNA and effectively terminate its protease activity (30, 31). The auto-cleavage is essential for restraining the substrate-cleavage activity of SPRTN and preventing undamaged chromatin from unnecessary cleavage by SPRTN.

1.2.3 the role of SPRTN in DPC Repair Pathways

SPRTN is the essential DPC enzyme for cellular viability. Cells carrying SPRTN mutations or deficient in expression exhibit increased sensitivity to DPC-inducing agents such as formaldehyde, CPT, and Etoposide (32). Mass spectrometry analysis of SPRTN degradation products did not reveal any specific cleavage sites, indicating SPRTN does not preferentially recognize any specific amino acid sequence when identifying protein substrates (33). This also explains that SPRTN has the broadest substrate specificity among all the DPC repair pathways. Therefore SPRTN need to be well regulated to avoid potential threat to normal proteins in cells. As mentioned in 1.2.2, the protease activity of SPRTN is strictly dependent on activation by DNA.

In addition to DNA switch, The activity of SPRTN in cleaving DPCs is also regulated by ubiquitin switch (34). The UBZ domain of SPRTN is involved in regulating the chromatin access of SPRTN. Unmodified SPRTN can localize to chromatin, whereas mono-ubiquitinated SPRTN cannot. The induction of DPCs upon exposure to formaldehyde leads to the deubiquitination of SPRTN and hence regulates its access to chromatin (30). This ubiquitin switch is seen as upstream regulatory switch for SPRTN activity. Once unmodified SPRTN is localized to chromatin, the DNA switch is initiated by binding to DNA. The dual regulation by the ubiquitin switch and the DNA switch enables SPRTN protease to cleave the targeted substrates and prevents the unwanted proteolysis. However, the specific ubiquitination signal, E3 ligases, and deubiquitinases involved in regulating the ubiquitin switch remain largely unknown.

Given the important role of SPRTN in DPCs repair pathways, inhibition of SPRTN activity could induce cancer cells sensitive to many types of DPCs inducing agents. Liver is the primary detoxification organ and is involved in the extensive metabolism of active aldehydes. Liver cells are exposed to a significantly higher amount of DPCs compared to other tissues. Therefore, inhibiting SPRTN activity may have a more pronounced therapeutic effect on liver cancer cells.

1.3 Histone H1

Chromatin is the higher-order structure responsible for protecting and condensing the eukaryotic genome in the nucleus. It is composed of repeating units called chromatosomes, which are made up of nucleosome cores, DNA and linker histone H1 (H1) (or H5 in birds) (35). The structure of the nucleosome core is highly conserved from yeast to metazoans (36). A nucleosome core is composed of particles which consist of ~147 bp segment of DNA winding in a left-handed manner around the histone octamer (37). The histone octamer is an assembly of eight core histone proteins, including two copies of each H2A, H2B, H3, and H4. The core histone proteins are small (11-15kDa) and conserved among eukaryotic species (38). Linker DNA has a variable length of ~50bp and connect nucleosome core particles (NCP) to become nucleosome cores (39). Histone H1 binds to linker DNA regions on the surface of the nucleosome cores (40).

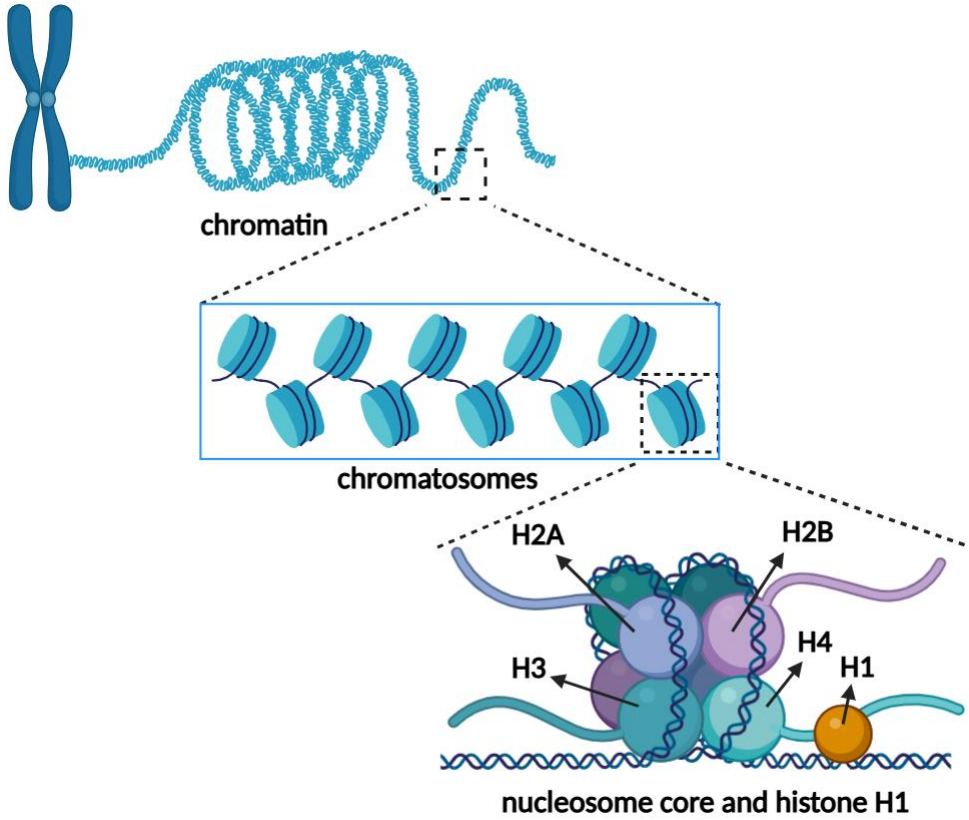


Figure 1.4 Structural details of chromatin. Color code: *H2A in blue, H2B in purple, H3 in dark green, H4 in light green, H1 in orange.*

Histone H1 is the most diverse class of histones, comprising eleven different subtypes in mammals (41). Histone H1 is lysine-rich and associated with the formation of higher-order chromatin structure. Studies have shown that H1 binding suppresses nucleosome unwrapping by associating with the two linker DNA segments, as well as by binding to the DNA that crosses the dyad, therefore limiting chromatin accessibility (42). There is a general correlation between the amount of H1 in chromatin and chromatin transcriptional activity (43).

1.3.1 Molecular Structure of Histone H1

In higher eukaryotes, H1 contains three structural domains: a short N-terminal domain (NTD), a central globular domain (GD), and a long C-terminal domain (CTD) (44). Mutations in these three domains affect the turn-over kinetics of H1 on DNA, suggesting that all the 3 domains contribute to the binding affinity of H1 to chromatin (45).

The globular domain has ~80 residues. It contains a winged helix-turn-helix DNA binding motif and two antiparallel β-strands (β1 and β2). The globular domain is highly conserved across species and responsible for binding to nucleosomal DNA. It protects an extra 20 bp of nucleosomal DNA from nuclease digestion.

Both terminal domains of H1 determine subtype specificity and carry a positive net charge under physiological conditions due to their high lysine content (47, 48). These domains are intrinsically disordered, with variability in sequence, length, and post-translational modifications (PTMs). Intrinsic disorder can provide several functional advantages, including the ability to interact with various partners, the potential to partially or completely fold when binding, and faster interaction rate (49, 50). These features collectively facilitate the regulation of H1's secondary structure and chromatin dynamics.

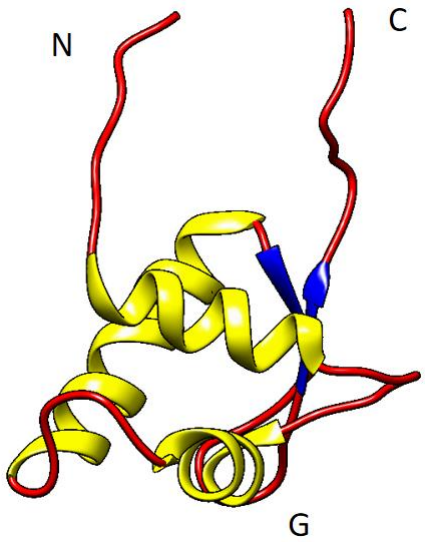


Figure 1.5 Protein structure of histone H1. Histone H1 contains an N-terminal domain, a globular domain and a C-terminal domain *(PDB: 6LAB)* (44). *The α-helices are colored yellow, while the β-strands are colored blue.*

1.3.2 Histone H1 Acts as a Model Substrate of SPRTN

Due to the high affinity to DNA, H1 and H3 easily form DNA-histone crosslinks by potent stimulus *in vivo* (51). In Ramadan's group, H1 is found largely enriched in DPCs induced by formaldehyde (FA) *in vivo* (unpublished data).

It has been characterized that SPRTN plays a role in DNA damage repair (DDR), such as degrading histone H1, H3 or Topo-2 in DNA replication (52). In the presence of DNA, SPRTN cleaves H1 in disordered regions near arginine, lysine, and serine residues (31), and can only cleave loose tails of H1 (27).

1.4 Ubiquitin

1.4.1 Ubiquitin and Ubiquitin Chains

Ubiquitin is a small protein at the molecular weight of 8.5kDa. Its conjugation to target proteins is known as ubiquitination (52). Ubiquitination is an important type of post-translational modifications (PTMs) that regulates various cellular processes, including cell proliferation, DDR, protein degradation, signaling pathways and so on (53, 54). Dysregulation of the ubiquitination pathway can lead to severe diseases, including cancer and neuro-degenerative disorders (55). Therefore, the study on the ubiquitin system, is of great significance to advance our understanding of the molecular mechanisms underlying these diseases and therefore, to develop potential therapeutic strategies.

Ubiquitin is highly conserved in eukaryotic organisms, which differs by only 3 amino acids from yeast to human (56). The diversity of ubiquitination depends on the lengths and linkage types of ubiquitin chains (57). Monoubiquitination refers to the covalent attachment of a single ubiquitin molecule to a substrate. These monoubiquitination events are highly prevalent in cells and perform various functions (58). Multi-monoubiquitination can happen at multiply sites on a single molecule of substrate (59). The attachment of ubiquitin chains to a substrate can result in protein polyubiquitination (60).

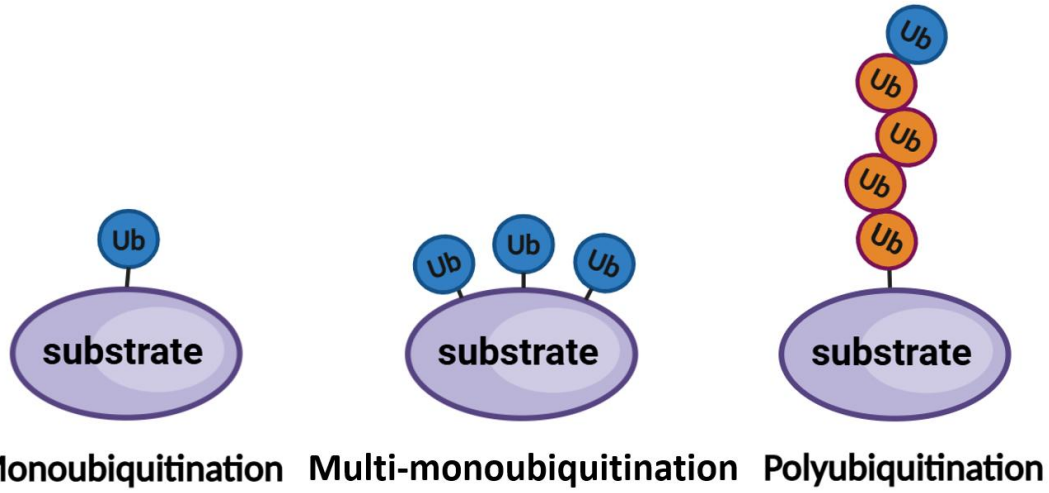


Figure 1.6 Classification of ubiquitination. *Ubiquitin can be conjugated as a monomer on one (monoubiquitination) or more sites on a substrate (multi-monoubiquitination) or as a polymer (polyubiquitination) by the sequential addition of ubiquitin molecules via lysine residues on ubiquitin itself (58-60).*

There are eight ubiquitin linkages, including Met1 (M1), Lys6 (K6), Lys11 (K11), Lys27 (K27), Lys29 (K29), Lys33 (K33), Lys48 (K48), and Lys63 (K63), that can be conjugated to the C-terminus of another ubiquitin molecule, forming structurally complex ubiquitin chains (61). Different types of ubiquitin chains have different functions in

cells. For example, K48-linked polyubiquitination primarily directs the substrate to the 26S proteasome for degradation (62). K63-linked polyubiquitination is primarily associated with DNA damage repair, protein trafficking, endocytosis, and signal transduction (63). K33-linked polyubiquitination promote kinase modifications. K6-linked polyubiquitination is involved in the regulation of DNA repair. K11-linked polyubiquitination is related to endoplasmic reticulum (ER)-associated degradation (64).

The ubiquitination process requires the catalytic activity of ubiquitin enzymes. First, ubiquitin-activating enzyme (E1) uses energy from ATP to form a high-energy thioester bond between its catalytic cysteine residue and the lysine residue at the C-terminus of ubiquitin, thereby activating the ubiquitin molecule. Next, the activated ubiquitin is transferred via a thioester bond to the catalytic cysteine residue of the ubiquitin-conjugating enzymes (E2). Finally, the activated ubiquitin is either directly attached to the protein substrate from E2 or transferred to the target protein under the action of ubiquitin-ligase enzymes (E3), forming an isopeptide bond between the C-terminal carboxyl group of ubiquitin and the ε-amino group of Lys residue from the target protein (65). E3 plays the most crucial role in the specific recognition of target proteins and the regulation of the activity of the ubiquitination system. E3 confers specificity by selecting target proteins and determining the type of ubiquitin modification in this cascade of enzymatic reactions (66). The process is reversible. Deubiquitinases (DUBs) can remove ubiquitin from the target proteins (67).

1.4.2 Ubiquitination-dependent Degradation of DPCs

Post-translational modifications (PTMs), such as SUMOylation, ubiquitylation, and poly-ADP-ribosylation (PARylation), play an important role in DPCs proteolysis (68, 69).

In the DPC degradation pathway disassociated with DNA replication, DPCs are SUMOylated by E3 SUMO ligases, such as protein inhibitor of activated STAT 4 (PIAS4). This process subsequently directs DPCs to be polyubiquitinated by SUMO-targeted E3 ubiquitin ligases, such as RNF4, leading to proteasomal degradation (70). This degradation process allows downstream DNA repair mechanisms to proceed after DNA damage occurs (71).

DPCs can be degraded by the proteasomes and/or the protease ACRC/GCNA, DDI1 and DDI2, FAM111A, FAM111B, and SPRTN. In the DPC-degradation pathway associated with DNA replication, DPCs are ubiquitinated by the E3 ubiquitin ligase TRAIP (72). This process stimulates the replicative CMG helicase to bypass the DPCs (75). Then the protein adduct becomes a barrier to replicative polymerases (73). After the CMG bypass, a short stretch of ssDNA is generated downstream of the DPCs, likely serving as a target for the E3 ubiquitin ligase RFWD3 (74). RFWD3 further ubiquitylates DPCs on ssDNA gaps, promoting DPCs degradation by the proteasome (74). In SPRTN-mediated DPCs-degradation pathway, UBZ domain of SPRTN recruits SPRTN to proper damage site by direct interaction with ubiquitin conjugates (67). Then, the protease activity is activated by the extension of the nascent strand to the DPCs (29, 72) and switched off by self-cleavage when DPC degradation is complete (30). After the DPC is degraded to a short peptide-DNA adduct by the proteasome and/or the SPRTN, translesion DNA synthesis (TLS) polymerases are allowed to bypass (75).

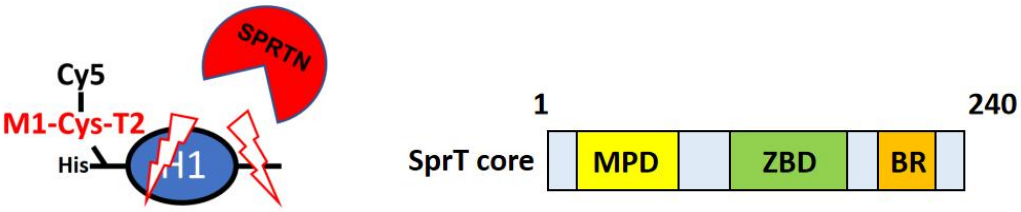
1.4.3 Ubiquitin and SPRTN

SPRTN protease activity is activated by DNAs to resolve DPCs. SPRTN cleavage is independent of amino acid sequence. This is beneficial to cleave different DNA-protein adducts, but it also needs strict regulation to avoid uncontrolled protein cleavage. Ubiquitination is a key mechanism in regulating and controlling SPRTN activity, ensuring its specificity in DPC repair (76). In cells, 50% of SPRTN is monoubiquitylated, but this is shielded by the UBZ domain (77, 78). Monoubiquitination on SPRTN negatively regulates the activity of SPRTN by inducing autocleavage and priming polyubiquitination for proteasomal degradation (79). This "off switch" on SPRTN can be reversed by deubiquitinases (DUBs), such as ubiquitin-specific protease 7 (USP7), valosin containing protein interacting protein 1 (VCPIP1), and ubiquitin-specific protease 11 (USP11) (80). Under DPC induction, SPRTN is deubiquitinated, which decreases its auto-cleavage and increases the number of enzymes involved in DPC damage repair (81, 82).

While it has been established that SPRTN ubiquitination can regulate its cleavage, free ubiquitin chains also play a critical role for SPRTN activity, thus highlighting the complex interplay between ubiquitin and SPRTN. However, the mechanism of SPRTN cleavage regulated by poly-ubiquitin or ubiquitin-like protein are still largely unknown.

Prof. Ramadan's group has developed a Cy5-labelled H1 as the model substrate to study SPRTN proteolysis (Fig 1.7 a). Dr. Wei Song has found that the substrate-cleavage activity from the SPRTN SprT core towards H1 is enhanced by the addition of Met1-linked (linear) tetra-ubiquitin chains (M1-Ub_4) in the presence of DNA, but not monoUb, linear di-ubiquitin chains (M1-Ub_2) or the ubiquitin-like protein ISG15 (Fig 1.7 b) (unpublished data).

(a) (b)



(c)

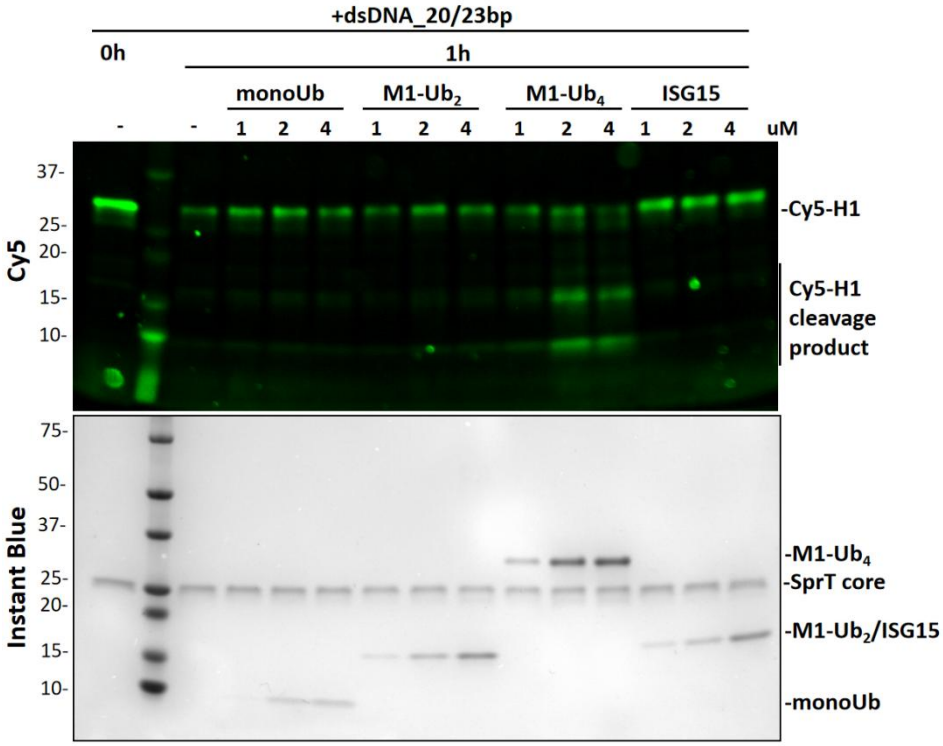


Figure 1.7 M1-Linear Ub$_4$, but not monoUb, M1-Ub$_2$ or ISG15, activates SprT core cleavage towards H1 (Courtesy of Dr. Wei Song). *(a) Cy5-labelled H1 as the substrate model. (b) Domain structure of SPRTN N-terminal. (c) SprT domain is activated by M1-Ub$_4$ in the presence of DNA, but not monoUb, M1-Ub$_2$ or ISG15.*

1.5 Objective of Thesis

Proteases which participate in DNA damage repair are seen as potential drug targets for cancer treatment. Prof. Ramadan's lab found that metalloprotease SPRTN and its binding partner ATPase p97 have essential effect on the DPCs repair of trapped PARP1. This type of DPCs is highly relevant to the treatment effect of PARP inhibitors in the colorectal and breast cancer. Studying the mechanism of DPC repair is beneficial to understand how cells maintain genome stability during the induction of DPCs. It will also help to develop new methods to target DPC repair pathways to enhance the sensitivity of tumours to chemotherapy.

SPRTN lacks substrate specificity, which necessitates tight regulation to prevent unintended cleavage of DNA-bound proteins as well as SPRTN itself. This highlights the importance of studying the mechanism of SPRTN regulation. As mentioned in 1.4.3, free M1-Ub_4 activates SPRTN activity. It may indicate that ubiquitin is involved in regulating SPRTN cleavage and its specificity towards substrates. To test this hypothesis, we further explore SPRTN proteolysis in the context of different substrates. We designed and purified H1-DPC, along with modified DPCs in which ubiquitin and ubiquitin-like protein are fused to histone1. This work is of great benefit to understand the mechanism of ubiquitin on SPRTN cleavage, and also promising to develop new cancer treatment strategies.

2. Materials & Methods

2.1 Plasmids Construction

2.1.1 Gene Cloning

Q5 High-Fidelity DNA Polymerase (NEB) and PCRBIO Taq Mix Red are used to amplify target fragments. The PCR reaction for both polymerases is as follows:

PCR reaction (Use PCRBIO Taq Mix Red):

2×PCRBIO Taq Mix Red	25 μL
Template	0.5 μL
Forward primer (10 uM)	2 μL
Reverse primer (10 uM)	2 μL
ddH_2O	20.5 μL

PCR reaction (Use Q5 High-Fidelity DNA Polymerase):

2×Q5 Master Mix	25 μL
Forward and Reverse Primers	5 μL
Template DNA	0.5 μL
ddH_2O	20 μL

PCR products were separated by agarose gel electrophoresis. DNA bands with the right size from TAE agarose gel were excised, then purified by Gel Extraction Kit. The DNA concentration is quantified by NanoDrop™ One (Thermo Fisher Scientific).

Restriction Enzyme Digestion:
NcoI (NEB) and BamHI (NEB) were used to digest plasmid pNIC28-Bsa4_DVC1 1-240 to generate linear backbones. The reaction is detailed below. According to the manufacturer's instructions. Incubate at 37℃ for ~2h. After digestion, the cleaved product with right size was cut from the TAE agarose gel, then purified by Gel Extraction Kit (NEB). DNA concentration is quantified by NanoDrop™ One (Thermo Fisher Scientific).

Double Digestion:

NcoI	1μL
BamHI	1μL
plasmid	5μL
rCutSmart™ Buffer	5μL
H_2O	38μL

Gene Ligation:
Purified gene fragments were ligated into the linearized plasmid backbones by using NEBuilder HiFi DNA Assembly Master Mix (NEB). Gene-specific primers were

designed for gene amplification. The gene fragments can be assembled to linear vectors by homologous recombination. The ligation mixture was incubate at 50 °C for 20 min.

Ligation Reaction:

NEBuilder HiFi DNA Assembly Master Mix	10µL
insert DNA:	0.2 pmol
digested vector backbone:	0.1 pmol
ddH_2O: to a final volume of 20µL	
DNA molar ratio is vector : insert = 1 : 2	

Transformation:

2 µL ligation product was applied for transformation into 25 µL DH5a competent cells (NEB). Incubate on ice for 30 min. Then heat shock at 42°C for 30 seconds. the cells were immediately placed on ice for 5 min, followed by adding 450 µL SOC. The cells were incubated with constant shaking at 37°C for 1 hour. Subsequently, 200 µL of the transformed cells were spread onto the agar plates containing the corresponding antibiotics and incubated overnight at 37°C.

Colony PCR and Sequencing:

Colony PCR was performed to screen the plasmids carrying the correct insertion using OneTaq polymerase (NEB) with specific primers. Positive plasmids were extracted by a mini-prep kit and sent for sequencing.

2.2 SDS-PAGE

Protein samples were mixed with 10 µL of 5x Laemmli Loading Dye and denatured at 95°C for 10 min. The samples were loaded onto the precast gel (Thermo Scientific) along with the protein marker (Precision Plus Protein™ Dual Xtra Prestained Protein Standards Biorad). Electrophoresis was performed at 120-140V in a mini-PROTEAN Tetra electrophoresis tank (Bio-Rad) filled with 1x MOPS buffer (Invitrogen). After electrophoresis, the gel was stained with InstantBlue (Expedeon) to visualize the proteins. iBright 1500 (Invitrogen) was used for imaging and analysis.
5x Laemelli Loading Dye: 15% β-mercaptoethanol, 9% SDS, 30% glycerol, 0.3 mg/ml bromophenol blue, 150 mM Tris-HCl, pH 6.8

2.3 Protein Expression

The BL21 *E.coli* glycerol stock carrying the target plasmid was incubated to 5ml LB with corresponding antibiotics. Incubate at 37℃, overnight. On the second day, add the seeding culture was added to 1L - 2L LB. Incubate at 37℃, 220 rpm until OD_{600} reaches to 0.8 ~ 1.0. For proteins expressed from the pGEX and pNIC28 vectors, induction was carried out by directly adding Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.25-0.5 mM, followed by incubation overnight at 20°C with constant shaking at 180 rpm. For proteins expressed from the pCold I and pCold-TF vectors, cold induction was performed by placing the culture at 15°C for 30 min. IPTG was then added to a final concentration of 0.25-0.5 mM, followed by incubation at 15°C with constant shaking at 180 rpm for 24 hours. The next day, the bacterial cultures were harvested by centrifugation at 5000 rpm for 10 min. The bacterial pellet was snap-frozen on dry ice, and stored at -20°C.

2.4 Protein Purification

2.4.1 Sonication

Bacteria pellets were resuspended in H1 Resuspension Buffer with additionally 200 uM PMSF and one EDTA-free protease inhibitor cocktail tablet (ROCHE). It was then processed by sonication and kept in an ice-water mixture. The lysate is centrifuged at 20Krpm for 30 min. The supernatant is collected for purification.
Sonication condition: 40% power 1 min-on/ 1min-off, 3 cycles
H1 Resuspension Buffer: 50 mM Tris-HCl, pH7.4, 500 mM NaCl, 10% Glycerol, 0.5 mM TCEP-HCl

2.4.2 Ni-NTA Affinity Chromatography and Tag Removal

The supernatant was loaded onto the pre-equilibrated Ni-NTA affinity chromatographic column. Then the column was washed with Wash Buffer A at 2ml/min in the AKTA pure system until the UV absorbance became steady. Wash Buffer B was used to elute the target protein. Selected peak fractions were examined by SDS-PAGE. Fractions containing the pure target protein were concentrated to desired.

To remove His tag, the protein was transferred to a dialysis bag and incubated with TEV protease. The sample was dialysed against the Dialysis Buffer for overnight cleavage and buffer exchange to reduce imidazale concentration.
Dialysis Buffer: 50 mM HEPES, pH 7.4, 150 mM NaCl, 0.5 mM TCEP-HCl

After dialysis, the free His tag is removed by Ni-NTA affinity chromatography again from the cleavage mixture. The flow-through containing the tag-free target protein was concentrated to desired volume for size-exclusion chromatography.
Column: 5 mL HisTrap crude FF (Cytiva)
Wash Buffer A: H1 Resuspension Buffer containing 30 mM imidazole
Wash Buffer B: H1 Resuspension Buffer containing 300 mM imidazole

2.4.3 Size-Exclusion Chromatography (SEC)

To further polish the protein, SEC was performed on an AKTA pure system. Concentrated sample was injected onto the SEC column and eluted with SEC buffer. Peak fractions were examined by SDS-PAGE. Then the pure target protein was collected and concentrated. Protein aliquots were snap-frozen on dry ice and stored at -80℃.

SEC column: Superose 6 Increase (24 mL) or HiLoad™ 16/600 Superdex™ 75pg (120 mL)

SEC Buffer: 50 mM HEPES, pH 7.4, 500 mM NaCl, 0.5 mM TCEP-HCl

2.4.4 Protein Quantification

Protein concentration was determined at A280 by NanoDrop. The calculation formula is as follows.

M=A280/Ext.coefficient

Concentration (mg/ml)=A280×Molecular weight/Ext.coefficient

2.4.5 Proteins Purified in This Thesis

Table 1. List of molecular weight and purification method of the proteins

Plasmid	Protein Name	Molecular Weight (kDa)	Tag	Tag Removed	Purification Method
pNIC28-ISG15	Tag-free ISG15	17.2	His	Yes	His*/SEC/His*/SEC
pCold I_2xISG15-C78S-mH1-Ncys	2×ISG15-H1	57	His	No	His*/SEC
pCold I_Ub-mH1-Ncys	Ub-H1	31.5	His	No	His*/SEC

His*: Affinity chromatography by Ni-NTA column

SEC: Size-exclusion chromatography

2.5 Cy5 Labeling

The H1 substrates were labeled with Cy5 maleimide dye for fluorescence detection in subsequent SPRTN assay. 5 µL Cy5 maleimide dye (dissolved in DMSO, Cytiva) was added into the protein sample, incubate at 4℃ for 2h or overnight. The samples were buffer-exchanged by the concentrator with the right cut-off until the flow-through has no visible blue color from the dye and concentrated to desired concentration. Protein aliquots were snap-frozen on the dry ice and stored at -80℃.

2.6 Making Model dsDNA-H1 DPC

The ssDNA (sequence provided below) was diluted in DEPC-treated water to a final concentration of 100 µM. To make FAM-Amide-dsDNA, 200 µL each below ssDNA was mixed together and incubated at 95℃ for 10 min and then cooled down to room temperature naturally. The product was then purified using a Zeba Spin desalting column (7 kDa, 3 mL, Thermo Scientific, 89889) (pre-equilibrated with Crosslink Buffer) to remove any unbound ssDNA. Centrifuge at 1,000g for 2 min. The flow-through containing the purified dsDNA was collected.

Table 2. List of sequence of ssDNA

ssDNA	Sequence
FAM-Turner-20bp-F	5'-FAM-ACGCCTGAAGAGTCTGGTGA-3'
Amide-WS-23bp-R	3'-cctTGCGGACTTCTCAGACCACT-Amide-5'

Crosslink Buffer: 100 mM NaH_2PO4, 150 mM NaCl, 5 mM EDTA, pH7.3

To crosslink dsDNA with succinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC), 20 µL of 0.1 M SMCC was added to the above dsDNA, resulting in a final concentration of 5 mM. The mixture was left at room temperature in the dark for 2 hours. After the incubation, a new Zeba Spin desalting column (7 kDa, Thermo Scientific) was used to remove excess SMCC. The flow-through was collected as dsDNA-SMCC.

To crosslink H1 with dsDNA-SMCC, 5-fold excess of dsDNA-SMCC was mixed with H1 and incubated overnight at 4℃ in the dark. The next day, the mixed product was loaded onto a pre-equilibrated Resource S column for further purification.

Ion Exchange (IEX):
The above crude mixture was loaded onto a 5ml HiTrap SP FF column by 100% IEX Buffer A. The flow-through was collected. Then 20% IEX Buffer B was used to wash column until UV curve and conductivity curve became steady. The proportion of IEX buffer B was gradually increased to 100% to elute the target protein. The purity of the dsDNA-H1 DPC was confirmed by SDS-PAGE analysis, and the concentration was measured by the absorption at A280. Protein aliquots were snap-frozen on dry ice and stored at -80℃.
IEX column: HiTrap SP FF (5ml)
IEX Buffer A: 50mM HEPES pH7.4, 150mM NaCl, 10% Glycerol, 0.5mM DTT
IEX Buffer B: 50mM HEPES pH7.4, 1M NaCl, 10% Glycerol, 0.5mM DTT

2.7 SPRTN Assay

SPRTN assay is performed to compare the activity of SPRTN towards different substrates. Reactions were prepared on ice to inhibit SPRTN activity temporarily. Mastermix was prepared for the reaction, and aliquoted according to the numbers of time points. At 0 min time point, put sample immediately on dry ice to snap freeze and terminate the reaction. The rest of the samples were all incubated at 30℃ in a PCR cycler. When time's up, each sample was immediately kept on dry ice to stop the reaction. Each sample was mixed with 6 µL 2X Loading buffer (4XNuPAGE buffer diluted with equal amount of 1M DTT). Then the samples were denatured at 95℃ for 10 min using the PCR cycler and 6 µL was loaded onto the precast SDS-PAGE gels (NuPAGE™ 4 to 15%, Bis-Tris, 1.0 mm, Midi Protein Gels, 15 wells).
Assay buffer: 25 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.5 mM TCEP-HCl

Table 3. List of DNA, proteins and time points used in SPRTN assay

SPRTN assay	SPRTN activity assay towards H1-DPC	SPRTN kinetics towards H1 and Ub_4-H1 and 2xISG15-H1		
substrates	FAM-dsDNA-H1-Ncys	Cy5-His-H1-Ncys	Cy5-Ub_4-H1	Cy5-2xISG15-H1
enzyme	Tag-free SprT core (26-240)	Tag-free SprT core (26-240)		
ubiquitin or ubiquitin-like protein	M1-Ub_4, ISG15	-		
dsDNA	-	*Turner-20bp-F* 5'-ACGCCTGAAGAGTCTGGTGA-3' *WS-23bp-R* 3'-cctTGCGGACTTCTCAGACCACT-5'		
Time points	0/0.5/1h	0/1/2.5/5 min		

3. Results

3.1 The Purification of Ubiquitin-fused H1 Substrates

His-tag contains six histidine residues, which form a unique structure that facilitates protein purification. The His-Tag at N-terminus is compatible with the bacterial transcription and translation machinery. Additionally, its small molecular weight of approximately 0.84 kDa usually does not interfere with the structure and function of the recombinant protein. Therefore, His tag was used for protein purification in this study.

Through protein expression in *E. coli* and various protein purification techniques, the purification of ISG15, 2×ISG15-H1, Ub-H1 and dsDNA-H1 DPC was well optimised. Table 1 shows the information of each protein.

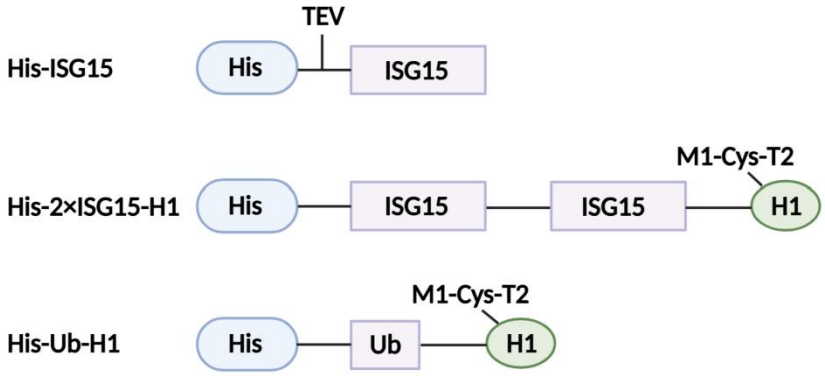


Fig 3.1 Schematics of His-ISG15, His-2×ISG15-H1 and His-Ub-H1. *Cys is introduced to the N-terminal of H1 to assist Cy5-labeling.*

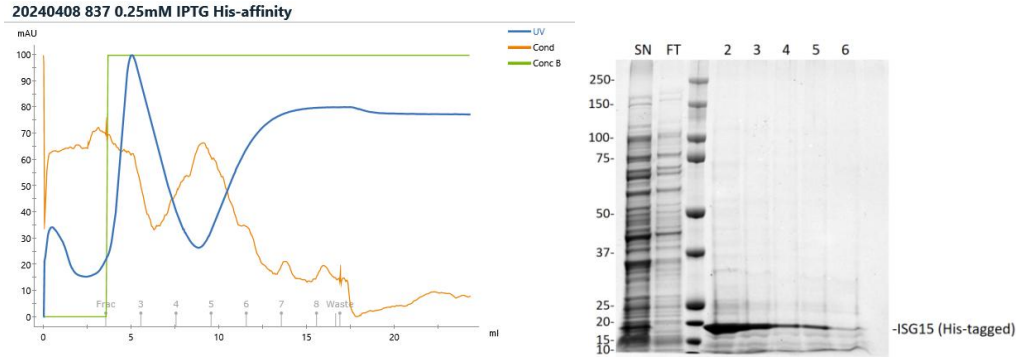
3.1.1 Purification of ISG15

The ubiquitin-like protein ISG15 serves as a negative control to compare whether SPRTN is activated by chains formed by ubiquitin or ubiquitin-like protein. The molecular weight of His-tagged ISG15 is 19.7 kDa. The UV curve of Ni-NTA affinity chromatography showed a prominent peak, indicating a good expression of His-tagged ISG15. The SDS-PAGE gel shows that most of the His-tagged ISG15 can be bound to the Ni-NTA column and eluted by a high concentration of imidazole. SDS-PAGE gel showed some impurities, which need to be further polished (See Fig. 3.2 a).

Subsequently, the crude products from Ni-NTA affinity chromatography were purified using SEC. The UV curve of the SEC showed six peaks. SDS-PAGE gel indicates that the target product was mainly accumulated in peak 4 (fractions 12-14 as shown in Fig. 3.2 b). However, the impurities can not be removed very well. A band between 10-15 kDa in SEC step which not present in the affinity purification step. this may be due to the inability of the 4-20% gel used in affinity chromatography to separate ISG15 from the impurities, whereas a 4-12% gel could achieve better separation. Alternatively, it is possible that ISG15 was partially degraded *in vitro*.

To further assist the purification, TEV cleavage of the His tag was performed. Products from peak 4 were incubated with TEV protease overnight for tag-removal.

(a)



(b)

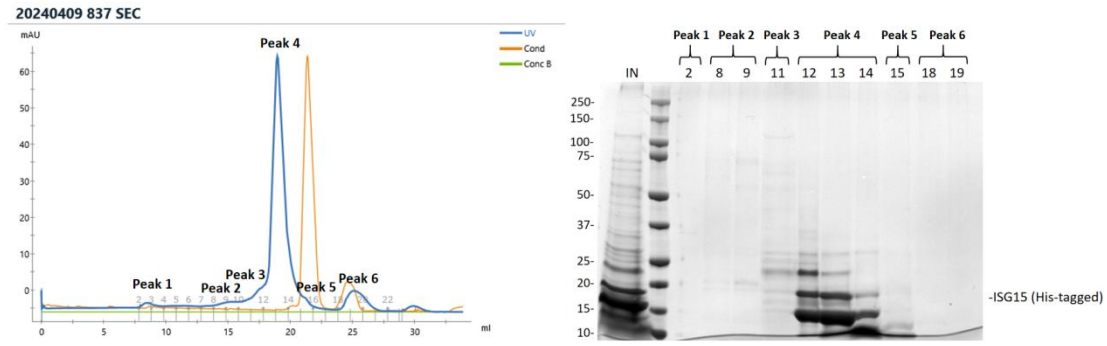


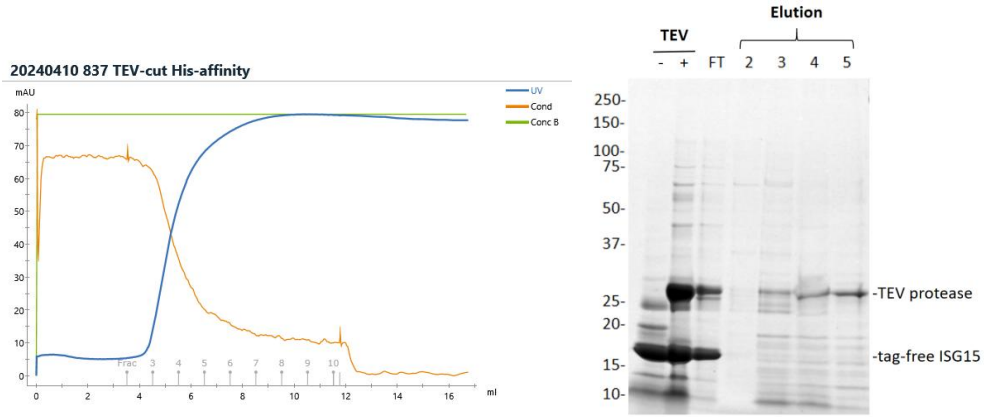
Figure 3.2 Purification of His-tagged ISG15. *Purification profile and SDS-PAGE analysis of (a)*

Ni-NTA affinity chromatography and (b) SEC of His-tagged ISG15. The SEC column used was Superose 6 Increase (24 mL). The band between 15 and 20 kDa is the His-tagged ISG15 (~19.7kDa). The products from peak 4 were concentrated to approximately 300 µL using a 10K concentrator.

To trap free His tag and any un-cleaved His-ISG15, Ni-NTA affinity chromatography was performed again. The UV curve of Ni-NTA affinity chromatography by imidazole elution shows a small peak, which contains only the free His tag and His-TEV, indicating the successful cleavage of almost all the His-tagged ISG15. The flow-through contains mostly the tag-free ISG15 with some TEV contamination (See Fig. 3.3 a). Therefore, the following SEC procedure is needed.

From the SEC profile, 4 peaks appeared. SDS-PAGE analysis showed that tag-free ISG15 was mostly eluted in peak 3 (fraction 12-14 in Fig. 3.3 b). Although a small amount of impurities remained, the purity is sufficient for the subsequent SPRTN assay.

(a)



(b)

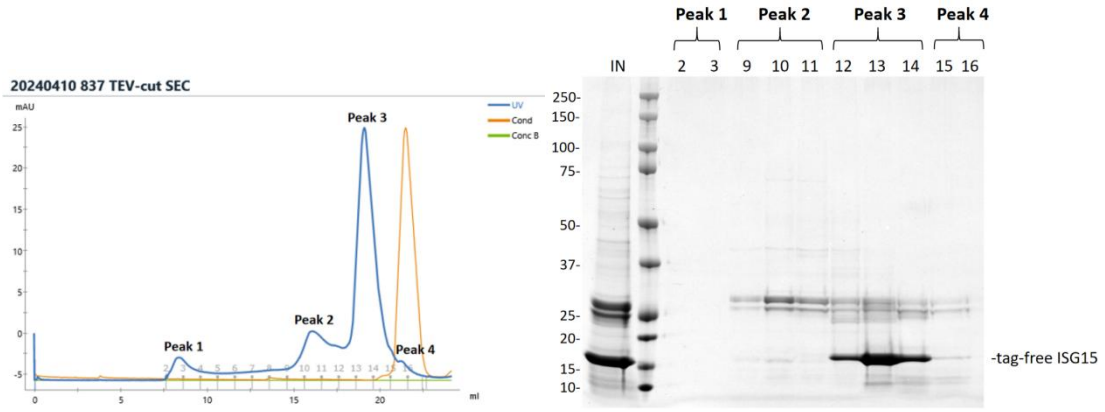


Figure 3.3 Purification of tag-free ISG15. *Purification profile and SDS-PAGE analysis of (a) Ni-NTA affinity chromatography and (b) SEC of tag-free ISG15 after TEV-cleavage. The SEC column used was Superose 6 Increase (24 mL). The band between 15 and 20 kDa is the tag-free ISG15 (~17.2kDa).*

3.1.2 Purification of 2×ISG15-H1

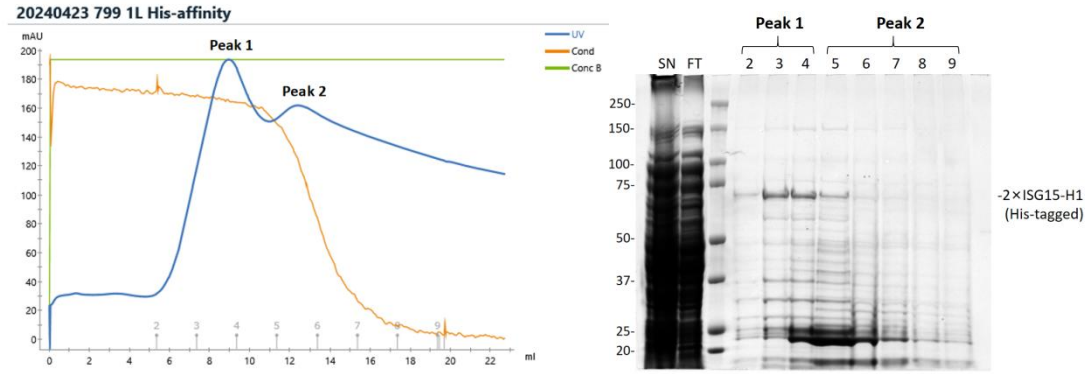
The 2×ISG15-H1 is derived from a plasmid that expresses His-tagged H1 fused with 2×ISG15 at the N-terminus of H1. The molecular weight is 57.2 kDa.

Given the high proportion of positively charged residues in H1, a high-salt elution was used for better separation. The UV curve of Ni-NTA affinity chromatography showed two small peaks. From the SDS-PAGE analysis, 2×ISG15-H1 is mostly enriched in peak 1. Fractions from peak 1 were further purified by SEC (See Fig. 3.4 a).

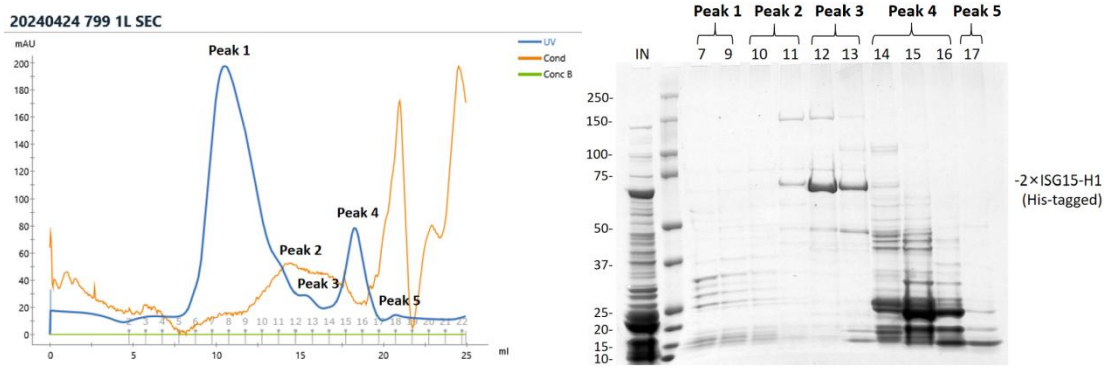
From the SEC profile, there were 5 peaks. SDS-PAGE analysis showed that 2×ISG15-H1 was mostly eluted in peak 3 with less impurity (fraction 11-13 in Fig. 3.4 b). Considering that the H1, Ub-H1, and Ub4-H1 substrates used in subsequent SPRTN assay have not been stripped of their His-tags, there is no need to remove the His-tag from ISG15-H1 either. Fraction from peak 3 was selected and concentrated for Cy5 labeling.

As seen in Fig 3.4 c, Cy5 labeling of 2×ISG15-H1 is very successful. The majority of the target protein can be labeled with Cy5 fluorescence, but not the impurities (Fig. 3.4 c).

(a)



(b)



(c)

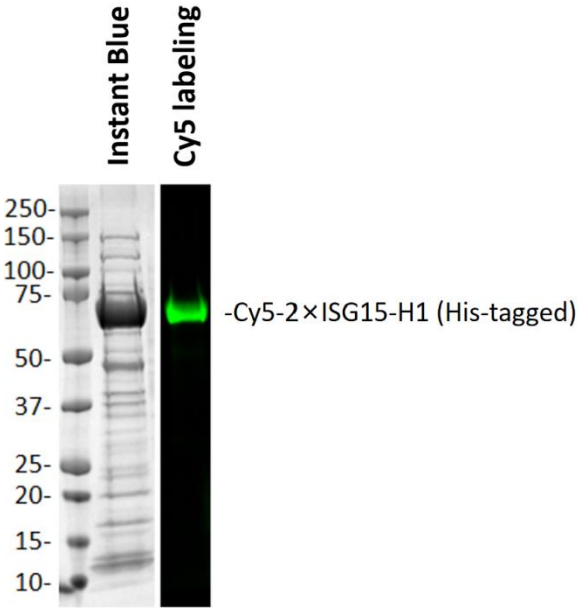


Figure 3.4 Purification of His-tagged 2×ISG15-H1. *Purification profile and SDS-PAGE analysis of (a) Ni-NTA affinity chromatography, (b) SEC of 2×ISG15-H1 (SEC column: Superose 6 Increase (24 mL)) and (c) Cy5 labeling of 2×ISG15-H1. The band between 50 and 75 kDa is the 2×ISG15-H1 (~57.2kDa).*

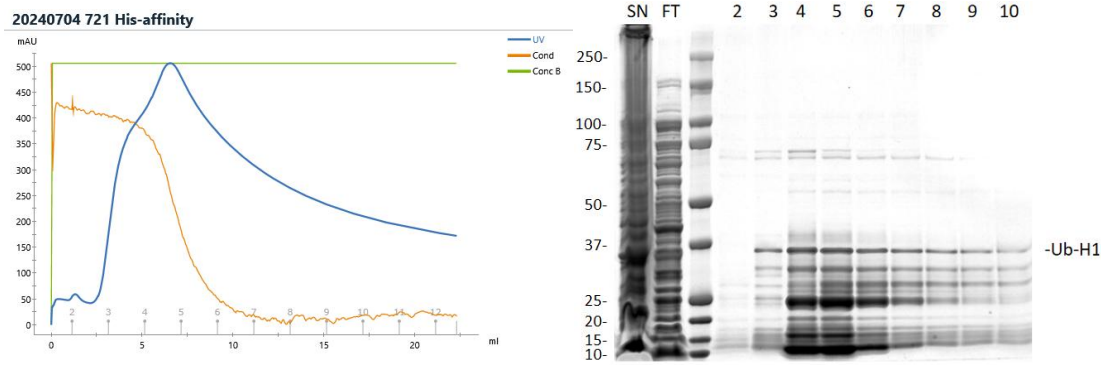
3.1.3 Purification of Ub-H1

The Ub-H1 is derived from fusing monoUb to the N-terminus of His-tagged H1, with the His tag located at the C-terminus of H1. The molecular weight is 31.5 kDa. The purification of Ub-H1 was also performed by Ni-NTA affinity chromatography and SEC.

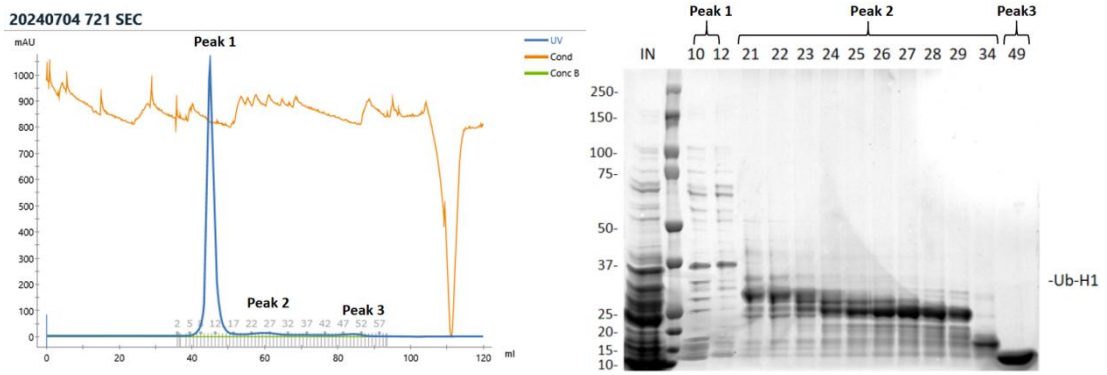
The SDS-PAGE analysis of Ni-NTA affinity chromatography showed that a distinct band of full-length Ub-H1 but with a large amount of impurities. This indicated that Ub-H1 is quite unstable.

Nevertheless, SEC was performed to remove impurities as much as possible. Three peaks appeared in the SEC profile. However, SDS-PAGE analysis showed that the purification by SEC could not separate the target protein from degradation products that have closer molecular weight to Ub-H1 (Fig. 3.5 b). Most of the Ub-H1 was eluted in fraction 21-22 which was the applied to Cy5 labelling (Fig. 3.5 b). The results from Cy5 labelling were encouraging, where Ub-H1 has been successfully labeled with Cy5 (Fig. 3.5 c).

(a)



(b)



(c)

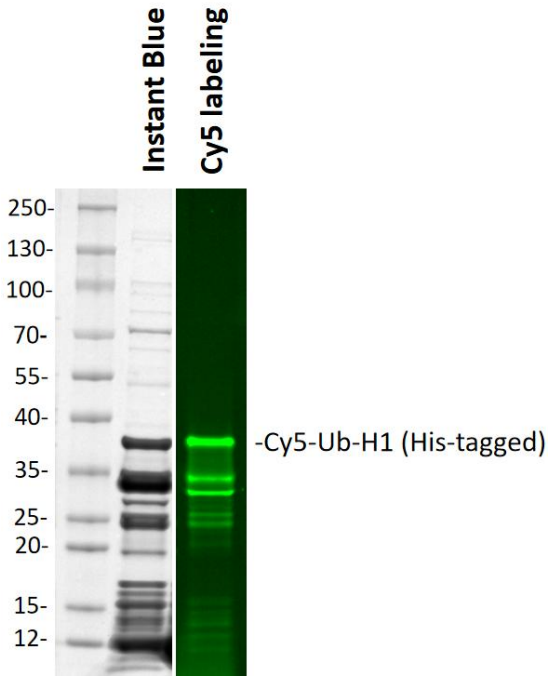


Figure 3.5 Purification of His-tagged Ub-H1. *Purification profile and SDS-PAGE analysis of (a) Ni-NTA affinity chromatography, (b) SEC of Ub-H1 (SEC column: HiLoad 16/600 Superdex™ 75pg (120 mL)) and (c) Cy5 labeling of Ub-H1 (Courtesy of Dr. Wei Song).*

3.1.4 Purification of dsDNA-H1 DPC

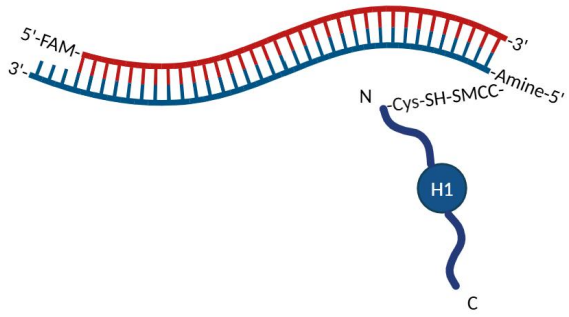
To covalently attach H1 to dsDNA as the model DPC for SPRTN proteolysis, FAM labeled dsDNA was crosslinked to the N-terminus of H1 via the Succinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC) linker (Fig 3.6 a).

To separate dsDNA-H1 DPC from any unreacted dsDNA, cation exchange chromatography was performed (Fig. 3.6 b, c). The UV curve from Resource S column washed by low-salt buffer revealed a significant peak in the flow-through, indicating an excess of free dsDNA that could not bind to the column due to its negative charge. dsDNA was labeled by FAM. This observation was corroborated by SDS-PAGE analysis with FITC scanning, where only a signal corresponding to free dsDNA was detected.

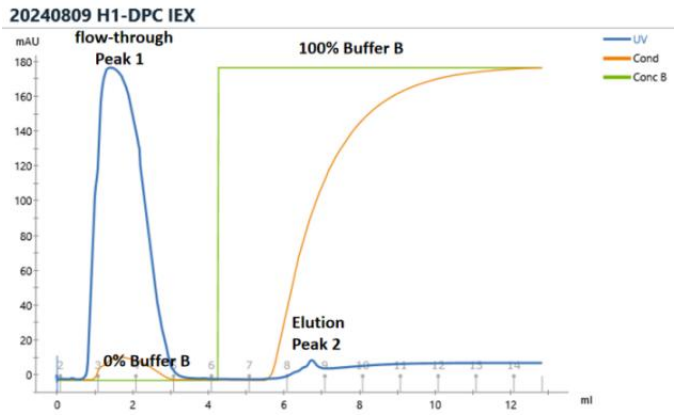
Elution with high-salt buffer produced a small peak, suggesting the presence of the crosslinked dsDNA-H1 species. It was further validated by SDS-PAGE, which displayed a single FITC-visible band (Fig. 3.6 b, c).

Instant Blue staining revealed two bands in the eluent (fraction 8): one corresponding to dsDNA-H1 DPC and another to free H1, indicating that not all H1 was crosslinked to the dsDNA. Additionally, the FITC-visible dsDNA in flow-though is strong but DPC band in eluent is weak. These results indicated that the yield of dsDNA crosslinked with H1 was relatively low and required further optimization.

(a)



(b)



(c)

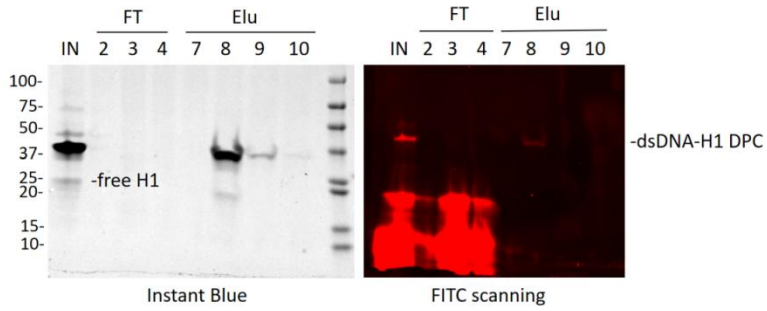


Figure 3.6 Purification of dsDNA-H1 DPC. *(a) Model of dsDNA crosslinking to N-terminal of H1. Cys was introduced to H1 N-terminal for the crosslinking. (b) Purification profile and (c) SDS-PAGE analysis of IEX of dsDNA-H1 DPC. (The IEX column: 1 mL Resource S column) The band around 37 kDa is the dsDNA-H1 DPC. Fractions selected from peak 2 (including fraction 8 and 9) from IEX were concentrated as the final product.*

3.2 Linear Ub_4, but not ISG15, Activates SprT Cleavage Towards H1-DPC

H1-dsDNA conjugate is used here as a model DPC substrate for SPRTN proteolysis in combination with the addition of M1-Ub_4 or ISG15 (Fig. 3.7 a).

In comparison to condition a and b, FITC analysis shows that there were only a limited amount of DPC cleavage products (about 10kDa) when SprT core was added. This finding indicates that the *in vitro* purified SprT core is active and capable of cleaving DPCs.

The addition of M1-Ub_4 in group c shows an increased amount of DPC cleavage products (about 10kDa). In contrast, ISG15 (group d) did not exhibit this effect. This result indicated that SPRTN proteolysis towards the model H1-DPC substrate can be activated with the addition of M1-Ub_4, but not ISG15 which is in line with the observation as seen in Fig 1.7.

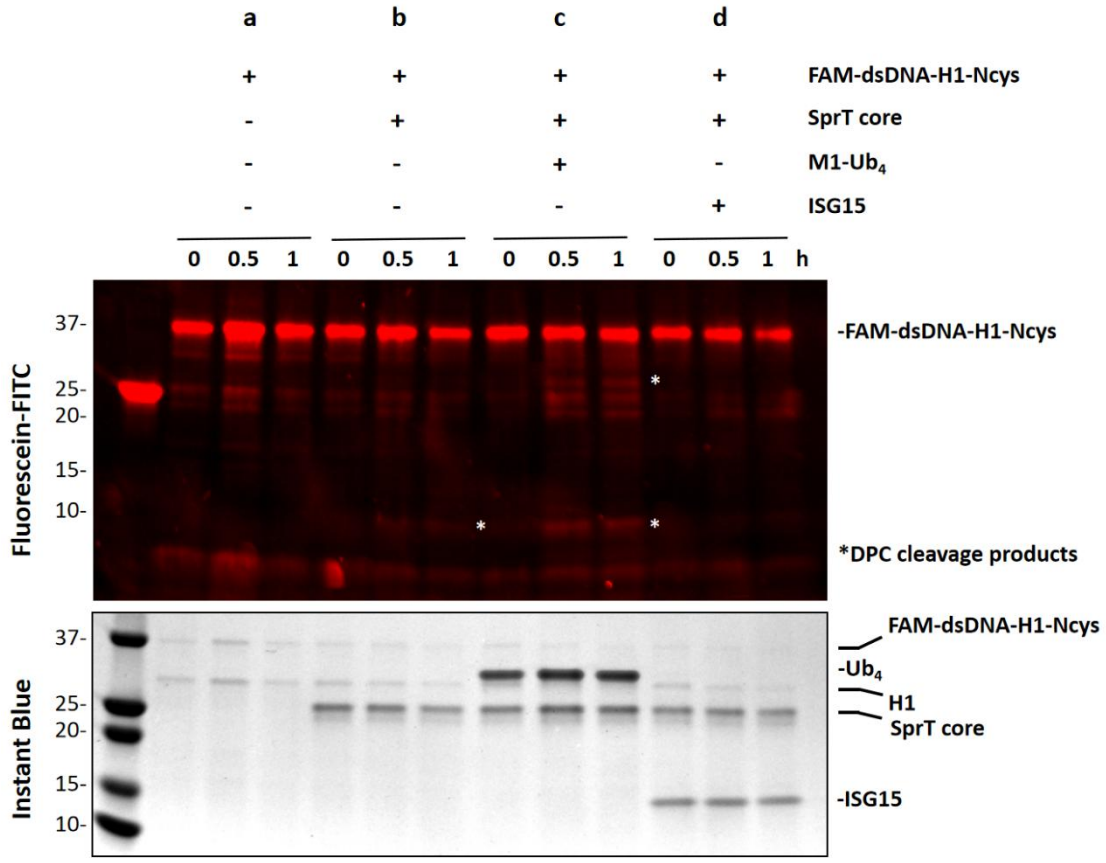


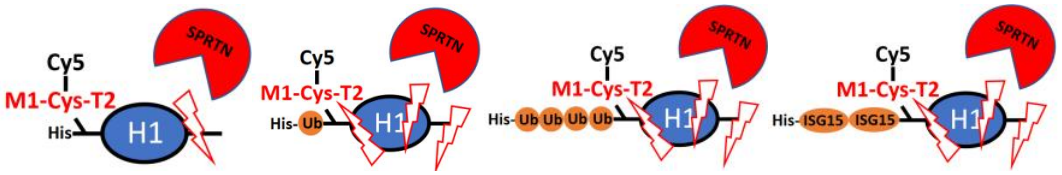
Figure 3.7 Linear Ub_4, but not ISG15, activates SprT core cleavage towards H1-DPC. *SPRTN assay towards H1-DPC. The dsDNA is labeled with FAM (red). Ncys is a cysteine (cys) labeled at the N-terminus of H1.*

3.3 SprT Core Cleaves Ub_4-H1 within 5 minutes, but not H1, Ub-H1 or 2xISG15-H1

Previous work showed that SPRTN proteolysis towards H1 is activated by the addition of M1-Ub_4, but not monoUb, M1-Ub_2 or ISG15 in the presence of DNA (Fig 1.7). To better understand SPRTN activation on ubiquitin-modified substrates, H1 constructs fused with monoUb or linear tetra-ubiquitin or 2xISG15 were designed and purified.

Compared to the control group without dsDNA, the addition of dsDNA could initiate the reaction of SprT proteolysis when Cy5-Ub_4-H1 was used as a substrate. SPRTN assay showed the full length of Cy5-Ub_4-H1 can be cleaved by SprT within 5 min in the presence of dsDNA, but the cleavage of H1, Ub-H1 and 2xISG15-H1 cannot be observed within such a short time course.

(a)



(b)

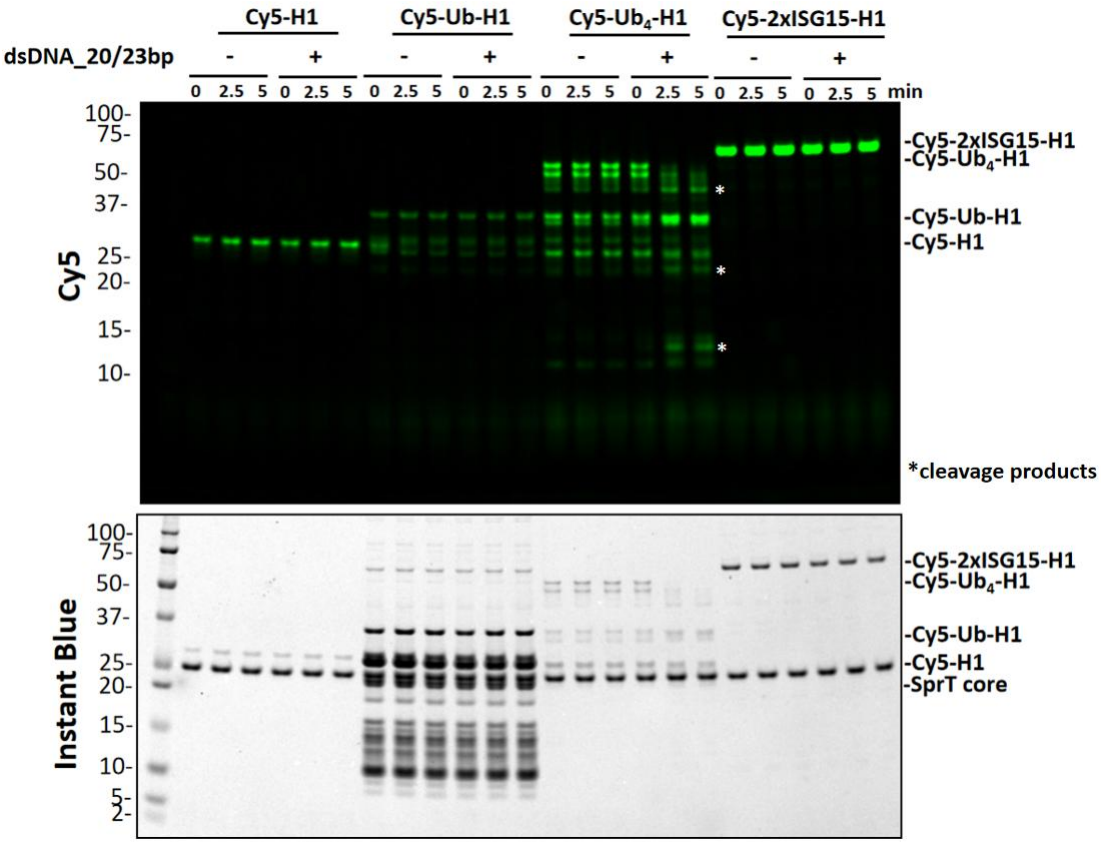


Figure 3.8 SprT core cleaves Ub_4-H1 in a faster manner than H1, Ub-H1 and 2xISG15-H1. *(a)*

Model of SPRTN cleavage on Cy5-H1, Cy5-Ub_4-H1 and Cy5-2xISG15-H1. Cy5 was introduced to H1 N-terminal. (b) SDS-PAGE analysis of SPRTN assay. The full-length Cy5-Ub_4-H1 can be cleaved by SprT core within 5 min, whereas H1, Ub-H1 and 2xISG15-H1 cannot.

4. Discussions

DPC repair pathway is considered as an important mechanism for maintaining genome stability. SPRTN is an essential enzyme for the repair. However, how substrate specificity regulation of SPRTN. To better understand SPRTN activation by modified substrates, I purified H1-DPC substrate, H1 constructs fused with mono-Ub, linear Ub_4 or 2xISG15 *in vitro*. The results show that SPRTN activity towards the model H1-DPC substrate is enhanced by the addition of Ub_4 chains on substrate, but not the ubiquitin-like protein ISG15. Furthermore, Cy5-Ub_4-H1 can be rapidly cleaved by SPRTN , while Cy5-H1, Cy5-Ub-H1 and Cy5-2xISG15-H1 cannot. These findings suggest that poly-ubiquitin chains are involved in regulating SPRTN cleavage on substrates.

4.1 Protein Purification

During the purification of ISG15, some TEV contamination was observed in the final product (Fig 3.3 b). We tested and optimized several purification methods such as IEX, but the TEV contamination persisted. However, the SEC profile (Fig 3.3 b) revealed that a significant portion of the free ISG15, along with a small amount of residual TEV protease was eluted in peak 3. Given that the concentration of TEV protease is quite low, the ISG15 obtained was sufficiently pure for subsequent SPRTN assays. Additionally, the TEV protease is site-specific, meaning that it does not interfere with the SPRTN proteolysis on substrates. It ensures that the assays were conducted with minimal interference from contaminants, allowing for accurate and reliable results in SPRTN proteolysis activity.

In Figure 3.4 c, there were several bands in the instant blue staining, while only one band with Cy5 signal. Inpurities were observed in the final product. They are likely to be unspecific *E. coli* protein co-purified with the intended product. Additionally, some extra bands may correspond to *E. coli* proteins that bind non-specifically to the His tag.

During the experiment, we found that protein induction in *E.coli* by using different volume of LB resulted in varying amounts of degradation products. Additionally, Cy5-Ub-H1 stored at 4℃ overnight was nearly completely degraded. It shows that the degradation of Cy5-Ub-H1 occurs both in *E.coli* and *in vitro*, making it hard to remove degradation products. During the protein induction process, we optimized IPTG concentration and shortened protein expression time to minimize the target protein degradation in *E.coli*. To further optimize it, we purified the Ub-H1 within one day to reduce extracellular degradation. Given the high proportion of positively charged residues in H1, we used high-salt elution buffer for better separation. Despite these efforts, some impurities remained. We also tested various purification methods, including SEC and IEX, but the issue persisted.

An important observation is the absence of the the FITC-visible dsDNA-H1 DPC band in both the flow-through and elution fractions. This could due to the fact that a small amount of very diluted elution sample was loaded on the gel. This can be addressed by running a more concentrated sample. However, it is also possible that the yield of crosslinked H1-dsDNA in the elution is low. One potential reason could be an issue with the binding of the H1-DNA crosslink to the cation exchange column. As shown in the UV profile from the Resource S column, a significant peak in the flow-through indicates that dsDNA binds to unspecific proteins, which carries a negative charge, did not bind effectively to the column during the low-salt wash. This suggests that the crosslinked H1-DNA complex might have a weaker interaction with the column, possibly due to incomplete crosslinking or the formation of a less stable complex that is prone to dissociation under the conditions used. To address this issue, adjusting the salt concentration in the wash buffer may improve the binding of the dsDNA-H1

DPC to the column, thereby reducing its loss in the flow-through. Several approaches can also be considered. First, optimizing the ratio of dsDNA-SMCC to H1 during the crosslinking reaction is crucial. Currently, a 5-fold excess of dsDNA-SMCC is used. Increasing this excess may enhance the chances of successful crosslinking. Furthermore, adjusting the incubation time and temperature may improve the efficiency of the crosslinking reaction.

4.2 Ubiquitin Chains Activate SPRTN Proteolysis of H1-DNA Conjugate

SPRTN can target various DPC substrates. In the previous work by Vaz et al., mass spectrometry analysis of SPRTN degradation products did not detect any specific cleavage sites (31). Moreover, the active site was found on SPRTN, but very little specificity was observed in SPRTN. This may indicate that SPRTN lacks substrate specificity and requires more strict regulation. Fig 1.7 and Fig 3.8 showed that DNA is required to initiate the SPRTN proteolysis indicating an important role that DNA's presence regulates SPRTN activity.

As the active site of SPRTN catalytic domain is located in a narrow groove, the proteolysis of the structured substrates by SPRTN is considered to be not efficient (28, 31). However, due to the covalent attachment between the protein and DNA in the DPCs, the recognition sites of protease may be obstructed by the steric structure or crosslinking sites, making DPCs more difficult to be recognized and resolved. In this study, a well-defined model dsDNA-H1 DPC substrate was produced *in vitro* (Fig 3.5). DPCs are formed by covalent attachment between FAM labeled dsDNA and the N-terminal of H1, providing a crucial tool for investigating SPRTN protease activity. Similar to the unconjugated H1 substrate (Fig 1.7), free M1-Ub_4 could also stimulate SPRTN activity towards the DPC (Fig 3.8). It demonstrates that ubiquitin also plays a regulatory role in SPRTN proteolysis of DPCs.

The N-terminal truncation of SPRTN (SprT core) was used in this study, indicating that the N-terminal region of SPRTN mediates the Ub-activation effect. All the data so far suggest possibly a dynamic regulation of SPRTN cleavage on substrates with post-translational modifications. Therefore, model H1 substrates with ubiquitination were designed to validate this theory.

The ubiquitin-fused H1 serves as a model substrate for the SPRTN proteolysis. The results from section 3.2 and 3.3 demonstrated that M1-Ub_4 can not only activate SPRTN proteolysis in its free (unbound) form but also function when conjugated to the substrates. It suggests that there may be ubiquitin-interacting elements within the SprT domain, aiding SPRTN in better recruiting ubiquitinated substrates for proteolysis.

The results from section 3.2 and 3.3 showed that the ubiquitin-like protein ISG15 and 2×ISG15 fused to H1 cannot activate SPRTN proteolysis. ISG15 is structurally similar to M1-diUb. It also contains a β-grasp protein fold along with a conserved C-terminal Gly-Gly motif (see Fig. 4.1) (83, 84). The molecular weight and structure of 2×ISG15 are similar to M1-Ub_4, which is why 2×ISG15-H1 was used as a control to study SPRTN proteolysis in section 3.3. It is evident that the sequences or structures in tetra-ubiquitin which are distinct from 2×ISG15, are involved in the activation of the SprT proteolysis. However, to determine the specific mechanism of how poly-ubiquitin activates SprT requires further investigation.

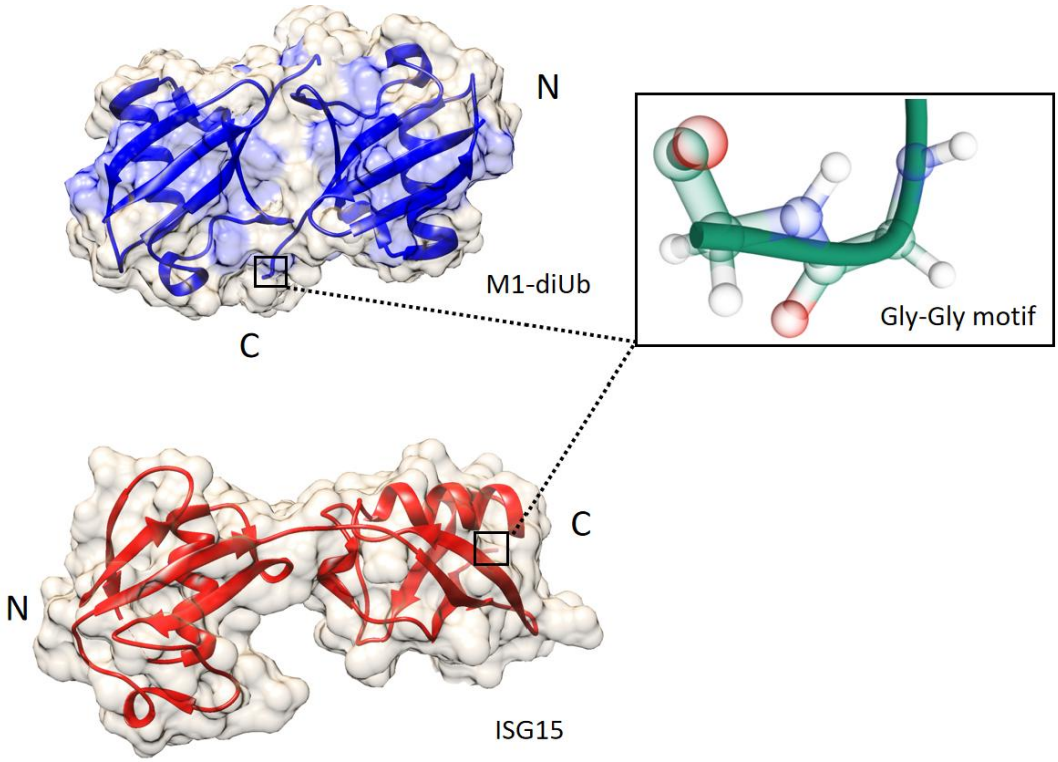


Figure 4.1 Secondary structure of M1-diUb and ISG15. *ISG15 (PDB: 3PSE) is structurally similar to M1-diUb (PDB: 2BGF). ISG15 consists of two ubiquitin-like domains connected by a flexible linker (84). Each ubiquitin-like domain contains a β-grasp protein fold and a conserved C-terminal Gly-Gly motif, which is similar to M1-diUb (83). The M1-diUb is colored blue, while the ISG15 is colored red.*

As shown in Fig 3.8, M1-Ub_4 enhances SPRTN activity towards the DPC substrate. I hypothesize that M1-Ub_4 could bind to the SPRTN protease in a manner that induces a conformational change in the SprT core, making the active site of the SprT core more accessible to its substrate. Polyubiquitin binding acts as a molecular switch, facilitating more efficient cleavage of the tightly crosslinked DNA-protein complex. However, the exact mechanism needs further investigation.

The C terminal of SPRTN contains UBZ domain, which regulates SPRTN mono-ubiquitination (35). Monoubiquitinated SPRTN is excluded from chromatin. The induction of DPCs triggers SPRTN deubiquitination and its subsequent relocalization to chromatin (36). The E3 ligase that regulates this ubiquitin switch remain unidentified. SPRTN assay (Fig 1.7) revealed that free polyubiquitin promotes the N-terminus of SPRTN (SprT) in substrate cleavage, whereas free monoubiquitin or ubiquitin-like proteins do not. This may indicate that the N-terminus of SPRTN regulates the cleavage efficiency of ubiquitinated substrates, while the C-terminus is responsible for the localization of SPRTN protease. It is worthwhile studying different types of ubiquitination modifications on substrates regulating SPRTN protease activity in the cleavage of DPC substrates.

5. Conclusion and Outlook

In this study we have purified the ubiquitin-like protein ISG15, ubiquitin-fused H1 substrates and dsDNA-H1 DPC substrate for SPRTN proteolysis. Several reproducible SPRTN assay results were obtained, described as follows:

1) In the presence of DNA, free M1-Ub_4 chains can activate SPRTN proteolysis towards H1 and H1-DPC, but not mono-ubiquitin, short ubiquitin chains or ISG15.

2) SPRTN activity can be rapidly activated by tetra-ubiquitinated chains on the substrates.

This study focuses on investigating the effect of ubiquitin and ubiquitin-like proteins on the activation of SPRTN cleavage in processing H1 or H1-DPC substrates. These findings will advance research on the role of ubiquitin and ubiquitination in DDR pathways, particularly in DPC repair.

Potential direction:

1) Investigate how the substrate ubiquitination influences SPRTN protease activity *in vivo*. It could not only validate our findings *in vitro*, but also provide deeper insights into the physiological relevance of SPRTN activation and its potential regulation by cellular signaling pathways. Additionally, studying SPRTN activity *in vivo* could reveal how the interplay between ubiquitination and other post-translational modifications governs genome stability and orchestrates the cellular response to DNA damage.

2) Explore the effects of other ubiquitin-like proteins on SPRTN. In addition to ubiquitin, ubiquitin-like proteins such as neural precursor cell expressed developmentally down-regulated 8 (NEDD8), small ubiquitin-like modifier (SUMO) and HLA-F-adjacent transcript 10 (FAT10) are also involved in many key biological processes within cells. Investigating the roles of these ubiquitin-like proteins in regulation of SPRTN and the impact on the repair of DNA-protein crosslinks (DPCs) will be an important direction for future research.

3) Explore the interaction between SprT core and ubiquitin by biophysical and structural approaches. SprT core consists of MPD domain, ZBD domain and BR domain. Further experiments can be conducted to narrow down the functional domain that interacts with ubiquitin.

4) Explore the role of SPRTN in disease. As more and more studies on the proteolytic role of SPRTN in DPC repair have been reported, the significance of SPRTN in both physiological and pathological processes is increasingly recognized. The mechanisms of SPRTN in various diseases need further study. It could provide new insights into future protein-targeted therapies.

Bibliography

1. Antonio Carusillo, Claudio Mussolino. DNA Damage: From Threat to Treatment. Cells. 2020;9(7):1665.
2. Jackson, S., Bartek, J. The DNA-damage response in human biology and disease. Nature. 2009;461:1071–1078.
3. Klages-Mundt NL, Li L. Formation and repair of DNA-protein crosslink damage. Sci China Life Sci. 2017;60(10):1065-76.
4. H.K. Kuo, J.D. Griffith, K.N. Kreuzer. 5-Azacytidine induced methyltransferase-DNA adducts block DNA replication *in vivo*. Cancer Res. 2007;67:8248-825.
5. Pedro Weickert and Julian Stingele. DNA-Protein Crosslinks and Their Resolution. Annu. Rev. Biochem. 2022;91:157-81.
6. Kiianitsa K, Maizels N. A rapid and sensitive assay for DNA-protein covalent complexes in living cells. Nucleic Acids Res. 2013;41(9):e104.
7. Stingele J, Bellelli R, Boulton SJ. Mechanisms of DNA-protein crosslink repair. Nat. Rev. Mol. Cell Biol. 2017;18:563-73.
8. Christman JK. 5-Azacytidine and 5-aza-2 -deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. Oncogene. 2002;21:5483-95.
9. O'Brien PJ, Siraki AG, Shangari N. Aldehyde sources, metabolism, molecular toxicity mechanisms, and possible effects on human health. Crit. Rev. Toxicol. 2005;35:609-62.
10. Swenberg JA, Lu K, Moeller BC, Gao L, Upton PB, Nakamura J, et al. Endogenous versus exogenous DNA adducts: their role in carcinogenesis, epidemiology, and risk assessment. Toxicol Sci. 2011;120(Suppl 1):S130-45.
11. Nakano T, Morishita S, Katafuchi A, Matsubara M, Horikawa Y, Terato H, et al. Nucleotide excision repair and homologous recombination systems commit differentially to the repair of DNAprotein crosslinks. Mol Cell. 2007;28(1):147-58.
12. Pommier Y, Huang SY, Gao R, Das BB, Murai J, Marchand C. Tyrosyl-DNA-phosphodiesterases (TDP1 and TDP2). DNA Repair. 2014;19:114-29.
13. Pouliot JJ, Yao KC, Robertson CA, Nash HA. Yeast gene for a Tyr-DNA phosphodiesterase that repairs topoisomerase I complexes. Science. 1999;286:552-55.
14. Matthew J. Schellenberg et al. ZATT (ZNF451)-mediated resolution of topoisomerase 2 DNA-protein cross-links. Science. 2017;357:1412-1416.
15. Stingele J, Schwarz MS, Bloemeke N, Wolf PG, Jentsch S. A DNA-dependent protease involved in DNA-protein crosslink repair. Cell. 2014;158:327-38.
16. Balakirev MY, Mullally JE, Favier A, Assard N, Sulpice E, et al. Wss1 metalloprotease partners with Cdc48/Doa1 in processing genotoxic SUMO conjugates. eLife. 2015;4:e06763.
17. Reinking HK, Hofmann K, Stingele J. Function and evolution of the DNA-protein crosslink proteases Wss1 and SPRTN. DNA Repair. 2020;88:102822.
18. Kottemann, M.C., and Smogorzewska, A. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. Nature. 2013;493:356-363.

19. Orta, M.L., Calderón-Montaño, J.M., Domínguez, I., Pastor, N., Burgos-Morón, E., López-Lázaro, M., Cortés, F., Mateos, S., and Helleday, T. 5-Aza-2'-deoxycytidine causes replication lesions that require Fanconi anemia-dependent homologous recombination for repair. Nucleic Acids Res. 2013;41:5827-5836.
20. Mechilli, M., Schinoppi, A., Kobos, K., Natarajan, A.T., and Palitti, F. DNA repair deficiency and acetaldehyde-induced chromosomal alterations in CHO cells. Mutagenesis. 2008;23:51-56.
21. Karanja, K.K., Lee, E.H., Hendrickson, E.A., and Campbell, J.L. Preventing over-resection by DNA2 helicase/nuclease suppresses repair defects in Fanconi anemia cells. Cell Cycle. 2014;13:1540-1550.
22. Jasmine D Peake, Eishi Noguchi. Fanconi anemia: current insights regarding epidemiology, cancer, and DNA repair. Hum Genet. 2022;141(12):1811-1836.
23. Kojima, Y. et al. FAM111A protects replication forks from protein obstacles via its trypsin-like domain. Nat. Commun. 2020;11:1-14.
24. Svoboda, M., Konvalinka, J., Trempe, J. F. & Grantz Saskova, K. The yeast proteases Ddi1 and Wss1 are both involved in the DNA replication stress response. DNA Repair. 2019;80:45-51.
25. Lessel D, Vaz B, Halder S, Lockhart PJ, Marinovic-Terzic I, Lopez-Mosqueda J, et al. Mutations in SPRTN cause early onset hepatocellular carcinoma, genomic instability and progeroid features. Nat Genet. 2014;46(11):1239-44.
26. Maskey, R., Kim, M., Baker, D. et al. Spartan deficiency causes genomic instability and progeroid phenotypes. Nat Commun. 2014;5:5744.
27. Mosbech A, Gibbs-Seymour I, Kagias K, Thorslund T, Beli P, et al. DVC1 (C1orf124) is a DNA damage - targeting p97 adaptor that promotes ubiquitin-dependent responses to replication blocks. Nat. Struct. Mol. Biol. 2012;19:1084-92.
28. Li F, Raczynska JE, Chen Z, Yu H. Structural Insight into DNA-Dependent Activation of Human Metalloprotease Spartan. Cell Rep. 2019;26(12):3336-46.e4.
29. Reinking HK, Kang HS, Gotz MJ, Li HY, Kieser A, et al. DNA structure-specific cleavage of DNA-protein crosslinks by the SPRTN protease. Mol. Cell. 2020;80:102-13.e6.
30. Stingele J, Bellelli R, Alte F, Hewitt G, Sarek G, Maslen SL, et al. Mechanism and Regulation of DNA-Protein Crosslink Repair by the DNA-Dependent Metalloprotease SPRTN. Mol Cell. 2016;64(4):688-703.
31. Vaz B, Popovic M, Newman JA, Fielden J, Aitkenhead H, Halder S, et al. Metalloprotease SPRTN/DVC1 Orchestrates Replication-Coupled DNA-Protein Crosslink Repair. Mol Cell. 2016;64(4):704-19.
32. Lopez-Mosqueda J, Maddi K, Prgomet S, Kalayil S, Marinovic-Terzic I, Terzic J, Dikic I. SPRTN is a mammalian DNA-binding metalloprotease that resolves DNA-protein crosslinks. Elife. 2016;5:e21491.
33. Fielden J, Ruggiano A, Popović M, Ramadan K. DNA protein crosslink proteolysis repair: From yeast to premature ageing and cancer in humans. DNA Repair (Amst). 2018;71:198-204.
34. Zhao S, Kieser A, Li HY, Reinking HK,Weickert P, et al. A ubiquitin switch controls autocatalytic inactivation of the DNA–protein crosslink repair protease SPRTN.

Nucleic Acids Res. 2021;49:902–15.
35. Simpson, et al. Structure of the chromatosome, achromatinparticle containing 160 base pairs of DNA and all the histones. Biochemistry. 1978;17:5524–5531.
36. K. Luger, A.W. Mader, R.K. Richmond, D.F. Sargent, T.J. Richmond, Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 1997;389:251–260.
37. T.J. Richmond, J.T. Finch, B. Rushton, D. Rhodes, A. Klug, Structure of the nucleosome core particle at 7 A resolution. Nature. 1984;311:532– 537.
38. G. Arents, R.W. Burlingame, B.C. Wang, W.E. Love, E.N. Moudrianakis, The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. Proc. Natl. Acad. Sci. U.S.A. 1991;88:10148-10152.
39. Chen R, Kang R, Fan XG, Tang D. Release and activity of histone in diseases. Cell Death Dis. 2014;5(8):e1370.
40. Ivic N, Bilokapic S, Halic M. Preparative two-step purification of recombinant H1.0 linker histone and its domains. PLOS One. 2017;12(12):e0189040.
41. Nicole Happel, Detlef Doenecke. Histone H1 and its isoforms: Contribution to chromatin structure and function. Gene. 2009;431(1-2):1-12.
42. Bernier, M., Luo, Y., Nwokelo, K.C., Goodwin, M., Dreher, S.J., Zhang, P., Parthun, M.R., Fondufe-Mittendorf, Y., Ottesen, J.J., Poirier, M.G. Linker histone H1 and H3K56 acetylation are antagonistic regulators of nucleosome dynamics. Nat. Commun. 2015;6:10152.
43. Zlatanova J, Caiafa P, Van Holde K. Linker histone binding and displacement: versatile mechanism for transcriptional regulation. Faseb j. 2000;14(12):1697-704.
44. Chapman GE, Hartman PG, Bradbury EM. Studies on the role and mode of operation of the very-lysine-rich histone H1 in eukaryote chromatin. The isolation of the globular and nonglobular regions of the histone H1 molecule. Eur J Biochem. 1976;61(1):69-75.
45. P. Vyas, D.T. Brown N- and C-terminal domains determine differential nucleosomal binding geometry and affinity of linker histone isotypes H1(0) and H1c J. Biol. Chem. 2012;287:11778-11787.
46. Zhou B-R, Feng H, Ghirlando R, Li S, Schwieters CD, Bai Y. A small number of residues can determine if linker histones are bound on or off dyad in the chromatosome. J Mol Biol. 2016; 428: 3948-3959.
47. Bradbury, E.M.; Cary, P.D.; Chapman, G.E.; Crane-Robinson, C.; Danby, S.E.; Rattle, H.W.; Boublik, M.; Palau, J.; Aviles, F.J. Studies on the role and mode of operation of the very-lysine-rich histone H1 (F1) in eukaryote chromatin. The conformation of histone H1. Eur. J. Biochem. 1975;52:605-613.
48. Lu, X.; Hansen, J.C. Identification of specific functional subdomains within the linker histone H10 C-terminal domain. J. Biol. Chem. 2004;279:8701-8707.
49. Hansen, J., M. Sato, R. Ruedy, K. Lo, D.W. Lea, and M. Medina-Elizade. Global temperature change. Proc. Natl. Acad. Sci. 2006;103:14288-14293.
50. Vladimir N. Uversky. Unusual biophysics of intrinsically disordered proteins. chim Biophys Acta. 2013;1834(5):932-51.
51. Stingele, J., Schwarz, M. S., Bloemeke, N., Wolf, P. G. & Jentsch, S. A DNA

dependent protease involved in DNA-protein crosslink repair. Cell. 2024;158: 327-338.
52. Tanya R Porras-Yakushi, Sonja Hess. Recent advances in defining the ubiquitylome. Expert Rev Proteomics. 2014;11(4):477-90.
53. Damgaard, R.B. The ubiquitin system: from cell signalling to disease biology and new therapeutic opportunities. Cell Death Differ. 2021;28:423-426.
54. Pohl, C. & Dikic, I. Cellular quality control by the ubiquitin-proteasome system and autophagy. Science. 2019;366:818-822.
55. Doris Popovic, Domagoj Vucic, Ivan Dikic. Ubiquitination in disease pathogenesis and treatment. Nat Med. 2014;20(11):1242-53.
56. Goldstein G, Scheid M, Hammerling U, et al. Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. Proc Natl Acad Sci USA. 1975;72:11–15.
57. Akutsu, M., Dikic, I. & Bremm, A. Ubiquitin chain diversity at a glance. J. Cell Sci. 2016;129:875-880.
58. Komander D, Rape M. The ubiquitin code. Annu Rev Biochem. 2012;81:203-29.
59. Swatek, K.N., Komander,D.. Ubiquitin modifications. Cell Res. 2016;26:399-422.
60. Kim W, Bennett EJ, Huttlin EL, et al. Systematic and quantitative assessment of the ubiquitin-modified proteome. Mol Cell. 2011;44:325-340.
61. Rieser E, Cordier S M, Walczak H. Linear ubiquitination: a newly discovered regulator of cell signalling. Trends Biochem Sci. 2013;38:94-102.
62. Chau, V., Tobias, J.W., Bachmair, A., Marriott, D., Ecker, D.J., Gonda, D.K., Varshavsky, A. A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. Science. 1989;243:1576-1583.
63. Zhou, H., Wertz, I., O'Rourke, K., Ultsch, M., Seshagiri, S., Eby, M., Xiao, W., Dixit, V.M. Bcl10 activates the NF-kappaB pathway through ubiquitination of NEMO. Nature. 2004;427:167-171.
64. Miranda, M., Sorkin, A. Regulation of receptors and transporters by ubiquitination: New insights into surprisingly similar mechanisms. Mol. Interv. 2007;7: 157-167.
65. Deshaies RJ, Joazeiro CAP. RING domain E3 ubiquitin ligases. Annu Rev Biochem. 2009;78:399-434.
66. Buetow L, Huang DT. Structural insights into the catalysis and regulation of E3 ubiquitin ligases. Nat. Rev. Mol. Cell Biol. 2016;17:626-642.
67. Heideker J, Wertz IE. DUBs, the regulation of cell identity and disease. Biochem. J. 2015;467:191.
68. Sun, Y., Chen, J., Huang, Sy.N. et al. PARylation prevents the proteasomal degradation of topoisomerase I DNA-protein crosslinks and induces their deubiquitylation. Nat Commun. 2021;12:5010.
69. Fábián, Z., Kakulidis, E.S., Hendriks, I.A. et al. PARP1-dependent DNA-protein crosslink repair. Nat Commun. 2024;15:6641.
70. ANNAMARIA RUGGIANO, BRUNO VAZ , SUSAN KILGAS, et al.. The protease SPRTN and SUMOylation coordinate DNA-protein crosslink repair to prevent genome instability. Cell Rep. 2021;37:110080.

71. LIU et al.,Mechanism and function of DNA replication-independent DNA-protein crosslink repair via the SUMO-RNF4 pathway. EMBO. 2021;40:e107413.
72. Larsen, N. B. et al. Replication-coupled DNA-protein crosslink repair by SPRTN and the proteasome in Xenopus egg extracts. Mol. Cell. 2019;73:574-588.e7.
73. Sparks, J. L. et al. The CMG helicase bypasses DNA-protein crosslinks to facilitate their repair. Cell. 2019;176:167-181.e21.
74. Gallina, I. et al. The ubiquitin ligase RFWD3 is required for translesion DNA synthesis. Mol. Cell. 2021;81:442-458.e9.
75. Duxin, J.P., Dewar, J.M., Yardimci, et al.. Repair of a DNA-protein crosslink by replication-coupled proteolysis. Cell. 2014;159:346-357.
76. Alexander Kröning, Johannes van den Boom, et al. Ubiquitin-directed AAA+ ATPase p97/VCP unfolds stable proteins crosslinked to DNA for proteolysis by SPRTN. J Biol Chem. 2022;298(6):101976.
77. Mosbech A, Gibbs-Seymour I, Kagias K, Thorslund T, Beli P, et al. DVC1 (C1orf124) is a DNA damage-targeting p97 adaptor that promotes ubiquitin-dependent responses to replication blocks. Nat. Struct. Mol. Biol. 2012;19:1084-92
78. Davis EJ, Lachaud C, Appleton P, Macartney TJ, Nathke I, Rouse J. DVC1 (C1orf124) recruits the p97 protein segregase to sites of DNA damage. Nat. Struct. Mol. Biol. 2012;19:1093-100.
79. Rune Busk Damgaard. The ubiquitin system: from cell signalling to disease biology and new therapeutic opportunities. Cell Death Differ. 2021;28:423-426.
80. Perry, M., Biegert, M., Kollala, S. S., Mallard, H., Su, G., Kodavati, M., et al. USP11 Mediates Repair of DNA-Protein Cross-Links by Deubiquitinating SPRTN Metalloprotease. J. Biol. Chem. 2021;296:100396.
81. Perry M, Biegert M, Kollala SS, Mallard H, Su G, et al. USP11 mediates repair of DNA-protein crosslinks by deubiquitinating SPRTN metalloprotease. J. Biol. Chem. 2021;296:100396.
82. Huang J, Zhou Q, Gao M, Nowsheen S, Zhao F, et al. Tandem deubiquitination and acetylation of SPRTN promotes DNA-protein crosslink repair and protects against aging. Mol. Cell. 2020;79:824-35.e5.
83. Hochstrasser, M. Origin and function of ubiquitin-like proteins. Nature. 2009;458:422-429.
84. Kerscher, O.; Felberbaum, R.; Hochstrasser, M. Modification of proteins by ubiquitin and ubiquitin-like proteins. Annu. Rev. Cell Dev. Biol. 2006;22:159-180.
85. Serrano, I. Structural insights into the specific recognition and transfer of ISG15. Nat Struct Mol Biol. 2024;31:214.

Appendix Ⅰ: Information of Plasmids and Proteins

Plasmid	Plasmid Source	Protein name	Protein Source
pGEX6P1-hUb	Wei Song	Mono-Ub	Wei Song
pET-24b(+)-C-Cys-His-Ub2(G76C)	Wei Song	M1-Ub_2	Wei Song
pET24-tetraUb	Wei Song	M1-Ub_4	Wei Song
pNIC28-ISG15	Wei Song	ISG15	Xiaosheng Zhu
pNIC28_DVC1 26-240	Wei Song	SprT	Wei Song
pCold I_mH1-Ncys	Wei Song	Cy5-H1-Ncys	Wei Song
pCold I_Ub4-mH1-Ncys	Wei Song	Cy5-Ub_4-H1	Wei Song
pCold I_2xISG15-C78S-mH1-Ncys	Wei Song	Cy5-2×ISG15-H1	Xiaosheng Zhu
pCold I_Ub-mH1-Ncys	Wei Song	Cy5-Ub-H1	Wei Song