Electronic Thesis and Dissertation Repository

6-22-2017 12:00 AM

Investigation of the Interactions between Elg1 and PCNA Involved in the DNA Damage Response

Jingwei Yan, *The University of Western Ontario*

Supervisor: Dr. Hong Ling, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Biochemistry
© Jingwei Yan 2017

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Biochemistry Commons

Recommended Citation

Abstract

Elg1, the major component of an alternative PCNA loader, is vital in maintaining chromosomal stability. PCNA is an essential coordinator for DNA replication and the damage response through protein-protein interactions. The N-terminus of Elg1 interacts with PCNA and preferentially SUMOylated PCNA (sPCNA) and contains five putative motifs for PCNA and SUMO binding. How these motifs contribute to PCNA interactions is unclear. In this study, we biochemically characterized the interactions between Elg1 and PCNA and found that i) two PCNA-interacting protein boxes (PIP boxes) participate in the interaction with PCNA, which may be interfered by the SUMO-interacting motifs (SIMs); ii) the Elg1 N-terminus binds to sPCNA with micromolar affinity, and mutations on SIMs have subtle effects on the binding affinity; iii) all three SIMs interact with SUMO. Given these findings, we propose that Elg1's interaction with PCNA is promoted by SIMs upon SUMOylation and inhibited by SIMs when lacks PCNA SUMOylation.

Dedicated to my family, friends, and lab mates for their tireless support and encouragement.

Acknowledgements

First and foremost, I would like to thank Dr. Hong Ling for allowing me an opportunity to pursue the Master's degree in Biochemistry. The experience has been exceptional and rewarding. Dr. Hong Ling's dedication in guiding me through the research and her support for the writing of the thesis was admirable in every sense of the word. Besides Dr. Hong Ling, I would also like to send my sincere gratitude to my advisory committee: Dr. Gary Shaw, for providing help on NMR experiments and perspectives on the project, and Dr. Chris Brandl, for his insightful advice and perspectives. It was a great enjoyment to work with them. Without these professors' encouragement and support, it would not have been possible to conduct this research.

Secondly, I would like to thank Dr. Guangxin Xing for generating most of the plasmids used in this study and to him and Dr. Vikash Jha in our lab for their aid in troubleshooting and advice. They have helped me overcome many obstacles, and it has been a pleasure to learn from them. To the members of the Lab, Dr. Chuanbing Bian, Lizhen Guo, Andy Liu, Rebecca Earnshaw, Sam Chu, Derek Tse, Harrison Taylor, Rashiq Shahad, and all the other undergraduate students. It was a great time to be working with you all and thank you for making my time here a unique one. My special thanks to Tam Bui, who trained me when I first started and has taught me so much. I would also like to thank Dr. Lynn Weir for her help on the thesis writing and Lee-Ann Briere for her training and kind advice.

Finally, to my parents, I am thankful to have your love and support. To my friends and loved ones, thank you for being there whenever I needed you and for always putting a smile on my face.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations and Acronyms

Amino Acids

Ala (A)	alanine
Arg (R)	arginine
Asn (N)	asparagine
Asp (D)	aspartic acid
Cys (C)	cysteine
Gln (Q)	glutamine
Glu (E)	glutamate
Gly (G)	glycine
His (H)	histidine
Ile (I)	isoleucine
Leu (L)	leucine
Lys (K)	lysine
Met (M)	methionine
Phe (F)	phenylalanine
Pro (P)	proline
Ser (S)	serine
Thr (T)	threonine
Trp (W)	tryptophan
Tyr (Y)	tyrosine
Val (V)	valine

°C	degrees Celsius
µl	microliter
µM	micromolar
Å	angstrom
AAA+	ATPases associated with diverse cellular activities
ATAD5	ATPase family, AAA domain-containing protein 5
ATP	adenosine triphosphate
$CaCl_2$	calcium chloride
D_2O	deuterium oxide, water composed of deuterium
DNA	deoxyribonucleic acid
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
EDTA	ethylenediaminetetraacetic acid
Elg1	enhanced level of genome instability
FT	flow through
GST	glutathione S-transferase
hPCNA	human PCNA
IDCL	interdomain connecting loop
IPTG	isopropyl β-D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry

K_d	dissociation constant
kDa	kilodalton
LB	lysogeny broth
M	molar
MBP	maltose-binding protein
$MgCl_2$	magnesium chloride
mM	millimolar
MMS	methyl methanesulfonate
Mocr	modified ocr tag
NaCl	sodium chloride
NH_4Cl	ammonium chloride
NLS	nuclear localization signal
nM	nanomolar
NMR	nuclear magnetic resonance
NP-40	nonyl phenoxypolyethoxylethanol
PCNA	proliferating c6fell nuclear antigen
PCR	polymerase chain reaction
PEG	polyethylene glycol
pH	potential of hydrogen
PIP box	PCNA-interacting motif
PIPE	polymerase incomplete primer extension
PMSF	phenylmethylsulfonyl fluoride
Pol η	polymerase η
RFC	replication factor C
RLC	replication factor C-like complex
rpm	revolutions per minute
S. cerevisiae	*Saccharomyces cerevisiae*
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SIM	SUMO-interacting motif
Smt3	suppressor of mif two 3
sPCNA	SUMOylated PCNA
SUMO	small ubiquitin-like modifier
TCEP	tris (2-carboxyethyl) phosphine
β-me	β-mercaptoethanol

1 Introduction

Cells are constantly subjected to intracellular and/or extracellular DNA damage factors. In the cell, DNA replication and repair are regulated by a complex network that involves multitudes of protein-protein interactions. Enhanced level of genome integrity protein (Elg1) was first found as a contributor to chromosomal stability in yeast [1]. Elg1 has been shown to form a replication factor complex (RFC)-like complex (Elg1-RLC) and functions in DNA replication. *Elg1* mutants display defects not only in DNA replication but also in S phase progression with the presence of methyl methanesulfonate (MMS), suggesting that Elg1 has a role in DNA damage response [1-2]. Studies have shown that Elg1-RLC mediates the process of proliferating cell nuclear antigen (PCNA) loading/unloading via direct protein-protein interactions [3-5].

PCNA lies at the center of DNA replication and orchestrates DNA repair pathways to handle various errors made in the genome via interactions with PCNA's numerous partner proteins [6-8]. As the processivity factor for polymerases, PCNA needs to be loaded onto the DNA during DNA synthesis [9-11]. On the other hand, PCNA must also be timely unloaded from DNA after DNA synthesis, which is mediated by Elg1-RLC in yeast [12-13]. Because of its roles in fundamental cellular functions, including genome replication, repair, and maintenance, PCNA is recognized as a potential anticancer target. Compounds targeting PCNA have been found to inhibit tumor cell growth [14-15].

Defects in PCNA unloading activity mediated by Elg1 result in various types of chromosome instability, such as DNA damage sensitivity, replication defects, and enhanced homologous recombination [5, 16]. Although Elg1 has been shown to have a role in unloading the DNA-bound PCNA from the template, the mechanism of how Elg1 mediates this process remains unclear. Investigations on interactions between Elg1 and PCNA will shed some light on the mechanism of PCNA unloading and its role in the DNA damage response, which would further extend the understanding of the complex network around PCNA and help future PCNA-targeted drug design.

1.1 Elg1

1.1.1 Clamp loader and Elg1-RLC

In order to achieve high-speed replication, DNA polymerases must attach to a sliding clamp (PCNA is the major one in eukaryotes) [17]. A clamp loader loads the sliding clamp onto DNA and timely unloads it from DNA during replication. Clamp loaders all belong to the AAA+ (ATPases associated with various cellular activities) family of ATPases [18] and interact with the sliding clamp during DNA replication and damage responses. Several protein complexes are responsible for loading/unloading the sliding clamp. RFC is the major loader for the loading and unloading of PCNA in yeast [19].

In addition to RFC, other non-canonical replication factor C-like complexes (RLCs) exist that are conserved in eukaryotes, such as Ctf18-RLC and Rad24-RLC [20-22]. Elg1 is also found to form an alternative RLC in yeast and is composed of five subunits - four small subunits (Rfc 2, 3, 4, and 5) and one large subunit Elg1, instead of Rfc1 in RFC. All these subunits contain a Walker motif that can hydrolyze ATP to drive the loading/unloading process [1].

1.1.2 Elg1 protein

Elg1 was first identified as a contributor to chromosomal stability in *S. cerevisiae* [5]. Elg1 shares limited sequence homology with its human ortholog ATPase family, AAA domain-containing protein 5 (ATAD5). Elg1 contains several conserved cores of AAA+ family (Figure 1A), and Elg1 can form RLCs that interact with PCNA, similar to Rfc1 [1, 23]. Although the N-terminus of Elg1 is predicted to be disordered, several functional motifs are found in the N-terminus, including a nuclear localization signal (NLS), three SUMO-interacting motifs (SIMs) (small ubiquitin-related modifier, SUMO), and a proposed PCNA-interacting motif (PIP box) (Figure 1B) [2, 24]. Elg1's central domain contains a non-canonical Walker motif that does not utilize ATP effectively, suggesting that the unloading activity is performed through the energy provided by the other four subunits [24]. Elg1's C-terminus is established to be significant in the formation of the RLC complex.

Yeast strains that lack amino acids 521–731 of Elg1 are not able to form RLC. Elg1-RLC assembly is likely to resemble the canonical RFC complex formation, where deletion of the C-terminus of any of the subunits (Rfc2, 3, 4, or 5, Elg1) abrogates complex formation [24].

Elg1 has been shown to bind PCNA and, preferentially, SUMOylated PCNA. A sequence of a PIP box has been identified in Elg1, and all three SIMs are able to bind to Smt3 (SUMO in yeast) alone *in vivo* [2]. However, little is known about how those SIMs promote Elg1's interaction with SUMOylated PCNA (sPCNA). Nonetheless, the interaction between Elg1 and PCNA/sPCNA is critical for normal cell growth. The nature of the interaction still needs to be further characterized for a better understanding of the role of Elg1.

1.1.3 SIM and SUMO

Elg1 contains three SIMs in its N-terminus, and their roles in the interaction with sPCNA are not well understood. SIM is a short stretch of residues that interacts with a specific surface groove of SUMO. Most of the SIMs that have been identified contain two important elements: a core of 3-4 hydrophobic residues (usually V or I) and a nearby acidic region, which contains a stretch of E or D or phosphorylated S or T [25]. In the case of Srs2, the yeast DNA helicase, the PIP box and SIM in its C-terminus can bind to PCNA and SUMO independently. The presence of both SIM and the PIP box in Srs2 enhances its interaction with sPCNA at least 7-fold compared to that with PCNA alone [26].

SUMO proteins are the ubiquitin-like modifiers that conjugate with certain substrates and regulate cellular processes, such as gene expression, DNA repair, chromosome assembly, and cellular signaling [27]. While yeast has only one SUMO, four SUMO isoforms have been identified in humans [28]. Even though the functions of SUMO can be distinctive, their main role is to regulate the interactions of the modified protein with other proteins. Most proteins that contain SIM would non-covalently bind to SUMO or the SUMO that is conjugated to the substrate.

It has been revealed that Elg1-RLC preferentially unloads sPCNA, as Elg1-RLC preferenti-

A

N terminal   Central domain   C terminal

NLS PIP box

Elg 1   II   V VI   VIII   791 a.a.

SIM1, 2, 3

B

MKRHVSLSDILTGNKRKVRRQDALQITIDDENDTESGTFDARTAKHDDSS

1 ——————————— NLS —— SIM1 ——————————— 50

VIFLNHSVVKPIEAVSTNHKSAKEFLMTKRTKEKCDDDDDDLIVISDKSP

51 —— PIP1 —————— PIP2 ——————— SIM2 —— 100

KSETNCSKIALSQEHEDDISIISTSRIKSSLLNERASKIKNFLKHETTDT

101 ——————— SIM3 ——————————————————— 150
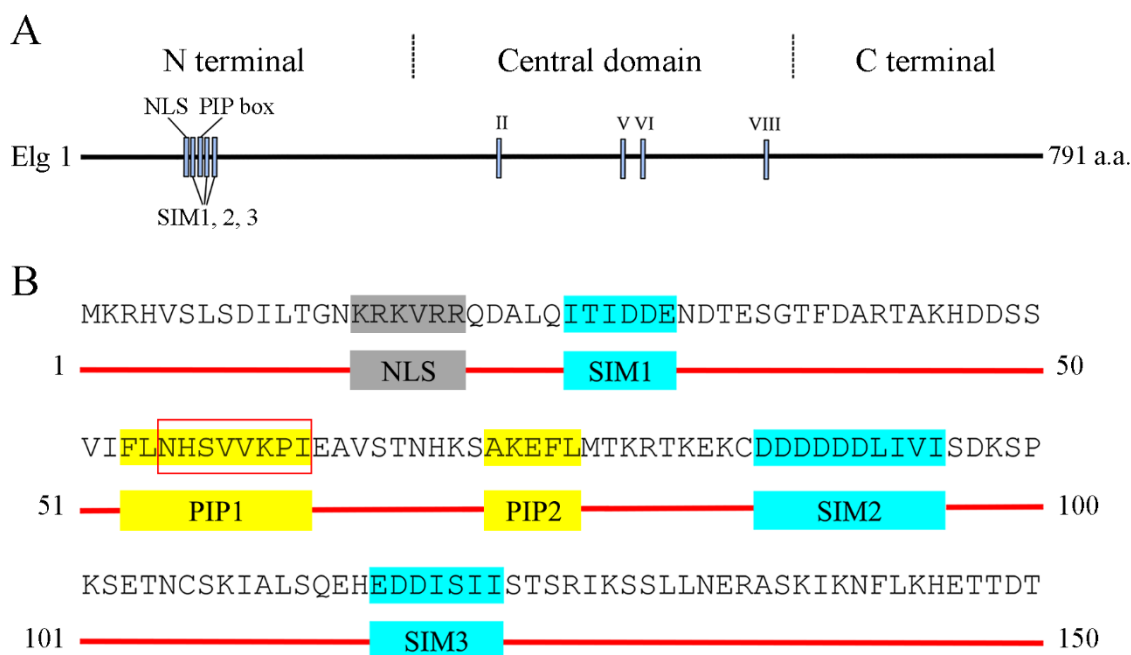
Figure 1: Schematic diagram of Elg1 functional motifs.

A) The conserved sequence in *S.cerevisiae* is shown in blocks. All square blocks that marked with roman numerals in the central domain represent the conserved cores for the AAA+ family. (Adapted from Bellaoui *et al*. [1]). These square blocks in the Elg1 N-terminus represent the functional motifs that have been identified;

B) The amino acid sequence of Elg1 1-150 is shown, along with a bar diagram that shows the functional motifs involved in the interactions. The NLS, PIP boxes (described later in preliminary research), and SIMs are marked in grey, yellow, and light blue, respectively. The red square outlines the proposed PIP box in Parnas *et al.* [2].

-ally interacts with sPCNA, and yeast strains that lack Elg1 show a significant accumulation of sPCNA on chromatin [2]. However, the mechanism of how the SIMs in Elg1's N-terminus work coordinately with the PIP box remains unclear. Elg1 may interact with SUMOylated PCNA in a similar manner like Srs2 to enhance binding affinity to PCNA, where the interaction with PCNA and the interaction with SUMO are independent of each other.

1.2 PCNA

PCNA is a homo-trimeric ring-shaped protein that coordinates with various proteins involved in DNA metabolism. It serves as a platform that facilitates the recruitment of proteins involved in DNA replication and DNA damage response pathways. During replication, PCNA interacts with DNA polymerase ε to facilitate leading-strand replication, while it works with DNA polymerase δ to synthesize DNA discontinuously in a series of Okazaki fragments on the lagging strand [29]. PCNA needs to be loaded repeatedly onto the DNA when each Okazaki fragment is synthesized [30]. Failure to unload PCNA can cause abnormalities in DNA replication and chromosome malformation.

PCNA also functions in many other processes, such as translesion synthesis, error-free damage bypass, homologous recombination, mismatch repair, and chromatin assembly [31-35]. Upon DNA damage, the major metabolic pathways to repair DNA include nucleotide excision repair, base excision repair, double-strand break repair, and mismatch repair, in which PCNA participates with its partners to coordinate the processes [36]. The diversified functions of PCNA are regulated by its interactions with different types of proteins.

1.2.1 Yeast PCNA

Recognized as the major sliding clamp, PCNA in yeast is encoded by the *POL30* gene, which codes for a protein of molecular weight of approximately 29 kDa. Yeast PCNA shows 35% homology in sequence with human PCNA [37]. Despite the poor similarity in the primary sequences of PCNAs among bacteria, archaea, and eukaryotes, the overall

structure of PCNAs is highly conserved. The first crystal structure of yeast PCNA was determined in 1994 [38]. The structure showed that PCNA exists as a homotrimer and that each monomer is composed of three repeated domains (Figure 2). The two domains, the N-terminus and the C-terminus, of each monomer are topologically identical and are connected to each other via the interdomain connecting loop (IDCL) (Figure 2). The three monomers of PCNA are arranged in a head-to-tail orientation, which gives them a closed ring-shaped structure.

The diameter of the center cavity of yeast PCNA is approximately 35 Å (Figure 2), which is sufficiently large to allow the double helix of DNA to pass through. The PCNA inner surface comprises 12 α-helices that are positively charged due to lysine and arginine residues, which allow the smooth passage of negatively charged DNA through the ring [39]. The PCNA outer surface comprises a set of β sheets, which stabilizes the interactions among the monomeric PCNAs. The head-to-tail arrangement of the PCNA monomers gives rise to two distinct faces, the front, and the back. The front face contains the IDCL, which bridges the amino-terminal domain and the carboxyl-terminal domain. The IDCL and the hydrophobic region on PCNA are responsible for making the contacts with binding partners. IDCL also contains sites for modification, which may alter its interaction with other partners and trigger a switching of functional pathways [40-41]. Several proteins have been shown to bind PCNA through interactions with IDCL, such as Fen1, Cdc9, and Srs2 [26, 42-43]. The role of PCNA's back face remains unclear. However, it has emerged as a site of PCNA post-translational modification [39].

1.2.2 PCNA-interacting protein motif (PIP Box)

Proteins that PCNA recruits share a sequence-consensus motif called the PIP-box. The sequence alignment of several PCNA-binding proteins reveals that this motif comprises the Q-X-X-(L/M/I/V)-X-X-(F/Y)-(F/Y) sequence (Figure 3) [36], where X represents any amino acid. The PIP motif forms a 3_{10} helix that acts as a hydrophobic anchor and facilitates binding to the surface of PCNA [43]. After the structure of the first PCNA in complex with a peptide (PCNA-p21) was determined [44], more PCNA-peptide complex structures were
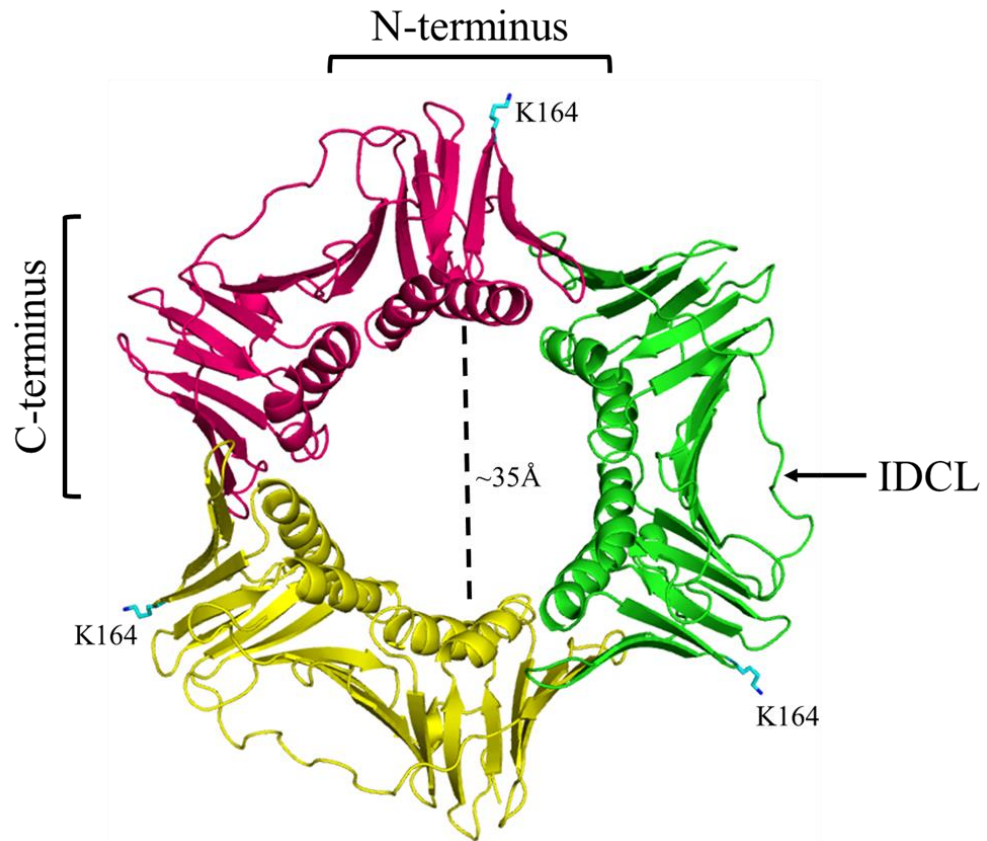
Figure 2: Crystal structure of yeast PCNA.
Figure constructed using 1PLQ structure from Protein Data Bank (PDB) shows the front view of PCNA. Each monomer is colored differently, and 3 monomers form the trimer in a head-to-tail manner. Also shown is that the center ring with a diameter of approximately 35Å (measured via Pymol). The position of one monomer's IDCL is indicated with an arrow at the front face of the PCNA trimer. The K164s (SUMOylation sites) are also indicated.

Canonical sequence:

Q–X–X–(L/M/I/V)–X–X–(F/Y)–(F/Y)

Canonical

```
Msh6:  MKQSSLLSFFSK
Pol2:  FKQTSLTKFFSK
Cdc9:  PKQATLESFFKR
Rrm3:  YRQQTLSSFFMG
Rad27: GIQGRLDGFFQV
Ung1:  RKQTTIEDFFGT
```

Non-canonical

```
Srs2:  SSQMDIFSQLSR
Mcm10: DFQHTLDVYIFG
Rad30: TSSKNILSFFTR
```

Figure 3: The sequence alignment of the PIP boxes in yeast proteins.
The alignment of several yeast proteins that have been shown to interact with PCNA is presented in this figure, including canonical and non-canonical ones. The canonical amino acids are highlighted.

subsequently determined, including Cdc9, Srs2, and TRAIP [26, 43, 45]. Moreover, some proteins have been found to contain multiple PIP boxes. For instance, human Pol η contains three PIP boxes that contribute differently to two distinct functions. One is to stimulate DNA synthesis, and another is to promote PCNA mono-ubiquitination [46]. Additionally, human Pol η not only possesses three PIP boxes, but also contains a ubiquitin-binding motif that results in its preferential interaction with mono-ubiquitinated PCNA and functions to synthesize DNA in an error-prone manner [46]. Some proteins possess another PCNA binding motif, instead of the PIP box, called the KA box. The consensus sequence of the KA box is K-A-(A/L/I)-(A/L/Q)-X-X-(L/V), where X represents any amino acid [47]. However, no consensus sequences of the PIP box and KA box were found in Elg1.

1.2.3 PCNA SUMOylation

PCNA can be modified in a multitude of ways, such as mono- and poly-ubiquitination, SUMOylation, phosphorylation, and acetylation (Figure 4) [48-52]. The different modifications on PCNA regulate its functions by altering its interactions with numerous binding partners. For example, ubiquitination of PCNA directs error-prone and error-free replication when the cell undergoes replication stresses. The mono-ubiquitination of PCNA facilitates the recruitment of translesion synthesis polymerases at the stalled replication fork and allows the replication to continue [50], while poly-ubiquitination of PCNA is believed to play a role in error-free damage avoidance through template switching [53].

In yeast, the SUMO modification of PCNA occurs mostly at K164 and to a small extent at K127 [48]. PCNA SUMOylation is not required for replication in undisrupted cells and is not a substitute in the absence of ubiquitination [54]. Although extensive research has been conducted to investigate the functions of PCNA SUMOylation, it is still not clearly understood. One major finding is that sPCNA can recruit Srs2 to the sites of replication where Srs2 can disrupt Rad51 filaments and prevent homologous recombination [55]. The C-terminus of Srs2 contains a non-canonical PIP box and a SIM which bind independently with PCNA and Smt3, respectively. Both motifs are required to recognize sPCNA [26]. In
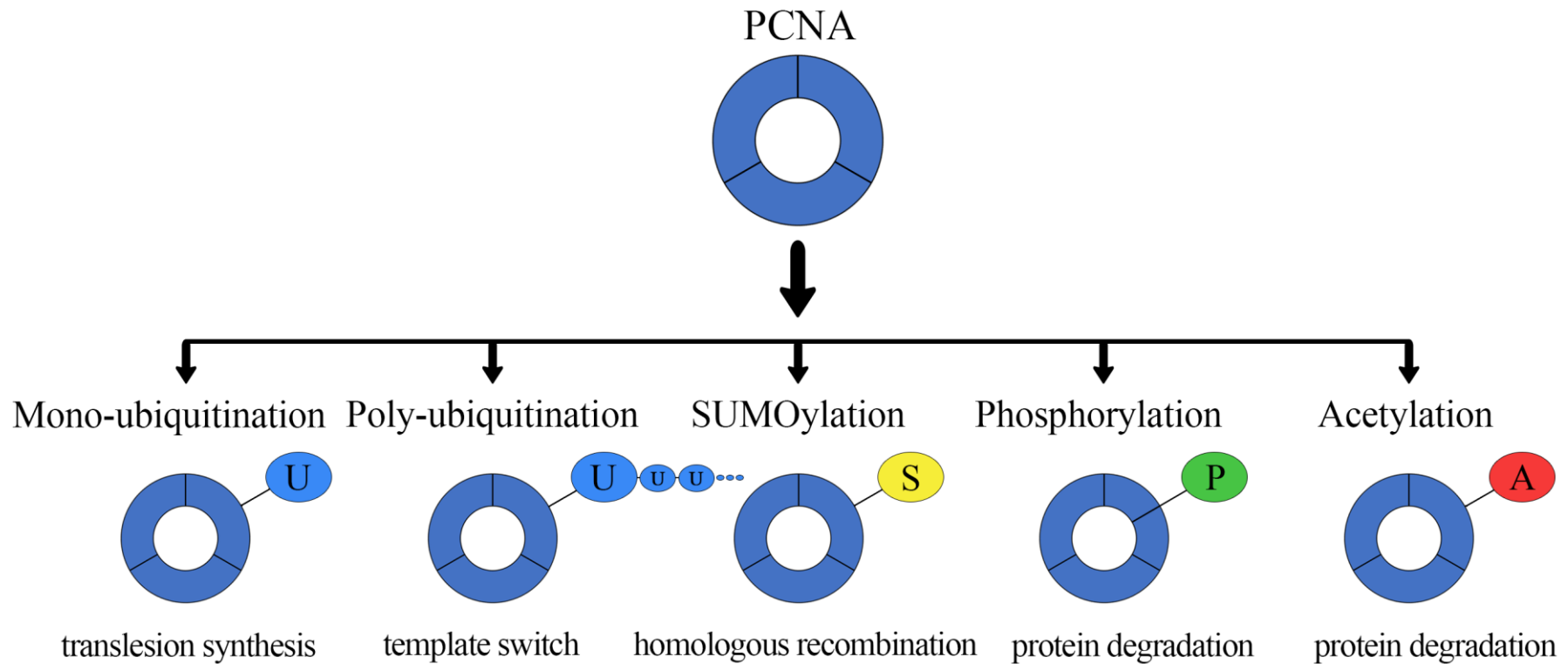
Figure 4: An overview of PCNA modifications and the pathways they direct.
Upon DNA damage, PCNA would undergo different types of modification, which subsequently results in preferred interactions with corresponding proteins and leads to different pathways. The PCNA trimer is marked in blue. Ubiquitin, SUMO, phosphate, and acetyl are represented in light blue, yellow, green, and red, respectively.

addition, it has also been shown that Elg1-RLC, which is the alternative clamp loader, preferentially binds to sPCNA and assists in unloading sPCNA from DNA [1-2]. It was proposed to interact with sPCNA in a similar manner as Srs2.

1.3 Research on Elg1-PCNA interaction

Substantial investigations have been conducted on Elg1-PCNA interactions [1-2, 5, 16]. Several attempts have been made to understand the mechanism of PCNA unloading facilitated by Elg1. It has been proposed that the N-terminus of Elg1 mainly mediates the interaction with PCNA, and SIMs would promote the interaction between Elg1 and PCNA upon SUMOylation. Parnas *et al.* [2] have confirmed the ability of Elg1's N-terminus to bind PCNA and proposed a PIP box from yeast-2-hybrid assays.

On the other hand, pull-down assays in our lab with Elg1L3+ protein (33-84), which contains only the predicted PIP box, but no SIMs, have validated its binding ability with PCNA. It was also observed that, in the Elg1L3+ (33-84) construct, the mutations not only on (F53L54, V58, P61I62), but also outside of the proposed PIP box reduced binding to PCNA. This region contains two hydrophobic amino acids (F75L76) (Figure 5A), which are equally important for the interaction with PCNA. We, therefore, realigned the sequence of the Elg1 N-terminus using Clustal Omega (Figure 5B) and found that the consecutive hydrophobic amino acids on the proposed PIP box and outside of the PIP box are possibly conserved for interaction with PCNA. We proposed that these two potential PIP boxes would both contribute to the interaction with PCNA. The positions of these two PIP boxes are also marked in Figure 1B.

1.4 Rationale, hypothesis, and objectives

The Elg1-PCNA interaction mediates the unloading activity of PCNA, thereby regulating PCNA's performance during DNA replication and repair. Elg1-RLC is well known as the major unloader of PCNA in yeast. Previous research has localized the "PIP box" of Elg1 on its N-terminus [2, 5], and our lab's preliminary investigations find two possible PIP boxes in the Elg1 N-terminus. Following our lab's previous research, both proposed PIP
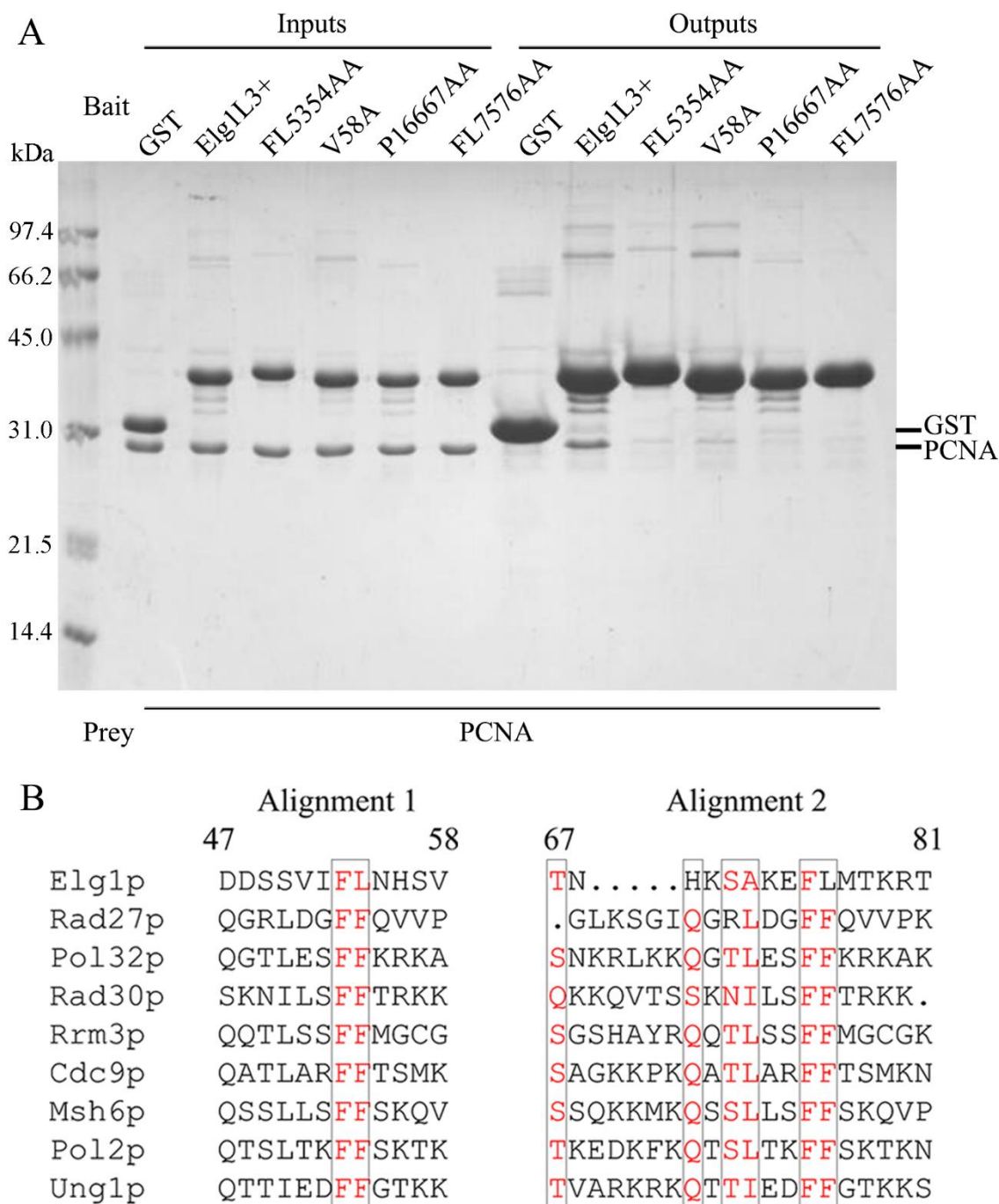
Figure 5: Previous results from our lab regarding Elg1-PCNA interaction.
These results come from Andy Liu's thesis. A) Representative GST pull-down assay on SDS-PAGE analysis. GST, GST-fused Elg1L3+ (33-84) and its mutants (contain point mutations) were used as bait, and PCNA (from yeast) was used as prey in GST pull-down assays. It shows the mutation on F53L54, V58, P61I62, F75L76 are important to the interaction with PCNA. The positions of yeast PCNA and GST are indicated on the right; B) Sequence alignments using Omega Cluster for the canonical amino acids with multiple yeast proteins that contain PIP boxes. The alignment results suggested that Elg1 may contain two potential PIP boxes.

boxes that interact with PCNA were further investigated in this study. Additionally, Elg1 shows preferential binding with sPCNA, based on pull-down assays, and yeast 2-hybrid studies have determined that all three SIMs of Elg1 can interact with Smt3 [2]. However, the mechanism of how SIMs promote Elg1's binding with sPCNA remains unclear, with the possibility that only one or two SIMs promote the interaction upon SUMOylation.

To address these questions, we hypothesized that both proposed PIP boxes interact with PCNA and all three SIMs contribute to the interaction with sPCNA. Two biochemical approaches, GST pull-down assays and isothermal titration calorimetry (ITC), were utilized in this thesis to evaluate the role of two proposed PIP boxes and three SIMs in the interactions between Elg1 and PCNA/sPCNA. In addition, nuclear magnetic resonance (NMR) was employed to better understand the SIMs' binding model with Smt3.

This study has two main aims:

- to identify the PIP box that plays the major role in the interactions with PCNA
- to understand the roles of Elg1 SIM(s) in the interactions with sPCNA

Better characterizations of Elg1-PCNA interaction would elucidate the structural basis of how Elg1 unloads PCNA from DNA and understand the role of this unloading activity in DNA replication and repair.

2 Materials and Methods

2.1 Bacterial strains and plasmids

The pET21b-PCNA K127G was a gift from Dr. Lima. The pMCSG9 (for tagging with MBP) [56] and pMCSG10 [57] (for tagging with GST) were purchased from the PlasmID Repository (Harvard Medical School). The plasmid pMocr was obtained from Michigan University for better protein expression and solubility [58]. The plasmid pCDFduet-1, the bacterial strains *Escherichia coli* BL21(DE3), and BL21(DE3) pRare were commercially purchased from Novagen. The bacterial strain *Escherichia coli* DH5α was purchased from Thermo Fisher Scientific. All plasmids used in this study are listed in the Table 1.

Table 1: List of proteins used in this study.

Name	Purified Protein	Plasmid Name	Description
Elg1N	His×6-GST-Elg1 21-150	pMCSG10-yElg1 21-150	Contains all five (2 PIP boxes, 3 SIMs) functional motifs
Elg1L1	His×6-GST-Elg1 21-102	pMCSG10-yElg1 21-102	SIM3 deleted
Elg1L2	His×6-GST-Elg1 21-84	pMCSG10-yElg1 21-84	SIM2 and SIM3 deleted
Elg1L3+	His×6-GST-Elg1 33-84	pMCSG10-yElg1 33-84	Contains PIP1 and PIP2 only, slightly extended than Elg1L3
Elg1L3	His×6-GST-Elg1 46-84	pMCSG10-yElg1 46-84	Contains PIP1 and PIP2 only
Elg1L4	His×6-GST-Elg1 64-102	pMCSG10-yElg1 64-102	Contains PIP2 and SIM2
Elg1L5	His×6-GST-Elg1 64-84	pMCSG10-yElg1 64-84	Contains PIP2 only
GST-p21	His×6-GST-p21	pMCSG10-p21 PIP	Sequence (SNASAVLQKKITDYFHPKK) of p21 PIP box is coded; used as a positive control for hPCNA interaction.
PCNA	Yeast PCNA K127G	pET21b-PCNA K127G	Contains a mutation at site 127, K to G
Smt3	ySmt3 Δ18K19R	pMCSG9-ySmt3 Δ18K19R	First 18 amino acids were deleted and the 19th amino acid was mutated (K to R)
His-Smt3	His- ySmt3 Δ18K19R		First 18 amino acids were deleted and the 19th amino acid was mutated (K to R); Six histidines were directly linked to the protein

E1	yAos1/yUba2 1-554	pCDFduet-yAos1 pMcor-yUba2 1-554	Used as E1 (activating enzyme) for SUMOylation
E2	yUbc9 K153R	pMCSG9-yUbc9 K153R	Contains a mutation at site 153, K to R; Used as E2 (conjugating enzyme) for SUMOylation
E3	ySiz1	pMocr-ySiz1 166-508	Contains the fragment of ySiz1 from 166 to 508
Elg1 WT	W20-Elg1 21-150	pMocr-yElg1 W20 Elg1 21-150	A tryptophan was added before Elg1 protein for the convenience of purification
SIM1m	W20-Elg1 21-150 SIM1 mutant	pMocr-yElg1 W20 SIM1m	Mutations on SIM1
SIM2m	W20-Elg1 21-150 SIM2 mutant	pMocr-yElg1 W20 SIM2m	Mutations on SIM2
SIM3m	W20-Elg1 21-150 SIM3 mutant	pMocr-yElg1 W20 SIM3m	Mutations on SIM3
SIM12m	W20-Elg1 21-150 SIM1 and SIM2 mutant	pMocr-yElg1 W20 SIM12m	Mutations on SIM1 and SIM2
SIM13m	W20-Elg1 21-150 SIM1 and SIM3 mutant	pMocr-yElg1 W20 SIM13m	Mutations on SIM1 and SIM3
SIM23m	W20-Elg1 21-150 SIM2 and SIM3 mutant	pMocr-yElg1 W20 SIM23m	Mutations on SIM2 and SIM3
SIM123m	W20-Elg1 21-150 SIM1, SIM2 and SIM3 mutant	pMocr-yElg1 W20 SIM123m	Mutations on SIM1, SIM2, and SIM3

2.2 List of buffers

Table 2: List of buffers used in this study.

Buffer Name	Used In	Components
His Buffer A	MBP/GST/Mocr-tagged protein purification	25 mM Tris pH 8.0, 500 mM NaCl, 5 mM β-Me
His Buffer B	GST/Mocr-tagged protein purification	25 mM Tris pH 8.0, 500 mM NaCl, 300 mM Imidazole, 5 mM β-Me
MBP Buffer B	MBP-tagged protein purification	25 mM Tris pH 8.0, 500 mM NaCl, 10 mM maltose, 5 mM β-Me
Dialysis Buffer	Protein purification	25 mM Tris pH 8.0, 200 mM NaCl, 0.5 mM EDTA, 5 mM β-Me
Anion Buffer A	MBP/GST/Mocr-tagged protein purification	25 mM Tris pH 8.0, 5 mM β-Me
Anion Buffer B	MBP/GST/Mocr-tagged protein purification	25 mM Tris pH 8.0, 1 M NaCl, 5 mM β-Me
Cation Buffer A	Protein purification	25 mM Tris pH 8.0, 5 mM β-Me
Cation Buffer B	Protein purification	25 mM Tris pH 8.0, 0.1 M NaCl, 5 mM β-Me
GST pull-down buffer	GST pull-down assay	20 mM Hepes pH 7.5, 100 mM NaCl, 5% glycerol, 1 mM EDTA, 5 mM β-Me
GST pull-down elution buffer	GST pull-down assay	20 mM Hepes pH 7.5, 100 mM NaCl, 10 mM reduced glutathione, 5% glycerol, 1 mM EDTA, 5 mM β-Me
ITC buffer	ITC	20 mM Hepes pH 7.5, 100 mM NaCl, 5 mM β-Me
NMR buffer	NMR	20 mM Hepes pH 7.5, 100 mM NaCl, 10% D_2O

2.3 Cloning of constructs for expression

We first cloned Elg1 N-terminus from 1 to 150 and hoped it could be used as the wild-type Elg1. However, we found this fragment have low expression and solubility. Therefore, we designed and tested more constructs in various lengths. The construct that expressed Elg1 21-150 with an extra W ahead of the protein was suitable for our purposes. This construct was used as the wild-type, and it was altered to generate various Elg1 SIM mutants. These mutants contain different combinations of SIM mutations, and M-PIPE cloning was used to introduce point mutations in the desired sites [59], where the hydrophobic amino acids of SIM(s) were mutated into alanines. The sequences of the primers used to introduce mutations into SIMs are listed in Appendix A.

In brief, each pair of primers requires approximately 15 bp of complementary sequences at the 5' primer end. The 3' primer end must be complementary to the flanks of the mutation site [60]. First, PCR reactions were set up to create a single SIM mutant. The reaction utilized Q5 polymerase to amplify and the extension of 25 cycles was conducted. PCR products were checked on a 1% agarose gel. After we confirmed that the products were of the proper size, DpnI digestion was used to remove the methylated original templates. Digested samples were transformed into DH5α competent cells via the heat shock method. Single clones were picked to grow in a 5 ml LB media, and minipreps were performed to extract the plasmids. The plasmids later were sent for sequencing to confirm the correct mutations. By using correct Elg1 SIM single mutants as templates, mutants containing double SIMs mutations and triple mutations were obtained in similar protocols. All plasmids were sent for sequencing to confirm correct mutations.

To investigate the role of different SIMs within the Elg1 N-terminus, we also generated the constructs for expression and purification of His-Smt3, using the pMocr Smt3 Δ18K19R as the template. Similar protocol and primer design were used as above. The primers used for mutagenesis were designed to remove the Mocr tag and TEV cutting

site. We mutated out the DNA code for the Mocr tag and TEV cutting site. This would directly link the protein with the 6 histidines for the convenience of purification. The primers were synthesized by UWO Oligo Factory (Appendix A).

2.4 Expression and purification

2.4.1 Expression and purification of yeast PCNA

The plasmid for expressing yeast PCNA was transformed into BL21(DE3) pRare competent cells via the heat shock method. All PCNAs in the following descriptions mean yeast PCNA unless stated otherwise. The transformants were later selected with 100 µg/mL ampicillin and 25 µg/mL chloramphenicol. Fresh colonies were transferred to LB media. After the cell culture reached an OD_{600} of 0.7, 1 mM IPTG was to induce the protein expression at 16 °C. Cells were harvested after overnight growth and frozen in -80°C for future usage.

The purification of yeast PCNA includes ammonium sulfate precipitation and two rounds of anion exchange chromatography [61]. Frozen cells were thawed with warm water and suspended in His Buffer A (Table 2), with 1mM protease inhibitor, phenylmethylsulfonyl fluoride (PMSF). Cells were lysed with the homogenizer and ultracentrifuged at 48,380 g for 30 minutes to obtain soluble protein. The soluble fractions of PCNA were then gradually spiked with ammonium sulfate to 45% saturation in an ice water bath. The solution was centrifuged and the supernatant was collected. After centrifugation at 12,000 g for 30 minutes, the new supernatant was separated and spiked to 75% saturation with more ammonium sulfate. The precipitant was dissolved with Dialysis Buffer (Table 2) and put into dialysis overnight. The dialyzed protein was directly loaded onto the Q columns (GE Healthcare). A gradient of Anion Buffer B (Table 2) (20% to 65% in 12 column volumes) was employed, and the protein was eluted at the conductivity of approximately 32 ms/cm. The fractions containing the protein of interest (checked with SDS-PAGE) were pooled and

concentrated with a 10 kDa cut-off concentrator at a speed of 2,000 g. The concentrated protein was flash frozen and stored in -80°C for future assay.

2.4.2 Expression and purification of SUMOylation enzymes and SUMO

a) E1 enzyme (yAos1/yUba2 1-544)

Plasmids expressing yAos1 and yUba2 1-544 were co-transformed into BL21(DE3) pRare competent cells via the heat shock method and selectively grown in 100 µg/ml ampicillin, 25 µg/ml chloramphenicol, and 50 µg/ml streptomycin. Expression was induced at OD_{600} 0.7-0.8 for 16°C overnight, and cells were frozen for storage. The targeted proteins were purified via the same protocol as for wild-type Elg1 purification (described in the later section 2.4.4).

b) E2 enzyme (yUbc9 K153R)

The plasmid to express yUbc9 was transformed using the same heat shock method as yeast PCNA. The cells were induced at OD_{600} 0.7-0.8 for 30°C overnight, then collected and frozen for storage. Expressed protein was purified using a similar protocol as for wild-type Elg1 purification (described in the later section 2.4.4), with a change of using cation exchange chromatography (SP columns from GE Healthcare), instead of anion exchange chromatography.

c) E3 enzyme (ySiz1 166-508)

The plasmid to express ySiz1 was transformed, and the cells were grown in the same conditions as for the E2. Expressed protein was purified by using the same protocol as for wild-type Elg1 purification (described in the later section 2.4.4).

d) SUMO (Smt3 Δ18K19R, His-Smt3 Δ18K19R)

The plasmid to express Smt3 Δ18K19R was transformed, and the cells grew in the same conditions as the E2. Expressed protein was purified with MBP affinity and anion exchange chromatography. After the soluble protein was obtained through the same

procedures as for Elg1, the sample was loaded onto MBP columns (GE Healthcare) and eluted with MBP Buffer B (Table 2). Those fractions containing the protein of interest were pooled, and a molar ratio of 1:40 TEV to protein was used for digestion, while dialysis occurred overnight. The digested and dialyzed sample was run through the MBP columns. The protein would remain in the flow through. The sample was then diluted to approximately 100 mM NaCl for anion exchange chromatography. A gradient of Anion Buffer B (10% to 30% in 12 column volumes) was employed. Fractions were checked by SDS-PAGE analysis. The fractions containing the protein of interest were pooled and concentrated with a 5 kDa cut-off concentrator with a speed of 2,000 g. The concentrated protein was flash frozen and stored in -80°C for future assay.

The plasmids to express His-Smt3 Δ18K19R were transformed and induced in the same conditions as for Smt3 Δ18K19R. Expressed protein was purified with nickel affinity chromatography and anion exchange chromatography as GST-fused Elg1 fragments (described in 2.4.3). The purified protein was concentrated to over 10 mg/ml and flash frozen for future assays.

2.4.3 Expression and purification of GST-fused Elg1 fragments

All plasmids for expressing GST-fused fragments were transformed in the same way as for PCNA. The cells were induced at OD_{600} 0.7-0.8 for 25°C overnight, then collected and frozen for storage. The protein was purified via nickel affinity chromatography and anion exchange chromatography.

Cells were lysed and centrifuged as described above as purification of yeast PCNA. The soluble protein underwent nickel affinity chromatography. Fractions containing the protein of interest were obtained via the same procedures as for Elg1 WT (in 2.4.4). Instead of setting up the protein for dialysis and digestion, the protein was directly diluted to approximately 100 mM NaCl and loaded onto the Q columns, followed by a gradient of Anion Buffer B (10% to 30% in 12 column volumes). The fractions containing the protein of interest (checked with SDS-PAGE) were pooled and

concentrated with a 10 kDa cut-off concentrator with a speed of 2,000 g. The protein was concentrated to over 20 mg/ml and stored with 50% glycerol (final protein concentration would be over 10 mg/ml) at -20°C for pull-down assays.

2.4.4 Expression and purification of Elg1 WT and Elg1 SIM mutants

Procedures to express Elg1 WT and Elg1 SIM mutants were the same as used for PCNA transformation. The transformants were induced with 1 mM IPTG when OD_{600} reached 0.4-0.6. The cells were grown at 37°C for 3 hours and collected for storage. Nickel affinity chromatography and anion exchange chromatography were applied to purify Elg1 WT and Elg1 mutant proteins.

Cells were lysed with the homogenizer, and the soluble protein was collected as described above. The samples were loaded onto the nickel columns (GE Healthcare) and washed with 10 column volumes of washing buffer that contained extra 15 mM imidazole in His Buffer A, followed by elution via His Buffer B (Table 2). Fractions that contained the protein of interest were pooled together, and a ratio of 1:40 TEV to protein was used to remove the tag while the sample underwent dialysis in the Dialysis Buffer (Table 2) overnight. After digestion and dialysis, the samples were run through the nickel columns again, and the protein of interest remained in the flow through. The flow-through samples were diluted to approximately 100 mM NaCl and loaded onto Q columns. The protein was eluted with a gradient of Anion Buffer B (10% to 30% in 12 column volumes). The fractions containing the protein of interest (checked with SDS-PAGE) were pooled and concentrated with a 10 kDa cut-off concentrator with a speed of 2,000 g. The concentrated protein was flash frozen and stored in -80°C for future assays.

2.5 SUMOylation of PCNA

Before setting up the SUMOylation reaction, the concentrations of the proteins were measured via the Bradford method, and the volume of protein needed for this reaction was calculated accordingly. Proteins - 15 µM PCNA, 60 µM Smt3 Δ18K19R, 90 nM

yAos1/yUba2 1-544 (E1), 300 nM Ubc9 K153R (E2), and 2 µM Siz1 166-508 (E3) - were mixed for SUMOylation [26]. The final buffer condition for the SUMOylation was 50 mM pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM TCEP, 5% glycerol, and 0.01% NP-40. A total of 2 mM ATP was added to initiate the reaction, followed by incubation at room temperature for 30 minutes. The reaction solution was later supplemented with excess EDTA to slow down the reaction and with NaCl to approximately 250 mM for loading onto the Q columns. The protein was eluted with a gradient of Anion Buffer B (25% to 65% in 12 column volumes). The fractions containing sufficiently pure sPCNA (checked with SDS-PAGE) were pooled and concentrated with a 30 kDa cut-off concentrator with a speed of 2,000 g. The concentrated protein was flash frozen and stored in -80°C for future experiments.

2.6 GST pull-down assay

GST-pull down assays were employed to investigate the roles of different functional motifs in the interaction with PCNA. Purified GST-tagged Elg1 protein fragments (bait) (200 µl, 1 mg/ml) were immobilized onto 20 µl of glutathione-sepharose beads. Also, GST and GST-p21 of the same concentration were used in this assay as the negative control and positive control (with hPCNA), respectively. After immobilizing the bait proteins, the beads were washed 3 times with 200 µl GST pull-down buffer (Table 2) to remove unbound protein. The prey proteins (Smt3, PCNA, sPCNA and hPCNA) were diluted with binding buffer to a molar concentration that was slightly over that of the bait. The preys were then added to the beads for incubation on ice. After an hour, the beads were washed 3 times with 200 µl binding buffer to remove unbound prey proteins. The bound proteins were eluted with 50 µl of GST pull-down elution buffer (Table 2). Supernatants were collected after centrifugation and resolved by SDS-PAGE.

To explore the role of SIMs, GST-tagged Elg1 protein fragments were used in pull-down assays with Smt3. Similar procedures were used as before, but with a higher protein concentration and higher molar ratio. The bait was prepared to 2 mg/ml, and the ratio of the prey's molar concentration to the concentration of the bait was increased to

3-fold. After using similar procedures as for investigating the interactions with PCNA and sPCNA of incubation, washing, and eluting, the eluted proteins were resolved on SDS-PAGE gels.

2.7 Isothermal titration calorimetry

ITC experiments were performed with proteins that were exchanged into the ITC buffer (Table 2) via size exclusion chromatography. ITC measurements were performed on the Nano ITC calorimeter (TA Instruments) at 25°C with a stirring speed of 250 rpm.

To start a test run for investigating the interaction between Elg1 and PCNA, 830 µM PCNA were titrated into 79 µM GST Elg1L3+. The titration was conducted in 34 x 1.5 µl injections, following a 0.5 µl initial injection. Control experiments were conducted by titrating sPCNA into the ITC buffer under the same experimental settings. Experimental data were analyzed by Nanoanalyze by subtracting the signals from control experiments and fitting into the independent binding model.

To investigate the role of SIMs in the interaction with sPCNA, 406 µM sPCNA were titrated into Elg1 WT and various Elg1 SIM mutants (60 µM) in 34 x 1.5 µl injections, following a 0.5 µl initial injection. Control experiments were conducted by titrating sPCNA into the ITC buffer under the same experimental settings. The same fitting strategy was performed to fit the data.

To investigate how mutations on SIMs affect the interaction with Smt3, Elg1 WT (0.72 mM) and Elg1 single mutants (0.56 mM Elg1 SIM1m, 0.78 mM Elg1 SIM2m, and 0.68 mM Elg1 SIM3m) were titrated into 0.1 mM His-Smt3. All titration experiments were performed with 26 x 2.0 µl injections, following a 0.5 µl initial injection. Control experiments were performed by titrating the same concentration of Elg1 single SIM mutants into the ITC buffer. As for the fitting strategy, we assumed that all three SIMs identically bind to the Smt3. Therefore, we multiplied the Elg1 concentration by 3 as the effective concentration when we fitted the titration data of Elg1 WT, and we

multiplied the Elg1 concentration by 2 when we fitted the titration data of Elg1 single SIM mutants, assuming that mutations on the SIMs abolished the binding ability. After the Elg1 concentration was multiplied, we fitted the data with an independent binding model same as above.

2.8 NMR spectroscopy

NMR spectroscopy was performed to further investigate how Elg1 interacts with Smt3. The ^{15}N labeled Smt3 was expressed in M9 minimal media that contains ^{15}NH$_4$Cl as the sole nitrogen source and was induced overnight. Unlabeled Elg1 WT and ^{15}N labeled Smt3 were purified as described in previous sections.

Purified Smt3 and Elg1 WT was exchanged into NMR buffer (Table 2) via size exclusion chromatography. Two samples for NMR were prepared by directly utilizing the exchanged sample. The sample containing only Smt3 was prepared by mixing the same volume of the Smt3 sample and the buffer, followed by adding D$_2$O (10% final). A total of 167 μM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) was added as an internal reference, and 3 mM imidazole was added as an internal pH indicator. The final concentration of Smt3 in the NMR sample was 61 µM. The other sample was prepared by mixing the same volume of the Smt3 sample and the Elg1 sample, followed by adding the same additives, which contain the highest Elg1: Smt3 ratio (5.5:1). Other samples with different ratios were made by mixing these two samples or their mixture with the same volume of Smt3 sample.

The samples were equilibrated on the Varian INOVA 600-MHz NMR spectrometer for 15 minutes after loading, and the signal was collected at 25°C for 30 minutes. The NMR spectrum was acquired with the following parameters: number of complex points= 1022, sweep width of 7002.8 Hz in ^{1}H and 1700 Hz in ^{15}N, 64 increments in second dimension and 32 scans for each increment. All data were processed using NMRPipe [62] and plotted using NMRViewJ [63]. Smt3 was previously assigned by NMR spectroscopy [64], and this was used to assign signals in the current spectra according

to the signals' relative positions. We presume the three SIMs are identical regarding their abilities to bind Smt3. Therefore, we set the n value as 3.0 and fit all data sets with shared K_d. The chemical shifts to calculate K_d are listed in Table 6. The curves to determine the K_d were plotted with the equations below:

$$Y = n * B * \frac{PL}{Pt}$$

$$PL = \frac{(nPt + nX + K_d) - \sqrt{(nPt + nX + K_d)^2 - 4n^2 XPt}}{2n^2}$$

(Y − chemical shift changes; PL - protein-ligand complex concentration; Pt - total protein concentration; X - added ligand concentration; K_d - dissociation constant (same units as Pt); n - the number of identical binding sites on ligand; B − normalization factor (a constant that used to adjust the data to the spectrum range))

The equation is used to measure changes in the chemical shift with the equation below:

$$\Delta Csp = \sqrt{(\Delta\delta(^1H))^2 + (0.2 * \Delta\delta(^{15}N))^2"}$$

The ΔCsp represents the overall chemical shift perturbations, $\Delta\delta(^1H)$ represents the chemical shift on ^{1}H, and $\Delta\delta(^1H)$ represents the chemical shift on ^{15}N. Those data were utilized to map the interacting surface. The basic structures were from PDB 3V62 [26] (contains the structure of Srs2 SIM binding to Smt3) and 2EKE [65] (contains the structure of yUbc9 binding to Smt3). The interacting surface of Smt3 was mapped in Pymol via locating the positions of the residues that have chemical shifts that are larger than average.

2.9 Crystallization

Two strategies were used to screen for crystals. One is to directly mix Elg1 fragments and PCNA. The proteins used for crystallization were exchanged in the ITC buffer via size exclusion chromatography. The protein complex of PCNA and Elg1 was prepared by directly mixing the two proteins at a ratio of 1: 1.2 and was concentrated for usage. Both hanging drop and sitting drop vapor diffusion were used to screen with commercially available and house-made crystallization kits.

Additionally, we used the synthesized peptide that contains the PIP box to complex PCNA/sPCNA for screening. The sequence of the synthesized peptide is DSVIFLNHSVVKPIEAVSK, which contains the proposed PIP box in Parnas *et al.* [2]. Unfortunately, the peptide is difficult to dissolve in aqueous solution. For this reason, we adjusted the pH of the peptide solution to the same as the buffer after dialysis. We directly mixed the dialyzed peptide with PCNA/sPCNA in an excess ratio (protein:peptide=1:3). The protein-peptide mixtures were used to screen with commercially available and home-made crystallization kits.

3 Results

3.1 Interactions between Elg1 and PCNA/sPCNA

3.1.1 Construct design of Elg1 SIM mutants

Several Elg1 SIM mutants have been generated to investigate the roles of different motifs within the Elg1 N-terminal fragment. To clone plasmids that have SIM mutations within Elg1 N-terminus, the M-PIPE cloning method was employed using the Elg1 WT construct as the template. The sequences of the primers used to generate Elg1 SIM mutants are listed in Appendix A, and the primers were synthesized by UWO Oligo Factory. Three sets of primers were designed to mutate SIM1, SIM2, and SIM3. We mutated the hydrophobic amino acids of SIMs into alanines and kept the acidic stretch (mutated sequences are shown in Figure 6). All 7 versions of SIM mutants were constructed via M-PIPE cloning. These detailed protocols are already described in the Materials and Methods.

3.1.2 Protein expression and purification

3.1.2.1 Expression and purification of yeast PCNA

The construct of PCNA we used has a mutation that changes K into G at the site of 127. That was made to abolish the extent of SUMOylation on site 127 to obtain more sPCNA with SUMOylation at K164. Yeast PCNA purification procedures include ammonium
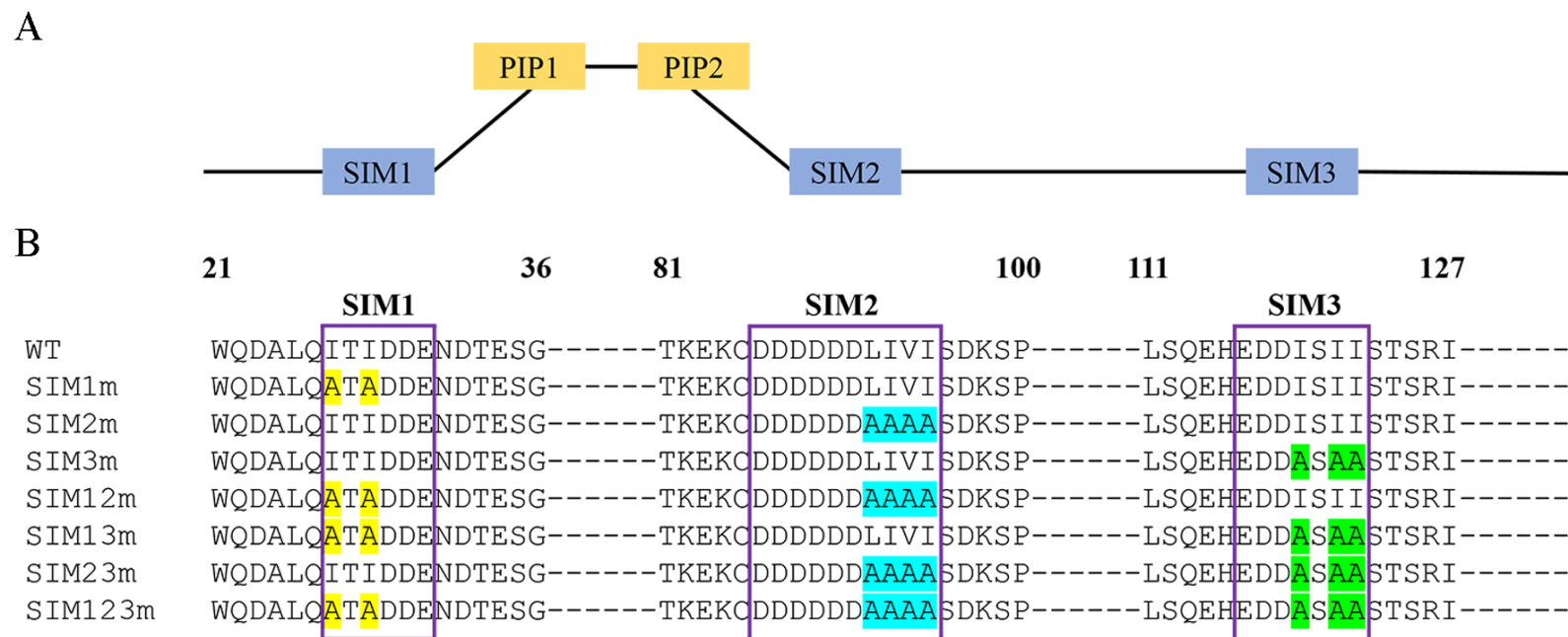
Figure 6: Construct design of Elg1 SIM mutants.

Panel A shows the order of five putative motifs in the Elg1 N-terminus, and panel B presents the mutations on the Elg1 WT with various combinations. The exact regions of SIMs are outlined with the purple boxes. The mutations on SIM1, SIM2, and SIM3 are highlighted in the yellow, light blue, and green, respectively.
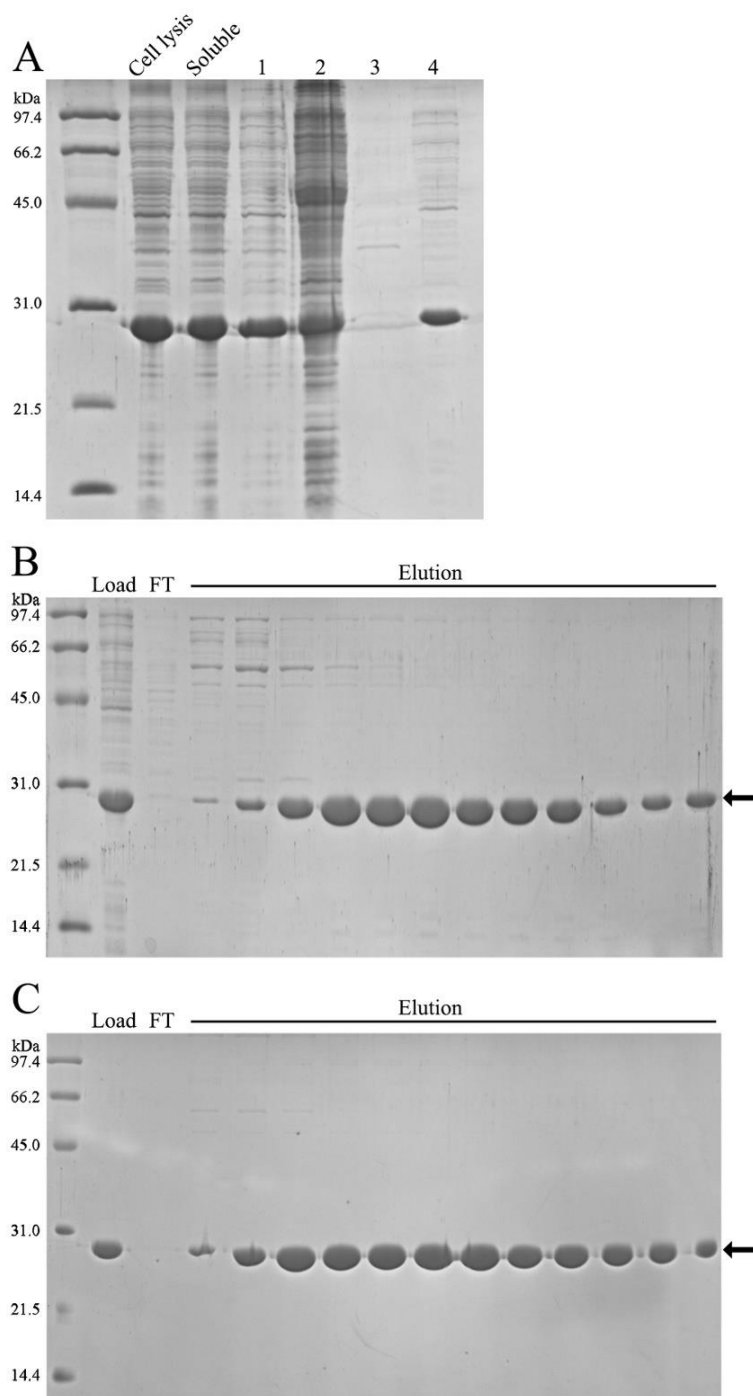
Figure 7: Purification of yeast PCNA.

SDS-PAGE analysis of purification samples after each step of chromatography. The position of yeast PCNA on SDS-PAGE is indicated with arrows.

A) Ammonium sulfate precipitation to purify yeast PCNA. Cell lysis (2 lanes to left of lane 1), soluble protein (to left of lane 1), 45% ammonium sulfate saturation supernatant (lane 1), 45% ammonium sulfate saturation precipitation (lane 2), 75% ammonium sulfate saturation supernatant (lane 3), and 75% ammonium sulfate saturation precipitation (lane 4) are indicated.

B) The first round of anion exchange chromatography using Q column.

C) The second round of anion exchange chromatography using Q column.

sulfate precipitation and two rounds of anion exchange chromatography.

As seen in Figure 7A, the dominant band in the cell lysis showed the position of PCNA on SDS-PAGE (slightly below 31.0 kDa marker). For ammonium sulfate precipitation, most PCNA remained in the supernatant (lane 1), while a lot of contaminants were precipitated with 45% ammonium sulfate saturation and removed via centrifugation (lane 2). With the addition of ammonium sulfate at 75% saturation in the second precipitation, most PCNA was precipitated (lane 3). Two rounds of anion exchange chromatography were later applied to further purify the protein (Figure 7B and 7C).

The PCNA starts to elute when the conductivity reached 36 ms/cm. Fractions that contain sufficiently pure PCNA were pooled. The yield of PCNA was over 100 mg for a one-liter culture. The PCNA has a molecule weight of approximately 28.9 kDa, and its purity after purification was checked with SDS-PAGE (Figure 8).

3.1.2.2 Expression and purification of SUMOylation enzymes

These enzymes were expressed and purified in various conditions as described in the section 2.4.2. The procedures to purify E1 and E3 includes nickel affinity chromatography, digestion and dialysis, a second nickel affinity chromatography, and anion exchange chromatography. The procedures to purify E2 are similar, with the last step changed to cation exchange chromatography, due to E2's high overall PI. All proteins were concentrated to over 10 mg/ml and marked with their molar concentration for the convenience of calculation. It is also critical to note the NaCl concentration in which the proteins were kept, for the SUMOylation reaction requires approximately 100 mM NaCl to avoid slow and insufficient SUMOylation. Those proteins-yUba2 1-544 (61.3 kDa) and yAos1 (39.3 kDa), yUbc9 K153R (17.9 kDa), and ySiz1 166-508 (38.8 kDa)- were resolved on SDS-PAGE after purification to check purity (Figure 8).

3.1.2.3 Expression and purification of Smt3 Δ18K19R

The plasmid pMCSG9-Smt3 Δ18K19R was transformed and induced as described in the

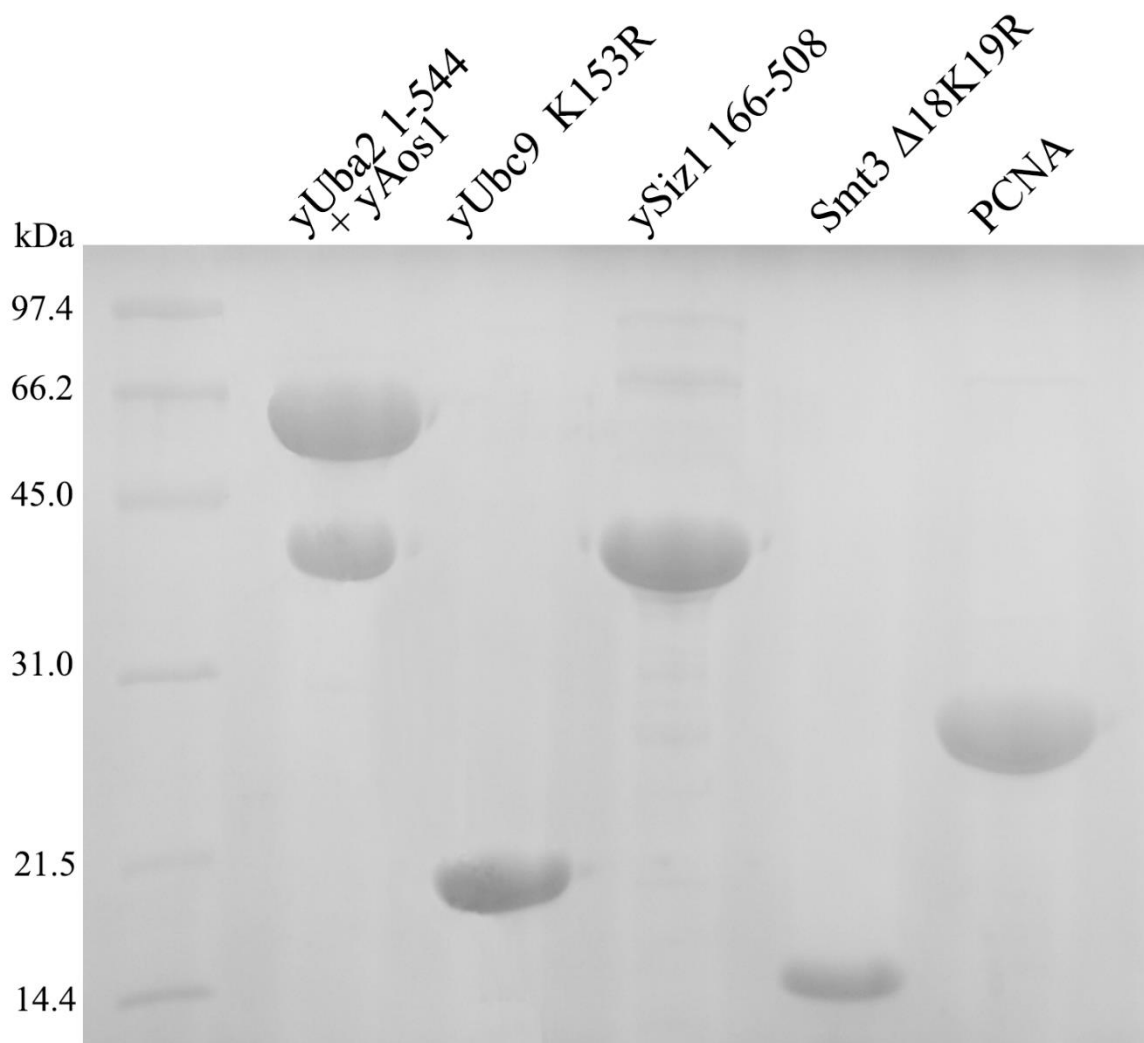**Figure 8: SDS-PAGE analysis of SUMOylation enzymes and substrates.**
The quality of purified SUMOylation enzymes and substrates were checked in this SDS-PAGE. Protein yUba2 1-544 (first lane upper band, 61.3 kDa) , yAos1 (first lane lower band, 39.3 kDa), yUbc9 K153R (17.9 kDa), ySiz1 166-508 (38.8 kDa), Smt3 Δ18K19R (9.3 kDa) and PCNA (28.9 kDa) are presented. All proteins were over 95% pure after purifications.

Materials and Methods. Expressed protein was purified with MBP affinity and anion exchange chromatography. Fractions that contain sufficiently pure Smt3 Δ18K19R was pooled after anion exchange chromatography and concentrated for future assays. Protein Smt3 Δ18K19R (9.3 kDa) was over 95% pure after purification, and it ran at a position of approximately 14.4 kDa on the SDS-PAGE gel (Figure 8).

3.1.2.4 Expression and purification of GST-tagged Elg1 fragments

As described in the Materials and Methods, these recombinant proteins were expressed in BL21(DE3) pRare and purified by nickel affinity chromatography and anion exchange chromatography. The representative purification (GST Elg1L1) is shown in Figure 9.

Each Elg1 N-terminal fragment was expressed together with a GST tag with 6 histidines added before the GST. Therefore, we can use the nickel columns to purify these proteins. After washing with the buffer that has a low concentration of imidazole, the protein eluted with His Buffer B (Figure 9). The eluted sample (contains approximately 500 mM NaCl) was diluted to approximately 100 mM NaCl and directly loaded onto Q HP columns (GE Healthcare) for anion exchange chromatography. The protein eluted with an increasing gradient of Anion Buffer B (Figure 9). Different Elg1 fragments eluted at different conductivities, within the range of 16 to 24 ms/cm. All purified proteins were concentrated and kept in -20°C with 50% glycerol (final concentration), while the proteins were kept over 10 mg/ml at final concentration. Normally, the yield of one fragment would be at least 6 mg for a one-liter culture. The purity of the protein was checked with SDS-PAGE during purification.

3.1.2.5 Expression and purification of Elg1 WT and Elg1 SIM mutants

As described in the Materials and Methods, the Elg1 WT and Elg1 mutant protein were expressed in BL21(DE3) pRare and purified by nickel affinity chromatography, digestion and dialysis, second nickel affinity chromatography and anion exchange chromatography. Shown in Figure 10A, the dominant band indicated the position of expressed protein (Elg1L1), and SDS-PAGE was used to check protein purity after nickel affinity chromatog-

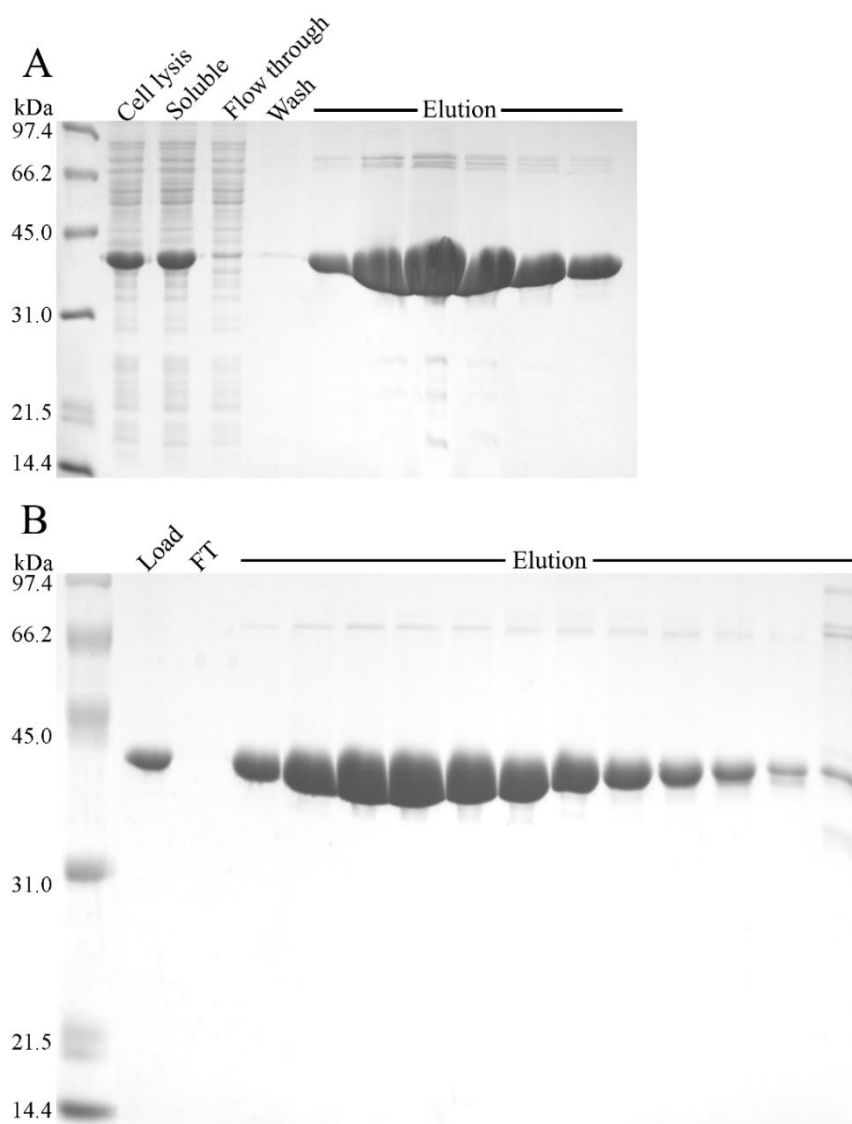**Figure 9: Representative purification of GST-tagged Elg1 fragments.**
This purification utilized cells expressing GST-fused Elg1L1. A) SDS-PAGE analysis of nickel affinity chromatography fractions. Cell lysis, soluble protein, flow through (marked as FT in the figure), wash, and fractions of elution are indicated; B) Anion exchange chromatography via Q column. The protein was over 90% pure after purification.
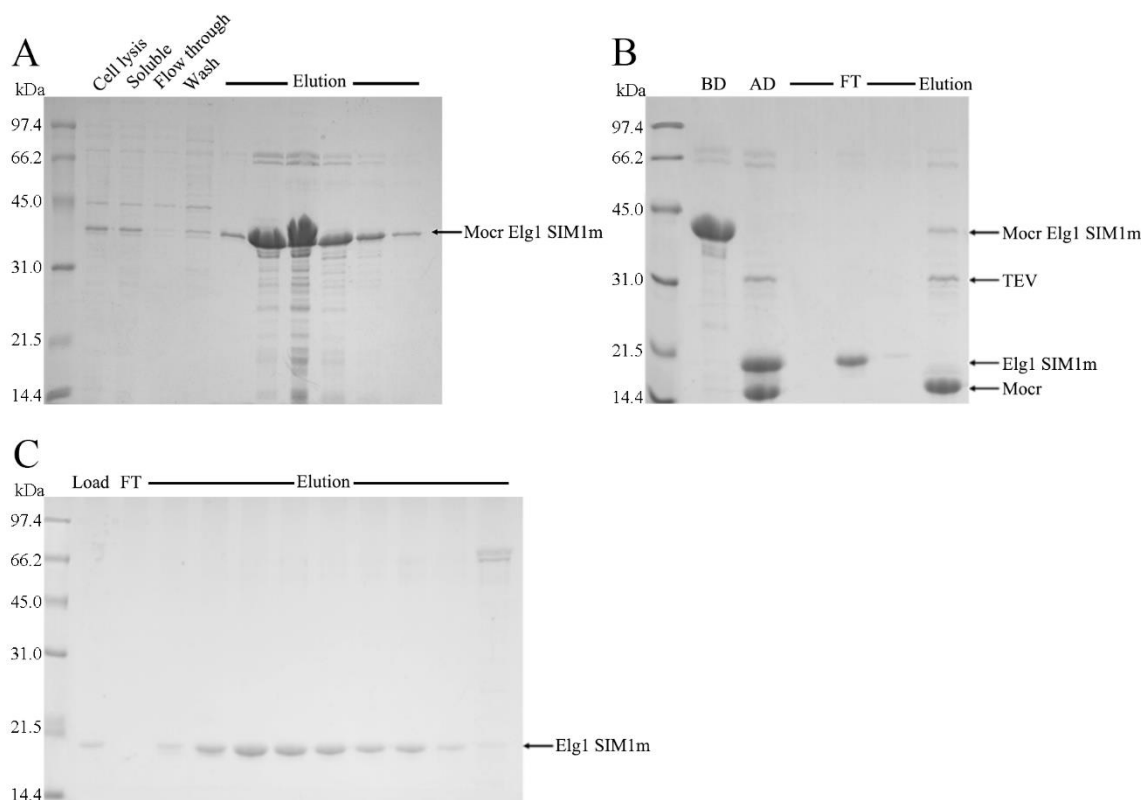
Figure 10: Representative purification of Elg1 SIM mutants.

This figure shows the purification of protein Elg1 SIM1m.

A) Nickel affinity chromatography of Elg1 SIM1m using the nickel column. Expressed tagged protein is indicated with the arrow; B) The second nickel affinity chromatography of Elg1 SIM1m. The lane "BD" represents the sample before digestion and dialysis; the lane "AD" represents the sample after digestion and dialysis. After loading onto the nickel columns, the protein of interest stayed in the flow through (indicated as "FT"); C) The anion exchange chromatography of Elg1 SIM1m using Q columns. The position of Elg1 SIM1m is indicated with arrow. The protein was over 95% pure after purification.

-raphy. Fractions that have protein of interest were pooled and put into dialysis, with the addition of TEV to cut off the tag. The digested sample was loaded onto the nickel columns that were pre-equilibrated with the buffer that contains extra 15 mM imidazole. The flow-through sample was subjected to the anion exchange chromatography (Figure 10C). Normally, we can get approximately 6 mg from a one-liter culture. The purified mutant proteins tend to aggregate in high protein concentrations compared to Elg1 WT. Therefore, the mutants were kept with cations via the flash frozen method. The protein was refreshed and exchanged into the required buffer when they were needed.

3.1.3 SUMOylation of PCNA

Proteins - 15 µM PCNA, 60 µM Smt3, 90 nM yAos1/yUba2 1-544 (E1), 300 nM yUbc9 K153R (E2), and 2 µM ySiz1 166-508 were mixed together in the SUMOylation buffer. The reaction was started with the addition of ATP (2 mM for final concentration). Different time points were tested to obtain the sufficiently pure sPCNA (Appendix B). Excess EDTA was added to the SUMOylation products to slow down the reaction, and the sample was centrifuged or filtered to remove any precipitation. Shown in Figure 11, the sample was purified by anion exchange chromatography, and sPCNA's purity was checked by SDS-PAGE. It is noteworthy that the SUMOylation products are not 100% pure and contains a small amount of unmodified PCNA and some double-SUMOylated PCNA. Normally, 20ml of reaction sample would produce about 9 mg of sPCNA. Protein sPCNA has a molecule weight of approximately 37.9 kDa, and it ran at the slightly lower position than 45.0 kDa on SDS-PAGE gels.

3.1.4 GST pull-down assay

To evaluate the functions of the motifs in the Elg1 N-terminus, GST pull-down assays were performed by using different GST-tagged Elg1 fragments containing various combinations of the protein binding motif (Figure 12A). Those fragments are named based on their lengths, such as Elg1L1 (longest, 21-102) and Elg1L5 (shortest, 64-84). The pull-down assay with GST as bait was performed as the negative control, and the pull-down assay that
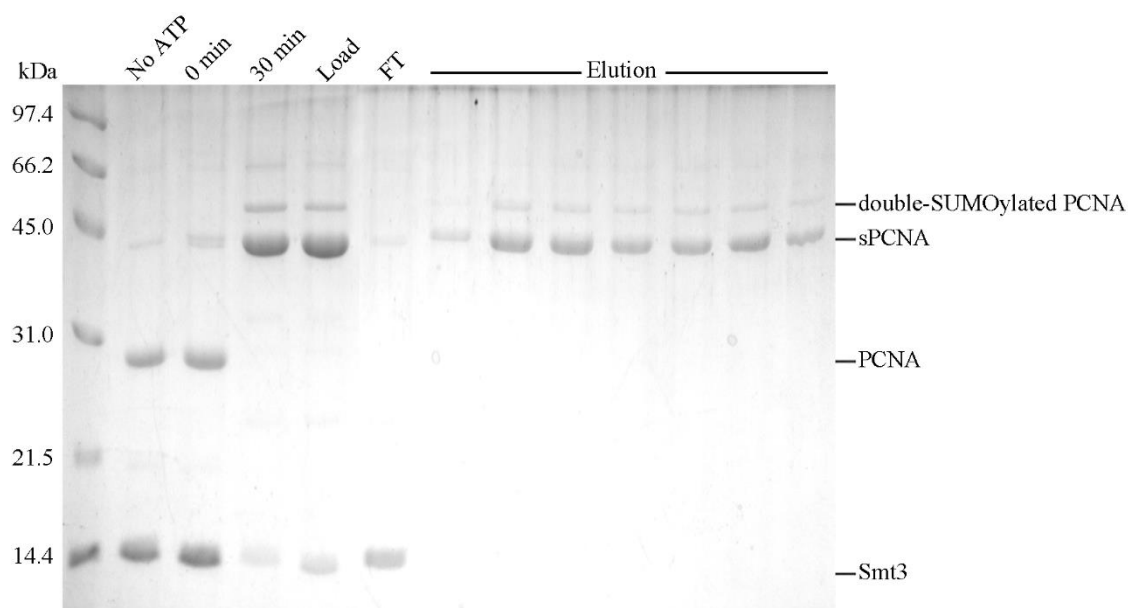
Figure 11: Purification of SUMOylated PCNA.
The lane marked "No ATP" represents the sample taken before the addition of ATP; the lane marked "0min" represents the sample taken right after the addition of ATP to stimulate SUMOylation; the lane marked "30min" represents the sample taken after 30min of room temperature incubation; the lane marked "Load" represents the sample before being loaded onto the Q column; "FT" and "Elution" represent the flow through of the Q column during sample loading and collection of eluted fractions, respectively. The positions of double-SUMOylated PCNA, sPCNA, PCNA, and Smt3 are marked on the right.

used GST-p21 as bait and hPCNA as prey was set up as the positive control (Figure 12B and 12C). A summary of the pull-down results with different Elg1 fragments is listed in Table 3. From the observation of the output lanes that used Smt3 as prey, Smt3 showed no binding with Elg1 N-terminus in this assay. This outcome may be a result of low binding affinity,

suggesting that a need for a higher molar concentration of prey (Smt3 in this case), or GST pull-down assay is not suitable for assessing Elg1 SIMs' binding abilities with Smt3.

Only one of the Elg1 N-terminal fragment pull-down assays showed a faint band (marked red in Figure 13A) that indicated interaction with PCNA on SDS-PAGE analysis. The fragment is Elg1L3 that contains no SIMs. Even though the band is weak, it still confirms the Elg1 N-terminus' binding ability with PCNA and also indicates low binding affinity between Elg1 and PCNA under normal conditions. However, other Elg1 N-terminal fragments that have more functional motifs than just PIP boxes did not show interaction with PCNA. The presence of SIM motifs in those longer constructs results in reduced interactions with PCNA, suggesting the SIMs may play an inhibitory role in the interaction with PCNA. Additionally, we confirmed the Andy Liu's results with our method that eluting proteins off the beads. Relatively strong PCNA binding was shown with the construct Elg1L3+ (Figure 14). Moreover, several Elg1 N-terminal fragments showed the ability to bind sPCNA, which is Elg1N (marked in blue in Figure 13A left), Elg1L3 (marked in blue in Figure 13A right) and Elg1L1 (marked in blue in Figure 13B). Comparing the detected sPCNA bands from Elg1N (21-150) and Elg1L1 (Elg1 21-102), a similar relative intensity is seen, suggesting SIM3, which is located between 103 and 150, may have subtle effects in the interaction with sPCNA. In addition, deletion of more SIMs results in no detection of sPCNA in GST pull-down assays, which indicates one single SIM may be not enough to promote the binding with sPCNA.

3.1.5 ITC

After the confirmation of the interactions between the Elg1 N-terminus and PCNA/sPCNA with GST pull-down assays, we attempted to further investigate these interactions. ITC
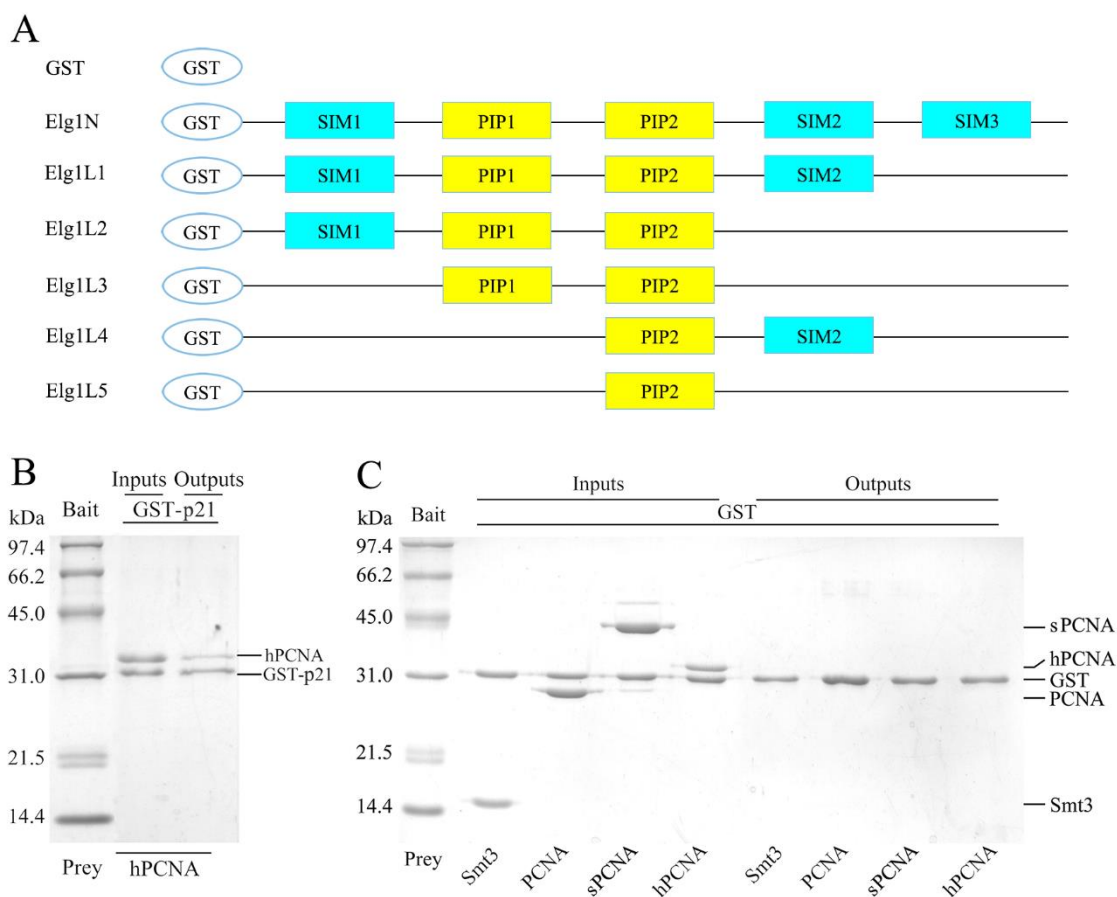
Figure 12: Structure of Elg1 fragments and GST pull-down controls.

A) The bar diagram represents the functional motifs of the Elg1 N-terminus that exist in the Elg1 constructs used in this assay. The missing motifs were deleted from the fragment.
B) Positive control experiment performed in this assay. The position of GST-p21 and hPCNA are indicated.
C) Negative control experiments performed in this assay. GST was used as the bait, while Smt3, PCNA, sPCNA, and hPCNA are used as prey. The result shows GST alone cannot bind to any of the preys.
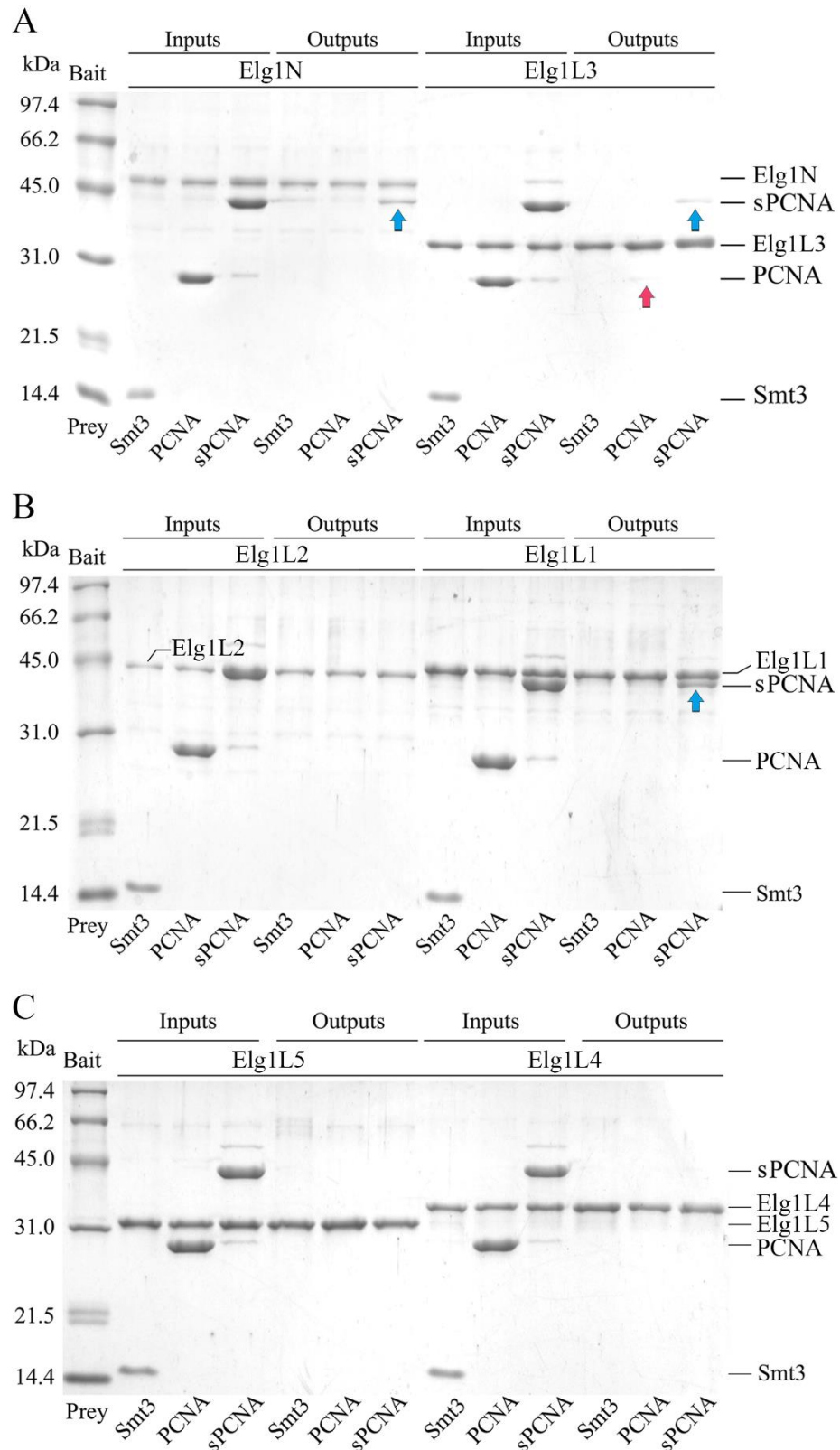
Figure 13: GST pull-down assays with various Elg1 N-terminal fragments.
A) Smt3, PCNA, and sPCNA as the prey and Elg1N (left)/ Elg1L3 (right) as the bait; B) Smt3, PCNA, and sPCNA as the prey and Elg1L2 (left)/ Elg1L1 (right) as the bait; C) Smt3, PCNA, and sPCNA as the prey and Elg1L5 (left)/ Elg1L4 (right) as the bait. Positive detections in the outputs are indicated with the red arrow for PCNA and the blue arrow for sPCNA.
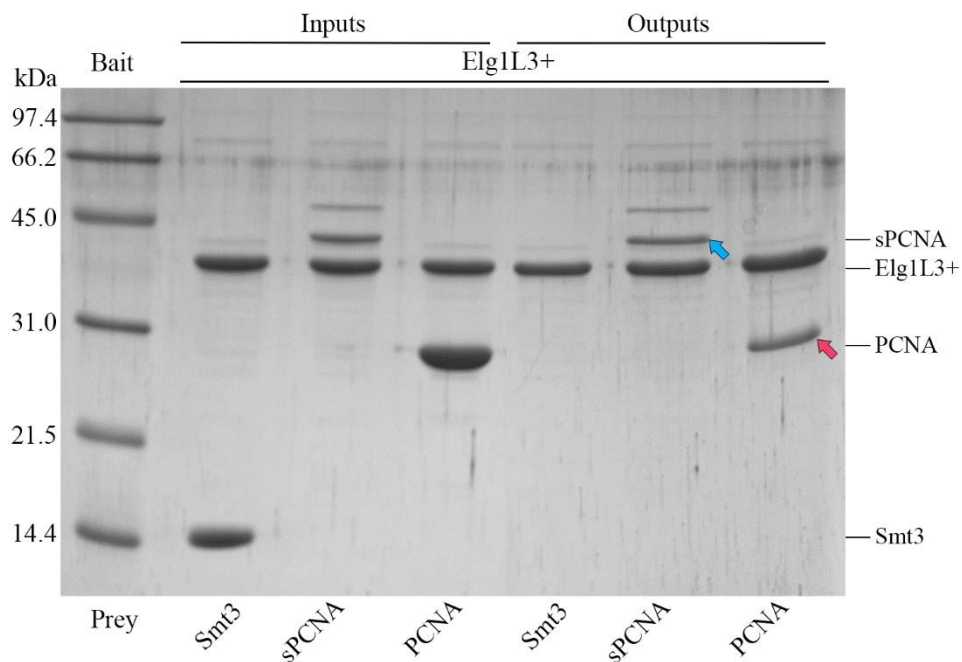
Figure 14: GST pull-down assay with Elg1L3+.
We used a different method from the one in Andy's thesis and confirm the results. Protein Elg1L3+ was used as the bait, and Smt3, PCNA, and sPCNA were used as the prey. Positive detection in the outputs is indicated with a red arrow for PCNA and a blue arrow for sPCNA. The results showed Elg1L3+ could interact with PCNA and sPCNA, but not with Smt3.

Table 3: A summary of GST pull-down assay results.

Short Name	Bait	Motifs	Prey		
			Smt3	PCNA	sPCNA
GST	GST	NA	-	-	-
Elg1N	Elg1 21-150	SIM1/PIP1/PIP2/SIM2/SIM3	-	-	+
Elg1L1	Elg1 21-102	SIM1/PIP1/PIP2/SIM2	-	-	+
Elg1L2	Elg1 21-84	SIM1/PIP1/PIP2	-	-	-
*Elg1L3+	Elg1 33-84	PIP1/PIP2	-	+	+
Elg1L3	Elg1 46-84	PIP1/PIP2	-	+	+
Elg1L4	Elg1 64-102	PIP2/SIM2	-	-	-
Elg1L5	Elg1 64-84	PIP2	-	-	-

All Elg1 fragments contain a GST tag at the N-terminus.

sPCNA is the short name of SUMOylated PCNA.

* The referred results were from Andy Liu's thesis, and these results were confirmed by us.

experiments were performed on the NanoITC. For an ITC titration, the energy absorbed or given off by protein-protein interactions are detected, and the analysis program can determine the thermodynamic data, such as the dissociation constant (K_d), stoichiometry (N), enthalpy (ΔH), and entropy (ΔS).

During test runs, the titration between the Elg1 N-terminus and PCNA was difficult to make an accurate analysis. We tried to titrate 830 µM PCNA into 79 µM Elg1L3+ (Appendix C). The construct Elg1L3+ (33-84) showed relatively strong binding with PCNA in GST pull-down assays among all constructs tested (Figure 14). Therefore, we used this construct for the test run. Although enough heat signals were produced for detection in ITC, an apparent K_d of over 424 µM was determined after analysis, which was significantly larger than the concentration of PCNA in the cell. Therefore, the accuracy of this analysis is questionable. Alternatively, we tried to characterize the interaction between the Elg1 N-terminus and sPCNA. The titration of sPCNA to Elg1 and titration of Elg1 to sPCNA were both tested. We observed that the titration of Elg1 to sPCNA tends to produce aggregation (samples became turbid), which reduced the effective protein concentration in binding. Therefore, the characterization of the binding was conducted via titration of sPCNA to Elg1. Shown in Figure 15A, 406 µM sPCNA were titrated into 60 µM Elg1 WT that contains all functional motifs for the binding at 25°C with a stirring speed of 250 rpm. The control experiment was performed by titrating the same concentration of sPCNA into the ITC buffer. The experimental data were subtracted with the control experiment and fitted using a one-site binding model. The main parameters of binding from the analysis are listed in Table 4. The apparent K_d for this titration between Elg1 WT and sPCNA is 8.23±3.31 µM, with an N value of 1.122±0.073, ΔH of -13.07±1.18 kcal/mol, ΔS of -20.57 kcal/mol·K.

To investigate the effects of SIMs on the interactions, constructs containing various combinations of SIMs mutations were made. Protein 406 µM sPCNA was titrated into 60 µM Elg1 SIM mutants that have various combinations of SIM mutations. The same setting as for Elg1 WT was employed to perform these titrations. However, it appears that Elg1

Figure 15: ITC titrations between sPCNA and Elg1/Elg1 SIM mutants.
All titrations were performed at 25°C with a stirring speed of 250 rpm. The control experiments were shown in green, and the protein to protein titration experiments were shown in red. A) 406 µM sPCNA titrated into 60 µM Elg1 WT; B) 406 µM sPCNA titrated into 60 µM SIM1m; C) 406 µM sPCNA titrated into 60 µM SIM2m; D) 406 µM sPCNA titrated into 60 µM SIM3m.

Continued Figure 15: ITC titrations between sPCNA and Elg1/ Elg1 SIM mutants.
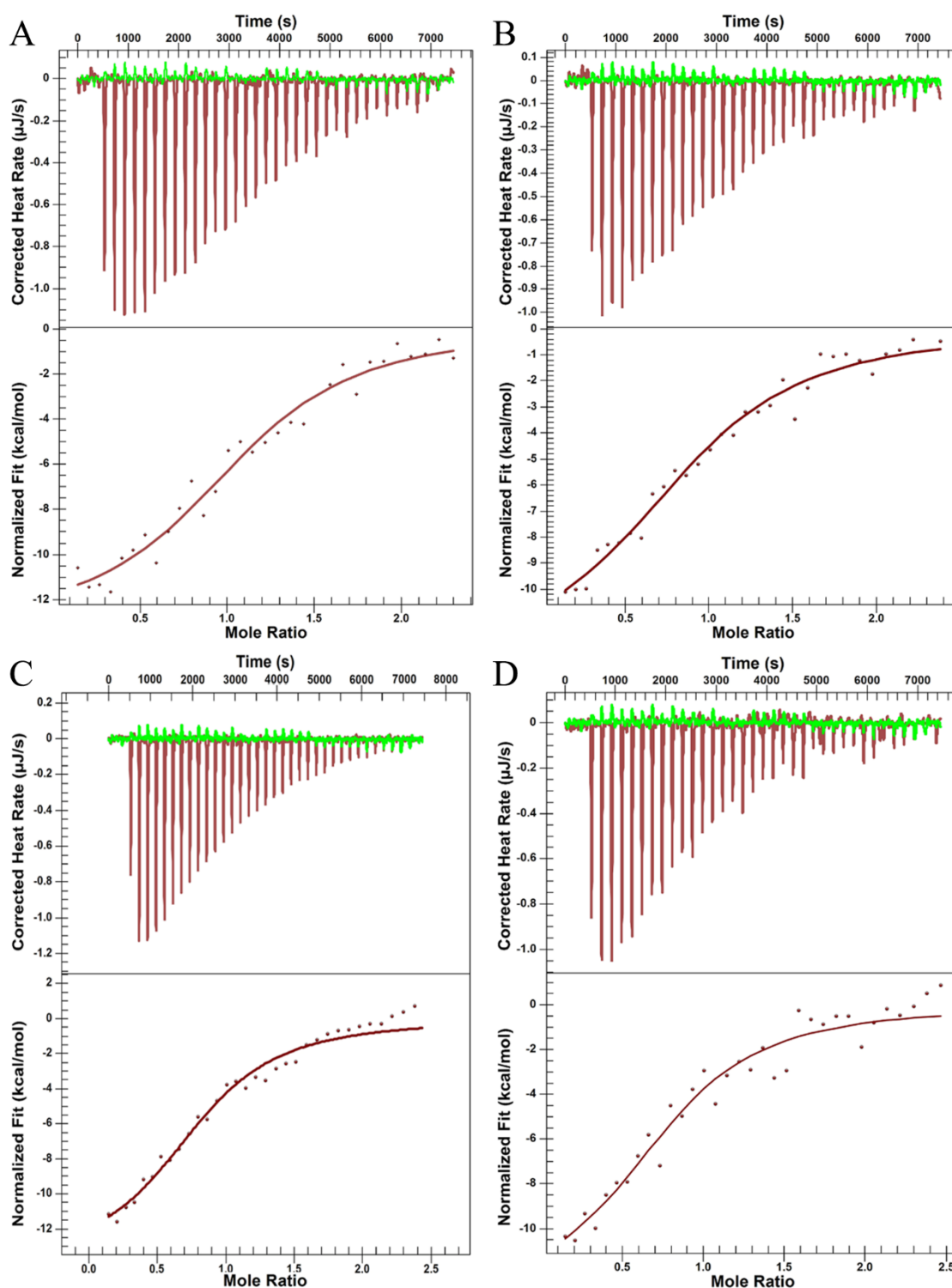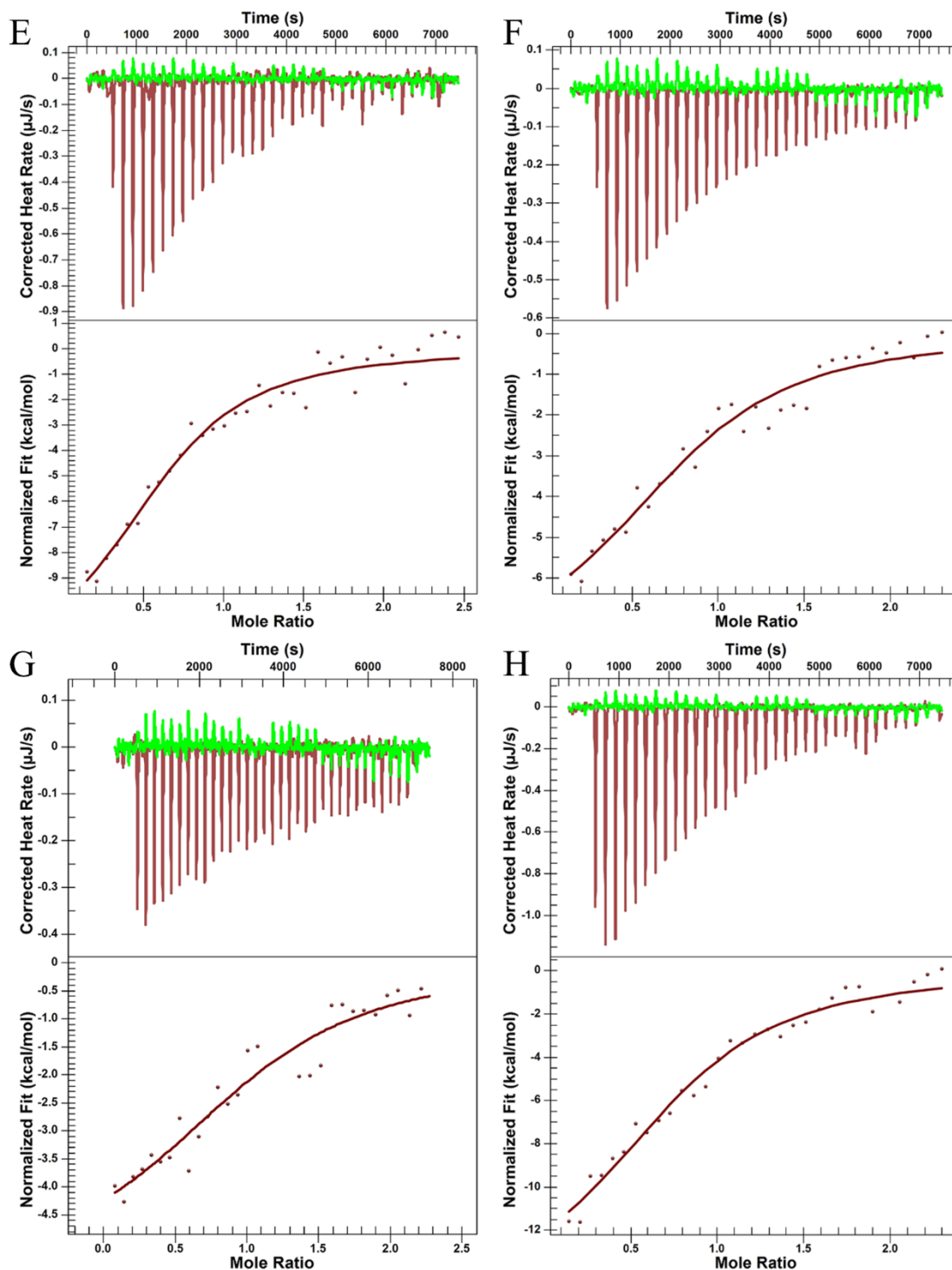All titrations were performed at 25°C with a stirring speed of 250 rpm. The control experiments were shown in green, and the protein to protein titration experiments were shown in red. E) 406 µM sPCNA titrated into 60 µM SIM12m; F) 406 µM sPCNA titrated into 60 µM SIM13m; G) 406 µM sPCNA titrated into 60 µM SIM23m; H) 406 µM sPCNA titrated into 60 µM SIM123m.

Table 4: A summary for ITC titrations between sPCNA and Elg1

	Protein in Cell	Protein in Syringe	N	$K_d(\mu M)$	ΔH(kcal/mol)	ΔS(kcal/mol·K)
A	Elg1 WT	sPCNA	1.122±0.073	8.23±3.31	-13.07±1.18	-20.57
B	SIM1m	sPCNA	0.950±0.086	12.13±4.07	-12.72±1.72	-20.18
C	SIM2m	sPCNA	0.850±0.060	8.87±3.36	-13.88±1.58	-23.43
D	SIM3m	sPCNA	0.818±0.112	9.59±4.43	-13.11±2.40	-21.02
E	SIM12m	sPCNA	0.658±0.112	12.97±6.80	-13.22±3.53	-21.99
F	SIM13m	sPCNA	0.849±0.126	14.17±6.62	-8.01±1.68	-4.68
G	SIM23m	sPCNA	1.115±0.186	20.21±14.22	-5.45±1.55	3.19
H	SIM123m	sPCNA	0.804±0.098	13.85±5.93	-15.27±2.86	-28.99

mutants have subtle effects on the binding with sPCNA after data fitting (Figure 15 B-H). As listed in Table 4, the binding affinity of Elg1 SIM mutants was not dramatically decreased, compared to the Elg1 WT. Shown in Figure 15B-D, sPCNA was titrated into Elg1 single SIM mutants (SIM1m, SIM2m, and SIM3m), and the K_d only slightly decreased to 12.13±4.07 µM, 8.87±3.36 µM, and 9.59±4.43 µM, respectively. No dramatic changes were observed in all analyzed parameters. For these Elg1 mutants containing two SIM mutations, titrations with sPCNA had notably weaker titrating signals, especially SIM13m and SIM23m. Titration to the SIM23m that have mutations on SIM2 and SIM3 had the weakest binding affinity (approximately 20 µM) among proteins tested in this study. The interactions for all Elg1 SIM mutants remain to be exothermic and have subtle alterations in apparent K_d. However, there was a significant change in ΔS for SIM13m and SIM23 when compared to the Elg1 WT. When sPCNA interacted with SIM13m, ΔS changed to -4.68 kcal/mol·K, and when sPCNA interacted with SIM23m, ΔS changed to 3.19 kcal/mol·K, comparing to the Elg1 WT (ΔS= -20.57 kcal/mol·K). Furthermore, titration to SIM123m (contains mutations on three SIMs) had a similar apparent K_d (13.85±5.93 µM) with double SIM mutants and maintained similar ΔH (15.27±2.86 kcal/mol) and ΔS (-28.99 kcal/mol·K) when compared to Elg1 WT, suggesting single SIM may not be enough to facilitate the interaction between Elg1 and sPCNA.

3.1.6 Crystallization

To determine the structural basis for the interactions between Elg1 and PCNA/sPCNA, we carried out crystallization of Elg1-PCNA complexes for three-dimensional structural information. Two methods were utilized to generate the complex of Elg1 with PCNA/sPCNA for crystallization.

One method was to mix purified Elg1 proteins directly with PCNA/sPCNA at 1.2: 1 molar ratio and then concentrate the mixture to 10mg/ml for crystallization screening. A drop ratio of 1:1 was used in screening. One hit appeared in the condition (0.1 M Tris/Bicine pH 8.5, 0.03 M MgCl$_2$, 0.03 M CaCl$_2$, 12.5% PEG 1000, 12.5% PEG 3350 and 12.5% MPD) from the Morpheus kit with the protein mixture of Mocr Elg1 21-84 and PCNA. The crystals

were cryoprotected and shot with X-ray, but they did not diffract. We also optimized around this condition, and a hit appeared in similar conditions (0.1 M Tris/Bicine pH 8.5, 0.03 M MgCl$_2$, 0.03 M CaCl$_2$ and 25% PEG 3350) (Figure 16A). Unfortunately, the crystals also did not diffract. The other method was to use synthesized peptides to form the complex with PCNA/sPCNA for screening. The sequence of the synthesized peptide is DSVIFLNHSVVKPIEAVSK, which contains the proposed PIP box in Parnas *et al.* [2] and original Elg1 residues 50-66. Unfortunately, the peptide was difficult to dissolve. We simply mixed PCNA with dialyzed peptide (still not dissolved, but the pH of the solution was adjusted to be the same as the buffer) and used the sample for screening. Several hits appeared (details are shown in Figure 16B-D), but these conditions need further optimization once we can confirm protein-complex crystals exist.

3.2 Interactions between Elg1 and Smt3

From the results above, the complexity of investigating the interaction with sPCNA hindered our progress. There are five putative motifs that are involved in the interactions with PCNA and SUMO, and little is known about how these motifs accommodate each other. Therefore, we altered the approach to investigate the interactions with Smt3, which could help us elucidate the roles of SIMs in the interaction with sPCNA. This may shed some light on the Elg1-sPCNA interaction and provides clues about Elg1-PCNA interaction.

3.2.1 Cloning, expression, and purification of His-Smt3

To investigate the role of different SIMs within the Elg1 N-terminus, we generated the constructs for expression and purification of His-Smt3, using the pMocr Smt3 Δ18K19R as the template. The plasmid pMocr Smt3 Δ18K19R codes the Smt3 mutant with 6 histidines and Mocr tag at the N-terminus. The primers used for mutagenesis were designed following the protocol [59] to remove the Mocr tag and TEV cutting site, which directly linked the protein with the 6 histidines for the convenience of purification. The primers
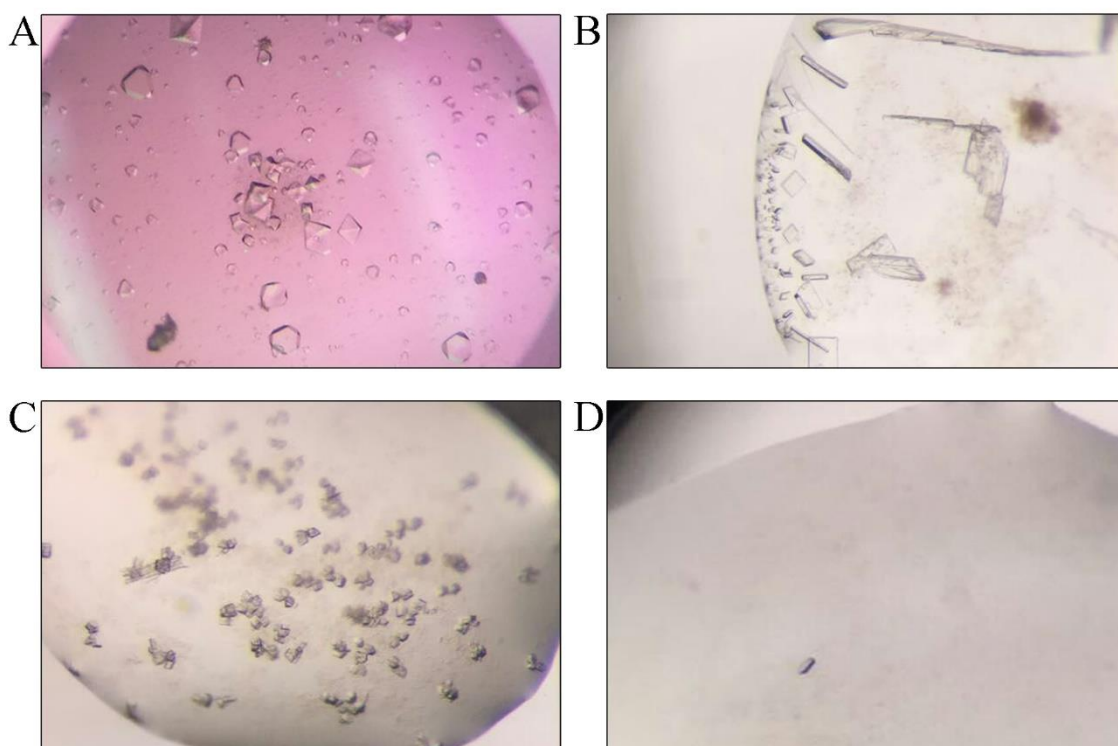
Figure 16: Crystal hits observed from kits.
A) Crystals observed in the condition (0.1 M Tris/Bicine pH 8.5, 0.03 M $MgCl_2$, 0.03 M $CaCl_2$, and 25% w/v PEG 3350), with protein mixture of Mocr Elg1 21-84 and PCNA; B) Crystals observed in the condition (0.1 M Tris pH 8.0, 0.2 M Magnesium chloride hexahydrate, and 20% w/v PEG 6000) with mixture of PCNA and peptide; C) Crystals observed in the condition (0.1 M Sodium acetate pH 5.0, 0.2 M Magnesium chloride hexahydrate, and 20% w/v PEG 6000), with mixture of PCNA and peptide; D) Crystals observed in the condition (0.1 M MMT pH 5.0 and 25% w/v PEG 1500), with mixture of sPCNA and peptide.

were synthesized by UWO Oligo Factory (Appendix A).

After we achieved the correct construct, it was expressed and induced as described in Materials and Methods. The expressed protein was purified via nickel affinity chromatography and anion exchange chromatography. As shown in Figure 17A, His-Smt3 was eluted with His Buffer B. The eluted protein was diluted to approximately 100 mM NaCl for loading onto the Q HP columns (GE Healthcare). A gradient of Anion Buffer B (10% to 30% in 10 column volumes) was applied to purify His-Smt3. The purity of His-smt3 was checked with SDS-PAGE, and fractions that contain sufficiently pure protein (Figure 17B) were pooled and concentrated to over 10 mg/ml. Flash frozen method with liquid nitrogen was used to store the protein for future assays. Normally, approximately 15 mg protein would be obtained from one-liter of culture.

3.2.2 GST pull-down assay

From the GST pull-down assay performed before, no interaction between Elg1 and Smt3 was detected by this method. For the sake of accuracy, we performed this assay again with a higher protein concentration and a higher molar ratio of concentration via the same method. In order to test the interaction of different SIMs with Smt3, we deliberately chose four Elg1 fragments: Elg1N (contains all three SIMs), Elg1L1 (contains SIM1 and SIM2), Elg1L2 (contains SIM1), and Elg1L3+ (contains no SIMs). Also, GST with Smt3 and GST-p21 with hPCNA were used as the negative control and positive control, respectively. Although we overloaded the SDS-PAGE gel, no obvious Smt3 bands were detected (Figure 18). This result suggests that the interaction between Elg1 and Smt3 is not suitable to be detected via GST pull-down.

3.2.3 ITC

Since the GST pull-down assay was not sensitive enough to detect the interaction with Smt3, we used ITC to verify the interaction between Elg1 N-terminus and Smt3, due to ITC's ability to detect a wide range of binding affinity, from mM to nM. Also, for the benefit
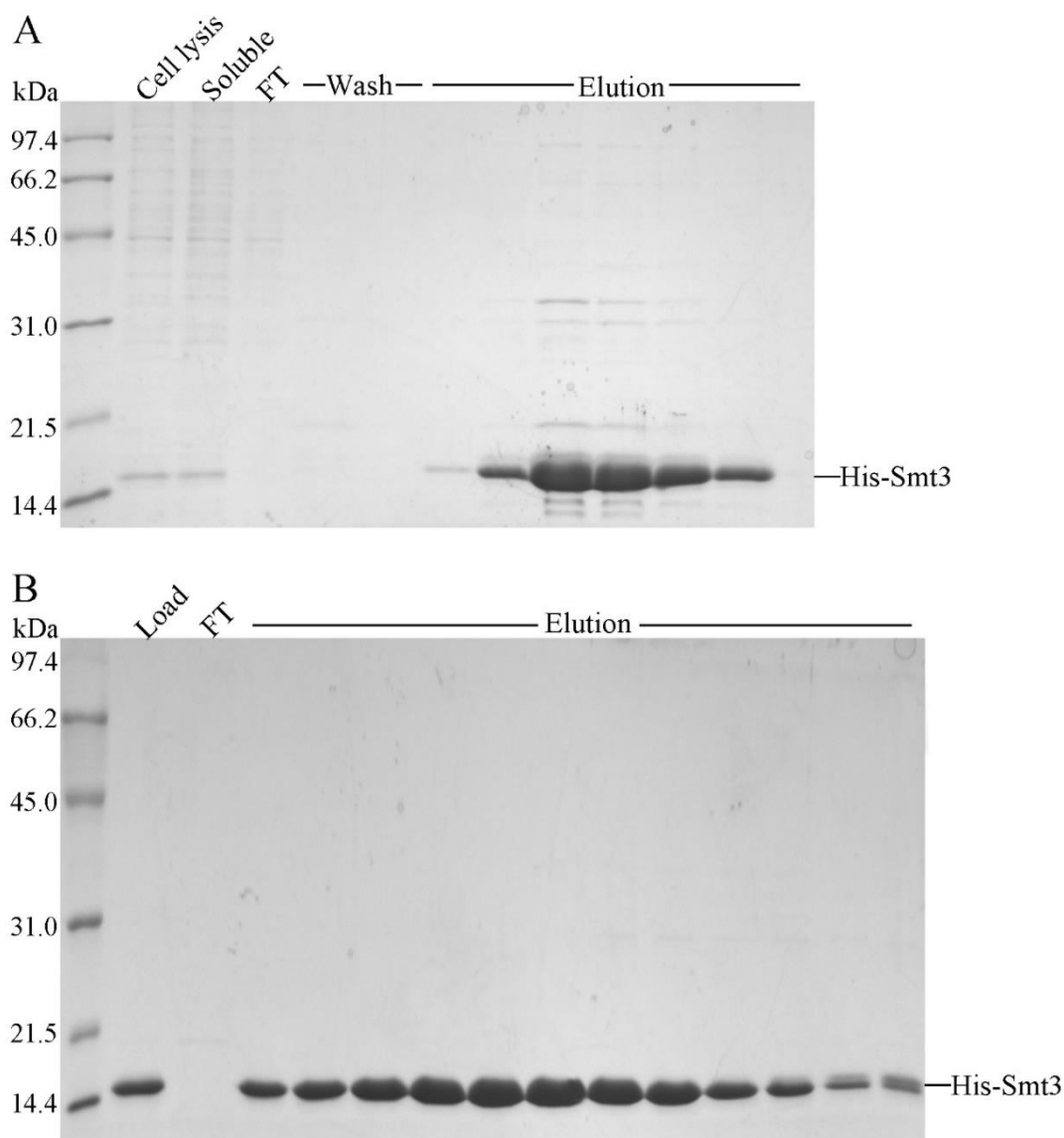
Figure 17: Purification of His-Smt3.

A) Nickel affinity chromatography of His-Smt3 purification. The experiments were performed with nickel columns. Cell lysis, soluble protein, loaded sample, wash, and elution are indicated.

B) Anion exchange chromatography of His-Smt3 purification. This procedure was performed on Q HP columns. The loaded sample, flow through (FT), and elution (with Anion Buffer B) are indicated.
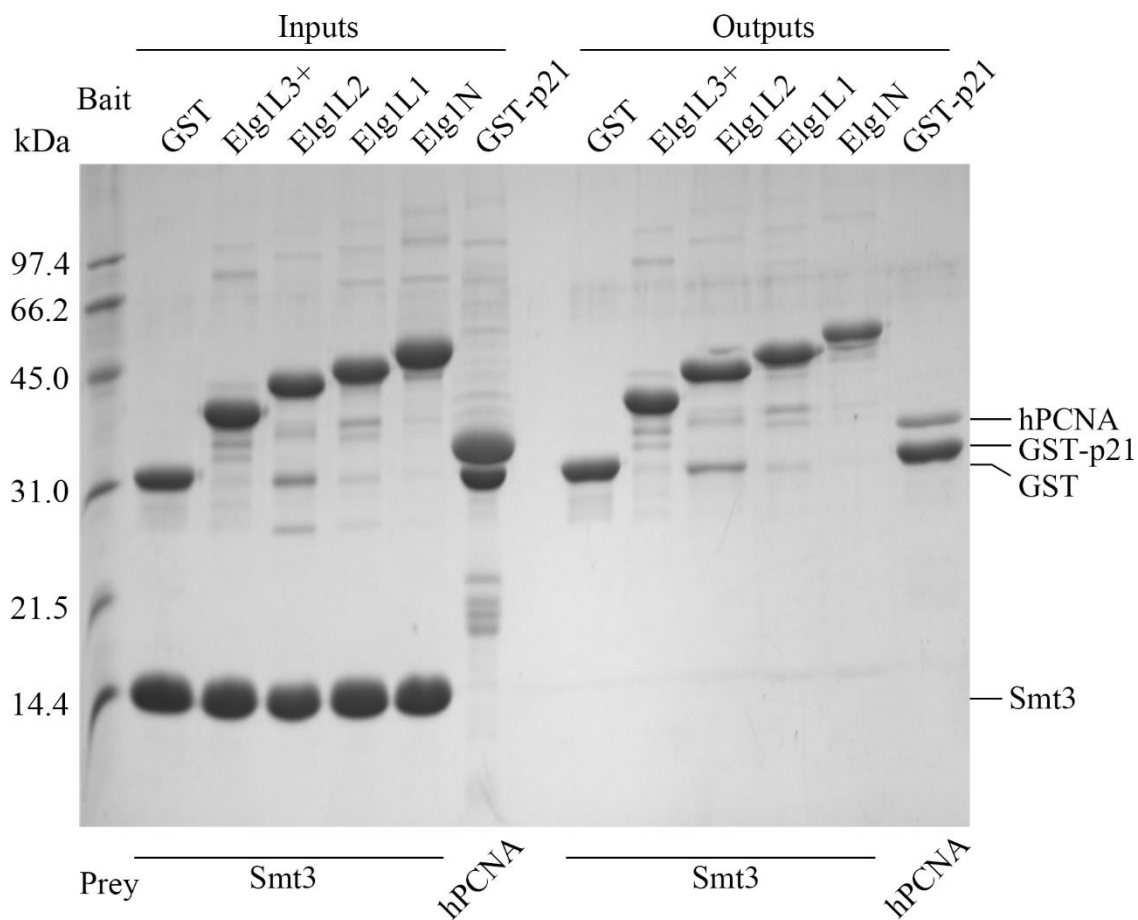
Figure 18: GST pull-down assay with Elg1 N-terminal fragments and Smt3.
The negative control experiment was prepared with GST as the bait. The positive control experiment was prepared with GST-p21 as the bait and hPCNA as the prey. This assay used a higher molar concentration ratio and higher protein concentration of GST-fused Elg1 fragments. No positive detection of Smt3 were shown with these Elg1 constructs used.
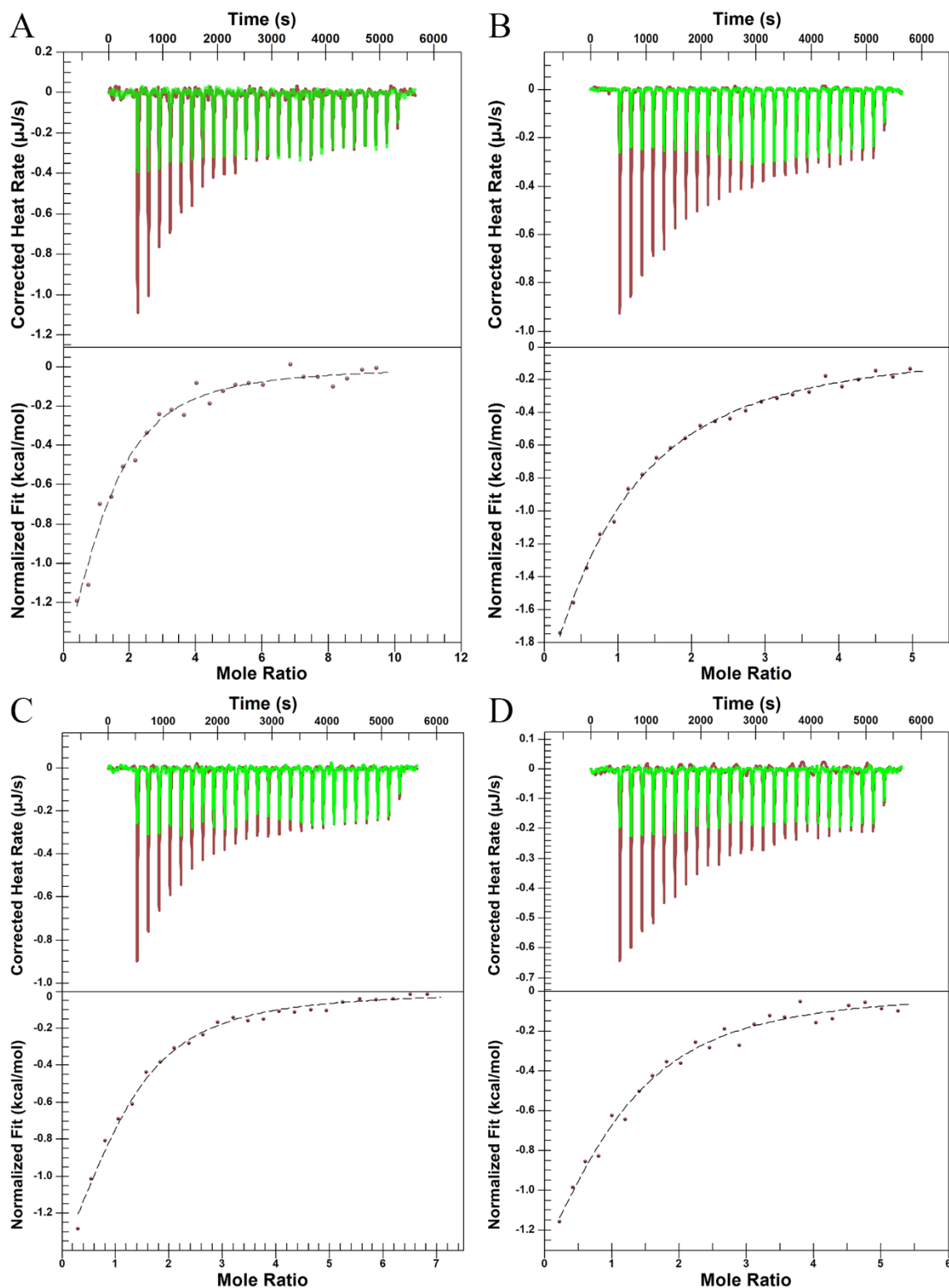
Figure 19: ITC titration between His-Smt3 and Elg1, or Elg1 SIM single mutants.
All titrations were performed at 25°C with a stirring speed of 250 rpm. The control experiments were shown in green, and the protein to protein titration experiments were shown in red. A) 0.72 mM Elg1 WT titrated to 0.1 mM His-Smt3; B) 0.56 mM SIM1m titrated to 0.1 mM His-Smt3; C) 0.78 mM SIM2m titrated to 0.1 mM His-Smt3; D) 0.68 mM SIM3m titrated to 0.1 mM His-Smt3.

Table 5: Summary of the ITC titrations between His-Smt3 and Elg1

	Protein in Cell	Protein in Syringe	Multiply by	N	$K_d(\mu M)$	ΔH(kcal/mol)	ΔS(kcal/mol·K)
A	His-Smt3	Elg1 WT	3	1.096±0.417	80.5±34.3	-2.66±1.58	9.82
B	His-Smt3	SIM1m	2	0.523±0.266	131.8±25.4	-8.11±6.03	-9.45
C	His-Smt3	SIM2m	2	0.686±0.185	77.8±16.4	-3.52±1.26	7.00
D	His-Smt3	SIM3m	2	0.991±0.315	73.8±29.1	-2.38±1.19	10.92

of purification, we had prepared another construct, His-Smt3 Δ18K19R, which expressed protein with 6 extra histidines at the N-terminus.

For the titration of 0.72 mM Elg1 WT to 0.1 mM His-Smt3, we subtracted the signal from the titration to buffer and used Nanoanalyze to fit the independent binding model. Assuming all three SIMs bind to Elg1 identically, we manually multiplied the concentration of Elg1 WT by 3 when we analyzed the data. Shown in Table 5, the apparent binding affinity for Smt3 is 80.5±34.3 μM, suggesting Smt3 binds to SIM in a weak manner. It gives an N value of 1.096±0.417, which corroborated our assumption that one Elg1 N-terminal fragment binds three Smt3. We assume that the mutation on SIMs would abolish their binding ability to Smt3. Therefore, with these Elg1 SIM mutants, we multiplied the concentration of Elg1 mutants by the number of the rest binding sites, 2 in this case. It seems that mutation on SIM1 has the largest impaired effects on the apparent K_d, which decreased 1.6-fold to 131.8±25.4 μM. This result suggests that SIM1 may be the dominant SIM for the interaction with Smt3. In addition, mutations on SIM2 and SIM3 have similar apparent binding affinity as Elg1 WT. The apparent K_d maintained to 77.8±16.4 μM for SIM2m and 73.8±29.1 μM for SIM3m, suggesting SIM2 and SIM3 have the ability to bind Smt3, but not as tight as SIM1. Also, they have some differences in the N value. The mutations on SIM3 did not affect the N value compared to Elg1 WT, but mutations on SIM2 have a relatively significant alteration in the N value, which decreased to 0.686±0.185. The change on the N values is likely caused by the possibility that alanine mutations on the SIMs did not fully abolish their binding abilities to Smt3. Subsequently, this interfered the analysis since we assume the otherwise. These results further indicate that all three SIMs contribute to the interaction with Smt3.

3.2.4 NMR

Although no interaction was detected in our GST pull-down assay, alternative research has revealed the binding between Smt3 and Elg1 via the yeast two-hybrid method [2]. Our ITC experiments also confirmed the existence of the interaction between the Elg1 N-terminus

and Smt3. However, the stoichiometry from ITC needs to be further investigated. Additionally, NMR experiments with titration can be employed to further validate protein interactions and provide clues about the binding model.

NMR was first used to calculate the apparent binding affinity. As described in the Materials and Methods section, the proteins used for this experiment were exchanged into the NMR buffer. The concentration of the protein samples was validated by amino acid analysis. Using 61 µM Smt3 as standard, we used NMR to measure the surface residues' chemical shifts while mixed with 42 µM, 84 µM, 168 µM, and 336 µM Elg1 WT. The signals were processed by NMRPipe and NMRviewJ. Since the Smt3 solution structure was determined [64], we were able to assign most of our residues based on the relative positions of the signal peaks. The 2D spectrum of Smt3 is shown in Figure 20, and those residues that exhibit large chemical shifts are shown in Figure 21. The residues showing relatively large shifts were used to calculate the apparent binding affinity (Table 6). We set the stoichiometry to 1:3 (Elg1: Smt3), which assumes one Elg1 molecule binds to three Smt3 with equal contributions of three SIMs, and the apparent binding affinity would be determined to approximately 150 µM (Figure 22). This binding affinity suggested that Smt3 binds to Elg1 weakly. The fitting curves in Figure 22 all have an R^2 of over 0.97. R^2 is known as the coefficient of determination, which is always between 0 and 1. In general, the higher the R^2, the better the model fits your data.

Furthermore, the chemical shifts data from NMR could be used to map the surface of Smt3 that is involved in the interaction between Elg1 and Smt3. Using the equation in the Materials and Methods section, we had minimized the artificial effects from the range difference between ^{1}H and ^{15}N. We applied these calculations to all assigned residues and discovered several residues on Smt3 that were shifted significantly with Elg1 binding. Shown in Figure 23, most residues that have large chemical shifts were located on the β sheet-1, β sheet-2, and α helix in the N-terminus of Smt3. Two Smt3 structures were utilized to map Smt3's interacting face. One is the structure of Smt3 binding to Srs2's SIM (PDB ID: 3V62) [26], those residues that were shifted with Elg1 binding (based on our NMR
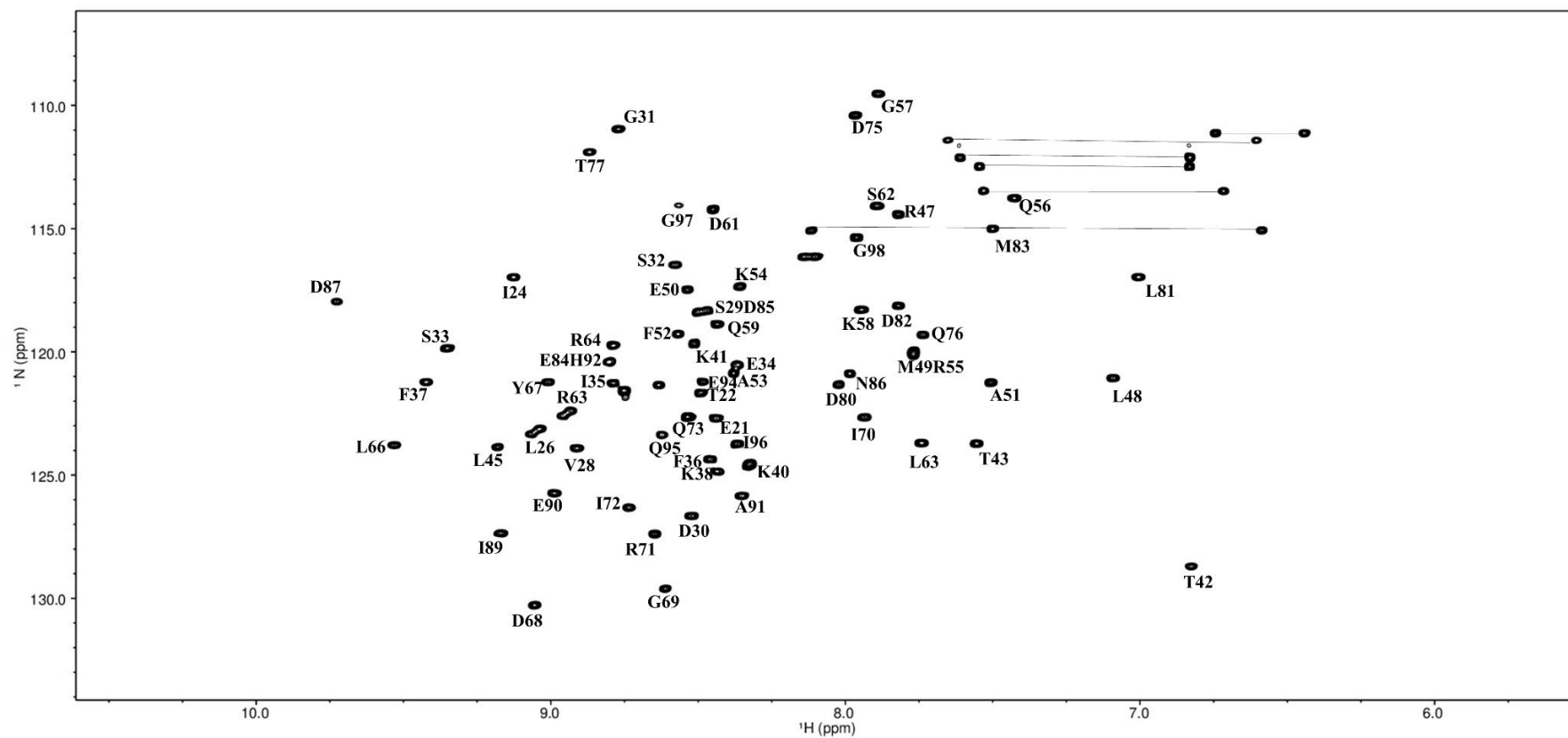
Figure 20: 2D ¹H-¹⁵N spectrum of Smt3 Δ18K19R with assignments.

This spectrum of Smt3 is obtained at 25°C and pH 7.5. Backbone amide resonances are labelled with the one-letter code and residue number in this 2D spectrum. (The spectrum also shows several additional peaks from asparagine, glutamine, and arginine side chain resonances, which are the peaks connected with solid lines.)
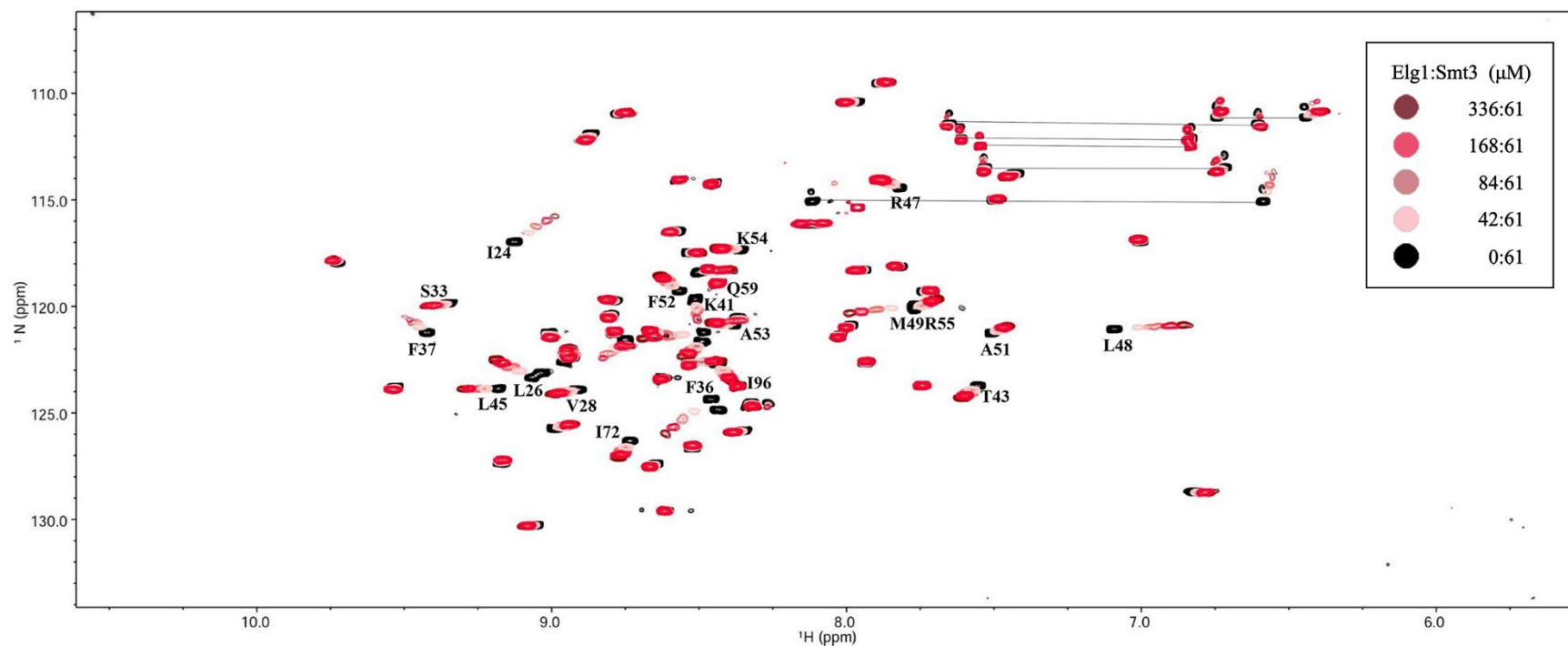
Figure 21: 2D ^{1}H-^{15}N spectrum of Smt3 Δ18K19R, free and in complex with various concentrations of Elg1 WT.
The spectrum of free Smt3 is shown in black, and the spectrums of the mixture with various ratios of concentration between Elg1 and Smt3 are indicated with different colors. This spectrum of Smt3 is obtained at 25°C and pH 7.5. Residues with relatively large chemical shifts (larger than average) are marked with the one-letter code and residue number. (The spectrum also shows several additional peaks from asparagine, glutamine, and arginine side chain resonances, which are the peaks connected with solid lines.)

A

B
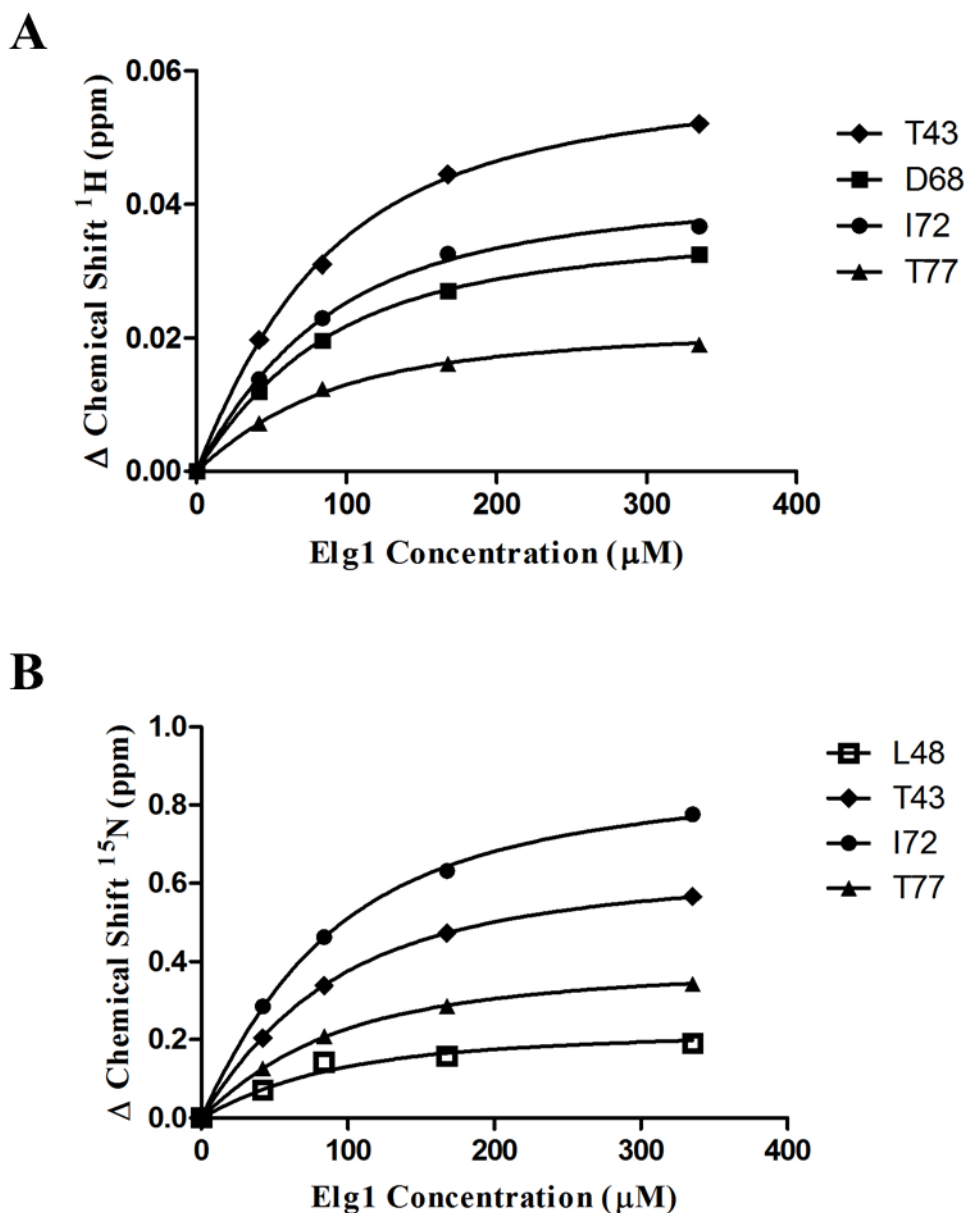
Figure 22: The fitting curves of chemical shifts used to calculate binding affinity.
The curves represent the fit of the equation to describe the situation that all three SIMs
bind to Smt3 identically and K_d was calculated in the same equations.
A) Based on the chemical shifts of ^1H, the curves correspond to $K_d=142.3\pm6.1$ µM.
B) Based on the chemical shifts of ^{15}N, the curves correspond to $K_d=153.4\pm7.8$ µM.

Table 6: Summary of chemical shifts for apparent K_d in NMR

Elg1 concentration	Smt3 concentration	Residues' chemical shifts in ^{1}H (ppm)			
(μM)	(μM)	T43	D68	I72	T77
42	61	0.020	0.012	0.014	0.007
84	61	0.031	0.020	0.023	0.012
168	61	0.044	0.027	0.033	0.016
336	61	0.052	0.033	0.037	0.019

Elg1 concentration	Smt3 concentration	Residues' chemical shifts in ^{15}N (ppm)			
(μM)	(μM)	T43	L48	I72	T77
42	61	0.20	0.07	0.29	0.13
84	61	0.34	0.14	0.46	0.21
168	61	0.47	0.16	0.63	0.29
336	61	0.57	0.19	0.78	0.34
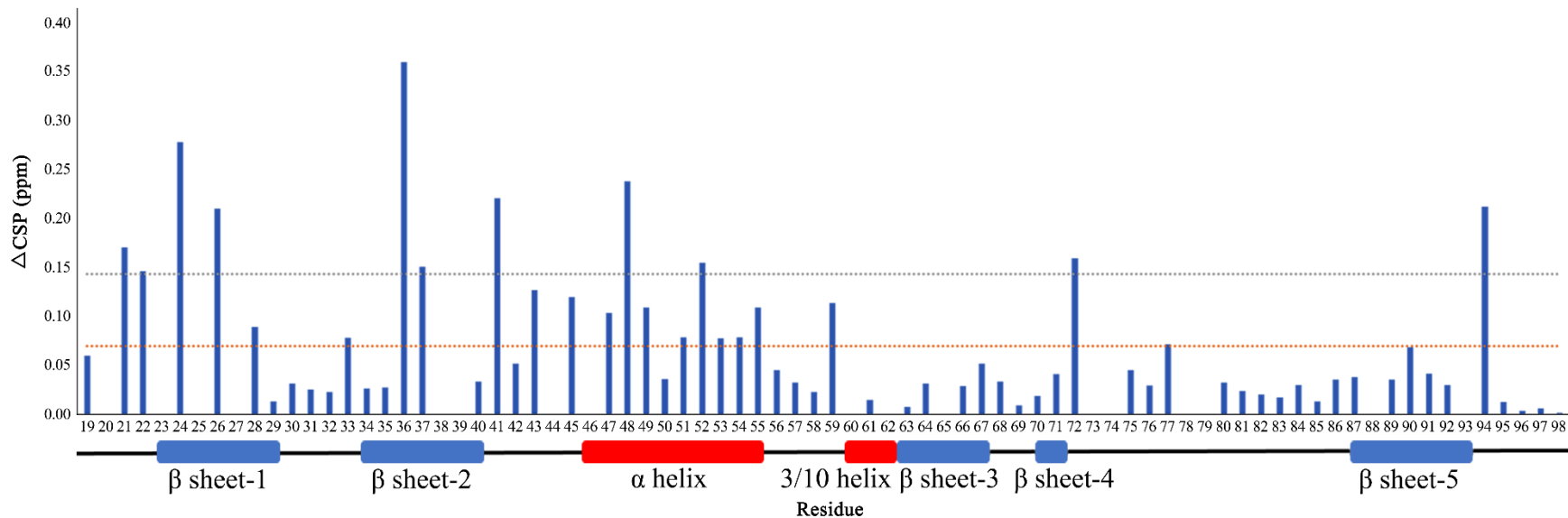
Figure 23: Chemical-shift perturbations of assigned residues in Smt3 Δ18K19R.
The chemical-shift perturbations are calculated with the equation described above and are plotted with their residue numbers. The secondary structure of Smt3 is also indicated in the bar diagram below. It shows most residues that have large chemical shifts are in the β sheet-1, β sheet-2, and α helix. The orange horizontal line represents the average overall chemical shifts (0.069 ppm), and the blue horizontal line represents the overall chemical shifts (0.143 ppm) that equal to the average overall chemical shifts plus one standard deviation.

A

B

C

D
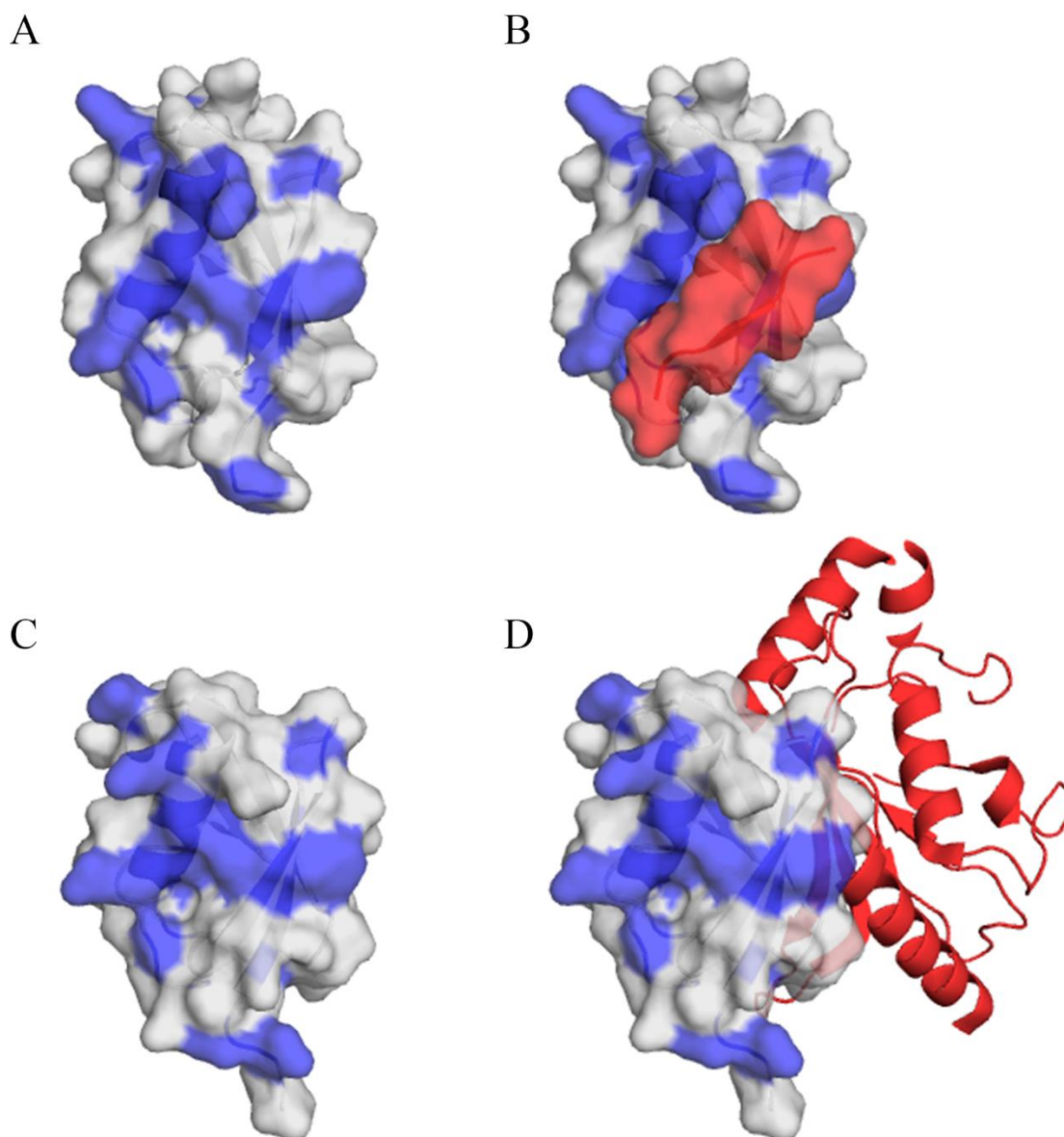
Figure 24: Elg1 binding surface mapped on the Smt3.

The Elg1 binding surface was mapped utilizing the Smt3 structure from PDB ID: 3V62 (panel A and B) and 2EKE (panel C and D). Smt3 structures are presented in gray. Residues with overall chemical shifts above the average based on our NMR data were mapped the Elg1-Smt3 interacting surface in blue.

A) The structure of Smt3 from 3V62. The Elg1-Smt3 interacting surface is mapped in blue.

B) The structure of Smt3 binding to Srs2 SIM. The surface of Srs2 SIM is presented in red. It shows that Elg1 binds to Smt3 in the same side as Srs2 binds to Smt3.

C) The structure of Smt3 from 2EKE. The Elg1-Smt3 interacting surface is mapped in blue.

D) The structure of Smt3 binding to yUbc9. The structure of yUbc9 is presented in red. It shows that Elg1 binds to Smt3 in the opposite side as yUbc9 binds to Smt3.

data) were mapped in blue (Figure 24A). Interestingly, it resembles the surface where Srs2 SIM binds (Figure 24B). Another structure we used is the structure of Smt3 binding to yUbc9 (PDB ID: 2EKE) [65]. Those shifted residues were also mapped in that structure (Figure 24C). We found out that the interface of Smt3 binding to yUbc9 is on the opposite face of where we mapped (Figure 24D). It is likely that Smt3 binds to Elg1 in a similar manner as Srs2, instead of yUbc9.

4 Discussion

4.1 Elg1's interaction with PCNA

4.1.1 Two PIP boxes participate in the interaction with PCNA

Despite the lack of a canonical PIP box sequence, Elg1 has been shown to bind to PCNA in the yeast-2-hybrid assay. Elg1 also preferentially binds to sPCNA [2]. In this work, two biochemical methods, GST pull-down assay and ITC, have been employed to assess the interaction between Elg1 and PCNA. In our hypothesis, we speculated that two PIP boxes are involved in the interaction with PCNA. The first one is very similar to the PIP box proposed by Parnas *et al.* [2], and the second one is based on our lab's previous investigations.

Parnas *et al.*'s analysis proposed the PIP box of Elg1 located between amino acids 43 and 67 using different Elg1 N-terminal fragments in yeast-2-hybrid assays [2]. Similarly, our results in the pull-down assays confirmed Elg1's binding ability with PCNA. In our GST pull-down assays, only Elg1L3 (46-84) and Elg1L3+ (33-84), the fragments containing two PIP boxes and excluding all three SIMs, demonstrate PCNA binding. Although the shorter fragment Elg1L5 (64-84) fails to show the interaction with PCNA, it does not exclude the weak binding of PIP2 and the possibility that two PIP boxes work together to facilitate the interaction with PCNA. In addition, we also used ITC for further investigations, where we titrated PCNA to Elg1L3+. But the ITC result fails to provide solid evidence for characterizing this interaction. There are two possibilities for this failure on ITC. One is

that the titration of the Elg1 fragment to PCNA is prone to aggregations. It would be similar to the case of titrating Elg1 fragment to sPCNA. Second is that the concentration we used may cause the Elg1 fragment to not fold properly. Although Elg1 fragment was expressed together with GST tag for better solubility, there remains the possibility that Elg1 fragment becomes less functional to the interaction with PCNA due to the improper folding at high concentration for titration. Collectively, our GST pull-down assays have confirmed Elg1 N-terminus's interaction with PCNA *in vitro*, and two PIP boxes participate in this interaction.

4.1.2 SIM1 and/or SIM2 may inhibit the interaction with PCNA

Elg1 contains three SIMs that were shown to interact with SUMO in yeast-2-hybrid assays [2]. Armstrong *et al.* [26] have demonstrated the PIP box's interaction with PCNA and SIM's interaction with SUMO are independent of each other in Srs2. But our GST pull-down assays suggest that the SIM(s) of Elg1 may interfere with the interaction with PCNA. With sPCNA as prey, multiple fragments, including Elg1L3 (46-84), Elg1N, and Elg1L1 (21-102), were detected to show binding with sPCNA. Elg1N contains all three SIMs, and Elg1L1 only contains SIM1 and SIM2, while Elg1L3 contains no SIMs. However, Elg1N and Elg1L1 did not show interactions with PCNA, while Elg1L3 showed binding with PCNA. We suspect that SIM1 and/or SIM2 inhibit the interaction with PCNA. When PCNA was SUMOylated, SIM1 and/or SIM2 bind to SUMO that is conjugated to PCNA and possibly changed the conformation of the Elg1 N-terminus, subsequently promoting the PIP box's interaction with PCNA.

In order to exclude the possible inhibitory effects of SIMs and test if the PIP box1 can bind to PCNA alone, we attempted to use synthesized peptide that contains original residues 50-66 of Elg1 for further investigations. Several past studies have determined the structure of the PCNA-peptide complex [26, 43]. We hoped that a higher concentration of peptides could allow for a more accurate analysis in the ITC and form the complex with PCNA for screening. Despite the prediction that the peptide would be soluble, this was not the case, especially for the aqueous buffers. We directly mixed the PCNA and the pH-adjusted

peptide solution, and the mixture was directly used to screen for crystals. Although several hits were observed, those crystals needed to be verified as protein-peptide complexes, instead of PCNA alone, and those conditions needed further optimization. Without a clear structure of the PCNA complex with the Elg1 PIP box, it is difficult to identify the PIP box residues that formed the hydrophobic plug that interacts with the IDCL of PCNA. Though our attempt to use the synthesized peptide did not work, other approaches, such as *in silico* alanine scanning, can be employed to derive more information.

4.2 Elg1's interaction with sPCNA

4.2.1 Elg1 N-terminus is confirmed to bind sPCNA with µM affinity

According to previous research [1-2], Elg1 would preferentially bind to sPCNA and this binding is mediated by the SIMs in the N-terminal. Our GST pull-down assays confirmed the binding with sPCNA, via using Elg1 46-84, Elg1 21-150, and Elg1 21-102 as bait (Figure 13). In addition, we determined the apparent K_d of interaction between the Elg1 N-terminus and sPCNA through ITC. Elg1 WT binds to sPCNA with an apparent K_d of 8.23±3.31 µM, and an N value of 1.122±0.073 (Table 4).

However, this binding affinity is much weaker than expected. Armstrong *et al.* [26] have established that the Srs2 C-terminal can bind to sPCNA with a K_d of approximately 25 nM via fluorescence polarization assays, which is a much tighter binding than what we observed. The reason for the large difference is probably because the PIP box of Elg1 binds to PCNA in a much weaker manner than the Srs2 C-terminus. We were able to determine that the Elg1 N-terminal binds to sPCNA in µM affinity with ITC, but further characterizations are required to uncover how the two protein-protein interactions, namely SIM binding to SUMO and the PIP box binding to PCNA, accommodated each other.

4.2.2 SIMs mutations have subtle effects on the interaction with sPCNA

Previous research has shown that all three SIMs are capable of interacting with Smt3 via yeast-2-hybrid assays [2]. From our GST pull-down results, SIM1 and SIM2 are needed for

the interaction with sPCNA. Deletion of SIM3 seems to have minor effects on the interaction with sPCNA. Moreover, ITC experiments with various SIMs mutants were conducted to shed some light on the mechanism of the binding with sPCNA. The results show that the interaction with sPCNA is not affected dramatically by mutations on SIMs. The apparent K_d decreased at most 2-fold with mutant Elg1 SIM23m. These results indicate that the interactions between SIMs and SUMO slightly contribute to Elg1's interaction with sPCNA. Since the mutations on all three SIMs only hinder the interaction with sPCNA to a small extent, we suspect that SIMs probably play a minor role in anchoring to PCNA. Only the mutation in the SIM1 exhibit a 1.5-fold decrease in the apparent binding affinity. However, with mutations on the double SIMs, their ΔS have significant differences, despite still having a similar level of apparent K_d. It implied that three SIMs in the Elg1 N-terminal may bind to sPCNA with different binding mechanisms. To be noticed, most ITC titrations have an N value of close to 1, except for the SIM12m. It is probably because this titration between sPCNA and SIM12m were conducted at last, which the protein may not be as effective as the fresh protein.

To discover the role of the SIMs in the interaction with sPCNA, various combinations of mutations on SIMs were generated and tested. Unexpectedly, SIMs mutations have subtle effects on the interaction with sPCNA.

4.3 Elg1's interactions with Smt3

4.3.1 Elg1 N-terminus is confirmed to bind Smt3 weakly

Parnas *et al.* have shown that the Elg1 N-terminus can bind Smt3 in yeast 2-hybrid experiments, and all three SIMs can also bind Smt3 [2]. To better characterize the interaction with Smt3, GST pull-down assays and ITC experiments were employed. However, our GST pull-down assays did not show a detectable interaction between the Elg1 N-terminus and Smt3.

Our ITC experiments determined that the Elg1 N-terminal binds to Smt3 with an apparent

K_d of approximately 80 μM (Table 5), while Elg1 binds to sPCNA with an apparent K_d of approximately 8 μM (Table 4). Additionally, our NMR experiments further confirmed that Elg1 binds to Smt3 weakly. As shown in Figure 21, some residues of Smt3 have chemical shift changes upon Elg1 binding. Those residues with large chemical shifts were used to calculate the K_d. With the assumption of an n value as 3.0 (n represents the number of binding sites on Elg1), the K_d is calculated to be approximately 150 μM (Figure 22). The reason why the K_d that measured via ITC and NMR are different is that ITC has large deviation when determining the K_d of low-affinity interaction. For low-affinity interaction, the inflections of ITC titrations are not well defined to accurately determine K_d. To be noted, the fitting strategy we used is not accurate, since all three SIMs would not be identical. Characterizing the interaction between the Smt3 and only one SIM would be the practical strategy for further investigation. The NMR experiments with peptides that only contain one SIM titrating to Elg1 would be suitable to better characterize this binding model.

4.3.2 All three SIMs contribute to the interaction with Smt3

In our ITC experiments, we were able to determine the apparent K_d of binding Smt3 is 80.5±34.3 μM (Table 5). This also gives an N value of 1.096±0.417, which support our assumption that one Elg1 binds three Smt3.

In order to understand the role of SIMs in the interaction with Smt3, various constructs that contain SIMs mutations were generated and the expressed proteins were utilized in the ITC. Also, as shown in Table 6, mutations on single SIMs dramatically affect the apparent K_d. Mutations on SIM1 have the largest impaired effects on the apparent K_d, which decreased 1.6-fold to approximately 132 μM. This result suggests that SIM1 may be the dominant SIM for the interaction with Smt3, likely the same case for interacting with sPCNA. Mutations on SIM2 and SIM3 maintain the similar apparent binding affinity, compared to the Elg1 WT, suggesting SIM2 and SIM3 bind to Smt3 in a manner weaker than SIM1. Collectively, all three SIMs contribute to the interaction with Smt3, and SIM1 probably is the major one that is responsible for the interaction.

4.3.3 Elg1's interaction with Smt3 is likely to resemble Srs2 interactions

Our NMR experiments not only provide us a clue about the apparent K_d, but also shed more light on the mechanism of how the SIM(s) of Elg1 bind(s) to Smt3. Based on the assignment of Smt3, the chemical shifts of most residues on the Elg1 N-terminus were calculated. Shown in Figure 23, most residues that have large chemical shifts were located on β sheet-1, β sheet-2, and the α helix in the Smt3 N-terminus. Based on the chemical shifts, we mapped the Elg1 binding surface using two known Smt3 structures, Srs2 bound by C-terminus and Smt3 in complex with yUbc9. Interestingly, the mapped surface resembles the one from Srs2's C-terminus (Figure 24B). In addition, it is unlikely that all three SIMs bind to one single Smt3. It is possible that all three SIMs play different roles in the interaction with Smt3 or sPCNA. Despite more questions having arisen during our investigations, determining the structure of Elg1 binding to Smt3 could provide insight into the SIM-SUMO interaction models. Approaches, such as NMR with Elg1 SIM mutants, can be employed to help determine the structure of Elg1's SIM binding to Smt3.

4.4 Proposed model for interactions with PCNA/sPCNA

Combining our results, we propose the following model for Elg1 and PCNA/sPCNA interaction (the schematic model is presented in Figure 25):

To maintain normal replication processes, Elg1 exhibits a weaker ability to bind PCNA than other PCNA partners, such as polymerases. SIM(s) may partially cover the PIP box on the Elg1 N-terminus and inhibit the interaction between Elg1 and PCNA. Upon DNA damage and/or other cofactors, PCNA undergoes SUMOylation to promote its binding with Elg1, which subsequently promoting Elg1-RLC's unloading function. Therefore, PCNA would not stay on the replication fork, and the replication process can proceed. With SUMOylation, SUMO is conjugated to PCNA, mostly at K164. The SIM(s) in the Elg1 N-terminal would bind to the Smt3, anchoring Elg1 and stabilizing Smt3's orientation, and the SIM(s) cease(s) to interfere with the interaction between the PIP box and PCNA, subsequently promoting the PCNA unloading activity.

4.5 Future directions

In this study, we have made progress to elucidate the interaction of the Elg1 N-terminus with PCNA, sPCNA, and Smt3. For the interaction with PCNA, we have shown that two PIP boxes participate in the PCNA binding. The binding is significantly weaker than that of most PIP boxes and PCNA interactions, and SIM(s) may play an inhibitory role in this interaction. Based on the results, we propose the model for Elg1's interactions with PCNA/sPCNA, which provides us with a better understanding about the role of Elg1 as a general unloader of PCNA. Although we currently cannot pinpoint the exact residues of the PIP box and determine the structural basis of Elg1-PCNA interaction, we have a clue on this interaction, which is probably from the inhibitory effect from SIM(s). Better construct design is required to investigate the interaction with PCNA, and experiments, such as alanine scanning and NMR, can be employed for further characterizations.

Upon DNA damage, PCNA would be SUMOylated. The feature that Elg1 preferentially binds to sPCNA implies a role of Elg1 in DNA damage. We determined that the interaction has an apparent K_d of approximately 8 μM. Mutations on SIMs have subtle effects on the apparent K_d. In addition, we have shown that all three SIMs in the Elg1 N-terminus contribute to the binding with Smt3. However, since one Smt3 can only accommodate one SIM, all three SIMs may contribute to different roles. These results indicate that SIMs play different roles in the interaction with sPCNA. Characterizing their different roles would help us understand how Elg1 regulates the PCNA unloading and how SUMOylation affects PCNA unloading. NMR experiments with peptide that contains only one SIM would better characterize the binding model between Elg1 and Smt3, and some *in vivo* work about the SIMs would be shed some light on Elg1's functions.

Although several aspects of this study require further investigations, progress has been made to understand the interactions between Elg1 and PCNA/sPCNA. Further investigations on these interactions will provide us with more insight into the PCNA unloading process and better understanding about the roles of Elg1 in DNA replication and damage response.
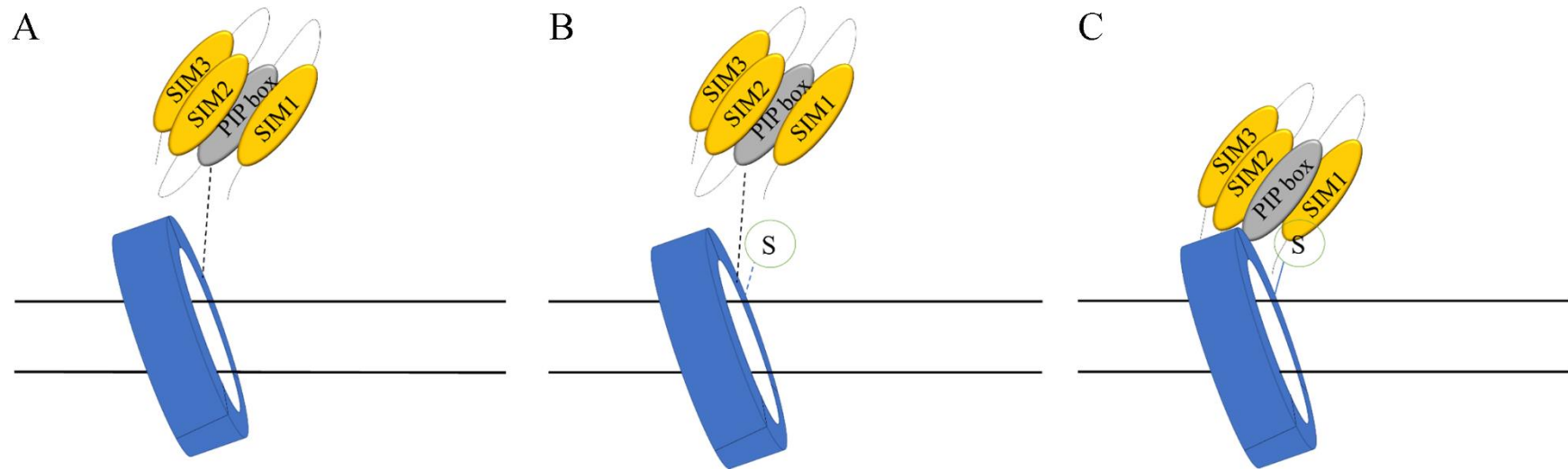
Figure 25: Proposed model of the interactions between Elg1 and PCNA/sPCNA.

Based on our results, we proposed a model to describe the interactions between Elg1 and PCNA/sPCNA. PCNA is shown as the blue ring. The motifs in the Elg1 N-terminus are colored as yellow for SIMs and gray for the PIP box. The S represents the SUMO.

A) During normal replication, the interaction between the PIP box of Elg1 and PCNA is inhibited by adjacent SIMs. The PIP box of the Elg1 binds to PCNA, which is represented by the black dashed line.

B) Upon DNA damage, PCNA would undergo SUMOylation and SUMO would be conjugated to PCNA at K164.

C) After PCNA is SUMOylated, SIM1 or SIM2 would bind to the SUMO, alleviating the inhibitory effects on the interaction between the PIP box and PCNA. The PIP box of Elg1 can bind to PCNA in a much stronger manner.

5 References

1. Bellaoui, M., Chang, M., Ou, J., Xu, H., Boone, C. and Brown, G. W. (2003) Elg1 forms an alternative RFC complex important for DNA replication and genome integrity. *EMBO J.* **22**, 4304-4313.

2. Parnas, O., Zipin-Roitman, A., Pfander, B., Liefshitz, B., Mazor, Y., Ben-Aroya, S., Jentsch, S. and Kupiec, M. (2010) Elg1, an alternative subunit of the RFC clamp loader, preferentially interacts with SUMOylated PCNA. *EMBO J.* **29**, 2611-2622.

3. Shiomi, Y. and Nishitani, H. (2013) Alternative replication factor C protein, Elg1, maintains chromosome stability by regulating PCNA levels on chromatin. *Genes Cells*. **18**, 946-959.

4. Kanellis, P., Agyei, R. and Durocher, D. (2003) Elg1 forms an alternative PCNA-interacting RFC complex required to maintain genome stability. *Curr Biol*. **13**, 1583-1595.

5. Kubota, T., Nishimura, K., Kanemaki, M. T. and Donaldson, A. D. (2013) The Elg1 replication factor C-like complex functions in PCNA unloading during DNA replication. *Mol Cell*. **50**, 273-280.

6. Moldovan, G. L., Pfander, B. and Jentsch, S. (2007) PCNA, the maestro of the replication fork. *Cell*. **129**, 665-679.

7. Shiomi, Y. and Nishitani, H. (2017) Control of genome integrity by RFC complexes; conductors of PCNA loading onto and unloading from chromatin during DNA replication. *Genes (Basel)*. **8**.

8. Boehm, E. M., Gildenberg, M. S. and Washington, M. T. (2016) The Many Roles of PCNA in Eukaryotic DNA Replication. *Enzymes*. **39**, 231-254.

9. Bruck, I. and O'Donnell, M. (2001) The ring-type polymerase sliding clamp family. *Genome Biol*. **2**, REVIEWS3001.

10. Bravo, R., Frank, R., Blundell, P. A. and Macdonald-Bravo, H. (1987) Cyclin/PCNA is the auxiliary protein of DNA polymerase-delta. *Nature*. **326**, 515-517.

11. Prelich, G., Tan, C. K., Kostura, M., Mathews, M. B., So, A. G., Downey, K. M. and Stillman, B. (1987) Functional identity of proliferating cell nuclear antigen and a DNA

polymerase-delta auxiliary protein. *Nature*. **326**, 517-520.

12. Johnson, C., Gali, V. K., Takahashi, T. S. and Kubota, T. (2016) PCNA Retention on DNA into G2/M Phase Causes Genome Instability in Cells Lacking Elg1. *Cell Rep*. **16**, 684-695.

13. Becker, J. R., Pons, C., Nguyen, H. D., Costanzo, M., Boone, C., Myers, C. L. and Bielinsky, A. K. (2015) Genetic Interactions Implicating Postreplicative Repair in Okazaki Fragment Processing. *PLoS Genet*. **11**, e1005659.

14. Wang, S. C. (2014) PCNA: a silent housekeeper or a potential therapeutic target? *Trends Pharmacol Sci*. **35**, 178-186.

15. Tan, Z., Wortman, M., Dillehay, K. L., Seibel, W. L., Evelyn, C. R., Smith, S. J., Malkas, L. H., Zheng, Y., Lu, S. and Dong, Z. (2012) Small-molecule targeting of proliferating cell nuclear antigen chromatin association inhibits tumor cell growth. *Mol Pharmacol*. **81**, 811-819.

16. Kubota, T., Katou, Y., Nakato, R., Shirahige, K. and Donaldson, A. D. (2015) Replication-Coupled PCNA Unloading by the Elg1 Complex Occurs Genome-wide and Requires Okazaki Fragment Ligation. *Cell Rep*. **12**, 774-787.

17. Kelch, B. A., Makino, D. L., O'Donnell, M. and Kuriyan, J. (2012) Clamp loader ATPases and the evolution of DNA replication machinery. *BMC Biol*. **10**, 34.

18. Neuwald, A. F., Aravind, L., Spouge, J. L. and Koonin, E. V. (1999) AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res*. **9**, 27-43.

19. Cullmann, G., Fien, K., Kobayashi, R. and Stillman, B. (1995) Characterization of the five replication factor C genes of Saccharomyces cerevisiae. *Mol Cell Biol*. **15**, 4661-4671.

20. Bermudez, V. P., Maniwa, Y., Tappin, I., Ozato, K., Yokomori, K. and Hurwitz, J. (2003) The alternative Ctf18-Dcc1-Ctf8-replication factor C complex required for sister chromatid cohesion loads proliferating cell nuclear antigen onto DNA. *Proc Natl Acad Sci U S A*. **100**, 10237-10242.

21. Parrilla-Castellar, E. R., Arlander, S. J. and Karnitz, L. (2004) Dial 9-1-1 for DNA damage: the Rad9-Hus1-Rad1 (9-1-1) clamp complex. *DNA Repair (Amst)*. **3**, 1009-

1014.

22. Murakami, T., Takano, R., Takeo, S., Taniguchi, R., Ogawa, K., Ohashi, E. and Tsurimoto, T. (2010) Stable interaction between the human proliferating cell nuclear antigen loader complex Ctf18-replication factor C (RFC) and DNA polymerase ϵ is mediated by the cohesion-specific subunits, Ctf18, Dcc1, and Ctf8. *J Biol Chem*. **285**, 34608-34615.

23. Lee, K. Y., Yang, K., Cohn, M. A., Sikdar, N., D'Andrea, A. D. and Myung, K. (2010) Human ELG1 regulates the level of ubiquitinated proliferating cell nuclear antigen (PCNA) through Its interactions with PCNA and USP1. *J Biol Chem*. **285**, 10362-10369.

24. Davidson, M. B. and Brown, G. W. (2008) The N- and C-termini of Elg1 contribute to the maintenance of genome stability. *DNA Repair (Amst)*. **7**, 1221-1232.

25. Hickey, C. M., Wilson, N. R. and Hochstrasser, M. (2012) Function and regulation of SUMO proteases. *Nat Rev Mol Cell Biol*. **13**, 755-766.

26. Armstrong, A. A., Mohideen, F. and Lima, C. D. (2012) Recognition of SUMO-modified PCNA requires tandem receptor motifs in Srs2. *Nature*. **483**, 59-63.

27. Geiss-Friedlander, R. and Melchior, F. (2007) Concepts in sumoylation: a decade on. *Nat Rev Mol Cell Biol*. **8**, 947-956.

28. Dohmen, R. J. (2004) SUMO protein modification. *Biochim Biophys Acta*. **1695**, 113-131.

29. Georgescu, R. E., Langston, L., Yao, N. Y., Yurieva, O., Zhang, D., Finkelstein, J., Agarwal, T. and O'Donnell, M. E. (2014) Mechanism of asymmetric polymerase assembly at the eukaryotic replication fork. *Nat Struct Mol Biol*. **21**, 664-670.

30. Beattie, T. R. and Bell, S. D. (2011) The role of the DNA sliding clamp in Okazaki fragment maturation in archaea and eukaryotes. *Biochem Soc Trans*. **39**, 70-76.

31. Jansen, J. G., Tsaalbi-Shtylik, A. and de Wind, N. (2015) Roles of mutagenic translesion synthesis in mammalian genome stability, health and disease. *DNA Repair (Amst)*. **29**, 56-64.

32. Hoege, C., Pfander, B., Moldovan, G. L., Pyrowolakis, G. and Jentsch, S. (2002) RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and

SUMO. *Nature*. **419**, 135-141.

33. Moldovan, G. L., Dejsuphong, D., Petalcorin, M. I., Hofmann, K., Takeda, S., Boulton, S. J. and D'Andrea, A. D. (2012) Inhibition of homologous recombination by the PCNA-interacting protein PARI. *Mol Cell*. **45**, 75-86.

34. Ortega, J., Li, J. Y., Lee, S., Tong, D., Gu, L. and Li, G. M. (2015) Phosphorylation of PCNA by EGFR inhibits mismatch repair and promotes misincorporation during DNA synthesis. *Proc Natl Acad Sci U S A*. **112**, 5667-5672.

35. Moggs, J. G., Grandi, P., Quivy, J. P., Jonsson, Z. O., Hubscher, U., Becker, P. B. and Almouzni, G. (2000) A CAF-1-PCNA-mediated chromatin assembly pathway triggered by sensing DNA damage. *Mol Cell Biol*. **20**, 1206-1218.

36. Choe, K. N. and Moldovan, G. L. (2017) Forging Ahead through Darkness: PCNA, Still the Principal Conductor at the Replication Fork. *Mol Cell*. **65**, 380-392.

37. Bauer, G. A. and Burgers, P. M. (1990) Molecular cloning, structure and expression of the yeast proliferating cell nuclear antigen gene. *Nucleic Acids Res*. **18**, 261-265.

38. Krishna, T. S., Kong, X. P., Gary, S., Burgers, P. M. and Kuriyan, J. (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell*. **79**, 1233-1243.

39. Dieckman, L. M., Freudenthal, B. D. and Washington, M. T. (2012) PCNA structure and function: insights from structures of PCNA complexes and post-translationally modified PCNA. *Subcell Biochem*. **62**, 281-299.

40. Gazy, I. and Kupiec, M. (2012) The importance of being modified: PCNA modification and DNA damage response. *Cell Cycle*. **11**, 2620-2623.

41. Matunis, M. J. (2002) On the road to repair: PCNA encounters SUMO and ubiquitin modifications. *Mol Cell*. **10**, 441-442.

42. Bruning, J. B. and Shamoo, Y. (2004) Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1. *Structure*. **12**, 2209-2219.

43. Vijayakumar, S., Chapados, B. R., Schmidt, K. H., Kolodner, R. D., Tainer, J. A. and Tomkinson, A. E. (2007) The C-terminal domain of yeast PCNA is required for physical and functional interactions with Cdc9 DNA ligase. *Nucleic Acids Res*. **35**,

1624-1637.

44. Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M. and Kuriyan, J. (1996) Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell*. **87**, 297-306.

45. Hoffmann, S., Smedegaard, S., Nakamura, K., Mortuza, G. B., Raschle, M., Ibanez de Opakua, A., Oka, Y., Feng, Y., Blanco, F. J., Mann, M., Montoya, G., Groth, A., Bekker-Jensen, S. and Mailand, N. (2016) TRAIP is a PCNA-binding ubiquitin ligase that protects genome stability after replication stress. *J Cell Biol*. **212**, 63-75.

46. Masuda, Y., Kanao, R., Kaji, K., Ohmori, H., Hanaoka, F. and Masutani, C. (2015) Different types of interaction between PCNA and PIP boxes contribute to distinct cellular functions of Y-family DNA polymerases. *Nucleic Acids Res*. **43**, 7898-7910.

47. Xu, H., Zhang, P., Liu, L. and Lee, M. Y. (2001) A novel PCNA-binding motif identified by the panning of a random peptide display library. *Biochemistry*. **40**, 4512-4520.

48. Hoege, C., Pfander, B., Moldovan, G. L., Pyrowolakis, G. and Jentsch, S. (2002) RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature*. **419**, 135-141.

49. Lee, K. Y. and Myung, K. (2008) PCNA modifications for regulation of post-replication repair pathways. *Mol Cells*. **26**, 5-11.

50. Stelter, P. and Ulrich, H. D. (2003) Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. *Nature*. **425**, 188-191.

51. Wang, S. C., Nakajima, Y., Yu, Y. L., Xia, W., Chen, C. T., Yang, C. C., McIntush, E. W., Li, L. Y., Hawke, D. H., Kobayashi, R. and Hung, M. C. (2006) Tyrosine phosphorylation controls PCNA function through protein stability. *Nat Cell Biol*. **8**, 1359-1368.

52. Yu, Y., Cai, J. P., Tu, B., Wu, L., Zhao, Y., Liu, X., Li, L., McNutt, M. A., Feng, J., He, Q., Yang, Y., Wang, H., Sekiguchi, M. and Zhu, W. G. (2009) Proliferating cell nuclear antigen is protected from degradation by forming a complex with MutT Homolog2. *J Biol Chem*. **284**, 19310-19320.

53. Chiu, R. K., Brun, J., Ramaekers, C., Theys, J., Weng, L., Lambin, P., Gray, D. A. and

Wouters, B. G. (2006) Lysine 63-polyubiquitination guards against translesion synthesis-induced mutations. *PLoS Genet*. **2**, e116.

54. Leach, C. A. and Michael, W. M. (2005) Ubiquitin/SUMO modification of PCNA promotes replication fork progression in Xenopus laevis egg extracts. *J Cell Biol*. **171**, 947-954.

55. Qiu, Y., Antony, E., Doganay, S., Koh, H. R., Lohman, T. M. and Myong, S. (2013) Srs2 prevents Rad51 filament formation by repetitive motion on DNA. *Nat Commun*. **4**, 2281.

56. Donnelly, M. I., Zhou, M., Millard, C. S., Clancy, S., Stols, L., Eschenfeldt, W. H., Collart, F. R. and Joachimiak, A. (2006) An expression vector tailored for large-scale, high-throughput purification of recombinant proteins. *Protein Expr Purif*. **47**, 446-454.

57. Eschenfeldt, W. H., Lucy, S., Millard, C. S., Joachimiak, A. and Mark, I. D. (2009) A family of LIC vectors for high-throughput cloning and purification of proteins. *Methods Mol Biol*. **498**, 105-115.

58. Roberts, G. A., Stephanou, A. S., Kanwar, N., Dawson, A., Cooper, L. P., Chen, K., Nutley, M., Cooper, A., Blakely, G. W. and Dryden, D. T. (2012) Exploring the DNA mimicry of the Ocr protein of phage T7. *Nucleic Acids Res*. **40**, 8129-8143.

59. Klock, H. E. and Lesley, S. A. (2009) The Polymerase Incomplete Primer Extension (PIPE) method applied to high-throughput cloning and site-directed mutagenesis. *Methods Mol Biol*. **498**, 91-103.

60. Klock, H. E., Koesema, E. J., Knuth, M. W. and Lesley, S. A. (2008) Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscreening to accelerate structural genomics efforts. *Proteins*. **71**, 982-994.

61. Yunus, A. A. and Lima, C. D. (2009) Purification of SUMO conjugating enzymes and kinetic analysis of substrate conjugation. *Methods Mol Biol*. **497**, 167-186.

62. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*. **6**, 277-293.

63. Johnson, B. A. and Blevins, R. A. (1994) NMR View: A computer program for the

visualization and analysis of NMR data. *J Biomol NMR*. **4**, 603-614.

64. Sheng, W. and Liao, X. (2002) Solution structure of a yeast ubiquitin-like protein Smt3: the role of structurally less defined sequences in protein-protein recognitions. *Protein Sci*. **11**, 1482-1491.

65. Duda, D. M., van Waardenburg, R. C., Borg, L. A., McGarity, S., Nourse, A., Waddell, M. B., Bjornsti, M. A. and Schulman, B. A. (2007) Structure of a SUMO-binding-motif mimic bound to Smt3p-Ubc9p: conservation of a non-covalent ubiquitin-like protein-E2 complex as a platform for selective interactions within a SUMO pathway. *J Mol Biol*. **369**, 619-630.

6 Appendices

Appendix A: Table of primers that were used to generate new constructs.

For Mutations	Sequence
I26A, I28A (SIM1)	5'- gccaccgcagatgacgaaaatgatactgaatcaggcac -3'
	5'- gtcatctgcggtggctgtagagcatcttgccaggattg -3'
L92A, I93A, V94A, I95A (SIM2)	5'- gcagcggcagccagtgataagagtcccaaaagtgagactaattg -3'
	5'- actggctgccgctgcatcatcatcgtcgtcatcgc -3'
I118A, I120A, I121A (SIM3)	5'- gcgtctgcagcctccacatcgagaatcaaatcatcgcttc -3'
	5'- ggaggctgcagacgcatcatcttcatgctcctgcg -3'
6 histidines were directly linked to Smt3	5'- catcatcattcccgccctgagactcacatcaatttaaaggtgtcc -3'
	5'- agggcgggaatgatgatgatgatggtgcatatgtatatctccttc -3'
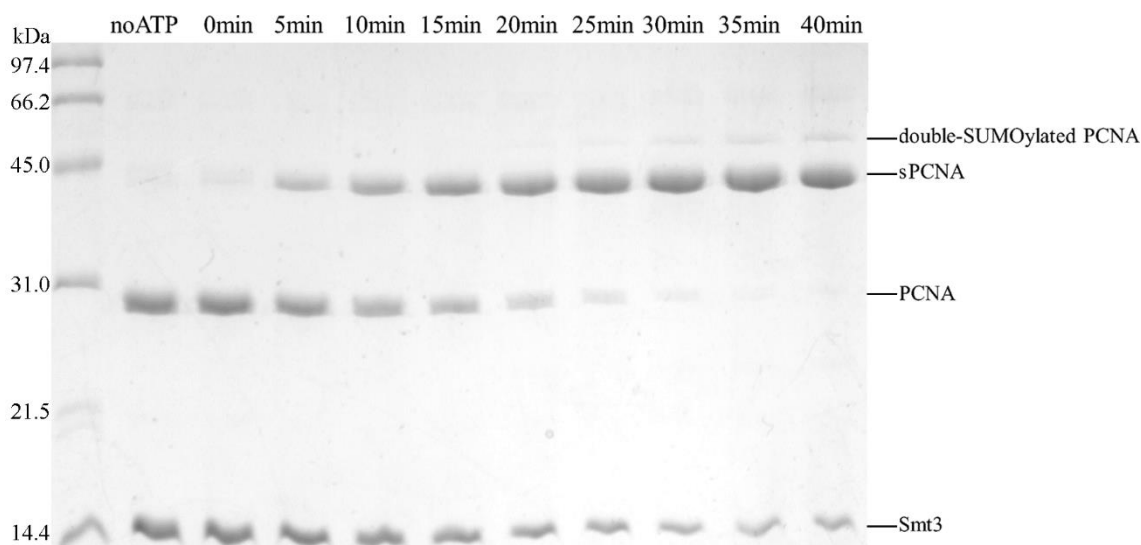
Appendix B: SUMOylation tested at different time points.
The reaction took place in a 1 ml system, and 10 µl of sample were taken at each time point for SDS-PAGE analysis. The lane labeled "noATP" represent the reaction sample that has not added ATP to initiate the reaction. The following lanes represent the situation of the reaction samples after various time points. The amount of PCNA and Smt3 gradually decreased, while more sPCNA was produced. The positions of double-SUMOylated PCNA, sPCNA, PCNA, and Smt3 are indicated on the right.
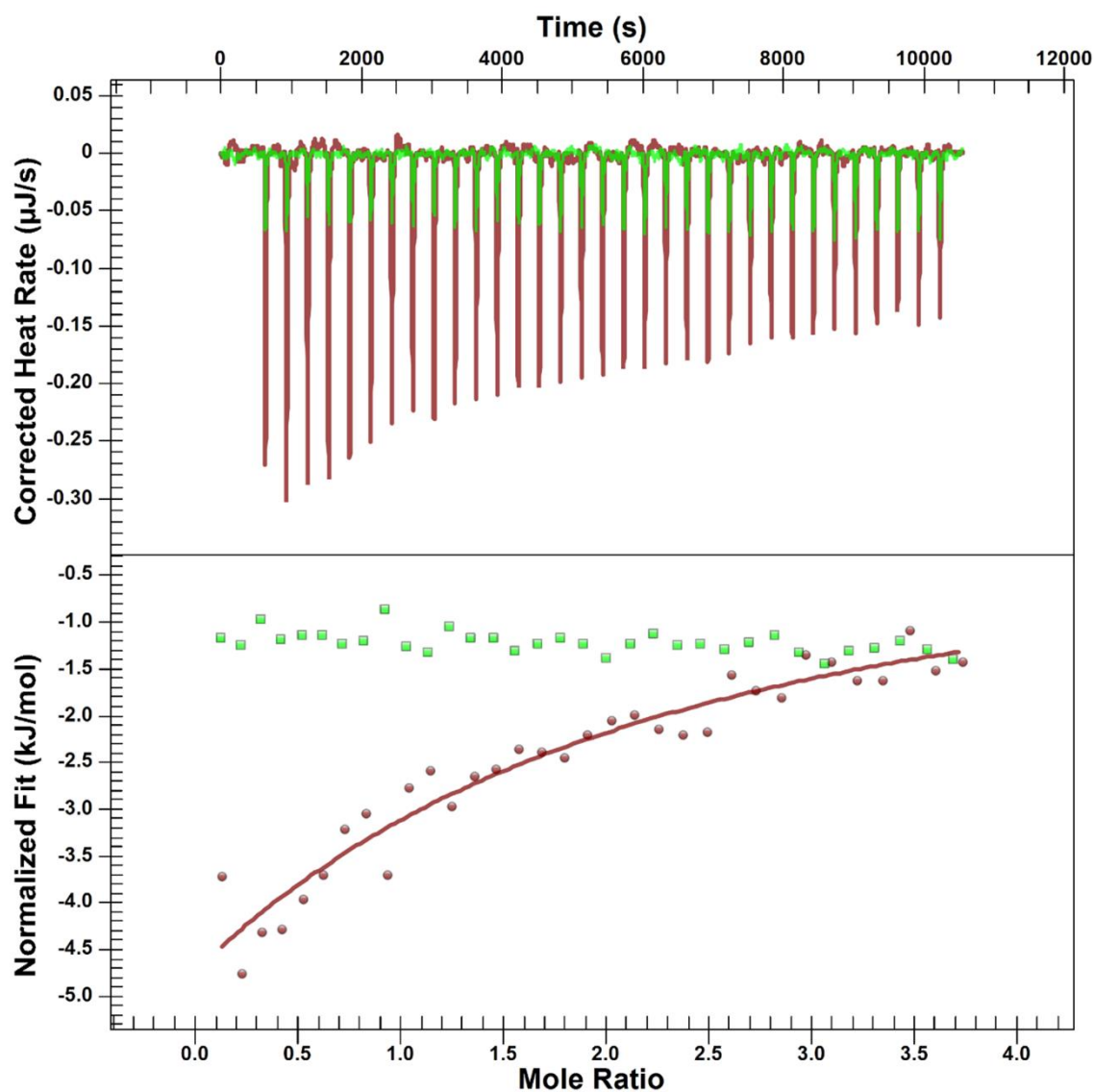
Appendix C: ITC titration of Elg1 and PCNA.

Protein 830 μM PCNA were titrated into 79 μM GST Elg1L3+ (33-84) at 25°C with a stirring speed of 250 rpm. This titration presented as a situation of no saturation after full injection, which makes the analysis inaccurate. This titration was analyzed by the program with an apparent K_d of over 424 μM and N of 0.775±1.206.

Curriculum Vitae

Name:	Jingwei Yan

Post-secondary Education and Degrees:

Central South University
Changsha, Hunan, China
2009-2013
Bachelor of Science in Life Science

University of Western Ontario
London, Ontario, Canada
2015-2017
Master of Science

Experiences and Honours:

Leader on the project *"the role of RTN3 in endothelial cells injury induced by reactive oxygen species and its mechanism"*, granted by University's Undergraduate Training Programs for Innovation and Entrepreneurship (2010-2013)

Honorable Award
Second National Colleges of Basic Medicine Innovation Experiment Design Competition (2012)

Western Graduate Research Scholarship (2015-2016)

Publication:

Journal article:
XIANG Rong, **YAN Jingwei,** FAN Liangliang, *et al*. Interactions between inflammation and endoplasmic reticulum stress in atherosclerosis[J]. Basic and Clinical Medicine, 253-256, 34(2), 2014.

Conference paper:
XIANG Rong, **YAN Jingwei,** CAO Beibei, *et al*. Reticulon 3 is involved in the endoplasmic reticulum stress of endothelial cells induced by lipopolysaccharide. The conference abstract collection of Genetics Society of Hunan Province, 34-35, Xiang Tan, Hunan, 2012.