# Investigating the Effects of Sumoylation on the Yeast General Transcription Factor TFIIF

**Yimo Dou**

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF

**Master of Science**

GRADUATE PROGRAM IN BIOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO
December 2018

## Abstract

Sumoylation is a post-translational modification that regulates numerous cellular processes including gene expression. We find that among the general transcription factors of RNA Polymerase II (RNAPII) in budding yeast, transcription factor IIF (TFIIF) shows a relatively high level of sumoylation, primarily on Lysine residues (Lys) 60/61 of its Tfg1 subunit. Fractionation analysis indicates that sumoylated Tfg1 is associated with chromatin, but its sumoylation level is not affected by stresses. Although Lys 60/61 were previously implicated in the Tfg1-RNAPII interaction, a Lysine-to-Arginine K60,61R mutation, which dramatically reduces Tfg1 sumoylation, does not affect that interaction. Chromatin immunoprecipitation analyses did not conclusively determine whether reducing Tfg1 sumoylation affects the recruitment of TFIIF or RNAPII to target gene promoters, but the expression of some genes is elevated in yeast expressing the K60,61R mutation. Our results demonstrate that chromatin-associated TFIIF is constitutively sumoylated which might be necessary for controlling the expression levels of some genes.

## Acknowledgements

First and foremost, I would like to express my heartfelt thanks to Dr. Emanuel Rosonina, for giving me the great opportunity of being a graduate student in my favorite field. As my supervisor, he showed not only his unparalleled expertise but also the extreme kindness and patience. I was encouraged to ask questions no matter how simple they were. His understanding and advices guided and motivated me to go further academically. One important thing he taught me was that a failed experimental result does not mean a failure of research. Without his help, I would not be able to regain confidence after setbacks, and this thesis would never have a chance to be accomplished. Moreover, as an international student who may require additional help, I received so much of his understanding and care. I also appreciate the help from my advisor, Dr. Peter Cheung, for giving me valuable tips and comments throughout my graduate study.

My gratitude should also be given to all the members of the Rosonina Lab, past and present, for their encouragements and warm-hearted help. I would like to thank: Akhi (for answering all my questions, no matter how naïve they were, and sharing life with me), Veroni (for being my patient mentor in the lab as well as one of my best friends with whom I always like to chat), John (for his erudition and humor), Marjan, Jessica, Areeb, Giovanni, Ben and Bryan. I would also like to thank all other friends from the 3rd and 4th floor for making me a home-like LSB.

A special thanks to Dr. Peter Cheung's Lab for allowing me to use their lab materials as well as providing me so much helpful directions in mammalian experiments, and Dr. Mark Bayfield's Lab for allowing me to utilize their equipment. My thanks should also be expressed to my thesis committee members Dr. Katalin Hudak and Dr. Gerald Audette for their patience and support.

# Table of Contents

## List of Tables

## List of Figures

## List of Abbreviations

| | |
|---|---|
| % w/v | Percent weight/volume |
| % (v/v) | Millilitre of solute/100 millilitre of solution |
| °C | Degree Celsius |
| 5-FOA | 5-Fluoroorotic Acid |
| ADH1 | Alcohol dehydrogenase 1 |
| Arg | Arginine |
| ATP | Adenosine Tri-Phosphate |
| Bp/bp | Base-pair or base-pairs |
| BPS | Branch point site |
| C | Cysteine |
| cDNA | Complementary deoxyribonucleic acid |
| ChIP | Chromatin immunoprecipitation |
| Chr V | Chromosome V |
| Chrom. | Chromatin-associated fraction |
| CTD | Carboxy-terminal domain |
| C-terminus/C-ter | Carboxy-terminus |
| D | Aspartic acid |
| DEPC | Diethyl pyrocarbonate |
| $dH_2O$ | Distilled water |
| DMEM | Dulbecco's Modified Eagle's Medium |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxyribonucleotide triphosphate |
| DTT | Dithiothreitol |

| | |
|---|---|
| E | Glutamic acid |
| EDTA | Ethylenediaminetetraacetic acid |
| EGTA | Ethylene glycol tetraacetic acid |
| EtOH | Ethanol |
| FBS | Fetal Bovine Serum |
| FWD | Forward PCR primer |
| g | Gram |
| *g* (italic) | x *g*-force (the number of times the gravitational force) |
| GAPDH | Glyceraldehyde 3-phosphate dehydrogenase |
| GAPDH | Glyceraldehyde 3-phosphate dehydrogenase |
| gDNA | Genomic deoxyribonucleic acid |
| GG | Di-glycine |
| GTF | General transcription factor |
| $H_2O_2$ | Hydrogen peroxide |
| HA | Human influenza hemagglutinin |
| HCl | Hydrochloric acid |
| HEPES | 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid |
| His | Histidine |
| hnRNP | Heterogeneous nuclear ribonucleoprotein |
| hrs | Hours |
| hs | Heat shock |
| HSP12 | Lipid-binding protein HSP12 |
| IB | Immunoblot |
| IgG | Immunoglobulin G |
| IMAC | Immobilized metal affinity chromatography |

| | |
|---|---|
| IP | Immunoprecipitation |
| kbp | Kilobase-pairs |
| KCl | Potassium chloride |
| kDa | Kilodalton |
| Kl::TRP1 | *K. lactis* TRP1 marker gene |
| LB | Luria broth |
| Lys/K | Lysine |
| M | Molar |
| min | Minute or minutes |
| mL | Millilitre |
| mM | Millimolar |
| mRNA | Messenger ribonucleic acid |
| MS | Mass spectrometry |
| MT | Mutant |
| NaCl | Sodium chloride |
| NEM | N-ethylmaleimide |
| Ni-NTA | Nickel-nitrilotriacetic acid |
| NP40 | Tergitol-type NP-40 |
| N-terminus/N-ter | Amine-terminus |
| O.D. | Optical density |
| ODCase | Orotidine 5'-phosphate decarboxylase |
| ORF | Open reading frame |
| PBS | Phosphate buffer saline |
| PBST | Phosphate buffer saline with Tween |
| PCR | Polymerase chain reaction |

| | |
|---|---|
| PEG | Polyethylene glycol |
| PHO5 | Acid phosphatase PHO5 |
| PIC | Pre-initiation complex |
| PIPES | Piperazine-N,N'-bis (2-ethanesulfonic acid) |
| PMA1 | Plasma membrane ATPase 1 |
| PMSF | Phenylmethylsulfonyl fluoride |
| PTM | Post-translational modification |
| PYK1 | Pyruvate kinase 1 |
| qPCR | Quantitative polymerase chain reaction |
| REV | Reverse PCR primer |
| RNA | Ribonucleic acid |
| RNAPII | RNA polymerase II |
| RPM | Revolutions per minute |
| rRNA | Ribosomal ribonucleic acid |
| RT-qPCR | Reverse transcription quantitative polymerase chain reaction |
| s | Second or seconds |
| SAE | SUMO-activating enzyme |
| SAH1 | S-adenosyl-l-homocysteine hydrolase 1 |
| SC/Sc | Synthetic Complete |
| SDS-PAGE | Sodium dodecyl sulphate - Polyacrylamide gel electrophoresis |
| SENP | SENtrin-specific proteases |
| Seq | Sequencing |
| SF3B1 | Splicing factor 3B subunit 1 |
| SM | Sulfometuron methyl |
| snRNP | Small nuclear ribonucleoprotein particle |

| | |
|---|---|
| Solub. | Soluble fraction |
| SR | Serine/arginine-rich |
| SS | Splice site |
| SSA4 | Hsp70 family chaperone SSA4 |
| SS-carrier DNA | Salmon sperm-carrier deoxyribonucleic acid |
| STL1 | Sugar transporter-like protein 1 |
| SUMO | Small ubiquitin-like modifier |
| TAF | TATA-binding protein-associated factors |
| TBP | TATA-binding protein |
| TF | Transcription factor |
| Tris | Trisaminomethane |
| Trp | Tryptophan |
| TSS | Transcription start site |
| URA | Uracil |
| WCE | Whole cell extract |
| WH | Winged helix |
| WT | Wild-type |
| x | Unspecified amino acid |
| X (Capitalized) | Times |
| YPD | Yeast extract peptone dextrose |
| µg | Microgram |
| µL | Microlitre |
| µM | Micromolar |

# Chapter 1: Introduction

## 1.1 RNA Polymerase II-mediated Transcription

As the central machinery involved in the expression of all eukaryotic protein-coding genes, RNA polymerase II (RNAPII) is important for all aspects of cell function, including regulating cell differentiation, maintaining cell identity as well as establishing the cellular response to environmental changes. Prior to initiating synthesis of RNA transcripts, RNAPII assembles on target promoters with multiple general transcription factors (GTFs), including TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH, to form a structure known as pre-initiation complex (PIC) (1). In mammalian cells, to begin the PIC assembly, TFIID, which contains the TATA box-binding protein (TBP) and multiple TBP-associated factors (TAFs), is required for recognizing the promoter region. In particular, the sequence-specific binding of TBP to the TATA box promoter element occurs ~30bp upstream of the transcription start site (TSS) (1). Although the consensus TATA sequence is not observed in most promoter regions in budding yeast, TBP can still bind to these promoters, which are located around 40-120bp upstream of the TSS in that organism (2). Therefore, this binding is considered functionally conserved from yeast to human. During the next step of PIC assembly, the binding of TFIIB and TFIIA to TBP results in the recruitment of RNAPII-TFIIF complex (1) (**Fig 1**). As a critical GTF, TFIIB plays a role in both RNAPII recognition and TSS selection (3). This promotes the formation of the core initiation complex which shares high similarity between human PIC and yeast PIC. The later involvement of TFIIE and TFIIH not only completes the PIC assembly but also contributes to the opening of the promoter region as well as to the stabilization of the melted DNA strands (1, 3).

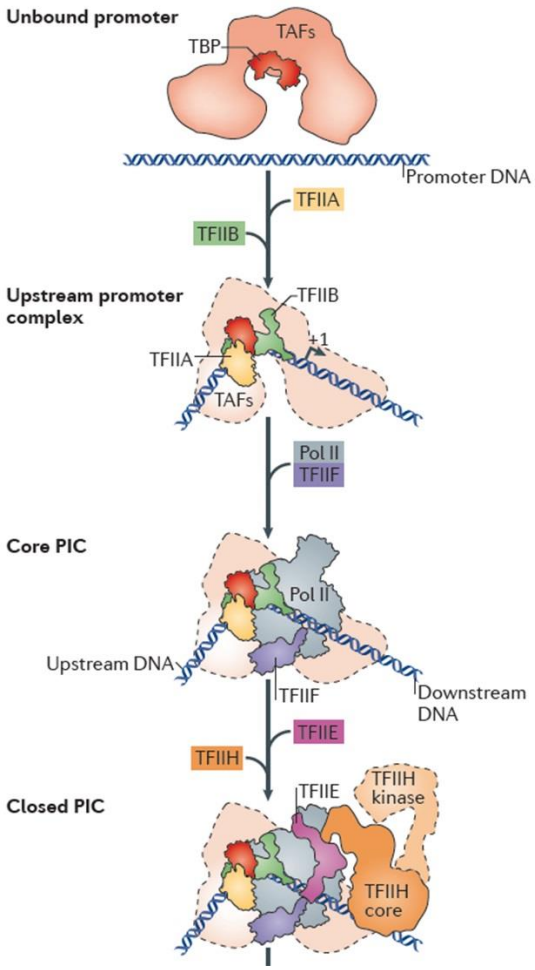



**Figure 1: Schematic of pre-initiation complex (PIC).** A canonical model for the early steps of PIC assembly from general transcription factors (TFIID, TFIIB, TFIIA, TFIIF, TFIIE and TFIIH; various colors) and RNA Polymerase II (Pol II; grey) on promoter DNA. The names for the intermediate complexes are shown on the left. Unbound promoter DNA is recognized by TATA box-binding protein (TBP) subunits of TFIID, inducing a bend of DNA. The recruitment of TFIIB and TFIIA stabilizes the TBP-DNA complex. The resulting upstream promoter complex and Pol II-TFIIF complex join together forming the core PIC. The later binding of TFIIE and TFIIH finish the PIC assembly. Reproduced from (1).

### 1.2 TFIIF Subunit Tfg1

Around 50% of all RNAPII is found in a complex with TFIIF in yeast (4). The functions of TFIIF include preventing the non-specific binding of RNAPII to DNA, stabilizing TFIIB binding within the PIC and TSS selection (1, 3). What's more, TFIIF also plays a role in stabilizing the RNA-DNA hybrid and stimulating

phosphodiester bond formation when the initial transcribing complex forms. During transcription elongation, TFIIF also has multiple functions, such as shortening RNAPII pausing time (5). In yeast, TFIIF contains three subunits. Tfg1, the largest subunit (82.2 kDa) of TFIIF, and Tfg2 are conserved from human to yeast. Their human counterparts are Rap74 and Rap30, respectively. The nonconserved subunit Tfg3 is also involved in several other gene-regulating complexes (6).

### 1.3 The Interaction between Tfg1 and RNAPII

The N-terminal regions of Tfg1 and Tfg2 form a dimerization domain while winged helix (WH) domains are structured at the C-terminal regions of both subunits when binding to RNAPII (5). Other than these structural similarities, the architectural difference is also observed. Tfg1 has a charged region between its dimerization domain and WH domain, whereas Tfg2 only has a linker region between the two domains (5) (**Fig 2**). Further cross-linking analysis reveals that Tfg1 is involved in both intra- and inter-protein interactions. Not only Tfg2, but also several RNAP II subunits, such as Rpb1, Rpb2 and Rpb9, can interact with Tfg1 (5) (**Fig 3**). For example, both Tfg1 Lys 60-Rpb2 Lys 606 and Tfg1 Lys 61-Rpb2 Lys 606 are cross-linking pairs with high confidence (**Appendix Table 1**).

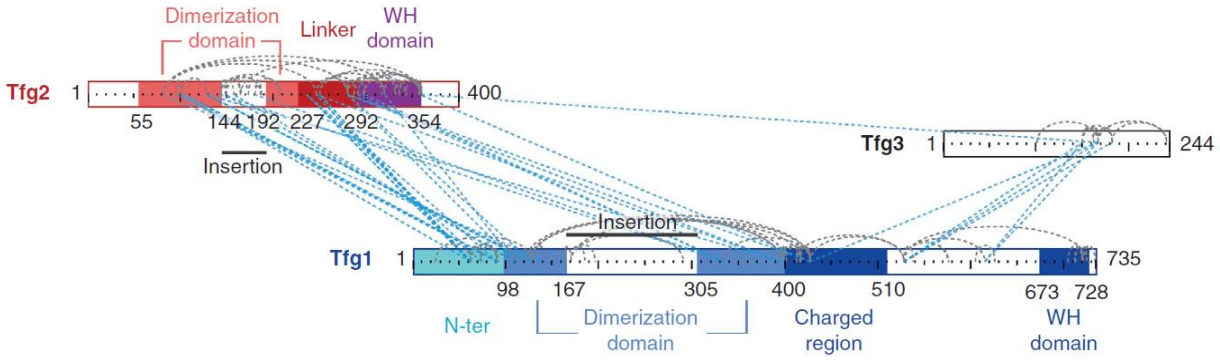



**Figure 2: Schematic of TFIIF subunits and domains.** TFIIF contains three subunits. Tfg1, the largest subunit in TFIIF, forms a dimer with Tfg2 via the interaction of dimerization domains of these two subunits. Both Tfg1 and Tfg2 have a C-terminal winged helix (WH) domain. A charged region is found in Tfg1, and a linker region is present in Tfg2. Dashed lines show cross-linked residues as determined from a cross-linking and mass spectrometry analysis of the complex. Reproduced from (5).

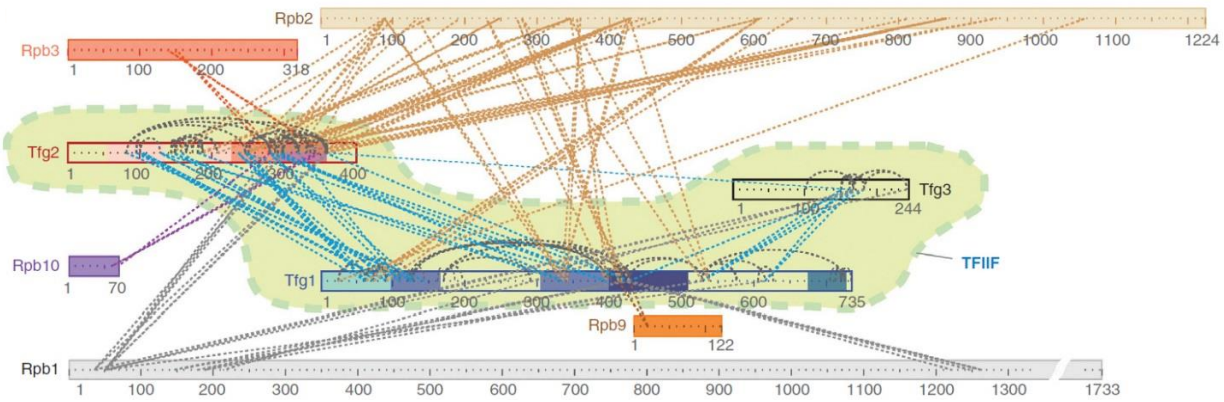



**Figure 3: Cross-linking analysis of the complete TFIIF-RNA Polymerase II (RNAPII) complex.** Map illustrates links within the TFIIF-RNAPII (dashed lines) identified by cross-linking and mass spectrometry analysis. Links from TFIIF to PNAPII are colour coded by the respective RNAPII subunits. Links within TFIIF subunits (grey) and between TFIIF subunits (blue) are also shown. Reproduced from (5).

### 1.4 Post-translational Modifications

Cellular processes are very dynamic and rely on appropriate protein activities (7). Changing of protein activities may affect the protein-protein interactions as well as the interactions between proteins and other molecules including carbohydrates and nucleic acids. This may lead to the disruptions of cellular events which then cause abnormal cell growth, cell differentiation and so on. As one way to regulate protein functions, the amount of protein is controlled by the rates of protein biosynthesis and degradation (7). Another way is through post-translational modifications (PTMs). PTMs are covalent chemical processes that alter protein structures enzymatically either by adding modifying groups to amino acids or by proteolytic cleavage (8). These modifications contribute to the heterogeneity of protein population and, as well, the complexity of the cell. Some commonly studied PTMs include phosphorylation, acetylation, methylation, glycosylation, ubiquitination and sumoylation. Dysregulating of these PTMs is found to be highly related to human diseases, for example, cancers, diabetes and a couple of neurological disorders (8, 9).

## 1.5 Sumoylation

### 1.5.1 Sumoylation and Ubiquitination

Sumoylation is a PTM which can affect protein stability, enzymatic activities, subcellular localization, protein-protein interactions, or protein-chromatin interactions (10). SUMOylated proteins have been involved in regulating cellular progresses such as cell cycle progression, transcription, nuclear translocation and DNA repair (9, 11, 12) (**Fig 4**). SUMO (small ubiquitin-like modifier) is reversibly conjugated to one or more Lysine (K) residues of the target protein giving rise to a monosumoylated or a polysumoylated form (9). In most cases, the targeting of SUMO to lysine is directed by the consensus sequence ΨKxE/D (where Ψ is a large hydrophobic residue and x any amino acid) (13). However, other studies showed multiple other consensus sequences, such as a phospho-dependent sequence, reverse consensus and non-consensus regions, can also direct SUMO to acceptor proteins (14, 15).

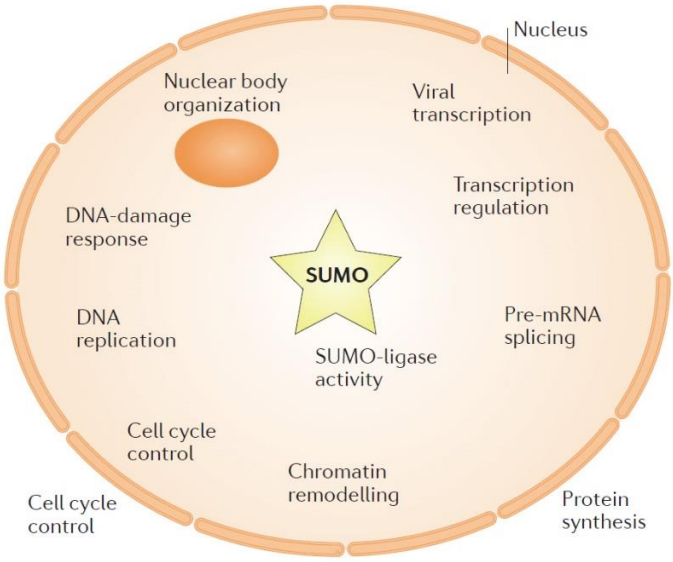



**Figure 4: Sumoylation functionality in nuclear processes.** Small ubiquitin-like modifiers (SUMOs) are nearly exclusively found in nucleus. Sumoylation, a process conjugating SUMO to substrate proteins, is highly involved in the dynamic nuclear processes including DNA repair, cell cycle regulation and transcription regulation. Reproduced from (28).

SUMO is conserved from human to yeast. In human, five SUMO paralogues are currently known (11, 16). SUMO1, 2 and 3 are the three main isoforms that are ubiquitously expressed, whereas SUMO4 and SUMO5 are only expressed in certain organs or cell types (11, 17). SUMO2 (also known as sentrin 2 and Smt3b) and SUMO3 (also known as sentrin 3 and Smt3a), which are often collectively referred to as SUMO2/3, show 97% sequence identity with each other, excepting the three-residue difference at their N-terminal regions (9, 16). Similar to SUMO2/3, SUMO1 (also known as UBL1, sentrin and Smt3c) is around 11 kDa. SUMO1 shares approximately 50% sequence similarity with SUMO2/3 (16, 18) (**Fig 5**). There is only one SUMO protein isoform in the yeast *Saccharomyces cerevisiae*, which is known as Smt3. Smt3 is ~45% similar to the three main human SUMO isoforms (16, 19) (**Fig 5**). Both Smt3 and SUMO2/3 are able to form a polysumoylated chains whereas SUMO1 cannot (19).


**Figure 5: SUMO family proteins.** The alignment shows the sequences of human ubiquitin, SUMO-1, SUMO-2, SUMO-3, SUMO-4 and yeast Smt3 (Smt3p). Residues identical in all proteins are shown on a blue background. The C-terminal di-glycine (GG) motif that is important for the maturation and conjugation of SUMO is also shown on a blue background. Purple background indicates residues that are only conservative changes across all four proteins. Residues showing semi-conservative changes among each of the proteins are shown on a yellow background. Reproduced from (16).

### 1.5.2 Sumoylation Machinery

The process of sumoylation involves a cascade of enzymatic activities which is similar to that of ubiquitination (16, 20). The three major enzymes are the activating enzyme E1, conjugating enzyme E2 and the ligase E3 (16) (**Table 1**).

**Table 1: Protein families involved in sumoylation machinery.** SUMO isoforms, SUMO E1 activating enzymes, SUMO E2 conjugating enzymes, SUMO E3 ligases and SUMO proteases in yeast and mammals are listed. Reproduced from (16).

| Protein family | Yeast (*Saccharomyces cerevisiae*) | Mammals |
|---|---|---|
| SUMO | Smt3 | SUMO-1<br>SUMO-2<br>SUMO-3 |
| SUMO E1 (activating enzyme) | Aos1<br>Uba2 | SAE1<br>SAE2 |
| SUMO E2 (conjugating enzyme) | Ubc9 | Ubc9 |
| SUMO E3 (ligase) | Siz1 | PIAS1 |
| | Siz2 | PIAS3 |
| | | PIASxα |
| | | PIASxβ |
| | | PIASy |
| | | RanBP2 |
| | | Pc2 |
| | Mms21 | Mms21 |
| | | HDAC4 |
| | | HDAC7 |
| | | MUL1 |
| | | Rhes |
| | | TOPORS |
| | | TLS |
| | | TRAF7 |
| SUMO proteases<br>(SUMO proteases and isopeptidase) | Ulp1 | SENP-1 |
| | Ulp2 | SENP-2 |
| | | SENP-3 |
| | | SENP-4 |
| | | SENP-5 |
| | | SENP-6 |
| | | SENP-7 |

Starting with an inactive precursor, the immature SUMO is cleaved at the C terminal region by members of the SENP family of sentrin/SUMO-specific protease (20) (**Fig 6**). In yeast, Ulp1 has the same protease activity as SENPs do in humans (16). This allows the exposure of a C-terminal di-glycine (GG) motif that is needed for SUMO conjugation (12). Then, this mature form of SUMO undergoes an activating step in an ATP-dependent manner. This activation of SUMO is carried out by E1 enzyme which reacts as a heterodimer of SAE1 (SUMO-activating enzyme E1) and SAE2 in human (21). The yeast form of the E1 enzyme is a heterodimer of Aos1 and Uba2. A thioester bond is formed between the active-site cysteine residue of SAE2 and the C-terminal glycine residue of SUMO before SUMO is transferred to Ubc9 (ubiquitin-conjugating 9) via another thioester bond (22). Ubc9, as the only known E2 enzyme in both human and yeast, can bind to specific lysine acceptors located in either SUMO consensus motif or several non-consensus regions of the target protein directly (16). Linking SUMO to Ubc9 promotes the conjugation of SUMO to the SUMO targets. In ubiquitination, there are two types of E3 ligases which are required for either transferring ubiquitin from the E2 enzyme to substrate proteins or facilitating this transfer without binding to ubiquitin (12, 16). However, SUMO E3 ligase activity is only observed to be required for sumoylation *in vivo*. These SUMO E3 ligases, for example Siz1 and Siz2 in yeast, serves as a scaffolds which promote the contact between SUMO-Ubc9 and the target protein (23). To achieve a highly dynamic reversible process, SENP enzymes also play a role in removing SUMO from sumoylated proteins. The yeast SENP homologues Ulp1 and Ulp2 are found to be in charge of the de-sumoylation process in that organism (24, 25).

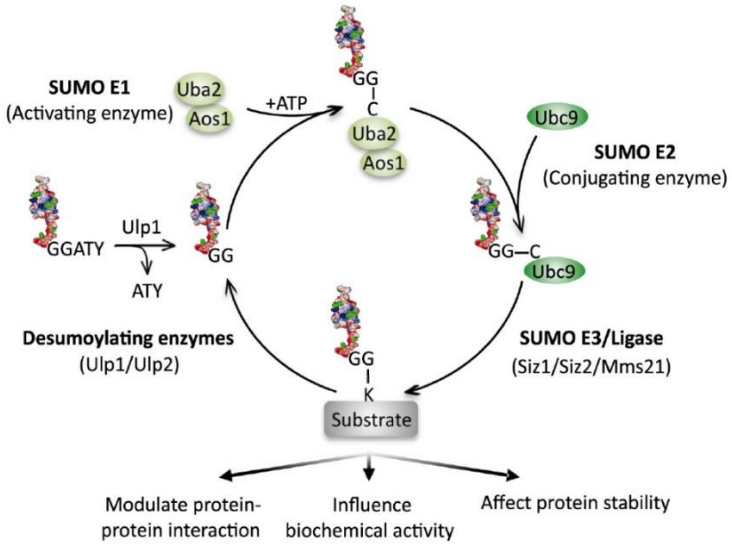



**Figure 6: Sumoylation machinery in yeast (*S. cerevisiae*) and examples of associated cellular effects.** SUMO precursor is catalyzed by SUMO protease Ulp1 to expose the C-terminal di-glycine (GG) motif. GG then undergoes adenylation which is followed by being transferred to Cysteine (C) residues in SUMO E1 (Uba2/Aos1 dimer) and then E2 (Ubc9) enzymes. Ubc9 targets SUMO to a lysine (K) residue located usually in a consensus motif. In vivo, SUMO E3 ligases (Siz1, Siz2 and Mms21) is observed to facilitate the SUMO conjugation step. Examples of the subsequent cellular effects resulting from sumoylation are provided. Reproduced from (20).

### 1.6 Sumoylation and Transcription

Sumoylation is a fundamental regulatory mechanism that controls a wide range of cellular processes such as cell cycle progression, DNA repair and chromatin segregation (9, 11, 12, 26). Although sumoylated proteins are found to be widespread, those involved in transcription, for example, GTFs, co-regulators and subunits of RNAPII, form one of the largest classes of SUMO targets both in yeast and mammalian cells (10, 26, 27). Through a number of proteomics studies, numerous SUMO-targeted proteins were identified, including GTFs and RNAPII subunits, were identified with high probability (28, 29, 30, 31) (**Table 2**). In the vast majority cases, sumoylation of transcription factors is associated with transcription repression through different SUMO-mediated mechanisms (27). For example, sumoylation of the transcription factor Elk-1 induces the recruitment of histone deacetylase HDAC-2. This results in reducing histone acetylation and subsequent transcription repression (32). However, sumoylation of some gene-specific transcription

factors can activate transcription (33). One of the well-illustrated studies found that sumoylation of transcription factor Ikaros serves as a derepression mechanism leading to transcription activation (34). In this case, the interaction between Ikaros and two co-repressors, Sin3 and NuRD, is blocked due to the two sumoylated lysine residues at the N-terminus of Ikaros.

**Table 2: High-probability SUMO acceptor sites on general transcription factors (GTFs) and RNAPII subunits.** Polypeptide sequences for GTF subunits that had been identified as SUMO targets in previous proteomics studies were analyzed by SUMOplot tool to identify likely sumoylation sites. Rpb1 and Rpb4 SUMO acceptor sites were identified in published studies without SUMOplot analysis (29, 31).

| GTF | SUBUNIT | SITE | SEQUENCE |
|---|---|---|---|
| TFIIA | Toa1 | 89 | NSSEFN**I**K**EE**NTGNE |
| | | 149 | ADVTSQ**P**K**IE**VKPEI |
| | | 153 | SQPKIE**V**K**PE**IELTI |
| TFIID | Taf7 | 146 | SGISIK**W**K**NE**RHAVV |
| | | 330 | VSSWEN**F**K**EE**PGEPL |
| | | 343 | PLSRPA**L**K**KE**EIHTI |
| | Taf8 | 205 | ITDLKT**I**K**KE**IVKES |
| | | 351 | LPKVQK**L**K**KE**KIRMA |
| TFIIF | Tfg1 | 60 | NSRGSL**V**K**KD**DPEYA |
| | | 91 | GRSNVK**V**K**DE**DPNEY |
| | | 497 | KESEQR**I**K**KE**MLQAN |
| | | 574 | KENESP**V**K**KE**EDSDT |
| | | 658 | NTVPSP**I**K**QE**EGLNS |
| | | 733 | GNDHME**L**K**KE*****|
| | Tfg2 | 296 | IKSIRM**P**K**KE**ILDYL |
| | | 359 | RPEYKK**L**K**EE**ERKAT |
| TFIIE | Tfa1 | 434 | SNTSND**V**K**QE**SINDK |
| RNAPII | Rpb1 | 1487 | VNADLD**V**K**DE**LMFSP |
| | Rpb4 | 83 | KHKKKH**L**K**HE**NANDE |

Although sumoylation of transcription factors is usually associated with transcription repression, surprisingly, chromatin immunoprecipitation (ChIP) studies revealed that SUMO is detected at transcriptionally active gene promoters, including both constitutive and activated inducible promoters (27). In contrast, no SUMO was detected at repressed or silent genes. Interestingly, Ubc9 was not found at the

promoter of constitutive genes, whereas activation of inducible genes is accompanied by recruitment of Ubc9 (27). This suggests that SUMO targets on constitutive promoters were sumoylated prior to their association with chromatin, whereas at induced genes, promoter-bound proteins are sumoylation during activation. It is paradoxical that a PTM that is associated with transcriptional repression is found at promoters of active genes. To resolve this, it will be important to identify SUMO targets associated with active gene promoters, and to examine how sumoylation regulates their specific functions at active promoters.

### 1.7 Sumoylation and Precursor mRNA Splicing

#### 1.7.1 Maturation of Precursor mRNA via Splicing

Precursor messenger RNAs (pre-mRNAs), which contain exons and introns, are synthesized by the RNAPII-mediated transcription machinery. Then a process known as pre-mRNA splicing is carried out. In this process, the spliceosome catalyzes the removal of introns as well as ligation of exons proceeding the formation of mature mRNA (35). As a macromolecular complex, the spliceosome consists of five small nuclear ribonucleoprotein particles (snRNPs), including U1, U2, U5 and U4/U6, and multiple non-snRNP splicing factors. Inside each snRNP, there is one small nuclear RNA (snRNA) or two in the case of U4/U6, a set of seven Sm proteins (B/B', D3, D2, D1, E, F and G) and many particle-specific proteins (36). The assembly of the spliceosome starts with recruitment of U1 snRNP to the 5' splice site (SS) and U2 snRNP to the branch site of pre-mRNA. This leads to the formation of pre-spliceosome (also known as complex A) (37) (**Fig 7**). In the next step, U4/U6-U5 tri-snRNP binds to the complex forming the pre-catalytic complex B. Complex B then undergoes numerous RNA and protein rearrangements, and consequently becomes a catalytically-active complex. This series of conversion of spliceosome allows the first step of the splicing reaction to begin. Moving on, more rearrangements are performed to convert complex B to complex C (also known as post-spliceosomal complex) in order to catalyze the second step. At this step, introns are removed and the flanking 5' and 3' exons are joined together. The spliceosome is then disassembled after completion

of the two catalytic steps (35, 37).

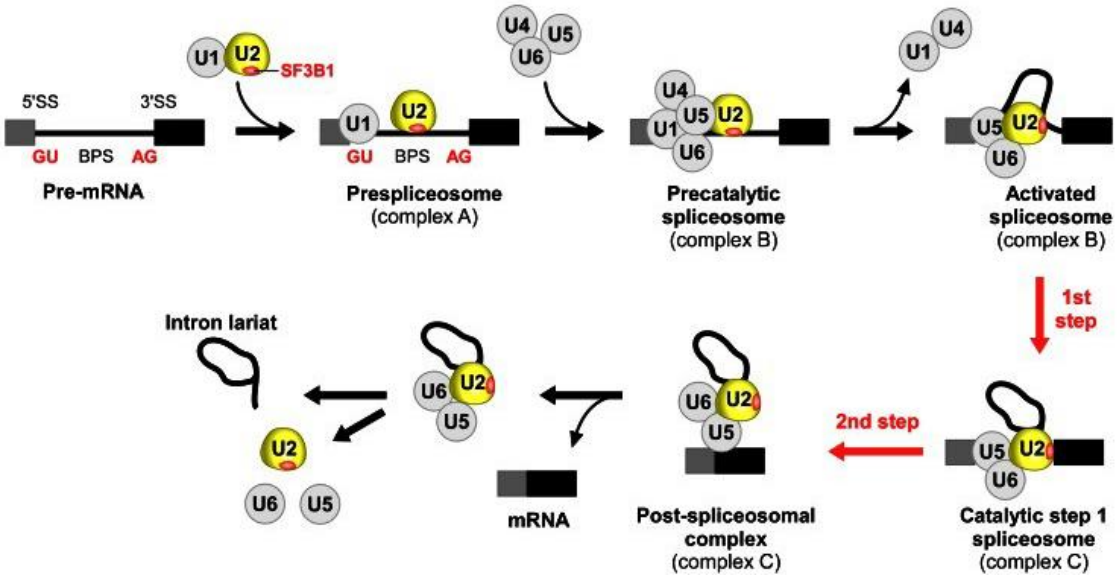



**Figure 7: The involvement of the essential U2 snRNP component SF3B1 in the stepwise pre-mRNA splicing process.** Binding of U1 snRNP to the 5' splice site (SS) and U2 snRNP to the branch point site (BPS) of pre-mRNA forms the prespliceosome (Complex A). Then, the U4/U6-U5 tri-snRNP is recruited forming the pre-catalytic complex B. Numerous RNA and protein rearrangements in complex B convert it to a catalytically activated complex B. This series of conversion of spliceosome allows the first step of splicing reaction to begin. More rearrangements are performed to convert complex B to post-spliceosomal complex (complex C) in order to catalyze the second step where introns are removed and the flanking 5' and 3' exons are joined together. The spliceosome is disassembled after these two catalytic steps. Reproduced from (37).

### 1.7.2 Sumoylation of Splicing Factors

In addition to its potential roles in regulating transcription initiation, SUMO conjugation has emerged as a critical regulatory mechanism of pre-mRNA splicing (35). In mammalian cells, several splicing factors, such as members of the heterogeneous nuclear ribonucleoprotein (hnRNP) and serine/arginine-rich (SR) protein families are SUMO substrates (35, 38, 39).

The splicing factor 3B subunit 1 (SF3B1) is a critical component of U2 snRNP. At an early step of the splicing reaction, base pairing of SF3B1 to both the 5' and 3' sides of the branch point sequence (BPS) contributes to the function of pre-spliceosome (37). Moreover, SF3B1 also interacts with other spliceosomal

components. These multiple interactions during splicing stabilizes the spliceosome as well as facilitates its recognition at the 3' SS (37). Intriguingly, a recent proteomics study identified SF3B1 as a SUMO target with high confidence (28). However, currently no follow-up study has successfully detected or characterized sumoylated SF3B1.

### 1.8 Objectives and Hypothesis

In the first of my two projects, I aimed to investigate how sumoylation of Tfg1 affects its properties and transcription in yeast. Previous work in the lab identified Lys residues 60 and/or 61 of Tfg1 as the major sumoylation site on TFIIF. To follow up on this, we examined Tfg1 sumoylation under different stress conditions, and found that a fraction of Tfg1 molecules are constitutively sumoylated under all conditions. Additionally, chromatin fractionation showed that both sumoylated and unsumoylated forms of Tfg1 are associated with chromatin. Based on the position of the SUMO-modified residues (Lys 60/61), we hypothesized that Tfg1 sumoylation may block the Tfg1-RNAPII interaction, which would then interrupt the formation of the PIC and eventually inhibit gene transcription. To address this, taking the advantage of previous lab work, several yeast mutants were generated that show either increased or decreased levels of Tfg1 sumoylation, compared to a wild-type (WT) strain. We tested whether modulated Tfg1 sumoylation levels had an effect on its interaction with RNAPII or on the association of Tfg1 and RNAPII with multiple constitutive and inducible genes by chromatin immunoprecipitation (ChIP). To further investigate the effects of Tfg1 sumoylation on transcription, we examined whether modulating Tfg1 sumoylation affects mRNA levels of various constitutive and stress-responding genes.

Indeed, our results showed that although a sumoylation-blocking mutation (K60,61R) does not alter the Tfg1-RNAPII interaction, reducing Tfg1 sumoylation level by deleting the SUMO E3 ligase Siz1 gene (*siz1*Δ) showed an influence on the association of Tfg1 with RNAPII. Additional pair of yeast strains, *ULP1* parental strain and *ulp1-1* mutant strain, were also used in order to test for the effect of increased Tfg1 sumoylation on this association. However, no clear result was obtained due to the faintness of immunoblot signals as well as the no-antibody control background. Notably, a common thing shared by these mutant

strains is that both deletion of *SIZ1* and mutation of *ULP1* affect numerous additional targets, besides Tfg1, indicating that the effects are not necessarily specific to sumoylation of Tfg1. To address this, additional work was carried out to generate a yeast strain that expresses an Smt3-Tfg1 fusion protein, which would mimic a permanently sumoylated form of Tfg1. Unfortunately, even after many trials and different approaches were attempted, the cloning procedures failed, and the fusion strain was not generated. Our attempts at examining how Tfg1 sumoylation affects its chromatin occupancy and the occupancy of RNAPII by ChIP were difficult to interpret because of inconsistent results after multiple trials. However, our analysis of how Tfg1 sumoylation affects mRNA abundancies did show a consist trend. At normal temperatures, the mRNA level of both heat shock (hs)-silenced and constitutive genes were higher in a yeast strain that shows reduced levels of Tfg1 sumoylation. Interestingly, upon heat shock treatment, the mRNA levels of hs-induced, hs-silenced and constitutive genes showed no difference in both WT and *tfg1-K60/61R* mutant (MT).

In my second project, we aimed to confirm that splicing factor SF3B1 is sumoylated in human cells, and further, if indeed confirmed, to test how sumoylation of SF3B1 affects splicing and gene expression. The hypothesis, based on the fact that SF3B1 was identified as a SUMO conjugate in proteomics studies, is that SF3B1 is sumoylated and its sumoylation alters gene expression by regulating splicing efficiency. However, our attempts at detecting SF3B1 by immunoblot, in both HEK 293T cells and in HeLa cells stably overexpressing expressing Histidine (His)-tagged SUMO, were initially not successful. Once the procedure was optimized, however, we were unable to detect sumoylated SF3B1. As a result, we did not proceed with this project.

# Chapter 2: Materials and Methods

## 2.1 Yeast Strains

All yeast (*Saccharomyces cerevisiae*) strains used in this study are listed in **Appendix Table 2**. Strains were derived from the W303 background strain. All chromosomally Tfg1-6HA-tagged yeast strains were generated by homologous recombination, as previously outlined (27). A *K. lactis TRP1* (Kl::TRP1) marker gene cassette was introduced into the genome adjacent to the 6HA-tag gene. The presence of the 6HA-tag was confirmed by PCR and Western blot analysis using an anti-HA antibody.

## 2.2 Yeast Media and Growth Conditions

Yeast cultures (10 – 50 mL) were grown at 30°C in yeast extract peptone dextrose (YPD; 1% yeast extract, 2% peptone & 2% glucose) or synthetic complete (SC; 0.17% YNB, 0.5% ammonium sulfate & 2% glucose) media to an optical density $A_{595}$ (O.D.) of 0.5-1.0.

To simulate several stress conditions, after the cultures reached an O.D. of 0.5-0.7, SC media was replaced with SC media containing different reagents, as indicated below, and yeast were allowed to continue growing for a certain time period depending on the type of the reagents. For high salt conditions, cells were treated with either 1 M sodium chloride (NaCl) or 1 M potassium chloride (KCl) for 5 min. For alcoholic conditions, cells were treated with 10% ethanol (EtOH) for 60 min. For amino acid starvation, cells were treated with 10 µg/mL sulfometuron-methyl (SM) for 20 min. Additionally, for oxidative stress, to the yeast culture was added 100 mM hydrogen peroxide ($H_2O_2$) for 5 min without replacing the original SC medium. For heat shock (hs) conditions, the culture was grown at 37°C for 12 min.

## 2.3 Yeast Immunoprecipitation (IP)

The strains were grown to saturation overnight at 37°C and diluted to an O.D. of 0.25 the next day. For IP, all steps were performed at 4°C. After reaching an O.D. between 0.5 – 1.0, the cells were collected by

centrifuging at 3000 *g* for 5 min, and then washed in IP-B+ buffer (20 mM HEPES, pH 7.5, 10% glycerol, 150 mM NaCl, 1 mM Ethylenediaminetetraacetic acid [EDTA], 1 mM phenylmethylsulfonyl fluoride [PMSF], yeast protease inhibitor cocktail, 1 M dithiothreitol [DTT] & 0.25 g N-ethylmaleimide [NEM]/100 mL buffer). After washing, re-collected yeast cells were resuspended in 0.5 mL of IP-B+ buffer and transferred to new microfuge tubes. 0.25 grams of acid washed glass beads were used in vortex homogenization for 30 min (vortexed for 15 min, 5 min on ice followed by another 15-minute vortex). The lysate was transferred to a new microfuge tube before centrifuging at top speed for 5 min. The supernatant was then transferred to a new tube and this step was repeated once. For preparing the input sample, 40 µL of lysate was diluted with an equal volume of 2X SDS-PAGE sample buffer (4% sodium dodecyl sulfate [SDS], 20% glycerol, 0.02% bromophenol blue, 10% beta-mercaptoethanol & 140 mM Tris-hydrochloric acid [HCl], pH 6.8) and was boiled for 3 min. 15 µL of Protein G sepharose beads were pre-conjugated with 1 µg of either rabbit anti-human influenza hemagglutinin (HA) epitope tag primary antibody (Novus) or mouse anti-8WG16 primary antibody (abcam) for at least 1 hr before mixing with the remaining yeast cell lysate. The lysate-bead mixture was then incubated on a rotator overnight at 4°C.

The next day, beads were washed 4 times with ice-cold IP-B+ buffer plus 1% NP40 followed by two washes with ice-cold IP-B+ buffer. Then, samples were boiled in 100 µL 2X SDS-PAGE sample buffer for 3 min to release proteins from the beads. Samples were cooled down on ice before loading.

**2.4 Western Blot and Antibodies**

Equal amounts of protein samples were resolved on a SDS-PAGE gel of an appropriate percentage (7.5%, 10% or 12.5%) and transferred to nitrocellulose membranes (BioShop) followed by incubating in blocking buffer (5% milk in 1X PBST [0.05% Tween-20 in 1% PBS (Fisher Scientific)]) for 30 min at room temperature. The membranes were then probed with an appropriate primary antibody (HA [1:3000, rabbit, Novus; for detection of HA-tagged Tfg1], SUMO1 [1:100, mouse, DSHB U of Iowa; for detection of mammalian SUMO1-involved sumoylation], SUMO2 [1:100, mouse, DSHB U of Iowa; for detection of mammalian SUMO2-involved sumoylation], SF3B1 [1:3000, rabbit, Abgent; for detection of splicing

component SF3B1], Smt3 [1:250, rabbit, Santa Cruz; for detection of sumoylation], 8WG16 [1:3000, mouse, abcam; for detection of RNA Polymerase II subunit Rpb1], Histone H3 [1:3000, rabbit, abcam; for detection of histone H3], or GAPDH [1:5000, rabbit, Sigma; for detection of GAPDH]), overnight at 4˚C. On the next day, the membranes were washed 3 times in 1X PBST for 5 min at room temperature. The blots were then incubated with the appropriate HRP conjugated secondary antibody (1:5000 for anti-rabbit IgG & anti-mouse IgG; Thermo Fisher Scientific) for 30 min at room temperature. The blots were then washed with 1X PBST three times again and the detection was done by incubating blots with chemiluminescence solution (ECL; Bio-Rad) for 5 min followed by exposing with a MicroChemi chemiluminescence imager (DNR Bio Imaging Systems).

**2.5 Chromatin Fractionation**

Yeast strains were incubated and grown to saturation overnight. On the next day, the cultures were diluted in 25 mL at an O.D. of 0.25. After reaching an O.D. of 0.5-0.6, the cells were harvested at 3000 *g* for 5 min at room temperature. They were then resuspended in 6.25 mL of Resuspension Buffer 1 (100 mM PIPES/ potassium hydroxide, pH 9.4 & 10 μM DTT), followed by incubation at 30°C for 10 min with agitation. The cells were collected (3000 *g* at room temperature for 5 min) in a 50-mL conical centrifuge tube and resuspended again in 2.5 mL of Resuspension Buffer 2 (0.6 M sorbitol & 25 mM Tris-HCl, pH 7.5 in YPD medium). Additionally, 10-15 mg/mL of the lytic enzyme zymolyase was added to each sample at this step, followed by 20-minute incubation at 30°C with agitation. The cells were again spun down (2000 RPM at room temperature for 3 min). The cell pellet was resuspended in 2.5 mL of Resuspension Buffer 3 (0.7 M sorbitol & 25 mM Tris-HCl, pH 7.5 in YPD medium), followed once again by incubation for 20 min at 30°C with agitation. Then, the cells were pelleted by centrifuging at 2000 RPM in room temperature for 3 min and then washed with 1 mL of lysis buffer (0.4 M sorbitol, 150 mM potassium acetate, 2 mM magnesium acetate, 20 mM PIPES/ potassium hydroxide, pH 6.8, 0.1% NP40, 1 mM PMSF, yeast protease inhibitor cocktail & 0.25 g NEM/100 mL buffer) three times. Pellets of each sample were then resuspended in 400 μL lysis buffer, followed by the addition of Triton-X 100 to a final concentration of 1%. Material

was mixed gently and then left on ice for 5 min. 90 µL of this lysed material was removed and an equal volume of 2X SDS-PAGE sample buffer (4% SDS, 20% glycerol, 0.02% bromophenol blue, 10% beta-mercaptoethanol & 140 mM Tris-HCl, pH 6.8) was added to it. This serves as the 'whole cell extract (WCE)' control. 100 µL of the remaining lysed material was removed and spun down in a microfuge tube at top speed for 15 min at 4°C. 95 µL of the supernatant obtained from this spin was transferred to a new tube and added an equal volume of 2X SDS-PAGE sample buffer. This serves as the 'soluble extract (Solub.)'. The pellet obtained from the spin was washed once with 100 µL lysis buffer and resuspended in 100µL of 2X SDS-PAGE sample buffer. This is the chromatin fraction (Chrom.). All of the three fractions (WCE, Solub. & Chrom.) were then boiled for 5 min and allowed to cool down. Samples were spun down at top speed for 1 min before being loaded directly to the SDS gels for western blot analysis, or vortexed well before freezing at -20℃.

**2.6 Chromatin Immunoprecipitation (ChIP) and Quantitative PCR (qPCR)**

Yeast strains were grown to saturation overnight. The next day, they were inoculated in 50 mL of YPD medium at 30°C and grown to an optical density $A_{595}$ (O.D.) of 0.5-0.7. The cultures were treated for cross-linking by 5 mL of a solution containing 1.1% formaldehyde, 100 mM NaCl, 1 mM EDTA and 50 mM HEPES-potassium hydroxide, pH 7.5 for 20 min (swirl briefly every 5 min). Cultures were then quenched with 7 mL of 2.5 M glycine for 5 min with occasional mixing. The samples were then transferred to a 50-mL conical tube and centrifuged at 3000 *g* for 5 min, 4°C. The obtained pellet was washed twice with 50 mL of ice-cold TBS (20 mM Tris-HCl, pH 7.5 & 150 mM NaCl), followed by washing with ChIP lysis buffer (50 mM HEPES- potassium hydroxide, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate & 0.1% SDS). Samples were then spun down at top speed for 5 min (4°C). The pellet was resuspended in 600 µL of ice-cold ChIP lysis buffer plus PMSF and yeast protease inhibitor cocktail and 400 µL (~0.3 g) of chilled acid-washed glass beads were then added.

To lyse the cells, a bead beating process (three times for 1 min with 1-minute breaks in between) was

performed using a Mini-Beadbeater. A 23G-guage needle was then used to poke a hole through the bottom of the tube containing the sample. Then poked tubes were placed on top of a 1.5 mL tube within a 14-mL culture tube and together spun at 950 *g* for 4 min (4°C). Samples obtained from this spin were then subjected to sonication aiming to shear the chromatin to fragments of ~500bp. Followed by centrifuging the samples in a microfuge at 14,000 *g*, 4℃, for 5 min, 600 µL of supernatant was transferred to a new chilled 1.5-mL tube and mixed with 15 µL of 5 M NaCl. 40 µL of the supernatant was stored as Input at -20°C, and the rest of the salt-adjusted supernatant was divided in two for immunoprecipitating with two different antibodies [1 µg of anti-HA rabbit antibody (Novus) or anti-8WG16 mouse antibody (abcam) pre-conjugated to 15 µL of washed magnetic Protein G beads], overnight at 4°C. The next day, a serial of bead washes (1 mL per wash per sample) was performed. The first wash used ChIP lysis buffer with 275 mM NaCl, followed by a second wash using the ChIP lysis buffer with 400 mM NaCl and then with the third washing buffer (10 mM Tris-HCl, pH 8.0, 0.25 M lithium acetate, 1 mM EDTA, 0.5% NP-40 & 0.5% sodium deoxycholate) and finally with TE buffer (10 mM Tris-HCl, pH 8.0 & 1 mM EDTA). Between each wash, the samples were incubated and rotated at room temperature for 4 min. After these washes, the beads were incubated with ChIP elution buffer (50 mM Tris-HCl, pH 7.5, 10 mM EDTA & 1% SDS) for 20 min at 65°C. Samples were separated from the beads using a magnet and then transferred to a new tube. Beads were then incubated with 250 µL of TE buffer again at room temperature for 4 min and supernatant obtained at this step was pooled with previous sample. This serves as the IP samples. Inputs (previously stored at -20˚C) were then thawed and mixed with 460 µL of TE buffer. Both IP samples and Input samples were treated with 10 µL of 10 mg/mL RNase A (EMD Millipore; 37°C, 30 min) and then 10 µL of 20 mg/mL proteinase K (BioShop; 42℃, 1 hr). After these treatments, all samples were incubated at 65°C for 4 hours to overnight to reverse cross-links. Each ChIP experiment was conducted at least three times. The average of quantitative PCR (qPCR) analyses are presented relative to the chromatin occupancy level at an untranscribed region of chromosome V (Chr V), with standard deviations shown as error bars. The value for each sample is calculated as follows:

Percent Chr V levels = 100*2^(-ΔCt[normalized ChIP to Chr V levels]).

Primer sequences used for qPCR are listed in **Appendix Table 3**.

### 2.7 RNA Extraction and Reverse Transcription-PCR (RT-PCR)

Yeast strains were grown in 10 mL YPD medium to an O.D. of 0.5-0.6 at 30°C. For heat shock treatment, cells were grown at 37℃ (instead of 30℃) for 20 min before harvesting. Samples were pelleted by centrifugation at 3000 *g*, 4°C, for 3 min and washed twice with chilled AE buffer (50 mM sodium acetate, pH 5.2 & 10 mM EDTA, pH 8.0; prepare with DEPC-treated [RNase-free] water). The pellet was then resuspended in 400 μL of ice-cold AE buffer and thoroughly mixed with 40 μL of 10% SDS, followed by the addition of 440 μL of phenol, pH 4.5. Samples were then chilled in a dry ice/ethanol bath (crushed dry ice powder mixed with 95% ethanol forming a thick slurry mixture) for 5 min and then transferred to a 65°C water bath for 5 min. Then a vortex-mixing step was performed for 30 s. After repeating this freeze/thaw/vortex cycle, samples were again flash frozen in a dry ice/ethanol bath before centrifuging at top speed, room temperature, for 7 min. The aqueous layers were transferred to new tubes for RNA precipitation using 50 μL of 3 M sodium acetate, pH 5.2 and 1 mL of chilled absolute ethanol. The RNA samples were then left on dry ice for 10 min before spinning down in a microfuge at 14,000 *g*, 4°C, for 15 min. The pellet was then washed with 1 mL of chilled 70% ethanol and centrifuged again at top speed for 5 min. Additional spinning and air-dry steps were performed to remove the residual liquid. The RNA pellets were resuspended in 100 μL of diethyl pyrocarbonate [DEPC]-treated water and the concentration was measured using the Nanodrop spectrophotometer.

For reverse-transcription, 24 μg of RNA samples were treated with DNase I (New England Biolabs) and approximately 0.72 μg of this DNase-treated RNA was used for cDNA synthesis using the iScript reverse transcriptase reaction (Bio-Rad) according to the manufacturer's instructions.

For quantification, qPCRs were performed using SYBR green mix (Froggabio) according to the manufacturer's instructions. Transcripts were normalized to 25S rRNA and the levels were calculated by

the $2^{-\Delta\Delta CT}$ method (40). All experiments were performed at least three times and the average values are presented with standard deviations shown as error bars. Primer sequences are listed in **Appendix Table 3**.

### 2.8 Genomic DNA Extraction

Yeast cultures (10 mL per strain) were incubated in the appropriate media overnight to saturation. The cells were collected at 3000 *g*, room temperature, for 5 min the next day. The pellet was washed in 0.5 mL of sterile water and spun down again at 3000 *g* for 2 min, then resuspended in 0.2 mL of extraction buffer (2% Triton-X 100, 1% SDS, 100 mM NaCl, 10 mM Tris-HCl, pH 8.0 & 1 mM EDTA). Then, 0.3 g of acid-washed baked glass beads and 0.2 mL of phenol:chloroform (1:1) were added to the samples, followed by a 3-min vortex. After adding 0.2 mL of TE, pH 7.4 to the samples, the cells were then centrifuged at top speed for 5 min. The aqueous layer was transferred to a new microfuge tube. 1 mL of absolute ethanol was then added and the sample was mixed by inversion followed by spinning at 16,100 *g* for 1 min. The supernatant was subsequently removed by decanting, and the pellet was resuspended in 0.4 mL of 1X TE (10 mM Tris-HCl, pH 8.0 & 1 mM EDTA, pH 8.0). The samples were then treated with 3µL of 10 mg/mL RNase A for 5 min at 37°C. After this incubation, 10 µL of 4 M ammonium acetate and 1 mL of absolute ethanol were added to re-precipitate the DNA. Samples were then inverted and centrifuged at 16,100 *g* for 2 min. Finally, the supernatant was removed and the pellet was resuspended in 50 µL of TE (stored at -20°C).

### 2.9 Spot Assay (Drop Test)

Cells were grown in appropriate liquid growth medium overnight, and the O.D. of each culture was determined the following morning. Approximately 10,000 cells of each strain were spotted next to each other in the first position, and serial five-fold dilutions were spotted in adjacent positions, on appropriate solid medium plates. All plates were then incubated at 30°C (unless otherwise stated) and were imaged daily for up to 3 days.

## 2.10 Cloning

Appropriate primers were used in basic PCR amplifications in order to generate desired gene copies required for cloning (**Appendix Table 4**). For each reaction, both forward and reverse PCR primers were used to amplify the sequence of one double-stranded DNA template. All reactions were carried out using 3-STEP standard PCR which contains a denaturation step, an annealing step and an extension step (detailed profile shown in **Appendix Table 5**). Before yeast transformation, PCR products were validated by agarose gel electrophoresis followed by purification using GeneJET Gel Extraction Kit (Thermo Fisher Scientific). The fusion PCR was performed similarly as described above. For each reaction, both forward and reverse PCR primers were used to amplify the sequences of two double-stranded DNA templates which were designed to have overlap sequences. First DNA template contains gene encoding Smt3 followed by a 36-basepair *TFG1* gene (downstream of *TFG1* start codon). Second DNA template contains gene encoding Tfg1-6HA.

For plasmid preparing, pMGL4 plasmid was first collected using GeneJET Plasmid Miniprep Kit (Thermo Fisher Scinetific). Then both pMGL4 (3 μg) and the insert DNA (2-3 μg) obtained from fusion PCR were digested by 1.5 μL of the restriction enzyme KpnI (BioLabs) in 50 μL reaction mix containing 5 μL of 10X NEBuffer 1.1 (BioLabs) and appropriate amount of $dH_2O$ under 37°C for 2 hrs. The obtained digested products were validated and purified (same as for PCR product) before ligation. Each ligation reaction was carried out in 20 μL of ligation mixture containing 2 μL of 10X T4 DNA ligase buffer, 2 μL of T4 DNA ligase, 1 μL of KpnI-digested pMGL4, 10 μL of KpnI-digested Insert DNA and appropriate amount of $dH_2O$. Mixture was incubated at room temperature for 10 min, and then, 2.5 μL of the mixture was transferred to 50 μL of competent *E. coli*. Next, the new mixture was incubated on ice for 20 min, followed by a 3-min incubation at 37°C and then a 5-min incubation on ice. Lastly, mixture was spread onto a Luria broth (LB) +ampicillin plate and incubated overnight at 37°C.

### 2.11 Yeast Transformation

Yeast cultures were inoculated overnight and grown to saturation. On the next day, cultures were diluted in 10 mL of SC medium and incubated at 30°C until reaching an O.D. of 0.5. The cells were then harvested at 3000 *g*, room temperature, for 5 min and then washed with 10 mL of sterile water. The pellet was resuspended in 500 µL of 0.1 M lithium acetate and transferred to a new microfuge tube. A quick spin (15 s) at 16,100 *g* was done to pellet cells. This was followed by resuspending the pellet in 500 µL of 0.1 M lithium acetate again. For each transformation, 100 µL of the sample was used. For each transformation, cells were mixed with Transformation Mix (240 µL of PEG 3500 50% w/v, 36 µL of 1.0 M lithium acetate, 25 µL of boiled SS-carrier DNA, 40 µL of water & 10 µL of a cleaned PCR product), vortexed vigorously for 1 minute and then incubated in a 42°C water bath for 40 min. After this incubation, the samples were centrifuged at top speed for 30 s, followed by removal of the transformation mix. The yeast cells were then plated onto appropriate solid medium and incubated at 30°C for 3 days.

### 2.12 Yeast Colony PCR

A small amount of the colony from the plate was picked using a pipet tip. Then these cells were transferred to a PCR tube containing 20 µL of 20 mM sodium hydroxide and mixed by agitation with the tip. After mixing, the PCR tube containing yeast cells was placed in the PCR machine and incubated at 95°C for 45 min. Meanwhile, for each sample, 25 µL of the PCR mix was prepare which contains14.75 µL of $dH_2O$, 2.5 µL of 10X PCR buffer (Taq Buffer with Magnesium chloride), 0.5 µL of 10 mM deoxyribonucleotide triphosphate, 1 µL of 5 µM Forward primer, 1 µL of 5 µM Reverse primer and 0.25 µL of Taq DNA Polymerase. 80 µL of $dH_2O$ was added to the incubated mixture followed by a spin at 3000 *g* for 5 min. 5 µL of the sample was taken from the top of the mixture and transferred to the PCR mix followed by the 3-STEP standard PCR (**Appendix Table 5**).

### 2.13 Mammalian Cell Lines and Media

HEK 293T cells, HeLa cells and *His-SUMO* HeLa cells (obtained through Dr. Patricia Richard, Columbia University) were maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS), 100 units/mL of penicillin/streptomycin antibiotics.

### 2.14 Mammalian Immunoprecipitation (IP)

Culture plates containing the growing cell line were washed twice with chilled 1x PBS before being treated with 1 mL of lysis buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 mM Ethylene glycol tetraacetic acid [EGTA], 0.1% Triton X-100 & 1 mM DTT) plus 100X mammalian protease inhibitor cocktail and NEM (0.03 g for 5 mL). After cells were scraped together and transferred to a microfuge tube, they were rotated in the cold room for 20 min. Samples were spun down at top speed, 4°C, for 20 min, followed by transferring the supernatant to a new tube. 50 μL of the supernatant of each sample was saved as Input with the addition of 50 μL of 2X SDS-PAGE sample buffer (4% SDS, 20% glycerol, 0.02% bromophenol blue, 10% beta-mercaptoethanol & 140 mM Tris-HCl, pH 6.8). 15 μL of lysis buffer-washed sepharose beads was added to the rest of each sample, followed by adding 2 μL of the appropriate antibodies (mouse anti-SF3B1 from LSBio) and rotating overnight at 4°C. On the next day, samples were removed by centrifuging at 2600 rpm for 1 min. The beads were washed three times in 1 mL of freshly-made lysis buffer with a 5-min incubation at 4°C in between. After washing, beads were collected by centrifugation, then 150 μL of 2X SDS-PAGE sample buffer was added to each tube, followed by boiling for 4 min and chilling on ice. These samples were then loaded on an SDS-PAGE gel for western blot analysis.

### 2.15 Immobilized Metal Affinity Chromatography (IMAC)

Three growing cell lines, including wild-type HeLa cell line, His-tagged SUMO1-expressing Hela cell line and His-tagged SUMO2-expressing HeLa cell line, were washed once with chilled 1x PBS. Cells were then collected in 1 mL chilled 1XPBS. Input control was prepared by adding 2X SDS-PAGE sample buffer to

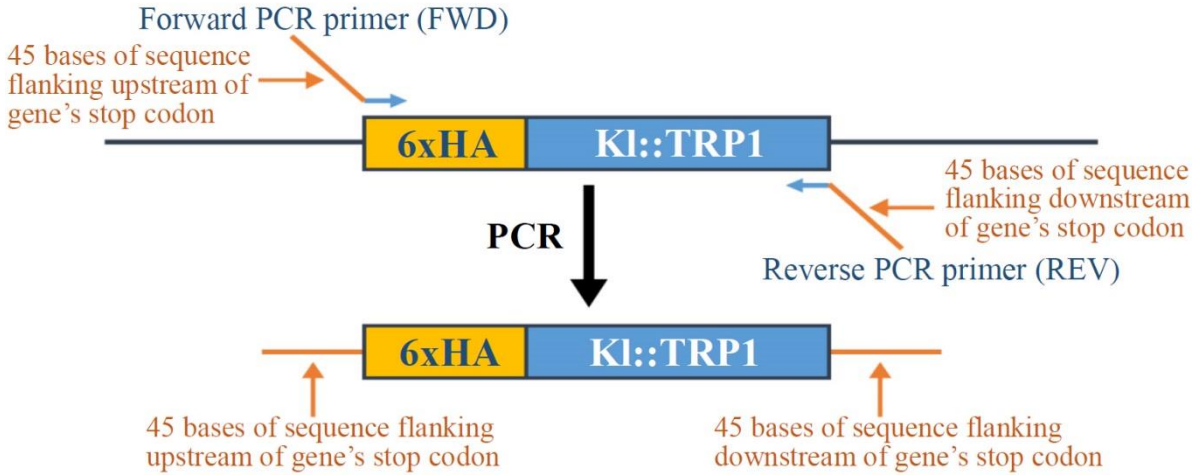
10% of the collected cells (110 µL) and then boiling for 6 min. The remaining cells were harvested by spinning down at 5000 rpm, 4°C, for 5 min. After removing the supernatant, cells were resuspended in 6 mL Denaturing Lysis Buffer (6 M Guanidinium-HCl, 100 mM disodium hydrogen phosphate-sodium phosphate monobasic, 10 mM Tris-HCl, pH 8.0, 5 mM imidazole & 10 mM beta-mercaptoethanol [add fresh at time of use]). 75 µL of nickel-nitrilotriacetic acid (Ni-NTA) agarose beads was added to the cell lysate before incubating at room temperature for 4 hrs. After incubation, four washing steps were performed. In each step, beads were washed for 5 min at room temperature in 750 µL of the appropriate buffer and collected by spinning at 3500 RPM, room temperature, for 4 min and then at 5600 RPM, for 5 min. Beads were first washed with Denaturing Lysis Buffer without imidazole followed by Buffer A pH 8.0 (8 M urea, 100 mM disodium hydrogen phosphate-sodium phosphate monobasic, 10 mM Tris-HCl, pH 8.0 & 10 mM beta-mercaptoethanol [add fresh at time of use]), Buffer A pH 6.3 (8 M urea, 100 mM disodium hydrogen phosphate-sodium phosphate monobasic, 10 mM Tris-HCl, pH 6.3, 10 mM beta-mercaptoethanol [add fresh at time of use] & 0.2% Triton X-100) and Buffer A pH 6.3 (8 M urea, 100 mM disodium hydrogen phosphate-sodium phosphate monobasic, 10 mM Tris-HCl, pH 6.3, 10 mM beta-mercaptoethanol [add fresh at time of use] & 0.1% Triton X-100). Proteins were eluted with 75 µL of Elution Buffer (200 mM imidazole, 150 mM Tris-HCl, pH 6.7, 30% glycerol, 5% SDS & 720 mM beta-mercaptoethanol [add fresh at time of use]) at room temperature for 20 min. Samples were added with 4X SDS-PAGE sample buffer (8% SDS, 40% glycerol, 0.04% bromophenol blue, 5% beta-mercaptoethanol & 240 mM Tris-HCl, pH 6.8) and boiled for 5 min before loading.

# Chapter 3: Results

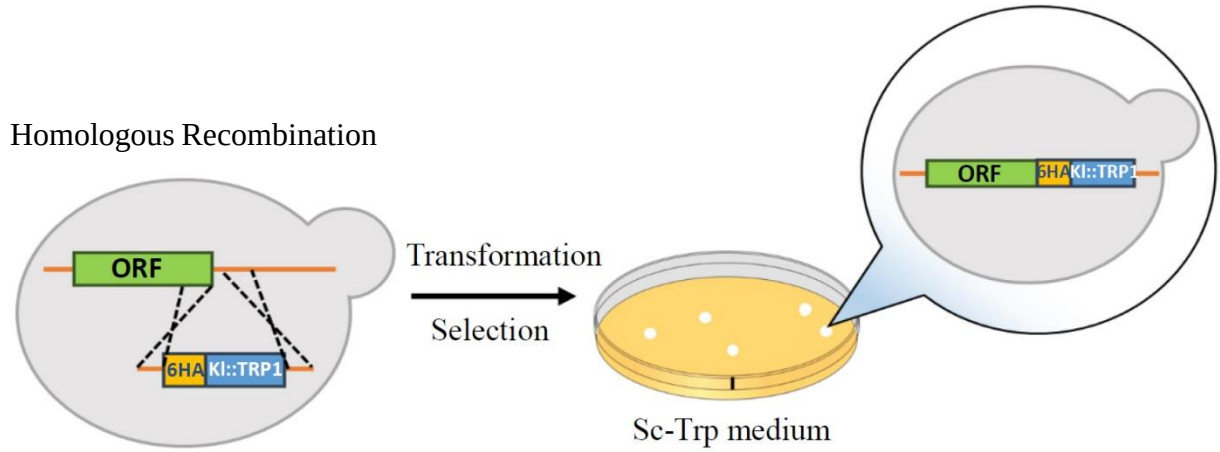
## 3.1 Sumoylation of GTF/RNAPII Subunits

To investigate the involvement of sumoylation in gene transcription, we examined a panel of GTF/RNAPII subunits for their sumoylation level. We used a Tagging - Immunoprecipitation - Immunoblot (Tag-IP-IB) procedure to confirm sumoylation of the protein of interest (**Fig 8**). Taking the advantage of HA-tagging system, we fused a 6xHA-tag (6HA) to each of the tested GTF/RNAPII subunits. This allows us to separately immunoprecipitate (IP) all proteins of interest through a single HA-IP experiment. The extracted proteins were then probed with Smt3 (yeast SUMO) and HA-tag antibodies. Both sumoylated and unmodified forms of proteins can be detected in HA blot whereas Smt3 blot reveals not only the sumoylated proteins of interest but, potentially, other associated sumoylated components that co-IP with the tagged proteins.

## STEP 1: PCR amplification of 6xHA tag-Kl::TRP1 marker




## STEP 2: Yeast transformation




## STEP 3: Immunoprecipitation (IP) and Immunoblot (IB) examination

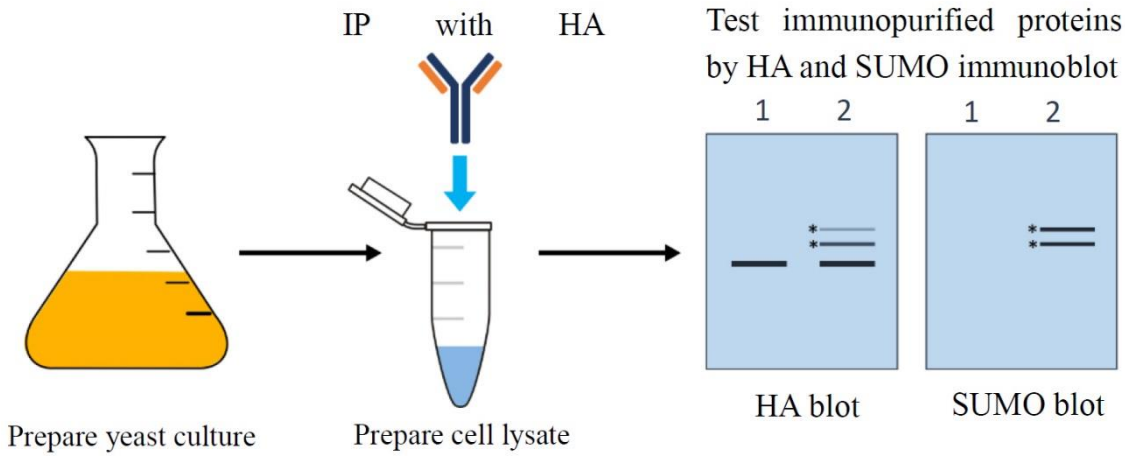


**Figure 8: Determination of whether the proteins of interest are sumoylated using Tagging-Immunoprecipitation-Immunoblot (Tag-IP-IB) procedure.** (**1**) STEP1: To generate a yeast strain expressing the protein of interest tagged with 6xHA tag, a "tagging cassette" that consist of 6xHA tag coding sequence and the *K. lactis TRP1* (Kl::TRP1) marker gene is amplified by PCR. Primers used in this amplification contain 45 bases of sequence matching sequence upstream and downstream of the stop codon of the gene of interest. (**2**) STEP2: Taking advantage of homologous recombination, the stop codon of the gene of interest is replaced by the amplicon during yeast transformation. Transformed yeast cells are selected using medium lacks Tryptophan. (**3**) STEP 3: The newly generated yeast strain undergoes IP procedure. The protein of interest fused with 6xHA tag is pulled down at this step and then examined by HA and SUMO immunoblot. In the example shown, lane 1 shows no sumoylated form of the protein of interest, while the protein in lane 2 shows two sumoylated isoforms (asterisks) as well as the unmodified form. Depending on the degree of sumoylation of the protein of interest, some sumoylated isoforms are not detectable by HA immunoblot.

Four GTF subunits, including Toa1 (subunit of TFIIA), Taf7 (subunit of TFIID), Tfg1 (subunit of TFIIF) and Tfa1 (subunit of TFIIE), and, as a positive control, the RNAPII largest subunit, Rpb1, were tagged with 6HA in our experiment. Each of these GTF subunits had been previously identified as putative SUMO targets in large-scale proteomics analyses, and Rpb1 sumoylation had previously been confirmed and studied (31) (**Table 2**). Immunoblotting for the HA-tag confirmed the presence of each of the subunits in the tagged strains (**Fig 9**). Among the panel of IPs, multiple sumoylated proteins were detected in the Smt3 blot, including relatively strong signals for proteins that corresponded in size to sumoylated Tfg1 (~142 kDa [130 kDa unmodified Tfg1 plus 12 kDa for the Smt3]) and the positive control, sumoylated Rpb1 (~287 kDa [275 kDa unmodified Rpb1 plus 12 kDa for the Smt3]; arrows in **Fig 9**). A smear or ladder of bands seen in the Smt3 immunoblot of Tfg1 likely represents multiple poly- or multi-sumoylated forms of Tfg1. Correspondingly, at least two modified forms of Tfg1 were detected in the HA blot in addition to the major, likely unmodified, form of the polypeptide. Further experiments, including the mutational analysis described below, confirmed that these bands correspond to sumoylated Tfg1 (see below). Interestingly, although Smt3-conjugated Toa1 was not detected, multiple sumoylated proteins were found in the Toa1-IP, possibly reflecting sumoylated proteins that interact with TFIIE (red-circled in **Fig 9**). Overall, our results show that among all the GTF subunits tested, Tfg1 has the highest potential to be a significant SUMO target.

We then focused our experiments on investigating sumoylation of this TFIIF subunit.

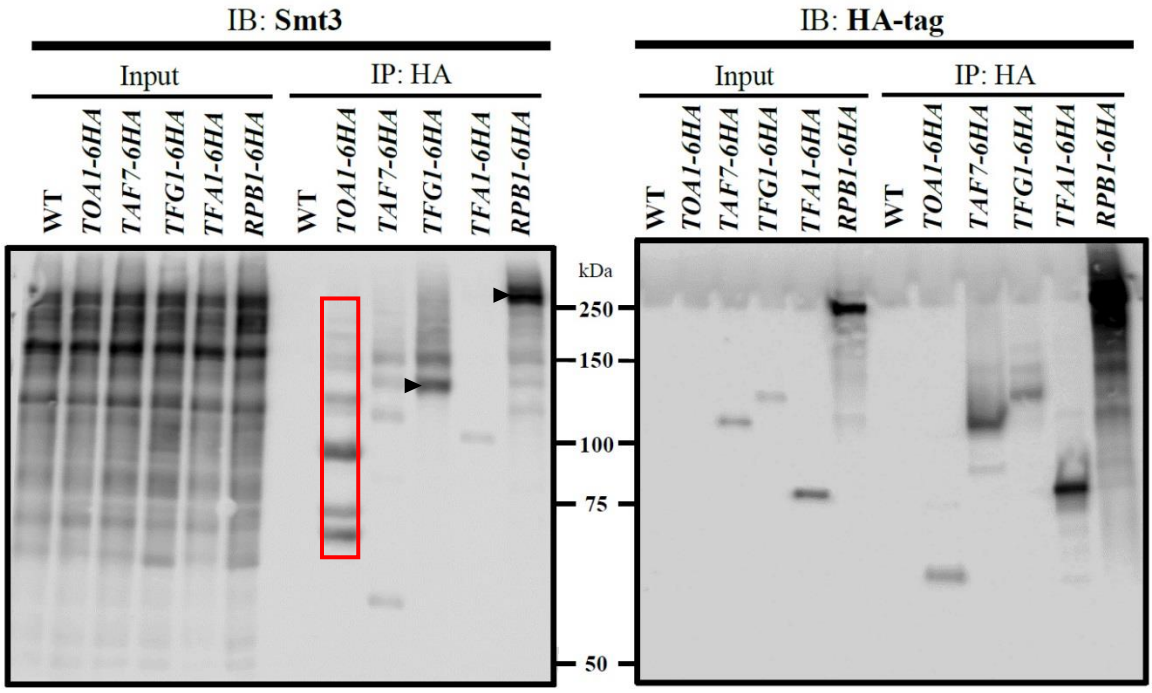



**Figure 9: Testing sumoylation level of a panel of GTF/RNAPII subunits.** 6HA-tagged GTF subunits (Toa1, Taf7, Tfg1 and Tfa1) and RNAPII largest subunit Rpb1 were pulled down by HA-IP. Western blot probing for SUMO (Smt3; left) revealed the level of sumoylation of each protein. Both Tfg1 and Rpb1 showed significant sumoylation levels (black arrow). Other sumoylated components that interact with these subunits were also detected (red rectangle). HA blot (right) confirmed the presence of each subunits. A wild-type (WT) yeast strain with no 6HA-tagged proteins was used as a control.

### 3.2 Sumoylation of Tfg1

#### 3.2.1 Tfg1 is the Major Sumoylated Subunit of TFIIF

Knowing that Tfg1 is one of three TFIIF subunits, we wanted to see whether the other two subunits, Tfg2 and Tfg3, are also highly sumoylated. Three yeast strains expressing Tfg1-6HA, Tfg2-6HA and Tfg3-6HA, respectively, were generated and tested by HA-IP. Immunoblot results indicated a successful pulldown of all three subunits (**Fig 10**). The control strain, which lacks any HA-tagged proteins, showed no band as expected, but both Tfg1 and Tfg2 sumoylation were observed in Smt3 blot. However, Tfg1 has a much

higher level of sumoylation compared with that of Tfg2. Additionally, we did not detect sumoylation of Tfg3. Thus we concluded that Tfg1 is the major sumoylated TFIIF subunit.

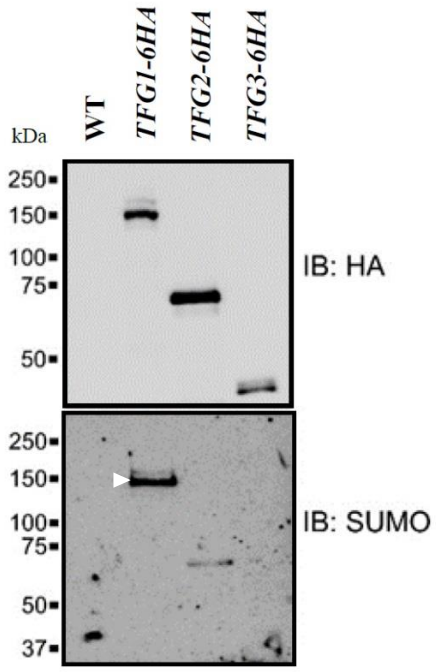



**Figure 10: Tfg1 is the major sumoylated subunit of TFIIF.** Yeast strains expressing Tfg1-6HA, Tfg2-6HA or Tfg3-6HA were used in HA-IP which was followed by immunoblotting for Smt3 (SUMO) and HA-tag. HA blots validated the successful immunopurification of each TFIIF subunit. Smt3 blot showed Tfg1 was sumoylated significantly higher than the other two subunits (white arrow). A wild-type (WT) yeast strain with no 6HA-tagged proteins was used as a control. (Result obtained from Russell Bahar)

As TFIIF is an essential component of the transcription machinery that assembles on gene promoters, we wished to determine whether sumoylated or unmodified forms of Tfg1 are associated with chromatin. Two yeast strains, both producing Tfg1-6HA, were tested by chromatin fractionation analysis, including an *ulp1-1* mutant strain and an isogenic *ULP1* wild-type (parental) strain. Whole cell extract (WCE), and soluble (Solub.) and chromatin-associated (Chrom.) fractions were isolated and analyzed by immunoblot with antibodies for HA, GAPDH, and histone H3 (**Fig 11**). A band at ~110 kDa in the HA immunoblot confirmed the presence of unmodified Tfg1 in WCE and in both soluble and chromatin fractions (**Fig 11**). A slower-

migrating HA signal corresponding to monosumoylated Tfg1 was also observed in these two fractions (Solub. and Chrom.), which indicates that both sumoylated and unsumoylated forms of Tfg1 are found on chromatin.

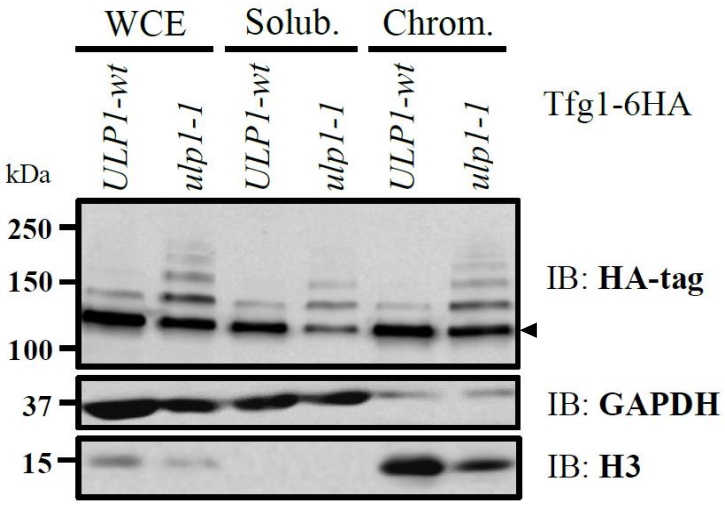



**Figure 11: Sumoylated Tfg1 presents in both soluble and chromatin fractions.** Two yeast strains, one expressing wild-type Ulp1 (*ULP1-wt*) and the other expressing a partially defective form of Ulp1 (*ulp1-1*) which results in increased global levels of sumoylation, were used for determining the location of both unmodified (black arrow) and sumoylated (higher molecular weight forms) Tfg1. Chromatin fractionation followed by immunoblotting for HA-tag shows both forms of Tfg1 can be found in soluble extract (Solub.) as well in the chromatin fraction (Chrom.). The overall amount of Tfg1 can be detected in whole cell extract (WCE). GAPDH and histone H3 serve as controls for soluble fraction and chromatin fraction, respectively.

To confirm that the higher molecular weight forms of Tfg1 correspond to sumoylated Tfg1, the *ulp1-1* strain was included. Ulp1 is one of two major yeast desumoylation enzymes, and the *ulp1-1* mutant displays significantly reduced protease activity, resulting in higher levels of sumoylation on many proteins. Consistent with this, in our HA immunoblot analysis, the ladder of bands that we attributed to sumoylated Tfg1 were indeed more prominent in the *ulp1-1* mutant strain (**Fig 11**). Sumoylation can affect protein localization and thus, by assessing *ulp1-1* mutant, we wished to see whether altered sumoylation affects the association of Tfg1 with chromatin. Our result shows that, comparing with the wild-type *ULP1* parental strain, the increased sumoylation level on Tfg1 had no effect on the localization of both forms of Tfg1 (**Fig**

**11**). Histone H3, a control for proteins associated with chromatin, was shown to be absent in soluble extract, and GAPDH, which is predominantly cytoplasmic, was mostly absent from the chromatin fraction. Unexpectedly, we also observed a weak band for GAPDH in chromatin fraction. GAPDH is commonly used as a control for the soluble fraction during chromatin fractionation. However, other study showed that, although most GAPDH is present in the cytoplasm, it can also be found in membranes, the nucleus and cytoskeletal structures (41). As a result, the detection of chromatin-associated GAPDH may due to the strongly sensitive GAPDH antibody. To conclude, we reasoned that both sumoylated and unmodified Tfg1 were located in soluble fraction and on chromatin to a similar extend. The chromatin association of Tfg1 is therefore not dependent on sumoylation or altered by artificially increasing its sumoylation level.

### 3.2.2 Effects of Different Stress Conditions on Tfg1 Sumoylation

Rapidly growing yeasts are sensitive to the changes in growth conditions (42). Under stress conditions, yeast cells have to balance the cellular growth and the stress response. To defend against different stresses, yeast cells correspondingly alter the transcriptome, such as by activating stress-response genes and repressing stress-silenced genes, in order to maintain proper cellular functions. As one of the essential components of the PIC, the functions of Tfg1 are critical for changes in gene transcription under stress conditions. Because sumoylation is known to have the ability to alter protein functions, it was surmised that the level of Tfg1 sumoylation might change in response to stress, in order to regulate its function and help overcome stress-induced defects.

To test this, first, a yeast strain expressing Tfg1-6HA was grown until reaching the exponential state, in which yeast cells are considered to be rapidly growing. Then, multiple stress treatments, including high salt concentration, alcohol, heat, oxidization and amino acid starvation, were performed to simulate different stress environments. Cells were cultured in SC medium with or without different treatments, for varied durations. We isolated the cell lysate immediately after finishing stress treatments and performed IP-immunoblot for detection of Tfg1 sumoylation. Compared with the control where yeast cells were cultured under normal growth conditions, high salt concentrations (SC+1 M NaCl/1 M KCl) did not affect Tfg1

sumoylation (**Fig 12A**). Similarly, no change was observed in the level of Tfg1 sumoylation under conditions with either heat shock (37 °C) or alcohol stress (SC+10% EtOH). This indicates that, under conditions where the transcriptome is changed due to the activation of heat shock-induced/alcohol-response genes and the repression of heat shock-silenced/alcohol-silenced genes, the level of Tfg1 sumoylation is not affected. Sulfometuron-methyl (SM) interrupts the biogenesis of isoleucine and valine leading to amino acid starvation (43). Our result showed that treating cells with 10 μg/mL SM did not change Tfg1 sumoylation levels (**Fig 12A**). Notably, alcohol stress led to a significant decrease in yeast growth rate and the level of both sumoylated and unmodified Tfg1 detected were much lower in the alcohol-treated sample than in the others. To further analyze the result, we performed a quantification of HA blot, and determined the ratio of sumoylated Tfg1 to total Tfg1 in each treatment (**Fig 12B**). Yeast cells under all stress conditions, even the 10% EtOH environment, showed a similar level of sumoylated Tfg1. Taking together, this experiment suggests that the sumoylation of Tfg1 is not affected by different stresses.

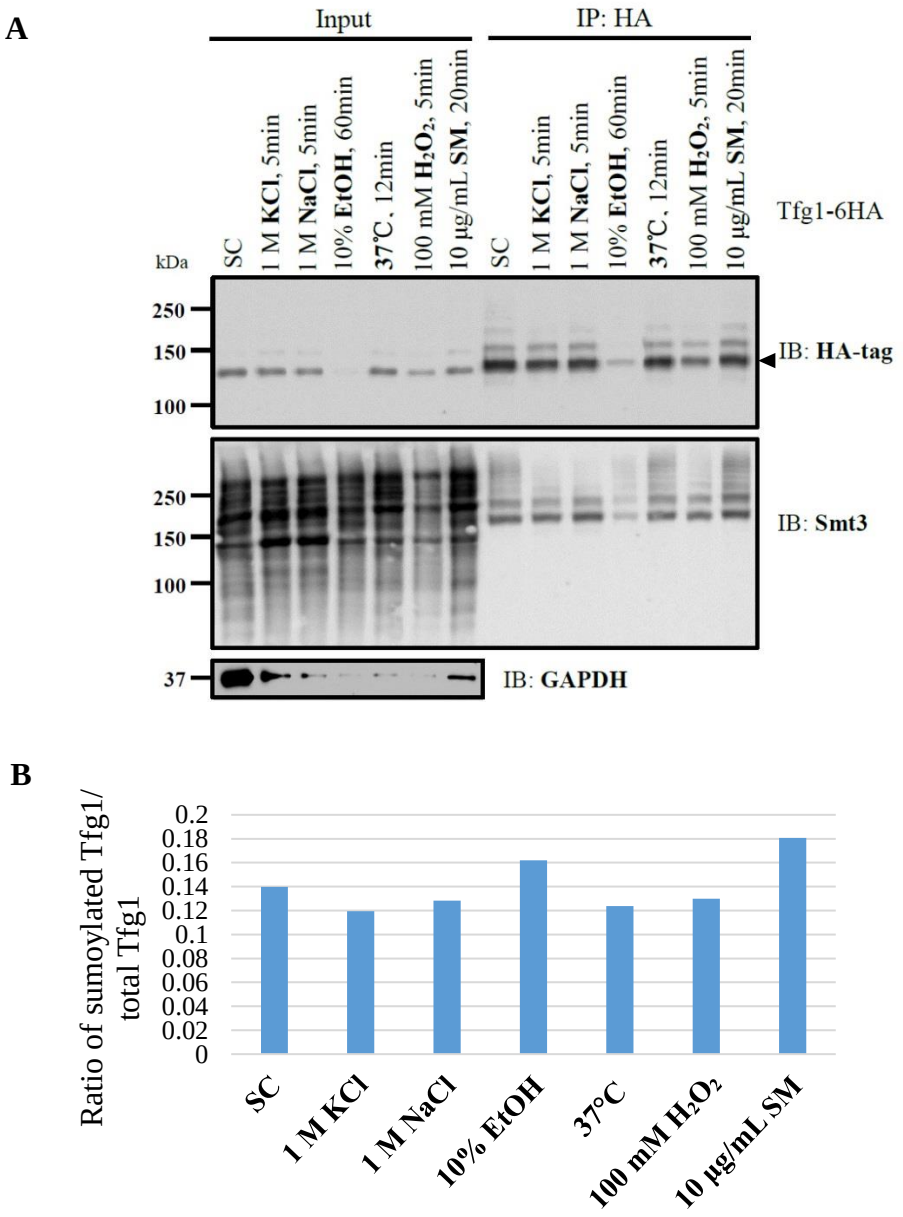



**Figure 12: Tfg1 sumoylation does not appreciably change under different stress conditions. A**) Yeast strain expressing 6HA-tagged Tfg1 was cultured and treated with different stress conditions: salt stress (1 M NaCl and 1 M KCl), alcohol stress (10% EtOH), heat shock (37°C) , oxidative stress (100 mM $H_2O_2$) and amino acid starvation stress (10 µg/mL SM), as well as the treatment duration are labeled above HA blot. Cell lysates were prepared after treatments followed by HA-IP and immunoblot. HA blot shows both sumoylated and unmodified (black arrow) Tfg1. Probing Smt3 further reveals Tfg1 sumoylation level. GAPDH serves as loading control. **B**) Bar graph shows the relative sumoylated Tfg1 level under different stress conditions. Quantification of immunoblot (panel **A**) was analyzed by ImageJ. The ratio of sumoylated Tfg1/total Tfg1 (y axis) was calculated using Excel. Stress conditions was labeled on x axis. No appreciably change in Tfg1 sumoylation was observed.

### 3.3 Identifying the Major Sumoylated Site of Tfg1

#### 3.3.1 Lysine 60 and 61 are the Predominant Sumoylated Residues of Tfg1

We focused on determining the major sumoylated site of Tfg1 in order to better understand the role of Tfg1 sumoylation. A previous cross-link-mass spectrometry study showed that Tfg1 Lysine residues 60 and 61 can cross-link with RNAPII subunit Rpb2, indicating that these lysines may be important for TFIIF-RNAPII interaction (5). It is possible that PTM of these two lysines can affect this interaction during transcription. Indeed, Lysine 61 was also found to be acetylated and to serve as a potential ubiquitination target with high confidence (44, 45). All these findings directed our focus to these two lysine residues. Aiming to test whether Tfg1 Lysine 60 and 61 (K60,61) can be sumoylated, a yeast strain with a mutated version of 6HA-tagged Tfg1 was generated. In this strain, both lysines were mutated to arginine (K60,61R), thus preventing these residues from becoming sumoylated. Then, this strain was compared with a yeast strain containing 6HA-tagged wild-type Tfg1. Immunoblotting of HA showed that similar levels of Tfg1 and Tfg1 mutant were pulled down through IP (**Fig 13**). However, as shown in the SUMO immunoblot of the IP, the K60,61R mutation almost abolished Tfg1 sumoylation, suggesting that lysine 60 and/or 61 is the major sumoylated site of Tfg1. Because we were unable to successfully generate yeast strains with either a K60R or K61R single mutation, it remains unclear whether both lysines are sumoylated or only one of them is modified. Nonetheless, we successfully narrowed down the major Tfg1 sumoylation site to Lysine 60 and 61.

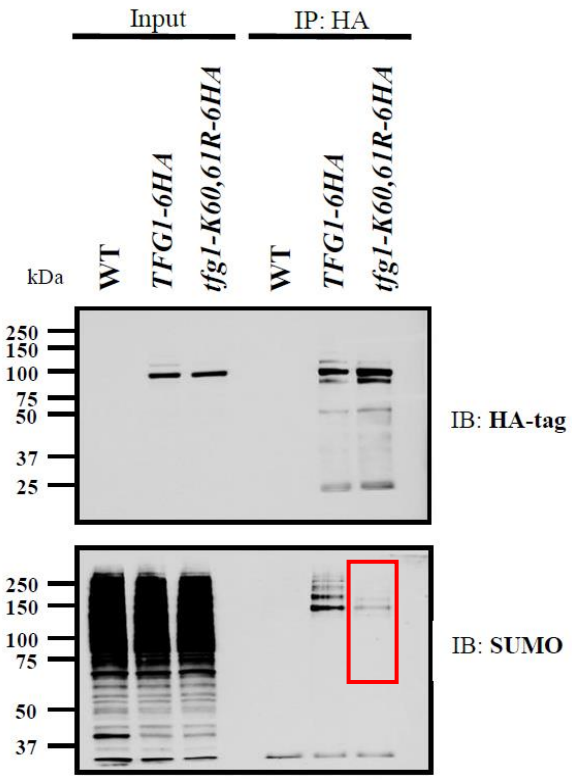



**Figure 13: Lysine 60 and/or 61 is the major sumoylated site of Tfg1.** Yeast strains with *TFG1-6HA* and *tfg1-K60,61R-6HA* genes respectively were used in an HA-IP which was followed by immunoblotting for Smt3 (yeast SUMO) and HA-tag. HA blots confirm that the immunopurification of Tfg1 was successful. The Smt3 (SUMO) blot shows that K60,61R (lysine to arginine) mutation dramatically reduces Tfg1 sumoylation (red rectangle). A wild-type yeast strain (W303) with no 6HA-tagged proteins was used as a control. (Result obtained from Akhi Akhter)

### 3.3.2 The K60,61R Mutation Does not Affect Yeast Growth

Previously we showed that different stresses do not affect Tfg1 sumoylation (**Fig 12**). Since Tfg1 K60,61 was identified as the predominant sumoylation site, we wished to study whether blocking Tfg1 sumoylation by the K60,61R mutation affects yeast cell growth patterns under normal and stress conditions. To answer this, a yeast growth assay (drop test) was performed. A pair of yeast strains, *TFG1-6HA* and *tfg1-K60,61R-6HA*, were grown to the same culture density and serial dilutions were prepared. Droplets of the serially-diluted yeast cultures were plated onto media plates (minimal medium, SC) with different stress reagents or on an SC plate incubated at 37°C. The stress reagents include salt (1 M NaCl and 1 M KCl), alcohol (7% EtOH), peroxide (1 mM $H_2O_2$) and an amino acid biogenesis inhibitor (10 µg/mL SM). We observed a very

similar growth pattern for both strains under all stress conditions, as well as in two control conditions (rich medium, YPD and minimal medium, SC) after incubating for three days (**Fig 14**).

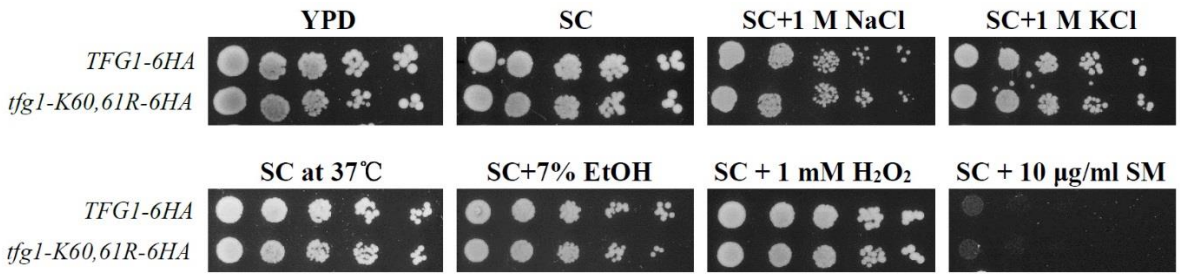



**Figure 14: Sumoylation-deficient Tfg1 mutant presents no defect in growth.** Yeast strains expressing Tfg1-6HA or Tfg1-K60,61R-6HA were cultured for drop test preparation. Cultures underwent a serial dilution and then were spotted either on SC medium plates with different stress reagents or on an SC plate incubated at 37°C. The stress conditions, including salt stress (1 M NaCl and 1 M KCl), alcohol stress (7% EtOH), heat shock (37°C), oxidative stress (1 mM $H_2O_2$) and amino acid starvation stress (10 μg/mL SM), are labeled. Images of yeast colonies were captured by scanning after incubation at 30°C or 37°C for 3 days. Colonies inoculated on YPD and SC plates serve as control growth conditions. Both yeast strains show similar growth pattern under each stress condition indicating that the K60,61R mutation does not affect yeast growth rate.

Interestingly, SM strongly inhibited the cell growth of both strains to a same extent. Aside from this observation, our drop test result shows that blocking Tfg1 sumoylation by mutating K60,61 does not affect the growth rate of yeast.

### 3.4 Effects of Tfg1 Sumoylation on the TFIIF-RNAPII Interaction

As mentioned above, Tfg1 K60,61 was revealed as one of the contact sites for TFIIF-RNAPII interaction during transcription initiation (1, 5). To test how the K60,61R mutation affects this interaction, in particular the Tfg1-RNAPII association, reciprocal IPs were conducted. HA-IP allowed us to pull down Tfg1 and associated proteins, such as RNAPII subunits Rpb1 and Rpb2. The 8WG16 antibody, which recognizes the RNAPII CTD repeat YSPTSPS of Rpb1, was used to IP and detects RNAPII. When comparing Tfg1-6HA and Tfg1-K60,61R-6HA IPs, although there is a lower level of sumoylated Tfg1 in the *tfg1-K60,61R-6HA*

strain, the amount of Tfg1-associated RNAPII was not different (**Fig 15A**). A similar result was observed when we pulled down Rpb1 and detected associated Tfg1. The *TFG1-6HA* wild-type strain and *tfg1-K60,61R-6HA* mutant strain exhibited the same level of RNAPII-interacted Tfg1 (**Fig 15B**). Thus we reasoned that blocking Tfg1 sumoylation has no effect on the Tfg1-RNAPII interaction.

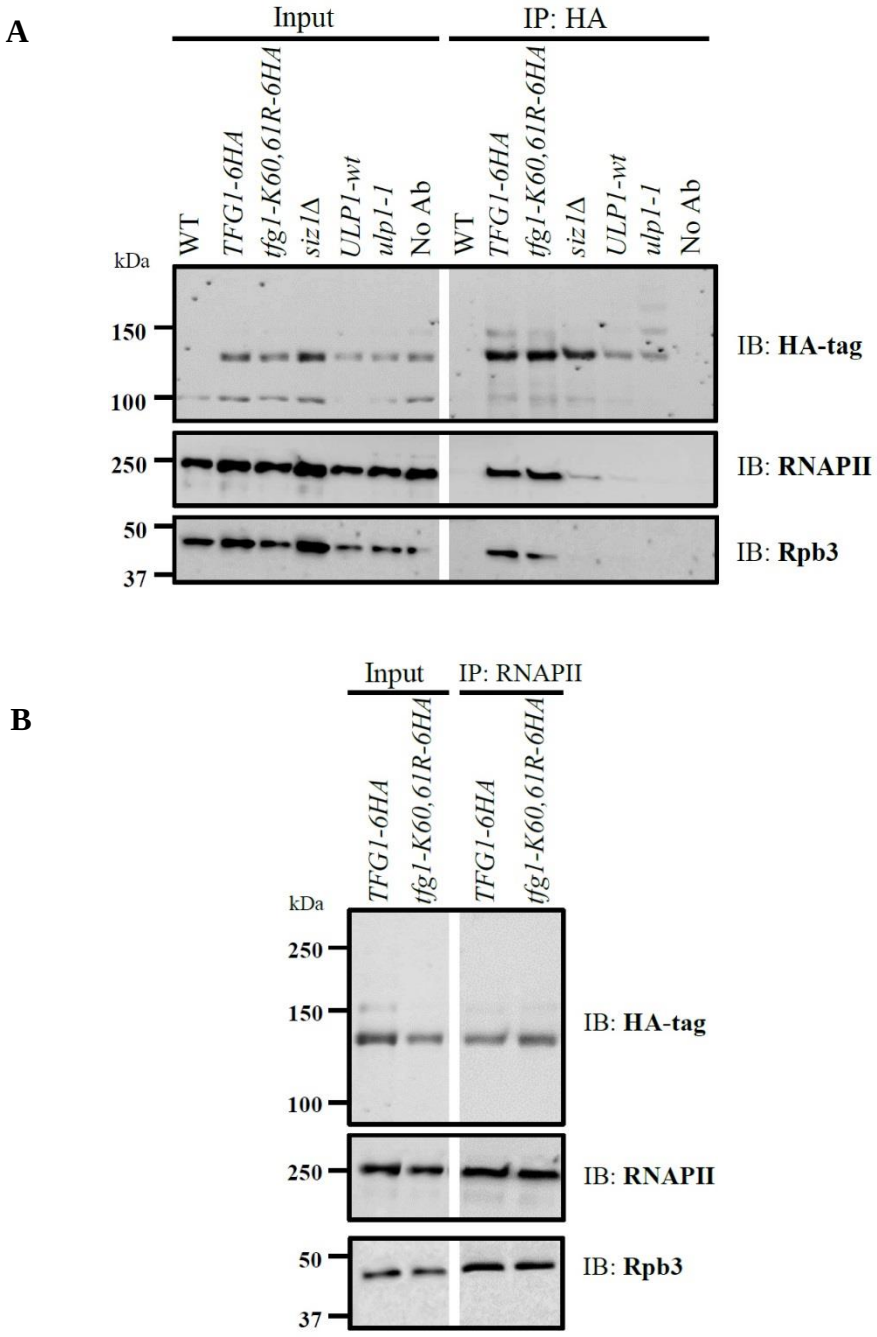



**Figure 15: The effects of Tfg1 sumoylation on TFIIF-RNAPII interaction.** Yeast strains with either increased (*ulp1-1 TFG1-6HA*) or decreased (*tfg1-K60,61R-6HA, siz1*Δ *TFG1-6HA*) Tfg1 sumoylation were tested by reciprocal IP, using an HA antibody and 8WG16 antibody, which recognizes the largest subunit of RNAPII, Rpb1 (HA-IP [panel **A**] and RNAPII-IP [panel **B**]). IPs were then analyzed by immunoblot with HA and 8WG16 antibodies, as well as with an antibody that recognizes the third largest subunit of RNAPII, Rpb3. Controls included a wild-type strain with no HA-tag (WT), the *ULP1-wt* strain, and a no antibody control (No Ab) performed using the *ulp1-1* strain.

Another way to potentially block Tfg1 sumoylation is through deleting *SIZ1*, a nonessential gene that encodes a major yeast SUMO E3 ligase. A yeast strain with a *SIZ1* deletion (*siz1*Δ) could theoretically inhibit sumoylation of Tfg1 at all acceptor sites, in addition to Lysine 60 and 61. In fact, the *siz1*Δ strain showed a near complete blocking of Tfg1 sumoylation, as determined by the apparent absence of higher-molecular weight forms of Tfg1 in an HA immunoblot of IPed Tfg1-6HA (**Fig 15A**). Interestingly, in the *siz1Δ* strain, a significantly lower level of Tfg1-associated RNAPII was detected (**Fig 15A**). To further determine the effects of altered Tfg1 sumoylation on TFIIF-RNAPII interaction, additional yeast strains were tested in this reciprocal IP: the *ulp1-1* strain that expresses a mutant form of the desumoylation enzyme Ulp1 and its isogenic *ULP1* wild-type counterpart. As expected in the *ulp1-1* mutant strain, a higher level of sumoylated Tfg1 was detected, as determined by increased levels of higher molecular weight forms of Tfg1 in the HA blot (**Fig 15A**). Notably, unlike other HA-tagged strains used, less Tfg1, both unmodified and modified forms, was pulled down in *ULP1* parental strain and *ulp1-1* mutant strain (**Fig 15A**). Consequently, in the *ULP1* parental strain, a faint coIPed RNAPII signal was detected in the Tfg1-6HA IP, whereas no RNAPII band was detected in the *ulp1-1* strain. Because the signal was too faint to be detected, quantification of immunoblot results was not able to be carried out. In order to improve this results and to examine the effect of elevated Tfg1 sumoylation on Tfg1-RNAPII interaction, following this experiment, several trials of HA-IP were repeated aiming to obtain quantifiable results. However, in all trials, a strong background was observed in no-antibody control (No Ab), on both HA blot and RNAPII blot, which largely reduces the confidence of quantification analysis. As a result, the clear effect of increased Tfg1 sumoyaltion by mutating desumoylation activity of Ulp1 on Tfg1-RNAPII interaction remains unrevealed.

In spite of the alterations of Tfg1-RNAPII interaction observed in the *siz1*Δ mutant strain, we have to consider the global effects induced by this mutation. In this case, not only Tfg1 sumoylation is changed, but sumoylation of numerous other target proteins will also be affected. These proteins include, for example, other GTFs and RNAPII subunits. Thus the changes in Tfg1-RNAPII interaction in these yeast mutants may be contributed by the alteration of sumoylation of other substrates. Although our data suggests that sumoylation of Tfg1 specifically does not affect the TFIIF-RNAPII interaction, it appears that global

changes to sumoylation levels can regulate the interaction and, therefore, possibly transcription.

### 3.5 Effects of Tfg1 Sumoylation on Gene Transcription

#### 3.5.1 Effects of Tfg1 Sumoylation on TFIIF and RNAPII Association with Chromatin

As described previously, both TFIIF and RNAPII bind to promoters to induce gene transcription. RNAPII is associated with the DNA template during the entire transcription process (initiation, elongation and termination) whereas TFIIF mainly functions during initiation and transiently facilitates early elongation steps (3, 46). According to this, we wanted to determine the effect of changed Tfg1 sumoylation on TFIIF and RNAPII chromatin occupancy. Chromatin immunoprecipitation (ChIP) was performed followed by quantitative polymerase chain reaction (qPCR) analysis. We tested relative chromatin occupancies of Tfg1-6HA, Tfg1-K60,61R-6HA, and RNAPII on constitutively transcribed genes *PMA1*, at its promoter and ORF (open reading frame), and *ADH1*, at its promoter. Additionally, a constitutively silent gene, *STL1* (osmosis-responding gene), was included as a negative control. At *PMA1* and *ADH1* promoters, both Tfg1-6HA and Tfg1-K60,61R-6HA showed significantly high occupancy compared to background signals detected at input samples and in a control strain that lacks an HA tag, indicating that 6HA-tag was efficiently IPed, as expected (**Fig 16**). Further supporting that the ChIP procedure was effective, only background levels of Tfg1 and RNAPII were detected on the silent *STL1* promoter (i.e. equal to the input control). Blocking Tfg1 sumoylation, however, did not have an obvious implication on TFIIF-DNA interaction at both *PMA1* and *ADH1* promoters, as the level of occupancy for wild-type Tfg1 and the K60,61R mutants were not statistically significantly different with the exception of a minor reduction of Tfg1-PMA1 ORF association in *tfg1-K60,61R-6HA* mutant strain (**Fig 16**). Similarly, RNAPII-DNA association at the *ADH1* promoter and the *PMA1* promoter and ORF was not affected by the sumoylation-blocking mutations of Tfg1. In summary, our data suggests that Tfg1 sumoylation has no apparent significant effect on both TFIIF- and RNAPII-chromatin association. However, the ChIP quantification results include large error bars which reflect a lack of consistency among the six independent ChIP-qPCR replicates that we performed. This

prevents us from drawing firm conclusions from this analysis.

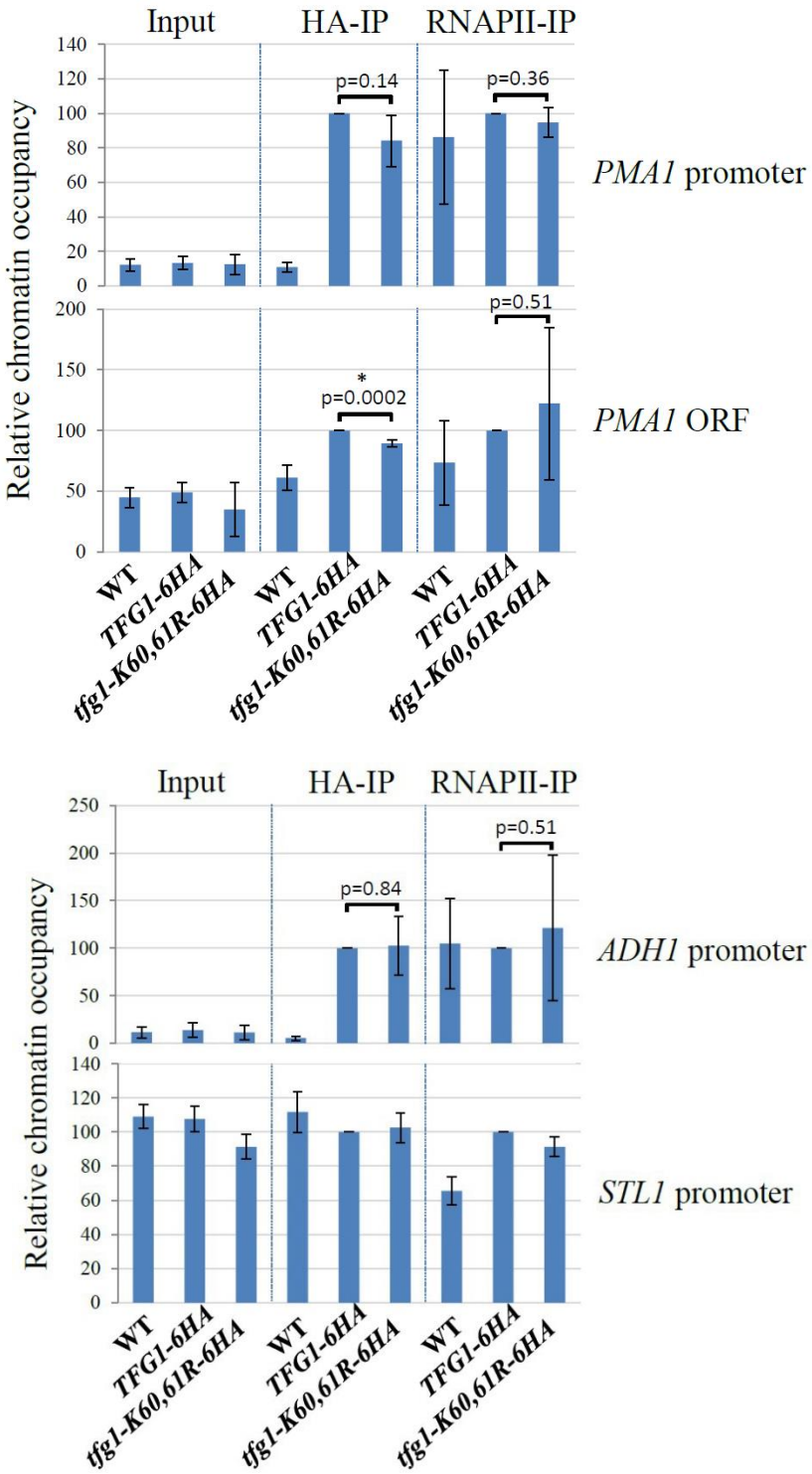


**Figure 16: Effect of Tfg1 sumoylation on TFIIF and RNAPII chromatin occupancy.** ChIP followed by qPCR amplification was used to analyze how altered Tfg1 sumoylation levels affect TFIIF and RNAPII association with chromatin, using HA and 8WG16 antibodies. Yeast strains used include a wild-type strain with no HA-tag (WT) and the *TFG1-6HA* and *tfg1-K60,61R-6HA* strains. Bar graphs show relative chromatin occupancies of TFIIF and RNAPII at the *PMA1* (housekeeping gene) promoter and ORF, and the *STL1* (osmosis-responding gene) and *ADH1* (constitutive gene for alcohol dehydrogenase) promoters. *STL1* was used as a negative control, as it is repressed under normal growth conditions.promoter was the control for repressed gene. Results were normalized to a non-transcribed region of Chromosome V (ChrV). ChIP-qPCR data and analyzed by Excel. Data are represented as mean +/- standard deviation of six independent experiments with p values labeled above. Asterisk indicates statistically significant values (Student's t-test with $p < 0.05$).

### 3.5.2 Effects of Tfg1 Sumoylation on RNA Expression

To explore how Tfg1 sumoylation alters gene transcription, we measured steady state mRNA levels of a number of sample genes in yeast strains expressing Tfg1-6HA or Tfg1-K60,61R-6HA by RT-qPCR. Strains were grown either at the normal growth temperature (30°C) or at an elevated temperature (37°C), to examine effects of both constitutively expressed genes and genes induced by heat shock (hs). Under normal growth conditions, blocking Tfg1 sumoylation resulted in an apparent higher level of RNA expression of all genes of interest (**Fig 17**). Statistical analysis (Student's *t*-test) of three replicates, however, indicated that only *SSA4* (hs-induced gene) and *PYK1* (constitutive gene) showed a statistically significant difference when grown at 30°C. As expected, heat shock treatment led to strong activation of hs-induced genes (*HSP12* and *SSA4*) and a significantly lower RNA expression of hs-silenced genes (*PHO5* and *SAH1*) in both yeast strains compared with the result of normal growth condition. In contrast, the RNA abundance of constitutive genes (*PMA1* and *PYK1*) were not changed to the same extent. Although our data suggests that Tfg1 sumoylation has an inhibitory role in RNA expression at some genes, additional repeats of this analysis will be needed to determine whether this is a statistically significant general effect of Tfg1 sumoylation.

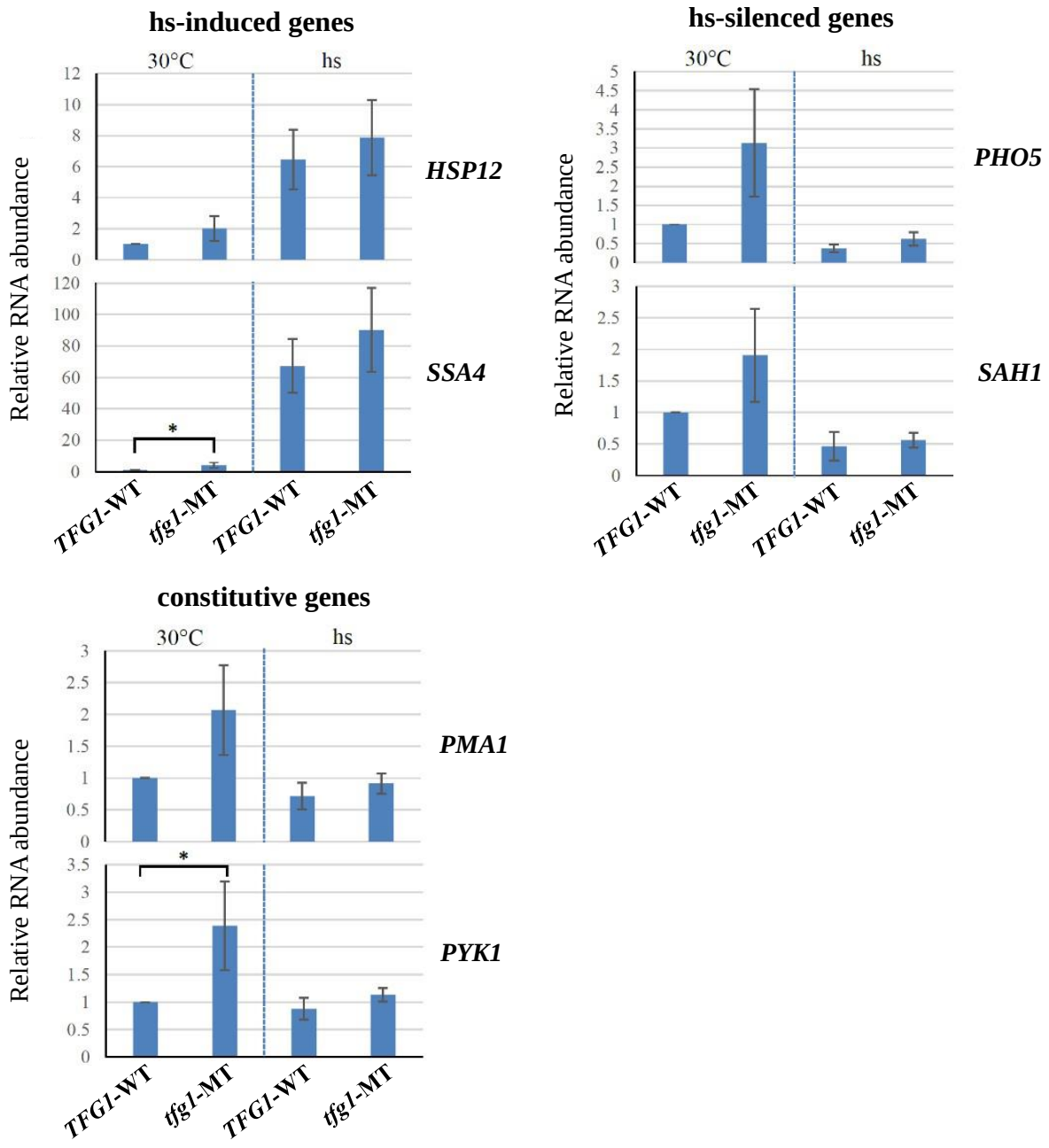



**Figure 17: Effect of Tfg1 sumoylation on RNA expression under normal growth and heat shock (hs) conditions.** Yeast RNA extraction followed by RT-qPCR amplification was done to analyze how RNA abundance changes as a result of decreased Tfg1 sumoylation. Yeast strains used include *TFG1-6HA* (*TFG1-WT*) and *tfg1-K60,61R-6HA* (*tfg1-MT*) strains. RNA levels were examined at the ORF regions of heat shock-induced genes (*HSP12* and *SSA4*), heat shock-silenced genes (*PHO5* and *SAH1*) and constitutive genes (*PMA1* and *PYK1*). Normal growth temperature is 30°C and heat shock temperature is 37°C. RNA levels are presented normalized to the 25S rRNA and relative to their level in the Tfg1-WT-expressing strain. Data are represented as mean +/- standard deviation of three independent experiments with asterisk indicating statistically significant values (Student's t-test with $p < 0.05$).

### 3.6 Attempting to Generate an *SMT3-TFG1* Fusion Yeast Strain

An *ulp1-1* strain was used to study the effects of increased Tfg1 sumoylation levels in our previous experiments. However, no clear result of how reduced desumoylation activity in the *ulp1-1* mutant affects Tfg1-RNAPII interaction was obtained. In addition, the effect of mutating Ulp1 is very likely to be widespread, and not limited to just Tfg1. To improve our results by ruling out the global effects caused by malfunctioning Ulp1, we attempted to generate a yeast strain expressing an Smt3-Tfg1 fusion protein. This fusion would allow us to simulate permanently sumoylated Tfg1 without affecting other SUMO targets. Thus, this new strain, once obtained, could be used for investigating the effects of increased Tfg1 sumoylation on multiple aspects of gene transcription, including TFIIF-RNAPII interaction, chromatin association and RNA expression.

The first strategy used to generate the fusion is shown in **Fig 18A**, and involves two yeast transformations (**Fig 18A**). A *URA3* marker gene was amplified by PCR and then transformed into diploid yeast strain W303 which contains two copies of wild-type *TFG1* gene, aiming to replace one copy of *TFG1* gene through homologous recombination. The presence of *URA3*, which encodes orotidine 5'-phosphate decarboxylase (ODCase), allows successfully transformed yeast cells to grow on standard SC medium plates lacking uracil (SC-URA). Thus, positive selection was performed by taking the advantage of *URA3* marker. Three colonies were taken from the SC-URA transformation plate and tested for the presence of *URA3* gene by yeast colony PCR. An image of the agarose gel showed that all three colonies (named yYD004A, B and C) were successful transformed (**Fig 19A**).

Before the second transformation, we needed to construct a DNA fragment that consists of the *SMT3* open reading frame fused to the *TFG1-6HA* gene sequence (*SMT3-TFG1-6HA*) and including flanking sequences ("homology arms") that match sequences surrounding the *TFG1* locus, which are needed for homologous recombination at that site (**Fig 18A**). *SMT3* and *TFG1-6HA* genes were amplified respectively by PCR and validated using agarose gel electrophoresis (**Fig 19B**). The products were then used as templates for fusion PCR. To gain a best fusion PCR efficiency, three amplifications with different *SMT3*/*TFG1-6HA* template

ratios were carried out. Fusion PCR product from all three reactions were examined by gel electrophoresis (**Fig 19B**). Among all amplifications, a template ratio of 1:2 (*SMT3* to *TFG1-6HA*) resulted in the best PCR efficiency. After purification, the fusion PCR product was introduced into yYD004A cells which were randomly selected from the first transformation.

Since the *URA3* gene product ODCase converts 5-fluoroorotic acid (5-FOA) into a toxic compound leading to cell death, SC medium containing 5-FOA was used to select cells in which the *URA3* marker was successfully replaced by the *SMT3-TFG1-6HA* gene. After transformation, two colonies (named Colony A and B) were observed on the 5-FOA-containing medium plate and were tested for the presence of either the *SMT3-TFG1-6HA* fusion gene or *URA3* by isolating genomic DNA (gDNA) and performing gene-specific PCRs. In additional to these colonies, gDNA of yYD004A was examined as a positive control for *URA3* gene and a negative control for the fusion gene. Neither of the colonies derived from successful transformation, as indicated by the absence of fusion gene (**Fig 19C**). Moreover, the absence of *URA3* gene in both colonies indicates that the *URA3* marker was indeed lost during the transformation, but not through replacement with the *SMT3-TFG1-6HA* cassette. As a result of this failed transformation, and multiple additional unsuccessful attempts, this strategy for generating the Smt3-Tfg1 fusion protein-expressing yeast strain was not pursued.

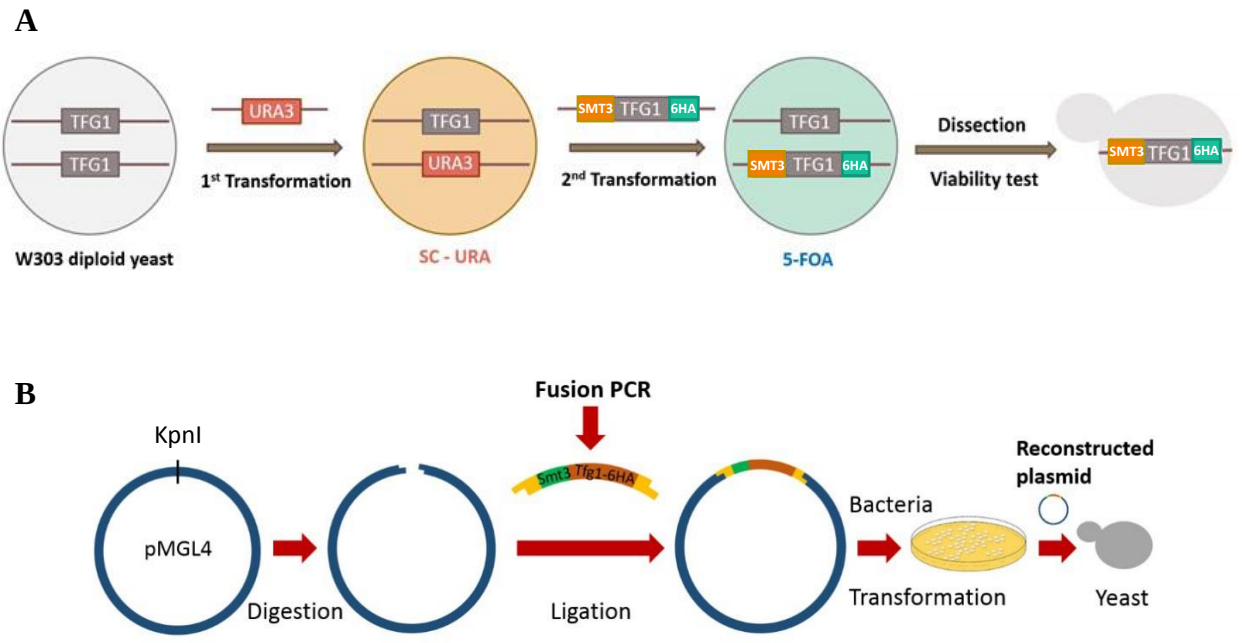



**Figure 18: Schematics for generating Smt3-Tfg1 fusion protein-expressing yeast strain.** Two cloning methods were attempted to generate a new yeast strain expressing Smt3-Tfg1-6HA. **A**) A diploid yeast strain W303 containing two copies of wild-type *TFG1* gene is used in the first transformation. The *URA3* marker, which encodes orotidine 5'-phosphate decarboxylase (ODCase), is introduced to replace one copy of the *TFG1* gene. *URA3*-expressing yeast is then selected on medium lacking Uracil (SC-URA). The second transformation replaces the *URA3* marker with a copy of the *SMT3-TFG1-6HA* fusion gene. Successfully transformed yeast cells are then selected using medium containing 5-fluoroorotic acid (5-FOA) because ODCase converts 5-FOA into a toxic substrate leading to yeast cell death. Selected diploid cells are then sporulated and haploid progeny are tested for viability. **B**) Vector plasmid pMGL4 is digested with the KpnI restriction enzyme. The fusion PCR product (*SMT3-TFG1-6HA*), with KpnI sticky ends, is then be ligated to the digested plasmid. Reconstructed plasmid is transformed into bacteria. The fusion gene is then amplified in bacteria and collected by miniprep-scale DNA preparation. Lastly, collected products are transformed into yeast cells allowing these cells to express Smt3-Tfg1-6HA fusion protein.

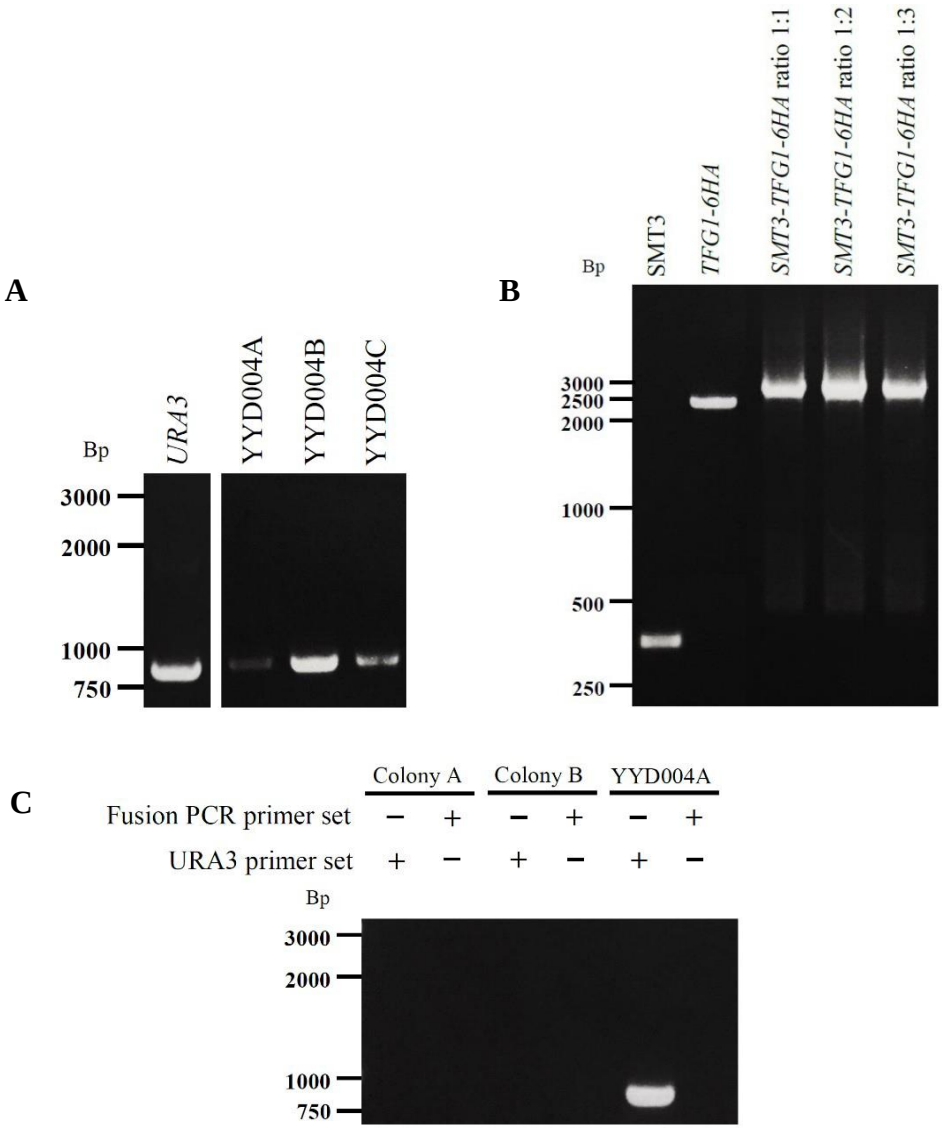



**Figure 19: Generation of *SMT3-TFG1-6HA* yeast strain by transformation with fusion PCR product.** Two yeast transformations were involved in this cloning procedure. **A**) The *URA3* marker (804bp) was amplified by PCR using plasmid pRS316 as template and validated by agarose gel electrophoresis. This PCR product was then transformed into the diploid yeast strain W303, and transformation product strains, named YYD004A, B and C, were tested for the presence of *URA3* gene by colony PCR. All three strains showed successful transformation. **B**) Genes for *SMT3* (~306bp) and *TFG1-6HA* (~2.5kbp) were amplified by PCR separately and then used as templates for fusion PCR. Three Smt3 to Tfg1-6HA template ratios were used: 1:1, 1:2 and 1:3. Obtained fusion PCR products were run on an agarose gel to validate the size of the DNA fragments. A band approximately 2.7kbp in length was observed indicating a successful fusion PCR. **C**) Fusion PCR product (*SMT3*/*TFG1-6HA* template ratio 1:2) was used to transform strain YYD004A. Two colonies (Colony A and B) were observed on plate and tested for the presence of *SMT3-TFG1-6HA* fusion DNA sequence. Neither *URA3* marker gene nor the fusion gene was present in either colony. Genomic DNA of YYD004A served as a positive control of for the presence of the *URA3* gene.

An alternative procedure was attempted, in which the *SMT3-TFG1-6HA* cassette would be cloned into an expression plasmid, which could then be transformed into yeast (**Fig 18B**). In this strategy, the fusion cassette would be generated with flanking sequences containing the KpnI restriction site, and the product would be subcloned into plasmid vector pMGL4 at the KpnI site. To validate KpnI digestion of the vector, uncut and digested versions of the plasmid were examined by agarose gel electrophoresis side by side. KpnI digestion resulted in a linear form of pMGL4 (~7.14kbp) which moved slight slower than the uncut supercoiled form (**Fig 20A**). A DNA band shown at the very top of the gel indicated an open-circular form of plasmid. Then, *TFG1-6HA* (~2.5kbp) and *SMT3-TFG1-6HA* (~2.7kbp) gene, both with flanking KpnI restriction sites, were obtained by normal PCR and fusion PCR, respectively, as we did in the first cloning procedure. *TFG1-6HA* gene will serve as a control in later steps. After purification, both PCR products were digested with KpnI, but analysis by electrophoresis showed that, although *TFG1-6HA* and *SMT3-TFG1-6HA* genes were successfully amplified, KpnI digestion led to the appearance of unexpected products, indicated by smears of bands located below 2000bp marker (**Fig 20B**). To eliminate these unwanted DNA products, we performed gel extraction to purify the desired gene fragments. These purified DNAs would then be used to ligate with KpnI-digested pMGL4. However, the concentration of the purified insert was too low to be useful for ligation. Because of the low yield of the purified insert fragments, and because of the appearance of the unidentified digestion products, we decided not to proceed with this strategy until these issues are resolved.

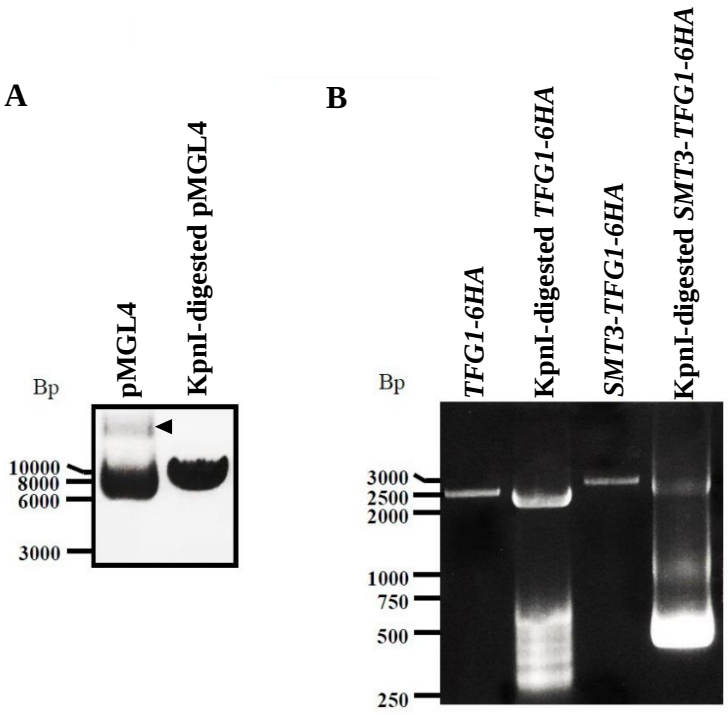



**Figure 20: Generation of *SMT3-TFG1-6HA* yeast strain by transformation with plasmid.** Inserts were prepared for generating a recombinant plasmid in this cloning procedure. **A**) Plasmid pMGL4 (~7.14kbp) was used as vector and digested by the KpnI restriction enzyme. Both intact plasmid and the digested version were analyzed by agarose gel electrophoresis. The open-circular form (black arrow) and the supercoiled form of pMGL4 were observed on gel. KpnI-digested plasmid was in a linear form and thus ran slightly slower than the supercoiled form. **B**) *TFG1-6HA* (~2.5kbp) and *SMT3-TFG1-6HA* (~2.7kbp) DNA fragments were amplified by PCR and then digested by KpnI. Both undigested and digested DNA pieces were validated by an agarose gel where digested DNA fragments ran slightly faster than the original copies. Unspecific bands of unknown origin were detected, possibly indicating unexpected cutting occurred during KpnI digestion.

### 3.7 Characterizing SF3B1 Sumoylation

The experiments described above focused on effects of sumoylation during transcription initiation in yeast, by examining GTFs and RNAPII. Moving on, we were also curious about eukaryotic cells use SUMO modifications to regulate pre-mRNA splicing. Specifically, we were interested in investigating the effects of sumoylation of a splicing factor on pre-mRNA processing events. A previous proteomics study, in which numerous putative SUMO conjugates were identified in HeLa cells, determined that the human protein SF3B1, a critical component of the U2 snRNP splicing complex, has a high probability of being sumoylated

(28, 37). Thus, we intended to confirm and characterize SF3B1 sumoylation in HeLa cells. We obtained an antibody for SF3B1, and performed IP-immunoblot analysis to detect sumoylated forms of the splicing factor. A conspicuous band around 150kDa was observed in both input and IP samples of SF3B1 blot confirming the successful pulldown of SF3B1 (**Fig 21**). However, although sumoylated proteins were detected in input samples, immunblots of the SF3B1 IP with SUMO1 and SUMO2 antibodies showed no signals, indicating that there is no detectable modified SF3B1 in these cells. Because the absence of a sumoylated SF3B1 band may due to the undetectable low level of endogenous modified SF3B1, we repeated the analysis in other HeLa-derived cell lines that overexpress His-tagged SUMO1 or SUMO2, which we anticipated could elevate SF3B1 sumoylation levels. First, both His-tagged cell lines were validated for their SUMO overexpression level. Because of the contamination issue of the wild-type HeLa cell culture, we used wild-type HEK 293T cell line as the control in immunoblot analysis instead. Compared with *His-SUMO1* HeLa cells, which showed a slightly higher level of SUMO1 than the wild-type HEK 293T cells, *His-SUMO2* HeLa cells showed a significant higher level of SUMO2 expression than wild-type cells (**Fig 22A**). As a result, a following IMAC taking the advantage of His-tag affinity was carried out using *His-SUMO2* HeLa cell line and wild-type HeLa cell line. Enrichment of His-tag signal observed on His-tag blot indicated a successful purification of His-tagged targets (**Fig 22B**). A higher level of SUMO2, although shown as a smear, was detected in *His-SUMO2* HeLa cells indicating the successful overexpression. Unfortunately, no sumoylated SF3B1 was IPed in either wild-type HeLa cells or SUMO2-overexpressing HeLa cells indicating using SUMO2-overexpressing cells did not improve the IP procedure in this experiment. In summary, we were unable to confirm that SF3B1 is sumoylated in mammalian cells.

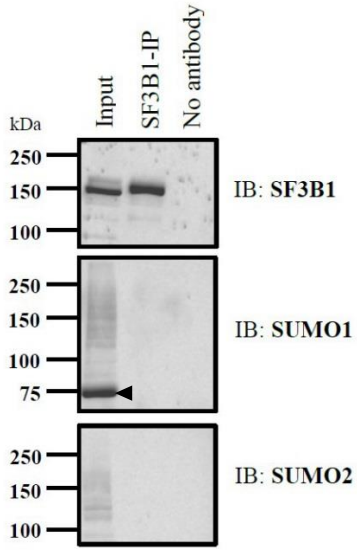



**Figure 21: Characterization of SF3B1 sumoylation.** SF3B1 was immunopurified from Hela cells and analyzed by immunoblot with antibodies for SF3B1, SUMO1 and SUMO2. Although SF3B1 was successfully purified, neither SUMO1 nor SUMO2 signals were observed in the IP, indicating that no detectable sumoylated SF3B1 was found. Typical SUMO1 and SUMO2 smears are present in input samples, including a strong band (black arrow) which corresponds in size to Ran GTPase Activating Protein 1, which is one of the most heavily sumoylated proteins known. A no antibody IP control was included.

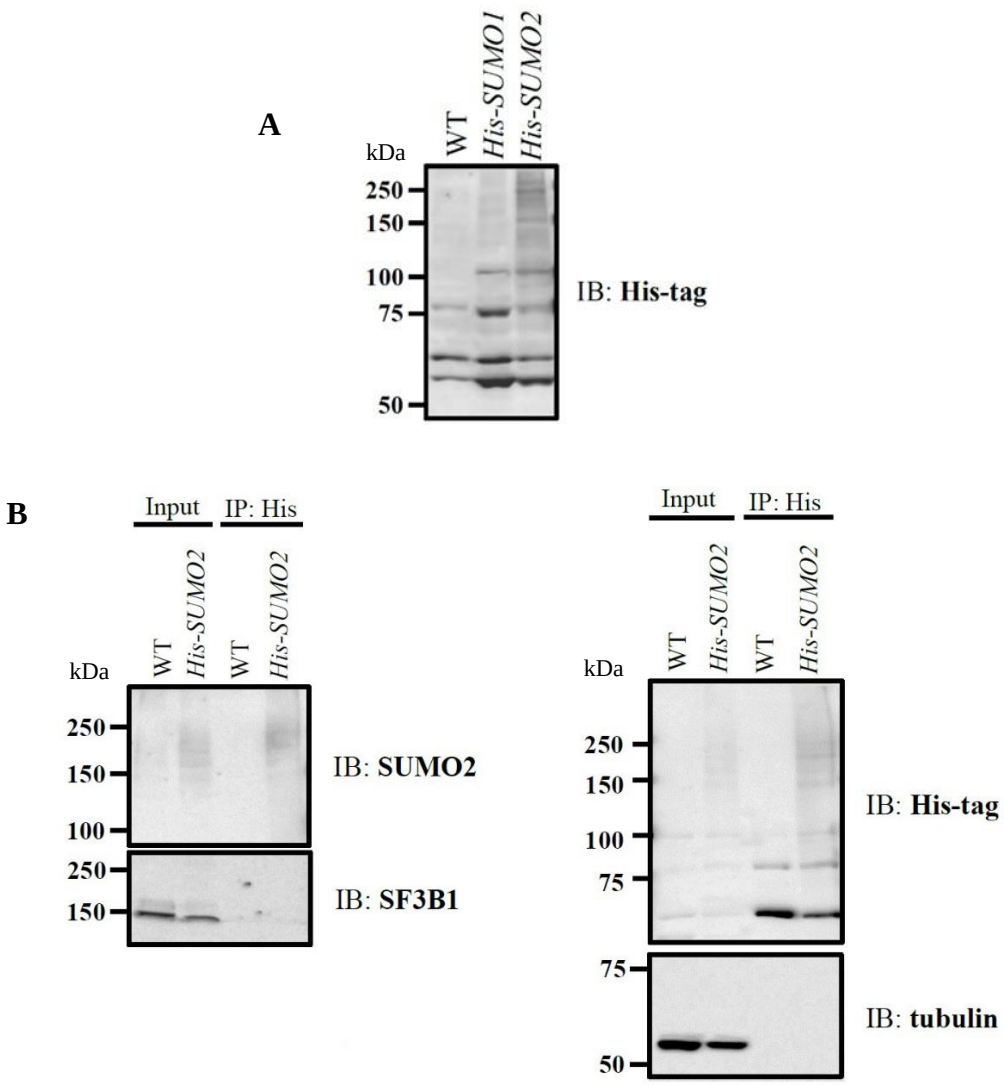



**Figure 22: Characterization of SF3B1 sumoylation in SUMO2-overexpressing cell line. A**) Validation of *His-SUMO1* and *His-SUMO2* HeLa cell lines by western blotting His-tag. Cell lysate of HEK 293T cell (WT) served as control. *His-SUMO2* HeLa cells showed a significant overexpression of SUMO2. **B**) Proteins modified by His-SUMO2 were purified by IMAC and analyzed by immunoblot with antibodies for SUMO2, SF3B1, His-tag and tubulin (loading control). Wild-type HeLa cell line (WT) served as a control. *His-SUMO2* HeLa cells showed an enriched His-tag signal indicating a successful purification of His-tagged targets. Overexpression of SUMO2 was observed as expected in *His-SUMO2* HeLa cells. However, SF3B1 was only observed in input indicating a failed IP. Tubulin signal was only observed in input as expected.

# Chapter 4: Discussion

Sumoylation is a post-translational modification (PTM) modifies proteins by a reversible covalent conjugation process. Through attachment of SUMO to acceptor proteins, protein-protein interactions, protein stability and intracellular localization can be altered. This modification is largely involved in regulating multiple cellular events including cell cycle progression, transcription and DNA repair. In particular, one of the largest groups of SUMO target includes proteins involved in transcription and its regulation. In most cases, the effect of sumoylation of transcription factors is the repression of gene expression. However, it may also activate transcription in a gene-dependent manner. Previous proteome-wide studies identified multiple GTF subunits, including TFIIF subunit Tfg1, as highly confident SUMO targets (**Table 2**). Thus, we tested multiple GTF subunits for their level of sumoylation. We used a previously-studied and confirmed SUMO target, Rpb1, as a positive control (31). Our result showed Tfg1 has a much higher sumoylation level compared with other GTF subunits in normally growing yeast cells (**Fig 9**). This gives us a clue of the involvement of sumoylation in regulating RNAPII-mediated transcription.

To benefit our investigation on how changes in cellular sumoylation levels can affect transcription, in particular its effect on the TFIIF-RNAPII interaction, we aimed to characterize Tfg1 sumoylation first. We found that, among all three TFIIF subunits, Tfg1 is the most sumoylated (**Fig 10**). This not only supports our result which showed Tfg1 has a higher sumoylation level compared with other GTF subunits, but also validated the prediction of others that Tfg1 is a SUMO target with high potential (**Table 2**). This result suggests that sumoylation regulates TFIIF function mainly through Tfg1. As Tfg1 interacts directly with RNAPII, and since the interaction of TFIIF and RNAPII is critical during transcription, much of our research focused on assessing Tfg1 sumoylation and its effect on the Tfg1-RNAPII interaction.

To deeply understand Tfg1 sumoylation, we examined it from multiple angles. We determined that the same fraction of both chromatin-associated and soluble Tfg1 is sumoylated (**Fig 11**). Furthermore, unlike what others have found, in which the dynamics of intracellular localization of certain transcription factors is

regulated by sumoylation, increased Tfg1 sumoylation did not significantly alter its chromatin-association pattern (47). As a result, we do not believe that sumoylation affects the subcellular localization of Tfg1, nor is it restricted to transcriptionally-engaged forms of TFIIF.

Moreover, we tested the effects of multiple stress conditions on Tfg1 sumoylation. The transcriptome is known to be significantly affected by different environmental factors, such as temperature and oxidizing compounds (48). The alteration of gene transcription patterns may be due to the modification of the association of TFIIF with RNAPII during transcription initiation, which we speculated might be affected by the level of sumoylated Tfg1. The stresses we performed included high salt, alcohol, heat shock, oxidative stress and amino acid starvation. By measuring the ratio of sumoylated Tfg1 over total Tfg1 level, we showed that the proportion of modified Tfg1 was not influenced by these stress factors (**Fig 12**). Importantly, even 10% EtOH, under which the cell growth rate was significantly reduced, did not lead to an obvious change in this ratio compared with all other conditions. This revealed that the level of Tfg1 sumoylation is likely not involved in stress-induced changes to the transcriptome. Additionally, we used the same stress conditions to test for differences in yeast growth rates caused by inhibition of Tfg1 sumoylation. This took the advantage of our previous generated yeast strain which contains a Tfg1-K60,61R mutated protein. In this strain, the Tfg1 sumoylation is almost completely blocked due to lose of two lysine residues at the major SUMO site and thus it served as a comparison that has depleted Tfg1 sumoylation (**Fig 13**). After inoculating for three days, the growth pattern, mainly the size of the yeast colonies, was observed to have no difference between *TFG1* wild-type and mutant strains under all stress conditions (**Fig 14**). Conclusively, we have demonstrated that altering the level of Tfg1 sumoylation does not affect the cellular response to a number of stress conditions, which is consistent with our finding that Tfg1 sumoylation levels do not change under those conditions.

Intriguingly, Lysine 60 and 61 of Tfg1, which is the major site of sumoylation, was previously found to be the cross-linking site between Tfg1 and RNAPII subunit Rpb2 (5). We hypothesized that sumoylation of Tfg1 Lys 60 and 61 would interrupt its interaction with RNAPII leading to repression of gene transcription. Furthermore, we expected that the K60,61R mutation would also impair the interaction and impair

transcription. To validate this, we performed reciprocal Tfg1 and RNAPII IPs. Surprisingly, blocking Tfg1 sumoylation by K60,61R mutation did not affect Tfg1-RNAPII association (**Fig 15**). This suggests that, although Lys 60 an 61 may be involved in the association of TFIIF with RNAPII, they are not critical for the interaction which must be stabilized by additional contacts between the proteins.

To assess how modulating Tfg1 sumoylation can affect the TFIIF-RNAPII interaction without altering the Tfg1 amino acid sequence, we performed IP experiments in strains lacking the SUMO ligase Siz1 or expressing a partially defective form of the desumoylation enzyme Ulp1 (**Fig 15**). Deletion of *SIZ1* gene effectively blocked Tfg1 sumoylation and, interestingly, also resulted in a significant decrease in the Tfg1-RNAPII interaction. Although this might indicate that Tfg1 sumoylation has a positive role in maintaining the interaction, which is opposite to our prediction, it is more likely due to the widespread effect of deleting *SIZ1*, which has numerous targets in addition to Tfg1 and RNAPII. As such, reduced sumoylation of some other factor(s) in the *siz1Δ* strain might indirectly reduce the Tfg1-RNAPII interaction. The *ulp1-1* mutant strain with an increased level of Tfg1 sumoylation was used to examine the effect of elevated Tfg1 sumoylation on Tfg1-RNAPII interaction. Unfortunately, it's hard to give a conclusion for this examination because of the lack of confident results, which is due to the faintness of immunoblot signals as well as the no-antibody control background.

As another strategy to examine the role of sumoylation of Tfg1, we attempted to generate a new yeast strain that expresses Tfg1 as an Smt3-Tfg1 fusion protein (**Fig 18**). This strategy has been used by others to study the effects of "permanent sumoylation" (49), and, in our case, because all Tfg1 molecules in the cell would be covalently attached to SUMO, this theoretically could simulate significantly elevated the level of SUMO modified Tfg1 in the cell. Furthermore, the strain would allow us to examine elevated Tfg1 sumoylation without affecting other SUMO target proteins. If we had succeeded in generating the strain, we would have used it to get an idea of how elevated Tfg1 sumoylation might affect the Tfg1-RNAPII interaction and gene transcription. Unfortunately, currently the fusion strain was not able to be generated due to failure of the cloning process.

The final part of our study on Tfg1 sumoylation focused on determining whether the modification affects

the association of Tfg1 with its chromatin binding sites or the expression levels of its target genes. Multiple ChIP experiments were performed, but because of a high level of inter-experiment variability, results were not conclusive in determining whether sumoylation affects binding of Tfg1 or RNAPII to active gene promoters. However, when the repeat experiments were taken together, the average occupancies of Tfg1 and RNAPII were not statistically significantly different in the *TFG1* wild-type and *tfg1-K60,61R* strains (**Fig 16**). We explored whether blocking sumoylation of Tfg1 through the K60,61R mutation might affect the expression levels of transcribed genes. Indeed, for a subset of the tested genes, inhibiting Tfg1 sumoylation resulted in a significant elevation of RNA abundance, although no change was detected at other genes (**Fig 17**). Overall, result of our RNA experiment indicated that, although Tfg1 sumoylation did not alter the TFIIF-RNAPII interaction, it may contribute to the repression of transcription through a RNAPII-independent mechanism.

To explore additional roles for sumoylation in the regulation of gene expression, we wished to examine the role of sumoylation in pre-mRNA splicing. According to proteomics studies which predicted that the human U2 snRNP splicing factor component SF3B1 is a SUMO target with high confidence, we were interested in confirming that it is sumoylated and characterize the role of sumoylation in regulating its function. However, we were unable to detect SF3B1 sumoylation in cultured cells with either endogenous or overexpressed levels of SUMO (**Fig 21 & 22**). One reason for this is that the level of endogenous sumoylated SF3B1 may be cell type-specific or condition-specific. To be clear, this means the endogenous sumoylated SF3B1 may be highly expressed in certain cell types or under particular conditions. No detection of SF3B1 sumoylation in our results may due to the low level of modified SF3B1 expressed in the cell lines tested, or the improper culturing and experimental conditions that do not favor SF3B1 sumoylation. The lack of anti-SF3B1 with high sensitivity can be another reason for the failed detection of SF3B1 sumoylation. Although we attempted to enhance this modification by using HeLa cells that overexpress SUMO, it seems to be insufficient to promote SF3B1 sumoylation to a detectable level.

To summarize our findings, amongst GTFs, TFIIF is relatively highly sumoylated in normally growing yeast cells, specifically its largest subunit, Tfg1. A fraction of Tfg1 molecules in the cell is constitutively

sumoylated, which is unaffected by exposure to stress. Changes in Tfg1 sumoylation, therefore, are not likely the key contributor for the altered transcription that occurs as part of the stress defense. Blocking Tfg1 sumoylation, through mutation of Lysine 60 and 61, does not alter the TFIIF-RNAPII interaction, but a general decrease in cellular sumoylation levels, through deletion of the SUMO ligase Siz1, does reduce the interaction. Although sumoylated Tfg1 is detected on chromatin, blocking the modification does not appear to affect chromatin occupancy of Tfg1 or RNAPII at promoters of active genes. However, the expression of some genes is elevated in strains expressing the Tfg1 K60,61R mutant, suggesting that the modification can have a gene-specific repressive role, through unknown mechanisms as well as the chromatin association of TFIIF and RNAPII. TFIIF is highly involved in RNAPII-mediated transcription by directly interacting with RNAPII and recruiting it to activated gene promoters. Our results now demonstrate that sumoylation can regulate chromatin-associated TFIIF, but since the modification does not affect the RNAPII-TFIIF interaction, sumoylation must control other properties of the TFIIF as a mechanism of regulating gene expression.

## Chapter 5: Future Directions

Although there are limitations to using a SUMO fusion for studying the role of sumoylation, we expect that generating an Smt3-Tfg1-expressing strain would be beneficial to determine what properties of TFIIF are enhanced or impaired by the modification. As part of the future directions, more efforts could be put in troubleshooting the cloning procedure in order to generate the Smt3-Tfg1 fusion yeast strain. Especially, the poor gel purification efficiency should be improved. By solving this, we will be able to get purified fusion PCR product at a high enough concentration that can be used in generating the desired reconstructed plasmids which then are used to transform yeast cells. If we successfully obtained this new strain, it can be examined in comparison to wild-type and *tfg1-K60,61R* strains in a number of experiments. This would allow us to test how dramatically decreased and increased Tfg1 sumoylation affects growth rates, the TFIIF-RNAPII interaction, recruitment of TFIIF and RNAPII to target gene promoters, and the expression level of a number of genes.

As a result of inconsistent ChIP results, we failed to show a clear change in chromatin association of Tfg1 and RNAPII at selected gene promoters and ORFs. Though technique troubleshooting may help to improve the results, an alternative method that also tests transcription factor/RNAPII-chromatin association may be used. Specifically, ChIP-Seq (chromatin immunoprecipitation followed by next generation shot-gun sequencing) can be performed. Ideally, this would be performed using *TFG1* wild-type, *tfg1-K60,61R*, and the fusion strain described above, to examine how Tfg1 sumoylation alters its global DNA-binding patterns. A number of previous studies have found that sumoylation affects the chromatin occupancy of transcription factors, and it has been suggested that this is a major role for the modification and a common mechanism of regulating gene expression (50). Additionally, RNA-Seq (whole-transcriptome sequencing) would also contribute to a wider understand of the regulatory role of Tfg1 sumoylation by examining how altered Tfg1 sumoylation levels affects the composition of the transcriptome. This will help determine whether effects of Tfg1 sumoylation are general or specific to certain groups of genes.

Much of our analysis focused on the interaction between Tfg1 and RNAPII. However, it will be interesting

to look into how sumoylation of Tfg1 might influence protein-protein interactions between TFIIF and other proteins, particularly those that compose the transcription pre-initiation complex, including other general transcription factors and components of the Mediator complex. To achieve this, we can take the advantage of an analytical technique, mass spectrometry (MS) to identify proteins that associate with Tfg1, Tfg1-K60/61R, or Smt3-Tfg1 after immunopurification from yeast cultures. The MS results will allow us to identify, upon decreased or increased Tfg1 sumoylation, whether binding of specific proteins is regulated by the modification.

Overall, the future directions mainly aim to study the effects of Tfg1 sumoylation in a global view, instead of focusing on individual genes or specific proteins. We hope that by looking broadly we can shed light on the function of Tfg1 sumoylation-directed transcription regulation, which may serve as a clue when more generally investigating the effect of sumoylation of other GTFs on gene expression.

## References


1. Sainsbury S, Bernecky C & Cramer P. (2015). Structural basis of transcription initiation by RNA polymerase II. *Nat Rev Mol Cell Biol*. 16(3):129-43.
2. Li Y, Flanagan PM, Tschochner H & Kornberg RD. (1994). RNA polymerase II initiation factor interactions and transcription start site selection. *Science*. 263(5148):805-7.
3. Hahn S & Young ET. (2011). Transcriptional regulation in Saccharomyces cerevisiae: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. *Genetics*. 189(3):705-36.
4. Rani PG, Ranish JA & Hahn S. (2004). RNA polymerase II (Pol II)–TFIIF and Pol II–Mediator complexes: the major stable Pol II complexes and their activity in transcription initiation and reinitiation. *Mol Cell Biol*. 24, 1709-20.
5. Chen ZA, Jawhari A, Fischer L, et al. (2010). Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J*. 29(4):717-26.
6. Henry NL, Campbell AM, Feaver WJ, Poon D, Weil PA & Kornberg RD. (1994). TFIIF-TAF-RNA polymerase II connection. *Genes Dev.* 8:2868–2878.
7. Jensen ON. (2004). Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr Opin Chem Biol*, 8(1):33-41.
8. Mann M & Jensen ON. (2003). Proteomic analysis of post-translational modifications. *Nat Biotechnol*. 21(3):255-61.
9. Miteva M, Keusekotten K, Hofmann K, Praefcke GJ & Dohmen RJ. (2010). Sumoylation as a signal for polyubiquitylation and proteasomal degradation. *Subcell Biochem*. 54:195-214.
10. Cubeñas-potts C & Matunis MJ. (2013). SUMO: a multifaceted modifier of chromatin structure and function. *Dev Cell*. 24(1):1-12.
11. Lamoliatte F, Mcmanus FP, Maarifi G, Chelbi-Alix MK & Thibault P. (2017). Uncovering the SUMOylation and ubiquitylation crosstalk in human cells using sequential peptide immunopurification.

*Nat Commun*. 8:14109.

12. Gill G. (2004). SUMO and ubiquitin in the nucleus: different functions, similar mechanisms?. *Genes Dev*. 18(17):2046-59.
13. Rodriguez MS, Dargemont C & Hay RT. (2001). SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. *J Biol Chem*. 276:12654–12659.
14. Hietakangas V, et al. (2006). PDSM, a motif for phosphorylation-dependent SUMO modification. *Proc Natl Acad Sci USA*. 103:45-50.
15. Hay RT. (2013). Decoding the SUMO signal. *Biochem Soc Trans*. 41:463-73.
16. Wilkinson KA & Henley JM. (2010). Mechanisms, regulation and consequences of protein SUMOylation. *Biochem J*. 428(2):133-45.
17. Liang YC, Lee CC, Yao YL, Lai CC, Schmitz ML & Yang WM. (2016). SUMO5, a Novel Poly-SUMO Isoform, Regulates PML Nuclear Bodies. *Sci Rep*. 6:26509.
18. Kamitani T, Kito K, Nguyen HP, Fukuda-kamitani T & Yeh ET. (1998). Characterization of a second member of the sentrin family of ubiquitin-like proteins. *J Biol Chem*. 273(18):11349-53.
19. Wang Y & Dasso M. (2009). SUMOylation and deSUMOylation at a glance. *J Cell Sci*. 122(23):4249-52.
20. Cremona CA, Sarangi P & Zhao X. (2012). Sumoylation and the DNA damage response. *Biomolecules*. 2(3):376-88.
21. Gong L, Li B, Millas S & Yeh ET. (1999). Molecular cloning and characterization of human AOS1 and UBA2, components of the sentrin-activating enzyme complex. *FEBS Lett*. 448:185-9
22. Johnson ES, Schwienhorst I, Dohmen RJ & Blobel G. (1997). The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer. *EMBO J*. 16(18):5509-19.
23. Johnson ES & Gupta AA. (2001). An E3-like factor that promotes SUMO conjugation to the yeast septins. *Cell*. 106(6):735-44.
24. Li SJ & Hochstrasser M. (1999). A new protease required for cell-cycle progression in yeast. *Nature*. 398:246-251.

25. Li SJ & Hochstrasser M. (2000). The yeast *ULP2 (SMT4)* gene encodes a novel protease specific for the ubiquitin-like Smt3 protein. *Mol Cell Biol*. 20:2367-77.
26. Makhnevych T, Sydorskyy Y, Xin X, et al. (2009). Global map of SUMO function revealed by protein-protein interaction and genetic networks. *Mol Cell*. 33(1):124-35.
27. Rosonina E, Duncan SM & Manley JL. (2010). SUMO functions in constitutive transcription and during activation of inducible genes in yeast. *Genes Dev*. 24(12):1242-52.
28. Hendriks IA & Vertegaal AC. (2016). A comprehensive compilation of SUMO proteomics. *Nat Rev Mol Cell Biol*. 17(9):581-95.
29. Esteras M, Liu IC, Snijders AP, Jarmuz A & Aragon L. (2017). Identification of SUMO conjugation sites in the budding yeast proteome. *Microb Cell*. 4(10):331-41.
30. Panse VG, Hardeland U, Werner T, Kuster B & Hurt E. (2004). A proteome-wide approach identifies sumoylated substrate proteins in yeast. *J Biol Chem*. 279(40):41346-51.
31. Chen X, Ding B, Lejeune D, Ruggiero C & Li S. (2009). Rpb1 sumoylation in response to UV radiation or transcriptional impairment in yeast. *PLoS ONE*. 4(4):e5267.
32. Yang SH & Sharrocks AD. (2004). SUMO promotes HDAC-mediated transcriptional repression. *Mol Cell*. 13(4):611-7.
33. Lyst MJ & Stancheva I. (2007). A role for SUMO modification in transcriptional repression and activation. *Biochem Soc Trans*. 35(Pt 6):1389-92.
34. Gómez-del arco P, Koipally J & Georgopoulos K. (2005). Ikaros SUMOylation: switching out of repression. *Mol Cell Biol*. 25(7):2688-97.
35. Pozzi B, Bragado L, Will CL, et al. (2017). SUMO conjugation to spliceosomal proteins is required for efficient pre-mRNA splicing. *Nucleic Acids Res*. 45(11):6729-45.
36. Will CL & Lührmann R. (2011). Spliceosome structure and function. *Cold Spring Harb Perspect Biol*. 3(7): a003707
37. Wan Y & Wu CJ. (2013). SF3B1 mutations in chronic lymphocytic leukemia. *Blood*. 121(23):4627-34.
38. Nuro-Gyina PK & Parvin JD. (2016). Roles for SUMO in pre-mRNA processing. *Wiley Interdiscip Rev*

*RNA*. 7(1):105-12.

39. Vassileva MT & Matunis MJ. (2004). SUMO modification of heterogeneous nuclear ribonucleoproteins. *Mol Cell Biol*. 24(9):3623-32.
40. Livak KJ & Schmittgen TD. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*.25:402-8.
41. Tomm JM, Rockstroh M, Müller S, Jende C, Kerzhner A & Bergen MV. (2011). Cell fractionation - an important tool for compartment proteomics. *JIOMICS.* 1(1):135-43.
42. Ho YH & Gasch AP. (2015). Exploiting the yeast stress-activated signaling network to inform on stress biology and disease signaling. *Curr Genet*. 61(4):503-11.
43. Bae NS, Seberg AP, Carroll LP & Swanson MJ. (2017). Identification of Genes in that Are Haploinsufficient for Overcoming Amino Acid Starvation. *G3 (Bethesda)*. 7(4):1061-84.
44. Downey M, Johnson JR, Davey NE, Newton BW, Johnson TL, Galaang S, Seller CA, Krogan N & Toczyski DP. (2015). Acetylome profiling reveals overlap in the regulation of diverse processes by sirtuins, gcn5, and esa1. *Mol Cell Proteomics*. 14(1):162-76.
45. Fang NN, Chan GT, Zhu M, Comyn SA, Persaud A, Deshaies RJ, Rotin D, Gsponer J & Mayor T. (2014). Rsp5/Nedd4 is the main ubiquitin ligase that targets cytosolic misfolded proteins following heat stress. *Nat Cell Biol*. 16(12):1227-37.
46. Schweikhard V, Meng C, Murakami K, Kaplan CD, Kornberg RD & Block SM. (2014). Transcription factors TFIIF and TFIIS promote transcript elongation by RNA polymerase II by synergistic and independent mechanisms. *Proc Natl Acad Sci USA*. 111(18):6642-7.
47. Hamard PJ, Boyer-guittaut M, Camuzeaux B, Dujardin D, Hauss C, Oelgeschläger T, Vigneron M, Kedinger C & Chatton B. (2007). Sumoylation delays the ATF7 transcription factor subcellular localization and inhibits its transcriptional activity. *Nucleic Acids Res*. 35(4):1134-44.
48. Morano KA, Grant CM & Moye-rowley WS. (2012). The response to heat shock and oxidative stress in Saccharomyces cerevisiae. *Genetics*. 190(4):1157-95.
49. Yurchenko V, Xue Z & Sadofsky MJ. (2006). SUMO modification of human XRCC4 regulates its

localization and function in DNA double-strand break repair. *Mol Cell Biol*. 26(5):1786-94.

50. Rosonina E, Akhter A, Dou Y, Babu J & Sri theivakadadcham VS. (2017). Regulation of transcription factors by sumoylation. *Transcription*. 8(4):220-31.

# Appendix

**Table 1: List of selected cross-linked residues within the TFIIF-RNAPII complex.** Examples of inter-protein links along with the interacting lysine residues are shown. Adapted from (5).

| | | | | | | |
|---|---|---|---|---|---|---|
| Inter-protein | RNAPII-TFIIF | Tfg1 | Lys 60 | N-terminal region | Rpb2 | Lys 606 |
| Inter-protein | RNAPII-TFIIF | Tfg1 | Lys 61 | N-terminal region | Rpb2 | Lys 606 |
| Inter-protein | RNAPII-TFIIF | Tfg1 | Lys 61 | N-terminal region | Rpb2 | Lys 652 |

**Table 2: *S. cerevisiae* strains used in this study.**

| Strain | Genotype |
|---|---|
| ERYM552B | *rpb1-6HA::RPB1 Kl TRP1* |
| YER023 (W303a) | *MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 [phi+]* |
| YFK001I | *TFA1-6HA::Kl TRP1* |
| YJB001A | *TOA1-6HA::Kl TRP1* |
| YJB002A | *TFG1-6HA::Kl TRP1* |
| YJP001A | *TAF7-6HA::Kl TRP1* |
| YRB013A | *tfg1-K60/61R-6HA::Kl TRP1* |
| YYD001A | *siz1Δ::KanMX TFG1-6HA::Kl TRP1* |
| YYD002D | *TFG1-6HA::Kl TRP1* |
| YYD003A | *TFG1-6HA::Kl TRP1 ulp1-1* |
| YYD004A,B,C | *TFG1 / tfg1Δ::URA3* |

**Table 3: Primers used for qPCR (ChIP and RT-qPCR).**

| Target Gene | Region | Primer Sequence ( from 5' to 3') | Direction |
|---|---|---|---|
| *ADH1* | Promoter | GTTTGCTGTCTTGCTATCAA | Sense |
| | | GAAAAAGAAACAAGGAAGAA | Anti-sense |
| Chromosome V (Chr V) | Untranscibed | CATTATCCGTAACGCCACTTT | Sense |
| | | CGATCTTAGTTCCAATGGTGAAA | Anti-sense |
| *HSP12* | ORF | CTCTGCCGAAAAAGGCAAGG | Sense |
| | | GACGGCATCGTTCAACTTGG | Anti-sense |
| *PHO5* | ORF | CACTTTTGCCAACTCGGACG | Sense |
| | | AGCATGTCTCTTGGCGTTCA | Anti-sense |
| *PMA1* | Promoter | CAATTATGACCGGTGACGAAAC | Sense |
| | | AATCGAAACTAATGGAGGGGAG | Anti-sense |
| *PMA1* | ORF | CTGGTCCATTCTGGTCTTCTATC | Sense |
| | | TCAGACCACCAACCGAATAAG | Anti-sense |
| *PYK1* | ORF | AACCCAAGACCAACCAGAGC | Sense |
| | | ACGGCGTTGATTGGGTAGTT | Anti-sense |
| *SAH1* | ORF | GGTAAGGTTGCCGTTGTTGC | Sense |
| | | TCGGTAACCAAGACACGAGC | Anti-sense |
| *SSA4* | ORF | GGCAGAAAAGTTCAAGGCCG | Sense |
| | | GCATCCTCTTCACCCACCTT | Anti-sense |
| *STL1* | Promoter | TGAAACCGCCGGAAGAAGTT | Sense |
| | | CCATGGTTGAGTGCCATCCT | Anti-sense |
| 25S | ORF | TCTAGCATTCAAGGTCCCATTC | Sense |
| | | CCCTTAGGACATCTGCGTTATC | Anti-sense |

**Table 4: Primers used for PCR during cloning procedure.**

| Target Gene | Primer Sequence ( from 5' to 3') | Direction |
|---|---|---|
| *SMT3* | AAGGTTGAAGAATAGCTAGGGAGGCCTATTCGAG TAAGCGCAAGGATGTCGGACTCAGAAGTCAATCA AGAAGC | Sense |
| | CCCGTTTCTACTGCCTGGTGGATTGCGTCTGGACA TAATCTGTTCTCTGTGAGCCTCAATAATATCG | Anti-sense |
| *TFG1-6HA* | ATGTCCAGACGCAATCCACCAGGCAGTAGAAACG GGGGAGGTCC | Sense |
| | TTTAAAGTTTCATTCCAGCATTAGC | Anti-sense |
| *URA3* | AAGGTTGAAGAATAGCTAGGGAGGCCTATTCGAG TAAGCGCAAGGATGTCGAAAGCTACATATAAG | Sense |
| | ACGAAAACTAAATAACCTATTAAGTACATAACATT ATAAACTAACTTAGTTTTGCTGGCCGCATC | Anti-sense |
| *SMT3-TFG1-6HA* | AAGGTTGAAGAATAGCTAGGGAGGCCTATTCGAG TAAGCGCAAGGATGTCGGACTCAGAAGTCAATCA AGAAGC | Sense |
| | ACGAAAACTAAATAACCTATTAAGTACATAACATT ATAAACTAACTTAGCTAGAAGCGTAATCTGGAACG TC | Anti-sense |

**Table 5: 3-STEP standard PCR profile.**

| | TEMPERATURE | DURATION |
|---|---|---|
| INITIAL DENATURATION | | |
| | 94˚C | 1:00 |
| REPEAT FOLLOWING THREE STEPS 30X | | |
| | 92˚C | 0:40 |
| | 55˚C | 0:40 |
| | 72˚C | 1:00 |
| FINAL EXTENSION | | |
| | 72˚C | 5:00 |
| COOL UNTIL SAMPLES REMOVED | | |
| | 4˚C | - |