The Sumoylation of General Transcription Factors Regulates Transcription in *Saccharomyces cerevisiae*

Mohammad Sohaib Baig

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 2021

Abstract

Sumoylation is a post-translational modification which primarily targets nuclear substrates and can result in altered target protein activity, stability, conformation, and/or localization. Out of the six general transcription factors (GTFs), we have determined that transcription factor IID (TFIID) and transcription factor IIF (TFIIF) are targets of sumoylation. The functional impacts of GTF sumoylation were assayed using strains with elevated or reduced TFIID or TFIIF sumoylation. Using chromatin immunoprecipitation (ChIP), we found that GTF sumoylation does not substantially impact GTF-chromatin interactions but can impact the promoter-wide sumoylation level. Notably, TFIID and TFIIF sumoylation enhances RNAPII recruitment at various promoters likely via altered protein-protein interactions but does so through seemingly distinctive mechanisms. A ChIP-seq and an RNA-seq revealed that reduced GTF sumoylation results in diminished RNAPII density at thousands of genes but this change does not substantially impact steady-state mRNA levels. This research has uncovered novel characteristics of transcriptional regulation in eukaryotes.

Dedication

This thesis is dedicated to Dr. Jyoti Garg. Due to complications with COVID-19, she has been bedridden for nearly half a year.

When I first entered Dr. Ron Pearlman's lab to simply fulfil the thesis course requirement for my undergraduate degree, I had no intention of spending the next couple years of my life as a complete lab rat. As the lab manager, it was your guidance, instruction, and encouragement which made me fall in love with the research process and made me want to continue down this path. Ultimately based on your (fantastic) advice, I chose to do my master's with Emanuel. If it was not for you, I would not be here today.

While the situation is still uncertain, I know that you will power through this and get back on your feet so that you can continue to radiate positivity wherever you go. I have said it before, but I am still eternally grateful for all your help.

Acknowledgements

Over the course of this project, I received a tremendous amount of help from my teachers, advisors, and peers. It was a privilege to work alongside the following people.

First of all, I would like to thank my supervisor, Dr. Emanuel Rosonina. Thank you so much for giving me the opportunity to be a part of your lab. It was truly an honour to learn from you and to work under your direction. You did everything you could to bolster my project and to foster my growth, both as a researcher and as a person. You not only allowed, but in fact encouraged me to tackle problems the way I saw fit, and you never once refused when I asked to take breaks from the lab to focus on other aspects of my life. You were truly the best supervisor I could have ever asked for and are one of the kindest people I have ever met.

I had the pleasure of being accompanied by some awesome lab mates on this journey. Su, you are one of my best friends. I know that our friendship will outlast the convenience of proximity, and I hope it lasts forever. Ben, I owe you so much. You always went out of your way to extend a helping hand and made my problems your own. I could not have done this without you. Mostafa, I still occasionally think about how our differences brought us closer together. We somehow went from not being on talking terms following a set of fierce arguments, to hanging out nearly every single weekend. I will not forget the countless hours we spent together (virtually) during the early days of the pandemic. Marjan, thank you for showing me the ropes and for enduring years of lab sessions with me. Anthea, while it feels like I barely even met you, I still miss those deep questions you would drop out of nowhere, and the ensuing late night lab sessions we would spend discussing them.

I would like to thank my committee: Dr. Peter Cheung, Dr. Terrance Kubiseski, and Dr. Derek Wilson for their support and cooperation. I would also like to extend my thanks to Dr. James Bryan McNeil, Kyra Kerhofs, Dr. Ronald Pearlman, Dr. Vivian Saridakis, Yimo Dou, Sajjad Ahmed, Verne Cox, York University, and again Dr. Jyoti Garg.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

°C	Degrees Celsius
3d	Three days
AA	Amino acid(s)
ATP	Adenosine triphosphate
cDNA	Complementary deoxyribonucleic acid
ChIP	Chromatin immunoprecipitation
Chr V	Chromosome V
Chrom.	Chromatin bound/associated fraction
CPM	Counts per million
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CTD	Carboxy-terminal domain
C-terminal	Carboxy-terminal
C-terminus	Carboxy-terminus
D	Aspartic acid
dH_2O	Distilled water
ddH_2O	Double distilled water
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
E	Glutamic acid
EDTA	Ethylenediaminetetraacetic acid
EtOH	Ethanol
Fig	Figure
FDR	False discovery rate
g	Gram
g (italic)	*g*-force
GO	Gene Ontology
H_2O_2	Hydrogen peroxide
HA	Human influenza hemagglutinin
HCl	Hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HRP	Horseradish peroxidase
IB	Immunoblot
IgG	Immunoglobulin G
IP	Immunoprecipitation
KCl	Potassium chloride
kDa	Kilodalton(s)
Kl::TRP1	*Kluyveromyces lactis* TRP1 marker gene
Lys/K	Lysine
M	Molar
Min	Minute
mL	Millilitre(s)
mM	Millimolar(s)
mRNA	Messenger ribonucleic acid
NaCl	Sodium chloride
NEM	N-ethylmaleimide
NGS	Next-generation sequencing
NP40	Nonylphenoxypolyethoxyethanol
n.s.	Not significant
N-terminal	Amino-terminal
N-terminus	Amino-terminus
OD	Optical density
ORF	Open reading frame
PBS	Phosphate buffer saline
PBST	Phosphate buffer saline with Tween
PCR	Polymerase chain reaction
PDC1	Pyruvate decarboxylase (gene)
PEG	Polyethylene glycol
PGK1	3-phosphoglycerate kinase (gene)
pI	Isoelectric point
PIC	Preinitiation complex
PIPES	Piperazine-N,N’-bis (2-ethanesulfonic acid)
PMSF	Phenylmethylsulfonyl fluoride
PTM	Post-translational modification
PYK1	Pyruvate kinase 1 (gene)

qPCR	Quantitative polymerase chain reaction
R	Arginine
RNA	Ribonucleic acid
RNAPII	RNA polymerase II
Rpb1	RNA polymerase II subunit 1
Rpb3	RNA polymerase II subunit 3
RPM	Revolutions per minute
RPS20	Ribosomal protein of the small subunit S20 (gene)
RPS27b	Ribosomal protein of the small subunit S27 (gene)
rRNA	Ribosomal ribonucleic acid
RT	Room temperature
RT-qPCR	Real-time quantitative polymerase chain reaction
S. cerevisiae	*Saccharomyces cerevisiae*
SAE	SUMO-activating enzyme
SC	Synthetic Complete
SDS-PAGE	Sodium dodecyl sulphate - polyacrylamide gel electrophoresis
Sec/s	Second(s)
SENP	Sentrin-specific proteases
Seq	Sequencing
Solub.	Soluble fraction
SS-carrier DNA	Salmon sperm-carrier deoxyribonucleic acid
SUMO	Small ubiquitin-like modifier
TAF	TATA-binding protein associated factors
TBP	TATA-binding protein
TDH1	Triose-phosphate dehydrogenase isozyme 1 (gene)
TDH3	Triose-phosphate dehydrogenase isozyme 3 (gene)
TF	Transcription factor
Tfg1	Transcription initiation factor IIF subunit alpha
TFIID	Transcription factor II D
TFIIF	Transcription factor II F
Tris	Trisaminomethane
Trp	Tryptophan
TSS	Transcription start site
UTR	Untranslated region

WB	Western blot
WCE	Whole cell extract
WH	Winged helix
YPD	Yeast extract peptone dextrose
yPIC	Yeast protease inhibitor cocktail
µg	Microgram
µL	Microlitre
µM	Micromolar

Section 1: Introduction

1.1 Sumoylation

Sumoylation is the reversible attachment of the conserved eukaryotic small ubiquitin-like modifier (SUMO) protein to various other cellular proteins. With assistance from the Ubc9 enzyme, the approximately 100 amino acid SUMO protein is attached to a certain consensus sequence within a target protein, and this post-translational modification (PTM) can result in altered target protein activity, stability, conformation, and localization, among other effects [1]. So far, SUMO has been found to conjugate with more than 6000 proteins in human cells, including more than 70% of all nuclear proteins [2]. SUMO is undoubtedly important for cellular functioning, as deletion mutants often have either a severe growth defect or are outright inviable, depending on the species. Despite this seemingly extensive presence within cellular systems and pathways, the ultimate role of SUMO is still undetermined [2].

While much is still unknown about the function or purpose of SUMO, the past twenty-five years have given insight into its rudimentary molecular effects and biochemical aspects. Even though the ~100AA (~12 kDa) SUMO homolog is conserved across various eukaryotes, billions of years of evolution have resulted in differences in isoform count and structure between species (**Figs 1A and B**) [2]. Some species have a single isoform, such as the single Smt3 protein (encoded by the *SMT3* gene) in *Saccharomyces cerevisiae* which was used in this study. Some plants, meanwhile, such as *Arabidopsis thaliana*, can have up to eight isoforms [3]. Within humans, five isoforms have been identified so far: SUMO-1, SUMO-2/3, SUMO-4, and SUMO-5. SUMO-1 and the two highly similar and often interchangeable SUMO-2 and SUMO-3 proteins seem to fulfil slightly different cellular niches, as paralogue specificity has been observed in certain cases [1]. Though it remains possible that paralogue preference is partly due to differences in the topological presence and overall abundance rather than protein identity [2,4]. SUMO-4 and SUMO-5 are less studied and tissue-specific [5,6].

The mechanism of SUMO conjugation resembles that of ubiquitination (**Fig 1C**) [7]. Briefly, an isopeptidase first cleaves the short C-terminal extension of the SUMO precursor, exposing a C-terminal Gly-Gly (diglycine) motif. Then, a SUMO E1 heterodimer enzyme composed of Aos1 and Uba2 in yeast (SAE1 and SAE2 in mammals) activates SUMO in an ATP-dependent manner, resulting in a high energy thioester bond. The SUMO protein is then passed on to, and forms another thioester bond with, the single SUMO E2 conjugating enzyme Ubc9. This conjugating enzyme selects and attaches SUMO to the target substrate, generally with the help of a SUMO E3 ligase. There are three SUMO E3 ligases in yeast and fifteen in mammals. The SUMO E3 factors typically guide the target substrate to the active site of Ubc9, where an isopeptide bond is formed between SUMO and a specific lysine residue on the target. This attachment typically occurs at a Ψ-K-X-D/E consensus motif on the substrate, where Ψ represents a large hydrophobic residue and X represents any residue. The human SUMO-2/3 isoforms and the yeast Smt3 themselves contain the SUMO attachment consensus motif and can be polysumoylated, creating a SUMO chain [1,2]. Desumoylation occurs via SUMO-specific proteases: Ulp1 and Ulp2 in yeast, and SENP1-7 in mammals. Due to the constitutive activity of desumoylating enzymes, sumoylation and desumoylation both occur rapidly, often leaving the substrate sumoylated for only a short period of time. Because of this, only a small fraction of a particular target protein is sumoylated at any given moment. It is possible that only the topologically relevant fraction is modified, or potentially the cell strategically designates sumoylation to be short lived as a mechanism of imparting a certain (as of yet unknown) transient effect. Notably, however, temporary sumoylation has been reported to have left permanent changes in target substrates in some cases [2].

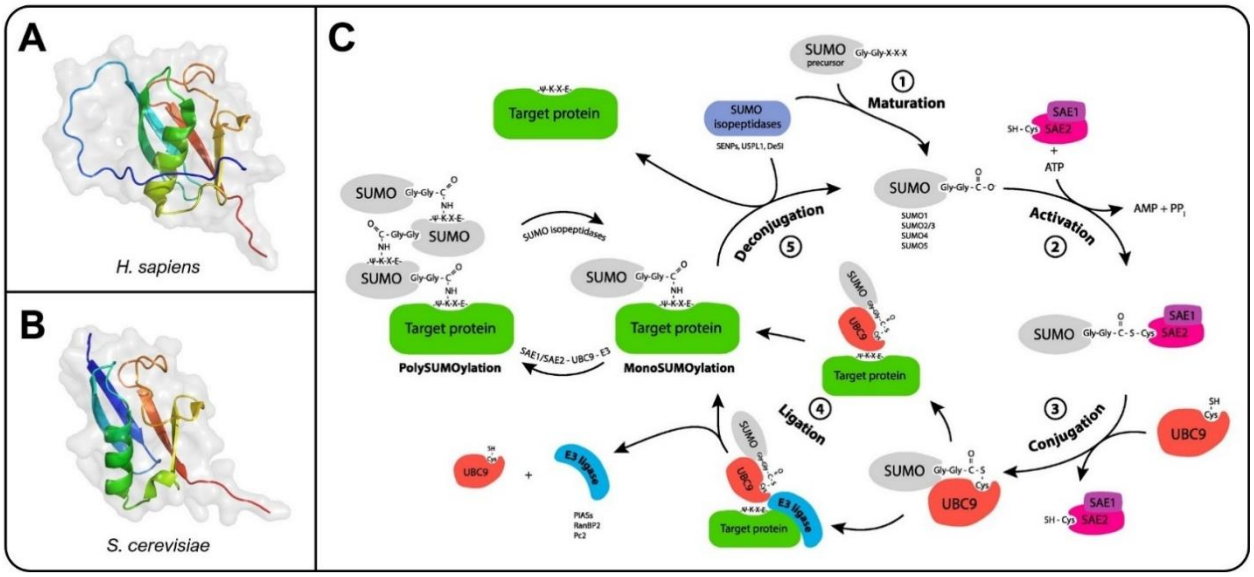

Figure 1. The SUMO protein and its reversible mechanism of attachment. (A) SUMO-1, the 101 amino acid isoform of the protein in humans. Unlike human SUMO-2/3 and the yeast Smt3, does not contain a SUMO attachment consensus motif to facilitate polysumoylation, and thereby acts as a chain terminator (PDB ID: 2N1V) [8]. **(B)** Smt3, the single 101 amino acid homolog of SUMO in *S. cerevisiae*. Contains the SUMO consensus motif surrounding lysines 11, 15, and 19, which allows for polysumoylation (PDB ID: 1EUV) [9]. **(C)** The SUMO conjugation and deconjugation cycle in humans. This mechanism of attachment is highly conserved between species. The SUMO precursor is matured through the action of an isopeptidase, revealing the C-terminal diglycine motif. ATP-dependent activation occurs via glycine attachment to an E1 enzyme, and SUMO is transferred to the sole E2 conjugating enzyme in humans and yeast, Ubc9. Either with or without assistance from an E3 ligating enzyme, SUMO is attached to a lysine residue on the target via an isopeptide bond. Polysumoylation can occur, extending the chain, or isopeptidases can reverse the attachment, releasing SUMO (image from [10]).

More than 500 proteins in budding yeast and more than 6000 proteins in human cells are known to be SUMO targets. Despite playing a part in numerous cellular pathways and processes, the ultimate role of sumoylation is still undetermined. Decades of research have, however, established the molecular impacts of sumoylation in particular cases and situations. Sumoylation seems to impart its effect on target substrate activity by changing the substrate's protein interactions, conformation (and thereby any interlinked protein interactions), or by competing for lysine residue occupancy with other PTMs [2,7]. For example, the MEK1 protein which is necessary for chemotaxis in *Dictyostelium* was found to be rapidly (but transiently) sumoylated in response to stimulation with a chemoattractant. Researchers found that when MEK1 was mutated to be nonactivatable, it no longer became sumoylated either [11]. For conformational changes, researchers found that covalent attachment of SUMO-1 to the human thymine-DNA glycosylase

(TDG) did not immediately affect its conformation, but the ensuing intramolecular interactions between the SUMO-1 and a SUMO-interacting motif (SIM) on TDG did affect TDG conformation [12]. In terms of PTM competition, the literature suggests that sumoylation of a transcription factor such as STAT5 can prevent subsequent acetylation and phosphorylation [13]. The SUMO PTM is relatively large at ~12 kDa, while other common PTMs tend to be smaller: the acetyl group is ~42 Da, the phosphate group is ~94 Da, and ubiquitin is ~8.5 kDa. It is postulated that the bulky size of SUMO might introduce steric hindrance to a localized area, preventing additional modifications at nearby residues or otherwise limiting protein-protein or protein-chromatin interactions [14]. Other molecular impacts of sumoylation include modified target stability, solubility, and localization [7,13].

It should be noted, however, that studying the effects of substrate sumoylation has proven to be a challenge. The current go-to approach employed by most labs is to mutate the lysine residue in the Ψ-K-X-D/E consensus motif to an arginine, thereby preventing sumoylation. But this mutation would also affect any other post-translational modifications at or near that site, such as phosphorylation, acetylation, or ubiquitination. Because of this, the caveat of potentially observing indirect effects must be continually considered [2]. Even when attempting to artificially sumoylate or increase the sumoylation level of a specific substrate (as discussed below in section 3.1), researchers cannot typically choose specific target residues. Instead, they are limited to either N-terminal or C-terminal attachment due to limits in contemporary cloning technology, which may not be representative of *in vivo* sumoylation. Furthermore, this results in an unnatural constitutive level of substrate sumoylation, rather than emulating the biologically accurate transient surge of activity.

1.2 Sumoylation and Transcription

Sumoylation primarily targets nuclear substrates, and transcription-related proteins form one of the largest classes of SUMO targets across species [15]. Initially, sumoylation was understood

as having a generally repressive role in transcription. This would align neatly with the fact that both SUMO (pI = 5.2) and DNA are negatively charged, which would support TF dissociation upon modification. Initial studies in the literature supported this notion for the most part. For instance, one group found that targeting the SUMO E2 Ubc9 enzyme to a promoter by fusing it with a Gal4 DNA-binding domain resulted in transcriptional repression at a Gal4 reporter gene. In another case, it was found that a reduction in histone H2B sumoylation caused an increase in basal gene expression [2]. As more studies have emerged, however, it has become increasingly clear that sumoylation has a consistent (i.e. reproducible) but context dependent role. Depending on the situation, the modification can either repress or promote transcription (**Figs 2A-D**) [13]. One example of this was the discovery that p53 sumoylation leads to transcriptional activation at reporter genes, likely by blocking ubiquitination at the same lysine residues [16]. In another example, it was found that sumoylation of the Bcl11b transcription factor in T cells enables its interaction with p300, thereby resulting in the derepression of genes which p300 normally represses [13]. SUMO attachment can also control the transcription rate by affecting TF-chromatin interactions, either positively or negatively (**Figs 2E and F**) [13]. Overall, while repression remains the more dominant consequence, sumoylation can promote transcription as well.

In addition to merely repressing or promoting transcription, an exciting third possibility has recently emerged: SUMO reinforces ‘correct’ TF-chromatin binding while simultaneously antagonizing ‘incorrect’ binding (**Fig 2G**) [17]. This binding-site selection model was proposed in a 2019 review paper based on the realization that various SUMO-deficient mutant TFs bind to several additional non-specific sites when compared to their wild-type counterparts. In this model, non-sumoylated TFs initially bind to chromatin with reduced specificity to allow for the sufficient binding of all functional sites and a subsequent sumoylation mediated conformational change either reinforces their binding if bound at appropriate locations or otherwise instigates their

dissociation. Thus, the fraction of TFs found to be sumoylated while chromatin-bound are likely correct binders while those TFs which are sumoylated while soluble would be expected to spurious binders which were ejected from chromatin [17]. Circling back, later studies on p53 sumoylation found that sumoylation at a different lysine contrarily blocks p53 acetylation, which is required for its activation, but importantly only when p53 was in solution. Once already DNA-bound, sumoylation at the same lysine was unable to dislodge p53 from DNA, which can be interpreted as being in agreement with the binding-site selection model [18,19]. It should be noted however, as the review paper highlighted, that the mechanism and cellular prevalence of this model of TF sumoylation are still undetermined and thus remains an active area of research [17]. While still incompletely understood, the effect of transcription factor sumoylation is evidently context dependent.

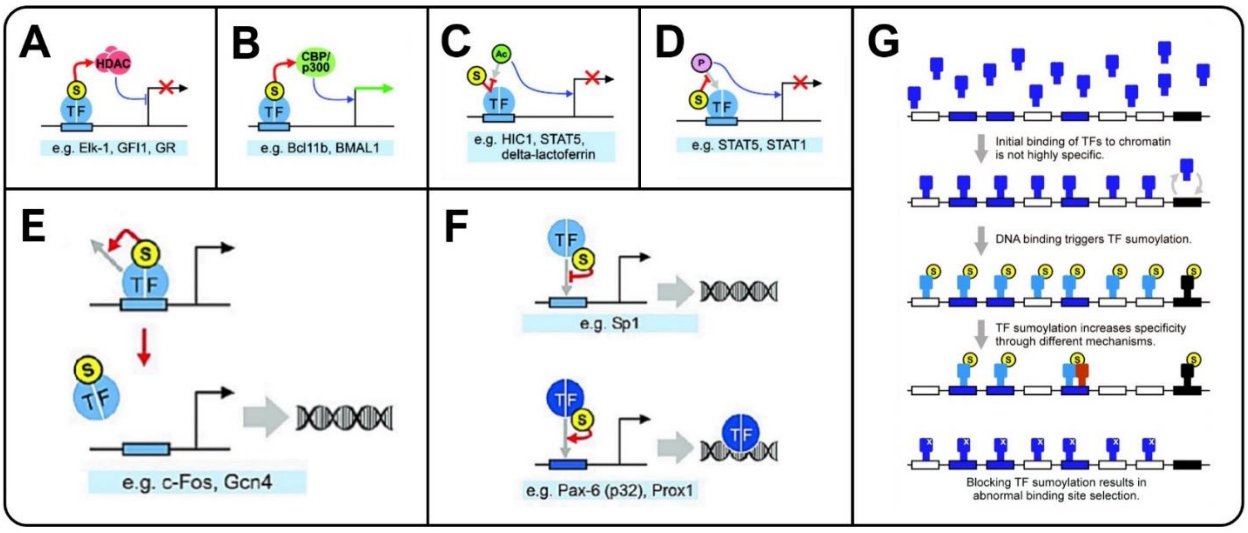

Figure 2. The sumoylation of transcription factors can result in an array of context-dependent outcomes. Depending on the gene or transcription factor in question, the effect of TF sumoylation can be distinct. **(A)** Sumoylation of TFs can repress transcription by recruiting histone deacetylases, or **(B)** can promote transcription by recruiting coactivators like CBP/p300. **(C)** SUMO can compete with acetylation for target lysine residue or **(D)** interfere with TF phosphorylation at a nearby residue, inhibiting transcription in both cases. **(E)** SUMO can impact TF-chromatin interactions by instigating TF clearance from promoters or **(F)** by directly repressing or enhancing their chromatin occupancy. **(G)** In certain cases, SUMO can play a role in enhancing the binding site selection of TFs by reinforcing specific binding and prompting TF dissociation when non-specifically bound. In each panel, SUMO is represented by an encircled *S*, and exemplary transcription factors are listed where appropriate (Adapted from [13] and [17]).

Numerous components of the transcriptional machinery are known or predicted to be SUMO targets. These include hundreds of sequence-specific transcription factors (SSTFs), the general transcription factors (GTFs), as well as multiple subunits of RNA Polymerase II (RNAPII). This is in line with the “group sumoylation” notion [2,20]. This emerging principle asserts that multiple components of a certain pathway, complex, or cluster are sumoylated (simultaneously and transiently) to impart a synergistic effect. This widespread sumoylation can either serve to stabilize complexes or act as a sort of ‘glue’ tethering pertinent components together. As proposed, the SUMO PTM on one protein in the group would attract other proteins with SIMs (SUMO-interacting motifs) to the group, via intermolecular interactions. These newly arriving proteins would then be sumoylated as well, resulting in a snowball effect whereby the group/complex would continue to grow in size and firmness as an increasing number of entanglements are established. A secondary advantage from a cellular perspective is that not every single potentially sumoylatable site in the group will need to be occupied to meet the threshold required to impart an effective result, as the presence of multiple SUMO proteins in the local vicinity provides a degree of redundancy. From a research point of view, however, this is a disadvantage. Mutating individual lysines to observe particular effects might be unfruitful as other SUMO interactions in the local region might persist or additional sites which are usually vacant might become sumoylated to compensate for the reduction. To counteract this, all potential SUMO sites within a topological area can be mutated, but this increases the chances of observing indirect effects as we stray further away from the natural state with each additional mutation [20]. Regardless, group sumoylation remains hypothetical. While examples of this phenomenon have been documented, its prevalence within the cell is unknown. And while group sumoylation is frequently referenced within the literature, it has rarely been explored experimentally [2].

1.3 Overview of Transcription Initiation

For protein synthesis to occur in all living systems, information encoded within DNA must first be transcribed into its corresponding messenger RNA (mRNA) form. Within eukaryotes, this is achieved through the activity of RNA Polymerase II. Before RNA Polymerase II mediated transcription can begin, the preinitiation complex (PIC) is formed at a gene promoter [21]. The PIC is composed of many general transcription factors and ultimately serves to recruit RNAPII and subsequently initiate transcription. Thereby, it is ultimately the formation of PIC components that control which gene will be expressed, rather than the polymerase itself. The six multi-subunit GTFs each play a distinct role in the process and have a particular order of assembly: TFIID, TFIIA, TFIIB, TFIIF, TFIIE, and TFIIH (**Fig 3**). To begin, the essential TATA-binding protein (TBP) subunit of TFIID arrives at the gene promoter and creates a sharp bend in the DNA, around 30 bp upstream of the transcription start site (TSS). Conventionally, TBP binds the TATA box element but if a TATA box is not present then TBP will still be required but may be coerced into binding by the sequence-specific TBP-associated factors (TAFs), which comprise the remainder of TFIID. In these cases, TBP will instead bind somewhere within the 200 bp region upstream of the TSS. TFIIA reinforces TBP binding and helps clear away repressive factors, which the cell employs to prevent inadvertent initiation. Next, TFIIB interacts with TBP and facilitates the recruitment of the polymerase. Following that, TFIIF escorts RNAPII to the growing complex at the promoter. TFIIF remains tethered to the polymerase when it is not DNA-bound to prevent inadvertent activity. TFIIE arrives at this stage. It is known to stimulate the reconfiguration of RNAPII to an activated state and might even possess DNA melting capabilities. TFIIH is recruited next; it has helicase activity to unravel the DNA strands ahead of the PIC and possesses kinase activity to phosphorylate the CTD of RNAPII, stimulating its release from the PIC. Another important non-GTF component of the PIC is the mediator complex. This large multi-subunit transcriptional coactivator relays information from various transcription factors to the polymerase at protein-coding genes [21].

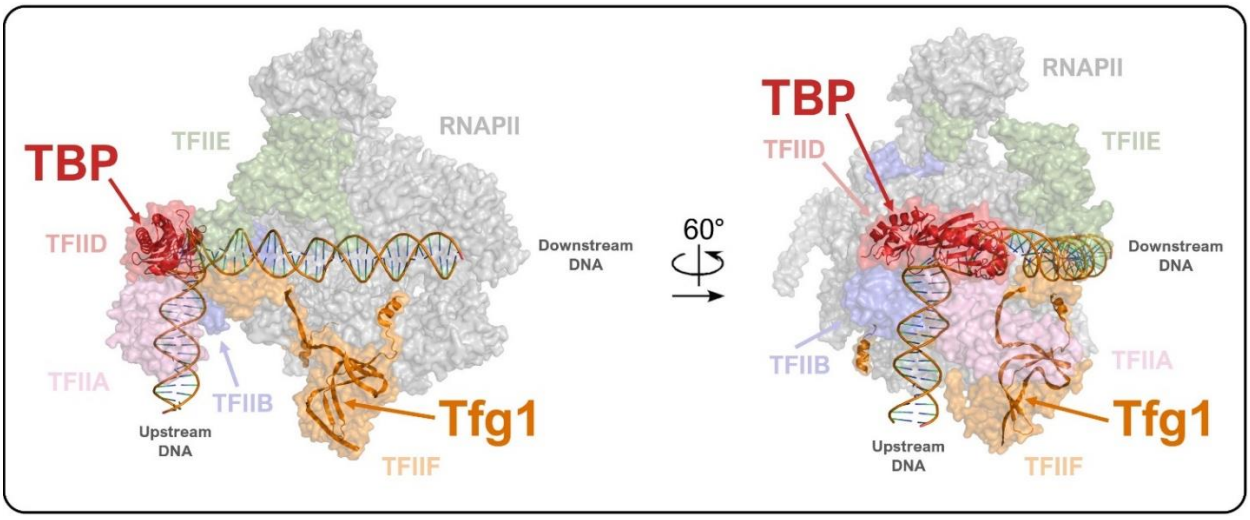

Figure 3. The transcription preinitiation complex is composed of six general transcription factors. Partial representation of *S. cerevisiae* preinitiation complex with relevant sections highlighted. Model constructed via cryo-electron microscopy (cryo-EM) [22]. The GTFs assemble in a particular order, and each plays a distinctive role to ultimately recruit RNAPII and begin transcription: TFIID, TFIIA, TFIIB, TFIIF, TFIIE, and TFIIH. TFIID is composed of TBP and many TAFs and serves as scaffolding for much of the remainder of the complex. TFIIF escorts RNAPII to the growing PIC and contains the Tfg1 subunit. TFIIH, which is located further downstream, and certain subunits of the other GTFs are not present in this diagram (PDB ID: 5FZ5) [21,22].

As will be explained in detail below (sections 1.6 and 1.7), the transcription initiation factor IIF subunit alpha (Tfg1) subunit of TFIIF is a known SUMO target in *S. cerevisiae* [15]. Meanwhile, large scale proteomics studies and other lines of evidence suggest that the TBP subunit of TFIID is also a SUMO target in *S. cerevisiae* [23]. Both of these GTF subunits have well understood functions and extensive regulatory control mechanisms.

1.4 Transcription Initiation Factor IIF Subunit Alpha (Tfg1)

TFIIF is the GTF responsible for shuttling RNAPII to the PIC. It has other functions such as preventing the non-specific binding of the polymerase with DNA, stabilizing TFIIB, enhancing the TSS selection capabilities of the PIC, stimulating the initial phosphodiester bond formation, and reducing the time RNAPII spends paused [24]. TFIIF has three subunits: Tfg1, Tfg2, and Tfg3 in *S. cerevisiae*. Tfg1 is the largest subunit of the three (**Figs 4A and B**). Within *S. cerevisiae*, Tfg1 is encoded by the *TFG1* gene, contains 735 residues, and has a mass of ~82 kDa. Previous

research in our lab has shown that Tfg1 is the primarily sumoylated subunit of the three, as will be discussed below.

Tfg1 either forms direct interactions or coordinates interactions via the other two TFIIF subunits with multiple RNAPII subunits: Rpb1, Rpb2, Rpb3, Rpb9, and Rpb10 (**Fig 4C**) [25]. The N-terminal regions of both Tfg1 and Tfg2 form a dimerization domain which binds the active centre cleft of RNAPII, near the downstream DNA portion. The C-terminal regions of Tfg1 and Tfg2 contain mobile winged helix (WH) domains close to the upstream DNA portion. Additionally, Tfg1 can become phosphorylated, and its phosphorylation correlates with a high level of transcriptional activity [26].

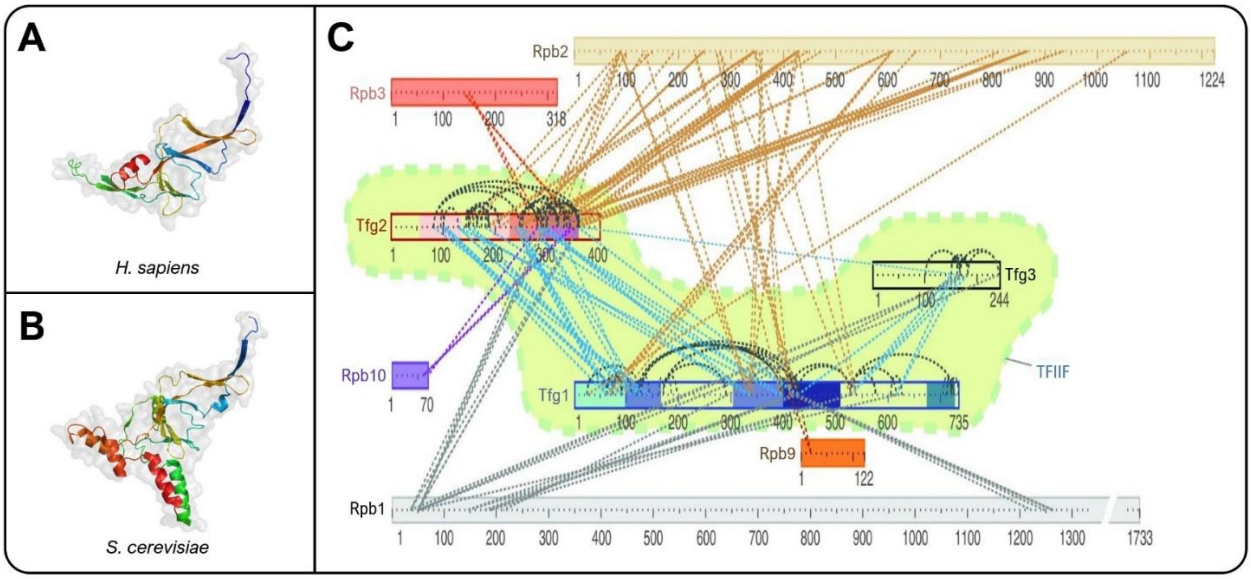

Figure 4. The Transcription Initiation Factor IIF Subunit Alpha (Tfg1) and its known interactions with RNAPII. **(A)** The RAP74 homolog of Tfg1 is a ~58 kDa protein in humans. Within the human TFIIF tetramer ($\alpha_2\beta_2$; RAP74 and RAP30), it is comprised of two distinctive N- and C-terminal domains which are connected by a central flexible loop (PDB ID: 1F3U) [25,27]. **(B)** Tfg1 in *S. cerevisiae* is a ~82 kDa protein and is a part of the yeast TFIIF trimer composed of Tfg1, Tfg2, and Tfg3 (PDB ID: 7O73) [25,28]. **(C)** A TFIIF cross-linking map shows extensive interactions between its three subunits and with RNAPII in *S. cerevisiae*. Of the three TFIIF subunits, Tfg1 is the one which primarily interacts with the largest subunit of the polymerase, Rpb1 (image from [26]).

1.5 TATA-Binding Protein (TBP)

TFIID is the first GTF to arrive at the core promoter and serves as scaffolding for much of the remainder of the PIC components. It orchestrates the interactions of more than 70 polypeptides and thereby establishes the general layout of the complex. In addition, it acts as a hub for regulatory signals. TFIID is composed of the aforementioned TBP and several TAF proteins which facilitate and enhance TBP function, although it has been demonstrated that the TAFs are not required to initiate transcription (**Fig 5**) [29]. In *S. cerevisiae*, TBP is encoded by the *SPT15* gene, has a length of 240 amino acids, and has a mass of ~27 kDa [30]. TBP is required for transcriptional initiation for all three major RNA polymerases: RNAPI, RNAPII, and RNAPIII. Furthermore, it is required for transcription regardless of whether the promoter contains a TATA-box element or not. In these cases, other GTF or coactivator subunits assist with promoter recognition [21]. Within certain higher eukaryotes, at certain promoters, and under specific conditions, TBP can be replaced by one of two similar variants: TBPL1/TRF2 and TBPL2/TRF3. Within *S. cerevisiae*, however, there is only a singular TBP and no analogues [31].

In addition to starting the process of transcriptional initiation, TBP has other functions. To bind DNA, TBP contains positively charged arginine and lysine residues, which interact with the negatively charged phosphates present on the DNA backbone. Upon binding, TBP creates an almost squared-off, 80° bend in an AT-rich stretch of DNA. It does this by kinking DNA ~40° at two close but distinct locations, creating a saddle shape. In each case, pairs of bulky phenylalanine residues intercalate within a minor groove of DNA. Then, additional amino acid side chains of TBP form hydrogen bonds with DNA bases, stabilizing the bent complex [32].

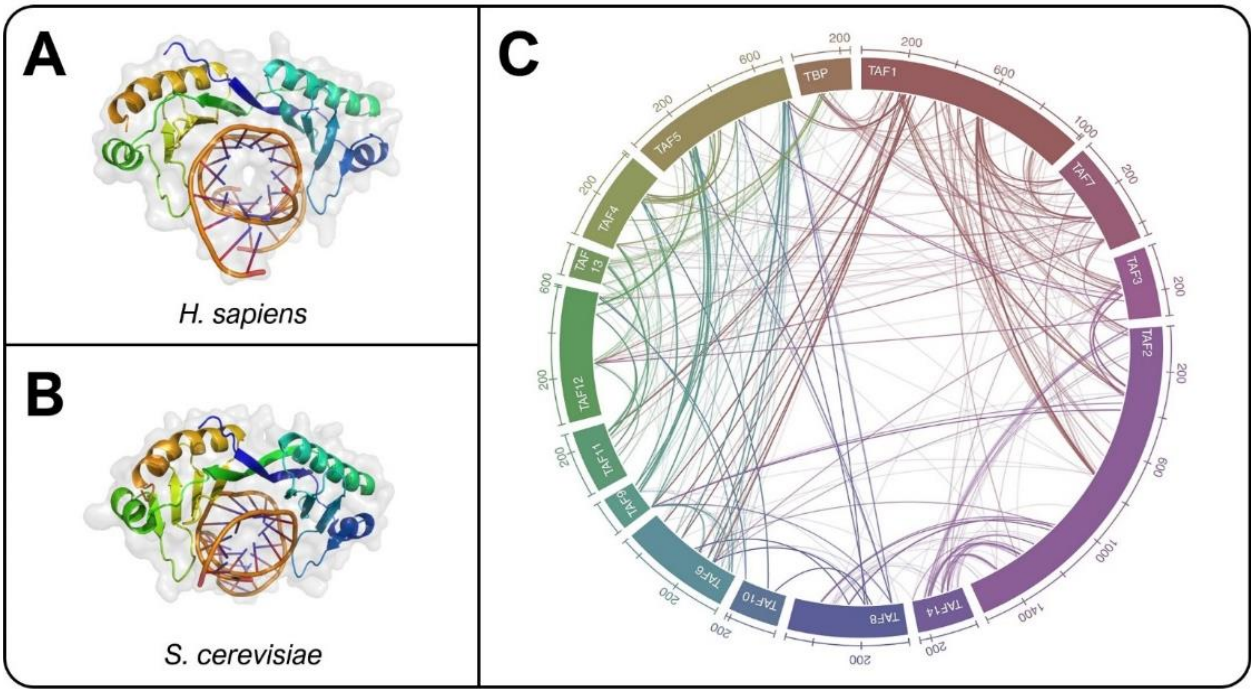

Figure 5. The TATA-Binding Protein (TBP) and its known interactions with TBP-associated factors (TAFs). (A) TBP is a ~38 kDa protein in humans (PDB ID: 1CDW) [33]. It is the first PIC component to arrive at a target promoter and it intercalates within DNA, creating a sharp bend for polymerase access to the genetic code [31,32]. **(B)** In *S. cerevisiae*, TBP is a ~27 kDa protein with a high level of functional similarity to the human counterpart (PDB ID: 1YTB) [34]. **(C)** Cross-linking map of TFIID subunits, including TBP and TAFs. TBP and the TAFs demonstrate extensive interactions to facilitate TBP-chromatin binding (image from [35]).

1.6 Insights From a Landmark SUMO ChIP-seq

Over the years, various studies have investigated the role of SUMO on chromatin. These include multiple genome-wide analyses examining the precise location and binding sites/targets of SUMO on chromatin. So far, it has been discovered that in log phase cells the majority of the SUMO signal on promoters originates from genes involved in translation, such as ribosomal protein genes (RPGs), histone-encoding genes, rRNA genes, and tRNA genes [15]. Further, in 2015, it was found that the SUMO signal at these translation-related genes specifically originated from sumoylation of the Rap1 transcription factor. Rap1 is multifunctional SSTF which generally binds the promoters of various classes of genes at a region which is modestly (~400 bp on average) upstream of the TSS. Rap1 and SUMO peaks align almost perfectly at virtually all RPGs. The investigators demonstrated that Rap1 sumoylation mediated its interactions with the basal

transcription machinery, and thereby enhanced GTF recruitment near the TSS of target gene promoters. Without Rap1 sumoylation, RPG expression was significantly diminished, and cells were found to be inviable when grown at slightly elevated (37°C) temperatures, signifying an inability to survive under stressful heat shock conditions [36].

Despite being primarily associated with translation-related genes, multiple chromatin immunoprecipitation sequencing (ChIP-seq) analyses have simultaneously demonstrated that there are robust SUMO signals originating from the promoter regions of hundreds of protein-coding genes which are not RPGs (hereafter referred to as non-RPGs) [15,36]. Our previous SUMO ChIP-qPCRs (quantitative polymerase chain reactions) reinforced this notion when we discovered that sumoylated proteins are found at many transcriptionally active genes (but not repressed or silenced genes), including protein-coding genes [37]. Based on these observations, our lab did a SUMO ChIP-seq and discovered that while SUMO does primarily associate with RPGs and tRNA genes, there are 147 distinct SUMO peaks at non-RPG protein-coding genes within *S. cerevisiae*. Intriguingly, the SUMO peaks at the majority of these non-RPGs were found to align with the core promoter rather than the binding sites of any known SSTFs and other promoter-associated factors (**Fig 6**). Precisely, this subset of SUMO peaks aligned almost perfectly with the TATA-box element just upstream of the TSS, where TBP and roughly where other GTFs bind. This indicated that GTFs are likely SUMO targets, at least when bound to these specific promoters. Of the 147 SUMO peaks at non-RPGs, 14 aligned with Rap1 binding sites, and 22 were located at overlapping promoters of two adjacent but divergent genes. A small number of the remaining 111 unique non-RPGs with SUMO peaks aligning with their core promoters did contain Rap1 binding sites, but the SUMO peaks found here did not align with the Rap1 binding sites, and instead aligned with the core promoter. The only exception was *GCN4*, which was found to have two SUMO peaks, one at its Rap1 binding site and one at its core promoter. The vast majority of the 111 non-RPGs did not contain a Rap1 binding site, and

therefore Rap1 does not regulate their activity. A full list of the non-RPGs containing SUMO peaks can be found within the supplementary tables of our published data (Table S9 at https://doi.org/10.1371/journal.pgen.1009828) [15].

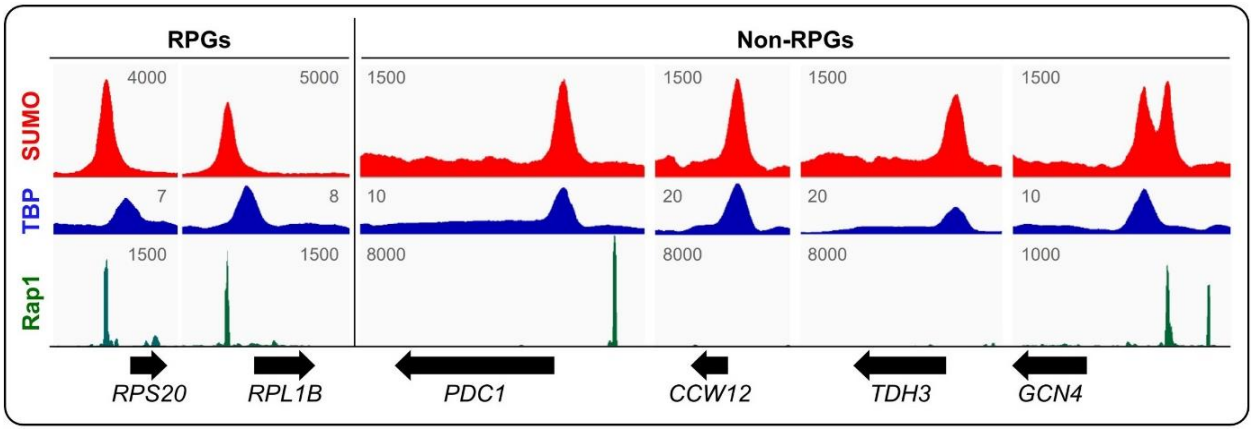

Figure 6. A previous SUMO ChIP-seq revealed that there are 147 SUMO peaks at non-RPG protein-coding genes. Of these, only 14 were found to align with the binding site of the multifunctional Rap1 transcription factor, which is responsible for facilitating RPG transcription once sumoylated. Another 22 of these genes were located at overlapping promoters of two nearby genes. The remaining 111 unique non-RPGs each contained a GTF-derived SUMO peak. The SUMO ChIP-seq signal (red) was aligned and compared with published TBP ChIP-seq (blue) and Rap1 ChIP-exo (green) datasets (NCBI GEO database accession numbers GSM2870615 and GSE93662, respectively) [15]. Due to their strong interactions, the TBP signal was indicative of GTF binding sites, not just TBP's binding site. At the *RPS20* and *RPL1B* ribosomal protein coding genes, the SUMO signal was derived from Rap1. At the *PDC1*, *CCW12*, and *TDH3* non-RPGs, the SUMO signal aligned with the TATA-box, where GTFs bind. At the exceptional *GCN4* non-RPG, two SUMO peaks were observed: one aligning with Rap1 and one with TBP (image from [15]).

Out of the 111 non-RPGs with SUMO peaks, it was found that over half of them fell into the top 10% of most actively transcribed genes as measured by RNAPII occupancy. While the intensity (i.e. read count) of the SUMO peaks themselves did not necessarily correlate with RNAPII density, these findings revealed that a common characteristic of many highly transcribed genes was promoter-associated sumoylation. A gene ontology (GO) analysis unveiled that many of these genes (around ~25%) were involved in the production of small molecules such as certain amino acids, implicating a potential pathway-specific regulatory system [15]. As will be discussed below, a handful of metabolic non-RPGs with prominent SUMO read count signals and high transcriptional activity were picked for individual ChIP analysis. These included *TDH3*, *PGK1*, *PYK1* (alias: *CDC19*), and *PDC1*. The expression products of these genes are enzymes involved

in glycolysis, gluconeogenesis, and fermentation. For comparison, two RPGs with Rap1-derived SUMO signals were selected as well: *RPS20* and *RPS27b*, both of which code for protein components of the small 40S ribosomal subunit. As a background level control, the *TDH1* gene coding for a stationary-phase specific isoform of the GAPDH protein was selected [38]. Unlike the other four chosen non-RPGs, the *TDH1* gene is reportedly not expressed in log-phase cells, which were used for the experimentations in this study, and does not have a SUMO peak according to our ChIP-seq data. Taken together, the past discoveries seem to imply a role for GTF sumoylation in regulating transcription at hundreds of genes. What remains unclear is why only 111 distinct non-RPGs out of the ~5800 non-RPGs in yeast have SUMO peaks, despite the fact that RNAPII occupancy is detected at most of them [15].

1.7 Overview of Previous Findings

Previous researchers in our lab have made good progress in investigating the link between GTFs and sumoylation. Large scale proteomics studies have suggested that TFIIA, TFIID, TFIIE, and TFIIF are putative SUMO targets [39]. By attaching HA epitope tags to various GTF subunits via homologous recombination, immunoprecipitating (IPing) them, and then immunoblotting them on HA and SUMO blots, the GTFs were individually assayed (**Fig 7A**). Specifically, Toa1, TAF7, TAF8, Tfg1, and Tfa1 were tagged. It was determined that while multiple GTFs display low levels of SUMO modifications, the Tfg1 subunit of TFIIF is a prominent target of SUMO modifications. Further investigation by means of IPing samples under denaturing conditions to dislodge any non-covalent interactions revealed that the largest Tfg1 subunit was the primarily sumoylated subunit, significantly more so than Tfg2 or Tfg3. A ladder pattern in SUMO immunoblots indicated that Tfg1 has multiple SUMO proteins attached to it, either at multiple locations and/or in a chain (multi- and/or polysumoylated). A densitometry analysis of HA immunoblots revealed that approximately 10% to 15% of Tfg1 is sumoylated at a given time in log phase cells, and a chromatin fractionation experiment revealed that Tfg1 can be sumoylated while soluble or chromatin bound [15].

To assess the functional impact of Tfg1 sumoylation, a mutant strain was constructed harboring two point mutations at the suspected lysine residues for SUMO attachment: Lys60 and Lys61. A computational prediction algorithm suggested that Lys60 in the N-terminal region of Tfg1 was a good candidate for SUMO attachment based on the identity of surrounding residues, which resembled the SUMO attachment consensus motif (Ψ-K-X-D/E). To proactively prevent SUMO from attaching to the nearby Lys61, both lysine residues were mutated to arginine residues, disrupting SUMO attachment. The resulting *Tfg1-K60,61R-HA* strain also contained a 6xHA epitope tag for immunodetection. When IPing the HA tag and immunoblotting for the presence of SUMO within this new mutant strain, it was discovered that the K60R and K61R mutations significantly reduced but did not abolish Tfg1 sumoylation (**Fig 7B**). This reduction in Tfg1 sumoylation did not impact cell survival or stress response capability. Somewhat unexpectedly, this reduction in Tfg1 sumoylation did not noticeably impact RNAPII-TFIIF interactions, either. This was evident when no appreciable difference was seen when co-IPing the Rpb1 subunit of RNAPII while using either a wildtype or the *Tfg1-K60,61R-HA* strain (**Fig 7C**, left side).

Next, the potential impact of elevated Tfg1 sumoylation levels on RNAPII-TFIIF interactions was explored. A *ulp1-1* strain was utilized, in which the desumoylating protease Ulp1 was partially defective, resulting in increased global sumoylation. SUMO immunoblots of epitope tagged Tfg1 within the *ulp1-1* strain verified the increase in Tfg1 sumoylation (data not shown). Elevated levels of Tfg1 sumoylation in the *ulp1-1* strain disrupted RNAPII-TFIIF interactions. This was confirmed by IPing the Rpb1 subunit of RNAPII and observing a reduction in co-IPed Tfg1 when using the mutant strain (**Fig 7C**, right side). A reciprocal co-IP experiment confirmed these results. For these experiments, an isogenic *ULP1-wt* strain was used as the wild-type parent, and endogenous Tfg1 contained the 6xHA epitope tag in both strains. In summary, only an increase and not a decrease in Tfg1 sumoylation disrupted RNAPII-TFIIF interactions [15]. However, due to the large number of cellular consequences that can stem from increased global sumoylation,

the results obtained using the *ulp1-1* mutant strain could not be attributed to increased Tfg1 sumoylation alone. To address this, and to confirm the previous results, a new strain would be needed to exclusively increase the sumoylation level of Tfg1.

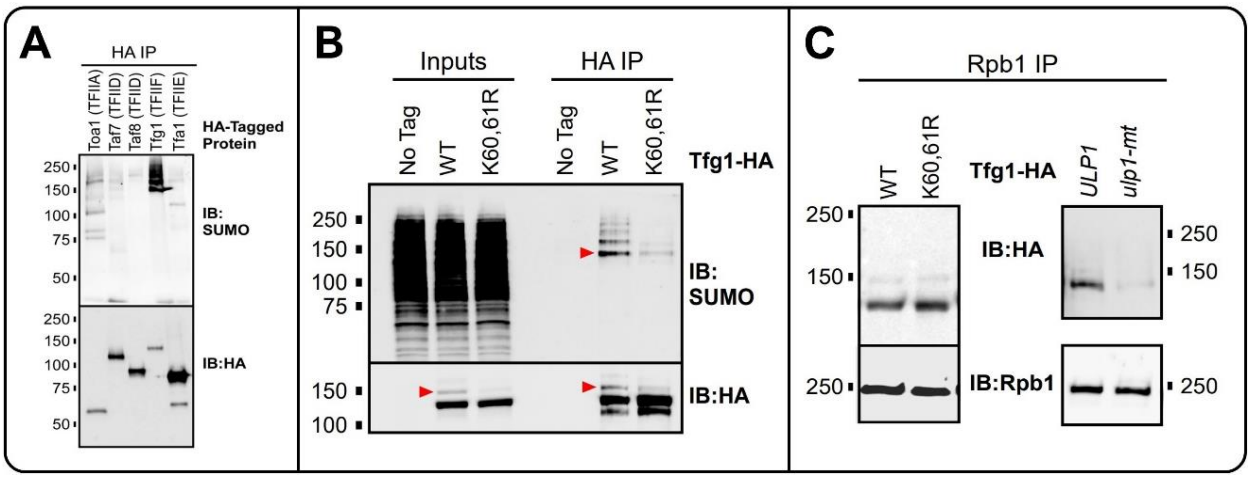

Figure 7. The Tfg1 subunit of TFIIF can be sumoylated at Lys60/61, and globally increased sumoylation reduces RNAPII-TFIIF interactions. These findings were produced from previous research in our lab. **(A)** Five GTF subunits were assayed for SUMO conjugation, and Tfg1 had the highest level of SUMO signal. An anti-HA antibody was used to IP Toa1 (TFIIA), TAF7 (TFIID), TAF8 (TFIID), Tfg1 (TFIIF), and Tfa1 (TFIIE), and anti-HA and anti-SUMO antibodies were used for immunodetection. **(B)** The Lys60/61 residues of Tfg1 are its primary SUMO conjugation sites. The HA Tag was IPed in the wild-type (no tag), *Tfg1-HA*, and *Tfg1-K60,61R-HA* strains. With the *Tfg1-HA* strain, the typical ladder pattern of multiple attached SUMO moieties separated by ~12 kDa could be seen in the SUMO immunoblot. A reduction in SUMO band count and band intensity of monosumoylated Tfg1 (red arrowhead) in HA and SUMO immunoblots demonstrated the reduction in sumoylation when Lys60 and Lys61 were mutated to arginines. **(C)** A reduction in Tfg1 sumoylation does not affect RNAPII-TFIIF interactions, whereas an increase in Tfg1 sumoylation via globally increased sumoylation does. When IPing with an anti-Rpb1 antibody, there was no difference between the *Tfg1-HA* and *Tfg1-K60,61R-HA* strains in HA or control Rpb1 immunoblots. When using the *ulp1-1* (alias: *ulp1-mt*) strain with globally elevated sumoylation levels, there was a decrease in Rpb1-Tfg1 co-IP amounts in the HA immunoblot, when compared to the parent *ULP1-wt* (alias: *ULP1*) strain. Both the *ULP1-wt* and *ulp1-1* strains contained 6xHA epitope tagged Tfg1 (all images from [15]).

1.8 Research Objectives

In order to investigate the uncharacterized role of SUMO at hundreds of non-RPGs, I aimed to continue the research on Tfg1 sumoylation. First, I would need to validate the results of the *ulp1-1* mutant by IPing Tfg1 in a strain with increased Tfg1 sumoylation levels specifically. This way, the reduction in RNAPII-TFIIF interactions could be attributed to Tfg1 SUMO modifications rather than indirect effects deriving from globally increased levels of sumoylation. Then I would utilize chromatin immunoprecipitation to identify changes to Tfg1-chromatin interactions, overall

changes to SUMO levels at promoters, and RNAPII-chromatin occupancy within three strains corresponding to: normal Tfg1, Tfg1 with reduced sumoylation, and Tfg1 with increased sumoylation. With this approach, the impact of Tfg1 sumoylation can be more precisely studied at individual promoters rather than being limited to a genome-wide average with standard immunoprecipitations. This will be integral as Tfg1 sumoylation is not expected to be a regulatory mechanism of transcriptional control at all of the genes transcribed by RNAPII. Particularly, differences can be unveiled between the regulation of RPGs versus select non-RPGs. Following that, I would analyze tangible changes to transcription levels by measuring changes in steady-state RNA levels between strains using real-time quantitative PCR (RT-qPCR). If definitive trends are observed using a handful of benchmark genes, the ultimate goal will be to validate the observations on a genome wide scale using ChIP-seq and RNA-seq. Additionally, chromatin fractionation can be used to gain a better understanding into the mechanism and dynamics of GTF SUMO conjugation by separately analyzing soluble and chromatin bound fractions of target proteins.

In addition to Tfg1 sumoylation, large scale proteomics studies [39], the overlap of SUMO peaks with the TATA-box [15], and published *in vitro* results utilizing human TBP [40], we suspected that the TBP subunit of TFIID was also a SUMO target in *S. cerevisiae*. Despite only detecting faint and diffuse bands in SUMO immunoblots when exploring TFIID sumoylation by IPing the TAF7 and TAF8 subunits (**Fig 7A**), we reasoned that weak interactions or harsher-than-anticipated immunoprecipitation conditions might have dislodged TBP from the complex, masking the detection of its SUMO modifications. In contrast, when detecting TFIIF sumoylation, lab members tagged Tfg1 directly and did not depend on successful co-IPing of Tfg1 by tagging and IPing other subunits of TFIIF. With this in mind, the results of the TAF7 and TAF8 IPs for the detection of TFIID (and thereby TBP) sumoylation were put into question, and we aimed to investigate TBP sumoylation and any resulting functional impacts using the same contemporary

lab techniques described above. Given that TBP (or an analogue) is required for the transcription of nearly all eukaryotic genes [31], the cellular consequences of TBP sumoylation might prove to be even more impactful than Tfg1 sumoylation.

This aim of this study was to understand how Tfg1 and TBP sumoylation at a substantial subset of non-RPGs impacts the regulation of transcription within *S. cerevisiae*. This would be accomplished by utilizing mutant strains with variable levels of sumoylation to assay the functional impacts of Tfg1 and TBP sumoylation on their protein-protein and protein-chromatin interactions, as well as measuring any ensuing changes to RNAPII recruitment and steady-state RNA levels. Given the conserved nature of SUMO modifications, this research could uncover novel characteristics of transcriptional regulation across the Eukarya domain.

Section 2: Materials and Methods

2.1 Yeast Strains and Growth Conditions

All cell growth and subsequent experimentation was performed under sterile conditions: using clean and/or autoclaved equipment beside a lit Bunsen burner, while utilizing appropriate lab etiquette. Strains of *Saccharomyces cerevisiae* were stored in glycerol stocks in -80°C freezers for long term storage. When in active use, strains were streaked out onto synthetic complete (SC; 0.17% YNB, 0.5% ammonium sulfate, and 2% glucose) media plates. The plates were composed of either SC media ("just SC") or a selective dropout formulation variant of the SC media, which was devoid of a certain amino acid, such as tryptophan. After 2-3 days of growth at 30°C, single colonies were inoculated into SC liquid media, with or without a selective dropout formulation, depending on the experiment and strain in question. Plates were then stored at 4°C and used for 5-6 weeks before restreaking from fresh stocks. Inoculated overnight cultures were grown to saturation in a 30°C shaker at at least a 1:5 culture-to-air ratio for sufficient aeration of cells during growth. When appropriate, a second 'day culture' was inoculated at an Optical Density (OD) of 0.25 at 595nm using cells from the overnight culture. This was done in order to collect actively growing (log phase) cells, and to match the cell density of all involved strains despite their innately different growth rates. Following approximately two doublings, when the OD_{595} was typically in the range of 0.6-0.7 for all strains, cells were collected for further experimentation. For reference, cells enter a stationary phase of growth around an OD_{595} reading of 1.0, when grown in standard SC media. All yeast strains used in this study are listed in **Table S1**.

2.2 Cloning Framework

The construction of mutant yeast strains followed an established cloning procedure. In brief, genomic DNA (gDNA) was obtained and purified from an existing strain of yeast, which was typically the 'parental' W303a wild-type lab strain. Then, the relevant section of DNA was PCR amplified and mutations or modifications were introduced to the DNA. For this study, these

encompassed the introduction of point mutations, the insertion of cassettes containing epitope tags and selection markers, and the insertion of a cassette coding for the SUMO protein into the ORF of another protein-coding gene. Cloned DNA was then transformed into fresh cells and genomic metagenesis occurred via homologous recombination. Following growth and selection, the mutations were confirmed via DNA sequencing of selected colonies. The newly established strains of yeast were then stored in -80°C glycerol stocks for long term storage.

2.3 Genomic DNA Extraction

In order to extract genomic DNA, strains of yeast were first inoculated in 15 mL of SC media and incubated in a 30°C shaker overnight. The following day, saturated cultures were spun at 3000 *g* for 5 minutes in a large bench-top centrifuge at room temperature (RT; 24°C). The growth media comprising the supernatant was discarded and the cell pellets were resuspended in 0.5 mL of dH_2O. The suspensions were transferred to microfuge tubes, spun down (13,000 RPM, 2 min, RT), and the supernatants were again discarded. In order to lyse the cells, the pellets were resuspended in 0.2 mL of an Extraction Buffer (2% Triton-X100, 1% SDS, 100 mM NaCl, 10 mM Tris-HCl at pH 8.0, and 1 mM ethylenediaminetetraacetic acid; EDTA), 0.3 mL of acid-washed glass beads (Sigma-Aldrich), and 0.2 mL of a 1:1 mixture of phenol/chloroform (pH 7.9). Samples were then vigorously vortexed for 3 min. To solubilize the DNA, 0.2 mL of TE buffer (10 mM Tris-HCl at pH 7.4 and 1 mM EDTA) was added, samples were mixed, and spun down (13,000 RPM, 5 min, RT). The aqueous layer was transferred to a fresh tube and 1 mL of absolute ethanol was added to precipitate the DNA. The samples were spun down (13,000 RPM, 1 min, RT), the supernatant was discarded, and the DNA was resuspended in 0.4 mL TE buffer (pH 7.4). To remove any RNA contamination, samples were incubated with 3 µL of 10 mg/mL RNase A (EMD Millipore) at 37°C for 5 min. Then, to facilitate re-precipitation of the DNA, 10 µL of 4 M ammonium acetate and 1 mL of absolute ethanol was used. Samples were spun again (13,000 RPM, 2 min, RT), the supernatant was discarded, and the DNA pellet was resuspended in TE buffer (pH 7.4).

A Nanodrop spectrophotometer (Thermo Fisher Scientific) was used to measure gDNA concentration, and all samples were equalized by diluting to 100 ng/μL with TE buffer.

2.4 PCR Amplification and Genomic Mutagenesis

The amplification of DNA obtained directly from NaOH-treated colonies or purified gDNA was accomplished using a three-step standard Polymerase Chain Reaction (PCR). This included denaturing the two strands of DNA apart at 98°C, annealing the primers with the ssDNA at 58°C, and then extending the complementary strand with the use of a Taq or Q5 DNA polymerase at 72°C. After some optimization, the optimal Bio-Rad thermal cycler PCR profile was established: 30 sec of initial denaturation, and then 30 cycles of replication with 10 sec of denaturation, 20 sec of annealing, and 2 min of extension each time, followed by a 5 min final extension. Successful PCR amplification was confirmed by assessing the size of the resulting amplicons after agarose gel electrophoresis. In order to purify the resulting PCR amplicons, a GeneJET Gel Extraction Kit (Thermo Fisher Scientific) was used.

To introduce point mutations, mutagenesis primers were used. These ~40 bp primers were complementary to the region of interest but had one nucleotide swapped out for a different one. This way, the primer still easily annealed to the relevant section of DNA, but following multiple cycles of PCR, the ‘mutant’ amplicons vastly outnumbered the original DNA sequence in the mixture. For all cases in this study, the mutagenesis primers were used to change AAA codons within the yeast genome at the *SPT15* gene (codes for TBP) to AGA codons. In effect, this resulted in specific lysine residues within TBP being mutated to arginine residues. In some cases, fusion PCR was used to combine two overlapping cassettes into a single large cassette using both templates in a single reaction with only two flanking primers. This was done to extend the length of the final PCR product and to enhance the recombination efficiency of point mutations, as some cassettes were found to only recombine partially. All primers used in this study are listed in **Table S2**.

2.5 Transformation and Selection

Once a PCR product was ready for transformation into yeast, the target strain was inoculated in SC medium for overnight growth (30°C) to saturation, followed by a day culture until the OD was in an optimal range (0.6-0.7). Once ready, the cells were harvested via centrifugation (3000 *g*, 5 min, RT), and washed in sterile dH_2O at a volume equal to which the cells were grown in, typically 15 mL. The solution was spun (3000 *g*, 5 min, RT) and the water was discarded. Next, the cell pellet was resuspended in 500 µL of 0.1 M lithium acetate (LiAc), transferred to a microfuge tube, spun (13,000 RPM, 15 sec, RT), and the supernatant was discarded. The pellet was once again resuspended in 500 µL of 0.1 M LiAc to induce competency and a 100 µL aliquot of the cells was spun (13,000 RPM, 15 sec, RT), the LiAc was removed. For the remainder of the transformation mixture, 240 µL of PEG 3500 (50% w/v), 36 µL of 1.0M LiAc, 25 µL of boiled SS-carrier DNA, and 40 µL of ddH_2O were used to facilitate uptake of 10 µL of DNA. The tube containing the cells, DNA, and transformation mixture was vortexed extensively and incubated at 42°C for 30 min to allow for sufficient levels of transformation to take place. The mixture was spun (13,000 RPM, 30 sec, RT), and the transformation mixture was discarded. The cell pellet was resuspended in 80 µL of ddH_2O and plated onto an appropriate solid medium and incubated at 30°C for 2-3 days for selection to take place.

Upon entry inside the cells, the PCR product integrated with the yeast genome via homologous recombination. In order to select for colonies containing the correct insert, a selection marker was present within the PCR product. In this study, the selection marker used was the *Kluyveromyces lactis TRP1* (*K. lactis::TRP1*) marker gene, which coded for the amino acid tryptophan. Therefore, cells were plated onto solid ‘SC-TRP’ plates to ensure exclusive growth of colonies with the insert. Resultant colonies were treated to a Colony PCR protocol, whereby single colonies were assessed for the presence of the genomic mutations. Colonies were transferred into 20 µL of 20 mM sodium hydroxide in a PCR tube and lysed by heating at 95°C for 45 min.

The solution was resuspended in 80 µL of dH_2O, spun in a small microfuge for 5 min, and 5 µL of solution was used as DNA template for a PCR reaction. Once successful amplification was confirmed via agarose gel electrophoresis, the samples were sent for DNA sequencing as a final check.

2.6 Spot Assay

The spot assay was used to assess strain sensitivity to various stressors on solid media plates. Strains of yeast were grown to saturation overnight at 30°C. The following day, the OD was measured, and the cultures were diluted down to equal amounts of cells with ddH_2O. Using a sterile technique, cells were adjacently spotted onto appropriate solid media plates five times, with five-fold dilutions separating each spot of cells horizontally. A multichannel pipette was used to spot multiple strains at once, while keeping all strains vertically equidistant. To stimulate a particular stress pathway, the solid SC media plates either contained additional compounds (incubated at 30°C) or were placed in 37°C to generate heat stress. The stress compounds included 1 M potassium chloride (KCl) for osmotic stress, 1 mM hydrogen peroxide (H_2O_2) for oxidative stress, and 7% ethanol (EtOH) for alcohol stress. For the next 2-3 days of growth, the plates were scanned daily to track and compare growth.

2.7 Growth Curves

For the precise monitoring of culture growth over time, an accuSkan FC (Thermo Fisher Scientific) automated spectrophotometer was utilized. Strains of yeast were inoculated for overnight growth at 30°C, followed by a short-lived and highly diluted day culture. In order to capture the transition from the lag phase to the log phase of growth, the day cultures were only incubated in a 30°C shaker for ~40 min. Once ready, 1 mL of a particular culture was mixed with 250 µL of a 5X stress solution. From this mixture, 300 µL was transferred in vertical triplicates (900 µL in total) into wells of a 96-well microtiter dish. The stress solutions used in this study included: a control SC media-

only (no stress), 1 M sodium chloride for osmotic stress, 1 mM hydrogen peroxide for oxidative stress, and 10% ethanol for alcohol stress. To set a baseline, blanks which contained only SC media were included as well, without any cells or stress solutions. The accuSkan machine gently rocked and incubated the cultures at 30°C and took 80 readings of the OD_{595} of each well, once every 15 min, for a total of 20 hours. Afterwards, the triplicates were averaged, and the blank signal was subtracted.

2.8 Immunoprecipitation (IP)

Immunoprecipitation was done to isolate and purify particular proteins from yeast cells. Strains were inoculated for 10 mL overnight and then 50 mL day culture growth at 30°C. Once the OD was in the optimal range of 0.7-0.8, cultures were harvested in a large bench-top centrifuge (3000 *g*, 5 min, 4°C), and the supernatant was discarded. For the remainder of the IP, samples were kept on ice as much as possible. The cell pellet was washed once with an IP Buffer containing 50 mM Tris-HCl (pH 8), 150 mM sodium chloride (NaCl), 1 mM EDTA (pH 8), along with 0.25% (w/v) N-Ethylmaleimide (NEM), 0.1% Nonidet P-40 (NP-40), 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.1 mM Dithiothreitol (DTT), and a 1X yeast protease inhibitor cocktail (yPIC; Sigma-Aldrich). The samples were spun (3000 *g*, 5 min, 4°C) and the supernatant was discarded. The samples were resuspended in 0.5 mL of IP Buffer again and transferred to a 2 mL microfuge tube. Approximately 0.3 grams of glass beads were added, and samples were vigorously vortexed for 30 min in a 4°C cold room to lyse the cells. The protein lysate was then separated from the glass beads by pipetting to a fresh 1.5 mL tube, and an additional 0.5 mL of IP Buffer was added to the 2 mL tube containing the glass beads. The tube was inverted several times to collect any residual lysate and this solution was then pooled with the lysate in the 1.5 mL tube. The lysate was then spun twice (14,000 RPM, 5 min, 4°C), and each time the supernatant was transferred to a fresh tube. At this point, 50 µL of lysate was removed to serve as input and was mixed with an equal volume of 2X SDS sample buffer (140 mM Tris-HCl at pH 6.8, 4% SDS, 20% glycerol, 0.02%

Bromophenol Blue dye, and 10% 2-mercaptoethanol) and then boiled at 95°C for 4 min. The remainder of the lysate was transferred to a tube containing either 15 µL of pre-conjugated HA antibody-beads (Sigma-Aldrich) or 15 µL of Protein G beads (Sepharose) along with 1 µg of 8WG16 antibody (Abcam) for detecting RNAPII and was left to rotate overnight in a 4°C cold room for the IP to occur. In either case, the beads were first washed in IP Buffer (plus 0.1% NP-40) and the supernatant was removed after a short spin (2500 RPM, 1 min, 4°C). The following day, the samples were washed in ice cold IP buffer (plus 1% NP-40 and 1 mM PMSF) three times and then twice with just IP Buffer to remove any detergent. To elute, the samples were resuspended in 100 µL of 2X SDS sample buffer and boiled for 4 min at 95°C.

2.9 Immunoprecipitation Under Denaturing Conditions

Under denaturing conditions, the IP protocol was modified substantially. After the initial pelleting, the cells were washed in 10 mL of denaturing 20% Trichloroacetic acid (TCA; Sigma-Aldrich) instead of IP buffer, and then resuspended in 250 µL of TCA. Approximately 0.3 grams of glass beads were added, and the samples were vigorously vortexed for 3 min. The lysates were then transferred to fresh tubes, washed with 500 µL of 5% TCA twice, and the TCA was then removed. All samples were resuspended in 750 µL of a modified 1X SDS sample buffer composed of 60 mM Tris (pH 6.7), 5% 2-mercaptoethanol, 10% SDS, and a few drops of bromophenol blue dye. After resuspension, an additional 100 µL of 1 M Tris (un-pH'ed) was added and samples were boiled at 95°C for 5 min. Samples were spun (13,000 RPM, 10 min, RT), and 40 µL was set aside for input. Of the remaining lysate, 400 µL was added to 4 mL of a special IP Buffer that contained Bovine Serum Albumin (BSA) to retain denatured protein solubility. Precisely, this buffer was composed of 50 mM Tris (pH 7.4), 150 mM NaCl, 0.5% NP-40, and 0.5 mg/ml BSA. To this 1:10 lysate dilution, 25 µL of appropriate antibody-conjugated beads (which had been washed in IP Buffer containing BSA) were added and the tubes were rotated overnight in a 4°C cold room.

Samples were washed in standard IP buffer thrice the next day, resuspended in 60 µL of normal 2X SDS sample buffer, and then boiled at 95°C for 4 min to elute.

2.10 Western Blotting

Western blots were performed to detect and identify proteins in cell extract (typically IP) samples. Fresh SDS-PAGE gels with variable levels of 30% Acrylamide/Bis Solution (Bio-Rad) were prepared each time, resulting in 7.5%, 10%, or 12.5% gels. Protein samples were re-boiled for 90 seconds prior to loading into gels, and the gels were run at 100 volts (V) for approximately 2 hours. The proteins were then transferred to either nitrocellulose (0.45 µm; Bio-Rad) or activated PVDF membranes (0.45 µm; Millipore) via semi-dry transfer (25 V, 7 min) if the proteins had a low molecular weight or wet transfer (100 V, 55 min) if the proteins had a high molecular weight. The membranes were then blocked with 5% milk on a shaker for 60 min. The milk was composed of skim milk powder (BioShop) which had been homogenized in a 1X PBST solution with 1% Phosphate-buffered saline (PBS; Thermo Fisher Scientific) and 0.05% Tween-20 (Thermo Fisher Scientific). Membranes were sealed in plastic bags with 2 mL solutions of appropriately diluted primary antibodies in 5% milk and left to incubate on a nutator in a 4°C cold room overnight. The primary antibodies used in this study were: anti-HA (Novus; rabbit, 1:3000), anti-Smt3 (Santa Cruz, rabbit, 1:500), 8WG16 (anti-Rpb1; Abcam, mouse, 1:3000), anti-Ser2p (Abcam, rabbit, 1:3000), anti-Ser5p (Abcam, mouse, 1:3000), anti-Histone H3 (Abcam, rabbit, 1:3000), and anti-GAPDH (Sigma-Aldrich, rabbit, 1:5000). The next day, the membranes were washed with 1X PBST three times and incubated with a 20 mL solution of horse-radish peroxidase (HRP) conjugated secondary antibodies (1:5000 anti-rabbit or anti-mouse IgG; Thermo Fisher Scientific) in 1% milk for 60 min. Before detection, the membranes were washed with 1X PBST three times again and incubated in two chemiluminescent reagents simultaneously: an oxidizing reagent and a luminol reagent (Bio-Rad). After a 4 min incubation, the membranes were imaged using a

MicroChemi chemiluminescence imager (DNR Bio Imaging Systems) for various exposure lengths.

2.11 RNA Extraction and cDNA Synthesis

To detect changes in steady-state RNA levels, RNA was extracted from yeast cells and converted to complementary DNA (cDNA) to measure differences with real-time quantitative PCR. Following overnight growth, 10 mL day cultures were grown at 30°C until an OD of 0.6-0.7 was reached. Cells were harvested by spinning (3000 *g*, 5 min, 4°C), and were resuspended in 1 mL of ice-cold AE Buffer: 50 mM sodium acetate (NaOAc; pH 5.2) and 10 mM EDTA (pH 8.0), prepared with nuclease-free water. Cells were washed in 1 mL of AE Buffer twice and then resuspended in 400 µL of AE Buffer again. As a detergent, 40 µL of 10% SDS was added and mixed. To denature DNA, 440 µL of phenol (pH 4.5) was added and mixed. To lyse, the following sequence was carried out twice: cells were flash frozen in a dry ice and ethanol bath for 5 min, incubated at 65°C for 5 min, and then vortexed vigorously for 30 sec. The resulting lysates were then flash frozen for another 5 min. Samples were spun (13,000 RPM, 7 min, RT) and the aqueous layer was transferred to new microfuge tubes. Next, 600 µL of phenol-chloroform (pH 4.5) was added, samples were vortexed and spun (13,000 RPM, 5 min, RT), and the aqueous layer was again transferred to fresh tubes. To precipitate the RNA, 50 µL of 3M NaOAc (pH 5.2) was added along with 1 mL of absolute ethanol, and the samples were left at -20°C overnight. Later on, samples were centrifuged (14,000 RPM, 15 min, 4°C) and the RNA pellet was washed in 70% ethanol by mixing, spinning (14,000 RPM, 5 min, 4°C), and decanting the supernatant. To completely remove the ethanol, samples were spun again (14,000 RPM, 2 min, 4°C), the residual liquid was removed by pipetting, and the pellet was allowed to air dry for approximately 5 min. The samples were resuspended in 100 µL of nuclease-free water.

Any potential DNA contaminant was eliminated via DNase treatment. For this 200 µL reaction, 24 µg of an RNA sample was incubated with 20 µL of DNase I Buffer (Bio-Rad), 2 µL of

DNase I enzyme (Bio-Rad), and an appropriate amount of DEPC-treated water at 37°C for 10 min. Next, 2 µL of 0.5 M EDTA (pH 8) was added and samples were heated at 75°C for 10 min to inactivate the DNase I enzyme. Complementary DNA synthesis was performed using an iScript cDNA Synthesis Kit (Bio-Rad). In this reaction, 6 µL of DNase-treated RNA was combined with 6 µL of 5x reverse-transcription reaction mix, 1.5 µL of the reverse transcriptase enzyme, and 16.5 µL of nuclease-free water in a small PCR tube. The tube was placed in a thermocycler and the following template was used: 5 min at 25°C, 30 min at 42°C, 5 min at 85°C, and a final 4°C hold step. The final cDNA sample was diluted in water and stored at -20°C.

2.12 Real-Time Quantitative PCR (RT-qPCR)

The relative levels of cDNA across samples were compared using RT-qPCR. For each qPCR reaction, 5 µL of the cDNA sample, 0.7 µL of each primer, and 7 µL of SYBR Green Master Mix (Froggabio; contained dNTPs, $MgCl_2$, and DNA polymerase) were mixed, and placed in a RotorGene-Q quantitative PCR machine (Qiagen). Each reaction was done in duplicates to ensure consistency. The thermal cycling template used had an initial 95°C denaturation step for 5 min, followed by 40 cycles of 95°C denaturation for 5 seconds and 60°C annealing + extension for 10 seconds. To normalize, primers amplifying the constitutively expressed 25S rRNA were included in each qPCR trial. RT-qPCR experiments were repeated at least three times with fresh cultures to ensure reproducibility, and the results of multiple trials were averaged.

2.13 RNA Sequencing (RNA-seq)

For the transcriptome-wide analysis of wild-type vs. mutant yeast strains, RNA was extracted and then sequenced in parallel via high-throughput RNA sequencing. For this protocol, every precaution was taken to prevent RNAse contamination of the samples. Before starting, the work area and pipettes were cleaning with RNAseZap (Sigma-Aldrich), and filter tips were used whenever appropriate. RNA was extracted as described above, again with 10 mL day cultures.

To further purify the RNA, however, a RNeasy Kit (Qiagen) was used. An aliquot of the final eluate was separated to quantify the concentration of the samples using a Qubit Fluorometer (Thermo Fisher Scientific) and to assess the samples on an agarose gel. Once ready, samples were sent to a lab at SickKids Toronto for RNA sequencing, and the exact NGS protocol and data analysis tools used can be found in **Table S3.**

2.14 Chromatin Immunoprecipitation (ChIP)

Chromatin Immunoprecipitation, coupled with quantitative PCR, was used to assess differential protein-chromatin interactions. For this study, differential Tfg1-, TBP-, SUMO-, and RNAPII-chromatin interactions across various strains were assessed. For ChIP, 50 mL day cultures were grown at 30°C until cells were in an exponential growth state. Flasks were transferred to a lab bench and incubated with 5 mL of an 11% formaldehyde solution (37% formaldehyde stock; Thermo Fisher Scientific) diluted in 140 mM NaCl, 1 mM EDTA, and 70 mM HEPES-KOH at pH 7.5) at RT for 20 min with occasional rocking to induce cross-linking. Samples were quenched with 7 mL of 2.5 M Glycine for 5 min and transferred to 50 mL falcon tubes on ice. Cells were pelleted in a large centrifuge (3000 *g*, 5 min, 4°C), and the pellet was washed in 40 mL of 1X Tris-buffered saline (TBS; 20 mM Tris-HCl at pH 7.5, 150 mM NaCl) twice and the supernatant was discarded. The cells were resuspended in 1 mL of ChIP lysis buffer (50 mM HEPES-OH at pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, and 0.1% SDS), transferred to a 2 mL screw-top tube, spun down (14,000 RPM, 5 min, 4°C), and the supernatant was then discarded. To flash freeze, samples were briefly placed in a tank of liquid nitrogen, and then transferred to a -80°C freezer.

To lyse, samples were resuspended in 600 µL of ChIP lysis buffer containing an additional 20mM PMSF and 1X yPIC along with 400 µL of acid-washed glass beads and then placed in a Mini-Beadbeater (BioSpec). The machine was run at the “homogenize” setting three times for a minute each time. Lysates were separated from the beads by using a sterile needle to poke a

small hole under the tubes, placing the tubes on top of polystyrene sonication tubes within a 14 mL culture tube, and spinning (950 *g*, 4 min, 4°C) to transfer the lysates to the sonication tubes. The chromatin was then sheared into ~500 bp fragments using a cup horn sonicator (Sonic Dismembrator Q700/FB-705) at 4°C. After an initial de-gas of the water bath, samples were subjected to twelve (30 sec) rounds of sonication at 40% amplitude. Samples were spun (14,000 RPM, 5 min, 4°C), and the supernatant was transferred to a fresh tube. The volume was adjusted to 1 mL via the addition of ChIP lysis buffer (plus PMSF and yPIC) and 15 µL of 5 M NaCl was added before aliquoting out 40 µL of each sample for input. The remainder of the lysate was used whole or split into parts and incubated at 4°C overnight with antibodies pre-conjugated to magnetic Protein G beads (Thermo Fisher Scientific).

For the preparation of the antibody-beads complex, 15 µL of beads and 1 µg of various antibodies (anti-HA, anti-Smt3, anti-Rpb1, or anti-Rpb3) were rotated together at 4°C in 1 mL of ChIP lysis buffer. Before use, the beads (with bound antibodies) were washed thrice with ChIP lysis buffer, and each time the beads were secured by placing the tubes on a magnet while the supernatant was removed. The following day, the samples were sequentially washed in four wash buffers at RT with four-minute rotations each time: (1) ChIP lysis buffer + 150 mM NaCl, (2) ChIP lysis buffer + 300 mM NaCl, (3) 10 mM Tris-HCl (pH 8.0), 0.25 M lithium acetate, 1 mM EDTA, 0.5% NP-40, and 0.5% sodium deoxycholate, and (4) TE buffer at pH 8 (10 mM Tris-HCl at pH 8.0 and 1 mM EDTA). After the final wash was removed, the beads were resuspended in 250 µL of a ChIP elution buffer (50 mM Tris-HCl at pH 7.5, 10 mM EDTA, and 1% SDS) and tubes were incubated at 65°C for 20 min to facilitate elution. The supernatants were transferred to a fresh set of tubes and the beads were rewashed in 250 µL of TE buffer (pH 8), and this was pooled with the previous eluate of each sample. To eliminate any RNA contaminants, samples were incubated with 10 µL of 10 mg/mL RNase A (EMD Millipore) at 37°C for 30 min. To eliminate any protein contaminants, samples were incubated with 10 µL of 20 mg/mL Proteinase K (BioShop) at 42°C

for 60 min. Cross-links were reversed by incubating samples at 65°C overnight. DNA samples were purified with a GeneJET Gel Extraction Kit (Thermo Fisher Scientific), and 150 µL of TE buffer (pH 7.4) was used as the final elution buffer.

The relative levels of specific DNA segments across ChIP samples were compared using qPCR. The same protocol, reagent volumes, and thermal cycling template as described above were used. To normalize, however, primers amplifying an untranscribed region of yeast chromosome V (Chr. V) were included in each qPCR trial. A 'fold over background' signal of 1 represented this background level of signal. The ChIP input samples were not typically used for normalization but were continually collected and stored regardless. Each ChIP-qPCR experiment was repeated at least three times with fresh cultures to ensure reproducibility, the results of multiple trials were averaged, and any deviations were represented as standard error bars.

2.15 Chromatin Immunoprecipitation Sequencing (ChIP-seq)

To analyze genome-wide trends, a slightly modified ChIP-seq protocol was used. Unlike with standard ChIP, relatively large 250 mL day cultures were grown (and all other reagent amounts were proportionally increased), the chromatin was sheared into ~250 bp fragments with sixteen rounds of sonication for increased precision (instead of ~500 bp with twelve rounds), and input samples were processed along with the eluates starting from the RNase A step in the ChIP protocol detailed above. Multiple trials of ChIP using this ChIP-seq protocol were performed and qPCR of a select number of genes was used to quickly test and compare trials. Once ready, input and eluate samples from two separate replicates were sent to a lab at SickKids Toronto to sequence the samples using next-generation sequencing (NGS). Details regarding the exact NGS protocol and data analysis tools used can be found in **Table S4**.

2.16 Chromatin Fractionation

Chromatin fractionation was used to isolate and analyze TFs in soluble versus chromatin-bound states. Wild-type and mutant yeast strains were grown at 30°C in 25 mL day cultures until an OD of ~0.5 was reached. Cultures were harvested by spinning (3000 *g*, 5 min, RT) and then resuspended in 6.25 mL of Resuspension Buffer 1 (0.1 M PIPES/KOH at pH 9.4 and 0.01 M DTT). Cultures were incubated in a 30°C shaker for 10 min. Cells were spun down again (3000 *g*, 5 min, RT), resuspended in 2.5 mL of Resuspension Buffer 2 (600 mM Sorbitol and 25 mM Tris at pH 7.5 in YPD medium), a 50 µL solution of the enzyme Zymolyase (5 U/µL; BioShop) was added, and the samples were incubated in a 30°C shaker for 20 min. Samples were spun down (2000 RPM, 3 min, RT) and the pellet was resuspended in Resuspension Buffer 3 (0.7 M Sorbitol and 25 mM Tris at pH 7.5 in YPD medium). The samples were again incubated in a 30°C shaker for 20 min, spun down (3000 *g*, 3 min, RT), and the pellet was washed three times with a lysis buffer (400 mM Sorbitol, 150 mM potassium acetate at pH 7.0, 2 mM magnesium acetate, 20 mM PIPES/KOH at pH 6.8, 1 mM PMSF, 0.25% NEM, and 1X yPIC) before the cells were resuspended in 400 µL of the same lysis buffer. As a detergent, 1% Triton X-100 was added. Samples were mixed gently and incubated on ice for 5 min.

For the whole cell extract (WCE), 90 µL of lysed material was removed and mixed with 90 µL of 2X SDS sample buffer. An additional 100 µL of the lysed material was removed and spun down (14,000 RPM, 15 min, 4°C) to separate the soluble and chromatin fractions. The ~95 µL of supernatant containing the soluble proteins was removed and mixed with 95 µL of 2X SDS sample buffer, this formed the soluble extract (SE). The remaining chromatin pellet was washed in 100 µL of the lysis buffer, the supernatant was removed, and the low-salt chromatin (LSC) pellet was resuspended in 100 µL of 2X SDS sample buffer. The WCE, SE, and LSC extracts were boiled at 95°C for 5 min and examined by immunoblotting.

Section 3: Results

3.1 SUMO-Tfg1 fusion results in elevated and constitutive Tfg1 sumoylation

In order to gauge the effects of increased levels of Tfg1 sumoylation, past lab members used a *ulp1-1* mutant yeast strain. In that strain, the desumoylating protease Ulp1 was functionally defective and global levels of sumoylation were increased, not just Tfg1's levels specifically. Due to the multitude of cellular consequences of increased global sumoylation, the results obtained using the *ulp1-1* mutant strain could not be attributed to increased Tfg1 sumoylation alone. To address this, a new strain was created to exclusively increase the sumoylation level of endogenous Tfg1. Lab members Yimo Dou and Dr. J. Bryan McNeil worked to create a strain in which a SUMO peptide was artificially fused to the N-terminus of Tfg1, and this new strain was called *SUMO-Tfg1-HA*. For detection, Tfg1 in this strain also contained a 6xHA epitope tag at its C-terminus. This version of Tfg1 could still be sumoylated at its normal K60/61 (and other) sites, but now had an uncleavable SUMO peptide continuously attached to its N-terminus. It was uncleavable because the fused SUMO had its C-terminal diglycine motif deleted, preventing its enzymatic cleavage, and thereby resulting in constitutive sumoylation.

To confirm the increase in Tfg1 sumoylation using the *SUMO-Tfg1-HA* strain, I performed an IP to isolate and purify Tfg1 using an anti-HA antibody. The resulting proteins were immunodetected by western blot using anti-HA and anti-SUMO antibodies. For comparison, four strains were used: the *W303a* wild-type parent, just tagged *Tfg1-HA*, the *Tfg1-K60,61R-HA* strain with reduced sumoylation, and the *SUMO-Tfg1-HA* strain with increased sumoylation. As expected, attaching SUMO directly to Tfg1 increased its sumoylation level and, notably, the level of sumoylation on the SUMO-fused form was dramatically higher than normal Tfg1 sumoylation levels (**Fig 8A**). To determine whether the very intense 'hypersumoylated' SUMO signal derives from SUMO chains forming on the fused SUMO moiety, an additional strain was created by Dr. J. Bryan McNeil. This new *monoSUMO-Tfg1-HA* strain still had a SUMO moiety artificially attached to its N-terminus (with a deleted diglycine motif), but this particular SUMO peptide had

all of its lysine residues mutated to arginines, in an attempt to prevent the rampant polysumoylation seen with the *SUMO-Tfg1-HA* strain. In theory, we could use this new strain to study the functional impact of constitutively, but not necessarily hyper-, sumoylated Tfg1. Another set of IPs and WBs were done using three strains: just tagged *Tfg1-HA*, the *Tfg1-K60,61R-HA* strain with reduced sumoylation, and the *monoSUMO-Tfg1-HA* strain with a (N-terminally) non-polysumoylatable SUMO peptide constitutively fused to it (**Fig 8B and C**).

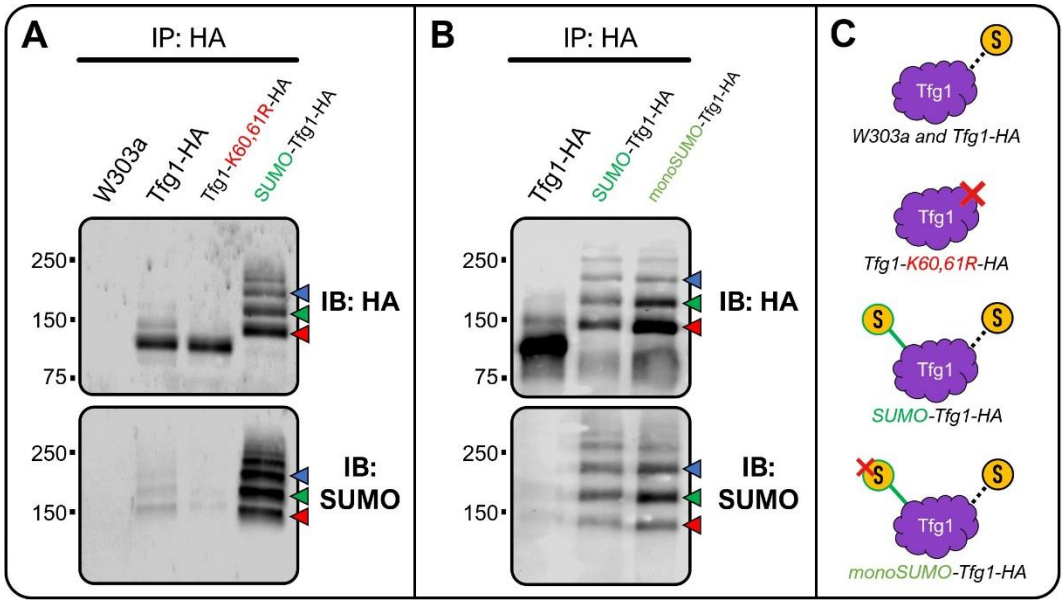

Figure 8. Fusing SUMO to Tfg1 results in Tfg1-specific hypersumoylation. In these blots, Tfg1 runs at ~110 kDa, monosumoylated Tfg1 runs at ~140 kDa (red arrowhead), doubly-sumoylated Tfg1 runs at ~170 kDa (green arrowhead), and triply-sumoylated Tfg1 runs at ~200 kDa (blue arrowhead). **(A)** Proteins from four strains were IPed for the HA epitope tag and immunodetected with anti-HA and anti-SUMO antibodies: *W303a* wild-type parent, just tagged *Tfg1-HA*, the *Tfg1-K60,61R-HA* strain with reduced sumoylation, and the *SUMO-Tfg1-HA* strain with increased sumoylation. The SUMO-Tfg1 fusion protein was found to be highly sumoylated, displaying intense SUMO bands in both HA and SUMO immunoblots. Since endogenous Tfg1 was constitutively sumoylated in the *SUMO-Tfg1-HA* strain, a nonsumoylated variant of Tfg1 was not present in that sample. **(B)** The addition of the *monoSUMO-Tfg1-HA* strain containing a SUMO peptide which could not be polysumoylated did not remedy the hypersumoylated status of the Tfg1 fusion protein. This meant that polysumoylation on the artificially fused SUMO was not entirely responsible for the SUMO bands seen with the *SUMO-Tfg1-HA* strain. The just tagged *Tfg1-HA* strain had relatively dim SUMO bands. **(C)** Cartoon illustrations of Tfg1 in the five strains utilized. Wild-type Tfg1 in the *W303a* and *Tfg1-HA* strains can be sumoylated at a particular location (Lys60/61), as indicated with a dashed line. When those residues are mutated to arginines in the *Tfg1-K60,61R-HA* strain, SUMO conjugation is blocked. SUMO is constitutively fused to the N-terminus of Tfg1 in the *SUMO-Tfg1-HA* strain, as indicated with a solid green line, but the fused SUMO peptide cannot be polysumoylated in the *monoSUMO-Tfg1-HA* variant. In both fusion strains, sumoylation can still occur at Lys60/61.

Surprisingly, preventing N-terminal polysumoylation had little impact on Tfg1 sumoylation, and it was found to retain its hypersumoylated status. A closer look revealed that the HA and SUMO

immunoblot bands in the *monoSUMO-Tfg1-HA* samples were more intense than in the *SUMO-Tfg1-HA* samples. A possible explanation for this was that the hypersumoylation within the *SUMO-Tfg1-HA* strain was negatively impacting Tfg1 stability and was thereby reducing the SUMO signal from those samples as well. This was supported by Dr. J. Bryan McNeil's growth curve assays, which revealed that the *SUMO-Tfg1-HA* strain suffers from a slight growth defect, compared to the just tagged *Tfg1-HA* strain. The *monoSUMO-Tfg1-HA* yeast do not suffer from a growth defect and grow at about the same rate as the *Tfg1-HA* yeast (**Fig S1**). A GAPDH normalization control using the input samples supported this notion, as the GAPDH signal from the *monoSUMO-Tfg1-HA* samples was slightly more intense than in the *SUMO-Tfg1-HA* samples, indicating superior Tfg1 stability in the *monoSUMO-Tfg1-HA* strain (**Fig S2**).

A secondary observation via plain visual examination was that the doubly- (green arrowhead) and triply-sumoylated (blue arrowhead) bands in both fusion strains seemed to be more concentrated than the singly-sumoylated bands (red arrowhead; SUMO immunoblot in **Fig 8B**). This could have been attributed to an increased number of SUMO epitopes on the multi- or poly-sumoylated proteins, however an exploration of various SUMO immunoblots (such as the one in **Fig 7B)** established that this phenomenon is not typically observed. The limited *in vivo* presence of multi- and/or polysumoylated substrates generally translates to dimmer multi- and/or polysumoylated bands on immunoblots. Perhaps in line with the group sumoylation hypothesis (as described in section 1.2), the presence of even a single constitutive, uncleavable SUMO modification on Tfg1 might have induced cooperative SUMO conjugation on other parts of Tfg1 (as discussed in section 4.1 below). Hypersumoylation resulting from fusing SUMO to other proteins has been reported in the literature before, such as when SUMO was artificially fused to the Gcn4 transcription factor [41]. Regardless, the *ulp1-1* strain with globally increased sumoylation indicated that increased Tfg1 sumoylation disrupted RNAPII-TFIIF interactions, and this could now be validated using strains which were more representative of *in vivo* conditions.

3.2 Increased Tfg1 sumoylation impairs Tfg1-RNAPII interactions

To assess the impact of increased levels of Tfg1-specific sumoylation on RNAPII-TFIIF interactions, a standard co-IP and subsequent WB protocol was followed. Cells from the following strains were cultivated for IPs with an anti-Rpb1 antibody: the *W303a* wild-type parent, just tagged *Tfg1-HA*, the *Tfg1-K60,61R-HA* strain with reduced sumoylation, and the *SUMO-Tfg1-HA* strain with increased sumoylation. An anti-Rpb1 antibody was used to IP RNAPII and co-IP any interacting Tfg1. For immunoblotting, an anti-Rpb1 antibody was used to confirm successful pulldowns, and an anti-HA antibody was used to assess differential RNAPII-TFIIF interactions. The results validated indications from previous lines of evidence: elevated Tfg1 sumoylation was found to correlate with a reduction in RNAPII-TFIIF interactions (**Fig 9A**). Only nonsumoylated Tfg1 was found to interact with Rpb1. After the *monoSUMO-Tfg1-HA* strain was created, it was analyzed using the same procedure and demonstrated the same effect (**Fig 9B**).

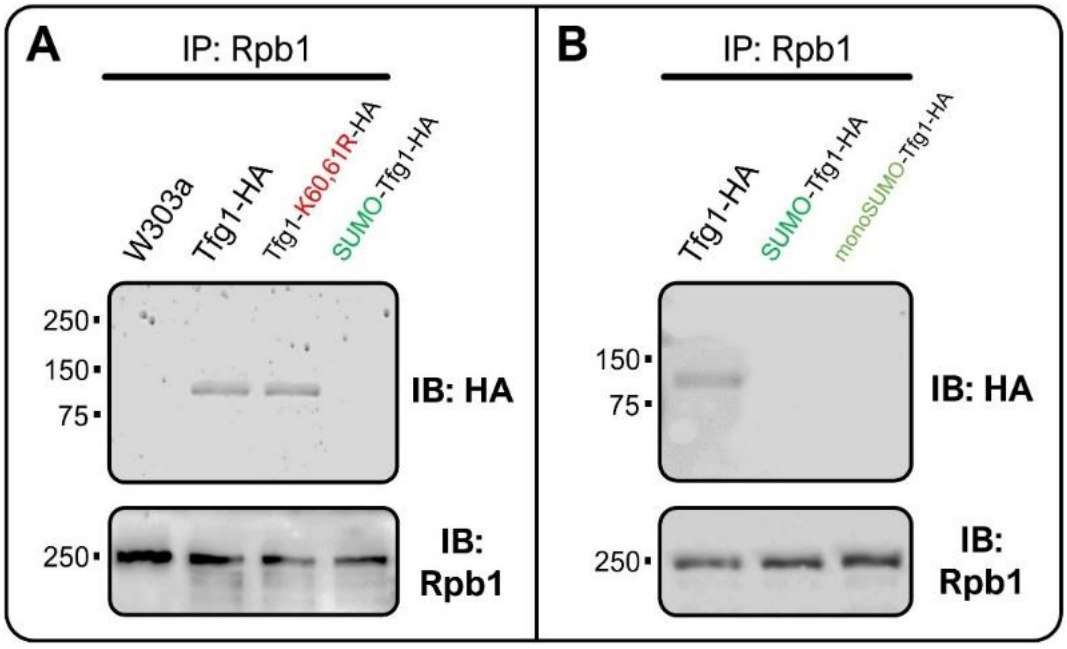

Figure 9. Fusing SUMO to Tfg1 disrupts RNAPII-TFIIF interactions, validating *ulp1-1* results. In these blots, nonsumoylated Tfg1 runs at ~110 kDa and Rpb1 runs at ~250 kDa. **(A)** A reduction in Tfg1 sumoylation does not impact RNAPII-TFIIF interactions, only a constitutive increase in Tfg1 sumoylation via the SUMO-Tfg1 fusion protein reduces Rpb1-Tfg1 co-IP amounts, and thereby RNAPII-TFIIF interactions. In general, sumoylated Tfg1 was not detected, implying that RNAPII does not interact with SUMO modified Tfg1. Proteins from the *W303a*, *Tfg1-HA*, *Tfg1-K60,61R-HA*, and *SUMO-Tfg1-HA* strains were IPed with an anti-Rpb1 antibody and HA and Rpb1 immunoblots were used to assess differential pulldown. **(B)** The addition of the *monoSUMO-Tfg1-HA* strain demonstrated that whether the artificially fused SUMO peptide could be polysumoylated or not did not affect the reduction in RNAPII-TFIIF interactions.

Interestingly, even the conjugation of a single constitutive SUMO peptide was enough to completely disrupt RNAPII-TFIIF interactions within immunoblots. Though, as stated before, even the presence of a single uncleavable SUMO modification seemed to be sufficient to induce or maintain a very high level of Tfg1 sumoylation. For this reason, only the *SUMO-Tfg1-HA* strain was used for subsequent experimentation, not the *monoSUMO-Tfg1-HA* strain. Given that only a minor growth defect was seen with both fusion strains, it is likely that some level of RNAPII-TFIIF interactions persisted but were too weak to be detectable within these co-IPs, as a complete abolishment of RNAPII-TFIIF interactions would likely be fatal. This is because TFIIF activity is essential for cell viability, as deletion mutants do not survive [42]. It is also possible that additional cellular regulatory mechanisms compensated for the lack of appropriate RNAPII shuttling via TFIIF. Due to the limits associated with using the IP-WB method of assessing transcription machinery protein-protein interactions, it could not be asserted how Tfg1 sumoylation was impacting interactions and activity at the relevant non-RPGs in particular. With IP-WBs, only the overall cellular average can be assessed. While IP-WBs indicated no noticeable decrease in RNAPII-TFIIF interactions with decreased Tfg1 sumoylation (using the *Tfg1-K60,61R-HA* strain), it remained possible that the interactions were reduced exclusively at the small subset of non-RPGs with detected SUMO peaks. Similarly, IP-WBs could not confirm whether increased Tfg1 sumoylation (via the *SUMO-Tfg1-HA* strain) necessarily or exclusively disrupted interactions at the same subset of non-RPGs. To address these gaps, the next step was to utilize ChIP to examine differential Tfg1-chromatin association, overall promoter sumoylation, and RNAPII density at specific gene promoters.

3.3 Tfg1 – HA ChIP-qPCR: fluctuating but overall unchanged Tfg1 association

Research in the field has indicated that transcription factor sumoylation can affect TF-chromatin interactions [13]. ChIP-qPCR using an anti-HA antibody was utilized to investigate how variable Tfg1 sumoylation affected Tfg1-chromatin association. For comparison, three strains were used:

just tagged *Tfg1-HA*, the *Tfg1-K60,61R-HA* strain with reduced sumoylation, and the *SUMO-Tfg1-HA* strain with increased Tfg1-specific sumoylation. Five gene promoters were selected for convenient analysis due to the sample limit of the RotorGene-Q qPCR machine: the *TDH1* non-RPG with no SUMO peak as a control, the *TDH3*, *PYK1*, and *PDC1* non-RPGs with GTF associated SUMO peaks, and the *RPS20* RPG with a Rap1 associated SUMO peak, rather than a GTF associated peak. Four independent trials using fresh cultures were completed and the results were averaged. For the normalization of all ChIP-qPCRs in this project, an untranscribed region of chromosome V was included to establish background levels of activity. We did not expect any deliberate transcription machinery or SUMO presence at this region of ChrV. A fold value of 1 represented this background level of signal. Overall, no significant change in Tfg1-chromatin interactions was observed based upon variable Tfg1 sumoylation at all tested gene promoters (**Fig 10**).

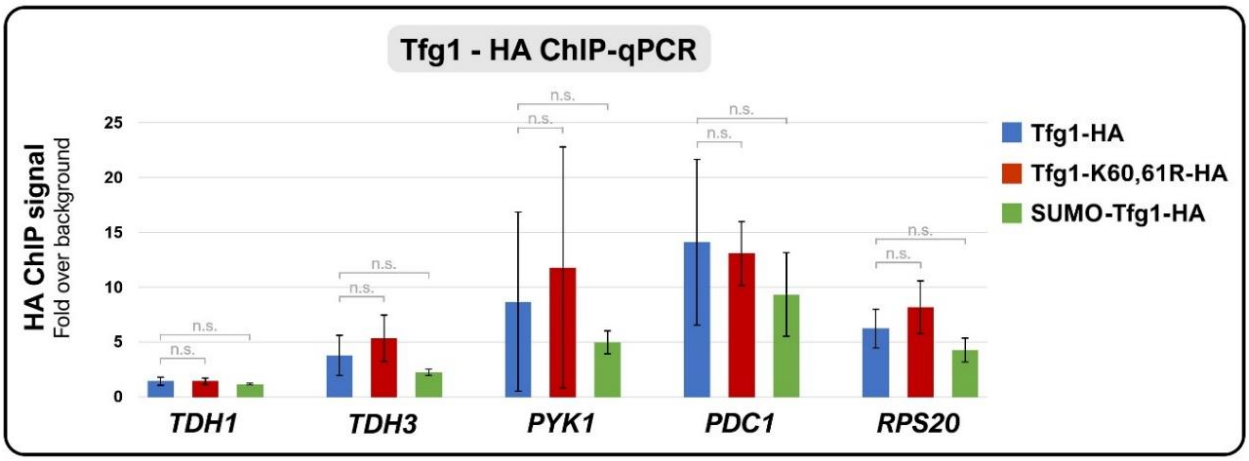

Figure 10. Tfg1 - HA ChIP-qPCR: Tfg1 sumoylation does not impact Tfg1's chromatin occupancy. While great variability was seen between individual trials, presumed to be due to the highly transient nature of TF-chromatin interactions, no significant changes were seen once the results were averaged. For this ChIP, an anti-HA antibody was used to pull down DNA segments cross-linked to Tfg1 in the following strains: *Tfg1-HA*, *Tfg1-K60,61R-HA* with reduced sumoylation, and *SUMO-Tfg1-HA* with increased sumoylation. Five gene promoters were assessed for differential Tfg1 binding to chromatin: the *TDH1* non-RPG with no SUMO peak as a background control, the *TDH3*, *PYK1*, and *PDC1* non-RPGs with GTF associated SUMO peaks, and the *RPS20* RPG with a Rap1 associated SUMO peak. The average of four trials was calculated, the standard deviation was represented with error bars, and the Student's *t*-test was used to assess significant changes (*p* value less than 0.05).

While no significant change was evident upon averaging the ChIP-qPCR trials, there was a great deal of variability between samples within individual trials, as suggested by the large error bars. Before the full scope of the project and the optimized ChIP protocol were established, preliminary ChIP trials using only two of the three strains (*Tfg1-HA* and *Tfg1-K60,61R-HA*) and a different set of genes were conducted and displayed a similarly high degree of inconsistency. Depending on the trial in question, decreased Tfg1 sumoylation was found to have either a positive or negative effect on Tfg1-chromatin association. Likewise, when the *SUMO-Tfg1-HA* strain was added, increased sumoylation also gave contradictory results between trials.

There were three potential explanations for this. First, this variability could have simply been indicative of Tfg1 sumoylation not having any particular effect on Tfg1's chromatin interactions, and thereby the differences seen between trials were simply naturally occurring fluctuations in protein density at the selected promoters. Thereby, these random fluctuations might have manifested as contradictory results between individual trials but could have simply been representative of the absence of a meaningful influence/effect. Second, it remained possible that our ChIP protocol was inadequate at accurately capturing the TF-chromatin interactions, which tend to be highly transient and in rapid flux by their nature. To counteract this lab-wide concern, dual crosslinking ChIP (dual-X-ChIP) was briefly explored but not pursued in this project due to time concerns [43]. The other speculative possibility was that there may have been an unknown variable at play, such as the always-in-flux current average rate of transcription of the culture. Not all yeast cultures grow at identical rates, and there can be a great deal of technical variability ("batch effects") in the growth of even the same strain in successive trials due to a multitude of molecular factors [44]. Even though attempts were made to equalize culture density between strains and between trials, some level of irregularity was unavoidable.

Regardless, the averaged results indicated that SUMO conjugation to Tfg1 did not affect its chromatin occupancy. Given this, the actual influences of variable Tfg1 sumoylation on protein-protein interactions at the target promoters could be studied with higher confidence as Tfg1 levels on chromatin were understood as being more-or-less unchanged between strains.

3.4 Tfg1 – SUMO ChIP-qPCR: Tfg1 sumoylation definitively impacts overall promoter sumoylation

How much of the SUMO signal detected at promoters is due to Tfg1 sumoylation? Or alternatively, is Tfg1 sumoylation prominent enough to impact the overall SUMO signal observed at target promoters when altered with mutant strains? To address these questions, five independent trials of ChIP-qPCR using an anti-SUMO antibody were conducted and averaged. As before, three strains were used: just tagged *Tfg1-HA*, the *Tfg1-K60,61R-HA* strain with reduced sumoylation, and the *SUMO-Tfg1-HA* strain with increased Tfg1-specific sumoylation. Five gene promoters were selected: the *TDH1* non-RPG with no SUMO peak as a control, the *TDH3*, *PYK1*, and *PDC1* non-RPGs with GTF associated SUMO peaks, and the *RPS20* RPG with a Rap1 associated SUMO peak. There was a clear correlation between Tfg1 sumoylation and overall promoter sumoylation at the three tested non-RPGs, further verifying that TFIIF is one of, if not the most, SUMO modified GTFs (**Fig 11**). The RPS20 RPG did not have a statistically significant change to overall promoter sumoylation levels, which was anticipated as its SUMO peak originated from Rap1, not GTFs. This result strongly suggested that the SUMO peaks at non-RPGs previously suspected to be derived from GTFs were in fact derived from GTF sumoylation.

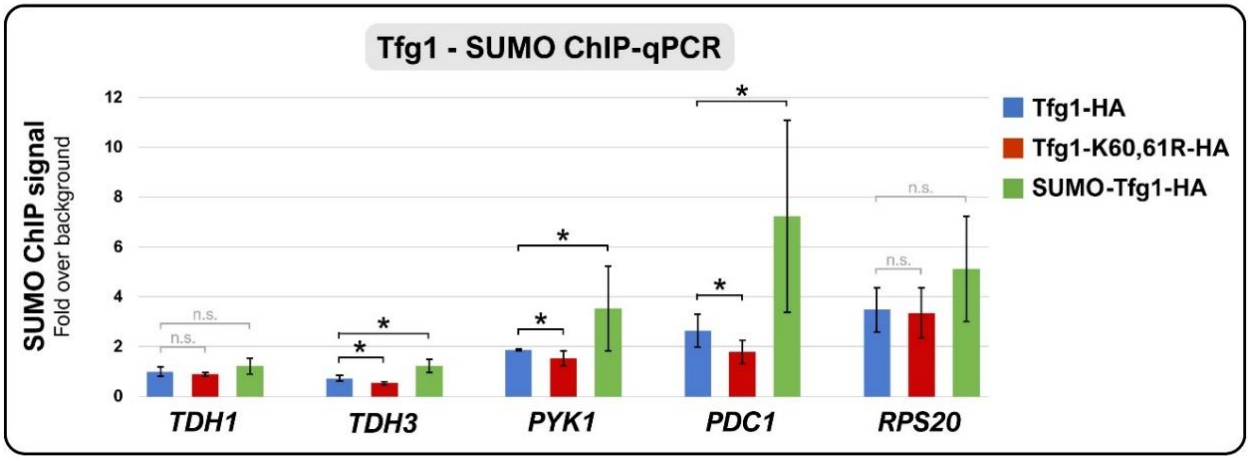

Figure 11. Tfg1 - SUMO ChIP-qPCR: Tfg1 significantly impacts overall promoter sumoylation at non-RPGs. There was a straightforward correlation between Tfg1 sumoylation and observed promoter-wide sumoylation, indicating that Tfg1 sumoylation alone is enough to significantly impact the overall SUMO signal. This significant change was only present at the tested non-RPGs and not at the tested RPG. Further, this result strongly supported the notion of the SUMO signal at a subset of non-RPGs being derived from GTFs, as the historic SUMO ChIP-seq suggested. For this ChIP, an anti-SUMO antibody was used to pull down DNA segments cross-linked to Tfg1 in the following strains: *Tfg1-HA*, *Tfg1-K60,61R-HA* with reduced sumoylation, and *SUMO-Tfg1-HA* with increased sumoylation. Five gene promoters were assessed for altered SUMO signal: the *TDH1* non-RPG with no SUMO peak as a background control, the *TDH3*, *PYK1*, and *PDC1* non-RPGs with GTF associated SUMO peaks, and the *RPS20* RPG with a Rap1 associated SUMO peak. The average of five trials was calculated, the standard deviation was represented with error bars, and the Student's *t*-test was used to assess significant changes (*p* value less than 0.05), indicated with an asterisk.

The key takeaway of this result was that while artificially diminished or elevated Tfg1 sumoylation did not impact Tfg1-chromatin interactions, it did successfully translate to altered chromatin bound Tfg1 SUMO conjugation, which could lead to changes in protein-protein interactions at the promoter. The most important of these interactions for transcriptional regulation was Tfg1's (and thereby TFIIF's) interactions with RNAPII. While standard ChIP cannot directly reveal differential RNAPII-TFIIF interactions at particular promoters, any resulting differences in RNAPII density between the three strains could be safely attributed to differential RNAPII-TFIIF interactions. At the least, it could be assumed that Tfg1 sumoylation coordinates interactions with RNAPII, even if the interaction is not direct.

3.5 Tfg1 – RNAPII ChIP-qPCR: altered Tfg1 sumoylation suppresses RNAPII recruitment

To assess whether RNAPII recruitment to promoters of SUMO peak-containing non-RPGs is affected by Tfg1 sumoylation levels, multiple independent trials of ChIP-qPCR using anti-RNAPII antibodies were performed, and the average was computed. Three strains were used: *Tfg1-HA*, *Tfg1-K60,61R-HA*, and *SUMO-Tfg1-HA*. First, the effect of reduced Tfg1 sumoylation was investigated using an anti-Rpb1 antibody. As this experiment was completed before any of the other ChIPs in this report, the full scope of the project had not yet been fully realized and a slightly different set of primers were used, although the principle remained the same. The *TDH3*, *PYK1*, and *PGK1* metabolic genes served as non-RPGs containing GTF associated SUMO peaks and *RPS27b* was the RPG with a Rap1 associated SUMO peak. While the *TDH1* gene promoter was not analyzed, all of the samples were still normalized to the background ChrV signal as usual. A clear trend was seen: a reduction in Tfg1 sumoylation correlated with a reduction in RNAPII recruitment at non-RPGs (**Fig 12A**). This was especially true for the tested non-RPG with the highest baseline activity, *TDH3*.

Later on in the project, the *SUMO-Tfg1-HA* strain was utilized to test the effects of constitutive Tfg1 SUMO conjugation on RNAPII recruitment, using an anti-Rpb3 antibody. While the results using the *SUMO-Tfg1-HA* strain were initially found to be not significant, a closer look at fold values within individual trials uncovered a clear trend within the overarching inconsistency. Ultimately, this was narrowed down to differences in experimental conditions affecting the growth rates, despite attempts to minimize variability. For instance, all of the samples had low levels of transcriptional activity and thereby RNAPII density in one trial while they all had high levels of density in another trial, resulting in large, overlapping error bars once averaged. Within each trial, however, the relative differences between samples maintained a clear trend, but this was not reflected in the calculated average of fold values. To address this, relative values were calculated as an alternative by dividing the *Tfg1-HA* and *SUMO-Tfg1-HA* fold values by *Tfg1-HA* fold values.

This resulted in a value of 1 for all *Tfg1-HA* samples and relative values for the *SUMO-Tfg1-HA* samples. Based on relative differences, constitutively elevated Tfg1 sumoylation was also found to negatively affect RNAPII recruitment at non-RPGs (**Fig 12B**). The original, non-relative chart can be found in the supplementary data (**Fig S3**).

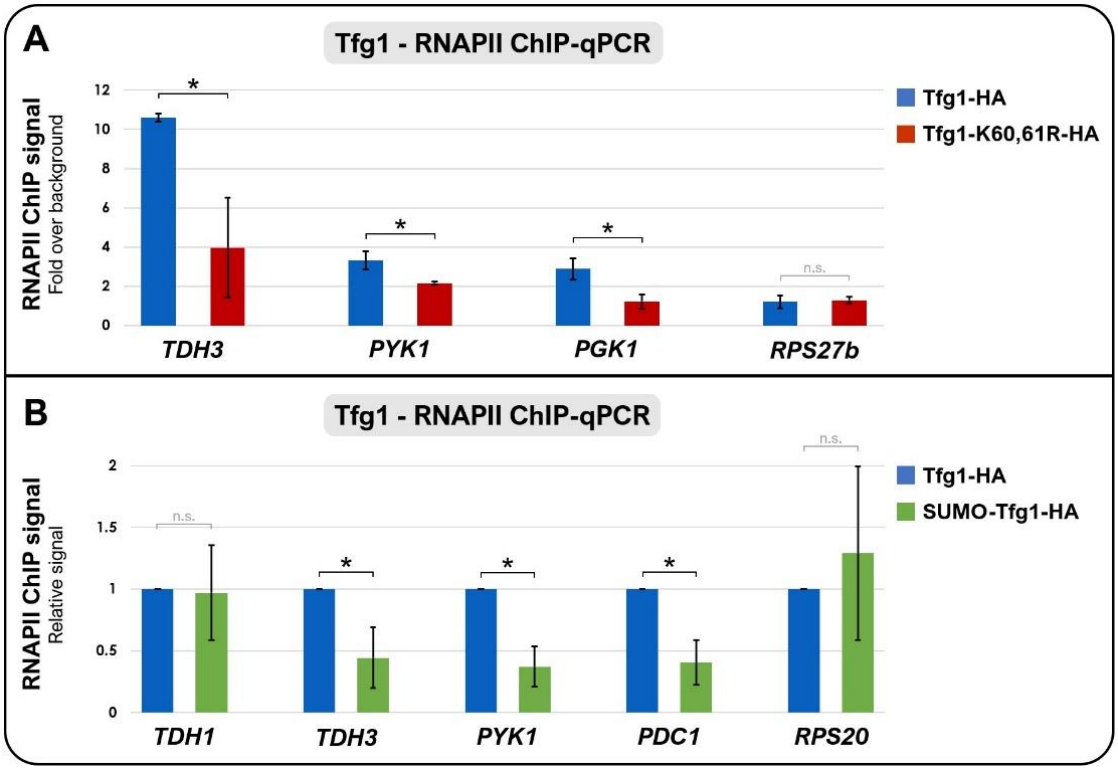

Figure 12. Tfg1 - RNAPII ChIP-qPCR: both reducing and increasing Tfg1 sumoylation impedes RNAPII recruitment. (A) Reducing promoter sumoylation levels via Tfg1 interrupts RNAPII-chromatin interactions at non-RPGs. For this ChIP, an anti-Rpb1 antibody was used to pull down DNA segments cross-linked to RNAPII in the *Tfg1-HA* and *Tfg1-K60,61R-HA* strains. Four gene promoters were assessed for differential RNAPII occupancy: the *TDH3*, *PYK1*, and *PGK1* non-RPGs with GTF associated SUMO peaks, and the *RPS27b* RPG with a Rap1 associated SUMO peak. The average of two trials was calculated, the standard deviation was represented with error bars, and the Student's *t*-test was used to assess significant changes (*p* value less than 0.05), indicated with an asterisk. **(B)** Constitutively increased Tfg1 and thereby promoter sumoylation levels also hindered RNAPII-chromatin interactions at non-RPGs. This was only observed in relative charts, as large error bars were present in the original charts (found in the supplementary data). For this ChIP, an anti-Rpb3 antibody was used to pull down DNA segments cross-linked to Tfg1 in the *Tfg1-HA* and *SUMO-Tfg1-HA* strains. Five gene promoters were assessed for differential RNAPII occupancy: the *TDH1* non-RPG with no SUMO peak as a background control, the *TDH3*, *PYK1*, and *PDC1* non-RPGs with GTF associated SUMO peaks, and the *RPS20* RPG with a Rap1 associated SUMO peak. The average of three trials was calculated, the standard deviation was represented with error bars, and the Student's *t*-test was used to assess significant changes (*p* value less than 0.05), indicated with an asterisk.

These findings revealed that both artificially reduced and (constitutively) elevated Tfg1 sumoylation resulted in reduced RNAPII recruitment at non-RPGs. While initially puzzling, we

reasoned that any disturbance to the normal level of SUMO modifications diminished optimal transcriptional activity. SUMO modifications are reversible and dynamic. Modified sumoylation levels in either direction might thereby disrupt the sophisticated balance and transient control associated with the numerous deeply interconnected regulatory mechanisms of transcription (as discussed below in section 4.2). What remained unclear was how representative these observed trends with select target genes were of changes within the pool of 111 unique non-RPGs, and of global changes across the genome. For global analysis, only the *Tfg1-HA* and the *Tfg1-K60,61R-HA* yeast strains were used, and the *SUMO-Tfg1-HA* strain was not used.

3.6 RNAPII ChIP-seq with *Tfg1-K60,61R-HA* strain: globally reduced recruitment

While clear reductions were observed in RNAPII density at a handful of non-RPG promoters, we aimed to measure the global changes to RNAPII density with reduced Tfg1 sumoylation. An RNAPII ChIP-seq was performed using two replicates of large-scale ChIPs using an anti-Rpb1 antibody for pulldown. Two strains were used: just tagged *Tfg1-HA* and the *Tfg1-K60,61R-HA* strain with reduced sumoylation. The *SUMO-Tfg1-HA* strain was not utilized. Purified samples were sent to an external lab for Next Generation Sequencing. After library synthesis, 10 million reads were taken per sample, and a differential binding assay (DiffBind) was performed to assess differential RNAPII density across gene ORFs. For normalization, the False Discovery Rate was calculated, which is generally heralded as a superior method of uncovering biologically appropriate significant changes, rather than the simple Student's *t*-test [45]. We discovered that Tfg1 sumoylation significantly impacted RNAPII density at thousands of genes. A reduction in Tfg1 sumoylation resulted in a genome-wide decrease in RNAPII density (**Fig 13A**). This change significantly impacted some non-RPGs containing GTF-derived SUMO peaks (**Fig 13B**), but unexpectedly also impacted many other non-RPGs without promoter-associated sumoylation. This demonstrated that SUMO likely controls transcription at more non-RPGs than the subset detected by the landmark SUMO ChIP-seq (as described in section 1.6).

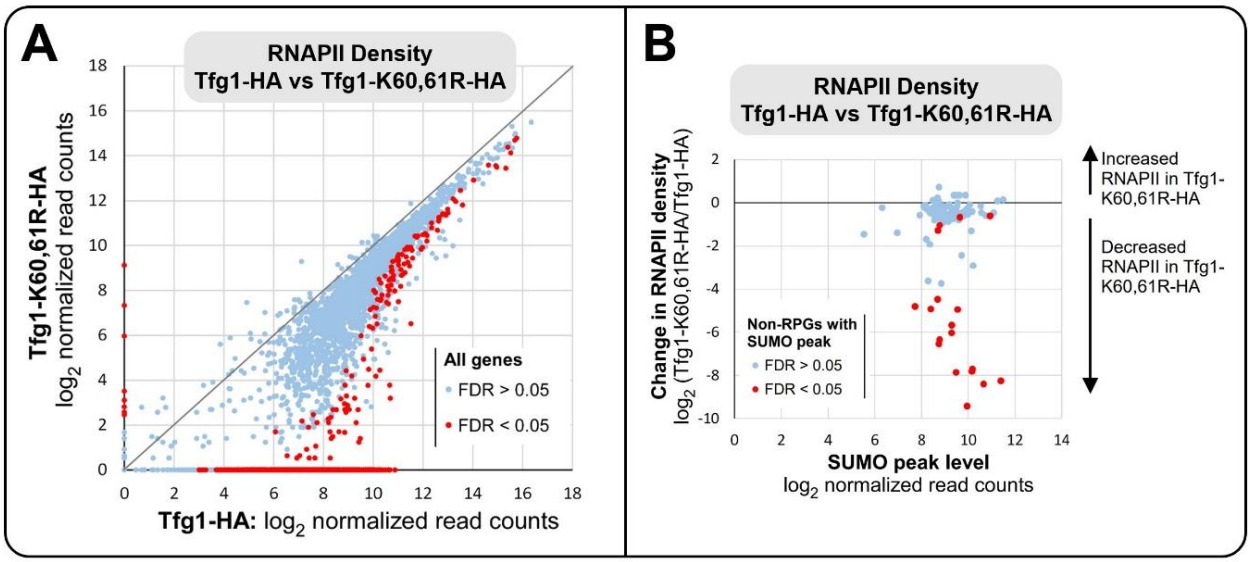

Figure 13. Tfg1 - RNAPII ChIP-seq: Tfg1 sumoylation impacts global RNAPII recruitment. Two independent replicates were utilized. A differential binding assay, DiffBind, was used to assess significant changes. Red dots represented significantly changed densities (FDR less than 0.05). **(A)** Log_2 normalized read counts of RNAPII density in the *Tfg1-HA* versus *Tfg1-K60,61R-HA* strains were measured at each gene ORF and compared. Almost half of all genes (47%) saw a significant decreases in the *Tfg1-K60,61R-HA* strain, indicating that Tfg1 sumoylation promotes RNAPII density genome wide. **(B)** Changes in RNAPII densities at the subset of non-RPGs associated with unique SUMO peaks. Vertical axis displays log_2 normalized ratio of *Tfg1-K60,61R-HA/ Tfg1-HA* signal, while horizontal axis shows SUMO peak intensity. Only 18 of the non-RPGs showed significantly reduced RNAPII density, and there was not a correlation between SUMO peak intensity and change in RNAPII density between strains (images from [15]).

Taking a closer look at the ChIP-seq results revealed that out of the 5990 genes assayed, 2828 had significantly reduced RNAPII density (FDR less than 0.05), which equated to about 47% of genes. At the majority of the affected genes, there was no RNAPII detected at the promoters in the mutant strain, implying a halting of transcriptional activity there. In contrast, only 9 genes had significantly higher RNAPII density in the *Tfg1-K60,61R-HA* strain, which could be attributed to chance or natural fluctuations. As implied by the small-scale ChIPs, there generally were not significant changes at RPGs, which are known to have Rap1-derived SUMO peaks. However, only 18 of the 111 unique non-RPGs with GTF-derived SUMO peaks had significantly reduced RNAPII density. This equated to a reduction at only about 16% of target non-RPGs. Notably, the three previously tested non-RPGs (*TDH3*, *PYK1*, and *PGK1*) had reduced RNAPII density while the previously tested RPG (*RPS27b*) did not show any significant change, validating those observed trends (from section 3.5, **Fig 12A**). There was not any obvious correlation between the

concentration of RNAPII (i.e. gene activity) at a particular gene and the severity of the change due to a reduction in Tfg1 sumoylation. Genes with low, medium, and high levels of activity were all affected. A GO analysis was performed on all significantly affected genes, but no obvious link emerged between gene identity and the impact of the mutation, as nearly half of all genes were affected. Even the 9 genes with higher RNAPII density in the *Tfg1-K60,61R-HA* strain did not belong to any particular gene group. This implied that GTF sumoylation impacts the general association of RNAPII with numerous promoters genome wide.

At almost all significantly affected genes, the *Tfg1-HA* samples had higher $\log_2$ normalized read counts of RNAPII when compared to the *Tfg1-K60,61R-HA* samples. While this result was not anticipated, it meant that this SUMO modification was even more important for transcriptional regulation than initially hypothesized, in a manner that remains unknown. In effect, Tfg1 sumoylation appears to be required for elevated, but not necessarily basal, levels of transcription in log phase yeast cells, as the mutant cells did not demonstrate a growth defect. In order to examine how this impacted subsequent transcription rates and thereby steady-state mRNA levels following the polymerase's release from the PIC, we used RT-qPCR but were not able to detect significant changes (**Fig S4**). Instead, we decided to examine global changes to RNA levels in the *Tfg1-K60,61R-HA* strain.

3.7 RNA-seq with *Tfg1-K60,61R-HA* strain: preserved steady-state expression levels

Changes to the overall transcriptome due to a reduction in Tfg1 sumoylation were measured by performing an RNA-seq. For this assay, RNA was extracted from the *Tfg1-HA* and *Tfg1-K60,61R-HA* strains and again sent to an external lab for NGS. For this experiment, the polyadenylated mRNAs were enriched from total RNA, sequencing libraries were prepared, and then 30 million reads were taken per sample. Despite seeing vast changes to RNAPII density across the genome, reducing Tfg1 sumoylation did not have a very significant effect on steady-state mRNA levels (**Fig 14**). While many genes showed changes, they were mostly small changes. Only 16 genes had

significantly changed expression. When Tfg1 sumoylation was reduced, 3 genes had reduced expression while 13 genes had increased expression. Given that mRNAs from 6120 genes were sequenced, these significant changes were attributed to chance or natural fluctuations.

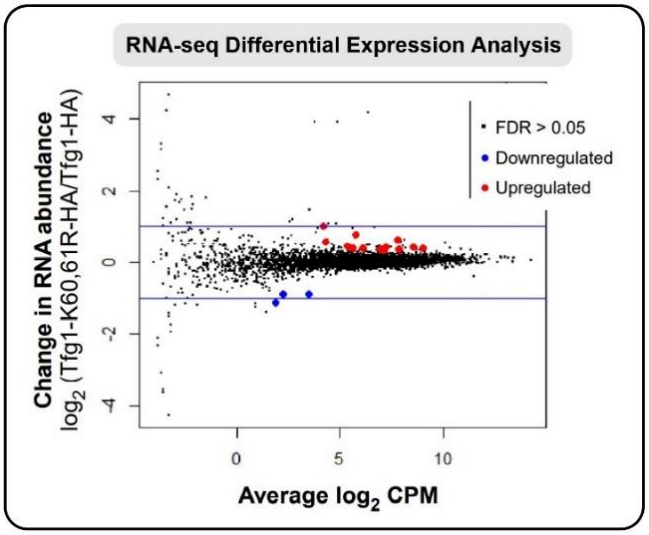

Figure 14. Tfg1 - RNA-seq: steady-state mRNA levels remain stable despite reduction in overall transcription. After PolyA enrichment, 30 million reads were taken per sample, and mRNAs from 6120 genes were sequenced for changes. A differential expression analysis using edgeR was plotted as a $\log_2$ ratio of *Tfg1-K60,61R-HA* over *Tfg1-HA* along the vertical axis. The horizontal axis showed the difference in mRNA abundance per gene, as the average $\log_2$ counts per million (CPM). Insignificant genes with FDR values above 0.05 were represented with black dots. Only 16 genes had significantly changed mRNA levels between strains (FDR less than 0.05). In the *Tfg1-K60,61R-HA* strain, 3 genes had decreased expression while 13 genes had increased expression, compared to the *Tfg1-HA* strain (image from [15]).

It remains unknown why virtually no significant changes were observed at the transcriptome level despite the vast impact on RNAPII density and thereby transcription. Two main speculative explanations were considered. First of all, it was possible that the reduction in transcription from reduced Tfg1 sumoylation was not sufficient enough to substantially impact steady-state mRNA levels. Biological systems tend to have vast buffering capacities and interconnected feedback loops to limit changes, and it was possible that a similar mechanism was at play here. The cells could have sensed a reduction in nascent mRNA production and in response employed systems to stabilize existing transcripts or could have deliberately reduced the expression/activity of mRNA degrading factors. A second possibility was that the vast reduction in transcription genome-wide

also negatively impacted the expression of mRNA degrading factors, leading to an inadvertent stabilization of the remaining transcripts. So even if the expression levels were overall low, the lack of turnover would have cancelled out any change, resulting in a similar level of steady-state mRNA. While the reason behind this observed phenomenon could not be ascertained, we next wished to establish if any functional change could be observed due to changes in any intermediate cellular pathways.

3.8 Functional analysis of reduced Tfg1 sumoylation

To conclude the analysis on Tfg1 sumoylation, the functional impact due to any alterations in cellular interactions or the regulation of gene expression were studied. For this assay, two exemplary SUMO peak containing non-RPGs with reportedly reduced RNAPII levels according to our ChIP-seq were focused on, as they both belonged to a similar pathway: the biosynthesis of the amino acid lysine. Although the *LYS1* and *LYS2* genes had substantially reduced RNAPII density, like most genes, their transcript levels were unaffected by the K60,61R mutation of Tfg1. This implied that their expression, and therefore lysine biosynthesis, is normal in the mutant strain. To confirm this, lysine auxotrophy was assessed via a drop test. Five strains were spotted onto a control SC medium plate and a plate of SC medium lacking lysine: the *W303a* wild-type parent, the *ubc9-6* strain with reduced global sumoylation [15], just tagged *Tfg1-HA*, the *Tfg1-K60,61R-HA* strain with reduced Tfg1 sumoylation, and a positive control *lys1*Δ strain which was unable to synthesize lysine due a deleted *LYS1* gene. As miniscule changes to mRNA-level changes were observed within the RNA-seq, we did not expect any notable growth defect in this assay. As anticipated, the *Tfg1-K60,61R-HA* yeast strain was not impacted by lysine deficiency and grew normally on media lacking lysine at 30°C for three days (**Fig 15**). The *ubc9-6* strain with globally reduced sumoylation was not particularly sensitive to the lysine deficiency either. This reinforced the findings of the RNA-seq: reducing Tfg1's sumoylation does not greatly impact the mRNA levels of RNAPII-mediated transcripts, as the cells retained their ability to express lysine.

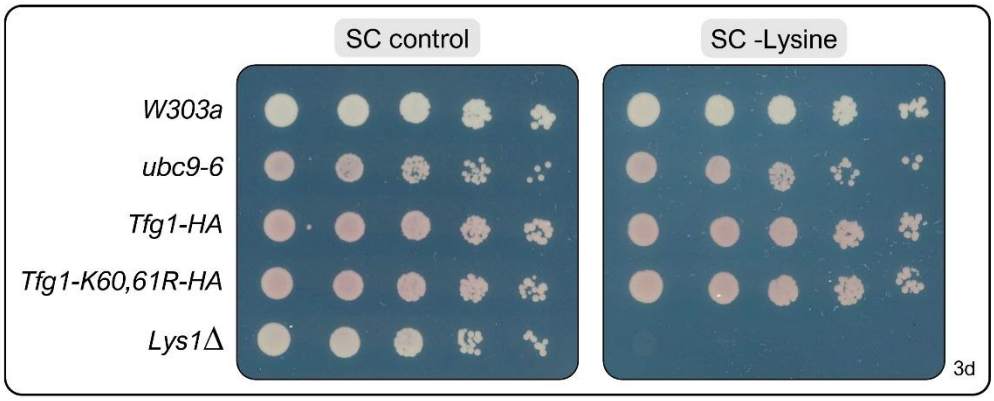

Figure 15. Reduced Tfg1 sumoylation does not result in lysine auxotrophy. A spot assay was used to assess whether the reduced transcription of the *LYS1* gene in the *Tfg1-K60,61R-HA* strain impacted the strain's ability to grow on lysine deficient medium. Neither the *Tfg1-K60,61R-HA* strain nor the *ub9-6* strain with globally reduced sumoylation resulted in lysine auxotrophy. Five strains were grown on a control SC media plate and a SC -Lysine (minus lysine) plate: the *W303a* wild-type parent, the *ubc9-6* strain with reduced global sumoylation, just tagged *Tfg1-HA*, the *Tfg1-K60,61R-HA* strain with reduced Tfg1 sumoylation, and a positive control *lys1*Δ strain with a deleted *LYS1* gene.

3.9 TBP is a SUMO Target in *S. cerevisiae*

While previous data from our lab has suggested that TFIIF is the primarily sumoylated GTF, we suspected that the TBP subunit of TFIID might also be an important target of SUMO modifications in *S. cerevisiae* (as described in section 1.8). There were four primary reasons for this prospect: (1) inconclusive past results, (2) hints from large scale proteomics studies, (3) the overlap of 147 SUMO peaks at non-RPGs with the TATA-box to which TBP binds directly, and (4) published *in vitro* results utilizing human cells. First, former lab members were unable to detect TBP sumoylation when IPing the TAF7 and TAF8 subunits of TFIID, but those findings were questioned due to potentially harsher-than-anticipated IP conditions which might have dislodged TBP from the complex. Otherwise, it remained possible that TBP was only sumoylated when not complexed with the remainder of TFIID, as that would also mask its sumoylation in such a co-IP. Second, large scale proteomics studies had identified the presence of SUMO attachment consensus motifs on TBP, which implied that SUMO binds to TBP in budding yeast [39]. Third, our landmark ChIP-seq identifying SUMO conjugation across the *S. cerevisiae* genome discovered that there are 147 distinct SUMO peaks at non-RPGs [15]. A closer look revealed that these peaks aligned virtually exactly with the TATA-box element. However, TBP serves as

scaffolding for much of the PIC and thereby multiple GTF subunits could have been responsible for the observed SUMO peaks. Fourth, a published *in vitro* assay using human-derived recombinant TBP found that free-form TBP could become sumoylated when incubated with SUMO proteins, along with E1 and E2 enzymes, however when complexed with TAFs, its SUMO conjugation motif was hidden and sumoylation was no longer observed [40]. The authors of the paper, and any subsequent literary research to my knowledge, were unable to experimentally confirm whether TBP is sumoylated *in vivo*, in any species (despite attempts [40,46]).

To determine whether sumoylated TBP can be detected *in vivo*, we utilized homologous recombination to insert a 6xHA epitope tag at the C-terminus of the *SPT15* gene, which codes for TBP in *S. cerevisiae*. Once sequencing confirmed successful attachment of the tag to endogenous TBP, creating the *TBP-HA* strain, an anti-HA antibody was used to immunoprecipitate TBP and anti-HA and anti-SUMO antibodies were used to immunodetect for the presence of SUMO conjugation. Based on the resulting western blots, we determined that TBP is sumoylated in *S. cerevisiae*, and based on the presence of multiple equally spaced SUMO bands, TBP is multi- and/or polysumoylated (**Fig 16A**). This was evidenced by the presence of bands at the appropriate additive weight (in kDa) of TBP + the 6xHA epitope tag + the SUMO proteins. To confirm that the observed SUMO bands derived from TBP rather than an interacting protein, an additional IP was performed under denaturing conditions using TCA (as described in section 2.9) to dislodge any non-covalent interactions. Under denaturing conditions, TBP was still found to be sumoylated (**Fig 16B**). Although only a single SUMO band was seen this time, this was attributed to the lower yield and potential desumoylation stemming from the harsher IP conditions. TBP is likely multi- and/or polysumoylated based on the equally spaced SUMO bands seen earlier.

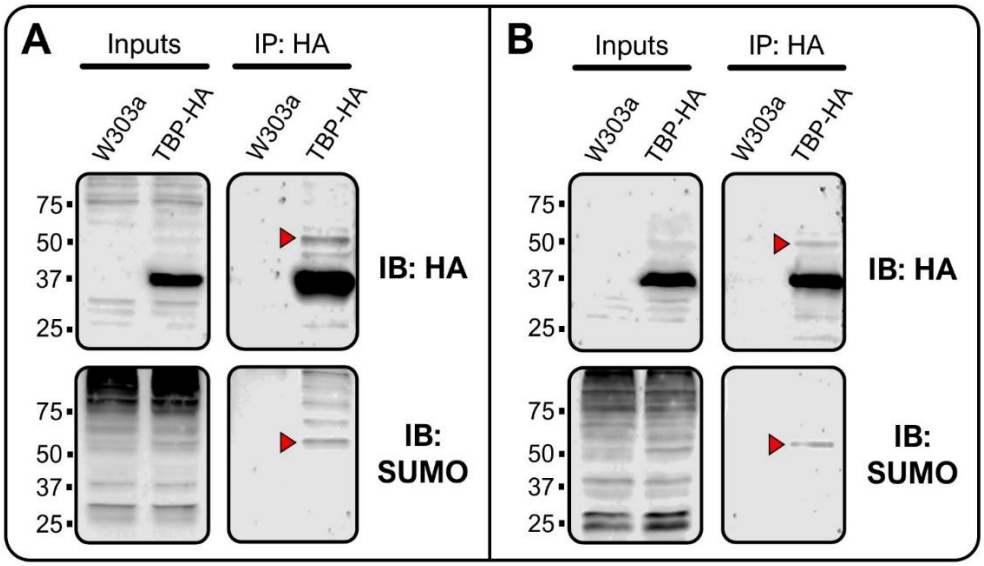

Figure 16. TBP is sumoylated in *S. cerevisiae*. In these blots, TBP runs at ~37 kDa and monosumoylated TBP (red arrowhead) runs at ~55 kDa. **(A)** The HA Tag was IPed in standard non-denaturing conditions in the wild-type *W303a* and *TBP-HA* strains, and the purified proteins were immunodetected on HA and SUMO immunoblots. Based on the multiple equally spaced SUMO bands, it was determined that TBP is multi- and/or polysumoylated. **(B)** An HA Tag IP under denaturing conditions confirmed that TBP, rather than an interacting protein, was sumoylated.

3.10 Generation of TBP SUMO-deficient mutant strains

Very much like with the Tfg1 part of this study, the functional impacts of TBP sumoylation were studied using a strain with reduced levels of TBP-specific sumoylation. In order to generate this strain, I used two separate prediction algorithms which identified potential SUMO conjugation sites on TBP based upon the presence of the SUMO attachment consensus motif. The GPS-SUMO tool identified lysine 127 as a potential SUMO target, and the Joined Advanced Sumoylation Site and Sim Analyzer (JASSA) tool identified both lysine 127 (again) and an inverted consensus motif surrounding lysine 167. Additionally, published mass spectrometry data from 2017 indicated that lysine 47 was a likely target of sumoylation [47]. Collectively, it was decided that all three residues should be explored for *in vivo* TBP SUMO conjugation: Lys47, Lys127, and Lys167. Homologous recombination-based site-directed mutagenesis was used to introduce point mutations. In each case, lysines were changed to arginines, resulting in three new mutant strains: *TBP-K47R-HA*, *TBP-K127R-HA*, and *TBP-K167R-HA*. However, only minor reductions in sumoylation were observed in each case when IP-WBs were done. This implied that all three sites

are SUMO targets, and that TBP is multisumoylated. To address this, a dual mutant encompassing the Lys127 and Lys167 mutations, *TBP-K127,167R-HA*, and a triple mutant encompassing all three mutations, *TBP-3KR-HA*, were made. These new mutants were immunoprecipitated and immunodetected along with the *W303a* parent, *TBP-HA*, and *TBP-K47R-HA* strains for comparison. The triple mutant *TBP-3KR-HA* strain was found to have the most SUMO-deficient TBP (**Fig 17**).

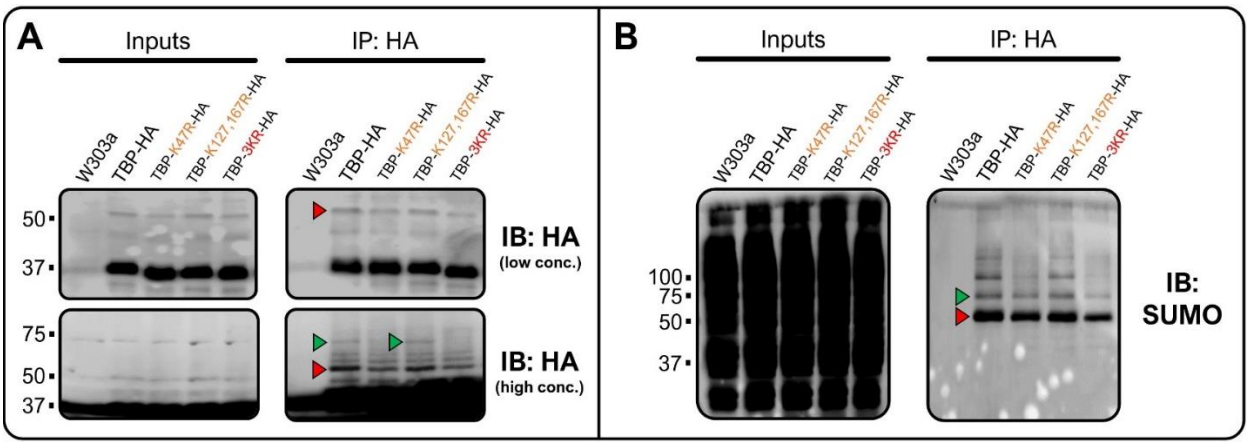

Figure 17. Point mutations within the *TBP-3KR-HA* strain reduce TBP SUMO conjugation. In these blots, TBP runs at ~37 kDa, monosumoylated TBP (red arrowhead) runs at ~55 kDa, and doubly-sumoylated TBP (green arrowhead) runs at ~70 kDa. Proteins from five strains were IPed for the HA Tag and were then HA and SUMO immunodetected: the *W303a* parent, just tagged *TBP-HA*, a single mutant *TBP-K47R-HA*, a double mutant *TBP-K127,167R-HA*, and a triple mutant *TBP-3KR-HA* strain. **(A)** Probing for the HA tag revealed a successful IP. Using a higher anti-HA antibody concentration, additional multi- or polysumoylation bands were made visible, but the nonsumoylated TBP bands (~37 kDa) were obscured by a large blotch (black area along bottom). The doubly sumoylated TBP band (green arrowhead) at ~70 kDa was not seen in high concentration HA immunoblots with the *TBP-K47R-HA* and *TBP-3KR-HA* strains, indicating particularly reduced sumoylation within those two strains. **(B)** Probing for SUMO showed variable sumoylation levels for each strain. Both *TBP-K47R-HA* and *TBP-3KR-HA* strains resulted in significantly reduced sumoylation levels with less, and of lower intensity, SUMO bands. The K127,167R double mutant did not reduce sumoylation as much, only slightly reducing band intensity.

While sumoylation was not abolished even in the triple mutant (like with the *Tfg1-K60,61R-HA* mutant strain), it was deemed low enough to likely have a functional impact. The Lys47 mutation was found to be the most effective at reducing SUMO conjugation and had the same SUMO band count as the triple mutant, but the triple mutant had slightly lower band intensity overall. While additional mutations could have been made to continually reduce TBP sumoylation, that would introduce the risk of venturing too far away from natural, endogenous TBP and therefore

additional mutations were not pursued. Given that SUMO modifications often interplay or compete with other PTMs, we wanted to minimize the impact on TBP's cellular interaction. Even the three chosen lysines were already at or near reported PTMs. Lys42 is a reported phosphorylation site [48], and it is possible that sumoylation at Lys47 obstructs this nearby phosphorylation location. TBP has four reported ubiquitination sites: K127, K133, K151, and K167, and two of these were chosen to be mutated due to SUMO conjugation potential [49] (data from: https://www.yeastgenome.org/locus/S000000950/protein). Given this, we cannot be completely confident that the observed downstream effects of the mutations were due to altered sumoylation or an alteration in the activity of another PTM. However, SUMO's biological role might be to regulate these exact PTMs, and thereby (indirectly) affecting other PTMs could be advantageous in terms of studying SUMO's role, in a roundabout way. For this analysis, no 'increased TBP sumoylation' counterpart strain was generated along the lines of our *SUMO-Tfg1-HA* strain. This was due to time/resource concerns and due to the debatable recapitulation of *in vivo* sumoylation provided by such a fusion system. The functional impact of only reduced TBP sumoylation was assessed. But before any functional assays could be performed, the growth and stress response capabilities of the new strains had to be established.

3.11 TBP SUMO-deficient mutants grow normally

In order to evaluate whether the SUMO-deficiency of the TBP mutants affected the overall rate of cell growth and survival, growth rates were measured using an accuSkan FC automated spectrophotometer. The machine took regular readings of the liquid cell culture density via optical density over a 20-hour time frame. There was only a minor growth delay observed for any strain containing a 6xHA epitope tag, and the presence of point mutations did not impact growth in any meaningful way (**Fig 18A**). There was no notable growth delay, crash, or other abnormality due to a reduction in TBP sumoylation when different concentrations of strains were spotted onto solid

media plates containing osmotic (1 M KCl), oxidative (1 mM H_2O_2), or alcoholic (7% EtOH) stress compounds, or when SC media plates were grown in heat stress conditions (37°C) (**Fig 18B**).

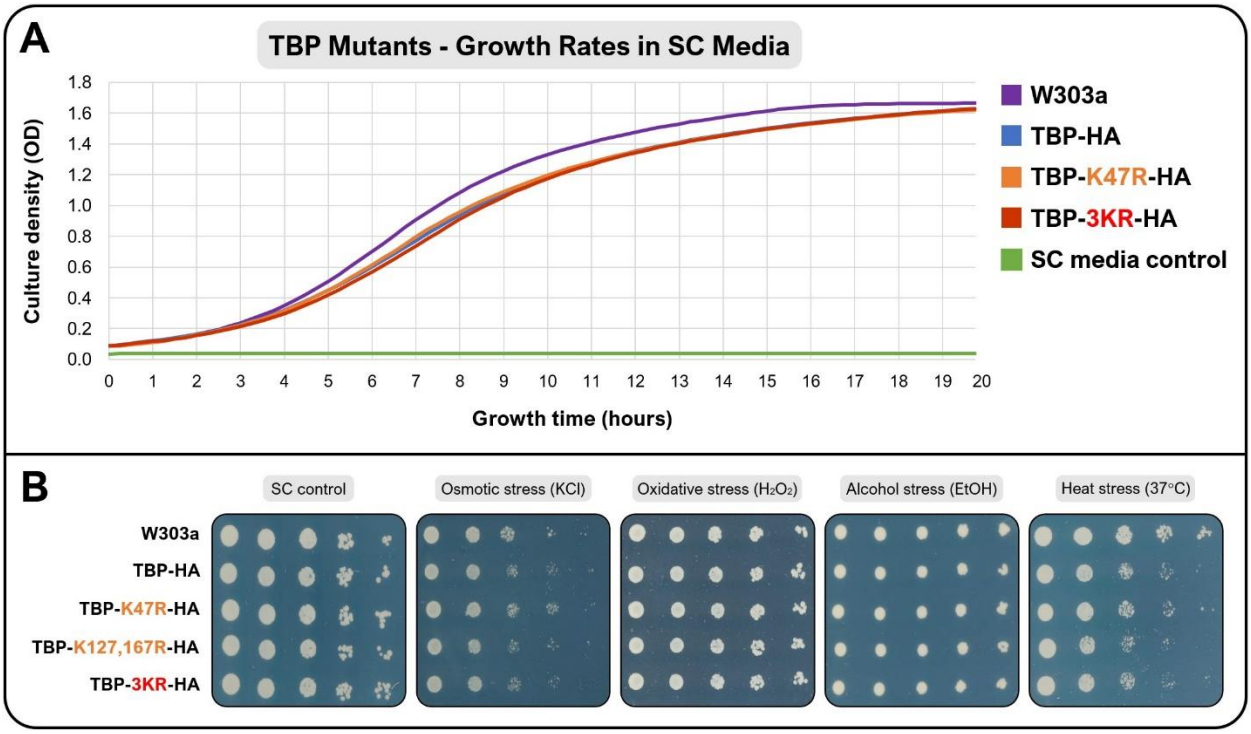

Figure 18. The TBP mutant strains do not suffer from any major growth defects. (A) Growth rates of TBP mutants as measured by an accuSkan FC automated spectrophotometer. Monitored the growth rates of the following strains in typical SC media: wild-type *W303a*, just tagged *TBP-HA*, and the SUMO-deficient mutants *TBP-K47R-HA* and *TBP-3KR-HA*. The *W303a* strain grew the quickest, with the other three strains trailing it closely. Triplicates of each strain were incubated at 30°C while changes in culture density were tracked over time. Triplicate results were then averaged and graphed. Wells containing only SC media and no culture were used as a negative control (green SC media control line along bottom). **(B)** Spot assays of TBP mutants following three days of 30°C incubation (unless specified). Each row of colony groups represented a different strain while each column was a different concentration of cells, decreasing five-fold from left to right. All strains grew without fault in the SC control, osmotic stress (1 M KCl), oxidative stress (1 mM H_2O_2), and alcohol stress (7% EtOH) conditions. With the heat stress (37°C) condition, the *W303a* strain grew fine while the *TBP-HA*, *TBP-K47R-HA*, *TBP-K127,167R-HA*, and *TBP-3KR-HA* strains had a slight growth defect and lagged behind their wild-type counterpart.

With the growth stability and stress response capability of the strains established as being unaffected, the changes to transcriptional regulation due to a reduction in TBP sumoylation could begin. As with Tfg1, we aimed to do ChIPs with anti-HA, anti-SUMO, and anti-RNAPII antibodies. This way any changes to TBP-chromatin association, overall promoter sumoylation,

and RNAPII density at particular promoters based upon TBP's sumoylation level could be established.

3.12 TBP – HA ChIP-qPCR: Unaffected chromatin interactions

As described above (in section 1.2), transcription factor sumoylation can reportedly affect TF-chromatin interactions [13]. Three fresh trials of ChIP-qPCR using an anti-HA antibody were utilized to investigate how variable TBP sumoylation affected TBP-chromatin association. For comparison, two strains were used: *TBP-HA* and *TBP-3KR-HA*. Five gene promoters were examined: the *TDH1* non-RPG without a SUMO peak as a control, the *TDH3*, *PYK1*, and *PDC1* non-RPGs with GTF associated SUMO peaks, and the *RPS20* RPG with a Rap1 associated SUMO peak. Like with the Tfg1 strains, there was inconsistency between trials, but when averaged there were no significant changes between strains, meaning that a reduction in TBP sumoylation did not particularly affect TBP's chromatin interactions (**Fig 19**). How an increase in TBP sumoylation beyond normal levels would fare remains unknown.

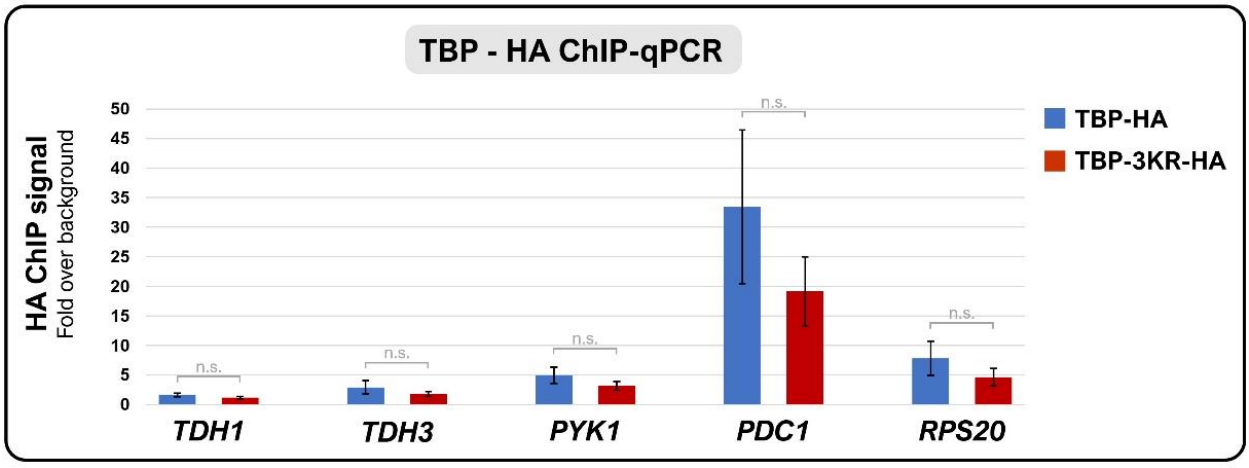

Figure 19. TBP - HA ChIP-qPCR: TBP sumoylation does not impact TBP's chromatin occupancy either. For this ChIP, an anti-HA antibody was used to pull down DNA segments cross-linked to TBP in the *TBP-HA* and *TBP-3KR-HA* strains. Five gene promoters were assessed for variable TBP binding to chromatin: the *TDH1* non-RPG with no SUMO peak as a background control, the *TDH3*, *PYK1*, and *PDC1* non-RPGs with GTF associated SUMO peaks, and the *RPS20* RPG with a Rap1 associated SUMO peak. The average of three trials was calculated, the standard deviation was represented with error bars, and the Student's *t*-test was used to assess significant changes (*p* value less than 0.05).

3.13 TBP – SUMO ChIP-qPCR: Maintained overall promoter sumoylation

The effect of TBP sumoylation on overall promoter sumoylation was determined by performing four trials of ChIP-qPCR using an anti-SUMO antibody. SUMO signals from the *TBP-HA* and *TBP-3KR-HA* strains were compared at the usual gene promoters: *TDH1*, *TDH3*, *PYK1*, *PDC1* and *RPS20*. In contrast to the strong Tfg1 results, however, no significant change was observed. TBP sumoylation had little impact on the overall promoter sumoylation at the tested genes (**Fig 20**).

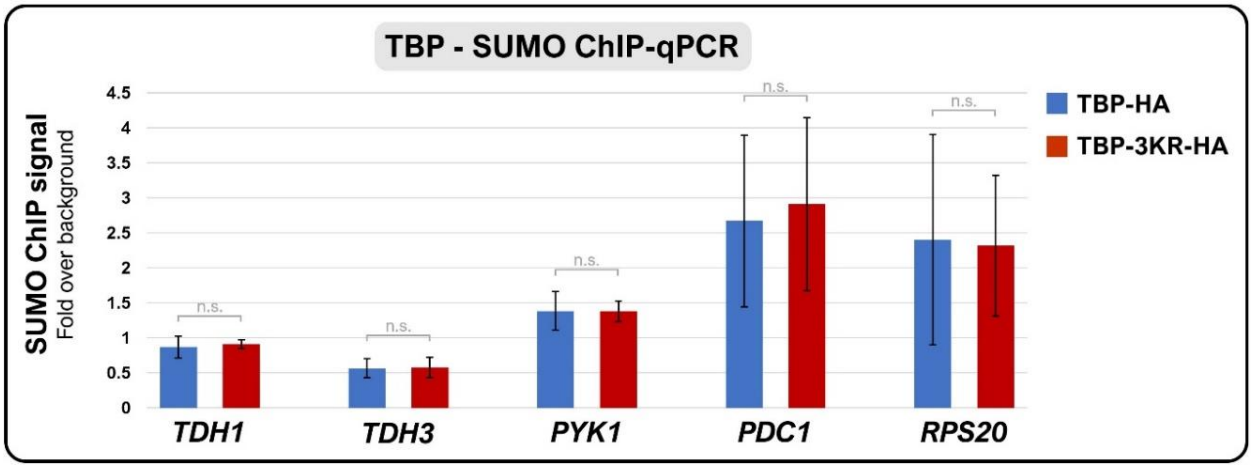

Figure 20. TBP - SUMO ChIP-qPCR: TBP sumoylation does not impact overall promoter sumoylation. In this ChIP, an anti-SUMO antibody was used to pull down DNA segments cross-linked to TBP in the *TBP-HA* and *TBP-3KR-HA* strains. Five gene promoters were assessed for altered SUMO signal: the *TDH1* non-RPG with no SUMO peak as a background control, the *TDH3*, *PYK1*, and *PDC1* non-RPGs with GTF associated SUMO peaks, and the *RPS20* RPG with a Rap1 associated SUMO peak. The average of four trials was calculated, the standard deviation was represented with error bars, and the Student's *t*-test was used to assess significant changes (*p* value less than 0.05).

There were three main explanations for these findings. First, this result did not necessarily mean that TBP sumoylation was unaffected at these promoters. As indicated by various preliminary indications and as shown above, the Tfg1 subunit of TFIIF is likely the primarily sumoylated GTF subunit, dominating the overall SUMO signature at promoters. Further, there are multiple other GTF subunits which are confirmed SUMO targets, masking changes in individual subunits [15]. While chromatin bound TBP may have been sumoylated, it is possible that its comparative level of sumoylation was low (but still efficacious, as will be discussed in section 3.14). The second explanation was that TBP is primarily or potentially exclusively sumoylated when soluble, and not

when chromatin bound. This is what was suggested by the previously discussed *in vitro* assay [40]. This would simultaneously explain a reduction in overall TBP sumoylation (as was seen in IP-WBs) and the absence of a reduction in chromatin bound TBP. To address this, a chromatin fractionation was performed to compare soluble vs chromatin bound TBP (section 3.15 below).

A third speculative possibility was that group sumoylation (as discussed in section 1.2) was compensating for a reduction in TBP sumoylation within the mutant strain. In short, group sumoylation entails that not all SUMO targets/epitopes within a vicinity or complex need to be sumoylated to elicit the complete functional output, introducing redundancy, and reducing the importance of individual SUMO binding sites [2,20]. Perhaps TBP binds to a certain highly localized area of the PIC where group sumoylation controls and instigates cooperative SUMO binding. In that case, the reduction in TBP sumoylation could have been masked by a compensatory increase in the sumoylation of a nearby substrate. TFIIF would likely not be included in this cluster/group, as changes in its sumoylation level while chromatin bound were observed. Otherwise, TFIIF sumoylation could be prevalent enough to overcome the buffering of group sumoylation, unlike TBP sumoylation.

3.14 TBP – RNAPII ChIP-qPCR: A modest reduction at tested non-RPGs

While no change in chromatin association or overall promoter sumoylation was observed with the *TBP-3KR-HA* strain, it was still possible that missing SUMO modifications on TBP could have affected protein-protein interactions, especially at the relevant subset of non-RPG promoters. To assess differential RNAPII recruitment due to reduced TBP sumoylation, four trials of ChIP-qPCR using an anti-Rpb3 antibody were done using the *TBP-HA* and *TBP-3KR-HA* strains. Changes in RNAPII density were measured at the *TDH1*, *TDH3*, *PYK1*, *PDC1* and *RPS20* gene promoters. A significant reduction in RNAPII occupancy at two of the three non-RPGs was seen, but not at the relatively inactive *PYK1* non-RPG (**Fig 21**). Given its relatively low transcriptional activity by default, based upon the signal within the *TBP-HA* strain, the relative drop in RNAPII density at

the *PYK1* gene was computed, and was found to be significant in that case (**Fig S5**). Given this caveat, we could conclude that a reduction in TBP sumoylation also represses RNAPII recruitment at the tested non-RPGs with GTF derived SUMO peaks like Tfg1, but perhaps to a slightly lesser degree. However, given the unexpected findings of the RNAPII ChIP-seq with the Tfg1 strain, it cannot be concluded whether this trend would inductively translate to the remaining subset of non-RPGs and the entire genome. Further, how this mutation affects steady-state mRNA levels remains subject to further exploration.

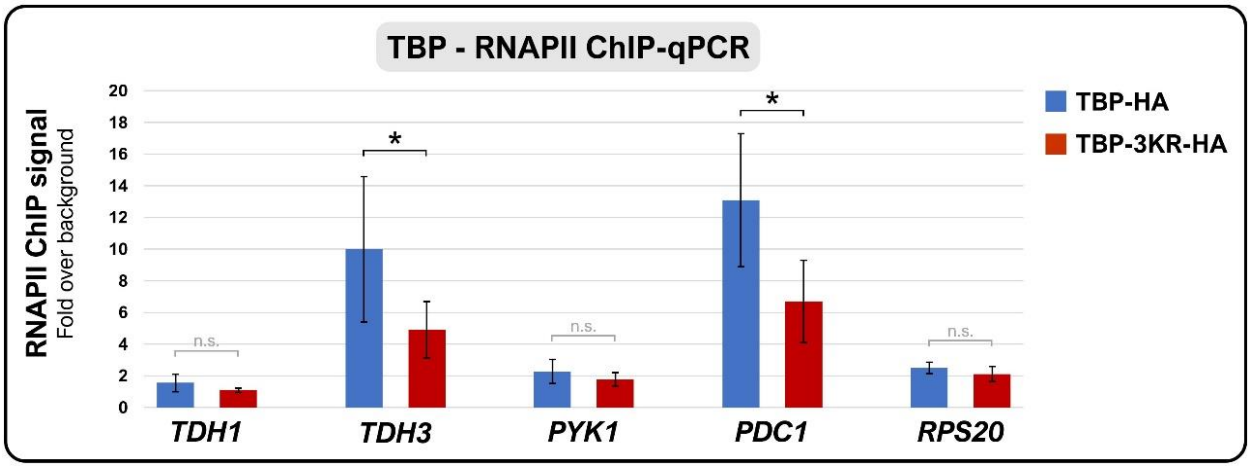

Figure 21. TBP - RNAPII ChIP-qPCR: Reduced TBP sumoylation represses RNAPII recruitment at non-RPGs. An anti-Rpb3 antibody was used in this ChIP to pull down DNA segments cross-linked to RNAPII in the *TBP-HA* and *TBP-3KR-HA* strains. Five gene promoters were assessed for differential RNAPII occupancy: the *TDH1* non-RPG with no SUMO peak as a background control, the *TDH3*, *PYK1*, and *PDC1* non-RPGs with GTF associated SUMO peaks, and the *RPS20* RPG with a Rap1 associated SUMO peak. The average of four trials was calculated, the standard deviation was represented with error bars, and the Student's *t*-test was used to assess significant changes (*p* value less than 0.05), indicated with an asterisk. The *TDH3* and *PDC1* non-RPGs had significantly reduced RNAPII occupancy based on fold values, but the third *PYK1* non-RPG with a low basal transcription rate did not. When a relative calculation of the *PYK1* gene results was done by comparing the percentage of change rather than fold values, it was found to be significant as well (**Fig S5**). Regardless, this distinction implied that Tfg1 sumoylation is likely more important than TBP sumoylation for RNAPII recruitment.

3.15 Chromatin Fractionation: TBP is only sumoylated when chromatin bound

Transcription factors can be found in both soluble and chromatin bound states. Chromatin or subcellular fractionation can be used to separate fractions of TFs based upon the principle that chromatin bound or associated proteins are relatively insoluble due to extensive charge-based interactions with DNA or histones [50]. After enzymatic cell wall breakdown, centrifugation can be used to split the soluble fraction of proteins in the supernatant from the chromatin bound fraction of proteins within the pellet. In this case, chromatin fractionation was used to isolate Tfg1 in both soluble and chromatin bound states, and the same procedure was used to study TBP later on as well. A previous lab member assessed the impact of Tfg1 sumoylation on its chromatin binding capability using fractionation. For that experiment, the *ULP1-wt* and *ulp1-1* (alias: *ulp1-mt*) strains, both containing epitope tagged Tfg1, were used. This way the effect of elevated global sumoylation on Tfg1-chromatin interactions could be examined as well. The results revealed that Tfg1 was about equally sumoylated in both soluble and chromatin bound states (**Fig 22A**). A preliminary analysis using the *Tfg1-K60,61R-HA* and *SUMO-Tfg1-HA* strains similarly showed a decrease and increase, respectively, in both soluble and chromatin fractions (**Fig S6**). There was no notable distinction or bias in effect linked to either fraction of Tfg1. In contrast, when a chromatin fractionation was done using the parent *W303a*, tagged *TBP-HA*, and mutant *TBP-3KR-HA* strains, no sumoylated TBP was detected in the soluble fraction (~55kDa band in HA immunoblots) (**Fig 22B**). Even at higher exposures, no sumoylated TBP was detected in the soluble fraction, and was only detected in the chromatin bound fraction. It is highly unlikely that soluble TBP was desumoylated during the procedure as a biological replicate displayed identical results and the same protocol was used to successfully observe soluble Tfg1 sumoylation. Further, while an appropriate reduction in TBP SUMO conjugation was seen within the whole cell extract in the mutant strain, there was seemingly no reduction in sumoylation on the chromatin bound fraction of TBP, despite the three point mutations.

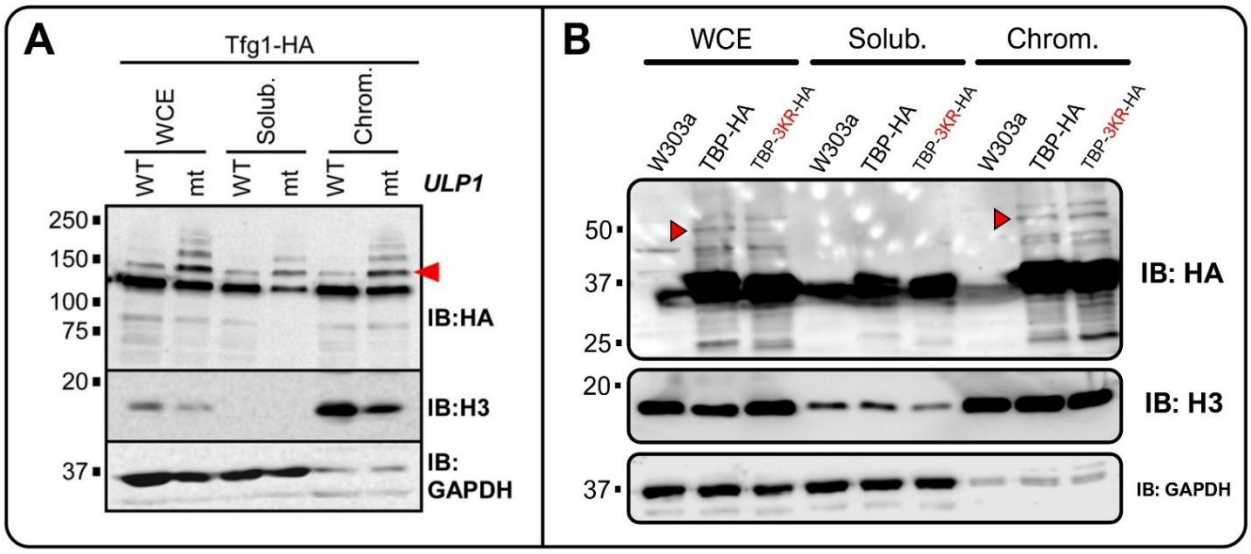

Figure 22. TBP is only sumoylated when chromatin bound, according to a chromatin fractionation analysis. Following fractionation into whole cell extract (WCE), soluble, and chromatin bound/associated fractions, samples were immunodetected with HA immunoblots to compare Tfg1 or TBP SUMO modifications between fractions. Samples were also run on Histone H3 and GAPDH immunoblots as chromatin bound and soluble controls, respectively. **(A)** A previous fractionation analysis from our lab demonstrated that Tfg1 can be modified by SUMO when soluble or chromatin bound. Monosumoylated Tfg1 is indicated with a red arrowhead. A strain with a defective Ulp1 desumoylating protease (*ulp1-1* or *ulp1-mt* strain) resulted in increased Tfg1 sumoylation in all fractions (image from [15]). **(B)** Unlike Tfg1, TBP was only sumoylated when chromatin bound, and was not sumoylated when soluble. Within the whole cell extract, there was a decrease in monosumoylated TBP (red arrowhead) in the mutant *TBP-3KR-HA* strain compared to the *TBP-HA* strain, but the monosumoylation level was seemingly retained when TBP was chromatin bound. Based on the control Histone H3 and GAPDH blots, the fractionation experiment itself was deemed largely successful.

While the retention of sumoylation on chromatin bound TBP within the triple mutant *TBP-3KR-HA* strain was unexpected, it did coincide with the SUMO ChIP-qPCR results above (in section 3.13). There was no observed reduction in overall promoter sumoylation when TBP's sumoylation level was reduced. Back then, this was attributed to either TFIIF's domination in overall promoter sumoylation amount or group sumoylation compensating for the reduction in TBP SUMO conjugation by attaching SUMO to another nearby residue. With the addition of the chromatin fractionation, the lack of a change in SUMO ChIP-qPCRs could now be attributed to TBP maintaining its sumoylation level even with three mutated lysine residues, by an unknown mechanism. The implications of this and the evident contrast with the Tfg1 results will be discussed below (in section 4.3).

There were, however, two technical issues with this finding. First, we cannot assert with complete certainty that the ~55kDa band belonged to SUMO modified TBP. While the size was approximately correct, multiple bands of various sizes belonging to either modified, complexed, or degraded TBP were present in the HA immunoblot. The ~55kDa band could have derived from a different modification or complex. This could be addressed by modifying the protocol to include an affinity chromatography step or by pairing the chromatin fractionation with an immunoprecipitation, as TBP contains the 6xHA epitope tag in the *TBP-HA* and *TBP-3KR-HA* strains. While such a procedure has not been reported in the literature before, we believe that a recently published protocol for enriching chromatin associated proteins permits this hybrid assay and can be pursued in the future [51]. The resulting proteins would be immunoblotted with an anti-SUMO antibody for confirmation of the observed trends in SUMO modifications. The second technical issue was regarding the location or presence of the sumoylated form of TBP in the mutant strain. If there was no change in soluble nor chromatin bound TBP based upon its sumoylation status, why was there a change in the whole cell extract? What other cell location or TBP state manifests the reduction in sumoylated TBP in the *TBP-3KR-HA* strain? This question remains open for future exploration.

Section 4: Discussion and Next Steps

4.1 Fusing SUMO to Tfg1 results in hypersumoylation of Tfg1

When past lab members attempted to study the effects of variable Tfg1 sumoylation via IP-WBs, two strains were used: *Tfg1-K60,61R-HA* and *ulp1-1*. The *Tfg1-K60,61R-HA* strain included a mutant variant of endogenous Tfg1 that had its primary SUMO attachment sites changed to prevent SUMO conjugation, resulting in a Tfg1-specific decrease in sumoylation. However, reduced Tfg1 sumoylation was not found to result in any noticeable decrease in RNAPII-TFIIF interactions based upon unchanged co-IP amounts. When the *ulp1-1* strain was used to simulate increased levels of global sumoylation, a reduction in RNAPII-TFIIF interactions was seen. However, given the vast array of cellular consequences which can stem from globally increased sumoylation, the observed effect could not be attributed to a Tfg1-specific increase in sumoylation. To address this, the *SUMO-Tfg1-HA* strain with a SUMO peptide artificially fused to the N-terminus of endogenous Tfg1 was used. With this strain, the reduction in RNAPII-TFIIF interactions due to elevated Tfg1 sumoylation was confirmed (**Fig 8**).

It should be noted, however, that the SUMO-Tfg1-HA fusion protein was not a perfect recapitulation of *in vivo* sumoylated Tfg1. Normally, Tfg1 is (primarily) sumoylated at Lys60 and/or Lys61 which lie within a disordered region, and not at the N-terminus. This meant that the artificially fused SUMO could have introduced a physical obstruction which diminished RNAPII-TFIIF or Tfg1-chromatin interactions. That being said, Tfg1 has 735 amino acids in *S. cerevisiae*, meaning the artificial SUMO was not relatively far from its natural attachment site. In order to verify that a physical obstruction was not responsible for the observed change to RNAPII-TFIIF interactions, a new strain with a functionally defective form of SUMO could be created. In a 2017 paper, researchers reported that the R47E mutation of SUMO is a lethal mutation in that it is not able to complement a *SMT3* deletion (*smt3*Δ*)* mutation [52]. The researchers concluded that SUMO might use this arginine residue to intermolecularly interact (not attach) with its targets. If a new strain is created in which this functionally inactive or “dead” variant of SUMO is N-terminally

fused to Tfg1, we can assess whether the observed reduction in RNAPII-TFIIF interactions is due to a physical obstruction or legitimate changes to protein-protein interactions. Importantly, the “dead” SUMO could still be polysumoylated (at K11, K15, or K19) by regular “alive” SUMO proteins [53]. And given the hypersumoylation which was observed with even the *monoSUMO-Tfg1-HA* strain, it is likely that a very high level of polysumoylation will occur by functionally active or “alive” SUMO proteins within this new strain. To counteract this, the lysine residues on this “dead” SUMO should be mutated to arginines, preventing its polysumoylation. If RNAPII-TFIIF interactions are disrupted in this new strain in a similar manner to the *SUMO-Tfg1-HA* and *monoSUMO-Tfg1-HA* strains, then a physical obstruction must have been responsible for the observed phenomena. If RNAPII-TFIIF interactions are not disrupted, then we can confidently rule out the physical obstruction possibility. However, given that increased Tfg1 sumoylation (at endogenous sites) via increased global sumoylation within the *ulp1-1* mutant strain also reduced RNAPII-TFIIF interactions, it remains unlikely that a physical obstruction was behind the observed disruption of interactions.

Another interesting takeaway from the use of the *SUMO-Tfg1-HA* and *monoSUMO-Tfg1-HA* strains (as described in section 3.1) was the observed hypersumoylation, or very high level of sumoylation, for Tfg1 in those strains. As suggested above, it is possible that the presence of even a single constitutive, uncleavable SUMO modification on Tfg1 might have induced enzymatically mediated cooperative SUMO conjugation on other parts of Tfg1. In essence, this would occur because the cellular systems would recognize the presence of the initial SUMO moiety and appropriately facilitate the supposedly queued surge in substrate (Tfg1) sumoylation. This surge would be transient per individual substrate as desumoylation enzymes would cleave all the additional SUMO proteins shortly thereafter. However, due to the everlasting presence of the initial fused SUMO, the cycle would restart soon after, and the overall fraction of substrate sumoylated at any given time would increase dramatically, as observed. An alternative

explanation for the why hypersumoylation was observed in the SUMO fusion strains is that the artificial attachment of the SUMO protein dramatically increased Tfg1 stability. Because of this increased stability, a significantly higher amount of Tfg1 was present, resulting in a stronger SUMO signal in western blots. This explanation is based on the fact that the primary use of the SUMO fusion system within the literature is as a biotechnology tool used to enhance target protein stability and solubility, particularly for difficult-to-express proteins. The SUMO fusion system is not typically used to study the functional impact of sumoylation itself, and in biotech applications, the fused SUMO is traditionally removed following protein expression [54].

However, we have reason to doubt the increased stability explanation. Even if fusing SUMO to Tfg1 stabilizes it, that alone does not adequately explain our observations. Within the biotechnology framework, fusing SUMO to Tfg1 would result in the stabilization of the monosumoylated Tfg1 species in particular, so why would that lead to an increase in the presence and stabilization of doubly-, triply-, and further sumoylated Tfg1, other than through the induction of successive sumoylation resulting from the initial modification? In fact, the doubly- and triply-sumoylated SUMO immunoblot bands were found to be more concentrated than the monosumoylated bands (**Fig 8B**). If cooperative sumoylation did not exist, we would only really expect to see the stabilization of the monosumoylated species, as that would be the only one containing an uncleavable SUMO modification. With these considerations, cooperative SUMO conjugation remains the primary explanation for these observations, and not increased stability, though it remains possible that both are at play here. In a self-propagating manner, perhaps cooperatively elevated Tfg1 sumoylation consequently resulted in higher Tfg1 stability.

This finding has significant implications. As far as I am aware, the concept of cooperative SUMO binding has not been directly explored in the literature before, other than being an implication of group sumoylation. If the SUMO fusion system induced cooperative sumoylation on other binding sites of Tfg1, did hypersumoylated Tfg1 consequently induce a cascade of

sumoylation events on other nearby PIC components, a la group sumoylation? This could be validated with IP-WBs. Going a step further, can cooperative sumoylation be induced within other SUMO targets and clusters via the SUMO fusion system, to study the role of SUMO in other cellular contexts? This topic could also be explored in the future.

4.2 Both reduced and increased Tfg1 sumoylation levels diminish RNAPII recruitment

When ChIPs were done using the *Tfg1-HA*, *Tfg1-K60,61R-HA*, and *SUMO-Tfg1-HA* strains, variable Tfg1 SUMO conjugation was not found to impact Tfg1-chromatin association in HA ChIPs (**Fig 10**). SUMO ChIPs then revealed that the overall promoter sumoylation can be affected by altering Tfg1 sumoylation alone (**Fig 11**). This implied that Tfg1 is one of the most sumoylated PIC components and thereby TFIIF is one of the most sumoylated GTFs, if not the most. This effect was only significant at the three tested metabolic non-RPGs with GTF-derived SUMO peaks. The tested RPG did not have a significantly changed SUMO signal. To assess how differential SUMO modification presence at the promoters impacted protein-protein interactions and translated into altered RNAPII recruitment, RNAPII ChIPs were done. Intriguingly, both decreased and constitutively increased Tfg1 sumoylation were found to significantly repress RNAPII recruitment at non-RPGs (**Fig 12**). We reasoned that any disturbance to the normal level of SUMO modifications disrupted the refined balance associated with the numerous interconnected regulatory mechanisms of transcriptional control. A ChIP-seq with the *Tfg1-K60,61R-HA* strain revealed that diminished Tfg1 SUMO conjugation resulted in a widespread decrease in RNAPII density across the genome, at many non-RPGs beyond the previously suspected subset, but expectedly not at RPGs (**Fig 13**). Intriguingly, this was in line with another recent RNAPII ChIP-seq from our lab using the *ubc9-6* strain with globally reduced sumoylation levels. With that strain, RNAPII levels were affected at hundreds of genes, far beyond the subset of 111 unique non-RPGs with detectable SUMO peaks [15]. A subsequent RNA-seq with the

Tfg1-K60,61R-HA strain demonstrated that this reduced RNAPII presence did not correlate with changes to steady-state RNA levels, which were found to be more or less unaffected (**Fig 14**).

There was a discrepancy between the IP-WBs and the ChIPs with regards to the impact of decreased Tfg1 sumoylation in the *Tfg1-K60,61R-HA* strain and resulting interactions with RNAPII. Within IP-WBs, decreased Tfg1 sumoylation did not noticeably impact RNAPII-TFIIF interactions. But within ChIPs, and especially within the ChIP-seq, there was a substantial decrease in RNAPII recruitment at genes, which indicated hindered or otherwise dysregulated RNAPII-TFIIF interactions during transcription. Our explanation for these seemingly contradictory findings is the following model. Initially, nonsumoylated TFIIF would bind RNAPII and escort it to the PIC. This is line with the IP-WBs: a reduction in Tfg1 sumoylation does not notably impact RNAPII-TFIIF interactions. Then, once the full PIC is assembled or at some stage during its assembly, various PIC components become sumoylated at particular genes. Exactly what types of genes are modified and why remains unclear, but certainly extends beyond the 111 unique non-RPGs suspected initially, as indicated by the ChIP-seq. A simultaneous surge in the sumoylation of PIC components like Tfg1 triggers an elevated rate of transcription by currently unknown mechanisms. It is plausible that Tfg1 sumoylation in particular instigates its dissociation from the PIC following promoter clearance by RNAPII in an effort to quickly reset the promoter for the reinitiation of transcription (**Fig 23A**). This way, Tfg1 sumoylation both promotes transcription and reduces RNAPII-TFIIF interactions [15]. Furthermore, as indicated by both IP-WBs and ChIPs, constitutive sumoylation would limit TFIIF from interacting with RNAPII in the first place (before PIC formation in this model) inhibiting transcription and explaining why both reduced and elevated Tfg1 SUMO modifications disrupt RNAPII recruitment (**Figs 23B and C**). Our previously published findings indicated that SUMO is found at the promoters of many active and induced gene promoters but not at repressed or silenced ones [37]. This aligns with the proposed model

because reinitiation as a regulatory control system would be most prevalent at highly transcribed genes.

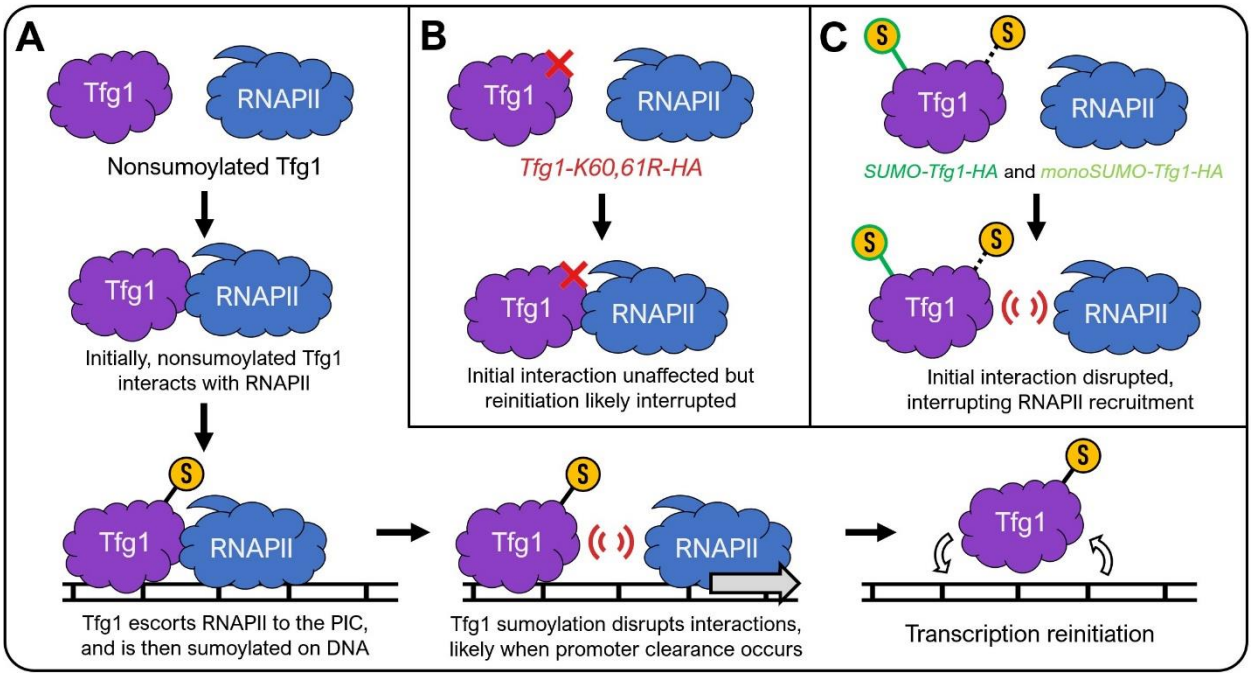

Figure 23. Proposed model of Tfg1 sumoylation. (A) Initially, nonsumoylated Tfg1 (along with the remainder of the TFIIF complex) binds RNAPII and escorts it to the growing preinitiation complex at a particular gene promoter. Likely in conjunction with other PIC components, Tfg1 gets sumoylated once bound to chromatin. This sumoylation is projected to disrupt RNAPII-TFIIF interactions at many genes, possibly around the same time as promoter clearance. Sumoylated Tfg1 is recycled for transcription reinitiation, increasing the overall transcription rate. **(B)** Reduced Tfg1 sumoylation in the *Tfg1-K60,61R-HA* strain does not affect initial RNAPII-TFIIF interactions, but an interruption in the reinitiation process due to blocked sumoylation reduces the overall transcription rate. **(C)** Constitutive Tfg1 sumoylation in the *SUMO-Tfg1-HA* strain likely limits the initial RNAPII-TFIIF interactions, disrupting initial RNAPII recruitment and reducing the overall transcription rate.

4.3 Differential regulation of transcription via Tfg1 vs TBP sumoylation

After the *TBP-3KR-HA* strain was generated, it was found that reducing TBP's sumoylation level did not impact cell viability (**Figs 16, 17, and 18**). ChIP experiments demonstrated that while this mutation did not affect TBP's chromatin occupancy or overall promoter sumoylation, it did result in significantly reduced RNAPII recruitment at target non-RPG promoters (**Fig 19, 20, and 21**). While not in line with *in vitro* suggestions utilizing human cells [40], a chromatin fractionation revealed that TBP is sumoylated essentially exclusively when chromatin bound (**Fig 22**). While it remains possible that some low level of sumoylated soluble TBP exists, it was not detected in

immunoblots. As with the Tfg1 analysis, future steps in the study of how TBP sumoylation regulates transcription in *S. cerevisiae* include an RNAPII ChIP-seq to assess globally modulated RNAPII density, and then RT-qPCRs and an RNA-seq to measure resultant transcriptome level changes. Additionally, as described above (in section 3.15), a chromatin fractionation paired with an immunoprecipitation could validate the confounding fractionation observations, and thereby could highlight an interesting distinction between human versus yeast TBP sumoylation. Alternatively, if a new strain is created with SUMO artificially and endogenously fused to TBP, and this constitutively sumoylated variant of TBP is similarly found to not be present within the soluble fraction, despite the increased stability that sumoylation seems to provide [54], that would further support the notion of TBP being sumoylated exclusively or at least predominantly when chromatin bound.

The TBP chromatin fractionation result was in line with the SUMO ChIP-qPCRs which uncovered no notable reduction in overall promoter sumoylation when the *TBP-3KR-HA* strain was used but was not in line with the RNAPII ChIP-qPCRs, which uncovered a positive correlation between TBP sumoylation and RNAPII recruitment at non-RPGs. If TBP sumoylation at promoters was unchanged between strains, and therefore protein-protein interactions were supposedly unchanged, why was a difference in RNAPII recruitment observed? There are two explanations to reconcile this apparent contradiction. First of all, only monosumoylated TBP bands were seen within the HA immunoblot of the chromatin bound fraction. It is possible that there was truly a reduction in doubly-, triply-, and higher SUMO bands, which simply was not captured. Second, since endogenous TBP had its lysines mutated in the *TBP-3KR-HA* strain, the chromatin bound mutant form of TBP which somehow maintained its sumoylated status must have been sumoylated elsewhere, on a different lysine residue. It is possible that this residue (or residues) only becomes accessible for SUMO conjugation following TBP attachment to chromatin and a subsequent conformational change, exposing a new binding pocket on TBP. In such a case,

sumoylation would occur at a non-traditional site on TBP, and even though TBP may still be "sumoylated", this sumoylation at a different location would result in a different impact on protein-protein interactions compared to the traditional sumoylation sites. In this manner, TBP would still maintain its sumoylated status, but its indirect interactions with RNAPII would be disrupted.

Decades of research on SUMO have uncovered that sumoylation seems to have a context-dependent role, and our findings on Tfg1 versus TBP sumoylation evidently reinforce that notion. Despite their complementary roles in transcriptional initiation, SUMO modifications seem to control Tfg1 and TBP interactions in slightly different ways. While the sumoylation of both appears to promote RNAPII recruitment and thereby transcription at gene promoters, the exact mechanism of action seems to be distinct. As described in our model for TFIIF, Tfg1 sumoylation is believed to promote transcription by prompting the release of TFIIF from chromatin for the rapid reinitiation of transcription. However, given that a soluble form of sumoylated TBP was not detected, a different model/mechanism must be employed to explain how SUMO modifications control TBP, and thereby how TBP sumoylation regulates transcription. Also, given that a soluble form of sumoylated TBP was not detected, the binding site selection model of TF-chromatin interactions (as discussed in section 1.2) likely does not control it either [17]. Within the binding site selection model, spuriously bound TBP would be sumoylated to be released from chromatin and given that TBP often has to bind TATA-less promoters, we would expect to see substantial spurious binding and thereby soluble sumoylated TBP [29,30]. Given all this, a potential model which encompass these findings is as follows. Initially, nonsumoylated TBP scans chromatin and binds to promoters, along with associated TAFs. Subsequent sumoylation of promoter bound TBP triggered by as of yet unknown conditions or at particular genes then enhances the coordination of interactions with RNAPII, increasing its rate or proficiency of recruitment. Before release from the PIC, TBP's sumoylation is reversed (**Fig 24**). This model and the experimental findings which

preceded it highlight a subtle but important distinction between Tfg1 versus TBP sumoylation, and how their modifications promote an elevated rate of transcription.

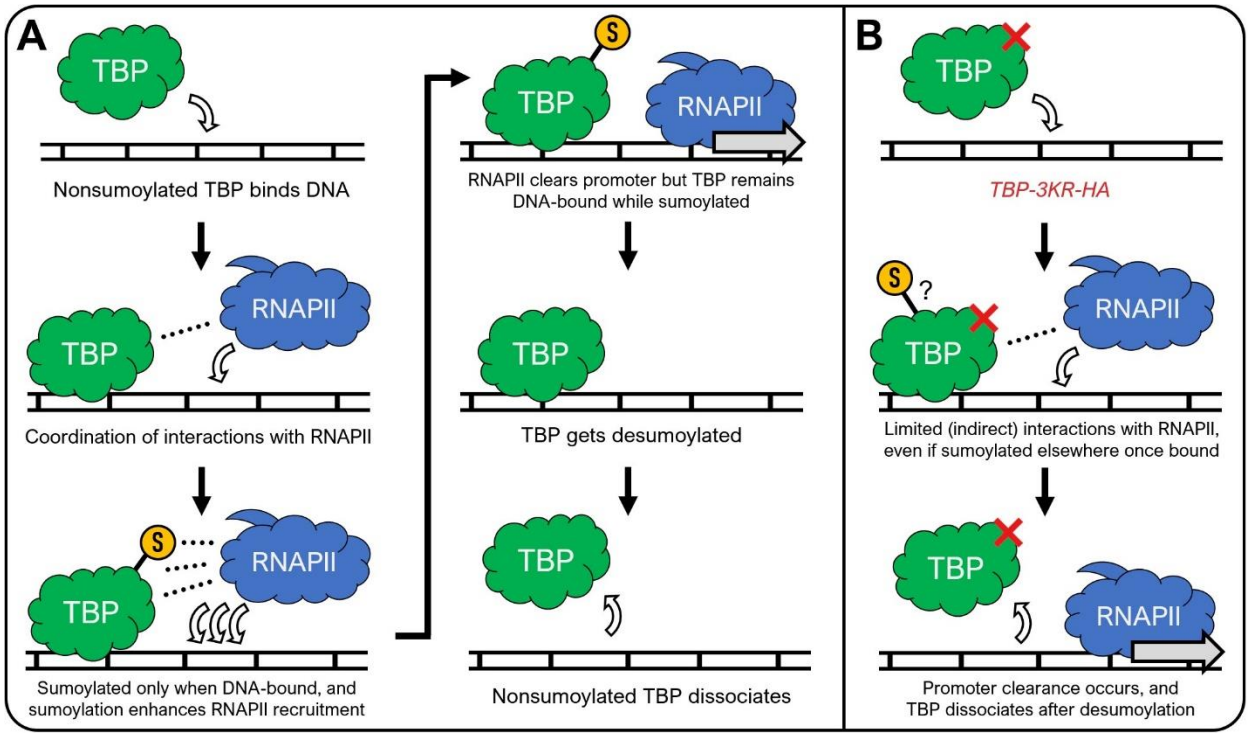

Figure 24. Proposed model of TBP sumoylation. (A) To begin, nonsumoylated TBP scans chromatin and binds to a gene promoter, prompting the formation of a new PIC. In doing so, TBP contributes to RNAPII recruitment by coordinating interactions within the network of GTFs. Once chromatin bound, TBP can be sumoylated, which is correlated with an increase in RNAPII recruitment at particular genes. This likely occurs because of a newly masked, unmasked, or introduced (i.e. the SUMO peptide) interaction motif on TBP resulting in different or additional protein-protein interactions which indirectly enhance RNAPII recruitment. Following promoter clearance by RNAPII, TBP is desumoylated and then dissociates. **(B)** When TBP sumoylation at the three primary sumoylation sites is blocked within the *TBP-3KR-HA* strain, RNAPII recruitment is limited. TBP is suspected to be sumoylated even within the mutant strain, as indicated by the chromatin fractionation (**Fig 22**), but at a different, unknown location which is unable to functionally compensate for the impact of sumoylation at the standard sites: Lys47, Lys127, and Lys167.

4.4 Group sumoylation likely coordinates the initiation of transcription

An emerging concept from our research on GTF sumoylation is the notion of group sumoylation controlling PIC sumoylation, and thereby regulating transcription. This would explain why blocking the sumoylation of Tfg1 or TBP alone did not have a major impact on transcription, at least when compared to how the sumoylation of Rap1 at RPGs was found to be very important for gene expression and general cell survival [36]. At the non-RPGs in the small-scale ChIPs (**Figs 12A and 21**), and at hundreds of genes in the RNAPII ChIP-seq (**Fig 13A**), RNAPII recruitment was

only diminished partly, and was not completely repressed. Furthermore, even in the cases where RNAPII recruitment was more or less abolished, the changes to steady-state mRNA levels were minimal, as revealed by the RNA-seq (**Fig 14**). It is suspected that cellular mechanisms buffered any changes to the transcriptome by stabilizing existing transcripts (as described in section 3.7), implying that the changes to transcriptional initiation were not substantial enough to bypass a certain buffering threshold. Given this, we believe it is likely that multiple GTF components become sumoylated coordinately to impact transcription synergistically.

In order to test and experimentally confirm this notion, future research will need to focus on mutating entire clusters or complexes, such as the entire PIC in this case. This would be accomplished by identifying any potential SUMO attachment sites on GTF subunits, RNAPII, and any other peripheral transcription initiation components, and mutating all of them (to arginine residues) simultaneously, within the same strain. For the precise detection of changes to the multiple involved proteins, each modified protein should be epitope tagged. All, or at least a large number of, pertinent and interrelated components would need to be mutated, as the compensatory sumoylation of components which are not typically sumoylated cannot be ruled out, given the seemingly indiscriminate nature of group sumoylation. Furthermore, given that a reduction in Tfg1 sumoylation resulted in drops in RNAPII densities at thousands of genes lacking detectable SUMO peaks according to the ChIP-seq (as described in section 3.6), it is possible that many more transcriptional components can be SUMO modified than we are currently able to detect. As such, the approach of preemptively mutating all PIC components is likely worthwhile. Because introducing such a large number of mutations might result in imperfect PIC association, any changes to the efficacy of PIC assembly must be continually considered and documented. Once the sumoylation level of the entire complex is abolished or at least considerably reduced, we can deduce the true impact of variable GTF and thereby PIC sumoylation, without group sumoylation potentially obscuring the conclusions.

An alternative approach would be to study the role of group sumoylation in PIC assembly and activity by simply epitope tagging all PIC components without introducing any mutations. Then, point mutations or the constitutive attachment of SUMO could be introduced on a one-by-one basis, creating new strains containing individually modified PIC components. Following the creation of these strains, the effects of the variable sumoylation of individual GTF components on the sumoylation level of particular nearby PIC components could be assessed. For instance, as alluded to before, does fusing SUMO to Tfg1 impact the sumoylation level of the other GTFs and RNAPII via a SUMO attachment cascade stemming from cooperative sumoylation? This could be experimentally verified with standard co-IPs and immunoblots to examine the overall changes to interactions between PIC components based on sumoylation status.

To go a step further, CRISPR-Cas9-based technologies such as ‘engineered ChIP’ could be utilized to detect changes in protein-protein interactions at particular promoters, rather than being limited to the cellular average [55]. In brief, this approach would rely on using guide RNA to direct the dead Cas9 (dCas9) enzyme to particular segments of sheared genomic DNA with cross-linked DNA-associated molecules like transcription factors. Then, dCas9 would be immunoprecipitated and, after the removal of cross-links and a purification step, the co-IPed DNA-associated proteins could be immunoblotted or analyzed via mass spectrometry. Using this approach, the impact of Tfg1 sumoylation on RNAPII sumoylation at genes with low vs high rates of gene expression could be compared, for instance. Taken together, the potential role of group sumoylation in controlling PIC assembly and RNAPII recruitment could be explored using contemporary lab techniques.

4.5 Conclusion

In conclusion, our findings demonstrate that the sumoylation of the Tfg1 subunit of TFIIF and the TBP subunit of TFIID promote transcription by enhancing RNAPII density at various genes. Mutant strains with variable levels of Tfg1 or TBP sumoylation were used to assess the functional impact of GTF SUMO conjugation. We found that Tfg1 and TBP sumoylation does not impact their chromatin interactions but variable Tfg1 sumoylation does impact the promoter wide SUMO signal. The sumoylation of both Tfg1 and TBP promotes RNAPII recruitment likely via altered protein-protein interactions but does so through seemingly different mechanisms. A ChIP-seq and an RNA-seq with the *Tfg1-K60,61R-HA* strain revealed that the effects of GTF sumoylation go beyond the initially suspected 111 unique non-RPGs but are not sufficient enough to impact steady-state mRNA levels. While it remains unknown why thousands of genes without detectable GTF-derived SUMO peaks were affected, this research has uncovered novel characteristics of transcriptional regulation in eukaryotes.

Appendix: Supplementary Data

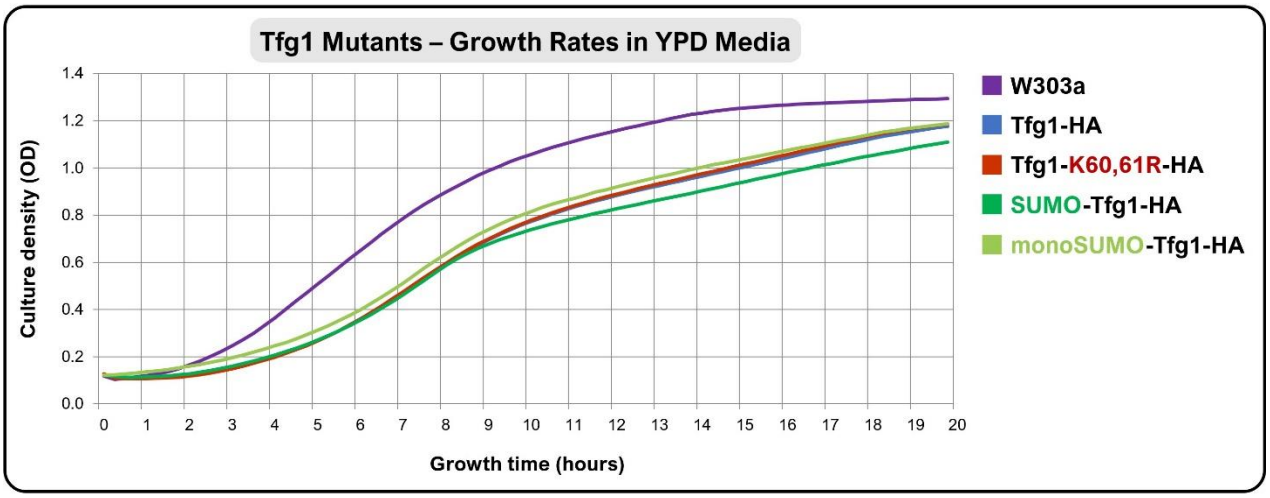

Figure S1. The *SUMO-Tfg1-HA* strain suffers from a slight growth defect. Growth rates of Tfg1 mutant strains as measured by an accuSkan FC automated spectrophotometer. Monitored the growth rates of the following strains in YPD media: wild-type *W303a*, just tagged *Tfg1-HA*, SUMO-deficient *Tfg1-K60,61R-HA*, the *SUMO-Tfg1-HA* strain with an artificially fused SUMO, and the *monoSUMO-Tfg1-HA* strain with an artificially fused SUMO which could not be polysumoylated. The *W303a* strain grew the quickest, with the *Tfg1-HA*, *Tfg1-K60,61R-HA*, and the *monoSUMO-Tfg1-HA* strains trailing it close together. The slight growth defect of these three strains likely derived from the presence of the HA Tag. The *SUMO-Tfg1-HA* strain grew the slowest. Triplicates of each strain were incubated at 30°C while changes in culture density were tracked over time. Triplicate results were then averaged and graphed. This data was kindly provided by Dr. J. Bryan McNeil.

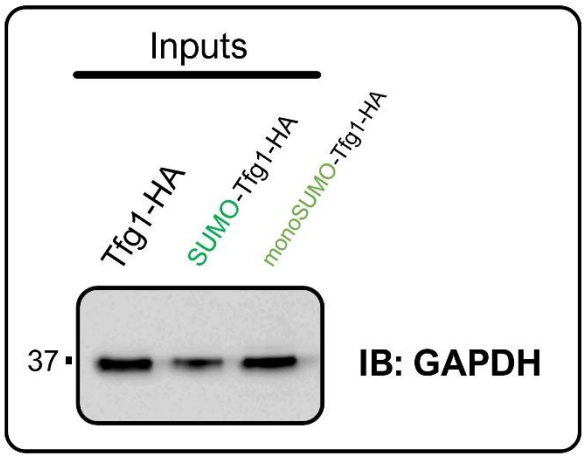

Figure S2. The *SUMO-Tfg1-HA* strain has slightly lower overall expression levels. GAPDH immunoblot of input (lysate) samples from HA Tag immunoprecipitation shown in **Fig 8B**. Compared to the *Tfg1-HA* strain, the *SUMO-Tfg1-HA* strain had slightly reduced band intensity upon plain visual examination when detecting the constitutively expressed GAPDH protein, indicating hindered overall expression or cell growth and survival. Preventing polysumoylation on the fused SUMO moiety by using the *monoSUMO-Tfg1-HA* strain rescued the reduction in GAPDH levels.

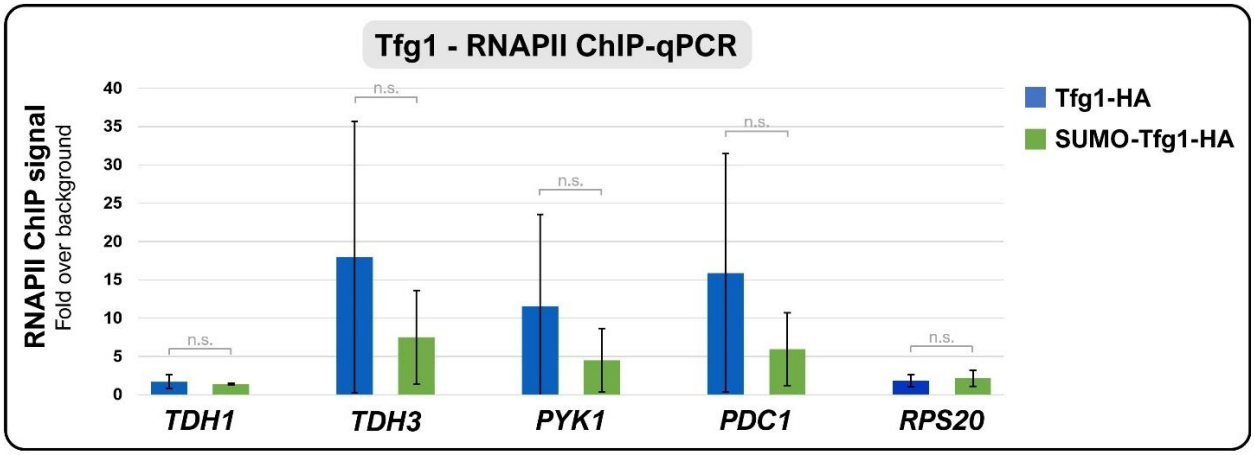

Figure S3. Tfg1 - RNAPII ChIP-qPCR: Original *SUMO-Tfg1-HA* data illustrates high level of variability. In this non-relative version of **Fig 12B**, the large error bars which prompted the switch to relative values can be seen. For this ChIP, an anti-Rpb3 antibody was used to pull down DNA segments cross-linked to Tfg1 in the *Tfg1-HA* and *SUMO-Tfg1-HA* strains. Five gene promoters were assessed for differential RNAPII occupancy: the *TDH1* non-RPG with no SUMO peak as a background control, the *TDH3*, *PYK1*, and *PDC1* non-RPGs with GTF associated SUMO peaks, and the *RPS20* RPG with a Rap1 associated SUMO peak. The average of three trials was calculated and the standard deviation was represented with error bars. There was great variability between trials, likely resulting from a fluctuating rate of transcription and resulting in large error bars, but overall trends remained consistent.

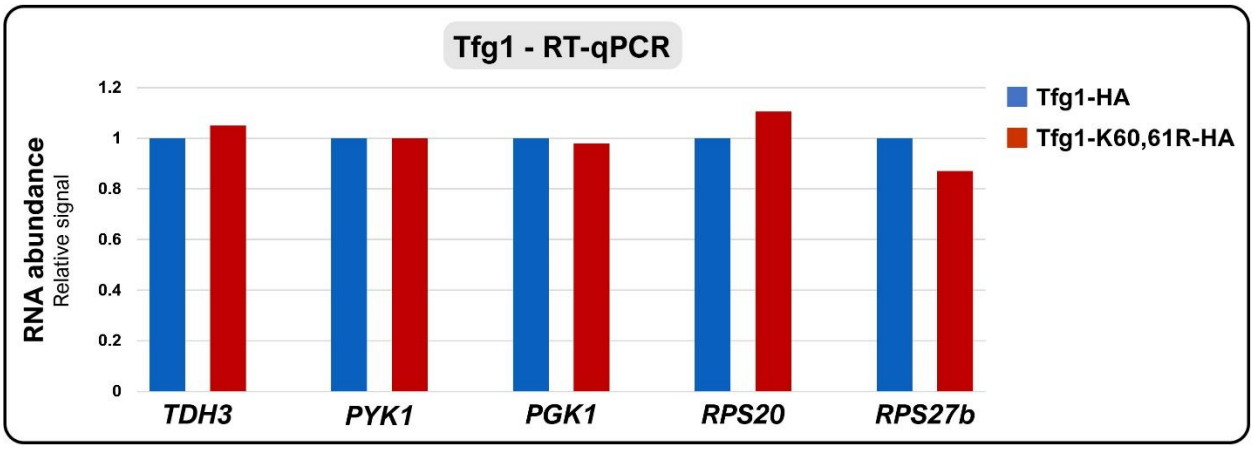

Figure S4. Tfg1 – RT-qPCR: Reduced Tfg1 sumoylation does not substantially impact the expression levels of select genes. In this exemplary, amalgamated trial, the minor changes to mRNA levels between the *Tfg1-HA* and *Tfg1-K60,61R-HA* strains within RT-qPCRs were evident. The relative values in these charts were based on an amalgamated trial, as different combinations of gene products were analyzed on separate days. All mRNA signals were normalized to the constitutively expressed 25S rRNA signal prior to calculating relative values. Three non-RPGs with GTF-derived SUMO peaks were chosen: *TDH3*, *PYK1*, and *PGK1*. Two RPGs with Rap1-derived SUMO peaks were chosen: *RPS20* and *RPS27b*. Because substantial changes were not observed for transcripts from any of these selected genes, even after multiple trials with certain gene products, we instead chose to examine any potential global changes to steady-state mRNA levels via an RNA-seq (**Fig 14**), but again observed only miniscule significant differences.

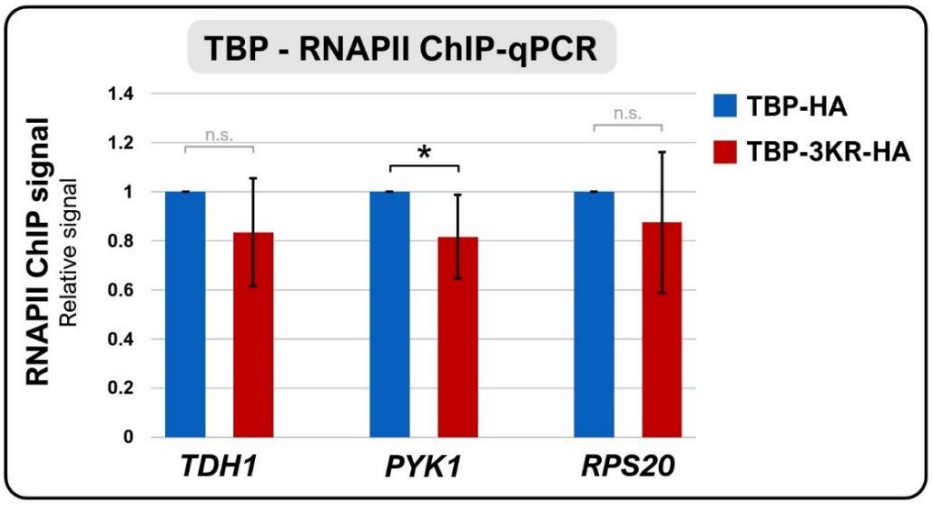

Figure S5. TBP – RNAPII ChIP-qPCR: *PYK1* gene results are significant when using relative values. When the sumoylation level of TBP was reduced using the *TBP-3KR-HA* strain, the *PYK1* gene was not found to have a significant drop in RNAPII presence when analyzing the original (ChrV normalized) values. In this relative version of **Fig 21**, however, *PYK1* was found to have a significant drop in RNAPII levels when TBP's sumoylation levels were reduced, although by only a slight margin. This indicated that while reduced TBP sumoylation impacted the three select non-RPGs with GTF-derived SUMO peaks, its impact was not as drastic as reduced Tfg1 sumoylation. Even with relative values, the *TDH1* background control and *RPS20* RPG did not have significantly impacted RNAPII occupancies. The average of four anti-Rpb3 ChIP trials was calculated, the standard deviation was represented with error bars, and the Student's *t*-test was used to assess significant changes (*p* value less than 0.05), indicated with an asterisk.

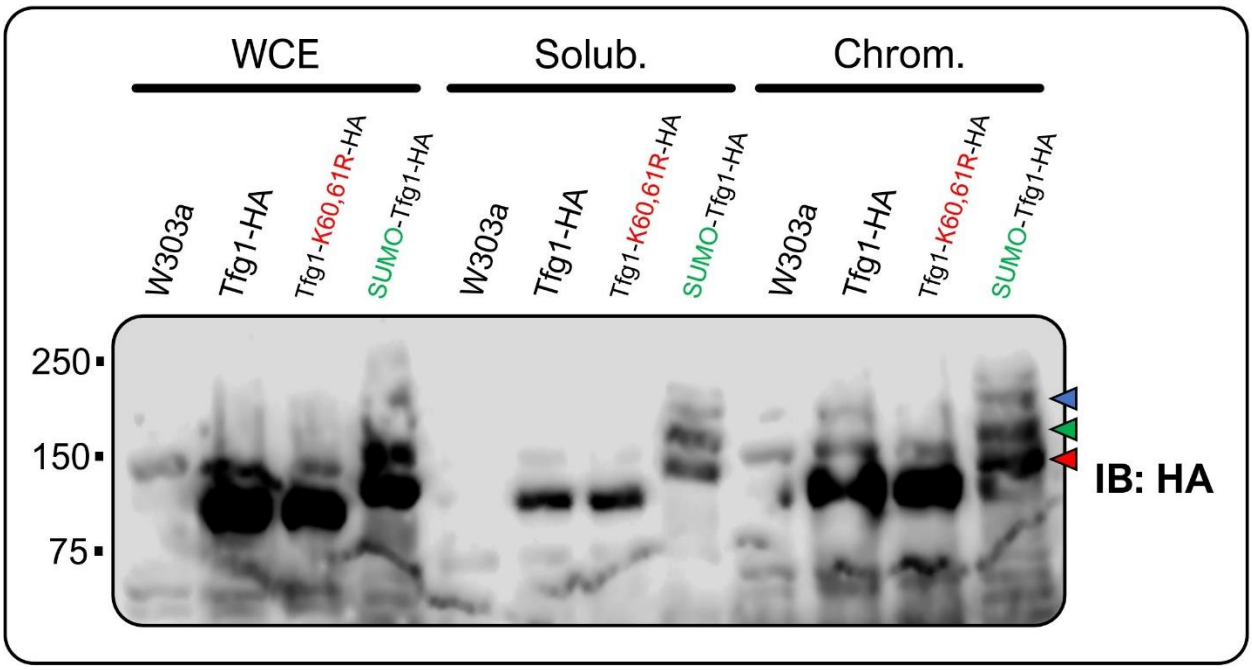

Figure S6. Increased or decreased Tfg1 sumoylation impacts both soluble and chromatin bound Tfg1 approximately equally. Tfg1 was found to be sumoylated when soluble or chromatin bound, this was in contrast to TBP, which was not found to be sumoylated when soluble. Within this HA Tag immunoblot, Tfg1 runs at ~110 kDa, monosumoylated Tfg1 runs at ~140 kDa (red arrowhead), doubly-sumoylated Tfg1 runs at ~170 kDa (green arrowhead), and triply-sumoylated Tfg1 runs at ~200 kDa (blue arrowhead). GAPDH and Histone H3 control blots were not run for this preliminary trial. In this chromatin fractionation using the parental wild-type *W303a*, the just tagged *Tfg1-HA*, the SUMO-deficient *Tfg1-K60,61R-HA*, and the *SUMO-Tfg1-HA* strain with constitutively sumoylated Tfg1, the sumoylation status of Tfg1 was assessed in various cellular fractions. Within the chromatin bound fraction, there was a plain correlation between theoretical Tfg1 sumoylation and the observed SUMO band count, implying that altering Tfg1 sumoylation directly impacted the sumoylation level of Tfg1 while it was chromatin bound, unlike with TBP. Unfortunately, there was a background band within the whole cell extract and chromatin bound fractions for the *W303a* negative control sample which ran at an almost identical rate to sumoylated Tfg1 bands, making those results less reliable. Notably, however, the background band was not present within the soluble fraction, and furthermore multiple sumoylated Tfg1 bands were visible for the *SUMO-Tfg1-HA* sample in the soluble fraction. Both of these observations indicated that Tfg1, unlike TBP, can be sumoylated while soluble.

Table S1. List of *Saccharomyces cerevisiae* strains used or referenced in this study.

Name	Lab Code	Genotype	Source
W303a	YER058	*MAT a ura3-52 trp1Δ2 leu2-3_112 his3-11 ade2-1 can1-100*	Dharmacon
Tfg1-HA	YJB002A	*TFG1-6xHA::K. lactis TRP1*	Justin Burgener
Tfg1-K60,61R-HA	YRB013A	*tfg1-K60,61R-6xHA::K. lactis TRP1 (*with inadvertent N462Q substitution)	Russell Bahar
ULP1-wt (with *Tfg1-HA*)	YYD002D	*ULP1 TFG1-6xHA::K. lactis TRP1*	Yimo Dou and Martine Heude
ulp1-1 (with *Tfg1-HA*)	YYD003A	*ulp1-I615N TFG1-6xHA::K. lactis TRP1*	Yimo Dou and Martine Heude
SUMO-Tfg1-HA	YJBM041	*SMT3(1-96ΔGG)-TFG1-6xHA::K. lactis TRP1*	J. Bryan McNeil and Yimo Dou
monoSUMO-Tfg1-HA	YJBM043	*smt3-KallR-(1-96ΔGG)-TFG1-6HA::K. lactis TRP1*	J. Bryan McNeil
Lys1Δ	YER042B	*lys1Δ::NAT-6MX tor1-1 fpr1::loxP-LEU2-loxP RPL13A-2×FKBP12::loxP*	Emanuel Rosonina
ubc9-6	YER050	*ubc9-1:Tadh1:TRP1:ubc9-1*	Damien D'Amours
TBP-HA	YMB001	*SPT15-6xHA::K. lactis TRP1*	This study
TBP-K47R-HA	YMB008	*spt15-K47R-6xHA::K. lactis TRP1*	This study
TBP-K127,167R-HA	YMB009	*spt15-K127,167R-6xHA::K. lactis TRP1*	This study
TBP-3KR-HA	YMB010B	*spt15-K47,127,167R-6xHA::K. lactis TRP1*	This study

Table S2. Sequence of primers used in this study. Codons resulting in point mutations highlighted in red.

Gene	Region(s)	Use case	Direction	Sequence (5' to 3')
SPT15	Promoter, ORF, and UTR	Regular PCR (flanking primers)	Forward	ACATATAAAACATGGCTTCAAAGGA
			Reverse	GAAAGTATGAAAAAGAATTGTCGGA
SPT15	ORF and UTR	K47R mutagenesis	Forward	GATGGTACAAAACCAGCAACTACTTT CCAGAGTGAAGAGGACATAAGAAGA GCTGCCCCAGAATCT
			Reverse	GAAAGTATGAAAAAGAATTGTCGGA
SPT15	Promoter and ORF	K47R and 3KR fusion PCR upstream product	Forward	ACATATAAAACATGGCTTCAAAGGA
			Reverse	CTGATGTGGCGGAGGTGTCTTTTTCA GATTCTGGGGCAGCTCTTCTTATGTC CTCTTCACTCTG
SPT15	Promoter and ORF	K47R and 3KR fusion PCR	Forward	ACATATAAAACATGGCTTCAAAGGA
			Reverse	GAAAGTATGAAAAAGAATTGTCGGA
SPT15	ORF and UTR	K127R mutagenesis	Forward	AGCTTTAATTTTTGCCTCAGGGAAAA TGGTTGTTACCGGTGCAAGAAGTGA GGATGACT
			Reverse	GAAAGTATGAAAAAGAATTGTCGGA
SPT15	ORF and UTR	K167R mutagenesis	Forward	CAGACTTCAAAATACAAAATATTGTC GGTTCGTGTGACGTTAGATTCCCTAT ACGTC
			Reverse	GAAAGTATGAAAAAGAATTGTCGGA
TFG1	Promoter, ORF, and UTR	Regular PCR (flanking primers)	Forward	TCAATTAGAGTATGCGGTTATGGAT
			Reverse	GAGTGAAGAAGAAGAAGATGACGAC
TFG1	ORF and UTR	K60,61R mutagenesis	Forward	TCCCCAGGCGTACCAAATGGAGACA ACAGCAGGGGATCTCTTGTCAGAAG GGATGATCCTGAGTATGC
			Reverse	GAGTGAAGAAGAAGAAGATGACGAC
(Chr. V)	UTR	qPCR normalization	Forward	CATTATCCGTAACGCCACTTT
			Reverse	CGATCTTAGTTCCAATGGTGAAA
TDH1	Promoter	qPCR	Forward	TGACCAAAACTGGAGTCTCG
			Reverse	TGCAAGAGAGAGAATAGAACTG
TDH3	Promoter	qPCR	Forward	TTAACGGTTTCGGTAGAATCGG
			Reverse	AACAACTTCGACGTTTGGTCTA
PYK1 (*CDC19*)	Promoter	qPCR	Forward	CTCCTTGCAATCAGATTTGG
			Reverse	TTGCGTGAGGTTATGAGTAG
PDC1	Promoter	qPCR	Forward	CTCCTTGCAATCAGATTTGG
			Reverse	TTGCGTGAGGTTATGAGTAG
PGK1	Promoter	qPCR	Forward	GAATCGTGTGACAACAACAGCC
			Reverse	AACCACCCCCTTGGTTAGA
RPS20	Promoter	qPCR	Forward	CGCGACTAGCCTCAGAGATT
			Reverse	GCTGAGCTTGAATGAAATAACCC
RPS27b	Promoter	qPCR	Forward	AATTCCCCTCTGCTTCCCG
			Reverse	ACAGCACACATGAAAGATGAGA

Table S3. *Tfg1-HA* vs *Tfg1-K60,61R-HA* RNA-seq NGS data analysis tools and parameters utilized. This data was sourced from our recent publication [15].

Samples and conditions	Two independent replicates were prepared from cultures grown in SC medium at 30°C. 1) Tfg1-HA 2) Tfg1-K60,61R
Library synthesis	NEB Ultra II directional mRNA library prep (PolyA enrichment)
Sequencing	Illumina HiSeq 2500; Paired-end reads; 2x 126 nt; 30 million reads/sample
Quality control	FastQC (0.11.9)
Genome alignment	HISAT2 (2.1.0) Parameters: Default options; UCSC annotaion file: sacCer3.ncbiRefSeq.gtf
Transcript counting	Transcripts counted using featureCounts from Subread (2.0.0) Parameters: Default options; -p (paired)
Differential expression analysis	edgeR (3.32.1) Parameters: Library sizes were normalized using calcNormFactors; dispersions were estimated with estimateDisp with robust argument; likelilhood ratio tests for differential expression were performed with glmFit and glmLRT.

Table S4. *Tfg1-HA* vs *Tfg1-K60,61R-HA* RNAPII ChIP-seq NGS data analysis tools and parameters utilized. This data was sourced from our recent publication [15].

Samples and conditions	Two independent replicates were prepared from cultures grown in SC medium at 30°C. Inputs and 8WG16 (Rpb1 antibody) IPs were sequenced. 1) Tfg1-HA 2) Tfg1-K60,61R-HA
Library synthesis	NEBNext Ultra II DNA library prep kit (New England Biolabs)
Sequencing	Illumina HiSeq 2500; Paired-end reads; 2x 126 nt; 10 million reads/sample
Quality control	FastQC (0.11.9)
Trimming	TrimGalore (0.6.6) Parameters: paired, length 40, stringency 5, illumina –q 25, clipped 6 bp from 5' end
Genome alignment	Bowtie2 (2.3.5.1) with sacCer3 reference genome
Peak calling	MACS (2.2.7.1) Parameters: paired-end; input as control; broad region calling; effective genome size 1.2e7; q-value cut-off: 0.1
Differential binding analysis	DiffBind (2.16.0) Parameters: minMembers=2 for dba.contrast; th=1 for dba.report; see notes below
Peak analysis and annotation	ChIPpeakAnno (3.22.0) from Bioconductor Parameters: TxDb.Scerevisiae.UCSC.sacCer3.sgdGene genome annotation package was used and the closest feature to the middle of each peak was used for annotation.
Notes	To determine RNAPII density at each ORF: DiffBind was applied to the four samples using a pre-defined peak-set that corresponds to ORF regions of all protein-coding genes. The "th=1" parameter was applied to determine the "concentration" (log_2 normalized ChIP read counts with control read counts subtracted) of RNAPII at all ORFs in both conditions. Here, these are referred to as RNAPII densities.

Additional supplementary information can be found within the 'supporting information' section of our published data (https://doi.org/10.1371/journal.pgen.1009828) [15]. This includes the full list of RNAPII ORF densities from the ChIP-seq, and the full list of mRNA concentrations from the RNA-seq.

References

1. Hay RT. SUMO: A History of Modification. Molecular Cell. 2005; 18(1): 1-12. https://doi.org/10.1016/j.molcel.2005.03.012

2. Boulanger M, Chakraborty M, Tempé D, Piechaczyk M, Bossis G. SUMO and Transcriptional Regulation: The Lessons of Large-Scale Proteomic, Modifomic and Genomic Studies. Molecules. 2021; 26(4): 828. https://doi.org/10.3390/molecules26040828

3. Novatchkova M, Budhiraja R, Coupland G, Eisenhaber F, Bachmair A. SUMO conjugation in plants. Planta; 2004; 220: 1-8. https://doi.org/10.1007/s00425-004-1370-y

4. Seifert A, Schofield P, Barton GJ, Hay RT. Proteotoxic stress reprograms the chromatin landscape of SUMO modification. Sci Signal. 2015; 8(384): rs7. https://doi.org/10.1126/scisignal.aaa2213

5. Baczyk D, Audette MC, Drewlo S, Levytska K, Kingdom JC. SUMO-4: A novel functional candidate in the human placental protein SUMOylation machinery. PLOS One. 2017; 12(5): e0178056. https://doi.org/10.1371/journal.pone.0178056

6. Liang YC, Lee CC, Yao YL, Lai CC, Schmitz ML, Yang WM. SUMO5, a Novel Poly-SUMO Isoform, Regulates PML Nuclear Bodies. Scientific Reports. 2016; 6: 26509. https://doi.org/10.1038/srep26509

7. Gareau JR, Lima CD. The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. Nat Rev Mol Cell Biol. 2010; 11(12): 861–871. https://doi.org/10.1038/nrm3011

8. Naik MT, Naik N, Shih H, Huang T. Solution structure of human SUMO1. 2016. https://doi.org/10.2210/pdb2n1v/pdb

9. Mossessova E, Lima CD. X-ray structure of the C-terminal ULP1 protease domain in complex with Smt3, the yeast ortholog of SUMO. 2011. https://doi.org/10.2210/pdb1euv/pdb

10. Boulanger M, Paolillo R, Piechaczyk M, Bossis G. The SUMO Pathway in Hematomalignancies and Their Response to Therapies. Int. J. Mol. Sci. 2019; 20(16): 3895. https://doi.org/10.3390/ijms20163895

11. Sobko A, Ma H, Firtel RA. Regulated SUMOylation and Ubiquitination of DdMEK1 Is Required for Proper Chemotaxis. Developmental Cell. 2002; 2(6): 745-756. https://doi.org/10.1016/S1534-5807(02)00186-7

12. Smet-Nocca C, Wieruszeski JM, Léger H, Eilebrecht S, Benecke A. SUMO-1 regulates the conformational dynamics of Thymine-DNA Glycosylase regulatory domain and

competes with its DNA binding activity. BMC Biochem. 2011; 12: 4. https://doi.org/10.1186/1471-2091-12-4

13. Rosonina E, Akhter A, Dou Y, Babu J, Sri Theivakadadcham VS. Regulation of transcription factors by sumoylation. Transcription. 2017; 8(4): 220–231. https://doi.org/10.1080/21541264.2017.1311829

14. Knorre DG, Kudryashova NV, Godovikova TS. Chemical and Functional Aspects of Posttranslational Modification of Proteins. Acta Naturae. 2009; 1(3): 29–51. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3347534/

15. Baig MS, Dou Y, Bergey BG, Bahar R, Burgener JM, et al. Dynamic sumoylation of promoter-bound general transcription factors facilitates transcription by RNA polymerase II. PLoS Genet. 2021; 17(9): e1009828. https://doi.org/10.1371/journal.pgen.1009828

16. Lyst MJ, Stancheva I. A role for SUMO modification in transcriptional repression and activation. Biochem Soc Trans. 2007; 35(6): 1389–1392. https://doi.org/10.1042/BST0351389

17. Rosonina E. A conserved role for transcription factor sumoylation in binding-site selection. Current Genetics. 2019; 65: 1307-1312. https://doi.org/10.1007/s00294-019-00992-w

18. Tang Y, Zhao W, Chen Y, Zhao Y, Gu W. Acetylation Is Indispensable for p53 Activation. Cell. 2008; 133(4): 612–626. https://doi.org/10.1016/j.cell.2008.03.025

19. Wu SW, Chiang CM. p53 sumoylation: mechanistic insights from reconstitution studies. Epigenetics. 2009; 4(7): 445–451. https://doi.org/10.4161/epi.4.7.10030

20. Psakhye I, Jentsch S. Protein Group Modification and Synergy in the SUMO Pathway as Exemplified in DNA Repair. Cell. 2012; 151(4): 807-820. https://doi.org/10.1016/j.cell.2012.10.021

21. Luse DS. The RNA polymerase II preinitiation complex. Transcription. 2014; 5(1): e27050. https://doi.org/10.4161/trns.27050

22. Plaschka C, Hantsche M, Dienemann C, Burzinski C, Plitzko J, Cramer P. Transcription initiation complex structures elucidate DNA opening (CC). 2019. https://doi.org/10.2210/pdb5fz5/pdb

23. Hendriks IA,Lyon D, Young C, Jensen LJ, Vertegaal ACO, et al. Site-specific mapping of the human SUMO proteome reveals co-modification with phosphorylation. Nature Structural & Molecular Biology. 2017; 24: 325–336. https://doi.org/10.1038/nsmb.3366

24. Rani PG, Ranish JA, Hahn S. 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. 2004; 24(4): 1709–1720. https://doi.org/10.1128/MCB.24.4.1709-1720.200

25. Akashi S, Nagakura S, Yamamoto S, Okuda M, Ohkuma Y, et al. Structural characterization of human general transcription factor TFIIF in solution. Protein Sci. 2008; 17(3): 389-400. https://doi.org/10.1110/ps.073258108

26. Chen ZA, Jawhari A, Fischer L, Buchen C, Tahir S, et al. Architecture of the RNA polymerase II–TFIIF complex revealed by cross-linking and mass spectrometry. EMBO J. 2010; 29(4): 717–726. https://doi.org/10.1038/emboj.2009.401

27. Gaiser F, Tan S, Richmond TJ. Crystal structure of the RAP30/74 interaction domains of human TFIIF. 2000. https://doi.org/10.2210/pdb1f3u/pdb

28. Schilbach S, Aibara S, Dienemann C, Grabbe F, Cramer P. Yeast RNA polymerase II transcription pre-initiation complex with closed distorted promoter DNA. 2021. https://doi.org/10.2210/pdb7o73/pdb

29. Patel AB, Greber BJ, Nogales E. Recent insights into the structure of TFIID, its assembly, and its binding to core promoter. Current Opinion in Structural Biology. 2020; 61: 17-24. https://doi.org/10.1016/j.sbi.2019.10.001

30. Eisenmann DM, Dollard C, Winston F. *SPT15*, the gene encoding the yeast TATA binding factor TFIID, is required for normal transcription initiation in vivo. Cell. 1989; 58(6): 1183-1191. https://doi.org/10.1016/0092-8674(89)90516-3

31. Akhtar W, Veenstra GJC. TBP-related factors: a paradigm of diversity in transcription initiation. Cell & Bioscience. 2011; 1: 23. https://doi.org/10.1186/2045-3701-1-23

32. Tolić-Nørrelykke SF, Rasmussen MB, Pavone FS, Berg-Sørensen K, Oddershede LB. Stepwise Bending of DNA by a Single TATA-Box Binding Protein. Biophys J. 2006; 90(10): 3694–3703. https://doi.org/10.1529/biophysj.105.074856

33. Nikolov DB, Chen H, Halay ED, Hoffmann A, Roeder RG, et al. Human TBP core domain complexed with DNA. 1996. https://doi.org/10.2210/pdb1cdw/pdb

34. Kim Y, Geiger JH, Hahn S, Sigler PB. Crystal structure of a yeast TBP/TATA-box complex. 1995. https://doi.org/10.2210/pdb1ytb/pdb

35. Kolesnikova O, Ben-Shem A, Luo J, Ranish J, Schultz P, et al. Molecular structure of promoter-bound yeast TFIID. Nature Communications. 2018; 9: 4666. https://doi.org/10.1038/s41467-018-07096-y

36. Chymkowitch P, Nguéa AP, Aanes H, Koehler CJ, Thiede B. Sumoylation of Rap1 mediates the recruitment of TFIID to promote transcription of ribosomal protein genes. Genome Res. 2015; 25(6): 897-906. https://doi.org/10.1101/gr.185793.114

37. Rosonina E, Duncan SM, Manley JL. SUMO functions in constitutive transcription and during activation of inducible genes in yeast. Genes Dev. 2010; 24(12) :1242-52. https://doi.org/10.1101/gad.1917910

38. Delgado ML, O'Connor JE, Azorín I, Renau-Piqueras J, Gil ML, et al. The glyceraldehyde-3-phosphate dehydrogenase polypeptides encoded by the *Saccharomyces cerevisiae* *TDH1*, *TDH2* and *TDH3* genes are also cell wall proteins. Microbiology. 2001; 147(2). https://doi.org/10.1099/00221287-147-2-411

39. Esteras M, Liu IC, Snijders AP, Jarmuz A, Aragon L. Identification of SUMO conjugation sites in the budding yeast proteome. Microbial Cell. 2017; 4(10): 331-341. https://doi.org/10.15698/mic2017.10.593

40. Boyer-Guittaut M, Birsoy K, Potel C, Elliott G, Jaffray E, et al. SUMO-1 Modification of Human Transcription Factor (TF) IID Complex Subunits: inhibition of TFIID promoter-binding activity through SUMO-1 modification of hsTAF5. Journal of Biological Chemistry. 2005; 280(11): 9937-9945. https://doi.org/10.1074/jbc.M414149200

41. Akhtar A, Rosonina E. Chromatin Association of Gcn4 Is Limited by Post-translational Modifications Triggered by its DNA-Binding in *Saccharomyces cerevisiae*. Genetics. 2016; 204(4): 1433-1445. https://doi.org/10.1534/genetics.116.194134

42. Khaperskyy DA, Ammerman ML, Majovski RC, Ponticelli AS. Functions of *Saccharomyces cerevisiae* TFIIF during Transcription Start Site Utilization. Mol Cell Biol. 2008; 28(11): 3757-3766. https://doi.org/10.1128/MCB.02272-07

43. Zeng PY, Vakoc CR, Chen ZC, Blobel GA, Berger SL. In vivo dual cross-linking for identification of indirect DNA-associated proteins by chromatin immunoprecipitation. Biotechniques. 2018; 41(6): 694-698. https://doi.org/10.2144/000112297

44. Talia SD, Skotheim JM, Bean JM, Siggia ED, Cross FR. The effects of molecular noise and size control on variability in the budding yeast cell cycle. Nature. 2007; 448(7156): 947-51. https://doi.org/10.1038/nature06072

45. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Statist. Soc. B. 1995; 57(1): 289-300. https://doi.org/10.2307/2346101

46. Wang Z, Jones GM, Prelich G. Genetic Analysis Connects SLX5 and SLX8 to the SUMO Pathway in *Saccharomyces cerevisiae*. Genetics. 2006; 172(3): 1499–1509. https://doi.org/10.1534/genetics.105.052811

47. Esteras M, Liu IC, Snijders AP, Jarmuz A, Aragon L. Identification of SUMO conjugation sites in the budding yeast proteome. Microbial Cell. 2017; 4(10): 331-341. https://doi.org/10.15698/mic2017.10.593

48. Lanz MC, Yugandhar K, Gupta S, Sanford EJ, Faça VM, et al. In-depth and 3-dimensional exploration of the budding yeast phosphoproteome. EMBO Rep. 2021; 22: e51121. https://doi.org/10.15252/embr.202051121

49. Swaney DL, Beltrao P, Starita L, Guo A, Rush J, et al. Global analysis of phosphorylation and ubiquitylation cross-talk in protein degradation. Nature Methods. 2013; 10: 676-682. https://doi.org/10.1038/nmeth.2519

50. Herrmann C, Avgousti DC, Weitzman MD. Differential Salt Fractionation of Nuclei to Analyze Chromatin-associated Proteins from Cultured Mammalian Cells. Bio Protoc. 2017; 7(6): e2175. https://doi.org/10.21769/BioProtoc.2175

51. Cuevas-Bermúdeza A, Garrido-Godinoa AI, Navarro F. A novel yeast chromatin-enriched fractions purification approach, yChEFs, for the chromatin-associated protein analysis used for chromatin-associated and RNA-dependent chromatin-associated proteome studies from *Saccharomyces cerevisiae*. Gene Reports. 2019; 16: e100450. https://doi.org/10.1016/j.genrep.2019.100450

52. Newman HA, Meluh PC, Lu J, Vidal J, Carson C, et al. A high throughput mutagenic analysis of yeast sumo structure and function. PLoS Genet. 2017; 13(2): e1006612. https://doi.org/10.1371/journal.pgen.1006612

53. Horigome C, Bustard DE, Marcomini I, Delgoshaie N, Tsai-Pflugfelder M, et al. PolySUMOylation by Siz2 and Mms21 triggers relocation of DNA breaks to nuclear pores through the Slx5/Slx8 STUbL. Genes Dev. 2016; 30(8): 931–945. https://doi.org/10.1101/gad.277665.116

54. Panavas T, Sanders C, Butt TR. SUMO fusion technology for enhanced protein production in prokaryotic and eukaryotic expression systems. Methods Mol Biol. 2009; 497: 303-317. https://doi.org/10.1007/978-1-59745-566-4_20

55. Hajian R, Balderston S, Tran T, DeBoer T, Etienne J, et al. Detection of unamplified target genes via CRISPR–Cas9 immobilized on a graphene field-effect transistor. Nature Biomedical Engineering. 2019; 3: 427-437 https://doi.org/10.1038/s41551-019-0371-x