The Contribution of Alternative Splicing Toward the Global Control of SUMO1 SUMOylation

Myriah Lorraine Acuna

University of Texas at El Paso, mlacuna@miners.utep.edu

Follow this and additional works at: https://digitalcommons.utep.edu/open_etd

Part of the Biology Commons

Recommended Citation
https://digitalcommons.utep.edu/open_etd/392

This is brought to you for free and open access by DigitalCommons@UTEP. It has been accepted for inclusion in Open Access Theses & Dissertations by an authorized administrator of DigitalCommons@UTEP. For more information, please contact lweber@utep.edu.
THE CONTRIBUTION OF ALTERNATIVE SPICING TOWARD THE GLOBAL CONTROL OF SUMO1 SUMOYLATION

MYRIAH LORRAINE ACUÑA

Master’s Program in Biological Sciences

APPROVED:

________________________________________________________________________
German Rosas-Acosta, Ph.D., Chair

________________________________________________________________________
Manuel Llano, Ph.D.

________________________________________________________________________
Chuan Xiao, Ph.D.

________________________________________________________________________
Charles Ambler, Ph.D.
Dean of the Graduate School
THE CONTRIBUTION OF ALTERNATIVE SPLICING TOWARD THE GLOBAL CONTROL OF SUMO1 SUMOYLATION

by

MYRIAH LORRAINE ACUÑA, B.S.

THESIS

Presented to the Faculty of the Graduate School of The University of Texas at El Paso in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

Department of Biological Sciences THE UNIVERSITY OF TEXAS AT EL PASO

December 2017
ACKNOWLEDGEMENTS

This journey has not been easy by any means necessary, however without the help and guidance of the many people in my life, it simply would not have been possible and nothing that comes easy is ever worth having. My first appreciation is directed towards my mentor, Dr. German Rosas-Acosta, for not only his incredible knowledge, expertise and simple fascination for all the scientific disciplines which has continued to inspire my own; but also, his attitude and outlook on life which has given me hope even on the darkest days. Secondly, I would like to thank my committee: Dr. Manuel Llano for simply wanting to help me achieve this goal by any means which is quite refreshing and has meant the world to me. Dr. Xiao, for taking this project on without hesitation and together giving me truly valuable advice on the directions I should take.

Another enormous thank you to the UTEP SUMO/Influenza laboratory and all who work there for the unfailing help and contribution of statistical analysis: Carlos Ontiveros, the laughs, encouragement and understanding: Arely Diaz and Sullenly Medina, and lastly, Amie Kern for having my back and the continuous support. I would also like to extend a thank you to Miguel Beltran, who patiently taught me most of what I know, who always lends a helping hand in whatever materials I need, who tells me exactly what I need to hear even if it’s not what I want to hear, and the dire editing of this thesis. I would also like to thank Cynthia Carreon, we finished our undergraduate journey together, embarked on graduate school together and are now finishing together as well. This journey would be truly lonely without your help and light.

Without the continuous love and support of one person in particular, none of this would be possible and that is my father. He’s been my biggest fan and always encourages me to aim for the stars, reminding me that anything is possible with pure focus, commitment and sheer will. My friends have meant the world to me whom I truly regard as my family: Yuridia, Amanda and Daniel who always lend a listening ear, give nothing but constant positivity and never let me forget how much they truly believe in me. Lastly, I would like to thank my coaches: Cristobal and Andy, as well as all of my team mates at 10th Planet Jiu-Jitsu, you have helped me accomplish this in more ways than you will ever know by simply showing me no situation is permanent and perseverance conquers all.
ABSTRACT

SUMOylation, the attachment of the Small Ubiquitin-like Modifier (SUMO) to a target protein, is a post-translational modification similar to that of Ubiquitination in terms of the cascade of events required to produce the final modified target. However, instead of targeting proteins for degradation, as Ubiquitination usually does, SUMOylation appears to regulate many vital cellular processes including nucleocytoplasmic transport, transcription, apoptosis, protein stability, response to stress, and progression through the cell cycle (Hay, R.T, 2005). This versatility exhibited by the SUMOylation system makes it an optimal target for viral manipulation, as our laboratory has previously described for Influenza A virus (Pal, S., Rosas-Acosta, G., 2009). It is presently acknowledged that the human genome contains at least five clearly distinct functional SUMO genes, namely SUMO1, 2, 3, 4 and 5 as well as various SUMO pseudogenes. Importantly, analysis of Homo sapiens SUMO1 sequences obtained from the NCBI database indicate the existence of three different processed mature mRNA transcripts for the SUMO1 gene, produced by alternative splicing, hereafter referred to as SUMO1 variants (var). While var1 and var2 are predicted to code for the protein typically thought of as SUMO1, var3 lacks exon 2 coding for an alternative SUMO1 isoform lacking a relatively short stretch of amino acids toward the N-terminus; this isoform will hereafter be referred to as SUMO1α. After performing an extensive review of published literature, we could not find any reports describing the existence of alternative SUMO1 isoforms. Furthermore, the existence of different mature mRNAs for SUMO1, while strongly supported by available RNA-Seq data, has not been actually reported in the primary scientific literature either. Hence, experimental confirmation of the existence of both, the different SUMO1 transcripts and the alternative SUMO1 isoform will be required to achieve a more thorough understanding of the molecular biology of SUMO1 and the mechanisms governing SUMO1 levels within the cell. Given the known “cyto-protective” properties of protein SUMOylation (Luo, J., Guo, C., 2017), an improved understanding of the contribution of such post-translational regulatory mechanisms to the global control of cellular SUMOylation may lead to the development of innovative therapies against conditions in which the cyto-protective role of protein SUMOylation plays a critical function in
host survival, including infectious diseases such as Influenza A virus infection, and cardiovascular ischemic events such as strokes and heart attacks.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS........................................................................................................... iv  
ABSTRACT............................................................................................................................... v  
TABLE OF CONTENTS.............................................................................................................. vii  
LIST OF TABLES....................................................................................................................... viii  
LIST OF FIGURES.................................................................................................................... ix  
Chapter 1: INTRODUCTION........................................................................................................ 1  
  1.1 Sumo Proteins and Conjugation ...................................................................................... 2  
  1.2 Important Regulatory Roles of SUMO........................................................................... 3  
  1.3 The Effects of SUMOylation within the Cell................................................................... 6  
  1.4 SUMO Gene Expression and Regulation....................................................................... 8  
Chapter 2: AIMS........................................................................................................................ 12  
Chapter 3: METHODS AND MATERIALS.............................................................................. 14  
  3.1 Primer Design Approach................................................................................................. 14  
  3.2 Cell Culture.................................................................................................................... 15  
  3.3 RNA Purification............................................................................................................ 16  
  3.4 Assessment of Purified RNA Quality and Quantity....................................................... 17  
  3.5 cDNA Synthesis............................................................................................................ 18  
  3.6 qRT-PCR: Two Step vs One Step Methods................................................................... 18  
  3.7 Cloning of SUMO1 Variants........................................................................................... 20  
  3.8 Transformation............................................................................................................... 20  
  3.9 Mini Prep......................................................................................................................... 21  
  3.10 Maxi Prep...................................................................................................................... 22  
  3.11 T7 RNA Polymerase In Vivo Transcription................................................................... 23  
  3.12 Generating Calibration Curves..................................................................................... 24  
Chapter 4: RESULTS.................................................................................................................. 25  
  4.1 Preliminary Data.............................................................................................................. 25  
  4.2 Final Data Collection Using Calibration Curves............................................................. 28  
Chapter 5: CONCLUSIONS...................................................................................................... 35  
Chapter 6: DISCUSSION........................................................................................................... 36  
Chapter 7: STRATEGIES AND PITFALLS............................................................................ 37  
Chapter 8: REFERENCES......................................................................................................... 38  
Chapter 9: VITA....................................................................................................................... 40
LIST OF TABLES

Table 3.6: Comparison of Running Conditions for the One-step vs Two-step qRT-PCR Methods.................19
LIST OF FIGURES

Figure 1.4: Three Different Proteins Produced through Alternative Splicing.................................11

Figure 3.1: SUMO1 Variants Detection Approach.................................................................14

Figure 4.1: First RNA Purification of A549 and HEK293A Cell Lines...........................................26

Figure 4.2 First Confirmation of SUMO1 Variants using RT-PCR................................................27

Figure 4.3: First Round of qRT-PCR of HEK293A and A549 Cell Lines.....................................27

Figure 4.4: Agarose Gel of RT-qPCR products Generated Ran in Triplicates...............................28

Figure 4.5: Agarose Gel Electrophoresis Confirmation of the Cloned SUMO1 Variants...............30

Figure 4.6: Panel A-Confirmation of pJET1.2 SUMO1 Variant Plasmids.....................................31

Figure 4.7: Calibration Curve for all Three SUMO1 Variants (Var).........................................32

Figure 4.8: Gel Electrophoresis Confirmation of all SUMO1 Variants Calibration Curves.........33

Figure 4.9: Quantitative One-step RT-PCR of all Three SUMO1 Variants..............................34
1. INTRODUCTION

1.1 SUMO Proteins and Conjugation

SUMO, the (Small Ubiquitin-like Modifier) protein and Ubiquitin have similar 3-D structure and undergo similar cascades of events, ending in their post-translational conjugation to a plethora of protein targets. However, unlike Ubiquitination, which frequently targets proteins for degradation, SUMOylation regulates a large number of vital cellular processes such as nucleocytoplasmic transport, transcription, apoptosis and the detection and repair of DNA damage (Dohmen, 2004; Geiss-Friedlander and Melchior, 2007). All SUMO proteins from yeast to human share the conserved Ubiquitin domain and the C-terminal di-glycine cleavage/attachment site (Laoong-u-thai, Y., Yang, J., 2009). The most prominent difference, however, between the SUMO proteins and Ubiquitin is the presence of a relatively large unstructured variable N-terminal extension in the SUMO proteins; this region ranges in length between 11 and 35 amino acids among the different SUMO proteins (Garaeu, J., Lima, C., 2011). Its unstructured nature suggests that it likely contributes to mediating specific protein–protein interactions (PPI) (Garaeu, J., Lima, C., 2011).

While many cellular proteins are covalently modified by SUMO1/2/3 proteins (Su, HL., Li, SS., 2002), the various SUMO proteins are predominantly localized to different locations such as the nuclear membrane, nuclear bodies, and cytoplasm, respectively (Su, HL., Li, SS., 2002). Although such localization is likely to correspond primarily to conjugated SUMO species, it should be noted that some free form of SUMO1/2/3 proteins can also be detected within the cell (Su, HL., Li, SS., 2002). Plants can synthesize up to 12 different SUMO forms and are considered to be the “master SUMOylators.” Plant SUMOylation is documented to regulate flowering, the bulbing of roots in germinating seeds, vascular development, as well as
tolerance to salt, pH, and humidity (Park, HJ., Dae-Kin, Y., 2011). For example, eight SUMO isoforms are encoded by the genome of the model plant *Arabidopsis Thaliana* (Dohman, R., 2004). Interestingly, similar to mammals, *Arabidopsis’ SUMO1 and SUMO2 conjugation is induced when subjected to heat stress (Kurepa, J., Vierstra, R., 2003)

As with many Ubiquitin and Ubiquitin-like proteins, all eukaryotic SUMO proteins are initially translated as immature precursors that require a protease to cleave them in order to produce the mature form (Gareau, J., Lima, C., 2010). This mature form contains a C-terminal diglycine motif that is necessary for adenylation by a SUMO E1(activation) enzyme (Gareau, J., Lima, C., 2010). The adenylate is then attacked by an E1(activation) enzyme and then transferred to a SUMO E2(conjugating) enzyme also known as UBC9 (Gareau, J., Lima, C., 2010). The enzyme can either interact directly with some substrates to transfer the SUMO to acceptor Lysine residues, or E3 SUMO (ligases) proteins can facilitate this activity (Gareau, J., Lima, C., 2010). The actual conjugation process frequently occurs through a SUMOylation consensus site (CS) which consists of a motif of residues surrounding the modified lysine within the substrate. This motif conforms to the sequence “ΨKXE,” where Ψ is a hydrophobic residue, K is the lysine site in which the actual conjugation occurs, X is any amino acid, and E is glutamic acid (Sampson, D., Matunis, M., 2001). It has been shown that for some substrates, the SUMO CS can function as the major determinant of Ubc9 binding and SUMO1 modification, as the mutation of residues within the SUMO CS resulted in a complete decrease of both Ubc9 binding and substrate modification (Sampson, D., Matunis, M., 2001). These findings revealed important implications for SUMO substrate recognition and ultimately the transfer to specific lysine residues (Sampson, D., Matunis, M., 2001).
While all SUMO1/2/3 proteins utilize the same conjugation pathway, the SUMO1 protein exhibits a mere 44% sequence identity with SUMO2 and SUMO3 proteins. In sharp contrast, SUMO2 and SUMO3 proteins share 86% sequence identity (Su, HL., Li, SS., 2002). Phylogenetic analyses indicate that the SUMO3 gene was derived from the SUMO2 gene (Su HL., Li SS., 2002), with evidence now indicating the existence of SUMO4 and SUMO5 proteins. SUMO4 is encoded by an intron within the human TAB2 gene (Bohren, K., Owerbach, D., 2004), and is expressed differently in various organs and tissues, with strongest expression occurring within kidney cells (Dohman, R., 2004) Importantly, SUMO2, SUMO3, and SUMO4 contain a SUMOylation consensus site within their primary structure, which allows them to become SUMOylated, thus triggering the formation of SUMO chains. In contrast, the SUMOylation consensus sequence is not present in SUMO1 (Dohman, R., 2004), which allows SUMO1 to act as a chain terminator \textit{in vitro} and \textit{in vivo} (Bohren, K., Owerbach, D., 2004).

1.2 Important Regulatory Roles of SUMO

Although SUMOylation is tied to numerous vital cellular processes, its effects upon transcriptional events appears to be one of the most critical, particularly, SUMO’s role in dictating what gets transcribed, ultimately acting as a transcriptional regulator. The system in which chromatin is packaged within the nucleosome within the four histone arrangements undeniably indicates there are “higher-order patterns that contribute to DNA’S sophisticated structure (Potts, C., Matiunis, M., 2016).” These “higher-order structures” have indicated the critical roles posttranslational protein modifications play in modulating DNA’s availability, in particular, how regulation of SUMO proteins affects chromatin structure (Potts, C., Matiunis, M., 2016). SUMO modification on chromatin structures have shown to effect transcriptional
activation and repression, DNA replication/repair, and influences on chromosome segregation (Golebiowski et al., 2009; Makhnevych et al., 2009).

SUMOylation has acquired a repressive reputation in some cases even causing a competitive/antagonistic effect on other post translational modifications such as acetylation. Such was the case in S. cerevisiae, in which all four core histones demonstrated SUMOylation detection (Nathan, D., Berger, S., 2006), specifically there are SUMOylated sites in H2A, H2B, and H4 which revealed a correlation between histone SUMOylation and transcriptional repression (Nathan, D., Berger, S., 2006). The results further insinuate histone SUMOylation exists in a dynamic state with acetylation and ubiquitination through competitive post-translational modification of specific lysine residues: positions 6, 7, 16 and 17 of both histones H3 and H4 (Nathan, D., Berger, S., 2006), typically thought of as novel acetylation and SUMOylation target sites that can only be occupied by one PTM at a time as determined by tandem mass spectrophotometry MS/MS (Nathan, D., Berger, S., 2008). While this is just one example of SUMO exerting a repressive transcriptional effect, it is incorrect to assume repression is the only role played by SUMO on chromatin. In fact, SUMO is now predicted to exert a repressive effect on only 60% of genes, while having no effect on 20% of genes and, even more interestingly, upregulating 20% of the rest of the genome. For example, SUMOylation of the 12th lysine residue in H4 by SUMO3 (Shiio and Eisenman, 2003) actually inhibits chromatin compaction as well as oligomerization (Dhall, A., Chetterjee, C., 2014), producing an open chromatin structure that enhances transcription thus increasing gene expression.

There are also slightly greater SUMOylation levels present at sub-telomeric regions compared to the internal regions of chromosomes (Nathan, D., Berger, S., 2008). A double-
affinity ChIP method used to infer sequences adjoining telomeric repeats demonstrating that regions within the right arm of chromosome VI and within the left arm of chromosome III are two times more enriched in H2B SUMOylation than other regions within their respective chromosomes (Nathan, D., Berger, S., 2008). Interestingly, the SUMO levels decreased as the ChIP probe came closer to the internal regions of chromosomes, suggesting a complete opposite profile compared to that of Ubiquitination (Nathan, D., Berger, S., 2008). Other, separate ChIP experiments involving exogenously expressed yeast SUMO-histone proteins also revealed enrichment at sub-telomeric regions (Nathan et al., 2006; Xhemalce et al., 2004; Zhao and Blobel, 2005). Thus, by serving as a chromatin protein recruitment signal for repressive transcription factors, histone SUMOylation not only coordinates, but integrates stimulatory signals to catalyze an extremely complex network of gene regulation events, as exemplified by telomeric silencing, in which SUMO helps the cell to maintain basal levels of transcription.

In addition to protein conjugation and the post-translational modifications of histones, SUMO has also been associated with the processing of pre-mRNA transcripts as seen by its localization to proteins both bound to promoters and on DNA-encoding exons of extremely active genes (Nuro-Gyina, P., Parvin, J., 2015). SUMO1 modification inhibition has also been proved to reduce the initiation of mRNA synthesis as well as splicing (Srikanth, C., Verma, S., 2017). In contrast to when factors associated with the promoters of highly active genes are SUMOylated, they then trigger the formation of an RNA polymerase II complex, influencing mRNA processing. Thus, SUMOylation may be the glue that binds together transcription initiation, elongation, and splicing (Nuro-Gyina, PK., Parvin, JD., 2016). Of even more relevance, SUMOylation of chromatin bound to promoters is associated with the epigenetic mark H3K4me3 found downstream of the transcription start site. The H3K4me3 mark is bound
by CHD1, which in turn stimulates RNAPII elongation and recruits the splicing machinery to the chromatin (Reinberg, D., Sims, R., 2006). Thus, SUMO itself may be considered an epigenetic mark based not only on its modification of the nucleosomal histones, but also the effects exerted on mRNA transcript processing.

1.3 The Effects of SUMOylation Within the Cell

There is much evidence now for SUMO engaging in a cyto-protective role when cells experience stress. For example, stress conditions such as arsenic poisoning, hibernation, heat shock as seen with the upregulation of certain SUMO isoforms in *Aradopsis Thaliana*, ischemic events such as strokes and heart attacks, as well as Influenza A viral infections have all been linked with massive increases in cellular SUMOylation.

Arsenic poisoning, for example, is unique in that it was proven to induce a SUMO1 modification of nucleolin at the 294th lysine residue. Nucleolin is a protein expressed ubiquitously that participates in processes such as cell-cycle regulation and ribosomal biogenesis (Zhang, D., Huang, C., 2015). Nucleolin SUMOylation then causes it to bind and stabilize GADD45α mRNA to increase its nuclear localization eventually causing GADD45α-mediated cell death (Zhang, D., Huang, C., 2015). Normal cellular nucleolin activity is regulated by post-translational modifications including phosphorylation (Olsen, M., 2011), methylation (Olsen, M., 2011), ADP-ribosylation (Olsen, M., 2011), and now, SUMOylation (Zhang, D., Huang, C., 2015). Thus, it is likely that nucleolin SUMOylation is important for regulating specific stress-induced mRNA stabilization (Zhang, D., Huang, C., 2015). As previously explained, genes exhibiting the highest level of expression (i.e. ribosomal proteins and translation factors) contain an extremely high level of SUMO1 modification (Kahn, H., Dejean, A., 2013).
The cellular SUMOylation system has also been established as an essential viral factor for the Influenza A life-cycle (Boggio and Chiocca, 2006; Rosas-Acosta and Wilson, 2004), in particular, SUMO is important for viral regulation of replication and gene expression during infections. It has been established that many viral proteins are modified and regulated by the host cell’s SUMOylation coordination, and logically a high likelihood the virus in turn modulates the system’s activity. (Santos, A., Rosas-Acosta, G., 2013). There are at least 4 different Influenza A viral proteins established as SUMO substrates in vivo (Santos, A., Rosas-Acosta, G., 2013): NS1, PB1, NP and M1 (Pal, S. Rosas-Acosta, G., 2009). Given the variety of viral proteins that are SUMOylated, viral infections subsequently affect and alter the cellular SUMOylation system to enhance these viral proteins to gain functions they wouldn’t normally have (Rosas-Acosta, G., Wilson, V. 2005). Such is the case for Human Papilloma Viruses (HPV) in which SUMO contributes to the full functionality of the helicase (Rangasamy, D., Wilson, V., 2000). Thus, molecules/compounds that modulate the cellular SUMO system may act as potentially effective Influenza antivirals.

Influenza A virus infections are known to trigger a “global increase in cellular SUMOylation (GICS)” (Pal, S., Rosas-Acosta, G., 2011), however it is important to understand if this GICS is exclusive of immortalized cell lines (manuscript under preparation) because if the GICS does take place in vivo, then it’s ubiquitous to other systems as well. Initially, it was proposed that interferon alone was responsible for the increase in SUMOylation, however, our laboratory was able to prove the viral non-structural protein, NS1, stimulated the global increase in SUMOylation (manuscript under preparation). While the viral protein NS1 triggers GICS, viral pathogenicity appears inversely related to NS1’s ability to trigger GICs. Therefore, specific characteristics of the NS1 protein, in particular, the ability to trigger GICS are
extremely important. However, nothing is known about what triggers the increase at the molecular level.

1.4 SUMO Gene Expression and Regulation

The SUMO1 gene is 32,500 bps in length and it is located on chromosome 2. It contains five exons interrupted by four introns, while SUMO2/3 transcripts contain only three introns (Su, H., Li, S., 2002). SUMO1 mRNA appears to be most abundant in human epithelial derived HeLa, kidney derived HEK293 and Neuronal NT2 cells, while SUMO3 mRNA appears less abundant than SUMO2 mRNA, specifically in HeLa and HEK293 cells (Su, H., Li, S., 2002). Genome-wide analysis of SUMO1 conjugation to chromatin-associated proteins revealed a striking enrichment of the SUMO1 mark on exons and no enrichment on introns (Neyret-Kahn, H., Dejean, A., 2013). Therefore, it is likely that the SUMO1-conjugated proteins associated with exons may act as splicing factors (Nuro-Gyina, P., Parvin, JD., 2016). The mechanism in which the exons are recognized at the DNA level, however, is unclear. Essential splicing factors, such as serine-arginine rich (SR) proteins are recognized as containing SUMO conjugation sites (Twyffels, L. Kruys, V., 2011). The SR protein SF2/ASF interacts with UBC9 and enhances SUMOylation of specific substrates as well as promoting SUMO conjugation to RNA processing factors (Pelisch, F., Srebrow, A., 2010) serving as an important regulator mechanism with implications for potential targets of drug therapy.

Changes in the level of cellular SUMO conjugation are thought to be related to and are not believed to be associated to meaningful changes in the total amount of SUMO present in the cell. Increased SUMOylation frequently occurs during health-threatening situations. In one study, massive increases in SUMOylation were produced in primary cortical neuronal cells upon the induction of ischemic conditions (Lee, Y. Hallenbeck, J., 2009). Such is the case for
stroke victims whose brain cells also experience a dramatic increase in cellular SUMOylation. Thus, increasing SUMOylation could initiate a form of protection from the damage induced by hypoxia and other ischemic events. Interestingly, hibernating mammals, which experience a decrease in cardiac frequency that under normal conditions would be lethal for the animal, exhibit a dramatic increase in cellular SUMOylation in the most critical organs for survival (Quinones, Q, Podgoreanu, M., 2014). Hibernation torpor is a model in which there is a profound tolerance to reduced blood flow and oxygen delivery to the brain with links to massively increased amounts of SUMOylation. Thus, understanding SUMO’s regulation is now urgent more than ever given its protective function due to its pertinence to the survival of ischemic events that fully depend on a global increase in cellular SUMO events including Influenza A viral infection and hypoxic conditions such as heart attacks and strokes.

The alignment of H. sapiens mRNA sequences acquired from the NCBI database result in three different forms of mature mRNA transcripts for SUMO1. Var1 contains all exons, the fifth exon within var2 is spliced into two different fragments designated here as exon5 and exon 5’. Of the three variants, var1 and var2 code for what is considered to be the prototypical SUMO1 protein. Var3 is of particular interest because the second exon is spliced out, thus the mature mRNA codes for a substantially shorter version of the protein that lacks a considerable portion of the N-terminal region of the protein. Whether this variant form of the transcript is produced at all, and whether it might contribute a slightly different function from that of the "normal" SUMO1 in the cell, is not known. Preliminary data suggest the three alternatively spliced variants are produced: var2 being the most abundant while var1 and var3 are produced at relatively equal but lower levels. The production of var3 is relevant due to a lack of residues
(5-29) resulting in a truncated form of the protein, thus, yielding a different shape feasibly designating a difference in function.

It is surprising that although evidence of the existing SUMO1 variants is strongly supported, there has been no specific investigation of these variant levels within the cell, as well as whether the variants and isoforms are expressed. While miRNA targets that reside within the 3’ UTR are proposed to regulate the SUMO1 gene. It is then critical to know if these variant levels are present in both immortalized cells and live tissue as well as whether the variant levels change during certain stress conditions to help us achieve our final goal of understanding how SUMOylation is regulated within the cell, particularly the abundance of SUMOylation and the modulator: SUMO.

There are many genes within the human genome known to be under the regulation of alternative splicing, in fact, it’s been discovered that 92-94% of human genes actually undergo alternative splicing which can vary amongst different tissues as well as between individuals (Wang, E., Curge, C., 2008). Up until recently, alternative splicing was considered a rare event in the expression of genes, however, now of the 92-94% genes that do undergo alternative splicing, it is predicted that 86% of those genes also produce a minor isoform (Wang, E., Burge, C., 2008). There are major housekeeping genes under the regulation of alternative splicing including telomerase, actin, and tubulin. It is peculiar that the telomeric ends of chromosomes not only contain an increase in SUMOylation sites, but the telomerase gene itself is under the regulation of alternative splicing.

Given what is now known about how common alternative splicing is to produce isoforms and the RNA-seq data available on the NCBI database, it is a logical summation that alternative splicing may be a potential candidate for the regulation of the SUMO1 gene,
ultimately regulating itself through a negative feedback mechanism (based on the existence of different SUMO1 variants) which is the result of multiple proteins being coded from the same gene (figure 1). Thus, it is hypothesized here that alternative splicing not only contributes to the regulation of SUMO1 expression, but also contributes to the upregulation of the expression of this group of cellular proteins during specific stress conditions: heat shock, Influenza A virus infections, and ischemic events such as strokes and heart attacks engaging in a time specific cyto-protective role.

Figure 1.2: Three different Proteins Produced through Alternative Splicing
2. AIMS

To achieve our long-term goal of understanding the important regulatory roles played by the expression of SUMO1 to specifically exert cyto-protective effects on stress, it is critical to develop a more detailed understanding of the molecular mechanisms contributing to the SUMO1 gene expression regulation. To test our hypothesis of alternative splicing regulating the SUMO1 gene in the pursuit of the long-term goal, two specific aims will be carried out 1) Confirmation of the existence of the three alternatively spliced variant forms of the mature SUMO1 transcript (mRNA). RNA-seq data obtained from the NCBI database indicate the existence of three alternatively spliced forms of the SUMO1 gene. Therefore, experimental confirmation of the existence of the different transcripts and the potential demonstration of the production of an alternate isoform of SUMO1 in the cell will be pursued by developing a set of primers to amplify individually each specific mRNA variant using qRT-PCR, followed by final confirmation by DNA sequencing. 2) Determine relative abundance of the variants to determine if there is an overall preferred variant. Once experimental confirmation is obtained for all three variants within immortalized cell lines, it is important to acquire a representation of the variants present throughout different tissues. We will then survey relative abundance of transcripts in different cell lines to fully assess transcript abundance from peripheral blood mononuclear cells purified using phycol gradients. Using both tissue culture cell lines as well as primary tissues to infer whether the same preference observed in tissue culture cell lines is also true in vivo. Once these two goals are achieved, future studies will determine if variant levels change and/or fluctuate during stress conditions. During stress conditions, there may be a change in the abundance of the transcripts and those changes might be associated with global increase cellular SUMOylation (GICS) levels. Using the experimental confirmation of these variants to further
deduce how cellular SUMOylation is regulated to ultimately treat: Influenza A viral infections, as well as hypoxic and heat shock conditions. Tissue culture cells will then be exposed to different conditions as triggers of cellular stress including: heat shock, hypoxia, low glucose levels and influenza A viral infection.
3. METHODS AND PROCEDURES

3.1 Primer Design Approach

An overall map of the RNA-seq data obtained from the NCBI database was first constructed using the three SUMO1 variant sequences to infer their distinctions. The map was drawn based on a series of alignments performed by ClustalW between the full genomic sequences for the SUMO1 gene located within Chromosome 2, as well as the different mature mRNA transcripts for SUMO1. The transcripts’ database accession numbers are: var1: NM_00352.4, var2: NM_001005781.1, var3: NM_001005782.1, SUMO1 gene: NG_011679.1. Afterwards, a more detailed map of each transcript was assembled to focus on the boundaries between exons to simplify the design of primers targeting exon/exon junctions (Figure 2).

**SUMO1 Variants Detection Approach**

![Diagram of SUMO1 Variants Detection Approach](image)

Figure 3.1: SUMO1 Variants Detection Approach- Representation of the primer approach to detect the three different SUMO1 variant spliced sequences obtained from the NCBI database.

The primers themselves were designed using the following guidelines: 1) Primers should have similar TM’s (melting temperatures) 2) the expected PCR products produced should be
between 200-300bp in length. 3) Primers should target exon/exon junctions. 4) There should be specific primer sets for every mature mRNA. 5) Any given primer pair should not amplify more than one mature mRNA. 6) Primers should contain as close as possible 50:50 [GC] : [AT] content with no more than 40:60 content either way. 7) Primers should end on a “clamping sequence” (CG, GC, GG, CC) and lastly, 8) primers should be free of sequences likely to form strong secondary structures. The primers designed also included a set of “Positive Control” primers (PC) designed to amplify all mature mRNAs for SUMO1, thus it is predicted that a fraction of every transcript may be correlated to the abundance of the other transcripts by comparing it to the products amplified by the positive control (PC) primers. Once designed and ordered from IDT (Integrated DNA Technologies), all of the primers were reconstituted in sterile TE following the volumes indicated to achieve a 100mM concentration. The 100mM primer stocks were mitigated once more into 1:10 dilution (20uL+180uL TE) aliquots to produce working solutions of 10uM concentrations for direct use in RT-PCR and qRT-PCR reactions.

3.2 Cell Culture

To test if the three SUMO1 Variants are indeed produced within mammalian cells, HEK293A and A549 cells were grown in tissue culture in DMEM media containing 10% Fetal Bovine Serum. Human embryonic kidney cells (HEK293A) are a suitable cell model due to their ease of growth, high transfection rates, and also produce many of the natural exogenous proteins. A549 cells are adeno-carcinomic human alveolar basal epithelial cells who can form an abundance of intercellular connections. The A549 cell lines are often used as an in vitro model as a Type II pulmonary epithelial cell model for drug metabolism as well as serving as an efficient transfection host. A third cell line type was incorporated to avoid any results being specific to that of immortalized cell lines, thus, Peripheral Blood Mononuclear Cells (PBMC)
were an ideal candidate being that these samples came directly from patients and contain nuclei, consisting of both lymphocytes (T cells, B cells, NK cells) and monocytes.

The immortalized cell lines were plated in 6-well plates at a concentration of 0.3x10^6 c/mL per well and demonstrated proper confluency roughly 34-hours post plating. The A549 and HEK293A cell lines were collected in groups of 3 wells per each sample for subsequent RNA purification. The PBM cells were directly isolated from patients and stored in cryovials awaiting RNA purification.

3.3 RNA Purification

Cells were washed in 1mL PBS and total RNA was purified using the Qiagen RNeasy Mini Kit® via the Qiashredder® method. To begin, 200uL of buffer RLT was added to each well and incubated for five minutes at 23 °C. The lysates from each of the three wells were combined into a single well, thus resulting in two samples of each cell line per plate. The samples were then transferred into a microfuge tube containing a QIAshredder® cartridge, mixed until there were no visible clumps, and centrifuged at full speed for two minutes at 23°C. Afterwards, one volume of diluted ethanol (70% EtOH) was added to the homogenized lysate followed by careful pipette re-suspension to ensure proper mixing. Up to 700uL of each sample was added to an RNeasy® spin column and centrifuged for thirty seconds at 8000 x g. The flow through was discarded and 700uL of Buffer RW1 was added to the RNeasy spin column, and again centrifuged for thirty seconds with a speed of 8000 x g at 23 °C. Next, 500uL of Buffer RPE was added and centrifuged for thirty seconds at a speed 8000 x g at 23 °C. Another 500uL of Buffer RPE was added but centrifuged for two minutes at 8000 x g to dry the cartridge. Subsequently, the cartridge was placed in a new 2mL collection tube and centrifuged for one minute at full speed, then placed into a new 1.5mL collection tube, eluted in 50uL of RNase-
free DEPC-treated H2O and centrifuged for 1 minute at 8000 x g. The eluted RNA was finally then aliquoted in 9uL and stored at -80°C.

3.4 Assessment of Purified RNA Quality and Quantity

To assess the quantity of the purified RNA a Qubit Fluorometer 3.0® was used, while formaldehyde-agarose gel electrophoresis was to determine the quality following these recipes and procedures: a 10x FA gel buffer consisting of 200mM 3-\{N-morpholino\} propane sulfonic acid (MOPS), 50mM Sodium Acetate, and 10mM EDTA whose pH was brought down to 7.0 using roughly 68mL of 1M-NaOH per one liter of 10x FA Buffer. A 1x FA gel running buffer could then be made using 100mL of the 10x FA gel buffer previously described, 74mL of 10% Formaldehyde and 826 mL of DEPC-H20. To load the gel while ensuring proper loading of the total purified RNA samples, we made a 4X RNA loading buffer consisting of: 20uL saturated bromophenol blue, 800uL of 500mM EDTA, pH 8.0, 2.664mL of 10% Formaldehyde, 2mL of 100% glycerol, 3,080 mL formamide, 4mL of 10X FA gel buffer and lastly 656uL of DEPC-H20. To produce an FA gel mix, we mixed 3.6g Agarose, 30mL 10x FA gel buffer, and 280mL of DEPC-H20. The contents were then heated up to a boil and the agarose was allowed to solubilize by cooling to 65°C. 20mL of 10% Formaldehyde was added to the mixture and finally stored at 65°C. It should be noted that problems arose using DEPC-H20 in both gel electrophoresis and qPCR reactions, we then switched to using QH20.

To Qubit Fluorometer quantification began by initially making two standards that generate a curve for the purified RNA to measure against. A working solution must first be made for both standards and samples by using 199uL of Qubit buffer with 1uL of Qubit Reagent. Both standard one and standard two use 190uL of the working solution combined with 10uL of Standard 1 and Standard 2. Once the standards were properly vortexed for two to three seconds,
the samples must sit for two minutes before the reading commences. To begin the assembly of reading the actual samples, 198uL of the working solution is combined with 2uL of the purified RNA followed by the mixing and settling of the samples.

3.5 cDNA Synthesis

Following the purified RNA extraction from both HEK293A and A549 cell lines, cDNA synthesis was performed using an NEB-M-MuLv® kit. This system required a total of 2ug of RNA added to 2uL of 50mM Poly-T primer, 1uL of 10mM DNTPs, and lastly qH20 to bring the reaction volume up to 10uL. The reaction was then heated for five minutes at 65°C and kept on ice while the other components were added: .5uL RNase inhibitor, 2uL 10x M-MuLv Buffer, 1uL M-MuLv Reverse transcriptase, and lastly qH20 to bring the reaction volume up to 20uL. The final steps of the cDNA synthesis included incubations at 42°C for one hour, followed by a cool down to 4°C. Negative control samples were produced using all of the ingredients minus the Poly-T enzyme and the Reverse transcriptase, qH20 was used in place of these components for the negative control reactions.

3.6 qRT-PCR: Two Step vs One Step Methods

To perform the initial PCR reactions, a two-step qRT-PCR method was first used, assuming the concentration of the cDNA produced was directly proportional to the amount of total RNA used as a template, meaning, 2ug of RNA in a final volume of 20uL is equal to 0.1ug/uL. We then proceeded to further dilute the cDNA in a 1:20 dilution (5uL of cDNA into 95uL qH20) totaling 5ug/uL cDNA template. The 50uL qPCR reactions consisted of assembling 20uL primer master mix (2uL of both forward and reverse primers within 16uL qH20), 5uL of the diluted template cDNA (5ug/uL) and lastly 25uL of DYNAMO SYBR Green® master mix. The
negative controls were compiled using the negative control samples from the cDNA synthesis lacking the Poly-T enzyme and the reverse transcriptase.

However, upon experiencing many contamination issues within the negative control samples using the two-step cDNA qRT-PCR method, it was decided to switch to the Bio-Rad iTaq™ Universal SYBR® Green One-Step Kit which completely avoids the need to synthesize cDNA altogether. Also, instead of using 2ug of RNA, this new method was able to carry out the reactions with a mere 100ng RNA coupled with 10uL reaction mix, 1uL of both forward and reverse primers, .5uL of reverse transcriptase, and lastly, water to bring the reaction volume up to 20uL. Negative controls were assembled using all components minus RNA template. Subsequent runs of both qRT-PCR methods were carried out using a MyGo thermocycler, along with MyGo programming ran on MAC software and machinery using the following running conditions:

<table>
<thead>
<tr>
<th>PCR Running Conditions</th>
<th>1-Step qRT-PCR</th>
<th>2-Step qRT-PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Hold Stage</td>
<td>50°C x 10’00”</td>
<td>------</td>
</tr>
<tr>
<td>Hold Stage</td>
<td>95°C x 3’00”</td>
<td>95°C x 7’00”</td>
</tr>
<tr>
<td>PCR Stage</td>
<td>95°C x 0’10”</td>
<td>95°C x 0’10”</td>
</tr>
<tr>
<td>Amplification Stage</td>
<td>60°C x 0’30” X 40 cycles</td>
<td>59°C x 0’30” X 44 cycles</td>
</tr>
<tr>
<td>Melt Hold</td>
<td>95°C x 0’15”</td>
<td>------</td>
</tr>
<tr>
<td>Melting</td>
<td>60°C x 0’60”</td>
<td>68°C x 10’00”</td>
</tr>
<tr>
<td>Meltlng</td>
<td>95°C x 0’01”</td>
<td>98°C x 0’0”</td>
</tr>
</tbody>
</table>

3.7 Cloning of SUMO1 Variants

Clones of each of the SUMO1 variants were generated first by using a Taq-Polymerase PCR reaction whose products were necessary to ultimately produce calibration curves in hopes of
detecting the exact copy number of each SUMO1 variant produced within the cell. The PCR reaction consisted of: 5µL cDNA, 2.5µL 10mM DNTP, 5µL Taq Polymerase Buffer, 2µL of each primer, 1µL of Taq Polymerase, and 32.5µL QH20 to bring the reaction volume up to 50µL. The running conditions were the same as in table one following the two step directions.

The PCR products generated were then subjected to restriction digests on ice using a Thermo-Scientific CloneJET PCR Cloning Kit®: 10µL of 2x Reaction Buffer, 1µL of each PCR product and 7µL of QH20 for a total of an 18µL reaction volume. The reaction was then vortexed briefly and centrifuged for three to five seconds, followed by an incubation at 70°C for five minutes. Finally, ligation reactions were compiled on ice by adding 1µL of the pJET1.2/blunt cloning vector® and 1µL of T4 DNA ligase to the 18µL blunt-end restriction digest reactions.

3.8 Transformation of SUMO1 Clones into NEB 10-Beta Competent E. Coli

Upon the ligation reaction setup, transformation of NEB 10-beta competent E. coli cells were then necessary. The transformation entailed thawing of the cells on ice for ten minutes followed by splitting the cell mixture into 2x 25µL aliquots. From this reaction, 1.5µL of each ligation reaction were then added to the cell mixture, carefully mixed, and incubated on ice for thirty minutes. Heat shock was performed for exactly thirty seconds at 42°C and the samples were then placed on ice again for five minutes. The NEB 10-beta/ Stable Outgrowth Media (SOC) was added to the mixture (450µL) and incubated at 37°C for sixty minutes at 250rpm.

Two different volumes of each SUMO1 variant mixtures were plated (75µL and 300µL, respectively) and incubated overnight at 37°C. To screen for the colonies with the correct SUMO1 variant sequences, a P200 pipette tip was used to pick five colonies of each variant and
placed into 300uL of H2O. Of the 300uL bacterial solution: 5uL was used to set up a replication plate, and another 5uL was used to set up another Taq Polymerase PCR reaction (Figure 7) for the screening process along with 2.5uL DNTP, 5uL Taq Polymerase Buffer, 2uL of the forward primer provided by the cloneJet PCR Cloning Kit, 2uL of the respective SUMO1 variant reverse primers, 1uL Taq Polymerase and 32.5uL of QH20 with the same running conditions as the two-step method listed in Table 1.

3.9 Mini-Prep Procedures

Of the fifteen colonies chosen for screening, six samples appeared to have the correct profile as determined via gel electrophoresis (Figure 7). However, sequencing analysis was needed for a complete confirmation. To achieve sequencing confirmation, the six chosen samples were first subjected to boiling lysis mini prep procedures. An overnight bacterial culture was inoculated from a single colony was grown in 5mL of sterile 2xYT + 10uL of ampicillin. After the overnight culture, the 5mL bacterial culture was split into two 1.5mL aliquots using two microfuge tubes. The remaining 1mL of the culture was transferred into a cyrovial with 110uL of bacterial freezing solution, vortexed at mid speed for ten seconds and stored at -80°C. The transferred 1.5mL of culture to the microfuge tubes were centrifuged for three minutes at full speed and the supernatant generated was suctioned out leaving only the pellet. Subsequently, 700uL of STET was then added to the bacterial pellet followed by 25uL of lysozyme stock solution. The sample was vortexed to re-suspend the pellet and placed on ice for five to ten minutes to allow the lysozyme digestion of the bacterial wall. After the lysozyme digestion, the tubes containing the re-suspended bacteria were boiled for exactly one minute followed by a high-speed spin for ten minutes at 23°C. The pellet was removed from the mixture and 50uL of
2.5 M sodium acetate was added to the remaining supernatant followed by the addition of 700uL of room temperature isopropanol. The samples were mixed by gentle inversion and centrifuged again for 10 minutes at full speed at 21°C to prevent the suspension from generating too much heat, thus preventing any degradation. The observation of a white/clear pellet should form at this point followed by removal of the supernatant. 200uL of 70% EtOH was added to the pellet and centrifuged again at maximum speed for five minutes at 4°C. The supernatant was discarded again followed by the air drying of the pellet by allowing the tube to remain inverted for ten minutes. The pellet was re-suspended in 50uL sterile TE containing 0.2ug/mL RNase A, incubated at 21°C for about ten minutes and placed on ice.

3.10 Maxi- Prep Procedures

The mini-prep samples were then suitable for sequencing resulting in one positive sample produced for each SUMO1 variant (Figure 7). The positive sequence samples were then subjected to Maxi-Prep procedures using a Qiagen Plasmid Endo-Free Maxi Prep® kit to continue with the ultimate goal of producing the respective SUMO1 variant calibration curves. The maxi-prep protocol began with the inoculation of the three selected sequenced samples initially stored in cryovials (-80°C) into 500uL 2xYT. The bacterial cells were then harvested by centrifugation at 6000 x g for fifteen minutes at 4°C. The supernatant was pipetted out and 10mL of Buffer P1 containing RNAse A as well as LyseBlue was added to the pellet followed by complete mixing until no cell clumps remained. 10mL of Buffer P2 was then added and mixed thoroughly by inverting the sealed tube four to six times and incubated at 23°C for 5 minutes. 10mL of chilled buffer P3 was added to the lysate and mixed immediately by inverting four to six times. The lysate was poured into the barrel of the QIAfilter cartridge and incubated
again for ten minutes. The cap was removed from the cartridge while gently inserting the plunger inside. The cell lysate was then filtered into a 50mL conical tube. 2.5mL of Buffer ER was added to the filtered lysate and inverted ten times followed by a thirty-minute incubation on ice. The QIAGEN-tip 500 was equilibrated by applying 10mL of Buffer QBT and the column was emptied by gravity flow to reduce surface tension. The filtered lysate was then added to the QIAGEN-tip and allowed to enter the resin by gravity flow. The QIAGEN-tip was washed with 2 x 30mL buffer QC. Elution of DNA was carried out by adding 15mL buffer QN followed by the precipitation of DNA by adding 10.5mL of room temperature isopropanol. The samples were mixed and centrifuged immediately at 15000 x g for thirty minutes at 4°C. The supernatant was carefully decanted and the DNA pellet was washed with 5L of endotoxin free room temperature 70% EtOH followed by centrifugation at 15000 x g for ten minutes. The supernatant was removed again followed by an air dry of the pellet for five to ten minutes followed by the dissolving of the DNA in 80uL TE. The maxi-prep samples were quantified using a Qubit Fluorometer and calculations were performed to measure precisely 1ug/uL of each of the SUMO1 variants plasmid DNA.

3.11 T7 RNA Polymerase In Vivo Transcription

Using a Megascript RNA synthesis kit, it was necessary to generate an in vivo transcription using the plasmid DNA of each SUMO1 variant beginning with a HindIII restriction digest, an enzyme that doesn’t cut SUMO itself but only has a cut site within the plasmid itself. This 50uL reaction resulted with 5uL being used for quality verification by gel electrophoresis (Figure 8 Panel B), and the rest precipitated by adding 4.5uL Ammonium Acetate and 90uL of chilled 100% EtOH. The three SUMO1 variant samples were centrifuged for thirty minutes at full
speed and the pellet was re-suspended in 20uL of sterile TE. The precipitated DNA was quantified again using the Qubit to ensure exactly 1ug of DNA would be used for the following in vivo transcription which consisted of assembling an NTP mixture (2uL of each), 2uL of 10X Reaction Buffer, 2uL of enzyme mix, and the respective SUMO1 variant 1ug amount at 23°C. The mixes were then incubated for four hours at 37°C and subjected again to gel electrophoresis to ensure quality (Figure 8 Panel C).

3.12 Generating Calibration Curves

The sequences for each SUMO1 variant were mapped out and inserted into the pJET1.2 plasmid sequence in between the T7 promotor site and the HindII cut site (AAGCTT). The exact number of nucleotides were entered in an oligonucleotide calculator to determine the exact grams/mol of each transcript. Each SUMO1 variant weight was then multiplied by 10^{11} (initial copy number dilution) and divided by Avogadro’s number (6.02 x 10^{23} molecules) to generate the needed amount of transcribed RNA in nanograms based on the molecular weight of each SUMO1 transcript. Based on the nanogram amount required of each RNA sample, dilutions were made following the initial Qubit quantifications. From the first dilution which contained 10^{11} copies, serial dilutions loading 10uL of each previous dilution were performed all the way down to 10^0 copies serving as a negative control. The concentration of each SUMO1 variant transcribed RNA was the amount of each serial dilution loaded into subsequent qRT-PCR triplicate reactions using the one step method. However, the qRT-PCR reactions started with the 10^7 copy number dilution all the way down to 10^0 serial dilutions being loaded. Statistical analysis involved using the Cq values ran in triplicates for each SUMO1 variant, the R-values were then calculated and compared amongst each other using excel spreadsheets.
4. RESULTS

4.1 Preliminary Data

To confirm the existence of the three SUMO1 variants (AIM1), total RNA was initially purified from both HEK293A and A549 cell lines using a Qiagen RNeasy Kit as outlined in materials and methods. Before proceeding to the next steps, a formaldehyde- agarose gel was used to assess the quality of the purified RNA (figure 3).

![Image](image)

Figure 4.1: First RNA Purification of A549 and HEK293A Cell Lines - Formaldehyde-agarose gel electrophoresis is used to assess the quality of the purified RNA, 6uL of purified RNA and 2uL of 4x RNA loading buffer was initially mixed then heated up to 65°C for 5’00” then placed on ice.

The purified RNA has a significantly abundant overall appearance, the 28S and 18S rRNA bands are copious at roughly ≥ 2.5kbp and 1.4kbp, respectively, and also run all the way up past 15kbp, an indication of good RNA quality. The apparent intensity of the 18S bands were substantially higher for the total RNA isolated from the HEK293A cells, which are consistent with the purification data obtained from the Qubit fluorometer (not shown). After the RNA quality assessment, we proceeded with the establishing of the SUMO1 variant’s existence using a p50x enzyme based RT-PCR reaction by first converting the total purified RNA into cDNA
using a Reverse Transcriptase using the two-step method running conditions outlined in Table 1. The PCR products were visualized using a 1.5% DNA agarose gel (figure 4) followed by sequencing to achieve a full confirmation of their existence (sequencing data not shown).

Figure 4.2 First confirmation of SUMO1 Variants using RT-PCR.

Following the initial round of RT-PCR and the additional confirmation provided by sequencing the PCR products, it was logical to then proceed with qRT-PCR to assess exactly when each variant is being produced in hopes of determining if there is an overall preference within the cell. These qRT-PCR reactions were assembled in MyGo strips and ran on a MyGo thermocycler, with data acquisition conducted by MyGo software (figure5) running on MAC machinery.

Figure 4.3: First round of qRT-PCR of HEK293A and A549 cell lines

The qRT-PCR performed using the previously synthesized cDNA appears to have worked (Figure 4). The melting curves (raw data not shown) suggested the products produced were specific and electrophoresis (Figure 4) results indeed confirmed the amplification of the
right SUMO1 variant targets. Figure 5 thus represents the Cq values obtained indicating the cycle in which the amplification begins its logarithmic growth. It was initially expected that the positive control primers to amplify all the variants present within the samples first because the sequences targeted by these primers are present in all the variants. Therefore, it was logical to assume the positive control would be the first to be amplified having the lowest Cq. Also, based on the intensity of the RT-PCR products obtained in the initial RT-PCR round (Figure 4), we expected var1 to be the most abundant of the three variants and var3 to be the least abundant with the Cq values ranging in order from: var1 > var2 > var3. Lastly, based on the first RT-PCR reaction, we expected the results to be similar for both cell lines. However, based on the actual Cq values obtained, this was not the case. Instead, var2 was the product with the lowest Cq while both var2 and the positive control Cqs were very similar, and the Cq for var1 and var3 were very similar. Figure 6 is the 1.5% DNA agarose visualization of the three SUMO1variant PCR products ran in triplicates:

![Agarose Gel of RT-qPCR products generated ran in triplicates.](image)

The subsequent runs of qRT-PCR (Figure 6) aid in the establishing a relative Cq value for each variant using newly purified RNA from a fresh batch of tissue culture cell lines: A549 and HEK293A. The signals were higher in A549 cell lines consistent with the qRT-PCR values. Thus, if the relative levels for each variant can be established, they can then be compared to
SUMO1 variant levels within live tissue samples such as PBM cells (AIM2). The comparison can be used to assess whether these transcript levels fluctuate when cells are exposed to different stress conditions, meaning one SUMO1 variant may become upregulated over another during these different conditions.

4.2 Final Data Collection Using Calibration Curves

However, while the preliminary date generated was useful in the simple establishment of the three SUMO1 variants, it was not enough to simply rule which SUMO1 variant is abundant over another. Therefore, calibration curves were absolutely essential to provide the exact copy amount of each SUMO1 variant produced in the cell. The first step required the cloning of each SUMO1 variant within a pJET1.2 plasmid by first conducting another PCR reaction of each SUMO1 variant by means of Taq-Polymerase followed by blunt-end restriction digests using the Thermo-Scientific® CloneJET PCR Cloning Kit as outlined in materials and methods. Ligation reactions then followed via use of the pJET1.2/blunt cloning vector and a T4 DNA ligase. The transformation of NEB10 b competent cells with said ligation reactions were performed followed by subsequent plating which resulted in the selection of five colonies from each SUMO1 variant plate. Another Taq Polymerase PCR reaction was implemented to continue the screening process using the pJET1.2 forward primer, along with the three respective SUMO1 variant reverse primers. Due to the added weight of the T7 polymerase start site (45bp), the predicted molecular weights of each SUMO1 variant increased. Var1 was predicted to weigh 399bp, var2 had an expected molecular weight of 301bp, and the last variant, var3, was expected to weigh 400bp. Of the fifteen colonies chosen for the PCR screening, three var1 clone colonies produced the increased predicted weights, two products for var2, and only
one var3 colony produced the weight as depicted using gel electrophoresis as depicted in figure 7.

![Agarose Gel Electrophoresis Confirmation of the Cloned SUMO1 Variants.](image)

Figure 4.5: Agarose Gel Electrophoresis Confirmation of the Cloned SUMO1 Variants.

The six screened colony samples that were chosen were then subjected to Boiling Lysis Mini Prep procedures and thereafter sent out for further sequencing analysis. Careful analysis of the sequencing results (data not shown) indicated there was one positive sample produced for each SUMO1 variant. The positive sequence samples were then subjected to Maxi-Prep procedures using a Qiagen Plasmid EndoFree Maxi Prep kit to continue with the production of the respective SUMO1 Variant calibration curves. The maxi-prep samples were quantified using a Qubit Fluorometer and calculations were performed to measure precisely 1ug/uL of each of the SUMO1 variants pJET1.2 plasmid DNA. Figure 8 Panel A is the DNA-Agarose gel electrophoresis used to check the quality of the pJET1.2 SUMO1 variant plasmid samples. The confirmation of the SUMO1 variant cloned plasmids made it possible to proceed with a T7 promoter in vivo transcription. This began by performing a restriction digest using HINDIII resulting with 5uL being used for quality verification by gel electrophoresis (Figure 9 Panel B), and the rest of the HINDIII digested DNA sampled were precipitated. The precipitated DNA
was quantified again using the Qubit to ensure exactly 1ug of DNA would be used for the following *in vivo* transcription (Figure 9 Panel C). The Formaldehyde-Agarose gel shows that SUMO1 var1 and var3 were in fact transcribed by the T7 polymerase, however, there was very faint signal for var 2 which was inconsistent with the HINDIII digestion visualization (Figure 8 Panel B). However, Qubit Quantification indicated that there was sufficient concentration for proceeding with the calibration curves.

Generating the SUMO1 Variant calibration curves commenced with calculating the predicted weights for each variant as stated in Methods and Materials. It was these calculations that aided in detecting the exact copy number for each variant by then performing serial dilutions. These dilutions were carefully calculated using the Qubit estimated quantification of the T7 transcribed RNA and subsequently indicated the amount of each dilution required for the new one-step qRT-PCR ran in triplicates. Statistical analysis of each Cq generated resulted with all three $R^2$ values coming close to a 0.99 value: (SUMO1Var1$R^2$=0.9896, SUMO1Var2$R^2$=0.9897, and SUMO1Var3$R^2$=0.9961). Thus, the percent of variance can be explained making these regression models valid (Figure 9).
Figure 4.7: Calibration Curve for all three SUMO1 variants (Var). Regression model stands valid given the predicted $R^2$ values.

The quantification of the SUMO1 Variants indicated Var3 calibration curves began their logarithmic growth first followed by Var 2 and lastly Var 3. However, Var3 produced the $R^2$ value closest to the ideal value of 1 with the steepest slope ($y=3.454x + 6.9033$), meaning the most consistent or greater range of change within the orders of magnitude provided by the conducted serial dilutions. To ensure the order of magnitudes were a correct reflection of the $R^2$ values obtained, gel electrophoresis was performed as a last step quality check (Figure 10 Panels B-D).
Figure 4.8: Gel electrophoresis Confirmation of all SUMO1 Variants Calibration Curves. Panel A is the calibration curve quantification. Panels B-D are each SUMO 1 Variant serial dilutions performed in triplicates to ensure the correct Cq value was used by achieving an average.

Once the calibration curves were generated it was then possible to detect the exact copy number of each variant found within A549 and HEK293 cells, both of which are immortalized cell lines. However, to avoid the detection being based on this characteristic alone, total RNA was purified from PBM cell lines. All three-cell line purified RNA were used to carry out a final one-step qRT-PCR reaction again performed in triplicates. The data obtained indicated a subtle, much less drastic preference for SUMO1Var2 production exclusively within immortalized cell lines, while the PBM cell line indicates a SUMO1Var1 preference (Figure 11 Panels C and D) as compared with the preliminary data. The data for SUMO1Var3, however, resulted in the lowest signal amongst the SUMO1 Variants, thus the need for the construction of an independent SUMO1Var3 graph (Figure 11 Panel B).
Figure 4.9 Quantitative one step RT-PCR of all three SUMO1 Variants. These triplicate samples were obtained from three different cell lines: A549, HEK293A, and PBMC using three different rounds of RNA purification. Panel A is the quantification of all SUMO1 Variants. Panel B depicts SUMO1 Var3 alone due to low signal compared to that of Var 1 and Var 2. Panel C is the quantification sorted out by different SUMO1 Variants with a logarithmic Y-axis to demonstrate a more general relativity between the three SUMO1 Variants. Panel D is also depicted with a logarithmic Y axis grouped by cell line.
5. CONCLUSIONS

Up until now, cellular SUMO regulation has typically been thought to occur through the specific conjugation enzymes, in particular the activity of the E3 ligase. However, our results appear to challenge this dogma, indicating there may be some regulation through alternative splicing, thus making it a post-transcriptional event. Initially following the preliminary data assumptions, the Cq values indicate var2 would be the most predominant SUMO1 transcript in both A549 and HEK293A cell lines, var1 and var3 would be present equally, each being less abundant than var2. However, upon the establishment of the SUMO1 variant calibration curves using the new one-step generated PCR products this was not the case for var 1 and var 2 which demonstrated equal amounts of transcript. However, var 3 was the only SUMO1 variant produced in dramatically lower levels in both immortalized cell lines as well as PBM cell lines. It is logical to proceed by imposing various cellular stress states on the different cell lines to determine if there are certain SUMO1 variants upregulated over others.

However, due to the dramatically lower signal for SUMO1 var3 (a rough 100-fold difference as determined through qRT-PCR results), we can then begin to think the N-terminally truncated form of SUMO1 (encoded by var3) must be expressed in the cell, although probably at lower levels than the full-length form of SUMO1 (encoded by var1 and var2), perhaps serving as a negative feedback mechanism when the SUMO genes are upregulated in response to different cellular stress environments. To provide well-rounded results, it is first necessary to prove SUMO1 alpha is in fact expressed in the cell using different proteomic analyses which might lead to SUMO1 containing some regulation at the post-translational level as well.
6. DISCUSSION

It is generally curious, given the versatility of this system along with the many roles and effects of so many processes with the cell, that generally nothing is known as to how this vital group of proteins is regulated. Given the way the three SUMO1 variants appear to be alternatively spliced, it is hypothesized here that the SUMO gene perhaps has the ability to regulate itself. SUMO1 var1 appears to be the generic prototype of the gene, it is curious that the third SUMO variant (var 3) is missing the second exon which may contain sequence elements within the ORF causing it to be upregulated in times of stress. While the second variant, which is indeed missing a portion of the fifth exon, causing it to generate the 5’ exon might be an area within the 3’UTR where miRNA targets reside are possibly spliced out giving further clues as to how the gene is regulated.

However, the most important conclusions nevertheless are 1) the primers designed appear to allow the specific amplification of the targeted regions 2) the procedures performed produced the expected results and therefore can be used for the intended goals 3) however, accepting Cqs at face value are incorrect assumptions and must be compared to the Calibration Curves to generate a valid copy number of each SUMO1 variant transcript. Thus, we have successfully developed an experimental approach to detect and quantify the different SUMO1 transcript variants produced within the cell. The upregulation of SUMO during different stress conditions may attribute to SUMO’s cyto-protective function in a potential time-specific manner, if further information can be discovered as to how the cell regulates SUMOylation, there might be an opportunity to develop chemicals that have the ability to increase SUMOylation and perhaps those chemicals could be used as potential treatments for viral infections, strokes and heart attacks.
7. STRATEGIES AND PITFALLS

It is difficult to determine whether antibodies typically directed against "SUMO1" would also recognize this shorter version of SUMO1. This would be an important brief project that would require the expression of the shorter SUMO1 version using a bacterial system, purifying the protein and then determining whether it is detected by antibodies. It would also be important to assess whether it is conjugated both in vitro and in vivo using site directed mutagenesis to create these constructs. Problems encountered using protocols were significant including a major source of contamination found within the respective negative control samples which then skewed the qRT-PCR data using the two-step method. It is crucial when setting up PCR reactions to do so under tissue culture aseptic conditions, using separate sterile micro pipettes, as well as distinct tip boxes for both positive and negative controls. Better results were obtained when using a mask as well as separate water reservoirs for the distinct samples generating more distinct Cq values with the negative control reactions being produced at a later time than the respective positive control samples. Another source of contamination identified was the use of shared RNA purification kits. Better results were obtained once a new kit was opened and each buffer aliquoted.

Once contamination issues were sorted out, a smeared profile began to appear on the DNA agarose gels of the qRT-PCR products generated. It was noticed that it was happening only within the var3 reactions. After careful re-examination of the mRNA transcripts for each variant, it was decided to remove the positive control and switch primers. Instead of var3 using the positive control reverse primer, the reverse primer for var1 was used. At first, the smeared profile disappeared but quickly returned.
8. REFERENCES


[Wilson, Van G., “Sumoylation Molecular Biology and Biochemistry.” Department of Medical Microbiology and Immunology, Texas A&M University System Health Science Center, Texas, USA. 2004. Horizon Bioscience. 32 Hewitt’s Lane Wymondham. Norkolf NR 18 0JA, U.K.]


[Pal S, Rosas JM, Rosas-Acosta G. “Identification of the non-structural influenza A viral protein NS1A as a bona fide target of the Small Ubiquitin-like MODifier by the use of...”]

[Nuro-Gyina, PK., Parvin, JD., “Roles for SUMO in pre-mRNA processing.” Department of Biomedical Informatics and the Ohio State University Comprehensive Cancer Center, The Ohio State University, Columbus, OH, USA. Wiley Interdiscip Rev RNA. 2016 Jan-Feb;7(1):105-12. doi: 10.1002/wrna.1318. Epub 2015 Nov 13.]


Dafna Nathan, Shelley L. Berger. Histone sumoylation is a negative regulator in Saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone modifications. Genes & Dev. 2006 20: 966-976; originally published online Apr 5, 2006


9. VITA

Myriah Acuña is a native El Pasoan, who earned her Associates of Science in Chemistry from El Paso Community College in 2013. She enrolled in the University of Texas at El Paso and earned her Bachelor’s of Science in Forensic Biology in 2014. That same year, she entered graduate school at UTEP in the Master of Biological Science program. She taught two molecular cell biology laboratory sections in her last semester. She plans to return for her Ph.D. in the fall of 2018.

Permanent address: 233 Orndorff Dr.
El Paso, Texas 79915

This thesis was typed by Myriah Lorraine Acuña