The Thesis Committee for Sarah Louise Milligan Certifies that this is the approved version of the following dissertation:

Ehrlichia chaffeensis Type 1 Secreted Effector TRP47 is a Novel SUMOylated Nucleomodulin

Committee:

Jere W. McBride, Mentor

David H. Walker

Sanjeev K. Sahni

Dean, Graduate School

Ehrlichia chaffeensis Type 1 Secreted Effector TRP47 is a Novel SUMOylated Nucleomodulin

by

Sarah L. Milligan, B.S.

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas Medical Branch

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science

The University of Texas Medical Branch

November 2015

ACKNOWLEDGEMENTS

I wish to thank:

Dr. Jere W. McBride, my supervising professor, for instilling in me the confidence to accomplish what I thought impossible.

Dr. David H. Walker and Dr. Sanjeev K. Sahni for their flexibility and insight as members of my supervising committee.

Tierra Farris for her immeasurable assistance in the development and execution of this project and for always having faith in my abilities as a student and a scientist.

Xiaofeng Zhang for her generosity in sharing her kindness and professional expertise.

Dr. Shubhajit Mitra for his patience and mentorship.

Dr. Joan E. Nichols and Dr. Jose M. Barral for their support and guidance through a stressful transition.

Claire Smalley for her steadfast friendship when I needed it most.

All past and present members of the McBride lab for their friendship and contributions to my research.

My family, especially my mother, father, and sister, for their unconditional love and understanding.

My husband, Larry, for keeping me grounded and focused on my goals.

This work has been funded by the National Institutes of Health.

Ehrlichia chaffeensis Type 1 Secreted Effector TRP47 is a Novel SUMOylated Nucleomodulin

Publication No._____

Sarah L. Milligan

The University of Texas Medical Branch, 2015

Supervisor: Jere W. McBride

Ehrlichia chaffeensis secretes immunodominant and immunoprotective protein effectors into the host cell using a type I secretion system. The best-characterized of these effectors are the tandem repeat proteins (TRPs). In this study, *E. chaffeensis* TRP47 was identified as a target for host cell SUMOylation and as a novel nucleomodulin with DNA-binding activity and host nuclear localization. Using a microfluidic chip peptide array, a SUMOylation site in the TRP47 N-terminal domain was identified. A SUMO-2 modification was further confirmed using an *in vitro* SUMOylation assay with recombinant TRP47, and native polySUMOylated TRP47 was immunoprecipitated from *E. chaffeensis*-infected cell lysates. TRP47 was detected in host cell nuclei primarily at 24 and 48 hours post-infection by immunofluorescence microscopy. A domain in the N-terminal portion of the protein containing a variant MYND-binding motif was identified as the region responsible for TRP47 nuclear localization using ectopically expressed GFP-tagged TRP47 truncation constructs. Recombinant TRP47 also binds human genomic DNA in an electrophoretic mobility shift assay. Taken together, these data demonstrate that *E. chaffeensis* TRP47 is a nucleomodulin with DNA-binding activity and a non-canonical nuclear localization signal, and is post-translationally modified by SUMOylation, which is a novel modification among bacteria.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Ehrlichiae and ehrlichiosis

Ehrlichiae are obligately intracellular gram-negative bacteria belonging to the order Rickettsiales and are tick-bone zoonoses capable of causing serious and fatal disease in humans and a wide variety of animals (1). The genus *Ehrlichia* currently has six members: *Ehrlichia chaffeensis*, the causative agent of human monocytotropic ehrlichiosis (HME); *Ehrlichia canis*, which causes canine monocytic ehrlichiosis; *Ehrlichia ewingii*, which infects humans and dogs; *Ehrlichia ruminantium*, the etiological agent of heartwater in ruminants (2); *Ehrlichia muris*, which causes persistent infection in mice (3); and the newly characterized *Ehrlichia muris*-like agent (EMLA), which infects both humans and mice (4). *Ehrlichia chaffeensis*, the focus of this study, is transmitted by the lone star tick (*Amblyomma americanum*) and is maintained in nature through persistent infection of white-tailed deer (*Odocoileus virginianus*) (1).

E. chaffeensis is only maintained transstadially in ticks, and thus relies on persistently infected white-tailed deer as a primary reservoir in nature. Humans are an incidental host and can become persistently infected, but this is extremely uncommon and occurs primarily in the immunocompromised (5). Acute *E. chaffeensis* infection, known as HME, is a serious and potentially fatal disease and one of the most prevalent life-threatening tick-borne illnesses in the United States. HME initially presents as an undifferentiated febrile illness with nonspecific flu-like symptoms (6). Administration of doxycycline or tetracycline improves patient outcome, but only if treatment is initiated very early during the course of illness (7). Undiagnosed HME, especially in elderly or immunocompromised patients, can progress to multi-organ failure and result in death. Unfortunately, prompt diagnosis is difficult and is often based primarily on clinical suspicion. It is likely that many HME infections are undiagnosed and go unreported, and

seroepidemiologic and prospective studies suggest that infection rates in a variety of populations are much higher than what would be expected based on the prevalence rate reported by the CDC (2.5 cases per million persons in 2010) (8, 9).

Ehrlichia chaffeensis life cycle and pathobiology

Two morphologically distinct forms of *E. chaffeensis* with unique roles in its life cycle have been characterized: large reticulate cells (RCs) with loosely dispersed nucleoid filaments and small dense-cored cells (DCs) with a condensed nucleoid (10). When *E. chaffeensis* is transmitted to a host, DCs attach to host cells (usually mononuclear phagocytes in human infections) and trigger their own phagocytosis in an early endosome-like vacuole. The DCs then transition to RCs and begin to replicate. After roughly 72 hours, the bacteria transition back into infectious DCs, exit the dying host cell, and infect new cells (11). Throughout this process, *E. chaffeensis* must prevent the destruction of the vacuole in which it resides, evade immune recognition, delay host cell apoptosis, and create an intracellular environment advantageous to its own survival and proliferation. This is accomplished in part through modulation of host gene expression and signaling pathways, manipulation of intracellular endosomal trafficking, and exploitation of host post-translational machinery, and likely by other mechanisms yet to be defined (12-15). Many bacterial pathogens utilize effector proteins to enhance infection and subvert host response. Characterization of the major immunoreactive proteins produced by *E. chaffeensis* has identified a family of tandem repeat proteins (TRPs) acting as type 1 secretion system (T1SS) substrates (16). Members of this family, which includes TRP120, TRP75, TRP32, and TRP47, have an acidic pI, an internal tandem repeat region, a C-terminal T1SS signal, and are immunoprotective (17). TRP120, the best-characterized of the TRPs, interacts with multiple host proteins, binds a specific DNA motif to modulate host gene transcription, and is SUMOylated (15, 18, 19).

Ehrlichia chaffeensis tandem repeat protein 47

Tandem repeat protein 47 (TRP47) is an immunoreactive secreted protein which was identified while investigating ehrlichial antigens, and is an ortholog of the *Ehrlichia canis* protein TRP36 (previously gp36) (20). The *E. chaffeensis* Arkansas strain TRP47 is composed of an N-terminal domain 158 amino acids in length, an internal tandem repeat domain consisting of 7 identical 19-mer repeats (the major epitope), and a short 26 amino acid C-terminal domain. The sequence of TRP47 can vary among strains, usually in the number of tandem repeats as well as TR sequence. A peculiar feature of TRP47 is the considerable difference between its predicted size (32.9 kDa) and the size at which it migrates during gel electrophoresis (47 kDa). This was originally attributed to O-linked glycosylation of the tandem repeats, but a later study utilizing mass spectrometry and two-dimensional gel electrophoresis demonstrated that glycosylation of TRP47 does not occur and that its highly acidic pI (4.2) is responsible for its anomalous migration (21). The same study identified serine and tyrosine phosphorylation sites in the N-terminal domain. TRP47 is only produced by DCs and is differentially expressed in tick and human cells (22). In human monocytes, it is the most highly transcribed *E. chaffeensis* gene, and protein is easily detected by immunoblot; however, in tick cells, it is transcribed at much lower levels and protein cannot be detected, indicating a high level of post-transcriptional regulation. The C-terminal portion of TRP47 is LDAVTSIF-enriched, which serves as a signal for secretion into the host cell by the *E. chaffeensis* type 1 secretion system (16). The role of TRP47 in infection has not been fully characterized, but several host proteins that interact strongly with TRP47 have been identified by yeast two-hybrid analysis and were confirmed by chemiluminescent co-immunoprecipitation (23). The host protein binding partners identified were polycomb group ring finger 5 (PCGF5), proto-oncogene tyrosine-protein kinase Fyn (Fyn), and protein tyrosine phosphatase, non-receptor type 2 (PTPN2). PCGF5 is associated with Polycomb Repressive Complex 2 (PRC2), which promotes chromatin silencing through histone methyltransferase activity (24). PCGF5 also interacts with

SUMOylated TRP120 (15). Fyn is a tyrosine kinase specifically targeting caveolin-1 and has been implicated in viral internalization (25). PTPN2 dephosphorylates phosphotyrosine peptides and regulates tyrosine phosphorylation levels in a wide variety of cellular signaling pathways, including the Jak/Stat pathway, which is inhibited by *E. chaffeensis* infection (26, 27). Analysis of the amino acid sequences of PCGF5, Fyn, and PTPN2 by the GPS-SUMO program predicted SUMO-interacting motifs (SIMs) in all three proteins, suggesting that SUMOylation of TRP47 may play an important role in TRP47-host protein interactions, in a manner similar to the interaction between TRP120 and PCGF5.

The SUMO pathway and exploitation by pathogens

The small ubiquitin-like modifier (SUMO) is a post-translational modification (PTM) with a pathway similar to that of ubiquitination, requiring multiple enzymes to activate the SUMO protein and covalently attach it to a target protein. Higher eukaryotes usually encode multiple SUMO isoforms, while lower eukaryotes typically only encode one isoform. Three isoforms have been identified in humans: SUMO-1, SUMO-2, and SUMO-3. SUMO-1 is associated with nuclear translocation and is constitutively conjugated to target proteins. SUMO-2 and SUMO-3, which have roughly 50% identity with SUMO-1 but 97% identity with each other, are conjugated in response to signaling or cellular stress (28). The SUMO enzymatic cascade occurs in a manner similar to that of ubiquitination. SUMO isoforms are expressed as inactive precursors and must be matured by sentrin-specific proteases, then activated by Uba2/Aos1, an E1 heterodimer (29). Activated SUMO is then conjugated to the E2 enzyme Ubc9 and transferred to the target protein, usually in the presence of an E3 ligase. SUMOylation only occurs on lysine residues, which are normally part of the canonical motif ψ-K-X-D/E, where ψ is a hydrophobic amino acid and X is any amino acid. However, it is estimated that roughly 40% of SUMOylation sites do not correspond to this motif (30). Unlike ubiquitination,

SUMOylation has only been observed in a small number of eukaryotic proteins, and a small proportion of a protein population is SUMOylated at any one time, a phenomenon known as the "SUMO enigma" (31).

Protein SUMOylation has been linked to subcellular localization, protein stability, protein-protein interactions, and regulation of apoptosis, among a number of other cellular processes. SUMO-2 and SUMO-3 can monoSUMOylate proteins, as well as form polySUMO chains via self-conjugating motifs (SCMs); however, SUMO-1 lacks an SCM and only appears as a monomer or as a cap on the end of a polySUMO chain. Elucidation of the functional implications of different polySUMO linkages has not yet taken place, but evidence exists supporting the hypothesis that monoSUMOylation and polySUMOylation have distinct functions in the cell. PolySUMOylation can act as a scaffold for protein-protein interactions, as well as target proteins for ubiquitin-mediated proteolysis (32). MonoSUMOylation has been implicated in mitochondrial fission, nuclear localization, and transcription factor function, in addition to other pathways (33-35).

SUMOylation machinery has only been found in eukaryotic cells, but some pathogens hijack the host SUMOylation pathway, either by encoding proteins with functional SUMOylation sites, or, more commonly, encoding proteins with SUMO-interacting motifs (SIMs) (36-38). Viral pathogens often target the formation of promyelocytic leukemia protein-nuclear bodies (PML-NBs), a previously described antiviral defense mechanism (39). PML possesses a SIM that recruits it to viral proteins as part of the innate immune response. HSV-1 exploits host SUMOylation by encoding a protein that induces degradation of SUMO-modified PML. Additionally, SUMOylation of viral proteins has been observed. Epstein-Barr virus manipulates host SUMOylation through outcompeting PML for endogenous SUMO. This decreases the antiviral capacity of PML and disrupts the formation of PML-NBs (40). Extracellular and intracellular bacterial pathogens have also been shown to manipulate host SUMOylation, although this is not as common as it is among viral pathogens. Thus far, most bacterial pathogens appear

to manipulate the host by decreasing global SUMOylation levels. The conserved *Yersinia* species outer protein YopJ, which is structurally similar to host enzymes required for SUMO maturation, interferes with host SUMOylation by decreasing the levels of conjugated and nonconjugated SUMO-1 (41). *Listeria monocytogenes* also decreases global SUMOylation levels, but utilizes a different approach; the secreted effector listeriolysin O targets the E2 enzyme Ubc9 and some SUMOylated proteins for degradation (42). Intriguingly, *E. chaffeensis* does not alter host SUMOylation levels, suggesting a unique mechanism for exploitation of host SUMOylation. Before this study, the only bacterial protein definitively shown to be SUMOylated was TRP120 (15).

Nuclear localization

The nucleus is a unique microenvironment with complex, tightly regulated import and export mechanisms. Larger proteins translated in the cytoplasm are required for signaling, gene expression, DNA remodeling and repair, and a myriad of other processes, but they need a nuclear localization signal (NLS) in order to enter the nucleus. Once they enter the nucleus, other signals are required to prevent export. Canonical NLSs are recognized by an α/β importin heterodimer and fall into one of three classes. The first class consists of a short stretch of basic amino acids, while the second class is bipartite with two stretches of basic amino acids separated by 10-12 nonspecific residues. The third class of canonical NLS consists of interspersed polar and nonpolar amino acids (43).

TRP47, like the other ehrlichial TRPs, lacks a canonical NLS. Additionally, multiple algorithms for detection of known canonical and non-canonical NLSs have failed to predict a motif associated with nuclear localization. This suggests an alternative mechanism is responsible. Signal sequence-independent mechanisms for nuclear localization include PTMs and interaction with a NLS-containing protein ("piggy-backing"). Analysis of the TRP47 amino acid sequence using the Eukaryotic Linear Motif

resource, which compares user-submitted sequences to a database of functional sites (44), revealed the presence of a variant MYND motif used by the heat shock protein 90 (HSP90) co-chaperones p23 and FKBP38 to associate with a protein complex that enters the nucleus (45).

Bacterial nucleomodulins

The ability to regulate host gene expression is a valuable tool, particularly for intracellular pathogens, which must manipulate the host cell to create an environment suitable for survival and reproduction. Effector proteins which enter and reprogram the host nucleus are known as nucleomodulins and are an emerging field of study in bacterial pathogenesis. Nucleomodulins operate in a variety of mechanisms, usually involving indirect interference with host gene transcription through interactions with host nuclear proteins or histone modification (46). More rarely, nucleomodulins can act as eukaryotic transcription factors and alter host gene expression by directly binding DNA and modulating its transcription. This has been observed in a small number of plant pathogens and endosymbionts, but not in a mammalian pathogen. Further, the ehrlichial TRPs are unique even among this small group of proteins because they directly bind host DNA using internal tandem repeat domains. The only similar family of proteins discovered thus far is the Transcription-activator-like (TAL) effectors of the plant pathogen *Xanthomonas* (46). TAL effectors bind host DNA via their internal tandem repeat domains, and sequence specificity is mediated by the 12^{th} and 13^{th} amino acid residues of each repeat. This allows variation in the DNA target sequence via variation of the amino acid residues which directly come into contact with DNA. DNA-binding proteins similar to the TAL effectors have been described in the plant pathogen *Ralstonia solanacearum*, but the ability of these proteins to act as transcription factors has not been determined (47).

MATERIALS AND METHODS

Maintenance of cell lines and cultivation of *E. chaffeensis*

Human cervical epithelial adenocarcinoma cells (HeLa; ATCC, Manassas, Va.) were cultured in minimal essential medium (MEM; Gibco, Carlsbad, Calif.) supplemented with 10% fetal bovine serum (FBS; Hyclone, Logan, Utah). Human monocytic leukemia cells (THP-1; ATCC) were cultured in RPMI 1640 medium with HEPES (Gibco) supplemented with 1 mM sodium pyruvate (Sigma-Aldrich, St. Louis, Mo.), 2.5 g/liter D-(+)-glucose (Sigma-Aldrich), and 10% FBS (Hyclone). *E. chaffeensis* (Arkansas strain) was maintained in THP-1 cells as previously described (22). The level of infection was determined by cytocentrifuging infected cells onto glass slides and staining using a modified Wright-Giemsa method (PROTOCOL™ Hema 3™; Thermo Fisher Scientific, Waltham, Mass.). Cell-free *E. chaffeensis* was obtained by passing infected cells through a 27 gauge needle ten times.

Antibodies

Polyclonal rabbit anti-TRP47 antibody was generated against a peptide from the TRP47 tandem repeat region (ASVSEGDAVVNAVSQETPA) by a commercial vendor (Genscript, Piscataway, N.J.). Primary antibodies used in this study were rabbit anti-SUMO-1 (Enzo Life Sciences, Farmingdale, N.Y.), rabbit anti-SUMO-2/3 (Enzo Life Sciences), mouse anti-SUMO-2/3 conjugated with Alexa Fluor® 647 (Abcam, Cambridge, United Kingdom), rabbit anti-polycomb group ring finger 5 (anti-PCGF5; Genscript), mouse anti-protein tyrosine phosphatase, non-receptor type 2 (anti-PTPN2; Santa Cruz Biotechnology, Dallas, Tex.), and rabbit anti-proto-oncogene tyrosine-protein kinase Fyn (anti-Fyn; Santa Cruz Biotechnology). Secondary antibodies used in this study were horseradish peroxidase-conjugated goat anti-rabbit IgG(H+L) (Kirkegaard & Perry Laboratories, Gaithersburg, Md.), Alexa Fluor® 594-conjugated goat anti-rabbit IgG(H+L)

and goat anti-mouse IgG(H+L) (Life Technologies, Carlsbad, Calif.), and Alexa Fluor® 488-conjugated goat anti-rabbit IgG(H+L) (Life Technologies).

Immunofluorescence microscopy

Uninfected or *E. chaffeensis*-infected (24, 48, and 72 hours post infection with cell-free bacteria) THP-1 cells were cytocentrifuged onto glass slides, fixed in acetone at -20°C for 10 min, incubated with anti-TRP47 diluted 1:100 with 2% bovine serum albumin (BSA; Sigma-Aldrich) in phosphate-buffered saline (PBS; Sigma-Aldrich) for 1 h, washed, then incubated with Alexa Fluor® 488-conjugated goat anti-rabbit IgG(H+L) for 30 min. HeLa cells were fixed 24 h after transfection for 25 min in a 4% paraformaldehyde solution (Alfa Aesar, Lancashire, United Kingdom), and permeabilized and blocked with 0.1% Triton X-100 (Bio-Rad Laboratories, Hercules, Calif.) and 2% BSA in PBS for 30 min. Slides were then washed and incubated for 1 h with anti-PCGF5 (1:100), anti-Fyn, or anti-PTPN2 (1:50) primary antibody diluted with 2% BSA in PBS, washed, and incubated for 30 min with Alexa Fluor® 594-conjugated goat anti-rabbit IgG(H+L) (PCGF5, Fyn) or goat anti-mouse IgG(H+L) (PTPN2) diluted 1:100 with 2% BSA in PBS. All slides were mounted with ProLong Gold antifade reagent with DAPI (4′,6-diamidino-2-phenylindole) (Cell Signaling Technology, Danvers, Mass.) and dried at room temperature overnight. Images were obtained using an Olympus BX61 epifluorescence microscope and were analyzed with Slidebook software (version 5.0; Intelligent Imaging Innovations, Denver, CO).

Generation and analysis of fluorescent signal intensity profiles

Distribution of TRP47 in *E. chaffeensis*-infected THP-1 cells was determined by generating fluorescent signal intensity profiles for each time point (uninfected and 24, 48, and 72 hpi) and comparing the intensity of FITC (TRP47) and DAPI signal. Line segments were drawn through the cytoplasm and nuclei of target cells, excluding morulae, and the

resulting signal values were exported from Slidebook and analyzed using Microsoft Excel. Fluorescent signal intensity profiles were created by plotting intensity (measured in ADU; Analogue to Digital Units) of data points against their position on the line segment (measured in pixels). Nuclear and cytoplasmic signal values were differentiated based on the DAPI signal intensity, which is high in nuclei and low in the cytoplasm. FITC signal values (n=50) from each region were averaged to determine average nuclear and cytoplasmic signal intensity for each time point. Because anti-TRP47 background staining of healthy cells was observed, the average signal intensities from healthy cells were subtracted from values obtained from infected cells before further analysis. Adjusted average signal values for each time point were plotted on a bar graph.

Microfluidic peptide chip array

Detection of SUMO 2/3 modification of peptides corresponding to potential TRP47 SUMOylation sites was performed by incubating a microfluidic peptide chip array with *E. chaffeensis*-infected cell lysate, then probing with a fluorescent anti-SUMO antibody. 12-mer peptides for the array consisted of a target lysine residue (K49 or K71) and the flanking 6 N-terminal residues and 5 C-terminal residues from the TRP47 amino acid sequence. Peptides with an alanine substituted for the target lysine residue served as negative controls. Peptides were synthesized directly on the chip with the mParaflo® Microchip System (LC Sciences, Houston, Tex.) using previously described methods (48-50). Cell lysate was extracted from *E. chaffeensis*-infected THP-1 cells using an Abcam Whole Cell Extraction Kit (ab113475) in the presence of 20 mM N-ethylmaleimide (NEM, covalent isopeptidase inhibitor; Sigma-Aldrich) and 20mM iodoacetamide (isopeptidase inhibitor; Sigma-Aldrich) to prevent deSUMOylation. The peptide chip was incubated with cell lysate, E1 enzyme, E2 enzyme, SUMO-2, SUMO-3, and Mg-ATP for 3 h at 37°C, and nonspecifically binding proteins were removed with washing. The peptide chip was then

incubated with Alexa Fluor® 647-conjugated mouse anti-SUMO-2/3 (1:500) and analyzed using the Axon GenePix® 4400A Microarray Scanner and GenePix® Pro 7 software (Molecular Devices, Sunnyvale, Calif.). Significant positive signals were determined by comparison with the corresponding negative control peptide.

Immunoprecipitation of native polySUMOylated TRP47

Cell lysate from *E. chaffeensis*-infected THP-1 cells was prepared as described in the above section. Protein concentration was determined using the DC™ protein assay (Bio-Rad Laboratories). PolySUMOylated protein was enriched from 200 µg cell lysate in three separate reactions using the POLYSUMO-QAPTURE® kit (Enzo Life Sciences) according to the manufacturer's recommendations. After overnight incubation with the affinity matrix, unbound protein was removed and the reactions were washed. PolySUMOylated protein bound by the affinity matrix was then eluted with either 1x NuPAGE® lithium dodecyl sulfate (LDS) denaturing sample buffer (Thermo Fisher Scientific) or the nondenaturing polySUMO-Qapture Elution Buffer included with the kit. Anti-TRP47 immunoprecipitation was then performed on the eluates collected under nondenaturing conditions using the Pierce™ Crosslink IP Kit (Thermo Fisher Scientific) according to the manufacturer's recommendations. Reactions were pre-cleared using Pierce™ Control Agarose Resin and incubated overnight with Pierce™ Protein A/G Plus Agarose crosslinked to anti-TRP47 antibody. After washing, bound protein was eluted and boiled with 4x LDS sample buffer. Cell lysate, eluate from poly-SUMO enrichment, and eluate from the anti-TRP47 IP were separated by size using SDS-PAGE and transferred onto nitrocellulose membrane (Genscript). The membrane was cut in half, and immunoblotting was performed against TRP47 and SUMO-2/3. Blots were blocked for one hour in 5% nonfat dry milk (TRP47) or 1% BSA (SUMO-2/3) in 0.1% Tween-20/Tris-buffered saline solution (TBST; Sigma-Aldrich) at room temperature, incubated with

primary antibody (anti-TRP47 or rabbit anti-SUMO 2/3, 1:1000) overnight at 4°C, washed with TBST, incubated with horseradish peroxidase-conjugated goat anti-rabbit secondary antibody (1:25,000) for one hour at room temperature, and washed again with TBST. Bound antibody was detected by chemiluminescence using SuperSignal™ West Dura Extended Duration Substrate (Thermo Fisher Scientific) and autoradiography film.

Expression constructs and site-directed mutagenesis

E. chaffeensis genomic DNA was extracted from infected THP-1 cells using the DNeasy Blood & Tissue Kit (QIAGEN, Hilden, Germany), and DNA concentration was determined with a NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific). Oligonucleotide primers (Table 1) were designed to amplify various regions of TRP47 for cloning into the pGEX-6P-1 N-terminal glutathione S-transferase (GST) fusion bacterial expression vector (GE Healthcare Life Sciences, Pittsburgh, Penn.) or the pAcGFP1-C In-Fusion® Ready N-terminal Green Fluorescent Protein (GFP) fusion bacterial/mammalian expression vector (Clontech, Mountain View, Calif.). PCR products were generated from *E. chaffeensis* genomic DNA with HotMasterMix (5 PRIME, Gaithersburg, Md.) using the following thermal cycling program: 94°C for 2 min; 30 cycles of 94°C for 30 s, annealing temperature (5°C less than the lowest primer T_m) for 30s, and 65°C for the appropriate extension time (30 s per 500 product base pairs); and 65°C for 7 min. Correct size was verified using the FlashGel DNA electrophoresis system (Lonza, Basel, Switzerland), and PCR products were purified with the MinElute PCR Purification Kit (QIAGEN). PCR products for pGEX-6P-1 constructs were digested with EcoRI and SalI high-fidelity restriction enzymes (New England Biolabs, Ipswich, Mass.) and ligated into digested vector (Fast-Link™ DNA Ligase; Epicentre, Madison, Wis.). Constructs were transformed into TOP10 *E. coli* (Invitrogen, Carlsbad, Calif.), and transformants were selected by growth on LB agar with ampicillin. PCR products for pAcGFP1-C constructs were cloned

into pre-linearized vector using the In-Fusion PCR Cloning Kit (Clontech), and transformants were selected by growth on LB agar with kanamycin. The DNA fragment encoding the His-tagged MYND-binding domain (His-MBD) was created by annealing the forward and reverse oligonucleotides (Table 1). A PCR product encoding TRP47 without the MYND-binding domain (No Motif construct) was amplified from a commercially synthesized pUc57 plasmid (Genscript) using the forward and reverse primers for the full-length GFP construct. PCR screening was used to verify correct insert size, and plasmids

Table 1: Oligonucleotide primers used to create *E. chaffeensis* TRP47 expression constructs

Target*	Forward Oligonucleotide Sequence	Reverse Oligonucleotide Sequence	Product Size (bp)
GST constructs			
FL_{1-316}	aaaagaattcatgcttcatttaacaacag	aaaagtcgacgaaataaaagtatctattacc	968
$5TR_{159-253}$	aaaagaattcgctagtgtatctgaaggag	aaaagtcgactgcaggagtttcttggctt	381
GFP constructs			
FL_{1-316}	aaggcctctgtcgacatgcttcatttaacaacagaaatt	agaattcgcaagcttttagaaataaaagtatctattaccaa	981
N_{1-158}	aaggcctctgtcgacatgcttcatttaacaacagaaatt	agaattcgcaagcttatttccttcaagaactggaac	501
$Ntrunc_{1-150}$	aaggcctctgtcgacatgcttcatttaacaacagaaatt	agaattcgcaagcttcactatactgtcacttaaagat	477
$TRC_{152-316}$	aaggcctctgtcgacggaaatgctagtgtatctgaa	agaattcgcaagcttttagaaataaaagtatctattaccaa	495
$C_{291-316}$	aaggcctctgtcgacactcaaccacaatctagagat	agaattcgcaagcttttagaaataaaagtatctattaccaa	111
$His\text{-}MBD_{151-155}$	aaggcctctgtcgacgttccagttcttgaacaccaccaccaccaccac	agaattcgcaagcttgtggtggtggtggtggtgttcaagaactggaac	48
Mutagenesis			
K49R	ggaagtgaacctgatcatggttatcatattttattagaaacaatggtcatgttatat	atataacatgaccattgtttctaaataaaatatgataaccatgatcaggttcacttcc	-
K71R	ggtgtacaagctgaaaactttgtatttgatataagaaatcacaatttaagagct	agctcttaaattgtgatttcttatatcaaatacaaagttttcagcttgtacacc	-

*Subscripts represent the amino acids contained in the fragment.

from positive colonies were isolated (QIAprep® Spin Miniprep Kit, QIAGEN) and sequenced to verify proper orientation and frame. Lysine-to-arginine mutation of the full-length TRP47 pAcGFP1-C construct was performed using the QuikChange II Site-Directed Mutagenesis Kit (Agilent, Santa Clara, Calif.) according to the manufacturer's instructions. Mutating primers (Table 1) were created using the QuikChange Primer Design program. Plasmids from positive transformants were isolated and sequenced to verify presence of the correct mutation.

Expression and purification of recombinant TRP47

Plasmids encoding full-length TRP47 and TRP47 tandem repeat region GST-fusion constructs were transformed into TurboCells® BL21(DE3) *E. coli* and plated on LB agar with ampicillin. Colonies were PCR-screened to verify presence of the insert, and positive transformants were regrown overnight in LB broth with ampicillin. Overnight cultures were diluted 1:20 in fresh growth medium and incubated with shaking at 37°C until an OD_{600} of 0.5 was achieved. Protein expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; Sigma-Aldrich) for 4 h, after which cultures were pelleted and stored at -80°C. Cell pellets were resuspended in PBS and lysed by sonication in the presence of protease inhibitors (cOmplete™ mini, EDTA-free; Roche, Basel, Switzerland). Lysates were cleared by centrifugation at 10,000xg and 4°C for 1 h, and the supernatant was collected and incubated with Glutathione Sepharose 4B affinity resin (GE Healthcare Life Sciences) at 4°C with rotation overnight. After washing with PBS, bound protein was eluted from the resin with 25 mM glutathione and the concentration measured using the Bio-Rad Protein Assay (Bio-Rad Laboratories). Recombinant protein expression and purification were verified using SDS-PAGE and protein staining (AcquaStain Protein Gel Stain; Bulldog Bio, Portsmouth, N.H.).

In vitro SUMOylation assay

SUMOylation of recombinant GST-tagged full-length TRP47 was performed with an *in vitro* SUMOylation kit (Enzo Life Sciences). TRP47 (200 nm) was added to SUMOylation buffer, free SUMO protein (isoform 1, 2, or 3), E1 enzyme, and E2 enzyme, with or without Mg-ATP. Reactions without Mg-ATP served as negative controls. All reactions were incubated at 37°C for 1 h then boiled with 1x LDS sample buffer and 1x NuPAGE® Sample Reducing Agent (Thermo Fisher Scientific). Samples were analyzed by SDS-PAGE and western blotting with anti-TRP47 (1:1,000), rabbit anti-SUMO-1 (1:1,000), or rabbit anti-SUMO-2/3 (1:1,000) and goat anti-rabbit horseradish peroxidase-conjugated secondary antibody (1:25,000). Bound antibody was detected by chemiluminescence using SuperSignal™ West Dura Extended Duration Substrate and autoradiography film.

Ectopic expression of GFP-tagged TRP47 constructs

Plasmids for transfection (pAcGFP1-C FL, N, Ntrunc, TRC, C, MBD, No Motif, K49R, K71R, GFP control) were isolated from 150 mL overnight *E. coli* cultures with the PerfectPrep EndoFree Plasmid Maxi Kit (5 PRIME) and quantitated using a spectrophotometer. HeLa cells were seeded at a density of 0.8 x 10^5 cells/mL in 8-well chamber slides 24 h prior to transfection. Transfections were performed in duplicate with Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's recommendations. Twenty-four hours after transfection, cells were fixed with acetone at -20°C for 10 min and mounted with ProLong Gold antifade reagent with DAPI. Slides were dried at room temperature overnight and viewed using fluorescent microscopy.

Electrophoretic mobility shift assay

The electrophoretic mobility shift assay (EMSA) was performed with recombinant GST-tagged full-length (FL) or 5 tandem repeat (5TR) TRP47 protein constructs, sheared

THP-1 genomic DNA (gDNA), biotin-labeled gDNA, and the LightShift™ Chemiluminescent EMSA Kit (Thermo Fisher Scientific). Sheared THP-1 genomic DNA was prepared using the ChIP-IT® Express Enzymatic Shearing Kit (Active Motif, Carlsbad, Calif.), purified, and quantitated. Complete shearing of chromatin into fragments of 200 base pairs or less was verified by gel electrophoresis and ethidium bromide staining. Sheared gDNA was biotin-labeled using the *Label* IT® Nucleic Acid Biotin Labeling Kit (Mirus Bio, Madison, Wis.). Reactions were performed in duplicate and contained a base mix of 5 ng of biotin-labeled gDNA, 10 mM Tris, 50 mM KCl, 1 mM DTT, 2.5% glycerol, 5 mM $MgCl_2$, 25 ng/µL poly (dI·dC), and 0.05% NP-40. Reactions containing only the base mix served as negative controls. 4 µg of 5TR TRP47 or 4 and 8 µg of full-length TRP47 was added to all other reactions. Competition reactions also contained 2.5 µg of unlabeled gDNA, which was incubated with the protein for 30 min before addition of biotin-labeled gDNA. The EMSA using 5TR TRP47 included additional reactions containing anti-TRP47 antibody diluted 1:20, 1:50, 1:100, and 1:200 in PBS. In these reactions, antibody was incubated with the protein for 30 min before addition of biotin-labeled gDNA. All reactions were incubated for a total of 1 h at 4°C. After incubation, non-denaturing loading buffer was added to the reactions, and they were electrophoresed for 90 min at 100V on a 6% native polyacrylamide gel in 0.5X Tris-borate-EDTA (TBE; Corning Life Sciences, Corning, N.Y.) buffer. Protein-DNA complexes were transferred onto Biodyne® B modified nylon membrane (Thermo Fisher Scientific) for 60 min at 20V and crosslinked for 5 min using a CL-1000 Ultraviolet Crosslinker (254 nm wavelength; UVP, Upland, Calif.). Protein-DNA complexes were visualized using the Chemiluminescent Nucleic Acid Detection Module (Thermo Fisher Scientific) and autoradiography film.

RESULTS

Identification of the TRP47 SUMOylation site

The TRP47 amino acid sequence (Fig. 1A) was analyzed using the GPS-SUMO program (30) to identify possible SUMOylation sites. Two lysine residues (K49 and K71) were identified as candidates. Peptides containing these residues and the flanking amino acids, as well as mutant peptides with an alanine residue substituted for the target lysine residue, were synthesized on a high-density microfluidic chip. After incubation with infected cell lysate and enzymes E1 and E2, peptides were probed with anti-SUMO fluorescent antibody and the fluorescent signal intensity analyzed (Fig. 1B). The wild-type and mutant peptides containing the K49 candidate SUMOylation site had similar and low signal intensities. The signal intensity of the wild-type peptide containing the K71 candidate SUMOylation site was robust and significantly higher than the signal intensity of the corresponding mutant peptide, and thus K71 was identified as the TRP47 SUMOylation site.

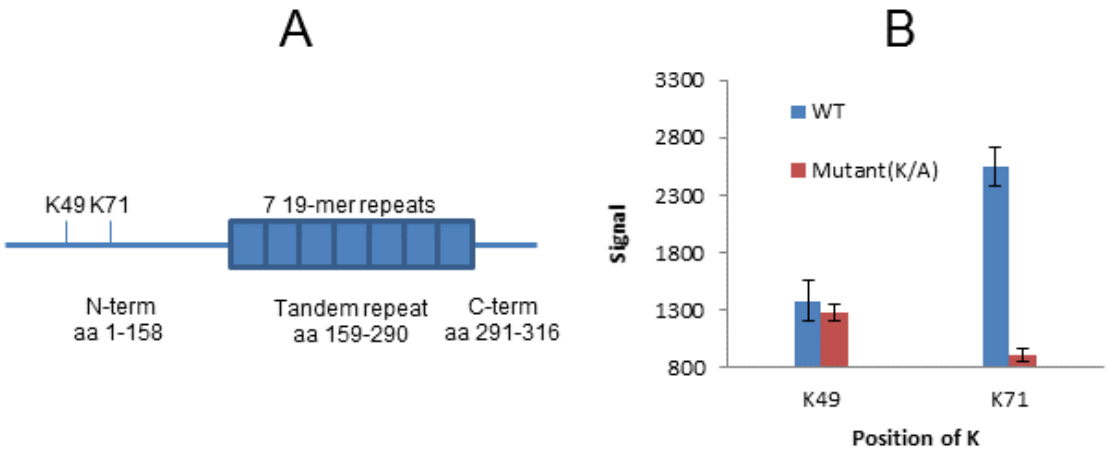

Figure 1: TRP47 protein domains and predicted SUMOylation sites. (A) Schematic showing the TRP47 protein domains and predicted SUMOylation sites. (B) Fluorescent signal intensity of K49 and K71 wild-type and mutant control peptides after incubation with fluorescent anti-SUMO 2/3 antibody.

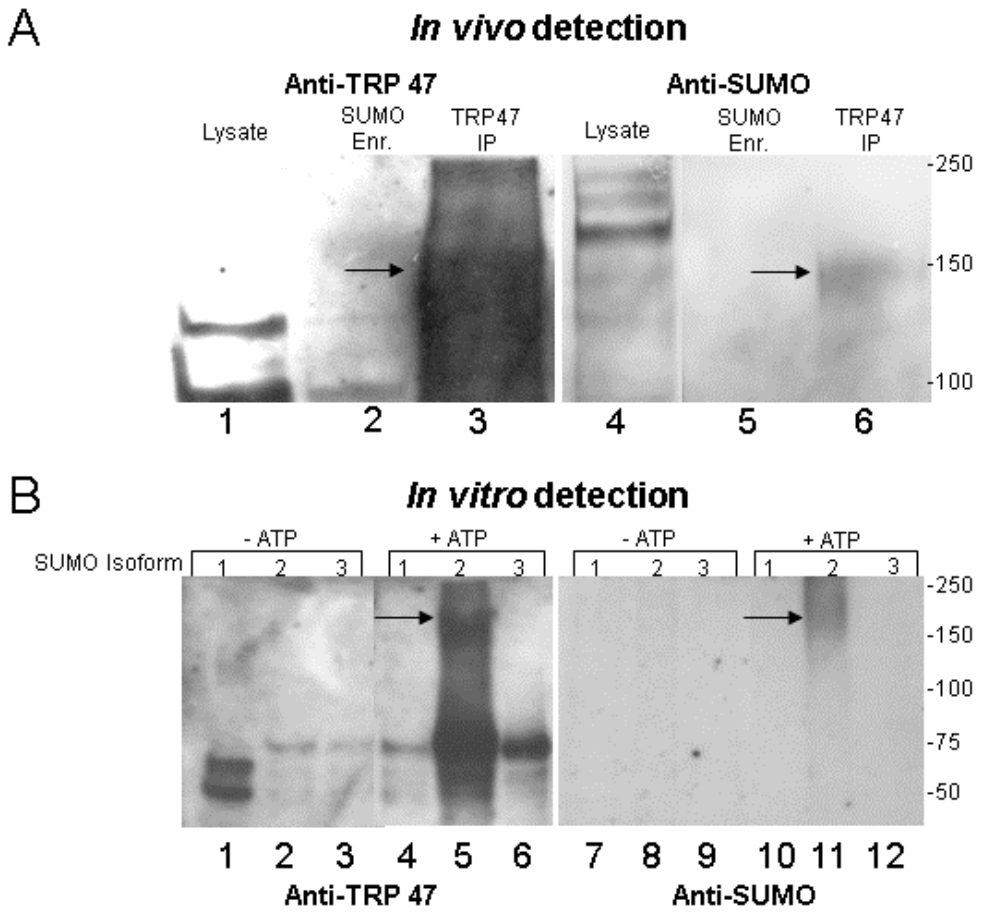

Figure 2: PolySUMOylation of TRP47 during infection and *in vitro*. (A) Immunoprecipitation of native polySUMOylated TRP47. Lanes 1-3 were probed with anti-TRP47 antibody. Lanes 4-6 were probed with anti-SUMO-2/3 antibody. Lanes 1 and 4 ("Lysate") are 4 µg infected cell lysate. Lanes 2 and 5 ("SUMO-Enr.") are infected cell lysate after enrichment with polySUMOylated protein. Lanes 3 and 6 ("TRP47 IP") are eluate from immunoprecipitation of TRP47 from polySUMO-enriched cell lysate. Arrows at lanes 3 and 6 indicate polySUMOylated TRP47. (B) *In vitro* SUMOylation assay. Top numbers indicate the SUMO isoform added to the reaction and bottom numbers indicate the lane number. Lanes 1-6 were probed with anti-TRP47 antibody. Lanes 7-12 were probed with anti-pan-SUMO antibody. Lanes 1-3 and 7-9 ("-ATP") are ATP-negative control reactions. Arrows at lanes 5 and 11 indicate TRP47 polySUMOylated by SUMO-2. Unmodified TRP47 migrated at 75 kDa in lanes 1-6.

Detection of SUMOylated TRP47 during infection and *in vitro*

TRP47 polySUMOylation was observed using two different methods, and the SUMO isoform responsible for TRP47 modification was identified. PolySUMOylation of native TRP47 was detected by enrichment and immunoprecipitation. PolySUMOylated protein was enriched from *E. chaffeensis*-infected THP-1 cell lysate and

immunoprecipitated using anti-TRP47 antibody. The resulting eluate was analyzed using western blotting and anti-TRP47 and anti-SUMO2/3 antibody. Although there was a large amount of background staining in the lane containing the anti-TRP47 immunoprecipitation on the anti-TRP47 blot, a distinct band at 150 kDa exhibited more intense antibody staining and was wider than the rest of the lane, indicating a higher abundance of protein. "Smearing" of lanes is a common occurrence when probing for SUMOylated and ubiquitinated protein (51-53). This 150 kDa band was also observed on the anti-SUMO blot (Fig. 2A, indicated by arrows). A single SUMO modification adds approximately 15 kDa to the mass of a target protein. Native TRP47 has a mass of 47 kDa when electrophoresed, and the 100 kDa increase in mass observed here is further evidence of polySUMOylation. Additionally, an *in vitro* SUMOylation assay was performed with recombinant GST-tagged full-length TRP47 to determine which SUMO isoforms are responsible for TRP47 SUMOylation. A 175 kDa band was observed when SUMO-2 and Mg-ATP were incubated with TRP47 (Fig. 2B, indicated by the arrows), demonstrating SUMO-2 modification. The GST tag adds 27 kDa to the mass of a fusion protein, and when the mass of this tag is subtracted, the polySUMOylated TRP47 bands indicated by arrows in Fig. 2A and Fig. 2B are similar in size. Bands corresponding to unmodified GST-tagged TRP47 can be observed in the left-hand blot at roughly 75 kDa (Fig. 2B, lanes 1-6).

SUMOylation and TRP47 protein-protein interactions

To determine the effect of SUMOylation on the previously observed colocalization of TRP47 with the host protein binding partners PCGF5, Fyn, and PTPN2 (23), plasmids encoding GFP and full-length (FL) TRP47, K49R mutant, and K71R mutant GFP-tagged constructs (detailed schematic of constructs in Fig. 6) were transfected into HeLa cells, and IFA was performed using antibodies against the host proteins of interest (Fig. 3). Full-length and K71R mutant TRP47 exhibited mixed nuclear and cytoplasmic localization,

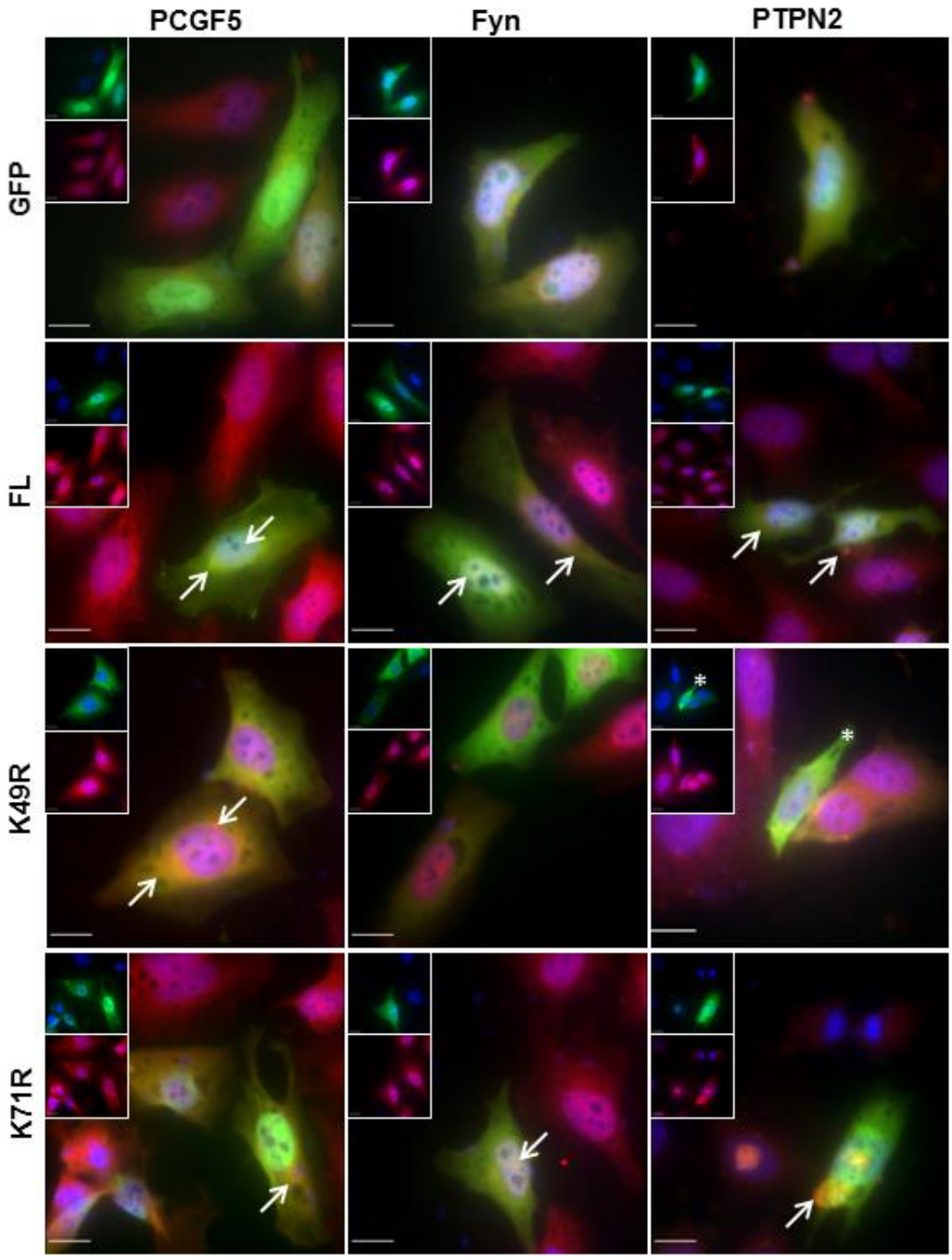

Figure 3: Effect of TRP47 SUMOylation on colocalization with host proteins. Row labels indicate the ectopic expression construct (FL = full-length TRP47) and column labels indicate the host protein stained by antibody. Areas of colocalization are indicated by arrows and appear yellow or orange. The asterisks in the K49R-PTPN2 panel indicate GFP expression. Nuclei are stained blue by DAPI, GFP-tagged ectopic expression constructs fluoresce green (FITC), and host protein bound by antibody fluoresces red (TRITC). Inset panels are FITC and DAPI (top) and TRITC and DAPI (bottom). Line segments measure a distance of 10 µm.

while the K49R mutant was observed exclusively in the cytoplasm (Figs. 3 and 7). Areas of colocalization between TRP47 and host proteins in Figure 3 are indicated by arrows and appear yellow or orange. This color variation was caused by differences in signal intensity among the ectopic expression constructs and among the antibodies against host proteins. PCGF5, which was detected diffusely throughout cells and concentrated in nuclei, colocalized with full-length TRP47 in nuclear and perinuclear puncta. Although the K71R mutant construct exhibited both nuclear and cytoplasmic localization, punctate colocalization with PCGF5 was observed in the cytoplasm only. The K49R mutant also showed punctate perinuclear and diffuse cytoplasmic colocalization with PCGF5. Fyn, which was detected mainly in nuclei and at lower levels in the cytoplasm, exhibited diffuse nuclear and cytoplasmic colocalization with full-length and K71R mutant TRP47, with more intense colocalization occurring in the nucleus. Similarly to what was observed with PCGF5, K49R mutant TRP47 diffusely colocalized with Fyn only in the cytoplasm. PTPN2 exhibited primarily punctate nuclear and cytoplasmic distribution in cells and large punctate colocalization with full-length and K71R mutant TRP47 in the cytoplasm. This pattern was not observed in cells expressing the K49R mutant TRP47 (designated by the asterisk in Fig. 3), which did not colocalize with PTPN2 in the large puncta exhibited by full-length and K71R TRP47. In summary, mutation of the SUMOylated lysine residue at position 71 did not affect TRP47 colocalization with PCGF5, Fyn, or PTPN2; however, mutation of the lysine residue at position 49 abrogated TRP47 nuclear localization and TRP47-PTPN2 colocalization. There do not appear to be any differences mediated by the presence or lack of TRP47 SUMOylation in interaction with these host proteins using this method, although the lysine at position 49 appears to play a role in interaction with PTPN2. Additionally, mutation of the SUMOylation site of the K71R mutant does not decrease the amount of TRP47 observed in the host cell nucleus.

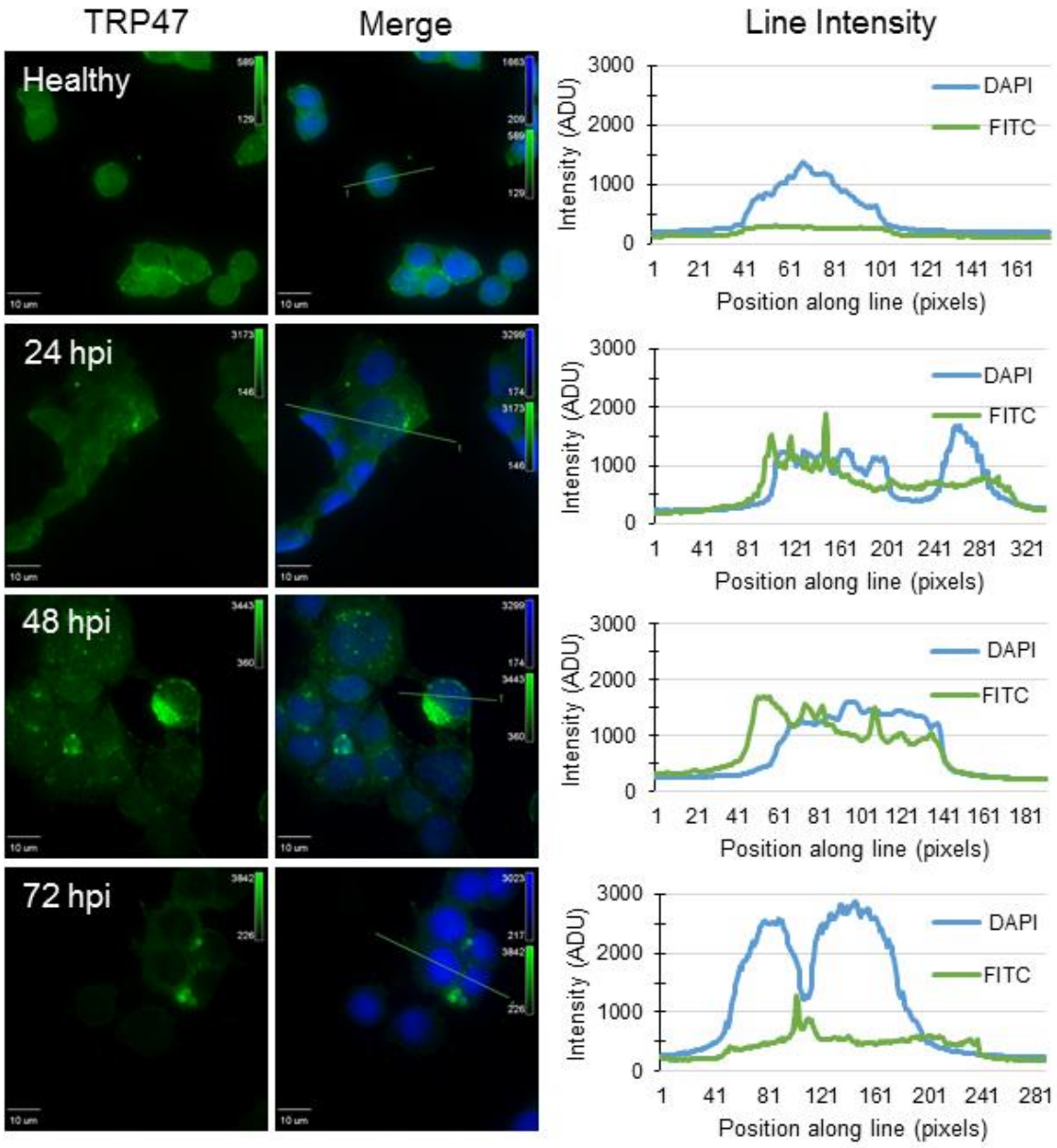

Figure 4: Detection of native TRP47 by immunofluorescence microscopy. Healthy cells are uninfected. Nuclei are stained blue with DAPI and TRP47 bound by antibody fluoresces green. The column of panels labeled "Merge" are combined FITC (TRP47) and DAPI signal, and the green line indicates where signal intensity values were collected to generate the intensity profiles shown on the right. Signal intensity was measured in Analogue to Digital Units (ADU).

Subcellular localization of TRP47 in host cells during infection

TRP47 is expressed and secreted by the infectious dense-cored form of *E. chaffeensis*, which transitions to the reticulate cell form soon after engulfment by the host cell and then back to the dense-cored form after roughly 72 hours of infection. Although

nuclear localization of TRP47 in infected host cells has been previously documented in the literature (23), a study examining temporal changes in TRP47 localization during infection had not been performed. To address this gap in knowledge, the amount and distribution of TRP47 in the host cell were compared at 24, 48, and 72 h post-infection (hpi) by immunofluorescence microscopy and fluorescent signal intensity analysis (Fig. 4). TRP47-expressing morulae were observed at 24, 48, and 72 hpi. The highest number of TRP47-expressing morulae was observed at 48 hpi, as well as the greatest amount of TRP47 visible as puncta in the host nucleus, rather than the more diffuse distribution observed at 24 hpi. The least amount of TRP47 was apparent at 72 hpi. A low and uniform amount of anti-TRP47 background staining was observed in the nuclei and cytoplasm of healthy, uninfected cells.

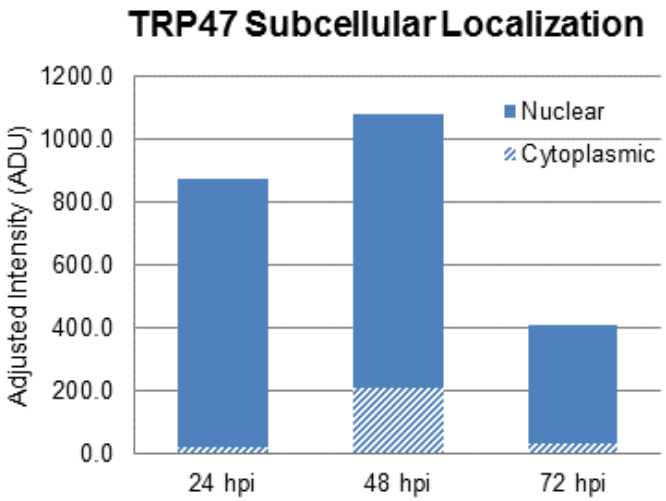

Figure 5: Cytoplasmic and nuclear distribution of TRP47 in host cells. Nuclear and cytoplasmic FITC (TRP47) data points (determined by strength of DAPI signal from Figure 4) were averaged to generate typical signal intensity values for each region at 24, 48, and 72. Intensity values were adjusted to remove background by subtracting values obtained from healthy cells and are measured in Analogue to Digital Units (ADU).

Fluorescent signal intensity was quantitatively measured by drawing line segments (shown in green on the "Merge" panels in Fig. 4) across cells, passing through the nucleus and the cytoplasm, and recording the intensity values for FITC (TRP47) and DAPI (nucleus) at regular intervals along the line. Intensity profiles for healthy cells and for each

time point were generated to display the variation in DAPI and FITC intensity values along the line segment (Fig. 4). Nuclear TRP47 signal was evident at 24 and 48 hpi, and very much decreased at 72 hpi. The overall intensity of TRP47 signal was also much higher at 24 and 48 hpi compared to 72 hpi. Nuclear or cytoplasmic origination of signal values was determined based on the strength of DAPI signal. Morulae were excluded from this portion of the analysis because they also stain with DAPI. To compare nuclear and cytoplasmic distribution of TRP47 between time points, 50 FITC data points from each region were averaged to create typical signal intensity values. After subtracting the nuclear and cytosolic averages of healthy cells to remove background signal, the values from each time point were plotted on a bar graph (Fig. 5). Consistent with the previously presented intensity profiles, combined nuclear and cytosolic TRP47 signal was much higher at 24 and 48 hpi than at 72 hpi. While the average nuclear TRP47 signal intensity was similar between 24 and 48 hpi, the average cytoplasmic TRP47 signal intensity was considerably lower at 24 hpi. This pattern was also observed at 72 hpi, but with a much lower nuclear TRP47 signal intensity value.

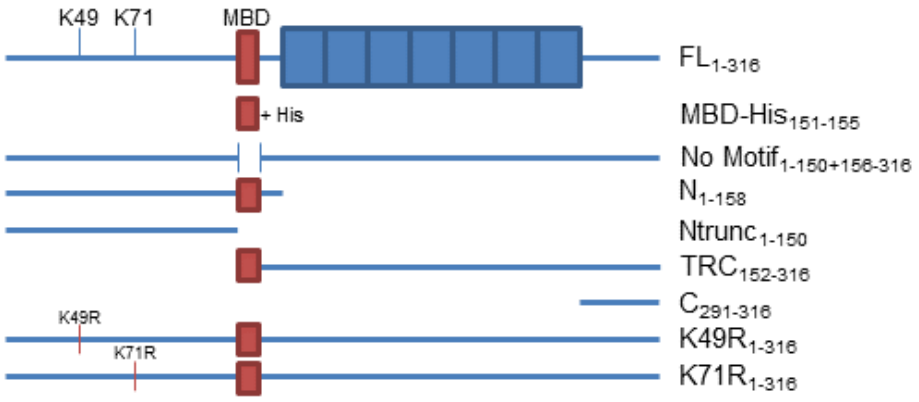

Figure 6: Schematic of GFP-tagged TRP47 expression constructs. Red boxes indicate presence of the MYND-binding domain (MBD). Open space between two lines indicates a deletion mutation. Subscripts on construct labels indicate which amino acid residues are included. Constructs shown are full-length (FL), His-tagged MBD (MBD-His), full-length TRP47 without the MBD (No Motif), N-terminal (N), truncated N-terminal without the MBD (Ntrunc), tandem repeat-C-terminal overlapping the MBD (TRC), C-terminal (C), K49R mutant (K49R), and K71R mutant (K71R).

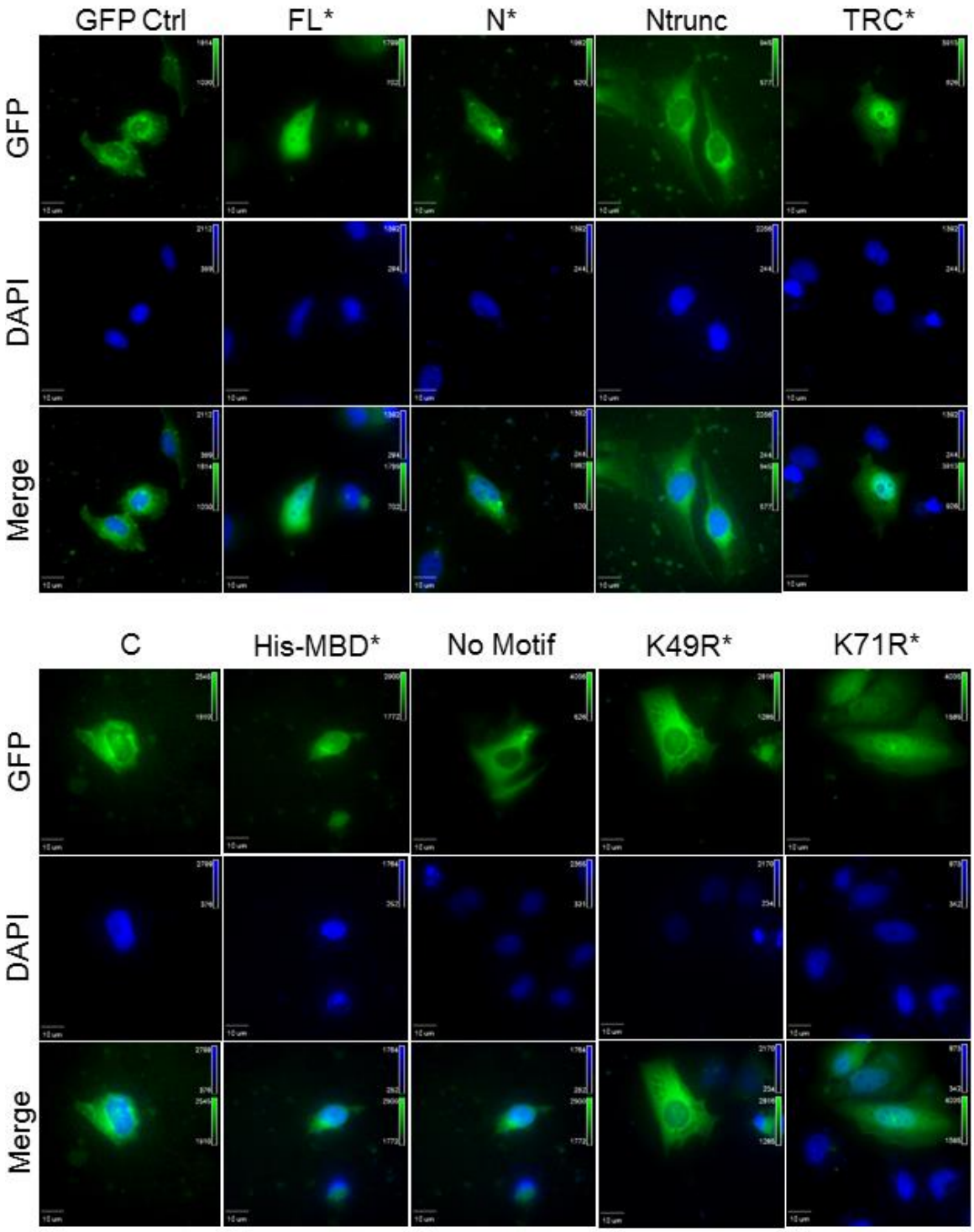

Figure 7: Localization of ectopically expressed TRP47 truncation constructs. Asterisks on column labels indicate presence of the MYND-binding domain. Column labels designate ectopically expressed constructs: full-length (FL), N-terminal (N), truncated N- terminal (Ntrunc), tandem repeat-C-terminal (TRC), C-terminal (C), His-MYND-binding domain (His-MBD), MBD deletion mutant (No Motif), K49R mutant, and K71R mutant. GFP-tagged constructs fluoresce green and DAPI-stained nuclei fluoresce blue.

Identification of the domain required for TRP47 nuclear localization

Analysis of the TRP47 amino acid sequence by the Eukaryotic Linear Motif resource identified a variant MYND-binding motif associated with nuclear localization of members of the Hsp90 complex. To determine the role of this domain in the subcellular localization of TRP47, a GFP control and TRP47 constructs (displayed in Fig. 6) encoding GFP-tagged full-length (FL), N-terminal (N), truncated N-terminal (Ntrunc), tandem repeat-C-terminal (TRC), C-terminal (C), His-MYND-binding domain (His-MBD), MBD deletion mutant (No Motif), K49R mutant, and K71R mutant TRP47 were ectopically expressed in HeLa cells and their localization observed by fluorescent microscopy (Figure 7). The FL, N, TRC, His-MBD, K49R mutant, and K71R mutant constructs contained the identified motif, while the GFP, Ntrunc, and C constructs did not. The FL and His-MBD TRP47 constructs exhibited strong nuclear localization and the N, TRC, and K71R TRP47 constructs exhibited diffuse cytoplasmic and nuclear localization. Exclusively cytoplasmic localization was observed with the Ntrunc, C, No Motif, and K49R mutant TRP47 constructs. The GFP control construct, which was observed only in the cytoplasm, was used as a negative control for nuclear localization. These results, with the exception of the K49R mutant TRP47 construct, are consistent with the original hypothesis that the MYND-binding motif is responsible for TRP47 nuclear localization. The inability of the K49R mutant TRP47 construct to enter or be retained in the nucleus could be due to a number of factors, such as unmasking of a nuclear export signal or loss of an important protein-protein interaction.

Confirmation of TRP47 DNA-binding activity and identification of functional domain

Two of the TRPs identified thus far (TRP120 and TRP32) are able to bind human genomic DNA, and do so via their internal tandem repeat domains. To determine if TRP47 follows this trend, a purified recombinant full-length and 5 tandem repeat (5TR) GST-

tagged TRP47 constructs were used in an electrophoretic mobility shift assay (EMSA) with biotin-labeled sheared human genomic DNA. Signal detection in this assay is based on binding of biotin by horseradish peroxidase-conjugated streptavidin; biotin-labeled DNA in complex with protein travels more slowly during electrophoresis, creating a shifted band. Unlabeled DNA-protein complexes also travel more slowly, but a band cannot be detected during visualization because no biotin-labeled gDNA is present in the complex. When an EMSA was performed with 4 µg of TRP47 5TR, formation of a protein-DNA complex was observed (Fig. 8A, lane 2; band height indicated by arrow) and addition of unlabeled competitor gDNA prevented detection of a band (Fig. 8A, lane 3), indicating that TRP47 specifically interacts with DNA. Addition of anti-TRP47 antibody resulted in loss of this band (Fig. 8A, lane 4), most likely due to interference with the ability of TRP47 to bind DNA. Band intensity was gradually restored when antibody concentration was decreased (Fig. 8A, lanes 5-7), confirming that TRP47 is present in the protein-DNA complex. An EMSA using two different amounts of full-length TRP47 (Fig. 8B, lanes 2 and 3; 4 µg and

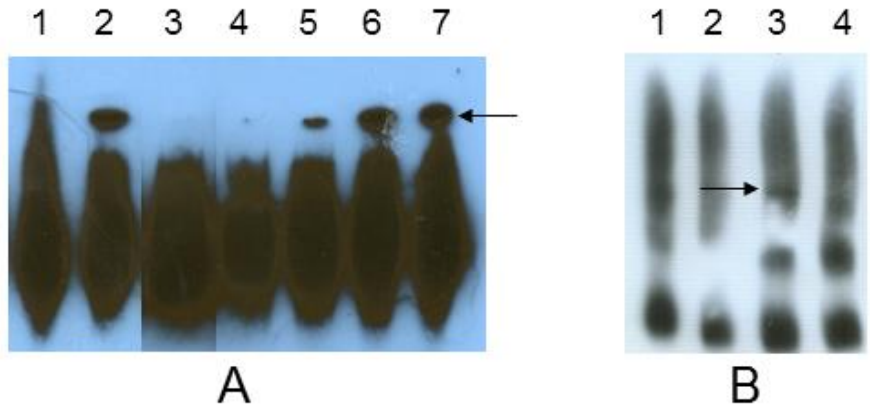

Figure 8: TRP47 DNA-binding activity. (A) EMSA with TRP47 5TR construct. Lane 1 is biotin-labeled gDNA only (negative control), lane 2 is biotin-labeled gDNA and TRP47 5TR, lane 3 is with the addition of unlabeled competitor in 500x excess of labeled gDNA, and lanes 4-7 are with the addition of anti-TRP47 antibody at dilutions of 1:20, 1:50, 1:100, and 1:200 respectively. Arrow indicates height of shifted band. (B) EMSA with full-length TRP47 construct. Lane 1 is biotin-labeled gDNA only (negative control), lanes 2 and 3 contain 4 µg and 8 µg of protein, respectively, and labeled gDNA, and lane 4 is labeled gDNA and 8 µg of protein with the addition of unlabeled competitor gDNA. Arrow indicates height of shifted band.

8 μg) showed formation of a protein-DNA complex (band height indicated by arrow) that could not be detected after the addition of unlabeled competitor gDNA (Fig. 8B, lane 4). In this study, the TRP47 tandem repeat region was confirmed as the functional domain for DNA binding. The amount of full-length protein required to visualize a band was twice the amount required for the tandem repeat construct, suggesting that full-length TRP47 has lower affinity for DNA than the tandem repeat region.

DISCUSSION

Nucleomodulins are an emerging class of bacterial effectors and invaluable tools for pathogens with small genomes such as *E. chaffeensis*. The TRPs are a family of multifunctional secreted proteins with diverse roles in infection. TRP47 was previously shown to interact with thirty-three host proteins, including PCGF5, Fyn, and PTPN2, which were investigated in this study. Although these TRP47-interacting host proteins are involved in diverse host cell functions, the majority participate in host cell signaling, transcriptional regulation, and vesicle trafficking (23). In this study, we attempted to gain further insight into the role of TRP47 in the host cell and to clarify the molecular basis for the multifarious interactions that occur with host proteins.

In this study, TRP47 was confirmed as a target for polySUMOylation through *in vivo* and *in vitro* methods. This is only the third instance in the literature that a bacterial effector has been shown to undergo SUMOylation, and the first time that SUMOylation of native protein has been conclusively demonstrated during infection. *Ehrlichia chaffeensis* TRP120 was the first described example of a SUMOylated bacterial effector, and the *Anaplasma phagocytophilum* secreted effector AmpA was the second (54). However, the results of the AmpA study are somewhat unclear, as the authors were unable to demonstrate *in vitro* SUMOylation of recombinant AmpA or present convincing evidence that AmpA is polySUMOylated *in vivo*. Previously, it was shown that SUMOylation of TRP120 mediated interaction with diverse host proteins containing SUMO-interacting motifs (SIMs), including PCGF5 (15). TRP47 has also been demonstrated to interact with PCGF5, along with several other SIM-containing proteins, suggesting that, like TRP120, TRP47 may utilize SUMOylation to facilitate interaction with host proteins. When the most strongly interacting TRP47 binding partners were analyzed with the program GPS-SUMO, predicted SIMs were identified in all three host proteins included in this study (PCGF5, Fyn, and PTPN2). Despite this, this study was not able to confirm SUMO-dependent

interactions using IFA or traditional co-immunoprecipitation (data not shown). It is possible, however, that other more sensitive methods such as chemiluminescent co-immunoprecipitation might be able to detect SUMO-dependent interaction. Even if SUMOylation is not required for TRP47-host protein interactions, its nature as a multifunctional modification presents a wide variety of other possible roles in the function of TRP47. SUMOylation also has been associated with regulation of protein localization, modulation of the apoptosis response, protein stability, and transcription factor function, any of which could potentially play an important role in TRP47 function.

When the requirements for TRP47 nuclear localization were examined, it was found that constructs containing a variant eukaryotic MYND-binding motif, which is associated with nuclear localization of members of the Hsp90 complex, were trafficked to the nucleus when ectopically expressed. Constructs not possessing this domain, or in which the N-terminal lysine 49 residue was mutated to arginine, were only observed in the cytoplasm. The identification of a traditionally eukaryotic motif as the domain responsible for TRP47 nuclear localization is unsurprising, given the demonstrated propensity of *E. chaffeensis* for subversion of host mechanisms and pathways. However, the inability of the K49R mutant to enter (or remain inside) the nucleus is intriguing. K49 was identified as a potential ubiquitination site by the BDM-PUB algorithm (Prediction of Ubiquitination site with Bayesian Discriminant Method), which bases its predictions on previously published ubiquitination sites. Ubiquitination has been shown to regulate the nuclear localization of p53, the transcription factor MyoD, and the lipogenic enzyme cytidylyltransferase through targeting ubiquitinated protein for degradation and masking nuclear import or export signals (55-57). It is possible that ubiquitination of K49 serves as a nuclear retention signal or masks a nuclear export signal found in the full-length protein, in which case both the MYND-binding motif and the putative ubiquitination site would need to be present in order for TRP47 to remain in the nucleus. However, this would not explain why one construct not containing the site (tandem repeat-C-terminal) still exhibited nuclear localization.

MonoSUMOylation is associated with nuclear localization, and although unlikely, it is also possible that K49 is a secondary SUMOylation site acting in concert with the MYND-binding domain to promote nuclear localization. Mutation of this lysine residue also abrogated colocalization with PTPN2, suggesting that interaction between TRP47 and PTPN2 may be mediated by ubiquitination.

An earlier study utilizing the same full-length TRP47 construct used here demonstrated primarily cytoplasmic localization of ectopically expressed TRP47, although a small amount of nuclear TRP47 was still observed (23). This could be due to differences in imaging methods; wide-field epifluorescence microscopy was used in this study, whereas confocal laser-scanning microscopy was used previously. An underlying principle of confocal microscopy is optical sectioning, in which images are acquired at specific depths in a specimen. This is in contrast to epifluorescence microscopy, which detects the total fluorescence emitted by a specimen. It is possible that the image depths chosen to best illustrate cytoplasmic colocalization of TRP47 and host proteins by confocal microscopy were not representative of actual TRP47 nuclear localization.

Previously, TRP47 was observed in the nuclei of infected host cells; however, this was never studied in depth. TRP47 is only expressed by DCs, which enter the host cell and quickly transition to RCs. Any TRP47 found in the host cell nucleus or cytoplasm would have to have been produced and secreted by *E. chaffeensis* before the DC-to-RC transition. This study sought to examine the temporal variation in subcellular localization of TRP47 during infection. Infected cells were viewed by immunofluorescence microscopy after 24, 48, and 72 hours and the distribution of TRP47 in host cell nuclei and cytoplasm was compared. TRP47-staining morulae were observed at all three time points, with the least number of morulae observed at 72 hours. A comparison of fluorescent signal line intensity profiles indicated that nuclear TRP47 signal was highest at 24 and 48 hpi and appreciably lower at 72 hpi, which is consistent with the previous body of knowledge regarding regulation of its transcription (22). A notable difference between the 24 hpi and 48 hpi time

points was the ratio of cytoplasmic to nuclear TRP47 signal. At 24 hpi, little TRP47 signal was detected in the cytoplasm as compared to the nucleus, whereas at 48 hpi, this proportion increased considerably. This would suggest that TRP47 enters the nucleus quickly after infection, then a small proportion slowly exits. The mechanism for this change in localization is unclear, but it could be due to nuclear export triggered by interactions with host proteins or post-translational modification.

When TRP47-DNA interactions were analyzed by EMSA, it was shown that TRP47 is able to bind human genomic DNA and that the tandem repeat domain is responsible for this interaction. N-terminal and C-terminal TRP47 were also used in an EMSA and failed to produce a band shift (data not shown). The decreased affinity of full-length TRP47 for DNA as compared to the tandem repeat construct indicates that there is probably an inhibitory domain in the N-terminal or C-terminal portion of the protein. It is possible that the conformation of the tandem repeat domain is more amenable to DNA binding when it is expressed on its own. Full-length TRP47 may require a post-translational modification or interaction with another protein to adopt an ideal conformation for DNA binding. The EMSAs using full-length and tandem repeat TRP47 were performed with two different preparations of biotin-labeled gDNA, which most likely caused the inconsistency in free gDNA migration observed between the two assays. Although a classical supershift of the TRP47-DNA complex with addition of anti-TRP47 antibody was not observed in this study, a gradual inhibition of DNA-binding activity, indicated by decrease in band intensity, occurred with increasing concentration of antibody. This phenomenon has been described in the literature for several transcription factors (58-61) and may be due to competition between the antibody and DNA to bind the TRP47 tandem repeat region. The antibody used in the supershift assay was directed against a molecularly determined epitope in the tandem repeat domain, which supports this idea. It is intriguing that the tandem repeat domain is involved in DNA binding. This is similar to TRP120, which also uses its tandem repeat domain for interaction with host genomic DNA (19), and suggests that tandem

repeats of all the TRP proteins may function as DNA-binding domains. This is similar to the TAL family of *Xanthomonas* effectors, which also interact with host DNA via internal tandem repeat DNA-binding domains (46). These TAL effectors act as transcription factors and regulate the transcription of host genes important to infection. The ability of TRP47 to enter the nucleus and bind human genomic DNA is very convincing evidence that it may act as a host transcription factor, similarly to TRP120 and the TAL effectors.

REFERENCES

1. **Paddock CD, Childs JE.** *Ehrlichia chaffeensis*: a prototypical emerging pathogen. Clin Microbiol Rev. 2003;16(1):37-64.

2. **Yunker CE.** Heartwater in sheep and goats: a review. Onderstepoort Journal of Veterinary Research. 1996;63(2):159-70.

3. **Olano JP, Wen G, Feng HM, McBride JW, Walker DH.** Histologic, serologic, and molecular analysis of persistent ehrlichiosis in a murine model. The American journal of pathology. 2004;165(3):997-1006. doi: 10.1016/S0002-9440(10)63361-5. PubMed PMID: 15331423; PMCID: 1618610.

4. **Saito TB, Walker DH.** A tick vector transmission model of monocytotropic ehrlichiosis. J Infect Dis. 2015;212(6):968-77. Epub 2015/03/05. doi: 10.1093/infdis/jiv134. PubMed PMID: 25737562; PMCID: Pmc4548458.

5. **Safdar N, Love RB, Maki DG.** Severe *Ehrlichia chaffeensis* infection in a lung transplant recipient: a review of ehrlichiosis in the immunocompromised patient. Emerg Infect Dis. 2002;8(3):320-3. doi: 10.3201/eid0803.010249. PubMed PMID: 11927032; PMCID: Pmc2732464.

6. **Fishbein DB, Dawson JE, Robinson LE.** Human ehrlichiosis in the United States, 1985 to 1990. Annals of Internal Medicine. 1994;120(9):736-43. PubMed PMID: 8147546.

7. **Hamburg BJ, Storch GA, Micek ST, Kollef MH.** The importance of early treatment with doxycycline in human ehrlichiosis. Medicine (Baltimore). 2008;87(2):53-60.

8. **Marshall GS, Jacobs RF, Schutze GE, Paxton H, Buckingham SC, DeVincenzo JP, Jackson MA, San Joaquin VH, Standaert SM, Woods CR.** *Ehrlichia chaffeensis* seroprevalence among children in the southeast and south-central regions of the United States. Archives of Pediatrics & Adolescent Medicine. 2002;156(2):166-70. Epub 2002/01/30. PubMed PMID: 11814379.

9. **Yevich SJ, Sanchez JL, DeFraites RF, Rives CC, Dawson JE, Uhaa IJ, Johnson BJ, Fishbein DB.** Seroepidemiology of infections due to spotted fever group rickettsiae and *Ehrlichia* species in military personnel exposed in areas of the United States where such infections are endemic. Journal of Infectious Diseases. 1995;171(5):1266-73.

10. **Popov VL, Chen SM, Feng HM, Walker DH.** Ultrastructural variation of cultured *Ehrlichia chaffeensis*. JMedMicrobiol. 1995;43(6):411-21.

11. **Rikihisa Y.** *Anaplasma phagocytophilum* and *Ehrlichia chaffeensis*: subversive manipulators of host cells. Nature Reviews Microbiology. 2010;8(5):328-39. doi: 10.1038/nrmicro2318. PubMed PMID: 20372158.

12. **Zhang J-Z, Sinha M, Luxon BA, Yu X-J.** Survival strategy of obligately intracellular *Ehrlichia chaffeensis*: novel modulation of immune response and host cell cycles. Infection and Immunity. 2004;72(1):498-507. doi: 10.1128/iai.72.1.498-507.2004.

13. **Lin M, Rikihisa Y.** *Ehrlichia chaffeensis* downregulates surface Toll-like receptors 2/4, CD14 and transcription factors PU.1 and inhibits lipopolysaccharide activation of NF-kappa B, ERK 1/2 and p38 MAPK in host monocytes. Cellular Microbiology. 2004;6(2):175-86. PubMed PMID: 14706103.

14. **Kumagai Y, Cheng Z, Lin M, Rikihisa Y.** Biochemical activities of three pairs of *Ehrlichia chaffeensis* two-component regulatory system proteins involved in inhibition of lysosomal fusion. Infection and Immunity. 2006;74(9):5014-22.

15. **Dunphy PS, Luo T, McBride JW.** *Ehrlichia chaffeensis* exploits host SUMOylation pathways to mediate effector-host interactions and promote intracellular survival. Infection and immunity. 2014;82(10):4154-68. doi: 10.1128/IAI.01984-14. PubMed PMID: 25047847; PMCID: 4187855.

16. **Wakeel A, den Dulk-Ras A, Hooykaas PJ, McBride JW.** *Ehrlichia chaffeensis* tandem repeat proteins and Ank200 are type 1 secretion system substrates related to the repeats-in-toxin exoprotein family. Frontiers in Cellular and Infection Microbiology. 2011;1:22. doi: 10.3389/fcimb.2011.00022. PubMed PMID: 22919588; PMCID: 3417381.

17. **Kuriakose JA, Zhang X, Luo T, McBride JW.** Molecular basis of antibody mediated immunity against *Ehrlichia chaffeensis* involves species-specific linear epitopes in tandem repeat proteins. Microbes and Infection / Institut Pasteur. 2012;14(12):1054-63. doi: 10.1016/j.micinf.2012.05.012. PubMed PMID: 22658957; PMCID: 3445803.

18. **Luo T, Kuriakose JA, Zhu B, Wakeel A, McBride JW.** *Ehrlichia chaffeensis* TRP120 interacts with a diverse array of eukaryotic proteins involved in transcription, signaling, and cytoskeleton organization. Infection and Immunity. 2011;79(11):4382-91. doi: 10.1128/IAI.05608-11. PubMed PMID: 21859857; PMCID: 3257936.

19. **Zhu B, Kuriakose JA, Luo T, Ballesteros E, Gupta S, Fofanov Y, McBride JW.** *Ehrlichia chaffeensis* TRP120 binds a G+C-rich motif in host cell DNA and exhibits eukaryotic transcriptional activator function. Infection and Immunity. 2011;79(11):4370-81. doi: 10.1128/IAI.05422-11. PubMed PMID: 21859854; PMCID: 3257946.

20. **Doyle CK, Nethery KA, Popov VL, McBride JW.** Differentially expressed and secreted major immunoreactive protein orthologs of *Ehrlichia canis* and *E. chaffeensis* elicit early antibody responses to epitopes on glycosylated tandem repeats. Infection and Immunity. 2006;74(1):711-20. doi: 10.1128/IAI.74.1.711-720.2006. PubMed PMID: 16369028; PMCID: 1346619.

21. **Wakeel A, Zhang X, McBride JW.** Mass spectrometric analysis of *Ehrlichia chaffeensis* tandem repeat proteins reveals evidence of phosphorylation and absence of glycosylation. PloS One. 2010;5(3):e9552. doi: 10.1371/journal.pone.0009552. PubMed PMID: 20209062; PMCID: 2832021.

22. **Kuriakose JA, Miyashiro S, Luo T, Zhu B, McBride JW.** *Ehrlichia chaffeensis* transcriptome in mammalian and arthropod hosts reveals differential gene expression and post transcriptional regulation. PloS One. 2011;6(9):e24136. doi: 10.1371/journal.pone.0024136. PubMed PMID: 21915290; PMCID: 3167834.

23. **Wakeel A, Kuriakose JA, McBride JW.** An *Ehrlichia chaffeensis* tandem repeat protein interacts with multiple host targets involved in cell signaling, transcriptional regulation, and vesicle trafficking. Infection and Immunity. 2009;77(5):1734-45. doi: 10.1128/IAI.00027-09. PubMed PMID: 19273555; PMCID: 2681728.

24. **Saurin AJ, Borden KL, Boddy MN, Freemont PS.** Does this have a familiar RING? Trends in Biochemical Sciences. 1996;21(6):208-14. Epub 1996/06/01. PubMed PMID: 8744354.

25. **Coyne CB, Bergelson JM.** Virus-induced Abl and Fyn kinase signals permit coxsackievirus entry through epithelial tight junctions. Cell. 2006;124(1):119-31. Epub 2006/01/18. doi: 10.1016/j.cell.2005.10.035. PubMed PMID: 16413486.

26. **Lee EH, Rikihisa Y.** Protein kinase A-mediated inhibition of gamma interferon-induced tyrosine phosphorylation of Janus kinases and latent cytoplasmic transcription factors in human monocytes by *Ehrlichia chaffeensis*. Infection and Immunity. 1998;66(6):2514-20.

27. **Stuible M, Doody KM, Tremblay ML.** PTP1B and TC-PTP: regulators of transformation and tumorigenesis. Cancer Metastasis Reviews. 2008;27(2):215-30. Epub 2008/02/01. doi: 10.1007/s10555-008-9115-1. PubMed PMID: 18236007.

28. **Wang YE, Pernet O, Lee B.** Regulation of the nucleocytoplasmic trafficking of viral and cellular proteins by ubiquitin and small ubiquitin-related modifiers. Biology of the cell / under the auspices of the European Cell Biology Organization. 2012;104(3):121-38. doi: 10.1111/boc.201100105. PubMed PMID: 22188262; PMCID: 3625690.

29. **Mukhopadhyay D, Dasso M.** Modification in reverse: the SUMO proteases. Trends in Biochemical Sciences. 2007;32(6):286-95. Epub 2007/05/15. doi: 10.1016/j.tibs.2007.05.002. PubMed PMID: 17499995.

30. **Zhao Q, Xie Y, Zheng Y, Jiang S, Liu W, Mu W, Liu Z, Zhao Y, Xue Y, Ren J.** GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs. Nucleic Acids Research. 2014;42(Web Server issue):W325-30. Epub 2014/06/02. doi: 10.1093/nar/gku383. PubMed PMID: 24880689; PMCID: Pmc4086084.

31. **Gareau JR, Lima CD.** The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. Nature Reviews Molecular Cell Biology. 2010;11(12):861-71. doi: 10.1038/nrm3011. PubMed PMID: 21102611; PMCID: 3079294.

32. **Tatham MH, Geoffroy MC, Shen L, Plechanovova A, Hattersley N, Jaffray EG, Palvimo JJ, Hay RT.** RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. Nature Cell Biology. 2008;10(5):538-46. Epub 2008/04/15. doi: 10.1038/ncb1716. PubMed PMID: 18408734.

33. **Harder Z, Zunino R, McBride H. Sumo1 conjugates mitochondrial substrates and participates in mitochondrial fission.** Current Biology : CB. 2004;14(4):340-5. Epub 2004/02/20. doi: 10.1016/j.cub.2004.02.004. PubMed PMID: 14972687.

34. **Gill G.** SUMO and ubiquitin in the nucleus: different functions, similar mechanisms? Genes & Development. 2004;18(17):2046-59. doi: 10.1101/gad.1214604. PubMed PMID: 15342487.

35. **Priyanka, Kotiya D, Rana M, Subbarao N, Puri N, Tyagi RK.** Transcription regulation of nuclear receptor PXR: role of SUMO-1 modification and NDSM in receptor function. Molecular and Cellular Endocrinology. 2015. Epub 2015/11/10. doi: 10.1016/j.mce.2015.11.001. PubMed PMID: 26549688.

36. **Chang PC, Kung HJ.** SUMO and KSHV Replication. Cancers. 2014;6(4):1905-24. Epub 2014/10/01. doi: 10.3390/cancers6041905. PubMed PMID: 25268162; PMCID: Pmc4276950.

37. **Gonzalez-Santamaria J, Campagna M, Garcia MA, Marcos-Villar L, Gonzalez D, Gallego P, Lopitz-Otsoa F, Guerra S, Rodriguez MS, Esteban M, Rivas C.** Regulation of vaccinia virus E3 protein by small ubiquitin-like modifier proteins. Journal of Virology. 2011;85(24):12890-900. Epub 2011/10/01. doi: 10.1128/jvi.05628-11. PubMed PMID: 21957283; PMCID: Pmc3233166.

38. **Wimmer P, Schreiner S, Dobner T.** Human pathogens and the host cell SUMOylation system. Journal of Virology. 2012;86(2):642-54. doi: 10.1128/JVI.06227-11. PubMed PMID: 22072786; PMCID: 3255802.

39. **Smith MC, Box AC, Haug JS, Lane WS, Davido DJ.** A Phospho-SIM in the Antiviral Protein PML is Required for Its Recruitment to HSV-1 Genomes. Cell. 2014;3(4):1131-58. Epub 2014/12/17. doi: 10.3390/cells3041131. PubMed PMID: 25513827; PMCID: Pmc4276917.

40. **Murata T, Hotta N, Toyama S, Nakayama S, Chiba S, Isomura H, Ohshima T, Kanda T, Tsurumi T.** Transcriptional repression by sumoylation of Epstein-Barr virus BZLF1 protein correlates with association of histone deacetylase. The Journal of Biological Chemistry. 2010;285(31):23925-35. Epub 2010/06/03. doi: 10.1074/jbc.M109.095356. PubMed PMID: 20516063; PMCID: Pmc2911316.

41. **Cornelis GR, Denecker G.** Yersinia lead SUMO attack. Nature Medicine. 2001;7(1):21-3. Epub 2001/01/03. doi: 10.1038/83298. PubMed PMID: 11135606.

42. **Citro S, Chiocca S.** Listeria monocytogenes: a bacterial pathogen to hit on the SUMO pathway. Cell Research. 2010;20(7):738-40. Epub 2010/06/10. doi: 10.1038/cr.2010.76. PubMed PMID: 20531377.

43. **Jans DA, Chan CK, Huebner S.** Signals mediating nuclear targeting and their regulation: application in drug delivery. Medicinal Research Reviews. 1998;18(4):189-223. Epub 1998/07/17. PubMed PMID: 9664290.

44. **Dinkel H, Michael S, Weatheritt RJ, Davey NE, Van Roey K, Altenberg B, Toedt G, Uyar B, Seiler M, Budd A, Jodicke L, Dammert MA, Schroeter C, Hammer M, Schmidt T, Jehl P, McGuigan C, Dymecka M, Chica C, Luck K, Via A, Chatr-Aryamontri A, Haslam N, Grebnev G, Edwards RJ, Steinmetz MO, Meiselbach H, Diella F, Gibson TJ.** ELM--the database of eukaryotic linear motifs. Nucleic Acids Research. 2012;40:D242-51. Epub 2011/11/24. doi: 10.1093/nar/gkr1064. PubMed PMID: 22110040; PMCID: Pmc3245074.

45. **Song D, Li LS, Heaton-Johnson KJ, Arsenault PR, Master SR, Lee FS.** Prolyl Hydroxylase Domain Protein 2 (PHD2) binds a Pro-Xaa-Leu-Glu motif, linking it to the heat shock protein 90 pathway. The Journal of Biological Chemistry. 2013;288(14):9662-74. doi: 10.1074/jbc.M112.440552. PubMed PMID: 23413029; PMCID: Pmc3617269.

46. **Bierne H, Cossart P.** When bacteria target the nucleus: the emerging family of nucleomodulins. Cellular Microbiology. 2012;14(5):622-33. Epub 2012/02/01. doi: 10.1111/j.1462-5822.2012.01758.x. PubMed PMID: 22289128.

47. **Li L, Atef A, Piatek A, Ali Z, Piatek M, Aouida M, Sharakuu A, Mahjoub A, Wang G, Khan S, Fedoroff NV, Zhu JK, Mahfouz MM.** Characterization and DNA-binding specificities of *Ralstonia* TAL-like effectors. Molecular Plant. 2013;6(4):1318-30. doi: 10.1093/mp/sst006. PubMed PMID: 23300258; PMCID: Pmc3716395.

48. **Gao X, Zhou X, Gulari E.** Light directed massively parallel on-chip synthesis of peptide arrays with t-Boc chemistry. Proteomics. 2003;3(11):2135-41. doi: 10.1002/pmic.200300597. PubMed PMID: 14595812.

49. **Pellois JP, Wang W, Gao X.** Peptide synthesis based on t-Boc chemistry and solution photogenerated acids. J Comb Chem. 2000;2(4):355-60. PubMed PMID: 10891103.

50. **Pellois JP, Zhou X, Srivannavit O, Zhou T, Gulari E, Gao X.** Individually addressable parallel peptide synthesis on microchips. Nat Biotechnol. 2002;20(9):922-6. doi: 10.1038/nbt723. PubMed PMID: 12134169.

51. **Klug H, Xaver M, Chaugule Viduth K, Koidl S, Mittler G, Klein F, Pichler A.** Ubc9 sumoylation controls SUMO chain formation and meiotic synapsis in *Saccharomyces cerevisiae.* Molecular Cell. 2013;50(5):625-36. doi: http://dx.doi.org/10.1016/j.molcel.2013.03.027.

52. **Gonzalez-Santamaria J, Campagna M, Ortega-Molina A, Marcos-Villar L, de la Cruz-Herrera CF, Gonzalez D, Gallego P, Lopitz-Otsoa F, Esteban M, Rodriguez MS, Serrano M, Rivas C.** Regulation of the tumor suppressor PTEN by SUMO. Cell Death & Disease. 2012;3:e393. doi: http://www.nature.com/cddis/journal/v3/n9/suppinfo/cddis2012135s1.html.

53. **Franco M, Seyfried NT, Brand AH, Peng J, Mayor U.** A novel strategy to isolate ubiquitin conjugates reveals wide role for ubiquitination during neural development. Molecular & Cellular Proteomics : MCP. 2011;10(5):M110.002188. Epub 2010/09/24. doi: 10.1074/mcp.M110.002188. PubMed PMID: 20861518; PMCID: Pmc3098581.

54. **Beyer AR, Truchan HK, May LJ, Walker NJ, Borjesson DL, Carlyon JA.** The *Anaplasma phagocytophilum* effector AmpA hijacks host cell SUMOylation. Cellular Microbiology. 2015;17(4):504-19. Epub 2014/10/14. doi: 10.1111/cmi.12380. PubMed PMID: 25308709.

55. **Lee J, Gu W.** The multiple levels of regulation by p53 ubiquitination. Cell Death and Differentiation. 2010;17(1):86-92. doi: 10.1038/cdd.2009.77. PubMed PMID: 19543236; PMCID: Pmc3690487.

56. **Lingbeck JM, Trausch-Azar JS, Ciechanover A, Schwartz AL.** Determinants of nuclear and cytoplasmic ubiquitin-mediated degradation of MyoD. The Journal of Biological Chemistry. 2003;278(3):1817-23. Epub 2002/10/25. doi: 10.1074/jbc.M208815200. PubMed PMID: 12397066.

57. **Chen BB, Mallampalli RK.** Masking of a nuclear signal motif by monoubiquitination leads to mislocalization and degradation of the regulatory enzyme cytidylyltransferase. Molecular and Cellular Biology. 2009;29(11):3062-75. Epub 2009/04/01. doi: 10.1128/mcb.01824-08. PubMed PMID: 19332566; PMCID: Pmc2682000.

58. **Wang J, Gu Z, Ni P, Qiao Y, Chen C, Liu X, Lin J, Chen N, Fan Q.** NF-kappaB P50/P65 hetero-dimer mediates differential regulation of CD166/ALCAM expression via interaction with micoRNA-9 after serum deprivation, providing evidence for a novel negative auto-regulatory loop. Nucleic Acids Research. 2011;39(15):6440-55. Epub 2011/05/17. doi: 10.1093/nar/gkr302. PubMed PMID: 21572107; PMCID: Pmc3159468.

59. **Rauch T, Zhong X, Pfeifer GP, Xu X.** 53BP1 is a positive regulator of the BRCA1 promoter. Cell Cycle. 2005;4(8):1078-83. Epub 2005/06/23. PubMed PMID: 15970701.

60. **Watanabe M, Nakajima S, Ohnuki K, Ogawa S, Yamashita M, Nakayama T, Murakami Y, Tanabe K, Abe R.** AP-1 is involved in ICOS gene expression downstream of TCR/CD28 and cytokine receptor signaling. Eur J Immunol. 2012;42(7):1850-62. Epub 2012/05/16. doi: 10.1002/eji.201141897. PubMed PMID: 22585681.

61. **Fralix KD, Zhao S, Venkatasubbarao K, Freeman JW.** Rap1 reverses transcriptional repression of TGF-beta type II receptor by a mechanism involving AP-1 in the human pancreatic cancer cell line, UK Pan-1. Journal of Cellular Physiology. 2003;194(1):88-99. Epub 2002/11/26. doi: 10.1002/jcp.10192. PubMed PMID: 12447993.

VITA

Sarah L. Milligan was born on February 22, 1992 to David Schmidt and Megan Peters in Lawrenceville, Georgia. After graduating from high school in 2010 with her future husband, she attended Texas A&M University in College Station, Texas. While at Texas A&M, she worked as a research assistant in the Department of Biology and wrote an undergraduate thesis on site-directed mutagenesis of a *Clostridium difficile* protein-coding gene. In May of 2013, she graduated with a Bachelor's degree in Biology. She began graduate school at the University of Texas Medical Branch in the fall of 2013 and married Larry T. Milligan the following spring.

During her graduate education at the University of Texas Medical Branch, Sarah presented her research at several regional and national scientific conferences. Her presentation at the 2014 meeting of the Texas Branch of the American Society for Microbiology was given the Samuel L. Kaplan Award for outstanding graduate student poster in gram-negative microbiology. She was an active member in the UTMB Graduate Student Organization, serving as the Experimental Pathology Department Representative from the spring of 2014 to the spring of 2015. She also mentored a student during the summer of 2015 in the Biomedical Research Training for High School Students program.

Permanent address: 6315 Central City Blvd., Apartment 818, Galveston, Texas 77551

This dissertation was typed by Sarah L. Milligan.