Shigella Effector IpaH9.8 Modulates a Host Ubiquitin-Like Pathway and Lysosomal Biogenesis

by

Lucas Jarche

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
November 20, 2018

© Copyright by Lucas Jarche, 2018

Table of Contents

List of Tables

List of Figures

Abstract

The bacterium *Shigella* causes bacterial dysentery in developing countries and represents a persistent burden to healthcare systems due to an increase in antibiotic-resistant strains and lack of a vaccine strategy. *Shigella* encodes a type III secretion system to inject effector proteins into the eukaryotic cytosol to hijack and disrupt host processes. One of these effectors is IpaH9.8, a Novel E3 ubiquitin Ligase (NEL) that can target host proteins for post-translational modification with ubiquitin. Cellular targets for IpaH9.8 are unknown despite more than a decade of searching. My results are consistent with a model where IpaH9.8 interacts with a ubiquitin-like protein, UFM1. By co-immunoprecipitation, mass spectrometry, and *in vitro* UFMylation I have identified Galectin-7, a sugar-binding protein, as an IpaH9.8 target. Additionally, I have characterized another IpaH9.8 interactor, ZKSCAN3, a negative regulator of lysosomal biogenesis, and have provided preliminary evidence that *Shigella* manipulate host lysosomes through ubiquitination of ZKSCAN3.

List of Abbreviations Used

4E-BP	4E-binding protein
ABIN-1	A20 binding inhibitor of NF- κB
AMP	Adenosine monophosphate
AMPK	AMP-activated protein kinase
AraC	Arabinose C
ASC1	Activator signal cointegrator 1
ATF	Activating transcription factor
ATG	Autophagy-related protein
ATP	Adenosine triphosphate
ATP6V1A	ATPase H+ Transporting V1 Subunit A1
AvrPtoB	Avirulence Pto protein kinase B
BCV	Bacteria-containing vacuole
BEL	Bacterial-encoded E3 Ligase
BioID	Biotin Identification
BL21	*E. coli* strain B, designation BL21
BRCA1	Breast cancer-asscociated 1
BSA	Bovine serum albumin
CD	Catalytically-dead
CDK5RAP3	Cyclin-dependent kinase regulatory subunit-associated protein 3
CID	Collision-induced dissociation
CLEAR	Coordinated lysosomal expression and regulation
CMV	Cytomegalovirus
CORE	Centralized Operation of Research Equipment
cRAP	common Repository of Adventitious Proteins
CRL	Cullin RING Ligase
CTSA	Cathepsin A
DAPI	4′,6-diamidino-2-phenylindole
ddH2O	Double-distilled H_2O
DDRGK1	DDRGK domain-containing protein 1
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
DUB	Deubiquitinase
E. coli	*Escherichia coli*
E6-AP	E6-Associated protein
EDTA	Ethylenediaminetetraacetic acid
EHEC	Enterohemorrhagic Escherichia coli

ER	Endoplasmic Reticulum
FAT10	HLA-F adjacent transcript 10
FBS	Fetal bovine serum
GALS7	Galectin-7
GBP	Guanylate binding protein
GFP	Green fluorescent protein
GTP	Guanosine triphosphate
H2B	Histone 2B
HA	Hemagglutinin
HCD	Higher-energy collisional dissociation
HECT	Homologous to the E6-AP carboxyl terminus
HEK	Human embryonic kidney
HNRNPF	Heterogeneous Nuclear Ribonucleoprotein F
HOIL	Heme-oxidized IRP2 ubiquitin ligase
HOIP	HOIL-interacting protein
HRAS	Human Ras protein
HRP	Horseradish peroxidase
IcsB	Intracellular spread
IF	Immunofluorescence
IκB	NF-κB inhibitor
IKK	IκB Kinase
IP Buffer	Immunoprecipitation buffer
Ipa(B/C/H)	Invasion plasmid antigen
IpgC	Invasion plasmid gene C
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRE1α	Inositol-requiring enzyme 1α
ITCH	Itchy (Mouse Homolog) E3 Ubiquitin Protein Ligase
JNK1	c-Jun N-terminal kinases
LAMP1	Lysosomal-associated membrane protein 1
LB	Lysogeny broth
LC-MS/MS	Liquid chromatography tandem mass-spectrometry
LC3	Light chain 3B
LRR	Leucine-rich repeat
LRSAM1	Leucine rich repeat and sterile alpha motif containing 1
LUBAC	Linear ubiquitin chain assembly complex
M cell	Microfold cell
MAPK	Mitogen-activated protein kinase
MAPKK	Mitogen-activated protein kinase kinase
mTOR	Mechanistic target of rapamycin
mTORC1	Mechanistic target of rapamycin complex 1

Mxi	Membrane expression of invasion plasmid antigens
Myc	Myelocytomatosis
MyD88	Myeloid differentiation primary response 88
NAG	N-actetylglucosaminidase
NCAM	Neural cell adhesion molecule
NDP52	Nuclear domain 10 protein 52
NEDD8	Neural precursor cell expressed developmentally down-regulated protein 8
NEL	Novel E3 ligase
NEMO	NF-κB essential modulator
NETN	NaCl, EDTA, Tris, Nonidet Buffer
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NleG/L	Non-locus-of-enterocyte-effacement-encoded protein
NLRC4	Nod-like receptor family CARD domain-containing protein 4
NLRP3	NACHT, LRR, and PYD domains-containing protein 3
Nod1	Nucleotide-binding oligomerization domain-containing protein 1
NONO	Non-POU domain-containing octamer-binding protein
NopM	Nodulation protein M
OD	Optical density
Osp(E/I/G/F)	Outer surface protein
PAMP	Pathogen-associated molecular pattern
PBS	Phosphate-buffered saline
PEI	Polyethylenimine
PERK	Protein kinase R-like endoplasmic reticulum kinase
PFA	Paraformaldehyde
PKC	Protein kinase C
PKN1	Protein kinase N1
PRR	Pattern recognition receptor
PTM	Post-translational modification
PVDF	Polyvinylidene difluoride
RBR	RING-between-RING
RCAD	Regulator of C53/LZAP and DDRGK1
Rel	Reticuloendotheliosis oncogene cellular homolog
RING	Really interesting new gene
RIPA buffer	Radioimmunoprecipitation buffer
RPL	Ribosomal protein L(arge)
RPS	Ribosomal protein S(mall)
SCF	SKP1 Cullin F Box protein complex
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SFPQ	Splicing factor proline- and glutamine-rich

SGSH	N-Sulfoglucosamine Sulfohydrolase
SidE	Substrate of Icm/Dot Transporter protein E
SlrP	*Salmonella* leucine-rich repeat protein
SMURF2	SMAD ubiquitination regulatory factor 2
SopA	*Salmonella* outer protein A
Spa	Surface presentation of antigens
SQSTM1	Sequestosome 1
SspH1/2	Salmonella secreted protein H
Ste7	Serine/Threonine-protein kinase 7
STRING	Search Tool for the Retrieval of Interacting Genes
StUbL	SUMO-targeted ubiquitin ligase
SUMF1	Sulfatase-modifying factor 1
SUMO	Small ubiquitin-like modifier
T3SS	Type III secretion system
TBK1	TRAF family member-associated NF-κB activator-binding kinase 1
TBST	Tris-buffered saline with tween-20
TFEB	Transcription factor EB
TLR	Toll-like receptor
TNFα	Tumour-necrosis factor alpha
TNFR	Tumour-necrosis factor receptor
TRAF	Tumour-necrosis factor receptor associated factors
TRIM(65/56)	Tripartite motif-containing protein
TRIP4	Thyroid hormone receptor interactor 4
U-box	UFD2-homology domain
Ub	Ubiquitin
UBA	Ubiquitin-like modifier activating enzyme
UBC	Ubiquitin conjugating enzyme
Ubl	Ubiquitin-like protein
UFBP1	UFM1-binding protein 1 with a PCI domain
UFC1	Ubiquitin-fold modifier-conjugating enzyme 1
UFL1	Ubiquitin-fold modifier ligating enzyme 1
UFM1	Ubiquitin-fold modifier 1
UFSP1/2	Ubiquitin-fold modifier-specific protease
ULK	Unc51-like autophagy activating kinase
UPR	Unfolded protein response
URM1	Ubiquitin-related modifier 1
v/v	volume by volume
VirA/F/B	Virulence factor A
WCL	Whole cell lysate
WIPI	WD-repeat protein interacting with phosphoinositides

WT	Wildtype
Xbp-1	X-box binding protein 1
XopD	*Xanthomonas* outer protein D
YopM	*Yersinia* outer membrane protein M
ZKSCAN3	Zinc finger with KRAB and SCAN domains 3

Acknowledgements

A big thank you to all members of the Rohde lab past and present who made this work possible. Especially to Adrian Herod, the only one who stuck around. Thanks to Julie Ryu and Angela Daurie who trained me in my early days in the lab, and to Jeremy Benjamin and Kaitlyn Tanner whose work I built upon.

Thanks to the McCormick lab including Carolyn Robinson, Eric Pringle, Ben Johnston, Mariel Kleer and Patrick Slaine. They are all much smarter than me and helped me out in some capacity along the way, especially Carolyn who built several of the tools I used in this thesis. I would be remiss without giving a big thank you to Craig McCormick, my co-supervisor who took me under his wing for a whole year. His time, thoughts, and casual comments helped to shape my thesis in an instrumental way.

Thank you to my committee members Dr. Zhenyu Cheng and Dr. Melanie Dobson who push me to be a better scientist each meeting and ask questions the rest of us forgot to think of. Thank you as well to Dr. Don Stoltz and Mary Ann Trevors for their help with Electron Microscopy, and Stephen Whitefield for training and advice on microscopy.

The biggest thank you to my supervisor John Rohde. The breadth of his knowledge seriously astounds me at times, as does his humour, helpfulness, Simpsons references, and casual, down-to-earth attitude in all things big and small. I would not have even considered this degree under a different supervisor.

Lastly, I want to thank my family, and my partner Emma for their support throughout the years, Andy, whose writing gets me out of bed in the morning, and Tupper Mixer Fridays that put me to bed at night.

Chapter 1: Introduction

The human body encodes approximately 20,000 protein coding genes (Ezkurdia et al., 2014), whereas the total proteome complexity is thought to exceed 1,000,000 unique protein variants (Harper & Bennett, 2016). This discrepancy arises from the fact that multiple proteins isoforms can be encoded from the same gene, and that many proteins are differentially modified after being translated. These so-called post-translational modifications (PTMs) do not result in entirely new proteins, but rather increase the functional complexity of the proteome and help regulate this diverse network of proteins.

1.1: Post-Translational Modifications

PTMs consist of a variety of functional groups or proteins that are covalently attached to already-formed proteins to change how these proteins function inside of the cell, including changes in localization, enzymatic activity, and stability. The exhaustive suite of PTMs comprises many that are poorly understood; however, comparisons among the best understood PTM systems identifies shared themes and paradigms between them. The top ten most represented PTMs in known proteomes are as follows: phosphorylation, acetylation, N-linked glycosylation, amidation, hydroxylation, methylation, O-linked glycosylation, ubiquitination, cyclization of glutamic acid, and sulfation (Khoury, Baliban, & Floudas, 2011).

Phosphorylation is the most abundant PTM and is catalyzed by a group of enzymes known as kinases. Kinases assist in the transfer of a phosphoryl group from a molecule of ATP (or another protein) to serine, threonine, or tyrosine residues in target proteins in eukaryotic cells (see review (Ubersax & Ferrell, 2007)). In prokaryotic cells,

phosphorylation also includes histidine phosphorylation, and involves receptors that recognize external stimuli and autophosphorylate a histidine residue before transferring the phosphate to an aspartate on another protein, the response regulator, to form a two-component system (Hoch, 2000). Phosphorylation can also occur, albeit rarely, on arginine, lysine, or cysteine residues (Ciesla, Fraczyk, & Rode, 2011; Sun et al., 2012). Protein phosphorylation is reversible, a role that is handled by phosphatases. Both kinases and phosphatases can be specific, (de)phosphorylating a handful of sites each, or (de)phosphorylating several hundred sites (Ptacek et al., 2005). Phosphorylation of proteins can result in their activation or inhibition, which is usually a consequence of structural changes brought about by phosphorylation. These structural changes can occur proximal to the site of phosphorylation or farther away in the protein and influence protein-interaction surfaces. For example, in the first protein shown to be regulated through phosphorylation, glycogen phosphorylase, Serine14 phosphorylation disrupts acidic α-helices, bringing order to the intrinsically disordered N-terminus tail. These quaternary structural changes in the N-terminal lead to displacement of residues further along in the protein resulting in an ordered-to-disordered transition in the C-terminus, and allosteric activation of the enzyme (Johnson & Barford, 1993). Since the discovery of phosphorylation, this PTM has been implicated in metabolism, cell signalling, motility, immunity, cell cycle control, and more (Ubersax & Ferrell, 2007). While phosphorylation is the most studied PTM, acetylation has emerged as an area of intense study in the last 20 years.

Acetylation is a PTM that is closely linked with chromatin and gene expression. This process involves the donation of an acetyl group from acetyl-coenzyme A.

Acetylation occurs at the N-terminus of proteins, or at the exposed amino group of a lysine residue. Like phosphorylation, acetylation is reversible. In eukaryotes, the enzymes responsible for acetylation/deacetylation were discovered and named histone acetyltransferases/deacetylases, due to the target of acetylation: histones. Histones are proteins that wrap DNA in bundles called the nucleosome and assemble in defined structures to form chromatin. Lysine acetylation of histones was shown to influence the accessibility of DNA wrapped around the histone core of the nucleosome as well as assembly of histones into nucleosomes (Krajewski & Becker, 1998). Histones were also demonstrated to be methylated and phosphorylated, with one modification being able to influence another. For example, phosphorylation of Ser10 on histone H3 acts as a signal for subsequent acetylation at a downstream lysine (Cheung et al., 2000). We now know that lysines in proteins other than histones are targeted for acetylation, examples being general transcription factors and the proteins in NF-κB signalling pathway (Greene & Chen, 2004). Histone deacetylases have since been appropriately renamed lysine deacetylases. Likewise, the function of N-terminal acetylation has been implicated in the sub-cellular localization for a variety of proteins (see review (Drazic, Myklebust, Ree, & Arnesen, 2016)). While phosphorylation and acetylation involve the addition of a relatively small functional group, glycosylation is an example of a PTM that involves a considerably larger modification.

Glycosylation refers to the addition of a carbohydrate chain to a protein, often co-translationally in addition to post-translationally. Glycosylation is considerably more complex than the addition of a phosphoryl or acetyl group, and can occur on an asparagine residue (*N*-glycosylation), on a serine/threonine/hydroxy-lysine/tyrosine

residue (*O*-glycosylation), or on a tryptophan residue (*C*-mannosylation) to name a few. These glycosylation events are catalyzed by glycosyltransferases, usually accepting sugars from sugar nucleotide donors, and are broken down by glycoside hydrolases. To add another layer of complexity, glycosylation can lead to the addition of monosaccharides as well as branching glycan chains. These diverse modifications aid in protein folding and structure; cell-signalling through interaction with lectins; protein localization and secretion; protection from proteolytic degradation; and broader cellular effects like immunity, inflammation, and cellular regulation (see review (Moremen, Tiemeyer, & Nairn, 2012)).

Finally, another PTM that is important to consider due to its presence in both bacteria and eukaryotes is ADP-ribosylation, which is the transfer of ADP-ribose from NAD^+ to target proteins. ADP-ribosylation is carried out by ADP-ribosyltransferases that constitute a large group of proteins including the bacterial cholera and diphtheria toxins. These toxins use ADP-ribosylation machinery to inactivate G proteins and translational elongation factors respectively (M. S. Cohen & Chang, 2018). ADP-ribosylation results in mono-ADP-ribosylation or poly-ADP-ribosylation chains, attached to amino acids with oxygen-, nitrogen-, or sulfur-containing side chains (Luscher et al., 2018). ADP-ribosylation is also reversible through specific hydrolases. Established roles for ADP-ribosylation in eukaryotic cells include: immunity, cellular stress responses, transcriptional and translational regulation, and cell division (M. S. Cohen & Chang, 2018).

PTMs share common themes such as: reversibility (though some PTMs are irreversible, such as sulfation or proteolytic cleavage), donor-acceptor paradigms, shared

target residues, and most importantly, cooperation and antagonism. PTMs are capable of acting individually to increase the functionality and regulation of the proteome, but they further influence this complexity through interactions between different PTMs. In this next chapter I will describe how ubiquitination, the particular concern of this thesis, communicates with other PTMs, but also how it fills a unique role in our cellular system.

1.2: The Ubiquitin System

Ubiquitin is an 8.5kDa, 76-amino acid regulatory protein that adopts a distinctive ß-grasp fold (also called the "ubiquitin fold") structure characterized by five anti-parallel ß-strands grasping a single helical segment (Vijay-Kumar, Bugg, & Cook, 1987). Unlike the other PTMs discussed in this thesis, ubiquitin is exclusive to eukaryotic organisms. Although two analogous ubiquitin programs exist in bacteria: prokaryotic ubiquitin-like protein (Pup) in the Gram-positive Actinobacteria phylum (Pearce, Mintseris, Ferreyra, Gygi, & Darwin, 2008), and the homologous ubiquitin bacterial (UBact) system found in a variety of Gram-negative bacteria phylums such as Nitrospirae and Verrucomicrobia (Lehmann, Udasin, Livneh, & Ciechanover, 2017). An additional ubiquitin-like system also exists in Archaea (Humbard et al., 2010).

Ubiquitin is distinct from the previously-mentioned PTMs in that it involves the addition of an entire protein to (usually) a lysine residue on the target protein. Similar to glycosylation and ADP-ribosylation, ubiquitination can be more complex than one modification on one residue. Ubiquitination can happen singly (monoubiquitination), singly on multiple lysine residues (multi-monoubiquitination), or can form both linear and branched polyubiquitin chains by being conjugated to internal lysines on another ubiquitin protein. Ubiquitination can lead to protein destruction, activation, or changes in

cellular localization, depending on the manner of its attachment (Swatek & Komander, 2016). The discovery of ubiquitin, and with it that proteins are regulated through the covalent attachment of other proteins was a pivotal finding (Ciechanover, Heller, Elias, Haas, & Hershko, 1980; Hershko, Ciechanover, Heller, Haas, & Rose, 1980) and has led to the study of other ubiquitin-like proteins that will be discussed later in this thesis.

1.2.1: Transfer of Ubiquitin – E1s, E2s, E3s

Ubiquitin is, in point of fact, not produced in the readily-conjugable 76-amino acid form. In humans, ubiquitin is produced from four different genes. UBB and UBC code for multiple copies of ubiquitin and produce a four and nine polyubiquitin product respectively (Wiborg et al., 1985). On the other hand, UBA52 and RPS27A produce ubiquitin fused to ribosomal proteins (Finley, Bartel, & Varshavsky, 1989; Redman & Rechsteiner, 1989). All four of these gene products must first be processed by deubiquitinases (DUBs) to release the mature monomeric form of ubiquitin required for ubiquitination. This leads to multiple areas of regulation: transcription can be individually tuned for each gene (some of which are induced under specific circumstances like stress (Finley, Ozkaynak, & Varshavsky, 1987)), in addition to the regulation of the DUBs themselves. Furthermore, the addition of ubiquitin to its substrate, canonically via an isopeptide linkage between the ubiquitin C-terminal glycine and the substrate lysine side chain, requires a cascade of three classes of enzymes.

Free ubiquitin is conjugated to target proteins by a three-enzyme cascade consisting of an E1 activating enzyme, E2 conjugating enzyme, and E3 ligase (Figure 1). Monomeric ubiquitin is first activated, in an ATP-dependent fashion, by the E1 enzyme. This is a relative unspecific process, as there are only two human E1 enzymes that

recognize ubiquitin: UBA1, UBA6 (Schulman & Harper, 2009). The E1 enzyme first binds ATP and ubiquitin, and then catalyzes the acyl-adenylation of ubiquitin on its C-terminus, resulting in a high energy ubiquitin-AMP thioester-linked intermediate. Next, a conserved cysteine residue in the E1 enzyme catalytic core performs a nucleophilic attack on this thioester bond. This reaction ends with a ubiquitin-E1 thioester-linked complex, and a free AMP, though E1s are known to adenylate another ubiquitin before discharging the first protein (Schulman & Harper, 2009). This process is known as ubiquitin activation.

Ubiquitin is then transferred through transthiolation to another conserved cysteine residue, this time on the E2 conjugating enzyme. E2 enzymes are more varied than E1s, with at least 38 encoded in humans (Ye & Rape, 2009). The ubiquitin-E1 conjugate interacting with the E2 is conjugated with two ubiquitin proteins, one attached to the active site cysteine, and another attached to its adenylation domain ready to undergo the next round of ubiquitination (Schafer, Kuhn, & Schindelin, 2014).

The E3 enzyme confers substrate specificity to the ubiquitination reaction; 600-1,000 E3s exist in the human genome (Deshaies & Joazeiro, 2009), which is more than the estimated <500 human kinases (P. Cohen & Tcherpakov, 2010). E3 enzymes are split into two main categories. HECT (Homologous to E6-AP Carboxyl Terminal) E3s have a catalytic cysteine that accepts ubiquitin from the E2 (again, through a transthiolation reaction) before ubiquitinating targets. RING (Really Interesting New Gene) E3s on the other hand do not contain this catalytic cysteine but act instead as scaffolds, bringing the E2 and the substrate into close proximity to direct exchange of ubiquitin from one to the other. A comprehensive examination of various E3 groups follows in the next section.

Ubiquitin is attached to its substrate typically on a lysine residue, specifically via amide isopeptide bond between the lysine amide group and the C-terminal glycine of ubiquitin (Hershko & Ciechanover, 1998). Addition of ubiquitin to a protein, and subsequent digestion with trypsin, therefore leaves a diagnostic Gly-Gly dipeptide attached to an internal lysine on the target protein. This di-Gly tag is detectable via mass spectrometry and used as evidence of ubiquitin modification (W. Kim et al., 2011). A growing appreciation for non-canonical ubiquitination where ubiquitin is coupled not to a lysine, but rather to the N-terminal amine group (Vosper et al., 2009) of a protein, or the hydroxyl group of threonine and serine (Tait et al., 2007), or cysteine residues (Vosper et al., 2009). Recall that this diversity of attachment was hinted by glycosylation, whose chemistry is also dependent on the free amino group of lysine residues. So far, the consequences of these non-canonical events are poorly understood, but most examples target proteins for degradation similarly to canonical ubiquitination (McDowell & Philpott, 2013).

1.2.2: Types of E3s

At minimum, 616 RING E3s are expressed in humans, which is more than the suite of specific kinases in a cell (Deshaies & Joazeiro, 2009). The RING protein domain is responsible for the recruitment of both ubiquitin-E2 complexes and substrate proteins. Interestingly, the ability of a RING E3 to ubiquitinate its target is not correlated with the strength of its interaction with the charged E2 (Lorick et al., 1999). Instead, it is thought that the RING E3 binds both substrate and charged E2 weakly, to prevent product inhibition (Deshaies & Joazeiro, 2009). RING E3s do, however, promote the discharge of ubiquitin from the E2 (ideally to the substrate), up to 87-fold (Ozkan, Yu, & Deisenhofer,

2005) through binding-induced conformational changes. RING-type E3s are defined structurally and possess conserved cysteine residues that co-ordinate zinc and are required for RING stability. E3s with RING-like domains also exist, a group that includes the U-box proteins (Aravind & Koonin, 2000). These RING-like domains are similar to the RING domain in structure but do not co-ordinate zinc. Both RING E3 and RING-like E3s tend to form homo and heterodimeric complexes (Metzger, Pruneda, Klevit, & Weissman, 2014), but there is also a class of E3s that form even larger complexes.

Cullin RING ligases (CRLs) are a subset of RING E3s that are composed of multiple proteins assembled in a specific complex (Feldman, Correll, Kaplan, & Deshaies, 1997; Skowyra, Craig, Tyers, Elledge, & Harper, 1997). CRLs are composed of a RING protein in complex with a cullin-family protein that acts as a scaffold for the protein complex. These complexes have additional proteins that are used to recruit and interact with regulatory proteins that dictate substrate specificity (Petroski & Deshaies, 2005). The prototypic CRL, SCF, is made up of SKP1, a Cullin-family scaffold, and an F box protein and is responsible for regulating a large variety of cellular processes such as: DNA replication, cell growth, transcription, and circadian rhythm (Cardozo & Pagano, 2004). Notably, CRLs are regulated by the attachment of a ubiquitin-like protein called NEDD8, an example of the regulatory overlap between the ubiquitin system and other ubiquitin-like systems.

HECT E3s are less numerous than RING E3s, and are named after the first enzyme of their class: Homologous to the E6-AP Carboxyl-terminus (Scheffner, Nuber, & Huibregtse, 1995). E6-AP, a human papillomavirus protein, was the first E3 shown to accept ubiquitin directly through transthiolation from the E2 to a conserved cysteine on

the E3. Since that pivotal finding, approximately 30 HECT E3s have been discovered in mammals (Rotin & Kumar, 2009). The HECT domain is split into two lobes, the C-terminal lobe interacts with ubiquitin, the N-terminal lobe interacts with the E2, whereas substrate recognition and binding occurs outside of the HECT domain (Rotin & Kumar, 2009). The activity of HECT E3s can be governed by phosphorylation as is the case with ITCH by JNK1 (Gallagher, Gao, Liu, & Karin, 2006). HECT activity may also be controlled by autoinhibitory intramolecular domains as is the case with SMURF2 (Wiesner et al., 2007), as well as through oligomerization and a variety of other PTMs (Sluimer & Distel, 2018).

Another class of E3 ligases that share characteristics with both the HECT E3s and RING E3s are called the RING between RING (RBR) E3s. RBR E3s are characterized by two RING domains, RING1 and RING2, with an in-between-RING domain between the two. The RING2 domain contains the active site cysteine that accepts ubiquitin via a thioester bond just like HECT E3s. The first RBR to be discovered was parkin, an E3 associated with mitochondrial damage and whose mutation contributes to Parkinson's disease (Wenzel, Lissounov, Brzovic, & Klevit, 2011). Our understanding of RBRs has since expanded to include 14 RBR E3s in humans including HOIP and HOIL-1L, components of LUBAC, the linear ubiquitin chain assembly complex. RBR are often autoinhibited and can dictate the nature of the ubiquitin chains formed on the target substrate (Dove & Klevit, 2017).

1.2.3: Deubiquitinases

As previously mentioned, deubiquitinases (DUBs) process nascent polyubiquitin and ubiquitin-ribosomal protein fusion proteins; however, DUBs also cleave ubiquitin

from ubiquitinated proteins as a means of recycling, regulating, or reversing ubiquitination. In this way DUBs act analogously to phosphatases in regulating phosphorylation. DUBs are proteases of two sorts; the majority of the 100 or so DUBs encoded in humans (Nijman et al., 2005) are cysteine proteases, while the rest are metalloproteases. DUBs are specific for particular ubiquitin chain types, a specificity that is contributed by a variety of protein-binding domains (Mevissen & Komander, 2017). Additionally, the activity of DUBs is regulated by PTMs, including phosphorylation, ubiquitination, and other ubiquitin-like covalent modifications (Reyes-Turcu, Ventii, & Wilkinson, 2009).

1.2.4: Linkages & Consequences of Ubiquitination

The number of ubiquitins and the manner by which they are attached to a substrate will alter that protein's fate within the cell in different ways (Behrends & Harper, 2011) (Figure 1).

Monoubiquitination is typically associated with regulatory events. For example, monoubiquitination of endocytic proteins can lead to inhibition (Hoeller et al., 2006). Often, functions associated with monoubiquitination are due to intracellular localization as demonstrated by HRAS, a GTPase involved in MAPK signalling (Jura, Scotto-Lavino, Sobczyk, & Bar-Sagi, 2006). Furthermore, monoubiquitination has been shown to affect the formation of protein complexes, the canonical example being the requirement of ubiquitination of histone H2B for transcriptional elongation (Pavri et al., 2006). The particular sites of monoubiquitination can be well defined, as is the case for IκBα (Baldi, Brown, Franzoso, & Siebenlist, 1996) but may also occur on no lysine in particular (Treier, Staszewski, & Bohmann, 1994). Multi monoubiquitination on the other hand has

been linked to degradation via endocytosis and the lysosome, as is demonstrated by receptor tyrosine kinases (Haglund et al., 2003).

Ubiquitin has seven internal lysines (K6, K11, K27, K29, K33, K48, K63) that can serve as the basis for polyubiquitin chain formation. The first polyubiquitin chain to be discovered was the K48-linked chain (Chau et al., 1989), which canonically signals proteins for degradation via the proteasome, a barrel-like structure that recycles proteins through proteolysis. The entry to the proteases within the barrel core of the proteasome is regulated by components of the "lid" that bind to polyubiquitin signals on proteins (Elsasser & Finley, 2005). Phosphorylation of specific domains (Kornitzer, Raboy, Kulka, & Fink, 1994; Yaglom et al., 1995) and sequence specific motifs called degrons (Varshavsky, 1996) are both signals for polyubiquitination and subsequent degradation of these proteins. It is also worth mentioning that K11-linked and most other polyubiquitin chains (with the exception of K63-linked chain) (Xu et al., 2009); monoubiquitination; or ubiquitin-independent signals have also been linked to destruction via the proteasome (Kravtsova-Ivantsiv & Ciechanover, 2012), overturning previous thinking that there existed a minimum proteasome signal (Thrower, Hoffman, Rechsteiner, & Pickart, 2000). While K48-linked polyubiquitin chains are the most common in human cells, other polyubiquitin linkages have important physiological functions.

K63-linked ubiquitin chains have two main established roles to date. They are involved in the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) response, a signalling pathway involved in defense against pathogens. During this response, a specific set of E3 ligases called the TRAFs are known to promote K63 chains (also called regulatory ubiquitin) in order to activate IκB Kinase (IKK), which

phosphorylates downstream signal adaptors (Windheim, Stafford, Peggie, & Cohen, 2008). Additionally, K63-linked chains play a role in DNA-damage response, where they help to recruit the E3 ligase BRCA1 (Bennett & Harper, 2008). K11-linked chains, on the other hand, have been implicated in cell cycle regulation (Jin, Williamson, Banerjee, Philipp, & Rape, 2008), endoplasmic reticulum-associated degradation (Xu et al., 2009), and TNFα signalling (Dynek et al., 2010).

K48, K11, and K63-linked polyubiquitin chains are all examples of homotypic ubiquitin chains, where each ubiquitin is linked by the same lysine. Heterotypic ubiquitin chains also exist, where multiple linkages appear in a single polyubiquitin chain. While heterotypic ubiquitin chain formation is still a relatively new area of study, some functions have been elucidated for specific chain types. For example, K29/K33 polyubiquitin chains prevent phosphorylation of the energy-sensing protein AMPK (Al-Hakim et al., 2008). Other non-canonical chains include branched polyubiquitin chains, where one ubiquitin is attached to two others through various linkages, and they are thought to enhance recognition by the proteasome (Meyer & Rape, 2014). Additionally, polyubiquitin chains have been found that incorporate other ubiquitin-like proteins such as SUMO and function analogous to K48-linked chains to target substrates to the proteasome (Tatham et al., 2008). These other ubiquitin-like proteins will be discussed in Chapter 1.5.

Finally, the latest polyubiquitin chains to be discovered are linear ubiquitin chains, joined head-to-tail, using the N-terminal methionine amino group rather than a side chain amino group from an internal lysine (Kirisako et al., 2006). These linear ubiquitin chains (also known as met-ubiquitin) are assembled by the linear ubiquitin

assembly complex (LUBAC), a multiprotein complex containing two RBR E3s mentioned earlier. The consequences of linear ubiquitination are understood in NF-κB signalling, where they act in a regulatory capacity to activate the pathway (Tokunaga et al., 2009). Clearly, polyubiquitin chains dictate a variety of functions, but what dictates how polyubiquitin chains are formed?

There are multiple models that account for ubiquitin chain assembly and linkage type. The classical model of polyubiquitin chain assembly follows a sequential addition explanation. After the first ubiquitination event, the lysine from ubiquitin on an already-ubiquitinated substrate attacks another ubiquitin-E2 or ubiquitin-HECT-E3 complex, leading to the addition of another ubiquitin on the substrate chain. This model becomes less tenable for longer ubiquitin chains, but could be explained by the ubiquitin chain looping out to accommodate the required machinery (Rape, Reddy, & Kirschner, 2006). However, this model does not account for E2s which seem to be able to able to synthesize chains of ubiquitin without the presence of an E3 enzyme or a substrate protein such as E2-25K (Z. J. Chen, Niles, & Pickart, 1991).

Three other proposed models for polyubiquitin chain assembly include: the chain elongation model where another set of enzymes (sometimes called E4s) help build the growing chain (Koegl et al., 1999), the indexation model where the ubiquitin chain is built on the E3 before transfer to the substrate (Verdecia et al., 2003), and the see-saw model where the ubiquitin chain is built on a dimeric E2 (or an E2 and HECT E3 pair) and passed back and forth to facilitate elongation (Hochstrasser, 2006).

More recently, determinants of chain type have been associated with whichever enzyme is the last to accept ubiquitin directly. For HECT E3s, the C-terminal HECT

domain lobe has been shown as solely responsible for ubiquitin chain type, independent of the recruited E2 (H. C. Kim & Huibregtse, 2009). Conversely, RING E3s do not direct ubiquitin chain types directly, but instead recruit different charged E2s that specifically dictate chain type (Williamson et al., 2011). One of the best examples of this is the recruitment of Ubc13, the only known E2 that constructs K63-linked chains (Deng et al., 2000). The reality is most likely that there is no single factor that dictates how the ubiquitin machinery forms polyubiquitin chains, but rather a variety of mechanisms that are used on a case-by-case basis.

1.3: *Shigella*

Many bacteria exploit the eukaryotic ubiquitin system in pathogenesis (Perrett, Lin, & Zhou, 2011), one of the best characterized examples being *Shigella. Shigella* is a genus of rod-shaped, non-motile, Gram-negative, intracellular pathogen. *Shigella* belongs to the phylum Proteobacteria and the family Enterobacteriaceae, and can survive for weeks on contaminated surfaces, food, and water (Kramer, Schwebke, & Kampf, 2006). *Shigella* is the causative agent of Shigellosis, a diarrheal disease in primates, including humans.

Shigellosis is characterized by mild to severe dysenteric symptoms including: abdominal discomfort, cramps, diarrhea, fever, and bloody stool. The Centers for Disease Control (CDC, 2017) estimates that the incidence of shigellosis is 80-165 million cases per year, predominantly in children. Though healthy individuals usually clear the infection without intervention, 600,000 deaths are attributed to shigellosis annually, often from secondary complications such as dehydration, malnutrition, and renal failure (Phalipon & Sansonetti, 2007). These numbers vary considerably from source to source,

but are nonetheless thought to be conservative estimates as traditional diagnosis methods underrepresent *Shigella* (Mani, Wierzba, & Walker, 2016). A landmark study that used modern molecular tools to identify *Shigella* in a multicentre study showed that while the morbidity due to shigellosis has declined in recent decades, the incidence remains unchanged (von Seidlein et al., 2006). Importantly, no vaccine currently exists for *Shigella* despite the emergence of multi-drug antibiotic resistance (Puzari, Sharma, & Chetia, 2018).

1.3.1: *Shigella* Species

The genus *Shigella* is divided into four species: *S. flexneri, S. sonnei, S. dysenteriae,* and *S. boydii. S. sonnei* is the primary species found in the developed countries, accounting for 77% of all cases (WHO, 2006), whereas it has a comparatively low impact in developing countries at 23.7% of all cases (Livio et al., 2014). *S. dysenteriae* is unique in producing the enterotoxin Shiga Toxin, and is responsible for epidemic outbreaks in conditions such as refugee camps, or where sanitation conditions are poor, facilitating the fecal-oral spread of the bacteria. *S. boydii* is largely relegated to the Indian subcontinent and is comparatively rare, accounting for only 5.4% of all cases (Kania, Hazen, Hossain, Nataro, & Rasko, 2016). *S. flexneri* is the most commonly isolated species both worldwide at 60% of all cases (WHO, 2006), and in developing countries at 65.9% of all cases (Livio et al., 2014). *Shigella flexneri* is the focus of investigation in the Rohde lab, and this thesis. By extension, mention of *Shigella* from this point on will generally refer to *Shigella flexneri.*

1.3.2: *Shigella* Infection

Shigella infects via the fecal-oral route: person-to-person, or through contaminated food and water. *Shigella* is resistant to stomach acid, and as such has a minutely small infectious dose in the range of 10 organisms (Gorden & Small, 1993). This makes the transmission of *Shigella* more effective than other enteric pathogens. *Shigella* invades through the intestinal epithelium. Specifically, specialized gut cells called M cells, found in the gut-associated lymphoid follicles called Peyer's patches, transcytose *Shigella* from the gut lumen to the underlying M cell pocket (Wassef, Keren, & Mailloux, 1989). While this process is usually used to present lumen antigens to the gut-associated lymphoid tissue, *Shigella* uses it as an entry point and subsequently infects macrophages and dendritic cells that reside in this pocket. Infection of macrophages results in the induction of pyroptosis, an inflammatory programmed cell death event, and the release of *Shigella* that subsequently enters epithelial cells via the basolateral membrane (Ashida, Mimuro, & Sasakawa, 2015; Neiman-Zenevich, Stuart, Abdel-Nour, Girardin, & Mogridge, 2017). These first infection steps result in cell damage and the production of proinflammatory mediators that act to recruit immune cells and cause inflammation.

Shigella triggers its own uptake into epithelial cells and resides briefly inside a bacteria-containing vacuole (BCV). *Shigella* escapes the ensuing vacuole to replicate in the cytosol (Cossart & Sansonetti, 2004). *Shigella* then uses the secreted protein IcsA to interact with N-WASP and induce Arp2/3-dependent actin polymerization to move intracellularly (Welch & Way, 2013). Remarkably, *Shigella* can use this same force to propel itself intercellularly into neighbouring cells, avoiding host immune surveillance

and disrupting/killing host epithelium (Goldberg & Theriot, 1995). This process is remarkably similar the actin-based movement of another pathogen, *Listeria* (Welch & Way, 2013).

1.3.3: Cellular Defences Against *Shigella* – The NF-κB Response

Bacterial infections are sensed by pattern-recognition receptors (PRRs) that recognize components of bacterial cells that should not usually be present in the human cellular milieu (Takeuchi & Akira, 2010). Toll-like receptors (TLRs) sense extracellular pathogen-associated molecular patterns (PAMPs) like lipopolysaccharides, peptidoglycans, and flagellins, while the intracellular receptors for these same patterns are Nod1 and other nod-like receptors (Takeuchi & Akira, 2010). As *Shigella* experiences both intra and extracellular phases during infection, each of these receptors has been implicated in innate immune response against *Shigella* (Rahman & McFadden, 2011).

Activation of PRRs initiates signal transduction pathways that result in changes in gene expression within the nucleus. PRR recognition of PAMPs causes oligomerization of these receptors that recruits receptor-associated signal transducer adaptor proteins such as MyD88 or TRAF proteins (TNF receptor-associated factor). These proteins link initial signalling events to the signal transduction pathways NF-κB and the mitogen-activated protein kinase (MAPK) cascades to induce inflammation and defences against pathogens. It is worth noting that the NF-κB response can also be activated by other signals originating from protein kinase c (PKC) and tumor necrosis factor receptors (TNFRs).

The NF-κB signal cascade is perhaps the most important cellular defense linked to *Shigella* infection. In turn, *Shigella* devotes a significant portion of its virulence toolkit to subverting this signalling program. Transduction of PRR signals activates a protein

complex called IκB kinase or IKK. IKK is composed of three proteins: IKKα, IKKβ, and IKKγ (also known as NEMO). The IKK complex receives signals, such as K63-ubiquitination of NEMO from TRAF6, that activate its kinase activity. As its name implies, IKK phosphorylates IκB, also known as the inhibitor of κB. Phosphorylation of this inhibitor results in its subsequent K48-linked polyubiquitination and degradation by the proteasome, freeing NF-κB from its inhibition and allowing its migration to the nucleus.

NF-κB itself is a dimeric protein complex consisting of some combination of p50, p52, p65 (RelA), c-Rel, and RelB. IκB sterically blocks the NF-κB nuclear localization signal and DNA-binding motif until it is destroyed. Depending on what variants of NF-κB is activated, these transcription factors enter the nucleus, recruit various co-factors, and turn on the expression of a large variety of inflammatory genes (Hoffmann, Leung, & Baltimore, 2003).

While the NF-κB inflammatory response and its subversion is the best characterized host-*Shigella* interface, other factors play a role in modulating the spread of *Shigella* including: antimicrobial peptides, physical barriers to infection, and autophagy (which will be discussed at length in Chapter 1.6).

1.3.4: The Type III Secretion System

Early pioneering work in the *Shigella* field showed that a 37kB region of a large 220kB virulence plasmid was necessary and sufficient for invasion of epithelial cells (Maurelli, Baudry, d'Hauteville, Hale, & Sansonetti, 1985). This so-called entry region was later shown to express components of a type III secretion system (T3SS), a needle-like apparatus capable of translocating effector proteins into the host cytosol, and these

very-same effector proteins. *Shigella* uses its T3SS to inject a variety of effector proteins into the cytosol of infected cells that aid in its pathogenesis (Coburn, Sekirov, & Finlay, 2007).

The more than twenty Mxi/Spa proteins that comprise the T3SS assemble to form a structure that is evolved from bacterial flagella (Abby & Rocha, 2012). The T3SS is comprised of a basal body structure that spans the bacteria inner and outer membranes. Attached to the basal body and extending outside *Shigella* is a needle-like apparatus that ends in a protein complex that is used to form a translocon, or pore, in the host cell (Marlovits & Stebbins, 2010). The needle-like apparatus is hollow, and it is through this channel that effector proteins transit from the bacterium into the cytosol of the infected human cell. The hollow nature of the T3SS and transit of effector proteins was demonstrated by real-time tracking of fluorescently-labeled effector proteins (Enninga, Mounier, Sansonetti, & Tran Van Nhieu, 2005), and recent structural studies (Hu et al., 2018; Nans, Kudryashev, Saibil, & Hayward, 2015).

1.3.5: *Shigella* T3SS Effectors

Production of the effectors is controlled by temperature and the activity of the T3SS. At 37°C, the AraC family transcriptional activator, VirF, activates transcription of *VirB*. VirB then directs transcription of the *mxi-spa* regulon, which encodes the structural components of the T3SA and some of the early effectors (Mavris et al., 2002). The early effectors include IpaA, VirA, and IpgD that aid in bacterial uptake and regulating actin dynamics (Parsot, 2009). These effectors also include the components of the translocon IpaB and IpaC which are maintained in a secretion-ready state through their interaction with the chaperone IpgC. Contact with host cells triggers delivery of IpaB and IpaC

through the T3SA where they are believed to oligomerize and form a pore in the host membrane. Delivery of IpaB and IpaC frees IpgC which now acts as a transcriptional coactivator. IpgC forms a complex with MxiE and binds to the promoters of several genes and promotes their transcription (reviewed in (Phalipon & Sansonetti, 2007)). These MxiE-IpgC dependent (late) effectors are then injected into host cells where they interfere with host cell processes (reviewed in (Parsot, 2009)). Some of the late *Shigella* effectors are outlined in Table 1. Not included are the IpaH family of late effectors, which will be described in detail in the next section of this thesis.

Table 1: *Shigella* late effector functions.

Effector	Enzymatic activity	Process affected	Reference
OspG	Ubiquitin-dependent kinase	NF-κB	(D. W. Kim et al., 2005)
OpsI	Ubc13 deamidase	NF-κB	(Sanada et al., 2012)
OspF	Phosphothreonine lyase	MAPKs	(H. Li et al., 2007)
OspE1/E2	Unknown	Host cell adherence	(M. Kim et al., 2009)

1.4: Bacterial-Encoded E3 Ligases (BELs)

Despite lacking their own ubiquitin system, prokaryotic pathogens encode proteins that interact with ubiquitin. OspG and OspI, for example, block the functional activity of two different E2 enzymes (D. W. Kim et al., 2005; Sanada et al., 2012). Other methods of interfering with the host ubiquitin system include encoding DUBs to remove ubiquitin, or sequestering host DUBs to disrupt their function (Anderson & Frank, 2012;

Patel, Hueffer, Lam, & Galan, 2009). Bacteria have even evolved enzymes that require ubiquitin to function, or those whose functions change with ubiquitination, ensuring that their pathogenic activity is only turned on in host cells (Anderson & Frank, 2012). The most blatant example of pathogens hijacking the host ubiquitin system has been the evolution of bacterial-encoded E3 ligases (BELs), to specifically direct host ubiquitination (Huibregtse & Rohde, 2014).

1.4.1: Non-*Shigella*-Encoded BELs

The first BEL to be discovered was AvrPtoB found in the tomato plant pathogen *Pseudomonas syringae* (Abramovitch, Janjusevic, Stebbins, & Martin, 2006). AvrPtoB was later shown to be structurally similar to a RING E3, and exerts its function by ubiquitinating and destroying components of the plant immune system (Rosebrock et al., 2007). Other RING-like BELs were soon discovered, including the Enterohemorrhagic *E. coli* (EHEC) U-box-containing protein NleG (B. Wu et al., 2010), and the *Legionella* U-box type BEL LubX that ubiquitinates and degrades its own effectors as a form of regulation (Kubori, Shinzawa, Kanuka, & Nagai, 2010).

In addition to RING-like BELs, there are also HECT-like BELs. *Salmonella* effector SopA function depends on a catalytic cysteine and formation of a Ub-SopA catalytic intermediate (Y. Zhang, Higashide, McCormick, Chen, & Zhou, 2006). Like its U-box-containing counterparts, SopA downregulates the host innate immune response against pathogens by targeting, in this case, specific immunity-related TRIM56 and TRIM65 proteins for degradation (Fiskin et al., 2017). Another HECT-type BEL is the EHEC protein NleL that promotes formation of actin pedestals that are crucial for

successful infection (Piscatelli et al., 2011). The discovery of the function of the IpaH family of proteins led to their designation as a new type of E3 ligase: the NELs.

1.4.2: The IpaHs

The Invasion Plasmid Antigen H, or IpaH family of proteins are second-phase effectors (expression controlled by MxiE) in *Shigella,* and were shown to possess E3 ligase activity (Rohde, Breitkreutz, Chenal, Sansonetti, & Parsot, 2007). The structure of the IpaHs was solved independently by three groups, and demonstrated that these enzymes represent a structurally distinct class of E3s (Quezada, Hicks, Galan, & Stebbins, 2009; Singer et al., 2008; Zhu et al., 2008). The IpaHs instead have two protein domains: an N-terminal leucine-rich repeat (LRR) domain and a C-terminal novel E3 ligase (NEL) domain. The LRR domain contains seven 21-amino acid repeats that are highly variable amongst different IpaH members. This variation may explain their ability to recognize different substrates (Haraga & Miller, 2006; Keszei et al., 2014).

The LRR domain was also shown to be involved in autoinhibiting the activity of the NEL domain in the absence of substrates, as N-terminal truncations lead to an increase in activity, including autoubiquitination which has been shown to regulate E3 enzymes through proteasome-mediated degradation (Galan & Peter, 1999). The LRR masks the catalytic C-terminal domain until substrate binding induces a conformational change (Chou, Keszei, Rohde, Tyers, & Sicheri, 2012). Autoinhibited IpaHs are still able to accept ubiquitin from Ub-E2 conjugates but are blocked from subsequent ubiquitination of substrates. Instead they are "short-circuited" and discharge ubiquitin non-productively into the surrounding space (Keszei & Sicheri, 2017). The NEL domain, on the other hand, while structurally dissimilar from other E3s, functions analogously to a

HECT E3 in that its activity is dependent on a catalytic cysteine residue and formation of Ub-IpaH thioester-linked intermediates (Singer et al., 2008; Zhu et al., 2008).

The IpaH family spans several bacterial genera. In the enteric pathogen *Salmonella,* three IpaH family members (SspH1, SspH2, and SlrP) are secreted as effectors. Only two of these family members are well defined. SspH1 is one of the best characterized IpaH family members, whose target, PKN1, a kinase involves in actin cytoskeleton regulation, has been shown to interact directly with the SspH1 LRR domain. Ubiquitination of PKN1 leads to its proteasomal destruction, presumably to suppress host immune responses to *Shigella* through PKN1-mediated regulation of the androgen receptor (Keszei et al., 2014). The mechanism of ubiquitin chain formation for NEL-domain E3s follows a "see-saw" model where a polyubiquitin chain is extended by passing the chain from the E2 to the E3 then back again (Levin et al., 2010). SlrP, targets an altogether different host protein, thioredoxin-1. Thioredoxin is responsible for controlling redox signalling events; its targeting and destruction is thought to lead to increased susceptibility to apoptosis (Bernal-Bayard & Ramos-Morales, 2009). These examples demonstrate the scope of the targets of the IpaHs in a single bacterial genus; however, IpaH family members exist in other genera.

The nitrate-fixing bacterium *Rhizobium* encodes an IpaH family effector, NopM (previously Y4FR). Interestingly, this E3 ligase aids in symbiosis not pathogenicity, facilitating the interaction between *Rhizobium* and the legumes it nodulates by suppressing the plant immune system (Xin et al., 2012). *Yersiniae* encode an effector, YopM, that has considerable homology to the LRR domain of the IpaHs, but has a truncated C-terminus lacking a NEL domain. Recent research has reportedly ascribed E3

ligase function to a cysteine-lysine-aspartic acid motif in YopM (Wei et al., 2016) though the researchers did not specify the E2 responsible for the reported K63-linkages they observed, and used YopM isolated from a specific strain of *Yersinia pestis* rather than the original *Yersinia enterocolitica* sequence first ascribed to YopM. Additionally, *Pseudomonas, Edwardsiella,* and *Bradyrhizobium* bacteria encode poorly characterized IpaH family member proteins (Rohde et al., 2007).

While IpaH family members are found occasionally in other species of bacteria, they are a hallmark of *Shigella* infection, where there are twelve different IpaH variants (Ashida, Toyotome, Nagai, & Sasakawa, 2007). The virulence plasmid encodes the five best characterized IpaHs (IpaH1.4, IpaH2.5, IpaH4.5, IpaH7.8, and IpaH9.8), whereas there are seven poorly-characterized chromosomal IpaHs (IpaH0722, IpaH0887, IpaH1383, IpaH1880, IpaH2022, IpaH2202, and IpaH2610). The chromosomal IpaHs are, nonetheless, found to be secreted and act as effectors; though individual deletion of each chromosomal IpaH shows no effect on virulence, deletion of all seven effectors attenuates *Shigella* lethality in a mouse infection model (Ashida et al., 2007). Of the chromosomal IpaHs, only IpaH0722 has been ascribed a role in infection, specifically by dampening NF-κB pathogen responses through modulation of proteins downstream to Protein Kinase C (Ashida, Nakano, & Sasakawa, 2013).

The smallest plasmid-borne IpaHs, IpaH1.4 and IpaH2.5 are 98% identical, and as such are hypothesized to have arisen from a genetic duplication event (Buchrieser et al., 2000). Previously these genes were thought to be truncated, inactive forms of IpaH7.8 (Venkatesan, Buysse, & Hartman, 1991), but have since been shown to play a role in targeting LUBAC. Specifically, both proteins (though IpaH1.4 is thought to be the

essential ligase) lead to K48-mediated proteosomal degradation of the HOIL subunit of LUBAC, abrogating linear ubiquitination chain formation and dampening the NF-κB response to PAMPs (de Jong, Liu, Chen, & Alto, 2016). This phenomenon is described in detail by Noad et al. (2017) who show that linear ubiquitin acts as a signal for destruction on bacteria through autophagic clearance or NF-κB inflammation activation.

IpaH4.5 is another example of an effector that downregulates the NF-κB pathway. Wang et al. (2013) demonstrated *in vitro* and in a murine model that IpaH4.5 knockout is characterized by an increase in inflammation. They showed that the NEL domain of IpaH4.5 specifically interacts with and ubiquitinates the p65 subunit of NF-κB. More recently, another substrate of IpaH4.5 was proposed, TANK-binding kinase 1 (TBK1). TBK1 responds to PAMPs to clear bacterial infection through IFN regulatory factor 3. *Shigella* subverts this signalling axis by degrading TBK1 through IpaH4.5 directed, K48-dependent proteasome-mediated destruction (Z. Zheng et al., 2016).

IpaH7.8 was shown to be involved in the initial *Shigella* infection of macrophages, where it induces pyroptosis to release bacteria and subsequently infect epithelial cells. Mice infected with IpaH7.8-deficient *Shigella* were shown to have accumulated bacteria in the macrophages and were delayed for successful infection. This phenotype was ascribed to loss of IpaH7.8-dependent ubiquitination and subsequent degradation of glomulin (Suzuki et al., 2014). Glomulin is involved in negatively regulating the NLRP3/NLRC4 inflammasome, a multi-protein complex that activates pyroptosis.

The focus of this thesis is, however, on the last virulence plasmid-encoded IpaH: IpaH9.8.

1.4.2.1: IpaH9.8

IpaH9.8 was the first IpaH to be characterized; establishing a paradigm for the IpaHs being E3 ubiquitin ligases when it was shown to be capable of ubiquitinating the MAPKK Ste7 in a *Saccharomyces cerevisiae* model system (Rohde et al., 2007). Early data suggested that IpaH9.8 was shuttled to the nucleus in a microtubule-dependent fashion (Toyotome et al., 2001). Analysis of IpaH9.8 mutants revealed that IpaH9.8 dampened inflammation in a murine model (Okuda et al., 2005). Later, another substrate for IpaH9.8 was identified by the same lab. Together, with a ubiquitin-binding adaptor protein ABIN-1, they reported that IpaH9.8 led to proteasomal-dependent degradation of NEMO (also known as IKK gamma), dampening the NF-κB response (Ashida et al., 2010). However, our lab has been unable to replicate this result. In support of our results, Alto and coworkers found no evidence that IpaH9.8 ubiquitinates NEMO (de Jong et al., 2016).

Partway into my research program at Dalhousie, another substrate for IpaH9.8, guanylate binding proteins (GBPs), was discovered independently by three groups (P. Li et al., 2017; Piro et al., 2017; Wandel et al., 2017). GBPs are interferon-inducible GTPases. Type I or Type II interferons recognize invading pathogens in the cytosol and upregulate a variety of interferon-stimulated genes including the GBPs. GBP1 is recruited to intracellular *Shigella,* which coats the bacteria and recruits other GBPs in a hierarchical manner. GBP-coated bacteria act as targets of ubiquitination, which can lead to autophagic clearance of the bacteria (Al-Zeer, Al-Younes, Lauster, Abu Lubad, & Meyer, 2013), killing of the bacteria directly (Man et al., 2015), or inhibition of *Shigella* actin-based motility (Piro et al., 2017; Wandel et al., 2017). Three studies have

independently shown that IpaH9.8 is responsible for disrupting GBPs. They collectively demonstrate that IpaH9.8 catalyzes K48-linked polyubiquitin chains on GBP1, GBP2, and GBP4 to facilitate their proteasomal destruction.

Our lab was interested in finding additional substrates for IpaH9.8, as the LRR domain is a promiscuous protein-protein interaction domain and can mediate interactions with multiple proteins. Despite more than a decade of research from many labs of note, discovery of other substrates of IpaH9.8 has been strangely lacking. I hypothesized that this was because IpaH9.8 cooperates with another ubiquitin-like protein.

1.5: Ubiquitin-Like Systems

Ubiquitin-like proteins (Ubls) are proteins that are covalently attached to other proteins in a fashion similar to ubiquitin. They share the beta-grasp fold domain with ubiquitin, but do not necessarily show amino acid sequence similarities. Ubls are distinct from proteins that share the beta-grasp fold but are not conjugated to other proteins, and these proteins are instead termed ubiquitin-domain proteins (Jentsch & Pyrowolakis, 2000). Ubls include: SUMO, NEDD8, FAT10, ATG8/12, URM1, and UFM1. Ubls are transferred in a remarkably similar manner to ubiquitination, proceeding via the same three enzyme (E1, E2, E3) cascade, though each family of Ubl usually has its own distinct protein machinery (van der Veen & Ploegh, 2012). Additionally, attachment of Ubls to a substrate occurs in the same fashion, as the C-terminal moieties of most Ubls end in the distinctive Gly-Gly repeat (Jentsch & Pyrowolakis, 2000). Similarly, Ubls often form polymer chains and modification of target proteins with them is reversible by a distinct class of proteases.

Ubls perform a diverse set of roles inside the cellular milieu. Nedd8 (or RUB1 in non-mammals) is the closest related Ubl to ubiquitin (Singh et al., 2012). As previously mentioned, cullin-RING E3 ubiquitin ligases require neddylation in order to properly recognize targets for ubiquitination (Read et al., 2000). The ATG family of Ubls play a role in the formation of autophagosomes (Kabeya et al., 2004). FAT10 is thought to target proteins for proteasomal degradation in a manner similar to ubiquitin (Hipp, Kalveram, Raasi, Groettrup, & Schmidtke, 2005). While the literature for most Ubls is sparse compared to ubiquitin, SUMO (small-ubiquitin-related modifier) has been extensively studied.

The first SUMOylation substrate was RanGAP1, in which SUMOylation was shown to affect its cellular localization (Matunis, Wu, & Blobel, 1998). Since this discovery, nearly 1,000 SUMO substrates have been identified, much more than for any other Ubl (Cappadocia & Lima, 2018). Interestingly, SUMO provides unique insight into the interplay between Ubls and ubiquitin. First, polyubiquitin-SUMO heterotypic chains are found *in vivo* and contribute to proteasomal degradation, DNA repair, and NF-κB activation (Guzzo & Matunis, 2013). The enzymes responsible for SUMO-ubiquitin heterotypic chains, are E3s known as StUbLs – or SUMO-targeted ubiquitin ligases. SUMOylation is required for substrate recognition and ubiquitination by StUbLs, which represents an important cooperation between these two PTMs (Prudden et al., 2007). Finally, SUMOylation and ubiquitination are shown to be at odds during NF-κB signalling, where SUMOylation of IκB prevents ubiquitination of a lysine that would activate signalling (Desterro, Rodriguez, & Hay, 1998). Where SUMO is the best characterized Ubl, UFM1 is perhaps the least well characterized.

1.5.1: UFM1

Ubiquitin fold modifier 1 (UFM1) is a 9.1kDa ubiquitin-like protein that, despite only 14% amino acid sequence similarity with ubiquitin, shares a similar tertiary structure (including the beta grasp "ubiquitin fold" domain). It is one of the more recently discovered Ubls and is conserved amongst plants and animals, but not fungi (Komatsu et al., 2004).

Like ubiquitin, UFM1 is produced first in a pro-UFM1 form that must be processed before conjugation. This processing is achieved by two UFM-specific proteases UFSP1 or UFSP2 to cleave off a serine-cysteine dipeptide at the C-terminal end of the protein (Kang et al., 2007). UFSP2 is the more conserved protease amongst species and is thought to be responsible for the majority of cleavage events. Unlike most other Ubls that end with two glycine residues after processing, the newly exposed C-terminal end of UFM1 is characterized by a valine followed by a glycine. UFSP1 and UFSP2 are also responsible for cleaving UFM1 from substrates, indicating that UFM1 is a reversible PTM.

My interest in UFM1 stems from a yeast two-hybrid protein interaction screen (Appendix A) from the Rohde lab where IpaH9.8 was shown to interact with UFM1 (Jeremy Benjamin, unpublished data). This suggested that finding substrates for IpaH9.8 might involve delving out of the ubiquitin field and into the ubiquitin-like protein field.

1.5.1.1: UFM1 Machinery & Transfer

The protein machinery involved in the UFMylation cascade is much less complex than that of ubiquitin. To date, UFM1 conjugation is understood to involve only a single E1 (UBA5), E2 (UFC1), and E3 (UFL1).

The E1 for UFM1 is UBA5, one of eight total E1s that are responsible for Ubl activation in humans (Schulman & Harper, 2009). Interestingly, UBA5 has been shown to activate both UFM1 and SUMO, which is atypical as E1s are usually highly specific for a single Ubl (M. Zheng et al., 2008). UBA5 is significantly smaller (44.7kDa) than most other E1s with the catalytic located in the adenylation domain (Schulman & Harper, 2009). However, recent mechanistic studies have shown that UBA5 acts a homodimer consistent with other non-canonical E1 enzymes. UBA5 activation is different from canonical activation in a few regards: 1) UBA5 dimerizes and non-covalently associates with UFM1 before activation (Bacik, Walker, Ali, Schimmer, & Dhe-Paganon, 2010), 2) UFM1 binds to a discrete site on each UBA5 in the dimeric complex, which stabilizes the dimer and recruits ATP (Mashahreh et al., 2018), and 3) UFM1 is transferred to the E2 enzyme, UFC1, in a *trans* fashion. That is, UFC1 interacts with one monomer of UBA5 and accepts UFM1 from the other (Oweis et al., 2016).

The E2 enzyme UFC1 shares only an approximately 10-amino acids motif with other E2s, centered around its catalytic cysteine (Komatsu et al., 2004). It accepts UFM1 on this catalytic cysteine through transesterification, in a fashion that is assumed to be similar to other known E2s (T. Mizushima et al., 2007).

The E3 enzyme UFL1, or RCAD, was the latest enzyme in the UFM1 cascade to be discovered (Tatsumi et al., 2010), as it does not resemble any known E3s including HECT or RING or U-box type E3s. The N-terminus of UFL1 is conserved across species and is sufficient for the recruitment of E2 and substrates for UFMylation. Because the N-terminal domain does not contain a catalytic cysteine, UFL1 is thought to act a scaffold-type E3 that does not bind UFM1 directly (Tatsumi et al., 2010).

1.5.1.2: Known Substrates & Function

The first substrate of UFMylation to be discovered was UfBP1, also known as C20orf116, DDRGK1, or Dashurin (Tatsumi et al., 2010). While initially lacking in physiological consequences for UfBP1 UFMylation, recent studies have shown that UFM1 and UfBP1 colocalize to the endoplasmic reticulum (ER) and help prevent ER stress-induced apoptosis – a response to the perturbation of protein folding in the ER (Lemaire et al., 2011). When the ER is unable to maintain homeostasis of protein folding, it activates the unfolded protein response (UPR) to halt general protein translation, express chaperones to help with protein folding, clear irreversibly incorrectly folded proteins, or induce apoptosis if homeostasis cannot be re-established. The UPR uses three different proteins to regulate these processes by turning on specific genes: PERK, IRE1α (through splicing of Xbp-1), and ATF6. A year after the initial finding implicating UFM1 in ER-stress, transcription of UFM1 was shown to be activated by Xbp-1, and that knockdown of UFM1 induced ER Stress and UPR activation (Y. Zhang, Zhang, Wu, Lei, & Li, 2012). Finally, Liu et al. (2017) demonstrated that UfBP1 plays a definitive role in this ER-homeostasis phenotype through UFMylation-dependent stabilization of IRE1α.

The second well-characterized target of UFMylation is ASC1 (TRIP4). ASC1 is a transcriptional co-activator that works alongside the dimeric Estrogen-Receptor Alpha to bind to estrogen receptor responsive elements in genes and turn on their expression (Jung et al., 2002). These genes are involved in cell growth and are essential in the proliferation of a wide variety of breast tumor cells. ASC1 is polyUFMylated, which acts as a scaffold to recruit additional transcriptional co-activators required to turn on these genes (Yoo et al., 2014). Interestingly, UFMylation of ASC1 is dependent on UfBP1, revealing a role

for UfBP1 as an adaptor protein, which has helped other groups reveal potential targets of UFMylation (Zhu et al., 2018).

Additionally, both UfBP1 and UFL1 have been shown to form a large complex with a putative tumor-supressor protein C53 (itself a target of UFMylation, also known as CDK5RAP3) that is thought to downregulate NF-κB signalling (Kwon et al., 2010; Wu, Lei, Mei, Tang, & Li, 2010; Xi, Ding, Zhou, Wang, & Cong, 2013). C53 is thought to be required for some of UfBP1's function in ER-homeostasis (Y. Zhang et al., 2012). In the initial study, knockdown of UFL1 was shown to increase proteasomal degradation of UfBP1, whereas overexpression of UFL1 decreased ubiquitination of UfBP1. This suggests a competitive relationship between UFMylation and ubiquitination, similar to that seen between ubiquitination and SUMOylation. Finally, another recent study mapping the ribosomal interactome suggested that UFMylation of ribosomal proteins was commonplace (Simsek et al., 2017).

Various systemic studies have examined the effect of knocking out components of the UFMylation machinery and their effect on the development of blood cells. UBA5 knockdown in mice leads to defects in erythroid development, anemia, and death in utero (Tatsumi et al., 2011). In another murine model, UFL1 knockdown led to impaired hematopoietic development, most likely by elevating ER stress and UPR activation in bone marrow cells (M. Zhang et al., 2015). Cai et al. showed that knocking out UfBP1, UBA5, and ASC1 all led to similar ER-stress-dependent cell death phenotypes in hematopoietic stem cells (Cai et al., 2015).

Mutations in and dysregulation of UFM1 have also been implicated in other human diseases: ischemic heart disease (Azfer, Niu, Rogers, Adamski, & Kolattukudy,

2006), type II diabetes (Lu, Yang, Allister, Wijesekara, & Wheeler, 2008), hip dysplasia (Watson et al., 2015), encephalopathy (Low et al., 2018; Nahorski et al., 2018), and schizophrenia (Rubio, Wood, Haroutunian, & Meador-Woodruff, 2013). Type II diabetes and ischemic heart disease are both typically associated with induction of ER stress (Y. Zhang et al., 2012). The effect of UFM1 on brain development on the other hand may be linked to UFC1, which has been shown to interact with NCAM, a protein involved in neurodevelopment (Homrich et al., 2014). The importance of UFM1 in these other human diseases, while demonstrable, has yet to be fully elucidated. Finally, UFM1 has been implicated in infectious disease, identified as a gene that may play a role in susceptibility to *Helicobacter pylori* infection (Maran et al., 2013).

I hypothesized that the reason why our lab has not been able to find targets of IpaH9.8 is because IpaH9.8-directed ubiquitination may be dependent, in some fashion, upon UFMylation, as suggested by the two-hybrid assay. However other potential targets were also identified in the two-hybrid screen, one of which is intricately linked to autophagy and is called ZKSCAN3.

1.6: Autophagy

Cell survival is a balancing act between breaking down old and unneeded material (catabolism) and building new structures required for survival (anabolism). Autophagy is a catabolic cellular process that is used to break down unwanted material into building blocks for other anabolic processes (reviewed in (Kaur & Debnath, 2015)). Autophagy is a highly regulated process induced under conditions of nutrient shortage, infection, or various other cellular stresses. There are three main types of autophagy. Microautophagy involves invaginations of the lysosome to directly engulf substrates. Chaperone-mediated

autophagy involves substrates shuttled directly to the lysosome by chaperones and without the formation of the autophagosome. Macroautophagy is the focus of this thesis, and, briefly, proceeds via the generation of a double membrane vesicle called the autophagosome around cargo. The autophagosome fuses with the lysosome and results in the degradation of this cargo into constituent building blocks that are recycled for other uses. The proteins involved in autophagy are encoded by the ATG family genes and they are responsible not just for bulk cytosolic recycling autophagy, but also autophagy of specific cargoes.

The first step in autophagy, induction, proceeds via signalling of some cytosolic stress, such as reduced signalling through the mammalian target of rapamycin (mTOR) program to recruit a variety of ATG proteins, usually starting with the ATG1/ATG13/ATG17 complex (also known as ULK proteins in mammals) (N. Mizushima, 2007). Formation of the ATG/ULK complex leads to signalling through more ATG protein complexes (including Beclin1 and WIPI proteins) and the beginning of the autophagosome, called the phagophore. The next step, nucleation consists of the recruitment of more proteins to the growing phagophore. Importantly, another ATG protein ATG8/LC3-I becomes lipidated by phosphatidylethanolamine to form LC3-II and studs the growing phagophore. Next, in the case of selective autophagy, proteins, aggregates, or organelles are targeted to the growing phagophore by specific adaptor proteins. For example, p62/SQSTM1 binds to ubiquitinated proteins and directs them by binding to LC3-II (Shaid, Brandts, Serve, & Dikic, 2013). The next steps consist of expansion of the phagophore and the formation of the enclosed autophagosome and

subsequent fusion to the lysosome to form the autolysosome, which is also highly regulated by a variety of ATG proteins.

1.6.1: TFEB ZKSCAN Axis

The signals for autophagy are varied but largely dependent on stresses such as oxidative stress, ER stress, pathogen infection, or nutrient starvation. A key sensor of both nutrient starvation and pathogen infection is the kinase mTOR (Tattoli et al., 2012). mTOR regulates and links lysosomal biogenesis genes with autophagy. The mTOR-containing complex mTORC1 senses nutrient levels by measuring digestion-generated amino acids from the lysosome, which result in its activation. Under nutrient-rich conditions mTORC1 phosphorylates a transcription factor TFEB (Settembre et al., 2011; Settembre et al., 2012). Phosphorylated TFEB is unable to transit into the nucleus due to its retention in the cytoplasm by 14-3-3 proteins. During nutrient starvation, mTORC1 is no longer bound to the lysosome which means TFEB is not phosphorylated and is free to enter the nucleus. TFEB activates the transcription of genes containing a co-ordinated lysosomal expression and regulation (CLEAR) sequence in their regulatory region. The protein products of these genes are involved in lysosomal biogenesis and autophagy initiation.

Another transcription factor, ZKSCAN3, plays an antagonistic role to TFEB (Chauhan et al., 2013). Phosphorylation of ZKSCAN3 is analogous to that of TFEB, leading to its exclusion from the nucleus. Under nutrient-rich conditions, TFEB is sequestered in the nucleus, but ZKSCAN3 is free to move into the nucleus where it represses these very same CLEAR sequence-containing genes. In this way TFEB and ZKSCAN3 play opposing roles to regulate lysosomal biogenesis and the induction of

autophagy (reviewed in (Saftig & Haas, 2016)). Furthermore, recently ZKSCAN3 and TFEB have been shown to be induced through an mTOR independent pathway: indirectly by PKC, through the MAPKs JNK and p38 (Y. Li et al., 2016).

1.6.2: Autophagy and *Shigella*

Autophagy is not only a way to recycle cellular building blocks, but also plays a role in the clearance of intracellular pathogens. This branch of autophagy is called xenophagy. *Shigella* has evolved specific mechanisms to modulate the xenophagic response to infection in the form of its repertoire of protein effectors.

Xenophagy of *Shigella* proceeds after rupture of the bacterial-containing vacuole (BCV). Cytosolic *Shigella* is ubiquitinated by specific E3 ligases such as LRSAM1 that direct the formation of K27-linked polyubiquitin chains on the surface of the bacteria. These chains interact with adaptor-specific adaptor proteins such as SQSTM1 or NDP52. (von Muhlinen, Thurston, Ryzhakov, Bloor, & Randow, 2010). These adaptor proteins link the process of ubiquitination and autophagy because they contain motifs that bind to ubiquitinated *Shigella* as well as LC3 on the phagophore to specifically target bacteria to the autophagosome (Cemma, Kim, & Brumell, 2011). More recently, LC3-positive autophagosome recruitment has been linked to *Shigella* that have yet to escape the BCV. *Shigella* T3SS effectors IcsB and VirA, as well as host factor Toca-1 prevent this LC3 recruitment (Baxt & Goldberg, 2014; Campbell-Valois, Sachse, Sansonetti, & Parsot, 2015). Galectins, host sugar binding proteins, have also been implicated in targeting intracellular bacteria for xenophagic clearance. Galectin-8 has been shown to accumulate at sites of BCV membrane damage and also to recruit NDP52 in the case of *Salmonella* (Thurston, Wandel, von Muhlinen, Foeglein, & Randow, 2012). Additionally, NOD-like

receptors have been implicated in recruiting autophagic machinery to clear *Shigella* (Travassos et al., 2010).

Salmonella has also been shown to *increase* autophagy during infection (Yu et al., 2014). The authors of this study propose that an increase in autophagy may provide the necessary building blocks for *Salmonella* intracellular replication. This demonstrates that while some bacteria may downregulate the xenophagic response during infection, increasing generalized macroautophagy may be an alternate productive pathogenic strategy.

1.7 Research Described in Thesis

The research in this thesis focuses on the characterization of *Shigella* protein IpaH9.8 and its interaction with two binding-partners: UFM1 and ZKSCAN3. To determine whether IpaH9.8 and UFM1 interact, as suggested by the yeast two-hybrid assay and to determine the nature of this interaction, *in vitro* ufmylation assays were conducted using purified E1, E2, and E3 enzymes. IpaH9.8 was shown to catalyze the formation of high-molecular weight species consistent with ufmylation. Furthermore, mass spectrometry was used to determine the suite of ufmylated human proteins with and without IpaH9.8 present. From this list of proteins, Galectin-7 was examined for ufmylation *in vitro* and in a cell culture model and results suggest that Galectin-7 is UFMylated in an IpaH9.8-dependent fashion. A fluorescent-tagged construct of UFM1 was created and used to examine distribution of IpaH9.8 and UFM1 inside human cells. This approach was also applied to ZKSCAN3 IpaH9.8 induced foci formation of ZKSCAN3. The effect of IpaH9.8 on ZKSCAN3 lysosomal gene targets was also indirectly examined through the lysosome-tracking dye lysotracker and electron

microscopy. Lysotracker but not electron microscopy suggested that IpaH9.8 might increase lysosome number or size. I propose that UFM1 and IpaH9.8 functionally interact in one of three different potential models to represent as-yet-unknown communication between ubiquitination and ufmylation. Additionally, I propose that IpaH9.8 might ubiquitinate ZKSCAN3 to modulate lysosome and autophagic levels inside human cells by affecting the ability of ZKSCAN3 to re-localize and repress its target genes.

Chapter 2: Methods

2.1: Bacterial Strains & Culture Maintenance

Escherichia coli strains DH5α (F^- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(r_K^-, m_K^+) *pho*A *sup*E44 λ^- *thi*-1 *gyr*A96 *rel*A1) and BL21 (F^- *omp*T *hsd*S_B (r_B^-, m_B^-) *gal dcm* (DE3)) were used for propagation of plasmids and expression of proteins respectively. *E. coli* was grown overnight at 37°C with shaking in liquid, or on solid lysogeny broth (LB) medium (10mg/mL tryptone, 5mg/mL yeast extract, 10mg/mL sodium chloride, with/without 20mg/mL agar). As required, the following antibiotics were used: ampicillin 100µg/mL, kanamycin 25µg/mL. Frozen stocks (1mL overnight culture and 0.5mL 50% glycerol) were kept at -80°C.

Calcium chloride-competent cells were prepared for transformation according to Sambrook and Russell (2001). Briefly, overnight cultures were subcultured 1:100 into fresh LB media and grown until mid-log phase (OD_{600} = 0.4 - 0.6). The cultures were cooled on ice to halt growth and pelleted at 5000xg for 15 minutes at 4°C. Cells were washed twice in ice cold 100mM $CaCl_2$ and pelleted after each wash. Cells were resuspended in 1/100 of starting volume in 100mM $CaCl_2$, 10% glycerol and stored at -80°C.

Heat shock transformation proceeded as follows. Calcium chloride-competent BL21 cells (100µL) were thawed on ice per plasmid. Approximately 100ng of DNA was added to the cells and incubated for 30 minutes. The cells were then incubated for 45 seconds at 42°C and then incubated for 1-2 minutes on ice. LB medium (1mL) was added and cells were recovered at 37°C for 45-60 minutes. This mixture (100µL) was plated on

selective LB solid medium with the appropriate antibiotic and incubated overnight at 37°C.

2.2: Purification of Proteins from *E. coli*

The following plasmids were created by Rohde lab member Julie Ryu: pET303/CT-His-UFC1, pET303/CT-His-UBA5, pGEX-6P-1-IpaH9.8, and pGEX-6P-1-IpaH9.8(C337A). These plasmids were used to express hexahistidine-tagged (His) UFC1 and UBA5 and Glutathione S-Transferase-tagged (GST) IpaH9.8 and IpaH9.8 (C337A) respectively. A pET303/CT-His-UFL1 plasmid encoding His-tagged UFL1 was also created by McCormick lab member Mariel Kleer. Cobalt bead affinity purification was used to purified His-tagged proteins, while reduced glutathione agarose bead affinity purification was used to purify GST-tagged proteins.

A single colony of a BL21 *E. coli* transformant was used to inoculate 5mL of LB medium. The cultures were incubated overnight and then diluted 1:100 into 500mL of LB medium and grown until mid-log phase. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to 1mM final concentration and the cultures were allowed to grow for an additional 2 hours. The cultures were cooled on ice and cells were pelleted at 5000xg for 15 minutes at 4°C. The pellets were resuspended in 5mL of NETN buffer (20mM Tris pH 8.0, 100mM NaCl, 1mM EDTA, 0.5% NP-40, Pierce protease inhibitor tablet [ThermoFisher CAT#88666], lysozyme 3mg/mL). The resultant cells were subjected to freeze-thawing twice and then sonicated 3x 10 seconds on ice using a probe sonicator. Lysates were centrifuged for 30 minutes at 5000xg at 4°C and the supernatant was collected.

For purification of GST-tagged proteins, lysates were incubated with 1mL of pre-NETN-washed Pierce Glutathione Agarose beads [ThermoFisher CAT#16100] for 1 hour with end-over-end rotation at 4°C. The beads were washed 3x with RIPA buffer [150mM NaCl, 1%NP-40, 0.1% sodium dodecyl sulfate (SDS), 0.5% sodium deoxycholate, 50mM Tris-HCl pH 8, 1mM EDTA] and 1x with NETN and fractions were collected by pelleting for 2 minutes at 700xg and 4°C. The bound fraction was eluted with 1mL of elution buffer containing reduced glutathione (10mM glutathione, 50mM Tris-HCl, pH 8).

For purification of His-tagged fusion proteins, lysates were resuspended in 10mL His wash buffer (20mM Na_2PO_4, 500mM NaCl, pH 8.0) and incubated with 1mL HisPur Cobalt Resin [ThermoFisher CAT#89964] for 1 hour with end-over-end rotation at 4°C. Beads were washed 3x with wash buffer and fractions were collected by pelleting beads at 700xg for 2 minutes at 4°C. The eluted fraction was collected by eluting in 1mL His Elution Buffer (150mM imidazole, 20mM Na2PO4, 500mM NaCl).

For both purification of His-tagged and GST-tagged fusion proteins, total protein stain visualization was used to determine protein purity. These purified proteins are hereafter referred to as 'homemade'.

2.3: SDS PAGE & Immunoblotting

SDS-polyacrylamide gels (6%-15%) were made as described by Sambrook and Russell (2001) and cast using a Hoefer gel caster system with running buffer (25mM Tris, 192mM glycine, 0.1% SDS). Proteins samples were quantified using the DC Protein Assay system [Bio-Rad CAT# 5000111] and were combined with an equal volume protein sample buffer (60mM Tris-HCl pH6.8, 10% glycerol, 1% SDS, 0.1M DTT).

Protein samples were boiled for 5 minutes prior to loading or storage at -20°C and were resolved at approximately 10-20µg of total protein per lane. NEB Color Prestained Broad Range Protein Standard [CAT#P7712S] was used as a protein ladder. Gels were run for 1.5 to 2.5 hours at 110-150V or as long as it took the loading front to reach the end of the gel.

For immunoblot analysis, a Hoefer wet transfer apparatus was used with Towbin transfer buffer (25mM Tris, 192mM glycine, 20% (v/v) methanol, 0.1% SDS) to transfer proteins to polyvinylidene difluoride (PVDF) membranes. Membranes were first soaked in methanol. Proteins were transferred at constant amperage 0.15A for 2 hours or 0.05A overnight. Membranes were subsequently blocked for 1 hour at room temperature on a shaking incubator with 5% skim milk powder in Tris-buffered saline with Tween-20 (TBST, 145mM NaCl, 5mM Tris-HCl pH7.5, 0.1% Tween-20). Primary antibody was used in the concentration described in Table 2 diluted in the same 5% skim milk solution as above. Membranes were incubated in primary antibody for 1 hour.

Table 2: Antibody dilution and purchasing information.

Abbreviations: (WB) western blotting, (IF) immunofluorescence.

Antibody	Dilution	Source
UFM1	1:1000	BostonBiochem CAT#A-500
IpaH	1:1000	Eurogentec see reference (Mavris et al., 2002)
Penta-His	1:1000	Qiagen CAT#34660
GST	1:1000	SantaCruz CAT#sc33613
UBCH5B	1:1000	BostonBiochem CAT#A-615
Galectin-7	1:7000	R&DSystems CAT#AF1339
UFL1 (K0776)	1:2000	Abcam CAT#ab226216
UFC1	1:1000	Abcam CAT#ab189252
Myc 71D10 (rabbit)	1:1000	Cell Signalling CAT#2278
Myc 9B11 (mouse)	1:1000 (1:8000 IF)	Cell Signalling CAT#2276

Antibody	Dilution	Source
ZKSCAN3	1:1000	Sigma CAT#SAB2700902
HA	1:1000	Cell Signalling CAT#3724
B-Actin	1:1000	Cell Signalling CAT#4970
Anti-Mouse HRP (WB)	1:2000	Cell Signalling CAT#7076
Anti-Rabbit HRP (WB)	1:2000	Cell Signalling CAT#7074
Donkey anti-mouse 488 (IF)	1:1000	ThermoFisher CAT#R37114
Anti-Goat HRP (WB)	1:2000	Sigma CAT#A8919

Membranes were washed for 3x10 minutes with TBST. Secondary antibody was diluted as before and incubated with membranes for 1 hour. Membranes were washed again 3x10 minutes with TBST. Membranes were developed using Pierce ECL Plus Substrate kit [ThermoFisher CAT#32132] and imaged using a Bio-Rad ChemiDoc Touch Imaging System [Bio-Rad CAT# 17001401].

Gels that were not subject to immunoblot analysis were instead stained for total protein analysis using Imperial Protein Stain [ThermoFisher CAT#24615]. Gels were washed for 5 minutes with ddH_2O then covered with Imperial Protein Stain. Gels were stained with agitation at room temperature for 1-2 hours or overnight. Gels were destained using ddH_2O for 2-4 hours with shaking at room temperature and imaged using a Bio-Rad ChemiDoc Touch.

2.4: *In Vitro* Ubiquitination / UFMylation Assays

The following amounts of proteins were used for *in vitro* ubiquitination and UFMylation assays: 2 µg HA-tagged ubiquitin (HA-Ub) [BostonBiochem CAT# U-110], 0.5 µg UBA1 [BostonBiochem CAT# E-305], 2 µg UBCH5B [BostonBiochem CAT# E2-622], 1 µg homemade GST-IpaH9.8 or GST-IpaH9.8(C337A), 1 µg GST-ZKSCAN3 (purified by Rohde lab member Kaitlyn Tanner), 2 µg His-UFM1 [BostonBiochem CAT# UL-500], 0.5 µg His-UBA5 [BostonBiochem CAT# E-319] or 0.5 µg homemade

UBA5-His, 2 µg His-UFC1 [BostonBiochem CAT# E2-675] or 2 µg homemade UFC1-His, and 1 µg Galectin-7 [Cedarlane CAT# CLCYT016-2] (Figure 2). Amounts of protein were also increased if multiple western blot analyses were needed. Proteins were combined with charging buffer (25mM Tris pH 7.5, 50mM NaCl, 5mM MgCl2, 4mM ATP, 0.25mM DTT). Reactions were started by brief centrifugation to combine reaction components. Reactions proceeded at room temperature for 1-2 hours as indicated and were then split equally in two. One half was stopped with an equal volume of Urea Stop Buffer (50mM Tris pH 6.8, 2% SDS, 4M Urea, 10% glycerol) and the other stopped with an equal volume of dithiothreitol (DTT) Stop Buffer (50mM Tris pH 6.8, 2% SDS, 100mM DTT, 10% glycerol). Samples were immediately boiled for 5 minutes and analyzed via SDS-PAGE and western blotting or stored at -20°C.

2.5: Mass Spectrometry of *In Vitro* UFMylation Gel Bands

In vitro UFMylation reactions were separated via SDS-PAGE and visualized using Imperial Protein Stain as described in section 2.2. Six gel slices of approximately 1mm x 6.5 mm were excised using a gel cutting pipette tip. The slices were incubated in dH_2O for 2 hours and then cut into 1mm x 1mm slices and washed in dH_2O twice. Samples were reduced with 10mM DTT at 56°C, then alkylated with 55mM iodoacetamide in the dark at room temperature, both for 30 minutes. Slices were dried with 200µl of acetonitrile and immersed in 20ug/mL of trypsin [ThermoFisher CAT#90057] for 2 hours. Samples were then incubated at 37°C overnight with the addition of 50mM ammonium bicarbonate. A mix of of 50% acetonitrile in 5% formic acid (100µl) was used to extract peptides, which were subsequently dried using a vacuum centrifuge. Pellets were resuspended in 20µl 3% acetonitrile and 0.5% formic acid.

Processed gel slices were analyzed via electrospray liquid chromatography-tandem mass spectrometry (LC-MS/MS) using a VelosPRO orbitrap mass spectrometer with an UltiMate 3000 Nano-LC. The initial separation occurred at 300nL/min using a PicoFRIT C18 self-packed 75mm x 60cm capillary column [New Objective CAT#PF360-50-##-CE-5]. Mass spectrometry data was acquired at a resolution of 30,000 with 10 succesive MS/MS spectra in HCD (higher-energy collisional dissociation) and CID (collision-induced dissociation) mode.

Raw data was analyzed by Alejandro Cohen of the Dalhousie CORE proteomics facility using Proteome Discoverer 2.2. Spectra were searched against a manually generated list of proteins present in the reaction mixtures as well as the cRAP database of common MS contaminants. The following custom modifications were set: methionine (Met) oxidation, N-terminal Met loss, and UFMylation (VG) on lysines or cysteines. A second round of data analysis included UFMylation on additional amino acid residues as an additional modification.

2.6: Cell Culture & Maintenance

HeLa, HEK293T, HEK293A cell lines, and HEK293T cells stably transduced with a retrovirus expressing human codon-optimized IpaH9.8, IpaH9.8(C337A), or empty vector (Tanner, 2014) were obtained from Rohde lab frozen stocks. Cells were maintained in a 5% CO_2 incubator at 37°C and passaged every 3-5 days upon reaching 80-90% confluency. When cells reached 80% confluence, medium was removed and cell culture flasks were washed with 5mL PBS (phosphate buffered saline). Trypsin was used to lift adherent cells, specifically 0.05% trypsin with 0.5M EDTA [Invitrogen CAT#25300-054] with a contact time of 10-15 minutes at 37°C. The flask was washed

repeatedly in 5mL medium until cells were evenly resuspended and 0-3 drops of media were left in the flask. Fresh medium (10mL) was added to the flask and incubated as described earlier.

For HeLa, HEK293T and HEK293A cells, Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS) was used. For stably-transduced cells, puromycin (1μg/mL) was added to the medium to select and maintain stable integration.

2.6.1: PEI Transfection

The following mammalian-expression plasmids were used to transfect HEK293T, HEK293A, HeLa, or stably-transduced HEK293T cells: Myc-tagged IpaH9.8 (prK5-myc-IpaH9.8 – constructed by John Rohde), His-tagged UFM1 (pCR3.1-His-UFM1 – constructed by Carolyn Robinson), empty vector (pCR3.1), N-terminal mRuby3-tagged UFM1 (mRuby3-N-UFM1), N-terminal mRuby3-tagged ZKSCAN3 (mRuby3-N-ZKSCAN3), and a GFP control (pcDNA3.1-eGFP).

Cells were seeded onto 12-well, 10cm, or 15cm cell culture plastic dishes at a density of 0.1×10^6, 2.2×10^6, and 5.0×10^6 (per well) cells respectively. After 24 hours, the appropriate amount of polyethylinamine (PEI) was mixed with half with required Opti-MEM reduced-serum media, while the appropriate amount of mammalian expression vector was incubated with the other half of the required volume of Opti-MEM (Table 3). Both solutions were incubated at room temperature for 5 minutes before combining and incubating for another 15 minutes. Meanwhile, seeded plates were washed 3 times with serum-free DMEM and the appropriate amount of medium was replaced in the dish (Table 3). Dropwise, the PEI-DNA-Opti-MEM mixture was added to

the seeded plates and incubated at 37°C 5% CO_2 for 4-6 hours. After the allotted time, medium was aspirated and replaced with DMEM with 10% FBS. Transfection efficiency was verified 24 hours post-transfection with a GFP control.

Table 3: List of transfection reagents.

Required volumes of DNA, PEI, Opti-MEM and Serum-Free DMEM required for optimum transfection of different sized dishes of cells.

Size of Dish	Amount of DNA (μg)	Amount of PEI (μL)	Volume of Opti-MEM (mL)	Volume of Serum-Free Media (mL)
12-Well	0.5	1.5	0.2	0.8
10 cm	6.0	18.0	1.0	9.0
15 cm	18.0	54.0	3.0	20.0

2.7: Co-Precipitation of UFM1 and IpaH9.8

HEK293T cells were transfected with His-tagged UFM1 and Myc-tagged IpaH9.8. Cobalt bead affinity purification was used to detected proteins that interact with His-tagged UFM1. Likewise, protein A agarose bead-Myc antibody immunopurification was used to detect proteins that interact with Myc-tagged IpaH9.8.

2.7.1: 6x-His Fusion Protein Co-Affinity Precipitation

Six 15-cm dishes were seeded with 5.0 x 10^6 HEK293T cells stably-transduced with IpaH9.8. Three plates were transfected with pCR3.1-His-UFM1 and three plates were transfected with a negative control empty pCR3.1 vector. After 24 hours, medium was removed and cells were washed twice in cold PBS. One to two mL of 20mM Tris-Cl pH7.4 was added to one plate and cells were collected using a cell scraper. This sample was transferred to one of the other plates and the process repeated twice so that the total

sample volume from 3 plates did not exceed 2mL. Samples were washed twice with 20mM Tris-Cl and cells were collected by centrifugation at 500xg for 2 minutes in between washes. Samples were left to swell on ice for 30-60 minutes and lysed using a 21G needle. Lysates were clarified by centrifugation for 10 minutes at 4°C at maximum speed in a benchtop microfuge and the supernatant was collected.

The supernatant was applied to 1mL of HisPur Cobalt Resin and His-tagged proteins were affinity purified as described in section 2.3.2 with the exception that the wash buffer (20mM Na2PO4, 500mM NaCl) was supplemented with 0, 5, 7.5, or 10mM imidazole and optimized at 5mM. Whole cell lysate, unbound, washes, and eluted fractions were collected.

2.7.2: Myc Fusion Protein Co-Immunoprecipitation

Four 15-cm dishes were seeded with HEK293T cells at 5.0 x 10^6 cells per plate. Two plates were transfected with equal amounts prK5-Myc-IpaH9.8 and pCR3.1-UFM1-His, while the remaining plates were transfected with only pCR3.1-UFM1-His. Transfections were carried out as described in section 2.6. At 24 hours post-transfection, medium was removed and cells were collected in PBS. Samples were centrifuged at 300xg for 5 minutes at 4°C and washed once with PBS by repeating centrifugation. Samples were resuspended in 500µl – 1mL lysis buffer (2% SDS, 10% glycerol, 62.5 mM Tris pH6.8) and lysed with a 21G needle or with sonication. Lysates were clarified by centrifugation at 13,000xg for 15 minutes at 4°C and the supernatant was collected.

Lysates were diluted 5-fold with cold IP Buffer (50mM Tris pH8.0, 150mM NaCl, 1% NP-40) and αMyc [Cell Signalling CAT#71D10] antibody was added at a 1:1000 dilution. Lysates were incubated with αMyc antibody at 4°C overnight. The next

day, lysates were combined with 1mL Protein A Agarose Beads [Sino Biological CAT#10600-P07E-RN] pre-washed with IP Buffer. Lysates were incubated with beads for 1-3 hours. Beads were collected by centrifugation at 700xg for 2 minutes and washed three times with 1 volume IP buffer for 10 minutes, with centrifugation as before to collect beads. To elute bound proteins, beads were resuspended in 2x Laemmli buffer (4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.125M Tris HCl, pH 6.8) and boiled for 5 minutes. Supernatant was collected by centrifugation at 300xg for 2 minutes and analyzed via SDS-PAGE.

2.8: Mass Spectrometry of Co-precipitated proteins

His-tagged proteins were affinity-purified as described previously from two samples: HEK293T cells stably expressing untagged-IpaH9.8 transfected with pCR3.1-UFM1-His, and HEK293T cells stably expressing empty pBMN1 vector transfected with pCR3.1-UFM1-His. The lysis and wash buffers were supplemented with 1% and 0.5% SDS respectively. Otherwise the co-precipitation process was the same. Eluted fractions for both samples were subjected to cold acetone precipitation. Briefly, 4 times the sample volume of -20°C acetone was added to each sample, vortexed, and incubated for 60 minutes at -20°C. Samples were subject to centrifugation for 10 minutes at 13,000-15,000xg at 4°C, and the supernatant was decanted. An additional acetone wash was repeated as necessary. The protein pellets were dried by evaporating the acetone at room temperature for 30 minutes. The pellet was resuspended in a low volume (200-300µl) of buffer containing 8M urea, 0.4M ammonium bicarbonate and sonicated for 3 x 10 seconds pulses. DTT (10µl 0.5M) was added to each sample and incubated for 30 minutes at 60°C before cooling for 5 minutes at room temperature. $IAcNH_2$ (20µl 0.7M)

was added and samples were incubated for 30 minutes at room temperature before adding to 1.2mL of dH_2O.

Processed samples were digested with 100µl of 0.02µg/µl trypsin in 50mM ammonium bicarbonate overnight at 37°C with shaking. The next day 1µl of trifluoroacetic acid was added to the sample and the pH was lowered to pH 3.0 with formic acid. Samples were desalted according to the PicoFrit column specifications and dried via vacuum centrifugation. Pellets were resuspended in 20µl of 3% acetonitrile in 0.1% formic acid and sonicated for 15 minutes until resuspended. Samples were processed via LC-MS/MS (section 2.5) and were analyzed as before with an additional analysis step that included comparison to a human and bacterial proteome database.

2.9: Fluorescent Microscopy

2.9.1: Cloning of Constructs

To create plasmids that would express UFM1 and ZKSCAN3 tagged with mRuby3 and mClover 3, four plasmids were donated by the McCormick lab: a C-terminal mCherry tagging vector (pmCherry-C1), an N-terminal mCherry tagging vector (pmCherry-N1), an mRuby3 Histone H2B fusion protein vector (pKanCMV-mRuby3-10aa-H2B), and an mClover3 Histone H2B fusion protein vector (pKanCMV-mClover3-10aa-H2B). mRuby3 and mClover 3 coding regions were amplified using primers described in Table 4. Each gene sequence was amplified twice, using the same forward primer and a different reverse primer for insertion into a C-tagging or N-tagging pmCherry vector. Additionally, ZKSCAN3 and UFM1 were amplified from pcDNA3-ZKSCAN3 and pCR3.1-UFM1-His (primers in Table 4). After amplification PCR products were purified using a QIAquick PCR Purification Kit (Qiagen CAT#28104).

mRuby and mClover 3 PCR products and both pmCherry plasmids were digested using AgeI-HF and XhoI for C-tagging constructs, and AgeI-HF and NotI-HF for N-tagging constructs. ZKSCAN3 and UFM1 were digested with EcoRI-HF and BamHI-HF for UFM1 and EcoRI-HF and SalI for ZKSCAN3. Digests proceeded for 1 hour at 37°C in the supplied buffer. All restriction enzymes were purchased from New England Biolabs (NEB). Restriction products were resolved by electrophoresis in 1% agarose gels in TBE (45mM Tris-borate, 1mM EDTA) run at 120V for 30-45 minutes and with NEB 1kb DNA ladder (NEB CAT#N3232L) as a molecular weight standard. Products of the appropriate size were gel extracted using the QIAquick kit. Ligation of vectors with appropriate PCR inserts was done overnight at room temperature using T4 DNA ligase (NEB CAT#M0202L) in the supplied buffer at a ratio of 3 x insert: 1 x vector. Ligation products were used to transform $CaCl_2$-competent DH5α *E. coli* cells as described in section 2.1, and selected for overnight on LB-kanamycin agar plates.

Table 4: Primer sequences for amplification of mRuby3, mClover3, ZKSCAN3, and UFM1 coding regions.

Construct	Primers
mRuby3/mClover3 C-Tagging Construct	FWD: 5'-GATCACCGGTCGCCACCATGGT-3' REV: 5'-GATCGCGGCCGCTTTACTTGTACAGCTCGTCCATG-3'
mRuby3/mCover3 N-Tagging Construct	FWD: 5'-GATCACCGGTCGCCACCATGGT-3' REV: 5'-GATCCTCGAGATCTGAGTCCGGACTTGTACAGCTCGTCCATG-3'
N-Tagged Ruby/Clover ZKSCAN3	FWD: 5'-ATAGAATTCAGCTAGAGAATTAAGTGAAAGCA-3' REV: 5'-ATAGTCGACCTACTGTGATAGGATGTTTTTCCC-3'
N-Tagged Ruby/Clover UFM1	FWD: 5'-ATAGAATTCTCGAAGGTTTCCTTTAAGATC-3' REV: 5'-ATAGGATCCTTATCCAACACGATC

Construct	Primers
	TCTAGG-3'

Plasmids were extracted from transformants and purified using a QIAprep spin miniprep kit (Qiagen CAT#27104). Successful clones were identified via diagnostic restriction digest and/or sequencing by Genewiz and maintained as plasmids stocks at -20°C or frozen bacterial stocks at -80°C.

2.9.2: Preparation of Slides for Imaging Fusion Proteins

HEK293A cells were seeded at 0.1 x 10^6 cells per well on 18mm circular coverslips [VWR CAT#631-0153] pre-washed with PBS in a 12-well dish. Cells were transfected with N-terminal Ruby3 fusion proteins UFM1 and ZKSCAN3, or an empty vector control according to Chapter 2.6.1. 24 hours later, medium was removed from wells and coverslips were washed once with PBS. 500µl of 4% paraformaldehyde (PFA) solution was added to each well and cells were fixed by incubation for 15 minutes at room temperature. PFA was then aspirated and coverslips were washed with PBS once more. PBS was replaced and cells were either stored for up to one week at 4°C or immediately stained with Hoechst nuclear stain [ThermoFisher CAT#62249]. Hoechst was diluted 1:2000 in PBS and 500µl was added to each well. Coverslips were incubated for 15 minutes at room temperature in the dark. Next, Hoechst solution was removed and coverslips were washed three times with PBS. Coverslips were mounted face down in 15µl ProLong Gold Antifade reagent [Invitrogen CAT#P36930] on 25 x 75 x 1mm microscope slides [ThermoFisher CAT#12-555-3]. Coverslips were adhered overnight and imaged as described in Chapter 2.9.4.

2.9.3: Immunofluorescence of IpaH9.8-Myc

HEK293T cells were transfected with prK5-myc-IpaH9.8 and fixed as described in section 2.92. Coverslips were blocked using IF blocking buffer (1% human serum, 0.1% Triton X-100, in PBS) for 1 hour at room temperature with agitation. Coverslips were washed with PBS and incubated in primary antibody by inversion onto 50µl of solution (αMyc antibody 9B11 diluted 1:8000 in IF blocking buffer) overnight at 4°C. Following incubation, coverslips were returned to the 12-well dish and washed three times with PBS for 5 minutes with agitation at room temperature. Secondary antibody solution (donkey anti-mouse Alexa 488 diluted in IF blocking buffer) was added to each well and incubated for 1 hour with agitation. Coverslips were washed with PBS then incubated with 4′,6-diamidino-2-phenylindole (DAPI) nuclear stain (diluted 2:100, then 1:1000 into PBS) for 5 minutes in the dark. Finally, coverslips were washed three times with PBS and mounted as before.

2.9.4: Microscopy

Slides were imaged at the Dalhousie University Miscroscopy CORE facility using a Zeiss Axiovert 200M and Hamamatsu Orca R2 Camera. Images were prepared using GNU Image Manipulation Program (GIMP) and Microsoft Powerpoint.

2.10: Fluorescent Imaging of Lysosomal Compartments

HEK293T cells were seeded on coverslips as described in Chapter 2.9.2. To induce lysosomal biogenesis, after 24 hours, the medium was replaced with media containing either 250nM of the drug Torin [InvivoGen CAT#inh-tor1] diluted in DMSO or DMSO alone. Cells were treated for 4 hours at 37°C, washed with Torin-free medium before replacement with Lysotracker [ThermoFisher CAT# L7528]-supplemented media

at a concentration of 75nM to fluorescently stain the lysosomal compartment. Lysotracker treatment proceeded for 2 hours at 37°C before washing and replacement with either DMEM or PBS to reduce background signal. Cells were imaged on a EVOS FL Cell Imaging System [ThermoFisher CAT#AMF4300].

2.11: Electron Microscopy

Stably-transduced HEK293T cells treated with Torin or DMSO (as described in Chapter 2.10) were grown to 80% confluency and trypsinized with 1mL trypsin at 37°C for 10 minutes. Cells were resuspended in 10mL DMEM and pelleted at 700xg for 3 minutes. Pellets were fixed for 1 hour at room temperature with 3% glutaraldehyde in a fixation buffer (0.05M sodium cacodylate, 0.25M sucrose). Following fixation, samples were incubated for 2 hours in 2% OsO_4 diluted in fixation buffer, and stained overnight in 0.1% aqueous uranyl acetate as a negative stain. Cell pellets were dehydrated in ethanol and embedded in TAAB embedding resin [TAAB CAT#E037]. Micrographs were taken using a 120kV JEOL 1230 transmission electron microscope at an operating voltage of 80.0 kV.

Chapter 3: Results

3.1: Substrates for IpaH9.8

Our lab has previously shown that mice infected with *Shigella* lacking IpaH9.8 have decreased bacterial burdens and improved outcomes (Tanner, 2014), suggesting a role for IpaH9.8 in virulence. In an effort to find substrates for the *Shigella* protein IpaH9.8, Jeremy Benjamin, a previous member of the Rohde lab, conducted a yeast two-hybrid genetic protein-interaction screen using a catalytically dead variant (C337A) of

IpaH9.8 as bait (Appendix A). Of the library of mammalian prey proteins that interacted with IpaH9.8, two were chosen for further exploration based on two criteria. ZKSCAN3 was chosen because it was one of the highest confidence interactions (i.e. multiple clones appeared in the screen) besides known contaminants and potential false-positive DNA-binding proteins such as PCNA (Bowman, O'Donnell, & Kuriyan, 2004). Secondly, AvrPtoB, the first BEL to be discovered, was shown to interact with ubiquitin in an analogous yeast 2-hybrid assay (Abramovitch et al., 2006). As such, UFM1 was chosen because it represented the only UBL to appear in the screen.

3.2: UFM1

3.2.1: Protein Purification

To confirm the interaction between IpaH9.8 and UFM1, these proteins were expressed in *E. coli* and purified from the cell lysates as were the requisite E1 (UBA5) and E2 (UFC1) enzymes. Plasmids capable of expression of His-tagged UBA5 and UFC1, and GST-tagged IpaH9.8 were created by previous Rohde lab member Julie Ryu (Figure 2). Additionally, a plasmid encoding an active site cysteine to alanine point mutant (C337A) abrogating the catalytic function of IpaH9.8 was constructed. These tagged proteins were purified using cobalt and glutathione agarose affinity chromatography respectively (Figure 3). Additionally, His-tagged UBA5, UFC1, and UFM1 were purchased from BostonBiochem. For a schematic of each protein construct, see Figure 2. It is worth noting that the homemade UBA5 is a naturally-occurring, uncharacterized variant that lacks the first 56 amino acids compared to the full-length BostonBiochem protein. The expected mobility of each protein is described in Table 5.

Table 5: Approximate observed sizes of proteins detected by immunoblotting.

Protein Name	Approximate Size
His-UFM1	9 kDa
HA-Ubiquitin	9 kDa
Full-length His-UBA5 (BostonBiochem)	52 kDa
Truncated His-UBA5 (Homemade)	45 kDa
His-UFC1 (HM or BB)	22 kDa
His-UFL1	100 kDa
GST-IpaH9.8/C337A	80 kDa
Myc-IpaH9.8/C337A (or untagged)	58 kDa
UBA1	113 kDa
UBCH5B	25 kDa
ZKSCAN3	61 kDa
Galectin-7	15 kDa

3.2.2: *In Vitro* UFMylation

My first goal was to establish that the purified ubiquitination proteins (UBA1 and UBCH5B, previously purified in the Rohde lab) were active and when incubated with IpaH9.8 and UFM1 gave the expected banding pattern indicative of a functional interaction (Rohde et al., 2007). In that case, as demonstrated in Figure 4A, the addition of IpaH9.8 disrupts a prominent ~20kDa band corresponding to charged E2 (UBCH5B~Ub). The bulk of this band is replaced by a higher molecular weight smearing or laddering consistent with ubiquitin bound to IpaH9.8. These smears are sensitive to DTT, a reducing agent that disrupts thiol linkages – such as those expected between the IpaH9.8 catalytic cysteine and ubiquitin (Singer et al., 2008; Zhu et al., 2008). I performed similar experiments using purified components of the UFM1 program and observed that addition of IpaH9.8 into reaction mixtures resulted in the loss and addition of similar new species (Figure 4B). An abundant species of ~30kDa in lane 4 is the size expected for a charged UFC1 species and disappears with the addition of IpaH9.8, replaced by a high molecular weight smearing. Incubation of UBA5 and IpaH9.8 in the

absence of UFC1 was also sufficient to induce similar smearing. These data suggest that IpaH9.8 *functionally* interacts with the UFM1 system *in vitro,* consistent with a model where IpaH9.8 acts as a dual-function E3 ligase.

Furthermore, both homemade and BostonBiochem UBA5 and UFC1 enzymes reacted with IpaH9.8 (Figure 5). The homemade His-tagged UBA5 and UFC1 gave high molecular weight smearing when incubated with His-UFM1 and blotted for the His tag (Figure 5A). This smearing is decreased when IpaH9.8 is added, suggesting that IpaH9.8 may be disrupting UFM1-UBA5 or UFM1-UFC1 conjugates. The IpaH9.8-dependent collapse of His-tagged material is also visible in Figure 6B. These high molecular weight species are not observed with the BostonBiochem enzymes. However, when blotting for IpaH, the BostonBiochem enzymes result in high molecular weight IpaH9.8 conjugates, which are decreased with DTT as well as, to a lesser degree, the C337A IpaH9.8 mutants (Figure 5B).

IpaH9.8 also shows activity when incubated with BostonBiochem UBA5 and not UFC1. In this case, Figure 6A (observed also in Figure 6B), the ~46kDa band corresponding most likely to UBA5 appears only in reactions stopped with DTT. However, addition of IpaH9.8 (lane 4 and 5) stabilizes the species' appearance in urea-stopped reactions. Total protein stains show that the intensity of the band in question is decreased but still apparent (Figure 6B, lane 3).

All of this goes to show that puzzling out the pattern of bands from *in vitro* ubiquitination and UFMylation reactions is a complex undertaking. To gain insight to the nature of the species that were created in the *in vitro* reaction mixtures, I turned to mass spectrometry.

3.2.2.1: Mass Spectrometry analysis of In Vitro *Reaction Species*

One of the challenges interpreting *in vitro* UFMylation reactions is the uncertainty involved in assigning a species observed on SDS-PAGE analysis to a specific protein. Despite immunoblotting with controls, unknown species often appear. To aid in this identification, mass spectrometry of excised gel bands from an *in vitro* UFMylation reaction was used to identify their composition. To maximize my findings, I examined 6 different gel bands from a variety of reactions including UFM1, IpaH9.8, and the homemade and BostonBiochem varieties of UBA5, UFC1 (Figure 7).

Samples 1 and 2 were two high molecular weight (100kDa+) species that arose when homemade UFC1, UBA5, UFM1, and IpaH9.8 were combined. Sample 6 was a similar-sized but distinct species that was produced when the BostonBiochem variants of the same enzymes were combined. Samples 3, 4 were lower molecular weight (46-60kDa) species that were produced when the homemade UBA5 and UFC1 were combined with UFM1 but without IpaH9.8, whereas sample 5 was another similar sized but distinct species for a reaction containing the BostonBiochem enzymes.

UFM1 and IpaH9.8 were found in 2 of the 3 high molecular weight slices (Samples 1, 2, and 6) that I previously ascribed to IpaH9.8-UFM1 conjugates. This was true for both homemade UBA5/UFC1 and BostonBiochem UBA5/UFC1, suggesting that both isoforms are active despite the difference in banding patterns observed in previous *in vitro* reactions. Surprisingly, however, UBA5 and UFC1 were found in the high molecular weight species as well, suggesting that this species is more complex than just IpaH9.8 and UFM1. Additionally, IpaH9.8 was shown to be present in bands that

corresponded to molecular weights below 80kDa – the observed mobility for the full-length protein – and in sample 4 that was not designated as containing IpaH9.8.

One of the more interesting findings was the detection of a UBA5 peptide with a VG modification on a non-lysine residue (bands 3, 4, and 5). VG modifications are diagnostic of UFMylation in the same way that GG modifications are diagnostic of ubiquitination in a mass spectrometry analysis. One of the processing steps in mass spectrometry involves the digestion of proteins using trypsin, which digests peptide bonds after a lysine or arginine amino acid, except when they are followed by a proline residue, or if the amino acid is otherwise blocked (i.e. ubiquitinated or ufmylated). Because the last three amino acids of the activated form of UFM1 are RVG (or RGG in the case of ubiquitin), trypsin digestion generates a VG tag attached to the substrate protein peptide. When the mass-to-charge ratio of this peptide is converted into a spectrum, it is searched against a database of theoretical spectra. The weight of the additional VG shifts the spectrum ever-so-slightly so that it scores more highly against a theoretical spectrum where these modifications are present. Unfortunately, the spectrum of a protein where one amino acid has a VG modification is extremely similar to the spectrum of the same protein where the following or preceding amino acid has the VG modification instead. For that reason, mass spectrometry analysis of the UBA5 peptide "Q**EDS**V**T**ELTVEDSGESLEDLMAK" (residues 391 to 414 in the UBA5 protein) inconclusively determined that a VG modification was present on either a glutamic acid (computationally most likely), aspartic acid, serine, or threonine residue (underlined and bolded in the peptide sequence).

3.2.3: Co-Purification of IpaH9.8 and UFM1

Yeast 2-hybrid hits need to be validated by demonstrating that the two proteins also interact in their natural cellular context. Common approaches include co-immunoprecipitation or co-affinity precipitation assays, where one protein is expressed fused to an affinity or epitope tag that allows it to be purified from the cell extract, and the presence of the other interacting protein is determined. In this case, I tackled the problem from both ends using Myc-tagged IpaH9.8 and 6xHis-tagged UFM1.

Lysates from HEK293T cells expressing His-UFM1 and untagged IpaH9.8 or untagged IpaH9.8 alone were subject to cobalt bead His-affinity purification, competitively eluted from the beads with imidazole, and analyzed via SDS-PAGE and Western blotting to detect the presence of IpaH9.8. HEK293T cells expressing both His-tagged UFM1 and IpaH9.8 were compared to a negative control expressing only IpaH9.8. Expression of His-UFM1 increased the recovery of IpaH9.8 as measured by anti-IpaH immunoblot (Figure 8A). However, as measured by UFM1 immunoblot, His-UFM1 did not seem to be enriched despite the recovery of its supposed interaction partner. This indicates high level of background untagged-UFM1 in the negative control. Contrarily, in a separate experiment His-tagged UFM1 interacted with cobalt beads, while endogenous UFM1 did not (Figure 8B).

To clarify this experiment, epitope-tagged Myc-IpaH9.8 was used in the place of His-UFM1. Lysates from HEK293T cells expressing Myc-tagged IpaH9.8 and His-UFM1 or lysates from cells expressing only His-UFM1 were subject to protein A agarose beads coupled to a Myc antibody immunopurification and bound proteins were eluted from the beads by boiling, and analyzed via SDS-PAGE and Western blotting.

Expression of Myc-IpaH9.8 enhanced the recovery of low molecular weight species consistent with UFM1 as demonstrated with anti-UFM1 immunoblotting (Figure 8C). Additional species were shown in the UFM1, IpaH9.8 and Myc (data not shown) western blots that represented, most likely, detected Myc antibody, breakdown products of the expressed proteins, or cross-reacting proteins.

3.2.4: Mass Spectrometry Screen for UFMylation Substrates of IpaH9.8

After this characterization of the UFM1-IpaH9.8 interaction, I began a search to determine if IpaH9.8 expression changed the suite of UFMylated proteins in a host cell. A schema for the identification of IpaH9.8-directed UFMylation targets via co-affinity purification and tandem-mass spectrometry is described in Figure 9. Briefly, His-tagged proteins were purified as before, from lysates of HEK293T cells co-expressing His-UFM1 and IpaH9.8, or as a control, just His-UFM1. Eluted fractions were digested with trypsin and analyzed via liquid chromatography tandem mass spectrometry (LC-MS/MS). The aim of this experiment was to differentiate between proteins that are UFMylated under steady-state cellular conditions and proteins that are specifically UFMylated when IpaH9.8 is present.

The resulting list of proteins was largely similar between both samples. Therefore, all proteins that showed up in both lists were removed from the UFM1+IpaH9.8 sample so that it included only proteins that were unique to IpaH9.8-induced UFMylation. Both lists were then ordered by the number of unique peptides captured by LC-MS/MS. In the IpaH9.8-expressing sample, the unique peptide count cut-off was 3, which included 70 different proteins. For the control UFM1-only sample, the unique peptide count cut-off was 10 (higher because it included all of the most abundant proteins present in both

samples), which included 72 different proteins. The total list of proteins is described in Table 6. Additionally, a STRING protein-protein interaction map was constructed for each set of proteins to demonstrate the known interactions between these groups of proteins (Szklarczyk et al., 2017) (Figure 10, Figure 11).

Additional information from mass spectrometry came in the form of VG modifications. As before, VG modifications indicated proteins that were mostly likely UFMylated. However, a complication arose because a VG mass shift is nearly identical to the mass shift of a miscleaved peptide. Similar to a polymerase that stalls on highly repetitive nucleobases, trypsin also miscleaves a peptide when it is exposed to consecutive lysines or arginines. If a peptide starts with multiple arginines, the N-terminal end of that peptide may be XXX or RXXX or RRXXX depending on the efficiency of the trypsin cleavage. A single miscleaved arginine at the beginning or end of a peptide is nearly the same mass shift as that same peptide with an additional VG modification. For that reason, in Table 6, potential false-positive VG-modifications are denoted by an underline.

Of the total list of proteins, some of which will be examined in the discussion, galectin-7 was chosen for further characterization as a putative IpaH9.8-induced target of ufmylation. Galectins are sugar-binding proteins involved in recognizing intracellular pathogen escape from vacuoles and have previously been implicated in defense against *Shigella*. For that reason, I chose Galectin-7 as a potential target to further investigate.

Table 6: Mass spectrometry hits from co-affinity purification of His-UFM1 and IpaH9.8.

On the left are the hits from the UFM1-expressing sample. On the right are hits from the UFM1+IpaH9.8-expressing sample. Discovered VG modifications are denoted by 'VG' and underlined if they represent a potential false positive.

UFM1 Sample			UFM1 + 9.8 Sample		
Unique Peptides	Genes Name	VG Tag	Unique Peptides	Gene Name	VG Tag
80	FLNA		11	DDX42	
64	KRT1		10	RPS19	
62	KRT2	VG	9	KRT18	
54	KRT9		9	RPL5	
51	KRT10	VG	8	RPS5	
50	NONO	VG	8	DSG1	
45	KRT5	VG	8	EEF2	
44	KRT6A	VG	7	HINT1	
43	DHX15		**7**	**LGALS7**	
38	KRT14		7	S100A7	
37	SFPQ	VG	6	ATXN2	
35	RPL4		6	RPS7	VG
32	KRT16		6	ASS1	
32	VIM	VG	6	CHD1	
30	KRT17		6	PGK1	
28	ALB		6	RPL37	
27	ACTB		5	KRT13	
26	NUFIP2		5	YWHAZ	
25	DSP	VG	5	RPS20	
23	RPL3		5	BASP1	
23	ENO1	VG	5	SEC31A	
23	HRNR		5	FLG	
23	NCL		5	TPI1	
23	TKT		5	ALDOA	
22	HSPA8	VG	5	BCLAF1	
22	KRT8	VG	5	CAPRIN1	
22	SRRM2		4	KRT78	
21	ATXN2		4	ARGLU1	

UFM1 Sample			UFM1 + 9.8 Sample		
Unique Peptides	Genes Name	VG Tag	Unique Peptides	Gene Name	VG Tag
20	RPL8		4	DAP	
20	NPM1		4	UFM1	
20	PKM		4	FUBP1	
19	RPL7A		4	RPS3	
19	CCNT1	VG	4	RPS16	
18	HNRNPL	VG	4	MEPCE	
17	RPL13		4	CKB	
17	KRT76		4	HNRNPK	
17	PLRG1		4	RAN	
16	RP39		4	UBAP2	
16	GAPDH		4	ATN1	
16	SF1		4	SERPINB3	
15	RPS18		4	ACTN4	
15	RPL13A		4	PDCD5	
15	HSP90AB1		3	HSP90AA1	
15	PRDX1	VG	3	YWHAG	
15	SF3A2		3	RPS27	
15	TUBB		3	B2M	
14	RPL28		3	S100A9	
14	HNRNPU		3	BOLA2	
13	RPS6		3	NACA	
13	RPS27A		3	SMAP	
12	RPS13		3	CRIP1	
12	RPS4X		3	GSTP1	
12	RPL10	VG	3	TAGLN2	
12	RPL31		3	CBLL1	
12	EEF1A1		3	FASN	
11	RPS11		3	KPRP	
11	RPL10A		3	SNRPD3	
11	RPL24		3	CSN1S1	
11	RPL27		3	SUB1	
11	RPL6		3	YTHDC1	
11	HSPA5	VG	3	CNBP	
11	YBX1		3	WDR12	
10	RPS8		3	CASP14	
10	RPL14		3	WAC	
10	RPL15		3	PTGES3	

UFM1 Sample			UFM1 + 9.8 Sample		
Unique Peptides	Genes Name	VG Tag	Unique Peptides	Gene Name	VG Tag
10	RPL18A		3	NMT1	
10	RPL26		3	DYNLL1	
10	RPL7		3	DDX46	
10	ACTBL2		3	WDR1	
10	CSRP1		3	CSTA	
10	JUP				
10	PPIA				

3.2.5: Galectin-7 as a IpaH9.8-Directed UFMylation Target

To determine Galectin-7 levels in response to overexpression of IpaH9.8 and UFM1, a galectin-7 polyclonal antibody was purchased from R&DSystems (Table 2). HEK293T cells were transfected with a combination of plasmids that would express UFM1, IpaH9.8, or IpaH9.8 C337A mutant, or empty vector control, lysed, and examined via Western blotting for galectin-7 levels. Expression of a species corresponding to the approximate size of galectin-7 (15kDa) was decreased in the absence of UFM1 and IpaH9.8, suggesting that these two proteins help to stabilize galectin-7 (Figure 12A). Furthermore, this stabilization appeared to be dependent on the expression of both proteins together, and this phenotype was not recapitulated with the C337A IpaH9.8 mutant (Figure 12B).

To confirm this observation, purified galectin-7 was purchased from Cedarlane labs and used as a substrate for an in-vitro UFMylation assay with purified UFM1, UBA5, UFC1, and IpaH9.8. High molecular weight smears of galectin-containing material were detected by western blotting, consistent with galectin-7 UFMylation (Figure 13). These smears were dependent on the addition of IpaH9.8, E1, and E2. Additional discrete species at ~30kDa, ~45Da, and 100+kDa were also dependent on E1,

E2, and IpaH9.8 and may represent polyUFMylated galectin-7 or aggregates of Galectin-7 and the UFMylation machinery (UBA5, UFC1, UFM1).

3.2.6: Fluorescent Microscopy

As an additional tool for future inquiry, I created fluorescent fusions of mClover3 and mRuby3 to the amino terminal of UFM1. As a proof of concept and to investigate any broad trends in localization of UFM1 in response to IpaH9.8, I transfected cells with this fluorescent construct in the presence of IpaH9.8 or a negative control empty vector, and also examined IpaH9.8 localization via immunofluorescence of a Myc-tagged IpaH9.8 construct. While no striking changes were observed in UFM1 localization in response to IpaH9.8, UFM1 signal was diffuse and cytoplasmic, whereas nuclear strong nuclear staining was observed (Figure 14A, B). UFM1 signal may appear more diffuse within a cell when IpaH9.8 is introduced, though these results are not sufficient to confirm this finding. I also observed that IpaH9.8 localization is similarly diffused throughout the cytoplasm, but with more intense nuclear staining.

3.3: ZKSCAN3

To supplement my findings with UFM1, I also chose to conduct a preliminary investigation on another yeast 2-hybrid hit, ZKSCAN3, and its interaction with IpaH9.8. Previous Rohde lab member Kaitlyn Tanner demonstrated that ZKSCAN3 induced autophagy *in vitro*, which was dependent on IpaH9.8. I began my investigation by confirming that IpaH9.8 ubiquitinates ZKSCAN3 *in vitro*.

3.3.1: *In Vitro* Ubiquitination Assays

To confirm ZKSCAN3 as a potential substrate for IpaH9.8 I used a similar approach to that which I took with UFM1 using purified protein from BostonBiochem

and GST-IpaH9.8 (see section 3.1). Upon addition of E1, E2, HA-Ubiquitin, and WT GST-IpaH9.8 to ZKSCAN3, additional 61kDa+ high molecular weight species consistent with ubiquitinated ZKSCAN3 (61kDa) were detected by HA immunoblotting (Figure 15). These species were not observed when ZKSCAN3 was incubated with the C337A IpaH9.8 mutant, indicating they were dependent on IpaH9.8 catalytic activity. The appearance of this species also corresponded with the disappearance of a smaller species consistent with charged E2-Ub. While these species were not apparent without the E1 or E2 or IpaH9.8, there were additional high molecular weight species that appeared without wild type IpaH9.8. Analysis via ZKSCAN3 antibody did not yield any discernible differences between the same conditions.

3.3.2: Visualization of Lysosomes Using LysoTracker

ZKSCAN3 is a negative transcriptional regulator of lysosomal biogenesis. To determine if IpaH9.8 might be affecting lysosomal biogenesis through modulation of ZKSCAN3 I used a dye called LysoTracker. LysoTracker is a fluorescent probe attached to a weakly basic amine molecule and has been shown to accumulate in acidic organelles through an as-yet-unidentified mechanism (Chazotte, 2011). As such, it is a powerful tool for visualizing lysosomes.

The drug Torin was used as a positive control, as it inhibits the phosphorylation of mTORC1 substrates including TFEB – leading to lysosomal biogenesis (Thoreen et al., 2009). 4E-BP is a well-characterized mTORC1 substrate, and collapses from several phosphorylated species to a single species upon Torin treatment (Thoreen et al., 2009). I detected this substrate via immunoblotting to verify the validity of Torin treatment (Figure 16). An empty vector control was used to illustrate steady-state lysosome levels

in HeLa cells. Additionally, plasmids that would express IpaH9.8 and C337A mutant IpaH9.8 were transfected into HeLa cells and compared to both controls. The addition of WT IpaH9.8 induced levels of lysosomes similar to Torin-treated cells (Figure 16). C337A IpaH9.8 The lysosomal levels in cells expressing an empty vector control were also similar to cells expressing C337A IpaH9.8 (Figure 16). This demonstrates that IpaH9.8 increases lysosomal biogenesis as has previously been observed for Torin treatment, which is consistent with a model where IpaH9.8 interferes with ZKSCAN3 function.

3.3.3: Electron Microscopy

To corroborate the idea that IpaH9.8 increases lysosome number, I used electron microscopy to examine lysosomal size and number directly. The same four conditions were used as those in the LysoTracker assay to see if over-expression of IpaH9.8 recapitulated Torin-induced lysosomal changes. While lysosomes are described as electron-dense roughly-circular shapes, they are difficult to distinguish from mitochondria when their cristae are not entirely visible. For this reason, I was unable to conclusively determine if lysosome levels increased with both Torin treatment or IpaH9.8 overexpression without a direct lysosome stain. Nonetheless, a larger number of smaller, electron-dense organelles were present in IpaH9.8-expressing and Torin-treated cells compared to an empty vector control (Figure 17).

3.3.4: Fluorescent Microscopy

I constructed a fusion protein between the fluorescent proteins mClover3 and mRuby3 and the N-terminus of ZKSCAN3. Cells were transfected with this construct and a mammalian expression vector for IpaH9.8 or an empty vector control. Using

fluorescent microscopy, the effect of IpaH9.8 on ZKSCAN3 localization was observed to be more striking than the effect of IpaH9.8 on UFM1. As seen in Figure 18, expression of IpaH9.8 induced perinuclear ZKSCAN3 foci formation. These foci were more intense and less diffuse than ZKSCAN3 distribution in cells without IpaH9.8.

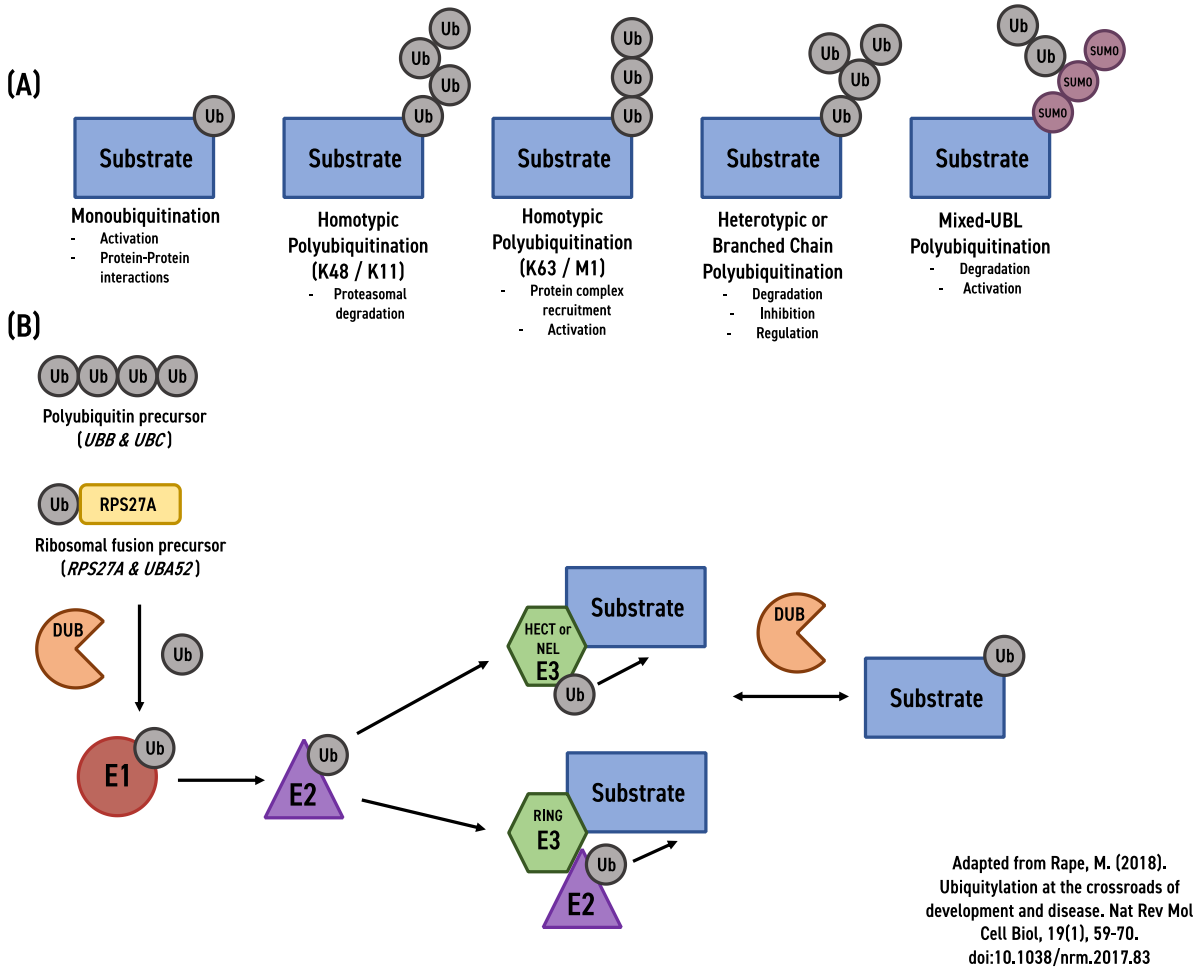

Figure 1: The ubiquitin conjugation system and the consequences thereof.
(A) Ubiquitin (ub) can be attached to substrates as a monomer or a polyubiquitin chain. Polyubiquitin chains can arise through attachments to one of seven internal ubiquitin lysine residues, or the first methionine residue (M1) of ubiquitin, and include mixed ubiquitin-like (UBL) chains such as polySUMO-ubiquitin chains. Each chain type dictates various cellular fates (B) Ubiquitin is produced as a polypeptide or ribosomal protein fusion and must first be processed by deubiquitinase (DUB) enzymes before conjugation. Conjugation proceeds via a E1-E2-E3 thioester cascade where HECT or NEL E3s accept ubiquitin directly versus RING E3s that act as a scaffold to bring the charged E2 and substrate together. Ubiquitination of substrates can be reversed through DUBs. Figure adapted from (Rape, 2018).

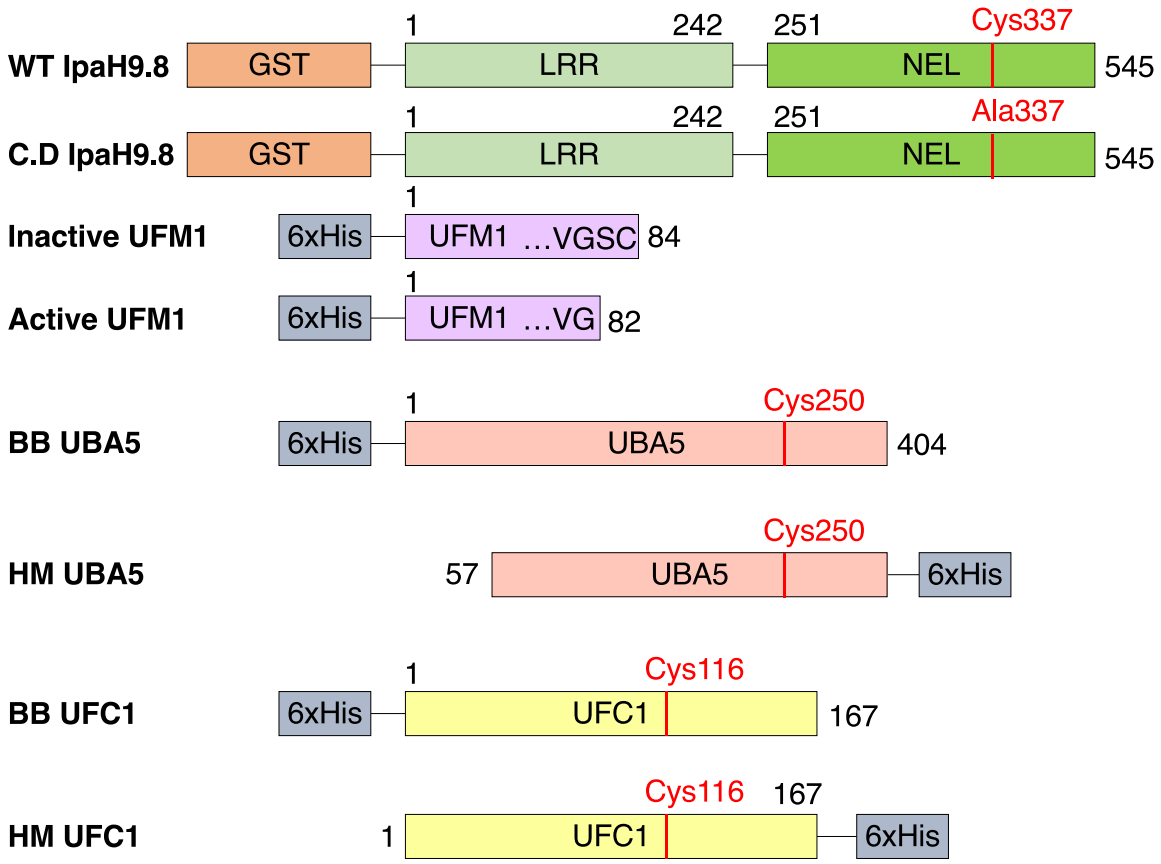

Figure 2: Schematics of purified proteins used for *in vitro* ubiquitination and UFMylation.
Schematic demonstrating the leucine rich repeat (LRR) and Novel E3 Ligase (NEL) domain of wild type (WT) IpaH9.8, or Catalytically-Dead (C.D) IpaH9.8. Also shown are is UFM1 (the active form was used in all instances in this thesis), and BostonBiochem (BB), or Homemade (HM) variants of UBA5 and UFC1. Numbers denote the position of amino acids, and highlighted residues denote the catalytic cysteine residue of the enzyme.

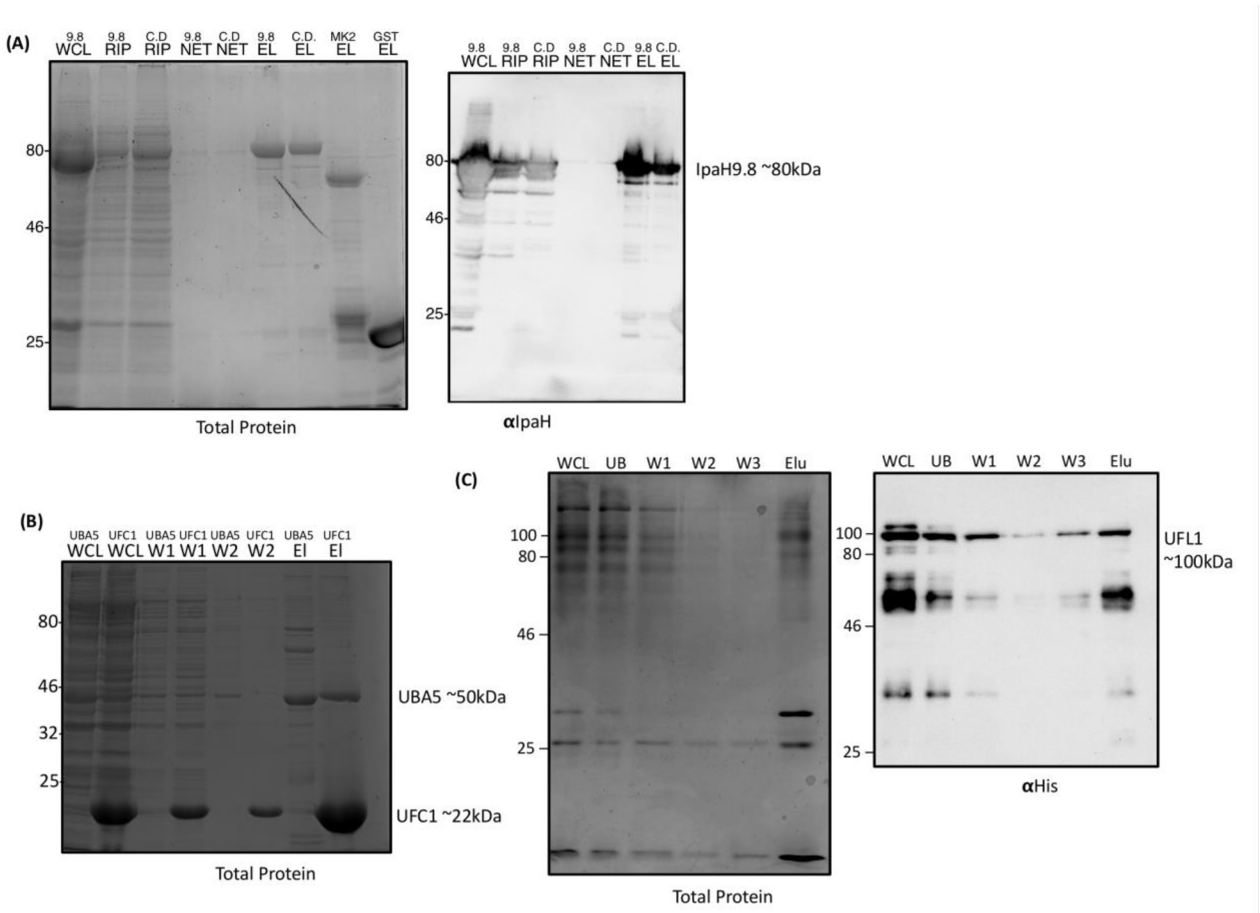

Figure 3: Purification of bacterially-expressed IpaH9.8, UBA5, UFC1, and UFL1 proteins.

(A) *E. coli* expressing GST-tagged IpaH9.8 and C337A IpaH9.8 (C.D) were lysed (WCL) and passed through a column containing glutathione agarose beads. Columns were washed with two buffers (RIP, NET), and eluted (EL) with reduced glutathione. Fractions were analyzed via SDS-PAGE. On the left is a total protein stain of these fractions, with GST-tagged MK2 and free GST as a control. On the right is an anti-IpaH immunoblot on these same fractions. (B) His-tagged UBA5 and UFC1 were purified using cobalt agarose beads. Whole cell lysate (WCL), washes (W1, W2), and eluted (El) fractions were analyzed via SDS-PAGE and subject to a total protein stain. (C) His-tagged UFL1 was purified using cobalt agarose beads. Whole cell lysate (WCL), unbound lysate (UB), washes (W1, W2, W3) and eluted (Elu) fractions were analyzed via SDS-PAGE and subjected to a total protein stain (left) and anti-PentaHis immunoblotting (right). Mobility of MW standards is shown at left in kDa.

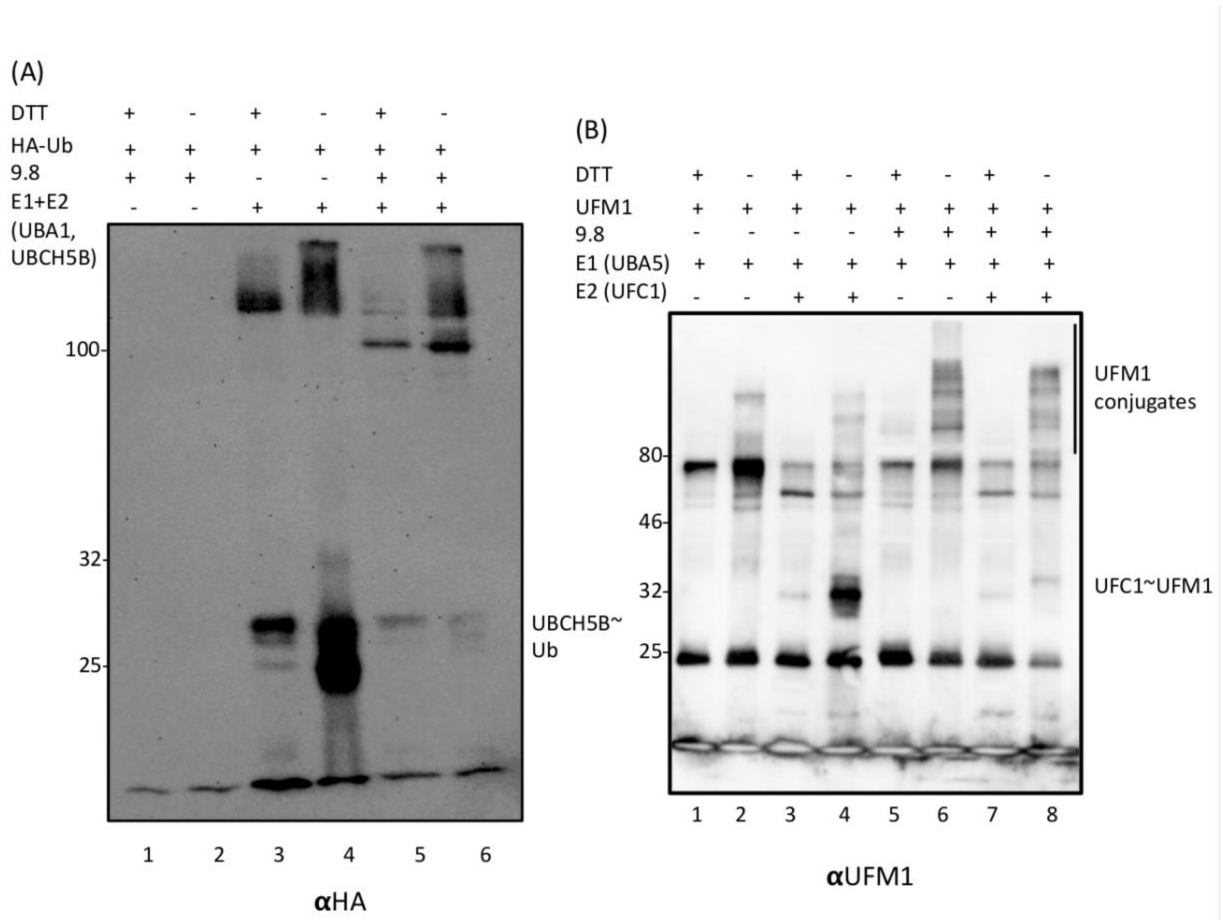

Figure 4: IpaH9.8 incubated with the ubiquitination and UFMylation machinery leads to changes in DTT-sensitive high molecular weight conjugates.
(A) UBA1 (E1), UbcH5b (E2), and full length wild type GST-IpaH9.8 (9.8) were incubated with HA-ubiquitin (Ub) and ATP for 1 hour, then stopped with a buffer containing DTT or urea. Reactions were analyzed via SDS-PAGE and subsequent anti-HA immunoblotting. Addition of IpaH9.8 resuled in the ablation of a species that corresponds to charged E2 (UbcH5b-Ub) and changes in DTT-sensitive high molecular weight ubiquitin conjugates, including a new DTT-insensitive species. (B) His-UBA5 (E1), His-UFC1 (E2), and full length, wild type GST-IpaH9.8 (FL-WT) were incubated with His-UFM1, and ATP for 1 hour, then stopped with a buffer containing DTT or urea. Reactions were analyzed via SDS-PAGE and subsequent anti-UFM1 immunoblotting. Addition of IpaH9.8 results in the ablation of a species that corresponds to charged E2 (UFC1-UFM1) and the appearance of DTT-sensitive high molecular weight UFM1 conjugates. Representative blots from several independent experiments are shown. Mobility of MW standards shown at left are in kDa.

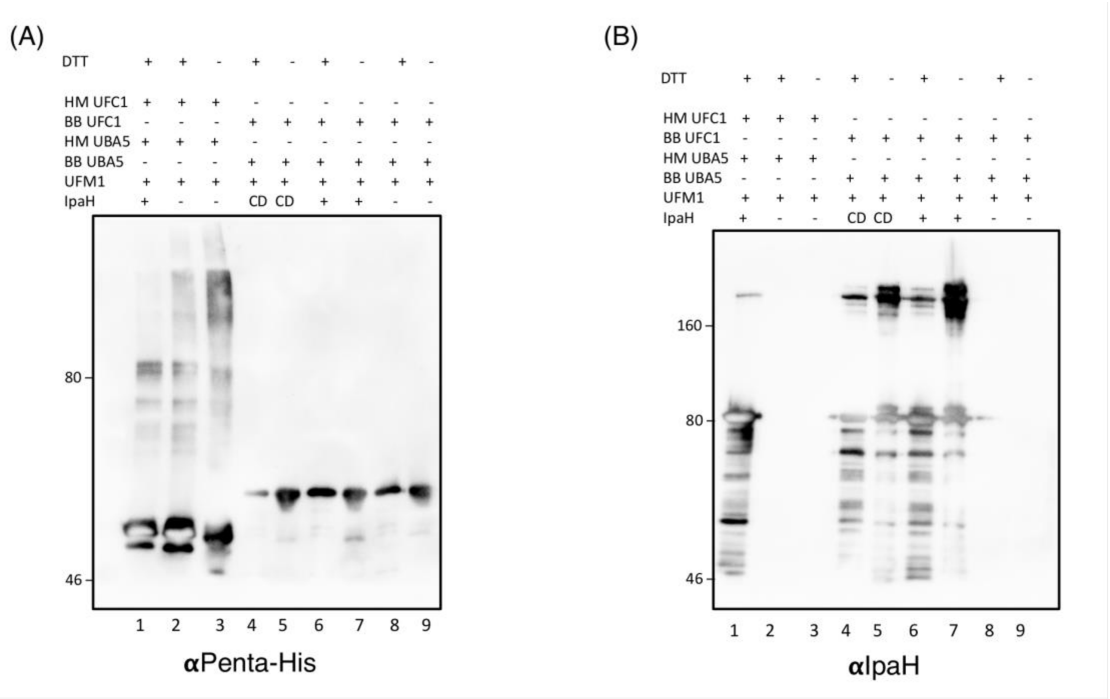

Figure 5: Homemade and BostonBiochem UFC1 and UBA5 react differently with IpaH9.8 and UFM1 *in vitro*.

Purified His-UFC1 and His-UBA5 (homemade HM or BostonBiochem BB) were combined with His-UFM1, and IpaH9.8 for 1 hour and reactions were stopped with a buffer containing either DTT or Urea. Reactions were analyzed via SDS-PAGE and subject to anti-Penta-His or anti-IpaH immunoblot. (A) Incubation of HM enzymes but not BB enzymes with IpaH and UFM1 results in high molecular weight His-containing smearing (B) Incubation of BB enzymes with IpaH and UFM1 results in multiple high molecular weight IpaH-containing species compared to HM enzymes. Representative blots from several independent experiments are shown. Mobility of MW standards shown at left are in kDa.

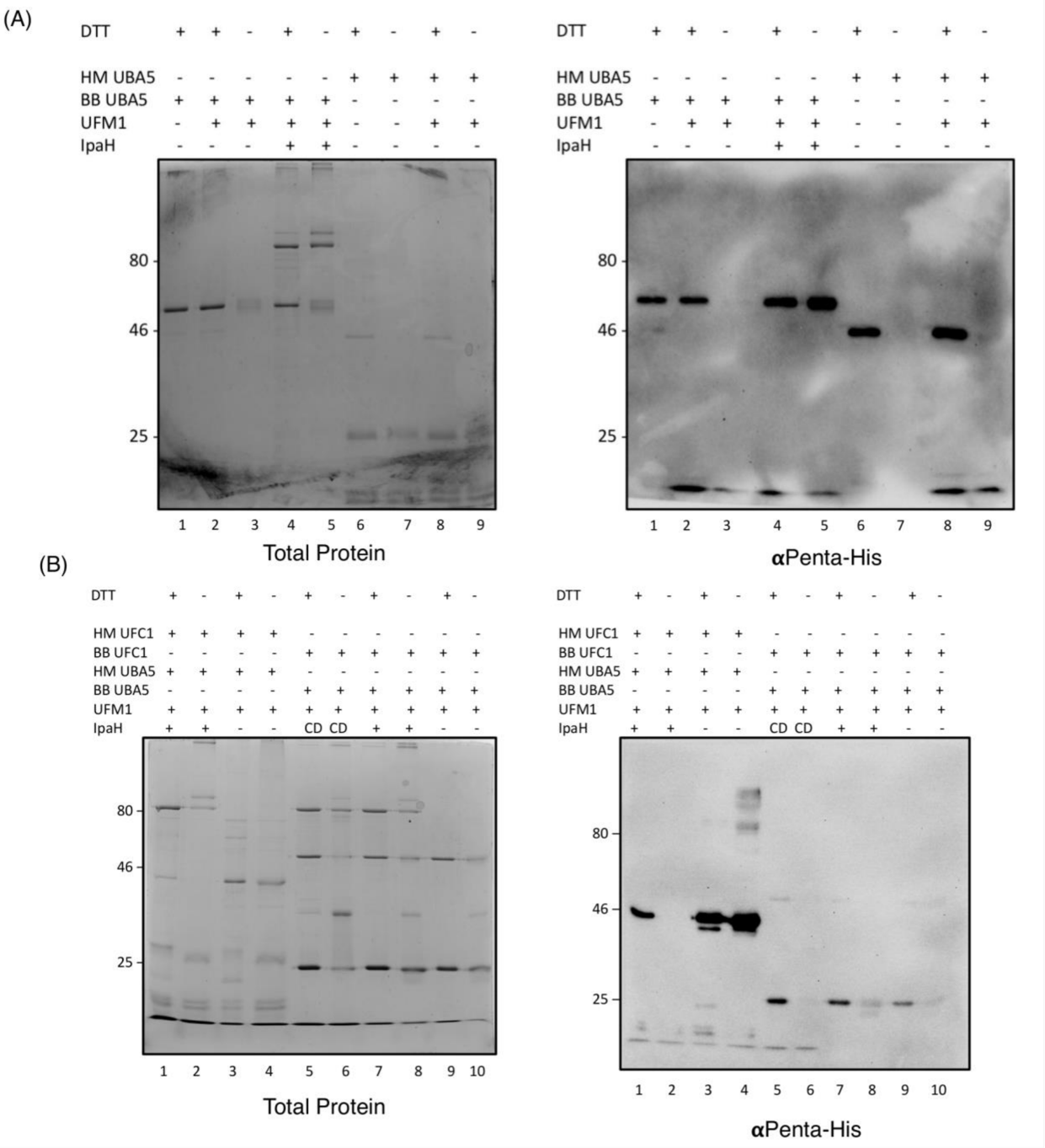

Figure 6: Further differences between homemade and BostonBiochem UFC1 and UBA5.

Purified His-tagged UFC1 and UBA5 (homemade HM or BostonBiochem BB) were combined with His-UFM1, and GST-IpaH9.8 for 1 hour and reactions were stopped with a buffer containing either DTT or Urea. Reactions were analyzed via SDS-PAGE and the same reactions were subjected to either total protein stain (left) or anti-penta-his immunoblotting (right). Representative blots from several independent experiments are shown. Mobility of MW standards shown at left are in kDa.

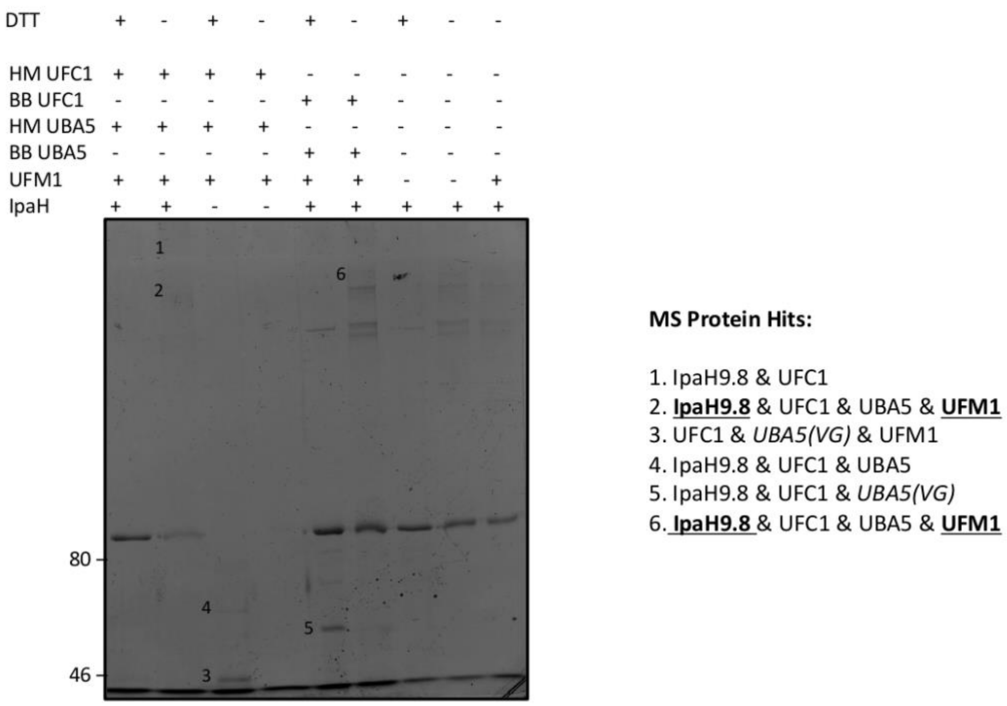

Figure 7: Mass Spectrometry identifies high molecular weight conjugates between IpaH9.8 and UFM1 *in vitro*.
Purified UFC1 and UBA5 (homemade HM or BostonBiochem BB) were combined with UFM1, and IpaH9.8 for 1 hour and reactions were stopped with a buffer containing either DTT or Urea. Reactions were analyzed via SDS-PAGE and subjected to total protein stain. Six gel slices were excised as denoted by the numbers 1-6, and analyzed via tandem liquid chromatography mass spectrometry. The proteins that appeared in each gel slice are listed next to the figure. In bold and underlined are the high molecular weight IpaH9.8 – UFM1 conjugates. Additionally, a VG modification indicative of UFMylation was found on UBA5 in samples 3, 4, and 5. Mobility of MW standards shown at left are in kDa.

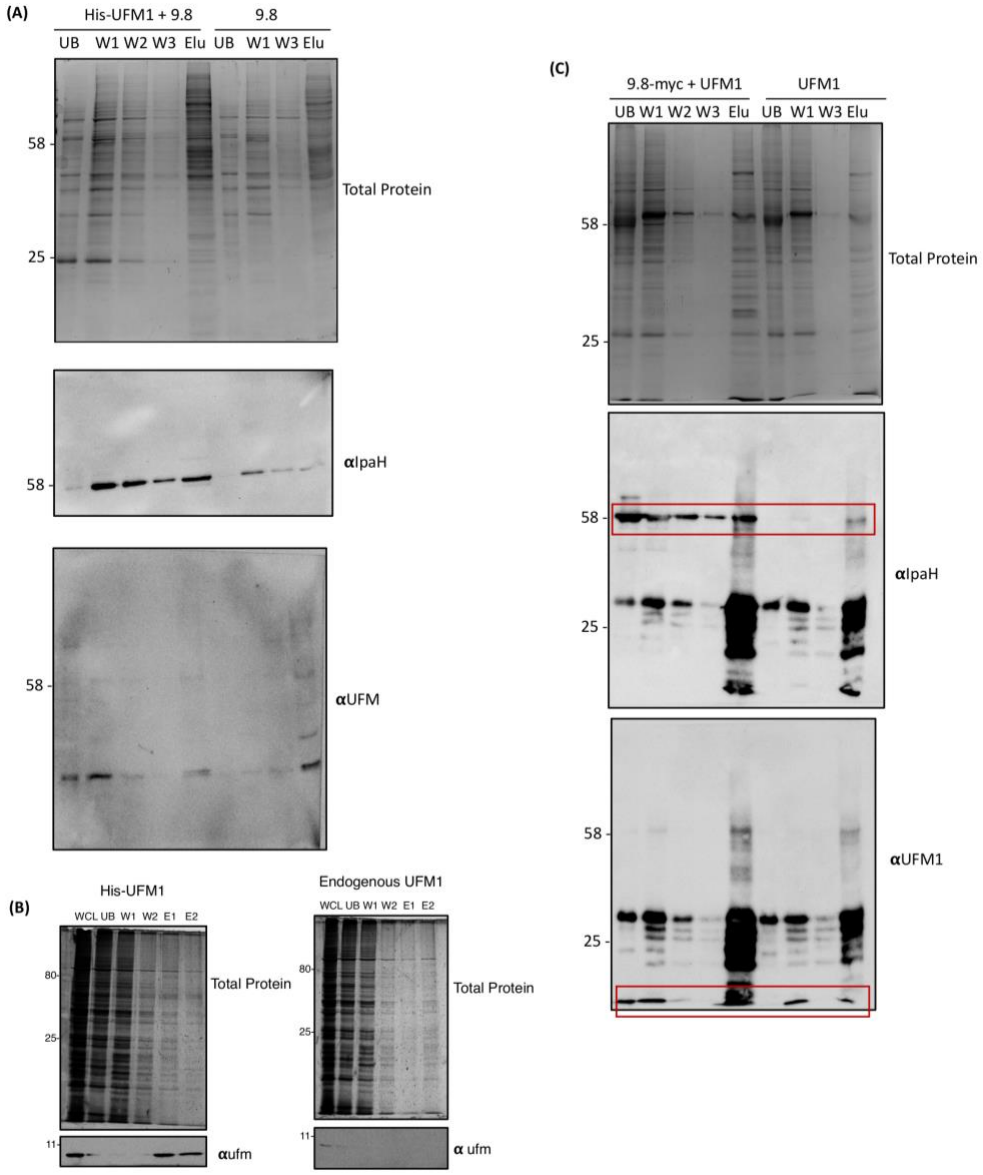

Figure 8: Immunoprecipitation of Myc-IpaH9.8 or His-UFM1 affinity-purification enhances recovery of the other protein.
HEK293T cells were transfected with plasmids that express Myc-IpaH9.8 and/or His-UFM1 and harvested. Unbound (UB), wash (W1, W2, W3), and eluted (Elu) fractions were analyzed via SDS-PAGE followed by total protein stain or immunoblot. (A) His-UFM1 was purified using cobalt bead affinity chromatography from cells co-expressing His-UFM1 and IpaH9.8 and, as a control, cells expressing only IpaH9.8. (B) Affinity purification of cells expressing either His-UFM1 (left) or endogenous UFM1 (right) shows that only His-UFM1 binds efficiently to cobalt beads. (C) His-UFM1 was co-purified with Myc-IpaH9.8 using protein A-myc antibody agarose beads, versus a His-UFM1-only control. Expected position of UFM1 or IpaH9.8 are indicated by red boxes. Mobility of MW standards shown at left are in kDa.

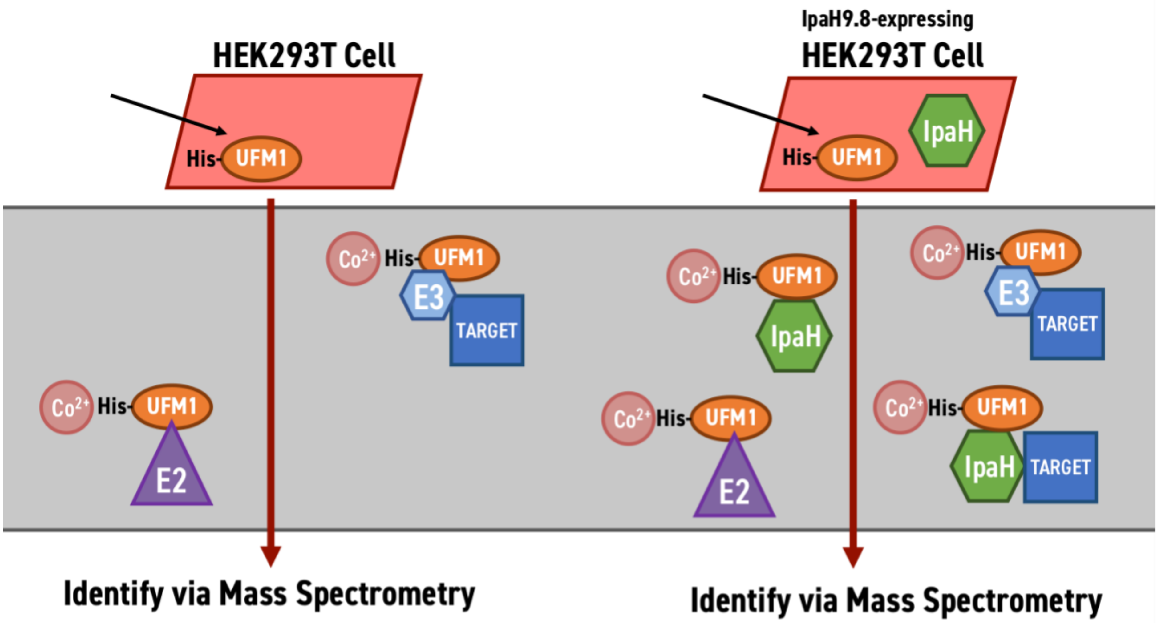

Figure 9: Schema for Co-immunoprecipitation mass spectrometry screen for IpaH9.8-induced UFMylation targets.
HEK293T cells expressing IpaH9.8 are transfected with His-tagged UFM1, lysed and passed through a column on cobalt beads. These lysates are compared to HEK293T cells expressing only His-tagged UFM1. Both columns are washed and the proteins stuck to the cobalt beads are competitively eluted with imidazole. The resulting mix of proteins are compared to determine the suited of UFMylated proteins in human cells with and without the bacterial protein IpaH9.8.

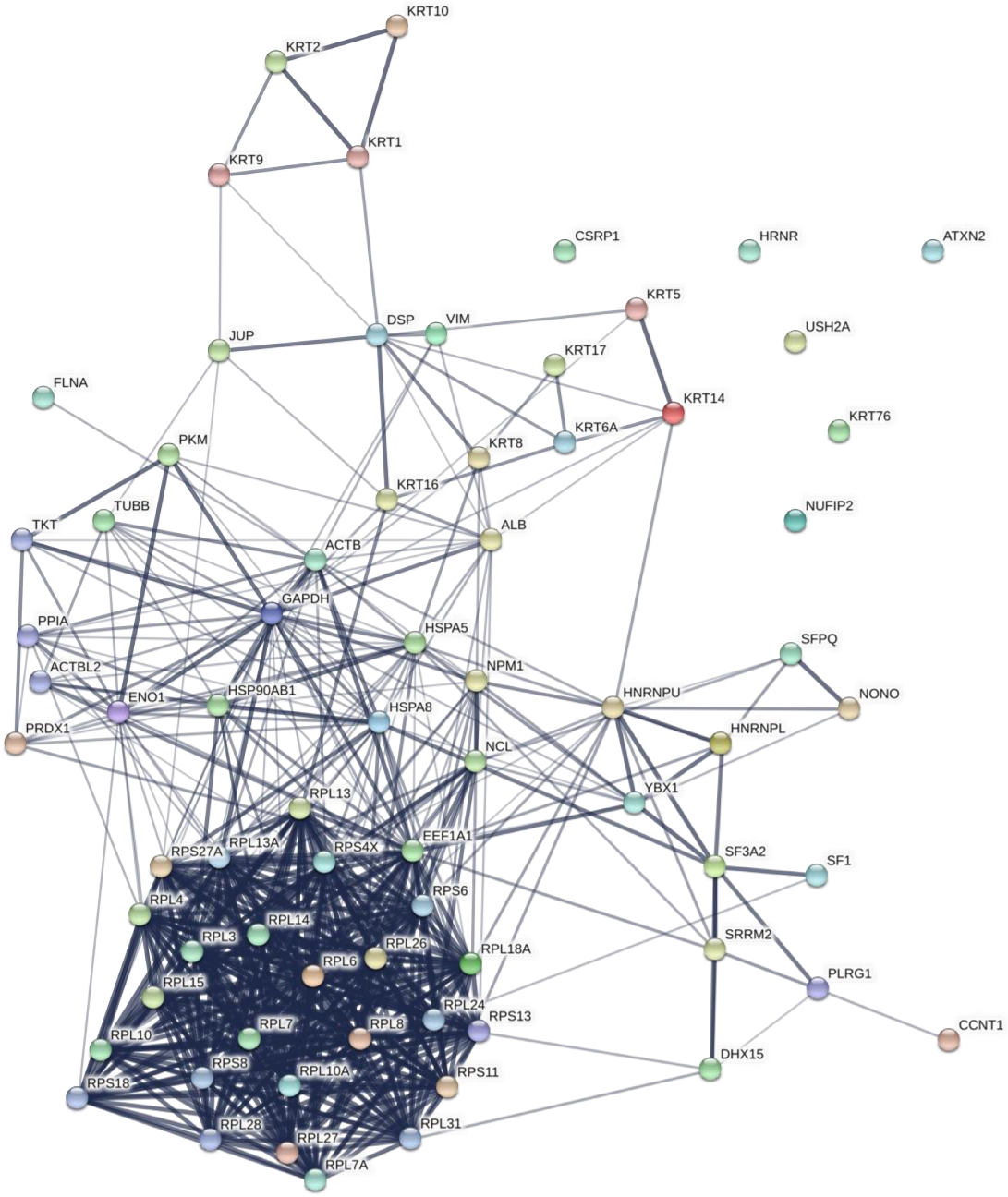

Figure 10: STRING functional protein-protein association map for the top 70 targets of UFMylation as determined by a coupled co-immunoprecipitation mass spectrometry screen.
HEK293T cells expressing His-UFM1 were subject to co-immunoprecipitation followed by mass spectrometry to determine targets of UFMylation. The resulting list of proteins was ordered by number of unique peptides and cut off at 10. STRING map was created using the STRING API at string-db.org. Solid lines represent known relationships between proteins with line thickness indicating the confidence of empirical evidence for this interaction.

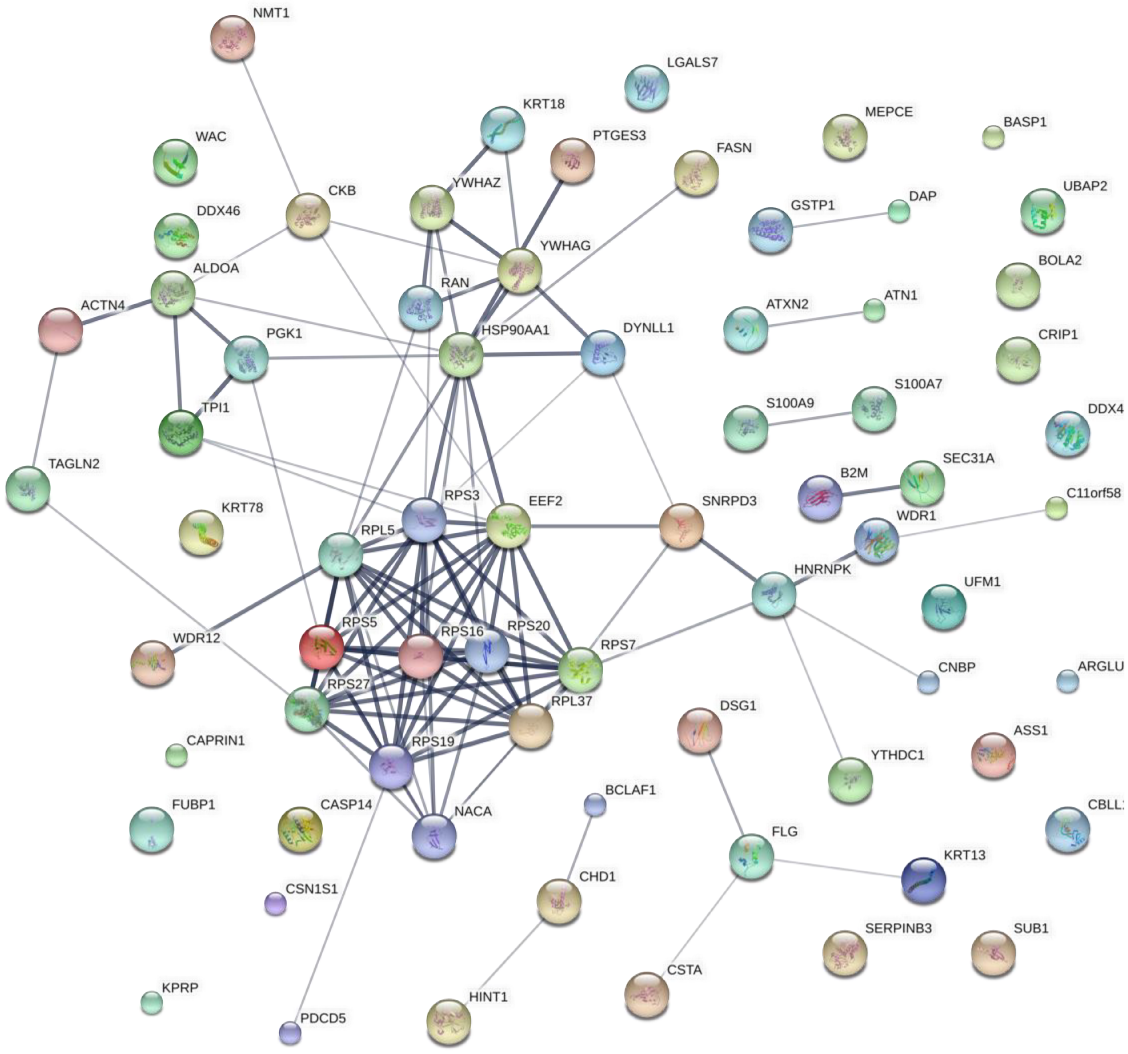

Figure 11: STRING functional protein-protein association map for the top 70 targets of IpaH9.8-induced UFMylation as determined by a coupled co-immunoprecipitation mass spectrometry screen.
HEK293T cells expressing His-UFM1 and IpaH9.8 were subject to co-immunoprecipitation followed by mass spectrometry to determine targets of IpaH9.8-induced UFMylation. The resulting list of proteins was narrowed to proteins only present in this sample and ordered by number of unique peptides and cut off at 3. STRING map was created using the STRING API at string-db.org. Solid lines represent known relationships between proteins with line thickness indicating the confidence of empirical evidence for this interaction.

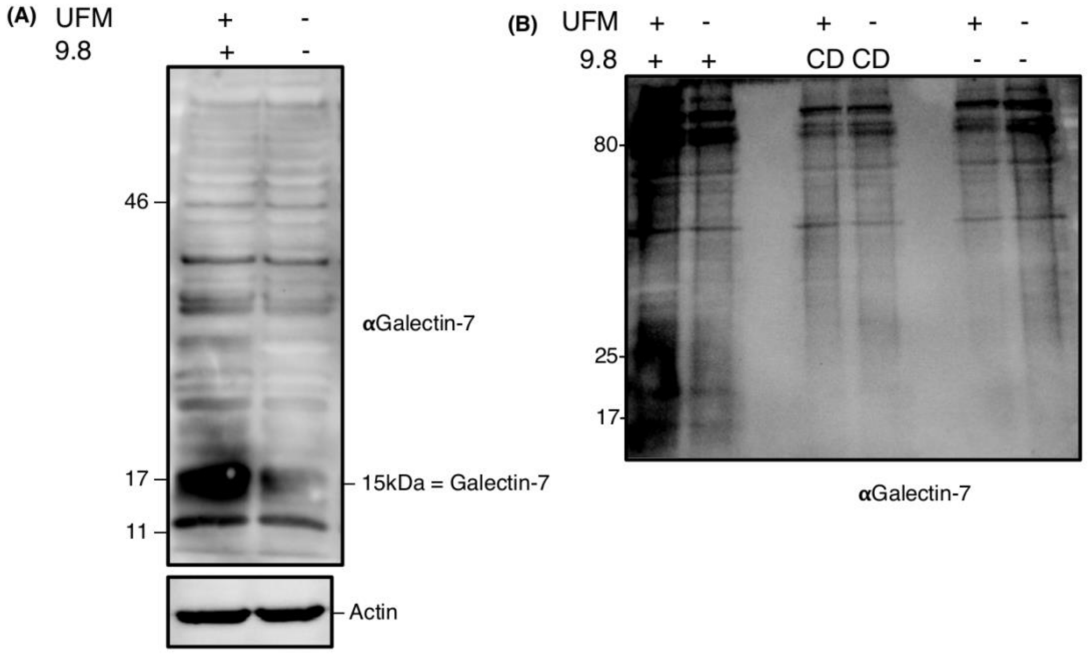

Figure 12: Expression of UFM1 and IpaH9.8 stabilizes Galectin-7 in HEK293T cells. (A) HEK293T cells expressing IpaH9.8 and transfected with UFM1 were compared to HEK293T cells transfected with an empty vector control. Lysates were analyzed via SDS-PAGE and subsequent anti-Galectin-7 immunoblotting, alongside an actin loading control. (B) HEK293T cells expressing a combination of IpaH9.8, C337A IpaH9.8 (CD), UFM1, or empty vector were lysed and analyzed via SDS-PAGE and anti-Galectin-7 immunoblotting. Representative blots of several independent experiments. Mobility of MW standards shown at left are in kDa.

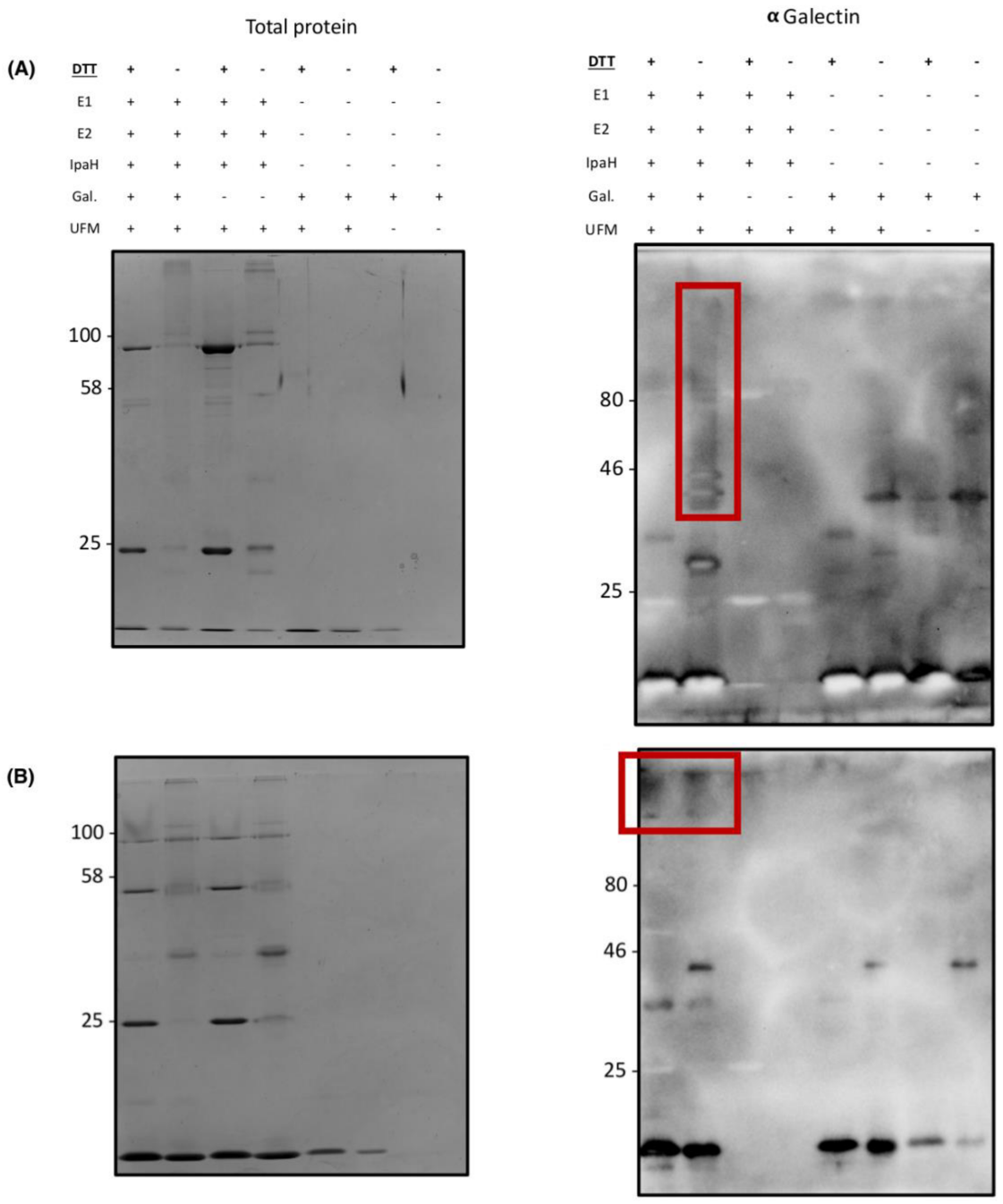

Figure 13: Incubation of Galectin-7 with of IpaH9.8 and UFMylation machinery *in vitro* results in high molecular weight Galectin conjugates.
(A) Purified Galectin-7 was incubated with UBA5 (E1), UFC1 (E2), IpaH9.8, UFM1, and ATP for 1 hour and the reactions stopped with a buffer containing DTT or urea. Samples were analyzed via SDS-PAGE and total protein stain (left) or anti-galectin-7 immunoblotting (right). (B) An independent repeat of the experiment using less Galectin-7 demonstrating a different high molecular weight banding pattern. High molecular weight conjugates highlighted in red boxes. Representative blots of several independent experiments are shown. Mobility of MW standards shown at left are in kDa.

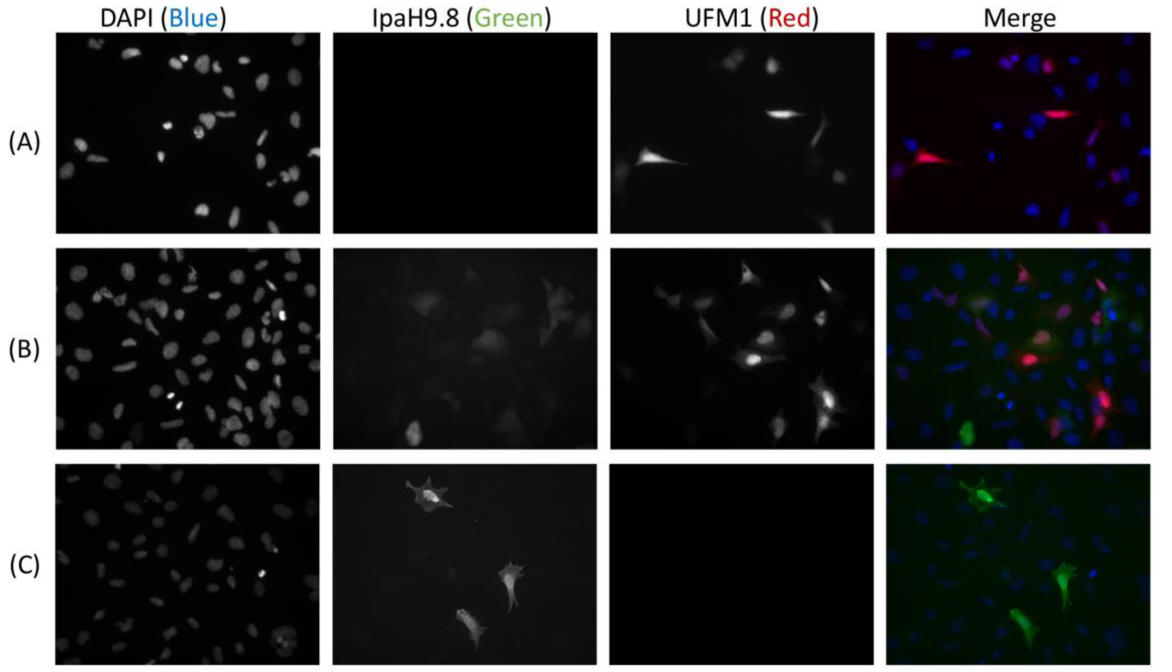

Figure 14: IpaH9.8 expression has no discernable effect on UFM1 levels or distribution as measured by fluorescent microscopy.
HEK293A cells were transfected with mRuby-tagged UFM1 (red) and/or myc-tagged IpaH9.8 detected by αMyc immunofluorescence (green), and a DAPI nuclear stain (blue). (A) UFM1 localization in cells that do not express IpaH9.8 (B) UFM1 localization in cells that express IpaH9.8 (C) IpaH9.8 localization (contains endogenous untagged UFM1). Representative fields of several independent experiments.

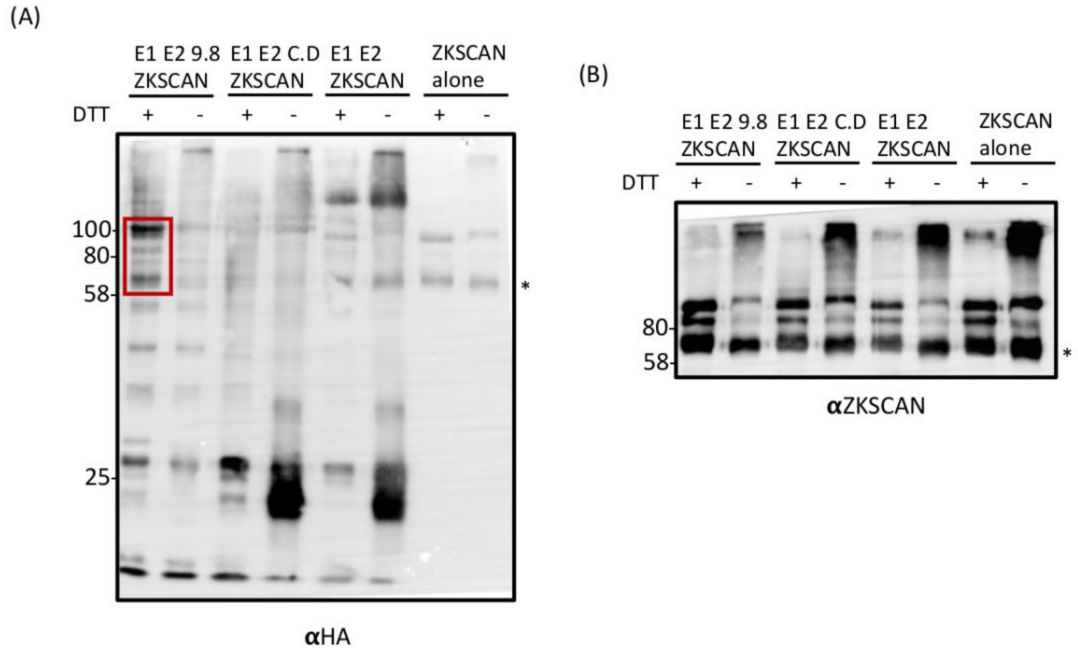

Figure 15: *In vitro* incubation of IpaH9.8 with ZKSCAN3 results in high molecular weight species consistent with ZKSCAN3 ubiquitination.
Purified ZKSCAN3, UBA1 (E1), UBCH5B (E2), IpaH9.8, and HA-Ubiquitin were incubated with ATP for 1 hour and stopped with a buffer containing DTT or urea. Reactions were analyzed via SDS-PAGE and subsequent (A) anti-HA immunoblotting or (B) anti-ZKSCAN immunoblotting. Addition of IpaH9.8 to ZKSCAN3 and the requisite ubiquitination machinery results in the ablation of a charged E2 species (~25kDa), and the emergence of additional HA-containing high molecular weight (~90kDa) species indicative of ubiquitination (indicated by red box). * indicates unmodified ZKSCAN3. Mobility of MW standards shown at left are in kDa.

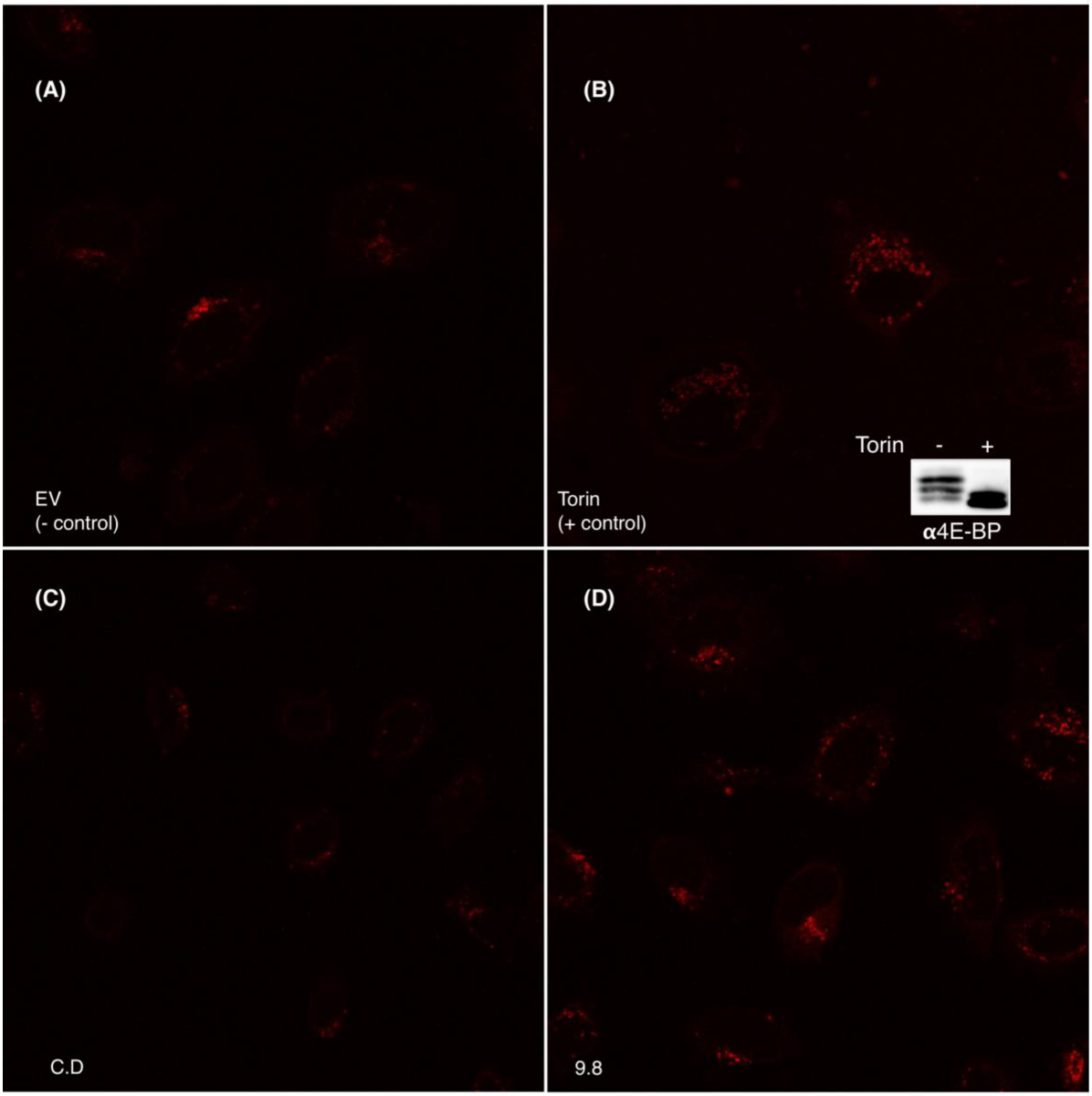

Figure 16: Lysosome levels as measured by Lysotracker are similar between Torin-treated and IpaH9.8-expressing HeLa cells.

HeLa cells were treated with Torin, a drug that inhibits the phosphorylation of mTOR substrates, or transfected with IpaH9.8, C337A catalytically dead IpaH9.8 (C.D), or an empty vector control. Lysosome level in the negative control (A) was comparable to those in the catalytically dead IpaH9.8-expressing cells (C). Likewise, Torin-treated (B) lysosomal levels were comparable to wildtype IpaH9.8-expressing cells (D). Collapse of phosphorylated mTORC1 target 4E-BP shown to validate Torin treatment in (B). Representative image of several independent experiments.

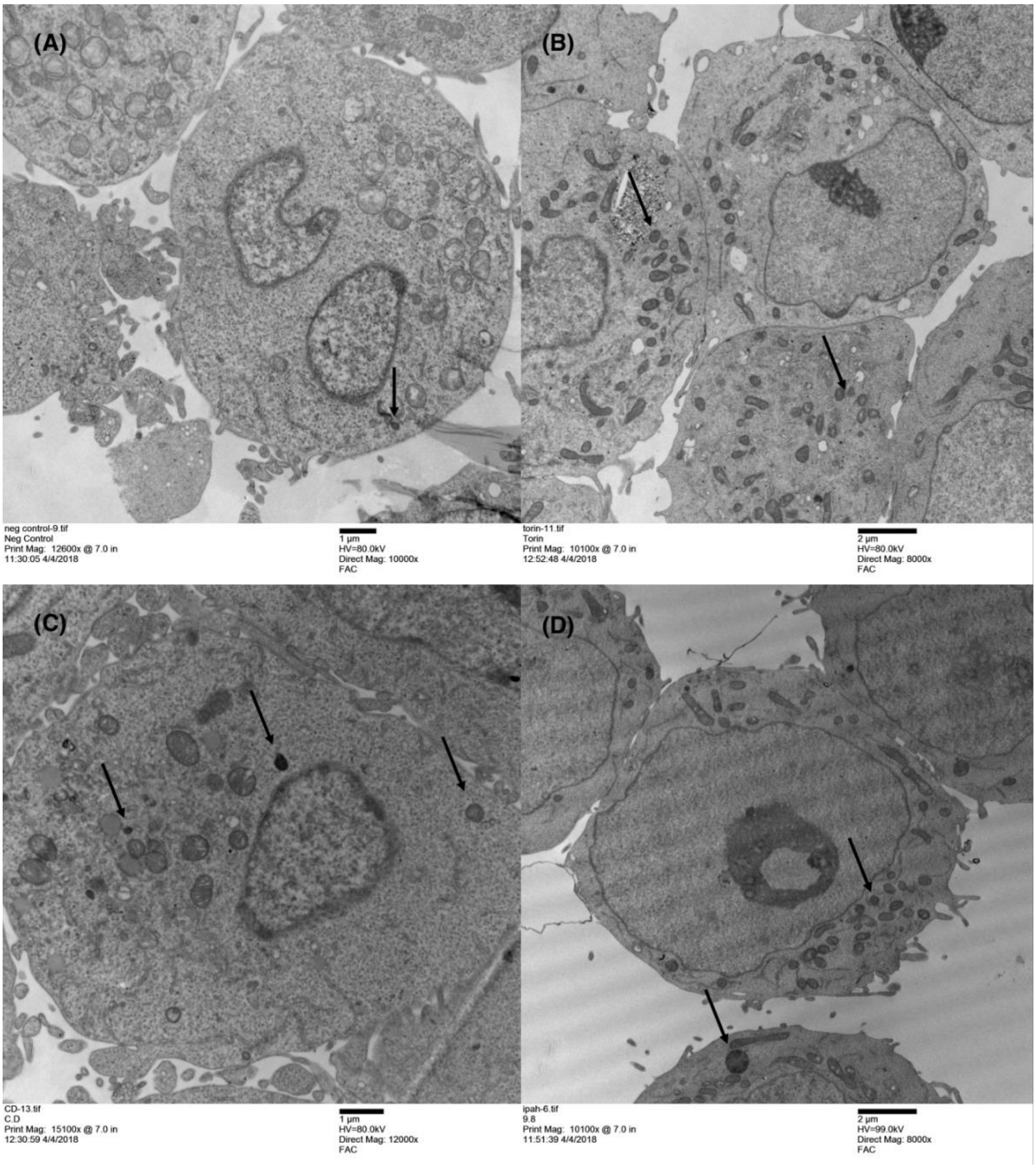

Figure 17: IpaH9.8 increases the amount of electron dense bodies in HEK293T cells, visible by transmission electron micrograph.
HEK293T cells expressing IpaH9.8, C337A IpaH9.8 (C.D), empty vector control, or Torin-treated cells were examined by transmission electron micrograph for lysosomes (likely candidates indicated by black arrows). (A) HEK cells expressing empty vector negative control. (B) HEK cells treated with Torin. (C) HEK cells expressing C337A IpaH9.8 (D) HEK cells expressing wildtype IpaH9.8.

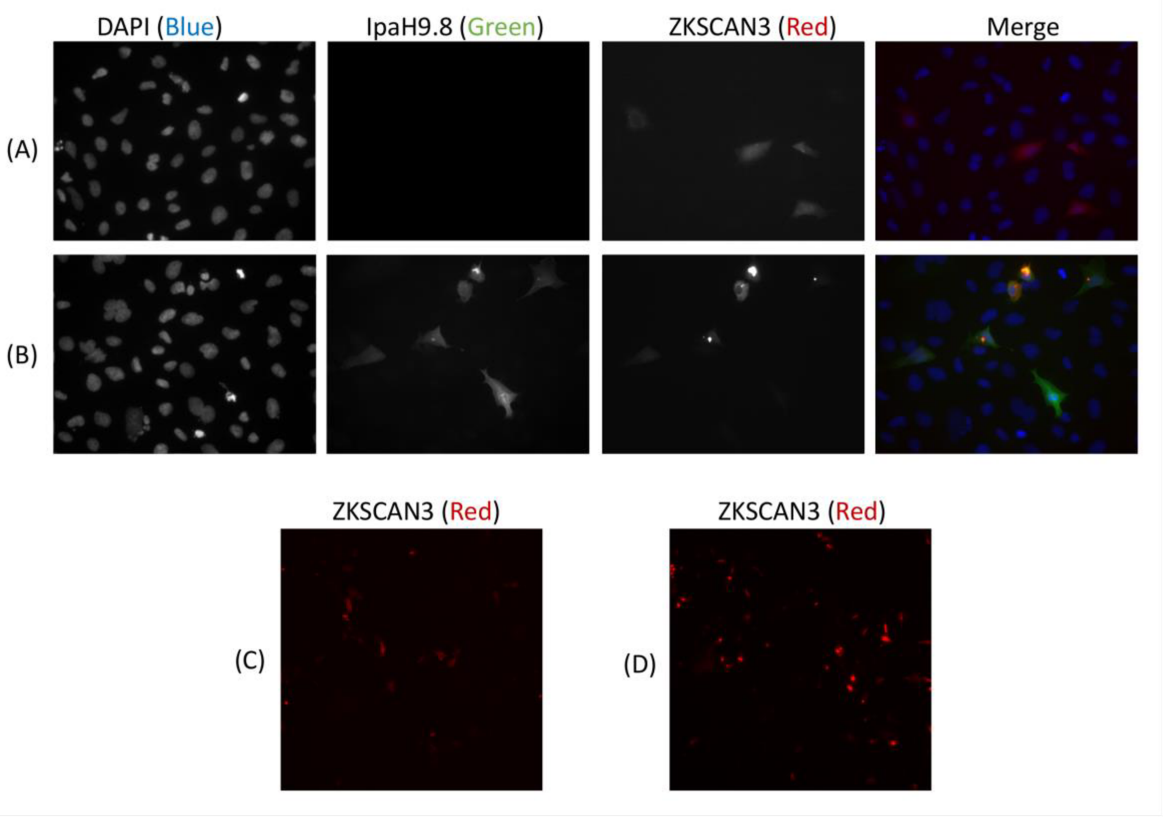

Figure 18: IpaH9.8 expression induces ZKSCAN3 perinuclear foci formation. HEK293A cells were transfected with mRuby-tagged ZKSCAN3 (red) and/or myc-tagged IpaH9.8 detected by αMyc immunofluorescence (green), and a DAPI nuclear stain (blue). (A) ZKSCAN3 localization in cells that do not express IpaH9.8 is diffuse (B) ZKSCAN3 perinuclear foci in cells that express IpaH9.8 (C) ZKSCAN3-expressing cells under lower magnification to show the widespread diffuse nature of ZKSCAN3. (D) ZKSCAN3- and IpaH9.8-expressing cells under lower magnification to show widespread IpaH9.8-induced perinuclear foci. Representative fields of several independent experiments.

Chapter 4: Discussion

4.1: Confirmation of Interaction Between UFM1 and IpaH9.8

The ubiquitin system is a key host defence and homeostasis system that is hijacked by pathogens. Interactions between pathogens and UBLs is a recently discovered phenomenon but nonetheless includes several prominent examples mostly involving SUMO and NEDD8 deconjugation. For example, the *Xanthomonas euvesicatoria* XopD protein acts as a deSUMOylase (J. G. Kim, Stork, & Mudgett, 2013), while *Chlamydia trachomatis* ChlaDUB1 acts as a deNEDDylase (Le Negrate et al., 2008). My project began with a yeast 2-hybrid hit that suggested that the *Shigella* effector IpaH9.8 interacted with the UBL UFM1 and I began to characterize this interaction. I proposed three models for this interaction (Figure 19). The simplest explanation is that IpaH9.8 is a dual-function E3 ligase that is capable of ubiquitination and UFMylation alike. Secondly, IpaH9.8 itself could be a target of UFMylation, where conjugation with UFM1 activates/modifies IpaH9.8 function in a fashion similar to Cullin-Ring Ligase NEDDylation. Thirdly, IpaH9.8 might be an E3 ligase that preferentially targets UFMylated substrates as is seen with SUMO for SUMO-targeted ubiquitin ligases (StUbLs).

Evidence to support the first model, that IpaH9.8 is a dual function (Ubiquitin and UFM1) E3 ligase, comes from my initial *in vitro* UFMylation experiments (Figure 4). Combining UFM1, IpaH9.8, UFC1, and UBA5 resulted in specific high molecular weight species that were consistent with those seen in the case of IpaH9.8 ubiquitination. These species were sensitive to DTT, suggesting a cysteine linkage rather than lysine ubiquitination. Furthermore, these species were also dependent on both IpaH9.8 and

UFM1, and also decreased the intensity of species corresponding to charged UFC1-UFM1 conjugates. These results are consistent with a model where IpaH9.8 accepts UFM1 from UFC1 for subsequent ligation, but not sufficient to discount a more regulatory function of UFM1 consistent with the other two models.

I used mass spectrometry to identify these putative IpaH9.8-UFM1 conjugates (high molecular weight) and UFC1-UFM1 conjugates (low molecular weight). IpaH9.8 and UFM1 were identified in two of the three high molecular weight species, whereas UFC1 and UFM1 were identified in all three of the lower molecular weight species (Figure 7). However, to complicate matters, UBA5 was also identified in several of these species, suggesting a more complex aggregate of proteins. Furthermore, IpaH9.8 was identified in species that are too small to explain the presence of this ~80kDa protein. I attribute this finding to the fact that many proteins in these reactions are biochemically poorly defined, and the presence of IpaH9.8 may in fact be a truncated or cleaved product. Evidence for the presence of multiple truncated IpaH9.8 species was also identified by western blotting (Figure 5B). Nonetheless, this does not explain the fact that IpaH9.8 was identified in a lane where no IpaH9.8 was added to begin with. This may be due to the fact that IpaH9.8 was loaded into the adjacent wells, and mass spectrometry is sensitive enough to pick up these proteins in such a small gel format.

4.2: Non-Canonical UFMylation

UFMylation of a protein is determined via mass spectrometry by the presence of a diagnostic VG motif. The motif arises from the last two amino acids in UFM1 that remain attached to a digested peptide due to the nature of trypsin cleavage. A particularly interesting finding from these mass spectrometry experiments was the presence of a non-

canonical (non-lysine) UFM1 modification on UBA5, either on an aspartic acid, glutamic acid, serine, or threonine. While non-canonical UFMylation is an unreported phenomenon, non-canonical ubiquitination has been reported on these very same residues (McDowell & Philpott, 2013). The uncertainty as to the specific residue of UFMylation is caused by the highly similar spectra produced by modification at the end of these residues. At the cost of time and money, the resolution of the mass spectrometry instrument can be tuned to eliminate the noisiness of the peaks on these spectra to better distinguish one species from another, and is a worthwhile exploration for a second attempt. Furthermore, there is an element of luck, as peptides can be fragmented in such a way as to suggest which residue is most likely UFMylated by either including or excluding certain amino acids. For example, a fragmentation in the middle of these uncertain residues would result in a VG mass shift in one peptide and not the other. A biochemical approach is also possible, by mutating these specific residues and searching for the continued presence of the VG-tag. Further experiments to see if this modification is absolutely dependent upon UFM1, or IpaH9.8 are necessary to rule out other possible explanations for the presence of this tag. Non-canonical UFMylation, and even canonical UFMylation of UBA5 would both prove to be novel findings in the UFM1 field. E1 regulation through PTMs is not unheard of, however, as UBA1, the canonical ubiquitin E1, is regulated through phosphorylation (Stephen, Trausch-Azar, Ciechanover, & Schwartz, 1996).

4.3: UBA5 Isoforms, Co-Immunoprecipitation, & Immunofluorescence

A factor that precipitated my mass spectrometry approach is that many species in these *in vitro* UFMylation reactions are puzzling and difficult to ascribe to specific

proteins. Differences in high molecular weight banding patterns were observed between purchased versus homemade enzymes (Figure 5 and 6). Additionally, certain species were stabilized differently by IpaH9.8 depending on the presence of UBA5 or UFC1. Nonetheless, seeing these different patterns and that UBA5 and UFC1 were identified by mass spectrometry to be present in the high molecular weight species suggests that both the purchased full length UBA5 and the homemade truncated UBA5 were functional to some degree. Very recently, the 56-amino acid extension to UBA5 has now been implicated in increasing the efficiency of UFM1 activation, and both isoforms were shown to be active (Soudah et al., 2018). The *in vitro* UFMylation assays (Figure 4) also suggested that these high molecular weight UFM1-IpaH9.8 conjugates might form without the E2 UFC1. Recently, ubiquitination independent of E1 and E2 was demonstrated by the *Legionella* effector SidE (Qiu et al., 2016). With the purification of the UFM1 E3 UFL1 at the end of my project, I have created an opportunity for additional biochemical investigation that may shed light on this complex process.

An additional explanation for species that did not fit previously observed patterns is the impurity of my purified enzymes. Additional species are present in the eluted fraction of UBA5, UFC1, and IpaH9.8 purification (Figure 8) that may have altered the function of these enzymes, or may be cross reactive with the antibodies used for detection. These impurities are also a likely explanation for the presence of less abundant species detected by total protein stain.

To further characterize the interaction between IpaH9.8 and UFM1, I determined if the proteins interacted in mammalian cells. I approached this from both ends, by affinity purification of either Myc-tagged IpaH9.8 or His-tagged UFM1, and found that

expression of one tagged protein enhanced the recovery of the other protein (Figure 8). No band shifts consistent with IpaH9.8 UFMylation were observed in these lysates, although two discrete species were observed (58kDa and 5-10kDa). This suggests that IpaH9.8 may not be directly UFMylated in a manner similar to CRLs, but rather that UFM1 interacts with IpaH9.8 to direct targets in a StUbL-like manner, or as a UFM1 E3.

Finally, in an attempt to gain localization information of UFM1 in response to IpaH9.8 (or vice-versa), I examined both proteins via fluorescent microscopy. Both proteins showed diffuse staining throughout the cell, and while expression of IpaH9.8 may have further diffused UFM1 localization within the cellular milieu, the results were not striking enough to distinguish sufficiently (Figure 14). UFM1 has been previously shown to localize to the nucleus and cytoplasm alike (Yoo et al., 2014), whereas IpaH9.8 was shown to localize to the nucleus (Toyotome et al., 2001).

4.4: IpaH9.8-Directed UFM1 Targets

Regardless of which of the 3 models for IpaH9.8-UFM1 interaction is correct, the suite of UFMylated proteins inside of a human cell might change if IpaH9.8 is expressed. This list of proteins could represent substrates for IpaH9.8-directed UFMylation, or proteins that are UFMylated and subsequently recognized by IpaH9.8 for ubiquitination. To determine this list of proteins, I compared cells that expressed only UFM1 with those that expressed UFM1 and IpaH9.8. I enriched both samples for UFM1-interacting proteins by metal-ion affinity purification of His-tagged UFM1 and analyzed both samples by mass spectrometry (Figure 9). While both samples shared considerable overlap between the most abundant proteins, I was most interested in the proteins that showed up exclusively in the sample containing IpaH9.8. It is, however, worth noting

that all UFM1-interacting proteins from this screen are of potential interest as the list of validated UFM1 substrates is a short one (Figures 10 and 11).

To that end, and as a validation for this screen, a considerable-sized node of ribosomal proteins was identified through the mass spectrometry screen (Figure 10 and 11). Ribosomal proteins along with ribosomal RNA make up both the large and small subunit of the ribosome and are designated RPLxx (large subunit) and RPSxx (small subunit) accordingly. During the course of this study, a group from Stanford conducted an affinity purification screen to identify ribosome-associated proteins and discovered UFM1 as a direct modification of many ribosomal proteins (Simsek et al., 2017). Specifically, my mass spectrometry approach identified RPS3 and RPL10, two of the three proteins that Simsek and colleagues showed to be direct targets of UFMylation. A VG modification was also found on RPL10 in my analysis.

It must be noted that the lack of a VG modification on other proteins does not necessarily imply that a protein is not UFMylated. Rather, lack of this VG modification may reflect lack of the appropriate diagnostic peptide, whether that be for stoichiometric reasons, incomplete fragmentation, or incomplete coverage. This also means that a VG modification is less likely to be discovered for a less abundant protein. It is possible to tune LC-MS/MS to render an increased coverage of certain proteins, but that comes at the expense of time, cost, and a less complete coverage of the proteins in a given sample, especially one as complex as those we provided.

The top hit from the mass spectrometry screen, present in both samples, was Filamin A. Filamin A crosslinks polymerized actin filaments and participates in the remodelling of the actin cytoskeleton. *Shigella* induces actin cytoskeletal rearrangement

during infection, and Filamin A was one of the top two hits from a yeast 2-hybrid assay looking for targets of IpaH7.8, but was ultimately discarded in favour of glomulin (Suzuki et al., 2014). Furthermore, Filamin A has very recently been shown to be controlled by IRE1α, one of the three branches of the Unfolded Protein Response (UPR) (Urra et al., 2018). As mentioned previously, one of the two known substrates of UFM1, UfBP1 stabilizes the expression of IRE1α and is intimately involved in the UPR.

Other proteins I considered but ultimately did not have time to pursue included: NONO (non-POU domain containing octamer), the highest abundance protein with a VG modification found in both samples. NONO forms a heterodimer with SFPQ (also found in high abundance in both samples, with a VG modification) to modulate pre-mRNA splicing and other transcriptional regulation events (Knott, Bond, & Fox, 2016). Additionally, a proximity-based biotin ligase screen (BioID) for UFM1-interacting partners identified NONO, albeit not as a high-confidence hit (Pirone et al., 2017). Of the 20 high-confidence hits from the BioID screen, the only one to show up in my mass spectrometry screen was another protein involved in pre-mRNA splicing, HNRNPF, which was just below the cut-off of 3 unique peptides for the IpaH9.8 experimental sample. Instead, I decided to focus on Galectin-7, as it was found only in the IpaH9.8 sample and represented a IpaH9.8-induced UMF1 target with previous links to intracellular pathogens.

4.5: Galectin-7 as an IpaH9.8-Directed UFMylation Target

Galectin-7 is a poorly characterized member of the galectin family of sugar-binding proteins. Specifically, Galectin-7 has been associated with epithelial cell maturation, migration, apoptosis, and various cancers (Saussez & Kiss, 2006). The role of

galectin-7 in epithelial cell maturation stems from observations that it prevents ubiquitination and degradation of the MAPK JNK1 (H. L. Chen et al., 2016). The galectin family contains two members whose functions have been implicated in pathogenic bacterial defence. Galectin-8 links pathogen invasion to autophagic clearance by binding to glycans exposed on bacteria-containing vesicles and the selective autophagy receptor NDP52 (Thurston et al., 2012). Galectin-3 senses T3SS-destabilized bacteria-containing vesicles and recruits GBPs to direct host innate immune responses (Feeley et al., 2017). Galectins as a whole might therefore represent an ideal target for pathogenic effector proteins. IpaH9.8 is an ideal candidate for this role as it already has an established role in targeting GBPs.

Galectin-7 was identified as one of the top 10 proteins that was enriched in UFM1 affinity purification only when IpaH9.8 was present. To further characterize this interaction, I examined the effect of UFM1 and IpaH9.8 expression on Galectin-7 levels in human cells. I found that expression of both UFM1 and IpaH9.8 stabilized Galectin-7 levels to a degree that was not recapitulated with the expression of just one of these proteins (Figure 12). This was, at first, counterintuitive, as ubiquitination of a target is often associated with decreased levels of a protein. However, monoubiquitination of a protein is often associated with non-destructive fates, and the fate of UFMylation is still a relatively uncharacterized process. To see if this stabilization was dependent on UFMylation and not ubiquitination, I tested to see if Galectin-7 was UFMylated *in vitro*. Using purified protein, I showed high molecular weight Galectin-7 positive species consistent with UFMylated Galectin-7 upon the addition of IpaH9.8 and the requisite UFMylation machinery (Figure 13). Due to the poor quality of Galectin-7 reagents

(namely the polyclonal antibody) the nature of this interaction is not entirely clear; however, if validated, UFMylation of Galectin-7 might be working similarly to SUMOylation of IκB, where SUMOylation prevents subsequent ubiquitination and destruction.

If Galectin-7 function in opposition to Galectin-3 or Galectin-8 during infection, it might make sense to stabilize Galectin-7 while downregulating fates associated with the other galectins to direct a more favourable immune response for *Shigella* infection (Figure 20). This immune response might be in line with previous observations that galectin-7 modulates JNK1 (H. L. Chen et al., 2016). Though it is entirely possible that Galectin-7 function in *Shigella* infection is not related to other Galectins, and its stabilization is for another reason altogether.

Future analysis of Galectin-7 would benefit from the construction of an epitope-tagged Galectin-7 protein to ease visualization via immunoblotting, allow cost-effective purification, and to allow visualization via immunofluorescence. Furthermore, the effect of *Shigella* infection on Galectin-7 levels would allow for more physiologically relevant observations.

4.6: ZKSCAN3 Interaction with IpaH9.8

My second aim was to characterize the interaction between IpaH9.8 and another putative interactor: ZKSCAN3. ZKSCAN3 is an mTORC1-directed negative regulator of lysosomal biogenesis, directly opposed by the transcriptional activator TFEB. Both ZKSCAN3 and TFEB are regulated by exclusion from the nucleus through phosphorylation (Chauhan et al., 2013). To see if ZKSCAN3 is subject to another level of regulation through IpaH9.8-directed ubiquitination, I purified the required proteins for an

in vitro ubiquitination assay. I showed that combining IpaH9.8, ZKSCAN3, and the requisite ubiquitination machinery results in a characteristic ubiquitination banding pattern (Figure 15). That is, the disruption of a species indicating a charged E2-Ub, and the appearance of additional (or more pronounced) higher molecular weight species. However, the higher molecular weight forms seemed to be discrete species that were not consistent with polyubiquitination, indicating that ZKSCAN3 might be multi-monoubiquitinated or diubiquitinated instead, fates not typically associated with proteasomal degradation. Nonetheless, I hypothesized that IpaH9.8 was disrupting the function of ZKSCAN3, which would lead to an increase in lysosomal biogenesis.

To see if ubiquitination of ZKSCAN3 affected lysosome level, I used an acid-tracking dye called Lysotracker. IpaH9.8-expressing cells appeared to have a higher number of lysosomes compared to C337A IpaH9.8 mutant, and an empty vector and control. This was in line with the results obtained from cells treated with Torin, a drug that increases lyosomal biogenesis in an mTORC1-dependent fashion (Figure 16). Other authors have attempted to quantify lysosomes using flow cytometry (Y. Li et al., 2016; J. Zhang et al., 2018), despite warnings in the LysoTracker manual suggesting that quantification of LysoTracker should not be attempted via Flow Cytometry or Fluometry due to lysosome staining only accounting for a small percentage of the overall fluorescence of a cell. Automated quantification through ImageJ and cell profiler could not distinguish discrete lysosomes, and initial experiments did not provide enough numbers for statistical power by manual counting. These factors indicate that this experiment would benefit from repetition with quantification.

As another independent line of inquiry to see changes in number or size of lysosomes I examined these same conditions using electron microscopy (Figure 17). Instead of discrete electron-dense bodies easily ascribed to lysosomes, I saw a variety of electron dense bodies inside cells whose characteristics could be ascribed to lysosomes, mitochondria, or other organelles. Regardless of these differences, there did seem to be a difference between the number of these electron-dense bodies recapitulating the LysoTracker assay. Wildtype IpaH9.8 and C337A IpaH9.8 both showed more electron dense bodies in line with Torin-treated cells compared to empty vector controls. It was not possible to definitively ascribe these structures as lysosomes, and as such, repetition of this experiment with an electron-dense lysosome dye, or immunogold staining is merited.

While LysoTracker and electron microscopy are indicators of lysosome size and number, neither give any information as to the functionality of the lysosome. To determine if IpaH9.8 influenced lysosomal activity and for a more quantitative assay I turned to a biochemical N-actetylglucosaminidase (NAG) assay. NAG is found in the lysosome and degrades glycolipids and glycoproteins. Its activity is measured colorimetrically by cleaving p-nitrophenol coupled to a NAG substrate. Unfortunately, the results from this NAG assay were not reproducible across different starting concentration of proteins despite the supposed dynamic range of the assay. Slight increases in initial material led to drastic changes in final NAG-activity values across a variety of lysing conditions and cell lines, and for that reason the data was not included in this thesis. A variety of other functional lysosomal assays exist including: cathepsin measurement assays, and the DQ-red BSA protease assay. These assays both rely on

cleaving lysosomal protease substrates that fluoresce in their cleaved state and would provide a good alternative future line of investigation.

Despite these set-backs in the functional lysosomal assay, fluorescent microscopy analysis of ZKSCAN3 showed a striking change in localization in the presence of IpaH9.8 (Figure 18). In the presence of IpaH9.8, ZKSCAN3 localization changed from relatively diffuse to perinuclear puncta. These perinuclear hubs may represent the sites of exclusion of ZKSCAN3 from the nucleus. If IpaH9.8 is not polyubiquitinating ZKSCAN3 and targeting it for degradation, IpaH9.8-dependent ubiquitination might instead exclude ZKSCAN3 from the nucleus as phosphorylation does. Nuclear exclusion of a negative lysosomal regulator would explain the increase in lysosome number shown by LysoTracker and electron microscopy. There is precedent for ubiquitination of a transcription factor affecting its activity independent of proteolysis as demonstrated by ubiquitination of Met4 (Flick et al., 2004). Additionally, it is possible that these puncta represent the aggresome, an amalgamation of protein destined for degradation when the normal protein degradation system (the proteasome) is overwhelmed. Regardless, ZKSCAN3 is known to be regulated through exclusion from the nucleus, and the fact that IpaH9.8 induces similar localization of ZKSCAN3 suggests that there is indeed a functional interaction between these two proteins.

ZKSCAN3 as a putative target for IpaH9.8 ubiquitination represents a mechanism by which IpaH9.8 may be turning on lysosomal biogenesis, a key component of autophagy. This may seem contradictory, as *Shigella* has effectors (namely IcsB) that it uses to downregulate autophagy. To explain this, we turn to *Salmonella* which has been shown to turn on generalized autophagy, thereby producing free amino acids and building

blocks to facilitate intracellular survival and growth (Yu et al., 2014). *Shigella* has mechanisms to avoid its specific autophagic (or more accurately, xenophagic) clearance, but IpaH9.8 ubiquitination of ZKSCAN3 might represent a mechanism by which it is turning on generalized autophagy to help with intracellular growth (Figure 21).

To further characterize the functional relationship between IpaH9.8 and ZKSCAN3, a similar approach to UFM1 could be taken with a co-purification confirmation of interaction and identification of IpaH9.8-induced ZKSCAN3 ubiquitination via mass spectrometry. Additionally, to determine if IpaH9.8 modulates the expression of the ZKSCAN3-induced CLEAR genes, qPCR could be used to measure their expression. These genes include: CTSA, SGSH, LAMP1, SUMF1, and ATP6V1A. Primers for the quantification of mRNAs for these genes were designed and ordered were but these analyses were not carried out in the interest of time.

4.7: Significance

In this thesis, I have characterized a functional interaction between the *Shigella* protein IpaH9.8 and the ubiquitin-like protein UFM1. The importance UFMylation in *Shigella* pathogenesis is two-fold. First, UFM1 is one of the poorest characterized PTMs. Its elucidation represents a basic scientific understanding into the regulation of human processes. Already UFMylation has been implicated in a variety of human conditions like cancer, neurodevelopmental diseases, and heart disease. Secondly, *Shigella* is a pathogenic bacterium that is responsible for morbidity and mortality in the developing world and for which there is no vaccine. Characterization of the *Shigella*-UFM1 interface represents the potential for future therapeutic interventions. This project has led to a better understanding of *Shigella* effector function, characterization of an poorly defined

UBA5 isoform, identification of a potentially novel UFMylation residue, confirmation of IpaH9.8-UFM1 interaction, and has identified a list of UFM1-interacting proteins including a putative substrate of IpaH9.8 dependent UFMylation: Galectin-7.

In my characterization of ZKSCAN3, the second target of IpaH9.8, I am further contributing to the understanding of a critical *Shigella* effector, and implicating *Shigella* in modulation of lysosomal biogenesis and autophagy, two processes that lie at the core of homeostasis of the human cell. I have shown that ZKSCAN3 is a putative IpaH9.8 interactor, that this interaction results in widespread cellular localization changes, and also shown that cellular lysosomal levels are influences by IpaH9.8 expression.

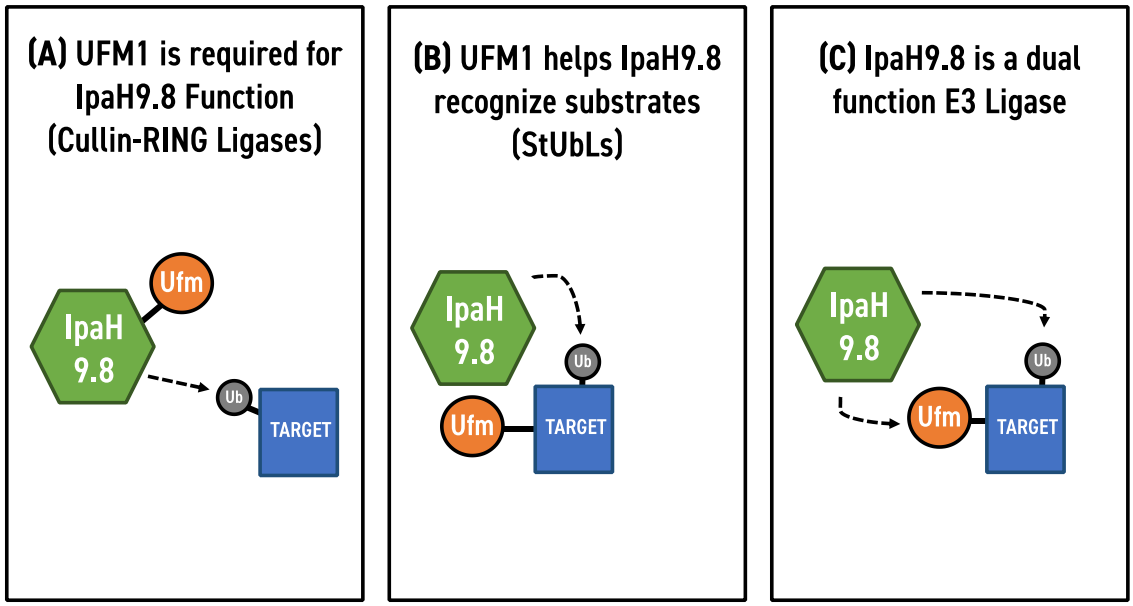

Figure 19: Three tentative models of interactions between IpaH9.8 and UFM1. (A) IpaH9.8 ubiquitin ligase enzymatic activity relies on regulation through UFMylation. This is analogous to Cullin-RING ligases that are NEDDylated before functioning as E3 ubiquitin ligases. (B) IpaH9.8 preferentially recognizes targets that are UFMylated. This is analogous to SUMO-targetted ubiquitin ligases (StUbLs) that are E3 ubiquitin ligases that preferentially recognize SUMOylated targets. (C) IpaH9.8 is an E3 ligase that is capable of ligating both ubiquitin and UFM1 to target proteins.

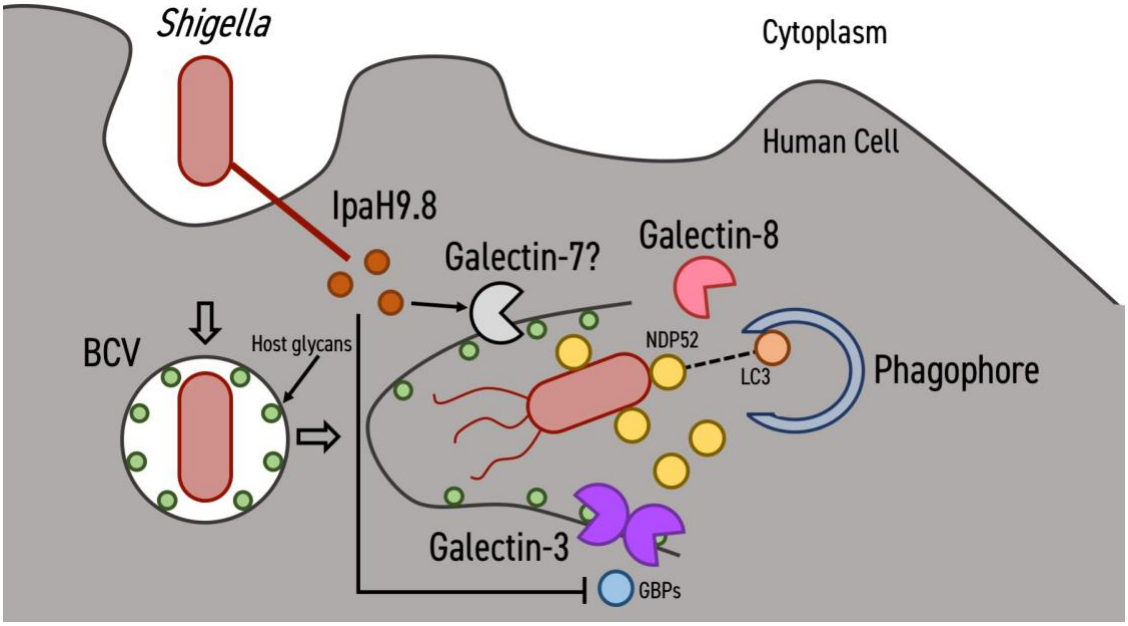

Figure 20: IpaH9.8 UFMylates Galectin-7: A tentative model.
Galectins recognize exposed glycans on the internal surface of bacteria-containing vacuole (BCV). These glycans are exposed when bacteria like *Shigella* escape the vacuole, or destabilize the vacuole through their type III secretion system (T3SS). Galectin-3 recruits Guanylate-binding proteins (GBPs) to the vacuole to clear *Shigella*. IpaH9.8 has been shown to ubiquitinate and degrade GBPs. Galectin-8 is also recruited to this vacuole and communicates with the autophagy machinery through the LC3-binding NDP52 protein. *Shigella* uses its T3SS to inject IpaH9.8 into the host cytosol. I propose that IpaH9.8 might stabilize Galectin-7 through UFMylation to shift the Galectin response away from Galectin-3 and Galectin-8 mediated fates. The function of Galectin-7 in *Shigella* early entry is unknown.

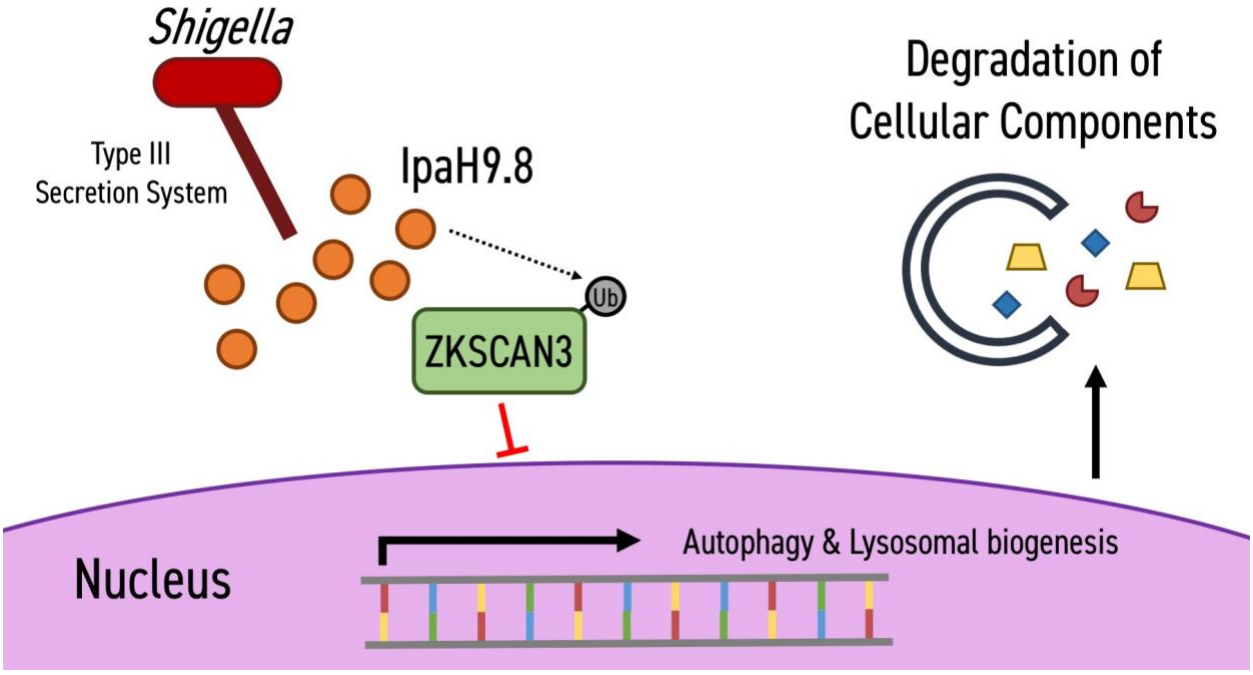

Figure 21: ZKSCAN3 is ubiquitinated by IpaH9.8: A proposed model.
Shigella uses its type III secretion system to inject IpaH9.8 into the cytosol of the infected cell. IpaH9.8 ubiquitinates ZKSCAN3 leading to its exclusion from the nucleus either through degradation or localization signals. ZKSCAN3 is a negative transcriptional regulator of autophagy and lysosomal biogenesis; its exclusion from the nucleus leads to an increase in autophagy, breaking down cellular structures into building blocks that *Shigella* needs for intracellular growth.

References

Abby, S. S., & Rocha, E. P. (2012). The non-flagellar type III secretion system evolved from the bacterial flagellum and diversified into host-cell adapted systems. *PLoS Genet, 8*(9), e1002983. doi:10.1371/journal.pgen.1002983

Abramovitch, R. B., Janjusevic, R., Stebbins, C. E., & Martin, G. B. (2006). Type III effector AvrPtoB requires intrinsic E3 ubiquitin ligase activity to suppress plant cell death and immunity. *Proc Natl Acad Sci U S A, 103*(8), 2851-2856. doi:10.1073/pnas.0507892103

Al-Hakim, A. K., Zagorska, A., Chapman, L., Deak, M., Peggie, M., & Alessi, D. R. (2008). Control of AMPK-related kinases by USP9X and atypical Lys(29)/Lys(33)-linked polyubiquitin chains. *Biochem J, 411*(2), 249-260. doi:10.1042/BJ20080067

Al-Zeer, M. A., Al-Younes, H. M., Lauster, D., Abu Lubad, M., & Meyer, T. F. (2013). Autophagy restricts *Chlamydia trachomatis* growth in human macrophages via IFNG-inducible guanylate binding proteins. *Autophagy, 9*(1), 50-62. doi:10.4161/auto.22482

Anderson, D. M., & Frank, D. W. (2012). Five mechanisms of manipulation by bacterial effectors: a ubiquitous theme. *PLoS Pathog, 8*(8), e1002823. doi:10.1371/journal.ppat.1002823

Aravind, L., & Koonin, E. V. (2000). The U box is a modified RING finger - a common domain in ubiquitination. *Curr Biol, 10*(4), R132-134.

Ashida, H., Kim, M., Schmidt-Supprian, M., Ma, A., Ogawa, M., & Sasakawa, C. (2010). A bacterial E3 ubiquitin ligase IpaH9.8 targets NEMO/IKKgamma to dampen the host NF-kappaB-mediated inflammatory response. *Nat Cell Biol, 12*(1), 66-73; sup pp 61-69. doi:10.1038/ncb2006

Ashida, H., Mimuro, H., & Sasakawa, C. (2015). *Shigella* manipulates host immune responses by delivering effector proteins with specific roles. *Front Immunol, 6*, 219. doi:10.3389/fimmu.2015.00219

Ashida, H., Nakano, H., & Sasakawa, C. (2013). *Shigella* IpaH0722 E3 ubiquitin ligase effector targets TRAF2 to inhibit PKC-NF-kappaB activity in invaded epithelial cells. *PLoS Pathog, 9*(6), e1003409. doi:10.1371/journal.ppat.1003409

Ashida, H., Toyotome, T., Nagai, T., & Sasakawa, C. (2007). *Shigella* chromosomal IpaH proteins are secreted via the type III secretion system and act as effectors. *Mol Microbiol, 63*(3), 680-693. doi:10.1111/j.1365-2958.2006.05547.x

Azfer, A., Niu, J., Rogers, L. M., Adamski, F. M., & Kolattukudy, P. E. (2006). Activation of endoplasmic reticulum stress response during the development of ischemic heart disease. *Am J Physiol Heart Circ Physiol, 291*(3), H1411-1420. doi:10.1152/ajpheart.01378.2005

Bacik, J. P., Walker, J. R., Ali, M., Schimmer, A. D., & Dhe-Paganon, S. (2010). Crystal structure of the human ubiquitin-activating enzyme 5 (UBA5) bound to ATP: mechanistic insights into a minimalistic E1 enzyme. *J Biol Chem, 285*(26), 20273-20280. doi:10.1074/jbc.M110.102921

Baldi, L., Brown, K., Franzoso, G., & Siebenlist, U. (1996). Critical role for lysines 21 and 22 in signal-induced, ubiquitin-mediated proteolysis of I kappa B-alpha. *J Biol Chem, 271*(1), 376-379.

Baxt, L. A., & Goldberg, M. B. (2014). Host and bacterial proteins that repress recruitment of LC3 to *Shigella* early during infection. *PLoS One, 9*(4), e94653. doi:10.1371/journal.pone.0094653

Behrends, C., & Harper, J. W. (2011). Constructing and decoding unconventional ubiquitin chains. *Nat Struct Mol Biol, 18*(5), 520-528. doi:10.1038/nsmb.2066

Bennett, E. J., & Harper, J. W. (2008). DNA damage: Ubiquitin marks the spot. *Nat Struct Mol Biol, 15*(1), 20-22. doi:10.1038/nsmb0108-20

Bernal-Bayard, J., & Ramos-Morales, F. (2009). *Salmonella* type III secretion effector SlrP is an E3 ubiquitin ligase for mammalian thioredoxin. *J Biol Chem, 284*(40), 27587-27595. doi:10.1074/jbc.M109.010363

Bowman, G. D., O'Donnell, M., & Kuriyan, J. (2004). Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. *Nature, 429*(6993), 724-730. doi:10.1038/nature02585

Buchrieser, C., Glaser, P., Rusniok, C., Nedjari, H., D'Hauteville, H., Kunst, F., . . . Parsot, C. (2000). The virulence plasmid pWR100 and the repertoire of proteins secreted by the type III secretion apparatus of *Shigella flexneri*. *Mol Microbiol, 38*(4), 760-771.

Cai, Y., Pi, W., Sivaprakasam, S., Zhu, X., Zhang, M., Chen, J., . . . Li, H. (2015). UFBP1, a key component of the ufm1 conjugation system, is essential for ufmylation-mediated regulation of erythroid development. *PLoS Genet, 11*(11), e1005643. doi:10.1371/journal.pgen.1005643

Campbell-Valois, F. X., Sachse, M., Sansonetti, P. J., & Parsot, C. (2015). Escape of actively secreting *Shigella flexneri* from ATG8/LC3-positive vacuoles formed during cell-to-cell spread is facilitated by IcsB and VirA. *MBio, 6*(3), e02567-02514. doi:10.1128/mBio.02567-14

Cappadocia, L., & Lima, C. D. (2018). Ubiquitin-like protein conjugation: Structures, chemistry, and mechanism. *Chem Rev, 118*(3), 889-918. doi:10.1021/acs.chemrev.6b00737

Cardozo, T., & Pagano, M. (2004). The SCF ubiquitin ligase: Insights into a molecular machine. *Nat Rev Mol Cell Biol, 5*(9), 739-751. doi:10.1038/nrm1471

CDC. (2017, May 31, 2017). Shigellosis. *Infectious Diseases Related to Travel.* Retrieved from https://wwwnc.cdc.gov/travel/yellowbook/2018/infectious-diseases-related-to-travel/shigellosis

Cemma, M., Kim, P. K., & Brumell, J. H. (2011). The ubiquitin-binding adaptor proteins p62/SQSTM1 and NDP52 are recruited independently to bacteria-associated microdomains to target *Salmonella* to the autophagy pathway. *Autophagy, 7*(3), 341-345.

Chau, V., Tobias, J. W., Bachmair, A., Marriott, D., Ecker, D. J., Gonda, D. K., & Varshavsky, A. (1989). A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science, 243*(4898), 1576-1583.

Chauhan, S., Goodwin, J. G., Chauhan, S., Manyam, G., Wang, J., Kamat, A. M., & Boyd, D. D. (2013). ZKSCAN3 is a master transcriptional repressor of autophagy. *Mol Cell, 50*(1), 16-28. doi:10.1016/j.molcel.2013.01.024

Chazotte, B. (2011). Labeling lysosomes in live cells with LysoTracker. *Cold Spring Harb Protoc, 2011*(2), pdb prot5571. doi:10.1101/pdb.prot5571

Chen, H. L., Chiang, P. C., Lo, C. H., Lo, Y. H., Hsu, D. K., Chen, H. Y., & Liu, F. T. (2016). Galectin-7 regulates keratinocyte proliferation and differentiation through JNK-miR-203-p63 signaling. *J Invest Dermatol, 136*(1), 182-191. doi:10.1038/JID.2015.366

Chen, Z. J., Niles, E. G., & Pickart, C. M. (1991). Isolation of a cDNA encoding a mammalian multiubiquitinating enzyme (E225K) and overexpression of the functional enzyme in *Escherichia coli*. *J Biol Chem, 266*(24), 15698-15704.

Cheung, P., Tanner, K. G., Cheung, W. L., Sassone-Corsi, P., Denu, J. M., & Allis, C. D. (2000). Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. *Mol Cell, 5*(6), 905-915.

Chou, Y. C., Keszei, A. F., Rohde, J. R., Tyers, M., & Sicheri, F. (2012). Conserved structural mechanisms for autoinhibition in IpaH ubiquitin ligases. *J Biol Chem, 287*(1), 268-275. doi:10.1074/jbc.M111.316265

Ciechanover, A., Heller, H., Elias, S., Haas, A. L., & Hershko, A. (1980). ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. *Proc Natl Acad Sci U S A, 77*(3), 1365-1368.

Ciesla, J., Fraczyk, T., & Rode, W. (2011). Phosphorylation of basic amino acid residues in proteins: Important but easily missed. *Acta Biochim Pol, 58*(2), 137-148.

Coburn, B., Sekirov, I., & Finlay, B. B. (2007). Type III secretion systems and disease. *Clin Microbiol Rev, 20*(4), 535-549. doi:10.1128/CMR.00013-07

Cohen, M. S., & Chang, P. (2018). Insights into the biogenesis, function, and regulation of ADP-ribosylation. *Nat Chem Biol, 14*(3), 236-243. doi:10.1038/nchembio.2568

Cohen, P., & Tcherpakov, M. (2010). Will the ubiquitin system furnish as many drug targets as protein kinases? *Cell, 143*(5), 686-693. doi:10.1016/j.cell.2010.11.016

Cossart, P., & Sansonetti, P. J. (2004). Bacterial invasion: The paradigms of enteroinvasive pathogens. *Science, 304*(5668), 242-248. doi:10.1126/science.1090124

de Jong, M. F., Liu, Z., Chen, D., & Alto, N. M. (2016). *Shigella flexneri* suppresses NF-kappaB activation by inhibiting linear ubiquitin chain ligation. *Nat Microbiol, 1*(7), 16084. doi:10.1038/nmicrobiol.2016.84

Deng, L., Wang, C., Spencer, E., Yang, L., Braun, A., You, J., . . . Chen, Z. J. (2000). Activation of the IkappaB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. *Cell, 103*(2), 351-361.

Deshaies, R. J., & Joazeiro, C. A. (2009). RING domain E3 ubiquitin ligases. *Annu Rev Biochem, 78*, 399-434. doi:10.1146/annurev.biochem.78.101807.093809

Desterro, J. M., Rodriguez, M. S., & Hay, R. T. (1998). SUMO-1 modification of IkappaBalpha inhibits NF-kappaB activation. *Mol Cell, 2*(2), 233-239.

Dove, K. K., & Klevit, R. E. (2017). RING-between-RING E3 ligases: Emerging themes amid the variations. *J Mol Biol, 429*(22), 3363-3375. doi:10.1016/j.jmb.2017.08.008

Drazic, A., Myklebust, L. M., Ree, R., & Arnesen, T. (2016). The world of protein acetylation. *Biochim Biophys Acta, 1864*(10), 1372-1401. doi:10.1016/j.bbapap.2016.06.007

Dynek, J. N., Goncharov, T., Dueber, E. C., Fedorova, A. V., Izrael-Tomasevic, A., Phu, L., . . . Vucic, D. (2010). c-IAP1 and UbcH5 promote K11-linked polyubiquitination of RIP1 in TNF signalling. *EMBO J, 29*(24), 4198-4209. doi:10.1038/emboj.2010.300

Elsasser, S., & Finley, D. (2005). Delivery of ubiquitinated substrates to protein-unfolding machines. *Nat Cell Biol, 7*(8), 742-749. doi:10.1038/ncb0805-742

Enninga, J., Mounier, J., Sansonetti, P., & Tran Van Nhieu, G. (2005). Secretion of type III effectors into host cells in real time. *Nat Methods, 2*(12), 959-965. doi:10.1038/nmeth804

Ezkurdia, I., Juan, D., Rodriguez, J. M., Frankish, A., Diekhans, M., Harrow, J., . . . Tress, M. L. (2014). Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes. *Hum Mol Genet, 23*(22), 5866-5878. doi:10.1093/hmg/ddu309

Feeley, E. M., Pilla-Moffett, D. M., Zwack, E. E., Piro, A. S., Finethy, R., Kolb, J. P., . . . Coers, J. (2017). Galectin-3 directs antimicrobial guanylate binding proteins to vacuoles furnished with bacterial secretion systems. *Proc Natl Acad Sci U S A, 114*(9), E1698-E1706. doi:10.1073/pnas.1615771114

Feldman, R. M., Correll, C. C., Kaplan, K. B., & Deshaies, R. J. (1997). A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p. *Cell, 91*(2), 221-230.

Finley, D., Bartel, B., & Varshavsky, A. (1989). The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis. *Nature, 338*(6214), 394-401. doi:10.1038/338394a0

Finley, D., Ozkaynak, E., & Varshavsky, A. (1987). The yeast polyubiquitin gene is essential for resistance to high temperatures, starvation, and other stresses. *Cell, 48*(6), 1035-1046.

Fiskin, E., Bhogaraju, S., Herhaus, L., Kalayil, S., Hahn, M., & Dikic, I. (2017). Structural basis for the recognition and degradation of host TRIM proteins by *Salmonella* effector SopA. *Nat Commun, 8*, 14004. doi:10.1038/ncomms14004

Flick, K., Ouni, I., Wohlschlegel, J. A., Capati, C., McDonald, W. H., Yates, J. R., & Kaiser, P. (2004). Proteolysis-independent regulation of the transcription factor Met4 by a single Lys 48-linked ubiquitin chain. *Nat Cell Biol, 6*(7), 634-641. doi:10.1038/ncb1143

Galan, J. M., & Peter, M. (1999). Ubiquitin-dependent degradation of multiple F-box proteins by an autocatalytic mechanism. *Proc Natl Acad Sci U S A, 96*(16), 9124-9129.

Gallagher, E., Gao, M., Liu, Y. C., & Karin, M. (2006). Activation of the E3 ubiquitin ligase Itch through a phosphorylation-induced conformational change. *Proc Natl Acad Sci U S A, 103*(6), 1717-1722. doi:10.1073/pnas.0510664103

Goldberg, M. B., & Theriot, J. A. (1995). *Shigella flexneri* surface protein IcsA is sufficient to direct actin-based motility. *Proc Natl Acad Sci U S A, 92*(14), 6572-6576.

Gorden, J., & Small, P. L. (1993). Acid resistance in enteric bacteria. *Infect Immun, 61*(1), 364-367.

Greene, W. C., & Chen, L. F. (2004). Regulation of NF-kappaB action by reversible acetylation. *Novartis Found Symp, 259*, 208-217; discussion 218-225.

Guzzo, C. M., & Matunis, M. J. (2013). Expanding SUMO and ubiquitin-mediated signaling through hybrid SUMO-ubiquitin chains and their receptors. *Cell Cycle, 12*(7), 1015-1017. doi:10.4161/cc.24332

Haglund, K., Sigismund, S., Polo, S., Szymkiewicz, I., Di Fiore, P. P., & Dikic, I. (2003). Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation. *Nat Cell Biol, 5*(5), 461-466. doi:10.1038/ncb983

Haraga, A., & Miller, S. I. (2006). A *Salmonella* type III secretion effector interacts with the mammalian serine/threonine protein kinase PKN1. *Cell Microbiol, 8*(5), 837-846. doi:10.1111/j.1462-5822.2005.00670.x

Harper, J. W., & Bennett, E. J. (2016). Proteome complexity and the forces that drive proteome imbalance. *Nature, 537*(7620), 328-338. doi:10.1038/nature19947

Hershko, A., & Ciechanover, A. (1998). The ubiquitin system. *Annu Rev Biochem, 67*, 425-479. doi:10.1146/annurev.biochem.67.1.425

Hershko, A., Ciechanover, A., Heller, H., Haas, A. L., & Rose, I. A. (1980). Proposed role of ATP in protein breakdown: Conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proc Natl Acad Sci U S A, 77*(4), 1783-1786.

Hipp, M. S., Kalveram, B., Raasi, S., Groettrup, M., & Schmidtke, G. (2005). FAT10, a ubiquitin-independent signal for proteasomal degradation. *Mol Cell Biol, 25*(9), 3483-3491. doi:10.1128/MCB.25.9.3483-3491.2005

Hoch, J. A. (2000). Two-component and phosphorelay signal transduction. *Curr Opin Microbiol, 3*(2), 165-170.

Hochstrasser, M. (2006). Lingering mysteries of ubiquitin-chain assembly. *Cell, 124*(1), 27-34. doi:10.1016/j.cell.2005.12.025

Hoeller, D., Crosetto, N., Blagoev, B., Raiborg, C., Tikkanen, R., Wagner, S., . . . Dikic, I. (2006). Regulation of ubiquitin-binding proteins by monoubiquitination. *Nat Cell Biol, 8*(2), 163-169. doi:10.1038/ncb1354

Hoffmann, A., Leung, T. H., & Baltimore, D. (2003). Genetic analysis of NF-kappaB/Rel transcription factors defines functional specificities. *EMBO J, 22*(20), 5530-5539. doi:10.1093/emboj/cdg534

Homrich, M., Wobst, H., Laurini, C., Sabrowski, J., Schmitz, B., & Diestel, S. (2014). Cytoplasmic domain of NCAM140 interacts with ubiquitin-fold modifier-conjugating enzyme-1 (Ufc1). *Exp Cell Res, 324*(2), 192-199. doi:10.1016/j.yexcr.2014.04.003

Hu, J., Worrall, L. J., Hong, C., Vuckovic, M., Atkinson, C. E., Caveney, N., . . . Strynadka, N. C. J. (2018). Cryo-EM analysis of the T3S injectisome reveals the structure of the needle and open secretin. *Nat Commun, 9*(1), 3840. doi:10.1038/s41467-018-06298-8

Huibregtse, J., & Rohde, J. R. (2014). Hell's BELs: Bacterial E3 ligases that exploit the eukaryotic ubiquitin machinery. *PLoS Pathog, 10*(8), e1004255. doi:10.1371/journal.ppat.1004255

Humbard, M. A., Miranda, H. V., Lim, J. M., Krause, D. J., Pritz, J. R., Zhou, G., . . . Maupin-Furlow, J. A. (2010). Ubiquitin-like small archaeal modifier proteins (SAMPs) in *Haloferax volcanii*. *Nature, 463*(7277), 54-60. doi:10.1038/nature08659

Jentsch, S., & Pyrowolakis, G. (2000). Ubiquitin and its kin: How close are the family ties? *Trends Cell Biol, 10*(8), 335-342.

Jin, L., Williamson, A., Banerjee, S., Philipp, I., & Rape, M. (2008). Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex. *Cell, 133*(4), 653-665. doi:10.1016/j.cell.2008.04.012

Johnson, L. N., & Barford, D. (1993). The effects of phosphorylation on the structure and function of proteins. *Annu Rev Biophys Biomol Struct, 22*, 199-232. doi:10.1146/annurev.bb.22.060193.001215

Jung, D. J., Sung, H. S., Goo, Y. W., Lee, H. M., Park, O. K., Jung, S. Y., . . . Lee, Y. C. (2002). Novel transcription coactivator complex containing activating signal cointegrator 1. *Mol Cell Biol, 22*(14), 5203-5211.

Jura, N., Scotto-Lavino, E., Sobczyk, A., & Bar-Sagi, D. (2006). Differential modification of Ras proteins by ubiquitination. *Mol Cell, 21*(5), 679-687. doi:10.1016/j.molcel.2006.02.011

Kabeya, Y., Mizushima, N., Yamamoto, A., Oshitani-Okamoto, S., Ohsumi, Y., & Yoshimori, T. (2004). LC3, GABARAP and GATE16 localize to autophagosomal membrane depending on form-II formation. *J Cell Sci, 117*(Pt 13), 2805-2812. doi:10.1242/jcs.01131

Kang, S. H., Kim, G. R., Seong, M., Baek, S. H., Seol, J. H., Bang, O. S., . . . Chung, C. H. (2007). Two novel ubiquitin-fold modifier 1 (Ufm1)-specific proteases, UfSP1 and UfSP2. *J Biol Chem, 282*(8), 5256-5262. doi:10.1074/jbc.M610590200

Kania, D. A., Hazen, T. H., Hossain, A., Nataro, J. P., & Rasko, D. A. (2016). Genome diversity of *Shigella boydii*. *Pathog Dis, 74*(4), ftw027. doi:10.1093/femspd/ftw027

Kaur, J., & Debnath, J. (2015). Autophagy at the crossroads of catabolism and anabolism. *Nat Rev Mol Cell Biol, 16*(8), 461-472. doi:10.1038/nrm4024

Keszei, A. F., & Sicheri, F. (2017). Mechanism of catalysis, E2 recognition, and autoinhibition for the IpaH family of bacterial E3 ubiquitin ligases. *Proc Natl Acad Sci U S A, 114*(6), 1311-1316. doi:10.1073/pnas.1611595114

Keszei, A. F., Tang, X., McCormick, C., Zeqiraj, E., Rohde, J. R., Tyers, M., & Sicheri, F. (2014). Structure of an SspH1-PKN1 complex reveals the basis for host substrate recognition and mechanism of activation for a bacterial E3 ubiquitin ligase. *Mol Cell Biol, 34*(3), 362-373. doi:10.1128/MCB.01360-13

Khoury, G. A., Baliban, R. C., & Floudas, C. A. (2011). Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci Rep, 1*. doi:10.1038/srep00090

Kim, D. W., Lenzen, G., Page, A. L., Legrain, P., Sansonetti, P. J., & Parsot, C. (2005). The *Shigella flexneri* effector OspG interferes with innate immune responses by targeting ubiquitin-conjugating enzymes. *Proc Natl Acad Sci U S A, 102*(39), 14046-14051. doi:10.1073/pnas.0504466102

Kim, H. C., & Huibregtse, J. M. (2009). Polyubiquitination by HECT E3s and the determinants of chain type specificity. *Mol Cell Biol, 29*(12), 3307-3318. doi:10.1128/MCB.00240-09

Kim, J. G., Stork, W., & Mudgett, M. B. (2013). *Xanthomonas* type III effector XopD desumoylates tomato transcription factor SlERF4 to suppress ethylene responses and promote pathogen growth. *Cell Host Microbe, 13*(2), 143-154. doi:10.1016/j.chom.2013.01.006

Kim, M., Ogawa, M., Fujita, Y., Yoshikawa, Y., Nagai, T., Koyama, T., . . . Sasakawa, C. (2009). Bacteria hijack integrin-linked kinase to stabilize focal adhesions and block cell detachment. *Nature, 459*(7246), 578-582. doi:10.1038/nature07952

Kim, W., Bennett, E. J., Huttlin, E. L., Guo, A., Li, J., Possemato, A., . . . Gygi, S. P. (2011). Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol Cell, 44*(2), 325-340. doi:10.1016/j.molcel.2011.08.025

Kirisako, T., Kamei, K., Murata, S., Kato, M., Fukumoto, H., Kanie, M., . . . Iwai, K. (2006). A ubiquitin ligase complex assembles linear polyubiquitin chains. *EMBO J, 25*(20), 4877-4887. doi:10.1038/sj.emboj.7601360

Knott, G. J., Bond, C. S., & Fox, A. H. (2016). The DBHS proteins SFPQ, NONO and PSPC1: A multipurpose molecular scaffold. *Nucleic Acids Res, 44*(9), 3989-4004. doi:10.1093/nar/gkw271

Koegl, M., Hoppe, T., Schlenker, S., Ulrich, H. D., Mayer, T. U., & Jentsch, S. (1999). A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. *Cell, 96*(5), 635-644.

Komatsu, M., Chiba, T., Tatsumi, K., Iemura, S., Tanida, I., Okazaki, N., . . . Tanaka, K. (2004). A novel protein-conjugating system for Ufm1, a ubiquitin-fold modifier. *EMBO J, 23*(9), 1977-1986. doi:10.1038/sj.emboj.7600205

Kornitzer, D., Raboy, B., Kulka, R. G., & Fink, G. R. (1994). Regulated degradation of the transcription factor Gcn4. *EMBO J, 13*(24), 6021-6030.

Krajewski, W. A., & Becker, P. B. (1998). Reconstitution of hyperacetylated, DNase I-sensitive chromatin characterized by high conformational flexibility of nucleosomal DNA. *Proc Natl Acad Sci U S A, 95*(4), 1540-1545.

Kramer, A., Schwebke, I., & Kampf, G. (2006). How long do nosocomial pathogens persist on inanimate surfaces? A systematic review. *BMC Infect Dis, 6*, 130. doi:10.1186/1471-2334-6-130

Kravtsova-Ivantsiv, Y., & Ciechanover, A. (2012). Non-canonical ubiquitin-based signals for proteasomal degradation. *J Cell Sci, 125*(Pt 3), 539-548. doi:10.1242/jcs.093567

Kubori, T., Shinzawa, N., Kanuka, H., & Nagai, H. (2010). *Legionella* metaeffector exploits host proteasome to temporally regulate cognate effector. *PLoS Pathog, 6*(12), e1001216. doi:10.1371/journal.ppat.1001216

Kwon, J., Cho, H. J., Han, S. H., No, J. G., Kwon, J. Y., & Kim, H. (2010). A novel LZAP-binding protein, NLBP, inhibits cell invasion. *J Biol Chem, 285*(16), 12232-12240. doi:10.1074/jbc.M109.065920

Le Negrate, G., Krieg, A., Faustin, B., Loeffler, M., Godzik, A., Krajewski, S., & Reed, J. C. (2008). ChlaDub1 of *Chlamydia trachomatis* suppresses NF-kappaB activation and inhibits IkappaBalpha ubiquitination and degradation. *Cell Microbiol, 10*(9), 1879-1892. doi:10.1111/j.1462-5822.2008.01178.x

Lehmann, G., Udasin, R. G., Livneh, I., & Ciechanover, A. (2017). Identification of UBact, a ubiquitin-like protein, along with other homologous components of a conjugation system and the proteasome in different gram-negative bacteria. *Biochem Biophys Res Commun, 483*(3), 946-950. doi:10.1016/j.bbrc.2017.01.037

Lemaire, K., Moura, R. F., Granvik, M., Igoillo-Esteve, M., Hohmeier, H. E., Hendrickx, N., . . . Schuit, F. (2011). Ubiquitin fold modifier 1 (UFM1) and its target UFBP1 protect pancreatic beta cells from ER stress-induced apoptosis. *PLoS One, 6*(4), e18517. doi:10.1371/journal.pone.0018517

Levin, I., Eakin, C., Blanc, M. P., Klevit, R. E., Miller, S. I., & Brzovic, P. S. (2010). Identification of an unconventional E3 binding surface on the UbcH5 ~ Ub conjugate recognized by a pathogenic bacterial E3 ligase. *Proc Natl Acad Sci U S A, 107*(7), 2848-2853. doi:10.1073/pnas.0914821107

Li, H., Xu, H., Zhou, Y., Zhang, J., Long, C., Li, S., . . . Shao, F. (2007). The phosphothreonine lyase activity of a bacterial type III effector family. *Science, 315*(5814), 1000-1003. doi:10.1126/science.1138960

Li, P., Jiang, W., Yu, Q., Liu, W., Zhou, P., Li, J., . . . Shao, F. (2017). Ubiquitination and degradation of GBPs by a *Shigella* effector to suppress host defence. *Nature, 551*(7680), 378-383. doi:10.1038/nature24467

Li, Y., Xu, M., Ding, X., Yan, C., Song, Z., Chen, L., . . . Yang, C. (2016). Protein kinase C controls lysosome biogenesis independently of mTORC1. *Nat Cell Biol, 18*(10), 1065-1077. doi:10.1038/ncb3407

Liu, J., Wang, Y., Song, L., Zeng, L., Yi, W., Liu, T., . . . Cong, Y. S. (2017). A critical role of DDRGK1 in endoplasmic reticulum homoeostasis via regulation of IRE1alpha stability. *Nat Commun, 8*, 14186. doi:10.1038/ncomms14186

Livio, S., Strockbine, N. A., Panchalingam, S., Tennant, S. M., Barry, E. M., Marohn, M. E., . . . Levine, M. M. (2014). *Shigella* isolates from the global enteric multicenter study inform vaccine development. *Clin Infect Dis, 59*(7), 933-941. doi:10.1093/cid/ciu468

Lorick, K. L., Jensen, J. P., Fang, S., Ong, A. M., Hatakeyama, S., & Weissman, A. M. (1999). RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proc Natl Acad Sci U S A, 96*(20), 11364-11369.

Low, K. J., Baptista, J., Babiker, M., Caswell, R., King, C., Ellard, S., & Scurr, I. (2018). Hemizygous UBA5 missense mutation unmasks recessive disorder in a patient with infantile-onset encephalopathy, acquired microcephaly, small cerebellum, movement disorder and severe neurodevelopmental delay. *Eur J Med Genet*. doi:10.1016/j.ejmg.2018.06.009

Lu, H., Yang, Y., Allister, E. M., Wijesekara, N., & Wheeler, M. B. (2008). The identification of potential factors associated with the development of type 2 diabetes: A quantitative proteomics approach. *Mol Cell Proteomics, 7*(8), 1434-1451. doi:10.1074/mcp.M700478-MCP200

Luscher, B., Butepage, M., Eckei, L., Krieg, S., Verheugd, P., & Shilton, B. H. (2018). ADP-ribosylation, a multifaceted posttranslational modification involved in the control of cell physiology in health and disease. *Chem Rev, 118*(3), 1092-1136. doi:10.1021/acs.chemrev.7b00122

Man, S. M., Karki, R., Malireddi, R. K., Neale, G., Vogel, P., Yamamoto, M., . . . Kanneganti, T. D. (2015). The transcription factor IRF1 and guanylate-binding proteins target activation of the AIM2 inflammasome by *Francisella* infection. *Nat Immunol, 16*(5), 467-475. doi:10.1038/ni.3118

Mani, S., Wierzba, T., & Walker, R. I. (2016). Status of vaccine research and development for *Shigella*. *Vaccine, 34*(26), 2887-2894. doi:10.1016/j.vaccine.2016.02.075

Maran, S., Lee, Y. Y., Xu, S., Rajab, N. S., Hasan, N., Syed Abdul Aziz, S. H., . . . Zilfalil, B. A. (2013). Gastric precancerous lesions are associated with gene variants in *Helicobacter pylori*-susceptible ethnic Malays. *World J Gastroenterol, 19*(23), 3615-3622. doi:10.3748/wjg.v19.i23.3615

Marlovits, T. C., & Stebbins, C. E. (2010). Type III secretion systems shape up as they ship out. *Curr Opin Microbiol, 13*(1), 47-52. doi:10.1016/j.mib.2009.11.001

Mashahreh, B., Hassouna, F., Soudah, N., Cohen-Kfir, E., Strulovich, R., Haitin, Y., & Wiener, R. (2018). Trans-binding of UFM1 to UBA5 stimulates UBA5 homodimerization and ATP binding. *FASEB J, 32*(5), 2794-2802. doi:10.1096/fj.201701057R

Matunis, M. J., Wu, J., & Blobel, G. (1998). SUMO-1 modification and its role in targeting the Ran GTPase-activating protein, RanGAP1, to the nuclear pore complex. *J Cell Biol, 140*(3), 499-509.

Maurelli, A. T., Baudry, B., d'Hauteville, H., Hale, T. L., & Sansonetti, P. J. (1985). Cloning of plasmid DNA sequences involved in invasion of HeLa cells by *Shigella flexneri*. *Infect Immun, 49*(1), 164-171.

Mavris, M., Page, A. L., Tournebize, R., Demers, B., Sansonetti, P., & Parsot, C. (2002). Regulation of transcription by the activity of the *Shigella flexneri* type III secretion apparatus. *Mol Microbiol, 43*(6), 1543-1553.

McDowell, G. S., & Philpott, A. (2013). Non-canonical ubiquitylation: Mechanisms and consequences. *Int J Biochem Cell Biol, 45*(8), 1833-1842. doi:10.1016/j.biocel.2013.05.026

Metzger, M. B., Pruneda, J. N., Klevit, R. E., & Weissman, A. M. (2014). RING-type E3 ligases: Master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. *Biochim Biophys Acta, 1843*(1), 47-60. doi:10.1016/j.bbamcr.2013.05.026

Mevissen, T. E. T., & Komander, D. (2017). Mechanisms of deubiquitinase specificity and regulation. *Annu Rev Biochem, 86*, 159-192. doi:10.1146/annurev-biochem-061516-044916

Meyer, H. J., & Rape, M. (2014). Enhanced protein degradation by branched ubiquitin chains. *Cell, 157*(4), 910-921. doi:10.1016/j.cell.2014.03.037

Mizushima, N. (2007). Autophagy: Process and function. *Genes Dev, 21*(22), 2861-2873. doi:10.1101/gad.1599207

Mizushima, T., Tatsumi, K., Ozaki, Y., Kawakami, T., Suzuki, A., Ogasahara, K., . . . Yamane, T. (2007). Crystal structure of Ufc1, the Ufm1-conjugating enzyme. *Biochem Biophys Res Commun, 362*(4), 1079-1084. doi:10.1016/j.bbrc.2007.08.129

Moremen, K. W., Tiemeyer, M., & Nairn, A. V. (2012). Vertebrate protein glycosylation: Diversity, synthesis and function. *Nat Rev Mol Cell Biol, 13*(7), 448-462. doi:10.1038/nrm3383

Nahorski, M. S., Maddirevula, S., Ishimura, R., Alsahli, S., Brady, A. F., Begemann, A., . . . Alkuraya, F. S. (2018). Biallelic UFM1 and UFC1 mutations expand the essential role of ufmylation in brain development. *Brain, 141*(7), 1934-1945. doi:10.1093/brain/awy135

Nans, A., Kudryashev, M., Saibil, H. R., & Hayward, R. D. (2015). Structure of a bacterial type III secretion system in contact with a host membrane in situ. *Nat Commun, 6*, 10114. doi:10.1038/ncomms10114

Neiman-Zenevich, J., Stuart, S., Abdel-Nour, M., Girardin, S. E., & Mogridge, J. (2017). *Listeria monocytogenes* and *Shigella flexneri* Activate the NLRP1B Inflammasome. *Infect Immun, 85*(11). doi:10.1128/IAI.00338-17

Nijman, S. M., Luna-Vargas, M. P., Velds, A., Brummelkamp, T. R., Dirac, A. M., Sixma, T. K., & Bernards, R. (2005). A genomic and functional inventory of deubiquitinating enzymes. *Cell, 123*(5), 773-786. doi:10.1016/j.cell.2005.11.007

Noad, J., von der Malsburg, A., Pathe, C., Michel, M. A., Komander, D., & Randow, F. (2017). LUBAC-synthesized linear ubiquitin chains restrict cytosol-invading bacteria by activating autophagy and NF-kappaB. *Nat Microbiol, 2*, 17063. doi:10.1038/nmicrobiol.2017.63

Okuda, J., Toyotome, T., Kataoka, N., Ohno, M., Abe, H., Shimura, Y., . . . Sasakawa, C. (2005). *Shigella* effector IpaH9.8 binds to a splicing factor U2AF(35) to modulate host immune responses. *Biochem Biophys Res Commun, 333*(2), 531-539. doi:10.1016/j.bbrc.2005.05.145

Oweis, W., Padala, P., Hassouna, F., Cohen-Kfir, E., Gibbs, D. R., Todd, E. A., . . . Wiener, R. (2016). Trans-binding mechanism of ubiquitin-like protein activation revealed by a UBA5-UFM1 complex. *Cell Rep, 16*(12), 3113-3120. doi:10.1016/j.celrep.2016.08.067

Ozkan, E., Yu, H., & Deisenhofer, J. (2005). Mechanistic insight into the allosteric activation of a ubiquitin-conjugating enzyme by RING-type ubiquitin ligases. *Proc Natl Acad Sci U S A, 102*(52), 18890-18895. doi:10.1073/pnas.0509418102

Parsot, C. (2009). *Shigella* type III secretion effectors: how, where, when, for what purposes? *Curr Opin Microbiol, 12*(1), 110-116. doi:10.1016/j.mib.2008.12.002

Patel, J. C., Hueffer, K., Lam, T. T., & Galan, J. E. (2009). Diversification of a *Salmonella* virulence protein function by ubiquitin-dependent differential localization. *Cell, 137*(2), 283-294. doi:10.1016/j.cell.2009.01.056

Pavri, R., Zhu, B., Li, G., Trojer, P., Mandal, S., Shilatifard, A., & Reinberg, D. (2006). Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell, 125*(4), 703-717. doi:10.1016/j.cell.2006.04.029

Pearce, M. J., Mintseris, J., Ferreyra, J., Gygi, S. P., & Darwin, K. H. (2008). Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. *Science, 322*(5904), 1104-1107. doi:10.1126/science.1163885

Perrett, C. A., Lin, D. Y., & Zhou, D. (2011). Interactions of bacterial proteins with host eukaryotic ubiquitin pathways. *Front Microbiol, 2*, 143. doi:10.3389/fmicb.2011.00143

Petroski, M. D., & Deshaies, R. J. (2005). Function and regulation of cullin-RING ubiquitin ligases. *Nat Rev Mol Cell Biol, 6*(1), 9-20. doi:10.1038/nrm1547

Phalipon, A., & Sansonetti, P. J. (2007). *Shigella*'s ways of manipulating the host intestinal innate and adaptive immune system: A tool box for survival? *Immunol Cell Biol, 85*(2), 119-129. doi:10.1038/sj.icb7100025

Piro, A. S., Hernandez, D., Luoma, S., Feeley, E. M., Finethy, R., Yirga, A., . . . Coers, J. (2017). Detection of cytosolic *Shigella flexneri* via a C-terminal triple-arginine motif of GBP1 inhibits actin-based motility. *MBio, 8*(6). doi:10.1128/mBio.01979-17

Pirone, L., Xolalpa, W., Sigurethsson, J. O., Ramirez, J., Perez, C., Gonzalez, M., . . . Sutherland, J. D. (2017). A comprehensive platform for the analysis of ubiquitin-like protein modifications using in vivo biotinylation. *Sci Rep, 7*, 40756. doi:10.1038/srep40756

Piscatelli, H., Kotkar, S. A., McBee, M. E., Muthupalani, S., Schauer, D. B., Mandrell, R. E., . . . Zhou, D. (2011). The EHEC type III effector NleL is an E3 ubiquitin ligase that modulates pedestal formation. *PLoS One, 6*(4), e19331. doi:10.1371/journal.pone.0019331

Prudden, J., Pebernard, S., Raffa, G., Slavin, D. A., Perry, J. J., Tainer, J. A., . . . Boddy, M. N. (2007). SUMO-targeted ubiquitin ligases in genome stability. *EMBO J, 26*(18), 4089-4101. doi:10.1038/sj.emboj.7601838

Ptacek, J., Devgan, G., Michaud, G., Zhu, H., Zhu, X., Fasolo, J., . . . Snyder, M. (2005). Global analysis of protein phosphorylation in yeast. *Nature, 438*(7068), 679-684. doi:10.1038/nature04187

Puzari, M., Sharma, M., & Chetia, P. (2018). Emergence of antibiotic resistant *Shigella* species: A matter of concern. *J Infect Public Health, 11*(4), 451-454. doi:10.1016/j.jiph.2017.09.025

Qiu, J., Sheedlo, M. J., Yu, K., Tan, Y., Nakayasu, E. S., Das, C., . . . Luo, Z. Q. (2016). Ubiquitination independent of E1 and E2 enzymes by bacterial effectors. *Nature, 533*(7601), 120-124. doi:10.1038/nature17657

Quezada, C. M., Hicks, S. W., Galan, J. E., & Stebbins, C. E. (2009). A family of *Salmonella* virulence factors functions as a distinct class of autoregulated E3 ubiquitin ligases. *Proc Natl Acad Sci U S A, 106*(12), 4864-4869. doi:10.1073/pnas.0811058106

Rahman, M. M., & McFadden, G. (2011). Modulation of NF-kappaB signalling by microbial pathogens. *Nat Rev Microbiol, 9*(4), 291-306. doi:10.1038/nrmicro2539

Rape, M. (2018). Ubiquitylation at the crossroads of development and disease. *Nat Rev Mol Cell Biol, 19*(1), 59-70. doi:10.1038/nrm.2017.83

Rape, M., Reddy, S. K., & Kirschner, M. W. (2006). The processivity of multiubiquitination by the APC determines the order of substrate degradation. *Cell, 124*(1), 89-103. doi:10.1016/j.cell.2005.10.032

Read, M. A., Brownell, J. E., Gladysheva, T. B., Hottelet, M., Parent, L. A., Coggins, M. B., . . . Palombella, V. J. (2000). Nedd8 modification of cul-1 activates SCF(beta(TrCP))-dependent ubiquitination of IkappaBalpha. *Mol Cell Biol, 20*(7), 2326-2333.

Redman, K. L., & Rechsteiner, M. (1989). Identification of the long ubiquitin extension as ribosomal protein S27a. *Nature, 338*(6214), 438-440. doi:10.1038/338438a0

Reyes-Turcu, F. E., Ventii, K. H., & Wilkinson, K. D. (2009). Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. *Annu Rev Biochem, 78*, 363-397. doi:10.1146/annurev.biochem.78.082307.091526

Rohde, J. R., Breitkreutz, A., Chenal, A., Sansonetti, P. J., & Parsot, C. (2007). Type III secretion effectors of the IpaH family are E3 ubiquitin ligases. *Cell Host Microbe, 1*(1), 77-83. doi:10.1016/j.chom.2007.02.002

Rosebrock, T. R., Zeng, L., Brady, J. J., Abramovitch, R. B., Xiao, F., & Martin, G. B. (2007). A bacterial E3 ubiquitin ligase targets a host protein kinase to disrupt plant immunity. *Nature, 448*(7151), 370-374. doi:10.1038/nature05966

Rotin, D., & Kumar, S. (2009). Physiological functions of the HECT family of ubiquitin ligases. *Nat Rev Mol Cell Biol, 10*(6), 398-409. doi:10.1038/nrm2690

Rubio, M. D., Wood, K., Haroutunian, V., & Meador-Woodruff, J. H. (2013). Dysfunction of the ubiquitin proteasome and ubiquitin-like systems in schizophrenia. *Neuropsychopharmacology, 38*(10), 1910-1920. doi:10.1038/npp.2013.84

Saftig, P., & Haas, A. (2016). Turn up the lysosome. *Nat Cell Biol, 18*(10), 1025-1027. doi:10.1038/ncb3409

Sambrook, J., & Russell, D. W. (2001). *Molecular cloning : a laboratory manual* (3rd ed.). Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Sanada, T., Kim, M., Mimuro, H., Suzuki, M., Ogawa, M., Oyama, A., . . . Sasakawa, C. (2012). The *Shigella flexneri* effector OspI deamidates UBC13 to dampen the inflammatory response. *Nature, 483*(7391), 623-626. doi:10.1038/nature10894

Saussez, S., & Kiss, R. (2006). Galectin-7. *Cell Mol Life Sci, 63*(6), 686-697. doi:10.1007/s00018-005-5458-8

Schafer, A., Kuhn, M., & Schindelin, H. (2014). Structure of the ubiquitin-activating enzyme loaded with two ubiquitin molecules. *Acta Crystallogr D Biol Crystallogr, 70*(Pt 5), 1311-1320. doi:10.1107/S1399004714002910

Scheffner, M., Nuber, U., & Huibregtse, J. M. (1995). Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. *Nature, 373*(6509), 81-83. doi:10.1038/373081a0

Schulman, B. A., & Harper, J. W. (2009). Ubiquitin-like protein activation by E1 enzymes: The apex for downstream signalling pathways. *Nat Rev Mol Cell Biol, 10*(5), 319-331. doi:10.1038/nrm2673

Settembre, C., Di Malta, C., Polito, V. A., Garcia Arencibia, M., Vetrini, F., Erdin, S., . . . Ballabio, A. (2011). TFEB links autophagy to lysosomal biogenesis. *Science, 332*(6036), 1429-1433. doi:10.1126/science.1204592

Settembre, C., Zoncu, R., Medina, D. L., Vetrini, F., Erdin, S., Erdin, S., . . . Ballabio, A. (2012). A lysosome-to-nucleus signalling mechanism senses and regulates the lysosome via mTOR and TFEB. *EMBO J, 31*(5), 1095-1108. doi:10.1038/emboj.2012.32

Shaid, S., Brandts, C. H., Serve, H., & Dikic, I. (2013). Ubiquitination and selective autophagy. *Cell Death Differ, 20*(1), 21-30. doi:10.1038/cdd.2012.72

Simsek, D., Tiu, G. C., Flynn, R. A., Byeon, G. W., Leppek, K., Xu, A. F., . . . Barna, M. (2017). The mammalian ribo-interactome reveals ribosome functional diversity and heterogeneity. *Cell, 169*(6), 1051-1065 e1018. doi:10.1016/j.cell.2017.05.022

Singer, A. U., Rohde, J. R., Lam, R., Skarina, T., Kagan, O., Dileo, R., . . . Savchenko, A. (2008). Structure of the *Shigella* T3SS effector IpaH defines a new class of E3 ubiquitin ligases. *Nat Struct Mol Biol, 15*(12), 1293-1301. doi:10.1038/nsmb.1511

Singh, R. K., Zerath, S., Kleifeld, O., Scheffner, M., Glickman, M. H., & Fushman, D. (2012). Recognition and cleavage of related to ubiquitin 1 (Rub1) and Rub1-ubiquitin chains by components of the ubiquitin-proteasome system. *Mol Cell Proteomics, 11*(12), 1595-1611. doi:10.1074/mcp.M112.022467

Skowyra, D., Craig, K. L., Tyers, M., Elledge, S. J., & Harper, J. W. (1997). F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. *Cell, 91*(2), 209-219.

Sluimer, J., & Distel, B. (2018). Regulating the human HECT E3 ligases. *Cell Mol Life Sci, 75*(17), 3121-3141. doi:10.1007/s00018-018-2848-2

Soudah, N., Padala, P., Hassouna, F., Kumar, M., Mashahreh, B., Lebedev, A. A., . . . Wiener, R. (2018). An N-terminal extension to UBA5 adenylation domain boosts UFM1 activation: Isoform-specific differences in ubiquitin-like protein activation. *J Mol Biol*. doi:10.1016/j.jmb.2018.10.007

Stephen, A. G., Trausch-Azar, J. S., Ciechanover, A., & Schwartz, A. L. (1996). The ubiquitin-activating enzyme E1 is phosphorylated and localized to the nucleus in a cell cycle-dependent manner. *J Biol Chem, 271*(26), 15608-15614.

Sun, F., Ding, Y., Ji, Q., Liang, Z., Deng, X., Wong, C. C., . . . He, C. (2012). Protein cysteine phosphorylation of SarA/MgrA family transcriptional regulators mediates bacterial virulence and antibiotic resistance. *Proc Natl Acad Sci U S A, 109*(38), 15461-15466. doi:10.1073/pnas.1205952109

Suzuki, S., Mimuro, H., Kim, M., Ogawa, M., Ashida, H., Toyotome, T., . . . Sasakawa, C. (2014). *Shigella* IpaH7.8 E3 ubiquitin ligase targets glomulin and activates inflammasomes to demolish macrophages. *Proc Natl Acad Sci U S A, 111*(40), E4254-4263. doi:10.1073/pnas.1324021111

Swatek, K. N., & Komander, D. (2016). Ubiquitin modifications. *Cell Res, 26*(4), 399-422. doi:10.1038/cr.2016.39

Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., . . . von Mering, C. (2017). The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res, 45*(D1), D362-D368. doi:10.1093/nar/gkw937

Tait, S. W., de Vries, E., Maas, C., Keller, A. M., D'Santos, C. S., & Borst, J. (2007). Apoptosis induction by Bid requires unconventional ubiquitination and degradation of its N-terminal fragment. *J Cell Biol, 179*(7), 1453-1466. doi:10.1083/jcb.200707063

Takeuchi, O., & Akira, S. (2010). Pattern recognition receptors and inflammation. *Cell, 140*(6), 805-820. doi:10.1016/j.cell.2010.01.022

Tanner, K. G. (2014). *Shigella effector IpaH9.8 interacts with autophagy transcription factor ZKSCAN3 and increases autophagy during infection*. Master's Thesis. Dalhousie University.

Tatham, M. H., Geoffroy, M. C., Shen, L., Plechanovova, A., Hattersley, N., Jaffray, E. G., . . . Hay, R. T. (2008). RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nat Cell Biol, 10*(5), 538-546. doi:10.1038/ncb1716

Tatsumi, K., Sou, Y. S., Tada, N., Nakamura, E., Iemura, S., Natsume, T., . . . Komatsu, M. (2010). A novel type of E3 ligase for the Ufm1 conjugation system. *J Biol Chem, 285*(8), 5417-5427. doi:10.1074/jbc.M109.036814

Tatsumi, K., Yamamoto-Mukai, H., Shimizu, R., Waguri, S., Sou, Y. S., Sakamoto, A., . . . Komatsu, M. (2011). The Ufm1-activating enzyme Uba5 is indispensable for erythroid differentiation in mice. *Nat Commun, 2*, 181. doi:10.1038/ncomms1182

Tattoli, I., Sorbara, M. T., Vuckovic, D., Ling, A., Soares, F., Carneiro, L. A., . . . Girardin, S. E. (2012). Amino acid starvation induced by invasive bacterial pathogens triggers an innate host defense program. *Cell Host Microbe, 11*(6), 563-575. doi:10.1016/j.chom.2012.04.012

Thoreen, C. C., Kang, S. A., Chang, J. W., Liu, Q., Zhang, J., Gao, Y., . . . Gray, N. S. (2009). An ATP-competitive mammalian target of rapamycin inhibitor reveals rapamycin-resistant functions of mTORC1. *J Biol Chem, 284*(12), 8023-8032. doi:10.1074/jbc.M900301200

Thrower, J. S., Hoffman, L., Rechsteiner, M., & Pickart, C. M. (2000). Recognition of the polyubiquitin proteolytic signal. *EMBO J, 19*(1), 94-102. doi:10.1093/emboj/19.1.94

Thurston, T. L., Wandel, M. P., von Muhlinen, N., Foeglein, A., & Randow, F. (2012). Galectin 8 targets damaged vesicles for autophagy to defend cells against bacterial invasion. *Nature, 482*(7385), 414-418. doi:10.1038/nature10744

Tokunaga, F., Sakata, S., Saeki, Y., Satomi, Y., Kirisako, T., Kamei, K., . . . Iwai, K. (2009). Involvement of linear polyubiquitylation of NEMO in NF-kappaB activation. *Nat Cell Biol, 11*(2), 123-132. doi:10.1038/ncb1821

Toyotome, T., Suzuki, T., Kuwae, A., Nonaka, T., Fukuda, H., Imajoh-Ohmi, S., . . . Sasakawa, C. (2001). *Shigella* protein IpaH(9.8) is secreted from bacteria within mammalian cells and transported to the nucleus. *J Biol Chem, 276*(34), 32071-32079. doi:10.1074/jbc.M101882200

Travassos, L. H., Carneiro, L. A., Ramjeet, M., Hussey, S., Kim, Y. G., Magalhaes, J. G., . . . Philpott, D. J. (2010). Nod1 and Nod2 direct autophagy by recruiting ATG16L1 to the plasma membrane at the site of bacterial entry. *Nat Immunol, 11*(1), 55-62. doi:10.1038/ni.1823

Treier, M., Staszewski, L. M., & Bohmann, D. (1994). Ubiquitin-dependent c-Jun degradation in vivo is mediated by the delta domain. *Cell, 78*(5), 787-798.

Ubersax, J. A., & Ferrell, J. E., Jr. (2007). Mechanisms of specificity in protein phosphorylation. *Nat Rev Mol Cell Biol, 8*(7), 530-541. doi:10.1038/nrm2203

Urra, H., Henriquez, D. R., Canovas, J., Villarroel-Campos, D., Carreras-Sureda, A., Pulgar, E., . . . Hetz, C. (2018). IRE1alpha governs cytoskeleton remodelling and cell migration through a direct interaction with filamin A. *Nat Cell Biol, 20*(8), 942-953. doi:10.1038/s41556-018-0141-0

van der Veen, A. G., & Ploegh, H. L. (2012). Ubiquitin-like proteins. *Annu Rev Biochem, 81*, 323-357. doi:10.1146/annurev-biochem-093010-153308

Varshavsky, A. (1996). The N-end rule: Functions, mysteries, uses. *Proc Natl Acad Sci U S A, 93*(22), 12142-12149.

Venkatesan, M. M., Buysse, J. M., & Hartman, A. B. (1991). Sequence variation in two ipaH genes of *Shigella flexneri* 5 and homology to the LRG-like family of proteins. *Mol Microbiol, 5*(10), 2435-2445.

Verdecia, M. A., Joazeiro, C. A., Wells, N. J., Ferrer, J. L., Bowman, M. E., Hunter, T., & Noel, J. P. (2003). Conformational flexibility underlies ubiquitin ligation mediated by the WWP1 HECT domain E3 ligase. *Mol Cell, 11*(1), 249-259.

Vijay-Kumar, S., Bugg, C. E., & Cook, W. J. (1987). Structure of ubiquitin refined at 1.8 A resolution. *J Mol Biol, 194*(3), 531-544.

von Muhlinen, N., Thurston, T., Ryzhakov, G., Bloor, S., & Randow, F. (2010). NDP52, a novel autophagy receptor for ubiquitin-decorated cytosolic bacteria. *Autophagy, 6*(2), 288-289.

von Seidlein, L., Kim, D. R., Ali, M., Lee, H., Wang, X., Thiem, V. D., . . . Clemens, J. (2006). A multicentre study of *Shigella* diarrhoea in six Asian countries: Disease burden, clinical manifestations, and microbiology. *PLoS Med, 3*(9), e353. doi:10.1371/journal.pmed.0030353

Vosper, J. M., McDowell, G. S., Hindley, C. J., Fiore-Heriche, C. S., Kucerova, R., Horan, I., & Philpott, A. (2009). Ubiquitylation on canonical and non-canonical sites targets the transcription factor neurogenin for ubiquitin-mediated proteolysis. *J Biol Chem, 284*(23), 15458-15468. doi:10.1074/jbc.M809366200

Wandel, M. P., Pathe, C., Werner, E. I., Ellison, C. J., Boyle, K. B., von der Malsburg, A., . . . Randow, F. (2017). GBPs inhibit motility of *Shigella flexneri* but are targeted for degradation by the bacterial ubiquitin ligase IpaH9.8. *Cell Host Microbe, 22*(4), 507-518 e505. doi:10.1016/j.chom.2017.09.007

Wang, F., Jiang, Z., Li, Y., He, X., Zhao, J., Yang, X., . . . Yuan, J. (2013). *Shigella flexneri* T3SS effector IpaH4.5 modulates the host inflammatory response via interaction with NF-kappaB p65 protein. *Cell Microbiol, 15*(3), 474-485. doi:10.1111/cmi.12052

Wassef, J. S., Keren, D. F., & Mailloux, J. L. (1989). Role of M cells in initial antigen uptake and in ulcer formation in the rabbit intestinal loop model of shigellosis. *Infect Immun, 57*(3), 858-863.

Watson, C. M., Crinnion, L. A., Gleghorn, L., Newman, W. G., Ramesar, R., Beighton, P., & Wallis, G. A. (2015). Identification of a mutation in the ubiquitin-fold modifier 1-specific peptidase 2 gene, UFSP2, in an extended South African family with Beukes hip dysplasia. *S Afr Med J, 105*(7), 558-563. doi:10.7196/SAMJnew.7917

Wei, C., Wang, Y., Du, Z., Guan, K., Cao, Y., Yang, H., . . . He, X. (2016). The Yersinia Type III secretion effector YopM Is an E3 ubiquitin ligase that induced necrotic cell death by targeting NLRP3. *Cell Death Dis, 7*(12), e2519. doi:10.1038/cddis.2016.413

Welch, M. D., & Way, M. (2013). Arp2/3-mediated actin-based motility: A tail of pathogen abuse. *Cell Host Microbe, 14*(3), 242-255. doi:10.1016/j.chom.2013.08.011

Wenzel, D. M., Lissounov, A., Brzovic, P. S., & Klevit, R. E. (2011). UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids. *Nature, 474*(7349), 105-108. doi:10.1038/nature09966

WHO. (2006). *State of the art of new vaccine research and development*. Geneva, Switzerland: World Health Organization.

Wiborg, O., Pedersen, M. S., Wind, A., Berglund, L. E., Marcker, K. A., & Vuust, J. (1985). The human ubiquitin multigene family: Some genes contain multiple directly repeated ubiquitin coding sequences. *EMBO J, 4*(3), 755-759.

Wiesner, S., Ogunjimi, A. A., Wang, H. R., Rotin, D., Sicheri, F., Wrana, J. L., & Forman-Kay, J. D. (2007). Autoinhibition of the HECT-type ubiquitin ligase Smurf2 through its C2 domain. *Cell, 130*(4), 651-662. doi:10.1016/j.cell.2007.06.050

Williamson, A., Banerjee, S., Zhu, X., Philipp, I., Iavarone, A. T., & Rape, M. (2011). Regulation of ubiquitin chain initiation to control the timing of substrate degradation. *Mol Cell, 42*(6), 744-757. doi:10.1016/j.molcel.2011.04.022

Windheim, M., Stafford, M., Peggie, M., & Cohen, P. (2008). Interleukin-1 (IL-1) induces the Lys63-linked polyubiquitination of IL-1 receptor-associated kinase 1 to facilitate NEMO binding and the activation of IkappaBalpha kinase. *Mol Cell Biol, 28*(5), 1783-1791. doi:10.1128/MCB.02380-06

Wu, B., Skarina, T., Yee, A., Jobin, M. C., Dileo, R., Semesi, A., . . . Savchenko, A. (2010). NleG Type 3 effectors from enterohaemorrhagic *Escherichia coli* are U-Box E3 ubiquitin ligases. *PLoS Pathog, 6*(6), e1000960. doi:10.1371/journal.ppat.1000960

Wu, J., Lei, G., Mei, M., Tang, Y., & Li, H. (2010). A novel C53/LZAP-interacting protein regulates stability of C53/LZAP and DDRGK domain-containing protein 1 (DDRGK1) and modulates NF-kappaB signaling. *J Biol Chem, 285*(20), 15126-15136. doi:10.1074/jbc.M110.110619

Xi, P., Ding, D., Zhou, J., Wang, M., & Cong, Y. S. (2013). DDRGK1 regulates NF-kappaB activity by modulating IkappaBalpha stability. *PLoS One, 8*(5), e64231. doi:10.1371/journal.pone.0064231

Xin, D. W., Liao, S., Xie, Z. P., Hann, D. R., Steinle, L., Boller, T., & Staehelin, C. (2012). Functional analysis of NopM, a novel E3 ubiquitin ligase (NEL) domain effector of *Rhizobium* sp. strain NGR234. *PLoS Pathog, 8*(5), e1002707. doi:10.1371/journal.ppat.1002707

Xu, P., Duong, D. M., Seyfried, N. T., Cheng, D., Xie, Y., Robert, J., . . . Peng, J. (2009). Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. *Cell, 137*(1), 133-145. doi:10.1016/j.cell.2009.01.041

Yaglom, J., Linskens, M. H., Sadis, S., Rubin, D. M., Futcher, B., & Finley, D. (1995). p34Cdc28-mediated control of Cln3 cyclin degradation. *Mol Cell Biol, 15*(2), 731-741.

Ye, Y., & Rape, M. (2009). Building ubiquitin chains: E2 enzymes at work. *Nat Rev Mol Cell Biol, 10*(11), 755-764. doi:10.1038/nrm2780

Yoo, H. M., Kang, S. H., Kim, J. Y., Lee, J. E., Seong, M. W., Lee, S. W., . . . Chung, C. H. (2014). Modification of ASC1 by UFM1 is crucial for ERalpha transactivation and breast cancer development. *Mol Cell, 56*(2), 261-274. doi:10.1016/j.molcel.2014.08.007

Yu, H. B., Croxen, M. A., Marchiando, A. M., Ferreira, R. B., Cadwell, K., Foster, L. J., & Finlay, B. B. (2014). Autophagy facilitates Salmonella replication in HeLa cells. *MBio, 5*(2), e00865-00814. doi:10.1128/mBio.00865-14

Zhang, J., Wang, J., Zhou, Z., Park, J. E., Wang, L., Wu, S., . . . Shen, H. M. (2018). Importance of TFEB acetylation in control of its transcriptional activity and lysosomal function in response to histone deacetylase inhibitors. *Autophagy, 14*(6), 1043-1059. doi:10.1080/15548627.2018.1447290

Zhang, M., Zhu, X., Zhang, Y., Cai, Y., Chen, J., Sivaprakasam, S., . . . Li, H. (2015). RCAD/Ufl1, a Ufm1 E3 ligase, is essential for hematopoietic stem cell function and murine hematopoiesis. *Cell Death Differ, 22*(12), 1922-1934. doi:10.1038/cdd.2015.51

Zhang, Y., Higashide, W. M., McCormick, B. A., Chen, J., & Zhou, D. (2006). The inflammation-associated *Salmonella* SopA is a HECT-like E3 ubiquitin ligase. *Mol Microbiol, 62*(3), 786-793. doi:10.1111/j.1365-2958.2006.05407.x

Zhang, Y., Zhang, M., Wu, J., Lei, G., & Li, H. (2012). Transcriptional regulation of the Ufm1 conjugation system in response to disturbance of the endoplasmic reticulum homeostasis and inhibition of vesicle trafficking. *PLoS One, 7*(11), e48587. doi:10.1371/journal.pone.0048587

Zheng, M., Gu, X., Zheng, D., Yang, Z., Li, F., Zhao, J., . . . Mao, Y. (2008). UBE1DC1, an ubiquitin-activating enzyme, activates two different ubiquitin-like proteins. *J Cell Biochem, 104*(6), 2324-2334. doi:10.1002/jcb.21791

Zheng, Z., Wei, C., Guan, K., Yuan, Y., Zhang, Y., Ma, S., . . . He, X. (2016). Bacterial E3 ubiquitin ligase IpaH4.5 of *Shigella flexneri* targets TBK1 to dampen the host antibacterial response. *J Immunol, 196*(3), 1199-1208. doi:10.4049/jimmunol.1501045

Zhu, Y., Lei, Q., Li, D., Zhang, Y., Jiang, X., Hu, Z., & Xu, G. (2018). Proteomic and biochemical analyses reveal a novel mechanism for promoting protein ubiquitination and degradation by UFBP1, a key component of ufmylation. *J Proteome Res, 17*(4), 1509-1520. doi:10.1021/acs.jproteome.7b00843

Zhu, Y., Li, H., Hu, L., Wang, J., Zhou, Y., Pang, Z., . . . Shao, F. (2008). Structure of a *Shigella* effector reveals a new class of ubiquitin ligases. *Nat Struct Mol Biol, 15*(12), 1302-1308. doi:10.1038/nsmb.1517

Appendix A: IpaH9.8 Protein-Protein Interaction Results

Top 50 mammalian proteins shown to interact with IpaH9.8 by yeast two-hybrid protein-protein interaction screen, order the number of clones detected. Data collected by Jeremy Benjamin.

Frequency	Gene	Description
23	ATP1B3	ATPase, Na+/K+ transporting, beta 3 polypeptide
11	ARHGDIB	Rho GDP dissociation inhibitor (GDI) beta
7	KRT222	keratin 222
7	CCT3	chaperonin containing TCP1, subunit 3 (gamma)
6	ATP1B1	ATPase, Na+/K+ transporting, beta 1 polypeptide
6	MRVI1	murine retrovirus integration site 1 homolog
5	PCNA	proliferating cell nuclear antigen
5	**ZKSCAN3**	**zinc finger with KRAB and SCAN domains 3**
4	C17orf46	chromosome 17 open reading frame 46
4	PGM1	phosphoglucomutase 1
3	C9orf78	chromosome 9 open reading frame 78
3	GBP4	guanylate binding protein 4
3	EIF1B	eukaryotic translation initiation factor 1B
3	MAP3K1	mitogen-activated protein kinase kinase kinase 1
2	TMED7	transmembrane emp24 protein transport domain containing 7
2	MPHOSPH8	M-phase phosphoprotein 8
2	MYBPC1	myosin binding protein C, slow type
2	RPS20	ribosomal protein S20
2	ZBTB38	zinc finger and BTB domain containing 38
1	LOC284412	hypothetical LOC284412, non-coding RNA
1	PRDX6	peroxiredoxin 6
1	AP1G1	adaptor-related protein complex 1, gamma 1 subunit
1	C14orf45	chromosome 14 open reading frame 45
1	COPS5	COP9 constitutive photomorphogenic homolog subunit 5
1	CUL1	cullin 1
1	MRPS9	mitochondrial ribosomal protein S9
1	N4BP2L2	NEDD4 binding protein 2-like 2
1	ZNF350	zinc finger protein 350
1	ANKRD7	ankyrin repeat domain 7
1	FMO2	flavin containing monooxygenase 2 (non-functional)
1	PEAK1	NKF3 kinase family member
1	PNISR	PNN-interacting serine/arginine-rich protein
1	SV2B	synaptic vesicle glycoprotein 2B
1	intergenic	chromosome 3 genomic contig

Frequency	Gene	Description
1	ANAPC1	anaphase promoting complex subunit 1
1	ATP2C1	ATPase, Ca++ transporting, type 2C, member 1
1	CD109	CD109 molecule
1	COMMD8	COMM domain containing 8
1	F13A1	coagulation factor XIII, A1 polypeptide
1	OLFML3	olfactomedin-like 3
1	SSR1	signal sequence receptor, alpha
1	**UFM1**	**ubiquitin-fold modifier 1**
1	VDAC2	voltage-dependent anion-selective channel protein 2
1	WWOX	WW domain containing oxidoreductase
1	ZBRK1	zinc-finger protein ZBRK1
1	ZNF177	zinc finger protein 177
1	ZNF237	zinc finger protein 237
1	ZNF251	zinc finger protein 251