BROAD-SPECTRUM PROTEOME EDITING WITH AN ENGINEERED BACTERIAL UBIQUITIN LIGASE MIMIC

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Morgan Baltz Ludwicki

August 2019

BROAD-SPECTRUM PROTEOME EDITING WITH AN ENGINEERED BACTERIAL UBIQUITIN LIGASE MIMIC

Morgan Baltz Ludwicki, Ph. D.

Cornell University 2019

Manipulation of the ubiquitin-proteasome pathway to achieve targeted silencing of cellular proteins has emerged as a reliable and customizable strategy for remodeling the mammalian proteome. One such approach involves engineering bifunctional proteins called ubiquibodies that are comprised of a synthetic binding protein fused to an E3 ubiquitin ligase, thus enabling post-translational ubiquitination and degradation of a target protein. Here, we have designed a panel of new ubiquibodies based on E3 ubiquitin ligase mimics from bacterial pathogens that are capable of effectively interfacing with the mammalian proteasomal degradation machinery for selective removal of proteins of interest. One of these, the *Shigella flexneri* effector protein IpaH9.8 fused to a fibronectin type III (FN3) monobody that specifically recognizes green fluorescent protein (GFP), was observed to potently eliminate GFP and its spectral derivatives as well as 15 different FP-tagged mammalian proteins that varied in size (27–179 kDa) and subcellular localization (cytoplasm, nucleus, membrane-associated, and transmembrane). We further demonstrated the modularity and flexibility of IpaH9.8 by redirecting its activity towards the disease relevant proteins SHP2, KRas, and ERK2 through the use of binding domains identified in literature as well as novel binding domains we isolated using yeast surface display.

To demonstrate therapeutically relevant delivery of ubiquibodies, we investigated two approaches: endowment of ubiquibodies with a cell-penetrating peptide

domain and packaging of mRNA encoding the GFP-specific ubiquibody in bioinspired nano-sized complexes. The resulting nanoplexes delivered ubiquibody mRNA in a manner that caused efficient target depletion in cultured mammalian cells stably expressing GFP as well as in transgenic mice expressing GFP ubiquitously. Lastly, we discuss preliminary designs for controlling ubiquibody behavior with spatial and temporal control. Overall, our results suggest that IpaH9.8-based ubiquibodies are a highly modular proteome editing technology with the potential for pharmacologically modulating disease-causing proteins.

BIOGRAPHICAL SKETCH

Morgan Baltz Ludwicki was born and raised in Mt. Juliet, Tennessee to Larry and Claire Baltz. She attended Lakeview Elementary School where her mother was her 5th grade science teacher. She then attended Holy Rosary Academy for the 6th through 8th grades where she was encouraged to explore her passion and aptitude for science and math by Mrs. Reynolds and Mrs. Benim. She then attended Mt. Juliet High School, graduating as valedictorian in 2008. Morgan then attended the University of Tennessee, Knoxville as a Haslam Scholar. There, she was fortunate to obtain an undergraduate research position in the lab of Dr. Eric T. Boder under the guidance of graduate student Vincent Price. Upon graduation from UT-Knoxville with a B.S. in Chemical and Biomolecular Engineering in 2013, Morgan began her graduate degree at Cornell University where she joined the lab of Dr. Matthew P. DeLisa. During her time at Cornell, Morgan was awarded a Presidential Life Sciences Fellowship, NSF Graduate Research Fellowship, and Samuel C. Fleming Family Graduate Fellowship. At Cornell, Morgan also met and married her husband, Jonathan Ludwicki.

This work is dedicated to

my father, whose memory inspires,

my mother, whose words encourage,

my husband, whose arms have become my home,

ACKNOWLEDGMENTS

First, I must thank my advisor Matt DeLisa for his unwavering confidence in both my research and myself. The members of the DeLisa research group have also been an invaluable source of support, knowledge, and laughs. Additionally, I must also thank my first research advisor, Eric Boder for instilling in me a love of research, well-designed experiments, and precise storytelling. I am thankful for the true and brutally honest friendships that I have with my incoming PhD cohort., but in particular – Emily, Dana, Arna, Laura, and Carolyn thank you for being my real-life inspiration for the diversity, strength, and supportiveness of women in science. Lastly, I would like to thank Maya Angelou for providing me with the manta, "I did then what I knew how to do. Now that I know better, I do better."

This work was supported by the National Science Foundation Grant CBET-0449080 (to M.P.D.), the National Institutes of Health Grant Numbers CA132223 (to M.P.D.), CA170820 and GM083898 (to R.W.R.), CA194864 and CA212608 (to S.K.) the New York State Office of Science, Technology and Academic Research Distinguished Faculty Award (to M.P.D.), the Cornell Technology Acceleration and Maturation (CTAM) Fund, the National Science Foundation Graduate Research Fellowship Program Grants DGE-1650441 (to M.B.L.), Cornell Presidential Life Science Fellowships (to M.B.L.), a Cornell Fleming Graduate Scholarship (to M.B.L).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

A, Ala	Alanine
Abs	Absorbance
AF488	AlexaFluor-488
Amp^R	Ampicillin resistance
ATCC	American Type Culture Collection
ATP	Adenosine tri-phosphate
βMe	Beta-mercaptoethanol
BSA	Bovine serum albumin
CPP	Cell penetrating peptide, cell permeable peptide
C, Cys	Cysteine
CHIP	Carboxyl terminus of Hsc70-interacting protein
Cm^R	Chloramphenicol resistant
CMV	Cytomegalovirus
CO_2	Carbon dioxide
CRL	Cullin-ring e3 ubiquitin ligase
C-terminal	Carboxyl terminus
Da	Dalton
DARPin	Designed ankyrin repeat protein
DBP	Designer binding protein
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide
DTT	DL-Dithiothreitol
E1	Ubiquitin-activating enzyme
E2	Ubiquitin-conjugating enzyme

E3	E3 ubiquitin ligase
E. coli	*Escherichia coli*
ECV	Elongin B/C, Cullen-2, VHL
EGFP	Enhanced green fluorescent protein
ELISA	Enzyme=linked immunosorbent assay
FBS	Fetal bovine serum
FN3	Fibronectin type III domain
FP	Fluorescent protein
HECT	Homologous to EGAP C terminus
HEK293T	Human Embryonic Kidney 293 cells with SV40 large T antigen
HRP	Horseradish peroxidase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K, Lys	Lysine
K_D	Equilibrium dissociation constant
Kan^R	Kanamycin resistant
LP	Lysosomal pathway
LRR	Leucine-rich repeat
MFI	Mean fluorescence intensity
mRNA	Messenger ribonucleic acid
NEL	Novel E3 Ligase
Ni-NTA	Nickel nitrilotriacetic acid
NP40	Nonidet P-40
N-terminal	Amine terminus
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction

RPM	Rotations per minute
S. cerevisiae	*Saccharomyces cerevisiae*
SCF	Skp1/Cdc53 or Cullen-1/F-box protein
scFv	Single-chain Fv (variable fragment)
SD	Standard deviation from the mean
SDS	Sodium dodecyl sulfate
$Strep^R$	Streptomycin resistant
TPR	Tetratricopeptide repeat
TrisHCL	Tris(hydroxymethyl)aminomethane hydrochloride
uAb	Ubiquibody
Ub	Ubiquitin
UPP	Ubiquitin proteasome pathway
VHL	von Hippel-Lindau

CHAPTER 1 – INTRODUCTION

Natural protein degradation mechanisms

In response to different environmental cues or as part of their developmental cycle, all cells change their cytosolic protein content by modifying the balance between rates of protein synthesis and degradation [1, 2]. This feat is accomplished by an integrated surveillance system comprised of chaperones and protein degradation machineries that are required to maintain protein homeostasis (proteostasis). Two of the major quality control pathways responsible for regulating proteostasis are the ubiquitin proteasome pathway (UPP) and the lysosomal pathway (LP). In eukaryotic cells, the UPP is the primary mechanism through which proteins are tagged and degraded by the 26S proteasome [3, 4]. The UPP system is highly conserved in eukaryotes and consists of an enzymatic cascade involving, at a minimum, a ubiquitin-activating enzyme (E1), a ubiquitin-conjugating enzyme (E2), and a ubiquitin ligase (E3) (**Fig. 1.1**). In brief, an E1 is activated by the addition of a ubiquitin molecule through an ATP-dependent thioester bond between the carboxyl-terminal glycine residue in ubiquitin and a cysteine residue in E1. Next, the activated ubiquitin molecule is transferred to the active site cysteine residue of an E2. Finally, an E3 catalyzes the transfer of the ubiquitin molecule to the e-amino group of a lysine residue in the target protein. The attached ubiquitin molecule serves as an acceptor for further cycles of ubiquitination, generating a polyubiquitin chain. Each subsequent ubiquitin molecule can be attached to any of the seven lysine residues in the preceding ubiquitin, creating a diverse variety of polyubiquitin chain topologies that determine different protein fates [5] and give rise to the concept of a ubiquitin code [6, 7]. For example, polyubiquitin chains linked through Lys48 of ubiquitin are a signal that is

recognized by the proteasome, which proteolytically cleaves the tagged protein and yields intact ubiquitin molecules and short (7 to 8-residue) peptides. In contrast, chains linked through ubiquitin Lys63 serve as signals for the NF-kB activation [8] and DNA repair [9] pathways.

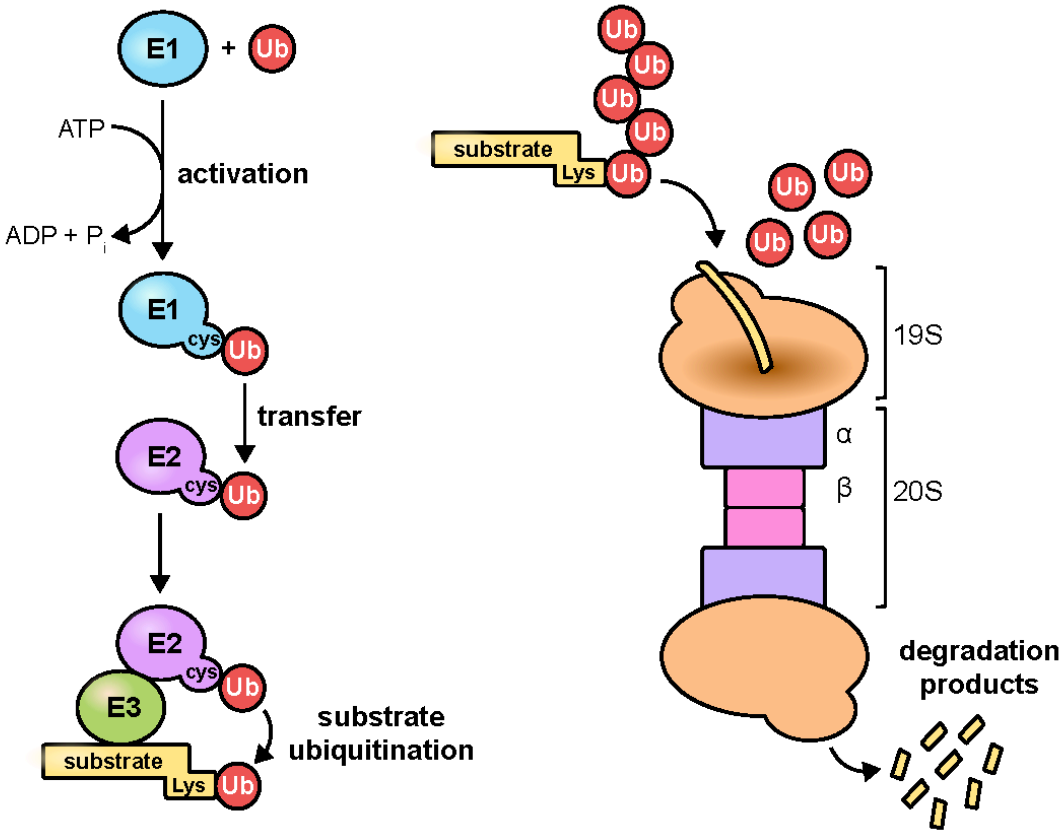


Figure 1.1 Ubiquitin proteasome pathway.
The ubiquitin proteasome pathway (UPP) is one of the major quality control pathways responsible for regulating proteostasis in mammalian cells. In the UPP, a cascade of enzymatic reactions for transferring ubiquitin (Ub) to target proteins is carried out by three enzymes: a ubiquitin-activating enzyme (E1), a ubiquitin-conjugating enzyme (E2) and a ubiquitin ligase (E3). The polyubiquitination of a protein directs it to the 26S proteasome, a 2.5-MDa molecular machine built from approximately 31 different subunits, which catalyzes protein degradation. Figure modified from Lopez-Barbosa et al., 2019.

Engineered protein knockdown techniques

Delineation of protein function has traditionally been investigated at the pre-translational level through genetic strategies that typically involve disrupting genomic DNA or mRNA transcripts, leading to complete or partial elimination of the gene product. Such loss-of-function approaches include clustered regularly interspaced short palindromic repeat (CRISPR)-Cas systems [10], zinc finger nucleases (ZFNs) [11], transcription activator-like effector nucleases [12], RNA interference (RNAi) [13], and antisense oligonucleotides (AONs).[14] Although these methods have become commonplace in basic research and are showing promise in the treatment of disease [15-20], gene silencing and genomic editing strategies are limited by a number of challenges including lack of temporal control and unpredictable off-target effects. Other issues include the inability to remove essential genes in the case of genome editing and the inability to decrease levels of proteins already present within cells, thereby leaving stable, long-lived proteins unaffected in the case of gene silencing. Also, due to the genetic basis of these strategies, protein knockout cannot be controlled through dosing or reversed easily [21], which is an important consideration for pharmacological applications.

Inhibition by binding via designer binding proteins

For post-translational targeting, small molecules remain the workhorses in the laboratory and the clinic owing to their proven ability to achieve occupancy-based inhibition with high precision. However, despite many successes, organic synthesis is still a rate-limiting factor in drug discovery projects [22], Indeed, the identification of lead candidates often requires lengthy campaigns involving the creation and screening of small-molecule libraries, with initial hits often lacking the specificity and/or affinity necessary

for efficient squelching of protein activity. Another challenge is that of the ~30,000 genes in the human genome, only ~10% encode proteins that are considered suitable for pharmacological manipulation via classical rule-of-five-compliant small-molecule drugs, leaving a large proportion of the proteome 'undruggable' [23-25].

Engineered antibodies produced from recombinant DNA technology such as single-chain Fv (scFv) antibody fragments can be used to achieve inhibition-by-binding and are especially attractive alternatives when the target lacks a suitable small molecule-binding site. Recombinant antibodies have several advantages including exquisite affinity and specificity, as well as more straightforward development and potential for proteome-wide coverage [26]. However, the use of such proteins for intracellular applications is limited due to the reducing environment of the cytoplasm that disfavors the formation of stable disulfide bonds, a necessary step in the folding of most immunoglobulin domains [27]. This has prompted the development of intracellular antibody fragments (intrabodies) that do not depend on disulfide bonds for folding [28-30] and non-antibody scaffolds such as the human fibronectin type III domain (FN3) [31, 32] and designed ankyrin repeat proteins (DARPins) that natively lack disulfides [33, 34]. Importantly, all of these modalities have been developed for antigen binding inside living cells [28, 31, 33]. Unfortunately, a drawback of these approaches is that attaining maximal effect requires 1:1 stoichiometry between inhibitor and target as well as sustained equilibrium target occupancy. Moreover, methods based on induced proteolysis can degrade proteins regardless of their function, a feat that is not possible with conventional occupancy-based inhibitors. Such function-independent knockout can also be achieved with engineered antibodies/antibody mimics; however, target binding alone does not always lead to inactivation.

Targeted protein degradation techniques

Proteolysis-targeting chimeric molecules

Proteome editing technology represents an orthogonal strategy for post-translational control of protein function based on the principle of inhibition-by-degradation whereby the "inhibitor" hijacks the cellular quality control machinery to selectively degrade target proteins. Owing to the catalytic nature of this approach, induced proteolysis overcomes many of the drawbacks associated with traditional occupancy-based inhibitors in terms of potency, efficacy, duration of action, and target selection. For example, proteome editing can be very efficient because both old and new protein molecules are targeted concomitantly, and it can be readily reversed. Because proteome editing operates at the post-translational level, it has the potential to delineate the functional consequences associated with specific post-translational modifications by preferentially removing modified proteins while leaving the unmodified subpopulations intact. Such distinctions are difficult or impossible to achieve with other silencing technologies.

One of the most advanced strategies for controlling protein levels involves small molecules called proteolysis targeting chimeras, or proteolysis-targeting chimeric molecules (PROTACs), that bring together an E3 ubiquitin ligase with the target protein to promote its degradation [35-37]. PROTACs are heterobifunctional small molecules consisting of two specific ligands—one directed to the protein target and one to an E3 ubiquitin ligase—joined together by a linker. Importantly, since PROTACs are not degraded with the protein, they have the ability to reach complete knockdown at sub-stoichiometric ratios, alleviating the need for sustained target occupancy that limits traditional small-molecule inhibitors.

The first PROTACs, described in 2001, were generated by linking a 10 amino-acid phosphopeptide derived from IκBα, which binds the Skp1-Cullin-F-box (SCF) E3 ubiquitin ligase complex of the cullin-RING ubiquitin ligase (CRL) family, to the small molecule ovalicin, which covalently binds to the active site of methionine aminopeptidase 2 (MetAP-2) [38]. In the presence of this chimeric compound, MetAP-2 was recruited to SCF, ubiquitinated, and degraded in a PROTACs-dependent manner. Using the same phosphopeptide linked to estradiol and dihydroxytestosterone enabled chemical knockdown of estradiol receptor (ER) and androgen receptor (AR), respectively [39]. To develop conditionally-active degraders, a 10-residue tyrosine kinase substrate sequence was linked to a peptide ligand for the von Hippel-Lindau (VHL) protein, a substrate receptor of the Cullin2 (CUL2) E3 ligase complex from the CRL superfamily. The resulting "phosphoPROTACs" displayed inducible silencing activity that was activated by a given receptor tyrosine kinase [40]. Despite these and other early successes, first-generation, peptide-based designs suffered from multiple limitations, including large size, poor cellular permeability and metabolism, thereby limiting their therapeutic viability.

Efforts to overcome these permeability issues have focused on entirely small-molecule based designs. The first such example conjugated an imidazoline derivative that binds the E3 ubiquitin ligase MDM2 to small molecule ligands of AR [41], which addressed the permeability issue but were not very potent. Along similar lines, Hashimoto and coworkers created hybrids between methyl bestatin, which binds the E3 ligase called inhibitor of apoptosis protein (IAP), and all-trans retinoic acid (ATRA), an endogenous ligand of cellular retinoic acid binding proteins CRABP-I and -II. These molecules, named SNIPERS (specific and non-genetic IAP-dependent protein erasers), promoted efficient

degradation of their protein targets; however, high doses (≥1 μM) were required for efficacy and numerous off-target effects were observed due to the lack of specificity of bestatin [42]. More recently, Crews and coworkers described a new generation of nonpeptidic PROTACs in which the peptide moiety of earlier designs was replaced with a high-affinity, small-molecule ligand for the VHL substrate receptor of the CUL2 E3 complex [43]. When linked to ligands for various targets such as the estrogen-related receptor alpha (ERRα) and the serine-threonine kinase RIPK2, efficient target degradation was observed with half-maximal effective concentration (EC_{50}) values ranging from <1 to 100 nM. Similar results were obtained with nonpeptidic PROTACs comprised of thalidomide and related molecules that bind with high affinity to the cereblon (CRBN) component of CRL complexes [44, 45].

Overall, PROTACs have evolved from early peptide-based compounds to more drug-like, small-molecule-based designs that enable chemical knockout in cells and in mice, and are just now entering the clinic [46]. Nonetheless, enhancements to potency, selectivity, delivery, metabolic stability, and bioavailability are still needed. These efforts will likely focus on optimizing linker length/composition, expanding the diversity of E3s recruited and protein substrates degraded, and developing targeted delivery strategies such as antibody-PROTACs conjugates that use the specificity of monoclonal antibodies to transport conjugated PROTACs payloads into desired cells and tissues. Such efforts will likely benefit from a recent structural study of the ternary complex between the E3 ligase, a PROTAC, and its target [47], shedding light on the mechanism and offering new opportunities for rationally optimizing PROTACs. Beyond their clinical development, PROTACs have also been adapted for preclinical target validation. For example, the dTAG

[48] and HaloPROTACs [49] schemes leveraged small-molecule ligands for FKB12 and HaloTag to selectively degrade fusion proteins comprised of FKB12 or HaloTag, respectively. Due to the wide availability of such fusion proteins in biological studies, the dTAG and HaloPROTACs degraders should prove useful as chemical genetic tools.

Heterobifunctional protein-based degraders

Although PROTACs represent a promising approach for degrading proteins of interest including essential drivers of human disorders, their generation is still limited by a lack of available ligands that link an E3 ubiquitin ligase with a desired protein as well as by technical difficulties associated with creating such ligands *de novo* [50]. Protein-based chimeras bypass these challenges by leveraging recombinant DNA technology to create hybrid proteins comprised of one domain having E3 ubiquitin ligase activity and another domain having target-binding activity. The resulting chimeras promote polyubiquitination and subsequent proteasomal degradation of their protein substrates when ectopically expressed in cells. An advantage of this approach is the potential to effectively bridge ubiquitin conjugation to virtually any intracellular target provided that a protein-binding domain is available or can be engineered.

In the earliest examples, target recognition was achieved by leveraging pairwise interactions that involve the protein of interest. For example, Zhou and coworkers demonstrated efficient protein knockout by genetically fusing the F-box protein β-TrCP with a peptide derived from the E7 protein encoded by human papillomavirus type 16 that is known to interact with retinoblastoma protein pRB [51, 52]. Following ectopic expression, the engineered chimera recruited pRB to the SCF machinery, a multi-protein E3 complex from the CRL superfamily, for ubiquitination and destruction. A handful of

other studies have similarly exploited natural protein-protein interactions, whereby fusion of an interacting partner protein to an E3 (or a component of an E3 ligase complex) yielded chimeras that silenced different oncoproteins including c-Myc [53], ErbB [54], HIF-α [55], and KRAS [56]. Several of these chimeras were even able to reduce angiogenesis and/or inhibit tumor growth following delivery by viral vectors [53, 55, 56], revealing the promise of this approach for therapeutic knockout of cancer targets.

Antibody-like heterobifunctional protein-based degraders

In a notable departure from these earlier efforts, our group reported a customizable proteome editing technology that could be extended beyond naturally occurring binding pairs. Specifically, we created heterobifunctional hybrids, called "ubiquibodies" (uAbs), by genetically fusing an E3 to an engineered antibody or to non-antibody scaffolds that bind specifically to proteins of interest [57, 58]. In the earliest designs, the flexible ubiquitin-tagging capacity of the human RING/U-box-type E3 ligase CHIP (carboxyl terminus of Hsc70-interacting protein) was tethered to different synthetic binding proteins such as a single-chain antibody fragment (scFv), a designed ankyrin repeat protein (DARPin), or a fibronectin type III (FN3) monobody. The resulting ubiquibodies efficiently redirected structurally diverse protein substrates, including β-galactosidase (β-gal) and maltose-binding protein (MBP) both from *E. coli*, to the UPP for proteolytic degradation independent of their biological function or native interactions [57]. Around the same time, Affolter and coworkers developed a similar strategy called "deGradFP" for shunting GFP-tagged proteins to the proteasome using a chimera between the F-box domain of *Drosophila melanogaster* Slmb that recruits the SCF machinery and a camelid-derived VHH nanobody specific for GFP [59]. Likewise, Sapkota and coworkers developed

ubiquibodies (under the name AdPROM for affinity-directed protein missile) that also degraded GFP-tagged proteins using the same anti-GFP nanobody attached to the VHL substrate receptor of the CRL system [60].

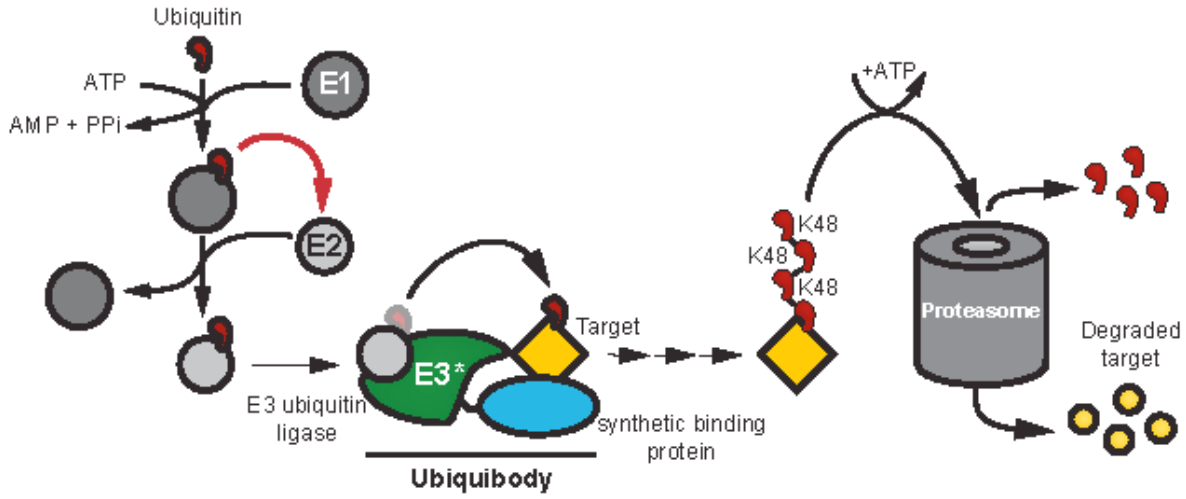


Figure 1.2 Ubiquibodies hijack the UPP for targeted protein degradation.
Schematic representation of ubiquibodies, chimeric fusions comprised of an E3 ubiquitin ligase catalytic domain (E3*) and a designer binding proteins (DBP) that hijack the mammalian ubiquitin proteasome pathway for degradation of desired protein targets (T).

A distinct advantage of ubiquibodies is their highly modular architecture, which enables target selection to be rewired by simply swapping the synthetic protein-binding domain. For example, targeted proteolysis has been achieved for a diverse array of protein substrates including eukaryotic proteins (ASC [61], Lck,[62] and SHP2 [61, 63]), intraneuronal bacterial proteins (Clostridium botulinum neurotoxin (BoNT) proteases) [64], and fluorescent proteins (FPs) [60, 61, 63, 65, 66]. Moreover, by incorporating synthetic binding proteins that recognize particular protein states (e.g., active vs. inactive conformation, mutant vs. wild-type, post-translationally modified, etc.), it becomes possible to deplete certain protein subpopulations while sparing others [52, 54, 58, 63] Also, in contrast to PROTACs where ligands for new substrates are hard to come by,

expanding the number of targets for ubiquibodies including those deemed undruggable is made relatively straightforward by library-based screening methods such as phage display [67], ribosome display [68], or yeast surface display [69]. Similarly, changing the E3 domain can readily reprogram ubiquitination kinetics or the mechanism of ubiquitin transfer. Indeed, a number of different E3 domains have been developed for ubiquibody-mediated knockdown including stand-alone E3s such as CHIP [57, 63, 66], E3 substrate adaptors/receptors that associate with larger E3 ligase complexes [59-63, 65].

Bacterial and viral E3 ubiquitin ligases

In the continuous battle between pathogens and their hosts, it is now firmly established that many bacteria and viruses counteract immune defense by inhibiting or redirecting the ubiquitination machinery of the host. One common immune evasion tactic used by these invaders is the deployment of their own E3 ubiquitin ligases that selectively target host immunoproteins for proteasomal degradation. By degrading important host adaptor and signaling proteins, bacteria and viruses effectively disable multiple innate immune pathways including the production of and response to interferons as well as other innate host defense mechanisms as discussed below.

Although some bacterial species contain a prokaryotic ubiquitin-like protein (Pup) system for the post-translational modification and degradation of proteins [70], most lack a canonical UPP and LP. Nonetheless, the versatility of these systems in regulating protein function and cell behavior makes them a particularly attractive target for bacteria. Indeed, to dampen the innate immune response during infection, many pathogenic bacteria encode effector proteins that either mimic enzymes of the UPP or modulate the activity of these host enzymes. Of relevance here are bacterial E3 ligases (or E3 ligase substrate adaptor

proteins) that redirect host immunoproteins to the UPP machinery for degradation [71-73]. These E3 mimics typically gain access to the host cytosol via direct injection through the bacterium's type III or type IV secretion system [74-76] although some appear to have inbuilt protein-transduction domains that facilitate internalization in host cells [77].

One of the earliest and most notable examples comes from studies of Shigella flexneri, the causative agent of bacillary dysentery. After being phagocytosed by the host, S. flexneri bacteria deliver a cocktail of type III effector proteins, mainly consisting of OspG, OspI and the 10-member IpaH family, that collectively function to inhibit the activation of the NF-κB pathway. The IpaH family members exhibit similarities to eukaryotic HECT-type E3s but are classified as novel E3 ligases (NELs) due to the absence of sequence and structural homology with any eukaryotic E3s [78, 79]. NELs contain N-terminal leucine-rich repeat (LRR) regions for substrate recognition and conserved C-terminal domains (CTDs) with E3 ubiquitin ligase activity. Among these, IpaH9.8, IpaH4.5, and IpaH0722 modulate the host inflammatory response by promoting proteasome-dependent degradation of NF-κB essential modulator (NEMO), the NF-κB p65 subunit, and tumor necrosis factor receptor-associated factor 2 (TRAF2), respectively. [72, 73] Interestingly, OspG and OspI proteins interact with the UPP in the opposite way, by preventing protein degradation. Specifically, OspG binds to activated E2s to bypass the ubiquitination of IκBα [80], an NF-κB inhibitor, while OspI disrupts the NF-κB pathway by preventing the polyubiquitination of UBC13 [81].

Viruses have also evolved elegant strategies for exploiting or avoiding the ubiquitin-proteasome system [82, 83]. Like their bacterial counterparts, viral genomes encode substrate adaptor proteins that recruit host E3 ligases or stand-alone E3 ligases,

which induce proteolysis of host immune factors that might have harmful effects on the viral life cycle. For example, Kaposi sarcoma herpesvirus (KSHV) encodes a viral E3 ligase of the RING-CH family named K3 that utilizes the UPP to ubiquitinate major histocompatibility complex I (MHC-I), leading to its lysosomal degradation in a ubiquitin-proteasome-dependent manner [84]. Likewise, the infected cell protein 0 (ICP0) of herpes simplex virus type 1 (HSV-1) is a viral E3 that induces UPP-dependent degradation of promyelocytic leukemia protein (PML) and Sp100, thereby avoiding antiviral sensing [85]. The E6 oncoprotein from human papillomavirus (HPV) is an example of a viral adaptor protein that recruits the cellular E3 ubiquitin ligase E6-AP to induce ubiquitination and degradation of p53, thereby allowing viral replication [86].

Therapeutic delivery challenges

Finally, engineered degraders must be able to efficiently enter cells if they are to be translated into the clinic. Unfortunately, intracellular delivery of protein-based therapeutics represents a significant obstacle as most globular protein drugs do not spontaneously cross plasma membranes due to their relatively large size and biochemical properties [50]. Efforts to overcome the delivery challenge in the context of targeted protein degradation have primarily focused on viral delivery methods [53, 55, 56]. In another recent study, Hantschel and coworkers used a chimeric bacterial toxin to deliver ubiquibodies in cancer cells [62].

Approximately 30 cell-permeable proteins (CPPs or cell penetrating peptides) , which are able to cross cellular membranes, have been identified from natural sequences, modified natural sources, or isolated de novo [87]. CPPs have become widely manipulated for the delivery of covalently attached and non-covalently associated cargos including

nucleic acids, proteins, small molecules, and imaging agents. Recently, it was discovered that the LRR motif-containing bacterial effectors YopM from Yersinia enterocolitica [88], SspH1 from Salmonella typhimurium [89], and IpaH effectors from Shigella flexneri [90] are autonomously translocated into mammalian cells as a direct result of their N-terminal 2αH protein domains. Furthermore, they showed that the 2αH domain is necessary and sufficient for cell-penetration, as genetic fusion to other proteins, like GFP, resulted in their cellular uptake as well.

CHAPTER 2 - ENGINEERING BACTERIAL E3 UBIQUITIN LIGASES WITH BROAD-SPECTRUM SILENCING

Introduction

Manipulation of the ubiquitin-proteasome pathway to achieve targeted silencing of cellular proteins has emerged as a reliable and customizable strategy for remodeling the mammalian proteome. One such approach involves engineering bifunctional proteins called ubiquibodies that are comprised of a synthetic binding protein fused to an E3 ubiquitin ligase, thus enabling post-translational ubiquitination and degradation of a target protein independent of its function. Here, we attempted to broaden the range of E3s that can be functionally reprogrammed as bifunctional ubiquibody chimeras. However, in a notable departure from previous efforts involving mammalian E3s, we focused instead on a set of effector proteins from microbial pathogens that mimic host E3 ubiquitin ligases and hijack the UPP machinery to dampen the innate immune response during infection [71, 91]. The intrinsic plasticity of these enzymes led us to hypothesize that bacterial E3s could be manipulated for targeted proteolysis just like their mammalian counterparts.

Indeed, robust target silencing was achieved with a ubiquibody comprised of the *Shigella flexneri* E3 ligase IpaH9.8, which exhibits similarities to eukaryotic HECT-type E3s but is classified as a novel E3 ligase (NEL) due to the absence of sequence and structural homology with any eukaryotic E3s [71, 74, 79, 91, 92]. Specifically, when the C-terminal catalytic NEL domain of IpaH9.8 was fused to the GFP-specific FN3 monobody GS2 that specifically recognizes green fluorescent protein (GFP), was observed to potently eliminate GFP and its spectral derivatives as well as 15 different FP-tagged mammalian

proteins that varied in size (27–179 kDa) and subcellular localization (cytoplasm, nucleus, membrane-associated, and transmembrane). Furthermore, GS2-IpaH9.8 showed equivalent activity regardless of its mode of expression or that of its target. Together, these results demonstrate that GS2-IpaH9.8 is a flexible, modular, and robust methodology for the targeted degradation of GFP-tagged substrates in mammalian cells.

Results

Engineered IpaH9.8 potently silences GFP in mammalian cells

To determine whether E3 ubiquitin ligase mimics from pathogenic bacteria could be redesigned for silencing of non-native targets, we focused on a panel of 14 candidate enzymes representing the major classes of E3s found in bacteria to date (**Appendix A**) [71, 91]. This panel included E3 mimics with folds similar to eukaryotic E3s such as HECT-type, RING or U-box (RING/U-box)-type, and F-box domains, as well as unconventional E3s with folds unlike any other eukaryotic E3s such as NEL, XL-box-containing, and SidC. In general, ubiquibodies were engineered by removing the native substrate-binding domain from each E3 mimic and replacing it with a synthetic binding protein (**Fig. 2.1a**), akin to our previously designed ubidquibodies based on human CHIP [57]. For example, S. flexneri IpaH9.8 consists of an N-terminal domain with eight 20-residue leucine-rich repeats (LRRs) that mediate binding and specificity to native substrate proteins such as NF-κB essential modulator (NEMO) [93] and guanylate-binding proteins (GBPs) [94], while the C-terminal domain adopts a novel E3 ubiquitin ligase architecture [79, 92]. Hence, we replaced the N-terminal LRR domain of IpaH9.8 with GS2, an FN3 monobody that binds GFP with nanomolar affinity (K_d = 31 nM) [95]. By swapping the natural substrate recognition function of these enzymes with the GS2 monobody, synthetic E3

ligases were created that we hypothesized would target GFP and promote its proteasomal degradation. To test this hypothesis, the different GS2-E3 chimeras were transiently co-expressed along with enhanced GFP (EGFP) in mammalian cells and fluorescence activity was monitored by flow cytometric analysis. By far, the most striking depletion of EGFP was achieved with GS2-IpaH9.8, which reduced EGFP fluorescence to near background levels (**Fig. 2.1b**). All of the other ubiquibodies showed relatively weak silencing activity under the conditions tested here. GS2-NleG5-1, GS2-SspH1, SidC-GS2, and GS2-SopA were the most active among these, reducing EGFP fluorescence by ~20-40% (**Fig. 2.1b**).

In light of this robust silencing activity, we decided to focus our attention on the GS2-IpaH9.8. In cells expressing this chimera, the elimination of EGFP was efficient, with removal of up to 90% of the fluorescence activity (**Fig. 2.2a and b**) and no detectable EGFP protein in cell lysates (**Fig. 2.2c**). Importantly, silencing activity was completely abrogated when the catalytic cysteine of IpaH9.8 [74] was mutated to alanine (GS2-IpaH9.8^{C337A}) and when the non-cognate FN3 monobody AS15, which is specific for the Abl SH2 domain [96], was substituted for GS2 (**Fig. 2.2a-c**), indicating that target degradation was dependent on cooperation of both ubiquibody domains. In the case of GS2-IpaH9.8^{C337A}, expression in mammalian cells and EGFP binding activity *in vitro* were unaffected by the alanine substitution (**Fig. 2.3**), confirming that loss of silencing activity was due to catalytic inactivation. It should also be noted that removal of the LRR domain was essential for knockdown activity, as direct fusion of GS2 to full-length IpaH9.8 that had not been truncated resulted in no measurable silencing activity (data not shown). Interestingly, the genome sequences of S. flexneri strains indicate that several IpaH family members, namely IpaH1.4, IpaH2.5, IpaH4.5, IpaH7.8 and IpaH9.8, are encoded on the

220-kb virulence plasmid pWR100 while seven additional ipaH cognate genes are present on the chromosome [71]. To determine whether these family members were as proficient as IpaH9.8 at degrading EGFP in the ubiquibody context, we generated chimeras between GS2 and the catalytic domains derived from each of the pWR100-encoded IpaH family members as well as one chromosomally encoded member, IpaH0722. When expressed ectopically in cultured cells, all of the IpaH-based ubiquibodies were capable of efficient (~90% or greater) EGFP knockdown in mammalian cells (**Fig. 2.2c**). This result was not entirely surprising in light of the high homology shared by the different catalytic domains. Indeed, whereas the different IpaH family members were only ~70% similar to IpaH9.8 overall, the catalytic domains were much more similar (>99%) with just 1-3 amino acid substitutions and, in the case of IpaH1.4 and IpaH4.5, minor C-terminal truncations (**Appendix A**).

To benchmark the potency of our engineered bacterial ligase, we compared the GFP silencing activity catalyzed by GS2-IpaH9.8 with that of other synthetic ligases based on eukaryotic E3 machinery that have previously been reconfigured for targeted proteolysis [51-54, 56, 59-61, 65, 66]. Specifically, the natural substrate-binding domains for several eukaryotic E3 ubiquitin ligases from humans including carboxyl terminus of Hsc70-interacting protein (CHIP), speckle-type POZ protein (SPOP), β-transducing repeat-containing protein (βTrCP), and von Hippel-Lindau protein (VHL), as well as the Drosophila melanogaster supernumerary limbs (Slmb) protein were replaced with the GS2 monobody, resulting in a panel of synthetic ligases analogous to GS2-IpaH9.8. When the resulting panel of GFP-specific ubiquibodies was transiently co-expressed with EGFP in mammalian cells, all were capable of measurably reducing EGFP levels, but silencing

activity for each was relatively inefficient (~25-45%) under the conditions tested here (**Fig. 2.c**), reminiscent of previous results with a Slmb-nanobody chimera that was similarly ineffective at reducing unfused GFP levels [59]. The weak EGFP knockdown observed here for Slmb-GS2 was actually an improvement over previous results obtained with a chimera between Slmb and a GFP-specific VHH nanobody, cAbGFP4, that was incapable of promoting degradation of unfused GFP [59]. It should be noted, however, that the Slmb-cAbGFP4 fusion eliminated the fluorescence associated with larger GFP fusion proteins, suggesting that the data reported here are not necessarily indicative of ubiquibody dysfunction but instead may reflect differences in substrate preference/compatibility or extent of ubiquitin decoration. Regardless, none of the engineered chimeras involving eukaryotic E3s displayed the potency and robustness of GS2-IpaH9.8, which reproducibly degraded 90-95% of cellular fluorescence.

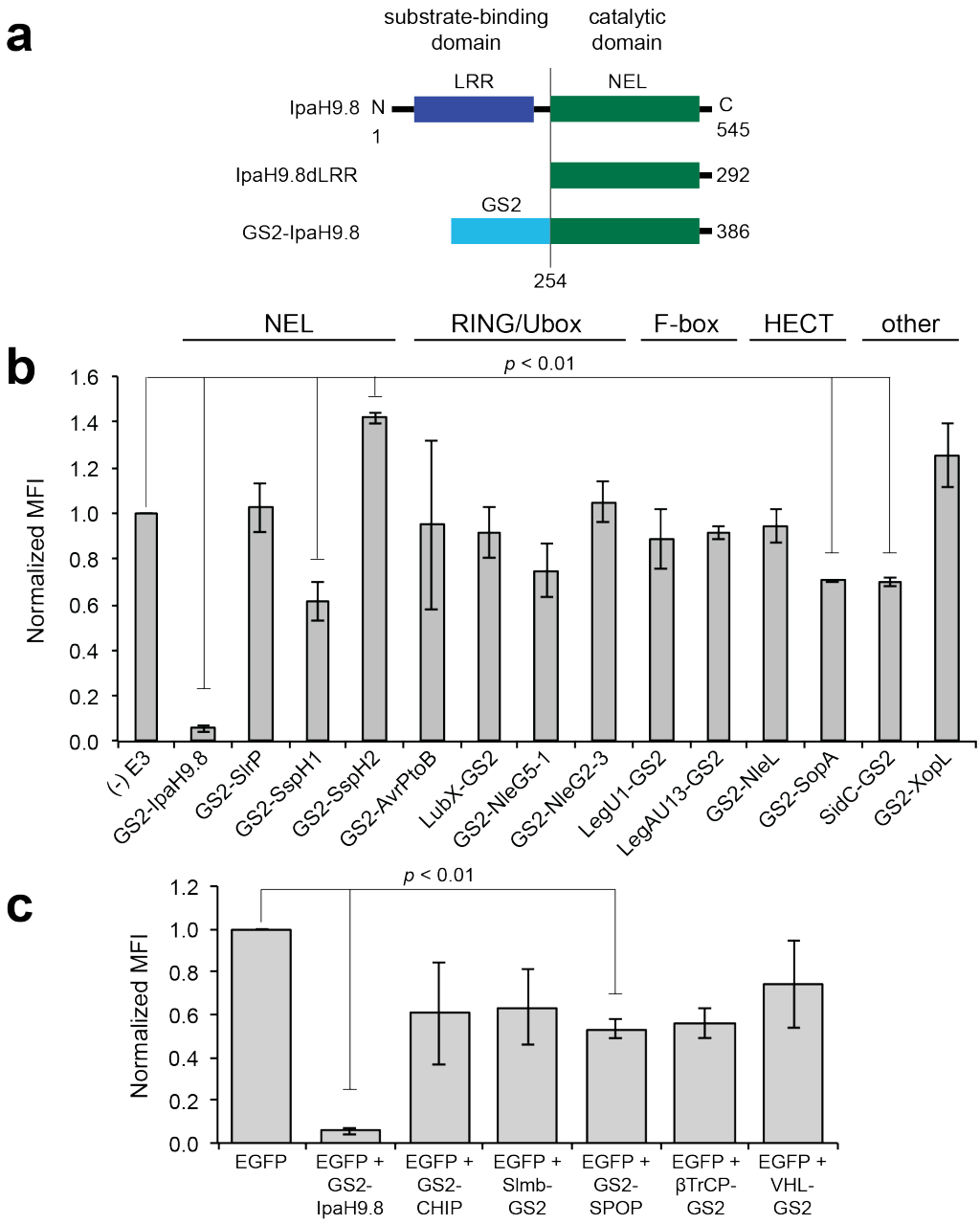


Figure 2.1 Engineering bacterial E3 ligase IpaH9.8 as a GFP-specific ubiquibody.
(a) Linear representation of IpaH9.8, IpaH9.8dLRR, and GS2-IpaH9.8. Numbers refer to amino acid positions from N terminus (N) to C terminus (C). The proteins are aligned vertically with the LRR and NEL domains of IpaH9.8. IpaH9.8dLRR is a truncated version of IpaH9.8 lacking the LRR domain. (b) Flow cytometric analysis of EGFP fluorescence activity in HEK293T cells transfected with plasmid pcDNA3-EGFP alone or co-transfected with pcDNA3-EGFP and a plasmid encoding one of the bacterial E3-based ubiquibodies as indicated. (c) Same as in (b) but with mammalian E3-based ubiquibodies as indicated. Flow cytometry data are representative of biological triplicates (three separately transfected wells) of the geometric MFI normalized to MFI measured for HEK283T cells expressing EGFP alone Error bars represent standard deviation (SD) of the mean. *p* values were determined by paired sample *t*-test.

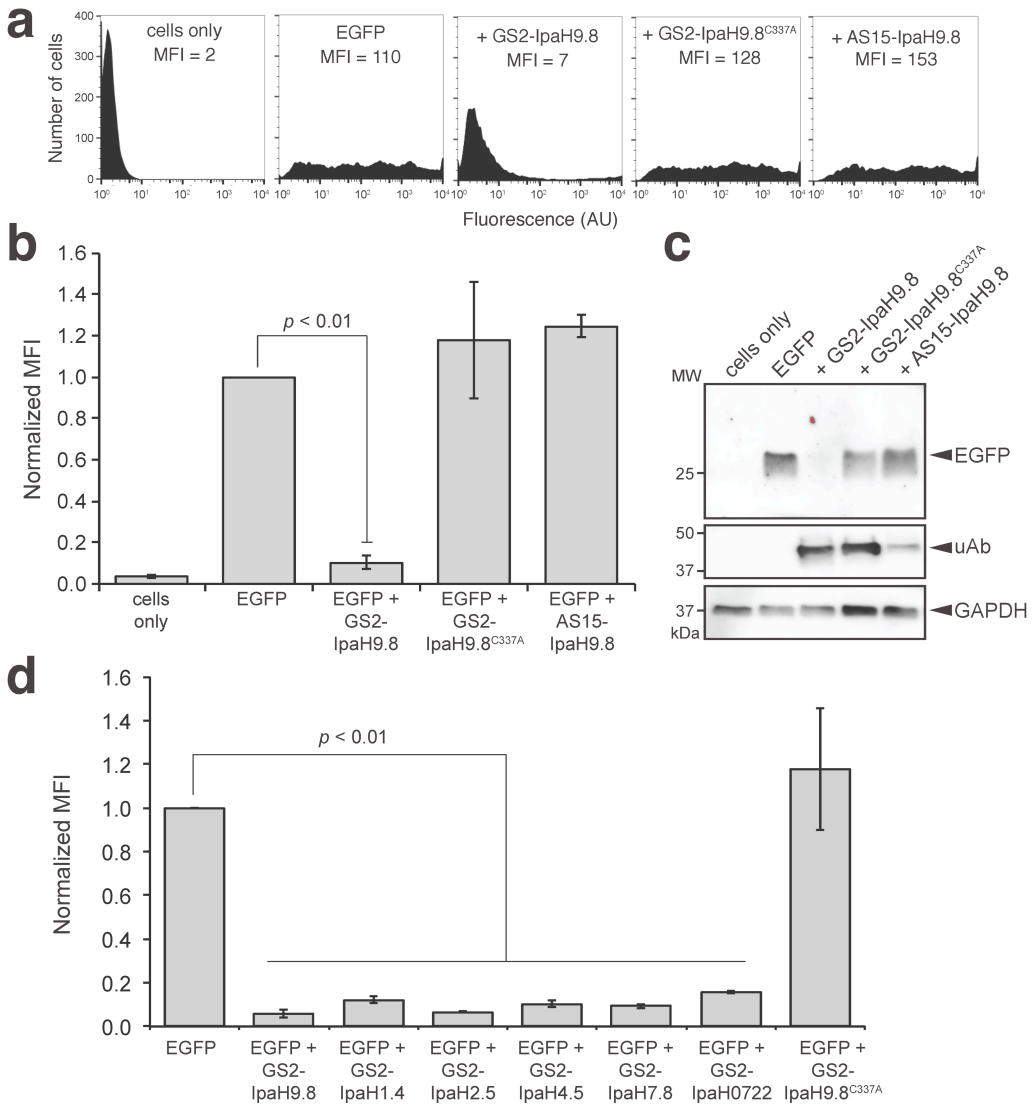


Figure 2.2 Catalytic domain of IpaH.8 essential for ubiquibody function.
(a) Representative fluorescence histograms obtained by flow cytometric analysis of EGFP fluorescence activity in HEK293T cells transfected with pcDNA3-EGFP alone or co-transfected with pcDNA3-EGFP and a plasmid encoding one of the following: GS2-IpaH9.8^{C337A}, AS15-IpaH9.8, or GS2-IpaH9.8. (b) Flow cytometric quantification of EGFP fluorescence activity for cells described in (a). (c) Western blot analysis of HEK293T cell lysates transfected as in (a) and (b). Blots were probed with antibodies specific for GFP, 6x-His (that detected tag on each ubiquibody), and GAPDH as indicated. An equivalent amount of total protein was loaded in each lane as confirmed by immunoblotting with anti-GAPDH. Molecular weight (MW) markers are indicated on left. (d) Flow cytometric quantification of EGFP fluorescence activity for HEK293T cells co-transfected with pcDNA3-EGFP and a plasmid encoding GS2 fused to one of the IpaH9.8 homologs as indicated. Flow cytometry data are representative of biological triplicates (three separately transfected wells) of the geometric MFI normalized to MFI measured for HEK283T cells expressing EGFP alone. Error bars represent standard deviation (SD) of the mean. *p* values were determined by paired sample *t*-test.

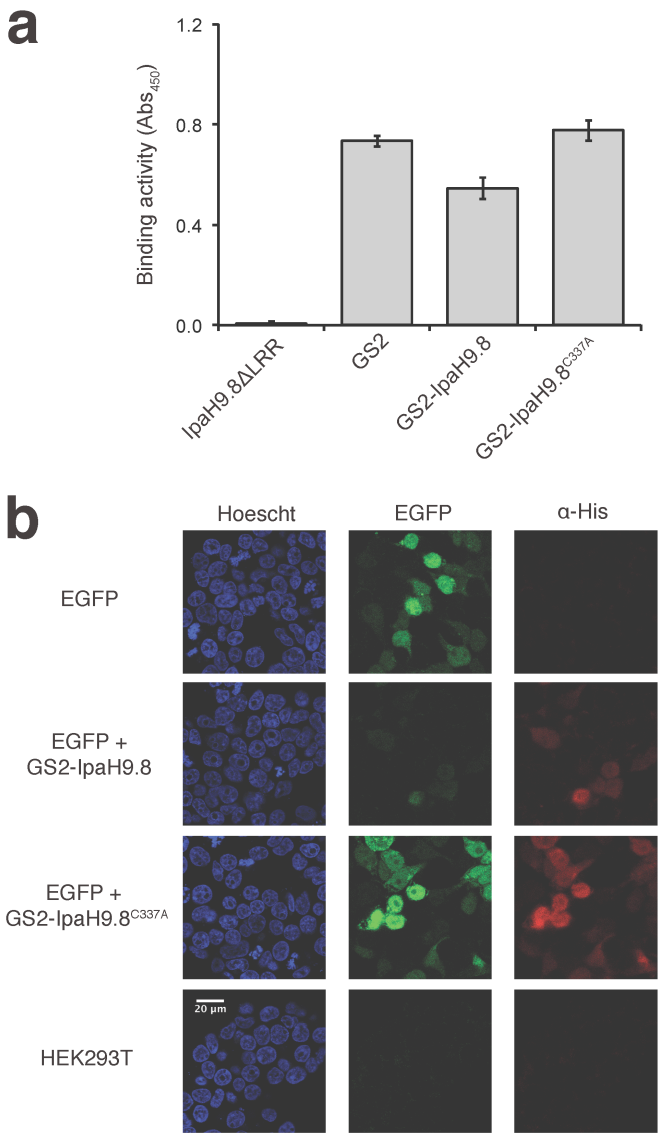


Figure 2.3 Characterization of GS2-IpaH9.8 binding activity and expression.
(a) Binding activity of GS2-IpaH9.8 compared to GS2 alone, IpaH9.8 lacking the LRR domain (IpaH9.8 LRR), or catalytically inactive GS2-IpaH9.8^{C337A} as indicated. Activity was measured by ELISA using GFP as immobilized antigen and 15 mg/mL of each protein applied per well. Detection was performed using anti-FLAG antibody conjugated to horseradish peroxidase (HRP). The quenched plate was read at 450 nm (Abs450). Data is the average of three biological replicates and error bars are the standard deviation (SD) of the mean. (b) Confocal microscopy images corresponding to HEK293T cells transfected with plasmid DNA encoding EGFP or co-transfected with plasmid DNA encoding EGFP and either pcDNA3-GS2-IpaH9.8^{C337A} or pcDNA3-GS2-IpaH9.8 as indicated. Non-transfected HEK293T control cells are also depicted. Hoescht stain (blue) denotes cell nuclei, EGFP signal (green) denotes FP target expression, and α-His signal (red) corresponds to immunolabeling of expressed ubiquibody in permeabilized cells.

A broad range of substrate proteins is degraded by GS2-IpaH9.8

To more deeply explore the substrate compatibility issue, we tested the ability of GS2-IpaH9.8 to degrade a range of different substrates. A growing number of GFP-derived fluorescent proteins (FPs) have been developed and optimized over the years, providing a diverse collection of new tools for biological imaging [97, 98]. To determine the extent to which different FP targets could be degraded, GS2-IpaH9.8 was transiently co-expressed in mammalian cells with monomeric versions of Emerald, Venus and Cerulean, as well as enhanced cyan fluorescent protein (ECFP). Approximately 65-85% of the cellular fluorescence activity associated with each of the FPs was ablated by GS2-IpaH9.8, whereas the structurally unrelated mCherry protein was not targeted by GS2-IpaH9.8, which was expected given the specificity of GS2 for the GFP fold (**Fig. 2.4**). Interestingly, the fluorescence activity of superfolder GFP (sfGFP), a rapidly folding and robustly stable mutant of EGFP, was unaffected by GS2-IpaH9.8, consistent with recent findings that sfGFP is resistant to proteasomal degradation [99].

Encouraged by the ability of GS2-IpaH9.8 to degrade different FPs, we next evaluated the ability of GS2-IpaH9.8 to degrade structurally diverse, FP-tagged substrate proteins. GS2-IpaH9.8 proficiently degraded 15 unique target proteins that varied in terms of their molecular weight (27-179 kDa) and subcellular localization (i.e., cytoplasm, nucleus, membrane-associated, and transmembrane) (**Fig. 2.5a and Fig. 2.6**). For example, GS2-IpaH9.8 triggered degradation of 80-92% of the fluorescence activity associated with FP fusions involving the cytoplasmic proteins α-actinin, α-synuclein (α-syn), extracellular signal-regulated kinase 2 (ERK2), focal adhesion kinase (FAK), F-tractin, paxillin (PXN), and vinculin (VCL) as determined by flow cytometric analysis (**Fig. 2.5 and Fig. 2.6**).

Similarly robust silencing was observed for: nuclear-targeted FP fusions involving histone H2B and the nuclear localization signal (NLS) derived from SV40 Large T-antigen; membrane-associated FP fusions involving Harvey rat sarcoma virus oncogene homolog carrying the oncogenic G12V mutation ($HRas^{G12V}$), Src-homology 2 domain-containing phosphatase 2 (SHP2), and the farnesyl sequence derived from HRas; and transmembrane FP fusions involving epidermal growth factor receptor (EGFR), avian erythroblastic leukemia viral oncogene homolog 2 (ErbB2), and mucin 1 (MUC1) (**Fig. 2.5a and Fig. 2.6**). Microscopy analysis of representative substrate proteins α-actinin-mEmerald, EGFP-NLS, farnesyl-mEmerald, and EGFR-mEmerald confirmed the expected subcellular localization of each fusion and corroborated the efficient degradation activity measured by flow cytometric analysis (**Fig. 2.5b**). The transmembrane protein EGFR-mEmerald was examined by immunolabeling with an antibody specific to the extracellular domain of EGFR. Importantly, the α-EGFR signal decreased concomitantly with GFP disappearance (**Fig. 2.5b**), indicating that degradation of the entire transmembrane protein was achieved. Taken together, these results establish GS2-IpaH9.8 as a robust proteome editing tool that is capable of silencing a broad spectrum of substrates that span several distinct subcellular locations.

GS2-IpaH9.8-mediated proteome editing is flexible and modular

An attractive feature of ubiquibodies is their highly modular architecture – the E3 catalytic domain and synthetic binding protein domain can be interchanged to reprogram the activity and specificity. Indeed, our results above revealed the ease with which different bacterial and eukaryotic E3 domains can be chimerized to form functional ubiquibodies. To investigate the interchangeability of the synthetic binding protein domain

in IpaH9.8-based ubiquibodies, we first replaced GS2 with other high-affinity GFP-binding proteins such as the FN3 monobody GS5 (K_d = 62 nM) [95] or cAbGFP4 (K_d = 0.32 nM) [100]. For these constructs, efficient EGFP silencing activity was observed that rivaled that seen with the GS2 monobody (Fig. 2.7). Interestingly, introduction of weaker affinity (~200-500 nM) FN3 monobodies [95] resulted in less efficient EGFP elimination (**Fig. 2.7**), suggesting that silencing activity may be a function of the affinity for the target protein. Although, because spatial arrangements and surface complementarity prioritize lysine sites for ubiquitination [101], an equally plausible explanation for these findings is that the various FN3 domains may differentially orient the ubiquibody with respect to GFP in a manner that affects how the substrate is ubiquitinated.

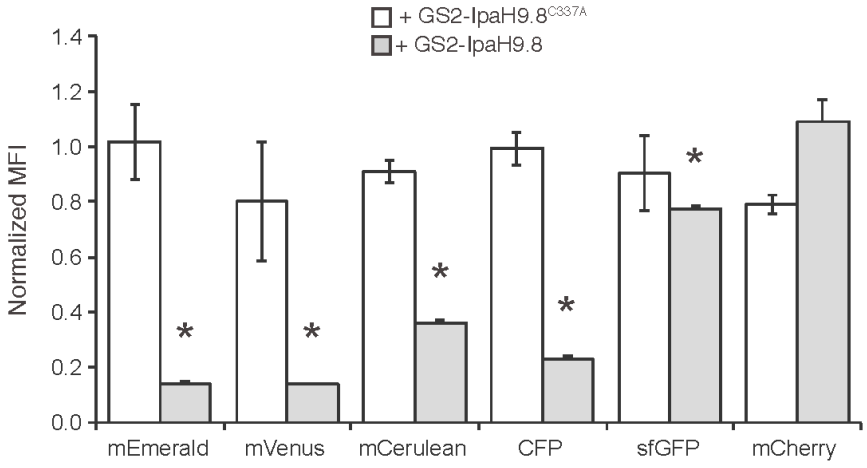


Figure 2.4 Ubiquibody-mediated silencing of FP variants.
Flow cytometric quantification of fluorescence activity in HEK293T cells co-transfected with plasmids encoding the FP variant and either pcDNA3-GS2-IpaH9.8^{C337A} (white) or pcDNA3-GS2-IpaH9.8 (grey) as indicated. mCherry served as negative control. Data are biological triplicates (three separately transfected wells) of the geometric MFI normalized to MFI measured for HEK293T cells expressing the corresponding FP alone. Error bars represent standard deviation (SD) of the mean. p values were determined by paired sample t-test. *, $p < 0.01$ (compared to HEK293T cells expressing the FP variant only).

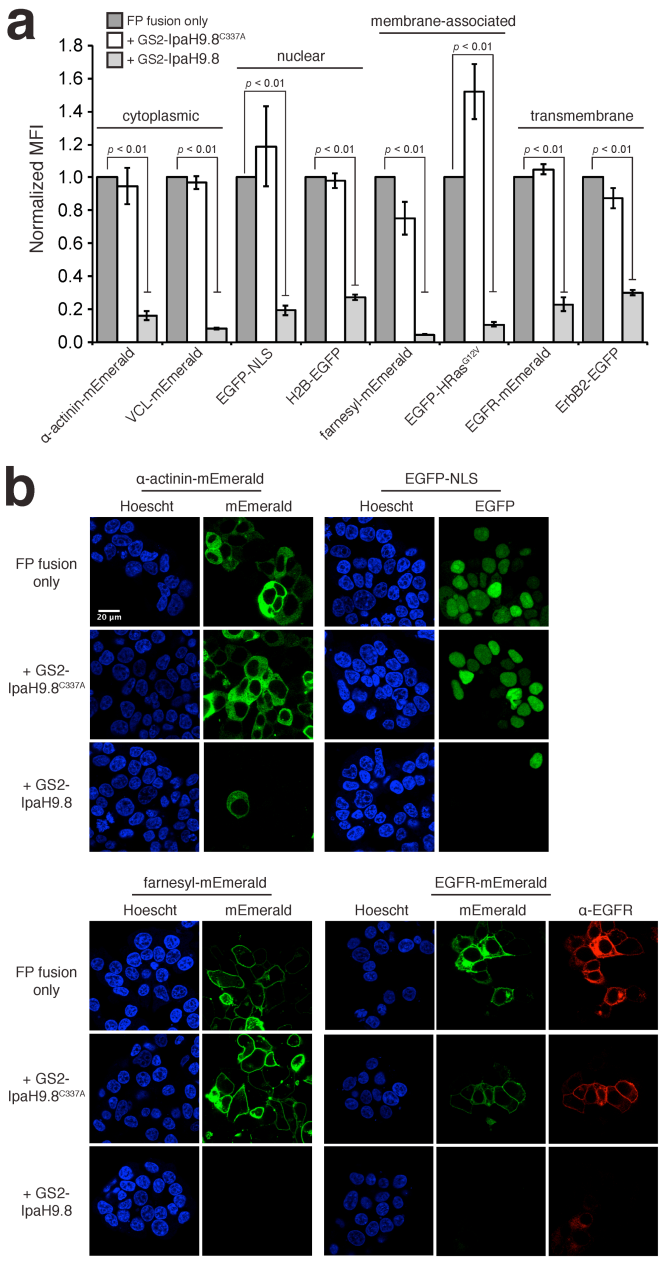


Figure 2.5 GS2-IpaH9.8 degrades structurally diverse fluorescent protein fusions.
(a) Flow cytometric quantification of fluorescence activity in HEK293T cells transfected with a plasmid encoding the indicated FP fusion alone (dark grey) or co-transfected with the FP fusion plasmid and either pcDNA3-GS2-IpaH9.8^{C337A} (white) or pcDNA3-GS2-IpaH9.8 (light grey). Data are biological triplicates (three separately transfected wells) of the geometric MFI normalized to MFI measured for HEK283T cells expressing the corresponding FP fusion protein alone. Error bars represent standard deviation (SD) of the mean. (b) Confocal microscopy images corresponding to select FP targets expressed in HEK293T cells transfected/co-transfected as described in (a). Hoescht stain (blue) denotes cell nuclei and EGFP signal (green) denotes fluorescent proteins. For the EGFR-mEmerald fusion, immunostaining with an EGFR-specific antibody (red) is also depicted. *p* values were determined by paired sample *t*-test.

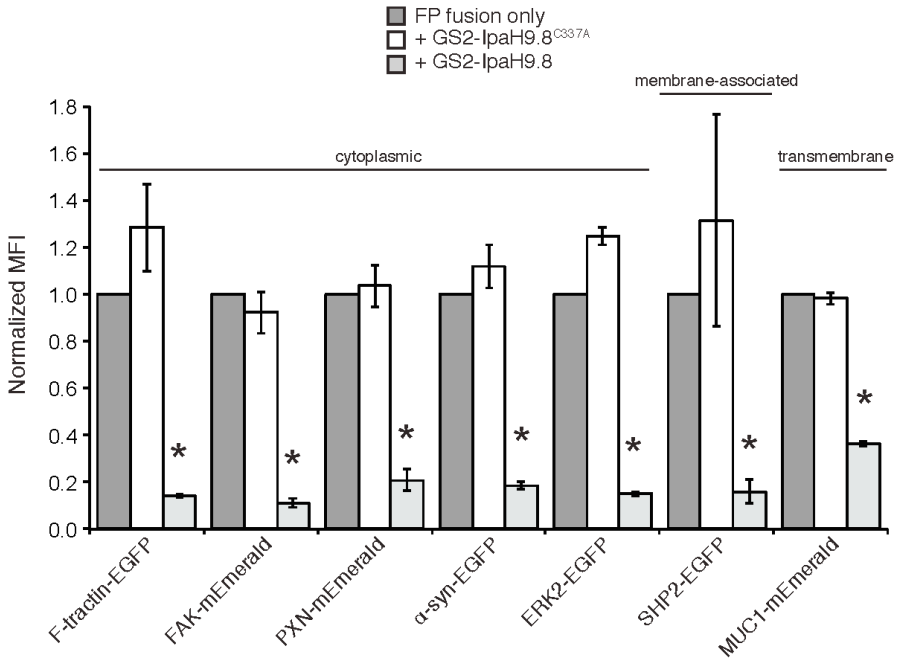


Figure 2.6 Ubiquibody-mediated silencing of additional FP fusion protein targets.
Flow cytometric quantification of fluorescence activity in HEK293T cells transfected with a plasmid encoding the indicated FP fusion alone (dark grey) or cotransfected with the FP fusion plasmid and either pcDNA3-GS2-IpaH9.8C337A (white) or pcDNA3-GS2-IpaH9.8 (light grey). Data are biological triplicates of the geometric MFI normalized to MFI measured for HEK283T cells expressing the corresponding FP alone. Error bars represent standard deviation (SD) of the mean. *, $p < 0.01$ (compared to HEK293T cells expressing the FP fusion only).

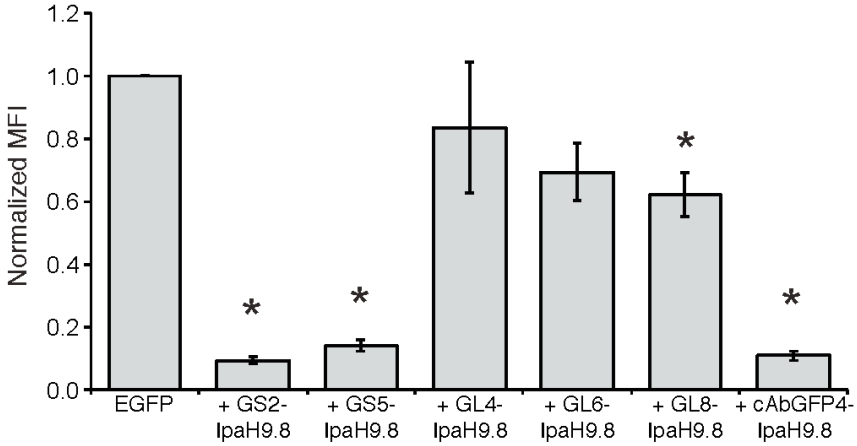


Figure 2.7 Modularity of ubiquibody platform.
Flow cytometric quantification of EGFP fluorescence activity in HEK293T cells transfected with plasmid DNA encoding EGFP or co-transfected with a plasmid encoding ubiquibody chimeras comprised of IpaH9.8 fused to a different GFP-directed binding protein as indicated. Data are biological triplicates of the geometric MFI normalized to MFI measured for HEK283T cells expressing the corresponding FP alone. Error bars represent standard deviation (SD) of the mean. *, $p < 0.01$ (compared to HEK293T cells expressing EGFP

It is worth noting, however, that when each of the bacterial or mammalian E3-based ubiquibodies was tested against a different substrate, namely EGFP-HRas^{G12V}, very similar results were observed. That is, a handful (e.g., XopL, CHIP, SPOP, and βTrCP) were able to reduce EGFP-HRas levels by as much as 60% but none were as effective as IpaH9.8 or its homologs, which all degraded ~80-90% of EGFP-HRas^{G12V} (**Figure 2.8**).

In all the experiments described above, efficient knockdown was achieved when GS2-IpaH9.8 and its corresponding target were transiently expressed. However, transient expression is not always an option, due to the experimental timescale, necessity for a precise expression profile, or the use of a recalcitrant mammalian cell line. Thus, to demonstrate the flexibility of GS2-IpaH9.8-mediated silencing, we evaluated degradation activity against target proteins that were expressed as stably integrated transgenes. Specifically, when GS2-IpaH9.8 was transiently expressed in cells that stably co-expressed EGFP, reduction of fluorescence activity was virtually identical to that observed for transiently expressed EGFP (**Fig. 2.9a**). Robust degradation was also observed for ERK2-EGFP, H2B-EGFP, and EGFP-HRas^{G12V}, regardless of their mode of expression (**Fig 2.9a**). When the ubiquibody and the target were both expressed as stable transgenes, thereby eliminating the need for transfection entirely, strong silencing activity was again observed for GS2-IpaH9.8 but not its inactive GS2-IpaH9.8^{C337A} counterpart (**Fig. 2.9b**).

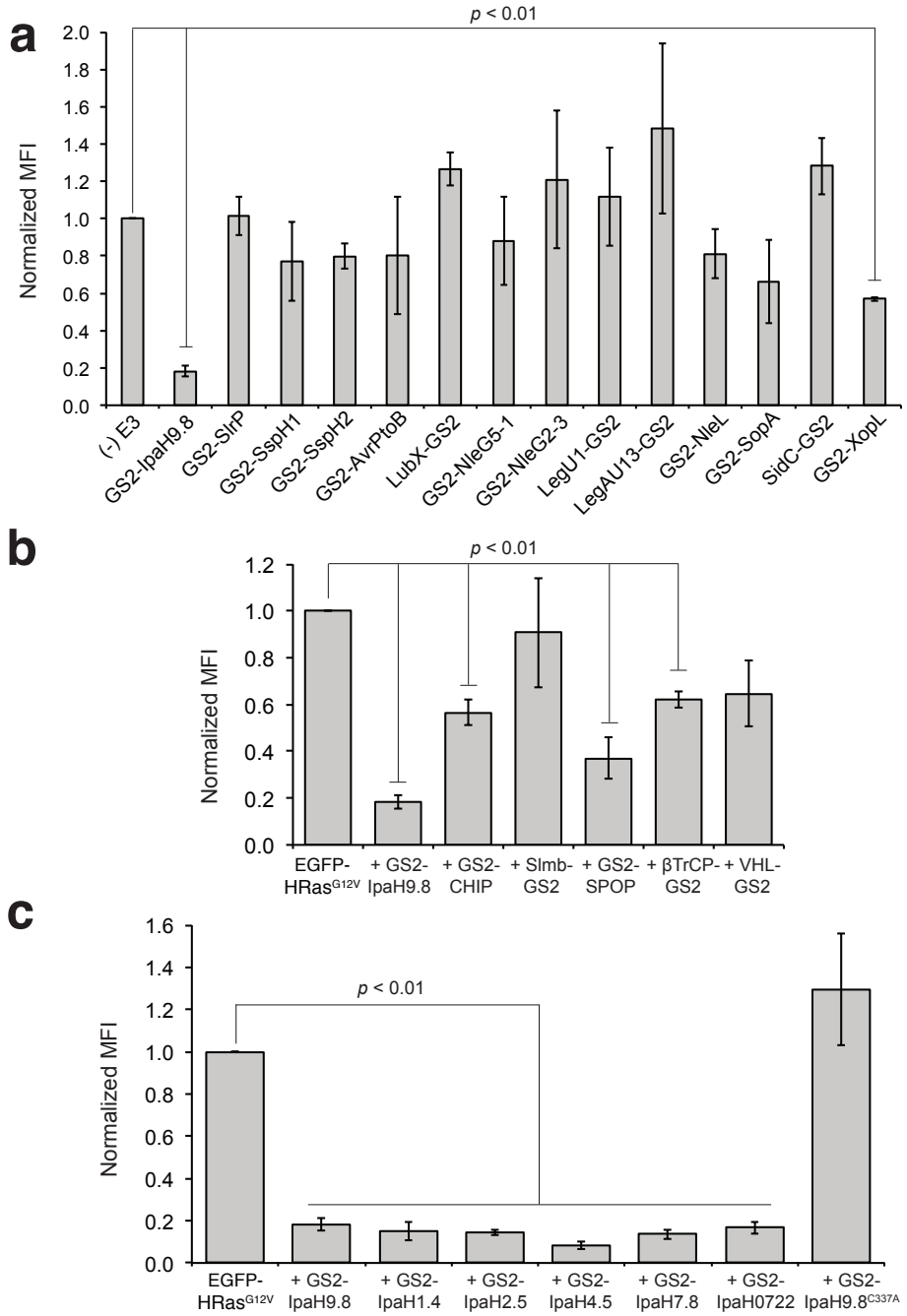


Figure 2.8 EGFP-HRas^{G12V} silencing by bacterial- and mammalian-based ubiquibodies. (a) Flow cytometric analysis of EGFP fluorescence activity in HEK293T cells transfected with plasmid pcDNA3-EGFP-HRas^{G12V} alone or co-transfected with pcDNA3-EGFP-HRas^{G12V} and a plasmid encoding a ubiquibody comprised of GS2 fused to one of the (a) bacterial or (b) mammalian E3 ubiquitin ligases as indicated. Values for geometric mean fluorescence intensity (MFI) are shown. (c) Same as in (a) but with plasmid DNA encoding a ubiquibody comprised of GS2 fused to one of the IpaH homologs as indicated. Data are biological triplicates (three separately transfected wells) of the geometric mean fluorescence intensity (MFI) normalized to MFI measured for HEK293T cells expressing EGFP-HRas^{G12V} alone. Error bars represent standard deviation (SD) of the mean. p values were determined by paired sample t-test.

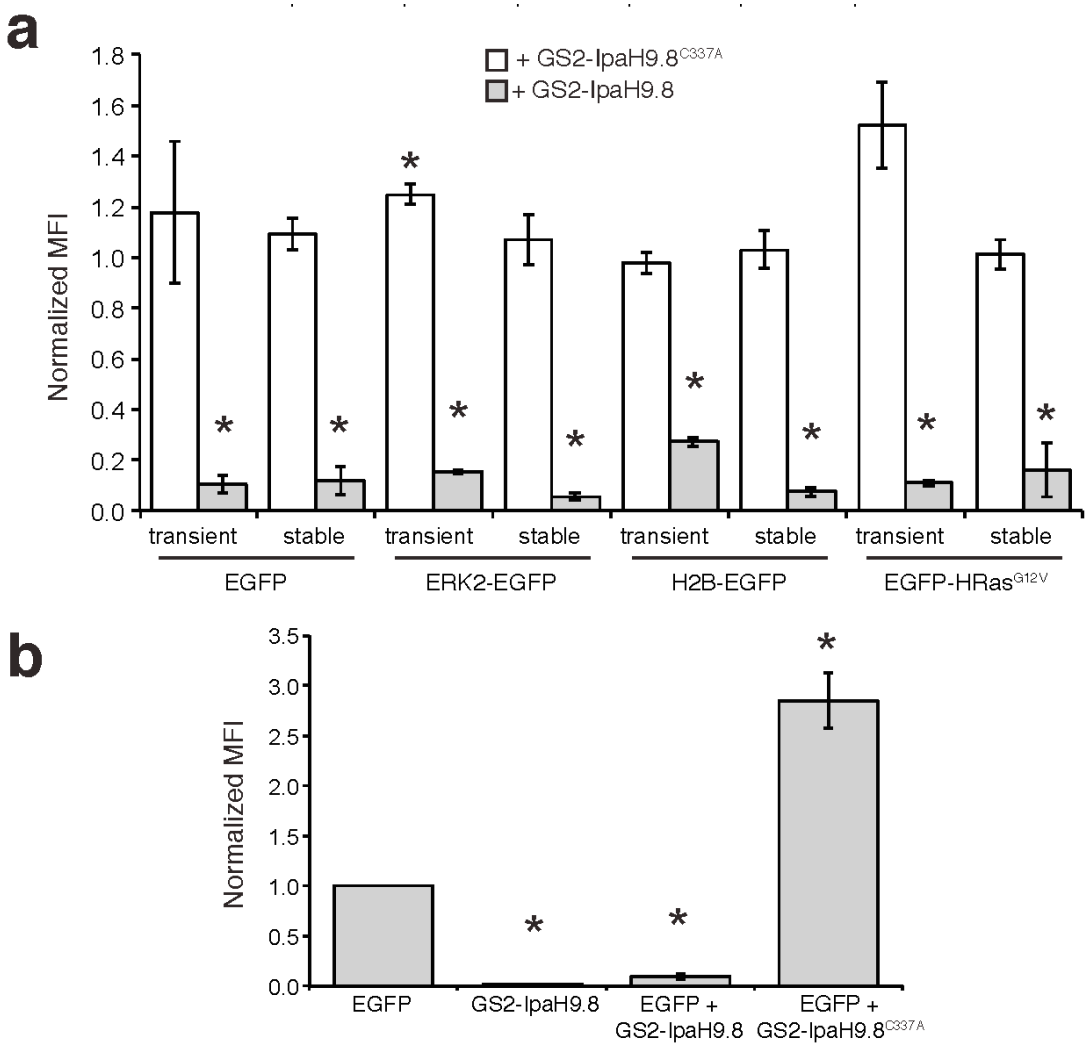


Figure 2.9 Effect of expression modality on GS2-IpaH9.8 efficacy.
(a) Flow cytometric quantification of EGFP fluorescence activity in HEK293T cells that transiently or stably expressed EGFP, ERK2-EGFP, H2B-EGFP, or EGFP-HRas^{G12V} as indicated. In all cases, cells were transiently transfected with plasmid DNA encoding either pcDNA3-GS2-IpaH9.8^{C337A} (white) or pcDNA3-GS2-IpaH9.8 (grey). (b) Flow cytometric quantification of EGFP fluorescence activity in MCF10a cells stably integrated with DNA encoding only EGFP-HRas^{G12V}, EGFP- HRas^{G12V} and GS2-IpaH9.8, EGFP- HRas^{G12V} and GS2-IpaH9.8^{C337A}, or GS2-IpaH9.8 alone. All data are biological triplicates of the geometric MFI normalized to MFI measured for HEK283T cells expressing the EGFP alone. Error bars represent standard deviation (SD) of the mean. *, $p < 0.01$ (compared to HEK293T cells expressing the FP fusion only).

Discussion

Ubiquibodies are a relatively new proteome editing modality that enable selective removal of otherwise stable proteins in somatic cells [57], with potential applications in basic research, drug discovery, and therapy. In this study, we created a new class of ubiquibodies that feature bacterial E3 ubiquitin ligases, thereby opening the door to a previously untapped source of ubiquitination activity for ubiquibody development. Specifically, we evaluated 14 bacterial E3 ligases belonging to a growing class of effector proteins that mimic host cell E3 ligases to exploit the ubiquitination pathway [71, 91]. Most notable among these was IpaH9.8 from S. flexneri, which proved to be a remarkable catalyst of protein turnover when directed to target substrates via a genetically fused synthetic binding domain. This silencing activity was found to be independent of the substrate's subcellular localization (i.e., cytoplasm, nucleus, plasma membrane) or expression modality (i.e., transient versus stable). The only other E3 ligases that functioned comparably were homologs of IpaH9.8 found in S. flexneri, either on the pWR100 virulence plasmid or the chromosome [71]. The N-terminal catalytic NEL domains of these enzymes share striking homology (99-100%), which explains their similar performance in the ubiquibody context. Accordingly, the next best functioning bacterial E3 ubiquitin ligase was S. typhimurium SspH1, which is also a NEL type enzyme with 38% identity to IpaH9.8 overall and 42% identity within the NEL domain [90]. It should also be pointed out that none of the mammalian E3 ubiquitin ligases were able to degrade EGFP levels below 60% under the conditions tested here. While the reasons for this are not entirely clear, given the successful knockdown results reported previously for these different E3 ligases in the ubiquibody format [57, 59-61, 65, 66], it is possible that EGFP may represent

a poor substrate for these engineered chimeras. We believe that the inherent conformational flexibility required to ubiquitinate these structurally diverse substrates helps to explain the NEL motif's remarkable ability for customizable target degradation. It should also be pointed out that while the work here leveraged previously confirmed E3 ubiquitin ligases, an analogous swapping strategy could be used to create GS2-based ubiquibodies for identifying novel E3 ligases. Such an approach could enable systematic identification of E3 ligases, which is an important objective given that the human genome encodes over 600 putative E3 ligases [102] and bacterial genomes likely encode hundreds of others, many of which remain to be validated as catalysts of ubiquitin transfer.

From a drug development standpoint, pharmacological control of gene products has traditionally been achieved using small molecule inhibitors that target enzymes and receptors having well-defined hydrophobic pockets where the small molecules are tightly bound. Unfortunately, a majority (~80-85%) of the human proteome is comprised of intractable targets, such as transcription factors, scaffold proteins, and non-enzymatic proteins, which have yet to be inhibited pharmacologically and thus have been deemed 'undruggable' [23, 24]. As an alternative, a number of techniques for silencing proteins at the DNA or RNA level are now available such as CRISPR, RNAi, TALENs, and ZFNs, with the first RNAi therapy, patisiran, gaining approval in 2018 for hereditary transthyretin amyloidosis [103]. Nonetheless, new adaptable technologies, such as ubiquibodies and the related PROTACs technology, that offer temporal and post-translational control over protein silencing are desirable especially because of their potential to overcome some of the limitations associated with nucleic acid targeting-based approaches such as irreversibility, lack of temporal control, and off-target effects [19, 104-106]. In principle,

both ubiquibodies and PROTACs can degrade proteins regardless of their function, including the currently undruggable proteome. Moreover, unlike conventional 'occupancy-based' therapeutics, ubiquibodies and PROTACs act catalytically, making them substantially more potent than the target-binding antibody mimetics and small molecule inhibitors, respectively, from which they are built.

Materials and Methods

Plasmids

All plasmids used in this study are provided in **Appendix C**. *E. coli* strain DH5α was used for the construction and propagation of all plasmids. To construct pcDNA3-EGFP, EGFP was PCR amplified using primers that introduced a 5' Kozak sequence and the resulting PCR product was ligated into pcDNA3. Plasmid pCDH1-ERK2-EGFP was created by gene assembly of ERK2 and EGFP using overlap extension PCR with primers that introduced a 5' Kozak sequence followed by ligation into pCDH1. Plasmid pcDNA3-EGFP-NLS was created by PCR amplification of EGFP with primers that added a 5' Kozak sequence and 3' SV40 NLS sequence and then ligation of the PCR product into pcDNA3. Plasmid pcDNA3-SHP2-EGFP was created by PCR amplification of SHP2 with a 5' Kozak sequence followed by ligation into pcDNA3-EGFP. Plasmid pcDNA3-EGFP-HRas^{G12V} was generated by PCR amplification of EGFP-HRas^{G12V} from plasmid mEGFP-HRas^{G12V} and the PCR product was subsequently ligated into pCDH1.

For creation of GFP-directed ubiquibodies, plasmid pcDNA3-HF-GS2 was created by PCR amplification of GS2 from pHFT2-GS2 [95] using primers that introduced upstream Kozak, 6x-His, and FLAG sequences followed by ligation into pcDNA3 such that BamHI and EcoRI restriction sites were available upstream of GS2 for generating N-

terminal fusions. For C-terminal fusions, plasmid pcDNA3-GS2-FH was created by PCR amplifying GS2 with primers that introduced an upstream Kozak sequence and downstream NheI and SbfI restriction sites followed by ligation into pcDNA3. The genes encoding AvrPtoB, IpaH9.8, NleG2-3, NleG5-1, NleL, SlrP, SopA, SPOP, SspH1, SspH2, and XopL were PCR amplified with primers introducing NheI and SbfI sites, after which the resulting PCR products were ligated in pcDNA3-GS2-FH. The genes encoding LegAU13, LegU1, and SidC were PCR amplified with primers that introduced BamHI and EcoRI sites, after which the resulting PCR products were ligated in pcDNA3-HF-GS2. Plasmid pcDNA3-GS2-CHIP was created by PCR amplification of GS2 from pHFT2-GS2 using primers that introduced an upstream HindIII and Kozak sequence and downstream NheI site, followed by ligation into pcDNA3-R4-CHIPdTPR in place of scFvR4. Plasmids pcDNA3-VHL-GS2 and pcDNA3-βTrCP-GS2 were created by PCR amplification of genes encoding VHL and βTrCP with primers that introduced HindII and XhoI (VHL) or BamHI and XhoI (βTrCP) sites after which the resulting PCR products were ligated in place of NSlmb in pcDNA3-NSlmb-GS2. Plasmids pcDNA3-GS2-IpaH9.8^{C337A}, pcDNA3-GS2-IpaH0722, pcDNA3-GS2-IpaH1.4, pcDNA3-GS2-IpaH2.5, pcDNA3-GS2-IpaH4.5, and pcDNA3-GS2-IpaH7.8 were created by site-directed mutagenesis of pcDNA3-GS2-IpaH9.8. The following genes were purchased: SspH1 (Twist Biosciences), IpaH9.8 (Twist Biosciences), VHL (GenScript, Ohu23297D), LubX (Twist Biosciences), LegU1(IDT), and LegAU13 (IDT). All others were amplified from existing plasmids in our laboratory stocks or from genomic DNA.

Plasmid pET24d-GS2-IpaH9.8 and pET24d-IpaH9.8dLRR were created by PCR amplification of full-length GS2-IpaH9.8 and truncated IpaH9.8dLRR, respectively, with

primers that introduced NcoI and NotI (GS2-IpaH9.8) or NheI and NotI (IpaH9.8dLRR) sites, after which the resulting PCR products were ligated into pET24d(+). Plasmid pET28a-GS2 was created by PCR amplification of GS2 from pHFT2-GS2 using primers that introduced an upstream NcoI site and downstream FLAG, 6x-His, and HindIII sequences, after which the resulting PCR product was ligated into pET28a(+). Plasmid pTriEx-3-GS2-IpaH9.8^{C337A} was created by PCR amplification of GS2-IpaH9.8^{C337A} from pcDNA3-GS2-IpaH9.8^{C337A} with primers that introduced EcoRV and HindIII sites, after which the resulting PCR product was ligated into pTriEx-3. Plasmid pET28a-EGFP was created by PCR amplification of GFP with primers adding C-terminal 6x-His tag, after which the resulting PCR product was ligated in pET28a(+). All plasmids were verified by DNA sequencing at the Cornell Biotechnology Resource Center (BRC).

Cell Lines, Culture and Transfection

All cell lines used in this study are provided in **Appendix C**. Briefly, HEK293T and HeLa cells were obtained from ATCC, HeLa H2B-EGFP cells were kindly provided by Elena Nigg, and MCF10A rtTA cells were kindly provided by Matthew Paszek, HEK293T, HeLa, and HeLa H2B-EGFP cells were cultured in DMEM with 4.5 g/L glucose and L-glutamine (VWR) supplemented with 10% FetalCloneI (VWR) and 1% penicillin-streptomycin-amphotericin B (ThermoFisher) MCF-10a cells were grown in DMEM/F12 media (ThermoFisher) supplemented with 5% horse serum (ThermoFisher), 20 ng/mL EGF (Peprotech), 0.5 mg/mL hydrocortisone (Sigma), 100 ng/mL cholera toxin (Sigma), 10 μg/mL insulin (Sigma), and 1% penicillin-streptomycin-amphotericin B (ThermoFisher). All cells were maintained at 37°C, 5% CO_2 and and 90% relative humidity (RH).

Stable MCF10A rtTA cell lines were generated using Nucleofection Kit V (Lonza)

and HyPBase, an expression plasmid for the hyperactive version of the PiggyBac transposase. Transposition of MCF10A rtTA cells was performed to generate the following stable lines: MCF10A EGFP-HRas^{G12V}; MCF10A EGFP-HRas^{G12V}:GS2-IpaH9.8; MCF10A EGFP- HRas^{G12V}:GS2-IpaH9.8^{C337A}; and MCF10A GS2-IpaH9.8. Stable cell lines were selected using 200 μg/mL hygromycin B (ThermoFisher).

Stable HEK293T cell lines expressing EGFP, EGFP-HRas^{G12V}, ERK2-EGFP, d2EGFP were generated by lentiviral transformation. Specifically, pLV-IRES-eGFP, pcDH1-EGFP-HRas^{G12V}, pcDH1-ERK2-EGFP, or pHIV-d2EGFP were transfected into HEK293T cells along with psPAX2 and pMD2.G by calcium phosphate transfection. Media was replaced after ~16 h, followed by a 48-h incubation to allow virus production. Viral supernatant was removed and Polybrene (Sigma-Aldrich) added to a final concentration of 8 μg/mL, followed by clearance of cell debris by centrifugation at 2,000 rpm for 5 min. Resultant supernatant was diluted 1:6 with cell media and added to previously plated HEK293T cells for stable integration. HEK293T-EGFP and HEK293T ERK2-EGFP cell lines were selected by fluorescence activated cell sorting (BD FACSAria). The HEK293T-EGFP-HRas^{G12V} cell line was selected using 1 μg/mL puromycin (Sigma-Aldrich).

Western blot analysis

HEK293T cells were plated at 10,000 cells/cm^2 and transfected as describedabove before lysis with RIPA lysis buffer (Thermo Fisher). Lysates were separated on Any kD polyacrylamide gel (Bio-Rad) and transferred to PVDF membranes. α-HIS-HRP (Abcam), α-GFP (Krackeler) and α-GAPDH (Millipore) antibodies were diluted 1:5,000 and in TBST + 1% milk and incubated for 1 h at room temperature. Secondary antibody goat anti-

mouse IgG with HRP conjugation (Promega) was diluted at 1:2,500 and used as needed.

Microscopy

Cells were plated at 10,000 cells/cm^2 on a glass bottom 12-well plate pre-treated with poly-L-lysine (Sigma-Aldrich). After seeding for 16-24 h, cells were transfected with 1 μg total DNA at a 1:2 ratio of DNA:jetPrime (Polyplus Transfection). cells were transfected with 0.05 μg of target, 0.25 μg of ubiquibody or control, and balanced with empty pcDNA3 vector. Culture media was replaced 4-6 h post-transfection. Then, 24 h post-transfection, cells were fixed with 4% paraformaldehyde. For EGFR-EGFP samples, cells were blocked with 5% normal goat serum in PBS for 2 h at room temperature. The anti-EGFR antibody (Cell Signalling #4267) was diluted 1:200 in 5% normal goat serum in PBS and incubated overnight at 4°C. Cells were washed three times with PBS, then incubated for 1 h at room temperature with anti-rabbit-AF647 diluted 1:200 in 5% normal goat serum in PBS. Cells were washed three times with PBS. Cell nuclei were stained with Hoescht diluted 1:10,000 in PBS for 10 min, then washed three times in PBS. Samples were imaged on a inverted Zeiss LSM88-confocal/multiphoton microscope (i880) using a 40x water immersion objective. Images were analyzed with FIJI.

Flow cytometric analysis

Cells were passed into 12-well plates at 10,000 cells/cm^2. 16-24 h after seeding, cells were transiently transfected with 1 μg total DNA at a 1:2 ratio of DNA:jetPrime (Polyplus Transfection). Cells were transfected with 0.05 μg of target, 0.25 μg of ubiquibody or control, and balanced with empty pcDNA3 vector. Culture media was replaced 4-6 h post-transfection. Then, 24 h post-transfection, cells were harvested and resuspended in phosphate buffered saline (PBS) for analysis using a BD FACSCalibur or

BD FACSAria Fusion. FlowJo Version 10 was used to analyze samples by geometric mean fluorescence determined from 10,000 events.

Protein expression and purification

Purified proteins were obtained by growing *E. coli* BL21(DE3) cells containing a pET28a-based plasmid encoding the desired protein or Rosetta(DE3) cells containing a pTriEx-3-based plasmid in 200 mL of Luria-Bertani (LB) medium at 37°C. Expression was induced with 0.1 mM IPTG when the culture density (Abs600) reached 0.6-0.8 and growth continued for 6 h at 30°C. Cultures were harvested by centrifugation at 4,000xg for 30 min at 4°C. Cell pellets were stored at -20°C overnight. Thawed pellets were resuspended in 10 mL equilibration buffer (25 mM Tris-HCl, pH 7.4, 500 mM NaCl and 20 mM imidazole) and lysed with a high-pressure homogenizer (Avestin Emulsi-Flex C5). The insoluble fraction was cleared by centrifugation at 12,000xg for 30 min at 4°C. His-tagged protein was purified by gravity flow using 500 µL HisPur Ni-NTA resin (ThermoFisher). The soluble fraction was passed through the resin, after which the resin was washed with 3 mL of wash buffer (25 mM Tris-HCl, pH 7.4, 500 mM NaCl and 50 mM imidazole). Protein was eluted with 1.5 mL elution buffer (25 mM Tris-HCl, pH 7.4, 500 mM NaCl and 250 mM imidazole). Purified fractions were desalted and concentrated (Pierce PES Protein Concentrators).

ELISA

A 96-well enzyme immunoassay plate was coated with 100 µL of EGFP at 10 µg/mL in 0.05 M NaCO3 buffer, pH 9.6 at 4°C overnight. The plate was then washed three times 200 µL PBST (1x PBS + 0.1% Tween 20) per well and blocked with 250 µL PBS

with 3% milk per well at room temperature for 3 h, slowly mixing. The plate was washed three more times, followed by the addition of serial dilutions of purified proteins in blocking buffer at 60 µL per well. Plate was incubated at room temperature, slowly mixing for 1 h. The plate was washed three times to remove un-bound protein and then incubated with 50 µL/well of anti-FLAG (DDDYK) antibody conjugated to horseradish peroxidase (HRP) diluted 1:10,000 in PBST + 1% milk for 1 h with slow mixing. The plate was washed three times before the addition of 50 µL/well 1-Step Ultra TMB (3,3',5,5'-tetramethylbenzidine) (ThermoFisher). The reaction was incubated with slow mixing and then quenched with 50 µL/well of 3N H2SO4 and read at 450 nm.

Acknowledgements

We thank Anne Messer, Byung Kangwon, Cam Patterson, Elena Nigg, Gregory Martin, Karel Svoboda, Markus Affolter, Matthew Paszek, Melanie Cobb, Pengbo Zhou, Didier Trono, Valerie Weaver, Jon Lakins, Alan Bradley, Michael Davidson, Jan Lammerding, and Smita Nair for kindly providing plasmids encoding genes used in this study. We also thank Yuxin Mao for providing *L. pneumophila* gDNA, Yung Fu for providing EHEC and *S. typhimurium* gDNA, and Adam Bogdanave for providing *X. campestris* gDNA. We also thank Alyse Portnoff and Erin Stephens for the preliminary development of ubiquibodies and insightful conversations pertaining to these experiments. This work was supported by the National Science Foundation Grant CBET-0449080 (to M.P.D.), the National Institutes of Health Grant Numbers CA132223 (to M.P.D.), CA170820 and GM083898 (to R.W.R.), CA194864 and CA212608 (to S.K.) the New York State Office of Science, Technology and Academic Research Distinguished Faculty Award (to M.P.D.), the Cornell Technology Acceleration and Maturation (CTAM) Fund, the

National Science Foundation Graduate Research Fellowship Program Grants DGE-1650441 (to M.B.L.), Cornell Presidential Life Science Fellowships (to M.B.L.) a Cornell Fleming Graduate Scholarship (to M.B.L.), the Koch Institute Support Grant P30-CA14051 from the National Cancer Institute and the Department of Defense Ovarian Cancer Research Program Teal Innovator Award (to P.T.H.). Confocal microscopy was supported by NYSTEM Grant CO29155 and NIH Grant S10OD018516 and FACS analysis was supported by the Cornell Biotechnology Resource Center and the David H. Koch Institute Flow Cytometry Core at MIT.

CHAPTER 3 – EXPANDING UBIQUIBODY TARGETING WITH ENGINEERED DESIGNER BINDING PROTEINS

Introduction

Previously, we demonstrated the highly modular architecture of ubiquibodies – interchanging both the E3 catalytic domain and designer binding domain identified the ubiquibody GS2-IpaH9.8 that potently eliminated a broad-spectrum of GFP tagged proteins. To explore the therapeutic potential of our ubiquibodies, we wanted to expand the specificity of IpaH9.8 ubiquibodies beyond GFP. While there have been extensive numbers of binding domains isolated against disease-related proteins, many of these recognize extracellular protein epitopes which are not compatible with our ubiquibodies which require access to the intracellular ubiquitin proteasome pathway [107-111]. We were able to evaluate several well-characterized binding domains as ubiquibody fusions against the intracellular disease-relevant proteins SHP2, Ras, and ERK2.

When IpaH9.8 was paired with the SHP2-specific FN3 NSa5, efficient silencing of SHP2-EGFP was achieved that was comparable to that observed from the GFP-specific ubiquibody GS2-IpaH9.8. IpaH9.8 was also successfully reprogrammed to degrade EGFP-KRas when fused to the KRas-specific FN3s RasInI and RasInII as well as the KRas-specific sso7d R11.1.6. Furthermore, all three Ras-specific ubiquibodies – RasInI-IpaH9.8, RasInII-IpaH9.8, and R11.1.6-IpaH9.8 – showed preference for the cancer-prevalent isoforms $KRas^{G12C}$, $KRas^{G12D}$, and $KRas^{G12V}$ over wild-type KRas.

Interestingly, minimal elimination of ERK2-EGFP was observed when IpaH9.8 was paired with a panel of ERK2-specific DARPins. We hypothesized that the DARPin

scaffold might not be the ideal fusion partner for IpaH9.8, so we utilized yeast surface display to isolate ERK2-specific members from a naïve FN3 library. Twenty clones were then analyzed for their ability to degrade ERK2 in mammalian cells when fused to IpaH9.8. Interestingly, despite not exhibiting the strongest binding activity, clones aERK2-10 and aERK2-16 were the most potent ubiquibodies, depleting ERK2-eGFP to ~50%, a significant improvement over DARPin-based ERK2-specific ubiquibodies. Together, these results further emphasize the modularity of our ubiquibody design and the flexibility of IpaH9.8 to accommodate a range of binding domain scaffolds and protein targets. .

Results

Targeting disease relevant proteins with IpaH9.8-based ubiquibodies

To redirect IpaH9.8 activity towards disease-relevant proteins, we first investigated the compatibility of the IpaH9.8 catalytic domain with three different FN3s – NSa1, NSa5, and CS3 – which are specific for the Src-homology 2 (SH2) domain of SHP2 [112]. The resulting NSa1-IpaH9.8, NSa5-IpaH9.8, and CS3-IpaH9.8 chimeras were tested for their ability to silence SHP2-EGFP by flow cytometric analysis. NSa1-IpaH9.8 elicited no significant degradation and CS3-IpaH9.8 resulted in moderate, but highly variable depletion of SHP2-EGFP. However, NSa5-IpaH9.8 exhibited strong silencing activity, degrading SHP2-EGFP as efficiently as the GFP-specific GS2-IpaH9.8 ubiquibody (**Fig. 3.1**).

Next, we evaluated how IpaH9.8 performed when paired with Ras-specific FN3s RasInI and RasinII [113], as well as the Ras-specific sso7d R11.1.6 [114]. These three Ras-specific ubiquibodies – RasInI-IpaH9.8, RasInII-IpaH9.8, and R11.1.6-IpaH9.8 –

were tested for their ability to silence EGFP fusions of KRas, $KRas^{G12C}$, $KRas^{G12D}$, and $KRas^{G12V}$. R11.1.6-IpaH9.8 facilitated the most robust depletion of all four KRas isoforms, displaying preference for the G12C and G12D isoforms (**Fig. 3.2**). This preference was expected, as R11.1.6 was specifically selected to exhibit selectivity for G12D vs WT KRas specificity [114]. The extent of degradation did not strictly correlate to affinity, however, as $KRas^{G12V}$ depletion matched that of $KRas^{WT}$, despite R11.1.6's reported higher affinity for that isoform ($KRas^{WT}$ = 41 nM, $KRas^{G12C}$ = 21 nM, $KRas^{G12D}$ = 4 nM, $KRas^{G12V}$ = 12 nM [114]). RasInI-IpaH9.8 and RasInII-IpaH9.8, while not as potent, also demonstrated preference for the $KRas^{G12C, G12D, and G12V}$ mutant isoforms over $KRas^{WT}$, in line with their previously characterized specificities [113]. Together, these results demonstrate the powerful modularity of the ubiquibody design – simply swapping the GFP-specific GS2 binding domain with the protein binders NSa5 and R11.1.6 enabled potent depletion of their respective disease-relevant targets - SHP2 and KRas, respectively.

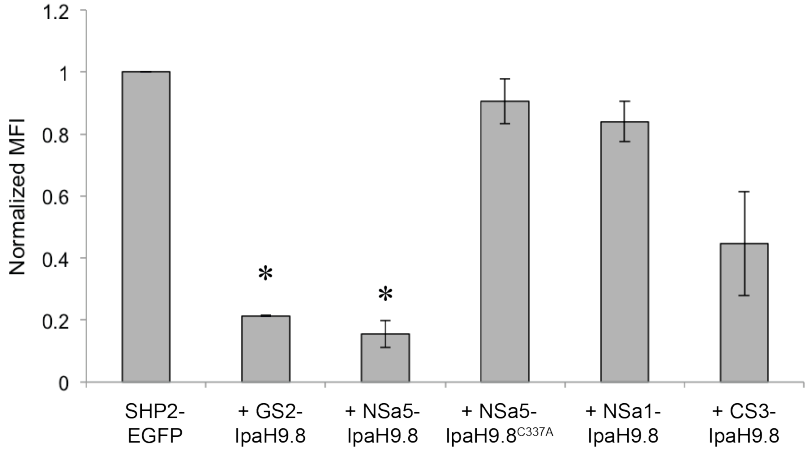


Figure 3.1 IpaH9.8 ubiquibodies against SHP2.
Flow cytometric quantification of EGFP fluorescence activity in HEK293T cells transfected with pcDNA3-SHP2-EGFP alone or cotransfected with pcDNA3-SHP2-EGFP and a plasmid encoding the indicated ubiquibody. Data are biological triplicates (three separately transfected wells) of the geometric MFI normalized to MFI measured for HEK293T cells expressing SHP2-EGFP alone. Error bars represent standard deviation (SD) of the mean. *, $p < 0.01$ (compared to HEK293T cells expressing SHP2-EGFP only).

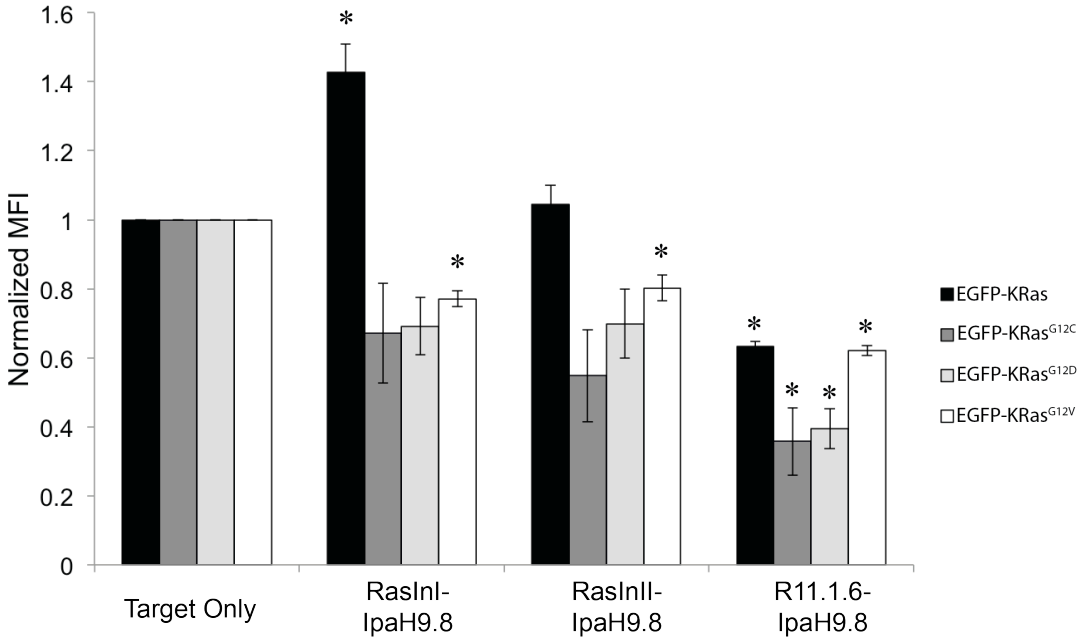


Figure 3.2 IpaH9.8 ubiquibodies against KRAS mutants.
Flow cytometric quantification of EGFP fluorescence activity in HEK293T cells transfected with pcDNA3-EGFP-KRas, pcDNA3-EGFP-KRasG12C, pcDNA3-EGFP-KRasG12D, or pcDNA3-EGFP-KRas^{G12V} alone or cotransfected with a plasmid encoding the indicated ubiquibody. Data are biological triplicates (three separately transfected wells) of the geometric MFI normalized to MFI measured for HEK293T cells expressing the indicated KRas protein alone. Error bars represent standard deviation (SD) of the mean. *, $p < 0.01$ (compared to HEK293T cells expressing the FP fusion only).

Isolation of novel ERK2-specific FN3s for improved ubiquibody targeting

We have previously characterized a panel of ERK2 targeting ubiquibodies, which fused pERK/ERK2-specific DARPins with the catalytic domain of the mammalian E3 ubiquitin ligase CHIP [115]. When these ubiquibodies were evaluated for their ability to degrade the fluorescent protein fusion ERK2-EGFP in mammalian cells, moderate degradation activity (~60-80% ERK2-EGFP remaining) was indeed observed, however co-expression of our GFP-targeting ubiquibody GS2-IpaH9.8 resulted in substantially higher depletion (~25% ERK2-EGFP remaining) (**Fig. 3.3**). Knowing that GS2-IpaH9.8 more potently depleted EGFP than GS2-CHIP (**Fig. 2.1c**),

we hypothesized that pairing the ERK2-specific DARPins with IpaH9.8 would improve our degradation efficiency of ERK2-EGFP. However, this was not true – in fact, only EpE89-IpaH9.8 showed moderate improvement over its CHIP counterpart (**Fig. 3.3**).

Consequently, these results led us to believe that minimally, IpaH9.8 does not pair well with the DARPin scaffold, but potentially ERK2 cannot be targeted with our current ubiquibodies. Combining our knowledge that IpaH9.8 exhibited potent activity against a variety of targets (EGFP, SHP2, and Ras) when fused to FN3 domains and our continued interest in evaluating the role of affinity and binding epitope on ubiquibody efficacy, we sought to harness the power of yeast surface display to isolate ERK2-specific FN3s that we could evaluate for ubiquibody activity.

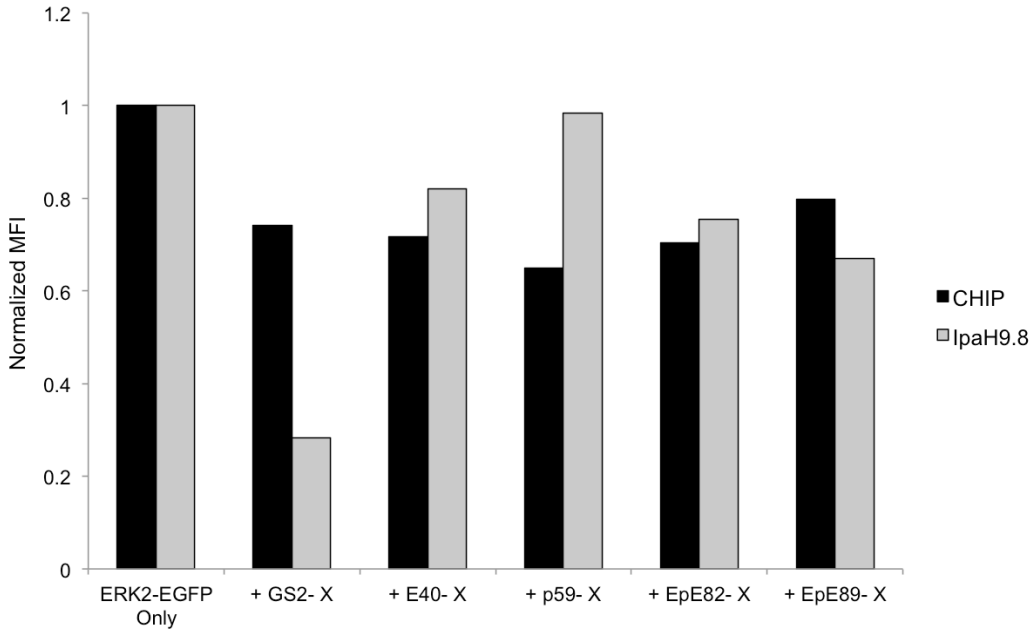


Figure 3.3 ERK2 degradation by DARPin-based ubiquibodies.
Flow cytometric quantification of EGFP fluorescence activity in HEK293T cells transfected with pcDNA3 ERK2-EGFP alone or cotransfected with pcDNA3-ERK2-EGFP and a plasmid encoding the indicated ubiquibody (e.g for GS2-X, black bar indicates GS2-CHIP and grey bar indicates GS2-IpaH9.8). Data are geometric MFI normalized to geometric MFI measured for HEK293T cells expressing ERK2-EGFP alone.

A naïve FN3 library [116] was screened twice using biotinylated ERK2 preloaded on Streptavidin magnetic beads, sorting by FACS of full-length clones, and then sorted by FACS for members exhibiting binding of ERK2 at 500 nM. The resultant population was then diversified by both gene and loop mutagenesis [116], followed by two additional rounds of isolation by FACS selections at 300 nM and 100 nM ERK2 (**Fig. 3.4**). Twenty clones were randomly selected for characterization. Sequencing analysis revealed that fifteen clones were unique – clones 8, 9, and 14 were unable to be sequenced and clones 1 and 4 were identical, as well as clones 2, 15, and 17. Those thirteen clones were then analyzed as individual cultures to confirm full-length expression and evaluate binding of ERK2 at 50 nM. Of the thirteen remaining clones, all the clones, with the exception of aERK2-5 showed binding above control (**Fig. 3.5**).

These hits were then cloned as IpaH9.8 fusions and evaluated for their ability to deplete ERK2-EGFP in mammalian cells. The ERK2-specific ubiquibodies elicited the reduction of ERK2-as determined by flow cytometry (**Fig. 3.6**). Specifically, aERK2-10-IpaH9.8 and aERK2-16-IpaH9.8 depleted ERK2-EGFP to ~50%, which was a substantial improvement over DARPin-based ubiquibodies. Analysis of the sequences revealed a high occurrence of the BC Loop sequence "YDPDGSD" and DE Loop sequence "FYHS". Interestingly, however, the two best clones as determined by ERK2-EGFP degradation, aERK2-10 and aERK2-16, had no similarity in their sequences.

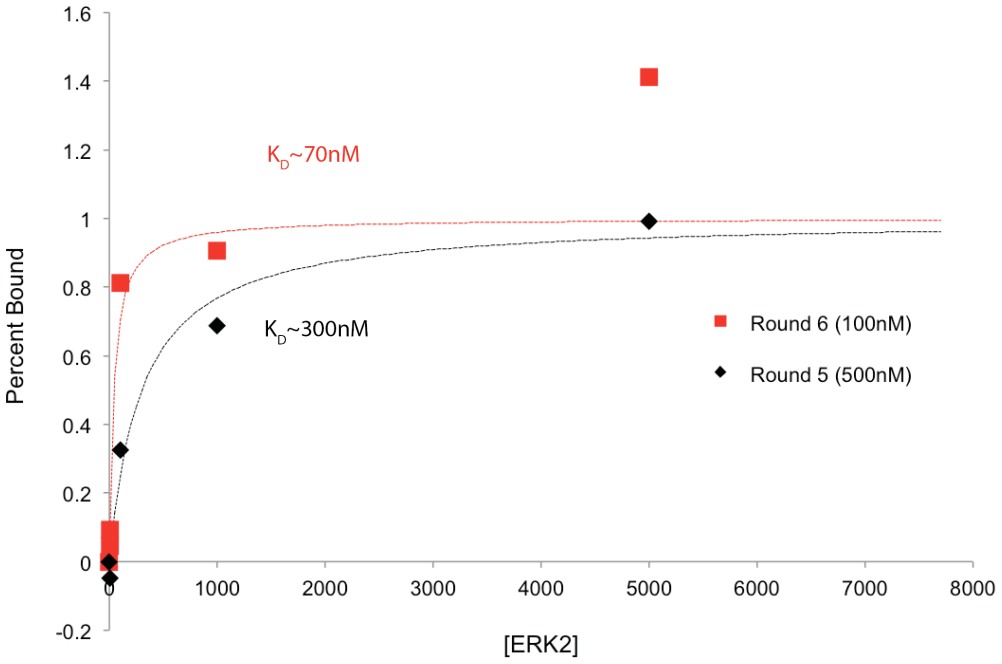


Figure 3.4 K_D measurement of ERK2-specific FN3 FACS libraries.
Flow cytometric quantification of ERK2 binding by sorted FN3 libraries. $5x10^6$ cells were incubated with biotinylated ERK2, followed by streptavidin-AF488. K_D was determined by fitting experimental data points by least squares regression.

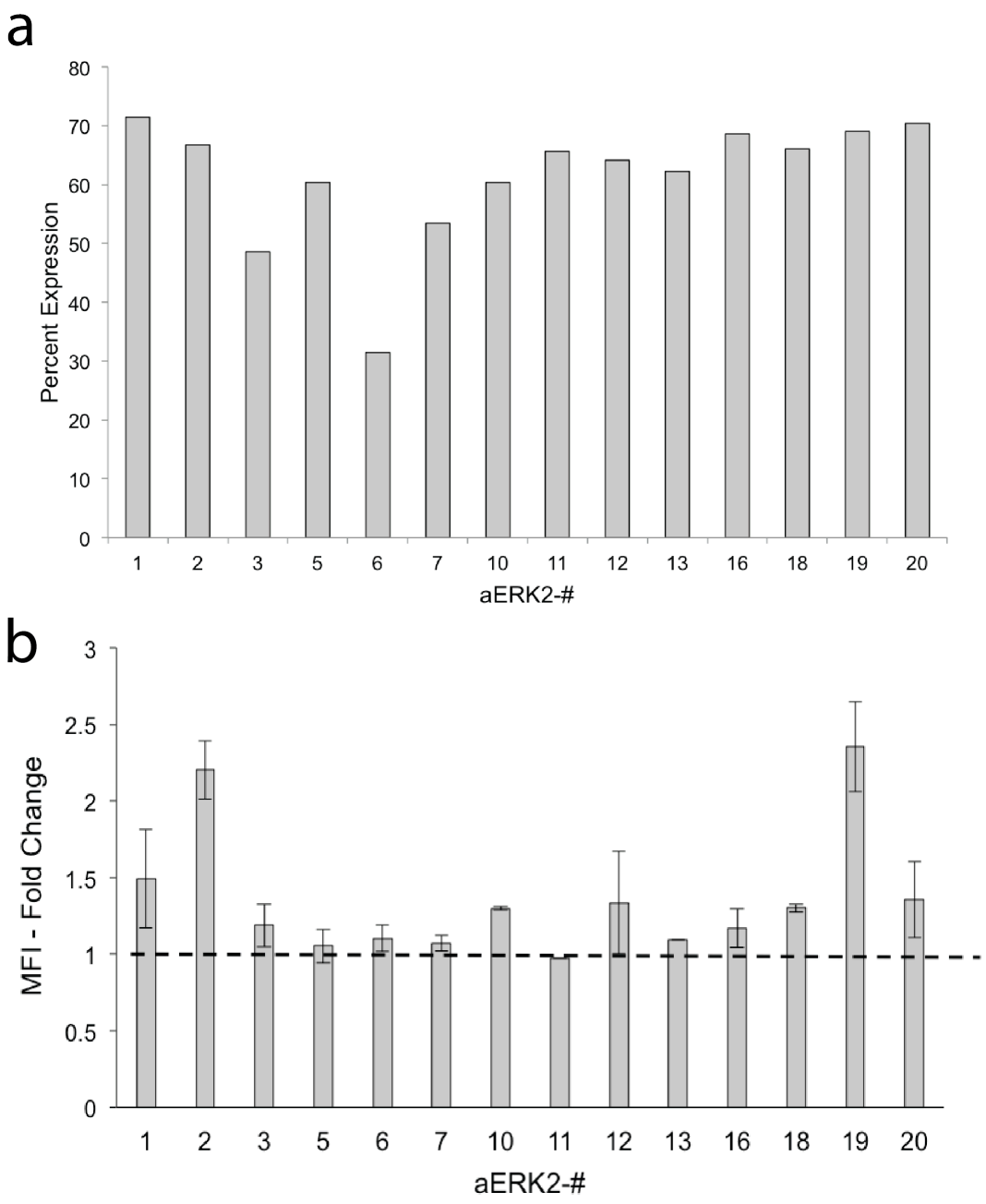


Figure 3.5 Characterization of ERK2-specific FN3 clones.
(a) Flow cytometric quantification of percentage of yeast cells expressing full-length FN3 clones on yeast surface as detected by α-cmyc antibody. (b) Flow cytometric quantification of ERK2 binding at 50 nM by individual FN3 clones displayed on yeast surface. Data are duplicates of geometric MFI normalized to MFI measured for binding observed by each clone at 0 nM ERK2. Error bars represent standard deviation (SD) of the mean. Horizontal dashed line indicates minimum threshold for positive binding signal.

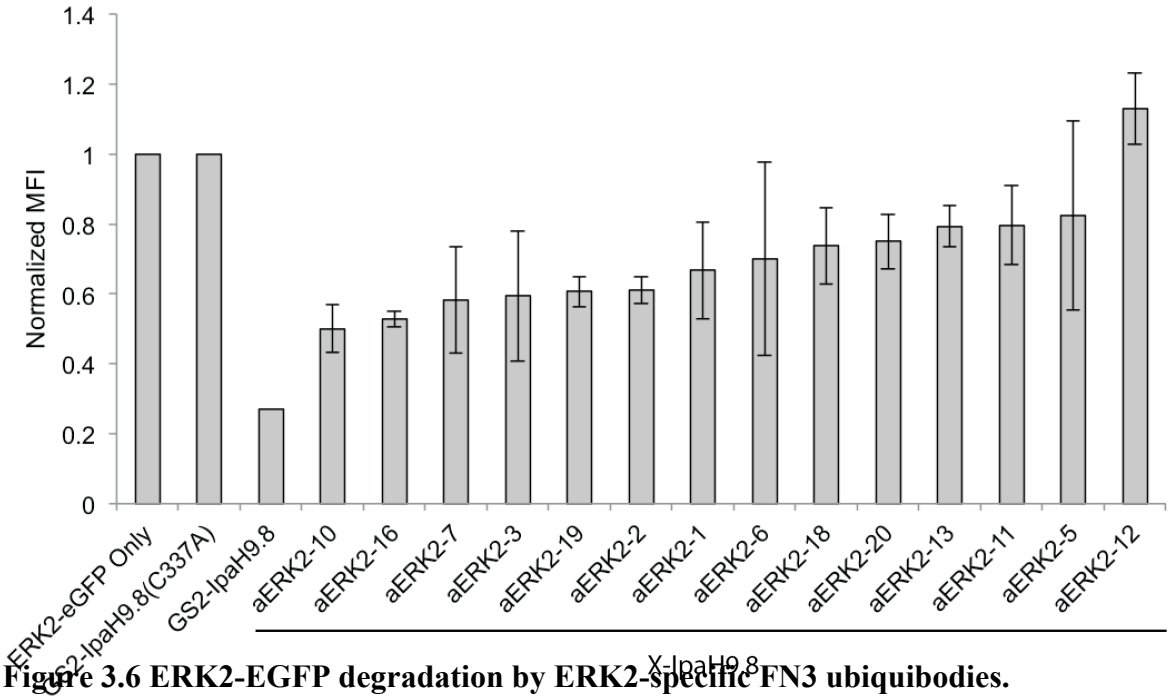


Figure 3.6 ERK2-EGFP degradation by ERK2-specific FN3 ubiquibodies.
Flow cytometric quantification of EGFP fluorescence activity in HEK293T cells transfected with pcDNA3-ERK2-EGFP alone or cotransfected with pcDNA3-ERK2-EGFP and a plasmid encoding the indicated ubiquibody (e.g. aERK2-10-IpaH9.8). Data are geometric MFI normalized to geometric MFI measured for HEK293T cells expressing ERK2-EGFP alone.

aERK2-#	BC Loop	DE Loop	FG Loop
10	Y D P D G S D -- -- --	F Y H S -- --	C Y D C D I S C S Y
16	S T P R Y Y Y L D S	G W C L F C	D H W S S D -- -- -- --
7	Q Y P C A Y S -- -- --	L L E T -- --	G R G E S E -- -- -- --
3	G H H F Y D S W -- --	G S R T -- --	F S Y F S F -- -- -- --
19	Y D P D G S D -- -- --	F Y H S -- --	F R S P S C -- -- -- --
2	Y D P D G S D -- -- --	F Y H S -- --	P G Y A N S S Y -- --
1	Y N R S V Y -- -- -- --	W Y Q S -- --	A C F D N D R C S C
6	Y M S C L A -- -- -- --	R S A T -- --	C R S N D Y S G -- --
18	Y N R S V Y -- -- -- --	W Y Q S -- --	D R D D G S N -- -- --
20	Y D P D G S D -- -- --	G S K S -- --	D T H G S C -- -- -- --
13	G S H C S Q S T -- --	G D S -- -- --	G Q D C S C -- -- -- --
11	Y P Y D G L S -- -- --	G S S T -- --	V D Y E C Y S S S L
5	T T S S F L Y -- -- --	G S K S -- --	D T H G S C -- -- -- --
12	C K P Y V Y -- -- -- --	A S M S -- --	F R R D S W -- -- -- --

Figure 3.7 Protein sequence alignments of loop regions from isolated ERK2-specific FN3s.
Fifteen unique clones are ordered according to their potency as ubiquibodies. Loop sequences that appear in multiple hits are highlighted.

Discussion

One of the advantages of the ubiquibody design is its modular architecture. Previously we demonstrated that choice of E3 ubiquitin ligase domain can dramatically affect catalytic activity. Here we further explored ubiquibody modularity – demonstrating the ease of reprogramming ubiquibodies to degrade disease-relevant targets such as SHP2, Ras, and ERK2, simply by swapping the synthetic designer binding protein. Furthermore, ubiquibodies retain the specificity of their designer binding protein, as in the case of R11.1.6-IpaH9.8 substrate preference for $KRas^{G12C}$ over $KRas^{WT}$. This high level of substrate discrimination is vital as we consider the use of ubiquibodies as therapeutic agents, as many diseases are driven by proteins with small amino acid mutations, as with $KRas^{G12C}$ and $KRas^{WT}$, or by post-translationally modified proteins.

Importantly, the ability to deplete these clinically important targets along with all of the other FP fusions (Chapter 3) serves to highlight the extraordinary modularity of the ubiquibody technology. Simply swapping the native substrate-binding domain of the E3 ubiquitin ligase can generate a made-to-order ubiquibody with specificity for a different substrate protein. Interestingly, Shigella have evolved a similar strategy for subverting host defenses during infection whereby plasmid and chromosomally-encoded IpaH proteins play a key role in dampening the host inflammatory response by mediating proteasomal degradation of NF-κB-related proteins [93, 117]. Specifically, by employing different LRR domains, which only share ~50% similarity [90], Shigella are able to redirect virtually identical catalytic NEL domains to an array of host proteins (e.g., NEMO, U2AF53 for IpaH9.8; Glomulin for IpaH7.8; p65 for IpaH4.5; HOIP for

IpaH2.5 and IpaH1.4; TRAF2 for IpaH0722) [71, 91, 117].

However, not all synthetic binding domains showed efficacy as ubiquibody fusions. While binding affinity is an obvious parameter to attribute these differences to, our limited data did not seem to support that hypothesis. Rather, we believe that while affinity plays a significant role in ubiquibody efficacy, other factors including the choice of binding domain scaffold and proximity of the target binding epitope to surface exposed lysines might be more predictive. For these reasons, we sought to isolate our own panel of ERK2-specific FN3 binders to test these hypotheses.

We did identify ERK2-specific binders that, when expressed as IpaH9.8 ubiquibody fusions, resulted in substantial depletion of ERK2-EGFP in mammalian cells. Interestingly, the clones that displayed the highest degradation activity as ubiquibodies, aERK2-10 and aERK2-16, did not demonstrate the highest binding activity as binding domains. This added more data to support our hypothesis that binding affinity is not the only parameter that defines ubiquibody potency. While further characterization of the isolated aERK2 FN3 clones is needed, particularly K_D calculation, these results underscore the importance for continued investigation into the binding domain requirements on efficient ubiquibody-mediated degradation.

Materials and Methods

Plasmids and strains

All plasmids used in this study are provided in **Appendix C**. *E. coli* strain DH5α was used for the construction and propagation of all plasmids. pcDNA3-GS2-IpaH9.8, pcDNA3-SHP2-EGFP, pCDH1-ERK2-EGFP was described previously (Chapter 2). To

construct pcDNA3-EGFP-KRas, a synthetic double-stranded gblock (IDT) encoding XhoI-KRas-NotI was ligated into pcDNA3-EGFP-HRas digested with XhoI and NotI. pcDNA3-EGFP-KRas^{G12C}, pcDNA3-EGFP-KRas^{G12D}, and pcDNA3-EGFP-KRas^{G12V} were created through site-directed mutagenesis of pcDNA3-EGFP-KRas.

pcDNA3-NSa1-IpaH9.8, pcDNA3-NSa5-IpaH9.8, and pcDNA3-CS3-IpaH9.8 were created by PCR amplification of NSa1, NSa5, and CS3 from pHFT2-SHP2-NSa1, pHFT2-SHP2-NSa5, and pHFT2-SHP2-CS3, respectively, with primers that introduced an N-terminal HindIII site and Kozak sequence and C-terminal NheI site, followed by ligation into pcDNA3-GS2-IpaH9.8 in place of GS2. Similarly, pcDNA3-NSa5-IpaH9.8^{C337A} was ligated into pcDNA3-GS2- IpaH9.8^{C337A} in place of GS2. pcDNA3-RasInI-IpaH9.8 and pcDNA3-RasInII-IpaH9.8 were created by PCR amplification of RasInI and RasInII from pET24a(+)-RasInI and pET24a(+)-RasInII, respectively, with primers that introduced an N-terminal HindIII site and Kozak sequence and C-terminal NheI site, followed by ligation into pcDNA3-GS2-IpaH9.8 in place of GS2. pcDNA3-R11.1.6-IpaH9.8 was created by PCR amplification of R11.1.6 from pE-Sumo-R11.1.6 with primers that introduced an N-terminal HindIII site and Kozak sequence and C-terminal NheI site, followed by ligation into pcDNA3-GS2-IpaH9.8 in place of GS2.

pcDNA3-E40-CHIP, pcDNA3-pE59-CHIP, pcDNA3-EpE82-CHIP, and pcDNA3-EpE89-CHIP were created as described previously [115]. pcDNA3-E40-IpaH9.8, pcDNA3-pE59-IpaH9.8, pcDNA3-EpE82-IpaH9.8, and pcDNA3-EpE89-IpaH9.8 were created by PCR amplification of E40, pE59, EpE82, and EpE89 with primers that introduced an N-terminal HindIII site and Kozak sequence and C-terminal NheI site, followed by ligation into pcDNA3-GS2-IpaH9.8 in place of GS2. All

plasmids were verified by DNA sequencing at the Cornell Biotechnology Resource Center (BRC).

Strain BL21(DE3) was used for cytoplasmic expression of ERK2 from pLK1_ERK2, which introduced an N-terminal Avi tags for biotinylation *in vivo* using pSPI02B-BirA-His and a C-terminal 6x-His tag for affinity purification and immunodetection. The Avi tags were used for avidin resin purification, after which 6 × -His tags were used for Ni-column purification to remove unbound biotin and enhance purity. Cultures were grown in LB medium supplemented with 50 μg/ml Ampicillin and 50 μg/ml Streptomycin and when OD_{600} reached ~0.3, IPTG (0.4 mM) and biotin (5 μM) were added for protein induction and biotinylation, respectively.

Flow cytometric analysis

HEK293T cells were maintained as described previously (Chapter 2). Cells were passed into 12-well plates at 10,000 cells/cm^2. 16-24 h after seeding, cells were transiently transfected with 1 μg total DNA at a 1:2 ratio of DNA:jetPrime (Polyplus Transfection). Cells were transfected with 0.05 μg of target, 0.25 μg of ubiquibody or control, and balanced with empty pcDNA3 vector. Culture media was replaced 4-6 h post-transfection. Then, 24 h post-transfection, cells were harvested and resuspended in phosphate buffered saline (PBS) for analysis using a BD FACSCalibur or BD FACSAria Fusion. FlowJo Version 10 was used to analyze samples by geometric mean fluorescence determined from 10,000 events.

Yeast media and plates

SD-CAA minimal media was prepared containing 20g/L D-glucose, 6.7 g/L yeast nitrogen base, 5 g/L casamino acids, 5.4 g/L Na_2HPO_4, 8.6 g/L $NaH_2PO_4(H_2O)$ at

pH 6.0. SG-CAA minimal induction media was prepared containing 18 g/L galactose, 2 g/L D-glucose, 6.7 g/L yeast nitrogen base, 5 g/L casamino acids, 5.4 g/L Na_2HPO_4, 8.6 g/L $NaH_2PO_4(H_2O)$ at pH 6.0. YPD media was prepared containing 20g/L dextrose, 20g/L peptone, 10 g/L yeast extract. SD-CAA plates were prepared containing 20g/L D-glucose, 6.7 g/L yeast nitrogen base, 5 g/L casamino acids, 5.4 g/L Na_2HPO_4, 8.6 g/L $NaH_2PO_4(H_2O)$ and 182 g/L sorbitol. YPD plates were prepared containing 20g/L dextrose, 20g/L peptone, 10 g/L yeast extract, 16 g/L agar. Liquid medias were sterile filtered. For yeast plates, Na_2HPO_4, $NaH_2PO_4(H_2O)$, and agar were autoclaved in 0.9X final volume and once cooled to room temperature, the remaining components in 0.1X final volume were sterile filtered into the autoclaved mixture.

Yeast library screening with magnetic beads

Protocols detailed here were based on those detailed in Chen et al., 2013. A frozen aliquot containing 10X diversity (1X diversity = 2.5 x 10^8) of the G4 library (EBY100 cells containing the pCT-CON-FN3 library members) was thawed into 1L of SD-CAA minimal media. The culture was grown at 30°C at 250 rpm overnight. The next day, stocks of the library were prepared by storing 10X diversity in 15% glycerol in SD-CAA in -80 C. 30X diversity was then pelleted at 2500xg for 5 min and resuspended in 1L of SG-CAA to induce expression. The induced library was incubated at 20°C for 24 hours.

Initial library screening was performed with Dynabeads MyOne Streptavadin T1 to allow for the isolation of weak affinity FN3 binders through avidity interaction. Two batches of beads were prepared by incubating 33pmoles of biotinylated ERK2 (bERK2) with 10 μL of Dynabeads in 100 μL of PBSA (1X phosphate-buffered

saline, 0.1% bovine serum albumin) in 2 mL microcentrifuge tubes. Tubes were incubated at 4°C for at least an hour rotating. Immediately before adding cells, the beads were incubated on magnet for 5 minutes, supernatant removed, and washed once on magnet with 1 mL of PBSA to remove any free antigen.

The induced naïve G4 library was measured at OD_{600} to determine cell density. Assuming OD_{600nm}=1 is ~10^7 cells/mL, 15X library diversity was pelleted at 3000xg for 5 min and washed with 1 mL of PBSA, then split into two aliquots in 2 mL microcentrifuge tubes. Cells were pelleted at 12,000xg for 1 min and resuspended in 1 mL PBSA. A negative sort was performed by adding 10 μL of bare Dynabeads to each tube and incubating at 4°C for at least two hours rotating. Cell and bead mixtures were placed on the magnet and unbound cells were immediately transferred to the tubes with the washed, ERK2-coated Dynabeads. This positive sort was incubated at 4°C for at least two hours rotating. The negative sort beads were washed with 1 mL PBSA on magnet, supernatant removed, and remaining beads and cells resuspended in 5 mL of fresh SD-CAA. Similarly, the positive sort was washed with 1 mL PBSA on magnet, supernatant removed, and beads and cells resuspended in 5 mL of fresh SD-CAA. After both positive and negative sorts were complete, cultures were grown at 30°C at 250 rpm overnight to recover isolated yeast. Serial dilutions of negative and positive sort cultures were plated on SD-CAA plates.

The culture that resulted from the positive sort was pelleted and 4 mL of media removed. The pelleted cells and beads were resuspended in the remaining 1 mL of culture and transferred to a 2 mL microcentrifuge tube. The mixture was added to the

magnet to remove the beads. Recovered cells were transferred back to the original culture tube. Culture density was measured at OD_{600} and original library diversity calculated by back-calculating growth, assuming a doubling time of 4 h. 10X library diversity (after magnetic bead screening) was pelleted at high speed for 1 minute and resuspended in 5 mL of SG-CAA. Culture incubated at 20°C at 250 rpm for 24 hours for protein induction.

Intermediate sorting was performed by pelleting 20x diversity of the induced population, washing in 1mL PBSA, pelleting at high speed for 1 min, and resuspending the cells in 1 mL of PBSA. Two negative sorts were performed sequentially - the first with 10uL of bare beads at 4°C for at least two hours rotating. Unbound cells from this first negative sort were then transferred to a new 2 mL microcentrifuge tube and once again incubated with 10 μL of bare beads at 4°C for at least two hours rotating. The bead and cell mixture was added to the magnet and unbound cells added to two tubes with washed ERK2-coated Dynabeads (as prepared previously). This positive sort was incubated at 4°C for at least two hours rotating. After incubation, mixture was washed on magnet with 1 mL PBSA to remove any non-bound yeast. Beads and cells were then resuspended in 5 mL of fresh SD-CAA. Serial dilutions of both negative sorts (washed as described for initial sort) and the positive sort were plated on SD-CAA plates. As before, cultures were grown at 30°C at 250 rpm overnight.

FACS of yeast surface displayed G4 library

Protocols detailed here were based on those detailed in Chen et al., 2013. For sorting of fully displayed clones, 10X library diversity (post-magnetic bead screening)

was washed with 1 mL of PBSA and resuspended in 100 μL of PBSA. Cells were labeled with 1 μL rabbit anti-c-myc (Sigma) for 1 hour at room temperature. Cells were washed with 1 mL of PBSA, pelleted at 3000xg for 1 min, and resuspended in 100 μL PBSA. Cells were then labeled with 10 μL of mouse anti-HA-FITC (Abcam) and 1 μL goat anti-rabbit AF647 for 1 hour on ice, maintaining a 10X excess of antibody:FN3. Estimating 10^5 FN3/yeast cell [116], for 10^7 cells in 100 μL, 2.5 μL of a 1mg/mL antibody stock is required. Cells were washed 3x in 1 mL PBSA and resuspended in 500 μL for sorting. Single-labeled and unlabeled cells were prepared in parallel and used to set gates on BD FACSAria flow cytometer. Sorted cells were collected in 15 mL tubes into 2 mL of SD-CAA media and recovered in the same tube overnight at 30°C at 250 rpm.

For sorting of binders, 10X library diversity (post-sorting of fully displayed clones) was washed with 1 mL of PBSA and resuspended in 100 μL of PBSA. Primary labeling was achieved by incubating cells with 1μL rabbit anti-c-myc (Sigma) and indicated concentration of biotinylated ERK2 (Round 3- 500 nM, Round 4 – 500 nM, Round 5- 100 nM) for 1 hour at room temperature. Cells were washed with 1 mL of PBSA, pelleted at high speed for 1 min, and resuspended in 100 μL of PBSA. Secondary labeling was achieved by incubating cells with Streptavidin-AF488 and goat anti-rabbit AF647 for 1 hour on ice. Cells were washed 3x in 1 mL PBSA and resuspended in PBSA for sorting. Incubation volumes usually ranged from 100 – 500 μL and varied based on the antigen (ERK2) concentration required and were optimized to ensure 10X excess antigen:FN3 and minimize antigen required. Single-labeled and unlabeled cells were prepared in parallel and used to set gates on BD FACSAria flow cytometer. Sorted

cells were collected in 15 mL tubes into 2 mL of SD-CAA media and recovered in the same tube overnight at 30°C at 250 rpm.

Plasmid Isolation from Yeast

Protocols detailed here were based on those detailed in Chen et al., 2013. Following isolation of full-length clones by FACS, 10^8 cells were pelleted at 300xg and plasmids isolated using Zymoprep Yeast Plasmid Miniprep Kit II (Zymo Research) with a slightly modified protocol. The pellet was resuspended in 200 µL Solution 1. 3 µL of Zymolase was added to the resuspended cells and incubated at 37°C for 1 hour. 200 µL of Solution was added with gentle mixing. Then 400 µL of Solution 3 was added and gently mixed. Samples were centrifuged at 12,000xg for 8 minutes. Supernatant was removed and centrifuged again at 12,000xg for 5 minutes. The cleared supernatant was transferred to an Epoch miniprep column. Spin column was centrifuged at 12,000xg for 1 minute and the flow through discarded. Column was washed with 550 µL of PE buffer (Qiagen) and centrifuged at 12,000xg for 1 minute. Flow through was discarded and the column was centrifuged again to clear any remaining PE buffer. Plasmid DNA was eluted with 40 µL of ddH_2O.

Random mutagenesis of pCT-Con-FN3 library

Zymoprepped plasmid DNA was subjected to random mutagenesis using the Random Mutagenesis Kit (Agilent) and primers detailed in Chen et al., 2013. The entire FN3 gene was targeted for low levels of mutagenesis and the three loops – BC, DE, and FG – were targeted for high levels of mutagenesis as detailed in the Random Mutagenesis Kit protocol. Mutagenized samples were run on a 1.5% agarose gel with DNA standards to quantify amount of amplification, and by extension level of

mutagenesis. Desired bands were purified by Qiagen DNA Gel Extraction Buffers and then further amplified with primers detailed in Chen et al., 2013 using standard Vent Polymerase conditions. Amplified DNA samples were then purified with Qiagen DNA PCR Clean-Up Buffers.

Yeast library electroporation

Protocols detailed here were based on those detailed in Chen et al., 2013. pCT-Con-Gene vector was linearized by first digesting with NdeI overnight at 37°C, followed by digestion with PstI and BamHI. pCT-Con-Loop vector was linearized by first digesting with SmaI overnight at 37°C, followed by digestion with NcoI and NdeI. After fully digested, linearized plasmids were purified using Qiagen PCR Clean-Up Buffers.

Protocols detailed in Benatuil et al., 2010 were used as a reference for the following yeast electroporation steps. EBY100 yeast cells were grown overnight in YPD at 225 rpm and 30°C. The next day, the EBY100 were subculture into 100 mL YPD at 0.3 OD_{600}. The subculture was grown at 30°C and 225 rpm until reaching OD_{600} = 1.6, when yeast cells were pelleted by centrifugation at 3000xg for 3 minutes. Cell pellet was washed twice with 50 mL ice-cold water and once with 50 mL ice-cold electroporation buffer (1M Sorbitol/1 mM $CaCl_2$). The yeast were then resuspended in 20 mL 0.1M LiAc/10 mM DTT in a culture flask and incubated for 30 minutes at 30°C with shaking at 225rpm. Then, cells were collected by centrifugation and washed with 50 mL ice-cold electroporation buffer. Finally, the cells were resuspended in 100 to 200 µL electroporation buffer to obtain a final volume of 1 mL – which is enough for two

electroporations of 400 μL each. 400 μL of prepared cells were added to 4 μg of linearized pCT-Con-Gene vector and 12 μg of mutagenized, amplified FN3 gene fragments in less than 50 μL. The cell and DNA mixture was then transfered to a pre-chilled 2 mm electroporation cuvette and incubated for 5 minutes on ice. The remaining aliquot of electrocompetent cells were prepared with 4 μg of linearized pCT-Con-Loop vector and 12 μg of mutagenized, amplified FN3 loop fragments (BC, DE, and FG fragments combined) Cells were electroporated at 2,500 V and 25 μF. Electroporated cells were transfered into 8 mL of 1:1 1M Sorbitol: YPD and incubated at 30°C with shaking at 225 rpm for 1 hour. Cells were centrifuged and resuspended in 250 mL SD-CAA. 10-fold dilutions were plated onto SD-CAA plates and incubated at 30°C for two days to determine library size.

Acknowledgements

Thank you to Richard Roberts, Shohei Koide, Dane Wittrup, Amy Karlsson, and Andreas Pluckthun for kindly providing plasmids encoding genes used in this study. Thank you to Erin Stephens for design of ERK2-specific CHIP ubiquibodies and valuable conversations. Thank you to Natalia Lopez Barbosa for performing experiments related to the isolation and characterization of ERK2-specific FN3s.

This work was supported by the National Science Foundation Grant CBET-0449080 (to M.P.D.), the National Institutes of Health Grant Numbers CA132223 (to M.P.D.), the New York State Office of Science, Technology and Academic Research Distinguished Faculty Award (to M.P.D.), the Cornell Technology Acceleration and Maturation (CTAM) Fund, the National Science Foundation Graduate Research

Fellowship Program Grants DGE-1650441 (to M.B.L.), Cornell Presidential Life Science Fellowships (to M.B.L.), and a Cornell Fleming Graduate Scholarship (to M.B.L). FACS analysis was supported by the Cornell Biotechnology Resource Center.

CHAPTER 4 – EFFECTIVE DELIVERY OF UBIQUIBODIES

INTRODUCTION

A major barrier for the therapeutic development of ubiquibodies is efficient intracellular delivery. Unfortunately, intracellular delivery of protein-based therapeutics represents a significant obstacle as most globular protein drugs do not spontaneously cross plasma membranes due to their relatively large size and biochemical properties [50]. Unlike smaller PROTACs, which can be designed for cell-permeability [118], ubiquibodies are relatively bulky proteins that do not effectively penetrate the cell membrane. To remedy this issue, we pursued two distinct therapeutic delivery strategies - protein delivery via native lytic peptides and mRNA delivery via polypeptide nanoplexes.

To explore the option of protein-based delivery, we genetically fused the native IpaH9.8 cell-penetrating peptide (CPP), 2αH, to the N-terminus of our previously described GS2-IpaH9.8 ubiquibody. The resulting 2αH-GS2-IpaH9.8 ubiquibody elicited robust GFP degradation when expressed via plasmid transfection. As a purified protein, 2αH-GS2-IpaH9.8 demonstrated autonomous cellular penetration consistent with endosomal pathway uptake, but GFP degradation was not observed.

More successfully, we also investigated the use of mRNA-based delivery in collaboration with the Hammond Lab at MIT. We implemented a bioinspired mRNA delivery strategy whereby mRNA encoding GS2-IpaH9.8 with an additional 3′-terminal polyadenosine (poly A) tail was stoichiometrically complexed with poly A binding proteins (PABPs), which served to improve mRNA stability and also stimulate mRNA

translation in eukaryotic cells [94]. The resulting ribonucleoproteins (RNPs) were stabilized with cationic polypeptides to protect the mRNA from degradation, enable its uptake by cells, and facilitate its endosomal escape. Importantly, these co-assembled nanoplexes delivered GS2-IpaH9.8 mRNA in a manner that caused efficient GFP silencing after introduction to cultured mammalian cells stably expressing GFP and after administration to transgenic mice expressing GFP ubiquitously.

Results

Delivery of proteins via native lytic peptides

To investigate whether our IpaH9.8-based ubiquibodies could be engineered to autonomously translocate into mammalian cells, we first genetically fused the native cell-penetrating peptide 2αH from IpaH9.8 [90] to the N-terminus of GS2-IpaH9.8. (**Fig. 4.1a**). The resulting ubiquibody 2αH-GS2-IpaH9.8 retained the domain architecture of the parental IpaH9.8 protein with the exception of the replacement of the native LRR domain with GS2, our GFP-specific FN3. 2αH-GS2-IpaH9.8 showed similarly robust EGFP degradation activity in HEK293T cells when expressed from our traditional mammalian expression plasmid (**Fig. 4.1b**).

To evaluate autonomous cellular uptake, proteins were expressed and purified from *E. coli*, followed by conjugation to the fluorophore AlexaFluor-488 to facilitate quantification and visualization (**Fig. 4.2a**). 2αH-GS2-IpaH9.8 did show autonomous cellular entry into HEK293T cells as observed by flow cytometry and immunofluorescence (**Fig. 4.2b-c**). No cellular uptake was observed for any proteins when incubated at 4°C, corroborating previous reports that CPP-mediated uptake relies on endocytic processes [90]. It remains unclear, however, why GS2 and IpaH9.8dLRR,

which lack the 2αH peptide, would elicit cellular uptake.

Although fluorescently labeled GS2-IpaH9.8 and 2αH-GS2-IpaH9.8 proteins showed robust entry into HEK293T cells, when unlabeled ubiquibodies were incubated with Hela EGFP cells, no EGFP degradation was observed (**Fig. 4.3**). Many factors could be contributing to this lack of observable activity, including entrapment in intracellular compartments, rapid degradation in lysosome, formation of inactive dimers, insufficient intracellular concentration, and insufficient incubation time.

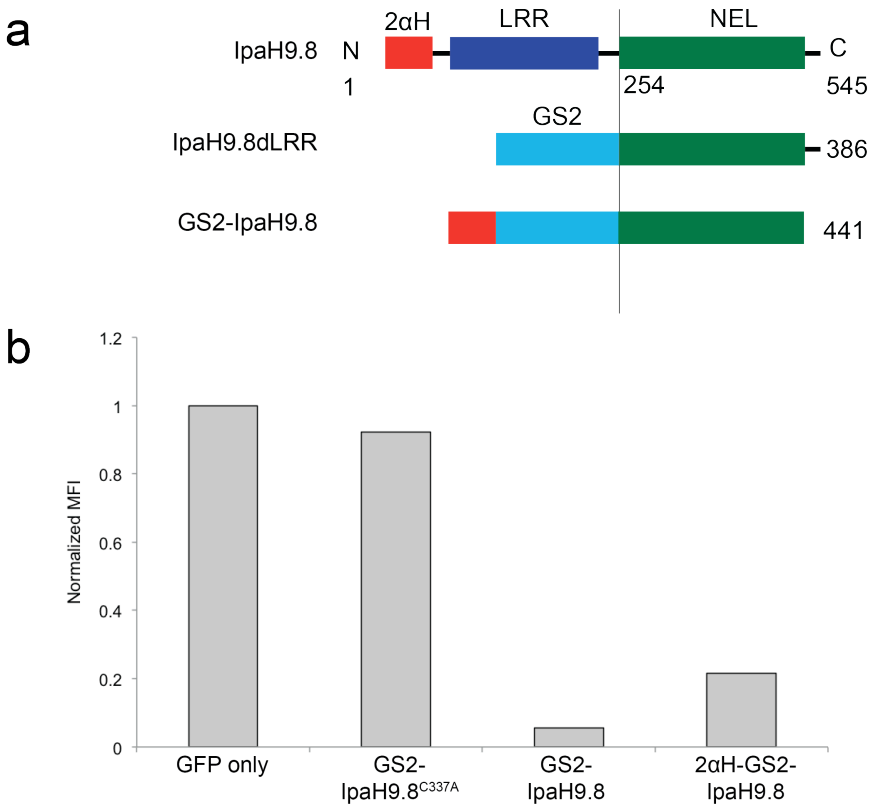


Figure 4.1 2αH-GS2-IpaH9.8dLRR still retains degradation activity.
(a) Linear representation of IpaH9.8, GS2-IpaH9.8, and 2αH-GS2-IpaH9.8. Numbers refer to amino acid positions from N terminus (N) to C terminus (C). The proteins are aligned vertically with the NEL domain of IpaH9.8. (b) Flow cytometric analysis of EGFP fluorescence activity in HEK293T cells transfected with plasmid pcDNA3-EGFP alone or co-transfected with pcDNA3-EGFP and a plasmid encoding GS2-IpaH9.8, GS2-IpaH9.8^{C337A} or 2αH-GS2-IpaH9 as indicated. Data are geometric MFI normalized to geometric MFI measured for HEK293T cells expressing ERK2-EGFP alone.

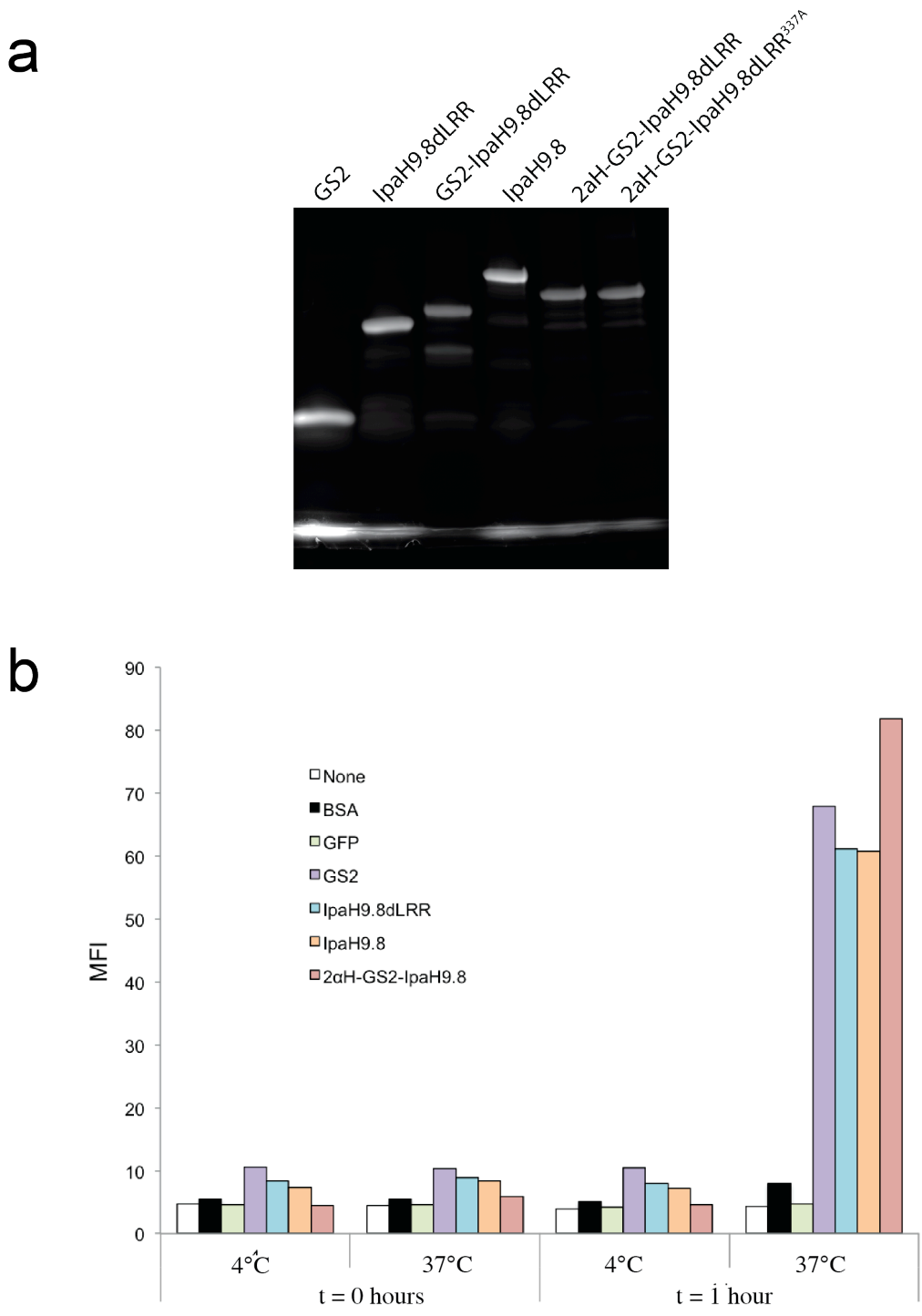


Figure 4.2 Mammalian cell uptake of 2αH-GS2-IpaH9.8.
(a) Fluorescently imaged SDS-PAGE gel of 5µg of AF-488 labeled proteins as indicated. (b) Flow cytometric analysis of HEK293T cells incubated with 400 nM of AF-488 labeled proteins for the indicated time and temperature. Data are represented as the geometric MFI.

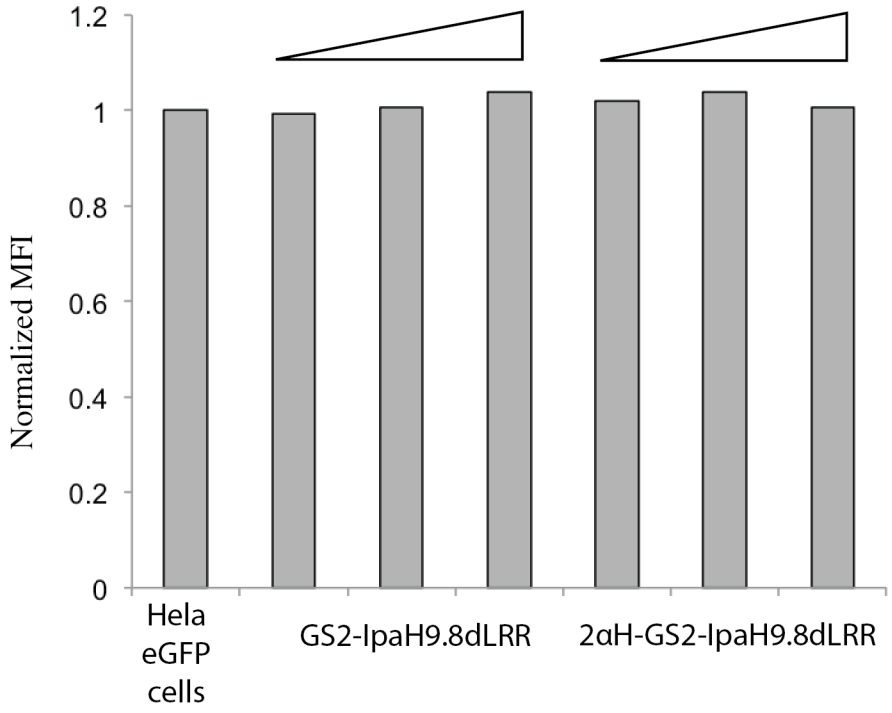


Figure 4.3 Degradation activity of 2αH-GS2-IpaH9.8dLRR in Hela EGFP cells.
Flow cytometric analysis of Hela EGFP cells incubated with 400 nM, 800 nM, or 1600 nM, as indicated by the triangles, of purified GS2-IpaH9.8 or 2αH-GS2-IpaH9.8 for 24 hours. Data are geometric MFI normalized to geometric MFI measured for HEK293T cells expressing EGFP alone

Delivery of mRNA encoding GS2-IpaH9.8 enables proteome editing in mice

From a therapeutic standpoint, one of the biggest challenges facing protein-based technologies such as ubiquibodies is intracellular delivery [50]. The Hammond group previously showed that co-assembled nanoplexes comprised of synthetic mRNA containing a poly A tail, PABPs, and biocompatible cationic polypeptides (**Fig. 4.4**) resulted in greatly enhanced mRNA expression *in vitro* and in mice [94]. Here, we hypothesized that delivery of GS2-IpaH9.8 mRNA/PABP nanoplexes to mammalian cells would result in significantly greater ubiquibody expression relative to mRNA

transfection alone by the same polyamine in HEK293T cells, thereby leading to potent protein degradation. To test this hypothesis, we first evaluated GS2-IpaH9.8 mRNA/PABP nanoplex delivery *in vitro* by quantifying the degradation of d2EGFP, a destabilized GFP variant that was expressed as a stable transgene in HEK293T cells. As expected, only when the cationic nanoplexes contained the active, target-specific GS2-IpaH9.8 mRNA and PABP was robust d2EGFP degradation achieved (**Fig. 4.5**). All other controls including catalytically inactive GS2-IpaH9.8^{C337A} mRNA/PABP nanoplexes, non-specific AS15-IpaH9.8 nanoplexes, and naked GS2-IpaH9.8 mRNA that was delivered without PABPs showed little to no silencing activity (**Fig. 4.5**). At 24 hours post-treatment, HEK293Td2EGFP cells receiving GS2-IpaH9.8 mRNA/PABP nanoplexes exhibited an 85% decrease in fluorescence activity, which was directly comparable to the knockdown activity achieved following DNA transfection seen above.

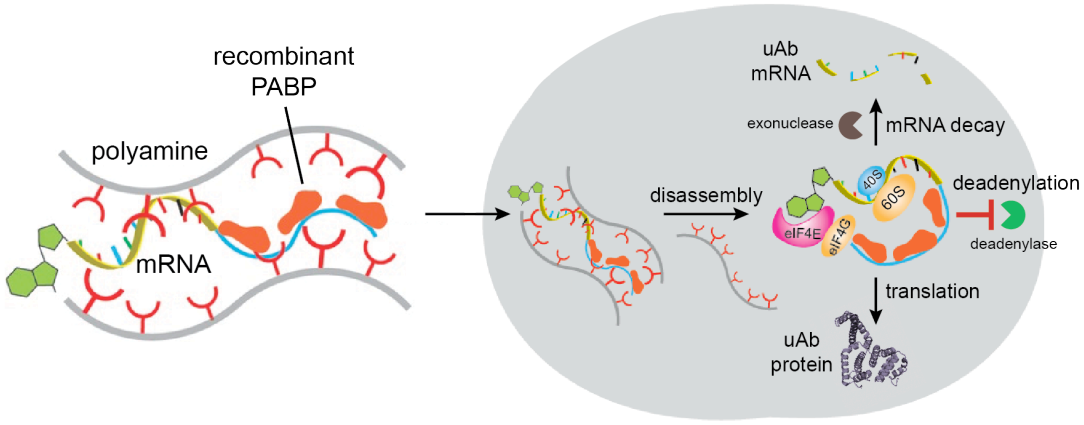


Figure 4.4 Schematic of polyamine -mediated stoichiometric assembly of ribonucleoproteins for enhanced mRNA delivery.
Following internalization in cells (grey circle), nanoplex disassembly results in the release of mRNA/PABP that is either degraded or translated to produce ubiquibody proteins.

Encouraged by these results, we next evaluated ubiquibody nanoplex-mediated delivery and silencing activity *in vivo*. Transgenic UBI-GFP/BL6 mice, which constitutively express EGFP in all tissues [119], were given subcutaneous injections of GS2-IpaH9.8 mRNA/PABP nanoplexes in ears. Note that although this mouse strain ubiquitously expresses EGFP, fluorescence is absorbed and undetectable in areas that are covered by hairs. Fluorescent imaging at 24 hours post-injection revealed that EGFP fluorescence in the left ears, which received GS2-IpaH9.8 mRNA/PABP nanoplex injections, was robustly ablated with a 70% decrease in ear fluorescence (**Fig. 4.6a and b**). In stark contrast, fluorescence in the right ears, which received either catalytically inactive GS2-IpaH9.8^{C337A} or non-specific AS15-IpaH9.8 nanoplex injections, was unaffected (**Fig. 4.5a and b**). Importantly, these results set the stage for therapeutic delivery of ubiquibodies as a viable strategy to post-translationally silence aberrantly expressed proteins in cancer and other

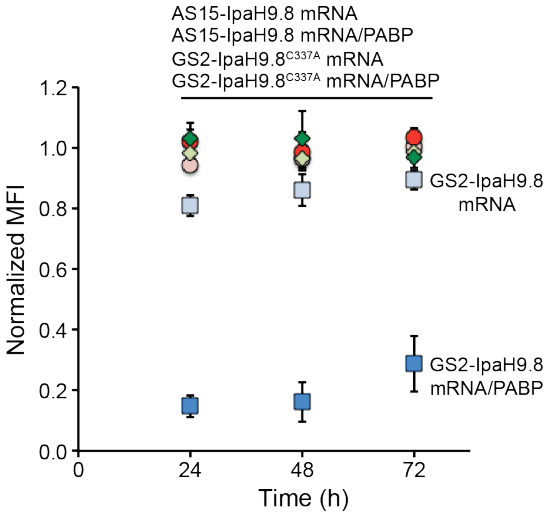


Figure 4.5 *In vitro* nanoplex delivery of ubiquibody mRNA.
Flow cytometric quantification of EGFP fluorescence activity in HEK293Td2EGFP cells incubated with: mRNA encoding GS2-IpaH9.8, GS2-IpaH9.8^{C337A}, or AS15-IpaH9.8; or with nanoplexes comprised of the same mRNAs formulated with PABP and TEP (N4) polyamine (mRNA:PABP weight ratio = 1:5). Measurements were taken at 24, 48, and 72 h post delivery. Data are expressed as the mean ± S.D. of biological triplicates (three separately treated wells). p values were determined by paired sample t-test.

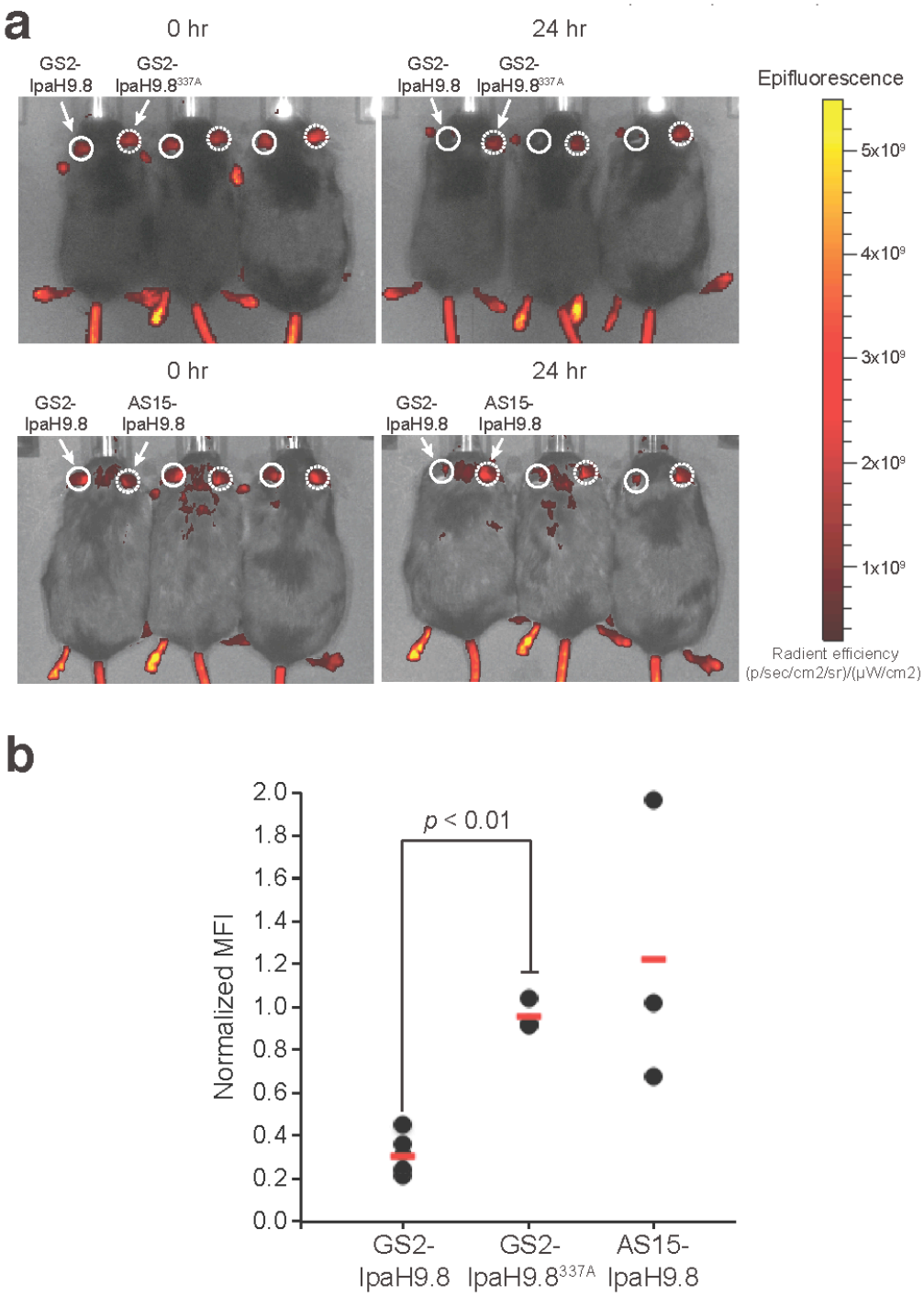


Figure 4.6 Proteome editing in mice via nanoplex delivery of ubiquibody mRNA.
(a) Epifluorescence imaging of UBC-GFP mice at 0 h (top) and 24 h (bottom) after ear injection of nanoplexes containing mRNA encoding GS2-IpaH9.8 (solid white circle), GS2-IpaH9.8^{C337A} (dashed white circle, top), or AS15-IpaH9.8 (dashed white circle, bottom). Numbers on the heat bar represent radiant efficiency (p/sec/cm^2/sr)/(μW/cm^2) (b) Quantification of GFP fluorescence in the ears of Ubi-GFP mice in (c). Data are reported as the mean radiant efficiency for each individual ear region (black circle) and the mean radiant efficiency (red bar) of each sample group (n = 6 for GS2-IpaH9.8, n = 3 for GS2-IpaH9.8^{C337A}, and n = 3 for AS15-IpaH9.8). p values were determined by paired sample t-test.

Discussion

A major advantage of ubiquibodies is the ease with which they can be rapidly adapted to hit a variety of intracellular targets due to their recombinant, modular design, which capitalizes on a large, preexisting repertoire of synthetic binding proteins as well as systematic, genome-wide efforts to generate and validate protein binders *de novo* against the human proteome [26]. Because obtaining antibody mimetics that bind with high specificity and affinity to a target should be easier than obtaining small molecules with the same properties, making custom-designed PROTACs is likely to be a much more challenging task [50]. Nonetheless, PROTACs holds great promise as a therapeutic approach because it is based on small molecules that have strong odds of getting into cells. Indeed, impressive preclinical *in vitro* and *in vivo* data are propelling the development of clinically viable PROTACs as evidenced by the founding of Arvinas in 2013 and C4 Therapeutics in 2016. It should be pointed out, however, that traditional medicinal chemistry approaches will be needed to improve the oral bioavailability, pharmacokinetics, and absorption, distribution, metabolism, excretion and toxicity (ADMET) properties of PROTACs [35, 36]. Compared to PROTACs, intracellular delivery of ubiquibody-based therapeutics is a much bigger hurdle as most globular protein drugs do not spontaneously cross plasma membranes due to their relatively large size and biochemical properties [50].

Cell penetrating peptides have emerged as a powerful tool for the delivery of proteins, nucleotides, and small molecules [120]. Furthermore, the identification of cell penetrating peptide sequences within bacterial effectors with known E3 ubiqutin ligase activity [88, 90, 121] made this an attractive strategy to pursue for the therapeutic

delivery of ubiquibodies. GS2-IpaH9.8 ubiquibody modified with the native cell-penetrating peptide 2αH from the N-terminus of the parental IpaH9.8 protein did show autonomous cellular uptake as analyzed by flow cytometry, but no degradation activity was observed. The specificity of uptake is yet to be confirmed, however, as other IpaH9.8-based proteins lacking the 2αH peptide were also able to enter mammalian cells. Interestingly, when 2αH-GS2-IpaH9.8 was expressed via plasmid transfection, robust GFP degradation was observed. This result adds to the growing evidence of the flexibility of our ubiquibody design - not only can the binding and catalytic domains be easy interchanged, but the design is also amenable to protein fusion at its N-terminus. While this work failed to achieve delivery of active ubiquibodies with a cell-penetrating peptide, it is likely a result of insufficient experimental conditions and not an inherent limitation of the ubiquibody design.

Another possible solution that we investigated here is the use of mRNA as a source of therapeutic gene product *in vivo.* In recent years, impediments to the use of mRNA, including its instability and immunogenicity, have been largely overcome through structural modifications, while issues related to delivery and protein expression profiles have been addressed through advances in nanotechnology and material science [122]. Here, we take advantage of our unique approach to create a first-in-kind therapeutic ubiquibody delivery strategy; this method involved a recently reported strategy of electrostatics to stabilize pre-formed protein-RNA complexes for delivery [94]. In terms of longevity of knockdown, this earlier study demonstrated that tail vein injection of nanoplexes formulated with luciferase mRNA resulted in the highest luciferase expression levels at 6 h post-injection, which remained statistically above

control at 24 h post-injection but significantly decayed to background after 48 h. While this previous work did not involve ubiquibodies directly, it does provide a sense for the longevity that might be possible using this delivery method. Here, synthetic mRNA encoding the GFP-directed GS2-IpaH9.8 chimera was co-assembled with PABPs and the assembled ribonucleoproteins were packaged into nanosized complexes using structurally defined polypeptides bearing cationic aminated side groups. The resulting nanoplexes achieved highly efficient silencing of GFP *in vitro* and *in vivo*, thereby demonstrating a new proteome editing paradigm and opening the door to clinical translation of ubiquibody -based therapeutics.

Materials and Methods

Cloning of 2αH constructs

All plasmids used in this study are provided in **Appendix C**. To construct pcDNA3-2αH-GS2-IpaH9.8, a synthetic double-stranded gblock (IDT) encoding HindIII-Kozak- 2αH –BamHI-GS2-NheI was ligated into pcDNA3-GS2-IpaH9.8 digested with HindII and NheI, to produce the final plasmid pcDNA3-2αH-GS2-IpaH9.8. To construct pET28a-2αH-GS2-IpaH9.8 and pET28a-2αH-GS2-IpaH9.8^{C337A}, the synthetic double-stranded gblock (IDT) encoding HindIII-NcoI/Kozak-2αH –BamHI-GS2-NheI was digested with NcoI and NheI and ligated into pET28a resulting in the intermediate plasmid pET28a-2αH-GS2. IpaH9.8dLRR and IpaH9.8dlRR^{C337A} were amplified with primers encoding 5' NheI and 3' Flag tag, His tag, and NotI sequences, followed by ligation into pET28a-2αH-GS2, resulting in the final plasmids pET28a-2αH-GS2-IpaH9.8 and pET28a-2αH-GS2-IpaH9.8^{C337A}. All plasmids were verified by DNA sequencing at the Cornell Biotechnology Resource Center (BRC).

Purification of proteins

GS2, IpaH9.8dLRR, IpaH9.8, GS2-IpaH9.8, and EGFP were purified as previously described (Chapter 2). 2αH-GS2-IpaH9.8 and 2αH-GS2-IpaH9.8^{C337A} purified proteins were obtained by growing *E. coli* BL21(DE3) cells containing a pET28a-based plasmid encoding the desired protein in 200 mL of Luria-Bertani (LB) medium at 37°C. Expression was induced with 0.1 mM IPTG when the culture density (Abs_{600}) reached 0.6-0.8 and growth continued for 6 h at 30°C. Cultures were harvested by centrifugation at 4,000xg for 30 min at 4°C. Cell pellets were stored at -20°C overnight. Thawed pellets were resuspended in 10 mL equilibration buffer (25 mM Tris-HCl, pH 7.4, 500 mM NaCl and 20 mM imidazole) and lysed with a high-pressure homogenizer (Avestin Emulsi-Flex C5). The insoluble fraction was cleared by centrifugation at 12,000xg for 30 min at 4°C. His-tagged protein was purified by gravity flow using 500 µL HisPur Ni-NTA resin (ThermoFisher). The soluble fraction was passed through the resin, after which the resin was washed with 3 mL of wash buffer (25 mM Tris-HCl, pH 7.4, 500 mM NaCl and 50 mM imidazole). Protein was eluted with 1.5 mL elution buffer (25 mM Tris-HCl, pH 7.4, 500 mM NaCl and 250 mM imidazole). Purified fractions were desalted and concentrated (Pierce PES Protein Concentrators). GS2, IpaH9.8dLRR, IpaH9.8, GS2-IpaH9.8, and EGFP were purified as previously described (Chapter 2).

Fluorophore conjugations

Purified proteins were fluorescently labeled with Alexa Fluor 488 NHS Ester (ThermoFisher). Briefly, ~1 µmol of protein was incubated with 7.5 µmol of Alexa Fluor-488 NHS Ester in PBS overnight at room temperature. Unreacted fluorophore was

removed using 10 MWCO desalting columns (Pierce). Degree of labeling was confirmed through quantification of A_{280} and A_{495} signals using ext. coefficient$_{495}$ = 71,000 mol/L cm^{-1}. Typical degrees of labeling were between 0.5-1.

Flow cytometric analysis

HEK293T cells were maintained as described previously (Chapter 2). For plasmid-based degradation experiments, cells were passed into 12-well plates at 10,000 cells/cm^2. 16-24 h after seeding, cells were transiently transfected with 1 μg total DNA at a 1:2 ratio of DNA:jetPrime (Polyplus Transfection). Cells were transfected with 0.05 μg of target, 0.25 μg of ubiquibody or control, and balanced with empty pcDNA3 vector. Culture media was replaced 4-6 h post-transfection. Then, 24 h post-transfection, cells were harvested and resuspended in PBS for analysis using a BD FACSCalibur. FlowJo Version 10 was used to analyze samples by geometric mean fluorescence determined from 10,000 events.

For cell uptake experiments, cells were passed into 12-well plates at 10,000 cells/cm^2 16-24 h after seeding, media was removed and 1 mL of fresh media containing the desired concentration of protein was added to the cells. After the desired incubation time, cells were harvested and washed with PBS. Immediately before analyzing on BD FACSCalibur, samples were diluted 1:1 in Trypan Blue to quench any proteins still externally associated with the cells.

Synthesis and characterization of cationic polypeptides

N4 (TEP) polyamines were synthesized as we described recently [94] according to a modified procedure of Uchida and coworkers [123]. Briefly, to a chilled solution of poly(β-benzyl-L-aspartate) in N-methyl-2-pyrrolidone (NMP) (Sigma) (2 mL) was

added dropwise with stirring 50 equivalents of tetraethylenepentamine (Sigma) diluted two-fold with NMP. After stirring for 2 h at 0°C, the pH was adjusted to 1 with dropwise addition while stirring of cold 6 N HCl. The resulting solution was dialyzed from a regenerated cellulose membrane bag (Spectrum Laboratories, 1 kDa MWCO) against 0.01 N HCl followed by distilled water, frozen, and lyophilized to give a white powder. Polyamines used in this study were characterized by ^{1}H NMR spectra in deuterium oxide (Cambridge Isotope Laboratories) using a Bruker Avance 400 MHz NMR spectrometer at 25°C: ^{1}H NMR (400 MHz, D2O) δ 4.72 (s, 1H), 3.64 – 3.39 (m, 9H), 3.37 – 3.05 (m, 5H), 3.00 – 2.62 (m, 4H).

Preparation of mRNA by in vitro transcription

cDNA encoding ubiquibodies was cloned into pGEM4Z/GFP/A64 by replacing the GFP fragment with XbaI and NotI sites. Additionally, the human α-globin 3' UTR sequence was placed between the cDNA and the poly A tail using NotI and EcoRI to improve mRNA translation. Linearization with SpeI, followed by *in vitro* transcription (IVT) with HiScribe™ T7 High Yield RNA Synthesis Kit (NEB), yielded a transcript containing 64 nucleotides of vector-derived sequence, the coding sequence, α-globin 3' UTR, and 64 A residues. In a typical 20 µl reaction, the following nucleotides were prepared: ATP (10 mM), pseudo-UTP (10 mM), methyl-CTP (10 mM), GTP (2 mM), anti-reverse Cap analog (8 mM, NEB). RNA was purified by RNeasy purification kit (Qiagen, Hilden, Germany). RNA quality was confirmed by running a 1% agarose gel. Concentration was determined by Abs_{260}.

Nanoplex transfection.

Polyamines were dissolved in 10 mM HEPES buffer (pH 7.4). For each well of a

96-well plate, 200 ng mRNA diluted in 5 µl OptiMEM (Thermo Fisher) was mixed with 5 µl OptiMEM containing PABP (mRNA:PABP weight ratio = 1:5) at room temperature for 10 min. Afterwards, 5 µl OptiMEM containing polyamines was added and incubated at room temperature for 15 min prior to transfection into HEK293T stably expressing d2EGFP. Polyamines were adjusted to achieve 50 to 1 (N/P) ratio for transfection. EGFP expression was measured by BD FACSCelesta at different time points after transfection.

Animal experiments

Mouse care and experimental procedures were performed under pathogen-free conditions in accordance with established institutional guidelines and approved protocols from the MIT Division of Comparative Medicine. C57BL/6-Tg(UBC-GFP)30Scha/J mice were purchased from Jackson Laboratory. 8-10 week-old mice were injected subcutaneously in ears with 5 µg mRNA and 25 µg PABP packaged with N4 (TEP) polyamines in a volume of 25 µl OptiMEM under anesthesia. Fluorescent imaging was performed with a CCD camera mounted in a light-tight specimen box (Xenogen). The exposure time was 1 s. Imaging and quantification of signals were controlled by Living Image acquisition and analysis software (Xenogen).

Acknowledgements

Thank you to Jiahe Li for his help designing, executing, and analyzing the mRNA nanoplex delivery experiments. Thank you to Alyse Portnoff and Erin Stephens for their insightful conversations. This work was supported by the National Science Foundation Grant CBET-0449080 (to M.P.D.), the National Institutes of Health Grant Numbers CA132223 (to M.P.D.), CA170820 and GM083898 (to R.W.R.), CA194864

and CA212608 (to S.K.) the New York State Office of Science, Technology and Academic Research Distinguished Faculty Award (to M.P.D.), the Cornell Technology Acceleration and Maturation (CTAM) Fund, the National Science Foundation Graduate Research Fellowship Program Grants DGE-1650441 (to M.B.L.), Cornell Presidential Life Science Fellowships (to M.B.L.) a Cornell Fleming Graduate Scholarship (to M.B.L.), the Koch Institute Support Grant P30-CA14051 from the National Cancer Institute and the Department of Defense Ovarian Cancer Research Program Teal Innovator Award (to P.T.H.). Confocal microscopy was supported by NYSTEM Grant CO29155 and NIH Grant S10OD018516 and FACS analysis was supported by the Cornell Biotechnology Resource Center and the David H. Koch Institute Flow Cytometry Core at MIT.

CHAPTER 5 - APPROACHES TOWARDS SPATIOTEMPORAL CONTROL OF UBIQUIBODY ACTIVITY

Introduction

While intracellular delivery poses a major challenge to the therapeutic application of ubiquibodies – for which we demonstrated successful strategies in Chapter 4 – it is also desirable to move beyond the use of perpetually active ubiquibodies towards a ubiquibodies endowed with spatial and temporal control. Recently, several studies have reported the use of genetically encoded photoreceptors to create engineered protein chimeras that are normally inactive, but that can be specifically "turned-on" by an external stimulus such as blue-light [124].

One light-gated strategy that has been used is the chimeric fusion of LOV domains. LOV domains are naturally found in plants, regulating a variety of photosynthetic processes [125]. LOV domains are members of the Per-Arnt-Sim (PAS) family, a class of signaling proteins, and as such adopt the canonical PAS domain fold [126]. They non-covalently bind a flavin nucleotide cofactor (FMN or FAD), but upon photoactivation, a covalent thioether bond is formed between the cofactor and a highly conserved cysteine residue on the LOV domain. This light absorption induces changes in and through the β -sheet of the LOV domain to the outer surface, where downstream effects depend on the specific LOV domain being used [126]. In a well-studied example, LOV2 from *Avena sativa* phototropin I (AsLOV2), a C-terminal amphipathic helix (Jα) undergoes significant light-induced conformational changes, while the LOV core experiences only small perturbations [127]. In the dark-state, the Jα helix interacts with

the β-sheet of the LOV core, but upon illumination, Jα undocks and unfolds, propagating signal to an effector domain [128]. The N-terminal A'α helix also exhibits light-induced unfolding, and triggers Jα undocking [129].

Several reports have shown that LOV domains can successfully be used as input domains to create engineered photoreceptors with novel light-gated biological functions. Strickland et al. developed LOV-TAP, by fusing AsLOV2 N-terminally to TrpR, a bacterial transcription factor, such that they shared a continuous helix. *In vitro* experiments initially showed a 5-fold change in activity between the light and dark states, but further site-directed mutagenesis improved this activity to a 64-fold change. Similarly, Wu et al. fused AsLOV2 N-terminally to the GTPase RacI, resulting in a 10-fold change in activity between the light and dark states. Moglich et al. successfully utilized a LOV domain from *Bacillus subtilis* YtvA to reprogram the histidine kinase FixL from an oxygen- to light-responsive protein, ultimately achieving 1000-fold higher activity in the light state. For each of the cases mentioned, library sizes were on the order of tens, and variants were screened manually.

AsLOV2 has also been used to develop methods for light-gated protein degradation. Bonger et al. created a blue-light inducible degradation (B-LID) domain, by fusing a small four amino acid peptide degron to the C-terminus of the Jα helix of AsLOV2, which becomes exposed following blue-light induction. When fused to various fluorescent proteins, B-LID facilitated light-gated proteasome-mediated degradation in zebra fish [130]. Renicke et al. utilized an analogous strategy, demonstrating light-gated proteasome-mediated degradation in *S. cerevisiae* [131].

Several efforts have been described which employ genetically encoded dimerizers based on plant photoreceptors to engineer light-dependent protein-protein interactions [132-135]. A notable example from Kennedy et al., utilized light-inducible protein –interaction modules based on *Arabidopsis thaliana* cryptochrome 2 and CIB1. Importantly, this system requires no exogenous ligand and dimerizes upon blue-light exposure with subsecond time resolution and subcellular spatial resolution. Furthermore, recent work from the same group has optimized the CRY2-CIB1 system, producing smaller modules with reduced dark interaction [136]. This system has been used to optically regulate transcription [137], recombinase activity [136], signalling [138], and other cellular functions.

Here, we present designs and preliminary characterization of proposed light-induced ubiquibodies. A panel of chimeric fusions of AsLOV2 in between the GS2 binding domain and IpaH9.8 catalytic domain of GS2-IpaH9.8 showed successful expression in mammalian cells. Further investigation of one fusion, GS2-AsLOV2-IpaH9.8-R1, revealed that IpaH9.8dLRR retained its ubiquitination activity and AsLOV2 exhibited characteristic spectral properties. Split ubiquibody designs were also evaluated using two different sets of fusion partners – light-gated CIB1/CRY2 and chemically induced FKBP/FRB. A variety of chimeric fusion combinations were explored with highly variable expression in mammalian cells. In particular, IpaH9.8dLRR fusions showed very poor expression, which we hypothesized was due to auto-ubiquitination and subsequent proteasomal degradation. To investigate if reduced auto-ubiquitination ubiquibodies could be engineered , lysine mutants of GS2-IpaH9.8 were generated and characterized for expression and degradation activity in mammalian

cells. All the individual lysine mutants showed robust degradation activity, but no significant improvement was observed for expression levels.

Together, these results indicate that such chimeric fusion of light-gated or chemically inducible modules can be executed without disrupting the individual functions of the ubiquibody binding and catalytic domains. However, continued research needs to be performed to more thoroughly characterize the interactions of these domains on the overall ubiquibody function to fully realize light-gated targeted protein degradation.

Results

Engineering light-gated chimeric ubiquibodies

Based on the success of AsLOV2 protein fusions in the literature, we hypothesized that a fusion between AsLOV2 and the C-terminal catalytic domain of IpaH9.8 could result in light-gated ubiquibody **(Fig. 5.1**). We designed and created a "register" of ubiquibody chimeras in which the AsLOV2 domain is sandwiched between the DBP and the IpaH9.8 catalytic domain. The C-terminal fusion between AsLOV2 and IpaH9.8 is varied by a single amino acid to attempt to effectively ratchet the proteins around each other, sampling all steric space (**Fig. 5.2a**). When expressed in HEK293T mammalian cells, all of the variants with the exception of GS2-AsLOV2-IpaH9.8-R3 showed observable expression (**Fig. 5.2b**). Register variants R1, R2, R4, and R5 showed the strongest signal, but also resulted in very strong bands below the anticipated protein molecular weight, indicative of a degradation product. Comparatively, R9 and R10 had much lower expression levels, but also had a significantly lower proportion of the degradation product.

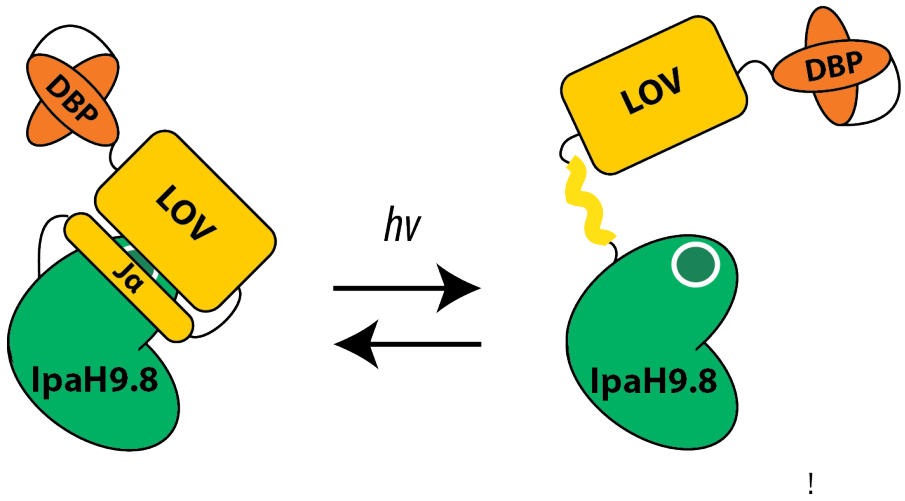


Figure 5.1 Schematic of light-gated AsLOV2 ubiquibody chimeras.
AsLOV2 fused C-terminal to DBP and N-terminal to IpaH9.8 catalytic domain can be stimulated by blue-light to reveal an active ubiquibody.

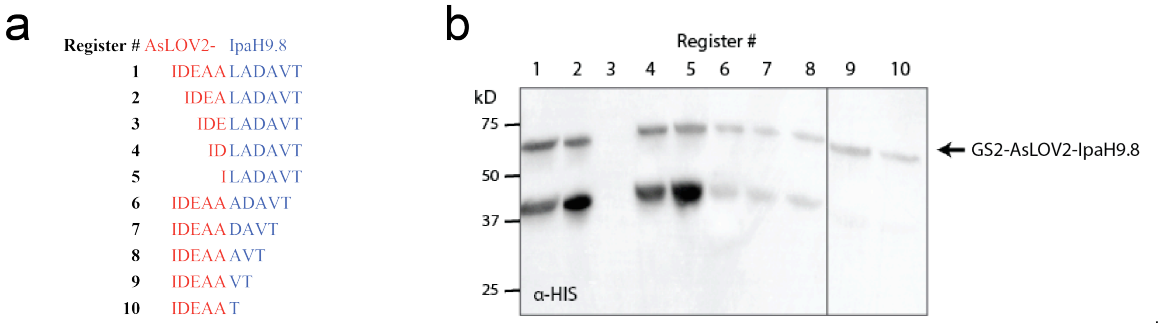


Figure 5.2 Mammalian expression of AsLOV2-ubiquibody chimeras.
(a) Detailed alignment of amino acid sequences of C-terminal AsLOV2 and N-terminal IpaH9.8 catalytic domain for the register of chimeric fusions. (b) Western blot analysis of HEK293T cell lysates transfected with indicated AsLOV2 chimeras expressed from pTriEx-3 vector. Blots were probed with antibodies specific for 6x-His (that detected tag on each ubiquibody). An equivalent amount of total protein was loaded in each lane. Molecular weight (MW) markers are indicated on left. Bands representing full-length protein are indicated with arrow.

AsLOV2 can be mutated to lock its conformation in the "open" or "closed" states [139]. To evaluate these light-gated ubiquibody designs without the added dynamic layer of light activation, we cloned "open" mutants of variants R1, R3, and R7 and evaluated their ubiquitination activity in *E. coli* expressing a reconstituted ubiquitination pathway (**Fig. 5.3**). While all three variants expressed in this system, but presented strong degradation bands. The high-molecular weight smearing detected for all three variants in the α-HIS blot is indicative of auto-ubiquitination. Similar laddering was observed in the α-ubiquitin blot, which could be indicative of ubiquibody auto-ubiquitination, but the low molecular weights also point to the generation of free poly-ubiquitin chains [140]. These results confirm that all three "open" variants contained active IpaH9.8 catalytic domains with R1 exhibiting the strongest degree of auto-ubiquitination, while R7 showed the strongest production of free poly-ubiquitin chains.

Next, the variant R1 was purified and analyzed by spectroscopy to evaluate AsLOV2. R1 displayed two strong fluorescence peaks at ~500 and ~525 nm, which were identical to the characteristic fluorescence profile observed from AsLOV2 alone [141]. No absorbance was observed from the buffer control or the GS2-IpaH9.8, confirming that the absorbance profile was specific to the AsLOV2 domain in the R1 protein.

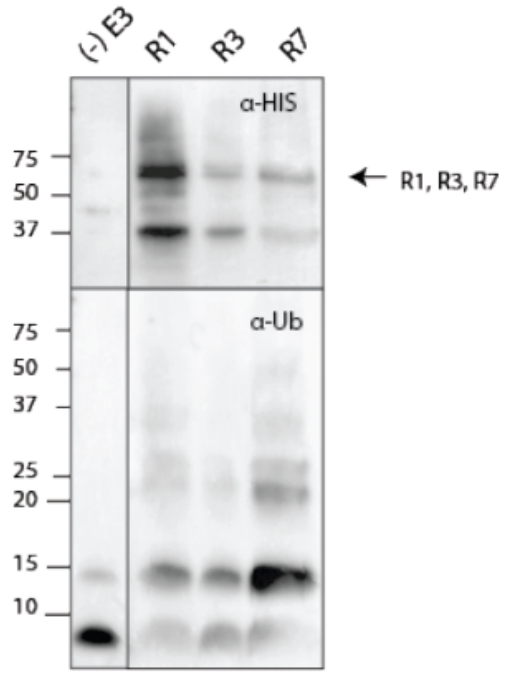


Figure 5.3 *E. coli* ubiquitination assay of light-gated AsLOV2 ubiquibody chimeras.
Western blot analysis of Rosetta(DE3) *E. coli* cell lysates co-transfected with AsLOV2 chimeras expressed from pTriEx-3 vector, .pCDF Duet UbcH5α/Ube1, and pACYC Ubiquitin Blots were probed with antibodies specific for 6x-His (that detected tag on each ubiquibody). and poly-ubiquitin An equivalent amount of total protein was loaded in each lane. Molecular weight (MW) markers are indicated on left. Bands representing full-length ubiquibody protein are indicated with arrow.

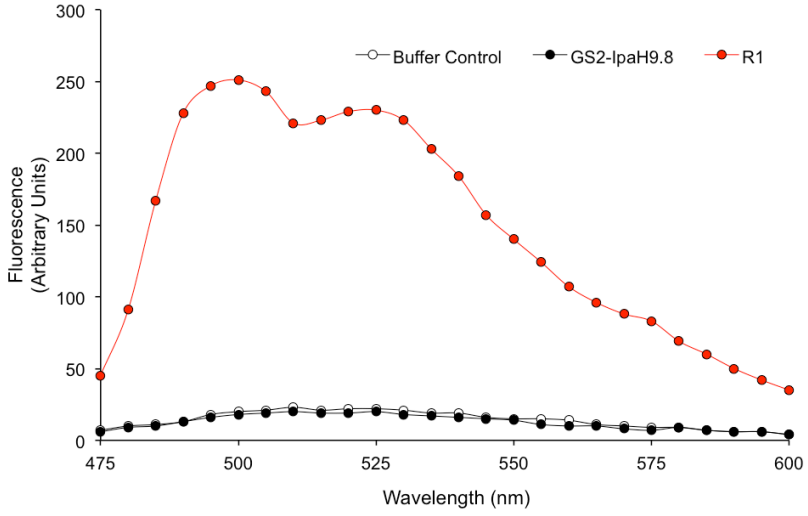


Figure 5.4 Fluorescence spectra of light-gated AsLOV2 ubiquibody fusion.
Absorption spectra of purified GS2-AsLOV2-IpaH9.8-R1, purified GS2-IpaH9.8, or buffer only. Samples were excited at 450 nm and emission fluorescence measured at the indicated wavelengths.

Designs for inducible ubiquibodies via protein interacting pairs

Another approach is to employ a split protein system in which the binding domain and catalytic domain are expressed separately and ubiquibody function is only achieved when the two halves are reassembled following an exogenous stimulus. Typically, the protein fragments are genetically fused to two known interacting proteins. We chose to look at two different systems: CRY2/CIB1 light-inducible pairing (**Fig. 5.3**) and FKBP/FRB rapamycin-inducible pairing (**Fig. 5.4**).

All CIB1/CRY2 fusions showed fairly equal expression in mammalian cells, but were not as well expressed as the GS2-IpaH9.8 ubiquibody (**Fig. 5.3b**). FRB/FKBP fusions had highly variable expression (**Fig. 5.4b**). FKBP-IpaH9.8dLRR and FKBP-IpaH9.8dLRR^{C337A} fusions expressed higher than GS2-IpaH9.8, but FRB-IpaH9.8dLRR was undetectable and FRB-GS2 was extremely low. These results highlight the important role of the fusion partner in moderating expression level. Across the two sets of designs, IpaH9.8dLRR chimeras had very low expression, potentially due to auto-ubiquitination leading to proteasomal degradation.

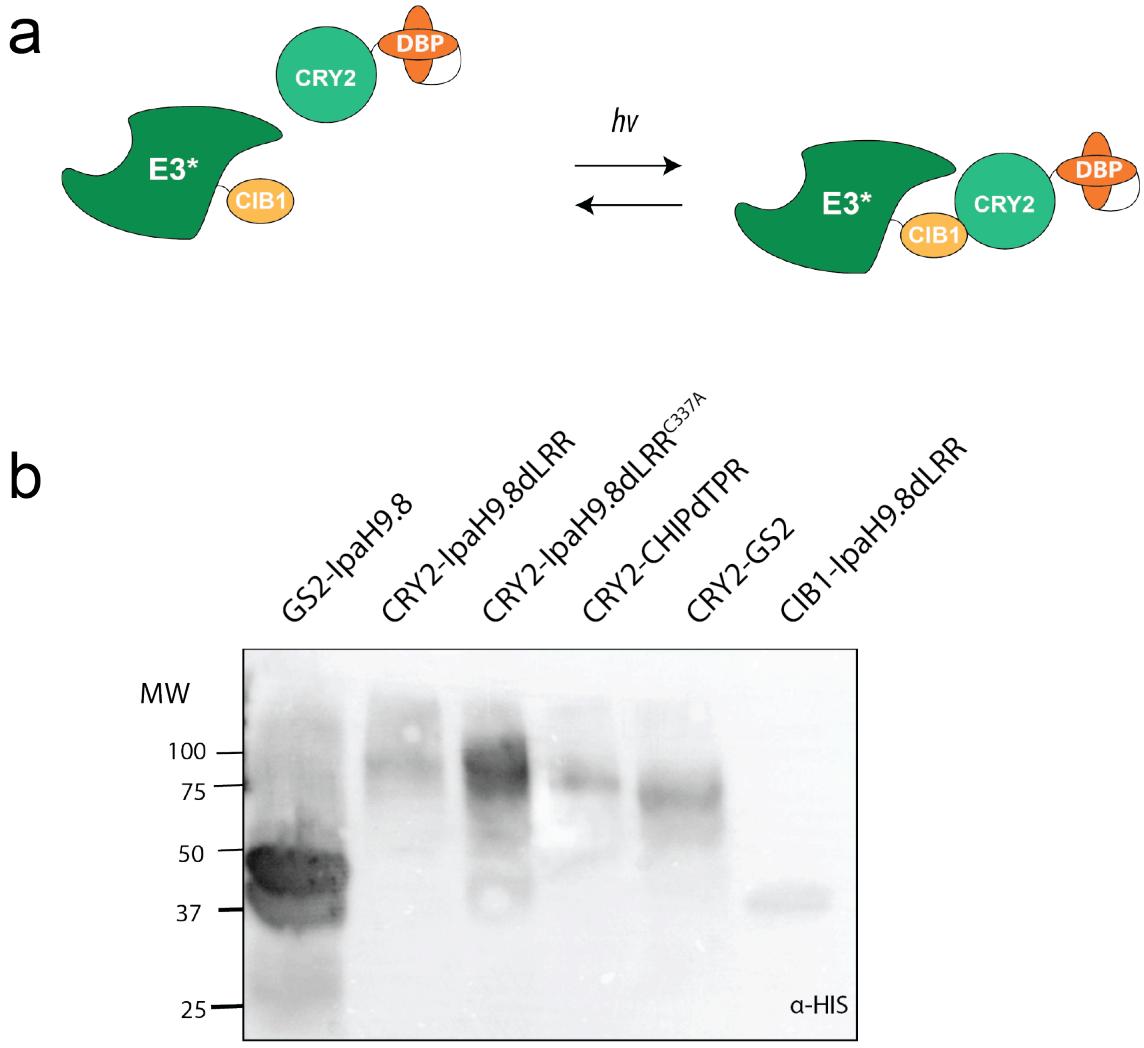


Figure 5.5 Mammalian expression of light-induced split ubiquibody designs.
Western blot analysis of HEK293T cell lysates-transfected with pcDNA3 plasmids encoding the indicated chimeras. Blots were probed with antibodies specific for 6x-His (that detected tag on each chimera). An equivalent amount of total protein was loaded in each lane. Molecular weight (MW) markers are indicated on left.

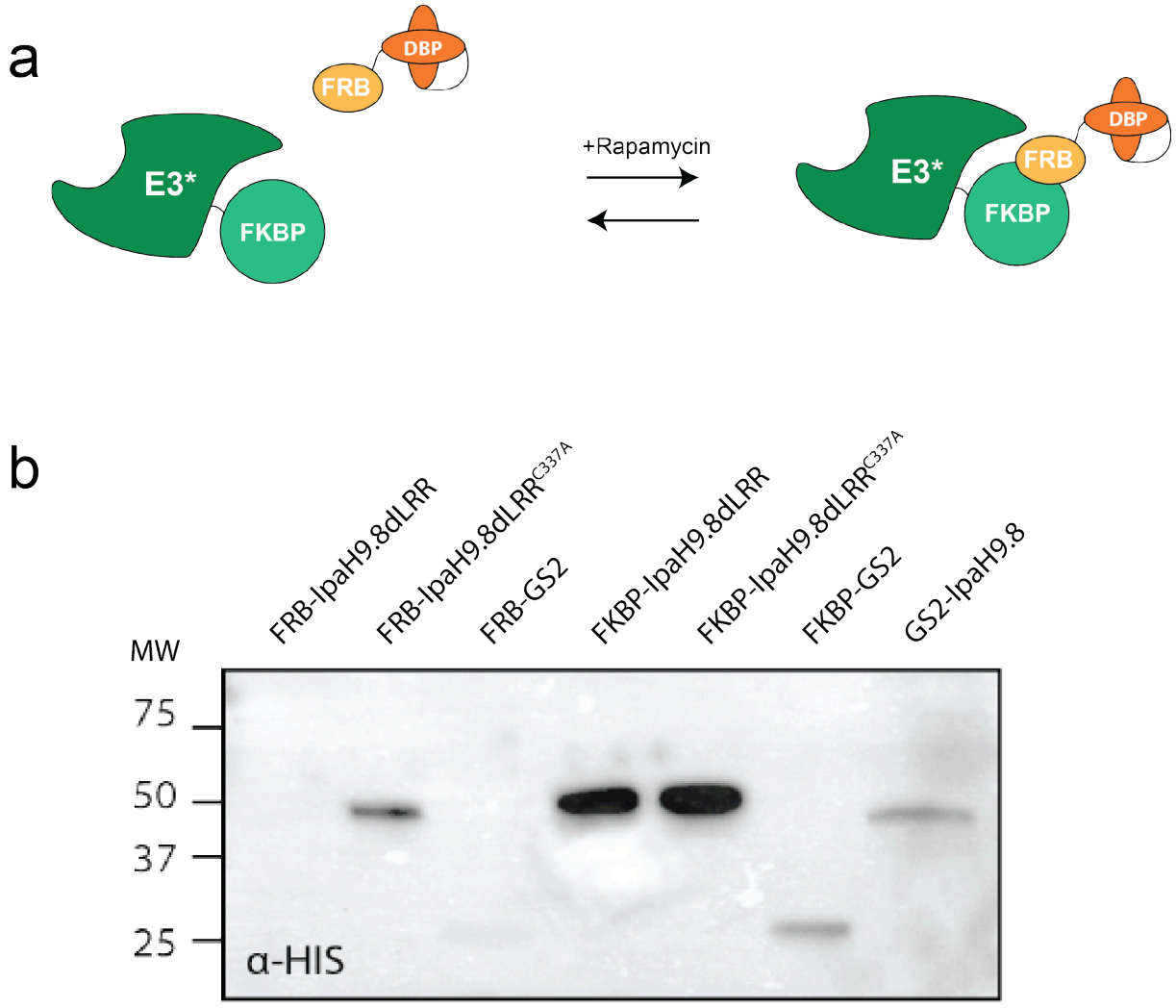


Figure 5.6 Mammalian expression of chemically-induced split ubiquibody designs. Western blot analysis of HEK293T cell lysates-transfected with pcDNA3 plasmids encoding the indicated chimeras. Blots were probed with antibodies specific for 6x-His (that detected a tag on each chimera). An equivalent amount of total protein was loaded in each lane. Molecular weight (MW) markers are indicated on left.

Reducing IpaH9.8 ubiquibody auto-ubiquitination

Previously, our lab had detailed the reduction in the auto-ubiquitination of the E3 ubiquitin ligase CHIP through mutation of lysine residues [142]. Here, we wanted to investigate if GS2-IpaH9.8 could maintain its exquisite degradation activity after lysine mutagenesis. GS2 contains two lysines, the C-terminal catalytic domain of IpaH9.8 contains nine (**Fig. 5.4**), and the recombinant Flag tag contains two. When single lysine mutants of GS2-IpaH9.8 were co-expressed in mammalian cells with EGFP, degradation activity equivalent to the parental GS2-IpaH9.8 ubiquibody was observed (**Fig. 5.5a**). These mutants cover every lysine except for IpaH9.8^{K350}. When these same constructs were evaluated for their expression in HEK293T cells, most mutants showed equivalent expression to the parental GS2-IpaH9.8, although GS2^{K55A}-IpaH9.8 appeared to have a significant increase.

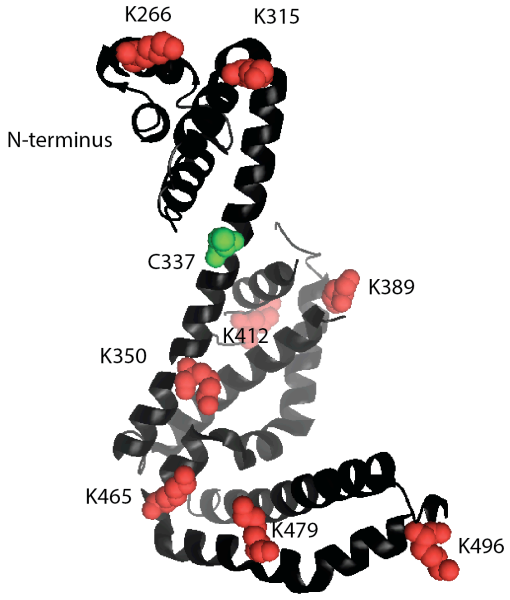


Figure 5.7 Crystal structure of IpaH9.8 highlighting lysine residues.
Crystal structure of the C-terminal NEL domain of IpaH9.8 (PDB 3l3p). Lysine (red) and cysteine (green) residues are colored for visualization.

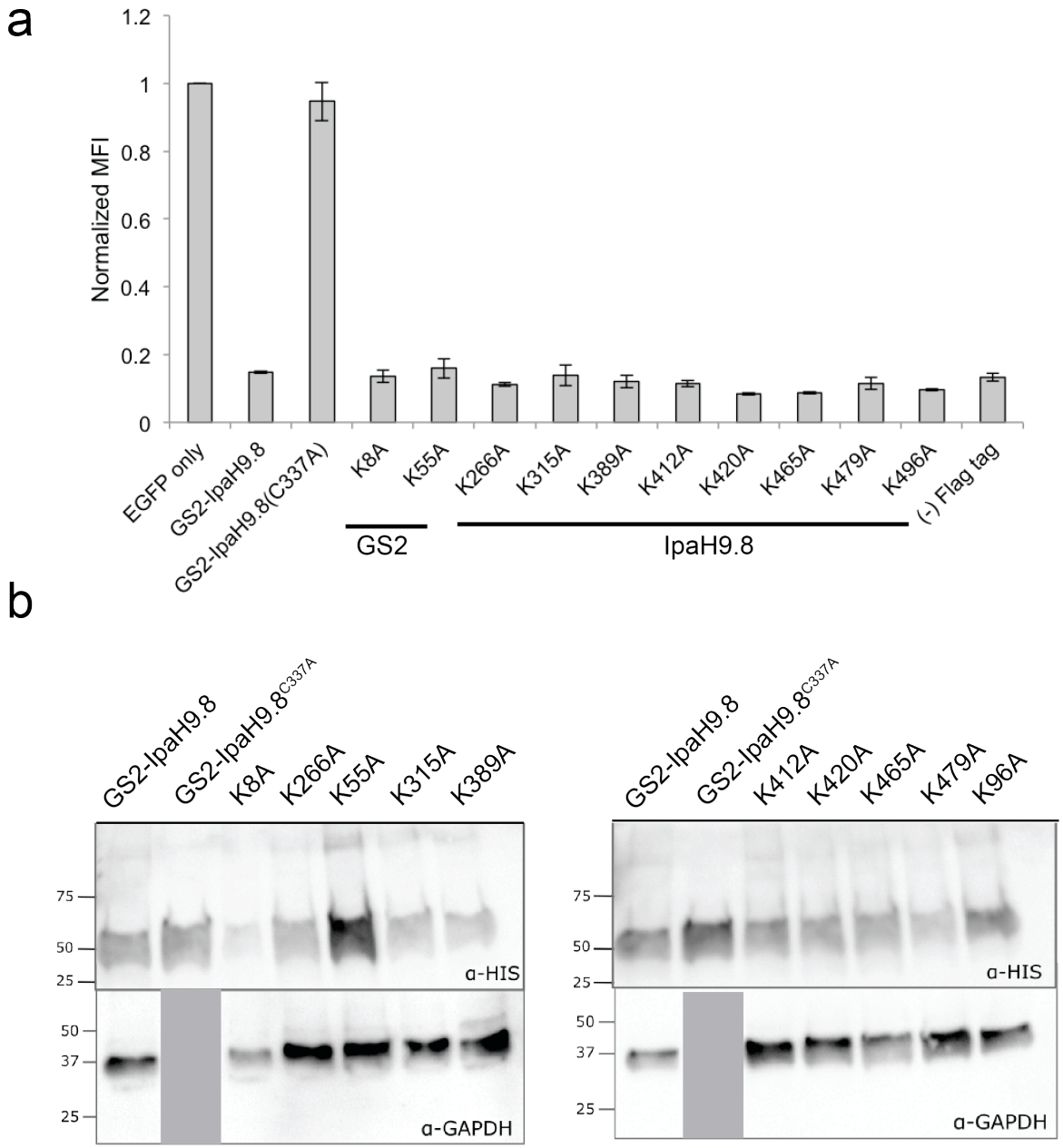


Figure 5.8 Characterization of GS2-IpaH9.8 lysine mutants in mammalian cells.
(a) Flow cytometric quantification of EGFP fluorescence activity in HEK293T cells transfected with pcDNA3-EGFP alone or cotransfected with pcDNA3-EGFP and a plasmid encoding the indicated ubiquibody. Data are biological triplicates (three separately transfected wells) of the geometric MFI normalized to MFI measured for HEK293T cells expressing EGFP alone. Error bars represent standard deviation (SD) of the mean. *, $p < 0.01$ (compared to HEK293T cells expressing EGFP only). (b). Western blot analysis of HEK293T cell lysates-transfected with pcDNA3 plasmids encoding GS2-IpaH9.8, GS2-IpaH9.8^{C337A}, or GS2-IpaH9.8 with the indicated lysine mutation. Blots were probed with antibodies specific for GFP, 6x-His (that detected tag on each ubiquibody), and GAPDH as indicated. An equivalent amount of total protein was loaded in each lane as confirmed by immunoblotting with anti-GAPDH. Molecular weight (MW) markers are indicated on left.

Discussion

Ubiquibodies are a powerful tool for the targeted degradation of protein targets in mammalian cells, and have ben shown to be effective when delivered as mRNA nanoplexes, highlighting their therapeutic potential. However, to reduce drug side effects, there has been a wide range of targeted drug delivery approaches developed [143]. However, another approach is to indiscriminately deliver therapeutic cargo, but employ light-based strategies for the activation only in the desired cells at a designated time [144].

Many have reported spatial and temporal control of specific protein function through the chimeric fusion of light-gated protein domains. We have preliminary evidence that the two most characterized approaches – AsLOV2 fusion and CIB1/CRY2 split protein fusion – are compatible with our ubiquibody technology. The use of a light-gated split ubiquibody really leverages the modularity of the ubiquibody design, separating binding and catalytic activity. Unfortunately this also means delivering two separate proteins and balancing the different expression levels of each to maintain an effective stoichiometry within the cell.

For this reason, the addition of the AsLOV2 light-gated domain to the ubiquibody design could be advantageous. While the designs we presented here attempt to light-gate the catalytic activity of the IpaH9.8 NEL domain, one could also envision moving AsLOV2 to the N-terminus of the binding domain, in an attempt to light-gate binding, hopefully prevent recognition of the target, except in the presence of a light stimulus. One of the major challenges associated with light-responsive proteins is the toxicity and lack of tissue penetration of blue-light wavelengths.

We observed highly variable expression levels of all light-induced fusions, particularly with IpaH9.8 NEL domain fusions. IpaH9.8 naturally exhibits auto-inhibitory functionality when its native LRR domain is not bound to substrate [145]. It is likely that the fusions explored here do not offer the same protection for the catalytic cysteine, leading to higher auto-ubiquitination and subsequent degradation. In a therapeutic context, this is a challenge as it necessitates the delivery of higher quantities of ubiquibody to achieve the necessary intracellular concentration. While a completely lysine-free GS2-IpaH9.8 mutant was not active, the single lysine mutants showed no negative impact to their activity. Thus, while more research must be done to investigate the effect on these mutations on the GS2-IpaH9.8 half-life, lysine modification provides another method that we can employ to tune improve ubiquibody function.

Materials and Methods

Strains and Plasmids

All plasmids used in this study are provided in **Appendix C**. For creation of light-gated ubiquibodies, plasmid pTriEx-3-GS2 was created by PCR amplification of GS2 from pHFT2-GS2 [95] using primers that introduced N-terminal Kozak and C-terminal 6x-His, and FLAG sequences followed by ligation into pTriEx-3 such that EcoRV and HindIII restriction sites were available downstream of GS2 for generating C-terminal fusions. The register of AsLOV2-IpaH9.8 fusions was created using overlap extension PCR which added N-terminal NheI site and C-terminal Flag and 6xHis tags and HindIII site. For the initial AsLOV2 amplification, the plasmid pBabe-CMV-puro-Tet-mVenus-Lov obtained from Addgene (#22033) served as the template.

Split ubiquibodies were created by PCR amplification of CRY2 from CRY2(535,L348F)-mCherry (Addgene #75372), CIB1 from pmGFP-CIB1 (Addgene #28240), FRB from pCherry-FRB (Addgene #25920), and FKBP from CFP-FKBP (Addgene #20160) with primers that introduced either 5' HindIII and Kozak sequences and 3' EcoRV site for C-terminal fusions or 5' EcoRV and 3' XbaI sites for N-terminal fusions. All plasmids were verified by DNA sequencing at the Cornell Biotechnology Resource Center (BRC).

Flow cytometry

Cells were passed into 12-well plates at 10,000 cells/cm^2. 16-24 h after seeding, cells were transiently transfected with 1 μg total DNA at a 1:2 ratio of DNA:jetPrime (Polyplus Transfection). Cells were transfected with 0.05 μg of target, 0.25 μg of ubiquibody or control, and balanced with empty pcDNA3 vector. Culture media was replaced 4-6 h post-transfection. Then, 24 h post-transfection, cells were harvested and resuspended in phosphate buffered saline (PBS) for analysis using a BD FACSCalibur or BD FACSAria Fusion. FlowJo Version 10 was used to analyze samples by geometric mean fluorescence determined from 10,000 events.

Western blotting

HEK293T cells were plated at 10,000 cells/cm2 and transfected as described above before lysis with RIPA lysis buffer (Thermo Fisher). MCF10A cells were plated at 20,000 cells/cm^2 and induced with 0.2 μg/mL doxycycline for 24 h before lysis with cell lysis buffer. Lysates were separated on Any kD polyacrylamide gels (Bio-Rad) and transferred to PVDF membranes. α-HIS-HRP (Abcam), α-GAPDH (Millipore), and mouse anti-ubiquitin (Millipore, P4D1-A11) antibodies were diluted 1:5,000 and in TBST

+ 1% milk and incubated for 1 h at room temperature. Secondary antibody goat anti-mouse IgG with HRP conjugation (Promega) was diluted at 1:2,500 and used as needed.

E. coli ubiquitination pathway reconstitution

pCDF Duet UbcH5a/Ube1 was obtained by Su et al., J. Immunol., 2006. Ubiquitin was cloned without any tags into the second MCS of pACYC, with the MCSI used for the designated E3. Targets were expressed from their own plasmid. The indicated plasmids were transformed into BL21(DE3) cells. Overnight cultures were subcultured into 5mL of media + antibiotics at 1:100 at 37°C. At OD_{600} = 0.5-0.8 cultures were induced with a final IPTG concentration of 1 mM and incubated for 3 hrs at 37°C. 1 mL of each culture was harvested by centrifugation at 13,200xg for 20 min. at 4°C. Supernatant was removed and pellets lysed with 200uL of BugBuster by rotating for 20 min. at room temperature. After centrifugation at 13,200xg for 20 min. at 4°C, the supernatant was collected as the soluble fraction and boiled with 2X SDS Loading Buffer for 15 min. at 100°C. Bradford Assay determined soluble fraction total protein quantity. Western Blots were loaded accordingly, normalizing for total protein content.

Protein expression and purification

Purified proteins were obtained by growing Rosetta(DE3) cells containing a pTriEx-3-based plasmid in 200 mL of Luria-Bertani (LB) medium supplemented with 50 μg/mL Ampicillin and 25 μg/mL Chloramphenicol at 37°C. Expression was induced with 0.1 mM IPTG when the culture density (Abs600) reached 0.6-0.8 and growth continued for 6 h at 30°C. Cultures were harvested by centrifugation at 4,000xg for 30 min at 4°C. Cell pellets were stored at -20°C overnight. Thawed pellets were resuspended in 10 mL equilibration buffer (25 mM Tris-HCl, pH 7.4, 500 mM NaCl and

20 mM imidazole) and lysed with a high-pressure homogenizer (Avestin Emulsi-Flex C5). The insoluble fraction was cleared by centrifugation at 12,000xg for 30 min at 4°C. His-tagged protein was purified by gravity flow using 500 µL HisPur Ni-NTA resin (ThermoFisher). The soluble fraction was passed through the resin, after which the resin was washed with 3 mL of wash buffer (25 mM Tris-HCl, pH 7.4, 500 mM NaCl and 50 mM imidazole). Protein was eluted with 1.5 mL elution buffer (25 mM Tris-HCl, pH 7.4, 500 mM NaCl and 250 mM imidazole). Purified fractions were desalted and concentrated (Pierce PES Protein Concentrators).

Spectra

Fluorescence spectra were obtained at an excitation wavelength of 450 nm using a TECAN Infinite M1000 Pro.

Blue-light device

A blue-light device was created following the instructions found in Support Protocol 1 in Tucker et al., 2015 [146].Twelve Blue (470 nm) Luxeon Reel LED on a SinkPAD-II 10 mm Square Base (SP-05-B4) and a 70 mm x 130 mm rectangular 20 mm High Alpha Heat Sink (2.1 °C/W) were ordered from luxeonstar.com. DFRobot DFR00009 LCD Shield for Arduino and Arduino Uno R3 Microcontroller A000066 were ordered fro Amazon. Images of the blue light device can be found in Appendix B.

CHAPTER 6 - FUTURE DIRECTIONS

Introduction

Protein knockout has always been an attractive way to elucidate cellular functions and holds therapeutic potential for the elimination of disease-relevant proteins. Recently, there has been a growing interest in targeted protein degradation techniques due to their catalytic activity, which requires lower intracellular concentrations than binding-through-inhibition techniques and their protein-level specificity, which enables the discrimination of post-translationally modified proteins

Our group has been developing ubiquibodies, which combine the catalytic ubiquitination activity of the mammalian E3 ubiquitin ligase CHIP and the specificity of designer binding proteins to eliminate protein targets from mammalian cells. In this work, we leveraged flow cytometry to rapidly screen a large variety of alternative E3 ubiquitin ligases for ubiquibody activity, which led to the discovery of the potent bacterial IpaH9.8 E3 ubiquitin ligase catalytic domain. Fusion of IpaH9.8 to the GFP-specific FN3 binding domain GS2- enabled the robust degradation of EGFP and EGFP-tagged mammalian proteins in all subcellular compartments and from all expression modalities. Furthermore, IpaH9.8 ubiquibodies with new specificities were rapidly generated by swapping the GFP-specific binding domain with designer binding domains that recognized SHP2, KRas, and ERK2. We also successfully demonstrated proteome editing in mice through the delivery of ubiquibodies packaged as mRNA nanoplexes. Furthremore, we detailed preliminary designs for the endowment of ubiquibodies with light-gated control. These results demonstrate modularity of the ubiquibody design and the therapeutic potential of ubiquibodies. However, there are still many facets of

ubiquibody-mediated protein degradation that necessitate more investigation and engineering.

Discussion

Continued exploration of bacterial E3 ubiquitin ligases

In this work, we analyzed eighteen bacterial E3 ubiquitin ligases for their ability to be engineered as GFP-specific ubiquibodies, which at the time of our work were all the bacterial E3 ligases with characterized binding and catalytic domains – necessary information their manipulation into ubiquibodies. The first bacterial E3 ubiquitin ligases was not identified until AvrPtoB in 2006 [147], therefore it is high likely that there are many more uncharacterized bacterial effectors that possess E3 ubiquitin ligase activity. Every E3 ubiquitin ligase is unique and could offer advantageous characteristics for future ubiquibodies. One could even envision adapting our flow cytometric based method for evaluating characterizing new ubiquibodies for the validation of purported new E3 ubiquitin ligases.

Therapeutic delivery for endogenous target degradation

Engineered degraders must be able to efficiently enter cells if they are to be used as therapeutics. Efforts to overcome the delivery challenge in the context of targeted protein degradation have primarily focused on viral delivery methods [53, 55, 56], although the work presented here demonstrated the potential of nanoplex-mediated delivery of synthetic mRNA encoding ubiquibodies to achieve highly efficient silencing *in vitro* and *in vivo* [63]. In another recent study, Hantschel and coworkers used a chimeric bacterial toxin to deliver ubiquibodies in cancer cells [62]. Although not yet demonstrated for protein-based degraders, a number of other strategies have been

described for internalizing intact proteins including the use of cationic lipids [148], cell-penetrating peptides [149], covalent attachment of oligosaccharides [150], esterification [151], and polymeric nanoparticles [152]. Additionally, continued efforts will need to be made to demonstrate the efficacy of ubiquibodies against endogenous targets. While in this work we were able to target overexpressed mammalian proteins as FP fusions, endogenous protein degradation warrants substantial investigation into the necessary ubiquibody characteristics such as affinity, half-life, delivery method, and host-cell effects.

Endowing ubiquibodies with spatiotemporal control

In addition to the continued development of delivery methods that can achieve therapeutically relevant concentrations of protein into cells, more work will also be needed to deliver proteins in a tissue-specific manner with minimal off-target effects. This could be accomplished by generating stimuli-sensitive switches that effectively control degrader activity. Along these lines, Ostermeier and coworkers have described methods for creating enzyme-binding protein hybrids whose enzymatic activity is activated only in the presence of antigen [153]. Another option would be the creation of light-inducible switches, such as the designs described in this work that could be conditionally activated in precise areas of the body, preferably using longer wavelengths of light. Towards this goal, Taxis and coworkers developed a generic photosensitive degron combining the light-reactive LOV2 domain with the mODC PEST sequence that provided synthetic light control of protein stability [131]. Overall, to advance from the lab to the clinic, these and other protein engineering advancements will be crucial for the development of protein-based degraders against virtually any protein of interest that

can be delivered efficiently and specifically in the body for the selective removal of disease-associated proteins.

Yeast surface display for development of new ubiquibodies

An outstanding question regarding ubiquibody design remains - is there an ideal designer binding domain scaffold for use in ubiquibody fusions? Results in Chapter 3 provide preliminary evidence that FN3s are more compatible than DARPins or scFvs and results from Chapter 2 comparing the degradation activity of EGFP by IpaH9.8 fused to a panel of GFP-specific FN3s indicates there may be an advantage to higher affinity, but there is not enough data to draw a definitive comparison. In this work, however, we demonstrated that yeast surface display could rapidly identify novel FN3 binders that can be evaluated for efficacy as ubiquibody fusions (Chapter 3). Similarly, yeast surface display could be used to identify novel binders to the same target protein from the range of available binding scaffolds: scFv, FN3, DARPin, sso7d, vhh, and VLR. Isolated binders could then be evaluated for their ability to degrade that target as ubiquibody fusions to IpaH9.8 or any other E3 ubiquitin ligase catalytic domain. Successful and unsuccessful ubiquibodies could be analyzed for characteristics such as target affinity, binding epitope, expression level, and half-life to elucidate a set of design rules for more efficient ubiquibody development in the future.

One could also envision the use of yeast surface display for the high-throughput characterization of the full-length ubiquibody, not just the binding domain. Directed evolution could be used to enhance the catalytic performance of the E3 domain or reprogram its affinity and/or selectivity for different E2 partners. For instance, Klevit and coworkers developed a high-throughput mutagenesis strategy to probe the

mechanism of E3-catalyzed transfer of ubiquitin from the E2 to the target protein [154]. By scoring the effect of nearly 100,000 mutations in an E3, they identified mutations that enhanced E3 activity both *in vitro* and in cellular degradation assays. This mutagenesis and screening approach not only uncovered activity-enhancing mutations but also helped to delineate the molecular basis of E3 ligase activity, which could be used in the future to rationally design even better catalysts. Finally, while ubiquibodies have focused exclusively on replacing the native substrate-binding domains (*e.g.*, 3-tetratricopeptide repeat (TPR) in human CHIP; LRR in *S. flexneri* IpaH9.8) of the E3 with synthetic binding proteins, one could instead change the specificity of these domains through directed evolution, thereby obviating the need for their removal. In support of this concept, combinatorial libraries based on TPR and LRR scaffolds have been constructed, and following high-throughput screening variants with custom binding specificities and affinities were isolated [155, 156].

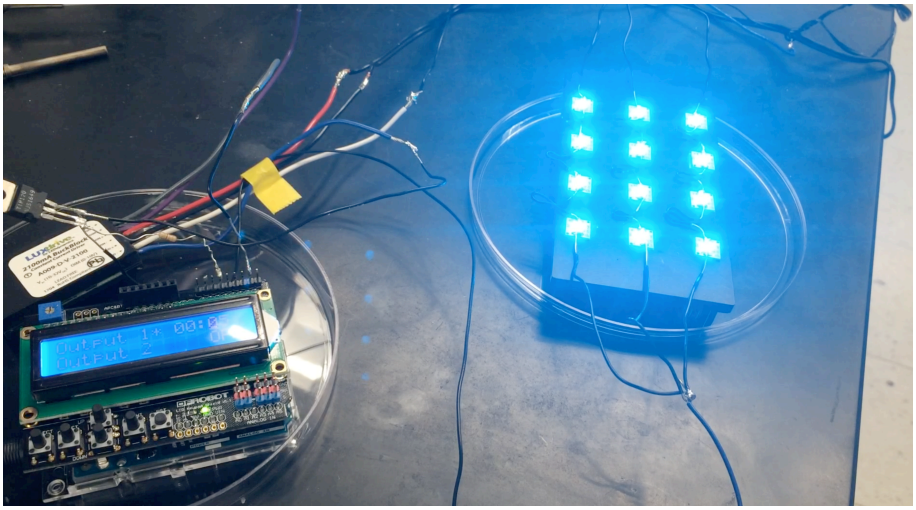
APPENDIX A – BACTERIAL E3 UBIQUITIN LIGASES

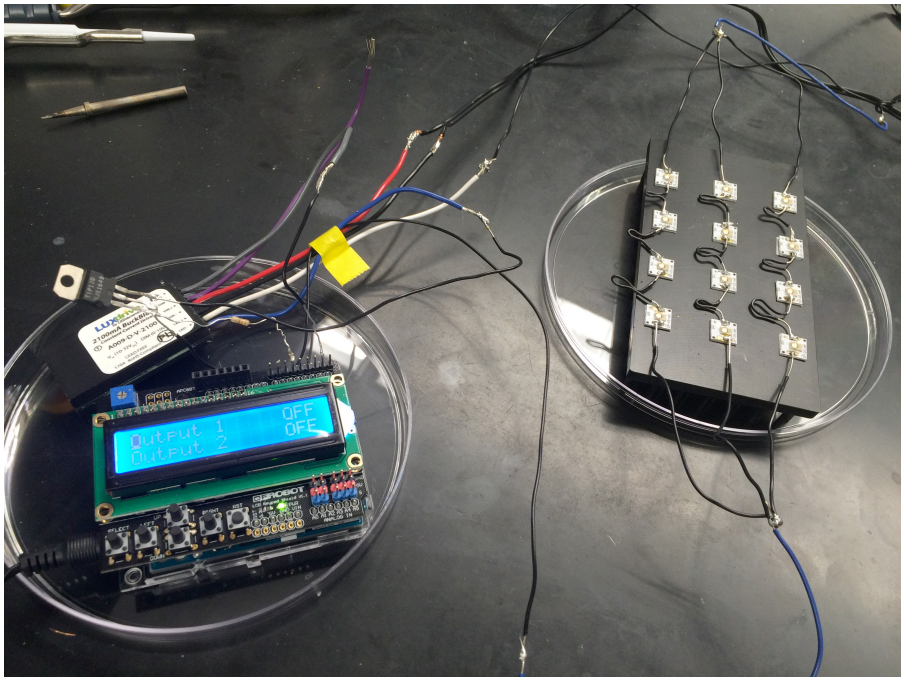
E3 ubiquitin ligase	Classification	Organism	Construction	Ref.[1]
AvrPtoB	U-box	*Pseudomonas syringae*	$AvrPtoB_{1-436}$ fused to N-terminus of GS2	[157]
βTrCP	F-box	*Homo sapiens*	$\beta TrCP_{1-568}$ fused to N-terminus of GS2 with $(GlySer)_{10}$ linker	[51]
CHIP	U-box	*H. sapiens*	$CHIP_{128-303}$ fused to C-terminus of GS2	[57]
IpaH0722	NEL	*Shigella flexneri*	$IpaH0722_{295-587}$ fused to C-terminus of GS2	[158]
IpaH1.4	NEL	*S. flexneri*	$IpaH1.4_{285-575}$ fused to C-terminus of GS2	[[158]
IpaH2.5	NEL	*S. flexneri*	$IpaH2.5_{292-570}$ fused to C-terminus of GS2	[158]
IpaH4.5	NEL	*S. flexneri*	$IpaH4.5_{296-574}$ fused to C-terminus of GS2	[158]
IpaH7.8	NEL	*S. flexneri*	$IpaH7.8_{274-565}$ fused to C-terminus of GS2	[158]
IpaH9.8	NEL	*S. flexneri*	$IpaH9.8_{254-545}$ fused to C-terminus of GS2	[159]
LegAU13	F-box	*Legionella pneumophila*	$LegAU13_{1-50}$ fused to N-terminus of GS2	[160]
LegU1	F-box	*L. pneumophila*	$LegU1_{1-56}$ fused to N-terminus of GS2	[160]
LubX	U-box	*L. pneumophila*	$LubX_{1-215}$ fused to N-terminus of GS2	[161]
NleG2-3	U-box	Enterohemorrhagic *Escherichia coli* (EHEC) O157:H7	$NleG2\text{-}3_{90-191}$ fused to C-terminus of GS2	[162]
NleG5-1	U-box	EHEC O157:H7	$NleG5\text{-}1_{113-213}$ fused to C-terminus of GS2	[162]
NleL	HECT	EHEC O157:H7	$NleL_{371-782}$ fused to C-terminus of GS2	[163]
SidC	Unconventional	*L. pneumophila*	$SidC_{1-542}$ fused to N-terminus of GS2	[75]
Slmb	F-box	*Drosophila*	$Slmb_{199-510}$ fused to C-terminus of GS2	[59]

		melanogaster		
SlrP	NEL	EHEC O157:H7	$SlrP_{465-765}$ fused to N-terminus of GS2	[164]
SopA	HECT	*Salmonella typhimurium*	$SopA_{370-782}$ fused to C-terminus of GS2	[165]
SPOP	F-box	*H. sapiens*	$SPOP_{167-374}$ fused to C-terminus of GS2	[65]
SspH1	NEL	*S. typhimurium*	$SspH1_{404-701}$ fused to N-terminus of GS2	[166]
SspH2	NEL	*S. typhimurium*	$SspH2_{492-788}$ fused to C-terminus of GS2	[167]
VHL	SCF-like ECV	*H. sapiens*	VHL_{1-213} fused to N-terminus of GS2	[168]
XopL	Unconventional	*Xanthomonas campestris*	$XopL_{474-660}$ fused to C-terminus of GS2	[92]

Table A.1 Bacterial E3 ubiquitin ligases used in this study.

APPENDIX B – BLUE-LIGHT DEVICE IMAGES





APPENDIX C – STRAINS AND PLASMIDS

Strain		
DH5α	F- (Φ80Δ*lacZ*DM15,) Δ(*lacIZYA-argF*)U169 *recA1 endA1 hsdR17*(r_k-,m_k+) *phoA supE44 λ- thi-1 gyrA96 relA1*	Laboratory stock
BL21(DE3)	F- *ompT gal dcm lon hsdS*$_B$(r_B^- m_B^-) λ(DE3)	Laboratory stock
Rosetta(DE3)	F- *ompT gal dcm lon hsdS*$_B$(r_B^- m_B^-) λ(DE3) pRARE (Cm^R)	Laboratory stock
EBY100	MATa AGA1::GAL1AGA1::URA3 ura352 trp1 leu2-delta200 his3delta200 pep4::HIS3 prbd1.6R can1 GAL	Laboratory stock
Cell line		
HEK293T		Laboratory stock
HEK293T-EGFP	HEK293T cells stably expressing EGFP	This study
HEK293T-ERK2-EGFP	HEK293T cells stably expressing ERK2-EGFP	This study
HEK293T-EGFP-$HRas^{G12V}$	HEK293T cells stably expressing EGFP-$HRas^{G12V}$	This study
HEK293T-d2EGFP	HEK293T cells stably expressing d2EGFP	This study
HeLa H2B-EGFP	HeLa cells stably expressing H2B-EGFP fusion	[59]
MCF-10A		Laboratory stock
MCF-10A rtTA		Laboratory stock
MCF-10A rtTA EGFP-$HRAS^{G12V}$	MCF-10A cells stably expressing EGFP-$HRAS^{G12V}$ fusion and rtTA	This study
MCF-10A rtTA GS2-IpaH9.8	MCF-10A cells stably expressing GS2-IpaH9.8 fusion and rtTA	This study
MCF-10A rtTA EGFP-$HRAS^{G12V}$ + GS2-IpaH9.8	MCF-10A cells stably expressing EGFP-$HRAS^{G12V}$ fusion, GS2-IpaH9.8, and rtTA	This study
MCF-10A rtTA EGFP $HRAS^{G12V}$ + GS2-IpaH9.8^{C337A}	MCF-10A cells stably expressing EGFP-$HRAS^{G12V}$ fusion, GS2-IpaH9.8^{C337A}, and rtTA	This study
Plasmid	**Relevant features**	**Source**
pcDNA3	CMV promoter, Amp^R	Laboratory stock
pCDH1-CMV-MCS-EF1α-Puro	CMV promoter; Pur^R, Amp^R	Laboratory stock
pET28a(+)	T7lac promoter; Kan^R	Novagen
pET24d(+)	T7lac promoter; Kan^R	Novagen
pTriEx-3	CMV promoter; T7 promoter; Amp^R	Novagen
psPAX2	Lentiviral packaging vector; CMV promoter; Amp^R	Laboratory stock
pMD2.G	Lentiviral packaging vector; CMV promoter; Amp^R	Laboratory stock
pPB TetOn Hygro	Transposon vector; SV40 promoter; Amp^R	Laboratory stock
Mammalian Cell Line Generation		
pLV rtTA-NeoR	Tetracycline inducible reverse transcriptional transactivator; CMV promoter; EF1α promoter; Neo^R; Amp^R	Laboratory stock
pCMV-hyPBase	PiggyBac transposase; CMV promoter; Amp^R	[169]
pHIV-d2EGFP	Lentiviral vector expressing d2EGFP EF-1^{α} promoter, IRES; Amp^R	[170]
pPB-GS2-IpaH9.8	GS2-IpaH9.8 with C-terminal Flag and 6x-His tags cloned in pPB TetOn Hygro; CMV promoter; SV40 promoter; Amp^R	This study
pPB-GS2-IpaH9.8^{C337A}	GS2-IpaH9.8^{C337A} with C-terminal Flag and 6x-His tags cloned in pPB TetOn Hygro; CMV promoter; SV40 promoter; Amp^R	This study

pCDH1-ERK2-EGFP	ERK2 fused to N-terminus of EGFP cloned in pCDH1-CMV-MCS-EF1α-Puro; CMV promoter; Pur^R, Amp^R	This study
pCDH1-EGFP	EGFP cloned in pCDH1-CMV-MCS-EF1α-Puro; CMV promoter; Pur^R, Amp^R	This study
pCDH1-EGFP-$HRas^{G12V}$	$HRas^{G12V}$ fused to C-terminus of EGFP cloned in pCDH1-CMV-MCS-EF1α-Puro; CMV promoter; Pur^R, Amp^R	This study
mRNA Production		
pGEM4Z/GFP/A64	*In vitro* transcription of GFP with 3' 64 residue poly A tail; T7 promoter; Amp^R	[171]
pGEM4Z/GS2-IpaH9.8/A64	In vitro transcription of GS2-IpaH9.8 with 3' human globin UTR and 64 residue polyA tail; T7 promoter; Amp^R	This study
pGEM4Z/GS2-IpaH9.8^{C337A}/A64	In vitro transcription of GS2-IpaH9.8^{C337A} with 3' human globin UTR and 64 residue polyA tail; T7 promoter; Amp^R	This study
pGEM4Z/AS15-IpaH9.8/A64	In vitro transcription of AS15-IpaH9.8 with 3' human globin UTR and 64 residue polyA tail; T7 promoter; Amp^R	This study
Bacterial Protein Expression		
pSPI02B-BirA-His	BirA expressed with C-terminal 6xHis tag; Tac promoter; $Strep^R$	Laboratory stock
pCDF Duet UbcH5α/Ube1	UBE2D1 expressed from MCS1 with N-terminal His tag; UBE1 expressed from MCS2 with C-terminal S-Tag; T7lac promoters; $Strep^R$	Laboratory stock
pACYC Ubiquitin	Ubiquitin expressed from MCS2; T7 promoter; Cm^R	This study
pETHis6MEK1 R4F+ERK2	pET-based expression of two proteins: constitutively active MEK1 and wild-type, His-tagged ERK2	[172]
pET28-EGFP	EGFP with C-terminal 6x-His tag in pET28a(+); T7lac promoter; Kan^R	This study
pET28(+)-GS2	GS2 with C-terminal Flag and 6x-His tags in pET28a(+); T7lac promoter; Kan^R	This study
pET24d(+)-GS2-IpaH9.8	GS2 fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags in pET24d(+); T7lac promoter; Kan^R	This study
pET24d(+)-IpaH9.8dLRR	IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags in pET24d(+); T7lac promoter; Kan^R	This study
pET24d(+)-IpaH9.8dLRR	IpaH9.8 domain with C-terminal Flag and 6x-His tags in pET24d(+); T7lac promoter; Kan^R	This study
pET28(+)-2αH-GS2-IpaH9.8	2αH fused to GS2-IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags in pET28a; T7lac promoter; Kan^R	This study
pET28(+)-2αH-GS2-IpaH9.8^{C337A}	Same as pET28a(+)-2αH-GS2-IpaH9.8, but includes cysteine to alanine mutation C337A in IpaH9.8.	This study
pTriEX-3-GS2-IpaH9.8^{C337A}	GS2 fused to IpaH9.8 lacking LRR domain and containing Cys337Ala mutation in pTriEx-3; CMV promoter; T7 promoter; Amp^R	This study
Synthetic Binding Domains		
pHFT2-GS2	GS2 monobody in pHFT2 plasmid; T7 promoter, Kan^R	[95]
pHFT2-GS5	GS5 monobody in pHFT2 plasmid; T7 promoter, Kan^R	[95]
pHFT2-GL4	GL4 monobody in pHFT2 plasmid; T7 promoter, Kan^R	[95]
pHFT2-GL6	GL6 monobody in pHFT2 plasmid; T7 promoter, Kan^R	[95]
pGalAga-GL8	GL8 monobody in pGalAgaCamR; Cam^R	[95]

pHFT2-AblSH2MB#AS15	AS15 monobody in pHFT2 plasmid; T7 promoter, Kan^R	[21]
pHFT2-SHP2-NSa1	NSa1 monobody in pHFT2 plasmid; T7 promoter, Kan^R	[112]
pHFT2-SHP2-NSa5	NSa5 monobody in pHFT2 plasmid; T7 promoter, Kan^R	[112]
pHFT2-SHP2-CS3	CS3 monobody in pHFT2 plasmid; T7 promoter, Kan^R	[112]
pET24a(+)-RasInI	RasInI with C-terminal GS linker, 6xHis, Avi, and Flag tags; T7lac promoter; Kan^R	[113]
pET24a(+)-RasInII	RasInII with C-terminal GS linker, 6xHis, Avi, and Flag tags; T7lac promoter; Kan^R	[113]
pE-Sumo-R11.1.6	R11.1.6 with N-terminal 6xHis and SUMOpro tags; T7 promoter; Amp^R	[114]
Protein Degradation Targets		
pLK1-ERK2	ERK2 with N-terminal Avi and C-terminal 6xHis tags; Amp^R	Laboratory stock
pHFT2-SHP2	Full-length SHP2 in pHFT2 plasmid; T7 promoter, Kan^R	[112]
pcDNA3-EGFP	EGFP cloned in pcDNA3; CMV promoter, Amp^R; Kozak	This study
mEmerald-C1	CMV promoter; Kan^R	Laboratory stock
mVenus-N1	CMV promoter; Kan^R	Laboratory stock
pcDNA3-mCerulean	CMV promoter; Kozak; Amp^R	This study
pcDNA3-CFP	CMV promoter; Kozak; Amp^R	This study
pcDNA3-sfGFP	CMV promoter; Kozak; Amp^R	This study
pcDH1-mCherry	CMV promoter; EF1α promoter; $Puro^R$; Amp^R	Laboratory stock
mEmerald-α-actinin-19	α-actinin fused to N-terminus of mEmerald; CMV promoter; Kan^R	Laboratory stock
pcDNA3.1(-)-α-synuclein-EGFP	α-synuclein fused to the N-terminus of EGFP cloned in pcDNA3.1; CMV promoter; Amp^R	[173]
mEmerald-FAK-5	FAK fused to C-terminus of mEmerald; CMV promoter; Kan^R	Laboratory stock
pEGFP-C1 F-tractin-EGFP	F-tractin fused to N-terminus of EGFP; CMV promoter; Kan^R	Addgene #58473
EGFR-mEmerald	EGFR fused to N-terminus of mEmerald; CMV promoter; Amp^R	Laboratory stock
pErbB2-EGFP	ErbB2 fused to N-terminus of EGFP; CMV promoter; Amp^R	Addgene #39321
pcDNA3-ERK2-EGFP	ERK2 fused to EGFP in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pLV-pGK-H2B-EGFP	H2B fused to C-terminus of EGFP; pGK promoter; Amp^R	Laboratory stock
mEGFP-$HRas^{G12V}$	$HRas^{G12V}$ (constitutively active) fused to mEGFP cloned in pCI; CMV promoter; Amp^R	[174]; Addgene #18666
mEmerald-Muc1-FL	Muc1 fused to N-terminus of mEmerald; CMV promoter; Kan^R	Laboratory stock
pcDNA3-EGFP-NLS	NLS sequence derived from C-terminus of SV40 fused to C-terminus of EGFP and cloned in pcDNA3; CMV promoter; Amp^R	This study
mEmerald-Paxillin-22	Paxillin fused to N-terminus of mEmerald; CMV promoter; Kan^R	Laboratory stock
pcDNA3-SHP2-EGFP	SHP2 fused to C-terminus of EGFP; CMV promoter; Amp^R	This study
mEmerald-Vinculin-23	Vinculin fused to C-terminus of mEmerald; CMV promoter; Kan^R	Laboratory stock
pcDNA3-EGFP-KRas	KRas fused to C-terminus of mEmerald; CMV promoter; Amp^R	This study

pcDNA3-EGFP-KRas^{G12C}	KRas^{G12C} fused to C-terminus of mEmerald; CMV promoter; Amp^{R}	This study
pcDNA3-EGFP-KRas^{G12D}	KRas^{G12D} fused to C-terminus of mEmerald; CMV promoter; Amp^{R}	This study
pcDNA3-EGFP-KRas^{G12V}	KRas^{G12V} fused to C-terminus of mEmerald; CMV promoter; Amp^{R}	This study
IpaH GFP-specific Ubiquibodies		
pcDNA3-HF-GS2	GS2 with N-terminal Kozak, Flag and 6xHis tags,and BamHI and EcoRI sites; CMV promoter; Amp^{R}	This study
pcDNA3-GS2-FH	GS2 with C-terminal NheI and SbfI sites, Flag tag, and 6xHis tag; CMV promoter; Amp^{R}.	This study
pcDNA3-GS2-IpaH9.8	GS2 fused to *Shigella flexneri* IpaH9.8 lacking LRR domain (IpaH9.8ΔLRR) with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^{R}	This study
pcDNA3-GS2-IpaH9.8^{C337A}	pcDNA3-GS2-IpaH9.8 containing Cys337Ala substitution	This study
pcDNA3-GS2-IpaH1.4	GS2 fused to *S. flexneri* IpaH1.4 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^{R}	This study
pcDNA3-GS2-IpaH2.5	GS2 fused to *S. flexneri* IpaH2.5 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^{R}	This study
pcDNA3-GS2-IpaH4.5	GS2 fused to *S. flexneri* IpaH4.5 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^{R}	This study
pcDNA3-GS2-IpaH7.8	GS2 fused to *S. flexneri* IpaH7.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^{R}	This study
pcDNA3-GS2-IpaH0722	GS2 fused to *S. flexneri* IpaH0722 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^{R}	This study
pcDNA3-GS5-IpaH9.8	GS5 fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^{R}	This study
pcDNA3-GL4-IpaH9.8	GL4 fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^{R}	This study
pcDNA3-GL6-IpaH9.8	GL6 fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^{R}	This study
pcDNA3-GL8-IpaH9.8	GL8 fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^{R}	This study
pcDNA3-vhhGFP4-IpaH9.8	vhhGFP4 fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^{R}	This study
pcDNA3-2αH-GS2-IpaH9.8	pcDNA3-GS2-IpaH9.8 with N-terminal 2αH sequence.	This study
pcDNA3-GS2-IpaH9.8^{K8A}	Same as pcDNA3-GS2-IpaH9.8 but includes K8A mutation in GS2 protein.	This study
pcDNA3-GS2-IpaH9.8^{K55A}	Same as pcDNA3-GS2-IpaH9.8 but includes K55A mutation	This study

	in GS2 protein.	
pcDNA3-GS2-IpaH9.8^{K266A}	Same as pcDNA3-GS2-IpaH9.8 but includes K266A mutation in IpaH9.8 protein.	This study
pcDNA3-GS2-IpaH9.8^{K315A}	Same as pcDNA3-GS2-IpaH9.8 but includes K315A mutation in IpaH9.8 protein.	This study
pcDNA3-GS2-IpaH9.8^{K389A}	Same as pcDNA3-GS2-IpaH9.8 but includes K389A mutation in IpaH9.8 protein.	This study
pcDNA3-GS2-IpaH9.8^{K412A}	Same as pcDNA3-GS2-IpaH9.8 but includes K412A mutation in IpaH9.8 protein.	This study
pcDNA3-GS2-IpaH9.8^{K420A}	Same as pcDNA3-GS2-IpaH9.8 but includes K420A mutation in IpaH9.8 protein.	This study
pcDNA3-GS2-IpaH9.8^{K465A}	Same as pcDNA3-GS2-IpaH9.8 but includes K465A mutation in IpaH9.8 protein.	This study
pcDNA3-GS2-IpaH9.8^{K479A}	Same as pcDNA3-GS2-IpaH9.8 but includes K479A mutation in IpaH9.8 protein.	This study
pcDNA3-GS2-IpaH9.8^{K496A}	Same as pcDNA3-GS2-IpaH9.8 but includes K496A mutation in IpaH9.8 protein.	This study
pcDNA3-GS2-IpaH9.8-H	Same as pcDNA3-GS2-IpaH9.8 but lacks C-terminal Flag tag.	This study
Other GFP-specific Ubiquibodies		
pcDNA3-flag-βTrCP-10GS-scFv13R4	βTrCP expressed with N-terminal Flag tag and C-terminal GS linker and scFv13R4 fusion; CMV promoter, Amp^R	Laboratory stock
pcDNA3-R4-CHIPdTPR	scFv13-R4 fused to CHIP lacking TPR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	[57]
pcDNA3_NSlmb-vhhGFP4	vhhGFP4 fused to F-box domain from *Drosophila melanogaster* Slmb cloned in pcDNA3; CMV promoter, Amp^R	[59]; Addgene #35579
pcDNA3-GS2-AvrPtoB	GS2 fused to *Pseudomonas syringae* AvrPtoB lacking N-terminal domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-LegAU13-GS2	GS2 fused to *L. pneumophila* LegAU13 F-box with N-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-LegU1-GS2	GS2 fused to *L. pneumophila* LegU1 F-box with N-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-LubX-GS2	GS2 fused to *Legionella pneumophila* LubX lacking CTD domain with N-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-GS2-NleG2-3	GS2 fused to EHEC O157:H7 NleG2-3 lacking N-terminal domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-GS2-NleG5-1	GS2 fused to EHEC O157:H7 NleG5-1 lacking N-terminal domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-GS2-NleL	GS2 fused to EHEC O157:H7 NleL lacking β-helix domain with C-terminal Flag and 6x-His tags cloned in pcDNA3;	This study

	CMV promoter; Kozak; Amp^R	
pcDNA3-SidC-GS2	GS2 fused to *L. pneumophila* SidC lacking N-terminal domain with N-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-GS2-SlrP	GS2 fused to Enterohemorrhagic *Escherichia coli* (EHEC) O157:H7 SlrP lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-GS2-SopA	GS2 fused to *S. typhimurium* SopA lacking β-helix domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-GS2-SspH1	GS2 fused to *Salmonella typhimurium* SspH1 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-GS2-SspH2	GS2 fused to *S. typhimurium* SspH2 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-GS2-XopL	GS2 fused to *Xanthomonas campestris* XopL lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA-βTrCP-GS2	GS2 fused to full length *H. sapiens* βTrCP with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-GS2-CHIPdTPR	GS2 cloned in place of scFvR4 in pcDNA3-R4-CHIPΔTPR; CMV promoter; Kozak; Amp^R	This study
pcDNA3-Slmb-GS2	GS2 cloned in place of vhhGFP4 in pcDNA3-NSlmb-vhhGFP4; CMV promoter; Kozak; Amp^R	This study
pcDNA3-GS2-SPOP	GS2 fused to *Homo sapiens* SPOP F-box with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-VHL-GS2	GS2 fused to *H. sapiens* VHL F-box with N-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
Ubiquibodies with disease-relevant specificity		
pcDNA3-AS15-IpaH9.8	AS15 fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-NSa1-IpaH9.8	NSa1 fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-NSa5-IpaH9.8	NSa5 fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-CS3-IpaH9.8	CS3 fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-NSa5-IpaH9.8^{C337A}	NSa5 fused to IpaH9.8^{C337A} lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study

pcDNA3-RasInI-IpaH9.8	RasInII fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-RasInII-IpaH9.8^{C337A}	RasInII fused to IpaH9.8^{C337A} lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-RasInII-IpaH9.8	RasInII fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-E40-IpaH9.8	E40 fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-pE59-IpaH9.8	pE59 fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-EpE82-IpaH9.8	EpE82 fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-EpE89-IpaH9.8	EpE89 fused to IpaH9.8 lacking LRR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-E40-CHIP	E40 fused to CHIP lacking TPR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-pE59-CHIP	pE59 fused to CHIP lacking TPR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-EpE82-CHIP	EpE82 fused to CHIP lacking TPR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
pcDNA3-EpE89-CHIP	EpE89 fused to CHIP lacking TPR domain with C-terminal Flag and 6x-His tags cloned in pcDNA3; CMV promoter; Kozak; Amp^R	This study
Inducible Ubiquibodies		
pBabe-CMV-puro-Tet-mVenus-Lov	AsLOV2 fused to the C-terminus of mVenus for mammalian expression; TetCMV promoter; Amp^R	Addgene (#22033)
CRY2(535,L348F)-mCherry	CRY2 (585 amino acid truncation) fused to N-terminus of mCherry for mammalian expression; CMV promoter; Kan^R	Addgene (#75372)
pmCIB1(dNLS)-pmGFP	CIB1 lacking NLS fused to the N-terminus of pmGFP; CMV promoter; Kan^R	Addgene (#28240)
pCherry-FRB	FRB fused to C-terminus of mCherry for mammalian expression; CMV promoter; Kan^R	Addgene (#25920)
CFP-FKBP	FKBP fused to the C-terminus of CFP for mammalian expression; CMV promoter; Kan^R	Addgene (#20160)
pTriEx-3-GS2-AsLOV2 -R1-IpaH9.8	GS2-IpaH9.8 ubiquibody with AsLOV2 insertion in Register 1 position; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-GS2-AsLOV2 –R2-IpaH9.8	GS2-IpaH9.8 ubiquibody with AsLOV2 insertion in Register 2 position; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-GS2-AsLOV2 –R3-	GS2-IpaH9.8 ubiquibody with AsLOV2 insertion in Register	This study

IpaH9.8	3 position; CMV promoter; T7 promoter; Amp^R	
pTriEx-3-GS2-AsLOV2 –R4-IpaH9.8	GS2-IpaH9.8 ubiquibody with AsLOV2 insertion in Register 4 position; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-GS2-AsLOV2 –R5-IpaH9.8	GS2-IpaH9.8 ubiquibody with AsLOV2 insertion in Register 5 position; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-GS2-AsLOV2 –R6-IpaH9.8	GS2-IpaH9.8 ubiquibody with AsLOV2 insertion in Register 6 position; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-GS2-AsLOV2 –R7-IpaH9.8	GS2-IpaH9.8 ubiquibody with AsLOV2 insertion in Register 7 position; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-GS2-AsLOV2 –R8-IpaH9.8	GS2-IpaH9.8 ubiquibody with AsLOV2 insertion in Register 8 position; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-GS2-AsLOV2 –R9-IpaH9.8	GS2-IpaH9.8 ubiquibody with AsLOV2 insertion in Register 9 position; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-GS2-AsLOV2 –R10-IpaH9.8	GS2-IpaH9.8 ubiquibody with AsLOV2 insertion in Register 10 position; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-CRY2-IpaH9.8dlRR	IpaH9.8dLRR expressed with C-terminal Flag and 6xHIS tags as a C-terminal fusion to CRY2; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-CRY2-IpaH9.8dlRR^{C337A}	IpaH9.8dLRR^{C337A} expressed with C-terminal Flag and 6xHIS tags as a C-terminal fusion to CRY2; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-CRY2-CHIPdTPR	CHIPdTPR expressed with C-terminal Flag and 6xHIS tags as a C-terminal fusion to CRY2; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-CRY2-GS2	GS2 with C-terminal Flag and 6xHIS tags expressed as a C-terminal fusion to CRY2; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-CIB1-IpaH9.8dlRR	IpaH9.8dLRR expressed with C-terminal Flag and 6xHIS tags as a C-terminal fusion to CIB1; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-FRB-IpaH9.8dLRR	IpaH9.8dLRR expressed with C-terminal Flag and 6xHIS tags as a C-terminal fusion to FRB; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-FRB-IpaH9.8dlRR^{C337A}	IpaH9.8dLRR^{C337A} expressed with C-terminal Flag and 6xHIS tags as a C-terminal fusion to FRB; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-FRB-GS2	GS2 with C-terminal Flag and 6xHIS tags expressed as a C-terminal fusion to FRB; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-FKBP-IpaH9.8dLRR	IpaH9.8dLRR expressed with C-terminal Flag and 6xHIS tags as a C-terminal fusion to FKBP; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-FKBP-IpaH9.8dlRR^{C337A}	IpaH9.8dLRR^{C337A} expressed with C-terminal Flag and 6xHIS tags as a C-terminal fusion to FKBP; CMV promoter; T7 promoter; Amp^R	This study
pTriEx-3-FKBP-GS2	GS2 with C-terminal Flag and 6xHIS tags expressed as a C-terminal fusion to FKBP; CMV promoter; T7 promoter; Amp^R	This study
Yeast Surface Display		
pCT-Con-G4	G4 FN3 Library expressed as AGA2 fusion; Gal1-10	Laboratory stock.

	promoter; Trp(+); Amp^R	
pCT-FN3-Gene	pCT-Con variant used for homologous recombination cloning containing HA and c-myc tags flanking PstI, NdeI, and BamHI sites; Trp(+); Amp^R	Laboratory stock
pCT-FN3-Loop	pCT-Con variant used for homologous recombination cloning containing FN3 gene, with DNA between BC and FG loop replaced by NcoI, SmaI, and NdeI sites; Trp(+); Amp^R	Laboratory stock

Table C.1 Strains, cell lines, and plasmids used in this study.

REFERENCES

1. Hochstrasser, M., *Ubiquitin, proteasomes, and the regulation of intracellular protein degradation*. Curr Opin Cell Biol, 1995. **7**(2): p. 215-23.
2. Balchin, D., M. Hayer-Hartl, and F.U. Hartl, *In vivo aspects of protein folding and quality control*. Science, 2016. **353**(6294): p. aac4354.
3. Ciechanover, A., *The unravelling of the ubiquitin system*. Nat Rev Mol Cell Biol, 2015. **16**(5): p. 322-4.
4. Hershko, A. and A. Ciechanover, *The ubiquitin system*. Annu Rev Biochem, 1998. **67**: p. 425-79.
5. Clague, M.J. and S. Urbe, *Ubiquitin: same molecule, different degradation pathways*. Cell, 2010. **143**(5): p. 682-5.
6. Komander, D. and M. Rape, *The ubiquitin code*. Annu Rev Biochem, 2012. **81**: p. 203-29.
7. Yau, R. and M. Rape, *The increasing complexity of the ubiquitin code*. Nat Cell Biol, 2016. **18**(6): p. 579-86.
8. Chen, Z.J., *Ubiquitin signalling in the NF-kappaB pathway*. Nat Cell Biol, 2005. **7**(8): p. 758-65.
9. Sun, L. and Z.J. Chen, *The novel functions of ubiquitination in signaling*. Curr Opin Cell Biol, 2004. **16**(2): p. 119-26.
10. Sander, J.D. and J.K. Joung, *CRISPR-Cas systems for editing, regulating and targeting genomes*. Nat Biotechnol, 2014. **32**(4): p. 347-355.
11. Gaj, T., et al., *Targeted gene knockout by direct delivery of zinc-finger nuclease proteins*. Nat Methods, 2012. **9**(8): p. 805-+.
12. Sun, N. and H. Zhao, *Transcription activator-like effector nucleases (TALENs): a highly efficient and versatile tool for genome editing*. Biotechnol Bioeng, 2013. **110**(7): p. 1811-21.
13. Fire, A., et al., *Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans*. Nature, 1998. **391**(6669): p. 806-11.
14. Rinaldi, C. and M.J.A. Wood, *Antisense oligonucleotides: the next frontier for treatment of neurological disorders*. Nat Rev Neurol, 2018. **14**(1): p. 9-21.
15. Cox, D.B., R.J. Platt, and F. Zhang, *Therapeutic genome editing: prospects and challenges*. Nat Med, 2015. **21**(2): p. 121-31.
16. Soutschek, J., et al., *Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs*. Nature, 2004. **432**(7014): p. 173-8.
17. Bumcrot, D., et al., *RNAi therapeutics: a potential new class of pharmaceutical drugs*. Nat Chem Biol, 2006. **2**(12): p. 711-9.
18. Wang, M., Z.A. Glass, and Q. Xu, *Non-viral delivery of genome-editing nucleases for gene therapy*. Gene Ther, 2017. **24**(3): p. 144-150.
19. Gaj, T., C.A. Gersbach, and C.F. Barbas, 3rd, *ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering*. Trends Biotechnol, 2013. **31**(7): p. 397-405.

20. Hebert, C.G., J.J. Valdes, and W.E. Bentley, *Beyond silencing - engineering applications of RNA interference and antisense technology for altering cellular phenotype*. Curr Opin Biotechnol, 2008. **19**(5): p. 500-505.
21. Beckwith, R.E.J., *Small-molecule-mediated Targeted Protein Degradation for Drug Discovery*. High Throughput Screening Methods: Evolution and Refinement, 2017. **1**: p. 252-274.
22. Blakemore, D.C., et al., *Organic synthesis provides opportunities to transform drug discovery*. Nat Chem, 2018. **10**(4): p. 383-394.
23. Crews, C.M., *Targeting the undruggable proteome: the small molecules of my dreams*. Chem Biol, 2010. **17**(6): p. 551-5.
24. Arkin, M.R. and J.A. Wells, *Small-molecule inhibitors of protein-protein interactions: progressing towards the dream*. Nat Rev Drug Discov, 2004. **3**(4): p. 301-17.
25. Hopkins, A.L. and C.R. Groom, *The druggable genome*. Nat Rev Drug Discov, 2002. **1**(9): p. 727-30.
26. Colwill, K., G. Renewable Protein Binder Working, and S. Graslund, *A roadmap to generate renewable protein binders to the human proteome*. Nat Methods, 2011. **8**(7): p. 551-8.
27. Jurado, P., et al., *Production of functional single-chain Fv antibodies in the cytoplasm of Escherichia coli*. J Mol Biol, 2002. **320**(1): p. 1-10.
28. Martineau, P., P. Jones, and G. Winter, *Expression of an antibody fragment at high levels in the bacterial cytoplasm*. J Mol Biol, 1998. **280**(1): p. 117-27.
29. Proba, K., et al., *Antibody scFv fragments without disulfide bonds made by molecular evolution*. J Mol Biol, 1998. **275**(2): p. 245-53.
30. Fisher, A.C. and M.P. DeLisa, *Efficient isolation of soluble intracellular single-chain antibodies using the twin-arginine translocation machinery*. J Mol Biol, 2009. **385**(1): p. 299-311.
31. Koide, A., et al., *Probing protein conformational changes in living cells by using designer binding proteins: application to the estrogen receptor*. Proc Natl Acad Sci U S A, 2002. **99**(3): p. 1253-8.
32. Koide, A., et al., *The fibronectin type III domain as a scaffold for novel binding proteins*. J Mol Biol, 1998. **284**(4): p. 1141-51.
33. Kawe, M., et al., *Isolation of intracellular proteinase inhibitors derived from designed ankyrin repeat proteins by genetic screening*. J Biol Chem, 2006. **281**(52): p. 40252-63.
34. Parizek, P., et al., *Designed ankyrin repeat proteins (DARPins) as novel isoform-specific intracellular inhibitors of c-Jun N-terminal kinases*. ACS Chem Biol, 2012. **7**(8): p. 1356-66.
35. Deshaies, R.J., *Protein degradation: Prime time for PROTACs*. Nat Chem Biol, 2015. **11**(9): p. 634-5.
36. Neklesa, T.K., J.D. Winkler, and C.M. Crews, *Targeted protein degradation by PROTACs*. Pharmacol Ther, 2017. **174**: p. 138-144.
37. Lai, A.C. and C.M. Crews, *Induced protein degradation: an emerging drug discovery paradigm*. Nat Rev Drug Discov, 2017. **16**(2): p. 101-114.

38. Sakamoto, K.M., et al., *Protacs: chimeric molecules that target proteins to the Skp1-Cullin-F box complex for ubiquitination and degradation*. Proc Natl Acad Sci U S A, 2001. **98**(15): p. 8554-9.
39. Sakamoto, K.M., et al., *Development of Protacs to target cancer-promoting proteins for ubiquitination and degradation*. Mol Cell Proteomics, 2003. **2**(12): p. 1350-8.
40. Hines, J., et al., *Posttranslational protein knockdown coupled to receptor tyrosine kinase activation with phosphoPROTACs*. Proc Natl Acad Sci U S A, 2013. **110**(22): p. 8942-7.
41. Schneekloth, A.R., et al., *Targeted intracellular protein degradation induced by a small molecule: En route to chemical proteomics*. Bioorg Med Chem Lett, 2008. **18**(22): p. 5904-8.
42. Sekine, K., et al., *Small molecules destabilize cIAP1 by activating auto-ubiquitylation*. Journal of Biological Chemistry, 2008. **283**(14): p. 8961-8968.
43. Bondeson, D.P., et al., *Catalytic in vivo protein knockdown by small-molecule PROTACs*. Nat Chem Biol, 2015. **11**(8): p. 611-7.
44. Lu, J., et al., *Hijacking the E3 ubiquitin ligase cereblon to efficiently target BRD4*. Chem Biol, 2015. **22**(6): p. 755-63.
45. Winter, G.E., et al., *DRUG DEVELOPMENT. Phthalimide conjugation as a strategy for in vivo target protein degradation*. Science, 2015. **348**(6241): p. 1376-81.
46. Mullard, A., *First targeted protein degrader hits the clinic*. Nat Rev Drug Discov, 2019.
47. Gadd, M.S., et al., *Structural basis of PROTAC cooperative recognition for selective protein degradation*. Nat Chem Biol, 2017. **13**(5): p. 514-521.
48. Nabet, B., et al., *The dTAG system for immediate and target-specific protein degradation*. Nat Chem Biol, 2018. **14**(5): p. 431-441.
49. Buckley, D.L., et al., *HaloPROTACS: use of small molecule PROTACs to induce degradation of HaloTag fusion proteins*. ACS Chem Biol, 2015. **10**(8): p. 1831-7.
50. Osherovich, L., *Degradation from within*. Science-Business Exchange, 2014. **7**: p. 10-11.
51. Zhou, P., et al., *Harnessing the ubiquitination machinery to target the degradation of specific cellular proteins*. Mol Cell, 2000. **6**(3): p. 751-6.
52. Zhang, J., N. Zheng, and P. Zhou, *Exploring the functional complexity of cellular proteins by protein knockout*. Proc Natl Acad Sci U S A, 2003. **100**(24): p. 14127-32.
53. Hatakeyama, S., et al., *Targeted destruction of c-Myc by an engineered ubiquitin ligase suppresses cell transformation and tumor formation*. Cancer Res, 2005. **65**(17): p. 7874-9.
54. Kong, F., et al., *Engineering a single ubiquitin ligase for the selective degradation of all activated ErbB receptor tyrosine kinases*. Oncogene, 2014. **33**(8): p. 986-95.
55. Sufan, R.I., et al., *Oxygen-independent degradation of HIF-alpha via bioengineered VHL tumour suppressor complex*. EMBO Mol Med, 2009. **1**(1): p. 66-78.

56. Ma, Y., et al., *Targeted degradation of KRAS by an engineered ubiquitin ligase suppresses pancreatic cancer cell growth in vitro and in vivo*. Mol Cancer Ther, 2013. **12**(3): p. 286-94.
57. Portnoff, A.D., et al., *Ubiquibodies, synthetic E3 ubiquitin ligases endowed with unnatural substrate specificity for targeted protein silencing*. J Biol Chem, 2014. **289**(11): p. 7844-55.
58. Baltz, M.R., E.A. Stephens, and M.P. DeLisa, *Design and functional characterization of synthetic E3 ubiquitin ligases for targeted protein depletion*. Curr Protoc Chem Biol, 2018. **10**(1): p. 72-90.
59. Caussinus, E., O. Kanca, and M. Affolter, *Fluorescent fusion protein knockout mediated by anti-GFP nanobody*. Nat Struct Mol Biol, 2011. **19**(1): p. 117-21.
60. Fulcher, L.J., et al., *An affinity-directed protein missile system for targeted proteolysis*. Open Biol, 2016. **6**(10).
61. Fulcher, L.J., et al., *Targeting endogenous proteins for degradation through the affinity-directed protein missile system*. Open Biol, 2017. **7**(5).
62. Schmit, N.E., K. Neopane, and O. Hantschel, *Targeted protein degradation through cytosolic delivery of monobody binders using bacterial toxins*. ACS Chem Biol, 2019. **14**(5): p. 916-924.
63. Ludwicki, M.B., et al., *Broad-spectrum proteome editing with an engineered bacterial ubiquitin ligase mimic*. ACS Cent Sci, 2019. **5**(5): p. 852-866.
64. Kuo, C.L., G.A. Oyler, and C.B. Shoemaker, *Accelerated neuronal cell recovery from Botulinum neurotoxin intoxication by targeted ubiquitination*. PLoS One, 2011. **6**(5): p. e20352.
65. Shin, Y.J., et al., *Nanobody-targeted E3-ubiquitin ligase complex degrades nuclear proteins*. Sci Rep, 2015. **5**: p. 14269.
66. Kanner, S.A., T. Morgenstern, and H.M. Colecraft, *Sculpting ion channel functional expression with engineered ubiquitin ligases*. Elife, 2017. **6**.
67. Sidhu, S.S., *Phage display in pharmaceutical biotechnology*. Curr Opin Biotechnol, 2000. **11**(6): p. 610-6.
68. Amstutz, P., et al., *In vitro display technologies: novel developments and applications*. Curr Opin Biotechnol, 2001. **12**(4): p. 400-5.
69. Gai, S.A. and K.D. Wittrup, *Yeast surface display for protein engineering and characterization*. Curr Opin Struct Biol, 2007. **17**(4): p. 467-73.
70. Pearce, M.J., et al., *Ubiquitin-like protein involved in the proteasome pathway of Mycobacterium tuberculosis*. Science, 2008. **322**(5904): p. 1104-7.
71. Maculins, T., et al., *Bacteria-host relationship: ubiquitin ligases as weapons of invasion*. Cell Res, 2016. **26**(4): p. 499-510.
72. Ashida, H. and C. Sasakawa, *Bacterial E3 ligase effectors exploit host ubiquitin systems*. Curr Opin Microbiol, 2017. **35**: p. 16-22.
73. Ashida, H., M. Kim, and C. Sasakawa, *Exploitation of the host ubiquitin system by human bacterial pathogens*. Nat Rev Microbiol, 2014. **12**(6): p. 399-413.
74. Rohde, J.R., et al., *Type III secretion effectors of the IpaH family are E3 ubiquitin ligases*. Cell Host Microbe, 2007. **1**(1): p. 77-83.

75. Hsu, F., et al., *The Legionella effector SidC defines a unique family of ubiquitin ligases important for bacterial phagosomal remodeling*. Proc Natl Acad Sci U S A, 2014. **111**(29): p. 10538-43.
76. Angot, A., et al., *Exploitation of eukaryotic ubiquitin signaling pathways by effectors translocated by bacterial type III and type IV secretion systems*. PLoS Pathog, 2007. **3**(1): p. e3.
77. Norkowski, S., et al., *Bacterial LPX motif-harboring virulence factors constitute a species-spanning family of cell-penetrating effectors*. Cell Mol Life Sci, 2018. **75**(12): p. 2273-2289.
78. Singer, A.U., et al., *Structure of the Shigella T3SS effector IpaH defines a new class of E3 ubiquitin ligases*. Nat Struct Mol Biol, 2008. **15**(12): p. 1293-301.
79. Zhu, Y., et al., *Structure of a Shigella effector reveals a new class of ubiquitin ligases*. Nat Struct Mol Biol, 2008. **15**(12): p. 1302-8.
80. Kim, D.W., et al., *The Shigella flexneri effector OspG interferes with innate immune responses by targeting ubiquitin-conjugating enzymes*. Proc Natl Acad Sci U S A, 2005. **102**(39): p. 14046-51.
81. Sanada, T., et al., *The Shigella flexneri effector OspI deamidates UBC13 to dampen the inflammatory response*. Nature, 2012. **483**(7391): p. 623-6.
82. Randow, F. and P.J. Lehner, *Viral avoidance and exploitation of the ubiquitin system*. Nat Cell Biol, 2009. **11**(5): p. 527-34.
83. Viswanathan, K., K. Fruh, and V. DeFilippis, *Viral hijacking of the host ubiquitin system to evade interferon responses*. Curr Opin Microbiol, 2010. **13**(4): p. 517-23.
84. Lorenzo, M.E., J.U. Jung, and H.L. Ploegh, *Kaposi's sarcoma-associated herpesvirus K3 utilizes the ubiquitin-proteasome system in routing class major histocompatibility complexes to late endocytic compartments*. J Virol, 2002. **76**(11): p. 5522-31.
85. Chelbi-Alix, M.K. and H. de The, *Herpes virus induced proteasome-dependent degradation of the nuclear bodies-associated PML and Sp100 proteins*. Oncogene, 1999. **18**(4): p. 935-41.
86. Scheffner, M., et al., *The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53*. Cell, 1990. **63**(6): p. 1129-36.
87. Fischer, R., et al., *Break on through to the Other Side—Biophysics and Cell Biology Shed Light on Cell-Penetrating Peptides*. ChemBioChem, 2005. **6**(12): p. 2126-2142.
88. Ruter, C., et al., *A newly identified bacterial cell-penetrating peptide that reduces the transcription of pro-inflammatory cytokines*. J Cell Sci, 2010. **123**(Pt 13): p. 2190-8.
89. Lubos, M.-L., et al., *Analysis of T3SS-independent Autonomous Internalization of the Bacterial Effector Protein SspH1 from Salmonella typhimurium*. Vol. 1. 2014.
90. Norkowski, S., M.A. Schmidt, and C. Ruter, *The species-spanning family of LPX-motif harbouring effector proteins*. Cell Microbiol, 2018. **20**(11): p. e12945.
91. Lin, Y.H. and M.P. Machner, *Exploitation of the host cell ubiquitin machinery by microbial effector proteins*. J Cell Sci, 2017. **130**(12): p. 1985-1996.
92. Singer, A.U., et al., *A pathogen type III effector with a novel E3 ubiquitin ligase architecture*. PLoS Pathog, 2013. **9**(1): p. e1003121.

93. Ashida, H., et al., *A bacterial E3 ubiquitin ligase IpaH9.8 targets NEMO/IKKgamma to dampen the host NF-kappaB-mediated inflammatory response*. Nat Cell Biol, 2010. **12**(1): p. 66-73; sup pp 1-9.
94. Li, J., et al., *Polyamine-Mediated Stoichiometric Assembly of Ribonucleoproteins for Enhanced mRNA Delivery*. Angewandte Chemie (International ed. in English), 2017. **56**(44): p. 13709-13712.
95. Koide, A., et al., *Teaching an old scaffold new tricks: monobodies constructed using alternative surfaces of the FN3 scaffold*. J Mol Biol, 2012. **415**(2): p. 393-405.
96. Koide, A., et al., *High-affinity single-domain binding proteins with a binary-code interface*. Proc Natl Acad Sci U S A, 2007. **104**(16): p. 6632-7.
97. Tsien, R.Y., *The green fluorescent protein*. Annu Rev Biochem, 1998. **67**: p. 509-44.
98. Shaner, N.C., P.A. Steinbach, and R.Y. Tsien, *A guide to choosing fluorescent proteins*. Nat Methods, 2005. **2**(12): p. 905-9.
99. Khmelinskii, A., et al., *Incomplete proteasomal degradation of green fluorescent proteins in the context of tandem fluorescent protein timers*. Mol Biol Cell, 2016. **27**(2): p. 360-70.
100. Saerens, D., et al., *Identification of a universal VHH framework to graft non-canonical antigen-binding loops of camel single-domain antibodies*. J Mol Biol, 2005. **352**(3): p. 597-607.
101. Buetow, L. and D.T. Huang, *Structural insights into the catalysis and regulation of E3 ubiquitin ligases*. Nat Rev Mol Cell Biol, 2016. **17**(10): p. 626-42.
102. Metzger, M.B., V.A. Hristova, and A.M. Weissman, *HECT and RING finger families of E3 ubiquitin ligases at a glance*. J Cell Sci, 2012. **125**(Pt 3): p. 531-7.
103. Adams, D., et al., *Patisiran, an RNAi Therapeutic, for Hereditary Transthyretin Amyloidosis*. N Engl J Med, 2018. **379**(1): p. 11-21.
104. Deleavey, G.F. and M.J. Damha, *Designing chemically modified oligonucleotides for targeted gene silencing*. Chem Biol, 2012. **19**(8): p. 937-54.
105. Fu, Y., et al., *High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells*. Nat Biotechnol, 2013. **31**(9): p. 822-6.
106. Fedorov, Y., et al., *Off-target effects by siRNA can induce toxic phenotype*. RNA, 2006. **12**(7): p. 1188-96.
107. Sha, F., et al., *Monobodies and other synthetic binding proteins for expanding protein science*. Protein Sci, 2017. **26**(5): p. 910-924.
108. Richter, A., E. Eggenstein, and A. Skerra, *Anticalins: Exploiting a non-Ig scaffold with hypervariable loops for the engineering of binding proteins*. FEBS Letters, 2014. **588**(2): p. 213-218.
109. Feldwisch, J. and V. Tolmachev, *Engineering of affibody molecules for therapy and diagnostics*. Methods Mol Biol, 2012. **899**: p. 103-26.
110. Pluckthun, A., *Designed ankyrin repeat proteins (DARPins): binding proteins for research, diagnostics, and therapy*. Annu Rev Pharmacol Toxicol, 2015. **55**: p. 489-511.
111. Muyldermans, S., *Nanobodies: natural single-domain antibodies*. Annu Rev Biochem, 2013. **82**: p. 775-97.

112. Sha, F., et al., *Dissection of the BCR-ABL signaling network using highly specific monobody inhibitors to the SHP2 SH2 domains*. Proc Natl Acad Sci U S A, 2013. **110**(37): p. 14924-9.
113. Cetin, M., et al., *RasIns: Genetically Encoded Intrabodies of Activated Ras Proteins*. Journal of molecular biology, 2017. **429**(4): p. 562-573.
114. Kauke, M.J., et al., *An engineered protein antagonist of K-Ras/B-Raf interaction*. Sci Rep, 2017. **7**(1): p. 5831.
115. Stephens, E., *Anatomy of a Ubiquibody*. 2018, Cornell University: eCommons.
116. Chen, T.F., et al., *Engineering fibronectin-based binding proteins by yeast surface display*. Methods Enzymol, 2013. **523**: p. 303-26.
117. Ashida, H. and C. Sasakawa, *Shigella IpaH family effectors as a versatile model for studying pathogenic bacteria*. Front Cell Infect Microbiol, 2015. **5**: p. 100.
118. Buckley, D.L. and C.M. Crews, *Small-molecule control of intracellular protein levels through modulation of the ubiquitin proteasome system*. Angew Chem Int Ed Engl, 2014. **53**(9): p. 2312-30.
119. Schaefer, B.C., et al., *Observation of antigen-dependent CD8+ T-cell/ dendritic cell interactions in vivo*. Cell Immunol, 2001. **214**(2): p. 110-22.
120. Stewart, K.M., K.L. Horton, and S.O. Kelley, *Cell-penetrating peptides as delivery vehicles for biology and medicine*. Org Biomol Chem, 2008. **6**(13): p. 2242-55.
121. Scharnert, J., et al., *Autonomous translocation and intracellular trafficking of the cell-penetrating and immune-suppressive effector protein YopM*. Cell Mol Life Sci, 2013. **70**(24): p. 4809-23.
122. Guan, S. and J. Rosenecker, *Nanotechnologies in delivery of mRNA therapeutics using nonviral vector-based delivery systems*. Gene Ther, 2017. **24**(3): p. 133-143.
123. Uchida, H., et al., *Modulated protonation of side chain aminoethylene repeats in N-substituted polyaspartamides promotes mRNA transfection*. J Am Chem Soc, 2014. **136**(35): p. 12396-405.
124. Moglich, A. and K. Moffat, *Engineered photoreceptors as novel optogenetic tools*. Photochem Photobiol Sci, 2010. **9**(10): p. 1286-300.
125. Taylor, B.L. and I.B. Zhulin, *PAS domains: internal sensors of oxygen, redox potential, and light*. Microbiol Mol Biol Rev, 1999. **63**(2): p. 479-506.
126. Moglich, A., R.A. Ayers, and K. Moffat, *Structure and signaling mechanism of Per-ARNT-Sim domains*. Structure, 2009. **17**(10): p. 1282-94.
127. Harper, S.M., L.C. Neil, and K.H. Gardner, *Structural basis of a phototropin light switch*. Science, 2003. **301**(5639): p. 1541-4.
128. Halavaty, A.S. and K. Moffat, *N- and C-terminal flanking regions modulate light-induced signal transduction in the LOV2 domain of the blue light sensor phototropin 1 from Avena sativa*. Biochemistry, 2007. **46**(49): p. 14001-9.
129. Zayner, J.P., C. Antoniou, and T.R. Sosnick, *The amino-terminal helix modulates light-activated conformational changes in AsLOV2*. J Mol Biol, 2012. **419**(1-2): p. 61-74.
130. Bonger, K.M., et al., *General method for regulating protein stability with light*. ACS Chem Biol, 2014. **9**(1): p. 111-5.
131. Renicke, C., et al., *A LOV2 domain-based optogenetic tool to control protein degradation and cellular function*. Chem Biol, 2013. **20**(4): p. 619-26.

132. Shimizu-Sato, S., et al., *A light-switchable gene promoter system*. Nat Biotechnol, 2002. **20**(10): p. 1041-4.
133. Levskaya, A., et al., *Spatiotemporal control of cell signalling using a light-switchable protein interaction*. Nature, 2009. **461**(7266): p. 997-1001.
134. Yazawa, M., et al., *Induction of protein-protein interactions in live cells using light*. Nat Biotechnol, 2009. **27**(10): p. 941-5.
135. Tyszkiewicz, A.B. and T.W. Muir, *Activation of protein splicing with light in yeast*. Nat Methods, 2008. **5**(4): p. 303-5.
136. Taslimi, A., et al., *Optimized second-generation CRY2-CIB dimerizers and photoactivatable Cre recombinase*. Nat Chem Biol, 2016. **12**(6): p. 425-30.
137. Kennedy, M.J., et al., *Rapid blue-light-mediated induction of protein interactions in living cells*. Nat Methods, 2010. **7**(12): p. 973-5.
138. O'Neill, P.R. and N. Gautam, *Subcellular optogenetic inhibition of G proteins generates signaling gradients and cell migration*. Mol Biol Cell, 2014. **25**(15): p. 2305-14.
139. Lungu, O.I., et al., *Designing photoswitchable peptides using the AsLOV2 domain*. Chem Biol, 2012. **19**(4): p. 507-17.
140. Edwards, D.J., et al., *Convergent evolution in the assembly of polyubiquitin degradation signals by the Shigella flexneri IpaH9.8 ligase*. J Biol Chem, 2014. **289**(49): p. 34114-28.
141. Kawano, F., et al., *Fluorescence imaging-based high-throughput screening of fast- and slow-cycling LOV proteins*. PLoS One, 2013. **8**(12): p. e82693.
142. Portnoff, A.D., *Ubiquibodies: Engineered Ubiquitin Ligases with Unnatural Substrate Specificity for Targeted Protein Silencing*. 2014, Cornell University.
143. Öztürk-Atar, K., H. Eroğlu, and S. Çalış, *Novel advances in targeted drug delivery*. Journal of Drug Targeting, 2018. **26**(8): p. 633-642.
144. Linsley, C.S. and B.M. Wu, *Recent advances in light-responsive on-demand drug-delivery systems*. Ther Deliv, 2017. **8**(2): p. 89-107.
145. Seyedarabi, A., et al., *A disulfide driven domain swap switches off the activity of Shigella IpaH9.8 E3 ligase*. Vol. 584. 2010. 4163-8.
146. Tucker, C.L., J.D. Vrana, and M.J. Kennedy, *Tools for Controlling Protein Interactions Using Light*. Current Protocols in Cell Biology, 2014. **64**(1): p. 17.16.1-17.16.20.
147. Abramovitch, R.B., et al., *Type III effector AvrPtoB requires intrinsic E3 ubiquitin ligase activity to suppress plant cell death and immunity*. Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(8): p. 2851-2856.
148. Zelphati, O., et al., *Intracellular delivery of proteins with a new lipid-mediated delivery system*. J Biol Chem, 2001. **276**(37): p. 35103-10.
149. Erazo-Oliveras, A., et al., *Protein delivery into live cells by incubation with an endosomolytic agent*. Nat Methods, 2014. **11**(8): p. 861-7.
150. Gabel, C.A. and S.A. Foster, *Mannose 6-phosphate receptor-mediated endocytosis of acid hydrolases: internalization of beta-glucuronidase is accompanied by a limited dephosphorylation*. J Cell Biol, 1986. **103**(5): p. 1817-27.

151. Mix, K.A., J.E. Lomax, and R.T. Raines, *Cytosolic delivery of proteins by bioreversible esterification*. J Am Chem Soc, 2017. **139**(41): p. 14396-14398.
152. Hasadsri, L., et al., *Functional protein delivery into neurons using polymeric nanoparticles*. J Biol Chem, 2009. **284**(11): p. 6972-81.
153. Nicholes, N., et al., *Modular protein switches derived from antibody mimetic proteins*. Protein Eng Des Sel, 2016. **29**(2): p. 77-85.
154. Starita, L.M., et al., *Activity-enhancing mutations in an E3 ubiquitin ligase identified by high-throughput mutagenesis*. Proc Natl Acad Sci U S A, 2013. **110**(14): p. E1263-72.
155. Cortajarena, A.L., et al., *Designed proteins to modulate cellular networks*. ACS Chem Biol, 2010. **5**(6): p. 545-52.
156. Tasumi, S., et al., *High-affinity lamprey VLRA and VLRB monoclonal antibodies*. Proc Natl Acad Sci U S A, 2009. **106**(31): p. 12891-6.
157. Janjusevic, R., et al., *A bacterial inhibitor of host programmed cell death defenses is an E3 ubiquitin ligase*. Science, 2006. **311**(5758): p. 222-6.
158. Ashida, H., et al., *Shigella chromosomal IpaH proteins are secreted via the type III secretion system and act as effectors*. Mol Microbiol, 2007. **63**(3): p. 680-93.
159. Okuda, J., et al., *Shigella effector IpaH9.8 binds to a splicing factor U2AF(35) to modulate host immune responses*. Biochem Biophys Res Commun, 2005. **333**(2): p. 531-9.
160. Ensminger, A.W. and R.R. Isberg, *E3 ubiquitin ligase activity and targeting of BAT3 by multiple Legionella pneumophila translocated substrates*. Infect Immun, 2010. **78**(9): p. 3905-19.
161. Quaile, A.T., et al., *Molecular Characterization of LubX: Functional Divergence of the U-Box Fold by Legionella pneumophila*. Structure, 2015. **23**(8): p. 1459-1469.
162. Wu, B., et al., *NleG Type 3 effectors from enterohaemorrhagic Escherichia coli are U-Box E3 ubiquitin ligases*. PLoS Pathog, 2010. **6**(6): p. e1000960.
163. Lin, D.Y., et al., *Biochemical and structural studies of a HECT-like ubiquitin ligase from Escherichia coli O157:H7*. J Biol Chem, 2011. **286**(1): p. 441-9.
164. Zouhir, S., et al., *The structure of the Slrp-Trx1 complex sheds light on the autoinhibition mechanism of the type III secretion system effectors of the NEL family*. Biochem J, 2014. **464**(1): p. 135-44.
165. Diao, J., et al., *Crystal structure of SopA, a Salmonella effector protein mimicking a eukaryotic ubiquitin ligase*. Nat Struct Mol Biol, 2008. **15**(1): p. 65-70.
166. Keszei, A.F., et al., *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, 2014. **34**(3): p. 362-73.
167. Bhavsar, A.P., et al., *The Salmonella type III effector SspH2 specifically exploits the NLR co-chaperone activity of SGT1 to subvert immunity*. PLoS Pathog, 2013. **9**(7): p. e1003518.
168. Maniaci, C., et al., *Homo-PROTACs: bivalent small-molecule dimerizers of the VHL E3 ubiquitin ligase to induce self-degradation*. Nat Commun, 2017. **8**(1): p. 830.
169. Yusa, K., et al., *A hyperactive piggyBac transposase for mammalian applications*. Proc Natl Acad Sci U S A, 2011. **108**(4): p. 1531-6.

170. Li, J., et al., *Structurally modulated codelivery of siRNA and Argonaute 2 for enhanced RNA interference*. Proc Natl Acad Sci U S A, 2018. **115**(12): p. E2696-e2705.
171. Boczkowski, D., et al., *Induction of tumor immunity and cytotoxic T lymphocyte responses using dendritic cells transfected with messenger RNA amplified from tumor cells*. Cancer Res, 2000. **60**(4): p. 1028-34.
172. Khokhlatchev, A., et al., *Reconstitution of mitogen-activated protein kinase phosphorylation cascades in bacteria. Efficient synthesis of active protein kinases*. J Biol Chem, 1997. **272**(17): p. 11057-62.
173. Butler, D.C., et al., *Bifunctional Anti-Non-Amyloid Component α-Synuclein Nanobodies Are Protective In Situ*. PLoS One, 2016. **11**(11): p. e0165964.
174. Yasuda, R., et al., *Supersensitive Ras activation in dendrites and spines revealed by two-photon fluorescence lifetime imaging*. Nat Neurosci, 2006. **9**(2): p. 283-91.