THE ANATOMY OF A UBIQUIBODY

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

Erin Ashley Stephens

May 2018

THE ANATOMY OF A UBIQUIBODY

Erin Ashley Stephens, Ph.D.

Cornell University, 2018

Protein silencing is an important aspect of both therapeutic targeting of aberrant protein activity and scientific investigation of native protein function. Many different techniques for silencing proteins at the DNA or RNA level exist, but new adaptable technologies are needed to effectively silence proteins at the post-translational level, and particularly with post-translational modification resolution. One such technology, developed by the DeLisa laboratory and termed ubiquibodies, hijacks natural cellular mechanisms to silence proteins post-translationally. A synthetic enzyme—the ubiquibody—functions by connecting two independent polypeptide domains, a target recognition domain and a catalytic domain via a flexible linker. This project has focused on mapping the tolerances of the ubiquibody technology. Expanding the range of silenced targets has been a key goal, as well as investigation into each of the three ubiquibody domains: target recognition, catalytic, and the linker between.

BIOGRAPHICAL SKETCH

Erin Ashley Stephens was born in Denver, CO to Karen Stephens and Mark Stephens; later to be joined by younger sister Brooke and step-father Steve Baker. She grew up Parker, CO and attended Douglas County High School. After graduating from the International Baccalaureate Program in 2008, she enrolled at the Colorado School of Mines on scholarship. The summer after her sophomore year, she was accepted into the Genetics and Genomic Sciences Summer Undergrad Research Program at the University of Alabama at Birmingham in the lab of Dr. Robert Kesterson. Upon returning to the School of Mines, she began research in the lab of Professor Matthew Posewitz, under the guidance of graduate student Victoria Work. After graduating with a B.S. in Chemistry from Colorado School of Mines in 2012, she began her graduate degree in Biochemistry, Cell and Molecular Biology at Cornell University. Upon entering the program, she was awarded a Presidential Life Sciences Fellowship. Between September 2012 and May 2013, she rotated between the laboratories of Professors Matthew DeLisa, Volker Vogt and Jerry Feigenson, and Anthony Brestcher, eventually joining the DeLisa laboratory in June 2013 and earning an NSF Graduate Research Fellowship for 2014-2016. During this time, she studied alongside Alyse Portnoff, and many other talented graduate students and post-doctoral researchers. Erin Stephens defended this dissertation on April 19, 2018, earning her Ph.D. in Cell and Molecular Biology.

Dedicated to my mother, father, and step-father, for their unwavering supported each and every crazy endeavor I have undertaken. Dedicated to my sister, for racing me into the fields of higher-education. And dedicated to Jenna, Mika, Hanna, Melissa, Becca, and Brittany, who have had a lasting influence from the very beginning of my education.

ACKNOWLEDGEMENTS

I must acknowledge Matt DeLisa as my supervisor for the entirety of the ubiquibody research and this dissertation. Furthermore, I recognize two colleagues, Alyse Portnoff and Morgan Baltz, for their parts in the discoveries and breakthroughs made on this project. I would like to thank the BMCB field as a whole for the fantastic program they have made, especially Volker Vogt who was instrumental in my decision to attend Cornell University and my mentor for writing my NSF-GRFP proposal. And finally, I thank my family, for supporting my move across the United States to become a doctor of philosophy.

This work was supported by the National Science Foundation Career Award CBET-0449080 (to M.P.D.), National Institutes of Health Grant CA132223A (to M.P.D.), New York State Office of Science, Technology and Academic Research Distinguished Faculty Award (to M.P.D.), National Science Foundation GRFP DGE-1144153 (to E.A.S.), the Cornell Presidential Life Science Fellowship Program (to E.A.S.).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviations

3’	3’ end of DNA or RNA
5’	5’ end of DNA or RNA
AA	amino acid
Abs_{600}	absorbance at 600nm
ACN	acetonitrile
ATCC	American Tissue Culture Center
ATP	adenosine triphosphate
BirA	bifunctional ligase/repressor BirA
BL21(DE3)	protein expressing strain of *E. coli*
BN-PAGE	blue native polyacrylamide gel electrophoresis
C2	protein calcium interacting domain
CHIP	carboxy-terminus Hsc70 interacting protein
CHIPΔTPR	C-terminus of CHIP starting at amino acid 128
CID	collision-induced dissociation
CO_2	carbon dioxide
CRISPr	clustered regularly interspaced short palindromic repeats
CTD	C-terminal domain
DARPin	designed ankyrin repeat protein, DARPins mentioned in this study: OFF7, E40, pE59, EpE82, EpE89, J1/2_2_25, J1/2_2_3, G1_4_27, J1_8_27, J2_3_29
DBP	designer binding protein
DI	deionization
DMEM	Dulbecco Modified Eagle Medium
DNA	deoxyribonucleic acid
DTT	1,4-dithiothreitol
E. coli	*Escherichia coli* (bacteria)
E1	ubiquitin activating enzyme one
E2	ubiquitin conjugating enzyme two
E3	ubiquitin ligase enzyme three
E6AP	ubiquitin protein-ligase E3A
ELISA	enzyme linked immunosorbent assay
ERK	extracellularly regulated kinase, a.k.a. mitogen-activated protein, refers to both ERK1 and ERK2
ESI	electrospray ionization
FA	4-Fluoroamphetamine
FBS	fetal bovine serum
Flag	protein tag, sequence: DYKDDDDK
FN3	fibronectin type III domains, mentioned in this study: YS1, YSX1, GS5, GS2, GS8, GS6, GS4, AL2, AL38

FT	Fourier transform
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GDP	guanine diphosphate
GFP	green fluorescent protein
-GG	glycine-glycine chemical moiety
gRNA	guide RNA
GST	glutathione S-transferase
GTP	guanine triphosphate
H_2SO_4	sulfuric acid
HCD	Higher-energy collisional dissociation
HECT	homologous to E6AP carboxyl terminus
HEK293T	human embryonic kidney cell line
His_6	protein tag, sequence: HHHHHH
HPLC	high performance liquid chromatography
HRP	horseradish peroxidase
Hsc70	heat shock cognate 71kDa
Hsp70	heatshock protein 70kDa
HTT	huntingtin protein
IBR	inbetween RING
IpaH9.8	invasion plasmid antigen 9.8 E3 ubiquitin ligase
IPTG	isopropyl β-D-1-thiogalactopyranoside
K11	lysine eleven of ubiquitin
K27	lysine twenty-seven of ubiquitin
K29	lysine twenty-nine of ubiquitin
K33	lysine thirty-three of ubiquitin
K48	lysine forty-eight of ubiquitin
K6	lysine six of ubiquitin
K63	lysine sixty-three of ubiquitin
KCl	potassium chloride
LB	Lysogeny broth/Luria-Bertani medium
LC	liquid chromatorgraphy
MAPK	mitogen activated protein kinase
MBP	maltose binding protein
MEK	dual specificity mitogen-activated protein kinase kinase, refers to MEK1 and MEK2
$MgCl_2$	magnesium chloride
Mnk	MAP kinase-interacting serine/threonine-protein kinase 2
MOPS	3-(N-morpholino)propanesulfonic acid
mRNA	messenger RNA
MS/MS	tandem mass spectrometry *(*MS/MS*)*
MW	molecular weight
N17	first seventeen amino acids of HTT
NEL	novel E3 ubiquitin ligase
Ni^{2+}	nickel ion
Ni-NTA	nickel nitrilotriacetic acid

OPD	o-phenylenediamine dihydrochloride
$P67^{phox}$	neutrophil cytosol factor 2 protein
Parkin	E3 ubiquitin-protein ligase parkin
PBS	phosphate buffered saline
PBST	phosphate buffered saline with 0.01% Tween20
PCA	protein-fragment complementation assay
PCR	polymerase chain reaction
PDB	RCSB Protein Data Bank
pERK	phosphorylated ERK
Plasmids	pET28a(+) (bacterial expression vector),pcDNA3 (mammalian expression vector)
$-PO_4$	phosphate chemical moiety
Poly-Q	polyglutamine
PROTAC	proteolysis targeting chimera
PTM	post-translational modification
PVD	polyvinylidene difluoride
R0	RING domain 0 of RBR E3s
R1	RING domain 1 of RBR E3s
R2	RING domain 2 of RBR E3s
Raf1	RAF proto-oncogene serine/threonine-protein kinase
Ras	GTPase HRas
RBR	RING-between-RING
RING	really interesting new gene
RNA	ribonucleic acid
RNAi	RNA interference
S. cerevisiae	*Saccharomyces cerevisiae* (yeast)
scFv	single chain variable fragment, scFvs mentioned in this study: scFv13, 13R4, D10, J21, GCN4, NAC32, C4, Hag, 3DX, gpD, JNK2, JNK1, c-Myc
SDS-PAGE	sodium dodecyl sulfate poly acrylamide gel electrophoresis
SEC	size exclusion chromatography
SH2	SH2B adaptor protein
Smurf2	E3 ubiquitin-protein ligase SMURF2
TPR	tetratricopeptide repeat, includes NCF-2
uAb	ubiquibody
Ub	ubiquitin
Ube1	human E1 activating enzyme
Ube2	human E2 activating enzyme (family includes members Ube2A, Ube2B, etc.)
Ubl	ubiquitin like domain
UPP	ubiquitin proteasome pathway
UV	ultraviolet
V_HH	nanobody comprised of the variable heavy chain of an immunoglobin G antibody, mentioned in this study: GFP4

WW	protein membrane binding domain containing two tryptophans
YKO	yeast knockout collection
α-helix	alpha helix protein secondary structure
β-gal	beta-galactosidase
β-sheet	beta sheet protein secondary structure

Amino Acids

Amino acid	3-letter code	1-letter code
alanine	ala	A
arginine	arg	R
asparagine	asn	N
aspartic acid	asp	D
cysteine	cys	C
glutamic acid	glu	E
glutamine	gln	Q
glycine	gly	G
histidine	his	H
isoleucine	ile	I
leucine	leu	L
lysine	lys	K
methionine	met	M
phenylalanine	phe	F
proline	pro	P
serine	ser	S
threonine	thr	T
tryptophan	trp	W
tyrosine	tyr	Y
valine	val	V

Ubiquibodies

Ubiquibody	Description
R4-uAb	uAb specific for β-galactosidase using scFv13R4
C4-uAb	uAb specific for N17 using scFvC4
YS1-uAb	uAb specific for MBP using FN3 YS1
OFF7-uAb	uAb specific for MBP using DARPin OFF7
E40-uAb	uAb specific for ERK1/2 using DARPin E40
pE59-uAb	uAb specific for pERK1/2 using DARPin pE59
EpE82-uAb	uAb specific for ERK1/2 and pERK1/2 using DARPin EpE82
EpE89uAb	uAb specific for ERK1/2 and pERK1/2 using DARPin EpE89
10AA	Derivative of EpE89-uAb, using the ten amino acid linker
15AA	Derivative of EpE89-uAb, using the fifteen amino acid linker
20AA	Derivative of EpE89-uAb, using the twenty amino acid linker
134AA	Derivative of EpE89-uAb, using the one hundred, thirty-four amino acid linker

Measurements

Symbol	Value
Da	Dalton
kDa	kiloDalton
M	molar
mM	millimolar
μM	micromolar
nM	nanomolar
sec	second
msec	millisecond
nsec	nanosecond
min	minute

hr	hour
g	gram
mg	milligram
µg	microgram
ng	nanogram
×g	gravitational constant
V	volt
kV	kilovolt
L	liter
mL	milliliter
µL	microliter

CHAPTER 1

INTRODUCTION: NEXT GENERATION UBIQUIBODIES

Introduction

Silencing a protein is a powerful action. Depending on the protein, a disease may be cured, a discovery may be made, or an entirely new biological system may be created. When defining 'protein silencing', the natural function of a protein must be abolished, either by removal or inhibition of the polypeptide. When a protein is silenced, the function it performs within a biologic system is effectively lost, resulting in phenotypic changes.

Following the central dogma, proteins are synthesized by the transcription of DNA into RNA, and subsequent translation of RNA into protein. Thus, if a protein need be silenced, there are three canonical stages at which that protein's existence can be disrupted: DNA, RNA, or protein (**Figure 1.1**). Deletion of specific genes from the genome can lead to the complete loss of all encoded proteins from a system. Alternatively, messenger RNA (mRNA) can be destroyed before being translated from nucleic to amino acid, preventing the synthesis of any new protein. The third canonical stage at which proteins may be silenced is the protein level, either by removal of all copies of the peptide or by effective inhibition of every copy's function.

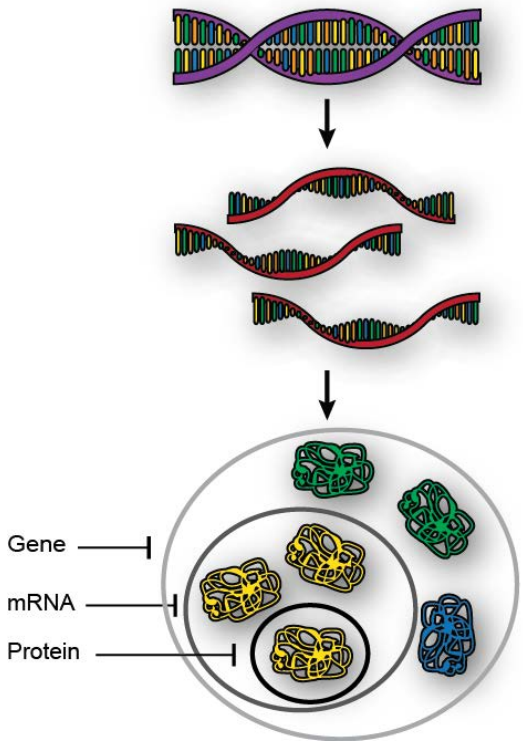


Figure 1.1: Silencing at different levels of protein synthesis.
The ratio of gene:mRNA transcript:protein product is not always 1:1:1. Alternative splicing can lead to multiple mRNA transcripts derived from the same gene. Post-translational modification can further diversify protein populations.

However, the nascent polypeptides released by the ribosome are rarely the final form of any protein. A fourth level of protein synthesis, beyond the central dogma, occurs with post-translational modification of the protein [1]. The vast majority of protein diversity comes from post-translational modifications (PTMs) (**Figure 1.2a**), and the types of PTM vary greatly, ranging from simple phosphorylation (the addition of a small phosphate group, $-PO_4$) to glycosylation (the addition of any number of complex glycan structures) (**Figure 1.2b**). The number of unique proteins is nearly three orders of magnitude greater than the number of unique genes.

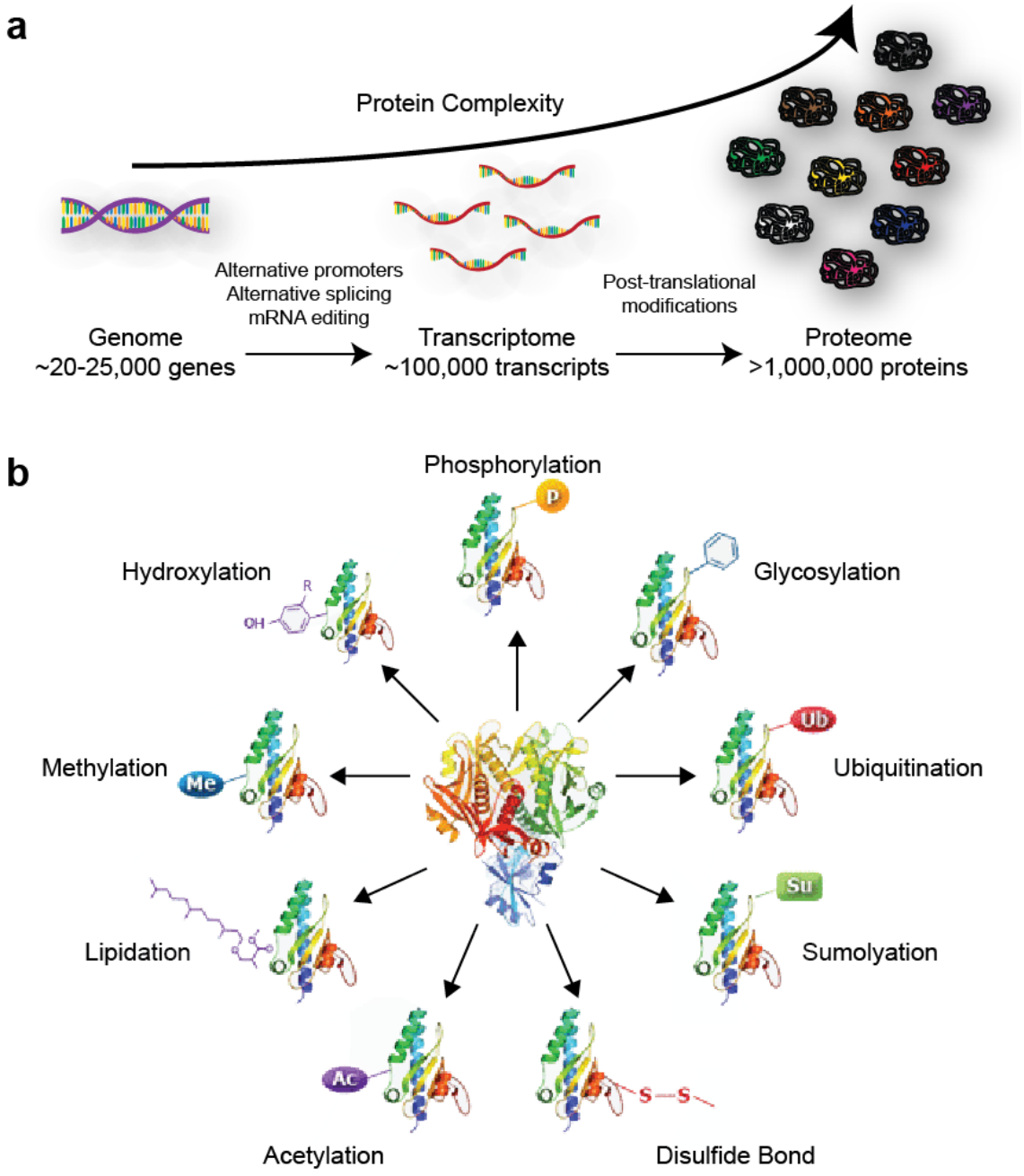


Figure 1.2: Prost translational modification of proteins.
(a) The complexity of proteins is increased by both post-transcriptional modification and post-translational modification. Adapted from ThermoFisher Scientific [2]. (b) Post translational modifications vary greatly in structure and complexity, ranging from addition of small chemical groups to cross-linkages of the polypeptide itself to conjugation of large macromolecular compounds. Adapted from Rockland [3].

Protein Silencing Levels

Current technologies developed to silence proteins span the first three stages of protein synthesis.

The ultimate silencing technology is the gene knockout; if the blueprint for a protein is deleted, that protein will never exist inside cells. Gene knockouts can be accomplished in many ways, depending on the cell type or organism being manipulated. Perhaps the most striking example of this type of protein silencing is the yeast knockout (YKO) collection, in which each gene in the *S. cerevisiae* genome was methodically knocked out one by one [4]. Notably, from this screen came a large percentage of genes, 18.7%, that were classified as essential for growth, due to loss of viability of the organism upon their deletion [5]. These essential genes are critical to biological function and often are the most in need of silencing, either for investigation or therapeutic treatment. Also, while quite powerful, gene knockouts were, for many years, inaccessible in model organisms less amenable to gene manipulation compared to yeast.

In answer to these difficulties, a revolutionary mechanism of gene knockout, clustered regularly interspaced short palindromic repeats, (CRISPr) which emerged near the start of this project, has blown open the door for protein silencing in many organisms that could previously not be studied at the DNA level [6]. CRISPr functions by repurposing a naturally occurring bacterial system, and has been streamlined to require

a single guide RNA (gRNA) and a single endonuclease, Cas9. When combined, these two elements make a powerful genome editing tool that can delete from, modify, or insert into a genome [7]. However, many of the same limitations that apply to traditional knockout methods also apply to CRISPr; essential genes cannot be completely silenced without catastrophic failure of the system's viability, and thus can only be targeted obliquely. Therefore, it is necessary to supplement DNA-level protein silence with other technologies that can parse essential gene function from organism survival.

The discovery of RNA-level silencing was quite by accident [8], though it has proved invaluable when moving to less genetically tractable systems, such as polyploid and eukaryotic systems. RNA interference (RNAi) works by hijacking a cell's natural defense mechanism against viruses. Introduction of double stranded RNA, complimentary to an mRNA of interest, can efficiently and effectively lead to the destruction of that mRNA, preventing its translation [9]. Further advancements in RNAi gave a huge boost to the temporal control of protein silencing [10]. Unlike with DNA-level silencing, when a biological system is manipulated at the RNA level, even if the loss of a protein causes the loss of system viability, the *ways* in which it shuts down become observable. This allows observation of phenotypes more specific than 'non-viable' to be connected to otherwise elusive essential genes.

Where RNAi struggles is with highly stable proteins that are turned over too slowly to observe measurable changes in phenotype caused by the inhibition of new protein synthesis. Proteins with half-lives on the order of hours or days are difficult to silence in a meaningful way with RNAi, to say nothing of proteins like collagen, which can survive for years [11]. Thus it is important to further complement DNA- and RNA-level silencing techniques with technologies that are adaptable to stable, essential proteins.

Protein-level silencing offers spatial and temporal control that is often difficult at earlier stages of protein synthesis. Most therapeutics function at the protein-level, due to the dangers of tampering with the genetic code of fully-developed, multicellular organisms. And although most protein-level approaches are susceptible to partial silencing (where intermediate phenotypes are observed due to incomplete inhibition of protein function) this is the only place where protein silencing can be achieved at post-PTM resolution. The purpose of this project has been to further develop a protein-level silencing technology, called ubiquibodies, that has the potential to silence native eukaryotic proteins by hijacking existing cellular mechanisms.

The Ubiquitin Proteasome Pathway

Eukaryotic cells naturally silence proteins effectively and efficiently; the involved mechanisms likely evolved due to the need for

rapid protein turnover to move through different stages of the cell's life cycle and to respond quickly to environmental changes. The two primary mechanisms by which cells silence proteins are the Ubiquitin Proteasome Pathway (UPP) and the lysosomal degradation via sorting and autophagy [12]. The targeted silencing potential of the lysosome has remained largely unexplored, though future generations of ubiquibodies and other technologies may soon tap into it (see Chapter 5).

The UPP utilizes a cascade of three enzymes: E1, E2, and E3, to covalently tag substrate proteins with a small peptide modifier, ubiquitin (**Figure 1.3**) [13]. Ubiquitination, the act of tagging with ubiquitin, occurs on surface lysine residues of the target; additionally, ubiquitin itself has seven lysine residues that can be ubiquitinated. Poly-ubiquitin chains of specific lysine linkages (K11 and K48) target marked substrates to the proteasome, where the ubiquitin tag is cleaved and the target protein destroyed by proteolysis. The proteasome is largely unspecific, allowing for a wide diversity of substrates to be recycled into their respective amino acids. Because the proteasome is non-specific for any particular polypeptide sequence, it falls to the E3s to determine which proteins should be tagged for degradation. In the human genome, there are approximately 700 E3s, many of which have additional chaperones, which adapt the UPP to an enormous range of substrates. This complex network of E3s allows for both tight and dynamic regulation of proteome through targeted protein degradation [14].

Ubiquibodies are synthetic E3 ubiquitin ligases, engineered to hijack the UPP. The ubiquibody recognizes a specific non-native target of interest and ubiquitinates it, so that it may meet an untimely end at the many peptidases of the proteasome.

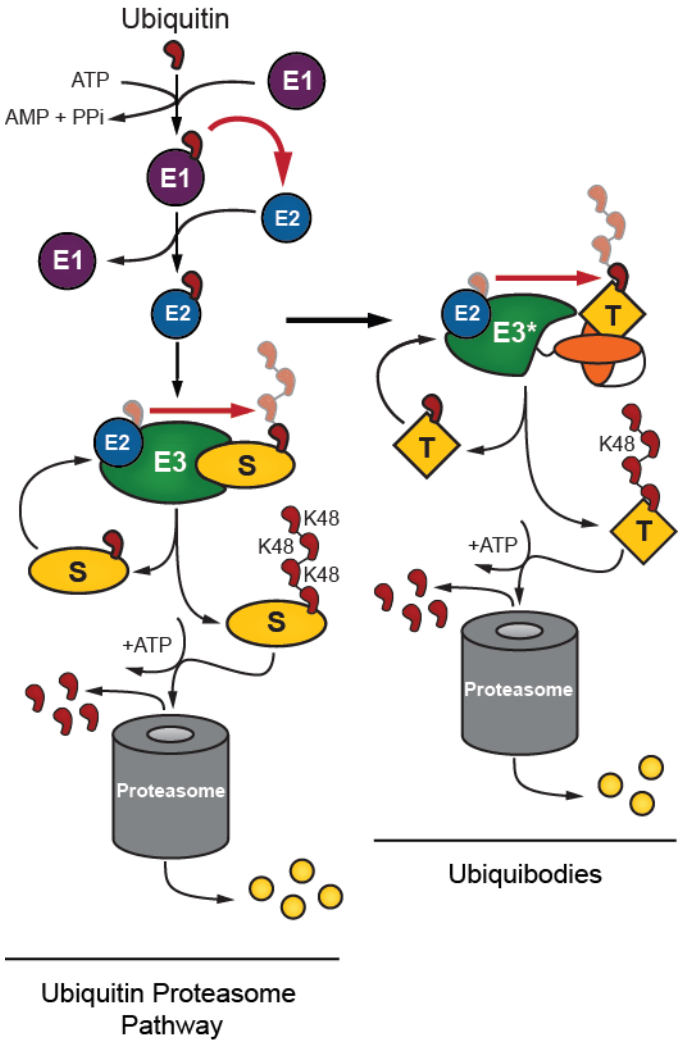


Figure 1.3: Hijacking the ubiquitin proteasome pathway.
Ubiquitin is activated in an ATP-dependent manner by the E1 activating enzyme. It is then transferred to an E2 conjugating enzyme, and then to a substrate protein (S) by an E3 activating enzyme. Poly-ubiquitination of a single substrate leads to recognition and subsequent degradation by the proteasome. Ubiquibodies (E3*) utilize existing E1 and E2 machinery to tag target proteins (T) for degradation.

Ubiquibody Design

An E3 ubiquitin ligase minimally consists of a target binding domain and a catalytic domain. The initial design of ubiquibodies used an N-terminal target recognition domain fused to a C-terminal catalytic domain via a short flexible linker (**Figure 1.4**). In previous work, ubiquibodies specific for two different bacterial targets were characterized based on their ability to bind their respective target, catalyze the formation of poly-ubiquitin chains, and facilitate the degradation of those targets in living cells [15]. The two targets, β-galactosidase (β-gal) and maltose binding protein (MBP) differ greatly in size and structure, but were effectively silenced by ubiquibodies.

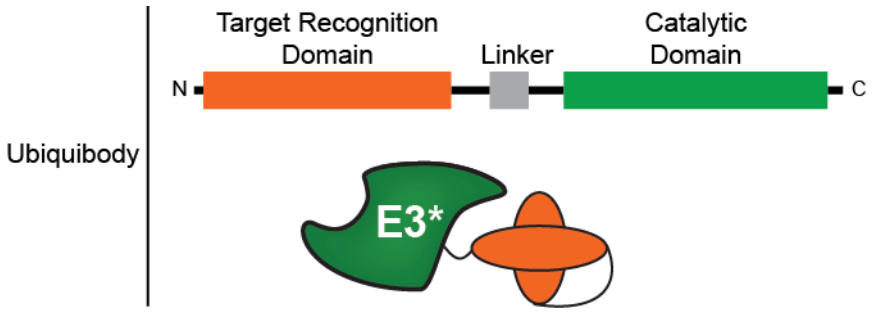


Figure 1.4: Modular design of ubiquibodies.
Ubiquibodies consist of three independent domains: target recognition, linker, and catalytic domain.

This project has focused on mapping the tolerances of the ubiquibody technology. Expanding the range of targets has been a key goal, as well as investigation into each of the three ubiquibody domains: target recognition, catalytic, and the linker between.

CHAPTER 2

THE DESIGNER BINDING PROTEIN

Introduction

The first requirement of a ubiquibody is target specificity.

Historically, and most famously, the antibody is the go-to protein for specific target recognition [16-18]. The primary function of an antibody is to bind to distinct polypeptide sequences or folds with high specificity. Ubiquibodies require a similar capability. However, unlike antibodies, ubiquibodies localize intracellularly, and, while cytoplasmic usage of antibodies has been shown [19], the cell's internal reducing environment is not ideal for these proteins. Antibodies are large, multi-chain proteins that require disulfide bonds and glycosylation for optimal function [20]. Additionally, the secondary functionality of immunogenic response activation that antibodies have is unnecessary to ubiquibody design. Therefore, the ubiquibody's target specificity must mimic that of an antibody, yet maintain a proclivity for the intracellular environment.

As an alternative to antibodies, many small protein scaffolds have been identified, some derived from the antibody itself and others from unrelated polypeptide sequences, but which, like the antibody, contain hypervariable regions for epitope recognition (**Figure 2.1**). With respect to the ubiquibody technology, the collective group of these binders have been termed designer binding proteins (DBPs).

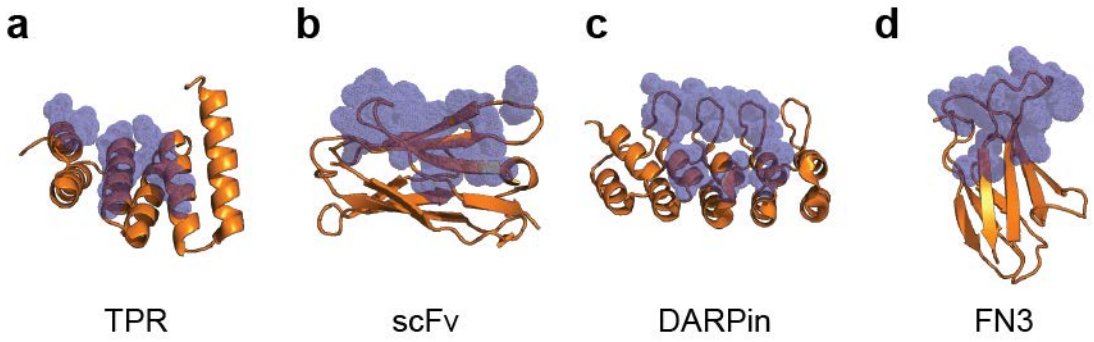


Figure 2.1: Diversity of DBPs.
While all DBPs are small, compact proteins, they span a wide range of secondary and tertiary structure; target binding paratopes are unique to each DBP scaffold. (a) TPR binds targets using contacts from the surface generated by α-helical repeats. PDB ID: 2C2L [21] (b) scFv binds targets using contacting the target via one face and the loop of three CDR regions. PDB ID: 1P4I [22] (c) DARPin contains variable residues in repeat loop and helix regions for target recognition. PDB ID: 5MA6 [23] (d) FN3 target diversity comes from randomization of the loop regions; though recent libraries also use one face. PDB ID: 3UYO [24]. Secondary structure (orange) of the DBP; DBP paratopes (blue) highlight the variable residues involved target recognition. Schematics were generated using PyMOL software.

From any given DBP scaffold, specific binders can be selected for from libraries randomized at the key residues responsible for cognate epitope interaction [25], independently from other domains of the ubiquibody. (Theoretically, libraries derived from the full ubiquibody could also be made, though that approach has not been used in this project.) The primary binders reported here were selected for by ribosome [26, 27] and phage [28] display. In every case, regardless of scaffold origin, the DBP endowed the ubiquibody with its respective target specificity; the full list of DBPs incorporated into the ubiquibodies as of this report spans five different scaffolds, mammalian and bacterial target proteins, and is shown in **Table 2.1**.

Table 2.1: DBPs incorporated into the ubiquibody scaffold

Scaffold	Name	Target	Selection	Ref
TPR	NCF-2	Rac	None[1]	[29]
scFv	scFv13	β-galactosidase	Immunization	[30]
	13R4	β-galactosidase	Bacterial selection	[31]
	D10	gpD	Bacterial selection/PCA	[32]
	J21	JNK2	PCA	
	GCN4	GCN4	Yeast two-hybrid	[33]
	NAC32	α-synuclein	Yeast surface display	[34]
	C4	Huntington	Phage display	[35]
	Hag	Hemagglutinin		[36]
	3DX	c-Myc		[37]
FN3	YS1	MBP		[38]
	YSX1	MBP		
	GS5	GFP		[24]
	GS2	GFP		
	GL8	GFP		
	GL6	GFP		
	GL4	GFP		
	AL2	Abl(SH2)		
	AL38	Abl(SH2)		
DARPin	OFF7	MBP	Ribosome display	[27]
	E40	ERK1/2		[26]
	pE59	pERK1/2		
	EpE82	ERK1/2 & pERK1/2		
	EpE89	ERK1/2 & pERK1/2		
	J1/2_2_25	JNK1/2		[39]
	J1/2_2_3	JNK1/2		
	J1_2_32	JNK1		
	J1_4_7	JNK1		
	J1_8_27	JNK1		
	J2_3_29	JNK2		
V_HH	GFP4	GFP	Immunization/ phage display	[40]

1NCF-2 is a naturally occurring TPR domain in human p67phox

Perhaps the most significant aspect of DBP-directed target specificity, is the capability to target subpopulations of protein families. Post-translationally modified proteins that could not be selectively studied by gene-knockout or RNAi are now differentiable with DBPs selected for specific protein isoforms [26].

Notably, all ubiquibodies studied in this project derive their catalytic activity from an E3 U-box scaffold (see Chapter 3). Briefly, the U-box used in ubiquibodies was taken from human carboxy-terminus Hsc70 interacting protein (CHIP), which contains a tetratricopeptide repeat (TPR) domain to direct its natural target specificity. The TPR domain is critical to E3 function, both for target recognition and for the E3 dimer conformational changes needed to transfer ubiquitin from a charged E2 to a target [41]. This natural DBP functions in tandem with other domains of the E3, allowing the U-box catalytic domain to correctly position a charged E2 and catalyze the transfer of ubiquitin to an available surface lysine residue. Any DBP incorporated into a ubiquibody that uses the CHIP catalytic domain must be able to do the same.

In previous work, we have shown that single-chain variable fragments (scFvs) and fibronectin type III domains (FN3s) are acceptable substitutes for CHIP's native TPR domain [15]. Here, the first usage of designed ankyrin repeat proteins (DARPins) in ubiquibodies is reported and plays a prominent role in expanding the functionality of the ubiquibody technology. In this work, multiple types of DBPs have been

characterized as replacements for the TPR domain of CHIP. They have been evaluated based on the ability to recognize cognate targets and to facilitate the transfer of ubiquitin from an E2 to the target protein *in vitro*. The lower limit of scFv cognate target size has been investigated; a direct comparison of two different DBPS, an FN3 and a DARPin, against the same target has been made; and exploration into PTM-specific DARPins against a conserved kinase has been focused upon.

Results

Targeted ubiquitination of N-terminal Huntington peptide

Poly-glutamine (poly-Q) containing proteins are interesting targets for directed ubiquitination, particularly with the elucidation of the ubiquitin code [42, 43], because poly-Q sequences are known to be resistant to degradation by eukaryotic proteasomes [44], which prefer proteolysis at acidic, basic, or hydrophobic residues [45]. Inclusion bodies that form from poly-Q aggregation, such as with mutant huntingtin or ataxin, also contain ubiquitin and components of the UPP [46]. Investigation into the ubiquitination patterns of these proteins may be important towards fully understanding diseases, such as Huntington's and spinocerebellar ataxia, that result from mutations extending poly-Q sequences [47].

For ubiquibodies, the huntingtin protein offered an opportunity to investigate the minimum size for cognate targets of scFv-driven

ubiquibodies. We sought to create a ubiquibody specific for the N-terminal product of the first exon of the Huntington (HTT) exon 1 gene by genetically fusing scFvC4 to CHIPΔTPR to generate C4-uAb. Residues 1-17 of the huntingtin protein product from exon 1 of the HTT gene (**Appendix A.3**) form an amphipathic helix, N17, that immediately precedes the poly-Q sequence beginning at Q18 [48]. The crystal structure of the scFvC4, in complex with the N17 peptide has been solved and the three lysine residues (K6, K9 and K15) of N17 are shown to be solvent accessible (**Figure 2.2a**).

scFvC4 directed binding of the N17 peptide was shown by enzyme-linked immunosorbent assay (ELISA) (**Figure 2.2b**). Both the scFv and ubiquibody bound well to the N17 peptide. Interestingly, the catalytic domain, CHIPΔTPR, also bound the peptide, despite the lack of a DBP. This suggests a secondary interaction of the peptide with CHIPΔTPR; theoretically, the peptide's amphipathic helix could interact with the amphipathic helices of the coiled-coil domain. To test whether N17 could be ubiquitinated by the C4-uAb, *in vitro* ubiquitination assays were performed (**Figure 2.2c**). No ubiquitination of the huntingtin N17 peptide was detected in the presence of C4-uAb. Anti-ubiquitin detection shows that the C4-uAb is an active E3 ubiquitin ligase, and the poly-ubiquitin bands likely correspond to auto-ubiquitination. Mono-auto-ubiquitination of C4-uAb was confirmed by anti-His detection of ubiquibody species.

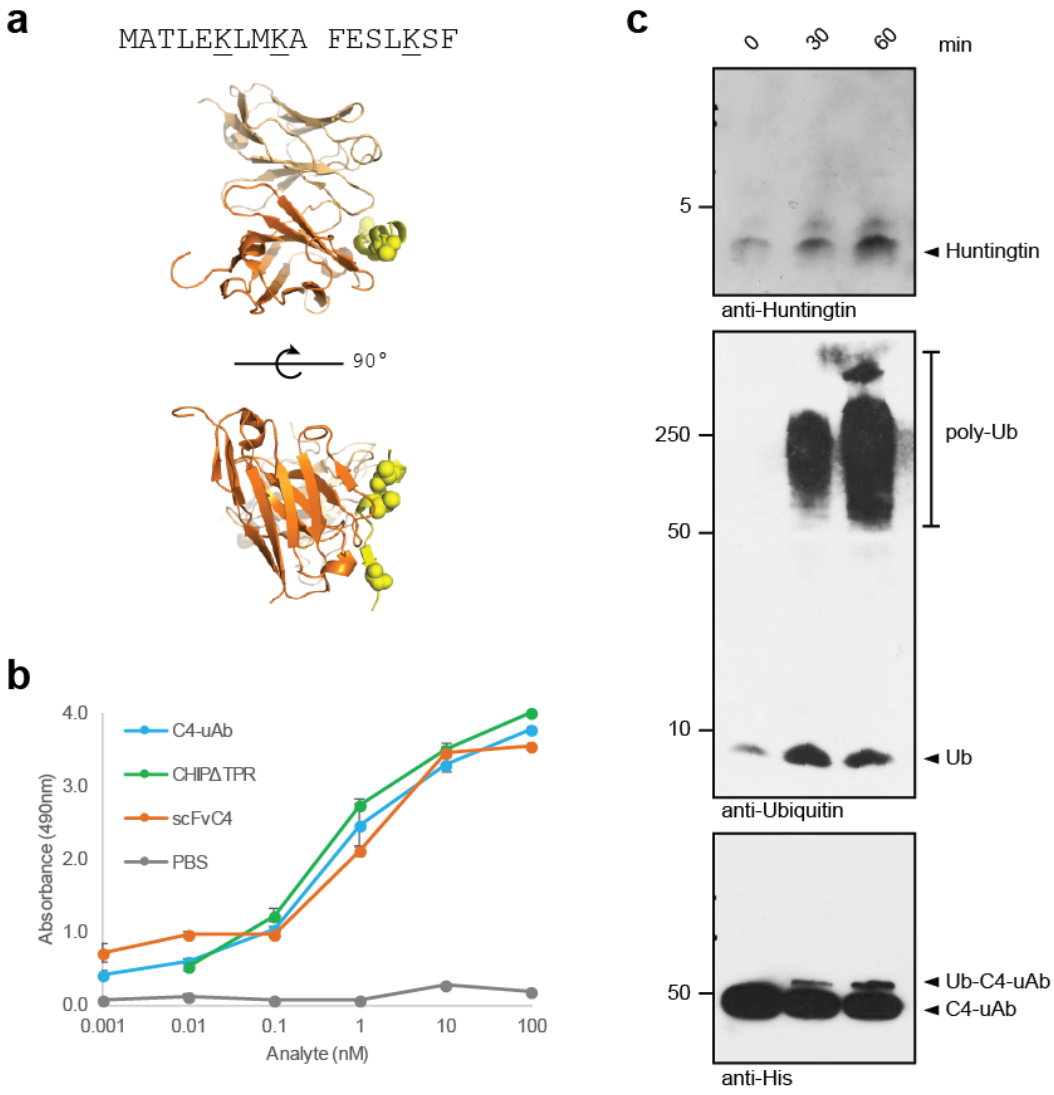


Figure 2.2: Characterization of the C4-uAb.
(a) Sequence and structure of the N17 huntingtin peptide (yellow) complexed with scFvC4 (orange). The V_H domain of the scFv (bright orange) interacts with the amphipathic helix of the N17 peptide; the three lysine residues of N17 (yellow spheres) are solvent accessible. PDB ID: 4RAV [48]. (b) ELISA analysis of N17 interaction with immobilized C4-uAb, scFvC4, and CHIPΔTPR. (c) *In vitro* ubiquitination assay of N17 peptide by C4-uAb. At the indicated times, reaction aliquots were removed and immunoblotted with the indicated antibodies.

This data suggests that seventeen residues may be outside the ubiquitinate-able size range for CHIPΔTPR, when guided by an scFv. To date, the largest protein complex to be successfully ubiquitinated by a ubiquibody is *E. coli* β-galactosidase, a 116 kDa protein that assembles

into a homotetramer of ~470 kDa. The smallest protein to be ubiquitinated with an scFv-based ubiquibody is bacteriophage capsid D (gpD), 16 kDa [49]. Taken together with the findings here, the minimum target size for successful ubiquitination by ubiquibodies—composed of an scFv DPB and the CHIPΔTPR catalytic domain—likely lies between 2-16 kDa, although the number of surface accessible lysine residues will play an important role.

Functional evaluation of an FN3- versus a DARPin

To address the question of which DBP scaffold is best suited for directed target specificity, while retaining E3 catalytic activity, a DARPin and an FN3 specific against the same target were tested *in vitro* and *in situ*. The FN3, YS1, was developed against bacterial maltose binding protein (MBP) using phage display by the Koide group [38]. YS stands for the tyrosine/serine binary code that was used to evolve binding; the annotated sequence for YS1 is shown in **Appendix A.4**. YS1 binds MBP with nanomolar affinity, primarily due to a slow off-rate (**Table 2.2**). The DARPin, OFF7, was also raised against MBP, but by ribosome display by the Plückthun group [27], (**Appendix A.5**). OFF7 has a similar on-rate against MPB, though it's off-rate is approximately thirty-fold slower than YS1, leading to K_D in the low nanomolar range. Due to the large structural differences between the two DBPs, it is conceivable that transfer of ubiquitin by the U-box catalytic domain will be better facilitated by one scaffold or the other.

Table 2.2: DBP affinity for MBP

DBP	k_a $M^{-1}s^{-1}$	k_d s^{-1}	K_D nM
YS1	7.5×10^5	5.4×10^{-2}	73
OFF7	4.2×10^5	1.9×10^{-3}	4.4

Binding data for YS1 from Gilbreth et al. [38] and for OFF7 from Binz et al. [27].

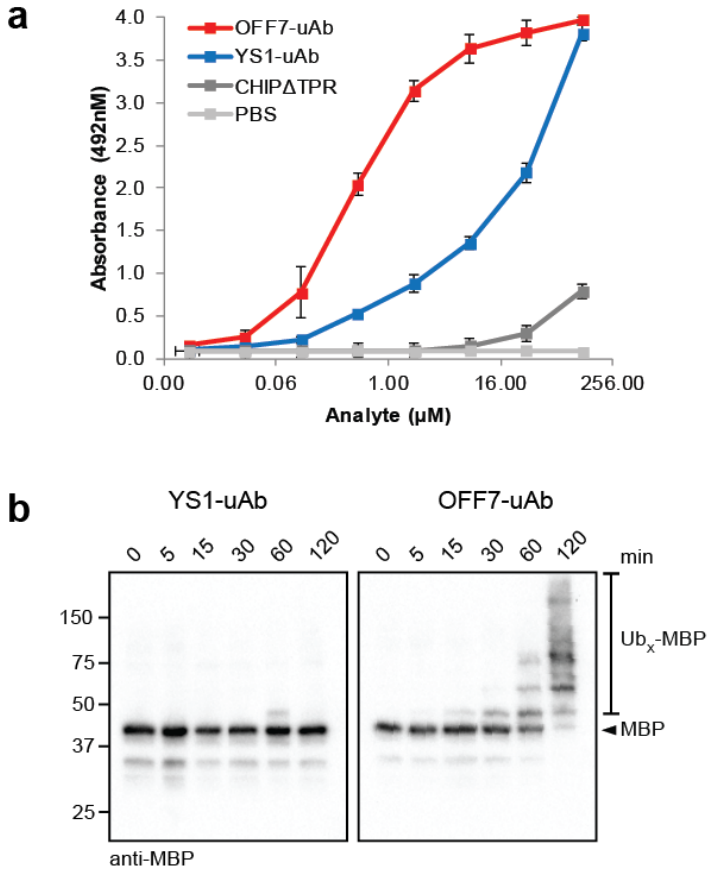


Figure 2.3: Functional evaluation of an FN3- versus a DARPin-uAb *in vitro*.
(a) DBP-directed binding of MBP measured by ELISA. Immobilized MBP was tested for interaction with YS1-uAb and OFF7-uAb, as compared to the catalytic domain alone. The DARPin-based ubiquibody, OFF7-uAb, bound with higher affinity than the FN3-based ubiquibody, YS1-uAb. The catalytic domain, CHIPΔTPR did not bind above background until concentrations greater than 10μM. (b) *In vitro* ubiquitination of MBP. At the indicated times, aliquots were boiled with loading buffer to halt the reaction, then immunoblotted with the indicated antibody.

To test whether FN3s and/or DARPins retained the ability to bind MBP when fused to CHIPΔTPR, ELISA analysis was used (**Figure 2.3a**). Purified OFF7-uAb bound MBP with higher affinity than YS1-uAb, though both DPBs were able to redirect ubiquibody specificity. The catalytic domain alone (CHIPΔTPR) did not bind MBP without the aid of a DBP. These results suggest that neither FN3 nor DARPin function is inhibited by C-terminal fusion to CHIPΔTPR.

Functional E3 catalytic activity was confirmed by incubation of OFF7-uAb and YS1-uAb with purified components of the UPP (E1, E2, ubiquitin, and ATP) and MBP, followed by immunoblot to visualize ubiquitination of the target (**Figure 2.3b**). With the OFF7-uAb, mono-ubiquitination of MBP was observed as early as 5 minutes and poly-ubiquitination after 30 minutes. With YS1-uAb, only mono-ubiquitination was observed, and only after 60 minutes. This data suggests that the DARPin was better suited—either structurally and/or functionally within the fusion protein—to facilitate the transfer of ubiquitin from the E2 to the target, across the ubiquibody scaffold.

Because MBP can be ubiquitinated in the presence of components of the UPP, MPB should be directed to the proteasome by ubiquibody activity *in vivo*. We hypothesized that the effectivity of the DARPin-uAb may be higher than the FN3-uAb, due to its higher rate of ubiquitination *in vitro*. In order to test whether the MBP can be redirected to the proteasome in mammalian cells, MBP alone, or with either YS1-uAb or

OFF7-uAb, was ectopically expressed in HEK293T cells (**Figure 2.4a** and **b**). Co-expression of MBP with either ubiquibody led to target degradation in a titratable manner. Oddly, silencing of MBP by the OFF7-uAb was inversely proportional to the DNA transfection levels; comparatively, YS1-uAb was able to silence MBP, in a manner directly proportional to DNA transfection levels [15]. Notably, OFF7-uAb required 100-fold lower amounts of transfected DNA, compared to YS1-uAb, to achieve optimal silencing, so when initial testing of the OFF7-uAb was performed, it was assumed to be inactive in mammalian cells [49], though the data here suggest otherwise. To ensure that the degradation levels were related to productive ubiquitination efficiency and not ectopic protein expression, the levels of OFF7-uAb were compared to YS1-uAb after transfection with equal molar amounts of DNA (**Figure 2.4c**). Both ubiquibodies were expressed from the same mammalian expression vector, pcDNA3, under the same promoter. YS1-uAb was expressed at a much higher level than OFF7-uAb, suggesting that OFF7-uAb's ability to silence MBP, even at much lower target:ubiquibody transfection ratios, was not due to superior protein expression. However, at optimal transfection ratios, both OFF7-uAb and YS1-uAb led to comparable silencing of MBP (**Figure 2.4d**). These results demonstrate that both DARPin- and FN3-based ubiquibodies are capable of targeted protein silencing in mammalian cells, but the necessary dynamic range of ectopic protein expression changes depending on the type of DBP used.

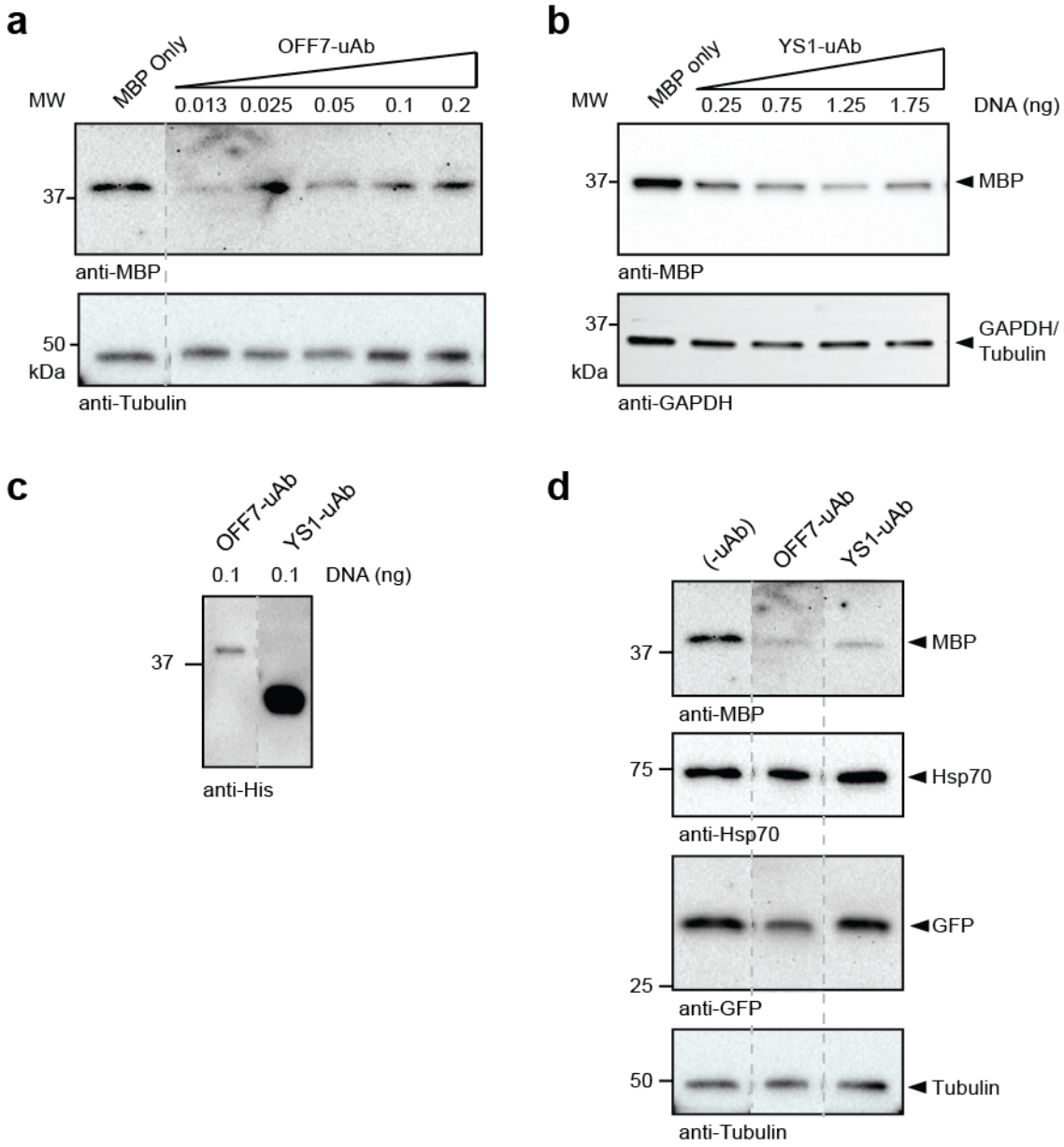


Figure 2.4: Silencing of MBP by YS1-uAb and OFF7-uAb *in situ*.
HEK293T cells were transfected with expression vectors encoding MPB, GFP, YS1-uAb, and/or OFF7-uAb; additional empty vector was used to balance the DNA levels. Cells were transfected with a 2:1 ratio of transfection agent:total DNA (ng) and incubated for 24hrs post-transfection. Samples were harvested and analyzed by immunoblot with the indicated antibodies; gel loading was normalized by total protein and confirmed by β-tubulin or GAPDH levels. (a) Silencing of MBP by OFF7-uAb. (b) Silencing of MBP by YS1-uAb. In (a) and (b), differing amounts of transfected ubiquibody DNA led to variable decreases in soluble MBP levels. (c) Expression of OFF7-uAb and YS1-uAb in HEK293T cells. (d) *In situ* silencing of MBP. Ectopically expressed MBP was silenced by both OFF7-uAb and the YS1-uAb. A native target of CHIP, Hsp70, was unaffected by uAb expression. Transfection uniformity was checked by GFP-expression.

Development of PTM-specific ubiquibodies

As outlined earlier, one of the primary milestones in the advancement of protein-level silencing is post-PTM resolution. The most common PTM is phosphorylation, with upwards of 25% or all proteins being phosphorylated [50]. In order to differentiate the non-phosphorylated from the phosphorylated isoforms of a protein, DBPs that selectively bind epitope or conformational differences between the two are needed. Near the beginning of this project, DARPins against a kinase, its non-phosphorylated isoform, and its phosphorylated form were identified by ribosome display [26]. Subsequently, we hypothesized that these phospho-specific DARPins could be used to generate phospho-specific ubiquibodies [51].

Extracellular-signal regulated kinase (ERK) is an essential protein in mammalian cells that activates upon a double phosphorylation event [52]. ERK1/2 are part of the mitogen activated protein kinase (MAPK) cascade that transduces and amplifies growth signals from the cell surface into the nucleus (**Figure 2.5**). The MAPK pathway tends to be upregulated in at least 30% of cancers, making its components ideal targets for protein silencing [53].

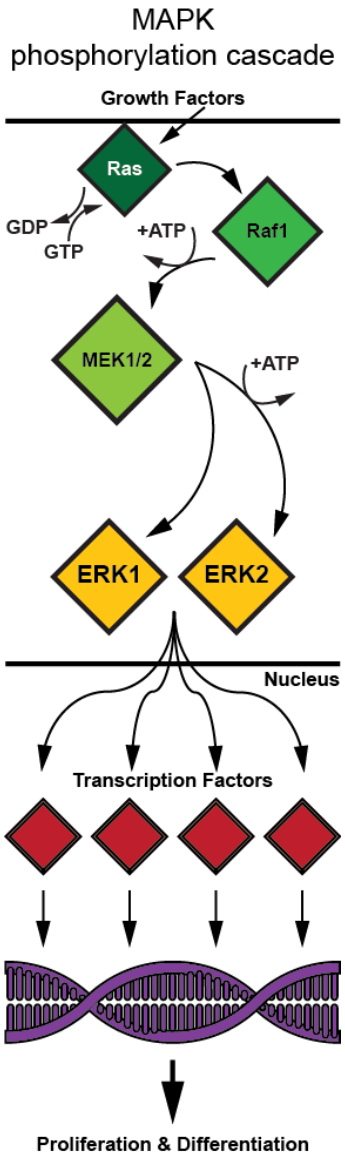


Figure 2.5: MAPK pathway.
Canonical MAPK activation upon signal detection includes three level (MAPKKK, MAPKK, and MAPK) phosphorylation signal transduction. ERK1/2 are kinases that function at the MAPK level of signaling.

Kummer et al. reported DARPins specific for inactive forms of the kinase (E40), specific for active doubly-phosphorylated kinase (pE59), and unspecific for PTM isoforms of ERK (EpE89 and EpE82), including co-crystal structures of E40 in complex with ERK2 and pE59 in complex with phosphorylated ERK2 (pERK2) [26]. Unsurprisingly, both DARPins

bound ERK near the Thr185/Tyr187 active site, where epitope differences between ERK2 and pERK2 are greatest. The respective affinities and specificities of these DARPins are shown in **Table 2.3**.

Table 2.3: Affinity of PTM-specific DARPins for ERK2/pERK2

DBP	K_D: ERK2 nM	K_D: pERK2 nM	Specificity
E40	6.6	1200	182
pE59	>8700	117	74
EpE89	-	-	None
EpE82	-	-	None

Data from Kummer et al. [26].

The DARPins identified in this screen were used to construct a suite of four ERK-specific ubiquibodies (E40-uAb, pE59-uAb, EpE89uAb, and EpE82-uAb). To test whether the ERK-ubiquibodies inherited their parent phospho-specificity, ELISA analysis of the ubiquibodies against ERK2 and pERK2 was performed (**Figure 2.6**). Isoform-specific DARPins retained their specificity when C-terminally fused to CHIPΔTPR; furthermore, the fusion seems to universally increase affinity, without affecting specificity. E40-uAb bound immobilized non-phosphorylated ERK2 with higher affinity than pERK2. Conversely, the pE59-uAb bound pERK2 with higher affinity than non-phosphorylated ERK2. EpE89-uAb bound both isoforms of ERK2 with comparable affinity. The three ubiquibodies bound their respective preferred target(s) with similar affinity. The unfused DARPins bound the expected targets, albeit at much lower affinities than the ubiquibodies (except for DARPin EpE89 which

bound pERK2 with significantly higher affinity than ERK2). Notably, both E40-uAb and pE59-uAb bound their non-cognate isoform, pERK2 and ERK2 respectively, with affinities higher than background. This is consistent with the findings of Kummer et al. that show E40 and pE59 retain micromolar affinities for their non-cognate isoform of ERK respectively [26].

The improvement of binding upon fusion with the catalytic domain CHIPΔTPR suggests that the DARPin-target interaction is stabilized by the addition of a C-terminal fusion. Ironically, the DARPins did not bind to ERK *in vitro* when a C-terminal Flag-His_6 was used (data not shown). Alternatively, dimerization by the ubiquibody may play a role in increased signal detection. The ability to preferentially detect one isoform versus another, differing only in two $–PO_4$ moieties of otherwise identical proteins, was an important first step in the generation of PTM-specific ubiquibodies.

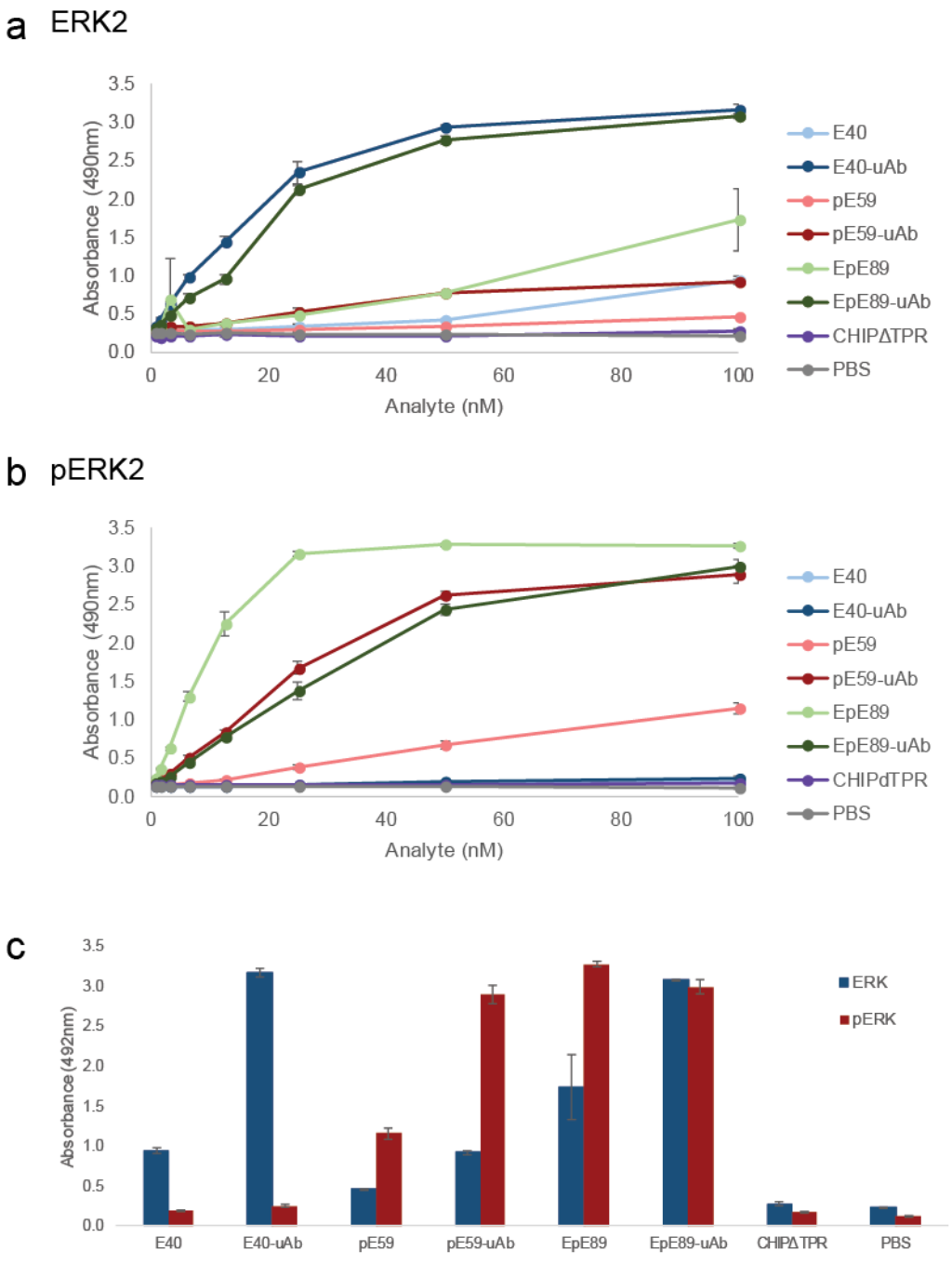


Figure 2.6: Binding specificity analysis of PTM-specific ubiquibodies.
Full ELISA curves of purified DARPins and ubiquibodies against immobilized (a) ERK2 or (b) pERK2. (c) Binding comparison of DARPins and DARPin-uAbs to ERK2 and pERK2 at 100nM. The E3 catalytic domain lacking a DBP, CHIPΔTPR, served as a negative control. [15]

Evaluation of PTM-specific ubiquibodies

To determine if phospho-specificity of the DBP leads to preferential ubiquitination of ERK and pERK, *in vitro* ubiquitination assays were performed (**Figure 2.7** and **Figure 2.8**). Ubiquitination of both isoforms was observed in the presence of either E40-uAb, pE59-uAb, or EpE89-uAb. Slightly stronger mono-ubiquitination of ERK2 was seen as early as 5 min with E40-uAb, compared to pE59-uAb or EpE89-uAb. Strong poly-ubiquitination smears were observed after 30 min for the pE59-uAb and EpE89-uAbs, and after 60min for the E40-uAb. Auto-ubiquitination of only the pE59-uAb was observed, though not of the E40-uAb or EpE89-uAb (though this may have been an artifact of detection of the ubiquibody through its Flag tag, which contains a ubiquitinatable lysine moiety). Ubiquitination of pERK2 yielded similar results, with pE59-uAb and 89-uAb showing faster formation of poly-ubiquitin species than the E40-uAb, though all three ubiquibodies showed target ubiquitination as early as 5min.

These results show that effective ubiquitination of ERK2 and pERK2 can be achieved in the presence of excess UPP pathway components, despite affinity differences between the DBPs. Furthermore, the dynamic range for ubiquitination by the U-box in general must minimally span nanomolar to micromolar target affinities. Future investigations into the kinetics of uAb-target-E2 interactions could illuminate the optimal DBP:target affinity, though secondary affinities for

non-cognate targets will need to be minimized lower than micromolar ranges (see Chapter 5).

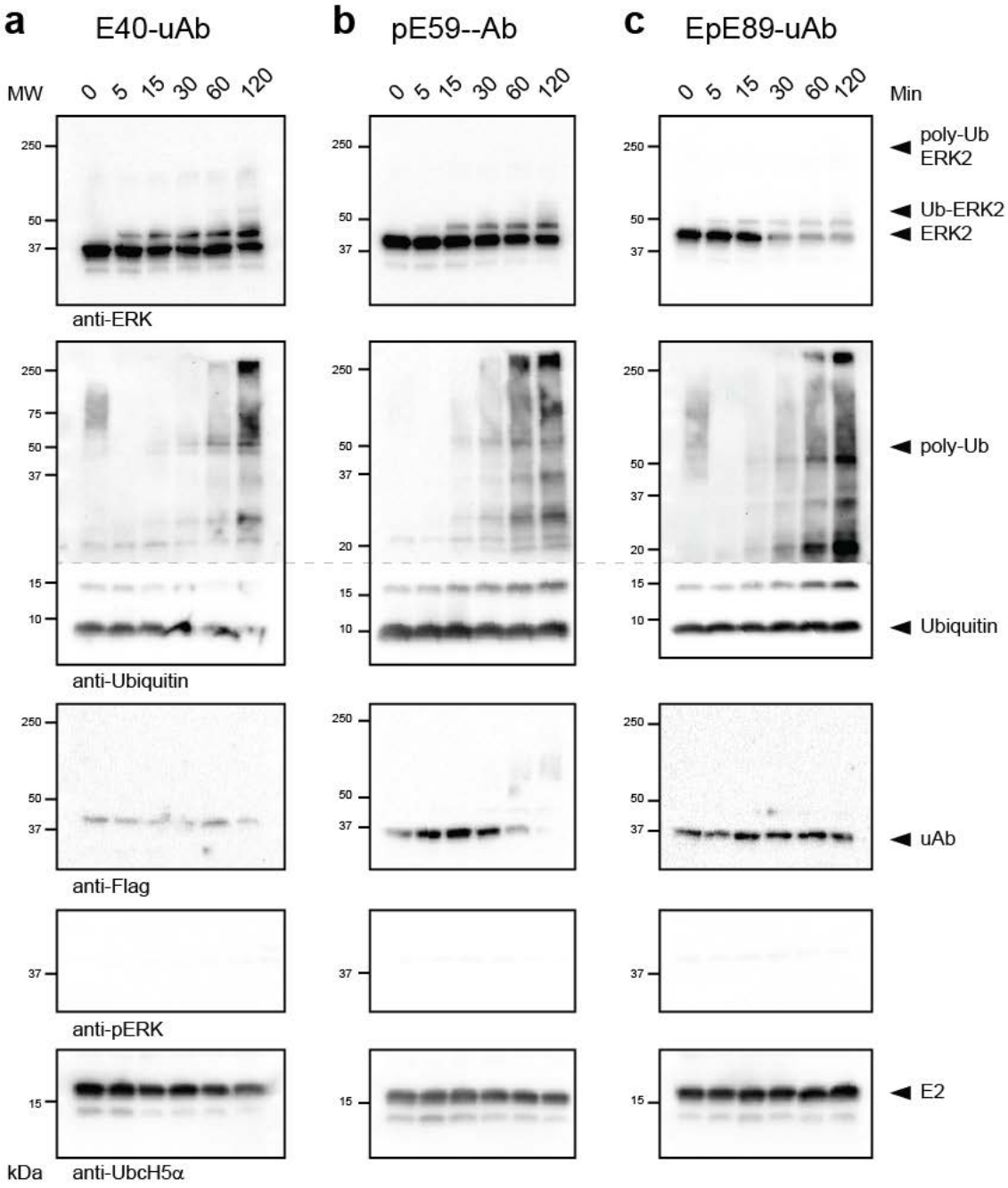


Figure 2.7: *In vitro* ubiquitination of ERK2.
Immunoblots of *in vitro* ubiquitination assays of ERK2 with (a) E40-uAb, (b) pE9-uAb, and (c) EpE89-uAb. Reactions were halted at the indicated time points and analyzed with the indicated antibodies.

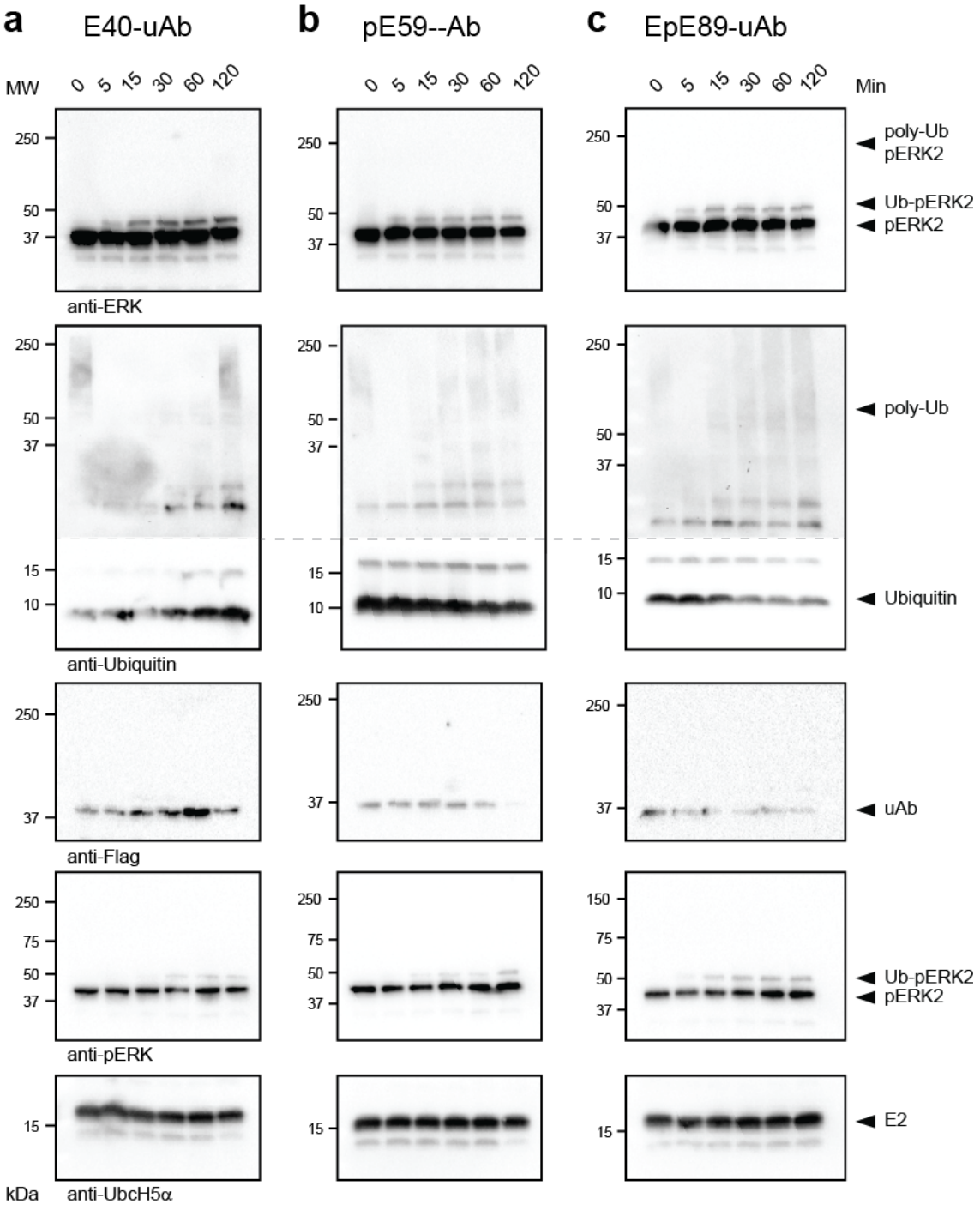


Figure 2.8: *In vitro* ubiquitination of pERK2.
Immunoblots of *in vitro* ubiquitination assays of pERK2 with (a) E40-uAb, (b) pE9-uAb, and (c) EpE89-uAb. Reactions were halted at the indicated time points and analyzed with the indicated antibodies.

The ability of the ubiquibodies to interact with endogenous ERK species was confirmed by pulldown assay of HEK293T cell lysate (**Figure 2.9**). Purified DARPins interacted strongly with endogenous ERK1 and ERK2; the corresponding ubiquibodies also interacted with ERK1/2 though at lower levels. E40-uAb showed a preference for ERK2 over ERK1, while pE59 and pE89 were unspecific for either homolog. EpE82-uAb did not pulldown endogenous target above background levels. Neither a non-specific ubiquibody, OFF7-uAb, nor the catalytic domain alone, CHIPΔTPR, interacted with endogenous ERK species.

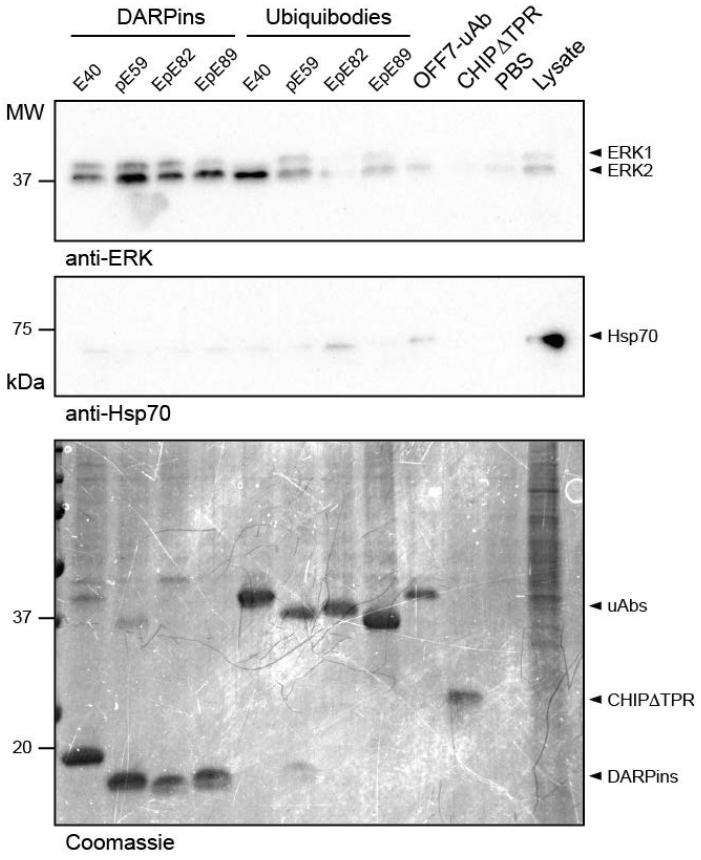


Figure 2.9: Pulldown assay of ERK-ubiquibodies from HEK293T cells.

Immunoblots of captured uAbs and control proteins incubated with HEK292T cell lysate. ERK-DARPins and –uAbs interacted with endogenous ERK1/2, while non-specific ubiquibody and CHIPΔTPR did not.

Next, to check for soluble expression of ubiquibodies in mammalian cells, HEK293T cells were transiently transfected with plasmid DNA encoding either E40-uAb, pE59-uAb or EpE89-uAb (**Figure 2.10**). Ubiquibody expression 24hrs post-transfection was detected by the His_6 tag of the ubiquibody; all three ubiquibodies were expressed at detectible levels, though E40-uAb was expressed at significantly lower levels than either pE59-uAb or EpE89-uAb. This lower expression may be tied to the basic structure of DARPin E40, which contains three ankyrin repeats, as opposed to two, as pE59 and EpE89 do. (OFF7-uAb, whose DBP is also a DARPin of three repeats, was similarly less well-expressed compared to other ubiquibodies, **Figure 2.4c**.)

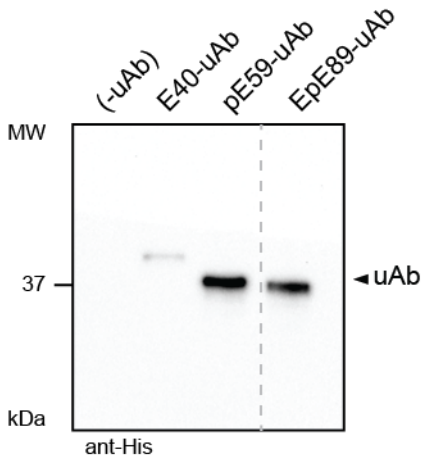


Figure 2.10: Expression of DARPin ubiquibodies in HEK293T cells.
Immunoblot of HEK293T cells transiently transfected with pcDNA3-uAb vectors and detected with the indicated antibody.

In order to test the ubiquibodies' ability to silence ERK *in vivo*, HEK293T cells were again transiently transfected with plasmid DNA encoding either EpE89-uAb, E40-uAb, or pE59-uAb (**Figure 2.11**). Samples were analyzed by immunoblot 24hrs after transfection for global, soluble ERK levels. Expression of the EpE89-uAb significantly lowered ERK levels in living cells (**Figure 2.11a** and **b**). Interestingly, ERK1/2 levels were inversely proportional to ubiquibody transfection levels, and scaled linearly with the amount of transfected DNA. This is similar to results seen with OFF7-uAb against ectopically expressed MBP, suggesting that the effect is an effect of DARPin-based ubiquibody mediated silencing and not related to the source of target protein expression.

Phospho-specific ubiquibodies were also tested for effect on endogenous ERK levels. Ectopic expression of the E40-uAb led to no measurable silencing of ERK1/2 (**Figure 2.11c**). Since the relative expression of E40-uAb was much lower than either of EpE89-uAb or pE59-uAb, it is possible that the level of ubiquibody expression that would lead to measurable silencing was outside the parameters of this assay. Surprisingly, ectopic expression of the pE59-uAb resulted in near complete silencing of global ERK1/2 levels in cells (**Figure 2.11d**). Ectopic expression of neither DARPins E40 and pE59 nor the U-box domain, CHIPΔTPR, alone affected endogenous ERK1/2 levels. Importantly, a downstream substrate of ERK1/2 signaling, Mnk, was

unaffected by the presence of pE59-uAb and resultant silencing of ERK1/2, indicating targeted specific activity of the ubiquibody in the presence of the full MAPK pathway. Collectively, these data show that effective silencing of endogenous targets is closely tied to expression levels of the ubiquibody, and the affective range of which may be tied to DBP-determined affinity. And, although pE59-uAb was theorized to specifically silence pERK, both the EpE89-uAb and pE59-uAb are functional silencers of an endogenous target protein [51].

Because *in vitro* ubiquitination, as visualized by immunoblot, did not correlate to equal silencing effects by the ubiquibodies *in vivo*, we sought to elucidate the ubiquitination profile of ERK by each ubiquibody. High molecular weight products (50-250kDa) from *in vitro* ubiquitination reactions with the E40-uAb, pE59-uAb, and EpE89-uAb were analyzed by mass spectrometry (**Figure 2.12**). As previously seen by immunoblot, ubiquitination sites on both ERK2 and pERK2 were identified for all three ubiquibodies. The ubiquitination profiles of the ubiquibodies overlapped significantly and preferentially ubiquitinated one face of (p)ERK2 (oriented forward in **Figure 2.12b** and **d**). This preferred face consisted of the plane formed by N- and C-lobes near the active site of ERK2, which correlated with the positioning of the C-terminus of the DARPins seen in co-crystal structures with ERK2 [26], and would thus be located in closest proximity to CHIPΔTPR when bound by the ubiquibody fusion protein (**Figure 2.13**).

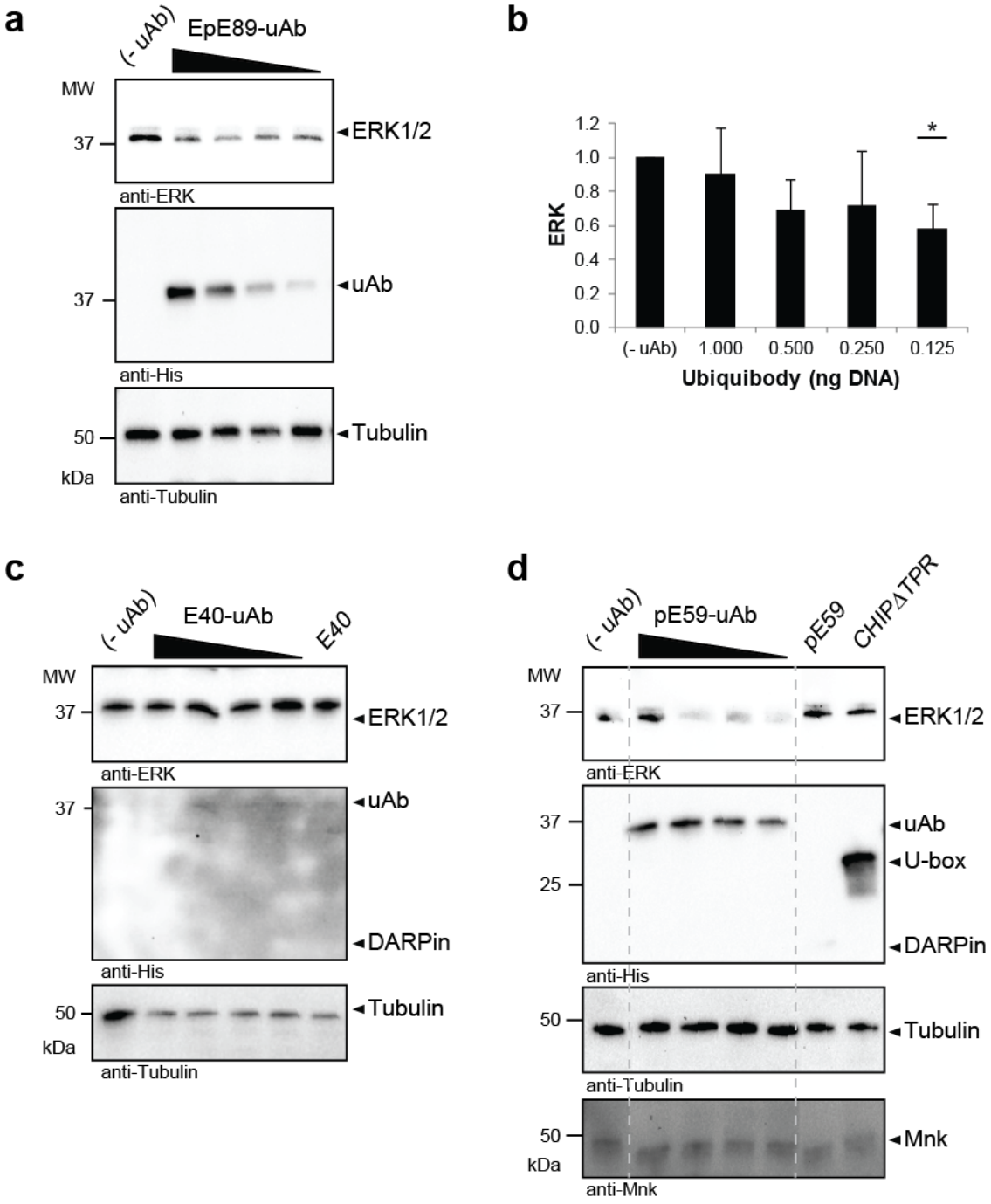


Figure 2.11: Silencing of endogenous ERK in HEK293T cells.
Representative immunoblots of extracts prepared from HEK293T cells transfected with, µg plasmid DNA per well, (a) pcDNA3-EpE89-uAb at 1.0, 0.5, 0.25 and 0.125 (c) pcDNA3-E40-uAb at at 1.0, 0.5, 0.25 and 0.125, and pcDNA3-E40 at 1.0, (d) pcDNA3-pE59-uAb at 2.0, 1.0, 0.5, and 0.25, and pcDNA3-pE59 and pcDNA3-CHIPΔTPR at 2.0. Transfection were balanced with empty pcDNA3 vector and cells were harvested 24 hr post-transfection. Lanes were normalized by total protein content and confirmed by β-Tubulin immunoblot. (b) Silencing of global ERK levels by the EpE89-uAb. Bars were generated by averaging densitometry analyses from three independent experiments using Image Lab software [15].

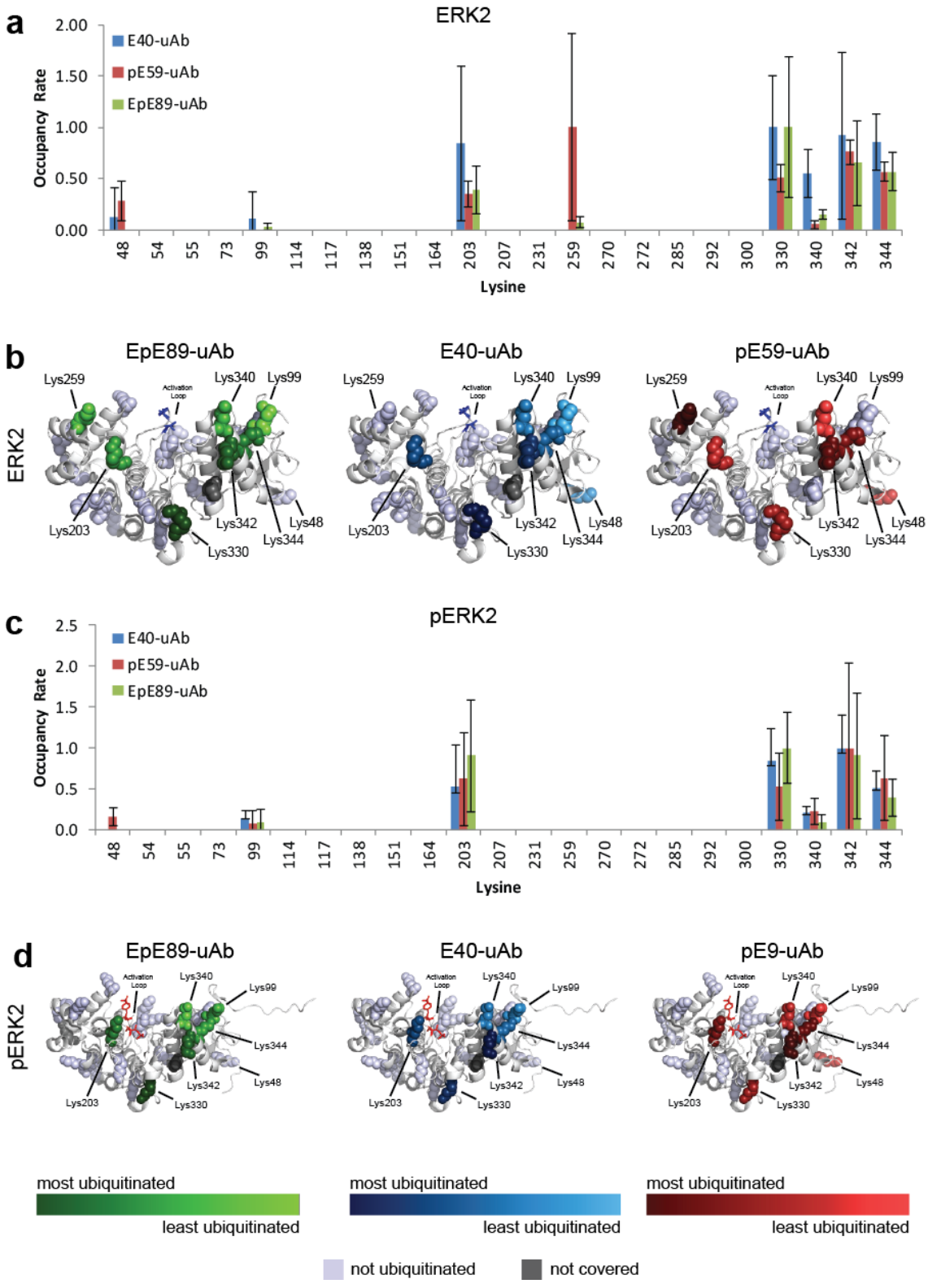

a
ERK2
E40-uAb
pE59-uAb
EpE89-uAb
Occupancy Rate
2.00
1.50
1.00
0.50
0.00
48 54 55 73 99 114 117 138 151 164 203 207 231 259 270 272 285 292 300 330 340 342 344
Lysine
b
ERK2
EpE89-uAb
E40-uAb
pE59-uAb
Activation Loop
Lys259
Lys340
Lys99
Lys203
Lys342
Lys48
Lys344
Lys330
c
pERK2
E40-uAb
pE59-uAb
EpE89-uAb
Occupancy Rate
2.5
2.0
1.5
1.0
0.5
0.0
48 54 55 73 99 114 117 138 151 164 203 207 231 259 270 272 285 292 300 330 340 342 344
Lysine
d
pERK2
EpE89-uAb
E40-uAb
pE9-uAb
Activation Loop
Lys340
Lys99
Lys344
Lys48
Lys342
Lys203
Lys330
most ubiquitinated
least ubiquitinated
most ubiquitinated
least ubiquitinated
most ubiquitinated
least ubiquitinated
not ubiquitinated
not covered

Figure 2.12: Ubiquitination profiles of ERK and pERK.
(a and b) Occupancy rate of –GG modification of ERK2/pERK2 lysine residues by mass spectrometry. Bars were generated by normalizing ubiquitinated residue counts relative to total residue counts then by overall target ubiquitination frequency, and by averaging across three independent experiments. Ubiquitinated peptide counts of peptides containing more than one non-C-terminal lysine residue were averaged over all non-C-terminal lysines. (c and d) Map of ubiquitination sites on ERK2 and pERK2 respectively. ERK2/pERK2 backbones are shown as ribbons (white), with lysines not covered by mass spectrometry analysis (dark grey spheres) and lysines not identified as ubiquitinated (light grey spheres). Lysines identified as ubiquitinated are represented as spheres colored by decreasing saturation with decreasing frequency of ubiquitination, with sites modified by EpE89-uAb (green), E40-uAb (blue), and pE59-uAb (red). Crystal structures from (b) PDB ID: 3ZU7 and (c) PDB ID: 3ZUV [26].

A total of eight lysine residues were found to be ubiquitinated on ERK2, and a subset of seven of those were ubiquitinated on pERK2, despite the presence of twenty-three surface exposed lysines in the crystal structure of ERK2. Of the ubiquitinated lysines, four (K99, K340, K342, and K344) were clustered in the three-dimensional structure of ERK2, suggesting an optimal orientation of the charged E2 (Ube2D1) relative to the target surface. Only one lysine, K259, was ubiquitinated in ERK2 and not pERK2, suggesting that the conformational changes upon phosphorylation reposition K259 away from the U-box-bound, ubiquitin-charged E2. K203 and K330 were among the most often ubiquitinated lysines, despite being located away from the lysine cluster, though both were located on the same face of ERK2. K48 was both one of the least often ubiquitinated residues and the only lysine not on the same face of ERK to be ubiquitinated.

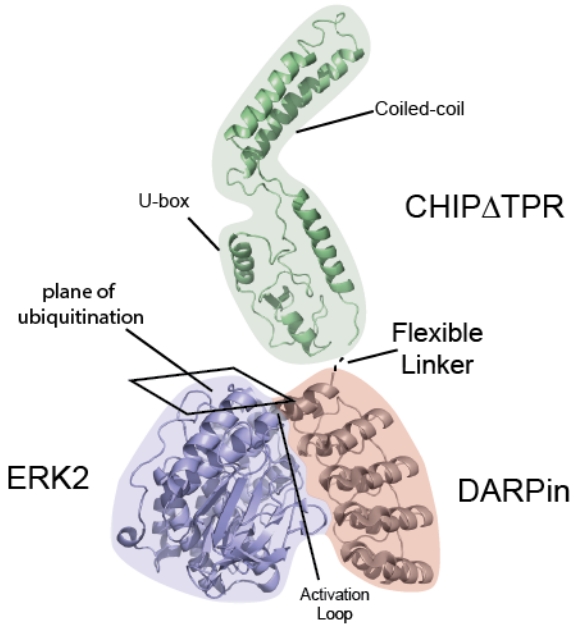


Figure 2.13: Orientation of U-box domain in ERK-uAb complex.
Cartoon diagram of potential domain arrangement of the ubiquibody relative to the target. The flexible linker is genetically fused to the C-terminus of the DARPin, and the N-terminus of CHIPΔTPR. The plane of ubiquitination is in direct proximity to the theoretical position of the catalytic domain. Schematics were generated using PDB ID: 2C2L [21] and 3ZU7 [26] and Pymol software.

Of the three ubiquibodies, pE59-uAb showed the most drastic change in its ubiquitination profile from ERK2 to pERK2. pE59-uAb showed high levels of ubiquitination of K259 that was completely abolished upon phosphorylation of the substrate. Conversely, pE59-uAb showed no ubiquitination of K99 in the non-phosphorylated form of ERK2, but showed low levels of ubiquitination of the same residue in pERK2. The E40-uAb was the only ubiquibody that did not ubiquitinate K259 in ERK2; the ubiquitination profile of E40-uAb also changed between ERK2 and pERK2, ubiquitinating K48 in ERK2 but not pERK2. Ubiquitination of K340 by the E40-uAb was also reduced between ERK2 and pERK2. Similar to the pE59-uAb, the EpE89-uAb did not

ubiquitinate K259 in the phosphorylated form of ERK2, although the rest of its ubiquitination profile remained similar between ERK2 and pERK2. Notably, all three ubiquibodies showed similar ubiquitination profiles of ubiquitin, preferentially forming K6, K11, K48, and K63 poly-ubiquitin linkages in the presence of the E2 Ube2D1 (see Chapter 3).

Taken together, these results suggest that even highly similar ubiquibodies can have significantly different target ubiquitination profiles, which may be responsible for the dramatically different silencing capabilities of the ubiquibodies *in vivo*. All three of the DARPins reported here bind the same epitope of ERK2 [26], suggesting that the observed differences in silencing and ubiquitination are not the result of radical repositioning of the U-box catalytic domain relative to the target. The small differences in catalytic efficiency in combination with large differences in affinity of each ubiquibody are likely responsible for the resultant silencing of endogenous targets.

As ERK1/2 are both essential, highly-regulated proteins, silencing them for long periods of time may prove difficult, as cells will naturally attempt to return their activity to base levels. To determine the effective time frame of ubiquibody-mediated silencing, HEK293T cells were transiently transfected with EpE89-uAb and incubated for 72hrs (**Figure 2.14**). At 24hr and 48hrs, effective EpE89-uAb dependent silencing of ERK1/2 was observed, though expression of the ubiquibody decreased marginally. After 72hrs, the amount of ERK1/2 protein returned to levels

comparable to cells transfected with empty vector. These results show that silencing of essential proteins can be ameliorated over time in living cells, despite continued presence of the ubiquibody.

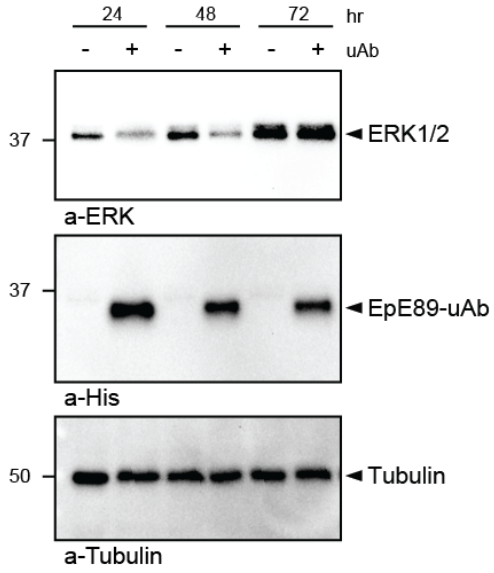


Figure 2.14: Silencing of ERK in mammalian cells over time.
Immunoblots of whole cell lysate of HEK293T cells transfected with either empty pcDNA3 or pcDNA3-EpE89-uAb. Cells were harvested at the indicated time points and analyzed with the indicated antibodies. [51]

Discussion

The designer binding protein is one of the two essential domains of engineered E3 ubiquitin ligases. Because DBPs are not limited to a single scaffold, they offer near unlimited variability to ubiquibody construction. In this report, limitations to the target size have been bracketed, comparisons between different DBP scaffolds against the same target substrate have been made, and the first attempt at post-PTM resolution silencing has been executed. The affinity of the DBP within the

ubiquibody scaffold has recurred as an important feature for successful target ubiquitination and subsequent silencing *in vivo*, though its effects have not been investigated directly.

Ironically, much of the ongoing research into DPBs focusses on identifying binders that have the highest possible affinities, often approaching low picomolar ranges [22, 54, 55]. The inability of the E40-uAb to silence a target to which it has low nanomolar affinity, while similar ubiquibodies with lower affinities to the same target are functional silencers, suggests that an optimal range of target affinity exists for ubiquibodies. The differential expression levels of OFF7-uAb and YS1-uAb required to silence the same target support this hypothesis. Indeed, most biological interactions take place in the high nanomolar to micromolar range [56]. Alternatively, in this case, there may be inherent differences in the ability of FN3-based ubiquibodies versus DARPin-based ubiquibodies to silence proteins that stem from differences in DBP structure and not affinity.

While perfect PTM-specific silencing by ubiquibodies has not been achieved as of yet, significant steps have been made towards generating phospho-specific silencers of endogenous ERK. Indeed, in the course of this project, the first ubiquibodies against an endogenous target have been successfully tested. In order to expand the range of PTM-specific silencing, further screening of DBPs will need to be made with careful attention to selection against non-preferred target isoforms. Notably,

PTM-specific antibodies are directly adaptable into the scFv format, so recent advancements made in PTM-specific antibodies are a rich source of potential for PTM-specific ubiquibodies [3]. Further exploration into the ERK versus pERK silencing effects of the phospho-specific ubiquibodies will help to better understand the expression and kinetic requirements of PTM-specific ubiquibodies

While this report focuses on the fusion of DBPs to a U-box catalytic domain, the modular nature of the ubiquibody allows for other catalytic domains to be used. The optimal type of DBP may change depending on the other domains of the ubiquibody. As for CHIPΔTPR, all tested types of DBP can function as substitutes of the native TPR, ranging from scFvs to FN3s to DARPins.

Materials and Methods

Genes and plasmid construction. Ubiquibodies, targets, and control proteins were cloned following the design and protocol outlined previously [15]. The following genes were provided as kind gifts from other labs: DARPin OFF7, from Marc Ostermeir (Johns Hopkins); YS1, from Shohei Koide (University of New York); DARPins E40, pE59, EpE89, EpE82, pLV-ERK2-Avi, pLV-MEK1R4F-ERK2-His-Avi, from Andreas Plückthun (University of Zurich); full length human CHIP, from Cam Patterson (Weill Cornell Medical Center); pSP10B BirA, from Amy

Karlsson (University of Maryland). scFvC4 was synthesized by GenScript based on NCBI sequences.

Bacterial Expression Vectors: All ubiquibodies were cloned into pET28a(+) between the NcoI and HindIII sites, in the format NcoI-ATG-DBP-EcoRI-GSGSG-CHIPΔTPR-SalI-Flag-His_6-Stop-HindIII. DBPs were PCR amplified, introducing a 5' NcoI site and start codon and 3' EcoRI codon, followed by double digestion of the product and backbone, pET28a(+)-R4-uAb, with NcoI/EcoRI, and subsequent ligation to yield pET28a(+)-DBP-uAb vectors. (Due to an internal EcoRI site in the FN3 monobody YS1, overlap extension PCR was used to add the GSGSG linker and N-terminus of CHIPΔTPR to the YS1 PCR product, until reaching a unique BstBI site within CHIPΔTPR. This overlap extension PCR product was then ligated between NcoI and BstBI sites of pET28a-R4-uAb, yielding pET28a-YS1-uAb.)

Control scFvs, FN3s, and OFF7 were similarly cloned into pET28a(+) between the NcoI and HindIII sites, in the format NcoI-ATG-DBP-SalI-Flag-His_6-Stop-HindIII. Cloning was the same as for the ubiquibodies, except during PCR amplification a 3' SalI site was introduced instead of the EcoRI, and subsequent double digestion, with NcoI/SalI, and ligation. Control DARPins (with the exception of OFF7) were cloned into pET28a(+) between NcoI and HindIII sites in the format NcoI-ATG-RGS-His_6-BamHI-DARPin-N-Stop-HindIII, using PCR amplification, followed by double digestion with NcoI/HindIII and ligation

to yield pET28a(+)-DBP vectors. CHIPΔTPR was cloned into pET28a(+) in the format NcoI-ATG-CHIPΔTPR-SalI-Flag-His_6-Stop-HindIII by PCR amplification of DNA corresponding to amino acids 128-303 of human CHIP, with introduction of a 5' NcoI site and a 3' SalI site. Double ligation was performed as above to insert this product into the backbone pET28a(+)-R4-uAb to yield pET28a(+)-CHIPΔTPR.

Mammalian expression vectors: All ubiquibodies and control proteins were cloned into pcDNA3 between the HindIII and XbaI sites by PCR amplification of the respective pET28a(+) vector, introducing a Kozak sequence at the start codon, and subsequent double digestion with HindIII/XbaI and ligation into backbone pcDNA3. The target substrate protein MBP was similarly cloned into pcDNA3 between HindIII and XbaI.

Protein expression and purification. All purified proteins were obtained from cultures of *E. coli* BL21(DE3) cells grown in Luria-Bertani (LB) medium. For biotinylated proteins, expression plasmids were co-transformed with pSP10B-BirA. Expression was induced with 0.1mM IPTG when the culture density (Abs_{600}) reached 0.6-0.8 and proceeded at 30°C for 6hr, after which cells were harvested by centrifugation at 4,000×g for 20min at 4°C. The resulting pellets were stored at -80°C overnight. Thawed pellets were resuspended in 15mL buffer (PBS, 10mM imidazole, pH 7.4) and lysed with a high-pressure homogenizer (Avestin EmulsiFlex C5). Lysates were cleared by centrifugation at 20,000×g for 20 min at 4°C. His_6-tagged proteins were subjected to gravity Ni^{2+}-affinity

purification HisPur™ Ni-NTA Resin (Thermo, 88221) following manufacturer's protocols. Samples were desalted into buffer (PBS, pH 7.4) using PD 10 Desalting Columns (GE Healthcare, 17-0851-01) following manufacturer's protocols. Samples were stored at 4°C for up to two weeks or diluted to 25% glycerol and stored at -80°C indefinitely. For biotinylated proteins, a second purification step was performed using Pierce Monomeric Avidin Agarose (ThermoFisher, 20228) according to manufacturer's protocols followed by desalting and storage as above.

Protein analysis. SDS-PAGE and immunoblotting of proteins was performed according to standard procedures. BioRad Coomassie R-250 stain was used to visualize proteins in SDS-PAGE (BioRad, Mini-PROTEAN® TGX). The following primary antibodies were utilized for immunoblotting: rabbit anti-huntingtin (Abcam, ab451669), mouse anti-MBP-HRP (NEB, E8038), rabbit anti-ERK (Cell Signaling Technology, 9102), rabbit anti-pERK (Cell Signaling Technology, 9101), mouse anti-ubiquitin (Millipore, P4D1-A11), rabbit anti-His_6-HRP (Abcam, ab1187), rabbit anti-β-tubulin (Cell Signaling Technology, 5346), rabbit anti-Flag-HRP (Abcam, ab49763). Secondary antibodies goat anti-rabbit IgG (H+L) and anti-mouse IgG (H+L) with HRP conjugation (Promega) were utilized as needed.

Enzyme linked immunosorbent assay (ELISA). For ELISA analysis of huntingtin and MBP, a previously established protocol was modified to detect binding to huntingtin and MBP [15]. Briefly, a 96-well

EIA plate was coated overnight at 4°C with 100µL C4-uAb, scFvC4, CHIPΔTPR, or PBS (for huntingtin) or 100µL MBP at 4µg/mL in PBS. The plate was then washed three times with 100µL PBST (1×PBS + 0.1% Tween20) per well for 5min with shaking and blocked with 100µL blocking buffer (PBS + 5% milk) per well at room temp, slowly mixing for 2hr. Purified protein samples were introduced in blocking buffer as serial dilutions with 50µL per well and incubated at room temp slowly mixing for 1hr. Three PBST washes were used to remove non-bound protein before introducing 50µL of anti-Flag-huntingtin (diluted 1:1000 in PBS) followed by anti-rabbit-HRP (1:2500) (huntingtin) or anti-His_6-HRP (1:10000) (MBP) and incubating at room temp with slow mixing for 1hr. Six final PBST washes were performed before incubation with 200µL OPD (Sigma Fast tablets) in the dark for 30min. The reaction was then quenched with 50µL 3N H_2SO_4 and absorbance read at 490nm.

For ERK and pERK ELISA analysis, a previously established protocol was used [26]. Briefly, all steps until detection were performed at 4°C. Biotinylated ERK2 and pERK2 diluted in coating buffer (PBS, 1mM DTT, 0.05% Tween20) were immobilized on NeutrAvidin coated 96-well plates (Pierce) overnight at 4°C. Plates were washed twice with 200µL coating buffer and then blocked with block buffer (coating buffer + 5% BSA) for 1hr. Plates were washed three times with 200µL coating buffer, introduction of the ubiquibody or control protein diluted in coating buffer to the indicated concentrations and incubated for 1hr. Plates were

washed three times with wash buffer (PBS, 0.05% Tween20), before introducing anti-RGS-His_4 (1:5000) for 1hr, followed by anti-rabbit-HRP (1:2500) for 1hr, washing three times between and after incubation with wash buffer. Incubation with 200µL OPD (Sigma Fast tablets) was performed in the dark for 30min. The reaction was then quenched with 50µL 3N H_2SO_4 and absorbance read at 490nm.

***In vitro* ubiquitination assays.** Ubiquitination assays were performed as previously described [41] in the presence of 0.1µM purified human recombinant Ube1 (R&D Systems, E-305), 4µM human recombinant UbcH5α/Ube2D1 (R&D Systems, E2-616), 3µM uAb, 1.5µM human ERK2-GST (ThermoFisher, PV3314) or MBP (NEB) or huntingtin (provided by Alyse Portnoff), 50µM human recombinant ubiquitin (R&D Systems, U100), 4mM ATP and 1mM DTT in 20mM MOPS, 100 mM KCl, 5 mM MgCl2, pH 7.2. Reactions were carried out at 37°C for 2hr, unless otherwise indicated, and stopped by boiling in 2x Laemmli loading buffer for analysis by immunoblotting.

Pulldown Assays. Purified ubiquibodies and control proteins were conjugated to HisPur™ Ni-NTA Resin (Thermo, 88221) by incubated 300µg protein with 1mL resin slurry for 30min at 4°C with end-over-end rotation. HEK293T cell lysate was prepared by harvesting cells in PBS, pelleting at 8000×g for 5min, and freezing at -20°C. Thawed pellets were lysed in NP40 lysis buffer (150mM NaCl, 1% Nonidet P-40, 50mM TrisHCl, pH 7.4) by pipetting and mixing at 4°C for 30min. Soluble

fractions were obtained by centrifugation of lysed cells at 18,000×g at 4°C for 20min. Prepared resin was incubated with 10µL lysate at 4°C overnight. Resin was washed with buffer (PBS, 25mM imidazole, pH 7.4), and proteins were eluted with buffer (PBS, 250mM imidazole, pH7.4). Samples were boiled with 2×Laemelli loading buffer and analyzed by immunoblot.

Cell culture, transfection, and lysate preparation. HEK293T cells were obtained from ATCC and cultured in standard medium at 37°C with 5% CO_2. HEK293T cells were cultured in DMEM with 10% heat inactivated FBS and 1% antibiotic-antimycotic (Cellgro). Cells were transfected in 6-well dishes at 60-80% confluency with 2µg total plasmid DNA using empty pcDNA3 plasmid to balance all transfections. jetPRIME® transfection (Polyplus Transfection) was performed according to manufacturer's instructions using a 1:2 ratio of jetPRIME® to DNA (w:v) and at 4hr post-transfection the growth media was refreshed. At 24hr post-transfection, cells were harvested by scraping in PBS and frozen at -20°C until analyzed by immunoblotting. Thawed pellets were lysed in NP40 lysis buffer (150mM NaCl, 1% Nonidet P-40 and 50mM TrisHCl, pH 7.4) by pipetting and mixing at 4°C for 30min. Soluble fractions were obtained by centrifugation of lysed cells at 18,000xg at 4°C for 20min. Samples were boiled in 2x Laemmli sample buffer for analysis by immunoblotting and were normalized using a detergent compatible total protein assay (Bio-Rad) such that 10µg total protein was loaded.

Mass spectrometry analysis. In preparation, ubiquitination assays were performed as above. Reactions were resolved by SDS-PAGE and stained with Coomassie R250 prior to gel excision. The products of each reaction were submitted to the Cornell Biotechnology Resource Center Services Proteomics and Mass Spectrometry Facility for detection of ubiquitination events. Briefly, the protein bands were cut from an SDS-PAGE gel and cut into ~1mm cubes. The gel bands were washed in 200µL DI water for 5min, followed by 200µL 100mM ammonium bicarbonate (ambic)/acetonitrile (ACN) (1:1) for 10min and finally 200µL ACN for 5min. The acetonitrile was discarded and the gel bands were dried in a speed-vac for 10min. The gel pieces were rehydrated with 70µL of 10mM DTT in 100mM ambic and incubated for 1hr at 56°C. The samples were allowed to cool to room temperature, after which 100µL of 55mM iodoacetamide in 100mM ambic was added and the samples were incubated at room temp in the dark for 60min. Following incubation, the gel slices were again washed as described above. The gel slices were dried and rehydrated with 50µL trypsin at 50 ng/µL in 45mM ambic, 10% ACN on ice for 30 min. The gel pieces were covered with an additional 25µL of 45mM ambic, 10% ACN and incubated at 37°C for 19hr. The digested peptides were extracted twice with 70µl of 50% ACN, 5% formic acid (FA) (vortexed 30min, sonicated 10min) and once with 70µl of 90% ACN, 5% FA. Extracts from each sample were combined and lyophilized. The lyophilized in-gel tryptic digest samples were reconstituted in 20µL of

nanopure water with 0.5% FA for nanoLC-ESI-MS/MS analysis, which was carried out by a LTQ-Orbitrap Velos mass spectrometer (Thermo-Fisher Scientific) equipped with a CorConneX nano ion source device (CorSolutions LLC). The Orbitrap was interfaced with a nano HPLC carried out by an UltiMate3000 UPLC system (Dionex). The gel extracted peptide samples (2-4 µL) were injected onto a PepMap C18 trap column-nano Viper (5µm, 100µm × 2cm, Thermo Dionex) at 20µL/min flow rate for on-line desalting and then separated on a PepMap C18 RP nano column (3µm, 75µm x 15cm, Thermo Dionex) which was installed in the "Plug and Play" device with a 10-µm spray at Cornell University Library on May 30, 2016 http://www.jbc.org/ Downloaded from Selective Protein Knockout Using Engineered Ubiquibodies 5 emitter (NewObjective). The peptides were then eluted with a 90min gradient of 5% to 38% ACN in 0.1% FA at a flow rate of 300nL/min. The Orbitrap Velos was operated in positive ion mode with nano spray voltage set at 1.5kV and source temperature at 275°C. Internal calibration was performed with the background ion signal at m/z 445.120025 as the lock mass. The instrument was operated in parallel data-dependent acquisition (DDA) mode using FT mass analyzer for one survey MS scan for precursor ions followed by MS/MS scans on top 7 highest intensity peaks with multiple charged ions above a threshold ion count of 7500 in both LTQ mass analyzer and HCD-based FT mass analyzer at 7,500 resolution. Dynamic exclusion parameters were set at repeat count 1 with a 15sec repeat

duration, exclusion list size of 500, 30sec exclusion duration, and ±10 ppm exclusion mass width. HCD parameters were set at the following values: isolation width 2.0m/z, normalized collision energy 35%, activation Q at 0.25, and activation time 0.1msec. All data were acquired using Xcalibur 2.1 operation software (Thermo-Fisher Scientific). All MS and MS/MS raw spectra were processed and searched using Proteome Discoverer 1.3 (PD1.3, Thermo-Fisher Scientific) against databases downloaded from NCBI-nr database. The database search was performed with two-missed cleavage site by trypsin allowed. The peptide tolerance was set to 10 ppm and MS/MS tolerance was set to 0.8Da for CID and 0.05Da for HCD. A fixed carbamidomethyl modification of cysteine, variable modifications on methionine oxidation, and ubiquitin modification of lysine were set. The peptides with low confidence score (with Xcorr score <2 for doubly charged ion and <2.7 for triply-charged ion) defined by PD1.3 were filtered out and the remaining peptides were considered for the peptide identification with possible ubiquitination determinations. All MS/MS spectra for possibly identified ubiquitination peptides from initial database searching were manually inspected and validated using both PD1.3 and Xcalibur 2.1 software.

Acknowledgements

We thank Marc Ostermeir, Cam Patterson, Shohei Koide, Andreas Plückthun, and Amy Karlsson for generously providing plasmids

encoding the genes in this study. I thank Alyse Portnoff for invaluable training in biochemical and molecular biology techniques, as well as superb mentorship in project design and execution. I thank Anthony Bretscher (Cornell University) for training in cell culture and transfection, Bunyarit Meksiriporn for initial training in surface plasmon resonance, and Morgan Baltz for in depth conversations on ubiquibody theory. This material is based upon work was supported by the National Science Foundation Career Award CBET-0449080 (to M.P.D.), National Institutes of Health Grant CA132223A (to M.P.D.), New York State Office of Science, Technology and Academic Research Distinguished Faculty Award (to M.P.D.), National Science Foundation GRFP DGE-1144153 (to E.A.S.), the Cornell Presidential Life Science Fellowship Program (to E.A.S.).

CHAPTER 3

THE E3 CATALYTIC DOMAIN

Introduction

The second requirement of a ubiquibody is E3 catalytic activity.

E3 catalysis results in transfer of a charged ubiquitin to an acceptor lysine on a substrate protein. Two distinct mechanisms for this transfer have been identified, though an intermediate mechanism can be found in a small subfamily of E3s [57]. In order to design synthetic enzymes with the same catalytic capability as natural E3s, we have focused on hijacking existing catalytic domains that behave modularly within the enzyme as a whole.

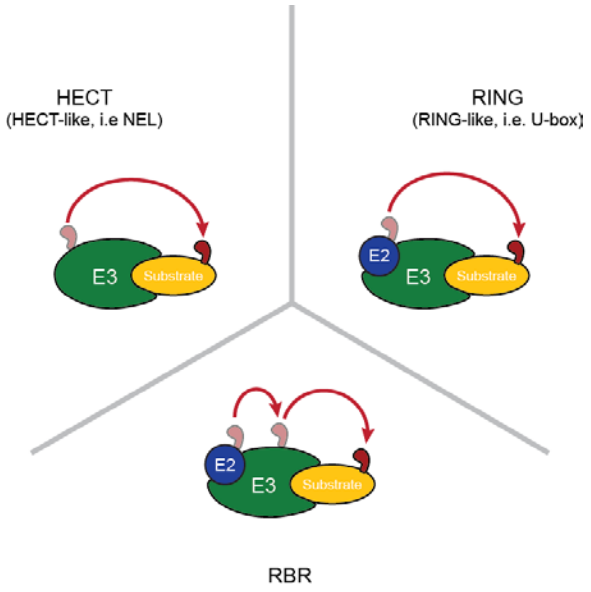


Figure 3.1: Types of E3 ubiquitin ligase.
E3 ubiquitin ligases can be sorted based on their catalytic mechanism, which usually stems from a conserved characteristic domain (i.e. HECT, RING, or RBR).

Classification of E3s based catalytic mechanism results in three families of enzyme: HECT, RING, and RBR (**Figure 3.1**).

The homologous to E6AP-C terminus (HECT) family is one of the smaller families of E3 ligase in the human genome, with a respectable 28 unique members. HECT E3s are characterized by incredibly diverse N-terminal regions, which for the 28 enzymes are divided into three subfamilies, and highly conserved C-terminal domains, for which the family is named [58, 59]. The C-terminal HECT domain is composed of two subdomains, N-lobe and C-lobe, that coordinate the transient conjugation of ubiquitin to a catalytic cysteine residue in the C-lobe [60]. The HECT domain—if not the entire E3—must interact with E2 and target sequentially, due to competition for access to the active cysteine site [61]. Unlike other E3 families where the E2 plays a critical role in product formation, HECT E3s themselves are responsible for the selection of acceptor lysine and subsequent poly-ubiquitin chain linkages [62].

The largest superfamily of E3s, with upwards of 600 known complexes in humans, is the really interesting new gene (RING) and RING-like E3s [57]. These ligases are often multi-subunit, large complexes with a variety of adaptors. Unlike HECT E3s, RING E3s do not contain a catalytic cysteine, and instead bring substrate proteins into direct proximity of the charged ubiquitin on the E2 [63]. This means that both substrate and charged E2 must be simultaneously bound for successful catalysis. While many RING E3s are multi-subunit, such as

Cullin-RING ligases or the anaphase promoting complex, a small sub-family of RING-like E3s, the U-box E3s, function as monomers or homodimers. The first generation of ubiquibodies was made using U-box E3 catalytic domains [15]; the features and functional parameters of this domain will comprise the majority of this report.

A final family of E3 ubiquitin ligase, RING-between-RING (RBR) E3s, have a catalytic mechanism somewhere between HECT and RING E3s [57]. The characteristic structure of an RBR E3 is two RING domains (R1 and R2) separated by an inbetween ring (IBR) domain [64]. Like HECT E3s, RBR E3s transfer ubiquitin to substrate proteins via a two-step mechanism, transiently conjugating the charged ubiquitin to a catalytic cysteine in the R2 domain, but like RING E3s, require zinc coordination to create a platform for E2 binding. Like HECT E3s and U-box E3s, RBR E3s tend to function as monomers or homodimers [57].

The monomeric nature of HECT, U-box, and RBR E3s, combined with well-defined consensus catalytic domains, make these types of E3 attractive for synthetic E3 design. While HECT and RBR E3s would allow stricter control over ubiquitination homogeneity, due to the lack of E2-independent poly-ubiquitin chain formation, U-box E3s offer the opportunity for tunable poly-ubiquitination, determined by the subset of E2 enzyme interactions.

The original purpose of ubiquibodies was proteasome-mediated silencing of selectable target proteins. This required ubiquibodies derived

from a U-box E3 to form catalytically active complexes by binding both E2 and target proteins, and to abide by the ubiquitin code [42, 43], forming lysine 11 (K11) and/or lysine 48 (K48) poly-ubiquitin linkages—the canonical linkages recognized by the proteasome. In this work, investigation into oligomeric state of ubiquibodies in solution has helped elucidate the important features required for productive catalytic activity. Because chain specificity is derived from the E2 in U-box catalyzed ubiquitination, the subset of human E2 interacting partners of U-box-based ubiquibodies has been determined. And further investigation into the types of poly-ubiquitin chain linkages has confirmed that ubiquibodies favor proteasome-targeting linkages in the presence of one of their cognate E2s. Finally, alternatives to the U-box scaffold have been designed and await testing.

Results

Rational design of a U-box-based ubiquibody.

Most U-box E3s follow a similar organization, with an N-terminal substrate binding domain, followed by a linker domain of varying purpose, and ending with the catalytic U-box domain [65]. Ubiquibodies as synthetic U-box-containing E3s, derive their catalytic activity from the C-terminus of carboxy-Hsc70 interacting protein (CHIP) (**Figure 3.2**). CHIP is a naturally occurring human E3, that functions as a homodimer to mediate the proteolysis of heat-shock and misfolded proteins [66, 67].

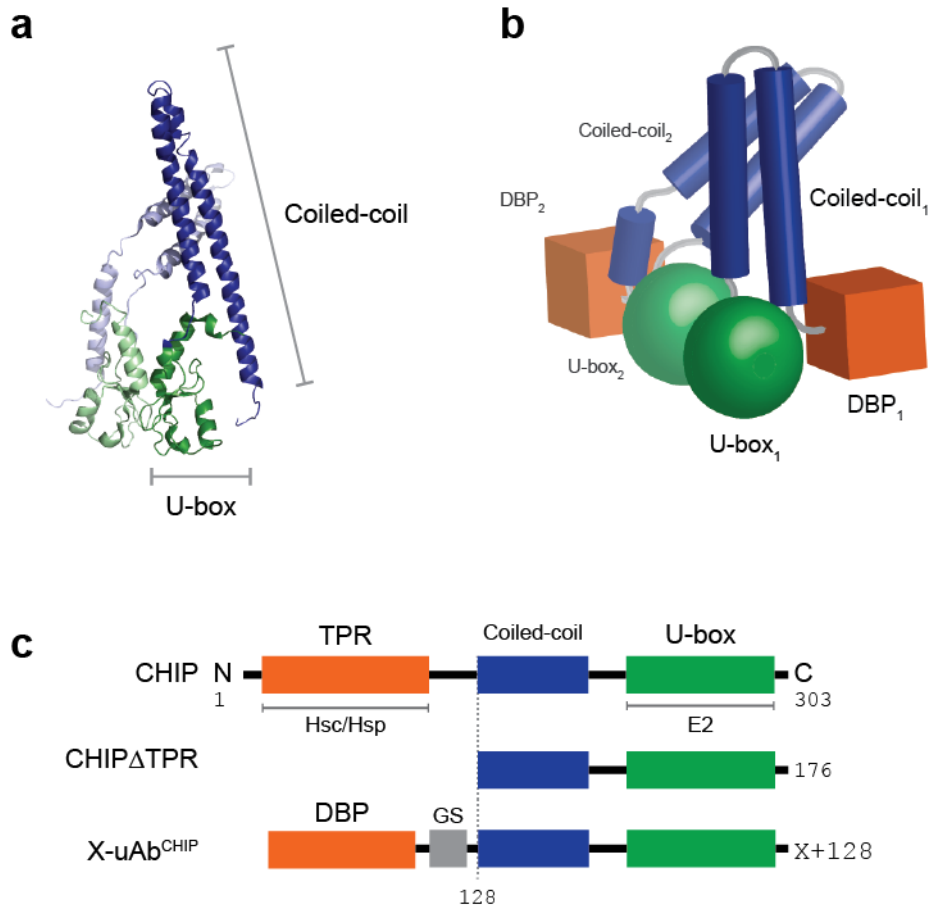


Figure 3.2: Rational design of U-box driven ubiquibodies.
(a) Crystal structure of the C-terminal domains of dimerized CHIP. Coiled-coil domains (blue) adopts different conformations in each monomer; U-box domains (green) are positioned near to each other. PDB ID: 2CL2 [66]. Schematic was generated using PyMol software. (b) Cartoon representation of dimerized ubiquibodies. DPS (orange) orient asymmetrically due to asymmetry in the coiled-coil domains (blue), resulting in blockage of one of the E2 binding sites in the U-box domains (green). Adapted from Qian et al. [41]; schematic was generated using Illustrator software. (c) Cartoon of the organization of wild type CHIP and CHIP-derived ubiquibodies. Adapted from Portnoff et al. [15].

The dimerization of the CHIP, via its coiled-coil domain, is essential for U-box function [41]. Ubiquitination occurs when one subunit of the CHIP homodimer interacts with a charged E2 ubiquitin conjugating enzyme, via its U-box domain, and a substrate protein via the TPR domain of the second subunit (**Figure 3.2b**). The CHIP homodimer is

asymmetric, such that only one E2 binding site is accessible at a time, due to blocking of the second site by the second TPR domain as a result of a secondary conformation of the coiled-coil region. Because the TPR domain is known to be important for the U-box catalytic mechanism, ubiquibodies that derive their catalytic activity from CHIP's U-box while replacing the TPR domain with a DBP must be able to either mimic or work around this requirement (see Chapter 2).

Ubiquibodies were designed by fusion of DBPs to the C-terminus of CHIP (CHIPΔTPR) via a short, flexible linker (**Figure 3.2c**). The design included the coiled-coil domain, and thus the predicted mechanism of ubiquibody catalysis would also depend on dimerization. To determine the oligomeric state adopted by ubiquibodies, size exclusion chromatography (SEC) was performed (**Figure 3.3a**). Wild type CHIP eluted at volumes expected of a trimer, or a dimer with a large water shell; CHIP was prone to aggregation even at ten-fold lower concentrations, which may correspond to its native ability to recognize and bind misfolded proteins. Ubiquibodies eluted homogenously at volumes expected of a tetrameric complex, suggesting the soluble forms could be comprised of either a tetramer or dimer of dimers. Unlike CHIP, ubiquibodies did not show the same tendency towards aggregation, indicating that this property is likely caused by native TPR activity. Comparatively, MBP, a known small monomeric protein, eluted at volumes expected of a monomer (**Table 3.1**). These results suggest that

ubiquibodies multimerize, forming dimers or larger complexes. To further investigate the quaternary structure of ubiquibodies in solution, blue native poly-acrylamide gel electrophoresis (BN-PAGE) was performed using purified ubiquibodies (**Figure 3.3b**). The unfused DBP, DARPin EpE89, ran at molecular weights corresponding to theoretical monomers and dimers, though detection of proteins of this size was near the lower limit of this assay. Conversely, CHIPΔTPR ran at multiple molecular weights higher than expected for either a monomer or dimer, indicating that this domain alone is sufficient for oligomerization. Bands of varying size were detected for CHIPΔTPR, suggesting that this protein may also be prone to aggregation. Full length ubiquibodies ran at the expected molecular weight of a dimer, and did not show the same tendency for multiple oligermized states. Taken together, these data suggest that ubiquibodies dimerize in solution, similar to wild-type CHIP. The C-terminal CHIPΔTPR region is responsible for this oligomerization, and potentially facilitates the formation of larger complexes.

Table 3.4: Oligomeric state of ubiquibodies

	Theoretical Monomeric MW (kDa)	Elution Vol. (mL)	Experimental MW (kDa)	Max. Oligomeric State
CHIP	37.13	12.03	136.73	3.7
E40-uAb	40.56	12.27	194.92	4.8
pE59-uAb	36.98	12.45	179.02	4.8
EpE89-uAb	37.01	12.37	185.92	5.0
MBP	43.39	14.80	58.93	1.4

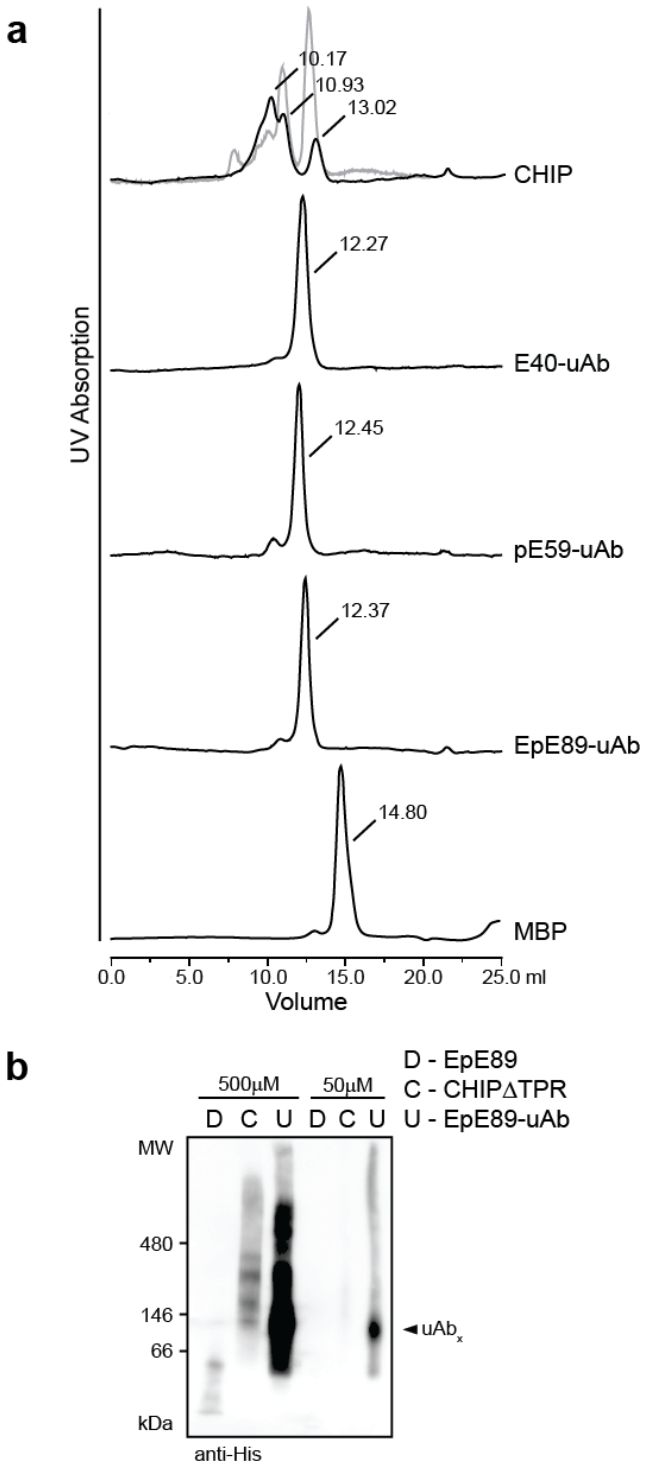


Figure 3.3: Oligomerization of U-box ubiquibodies.
(a) Size exclusion chromatography (SEC). Purified proteins were analyzed by SEC using a HiLoad Superdex 200 column. Elution profiles of three ubiquibodies are compared relative to wild type CHIP, which aggregated at high concentration, and MBP. (b) BN-PAGE of ubiquibodies. Immunoblot of purified ubiquibody and control proteins following BN-PAGE detected with the indicated antibody. [51]

In the course of this project, crystallization of the R4-uAb, OFF7-uAb, and EpE89-uAb was attempted, though no diffractable crystals formed under any of the tested conditions.

Functional characterization of U-box ubiquibodies

To demonstrate that U-box fusion results in functional synthetic E3 ubiquitin ligases, *in vitro* ubiquitination assays were performed (**Figure 3.4**). Incubation with components of the UPP (E1, E2, ubiquitin, and ATP) and EpE89-uAb resulted in the formation of higher molecular weight species of ERK2-GST. The appearance of these species correlated with the appearance of high molecular weight ubiquitin species, indicating that EpE89-uAb ubiquitinated ERK2-GST in a time dependent manner (**Figure 3.4a**). To confirm that each component of the pathway was necessary for ubiquitination of ERK2-GST, assays lacking one of each of the UPP proteins (or ATP) were performed (**Figure 3.4b**). Only when all components, including the full EpE89-uAb, were present was ubiquitination of ERK2-GST observed. Neither the binding domain, EpE89, nor catalytic domain, CHIPΔTPR, alone resulted in ubiquitination of ERK2-GST. Importantly, the presence of a catalytic domain, regardless of DPB presence or specificity, led to the formation of high molecular weight species of ubiquitin; similar molecular weight species were observed in detection of the ubiquibody, indicating that auto-ubiquitination is a prominent side reaction in this system. Auto-ubiquitination was significantly reduced in the presence of a cognate

target, indicating that the two reactions are in competition. Overall, the U-box domain has been confirmed to be an effective supplier of E3 catalytic activity to ubiquibodies against non-native targets.

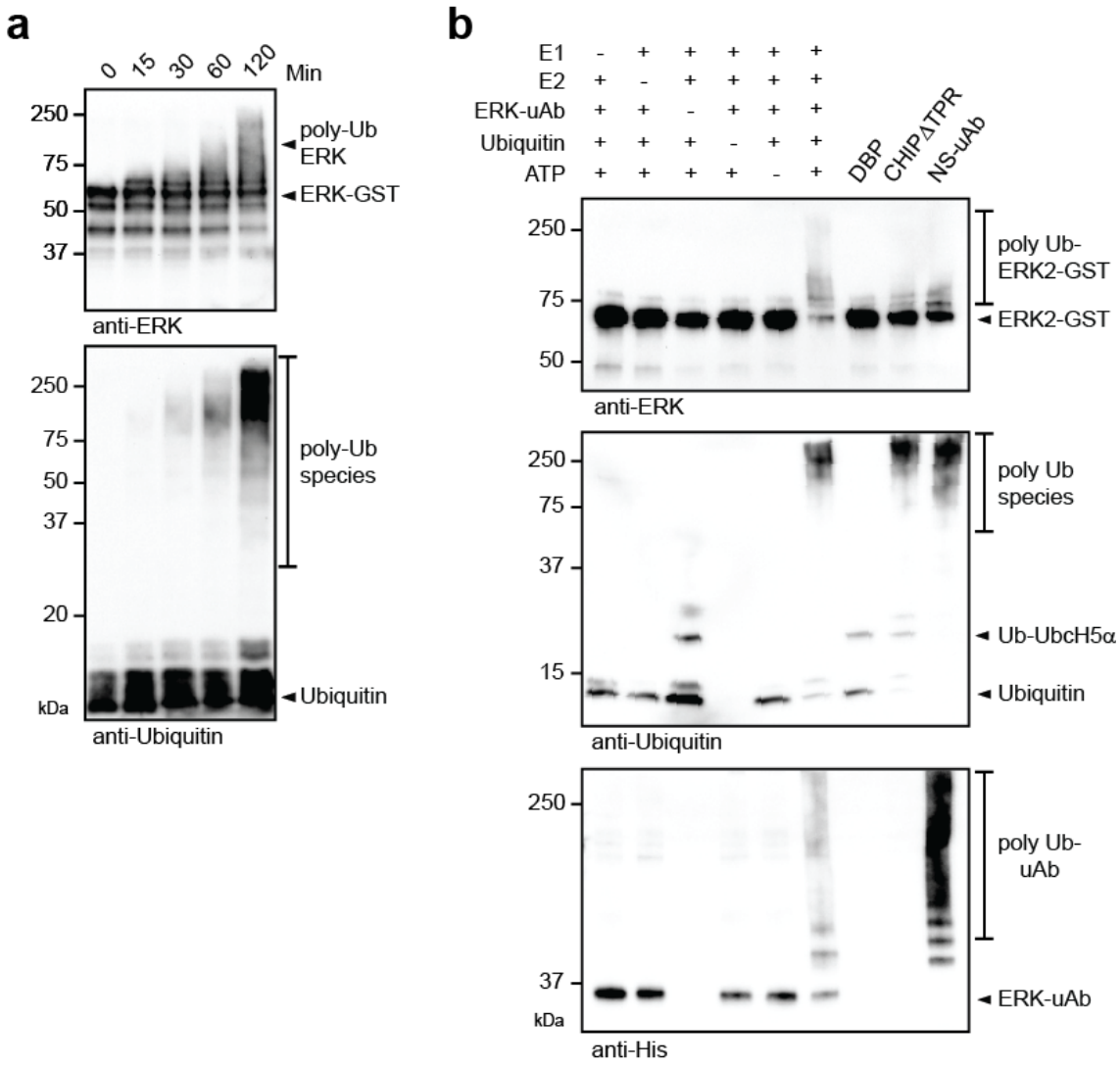


Figure 3.4: *In vitro* ubiquitination assay with U-box-containing ubiquibodies. (a) Time dependent ubiquitination of ERK2-GST. At the indicated time points, samples were removed and analyzed by immunoblot with the indicated antibodies. (b) UPP component necessity for catalytic activity of U-box-containing ubiquibodies. Reactions with the indicated components were incubated for 2hr, then analyzed by immunoblot. [51]

In order to determine the type of ubiquitin linkages formed by ubiquibodies, the above reactions were tested for presence of K48, K63 and K27 ubiquitin-ubiquitin linkages (**Figure 3.5a**). (Note that antibodies for the detection of K6, K11, K29, and K33 linkages were unavailable at the time of this project.) In the presence of a U-box catalytic domain, high molecular weight species of K48 and K63 linkages were detected, while no such K27 species were observed. While comparative levels of K48 linkages were observed regardless of the presence of target protein, the level K63 linkages decreased when auto-ubiquitination was the only catalyzable reaction. This suggests that K63 linkage formation is benefitted by the presence of target and that this type of poly-ubiquitination is more likely to be found on target substrates. Ubiquibody and target would be expected have similar levels of K48-linked ubiquitination.

All ubiquitination reactions reported up to this point were performed with the E2 enzyme family member Ube2D1 (a.k.a. UbcH5A), which was previously shown to interact with wild type CHIP [68]. To further characterize the poly-ubiquitination profile of ubiquibodies in the presence of Ube2D1, reaction products (60-250kDa) were analyzed by mass spectrometry. From samples containing EpE89-uAb and Ube2D1, ubiquitination on four ubiquitin lysine residues, K6, K11, K48, and K63, were found (**Figure 3.5b** and **c**). The relative abundance of these linkages was estimated across six independent experiments, three with ERK2 as

the target protein, and three with pERK2 as the target protein. K48 and K63 linkages were the most commonly found, while K6 linkages were least often found. The majority of ubiquibody poly-ubiquitination consisted of proteasome-specific linkages (K11 and K48).

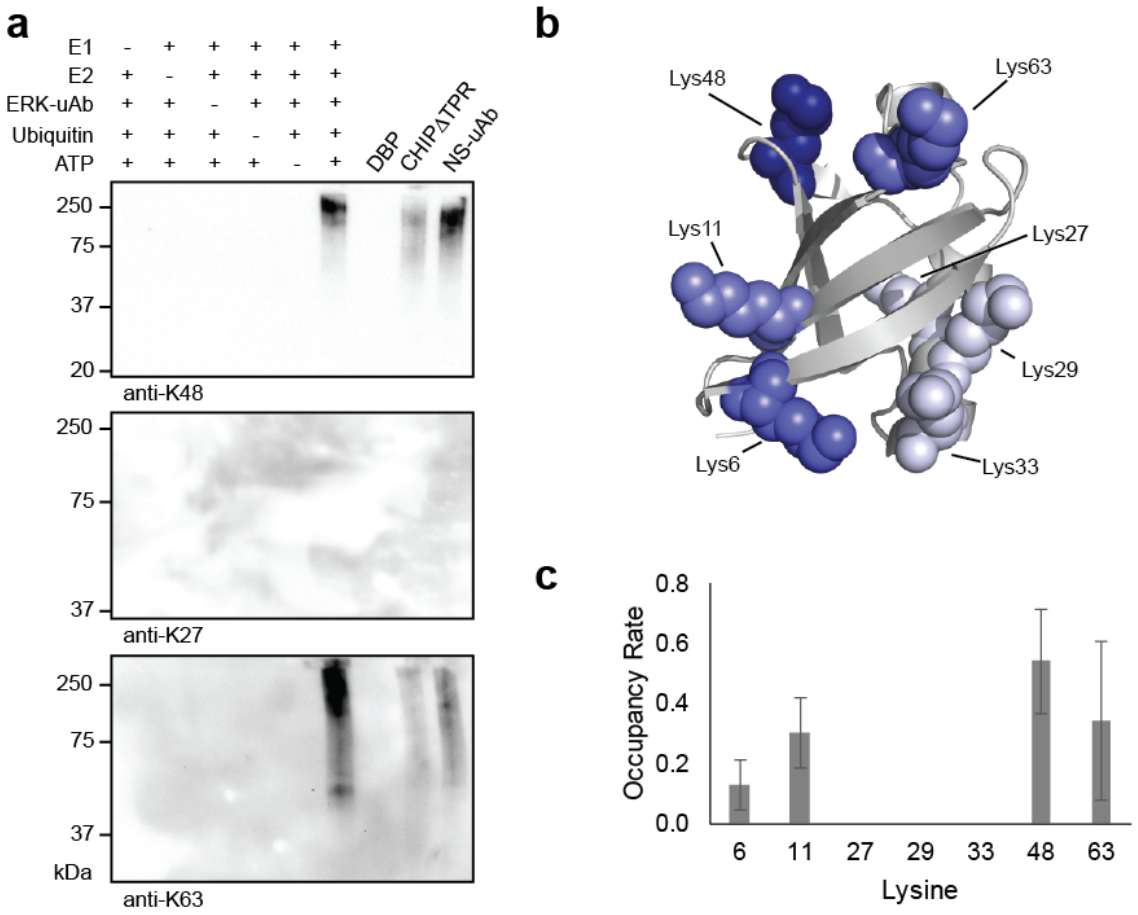


Figure 3.5: Poly-ubiquitin linkage profile of EpE89-uAb with UbcH5a.
(a) Formation of poly-ubiquitin chains by ubiquibodies. Reactions were halted after 2hrs and analyzed by immunoblot with the indicated antibodies. (b) Cartoon representation of lysine linkages mapped to the ubiquitin crystal structure. Ubiquitin (white ribbon) contains seven lysine residues (spheres); the relative ubiquitination levels of each lysine by the EpE89-uAb are represented in shadings of blue. PDB ID: 1UBQ [68]. Schematic was generated using PyMol Software. (c) Occupancy rate of –GG modification of ubiquitin by EpE89-uAb. Bars were generated by normalizing modified residue coverage with total residue coverage and averaging across six experiments, three with ERK2 and three with pERK2. [51]

To ensure that this ubiquitination profile was not a result of DBP or linker effects, similar experiments were performed while varying either the DBP or the linker (**Figure 3.6**). When the EpE89 was exchanged for E40 or pE59, no significant differences in the formation of poly-ubiquitination K6, K11, K48, or K63 linkages were observed. However, both E40-uAb and pE59-uAb demonstrated the ability to catalyze K27 linkages, abet at extremely low levels. Thus, DARPin DBPs do not appear to greatly affect the types of poly-ubiquitin chains formed by the U-box. When the 5AA linker was exchanged for one of four longer linkers (see Chapter 4), a significant decrease in K63 poly-ubiquitin linkages was observed, though the relative levels of K6, K11, and K48 linkages remained constant. No ubiquitination of novel residues was observed by ubiquibodies of varying linker composition. This suggests that linker length and flexibility plays at least a minimal role in the determination of lysine linkages, with the possibility of reducing non-proteasomal K63 linkages with varying linker composition. Overall however, the primary determination of poly-ubiquitination appears to be linked to the U-box/E2 combination, and remains mostly undisturbed by changes to other domains of the ubiquibody.

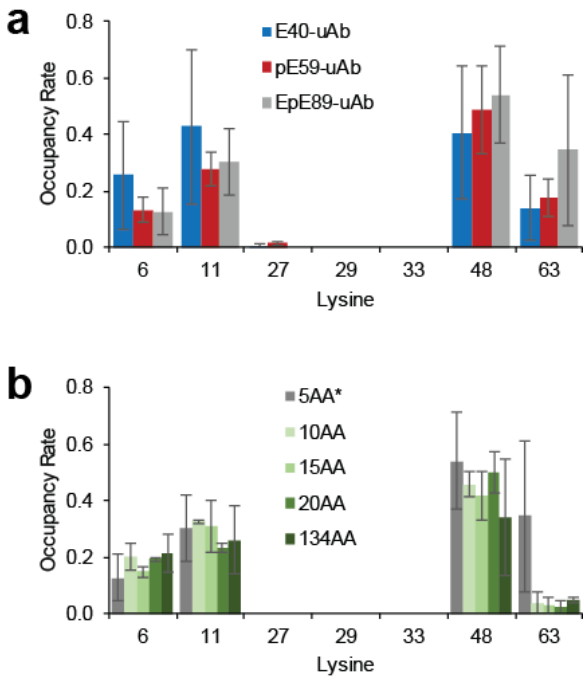


Figure 3.6: Effect of DBP and linker domains on poly-ubiquitin chain type. (a) Occupancy rate of ubiquitination on ubiquitin lysines by different ubiquibodies. (b) Occupancy rate of ubiquitination on ubiquitin lysines by ubiquibodies of varied linker length. Bars were generated by normalizing modified residue coverage with total residue coverage and averaging across (a) six independent experiments, three with ERK2 as the target and three with pERK2 as the target, or (b) across two independent experiments, with ERK2 as the target. [51]

To determine which other E2s ubiquibodies might interact with in cells, *in vitro* ubiquitination assays using a panel of human E2s (Ubiquigent) were performed. Forty-one E2s are encoded in the human genome [69]; a subset of thirty-two human E2s were used to screen the EpE89-uAb for catalytic activity (**Appendix B**). The ubiquibody catalyzed the formation of high molecular weight ubiquitin species with twenty-two of these E2s (**Figure 3.7** and **Table 3.2**), and of these, only nine corresponded to shifts of ERK2-GST to higher molecular weights, indicating that target ubiquitination by U-box E3s is mediated by a small subset of the E2s available in mammalian cells. Ube2D1, Ube2D2, and

Ube2D3 (all members of the UbcH5 E2 family), were the only E2s which catalyzed both ubiquitination of ERK2-GST and K48 poly-ubiquitin chains, indicating that proteasomal targeting by ubiquibodies in cells would likely be mediated by some combination of these E2s. (Other E2s that preferentially form K11 linkages could also contribute to proteasomal targeting.) Notably, many of the E2s that resulted in poly-ubiquitin species did not correspond to either ERK2-GST ubiquitination or auto-ubiquitination. This suggests that U-box fusions are capable of forming free ubiquitin chains, in some cases dozens of ubiquitin residues long. A subset of E2s, Ube2E2, Ube2E3, Ube2F and Ube2W, catalyzed mono-ubiquitination of ERK2-GST. In cells, this type of event often acts as a primer for downstream poly-ubiquitination by other E2s that are capable of ubiquitin chain extension [70, 71]. To fully elucidate the E2s which could function as ERK-Ub chain extenders, reactions containing mono-ubiquitin primed ERK should be performed. The only instance of K27 formation was with Ube2E1, but did not correspond to modification of either the target or ubiquibody. Should other linkage-specific antibodies become available, it will be interesting to determine the full profile of poly-ubiquitin chains formed by ubiquibodies against most of the E2 conjugating enzymes found in humans.

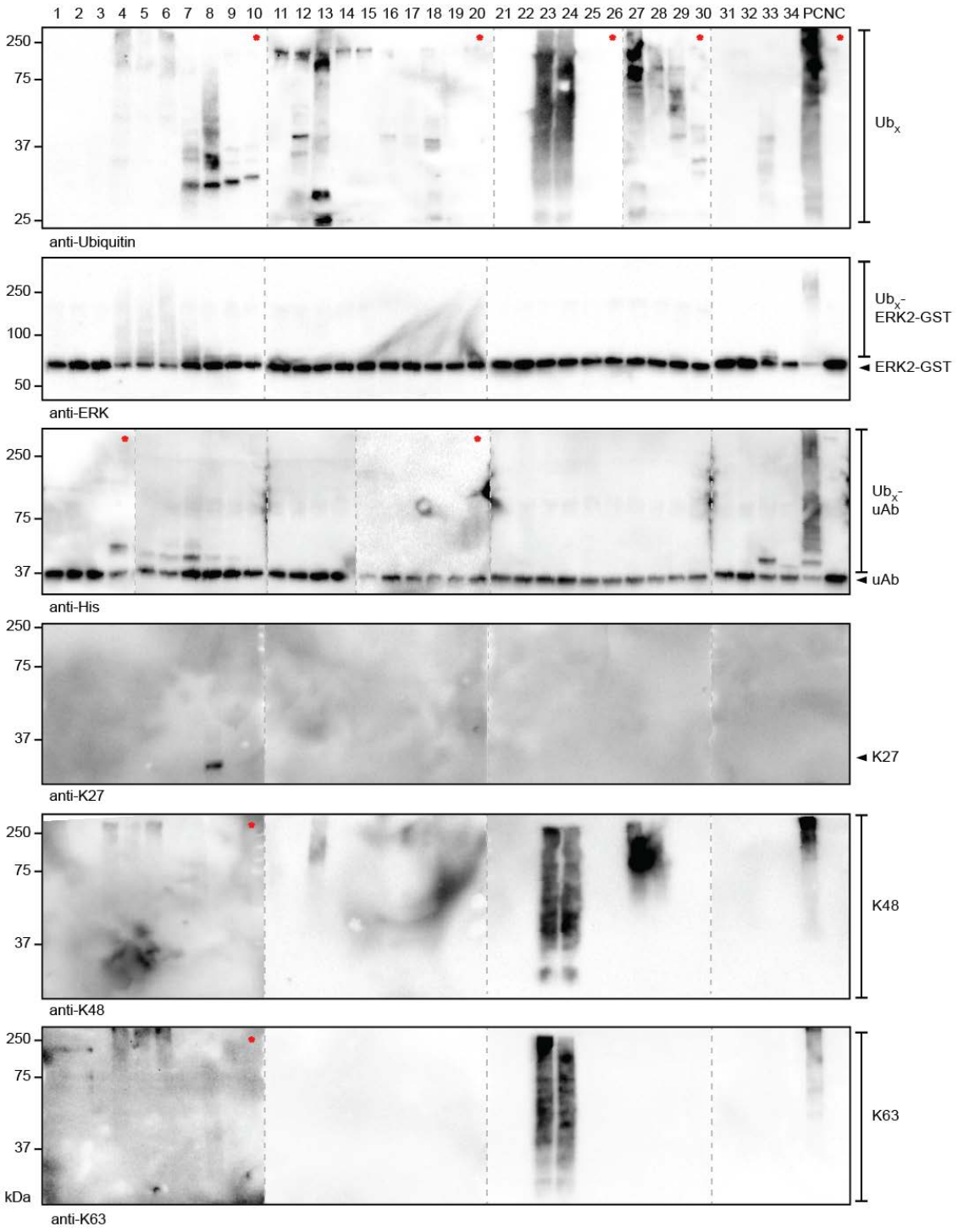
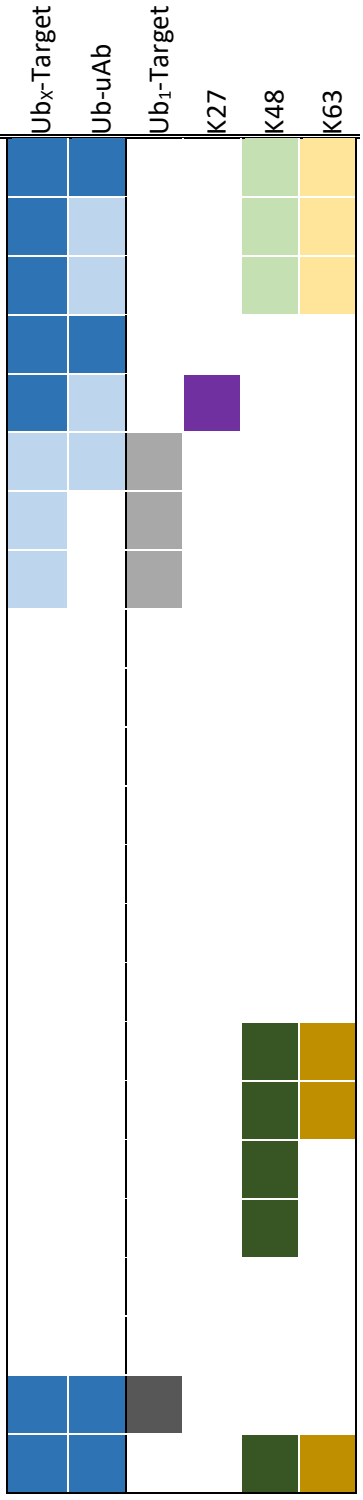


Figure 3.7: E2 scan of EpE89-uAb ubiquitination activity.
Immunoblot analysis of *in vitro* ubiquitination activity of the EpE89-uAb in the presence of different E2s (full list found in Appendix B) with the indicated antibodies. *Indicates different exposure of blots. [51]

Table 3.5: E2 interacting partners of EpE89-uAb.

Lane	E2	Alternate Name
4	Ube2D1	UbcH5A
5	Ube2D2	UbcH5B
6	Ube2D3	UbcH5C
7	Ube2D4	UbcH5D
8	Ube2E1	UbcH6
9	Ube2E2	UbcH8
10	Ube2E3	UbcH9
11	Ube2F	NCE2
12	Ube2G1	Ubc7
13	Ube2G2	Ubc7
14	Ube2H	UbcH2
15	Ube2I	Ubc9
16	6His-Ube2J1	NCUBE1
17	Ube2J2	NCUBE2
18	Ube2K	Ubc1
23	Ube2N/Ube2V1	Ubc13/Uev1A
24	Ube2N/Ube2V2	Ubc13/Mms2
27	Ube2R1	CDC34
28	Ube2R2	CDC34B
29	Ube2S	E2-EPF
30	Ube2T	HSPC150
33	6His-Ube2W	Ubc16
PC	Ube2D1	UbcH5A



As a final curiosity concerning the E2:uAb:target active catalytic complex, a brief investigation into the kinetics of the reaction was made (**Figure 3.8**). Simultaneously increasing E2 concentration and decreasing target concentration shifted reactions towards more extensive ubiquitination of ERK2-GST, as seen by the reduction of the unmodified ERK2-GST bands and increase of poly-ubiquitin conjugate smears in lanes 2 and 4. At standard reaction concentrations, ubiquitin is in great excess to the other reaction components; however, a two-fold reduction of available ubiquitin drastically reduced both ERK2-GST ubiquitination and ubiquitination in general. The effect of decreased ERK2-GST ubiquitination was ameliorated by the above changes to E2 and ERK2-GST concentrations, indicating that the reduction of poly-ubiquitin species in lane 3 was not the result of reducing ubiquitin concentrations to the level of limiting reagent, but rather, more complex interactions of between charged E2, target, and ubiquibody. Reduction of ubiquibody concentration similarly reduced global ubiquitination levels. These results suggest that ubiquitination is a highly sensitive to all components found in the catalytically active complex. (Because the ATP consumption of this reaction occurs several steps before transfer of ubiquitin to target lysine, common methods for determining rates of reaction based on ATP-depletion are not applicable.) Further investigations into the kinetics of ubiquitination could benefit from empirically trained computational

modeling, and should eventually be tuned to account for the concentrations of each of these reactants *in vivo*.

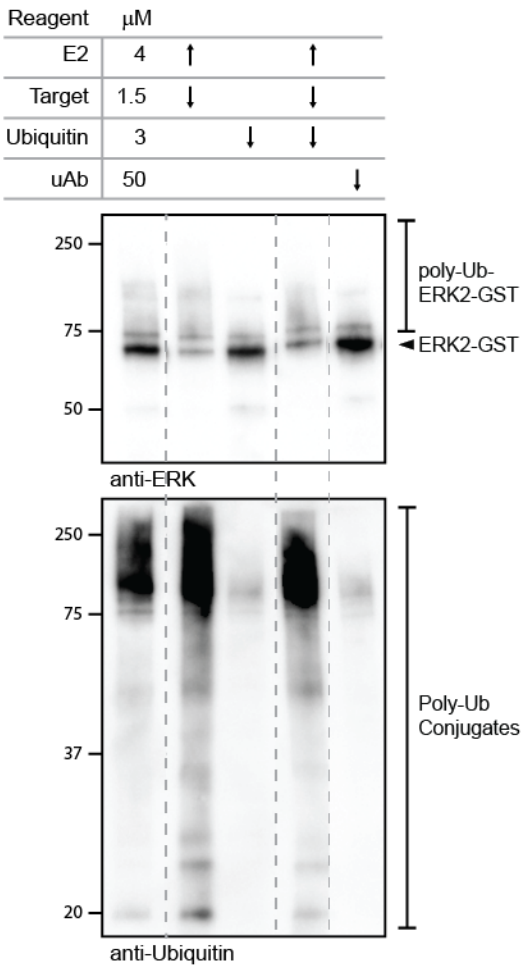


Figure 3.8: Effects of E2 and target on ubiquibody-mediated ubiquitination. Immunoblots of *in vitro* ubiquitination assays halted after 2hr, varying the relative concentrations of E2, target, ubiquitin, and uAb. The base concentrations for each reaction are shown above lane 1. In subsequent lanes an up-arrow represents a two-fold increase in component concentration and a down-arrow, a two-fold decrease in component concentration.

Rational design of alternative E3 Scaffolds

The first generation, as well as the ubiquibodies reported here, have relied on a single catalytic domain type. However, the wide diversity

of naturally occurring E3s suggests potential for both optimization and diversification of the ubiquibody scaffold. Towards this end, rational design of two different E3 ubiquitin ligase catalytic domains for incorporation into ubiquibodies are reported here. Test constructs of these ubiquibodies, using ERK DARPins as the DBPs, have been generated and await preliminary testing.

HECT and HECT-like E3s remove the variability of E2-dependent ubiquitin linkages and are therefore attractive as simplified silencing agents. One such ligase, Smurf2, is a 747 amino acid protein, containing an N-terminal membrane interacting domain and the canonical HECT C-terminal domain starting at residue 393 [72]; the full sequence, with overlaid domains, is shown in **Appendix A.7**. The HECT domain alone has been shown to be sufficient for thioester conjugation [73]; thus, in designing Smurf2 based ubiquibodies (uAb^{Smurf2}), residues Y368-E478 (Smurf2HECT) were deemed essential. Because flexibility has also been shown to be critical for HECT E3 activity, the unstructured linker between the last WW domain and the HECT domain was included, resulting in a 356 amino acid HECT catalytic domain (Smurf2CTD) (**Figure 3.9**)

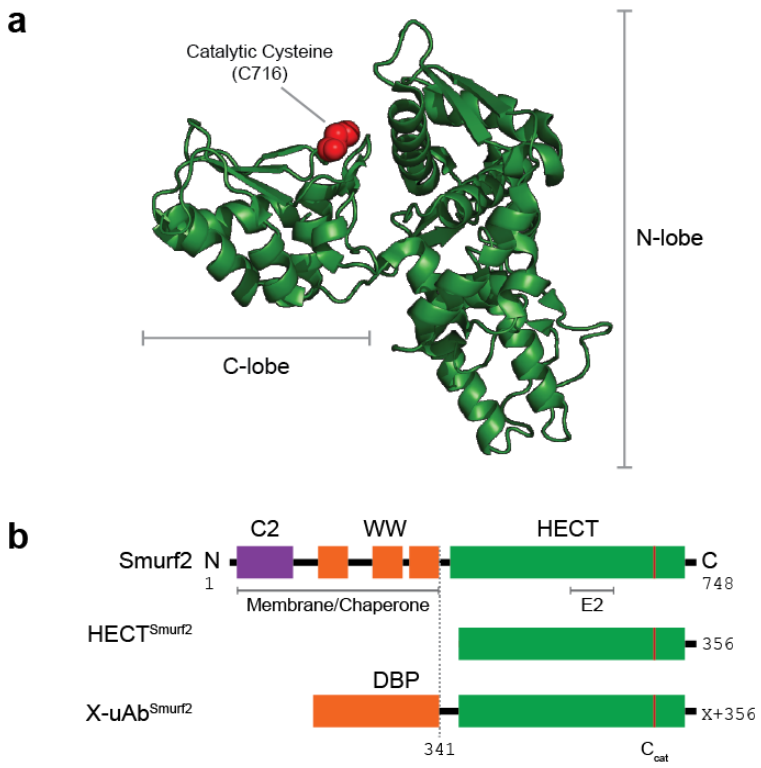


Figure 3.9: Rational design of HECT-based ubiquibodies.
(a) Crystal structure of the HECT domain of human Smurf2. Smurf2 (green) in its T-shape conformation; catalytic cysteine (red) is surface exposed. PDB ID: 1ZVD [74]. (b) Cartoon map of domain organization in wild type Smurf2 and uAb^{Smurf2}.

Alternatively, RBR E3 ubiquitin ligases exhibit similarities to both HECT and RING ubiquitin ligases, making them interesting sources of E3 catalytic activity. Auto-regulatory mechanisms for this family are of particular interest for potential RBR-ubiquibodies.

Parkin, one of the most famous RBR E3s because of its prominent role in Parkinson's disease, is structurally well-characterized, containing Ubl and R0 domains, in addition to the canonical RBR catalytic core [75]. Parkin exhibits both HECT-like and RING-like catalytic activities, dependent on the R1, IBR and R2 domains. The Ubl domain is important for proteome-co-localization of Parkin, and the R0 domain is thought to

regulate Parkin activation by interaction with the R2 domain, inhibiting its catalytic cysteine. C431, in the inactive state.

To design ubiquibodies that hijack Parkin's catalytic properties, truncations of the N-terminal domain up to residue 217 were made (**Figure 3.10**). This resulted in a 249 amino acid C-terminal domain of Parkin (ParkinCTD), that contained intact sequences of the R1, IRB, and R2 domains. Direct fusion of the DBP was used, generating a family of ubiquibodies, designated uAb^{Parkin}. Although this design did not include it, the Ubl domain could be incorporated in later iterations of uAb^{Parkin}, to aid in co-localization with the proteasome and better facilitate targeted protein silencing.

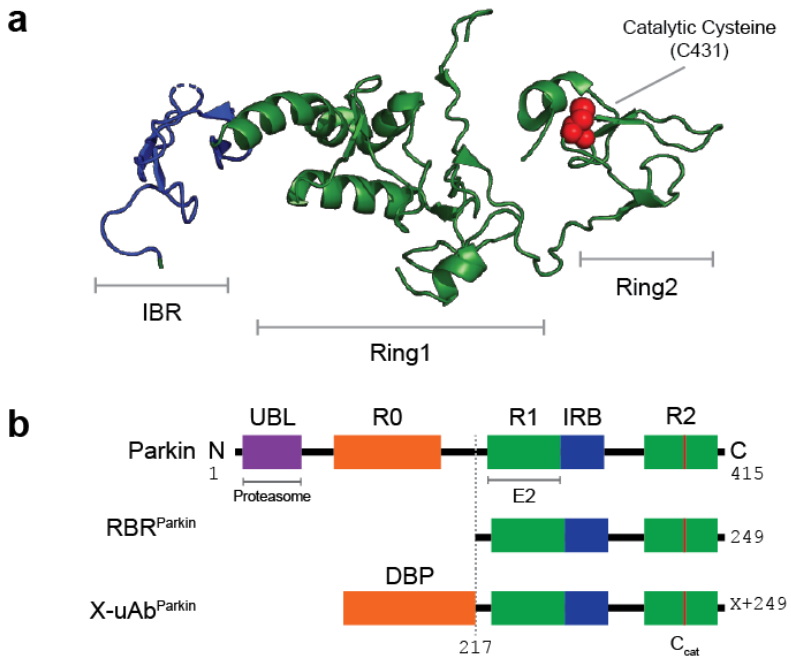


Figure 3.10: Rational design of RBR-based ubiquibodies.
(a) Crystal structure of the RBR domain of human Parkin. Parkin (green) adopts an extended conformation; catalytic cysteine (red) is surface exposed. PDB ID: 5C1Z [75]. (b) Cartoon map of domain organization in wild type Parkin and uAb^{Parkin2}.

As of this report, uAb^{Smurf2} and uAb^{Parkin} constructs have been generated (using the panel of ERK-specific DARPins as DBPs). Similar *in vitro* and *in vivo* characterization as outlined here for CHIP-based ubiquibodies could yield novel and exciting advancements to the control over the ubiquibody technology.

Discussion

In this study, the ubiquitination activity of U-box-containing ubiquibodies has been further elucidated.

For a U-box derived ubiquibody, dimerization is simultaneously interestingly intrinsic to ubiquibody function and frustratingly complicating to an already complex system. The fact that ubiquibodies behave in a manner similar to wild type CHIP is foretelling of their ability to integrate into the UPP without additional chaperones. Notably, dimerization introduces a second target binding site (see Chapter 5), that could lead to non-catalytically active ubiquibody:target complexes. This possibility is supported by the increase in ubiquitination seen when the E2 concentrations are increased relative to the target, suggesting that E2s must outcompete targets binding to the DBP to access the available U-box site in the same half of the dimer.

Ubiquitination profiles for ubiquibodies with the E2 Ube2DE1 have been determined by immunoblot and mass spectrometry. The prevalence of K11 and K48 poly-ubiquitin linkages offer explanation for successful

silencing of proteins observed *in vivo*. The E2 interacting partners of CHIPΔTPR fusion proteins have been determined, and the subset of E2s that lead to K48 linked chains have been found. One possibility for optimizing protein silencing by CHIP-derived ubiquibodies is directed evolution of the U-box domain for selective interaction with the subset of E2s that generate the preferred ubiquitin chains. Alternatively, as K63 linkages may be correlated to the flexibility of the ubiquibody, as seen by their reduction with more flexible ubiquibodies, linker optimization may aid in the preferential formation of proteasome-specific poly-ubiquitin linkages.

While ubiquibodies are no longer unique as a technology for synthetic E3-mediated protein silencing by the proteasome, they the older of the two platforms that utilize a U-box scaffold [15, 76], both derived from CHIP. Other groups have reported ubiquibody-like proteins that contain catalytic analogs to CHIPΔTPR, including subunits of the Cullin RING E3 complex [77, 78]. To date, no ubiquibody-like technology based on a mammalian HECT E3 has been published. However, coinciding with the progression of this project, another E3 catalytic domain, derived from bacterial IpaH9.8 [79], was introduced into ubiquibodies by another member of the DeLisa group. The NEL domain of IpaH9.8 has been proven to be a powerful mediator of targeted protein silencing, and many of the future generations of ubiquibodies will be built from it.

Methods

Genes and plasmid construction. Plasmids were constructed as described in Chapter 2 and Chapter 4. Human Smurf2 and Parkin genes were obtained from Addgene (Plasmid #13678: pRK-Myc-Smurf2 and Plasmid #17613: pRK5-HA-Parkin).

Protein expression and purification. All purified proteins were obtained from cultures of *E. coli* BL21(DE3) cells grown in of Luria-Bertani (LB) medium. Expression was induced with 0.1mM IPTG when the culture density (Abs_{600}) reached 0.6-0.8 and proceeded at 30°C for 6hr, after which cells were harvested by centrifugation at 4,000×g for 20 min at 4°C. The resulting pellets were stored at -80°C overnight. Thawed pellets were resuspended in 15mL buffer (PBS, 10mM imidazole, pH 7.4) and lysed with a high-pressure homogenizer (Avestin EmulsiFlex C5). Lysates were cleared by centrifugation at 20,000×g for 20min at 4°C. His_6-tagged proteins were subjected to gravity Ni^{2+}-affinity purification HisPurTM Ni-NTA Resin (Thermo 88221) following manufacturer's protocols. Samples were desalted into buffer (PBS, pH 7.4) using PD 10 Desalting Columns (GE Healthcare 17-0851-01) following manufacturer's protocols. Samples were stored at 4°C for up to two weeks or diluted to 25% glycerol and stored at -80°C indefinitely.

Protein analysis. SDS-PAGE and immunoblotting of proteins was performed according to standard procedures. The following primary

antibodies were utilized for immunoblotting: rabbit anti-ERK (Cell Signaling Technology, 9102), mouse anti-ubiquitin (Millipore, P4D1-A11), rabbit anti-6x-His-HRP (Abcam, ab1187), rabbit anti-Flag-HRP (Abcam, ab49763), rabbit anti-ubiquitin: linkage specific K27 (Abcam, ab181537 rabbit anti-ubiquitin: linkage specific K48 (Abcam, ab179434) rabbit anti-ubiquitin: linkage specific K63 (Abcam, ab140601). Secondary antibodies goat anti-rabbit IgG (H+L) and anti-mouse IgG (H+L) with HRP conjugation (Promega) were utilized as needed.

Size exclusion chromatography. SEC experiments were performed on an ÅKTA FPLC using a HiLoad Superdex200 column (GE Healthcare) according to manufacturer protocols. A standard curve for molecular weight was generated using Gel Filtration Molecular Weight Markers (blue dextran, cytochrome c, carbonic anhydrase, alcohol dehydrogenase, and β-amylate) (Sigma).

Blue native PAGE. Purified ubiquibodies and control proteins were diluted to the indicated concentrations with PBS, pH 7.4 and 10× loading buffer (0.1% Ponceau S (w/v), 5% glycerol (w/v)). 10μL sample were loaded into the lanes of a NativePAGE™ 4-16% Bis-Tris (Invitrogen). Electrophoresis was performed with anode buffer (25mM imidazole, pH7) and 10× cathode buffer (50mM tricine, 7.5mM imidaxole, 0.02mM Coomassie blue G-250, pH7) at 4°C for 30min at 100V, then 30min at 200V. Cathode buffer was exchanged to 1×, and electrophoresis continued at 200V for 3hr. Semidry transfer of the gel to PVD-membrane

was performed at 80mAmp for 75min. Membrane was destained in buffer (25% methanol, 10% acetic acid) for 1hr to overnight. NativeMark ladder (Thermo) was marked, and membranes were destained in 100% methanol for 5min then analyzed by immunoblot.

***In vitro* ubiquitination assays.** Unless otherwise indicated, *in vitro* ubiquitination reactions were performed as outlined in Chapter 2.

Mass spectrometry. Mass spectrometry analysis of ubiquitination reaction products was performed as described in Chapter 2.

Acknowledgements

Again, I thank Alyse Portnoff for mentorship and many in depth discussions on the topic of ubiquibodies. I thank Jason Boock and Michael Paul Robinson for training in biochemical techniques. And, of course, I thank Matthew DeLisa, for innumerable back-and-forth's in the design and execution of this project. This material is based upon work was supported by the National Science Foundation Career Award CBET-0449080 (to M.P.D.), National Institutes of Health Grant CA132223A (to M.P.D.), New York State Office of Science, Technology and Academic Research Distinguished Faculty Award (to M.P.D.), National Science Foundation GRFP DGE-1144153 (to E.A.S.), the Cornell Presidential Life Science Fellowship Program (to E.A.S.).

CHAPTER 4

THE LINKER

Introduction

Linkers are naturally of great importance in multi-domain proteins. Flexibility is a key attribute, determined by composition and length, as linkers must often allow surrounding domains to reposition for catalysis or with structural rearrangement [80, 81]. The average linker length in natural multi-domain proteins is around 10 residues and favor short or α-helical or coil structures [82]. Conversely, synthetic linkers are often optimized for solubility and flexibility, with secondary regard for secondary structure. scFvs, through treated as a single domain in the ubiquibody, actually consist of two domains, a V_H and a V_L, connected by a flexible linker engineered to increase its stability and folding [83]. Linkers like these are often rich in small or hydrophilic residues, such as glycine and serine, allowing for dramatic degrees of relative rotation of the two flanking poly-peptide regions while maintaining high solubility [84].

The initial design of ubiquibodies included a short linker between the DBP and the catalytic domain, to maintain flexibility between the first DBP, an scFv, and the coiled-coil region of CHIPΔTPR [15]. This inclusion was made in response to findings of Qian et al. that showed flexibility was critical for catalytic function of CHIP [41]. The first generation of ubiquibodies, based on this design, incorporated a five amino acid GS

linker, with the express purpose of maintaining the flexibility needed for catalytic conformation changes [15]. The full spacing between the flanking domains was seven residues, due to genetic incorporation of an EcoRI site intended DBP modularity. This restriction site abutted the GSGSG sequence, resulting an effective seven amino acid linker (EFGSGSG) between the DBP and CHIPΔTPR. In this report, preliminary targeted variations to this deceptively simple domain are explored.

Results

In an effort to further explore linker effects on ubiquibody function, a panel of varied linker lengths was designed (**Table 4.1**). Linkers of 5, 10, 15, 20, and 134 amino acids were genetically fused between the DBP and U-box domains of the ubiquibody scaffold. To maximize flexibility and reduce secondary structure effects, the EcoRI site (translated to EF) was exchanged for a BamHI (translated to GS). Three of the linkers, 5AA, 10AA, and 20AA, were modeled after the first generation linker, consisting of a GSGSG motif in singlet, duplicate, and quadruplicate respectively. The 15AA was derived from a eukaryotic inter-domain linker, native to human Rieske; the crystal structure of Rieske shows the linker to be unstructured, but important for flexibility between the catalytic domain and transmembrane anchor [85]. Unlike the other small linkers, the 15AA is not GS rich. The 134AA was taken from the C-terminal domain of Hsp70, a native interacting partner of CHIP [66].

Natively, this domain functions as both a lid for the substrate binding domain of Hsp70 and interacting partner for native CHIP, via the TPR domain. Because this α-helical rich domain has been shown to be important for ubiquitination of Hsp70/Hsc70 bound substrates [86], a direct fusion of this domain to the CHIPΔTPR was made Theoretically, longer flexible linkers should endow ubiquibodies with a larger radius of catalytic activity.

Table 4.6: Ubiquibody linker design.

Designation	Sequence	Length (AA)	Ref
5' BamHI	GS	2	
5AA	GSGSG	5	
10AA	GSGSGGSGSG	10	
15AA	TSMTATADVL AMAAA	15	[85]
20AA	GSGSGGSGSG GSGSGGSGSG	20	
134AA	itndkgrlsk ddidrmvqea eryksedean rdrvaaknal esytynikqt vedeklrgki seqdknkild kcqevinwld rnqmaekdey ehkqkelerv cnpiisklyq ggpgggsggg gsgasggpti eevd	134	[66]

The linker panel derived from EpE89-uAb was checked for soluble expression in *E. coli* (**Figure 4.1**). All five ubiquibody variants showed high levels of full-length expression, as seen by Coomassie staining and immunoblot, though the 134AA ubiquibody showed a proclivity for degradation *in vivo*.

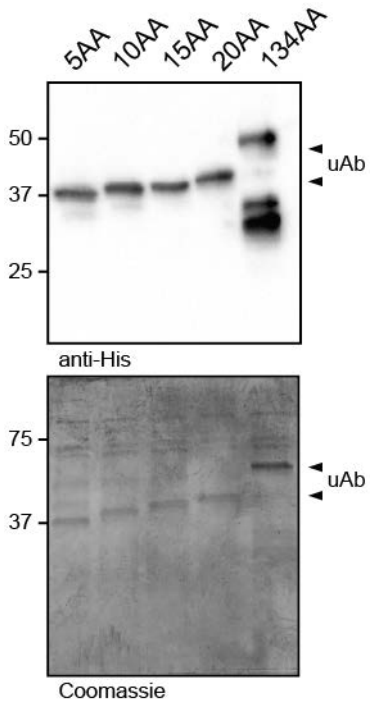


Figure 4.1: Soluble expression of EpE89-uAb linker panel in *E. coli*.
The linker panel of EpE89-uAb expressed well at expected molecular weights of 36.6, 37.0, 37.4, 37.6, and 51.2 kDa respectively, though 134AA showed a tendency for degradation.

To test the ability of the linker panel to ubiquitinate substrate proteins, *in vitro* ubiquitination assays were performed using purified components of the UPP and ERK2 (**Figure 4.2**). Ubiquitination of ERK2-GST was observed with all five ubiquibodies, though to a lesser extent with 134AA EpE89-uAb. Poly-ubiquitin species of high molecular weight were observed in all cases as well, though 10AA, 20AA, and 134AA ubiquibodies showed stronger ubiquitination. Presumably, the high levels of ubiquitination seen by anti-ubiquitin immunoblot for the 134AA ubiquibody correspond to increased auto-ubiquitination of the ubiquibody itself, rather than the substrate protein. Indeed, the 134AA

linker provides fourteen additional lysine residues. Conversely, the decrease in overall ubiquitination by the 5AA and 15AA ubiquibodies as seen by anti-ubiquitin immunoblot, suggest that auto-ubiquitination is decreased relative to the 10AA and 20AA ubiquibodies, while substrate ubiquitination is unaffected, as seen by anti-ERK immunoblot. Taken together, these results suggest that substrate ubiquitination is better facilitated by shorter, flexible linkers, while poly-ubiquitination in general is better supported by a combination of increased GS composition and longer length.

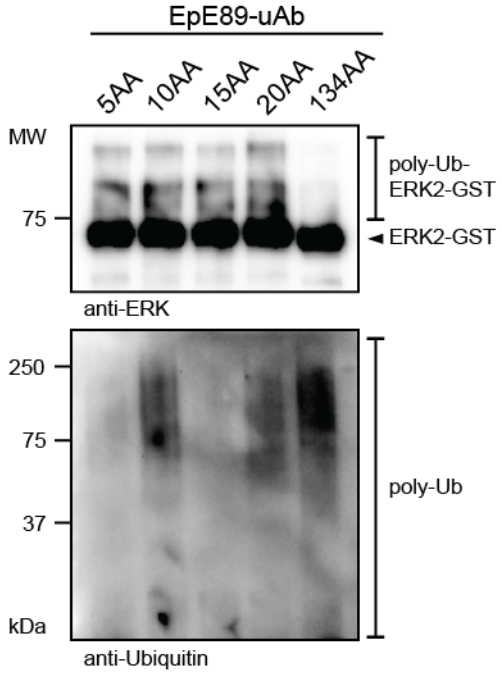


Figure 4.2: *In vitro* ubiquitination by the EpE89-uAb linker panel.
Ubiquibodies derived from the EpE89-uAb were incubated with E1, E2, ubiquitin, ERK2-GST and ATP for two hours. Shifts to higher molecular weight species of ERK2-GST were observed with all ubiquibodies, though to a lesser extent with 134AA EpE89-uAb. Poly-ubiquitination was seen in all cases, though most strongly with 10AA, 20AA, and 134AA ubiquibodies. [51]

To determine the lysine residues susceptible to ubiquitination by the linker panel of the ubiquibodies, the high molecular weight (60-250 kDa) reaction products of *in vitro* ubiquitination assays were analyzed by mass spectrometry. (The 5AA ubiquibody was excluded from this analysis, and the other linker panel members were compared to the original EpE89-uAb containing the EFGSGSG linker ubiquitination profile). Surprisingly, the number of ERK lysines modified with ubiquitin decreased with increasing linker length (**Figure 4.3**). Notably, all ubiquitination sites were contained within the subset of lysines found to be ubiquitinated by the original EpE89-uAb (see Chapter 3). The 15AA and 20AA ubiquibodies, despite extreme compositional differences, were found to ubiquitinate the same four lysine residues, clustered on the top surface of ERK2. The 10AA ubiquibody ubiquitinated the same set of lysines as EpE89-uAb, except for K258. The 134AA ubiquibody was found to ubiquitinate a single site on ERK2, outside of the C-terminal cluster seen with all other ERK-uAbs. The large α-helical structure of this linker may block the catalytic domain from the bound substrate, rather than allow the U-box domain to expand its access to the surface accessible lysines on other faces of ERK2.

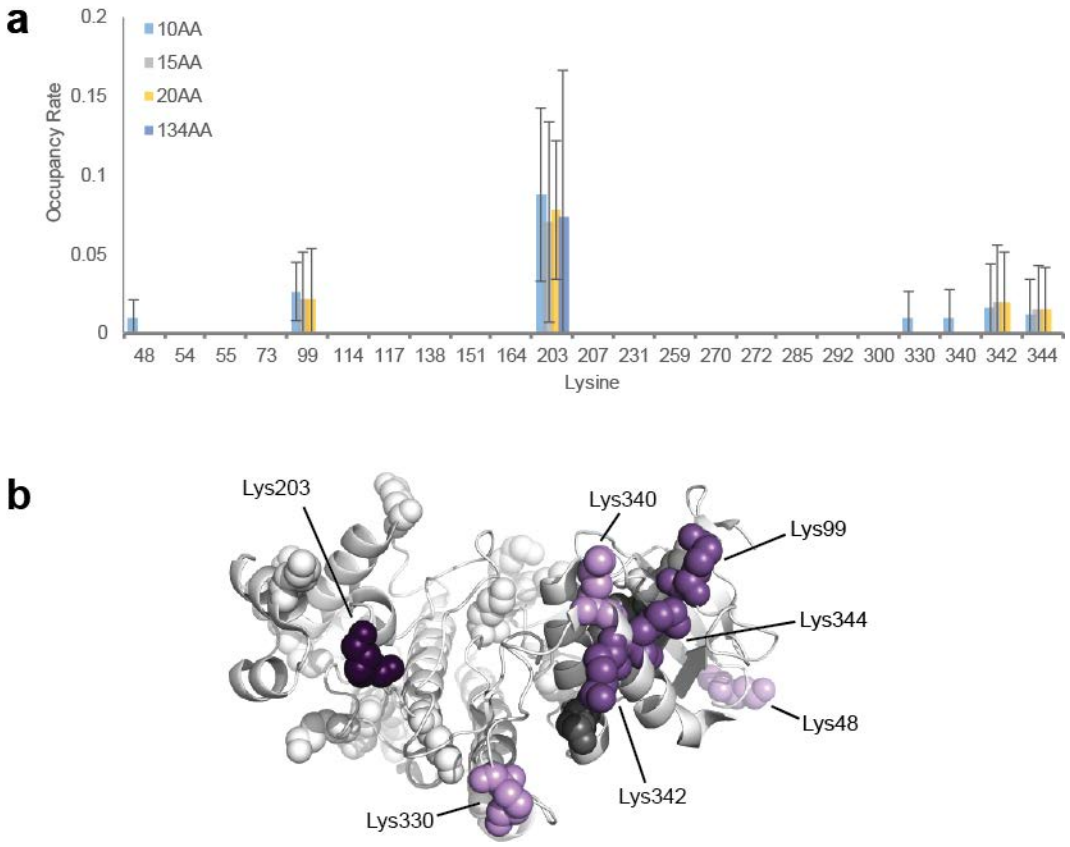


Figure 4.3: Substrate ubiquitination profiles of the EpE89-uAb linker panel. (a) Ubiquitination frequency of ERK2 lysine residues by EpE89-uAb variants. (b) Ubiquitination site map on ERK2 crystal structure: backbone (white ribbon), unmodified lysine residues (white spheres), lysines ubiquitinated by all linker panel ubiquibodies (dark purple), a subset of ubiquibodies (purple), or only the 10AA ubiquibody (light purple), lysines not covered by mass spec analysis (grey). Modified from PDB ID: 3ZU7 [26]. Schematic was generated using PyMol software. [51]

The pattern of auto-ubiquitination changed with varying linker length (**Figure 4.4**). Compared to the original EpE89-EF-GSGSG-uAb, the linker panel showed novel auto-ubiquitination patterns. For the 10AA and 15AA ubiquibodies, no auto-ubiquitination was observed by mass spectrometry analysis. For the 20AA ubiquibody, a single auto-ubiquitination site in C-terminal capping region of the DARPin (and found in the auto-ubiquitination pattern of the original EpE89-uAb) was

identified. The 134AA linker auto-ubiquitinated four sites, K7 and K101 of DARPin EpE89, and two novel sites, K10 and K55, within the linker itself. None of the linker panel ubiquibodies ubiquitinated sites in the U-box domain. The changes in auto-ubiquitination pattern within the series itself suggest that longer linkers allow for more flexible positioning of the U-box relative to its DARPin fusion partner, although, comparisons to the original construct belie that conclusion. However, the composition of the linker, when it includes novel lysine residues, plays an important role in the ability of a ubiquibody to auto-ubiquitinate.

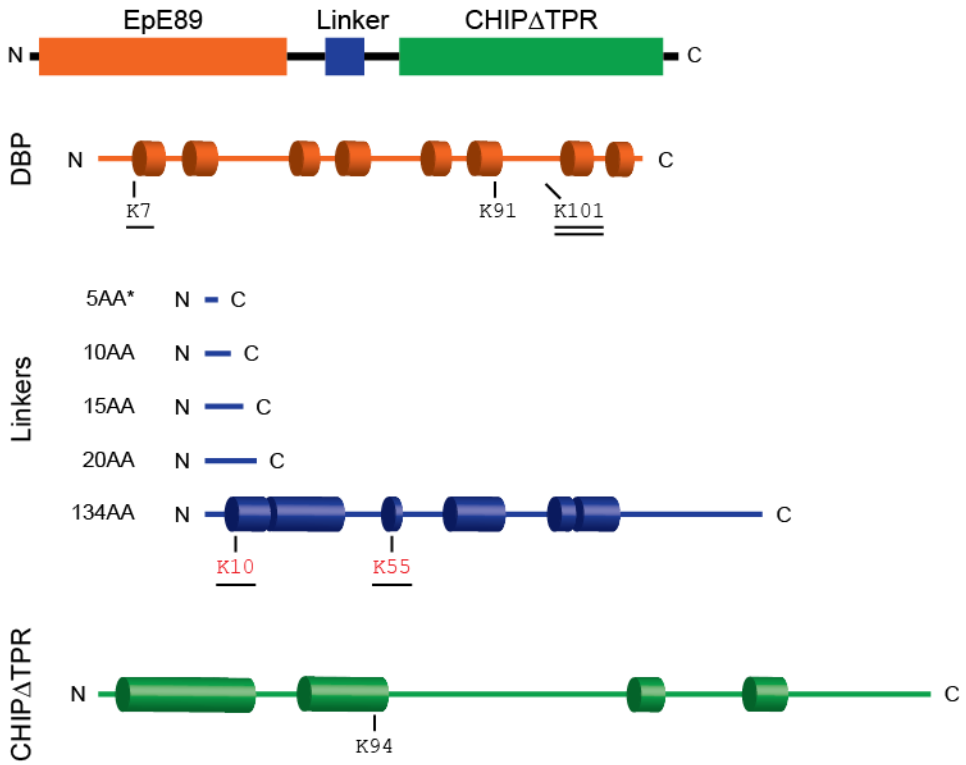


Figure 4.4: Auto-ubiquitination of linker EpE89-uAb.
(top) Cartoon diagram of the ubiquibody structure. (bottom) Site map of auto-ubiquitinated lysines (numbered from start of domain). Sites unique to 134AA linker are show in red. Underlined sites were conserved from the original auto-ubiquitination pattern by the 134AA uAb (single) or by both the 20AA and 134AA uAbs (double).

Discussion

Linker biology provides an important control point for engineered proteins. In its simplest form, the linker simple separates two domains in a polypeptide sequence; in more complex forms, the linker can act as a switch or important structural landmark within the fully folded protein. The linker incorporated into a ubiquibody will play a previously underestimated role in the lysine specificity of the E3 catalytic domain.

In this work, five variants of the first generation linker were characterized *in vitro*. While all led to active, full length constructs, differences in ubiquitination preferences became apparent in immunoblot results and confirmed by mass spectrometry analysis. Conversely to the original hypothesis, increasing linker length corresponded to a decrease in the number of accessible target lysine residues. While some of this may be attributed to the length, underlying structural difference in the linkers may play a role in the different ubiquitination profiles observed.

In regards to auto-ubiquitination, with the exception of the longest linker, no novel auto-ubiquitination sites were introduced by varying linker length and both new sites in the 134AA linker auto-ubiquitination profile were located in the linker itself. Since auto-ubiquitination is a key regulatory feature of E3 ubiquitin ligases, it may benefit ubiquibody stability to either prevent, or more tightly control, this modification. Within this panel of linker ubiquibodies, no ubiquitination of the

CHIPΔTPR domain was found; and only a single residue in this domain was ubiquitinated by the original ubiquibody construct. Linkers and DBPs lacking lysine residues, or containing K→R mutations, may lead to an increase in protein stability *in vivo* by preventing ubiquibody turnover by self-targeting to the proteasome. Additionally, the use more rigid or more structured linkers may prevent auto-ubiquitination by steric-enforced separation of the catalytic and DBP domains.

The potential of the linker domain is one of the most exiting areas of ubiquibody design. Because of its modular nature, more complex linkers, such as switches that can temporally regulate E3 activity, or localization signals, that could cause the ubiquibody to sub-cellularly localize could be incorporated into future generations of ubiquibodies.

Materials and Methods

Plasmid construction. Linker plasmids of the form pET28a(+)-NcoI-GS2-X-BamHI-CHIPΔTPR-SalI-Flag-His-TAA-HindIII, where X corresponds to each of the linkers in **Table 4.1**, were obtained by gene synthesis (Twist Bioscience). The EpE89 DBP was amplified from pET28a(+)-EpE89-uAb, with primers introducing an 5' NcoI sites and a 3' BamHI site; PCR product was double digested (NcoI/BamHI) and ligated into linker plasmid backbones.

Protein expression and purification. Proteins were purified as indicated in Chapter 2 and Chapter 3.

Protein analysis. SDS-PAGE and immunoblotting of proteins was performed according to standard procedures. BioRad Coomassie R-250 stain was used to visualize proteins in SDS-PAGE (BioRad, Mini-PROTEAN® TGX). The following primary antibodies were utilized for immunoblotting: rabbit anti-ERK (Cell Signaling Technolgy, 9102), mouse anti-ubiquitin (Millipore, P4D1-A11), rabbit anti-6x-His-HRP (Abcam, ab1187). Secondary antibodies goat anti-rabbit IgG (H+L) and anti-mouse IgG (H+L) with HRP conjugation (Promega) were utilized as needed.

Ubiquitination assays. Ubiquitination assays were performed as described in Chapter 2

Mass spectrometry analysis. Mass spectrometry analysis was performed as described in Chapter 2.

Acknowledgements

I thank Alyse Portnoff for working so hard to design the first carefully-constructed, functional ubiquibody and Matt DeLisa for the leeway given to explore these questions. We thank Elizabeth Anderson and Sheng Zhang of the Cornell Biotechnology Resource Center for performing the mass spectrometry analysis and database searches. This work was supported by the National Science Foundation Career Award CBET-0449080 (to M.P.D.), National Institutes of Health Grant CA132223A (to M.P.D.), New York State Office of Science, Technology and

Academic Research Distinguished Faculty Award (to M.P.D.), National Science Foundation GRFP DGE-1144153 (to E.A.S.), the Cornell Presidential Life Science Fellowship Program (to E.A.S.)

CHAPTER 5

THE ANATOMY OF A UBIQUIBODY

Introduction

Traditional protein silencing has been primarily achieved with the use of small molecule inhibitors. These compounds, often the active components of therapeutics, can be range greatly in potency and act either allosterically or competitively, to prevent native protein function. Unfortunately, the proteome has a high level of homology, often leading to off-target effects, that are only exacerbated by the fact that drugs must work in a one-to-one molar ratio with the proteins they silence. Higher concentrations of lead to higher rates of off-target effects. Furthermore, not all proteins are 'druggable', in that a useful binding cleft, where a small molecule could potentially inhibit protein function, does not exist [87]. These limitations of small molecules inhibitors necessitate the need for specific and enzymatic silencing.

The term 'ubiquibodies' was introduced by our lab in 2014, and it and other ubiquibody-like technologies have since been engineered to create tunable silencing of specific protein targets.

This work has shown ubiquibodies to be effective protein silencers. Ubiquibody-mediated silencing is postulated to be achieved by proteolytic cleavage of target proteins inside the proteasome. This generation of ubiquibodies relies on the simultaneous interaction of three independent

proteins, E2, ubiquibody, and target, where the ubiqiubody acts as scaffold bridge between the E2 and target proteins. (**Figure 5.1**).

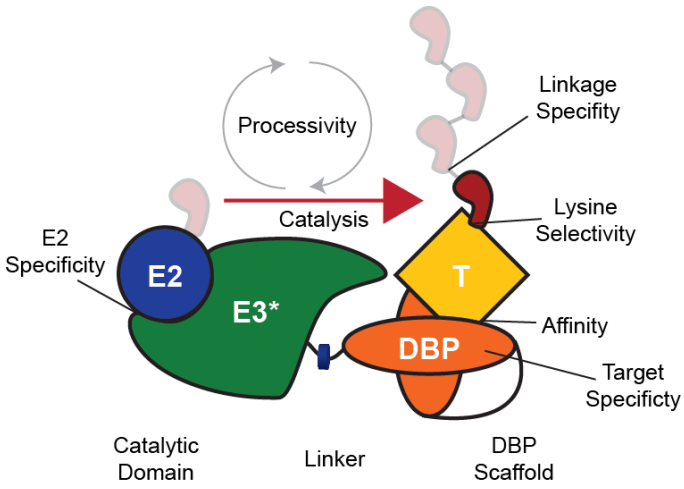


Figure 5.1: The anatomy of a ubiquibody.

The three domains of the ubiquibody, DBP, linker, and catalytic domain, function independently and form a modular platform that is amenable to rapid exchange of domains. In this report, ubiquibody-driven silencing has been found to be a function of several factors, including E2 specificity, catalytic effectivity and processivity, acceptor lysine selectivity, cognate target specificity and affinity, and poly-ubiquitin linkage specificity.

The Designer binding protein

All foreseeable generations of ubiquibodies will require directed-target specificity. By genetically encoding this specificity in the form of a DBP, large portions of the proteome that were previously inaccessible become targetable. Many types of DBP have been shown to be amenable to the ubiquibody scaffold, such that the ubiquibody technology is likely never to limited for wont of a DBP. At most, additional library screening may be necessary to acquire the desired binder.

One of the key considerations with DBP selection is its affinity for the cognate target. Affinities have played roles in PTM-specific silencing, as well as a possible cause of differential silencing by ubiquibodies specific for the same target. Natural PTM-specific silencing has been observed in several cases, including the MAPK pathway that was targeted by ubiquibodies here [88]. Further exploration into PTM-specific silencing will depend on both advancements to PTM binding capabilities (PTM-specific antibodies or DBPs) and characterization of the natural mechanisms with which cells differentiate protein subpopulations.

The E3 Catalytic domain

The type of E3 catalytic domain has been a common factor between all experimental ubiquibodies reported here. As an enzymatic process that requires the assembly of multiple protein subunits, the optimizing the kinetics of the association and dissociation of each subunit will likely

be an important part of pushing the reaction towards the kinetically active complex, when irreversible covalently linkage of ubiquitin occurs (**Figure 5.2**). Furthermore, the ability to hack the ubiquitin code (Figure 5.3) will depend on the E2/E3 interaction; the U-box offers advantage here, in that it naturally interacts with a large subset of E2s. Directed evolution of this domain could yield ubiquibodies capable of selectively interacting with E2s specific for different types of ubiquitin linkage.

The incorporation of a HECT, or HECT-like, E3 catalytic domain would greatly reduce the complexity of this system, as complexes of only two enzymes would be required to assemble to mediate ubiquitin ligation. Smurf2 has been proposed as a potential donor of catalytic activity, and investigations into bacterial IpaH have already begun. As always, if biologists want to try it, nature has beaten them to it [25]. Bacterial E3 ubiquitin ligases hijack mammalian degradation machinery to subvert host defense mechanisms and better effect their pathogenicity [89]; because bacteria evolved it first, many of these E3s come equipped with mechanisms which could become the focus of research into 'deliverable' ubiquibodies.

Not all ubiquibody-like technologies will depend upon engineered E3 catalytic activity. Proteolysis targeting chimeras (PROTACs) are another technology that hijack the UPP, but instead of containing catalytic activity themselves, form a bridge between native E3s and target proteins [90]. Alternatively, in Clift et al. proposed a much larger

approach (measured in kDa), using full length-antibodies as the bridge substrates and an E3 ubiquitin ligase [91]. These two technologies demonstrate the extremes between which UPP-hijacking technologies can function.

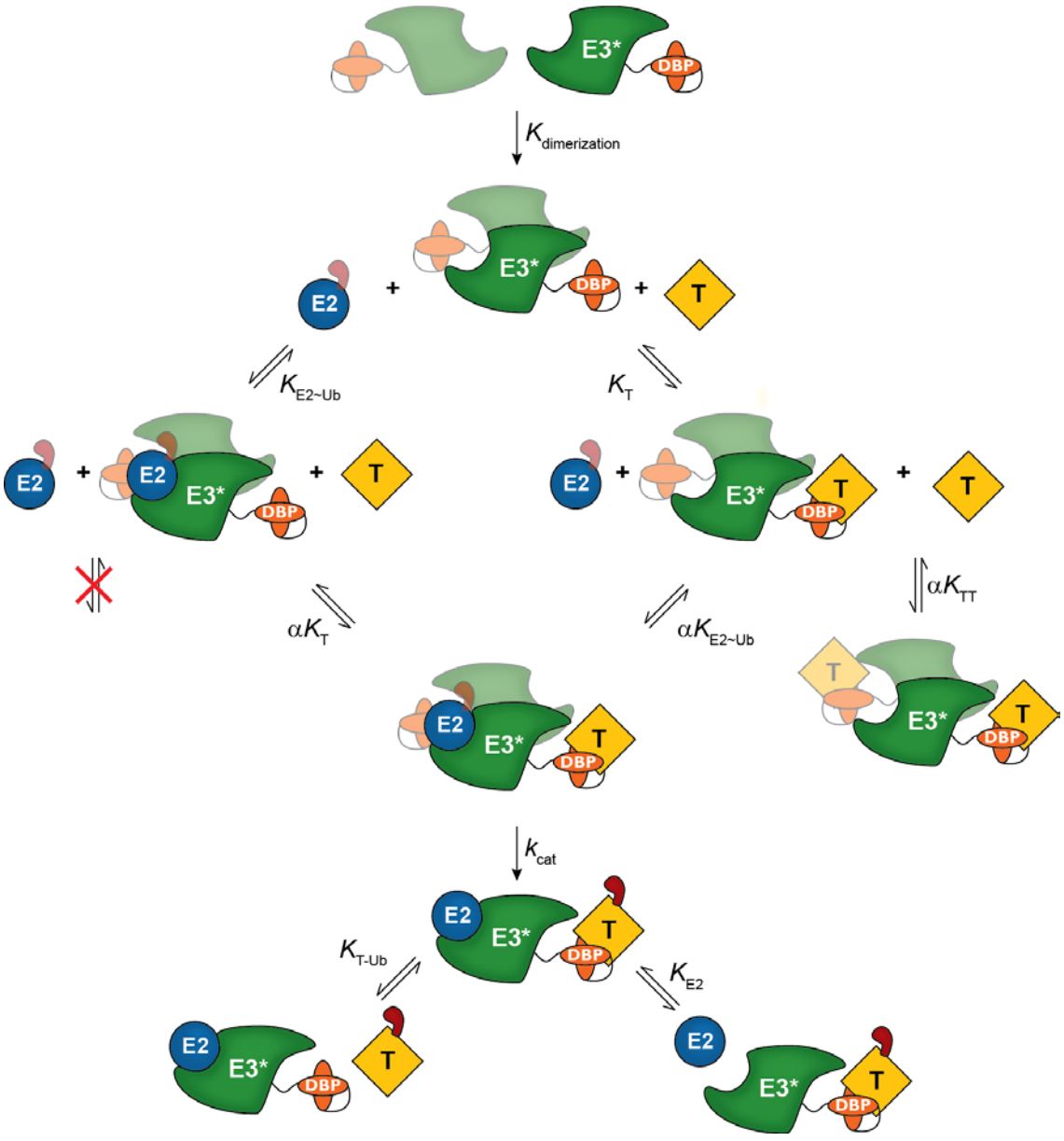


Figure 5.2: Kinetics of ubiquitination with dimeric ubiquibodies.

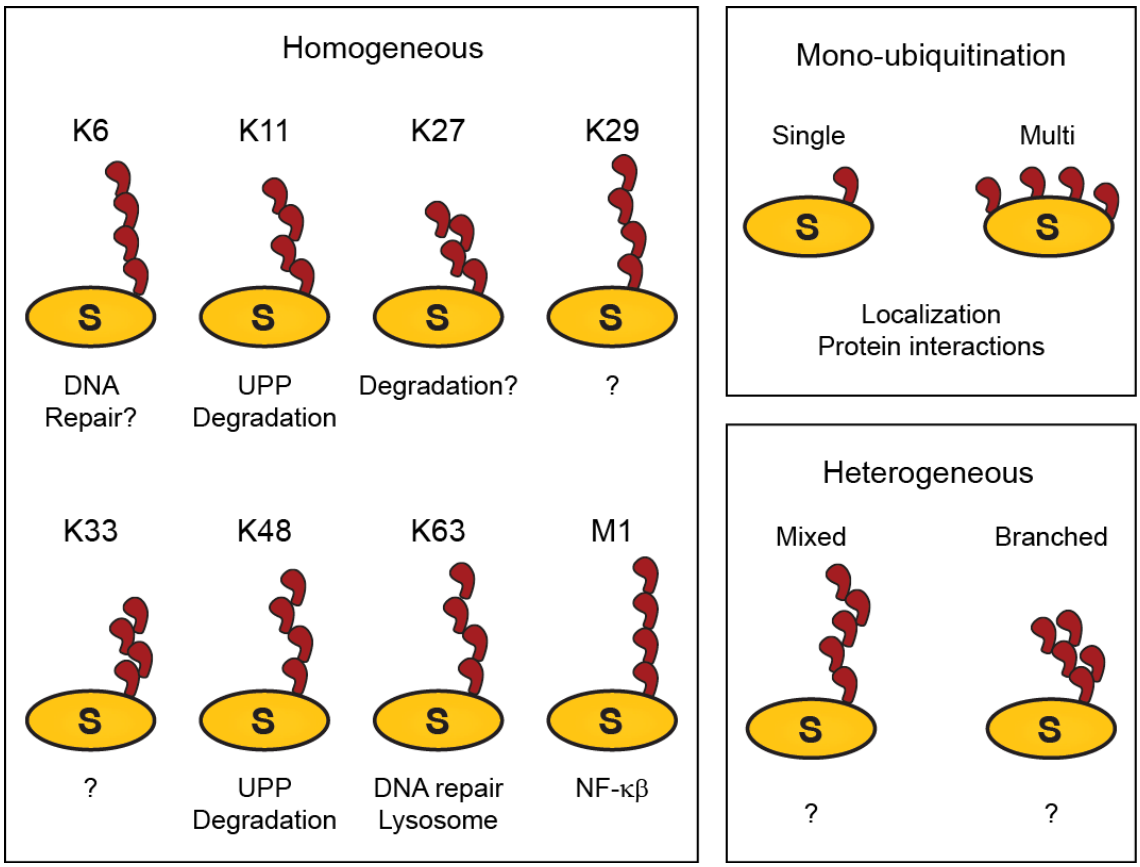


Figure 5.3: The ubiquitin code.
Lysine 11 and lysine 48 poly-ubiquitin linkages target substrates to the proteasome, while lysing 63 linkages are purported to be involved lysosomal targeting by autophagy. [42]

The Linker

The small flexible linker was not originally considered an independent domain of the ubiquibody. Here, the linker has been shown to play an important role in determining the ubiquitination profile of target substrate lysines. With the growing number of advancements in linker biology, the domain will likely provide an important control point for the ubiquitination profile of the ubiquibody and enzymatic activity overall.

One of the key considerations for ubiquibody effectivity, superficially explored here, is expression of the fusion protein, either in bacterial or in mammalian cells. Depending upon the application, significant levels of protein will be required. While all CHIPΔTPR fusion proteins reported here, expressed exceptionally well in a bacterial expression strains (*E. coli* BL21(DE3)) from high-copy plasmids, they did not all express equally well in mammalian cells from mammalian expression vector pcDNA3. In preliminary experiments, uAb^{Smurf2} and uAb^{Parkin} constructs similarly suffered from undetectable expression in mammalian cells (data not shown). Also, while ectopic expression has been the method of introduction of ubiquibodies to mammalian cells, other means of delivery will need to be explored going forward.

Conclusions

The future of ubiquibody, despite its name, need not be limited to ubiquitination. Other catalytic domains, such as deubiquitinating enzyme (DUB) or sumoylation catalytic domains, could be incorporated into the existing scaffolds. Modularity will play an important role in all rationally designed synthetic enzymes, and the knowledge concerning one domain gleaned here may facilitate the hijacking of completely unrelated biological systems.

Inducible enzymes have been designed for other applications. As they exist now, ubiquibodies are constitutively active in the presence of

their target and the UPP. Light or chemically activated domains could be added to the ubiquibody scaffold, allowing for *temporal* targeted protein silencing. Indeed, per the time of this report, the first steps into these uncharted realms of ubiquitination technology have already been taken by members of this lab.

The final piece to solving the puzzle of targeted protein silencing will be the full elucidation of the ubiquitin code. Understanding the effects of different linkages, and, then, selectable formation of those linkages on target proteins will constitute an invaluable advancement in protein silencing. And, as ubiquitination is involved in so many cellular processes, ubiquibodies, though born from the need for protein-level silencing technologies, could yet transcend their intended purpose to, unimagined possibilities.

Acknowledgements

I would like to acknowledge all the people who have worked on ubiquibody project, from the beginning until this point: Alyse Portnoff, for nigh single-handedly laying the groundwork and training myself, Morgan Baltz, for the explosive expansion of the ubiquibody technology, Kate Forsythe and Esther Chen, for the hard bench work needed to keep ubiquibodies chugging, Carolyn Snyder, for inquisitive and persistent dedication to understanding ubiquibody theory, Jeff Varner, for theoretical input into the conception of this project, and Matthew DeLisa,

for standing by each and every one of these people, and all those to come, with mentorship and perseverance. I humbly thank the many teachers who have guided me through the graduate program at Cornell University, particularly Linda Nicholson, Yimon Aye, Matthew Pascek, Julius Lucks, Bill Brown, Tony Brestcher and Volker Vogt.

This material is based upon work was supported by the National Science Foundation Career Award CBET-0449080 (to M.P.D.), National Institutes of Health Grant CA132223A (to M.P.D.), New York State Office of Science, Technology and Academic Research Distinguished Faculty Award (to M.P.D.), National Science Foundation GRFP DGE-1144153 (to E.A.S.), the Cornell Presidential Life Science Fellowship Program (to E.A.S.).

APPENDIX A

SEQUENCES

A.1. CHIPΔTPR

-Linker-hCHIPΔTPR-SalI-Flag-His6

```
GSGSGRLNFG DDIPSALRIA KKKRWNSIEE RRIHQESELH SYLSRLIAAE - 50
RERELEECQR NHEGDEDDSH VRAQQACIEA KHDKYMADMD ELFSQVDEKR -100
KKRDIPDYLC GKISFELMRE PCITPSGITY DRKDIEEHLQ RVGHFDPVTR -150
SPLTQEQLIP NLAMKEVIDA FISENGWVED YVDGADYKDD DDKGHHHHHH -200
```

A.2. scFv-C4

```
MAQVQLQESG GGLVQPGGSL RLSCAASGFT FSSYSMSWVR QAPGKGLEWVA - 50
                          VH
VISYDGSNKY YADSVKGRFT ISRDNSKNTL YLQMNSLRAE DTAVYYCARDR -100
                    Linker
YFDLWGRGTL VTVSSGGGGS GGGGSGGGGS QSALTQPASV SGSPGQSITIS -150
                          VL
CTGTSSDIGA YNYVSWYQQY PGKAPKLLIY DVSNRPSGIS NRFSGSKSGDT -200

ASLTISGLQA EDEADYYCSS FANSGPLFGG GTKVTVLGAA AYPYDVPDYA  -249
```

A.3. Huntington N17 Peptide

```
MATLEKLMKA FESLKSF    - 17
```

A.4. YS1

```
MVSSVPTKLE VVAATPTSLL ISWDASYSSS VSYYRITYGE TGGNSPVQEF - 50
TVPGSKSTAT ISGLKPGVDY TITVYAYSYY YYYYSSPISI NYRT       - 94
```

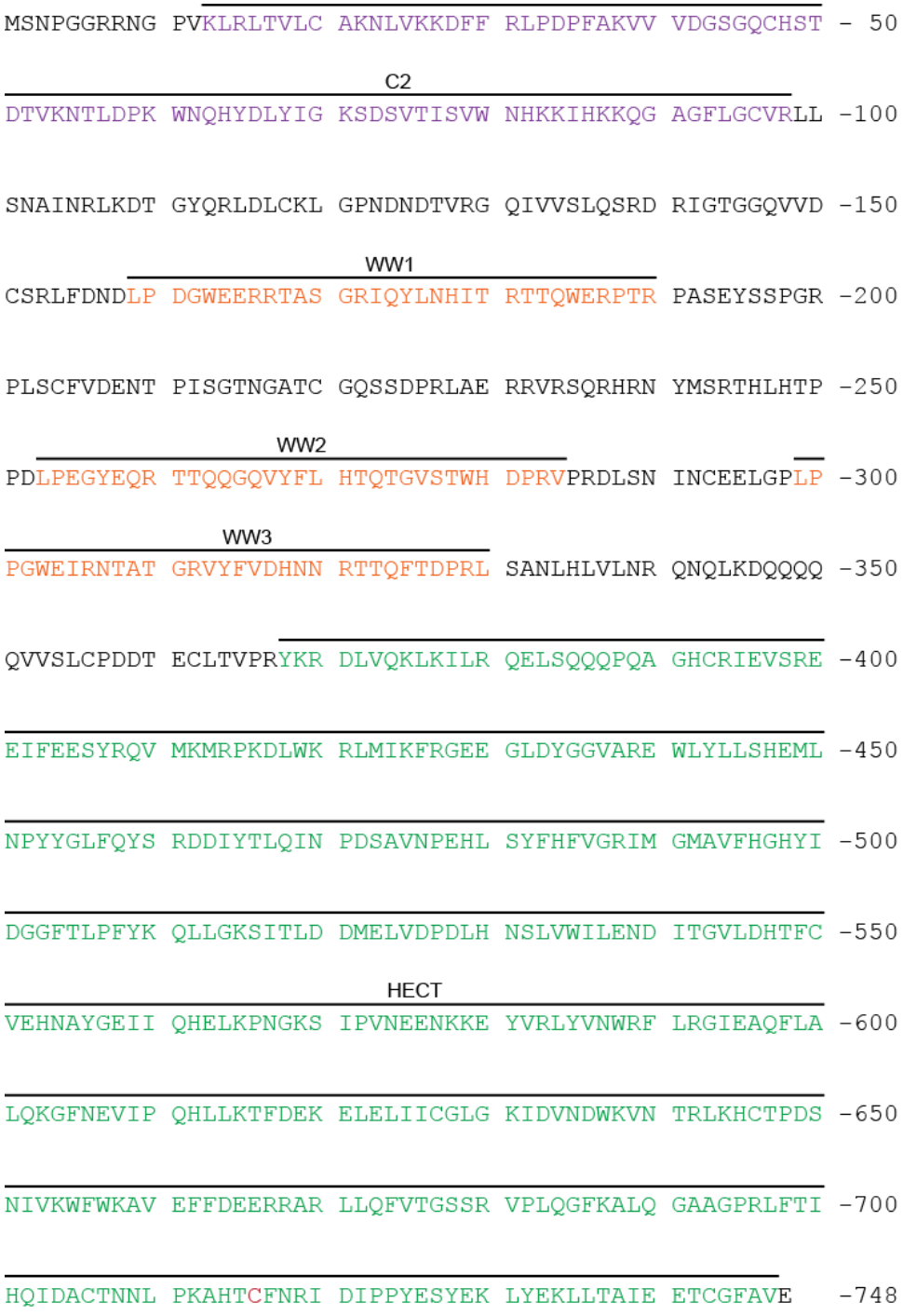
A.5. OFF7

```
DLGRKLLEAA RAGQDDEVRI LMANGADVNA ADNTGTTPLH LAAYSGHLEI - 50
VEVLLKHGAD VDASDVFGYT PLHLAAYWGH LEIVEVLLKN GADVNAMDSD -100
GMTPLHLAAK WGYLEIVEVL LKHGADVNAQ DKFGKTAFDI SIDNGNEDLA -150
EILQKLNEF                                               -159
```

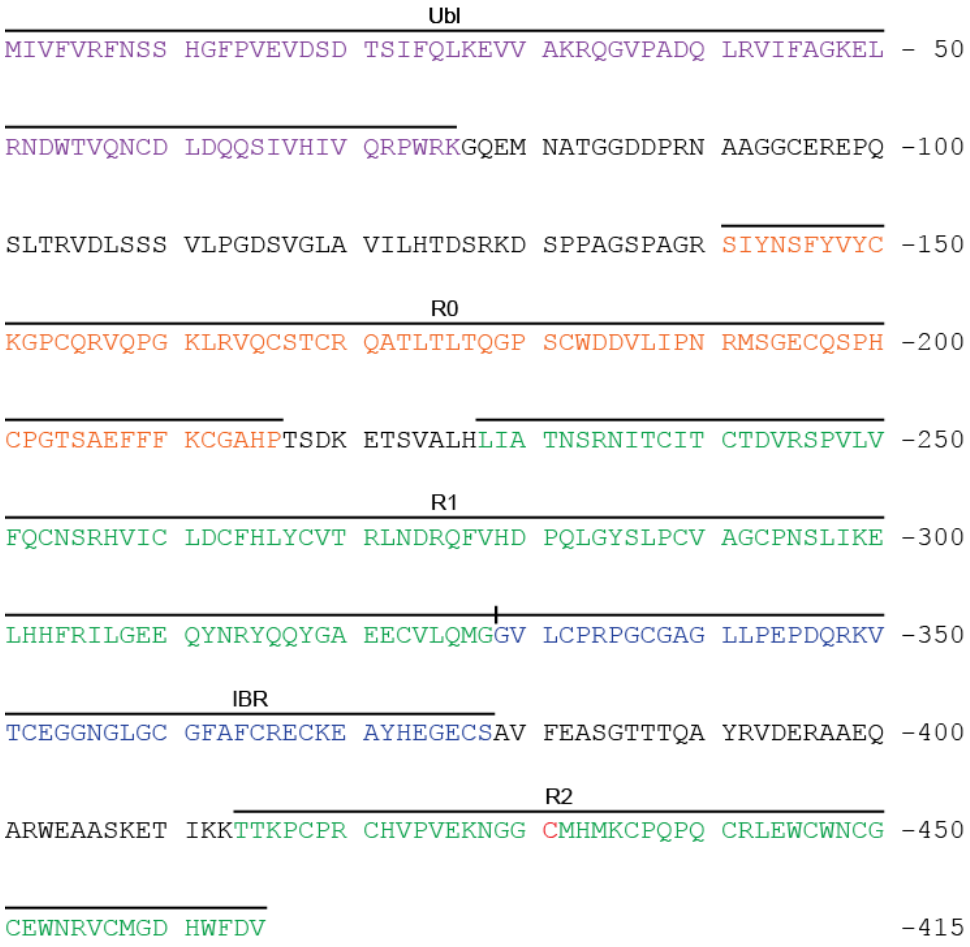
A.6. ERK DARPins

```
                N-Cap                          Repeat 1
DARPin   DLGKKLLEAARAGQDDEVRILMANGADVNA  XDXXGXTPLHLAAXXGHLEIVEVLLKXGADVNA
E40      ------------------------------  H-DQ-S-------WI--P---------H------
pE59     ------------------------------  L-ED-L-------QL-----------Y------
EpE89    ------------------------------  F-NI-L------SQW-----------H------
EpE82    ------------------------------  F-QI-L-------FE-----------Y------
         ******************************   *  *:******:  ** ********:******

                Repeat 2                       Repeat 3
DARPin   XDXXGXTPLHLAAXXGHLEIVEVLLKXGADVNA XDXXGXTPLHLAAXXGHLEIVEVLLKXGADVNA
E40      R-TD-W-------DN-----------Y------ Q-AY-L-------DR-----------H------
pE59     E-NF-I-------IR-----------H------
EpE89    K-IY-I-------AK-----------H------
EpE82    I-SY-I-------LH-----------Y------
          *  *:*******  ***********:******

                C-Cap
DARPin   QDKFGKTAFDISIDNGNEDLAEILQ
E40      -------------------------
pE59     -------------------------
EpE89    -------------------------
EpE82    -------------------------
         *************************
```

A.7. human Smurf2



A.8. *human Parkin*



APPENDIX B

E2 CONJUGATING ENZYMES

B.1. E2 Enzymes Screened

E2 Conjugating Enzyme[1]	Alternate Name	Tag
Ube2A	HR6A	No tag
Ube2B	HR6B	No tag
Ube2C	UbcH10	T7
Ube2D1	UbcH5A	T7
Ube2D2	UbcH5B	T7
Ube2D3	UbcH5C	No tag
Ube2D4	UbcH5D	T7
Ube2E1	UbcH6	No tag
Ube2E2	UbcH8	T7
Ube2E3	UbcH9	No tag
Ube2F	NCE2	T7
Ube2G1	Ubc7	T7
Ube2G2	Ubc7	No tag
Ube2H	UbcH2	No tag
Ube2I	Ubc9	No tag
6His-Ube2J1	NCUBE1	His-T7
Ube2J2	NCUBE2	T7
Ube2K	Ubc1	No tag
Ube2L3	UbcH7	No tag
Ube2L6	UbcH8	No tag
Ube2M	Ubc12	No tag
Ube2N	Ubc13	No tag
Ube2N/Ube2V1	Ubc13/Uev1A	No tag/T7
Ube2N/Ube2V2	Ubc13/Mms2	No tag/No tag
6His-Ube2Q	NICE-5	His-T7
Ube2Q2	-	No tag
Ube2R1	CDC34	T7
Ube2R2	CDC34B	T7
Ube2S	E2-EPF	T7
Ube2T	HSPC150	No tag

Ube2V1	Uev1A	T7
Ube2V2	Mms2	No tag
6His-Ube2W	Ubc16	His-T7
6His-Ube2Z	USE1	His-T7

[1]E2s were obtained from Ubiquigent

B.2. Human E2 Enzymes not covered

AKTIP	FLJ13258
BIRC6	BRUCE
UBE2E4P	UbcM2
UBE2L1	UBEL1, L-UBC, UBCH7N3
UBE2L2	
UBE2L4	
UBE2L5	
UBE2NL	
UBE2U	MGC35130

[2]Full list of human E2 conjugating enzymes from [69]

REFERENCES

1. Walsh, C., *Posttranslational modification of proteins : expanding nature's inventory*. 2006, Englewood, Colo.: Roberts and Co. Publishers.

2. Scientific, T. *Overview of Post-Translational Modifications (PTMs).* [Webpage] March 13, 2018]; Available from: https://www.thermofisher.com/us/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/overview-post-translational-modification.html.

3. Rockland. *Post-Translational Modification (PTM) Antibodies.* [Webpage] March 13, 2018]; Available from: https://rockland-inc.com/post-translational-modification-antibodies.aspx.

4. Winzeler, E.A., et al., *Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis.* Science, 1999. **285**(5429): p. 901-6.

5. Giaever, G., et al., *Functional profiling of the Saccharomyces cerevisiae genome.* Nature, 2002. **418**(6896): p. 387-91.

6. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.

7. Sander, J.D. and J.K. Joung, *CRISPR-Cas systems for editing, regulating and targeting genomes.* Nat Biotechnol, 2014. **32**(4): p. 347-55.

8. Fire, A., et al., *Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.* Nature, 1998. **391**(6669): p. 806-11.

9. Kim, D. and J. Rossi, *RNAi mechanisms and applications.* Biotechniques, 2008. **44**(5): p. 613-6.

10. Casey, J.P., R.A. Blidner, and W.T. Monroe, *Caged siRNAs for spatiotemporal control of gene silencing.* Mol Pharm, 2009. **6**(3): p. 669-85.

11. Rucklidge, G.J., et al., *Turnover rates of different collagen types measured by isotope ratio mass spectrometry.* Biochim Biophys Acta, 1992. **1156**(1): p. 57-61.

12. Kwon, Y.T. and A. Ciechanover, *The Ubiquitin Code in the Ubiquitin-Proteasome System and Autophagy.* Trends Biochem Sci, 2017. **42**(11): p. 873-886.

13. Ciechanover, A., *The unravelling of the ubiquitin system.* Nat Rev Mol Cell Biol, 2015. **16**(5): p. 322-4.

14. Ardley, H.C. and P.A. Robinson, *E3 ubiquitin ligases.* Essays Biochem, 2005. **41**: p. 15-30.

15. 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.

16. Orlandi, R., et al., *Cloning immunoglobulin variable domains for expression by the polymerase chain reaction.* Proc Natl Acad Sci U S A, 1989. **86**(10): p. 3833-7.

17. McCafferty, J., et al., *Phage antibodies: filamentous phage displaying antibody variable domains.* Nature, 1990. **348**(6301): p. 552-4.

18. Hoogenboom, H.R., *Selecting and screening recombinant antibody libraries.* Nat Biotechnol, 2005. **23**(9): p. 1105-16.

19. Marschall, A.L., et al., *Targeting antibodies to the cytoplasm.* MAbs, 2011. **3**(1): p. 3-16.

20. Murphy, K., et al., *Janeway's immunobiology*. 8th ed. 2012, New York: Garland Science. xix, 868 p.

21. Zhang, M., et al., *Chaperoned ubiquitylation--crystal structures of the CHIP U box E3 ubiquitin ligase and a CHIP-Ubc13-Uev1a complex.* Mol Cell, 2005. **20**(4): p. 525-38.

22. Zahnd, C., et al., *Directed in vitro evolution and crystallographic analysis of a peptide-binding single chain antibody fragment (scFv) with low picomolar affinity.* J Biol Chem, 2004. **279**(18): p. 18870-7.

23. Hansen, S., et al., *Design and applications of a clamp for Green Fluorescent Protein with picomolar affinity.* Sci Rep, 2017. **7**(1): p. 16292.

24. 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.

25. Binz, H.K., P. Amstutz, and A. Pluckthun, *Engineering novel binding proteins from nonimmunoglobulin domains.* Nat Biotechnol, 2005. **23**(10): p. 1257-68.

26. Kummer, L., et al., *Structural and functional analysis of phosphorylation-specific binders of the kinase ERK from designed ankyrin repeat protein libraries.* Proc Natl Acad Sci U S A, 2012. **109**(34): p. E2248-57.

27. Binz, H.K., et al., *High-affinity binders selected from designed ankyrin repeat protein libraries.* Nat Biotechnol, 2004. **22**(5): p. 575-82.

28. 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.

29. Lapouge, K., et al., *Structure of the TPR domain of p67phox in complex with Rac.GTP.* Mol Cell, 2000. **6**(4): p. 899-907.

30. Rotman, M.B. and F. Celada, *Antibody-mediated activation of a defective beta-D-galactosidase extracted from an Escherichia coli mutant.* Proc Natl Acad Sci U S A, 1968. **60**(2): p. 660-7.

31. 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.

32. Koch, H., et al., *Direct selection of antibodies from complex libraries with the protein fragment complementation assay.* J Mol Biol, 2006. **357**(2): p. 427-41.

33. der Maur, A.A., et al., *Direct in vivo screening of intrabody libraries constructed on a highly stable single-chain framework.* J Biol Chem, 2002. **277**(47): p. 45075-85.

34. Lynch, S.M., C. Zhou, and A. Messer, *An scFv intrabody against the nonamyloid component of alpha-synuclein reduces intracellular aggregation and toxicity.* J Mol Biol, 2008. **377**(1): p. 136-47.

35. Lecerf, J.M., et al., *Human single-chain Fv intrabodies counteract in situ huntingtin aggregation in cellular models of Huntington's disease.* Proc Natl Acad Sci U S A, 2001. **98**(8): p. 4764-9.

36. Krebber, C., et al., *Co-selection of cognate antibody-antigen pairs by selectively-infective phages.* FEBS Lett, 1995. **377**(2): p. 227-31.

37. Fujiwara, K., et al., *A single-chain antibody/epitope system for functional analysis of protein-protein interactions.* Biochemistry, 2002. **41**(42): p. 12729-38.

38. Gilbreth, R.N., et al., *A dominant conformational role for amino acid diversity in minimalist protein-protein interfaces.* J Mol Biol, 2008. **381**(2): p. 407-18.

39. 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.

40. Rothbauer, U., et al., *Targeting and tracing antigens in live cells with fluorescent nanobodies.* Nat Methods, 2006. **3**(11): p. 887-9.

41. Qian, S.B., et al., *Engineering a ubiquitin ligase reveals conformational flexibility required for ubiquitin transfer.* J Biol Chem, 2009. **284**(39): p. 26797-802.

42. Komander, D. and M. Rape, *The ubiquitin code.* Annu Rev Biochem, 2012. **81**: p. 203-29.

43. Yau, R. and M. Rape, *The increasing complexity of the ubiquitin code.* Nat Cell Biol, 2016. **18**(6): p. 579-86.

44. Venkatraman, P., et al., *Eukaryotic proteasomes cannot digest polyglutamine sequences and release them during degradation of polyglutamine-containing proteins.* Mol Cell, 2004. **14**(1): p. 95-104.

45. Kisselev, A.F., et al., *The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes. Implications for understanding the degradative mechanism and antigen presentation.* J Biol Chem, 1999. **274**(6): p. 3363-71.

46. Davies, S.W., et al., *Formation of neuronal intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation.* Cell, 1997. **90**(3): p. 537-48.

47. Zoghbi, H.Y. and H.T. Orr, *Glutamine repeats and neurodegeneration.* Annu Rev Neurosci, 2000. **23**: p. 217-47.

48. De Genst, E., et al., *Structure of a single-chain Fv bound to the 17 N-terminal residues of huntingtin provides insights into pathogenic amyloid formation and suppression.* J Mol Biol, 2015. **427**(12): p. 2166-78.

49. Portnoff, A.D., *Ubiquibodies : engineered ubiquitin ligases with unnatural substrate specificity for targeted protein silencing.* 2014. 1 online resource (144 pages).

50. Khoury, G.A., R.C. Baliban, and C.A. Floudas, *Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database.* Sci Rep, 2011. **1**.

51. Stephens, E.A., et al., *Ubiquibodies: Synthetic E3 Ubiquitin Ligases for PTM-resolution Targeted Silencing.* (in preparation), 2018.

52. Roskoski, R., Jr., *ERK1/2 MAP kinases: structure, function, and regulation.* Pharmacol Res, 2012. **66**(2): p. 105-43.

53. Cohen, P., *Protein kinases--the major drug targets of the twenty-first century?* Nat Rev Drug Discov, 2002. **1**(4): p. 309-15.

54. Hackel, B.J., A. Kapila, and K.D. Wittrup, *Picomolar affinity fibronectin domains engineered utilizing loop length diversity, recursive mutagenesis, and loop shuffling.* J Mol Biol, 2008. **381**(5): p. 1238-52.

55. Zahnd, C., et al., *A designed ankyrin repeat protein evolved to picomolar affinity to Her2.* J Mol Biol, 2007. **369**(4): p. 1015-28.

56. Sukenik, S., P. Ren, and M. Gruebele, *Weak protein-protein interactions in live cells are quantified by cell-volume modulation.* Proc Natl Acad Sci U S A, 2017. **114**(26): p. 6776-6781.

57. Morreale, F.E. and H. Walden, *Types of Ubiquitin Ligases.* Cell, 2016. **165**(1): p. 248-248 e1.

58. Lorenz, S., *Structural mechanisms of HECT-type ubiquitin ligases.* Biol Chem, 2018. **399**(2): p. 127-145.

59. Lorenz, S., et al., *Macromolecular juggling by ubiquitylation enzymes.* BMC Biol, 2013. **11**: p. 65.

60. Verdecia, M.A., et al., *Conformational flexibility underlies ubiquitin ligation mediated by the WWP1 HECT domain E3 ligase.* Mol Cell, 2003. **11**(1): p. 249-59.

61. Huibregtse, J.M., et al., *A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase.* Proc Natl Acad Sci U S A, 1995. **92**(7): p. 2563-7.

62. Sheng, Y., et al., *A human ubiquitin conjugating enzyme (E2)-HECT E3 ligase structure-function screen.* Mol Cell Proteomics, 2012. **11**(8): p. 329-41.

63. Metzger, M.B., et al., *RING-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination.* Biochim Biophys Acta, 2014. **1843**(1): p. 47-60.

64. Dove, K.K. and R.E. Klevit, *RING-Between-RING E3 Ligases: Emerging Themes amid the Variations.* J Mol Biol, 2017. **429**(22): p. 3363-3375.

65. Hatakeyama, S. and K.I. Nakayama, *U-box proteins as a new family of ubiquitin ligases.* Biochem Biophys Res Commun, 2003. **302**(4): p. 635-45.

66. Zhang, P., et al., *Crystal structure of the stress-inducible human heat shock protein 70 substrate-binding domain in complex with peptide substrate.* PLoS One, 2014. **9**(7): p. e103518.

67. Jiang, J., et al., *CHIP is a U-box-dependent E3 ubiquitin ligase: identification of Hsc70 as a target for ubiquitylation.* J Biol Chem, 2001. **276**(46): p. 42938-44.

68. Xu, Z., et al., *Interactions between the quality control ubiquitin ligase CHIP and ubiquitin conjugating enzymes.* BMC Struct Biol, 2008. **8**: p. 26.

69. van Wijk, S.J. and H.T. Timmers, *The family of ubiquitin-conjugating enzymes (E2s): deciding between life and death of proteins.* FASEB J, 2010. **24**(4): p. 981-93.

70. Stewart, M.D., et al., *E2 enzymes: more than just middle men.* Cell Res, 2016. **26**(4): p. 423-40.

71. Wu, K., J. Kovacev, and Z.Q. Pan, *Priming and extending: a UbcH5/Cdc34 E2 handoff mechanism for polyubiquitination on a SCF substrate.* Mol Cell, 2010. **37**(6): p. 784-96.

72. Kavsak, P., et al., *Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF beta receptor for degradation.* Mol Cell, 2000. **6**(6): p. 1365-75.

73. Wang, G., J. Yang, and J.M. Huibregtse, *Functional domains of the Rsp5 ubiquitin-protein ligase.* Mol Cell Biol, 1999. **19**(1): p. 342-52.

74. Ogunjimi, A.A., et al., *Regulation of Smurf2 ubiquitin ligase activity by anchoring the E2 to the HECT domain.* Mol Cell, 2005. **19**(3): p. 297-308.

75. Riley, B.E., et al., *Structure and function of Parkin E3 ubiquitin ligase reveals aspects of RING and HECT ligases.* Nat Commun, 2013. **4**: p. 1982.

76. Kanner, S.A., T. Morgenstern, and H.M. Colecraft, *Sculpting ion channel functional expression with engineered ubiquitin ligases.* Elife, 2017. **6**.

77. 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.

78. Fulcher, L.J., et al., *Targeting endogenous proteins for degradation through the affinity-directed protein missile system.* Open Biol, 2017. **7**(5).

79. Chou, Y.C., et al., *Conserved structural mechanisms for autoinhibition in IpaH ubiquitin ligases.* J Biol Chem, 2012. **287**(1): p. 268-75.

80. Gokhale, R.S. and C. Khosla, *Role of linkers in communication between protein modules.* Curr Opin Chem Biol, 2000. **4**(1): p. 22-7.

81. Reddy Chichili, V.P., V. Kumar, and J. Sivaraman, *Linkers in the structural biology of protein-protein interactions.* Protein Sci, 2013. **22**(2): p. 153-67.

82. George, R.A. and J. Heringa, *An analysis of protein domain linkers: their classification and role in protein folding.* Protein Eng, 2002. **15**(11): p. 871-9.

83. Bird, R.E., et al., *Single-chain antigen-binding proteins.* Science, 1988. **242**(4877): p. 423-6.

84. Argos, P., *An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion.* J Mol Biol, 1990. **211**(4): p. 943-58.

85. Nett, J.H., C. Hunte, and B.L. Trumpower, *Changes to the length of the flexible linker region of the Rieske protein impair the interaction of ubiquinol with the cytochrome bc1 complex.* Eur J Biochem, 2000. **267**(18): p. 5777-82.

86. Zhang, H., et al., *A bipartite interaction between Hsp70 and CHIP regulates ubiquitination of chaperoned client proteins.* Structure, 2015. **23**(3): p. 472-82.

87. Backus, K.M., et al., *Proteome-wide covalent ligand discovery in native biological systems.* Nature, 2016. **534**(7608): p. 570-4.

88. Nguyen, L.K., W. Kolch, and B.N. Kholodenko, *When ubiquitination meets phosphorylation: a systems biology perspective of EGFR/MAPK signalling.* Cell Commun Signal, 2013. **11**: p. 52.

89. Huibregtse, J. and J.R. Rohde, *Hell's BELs: bacterial E3 ligases that exploit the eukaryotic ubiquitin machinery.* PLoS Pathog, 2014. **10**(8): p. e1004255.

90. 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.

91. Clift, D., et al., *A Method for the Acute and Rapid Degradation of Endogenous Proteins.* Cell, 2017. **171**(7): p. 1692-1706 e18.