

Georgia State University

ScholarWorks @ Georgia State University

---

Chemistry Dissertations

Department of Chemistry

---

12-14-2021

## Profiling Substrate Proteins of Ring and RBR type E3 ligases by Orthogonal Ubiquitin Transfer and the Development of a Peptide Activator Targeting HECT-type E3 ligase

Li Zhou

*Georgia State University*

Follow this and additional works at: [https://scholarworks.gsu.edu/chemistry\\_diss](https://scholarworks.gsu.edu/chemistry_diss)

---

### Recommended Citation

Zhou, Li, "Profiling Substrate Proteins of Ring and RBR type E3 ligases by Orthogonal Ubiquitin Transfer and the Development of a Peptide Activator Targeting HECT-type E3 ligase." Dissertation, Georgia State University, 2021.

doi: <https://doi.org/10.57709/26658520>

This Dissertation is brought to you for free and open access by the Department of Chemistry at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Chemistry Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact [scholarworks@gsu.edu](mailto:scholarworks@gsu.edu).

Profiling Substrate Proteins of Ring and RBR type E3 ligases by Orthogonal Ubiquitin Transfer  
and the Development of a Peptide Activator Targeting HECT-type E3 ligase

by

Li Zhou

Under the Direction of Jun Yin, Professor/Ph.D.

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2021

## ABSTRACT

RING finger protein 38 (RNF38) is a member of the RING family of E3 ubiquitin (UB) ligases while HHARI is one of RBR family members: both regulate key processes in the cell. Through the E1-E2-E3 cascades, RNF38 and HHARI transfer UB to cellular proteins and regulate their stability, subcellular localization, and interaction with other proteins. Identifying the ubiquitination targets of RNF38 and HHARI holds the key to deciphering their roles in cell regulation. In this study, we used phage display to engineer the RING domains of the two E3 ligases for their incorporation into an orthogonal UB transfer (OUT) cascade to identify their substrate proteins in the cell. The OUT cascade consisting of engineered E1, E2 and E3 enzymes enable the exclusive transfer of an engineered UB (xUB) through a designated E3 to its substrate proteins. By affinity purification of xUB-conjugated proteins from cells expressing the OUT cascade of an E3 ligase and subsequent identification by proteomics, we hope to map the UB transfer pathways from the E3 ligase to its cellular targets. So far, we have successfully constructed the OUT cascade of RNF38 and HHARI and we used them to profile their substrates in HEK293T cells. The newly acquired substrate profiles of RNF38 and HHARI enables us to discover potential roles of the E3s in the cell. Additionally, we screened out a cyclic peptidomimetics P6 targeting E6AP, an E3 ubiquitinating p53 driven by the human papillomavirus and regulating pathways involved in neurodevelopmental diseases. The ubiquitin ligase activity of E6AP ubiquitinating its substrate proteins UbxD8, HHR23A, and  $\beta$ -catenin was substantially increased by P6 both *in vitro* and *in vivo*. Such results suggest that synthetic ligands may be used to enhance E3 activity in cells. Overall, we have demonstrated that the OUT cascades we generated with RING E3 and RBR E3 are powerful tools for profiling E3 substrates and mapping the cellular circuits mediated by the E3

enzymes. Additionally, the identified substrates can be further used as indicators to test the effects on the activity of E3 ligases targeted by specific peptides.

INDEX WORDS: Ubiquitination, Orthogonal UB Transfer, RING finger protein 38 (RNF38), HHARI, Phage Display, E6AP, P6

Copyright by

Li Zhou

2021

Profiling Substrate Proteins of Ring and RBR type E3 ligases by Orthogonal Ubiquitin Transfer  
and the Development of a Peptide Activator Targeting HECT-type E3 ligase

by

Li Zhou

Committee Chair: Jun Yin

Committee: Binghe Wang

Angela Mabb

Electronic Version Approved:

Office of Graduate Services

College of Arts and Sciences

Georgia State University

December 2021

## ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my advisor Prof. JUN YIN for the continuous support of my Ph.D. study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this dissertation. I could not have imagined having a better advisor and mentor for my Ph.D. study.

Besides my advisor, I would like to thank the rest of my dissertation committee, Prof. Binghe Wang and Prof. Angela Mabb, for their insightful comments and encouragement, as well as for asking hard questions which incited me to widen my research from various perspectives.

My sincere thanks also go to my mentor, Dr. Karan Bhuripanyo, who taught me firstly to have a comprehensive understanding of the ubiquitin research field. He also used his own experience to introduce the projects he was studying. Thanks to his patience and teaching, I received trained on primer design, molecular cloning, protein expression, phage selection, and *in vitro* ubiquitination. Eventually, I started to work on my own project for constructing the OUT cascade of RNF38 and CBI-b. By utilizing techniques I learned, I successfully engineered the interface between RNF38 and UbcH5B and generated the OUT cascade of xE1-xE2-xRNF38 *in vitro*. Dr. Yiyang Wang is a postdoc who joined our lab in 2014. I appreciated his help and training on techniques related to cellular work, which played a critical role in allowing us to get our work published. Essentially, he taught me how to culture mammalian cells, such as HEK293, HEK293T and Hela cells. Furthermore, I was able to do immunoprecipitation (IP) and immunoblotting (IB) for substrate verification in the cell. Also, I thank my fellow lab mates, Dr. Geng Chen, Dr. Han Zhou and Ruochuan Liu for the stimulating discussions, the sleepless nights we work together before deadlines, and all the fun we have had in the last four years.

I would also like to thank my undergraduate trainees In Ho Jeong and Geon Jeong, who have been working with me on all the OUT E3 ligase projects, and master student trainee, Nicolas A Rios, who helped on the initial engineering of RNF38 Ring engineering.

Last but not the least, I would like to thank my family: my parents Guohai Zhou and Guo'e Quan who encouraged me to apply to GSU for my Ph.D. program, my sister, Zhenyan Zhou, who made me realize my potential work on biochemistry area, and finally my wife, Farfalla Hu, who accompanied me to complete my Ph.D. training and supported me to overcome all the difficulties.

## TABLE OF CONTENTS

<b>ACKNOWLEDGEMENTS .....</b>	<b>IV</b>
<b>TABLE OF CONTENTS .....</b>	<b>VI</b>
<b>LIST OF TABLES .....</b>	<b>XIII</b>
<b>LIST OF FIGURES .....</b>	<b>XIV</b>
<b>LIST OF ABBREVIATIONS .....</b>	<b>XVIII</b>
<b>1 INTRODUCTION.....</b>	<b>1</b>
<b>1.1 Background.....</b>	<b>1</b>
<b>1.2 The challenge of substrate identification .....</b>	<b>4</b>
<b>1.3 The rationale and past development of orthogonal ubiquitin transfer (OUT) .....</b>	<b>5</b>
<b>1.4 The overview of E3s and their characteristics.....</b>	<b>10</b>
<b>1.4.1 HECT E3 ligase .....</b>	<b>11</b>
<b>1.4.2 RBR E3 ligase .....</b>	<b>12</b>
<b>1.4.3 RING E3 ligase .....</b>	<b>13</b>
<b>1.5 Discussion and future directions.....</b>	<b>14</b>
<b>2 ENGINEERING UBCH5B-RNF38 INTERFACE BY PHAGE SELECTION .....</b>	<b>17</b>
<b>2.1 Rationale on the engineering of UbcH5b-RNF38 interface .....</b>	<b>17</b>
<b>2.2 Direct Evolution Technique .....</b>	<b>18</b>
<b>2.3 Introduction of M13 Phage .....</b>	<b>22</b>
<b>2.4 Property of Phagemid vectors.....</b>	<b>23</b>

<b>2.5</b>	<b>Results .....</b>	<b>25</b>
2.5.1	Phage display of the RING domain of RNF38.....	25
2.5.2	Model selection for RNF38 Ring-displaying-phage .....	26
2.5.3	Library phage construction of RNF38 .....	28
2.5.4	Phage selection of Library .....	31
2.5.5	Confirmation of OUT activity of selected mutants .....	34
<b>2.6</b>	<b>Discussion.....</b>	<b>35</b>
<b>2.7</b>	<b>Method .....</b>	<b>37</b>
2.7.1	pComb-RNF38 RING (FLAG) construction and protein expression of wtRNF38 Ring .....	37
2.7.2	Autoubiquitination assay of recombinantly expressed RNF38 RING .....	38
2.7.3	RNF38 phage propagation.....	38
2.7.4	In vitro autoubiquitination assay of RNF38 Ring-displayed phage .....	39
2.7.5	Biopanning of RNF38 Ring phage and quantification by ELISA .....	40
<b>2.8</b>	<b>RNF38 library selection.....</b>	<b>41</b>
2.8.1	Model Selection.....	41
2.8.2	Library construction of pComb-RNF38 RING .....	41
2.8.3	Library phage preparation for RNF38 Ring .....	42
2.8.4	Biopanning and selection of library phage .....	43

2.8.5	Testing the selected mutants via biopanning assays to validate the OUT cascade of RNF38 Ring .....	44
2.8.6	Construction pGEX-RNF38 Ring mutants and protein expression.....	44
3	<b>VERIFYING THE OUT CASCADES OF RNF38 MUTANTS AND OTHER RING-TYPE E3S BOTH IN VITRO AND IN VIVO .....</b>	<b>46</b>
3.1	<b>Background information .....</b>	<b>46</b>
3.1.1	<b>The rationale.....</b>	<b>46</b>
3.2	<b>The structural and biological roles of Cbl-b, Pirh2 and RNF38 .....</b>	<b>46</b>
3.2.1	<b>CBL-B (Casitas B lymphoma-b).....</b>	<b>46</b>
3.2.2	<b>PIRH2 (p53-induced protein with a RING-H2 domain) .....</b>	<b>48</b>
3.2.3	<b>RNF38 (Ring Finger Protein 38).....</b>	<b>50</b>
3.3	<b>Results .....</b>	<b>52</b>
3.3.1	<b>Verifying the activity of RNF38 Full-length via in vitro ubiquitination .....</b>	<b>52</b>
3.3.2	<b>Examining the activity of xRNF38 Ring mutants in vitro.....</b>	<b>53</b>
3.3.3	<b>Engineering the both Cbl-b and Pirh2 Ring by for OUT activity .....</b>	<b>55</b>
3.3.4	<b>Examining the activity of both xCbl-b Ring and xPirh2 Ring mutants in vitro .....</b>	<b>56</b>
3.3.5	<b>Reconstitution of xRNF38 OUT activity in vivo.....</b>	<b>58</b>
3.3.6	<b>Reconstitution of xCbl-b and xPirh2 OUT activity in vivo .....</b>	<b>63</b>
3.3.7	<b>Verifying substrates of RNF38 by in vitro ubiquitination.....</b>	<b>64</b>

3.3.8	General design of verifying substrates for E3s in vivo for future work.....	65
3.4	Discussion.....	66
3.5	Methods.....	69
3.5.1	Generation of Lentiviral constructs.....	69
3.5.2	Construction of stable cell line .....	69
3.5.3	Production of lentiviral particles .....	70
3.5.4	Lentiviral Transduction.....	70
3.5.5	DNA Transient Transfection.....	71
3.5.6	In vivo Pulldown assays .....	72
3.5.7	In vitro ubiquitination of substrate proteins .....	75
4	<b>SUBSTRATES IDENTIFICATION OF RBR-TYPE E3 LIGASE HHARI/ARIH1</b>	<b>77</b>
4.1	Structure and Function of HHARI/ARIH1 .....	77
4.2	Results .....	79
4.2.1	Phage display of the RBR domain of HHARI .....	79
4.2.2	Verifying the OUT activity of HHARI RBR mutants via in vitro ubiquitination .....	83
4.2.3	Testing of xHHARI full-length OUT activity in vitro.....	84
4.2.4	Reconstitution of xHHARI OUT activity in vivo .....	86
4.2.5	Verifying substrates of HHARI by in vitro ubiquitination .....	90
4.3	Discussion and conclusion .....	91

<b>4.4</b>	<b>Methods.....</b>	<b>94</b>
4.4.1	Generation of Lentiviral constructs.....	94
4.4.2	Construction of stable cell line .....	94
4.4.3	Production of lentiviral particles .....	95
4.4.4	Lentiviral Transduction.....	95
4.4.5	DNA Transient Transfection.....	96
4.4.6	In vivo Pulldown assays .....	96
4.4.7	In vitro ubiquitination of substrate proteins .....	99
<b>5</b>	<b>EXTEND SUBSTRATE STUDIES OF E6AP BY DESIGNING A CYCLIC <math>\gamma</math>-AA PEPTIDE .....</b>	<b>101</b>
5.1	HECT-type E3 ligase E6AP.....	101
5.1.1	Structure and biological function .....	101
5.1.2	Rational for designing the Cyclic $\gamma$ -AA peptide targeting E6AP.....	102
5.2	Results .....	103
5.2.1	Library screening and characterizing the affinities of the $\gamma$ -AA peptide ligands targeting E6AP HECT domain completed by Dr. Jianfeng Cai's group .....	103
5.2.2	Effects of $\gamma$ -AApeptides on E6AP activity assayed by autoubiquitination ....	106
5.2.3	Effect of $\gamma$ -AApeptides on substrate ubiquitination catalyzed by E6AP .....	108
5.2.4	P6-mediated enhancement of ubiquitination and accelerated degradation of E6AP substrates in HEK293 cells. ....	111

5.2.5	Stability of P6.....	113
5.3	Discussion and conclusion .....	114
5.4	Methods.....	116
5.4.1	Library preparation.....	116
5.4.2	Library Screening .....	117
5.4.3	Cleavage and analysis. ....	117
5.4.4	FITC-labeled peptides preparation. ....	118
5.4.5	Binding affinity.....	118
5.4.6	Construction of protein expression plasmids.....	119
5.4.7	In vitro assays to test the effect of peptide ligands on self-ubiquitination of E6AP and substrate ubiquitination by E6AP .....	119
5.4.8	Co-immunoprecipitation and to examine the effects of P6 peptide on UbxD8 and Rad23a.....	120
5.4.9	E6AP induced protein degradation in HEK293T cell.....	121
5.5	Supplementary materials .....	121
	REFERENCES.....	123
	APPENDICES .....	135
	Appendix A .....	135
	Published and submitted Papers.....	135



**LIST OF TABLES**

<b>Table 1 The selection output phages and the enrichment for six rounds. ....</b>	<b>33</b>
<b>Table 2 Primary antibody information.....</b>	<b>45</b>
<b>Table 3 Primers used in this study .....</b>	<b>45</b>
<b>Table 4 Potential substrates of RNF38 identified by proteomics .....</b>	<b>62</b>
<b>Table 5 Primary antibody information.....</b>	<b>75</b>
<b>Table 6 Sequences of the primers used in this study .....</b>	<b>76</b>
<b>Table 7 Potential substrates of HHARI .....</b>	<b>89</b>
<b>Table 8 Primary antibody information.....</b>	<b>100</b>
<b>Table 9 The alignment of the peptide sequences from the library screening and the binding affinities of the identified <math>\gamma</math>-AA peptide ligands P1-8 with the HECT domain of E6AP as measured by fluorescence polarization assay .....</b>	<b>105</b>

## LIST OF FIGURES

<b>Figure 1.1 Native ubiquitin transfer cascade for proteasomal degradation.....</b>	<b>2</b>
<b>Figure 1.2 Different modes of ubiquitination .....</b>	<b>3</b>
<b>Figure 1.3 Current Approaches for E3 substrate identification.....</b>	<b>5</b>
<b>Figure 1.4 The OUT (orthogonal ubiquitin transfer) cascade .....</b>	<b>6</b>
<b>Figure 1.5 A scheme for bump and hole .....</b>	<b>8</b>
<b>Figure 1.6 The Crystal structure of the yeast E1 Uba1 in complex with UB (PDB ID 3CMM).....</b>	<b>9</b>
<b>Figure 1.7 The interface of the yeast E1 Uba1 and Ubc1 .....</b>	<b>10</b>
<b>Figure 2.1 Directed evolution overview .....</b>	<b>19</b>
<b>Figure 2.2 Schematic diagram of the structure of an M-13 filamentous bacteriophage .....</b>	<b>22</b>
<b>Figure 2.3 In vitro auto-ubiquitination assays of RNF38 RING .....</b>	<b>26</b>
<b>Figure 2.4 Biopanning assay for RNF38 Ring displayed on M13 bacteriophage.. ..</b>	<b>27</b>
<b>Figure 2.5 Model selection of RNF38 phage.....</b>	<b>28</b>
<b>Figure 2.6 Structure analysis of the Ring domains of RNF38 interacting with the UbcH5b. .....</b>	<b>29</b>
<b>Figure 2.7 Schematic representation of RNF38 Ring library construction.....</b>	<b>31</b>
<b>Figure 2.8 A scheme for library selection process .....</b>	<b>32</b>
<b>Figure 2.9 Activity testing of RNF38 Ring mutants from selection .....</b>	<b>34</b>

<b>Figure 3.1 Roles of Cbl E3s in cell regulation .....</b>	<b>48</b>
<b>Figure 3.2 A proposed model for CHK2 ubiquitination by Pirh2 in response to DNA damage response (DDR).....</b>	<b>48</b>
<b>Figure 3.3 The cellular regulation of RNF38.....</b>	<b>52</b>
<b>Figure 3.4 The activity test of full-length wild-type RNF38 .....</b>	<b>53</b>
<b>Figure 3.5 Activity assays of RNF38 Ring mutant variants.....</b>	<b>54</b>
<b>Figure 3.6 Structure comparison and sequence alignment of the Ring domains of RNF38, Cbl-b and Pirh2.....</b>	<b>55</b>
<b>Figure 3.7 <i>In vitro</i> ubiquitination assays of Pirh2 variants.....</b>	<b>56</b>
<b>Figure 3.8 <i>In vitro</i> ubiquitination assays of Cbl-b (38~427) variants.....</b>	<b>58</b>
<b>Figure 3.9 Expressing the OUT cascade of xRNF38 and HBT-xUB in HEK293T cells.....</b>	<b>59</b>
<b>Figure 3.10 Purification of xUB-modified proteins using affinity chromatography to determine the RNF38 substrates. ....</b>	<b>61</b>
<b>Figure 3.11 Volcano plot for substrates profile of RNF38 .....</b>	<b>63</b>
<b>Figure 3.12 Expressing the OUT cascade of xCbl-b&amp;HBT-xUB and xPirh2&amp;HBT-xUB in HEK293T cells, respectively .....</b>	<b>64</b>
<b>Figure 3.13 <i>In vitro</i> assays to test the ubiquitination of RNF38 substrates identified by OUT.....</b>	<b>65</b>
<b>Figure 4.1 <i>In vitro</i> auto-ubiquitination assays of HHARI RBR .....</b>	<b>80</b>

<b>Figure 4.2 Analysis by Modeled crystal structure of HHARI RBR and UbcH7.....</b>	<b>81</b>
<b>Figure 4.3 Enrichment of phage selection and sequence alignment of potential xHHARI mutants.....</b>	<b>83</b>
<b>Figure 4.4 Activity assays of HHARI RBR mutant variants. ....</b>	<b>84</b>
<b>Figure 4.5 <i>In vitro</i> ubiquitination assays of full-length HHARI mutants. ....</b>	<b>85</b>
<b>Figure 4.6 Expressing the OUT cascade of xHHARI in HEK293T cells .....</b>	<b>87</b>
<b>Figure 4.7 Purification of xUB-modified proteins using affinity chromatography to determine the HHARI substrates .....</b>	<b>88</b>
<b>Figure 4.8 Volcano plot for substrates profile of HHARI.....</b>	<b>89</b>
<b>Figure 4.9 OUT-identified HHARI substrates were tested for ubiquitination <i>in vitro</i>. .....</b>	<b>91</b>
<b>Figure 5.1 Screening of <math>\gamma</math>-AA peptide ligands with binding affinities with the E6AP HECT domain.....</b>	<b>105</b>
<b>Figure 5.2 Effect of <math>\gamma</math>-AA peptide ligands on the UB ligase activity of E6AP assayed by self-ubiquitination .....</b>	<b>108</b>
<b>Figure 5.3 Effect of the <math>\gamma</math>-AA peptide ligands on the UB ligase activity of E6AP assayed by substrate ubiquitination .....</b>	<b>110</b>
<b>Figure 5.4 The stimulatory effect of the P6 ligand on E6AP-catalyzed substrate ubiquitin in the cell and acceleration of the degradation of E6AP substrates .....</b>	<b>112</b>
<b>Figure 5.5 Analytic HPLC traces for peptide stability study.....</b>	<b>113</b>

**Figure 5.6 Procedure to synthesize the OBOC  $\gamma$ -AApeptide library ..... 121**

**Figure 5.7 A schematic illustration of the screening method..... 122**

**LIST OF ABBREVIATIONS**

UB - Ubiquitin

E1 - Ubiquitin activating enzyme

E2 - Ubiquitin conjugating enzyme

E3 - Ubiquitin ligase

PTM - post-translational modifications

HECT - homologous to E6AP carboxyl terminus

RING - Really Interesting New Gene

RBR - Ring between RING

Ab - Antibody

DUB - deubiquitinating enzyme

UBL - ubiquitin-like protein

RNF38 - RING finger protein 38

GEF - guanine nucleotide exchange factor

UBL - ubiquitin-like domain

UPD - unique parkin domain

CRL- Cullin RING Ligase

SCF complex - Skp1-cullin1-F-box

APC/C - anaphase-promoting complex/cyclosome

MAPK3 - mitogen-activated protein kinase 3

PRMT1 - protein arginine N-methyltransferase 1

PPP3CA - protein phosphatase 3 catalytic subunit alpha

PGAM5 - phosphoglycerate mutase 5

OTUB1 - ovarian tumor domain containing ubiquitin aldehyde binding protein 1

FACS - fluorescence-activated cell sorting

MOI - multiplicity of infection

NK - natural killer

RTK - receptor tyrosine kinases

PIRH2 - p53-induced protein with a RING-H2 domain

CBL-B - Casitas B lymphoma-b

polH - DNA polymerase Eta

HCC - hepatocellular carcinoma

NSCLC - non-small cell lung cancer

CRC - colorectal cancer

GC - gastric cancer

HBT-xUb - His-Biotin-xUb

MS - mass spectrometry

ELAV - embryonic lethal abnormal vision

SPFQ - Splicing factor, proline- and glutamine-rich

PCNA - proliferating cell nuclear antigen

HHARI - human homolog of Ariadne

ATM - ataxia telangiectasia mutant

PD-L1 - Programmed death-ligand 1

SEC31A - Protein transport protein Sec31A

NSF - N-ethylmaleimide sensitive factor

PPM1G - Protein phosphatase 1G

OS - osteosarcoma

PGM1 - Phosphoglucomutase 1

KPNA2 - Karyopherin  $\alpha 2$

OGT - O-Linked N-Acetyl Glucosamine Transferase

AS - Angelman Syndrome

PROTAC - proteolysis-targeting chimeras

HPV - human papillomavirus

GPS - global protein stability profiling

OUT - orthogonal ubiquitin transfer

xUB - engineered UB

xE1- engineered E1

AD - adenylation domain

UFD - ubiquitin fold domain

UBA-UB associated domains

fE1- forward mutation in AD on E1

bE1- three mutations in UFD of E1

Bio-UB – biotinylated UB

ELISA – enzyme-linked immunosorbent assay

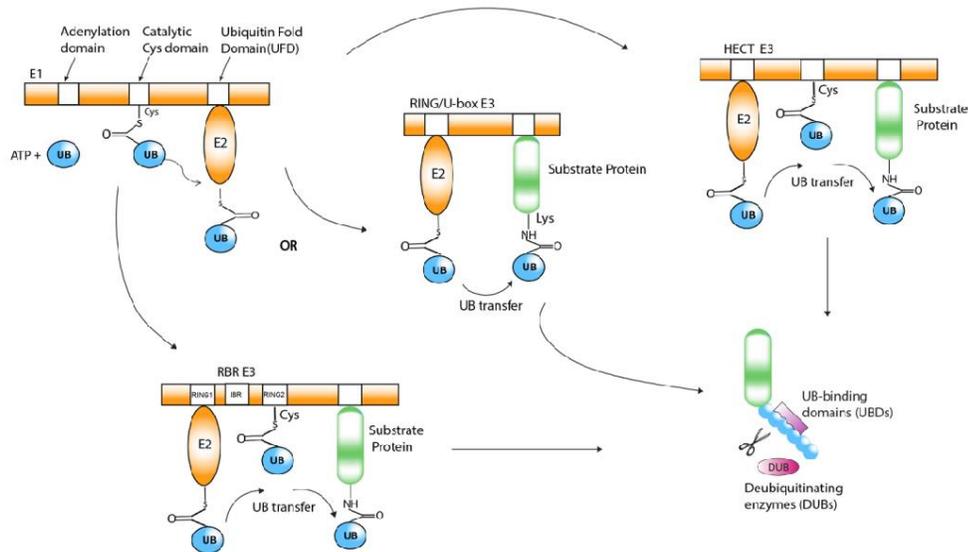
## 1 INTRODUCTION

### 1.1 Background

In the eukaryotic cell, there are many post-translational modifications( PTM ) such as ubiquitination, phosphorylation, glycosylation, methylation, acetylation and lipidation<sup>[1]</sup>. In 1977, Ubiquitination was discovered for modifying histones in the cell<sup>[2]</sup>. In 2004, three scientists, Aaron Ciechanover, Avram Hershko and Irwin Rose were awarded the Nobel Prize in Chemistry for their significant contributions in the field of ubiquitination. Furthermore, ubiquitination was studied as a protein modifier in the cell to initiate the signal of proteolysis by the 26S proteasome in mid-90s<sup>[3]</sup>. Although protein modification by ubiquitin is best known for targeting the modified proteins for proteasomal degradation to serve as a crucial regulator of cellular signaling pathways, it is also widely recognized for its roles as non-degradative signals to control cellular processes including endocytosis<sup>[4]</sup> and lysosomal targeting, subcellular localization of proteins, autophagy, DNA repair, and kinase activation<sup>[3,5]</sup>.

Ubiquitin (UB) is a highly conserved regulatory protein with 76 amino acids that may alter other proteins to drive a series of cellular signal transduction processes via an E1-E2-E3 enzymatic cascade<sup>[6]</sup>. The E1, E2, E3, and ubiquitin (UB) are required for ubiquitination to occur in a very precise manner in the eukaryotic cell and form large networks of cell signaling pathways affecting almost all proteins in the eukaryotic cell<sup>[7]</sup>. Briefly, UB is initially activated by the ubiquitin activating enzyme (E1), which forms a thioester bond between the catalytic Cys of E1 and the C-terminal Gly of UB. Uba1 and Uba6 are the two types of E1 encoded in the human genome. The UB linked to E1 is then loaded onto a UB conjugate enzyme (E2) via a thioester exchange process,

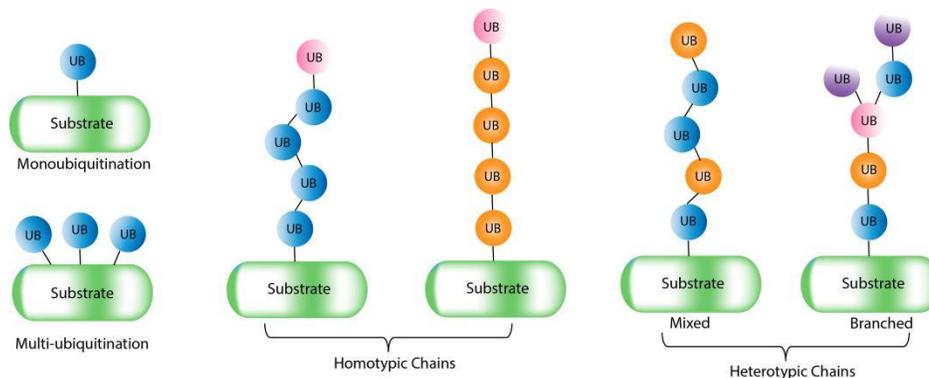
resulting in the formation of a UB-E2 conjugation. Finally, the UB ligase (E3) binds the UB-E2 conjugate to the substrate proteins that have been linked to the E3s via UB transfer (**Figure 1.1**).



**Figure 1.1 Native ubiquitin transfer cascade for proteasomal degradation** The UB transfer cascade is comprised of the UB-activating enzymes E1, UB conjugating enzymes E2, and UB ligases E3. E3s are divided into HECT and RBR UB ligases, which include catalytic Cys residues that catalyze the formation of thioester intermediates with UB, and RING-type ligases, which include the RING finger and U-box–containing protein families. E3s of the RING type activate E2; UB, directly transferring UB to substrate proteins. Autoubiquitination of E3s is also possible (shown for RING-type E3s). DUBs identify and cleave or alter UB chains placed on substrate proteins. Additionally, chains are identified by UBD-containing proteins, which recruit additional cellular partner proteins in many circumstances.

In most ubiquitination processes, E3 ligases determine substrate specificity and UB linkage type<sup>[8]</sup>. Ubiquitination can be classified into three different types based on its characteristics: monoubiquitination, polyubiquitination, and branched ubiquitination (**Figure 1.2**). Substrate proteins attached by a single UB molecule or multiple UB monomers are referred as the mono-ubiquitination or multi-ubiquitination<sup>[9]</sup>, respectively. In contrast, polyubiquitination refers to two

or more UB molecules are jointly loaded on the protein targets via the formation of an amide bond with the C-terminal carboxyl groups of UB<sup>[10,11]</sup>. Polyubiquitination linkage can be further divided into two categories: homotypic chains and heterotypic chains<sup>[12]</sup>. UB chains linked through the same lysine site of each UB are called homotypic, whereas heterotypic chains are linked through different lysine sites of UB molecules (K6, K11, K27, K29, K33, K48 and K63) or through its N-terminal Met<sup>[12]</sup>. Mixed chains are heterotypic chains that include more than one kind of connection, yet each UB monomer within the chain is linked through only one acceptor site. Branched chains, on the other hand, are made up of one or more UB subunits that are concurrently connected through least two distinct acceptor sites.



**Figure 1.2 Different modes of ubiquitination**

The diversity in chain positions, linkages, and topologies serves different purposes as different signals within the cell. For example, the well-studied K48-linked chains function as a degradation signal for protein substrates to be transferred to the 26S proteasome<sup>[11]</sup>. K6, K27 and M-1 linked chains are involved in DNA damage repair, NF- $\kappa$ B signaling and mitochondrial autophagy<sup>[13-15]</sup>; K11 polyubiquitination is involved in cell cycle and trafficking events<sup>[16]</sup>; K29 polyubiquitination modulates UFD-mediated protein degradation<sup>[17]</sup>; and K33 polyubiquitination

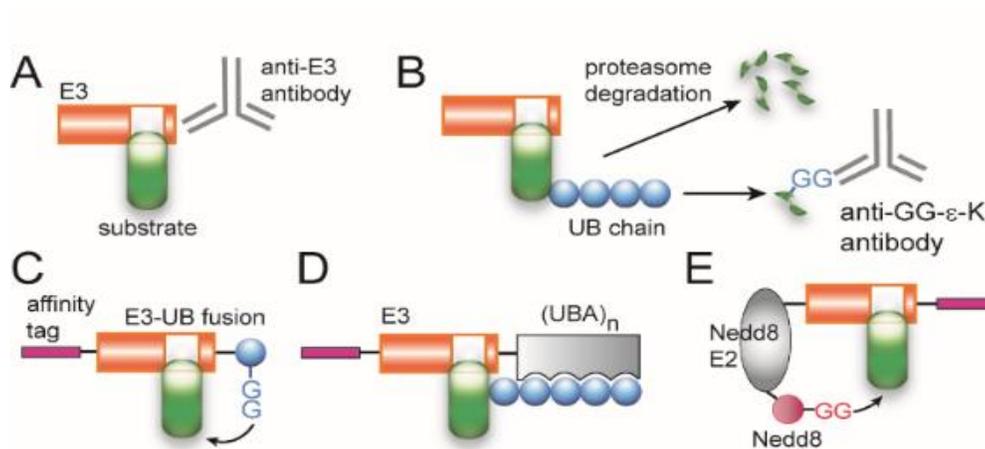
is involved in Toll receptor-mediated signaling pathways<sup>[18]</sup>. K63 polyubiquitination generally participates in protein-protein interactions<sup>[19,20]</sup>, protein activity, and protein transport, therefore controlling a variety of biological processes. The K63 ubiquitination modification is generally mediated by a specific E3 ligase.

E3s are thought to have a vast number of substrate proteins and have important regulatory roles in the eukaryotic cell via the E1-E2-E3 cascades. Furthermore, E3 dysfunction has been linked to a variety of illnesses, including cancer<sup>[21,22]</sup>. Meanwhile, there is a cross-reactivity among E3s in that most E3s share the same substrates on distinct binding residues and/or with different linkages. Profiling the substrates of E3s would thus be very useful in elucidating their biological roles.

## 1.2 The challenge of substrate identification

Many advanced approaches for identifying the substrates of E3s have been developed and improved over time (**Figure 1.3**). One strategy focuses on the changes in cellular protein stability caused by E3 disruption *in vivo*, taking advantage of the fact that the proteasomal degradation signal can be tracked<sup>[23]</sup>. Cullin E3s have been reported to use a method called global protein stability profiling (GPS) to identify their substrates<sup>[24]</sup>. Changing the activity of one E3, on the other hand, may influence the activity of other E3s, resulting in the cross regulation of the ubiquitination levels of the substrates among various E3. Another approach is to select possible substrates based on their high binding affinity for specific E3s. The yeast two-hybrid technique is one of the examples<sup>[25]</sup>. Furthermore, the binding affinity technique was employed to evaluate candidate targets using a protein microarray *in vitro*<sup>[26]</sup>. However, an affinity-based method is unable to distinguish some adapter or regulator proteins that can bind to E3 and perform distinct

tasks. Trapping substrates with E3s fused to UB or UB associated domains (UBA) is another approach of identifying substrates [27,28]. The E3s are turned into a trap in these fusions so that the substrates are coupled to E3s following UB transfer, allowing the substrates connected to E3s to be purified and further identified. However, because UB and UBA can function as binding motifs in many protein-protein interactions, this method may introduce some artificial identification of E3 substrates.

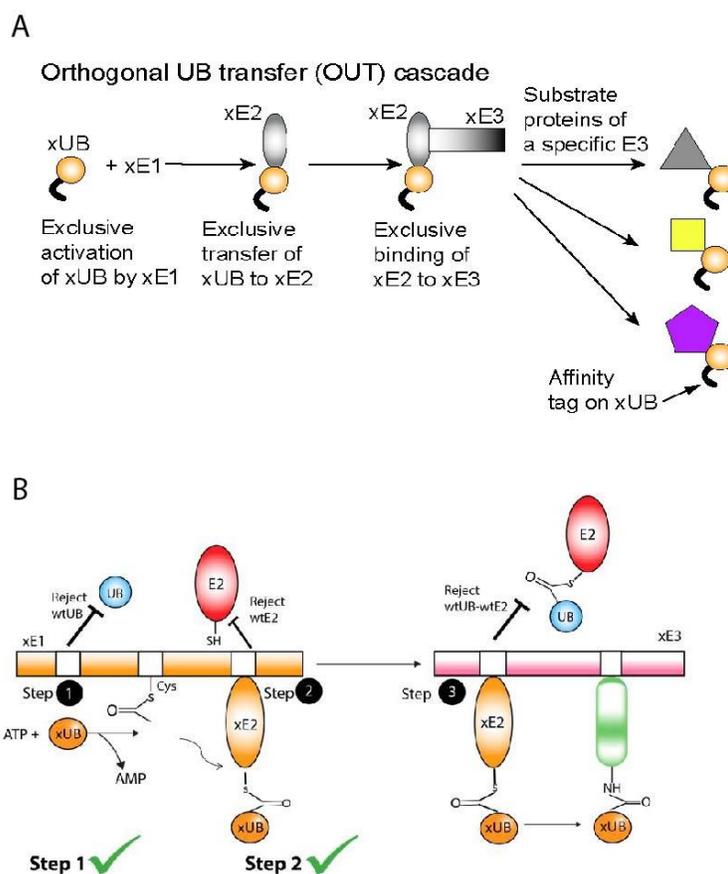


**Figure 1.3 Current Approaches for E3 substrate identification.** (A) Affinity purification of E3-substrate complexes. (B) Perturbation of E3 activity to read changes in protein degradation or ubiquitination level. After trypsin digestion, peptides covering the ubiquitination sites can be enriched by an anti-GG-ε-K antibody that binds to diGly conjugated Lys. (C) E3-UB fusion to react with substrates bound to E3 in cis (UBAIT). (D) E3-UBA fusion to bind to substrate with an extended UB chain. (E) Nedd8 E2 fused to E3 to transfer Nedd8 to E3 substrates (Neddylator).

### 1.3 The rationale and past development of orthogonal ubiquitin transfer (OUT)

In our laboratory, we proposed and developed an orthogonal ubiquitin transfer (OUT) technique for identifying the direct protein substrates of specific E3 ligases. We devise a unique

pathway in which an engineered UB variant (xUB) would be exclusively loaded onto the substrates of an E3 through the only feasible pathway comprising of engineered E1 (xE1), E2 (xE2), and E3 (xE3) enzymes. After that, we over-expressed the OUT cascade of xE1-xE2-xE3 in the cell and isolate the xUB-conjugated proteins that would be possible substrates of the corresponding E3s (**Figure 1.4**). Because the enzymes participating in the OUT cascade are only active with their designed partners, xUB transfer through the OUT cascade is free of cross-reactions with wild-type E2s and E3s. Using the well-developed OUT cascade, our lab has successfully identified direct substrates for HECT-type E3 E6AP, and U-box-type E3s E4B, and CHIP<sup>[29,30]</sup>.

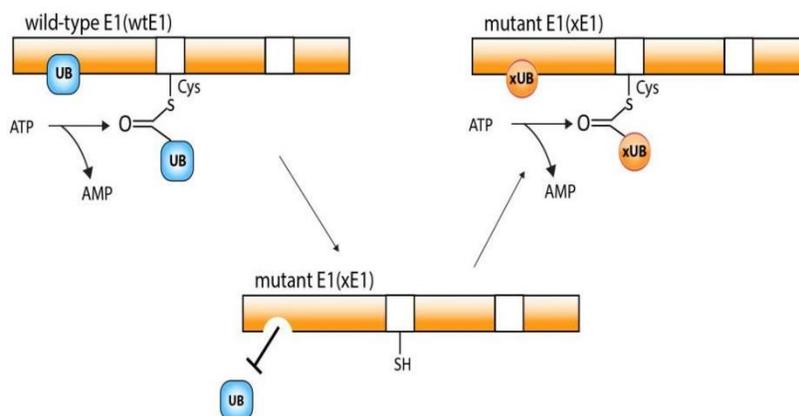


**Figure 1.4 The OUT (orthogonal ubiquitin transfer) cascade. (A)** A proposed scheme of Orthogonal UB transfer (OUT) cascade; **(B)** There are three steps to creating an OUT cascade. To generate the xUB-xE1 conjugate, the first step is to create an xUB-xE1 pair that can only activate xUB. The second stage is to make an xE1-xE2 pair that can only transfer xUB to a specially

constructed E2. Finally, the xE2-xE3 pair must be engineered to transfer xUB to the xE3 substrate proteins.

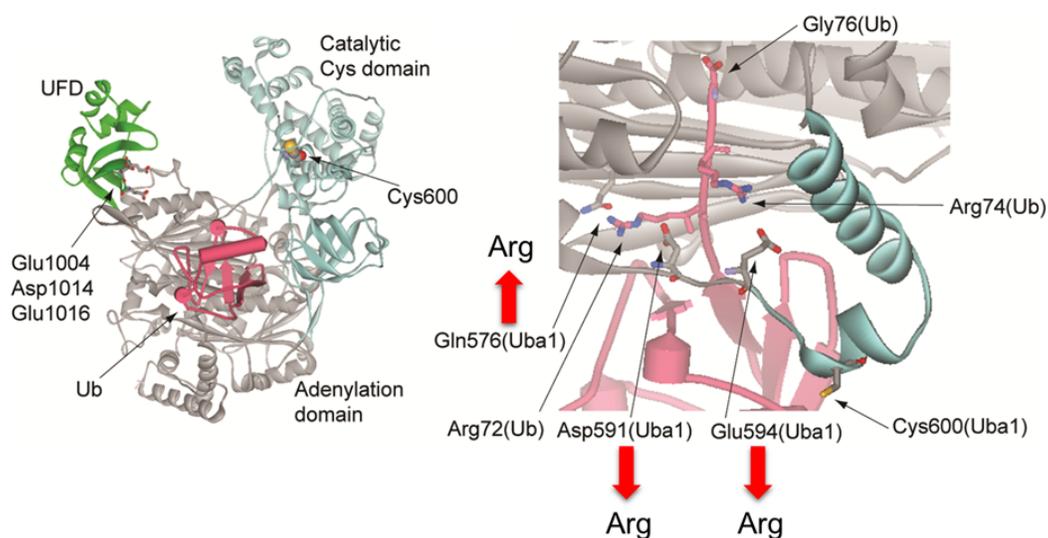
As previously reported, our group has developed the concept and engineered enzyme cascades for orthogonal ubiquitin transfer (OUT) as a way for uncovering the extremely complex ubiquitin enzyme network<sup>[31]</sup>. Our goal was to design an engineered ubiquitin transfer pathway that utilizes only engineered E1, E2, and E3 enzymes, while retaining the substrate-binding domains of the E3 enzymes, to identify native substrates of the E3s. All of the enzymes involved were engineered so that they are exclusively active with their engineered partners and inactive with their wild-type partners to make the whole pathway orthogonal<sup>[32]</sup>.

Essentially, constructing the complete OUT cascade requires the engineering of three protein interface interactions: those between UB and E1, between E1 and E2, and between E2 and E3. When working with those enzymes, we tried to find a way to modify an interaction interface that would eliminate the enzymes' activity to function with their wild-type companions while yet allowing the enzymes to function exclusively with one another. The “bump-and-hole” strategy has traditionally been applied to achieve this goal<sup>[31]</sup> (**Figure 1.5**). Basically, the charge reversal between several vital residues of the enzymes was used<sup>[33,34]</sup>. Previous study on the engineering of interface between UbcH5b (E2 conjugating enzyme) and CNOT4 (RING-type E3 ligase) by Winkler and Timmers demonstrated that a key lysine residue in Ubch5b was mutated to reversed charge<sup>[35,36]</sup>. Then, its activity with wild-type CNOT4 was abolished. However, the activity was restored when the aspartate was mutated to glutamate via the charge exchange in CNOT4.



**Figure 1.5 A scheme for bump and hole.** Engineering wt E1 by introducing mutations that removes original wild-type activity, while simultaneously incorporating mutations on wt UB to restore the interaction with engineered E1 by electrostatic attraction; similar results could conceivably also be achieved utilizing steric effects or hydrophobic and hydrophilic interactions instead. Finally, the engineered mutants would interact with each other without reactivity with native enzymes.

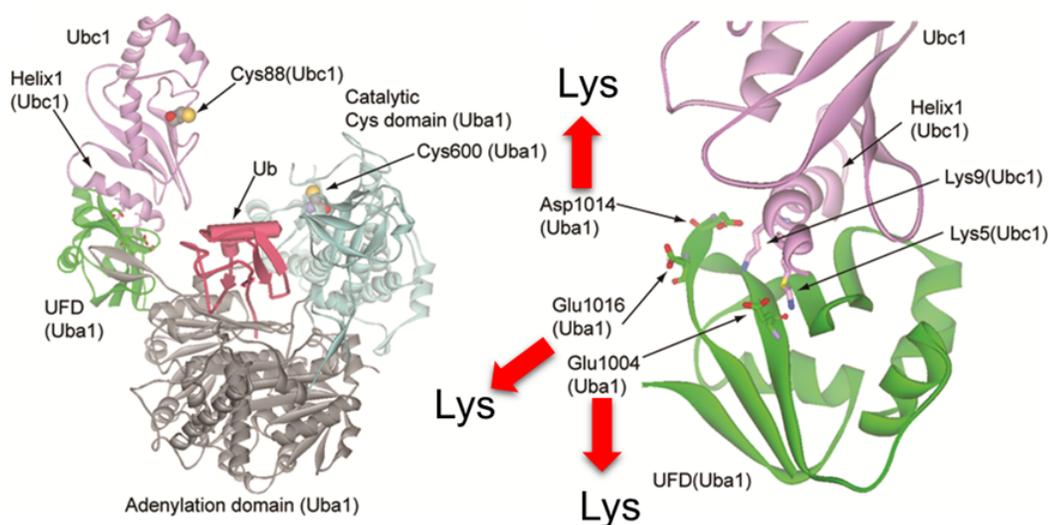
Using the bump and hole approach, we first introduced three mutations (Q576R, S589R, and D591R) into yeast E1's adenylation domain to prevent it from interacting with wt UB relying on a crystal structure of UB~E1 solved by Lee and Schindelin<sup>[37]</sup> (**Figure 1.6**). To restore the activity of engineered E1 with UB, we utilized a combination of phage selection and site-directed mutagenesis to induce complementary mutations (R42E and R72E) into wild-type UB, which allows engineered E1 to activate the engineered UB in the OUT cascade. Furthermore, we generated human xUba1 by introducing mutations Q608R, S621R, and D623R into the adenylation domain of human Uba1, which would similarly exhibit orthogonal activity, based on sequence homology between the human and yeast E1s. Thus, the engineered E1 and UB were defined as xE1 and xUB, respectively.



**Figure 1.6 The Crystal structure of the yeast E1 Uba1 in complex with UB (PDB ID 3CMM).** The binding of yeast UB with Uba1 was modeled in the structure. Cys600 is the catalytic Cys residue for UB conjugation in Uba1. To initiate UB transfer, UB is first bound to the adenylation domain of Uba1. The energy of ATP hydrolysis is then used to form a thioester conjugate between the UB C-terminal carboxylate and the catalytic Cys residue of Uba1. The targeted residues (red arrow) in the adenylation domain of the Uba1 E1 enzyme were mutated to Arg in xE1; and the R42 and R72 in wt UB were mutated to Glu as xUB.

We used a similar approach in the development of an xE1-xE2 pair (**Figure 1.7**). The negatively charged residues E1004, D1014, and E1016 in yeast Uba1 have previously been found to be critical for interaction with E2. By the charge reversal method, these three residues were mutated to lysine with positive charges to block its interaction with the native Ubc1 (yeast E2 enzyme)<sup>[31]</sup>. Then, a phage selection was carried out to identify compatible mutations in wild-type Ubc1 to restore interaction with xUba1(UFD). Correspondingly, we introduced mutations E1037K, D1047K, and E1049K into the UB fold domain (UFD) of human Uba1<sup>[31]</sup>. This human xUba1 stimulates xUB, resulting in the formation of xUB~xUba1 thioester conjugates. We also inserted the K4E and K8E mutations into UbcH5b<sup>[29]</sup>, which were derived from the mutations in yeast

xUbc1, and confirmed that the UbcH5b mutant (xUbcH5b) was capable of accepting xUB from xUba1. Thus, xUB could not be activated by wild-type human Uba1, and wt UB could not be activated by xUba1. Additionally, due to the modifications to the UFD domain, xUB could not be transferred from xUba1 to wild-type UbcH5b. Consequently, the human xUba1-xUbcH5b pair for xUB transfer is completely orthogonal to the native E1-E2 pair for the transfer of wt UB.



**Figure 1.7 The interface of the yeast E1 Uba1 and Ubc1.** Mutations E1004K, D1014K, and E1016K were inserted into the UFD domain of yeast Uba1 to create orthogonal xE1-xE2 pair (UFD). The UFD domain mutations prevented the binding of the N-terminal helix of the E2 enzymes. To reestablish E2 contact with xUba1(UFD), the mutations K5D, R6E, K9E, E10Q, and Q12L were introduced. The newly generated xUba1 mutant would transfer xUB to xUbc1 for the formation of xUBxUbc1 conjugate while xUba1 failed to transfer xUB to wt E2 enzymes. So, the xUba1-xUbc1 pair was orthogonal to the wt E1-E2 pairs in xUB transfer.

#### 1.4 The overview of E3s and their characteristics

E3 ligases play a critical role in determining substrate proteins for ubiquitination because they act as substrate recruiters for the E1-E2-E3 cascade<sup>[38]</sup>. On the basis of their ubiquitin transfer mechanism and the domain of E3 for binding to the E2~UIB conjugate, E3 ligases may be further

subdivided into three primary subtypes according to their molecular structure and mechanism: homologous to E6-associated protein C-terminus (HECT), really fascinating new gene (RING), and RING-in-between-RING (RBR) E3s<sup>[39,40]</sup>.

#### 1.4.1 HECT E3 ligase

The HECT-type E3s form a thioester-linked intermediate with UB before the modifier is ligated to the substrate. As a consequence, the phase of UB transfer from E2 to E3 involves a distinct chemical transthiolation rather than thioester aminolysis<sup>[41]</sup>. The HECT E3s, which have about 30 members in the human genome, have prominent roles in protein trafficking, the immune response, and in several signaling pathways that regulate cellular growth and proliferation. They are comprised of a larger N-terminal lobe (N-lobe) containing the E2 binding domain and a catalytic cysteine-containing C-terminal lobe (C-lobe)<sup>[42]</sup>. The two lobes are linked by a flexible hinge area that enables the C-lobe to move around to assist the passage of UB from the E2 to the E3<sup>[42,43]</sup>. The HECT E3s can be classified into three major groups based on the domain organization found in the N-terminal portion of the proteins: the NEDD4 family, consisting of a membrane/lipid-binding C2 domain, two to four WW domains for substrate recognition, and a C-terminal HECT domain<sup>[44]</sup>; the HERC family, comprised of one or more regulators of chromatin condensation 1 (RCC)-like domains (RLD), which function as a guanine nucleotide exchange factor (GEF) for a small GTPase involved in membrane trafficking activities<sup>[45,46]</sup>; the other 13 HECTs lack distinct N-terminal domains and are therefore categorized as "other" HECT ligases<sup>[47]</sup>. Surprisingly, the specificity of ubiquitin chain linkage seems to be controlled by the C-lobe of the HECT domain rather than E2s<sup>[48]</sup>. For instance, Rsp5, a member of the NEDD4 family mainly generates K63 chains<sup>[49]</sup>, while E6AP, the first identified member of the HECT family<sup>[50]</sup>,

is a K48-specific enzyme and HUWE1 produces polyubiquitin chains with K6-, K11-, and K48-linked polymers<sup>[51]</sup>.

#### 1.4.2 RBR E3 ligase

RBR E3s were initially identified based on the conserved sequence alignments in eukaryotic organisms that predicted a tripartite motif composed of three Zn<sup>2+</sup>-binding domains: a RING1 domain that binds UB-E2 conjugate, a RING2 domain that catalyzes a transthioesterification process and accepts UB molecule from RING1, and an In-Between-RING (IBR) domain<sup>[52]</sup>. The ubiquitination mechanism by RBR E3 ligases acts as a 'RING/HECT hybrid' way<sup>[53]</sup> in which RING1 first binds the E2 (from the E2-Ub conjugate) in a RING E3-like manner and subsequently transfers the UB onto the catalytic cysteine in RING2 to form the reactive HECT-like E3-Ub thioester intermediate, which promotes ubiquitin transfer to the substrate.

Generally, it is believed that RBR E3 ligases exist in autoinhibited forms, where the inhibitory domains that is not a part of the RBR core limits the enzyme's ubiquitination activity<sup>[54]</sup>. In human genome, there are 14 family members of RBR E3 ligases<sup>[55]</sup>. To date, many crystal structures of RBR E3 ligases have been solved in the autoinhibition states<sup>[54]</sup>. Among them, Parkin<sup>[56]</sup> is one of best studied candidates that the RING2 active site cysteine is hidden by the unique parkin domain (UPD), while the E2 binding site on the RING1 domain is blocked by the ubiquitin-like domain (UBL) and the repressor (REP) element, thereby causing the inactive state of Parkin. To transform Parkin to active state, a kinase, called PINK1, from the mitochondrial outer membrane phosphorylates the Ser65 of both ubiquitin and Parkin's N-terminal ubiquitin-like (UBL) domain to induce a conformational change to expose the Cys of RING2<sup>[57,58]</sup>.

RBR E3 ligase malfunction has been linked to neurological disorders, infection, inflammation, and malignancy, according to emerging data<sup>[59,60]</sup>. RBR E3 ligases have been shown to play a key role in carcinogenesis by regulating the degradation of tumor promoters and suppressors in a variety of malignancies.

### 1.4.3 RING E3 ligase

The human genome contains over 600 E3 ligases, and the RING family, which is encoded by 270 human genes, is the most common category of ubiquitin ligase<sup>[61]</sup>. The consensus sequence for the RING finger domain is as follows: C-X<sub>2</sub>-C-X<sub>[9-39]</sub>-C-X<sub>[1-3]</sub>-H-X<sub>[2-3]</sub>-C-X<sub>2</sub>-C-X<sub>[4-48]</sub>-C-X<sub>2</sub>-C<sup>[62]</sup>, where X is any amino acid residue. They are distinguished by the presence of a zinc-binding domain known as RING (Really Interesting New Gene) or a U-box domain, which has the identical RING fold but lacks zinc. The RING and U-box domains bind E2-UB conjugate and stimulate the UB transfer. RING E3s act as a scaffold to direct the UB-charged E2 with respect to the substrate protein, mediating a direct transfer of UB to the substrate. RING domains exist in a wide variety of structural settings. A distinguishing characteristic of RING E3s is their tendency to form homodimers and heterodimers<sup>[62]</sup>. Homodimeric RING E3s such as cIAP, RNF4 and TRIM5 $\alpha$ , can typically bind two E2s (one for each monomer); Although homodimeric RING E3s have the intrinsic ability to engage functionally with E2s, this does not appear to be the case for other heterodimeric RING E3s. U-box domains, on the other hand, can function as monomers or homodimers.

There are RING E3s acting as multi-component assemblies. The Cullin RING Ligase (CRL) family is the most representative type. CRLs are a broad class of UB ligases with numerous characteristics in common, consisting of a scaffold protein (cullin), a bridging protein, a substrate

receptor protein, and a RING protein that recruits E2 ligases. CRL proteins is comprised of the following eight members: CRL1, CRL2, CRL3, CRL4A, CRL4B, CRL5, CRL7, and CRL9<sup>[63,64]</sup>. The SCF complex (Skp1-cullin1-F-box) is the best illustrative example of CRL family proteins, also referred as CRL3<sup>[64]</sup>, which is composed of stable components, including S-phase kinase-associated protein 1 (Skp1), ligase Rbx1, and Cullin 1 (Cul1). Cul1, which connects the Skp1 and Rbx1 domains, is the primary structural scaffold for the SCF complex. Skp1 is a connexin that forms the horseshoe complex with Cul1 and is required for F-box recognition and binding. Rbx1 features a zinc-binding domain termed the RING Finger that interacts with the E2-UB conjugate and transfers UB to the target protein's lysine residues<sup>[65]</sup>. As the SCF complex's most critical component, the F-box protein is responsible for substrate recognition and establishes the SCF complex's specificity<sup>[66]</sup>. The anaphase-promoting complex/cyclosome (APC/C) is another significant multi-subunit E3. It is a big assembly of 11-13 subunits, including a RING subunit (Apc11) and a cullin-like subunit (Apc2)<sup>[67,68]</sup>. Additionally, it has two interchangeable co-activator subunits, Cdc20 and Cdh1, each of which recognizes a specific substrate and is active during a distinct phase of the cell cycle. APC/C is important in mitosis because it initiates the shift from metaphase to anaphase<sup>[69]</sup>.

## 1.5 Discussion and future directions

In the past years, our lab studied and engineered the interface between E1-UB and E1-E2, resulting the success construction of orthogonal xE1-xE2 pairs for the transfer of xUB and the orthogonal pairs do not share any cross reactivity with wt E1, E2 and UB. To accomplish this goal, we used a combination of rational site-directed mutagenesis and activity-based screening of phage libraries. Additionally, the mutations we generated are exclusively reactive between the designed Ub (xUb), E1(xE1), and E2 (xE2) enzymes, while the wt UB, E1, and E2 remain inactive.

Moreover, the mutations are transferable to other E1s and E2s as long as their structures are comparable, allowing for a larger pool of E3s, and consequently more potential substrates identified by OUT cascade. Thus, the last step to complete the ubiquitination cycle is the engineering of interface between E2 and E3 (**Figure 1.4**), which becomes my priority in Dr. Yin's lab. Before I started to study the E3 ligases, I reviewed all E3 ligase categories and gained an understanding of the functions of E3 ligases in the cell. Interestingly, RNF38, my primary study target E3 ligase (RING-type), functions as either a tumor inducer or suppressor in a variety of cancer cell lines. It further entices me to investigate the function of RNF38 in humans and to clarify the biological significance via the identification of protein substrates using the OUT approach. Thus, the creation of the xE2-xE3 pair for RNF38 is the last step in completing the orthogonal ubiquitin transfer cascade in the cell, which enables the xUB to reach its ultimate destination through the formation of an amide bond between the xUB and substrates mediated by xRNF38. Chapter 2 and Chapter 3 will mainly focus on engineering of UbcH5b-RNF38 interface and the confirmation of RNF38 OUT activity both *in vitro* and *in vivo*. HHARI, an RBR-type E3 ligase was initially conducted by a postdoc in our lab, whose name is Geng Chen. With the completion of her phage selection work on screening out the xHHARI that can specifically transfer xUB *in vitro*, I took over her project to carry out the verification of HHARI OUT activity in cells and profile the substrate proteins via proteomics and co-IP analysis. The detailed introduction about HHARI is described in Chapter 4. Lastly, Chapter 5 will discuss an accomplished project, waiting for publishing. Due to the success of the xUba1-xUbcH7-xE6AP-xUB cascade, we collaborated with Professor Jianfeng Cai's group on designing a cyclic peptide to test the effect on the activity of E6AP via *in vitro* ubiquitination assays. Eventually, we used the verified substrates (UbxD8,

HHR23A and  $\beta$ -catenin) by our E6AP OUT cascade as downstream targets to further examine the efficiency of peptides both *in vitro* and *in vivo*.

## 2 ENGINEERING UBCH5B-RNF38 INTERFACE BY PHAGE SELECTION

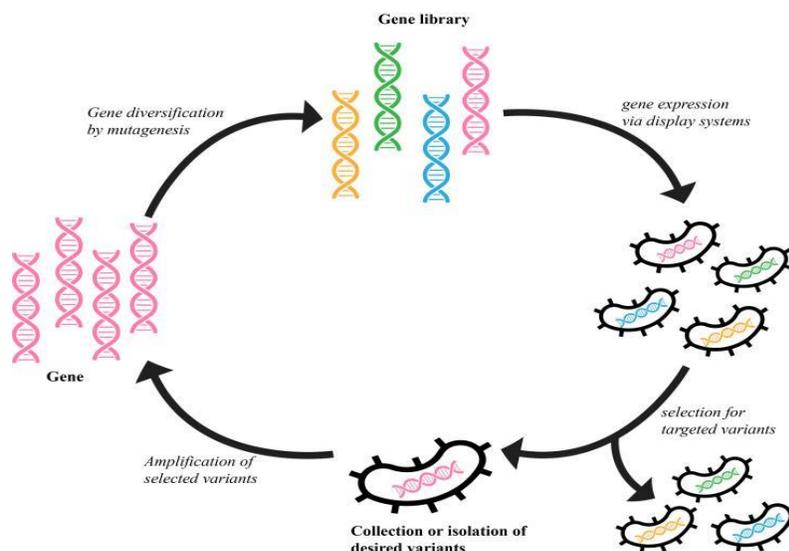
### 2.1 Rationale on the engineering of UbcH5b-RNF38 interface

As stated in previous chapter, our lab has successfully constructed the OUT cascade for CHIP, E4B and E6AP and used them to profile the substrate specificities of these E3s. By engineering the interface of the UbcH7-E6AP HECT domain via yeast surface display, we confirmed that MAPK1, CDK1, CDK4, PRMT5,  $\beta$ -catenin, and UbxD8 are all E6AP substrates. We discovered that E6AP may ubiquitinate either MAPK1 to stabilize it or CDK1, CDK4, PRMT5,  $\beta$ -catenin, and UbxD8 to facilitate their proteasomal destruction. E6AP-mediated ubiquitination and degradation of UbxD8 supports a function for E6AP in lipid metabolism, since UbxD8 controls lipid droplet size and abundance through a complex with p97/VCP<sup>[70]</sup>. Additionally, we screened xUbcH5b-xE4B and transplanted mutations from E4B into CHIP using another display technology known as phage display. As a result, the kinase MAPK3 (mitogen-activated protein kinase 3), the methyltransferase PRMT1 (protein arginine N-methyltransferase 1), and the phosphatase PPP3CA (protein phosphatase 3 catalytic subunit alpha) were discovered as common substrates of the two U-box E3s: CHIP and E4B. Importantly, E4B substrates PGAM5 (phosphoglycerate mutase 5) and OTUB1 (ovarian tumor domain containing ubiquitin aldehyde binding protein 1) were validated, as were CHIP substrates  $\beta$ -catenin and CDK4 (cyclin-dependent kinase 4). Moreover, we demonstrated that CHIP promotes CDK4 degradation in response to endoplasmic reticulum stress using the CHIP-CDK4 circuit described by OUT. Similar to U-box-type E3s, RING-type RNF38 acts as a bridge to recruit substrates and subsequently facilitates UB transfer to them. As a consequence of the success of the U-box CHIP and E4B OUT cascades, a direct evolution strategy is applied to modify the interface between UbcH5b and RNF38 for the construction of an OUT

cascade of Ring type E3 RNF38 to profile its substrate specificity and elucidate its functions in the cell.

## 2.2 Direct Evolution Technique

A directed evolution is a laboratory-based simulation of the natural evolutionary cycle. The cycle requires three stages<sup>[71,72]</sup>: 1) random mutagenesis and/or gene recombination to create a wide library of variations. Without comprehensive structural or function knowledge, random mutagenesis of a single gene would enable sparse sampling of the sequence space to find the key spots that are strongly associated with the desired protein characteristic. The more diverse the library, the more likely it will improve the variants with helpful or ideal functional spots<sup>[73]</sup>. Targeted libraries with a high ratio of beneficial to detrimental mutants often serve as a good starting point for the next round of screen or selection. *Error-prone* PCR and DNA shuffling techniques are the most popular approaches used for library construction of variants; 2) a suitable screening/selection method is used to obtain functional variants, and non-functional or useless variations are eliminated via plasmid, phage, or ribosome display, growth complementation, and reporter-based methods<sup>[71,74]</sup>. This technique can evaluate a huge library comprising more than  $10^9$  variations and distinguishing the proteins with the required characteristics. 3) In the final step, the optimal mutant sequence or pool of mutant sequences is selected for gene amplification. Multiple rounds including the whole mutagenesis, screening, selection, and gene amplification cycle are repeated using the enriched sequence mutants until the desired characteristics are identified<sup>[75]</sup>, as specified by the protein engineering objective (**Figure 2.1**).



**Figure 2.1 Directed evolution overview.** Shown here is one possible way of conducting an experiment based on directed evolution; note the aforementioned three important factors: variation (mutagenesis), identification and enrichment (selection), and heredity (amplification of enriched clones).

To create the mutant library, *in vitro* mutagenesis is the most often used method for efficient and well-controlled gene diversification. *Error-prone* PCR is one of *in vitro* gene mutagenesis methods by which random mutants may be inserted into any piece of DNA<sup>[76]</sup>. It doesn't require the gene structure and function information. Thus, when structural knowledge of the target protein is unclear, or when the impact of specific mutations throughout the whole gene need to be studied, error-prone PCR becomes a good choice, which utilizes a low-fidelity DNA polymerase such as Taq DNA polymerase (a non-proofreading enzyme) to produce mutations during PCR amplifications performed under non-standard conditions<sup>[77]</sup>. By altering the composition of the reaction buffer, such as changing the concentrations of  $MgCl_2$  ions, the reaction was supplemented with  $MnCl_2$  ions, unbalanced ratio of dNTP concentrations, or polymerase extension time, the base-pairing fidelity can be lowered that leads to increased mutation rates<sup>[78]</sup>. In these

circumstances, the polymerase makes mistakes in base pairing during DNA synthesis, causing errors in the newly produced complementary DNA strand. Consequently, the random mutagenesis library is constructed. This technique generates libraries with a high number of AG and TC transitions, which bias the resultant sequences toward high GC content<sup>[76]</sup>. Using a combination of Taq polymerase and Mutazyme DNA polymerase, which preferentially produces GC-AT transitions and GC-TA transversion mutations, a more balanced mutation distribution may be achieved<sup>[79]</sup>. In addition to studying unknown structural information about the protein of interest, most critical residues or regions in the protein of interest have already been defined via existing mutagenesis studies or crystal structures, in which the specific mutations identified can be utilized for targeting. Thus, the concept of focus mutagenesis is introduced<sup>[80]</sup>, which requires synthetic DNA oligonucleotides carrying one or more degenerate codons at the target residues to introduce mutations during oligo annealing or PCR<sup>[80]</sup>. PCR-based mutagenesis combined with either restriction enzyme-based cloning or other elegant DNA assembly methods is one of focus mutagenesis methods to generate a mutant library, which we have used to engineer the interface between xE1-xUB and xE1-xE2 in the previous studies.

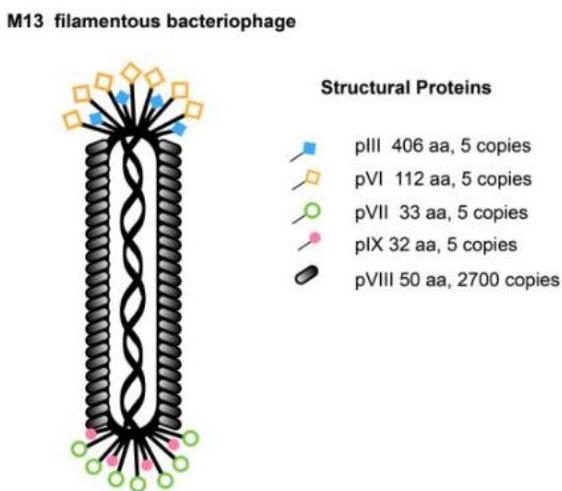
Following the successful creation of a library with more than a billion variations via mutagenesis, an effective selection/screening technique is required in order to identify the phenotype with desired characteristics and to allow for the enrichment of functional variants to be identified<sup>[81]</sup>. Currently, hundreds of screening or selection assays have been reported<sup>[72]</sup>, among them, the major principle of selection techniques would apply selective pressures on the protein libraries, allowing functional variants to be enriched for the subsequent rounds. On one hand, yeast display is one of the selection methods used for screening<sup>[82]</sup>. The protein library is expressed and displayed as a fusion to the Aga2p protein on the surface of yeast, then flow cytometry (also known

as fluorescence-activated cell sorting, or FACS) is utilized to identify yeast expressing desired proteins via the treatment with fluorescent probes that mark the cells differently according to the biochemical and biophysical characteristics of the exhibited protein<sup>[83]</sup>. Lastly, the yeast cells are screened through a FACS instrument's fluidics stream, and the functional mutants can be enriched. On the other hand, phage display with biopanning assay introduced by George P. Smith<sup>[84]</sup> has been also widely applied in the laboratory. In general, the library variants are displayed on filamentous phage by fusing the proteins or peptides of interest to pIII or pVIII gene of filamentous phage<sup>[85]</sup>. Random peptide libraries with phage-displayed phenotypes provide functional access to the proteins or peptides and establish a physical connection between the phenotype (the displayed protein fragment/peptide) and genotype (the encoded DNA), which needs to conduct a screening process that can identify the functional binders by affinity selection<sup>[72,86]</sup>. The affinity selection, also known as biopanning, is that a phage displayed library is incubated with an immobilized target, then the functional phage variants bound to the immobilized target can be kept, whereas the non-active phage variants with lower binding affinity would be washed away<sup>[86]</sup>. Acid or high concentration of salt or specific enzyme cleavage reaction (such as TEV cleavage) can elute the active phages for infecting bacterial host cells in order to enrich and amplify the high affinity binders<sup>[87]</sup>. Typically, conducting the phage selection through three or more rounds is recommended to screen out the targets with strong affinity binding.

In summary, both yeast display and phage display are useful and recommended to be applied for protein engineering, including protein-protein, protein-peptide, and protein-DNA interactions that uses a genotype-phenotype linkage between the displayed peptide or protein and the encoded DNA. Depending on the properties of the target proteins, we can screen them for protein engineering using either phage display or yeast display.

### 2.3 Introduction of M13 Phage

M-13 phage is one of the Ff phages (**Figure 2.2**), which are members of the bacteriophage family filamentous (inovirus) and can be generated via propagation in *Escherichia coli* (*E. coli*) carrying an F-pilus. It is made up of circular single-stranded DNA (ssDNA) that is about 6407 nucleotides long and contains 2700 copies of the main coat protein (pVIII) and five copies of each of the four minor coat proteins (pIII and pVI at one end and pVII and pIX at the other end)<sup>[88,89]</sup>. M13 phage is a non-lytic virus approximately 6.6 nm in diameter, 880 nm in length, which indicates that it is capable of releasing offspring without killing the host cells<sup>[90]</sup>. Infected cells exhibit a reduction in their rate of growth, resulting in turbid plaques in *E. coli* lawns.



**Figure 2.2 Schematic diagram of the structure of an M-13 filamentous bacteriophage.** The circular ssDNA is encapsulated by a proteinaceous coat. The coat comprises five types of structural proteins and six types of non-structural proteins

A virus's life cycle is often divided into four stages: infection, replication of the viral genome, assembly of new viral particles, and finally discharge of offspring particles from the host<sup>[91]</sup>. Upon M13 pIII invasion of host cells, the phage genome is subsequently transmitted to the bacterial cell's cytoplasm<sup>[92]</sup>, where resident proteins transform it to a double-stranded replicative form. This DNA then acts as a template for the phage genes to be expressed.

The next stage of the phage life cycle is called genome amplification. pII nicks the double-stranded form of the genome in order to initiate (+) strand replication. Host enzymes copy the replicated (+) strand, resulting in an increase in the number of double-stranded phage DNA copies. Then, pIV, pI, and pXI are involved in phage maturation<sup>[93]</sup>. The outer membrane forms a stable structure with several copies of pIV<sup>[94]</sup>. On the other hand, the C-terminal sections of pI and pXI interact with the N-terminal region of pIV in the periplasm, indicating that they are both involved in membrane assembly. The pI, pXI, and pIV complex together produce secretion channels for mature phage from the bacterial host. Additionally, two minor phage coat proteins, pIX and pVII<sup>[95]</sup>, are believed to engage with the pV-single stranded DNA complex at a portion of the DNA termed the packing sequence to begin phage secretion. The pV proteins that cover the single stranded DNA are then replaced by pVIII proteins in the bacterial membrane<sup>[96]</sup>, and the expanding phage filament is threaded through the pI, pXI, and pIV channels. Once the phage DNA is completely coated with pVIII, the secretion is completed by the addition of the pIII/pVI cap<sup>[97,98]</sup>, at which point the new phage detaches from the bacterial surface. The life cycle of phage will repeat continuously.

#### **2.4 Property of Phagemid vectors**

A phagemid is a plasmid that includes a F1 phage replication origin<sup>[99]</sup>. Typically, the phagemid can be utilized as a DNA cloning vector, which includes the bacterial replication origin, the selective marker (antibiotic resistance gene), a gene of a phage coat protein used to display foreign proteins, restriction enzyme recognition sites for cloning, a transcription promoter and a DNA segment encoding a signal peptide<sup>[100]</sup>. A phagemid may be packed as single-stranded DNA in viral particles or reproduced as a plasmid and phagemids have a double-stranded replication origin (ori) as well as a single-stranded replication and packing origin (f1 ori)<sup>[100]</sup>.

The phagemid DNA can be replicated independently of phage production for mutagenesis and amplification<sup>[101]</sup>. However, in order to display the protein on the phage, the remaining phage gene activities must be achieved by the infection of helper phage<sup>[99]</sup>. Due to the fact that the helper phage has its own origins of minus or plus strands, the replication proteins affect not only the phagemid but also the helper phage<sup>[100]</sup>. However, due to the lack of origins or replication in the helper phage, the virions are produced with lower efficiency than the phagemid itself<sup>[100]</sup>. Thus, when *E. coli* cells are transformed with the phagemid are cultured under selective resistance and infected with the helper phage at an acceptable multiplicity of infection (MOI), both phagemid and helper phage can replicate, but mature virions contain phagemid predominate<sup>[101]</sup>. Typically, pIII plays a central role in the development of phage display. It is widely used to display the protein of interest fused with pIII because there are less than five copies of fusion proteins in the progeny phages due to structural properties of the Fφ phages<sup>[102]</sup>. The presence of numerous copies of the displayed proteins would more likely cause the avidity effect rather than the affinity effect, which makes it difficult to discern between proteins with varying affinities. Therefore, the pIII-type phagemid is used to reduce the avidity effect during the selection<sup>[100]</sup>. In addition, the N-terminal domain of coat protein pIII is crucial for infecting host cells. If pIII-type phagemids are employed, fusion proteins containing the whole protein III propagated in the host cell may inhibit host cell F pili regeneration and hinder helper phage infectivity<sup>[103]</sup>. Overall, the phagemid are used with many advantages<sup>[103,104]</sup>: 1) phagemid genomes are smaller and can accept bigger DNA fragments; 2) phagemids are faster at transformation, allowing for a more diverse phage display library; 3) phagemids' genomes include a variety of restriction enzyme recognition sites useful for DNA recombination and gene engineering; 4) fusion

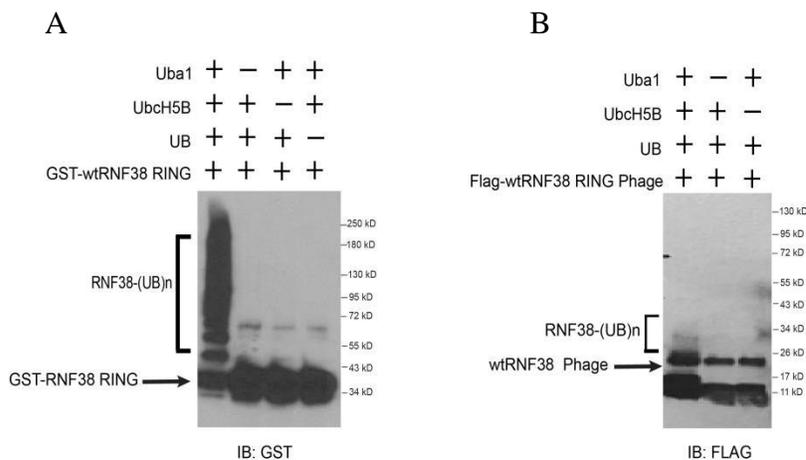
protein expression can be readily regulated;5) The genetic stability of phagemids is superior than recombinant phages.

With the detailed introduction of phage structure and biological function, pComb3H is one of mostly used phagemids for phage display<sup>[105]</sup>. The gene of interest is cloned into the N-terminal of pIII gene of phagemid pComb, and then transformed into bacteria for the propagation of phage displaying the protein fused with pIII encoded by the target gene and pIII gene, which has been applied in our previous E1-E2 interface engineering<sup>[31,32]</sup>. Due to the success of display system on phage, we still choose pComb vector as our phagemid to display the RNF38 Ring domain and further conduct phage selection via biopanning assay to engineer the interface between UbcH5b(E2)-RNF38 Ring(E3).

## 2.5 Results

### 2.5.1 Phage display of the RING domain of RNF38

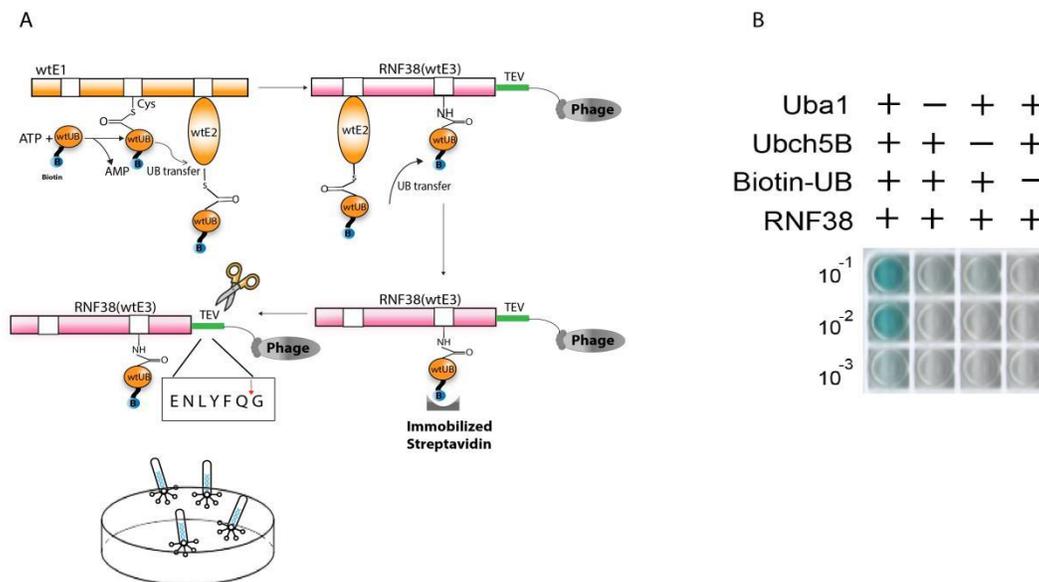
The RNF38 Ring domain in pGEX vector was expressed in *E.coli* system and the activity was reconstituted by in vitro autoubiquitination assay(**Figure 2.3A**). The RNF38 Ring domain of the E3 ligase RNF38 was cloned into the phagemid display vector pComb3H, with a FLAG epitope tag linked to the protein's N-terminus and a TEV cleave site fused to the RNF38 gene's C-terminus. Then, I prepared phages displaying the RNF38 Ring domain with a fused FLAG epitope and used Western blotting to prove that the Ring domain could be well displayed on the M13 phage. The reaction of the Ring domain displayed on phage with wt UB, Uba1 (E1), and UbcH5b (E2) resulted in the formation of polyUB chains (**Figure 2.3B**). This indicates that the RNF38 Ring's catalytic activity is retained on the phage.



**Figure 2.3** *In vitro* auto-ubiquitination assays of RNF38 RING. **(A)** Recombinant RNF38 Ring domain protein was tested for auto-ubiquitination activity with wild-type UB, Uba1(E1) and UbcH5b(E2) enzymes. The ubiquitin-conjugated RNF38 Ring protein was immunoblotted by anti-GST antibody. **(B)** Phage surface-displayed RNF38 Ring with FLAG epitope displayed by pComb3H system were tested for auto-ubiquitination activity with wtUB, Uba1(E1) and UbcH5b(E2) enzymes. The ubiquitin-conjugated RNF38 Ring protein and Ring phage was probed with anti-FLAG antibody.

### 2.5.2 Model selection for RNF38 Ring-displaying-phage

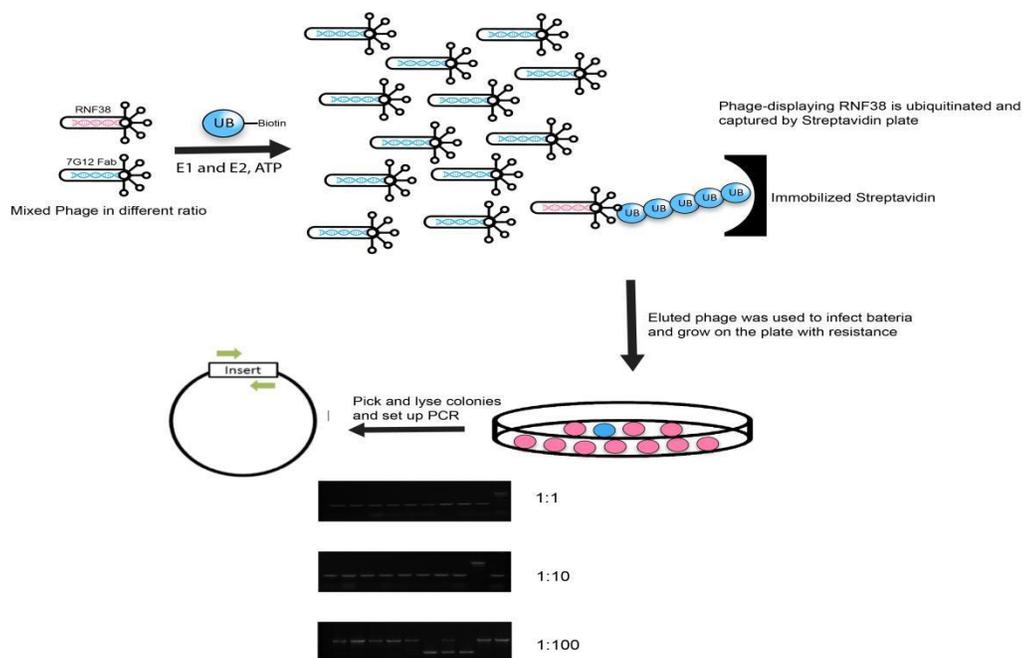
We then established an efficient phage selection protocol based on the standard biopanning technique that we used earlier for our UB and E2 phage selections<sup>[31]</sup>. Phage displaying RNF38 Ring domain were conjugated with biotin-wtUB through the transfer reaction catalyzed by Uba1(E1) and UbcH5b(E2) (**Figure 2.4A**). The reaction mixture was bound to a commercially available polystyrene plates coated with streptavidin to retain biotin-conjugated phage partials. Amount of phage bound to the plate was revealed by ELISA (**Figure 2.4B**) with a mouse anti-phage antibody and an anti-mouse IgG antibody conjugated to HRP. We found substantial binding of Ring domain displayed phage to the plate comparing to low binding levels of phage when Uba1, UbcH5b or biotin-wtUB was excluded from the reaction. Ultimately, the model selection was carried out to verify the enrichment of Ring domain phage.



**Figure 2.4 Biopanning assay for RNF38 Ring displayed on M13 bacteriophage.** (A) Phage particles derived from the pComb phagemid displaying the RNF38 Ring were ubiquitinated *in vitro*; (B) Ubiquitinated phages were captured and revealed by ELISA.

A library is typically composed of a very large number of inactive mutants (called background or noise) that compete for binding of the active ones via non-specific binding. To more closely resemble the conditions of a library selection, we performed another assay through the mixture of library phage and control phage (7G12 Fab) at different ratios to test the enrichment. According to this protocol, we mixed RNF38 Ring phage with a phage displaying the 7G12 Fab at varying ratios (1:1, 1:10 and 1:100). Phage mixtures were labeled with biotin-wt UB in the presence of Uba1 and Ubch5b and selected by binding to the streptavidin plate via the strong biotin-streptavidin affinity binding. After the selection, the composition of the phage pool was analyzed by PCR of *E.coli* colonies infected by phage. I found that a single round of selection can enrich catalytic active Ring phage over the 7G12 phage by at least 20~30 fold. In a 1:1 ratio, all output clones were RNF38; in a 1:10 ratio, 9 out of 10 clones are RNF38 while 1 clone is 7G12;

and even when the RNF38 phage was outnumbered by 1:100, we were still able to distinguish 7 out of 10 as RNF38 clones. Thus, I established an efficient selection platform to proceed the next step, which is to engineer E2-E3 interface using the developed selection protocol.



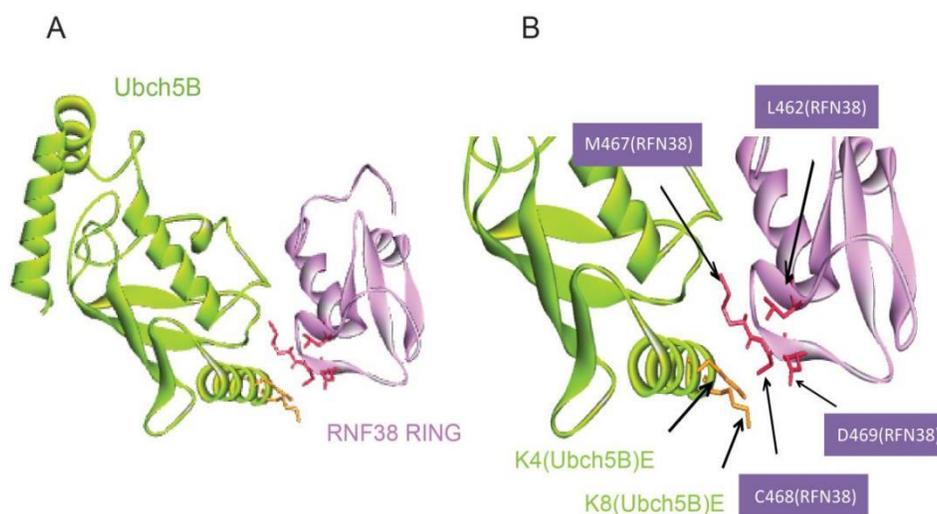
**Figure 2.5 Model selection of RNF38 phage.** A conventional ubiquitination and biopanning experiment were performed on the phage mixture (E1, E2 biotin-UB and phage mixed in tube), and bound phages were eluted with TEV cleavage (**Figure 2.4A**). Colony PCR of infected bacterial colonies verified the phages' identities.

## 2.5.3 Library phage construction of RNF38

### 2.5.3.1 General considerations for library mutagenesis

To conduct the real phage selection, a DNA library containing randomized mutants on the specific region (Ring domain) needs to be constructed. A reported crystal structure of RING domain of the E3 ligase RNF38 in complex with the E2 conjugating-enzyme UbcH5B was available (**Figure 2.6**)<sup>[106]</sup>. Since the mutations in xE2 are charge reversals in the N-terminal helix

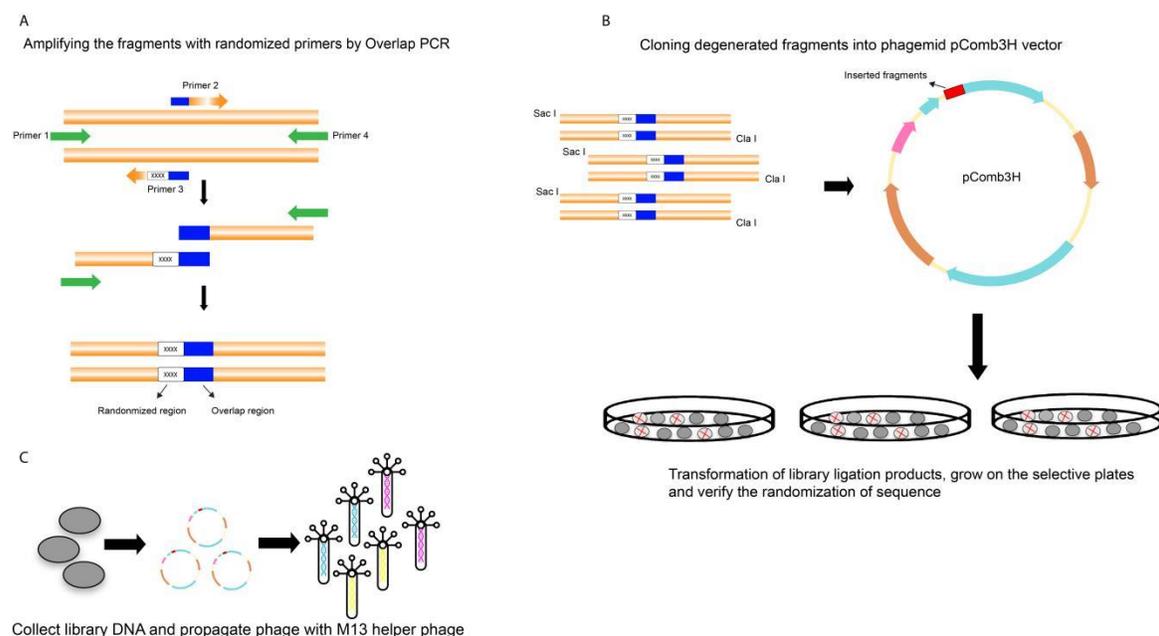
of E2, we were particularly interested in residues which interact with that region. The crystal structure of the Ubch5b-RNF38 Ring complex suggests that the L462, M467, C468 and D469 in RNF38 Ring domain are relatively close to the K4 and K8 of xUbch5B. Among them, the side chain carboxyl groups of D469 of RNF38 could be involved in electrostatic interactions with the side chains of K4 and K8 and mutagenesis of the N-terminal helix of E2 would lead to charge reversal of these residues and completely abrogated activity with wild-type RINGs and U-boxes, as was shown in the previous publication<sup>[29]</sup>. Also, C468 may potentially show interaction via a hydrogen bond or a salt bridge with K4 and K8. Based on the analysis of crystal structure between RNF38 and Ubch5B, we eventually focused on a series of residues spanning a loop structure that may interact with the N-terminal helix of E2. L462, M467, C468 and D469 in the loop of the RNF38 Ring domain were selected for randomization to construct the library through mutagenesis.



**Figure 2.6 Structure analysis of the Ring domains of RNF38 interacting with the Ubch5b.** (A) Crystal structure of Ubch5b in complex with the Ring domain of RNF38 (PDB ID 4V3L). N-terminal helix of Ubch5b interacts with loop1 of the RNF38 Ring domain. (B) Interaction between K4 and K8 of Ubch5b with the loop1 residues in the Ring domain of RNF38. K4 and K8 in Ubch5b were mutated to Glu to generate xUbch5b and they mainly engage loop1 residues L462, M467, C468, and D469 of the Ring. The K4 and K8 of Ubch5b may interact with C468 and D469 of RNF38 Ring via a hydrogen bond or a salt bridge. M467 and L462 were chosen due to their proximity to K4.

### 2.5.3.2 Library Construction for RNF38 Ring

We constructed our RNF38 phage library using generic restriction enzyme digestion and ligation into the pComb3H phagemid. The vector's original insert was 7G12 Fab which has been shown to be inactive *in vitro* ubiquitination assays and should be excluded from our selection if they are displayed in the starting library due to deficient activity with E1, E2 and UB and incomplete digestion. First, both Primer 1/Primer 3 pair and Primer 2/Primer 4 pair used the Ring domain of RNF38 approximately 280 bp as template, and there were two fragments amplified by PCR. The 4 residue mutations on the target loop of RNF38 Ring analyzed in **Figure 2.6** were introduced via a standard overlap extension PCR using Primer 1/Primer 3 pair with targeting positions randomly assigned as NNK codons<sup>[107]</sup>, which is to avoid codon bias as much as possible, as using standard NNN codons results in overrepresentation of certain amino acids. Then, further PCR using Primer 1/Primer 4 pair was proceeded for the whole RNF38 Ring fragment with randomized mutations. The PCR product was digested with Sac I and Cla I to release the sticky ends and ligated with pComb vector. Finally, the ligated products were transformed via electroporation into bacteria XL1-blue to screen out the intact colonies containing RNF38 Ring with selective resistance. Library DNA for RNF38 Ring was collected by maxi prep using Qiagen kit and dissolved in water with high purity (**Figure 2.7**), which could be used for phage propagation. The template DNA for the PCR process was an RNF38 Ring mutant in which all 4 desired residues were altered to alanine—a dead mutant that would not be active with E1, E2 and UB in order to reduce the impacts of beginning bias imposed by PCR. Additionally, we fused a FLAG tag to the N-terminal Ring sequence—the tag can be utilized to detect and remove frame shift or truncated mutants later in the selection process.

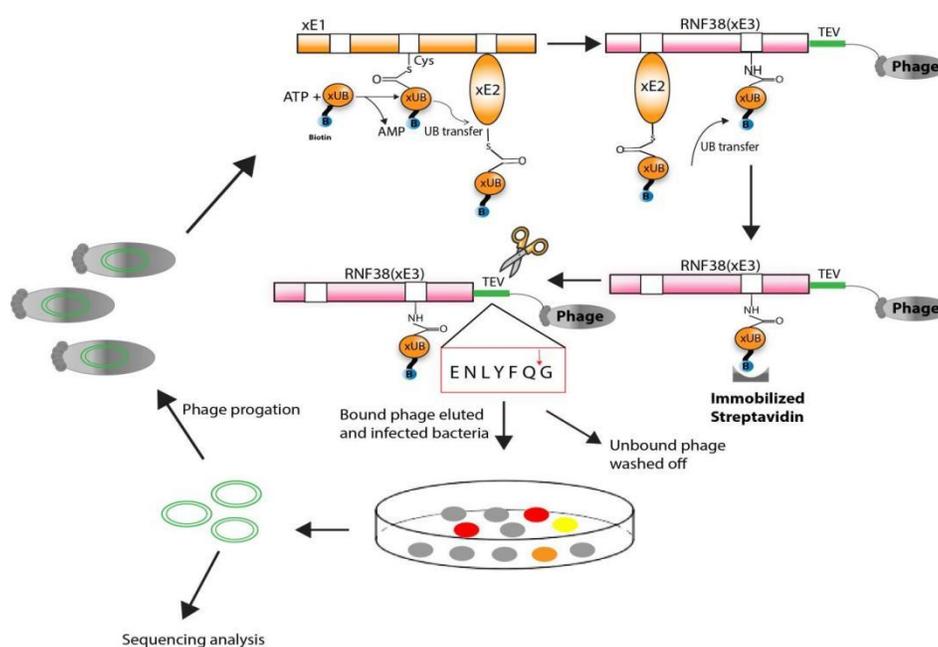


**Figure 2.7 Schematic representation of RNF38 Ring library construction.** (A) Amplifying the DNA with randomized primers by two-step PCR; (B) The library PCR products were ligated and transformed for library size titration and the clones were picked to verify the randomization of targeting region of RNF38 via sequencing; (C) Library DNA was collected and transformed into bacteria for phage propagation with M13 helper phage infection

#### 2.5.4 Phage selection of Library

We carried out library selection (**Figure 2.8**) via *in vitro* ubiquitination and biopanning assays as how we performed in model selection. For the first round of biopanning assay, we reacted library phage particles with commercially available N-terminal biotinylated wild-type UB, Back-Uba1 (E1 with C-terminal mutations interacting with xUbcH5b), and xUbcH5b (E2 with K4E and K8E mutations engineered previously). In the meantime, we would also set up three control groups that exclude the essential component E1, E2 and biotin-wtUB respectively in order to distinguish genuine activity from background binding. With an hour of *in vitro* ubiquitination reaction proceeded, we transferred each reaction mixture containing the phage particles to the commercially

available pre-blocked polystyrene plates coated with streptavidin and incubated the reaction mixture with the plates for 1 hr via biotin-streptavidin affinity binding. After 1 hr incubation, each reaction plate was thoroughly washed by TBST (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% v/v), the phages were eluted from the wells by the addition of a TEV (Tobacco Etch Virussolution) protease, a highly sequence-specific cysteine protease, which serves to recognize the cleavage TEV site and cleaves the phage bound to the plate. Eluted phages were subsequently titrated and amplified for the next round of selection via infection of a culture of XL1-blue (a popular male strain of *E.coli*) cells.



**Figure 2.8 A scheme for library selection process.** The naive library phage was produced via the transformation of library DNA of RNF38 Ring into XL1-blue cell; then, RNF38 library phages were charged with bUba1(back-E1), xUbcH5b(xE2) and biotin-wt UB (a commercially biotinylated wild-type Ub) and subjected to the plate coated with streptavidin. Through the ubiquitination in tube, the active phages were eluted by addition of TEV protease, and the eluted phage were used to infect the bacteria for sequencing (verifying the clones) and propagation for the next round of phage selection.

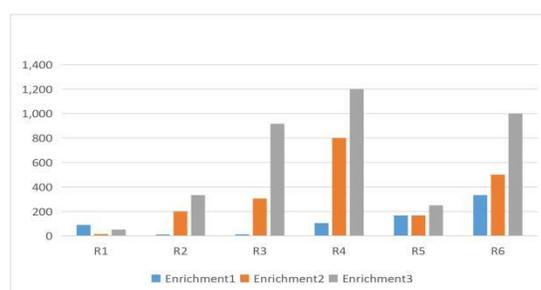
As indicated in **Table 1A-B**, we kept a close eye on the amount of output phage and the library's enrichment. The number of eluted phage particles from the ubiquitination process was divided by the number of eluted phage particles from each of the controls to determine the enrichment. Multiple rounds of selection were carried out as we changed the stringency of the selection conditions such as decreasing the input phage number and lowering biontin-wt UB concentration. With the stably increased enrichment after six rounds, we picked 20 clones from 6<sup>th</sup> library and sent them for sequencing. Four mutants were finally determined to test the activity by either *in vitro* ubiquitination assay or ELISA. The results were discussed in next section.

A

	R1	R2	R3	R4	R5	R6
Output titer(E1,E2,Ub)	$3.6 \times 10^6$	$1.0 \times 10^6$	$2.2 \times 10^6$	$4.8 \times 10^6$	$1.0 \times 10^6$	$1.0 \times 10^6$
control(-E1)	$4.0 \times 10^5$	$9.0 \times 10^4$	$3.6 \times 10^5$	$4.6 \times 10^3$	$6.0 \times 10^3$	$3.0 \times 10^3$
control(-E2)	$2.5 \times 10^5$	$5.0 \times 10^3$	$7.2 \times 10^3$	$6.0 \times 10^3$	$6.0 \times 10^3$	$2.0 \times 10^3$
control(-UB)	$7.0 \times 10^5$	$3.0 \times 10^3$	$2.4 \times 10^3$	$4.0 \times 10^3$	$4.0 \times 10^3$	$1.0 \times 10^3$

B

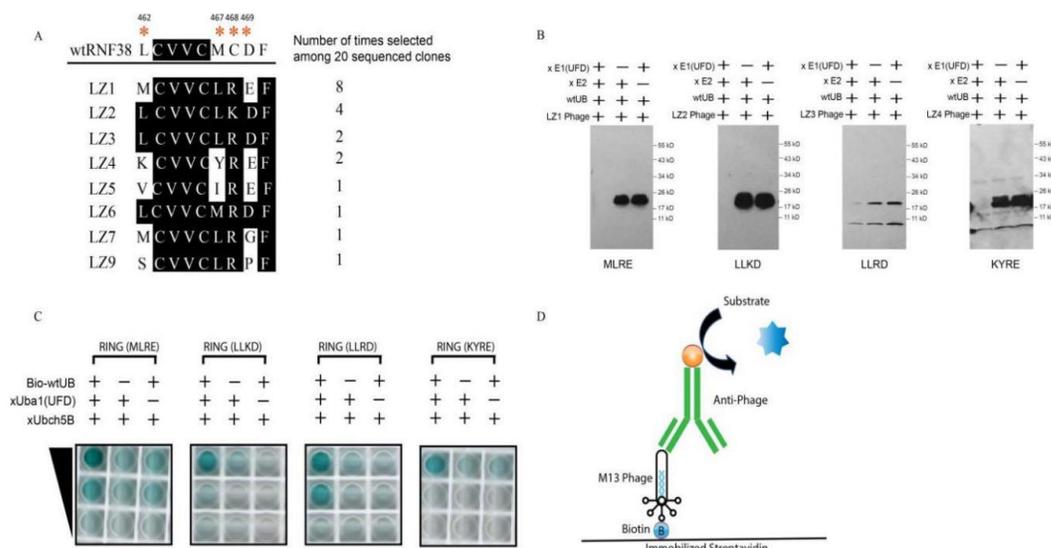
	R1	R2	R3	R4	R5	R6
Enrichment1	90	11	6	104	166	333
Enrichment2	14	200	306	800	166	500
Enrichment3	51	333	917	1200	250	1000



**Table 1 The selection output phages and the enrichment for six rounds.**

### 2.5.5 Confirmation of OUT activity of selected mutants

We analyzed the sequencing results in both round 5 and round 6 and did alignment for 20 clones we sequenced (**Figure 2.9A**). We picked 4 clones with highest frequency named as LZ1-LZ4 and prepared the phages for the selected mutants respectively. The propagated phages were all used to react with bE1, xE2 and biotin-wtUB to mimic the selection condition, as a consequence, all of them could be loaded with biotin-wt UB (**Figure 2.9B**), which demonstrated the activity in a standard ubiquitination assay. Moreover, a standard ELISA was performed as described in previous chapter, which involved bE1, xE2 and biotin-wtUB reacting with phage particles for each mutant and the reaction mixture was incubated with the streptavidin plate, then the active mutant phage particles conjugated with biotin-wtUb were captured and further developed by anti-M13 phage antibody (**Figure 2.9C**), showing that the mutant phages were all bound to the streptavidin plate. The scheme for ELISA was represented in **Figure 2.9D**



**Figure 2.9 Activity testing of RNF38 Ring mutants from selection.** (A) Sequence alignment of RNF38 Ring mutants; (B) Four RNF38 Ring mutants were prepared displaying on the phage respectively and charged with bUba1(back-E1), xUbcH5b(xE2) and biotin-wtUB (a commercially biotinylated wt ubiquitin) simulating the selection condition. The UB moiety linked to phage displaying mutants was immunoblotted by an anti-Flag antibody since a Flag epitope tag was fused to N-terminal of RNF38 Ring; (C-D) ELISA assay was performed to qualify the amount of active phage captured by the plate coated with streptavidin; a scheme for ELISA was represented.

## 2.6 Discussion

In this study, the goal of engineering the interface of xUbcH5b and xRNF38 was achieved. First, a proper display system, phage display, was chosen to screen out the E3 mutants that bind to xUbcH5b(E2). Since the Ring domain of RNF38 is commonly recognized as the crucial part interacting with the N-terminal helix of E2, the Ring domain of the ubiquitin ligase RNF38 was initially cloned onto a phagemid pComb3H vector, and it was successfully expressed on the phage particle. Furthermore, the displayed protein was tested by *in vitro* autoubiquitination assays, showing the activity on the phage particle due to the facts: 1) the Ring domain is a small domain with independent activity in autoubiquitination assays; 2) the active RNF38 Ring is around 77 amino acids in length, which is within the range of proteins that may be displayed on phage as a fusion protein to pIII.

Ring-type E3 ligases such as MDM2<sup>[108]</sup>, RNF4<sup>[109,110]</sup> and TRAF6<sup>[111,112]</sup> function either as heterodimers or homodimers in their active forms—the dimeric ligase is thought to engage asymmetrically with E2 that one monomer E3 interacts with the E2 enzyme, while the other monomer acts as a stimulant for the release of UB. RNF38 is currently one of a few E3s known to work efficiently as a monomer similar to Pirh2<sup>[113]</sup>, Cbl-b and c-Cbl<sup>[114,115]</sup>. Thus, RNF38 was thought to meet the standard to be displayed on the phage. To our delight, the RNF38 Ring had been shown to be efficiently expressed by the phage pIII and the activity of displayed Ring protein was confirmed via a standard *in vitro* ubiquitination assay (**Figure 2.3A**) and ELISA assay (**Figure 2.4B**).

Based on the directed evolution strategy, we modified the method and performed a library construction with 4 randomized residues in the loop1 of RNF38 Ring. To test the binding activity,

*in vitro* ubiquitination assay was carried out through E1-E2-E3 cascade and found the biotin- wtUB was loaded onto the library E3 variants and captured by the streptavidin plates to enrich the active mutants. Consequently, the obvious consensus among the amino acid sequence of the selected mutants was observed, thought to be potential xE3, which refers to xRNF38. Moreover, the four mutants were selected due to their high frequency among the selected variants and the activity were examined by *in vitro* ubiquitination assay and ELISA similar to the wt RNF38 Ring (**Figure 2.9B and 2.9C**). According to the results, of all the residues in the RNF38 Ring which interact with the E2 N-terminal helix<sup>[106]</sup> (**Figure 2.6**), we found C468 seemed to play the most prominent role by far, interacting with K4 and K8 of E2 (**Figure 2.6**). Since the mutation of these K4E and K8E in xUbcH5b(xE2), it removes the interaction with wild-type RNF38 Ring due to the unfavorable repulsive interactions. As was predicted, none of the mutants that were maintained following phage selection contained any negatively charged residues at position 468(either C468R or C468K) while some other positions had non-polar leucine (Leu), methionine (Met) and positively charged lysine (Lys). In conclusion, we have good reasons to hypothesize that these residues play crucial roles to restore the binding activity between Ring and the N-terminus of xE2 via favorable electrostatic interactions or hydrogen bonding. Interestingly, since we found the most converged residue was located at position C468, tending to be Arginine(R) residue, a positively charged amino acid, which extremely matched our expectation that induced the electrostatic interaction with K4E and K8E of xUbcH5b. We then generated a DNA construct of the single residue mutant RNF38 RING(C468R) as mutant LZ5 and confirmed the activity through *in vitro* ubiquitination assay (data not shown), showing a comparable activity with other mutants (LZ1-LZ4). The deep mechanism of why this residue can restore the activity with xUbcH5b dramatically

needs to be further investigated, such as crystallization by capturing the xUbcH5b-xRNF38 Ring C468R intermediate during ubiquitination.

To summarize, the sequences of the selected mutants confirmed our initial hypothesis that C468 was mutated to positive residues to compensate for the mutations of lysine residues in the N-terminal helix of xUbcH5b, while most other residues in RNF38 Ring retained or mutated to similar residues. The RNF38 OUT cascade with was constructed successfully the mutants of RNF38 the cascade can be used as a powerful tool to validate the potential substrate proteins of RNF38. Chapter 3 will explore these techniques and their outcomes.

## **2.7 Method**

### **2.7.1 pComb-RNF38 RING (FLAG) construction and protein expression of wtRNF38 Ring**

The original pGEX-wtRNF38 RING construct which was used as the template was a gift from Professor Danny Huang's group. The RING domain of RNF38 was PCR-amplified by FZ01 and FZ02 and inserted via sticky end ligation into pComb vector, fused to a N-terminal FLAG tag, in between Sac II and Spe I restriction sites.

For protein expression of pGEX-wt RNF38 Ring, GST affinity purification was done following a standard protocol. To express the GST-RNF38 Ring protein, the pGEX-RNF38 Ring plasmid was transformed into BL21(DE3) cells (Agilent). The cells were grown in 2x YT broth supplemented with kanamycin (70 mg/ml) at 37 °C. When the culture reached an optical density of 0.6~0.8, IPTG (isopropyl- $\beta$ -d-thiogalactopyranoside) was added to a concentration of 1 mM. The culture was shaken at 15 °C overnight. The cells were harvested by centrifugation at 5000 rpm for 30 minutes, resuspended in 5-10 ml of lysis buffer (20 mM Tris pH 7.5, 0.1 % Triton X-100, 1 mg/ml, lysozyme 1  $\mu$ g/ml DNase), and treated with 2 mg/ml lysozyme (Alfa Aesar) and

incubated on ice for 30 minutes, before lysis being completed by either French press or sonication. Then, remove the insoluble material by centrifugation at 10,000 rpm at 4°C for 25 min. Mix the clarified lysate in a falcon tube at 4°C for 14~18hrs with Pierce<sup>TM</sup> Glutathione Agarose (Thermo, Cat.16102BID) preequilibrated with 20 mM Tris pH 7.5, containing 0.1 % Triton X-100. Protein was washed and eluted from the agarose by washing buffer and elution buffer, and further was dialyzed overnight at 4 °C in a buffer containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl and 1 mM DTT. Lastly, the protein was concentrated, and the concentration was measured using Bradford assay according to the vendor's protocol (Bio-Rad). The success of the expression was confirmed via western blotting and Coomassie staining of SDS gel; the yield of the purified protein was assessed using a standard Bradford assay and taken as total protein concentration.

### **2.7.2 Autoubiquitination assay of recombinantly expressed RNF38 RING**

In a 50 µL reaction containing 50 mM Tris pH 7.5, 10 mM MgCl<sub>2</sub>, 2.5 mM ATP, and 1 mM DTT, 5 µM GST-tagged RNF38 protein was incubated with 1 µM Uba1, 5 µM Ubch5b, and 10 µM HA-tagged ubiquitin. The reaction was allowed to proceed at 37°C for 1 hr. SDS-PAGE of reaction and control was run and analyzed via western blotting using mouse antibody against GST (to monitor the formation of ubiquitin chains by RNF38).

### **2.7.3 RNF38 phage propagation**

The pComb-wtRNF38 Ring phagemid was then transformed into XL1-blue (Agilent) cells; transformants were screened on agar plates containing 100 µg/ml ampicillin and 2% glucose; a single colony was then inoculated into LB medium with 2% glucose, 10 µg/ml tetracycline and 100 µg/ml ampicillin and grown overnight at 37°C. The overnight culture was then inoculated into 20ml of 2xYT medium supplemented with 2% glucose, 10 µg/ml tetracycline and 50 µg/ml

ampicillin at a starting optical density of 0.1, and the culture was shaken at 37°C until the OD reached ~0.6, and the approximate number of cells was calculated; VCSM13 helper phage (Aligent) was then added to the culture at a multiplicity of infection of ~10; the culture was then incubated at 37°C in an incubator without shaking. The cells were then harvested by centrifugation at 3700 rpm for 10 minutes, and resuspended in 200 ml of fresh 2xYT media with 100 µg/ml ampicillin and 50 mg/ml kanamycin; the culture was then grown overnight at 30°C. The next morning, the culture was cleared by centrifugation at 5000 rpm for 20 minutes; the cell pellet was discarded. 25 ml of filter-sterilized 5x PEG/NaCl solution (200 g/L PEG-8000 (Sigma), 2.5 M NaCl) was thoroughly mixed with the 125 ml supernatant from the previous step, and the mixture was incubated on ice for 1 hour and afterwards centrifuged at 5000 rpm for 1 hr to precipitate phage particles, which appeared as a white streaky pellet. The phage pellet was thoroughly resuspended in 1 ml of sterile TBS and cleared of any residual cell debris by centrifugation at 10000 rpm for 10 minutes. The resulting supernatant was then titrated by infection and tested for the display of the RNF38 Ring by western blotting probed with an antibody against FLAG.

#### **2.7.4 *In vitro* autoubiquitination assay of RNF38 Ring-displayed phage**

Over  $1 \times 10^{11}$  phage particles derived from the pComb-Flag-RNF38 Ring construct were incubated with 1 µM Uba1, 5 µM UbcH5b and 20 µM N-terminal HA-ubiquitin, in 50 µL reaction containing 50 mM Tris pH 7.5, 10 mM MgCl<sub>2</sub>, 2.5 mM ATP, and 1 mM DTT and reacting for 1 hour at 37°C shaker. The reaction quenched by boiling in Laemmli buffer containing 100 mM DTT and analyzed via SDS-PAGE and immunoblotting using mouse antibody against FLAG.

### 2.7.5 Biopanning of RNF38 Ring phage and quantification by ELISA

For the first round, around  $10^{10}\sim 10^{11}$  phage particles displaying RNF38 Ring were charged with 1  $\mu\text{M}$  uba1, 5  $\mu\text{M}$  UbcH5b and 0.5  $\mu\text{M}$  N-terminal labelled biotin-ubiquitin (Boston Biochem) in 20  $\mu\text{L}$  total reaction volume containing 50 mM Tris pH 7.5, 10 mM  $\text{MgCl}_2$ , and 3 mM ATP, and incubated for 1 hour at room temperature; negative controls separately omitting E1, E2 and ubiquitin were also included. The reactions were diluted 10-fold into 3% BSA-TBST and incubated on streptavidin-coated polystyrene plates (Thermo) for 1 hour at room temperature. The plates were thoroughly washed 30 times with TBST, and 30 times with TBS. 100  $\mu\text{L}$  of an elution buffer containing 100 mM DTT in TBS was applied to each well and allowed to incubate for 15 minutes. Eluted phage particles from each well were added to 1 ml of mid-log culture of XL1-blue cells in LB medium and allowed to infect for 1 hour at  $37^\circ\text{C}$ . Quantification of infected cells was done by serial dilution and plating on selective agar media containing 100  $\mu\text{g}/\text{ml}$  ampicillin. After incubation of the plates overnight at  $37^\circ\text{C}$ , the number of colonies were used to extrapolate the total number of phage particles eluted from each well, assuming one phage particle per colony. In a similar manner, the amount of input phage per well was also calculated.

Alternatively, quantification was also performed via ELISA assay, following the reaction conditions described above. The reaction was serially diluted four times, with ten-fold dilutions in 3% BSA-TBST and loaded, incubated and washed in the same manner as the above protocol. The presence of phage particles was detected by 1 hour incubation with anti-M13 antibody-HRP (Fisher) in 3% BSA-TBS, and after washing 10 times with TBST and 10 times with TBS, the signal in each well was quantified using TMB substrate kit (Fisher).

## **2.8 RNF38 library selection**

### **2.8.1 Model Selection**

RNF38-displaying phage particles were mixed with phage particles displaying 7G12 in varying ratios of 1:1, 1:10 and 1:100 (titration of phage preps were done via infection and colony-counting); the different mixes, each containing  $\sim 10^{10}$  phage particles in total were subjected to the same reaction conditions, binding, and elution, and was used to infect XL1-blue cells in the same manner as described in the previous section. After infection, the culture was streaked onto selective ampicillin media, after overnight incubation, colony PCR was performed using the phagemid-specific Jun13 and Jun14 primer pair. The two different types of phage particles were distinguished and identified by the size of expected PCR product.

### **2.8.2 Library construction of pComb-RNF38 RING**

pGEX-RNF38 RING wild-type was used as the template to amplify the Ring sequence with FZ01 and FZ02 primers by PCR. The insert was then ligated into pComb3H to replace the original insert of 7G12, via Sac II and Spe I sites. The original vector has a FLAG tag at the N-terminal. Residues to be randomized in the RNF38 Ring (L462, M467, C468 and D469) were first mutated to Ala to generate phagemid pComb-RNF38 Ring 5Ala. For mutagenesis, the Ring gene was amplified by two sets of primers FZ01-FZ03 and FZ04-FZ02. The PCR fragments were assembled by overlap extension for cloning into the pComb phagemid. For randomization of the five residues, pComb-RNF38 Ring 5Ala mutant was used as the template for PCR reactions with primer pairs FZ01-FZ03 and FZ05-FZ02. The overlap extension of the amplified fragments was cloned into pComb phagemid to generate the library. The library DNA was transformed into XL1-blue cells (Agilent) by electroporation. The cells were plated on LB-agar plates containing

ampicillin (100 µg/ml) and 2% glucose and incubated overnight at 37°C. The phagemid DNA for the library was prepared with a DNA Maxiprep kit (Qiagen).

### **2.8.3 Library phage preparation for RNF38 Ring**

A quantity of library DNA was extracted and purified from bacteria (depending on the round), and then transferred into 50 µl of electrocompetent XL1-blue cells (Agilent). The transformed cells were put to 1.0 ml SOC medium and allowed to recover for 1 hour at 37 °C shaker. Following that, 20 ml of 2xYT broth containing 50 µg/ml ampicillin and 10 ug/ml tetracycline was added to the recovered culture and left to grow for another 2~4 hours at 37°C until the OD<sub>600</sub> reached 0.6~0.8. Then, the VCSM13 helper phage (Agilent) was added to superinfect the cells at >10-fold multiplicity of infection (the total number of cells was approximated using optical density). After 1hr infection, the culture was then transferred to 200 ml of 2xYT medium, and the concentration of ampicillin was increased to 100 µg/ml and tetracycline concentration was still the same to the original. Finally, another antibiotic, kanamycin, was added at a final concentration of 70 µg/ml, along with IPTG addition at a final concentration of 0.5 mM, the temperature was adjusted to 30°C, and the 200 ml cell culture was allowed to grow overnight under vigorous shaking at a speed of 300 rpm. The next morning, phage purification precipitated by PEG was proceeded by a standard protocol described earlier<sup>[36]</sup>. The purified library phage particles were suspended in sterile TBS or PBS buffer. The display of the library phage was then probed by western blotting with anti-FLAG antibody since N-terminal FLAG affinity tag was fused to RNF38 Ring.

#### 2.8.4 Biopanning and selection of library phage

The exact reaction conditions in which phage particles displaying mutant RNF38 Ring domains were subjected to react with commercial biotin-wtUB (Boston Biochem), back-Uba1 (bE1, with mutations E1004K, D1014K, E1016K at UFD domain) and xUbcH5b (xE2 with K4E and K8E mutations) vary from round to round. All library ubiquitination reactions were performed for 1 hour at 37°C incubator. The different conditions in each round are as given in the table below.

Generally, completed reactions were diluted 10-fold into 3% BSA-TBST solution before being loaded onto streptavidin plates (Thermo). Controls was set up without E1, E2 and UB, respectively similar to what we conduct in the biopanning assay of the wild-type RNF38 Ring phage. Each reaction mixture was incubated on the streptavidin plates for 1~2 hours at room temperature. After incubation, the plates were thoroughly washed ~20-30 times with 3% BSA-TBST. Elution was performed by TEV cleavage reaction per well. To proceed the cleavage reaction, we added TEV protease (5 units), DTT (1 mM) and ProTEV Buffer to each well, the plates were incubated at 30°C or lower temperature for 30-45 min since long time. Then, collect reaction mixture from each well. The eluted phage particles from one well were added to 1 ml of a high-density culture (OD around 0.6~0.8) of XL1-blue cells in 2xYT medium supplemented with 10 µg/ml tetracycline. The culture was incubated for 1.5 hours at 37°C with slow (~100 rpm) shaking. After incubation, they were plated onto LB-agar plates containing 100 µg/ml ampicillin and 2% glucose and allowed to grow overnight at 37°C shaker. The titers of the selected phage particles, taken to be equal to the number of infected colonies, were compared to the controls, and enrichment was calculated. The cells were collected from the plates and exact the selected library DNA using maxiprep kits (QIAGEN). The selected library DNA was then used in propagating the next round of phage.

### **2.8.5 Testing the selected mutants via biopanning assays to validate the OUT cascade of RNF38 Ring**

By analyzing the sequencing results of 20 clones we sent, top 4 phagemids were selected and labeled as LZ1, LZ2, LZ3 and LZ4. To further identify the activity of each clone, they were all transformed into XL1-blue cells and propagate phage particles following a standard protocol previously described. The phage particles for each clone with an amount of  $2 \times 10^{10}$  cfu were charged with 0.5  $\mu$ M bUba1), 5  $\mu$ M xUbch5b, and 0.5  $\mu$ M Biotin-wtUb (Boston Biochem) in 50  $\mu$ L of 50 mM Tris pH 7.5, 10 mM  $MgCl_2$ , 5 mM ATP and 0.1% BSA, and incubated for 1 hr at 37°C. In the meantime, we set up controls without E1, E2 and Ub respectively, following the same condition. After 1 hr reaction, the reaction of each well was loaded onto streptavidin-coated plates. Finally, the titration and the quantitation by ELISA were described earlier in 2.7.5.

### **2.8.6 Construction pGEX-RNF38 Ring mutants and protein expression**

The selected 4 phagemids were purified and used as template, so a PCR amplification was performed following the vendor's manual (NEB), then we digested both inserts and vector with restriction enzymes BamH I and Not I. Then, the ligation of the PCR products into the pGEX vector was performed using T4 ligase at 16°C overnight. In addition, the protein expression and purification are the same as those of the protocol described in 2.7.1.

**Table 2 Primary antibody information**

Antibody	Supplier	Catalog Number	Diluton
FLAG	Sigma	F4042	1:2000~1:4000
GST	Santa Cruz	sc-138	1:500~1:1000

**Table 3 Primers used in this study**

Primer	Sequence
FZ01	5'-GACTGCCCGCGGACTAAAGCAGATATTGAAC-3'
FZ02	5'-ACTAGTGCCCTGAAAATACAGGTTTTCTTCTGAATCCCGATG-3'
FZ03	5'-AGTCTGTTCTGACTGGTGGTTGTTAGG-3'
FZ04	5'-CAGTCAGAACAGACTGCATGTGTAGTATGCGCAGCAGCATTGAGTCAAGG- 3'
FZ05	5'-CAGTCAGAACAGACTNNKTGTGTAGTATGCNNKNNKNNKTTTGAGTCAAGG- 3'

### 3 VERIFYING THE OUT CASCADE OF RNF38 MUTANTS AND OTHER RING-TYPE E3S BOTH *IN VITRO* AND *IN VIVO*

#### 3.1 Background information

##### 3.1.1 The rationale

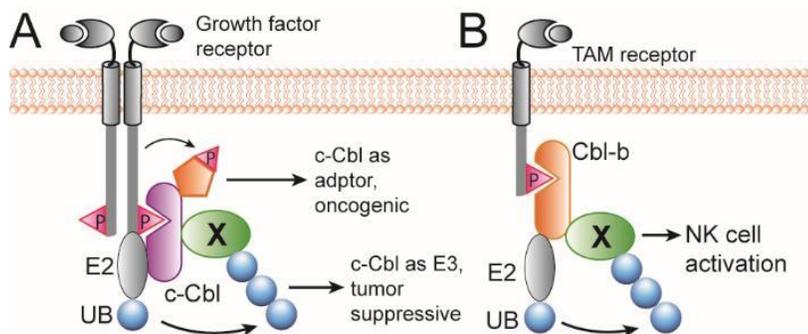
The directed evolution strategy effectively produced Ring mutants of RNF8 with reactivity to xUba1, xUbcH5b, and xUB. The restored interaction between the Ring mutants and the xUbcH5b encouraged us to move forward and explore whether the substrate can be ubiquitinated by xE1-xE2-xE3-xUB cascade both *in vitro* and *in vivo*. However, it was not immediately concluded if the mutations on the RNF38 Ring would influence substrate ubiquitination due to the change of a few residues. The conventional thinking is that substrate contact needs additional domains of the E3 RNF38. In addition, members in the Ring family E3s generally possess remarkable structural similarity as far as the Ring domain is concerned but show great variation as to the other protein-interaction modules fused to the Ring domain. This chapter is mainly to discuss the verification of activity of xRNF38 both *in vitro* and *in vivo* ubiquitination, as well as the confirmation of activity of xCbl-b and xPirh2 via the transplantation of xRNF38 Ring to other monomeric Ring E3 ligases, Cbl-b and Pirh2.

#### 3.2 The structural and biological roles of Cbl-b, Pirh2 and RNF38

##### 3.2.1 CBL-B (Casitas B lymphoma-b)

The Cbl family proteins are monomeric RING-type E3s that could ubiquitinate a bunch of cellular proteins and subsequently degrade them through the proteasome<sup>[114]</sup>. For these substrate proteins, CBL-b has been verified as a crucial gatekeeper that can limit T-cell activation and more specifically natural killer (NK) cell activation<sup>[116,117]</sup>(**Figure 3.1B**). A Cbl-b-deficient mice model

has exhibited the anti-tumor activity which has built a key link between Cbl-b and TAM-receptors (Tyro, Axl, Mer)<sup>[118]</sup>. With the activation of TAM receptors, it will induce the dissociation of Cbl-b from TAM and produce an inhibitory peptide or sequence to NK cell, thereby leading to the abrogation of anti-tumor responses<sup>[118]</sup>. Hence, developing a TAM inhibitor such as a synthesized small molecule may exert more specific inhibitory effects on cancer therapy. With high homology to Cbl-b, c-Cbl could regulate receptor tyrosine kinases (RTK) including EGFR, PDGFR and FGFR through the ubiquitination of RTKs<sup>[119,120]</sup>. Furthermore, the ubiquitination of RTKs would undergo endocytosis, endosome sorting, and eventually entering the lysosome for degradation. Besides, a handful of non-receptor tyrosine kinases such as ZAP-70 could be captured by c-Cbl that acts as an adaptor to prolong their signals in the cell<sup>[121,122]</sup>(**Figure 3.1A**). Hence, c-Cbl plays dual roles in reducing or promoting RTK signals by functioning as an E3 ligase or adaptor protein. In summary, the Cbl family proteins do have a huge network with RTKs. Identifying the substrates of Cbl proteins would help us to untangle the complicated networks of Cbl-mediated cell signaling and provide a more effective platform for cancer-drug design. The construction of the OUT cascade with Cbl-b can help to identify the substrates of Cbl-b in the cell. This would help us gain more profound insights into cell signaling mechanisms such as cell cycle regulation and cancer immunity and provide the basis of cancer drug design.

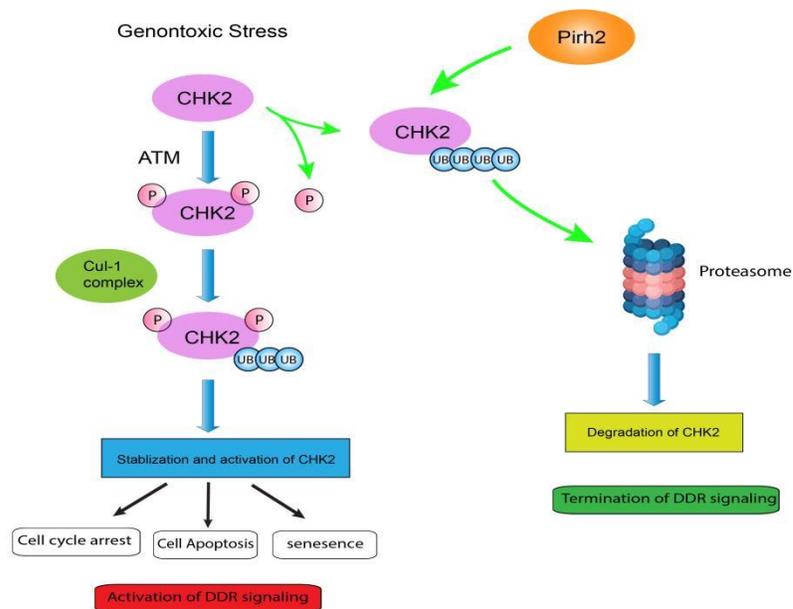


**Figure 3.1 Roles of Cbl E3s in cell regulation.** (A) c-Cbl inhibits tumor development and acts as an adaptor for RTKs. OUT will be used to distinguish c-Cbl substrates from their binding partners. (B) Cbl-b activates natural killer cells (NK cells) against cancer cells. OUT will be used to discover critical Cbl-b substrates for NK cell regulation.

### 3.2.2 PIRH2 (p53-induced protein with a RING-H2 domain)

Pirh2 was identified as a p53-induced protein with a RING-H2 domain, also known as Rchy1 (RING-finger and CHY-zinc-finger domain-containing protein 1). In 2002, it was firstly reported as an interactor with androgen receptor<sup>[123]</sup>. To date, there are at least five isoforms of Pirh2 protein, named as Pirh2A, B, C, C' (also called Pirh2b), and D<sup>[124]</sup>. Among them, Pirh2A is widely studied with full-length protein composed of 261 amino acids and it can be divided into the N-terminal CHY-Zn-finger domain, the central RING-finger domain, and the C-terminal domain (CTD)<sup>[113]</sup>. The RING domain is responsible for the ubiquitin ligase activity, while the N- and C-terminal domains function as protein-protein interactors, which are often utilized to recognize and bind substrates, hence inducing ubiquitination<sup>[125]</sup>. p53, the well-known tumor suppressor proteins which is involved in the control of cellular stress responses such as DNA damage, oncogene activation, and hypoxia<sup>[126]</sup>, was reported as a substrate protein of Pirh2 via the strong interaction with the CTD of Pirh2 targeting the p53 tetramerization domain (TET) and weak binding to the NTD of Pirh2<sup>[127,128]</sup>. The polyubiquitination of p53 by Pirh2 causes proteasomal degradation of

p53 and leads to the decrease in p53-mediated functions including apoptosis and cell cycle arrest<sup>[127]</sup>. Pirh2 was also shown to enhance the ubiquitination of AR corepressor HDAC1 (histone deacetylase 1), which leads to its degradation<sup>[124]</sup>. In addition, Pirh2 is also found to inhibit androgen-dependent secretion of PSA via ubiquitination and degradation of  $\gamma$ -COP ( $\gamma$ -subunit of coatmer complex)<sup>[129]</sup>. Furthermore, the stability of CHK2 protein was also regulated by Pirh2 by the proteasomal degradation in response to DNA damage<sup>[130]</sup>(**Figure 3.2**). Pirh2 also monoubiquitylates polH (DNA polymerase Eta) to prevent it from functioning as a polymerase that bypasses residual DNA lesions during the S phase of the cell cycle, and thereby inhibits the interaction between polH and PCNA<sup>[131,132]</sup>. Although Pirh2 ubiquitinates and degrades multiple tumor suppressors, suggesting the oncogenic activity, its cellular function remains elusive. Hakem et al., showed that the Pirh2 can polyubiquitylate c-Myc and mediates its proteolysis acting as a tumor suppressor protein in lung, ovarian, and breast cancers<sup>[133]</sup>. Recently, Daks et al., concluded that Pirh2 regulated the c-Myc expression via ubiquitinating HuR<sup>[125]</sup>, a RNA-binding protein involved in the down-regulation of splicing and protein stability of c-Myc. To summarize, Pirh2 is emerging as an E3 ligase with critical involvement in DNA damage response, cell proliferation, and cell cycle progression. Finding out the downstream targets of Pirh2 would help to elucidate the important regulatory roles of Pirh2 in cells.



**Figure 3.2 A proposed model for CHK2 ubiquitination by Pirh2 in response to DNA damage response (DDR)<sup>[130]</sup>.** **Left:** In response to genotoxic stress, CHK2 is phosphorylated by ATM, the fully activated CHK2 was polyubiquitinated by E3 ligase Cul-1 complex, thereby leading to the activation of DDR. **Right:** The E3 ligase Pirh2 catalyzes the polyubiquitylation of dephosphorylated CHK2, resulting in their proteasomal degradation, which contributes to the termination of DDR signaling.

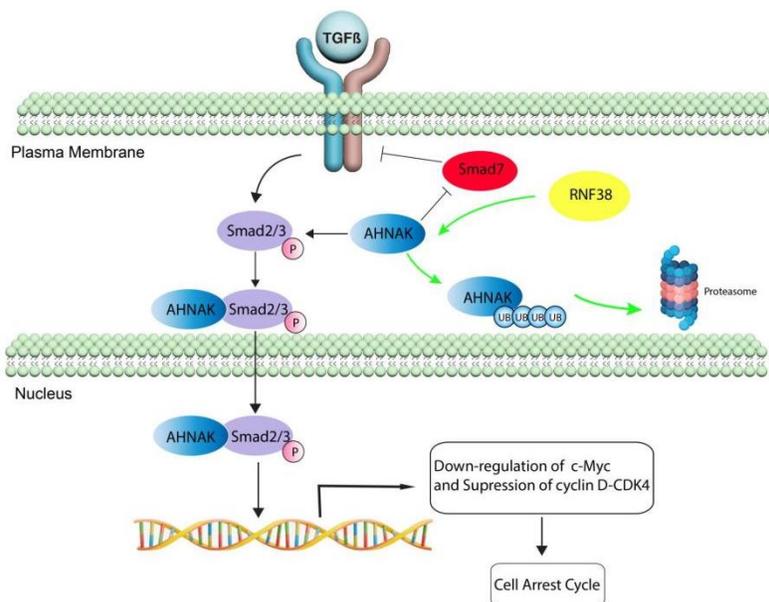
### 3.2.3 RNF38 (Ring Finger Protein 38)

There are three isoforms of RNF38 predicted. Its isoform 1 is longest in length (515 amino acids), while isoform 2 and 3 lack residues 5-54 and 1-84, respectively<sup>[134]</sup>. To date, RNF38 is found to have a coiled-coil motif and a RING-H2 domain (C3H2C2) at its carboxy-terminus that is in charge of E3 activity<sup>[135]</sup>. The roles of other domains in RNF38 are unknown. In addition to the structure of RNF38, its mRNA is extensively expressed in a range of human tissues, and its evolutionary conservation indicates that it performs a critical biological function<sup>[136]</sup>. Furthermore, RNF38 is peculiarly situated on a region of chromosome 9 (9p13) that is commonly deleted in a

variety of malignancies. The RING-type zinc finger motif is a zinc-binding domain present in RING ubiquitin ligase family members that is involved in a variety of biological activities such as protein degradation, signal transduction, cell proliferation, and oncogenesis. As of now, there are a few reports on RNF38's role in malignancies, including hepatocellular carcinoma (HCC), non-small cell lung cancer (NSCLC), colorectal cancer (CRC) and gastric cancer (GC). RNF38 was initially reported by Sheren et al<sup>[134]</sup> to ubiquitinate p53, indicating the potential oncogenic activity in the cells. In HCC patients, RNF38 is highly expressed that conferred the EMT phenotype on HCC cells via facilitating TGF- $\beta$  signaling by ubiquitinating a tumor suppressor AHNAK, thereby leading to the poor prognosis<sup>[137]</sup>(**Figure 3.3**). Besides, RNF38 was significantly expressed in gastric cancer and promoted STAT3 signaling by ubiquitinating SHP-1, therefore boosting gastric cancer cell proliferation<sup>[135]</sup>. Furthermore, it has been found that RNF38 increased the development of non-small cell lung cancer (NSCLC) through regulating EMT, implying a bad prognosis for NSCLC patients<sup>[138]</sup>. In contrast to HCC, NSCLC, and GC, RNF38 is downregulated in patients with colorectal cancer (CRC). RNF38 overexpression in CRC cells can ubiquitinate LDB1 and promote its degradation, therefore limiting the development of CRC cells, which indicates that RNF38 functions as a tumor suppressor in colorectal cancer (CRC) cells<sup>[139]</sup>. Thus, RNF38 has been shown to play critical roles in carcinogenesis across a range of cancer types by controlling the degradation of tumor promoters or suppressors.

As described in Chapter 1, the E3s recognize protein ubiquitination targets, they often play key regulatory roles, and their malfunction drives the development of many diseases including cancer, neurodegeneration, and inflammation<sup>[4,140]</sup>. Regardless of the nature of interaction with E2s, an E3 may take up UB from several E2s, and numerous E3s may transfer UB to a pool of substrates that overlaps. Due to the intricate cross-reactivities between E2, E3, and substrates,

profiling the substrates of a given E3 to map it on the cell signaling network is a substantial difficulty<sup>[30]</sup>.



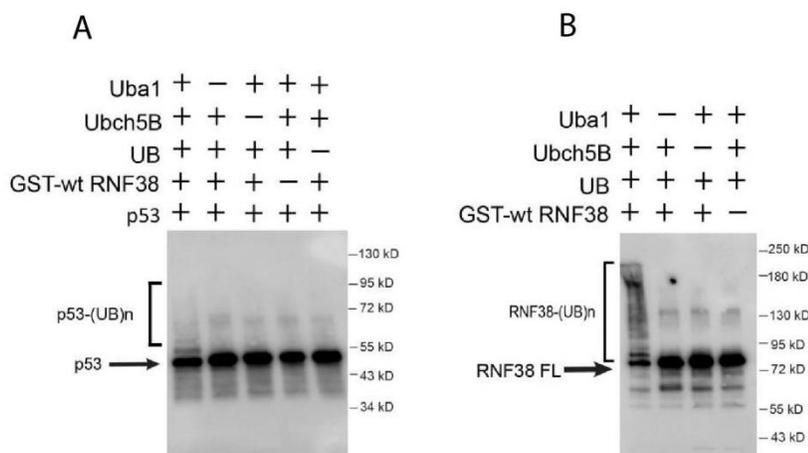
**Figure 3.3 The cellular regulation of RNF38.** The high-level expression of RNF38 ubiquitinates AHNAK for proteasomal degradation. AHNAK functions as binder to Smad 2/3 proteins to inhibit the activity of p-Smad encoding enzyme in the nucleus. The dysfunction of AHNAK leads to the cell arrest cycle.

### 3.3 Results

#### 3.3.1 Verifying the activity of RNF38 Full-length via *in vitro* ubiquitination

Previously, we have demonstrated that RNF38 Ring domain could be well expressed through the bacterial expression system (**Figure 2.3A**) and showed good activity by auto-ubiquitination. In theory, we should be able to use the same system to express RNF38 full-length protein. Unfortunately, we found we were not able to express RNF38 full-length in bacteria since RNF38 full-length protein is very fragile<sup>[141]</sup>. After going through literature review on protein

expression of full-length RNF38, we successfully expressed the RNF38 full-length with GST epitope tag on N-terminal in HEK293T cell line, which applies a mammalian expression system<sup>[142]</sup>. The activity of GST-RNF38 full-length was confirmed via *in vitro* autoubiquitination assays (**Figure 3.4B**). Furthermore, we had a known substrate of RNF38, called p53, commercially purchased from BD Biosciences included in the auto-ubiquitination reaction to initiate p53 ubiquitination *in vitro*<sup>[134]</sup>. As a result, p53 ubiquitinated by RNF38 was reconstituted (**Figure 3.4A**).

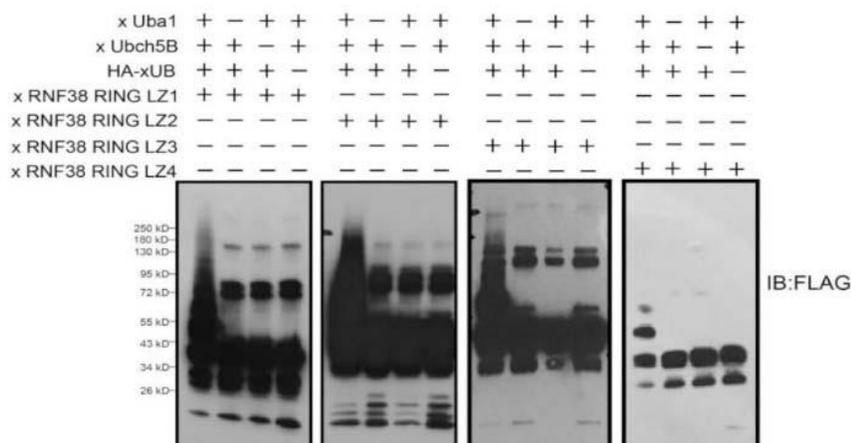


**Figure 3.4 The activity test of full-length wild-type RNF38.** (A) p53, as a known substrate of RNF38, was polyubiquitinated by RNF38 *in vitro*; (B) The successful expression of the full-length wild-type RNF38 protein was confirmed by Western blotting, and its activity was verified by *in vitro* autoubiquitination assay

### 3.3.2 Examining the activity of xRNF38 Ring mutants *in vitro*

As introduced in Chapter 2, four RNF38 Ring mutants were chosen to be cloned into pGEX vector using the phagemids as template so we can express the full length RNF38 mutants. A GST-tagged protein purification was performed using the standard protocol as described in Chapter 2.

In addition, a FLAG epitope tag was fused to the N-terminal of RNF38 Ring. The purified RNF38 Ring mutants were subjected to reaction with xUba1, xUbcH7 and xUB, respectively. After 1hr reaction, we ran SDS-PAGE, and the anti-FLAG antibody was used to immunoblot the RNF38 Ring mutants. The strong signal was observed by detecting the intensity of ubiquitination smear, indicating that all of 4 mutants were active (**Figure 3.5**). Among the 4 mutants, LZ1, LZ2, LZ3 and LZ4 mutants showed virtually the same levels of activity, as was also observed when only the wild-type Ring domain protein was charged with wtE1, wtE2 and wtUB (**Chapter 2, 2.3A**). The successful xUB transfer through the entire OUT cascade was observed via *in vitro* ubiquitination assay.

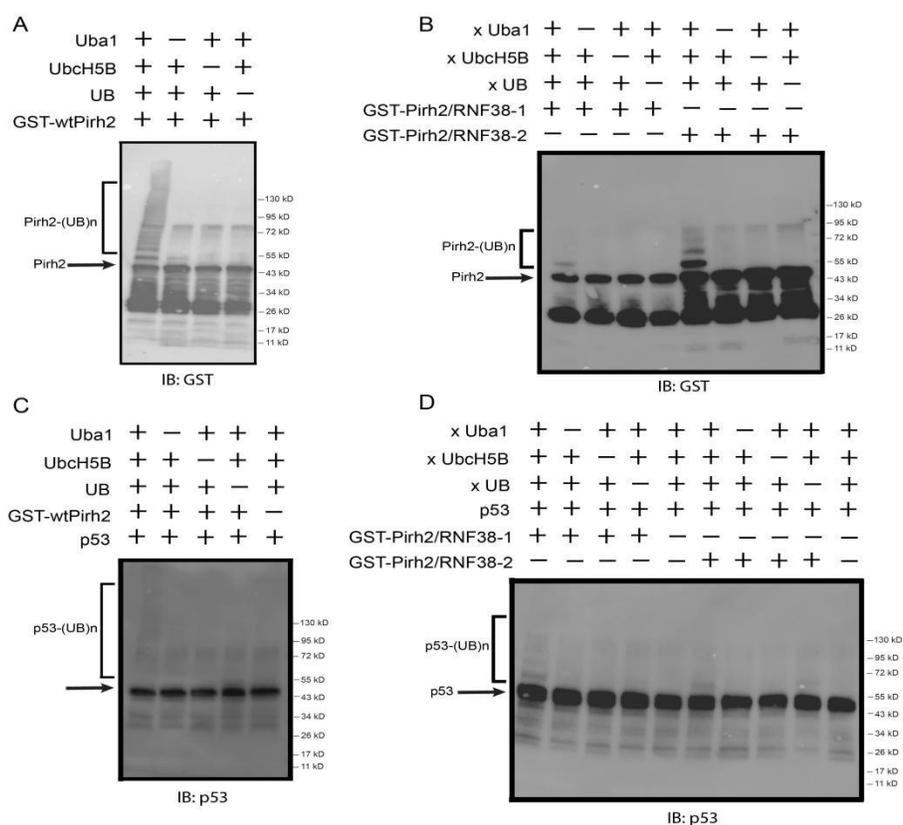


**Figure 3.5 Activity assays of RNF38 Ring mutant variants.** *In vitro* autoubiquitination assays of the RNF38 Ring mutants LZ1, LZ2, LZ3 and LZ4. Autoubiquitination was probed against the FLAG epitope on the RNF38 protein.



### 3.3.4 Examining the activity of both xCbl-b Ring and xPirh2 Ring mutants *in vitro*

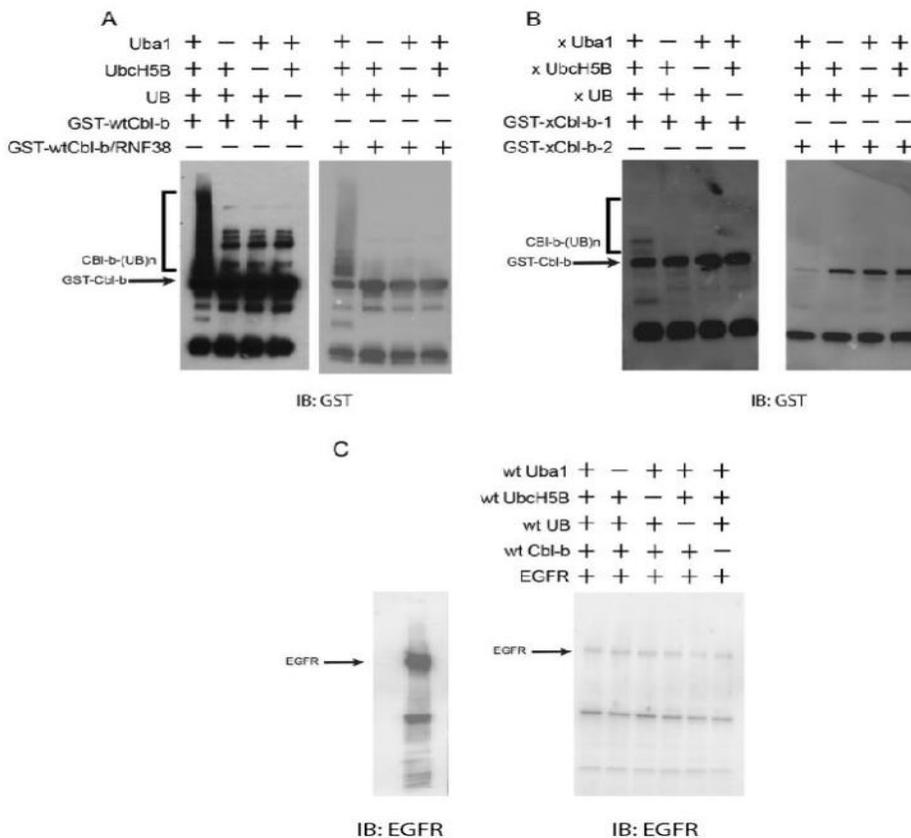
The success of this design was confirmed in xUB autoubiquitination reactions with the xUba1-xUbcH5b pair (**Figure 3.7 ad 3.8**). Furthermore, both Pirh2 mutants could transfer xUB to itself (**Figure 3.7A-B**) and p53 (**Figure 3.7C-D**), a known Pirh2 substrate<sup>[127]</sup>, at an efficiency comparable to wt UB transfer by wt Pirh2. These results demonstrated that either Pirh2-LZ1 and Pirh2-LZ5 could be used as an xPirh2 in an OUT cascade to profile its substrates.



**Figure 3.7** *In vitro* ubiquitination assays of Pirh2 variants. **(A)** Autoubiquitination of GST-tagged wild-type full-length Pirh2; **(B)** Autoubiquitination of GST-tagged Pirh2 mutants, xPirh2-LZ1 and xPirh2-LZ5; **(C)** Putative substrate (p53) ubiquitination by wtPirh2; **(D)** Ubiquitination of p53 by xPirh2 LZ1 and LZ5.

For Cbl-b mutants (TKBD-LHR-RING), we also discovered both Cbl-b-LZ1 and Cbl-b-LZ5 could be autoubiquitinated *in vitro*, which suggested that xUB was enabled to be loaded on the Cbl-b mutants via the xE1-xE2-xCbl-b cascade (**Figure 3.8A-B**). For a known substrate of Cbl-b verification, we found EGFR is a good candidate for Cbl-b<sup>[143]</sup>. However, it was very hard to express the protein EGFR from bacteria. Thus, we took an alternative approach to express EGFR, which used a mammalian expression system. As reported by Huang et al's group<sup>[144]</sup>, EGFR was highly expressed in A431 cells, a type of epidermoid carcinoma cell line. We purchased the cell line from ATCC and cultivate the cell as instructed, similar to HEK293T culture condition. Five 75cm<sup>2</sup> flasks of cells were precleared and collected, then the antibody against EGFR was added to bind to the protein EGFR in the cell. After overnight binding, protein agarose A/G beads were added into the cleared cell lysate expressing EGFR. After 4 hrs, EGFR in A431 cell line was precipitated by the anti-EGFR antibody and the agarose A/G beads. The beads were washed by PBS and was used for SDS-PAGE analysis. The protein EGFR on beads was probed with anti-EGFR antibody via western blotting (**Figure 3.8C left panel**). As a result, EGFR was well expressed showing the molecular weight around 135 kDa. With successful expression in A431, we started to verify the activity via *in vitro* ubiquitination assays using the wt UB transfer cascade through wtUba1-wtUbcH5b-wtCbl-b cascade. The reaction condition was the same to the autoubiquitination of Cbl-b. Unfortunately, the polyubiquitin chain was not formed on the EGFR (**Figure 3.8C right panel**), the result with the xUba1-xUbcH5b-xCbl-b (either LZ1 or LZ5) was also not successful. Although the substrate confirmation failed to by wtUB transfer through wtUba1-wtUbcH5b-wtCbl-b and xUba1-xUbcH5b-xCbl-b cascades, the autoubiquitination of

xCbl-b was verified. We have decided to generate the stable cell line expressing xUba1-xUbcH5b-xCbl-b. *In vivo* reconstitution assay will be further studied and verified.

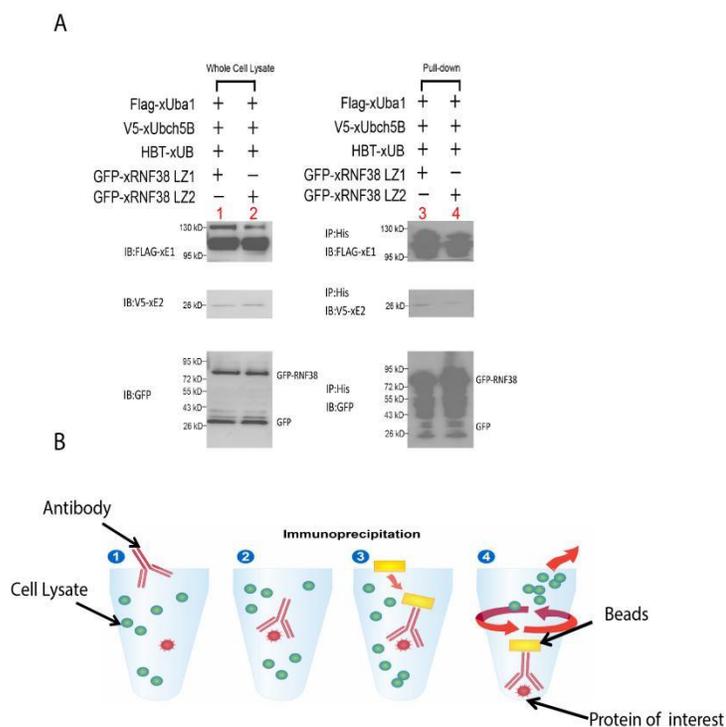


**Figure 3.8** *In vitro* ubiquitination assays of Cbl-b (38~427) variants. (A) Autoubiquitination of GST-tagged wild-type Cbl-b and Cbl-b/RNF38 Ring fusion mutant; (B) Autoubiquitination of GST-tagged Cbl-b mutants, xCbl-b-LZ1 and xCbl-b-LZ5. (C) Left panel: the expression of EGFR; right panel: *in vitro* ubiquitination of EGFR by wild-type Cbl-b.

### 3.3.5 Reconstitution of xRNF38 OUT activity *in vivo*

The mammalian HEK293T cell lines were employed to stably express each individual component of OUT including xUba1 and xUbcH5b. xRNF38 and HBT-xUb(His-Biotin-xUb) were subsequently co-transfected into the HEK293T to initiate the OUT cascade. xUba1, xUbcH5b and

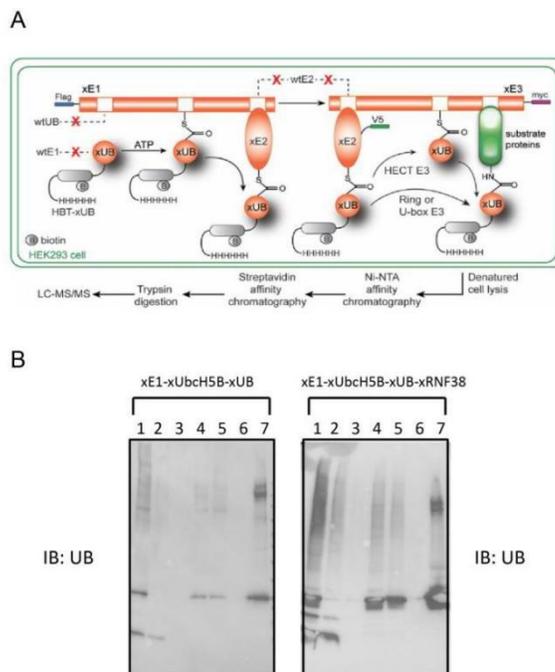
HBT-xUB were respectively cloned into mammalian pLenti vector for expression with various antibiotic resistances. Besides, xRNF38 full-length was introduced into pEBG vector( a gift from Dr. Kassenbrock's lab) to be expressed *in vivo* because RNF38 fused to pLenti vector failed to express after we tried 3 times of transient transfections in HEK923T cells. At 48 hr post-transfection of xRNF38 and HBT-xUB, the cells were incubated for 4 hours in 25  $\mu$ M MG132 (Sigma-Aldrich) then collected in lysis buffer (RIPA buffer) since each component was fused with a unique epitope tag for detection. According to results with *in vivo* pulldown assay, we found xUB could be transferred through xUba1-xUbch5b-xRNF38 cascade to form xUB-xRNF38 conjugate in the cell. As a result, GFP-xRNF38 could be co-purified with HBT-xUB in cells expressing the HBT-xUB and the xUba1-xUbch5b cascade (**Figure 3.9A-B**). Hence, the OUT cascade of xRNF38 was achieved *in vivo*.



**Figure 3.9 Expressing the OUT cascade of xRNF38 and HBT-xUB in HEK293T cells. (A)** Expression of the components of the xUba1- xUbch5b-xRNF38 cascade was confirmed in cells (lane 1-2). To confirm xUB transfer through the OUT cascade, lysates of the cells expressing the

OUT cascade of xRNF38 and HBT-xUB were immunoprecipitated by Ni-NTA agarose beads, and the co-purification of xUba1, xUbcH5b, and xRNF38 in the xUB-conjugated fraction was confirmed by Western blotting (lane 3-4); **(B)** a scheme for immunoprecipitation.

Having confirmed the orthogonality of the OUT cascades *in vivo* through immunoprecipitation and probing for the Flag, V5 and GFP tags, we next sought to enrich xUB-conjugated substrate proteins interacting with xRNF38 for proteomics analysis. We used the constructed HEK293T cell lines stably expressing the OUT cascades of RNF38. We transiently transfected the cell lines to express HBT-xUB and GFP-xRNF38. To isolate the substrate proteins, cells were lysed, and xUB-conjugated proteins were purified by tandem affinity chromatography with Ni-NTA and streptavidin resin under denaturing conditions. The substrates bound to the streptavidin resin were digested by trypsin and identified by mass spectrometry (MS) proteomics **(Figure 3.10A-B)**. To filter away proteins bound non-specifically to the resin or conjugated to xUB independent of xE3, HBT-xUB expression was also initiated in stable cell lines that expressing the xUba1-xUbcH5b pair without xRNF38. xUB-conjugated proteins were purified from the control cells and identified by LC-MS. Comparison of the profiles of xUB-conjugated proteins from cells expressing the OUT cascade and the control cells identified proteins that were dependent on xRNF38 for modification by xUB. These proteins are candidates that directly interacted with RNF38 in the cell and underwent ubiquitination to a greater extent.



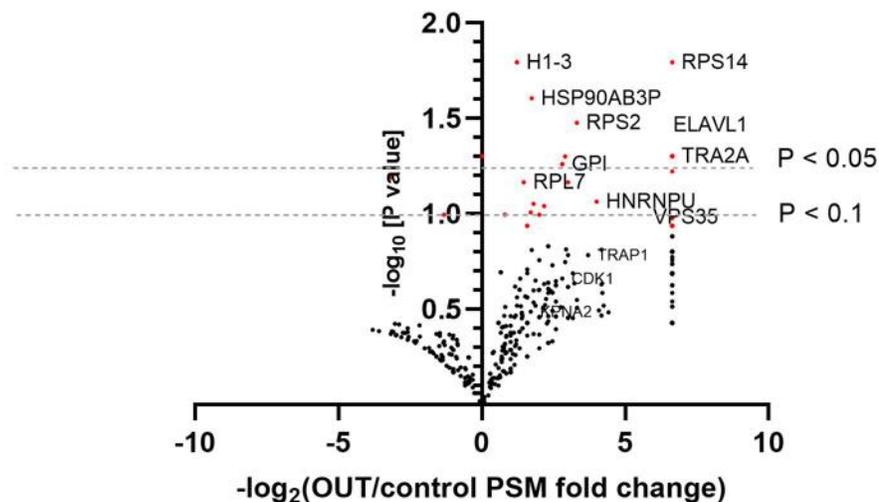
**Figure 3.10 Purification of xUB-modified proteins using affinity chromatography to determine the RNF38 substrates. (A)** Scheme to show cells expressing the OUT cascade of xRNF38, tandem affinity purification of xUB-conjugated proteins from the cell lysate, and identification of purified proteins by LC-MS/MS; **(B)** Purification of xUB conjugated proteins from lysates of cells expressing the full OUT cascade of RNF38 and the control without xRNF38 expression (lanes 1-7). The Western blots of the gels were probed with an anti-UB antibody. Lane 1, cell lysate; lane 2, flow-through from the Ni-NTA column; lane 3, wash solution of the Ni-NTA column; lane 4, elution from the Ni-NTA column; lane 5, flow-through from the streptavidin column; lane 6, wash solution of the streptavidin column; lane 7, protein bound to the streptavidin beads after washing.

Following this protocol, we developed xUB-conjugated protein profiles in positive cell lines expressing the RNF38 OUT cascade and in control cells expressing the xUba1-xUbcH5b pair lacking xRNF38. To compare the two profiles, we calculated the ratio of peptide spectrum matches (PSMs) for each protein isolated from cells with the OUT cascade to those sorted from control cells. Tandem purification was repeated three times to identify proteins showing 2-fold or higher PSM ratio between OUT and control cells in at least two of the repeats. **Table 4** lists the potential substrate proteins having a PSM ratio equal to or higher than two in 2 out of 3 repetitions. Currently,

only a few substrates for RNF38 were verified as known substrates, so the enriched targets from our OUT-screen profile would be novel substrates as shown in **Figure 3.11**. According to P value < 0.1 in volcano plot, indicating the statistical significance, some physiologically relevant targets in RNF38 OUT screen were evaluated and chosen for substrate verification, as mentioned in the following section. *In vitro* and *in vivo* substrate verification should be further carried out to confirm the authenticity of OUT cascade selected hits.

**Table 4 Potential substrates of RNF38 identified by proteomics**

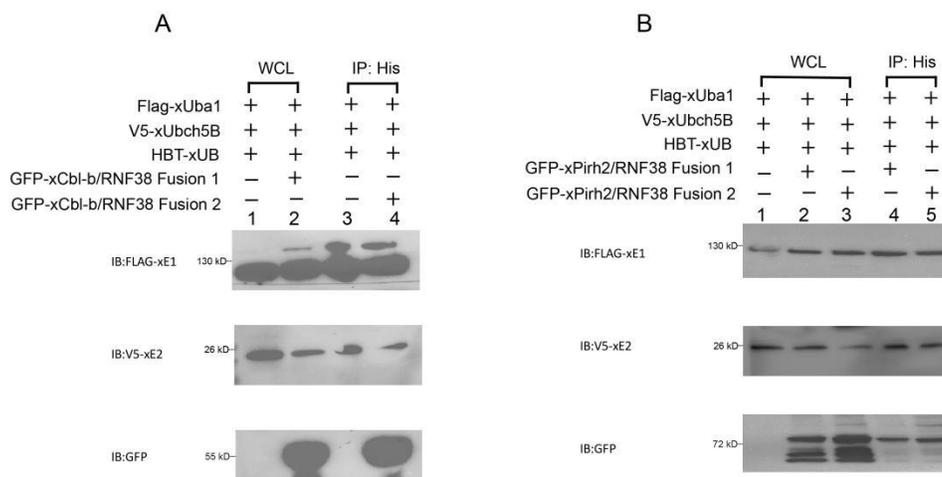
Description	Gene	xRNF38_PSM_1	xRNF38_PSM_2	xRNF38_PSM_3	xRNF38_Ctrl_PSM_1	xRNF38_Ctrl_PSM_2	xRNF38_Ctrl_PSM_3	Log2_PSM_Ratio_1	Log2_PSM_Ratio_2	Log2_PSM_Ratio_3
Adenylosuccinate synthetase isozyme 2	ADSS2	1	0	1	0	0	0	6.644	6.644	6.644
Cell cycle and apoptosis regulator protein 2	CCAR2	1	0	1	0	0	0	6.644	6.644	6.644
Cyclin-dependent kinase 1	CDK1	1	1	5	2	0	1	-1.000	6.644	2.322
Chloride intracellular channel protein 1	CLIC1	1	0	1	0	0	0	6.644	6.644	6.644
ELAV-like protein 1	ELAVL1	1	1	1	0	0	0	6.644	6.644	6.644
Neutral alpha-glucosidase AB	GANAB	1	0	1	0	0	0	6.644	6.644	6.644
Glutamate dehydrogenase 1	GLUD1	1	1	1	0	0	0	6.644	6.644	6.644
Glucose-6-phosphate isomerase	GPI	3	1	3	1	0	0	1.585	6.644	6.644
COP9 signalosome complex subunit 1	GPS1	1	0	1	0	0	0	6.644	6.644	6.644
Heterogeneous nuclear ribonucleoprotein H3	HNRNPH3	1	1	1	0	0	0	6.644	6.644	6.644
Isoleucine-tRNA ligase, cytoplasmic	IARS	1	0	1	0	0	0	6.644	6.644	6.644
Beta-Klotho	KLB	1	1	1	0	0	0	6.644	6.644	6.644
Importin subunit alpha-1	KPNA2	2	2	9	5	0	4	-1.322	6.644	1.170
Asparagine-tRNA ligase, cytoplasmic	NARS1	1	0	1	0	0	0	6.644	6.644	6.644
RNA cytosine C(5)-methyltransferase NSUN2	NSUN2	1	0	1	0	0	0	6.644	6.644	6.644
2-oxoglutarate dehydrogenase, mitochondrial	OGDH	1	0	1	0	0	0	6.644	6.644	6.644
Serine/threonine-protein phosphatase PGAM5	PGAM5	1	0	1	0	0	0	6.644	6.644	6.644
Phosphoglucosmutase-1	PGM1	2	0	4	0	1	0	6.644	-6.644	6.644
GTP-binding nuclear protein Ran	RAN	2	2	5	1	1	1	1.000	1.000	2.322
RING finger protein 44	RNF44	5	31	27	0	0	0	6.644	6.644	6.644
Splicing factor, proline- and glutamine-rich	SFPQ	5	2	2	0	1	1	6.644	1.000	1.000
Transformer-2 protein homolog alpha	TRA2A	1	1	1	0	0	0	6.644	6.644	6.644
Voltage-dependent anion-selective channel protein 2	VDAC2	1	0	1	0	0	0	6.644	6.644	6.644
Vacuolar protein sorting-associated protein 13D	VPS13D	1	0	1	0	0	0	6.644	6.644	6.644
Vacuolar protein sorting-associated protein 35	VPS35	1	0	1	0	0	0	6.644	6.644	6.644



**Figure 3.11** Volcano plot for substrates profile of RNF38

### 3.3.6 Reconstitution of xCbl-b and xPirh2 OUT activity *in vivo*

Like xRNF38, the stable cells expressing xCbl-b and xPirh2 were generated, respectively. Briefly, due to the success of Ring-type RNF38 expression in HEK923T cells, the xCbl-b (38~427) and full-length xPirh2 were incorporated into the pEBG vector for *in vivo* expression. At 48 hours post-transfection, xCbl-b/HBT-xUB and xPirh2/HBT-xUB cells were respectively treated for 4 hours in 25  $\mu$ M MG132 (Sigma-Aldrich) and then collected in lysis buffer (RIPA buffer) since each component was fused with a distinct epitope tag for detection. As a consequence, xUB can be transferred through either xUba1-xUbcH5b-xCbl-b (**Figure 3.12A**) or xUba1-xUbcH5b-xPirh2 cascade (**Figure 3.12B**), resulting in the formation of xUB-xE3(xCbl-b or xPirh2) conjugate.

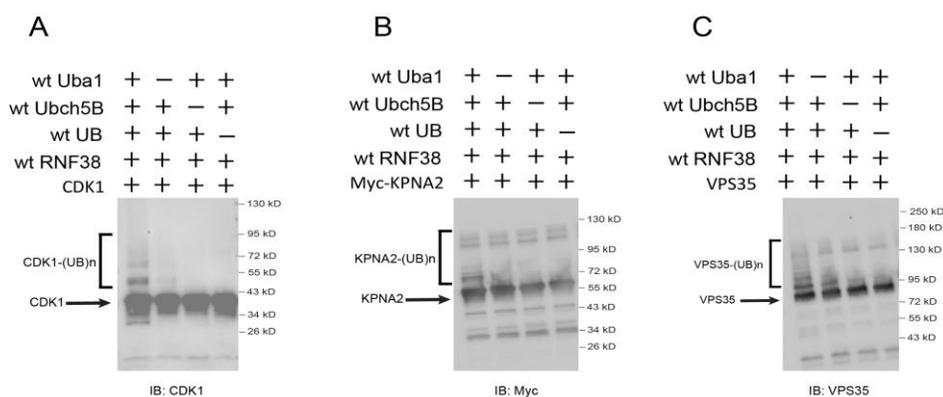


**Figure 3.12 Expressing the OUT cascade of xCbl-b&HBT-xUB and xPirh2&HBT-xUB in HEK293T cells, respectively.** (A) Expression of the components of the xUba1- xUbch5b-xCbl-b cascade was confirmed in cells (lane 1-2). To confirm xUB transfer through the OUT cascade, lysates of the cells expressing the OUT cascade of xCbl-b and HBT-xUB were immunoprecipitated by Ni-NTA agarose beads, and the co-purification of xUba1, xUbch5b, and xCbl-b in the xUB-conjugated fraction was confirmed by Western blotting (lane 3-4); (B) the same to xCbl-b, expression of the components of the xUba1- xUbch5b-xPirh2 cascade was confirmed in cells (lane 1-3). Immunoprecipitation by Ni-NTA agarose beads confirmed the xUB transfer through the xUba1-xUbch5b-xPirh2 cascade and the xUB-conjugated fraction was confirmed by Western blotting (lane 4-5)

### 3.3.7 Verifying substrates of RNF38 by *in vitro* ubiquitination

To validate the potential hits from our proteomics analysis, we first found some key signaling proteins including CDK1, VPS35, KPNA2, ELAVL1, MAGOH, MAGOHB, SFPQ, and PCNA are likely substrates of RNF38 (as analyzed in Table 2). We thus assayed if RNF38 targets them for ubiquitination and regulates the stability in the cell. We carried out similar protocols to the previously mentioned assays on putative substrate p53. At the time of writing the thesis, we only have expression constructs CDK1, VPS35 and KPNA2 available. So, we first expressed and purified the potential substrates CDK1, VPS35 and KPNA2 from *E. coli* by applying Ni-NTA purification method, and the potential purified proteins were subject to react with wtUba1-

wtUbcH5b-wtUB-wtRNF38 in the tube. Through the wtUb transfer by Uba1-UbcH5b-RNF38 cascade, the ubiquitin molecules were able to be loaded onto the potential substrate proteins CDK1, VPS35 and KPNA2 via forming the polyubiquitination chains (**Figure 3.13**), which suggesting the direct interaction between substrates and RNF38. Protein expressed in *E. coli* cells may not bear the proper post-translational modifications for RNF38 recognition, so the *in vitro* assays may not reflect the real activity of RNF38 with the substrate proteins. Further *in vivo* assays were essential to verify whether the potential substrates are ubiquitinated by RNF38 in the cell.



**Figure 3.13** *In vitro* assays to test the ubiquitination of RNF38 substrates identified by OUT. The wt UB was transmitted to the possible substrate proteins produced from *E. coli* cells through the wt Uba1-UbcH5b-RNF38 cascade. CDK1(A), KPNA2(B) and VPS35(C), were all verified to be ubiquitinated by RNF38.

### 3.3.8 General design of verifying substrates for E3s *in vivo* for future work

We will silence RNF38 in HEK293T cells with lentivirus delivering the anti-RNF38 shRNA. We will also overexpress RNF38 in blank HEK293T cells and cells harboring the anti-RNF38 shRNA. Then, cells will be treated with proteasome inhibitor MG132 before harvesting to

inhibit protein degradation. Ubiquitination levels of various substrates in different cell populations should be revealed by immunoprecipitation with substrate-specific antibodies and immunoblotting with an anti-UB antibody. Comparing to the parental HEK293 cells, cells expressing anti-RNF38 shRNA should presumably have lower levels of poly-ubiquitinated forms of substrates. The poly-ubiquitination of each target protein in the HEK293 cells harboring the anti-RNF38 shRNA can be restored by overexpressing RNF38 in the cell. Furthermore, HEK293T cells with overexpression of RNF38 will give rise to more intense poly-ubiquitination of substrates comparing to the parental HEK293T cells. Thus, these experiments will be established and conducted later.

### 3.4 Discussion

Up to now, by engineering the interface of xUbcH5b and RNF38 Ring via phage selection, we successfully restored the binding activity between xUbH5b and xRNF38 and completed the full platform termed as “orthogonal ubiquitin transfer (OUT)” *in vitro*. Furthermore, we transferred the xUba1-xUbcH5b-xRNF38-xUB into the mammalian cell system (HEK293T) and confirmed the exclusive capability of xUB transfer through xUba1-xUbcH5b-xRNF38 cascade *in vivo*, which convinced us of performing tandem purification via a sequential pull-down of Ni-NTA and streptavidin resins since 6-His and Biotin tags are fused to xUB. The xUB-conjugated proteins were analyzed by proteomics. As a result, we were able to identify approximately 200 potential substrates of RNF38. A key merit of OUT is that the substrates are modified by ubiquitination via the direct interaction with RNF38. Other methods based on E3 substrate binding or change of protein expression level upon perturbation of E3 activity, use indirect readouts of substrate ubiquitination to assign E3 substrates. In this project, we used phage display and screened out the RNF38 mutants to generate the xE2-xE3 pair for the OUT cascade. The OUT cascade in HE293T cells significantly enriched several targets of RNF38, such as CDK1, KPNA2, ELAVL1, MAGOH,

MAGOHB, SFPQ and PCNA. Among the potential substrates, we successfully verified that CDK1, KPNA2 and VPS35 could be ubiquitinated by RNF38 *in vitro* via monoubiquitination and polyubiquitination respectively (**Figure3.9**). CDK1 is required for all cell types for proliferation. The dysfunction of CDK1 results in embryonic death, whereas conditional CDK1 inhibition in postnatal animals inhibits cell growth<sup>[145]</sup>. Also, CDK1 is activated by A- or B-type cyclins and phosphorylates a vast variety of proteins involved in DNA replication, mitotic progression, and cell division<sup>[146,147]</sup>. VPS35 was firstly identified in yeast, which is responsible for the assembly of lysosome-like vacuoles and the sorting of vacuolar proteins<sup>[148]</sup>. The human homologue of yeast VPS35, termed as hVPS35, is a 796-residue protein. It has been reported that the mutations on VPS35 were associated with Parkinson's disease<sup>[148,149]</sup>. Currently, both E3 ligases, Triad3a<sup>[150]</sup> and Parkin<sup>[151]</sup>, can ubiquitinate VPS35 for cellular regulation. Thus, RNF38 may be potentially involved in neuro development. Interestingly, a concern has also be raised that why different E3s target the same substrate protein. On the one hand, this could be due to the structural similarities of the E3s; on the other hand, the expression of different E3s in diverse tissues could be different. Only the highly expressed E3 then acts as a ligase, transferring the ubiquitin molecules to the substrate proteins. MAGOH is involved in the EJC (exon junction complex) and the nonsense-mediated mRNA decay (NMD) pathways<sup>[152]</sup>, where it acts to preserve the integrity of mRNAs. Moreover, MAGOHB, a paralog of MAGOH in vertebrates, is also correlated to mRNAs via EJC and NMD<sup>[153,154]</sup>. Besides the function in mRNA processing, it has been discovered that the MAGOH paralogs (MAGOH and MAGOHB) are related to cancer development. In gastric cancer cells, silencing MAGOH-MAGOHB prominently displayed anti-tumor effects through the BRAF/MEK/ERK signaling pathway<sup>[155]</sup>. Thus, the potential ubiquitination of MAGOH paralogs identified by RNF38 OUT screen may assist to explain the roles of RNF38 E3 in cells. ELAVL1

encodes a protein known as HuR, which is a member of the embryonic lethal abnormal vision (ELAV)/Hu family<sup>[156]</sup>. HuR associates with U- and AU-rich motifs in the 3'-UTR region of a vast number of messenger RNAs and controls their nuclear cytoplasmic shuttling and splicing by modulating their stability<sup>[157]</sup>. To date, Pirh2 and TRIM21 are the only two E3 ligases that can ubiquitinate HuR for proteasomal degradation and induces the upregulation of c-Myc (a proto-oncogene that is often constitutively persistently expressed in cancers). Since ELAVL1 appeared in our 3 repetitions, showing significance ( $P < 0.1$ ) in RNF38 OUT screen, we then decide to start to express the protein in *E. coli* and set up ubiquitination *in vitro* for verification. RNF38 would be likely the third E3 ligase to regulate the expression of c-Myc and other downstream targets of HuR, thereby promoting carcinogenesis. Meanwhile, HuR will be recognized as a novel substrate of RNF38, suggesting RNF38 can be a therapeutic target in different carcinomas. In addition, we found another RNA binding protein SPFQ (Splicing factor, proline- and glutamine-rich) controls the translation of microRNAs at the local level and increases neurotrophin-dependent axon viability<sup>[158]</sup>. The perturbation of SPFQ activity, expression, or localization has been correlated to various diseases such as Alzheimer's disease<sup>[159]</sup>. PCNA (proliferating cell nuclear antigen) is also reported to be involved in DNA synthesis. By controlling the maturation and assembly of Okazaki fragments, the K63-linked polyubiquitinated-PCNA protects replication forks against DNA2-mediated destruction<sup>[160]</sup>. Through the overview of potential substrates of RNF38 OUT screen, it enlightens us that RNF38 are involved in many cellular regulation pathways, especially DNA/RNA replication and damage as well as cancer progression. Thus, the OUT screen indeed provides an efficient way to help identify the direct substrate proteins of an E3 or elucidate their biological functions.

Although the substrate verification of RNF38 *in vivo* will be delayed before my graduation, the success of RNF38 OUT cascade both *in vitro* and *in vivo* would be a solid base to continue my research on confirming the substrates of RNF38 during my postdoc stage. I believe all the efforts will pay off in the end with this fantastic story. As for Cbl-b and Pirh2, we will carry out more rounds of tandem purification until we got 3~4 datasets ready for each E3, and we will then start to analyze the potential substrates for them.

### **3.5 Methods**

#### **3.5.1 Generation of Lentiviral constructs**

To generate pLenti6-V5-D-TOPO-Asc1-hygromycin-HBT-(x)UB plasmids, HBT tag was sub-cloned from pQCXIP HBT-UB and fused with DNA fragments of human wt UB or xUB by PCR. The assembled DNA fragment was cloned into the pLenti6 plasmid with a hygromycin resistant gene. Genes of xUba1 and xUbcH5b were cloned into lentiviral vectors for the selection of stable cell lines. Flag-xUba1 gene was PCR amplified with primer WY15 and primer WY16 and cloned into pLenti6-V5-D-TOPO-Flag-Asc1-blasticidin vector between restriction sites EcoRI and AscI. V5-xUbcH5b gene was PCR amplified from pET-xUbcH5b with PCR primers WY17 and WY18, digested with restriction enzymes AfeI and Nhe I, and cloned into pLenti4-V5-D backbone with a zeocin-resistance gene. The gene of xRNF38 was PCR amplified with primers FZ06 and FZ07 and cloned into a pEBG vector with a GFP tag between restriction sites BsrG I and Sal I.

#### **3.5.2 Construction of stable cell line**

Virus packaging, virus infection and selection of stable cell lines were performed according to the manufacturer's protocol for the ViraPower Lentiviral Expression System. Stable HEK293T

cell lines expressing Flag-xUba1 and V5-xUbcH5b were selected with 10  $\mu\text{g}/\text{mL}$  blasticidin and 100  $\mu\text{g}/\text{mL}$  Zeocin, respectively. All recombinant DNA manipulations involving xE3 enzymes were accomplished in NEBStable cells (NEB), which possess greatly enhanced DNA stability and reduced possibility of recombination.

### **3.5.3 Production of lentiviral particles**

We generally followed the existing protocols according to the Virapower<sup>TM</sup> kit from Invitrogen. In 1.5 ml of Opti-MEM I medium without serum, 9  $\mu\text{g}$  of Virapower<sup>TM</sup> Packaging mix was mixed with  $\sim 3$   $\mu\text{g}$  of the appropriate pLenti construct and added to a 5-minute preincubated solution of 36  $\mu\text{L}$  of Lipofectamine<sup>TM</sup> 2000 reagent in 1.5 ml Opti-MEM without serum. The mixture was incubated for 20 minutes at room temperature to allow to formation of DNA-containing liposomes, and afterwards added dropwise to a plate of cultured 293T cells at 90-95% confluency in 5 ml of Opti-MEM. The cells were then allowed to incubate overnight at 37°C at 5% CO<sub>2</sub>. The next morning, the liposome-containing medium was discarded and replaced with 10 ml of fresh complete culture medium without antibiotics; the cells were then allowed to grow for a further 48-72 hours at 37°C at 5% CO<sub>2</sub>. Afterwards, the supernatant containing the lentiviral particles was harvested and cleared by filtration.

### **3.5.4 Lentiviral Transduction**

The lentiviral solution was diluted 2-fold in complete media and then used to for transfection by addition to plated non-confluent HEK293T cells, supplemented with 10  $\mu\text{g}/\text{ml}$  Polybrene®; the cells were incubated overnight at 37°C at 5% CO<sub>2</sub>. The following day, the virus-containing medium was replaced with fresh complete medium, and the cells were allowed to grow under the same conditions for one more overnight. Afterwards (on the third day of transduction),

the medium was once again replaced, this time with medium supplemented with the appropriate antibiotic (Depending on the construct, the concentrations were: blasticidin 10 µg/ml, hygromycin 200 µg/ml, zeocin 100 µg/ml, and puromycin 1 µg/ml); the antibiotic was maintained throughout the whole course of selection (lengths of selection were: blasticidin 5 days, hygromycin 7 days, zeocin 10 days, and puromycin 3-5 days) with the medium being changed every 3-4 days. Antibiotic-resistant colonies were identified, picked, and expanded, being now regarded as stable cell lines. The expression of OUT proteins of interest were confirmed by Western immunoblotting against the appropriate epitope antibody-to this end the cells could directly be lysed by boiling in Laemmli buffer supplemented with DTT or BME.

### **3.5.5 DNA Transient Transfection**

According to the manufacturer's manual (Horizon), for 75 cm<sup>2</sup> flask, seed cells in 10 mL of DMEM growth medium 24 hours prior to transfection at a density that will give a confluency of 70-90% for adherent cells on the day of transfection. Dilute DNA in 500 µL per well of serum-free DMEM or other serum-free growth medium. Gently mix DharmaFECT kb transfection reagent by pipetting. Dilute DharmaFECT kb reagent in 500 µL per well of serum-free DMEM or other serum-free growth medium. Add 500 µL diluted DharmaFECT kb reagent to the diluted DNA for a total of 1 mL per well. Incubate 20 minutes at room temperature. Gently aspirate medium from cell culture plate. Immediately dispense 1 mL of the transfection reagent/DNA mixture to flask. Gently rock the plate to achieve even distribution of the complexes. Incubate at 37°C in a CO<sub>2</sub> incubator. Begin to analyze transgene expression 24~72 hours later.

### **3.5.6 *In vivo* Pulldown assays**

#### **3.5.6.1 Lysis of cells**

In general, HEK293T cells were lysed in RIPA buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.1 percent SDS, 0.5 percent sodium deoxycholate, and 1% Triton-X) supplemented with 1 mM PMSF. Alternatively, as was the case with HBT tandem purification of xUb-conjugated proteins, the cells are lysed under denaturing conditions using Buffer A (50 mM sodium phosphate pH 8.0, 300 mM NaCl, 8 M Urea, 0.5 percent NP-40) and 1 mM PMSF. Following lysis, lysates were cleaned by sonication and centrifugation at a speed of 10,000+ rpm for at least 10 minutes to eliminate cellular debris.

#### **3.5.6.2 Immunoprecipitation for target proteins**

Preclear whole cell lysate (optional step) as follows: add 0.25 g of the appropriate control IgG (corresponding to the primary antibody's host species) to approximately 1 mL of whole cell lysate, followed by 20  $\mu$ L of the appropriate suspended (25 percent v/v) agarose conjugate (Protein A/G PLUS-Agarose: sc-2003). Incubate for 30 minutes at 4°C. Centrifugation for pelleting beads at 3,000 rpm (about 1,000xg) for 30 seconds at 4°C. Supernatant (cell lysate) should be transferred to a fresh microcentrifuge tube at 4°C. Add 10 g of primary antibody agarose conjugate to 1 mL of the above cell lysate, or roughly 100-1000g of total cellular protein, and incubate at 4°C for 1 hour to overnight with mixing. Alternatively, if the primary antibody agarose conjugate is not available, incubate 1 ml cell lysate for 1-2 hours at 4° C with 1-10  $\mu$ L primary antibody. Add 20  $\mu$ L of agarose conjugate suspension (Protein A/G PLUS-Agarose: sc-2003). Cap tubes and incubate for 1 hour to overnight at 4°C on a rocker platform or rotating device. Centrifugation at 3,000 rpm (about 1,000xg) for 30 seconds at 4°C to collect pellet. Aspirate and discard supernatant

carefully. The agarose was rinsed four times with PBS; samples for SDS-PAGE and Western blotting were prepared by boiling the agarose in Laemmli solution to liberate a fraction of the bound proteins.

### 3.5.6.3 Tandem affinity purification of xUB-conjugated proteins

To enrich the HBT-xUb ubiquitinated proteins, tandem purification of HBT-xUB-conjugated proteins was performed as previously described<sup>[142]</sup>. Briefly, ten 75cm<sup>2</sup>-flasks of HEK293T cells stably expressing the xUba1-xUbcH5b were acutely co-transfected with xRNF38 and HBT-xUB for 48 hours. To inhibit proteasome activity, cells were treated with 10  $\mu$ M MG132 for 4 hours at 37°C. Cells were then washed twice with ice-cold 1 $\times$  phosphate-buffered saline (PBS) (pH 7.4) and harvested by a cell scraper with buffer A [8 M urea, 300 mM NaCl, 50 mM tris, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 0.5% NP-40, 1 mM phenylmethylsulfonyl fluoride, and benzonase (125 U/ml) (pH 8.0)]. For Ni-NTA purification, cell lysates were centrifuged at 15,000g for 30 min at room temperature. Thirty-five microliters of Ni<sup>2+</sup> Sepharose beads (GE Healthcare) for each 1 mg of protein lysates was added to the clarified supernatant. After incubation overnight at room temperature in buffer A with 10 mM imidazole on a rocking platform, Ni<sup>2+</sup> Sepharose beads were pelleted by centrifugation at 100g for 3 min and washed sequentially with a 20-bead volume of buffer A (pH 8.0), buffer A (pH 6.3), and buffer A (pH 6.3) with 10 mM imidazole. After washing the beads, proteins were eluted twice with a 5-bead volume of buffer B [8 M urea, 200 mM NaCl, 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 2% SDS, 10 mM EDTA, 100 mM tris, and 250 mM imidazole (pH 4.3)]. For streptavidin purification, the elution solution was adjusted to pH 8.0. Fifty microliters of streptavidin Sepharose beads (Thermo Fisher Scientific) was added to the elution to bind HBT-xUB-conjugated proteins. After incubation on a rocking platform overnight at room temperature, streptavidin beads were pelleted and washed sequentially with 1.5 ml of buffer C [8 M urea, 200

mM NaCl, 2% SDS, and 100 mM tris (pH 8.0)], buffer D [8 M urea, 1.2 M NaCl, 0.2% SDS, 100 mM tris, 10% EtOH, and 10% isopropanol (pH 8.0)], and buffer E [8 M urea and 100 mM NH<sub>4</sub>HCO<sub>3</sub> (pH 8)].

#### **3.5.6.4 Sample digestion and LC-MS/MS analysis**

After washing, the streptavidin beads were spun down, residual urea was removed, and liquid chromatography coupled to tandem MS (LC-MS/MS) on an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific) was performed at the Emory Integrated Proteomics Core (EIPC) according to previously published methods<sup>[156,157]</sup>. The spectra collected were compared to the human UniProt database using Proteome Discoverer 2.0. (90,300 target sequences). Fully tryptic restriction and a parent ion mass tolerance of  $\pm 20$  parts per million were used as search criteria. Methionine oxidation (+15.99492 Da), asparagine and glutamine deamidation (+0.98402 Da), lysine ubiquitination (+114.04293 Da), and protein N-terminal acetylation (+42.03670 Da) were all variable modifications (up to three per peptide were permitted); cysteine received a fixed carbamidomethyl modification (+57.021465 Da). The PSMs were filtered using a percolator to achieve a 1% false discovery rate.

#### **3.5.6.5 Bioinformatics analysis**

IPA software ([www.ingenuity.com](http://www.ingenuity.com)) was used to map and identify the biological networks and molecular pathways with a significant proportion of genes having RNF38 ubiquitination targets. Fisher's exact test in IPA software was used to calculate the P values for pathways and networks. The level of statistical significance was set at  $P < 0.05$ . IPA was also used to visualize the identified biological networks.

### 3.5.7 *In vitro* ubiquitination of substrate proteins

Genes of potential substrate proteins CDK1, VPS35 and KPNA2 were cloned into pET vector for protein expression by our previous lab mate. The plasmids were transformed into BL21 (DE3) cells to express the protein. To assay ubiquitination by RNF38, 5 to 10  $\mu$ M substrate proteins were incubated with 1  $\mu$ M wt Uba1, 5  $\mu$ M wt UbcH5b, and 50  $\mu$ M wt UB in TBS supplemented with 10 mM MgCl<sub>2</sub> and 5 mM ATP. After 2-hour reaction at room temperature, the ubiquitination of substrates was analyzed by Western blotting probed with either substrate-specific antibodies or antibodies against the tags fused to the substrates

**Table 5 Primary antibody information**

Antibody	Supplier	Catalog Number	Diluton
$\beta$ -actin	Santa Cruz	sc-47778	1:1000
Ub	Santa Cruz	sc-8017	1:600
FLAG	Sigma	F4042	1:2000~1:4000
GST	Santa Cruz	sc-138	1:500~1:1000
p53	Santa Cruz	sc-126	1:1000
V5	Santa Cruz	sc-271944	1:500
GFP	Santa Cruz	sc-9996	1:600
RNF38	OriGene	TA330460	1:300
CDK1	Santa Cruz	sc-54	1:300
VPS35	Santa Cruz	sc-374372	1:500
KPNA2	Santa Cruz	sc-55538	1:1000

**Table 6 Sequences of the primers used in this study**

Primer	Sequence
WY 15	5'-ACTGCTGAATTCACCCCAAGTCTGGCGTCAAG-3'
WY 16	5'-GGCTAGTGGCGCGCCTCATCAGCGGATGGTGTATCG-3'
WY 17	5'-ACCATGAGCGCTATGGCGGCCAGCGAGAGGCTGATGGAGGAGC-3'
WY 18	5'-ACCCTTGCTAGCGTCCACAGGTCGCTTTTCC-3'
FZ06	5'-GCGAGCTGTACAAGTCCGGACTCAGATCTATGGCTTGTAAGATATCTC-3'
FZ07	5'-GTACCGTCGACTCATTCTGAATCCCGATGC-3'

## 4 SUBSTRATES IDENTIFICATION OF RBR-TYPE E3 LIGASE HHARI/ARIH1

### 4.1 Structure and Function of HHARI/ARIH1

HHARI (human homolog of Ariadne), also known as ARIH1, is one of RBR-type E3 ligases composed of 557 amino acid residues (64 kDa) distributed in the cytoplasm and the nuclei of cells<sup>[59]</sup>. As is introduced in Chapter 1, RBR-type E3 ligases share the conserved module called RING1-(in-between-Ring) IBR-RING2. The RING1 adopting a typical cross-brace fold similar to the canonical RINGs that can bind to E2~UB conjugate, whereas the RING2 contains a conserved Cys residue that is not involved in Zn<sup>2+</sup> coordination but serves as the active site to accept UB from RING1 via a transthioesterification reaction<sup>[161]</sup>. Although IBR domains do not contain an active-site Cys, their sequences vary greatly among RBR E3s (Parkin, HOIP and HHARI) leading to the flexibility, thus allowing large scale conformational rearrangements that are required for transitioning between autoinhibited and active forms<sup>[162,163]</sup>. Thus, the IBR domain also holds a crucial role in the UB transfer from E2 to the RING2 active site cysteine and then the substrate<sup>[161]</sup>. Besides the conserved domain, HHARI also contains a C-terminal Ariadne domain, mainly formed by  $\alpha$ -helices, which masks the catalytic Cys357 of RING2 and blocks the binding to the E2~UB conjugate with RING1, so HHARI is autoinhibited<sup>[54,164]</sup>. Although the reasons for E3 autoinhibition are unclear, it seems that this might be a mechanism for avoiding misdirected substrate ubiquitination and so guiding UB ligation to substrates in the proper biological context.

Due to the unique structural characteristics of RBR E3s, they have been implicated to be involved in neurological disorders, infection, inflammation, DNA damage response and cancer as a result of their dysfunctional activity. RBR E3 ligases have been discovered to play essential roles in carcinogenesis by controlling the degradation of tumor promoters or suppressors in a variety of

malignancies. HHARI is one of RBR E3s that binds to neddylated CUL-type RING E3 ligases (CRLs) and the binding of HHARI stimulates the RBR ligase activity, leading to the ligation of UB to the client substrates to regulate cell progression<sup>[165]</sup>. Scott et al., proved that the interaction between HHARI and CRLs is essential for the activation of HHARI and they jointly regulate the substrate ubiquitination<sup>[165]</sup>, which explained the significance of the autoinhibition mechanism of HHARI. What's more, HHARI is abundantly expressed in various cancer cells, particularly in breast, lung and hepatocellular carcinomas<sup>[59,166]</sup>. In addition, HHARI was reported to be involved in the ubiquitination of defective mitochondria, prompting mitophagy, a biological process that removes damaged mitochondria via degradation by lysosome<sup>[59]</sup>. Also, they established that the removal of HHARI might predispose tumor cells to chemotherapy-induced mortality. Using yeast 2-hybrid analysis and Co-IP assays, Tan et al.,<sup>[167]</sup> demonstrated that 4EHP, also known as EIF4E2, was ubiquitinated by HHARI for proteasomal degradation as a way for regulating the gene expression. Homologous to EIF4E<sup>[168]</sup>, 4EHP recognizes and binds the 7-methylguanosine-containing mRNA cap during an early step in the initiation, acting as a repressor of translation initiation<sup>[169]</sup>. With identification of interaction between HHARI and 4EHP, Stechow et al., further illustrated the relationship between them and reported that HHARI promotes and fine-tunes a 4EHP-mediated mRNA translation arrest, which protects stem and cancer cells against genotoxic stress<sup>[170]</sup>. Relying on ATM (ataxia telangiectasia mutant), DNA damage increases the enrichment of HHARI protein by inhibiting its degradation and then HHARI protein accumulates and triggers translation arrest through 4EHP<sup>[170]</sup>. Based on HHARI/ARIH1's biological roles, we may possibly conclude that HHARI is an oncogenic protein. On the other hand, a recent study debated that HHARI could exert anti-tumor activity targeting PD-L1 for proteasomal degradation via HHARI-

dependent ubiquitination<sup>[60]</sup>, which indicates that HHARI could function as a tumor suppressor to enhance the effect of cancer immunotherapies with the development of HHARI-activating agents.

Similar to Ring-type E3s (RNF38, Cbl-b and Pirh2), HHARI plays a crucial role in cellular processes as introduced. Thus, mapping its substrates by OUT cascade can assist to elucidate its roles in many cellular signaling pathways. By combination of crystal structure analysis and phage selection, we could engineer HHARI to assemble an OUT cascade with xE1 and xE2 for the transfer of xUB to its substrate proteins to enable their identification.

The HHARI project was initially carried out by Dr. Geng Chen, a postdoctoral fellow in the lab. She contributed to the construction of HHARI phagemid, model selection and library selection of HHARI, finally she was able to select out the xHHARI mutants connecting with our established OUT cascade of xE1-xE2 *in vitro*. Then, I took over the project in 2018 and started to verify the orthogonality of the xUba1-xUbcH7-xHHARI cascade in the cell and identify the substrate proteins of HHARI. In this chapter, the results for model selection and library selection of HHARI will be briefly discussed due to similarity to RNF38 selection and I will mainly focus on the activity of xHHARI verification and substrate identification in the HEK293T cells.

## 4.2 Results

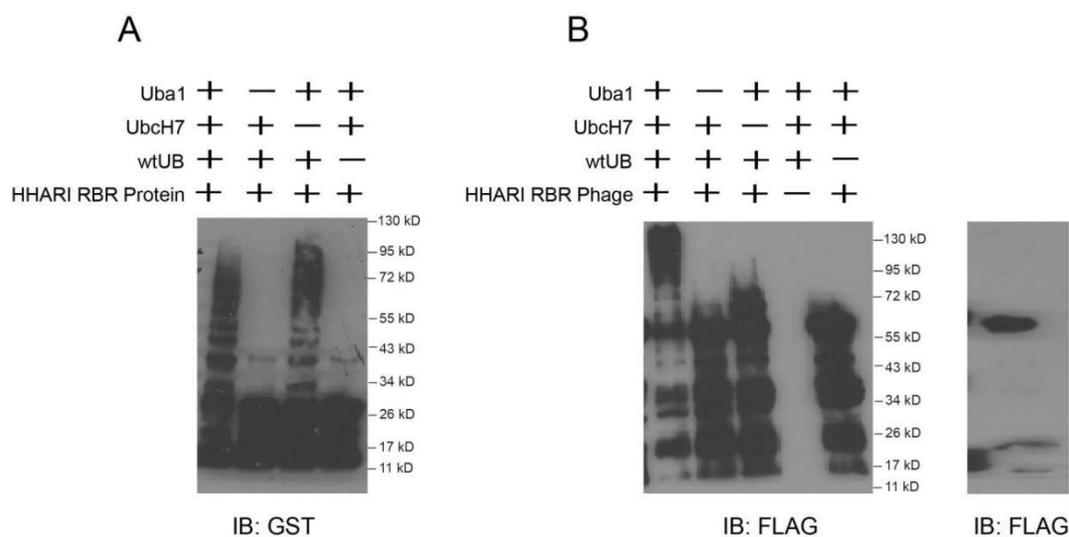
### 4.2.1 Phage display of the RBR domain of HHARI

#### 4.2.1.1 Testing activity of HHARI RBR domain *in vitro* autoubiquitination assays

The pGEX vector for the expression of HHARI RBR domain was transformed into BL21(DE3) and the RBR domain was well expressed and its activity was verified in an *in vitro* ubiquitination reaction with wt Uba1(E1), wt UbcH7(E2) and wt UB with an HA epitope tag. The

GST tag fused RBR domain was probed with anti-GST antibody, showing the strong ubiquitination signals in forming the polyubiquitination chain (**Figure 4.1A**). Reactions excluding E2 also showed ubiquitination activity because E2-independent ubiquitination is a common feature of RBR E3s<sup>[171,172]</sup>.

Similar to pComb-HHARI RBR cloning, the RBR domain of the E3 ligase HHARI was cloned into the phagemid vector pComb3H, together with a FLAG epitope tag at the protein's N-terminus and a TEV cleave site at the HHARI gene's C-terminus used for phage elution. The HHARI displaying phage particles were prepared and Western blotting of the phage probed with anti-Flag antibody showed the successful display of HHARI RBR on the M13 phage (**Figure 4.1C**). When the RBR-displaying phage was incubated with wild-type UB, Uba1 (E1), and UbcH5b (E2), poly UB chains were formed based on the Western blot analysis of the reaction mixture. This implies that the phage retains the enzymatic activity of the HHARI RBR (**Figure 4.1B**).

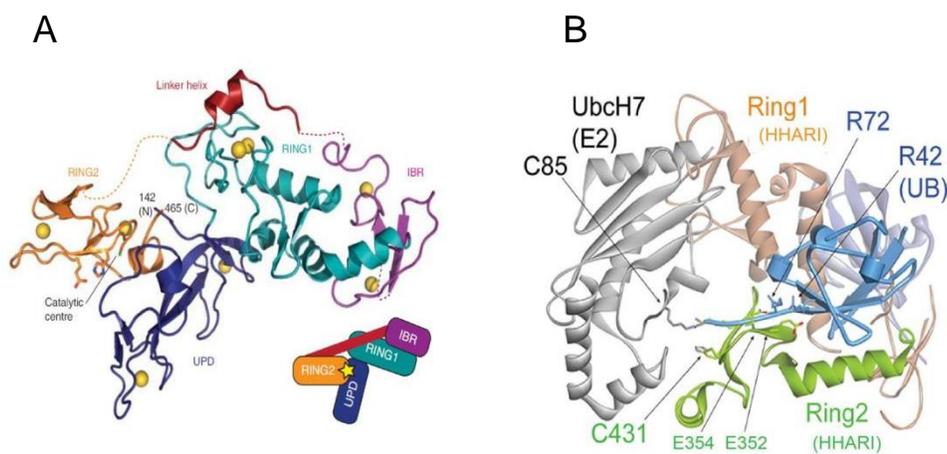


**Figure**

**4.1 *In vitro* auto-ubiquitination assays of HHARI RBR.** (A) Recombinant HHARI RBR domain protein (A) and phage displaying HHARI RBR with FLAG epitope displayed by the pComb3H system. (B), HHARI displaying phage particles were tested for auto-ubiquitination activity with wt UB, Uba1 (E1) and UbcH7 (E2) enzymes. The UB-conjugated HHARI RBR protein and RBR-displayed phage were immunoblotted by either anti-GST or anti-Flag antibodies.

#### 4.2.1.2 Library Construction of HHARI RBR through the analysis of modeled structure

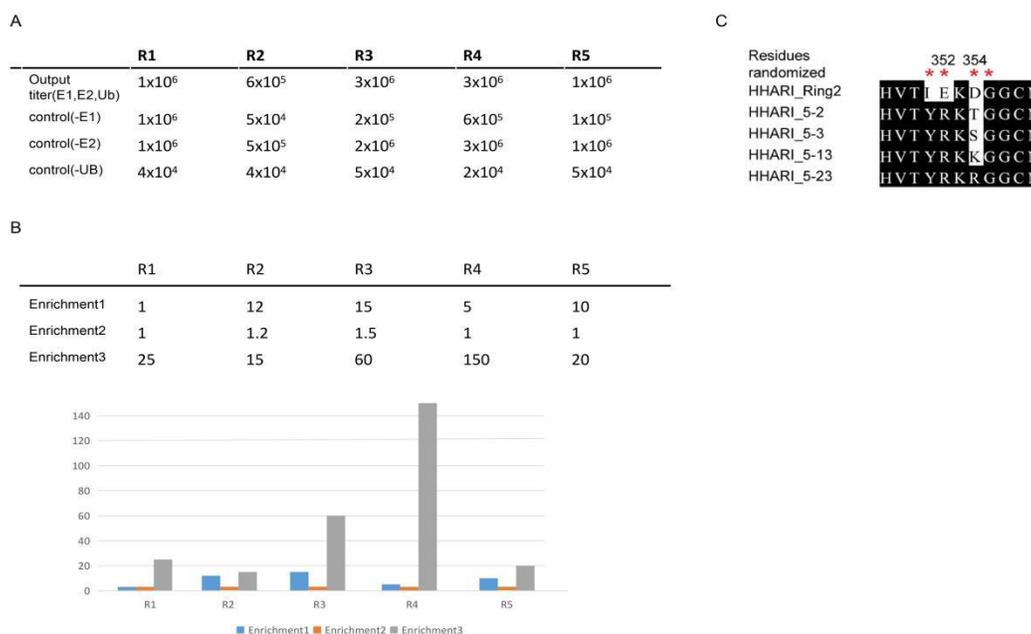
As described in 4.1, the crystal structure of HHARI showed the E3 in an autoinhibited state with the RBR domain in a complex with a C-terminal Ariadne domain that shields it from binding to the UB~E2 conjugate (**Figure 4.2A**)<sup>[54]</sup>. In contrast, the crystal structure of RBR E3 HOIP in complex with the UB~UbcH5b conjugate showed the E3 in a catalytically competent conformation with the RBR domain binding to UbcH5b before it transfers UB to the catalytic Cys residue of the RBR<sup>[171]</sup>. We assumed the RBR of HHARI would adopt a similar conformation as that of HOIP during UB transfer, so we modeled the RBR of HHARI into the HOIP structure (**Figure 4.2B**). The model showed that E352 and D354 of Ring 2 of HHARI RBR would interfere with the binding of xUB since they would repel the R42E and R72E mutations in xUB. We thus generated an RBR phage library of HHARI with randomized residues at I351, E352, D354 and G355 (**Figure 4.2B**).



**Figure 4.2 Analysis by Modeled crystal structure of HHARI RBR and Ubch7**

#### 4.2.1.3 Library Selection of HHARI RBR through biopanning assay

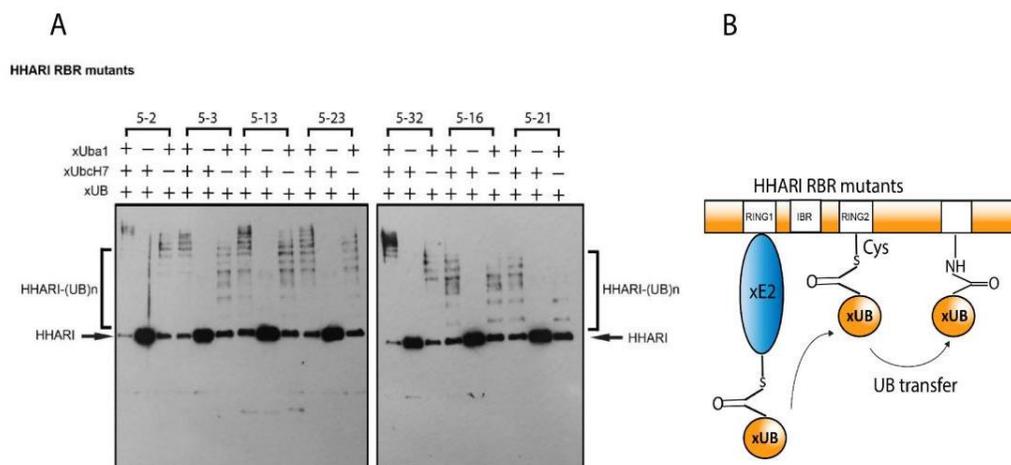
Following the same protocol as RNF38 selection, the phage library of HHARI RBR was charged with back-Uba1, xUbCH7 and biotin-wtUb, the phage reaction mixture was added on the plate with precoated streptavidin. Through the affinity binding between biotin and streptavidin, the active RBR phage particles conjugated with biotin-wtUb were captured by the plates, then a reaction for TEV cleavage was carried out to elute the active phage particles and they were used for further infection of the bacteria. We carried out 5 rounds for library selection and the mutants containing converged sequences were effectively enriched and selected out in fifth round, which suggesting the true activity of each mutant was retained to interact with xUba1 (mutations in UFD) and xUbCH7 (an engineered E2 that shares no cross-reactivity with wtE1) to be self-ubiquitinated via the addition of ubiquitin. We were ecstatic to learn that the sequences exhibit amazing convergence, especially after the fourth and fifth rounds of selection. this is also reflected in our enrichment table (**Figure 4.3A-B**), which improved significantly in the fourth round. Additionally, we identified 4 mutants HHARI 5-2, 5-3, 5-13 and 5-23 with that were most abundant in the selected phage pool (**Figure 4.3C**).



**Figure 4.3 Enrichment of phage selection and sequence alignment of potential xHHARI mutants.** (A) The selection conditions and titer numbers were listed the five rounds of selection. (B) The approximate enrichment values for each round of phage selection. Enrichment 1, 2 and 3 were referred as the ratio Output (E1, E2, Ub)/Output(-E1), Output (E1, E2, Ub)/Output(-E2) and Output (E1, E2, Ub)/Output(-Ub). (C) The sequence alignment after five rounds of phage selection

#### 4.2.2 Verifying the OUT activity of HHARI RBR mutants via *in vitro* ubiquitination

The 4 mutants with highest frequency displayed in both fourth and fifth rounds of selection were cloned into pGEX vector to express the protein of interest in bacterial system. A GST-tagged protein purification was performed using the standard protocol as described in chapter 3. The purified HHARI RBR mutants were tested in a reaction with xUba1, xUbcH7 and xUB, respectively. After the 2 hr reaction, we ran SDS-PAGE and Western blot of the gel was probed with an anti-GST antibody. Strong signals of RBR domain polyubiquitination was observed, indicating that all 4 mutants were active (**Figure 4.4A-B**).



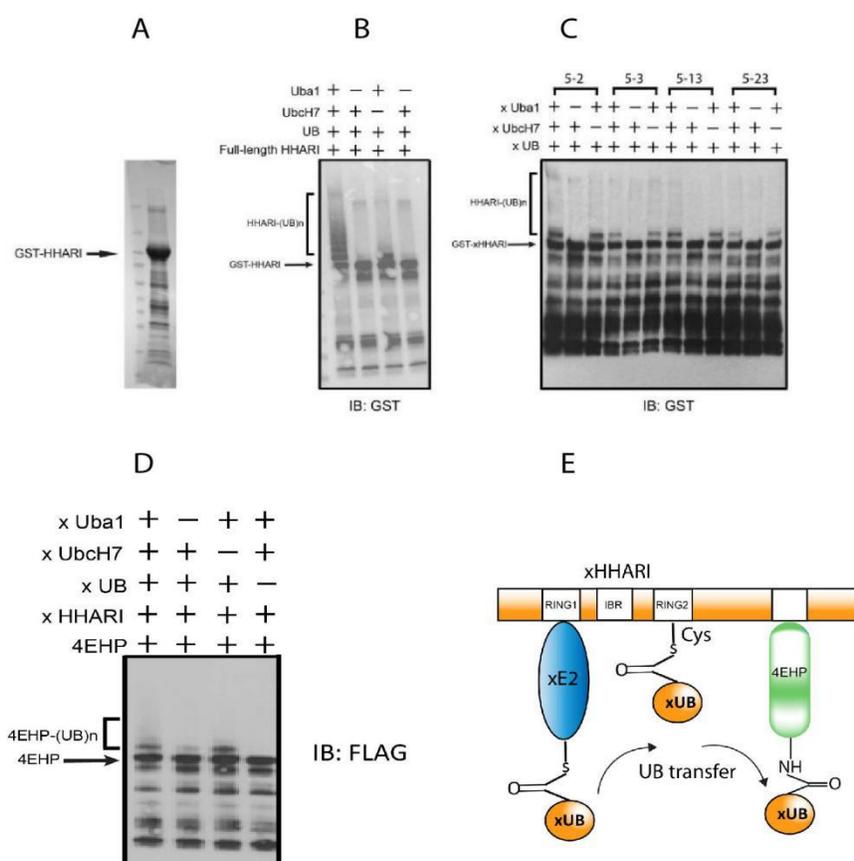
**Figure 4.4 Activity assays of HHARI RBR mutant variants. (A)** *In vitro* autoubiquitination assays of the HHARI mutants 5-2, 5-3, 5-13, 5-23, 5-32, 5-16 and 5-21. Autoubiquitination was probed against the GST epitope on the HHARI protein. **(B)** A scheme for autoubiquitination of HHARI RBR proteins.

#### 4.2.3 Testing of xHHARI full-length OUT activity *in vitro*

With the success in engineering the HHARI RBR mutants, we moved to introduce the clone 5-2's mutations into the full-length HHARI. The construction of pGEX-HHARI full-length was used as a template and the mutations were introduced by a combination of PCR (amplifying the fragment), digestion (cutting both vector and insert with restriction enzymes to release the sticky ends) and ligation (T4 ligase was employed to ligate the digested vector and insert). Finally, the pGEX-HHARI full-length 5-2 was completed, named as pGEX-xHHARI. Similarly, both full-length wild-type HHARI and xHHARI were expressed in BL21(DE3) strain and purified by GST tag, confirmed by coomassie blue staining (**Figure 4.5A**). The expressed wt HHARI was charged with wt Uba1, wt Ubch7 and wt UB, detected and confirmed by GST antibody, showing the full-length wt HHARI autoubiquitination was reconstituted (**Figure 4.5B**). By the same manner, the full-length xHHARI was mixed with xUba1, xUbch7 and xUB to initiate the self-ubiquitination

*in vitro*. By western blotting, the GST antibody was applied to probe the HHARI fused with a GST tag, as a result, the activity of xHHARI was verified (**Figure 4.5C**).

Afterwards, the mutants were tested for OUT ubiquitination on 4EHP, a known substrate of HHARI (**Figure 4.5 D-E**). The substrate 4EHP was monoubiquitinated by the full-length xHHARI via a detection of anti-FLAG antibody, suggesting the successful transfer of xUB through the xUba1-xUbcH7-xHHARI cascade. In summary, the xUB conjugation to the substrate through the engineered cascade was confirmed.



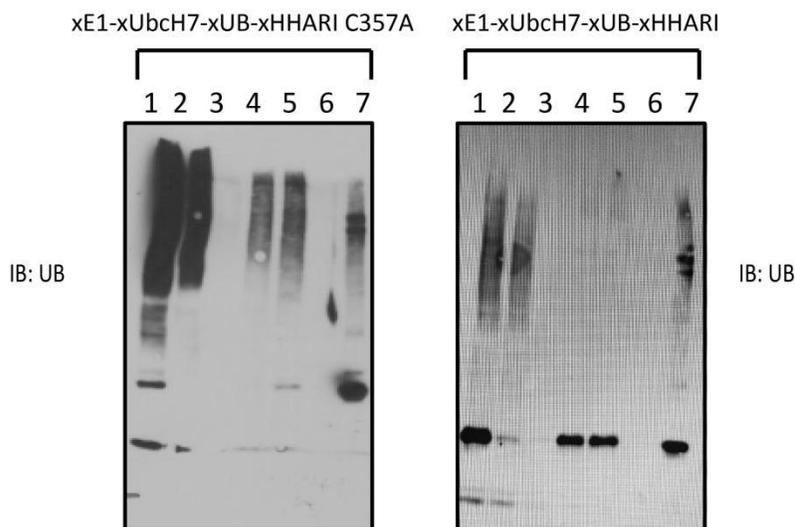
**Figure 4.5** *In vitro* ubiquitination assays of full-length HHARI mutants. (**A**) The expression of GST-tagged HHARI full-length was well expressed by coomassie blue staining (the major band was 89 kDa); (**B**) Autoubiquitination of GST-tagged wild-type HHARI full-length; (**C**) Autoubiquitination of GST-tagged HHARI mutants 5-2, 5-3, 5-13 and 5-23; (**D**) A known substrate, 4EHP ubiquitination of xHHARI mutant 5-2; (**E**) A scheme for *in vitro* substrate ubiquitination.

#### 4.2.4 Reconstitution of xHHARI OUT activity *in vivo*

The xHHARI OUT activity was confirmed in the same manner as xRNF38. Generally, the stable HEK293T cell lines that expressed each individual component of the OUT cascade was constructed except for xHHARI and HBT-xUB. Additionally, each individual component of OUT was fused with a unique epitope tag for detection. xHHARI was cloned into pLenti vector fused with a N-terminal Myc tag that could be used for protein signal detection. The pLenti plasmids of xHHARI and HBT-xUB were co-transfected into the cells with stable expression of xE1 and xE2, while the control cell line was co-transfected with xHHARI C357A and HBT-xUB. xHHARI C357A is a catalytically inactive mutant that is not able to catalyze the transfer of UB to the substrate proteins <sup>[173]</sup>. After being treated with 10  $\mu$ M MG132 for 4 hrs, the cells were lysed in RIPA buffer and cleared by sonication and centrifugation. After the removal of cell debris, we collected the supernatant and added Ni-NTA beads to bind to xUB-attached proteins, analyzed them by Western blotting. The antibodies for each epitope tag on xE1, xE2 and xE3 were probe against FLAG, V5 and Myc, respectively (**Figure 4.6 left panel**). The enzymes of the OUT cascade including xUba1, xUbcH7 and xHHARI were all expressed in the cell lysate. Interestingly, the xUB-conjugated proteins bound to the Ni-NTA beads were able to detect the xHHARI C357A mutant with lower signal compared to xHHARI pulled down by HBT-xUB (**Figure 4.6 right panel**). We think xHHARI C357A may still interacting with HBT-xUB via non-specific binding due to the complicated environment in cells. Further studies need to be carried out to identify the reason. Based on the pulldown results, we would expect to see the difference of substrate profiles between the cells expressing xUba1-xUbcH7-xHHARI-xUB and the control group that expressing xHHARI C357A with xUba1-xUbcH7-xUB.



between cells expressing the full HHARI OUT cascade and the control cells expressing xHHARI C357A. We performed affinity purification and proteomic screen three times (**Figure 4.7**).



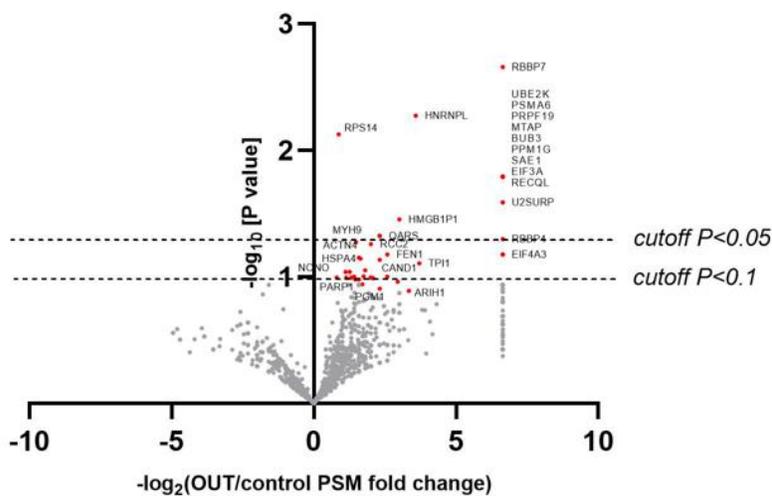
**Figure 4.7 Purification of xUB-modified proteins using affinity chromatography to determine the HHARI substrates.** Purification of xUB conjugated proteins from lysates of cells expressing the full OUT cascade of HHARI and the control with xHHARI C357A mutant expression (lanes 1-7). The Western blots of the gels were probed with an anti-UB antibody. Lane 1, cell lysate; lane 2, flow-through from the Ni-NTA column; lane 3, wash solution of the Ni-NTA column; lane 4, elution from the Ni-NTA column; lane 5, flow-through from the streptavidin column; lane 6, wash solution of the streptavidin column; lane 7, protein bound to the streptavidin beads after washing.

The volcano plot showed that 206 potential targets with  $P < 0.1$  were enriched from cells expressing the catalytic activity OUT cascade of HHARI (**Figure 4.8**). Thus, these proteins are likely the direct ubiquitination targets of HHARI with a high statistical significance. **Table 7** lists proteins having a PSM ratio of 2 or above in all three repetitions. The Ingenuity Pathway Analysis (IPA) of the HHARI substrates indicated that a protein network managing cancer progression and DNA cycle, recombination, and repair had the greatest significant network with possible HHARI

substrates including HNRNPL, CAND1, PPM1G, MMTAG2(C1orf35), ACTN4, ACTN1, PGM1, EIF3A, EIF4A3, and KPNA2. We carried out ubiquitination assays *in vitro* and *in vivo* to verify whether these proteins were targets of HHARI.

**Table 7 Potential substrates of HHARI**

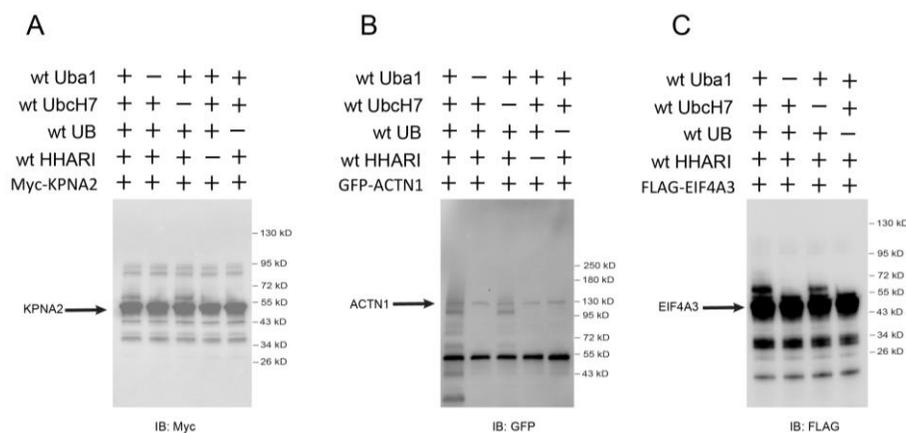
Protein	Gene	xHHARI C357A_1 PSMs	xHHARI C357A_2 PSMs	xHHARI C357A_3 PSMs	xHHARI_1 PSMs	xHHARI_2 PSMs	xHHARI_3 PSMs	Log2_ratio_1	Log2_ratio_2	Log2_ratio_3
Histone-binding protein RBBP7	RBBP7	0	0	0	2	2	3	6.644	6.644	6.644
Heterogeneous nuclear ribonucleoprotein L	HNRNPL	0	1	0	5	3	4	6.644	1.585	6.644
Pre-mRNA-processing factor 19	PRPF19	0	0	0	1	1	2	6.644	6.644	6.644
S-methyl-5-thioadenosine phosphorylase	MTAP	0	0	0	1	1	2	6.644	6.644	6.644
Mitotic checkpoint protein BUB3	BUB3	0	0	0	2	1	1	6.644	6.644	6.644
Protein phosphatase 1G	PPM1G	0	0	0	2	1	1	6.644	6.644	6.644
SUMO-activating enzyme subunit 1	SAE1	0	0	0	1	1	2	6.644	6.644	6.644
Eukaryotic translation initiation factor 3 subunit A	EIF3A	0	0	0	1	1	2	6.644	6.644	6.644
ATP-dependent DNA helicase Q1	RECQL	0	0	0	1	1	2	6.644	6.644	6.644
Multiple myeloma tumor-associated protein 2	MMTAG2	0	0	0	3	2	1	6.644	6.644	6.644
U2 snRNP-associated SURP motif-containing protein	UZSURP	0	0	0	3	1	2	6.644	6.644	6.644
Putative high mobility group protein B1-like 1	HMGB1P1	0	1	0	2	2	4	6.644	1.000	6.644
Heterogeneous nuclear ribonucleoprotein A0	HNRNPA0	0	1	0	2	2	1	6.644	1.000	6.644
Histone-binding protein RBBP4	RBBP4	0	0	0	4	1	5	6.644	6.644	6.644
Flap endonuclease 1	FEN1	0	1	0	3	1	2	6.644	0.000	6.644
Eukaryotic initiation factor 4A-III	EIF4A3	0	0	0	3	1	1	6.644	6.644	6.644
U5 small nuclear ribonucleoprotein 200 kDa helicase	SNRNP200	0	0	0	3	1	1	6.644	6.644	6.644
Poly(C)-binding protein 3	PCBP3	0	0	0	1	1	3	6.644	6.644	6.644
Triosephosphate isomerase	TPI1	1	0	0	6	1	6	2.595	6.644	6.644
Cullin-associated NEDD8-dissociated protein 1	CAND1	1	1	0	7	2	3	2.807	1.000	6.644
Vesicle-trafficking protein SEC22b	SEC22B	0	1	0	1	1	2	6.644	0.000	6.644
Programmed cell death 6-interacting protein	PDCD6IP	0	0	0	2	0	2	6.644	6.644	6.644
Ubiquitin thioesterase OTUB1	OTUB1	0	1	0	1	1	1	6.644	0.000	6.644
Vacuolar protein sorting-associated protein 35	VPS35	0	0	0	0	1	1	6.644	6.644	6.644
Importin subunit alpha-1	KPNA2	1	0	0	1	2	1	1.000	6.644	2.322
Alpha-actinin-4	ACTN4	2	1	0	4	2	3	1.000	1.000	6.644
Alpha-actinin-1	ACTN1	2	1	0	3	2	2	0.584	1.000	6.644



**Figure 4.8 Volcano plot for substrates profile of HHARI**

#### 4.2.5 Verifying substrates of HHARI by *in vitro* ubiquitination

At the time of writing the dissertation, I was still preparing the plasmids for the expression of the potential substrates of HHARI. I only have KPNA2 and EIF4A3 in hand and got both proteins expressed through *E.coli* system. For ACTN1 expression, I got the mammalian expression construct pEGFP-ACTN1 from Addgene, and I carried out the transfection of ACTN1 into HEK293T cells. After 48hr post-transfection, the GFP antibody was added to immunoprecipitate the ACTN1; lastly, the GFP antibody conjugated with ACTN1 was pulled down by protein A/G PLUS-Agarose for future assay. Similar to *in vitro* autoubiquitination, the expressed potential substrates KPNA2, ACTN1 and EIF4A3 were incubated with wt Uba1, wt UbcH7, wt HHARI and wt UB at 37°C respectively to facilitate the ubiquitination of the substrate by HHARI. The controls would be the reactions excluding E1, E2, E3 or UB. After 2 hrs, the reactions were quenched and analyzed by SDS-PAGE and Western blotting probed with either substrate-specific antibodies or an antibody against the epitope tag fused with the substrate. It's possible that *E.coli*-expressed substrates lack the necessary posttranslational modifications, such as phosphorylation, to enable E3 recognition, or that adaptor proteins required for UB transfer are lacking. To our delight, both KPNA2, ACTN1 and EIF4A3 were found to be mono-ubiquitinated by HHARI (**Figure 4.9**).



**Figure 4.9 OUT-identified HHARI substrates were tested for ubiquitination *in vitro*.** The putative substrate proteins produced from *E. coli* cells received wt UB through the wt Uba1-UbcH7-HHARI cascade. KPNA2(A), ACTN1(B), and EIF4A3(C) have all been shown to be ubiquitinated by HHARI.

### 4.3 Discussion and conclusion

Indeed, the construction of OUT is a long journey, which contains the engineering of three pairs of interfaces: xE1-xUB, xE1-xE2 and xE2-xE3. Here, we successfully constructed OUT cascades for RBR-type HHARI that interacts with xE1-xE2 and transfers the xUB to the substrates *in vitro*. The restoration of interaction between xUbcH7 and HHARI RBR domain convinced us to apply the OUT cascade in cells to identify the potential substrates that directly interact with HHARI. We then performed tandem purification of cellular proteins conjugated to HBT-xUB from cells expressing the HHARI OUT cascade and the catalytically inactive cascade. We then identified the potential substrates of HHARI via the assessment of the ratio of peptide spectrum matches (PSMs) between profiles of proteins identified by proteomics in the OUT and control cells.

As reported, there were only very few HHARI substrates in the literature including 4EHP<sup>[167]</sup>, PD-L1<sup>[60]</sup>(Programmed death-ligand 1), SEC31A<sup>[165]</sup>(Protein transport protein Sec31A) and NSF<sup>[174]</sup>(N-ethylmaleimide sensitive factor). We didn't find those proteins in our OUT screen; however, we did find some substrates in OUT screen share the homology with those proteins. For instance, the P value of EIF4A3(P=0.066<0.1) and EIF3A(P=0.016<0.1) were showing significant difference comparing the OUT cascade with the control without xHHARI expression, and both of them are in the same EIF (Eukaryotic initiation factor) family with 4EHP/EIF4E2. In addition, PPM1G was obviously enriched from HHARI OUT screen, and it has been implicated in a variety of biological activities, including transcriptional control, translational modulation, cell cycle

regulation, and alternative splicing. It is also reported that PPM1G is involved in NF- $\kappa$ B transcriptional activation via modulating the transcription elongation activity of 7SK snRNP and P-TEFb<sup>[175,176]</sup>. Recent study showed that the activation of PPM1G (Protein phosphatase 1G) would cause the abundant expression of PPM1G in HCC cells; the upregulation of PPM1G dephosphorylated the SRSF3, leading to the detachment of SRSF3 from RNA, thereby promoting cell proliferation and metastasis of the HCC cells<sup>[176]</sup>. The non-muscle  $\alpha$ -actinin isoforms ACTN1 and ACTN4 belong to the spectrin superfamily<sup>[177,178]</sup>. Both have many roles in non-muscle cells, including cellular motility and cell adhesion. For example, ACTN4 was overexpressed in gastric cancers leading to a poor prognosis<sup>[179]</sup>, moreover ACTN4 promotes the progression of osteosarcoma (OS) and enhances its invasive ability through the NF- $\kappa$ B Pathway<sup>[177]</sup>. ACTN1 studied by Blondelle et al<sup>[180]</sup> showed that it was regulated by Culling-3 (CUL3), a member of Cullin-RING E3 UB ligases (CRLs), for degradation via ubiquitination, thereby controlling the muscle development. The dysfunction of CUL3 leads to the accumulation of ACTN1 and ACTN4 in C2C12 myoblasts, promoting the nemaline myopathy<sup>[180]</sup>, a congenital neuromuscular disorder characterized by muscle weakness, fiber atrophy, and presence of nemaline bodies within myofibers<sup>[181,182]</sup>. HHARI was reported to be activated by CRL family via NEDD8 and transfer the first UB to CRL substrates<sup>[165]</sup>, thus we speculate that the ACTN1 and ACTN4 would be firstly monoubiquitinated by HHARI and then polyubiquitinated by CUL3 for proteolysis. C1orf35, also known as MMTAG2, was described as a candidate oncogene and the overexpression of MMTAG2 would promote carcinogenesis by binding to the i-motif of NHE III1 in the c-MYC promoter, thereby stimulating the transcription of c-MYC<sup>[183]</sup>. PGM1(Phosphoglucomutase 1) is an enzyme involved in glycogen metabolism. It reversibly converts glucose-1-phosphate to glucose-6-phosphate under physiological conditions<sup>[184]</sup>. It is also reported to promote progression in several

cancers, including lung cancer, gastric cancer and ovarian cancer, suggesting that it is correlated to cell viability, proliferation and metabolism<sup>[185]</sup>. KPNA2(Karyopherin  $\alpha$ 2) is an adaptor protein holding an important role in the transportation of proteins from the cytoplasm into the nucleus. It's also recognized as a potential biomarker in multiple forms of cancer and the expression of KPNA2 were associated with poor prognosis in patients with small hepatocellular carcinoma and osteosarcoma<sup>[186,187]</sup>. Our lab recently reported KPNA2 could be ubiquitinated by an OUT-based combination of E6AP E3 ligase and human papillomavirus E6; additionally, E6AP-E6 pair could significantly promote the degradation of KPNA1 in cells, leading to the inhibition of nuclear transport of phosphorylated STAT1 and the suppression of interferon- $\gamma$ -induced apoptosis in cervical cancer cells(CRC)<sup>[188]</sup>. Interestingly, we also found KPNA2 was enriched in HHARI OUT screen, and it was verified as a potential substrate of HHARI in vitro (**Figure 4.9**). To verify if it is genuinely ubiquitinated by HHARI, we need to carry out the *in vivo* ubiquitination, which will be completed after my graduation.

After going through the potential substrates in OUT screen, we may conclude that the HHARI is crucially involved in many cellular signaling processes, related to cancer progression, neuro diseases and DNA replication. Mapping the network of HHARI in the cell has been a challenge due to the complexity of its structure and function. Profiling the substrates of HHARI plays a key role on elucidating the role of HHARI in cells. Thus, the "Orthogonal Ubiquitin Transfer" (OUT) platform was established to identify the substrates of HHARI directly based on the catalytic transfer of UB.

## **4.4 Methods**

### **4.4.1 Generation of Lentiviral constructs**

To generate pLenti6-V5-D-TOPO-Asc1-hygromycin-HBT-(x)UB plasmids, HBT tag was sub-cloned from pQCXIP HBT-UB and fused with DNA fragments of human wt UB or xUB by PCR. The assembled DNA fragment was cloned into the pLenti6 plasmid with a hygromycin resistant gene. Genes of xUba1 and xUbcH7 were cloned into lentiviral vectors for the selection of stable cell lines. Flag-xUba1 gene was PCR amplified with primer WY15 and primer WY16 and cloned into pLenti6-V5-D-TOPO-Flag-Asc1-blasticidin vector between restriction sites EcoR I and Asc I. V5-xUbcH7 gene was PCR amplified from pET-xUbcH7 with PCR primers WY17 and WY18, digested with restriction enzymes AfeI and Nhe I, and cloned into pLenti4-V5-D backbone with a zeocin-resistance gene. The gene of xHHARI was PCR amplified with primers FZ08 and FZ09 and cloned into a pLenti vector with a Myc tag between restriction sites Nhe I and BssH II.

### **4.4.2 Construction of stable cell line**

Virus packaging, virus infection and selection of stable cell lines were performed according to the manufacturer's protocol for the ViraPower Lentiviral Expression System. Stable HEK293T cell lines expressing Flag-xUba1 and V5-xUbcH7 were selected with 10 µg/mL blasticidin and 100 µg/mL Zeocin, respectively. All recombinant DNA manipulations involving xE3 enzymes were accomplished in NEBStable cells (NEB), which possess greatly enhanced DNA stability and reduced possibility of recombination.

#### 4.4.3 Production of lentiviral particles

We generally followed the existing protocols according to the Virapower™ kit from Invitrogen. In 1.5 ml of Opti-MEM I medium without serum, 9 µg of Virapower™ Packaging mix was mixed with ~3 µg of the appropriate pLenti construct and added to a 5-minute preincubated solution of 36 µL of Lipofectamine™ 2000 reagent in 1.5 ml Opti-MEMR I without serum. The mixture was incubated for 20 minutes at room temperature to allow to formation of DNA-containing liposomes, and afterwards added dropwise to a plate of cultured 293T cells at 90-95% confluency in 5 ml of Opti-MEMR I. The cells were then allowed to incubate overnight at 37°C at 5% CO<sub>2</sub>. The next morning, the liposome-containing medium was discarded and replaced with 10 ml of fresh complete culture medium without antibiotics; the cells were then allowed to grow for a further 48-72 hours at 37°C at 5% CO<sub>2</sub>. Afterwards, the supernatant containing the lentiviral particles was harvested and cleared by filtration.

#### 4.4.4 Lentiviral Transduction

The lentiviral solution was diluted 2-fold in complete media and then used to for transfection by addition to plated non-confluent HEK293T cells, supplemented with 10 µg/ml Polybrene®; the cells were incubated overnight at 37°C at 5% CO<sub>2</sub>. The following day, the virus-containing medium was replaced with fresh complete medium, and the cells were allowed to grow under the same conditions for one more overnight. Afterwards (on the third day of transduction), the medium was once again replaced, this time with medium supplemented with the appropriate antibiotic (Depending on the construct, the concentrations were: blasticidin 10 µg/ml, hygromycin 200 µg/ml, zeocin 100 µg/ml, and puromycin 1 µg/ml); the antibiotic was maintained throughout the whole course of selection (lengths of selection were: blasticidin 5 days, hygromycin 7 days,

zeocin 10 days, and puromycin 3~5 days) with the medium being changed every 3~4 days. Antibiotic-resistant colonies were identified, picked, and expanded, being now regarded as stable cell lines. The expression of OUT proteins of interest were confirmed by Western immunoblotting against the appropriate epitope antibody-to this end the cells could directly be lysed by boiling in Laemmli buffer supplemented with DTT or BME.

#### **4.4.5 DNA Transient Transfection**

According to the manufacturer's manual (Horizon), for 75 cm<sup>2</sup> flask, seed cells in 10 mL of DMEM growth medium 24 hours prior to transfection at a density that will give a confluency of 70-90% for adherent cells on the day of transfection. Dilute DNA in 500 µL per well of serum-free DMEM or other serum-free growth medium. Gently mix DharmaFECT kb transfection reagent by pipetting. Dilute DharmaFECT kb reagent in 500 µL per well of serum-free DMEM or other serum-free growth medium. Add 500 µL diluted DharmaFECT kb reagent to the diluted DNA for a total of 1 mL per well. Incubate 20 minutes at room temperature. Gently aspirate medium from cell culture plate. Immediately dispense 1 mL of the transfection reagent/DNA mixture to flask. Gently rock the plate to achieve even distribution of the complexes. Incubate at 37°C in a CO<sub>2</sub> incubator. Begin to analyze transgene expression 24-72 hours later.

#### **4.4.6 *In vivo* Pulldown assays**

##### **4.4.6.1 Lysis of cells**

In general, HEK293T cells were lysed in RIPA buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.1 percent SDS, 0.5 percent sodium deoxycholate, and 1% Triton-X) supplemented with 1 mM PMSF. Alternatively, as was the case with HBT tandem purification of xUb-conjugated proteins, the cells are lysed under denaturing conditions using Buffer A (50 mM sodium phosphate pH 8.0,

300 mM NaCl, 8 M Urea, 0.5 percent NP-40) and 1 mM PMSF. Following lysis, lysates were cleaned by sonication and centrifugation at a speed of 10,000+ rpm for at least 10 minutes to eliminate cellular debris.

#### **4.4.6.2 Immunoprecipitation for target proteins**

Preclear whole cell lysate (optional step) as follows: add 0.25 g of the appropriate control IgG (corresponding to the primary antibody's host species) to approximately 1 mL of whole cell lysate, followed by 20  $\mu$ L of the appropriate suspended (25 percent v/v) agarose conjugate (Protein A/G PLUS-Agarose: sc-2003). Incubate for 30 minutes at 4°C. Centrifugation for pelleting beads at 3,000 rpm (about 1,000xg) for 30 seconds at 4°C. Supernatant (cell lysate) should be transferred to a fresh microcentrifuge tube at 4°C. Add 10 g of primary antibody agarose conjugate to 1 mL of the above cell lysate, or roughly 100-1000g of total cellular protein, and incubate at 4°C for 1 hour to overnight with mixing. Alternatively, if the primary antibody agarose conjugate is not available, incubate 1 ml cell lysate for 1-2 hours at 4° C with 1-10  $\mu$ L primary antibody. Add 20  $\mu$ L of agarose conjugate suspension (Protein A/G PLUS-Agarose: sc-2003). Cap tubes and incubate for 1 hour to overnight at 4°C on a rocker platform or rotating device. Centrifugation at 3,000 rpm (about 1,000xg) for 30 seconds at 4°C to collect pellet. Aspirate and discard supernatant carefully. The agarose was rinsed four times with PBS; samples for SDS-PAGE and Western blotting were prepared by boiling the agarose in Laemmli solution to liberate a fraction of the bound proteins.

#### **4.4.6.3 Tandem affinity purification of xUB-conjugated proteins**

To enrich the HBT-xUb ubiquitinated proteins, tandem purification of HBT-xUB-conjugated proteins was performed as previously described<sup>[142]</sup>. Briefly, ten 75cm<sup>2</sup>-flasks of

HEK293T cells stably expressing the xUba1-xUbcH7 were acutely co-transfected with xHHARI and HBT-xUB for 48 hours. To inhibit proteasome activity, cells were treated with 10  $\mu$ M MG132 for 4 hours at 37°C. Cells were then washed twice with ice-cold 1 $\times$  phosphate-buffered saline (PBS) (pH 7.4) and harvested by a cell scraper with buffer A [8 M urea, 300 mM NaCl, 50 mM tris, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 0.5% NP-40, 1 mM phenylmethylsulfonyl fluoride, and benzonase (125 U/ml) (pH 8.0)]. For Ni-NTA purification, cell lysates were centrifuged at 15,000g for 30 min at room temperature. Thirty-five microliters of Ni<sup>2+</sup> Sepharose beads (GE Healthcare) for each 1 mg of protein lysates was added to the clarified supernatant. After incubation overnight at room temperature in buffer A with 10 mM imidazole on a rocking platform, Ni<sup>2+</sup> Sepharose beads were pelleted by centrifugation at 100g for 3 min and washed sequentially with a 20-bead volume of buffer A (pH 8.0), buffer A (pH 6.3), and buffer A (pH 6.3) with 10 mM imidazole. After washing the beads, proteins were eluted twice with a 5-bead volume of buffer B [8 M urea, 200 mM NaCl, 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 2% SDS, 10 mM EDTA, 100 mM tris, and 250 mM imidazole (pH 4.3)]. For streptavidin purification, the elution solution was adjusted to pH 8.0. Fifty microliters of streptavidin Sepharose beads (Thermo Fisher Scientific) was added to the elution to bind HBT-xUB-conjugated proteins. After incubation on a rocking platform overnight at room temperature, streptavidin beads were pelleted and washed sequentially with 1.5 ml of buffer C [8 M urea, 200 mM NaCl, 2% SDS, and 100 mM tris (pH 8.0)], buffer D [8 M urea, 1.2 M NaCl, 0.2% SDS, 100 mM tris, 10% EtOH, and 10% isopropanol (pH 8.0)], and buffer E [8 M urea and 100 mM NH<sub>4</sub>HCO<sub>3</sub> (pH 8)].

#### **4.4.6.4 Sample digestion and LC-MS/MS analysis**

After washing, the streptavidin beads were spun down, residual urea was removed, and liquid chromatography coupled to tandem MS (LC-MS/MS) on an Orbitrap Fusion mass

spectrometer (Thermo Fisher Scientific) was performed at the Emory Integrated Proteomics Core (EIPC) according to previously published methods<sup>[156,157]</sup>. The spectra collected were compared to the human UniProt database using Proteome Discoverer 2.0. (90,300 target sequences). Fully tryptic restriction and a parent ion mass tolerance of  $\pm 20$  parts per million were used as search criteria. Methionine oxidation (+15.99492 Da), asparagine and glutamine deamidation (+0.98402 Da), lysine ubiquitination (+114.04293 Da), and protein N-terminal acetylation (+42.03670 Da) were all variable modifications (up to three per peptide were permitted); cysteine received a fixed carbamidomethyl modification (+57.021465 Da). The PSMs were filtered using a percolator to achieve a 1% false discovery rate.

#### **4.4.6.5 Bioinformatics analysis**

IPA software ([www.ingenuity.com](http://www.ingenuity.com)) was used to map and identify the biological networks and molecular pathways with a significant proportion of genes having HHARI ubiquitination targets. Fisher's exact test in IPA software was used to calculate the P values for pathways and networks. The level of statistical significance was set at  $P < 0.1$ . IPA was also used to visualize the identified biological networks.

#### **4.4.7 *In vitro* ubiquitination of substrate proteins**

Genes of potential substrate proteins KPNA2 and EIF4A3 were cloned into pET vector for protein expression. The plasmids were transformed into BL21 (DE3) cells to express the protein. pEGFP-ACTN1 was purchased from Addgene and was expressed in HEK293T cells. To assay ubiquitination by HHARI, 5 to 10  $\mu\text{M}$  substrate proteins were incubated with 1  $\mu\text{M}$  wt Uba1, 5  $\mu\text{M}$  wt UbcH7, and 50  $\mu\text{M}$  wt UB in TBS supplemented with 10 mM  $\text{MgCl}_2$  and 5 mM ATP. After 2-hour reaction at room temperature, the ubiquitination of substrates was analyzed by Western

blotting probed with either substrate-specific antibodies or antibodies against the tags fused to the substrates.

**Table 8 Primary antibody information**

<b>Antibody</b>	<b>Supplier</b>	<b>Catalog Number</b>	<b>Diluton</b>
$\beta$ -actin	Santa Cruz	sc-47778	1:1000
Ub	Santa Cruz	sc-8017	1:600
FLAG	Sigma	F4042	1:2000~1:4000
GST	Santa Cruz	sc-138	1:500~1:1000
KPNA2	Santa Cruz	sc-55538	1:1000
V5	Santa Cruz	sc-271944	1:500
Myc	Santa Cruz	sc-55475	1:600
ARIH1	Santa Cruz	sc-514551	1:500
GFP	Santa Cruz	sc-9996	1:500~1:1000

## 5 EXTEND SUBSTRATE STUDIES OF E6AP BY DESIGNING A CYCLIC $\Gamma$ -AA PEPTIDE

### 5.1 HECT-type E3 ligase E6AP

#### 5.1.1 Structure and biological function

E6AP, also known as UBE3A, is a 100 kDa cellular protein composed of 875 amino acids<sup>[189]</sup>. The E6 protein of HPV was reported to activate E6AP to enhance p53 degradation, which is a crucial step in the progression of HPV-driven carcinomas<sup>[189,190]</sup>. Also, we recently discovered that the stability of O-Linked N-Acetyl Glucosamine Transferase (OGT) was regulated by E6AP in the cell, suggesting that HPV may affect protein O-GlcNAcylation in the host cells by stimulating the UB ligase activity of E6AP<sup>[191]</sup>. The binding region for E6 is localized to the N-terminal of E6AP. Additionally, the last 84 amino acids of E6AP were required for p53 degradation<sup>[192]</sup>. The C-terminus of E6AP contains the HECT domain of 350 residues, a region shared by several E3 ligases structurally similar to E6AP<sup>[193]</sup>, and this domain is required for the ubiquitination function of E6AP. The HECT domain of E6AP is made up of two lobes that are joined by a three-residue hinge and pack loosely over a small interface<sup>[194]</sup>. Residue C843 in the HECT domain of E6AP is the catalytically active Cys for transferring UB through E6AP to its substrates<sup>[194]</sup>. Several studies showed that the HECT domain of E6AP are frequently mutated in AS (Angelman Syndrome) patients and the E6AP mutants are not able to form thioester bond with UB<sup>[195,196]</sup>, thereby the ubiquitination of the downstream targets of E6AP is affected. The deficiency of E6AP would induce the uncontrolled accumulation of protein(s) in the brain contributing to the disease pathogenesis of AS<sup>[197,198]</sup>. Thus, E6AP could be a target for AS treatment.

Recent work on repurposing E3s for induced protein degradation further fuels the effort for designing and screening E3 ligands for the assembly of bifunctional proteolysis-targeting chimeras

(PROTAC) or monovalent molecular glues to control protein stability in the cell<sup>[199,200]</sup>. E6 associated protein (E6AP) has been a prototypical E3 for probing the catalytic mechanism of the UB transfer reaction and the roles of E3s in cell regulation<sup>[201]</sup>. As introduced, E6AP was identified for its association with the E6 protein of the human papillomavirus (HPV) that would stir E6AP to ubiquitinate p53 for its degradation by the proteasome. This would allow the virus to subvert the antiviral response of the host cells and promote viral infection that eventually leads to tumorigenesis<sup>[193,202]</sup>. For counteracting E6AP that coalesce with HPV E6 to manifest its oncogenic activity, efforts have been devoted to the screening of cyclic peptides that would bind to E6AP and inhibit p53 ubiquitination<sup>[203]</sup> and small molecules and peptide ligands that would bind to E6 to prevent the formation of the E6-E6AP complex<sup>[204,205]</sup>. A variety of approaches has been developed to replenish E6AP activity in Angelman patients, including the use of a topoisomerase inhibitor, antisense oligonucleotides, and Cas9-mediated gene editing to activate the expression of UBE3A gene<sup>[206,207]</sup>. Furthermore, a recent screen yielded small molecules that can stimulate the UB ligase activity of E6AP as demonstrated by *in vitro* ubiquitination assays<sup>[208]</sup>.

### 5.1.2 Rational for designing the Cyclic $\gamma$ -AA peptide targeting E6AP

In this study, we screened a  $\gamma$ -AA peptide library for ligands binding to the HECT domain of E6AP.  $\gamma$ -AA peptides (**Figure 5.1A**), named for the oligomers of  $\gamma$ -substituted-N-acylated-N-aminoethyl amino acids, are derived from the backbone of chiral peptide nucleic acids (PNA). This new class of unnatural peptidomimetics possesses enormous chemical diversity, remarkable resistance to proteolytic degradation, and excellent capacities for cell delivery<sup>[209]</sup>. We developed an affinity-based screen for peptides binding to the HECT domain of E6AP and emerged from the screen are a series of  $\gamma$ -AA cyclic peptides with submicromolar affinity with the target HECT domain. One peptide ligand, enriched, known as P6, can significantly stimulate the activity of

E6AP in the ubiquitination of substrate proteins UbxD8, Rad23a, and  $\beta$ -catenin in reconstituted reactions. Interestingly, P6 increases the ubiquitination of E6AP substrates in the cell, allowing the proteasomal degradation. The discovery of the  $\gamma$ -AA peptide ligand for E6AP activation attests to the malleability of the non-conventional peptide scaffold for ligand discovery to target the protein ubiquitination cascade. Moreover, our results suggest a new therapeutic landscape based on the identification of E3-activating ligands and confirm the feasibility of using E3 activators alongside inhibitors and substrate recruiters for manipulating protein degradation pathways in the cell.

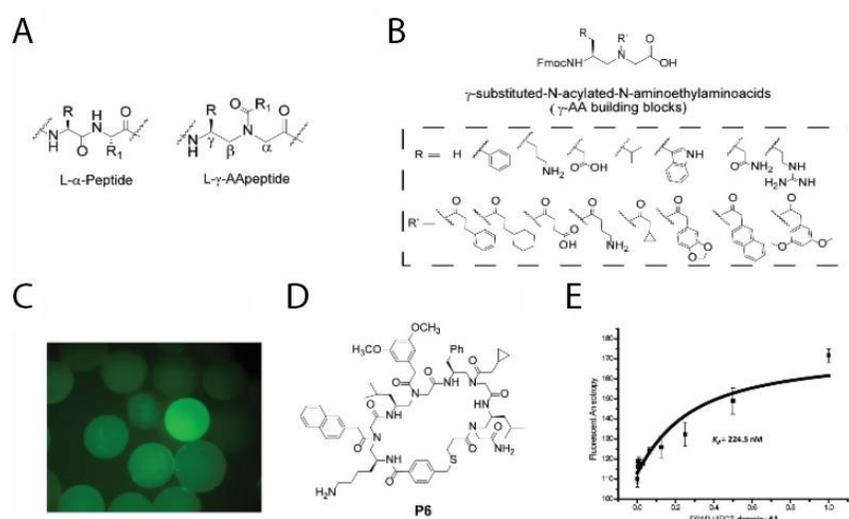
## 5.2 Results

### 5.2.1 Library screening and characterizing the affinities of the $\gamma$ -AA peptide ligands targeting E6AP HECT domain completed by Dr. Jianfeng Cai's group

E6AP runs the third relay of the E1-E2-E3 cascade that passes UB to the substrate proteins to form an isopeptide bond between the C-terminal carboxylate of UB and the Lys residues of the substrates. The HECT domain of E6AP engages the UB~E2 thioester conjugate to facilitate the delivery of UB from E2 to a catalytic Cys residue of the HECT domain before passing UB to the substrate proteins<sup>[210]</sup>. We recently engineered the HECT domain of E6AP for the assembly of an orthogonal UB transfer cascade (OUT) to profile the substrate specificity of the E3<sup>[30]</sup>. We posited that synthetic ligand binding to the HECT domain of E6AP might affect its catalytic activities so they could be further developed as inhibitors or activators of the E3. Toward such a goal, we designed a cyclic  $\gamma$ -AA peptide library comprising of a variety of  $\gamma$ -AA building blocks (**Figure 5.1A**). Each cyclic  $\gamma$ -AA peptide contained four  $\gamma$ -AA building blocks, and it matched the size of

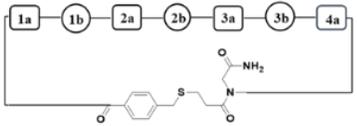
an 8-residue cyclic peptide. The choice of the macrocyclic ring size was based on our previous work that demonstrated high binding affinities of cyclic  $\gamma$ -AA peptides of four units with diverse biological targets<sup>[211]</sup>. In the current library design, both side chains of  $\gamma$ -AA building blocks (R and R') were selected from a pool of hydrophobic and charged groups, and a random combination of the building blocks would constitute a library of greater than  $2 \times 10^6$  in diversity (**Figure 5.1B**). A detailed protocol for the preparation and screening of the one-bead-two-compound (OBTC) library of the  $\gamma$ -AA peptides is included in the supplemental materials (**Supplementary Figure 5.6 and 5.7**). The library was screened based on the binding between the bead-anchored peptides with the HECT domain of E6AP with an N-terminal Flag tag that was recognized by an anti-Flag antibody labeled with AlexaFluor 488 that would emit green fluorescence. Five beads with strong green fluorescence were picked for their positive response to the binding of the E6AP HECT domain (**Figure 5.1C**), and the corresponding structures of the  $\gamma$ -AA peptides anchored on the beads were elucidated by the tandem MS/MS of MALDI (**Figure 5.1D**). We found two beads each yielded a single unambiguous structure, whereas the remaining three beads each was associated with two possible structures, leading to a total of eight  $\gamma$ -AA peptides as putative binders of the E6AP HECT domain. The peptide ligands from the screen were named P1-8 and they were resynthesized to measure their individual binding affinities with the HECT domain by fluorescence polarization (FP). P3-8 were found to have submicromolar binding affinities with the E6AP HECT with  $K_d$ 's ranging from 80 to 224.5 nM, while P1 and P2 did not show any measurable binding with the HECT domain (**Table 9 and Figure 5.1E**). Alignment of the selected peptide sequences revealed a preference for bulky hydrophobic residues such as naphthyl and 3,4-(methylenedioxy) phenyl at positions 1b, 2b and 3b, and small hydrophobic residues such as isopropyl and phenyl at positions 2a and 4a (**Table 9**). P4 showed a two-fold higher affinity than P3 with the HECT domain

while the main difference between the two peptides is that P4 has a phenyl residue at position 1a while P3 has a hydrogen atom at the corresponding position. Similarly, P7 with a phenyl residue at position 1a has a two-fold higher affinity than P8 of a similar structure but with a hydrogen atom at the same position. Insights from such comparison may improve the design of the  $\gamma$ -AA peptide ligands of the E6AP HECT domain in the future.



**Figure 5.1 Screening of  $\gamma$ -AA peptide ligands with binding affinities with the E6AP HECT domain.** (A) General structures of L- $\alpha$ -peptide and L- $\gamma$ -AA peptide. (B)  $\gamma$ -AA building blocks used for library preparation. (C) Beads bounded with AlexaFluor488-conjugated anti-Flag antibody emitted with green fluorescence. (D) The structure of P6 ligand from the affinity screen. (E) Fluorescence polarization assay for measuring the binding affinity of P6 with the HECT domain of E6AP. The P6 peptide was labeled with the FITC fluorophore and the  $K_d$  of the P6-HECT complex was determined to be 224.5 nM, Error bars=standard deviation from three independent experiments.

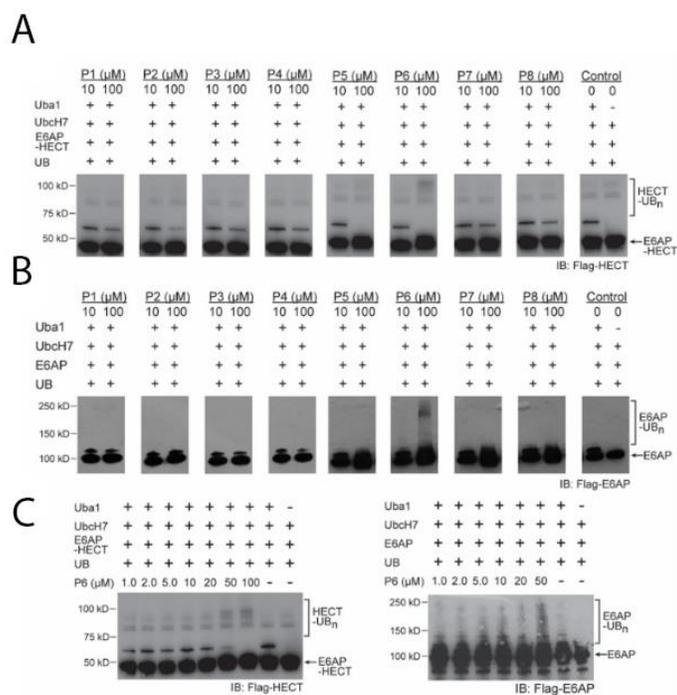
**Table 9 The alignment of the peptide sequences from the library screening and the binding affinities of the identified  $\gamma$ -AA peptide ligands P1-8 with the HECT domain of E6AP as measured by fluorescence polarization assay.** Positive-charged residues were shown in blue and negative-charged residues were shown in red. a and b denote the chiral and achiral side chains in a  $\gamma$ -AA building block, respectively.

								$K_d$ (nM)	
Peptide	Position	1a	1b	2a	2b	3a	3b		4a
P1		H				H			NA
P2						H			NA
P3		H							187.4
P4									82.2
P5									161.7
P6									224.5
P7						H			83.4
P8		H							171.1

### 5.2.2 Effects of $\gamma$ -AA peptides on E6AP activity assayed by autoubiquitination

This part was accomplished by my colleague, Ruochuan Liu. We first assayed the activities of the  $\gamma$ -AA peptides based on their effects on the self-ubiquitination of the HECT domain and full-length E6AP. We incubated E6AP HECT with 10  $\mu$ M and 100  $\mu$ M of each peptide from the affinity screen and added Uba1 (E1), UbcH7 (E2) and HA-tagged UB (HA-UB) to initiate the ubiquitination reaction. We found peptides P1-5 and 7-8 had little effect on the formation of UB-HECT conjugate while P6 peptide enhanced the formation of polyubiquitinated HECT species in the high-molecular weight range (**Figure 5.2A**). This suggests a stimulatory effect of P6 on the catalytic activity of the HECT domain. We then repeated the assay on the full-length E6AP with

the chosen condition that mainly generated mono-ubiquitinated species of the E3 without the addition of the peptides. Similar to the assay with the HECT domain, there was not much effect of P1-5 and 7-8 on autoubiquitination of full-length E6AP except that P3 and P4 showed weak inhibitory effect at 100  $\mu\text{M}$  concentration. P6 again stimulated E6AP self-ubiquitination, demonstrating its unique effect on E6AP activation (**Figure 5.2B**). We then assayed the self-ubiquitination of HECT and full-length E6AP in the presence of P6 peptide of varying concentrations and found the peptide can stimulate E3 self-ubiquitination at a concentration of 5  $\mu\text{M}$  (**Figure 5.2C**). Since HECT domain and the full-length E6AP showed a similar response to P6, it is likely that P6 activated E6AP by binding and stimulating the catalytic activity of the HECT domain. It is intriguing that P6 exhibited the weakest binding affinity (224.5 nM) for the E6AP HECT domain among P3-8 (**Figure 5.1E**), but it has a unique stimulatory effect on the UB transfer reaction catalyzed by the HECT domain. Alignment of the sequences of P6 and other peptides showed P6 has a distinctive positively charged Lys side chain at position 1a while there is a hydrogen (P3 and P8), phenyl (P4 and P7) or Asp sidechain (P5) at the same position (**Table 9**). The Lys side chain at 1a may contribute to the different binding mode of P6 with the HECT domain to activate the catalytic activity of E6AP.

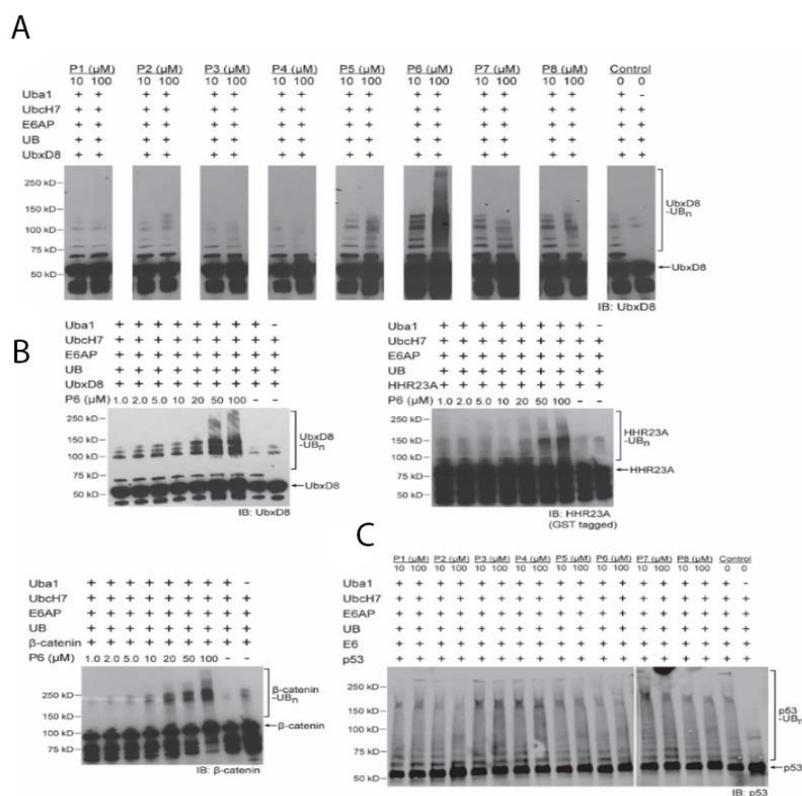


**Figure 5.2 Effect of  $\gamma$ -AA peptide ligands on the UB ligase activity of E6AP assayed by self-ubiquitination.** (A) Assaying the effects of ligands P1-8 on E6AP catalysis based on the self-ubiquitination of the HECT domain of E6AP. HECT domain self-ubiquitination was measured by Western blot probed with an anti-Flag antibody that binds to the N-terminal Flag tag fused to the HECT domain. The P6 ligand showed a stimulatory effect on the HECT domain activity. (B) Assaying the effect of P1-8 on E6AP activity based on the self-ubiquitination of the full-length E3. Ubiquitination of E6AP was followed by probing the N-terminal Flag tag fused to the E3 on the Western blot. (C) Dose-dependent activation of the HECT domain (left panel) and full-length E6AP (right panel) by the P6 ligand as measured by the self-ubiquitination reaction.

### 5.2.3 Effect of $\gamma$ -AA peptides on substrate ubiquitination catalyzed by E6AP

P6 stimulation of E6AP self-ubiquitination prompted us to assay the effect of P6 on the ubiquitination of E6AP substrates. We previously used an engineered-OUT cascade of E6AP to identify UbxD8, an adaptor protein regulating lipid droplet formation, and  $\beta$ -catenin, a transcription factor, as E6AP substrates. Other reports also verified  $\beta$ -catenin and HHR23A, a

protein involved in DNA repair, as E6AP substrates<sup>[212,213]</sup>. We thus measured the effect of the peptide ligands on E6AP-catalyzed ubiquitination of UbxD8, HHR23A, and  $\beta$ -catenin. We first assayed E6AP ubiquitination of UbxD8 in the presence of 10  $\mu$ M and 100  $\mu$ M peptides. P1-4 did not show much effect on UbxD8 ubiquitination compared to the control reactions with no addition of the peptides. In contrast, P6 showed a distinctive stimulatory effect on UbxD8 ubiquitination at both concentrations of the peptide, while P5, P7, and P8 showed a weaker stimulatory effect (**Figure 5.3A**). When the ubiquitination assay was performed with varying concentrations of the P6 peptide, all three substrates, including UbxD8, HHR23A, and  $\beta$ -catenin, showed enhanced ubiquitination in a dose-dependent manner in response to the amount of P6 in the reconstituted reaction (**Figure 5.3B**). These results suggest P6 would enhance the UB-transfer activity of E6AP to a broad range of substrates. The E6 protein from HPV virus can both enhance the activity of E6AP in substrate ubiquitination and stir its substrate specificity to new targets such as p53<sup>[214]</sup>. We thus assayed if the  $\gamma$ -AA peptides from the affinity screen would affect p53 ubiquitination catalyzed by E6AP. We set up the ubiquitination reaction with either 10  $\mu$ M or 100  $\mu$ M of the peptides and found none of the peptides, including P6, affected p53 ubiquitination with HPV E6 in the reconstituted reaction (**Figure 5.3C**). Such a result suggests HPV E6 may override the stimulatory effect of P6 on E6AP, so no additional enhancement of p53 ubiquitination by P6 is observed when both P6 peptide and HPV E6 protein were added to the reaction.

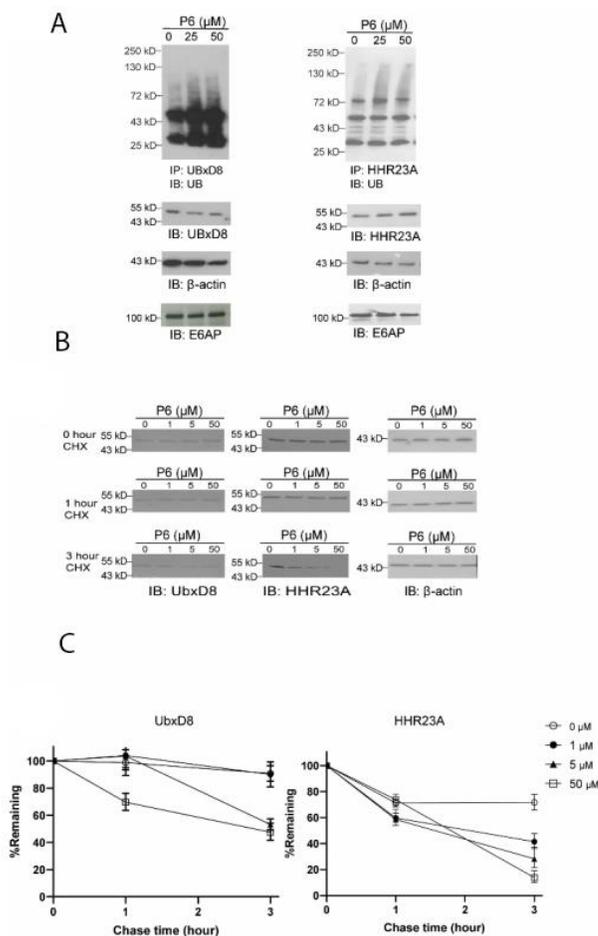


**Figure 5.3 Effect of the  $\gamma$ -AA peptide ligands on the UB ligase activity of E6AP assayed by substrate ubiquitination.** (A) Measuring the effects of ligands P1-8 on UbxD8 ubiquitination catalyzed by E6AP. The ubiquitination reaction contains 10 or 100  $\mu$ M of the ligands incubating with the Uba1-UbcH7-E6AP cascade and UbxD8 as the E6AP substrate. Ligand P6 showed the most significant stimulatory effect on UbxD8 ubiquitination catalyzed by E6AP. (B) Dose-dependent activation of E6AP-catalyzed substrate ubiquitination by the P6 ligand. Varying concentrations of P6 was added to the ubiquitination reaction containing the Uba1-UbcH7-E6AP cascade and the E6AP substrates UbxD8 (left panel), HHR23A (right panel) and  $\beta$ -catenin (bottom left panel). Ubiquitination of the substrates was measured by Western blots of the reaction probed with antibodies specific for each substrate. (C) Ligands P1-8 have no stimulatory effect on p53 ubiquitination catalyzed by E6AP pairing with the E6 protein of HPV. 10 or 100  $\mu$ M of the ligands were incubated with the E6AP enzymatic cascade, the viral E6 protein, and p53. The ubiquitination of p53 was followed by a Western blot of the reaction mixture probed with an anti-p53 antibody.

#### **5.2.4 P6-mediated enhancement of ubiquitination and accelerated degradation of E6AP substrates in HEK293 cells.**

We then assayed if P6 would affect the ubiquitination of E6AP substrates and their stabilities in the cell. We incubated HEK293 cells with 25  $\mu$ M and 50  $\mu$ M P6 peptide for 4 hours in the presence of MG132, a proteasome inhibitor to suppress the degradation of ubiquitinated proteins in the cell. We then lysed the cell, immunoprecipitated UbxD8 and HHR23A as E6AP substrates with specific antibodies and analyzed the ubiquitination levels of substrate proteins by Western blotting probed with an anti-UB antibody (**Figure 5.4A**). We found both UbxD8 and HHR23A showed an enhanced level of protein ubiquitination in cells with the addition of P6 compared to the control cells with no P6 added. This result suggests that the P6 peptide can enhance the ubiquitination of E6AP substrates in the cell.

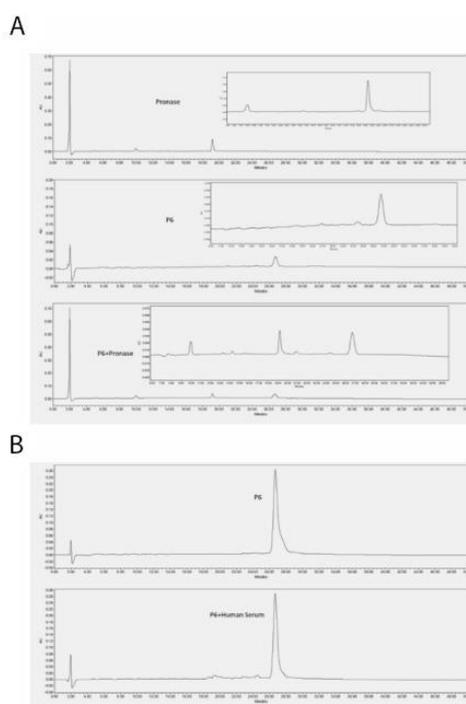
To measure if the enhanced E6AP activity due to P6 stimulation would accelerate the degradation of the substrate proteins in the cell, we carried out a cycloheximide (CHX) chase assay to follow the stability of UbxD8 and HHR23A in the presence of P6. We pretreated HEK293 cells with various concentrations of P6 to stimulate E6AP activity and then added CHX, a ribosome inhibitor to block the synthesis of new proteins. At various time points of CHX chase, we lysed the cells and probed the levels of UbxD8 and HHR23A to follow their degradation (**Figure 5.4B**). We found both substrates showed a faster degradation pattern in cells cultured with 50  $\mu$ M P6 than in cells with no P6 added. HHR23A also showed accelerated degradation in cells treated with 1 or 5  $\mu$ M of P6 (**Figure 5.4C**). These results confirm that P6 can promote the degradation of E6AP substrates by stimulating the catalytic activity of E6AP in cells.



**Figure 5.4 The stimulatory effect of the P6 ligand on E6AP-catalyzed substrate ubiquitin in the cell and acceleration of the degradation of E6AP substrates. (A) P6 enhanced the ubiquitination of E6AP substrates UbxD8 (left panel) and Rad23A (right panel) in HEK293 cells. P6 peptide of 0, 25 and 50  $\mu$ M was incubated with HEK293 cells for 14 h and the cells were treated with proteasome inhibitor MG132 for another 4 h. The cells were harvested for immunoprecipitating the substrate proteins with specific antibodies from the cell lysate, and the ubiquitination levels of the substrate proteins were measured by Western blots probed with an anti-UB antibody. (B) Cycloheximide (CHX) chase assay to measure the accelerated degradation of UbxD8 (left panels) and Rad23A (middle panels) in the presence of varying concentrations of P6.  $\beta$ -actin levels in the cell were measured in parallel as a loading control (right panels). Cells were incubated with varying concentrations of P6 peptide for 14 h and treated with ribosome inhibitor CHX for 0, 1, and 3 hours before harvesting to collect cell lysates. Levels of UbxD8, Rad23A and  $\beta$ -actin in the cell lysates were measured by Western blot probed with specific antibodies. (C) Levels of the E6AP substrates UbxD8 (left panel) and Rad23A (right panel) were plotted against the chase time after the cells were incubated with varying concentrations of the P6 peptide. P6 accelerated UbxD8 degradation at 5  $\mu$ M of the ligand and accelerated Rad23A degradation at 1  $\mu$ M of the ligand, Error bars=standard deviation from three independent experiments.**

### 5.2.5 Stability of P6

A distinctive characteristic of  $\gamma$ -AA peptides compared with canonical peptides is their remarkable resistance to protease. To assay the stability of P6 in presence of protease, we incubated it with 0.1 mg/mL pronase at 37 °C for 24 hrs. High-performance liquid chromatography and mass spectrometry of the P6 peptide before and after the exposure to pronase showed that P6 was resistant to protease cleavage (**Figure 5.5A**). We also incubated P6 in the human serum for 24 h and confirmed its stability (**Figure 5.5B**). The superior stability of P6 potentiates its use as a molecular probe for cell-based studies and therapeutic development.



**Figure 5.5 Analytic HPLC traces for peptide stability study.** (A) Analytic HPLC traces of P6 incubation with ammonium bicarbonate buffer or incubation with pronase (0.1 mg/mL) for 24 hours. Specifically, the top picture is the trace of pronase with a partial zoom from 8.5 min to 23.5 min, the middle picture is the trace of P6 after incubation with ammonium bicarbonate buffer and a partial zoom from 10 min to 34 min was included, and the bottom picture is the trace of P6 after incubation with pronase and a partial zoom from 6 min to 36 min was included. (B) The serum

stability of P6 was determined in 25% serum (v/v) at 37°C for 24 hours. Specifically, the top picture is the trace of P6 after incubation with water for 24h. the bottom picture is the trace of P6 after incubation with 25% serum for 24h.

### 5.3 Discussion and conclusion

$\gamma$ -AA peptides have exhibited promising potentials for biological application and drug discovery. Due to their propensity to form helical structures akin to  $\alpha$ -helices of proteins<sup>[215,216]</sup>,  $\gamma$ -AA peptides could mimic host-defense peptides<sup>[217,218]</sup> and modulate disease-related protein-protein interactions such as p53/MDM2,  $\beta$ -catenin/BCL9, and GLP-1/GLP-1R in vitro and in vivo<sup>[219,220]</sup>. Additionally, owing to the convenience for incorporating unnatural functionalities to their scaffolds with the modular synthesis protocol,  $\gamma$ -AA peptides are ideal for generating diverse libraries to screen for ligands of biological targets<sup>[221,222]</sup>. In this study, we carried out an affinity-based screen to identify  $\gamma$ -AA peptides that can bind to the HECT domain of E6AP with submicromolar binding affinity. Among the ligands we deduced from the library, one peptide, P6, stands out as a potent activator of E6AP – not only P6 can stimulate self-ubiquitination of E6AP and E6AP-catalyzed substrate ubiquitination in reconstituted reactions in vitro, but it can also enhance the ubiquitination of E6AP substrates in the cell and accelerate their degradation by the proteasome. Such an E3 ligand that can stimulate the UB ligase activity of an E3 is unique in that it can occupy a distinctive therapeutic landscape from the many E3 inhibitors that are being moved through the drug discovery pipeline<sup>[223]</sup>. Both an overactive E3 due to dysregulation or an underperforming E3 due to genetic mutations can be causative of diseases. The discovery of E3-stimulating ligands such as P6 for E6AP may boost the activity of mutated E3s in the cell to restore their normal functions. It would be interesting to assay if P6 and other peptide ligands from the screen may activate mutated E6AP that are causative for Angelman syndrome. Alternatively, the

screen of the  $\gamma$ -AA peptide library could be repeated with the mutated E6AP to identify ligands that can restore the activity of the E3 implicated in Angelman syndrome.

Previously phage-displayed libraries of UB variants (UbVs) have been selected for binding to the HECT domains of various E3s, and some of the UbVs from the selection were found to activate HECT E3s Nedd4 and Nedd4L<sup>[222]</sup>. Peptide and small molecule ligands of E3 would be better leads for drug development so N-methyl-cyclic peptides were identified for binding to the HECT domain of E6AP and inhibiting its UB ligase activity<sup>[203]</sup>. Yet the activity of the N-methyl-peptide in the cell was not characterized. In another report, small molecule ligands of flavin derivatives were found to activate E6AP to enhance its ubiquitination of substrate proteins<sup>[208]</sup>. Interestingly, the flavin-like ligands would induce a conformational change of E6AP close to the effect of HPV E6 on E6AP<sup>[208]</sup>. The effects of the flavin-like ligands on E6AP activity in the cell were not characterized. Still, it would be interesting to test if the P6  $\gamma$ -AA peptide would target the same binding site of the flavin ligand in E6AP and if P6 would induce a similar conformational change in E6AP to activate its ligase activity.

Our study also demonstrated the advantage of using the  $\gamma$ -AA peptide scaffold for developing ligands to affect E3 activities in the cell. The enzymatic cascades of UB transfer rely on sophisticated protein-protein interactions to deliver UB to the substrate proteins. Peptides and their structural mimics such as  $\gamma$ -AA peptides would have a better chance than small molecules to manipulate protein-protein interactions to inhibit or activate enzymes of the UB-transfer cascades. Indeed, recent success in designing stapled peptides to perturb E1-E2 interactions demonstrates the potential of peptide ligands as a privileged scaffold to target protein ubiquitinating enzymes<sup>[224]</sup>. The remarkable stability of the  $\gamma$ -AA peptides at physiological conditions would provide another advantage for developing them to activate or inhibit protein ubiquitination pathways.

## 5.4 Methods

### 5.4.1 Library preparation.

A detailed protocol was described in the supplementary information. Briefly, a split and pool method were used to prepare the  $\gamma$ -AA peptide-based OTBC library. Each bead was manipulated to have two layers: inner and outer layers. The  $\gamma$ -AA peptide was synthesized on the outer layer, in which the Fmoc protecting group of the  $\gamma$ -AA peptide building block was removed by 20% piperidine in DMF, and the exposed amino group reacted with the next  $\gamma$ -AA peptide building block using HOBt/DIC (6:6 equiv.) as the activation agents in DMF for 6 h. The Alloc protecting group in the  $\gamma$ -AA peptide building block was removed by 1% Pd (PPh<sub>3</sub>)<sub>4</sub> and 10% Me<sub>2</sub>NH·BH<sub>3</sub> in CH<sub>2</sub>Cl<sub>2</sub>, and the deprotected building block reacted with carboxylic acids in the presence of HOBt/DIC (6:6 equiv.) to introduce side chains. The decoding peptide was synthesized on the inner layer, in which the Dde group was removed using deprotection solution according to the previous report<sup>50</sup>. The Dde protected amino acids were coupled onto the solid phase in DMF for 4 h in the presence of PyBop (6 equiv.) and NEM (6 equiv.).

Library screening. The Tentagel beads were swelled in DMF for 1 h. After being washed with Tris buffer for three times, the beads were equilibrated in Tris buffer overnight at room temperature.

Prescreening. Firstly, the tantagel beads were incubated with blocking buffer (1% BSA in Tris buffer with a 1000 × excess of *E. coli* lysate) for 1h, then they were washed thoroughly with Tris buffer, and followed by the incubation with a AlexaFluor488-conjugated FLAG Epitope tag monoclonal antibody at a dilution of 1:500 for 2 h at room temperature. Then, the beads were washed with Tris buffer for three times and transferred into a 6-well plate to be observed under a

fluorescence microscope. The beads emitting green fluorescence were picked up and excluded from formal screening. The rest of the beads were pooled into a peptide vessel. After being washed with Tris buffer, the beads were treated with 8 M guanidine·HCl for 1 h to remove the bound protein at room temperature. Finally, the guanidine·HCl was washed away with water and Tris buffer. The beads were then incubated in DMF for 1 h and followed by washing and equilibration in Tris buffer overnight.

#### **5.4.2 Library Screening**

The beads were incubated with blocking buffer (1% BSA in Tris buffer with a 1000 × excess of *E. coli* lysate) for 1 h at room temperature. After being washed with Tris buffer for three times, the beads were incubated with Flag-tagged E6AP domain at a concentration of 50 nM for 4 h with 1% BSA in Tris buffer and 1000 × excess of *E. coli* lysate. After the thorough wash with Tris buffer, the library beads were incubated with 20 μL AlexaFluor488-conjugated FLAG Epitope tag monoclonal antibody in 10 mL Tris buffer at room temperature for 2 h. Then, the beads were washed with Tris buffer for three times and transferred into a 6-well plate to be observed under a fluorescence microscope. The beads emitting green fluorescence were picked up as putative hits.

#### **5.4.3 Cleavage and analysis.**

Each bead identified was transferred to a 1.5 mL microtube and denatured with 100 μL 8 M guanidine·HCl for 1 h at room temperature, respectively. Then, the bead was rinsed with Tris buffer, water, DMF, and ACN for three times in sequence. At last, the bead was placed in ACN overnight and then the ACN was evaporated. The bead was incubated in the solution of ACN: glacial acetic acid: H<sub>2</sub>O containing cyanogen bromide (CNBr) (v: v: v = 5:4:1) at a concentration of 50 mg/mL

overnight at room temperature. After the cleavage of the peptides from the bead, the solution was evaporated, and the cleaved peptide was dissolved in ACN: H<sub>2</sub>O (4:1) and decoded by MALDI/MS.

#### 5.4.4 FITC-labeled peptides preparation.

The Fmoc-Lys (Dde)-OH was first attached to the Rink amide resin. Then the Fmoc-protecting group was removed and followed by the coupling with desired building blocks. After cyclization, the Dde group was removed. FITC (2 equiv.) and DIPEA (10 equiv.) in DMF were added to the vessel and shaken overnight at room temperature. Then the FITC-labeled cyclic peptide was cleaved by 1:1 (v/v) DCM/TFA containing 2% triisopropylsilane. The crude was purified by the Waters HPLC, and the detailed structures can be found in the supplementary materials.

#### 5.4.5 Binding affinity.

The binding affinity (K<sub>d</sub>) of the peptides were measured by fluorescence polarization. Briefly, a constant amount of the 100 nM FITC-labeled cyclic peptide was incubated with a serial dilution of E6AP HECT domain. The K<sub>d</sub> values was calculated using the following equation, in which the L<sub>st</sub> and x refer to the concentration of peptide and protein, respectively.

$$y = [FP_{\min} + (FP_{\max} - FP_{\min})] \frac{(K_d + L_{st} + x) - \sqrt{(K_d + L_{st} + x)^2 - 4L_{st} \times x}}{2L_{st}}$$

#### **5.4.6 Construction of protein expression plasmids**

Uba1, UbcH7, HA tagged Ub and E6AP (HECT domain and full-length) were generated in the protein expression vector pET as described previously<sup>[142]</sup>. Also, the construction of substrate UbxD8, HHR23A and  $\beta$ -catenin were generated on the either pET or pGEX vector for protein expression in *E.coli*.

#### **5.4.7 *In vitro* assays to test the effect of peptide ligands on self-ubiquitination of E6AP and substrate ubiquitination by E6AP**

All assays were set up in 50  $\mu$ L reaction buffer supplemented with 50 mM Tris, 5 mM  $MgCl_2$ , 5 mM ATP and 1 mM DTT. 10  $\mu$ M or 100  $\mu$ M of each peptide was incubated with 0.5  $\mu$ M wt Uba1, 0.5  $\mu$ M wt UbcH7, 0.5  $\mu$ M wt N-terminal Flag-tagged E6AP HECT domain or full-length protein at 37°C for 30 min before 5  $\mu$ M wt UB was added to start the UB-transfer reaction. The reactions were incubated for another 2 h at 37°C and then quenched by boiling in the sample loading buffer of SDS-PAGE with DTT for 5 min and analyzed by SDS-PAGE and Western blot probed with anti-Flag antibody. In another experiment, P6 of gradient concentrations from 1  $\mu$ M to 100  $\mu$ M was added to the ubiquitination reaction mixture and the reactions proceeded similarly to assay substrate ubiquitination. For substrate ubiquitination, we use the same condition of *in vitro* autoubiquitination and 2  $\mu$ M substrates (UbxD8,  $\beta$ -catenin or Rad23A) were added into the reactions, respectively, analyzed by SDS-PAGE and Western blot probed with substrate-specific antibodies.

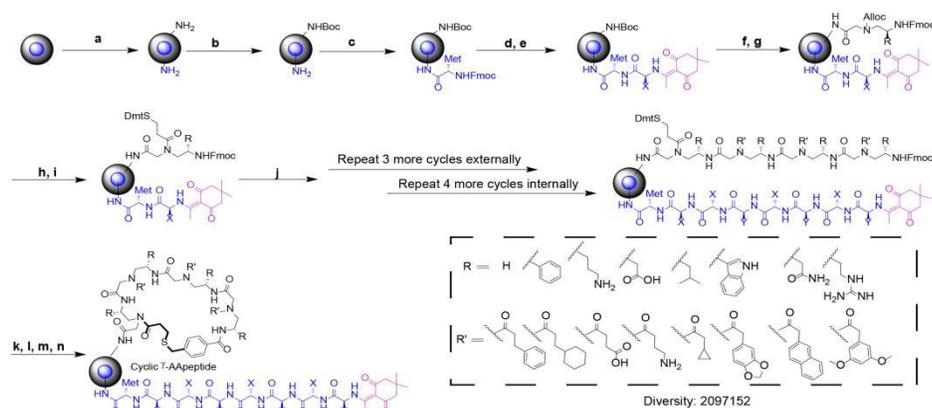
#### **5.4.8 Co-immunoprecipitation and to examine the effects of P6 peptide on UbxD8 and Rad23a**

HEK293T cells were treated with P6 peptide at 0 $\mu$ M, 25 $\mu$ M and 50 $\mu$ M for 14hrs. To immunoprecipitate substrate proteins, cells were treated with 0.5  $\mu$ M MG132 (American Peptide, Sunnyvale, CA) for extra 12hrs. Then, HEK293T cells (80-90% confluent monolayer in 75 cm<sup>2</sup> cell culture flask) with varying concentration of P6 peptide were washed twice with ice-cold PBS, pH 7.4. 1 mL ice-cold RIPA buffer was added to cell monolayer and incubated with cell for at 4 °C for 10 min. The cells were disrupted by repeated aspiration through a 21-gauge needle. The cell lysate was transferred to a 1.5 mL tube. The cell debris was pelleted by centrifugation at 13,000 r.p.m. for 20 min at 4 °C and the supernatant was transferred to a 1.5 mL centrifuge tube and precleared by adding 1.0  $\mu$ g of the appropriate control IgG (normal mouse or rabbit IgG corresponding to the host species of the primary antibody). 20  $\mu$ L of re-suspended volume of Protein A/G PLUS-agarose was added to the supernatant and incubation was continued for 30 min at 4 °C. From the cleared cell lysate, volume containing 2 mg total protein was transferred to a new tube. 30  $\mu$ L (i.e., 6  $\mu$ g) primary antibody was then added and incubation was continued for 1 hr at 4 °C. After incubation, 50  $\mu$ L of re-suspended volume of Protein A/G PLUS-Agarose was added. The tubes were capped and incubated at 4 °C on a rocking platform overnight. The agarose beads were pelleted by centrifugation at 350  $\times$  g for 5 min at 4 °C. The beads were then washed 3 times each time with 1.0 mL PBS. After the final wash, the beads were resuspended in 40 ml of 1 $\times$  Laemmli buffer with  $\beta$ -mercaptoethanol. The samples were boiled for 5 min and analyzed by SDS-PAGE and Western blotting probed with antibodies against UB.

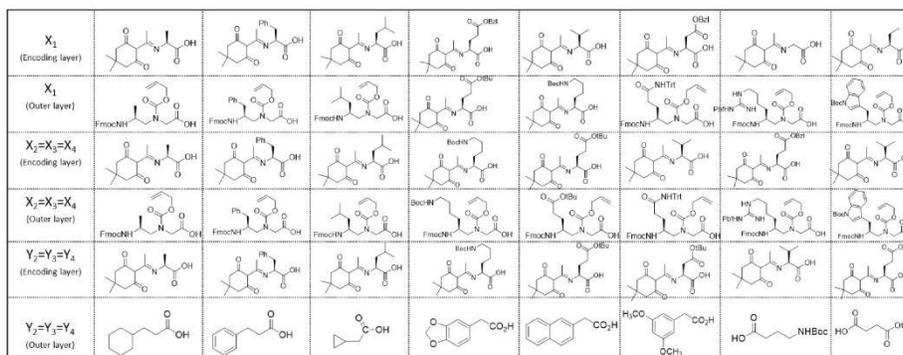
### 5.4.9 E6AP induced protein degradation in HEK293T cell

To examine the effects of P6 peptide on the levels of target proteins. Cycloheximide (CHX) chase assays were performed, HEK293T cells ( $5 \times 10^6$  cells) were treated with P6 peptide at 0uM, 1uM, 5uM and 50uM. After 14hrs, cells were treated with 100  $\mu\text{g}/\text{mL}$  CHX to block *de novo* protein synthesis and the cells were harvested after variable length of incubation time with CHX. The amount of substrate proteins in the cell were assayed by immunoblotting with antibodies against each substrate proteins. Protein levels were normalized to  $\beta$ -actin.

### 5.5 Supplementary materials

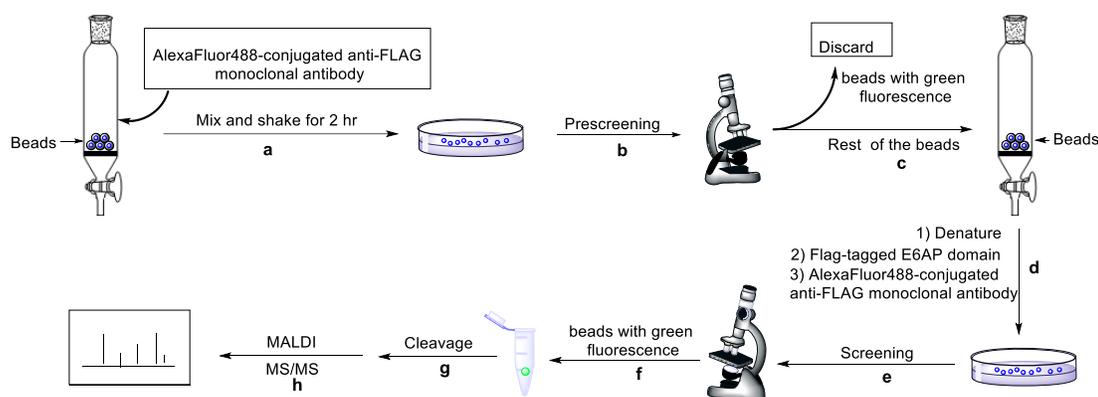


#### Decoding Map



**Figure 5.6 Procedure to synthesize the OBOC  $\gamma$ -AA peptide library.** a), Soak in water overnight, then washed with 1:1 (v/v) DCM/Et<sub>2</sub>O. b), (Boc)<sub>2</sub>O (0.5 equiv.) in 1:1 (v/v) DCM/Et<sub>2</sub>O for 3h, then washed with DCM and DMF. c), Fmoc-Met-OH (0.5 equiv.), HOBt (2 equiv.) and

DIC (2 equiv.) in DMF for 4h. d), 20% piperidine in DMF, then split the beads into five portions equally. e), Dde protected amino acids (2 equiv.) with PyBOP (6 equiv.) and NEM (6equiv.) in DMF were added to five peptide synthesis vessels, respectively. f), TFA/triisopropylsilane/H<sub>2</sub>O/Thioanisole (v:v:v:v 94:2:2:2). g), Fmoc protected  $\gamma$ -AApeptide/HOBt/DIC (3:6:6 equiv.) in DMF were added to five peptide synthesis vessels, respectively. for 4h. h), Pd (PPh<sub>3</sub>)<sub>4</sub> (1%) and Me<sub>2</sub>NH•BH<sub>3</sub> (10%) in DCM for 10 min. i), Dmt protected mercaptopropionic acid/HOBt/DIC (3:6:6 equiv) in DMF overnight. j), NH<sub>2</sub>OH•HCl and of imidazole in NMP, diluted with DCM. k), 20% piperidine in DMF. l), 4-(bromomethyl)benzoyl chloride (2 equiv.) and DIPEA (4 equiv.) in DCM for 1h. m), TFA/triisopropylsilane/DCM (2:2:96; v/v/v). n), (NH<sub>4</sub>)<sub>2</sub>CO<sub>3</sub> (10 equiv.) in 1:1 (v/v) DMF/H<sub>2</sub>O overnight.



**Figure 5.7 A schematic illustration of the screening method.** a), The beads were incubated with blocking buffer (1% BSA in Tris buffer with a 1000 × excess of *E. coli* lysate) for 1h, then washed with Tris buffer and incubated with the AlexaFluor488-conjugated FLAG Epitope tag monoclonal antibody for 2 h at room temperature. After being washed with Tris buffer, then the beads were transferred into a 6-well plate. b), Beads were observed under a fluorescence microscope. c), Beads emitting green fluorescence were picked up and discarded. The rest of the beads were pooled into a peptide vessel. d), The beads were treated with 8 M guanidine•HCl for 1 h to remove the bond protein. After wash and equilibration in Tris buffer, the beads were incubated with Flag-tagged E6AP domain at a concentration of 50 nM for 4 h with 1% BSA in Tris buffer and 1000 × excess of *E. coli* lysate. After the thorough wash with Tris buffer, the library beads were incubated with AlexaFluor488-conjugated FLAG Epitope tag monoclonal antibody at room temperature for 2 h. e), The beads were transferred into a 6-well plate to be observed under a fluorescence microscope. f), The beads emitting green fluorescence were picked up as putative hits. g), Cleavage the decoding sequence from the beads. h), The cleaved peptide was dissolved in ACN: H<sub>2</sub>O (4:1) and decoded by MALDI/MS.

## REFERENCES

- 1 Hunter, T. The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Molecular cell* **28**, 730-738 (2007).
- 2 Deng, L., Meng, T., Chen, L., Wei, W. & Wang, P. The role of ubiquitination in tumorigenesis and targeted drug discovery. *Signal transduction and targeted therapy* **5**, 1-28 (2020).
- 3 Hershko, A. & Ciechanover, A. The ubiquitin system. *Annual review of biochemistry* **67**, 425-479 (1998).
- 4 Wang, Y. *et al.* Regulation of the endocytosis and prion-chaperoning machineries by yeast E3 ubiquitin ligase Rsp5 as revealed by orthogonal ubiquitin transfer. *Cell Chemical Biology* (2021).
- 5 Nanduri, P., Hao, R., Fitzpatrick, T. & Yao, T.-P. Chaperone-mediated 26S proteasome remodeling facilitates free K63 ubiquitin chain production and aggresome clearance. *Journal of Biological Chemistry* **290**, 9455-9464 (2015).
- 6 Ciechanover, A. The unravelling of the ubiquitin system. *Nature reviews Molecular cell biology* **16**, 322-324 (2015).
- 7 O'Connor, H. F. & Huijbregtse, J. M. Enzyme–substrate relationships in the ubiquitin system: approaches for identifying substrates of ubiquitin ligases. *Cellular and Molecular Life Sciences* **74**, 3363-3375 (2017).
- 8 Berndsen, C. E. & Wolberger, C. New insights into ubiquitin E3 ligase mechanism. *Nature structural & molecular biology* **21**, 301-307 (2014).
- 9 Rajalingam, K. & Dikic, I. Expanding the ubiquitin code. *Cell* **164**, 1074-1074. e1071 (2016).
- 10 Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell research* **26**, 399-422 (2016).
- 11 Hochstrasser, M. Ubiquitin-dependent protein degradation. *Annual review of genetics* **30**, 405-439 (1996).
- 12 French, M. E., Koehler, C. F. & Hunter, T. Emerging functions of branched ubiquitin chains. *Cell Discovery* **7**, 1-10 (2021).
- 13 Michel, M. A., Swatek, K. N., Hospenthal, M. K. & Komander, D. Ubiquitin linkage-specific affimers reveal insights into K6-linked ubiquitin signaling. *Molecular cell* **68**, 233-246. e235 (2017).
- 14 Geisler, S. *et al.* PINK1/Parkin-mediated mitophagy is dependent on VDAC1 and p62/SQSTM1. *Nature cell biology* **12**, 119-131 (2010).
- 15 Akutsu, M., Dikic, I. & Bremm, A. Ubiquitin chain diversity at a glance. *Journal of cell science* **129**, 875-880 (2016).
- 16 Metzger, M. B. & Weissman, A. M. Working on a chain: E3s ganging up for ubiquitylation. *Nature cell biology* **12**, 1124-1126 (2010).
- 17 Hwang, C.-S., Shemorry, A., Auerbach, D. & Varshavsky, A. The N-end rule pathway is mediated by a complex of the RING-type Ubr1 and HECT-type Ufd4 ubiquitin ligases. *Nature cell biology* **12**, 1177-1185 (2010).
- 18 Ye, W. *et al.* TRIM8 negatively regulates TLR3/4-mediated innate immune response by blocking TRIF–TBK1 interaction. *The Journal of Immunology* **199**, 1856-1864 (2017).

- 19 Cohen, P. & Strickson, S. The role of hybrid ubiquitin chains in the MyD88 and other innate immune signalling pathways. *Cell Death & Differentiation* **24**, 1153-1159 (2017).
- 20 Emmerich, C. H. *et al.* Activation of the canonical IKK complex by K63/M1-linked hybrid ubiquitin chains. *Proceedings of the National Academy of Sciences* **110**, 15247-15252 (2013).
- 21 Paul, S. Dysfunction of the ubiquitin–proteasome system in multiple disease conditions: therapeutic approaches. *Bioessays* **30**, 1172-1184 (2008).
- 22 Bernassola, F., Karin, M., Ciechanover, A. & Melino, G. The HECT family of E3 ubiquitin ligases: multiple players in cancer development. *Cancer cell* **14**, 10-21 (2008).
- 23 Emanuele, M. J. *et al.* Global identification of modular cullin-RING ligase substrates. *Cell* **147**, 459-474 (2011).
- 24 Harper, J. W. & Tan, M.-K. M. Understanding cullin-RING E3 biology through proteomics-based substrate identification. *Molecular & Cellular Proteomics* **11**, 1541-1550 (2012).
- 25 Guo, Z. *et al.* Proteomics strategy to identify substrates of LNX, a PDZ domain-containing E3 ubiquitin ligase. *Journal of proteome research* **11**, 4847-4862 (2012).
- 26 Shimoji, T. *et al.* Identification of annexin A1 as a novel substrate for E6AP-mediated ubiquitylation. *Journal of cellular biochemistry* **106**, 1123-1135 (2009).
- 27 Mark, K. G., Loveless, T. B. & Toczyski, D. P. Isolation of ubiquitinated substrates by tandem affinity purification of E3 ligase–polyubiquitin-binding domain fusions (ligase traps). *Nature protocols* **11**, 291-301 (2016).
- 28 Mark, K. G., Simonetta, M., Maiolica, A., Seller, C. A. & Toczyski, D. P. Ubiquitin ligase trapping identifies an SCFSaf1 pathway targeting unprocessed vacuolar/lysosomal proteins. *Molecular cell* **53**, 148-161 (2014).
- 29 Bhuripanyo, K. *et al.* Identifying the substrate proteins of U-box E3s E4B and CHIP by orthogonal ubiquitin transfer. *Science advances* **4**, e1701393 (2018).
- 30 Wang, Y. *et al.* Identifying the ubiquitination targets of E6AP by orthogonal ubiquitin transfer. *Nature communications* **8**, 1-14 (2017).
- 31 Zhao, B. *et al.* Orthogonal ubiquitin transfer through engineered E1-E2 cascades for protein ubiquitination. *Chemistry & biology* **19**, 1265-1277 (2012).
- 32 Zhao, B. *et al.* Protein engineering in the ubiquitin system: tools for discovery and beyond. *Pharmacological reviews* **72**, 380-413 (2020).
- 33 Shah, K., Liu, Y., Deirmengian, C. & Shokat, K. M. Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. *Proceedings of the National Academy of Sciences* **94**, 3565-3570 (1997).
- 34 Liu, Y., Shah, K., Yang, F., Witucki, L. & Shokat, K. M. Engineering Src family protein kinases with unnatural nucleotide specificity. *Chemistry & biology* **5**, 91-101 (1998).
- 35 Winkler, G. S. & Timmers, H. T. M. Structure-based approaches to create new E2–E3 enzyme pairs. *Methods in enzymology* **399**, 355-366 (2005).
- 36 Bhuripanyo, K. *Identification of ubiquitin ligase substrates via orthogonal ubiquitin transfer*, The University of Chicago, (2016).
- 37 Lee, I. & Schindelin, H. Structural insights into E1-catalyzed ubiquitin activation and transfer to conjugating enzymes. *Cell* **134**, 268-278 (2008).
- 38 Zheng, N. & Shabek, N. Ubiquitin ligases: structure, function, and regulation. *Annual review of biochemistry* **86**, 129-157 (2017).

- 39 Wang, D., Ma, L., Wang, B., Liu, J. & Wei, W. E3 ubiquitin ligases in cancer and implications for therapies. *Cancer and Metastasis Reviews* **36**, 683-702 (2017).
- 40 Buetow, L. & Huang, D. T. Structural insights into the catalysis and regulation of E3 ubiquitin ligases. *Nature reviews Molecular cell biology* **17**, 626-642 (2016).
- 41 Pickart, C. M. Mechanisms underlying ubiquitination. *Annual review of biochemistry* **70** (2001).
- 42 Kamadurai, H. B. *et al.* Insights into ubiquitin transfer cascades from a structure of a UbcH5B~ ubiquitin-HECTNEDD4L complex. *Molecular cell* **36**, 1095-1102 (2009).
- 43 Morreale, F. E. & Walden, H. Types of ubiquitin ligases. *Cell* **165**, 248-248. e241 (2016).
- 44 Rotin, D. & Kumar, S. Physiological functions of the HECT family of ubiquitin ligases. *Nature reviews Molecular cell biology* **10**, 398-409 (2009).
- 45 Sánchez-Tena, S., Cubillos-Rojas, M., Schneider, T. & Rosa, J. L. Functional and pathological relevance of HERC family proteins: a decade later. *Cellular and Molecular Life Sciences* **73**, 1955-1968 (2016).
- 46 Huang, L. *et al.* Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade. *Science* **286**, 1321-1326 (1999).
- 47 Weber, J., Polo, S. & Maspero, E. HECT E3 ligases: a tale with multiple facets. *Frontiers in physiology* **10**, 370 (2019).
- 48 Kim, H. C. & Huibregtse, J. M. Polyubiquitination by HECT E3s and the determinants of chain type specificity. *Molecular and cellular biology* **29**, 3307-3318 (2009).
- 49 French, M. E., Kretzmann, B. R. & Hicke, L. Regulation of the RSP5 ubiquitin ligase by an intrinsic ubiquitin-binding site. *Journal of Biological Chemistry* **284**, 12071-12079 (2009).
- 50 Huibregtse, J. M., Scheffner, M., Beaudenon, S. & Howley, P. M. A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. *Proceedings of the National Academy of Sciences* **92**, 2563-2567 (1995).
- 51 Jäckl, M. *et al.*  $\beta$ -Sheet augmentation is a conserved mechanism of priming HECT E3 ligases for ubiquitin ligation. *Journal of molecular biology* **430**, 3218-3233 (2018).
- 52 Aguilera, M., Oliveros, M., Martínez-Padrón, M., Barbas, J. A. & Ferrús, A. Ariadne-1: a vital Drosophila gene is required in development and defines a new conserved family of ring-finger proteins. *Genetics* **155**, 1231-1244 (2000).
- 53 Wenzel, D. M., Lissounov, A., Brzovic, P. S. & Klevit, R. E. UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids. *Nature* **474**, 105-108 (2011).
- 54 Duda, D. M. *et al.* Structure of HHARI, a RING-IBR-RING ubiquitin ligase: autoinhibition of an Ariadne-family E3 and insights into ligation mechanism. *Structure* **21**, 1030-1041 (2013).
- 55 Smit, J. J. & Sixma, T. K. RBR E3-ligases at work. *EMBO reports* **15**, 142-154 (2014).
- 56 Wauer, T. & Komander, D. Structure of the human Parkin ligase domain in an autoinhibited state. *The EMBO journal* **32**, 2099-2112 (2013).
- 57 Narendra, D. P. *et al.* PINK1 is selectively stabilized on impaired mitochondria to activate Parkin. *PLoS biology* **8**, e1000298 (2010).
- 58 Trempe, J.-F. *et al.* Structure of parkin reveals mechanisms for ubiquitin ligase activation. *Science* **340**, 1451-1455 (2013).
- 59 Villa, E. *et al.* Parkin-independent mitophagy controls chemotherapeutic response in cancer cells. *Cell reports* **20**, 2846-2859 (2017).

- 60 Wu, Y. *et al.* ARIH1 signaling promotes anti-tumor immunity by targeting PD-L1 for proteasomal degradation. *Nature communications* **12**, 1-14 (2021).
- 61 Budhidarmo, R., Nakatani, Y. & Day, C. L. RINGs hold the key to ubiquitin transfer. *Trends in biochemical sciences* **37**, 58-65 (2012).
- 62 Borden, K. L. & Freemont, P. S. The RING finger domain: a recent example of a sequence—structure family. *Current opinion in structural biology* **6**, 395-401 (1996).
- 63 Deshaies, R. J. & Joazeiro, C. A. RING domain E3 ubiquitin ligases. *Annual review of biochemistry* **78** (2009).
- 64 Xie, J., Jin, Y. & Wang, G. The role of SCF ubiquitin-ligase complex at the beginning of life. *Reproductive Biology and Endocrinology* **17**, 1-9 (2019).
- 65 Schulman, B. A. *et al.* Insights into SCF ubiquitin ligases from the structure of the Skp1–Skp2 complex. *Nature* **408**, 381-386 (2000).
- 66 Zheng, N. *et al.* Structure of the Cul1–Rbx1–Skp1–F box Skp2 SCF ubiquitin ligase complex. *Nature* **416**, 703-709 (2002).
- 67 Schreiber, A. *et al.* Structural basis for the subunit assembly of the anaphase-promoting complex. *Nature* **470**, 227-232 (2011).
- 68 Metzger, M. B., Pruneda, J. N., Klevit, R. E. & Weissman, A. M. RING-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1843**, 47-60 (2014).
- 69 Satyanarayana, A. & Kaldis, P. Mammalian cell-cycle regulation: several Cdks, numerous cyclins and diverse compensatory mechanisms. *Oncogene* **28**, 2925-2939 (2009).
- 70 Olzmann, J. A., Richter, C. M. & Kopito, R. R. Spatial regulation of UBXD8 and p97/VCP controls ATGL-mediated lipid droplet turnover. *Proceedings of the National Academy of Sciences* **110**, 1345-1350 (2013).
- 71 Gargiulo, S. & Soumillion, P. Directed evolution for enzyme development in biocatalysis. *Current opinion in chemical biology* **61**, 107-113 (2021).
- 72 Wang, Y. *et al.* Directed evolution: Methodologies and applications. *Chemical reviews* (2021).
- 73 Dalby, P. A. Strategy and success for the directed evolution of enzymes. *Current opinion in structural biology* **21**, 473-480 (2011).
- 74 Arnold, F. H. Directed evolution: creating biocatalysts for the future. *Chemical engineering science* **51**, 5091-5102 (1996).
- 75 Leemhuis, H., Stein, V., Griffiths, A. D. & Hollfelder, F. New genotype–phenotype linkages for directed evolution of functional proteins. *Current opinion in structural biology* **15**, 472-478 (2005).
- 76 Braman, J. *In vitro mutagenesis protocols*. Vol. 634 (Springer, 2010).
- 77 Pritchard, L., Corne, D., Kell, D., Rowland, J. & Winson, M. A general model of error-prone PCR. *Journal of theoretical biology* **234**, 497-509 (2005).
- 78 Arnold, F. H. & Georgiou, G. Directed evolution library creation. *Methods in molecular biology* **231**, 231 (2003).
- 79 Vanhercke, T., Ampe, C., Tirry, L. & Denolf, P. Reducing mutational bias in random protein libraries. *Analytical biochemistry* **339**, 9-14 (2005).
- 80 Labrou, N. E. Random mutagenesis methods for in vitro directed enzyme evolution. *Current Protein and Peptide Science* **11**, 91-100 (2010).

- 81 Arnold, F. H. Directed evolution: bringing new chemistry to life. *Angewandte Chemie International Edition* **57**, 4143-4148 (2018).
- 82 Chen, I., Dorr, B. M. & Liu, D. R. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proceedings of the National Academy of Sciences* **108**, 11399-11404 (2011).
- 83 Cherf, G. M. & Cochran, J. R. Applications of yeast surface display for protein engineering. *yeast surface display*, 155-175 (2015).
- 84 Scott, J. K. & Smith, G. P. Searching for peptide ligands with an epitope library. *Science* **249**, 386-390 (1990).
- 85 Smith, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315-1317 (1985).
- 86 Wu, C.-H., Liu, I.-J., Lu, R.-M. & Wu, H.-C. Advancement and applications of peptide phage display technology in biomedical science. *Journal of biomedical science* **23**, 1-14 (2016).
- 87 Hetrick, K. J., Walker, M. C. & Van Der Donk, W. A. Development and application of yeast and phage display of diverse lanthipeptides. *ACS central science* **4**, 458-467 (2018).
- 88 Smeal, S. W., Schmitt, M. A., Pereira, R. R., Prasad, A. & Fisk, J. D. Simulation of the M13 life cycle I: Assembly of a genetically-structured deterministic chemical kinetic simulation. *Virology* **500**, 259-274 (2017).
- 89 Rakonjac, J., Das, B. & Derda, R. Filamentous Bacteriophage in Bio/Nano/Technology, Bacterial Pathogenesis and Ecology. *Frontiers in microbiology* **7**, 2109 (2016).
- 90 Roux, S. *et al.* Cryptic inoviruses revealed as pervasive in bacteria and archaea across Earth's biomes. *Nature microbiology* **4**, 1895-1906 (2019).
- 91 Dubuisson, J. & Cosset, F.-L. Virology and cell biology of the hepatitis C virus life cycle—An update. *Journal of hepatology* **61**, S3-S13 (2014).
- 92 Petrenko, V. & Smith, G. Vectors and Modes of Display in Phage Display in Biotechnology and Drug Discovery. *Bo Raton, FL, USA, CRC Pressw, Taylor & Francis Group*, 714 (2005).
- 93 van Wezenbeek, P. M., Hulsebos, T. J. & Schoenmakers, J. G. Nucleotide sequence of the filamentous bacteriophage M13 DNA genome: comparison with phage fd. *Gene* **11**, 129-148 (1980).
- 94 Weiss, G. A. & Sidhu, S. S. Design and evolution of artificial M13 coat proteins. *Journal of molecular biology* **300**, 213-219 (2000).
- 95 Løset, G. Å., Roos, N., Bogen, B. & Sandlie, I. Expanding the versatility of phage display II: improved affinity selection of folded domains on protein VII and IX of the filamentous phage. *PLoS One* **6**, e17433 (2011).
- 96 Albericio, F. & Kruger, H. G. Therapeutic peptides. *Future medicinal chemistry* **4**, 1527-1531 (2012).
- 97 Minard, P., Urvoas, A. & Soumillion, P. Protein Engineering by Phage display. (2008).
- 98 Smith, G. P. Phage display: Simple evolution in a petri dish (Nobel lecture). *Angewandte Chemie International Edition* **58**, 14428-14437 (2019).
- 99 Nafisi, P. M., Aksel, T. & Douglas, S. M. Construction of a novel phagemid to produce custom DNA origami scaffolds. *Synthetic Biology* **3**, ysy015 (2018).
- 100 Qi, H., Lu, H., Qiu, H.-J., Petrenko, V. & Liu, A. Phagemid vectors for phage display: properties, characteristics and construction. *Journal of molecular biology* **417**, 129-143 (2012).

- 101 Frei, J. & Lai, J. Protein and antibody engineering by phage display. *Methods in*  
*enzymology* **580**, 45-87 (2016).
- 102 Felici, F., Luzzago, A., Folgori, A. & Cortese, R. Mimicking of discontinuous epitopes  
by phage-displayed peptides, II. Selection of clones recognized by a protective  
monoclonal antibody against the Bordetella pertussis toxin from phage peptide libraries.  
*Gene* **128**, 21-27 (1993).
- 103 Sidhu, S. S. Phage display in pharmaceutical biotechnology. *Current opinion in*  
*biotechnology* **11**, 610-616 (2000).
- 104 Soltes, G. *et al.* A new helper phage and phagemid vector system improves viral display  
of antibody Fab fragments and avoids propagation of insert-less virions. *Journal of*  
*immunological methods* **274**, 233-244 (2003).
- 105 Barbas, C. F., Kang, A. S., Lerner, R. A. & Benkovic, S. J. Assembly of combinatorial  
antibody libraries on phage surfaces: the gene III site. *Proceedings of the National*  
*Academy of Sciences* **88**, 7978-7982 (1991).
- 106 Buetow, L. *et al.* Activation of a primed RING E3-E2-ubiquitin complex by non-  
covalent ubiquitin. *Molecular cell* **58**, 297-310 (2015).
- 107 Kay, B. K., Winter, J. & McCafferty, J. *Phage display of peptides and proteins: a*  
*laboratory manual*. (Elsevier, 1996).
- 108 Linke, K. *et al.* Structure of the MDM2/MDMX RING domain heterodimer reveals  
dimerization is required for their ubiquitylation in trans. *Cell death & differentiation* **15**,  
841-848 (2008).
- 109 Liew, C. W., Sun, H., Hunter, T. & Day, C. L. RING domain dimerization is essential for  
RNF4 function. *The Biochemical journal* **431**, 23 (2010).
- 110 Plechanovová, A. *et al.* Mechanism of ubiquitylation by dimeric RING ligase RNF4.  
*Nature structural & molecular biology* **18**, 1052-1059 (2011).
- 111 Yin, Q. *et al.* E2 interaction and dimerization in the crystal structure of TRAF6. *Nature*  
*structural & molecular biology* **16**, 658-666 (2009).
- 112 Deng, L. *et al.* Activation of the I $\kappa$ B kinase complex by TRAF6 requires a dimeric  
ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. *Cell* **103**, 351-  
361 (2000).
- 113 Halaby, M.-j., Hakem, R. & Hakem, A. Pirh2: an E3 ligase with central roles in the  
regulation of cell cycle, DNA damage response, and differentiation. *Cell cycle* **12**, 2733-  
2737 (2013).
- 114 Fang, D. *et al.* Cbl-b, a RING-type E3 ubiquitin ligase, targets phosphatidylinositol 3-  
kinase for ubiquitination in T cells. *Journal of Biological Chemistry* **276**, 4872-4878  
(2001).
- 115 Qingjun, L., Zhou, H., Langdon, W. & Zhang, J. E3 ubiquitin ligase Cbl-b in innate and  
adaptive immunity. *Cell cycle* **13**, 1875-1884 (2014).
- 116 Matalon, O. & Barda-Saad, M. Cbl ubiquitin ligases mediate the inhibition of natural  
killer cell activity. *Communicative & integrative biology* **9**, e1216739 (2016).
- 117 Lu, T. *et al.* Cbl-B is upregulated and plays a negative role in activated human NK cells.  
*The Journal of Immunology* **206**, 677-685 (2021).
- 118 Paolino, M. *et al.* The E3 ligase Cbl-b and TAM receptors regulate cancer metastasis via  
natural killer cells. *Nature* **507**, 508-512 (2014).
- 119 Naramura, M. *et al.* c-Cbl and Cbl-b regulate T cell responsiveness by promoting ligand-  
induced TCR down-modulation. *Nature immunology* **3**, 1192-1199 (2002).

- 120 Lyle, C. L., Belghasem, M. & Chitalia, V. C. c-Cbl: an important regulator and a target in  
angiogenesis and tumorigenesis. *Cells* **8**, 498 (2019).
- 121 Kales, S. C., Ryan, P. E., Nau, M. M. & Lipkowitz, S. Cbl and human myeloid  
neoplasms: the Cbl oncogene comes of age. *Cancer research* **70**, 4789-4794 (2010).
- 122 Schmidt, M. H. & Dikic, I. The Cbl interactome and its functions. *Nature reviews  
Molecular cell biology* **6**, 907-919 (2005).
- 123 Beitel, L. *et al.* Cloning and characterization of an androgen receptor N-terminal-  
interacting protein with ubiquitin-protein ligase activity. *Journal of molecular  
endocrinology* **29**, 41-60 (2002).
- 124 Logan, I. R. *et al.* Human PIRH2 enhances androgen receptor signaling through  
inhibition of histone deacetylase 1 and is overexpressed in prostate cancer. *Molecular  
and cellular biology* **26**, 6502-6510 (2006).
- 125 Daks, A. *et al.* The RNA-binding protein HuR is a novel target of Pirh2 E3 ubiquitin  
ligase. *Cell death & disease* **12**, 1-14 (2021).
- 126 Zilfou, J. T. & Lowe, S. W. Tumor suppressive functions of p53. *Cold Spring Harbor  
perspectives in biology* **1**, a001883 (2009).
- 127 Leng, R. P. *et al.* Pirh2, a p53-induced ubiquitin-protein ligase, promotes p53  
degradation. *Cell* **112**, 779-791 (2003).
- 128 Sheng, Y. *et al.* Molecular basis of Pirh2-mediated p53 ubiquitylation. *Nature structural  
& molecular biology* **15**, 1334-1342 (2008).
- 129 Maruyama, S. *et al.* Ubiquitylation of  $\epsilon$ -COP by PIRH2 and regulation of the secretion of  
PSA. *Molecular and cellular biochemistry* **307**, 73-82 (2008).
- 130 Bohgaki, M. *et al.* The E3 ligase PIRH2 polyubiquitylates CHK2 and regulates its  
turnover. *Cell Death & Differentiation* **20**, 812-822 (2013).
- 131 Jung, Y.-S., Liu, G. & Chen, X. Pirh2 E3 ubiquitin ligase targets DNA polymerase eta for  
20S proteasomal degradation. *Molecular and cellular biology* **30**, 1041-1048 (2010).
- 132 Jung, Y.-S., Hakem, A., Hakem, R. & Chen, X. Pirh2 E3 ubiquitin ligase  
monoubiquitinates DNA polymerase eta to suppress translesion DNA synthesis.  
*Molecular and cellular biology* **31**, 3997-4006 (2011).
- 133 Hakem, A. *et al.* Role of Pirh2 in mediating the regulation of p53 and c-Myc. *PLoS  
genetics* **7**, e1002360 (2011).
- 134 Sheren, J. E. & Kassenbrock, C. K. RNF38 encodes a nuclear ubiquitin protein ligase that  
modifies p53. *Biochemical and biophysical research communications* **440**, 473-478  
(2013).
- 135 Zhang, J. *et al.* RING finger protein 38 induces gastric cancer cell growth by decreasing  
the stability of the protein tyrosine phosphatase SHP-1. *FEBS letters* **592**, 3092-3100  
(2018).
- 136 Eisenberg, I. *et al.* Cloning and characterization of a novel human gene RNF38 encoding  
a conserved putative protein with a RING finger domain. *Biochemical and biophysical  
research communications* **294**, 1169-1176 (2002).
- 137 Peng, R. *et al.* Overexpression of RNF38 facilitates TGF- $\beta$  signaling by Ubiquitinating  
and degrading AHNAK in hepatocellular carcinoma. *Journal of Experimental & Clinical  
Cancer Research* **38**, 1-14 (2019).
- 138 Xiong, D. *et al.* Ring finger protein 38 promote non-small cell lung cancer progression by  
endowing cell EMT phenotype. *Journal of Cancer* **9**, 841 (2018).

- 139 Huang, Z. *et al.* RING finger protein 38 mediates LIM domain binding 1 degradation and regulates cell growth in colorectal cancer. *OncoTargets and therapy* **13**, 371 (2020).
- 140 Abou Zeinab, R. *et al.* Pirh2, an E3 ligase, regulates the AIP4–p73 regulatory pathway by modulating AIP4 expression and ubiquitination. *Carcinogenesis* **42**, 650-662 (2021).
- 141 Yonezawa, T. *et al.* The ubiquitin ligase RNF38 promotes RUNX1 ubiquitination and enhances RUNX1-mediated suppression of erythroid transcription program. *Biochemical and biophysical research communications* **505**, 905-909 (2018).
- 142 Hunter, M., Yuan, P., Vavilala, D. & Fox, M. Optimization of protein expression in mammalian cells. *Current protocols in protein science* **95**, e77 (2019).
- 143 Pennock, S. & Wang, Z. A tale of two Cbls: interplay of c-Cbl and Cbl-b in epidermal growth factor receptor downregulation. *Molecular and cellular biology* **28**, 3020-3037 (2008).
- 144 Dou, H. *et al.* Structural basis for autoinhibition and phosphorylation-dependent activation of c-Cbl. *Nature structural & molecular biology* **19**, 184-192 (2012).
- 145 Santamaría, D. *et al.* Cdk1 is sufficient to drive the mammalian cell cycle. *Nature* **448**, 811-815 (2007).
- 146 Berthet, C., Aleem, E., Coppola, V., Tessarollo, L. & Kaldis, P. Cdk2 knockout mice are viable. *Current Biology* **13**, 1775-1785 (2003).
- 147 Diril, M. K. *et al.* Cyclin-dependent kinase 1 (Cdk1) is essential for cell division and suppression of DNA re-replication but not for liver regeneration. *Proceedings of the National Academy of Sciences* **109**, 3826-3831 (2012).
- 148 Cutillo, G., Simon, D. K. & Eleuteri, S. VPS35 and the mitochondria: connecting the dots in Parkinson's disease pathophysiology. *Neurobiology of Disease*, 105056 (2020).
- 149 Seaman, M. N., Marcusson, E. G., Cereghino, J. L. & Emr, S. D. Endosome to Golgi retrieval of the vacuolar protein sorting receptor, Vps10p, requires the function of the VPS29, VPS30, and VPS35 gene products. *Journal of cell biology* **137**, 79-92 (1997).
- 150 Schwintzer, L., Roca, E. A. & Broemer, M. TRIAD3/RNF216 E3 ligase specifically synthesises K63-linked ubiquitin chains and is inactivated by mutations associated with Gordon Holmes syndrome. *Cell death discovery* **5**, 1-11 (2019).
- 151 Williams, E. T. *et al.* Parkin mediates the ubiquitination of VPS35 and modulates retromer-dependent endosomal sorting. *Human molecular genetics* **27**, 3189-3205 (2018).
- 152 Gehring, N. H. *et al.* Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. *Molecular cell* **20**, 65-75 (2005).
- 153 Singh, K. K., Wachsmuth, L., Kulozik, A. E. & Gehring, N. H. Two mammalian MAGOH genes contribute to exon junction complex composition and nonsense-mediated decay. *RNA biology* **10**, 1291-1298 (2013).
- 154 Rehman, A., Chandra, P. & Singh, K. K. The MAGOH paralogs-MAGOH, MAGOHB and their multiple isoforms. *Gene Reports* **24**, 101214 (2021).
- 155 Zhou, Y. *et al.* MAGOH/MAGOHB Inhibits the Tumorigenesis of Gastric Cancer via Inactivation of b-RAF/MEK/ERK Signaling. *OncoTargets and therapy* **13**, 12723 (2020).
- 156 De Silanes, I. L., Zhan, M., Lal, A., Yang, X. & Gorospe, M. Identification of a target RNA motif for RNA-binding protein HuR. *Proceedings of the National Academy of Sciences* **101**, 2987-2992 (2004).

- 157 Uren, P. J. *et al.* Genomic analyses of the RNA-binding protein Hu antigen R (HuR) identify a complex network of target genes and novel characteristics of its binding sites. *Journal of Biological Chemistry* **286**, 37063-37066 (2011).
- 158 Cosker, K. E., Fenstermacher, S. J., Pazyra-Murphy, M. F., Elliott, H. L. & Segal, R. A. The RNA-binding protein SFPQ orchestrates an RNA regulon to promote axon viability. *Nature neuroscience* **19**, 690-696 (2016).
- 159 Ke, Y. *et al.* Tau-mediated nuclear depletion and cytoplasmic accumulation of SFPQ in Alzheimer's and Pick's disease. *PloS one* **7**, e35678 (2012).
- 160 Thakar, T. *et al.* Ubiquitinated-PCNA protects replication forks from DNA2-mediated degradation by regulating Okazaki fragment maturation and chromatin assembly. *Nature communications* **11**, 1-14 (2020).
- 161 Wang, P., Dai, X., Jiang, W., Li, Y. & Wei, W. in *Seminars in cancer biology*. 131-144 (Elsevier).
- 162 Dove, K. K. *et al.* Structural studies of HHARI/UbcH7~ Ub reveal unique E2~ Ub conformational restriction by RBR RING1. *Structure* **25**, 890-900. e895 (2017).
- 163 Spratt, D. E., Walden, H. & Shaw, G. S. RBR E3 ubiquitin ligases: new structures, new insights, new questions. *Biochemical journal* **458**, 421-437 (2014).
- 164 Yuan, L., Lv, Z., Atkison, J. H. & Olsen, S. K. Structural insights into the mechanism and E2 specificity of the RBR E3 ubiquitin ligase HHARI. *Nature communications* **8**, 1-14 (2017).
- 165 Scott, D. C. *et al.* Two distinct types of E3 ligases work in unison to regulate substrate ubiquitylation. *Cell* **166**, 1198-1214. e1124 (2016).
- 166 Wang, F., Kuang, Y., Salem, N., Anderson, P. W. & Lee, Z. Cross-species hybridization of woodchuck hepatitis viral infection-induced woodchuck hepatocellular carcinoma using human, rat and mouse oligonucleotide microarrays. *Journal of gastroenterology and hepatology* **24**, 605-617 (2009).
- 167 Tan, N. G. *et al.* Human homologue of ariadne promotes the ubiquitylation of translation initiation factor 4E homologous protein, 4EHP. *FEBS letters* **554**, 501-504 (2003).
- 168 Rom, E. *et al.* Cloning and characterization of 4EHP, a novel mammalian eIF4E-related cap-binding protein. *Journal of Biological Chemistry* **273**, 13104-13109 (1998).
- 169 Morita, M. *et al.* A novel 4EHP-GIGYF2 translational repressor complex is essential for mammalian development. *Molecular and cellular biology* **32**, 3585-3593 (2012).
- 170 von Stechow, L. *et al.* The E3 ubiquitin ligase ARIH1 protects against genotoxic stress by initiating a 4EHP-mediated mRNA translation arrest. *Molecular and cellular biology* **35**, 1254-1268 (2015).
- 171 Lechtenberg, B. C. *et al.* Structure of a HOIP/E2~ ubiquitin complex reveals RBR E3 ligase mechanism and regulation. *Nature* **529**, 546-550 (2016).
- 172 Chew, K. C. *et al.* Parkin mediates apparent E2-independent monoubiquitination in vitro and contains an intrinsic activity that catalyzes polyubiquitination. *PloS one* **6**, e19720 (2011).
- 173 Dove, K. K., Stieglitz, B., Duncan, E. D., Rittinger, K. & Klevit, R. E. Molecular insights into RBR E3 ligase ubiquitin transfer mechanisms. *EMBO reports* **17**, 1221-1235 (2016).
- 174 Ramírez, J. *et al.* The ubiquitin ligase Ariadne-1 regulates neurotransmitter release via ubiquitination of NSF. *Journal of Biological Chemistry* **296** (2021).

- 175 Gudipaty, S. A., McNamara, R. P., Morton, E. L. & D'Orso, I. PPM1G binds 7SK RNA  
and Hexim1 to block P-TEFb assembly into the 7SK snRNP and sustain transcription  
elongation. *Molecular and cellular biology* **35**, 3810-3828 (2015).
- 176 Chen, D. *et al.* PPM1G promotes the progression of hepatocellular carcinoma via  
phosphorylation regulation of alternative splicing protein SRSF3. *Cell death & disease*  
**12**, 1-11 (2021).
- 177 Huang, Q. *et al.* ACTN4 promotes the proliferation, migration, metastasis of  
osteosarcoma and enhances its invasive ability through the NF- $\kappa$ B pathway. *Pathology &*  
*Oncology Research* **26**, 893-904 (2020).
- 178 Honda, K. *et al.* Actinin-4, a novel actin-bundling protein associated with cell motility  
and cancer invasion. *The Journal of cell biology* **140**, 1383-1393 (1998).
- 179 Liu, X. & Chu, K.-M.  $\alpha$ -Actinin-4 promotes metastasis in gastric cancer. *Laboratory*  
*Investigation* **97**, 1084-1094 (2017).
- 180 Blondelle, J. *et al.* Cullin-3-dependent deregulation of ACTN1 represents a pathogenic  
mechanism in nemaline myopathy. *JCI insight* **4** (2019).
- 181 Romero, N. B., Sandaradura, S. A. & Clarke, N. F. Recent advances in nemaline  
myopathy. *Current opinion in neurology* **26**, 519-526 (2013).
- 182 North, K. & Ryan, M. M. Nemaline myopathy. (2015).
- 183 Luo, S.-Q. *et al.* C1orf35 contributes to tumorigenesis by activating c-MYC transcription  
in multiple myeloma. *Oncogene* **39**, 3354-3366 (2020).
- 184 Whitehouse, D. B., Tomkins, J., Lovegrove, J. U., Hopkinson, D. A. & McMillan, W. O.  
A phylogenetic approach to the identification of phosphoglucomutase genes. *Molecular*  
*biology and evolution* **15**, 456-462 (1998).
- 185 Cao, B. *et al.* Knockdown of PGM1 enhances anticancer effects of orlistat in gastric  
cancer under glucose deprivation. *Cancer cell international* **21**, 1-15 (2021).
- 186 Jiang, P. *et al.* Aberrant expression of nuclear KPNA2 is correlated with early recurrence  
and poor prognosis in patients with small hepatocellular carcinoma after hepatectomy.  
*Medical oncology* **31**, 131 (2014).
- 187 Jiang, L., Liu, J., Wei, Q. & Wang, Y. KPNA2 expression is a potential marker for  
differential diagnosis between osteosarcomas and other malignant bone tumor mimics.  
*Diagnostic Pathology* **15**, 1-9 (2020).
- 188 Wang, Y. *et al.* Orthogonal ubiquitin transfer reveals human papillomavirus E6  
downregulates nuclear transport to disarm interferon- $\gamma$  dependent apoptosis of cervical  
cancer cells. *The FASEB Journal* **35**, e21986 (2021).
- 189 Beaudenon, S. & Huibregtse, J. M. HPV E6, E6AP and cervical cancer. *BMC*  
*biochemistry* **9**, 1-7 (2008).
- 190 Huibregtse, J. M., Scheffner, M. & Howley, P. M. Cloning and expression of the cDNA  
for E6-AP, a protein that mediates the interaction of the human papillomavirus E6  
oncoprotein with p53. *Molecular and cellular biology* **13**, 775-784 (1993).
- 191 Peng, K. *et al.* Regulation of O-Linked N-Acetyl Glucosamine Transferase (OGT)  
through E6 Stimulation of the Ubiquitin Ligase Activity of E6AP. *International journal*  
*of molecular sciences* **22**, 10286 (2021).
- 192 Huibregtse, J. M., Scheffner, M. & Howley, P. M. Localization of the E6-AP regions that  
direct human papillomavirus E6 binding, association with p53, and ubiquitination of  
associated proteins. *Molecular and cellular biology* **13**, 4918-4927 (1993).

- 193 Huibregtse, J. M., Scheffner, M. & Howley, P. M. A cellular protein mediates association  
of p53 with the E6 oncoprotein of human papillomavirus types 16 or 18. *The EMBO  
journal* **10**, 4129-4135 (1991).
- 194 Khatri, N. & Man, H.-Y. The autism and Angelman syndrome protein Ube3A/E6AP: the  
gene, E3 ligase ubiquitination targets and neurobiological functions. *Frontiers in  
molecular neuroscience* **12**, 109 (2019).
- 195 Tomaić, V. & Banks, L. Angelman syndrome-associated ubiquitin ligase UBE3A/E6AP  
mutants interfere with the proteolytic activity of the proteasome. *Cell death & disease* **6**,  
e1625-e1625 (2015).
- 196 El Hokayem, J., Weeber, E. & Nawaz, Z. Loss of Angelman syndrome protein E6AP  
disrupts a novel antagonistic estrogen-retinoic acid transcriptional crosstalk in neurons.  
*Molecular neurobiology* **55**, 7187-7200 (2018).
- 197 George, A. J., Hoffiz, Y. C., Charles, A. J., Zhu, Y. & Mabb, A. M. A comprehensive  
atlas of E3 ubiquitin ligase mutations in neurological disorders. *Frontiers in genetics* **9**,  
29 (2018).
- 198 Mabb, A. M., Judson, M. C., Zylka, M. J. & Philpot, B. D. Angelman syndrome: insights  
into genomic imprinting and neurodevelopmental phenotypes. *Trends in neurosciences*  
**34**, 293-303 (2011).
- 199 Gerry, C. J. & Schreiber, S. L. Unifying principles of bifunctional, proximity-inducing  
small molecules. *Nature chemical biology* **16**, 369-378 (2020).
- 200 Lai, A. C. & Crews, C. M. Induced protein degradation: an emerging drug discovery  
paradigm. *Nature reviews Drug discovery* **16**, 101-114 (2017).
- 201 Ronchi, V. P., Kim, E. D., Summa, C. M., Klein, J. M. & Haas, A. L. In silico modeling  
of the cryptic E2~ ubiquitin-binding site of E6-associated protein (E6AP)/UBE3A  
reveals the mechanism of polyubiquitin chain assembly. *Journal of Biological Chemistry*  
**292**, 18006-18023 (2017).
- 202 Scheffner, M., Werness, B. A., Huibregtse, J. M., Levine, A. J. & Howley, P. M. The E6  
oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation  
of p53. *Cell* **63**, 1129-1136 (1990).
- 203 Yamagishi, Y. *et al.* Natural product-like macrocyclic N-methyl-peptide inhibitors  
against a ubiquitin ligase uncovered from a ribosome-expressed de novo library.  
*Chemistry & biology* **18**, 1562-1570 (2011).
- 204 Malecka, K. A. *et al.* Identification and characterization of small molecule human  
papillomavirus E6 inhibitors. *ACS chemical biology* **9**, 1603-1612 (2014).
- 205 Cherry, J. J. *et al.* Structure based identification and characterization of flavonoids that  
disrupt human papillomavirus-16 E6 function. *PloS one* **8**, e84506 (2013).
- 206 Wolter, J. M. *et al.* Cas9 gene therapy for Angelman syndrome traps Ube3a-ATS long  
non-coding RNA. *Nature* **587**, 281-284 (2020).
- 207 Meng, L. *et al.* Towards a therapy for Angelman syndrome by targeting a long non-  
coding RNA. *Nature* **518**, 409-412 (2015).
- 208 Offensperger, F. *et al.* Identification of small-molecule activators of the ubiquitin ligase  
E6AP/UBE3A and Angelman syndrome-derived E6AP/UBE3A variants. *Cell chemical  
biology* **27**, 1510-1520. e1516 (2020).
- 209 Sang, P. *et al.* Sulfonyl- $\gamma$ -AApeptides as helical mimetics: crystal structures and  
applications. *Accounts of Chemical Research* **53**, 2425-2442 (2020).

- 210 Fajner, V., Maspero, E. & Polo, S. Targeting HECT-type E3 ligases—insights from  
catalysis, regulation and inhibitors. *FEBS letters* **591**, 2636-2647 (2017).
- 211 Shi, Y. *et al.* One-bead–two-compound thioether bridged macrocyclic  $\gamma$ -AApeptide  
screening library against EphA2. *Journal of medicinal chemistry* **60**, 9290-9298 (2017).
- 212 Kuslansky, Y. *et al.* Ubiquitin ligase E6AP mediates nonproteolytic polyubiquitylation of  
 $\beta$ -catenin independent of the E6 oncoprotein. *Journal of General Virology* **97**, 3313-3330  
(2016).
- 213 Jason, J. Y. *et al.* The autism-linked UBE3A T485A mutant E3 ubiquitin ligase activates  
the Wnt/ $\beta$ -catenin pathway by inhibiting the proteasome. *Journal of Biological Chemistry*  
**292**, 12503-12515 (2017).
- 214 Pol, S. B. V. & Klingelutz, A. J. Papillomavirus E6 oncoproteins. *Virology* **445**, 115-  
137 (2013).
- 215 She, F. *et al.* De Novo Left-Handed Synthetic Peptidomimetic Foldamers. *Angewandte  
Chemie* **130**, 10064-10068 (2018).
- 216 Shi, Y. *et al.* Helical sulfono- $\gamma$ -AApeptides with aggregation-induced emission and  
circularly polarized luminescence. *Journal of the American Chemical Society* **141**,  
12697-12706 (2019).
- 217 Wang, M. *et al.* Development of bis-cyclic imidazolidine-4-one derivatives as potent  
antibacterial agents. *Journal of Medicinal Chemistry* **63**, 15591-15602 (2020).
- 218 Wang, M. *et al.* Modular Design of Membrane-Active Antibiotics: From Macromolecular  
Antimicrobials to Small Scorpionlike Peptidomimetics. *Journal of medicinal chemistry*  
(2021).
- 219 Sang, P. *et al.* Inhibition of  $\beta$ -catenin/B cell lymphoma 9 protein– protein interaction  
using  $\alpha$ -helix–mimicking sulfono- $\gamma$ -AApeptide inhibitors. *Proceedings of the National  
Academy of Sciences* **116**, 10757-10762 (2019).
- 220 Sang, P. *et al.*  $\alpha$ -Helix-mimicking sulfono- $\gamma$ -AApeptide inhibitors for p53–  
MDM2/MDMX protein–protein interactions. *Journal of medicinal chemistry* **63**, 975-986  
(2020).
- 221 Teng, P. *et al.* The folding propensity of  $\alpha$ /sulfono- $\gamma$ -AA peptidic foldamers with both  
left-and right-handedness. *Communications Chemistry* **4**, 1-9 (2021).
- 222 Zhang, W. *et al.* System-wide modulation of HECT E3 ligases with selective ubiquitin  
variant probes. *Molecular cell* **62**, 121-136 (2016).
- 223 Wertz, I. E. & Wang, X. From discovery to bedside: targeting the ubiquitin system. *Cell  
chemical biology* **26**, 156-177 (2019).
- 224 Cathcart, A. M. *et al.* Targeting a helix-in-groove interaction between E1 and E2 blocks  
ubiquitin transfer. *Nature chemical biology* **16**, 1218-1226 (2020).

## APPENDICES

### Appendix A

#### *Published and submitted Papers*

Bo Huang\*, **Li Zhou\***, Ruochuan Liu\*, et al. "Activation of E6AP/UBE3A-Mediated Protein Ubiquitination and Degradation Pathways by a Cyclic  $\gamma$ -AA peptide". manuscript submitted.2021, co-first author, accepted by Journal of Medical Chemistry

Wang, Y., Liu, R., Liao, J., Jiang, L., Jeong, G. H., **Zhou, L.**, ... & Yin, J. (2021). Orthogonal ubiquitin transfer reveals human papillomavirus E6 downregulates nuclear transport to disarm interferon- $\gamma$  dependent apoptosis of cervical cancer cells. *The FASEB Journal*, 35(11), e21986.

Peng, K., Liu, R., Jia, C., Wang, Y., Jeong, G. H., **Zhou, L.**, ... & Zhao, B. (2021). Regulation of O-Linked N-Acetyl Glucosamine Transferase (OGT) through E6 Stimulation of the Ubiquitin Ligase Activity of E6AP. *International Journal of Molecular Sciences*, 22(19), 10286.

Wang, Y., Fang, S., Chen, G., Ganti, R., Chernova, T. A., **Zhou, L.**, ... & Yin, J. (2021). Regulation of the endocytosis and prion-chaperoning machineries by yeast E3 ubiquitin ligase Rsp5 as revealed by orthogonal ubiquitin transfer. *Cell Chemical Biology*.

Yin, J., Wang, Y., Bhuripanyo, K., Chen, G., **Zhou, L.**, Liu, R., & Zhou, H. (2018). Identifying the Substrate Proteins of E3 Ubiquitin Ligase by Orthogonal Ubiquitin Transfer (OUT). *The FASEB Journal*, 32, 654-13.

Bhuripanyo, K., Wang, Y., Liu, X., **Zhou, L.**, Liu, R., Duong, D., ... & Yin, J. (2018). Identifying the substrate proteins of U-box E3s E4B and CHIP by orthogonal ubiquitin transfer. *Science advances*, 4(1), e1701393.

Wang, Y., Liu, X., **Zhou, L.**, Duong, D., Bhuripanyo, K., Zhao, B., ... & Yin, J. (2017). Identifying the ubiquitination targets of E6AP by orthogonal ubiquitin transfer. *Nature communications*, 8(1), 1-14.