THE UNIVERSITY OF CHICAGO

IDENTIFICATION OF UBIQUITIN LIGASE SUBSTRATES VIA ORTHOGONAL UBIQUITIN TRANSFER

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMISTRY

BY

KARAN BHURIPANYO

CHICAGO, ILLINOIS

JUNE 2016

Dedicated to all my beloved family and friends, but above all,

to Him whose love has made me more than a conqueror.

TABLE OF CONTENTS

CHAPTER 4

OUT CASCADE THROUGH E3 MUTANTS AND

IN VITRO VERIFICATION OF ACTIVITY …….96

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

LIST OF TABLES

CHAPTER 2

CHAPTER 3

CHAPTER 4

\

CHAPTER 5

PREFACE AND ACKNOWLEDGEMENT

I must confess, the format of this thesis, including the rough number of pages and references, as well as the number of chapters, was virtually copied in its entirety from the thesis of Keya Zhang, my senior graduate student and mentor, who is now a PhD graduate of our group. In the preface of his thesis, he mentioned “frantically trying to wrap up the writing of the thesis” on the very eve of his dissertation defense—that part brought a smile to my face the first time I read it; I have often consulted his thesis for specific procedures and protocols, but yet, for some reason, this part, this particular sentence of his thesis was the part that left the deepest impression on me.

And here I am, in my own turn, a few months short of four years afterwards, frantically—perhaps not as frantically as my mentor had done, and yet frantically enough, putting together my own thesis three days before my dissertation defense—it is probably natural that one takes after his mentor in both the good things and the bad, but unfortunately, in my case it seems that I have somehow inherited the tendency to procrastinate—notwithstanding that I was probably already a procrastinator by nature at the very start. That said, looking back, I would very much like to say that the data contained in this thesis: our research, its challenges and setbacks, as well as its success and achievements, were very much the result of a so-called labor of love—but then I probably would be half lying. More accurately, it was more of a labor of a complicated juxtaposition of different seemingly contradictory things: love and hate, fulfillment and frustration, boredom and excitement, hope and disappointment: the things which normally make up everyday life. There were times I would be excited about my research and its results and work tirelessly into the wee hours of the morning to complete my experiments, and then on the other hand there were times I felt completely sick of my research—the mundane and routine repetition

of the same old techniques and experiments and the frustration of getting confounding results—I felt as if I would vomit if I had stayed in the lab a moment longer; I wanted to get away from it all; there were such times as well. I have never known any PhD student who has always been enamored by his or her own research, and I believe I probably never will. However, I am sure that when the time of convocation comes, we would all agree that ultimately the fruits of our labors—the sense of fulfillment, and the feeling of conquest which comes from having fully exerted oneself and emerge triumphant in the end—are exceedingly sweet. It is with this feeling that I would like to thank all the people in my life who have made my success possible.

I am extremely, extremely grateful to my research advisor, Prof. Jun Yin, first and foremost for his patience in dealing with many of my shortcomings, and secondly for his support and advice in all sorts of circumstances. I would like to thank him especially for instilling a sense of discipline into my research in a manner which no one else would have done; while his hands-on, detailed, micromanaging, and yet, realistic approach at times clashed with my tendency to want to be independent and explore new ideas and protocols, in the end, I feel that our tendencies offset each other to strike the perfect balance: I was able to learn how to think independently, how to obtain information and evaluate each piece of advice I received from various sources, and to ultimately solve research problems on my own judgment, while still staying focused and realistic, and getting the actual results expected of me. Again, thank you, Dr. Yin for putting up with me, and being patient and supportive throughout these five long years, even though we do not always see eye-to-eye, and I hope my stubbornness has not caused your hair to turn even more prematurely grey. I hope you realize that all my past and future success as a researcher is your success as well; it was a privilege to be your student and I will strive to make you proud in my future endeavors.

I would also like to thank my mentors, Keya Zhang, and Bo Zhao, a former postdoctoral fellow of our group, for allowing me to learn from them—even to the point of selflessly giving me hands-on experience during my earliest days as a member of our group, at the cost of sometimes ruining their experiments by my incompetence. In addition to the relevant laboratory techniques and knowledge, they have also taught me how to survive in the lab, how to negotiate with one's research advisor, and how to manage one's time to somehow have a fighting chance of retaining one's sanity while on a working schedule of sixty to seventy hours per week.

I would like to thank all the people who had joined our group when we moved to Georgia State University; especially Dr. Yiyang Wang, a valued collaborator and good friend, who is also the hardest-working person I have ever known, and through whose Herculean efforts the completion of my main project was possible. I would also like to express my thanks to the rest of the group: Li, Han, Geng, Jessica, Archana, and others, for providing their friendships and interesting conversations, and making everyday life in the lab tolerable, if not sometimes downright enjoyable, and above all for being partakers in my misery, as well as in my happiness.

Of course, I'd like to thank the numerous professors at the University of Chicago for their instruction and advice in the classes I have taken, and especially Profs Chuan He and Joseph Piccirilli for serving as my committee members for my dissertation defense. I would also like to thank the chemistry department, as well as the university itself for giving me, whose credentials were, in my opinion quite far from being lackluster, the opportunity to study in the prestigious chemistry PhD program. I would like to thank the department of Chemistry at Georgia State University as well, for giving Prof. Yin the opportunity and resources to continue his research there. I also would like to thank our collaborators from other institutions, those who have been kind enough to provide us with certain materials, as well as practical advice when we ventured

into an unfamiliar field of work, especially Prof. Hiroaki Kiyokawa and his group at Northwestern University who proved to be tremendous help in familiarizing our group with the cellular biology side of things, and in providing valuable constructs for our research.

In my personal life, I am eternally grateful for my parents for their indescribable love and support, and for their encouragement and discipline, and for being there to convince me not to give up when I had at times felt like quitting. I am especially grateful for their instruction, their life experiences and their advice in all the matters of life. I am also grateful for the kindness and love from my extended family, which consists of my grandparents, twenty or so aunts and uncles on both sides, as well as several dozens of cousins, many of whom grew up with me and are as my brothers and sisters. I am also grateful for the friendship of those in my church, my pastors, my friends, my teachers, and my classmates, and I would also like to extend my gratitude for all the people who have been in my life—I am thankful for all of my experiences, both good and bad. Above all, my ultimate thanks are to my God, the Lord Jesus Christ, without whom I can do and have done nothing. Through this thesis, my five-and-a-half seemingly long years of research as a graduate student has been condensed and summarized into a volume consisting of a mere one hundred and fifty pages, but if I had to write of the things my God has done personally for me, and of His Grace in all the aspects of my utterly undeserving life, perhaps I should quote the words of the beloved Apostle: the world would not be enough to contain all the books which would be written.

ABSTRACT

This Ph.D. dissertation describes the main project of my academic research at the University of Chicago as a graduate student: the identification of substrates of ubiquitin E3 ligases using our lab's unique system called orthogonal ubiquitin transfer (OUT). Each chapter describes different, sequential steps in how the goal was finally achieved; each chapter builds on the previous, akin to rungs in a ladder, and together they demonstrate a logical, sequential way of approaching and solving the problem in distinct steps. Chapter 1 describes background information about the ubiquitin-proteasome system, as well as the Yin lab's previous research and established information on OUT at the time of my joining the program, and also includes some parts of my work in the earliest year of my study. Chapter 2 describes my own part in the development of a ubiquitin ligase for OUT via rational design and specific point mutagenesis, as well as my work and analysis on the E2-E3 interface. Chapter 3 involves the construction of a phage display library displaying the U-box domain of the ubiquitin ligase E4B, as well as the screening for active mutants, and the analysis of the consensus of the sequences of the mutants selected. Chapter 4 involves the characterization of the selected mutants and the assaying of their activity with OUT, as well the application of the selected amino acid sequences onto corresponding residues on other U-box E3 ligases, namely, the yeast homolog Ufd2 and CHIP, and the activity of those mutant analogs on putative substrates. Chapter 5, the last chapter, describes the practical application of the whole OUT cascade *in vivo*, and the proteomics screening for substrates, and the verification of such substrates. The potential hits identified during our *in vivo* proteomics screens, which were run in duplicate, are as attached in the supplemental files available online.

LIST OF PUBLICATIONS

1 Zhang, K., Bhuripanyo, K., Wang, Y. & Yin, J. Coupling Binding to Catalysis: Using Yeast Cell Surface Display to Select Enzymatic Activities. *Methods in molecular biology (Clifton, N.J.)* **1319**, 245-260, doi:10.1007/978-1-4939-2748-7_14 (2015).

2 Zhang, K. *et al.* Engineering new protein-protein interactions on the beta-propeller fold by yeast cell surface display. *Chembiochem : a European journal of chemical biology* **14**, 426-430, doi:10.1002/cbic.201200718 (2013).

3 Zhang, K. *et al.* Engineering the substrate specificity of the DhbE adenylation domain by yeast cell surface display. *Chemistry & biology* **20**, 92-101, doi:10.1016/j.chembiol. 2012.10.020 (2013).

4 Zhao, B. *et al.* Specificity of the E1-E2-E3 enzymatic cascade for ubiquitin C-terminal sequences identified by phage display. *ACS chemical biology* **7**, 2027-2035, doi:10.1021/cb300339p (2012).

5 Zhao, B. *et al.* Orthogonal ubiquitin transfer through engineered E1-E2 cascades for protein ubiquitination. *Chemistry & biology* **19**, 1265-1277, doi:10.1016/ j.chembiol.2012.07.023 (2012).

6 Zhao, B. *et al.* Inhibiting the protein ubiquitination cascade by ubiquitin-mimicking short peptides. *Organic letters* **14**, 5760-5763, doi:10.1021/ol3027736 (2012).

7 Zhao, B. *et al.* SUMO-mimicking peptides inhibiting protein SUMOylation. *Chembiochem : a European journal of chemical biology* **15**, 2662-2666, doi:10.1002/cbic.201402472 (2014).

8 Zhao, B. *et al.* Profiling the cross reactivity of ubiquitin with the Nedd8 activating enzyme by phage display. *PLoS One* **8**, e70312, doi:10.1371/journal.pone.0070312 (2013).

9 Zhao, B. *et al.* Phage selection assisted by Sfp phosphopantetheinyl transferase-catalyzed site-specific protein labeling. *Methods in molecular biology (Clifton, N.J.)* **1266**, 161-170, doi:10.1007/978-1-4939-2272-7_11 (2015).

10 Zhao, B. *et al.* Phage display to identify Nedd8-mimicking peptides as inhibitors of the Nedd8 transfer cascade. *Chembiochem : a European journal of chemical biology* **14**, 1323-1330, doi:10.1002/cbic.201300234 (2013).

TO BE PUBLISHED

11 Orthogonal Ubiquitin Transfer to E4B to Identify its Ubiquitination Targets in the Cell (In preparation)

12 Profiling CHIP substrates by orthogonal ubiquitin transfer (OUT) (In preparation)

Chapter 1

Introduction: the rationale and development of orthogonal ubiquitin transfer (OUT)

1.1 The ubiquitin-proteasome pathway

Ubiquitination is known as one of the most common post-translational modification of proteins—second in abundance only to phosphorylation[1]; as the name implies, the ubiquitous nature of such a modification belies a vastly complicated network of checkpoints and regulations involving virtually all proteins in any eukaryotic cell. The most well-known function of ubiquitination is targeting proteins for proteasomal degradation [2-4]; although various types of ubiquitination: different linkages or ubiquitination on different sites, can function to convey a myriad of other non-degradative signals [5-10].

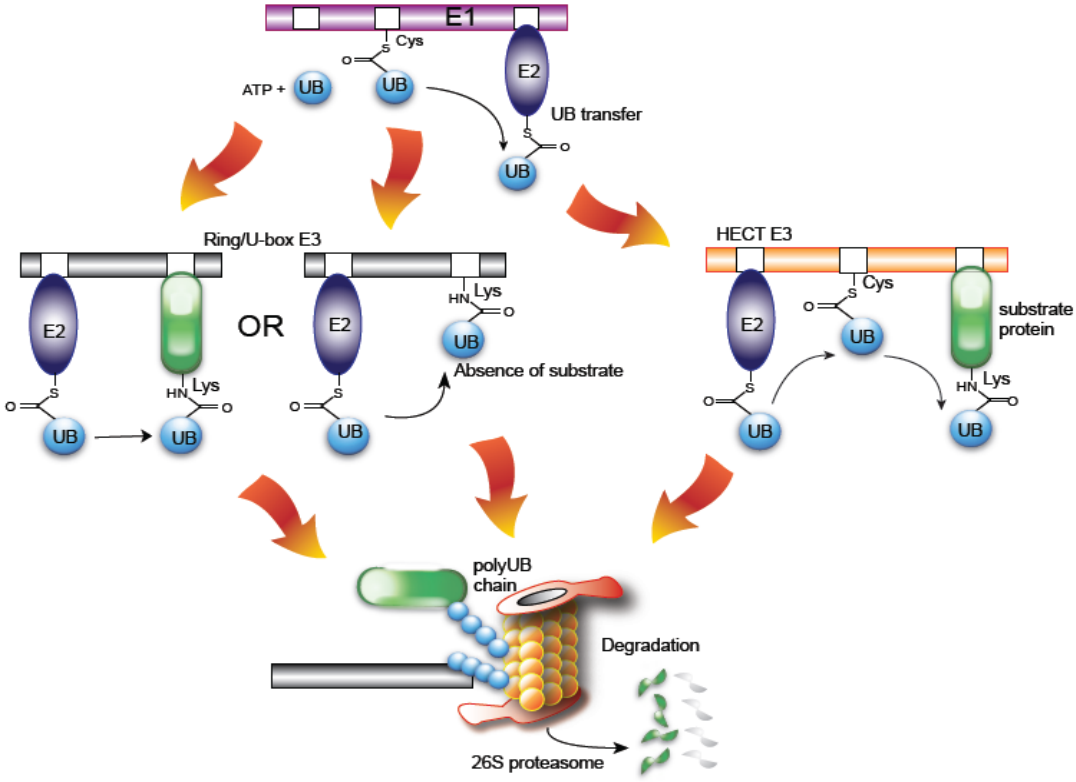

Figure 1-1 Brief overview of the ubiquitin (UB) transfer cascade
Although proteasomal degradation is the most common fate for ubiquitinated substrates proteins, as shown here, this is not always necessarily so.

The eukaryotic cell utilizes a sequential cascade of enzymes to transfer ubiquitin (UB) onto the target substrate protein[11] **(Figure 1-1)**; the first step involves a class of enzymes termed ubiquitin activating enzymes (E1) which, upon self-catalyzed adenylation of the ubiquitin C-terminal carboxyl group, which serves to activate the carbon for nucleophilic attack, subsequently transfers the ubiquitin terminal onto itself via a catalytic nucleophilic cysteine thiol. The second class of enzymes are known as ubiquitin conjugating enzymes (E2); an E2 receives the ubiquitin C-terminus from E1 enzymes through a similar mechanism involving its own catalytic nucleophilic cysteine through a transthioesterification reaction. The next step involves a class of enzymes which are called ubiquitin ligases (E3). These E3s are either of HECT-type, which receives ubiquitin onto its own catalytic cysteine residue from the E2-ub thioester conjugate, in a similar manner to E2s, or of the RING-type[12]. RING E3s, in the presence of a suitable substrate, do not directly load ubiquitin covalently onto themselves, but act as a bridge to bring the E2~UB conjugate and the final substrate into close proximity for ubiquitination to occur, forming a stable amide bond between, in the vast majority of cases, the ε-amino group of the substrate's lysine residue (with the exception of the extremely rare N-terminal ubiquitination[13-15], in which case it is with the amino group at the N-terminal of the target substrate protein) and the ubiquitin C-terminal carboxyl group. In the absence of substrates, both HECT[16] and RING-type[17] E3s are known to ubiquitinate themselves on their own lysine residues in a process termed autoubiquitination.

1.2 The challenge of substrate identification

A very peculiar characteristic of ubiquitination is that ubiquitin molecules, whether those conjugated to particular substrates, or free unbound ubiquitin, can act as substrates of

ubiquitination themselves, forming what are usually referred to as ubiquitin chains; in addition to its N-terminal amino group, ubiquitin has seven internal lysine residues (usually referred to by their position as K6, K11, K27, K29, K33, K48 and K63) the amino groups of all of which have been reported to be involved in chain formation via amide bonds with the ubiquitin C-terminal carboxyl groups[18-25]. Ubiquitin chains may be homogenous, in which all of the linkages are on the same lysine of ubiquitin, or mixed[26], where they are different; chains may also be straight or forked/branched[27], and may be formed on various different internal lysine residues in a particular substrate[28] **(Figure 1-2).**

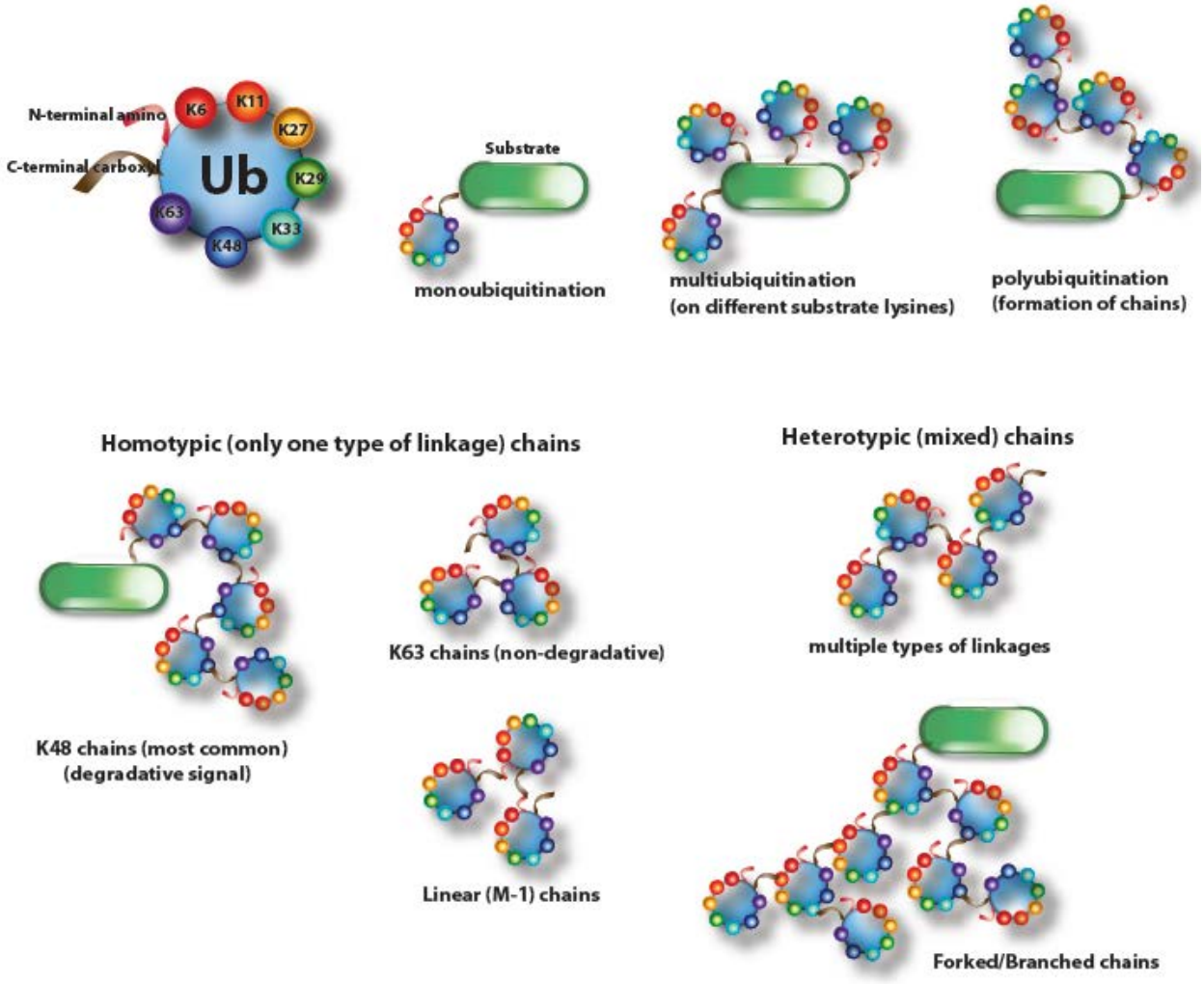

Figure 1-2 Different modes of ubiquitination and different chain types and linkages. Shown here are the three of the most common homotypic chains: K48, K63, and linear.

Different chain locations, linkages and topologies serve different functions as different signals inside the cell; K48 linked chains are by far the most well-known, and serve as a recognition site for the 26S proteasome, hence, proteins on which this of chain is attached is consigned to degradation[29]. Other types of chains may also function as proteasomal degradation signals as well, namely K11[30] and K29[22] chains. Indeed, a common indicator of the activity of E3s is the extent upon which their putative substrates are degraded[31]. Yet, other types of linkages, including mixed linkages, signify different things, from translocation[32], to stabilization[33] (by preventing degradation signals from being formed), to modulation of activity[34,35]. Additionally, substrates may only be labelled with a single ubiquitin residue—monoubiquitinated; while in many cases a certain substrate monoubiquitinated by an E3 enzyme may simply have the initial ubiquitin residue further extended into a chain conferring the appropriate biological signal by a fourth class of enzymes, termed E4s (ubiquitin chain elongation factors) some of which are also E3s, monoubiquitination has been observed to be a unique biological signal in its own right, as observed particularly in the case of histones[36-38].

While there are only a handful of known E1 enzymes characterized, there are dozens of known E2 enzymes, and at the time of writing it is estimated that over a thousand of E3s have been reported[39]. Naturally, the enzymes in this cascade are promiscuous; a particular enzyme must logically ubiquitinate a set consisting of multiple distinct targets; it is also established that these enzymes are somewhat functionally redundant[40], that is, these sets contain overlapping targets between different enzymes. Hence, the number of cascade of possible paths a ubiquitin molecule can travel along in order to reach a particular substrate is, conceivably, much greater than the even the number of proteins there are in the cell. Untangling this complicated web and

ascertaining which E1, E2, and E3s combinations are responsible for the ubiquitination of a particular substrate remains a daunting challenge **(Figure 1-3).**

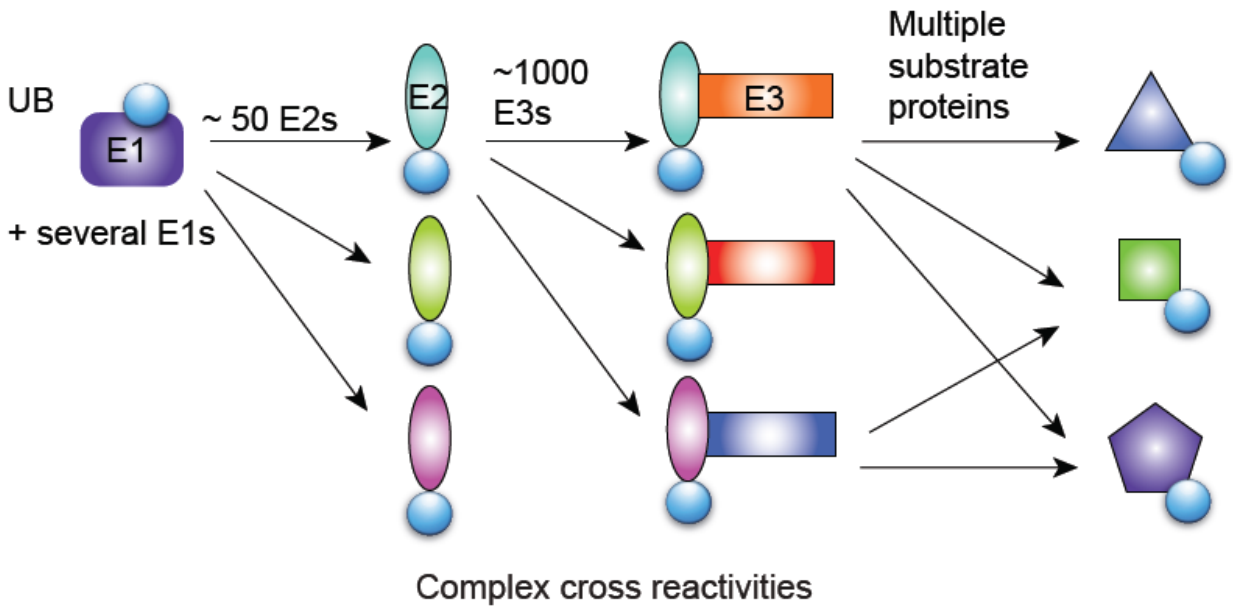

Figure 1-3. The presently characterized ubiquitin cascade by the numbers.
The vast number of ubiquitin-related enzymes in the cascade makes for a virtually limitless number of different possible paths for a ubiquitin molecule.

The sheer number of E3s in particular, have posed a challenge in ascertaining which E3s ubiquitinate a particular substrate. Also, while a particular substrate can be ubiquitinated by multiple E3s, and yet the different E3s may ubiquitinate that substrate on different residues and/or with different linkages; in other words, the specificity of ubiquitination sites and linkage types are determined by these manifold E3 enzymes[41], although it has been proposed that E2 enzymes may also play a part[42]. All in all, the identification of a particular E3's repertoire of substrates remains a challenge. Many proposed methods are in circulation; a handful of them rely on overexpression of a particular E3 *in vivo* with subsequent quantification of the modulation of stability of cellular proteins[43], making use of the fact that the majority of ubiquitination confers a

proteasomal degradation signal. Various other studies[44-48] utilize the fact that ubiquitination leaves a unique isopeptide bond between the ubiquitin's C-terminal carboxyl and the substrate's lysine ε-amino group; by overexpression of particular ubiquitin enzymes of interest, and then from the trypsinized lysate pulling down ubiquitinated substrates using an antibody which detects for that particular region (also known as anti-GlyGly antibody, one of the fragments from trypsin cleavage from ubiquitin's C-terminal would have two glycine residues, G75 and G76 remaining), ubiquitinated peptide fragments are enriched and characterized via subsequent mass spectrometry and identity of the protein elucidated by proteomics[49]; this approach is more thoroughly discussed in the final chapter. Still, other methods exist which attempt to screen for potential targets which contain domains that have high affinity for particular E3s; the yeast two-hybrid method is an example of such a screening technique[50]. The direct use of the ubiquitin transfer machinery in screening peptide libraries or microarrays *in vitro* has also been reported[51-53]. All of these methods, while effective in their own scope, have their own particular shortcomings and problems; there are many relevant facts to be considered: for example ubiquitination does not always confer a degradative signal, and ubiquitin enzymes. As mentioned earlier, E3s are functionally redundant, and overexpression of one may downregulate the activity of other E3s which act in a similar manner such that there is no observable overall change in the extent of the substrate's ubiquitination; affinity does not always translate to real biological activity; conditions *in vitro* may not always translate to actual activity *in vivo*, to name a few. The particular shortcomings of each method will not be discussed thoroughly in detail here; it suffices to say that there has been continuous interest and enthusiasm among researchers in the development of novel and effective strategies for E3 substrate identification.

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

Our lab has proposed, as reported in our previous research, the concept and development of an engineered artificial cascade of mutant enzymes which we have termed orthogonal ubiquitin transfer (OUT)[54] as a method of untangling the extremely complicated network of ubiquitin enzymes. Our goal was to construct a distinct and separate pathway in which an engineered variant of ubiquitin, travels through only one possible path consisting of engineered E1, E2, and E3 enzymes; all of the enzymes involved being engineered as such that they are exclusively active with only its engineered partner, and inactive to their wild-type partners (hence, orthogonal), save for the final substrate-binding domain of E3 enzymes, which are retained as such to identify native substrates of the whole cascade **(Figure 1-4).**

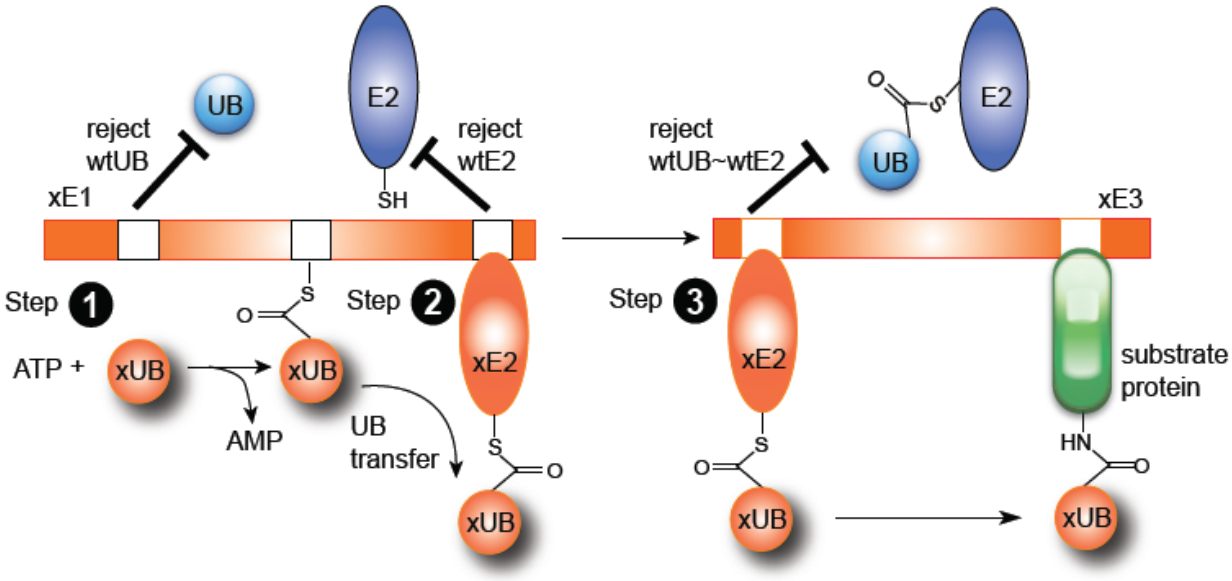

Figure 1-4 The OUT (orthogonal ubiquitin transfer) cascade
Engineered orthogonal enzymes reject their wild-type counterparts, hence forming an exclusive cascade through which an engineered ubiquitin variant travels to arrive at its destination.

Engineering the whole cascade basically comes down to (re)engineering three pairs of protein interface interactions, those between ubiquitin and E1, between E1 and E2, and between E2 and E3. In each case, we sought to come up with an interacting interface which abolishes the enzymes' activity with their wild-type partners, while retaining exclusive activity of the two partners with each other. This has generally been accomplished by the so-called bump-and-hole strategy[55] **(Figure 1-5),** which involves reversing the charge of crucial residues in the native interaction, such that a once-existing attractive electrostatic interaction is changed to a repulsive interaction instead. The successful engineering of a reversed-charge interface has previously been successfully demonstrated in several systems[56-58], of particular note and relevance is the preexisting work on enzymes in the ubiquitin pathway reported by Winkler and Timmers[59], involving the swap of critical charged residues between ubch5b (an E2 enzyme) and CNOT4 (a RING-type E3); when the charge of a critical lysine residue in ubch5b was reversed, its activity with wild-type CNOT4 was abrogated; however, the activity was restored by complementary reversal of the charges of crucial aspartate and glutamate residues in CNOT4.

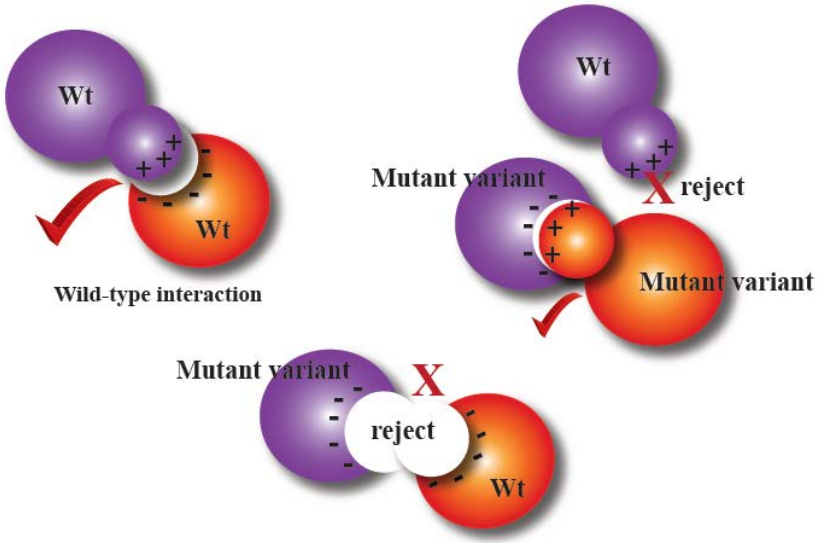

Figure 1-5 The bump-and-hole strategy
Creating charge reversal mutations of critical residues is one direct way of removing original wild-type activity by electrostatic repulsion, while simultaneously creating an orthogonal and analogous interaction by electrostatic attraction; similar results could conceivably also be achieved utilizing steric effects or hydrophobic and hydrophilic interactions instead.

In the same vein, our lab had attempted to engineer a charge-reversed ubiquitin and E1 interaction as the first step of our cascade. We relied mainly on a crystal structure showing ubiquitin in complex with the E1 enzyme uba1 reported by Lee and Schindelin[60], as the basis for our OUT engineering. Through rational-design mutagenesis in the E1 enzyme and phage-display screening for reactive ubiquitin partners, as well as further mutagenesis to enhance OUT activity as well as to abrogate any residual remaining wild-type cross reactivity, we ultimately succeeded in developing a completely orthogonal pair of UB-E1 which involves two charge reversal mutations of critical ubiquitin residues namely, R42E and R72E, and complementary mutations of negative or neutral residues, those located in what is called the adenylation domain of Uba1 enzyme, to positive residues, specifically Q576R, S589R, and D591R[54] **(Figure 1-6).**

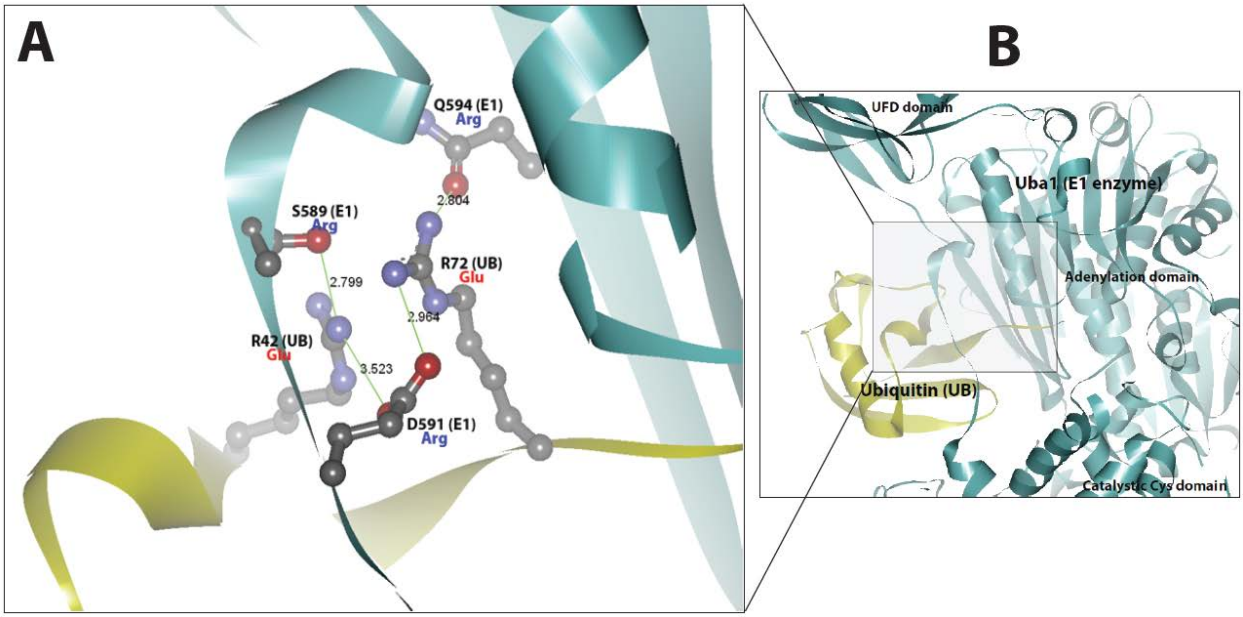

Figure 1-6 The ubiquitin-E1 interface. (A) A close-up view of the targeted residues in the adenylation domain of the Uba1 E1 enzyme (cyan), S589, D591 and Q594, which were mutated to arginine (blue text) in xE1; to match these mutations R72 and R42 of UB (yellow) were mutated to Glu (red text). The potential interactions between the residues (green line) are as shown with the distances between the appropriate atoms given in angstroms. (B) The UB-Uba1 structure showing ubiquitin (yellow) with the various domains of Uba1 (cyan), the close-up view shown in (A) is that of the shaded area (grey rectangle).

When these mutations were made on analogous residues in the adenylation domain of other E1 enzymes, namely, Ube1, the homolog of the yeast Uba1, and Uba6, a fundamentally different enzyme altogether, the resulting mutant E1s also displayed orthogonal activity (data awaiting publication). These engineered proteins had been given the name and commonly referred to in our previous communications as xE1 (and the R42E R72E double mutant ubiquitin as xUB)—the symbol "x" consisting of two perpendicular lines being used to denote orthogonality; xUB is inactive against wild-type E1 enzymes, but is activated by xE1, and likewise xE1 is inactive with wild-type ubiquitin but can activate xUB.

We employed a similar strategy in the engineering of an xE1-xE2 pair. At the time of our investigation, the negatively charged residues E1004, D1014 and E1016 in Uba1 had been previously reported to be crucial for interaction with E2[60]. As expected, charge reversal mutations of these residues to Lys diminish activity with the native E2 class of enzymes.

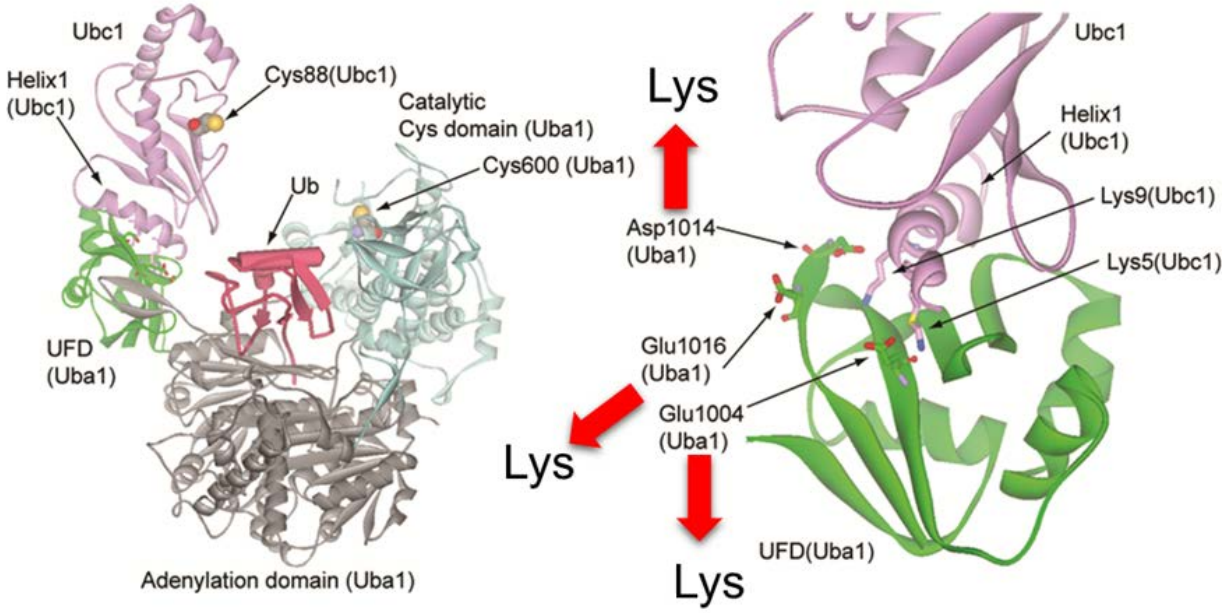

Figure 1-7 The E1-E2 interface. In contrast to the E1 adenylation domain (gray), and the catalytic cysteine domain (light blue), which interact with ubiquitin, interaction with E2 (lilac) involves the ubiquitin fold (UFD) domain (green) further to the C-terminal of the E1 enzyme, which is laced with acidic residues; these residues were mutated to Lys to disrupt the wild-type binding in our OUT system.

These mutated residues **(Figure 1-7)** are located in what is termed the ubiquitin fold domain (UFD) of E1 which lies further to the C-terminus than the first set of mutations in the adenylation domain (AD), hence, they are sometimes referred to as "back mutations", whereas the mutations which interact with xUB are referred to as "front mutations"; hence, E1 variants containing their respective sets of mutations are referred to as xE1(UFD) or (bE1), and xE1(AD) (fE1); combining the two sets of mutations together completes the full E1 variant which we intend to use in OUT, which is completely inactive against both wild-type ubiquitin and wild-type E2; this combined mutant is what is referred sometimes as (3+3)E1, (f+b)E1, or xE1(full), but most regularly to as simply xE1 (see appendix for a complete list of mutants and their respective terminologies). The terms bE1, fE1, and xE1 shall be regularly used in future chapters of this communication (**Figure 1-8**).

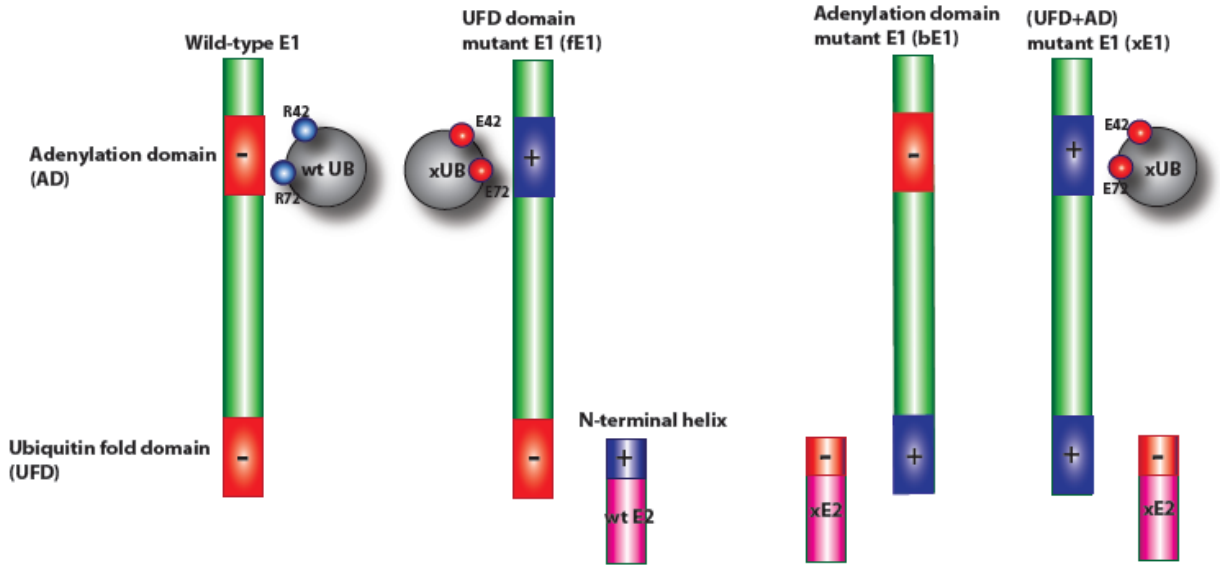

Figure 1-8 Summary of E1 variants. Charge reversals in the adenylation and ubiquitin fold domains of E1 (referred to as AD and UFD, respectively) abrogate preference for wild-type enzymes and establish activity with OUT variants of UB and E2.

The N-terminal helix of E2 enzymes have long been determined to be responsible for interaction with E1; this region comprises of a streak of positively charged residues as shown **(Figure 1-7)**, which interact with the acidic residues in the E1 UFD domain mentioned previously; obviously the favorable electrostatic salt bridge interactions would be replaced by an unfavorable electrostatic repulsion when bE1 is used instead. Conceivably, one might imagine that reversing the charges of Lys and Arg residues in the E2 N-terminal helix would restore activity. In a similar manner as the development of xUB, we screened a library consisting of phage particles displaying N-terminal randomized variants of the E2 enzyme Ubc1 for mutants with activity with bE1, the methodology of which will be discussed in the third chapter where a similar approach was taken to engineer the E3 enzyme for OUT activity.

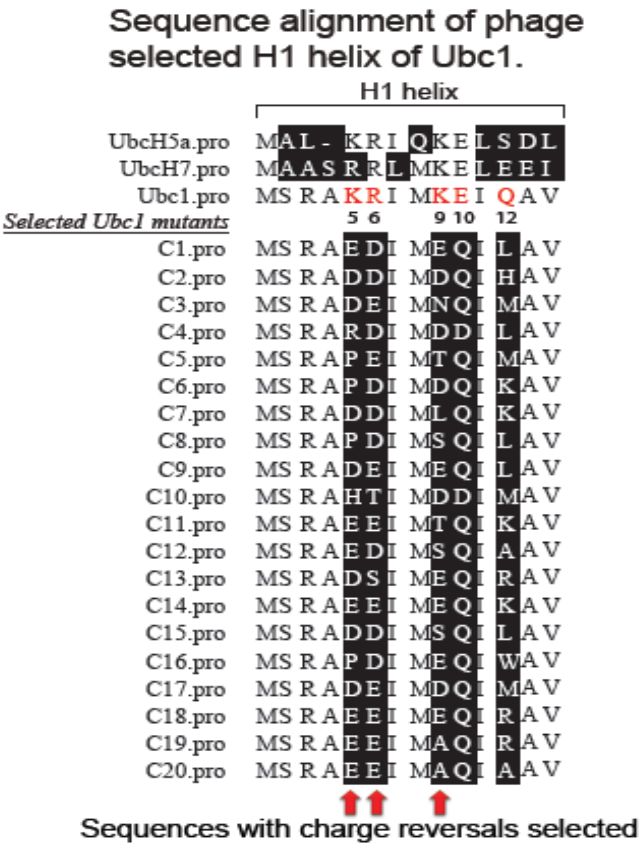

Figure 1-9

Sequence of selected E2 mutants

Five residues, including K5, R6, K9, E10 and Q12 in the yeast E2 ubc1 N-terminal helix were randomly mutated in a library pool of E2 variants displayed on phage particles; the mutants were then screened and enriched for activity with bE1. The sequences of the surviving mutants show a very strong consensus, with apparent charge reversal mutations being favored at positions 5, 6, and 9 (red arrows), consistent with our initial predictions.

The consensus of selected Ubc1 N-terminal sequences from our phage library demonstrates a clear preference for the charge reversals of the positive residues K5, R6 and K9 **(Figure 1-9)**—all consistent with what one would expect for favorable interaction with the now positively-charged UFD domain surface of bE1. The several selected mutants which were tested proved active both with bE1 and wild-type ubiquitin, per the selection conditions, and also with xE1 and xUB, as would be required in a complete OUT scenario—the overall activity of xE2 variants were seemingly unaffected as a whole by mutations in the E1 adenylation domain. Additionally, since most E2s generally have very similar structures and amino acid sequences, making analogous mutations in the N-terminal helices of other E2 enzymes, namely ubch7 and ubch5, proved successful in creating active xE2 variants of those enzymes[54].

1.4 Discussion and future directions

Our investigation has yielded active and orthogonal xub, xE1, and xE2 variants, which may prove useful in the identification of particular E3 substrates once the complete OUT cascade is constructed. We have utilized a combination of both ration design mutagenesis and activity-based screening of phage libraries to achieve our purpose. Moreover, we have shown that the mutations we designed and developed are not limited in their usefulness and applicability to merely the specific E1 and E2 we used in our studies, but are fully transferable to a variety of different E1 and E2 enzymes, as long as their general structures and modes of binding are somewhat similar; this theoretically allows us to target a much wider assortment of E3s, and consequently, potential substrates with OUT. Obviously, the next and final step was to develop an xE2-xE3 pair which would complete the OUT cascade and allow xUB to reach its final destination on the substrate lysine; this was the crux of my research with the Yin group. An

approach similar to the ones we had used for our work with the xub-xE1, and xE1-xE2 interfaces was initially adopted; however, throughout our whole endeavor, we discovered that the E2-E3 interface possessed many particular idiosyncrasies and fundamental differences with the other two interfaces; this made our task a considerable challenge in its own right. The details of our efforts will be discussed in the following chapters.

1.5 References

1 Hunter, T. The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Molecular cell* **28**, 730-738, doi:10.1016/j.molcel.2007.11.019 (2007).

2 Ciechanover, A. Proteolysis: from the lysosome to ubiquitin and the proteasome. *Nature reviews. Molecular cell biology* **6**, 79-87, doi:10.1038/nrm1552 (2005).

3 Amm, I., Sommer, T. & Wolf, D. H. Protein quality control and elimination of protein waste: the role of the ubiquitin-proteasome system. *Biochimica et biophysica acta* **1843**, 182-196, doi:10.1016/j.bbamcr.2013.06.031 (2014).

4 Reinstein, E. & Ciechanover, A. Narrative review: protein degradation and human diseases: the ubiquitin connection. *Annals of internal medicine* **145**, 676-684 (2006).

5 Nanduri, P., Hao, R., Fitzpatrick, T. & Yao, T. P. Chaperone-mediated 26S proteasome remodeling facilitates free K63 ubiquitin chain production and aggresome clearance. *The Journal of biological chemistry* **290**, 9455-9464, doi:10.1074/jbc.M114.627950 (2015).

6 Erpapazoglou, Z., Walker, O. & Haguenauer-Tsapis, R. Versatile roles of k63-linked ubiquitin chains in trafficking. *Cells* **3**, 1027-1088, doi:10.3390/cells3041027 (2014).

7 Iwai, K., Fujita, H. & Sasaki, Y. Linear ubiquitin chains: NF-kappaB signalling, cell death and beyond. *Nature reviews. Molecular cell biology* **15**, 503-508, doi:10.1038/nrm3836 (2014).

8 Ramanathan, H. N., Zhang, G. & Ye, Y. Monoubiquitination of EEA1 regulates endosome fusion and trafficking. *Cell & bioscience* **3**, 24, doi:10.1186/2045-3701-3-24 (2013).

9 Gatti, M. *et al.* A novel ubiquitin mark at the N-terminal tail of histone H2As targeted by RNF168 ubiquitin ligase. *Cell cycle (Georgetown, Tex.)* **11**, 2538-2544, doi:10.4161/cc.20919 (2012).

10 Wu, K. *et al.* Mono-ubiquitination drives nuclear export of the human DCN1-like protein hDCNL1. *The Journal of biological chemistry* **286**, 34060-34070, doi:10.1074/jbc.M111.273045 (2011).

11 Ciechanover, A. The ubiquitin-proteasome pathway: on protein death and cell life. *The EMBO journal* **17**, 7151-7160, doi:10.1093/emboj/17.24.7151 (1998).

12 Metzger, M. B., Hristova, V. A. & Weissman, A. M. HECT and RING finger families of E3 ubiquitin ligases at a glance. *Journal of cell science* **125**, 531-537, doi:10.1242/jcs.091777 (2012).

13 Vittal, V. *et al.* Intrinsic disorder drives N-terminal ubiquitination by Ube2w. *Nature chemical biology* **11**, 83-89, doi:10.1038/nchembio.1700 (2015).

14 Coulombe, P., Rodier, G., Bonneil, E., Thibault, P. & Meloche, S. N-Terminal ubiquitination of extracellular signal-regulated kinase 3 and p21 directs their degradation by the proteasome. *Molecular and cellular biology* **24**, 6140-6150, doi:10.1128/mcb.24.14.6140-6150.2004 (2004).

15 Schumacher, F. R., Wilson, G. & Day, C. L. The N-terminal extension of UBE2E ubiquitin-conjugating enzymes limits chain assembly. *Journal of molecular biology* **425**, 4099-4111, doi:10.1016/j.jmb.2013.06.039 (2013).

16 Nuber, U., Schwarz, S. E. & Scheffner, M. The ubiquitin-protein ligase E6-associated protein (E6-AP) serves as its own substrate. *European journal of biochemistry / FEBS* **254**, 643-649 (1998).

17 Amemiya, Y., Azmi, P. & Seth, A. Autoubiquitination of BCA2 RING E3 ligase regulates its own stability and affects cell migration. *Molecular cancer research : MCR* **6**, 1385-1396, doi:10.1158/1541-7786.mcr-08-0094 (2008).

18 Walczak, H., Iwai, K. & Dikic, I. Generation and physiological roles of linear ubiquitin chains. *BMC biology* **10**, 23, doi:10.1186/1741-7007-10-23 (2012).

19 Wu-Baer, F., Lagrazon, K., Yuan, W. & Baer, R. The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin. *The Journal of biological chemistry* **278**, 34743-34746, doi:10.1074/jbc.C300249200 (2003).

20 Rape, M. Assembly of k11-linked ubiquitin chains by the anaphase-promoting complex. *Sub-cellular biochemistry* **54**, 107-115, doi:10.1007/978-1-4419-6676-6_9 (2010).

21 Birsa, N. *et al.* Lysine 27 ubiquitination of the mitochondrial transport protein Miro is dependent on serine 65 of the Parkin ubiquitin ligase. *The Journal of biological chemistry* **289**, 14569-14582, doi:10.1074/jbc.M114.563031 (2014).

22 Kristariyanto, Y. A. *et al.* K29-selective ubiquitin binding domain reveals structural basis of specificity and heterotypic nature of k29 polyubiquitin. *Molecular cell* **58**, 83-94, doi:10.1016/j.molcel.2015.01.041 (2015).

23 Michel, M. A. *et al.* Assembly and specific recognition of k29- and k33-linked polyubiquitin. *Molecular cell* **58**, 95-109, doi:10.1016/j.molcel.2015.01.042 (2015).

24 Petroski, M. D. & Deshaies, R. J. Mechanism of lysine 48-linked ubiquitin-chain synthesis by the cullin-RING ubiquitin-ligase complex SCF-Cdc34. *Cell* **123**, 1107-1120, doi:10.1016/j.cell.2005.09.033 (2005).

25 Branigan, E., Plechanovova, A., Jaffray, E. G., Naismith, J. H. & Hay, R. T. Structural basis for the RING-catalyzed synthesis of K63-linked ubiquitin chains. *Nature structural & molecular biology* **22**, 597-602, doi:10.1038/nsmb.3052 (2015).

26 Nakasone, M. A., Livnat-Levanon, N., Glickman, M. H., Cohen, R. E. & Fushman, D. Mixed-linkage ubiquitin chains send mixed messages. *Structure (London, England : 1993)* **21**, 727-740, doi:10.1016/j.str.2013.02.019 (2013).

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

28 Yin, H., Gui, Y., Du, G., Frohman, M. A. & Zheng, X. L. Dependence of phospholipase D1 multi-monoubiquitination on its enzymatic activity and palmitoylation. *The Journal of biological chemistry* **285**, 13580-13588, doi:10.1074/jbc.M109.046359 (2010).

29 Mallette, F. A. & Richard, S. K48-linked ubiquitination and protein degradation regulate 53BP1 recruitment at DNA damage sites. *Cell research* **22**, 1221-1223, doi:10.1038/cr.2012.58 (2012).

30 Wickliffe, K. E., Williamson, A., Meyer, H. J., Kelly, A. & Rape, M. K11-linked ubiquitin chains as novel regulators of cell division. *Trends in cell biology* **21**, 656-663, doi:10.1016/j.tcb.2011.08.008 (2011).

31 Melvin, A. T., Woss, G. S., Park, J. H., Waters, M. L. & Allbritton, N. L. Measuring activity in the ubiquitin-proteasome system: from large scale discoveries to single cells analysis. *Cell biochemistry and biophysics* **67**, 75-89, doi:10.1007/s12013-013-9621-9 (2013).

32 Erpapazoglou, Z. *et al.* A dual role for K63-linked ubiquitin chains in multivesicular body biogenesis and cargo sorting. *Molecular biology of the cell* **23**, 2170-2183, doi:10.1091/mbc.E11-10-0891 (2012).

33 Rodrigues, L., Popov, N., Kaye, K. M. & Simas, J. P. Stabilization of Myc through heterotypic poly-ubiquitination by mLANA is critical for gamma-herpesvirus lymphoproliferation. *PLoS pathogens* **9**, e1003554, doi:10.1371/journal.ppat.1003554 (2013).

34 Ranaweera, R. S. & Yang, X. Auto-ubiquitination of Mdm2 enhances its substrate ubiquitin ligase activity. *The Journal of biological chemistry* **288**, 18939-18946, doi:10.1074/jbc.M113.454470 (2013).

35 Pettersson, S. *et al.* Non-degradative ubiquitination of the Notch1 receptor by the E3 ligase MDM2 activates the Notch signalling pathway. *Biochem J* **450**, 523-536, doi:10.1042/bj20121249 (2013).

36 Turco, E., Gallego, L. D., Schneider, M. & Kohler, A. Monoubiquitination of histone H2B is intrinsic to the Bre1 RING domain-Rad6 interaction and augmented by a second Rad6-binding site on Bre1. *The Journal of biological chemistry* **290**, 5298-5310, doi:10.1074/jbc.M114.626788 (2015).

37 Koegl, M. *et al.* A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. *Cell* **96**, 635-644 (1999).

38 Nordquist, K. A. *et al.* Structural and functional characterization of the monomeric U-box domain from E4B. *Biochemistry* **49**, 347-355, doi:10.1021/bi901620v (2010).

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

40 Morishima, Y. *et al.* CHIP deletion reveals functional redundancy of E3 ligases in promoting degradation of both signaling proteins and expanded glutamine proteins. *Human molecular genetics* **17**, 3942-3952, doi:10.1093/hmg/ddn296 (2008).

41 David, Y. *et al.* E3 ligases determine ubiquitination site and conjugate type by enforcing specificity on E2 enzymes. *The Journal of biological chemistry* **286**, 44104-44115, doi:10.1074/jbc.M111.234559 (2011).

42 David, Y., Ziv, T., Admon, A. & Navon, A. The E2 ubiquitin-conjugating enzymes direct polyubiquitination to preferred lysines. *The Journal of biological chemistry* **285**, 8595-8604, doi:10.1074/jbc.M109.089003 (2010).

43 Emanuele, M. J. *et al.* Global identification of modular cullin-RING ligase substrates. *Cell* **147**, 459-474, doi:10.1016/j.cell.2011.09.019 (2011).

44 Thompson, J. W. *et al.* Quantitative Lys--Gly-Gly (diGly) proteomics coupled with inducible RNAi reveals ubiquitin-mediated proteolysis of DNA damage-inducible transcript 4 (DDIT4) by the E3 ligase HUWE1. *The Journal of biological chemistry* **289**, 28942-28955, doi:10.1074/jbc.M114.573352 (2014).

45 Wagner, S. A. *et al.* A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. *Mol Cell Proteomics* **10**, M111.013284, doi:10.1074/mcp.M111.013284 (2011).

46 Udeshi, N. D. *et al.* Methods for quantification of in vivo changes in protein ubiquitination following proteasome and deubiquitinase inhibition. *Mol Cell Proteomics* **11**, 148-159, doi:10.1074/mcp.M111.016857 (2012).

47 Xu, G., Paige, J. S. & Jaffrey, S. R. Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nature biotechnology* **28**, 868-873, doi:10.1038/nbt.1654 (2010).

48 Kim, W. *et al.* Systematic and quantitative assessment of the ubiquitin-modified proteome. *Molecular cell* **44**, 325-340, doi:10.1016/j.molcel.2011.08.025 (2011).

49 Beaudette, P., Popp, O. & Dittmar, G. Proteomic techniques to probe the ubiquitin landscape. *Proteomics* **16**, 273-287, doi:10.1002/pmic.201500290 (2016).

50 Guo, Z. *et al.* Proteomics strategy to identify substrates of LNX, a PDZ domain-containing E3 ubiquitin ligase. *J Proteome Res* **11**, 4847-4862, doi:10.1021/pr300674c (2012).

51 Andrews, P. S. *et al.* Identification of substrates of SMURF1 ubiquitin ligase activity utilizing protein microarrays. *Assay Drug Dev Technol* **8**, 471-487, doi:10.1089/adt.2009.0264 (2010).

52 Kenten, J. H. *et al.* Assays for high-throughput screening of E2 AND E3 ubiquitin ligases. *Methods in enzymology* **399**, 682-701, doi:10.1016/s0076-6879(05)99045-9 (2005).

53 Ayad, N. G., Rankin, S., Ooi, D., Rape, M. & Kirschner, M. W. Identification of ubiquitin ligase substrates by in vitro expression cloning. *Methods in enzymology* **399**, 404-414, doi:10.1016/s0076-6879(05)99028-9 (2005).

54 Zhao, B. *et al.* Orthogonal ubiquitin transfer through engineered E1-E2 cascades for protein ubiquitination. *Chem Biol* **19**, 1265-1277, doi:10.1016/j.chembiol.2012.07.023 (2012).

55 Bishop, A. *et al.* Unnatural ligands for engineered proteins: new tools for chemical genetics. *Annual review of biophysics and biomolecular structure* **29**, 577-606, doi:10.1146/annurev.biophys.29.1.577 (2000).

56 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 of the United States of America* **94**, 3565-3570 (1997).

57 Liu, Y., Shah, K., Yang, F., Witucki, L. & Shokat, K. M. Engineering Src family protein kinases with unnatural nucleotide specificity. *Chem Biol* **5**, 91-101 (1998).

58 Hwang, Y. W. & Miller, D. L. A mutation that alters the nucleotide specificity of elongation factor Tu, a GTP regulatory protein. *The Journal of biological chemistry* **262**, 13081-13085 (1987).

59 Winkler, G. S. & Timmers, H. T. M. Structure-Based Approaches to Create New E2–E3 Enzyme Pairs. **399**, 355-366, doi:10.1016/s0076-6879(05)99024-1 (2005).

60 Lee, I. & Schindelin, H. Structural insights into E1-catalyzed ubiquitin activation and transfer to conjugating enzymes. *Cell* **134**, 268-278, doi:10.1016/j.cell.2008.05.046 (2008).

Chapter 2

Construction of xE3 through rational design mutagenesis

2.1 Background information

2.1.1 Rationale

As stated in the introductory chapter, the next step in the development of OUT was to develop the xE2-xE3 interface—to this end, we had initially considered using the same bump-and-hole strategy as we had done earlier with the xUB-xE1 and xE1-xE2 interfaces[1]; the first step of which was to identify key residues involved in the binding, and hence we started by first examining existing studies and crystal structures (more details in 2.1.2). Upon examination of such published structures, we found that the N-terminal helix region of the E2 enzyme also plays a big role in the interaction with E3s—we had made several charge-reversal mutations of positive residues in this region to interact with the C-terminal mutations on xE1; indeed, the xE2 enzyme variant developed through phage selection was found to be completely inactive with all the tested wild-type E3 enzymes **(Figure 2-1).**

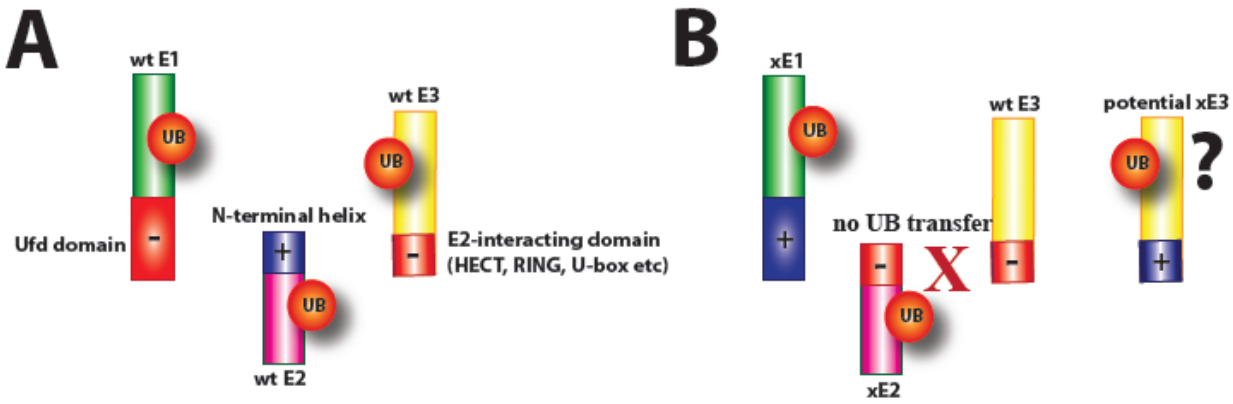

Figure 2-1 Rational design of xE3. The charge reversal mutations (A) in the N-terminal helix region of xE2 has arrested its activity with all tested wild-type E3 enzymes (B). Hence, our efforts centered on engineering a similarly charge-reversed xE3 to match those mutations.

Hence, in other words, for the xE2-xE3 interface we needed not come up with a novel method of breaking up the wild-type interaction as we had needed to in the previous xub-xE1 and xE1-xE2 steps—the mutations in E2 complementary to XE1 had already fulfilled that requirement; our challenge was simply this: to restore the already broken interface via complementary mutations in the E3 enzyme. Our initial efforts through rational design mutagenesis comprises the crux of this chapter.

2.1.2 The RING E3 family

For the development of xE3, we focused primarily on the RING (really interesting new gene) family of E3s, which comprises a vast majority, ~80-90% of all characterized E3 enzymes[2,3]; its members are involved in regulation of numerous biological pathways[4-12], and are involved in the pathogenesis or suppression of many different diseases[13-22]. These E3s are wildly divergent as to their amino acid sequences and domain compositions, but all members share the remarkably well-conserved RING domain which is responsible for E3 ligase activity[23].

The general structure of RING domains are as given in the following examples provided by Brzovic et.al.[24] **(Figure 2-2)**. RING domains contain a pair of interleaved zinc finger domains—zinc ions being coordinated by certain Cys and His residues (and in extremely rare cases, Asp[25])[26]; this is a central defining feature of the domain, upon which these enzymes can be categorized into various subfamilies—based on their zinc coordination patterns; it is this zinc coordination which is responsible for maintaining structural integrity and E3 function of RING domains. The vast majority of RINGs are known to be active as either heterodimers or homodimers[27]—examples of such are the breast-cancer related BRCA1/BARD1 heterodimer[24], or the RAG1 homodimer[28] (**Figures 2-2** and **2-3**) ; although RING domains which are believed to be active as monomers[29] also exist.

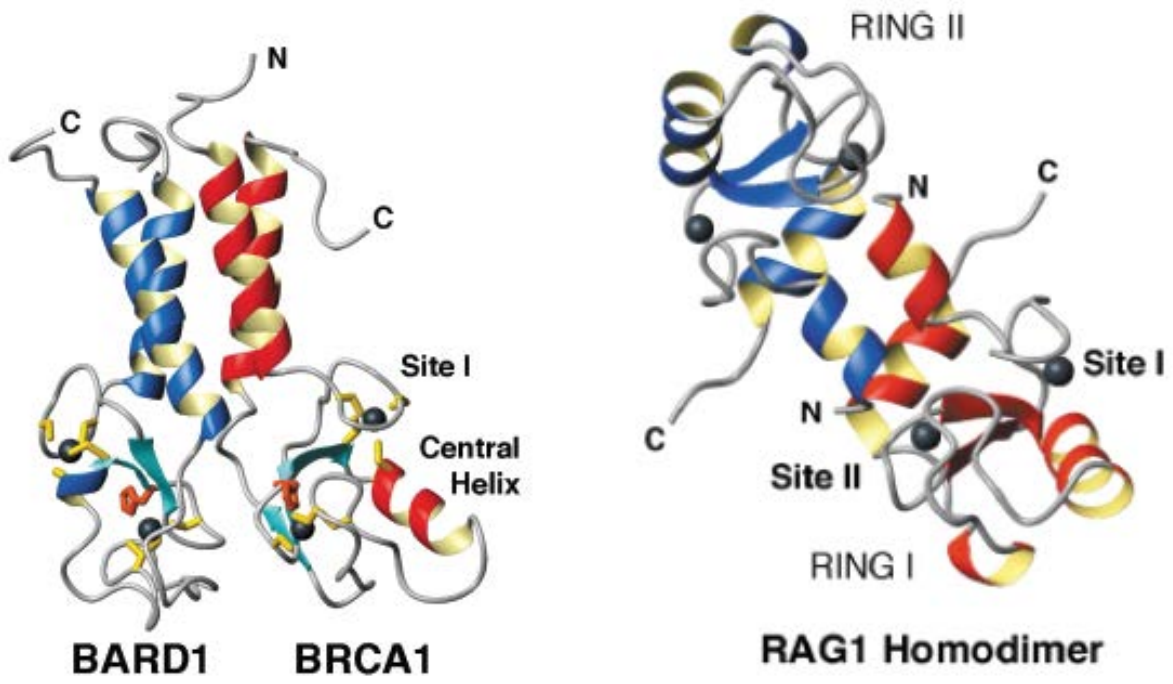

Figure 2-2 Example structures of RING dimers.
Shown here are structures of the BRCA1/BARD1 heterodimer and the RAG1 homodimer[24]; note the analogous helical domains responsible for dimerization, and the presence of a pair of zinc ions (black spheres) in each monomer, which comprises the core of the zinc fingers that confer structural stability. Site I represents the loop which plays a part in interacting with the E2 via the E2 N-terminal helix, and is critical to our mutagenesis (see further discussion).

As stated earlier, RING E3s are generally thought to function as bridges which bring the E2~UB conjugate and the substrate into proximity, and to activate the E2~ub thioester for release[30] onto the appropriate lysine residue. They typically do not load ubiquitin onto themselves, as in the case of HECT type E3s, but in the absence of substrate they are known to autoubiquitinate[31-38]—to put ubiquitin molecules on their own lysine residues. They also are believed to be responsible for the determination of the type and topology of the ubiquitin chains to be formed[39]; experimental evidence from kinetic studies supports the idea that the E3 transfers ubiquitin molecules in a one-by-one manner from the E2 enzyme onto the substrate[40,41],

elongating the polyubiquitin chain residue by residue, each time involving the chain extension point being further and further away, (and hence it is held that the E2-E3 complex must possess an extraordinary amount of flexibility, as they have been known to successfully ubiquitinate lysine residues across distances up to over 100Å[42]; the source of such flexibility being still poorly understood) while other competing, less popular views have been proposed, such as one involving the formation of the whole chain while it is still tethered to the E2 cysteine, and then the whole chain being transferred to the substrate[43] *en bloc*.

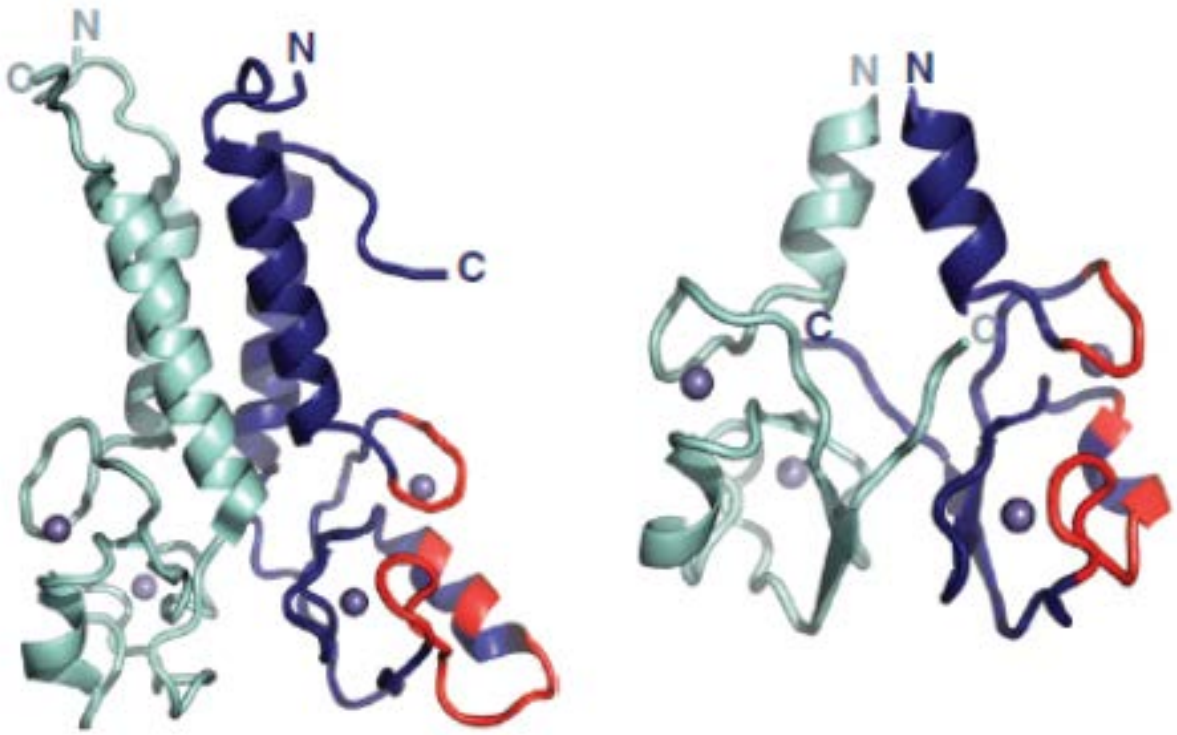

Figure 2-3 E2-interacting regions in RING domains. The structures of the BRCA1 (as a heterodimer with BARD1, left) and the BIRC7 homodimer (right), with their analogous E2-interacting regions highlighted in red. In general, in the case of heterodimers, only one dimer is active as a RING and interacts with E2 enzymes; however, in the case of homodimers, each of the dimers is autonomously capable of interacting with its own E2.

The interaction between E2 enzymes and RING domains, as shown in the structure provided by Metzger et.al.[27], involves a large interface comprised of two swaths of residues across the surface of the RINGs—the upper loop is the region which mainly interacts with the N-

terminal helix of E2, while the larger, lower stretches of residues interacts with other portions of the E2 enzyme **(Figure 2-3).**

2.2 Results

2.2.1 Engineering xE2-xE3: general considerations

It was clear that our focus had to be on the residues in E3 which interacted with the N-terminal helix; we had not touched any residues in any other domains in xE2, and to this end, we turned to a published crystal structure between Ubch5, an E2 enzyme and the RING E3 cIAP[44], which, at the time of investigation was to our knowledge the only one of its kind in circulation; many additional E2-E3 structures have been made available since then[45-48]. The Ubch5-cIAP structure, showing the residues most likely interacting with the basic residues, which were mutated to acidic residues in xE2, in the N-terminal helix, is as shown **(Figure 2-4).**

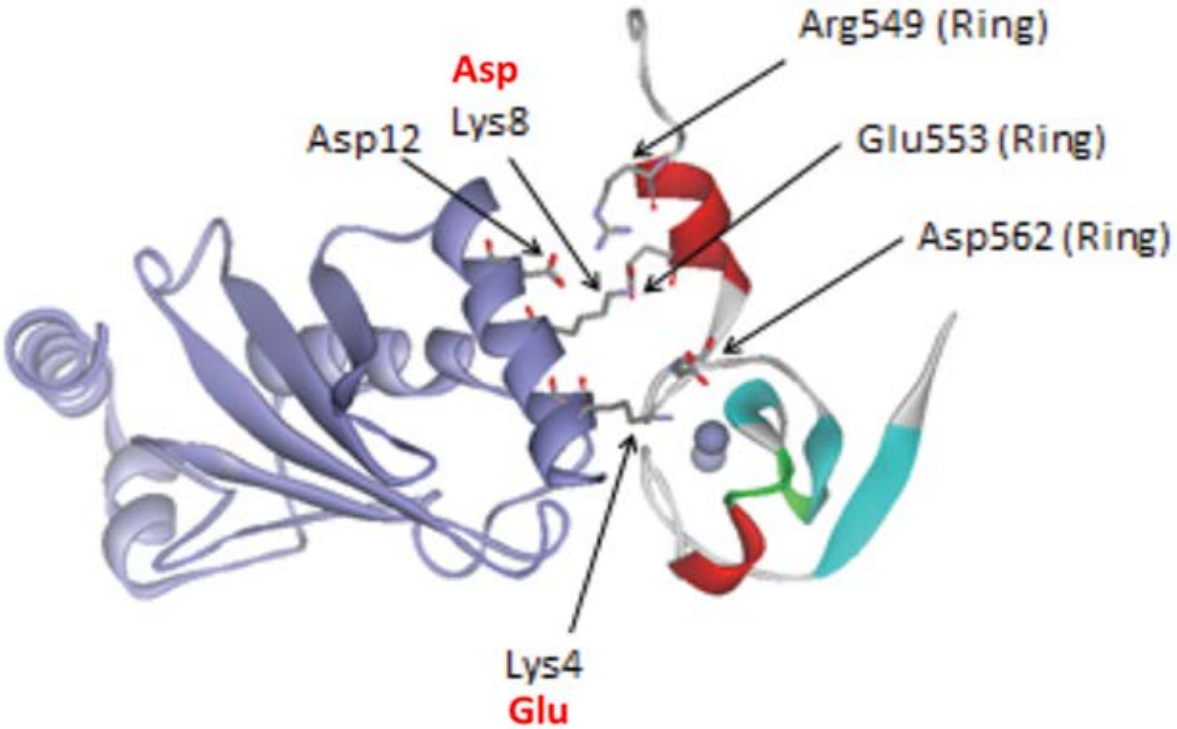

Figure 2-4 The ubch5b-cIAP RING complex. The E2 enzyme (left, purple) with K4 and K8 in the N-terminal helix, which were charge-reversed in xE2 (red text), shown interacting with the acidic residues E553 and D562 in the cIAP RING (right).

The potential interactions between K4 and D562, and between K8 and E553 were immediately obvious; since both K4 and K8 were mutated to glutamate and aspartate residues, respectively, we hence focused our efforts on making complementary charge reversal mutations in these relevant acidic residues in the RING domains.

2.2.2 Alignment of RINGs and mutagenesis of key acidic residues

Given the structural similarity among the members of the RING family as far as the RING domain was concerned, we identified residues which would most likely be equivalent to E553 and D562 in cIAP; the consensus of the alignment was readily obvious; the zinc-coordinating cysteine residues which were perfectly conserved among the RINGs served as an effective guide (**Figure 2-5**).

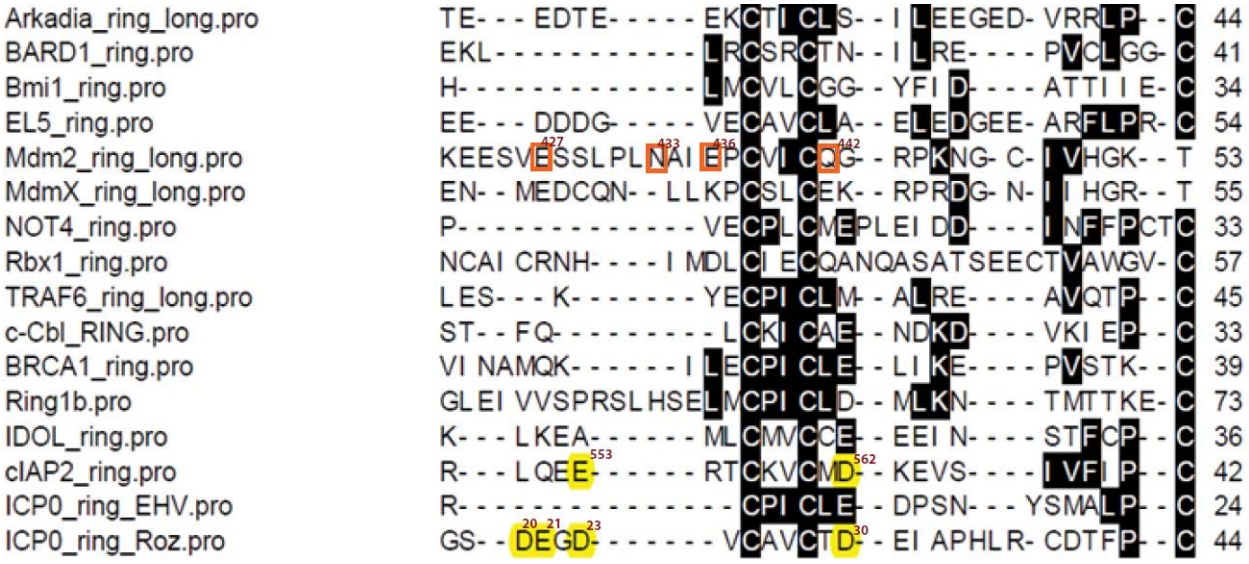

Figure 2-5 Alignment of RING amino acid sequences
The amino acid sequences of various RING domains were aligned, with many of the highly conserved residues (shaded in black) being zinc-coordinating cysteines. E553 and D562 (shaded in yellow, upper row) in cIAP being aligned with potentially analogous residues in the ICP0 RING (yellow shaded, lower row), namely (from left to right): D20, E21, D23 and D30. In the same manner, albeit less convincingly, in the Mdm2 RING domain one of either E427, N433 or E436 would probably align with D553, while Q442 aligned with D562; the aligned Mdm2 residues are boxed in orange.

From the alignment, we determined that ICP0 possessed the greatest level of sequence homology with the cIAP RING; D30 of the ICP0 RING most likely corresponded to D562, and that one of D20, E21 or D23 would most likely correspond to E553. We hence decided to try one single mutant and three double ICP0 RING mutants in which these acidic residues were mutated to lysine, as given below (**Figure 2-6**).

	20	21	22	23	24	25	26	27	28	29	30
Wt ICP0	D	E	G	D	V	C	A	V	C	T	D
ICP0 11	D	E	G	D	V	C	A	V	C	T	K
ICP0 21	K	E	G	D	V	C	A	V	C	T	K
ICP0 22	D	K	G	D	V	C	A	V	C	T	K
ICP0 23	D	E	G	K	V	C	A	V	C	T	K

Figure 2-6 ICP0 RING mutants

Alignment of the wild-type ICP0 amino acid sequence (blue text) with those of rationally-designed mutant RING domains, showing mutated positions (shaded in yellow), and mutated residues (red text).

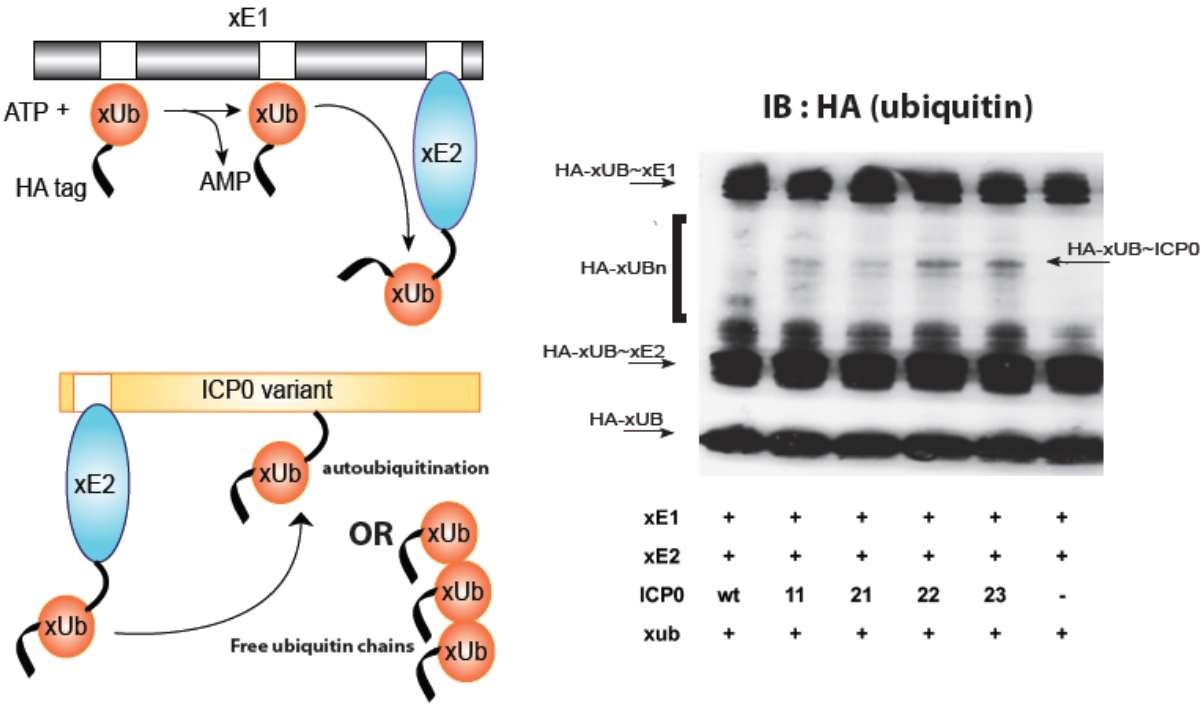

Figure 2-7 Ubiquitination assay of ICP0 mutants

Wild-type ICP0 RING, and engineered variants were charged with OUT enzymes xE1, xE2 and xUB, and monitored for either xub-ICP0 autoubiquitination conjugate, or formation of unbound xUB chains—a known activity of ICP0. The ubiquitination reactions were blotted against the HA epitope tag on xUB.

We then expressed these ICP0 RING mutants, and tested them in a ubiquitination assay with xE1, xE2 and xUB; the ICP0 RING is generally observed to form mostly free ubiquitin chains[49]—unanchored chains which are not attached to any substrate proteins—though relatively poorly understood, the biological function and significance of such chains have been reported[50-55], although it has been observed that the ICP0 RING also possesses some detectable residual autoubiquitination activity[56]; hence, through Western immunoblotting, we probed these xE3 reactions against the HA epitope tag fused to ubiquitin, with dubious results **(Figure 2-7).**

We also utilized this approach to make certain mutations in the RING domain of Mdm2 as well; although the alignment between cIAP and Mdm2 was less convincing than that with ICP0; Mdm2 did not possess acidic residues which were in the exact same location as D562 of cIAP, while E553 probably would correspond to either E427, N433 or E436 in the Mdm2 RING (see **Figure 2-5** above); nonetheless, using information from the standalone Mdm2 RING domain crystal structure[57], spatially speaking, these selected residues lay in the loop region which would be poised to interact with the N-terminal helix of the E2 enzyme. The sequence of these mutants were as given below (**Figure 2-8**); however, upon testing they showed no observable autoubiquitination activity (**Figure 2-9**).

	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442
wtmdm2	E	S	S	L	P	L	N	A	I	E	P	C	V	I	C	Q
mdm2 11	E	S	S	L	P	L	N	A	I	E	P	C	V	I	C	K
mdm2 21	K	S	S	L	P	L	N	A	I	E	P	C	V	I	C	K
mdm2 22	E	S	S	L	P	L	K	A	I	E	P	C	V	I	C	K
mdm2 23	E	S	S	L	P	L	N	A	I	K	P	C	V	I	C	K

Figure 2-8 Mdm2 RING mutants Alignment of the wild-type Mdm2 RING sequence (blue text) with those of rationally-designed mutants, showing mutated positions (shaded in yellow), and mutated residues (red text).

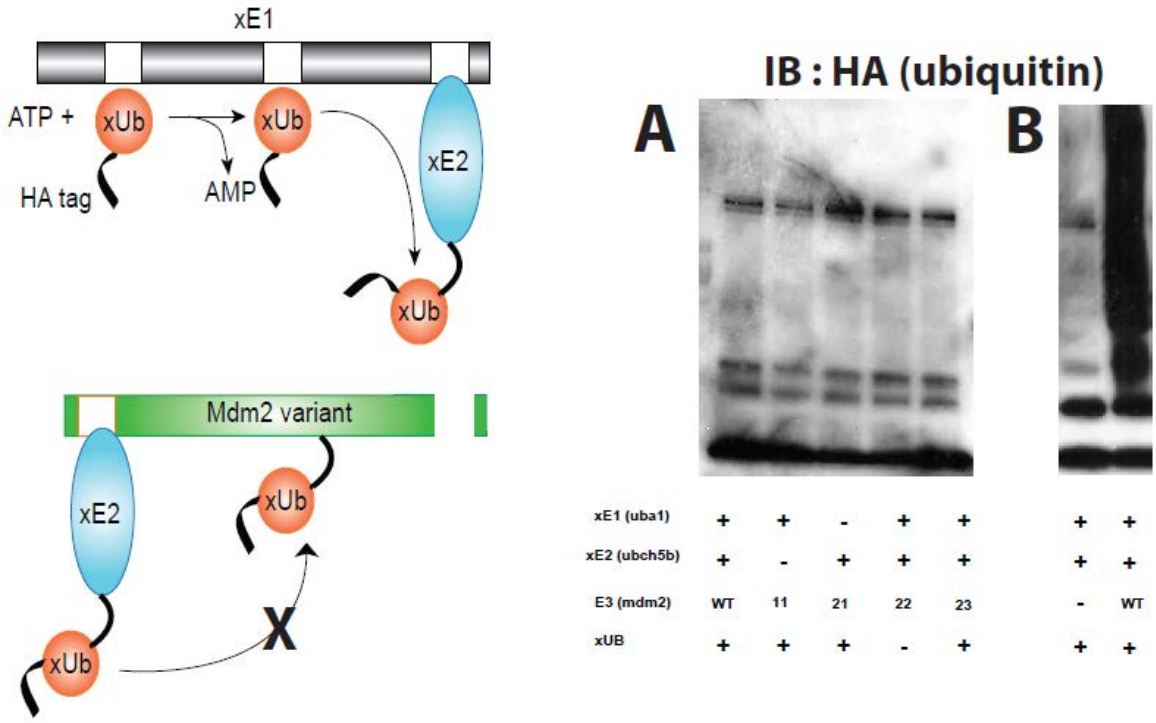

Figure 2-9 Ubiquitination assay of Mdm2 RING mutants. Wild-type Mdm2 RING, and engineered variants (**Figure 2-8**) were charged with OUT enzymes xE1, xE2 and xUB and activity was monitored by immunoblot against the HA epitope tag on xUB (A), with no significant ubiquitination observed. In comparison, the wild-type Mdm2 RING with wild-type E1, E2 and UB enzymes are also shown (B), compared to the control without the E3 enzyme.

2.2.3 Similar mutagenesis of the U-box domain

We turned our attention to a similar class of E3s: the U-box family, which is closely related to the RING family[58]. U-box E3s function in a manner identical to RING E3s[59] in that they act as bridges between the E2 and substrate; many U-box E3s also form dimers[60,61] whose structures bear strong resemblance with known RING dimers, and the general 3-dimensional conformation of the U-box greatly resembles that of the RING domains[62]; the main difference between U-boxes and RINGs is that U-box domains do not contain zinc fingers; they rely on salt bridges and hydrogen bonding[63] to achieve the similar structural stability conferred by the zinc-coordinating residues of the RINGs. The U-box E3s will be discussed in deeper detail in chapters 3 and 4.

At the start of our investigation, there had been a relative dearth of published E2-E3 crystal structures; hence, we also took great interest in a particular crystal structure published by Zhang et. al.[60], between a U-box E3, CHIP, and the E2 enzyme ubc13. The structure, showing similar acidic residues in CHIP interacting with residues in the N-terminal helix of E2 is as shown below **(Figure 2-10).**

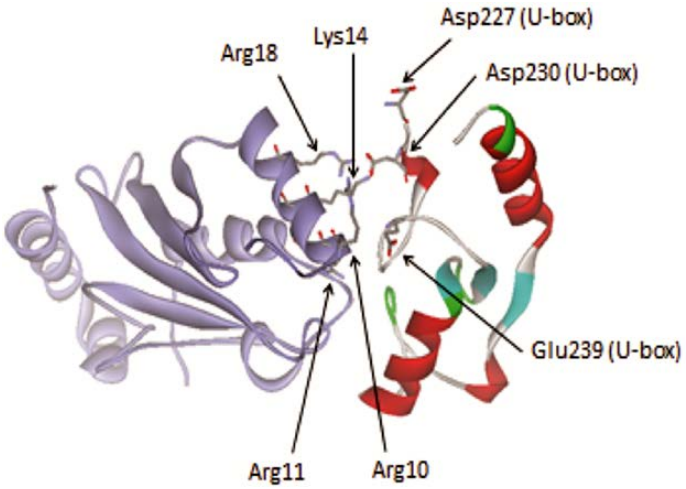

Figure 2-10
The Ubc13-CHIP U-box complex. The ubc13 E2 enzyme (left, purple) shown with four positively-charged residues in the N-terminal helix, which may potentially interact with the acidic residues D230 and E239 in the CHIP U-box (right), D227 not being likely involved given its orientation and relative position.

Naturally, we targeted D230 and E239 for mutagenesis; D227 being a bit further away and not in the proper orientation for E2 interaction; in an analogous manner to our work on ICP0, we constructed two single mutants and a double mutant of the CHIP U-box in which these residues were mutated to lysine. The sequences of which were as shown (**Figure 2-11**).

	230	231	232	233	234	235	236	237	238	239
wt CHIP	D	Y	L	C	G	K	I	S	F	E
CHIP 11	K	Y	L	C	G	K	I	S	F	E
CHIP 12	D	Y	L	C	G	K	I	S	F	K
CHIP 21	K	Y	L	C	G	K	I	S	F	K

Figure 2-11 CHIP U-box mutants. Alignment of the wild-type CHIP U-box sequence (blue text) with designed mutants, showing mutated positions (shaded in yellow), and residues (red text).

Testing the activity of these mutants with XE1, xE2 and xUB, in both autoubiquitination and in ubiquitination of CHIP's putative substrate Hsc70[64] gave somewhat encouraging results. These immunoblots were detected for the HA epitope on ubiquitin; possible bands corresponding in size to the CHIP~Ub and Hsc70-Ub conjugates were observed (**Figure 2-12**).

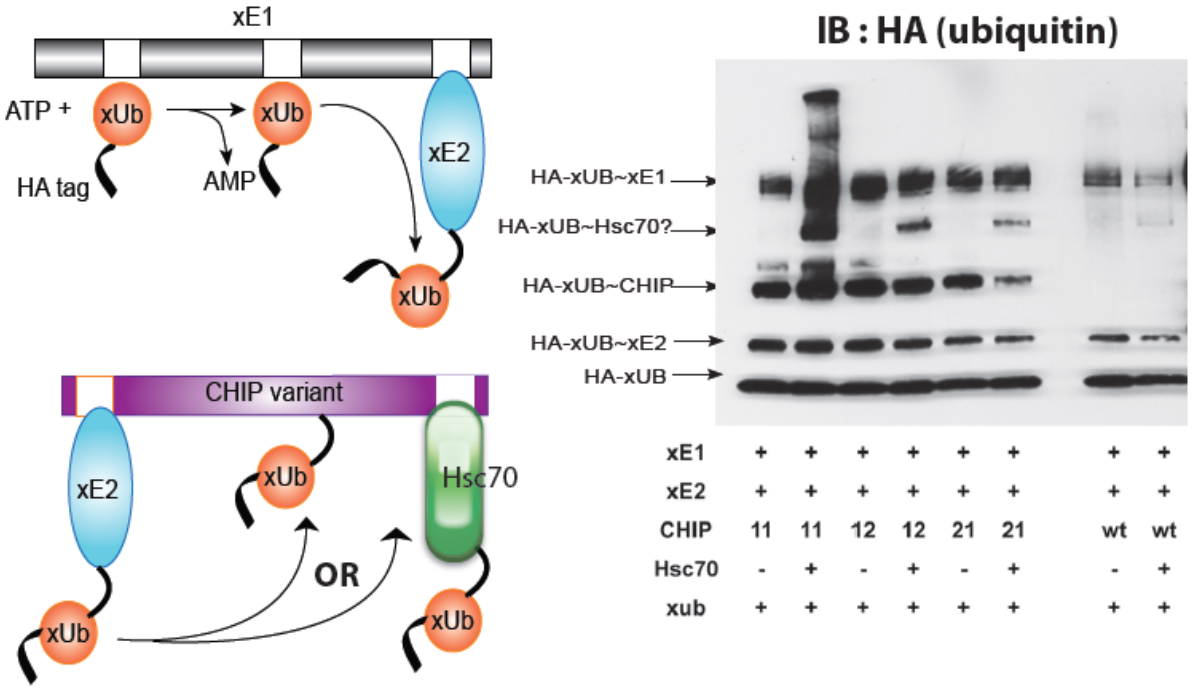

Figure 2-12 Ubiquitination assay of CHIP mutants. Wild-type CHIP and CHIP variants were charged with XE1, xE2 and xub, and tested for autoubiquitination, and putative substrate (Hsc70) ubiquitination. Blot was probed against HA epitope on xub.

We also probed the results of this assay using antibodies directly against the substrate Hsc70; Hsc70 was purchased commercially in the biotinylated form; however, upon immunoblotting and subsequent probing with Streptavidin-HRP, we did not detect any significant ubiquitination on the substrate Hsc70 (data not shown)—suggesting that the ubiquitin conjugates assigned as Hsc70~UB (**Figure 2-12**) were most likely higher molecular weight CHIP~UB conjugates and that the presence of Hsc70 may have enhanced CHIP autoubiquitination. Indeed, probing the same assay with an antibody against CHIP revealed that

the xCHIP variants were indeed capable of autoubiquitination with the OUT cascade; however, overall the ubiquitination activity of OUT was much lower than that of the wild-type cascade.

2.2.4 Partial restoration of E2 mutations

We had then observed that the creation of a RING or U-box domain which would be robustly active with the variant of xE2 (ubch5b 7-1 mutations from the previous ubc1 selection; see chapter 1) we had been using, was rather challenging at the least, and perhaps, although there was no way to be certain of this, theoretically impossible. We hence decided to partially restore some of the charge reversal mutations in the N-terminal helix back to the original wild-type sequence in hope of obtaining a more workable target for these E3 domains to accept, and to this end, we designed four different E2 variants given the names of 79-82, as shown (**Figure 2-13**).

	1	2	3	4	5	6	7	8	9
wtubch5b	M	A	L	K	R	I	H	K	E
xubch5b 7-1	M	A	L	E	D	I	H	E	Q
ubch5b 79	M	A	L	E	D	I	H	E	E
ubch5b 80	M	A	L	E	D	I	H	K	E
ubch5b 81	M	A	L	E	R	I	H	E	E
ubch5b 82	M	A	L	K	D	I	H	E	E

Figure 2-13 E2 mutant alignment. Shown is the alignment of the first nine residues in the N-terminal helix of the E2 enzyme ubch5b, with the wild type sequence (blue text), followed by mutant 7-1, which possesses four analogous mutations to the selected C1 mutant (refer to **Figure 2-5**); some residues were restored to their wild type amino acids in mutants 79-82.

A crucial requirement for these mutants to function as xE2 was that they could still accept ubiquitin from the E1 enzyme variants bearing the UFD fold mutations (namely, bE1 or xE1), and all four mutants proved to have retained that capability, as shown in the following ubiquitination assay **(Figure 2-14).**

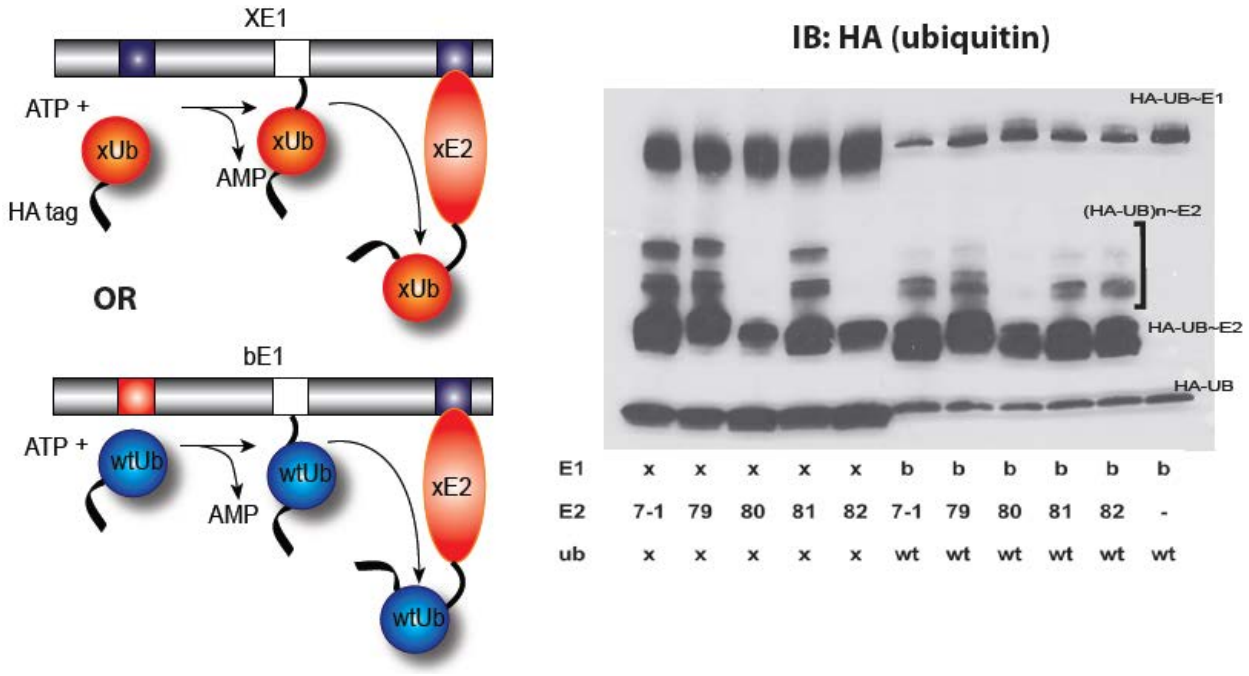

Figure 2-14 Ubiquitination assay of E2 mutants
The Ubch5b mutants 79-82 (**Figure 2-13**) were tested for ubiquitination activity against xE1/xUB (left set of reactions), and against bE1/wtUB. Immunoblot was probed for the HA epitope present on UB and any UB conjugates.

While the activity levels of these four mutants were somewhat similar, we decided to focus on mutant 81—a mutant which still retained the critical charge-reversed mutations on K4 and K8 of Ubch5, the two residues which seemed to be the most important in E2-E3 interaction shown in the crystal structures, and had all other residues reverted back to the wild-type E2 sequence. The vast majority of our work in the OUT system in the following chapters mainly involved this 81 mutant acting as xE2, and hence it will be referred to from this point on as such.

Upon confirmation by testing, the 81 mutant of ubch5b proved to be more active than the original 7-1 mutant from phage selection with the previously-mentioned CHIP mutants, both in autoubiquitination, and in substrate ubiquitination, as shown below. The assay methods were similar to what was described in 2.2.2 and 2.2.3 **(Figure 2-15)**; however, we were not able to

detect any significant activity on the engineered ICP0 and Mdm2 RING domains even with these new E2 mutants (data not shown).

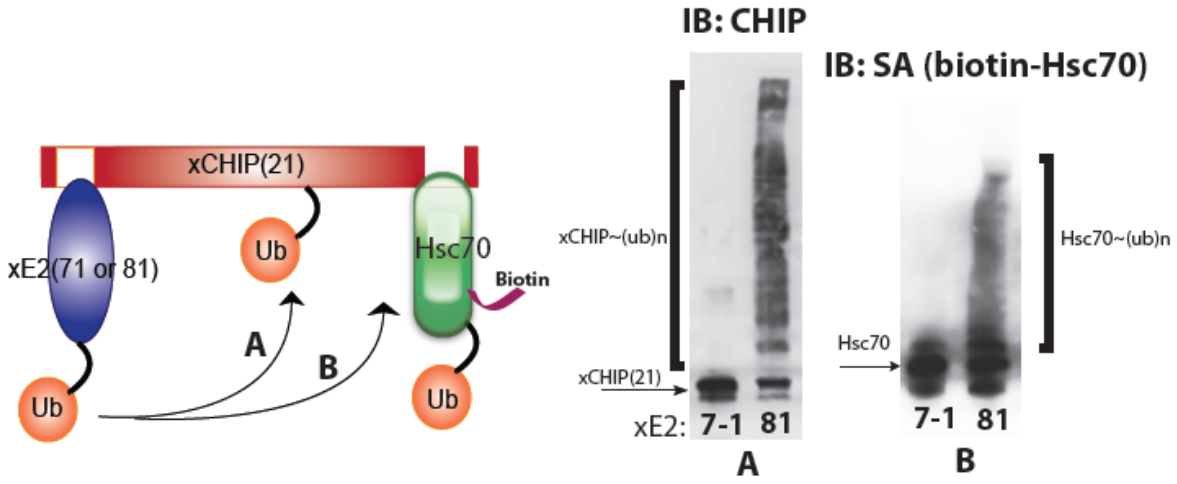

Figure 2-15 OUT activity of E2 mutants on E3 and putative E3 substrate
E2 mutants 7-1 and 81 were charged with bE1 and wild-type ubiquitin and tested for activity on xCHIP21 (**Figure 2-11**) and the putative CHIP substrate Hsc70. Smears on the immunoblots indicate the presence of various high molecular weight ubiquitin conjugates.

2.3 Discussion

2.3.1 The xE2-xE3 mutations

In reality, had the E2-E3 interface not intersected with the E1-E2 interface, a report involving the engineering of an orthogonal E2-E3 pair had already been previously published by Winkler and Timmers[65]; the charge reversal mutation site, K63E in E2, lay much further downstream of the N-terminal helix region—this mutation was complemented by a set of reciprocal charge reversal mutations D48K and E48K, in the RING E3 CNOT4, to create an orthogonal pair. Doubtless we would have taken their approach and used it for our platform, had it not been for the fact that we already had the broken interaction involving the E2 N-terminal helix to deal with.

It was not clear what the causes of the general lack of activity between the 7-1 xE2 and the engineered RING and U-box domains were. There were many possible reasons; perhaps the presence of four acidic residues in close proximity resulted in electrostatic repulsions which distorted the general conformation and recognition of the N-terminal helix, or that perhaps the presence of acidic residues in certain positions which in the wild-type enzyme did not interact with E3 generated non-native electrostatic interactions which prevented the E2 and E3 from coming into the proper orientation conducive to ubiquitin transfer.

Also it might have been the case, as mentioned earlier, that the binding between the 7-1 xE2 and the E3 mutants involved a greater number of possible electrostatic interactions, and in the end the binding was too strong for proper E2 dissociation—as mentioned previously, the E1-E2 interface and the E2-E3 interface overlap in the N-terminal helix, and hence as far as the E2 enzyme is concerned, E1 binding and E3 binding are believed to be mutually exclusive[66]; hence, formation of ubiquitin chains requires the E2 to periodically dissociate from the substrate complex to "recharge" by interaction with and ubiquitination by E1; in other words, tighter E2-E3 binding may yet end up to be detrimental to the level of overall ubiquitination[67].

It was also possible that the charge reversal mutations in E3 might have disrupted its native conformation, or that we did not target all the correct residues in E3 which actually interact with E2—indeed, a study later published on ICP0's interaction with the E2 enzyme[68] suggested that while we had correctly targeted D30 (numbered as D129 in the study), which supposedly interacts with K4 of E2, as well as E21 (numbered as E122) which interacts with both K4 and K8, we had not considered D16, which lay further upstream and was proposed to interact with K4; the software-modeled ICP0 structure given in the study differed slightly from the previously mentioned E2-cIAP structure we had used as a guide, hence our rational design

approach might very well have been intrinsically imperfect from the start (refer to **Figures 2-5** and **2-6**).

That said, for whatever reason, restoration of certain residues in xE2 back to their wild-type amino acids restored the activity of CHIP to a certain degree (**Figure 2-15**); while we could not say that we were completely satisfied with the results of our attempts, we were pleased with the promise the 81 ubch5b-xCHIP 21 mutant pair had shown in activity both E3 autoubiquitination and substrate ubiquitination assays, and that our rational design approach had yielded some degree of success; although we were unable to replicate the same results with the RING domain mutants. We hence concluded that significant differences probably existed between the RINGs and CHIP U-box domains, and that perhaps the RING domains tested had possessed greater selectivity.

That said, our success with the CHIP U-box notwithstanding, it was clear that the xE2-xE3 interface had not been and was far from optimized for activity; through rational design mutagenesis the best we could have aimed for was to develop an interface optimized for binding and interaction, but, as previously stated, tighter binding does not necessarily correlate to greater activity and our goal was to obtain an interface optimized for activity. Hence, this complicated our efforts in developing the xE2-xE3 interface through rational design; although in this chapter we were able to somewhat circumvent the problem through back-and-forth arbitrary mutations of residues in both the E2 and E3; nevertheless, we deemed that another approach which relies more heavily on actual observed activity than conjecture based on crystal structures was worth investigating. This approach will be discussed in Chapter 3.

2.4 Methods

The primers involved in this chapter are as given below. **(Table 2-1).**

Table 2-1 Primers used in chapter 2. Residues in red indicate restriction digest sites.

Primer name	Sequence (5'-3')
Kar49	GCCCCCCCTCGAGAGGACGGCGGGAGCGACGAGGGCGACGTGTGCGCCGTGTG CACGAAAGAGATCGCG
Kar50	GCCCCCCCTCGAGAGGACGGCGGGAGCGACGAGGGCAAAGTGTGCGCCGTGTGC ACGAAAGAGATCGCG
Kar51	GCCCCCCCTCGAGAGGACGGCGGGAGCGACAAAGGCGACGTGTGCGCCGTGTGC ACGAAAGAGATCGCG
Kar52	GCCCCCCCTCGAGAGGACGGCGGGAGCAAAGAGGGCGACGTGTGCGCCGTGTGC ACGAAAGAGATCGCG
Kar53	CCGCTCGCGGCCGCTCAGTCCAGGTCGTCGTCATC
Kar54	GGGAATTGTGAGCGGATAACAATTCC
Kar55	GGAAGAAGCGAGACATCCCCAAATACCTGTGTGGCAAGATCAGCTTTGAGCTGATG CGG
Kar56	GGAAGAAGCGAGACATCCCCGACTACCTGTGTGGCAAGATCAGCTTTAAACTGATG CGG
Kar57	GGGGATGTCTCGCTTCTTCC
Kar58	GTGGTGCTCGAGGTCTAGACTGTAGTCCTCCAC
Kar69	CCTTGTGTGATTTGTAAAGGTCGACCTAAAAATG
Kar70	ACAAATCACACAAGGTTTAATGGCATTAAGTGGCAAAC
Kar71	ACAAATCACACAAGGTTCAATGGCTTTAAGTGGCAAACTAGATTC
Kar72	ACAAATCACACAAGGTTCAATGGCATTAAGTGGCAAACTAGATTTCACACTCTCTTC
Kar73	ACAAATCACACAAGGTTCAATGGCATTAAGTGGCAAACTAGATTC
Kar74	ATGCATGCTCTCGAGTCAGGGGAAATAAGTTAGCACAATCATTTG

2.4.1 Construction of xE3 mutants pGEX wild-type ICP0 was a gift to our group. pET-GST-Mdm2-RING (wild-type) and pET-CHIP (wild-type) were constructs made by Bo Zhao, a former member of our group; these constructs were used as the template DNA for their respective PCR reactions.

The mutations in ICP0 were introduced via an internal XhoI restriction site of the gene; primers used were one of either Kar49, Kar50, Kar51 or Kar52 (for mutants 11, 23, 22 and 21 respectively) with the downstream primer Kar53. The PCR amplicon was double-digested and ligated into the original pGEX-ICP0 (wt) vector via XhoI/NotI.

The D230K and E239K mutations were introduced via overlap extension PCR by Kar55 and Kar56, respectively. The amplified fragment from the reaction with Kar54 (which binds to the upstream lac operator sequence in pET vectors) and either Kar55 or Kar56 was coupled with a complementary fragment from Kar57 and Kar58. The final PCR product was digested with, and ligated into pET-CHIP (wild-type) via NdeI/XhoI restriction sites.

Mdm2 mutations were also introduced by overlap extension PCR; Kar69 which bore the Q442K mutation common in all of the mutants, was paired with Kar74; this was coupled with a fragment from one of either Kar70, Kar71, Kar72 or Kar73 for clones 23, 22, 21, and 11, respectively, with Kar54. The final PCR product was digested with, and ligated into pET-GST-Mdm2 RING (wt) via NdeI/XhoI restriction sites.

2.4.2 Expression and purification of proteins

Uba1, ubch5b, ICP0, Mdm2 and CHIP variants were expressed in BL21(DE3) cells, grown in Luria-Bertani (LB) broth at 37˚C, with 1mM IPTG induction at OD ~0.6-0.8; induction was done at 15˚C overnight. The cells were harvested by centrifugation at 5000 rpm for 30 minutes, resuspended in 2-5 ml of lysis buffer (50 mM sodium phosphate pH 8.0, 300 mM NaCl, supplemented with 10 mM imidazole in the case of 6-His-affinity purification) per gram of wet pellet weight, 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.

Ubch5b, mdm2, and CHIP variants bearing 6-His tags, were affinity purified from cleared lysates using Ni-NTA agarose via a method developed by the manufacturer QIAGEN. GST-tagged ICP0 variants were purified using pre-packed columns containing glutathione resin (Pierce), according to the protocol given by the manufacturer. Eluted proteins were dialyzed overnight into 50 mM Tris pH 7.5-8.0, 50 mM NaCl and 0.5 mM DTT at 4˚C. Purified proteins were quantified using the Bradford assay, and the purity confirmed through Coomassie staining.

2.4.3 Ubiquitination assays

All ubiquitination reactions described in this chapter were performed in 50 µL of 50 mM Tris pH 7.5, 10 mM $MgCl_2$, 50 µM DTT, and 3 mM ATP.

2.4.3.1 E3 autoubiquitination assays

2 µM of E3 enzyme (ICP0, Mdm2, or CHIP) was charged with 0.5 µM XE1 (uba1), 2 µM xE2 (7-1 or 81 mutant of ubch5b) and 14 µM HA-tagged xub. The reactions were allowed to proceed for 90 minutes at 30˚C. The reactions were quenched by boiling in Laemmli buffer and analyzed by Western immunoblotting with either mouse anti-HA, or rabbit anti-CHIP (both from Santa Cruz).

2.4.3.2 Hsc70 ubiquitination with CHIP assays

0.1 µM of biotinylated Hsc70 (American Bio) was charged with 0.5 µM of xE1 (Uba1), 8 µM of xE2 (7-1 or 81 mutant of Ubch5b), 4 µM of xCHIP (11, 12 or 21 mutant) and 14 µM of xUB. The reactions were allowed to proceed overnight at 30˚C, diluted 10 fold in reaction buffer,

quenched by boiling in Laemmli buffer supplemented with DTT, and analyzed by Western immunoblotting with Streptavidin-HRP (Pierce).

2.4.3.3 E2 ubiquitination assay

2 μM of Ubch5b variant (7-1, or 79-82) was charged with either 0.1 μM of xE1 (Uba1) and 2 μM of HA-xUB or the same amounts of bE1 and HA-wild-type-UB. The reactions were allowed to proceed for 30 minutes at room temperature; immunoblotting and detection of HA-ub conjugates proceeded in an identical manner to that described in 2.4.3.1.

2.5 References

1 Zhao, B. *et al.* Orthogonal ubiquitin transfer through engineered E1-E2 cascades for protein ubiquitination. *Chem Biol* **19**, 1265-1277, doi:10.1016/j.chembiol.2012.07.023 (2012).

2 Budhidarmo, R., Nakatani, Y. & Day, C. L. RINGs hold the key to ubiquitin transfer. *Trends in biochemical sciences* **37**, 58-65, doi:10.1016/j.tibs.2011.11.001 (2012).

3 Metzger, M. B., Hristova, V. A. & Weissman, A. M. HECT and RING finger families of E3 ubiquitin ligases at a glance. *Journal of cell science* **125**, 531-537, doi:10.1242/jcs.091777 (2012).

4 Liu, Y. *et al.* RNF135, RING finger protein, promotes the proliferation of human glioblastoma cells in vivo and in vitro via the ERK pathway. *Sci Rep* **6**, 20642, doi:10.1038/srep20642 (2016).

5 Tomar, D. & Singh, R. TRIM family proteins: emerging class of RING E3 ligases as regulator of NF-kappaB pathway. *Biology of the cell / under the auspices of the European Cell Biology Organization* **107**, 22-40, doi:10.1111/boc.201400046 (2015).

6 Wu, W. *et al.* Targeting RING domains of Mdm2-MdmX E3 complex activates apoptotic arm of the p53 pathway in leukemia/lymphoma cells. *Cell death & disease* **6**, e2035, doi:10.1038/cddis.2015.358 (2015).

7 Ho, S. R., Mahanic, C. S., Lee, Y. J. & Lin, W. C. RNF144A, an E3 ubiquitin ligase for DNA-PKcs, promotes apoptosis during DNA damage. *Proc Natl Acad Sci U S A* **111**, E2646-2655, doi:10.1073/pnas.1323107111 (2014).

8 Kim, J. H. & Kim, W. T. The Arabidopsis RING E3 ubiquitin ligase AtAIRP3/LOG2 participates in positive regulation of high-salt and drought stress responses. *Plant physiology* **162**, 1733-1749, doi:10.1104/pp.113.220103 (2013).

9 Ma, P. *et al.* The ubiquitin ligase RNF220 enhances canonical Wnt signaling through USP7-mediated deubiquitination of beta-catenin. *Mol Cell Biol* **34**, 4355-4366, doi:10.1128/mcb.00731-14 (2014).

10 Mosbech, A., Lukas, C., Bekker-Jensen, S. & Mailand, N. The deubiquitylating enzyme USP44 counteracts the DNA double-strand break response mediated by the RNF8 and RNF168 ubiquitin ligases. *J Biol Chem* **288**, 16579-16587, doi:10.1074/jbc.M113.459917 (2013).

11 Park, E. S. *et al.* Tumor necrosis factor (TNF) receptor-associated factor (TRAF)-interacting protein (TRIP) negatively regulates the TRAF2 ubiquitin-dependent pathway by suppressing the TRAF2-sphingosine 1-phosphate (S1P) interaction. *J Biol Chem* **290**, 9660-9673, doi:10.1074/jbc.M114.609685 (2015).

12 Pedersen, S. M., Chan, W., Jattani, R. P., Mackie d, S. & Pomerantz, J. L. Negative Regulation of CARD11 Signaling and Lymphoma Cell Survival by the E3 Ubiquitin Ligase RNF181. *Mol Cell Biol* **36**, 794-808, doi:10.1128/mcb.00876-15 (2015).

13 Yu, Q. *et al.* E3 Ubiquitin ligase RNF183 Is a Novel Regulator in Inflammatory Bowel Disease. *Journal of Crohn's & colitis*, doi:10.1093/ecco-jcc/jjw023 (2016).

14 Burger, A. M. *et al.* A novel RING-type ubiquitin ligase breast cancer-associated gene 2 correlates with outcome in invasive breast cancer. *Cancer research* **65**, 10401-10412, doi:10.1158/0008-5472.can-05-2103 (2005).

15 Huang, J. *et al.* The poxvirus p28 virulence factor is an E3 ubiquitin ligase. *J Biol Chem* **279**, 54110-54116, doi:10.1074/jbc.M410583200 (2004).

16 Lee, J. & Zhou, P. Pathogenic Role of the CRL4 Ubiquitin Ligase in Human Disease. *Frontiers in oncology* **2**, 21, doi:10.3389/fonc.2012.00021 (2012).

17 Maejima, Y. *et al.* Muscle-specific RING finger 1 negatively regulates pathological cardiac hypertrophy through downregulation of calcineurin A. *Circulation. Heart failure* **7**, 479-490, doi:10.1161/circheartfailure.113.000713 (2014).

18 Taylor, S. J. *et al.* Loss of c-Cbl E3 ubiquitin ligase activity enhances the development of myeloid leukemia in FLT3-ITD mutant mice. *Experimental hematology* **43**, 191-206.e191, doi:10.1016/j.exphem.2014.11.009 (2015).

19 Wang, E. *et al.* Histone H2B ubiquitin ligase RNF20 is required for MLL-rearranged leukemia. *Proc Natl Acad Sci U S A* **110**, 3901-3906, doi:10.1073/pnas.1301045110 (2013).

20 Wang, Z. *et al.* RNF115/BCA2 E3 ubiquitin ligase promotes breast cancer cell proliferation through targeting p21Waf1/Cip1 for ubiquitin-mediated degradation. *Neoplasia (New York, N.Y.)* **15**, 1028-1035 (2013).

21 Xi, S. *et al.* Downregulation of ring-finger protein 43 in glioma associates with poor prognosis. *International journal of clinical and experimental pathology* **8**, 490-496 (2015).

22 Yagishita, N. *et al.* RING-finger type E3 ubiquitin ligase inhibitors as novel candidates for the treatment of rheumatoid arthritis. *International journal of molecular medicine* **30**, 1281-1286, doi:10.3892/ijmm.2012.1129 (2012).

23 Lorick, K. L. *et al.* RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 11364-11369 (1999).

24 Brzovic, P. S., Rajagopal, P., Hoyt, D. W., King, M. C. & Klevit, R. E. Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. *Nature structural biology* **8**, 833-837, doi:10.1038/nsb1001-833 (2001).

25 Zheng, N. *et al.* Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature* **416**, 703-709, doi:10.1038/416703a (2002).

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

27 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. *Biochim Biophys Acta* **1843**, 47-60, doi:10.1016/j.bbamcr.2013.05.026 (2014).

28 Rodgers, K. K. *et al.* A zinc-binding domain involved in the dimerization of RAG1. *J Mol Biol* **260**, 70-84, doi:10.1006/jmbi.1996.0382 (1996).

29 Dou, H., Buetow, L., Sibbet, G. J., Cameron, K. & Huang, D. T. Essentiality of a non-RING element in priming donor ubiquitin for catalysis by a monomeric E3. *Nat Struct Mol Biol* **20**, 982-986, doi:10.1038/nsmb.2621 (2013).

30 Das, R. *et al.* Allosteric activation of E2-RING finger-mediated ubiquitylation by a structurally defined specific E2-binding region of gp78. *Molecular cell* **34**, 674-685, doi:10.1016/j.molcel.2009.05.010 (2009).

31 Amemiya, Y., Azmi, P. & Seth, A. Autoubiquitination of BCA2 RING E3 ligase regulates its own stability and affects cell migration. *Molecular cancer research : MCR* **6**, 1385-1396, doi:10.1158/1541-7786.mcr-08-0094 (2008).

32 Chen, A., Kleiman, F. E., Manley, J. L., Ouchi, T. & Pan, Z. Q. Autoubiquitination of the BRCA1*BARD1 RING ubiquitin ligase. *The Journal of biological chemistry* **277**, 22085-22092, doi:10.1074/jbc.M201252200 (2002).

33 Walsh, M. C., Kim, G. K., Maurizio, P. L., Molnar, E. E. & Choi, Y. TRAF6 autoubiquitination-independent activation of the NFkappaB and MAPK pathways in response to IL-1 and RANKL. *PLoS One* **3**, e4064, doi:10.1371/journal.pone.0004064 (2008).

34 Singh, S. K. & Gellert, M. Role of RAG1 autoubiquitination in V(D)J recombination. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 8579-8583, doi:10.1073/pnas.1510464112 (2015).

35 Dueber, E. C. *et al.* Antagonists induce a conformational change in cIAP1 that promotes autoubiquitination. *Science (New York, N.Y.)* **334**, 376-380, doi:10.1126/science.1207862 (2011).

36 Ran, Y. *et al.* Autoubiquitination of TRIM26 links TBK1 to NEMO in RLR-mediated innate antiviral immune response. *Journal of molecular cell biology* **8**, 31-43, doi:10.1093/jmcb/mjv068 (2016).

37 Bourgeois-Daigneault, M. C. & Thibodeau, J. Autoregulation of MARCH1 expression by dimerization and autoubiquitination. *Journal of immunology (Baltimore, Md. : 1950)* **188**, 4959-4970, doi:10.4049/jimmunol.1102708 (2012).

38 Weissman, A. M., Shabek, N. & Ciechanover, A. The predator becomes the prey: regulating the ubiquitin system by ubiquitylation and degradation. *Nature reviews. Molecular cell biology* **12**, 605-620, doi:10.1038/nrm3173 (2011).

39 David, Y. *et al.* E3 ligases determine ubiquitination site and conjugate type by enforcing specificity on E2 enzymes. *The Journal of biological chemistry* **286**, 44104-44115, doi:10.1074/jbc.M111.234559 (2011).

40 Iwai, K. & Tanaka, K. Ubiquitin chain elongation: an intriguing strategy. *Molecular cell* **56**, 189-191, doi:10.1016/j.molcel.2014.10.009 (2014).

41 Pierce, N. W., Kleiger, G., Shan, S. O. & Deshaies, R. J. Detection of sequential polyubiquitylation on a millisecond timescale. *Nature* **462**, 615-619, doi:10.1038/nature08595 (2009).

42 Qian, S. B. *et al.* Engineering a ubiquitin ligase reveals conformational flexibility required for ubiquitin transfer. *The Journal of biological chemistry* **284**, 26797-26802, doi:10.1074/jbc.M109.032334 (2009).

43 Li, W., Tu, D., Brunger, A. T. & Ye, Y. A ubiquitin ligase transfers preformed polyubiquitin chains from a conjugating enzyme to a substrate. *Nature* **446**, 333-337, doi:10.1038/nature05542 (2007).

44 Mace, P. D. *et al.* Structures of the cIAP2 RING domain reveal conformational changes associated with ubiquitin-conjugating enzyme (E2) recruitment. *J Biol Chem* **283**, 31633-31640, doi:10.1074/jbc.M804753200 (2008).

45 Dou, H., Buetow, L., Sibbet, G. J., Cameron, K. & Huang, D. T. BIRC7-E2 ubiquitin conjugate structure reveals the mechanism of ubiquitin transfer by a RING dimer. *Nat Struct Mol Biol* **19**, 876-883, doi:10.1038/nsmb.2379 (2012).

46 Hodson, C., Purkiss, A., Miles, J. A. & Walden, H. Structure of the human FANCL RING-Ube2T complex reveals determinants of cognate E3-E2 selection. *Structure* **22**, 337-344, doi:10.1016/j.str.2013.12.004 (2014).

47 Plechanovova, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H. & Hay, R. T. Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature* **489**, 115-120, doi:10.1038/nature11376 (2012).

48 Pruneda, J. N. *et al.* Structure of an E3:E2~Ub complex reveals an allosteric mechanism shared among RING/U-box ligases. *Mol Cell* **47**, 933-942, doi:10.1016/j.molcel.2012.07.001 (2012).

49 Boutell, C., Sadis, S. & Everett, R. D. Herpes simplex virus type 1 immediate-early protein ICP0 and is isolated RING finger domain act as ubiquitin E3 ligases in vitro. *Journal of virology* **76**, 841-850 (2002).

50 Hao, R. *et al.* Proteasomes activate aggresome disassembly and clearance by producing unanchored ubiquitin chains. *Mol Cell* **51**, 819-828, doi:10.1016/j.molcel.2013.08.016 (2013).

51 Miranzo-Navarro, D. & Magor, K. E. Activation of duck RIG-I by TRIM25 is independent of anchored ubiquitin. *PloS one* **9**, e86968, doi:10.1371/journal.pone.0086968 (2014).

52 Ouyang, H. *et al.* Protein aggregates are recruited to aggresome by histone deacetylase 6 via unanchored ubiquitin C termini. *J Biol Chem* **287**, 2317-2327, doi:10.1074/jbc.M111.273730 (2012).

53 Rajsbaum, R. *et al.* Unanchored K48-linked polyubiquitin synthesized by the E3-ubiquitin ligase TRIM6 stimulates the interferon-IKKepsilon kinase-mediated antiviral response. *Immunity* **40**, 880-895, doi:10.1016/j.immuni.2014.04.018 (2014).

54 Strachan, J. *et al.* Insights into the molecular composition of endogenous unanchored polyubiquitin chains. *Journal of proteome research* **11**, 1969-1980, doi:10.1021/pr201167n (2012).

55 Zeng, W. *et al.* Reconstitution of the RIG-I pathway reveals a signaling role of unanchored polyubiquitin chains in innate immunity. *Cell* **141**, 315-330, doi:10.1016/j.cell.2010.03.029 (2010).

56 Boutell, C. & Davido, D. J. A quantitative assay to monitor HSV-1 ICP0 ubiquitin ligase activity in vitro. *Methods (San Diego, Calif.)* **90**, 3-7, doi:10.1016/j.ymeth.2015.04.004 (2015).

57 Kostic, M., Matt, T., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Solution structure of the Hdm2 C2H2C4 RING, a domain critical for ubiquitination of p53. *J Mol Biol* **363**, 433-450, doi:10.1016/j.jmb.2006.08.027 (2006).

58 Aravind, L. & Koonin, E. V. The U box is a modified RING finger - a common domain in ubiquitination. *Current biology : CB* **10**, R132-134 (2000).

59 Cyr, D. M., Hohfeld, J. & Patterson, C. Protein quality control: U-box-containing E3 ubiquitin ligases join the fold. *Trends in biochemical sciences* **27**, 368-375 (2002).

60 Zhang, M. *et al.* Chaperoned ubiquitylation--crystal structures of the CHIP U box E3 ubiquitin ligase and a CHIP-Ubc13-Uev1a complex. *Mol Cell* **20**, 525-538, doi:10.1016/j.molcel.2005.09.023 (2005).

61 Vander Kooi, C. W. *et al.* The Prp19 U-box crystal structure suggests a common dimeric architecture for a class of oligomeric E3 ubiquitin ligases. *Biochemistry* **45**, 121-130, doi:10.1021/bi051787e (2006).

62 Hatakeyama, S., Yada, M., Matsumoto, M., Ishida, N. & Nakayama, K. I. U box proteins as a new family of ubiquitin-protein ligases. *J Biol Chem* **276**, 33111-33120, doi:10.1074/jbc.M102755200 (2001).

63 Ohi, M. D., Vander Kooi, C. W., Rosenberg, J. A., Chazin, W. J. & Gould, K. L. Structural insights into the U-box, a domain associated with multi-ubiquitination. *Nature structural biology* **10**, 250-255, doi:10.1038/nsb906 (2003).

64 Soss, S. E., Rose, K. L., Hill, S., Jouan, S. & Chazin, W. J. Biochemical and Proteomic Analysis of Ubiquitination of Hsc70 and Hsp70 by the E3 Ligase CHIP. *PloS one* **10**, e0128240, doi:10.1371/journal.pone.0128240 (2015).

65 Winkler, G. S. & Timmers, H. T. M. Structure-Based Approaches to Create New E2–E3 Enzyme Pairs. **399**, 355-366, doi:10.1016/s0076-6879(05)99024-1 (2005).

66 Eletr, Z. M., Huang, D. T., Duda, D. M., Schulman, B. A. & Kuhlman, B. E2 conjugating enzymes must disengage from their E1 enzymes before E3-dependent ubiquitin and ubiquitin-like transfer. *Nature structural & molecular biology* **12**, 933-934, doi:10.1038/nsmb984 (2005).

67 Starita, L. M. *et al.* Activity-enhancing mutations in an E3 ubiquitin ligase identified by high-throughput mutagenesis. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E1263-1272, doi:10.1073/pnas.1303309110 (2013).

68 Vanni, E., Gatherer, D., Tong, L., Everett, R. D. & Boutell, C. Functional characterization of residues required for the herpes simplex virus 1 E3 ubiquitin ligase ICP0 to interact with the cellular E2 ubiquitin-conjugating enzyme UBE2D1 (UbcH5a). *Journal of virology* **86**, 6323-6333, doi:10.1128/jvi.07210-11 (2012).

Chapter 3

Identification of active E3 mutants via activity-based phage selection

3.1 Background information

3.1.1 Rationale

In the preceding chapter we have discussed our attempts at creating active xE3 RING/U-box mutants through rational design; the endeavor of which, we have had admittedly limited, yet, somewhat encouraging success. We have good reason to continue to believe that crucial salt bridges between positive residues in the E2 N-terminal helix and negative ones in a conserved loop regions in RING and U-box type E3s significantly contribute towards the interaction between the enzymes—as evidenced by the virtually complete cessation of activity when the charges of crucial arginine and lysine residues were reversed in xE2 and that reversal of acidic residues in the interacting loop in E3s to some extent restored the activity, and that in rational design engineering of these proteins, known crystal structures have given us somewhat useful clues as to which region of the E3 to effectively target for mutagenesis.

Yet, for all their usefulness, presently available crystal structures do not provide all the answers—at the time of writing, we have very good reason to believe that these crystal structures do not perfectly represent the E2-E3-UB complex in its active conformation, and that the whole complex most likely undergoes a transient, and yet drastic change to the so-called "closed conformation"[1]; it is in this closed conformation that the movement of the E2~UB thioester is restricted and coaxed into a favorable position for nucleophilic attack and in which actual ubiquitination takes place. In other words, a drastic, yet poorly-understood conformational

change of the whole E2-E3-UB complex which is not reflected in the present crystal structures is most likely required (**Figure 3-1**).

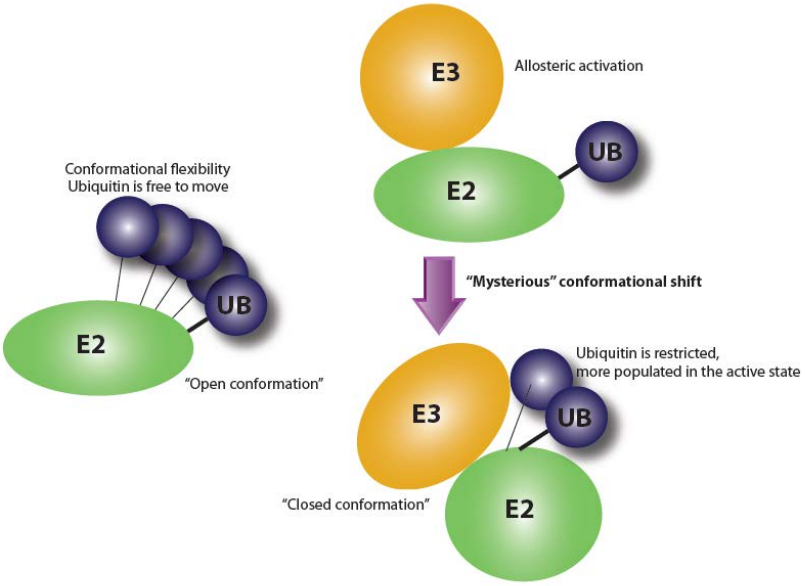

Figure 3-1.
The closed conformation for ubiquitin transfer. In the open conformation, the ubiquitin molecule is free to adopt a wide variety of conformations; however, upon the shift of the E2-E3-UB complex to the closed conformation, UB movement is much more restricted and hence the E2~UB thioester is more likely to be in the positions which are accessible to attack by nucleophiles.

Hence, information from existing crystal structures being imperfect, from a purely rational design standpoint, even with powerful tools such as sophisticated 3-D molecular modeling software, it is difficult to pinpoint the exact optimal combination of mutations that are to be made. In other words, to use an analogy, no matter how skilled the locksmith, he cannot make a key that works effectively if he does not know the exact shape of the keyhole in the lock. Naturally, when structural information is imperfect or lacking, a different approach which relies less on precise knowledge is worth considering: directed evolution; to use another analogy, if one struggles to catch a fish using a spear because one does not know the precise position of the fish, one might be inclined to switch to using a wide net instead.

3.1.2 Directed evolution approach

In contrast to the rational design approach, which involves the testing of specific mutants bearing well-defined, precise, engineered mutations, the so-called directed evolution approach involves a diversified population of mutants bearing randomized or semi-randomized mutations[2], generally referred to as a library—the library is subjected to an iterative selection process which is not altogether unlike artificial selection methods used in agriculture or animal husbandry, the end of which is an enriched population whose makeup consists of mutants with the desired characteristics.

The effectiveness of this method relies on three things: firstly, that there be suitable variation among, and even distribution of, mutants in the library[3]—the greater the diversity of the library, the greater chance of it containing something useful or optimal, and the more evenly distributed it is, the less likely that certain mutants are favored over others in terms of representation, which reduces the odds of obtaining false positives; secondly, that there be a suitable method in which useful mutants are positively identified and enriched and/or useless mutants are rejected and eliminated[4]—an effective selection or screening assay, which generally relies on catalytic activity or binding affinity between interactors, is most commonly used to achieve this; lastly, if multiple rounds of selection are to be performed, that there be a suitable way for the enriched mutants to propagate, to pass on the genetic information encoding their desired phenotypes[5], in other words, heredity—this is accomplished typically via a display system which makes use of already-existing natural machinery; a rough outline of one specific way this approach may to be taken is given in the figure below (**Figure 3-2**).

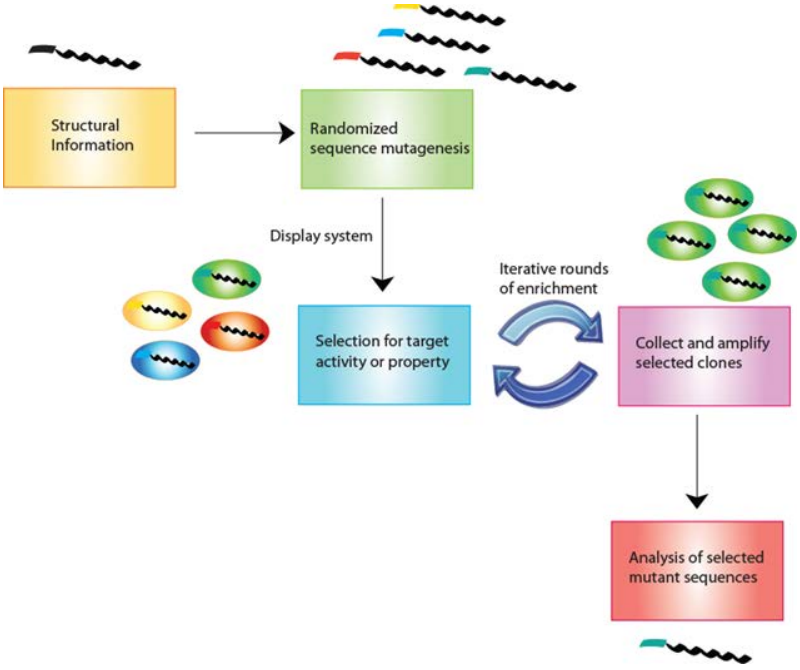

Figure 3-2
Directed evolution overview
Shown here is one possible way of conducting an experiment based on directed evolution; note the aforementioned three important factors: variation (mutagenesis), identification and enrichment (selection), and heredity (amplification of enriched clones).

When structural information of the target protein to be evolved is scarce, and it is unclear which region of the target protein is to be mutated, or when one wishes to investigate the effects of mutations throughout the whole gene and not just a particular region or domain, a library is often constructed through error-prone PCR[6], in which a non-proofreading polymerase, operating under certain non-optimal conditions, such as an unusually high divalent cation concentration in the buffer, or an unbalanced ratio of different dNTPs—conditions which can be adjusted to vary the rate of mutation frequency, sporadically incorporates mismatched nucleotides throughout the amplicon; this results in mutants consisting of mutations in different regions throughout the whole gene. The use of libraries generated by this method have been reported in numerous instances such those listed here. One drawback, however, is that the library, no matter what the display system, cannot possibly be large enough to accommodate the theoretical possible number of mutants—a number which would likely be astronomical; another drawback is the lack of evenness in the distribution of different mutants in the library: certain mismatches of base pairs

are more stabilized, and hence certain mutations are favored over others; this bias is inherent in the nature of the error-prone PCR reaction itself and can be affected by multiple variables in the reaction mixture[7]. On many occasions, however, critical residues or regions in the target protein have already been identified or proposed via existing mutagenesis studies or crystal structures, as was the case with our previous investigation involving the engineering of ubiquitin and E2, in which case mutations can be focused specifically at these residues or regions instead[8]; this is simply achieved by standard PCR using synthesized oligo nucleotide primers in which specific positions in the strand are randomized—this indeed, had been the main approach taken in our previous studies.

There have been virtually a countless number of published screening or selection assays—which will not be discussed thoroughly in detail here; generally, many assays rely on competition among mutants for "bait", in which mutants showing affinity for a target of interest are "caught" and subsequently enriched. The bait may be in the form of a target immobilized on solid support which can easily be isolated[9], such as on a polystyrene surface, or on agarose or magnetic beads, or in the form of a labelled or tagged target[10]: the bait, and whatever "prey" is caught by it, can be "reeled in" by the tag; while the selection assay is generally performed *in vitro*, it is possible to perform it *in vivo*, such as in the so-called two-hybrid method[11,12]. Also, "negative selections", which focus on eliminating undesired mutants rather than enriching desired ones have been reported[13]; one way this may be accomplished is by using a "predator", such as antibiotic, to cull away useless mutants, such as reported here[14,15].

A display system is essential for connecting the displayed mutant protein, which makes selection and screening possible, to its corresponding genetic information, which makes heredity and amplification possible[5]. Many systems involve the fusion of the target protein of interest to a

native protein on the surface or coat of a specific organism, where the target is exposed and can engage in substrate binding or enzymatic reactions. Phage display[16], yeast surface display[17], and bacterial surface display[18] are three of the most commonly used display systems. We have dealt extensively with M-13 phage display in our reported findings—the specifics of which will be discussed in the following section.

3.1.3 M-13 phage display

Background information in 3.1.3 is mostly summarized from a laboratory manual on phage display by Barbas et. al.[19]

3.1.3.1 M-13 phage biology

M-13 is a member of the Ff class of filamentous bacteriophages (genus *Inovirus*) which utilizes the host bacteria's F conjugative pilus as a receptor in infection (and hence, propagation must be done in *E.coli* strains containing the F factor, often referred to as "male" strains commercially, such as XL1-blue or NEBStable); these phages are lysogenic: they do not kill the host cell during productive infection[20]—cells that are infected merely exhibit slowed growth rates (the generation time of infected bacterial cells is around 50% longer than uninfected ones[19]), and are identifiable as plaques when plated. The general structure of the M-13 phage[21] is as shown below **(Figure 3-3).**

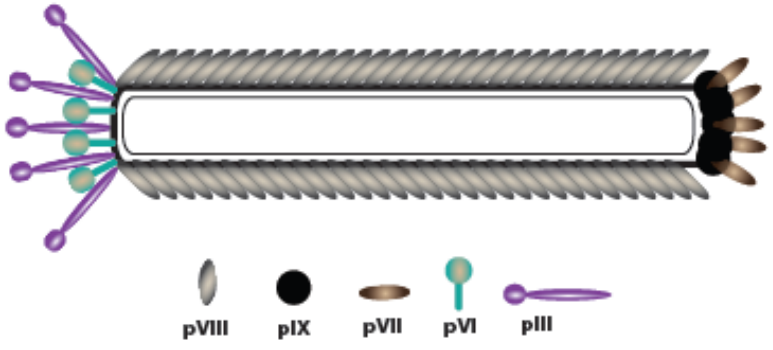

Figure 3-3. Structure of the M-13 phage.
Encapsulated within the protein coat comprised of 6 different proteins is the (+) single-stranded phage DNA. Phage display technology most commonly involves display on either pIII or pVIII.

The mature phage particle is approximately 6.5 nm in diameter and 930 nm in length[19]. Encapsulated within the cylinder-shaped protein coat is the circular single-stranded phage DNA; upon infection of the host bacteria, this single stranded DNA (referred to as the (+) strand) is translocated into the host cytoplasm, where host enzymes are employed to synthesize the complementary (-) strand—this (-) strand is the template for mRNA transcription of all of the phage proteins; however, only the (+) strand contains a hairpin region termed the "packaging signal", which is in turn recognized by certain phage proteins, and is the only strand packed into the protein coat upon nascent phage assembly[22].

The M-13 phage genome has been fully sequenced, and is found to encode for 11 proteins termed pI through pXI[23-25]. Out of these, the major capsid protein pVIII, and more often, the tail capsid protein pIII, have been commonly utilized in display proteins of interest on the phage surface[26]; although pVII[27] and pIX[28] display systems have also been reported, it is pIII display on M-13 phages which our research has been centered on: proteins of interest may be fused to the N-terminal portion of pIII, as first demonstrated by George Smith in 1985[16], and the resulting fusion protein is able to be packed and displayed on the phage tail, albeit at the cost of reduced infectivity—pIII plays a central role in host infection; however, they are still able to be propagated in a similar manner as wild-type phage[29].

3.1.3.2 Phagemid vectors

In the earlier days of the technology, the phage display vector in usage was simply the single (+) strand phage genomic DNA—with certain modifications such as the insertion of restriction sites for ease of cloning in genes of interest, or including a lacZ gene for blue-white

colony screening, such as in the case of the M13KE vector[30,31] widely used. Later on, advances in the technology led to the development of what is called the "phagemid"[32], which is basically a modified version of the typical bacterial plasmid used in routine research, containing things such as a bacterial origin of replication, transcription promoter, antibiotic resistance gene for selection, as well as certain affinity tags, and can be amplified and propagated as any typical plasmid[33]. The phagemid, however, does not contain the genes of all of the 11 phage proteins necessary for phage assembly—typically it only contains the display gene of interest a.k.a. pIII and whatever gene of interest fused to its N-terminal; the pIII fusion is also put under control of a lac promoter, making it possible to induce overexpression of the protein via galactose or its derivatives. Although the phagemid contains a packaging signal and can be packed in a single-stranded form into phage particles, a bacterial cell simply transformed with the phagemid cannot autonomously assemble phage particles, lacking most of the genes encoding the proteins necessary to do so; phage propagation must be initiated by the introduction of a "helper phage", which provides the cell with the rest of the necessary genetic information for phage assembly[34] (**Figure 3-4**)—the helper phage is essentially a wild-type phage whose genomic DNA bears a defective origin of replication which greatly diminishes the rate in which it is replicated by the host bacterial DNA polymerases[19], to ensure that the majority of the phages assembled contains the (+) strand of the phagemid, and not the helper phage DNA. The advantages of using a phagemid vector over a simple phage vector are numerous, among them greater ease of manipulation—the phagemid being in all aspects similar to a bacterial plasmid, greater control over the timing of which the phages are assembled, and the ability to either induce or inhibit expression of the fusion protein of interest.

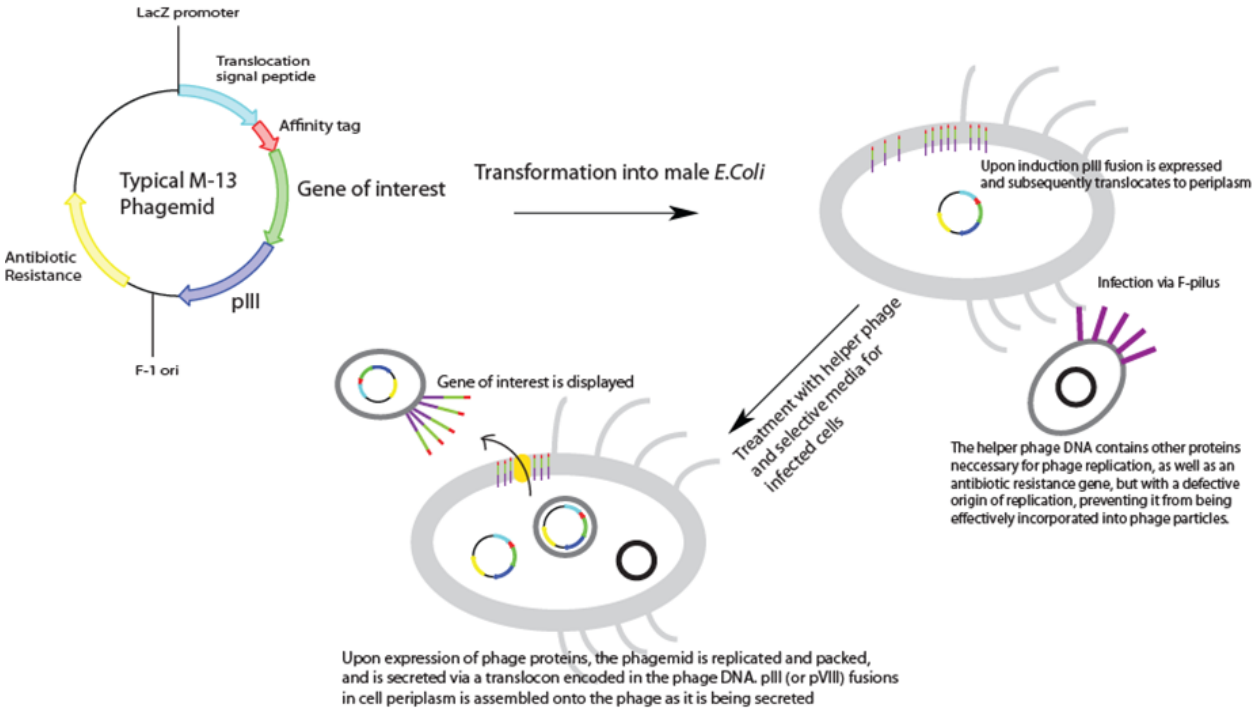

Figure 3-4 Phagemid vector and propagation of recombinant phage.
Shown is the general structure and composition of a generic pIII display phagemid; the double-stranded phagemid is transformed into a bacterial cell, upon which transcription and translation of the protein of interest can be regulated; the recombinant pIII fusion protein molecules produced are then translocated via a signal peptide sequence encoded in the phagemid to the periplasm of the cell, where in awaits incorporation into phage particles. Upon infection with helper phage, the (+) strand of the phagemid is packed into new phage particles, upon which the recombinant pIII is assembled onto, and which is secreted out of the cell through a protein pore comprised mostly of the phage protein pIV.

Of all the pIII display vectors, the pComb vectors are most well-known; the first of which was developed was pComb3[32], followed by pComb3X and pComb3H[35,36]—which was the backbone vector used in our previous E2 selection. Phages propagated from cells transformed with pComb vectors display the gene of interest as a direct N-terminal fusion to the pIII protein. However, in some cases it becomes necessary for certain proteins to have free unhindered C-terminal regions—in which cases pJuFo (pJF) vectors should be utilized instead. pJuFo vectors are derivatives of the original pComb3, and contain the leucine zipper protein Jun fused to the N-

terminal portion of pIII[37]. Under a separate ribosome binding site, the gene of interest to be displayed is fused to the C-terminal of the leucine zipper protein Fos—which is known to form a stable heterodimeric complex with Jun[38], being linked covalently through the spontaneous formation of a pair of disulfide bridges; the Jun-pIII fusion and the Fos-gene of interest fusion associate with each other inside the bacterial periplasm into which they are secreted, and the whole fusion is recognized and assembled onto nascent phage particles in the same manner as wild-type pIII; the compositions of the different phagemids are as given **(Figure 3-5).**

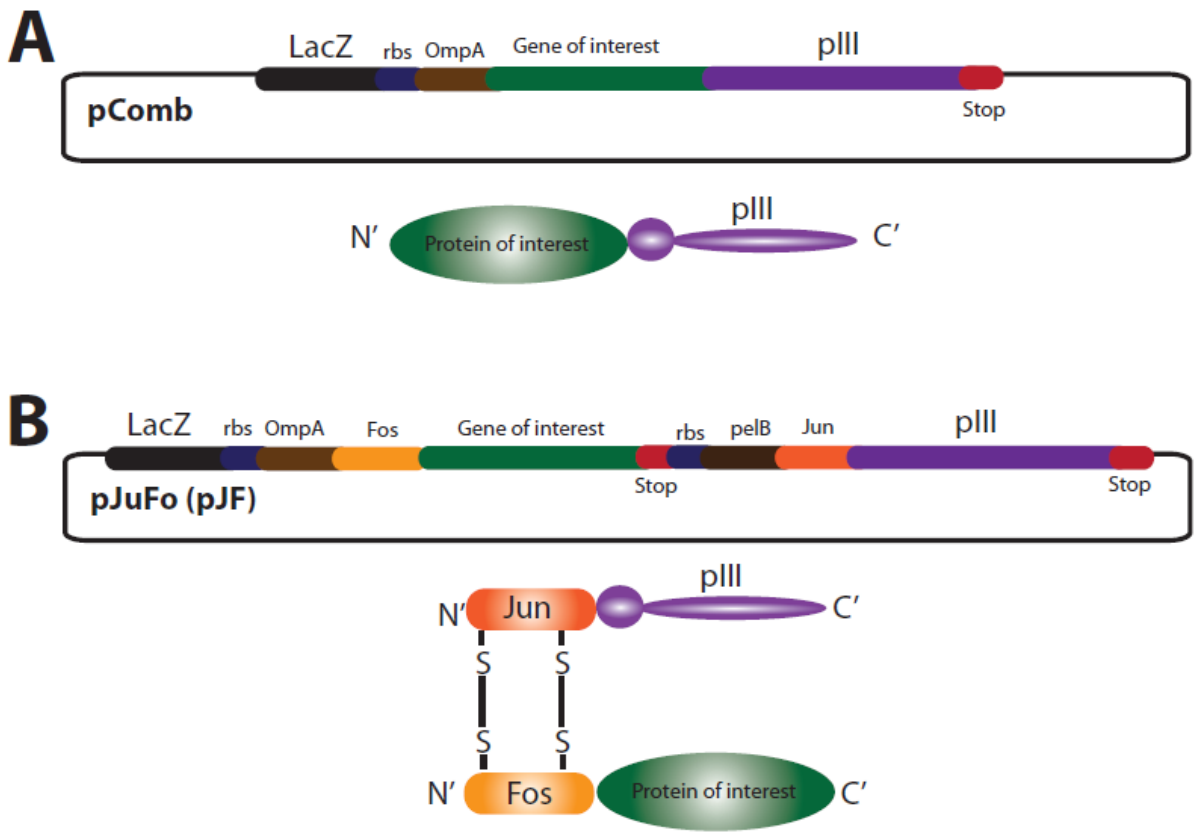

Figure 3-5 Comparison of phagemid systems. Partial vector maps for the pComb (A) and pJF (B) phagemids are as shown, along with the corresponding recombinant pIII proteins.

We indeed had utilized pJF3H in our ubiquitin selection as reported previously[39]; in which the phage involved would look somewhat similar to that shown above **(Figure 3-5)**, with

the Fos-ubiquitin fusion protein's C-terminal carboxyl being free to react with E1 enzymes, which would not have been possible under the pComb system. As also will be shown in this chapter, our library selection of the E4B U-box also utilized a pJF3H phagemid as the backbone.

3.2 Results

3.2.1 Past attempts at phage display of E3 ligases

We initially had many different E3 ligases as candidates for phage display; a successful display of a particular E3 ligase requires that the enzyme is able to be properly expressed, and tolerated by the host cell, and that the enzyme can survive the secreting process and be displayed in its active form upon the phage surface. While we have had past success in reconstituting the ubiquitin ligase activities of a variety of E3 enzymes when they were expressed as free proteins—E3s tested included HECT E3s, such as E6AP, Rsp5, and smurfs 1&2, and RING E3s, such as mdm2, ICP0, as well as the U-box E3 CHIP—when we tried testing the display of these ligases on the phage particles (data not shown), utilizing both pComb and pJF vectors, we were met with very little success—generally we found that one of the following problems occurred: either the E3 protein in the phagemid seems to be toxic to the host bacterial cell, resulting abrogated or greatly diminished proper culture growth (specifically in the case of Rsp5) and poor to non-existent phage yields, or the protein was not properly displayed on the phage surface, as evidenced by the lack of signal when probed either by Western immunoblotting and/or ELISA assays for the inserted affinity tag fused typically to the C-terminal end of the E3, or that the E3, although successfully displayed, had no detectable activity in ubiquitination assays (data not shown).

3.2.2 Phage display of the U-box domain of E4B

3.2.2.1 E4B U-box activity autoubiquitination assays

In a typical manner similar to our previous attempts with other E3 ligases, we cloned the U-box domain of the E3 ligase E4B[40] (the known biological roles and specifics of which will be discussed in the following chapter), into the phagemid display vector pJF3H, with FLAG epitope tag fused to the C-terminal of the protein. The resulting recombinant phage showed clear signal when probed for the FLAG epitope in western immunoblotting assays. More importantly, the displayed wild-type E4B U-box was also found to be active in a typical *in vitro* autoubiquitination assay with wild type ubiquitin, E1, and E2 enzymes. The level of activity of the phage-displayed U-box is similar to the level we observed when the free E4B U-box protein (expressed from a generic pET vector system) was used in a similar assay (**Figure 3-6**).

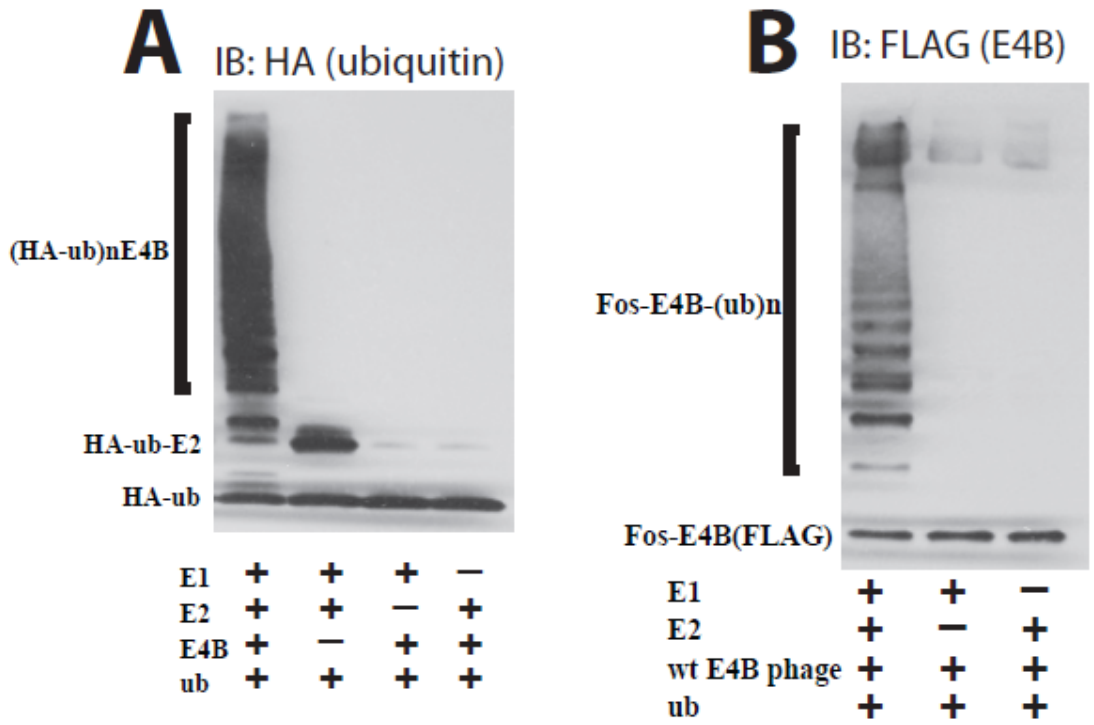

Figure 3-6 Ubiquitination assays of E4B U-box. Recombinantly expressed free E4B U-box protein (left) and phage surface-displayed E4B U-box with FLAG epitope, fused to the Fos protein in the pJF display system, were tested for autoubiquitination activity with wild-type ubiquitin, E1, and E2 enzymes. The presence of ubiquitin conjugates was immediately obvious.

3.2.2.2 E4B U-box-displaying-phage model selection assays

We next tested the pJF E4B phage for activity in a standard biopanning assay we previously used for our UB and E2 phage selections, as shown in the figure below. Biopanning usually has been carried out by our group on commercially-available polystyrene plates coated with streptavidin (Thermo); the phage undergoes a standard ubiquitination assay via E1 and E2 with commercial ubiquitin protein which has been labelled at the N-terminal amine group with biotin—any ubiquitinated phages are bound to the plate and their numbers can be quantified either directly by ELISA (**Figure 3-7A**), or by elution via disulfide reduction with DTT (which disrupts the Jun-Fos interaction) and direct titration of the phage number by infection of male bacterial cells. (**Figure 3-7B**).

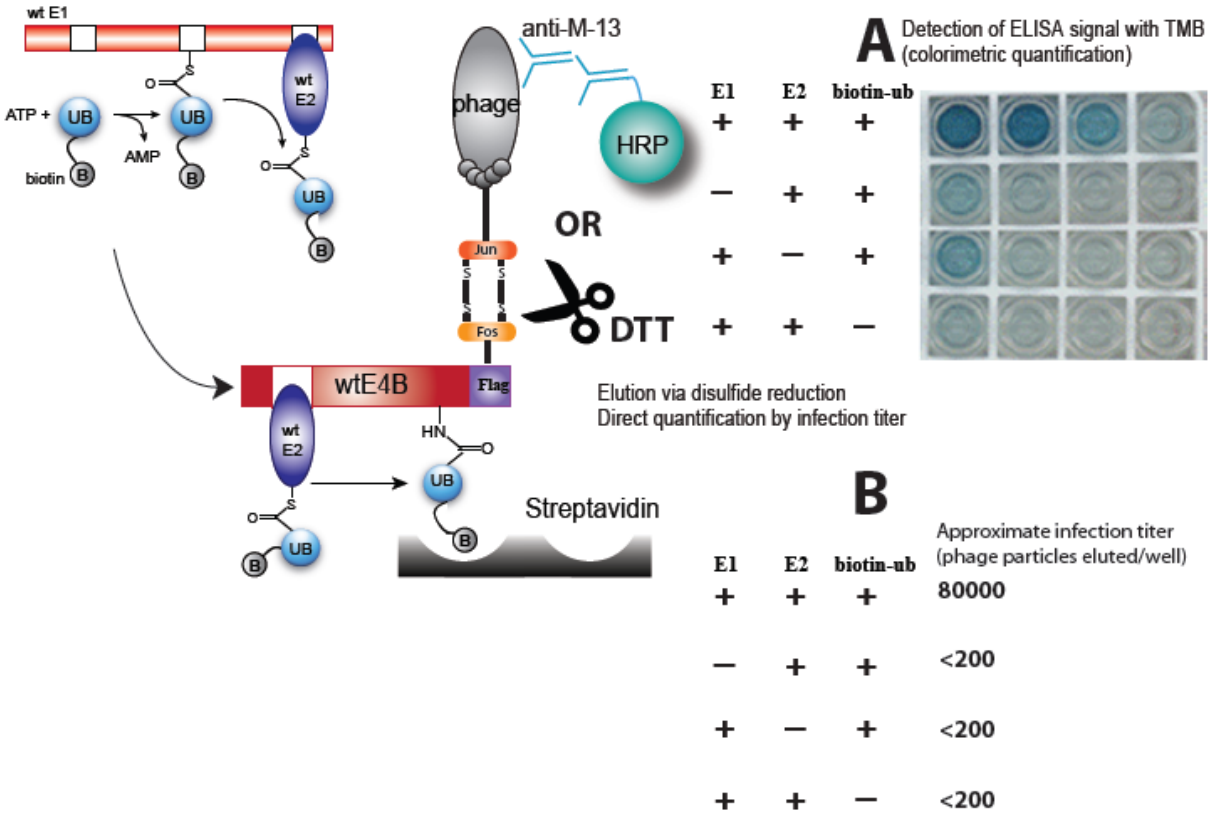

Figure 3-7 Biopanning assay of E4B phage. Phage particles derived from the pJF system displaying the E4B U-box were ubiquitinated in a similar manner as before (Scheme 3-1). Ubiquitinated phages were captured and quantified by ELISA (A) or direct titration (B).

We also tested the phage in another assay which more mimics the conditions of a library selection—a library is generally comprised of a very vast majority of inactive mutants which nonetheless compete via non-specific binding with the desirable, active mutants for binding of the bait molecule, which in this case, is biotin. In this assay we mixed varying ratios of E4B-displaying pJF-derived phage with another pJF-derived phage displaying the unrelated SV5V protein[41], which acts to simulate noise from inactive mutants in a real selection **(Figure 3-8).**

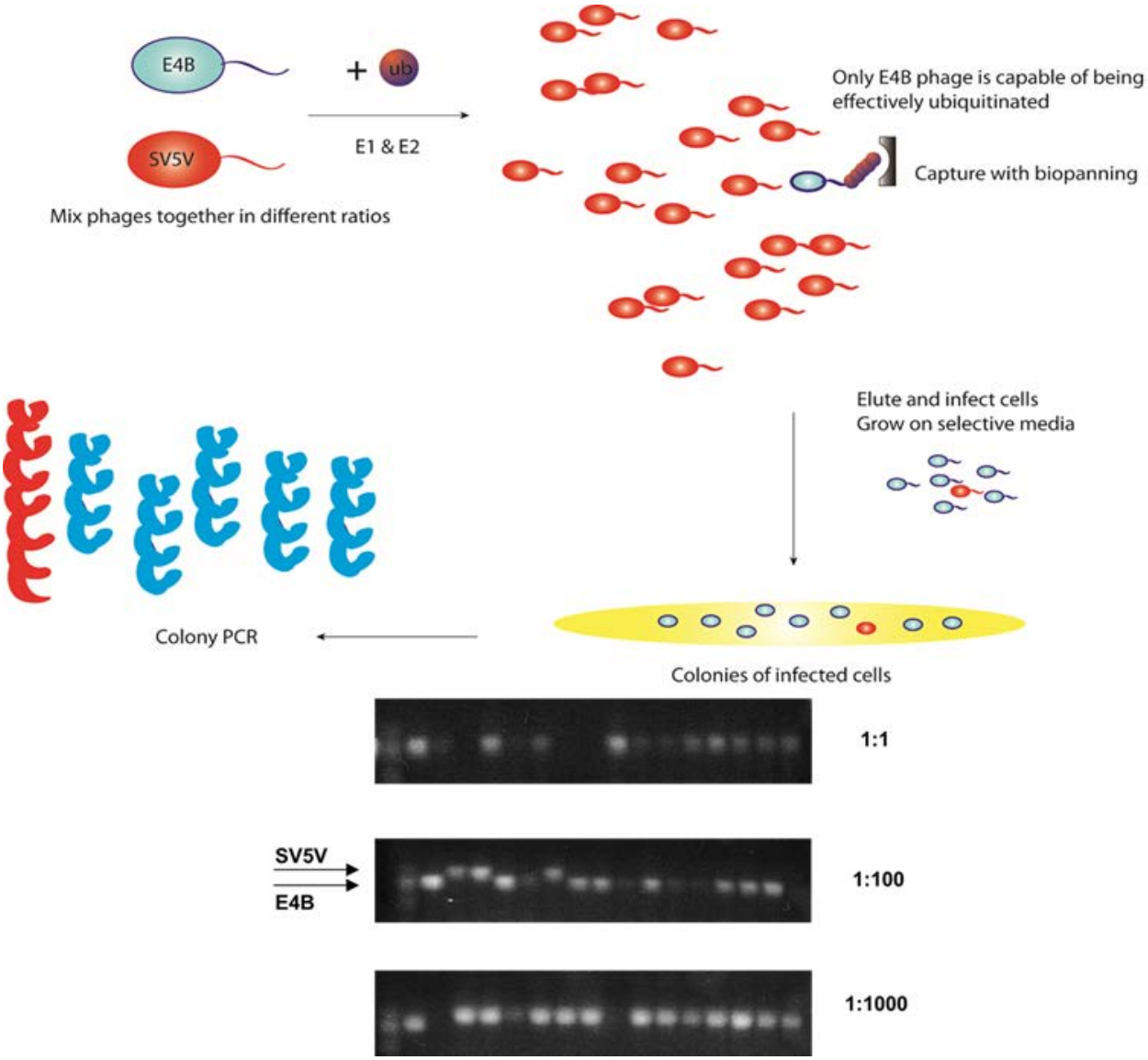

Figure 3-8 Model selection of E4B phage. The phage mixtures were subjected to a standard ubiquitination and biopanning assay, and bound phages were eluted with DTT (**Figure 3-7B**). Identity of the phages were confirmed by colony PCR of infected bacterial colonies.

We subjected the phage mixture to a standard biopanning assay as described above, after suitable washing to remove unbound phages, bound phages were eluted and used to infect male *E.coli* cells—upon infection the packed single-stranded form of the phagemid is translocated into the cell, conferring to it ampicillin resistance. Upon streaking on selective solid media, single colonies were screened via colony PCR using primers specific for the pJF vector—the E4B gene insert being noticeably smaller than SV5V (respective bands indicated by the arrows in Scheme 3-3). We discovered that within a single biopanning round of our assay we were able to enrich the active E4B phage over the SV5V phage by at least 100-1000 fold; in a 1:1 mixture all sampled output clones were E4B; in a 1:100 mixture the sampled output was 13 E4B clones and 3 SV5V clones; even when the E4B phage was outnumbered 1:1000, we still were able to detect one clone bearing the E4B insert out of sixteen. Hence, we then decided to proceed to the actual library selection to engineer the OUT variation of the E4B U-box, the first step of which, naturally would be to obtain the library containing randomized mutants, as will be discussed in the following section.

3.2.3 E4B Phage library construction

3.2.3.1 General considerations for library mutagenesis

A crystal structure previously reported of the U-box domain of the E3 ligase E4B in complex with the E2 enzyme Ubch5[42] (**Figure 3-9**) bears uncanny resemblance to other known E2-E3 structures mentioned in the previous chapter[43,44] (**Figures 2-4** and **2-10**). 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; from their findings, it appears that the side

chain carboxyl group of D1238 of E4B is possibly involved in electrostatic interactions with the side chain amino groups of K4 and K8 of E2—charge reversal of these residues completely abrogated activity with wild-type RINGs and U-boxes, as was shown in the previous chapter.

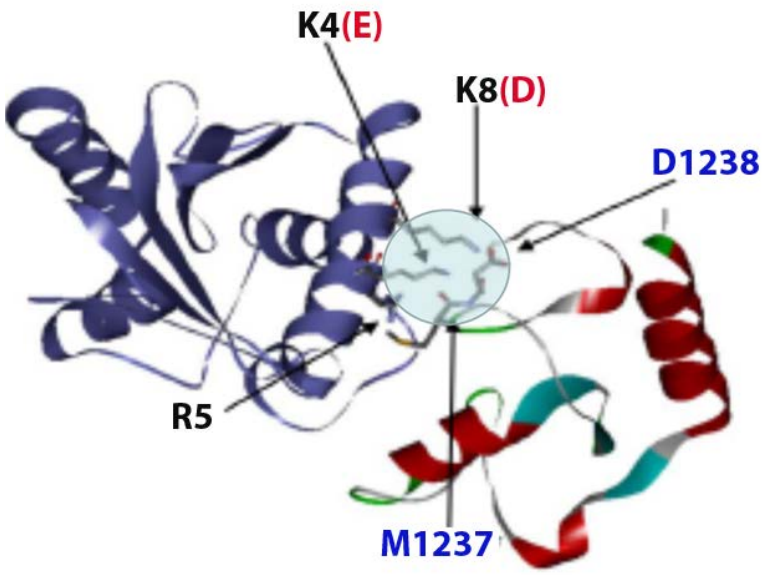

Figure 3-9
The Ubch5c-E4B U-box complex. Shown is the E2 enzyme ubch5c (left, purple) with crucial N-terminal basic residues as shown (black text) and their respective mutations in the 81 mutant (refer to **Figure 2-13**) of xE2 (red text), interacting with the E4B U-box (right, multicolored) residues (blue text) D1238 via likely salt bridges (light blue circle) through K4 and K8, and possibly, M1237 which is also in close proximity to R5.

With this in mind we ultimately decided to focus on a stretch of residues spanning a loop structure which seemed to be in closest proximity to the E2 N-terminus; R1234, L1236, M1237, D1238 and T1239 were targeted for randomized mutagenesis in library construction.

3.2.3.2 Library generation

We constructed our E4B phage library via generic restriction enzyme digestion and ligation in the pJF3H phagemid; the original insert of the vector was SV5V, which has proven to be inactive in our ubiquitination assays and should be eliminated from our selection should they somehow be overrepresented in the starting library due to incomplete digestion. We found dephosphorylation of the digested vector to be unnecessary—test ligations showed very little

background transformants. The approximately 300 bp U-box domain was inserted as a PCR product with digested sticky ends; the five point mutations on the target loop discussed in the previous section were introduced via standard overlap extension PCR utilizing primers which have the target positions randomized in primer synthesis as NNK codons[19]—this is to eliminate codon bias as much as possible, since using standard NNN codons leads to overrepresentation of certain amino acids. The template DNA for the PCR reaction was a mutant of the E4B U-box in which the five target residues were all mutated to alanine—a mutant which is sure to be inactive to minimize the effects of starting bias introduced by PCR as much as possible. We also introduced a FLAG tag at the C-terminal of the U-box sequence—this tag may be used as a way to identify and eliminate frameshift or truncated mutants later on in the selection.

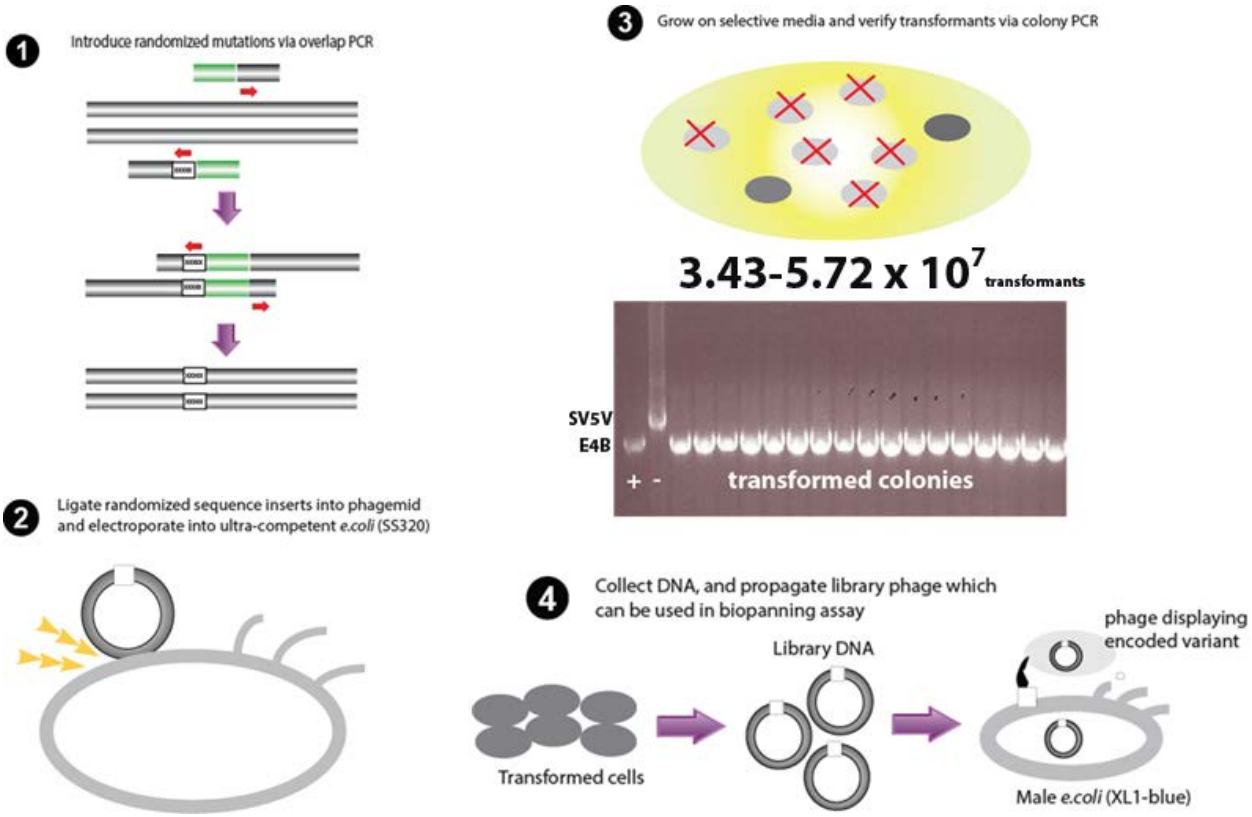

Figure 3-10 Construction of the library. A rough scheme describing the major steps in library construction. The number of transformants were titrated, and screened by colony PCR (step 3) similar to Scheme 3-3; the size of the amplicons were compared against control PCR reactions on known DNA templates (from left, first two); all 17 screened clones encoded for E4B.

After some test ligations, according to our titrations, we obtained a number of around 3.43-5.72 x 10^7 independent transformants from our large scale ligation—a number well above the theoretical maximum diversity of our library, which is the number of all possible combinations of five randomized amino acids, (20^5) which is 3.2 x 10^6—on the DNA level, the number of all possible NNK codon combinations was $(4x4x2)^5$ or ~3.3 x 10^7, or roughly the low-end of the likely number of transformants we had achieved. Hence, we had good reason to believe that every possible mutant had at least some level of representation in our constructed library. The mutants were confirmed via colony PCR and subsequent DNA sequencing. No transformants bearing the original SV5V insert were detected in either case (**Figure 3-10**).

3.2.4 Selection of the library

Selection of the library was done through ubiquitination and biopanning assays similar to our model selection assay; in our first attempt, we subjected our library through rounds of iterative selection; library phages were charged with N-terminal-labelled biotin-wtUB (Boston Biochem), bE1 (E1 with the C-terminal mutations for xE2), and xE2 (specifically mutant 81, which was discussed in the previous chapter); this combination was chosen over the full OUT cascade which would involve xUB and the xE1 with the full set of AD and UFD mutations (refer to **Figure 1-8**); our focus had been solely on the E2-E3 interface, and hence we deemed involving the R42E and R72E mutations in xUB with the selection process to be unnecessary; furthermore, this allowed the use of wild-type biotinylated UB which was commonly available commercially, whereas using xUB would have required us to synthesize and purify our own biotin-conjugated xUB—(although our group had previously developed a method to achieve this[39] involving fusion of a peptidyl carrier protein (PCP) onto the target, followed by enzymatic

conjugation of biotinylated Coenzyme A onto the PCP domain via Sfp phosphopantetheinyl transferase[45,46]). To distinguish genuine activity from background binding, we included three controls in which one component crucial to the ubiquitination of the U-box, namely E1, E2 or UB, was missing; after allowing the reaction to proceed, we incubated the reaction mixture containing the phages in commercially-available pre-blocked polystyrene wells coated with either streptavidin or neutravidin. After thorough washing, the phages were eluted from the wells by the addition of a DTT solution which serves to cleave the Jun-Fos disulfide bonds; 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 3-11**).

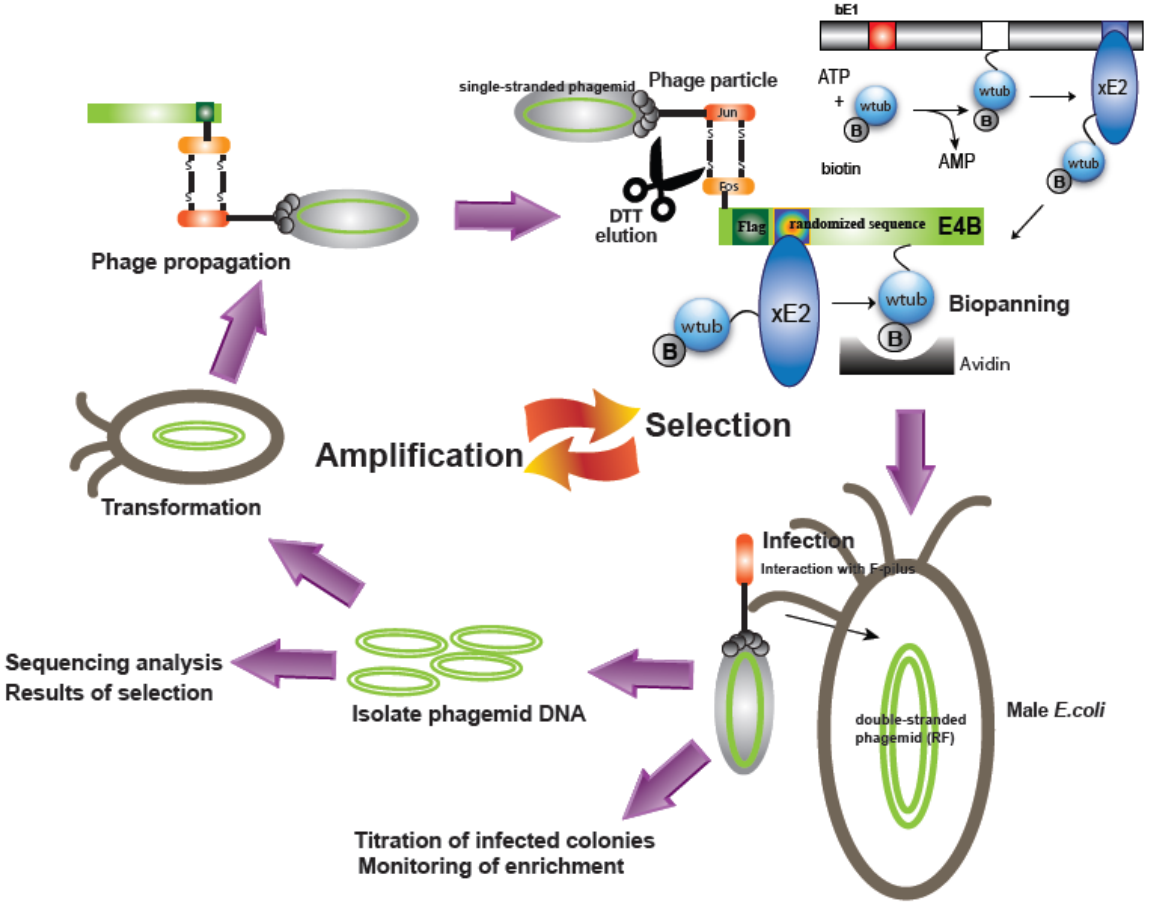

Figure 3-11 General schematic overview of the library selection process. While it was possible to propagate the next round of phage directly from the infected cells right after the biopanning stage (by adding helper phage), the method we chose better minimizes the effects of clones with growth and other non-activity-related advantages.

(A)	R1	R2	R3	R4	R5	R6
[bE1] (μM)	0.2	0.3	0.5	0.2	0.05	0.02
[xE2 81] (μM)	10	10	3	20	10	5
[biotin-wtub] (μM)	0.5	0.2	0.1	0.04	0.04	0.04
Input phage titer	$1.2x10^{11}$	$3.8x10^{11}$	$3.54x10^{10}$	$1.6x10^{10}$	$2.6x10^{10}$	$3.46x10^{10}$
Output titer (E1,E2,ub)	$2.4x10^{6}$	$1.76x10^{5}$	$7.21x10^{5}$	$1.95x10^{6}$	$8.0x10^{3}$	$5.6x10^{4}$
Control titer (-E2)	$6.2x10^{5}$	$5.6x10^{4}$	$3.89x10^{5}$	$1.5x10^{5}$	$2.2x10^{3}$	$9.26x10^{4}$
Control titer (-E1)	*$2.8x10^{5}$	$5.6x10^{4}$	$2.95x10^{5}$	$1.35x10^{4}$	$8.0x10^{2}$	$1.06x10^{4}$
Control titer (-ub)	$6.7x10^{4}$	$2.0x10^{4}$	$3.54x10^{4}$	$2.4x10^{3}$	$4.7x10^{2}$	$1.06x10^{4}$

(B)	R1	R2	R3	R4	R5	R6
Input titer/Output titer	$5x10^{4}$	$2.16x10^{5}$	$4.9x10^{4}$	$8.2x10^{3}$	$3.25x10^{6}$	$6.2x10^{5}$
Enrichment (E2)	4	3	2	13	4	0.5
Enrichment (E1)	8.5	3	3	150	10	5
Enrichment (ub)	35	8.8	20	800	17	5

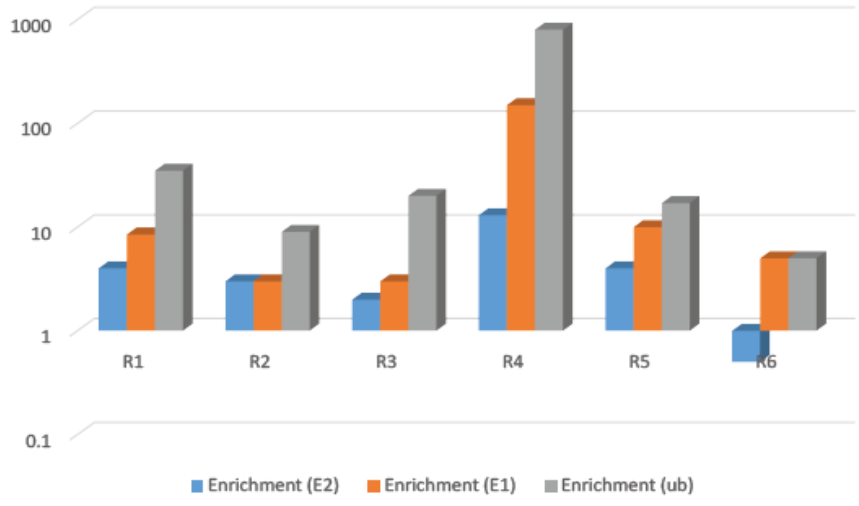

Table 3-1 Selection titers (first attempt)

The selection conditions and titer numbers were as given in (A) for the six rounds of selection; the approximate enrichment values were given in (B).

*The particular control was done in the absence of both E1 and E2 enzymes.

We constantly monitored the enrichment of the library (defined as the number of eluted phage particles from the ubiquitination reaction divided by the number of eluted phage particles from each of the controls) from round to round, and we either maintained or increased the

stringency of the selection conditions (by either decreasing the input phage number, or decreasing ubiquitin, E1, E2 enzyme concentrations) as the selection entered its latter rounds, for up to six rounds (**Table 3-1A**).

While the enrichment seemed to be steadily increasingly in the first few rounds, reaching the peak numbers in the fourth—which was probably our most productive round, towards the end of the selection the enrichment numbers seemed to have sharply decreased (**Table 3-1B**).To investigate why this was so, the DNA of remaining mutants in the last round were then sent for sequencing. Much to our chagrin, our sequencing results indicated that the surviving mutants were "junk", meaning they were either frameshift or truncated mutants, mutants which have lost the insert altogether, or mutants with recombination of sequences from bacterial genomic DNA (data not shown), in other words clones which had no activity, but might possess certain advantages in the form of faster growth and/or more efficient infection[47]. To ascertain whether or not there were still any "useful" surviving mutants, we PCR-amplified the E4B U-box sequence from the library DNA of the fourth, fifth and final rounds using primers specific for the E4B gene. The PCR product was then digested and religated into the original library backbone, and the resulting clones were sent for sequencing using phagemid-specific primers. The results indicate that among the "surviving" mutants, one particular clone, which we have given the name KB1, dominated, comprising up to 70% of all sequenced mutants in the final round. Four other mutants of note which were identified among these three rounds, named KB2-KB5, showed remarkable sequence similarity to each other and to KB1 (**Table 3-2**). Testing of these mutants, both on phage particles in similar biopanning assays as previous (**Figure 3-7B**), and as free expressed U-box protein from generic pET expression protocols, in ubiquitination assays confirmed that indeed they were ubiquitinated under selection conditions (**Figure 3-12 A&B**).

Position	1233	1236	1237	1238	1239	Frequency	
WT	R	L	M	D	T		
R4	R	I	L	H	T	5	KB1
	K	I	M	H	T	1	KB2
	I	I	T	Q	T	1	
	T	R	P	P	P	1	
	R	I	I	K	K	1	
R5	R	I	L	H	T	7	
	K	I	M	H	T	1	
	T	I	L	K	T	1	KB3
	K	I	L	H	T	1	KB4
R6	R	I	L	H	T	7	
	K	I	M	H	T	1	
	K	I	L	F	T	1	KB5
	T	P	P	Y	P	1	

Table 3-2
Sequence alignment of mutants (1st attempt)
Shown here are the amino acid sequences of remaining clones in the randomized positions 1233, and 1236-1239; Note the wild-type sequence (blue text), and the strong consensuses on certain notable residues (red text). Mutants of particular interest were given the names of KB1-KB5 (rightmost column); these were further characterized.

A **Biopanning with elution-titration**

	KB1	KB2	KB3
Input phage titer	1.8×10^9	4.4×10^9	6.6×10^9
Output phage titer (E1, E2, ub)	3.8×10^6	1.46×10^7	6.0×10^6
Control output titer (-E2)	2.1×10^6	7.6×10^6	3.2×10^6
Control output titer (-E1)	1.4×10^4	8.0×10^3	1.4×10^4
Enrichment (E2)	1.8	1.9	1.9
Enrichment (E1)	270	1800	430
Input/Output	470	300	1100

0.05 μM bE1, 5 μM xE2, 0.04 μM (bio)wt-ubiquitin

B **IB: FLAG (E4B U-box)**

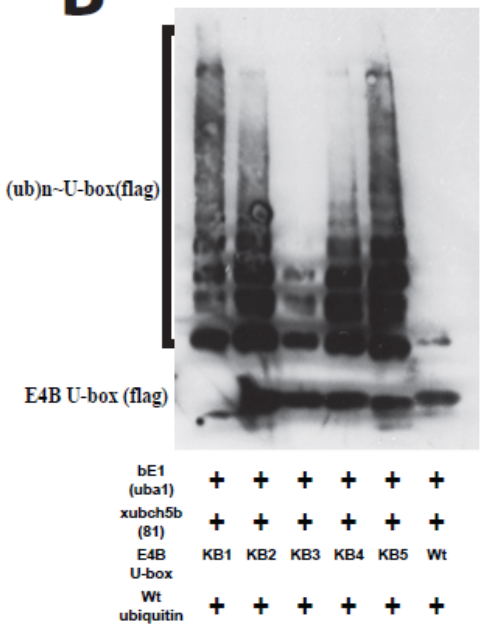

Figure 3-12. Testing of U-box mutants from selection (first attempt). Selected library mutants were tested in a biopanning assay similar to selection conditions (B), and recombinantly expressed as free proteins via generic pET systems with FLAG epitope tags; ubiquitination with bE1, xE2 and wild-type ubiquitin was tested by Western immunoblotting against FLAG, alongside with the wild-type U-box which showed virtually nonexistent activity (A).

We hence concluded that we had successfully converged the library through our selection and that these mutants were retained in the library because they were genuinely active; however, we decided to restart our selection from the beginning to determine if the library would also converge to similar sequences in our second run. Our second attempt very much mirrored our first attempt in protocol and strategy (**Table 3-3**).

(A)	R1	R2	R3	R4	R5	R6
[bE1] (μM)	1	0.1	0.02	0.1	0.06	0.06
[xE2 81] (μM)	10	5	10	5	10	5
[bio-wtub] (μM)	0.4	0.15	0.02	0.15	0.15	0.15
Input phage titer	1.3×10^{10}	3.4×10^{10}	2.4×10^{10}	3.7×10^{10}	2.0×10^{9}	1.0×10^{10}
Output titer (E1, E2, ub)	N/A*	3.1×10^{6}	6.1×10^{5}	3.7×10^{6}	4.7×10^{6}	9.4×10^{6}
Control titer (-E2)	N/A*	1.8×10^{6}	1.0×10^{5}	1.5×10^{5}	3.0×10^{4}	3.4×10^{4}
Control titer (-E1)	N/A*	2.2×10^{5}	1.0×10^{4}	1.7×10^{6}	1.5×10^{4}	2.8×10^{4}
Control titer (-ub)	N/A*	1.1×10^{5}	1.0×10^{4}	1.4×10^{4}	2.9×10^{3}	1.4×10^{4}

(B)	R1	R2	R3	R4	R5	R6
Input titer/Output titer	N/A*	1.1×10^{4}	3.9×10^{3}	1.0×10^{4}	425	1.1×10^{3}
Enrichment (E2)	N/A*	1.72	6.1	2.18	157	276
Enrichment (E1)	N/A*	14	61	24.7	313	335
Enrichment (ub)	N/A*	28	61	264	1620	671

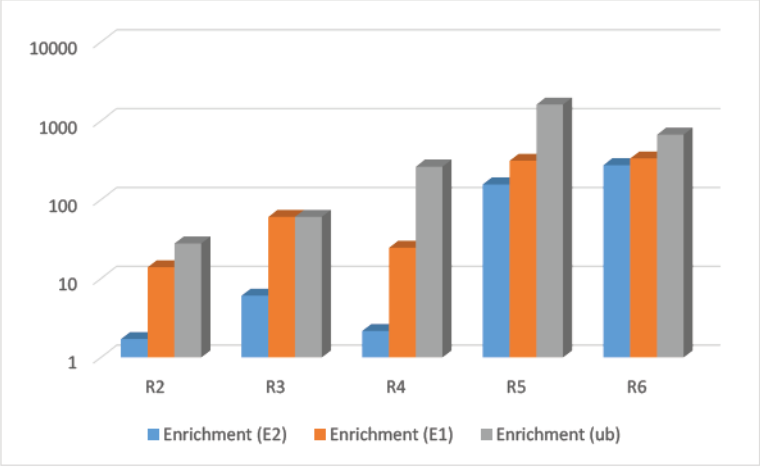

Table 3-3
Selection titers (Second attempt)
The selection conditions and titer numbers were as given in (A) for the six rounds of selection; the approximate enrichment values were given in (B). The library was purged of junk clones at the end of the 3rd round (hence, the selection conditions immediately afterwards became less stringent; note the drastically improved enrichment numbers after restoration of the library.)

*We did not expect any enrichment in the first round; output titer was not available.

In this attempt however, we were particularly watchful for the emergence of "junk" clones, which took over our library in our previous attempt; we screened the library phage in each round by western blotting for the signal of the C-terminal fused flag tag, which would only be displayed in clones which bore the full insert. When the signal began to fade after around three rounds of selection, we "purged" the library of junk clones by PCR amplification and religation, in a similar manner as was in our first attempt. Afterwards, we then continued the selection for three more rounds, making six rounds in total; the enrichment numbers from after library restoration were greatly improved, and the input/output ratio was much lower than our first attempt—which signified that a much greater percentage of phage particles were active. Clones from the fourth, fifth, and final rounds of selection were sequenced, and the results were as shown below (**Table 3-4 A-C**).

Position		1233	1236	1237	1238	1239	Frequency	
Wt		R	L	M	D	T		
R4	A	R	I	S	A	T	1	
		L	K	R	L	S	1	
		K	I	N	F	T	1	KB8
		T	Q	L	P	P	1	
		K	I	L	F	T	1	KB5
		R	P	G	A	P	1	
		I	I	M	F	T	1	
		A	H	W	D	F	1	
		M	I	I	N	R	1	
		L	I	L	F	T	1	
		R	I	M	Y	S	1	
R5	B	L	I	M	W	T	1	
		K	I	M	H	T	2	KB2
		K	I	L	F	T	3	KB5
		M	I	V	N	R	1	
		M	I	I	N	R	3	KB11
		I	I	M	R	T	1	KB9
		L	I	L	F	T	1	KB10

Table 3-4 Sequence alignment of mutants (2nd attempt)
Shown here are the amino acid sequences of clones in the 4th (A), 5th (B) and the final 6th (C) rounds; the wild-type sequence (blue text), and the consensus in certain positions (red text) are as given. The KB2 and KB5 mutants from the first attempt were also selected; other mutants of interest were given the names of KB6-KB12 (rightmost column.)

Position		1233	1236	1237	1238	1239	Frequency	
R6	C	K	I	M	H	T	7	KB2
		K	I	L	F	T	3	KB5
		L	I	T	Y	T	2	
		L	I	L	F	T	2	KB10
		I	I	M	R	T	2	KB9
		K	I	N	F	T	2	KB8
		K	I	S	Y	T	2	KB7
		K	I	M	Y	T	1	
		K	I	M	R	T	1	KB12
		R	I	Y	Y	T	1	KB6
		M	I	I	N	R	1	KB11

Table 3-4 (Continued)

We were delighted to discover that the sequences show remarkable convergence; this is especially true after the fifth and sixth rounds of selection; this was also reflected in our enrichment numbers which showed drastic improvement in the last two rounds, and was exceedingly superior to the numbers in our first attempt. Also, we were able to identify two mutants recovered our first selection: KB2 and KB5 as the two most abundant clones in the final round; this constituted strong evidence that our second attempt at selection was very successful, to a much greater extent than the first—the gradual, plainly-apparent convergence of the library from the fourth to sixth round is obvious, the general consensus among the surviving mutants in the last round is very strong—the unanimous consensus on the L1236I mutation—an activity enhancing mutation in general also reported by Starita et.al.[48] (to be discussed in the discussion section) at around the same time our selection was finished, and the virtually unanimous consensus to retain T1239 were especially convincing. The specific amino acids selected and how they most likely made an impact on the E2-E3 interface will be discussed in greater detail in the discussion section of this chapter.

3.2.5 Confirmation of OUT activity of selected mutants

Certain selected mutants which were tested, exhibited activity in a standard ubiquitination assay involving bE1, wtUB and xE2, which mimics the selection conditions (same as KB1-KB5 in Scheme 3-6), as well as with the completely orthogonal cascade comprising of full xE1 (AD+UFD), xUB and xE2; in contrast the wild-type E4B U-box had previously shown no appreciable activity with bE1, wtUB and xE2 (**Figure 3-12**) and it also did not show any detectable activity when charged with xE1, xUB and xE2. (**Figure 3-13**).

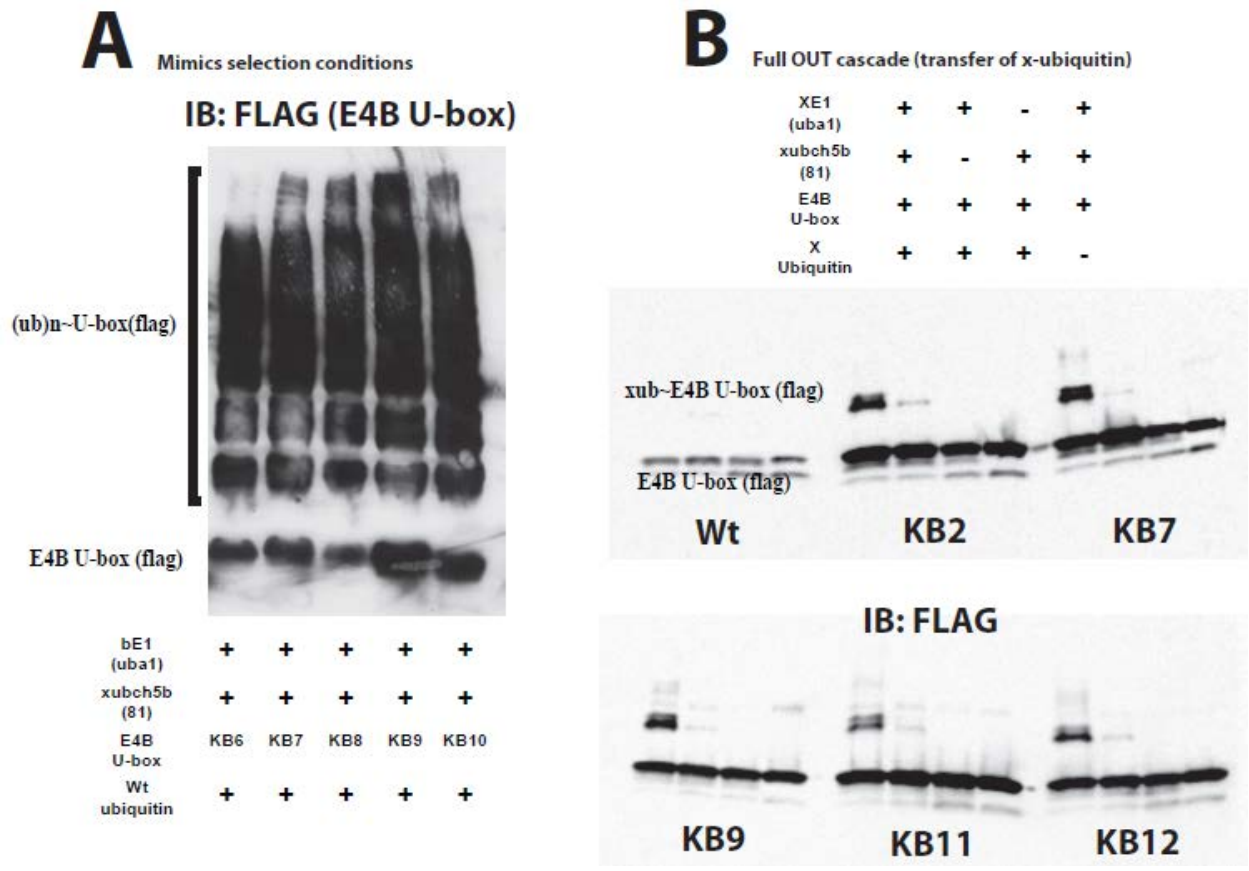

Figure 3-13. Testing of U-box mutants from selection (second attempt).
Certain selected mutants were tested for ubiquitination under conditions mimicking that in the library selection (see Scheme 3-6), with bE1, xE2 and wt-ubiquitin (A), and for ubiquitination with the full OUT cascade involving xub through XE1 and xE2, with controls each skipping one crucial component of the cascade, namely, from left to right, xE2, XE1, and x-ubiquitin (B); All the immunoblots were probed for the FLAG epitope, which was fused to the C-termini of the U-box proteins.

3.3 Discussion

3.3.1 Success of E4B display

The fact that the U-box of the ubiquitin ligase E4B was able to be successfully displayed on the phage particle and that the displayed protein is active in *in vitro* autoubiquitination assays on the phage particle implied that the enzyme was especially unique among the E3 tested, and this might have been for a handful of reasons—firstly, the U-box domain is relatively small, and possesses standalone activity in autoubiquitination assays; the active E4B U-box is around 75 amino acid residues in length[49], in contrast to the much larger HECT domains (which are typically ~350 residues)[50] which approach the general upper size limit of proteins which can be efficiently displayed by pIII—although there are some known exceptions to the rule[19].

Secondly, E4B is one of the handful of E3s known to efficiently function as monomers[40] (examples of other such E3s are the RING-type Cbl ligases[51]); the majority of RING-type E3s, as well as the U-box CHIP, have been known to function either as heterodimers or homodimers in their active forms—the interaction of the dimeric ligase with E2 is believed to be asymmetric: in general one of the monomers interacts with the E2 enzyme, while the other monomer serves to stimulate ubiquitin release[51], examples of such are the MDM2/MDMX[52], BRCA1/BARD1[53] heterodimers, as well as the CHIP homodimer[54]—all of our previous attempts displayed these ligases as monomers on the phage particle for various reasons including the generally accepted display size limit.

Lastly, the E4B U-box protein had been shown to be efficiently expressed in active form using generic recombinant *E.coli* protein expression systems—that the protein is compatible with being expressed through a bacterial system, the same system which is required for phage display;

in contrast, we have anecdotal as well as our own experimental evidence that many E3 ligases, especially RING-types, which contain zinc fingers, are not well-expressed in generic bacterial systems.

3.3.2 Selection challenges and solutions

Both attempts at our selection encountered the same problem: the emergence of "junk" clones—clones which have lost the U-box gene altogether or contain frameshifts or internal stop codons and hence show no signal when probed against the C-terminal affinity tag; theoretically these clones could arise through many different possibilities including recombination and DNA instability of the phagemid due to multiple rounds of amplification. Also, outside contamination from wild-type phage was possible; generally, such contamination is undetectable until the contaminants comprise a significant proportion of the library.

According to our protocol, the library was amplified in the presence of high concentration IPTG to induce overexpression of the U-box, with the intention of enhancing the level of display for more effective selection, however this came at the cost of intensifying the growth advantage of the bacterial cells bearing the junk clones, since the overexpressed pIII protein is believed to be toxic to the cell. Additionally, it has been reported that smaller phagemids are more rapidly packaged into phage particles; clones which bear deletions either in the target gene insert or elsewhere in the phagemid have a numerical advantage during propagation. Finally, as stated briefly earlier, clones which bear the wild-type pIII as opposed to those whose pIII proteins are fused to other genes are also more efficient in the infection process[19]. Hence, if the selection pressure in the biopanning assays was not strong enough to completely eliminate the junk clones from the library; though the "good" clones may be enriched through the assay—and we have good reason to believe so given our results, the enrichment might be more than offset by these

other advantages of the junk clones. In our second attempt at selection, we observed that our library started to become dominated by junk clones in as few as three rounds of selection, however, once the library DNA was reamplified and religated at the start of the fourth round of selection, we no longer detected any junk clones in the following three rounds carried out with similar protocols; this may be due to the fact that the representation of the active clones in the library was high enough at the later rounds to compete for all of the binding bait and prevent any junk clones from surviving in the biopanning process, and/or the fact that the stringency of the selection was increased in the latter rounds, eliminating all but the very most active mutants. In contrast the representation of active clones at the start of the selection was extremely low, while the selection conditions were more generous, hence there was greater possibility of junk clones being retained during the biopanning process, these clones then would utilize their growth advantages in such a way that the abundance of active clones were constantly kept down at a level where the junk clones could persist from round to round and eventually dominate the library if there had not been any intervention. This information may be useful to keep in mind for future such library selections.

Additionally, we also observed a significant amount of background signal which was generated when the phage particles were reacted with bE1 and ubiquitin in the absence of E2 (the enrichment based off of this control had mostly been stuck in the single digits throughout even the latter rounds of the first selection—although if such enrichment could be continuously maintained throughout several rounds, the vast majority of remaining mutants should still be those selected for activity with xE2; the background had also been similarly high, although to a slightly lesser extent, when biopanning was done with phage particles bearing wild-type E4B U-box with wild-type enzymes; however, when the free U-box proteins were expressed and tested,

neither the wild-type U-box nor the selected U-box mutants showed any semblance of detectable activity when the E2 enzyme is absent; we speculate that the E1-ub complex may possess some transthiolation activity with certain nucleophilic residues on the surface of the phage particle itself, such as the cysteine residues on the displayed Jun and Fos proteins. This problem was somewhat alleviated in the latter rounds of the second attempt of selection by including BSA in the reaction buffer.

3.3.3 Sequences of selected clones

Through our directed evolution approach with our randomized library, we have enriched mutants which have shown appreciable activity with our OUT system. We observed a very obvious consensus among the amino acid sequence of the selected mutants; as would be expected, the activity levels of all the mutants were quite similar (see **Figures 3-12, 3-13**).

Of all the residues in the E4B U-box which interact with the E2 N-terminal helix[42] (See **Figure 3-14A-C**), D1238 seemed to play the most prominent role by far, interacting with K4 and K8 of E2 (**Figure 3-14A**); upon mutation of these Lys residues to Asp, and Glu, respectively in xE2, this was replaced with unfavorable repulsive interactions. As would be expected, none of the mutants retained in our library bore any negatively-charged residues at position 1238—some mutants bore the slightly basic histidine substitution at this position (KB1 and KB2, which were the most abundant mutants in the first and second selection attempts, respectively) while some others had the more-strongly basic Lys (KB3) or Arg (KB9 and KB12)—it is quite reasonable to hypothesize that these residues restore the activity of the U-box with xE2 via favorable electrostatic interactions or hydrogen bonding; and yet still a good portion of the surviving mutants had the aromatic phenylalanine or tyrosine substitution at this position; these might have interacted with certain hydrophobic residues elsewhere in the E2 enzyme without affecting the

binding interface through any inappropriate steric hindrance; and yet it might have simply been the case that merely removing the unfavorable electrostatic repulsion was sufficient to restore a decent level of activity.

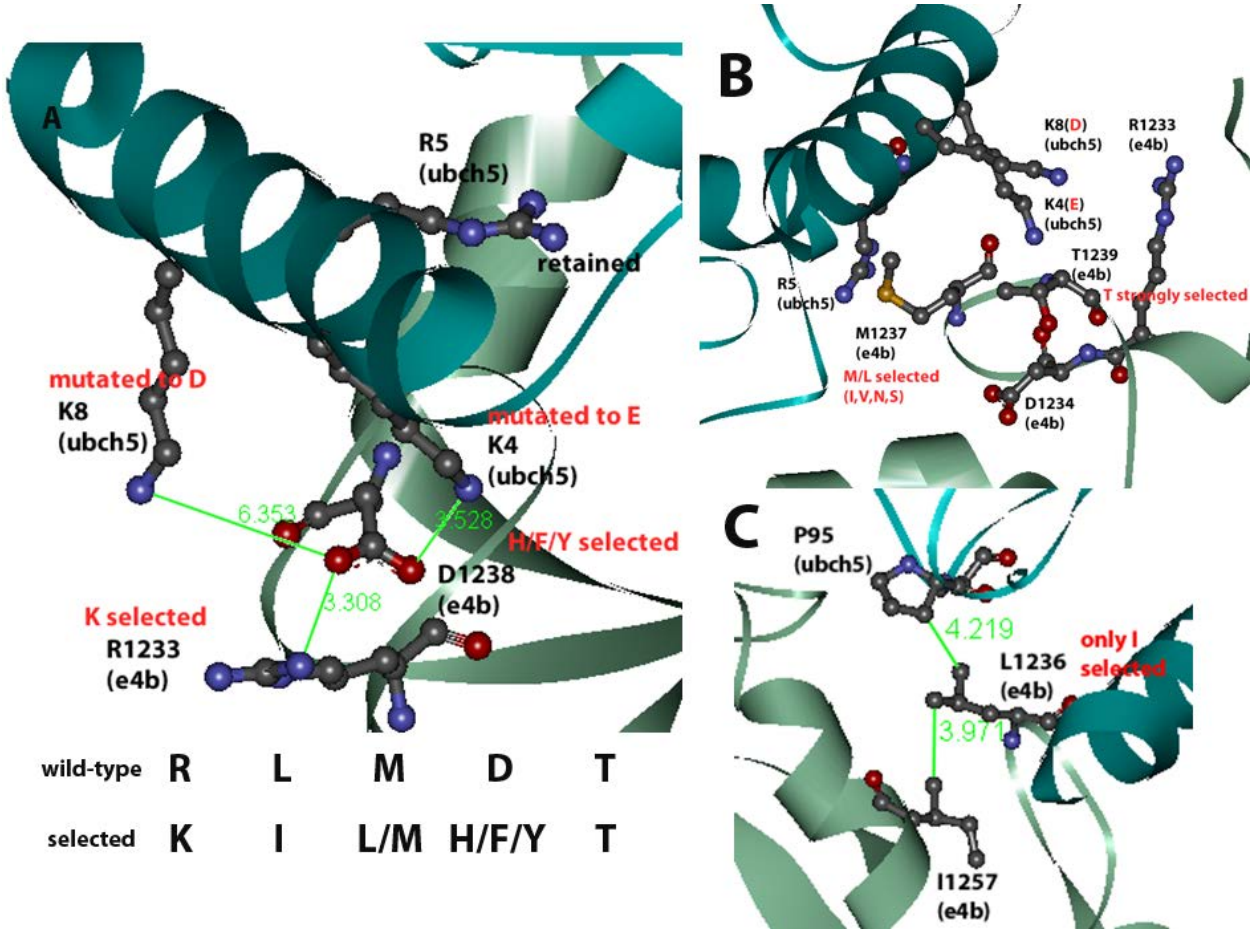

Figure 3-14 Analysis of effects of library mutations. The crystal structure of the ubch5-E4B U-box complex showing residues in N-terminal helix of ubch5 interacting with residues in the loop region of E4B; distances between atoms (green text), and specific mutations in xE2, along with those from the library consensus sequences (red text) are shown: (A) the potential salt bridges between K4 and K8 with D1238, (B) the close proximity between R5 and M1237, as well as the potential intramolecular hydrogen bond between T1239 and D1234, and (C) the considerable distance between L1236 and its most likely interacting partners.

Looking at other residues, the wild-type T1239 (**Figure 3-14B**) residue was conserved in virtually all of the mutants; the side chain hydroxyl being likely involved in intramolecular

hydrogen bonding with the backbone amide of D1234 which probably was crucial in retaining the active structure of the protein. In a similar manner, R1233 was retained in some mutants (KB1 and KB6), or replaced in the majority of surviving mutants with the similarly positive Lys residue; looking at the crystal structure, while this position could likely interact intramolecularly with D1238 in the wild-type enzyme (and hence in our mutants this residue could potentially interact with the substituted amino acid at position 1238, whether it be through hydrogen bonding with histidine, or cation-pi interactions with Phe or Tyr residues), it was questionable if this position was sufficiently in close enough proximity or properly aligned to interact with the N-terminal helix to interact with the negative residues in xE2 (**Figure 3-14A**), although that still remained a remote possibility—one would do well to remember that crystal structures are helpful, yet not infallibly so, and it might very well be that these mutants possess significantly different binding conformations with E2 from that of the wild-type U-box; as long as the autoubiquitination activity was still retained, they would still be selected regardless.

M1237 was probably the position with the most variation; while Met was retained in a handful of mutants, the enzyme also seemed to prefer Leu at this position, although it seemed to tolerate a wide variety of different residues, both hydrophobic and hydrophilic—the side chain of position could conceivably be in close proximity enough to interact with R5, which was not mutated in the 81 variant of xE2 used for this selection (**Figure 3-14B**), however, from our results there seemed to be no stringent requirement for this position, other than the avoidance of bulky or charged residues.

Lastly, in stark contrast, the L1236I mutation (**Figure 3-14C**) was a unanimous consensus among the mutants. This was somewhat surprising, given the fact that leucine and isoleucine are extremely similar residues, and yet this mutation seemed to be an absolute

requirement for any surviving mutant. While this particular position may interact with other particular hydrophobic residues, it was not obvious why other similar hydrophobic residues would not be tolerated, and generally hydrophobic interactions, especially at the distances shown in the crystal structure would be virtually nonexistent, and hence, in magnitude much weaker than, for example, hydrogen bonds or salt bridges; one would not expect the level of impact this residue would have in the selection of the library.

Interestingly enough, Starita et.al. have reported this L to I mutation (this particular leucine was referred to as L1107 in their numbering) as an activity enhancing mutation in their findings[48] also involving phage display and library selection of the E4B U-box, albeit with wild-type enzymes, and with mutations occurring throughout the whole U-box. This was the only activity-enhancing mutation reported in this loop region by this study; it is reported that this mutation enhanced the U-box's affinity for the E2 enzyme by up to 6 fold, although no precise structural explanation was proposed. It is exceedingly difficult to conceive of a scenario where a simple shift of a methyl group would cause such a big difference—although one may be well inclined to think of some very specific steric effect as the cause, although this is nowhere implied in the available crystal structure of the complex.

Nonetheless, in recent findings at the time of writing, it is believed that RING and U-box E3s not only act to recruit substrates on behalf of E2 for ubiquitination, but serves to allosterically prime the E2-ub thioester linkage for release: a "mysterious activation"[55]—it is believed that once the E3 binds allosterically to the E2, it restricts the movement of the tethered ubiquitin, driving it to be more likely in the so-called "closed" conformation, after this there is believed to be a shift in the binding of different components of the whole complex upon actual ubiquitination that is still poorly understood[1] (see **Figure 3-1**). It is likely that there is more to

the E2-E3 interaction than what is clearly apparent in the crystal structure; it may be that the active conformation of the complex exists in a very unstable and ephemeral manner which is extremely difficult to visualize—and that conformation is significantly different from that shown in the crystal structure; this notion is supported by several point mutagenesis studies on various different E3s—many of the reported mutations make little sense when viewed in the light of existing crystal structures.

In summary, the sequences of the selected mutants in part agreed with our initial expectations based off published crystal structures between the U-box of E4B—D1238 was mutated to positive or non-negative residues to counteract the mutations of lysine residues in the N-terminal helix, while generally other residues which do not seem to engage with the N-terminal helix have been retained or mutated to similar residues in the majority of surviving mutants—and yet in the other part also provided some rather surprising discoveries such as the activity enhancing L1236I mutation. This underscores the fact that there is more to the E2-E3 interface than meets the eye, and also the fact that when perfect knowledge is unavailable, is where the true strength of the directed evolution approach shines through—we would not expect to have discovered the activity-enhancing L to I mutation through any modeling or rational design approach. Directed evolution has been effectively used in the engineering and reengineering or many different proteins in countless biological pathways[56-70], and we now have good reason to believe that the U-box domain is no exception. With these mutants, theoretically we have completed a path for OUT transfer; we have successfully constructed a tool, the next step is to put the tool to use. Our next step would be to confirm the activity of our OUT system on putative substrate proteins, and afterwards, to utilize our system to screen for novel substrates in the living cell. These approaches and their results will be discussed in the following chapters.

3.4 Methods

The DNA primers involved in this chapter are as given in **Table 3-5** below.

Primer name	Sequence (5'-3')
Jun13	ACTTTTGCTTCCGGCTCGTATGT
Jun14	AATCAAAATCACCGGAACCAGAGC
Kar77	GGAAGTTCTGCGCGCCGAGCTCGGATCCATAGAGAAG
Kar78	TTTGATCTCGGTACCGTGGTCACTGCTCTG
Kar202	GGAAGTTCTGCGCGCCGAGCTCGGTTCCATAGAGAAG
Kar203	TCTCACGGGATCCGTCATCAGCGCAGCGGCTGCAGGG TCAGCGAACTCGTCCGGGGCATC
Kar204	CCTGCAGCCGCTGCGCTGATGACGGATCCCGTGAGACT GCCC
Karlib1	CCTTCTGGCACCGTCATGGAC
Karlib2	GTCCATGACGGTGCCAGAGGGTAACCTCACGGGATCC GTCATCAGMNNMNNMNNMNNAGGGTCMNNGAACTCGT CCGGGGCATC
Kar205	AGTGATCAGCGCGCCGAGCTCCATATGATAGAGAAGTT TAAACTTCTTGCAGAG
Kar206	TTCTAATGCGGCCGCGATCTCGAGCTTGTCATCGTCGTC CTTGTAGTCGGTACC

Table 3-5 Primers used in chapter 3.
Residues in red indicate restriction digest sites. Residues in blue indicate randomized positions, where N is any nucleotide and M is A or C. The reverse complement of MNN is NNK, where K is G or T; using NNK instead of NNN for randomized positions helps alleviate selection bias from certain amino acids which are encoded for by multiple codons

3.4.1 E4B U-box activity assays

3.4.1.1 pET E4B U-box (FLAG) construction and protein expression

The original E4B U-box construct which was used as the template was a gift from Rachel Klevit. The U-box domain of E4B was PCR-amplified by Kar205 and Kar206 and inserted via sticky end ligation into PET30-CHIP, which bore an N-terminal 6-His tag, in between NdeI and XhoI restriction sites.

Expression and 6-His affinity purification was done following our generic protocol previously described in chapter 2. The success of the expression was confirmed via western immunoblotting 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.

3.4.1.2 Autoubiquitination assay of recombinantly expressed E4B U-box

4 μM of E4B U-box (flag) protein was charged with 0.5 μM uba1, 3 μM ubch5a, and 5 M HA-tagged ubiquitin in a 50 μL reaction containing 50 mM Tris pH 7.5, 10 mM $MgCl_2$, 3 mM ATP and 50 μM DTT. The reaction was allowed to proceed at room temperature for 1 hour. SDS page of reaction and control was run and analyzed via western immunoblotting using mouse antibody against HA (to monitor the formation of ubiquitin chains by E4B).

3.4.2 Phage display of E4B

3.4.2.1 Construction of PJF E4B U-box

pET30a-E4B U-box wild-type was used as the template to PCR-amplify the U-box sequence with Kar77 and Kar 78 primers; the insert was then ligated into PJF3H, bearing the original insert of SV5V, via BssHII and KpnI sites; the original vector bore a flag-tag which had

been cloned in earlier at the C-terminal. PJF E4B-Ubox 5Ala mutant (library template) was constructed in a similar manner, involving the same two terminal primers Kar77 and Kar78 which were paired with two additional internal overlap extension primers Kar203 and Kar204, respectively

3.4.2.2 E4B phage propagation

The pJF E4B U-box wild-type phagemid DNA was then transformed into XL1-blue (Stratagene) 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 10 ml of 2YT medium supplemented with 2% glucose, 10 μg/ml tetracycline and 100 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 (Strategene) 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 3500 rpm for 10 minutes, and resuspended in 100 ml of fresh 2YT 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. 50 ml of filter-sterilized 5x PEG/NaCl solution (200 g/L PEG-8000 (Sigma), 2.5 M NaCl) was thoroughly mixed with the supernatant from the previous step, and the mixture was incubated on ice for 1 hour and afterwards centrifuged at 9000 rpm for 20 minutes 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 13000 rpm for 10 minutes. The resulting supernatant was then titrated by infection, and tested for the display of the E4B U-box by western immunoblotting against FLAG.

3.4.2.3 Autoubiquitination assay of E4B phage

Around > $2x10^{10}$ phage particles derived from the PJF E4B U-box-flag construct were charged with 0.5 μM uba1, 6 μM ubch5b and 14 μM HA-ubiquitin, in 50 μL reaction containing 50 mM Tris pH 7.5, 10 mM $MgCl_2$, 3 mM ATP, and 50 uM DTT and incubated for 1 hour at room temperature. The reaction quenched by boiling in Laemmli buffer containing 100 mM DTT, and analyzed via SDS page and immunoblotting using mouse antibody against FLAG.

3.4.2.4 Biopanning of E4B phage and quantification by infection-titration and ELISA

Around 10^8-10^9 phage particles displaying E4B U-box were charged with 1 μM uba1, 5 μM ubch5b and 0.2 μM N-terminal labelled biotin-ubiquitin (Boston Biochem) in 20 μL total reaction volume containing 50 mM Tris pH 7.5, 10 mM $MgCl_2$, 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 μL 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˚C. Quantification of infected cells was done by serial dilution and plating on selective agar media containing 100 μg/ml ampicillin. After incubation of the plates overnight at 37˚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).

3.4.3 E4B library selection

3.4.3.1 Model selection

E4B-displaying phage particles were mixed with phage particles displaying SV5V in varying ratios of 1:1, 1:10, 1:100 and 1:1000 (titration of phage preps were done via infection and colony-counting); the different mixes, each containing ~10^9 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.

3.4.3.2 Construction of the library

PCR amplification with the two terminal library primers Jun13 and Kar78, in combination with the two overlap extension primers Karlib2 and Karlib1, was performed using PJF E4B U-box 5Ala as the template. The backbone vector used was PJF SV5V—identical to the

one which was used to construct the PJF E4B U-box wild-type clone. The vector and insert PCR product were digested with EcoRI and KpnI restriction enzymes (NEB), gel purified using a standard gel extraction kit (QIAGEN), and further purified through phenol-chloroform extraction and ethanol precipitation. 5 μg of purified digested vector and 4 μg of purified digested library insert were ligated in 1 ml reaction volume using T4 DNA ligase (NEB) at room temperature overnight. The ligation mixture was again purified by ethanol purification, and the resulting DNA pellet was reconstituted into 75 μL water. Transformation of the reconstituted DNA was performed using electrocompetent SS320 cells (Strategene) in 14 reactions, each reaction containing 5 μL of reconstituted DNA and 95 μL cells; immediately after electroporation, the cells were diluted 10-fold into SOC medium and allowed to recover at 37C with vigorous shaking for 1 hour; afterwards, the cells were plated onto solid LB agar media containing 100 ug/ml ampicillin and 2% glucose and incubated overnight at 37C. The cells from all the plates were then collected, pooled together, and the starting library DNA was extracted using Maxi-prep (QIAGEN). Streaked single colonies were also sent for sequencing, as well as confirmed by colony-PCR using the Jun13/Jun14 primer pair.

3.4.3.3 Library phage propagation

An arbitrary amount of library DNA maxiprep or miniprep (depending on the round), which was approximated to yield at least 10 times as many transformants as the expected remaining library diversity, was electroporated into 50 volumes of electrocompetent XL1-blue cells (Agilent). The transformed cells were added to 20 volumes of SOC media and allowed to recover for 1 hour at 37˚C. Afterwards, the recovered culture was added to 50 mL of superbroth containing 100 μg/ml carbenicillin and 10 μg/ml tetracycline, and allowed to grow for 2 more hours at 37C. The culture was then poured into 200 mL of 2YT medium, while maintaining the

original ampicillin and tetracycline concentrations, and VCSM13 helper phage (Agilent) was then added to superinfect the cells at >10 fold multiplicity of infection (the total number of cells having been approximated by optical density), the cells were then allowed to shake at 37C for 2 additional hours. Finally, kanamycin was added to a final concentration of 70 ug/ml, and IPTG was added to a final concentration of 1 mM, the temperature was lowered to 30˚C, and the culture was allowed to grow overnight, under vigorous shaking (~300 rpm).

The next morning, phage purification by PEG precipitation was performed using standard protocols described earlier in 3.4.2.2. The purified library phage particles were suspended in sterile TBS buffer. The display of the library was tested periodically by western immunoblotting in the same manner as that described of wild-type phage against the C-terminal FLAG affinity tag.

3.4.3.4 Biopanning and selection of library phage

The exact reaction conditions in which phage particles displaying mutant E4B U-box domains were ubiquitinated with N-terminal biotin-ub (Boston Biochem), back mutant uba1 (bE1) and xE2 (clone 81) vary from round to round (**Tables 3-1A** and **3-3A**). All library ubiquitination reactions were performed for 1 hour at 37˚C. The different conditions in each round are as given in the table below. Rounds 4-6 of the second selection also involved a pre-incubation step, in which the unreacted phage particles were first incubated onto streptavidin plates to prevent selecting false positives—clones which merely have binding to either polystyrene or streptavidin, before the remaining unbound phage particles were used in the ubiquitination reactions.

Generally, completed reactions were diluted 10-fold into 0.1% BSA-TBST solution before being loaded onto streptavidin plates (Thermo). Controls skipping each component of the reaction similar to those described in the biopanning assay of the wild-type E4B phage were also set up. The reactions were allowed to incubate on the plates for 1-2 hours at room temperature, after which the plates were thoroughly washed ~30-50 times with 0.1% BSA-TBST. Elution was done by the addition of 100 μL of a solution containing 100 mM DTT in TBS per well, followed by a 15-minute incubation at room temperature. The eluted phage particles were added to 10 volumes of a high density culture (OD ~1.0-2.0) of XL1-blue cells in LB medium supplemented with 10 ug/ml tetracycline. The culture was incubated for 2 hours at 37˚C with slow (~100 rpm) shaking, before being plated onto LB-agar plates containing 100 μg/ml carbenicillin and 2% glucose and allowed to grow overnight at 37˚C. 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 as previously reported.

The cells were collected from the plates, and the selected library DNA was isolated using either DNA miniprep or maxiprep kits (QIAGEN). The selected library DNA was then used in propagating the next round of phage per protocol 3.4.3.3.

3.4.3.5 Rescue of the library via PCR amplification and religation

Isolated library DNA from the round of interest (as described in 3.4.3.3) was used as a template for PCR amplification using library primers; the PCR product was digested with EcoRI/KpnI, and further purified, before being ligated with previously prepared vector in an identical manner to that described in 3.4.3.2., save for the fact that the target number of

transformants was much lower, due to the library having lost diversity from previous rounds of selection. Transformed clones were screened by DNA sequencing and colony PCR, and the isolation of rescued library DNA was done by miniprep (QIAGEN). In our second attempt, the fourth round selection then proceeded from the recovered library DNA in a manner identical to that in 3.4.3.3.

3.4.4 OUT activity assays of E4B U-box mutants

3.4.4.1 Biopanning assays of selected mutants

The phagemid DNA of mutants KB1 through KB5, identified and isolated through sequencing, were used to propagate phage particles in a protocol previously described. In a much more stringent condition than the actual selection, around $2x10^9$ phage particles of each mutant U-box were charged with 0.05 μM bE1 (uba1), 5 μM xE2 (ubch5b clone 81), and 0.04 μM N-terminal biotin-ub (Boston Biochem) in 80 μL of 50 mM Tris pH 7.5, 10 mM $MgCl_2$, 3 mM ATP and 0.1% BSA, and incubated for 1 hour at 37˚C; separate controls skipping E1 and E2 were also set up for each mutant. 20 μL of 3% BSA-TBST was then added to each reaction, before loading onto streptavidin-coated plates; the rest of the protocol, including titration of cells infected by eluted phage particles follows that in 3.4.3.4.

3.4.4.2 Construction PET-E4B U-box mutants and protein expression

Using the selected phagemid DNA as template, PCR amplification, digestion and ligation of the PCR products into the digested expression vector, as well as the following protein

expression and purification steps are exactly the same as those of the protocol described in 3.4.1.1. The PET constructs of twelve mutants, KB1 through KB12, were all generated with C-terminal FLAG epitopes.

3.4.4.3 OUT ubiquitination assay of selected mutant U-box proteins

20 μM of E4B U-box-flag mutant proteins were charged with 1μM E1, 10 μM E2, and 50 μM HA-ubiquitin in 30 μL of 50 mM Tris pH 7.5, 10 mM $MgCl_2$ and 1.5 mM ATP. Reactions were incubated at 37˚C, for 15 minutes when using bE1 and wild-type ubiquitin, and for 2 hours when using XE1 and xub. Autoubiquitination activity was detected via western immunoblotting against the C-terminal FLAG tag of the U-box mutants.

3.5 References

1 Pruneda, J. N. *et al.* Structure of an E3:E2~Ub complex reveals an allosteric mechanism shared among RING/U-box ligases. *Molecular cell* **47**, 933-942, doi:10.1016/j.molcel.2012.07.001 (2012).

2 Voigt, C. A., Kauffman, S. & Wang, Z. G. Rational evolutionary design: the theory of in vitro protein evolution. *Advances in protein chemistry* **55**, 79-160 (2000).

3 Dalby, P. A. Strategy and success for the directed evolution of enzymes. *Current opinion in structural biology* **21**, 473-480, doi:10.1016/j.sbi.2011.05.003 (2011).

4 Arnold, F. H. Directed Evolution: Creating Biocatalysts for the Future. *Chemical Engineering Science* **51**, 5091-5102 (1996).

5 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, doi:10.1016/j.sbi.2005.07.006 (2005).

6 Kuchner, O. & Arnold, F. H. Directed evolution of enzyme catalysts. *Trends in biotechnology* **15**, 523-530, doi:10.1016/s0167-7799(97)01138-4 (1997).

7 McCullum, E. O., Williams, B. A., Zhang, J. & Chaput, J. C. Random mutagenesis by error-prone PCR. *Methods in molecular biology (Clifton, N.J.)* **634**, 103-109, doi:10.1007/978-1-60761-652-8_7 (2010).

8 Lutz, S. Beyond directed evolution--semi-rational protein engineering and design. *Current opinion in biotechnology* **21**, 734-743, doi:10.1016/j.copbio.2010.08.011 (2010).

9 Ehrlich, G. K., Berthold, W. & Bailon, P. Phage display technology. Affinity selection by biopanning. *Methods in molecular biology (Clifton, N.J.)* **147**, 195-208, doi:10.1385/1-59259-041-1:195 (2000).

10 Piggott, A. M. & Karuso, P. Identifying the cellular targets of natural products using T7 phage display. *Natural product reports*, doi:10.1039/c5np00128e (2016).

11 Young, K. H. Yeast two-hybrid: so many interactions, (in) so little time. *Biology of reproduction* **58**, 302-311 (1998).

12 Joung, J. K., Ramm, E. I. & Pabo, C. O. A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc Natl Acad Sci U S A* **97**, 7382-7387, doi:10.1073/pnas.110149297 (2000).

13 Carlson, J. C., Badran, A. H., Guggiana-Nilo, D. A. & Liu, D. R. Negative selection and stringency modulation in phage-assisted continuous evolution. *Nature chemical biology* **10**, 216-222, doi:10.1038/nchembio.1453 (2014).

14 Liao, H., McKenzie, T. & Hageman, R. Isolation of a thermostable enzyme variant by cloning and selection in a thermophile. *Proc Natl Acad Sci U S A* **83**, 576-580 (1986).

15 Stemmer, W. P. Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**, 389-391, doi:10.1038/370389a0 (1994).

16 Smith, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science (New York, N.Y.)* **228**, 1315-1317 (1985).

17 Boder, E. T. & Wittrup, K. D. Yeast surface display for screening combinatorial polypeptide libraries. *Nature biotechnology* **15**, 553-557, doi:10.1038/nbt0697-553 (1997).

18 Kenrick, S. A. & Daugherty, P. S. Bacterial display enables efficient and quantitative peptide affinity maturation. *Protein engineering, design & selection : PEDS* **23**, 9-17, doi:10.1093/protein/gzp065 (2010).

19 Barbas, C. F., Burton, D. R., Scott, J. K. & Silverman, G. J. *Phage Display: A Laboratory Manual*. (Cold Spring Harbor Laboratory Press, 2001).

20 Marvin, D. A., Symmons, M. F. & Straus, S. K. Structure and assembly of filamentous bacteriophages. *Progress in biophysics and molecular biology* **114**, 80-122, doi:10.1016/j.pbiomolbio.2014.02.003 (2014).

21 Moon, J. S. *et al.* M13 Bacteriophage-Based Self-Assembly Structures and Their Functional Capabilities. *Mini-reviews in organic chemistry* **12**, 271-281, doi:10.2174/1570193x1203150429105418 (2015).

22 Marvin, D. A. Filamentous phage structure, infection and assembly. *Current opinion in structural biology* **8**, 150-158 (1998).

23 Beck, E. & Zink, B. Nucleotide sequence and genome organisation of filamentous bacteriophages fl and fd. *Gene* **16**, 35-58 (1981).

24 Hill, D. F. & Petersen, G. B. Nucleotide sequence of bacteriophage f1 DNA. *Journal of virology* **44**, 32-46 (1982).

25 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).

26 Fagerlund, A., Myrset, A. H. & Kulseth, M. A. Construction of a filamentous phage display peptide library. *Methods in molecular biology (Clifton, N.J.)* **1088**, 19-33, doi:10.1007/978-1-62703-673-3_2 (2014).

27 Loset, G. A., Bogen, B. & Sandlie, I. Expanding the versatility of phage display I: efficient display of peptide-tags on protein VII of the filamentous phage. *PloS one* **6**, e14702, doi:10.1371/journal.pone.0014702 (2011).

28 Nilssen, N. R. *et al.* DeltaPhage--a novel helper phage for high-valence pIX phagemid display. *Nucleic acids research* **40**, e120, doi:10.1093/nar/gks341 (2012).

29 Frenzel, A., Kugler, J., Wilke, S., Schirrmann, T. & Hust, M. Construction of human antibody gene libraries and selection of antibodies by phage display. *Methods in molecular biology (Clifton, N.J.)* **1060**, 215-243, doi:10.1007/978-1-62703-586-6_12 (2014).

30 Scott, J. K. & Smith, G. P. Searching for peptide ligands with an epitope library. *Science (New York, N.Y.)* **249**, 386-390 (1990).

31 Cwirla, S. E., Peters, E. A., Barrett, R. W. & Dower, W. J. Peptides on phage: a vast library of peptides for identifying ligands. *Proc Natl Acad Sci U S A* **87**, 6378-6382 (1990).

32 Barbas, C. F., 3rd, Kang, A. S., Lerner, R. A. & Benkovic, S. J. Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proc Natl Acad Sci U S A* **88**, 7978-7982 (1991).

33 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, doi:10.1016/j.jmb.2012.01.038 (2012).

34 Kramer, R. A. *et al.* A novel helper phage that improves phage display selection efficiency by preventing the amplification of phages without recombinant protein. *Nucleic acids research* **31**, e59 (2003).

35 Yang, W. P. *et al.* CDR walking mutagenesis for the affinity maturation of a potent human anti-HIV-1 antibody into the picomolar range. *Journal of molecular biology* **254**, 392-403 (1995).

36 Rader, C. & Barbas, C. F., 3rd. Phage display of combinatorial antibody libraries. *Current opinion in biotechnology* **8**, 503-508 (1997).

37 Crameri, R. & Suter, M. Display of biologically active proteins on the surface of filamentous phages: a cDNA cloning system for selection of functional gene products linked to the genetic information responsible for their production. *Gene* **137**, 69-75 (1993).

38 Kerppola, T. K. & Curran, T. Fos-Jun heterodimers and Jun homodimers bend DNA in opposite orientations: implications for transcription factor cooperativity. *Cell* **66**, 317-326 (1991).

39 Zhao, B. *et al.* Orthogonal ubiquitin transfer through engineered E1-E2 cascades for protein ubiquitination. *Chem Biol* **19**, 1265-1277, doi:10.1016/j.chembiol.2012.07.023 (2012).

40 Nordquist, K. A. *et al.* Structural and functional characterization of the monomeric U-box domain from E4B. *Biochemistry* **49**, 347-355, doi:10.1021/bi901620v (2010).

41 Ulane, C. M. & Horvath, C. M. Paramyxoviruses SV5 and HPIV2 assemble STAT protein ubiquitin ligase complexes from cellular components. *Virology* **304**, 160-166 (2002).

42 Benirschke, R. C. *et al.* Molecular basis for the association of human E4B U box ubiquitin ligase with E2-conjugating enzymes UbcH5c and Ubc4. *Structure (London, England : 1993)* **18**, 955-965, doi:10.1016/j.str.2010.04.017 (2010).

43 Mace, P. D. *et al.* Structures of the cIAP2 RING domain reveal conformational changes associated with ubiquitin-conjugating enzyme (E2) recruitment. *J Biol Chem* **283**, 31633-31640, doi:10.1074/jbc.M804753200 (2008).

44 Zhang, M. *et al.* Chaperoned ubiquitylation--crystal structures of the CHIP U box E3 ubiquitin ligase and a CHIP-Ubc13-Uev1a complex. *Molecular cell* **20**, 525-538, doi:10.1016/j.molcel.2005.09.023 (2005).

45 Yin, J., Lin, A. J., Golan, D. E. & Walsh, C. T. Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nature protocols* **1**, 280-285, doi:10.1038/nprot.2006.43 (2006).

46 Zhao, B. *et al.* Phage selection assisted by Sfp phosphopantetheinyl transferase-catalyzed site-specific protein labeling. *Methods in molecular biology (Clifton, N.J.)* **1266**, 161-170, doi:10.1007/978-1-4939-2272-7_11 (2015).

47 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).

48 Starita, L. M. *et al.* Activity-enhancing mutations in an E3 ubiquitin ligase identified by high-throughput mutagenesis. *Proc Natl Acad Sci U S A* **110**, E1263-1272, doi:10.1073/pnas.1303309110 (2013).

49 Patterson, C. A new gun in town: the U box is a ubiquitin ligase domain. *Science's STKE : signal transduction knowledge environment* **2002**, pe4, doi:10.1126/stke.2002.116.pe4 (2002).

50 Rotin, D. & Kumar, S. Physiological functions of the HECT family of ubiquitin ligases. *Nature reviews. Molecular cell biology* **10**, 398-409, doi:10.1038/nrm2690 (2009).

51 Dou, H., Buetow, L., Sibbet, G. J., Cameron, K. & Huang, D. T. Essentiality of a non-RING element in priming donor ubiquitin for catalysis by a monomeric E3. *Nature structural & molecular biology* **20**, 982-986, doi:10.1038/nsmb.2621 (2013).

52 Leslie, P. L., Ke, H. & Zhang, Y. The MDM2 RING domain and central acidic domain play distinct roles in MDM2 protein homodimerization and MDM2-MDMX protein heterodimerization. *The Journal of biological chemistry* **290**, 12941-12950, doi:10.1074/jbc.M115.644435 (2015).

53 Brzovic, P. S., Rajagopal, P., Hoyt, D. W., King, M. C. & Klevit, R. E. Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. *Nature structural biology* **8**, 833-837, doi:10.1038/nsb1001-833 (2001).

54 Qian, S. B. *et al.* Engineering a ubiquitin ligase reveals conformational flexibility required for ubiquitin transfer. *The Journal of biological chemistry* **284**, 26797-26802, doi:10.1074/jbc.M109.032334 (2009).

55 Soss, S. E., Klevit, R. E. & Chazin, W. J. Activation of UbcH5c~Ub is the result of a shift in interdomain motions of the conjugate bound to U-box E3 ligase E4B. *Biochemistry* **52**, 2991-2999, doi:10.1021/bi3015949 (2013).

56 Axarli, I. *et al.* Directed evolution of Tau class glutathione transferases reveals a site that regulates catalytic efficiency and masks co-operativity. *The Biochemical journal* **473**, 559-570, doi:10.1042/bj20150930 (2016).

57 Bai, S., Wallis, J. G., Denolf, P. & Browse, J. Directed evolution increases desaturation of a cyanobacterial fatty acid desaturase in eukaryotic expression systems. *Biotechnology and bioengineering*, doi:10.1002/bit.25922 (2016).

58 Cheng, Q., Gao, H. & Hu, N. A trehalase from Zunongwangia sp.: characterization and improving catalytic efficiency by directed evolution. *BMC biotechnology* **16**, 9, doi:10.1186/s12896-016-0239-z (2016).

59 Jiang, H., Zhang, S., Gao, H. & Hu, N. Characterization of a cold-active esterase from Serratia sp. and improvement of thermostability by directed evolution. *BMC biotechnology* **16**, 7, doi:10.1186/s12896-016-0235-3 (2016).

60 Karpinski, J. *et al.* Directed evolution of a recombinase that excises the provirus of most HIV-1 primary isolates with high specificity. *Nature biotechnology*, doi:10.1038/nbt.3467 (2016).

61 Kawakami, T., Ogawa, K., Hatta, T., Goshima, N. & Natsume, T. Directed Evolution of a Cyclized Peptoid-Peptide Chimera against a Cell-Free Expressed Protein and Proteomic Profiling of the Interacting Proteins to Create a Protein-Protein Interaction Inhibitor. *ACS chemical biology*, doi:10.1021/acschembio.5b01014 (2016).

62 Molina-Espeja, P. *et al.* Synthesis of 1-Naphthol by a Natural Peroxygenase Engineered by Directed Evolution. *Chembiochem : a European journal of chemical biology* **17**, 341-349, doi:10.1002/cbic.201500493 (2016).

63 Oshima, S. *et al.* ASP2408 and ASP2409, novel CTLA4-Ig variants with CD86-selective ligand binding activity and improved immunosuppressive potency, created by directed evolution. *Protein engineering, design & selection : PEDS*, doi:10.1093/protein/gzw002 (2016).

64 Pandey, R. P., Parajuli, P., Koffas, M. A. & Sohng, J. K. Microbial production of natural and non-natural flavonoids: Pathway engineering, directed evolution and systems/synthetic biology. *Biotechnology advances*, doi:10.1016/j.biotechadv.2016.02.012 (2016).

65 Schutz, M. *et al.* Directed evolution of G protein-coupled receptors in yeast for higher functional production in eukaryotic expression hosts. *Scientific reports* **6**, 21508, doi:10.1038/srep21508 (2016).

66 Shivange, A. V., Roccatano, D. & Schwaneberg, U. Iterative key-residues interrogation of a phytase with thermostability increasing substitutions identified in directed evolution. *Applied microbiology and biotechnology* **100**, 227-242, doi:10.1007/s00253-015-6959-5 (2016).

67 Wang, M., Yu, C. & Zhao, H. Directed evolution of xylose specific transporters to facilitate glucose-xylose co-utilization. *Biotechnology and bioengineering* **113**, 484-491, doi:10.1002/bit.25724 (2016).

68 Xu, X., Liu, M. Q., Huo, W. K. & Dai, X. J. Obtaining a mutant of Bacillus amyloliquefaciens xylanase A with improved catalytic activity by directed evolution. *Enzyme and microbial technology* **86**, 59-66, doi:10.1016/j.enzmictec.2016.02.001 (2016).

69 Zhang, K. *et al.* Characterization and directed evolution of BliGO, a novel glycine oxidase from Bacillus licheniformis. *Enzyme and microbial technology* **85**, 12-18, doi:10.1016/j.enzmictec.2015.12.012 (2016).

70 Zhu, L., Wu, Z., Jin, J. M. & Tang, S. Y. Directed evolution of leucine dehydrogenase for improved efficiency of L-tert-leucine synthesis. *Applied microbiology and biotechnology*, doi:10.1007/s00253-016-7371-5 (2016).

Chapter 4

OUT cascade through E3 mutants and *in vitro* verification of activity

4.1 Background information

4.1.1 Rationale

Through the directed evolution approach, our phage selection process has yielded us U-box mutants of the E3 ligase E4B which possess the capability to autoubiquitinate with xUB through xE1 and xE2; this suggested that to some extent we had successfully restored E3 binding with the charge-reversed residues of xE2—obviously, since our goal was to use OUT for substrate identification, we would need to confirm the activity of the xE2-xE3 complex on substrate ubiquitination; it was not readily apparent if substrate ubiquitination would be affected by the mutations on the E3 U-box, although conventional wisdom would imply that substrate interaction involves other domains outside the U-box—members in the U-box family generally possess remarkable structural similarity as far as the U-box domain is concerned, but show great variation as to the other protein-interaction modules the U-box is fused to, such as the TPR[1], WD40[2] and armadillo domains[3]. This chapter discusses the testing of activity of xE4B in substrate ubiquitination, as well as potential application to another somewhat fundamentally different U-box, CHIP.

4.1.2 The U-box family of E3s

When compared to the existing information on HECT and RING families, relatively little is known about the U-box family of E3s, which received its name from its first formally-characterized member Ufd2[4], the yeast homolog of E4B. The U-box E3s comprise only around 3-4%[5] of characterized human ubiquitin ligases. The overall structure of approximately 75-amino

acid U-box domains bears striking similarities with RING domains, however, they lack the canonical zinc-coordinating cysteine and histidine residues of the RINGs[4]. As far as activity is concerned they resemble RINGs, in that in the presence of substrate they do not load ubiquitin onto themselves, but rather act as activating bridges between the E2 and substrate. At the time of writing, only a handful of U-box E3s have been comprehensively characterized; the overall domain composition of some of the better characterized enzymes are given as below by Marin[6] (**Figure 4-1**); there is great variety outside the U-box domain among these E3s, which suggest that different members of the family plays different roles in regulating different substrates, although in general it has been observed that many U-box enzymes play important roles in particular regarding protein quality control[7].

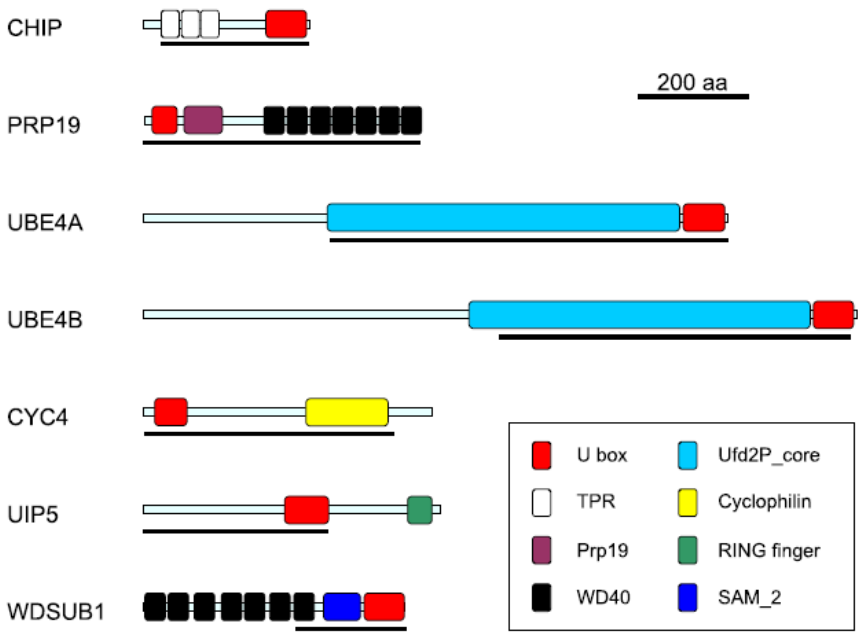

Figure 4-1
Domain compositions of U-box E3 ligases
While members of the U-box domain all contain the well-conserved U-box domain containing E3 ligase activity, there is great variation among the members in other regions of the proteins which is probably responsible for their different substrate preferences. UIP5 also possesses a RING finger, but only the U-box is believed to be active.

Certain U-box enzymes also possess unique idiosyncrasies which underscores the fact that there are truly fundamental differences in different members of this family, for example, there is excellent evidence that CHIP functions as an asymmetric dimeric complex, where only one monomer is active and the other monomer indispensably serves to stimulate ubiquitin

release[8], akin to the well-studied RING family dimers MDM2/MDMX and BRCA1/BARD1; another U-box, Prp19, has been proven to be active in the form of a homodimer of homodimers, a homotetramer[9]; in contrast, E4B and its homologs yeast Ufd2 and E4A, seem to possess sufficient activity to function as monomers[10], and are also, at the time of writing, the only U-box members definitively proven to possess the capability to function as E4 enzymes[11-13]—ubiquitin chain elongation factors; although recently CHIP has been proposed to be capable of acting as an E4 as well[14]; E4 enzymes act to elongate already-existing ubiquitin chains formed by other E3 enzymes[15]; this had been previously mentioned briefly in 1.2.

4.1.2.1 The E4B full-length structure and known biological roles

The construct of E4B, also known as mammalian Ufd2A (in contrast to yeast Ufd2), or as UBE4B[16] which we had previously dealt with in our phage selection and activity assays, was limited to the relatively small 75-amino acid U-box domain; the full length enzyme itself is appreciably much larger at around 1100 amino acid residues, and can be divided into three separate domains[10]: the N-terminal variable region, the armadillo-like repeats, and the U-box domain responsible for E3 activity. The N-terminal variable region shows non-existent conservation between homologs; the yeast Ufd2 N-terminal domain is over 200 amino acid residues shorter than the mammalian E4B; in contrast, the armadillo repeats comprising the vast majority of the protein's residues believed to be substrate-interacting domains, and the C-terminal U-box domains, are very well-conserved.

The crystal structure of full-length yeast Ufd2 containing all three domains, has been published[3], as shown below (**Figure 4-2**); in contrast, at the time of writing, existing structures of E4B have been limited only to the U-box domain; however one may safely assume that the overall E4B structure would be very similar to that of Ufd2, the only significant difference being

the variable N-terminal region. Modelling studies of E4B structure have generally relied on the Ufd2 structure as a guide.

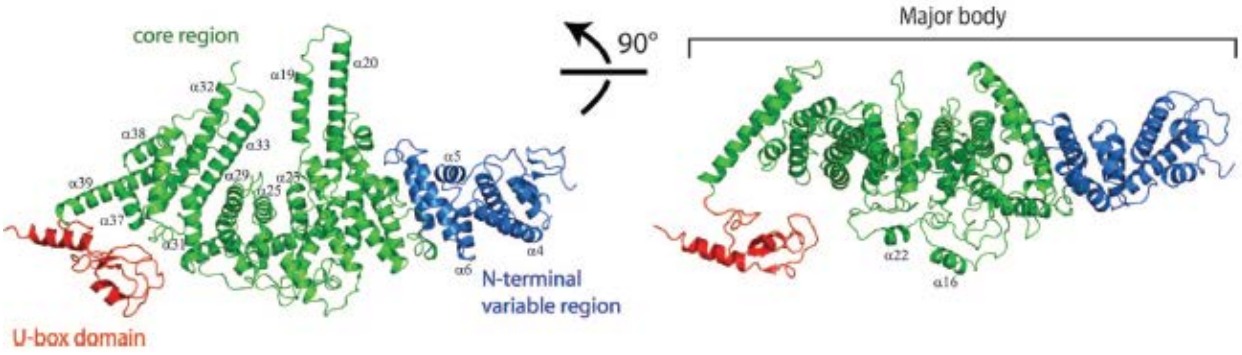

Figure 4-2 Structure of the E4B homolog Ufd2.
The full length 1000+ amino acid protein is shown with the U-box domain (red), armadillo repeats (green), and N-terminal variable region (blue).

Extremely little is known about the physiological functions of the mammalian E4B U-box ligase; only a small handful of substrates have been reported, including the widely-renowned and exhaustively-studied tumor suppressor protein p53[17], FEZ1[18], a protein involved in neuritogenesis, and ataxin-3[19], a protein involved in what is known as the Machado-Joseph disease. In specific cases depending on the substrate, the E4B U-box can act either solely as E3 (FEZ1), or E4 (ataxin-3), or as both (p53); in particular, upregulation of E4B activity has been correlated with decrease in p53 levels in certain medulloblastoma tumors[20]; *in vitro* ubiquitination of p53 with E4B acting as the sole E3 has been reported[20]; however, it is held, canonically, that in *vivo*, E4B acts as an E4 enzyme which polyubiquitinates p53 and any other potential substrates through the mediation of other E3s, namely CHIP[21], and Hdm2[22].

With exceedingly few reported ubiquitination substrates of E4B, and the fact that virtually nothing is known about the complications behind the mechanism which determines the

U-box's mode of action either as E3 or E4 or perhaps, both simultaneously, any findings we obtain from using our OUT system would likely be significant.

4.1.2.2 CHIP structure and known biological functions

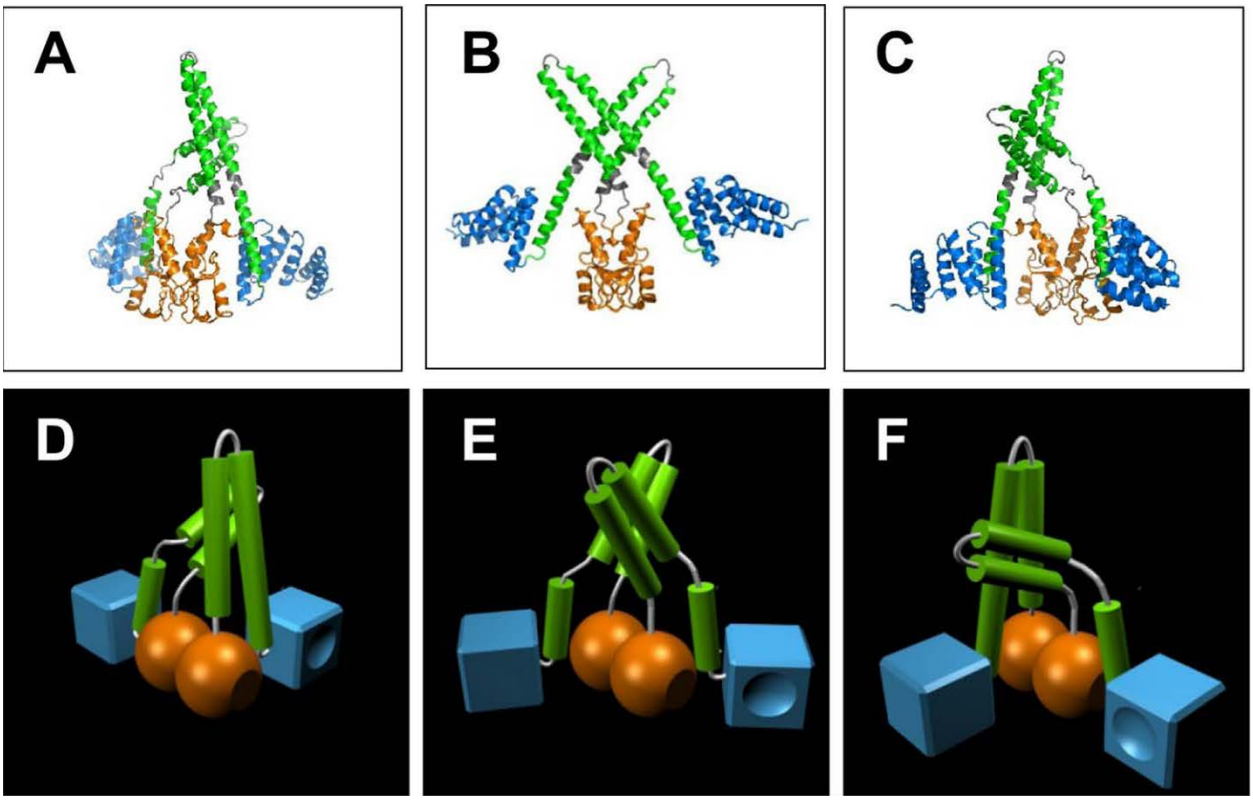

Figure 4-3 Computer-generated model structure of the CHIP homodimer. As previously reported by Qian et.al[8], the CHIP homodimer is believed to possess great flexibility crucial for activity; shown are the U-box domains (orange balls) showing the E2 binding site (flat surface), the TPR domains (blue cubes) showing substrate binding sites (blue holes), and the linker domains (green bars) showing flexible regions (silver). The inactive symmetric conformation is as shown (B&E) in which the active faces of the U-box domains are blocked, alongside the active, asymmetrical ones (A&D, C&F), in which one of the U-box domains becomes openly exposed.

In stark contrast to rather esoteric E4B enzyme, the CHIP (carboxyl terminus of Hsc70-interacting protein) E3 ligase, also known as STUB1, is one of the well-studied members of the U-box family, and of E3s in general. The ~300-amino acid full-length protein, whose rough structure was given as a complex with the E2 enzyme Ubc13[23] in chapter 2, is comprised of the

N-terminal TPR repeats, and linker domain, and the C-terminal U-box[24]. Crystal structures of the CHIP protein itself, or of the protein in complex with E2 or other proteins such as the molecular chaperone Hsc70, are readily available. It is well established that CHIP is active as a homodimer; the U-box domain and the linker domain are responsible for its dimerization. The dimer formed is believed to be asymmetric in its active conformation[25]; only one monomer is accessible by E2 and is active as E3 at any given time; the other monomer also likely serves a purpose and is crucial for stimulating ubiquitin release, as is also the case with dimeric RING enzyme complexes[26]. The complex is believed to be conformationally flexible and dynamic; the individual monomers "take turns" being the active one, this flexibility is conferred by certain regions in the linker domain[27]. (**Figure 4-3**)

CHIP possesses several scores of reported substrates and the known biological pathways regulated by CHIP are numerous[28-44]—CHIP has been known first, and foremost, as a cochaperone protein most well-known for its association with the chaperone Hsc70 (heat shock cognate 70kD) protein[45]—from which it gets its name—through its TPR domain, and interestingly Hsc70 itself as well as other chaperone proteins can also be targeted by CHIP for ubiquitination[46]; in other words, CHIP plays a role in protein quality control; it is believed to ubiquitinate misfolded or damaged proteins during various types of cell stress to consign them for proteasomal degradation[45,47,48]; although CHIP itself can target certain substrates in a chaperone-independent manner[43]. As would be expected, various oncogenes are among CHIP's numerous putative substrates, including but not limited to, c-myc[49], hypoxia-inducible factor 1[50], ErbB2[51], NF-kB[29] and EGFR[52]; somewhat paradoxically, the list of known CHIP substrates also include some tumor suppressor proteins as well, including p53[53-56] and PTEN[57], which may imply that CHIP itself may act as an oncogene in some instances. In whatever case, CHIP is

involved in the suppression of many different types of cancer[58-61], and in several neurological diseases[62-66], including Alzheimer's[67], making it a very interesting target of research.

4.2 Results

4.2.1 Full-length E4B activity assays

4.2.1.1 Expression and reconstitution of wild-type activity

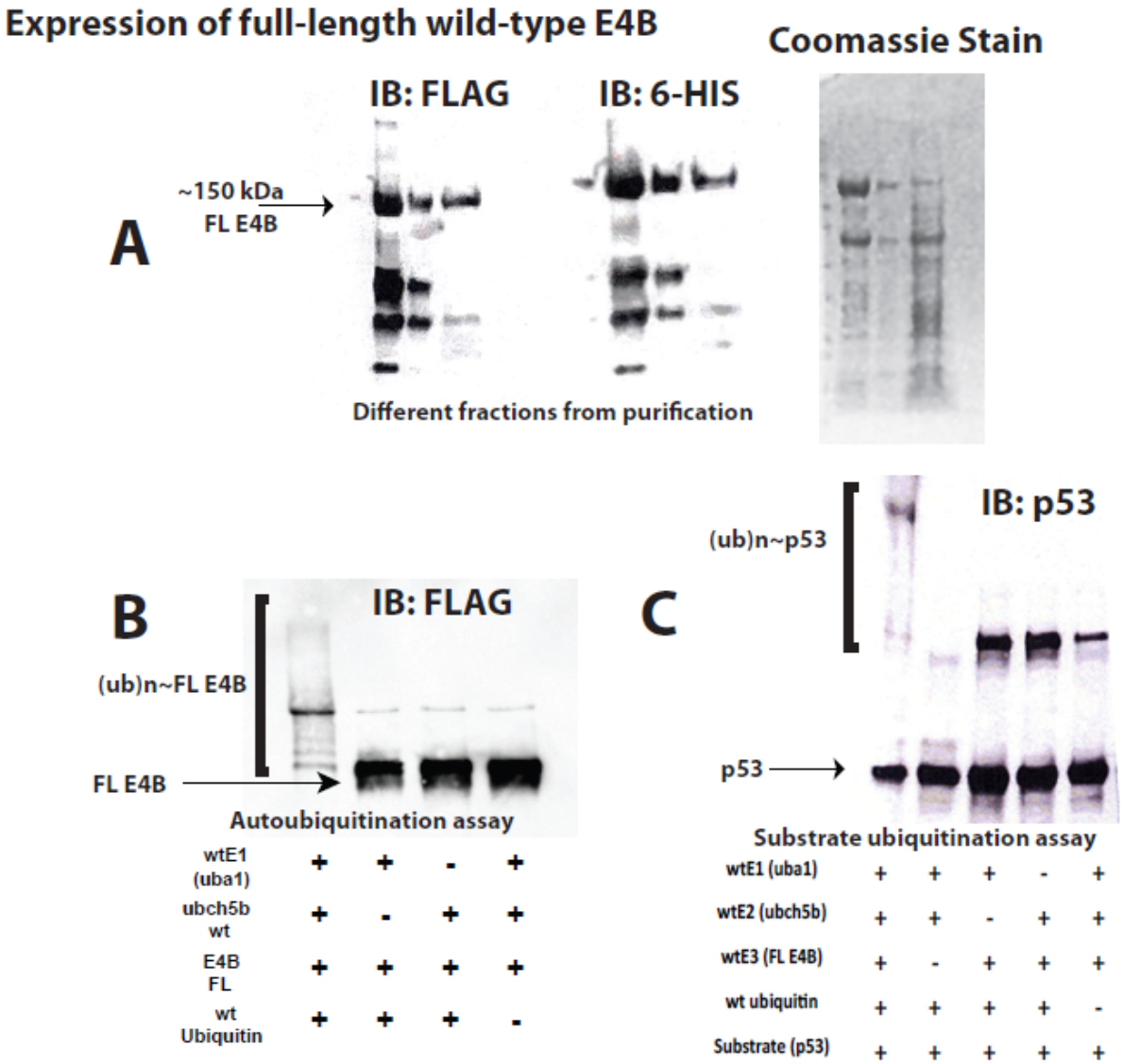

Figure 4-4 Full length wild-type E4B expression and activity. The successful expression of the full length wild-type E4B protein was confirmed by Western immunoblotting, and by Coomassie staining (A); activity of the expressed protein was assayed in both autoubiquitination (B) and putative substrate ubiquitination assays (C).

Our previous selection and activity assay solely involved only the U-box domain of E4B; to test substrate ubiquitination activity, the full-length enzyme, or at least a construct containing the armadillo repeats as well as the U-box would be required. Success of the bacterial expression of the full-length 1100+ amino acid protein, using slightly modified protocols, was confirmed through Coomassie staining, and through western immunoblotting against the N-terminal fused FLAG and 6-His tags (**Figure 4-4A**). Afterwards, we confirmed the activity of wild-type full length E4B (from hence referred to as fE4B) in autoubiquitination assays similar to previous assays on the U-box domain, and with the putative *in vitro* substrate p53 (**Figure 4-4B&C**).

4.2.1.2 Testing of xfE4B OUT activity

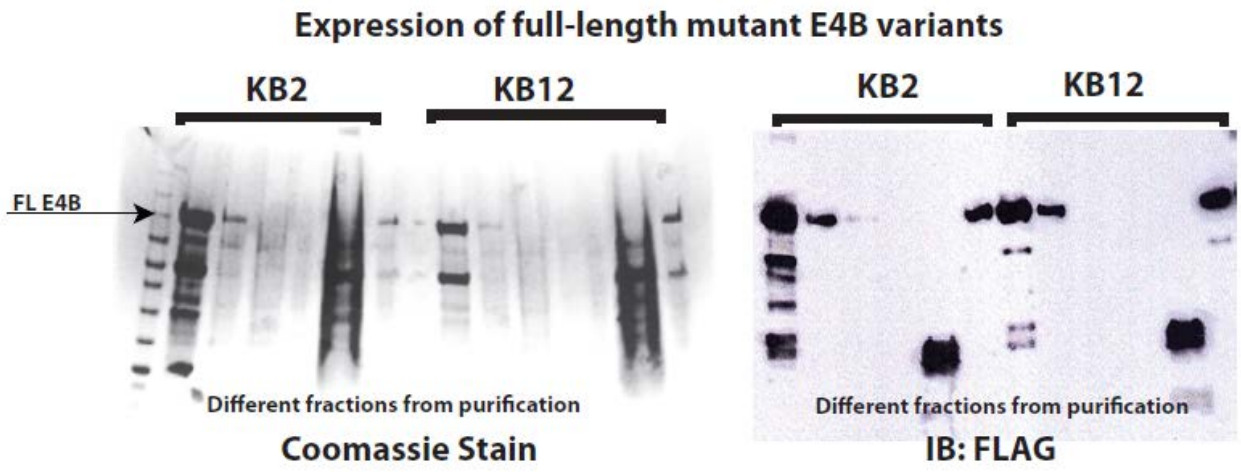

Figure 4-5 Expression of full length E4B mutant variants. The expression of these mutants were confirmed in a similar manner to the wild-type full length E4B enzyme.

Encouraged by our success which demonstrated that wild-type E4B activity was able to be reconstituted using our protocols, we then constructed full-length versions of two of the mutants selected from our library; although all tested selected clones were found to be active, we decided to focus on only these: KB2 (the most abundant clone in the 2nd selection), and KB12, which contained a crucial residue charge reversal in its D1238R mutation (see **Table 3-4**). The

expression of both variants (hence referred to as fKB2 and fKB12) were confirmed in the same manner as the wild-type, by Coomassie staining and Western immunoblotting against FLAG (**Figure 4-5**).

Afterwards, the mutants were tested for OUT ubiquitination involving xUB, xE1 and xE2 in both autoubiquitination, and p53 ubiquitination assays (**Figure 4-6A&B**), and were found, much to our delight, to be active; while as expected, the wild-type protein showed no observable activity in p53 ubiquitination. The two mutants showed virtually the same levels of activity, as was also observed when only the U-box domain constructs were tested (Chapter 3, Scheme 3-6B). This marked the first occasion where we have observed xub transferred successfully through the entire OUT cascade onto a substrate protein and to a certain extent, vindicated our approach.

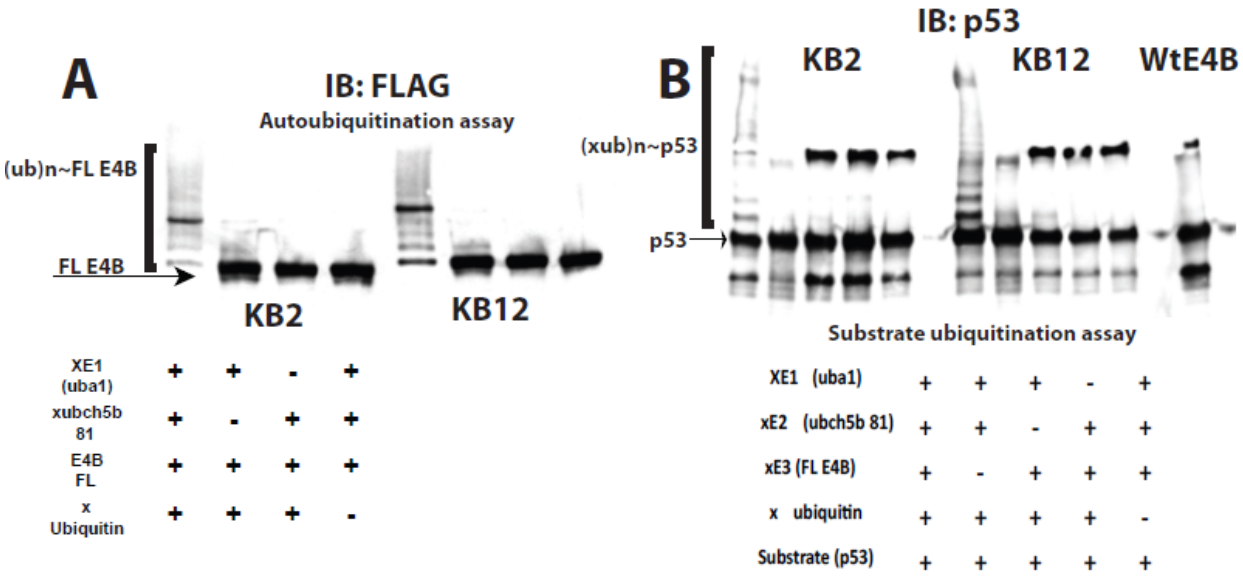

Figure 4-6 Activity assays of full length E4B mutant variants
Shown are the autoubiquitination (A) and substrate ubiquitination (B) assays of the full-length KB2 and KB12 mutants; autoubiquitination was blotted against the FLAG epitope on the E4B protein, while substrate ubiquitination was blotted with p53-specific antibodies; the wild-type E4B protein showed no appreciable activity in p53 ubiquitination when charged with OUT enzymes (rightmost lane).

4.2.2 Construction and activity assays of xCHIP mutants

4.2.2.1 Engineering the CHIP U-box for OUT activity

Encouraged by our success, we then sought to investigate whether or not our selected U-box mutations would also be useful on other E3 frames; as stated earlier, the overall structure of RING and U-box domains are remarkably similar, although their amino acid sequences are not so, which might make the task of determining which amino acid residue is analogous to which when comparing between starkly different E3s problematic.

We have previously discussed our efforts on the U-box ligase CHIP, and the rational design engineering thereof—which had some modest success. As mentioned earlier, there are certain fundamental differences between E4B and CHIP—differences which also exist in the U-box domain itself, which will be discussed more thoroughly in the discussion section; this notwithstanding, we have had great success previously in reconstituting CHIP E3 activity on its putative substrates p53[68], and Hsc70[69]; this, coupled with the fact the CHIP is a protein which possesses great biological significance, led to us deciding to focus our efforts in constructing active xCHIP mutants using the information we obtained from E4B selection.

	230	231	232	233	234	235	236	237	238	239
Wt CHIP	D	Y	L	C	G	K	I	S	F	E
CHIP 11	K	Y	L	C	G	K	I	S	F	E
CHIP 12	D	Y	L	C	G	K	I	S	F	K
CHIP 21	K	Y	L	C	G	K	I	S	F	K
xCHIP2	D	Y	L	K	D	P	I	M	H	T
xCHIP12	D	Y	L	K	D	P	I	M	R	T

Figure 4-7 Library selection-based CHIP mutants. The amino acid sequences of xCHIP2 and xCHIP12 mutants, with the transplanted loop (shaded green) as compared to those of wild-type (blue text) and rationally-designed mutants (Table 2-4) with their mutations (shaded yellow).

Using information from the existing crystal structures[23], we decided to target the whole stretch of amino acid residues 233-239 which comprise the loop region in CHIP interacting with the N-terminal helix of E2; this entire loop was replaced by the corresponding amino acid residues 1233 to 1239 of E4B mutants KB2 and KB12 to create mutants termed xCHIP2 and xCHIP12, respectively, as shown above (**Figure 4-7**).

4.2.2.2 xCHIP OUT activity assays

After confirming the success of expression, these xCHIP variants were then tested in autoubiquitination and p53 ubiquitination assays involving OUT components (**Figure 4-8A&B**), and were also found to be active at a level comparable to the wild-type enzyme (**Figure 4-8C**).

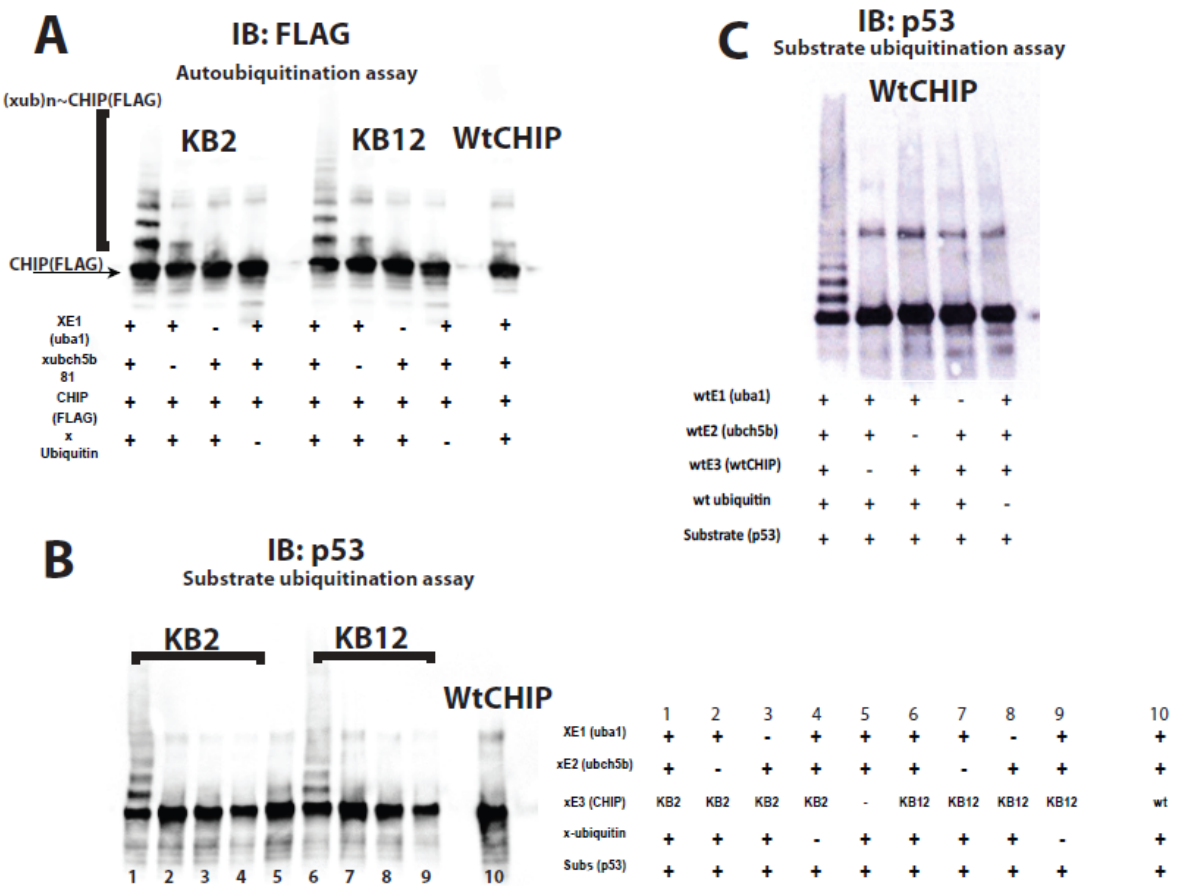

Figure 4-8 Activity assays of CHIP mutant variants. (A) Autoubiquitination of FLAG-tagged CHIP mutant variants, xCHIP2 and xCHIP12, as opposed to FLAG-wild-type CHIP (last lane); (B) in a similar manner, putative substrate (p53) ubiquitination of xCHIP KB2 and KB12 compared to wild-type (last lane) as shown and (C) Wild-type CHIP ubiquitination of p53.

In the same manner as E4B, the wild-type CHIP enzyme did not show any appreciable activity with OUT enzymes in the autoubiquitination and the substrate ubiquitination assays; this was consistent with our previous observations mentioned in Chapter 2.

4.3 Discussion

4.3.1 Full-length E4B expression

To our knowledge, at the time of writing there existed only one reported instance in which the full length E4B protein had been successfully isolated to be used in *in vitro* assays—involving expression from insect cell systems; the vast majority of studies on the full-length protein assayed the protein activity *in vivo* in mammalian cell systems. Expression of the full-length E4B protein via the generic bacterial expression protocols our group has been accustomed to, had proven problematic in our experience due to various reasons such as extremely low yield per culture volume, low purity due to the full length E4B protein copurifying with other unwanted proteins, the formation of inclusion bodies and insolubility in generic buffers, and observed protein truncation, due to incomplete transcription or translation, or degradation in the isolation process.

Our success in expression of the protein using the generic pET vector expression system was based off of a protocol previously reported in a similar bacterial expression of the full-length yeast homolog Ufd2 protein[3]; through many attempts in trial-and-error, we developed a significantly modified protocol which involved culturing the *E.coli* cells in rich media until very late in the log-growth phase, induction conditions using very high amounts of IPTG at low temperatures for extended times, and after standard affinity tag purification, a second purification step involving precipitation of the target protein at a specific ammonium sulfate gradient. When

our work is successfully published, we believe this would mark the first reported occasion in which this protein has been expressed in active form using bacterial systems.

4.3.2 Comparisons between E4B and CHIP U-box domains

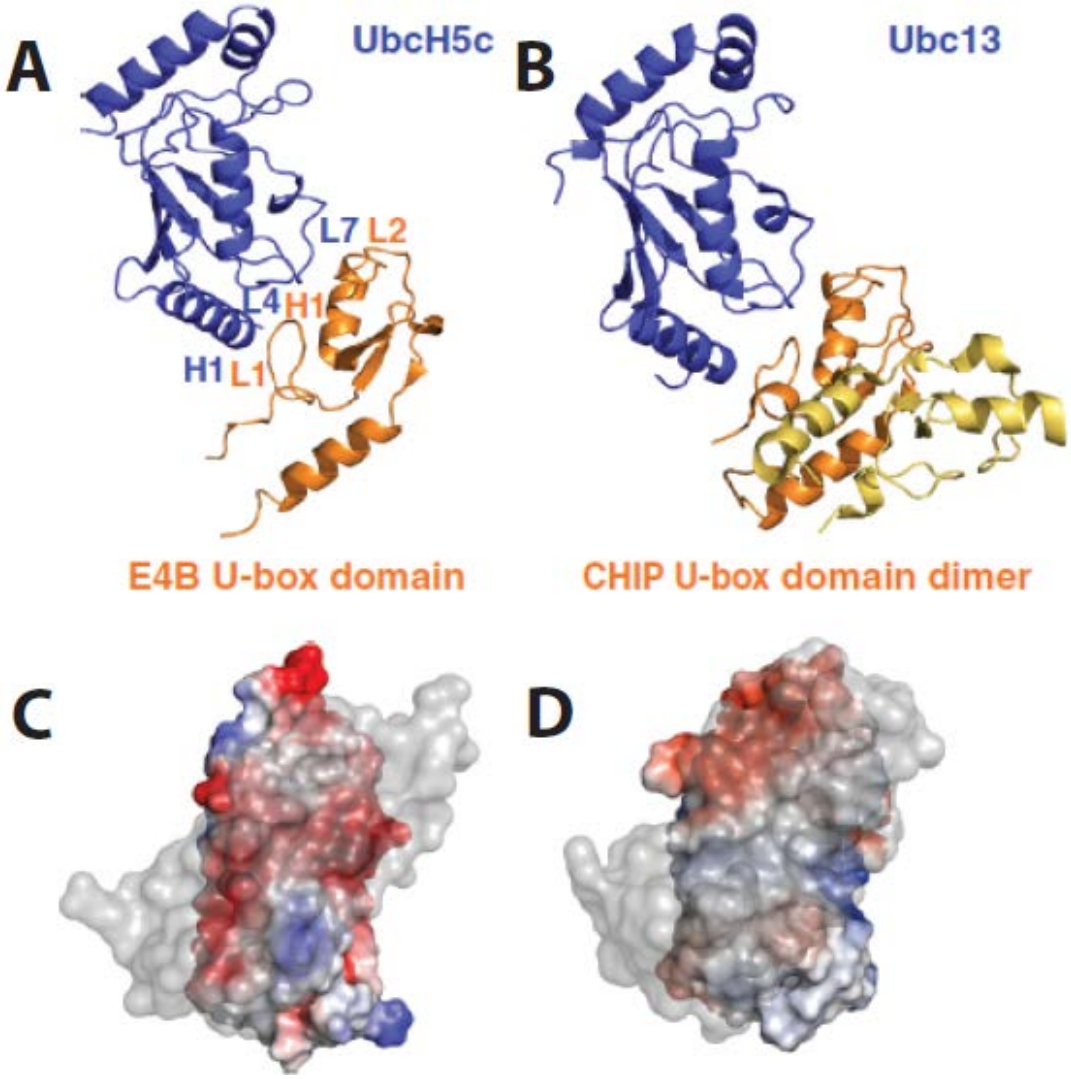

Figure 4-9 Comparison between the E4B and CHIP U-box domains[16]**.** The comparative structures of the E2-U-box complexes for (A) E4B monomer, and (B) asymmetric CHIP dimer; active U-box domains (orange), the E2 enzymes (purple), and the blocked, inactive CHIP U-box monomer (yellow), are as shown. The respective electrostatic surface potentials are as given in (C) and (D); the highly polarized E4B surface renders the theoretical dimer structure shown in (C) highly unfavorable due to negative electrostatic repulsion, whereas this effect is nowhere near as pronounced in CHIP (D).

Superficially, the U-box domains generally bear the same structural folds and motifs and can be easily superimposed to demonstrate the resemblance; however, there are certain fundamental differences between monomeric U-boxes such as E4B and dimeric ones such as CHIP which is readily apparent in the amino acid composition. Certain hydrophobic residues in the U-boxes of the dimeric CHIP, and the tetrameric Prp19, believed to responsible for dimerization via hydrophobic interactions, correspond to charged residues in E4B and its fellow monomeric homologs upon alignment; any theoretical E4B dimer in a similar conformation would hence be rendered unfavorable by electrostatic repulsions—this fact is readily apparent when one considers the drastically different surface potential characteristics of the two enzymes, as published by Benirschke et.al[16] shown above; the surface of CHIP contains large swaths of uncharged regions around the area where the monomers overlap and dimerization takes place, whereas the equivalent surface of the theoretical E4B dimer is markedly much more strongly polarized (**Figure 4-9**).

We have previously discussed using rational design mutagenesis which involved the charge reversal of certain acidic residues in RING domains and in the CHIP U-box in Chapter 2. The engineered RING domains exhibited nonexistent OUT activity, while one of the engineered CHIP mutants, in particular the D230K mutant—from the crystal structure this particular aspartate seemed to be in the closest proximity and have the highest likelihood of interacting with the N-terminal helix of E2—had been more promising; however, when we considered the amino acid sequence alignment of U-box domains based on secondary structure motifs (**Figures 4-10, 4-11**), between E4B and CHIP, it was clear that all of the loop residues which were randomized in our E4B selection with the exception for L1236 (which aligned to CHIP's similar I236), aligned to rather unconvincing counterparts in CHIP.

CHIP (human)	L	C	G	K	I	S	F	E	L	M	R	E	P
E4B (human)	F	R	D	P	L	M	D	T	L	M	T	D	P
Ufd2 (yeast)	F	L	D	P	L	M	Y	T	I	M	K	D	P
Prp19 (yeast)	M	L	C	A	I	S	G	K	V	P	R	R	P
NleG	I	Q	C	P	I	T	L	E	Q	P	E	K	G
AtPUB14	F	R	C	P	I	S	L	E	L	M	K	D	P

Figure 4-10 Alignment of U-box domains. Shown are the sequences in the E2 N-terminal helix-interacting loop region of different U-box domains. Conserved or semi-conserved residues are in red text. The loop region which was transplanted into the CHIP mutants: xCHIP2 and xCHIP12, is shaded in yellow.

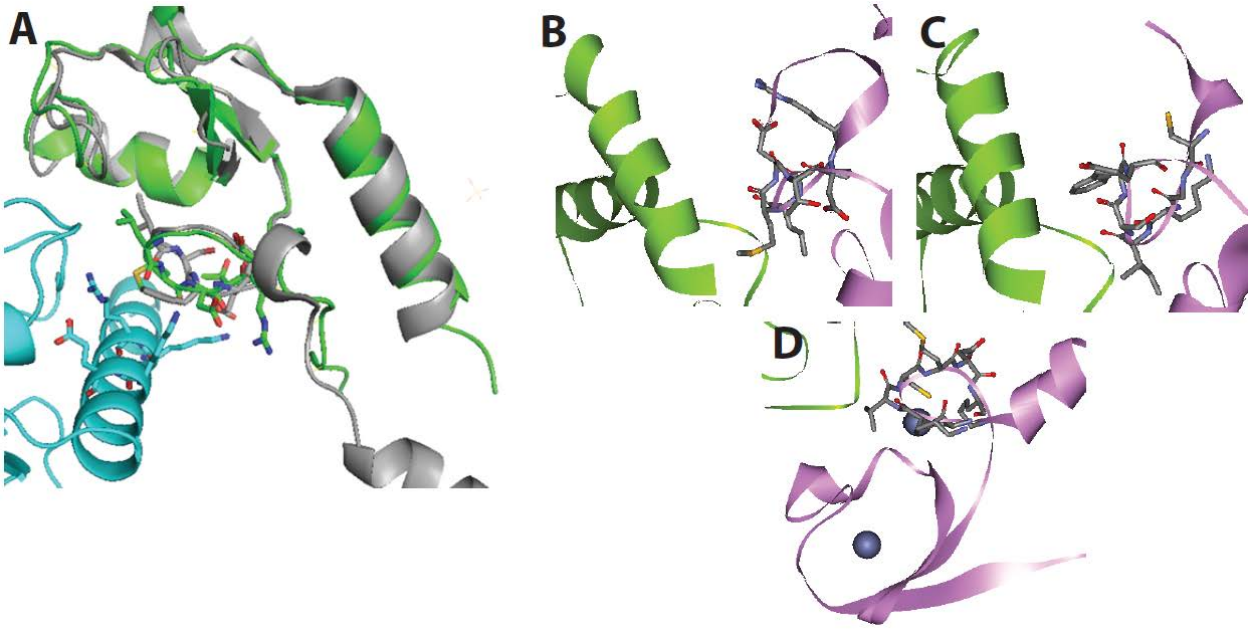

Figure 4-11 Comparison of E2-interacting loop domains. (A) Despite the apparent similarity between the E4B (green) and CHIP U-box domains (grey) when their respective structures taken in complex with the E2 enzyme (cyan) are interposed, (B-D) the amino acid compositions of the crucial loop in the E3 enzymes (lilac) interacting with the E2 enzyme (olive) for E4B (B) and CHIP (C) are considerably different in property and character. In comparison, the analogous loop domain in the TRAF6 RING (D) is also given, shown in complex with the crucial zinc ions (silver balls).

R1233 would align directly onto a cysteine residue; M1236 onto serine, the crucial salt bridge D1238 aligned onto phenylalanine, and T1239 onto glutamic acid. The unmutated D1234 and P1235 residues in E4B also do not have good matchings in, respectively, the glycine and lysine residues of CHIP. This suggested that despite the superficial structural similarity, the specific chemistry in the interactions between this loop region and the E2 N-terminal helix differs significantly between E4B and CHIP; most likely the binding modes and conformations, while probably similar to some extent as far as overall complex tertiary structure is concerned, are ultimately fundamentally different in the nature of the specific underlying chemical interactions. Hence, constructing the CHIP U-box to be perfectly analogous to the selected U-box mutants of E4B was far from being straightforward; however, to achieve our ultimate end of substrate ubiquitination, we had no need of a perfect analog; we merely needed an active one. And so, we came to the decision of replacing the entire loop region in the CHIP U-box with those from xE4B, which also included the unrandomized D1234 and P1235 residues; and as far as results were concerned, we seemed to be well-rewarded for our efforts, as these mutants readily ubiquitinated the putative p53 substrate with OUT components.

Admittedly, this approach had probably resulted in CHIP mutants which would most likely be recognized by, and interact with, the E2 enzyme as more of an E4B U-box than a CHIP U-box; however, since their TPR repeat domains remained unchanged, these mutants should still recognize their wild-type substrates; we also did not touch the hydrophobic residues responsible for CHIP dimerization in the linker domains, and hence we had good reason to expect that these mutants would also form stable, asymmetric dimers similar to those formed by the wild-type CHIP enzyme; as far as the whole E3 enzyme itself was concerned, we had no reason to suspect that these mutants would possess any significant structural differences with the wild-type CHIP;

however, further testing, including the acquisition of a crystal structure, is probably required to positively verify these claims.

4.3.3 Possible applications to other E3 enzymes

Given the obvious, though admittedly, superficial, structural similarities between the RING and the U-box families, it was definitely worth considering whether or not similar E2-interacting loop transplants would work on RINGs; it did not seem to matter that the amino acid sequences of E4B and CHIP were starkly different, and in many ways the sequence of CHIP's U-box bears more resemblance to the sequence of RINGs than that of some U-box domains.

However, in RING domains, which are comprised of zinc fingers, the analogous loop region contains a crucial and perfectly conserved zinc-coordinating cysteine residue (refer to region labelled as site I in the RING domains in **Figure 2-2**[70]) which would most likely affect the overall structure and activity of the enzyme if mutated; this residue, C233 is also apparently "conserved" in CHIP upon alignment (if one subscribes to the idea that the U-box family was derived from RINGs, as some do[71]), although that particular cysteine residue in the CHIP enzyme is oriented differently than those found in the RING enzymes (see **Figure 4-11**) and mutation of this residue in our xCHIP mutants did not seem to affect its activity—U-box domains are not known to coordinate zinc ions. A partial loop transplant conserving the cysteine residue while mutating all the other ones had been among one of the possibilities we had considered—afterwards members of our group then tested such mutations on the TRAF6 and Mdm2 RING domains; however the attempts yielded no fruitful results (data no shown).

Hence, for the time being, engineering the RING domains for OUT activity continues to be a challenge. Perhaps one possible way to circumvent this problem is to create chimeric proteins—fusing the entire engineered U-box domain onto the substrate binding regions of RINGs and testing for substrate ubiquitination activity. Similar techniques involving the creation of chimeric proteins to change substrate specificity have been reported on multiple occasions, including fusion even between conjugating enzymes in different pathways, such as ubiquitin and Nedd8 to transform what was formerly a ubiquitin E3 enzyme, into a neddylating one[72].

Conceivably, the monomeric E4B U-box may be fused to the frames of the handful of known monomeric RINGs such as the Cbl ligases[26], whereas the CHIP U-box, whose U-box amino acid composition is more amenable to dimerization, may be fused to the frames of the vast majority of RINGs which operate as dimers such as Mdm2[73] or BIRC7[74]; these chimaeras would theoretically interact with their respective native substrates normally with their intact would certainly be required to determine whether or not this would be possible, but if so, it would allow us to access ~90% of all known E3 ligases and open an innumerable number of avenues for future research.

Regardless, for the time being, our results confirmed that we had successfully reengineered the U-box domains of E4B and CHIP to allow these enzymes to ubiquitinate putative substrates with xUB through the newly completed cascade of XE1, xE2 and x-Ubox. The last and final step which remained was to reconstitute this entire orthogonal cascade *in vivo* and ultimately enrich and identify potential E3 substrates labelled with xUB. This would finally allow the OUT system, which our group had painstakingly endeavored over the years to construct, to fulfill the main purpose it was created for; the results of this will be discussed in the next and final chapter.

4.4 Methods

The list of primers used for this chapter are as given in **Table 4-1**.

Primer name	Sequence (5'-3')
Kar232	GACATCCCCGACTACCTGAAGGACCCTATTATGCATACGC TGATGCGGGAGCCGTGCATC
Kar233	GACATCCCCGACTACCTGAAGGACCCTATTATGAGGACGC TGATGCGGGAGCCGTGCATC
Kar234	CAGGTAGTCGGGGATGTCTCGCTTCTTCCTCTTCTCATC
CHIP NdeI up	AGAAGGAGATATACATATGAAGGGCAAGG
CHIP XhoI stop	TGGTGGTGCTCGAGGTCTAGTCAGTAGTCCTCCACCCAG CCATTCTCAGAGATGAATGCG
CHIP KpnI XhoI	TGGTGGTGCTCGAGGTCGGTACCGTAGTCCTCCACCCAG CC
Kar238	GAACTCGTCCGGGGC
Kar239	GCCCCGGACGAGTTCAAGGACCCTATTATGCATACGCTGA TGACCGATCCCG
Kar240	GCCCCGGACGAGTTCAAGGACCCTATTATGAGGACGCTG ATGACCGATCCCG
Kar241	TTCGATGATCTCGAGTCAGTGGTCACTGCTCTGTTTCTCTC
Kar242	GTCTCTGAAGCGGATCCATGAAGTGCAAGAAGAGATG

Table 4-1. Primers used in Chapter 4.
Residues in red indicate restriction sites.

4.4.1 Cloning and expression of full length E4B variants

pET30a-FLAG-full length wild-type E4B was a gift from Vanderbilt University; the vector had since been slightly modified by our group to correct a preexisting frameshift error. The mutant sequences of E4B KB2 and KB12 were introduced via overlap extension PCR using the wild-type full length as DNA template; Kar239 (KB2) or Kar240 (KB12) was paired with Kar241, and these were coupled with the PCR product from the Kar238/Kar242 pair; the final

product was digested with BamHI/XhoI, and ligated into the aforementioned pET30a wild-type E4B vector, utilizing an internal BamHI site in the E4B gene which lay several hundreds of base pairs upstream of the C-terminal mutated U-box region; hence, these mutants retained the N-terminal FLAG tag that had been present in the wild-type construct. The full-length E4B constructs (wild-type, KB2, and KB12) were transformed into ArcticExpress(DE3) cells (Strategene).

An overnight culture was inoculated into Terrific Broth medium supplemented with 70 mg/ml kanamycin and grown with vigorous shaking at 37˚C until the optical density at 600 nm reached ~1.0-1.5. The culture was cooled down to 13˚C for 10 minutes, before being induced by the addition of IPTG to a final concentration of 4 mM; induction was allowed to continue at 13˚C for 18 hours overnight.

The next morning, the cells were harvested by centrifugation at 5000 rpm for 10 minutes, and resuspended in lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl and 10 mM imidazole); the cells were then treated with ~5 mg/ml lysozyme and incubated at 30˚C for 1 hour; lysis of the cells was completed via sonication, and the lysate was cleared by centrifugation at 6000 rpm for 30 minutes. Ni-NTA agarose (QIAGEN) was incubated with the cleared lysate at 4˚C for 2 hours, and the resin was washed twice with 20 volumes of wash buffer (same as lysis buffer, but with 20 mM imidazole), and the bound protein was eluted twice, each time with two volumes of elution buffer (same as lysis buffer, but with 250 mM imidazole). The purified eluate was dialyzed overnight at 4˚C in a precipitation buffer containing 50 mM Tris pH 7.5, 1 M ammonium sulfate, 1 M KCl, 5 mM DTT and 10% glycerol. The pellet containing precipitated E4B protein was first reconstituted in TBS supplemented with 10 mM $MgCl_2$, and any remaining pellet afterwards was reconstituted in a solution containing 50 mM Tris pH 8.0; the

remaining supernatant from overnight dialysis was concentrated down further using Amicon centrifugal filters (EMD Millipore) with a MW cut off point of 100 kDa—the resulting pellet was also likewise reconstituted. The presence of full-length flag E4B in each of the different fractions was confirmed by Coomassie staining and Western immunoblotting against the FLAG affinity tag. Fractions found containing significant amounts of the target E4B protein (~130-150 kDa) were then pooled together and used for activity assays.

4.4.2 Cloning and expression of xCHIP

The loop-transplant mutations in the xCHIP2 and xCHIP12 clones were introduced via overlap extension PCR using the wild-type CHIP gene in pET vector (a construct made by a previous member of our group) as template, via the overlap extension primer Kar234 paired with "CHIP NdeI up", which was coupled with the products from either Kar232 (xCHIP2) or Kar233 (xCHIP12) with either "CHIP XhoI KpnI" for variants with the C-terminal FLAG epitope or "CHIP XhoI stop" for those without. The PCR product was digested with either NdeI/KpnI (for variants with FLAG) or NdeI/XhoI (for variants without) restriction enzyme pairs and ligated into similarly digested and dephosphorylated pET-mutant E4B U-box FLAG vector (from 3.4.4.2); the vector contained a FLAG sequence within the KpnI restriction site which would be removed if digested with XhoI. After confirmation by DNA sequencing, the constructs were then transformed into ArcticExpress(DE3) cells, and expression and purification of these proteins followed the generic protocol previously reported in 2.4.2. The success of protein expression was confirmed by Coomassie staining and by Western immunoblotting with rabbit antibody raised against the CHIP protein (Santa Cruz), or by antibodies against the FLAG epitope (for variants with FLAG).

4.4.3 Ubiquitination assays

All assays were done in TBS supplemented with 10 mM $MgCl_2$ and 1.5 mM ATP.

4.4.3.1 Autoubiquitination assays

5 μM of xCHIP2, xCHIP12 or wild-type CHIP with C-terminal FLAG epitope tags, or 5 μM of fKB2, fKB12, and wild-type fE4B which bore N-terminal FLAG affinity tags, were charged with 2 μM XE1 (uba1), 10 μM xE2 (ubch5b-81), and 60 μM xub in 30 uL reaction volume and incubated at 37˚C for 1.5 hours. The presence of ubiquitinated E3 species was detected using mouse anti-FLAG antibody (Fisher).

4.4.3.2 Substrate (p53) ubiquitination assays

For reconstitution of wild-type p53, in 30 μL reaction, 4 μM of wild-type E3, fE4B (N-terminal FLAG) or CHIP (no FLAG) was mixed with 0.4 μM p53 (Boston Biochem) and incubated at 4˚C for 30 minutes. Afterwards wild-type uba1 (0.2 μM), ubch5b (0.2 μM) and ubiquitin (30 μM) were added to the reaction, mixed, and allowed to incubate at 30˚C for 1 hour.

For OUT ubiquitination of p53, 5 μM of xCHIP2, xCHIP12 and wild-type CHIP (no FLAG), or 3 μM of fKB2, fKB12 and fE4B was mixed with 0.1 μM of p53 (Boston Biochem) and incubated at 4˚C for 30 minutes. Afterwards the reaction was charged with 2μM of XE1 (uba1), 20 μM xE2 (ubch5b-81) and 100 μM xub, and allowed to incubate at 30˚C for 1.5 hours.

All p53 ubiquitination reactions were immunoblotted and probed with mouse anti-p53 antibody (DO-1) from Santa Cruz Biotechnologies.

4.5 References

1 Paul, I. & Ghosh, M. K. The E3 ligase CHIP: insights into its structure and regulation. *Biomed Res Int* **2014**, 918183, doi:10.1155/2014/918183 (2014).

2 Vander Kooi, C. W. *et al.* The Prp19 WD40 domain contains a conserved protein interaction region essential for its function. *Structure* **18**, 584-593, doi:10.1016/j.str. 2010.02.015 (2010).

3 Tu, D., Li, W., Ye, Y. & Brunger, A. T. Structure and function of the yeast U-box-containing ubiquitin ligase Ufd2p. *Proc Natl Acad Sci U S A* **104**, 15599-15606, doi:10.1073/pnas.0701369104 (2007).

4 Hatakeyama, S., Yada, M., Matsumoto, M., Ishida, N. & Nakayama, K. I. U box proteins as a new family of ubiquitin-protein ligases. *J Biol Chem* **276**, 33111-33120, doi:10.1074/jbc.M102755200 (2001).

5 Patterson, C. A new gun in town: the U box is a ubiquitin ligase domain. *Science's STKE : signal transduction knowledge environment* **2002**, pe4, doi:10.1126/stke.2002.116.pe4 (2002).

6 Marin, I. Ancient origin of animal U-box ubiquitin ligases. *BMC evolutionary biology* **10**, 331, doi:10.1186/1471-2148-10-331 (2010).

7 Cyr, D. M., Hohfeld, J. & Patterson, C. Protein quality control: U-box-containing E3 ubiquitin ligases join the fold. *Trends in biochemical sciences* **27**, 368-375 (2002).

8 Qian, S. B. *et al.* Engineering a ubiquitin ligase reveals conformational flexibility required for ubiquitin transfer. *J Biol Chem* **284**, 26797-26802, doi:10.1074/jbc. M109.032334 (2009).

9 Vander Kooi, C. W. *et al.* The Prp19 U-box crystal structure suggests a common dimeric architecture for a class of oligomeric E3 ubiquitin ligases. *Biochemistry* **45**, 121-130, doi:10.1021/bi051787e (2006).

10 Nordquist, K. A. *et al.* Structural and functional characterization of the monomeric U-box domain from E4B. *Biochemistry* **49**, 347-355, doi:10.1021/bi901620v (2010).

11 Ferreira, R. T., Menezes, R. A. & Rodrigues-Pousada, C. E4-Ubiquitin ligase Ufd2 stabilizes Yap8 and modulates arsenic stress responses independent of the U-box motif. *Biology open* **4**, 1122-1131, doi:10.1242/bio.010405 (2015).

12 Hanzelmann, P., Stingele, J., Hofmann, K., Schindelin, H. & Raasi, S. The yeast E4 ubiquitin ligase Ufd2 interacts with the ubiquitin-like domains of Rad23 and Dsk2 via a novel and distinct ubiquitin-like binding domain. *J Biol Chem* **285**, 20390-20398, doi:10.1074/jbc.M110.112532 (2010).

13 Liu, C. *et al.* Ubiquitin chain elongation enzyme Ufd2 regulates a subset of Doa10 substrates. *J Biol Chem* **285**, 10265-10272, doi:10.1074/jbc.M110.110551 (2010).

14 Imai, Y. *et al.* CHIP is associated with Parkin, a gene responsible for familial Parkinson's disease, and enhances its ubiquitin ligase activity. *Mol Cell* **10**, 55-67 (2002).

15 Koegl, M. *et al.* A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. *Cell* **96**, 635-644 (1999).

16 Benirschke, R. C. *et al.* Molecular basis for the association of human E4B U box ubiquitin ligase with E2-conjugating enzymes UbcH5c and Ubc4. *Structure* **18**, 955-965, doi:10.1016/j.str.2010.04.017 (2010).

17 Wu, H. & Leng, R. P. UBE4B, a ubiquitin chain assembly factor, is required for MDM2-mediated p53 polyubiquitination and degradation. *Cell Cycle* **10**, 1912-1915, doi:10.4161/cc.10.12.15882 (2014).

18 Okumura, F., Hatakeyama, S., Matsumoto, M., Kamura, T. & Nakayama, K. I. Functional regulation of FEZ1 by the U-box-type ubiquitin ligase E4B contributes to neuritogenesis. *J Biol Chem* **279**, 53533-53543, doi:10.1074/jbc.M402916200 (2004).

19 Matsumoto, M. *et al.* Molecular clearance of ataxin-3 is regulated by a mammalian E4. *The EMBO journal* **23**, 659-669, doi:10.1038/sj.emboj.7600081 (2004).

20 Wu, H. *et al.* UBE4B promotes Hdm2-mediated degradation of the tumor suppressor p53. *Nature medicine* **17**, 347-355, doi:10.1038/nm.2283 (2011).

21 Hoppe, T. *et al.* Regulation of the myosin-directed chaperone UNC-45 by a novel E3/E4-multiubiquitylation complex in C. elegans. *Cell* **118**, 337-349, doi:10.1016/j.cell.2004.07.014 (2004).

22 Zhang, Y., Lv, Y., Zhang, Y. & Gao, H. Regulation of p53 level by UBE4B in breast cancer. *PLoS One* **9**, e90154, doi:10.1371/journal.pone.0090154 (2014).

23 Zhang, M. *et al.* Chaperoned ubiquitylation--crystal structures of the CHIP U box E3 ubiquitin ligase and a CHIP-Ubc13-Uev1a complex. *Mol Cell* **20**, 525-538, doi:10.1016/j.molcel.2005.09.023 (2005).

24 Graf, C., Stankiewicz, M., Nikolay, R. & Mayer, M. P. Insights into the conformational dynamics of the E3 ubiquitin ligase CHIP in complex with chaperones and E2 enzymes. *Biochemistry* **49**, 2121-2129, doi:10.1021/bi901829f (2010).

25 Schulman, B. A. & Chen, Z. J. Protein ubiquitination: CHIPping away the symmetry. *Mol Cell* **20**, 653-655, doi:10.1016/j.molcel.2005.11.019 (2005).

26 Dou, H., Buetow, L., Sibbet, G. J., Cameron, K. & Huang, D. T. Essentiality of a non-RING element in priming donor ubiquitin for catalysis by a monomeric E3. *Nat Struct Mol Biol* **20**, 982-986, doi:10.1038/nsmb.2621 (2013).

27 Nikolay, R. *et al.* Dimerization of the human E3 ligase CHIP via a coiled-coil domain is essential for its activity. *J Biol Chem* **279**, 2673-2678, doi:10.1074/jbc.M311112200 (2004).

28 Gil Lorenzo, A. F. *et al.* Heat Shock Protein 70 and CHIP Promote Nox4 Ubiquitination and Degradation within the Losartan Antioxidative Effect in Proximal Tubule Cells. *Cell Physiol Biochem* **36**, 2183-2197, doi:10.1159/000430184 (2015).

29 Jiang, B. *et al.* Carboxyl terminus of HSC70-interacting protein (CHIP) down-regulates NF-kappaB-inducing kinase (NIK) and suppresses NIK-induced liver injury. *J Biol Chem* **290**, 11704-11714, doi:10.1074/jbc.M114.635086 (2015).

30 Min, B. *et al.* CHIP-mediated degradation of transglutaminase 2 negatively regulates tumor growth and angiogenesis in renal cancer. *Oncogene*, doi:10.1038/onc.2015.439 (2015).

31 Petrucelli, L. *et al.* CHIP and Hsp70 regulate tau ubiquitination, degradation and aggregation. *Hum Mol Genet* **13**, 703-714, doi:10.1093/hmg/ddh083 (2004).

32 Teng, Y., Rezvani, K. & De Biasi, M. UBXN2A regulates nicotinic receptor degradation by modulating the E3 ligase activity of CHIP. *Biochem Pharmacol* **97**, 518-530, doi:10.1016/j.bcp.2015.08.084 (2015).

33 Zhang, H. T. *et al.* The E3 ubiquitin ligase CHIP mediates ubiquitination and proteasomal degradation of PRMT5. *Biochim Biophys Acta* **1863**, 335-346, doi:10.1016/j.bbamcr.2015.12.001 (2016).

34 Ferreira, J. V. *et al.* STUB1/CHIP is required for HIF1A degradation by chaperone-mediated autophagy. *Autophagy* **9**, 1349-1366, doi:10.4161/auto.25190 (2013).

35 Gao, B. *et al.* Inhibition of histone deacetylase activity suppresses IFN-gamma induction of tripartite motif 22 via CHIP-mediated proteasomal degradation of IRF-1. *Journal of immunology (Baltimore, Md. : 1950)* **191**, 464-471, doi:10.4049/jimmunol.1203533 (2013).

36 Jang, K. W. *et al.* Ubiquitin ligase CHIP induces TRAF2 proteasomal degradation and NF-kappaB inactivation to regulate breast cancer cell invasion. *Journal of cellular biochemistry* **112**, 3612-3620, doi:10.1002/jcb.23292 (2011).

37 Kao, S. H. *et al.* GSK3beta controls epithelial-mesenchymal transition and tumor metastasis by CHIP-mediated degradation of Slug. *Oncogene* **33**, 3172-3182, doi:10.1038/onc.2013.279 (2014).

38 Ko, H. R. *et al.* P42 Ebp1 regulates the proteasomal degradation of the p85 regulatory subunit of PI3K by recruiting a chaperone-E3 ligase complex HSP70/CHIP. *Cell death & disease* **5**, e1131, doi:10.1038/cddis.2014.79 (2014).

39 Lee, J. H., Khadka, P., Baek, S. H. & Chung, I. K. CHIP promotes human telomerase reverse transcriptase degradation and negatively regulates telomerase activity. *J Biol Chem* **285**, 42033-42045, doi:10.1074/jbc.M110.149831 (2010).

40 Lee, S. *et al.* Heat shock protein cognate 70-4 and an E3 ubiquitin ligase, CHIP, mediate plastid-destined precursor degradation through the ubiquitin-26S proteasome system in Arabidopsis. *The Plant cell* **21**, 3984-4001, doi:10.1105/tpc.109.071548 (2009).

41 Sarkar, S., Brautigan, D. L., Parsons, S. J. & Larner, J. M. Androgen receptor degradation by the E3 ligase CHIP modulates mitotic arrest in prostate cancer cells. *Oncogene* **33**, 26-33, doi:10.1038/onc.2012.561 (2014).

42 Shang, Y. *et al.* CHIP/Stub1 interacts with eIF5A and mediates its degradation. *Cellular signalling* **26**, 1098-1104, doi:10.1016/j.cellsig.2014.01.030 (2014).

43 Singh, A. K. & Pati, U. CHIP stabilizes amyloid precursor protein via proteasomal degradation and p53-mediated trans-repression of beta-secretase. *Aging cell* **14**, 595-604, doi:10.1111/acel.12335 (2015).

44 Yao, X., Li, G., Lu, C., Xu, H. & Yin, Z. Arctigenin promotes degradation of inducible nitric oxide synthase through CHIP-associated proteasome pathway and suppresses its enzyme activity. *International immunopharmacology* **14**, 138-144, doi:10.1016/j.intimp. 2012.06.017 (2012).

45 Rosser, M. F., Washburn, E., Muchowski, P. J., Patterson, C. & Cyr, D. M. Chaperone functions of the E3 ubiquitin ligase CHIP. *J Biol Chem* **282**, 22267-22277, doi:10.1074/jbc.M700513200 (2007).

46 Soss, S. E., Rose, K. L., Hill, S., Jouan, S. & Chazin, W. J. Biochemical and Proteomic Analysis of Ubiquitination of Hsc70 and Hsp70 by the E3 Ligase CHIP. *PLoS One* **10**, e0128240, doi:10.1371/journal.pone.0128240 (2015).

47 Murata, S., Minami, Y., Minami, M., Chiba, T. & Tanaka, K. CHIP is a chaperone-dependent E3 ligase that ubiquitylates unfolded protein. *EMBO reports* **2**, 1133-1138, doi:10.1093/embo-reports/kve246 (2001).

48 McDonough, H. & Patterson, C. CHIP: a link between the chaperone and proteasome systems. *Cell Stress & Chaperones* **8**, 303, doi:10.1379/1466-1268(2003)008 <0303:calbtc>2.0.co;2 (2003).

49 Paul, I., Ahmed, S. F., Bhowmik, A., Deb, S. & Ghosh, M. K. The ubiquitin ligase CHIP regulates c-Myc stability and transcriptional activity. *Oncogene* **32**, 1284-1295, doi:10.1038/onc.2012.144 (2013).

50 Isaacs, J. S. *et al.* Hsp90 regulates a von Hippel Lindau-independent hypoxia-inducible factor-1 alpha-degradative pathway. *J Biol Chem* **277**, 29936-29944, doi:10.1074/jbc.M204733200 (2002).

51 Xu, W. *et al.* Chaperone-dependent E3 ubiquitin ligase CHIP mediates a degradative pathway for c-ErbB2/Neu. *Proc Natl Acad Sci U S A* **99**, 12847-12852, doi:10.1073/pnas.202365899 (2002).

52 Wang, T. *et al.* CHIP is a novel tumor suppressor in pancreatic cancer through targeting EGFR. *Oncotarget* **5**, 1969-1986, doi:10.18632/oncotarget.1890 (2014).

53 Lim, J. H. & Woo, C. H. Laminar flow activation of ERK5 leads to cytoprotective effect via CHIP-mediated p53 ubiquitination in endothelial cells. *Anatomy & cell biology* **44**, 265-273, doi:10.5115/acb.2011.44.4.265 (2011).

54 Muller, P., Hrstka, R., Coomber, D., Lane, D. P. & Vojtesek, B. Chaperone-dependent stabilization and degradation of p53 mutants. *Oncogene* **27**, 3371-3383, doi:10.1038/sj.onc.1211010 (2008).

55 Naito, A. T. *et al.* Promotion of CHIP-mediated p53 degradation protects the heart from ischemic injury. *Circulation research* **106**, 1692-1702, doi:10.1161/circresaha. 109.214346 (2010).

56 Sisoula, C., Trachana, V., Patterson, C. & Gonos, E. S. CHIP-dependent p53 regulation occurs specifically during cellular senescence. *Free radical biology & medicine* **50**, 157-165, doi:10.1016/j.freeradbiomed.2010.10.701 (2011).

57 Ahmed, S. F. *et al.* The chaperone-assisted E3 ligase C terminus of Hsc70-interacting protein (CHIP) targets PTEN for proteasomal degradation. *J Biol Chem* **287**, 15996-16006, doi:10.1074/jbc.M111.321083 (2012).

58 Cao, Z. *et al.* MiR-1178 promotes the proliferation, G1/S transition, migration and invasion of pancreatic cancer cells by targeting CHIP. *PLoS One* **10**, e0116934, doi:10.1371/journal.pone.0116934 (2015).

59 Hiyoshi, H. *et al.* 2-(4-Hydroxy-3-methoxyphenyl)-benzothiazole suppresses tumor progression and metastatic potential of breast cancer cells by inducing ubiquitin ligase CHIP. *Scientific reports* **4**, 7095, doi:10.1038/srep07095 (2014).

60 Tsuchiya, M. *et al.* Ubiquitin ligase CHIP suppresses cancer stem cell properties in a population of breast cancer cells. *Biochemical and biophysical research communications* **452**, 928-932, doi:10.1016/j.bbrc.2014.09.011 (2014).

61 Wang, Y. *et al.* CHIP/Stub1 functions as a tumor suppressor and represses NF-kappaB-mediated signaling in colorectal cancer. *Carcinogenesis* **35**, 983-991, doi:10.1093/carcin/bgt393 (2014).

62 Chen, X. & Burgoyne, R. D. Identification of common genetic modifiers of neurodegenerative diseases from an integrative analysis of diverse genetic screens in model organisms. *BMC genomics* **13**, 71, doi:10.1186/1471-2164-13-71 (2012).

63 Dickey, C. A., Patterson, C., Dickson, D. & Petrucelli, L. Brain CHIP: removing the culprits in neurodegenerative disease. *Trends in molecular medicine* **13**, 32-38, doi:10.1016/j.molmed.2006.11.003 (2007).

64 Rao, S. N., Sharma, J., Maity, R. & Jana, N. R. Co-chaperone CHIP stabilizes aggregate-prone malin, a ubiquitin ligase mutated in Lafora disease. *J Biol Chem* **285**, 1404-1413, doi:10.1074/jbc.M109.006312 (2010).

65 Shi, C. H. *et al.* Ataxia and hypogonadism caused by the loss of ubiquitin ligase activity of the U box protein CHIP. *Hum Mol Genet* **23**, 1013-1024, doi:10.1093/hmg/ddt497 (2014).

66 Yan, W. Q., Wang, J. L. & Tang, B. S. [Functions of carboxyl-terminus of Hsc70 interacting protein and its role in neurodegenerative disease]. *Zhonghua yi xue yi chuan xue za zhi = Chinese journal of medical genetics* **29**, 426-430, doi:10.3760/cma.j.issn. 1003-9406.2012.04.010 (2012).

67 Tsvetkov, P., Adamovich, Y., Elliott, E. & Shaul, Y. E3 ligase STUB1/CHIP regulates NAD(P)H:quinone oxidoreductase 1 (NQO1) accumulation in aged brain, a process impaired in certain Alzheimer disease patients. *J Biol Chem* **286**, 8839-8845, doi:10.1074/jbc.M110.193276 (2011).

68 Esser, C., Scheffner, M. & Hohfeld, J. The chaperone-associated ubiquitin ligase CHIP is able to target p53 for proteasomal degradation. *J Biol Chem* **280**, 27443-27448, doi:10.1074/jbc.M501574200 (2005).

69 Jiang, J. *et al.* CHIP is a U-box-dependent E3 ubiquitin ligase: identification of Hsc70 as a target for ubiquitylation. *J Biol Chem* **276**, 42938-42944, doi:10.1074/jbc.M101968200 (2001).

70 Brzovic, P. S., Rajagopal, P., Hoyt, D. W., King, M. C. & Klevit, R. E. Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. *Nature structural biology* **8**, 833-837, doi:10.1038/nsb1001-833 (2001).

71 Aravind, L. & Koonin, E. V. The U box is a modified RING finger - a common domain in ubiquitination. *Current biology : CB* **10**, R132-134 (2000).

72 Zhuang, M., Guan, S., Wang, H., Burlingame, A. L. & Wells, J. A. Substrates of IAP ubiquitin ligases identified with a designed orthogonal E3 ligase, the NEDDylator. *Mol Cell* **49**, 273-282, doi:10.1016/j.molcel.2012.10.022 (2013).

73 Wu, W. *et al.* Targeting RING domains of Mdm2-MdmX E3 complex activates apoptotic arm of the p53 pathway in leukemia/lymphoma cells. *Cell death & disease* **6**, e2035, doi:10.1038/cddis.2015.358 (2015).

74 Dou, H., Buetow, L., Sibbet, G. J., Cameron, K. & Huang, D. T. BIRC7-E2 ubiquitin conjugate structure reveals the mechanism of ubiquitin transfer by a RING dimer. *Nat Struct Mol Biol* **19**, 876-883, doi:10.1038/nsmb.2379 (2012).

Chapter 5

In vivo OUT substrate identification via proteomics and confirmation

5.1 Background information

5.1.1 Rationale

Up until this point, we had successfully reconstituted a complete OUT cascade which was proven to be capable of ubiquitinating putative substrates with xUB through our engineered xE1, xE2 and xE3 conjugating enzymes. The construction of OUT was finished; the thing that remained, the final step in our journey, was to utilize OUT to fulfill the purpose it was created for—the identification of substrates of particular E2-E3 combinations. To this end, we utilized a relatively straightforward proteomics approach, consisting of enriching xUB-conjugated proteins from the lysate in which the entire OUT cascade had been introduced—which was the original intent we had in mind when the project was started; since then however, many more sophisticated protocols in ubiquitin proteomics have been developed—these will be discussed briefly in the next section, followed by our own rather relatively simple approach to the problem and the results thereof, and finally the assays which provide confirmation of our findings.

5.1.2 Common strategies in ubiquitin proteomics

While one might consider our OUT system approach as an unprecedented novelty to some extent, the proteomics approach towards identifying substrates of ubiquitination has long been in existence. In general, the crux of this approach involves first and foremost a method to enrich and identify ubiquitinated proteins from a cell lysate. An extremely common strategy to achieve this, is what is called "di-Gly proteomics[1,2]".

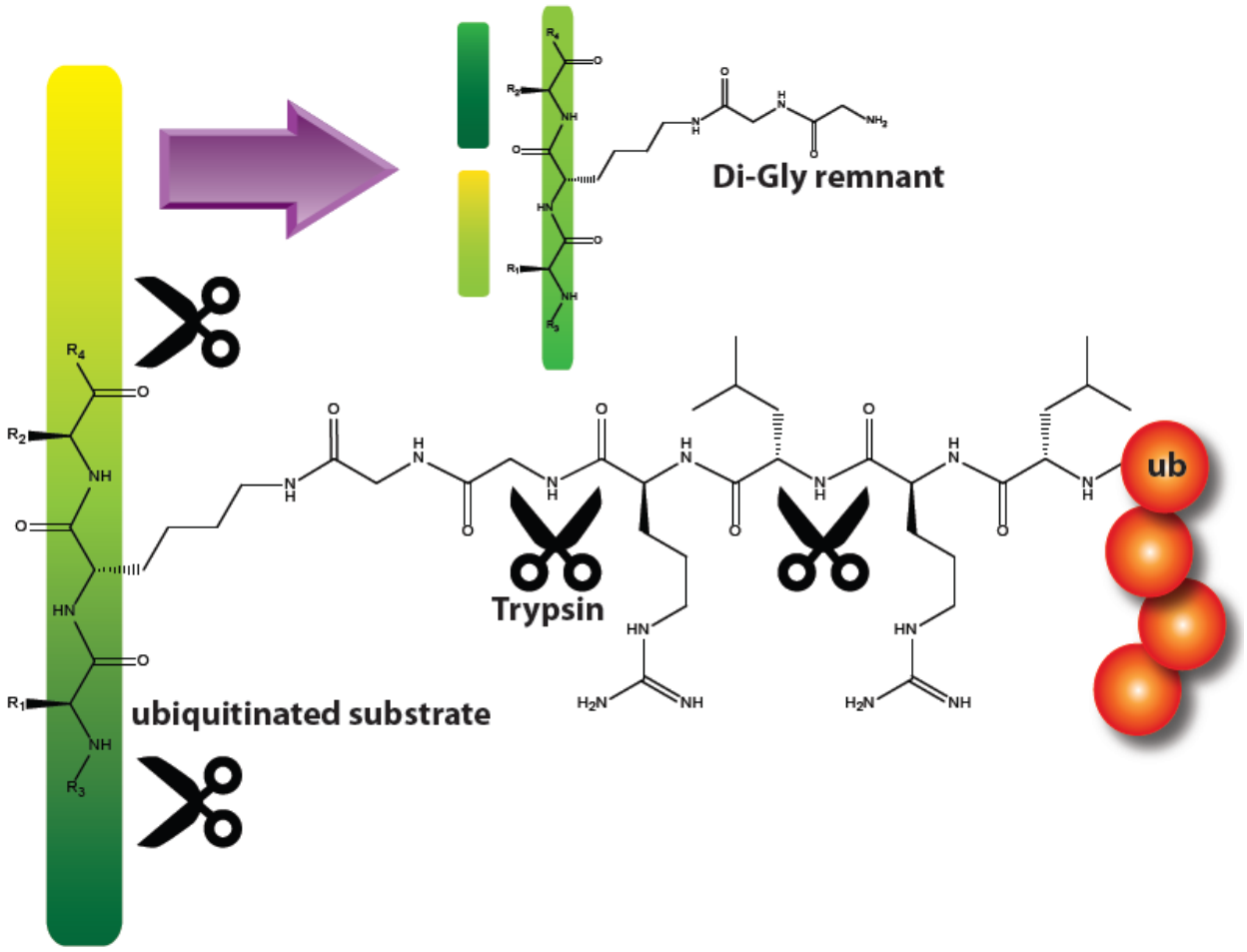

Figure 5-1 Trypsin cleavage of ubiquitinated proteins. The protease trypsin will cleave at the C-terminal of R74 of ubiquitin, exposing two glycine residues as a remnant, as well as other suitable places in the substrate. The resulting peptide fragment can then be enriched and characterized.

Ubiquitinated substrates (with possibly the exception in the case of the extremely rare N-terminal ubiquitination) contain certain lysine residues whose side chain ε-amino groups are linked through amide linkages with the C-terminal carboxyl groups of ubiquitin; (the C-terminal amino acid residues of ubiquitin have the sequence of LRLRGG) digestion with trypsin (which cleaves on the C-terminal end of either arginine or lysine residues; however it will not cleave a lysine residue which has been modified by an isopeptide bond with ubiquitin due to steric

hindrance), resulting in remnants consisting of the two terminal glycine residues from ubiquitin fused to cleaved peptide fragments from the substrate via the lysine ε-amino group[3] (**Figure 5-1**). These cleaved fragments generally are enriched via pulldown with bead-conjugated antibodies against the isopeptide-linked lysine and diglycine motif, often termed "di-Gly antibodies"[4-6] and the identity of the proteins the fragments belong to elucidated through mass spectrometry—this method can also be used to determine which lysine residues on a certain substrate are ubiquitinated[7]. Attempts to improve upon the rather mediocre affinity and specificity of the existing di-Gly antibodies have also led to the development of more sophisticated ubiquitin-specific and even chain-type specific antibodies[8], and of ubiquitin-binding domain peptides[9-12]; also, some approaches involve introduction and overexpression of exogenous ubiquitin bearing epitope tags[13,14]—some of which, such as the biotin tags in circulation[15-17]—peptide sequences biotinylated *in vivo* by endogenous eukaryotic enzymes, possess extremely strong binding with avidin derivatives such that it survives the most stringent wash conditions, and generic elution protocols are rendered inefficient—elution is generally done through enzymatic (such as TEV) cleavage though an inserted protease recognition site, or the bead-bound proteins are sent as-is for MS analysis; multiple different tags may be also used in tandem purification[18-20] to increase the specificity as well. The main drawback of this approach is that exogenous ubiquitin most likely would interfere with endogenous ubiquitin activity[9], leading to results that are not based on the natural default state of the cell.

Attempts to use the proteomics approach to determine substrates of particular E3 enzymes or systems have been reported[21-23]. These generally rely on observing the changes in the ubiquitinome—the ubiquitinated proteome when the E3 of interest is either overexpressed[21,23], or knocked down[19,24]; actually, this strategy can also be used to determine changes in the

ubiquitinome under a variety of conditions, such as under mitochondrial depolarization[25]. The comparison between the ubiquitinome of perturbed cells, and that of the control group generally involves monitoring the change in abundance of proteins enriched for ubiquitination by various different methods such as SILAC[26], which distinguishes between the two populations by different isotopic labelling and allows the simultaneous monitoring of thousands of proteins.

5.2 Results

5.2.1 Reconstitution of OUT *in vivo*

We first started by creating mammalian HEK293T stable cell lines in which each distinct component of OUT, save for xUB which was transiently transfected in the last step, was introduced sequentially in separate vectors bearing different antibiotic resistances. Each separate component of OUT was also fused with its own unique epitope tag for detection (**Figure 5-2**).

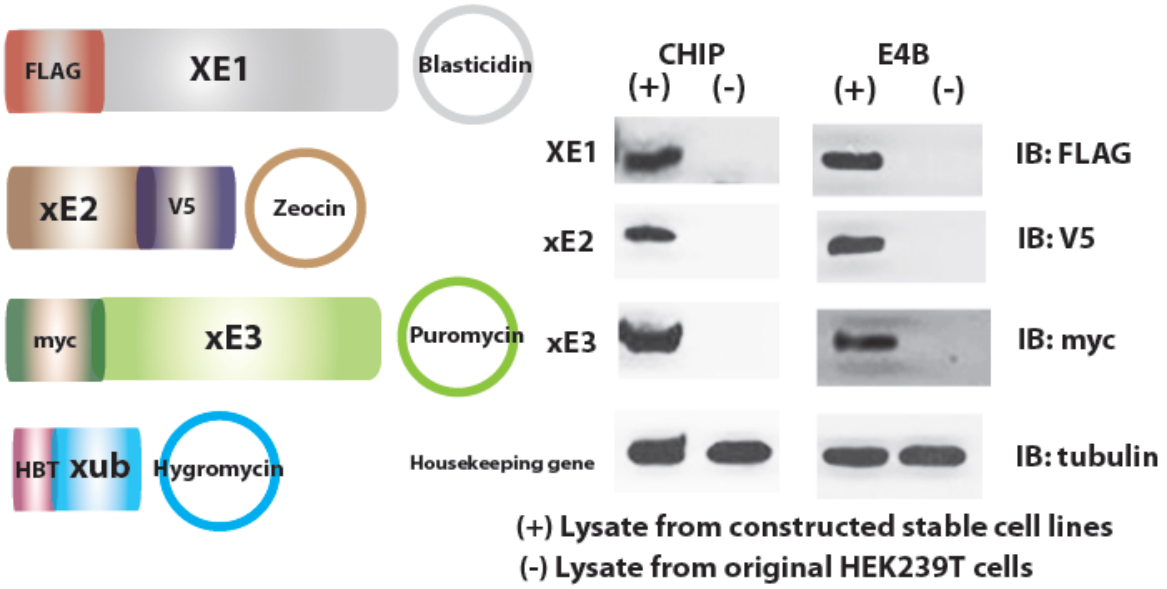

Figure 5-2 Lentiviral constructs and expression. OUT components introduced in vivo with their respective epitope tags, with Western immunoblotting confirmation of expression from cell lysate (+); controls from blank HEK293T cells on the right are signified with (-).

5.2.2 Enrichment of x-ubiquitinated proteins

Having confirmed the expression of our OUT system reconstituted *in vivo* by Western immunoblotting for each of the components, we then next sought to enrich xUB-conjugated substrate proteins to send for proteomics analysis.

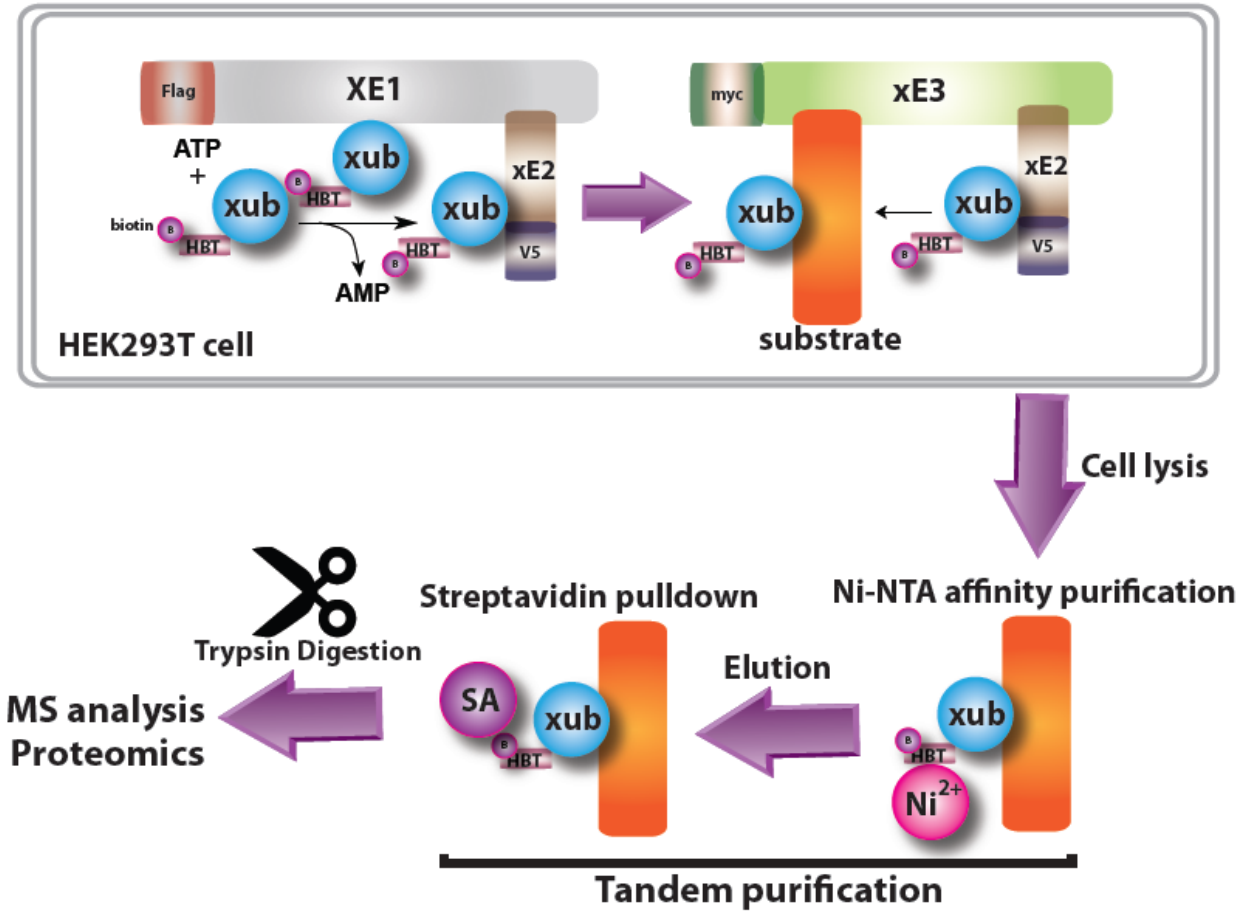

Figure 5-3 Enrichment of x-ubiquitinated proteins. From the lysate of cells in which the entire OUT cascade had been introduced, proteins ubiquitinated with xub were enriched via tandem affinity purification of the HBT tag before being sent for mass spectrometry analysis.

For this, we adopted a rather straightforward approach; our xUB construct contained an HBT tag[18], which is a 6-His tag fused to a biotin tag; these made two different ways of enriching x-ubiquitinated proteins. We performed a tandem purification of the cell lysates, first with Ni-

NTA beads to bind to the 6-His tag, and after elution with imidazole, this was followed by purification by streptavidin sepharose which allowed extremely harsh and denaturing washing conditions. We did not perform TEV cleavage, but sent the washed sepharose beads directly for analysis—in which the bound proteins were trypsinized and analyzed by mass spectrometry (**Figure 5-3**). The x-ubiquitinomes of the populations containing the whole OUT cascade with either xE4B or xCHIP, were to be compared against that of a control population, in which only xUB, xE1 and xE2 were present. This process was done in duplicate for each E3, and the potential hits of both runs were compiled together.

5.2.3 Proteomics analysis results

Protein	PSM (+E3)	PSM (-E3)	Brief Description
TP53	(31)	(0)	Tumor suppressor
MAPK1	14 (49)	0 (12)	MAP kinase
MAPK3	5	0	MAP kinase
MAPK14	(10)	(0)	MAP kinase
CDK1	5 (7)	1 (0)	Cyclin-dependent kinase
CDK2	4 (5)	0 (0)	Cyclin-dependent kinase
CDK4	6	0	Cyclin-dependent kinase
CDK6	(17)	(0)	Cyclin-dependent kinase
CDK8	(4)	0	Cyclin-dependent kinase
CDK11b	(36)	(1)	Cyclin-dependent kinase
CDK12	(13)	(0)	Cyclin-dependent kinase
CDK13	(18)	(0)	Cyclin-dependent kinase
SRC	7	0	Non-receptor tyrosine kinase
KPNA1	(3)	0	Importin subunit
KPNA2	6 (30)	0 (2)	Importin subunit
KPNA3	4 (10)	1 (2)	Importin subunit
KPNA4	6 (9)	0 (2)	Importin subunit
PGAM5	4 (35)	0 (2)	Serine/Threonine phosphatase
PPP3CA	5	0	Serine/Threonine phosphatase
OTUB1	3	1	Ubiquitin hydrolase (DUB)
PRMT1	(12)	(2)	Arginine methyltransferase
PRMT5	6	1	Arginine methyltransferase

Table 5-1 Potential substrates of E4B identified by proteomics. Shown is the list of hits from the proteomic analysis; putative substrates are in red text; PSM numbers outside the parentheses are the results from the first attempt; while the numbers inside the parentheses indicate results from the second attempt. Hits are defined as having two-fold or greater enrichment over the number of PSMs from the control cell lines without xE3, and having at least 3 PSMs identified.

Protein	PSM (+E3)	PSM (-E3)	Brief Description
TP53	3 (4)	0 (0)	Tumor suppressor
PRMT5	12 (13)	1 (0)	Arginine methyltransferase
FLNA	265 (240)	116 (3)	Actin-binding protein
HSPA1A	(91)	(18)	Heat shock 70 kDa protein
HSPA4	28 (28)	11 (0)	Heat shock 70 kDa protein
HSPA8	(57)	(12)	Heat shock cognate (Hsc70)
EIF4E	5	1	Transcription factor
FMR1	3	1	Fragile mental X retardation
ILKAP	3	1	Integrin-linked kinase
CDK4	(3)	(0)	Cyclin-dependent kinase
CDK12	(3)	(0)	Cyclin-dependent kinase
CDK13	3	1	Cyclin-dependent kinase
PPP1CA	11 (8)	3 (0)	Serine/Threonine phosphatase
PPP1CB	8	3	Serine/Threonine phosphatase
PPP1CC	9 (6)	1 (0)	Serine/Threonine phosphatase
PPP3CA	8 (12)	0 (0)	Serine/Threonine phosphatase
PPP3CB	5 (7)	1 (0)	Serine/Threonine phosphatase
PPP3CC	5 (3)	0 (0)	Serine/Threonine phosphatase
PGAM5	6	0	Serine/Threonine phosphatase
PRMT1	5	1	Arginine methyltransferase
OTUB1	5 (3)	1 (0)	Ubiquitin hydrolase (DUB)
KPNA2	21 (5)	0 (2)	Importin subunit
KPNA3	4 (4)	1 (2)	Importin subunit
KPNA4	6 (9)	0 (2)	Importin subunit
CTNNBL1	5	0	Beta-catenin like protein
HSPA5	(17)	(6)	Heat shock 70 kDa protein
HSPA9	(7)	(0)	Heat shock 70 kDa protein

Table 5-2 Potential substrates of CHIP identified by proteomics.
In the same manner as **Table 5-1**, putative substrates are given in red text; PSM numbers outside the parentheses are the results from the first attempt; while the numbers inside the parenthesis indicate results from the second attempt. Hits are defined as having two-fold or greater enrichment over the corresponding PSM from the control cell lines without xE3, and having at least 3 PSMs identified.

The analysis supplied us with potential substrates detected by mass spectrometry, which were reported in the form of PSMs (peptide spectrum matches)—which corresponded to the number of peptide fragments identified as potentially belonging to the particular proteins. The observed PSMs from enriched proteins from cells containing the full OUT cascade: XE1, xE2, and either xE4B (fKB2) or xCHIP (xCHIP12) were compared against the control derived from cells in which the xE3 component was missing; we arbitrarily screened potential hits by these two criteria: having three or more PSMs detected, and showing at least two-fold enrichment over the PSMs of the control group; this resulted in around 1000+ hits of potential substrates for both E3s. Some particularly biologically-significant hits in our analysis for each group were given in **Table 5-1** and **Table 5-2** above; the full lists of hits are included in the supplementary information section.

5.2.4 Verification of potential substrates

5.2.4.1 *In vitro* ubiquitination assays

In vitro verification of identified hits from our proteomics analysis was carried out using similar protocols to the previously mentioned assays on putative substrates such as p53 or Hsc70, with wild-type ubiquitin and ubiquitin conjugating enzymes to eliminate the possibility of false positives which might have arisen as the result of non-native interactions caused by the specific mutations in xE3 which conferred OUT activity. In a somewhat arbitrary manner, we focused our efforts generally on well-studied substrates which play well-elucidated roles in significant biological pathways. The majority of substrates tested were found to be ubiquitinated by their respective E3s *in vitro* as long as suitable conditions (which may vary from substrate to substrate) were used. The more significant assay results were as given for E4B (**Figure 5-4**) and CHIP (**Figure 5-5**).

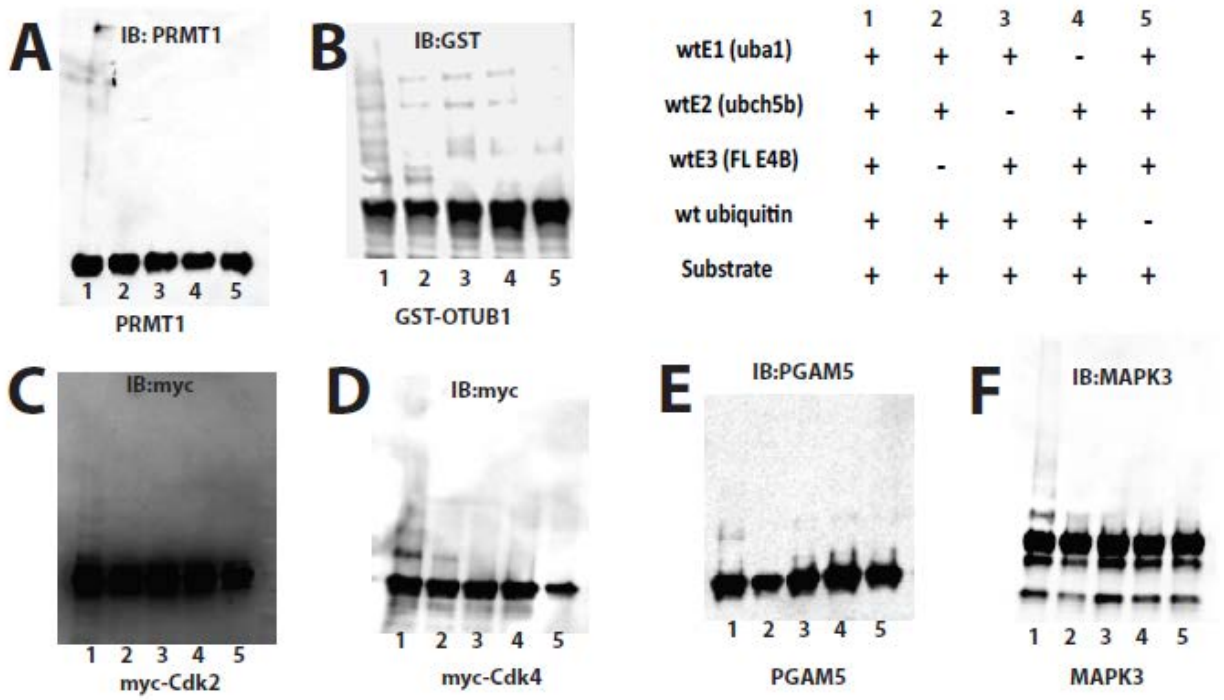

Figure 5-4 *In vitro* verification of E4B substrates. Potential substrates were tested for ubiquitination by wild-type E1, E2, ubiquitin and full-length E4B proteins; Western immunoblotting was done using antibodies either against the substrates themselves or their fused epitope tags.

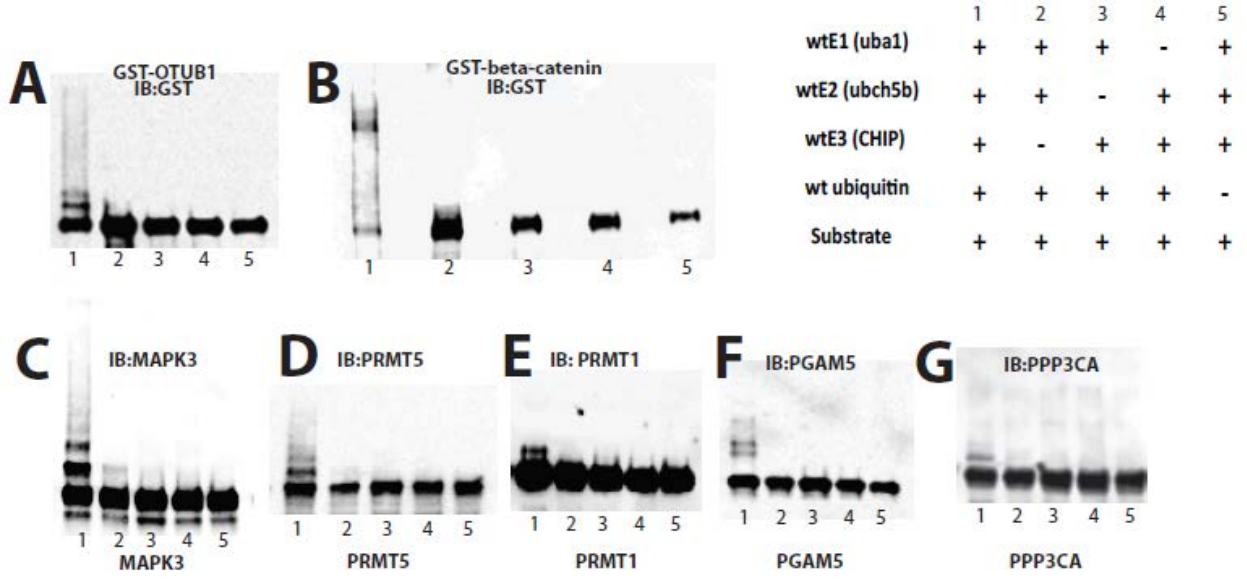

Figure 5-5 *In vitro* verification of CHIP substrates. Potential substrates were tested for ubiquitination by wild-type E1, E2, ubiquitin and CHIP proteins; Western immunoblotting was done using antibodies either against the substrates themselves or their fused epitope tags

5.2.4.2 *In vivo* verification of potential substrates

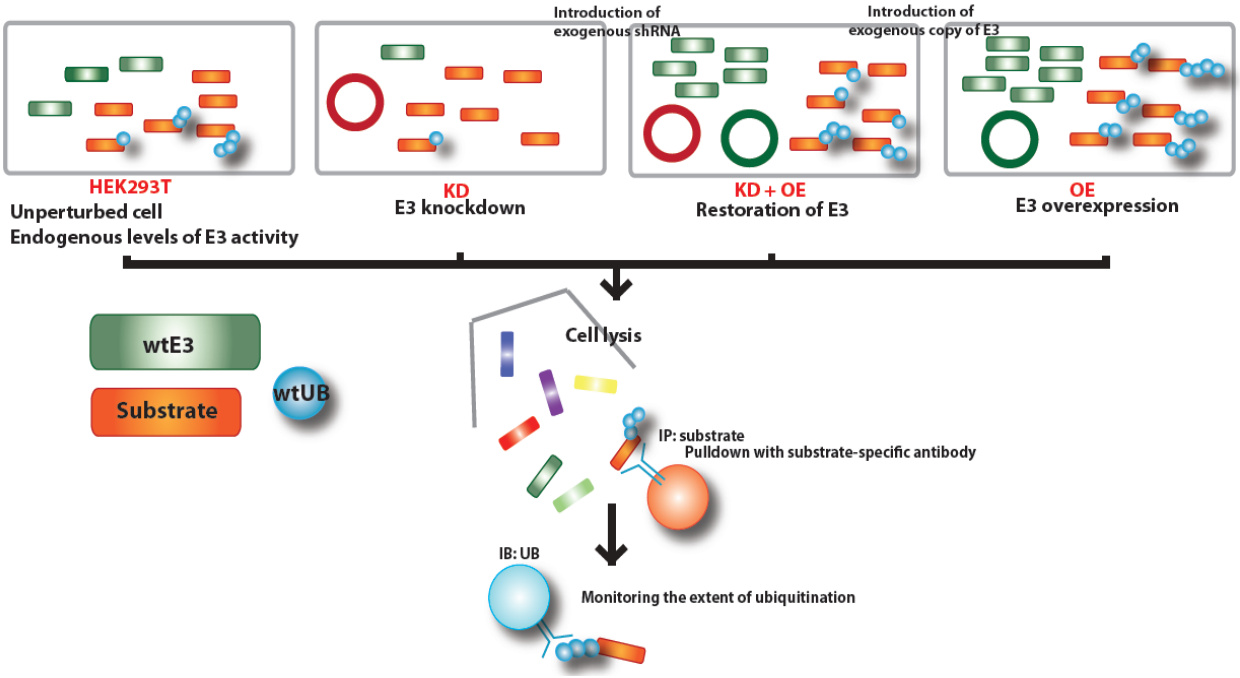

Figure 5-6 Probing the effects of cellular E3 levels on substrate ubiquitination.
Four different populations of cells were assayed for their levels of ubiquitination of the substrate of interest; they consist of unperturbed cells (HEK293T), cells in which the E3 of interest had been knocked down (KD) via introduction of exogenous shRNA, cells in which after E3 knockdown an exogenous copy of the E3 of interest had been introduced to restore E3 activity (KD+OE), and finally cells in which the E3 of interest had been simply overexpressed without knockdown (OE). Following cell lysis, the substrate of interest was enriched and assayed for ubiquitination levels.

In vivo assays, involving mammalian HEK293T cells, of certain promising potential substrates which showed high levels of activity *in vitro* was also carried out using a significantly more sophisticated protocol, which involved monitoring the extent of ubiquitination of the target substrate by first enriching the target of interest via affinity pulldown and probing with anti-ubiquitin antibodies after Western immunoblotting (**Figures 5-7** and **5-8**). We compared the levels of ubiquitination of potential substrates in four populations of cells (**Figure 5-6**): unperturbed cells, cells in which an introduced exogenous copy of the proper E3 (E4B or CHIP)

was overexpressed, cells in which the E3 had been knocked down (through the creation of a stable cell line in which the genes transcribing shRNAs were introduced in a vector), and finally cells in which the effects of the knockdown of the E3 in question had been reversed through the introduction and overexpression of exogenous E3.

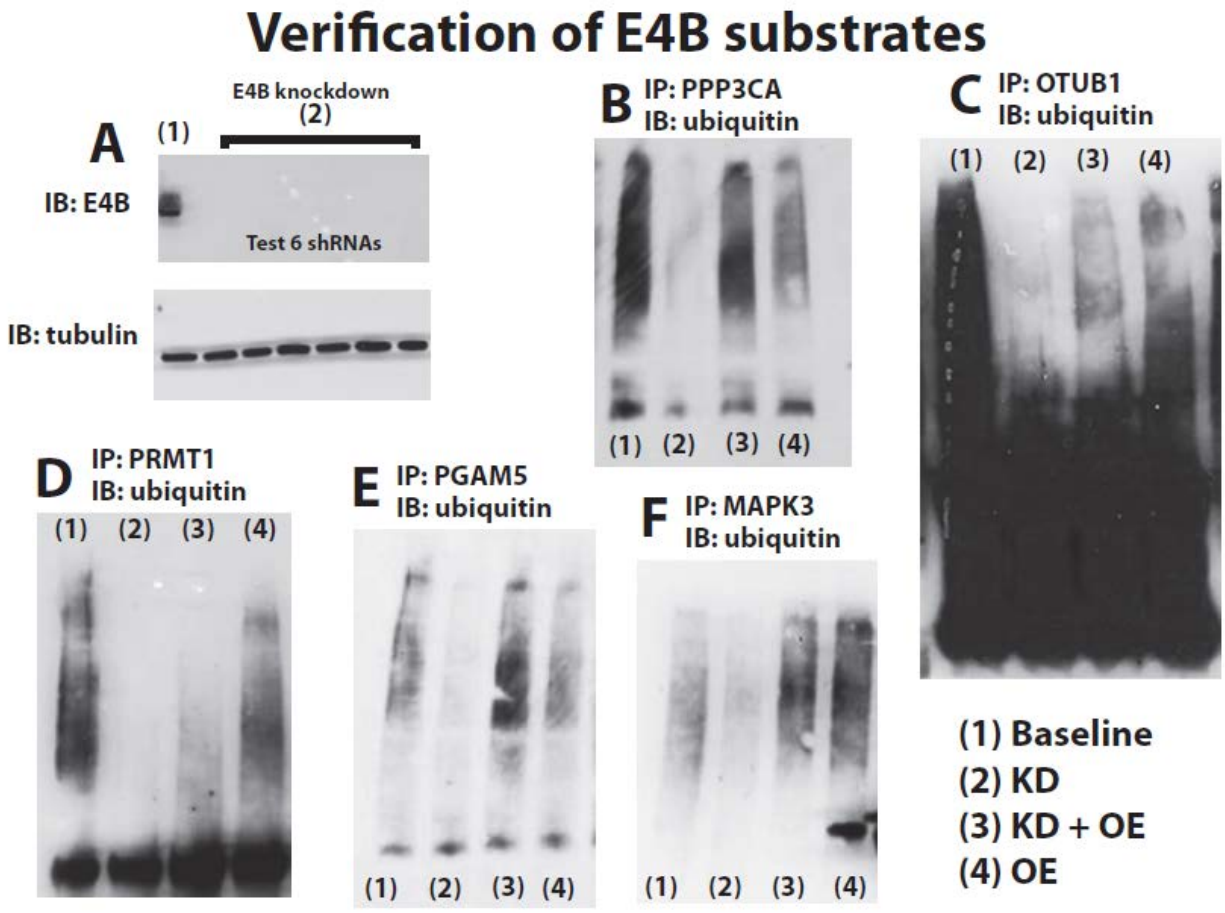

Figure 5-7 *In vivo* verification of E4B substrates.
To achieve knockdown of E4B, six stable cell lines containing six different commercial E4B shRNA constructs were created; the success of E4B knockdown was monitored by Western immunoblotting of lysates with antibodies specific against E4B (top left, boxed area). Shown is the control from unperturbed HEK293T cells (1), and the lysates from the different knockdown cell lines (2). Tubulin levels were monitored as standard controls. Potential substrates of E4B from proteomics analysis were enriched via immunoprecipitation out from the cell lysates of four different cell populations (1) to (4) (refer to **Figure 5-6**), using substrate-specific antibodies. The levels of ubiquitination of substrates were monitored by Western immunoblotting with antibody against ubiquitin.

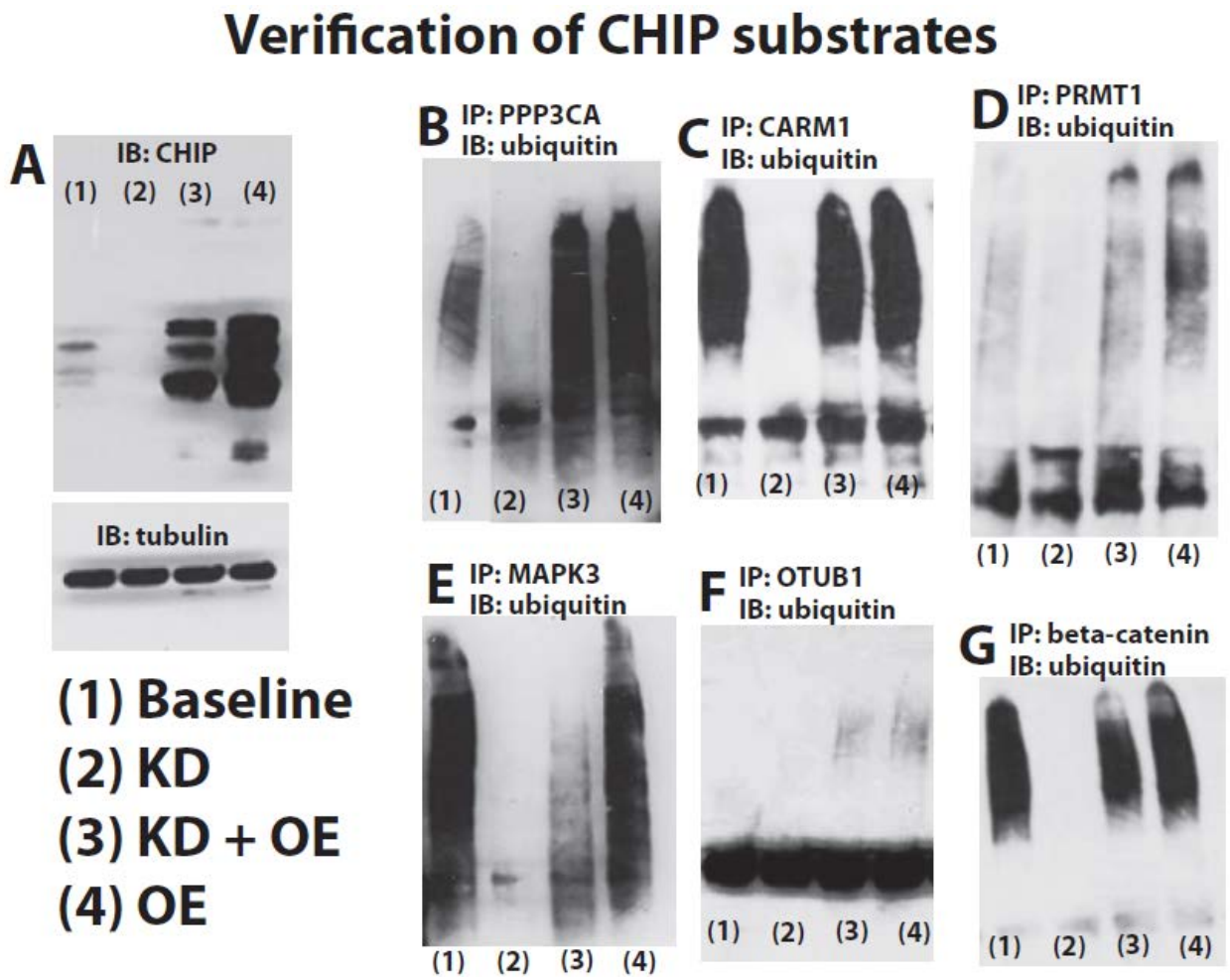

Figure 5-8 *In vivo* verification of CHIP substrates.
The cell lysates of four different cell populations (1) to (4) (refer to **Figures 5-6** and **5-7**), were monitored for the success of either knockdown or overexpression of the CHIP protein (boxed area), with tubulin levels included as standard controls. Potential substrates were also enriched via immunoprecipitation from these four different lysates using substrate-specific antibodies, and the levels of ubiquitination were monitored by Western immunoblotting against ubiquitin, in the same exact manner as **Figure 5-7**.

5.3 Discussion

5.3.1 The OUT cascade successfully reconstituted *in vivo*

Up until this point, all of our assays involving the individual steps of our OUT cascade had been verified solely *in vitro*; our results provided much needed confirmation that our system could also function efficiently *in vivo* as well; we identified the PSMs of the respective U-box

E3s in the list of proteins enriched, and hence this suggested that our transfected xE3s underwent some certain extent of autoubiquitination with xUB, which would be consistent with our previous *in vitro* results and would affirm that indeed each component, and each ubiquitin transfer step in the entire OUT cascade was functioning; the fact that there was a very obvious difference in the PSMs of the E3s between the xE3-transfected group the control group strongly suggested that these fragments indeed mostly came from the introduced xE3 and not from their endogenous wild-type counterparts, which in any case would not be expected to be active with the OUT system and hence, not enriched by our tandem purification methods.

5.3.2 Identified proteomics hits

Through our screening, we identified multiple reported putative substrates of CHIP among our hits, which included targets we had previously characterized such as p53[27], Hsc70[28] and Hsp70[29] (a homolog of Hsc70 which shares 94% sequence identity), as well as other putative substrates previously reported by other groups[30-34] and also some keratin proteins—mutant keratins have been reported to be targeted by CHIP as well[35]; the tentative list of identified putative substrates were given in the results section. We had good reason to believe that our approach with CHIP was successful, and that a good portion of the hits could very well contain *bona fide* substrate of CHIP, many of which would be novel discoveries.

In contrast, the success of the E4B substrate identification proteomics screen was more difficult to evaluate from the list of hits alone, due to the scarcity of reported substrates. To our knowledge, there had only been one confirmed substrate of E4B acting singlehandedly as an E3 enzyme *in vitro*, FEZ1[36]—we did not detect any PSMs belonging to this protein, nor ataxin-3[37], in any of our proteomics analyses, including the control group; however, we did detect p53[38-40]—

the ubiquitination of which with the full-length E4B protein we had successfully reconstituted *in vitro* in Chapter 4, as a very strong hit in one of our runs.

There are several possible explanations as to why certain putative substrates were not identified in our screens—we did not identify the majority of the precious few putative E4B substrates, and numerous putative CHIP substrates were also not detected. Firstly, that there may be insufficient sensitivity in our assay to detect very trace amounts of proteins; perhaps the proteins in question exist only at very low endogenous levels and/or the fraction of the proteins which exist in the ubiquitinated form in the cell at a given time is too small. Secondly, it may be that ubiquitination of these substrates required a different E2; our approach only involved one variant of xE2 which is ubch5b—generally, the ubch5 variants are the canonical E2s which interact with RINGs and U-boxes; however both CHIP[41] and E4B[42] have been reported to interact with other E2 enzymes. Finally, it may be possible that mutations in the OUT system, whether on xUB, xE2 or xE3 had an unexpected effect upon substrate recognition, or that the overexpression of OUT components disrupted the default balance of the cell by competition with existing endogenous wild-type counterparts, resulting in significant changes in the cellular levels of certain proteins—for instance, we have not positively confirmed whether or not x-ubiquitin is recognizable by the proteasome, or by the multitude of deubiquitinating enzymes inside the cell, or whether or not x-ubiquitin is capable of being incorporated into preexisting wild-type ubiquitin chains, in other words, whether the formation of wild-type/x mixed chains is possible.

As far as the hits themselves were concerned, both E3s seemed to target various different proteins across many different biological pathways; in particular, we noticed that E4B seemed to target a wide variety of kinase enzymes: the MAPKs, the CDKs, as well as the Src kinase, as well as some phosphatases; we also noted the relative abundance of proteins involved in

transmembrane transport such as the different subunits of importin as well as some pore proteins. In a similar manner, CHIP seemed to favor many of the same kinase and phosphatase enzymes, as well as some transport proteins as did E4B; which might suggest that these are common substrates in general for U-box type ligases; Also, CHIP seemed to target multiple heat shock response proteins, as well as many proteasome-associated proteins—consistent with its known role in protein quality control[43-47]. In any case, IPA analysis of the complete lists of our hits implied that E4B substrates are involved in remodeling of epithelial adherens junctions, integrin signaling and RhoA signaling; associated physiological systems include embryonic development and hair/skin development, whereas CHIP substrates are involved in tRNA charging and mTOR signaling, and two of the top protein networks involving CHIP substrates are related to cancer and neurological disorders—this is consistent with the conclusions of numerous other studies on the CHIP protein[48-58].

There had been a concern regarding the "leakiness" of the OUT cascade—that certain OUT components still retained some background activity with their wild-type partners, leading to false positives. The presence of certain E2 and E3 enzymes among the proteomics hits might be an indicator of this, but not necessarily so, as some E3 enzymes have been known to cross ubiquitinate other E3s[59,60], as well as catalyze autoubiquitination of E2s[61] to consign them for degradation; the biological significance of this phenomenon is not completely understood, but it is conceivable that the cell may utilize this to autoregulate its own ubiquitination machinery in a manner similar to the observed autoubiquitination of RINGs and U-boxes when they are no longer needed in the absence of substrates[62].

5.3.3 Verification of potential substrates

Ultimately, success in verification of E3 ubiquitination activity on identified hits would be the best indicator as to the effectiveness of our approach. We found that *in vitro* ubiquitination assays with wild-type enzymes were successful for the majority of particular hits which we decided to further investigate—generally we intended to choose proteins which had been well-studied and had their activities successfully reconstituted. Some proteins were not observed to be ubiquitinated using our generic protocols—among them the Src kinase, and the importin subunit KPNA2 (data not shown); this might have been due to a myriad of possible different reasons besides those hits being false positives in the first place, such as the fact that all of the tested potential substrate proteins were recombinantly expressed in E.coli using our generic protocol; there might have been a lack in proper chaperoning in ensuring the proteins fold properly, or necessary post translational modifications which were required for them to be recognized by ubiquitination enzymes, or that certain cellular conditions significantly different from the *in vitro* conditions in the assay were required.

We also performed *in vivo* verification on some particular selected substrates as well; and obtained some particularly convincing results. Specifically, we would take the sample from the unperturbed cells to represent the baseline level of ubiquitination of that particular substrate with existing endogenous enzymes—this would vary from substrate to substrate; when we overexpressed the E3 in question, ideally we would expect to see an increase in the ubiquitination level of the substrate, and conversely, when the E3 was knocked down, we would also expect the extent of ubiquitination to decrease; finally, when the E3 activity was restored in these knocked-down cells through introduction of an exogenous copy of E3, we would ideally

expect to see ubiquitination of the substrate restored again. We observed that this hypothesis was apparently confirmed in many cases (**Figures 5-7** and **5-8**).

In some instances we observed that the baseline levels of ubiquitination of certain potential substrates were already high, and/or that overexpression of E3s did not seem to significantly alter ubiquitination levels. However, if a marked decrease in ubiquitination was observed when the appropriate E3 was knocked down, we still had good reason to believe those proteins may yet be *bona fide* substrates—most likely the endogenous level of the E3 was enough to saturate the ubiquitination reactions of those proteins. In other instances, we observed that knockdown of the target E3 did not significantly decrease the ubiquitination levels of certain proteins; and yet, if we simultaneously observed increases in ubiquitination levels when the E3 was overexpressed, then those substrates should not be ruled out—while overexpression of the E3 of interest may result in the emergence of some promiscuous non-native activity on certain substrates it would not ubiquitinate under normal circumstances, there is also the possibility that when that particular E3 is knocked down, the cell responds by upregulating the activities of other E3s capable of targeting the particular substrate to compensate.

5.3.4 Conclusion and future directions

Through the development and utilization of OUT, we have successfully identified various potential substrates for the U-box E3s E4B and CHIP—many of which have never been reported as such. The results of our *in vitro* and *in vivo* assays in this chapter provides very strong evidence that these proteins are indeed true substrates of their respective E3s; with this, the efforts of our entire investigation were vindicated.

We have by no means experimented with all the possible different hits our proteomics analyses have yielded—we have tested less than 1% of such, and yet each one of the substrates we have verified would easily comprise its own topic of research in the future; there would still be questions to be answered such as those regarding the type of chain linkage, the fate of the ubiquitinated protein, the cellular signal its ubiquitination represents and the cellular response it generates, etc. It would also be interesting to apply the OUT approach to other families of E3s, or even to adapt it to systems involving other ubiquitin-like proteins which conjugate to substrates via similar enzymatic cascades such as Nedd8, SUMO, or ISG15. Stated here are but a few of the virtually innumerable possibilities.

The proteomics approach we have taken in general is also not limited to simply monitoring cells in unperturbed states; it is theoretically possible to utilize OUT substrate identification in a wide variety of different cell states and conditions, such as under different cell cycle stages, or different types of cell stress, or in diseased or cancerous tissue cells; it would be expected that the activity of enzymes in the ubiquitin cascade would undergo significant changes under different cell conditions, and hence our system may potentially be used to monitor and evaluate those changes.

In conclusion, the journey we have undertaken has come to an end—we have reached our destination; however, this destination is but a starting point for many more journeys to be undertaken. The success of the OUT system has opened up a multitude of avenues for future research; it is likely that it will continue to prove its usefulness to our group and the field of ubiquitin research as a whole for many years to come.

5.4 Methods

The primers used in this chapter are as given in **Table 5-3**. The shRNA constructs, as well as all the epitope-specific and substrate-specific antibodies were obtained commercially from Santa Cruz Biotechnologies; a full list of shRNA constructs and antibodies is given in the supplementary information section at the end of this chapter.

Primer name	Sequence (5'-3')
CHIP XhoI stop	TGGTGGTGCTCGAGGTCTAGTCAGTAGTCCTCCACCCAGCCATTCTCAGA GATGAATGCG
myc CHIP EcoRI up	CAGCCCAGAGAATTCTATGGAACAAAAACTCATCTCAGAAGAGGATCTG GCTAGCAAGGGCAAGGAGGAGAAGGAGGGCGGCGCACGG
For Yiyang 1 corr	CAGCCCAGACTCGAGTATGGAACAAAAACTCATCTCAGAAGAGGATCTG GCTAGCGAGGAGCTGAGCGTGAC
E4B BsrGI middle	TTTTATTGTACAATACTCTCCTCAGGTGCTTTATGAGCCC
E4B AgeI with stop	GATGATACCGGTTCAGTGGTCACTGCTCTGTTTCTCTCTCATCC
Cdk2 SacII up myc	GACAAGCCGCGGGAACAAAAGCTTATTTCTGAAGAAGACATGGAGAAC TTCCAAAAGGTG
Cdk2 stop NotI down	CCCGGCTGCGCGGCCGCTCAGAGTCGAAGATGGGGTACTGGCTTGG
Cdk4 ScaI up	GATATCGAGTACTATGGCTGCCACTCGATATGAACCCG
Cdk4 stop notI down	CCCGGGCTGCGGCCGCTCACTCTGCGTCGCTTTCCTCCTTGTGCAGG

Table 5-3 Primers used in Chapter 5.
Residues in red indicate restriction sites

5.4.1 Lentiviral constructs

Constructs of wild type and xE1 (ube1—the mammalian homolog of uba1) bearing N-terminal FLAG epitope tags in pLenti (blasticidin), as well as wild-type and xUB bearing N-terminal HBT tags in pLenti (hygromycin) were gifts from Hiroaki Kiyokawa of Northwestern University. Wild-type and xE2 (ubch5b) with C-terminal V5 epitope tags, were cloned into pLenti (Zeocin) via AfeI/NheI restriction sites by a postdoctoral fellow of our group.

Wild-type and xCHIP12 with N-terminal myc epitope tags introduced via primer sequence were PCR-amplified using their respective pET constructs (4.4.2) as DNA template, with the “CHIP XhoI stop” and “myc CHIP EcoRI up” primers, digested with ligated into similarly digested and dephosphorylated pLenti (puromycin) via EcoRI/XhoI restriction sites.

fKB2 (full-length xE4B) with N-terminal myc epitope tag introduced via primer sequence was PCR-amplified using its pET construct (4.4.1) as DNA template, with the “for Yiyang 1 corr” and “E4B AgeI with stop” primers, digested, and ligated into similarly digested and dephosphorylated pLenti (puromycin) via XhoI/AgeI restriction sites.

Wild-type E4B sequence was PCR amplified up using the pET construct (4.4.1) as DNA template with the “E4B BsrGI middle” and “E4B AgeI with stop” primers, digested and ligated into similarly digested pLenti-fKB2 (see above paragraph), utilizing the internal BsrGI site in the E4B gene and the terminal AgeI sites to construct pLenti-fE4B.

All recombinant DNA manipulations involving xE3 enzymes were accomplished in NEBStable cells (NEB), which possess greatly enhanced DNA stability and reduced possibility of recombination.

5.4.2 Creation of stable cell lines/transient transfection

For transient transfection, the protocols were exactly as described as follows in 5.4.2.1 and 5.4.2.2; the main difference between transient transfection and creating a stable cell line being the lack of antibiotic selection; the transiently-transfected cells were assayed immediately on the third day of lentiviral transduction.

5.4.2.1 Production of lentiviral particles

We generally followed the existing protocols according to the Virapower™ kit from Invitrogen. In 1.5 ml of Opti-MEMR 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 293FT 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_2. 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_2. Afterwards, the supernatant containing the lentiviral particles was harvested and cleared by filtration.

5.4.2.2 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 ug/ml Polybrene®; the cells were incubated overnight at 37˚C at 5% CO_2. 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, or by using a generic lysis protocol given in 5.4.3.1.

5.4.3 Pulldown and purification/enrichment assays

5.4.3.1 Lysis of cells

For generic assays, HEK293T cells were generally lysed by suspension in RIPA buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate and 1% Triton-X) supplemented with 1 mM PMSF.

Alternatively, lysis under harsher denaturing conditions, as was in the case of HBT tandem purification of xUB-conjugated proteins, could be carried out using Buffer A (50 mM sodium phosphate pH 8.0, 300 mM NaCl, 8 M Urea, 0.5% NP-40) with 1 mM PMSF.

After lysis, lysates were cleared by centrifugation at 10,000+ rpm for at least 10 minutes to remove cellular debris.

5.4.3.2 Generic pulldown of target proteins

For most pulldowns, cleared cell lysates were generally first pre-incubated for 30 minutes at 4˚C with the appropriate IgG (from the same animal sources as the antibody to be used) and protein A/G plus-agarose (Santa Cruz) to remove background proteins which bind to the IgG portion of the antibody; after centrifugation, the agarose along with any bound proteins was

discarded. An appropriate amount of the remaining supernatant (enough to contain up to ~500 μg total protein) was then charged with the appropriate pulldown antibody (concentrations vary between particular substrate-antibody pairs), and incubated for 1 hour at 4˚C; 20 μL of protein A/G plus-agarose slurry was then added and incubation at 4˚C was continued overnight. Afterwards, the agarose was washed 4 times with PBS; samples for subsequent electrophoresis on SDS-page and Western immunoblotting were obtained by boiling the agarose in Laemmli buffer to release a portion of bound proteins.

5.4.3.3 Tandem purification for HBT tag

To enrich proteins ubiquitinated with xUB (which bore an HBT tag), the cells were first lysed using the alternative protocol as described in 5.4.3.1 with buffer A; the subsequent lysate was then cleared by centrifugation at 15000 rpm for 30 minutes, and incubated overnight with Ni-NTA Agarose (Thermo), imidazole was also added to a final concentration of 10 mM to reduce non-specific binding. Afterwards, the agarose was washed once with buffer A (pH 8.0), then once with buffer A (pH 6.3) and finally with imidazole-supplemented buffer A (pH 6.3, 10 mM imidazole). Bound proteins were eluted twice, each time with 3 resin volumes of buffer B (50 mM sodium phosphate, 100 mM Tris pH 4.3, 200 mM NaCl, 8 M urea, 2% SDS, 10 mM EDTA, 250 mM imidazole); afterwards the pH of the eluate was adjusted to 8.0, and incubated with Streptavidin-sepharose overnight at room temperature. The sepharose beads were then washed sequentially once each with Buffers C (100 mM Tris pH 8.0, 200 mM NaCl, 8 M urea, 2% SDS), D (100 mM Tris pH 8.0, 1.2 M NaCl, 8 M urea, 0.2% SDS, 10% ethanol, 10% isopropanol), and E (8 M urea, 100 mM ammonium bicarbonate); the bead-bound proteins were then sent for proteomics MS analysis.

5.4.4 Potential substrate verification.

Consult the supplementary information section for the complete list and concentrations of antibodies used in Western immunoblotting, and pulldown assays.

5.4.4.1 Procurement of potential substrates.

The majority of potential substrates in this chapter tested in vitro were expressed directly from constructs purchased online from Addgene. A complete list of purchased substrates, as well as their purification methods are as given below in **Table 5-4.** The expression constructs of PRMT1 and PRMT5 were gifts to our lab.

Table 5-4 List of purchased constructs.

Substrate	Construct name	Addgene #	Purification
MAPK3	NpT7-5-Erk1	39299	6-His
PPP3CA	pET15b CnA CnB	11787	6-His
PGAM5	PGAM5	38976	6-His
OTUB1	pOPINK-OTUB1	61420	GST
β-catenin	pGEX-Bcatfl	24193	GST

5.4.4.2 Construction of expression constructs of potential substrates and protein expression.

Cdk2 and Cdk4 constructs in mammalian vectors were gifts from Hiroaki Kiyokawa from Northwestern University these were used as DNA templates for bacterial expression constructs in the pET system; the Cdk2 gene was PCR-amplified with the “Cdk2 SacII up myc” and the “Cdk2 stop NotI down” primers, digested, and ligated into similarly digested and dephosphorylated pET-myc-smurf2 via SacII/NotI, while Cdk4 was PCR-amplified up with

"Cdk4 ScaI up" and "Cdk4 stop NotI down" primers, digested, and ligated into the same mother vector, except through ScaI/NotI instead.

Protein expression was done in either BL21(DE3) or ArcticExpress(DE3) cells, and subsequent lysis and affinity purification (6-His or GST), as well as dialysis of purified proteins followed previously described protocols in 2.4.2. Success of protein expression was monitored by Coomassie staining, and by Western immunoblotting.

5.4.4.3 *In vitro* ubiquitination assays of potential substrates

Table 5-5 Reaction conditions for *in vitro* ubiquitination of substrates.

Substrate Name	[uba1] (μM)	[ubch5b] (μM)	[fE4B] (μM)	[ub] (μM)	[substrate] (μM)	Reaction time
MAPK3	1	2	1	100	10	2 hours
CDK2	1.5	10	1	100	10	Overnight
CDK4	1	10	1	80	10	2 hours
OTUB1	1.5	20	1	100	10	4 hours
PGAM5	1.5	20	1	120	10	Overnight
PRMT1	1.5	20	1	120	10	Overnight
Substrate Name	**[uba1] (μM)**	**[ubch5b] (μM)**	**[CHIP] (μM)**	**[ub] (μM)**	**[substrate] (μM)**	**Reaction time**
OTUB1	1	2	20	100	10	2 hours
β-catenin	1	2	20	100	5	2 hours
MAPK3	1	2	20	100	10	2 hours
PRMT5	1	2	20	100	10	2 hours
PRMT1	1	2	20	80	10	5 minutes
PGAM5	1.5	20	20	120	10	Overnight
PPP3CA	1	2	20	80	10	15 minutes

All assays were carried out at 30˚C, in 30 μL of TBS supplemented with 10 mM MgCl2 and 1.5 mM ATP. Specific assay conditions varied from substrate to substrate, depending on

observed reactivity. 5-10 μM of potential substrates were first incubated with their respective wild-type E3s (20 μM for CHIP, 1 μM for fE4B) for ~30-60 minutes on ice before being charged with 1-1.5 μM wild-type E1 (uba1), 2-20 μM wild-type E2 (ubch5b) and 80-120 μM wild-type ubiquitin; the reactions were allowed to incubate for varying durations from 10 minutes to overnight. The reactions were quenched by boiling in Laemmli buffer with BME, and analyzed by Western immunoblotting against either substrate-specific antibodies, or antibodies against certain epitope tags the substrates were fused to. The specific reaction conditions, including enzyme concentrations, for each substrate was as given above in **Table 5-5**; the specific antibodies used in Western immunoblotting are listed in the supplementary information section.

5.4.4.4 *In vivo* confirmation of activity

The construction of stable cell lines containing the genes for E3 shRNA followed the protocols described in 5.4.2.1 and 5.4.2.2; the list of shRNAs tested are given in the supplementary information. Overexpression of wild-type E3s from their pLenti (puromycin) constructs (described in 5.4.1) followed the protocol for transient transfection in 5.4.2.

Lysis of cells and subsequent pulldown of particular substrates followed protocols 5.4.3.1 and 5.4.3.2, respectively; consult the supplementary information section for specific antibodies and conditions. Beads were boiled in Laemmli buffer supplemented with BME, and samples were analyzed for ubiquitin conjugates by Western immunoblotting with a ubiquitin-specific antibody (Santa Cruz).

5.5 References

1 Thompson, J. W. *et al.* Quantitative Lys--Gly-Gly (diGly) proteomics coupled with inducible RNAi reveals ubiquitin-mediated proteolysis of DNA damage-inducible transcript 4 (DDIT4) by the E3 ligase HUWE1. *The Journal of biological chemistry* **289**, 28942-28955, doi:10.1074/jbc.M114.573352 (2014).

2 Xu, G., Deglincerti, A., Paige, J. S. & Jaffrey, S. R. Profiling lysine ubiquitination by selective enrichment of ubiquitin remnant-containing peptides. *Methods in molecular biology (Clifton, N.J.)* **1174**, 57-71, doi:10.1007/978-1-4939-0944-5_4 (2014).

3 Beaudette, P., Popp, O. & Dittmar, G. Proteomic techniques to probe the ubiquitin landscape. *Proteomics* **16**, 273-287, doi:10.1002/pmic.201500290 (2016).

4 Xu, G., Paige, J. S. & Jaffrey, S. R. Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nature biotechnology* **28**, 868-873, doi:10.1038/nbt.1654 (2010).

5 Wagner, S. A. *et al.* Proteomic analyses reveal divergent ubiquitylation site patterns in murine tissues. *Mol Cell Proteomics* **11**, 1578-1585, doi:10.1074/mcp.M112.017905 (2012).

6 Povlsen, L. K. *et al.* Systems-wide analysis of ubiquitylation dynamics reveals a key role for PAF15 ubiquitylation in DNA-damage bypass. *Nature cell biology* **14**, 1089-1098, doi:10.1038/ncb2579 (2012).

7 Danielsen, J. M. *et al.* Mass spectrometric analysis of lysine ubiquitylation reveals promiscuity at site level. *Mol Cell Proteomics* **10**, M110 003590, doi:10.1074/mcp.M110.003590 (2011).

8 Phu, L. *et al.* Improved quantitative mass spectrometry methods for characterizing complex ubiquitin signals. *Mol Cell Proteomics* **10**, M110.003756, doi:10.1074/mcp.M110.003756 (2011).

9 Gao, Y. *et al.* Enhanced Purification of Ubiquitinated Proteins by Engineered Tandem Hybrid Ubiquitin-binding Domains (ThUBDs). *Mol Cell Proteomics*, doi:10.1074/mcp.O115.051839 (2016).

10 Scott, D., Oldham, N. J., Strachan, J., Searle, M. S. & Layfield, R. Ubiquitin-binding domains: mechanisms of ubiquitin recognition and use as tools to investigate ubiquitin-modified proteomes. *Proteomics* **15**, 844-861, doi:10.1002/pmic.201400341 (2015).

11 Aillet, F. *et al.* Isolation of ubiquitylated proteins using tandem ubiquitin-binding entities. *Methods Mol Biol* **832**, 173-183, doi:10.1007/978-1-61779-474-2_12 (2012).

12 Lopitz-Otsoa, F. *et al.* Integrative analysis of the ubiquitin proteome isolated using Tandem Ubiquitin Binding Entities (TUBEs). *Journal of proteomics* **75**, 2998-3014, doi:10.1016/j.jprot.2011.12.001 (2012).

13 Oshikawa, K., Matsumoto, M., Oyamada, K. & Nakayama, K. I. Proteome-wide identification of ubiquitylation sites by conjugation of engineered lysine-less ubiquitin. *J Proteome Res* **11**, 796-807, doi:10.1021/pr200668y (2012).

14 Peng, J. *et al.* A proteomics approach to understanding protein ubiquitination. *Nature biotechnology* **21**, 921-926, doi:10.1038/nbt849 (2003).

15 Franco, M., Seyfried, N. T., Brand, A. H., Peng, J. & Mayor, U. A novel strategy to isolate ubiquitin conjugates reveals wide role for ubiquitination during neural development. *Mol Cell Proteomics* **10**, M110.002188, doi:10.1074/mcp.M110.002188 (2011).

16 Lectez, B. *et al.* Ubiquitin profiling in liver using a transgenic mouse with biotinylated ubiquitin. *J Proteome Res* **13**, 3016-3026, doi:10.1021/pr5001913 (2014).

17 Min, M., Mayor, U., Dittmar, G. & Lindon, C. Using in vivo biotinylated ubiquitin to describe a mitotic exit ubiquitome from human cells. *Mol Cell Proteomics* **13**, 2411-2425, doi:10.1074/mcp.M113.033498 (2014).

18 Tagwerker, C. *et al.* A tandem affinity tag for two-step purification under fully denaturing conditions: application in ubiquitin profiling and protein complex identification combined with in vivocross-linking. *Mol Cell Proteomics* **5**, 737-748, doi:10.1074/mcp.M500368-MCP200 (2006).

19 Fang, N. N., Ng, A. H., Measday, V. & Mayor, T. Hul5 HECT ubiquitin ligase plays a major role in the ubiquitylation and turnover of cytosolic misfolded proteins. *Nat Cell Biol* **13**, 1344-1352, doi:10.1038/ncb2343 (2011).

20 Kim, J. Y., Anderson, E. D., Huynh, W., Dey, A. & Ozato, K. Proteomic survey of ubiquitin-linked nuclear proteins in interferon-stimulated macrophages. *Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research* **31**, 619-628, doi:10.1089/jir.2011.0006 (2011).

21 Galligan, J. T. *et al.* Proteomic analysis and identification of cellular interactors of the giant ubiquitin ligase HERC2. *J Proteome Res* **14**, 953-966, doi:10.1021/pr501005v (2015).

22 Kim, T. Y. *et al.* Substrate trapping proteomics reveals targets of the betaTrCP2/FBXW11 ubiquitin ligase. *Molecular and cellular biology* **35**, 167-181, doi:10.1128/MCB.00857-14 (2015).

23 Magliozzi, R. *et al.* Datasets from an interaction proteomics screen for substrates of the SCF(betaTrCP) ubiquitin ligase. *Data Brief* **4**, 229-234, doi:10.1016/j.dib.2015.05.003 (2015).

24 Emanuele, M. J. *et al.* Global identification of modular cullin-RING ligase substrates. *Cell* **147**, 459-474, doi:10.1016/j.cell.2011.09.019 (2011).

25 Ordureau, A. *et al.* Quantitative proteomics reveal a feedforward mechanism for mitochondrial PARKIN translocation and ubiquitin chain synthesis. *Molecular cell* **56**, 360-375, doi:10.1016/j.molcel.2014.09.007 (2014).

26 Ong, S. E. *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **1**, 376-386 (2002).

27 Esser, C., Scheffner, M. & Hohfeld, J. The chaperone-associated ubiquitin ligase CHIP is able to target p53 for proteasomal degradation. *The Journal of biological chemistry* **280**, 27443-27448, doi:10.1074/jbc.M501574200 (2005).

28 Jiang, J. *et al.* CHIP is a U-box-dependent E3 ubiquitin ligase: identification of Hsc70 as a target for ubiquitylation. *The Journal of biological chemistry* **276**, 42938-42944, doi:10.1074/jbc.M101968200 (2001).

29 Soss, S. E., Rose, K. L., Hill, S., Jouan, S. & Chazin, W. J. Biochemical and Proteomic Analysis of Ubiquitination of Hsc70 and Hsp70 by the E3 Ligase CHIP. *PLoS One* **10**, e0128240, doi:10.1371/journal.pone.0128240 (2015).

30 Zhang, H. T. *et al.* The E3 ubiquitin ligase CHIP mediates ubiquitination and proteasomal degradation of PRMT5. *Biochimica et biophysica acta* **1863**, 335-346, doi:10.1016/j.bbamcr.2015.12.001 (2016).

31 Arndt, V. *et al.* Chaperone-assisted selective autophagy is essential for muscle maintenance. *Current biology : CB* **20**, 143-148, doi:10.1016/j.cub.2009.11.022 (2010).

32 Murata, T. & Shimotohno, K. Ubiquitination and proteasome-dependent degradation of human eukaryotic translation initiation factor 4E. *J Biol Chem* **281**, 20788-20800, doi:10.1074/jbc.M600563200 (2006).

33 Choi, Y. N., Jeong, D. H., Lee, J. S. & Yoo, S. J. Regulation of fragile X mental retardation 1 protein by C-terminus of Hsc70-interacting protein depends on its phosphorylation status. *Biochemical and biophysical research communications* **453**, 192-197, doi:10.1016/j.bbrc.2014.09.099 (2014).

34 Meacham, G. C., Patterson, C., Zhang, W., Younger, J. M. & Cyr, D. M. The Hsc70 co-chaperone CHIP targets immature CFTR for proteasomal degradation. *Nature cell biology* **3**, 100-105, doi:10.1038/35050509 (2001).

35 Loffek, S. *et al.* The ubiquitin ligase CHIP/STUB1 targets mutant keratins for degradation. *Human mutation* **31**, 466-476, doi:10.1002/humu.21222 (2010).

36 Okumura, F., Hatakeyama, S., Matsumoto, M., Kamura, T. & Nakayama, K. I. Functional regulation of FEZ1 by the U-box-type ubiquitin ligase E4B contributes to neuritogenesis. *The Journal of biological chemistry* **279**, 53533-53543, doi:10.1074/jbc.M402916200 (2004).

37 Matsumoto, M. *et al.* Molecular clearance of ataxin-3 is regulated by a mammalian E4. *The EMBO journal* **23**, 659-669, doi:10.1038/sj.emboj.7600081 (2004).

38 Wu, H. & Leng, R. P. UBE4B, a ubiquitin chain assembly factor, is required for MDM2-mediated p53 polyubiquitination and degradation. *Cell cycle (Georgetown, Tex.)* **10**, 1912-1915, doi:10.4161/cc.10.12.15882 (2014).

39 Wu, H. *et al.* UBE4B promotes Hdm2-mediated degradation of the tumor suppressor p53. *Nature medicine* **17**, 347-355, doi:10.1038/nm.2283 (2011).

40 Zhang, Y., Lv, Y., Zhang, Y. & Gao, H. Regulation of p53 level by UBE4B in breast cancer. *PLoS One* **9**, e90154, doi:10.1371/journal.pone.0090154 (2014).

41 Zhang, M. *et al.* Chaperoned ubiquitylation--crystal structures of the CHIP U box E3 ubiquitin ligase and a CHIP-Ubc13-Uev1a complex. *Molecular cell* **20**, 525-538, doi:10.1016/j.molcel.2005.09.023 (2005).

42 Benirschke, R. C. *et al.* Molecular basis for the association of human E4B U box ubiquitin ligase with E2-conjugating enzymes UbcH5c and Ubc4. *Structure (London, England : 1993)* **18**, 955-965, doi:10.1016/j.str.2010.04.017 (2010).

43 McDonough, H. & Patterson, C. CHIP: a link between the chaperone and proteasome systems. *Cell Stress & Chaperones* **8**, 303, doi:10.1379/1466-1268(2003)008<0303:calbtc>2.0.co;2 (2003).

44 Murata, S., Chiba, T. & Tanaka, K. CHIP: a quality-control E3 ligase collaborating with molecular chaperones. *The International Journal of Biochemistry & Cell Biology* **35**, 572-578, doi:10.1016/s1357-2725(02)00394-1 (2003).

45 Murata, S., Minami, Y., Minami, M., Chiba, T. & Tanaka, K. CHIP is a chaperone-dependent E3 ligase that ubiquitylates unfolded protein. *EMBO reports* **2**, 1133-1138, doi:10.1093/embo-reports/kve246 (2001).

46 Paul, I. & Ghosh, M. K. The E3 ligase CHIP: insights into its structure and regulation. *Biomed Res Int* **2014**, 918183, doi:10.1155/2014/918183 (2014).

47 Rosser, M. F., Washburn, E., Muchowski, P. J., Patterson, C. & Cyr, D. M. Chaperone functions of the E3 ubiquitin ligase CHIP. *The Journal of biological chemistry* **282**, 22267-22277, doi:10.1074/jbc.M700513200 (2007).

48 Dickey, C. A., Patterson, C., Dickson, D. & Petrucelli, L. Brain CHIP: removing the culprits in neurodegenerative disease. *Trends in molecular medicine* **13**, 32-38, doi:10.1016/j.molmed.2006.11.003 (2007).

49 Rao, S. N., Sharma, J., Maity, R. & Jana, N. R. Co-chaperone CHIP stabilizes aggregate-prone malin, a ubiquitin ligase mutated in Lafora disease. *J Biol Chem* **285**, 1404-1413, doi:10.1074/jbc.M109.006312 (2010).

50 Tsvetkov, P., Adamovich, Y., Elliott, E. & Shaul, Y. E3 ligase STUB1/CHIP regulates NAD(P)H:quinone oxidoreductase 1 (NQO1) accumulation in aged brain, a process impaired in certain Alzheimer disease patients. *J Biol Chem* **286**, 8839-8845, doi:10.1074/jbc.M110.193276 (2011).

51 Yan, W. Q., Wang, J. L. & Tang, B. S. [Functions of carboxyl-terminus of Hsc70 interacting protein and its role in neurodegenerative disease]. *Zhonghua yi xue yi chuan xue za zhi = Zhonghua yixue yichuanxue zazhi = Chinese journal of medical genetics* **29**, 426-430, doi:10.3760/cma.j.issn.1003-9406.2012.04.010 (2012).

52 Cao, Z. *et al.* MiR-1178 promotes the proliferation, G1/S transition, migration and invasion of pancreatic cancer cells by targeting CHIP. *PLoS One* **10**, e0116934, doi:10.1371/journal.pone.0116934 (2015).

53 Jang, K. W. *et al.* Ubiquitin ligase CHIP induces TRAF2 proteasomal degradation and NF-kappaB inactivation to regulate breast cancer cell invasion. *Journal of cellular biochemistry* **112**, 3612-3620, doi:10.1002/jcb.23292 (2011).

54 Min, B. *et al.* CHIP-mediated degradation of transglutaminase 2 negatively regulates tumor growth and angiogenesis in renal cancer. *Oncogene*, doi:10.1038/onc.2015.439 (2015).

55 Sarkar, S., Brautigan, D. L., Parsons, S. J. & Larner, J. M. Androgen receptor degradation by the E3 ligase CHIP modulates mitotic arrest in prostate cancer cells. *Oncogene* **33**, 26-33, doi:10.1038/onc.2012.561 (2014).

56 Tsuchiya, M. *et al.* Ubiquitin ligase CHIP suppresses cancer stem cell properties in a population of breast cancer cells. *Biochemical and biophysical research communications* **452**, 928-932, doi:10.1016/j.bbrc.2014.09.011 (2014).

57 Wang, T. *et al.* CHIP is a novel tumor suppressor in pancreatic cancer through targeting EGFR. *Oncotarget* **5**, 1969-1986, doi:10.18632/oncotarget.1890 (2014).

58 Wang, Y. *et al.* CHIP/Stub1 functions as a tumor suppressor and represses NF-kappaB-mediated signaling in colorectal cancer. *Carcinogenesis* **35**, 983-991, doi:10.1093/carcin/bgt393 (2014).

59 Inuzuka, H. *et al.* Phosphorylation by casein kinase I promotes the turnover of the Mdm2 oncoprotein via the SCF(beta-TRCP) ubiquitin ligase. *Cancer cell* **18**, 147-159, doi:10.1016/j.ccr.2010.06.015 (2010).

60 Itahana, K. *et al.* Targeted inactivation of Mdm2 RING finger E3 ubiquitin ligase activity in the mouse reveals mechanistic insights into p53 regulation. *Cancer cell* **12**, 355-366, doi:10.1016/j.ccr.2007.09.007 (2007).

61 Banka, P. A., Behera, A. P., Sarkar, S. & Datta, A. B. RING E3-Catalyzed E2 Self-Ubiquitination Attenuates the Activity of Ube2E Ubiquitin-Conjugating Enzymes. *Journal of molecular biology* **427**, 2290-2304, doi:10.1016/j.jmb.2015.04.011 (2015).

62 Weissman, A. M., Shabek, N. & Ciechanover, A. The predator becomes the prey: regulating the ubiquitin system by ubiquitylation and degradation. *Nature reviews. Molecular cell biology* **12**, 605-620, doi:10.1038/nrm3173 (2011).

5.6 Supplementary information

5.6.1 Antibodies used in this study

The antibodies used in this chapter are as given in **Table 5-6**. All antibodies were diluted in 5% nonfat milk/TBST.

Antibody Name	Epitope	Company	Catalog number	Approximate dilution factor used (WB/IP)
Goat Anti-Mouse IgG		Thermo	31438	5000
Penta-His		Qiagen	34660	1000
Anti-FLAG	M2	Sigma	F3165	2000
HA	F-7	Santa Cruz	SC-7392	500
CHIP	H-231	Santa Cruz	SC-66830	500
C-MYC	9E-10	Santa Cruz	SC-40	500
V5	C-9	Santa Cruz	SC-271944	500
P53	DO-1	Santa Cruz	SC-126	1000
UB	P4D1	Santa Cruz	SC-8017	200
Ufd2	C-1	Santa Cruz	SC-377072	500
ERK1	K-23	Santa Cruz	SC-94	1000/500
Tubulin	B-5-1-2	Santa Cruz	SC-23948	500
PRMT1	PRMT1-171	Santa Cruz	SC-23948	100/500
OTUB1	J-61	Santa Cruz	SC-130458	500
GST	B-14	Santa Cruz	SC-138	500
β-catenin	E-5	Santa Cruz	SC-7963	500
PP2B-A	H-209	Santa Cruz	SC-9070	1000/500
CARM1	G-2	Santa Cruz	SC-393381	500
PGAM5	K-16	Santa Cruz	SC-161156	500
Goat Anti-Rabbit IgG		Santa Cruz	SC-2004	5000

Table 5-6 List of antibodies used. The dilution factors given are applicable either for Western immunoblotting or immunoprecipitation, depending on the specific assay the antibody was used in; in the case which a particular antibody was used for both types of assay, the dilution factor for Western immunoblotting is given as the first number, while the dilution factor for immunoprecipitation is given as the second number.

5.6.2 Purchased shRNA constructs.

A complete list of shRNA constructs (Dharmacon) tested for the knockdown of E3 enzymes are as given in **Table 5-7**.

Kit used	Catalog number (Dharmacon)
GIPZ UBE4B Lentiviral shRNA Transduction Starter Kit	V2LHS_57030
	V2LHS_57065
	V2LHS_57066
	V2LHS_327169
	V2LHS_327170
	V2LHS_327173
GIPZ STUB1 shRNA Transfection Starter Kit	**V3LHS_304718**

Table 5-7 List of shRNAs tested.
Although six shRNAs were tested for the knockdown of E4B (see Figure 5-7), each displaying similar levels of efficacy, ultimately only one was used. The shRNAs used to construct stable E3 knockdown cell lines are as given in red text.

5.6.3 Full proteomics analysis results.

The enrichment of xUB-conjugated substrates as described in 5.2.2, and subsequent analysis through trypsinization and tandem mass spectrometry was conducted in duplicate for both E4B and CHIP cell lines. The full results of all potential hits for both attempts for each E3 are as attached in the Excel files available online.