Subunit Selective degradation of WIZ, a lenalidomide- and pomalidomide-dependent substrate of E3 ubiquitin ligase $CRL4^{CRBN}$

Thesis by
Helen Yu

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2019
Defended 8/25/2018

Helen Yu
ORCID: 0000-0001-9881-8414

ACKNOWLEDGEMENTS

First, I would have to thank the two people that I was lucky enough to be advised by during my time here. First, I need to thank my advisor, Prof. Raymond J. Deshaies. He tells me painful truths that I probably need to hear, but would strongly prefer not to hear. I learned how to be rigorous and skeptical, and think in a scientifically rigorous manner. But most importantly, for showing me that he cares, by always come through for me on things that had no tangible benefit to him, even during that times I was probably annoying the bejeesus out of him.

I have to thank Prof. Doug Rees, who served as my interim advisor when Ray decided to quit his job. The amount of time he committed to checking in with me, making sure that I was on track, and I was okay was probably far more than he initially agreed to committing. I've enjoyed being able to talk to him about my career, and be able to experience the wonderful nurturing environment that his lab provides.

I would have to thank Dr. Alice Huang for her infinite wisdom and support. And for administering the much-needed back hand when needed. It's cool to be mentored by someone who is kind and also tough as nails.

Of course, I would have to acknowledge Prof. Linda Hsieh-Wilson whose group I joined when I first came here. I was lucky to be able to receive her mentorship, and see how tenacious and persistant she was, and learn how much of science is 99% perspiration and 1% inspiration, as well as her belief in me.

I would have to thank my wonderful committee, Profs. Dave Tirrell, Shu-Ou Shan, Bil Clemmons. They've always been wonderful about making time to meet with me on short notice. They clearly don't have to have additional committee meetings to make sure students are on track, and I'm sure there are other things that are significantly much more exciting for them. However, it does mean a lot to me that they care enough to commit their time and energy outside of what is explicitly required of them to make sure that I'm successful.

I would have to thank Dr. Ron Zuckermann and Prof Carolyn Bertozzi, who mentored me while I was an undergrad. Ron checked in with me to make sure I was doing okay, and he was always just a phone call away in case I needed anything. Carolyn has always been a repository of support and inspiration.

I've been lucky to have been able to work with Dr. Sonja Hess, at the Proteome Exploration Laboratory to do some of the proteomics experiments. Annie Moradian, Mike Swederski, Tanya Yakushi, and Roxana were wonderful about teaching me about proteomics, and different instrumentation. I really am fortunate to have been able to receive their help. Much of the mass spec work was possible because of the wonderful facility.

I would have to thank Dr. Claude Rogers, who helped me a lot with my experiments when I was a younger student, who is a repository of knowledge, incisive feedback, and god-like programming skills.

I would have to thank the Deshaies lab! Rati Verma for sitting with me almost every Friday to go through my data figure by figure, and bond with me about how to deal with people. Dr. Jennifer Mamrosh and David Sherman who have been like my older siblings, checking in with me whenever I look upset, making the time for me when they clearly have other things that they're stressed out about to listen to me vent, and giving me scientific feedback. I would have to of course thank Emily Blythe for being my baymate through all of the years. Of course, so much of my life was easier because I didn't have to order things, because Robert Oania took care of so much of it. I didn't have to make any of my own buffers or my own water, because Heenam took care of it. And I didn't have to cut my own nitrocellulose because Daphne took care of it. I'm lucky to be able to be taken care of by so many support staff.

Of course I would have to thank Dr Thang Nguyen for being generous with his reagents, protocols, working with me to study CRBN. I've enjoyed being able to benefit from the groundwork that he set up.

I've been lucky to work in an environment where everyone has always been open about sharing their equipment and reagents, and have always shown an interest in getting to know me. Prof. Lea Goentoro, Prof. Kai Zinn, and more recently Prof. Rebecca Voorhees have always been generous in terms of sharing equipment and looking after me. All of my experiments were done on Lea's odyssey, and I've had a lot of fun being able to crash her lab's group meetings. It's always been nice to walk down the hall down the Zinn lab whenever I run out of a reagent, or have people ask me about how I'm doing. It makes the workplace seem more like a second home. (and thanks Prof. Pamela Bjorkman for letting me use your tissue culture hood!).

A significant portion of the time, projects end up not working out. But what's

important is that you get to learn from people who are extremely good. A significant portion of my time is saved because there is someone else who knows how to do things, and takes care of instruments. I have to thank Dr. Daniel Braas, at UCLA who I worked with to do the metabolomics experiments. He gave me meticulously annotated protocols, that always resulted in beautiful data as long as I followed each step precisely, and spent so much time working with me to come up with future steps. I have to thank Dr. Jost Vielmitter at the Protein Expression Center for working with me to express CRBN, and who committed his resources to helping me troubleshoot expression glutamine synthetase. Additionally, Igor Antoschenkin and Vijiya Kumar at the Genomics Facility who trained me how to work with RNA , do qPCR and use the bioanalyzer. I would have to thank Shelley Diamond at the Cell Sorting facility, who spent her time helping me make sure my experiment was done properly and safely, and who also coordinated my time when I needed to use the Amaxa nucleofector in the Rothenerg lab.

I would like to thank my wonderful parents who were brave enough to leave everything they knew to move to a foreign country, and put up with misery so I could have a better opportunity in life. So much of the things that are possible for me now only exist for me because of the sacrifices they made.

Finally I would have to thank my Uncle, Dr. Dingwei Yu who unfortunately passed away, who always believed in me and was the other troublemaker in the family. I wish you could have lived to see this.

ABSTRACT

This dissertation is focused on identifying novel targets of immunomodulatory (IMiD) drugs. IMiDs are a class of drugs that are used to treat multiple myeloma. The first chapter is an introduction to the clinical use of IMiDs, as well as the protein cereblon (CRBN), the primary target of IMiDs. The second chapter describes work towards the identification of a novel IMiD target, WIZ, that is regulated by CRBN in an IMiD dependent manner. Mass spectrometry was performed to identify novel binding partners, and IMiD dependent regulation by CRBN was validated using chemical and genetic methods. Understanding how these drugs work will inform the production of more potent and more selective drugs.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Chapter 1

INTRODUCTION

1.1 Protein Degradation as a Strategy to Treat Cancer

The proteins inside of the cell are like the metaphorical ship of Theseus: old, rotting components are constantly being replaced by newer functional ones (Fig. 1.1) (Bachmair and Varshavsky, 1989). The genetic code acts as the blueprint for how the ship is to be constructed: what components are to be used, and in what proportions they are to be combined. Just as a ship constructed from a malfunctioning part, or the incorrect number of parts, would not function properly, a cell produced with incorrect proteins or improper number of proteins can result in disease (Reinstein and Ciechanover, 2006). Therefore, the goal of any therapeutic is to target a component that is malfunctioning or produced in the wrong quantities.

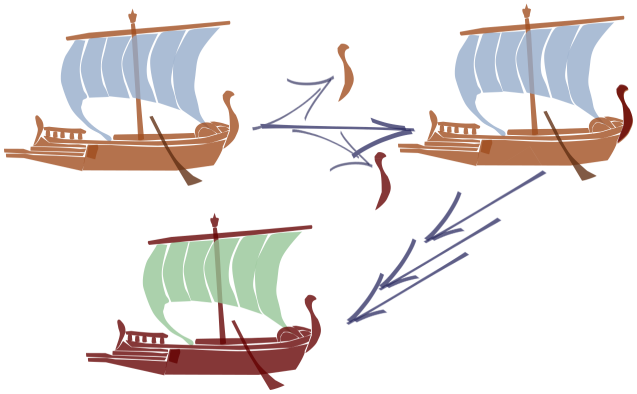

Figure 1.1: The proteome is like the metaphorical ship of Theseus. Old components are constantly being replaced by newer ones. Over time, the cell manages to replace all of its components with newer ones.

One therapeutic strategy is to directly modify the activity of a malfunctioning component. Small molecules can be used to achieve this, which work by binding to the target malformed protein and either activate or block the wildtype enzymatic activity. An example of how a small molecule inhibitor would accomplish this is the usage of all-trans retinoic acid in the treatment of acute promyelocytic leukemia (APL) (M. Huang et al., 1988). APL occurs when the retinoic acid receptor (RARA) undergoes a translocation with the promyelocytic leukemia gene on chromosome 15 to generate a chimeric protein. Normally, retinoic acid receptor is activated by

retinoic acid to transcribe genes promoting cellular differentiation. However, the retinoic acid receptor chimera is not responsive to retinoic acid and exerts a dominant negative effect, causing APL. In healthy individuals, white blood cells help the body fight off infection and depend on signaling through the retinoic acid receptor to mature correctly. Because the white blood cells in APL patients can no longer mature correctly, a large number of immature precursors end up accumulating. Therefore, a therapeutic regimen would rescue this deficit by restoring proper signaling through the retinoic acid receptor. Since the chimeric retinoic acid receptor in APL is responsive to all-trans retinoic acid, all-trans retinoic acid can be used to convert an improperly made protein into a functional one (M. Huang et al., 1988). All-trans retinoic acid is effective towards the treatment of acute promyelocytic leukemia primarily because RARA has a well defined ligand binding site (M. Huang et al., 1988). Although similar strategies are often effective for targeting enzymes, they are not effective for diseases caused by defects in other types of proteins, or when the protein is not produced Cromm and Crews, 2017.

An example of such a disease not treatable by modulating protein function is Duchenne Muscular Dystrophy, which is caused by a loss of Dystrophin function. Dystrophin scaffolds the cytoskeleton in muscle fibers, likely making it inaccessible to small molecule inhibitors (Fairclough, Wood, and Davies, 2013). Additionally, since Dystrophin in Duchenne Muscular Dystrophy is disrupted by a premature stop codon, no protein is produced, eliminating the rationale to design a small molecule modulator to treat the condition (Aartsma-Rus et al., 2009). The result of untreated Duchenne Muscular Dystrophy is muscle weakness and loss, often manifesting as death at a young age from cardiac related problems (Eagle et al., 2002). Diseases like Duchenne Muscular Dystrophy highlight the need for alternative therapeutic strategies not dependent on protein structure.

One method that has gained prominence in recent years is a genome editing method called CRISPR, which stand for clustered regularly interspaced palindromic repeats (Cong et al., 2013). CRISPR is used by prokaryotic organisms to protect them from attack from viral intruders, but has been repurposed for genome editing in mammalian cells. The system is comprised of an enzyme (Cas 9), which uses a cognate RNA to target it to specific sequences of DNA (Mali, Esvelt, and Church, 2013) to cause double strand breaks. These breaks can then be repaired with an exogenously supplied template via homology directed repair. The exogenously supplied template can be designed to repair the cleaved DNA sequence, providing a

mechanism for genomic editing. Since CRISPR modifies DNA, it is not restricted to enzymes, and therefore can be used more broadly for therapeutics. However, a limitation of genome-editing methods is it that the length of time it would take to reach a new equilibrium with the repaired protein depends on the half-life of the underlying protein. Additionally, the nuclease requires a proteospacer adjacent motif (PAM) to be adjacent to the target site in order to bind, which poses a restriction on the type of proteins that can be repaired based upon their primary sequence (Yilan Zhang et al., 2014). Finally, actively dividing cells that are easier to transfect are inherently going to be more amenable to be targeted by this method (Richner and Krook-Magnuson, 2018). As a result, this method would not be as effective for difficult to transfect cell types like neurons. Therefore, this illustrates the necessity of complementary therapeutic strategies.

Additionally, diseases that are caused by producing too many copies of a protein can be targeted by either adjusting the rate of production or the rate of degradation. Adjusting the rate of production would have similar limitations as genome editing based strategies, such as low effectiveness in treating conditions with protein targets having long half-lives. Additional technical limitations exist, particularly in delivery of these therapeutics. An example that highlights this limitation is Bevasiranib, which was developed for the treatment of age-related Macular Degeneration (AMD) (Garba and Mousa, 2010). As people age, fatty deposits gradually accumulate underneath the retina, which can cause blindness. This is exacerbated by release of cytokines like vascular endodermal growth factor (VEGF). Bevasirib is a siRNA therapeutic. siRNA targets the rate of protein production by catalyzing the decay of mRNA transcripts before they can be translated into protein. Although Bevasirib was extremely efficient in removing the mRNA transcripts of VEGF in cell culture and animal models, the drug was ineffective towards ameliorating the symptoms of the disease in humans (Reinstein and Ciechanover, 2006; Garba and Mousa, 2010; Shen et al., 2006). Because oligonucleotides are unstable in plasma, macular degeneration is one of the few diseases that can theoretically be targeted using siRNA based technology because the eye and skin are some of the few tissues that can directly uptake the oligonucleotide sequences. Therefore, siRNA based technologies are extremely restricted in terms of the type of diseases that they can target, and how effective they are depends on how long it takes to eliminate the existing copies of the protein. Therefore, it illustrates the practicality of a therapeutic strategy that works by catalyzing the rate of protein decay.

Catalyzing the rate of decay is a versatile therapeutic strategy without the limitations of many of the other tools. Unlike small molecule inhibitors which are primarily limited to targeting enzymes, drugs that catalyze the rate of decay would be applicable to a larger variety of targets. Since the time to a new equilibrium depends on the rate of decay, degradation-based methods show promise to work more rapidly than genome editing or translation based methods (Lai and Crews, 2017). Additionally, since the cell would have to restore protein levels to reverse the effects of treatment instead of waiting for a small molecule to dissociate, it would enhance the duration of the effect. Just as a ship that was constructed from the broken components, or components that were combined in the wrong proportions, cells undergo disease states like cancer when too much of a protein, or a bad version of a protein is produced. Therefore, the goal of a therapeutic strategy would be to identify and remove a malfunctioning or overproduced component. However, in order for this strategy to result in a viable therapeutic necessitates understanding the way proteins are removed from the body, as well as identifying a degradation method that is as specific and selective as possible.

1.2 Ubiquitin Proteasome System as the Optimal Mechanism to Repurpose for Drug Design

Just as the way a ship constructed from a malfunctioning component, or components combined in the wrong proportions, would not be functional, constructing a cell with a malfunctioning protein or the wrong number of proteins can result in a disease state. Therefore, one obvious therapeutic strategy would be to remove the component by redirecting the existing cellular machinery to modulate the rate of degradation. Since there are several ways that a protein can be removed from the body, the objective is to identify the method that would enable the most precise removal of a protein of interest. Autophagy is a method of removing proteins, where the cell cannibalizes itself to liberate biomolecules (Kaur and Debnath, 2015). Unsurprisingly, autophagy is the most efficient way to convert bulk amounts of organelles or proteins into individual building blocks (amino acids). Since autophagy is normally triggered under conditions of extreme stress or starvation, it would be difficult to trigger autophagy without cannibalizing the entire cell. During autophagy, target proteins are quarantined from the rest of the proteome by being contained in a double membrane bound compartment, called the autophagosome (Fig. 1.2a) (Xie and Klionsky, 2007). The autophagosome can undergo subsequent fusion with the lysosome, a membrane bound compartment full of hydrolytic enzymes that rapidly

digests organelles or protein complexes down into simple biomolecules (Fig. 1.2b and c).

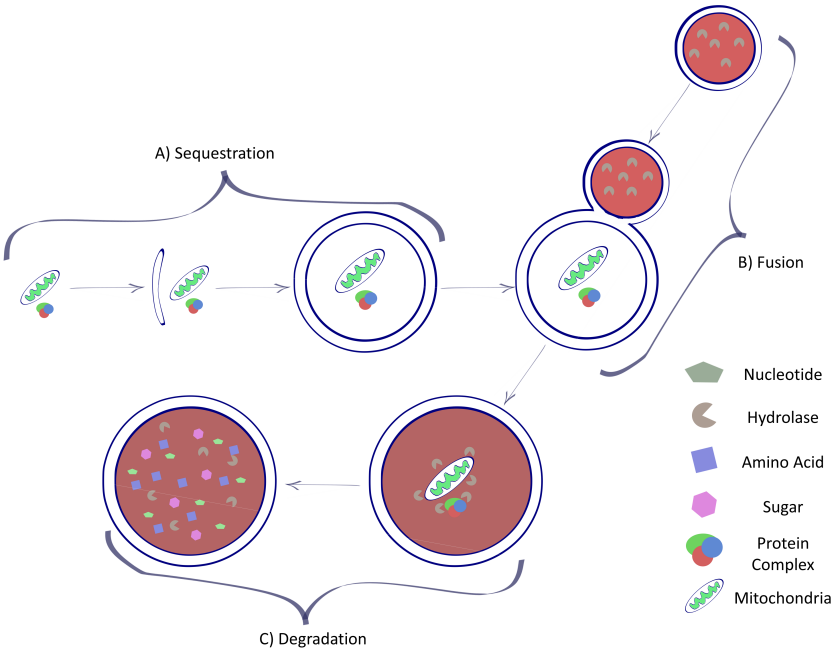

Figure 1.2: Degradation through autophagy. A) Sequestration: Organelles and complexes are engulfed within a lipid bilayer. B) Fusion: The organelles and complexes sequestered in a lipid bilayer undergoes subsequent fusion with the lysosome, a low pH compartment which contains hydrolases. C) Degradation: This fusion event results in the degradation of materials by hydrolases inside of the autophagosome.

However, the way that proteins are isolated and eliminated during autophagy make it a suboptimal pathway to redirect for therapeutic purposes. First, there is no known way to selectively target certain components for degradation, as it is physically impossible to selectively enclose one subunit of a protein complex or one part of an organelle. Similarly, since the engulfed material gets exposed to hydrolytic enzymes after fusion with the lysosome, there is no way to salvage functional protein complex members or organelle components. So although autophagy is useful under starvation conditions where the objective is the convert the existing biomass back into simple building blocks as efficiently as possible, it is too imprecise and difficult to control to be useful therapeutically. What is needed is a method of degradation that selectively exposes a single protein, part of a protein complex, or part of an organelle to a protease.

Unlike autophagy, where the degradative workhorse is the lysosome, the proteasome

is multisubunit enzyme that can degrade proteins one by one. The proteasome is comprised of two subparticles: a 20S core, and two 19S regulatory subunits, which combine to produce the 26S holoenzyme (Chu-Ping et al., 1994). All of the proteolytic activity is contained within 7 subunits which make up the two inner rings of the 20S core (Rubin et al., 1998; J. Lowe et al., 1995). These subunits can digest proteins by cleaving after basic residues like arginine and lysine (trypsin-like activity), aromatic/hydrophobic amino acids (chymotrypsin-like), and acidic residue (caspase-like activity), resulting in peptides that are 4-20 amino acids long (Goldberg, 2012; Arendt and Hochstrasser, 1997; Wenzel et al., 1994). In order for something to be digested, it must be inserted into the 20S core of the proteasome, which is normally blocked off by the N terminal regions of components of the heptameric α subunit ring (Groll et al., 2000). The protease active sites are normally sealed within the barrel of the proteasome, which would enable the ability to selectively degrade a subunit of a multi-protein complex (Tanaka, 2009; Seemuller et al., 1995). However, substrates cannot be degraded without a mechanism to open up the proteasome, to allow substrates to interact with the protease active sites.

The 19S particle can gate the 20S pore. The subunits that comprise the hexameric ring on the bottom of the 19S particle have N terminal tails with a HbYX motif that can insert into pockets between each monomeric unit of the 20S subunit to allow it to be opened up (widening the pore from 9 Å to 20 Å) (Rabl et al., 2008). This allows for the entry of substrates to be degraded. However, since the pore within the proteasome is 13 Å, which is more narrow than the width of most folded proteins, meaning that the channel would get clogged if a protein was fed in directly (Lucas and Ciulli, 2017; Unno et al., 2002; J. Lowe et al., 1995). Therefore, substrates must be unfolded before being fed into the 20S subunit (Smith et al., 2005; J. Lowe et al., 1995). The 19S subunit has a heterohexameric ring that is comprised of AAA ATPases that pull apart proteins before they are fed into the 20S subunit (Benaroudj and Goldberg, 2000). Nonspecific proteolysis by the proteasome could have disastrous results. Therefore, what is needed is a mechanism to selectively target the proteolytic machinery to specific proteins.

The proteasome uses ubiquitin to identify which proteins need to be degraded (Pickart and Eddins, 2004; Hershko and Heller, 1985). Ubiquitin is an 8 kDa protein that is ubiquitously expressed in most eukaryotic organisms Goldstein et al., 1975. The C terminal glycine on ubiquitin is conjugated to the amine on lysines of the target proteins via an isopeptide linkage. The proteasome has two 2 receptors, 26S

Proteasome Regulatory Subunit RPN13 (RPN13) and 26S Proteasome Regulatory Subunit RPN10 (RPN10) on the 19S particle, that enable it to selectively identify and enrich for proteins that have been targeted for degradation (Fig. 1.3) (Elsasser et al., 2004; Verma, Oania, et al., 2004). RPN10 only binds to an ubiquitin if it is polymerized into a chain with at least 4 monomeric units (Thrower et al., 2000; Wickliffe et al., 2011). Alternatively, polyubiquitinated proteins can be delivered to the proteasome via shuttle proteins such as UV Excision Repair Protein Rad23 Homolog A and B (RAD23A and RAD23B) and Ubiquilin-2 (UBQLN2) (Chen and Madura, 2002; E. D. Lowe et al., 2006). A protein to be degraded must have a polymer of ubiquitin with at least 4 monomeric units appended to a lysine residue (Chau et al., 1989). However, since each ubiquitin itself has 7 lysine residues, there are a minimum of 343 possible sub structures that can be formed. RPN10 is known to recognize one of the substructures where ubiquitin is polymerized through lysine 48 (Finley et al., 1994; Chau et al., 1989).

Although appending a protein with an ubiquitin polymer creates an affinity tag the proteasome can use to identify and enrich for proteins that are marked for degradation, it also could clog the 20S degradative machinery. Therefore, a metallo enzyme (RPN11) is needed to shear ubiquitin off of the unfolded protein before it gets fed into the 20S subunit (Verma, Aravind, et al., 2002; Yao and Cohen, 2002).

Although the proteasome is fairly effective at removing proteins that are cytosolic and soluble, it has problems digesting proteins in other compartments or embedded in the membrane of an organelle. Therefore, an additional factor, p97/VCP, is needed to extract the protein so it is accessible to the proteasome (Meyer, Bug, and Bremer, 2012). p97 has an ATPase domain to provide the energy needed to perform the extraction (Buchberger et al., 2001). p97 also has a number of adapters that can divide it into different subcomplexes to further increase the specificity of the interaction. The function of the proteasome and p97 together enables the removal of many proteins marked for degradation.

However, both p97 and the proteasome rely upon cellular machinery that only ubiquitinates proteins that need to be removed. Repurposing both for therapeutic purposes necessitates being able to selectively ubiquitinate a therapeutic target. However, it would take an extremely long time for the reaction to occur spontaneously because the terminal side chain amine on the lysine would need to react with the carboxylic acid on the C terminus of the ubiquitin to generate an amide bond. Carboxylic acids make poor leaving groups, meaning that it would be unlikely for

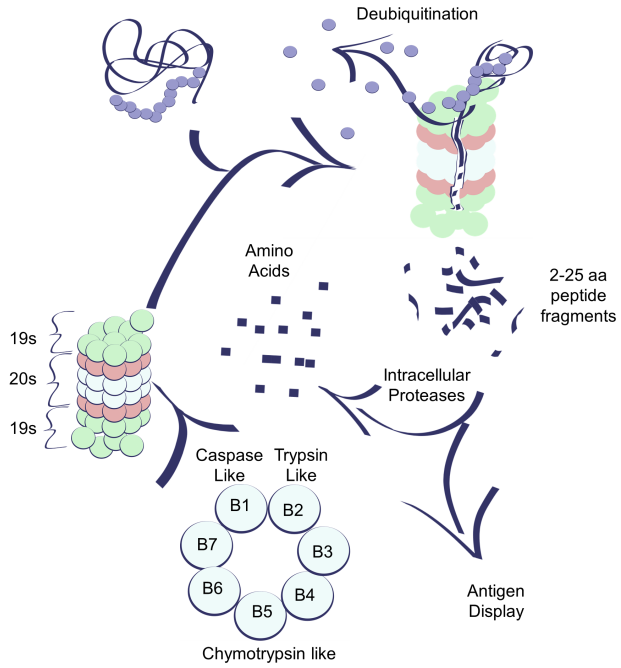

Figure 1.3: Ubiquitinated proteins are recruited by receptors RPN10 or RPN13 on the proteasome. The ubiquitin chain is removed *en bloc* by RPN11, a metalloprotease. The protein is unfolded by ATPases and digested into smaller peptides within the proteasome, by the 20S core peptidases.

the uncatalyzed reaction to occur on a time scale that would be compatible with life (Stevenson and Williamson, 1958; Hodgman, 1951). Therefore, an enzyme (ubiquitin-activating enzyme; E1) is needed to convert the carboxylic acid into a thioester, which is a better leaving group, for the reaction to occur on an appropriate timescale (Fig. 1.4a). The lysines on the proteins that need to be removed are chemically indistinguishable from the lysines on the proteins that are to be retained, meaning that a mechanism is needed to target chemically reactive ubiquitin to the protein of interest (Fig. 1.4c and d).

A large class of enzymes, E3 ubiquitin ligases, are responsible for targeting activated ubiquitin to specific proteins (R. Deshaies, 1999). Hundreds of E3 ubiquitin ligase enzymes tune the specificity of protein degradation by directing the activated ubiquitin thioester to different subsets of proteins. However, catalysis can only occur if the activated ubiquitin is oriented towards the target lysine. Therefore, an enzyme, ubiquitin-conjugating enzyme/E2, positions the activated ubiquitin to facilitate nucleophilic attack by the lysines on the protein of interest (Fig. 1.4b).

Degrading a protein through the ubiquitin proteasome system instead of through

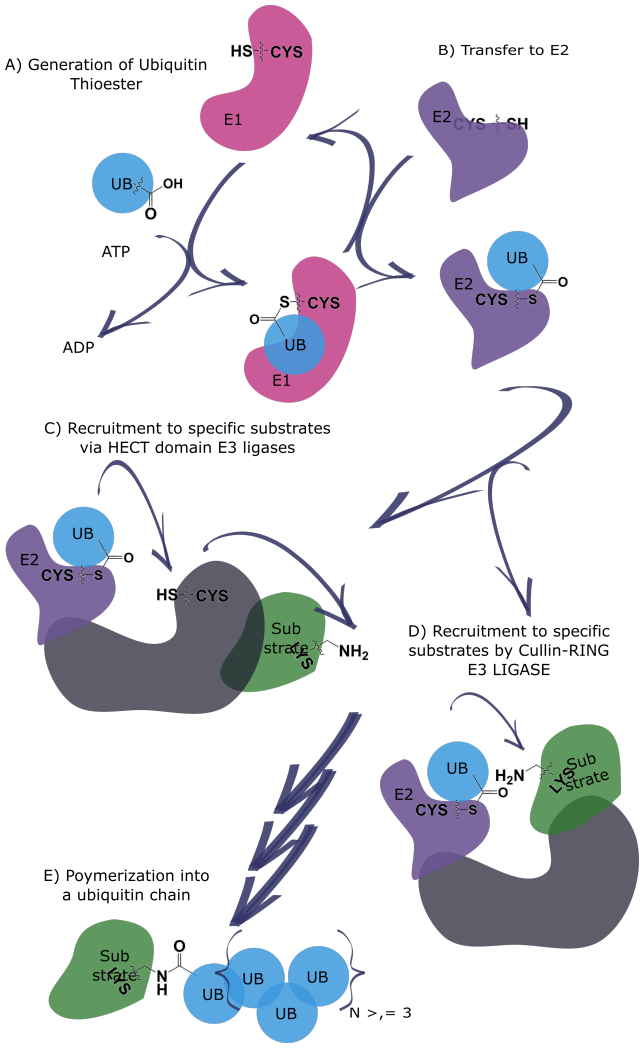

Figure 1.4: A) Generation of Ubiquitin Thioester by E1: The C terminal carboxylic acid is converted into a more reactive thioester by an enzyme called E1. B) Transfer to E2: For catalysis to occur, the ubiquitin thioester needs to be transferred to a different enzyme, E2, to orient it correctly. C) Recruitment to HECT domain E3 ubiquitin Ligase: HECT domain E3 ubiquitin ligases target the activated ubiquitin to the protein of interest. HECT domain ligases work by transferring the ubiquitin directly onto the E3 before transferring it to the substrate. D)Recruitment to cullin-RING Domain E3 ubiquitin ligase: cullin-RING Domain E3 ubiquitin ligases recruit the activated ubiquitin to the substrate. They work by positioning the E2 to transfer the activated ubiquitin directly from the E2 to the substrate. E) Polymerization of ubiquitin: Since the proteasome can only recognize polymers of ubiquitin with at least 4 monomeric units, the ligase must create a chain long enough for the proteasome to grab onto.

autophagy provides the selectivity necessary to degrade one subunit of a complex or one protein in an organelle. Since the entire protein or complex is bathed in hydrolytic enzymes during autophagy, it makes it impossible to selectively remove a protein that was one component of a complex, or one component of an organelle. Therefore, a more selective method like the ubiquitin proteasome system is more suited towards redirecting for therapeutic purposes. Modification of the target protein with an ubiquitin polymer creates an affinity enrichment handle that the proteasome can use to identify which proteins need to be degraded. Since E3 ubiquitin ligases are the enzymes that direct activated ubiquitin to specific proteins, identifying the optimal E3 ubiquitin ligase to target is necessary for redirecting the proteasome to catalyze the degradation of a therapeutic target.

1.3 Substrate Binding and Catalysis are Decoupled in cullin-RING Type E3 Ubiquitin Ligases

E3 ubiquitin ligases tag substrate proteins with ubiquitin, which serves as an affinity tag for subsequent enrichment and degradation by the proteasome. Redirecting the proteasome to digest a therapeutic target necessitates being able to modify the substrate specificity of an E3 to ubiquitinate a therapeutic target. However, some E3 ubiquitin ligases are more ideal than others to being redirected towards a therapeutic target. Since the catalytic ubiquitin is directly transferred from the E2 to the E3, in Homologous to the E6-AP Carboxyl Terminus (HECT) domain type E3 ligases, it would be more difficult to decouple substrate binding and catalysis (L. Huang et al., 1999; R. Deshaies, 1999). In contrast, cullin-RING domain type E3 ubiquitin ligases can have their substrate specificity redirected without antagonizing catalysis because the catalytic ubiquitin is transferred directly from the E2 to the substrate (Zheng, Schulman, et al., 2002). Therefore, it is more straightforward to modify substrate binding without antagonizing catalysis.

The specificity of a cullin-RING type E3 ubiquitin ligase is determined by the substrate receptor, an interchangeable subunit which recruits proteins to the enzymatic core (Skowyra et al., 1997). Once the target is brought to the enzymatic core, it is held in proximity to the activated ubiquitin (supplied by the Ub-E2). Both the substrate receptor and the E2 are fixed to a rigid scaffold (Cullin) to hold them in place, as well as adapters to facilitate the interaction with the scaffold (Ring Box protein 1; RBX1 for the E2 and DDB1 or SKP1 are examples of ones for the substrate receptor) (Fig. 1.5a) (Petroski and R. J. Deshaies, 2005a; Feldman et al., 1997). Recruiting the target protein to the enzymatic core effectively increases the

local concentration of ubiquitin from 10-20 μM to 3 mM. However, the complex is enzymatically inactive because the activated ubiquitin is at least 50 Å from the target lysine (Fig. 1.5a) (Zheng, Schulman, et al., 2002; Zheng, P. Wang, et al., 2000; Haas, 1988). Since the ubiquitin polymer is fairly flexible enough, the only thing that is needed is a way to traverse the 40 Å gap for the complex to be catalytically competent (Petroski, Kleiger, and R. J. Deshaies, 2006).

Modification of the cullin scaffold with the small ubiquitin-like protein Nedd-8 dislodges RBX1, the adapter for the E2, from the scaffold in a way that allows the activated ubiquitin to spring across the 50 Å gap (Fig. 1.5b). Once RBX1 is dislodged from the cullin scaffold, the activated ubiquitin would be able to traverse the 50 Å gap, and the complex has enough flexibility to polymerize ubiquitin through the distal end (Duda et al., 2008). However, each E2 only brings in 1 equivalents of activated ubiquitin (Fig. 1.5e). Therefore, a mechanism is needed to regenerate a complex with a charged E2 so ubiquitin can be polymerized into a chain long enough for the proteasome to grab onto. For the E3 ligase complex to be regenerated on a time scale that is compatible with life, the depleted E2 must dissociate rapidly enough so that the E3 ligase can reassemble with an ubiquitin-charged E2.

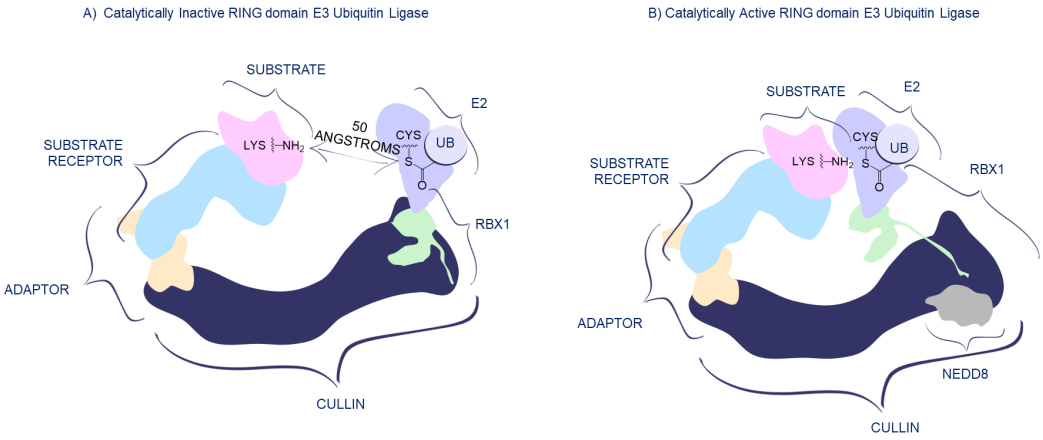

Figure 1.5: A) Catalytically inactive cullin-RING Ligase: Cullin-RING ligases are made up of a substrate receptor that brings in the target protein, the E2 which brings in the activated ubiquitin, and a rigid scaffold (cullin) which holds the two together. Both are anchored to the scaffold via adapters. The complex is normally enzymatically inactive because there is 50 Å between the target protein and the activated ubiquitin. B) Catalytically active cullin-RING ubiquitin ligase: Modification of the cullin scaffold with Nedd-8 dislodges RBX1, the adapter for the E2, to bridge the 50 Å gap.

Previous work on CDC34, the yeast ortholog of the E2 of cullin-RING type E3

ubiquitin ligases, showed that although the E2 binds tightly to the E3, the complex forms and falls apart quickly enough to enable rapid regeneration of the E3 ubiquitin ligase complex with activated ubiquitin, enabling chain elongation to occur on a time scale amenable to responding to rapid environmental changes (Kleiger et al., 2009). However, just because ubiquitin can be polymerized on the appropriate time scale does not mean it would be constructed with the linkage that enables recognition by the proteasome, as each ubiquitin has seven lysines. Since proteasome can recognize substructures where ubiquitin is polymerized through lysine 48 or lysine 63, a mechanism is needed to ensure that ubiquitin is polymerized into one of those substructures (Saeki et al., 2009; Jacobson et al., 2009; Petroski and R. J. Deshaies, 2005b).

Because the ubiquitin is transferred directly from the E2 to the target substrate in cullin-RING E3 ligases, the E2 selects one out of the seven lysines on the ubiquitin to polymerize (Petroski and R. J. Deshaies, 2005b). The E2 sterically restricts access to ubiquitin so that lysine 48 has the most access. However, even if the target substrate is recruited to the substrate receptor, it would only be productive if the cognate substrate receptor is assembled with the enzymatic core. This presents a problem as multiple substrate receptors can combine with the enzymatic core. Most substrate receptors are not assembled with the enzymatic core (Reitsma et al., 2017). Without a mechanism to adjust the relative stoichiometry of substrate receptors assembled with the enzymatic core, there would be no way to adjust to a sudden fluctuation in the amount of substrate. One solution is to recycle the enzymatic core from complexes that are enzymatically inactive, or not bound to substrate. However, it takes half a week for the substrate receptor to dissociate from the enzymatic core, which is far too slow to readjust the relative stoichiometry of different substrate receptors assembled with the enzymatic core on a time scale compatible with life (Pierce et al., 2013). Therefore, a mechanism is needed to accelerate the rate of dissociation of the substrate receptor so the enzymatic core can be recycled from inactive complexes.

Cullin-associated and neddylation-dissociated protein 1 (Cand1) accelerates the disassembly of complexes that are enzymatically inactive, or where the enzymatic core does not have the Nedd-8 modification (Pierce et al., 2013). As a result, the enzymatic core can rapidly reassemble with a different set of substrate receptors to accommodate a sudden shift in relative occupancy of substrate receptors. However, because none of the core enzymes recycled by Cand1 are modified with Nedd-8,

the complexes that would reform would be enzymatically inactive and vulnerable to disassembly by Cand1 before the substrate can be ubiquitinated. Previous work demonstrated that the complex is neddylated as the new complex is formed (X. Liu et al., 2018). CSN, the enzyme complex that cleaves the Nedd-8 modification, only works on proteins that are not in active complex with substrate (Fig. 1.6) (Mosadeghi et al., 2016). As a result,the relative stoichiometry of substrate receptors assembled with the enzymatic core can rapidly adjust to dynamic changes in relative occupancy of different substrate receptors.

1.4 CRBN as an Ideal Substrate Receptor to Design a Proteolysis Targeting Chimera

Cullin-RING E3 ligases are comprised of an enzymatic core made up of a rigid scaffold bound to an E2 that brandishes the chemically activated ubiquitin, and an interchangeable substrate receptor to direct the enzyme to subsets of proteins (R. Deshaies, 1999). Since the substrate binding and catalytic activity are uncoupled, it makes it the ideal type of E3 ubiquitin ligase to redirect towards degrading a therapeutic target. Redirecting a cullin-RING E3 ligase to ubiquitinate a therapeutic target can be accomplished by scaffolding the interaction between the substrate receptor and a novel protein with a small molecule. However, substrate receptors tend to have exposed extended surface that are well suited to mediating protein-protein interactions, but nightmarish from the perspective of small molecule design (Lucas and Ciulli, 2017).

The preferred motif is a small well defined hydrophobic pocket that would place therapeutic target in proximity to the charged E2. Therefore, it is unsurprising that all of the substrate receptors that have been successfully redirected to target different proteins have a small hydrophobic pocket that mediates the recognition of an amino acid sequence, or degron with a modified amino acid. An example of such a substrate receptor is Von Hippel Lindau (VHL), which normally binds to a hydroxyl proline in a process that regulates HIF-1α, a transcription factor that is responsive to molecular oxygen (Min et al., 2002). HIF-1α is hydroxylated in the presence of oxygen, which enables VHL to bind and ubiquitinate it (L. E. Huang and Bunn, 2003). VHL can distinguish between the hydroxylated form because of 2 hydrogen bonding interactions provided by a buried histidine that provide 1000-fold selectivity (Min et al., 2002; Hon et al., 2002). The hydroxylated peptide binds through an extended conformation on the surface of the protein, with the hydroxylated proline wedged into a gap. Therefore, it would be the type of binding that would be difficult

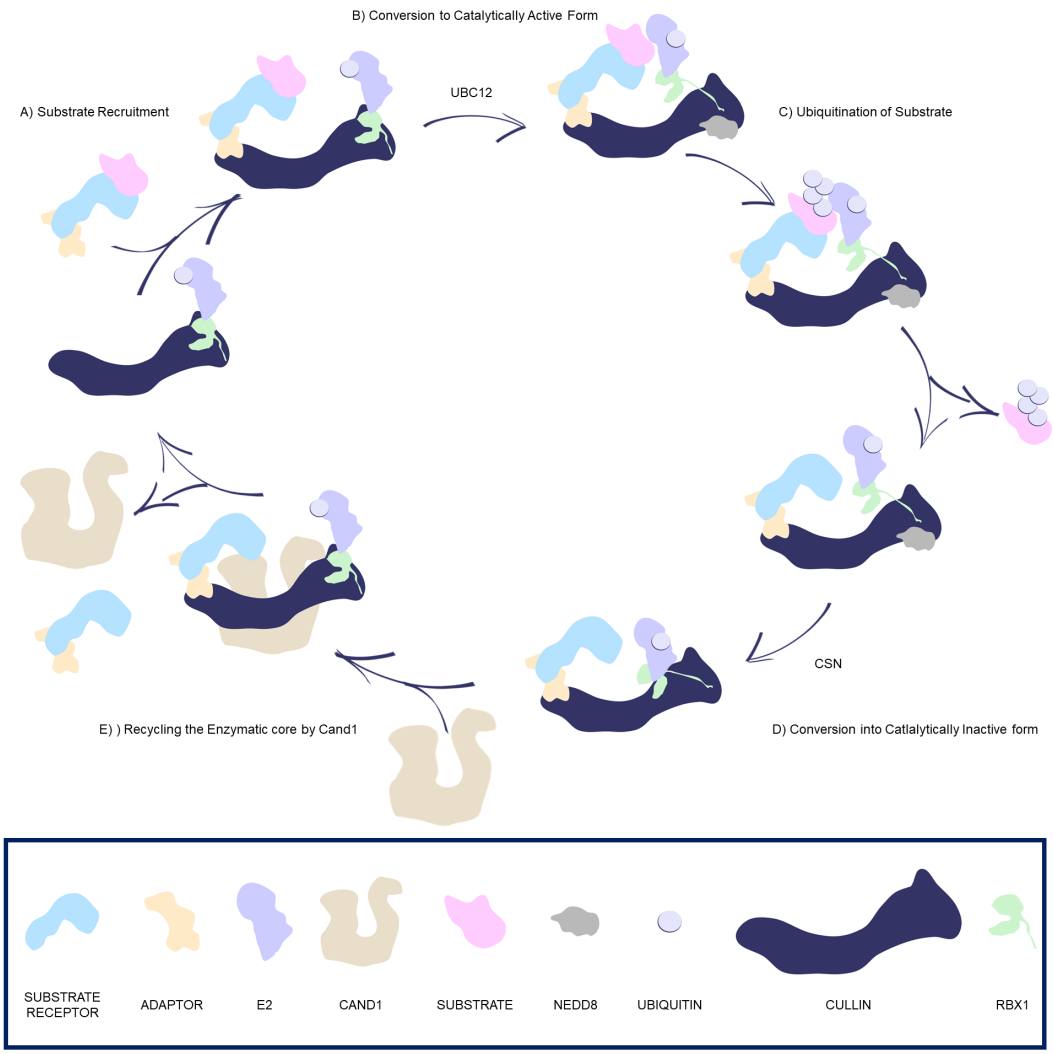

Figure 1.6: The cullin-RING Ligase cycle. A) Substrate Recruitment: Substrate receptors recruit the substrate to the enzymatic core. B) Conversion to Catalytically Active form: UBC12 neddylates the cullin scaffold, converting it into an enzymatically active form. C) Ubiquitination of Substrate: E3 ubiquitin ligase complex ubiquitinates substrate, earmarking it for degradation. D) Conversion into catalytically inactive form by CSN: CSN deneddylates the cullin scaffold, making the complex enzymatically inactive. C) Recycling the enzymatic core by Cand1: Cand1 binds to the enzymatically inactive core, and catalyzes the rate of dissociation of the substrate receptor, allowing the enzymatic core to combine with a new substrate receptor.

to replicate with a small molecule because it's mostly exposed, although the kinetics would enable the fast assembly and disassembly of the E3-substrate complex.

Similarly, cereblon (CRBN) is also a substrate receptor that contains a small well defined binding pocket. CRBN is one of the substrate receptors associated with an E3 ligase with an enzymatic core that is comprised of DDB1 (the substrate receptor adapter) and Cullin4a. Substrate receptors that assemble with the enzymatic core are referred to as DDB1 Cul4 Associated Factors (DCAFs) (Petzold, Fischer, and Thomä, 2016). Most DCAFs have a WD40 domain to facilitate protein-protein interactions that form a large extended surfaces. Instead, CRBN is comprised of a seven α helical bundle, an N terminal region with homology to a bacterial Lon protease, and a C terminus with a similar fold to pseudouridine synthetase or bacterial methionine sulfoxide reductase (Chamberlain et al., 2014; Petzold, Fischer, and Thomä, 2016). CRBN is anchored to DDB1, the adapter for Cul4a by the seven α helical bundle, which inserts between the two propellers that are produced by the WD40 domain (BPA, BPC) (Fig. 1.7).

The pseuduridine synthase/bacterial methionine sulfoxide reductase like domain in CRBN faces where the charged ubiquitin-E2 is, and is believed to be the substrate binding domain. Like VHL, CRBN has a well-defined hydrophobic pocket formed by three different tryptophans that is directly facing where the ubiquitinated E2 would be. Additionally, crystallographic studies indicated that there is a zone where a potential substrate could be ubiquitinated because the scaffold could rotate 150°. The hydrophobic binding domain has structural similarity to a domain known to recognize acetylated and methylated lysines.

The domain was shown to mediate the recognition of a substrate with an acetyl degron, and was also shown to bind to uridine (Fig. 1.8) (Hartmann et al., 2014; Van Nguyen et al., 2016). Studying a drug that is known to work through CRBN could serve as a model system to engineering newer, more potent, and more selective drugs.

1.5 Immunomodulatory Drugs as a Model System to Inform the Design of Proteolysis Targeting Chimeras

CRBN is a substrate receptor that has previously shown a lot of promise for redirecting to catalyze the degradation of a therapeutic target, because it has a well defined hydrophobic binding pocket that enables precise placement of a therapeutic target adjacent to the ubiquitinated E2. Therefore, understanding how a drug that is

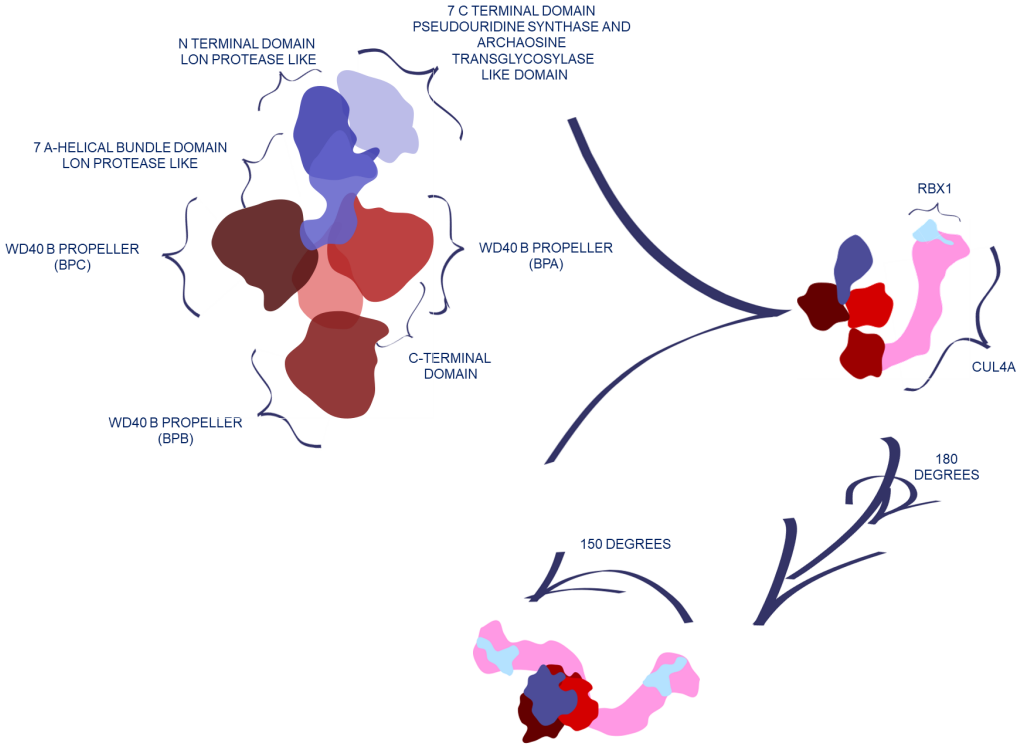

Figure 1.7: CRBN is comprised of an N terminal Lon protease like domain, a C terminal pseudouridine like domain, and a seven α helical domain. The seven α helical domain forms the interface with DDB1, the adapter to the Cullin, between the two WD40 domains.

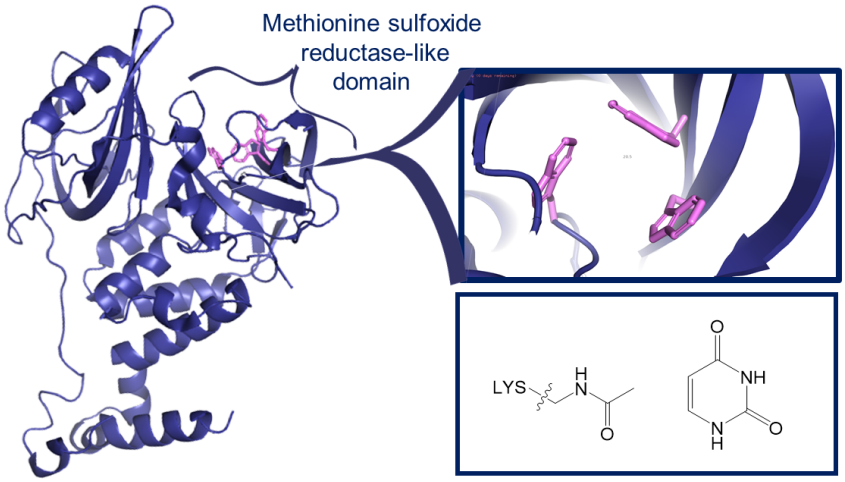

Figure 1.8: CRBN has a binding pocket that is flanked by three tryptophans. This binding pocket is in the methionine sulfoxide reductase-like domain, and is normally used to recognize acetyl lysine and uridine.

known to work through redirecting what proteins CRBN recognizes would act as a model system to facilitate the understanding of how drugs in this entire class might work. Thalidomide was previously demonstrated to be a drug that was effective towards redirecting CRBN to target novel proteins for degradation (Lu et al., 2014). Thalidomide was originally prescribed as a sedative to treat morning sickness, and was rapidly cleared for usage in humans because of its low toxicity in mice. However, children born to the women who took the drug had horrific birth defects, as their limbs did not form properly. As a result of the public outcry, the drug was abruptly withdrawn from the market.

Years later, thalidomide was approved for limited usage because it was found to be effective towards treating a specific side effect of leprosy, erythema nodosum leprosum. Erythema nodosum leprosum is a complication from leprosy that is caused by massive inflammation from overproduction of a specific cytokine, the first evidence that indicated that thalidomide would be effective towards regulating the immune response (Sheskin, 1965). In the late 1990s, it was serendipitously discovered that thalidomide was effective towards the treatment of multiple myeloma (Palumbo et al., 2008). Multiple myeloma is a cancer of the plasma cells, which are normally produced in the bone marrow and produce antibodies to fight off infection. However, during multiple myeloma, one clone hyperproliferates and produces defective versions of antibodies.

Therefore, one of the primary ways that multiple myeloma is diagnosed is through the increased viscosity of the blood from hyperproduction of antibodies (up to 30g/L) (Kopp, Beirne, and Burns, 1967). Treatment of patients with thalidomide decreases the production of aberrant antibodies, indicating that this drug is effective towards multiple myeloma. Subsequently, two different analogs, lenalidomide and pomalidomide, were shown to be significantly more potent and were approved as therapies (Bartlett, Dredge, and Dalgleish, 2004). Collectively, thalidomide and its analogs are referred to as IMiDs, short for immunomodulatory drugs, for their ability to adjust the immune system's response. However, the identity of the primary target of IMiDs was still unknown.

Within five years of approving thalidomide for the treatment of multiple myeloma by the FDA, CRBN was identified as its primary binding partner. In these experiments, Handa and coworkers devised a resin with thalidomide immobilized onto it to capture any interacting partners, and identify the binding partners with shotgun mass spectrometry (Takumi Ito et al., 2010). Thalidomide is a heterobifunctional

molecule with a glutarimide moiety that was shown to bind to the hydrophobic binding pocket of CRBN, and a phthalimide moiety (AEA Lopez-Girona et al., 2012). However, initially it was unknown if thalidomide primarily worked to block the degradation of endogenous substrates, or if it worked by recruiting novel proteins for degradation via the phthalimide moiety. Comparing the effect of thalidomide treatment *in utero* to the phenotype of the knock out animal could distinguish between the two possible outcomes.

In humans, truncating the C terminal domain of CRBN, widely believed to be the substrate binding domain, was shown to result in a mild mental retardation phenotype (Higgins et al., 2004). In mice, ablation of CRBN resulted in a pseudo starvation phenotype (K. M. Lee et al., 2013). Since in both of these cases, animals with CRBN knocked out or where the E3 ligase activity was inhibited were formed with complete limbs, it would stand to reason that the primary teratogenic effect of the drug was not through inhibition. Additionally, ablation of CRBN resulted in resistance to lenalidomide, indicating that the drug could not work through inhibition of CRBN (Zhu et al., 2011). One alternative strategy is to look at how addition of thalidomide or thalidomide analogs influence degradation of specific endogenous substrates. Additional work identified MEIS2 as an endogenous substrate, binding of which was shown to be blocked by addition of IMiDs (Fischer et al., 2014). Therefore, it appears thalidomide can act as both an inhibitor and to recruit novel proteins, which needs to be determined on a case by case basis, showing the importance of being able to identify specific substrates.

Thalidomide is made up of two portions: a glutarimide moiety that was previously shown to be necessary for binding to CRBN, and a phthalimide moiety. The phthalimide part of the thalidomide was shown to mediate interactions with *de novo* substrate, specifically recruiting proteins with a βhairpin loop. Proteins with this motif, like proteins with a C_2H_2 zinc finger motif (α helix and βhairpin, with two cysteines and 2 histidine coordinating a zinc), as well as Casein Kinase 1 α, have been shown to interact with this protein (Petzold, Fischer, and Thomä, 2016). Critically, many of the proteins that were identified to be recruited, like Ikaros, were proteins that were previously believed to be undruggable because of the lack of a well-defined binding site (Lu et al., 2014). Therefore, thalidomide works by recruiting a number of different proteins for degradation. Recent work pinpointed a transcription factor, SALL4, as the protein responsible for the teratogenic effects (Donovan et al., 2018). Additional analogs like lenalidomide and pomalidomide have been shown to be

significantly more potent and more selective, and primarily are modified at the phthalimide ring. Understanding the identity of specific substrates and how varying the substitution along the ring changes the identity of the substrates that are recruited for degradation would provide the information to inform the development of more potent and more selective analogs.

References

Aartsma-Rus, Annemieke et al. (2009). "Theoretic applicability of antisense-mediated exon skipping for Duchenne muscular dystrophy mutations". In: *Human mutation* 30.3, pp. 293–299.

Arendt, Cassandra S and Mark Hochstrasser (1997). "Identification of the yeast 20S proteasome catalytic centers and subunit interactions required for active-site formation". In: *Proceedings of the National Academy of Sciences* 94.14, pp. 7156–7161.

Bachmair, Andreas and Alexander Varshavsky (1989). "The degradation signal in a short-lived protein". In: *cell* 56.6, pp. 1019–1032.

Bartlett, J Blake, Keith Dredge, and Angus G Dalgleish (2004). "The evolution of thalidomide and its IMiD derivatives as anticancer agents". In: *Nature Reviews Cancer* 4.4, p. 314.

Benaroudj, Nadia and Alfred L Goldberg (2000). "PAN, the proteasome-activating nucleotidase from archaebacteria, is a protein-unfolding molecular chaperone". In: *Nature cell biology* 2.11, p. 833.

Buchberger, Alexander et al. (2001). "The UBX domain: a widespread ubiquitin-like module". In: *Journal of molecular biology* 307.1, pp. 17–24.

Chamberlain, Philip P et al. (2014). "Structure of the human Cereblon–DDB1–lenalidomide complex reveals basis for responsiveness to thalidomide analogs". In: *Nature Structural and Molecular Biology* 21.9, p. 803.

Chau, Vincent et al. (1989). "A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein". In: *Science* 243.4898, pp. 1576–1583.

Chen, Li and Kiran Madura (2002). "Rad23 promotes the targeting of proteolytic substrates to the proteasome". In: *Molecular and cellular biology* 22.13, pp. 4902–4913.

Chu-Ping, Ma et al. (1994). "Identification, purification, and characterization of a high molecular weight, ATP-dependent activator (PA700) of the 20 S proteasome." In: *Journal of Biological Chemistry* 269.5, pp. 3539–3547.

Cong, Le et al. (2013). "Multiplex genome engineering using CRISPR/Cas systems". In: *Science*, p. 1231143.

Cromm, Philipp M and Craig M Crews (2017). "Targeted protein degradation: from chemical biology to drug discovery". In: *Cell chemical biology* 24.9, pp. 1181–1190.

Deshaies, RJ (1999). "SCF and Cullin/Ring H2-based ubiquitin ligases". In: *Annual review of cell and developmental biology* 15.1, pp. 435–467.

Donovan, Katherine A et al. (2018). "Thalidomide promotes degradation of SALL4, a transcription factor implicated in Duane Radial Ray syndrome". In: *Elife* 7, e38430.

Duda, David M et al. (2008). "Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation". In: *Cell* 134.6, pp. 995–1006.

Eagle, Michelle et al. (2002). "Survival in Duchenne muscular dystrophy: improvements in life expectancy since 1967 and the impact of home nocturnal ventilation". In: *Neuromuscular disorders* 12.10, pp. 926–929.

Elsasser, Suzanne et al. (2004). "Rad23 and Rpn10 serve as alternative ubiquitin receptors for the proteasome". In: *Journal of Biological Chemistry*.

Fairclough, Rebecca J, Matthew J Wood, and Kay E Davies (2013). "Therapy for Duchenne muscular dystrophy: renewed optimism from genetic approaches". In: *Nature Reviews Genetics* 14.6, p. 373.

Feldman, RM Renny et al. (1997). "A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p". In: *Cell* 91.2, pp. 221–230.

Finley, Daniel et al. (1994). "Inhibition of proteolysis and cell cycle progression in a multiubiquitination-deficient yeast mutant." In: *Molecular and cellular biology* 14.8, pp. 5501–5509.

Fischer, Eric S et al. (2014). "Structure of the DDB1–CRBN E3 ubiquitin ligase in complex with thalidomide". In: *Nature* 512.7512, p. 49.

Garba, Adinoyi O and Shaker A Mousa (2010). "Bevasiranib for the treatment of wet, age-related macular degeneration". In: *Ophthalmology and eye diseases* 2, OED–S4878.

Goldberg, Alfred L (2012). *Development of proteasome inhibitors as research tools and cancer drugs*.

Goldstein, Gideon et al. (1975). "Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells". In: *Proceedings of the National Academy of Sciences* 72.1, pp. 11–15.

Groll, Michael et al. (2000). "A gated channel into the proteasome core particle". In: *Nature Structural and Molecular Biology* 7.11, p. 1062.

Haas, Arthur L (1988). "Immunochemical probes of ubiquitin pool dynamics". In: *Ubiquitin*. Springer, pp. 173–206.

Hartmann, Marcus D et al. (2014). “Thalidomide mimics uridine binding to an aromatic cage in cereblon”. In: *Journal of structural biology* 188.3, pp. 225–232.

Hershko, Avram and Hannah Heller (1985). “Occurrence of a polyubiquitin structure in ubiquitin-protein conjugates”. In: *Biochemical and biophysical research communications* 128.3, pp. 1079–1086.

Higgins, Joseph J et al. (2004). “A mutation in a novel ATP-dependent Lon protease gene in a kindred with mild mental retardation”. In: *Neurology* 63.10, pp. 1927–1931.

Hodgman, Charles D (1951). *Handbook of chemistry and physics*. Vol. 71. 3. LWW.

Hon, Wai-Ching et al. (2002). “Structural basis for the recognition of hydroxyproline in HIF-1α by pVHL”. In: *Nature* 417.6892, p. 975.

Huang, L Eric and H Franklin Bunn (2003). “Hypoxia-inducible factor and its biomedical relevance”. In: *Journal of Biological Chemistry* 278.22, pp. 19575–19578.

Huang, Lan et al. (1999). “Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade”. In: *Science* 286.5443, pp. 1321–1326.

Huang, ME et al. (1988). “Use of all-trans retinoic acid in the treatment of acute promyelocytic leukemia”. In: *Blood* 72.2, pp. 567–572.

Ito, Takumi et al. (2010). “Identification of a primary target of thalidomide teratogenicity”. In: *science* 327.5971, pp. 1345–1350.

Jacobson, Andrew D et al. (2009). “The lysine 48 and lysine 63 ubiquitin conjugates are processed differently by the 26S proteasome”. In: *Journal of Biological Chemistry*, jbc–M109.

Kaur, Jasvinder and Jayanta Debnath (2015). “Autophagy at the crossroads of catabolism and anabolism”. In: *Nature reviews Molecular cell biology* 16.8, p. 461.

Kleiger, Gary et al. (2009). “Rapid E2-E3 assembly and disassembly enable processive ubiquitylation of cullin-RING ubiquitin ligase substrates”. In: *Cell* 139.5, pp. 957–968.

Kopp, William L, Gregory J Beirne, and Robert O Burns (1967). “Hyperviscosity syndrome in multiple myeloma”. In: *The American journal of medicine* 43.1, pp. 141–146.

Lai, Ashton C and Craig M Crews (2017). “Induced protein degradation: an emerging drug discovery paradigm”. In: *Nature Reviews Drug Discovery* 16.2, p. 101.

Lee, Kwang Min et al. (2013). “Disruption of the Cereblon Gene Enhances Hepatic AMPK Activity and Prevents High-Fat Diet–Induced Obesity and Insulin Resistance in Mice”. In: *Diabetes* 62.6, pp. 1855–1864.

Liu, Xing et al. (2018). “Cand1-mediated adaptive exchange mechanism enables variation in F-box protein expression”. In: *Molecular cell* 69.5, pp. 773–786.

Lopez-Girona, AEA et al. (2012). “Cereblon is a direct protein target for immunomodulatory and antiproliferative activities of lenalidomide and pomalidomide”. In: *Leukemia* 26.11, p. 2326.

Lowe, Edward D et al. (2006). “Structures of the Dsk2 UBL and UBA domains and their complex”. In: *Acta Crystallographica Section D: Biological Crystallography* 62.2, pp. 177–188.

Lowe, Jan et al. (1995). “Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution”. In: *Science* 268.5210, pp. 533–539.

Lu, Gang et al. (2014). “The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins”. In: *Science* 343.6168, pp. 305–309.

Lucas, Xavier and Alessio Ciulli (2017). “Recognition of substrate degrons by E3 ubiquitin ligases and modulation by small-molecule mimicry strategies”. In: *Current opinion in structural biology* 44, pp. 101–110.

Mali, Prashant, Kevin M Esvelt, and George M Church (2013). “Cas9 as a versatile tool for engineering biology”. In: *Nature methods* 10.10, p. 957.

Meyer, Hemmo, Monika Bug, and Sebastian Bremer (2012). “Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system”. In: *Nature cell biology* 14.2, p. 117.

Min, Jung-Hyun et al. (2002). “Structure of an HIF-1α-pVHL complex: hydroxyproline recognition in signaling”. In: *Science* 296.5574, pp. 1886–1889.

Mosadeghi, Ruzbeh et al. (2016). “Structural and kinetic analysis of the COP9-Signalosome activation and the cullin-RING ubiquitin ligase deneddylation cycle”. In: *Elife* 5, e12102.

Palumbo, Antonio et al. (2008). “Thalidomide for treatment of multiple myeloma: 10 years later”. In: *Blood* 111.8, pp. 3968–3977.

Petroski, Matthew D and Raymond J Deshaies (2005a). “Function and regulation of cullin–RING ubiquitin ligases”. In: *Nature reviews Molecular cell biology* 6.1, p. 9.

– (2005b). “Mechanism of lysine 48-linked ubiquitin-chain synthesis by the cullin-RING ubiquitin-ligase complex SCF-Cdc34”. In: *Cell* 123.6, pp. 1107–1120.

Petroski, Matthew D, Gary Kleiger, and Raymond J Deshaies (2006). “Evaluation of a diffusion-driven mechanism for substrate ubiquitination by the SCF-Cdc34 ubiquitin ligase complex”. In: *Molecular cell* 24.4, pp. 523–534.

Petzold, Georg, Eric S Fischer, and Nicolas H Thomä (2016). “Structural basis of lenalidomide-induced CK1α degradation by the CRL4 CRBN ubiquitin ligase”. In: *Nature* 532.7597, p. 127.

Pickart, Cecile M and Michael J Eddins (2004). "Ubiquitin: structures, functions, mechanisms". In: *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* 1695.1-3, pp. 55–72.

Pierce, Nathan W et al. (2013). "Cand1 promotes assembly of new SCF complexes through dynamic exchange of F box proteins". In: *Cell* 153.1, pp. 206–215.

Rabl, Julius et al. (2008). "Mechanism of gate opening in the 20S proteasome by the proteasomal ATPases". In: *Molecular cell* 30.3, pp. 360–368.

Reinstein, Eyal and Aaron Ciechanover (2006). "Narrative review: protein degradation and human diseases: the ubiquitin connection". In: *Annals of internal medicine* 145.9, pp. 676–684.

Reitsma, Justin M et al. (2017). "Composition and regulation of the cellular repertoire of SCF ubiquitin ligases". In: *Cell* 171.6, pp. 1326–1339.

Richner, Thomas J and Esther Krook-Magnuson (2018). "Progress Toward Precise Genetic Repair in Neurons". In: *Epilepsy currents* 18.2, pp. 121–122.

Rubin, David M et al. (1998). "Active site mutants in the six regulatory particle ATPases reveal multiple roles for ATP in the proteasome". In: *The EMBO journal* 17.17, pp. 4909–4919.

Saeki, Yasushi et al. (2009). "Lysine 63-linked polyubiquitin chain may serve as a targeting signal for the 26S proteasome". In: *The EMBO journal* 28.4, pp. 359–371.

Seemuller, Erika et al. (1995). "Proteasome from Thermoplasma acidophilum: a threonine protease". In: *Science* 268.5210, pp. 579–582.

Shen, J et al. (2006). "Suppression of ocular neovascularization with siRNA targeting VEGF receptor 1". In: *Gene therapy* 13.3, p. 225.

Sheskin, Jakob (1965). "Thalidomide in the treatment of lepra reactions". In: *Clinical Pharmacology & Therapeutics* 6.3, pp. 303–306.

Skowyra, Dorota et al. (1997). "F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex". In: *Cell* 91.2, pp. 209–219.

Smith, David M et al. (2005). "ATP binding to PAN or the 26S ATPases causes association with the 20S proteasome, gate opening, and translocation of unfolded proteins". In: *Molecular cell* 20.5, pp. 687–698.

Stevenson, George W and Dallas Williamson (1958). "Base Strengths of Cyanoamines1". In: *Journal of the American Chemical Society* 80.22, pp. 5943–5947.

Tanaka, Keiji (2009). "The proteasome: overview of structure and functions". In: *Proceedings of the Japan Academy, Series B* 85.1, pp. 12–36.

Thrower, Julia S et al. (2000). "Recognition of the polyubiquitin proteolytic signal". In: *The EMBO journal* 19.1, pp. 94–102.

Unno, Masaki et al. (2002). “The structure of the mammalian 20S proteasome at 2.75 Å resolution”. In: *Structure* 10.5, pp. 609–618.

Van Nguyen, Thang et al. (2016). “Glutamine triggers acetylation-dependent degradation of glutamine synthetase via the thalidomide receptor cereblon”. In: *Molecular cell* 61.6, pp. 809–820.

Verma, Rati, L Aravind, et al. (2002). “Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome”. In: *Science* 298.5593, pp. 611–615.

Verma, Rati, Robert Oania, et al. (2004). “Multiubiquitin chain receptors define a layer of substrate selectivity in the ubiquitin-proteasome system”. In: *Cell* 118.1, pp. 99–110.

Wenzel, Thorsten et al. (1994). “Existence of a molecular ruler in proteasomes suggested by analysis of degradation products”. In: *FEBS letters* 349.2, pp. 205–209.

Wickliffe, Katherine E et al. (2011). “K11-linked ubiquitin chains as novel regulators of cell division”. In: *Trends in cell biology* 21.11, pp. 656–663.

Xie, Zhiping and Daniel J Klionsky (2007). “Autophagosome formation: core machinery and adaptations”. In: *Nature cell biology* 9.10, p. 1102.

Yao, Tingting and Robert E Cohen (2002). “A cryptic protease couples deubiquitination and degradation by the proteasome”. In: *Nature* 419.6905, p. 403.

Zhang, Yilan et al. (2014). “Comparison of non-canonical PAMs for CRISPR/Cas9-mediated DNA cleavage in human cells”. In: *Scientific reports* 4, p. 5405.

Zheng, Ning, Brenda A Schulman, et al. (2002). “Structure of the Cul1–Rbx1–Skp1–F box Skp2 SCF ubiquitin ligase complex”. In: *Nature* 416.6882, p. 703.

Zheng, Ning, Ping Wang, et al. (2000). “Structure of a c-Cbl–UbcH7 complex: RING domain function in ubiquitin-protein ligases”. In: *Cell* 102.4, pp. 533–539.

Zhu, Yuan Xiao et al. (2011). “Cereblon expression is required for the antimyeloma activity of lenalidomide and pomalidomide”. In: *Blood* 118.18, pp. 4771–4779.

Chapter 2

SUBUNIT SELECTIVE DEGRADATION OF WIZ, A NOVEL SUBSTRATE OF LENALIDOMIDE AND POMALIDOMIDE

2.1 Abstract

Immunomodulators (IMiDs) are an effective class of drugs used to treat blood cancers. IMiDs are believed to work by recruiting protein targets containing a -hairpin motif for ubiquitination by E3 ubiquitin ligase complexes composed of cereblon (CRBN), Cullin-4a (CUL4a), DNA Damage Binding protein-1 (DDB1), and Ring Box-1 (RBX1). The ubiquitinated protein is subsequently degraded by the proteasome. By characterizing the repertoire of proteins that show an increased physical association with CRBN after IMiD treatment, we identified a novel IMiD substrate, Widely Interspaced Zinc Finger Motifs (WIZ). WIZ contains a C_2H_2 zinc finger domain, like several other substrates that were previously characterized. We demonstrate that IMiDs with amine at the 4' position of the phthalimide ring (lenalidomide and pomalidomide) stabilize physical association of WIZ with CRBN, deplete WIZ steady state protein levels in a way that is dependent on E3 ligase activity, and enhance the rate of degradation. These findings illustrate the importance of systematically characterizing the full repertoire of proteins that are targeted for degradation by IMiD compounds, to better understand the mechanisms and consequences of IMiD-induced protein degradation.

2.2 Introduction

Redirecting the existing cellular degradation machinery to remove unwanted components rapidly and specifically is potentially a versatile therapeutic approach (Cromm and Crews, 2017). One possible strategy to achieve this goal is to redirect the proteasome, a large enzyme that digests proteins. Selective elimination of one subunit of a multi-protein complex by this approach is possible because the proteasome can extract a single subunit from a complex, unfold it, and thread it into its inner chamber, where the protease active sites reside (Finley, 2009). Therefore, what is needed is a strategy to redirect the proteasome to remove a therapeutic target in addition to its endogenous substrates.

The proteasome normally identifies and enriches for proteins that need to be remove by using ubiquitin as an affinity tag (Hershko and Heller, 1985; Pickart and Eddins,

2004). Ubiquitin is an 8 kDa protein that is ubiquitously expressed in eukaryotic organisms (Goldstein et al., 1975). The C-terminal glycine on ubiquitin is conjugated to the amine on lysines of the target proteins via an isopeptide linkage. Formation of a ubiquitin polymer with at least 4 monomeric units yields a signal that binds tightly to the proteasome (Thrower et al., 2000; Wickliffe et al., 2011). Therefore, to degrade a therapeutic target, a drug would need to catalyze the ubiquitination of the target to redirect it to the proteasome.

Three enzymes act in tandem to catalyze selective ubiquitination of proteins (Hershko and Heller, 1985; Jentsch, 1992; Hochstrasser and Varshavsky, 1990). E1 first converts the C-terminal carboxylic acid on ubiquitin into a more reactive thioester to enhance the rate of reaction, as the C-terminus of ubiquitin is fairly chemically inert. The activated ubiquitin is transferred as a thioester to an E2, and the E2 ubiquitin thioester then binds to an E3 enzyme, which also binds to substrate. RING-type E3 enzymes position the E2 for subsequent transfer of ubiquitin to a lysine residue on the bound substrate (Petroski and R. J. Deshaies, 2005). Since E3 ubiquitin ligases are the components of this pathway that identify which proteins get ubiquitinated, one strategy to redirect the proteasome to digest a therapeutic target is to modify the substrate specificity of an E3 ubiquitin ligase to recognize the protein of interest.

One class of E3 ubiquitin ligases that have been successfully redirected towards a therapeutic target (i.e. neosubstrate) are cullin-RING type E3 ubiquitin ligases (CRLs) (Angers et al., 2006; Sakamoto, Kim, Kumagai, et al., 2001; Winter et al., 2015; J. Lu et al., 2015; Sakamoto, Kim, Verma, et al., 2003; Schneekloth et al., 2004; Gandhi et al., 2014; Uehara et al., 2017; Han et al., 2017; Bondeson et al., 2015). A distinguishing feature of CRLs is that they contain interchangeable subunits that recruits specific substrates to the enzymatic core (Petroski and R. J. Deshaies, 2005). CRLs work by catalyzing the discharge of ubiquitin from E2 onto a lysine residue of a juxtaposed natural or neosubstrate.

Investigating molecules that work by re-directing CRL activity toward neosubstrates can inform the design of future drugs. One class of drugs that redirect CRL activity towards neosubstrates is referred to as IMiDs, or immunomodulatory drugs (Singhal et al., 1999). IMiDs have been shown to be effective for the treatment of hematological malignancies by redirecting cereblon (CRBN), the interchangeable substrate receptor that determines the specificity of the $CRL4^{CRBN}$ complex to recruit a set of neosubstrates (Fig. 2.1a) (Gandhi et al., 2014; G. Lu et al., 2014; Krönke, Fink, et al., 2015; Krönke, Udeshi, et al., 2014). CRBN is one of many substrate

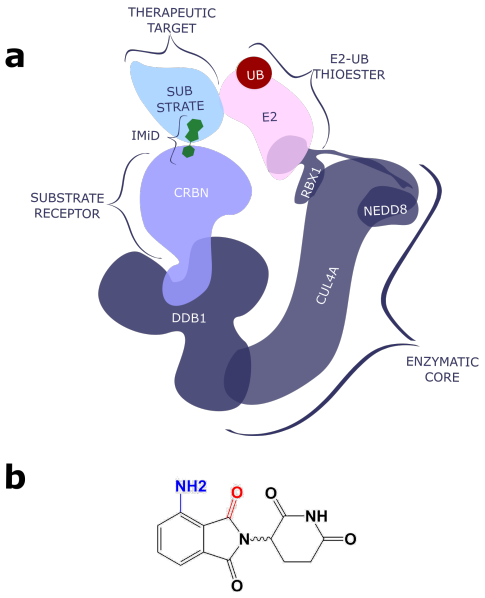

Figure 2.1: Mechanism of action for IMiDs. (a) CRBN is normally in complex with DDB1, which in turn interacts with CUL4-RBX1 to form the ubiquitin ligase $CRL4^{CRBN}$. IMiDs nucleate a novel protein-protein interaction between CRBN and a cellular protein (neo-substrate). Ubiquitin-thioesterified to a ubiquitin conjugating enzyme (E2) is recruited to the $CRL4^{CRBN}$ by RBX1, priming both for bringing activated ubiquitin into proximity of the neosubstrate. (b) General structure of IMiDs. Addition of the blue NH_2 group to thalidomide yields lenalidomide. Removal of the red carbonyl group from lenalidomide yields pomalidomide.

receptors that are recruited to CUL4-RBX1 by DDB1 to form CRL4 complexes (Fig. 2.1a) (Angers et al., 2006; Ito et al., 2010).

IMiDs are bipartite molecules that act as a molecular bridge to link neosubstrates to CRBN. The glutarimide moiety binds CRBN, and the pthalimide moiety binds to protein targets (Fig. 2.1b) (Chamberlain et al., 2014; Fischer et al., 2014). All known IMiD-dependent neosubstrates reported for $CRL4^{CRBN}$ are predicted to contain a β-hairpin motif like Casein Kinase 1α (CK1α) and C_2H_2 zinc finger proteins Ikaros (IKZF1), Aiolos (IKZF3), E3 ubiquitin-protein ligase ZFP91 (ZFP91), Sal-like protein 4 (SALL4) and Aiolos (IKZF3) (Fig. 2.1a) (Gandhi et al., 2014; G. Lu et al., 2014; Krönke, Fink, et al., 2015; Krönke, Udeshi, et al., 2014; Donovan et al., 2018; An et al., 2017). Crystal structures have shown that β-hairpin motif comes into contact with the pthalimide moiety of IMiDs, which explains how these drugs work (Petzold, Fischer, and Thomä, 2016; Matyskiela et al., 2016).

Because novel substrates are recruited to CRBN through the pthalimide moiety

of IMiDs, understanding how modification of the basic bicyclic IMiD scaffold influences the identity of the neosubstrates that are recruited to CRBN will be critical towards the development of new analogs with different patterns of selectivity (Fig. 2.1b). Interestingly, relatively modest alterations to the IMiD scaffold can have surprisingly dramatic effects. For example, substitution of an amine at the 4' position of thalidomide (Fig. 2.1b) (as occurs in lenalidomide and pomalidomide) results in a 100 to 1,000-fold increase in potency (Muller et al., 1999). Another dramatic example of specificity is seen with deletion of a simple carbonyl from pomalidomide to generate lenalidomide (Fig. 2.1b), which enables targeting of the neosubstrate CK1α for ubiquitination and degradation (Krönke, Fink, et al., 2015; Petzold, Fischer, and Thomä, 2016). Thus, relatively subtle modifications to the IMiD chemical scaffold result in dramatic changes in its biological properties. Additionally, it remains unknown how the effects of IMiDs are modified in different cell types. If we are to understand how to harness targeted protein degradation, it will be important to understand the physical and chemical basis for targeting by identifying and characterizing specific neosubstrates.

IMiDs enhance physical association of CRBN with proteins with Zinc Finger domains

Previous IMiD substrates have been identified in proteomics experiments by determining which protein accumulate as ubiquitinated intermediates or exhibit a shorter half life in the presence of IMiDs (G. Lu et al., 2014; An et al., 2017; Krönke, Udeshi, et al., 2014; Donovan et al., 2018; Krönke, Fink, et al., 2015). However, there are several disadvantages to both of these methods. First, ubiquitinated intermediates typically account for a tiny fraction of the total protein, and only a tiny fraction of all peptides from a protein have ubiquitin conjugated to them (R. Deshaies, 1999). Second, the rate of decay is nonlinear, and the half-life of proteins can range from minutes to months, potentially confounding a large-scale screen analyzed at a single time point (Toyama et al., 2013). To circumvent these issues, we chose to identify proteins whose physical association with CRBN was enhanced upon treatment with IMiD, as we reasoned that this was the most direct and sensitive method. To accomplish this, we used a THP-1 cell line that was transiently transfected to express FLAG-CRBN (Fig. 2.2a).

We identified candidate proteins by treating cells with or without lenalidomide for 2 hours, immunoprecipitating for CRBN via the FLAG epitope, and identifying which proteins showed enhanced physical association with CRBN in the presence of

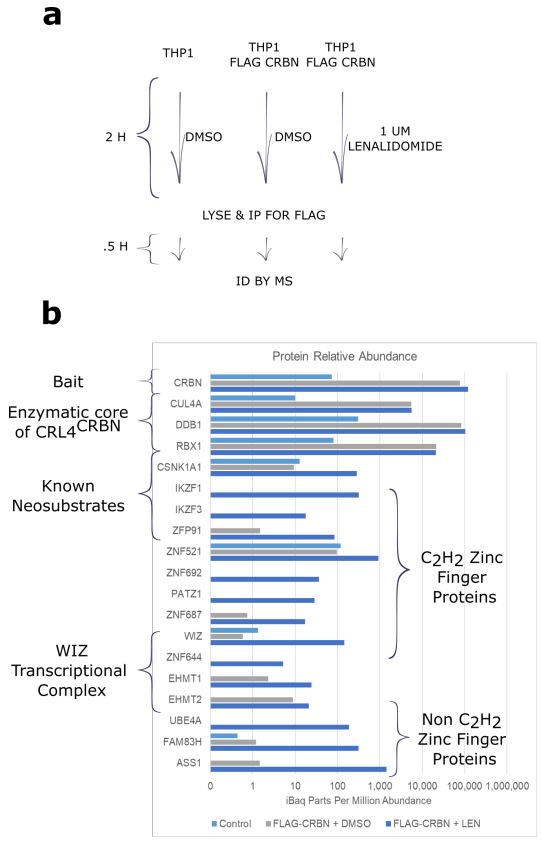

Figure 2.2: Proteins with Zinc Finger Domains show enhanced IMiD dependent enhanced physical association with CRBN. (a) Outline of IP-MS strategy to identify proteins that show an enhanced physical association. (b) Peptide count of proteins identified in the experiment. This includes those known to bind CRBN, known core components of the $CRL4^{CRBN}$ complex, zinc finger proteins, and novel binding partners.

lenalidomide via quantitative shotgun mass spectrometry (Fig. 2.2a). Components of the enzymatic core of $CRL4^{CRBN}$ (CUL4a and RBX1) and the adapter (DDB1) that scaffolds CRBN's interaction with the enzymatic core were recovered in equal amounts regardless of the presence or absence of lenalidomide (Fig. 2.2b).

Previously characterized neosubstrates (ZFP91, IKZF1, IKZF3, CK1α) showed IMiD-dependent association, validating this approach (Fig. 2.2b) (Gandhi et al., 2014; An et al., 2017; Krönke, Fink, et al., 2015; G. Lu et al., 2014; Krönke, Udeshi, et al., 2014). However, one limitation of this approach is that enhanced physical association with CRBN might not necessarily result in a productive degradation event (Eichner et al., 2016).

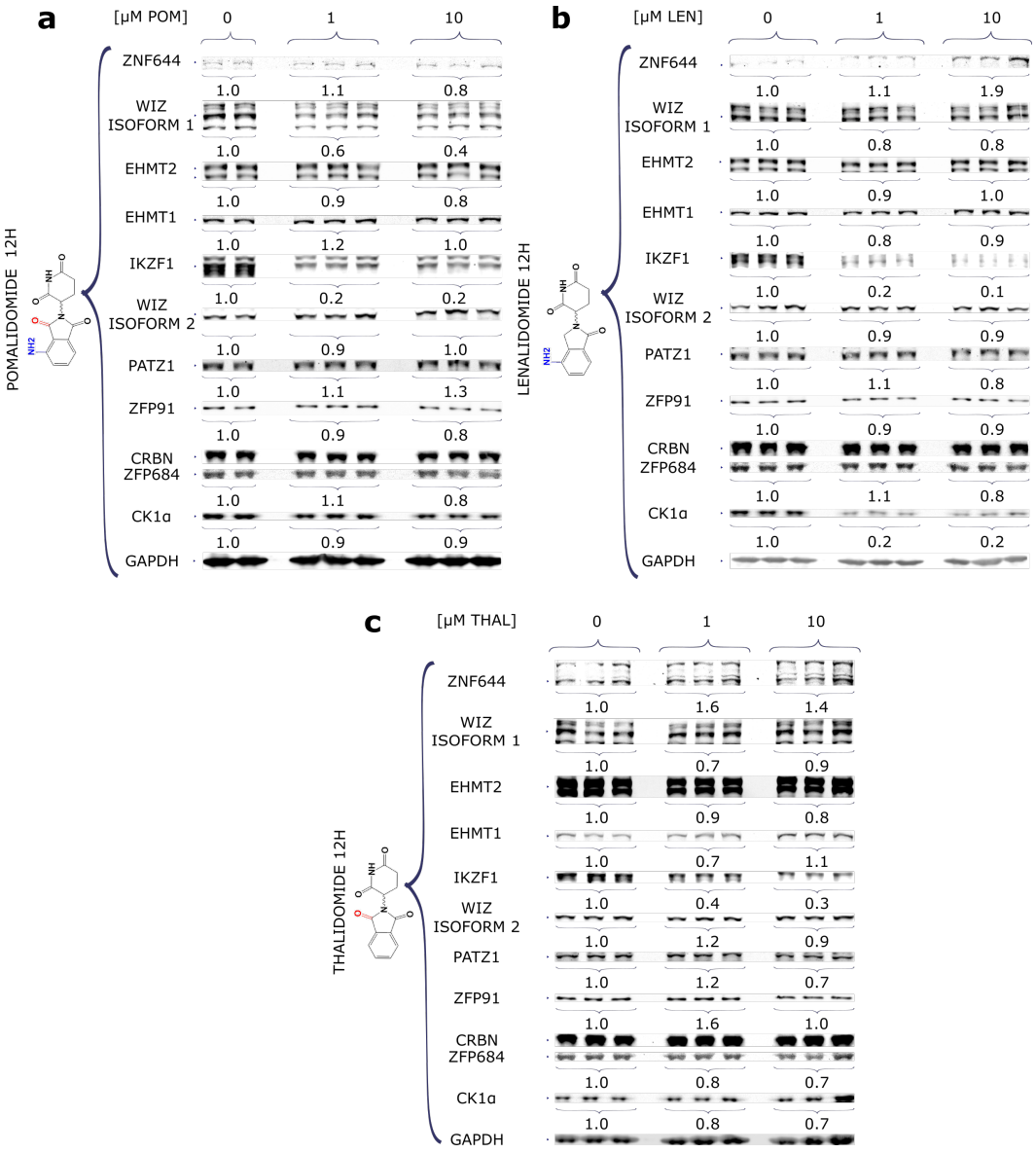

Figure 2.3: WIZ protein abundance shows IMiD dependent regulation. L363 cells were maintained in RPMI-1640 with 10% FBS and 2 μM glutamine before treatment with (a) pomalidomide, (b) lenalidomide, or (c) thalidomide for 12 hours at 0, 1 or 10 μM. Cell lysates were processed for immunoblotting and he signals were quantified on a LICOR odyssey. The median level of each protein normalized to no drug and the GAPDH signal in the same samples is shown below the immunoblot image for each set of triplicates. All protein levels were normalized to GAPDH, which served as a loading control. The previously identified substrates ZFP91, IKZF1, and CK1α serve as positive controls. PATZ1, ZFP684, and ZNF644 are the other zinc finger proteins screened.

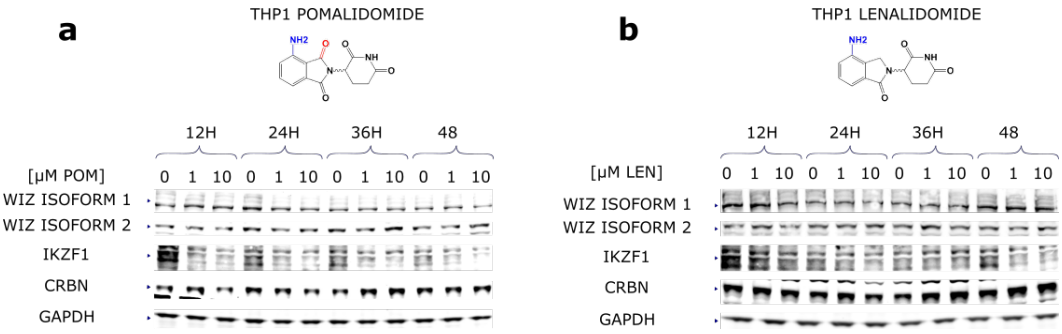

Figure 2.4: IMiD regulation is cell type specificity. THP-1 Cells were then treated with (a) pomalidomide or (b) lenalidomide 0, 1, or 10 μM. for 12, 24, 36, and 48 hours. Cell lysates were separated by SDS PAGE, and immunoblotted against antibodies for WIZ, CRBN, IKZF1, CK1α, and GAPDH.

2.3 IMiD regulation shows cell type dependence

To understand how generalizable IMiD-dependent depletion of WIZ was across different cell types, THP-1 cells were treated with IMiDs (0, 1, and 10 μM) with a 4' amine modification for 12, 24, 36, and 48 hours (Fig. 2.4a–b). Previously characterized IMiD dependent substrates like IKZF1 (G. Lu et al., 2014; Krönke, Udeshi, et al., 2014; Gandhi et al., 2014) were also immunoblotted as a positive control. Unlike what we observed in L363 cells (Fig. 2.3a), no reduction of WIZ was observed in THP-1 cells (Fig. 2.4a-b), indicating that the IMiD effect shows cell type dependence. Previous work had indicated that WIZ was downregulated in H9 hESC (Donovan et al., 2018). Thus, the IMiD effect on WIZ shows cell type dependence.

WIZ binding partners are not regulated by IMiD

WIZ is a protein that is normally in a transcriptional complex with EHMT1, EHMT2, and ZNF644 (Bian, Chen, and Yu, 2015; Ueda et al., 2006). It is believed to target methyltransferases to specific DNA sequences via its zinc finger motif. To investigate out how many of the IMiD-dependent binding events translated into productive degradation, we examined by immunoblotting a subset of the identified proteins to see if their steady state levels dropped after IMiD treatment(Fig. 2.3a–c). We selected proteins with a zinc finger domain, because some previously characterized IMiD-dependent substrates (e.g. IKZF1, Aiolos, ZFP91, ZFP692) are zinc finger proteins (Gandhi et al., 2014; An et al., 2017; Krönke, Udeshi, et al., 2014; G. Lu et al., 2014; Krönke, Fink, et al., 2015; Donovan et al., 2018). L363 cells were treated with pomalidomide, lenalidomide, or thalidomide for 12 hours at multiple doses (0, 1, and

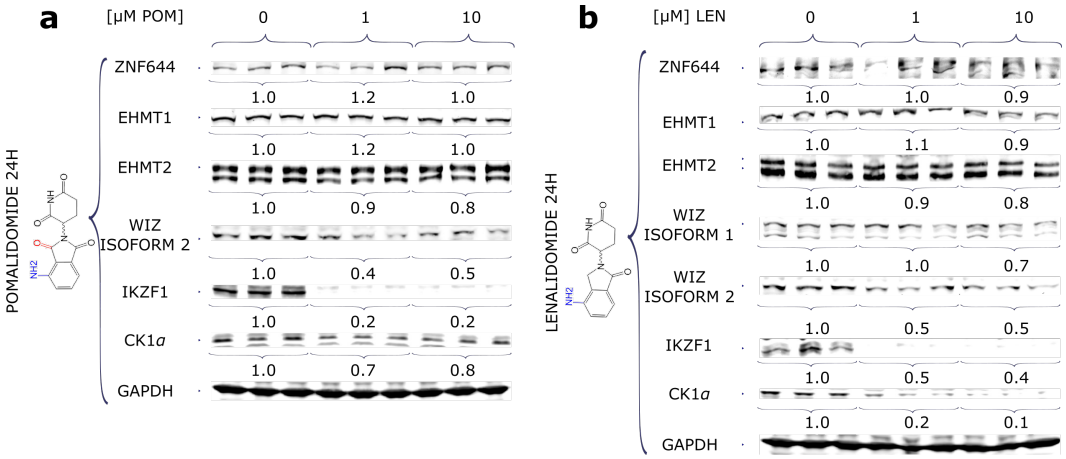

Figure 2.5: IMiDs regulate the abundance of WIZ, but not the other components of the complex EHMT1, EHMT2, and ZNF644. MM.1s cells were treated in triplicate for 24 hours with (a) pomalidomide or (b) lenalidomide at 0, 1, or 10 μM. Cell lysates were separated by SDS PAGE, and immunoblotted against antibodies against Wiz, EHMT1, EHMT2, and ZNF644.

10 μM), in biological triplicate (Fig. 2.3a–c). Both isoforms of WIZ (the identity of the isoforms was confirmed by knockdown) (Fig. 2.6) showed the largest fold change after treatment with pomalidomide (Fig. 2.3a). Previously characterized substrates (IKZF1) showed IMiD dependent depletion (Fig. 2.3a–c). Additionally, substrates such as CK1α that were known to be specific to lenalidomide, showed lenalidomide dependent depletion, (Fig. 2.3b) as had been reported previously (Krönke, Fink, et al., 2015).

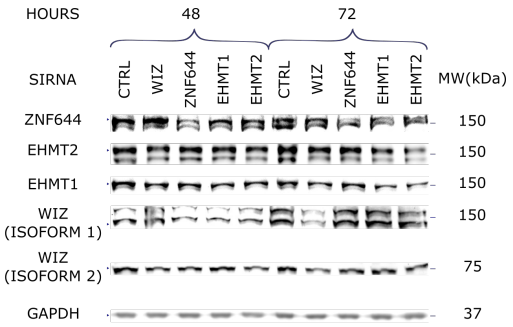

Figure 2.6: THP-1 cells were transiently transfected with an siRNA against WIZ, EHMT1, EHMT2, or ZNF644. Cells were harvested after 48 or 72 hours and immmunoblotted against different commercially available antibodies.

The reduction in WIZ increased over time and showed a larger effect after treatment pomalidomide than lenalidomide (Fig. 2.3a–c). WIZ is a protein that is normally

in a transcriptional complex with EHMT1, EHMT2, and ZNF644 (Bian, Chen, and Yu, 2015). It is believed to target methyltransferases to specific DNA sequences via its zinc finger motif. Our mass spec data revealed the entire complex was recruited to CRBN after treatment with IMiD (Fig. 2.2b). However, immunoblotting indicated that WIZ was the only component of this complex that was depleted (Fig. 2.5a,b).

WIZ binding partners are not regulated by IMiD

WIZ is a protein that is normally ina transcriptional complex with EHMT1, EHMT2, and ZNF644 (Bian, Chen, and Yu, 2015; Ueda et al., 2006) It is believed to target methyltransferases to specific DNA sequences via its zinc finger motif. Our mass spec data revealed the entire complex was recruited to CRBN after treatment with IMiD (Fig. 2.1b). To tease out if WIZ was the only component of the transcriptional complex that showed regulation, MM.1s cells were treated with lenalidomide or pomalidomide (0, 1 or 10 μM) for 24 hours and WIZ and EHMT1, EHMT2, and ZNF644 were evaluated by immunoblotting (Fig. 2.5a–c). Depletion of both WIZ and the validated substrate IKZF1 were observed, but the other components of the WIZ transcriptional complex remained stable.

Levels of WIZ are specifically modulated by IMiDs with an amine modification at the 4' position

To elucidate how subtle modifications in the IMiD scaffold influenced WIZ depletion, MM.1S cells were treated with pomalidomide, lenalidomide, or thalidomide (0, 1, and 10 μM) at 12 , 24, 36, and 48 hours (Fig. 2.7a–c). Both isoforms of WIZ showed the largest fold change upon treatment with IMiDs with an amine in the 4' position (lenalidomide and pomalidomide)(Fig. 2.7a–b).

IMiD dependent depletion of WIZ is dependent on CRBN's E3 ubiquitin ligase activity

IMiDs enhanced the physical association between WIZ and CRBN, resulting in the depletion of WIZ which is consistent with the hypothesis that WIZ is a new $CRL4^{CRBN}$ neosubstrate.

We next sought to directly test this hypothesis. Depletion of CRBN in MM.1s cells strongly blocked pomalidomide-induced depletion of WIZ (Fig. 2.7b, d), confirming that CRBN was indeed required for the effect. To address the role of CRL E3 activity and the proteasome, L363 cells were treated with 0, 1, or 10 μM of pomalidomide for 12 hours in the absence or presence of the Nedd8 conjugation inhibitor pevonedistat

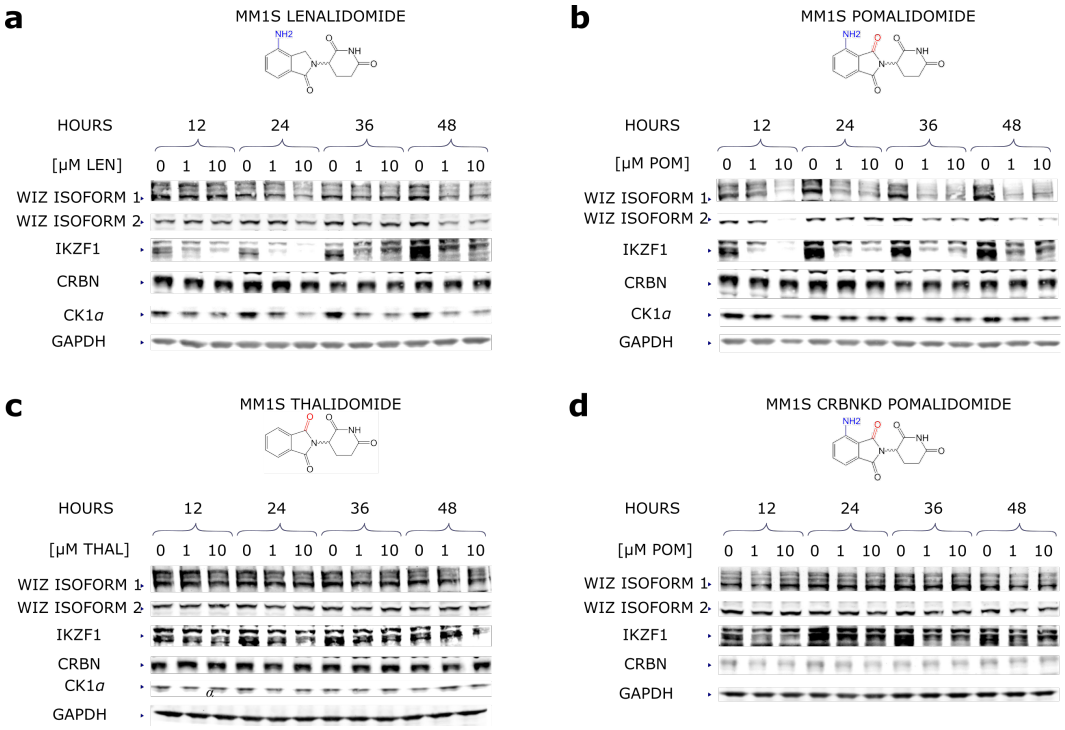

Figure 2.7: Wiz protein abundance is specific to IMiDs modified with an amine at the 4' position. (a-c) MM.1s or (d) MM.1s cells with CRBN knocked down with an shRNA were treated with (a) lenalidomide, (b) (d) pomalidomide, or (c) thalidomide at 0, 1, or 10 μM for 12- 48 hours. Cell lysates were separated by SDS PAGE, and immunoblotted against antibodies for WIZ, CRBN, IKZF1, CK1α, and GAPDH.

or the proteasome inhibitor bortezomib (Fig. 2.8a-c) (Adams et al., 1999; Soucy et al., 2009). IMiD depletion of WIZ and known substrates such as IKZF1 was blocked upon cotreatment with proteasome inhibitor, indicating the effect was dependent on proteasome activity (Fig. 2.8a,b).

Nedd8 is a ubiquitin-like protein that activates CRLs upon its conjugation to the cullin subunit (R. J. Deshaies, Emberley, and Saha, 2010). Co-treatment with 250 nM pevondistat showed accumulation of CUL4 in the de-neddylated version, indicating that the enzymatic core of the complex would be catalytically inactive (Fig. 2.8a, c). IMiD depletion of WIZ and known substrates such as IKZF1 was also blocked upon cotreatment with pevondistat, indicating IMiD induced depletion was dependent upon the core complex being enzymatically active (Fig. 2.8a,c).

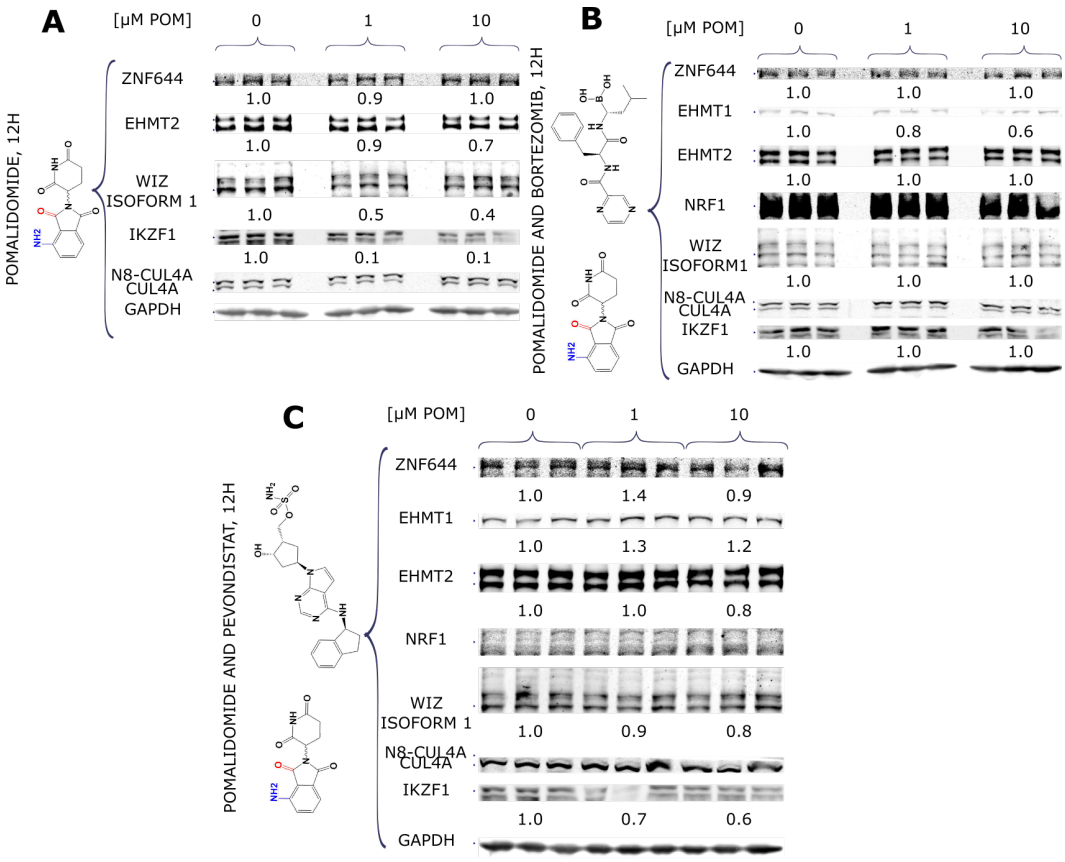

Figure 2.8: IMiD induced degradation of WIZ is dependent on E3 ubiquitin ligase activity and proteasome degradation. L363 cells were maintained in RPMI with 10% FBS before treatment with pomalidomide at 0, 1, or 10 μM in biological triplicate. They were cotreated with either (a) DMSO, (b) a proteasome inhibitor bortezomib , or (c) a Nedd-8 conjugation inhibitor pevodnestat. Cell lysates were separated by SDS-page and immunoblotted against the indicated antibodies. NRF1 was used as a positive control, for the proteasome inhibitor. IKZF1 was a positive control for IMiD treatment. GAPDH was the loading control. The action of pevondistat was confirmed by its effect on CUL4 neddylation.

IMiD dependent depletion of WIZ protein levels was due to catalyzing the rate of decay

Although we found IMiD dependent depletion of WIZ was dependent on both CRBN and CRL ubiquitin ligase activity (Fig. 2.7d) (Fig. 2.8a–c), it could be due to increased degradation of WIZ or inhibition of WIZ expression. To decouple these two possible outcomes, we compared the rate of decay by treating cells with or without IMiD in the presence of an inhibitor of translation (Fig. 2.8a-c) (Obrig et al., 1971). We saw that both isoforms of WIZ, as well as IKZF1, a known IMiD substrate, experienced a higher rate of decay upon addition of IMiD (Fig. 2.8b,c), consistent with this being due to an increased rate of degradation (G. Lu et al., 2014; Krönke, Udeshi, et al., 2014).

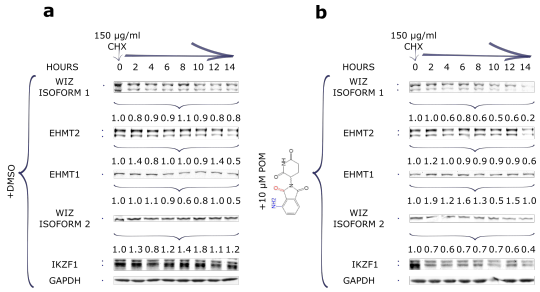

Figure 2.9: Pomalidomide stimulates the degradation of WIZ. MM.1s cells were maintained in RPMI 1640 before treatment with cycloheximide (150 µM/mL) with or without 10 µM of pomalidomide. Cells were harvested at 0, 2, 4, 6, 8, 12, and 14 hours before being separated by SDS-PAGE and immunoblotted for Wiz, IKZF1, and GAPDH.

Discussion

Targeted protein degradation is rapidly becoming an area of intense interest for the biopharmaceutical industry (Crews, 2010). The overall goal for researchers in this area is to design a drug that selectively targets a particular protein of interest for ubiquitin-dependent degradation via the proteasome. One particular class of drugs, the IMiDs, achieves its effects through selective destruction of specific neosubstrates via the $CRL4^{CRBN}$ ubiquitin ligase pathway (Ito et al., 2010; G. Lu et al., 2014; Krönke, Udeshi, et al., 2014). Interestingly, different IMiDs that differ from each other in subtle ways elicit the degradation of different sets of overlapping neosubstrates. To better understand the biological effects of IMiDs and to gain insight into the mechanisms by which IMiDs recruit substrates to CRBN, we employed an affinity purification-mass spectrometry approach to identify proteins that bound CRBN in an IMiD-dependent manner (Fig. 2.2a,b). This led to the identification

of a dozen proteins, including the previously-described $CRL4^{CRBN}$ neosubstrates IKZF1, IKZF3, ZFP91, ZNF692 and CK1α (Fig. 2.2a,b)(Krönke, Fink, et al., 2015; Krönke, Udeshi, et al., 2014; G. Lu et al., 2014; An et al., 2017; Sievers et al., 2018). In addition, we identified a number of zinc finger proteins, including WIZ. Note that while this work was being being prepared for submission, Thoma and colleagues reported the interaction of a large collection of zinc fingers with CRBN, including those reported here (Sievers et al., 2018).

We show here that WIZ is a *bonda-fide* neosubstrate for $CRL4^{CRBN}$ (Fig. 2.7a,d). In addition to increasing association of WIZ with CRBN, IMiDs also induced depletion of WIZ that was dependent on the CRBN protein, CRL activity (as determined with the pan-CRL inhibitor pevonedistat), and proteasome activity (Fig. 2.9a–c). Importantly, depletion of WIZ was due to an increase in its rate of degradation (Fig. 2.9a,b). The effect of IMiDs on WIZ degradation was selective, in that only IMiDs with an amino group at the 4' position (lenalidomide and pomalidomide) induced degradation (Fig. 2.7a,b). This is consistent with prior observations that subtle structural differences in IMiD structure can have powerful discriminating effects on substrates such as CK1α (Krönke, Udeshi, et al., 2014; Krönke, Fink, et al., 2015). In addition, we observed IMiD-dependent depletion of WIZ in L363 but not in THP-1 cells. We do not know the reason for this selectivity. However, it is not due to the inability of WIZ to bind transfected CRBN in an IMiD-dependent manner in these cells(Fig. 2.2b). This selectivity points to the potential to develop therapeutic agents that induce degradation of target proteins in a cell type specific manner.

Our data, combined with other recently reported results, raise the question of whether the effects of IMiDs *in vivo* are more complex than originally envisioned, and could potentially be influenced by depletion of multiple proteins in addition to the original neosubstrates IKZF1 and IKZF3 (Sievers et al., 2018).

A great deal of effort is being invested in developing proteolysis-targeting chimeric molecules (PROTACs) that, like IMiDs, induce the degradation of specific proteins to achieve a therapeutic effect (R. J. Deshaies, 2015). A result we observed here that is of particular relevance to those efforts is that IMiD induced depletion is highly selective for a single subunit for a multisubunit complex. WIZ associates with EHMT1, EHMT2, and ZNF644 to form a gene regulatory complex, and all four of these proteins were recruited to CRBN in the presence of IMiD (Fig. 2.2a,b) (An et al., 2017; Ueda et al., 2006). However, only WIZ was degraded. This is consistent

with the previously observed subunit selectivity of the ubiquitin proteasome pathway both *in vivo* and in reconstituted systems (Verma et al., 2001; Johnson, Gonda, and Varshavsky, 1990). Harnessing the power of subunit-selective degradation to remodel cellular machines to achieve a highly selective therapeutic outcome holds promise for the development of a suite of next-generation therapeutics.

Experimental Procedures

Materials

Bortezomib (B-1408) was purchased from LC Laboratories. Pomalidomide (P0018-25MG) and N-ethyl maleimide (E3876-5G) were purchased from Sigma Aldrich. Lenalidomide (HY-A0003) and pevonedistat (905579-51-3) were purchased from MedChem Express. All were prepared as single-use DMSO stocks, and stored at -20o C before usage. Laemmli Buffer (1610737) was purchased from Biorad. BSA (9998S) was purchased from Cell Signaling Technology.

Antibodies

GAPDH (SC-365062) antibody was from Santa Cruz Biotechnology. CUL4a (2699S), DDB1 (A300-462A), EHMT1 (A301-642A-M), and ZNF644 (A301-642a-M) antibodies were from Bethyl Laboratories. CRBN (HPA045910) and FLAG (F1804) antibodies, were from Sigma Aldrich. WIZ (ab92334), ZNF521 (ab156271), CK1α (ab108296), and EHMT1 (ab185050) antibodies were from Abcam. PATZ1 (PA5-30478), ZNF684 (PA5-40984), and ZFP91 (PA5-43064) antibodies were from Thermo Fischer Scientific. Anti-Rabbit IgG IR800 (926-32211) was purchased from Li-COR Biosciences. Anti-Mouse IgG IR680 (A10038) was purchased from Invitrogen.

Cell lines

MM.1s and THP-1 cells were purchased from ATCC. L363 cells were provided by Francesco Parlati, Calithera Biosciences, South San Francisco. Cells were grown in RPMI-1640 (ATCC formulation), with 10% heat-inactivated fetal bovine serum (FBS), 2 μM glutamine, and penicillin-steptomycin. THP-1 cells were supplemented with 50 μM β-mercaptoethanol. Cell lines were tested for mycoplasma using Lonza's MycoAlert Mycoplasma Detection kit and authenticated by Laragen using PowerPlex 16 system, as well as the Cell Line ID.

Western blot analysis

MM.1s cell for each sample, as well as a replicate to measure protein levels were harvested by centrifugation and resuspended in pre-warmed RPMI 16–40 at 1 million cells/mL. They were then seeded onto 24 well plates, for 1.5 mL total volume. Pomalidomide was added from a 1000x stock. Cells were spun down, and the medium was aspirated, and flash frozen for processing later. A separate replicate to normalize protein concentration was rinsed with PBS 2x, before being flash frozen. Samples were then solubilized in 200 μL of Laemmle buffer supplemented with fresh β-mercaptoethanol and protease inhibitor cocktail. Cell lysates were sonicated (5 s; 10%) and cleared by centrifugation before boiling for 3 minutes. Boiled lysates were cleared by centrifugation (14,000 xg). Relative concentrations were calculated based on the BCA assays, and the total lysate volume was adjusted according to that. Samples were loaded onto 4–12% protein gels and transferred for 3 hours at 70 V, with stirring.

RNAi-Mediated knockdown

siRNAs were purchased from Santa Cruz Biotechnology. Cells were seeded at a final concentration of 500,000/mL (THP-1 cells), and transfected with Opti-MEM and Lipofectamine RNAimax, as previously described. Cells were harvested after 48, and 72 hours, and analyzed for efficiency of knockdown.

Cycloheximide Chase

MM.1s cells were seeded at 10^6 cells/mL in complete RPMI 1640 with 10% FBS in 24 well plates. Cells were treated with 0 or 10 μM pomalidomide, along with 150 μg/mL of cycloheximide to initiate the chase. Cells were harvested at the indicated time, and subjected to immunoblot analysis.

Immunoprecipitation

THP-1 cells were transiently transfected with a plasmid expressing FLAG CRBN. After 36 hours, cells were treated with lenalidomide for 2 hours (1 μM). Cells were lysed in 50 mM HEPES, pH 7.5, 70 mM KOAc, 5 mM $Mg(OAc)_2$, 0.2% n-dodecyl-β-D-maltoside with protease inhibitor for 30 minutes at 4 °C. Lysates were pelleted for 15 minutes at 16,000 g to remove cellular debris. Supernatent was incubated with anti-FLAG M2 affinity gel for 2 hr at 4 °C. The anti-FLAG M2 affinity gel was washed with lysis buffer 3 times, then with 100 mM Tris-HCL (pH 8.5) two times. Samples were eluted with 100 mM Tris-HCl (pH 8.5) with 10M urea.

Mass Spectrometric analysis

Eluted samples were spiked with 4.48 mM DTT (1/278) to reach a final concentration of 1.25 M DTT, to be reduced. Reduced samples were alkylated with 10 mM NEM for 30 min at room temperature in the dark. Reduced and alkylated samples were digested with Lys-C (Wako) at a 1:200 ratio for 4 hr at RT. Partially digested lysates spiked with 1 mM $CaCl_2$ were diluted with 50 mM Tris (pH 8.0) down to 2 M urea before being digested with sequence grade trypsin (Promega) at 1:100 37°C overnight in the dark. The reaction was quenched after 15 hr by adding trifluoroacetic acid (TFA) to 0.1%. Lysates were cleared via centrifugation at 4000xg for 15 minutes. Peptides were were desalted using a 500 mg capacity Sep-pak column that initially was hydrated using 7 column volumes of ACN (21 ml), followed by an equilibration step with 7 column volumes of Buffer A (0.1% TFA in H2O) (21 ml). Cleared peptides were loaded onto the resin by gravity flow, washed with 7 column volumes of Buffer A, followed by 3 column volumes of Wash buffer (0.1% TFA, 5% ACN in H2O). Desalted peptides were eluted using 2 column volumes of Elution buffer (0.1% TFA, 40% ACN in H_2O) (6 ml). The resulting peptide sample was frozen by storing at 80°C for at least 1 hr and dried via lyphilization.

NanoLC-MS/MS analysis

The dried immunoprecipitated peptides were resuspend in Buffer A (0.2% Formic Acid, 2% ACN, nanoLC grade 97.8% H2O) and subjected to proteomic analysis using an EASY II nano-UPLC (Thermo Fisher Scientific) connected on-line to an Orbitrap Elite hybrid mass spectrometer with a nanoelectrospray ion source (Thermo Scientific) using settings similar to those previously described (Porras-Yakushi and Hess, 2014). Peptides were separated using a 15 cm silica analytical column with a 75 µm inner diameter packed in-house with reversed phase ReproSil-Pur C18AQ 3 µm resin (Dr Maisch GmbH, Amerbuch-Entringen, Germany). The flow rate was set to 350 nl/min, using a linear gradient of 2%-32% B (0.2% Formic Acid, 80% ACN, 19.8% nanoLC grade H_2O). Mass spectrometry detectable samples were analyzed on a 159 min gradient, while basic reversed phase immunoprecipitated samples were analyzed on a 90 min gradient. The mass spectrometer was set to collect data in a data-dependent mode, switching automatically between full-scan MS and tandem MS acquisition. All samples were analyzed by ETD and decision tree fragmentation. For ETD fragmentation, the fifteen most intense precursor ions were selected, while the 20 most intense ions were selected for fragmentation using

the decision tree method. Data acquisition was managed using Xcalibur 2.0.7 and Tune 2.4 software (Thermo Fisher Scientific).

Acknowledgements

Thanks to Rati Verma, Jennifer Mamrosh, Emily Blythe, Thang Nguyen, and David Sherman for scientific discussion. H.H.Y was supported by NSF GRFP fellowship. R.J.D. was an investigator at the Howard Hughes Medical institute, and this work was supported by HHMI and a gift from Amgen.

Data Analysis

Raw data was searched using MaxQuant (Jürgen Cox and Mann, 2008; Jurgen Cox et al., 2011), version 1.6.5.0. Precursor mass tolerance was set to 4.5 ppm after recalibration and fragment tolerance was 0.5 Da. Oxidation of methionine and protein N-terminal acetylation were specified as variable modifications and carbamidomethylation of cysteine was specified as a fixed modification. Trypsin was specified as the digestion enzyme, with up to two missed cleavages allowed. Spectra were searched against the UniProt human database (93591 entries) and a contaminant database (246 entries). Score thresholds were established so that peptide and protein false discovery rates were less than 1% as estimated by a target-decoy approach. Match between runs and iBAQ quantitation (Schwanhäusser et al., 2011) were enabled. iBAQ parts per million (ppm) abundances were calculated by multiplying individual protein iBAQ abundances by 10^6 and dividing by the sum of all non-contaminant protein iBAQ values

References

Adams, Julian et al. (1999). “Proteasome inhibitors: a novel class of potent and effective antitumor agents”. In: *Cancer research* 59.11, pp. 2615–2622.

An, Jian et al. (2017). “pSILAC mass spectrometry reveals ZFP91 as IMiD-dependent substrate of the CRL4 CRBN ubiquitin ligase”. In: *Nature Communications* 8, p. 15398.

Angers, Stephane et al. (2006). “Molecular architecture and assembly of the DDB1–CUL4A ubiquitin ligase machinery”. In: *Nature* 443.7111, p. 590.

Bian, Chunjing, Qiang Chen, and Xiaochun Yu (2015). “The zinc finger proteins ZNF644 and WIZ regulate the G9a/GLP complex for gene repression”. In: *Elife* 4.

Bondeson, Daniel P et al. (2015). “Catalytic in vivo protein knockdown by small-molecule PROTACs”. In: *Nature chemical biology* 11.8, p. 611.

Chamberlain, Philip P et al. (2014). "Structure of the human Cereblon–DDB1–lenalidomide complex reveals basis for responsiveness to thalidomide analogs". In: *Nature Structural and Molecular Biology* 21.9, p. 803.

Cox, Jürgen and Matthias Mann (2008). "MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification". In: *Nature biotechnology* 26.12, p. 1367.

Cox, Jurgen et al. (2011). "Andromeda: a peptide search engine integrated into the MaxQuant environment". In: *Journal of proteome research* 10.4, pp. 1794–1805.

Crews, Craig M (2010). "Targeting the undruggable proteome: the small molecules of my dreams". In: *Chemistry & biology* 17.6, pp. 551–555.

Cromm, Philipp M and Craig M Crews (2017). "Targeted protein degradation: from chemical biology to drug discovery". In: *Cell chemical biology* 24.9, pp. 1181–1190.

Deshaies, Raymond J (2015). "Protein degradation: prime time for PROTACs". In: *Nature chemical biology* 11.9, p. 634.

Deshaies, Raymond J, Ethan D Emberley, and Anjanabha Saha (2010). "Control of cullin-ring ubiquitin ligase activity by nedd8". In: *Conjugation and deconjugation of ubiquitin family modifiers*. Springer, pp. 41–56.

Deshaies, RJ (1999). "SCF and Cullin/Ring H2-based ubiquitin ligases". In: *Annual review of cell and developmental biology* 15.1, pp. 435–467.

Donovan, Katherine A et al. (2018). "Thalidomide promotes degradation of SALL4, a transcription factor implicated in Duane Radial Ray syndrome". In: *Elife* 7, e38430.

Eichner, Ruth et al. (2016). "Immunomodulatory drugs disrupt the cereblon–CD147–MCT1 axis to exert antitumor activity and teratogenicity". In: *Nature medicine* 22.7, p. 735.

Finley, Daniel (2009). "Recognition and processing of ubiquitin-protein conjugates by the proteasome". In: *Annual review of biochemistry* 78, pp. 477–513.

Fischer, Eric S et al. (2014). "Structure of the DDB1–CRBN E3 ubiquitin ligase in complex with thalidomide". In: *Nature* 512.7512, p. 49.

Gandhi, Anita K et al. (2014). "Immunomodulatory agents lenalidomide and pomalidomide co-stimulate T cells by inducing degradation of T cell repressors I karos and A iolos via modulation of the E 3 ubiquitin ligase complex CRL 4 CRBN". In: *British journal of haematology* 164.6, pp. 811–821.

Goldstein, Gideon et al. (1975). "Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells". In: *Proceedings of the National Academy of Sciences* 72.1, pp. 11–15.

Han, Ting et al. (2017). "Anticancer sulfonamides target splicing by inducing RBM39 degradation via recruitment to DCAF15". In: *Science* 356.6336, eaal3755.

Hershko, Avram and Hannah Heller (1985). "Occurrence of a polyubiquitin structure in ubiquitin-protein conjugates". In: *Biochemical and biophysical research communications* 128.3, pp. 1079–1086.

Hochstrasser, Mark and Alexander Varshavsky (1990). "In vivo degradation of a transcriptional regulator: the yeast α2 repressor". In: *Cell* 61.4, pp. 697–708.

Ito, Takumi et al. (2010). "Identification of a primary target of thalidomide teratogenicity". In: *science* 327.5971, pp. 1345–1350.

Jentsch, Stefan (1992). "The ubiquitin-conjugation system". In: *Annual review of genetics* 26.1, pp. 179–207.

Johnson, Erica S, David K Gonda, and Alexander Varshavsky (1990). "Cis-trans recognition and subunit-specific degradation of short-lived proteins". In: *Nature* 346.6281, p. 287.

Krönke, Jan, Emma C Fink, et al. (2015). "Lenalidomide induces ubiquitination and degradation of CK1α in del (5q) MDS". In: *Nature* 523.7559, p. 183.

Krönke, Jan, Namrata D Udeshi, et al. (2014). "Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells". In: *Science* 343.6168, pp. 301–305.

Lu, Gang et al. (2014). "The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins". In: *Science* 343.6168, pp. 305–309.

Lu, Jing et al. (2015). "Hijacking the E3 ubiquitin ligase cereblon to efficiently target BRD4". In: *Chemistry & biology* 22.6, pp. 755–763.

Matyskiela, Mary E et al. (2016). "A novel cereblon modulator recruits GSPT1 to the CRL4 CRBN ubiquitin ligase". In: *Nature* 535.7611, p. 252.

Muller, George W et al. (1999). "Amino-substituted thalidomide analogs: potent inhibitors of TNF-α production". In: *Bioorganic & medicinal chemistry letters* 9.11, pp. 1625–1630.

Obrig, Tom G et al. (1971). "The mechanism by which cycloheximide and related glutarimide antibiotics inhibit peptide synthesis on reticulocyte ribosomes". In: *Journal of Biological Chemistry* 246.1, pp. 174–181.

Petroski, Matthew D and Raymond J Deshaies (2005). "Function and regulation of cullin–RING ubiquitin ligases". In: *Nature reviews Molecular cell biology* 6.1, p. 9.

Petzold, Georg, Eric S Fischer, and Nicolas H Thomä (2016). "Structural basis of lenalidomide-induced CK1α degradation by the CRL4 CRBN ubiquitin ligase". In: *Nature* 532.7597, p. 127.

Pickart, Cecile M and Michael J Eddins (2004). "Ubiquitin: structures, functions, mechanisms". In: *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* 1695.1-3, pp. 55–72.

Porras-Yakushi, Tanya R and Sonja Hess (2014). “Recent advances in defining the ubiquitylome”. In: *Expert review of proteomics* 11.4, pp. 477–490.

Sakamoto, Kathleen M, Kyung B Kim, Akiko Kumagai, et al. (2001). “Protacs: chimeric molecules that target proteins to the Skp1–Cullin–F box complex for ubiquitination and degradation”. In: *Proceedings of the National Academy of Sciences* 98.15, pp. 8554–8559.

Sakamoto, Kathleen M, Kyung B Kim, Rati Verma, et al. (2003). “Development of Protacs to target cancer-promoting proteins for ubiquitination and degradation”. In: *Molecular & Cellular Proteomics* 2.12, pp. 1350–1358.

Schneekloth, John S et al. (2004). “Chemical genetic control of protein levels: selective in vivo targeted degradation”. In: *Journal of the American Chemical Society* 126.12, pp. 3748–3754.

Schwanhäusser, Björn et al. (2011). “Global quantification of mammalian gene expression control”. In: *Nature* 473.7347, p. 337.

Sievers, Quinlan L et al. (2018). “Defining the human C2H2 zinc finger degrome targeted by thalidomide analogs through CRBN”. In: *Science* 362.6414, eaat0572.

Singhal, Seema et al. (1999). “Antitumor activity of thalidomide in refractory multiple myeloma”. In: *New England Journal of Medicine* 341.21, pp. 1565–1571.

Soucy, Teresa A et al. (2009). “An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer”. In: *Nature* 458.7239, p. 732.

Thrower, Julia S et al. (2000). “Recognition of the polyubiquitin proteolytic signal”. In: *The EMBO journal* 19.1, pp. 94–102.

Toyama, Brandon H et al. (2013). “Identification of long-lived proteins reveals exceptional stability of essential cellular structures”. In: *Cell* 154.5, pp. 971–982.

Ueda, Jun et al. (2006). “Zinc finger protein Wiz links G9a/GLP histone methyltransferases to the co-repressor molecule CtBP”. In: *Journal of Biological Chemistry* 281.29, pp. 20120–20128.

Uehara, Taisuke et al. (2017). “Selective degradation of splicing factor CAPERα by anticancer sulfonamides”. In: *Nature chemical biology* 13.6, p. 675.

Verma, Rati et al. (2001). “Selective degradation of ubiquitinated Sic1 by purified 26S proteasome yields active S phase cyclin-Cdk”. In: *Molecular cell* 8.2, pp. 439–448.

Wickliffe, Katherine E et al. (2011). “K11-linked ubiquitin chains as novel regulators of cell division”. In: *Trends in cell biology* 21.11, pp. 656–663.

Winter, Georg E et al. (2015). “Selective target protein degradation via phthalimide conjugation”. In: *Science (New York, NY)* 348.6241, p. 1376.