Modulation of the CRL4CRBN E3 Ubiquitin Ligase by Thalidomide Analogs

Modulation of the $CRL4^{CRBN}$ E3 ubiquitin ligase by thalidomide analogs

A dissertation presented

by

Quinlan Lloyd Sievers

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biological and Biomedical Sciences

Harvard University

Cambridge, Massachusetts

May 2017

Dissertation Advisor: Benjamin Ebert Quinlan Lloyd Sievers

Modulation of the CRL4^{CRBN} E3 ubiquitin ligase by thalidomide analogs

ABSTRACT

Thalidomide and its derivatives, lenalidomide and pomalidomide, are effective therapies for the hematopoietic malignancies multiple myeloma and del(5q) myelodysplastic syndrome. Their therapeutic properties are a result of their ability to induce the ubiquitination and proteasomal degradation of Ikaros (IKZF1), Aiolos (IKZF3), and Casein Kinase 1α (CK1α) by mediating the interaction of these proteins with Cereblon (CRBN), the substrate receptor for the CRL4^{CRBN} E3 ubiquitin ligase. The following thesis details two projects which leverage high-throughput screens to further our understanding of this unique mechanism of action. In chapter 2 we utilize a genome-scale CRISPR-Cas9 screen to identify the cellular machinery which is required for lenalidomide-induced CRL4^{CRBN} function and specifically focus on the cullin neddylation enzymes UBE2M and the COP9 signalosome, and the E2 ubiquitin-conjugating enzymes UBE2D3 and UBE2G1. In chapter 3 we utilize a combination of saturation mutagenesis, crystallography, and motif-based screening to characterize the C2H2 zinc finger degron in IKZF3 and identify novel targets of thalidomide analogs. In aggregate these studies provide evidence that therapeutic modulation of the CRL4^{CRBN} ubiquitin ligase by thalidomide analogs requires multiple factors beyond the ligase itself and that these compounds induce the degradation of a greater than previously appreciated number of proteins which are recognized via C2H2 zinc finger degrons. These findings inform not only our understanding of thalidomide analogs, but also the growing field of targeted protein degradation.

Table of Contents

Acknowledgements

I would like to thank my advisor Ben Ebert for his kindness and for the opportunity to work on such interesting projects, my friends for being a source of support and distraction, the Ebert lab for providing a great environment to learn science, Jessica for being a fantastic scientific collaborator and friend, Jenny for her scientific help and friendship, Georg and Nico for being gracious hosts in Basel, and my family for their support, especially my Mom.

Chapter 1: Introduction

Preface

Over the past two decades thalidomide and its derivatives, lenalidomide and pomalidomide, have proven to be effective therapies for the inflammatory disorder erythema nodosum leprosum and the hematologic neoplasms multiple myeloma and del(5q) myelodysplastic system. Despite their established clinical utility, it was not until recently that we understood the unique molecular mechanism from which these compounds derive their therapeutic properties: the ubiquitination and degradation of specific protein targets by the $CLR4^{CRBN}$ E3 ubiquitin ligase. But the clinical utility of thalidomide analogs and their novel mechanism of action belie the origin story of these compounds, and how the thalidomide scandal led to the reformation of drug safety standards around the world.

Thalidomide's history as a teratogen[1]

The 1950s-era advertisements heralded thalidomide as a wonder drug, a treatment for anxiety, sleeplessness, and nausea that above all was safe. Indeed, Chemie Grünenthal, the German pharmaceutical company which first marketed the drug, could find no evidence of toxicity in mice and rats despite using remarkably high doses of thalidomide (a fact that would later be attributed to species-specific differences in the molecular target of thalidomide). In an era when clinical trials and drug safety studies in humans were not mandated by the German government, Grünenthal introduced thalidomide to the public in October of 1957 without rigorous human testing.

Following the end of the second world war there was a growing demand for medications that could reduce anxiety and improve sleep. Many of these medications were barbiturates, a class

of drugs which could be lethal if taken in excess. Thalidomide—sold over-the-counter and accompanied by an aggressive advertising campaign centering on its safety—became a blockbuster drug. Marketed across Europe and Australia, thalidomide sold second only to aspirin in some countries. Under at least 37 different trade names, thalidomide was used for colds, coughs, flu, asthma, headaches, anxiety, nausea, and insomnia.

In 1960 the Food and Drug Administration (FDA) in the United States received an application for the approval of thalidomide from the United States distributor, William S. Merrell and Company. Merrell saw the application as formality and, assuming a quick approval, already had 10 million tablets of thalidomide ready for distribution. Indeed, similar to Europe and Australia, the laws at the time for drug approval in the United States were lax. The FDA did not require pharmaceutical companies to demonstrate efficacy, and while they did require data on safety, the FDA did not oversee or provide guidance on clinical trials. At the time of the application Merrell's "clinical trial" had consisted of distributing over two and a half million tablets of thalidomide to physicians in the United States without any obligation to consent patients or report outcomes data. Merrell's confidence in thalidomide's eventual approval also stemmed from a history of corruption, bribery, and kickbacks which was endemic within the FDA to the highest levels of its organization.

In this context, Merrell's application for thalidomide was placed on the desk Dr. Frances Oldham Kelsey, who was one week into her new position at the FDA. Kelsey was 47 at the time, and had come to the FDA with a PhD in pharmacology and a medical degree from the University of Chicago. Dr. Kelsey was troubled by a number of elements of the application, including the lack of safety data for the drug. She cited the application as "incomplete" and requested more information from Merrell. Over the ensuing months, she asked Merrell to resubmit its application

a half-dozen times and stood her ground under mounting pressure both from Merrell and her superiors at the FDA to approve the drug.

In the intervening time, troubling reports began to accumulate of two major side effects of thalidomide. The first, peripheral neuropathy, occurred in 5-20% of patients who had taken thalidomide for several months and was characterized by the irreversible loss of sensation in the hands or feet. The second, and more significant, was an increase in the number of miscarriages and children born with a wide range of birth defects to mothers who had taken thalidomide. These birth defects were wide-ranging, including organ damage to the heart, kidneys, genitals, or gastrointestinal tract, the absence or malformation of the eyes and the inner and outer ears, facial palsies or asymmetry, and the most iconic and frequent side effect, a foreshortening or complete absence of the upper limbs known as phocomelia or amelia, respectively[2].

Because Grünenthal and the other distributors of thalidomide had not organized a centralized data collection effort, the causal relationship between thalidomide use early in pregnancy and birth defects was difficult to make and was left up to physicians. Adding to this difficulty was the element of timing of drug exposure during pregnancy; today we know that fetal exposure to the drug between the 20^{th} and 36^{th} day of pregnancy was sufficient to cause miscarriage or birth defects, a window of time when many women would not have known they were pregnant. However, exposure after day 36 appeared not to harm the fetus, meaning that many women who had taken thalidomide while knowingly pregnant gave birth to healthy children. By 1961, when physicians in Germany and Australia began to present and publish definitive epidemiologic reports linking thalidomide to phocomelia [3,4], the drug had been in use for 4 years, and an estimated 8,000-12,000 children had been affected. Thalidomide, now infamous, was recalled worldwide.

In the United States, Dr. Kelsey's willingness to remain steadfast in her requirement that Merrell present adequate safety data successfully delayed the approval of thalidomide long enough for the drug's teratogenic potential to be revealed. She was awarded the President's Award for Distinguished Federal Civilian Service in 1962 by President John F. Kennedy. The thalidomide scandal served as a catalyst for the reformation of the FDA's drug approval and clinical trial standards; in 1962 the Senate unanimously passed the Kefauver-Harris drug amendments to the Food, Drug, and Cosmetic Act, which were later signed into law by President Kennedy. The law still stands today, dictating the requirements and safeguards for the modern drug approval process.

Clinical uses of thalidomide analogs

Today, thalidomide and its derivatives, lenalidomide and pomalidomide, can be found in a different set of advertisements, ones which contain a long list of side effects and warnings of the drugs' teratogenic potential. These new advertisements indicate the drugs' utility in treating three disorders: the infectious/inflammatory disease erythema nodosum leprosum (ENL), and the hematopoietic neoplasms multiple myeloma and del(5q) myelodysplastic syndrome (del(5q) MDS) **(Table 1)**.

Table 1 | Clinical uses of thalidomide, lenalidomide, and pomalidomide.

Category	Disorder	Pathogenesis	Clinical Features	Approved Medications
Infectious/ Inflammatory	Erythema Nodosum Leprosum (ENL)	Believed to be a result of immune complex deposition secondary to an infection with *mycobacterium leprae*	Painful, red nodules across skin of extremities and face	Thalidomide
Hematologic	Multiple Myeloma (MM)	Malignancy of post-germinal plasma cells	Lytic bone lesions, pathologic fractures, kidney failure, anemia	Thalidomide Lenalidomide Pomalidomide
	del(5q) Myelodysplastic Syndrome (5q- MDS)	Clonal expansion of a hematopoietic stem cell with impaired differentiation	Anemia, immunodeficiency, fatigue, thrombocytopenia	Lenalidomide

Erythema nodosum leprosum (ENL) is a complication of infection with the mycobacterium that causes leprosy and is characterized by the appearance of red, painful nodules across the skin of the extremities and face. While its pathogenesis is not completely understood, ENL is believed to be a result of an improper immune response, notable for increases in certain pro-inflammatory cytokines and the deposition of immune complexes in tissues (collections of antibodies and complement). Thalidomide's efficacy in this disorder was found when a patient who could not sleep due to the pain associated with these lesions was given thalidomide as a sedative[5]. The patient slept, and his disease went into remission. FDA approval for this indication was granted in 1998[6,7], the first approved clinical use after thalidomide was found to be a teratogen.

Thalidomide had previously been shown to inhibit the secretion of tumor necrosis factor alpha (TNFα) from peripheral blood mononuclear cells (PBMCs)[7]. Under the suspicion that this property explained thalidomide's therapeutic benefit in ENL, lenalidomide and pomalidomide were optimized in a TNFα inhibition assay in 1999 by Celgene, the company which manufactures the drugs today[8]. These more potent TNFα inhibitors incidentally also showed promise in treating the hematopoietic malignancies multiple myeloma and del(5q) myelodysplastic syndrome.

Multiple myeloma is a malignancy of post-germinal center plasma cells, the antibody-producing cells of the immune system. Symptoms associated with the disorder stem from the overproduction of antibodies by the malignant clone as well as its proliferation in the bone marrow; osteolytic bone lesions, bone pain, pathologic fractures, hypercalcemia, kidney disease, immunosuppression, and anemia are common features. Thalidomide was originally trialed in patients with multiple myeloma owing to a discovery in 1994 that the drug was an inhibitor of angiogenesis[9]. Indeed, multiple myeloma is associated with extensive

neovascularization of the bone marrow, the site where the tumor resides. Thalidomide resulted in a reduction in paraprotein levels (a measurement of antibody production as a proxy for tumor burden) in 78% of patients who were refractory to existing therapies[10]. Notably, however, the vascularization of the bone marrow remained unchanged. FDA approval for thalidomide was granted in 2006 for newly diagnosed multiply myeloma, followed by approval for lenalidomide in 2006 for myeloma with one prior therapy and again in 2015 for newly diagnosed myeloma. Lastly, pomalidomide was approved in 2013 for myeloma that had progressed after two prior therapies.

Myelodysplastic syndromes are caused by the indolent proliferation of a malignant hematopoietic stem cell with impaired differentiation potential. Patients present with fatigue, recurrent infections, and easy bruising due to suppressed production of the three major hematopoietic lineages: lymphocytes, red blood cells, and platelets. In rare cases, the disease can progress to acute myeloid leukemia (AML). Lenalidomide was shown to be effective in a subset of this disorder that is defined by the deletion of chromosome arm 5q, yielding transfusion independence in 67% of patients in phase III trials. Lenalidomide was FDA approved for this indication in 2005[11-13].

Intriguingly, despite their clear clinical utility, the molecular mechanism of action underlying the therapeutic properties of thalidomide, lenalidomide, and pomalidomide remained a mystery for over a decade after they entered the clinic.

Phenotypic effects of thalidomide analogs

This situation is not unique; drug discovery has relied on empiric testing for longer than it has used a target-based approach. As a consequence, many medications that we use in the clinic

today have poorly defined or unknown molecular mechanisms of action. An understanding of the mechanism of action of a drug can pay large dividends, facilitating the medicinal chemistry required to improve potency and selectivity, guiding the appropriation for other diseases, and providing insight into the molecular pathways that underlie disease pathology.

Ever since thalidomide was found to be a teratogen, scientists have sought to explain the molecular basis of its clinical properties. These early studies established thalidomide analogs as drugs with pleiotropic effects, eliciting changes at both the cellular and organismal level **(Table 2).** The most notable effects occurred at the level of the central and peripheral nervous system (sedation, induction of sleep, peripheral neuropathy, constipation), and the hematopoietic lineages (altered immune responses, remission of hematopoietic malignancies, increased clotting risk). The ability of the drugs to alter the activity of the immune system prompted Celgene, the manufacturer, to name thalidomide analogs “immunomodulatory drugs”, or “IMiDs”.

While this early work served as a formative body of information, it consisted of phenotypic descriptions and did not possess a unifying molecular mechanism. This was in part due to the lack of availability of newer technology that allowed for unbiased assessments of a drug’s molecular targets. In the past several years advances in the areas of mass-spectrometry and quantitative proteomics have made more incisive insights into thalidomide analogs’ molecular target.

Table 2 | Phenotypes elicited by thalidomide analogs

Organ System		Phenotype
Hematopoietic System	Cancer	Cell cycle arrest and apoptosis, remission of the malignant clone
	Bone Marrow stroma	dec. IL-6 secretion
	Immune System	Monocytes - ↓TNFa, ↓IL12 T cells - ↑IL2, ↑IFNg, ↑costimulation, ↑proliferation
	Vascular	Increased risk of thrombosis, anti-angiogenic
Central and Peripheral Nervous System	Central nervous system	Hypnotic, sedative, anti-emetic, anti-nausea
	Peripheral nervous system	Irreversible peripheral neuropathy, constipation
	Fetal Development	Phocomelia and amelia, facial nerve palsy, organ damage and malformations (eye, ears, kidney, heart, CNS)
Other		Free-radical oxidative DNA damage
		↓FGF8 in developing limb bud

The CRL4^{CRBN} E3 ubiquitin ligase is the molecular target of thalidomide analogs

Today we understand that thalidomide analogs' therapeutic effects rely on a unique molecular mechanism of action: they mediate the ubiquitination and proteasomal degradation of Ikaros (IKZF1), Aiolos (IKZF3), and Casein Kinase 1α (CK1α) by facilitating the interaction of these proteins with Cereblon (CRBN), the substrate receptor for the CRL4^{CRBN} E3 ubiquitin ligase[14-17] **(Figure 1).**

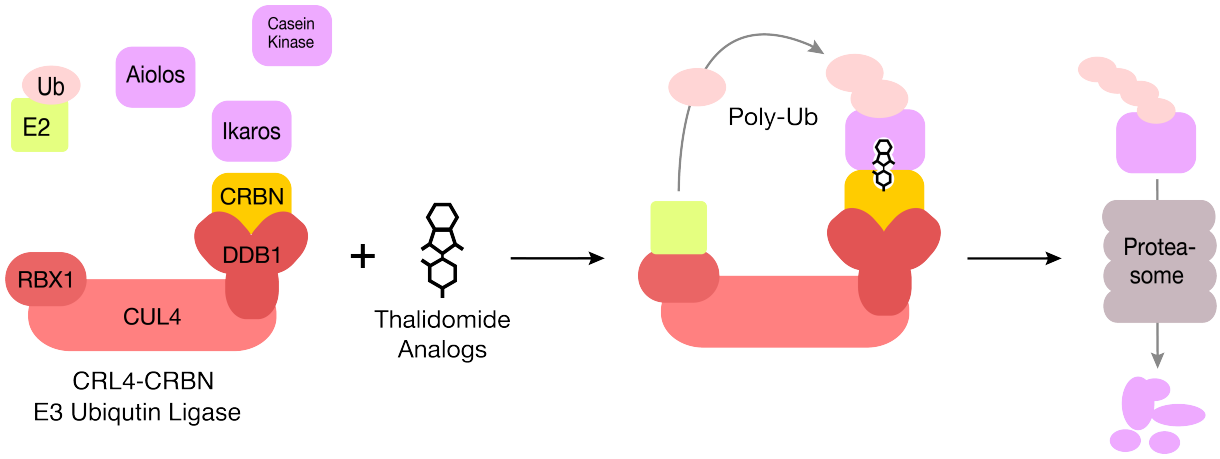


Figure 1 | Thalidomide analogs induce the ubiquitination and proteasomal degradation of protein targets via the CRL4^{CRBN} E3 ubiquitin ligase.

E3 ubiquitin ligases are a class of enzymes which regulate their protein substrates through ubiquitination, the covalent addition of the 8kD protein ubiquitin (Ub). E3 ligases come third in a cascade in which ubiquitin first forms a thioester bond with the catalytic cysteine on an E1 enzyme, then is passed to a cysteine on an E2 ubiquitin-conjugating enzyme, and then E3 ubiquitin ligases mediate the covalent transfer of ubiquitin to lysine residues on their target substrates either via an E3-Ub intermediate or by bringing the E2-Ub into proximity with the target substrate **(Figure 2)**. There is an increasing diversity and specificity of enzymes as you proceed in this cascade, with ~1-2 E1 enzymes, ~35 E2 enzymes, and an estimated ~600 E3 ubiquitin ligases encoded in the human genome[18]. Ubiquitination can have many effects on the target substrate depending on the number of ubiquitin moieties added and the lysine residues on ubiquitin which are used to build polyubiquitin chains. The canonical outcome of ubiquitination, polyubiquitination with lysine 48-linked ubiquitin, leads to rapid degradation of the protein by the 26s proteasome.

The first clue that pointed scientists towards the involvement of CRL4^{CRBN} in thalidomide's mechanism of action came in 2011 when ligand affinity chromatography and tandem mass

spectrometry was used to identify CRBN as the molecular target of thalidomide[14] **(Figure 3a)**. The authors additionally found DDB1 to associate with thalidomide in their study, a ubiquitin ligase component which more specifically implicated CRBN's role as the substrate receptor for the $CRL4^{CRBN}$ E3 ubiquitin ligase.

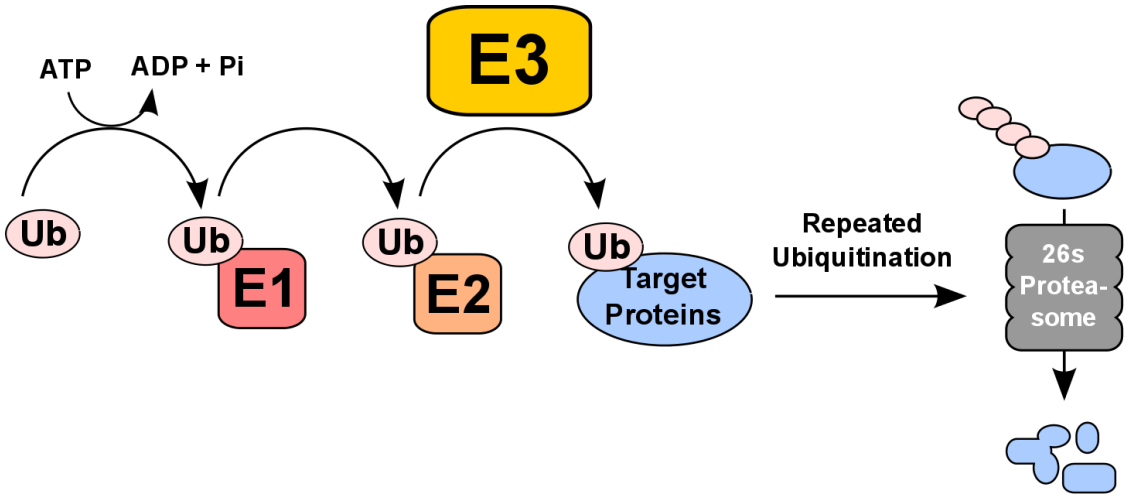


Figure 2 | The ubiquitin cascade.

$CRL4^{CRBN}$ belongs to the family of cullin-ring ligases (CRLs) which are defined by their multi-protein, modular architecture, and ability to facilitate donation of ubiquitin moieties directly from E2 enzymes onto target substrates[19]. $CRL4^{CRBN}$ specifically is comprised of four subunits which play distinct roles: CUL4 (scaffold), RBX1 (E2 binding), DDB1 (adaptor), and CRBN (substrate receptor). Upon recognition of a substrate by CRBN, the flexible cullin backbone brings an E2-Ub into proximity with the substrate, facilitating the covalent addition of the ubiquitin moiety to a lysine on the substrate.

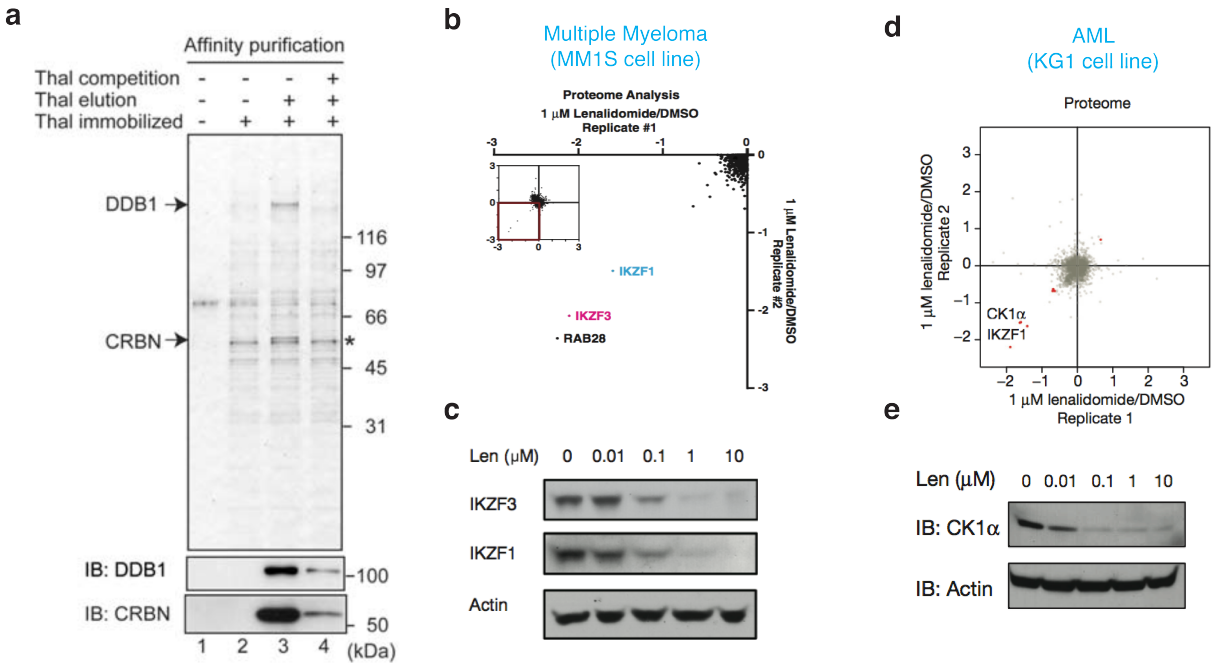


Figure 3 | Thalidomide analogs target CRBN, the substrate receptor for the CRL4^{CRBN} ubiquitin ligase, and induce the degradation of IKZF1, IKZF3, and CK1α. a, Thalidomide binding proteins were isolated from HeLa cell lysates using bead-immobilized thalidomide and identified using tandem mass-spectrometry. Asterisk indicates a non-specific band (excerpted from Ito et. al. 2010). **b,** MM1.S (multiple myeloma) were treated with DMSO or 1uM lenalidomide for 12h then protein lysates underwent mass spectrometry-based quantitative proteomic profiling (excerpted from Kronke et. al. 2014). **c,** Immunoblot of MM1.S cell lysates following 24h treatment with lenalidomide (excerpted from Kronke et. al. 2014). **d,** KG1 cells (AML) were treated with DMSO or 1uM lenalidomide for 12h then protein lysates underwent mass spectrometry-based quantitative proteomic profiling (excerpted from [20]). **e,** Immunoblot blot of KG1 cell lysate following 24h of treatment with lenalidomide (excerpted from Kronke et. al. 2015).

Thalidomide analogs induce the ubiquitination of IKZF1, IKZF3, and CK1α

Three years after the discovery that CRBN was the molecular target of thalidomide, studies from separate labs, respectively using quantitative mass-spectrometry and reporter-based ORF screens, found that treatment with lenalidomide resulted in the selective ubiquitination and proteasomal degradation of Ikaros (IKZF1), Aiolos (IKZF3), and Casein kinase 1α (CK1α)[15-17,20] **(Figure 3b-e)**.

These studies also found differences between the drugs; thalidomide is less potent for all three targets, pomalidomide is the most potent in degrading IKZF1 and IKZF3, and lenalidomide is the most potent degrader of CK1α.

IKZF1, IKZF3, and CK1α are therapeutic targets in multiple myeloma and del(5q) MDS

The degradation of IKZF1, IKZF3, and CK1α explains many of thalidomide analogs' therapeutic properties **(Table 3)**. IKZF1 and IKZF3 are members of the Ikaros family of transcription factors, a group of chromatin remodeling proteins with C2H2 zinc finger arrays which mediate sequence-specific DNA recognition and homo- and heterodimerization. Studies have established IKZF1 as a master regulator of lymphocyte specification and differentiation; IKZF1 tissue expression is limited to the lymphocyte lineage, and genetic inactivation of IKZF1 in mice results in the absence of T and B lymphocytes, as well as NK cells[21]. IKZF3 expression is limited to mature B cells, and its genetic inactivation in mice results in hyper-activated B cells, the formation of auto-antibodies, lymphoma, and the loss of B cells of the peritoneum, marginal zone, and bone marrow[22]. Additionally, expression of dominant negative isoforms of IKZF1 and IKZF3 are frequently observed in pediatric T- and B- acute lymphoblastic leukemia (ALL) in which the DNA binding region is lost and the dimerization domain is retained[23].

In keeping with their regulation of B-cell identity, IKZF1 and IKZF3 are essential genes in thalidomide analog-sensitive myeloma cell lines **(Figure 4a)**. Additionally, their degradation and the resulting alterations in transcription are thought to drive not only the anti-myeloma properties of thalidomide analogs, but also their immunomodulatory properties: IKZF1 and IKZF3 are negative regulators of IL2 transcription and increased IL2 secretion from T cells is thought to drive the therapeutic benefit in erythema nodosum leprosum **(Figure 4b)**.

CK1α is a serine-threonine kinase which negatively regulates the Wnt pathway via phosphorylation of Ser45 on beta catenin, a mark which in turn primes beta catenin for ubiquitination and degradation by the proteasome[24]. Beta catenin-driven transcription controls the proliferation of hematopoietic stem cells in a thresholded fashion; slight increases in beta catenin signaling results in the expansion of the hematopoietic stem cell compartment owing to an increased capacity for self-renewal, however beyond a certain range, beta catenin-driven transcription results in apoptosis of hematopoietic stem cells and bone marrow failure[25].

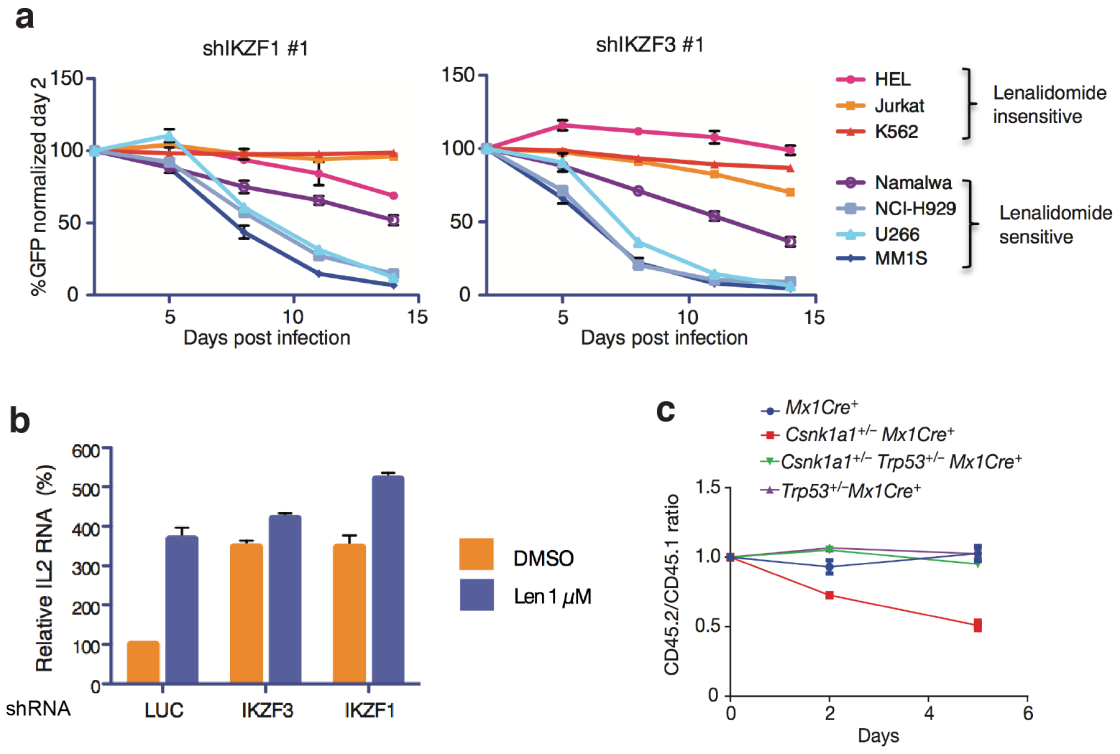


Figure 4 | IKZF1/IKZF3 and CK1α are therapeutic targets in multiple myeloma and myelodysplastic syndrome, respectively. a, Lenalidomide-sensitive and -insensitive myeloma cell lines were infected with shRNA expression constructs targeting either *IKZF1* or *IKZF3* and expressing EGFP. %EGFP expression was assessed via flow cytometry and normalized to day 2 post-infection (excerpted from Kronke et. al. 2014) **b,** T cells were infected with lentiviral vectors expressing shRNAs targeting LUC (control), *IKZF1*, or *IKZF3*, then stimulated for 12h with anti CD3/CD28 antibodies. mRNA was harvested from the cells and used for qPCR-based assessment of relative IL2 transcript levels (excerpted from Kronke et. al. 2014). **c,** CD45.2 mouse hematopoietic stem and progenitor cells with WT or +/- CK1α were transduced with I391V cereblon expression vectors and introduced in equal ratio with CD45.1 competitor cells into irradiated mice, which were then treated with lenalidomide (excerpted from Kronke et. al. 2015).

Table 3 | Therapeutic targets of thalidomide analogs

Degraded Target	Normal Function	Therapeutic Target in	Analog with greatest potency
Ikaros (IKZF1) and **Aiolos (IKZF3)**	Transcription factors required for proper lymphocyte differentiation and function	Multiple myeloma, Erythema nodosum leprosum	Pomalidomide
Casein Kinase 1 alpha	Kinase, regulator of Wnt pathway signalling and p53	del(5q) Myelodysplastic Syndrome	Lenalidomide

In keeping with this biology, CK1α resides on chromosome 5q and is lost in a heterozygous manner in del(5q) MDS. del(5q) MDS hematopoietic stem cells are therefore granted a competitive advantage due to a 50% reduction in CK1α levels and a concomitant increase in beta catenin-driven transcription. However this also provides a therapeutic window whereby del(5q) MDS cells are susceptible to the degradation of CK1α, which results in their apoptosis and repopulation of the bone marrow with normal HSCs[20,26] **(Figure 4c)**.

The discovery that thalidomide analogs could induce the degradation of IKZF1, IKZF3, and CK1α motivated work by structural biologists to understand how thalidomide analogs mediated CRBN's increased affinity for these proteins.

Thalidomide analogs bind at the CRBN-substrate interface

Thalidomide analogs are comprised of an identical glutarimide ring and a variable phthaloyl ring **(Figure 5a).** In 2014, a crystal structure of thalidomide bound to the CRBN-DDB1 complex revealed that these drugs bind to a pocket in CRBN's C-terminal substrate-interaction domain via the glutarimide ring while the phthaloyl ring remains solvent-exposed[27,28] **(Figure 5b-c)**.

Following the identification of CK1α as a thalidomide analog-induced substrate of $CRL4^{CRBN}$, a 2016 crystal structure of the CRBN-lenalidomide-CK1α ternary complex demonstrated that lenalidomide sits at the CK1α-CRBN interface, making contact with both CK1α and CRBN[29] **(Figure 5d)**. Specifically, CK1α interacts with the lenalidomide-CRBN complex via a β-hairpin**.**

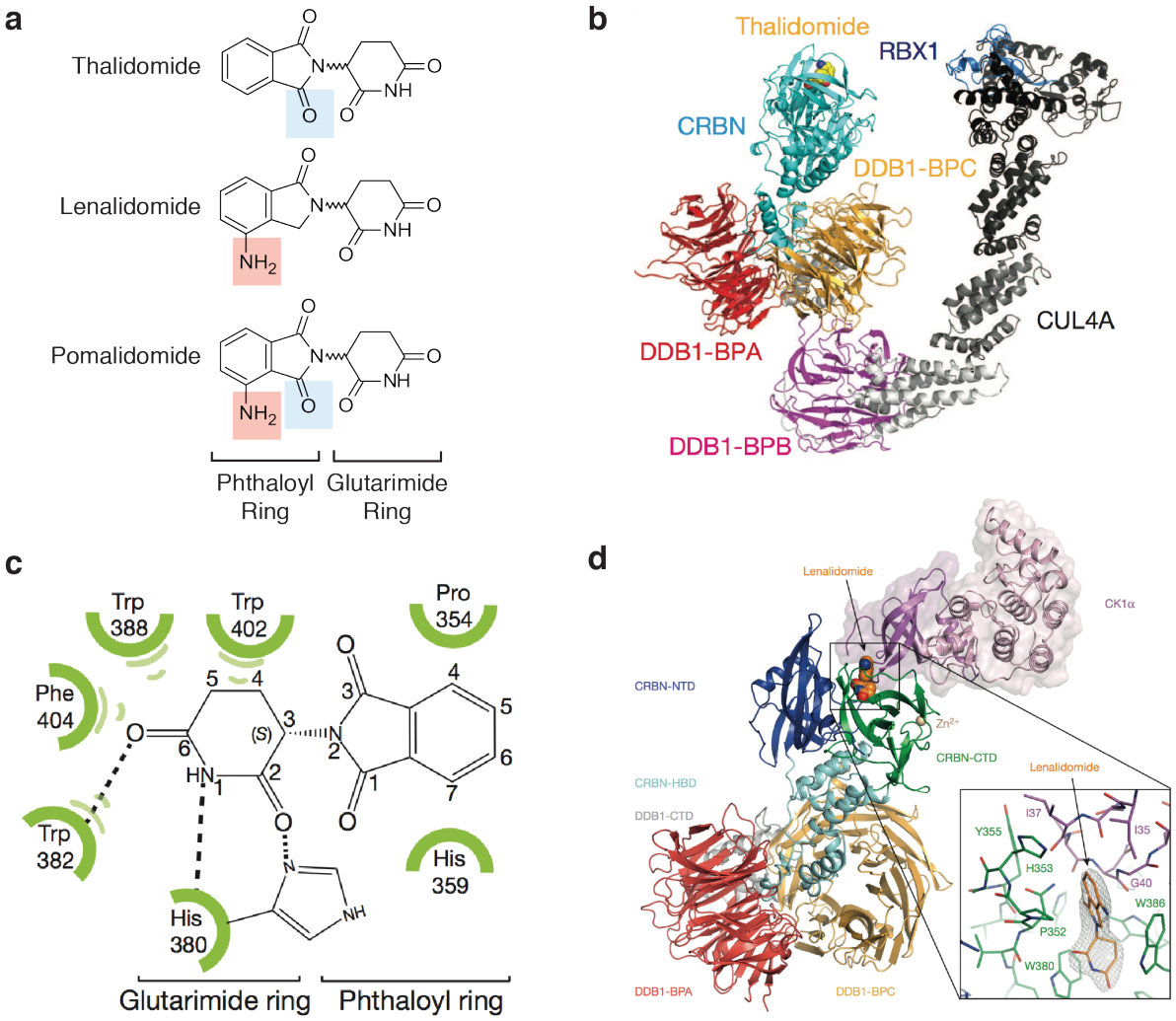


Figure 5 | Thalidomide analogs bind to a tri-tryptophan pocket on CRBN via their glutarimide ring and contact substrates via the phthaloyl ring. a, Structure of thalidomide, lenalidomide, and pomalidomide. **b,** Modeled structure of the $CRL4^{CRBN}$ ubiquitin ligase with thalidomide bound to the CRBN's C-terminal domain (Excerpted from Fischer et. al. 2014) **c,** Contacts between thalidomide's glutarimide ring and the tri-tryptophan binding pocket of CRBN (Excerpted from Fischer et. al. 2014). **d,** Crystal structure of casein kinase bound to the lenalidomide-CRBN-DDB1 complex (Excerpted from Petzold et. al. 2016).

The CK1α-lenalidomide-CRBN structure also offered an explanation for a long-standing question: why did mice and rats not exhibit the teratogenic effects of thalidomide? The answer is that there are several amino acid differences in mouse and human CRBN; one of these mutations, V391I , results in a steric clash with thalidomide-induced substrates that precludes ubiquitination and degradation in the context of mouse crbn. In keeping with this observation, thalidomide analogs do not degrade IKZF1, IKZF3, or CK1α in mouse cells, a feature which can be rescued by expressing the human variant of CRBN. However, to this date it is still unclear which substrates of $CRL4^{CRBN}$, known or unknown, are responsible for thalidomide analogs' teratogenicity.

The structure of CK1α bound to the lenalidomide-CRBN complex additionally explained why lenalidomide is a more potent degrader of CK1α than pomalidomide by pinpointing that the absence of the carbonyl on the phthaloyl ring of lenalidomide grants it the steric flexibility required to accommodate CK1α binding.

"Undruggable" targets in cancer

The discovery of the mechanism underlying thalidomide analogs' therapeutic properties comes at a critical time in our understanding of the molecular pathways which sustain cancers. Advances in high-throughput sequencing and genome-scale functional genetic screens have improved our ability to pinpoint the biological processes which drive malignancies, discoveries which in turn have allowed scientists to identify targets whose inactivation could impair oncogenic signaling and transcriptional programs. However, many of these therapeutic targets possess structural properties that render them "undruggable" by conventional means.

In fact, the “druggable” genome is quite small (estimated to be ~3.5-25% of gene products) and is populated mainly by proteins with discrete catalytic sites amenable to small-molecule binding, including ion channels, nuclear receptors, GPCRs, and enzymes[30]. The prototypical example of “druggable” targets is the kinase ABL, whose aberrant activation as a result of fusion to *BCR* is responsible for the majority of chronic myeloid leukemia (CML) [31] [32,33]. Medicinal chemists have become adept at designing small molecules capable of competitively binding to the ATP binding pocket of ABL, and as of today ABL kinase inhibitors have nearly rendered CML a chronic disease[34].

In contrast, the majority of candidate disease-modifying targets are proteins whose effector functions are carried out not by discrete enzymatic domains but rather via protein-protein interactions. Protein-protein interactions occur across comparatively vast (~1,500–3,000 $Å^2$) surface areas lacking viable small molecule binding pockets [35]. The transcription factor MYC is a prominent example: Overexpressed or amplified in many cancers[36], MYC is an experimentally validated therapeutic target however its effects on transcription are carried out via protein-protein interactions and therefore lacks a conventionally targetable catalytic site.

The growing imbalance between our ability to nominate candidate therapeutic targets and our ability to engineer small molecules which can manipulate these proteins has created a distinct need for novel approaches to small molecule design and mechanism of action. It is herein that the interest in thalidomide analogs lies; they establish the precedent that small molecules can induce ubiquitin-ligase mediated degradation of a select group of proteins, including the “undruggable” transcription factors IKZF1 and IKZF3. However, the strategy of targeted protein degradation existed prior to the discovery of thalidomide analogs’ mechanism of action, and an understanding of its history can help us contextualize the importance of thalidomide and its derivatives.

History of targeted protein degradation as a therapeutic strategy

The ability of ubiquitin ligases to selectively ubiquitinate and degrade protein targets has led scientists to propose that it would be an ideal system through which difficult-to-drug proteins could be "inhibited" via degradation. The first generation of chemical tools used to selectively target proteins for ubiquitination and degradation by an E3 ubiquitin ligase were developed in the early 2000s and consisted of chimeric molecules in which an E3 ligase-targeting peptide was conjugated to a small molecule with specificity for a protein target [37-39]. Referred to as Proteoloysis Targeting Chimeras (PROTACs), these bifunctional compounds were successful in harnessing the $CUL1^{\beta\text{-}TRCP}$ and von Hippel-Lindau ubiquitin ligases to degrade a number of protein targets, however the use of a peptide in the construct resulted in a number of limitations, namely reduced cell penetration and stability.

The discovery of small molecule ligands with binding specificity for the E3 ligase MDM2 led to a second generation of chimeric compounds consisting of two small molecules (E3 binder and target binder) adjoined by a chemical linker[40]. Owing to the absence of a peptide component these PROTACs exhibited improved cell penetration and stability, however, their size and the specific requirements of the chemical linker used to join the two small molecules (flexibility, lack of reactivity, preservation of the function of the pharmacophores) still continued to limit the solubility, cell penetration, stability, and bioavailability of these compounds.

Given the precedent set by PROTACs then, the novelty of thalidomide analogs more specifically is two-fold. First, they provide a small molecule through which CRL^{CRBN}-targeting PROTACs can be made, and indeed PROTACs consisting of thalidomide analogs linked to the BRD4 inhibitor JQ1 have successfully been used to degrade the therapeutic target BRD4 [41-43]. Second, and

more importantly, thalidomide analogs represent the pinnacle of what PROTACs hope to achieve: thalidomide analogs are low molecular weight compounds with drug-like properties (cell penetration and oral bioavailability) which can mediate stable, tertiary interfaces between a drug, substrate receptor, and “undruggable” protein targets. Indeed, the discovery of the mechanism of action of thalidomide analogs validates the principles of PROTACs and holds the potential to accelerate the discovery of other small molecules with similar abilities.

Scope of the thesis

The goal of my thesis work has been to deepen our understanding of the mechanism of action underlying thalidomide analogs’ therapeutic effects, with the dual intentions of informing the biology of $CRL4^{CRBN}$ pharmacologic modulation, but also the discovering rules, mechanisms, techniques, and experimental approaches which may serve as a foundation on which novel compounds utilizing this therapeutic paradigm maybe be discovered and investigated.

The following chapters detail two projects which share in common an attempt to leverage high throughput screens to provide novel insights into the molecular mechanism of action of these drugs **(Figure 6)**.

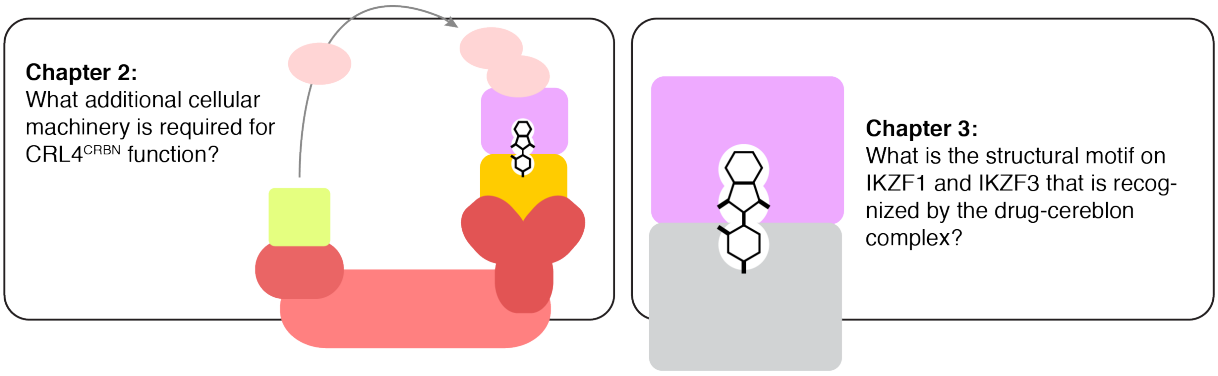


Figure 6 | Questions addressed by the thesis.

Chapter 2 discusses the results and follow-up of a genome-scale CRISPR-Cas9 screen in a lenalidomide-sensitive cell line which identified genes that regulate and are required for lenalidomide-mediated $CRL4^{CRBN}$ degradation of IKZF3. Specifically, we identify the cullin neddylation enzymes UBE2M and the COP9 signalosome, as well as the E2 enzymes UBE2D3 and UBE2G1 which cooperate to yield K48-linked polyubiquitination of lenalidomide induced $CRL4^{CRBN}$ substrates.

In chapter 3 we utilize a combination of saturation mutagenesis, crystallography and motif-based screening to define the second C2H2 zinc finger in IKZF3 as the drug-induced degron, and in turn identify additional C2H2 zinc finger degron-containing targets of thalidomide analogs.

In aggregate, these results improve our understanding of CRL regulatory biology as well as our grasp of the motifs on transcription factors targeted by ubiquitin ligases.

Chapter 2: Genome-wide screen identifies cullin-ring ligase neddylation machinery and E2 enzymes required for lenalidomide-mediated degradation of IKZF3 by the CRL4^{CRBN} ubiquitin ligase

Quinlan L. Sievers*[1,2,3], Jessica A. Gasser*[1,2], Glenn S. Cowley[1,a], Eric S. Fischer[5], Benjamin L. Ebert[1,2,4]

*Authors contributed equally

[1]Broad Institute of Harvard and MIT, Cambridge, MA
[2]Division of Hematology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA
[3]Harvard Medical School, MD/PhD program, Boston, MA
[4]Dana Farber Cancer Institute, Department of Medical Oncology, Boston, MA, USA
[5]Dana Farber Cancer Institute, Department of Chemical Biology, Boston, MA, USA
[a]Present Address: Discovery Science, Janssen Research and Development (Johnson & Johnson), Spring House, PA 19477

Corresponding Author:
Benjamin L. Ebert, MD/PhD
77 Avenue Louis Pasteur, HIM 743
Boston, MA 02115
bebert@partners.org

CONTRIBUTIONS

Quinlan Sievers and Benjamin Ebert worked collaboratively to form the concept and design of the genome-scale CRISPR-Cas9 screen, with guidance from Glenn Cowley. Quinlan Sievers carried out the genome scale screen and IKZF3 reporter screen with help from Glenn Cowley. Jessica Gasser and Quinlan Sievers worked on the validation experiments for the neddylation machinery and E2 enzymes, with guidance and essential reagents provided by Eric Fischer. Quinlan Sievers drafted the manuscript. All work was conducted under the direction of Benjamin Ebert.

ACKNOWLEDGEMENTS

Thank you to Jenny Chen of the Regev lab for helpful advice on the statistical analysis of the screens, Mudra Hegde and John Doench of the genetic perturbations platform for help with analysis and execution of the E2 library screen, and members of the Ebert lab for comments and suggestions throughout the course of the project.

ABSTRACT

Lenalidomide mediates the ubiquitination and degradation of Ikaros (IKZF1), Aiolos (IKZF3), and casein kinase 1α (CK1α) by facilitating their interaction with cereblon (CRBN), the substrate receptor for the $CRL4^{CRBN}$ E3 ubiquitin ligase. To identify the cellular machinery which regulates and is required for lenalidomide-induced $CRL4^{CRBN}$ activity we performed a positive selection, genome-scale CRISPR-Cas9 screen in a lenalidomide-sensitive myeloma cell line. *CRBN* was the top-ranking gene, with all 6 *CRBN*-targeting gRNAs ranking as the top 6 gRNAs in the genome-scale library. To determine which of the top 30 ranking genes are required for IKZF3 degradation, we performed a counterscreen using an IKZF3-EGFP reporter and identified the cullin ring ligase (CRL) neddylation enzymes UBE2M and the COP9 signalosome, as well as the E2-ubiquitin conjugating enzymes UBE2D3 and UBE2G1. We demonstrate that loss of UBE2M or members of the COP9 signalosome results in hypo- or hyper-neddylation of CUL4 respectively, and that this in turn influences $CRL4^{CRBN}$ activity and cereblon protein levels. Additionally, we establish UBE2G1 and UBE2D3 as requirements for K48-linked polyubiquitination of $CRL4^{CRBN}$ target substrates, and specifically define UBE2D3 as a monoubiquitinating “priming” E2 and UBE2G1 as a lysine-48 linked “chain-extending” E2. In aggregate, our identification of cullin-ring ligase regulators and E2 enzymes required for $CRL4^{CRBN}$ function extends our understanding of the molecular machinery which plays a critical role in shaping the therapeutic effects of lenalidomide.

INTRODUCTION

Over the past two decades lenalidomide has established itself as an effective therapy for the hematopoietic malignancies multiple myeloma and del(5q) myelodysplastic syndrome (MDS)[13,44,45]. Despite its clear clinical utility in these contexts, it was not until recently that we appreciated the molecular mechanism through which lenalidomide exerts its therapeutic effects: modulation of the $CRL4^{CRBN}$ E3 ubiquitin ligase.

E3 ubiquitin ligases are a large class of enzymes which play an important role in regulating the post-translational stability of their protein substrates via ubiquitination, the covalent addition of the 8kD protein ubiquitin. Specifically, upon recognition of a substrate, E3 ligases repeatedly mediate the transfer of ubiquitin from E2 ubiquitin-conjugating enzymes onto their target, forming polyubiquitin chains. This post-translational mark in turn leads to recognition and degradation by the 26s proteasome.

The $CRL4^{CRBN}$ E3 ubiquitin ligase belongs to the family of cullin-ring ligases (CRLs), which are defined by their multi-subunit, modular architecture based around a cullin backbone[19]. Via its glutarimide ring, lenalidomide binds to a tri-tryptophan binding pocket on the C-terminal domain of cereblon (CRBN), the $CRL4^{CRBN}$ subunit responsible for recognition of target substrates[27,28]. Lenalidomide's exposed phthaloyl ring increases CRBN's affinity for the lymphocyte lineage transcription factors Ikaros (IKZF1) and Aiolos (IKZF3), and the Wnt pathway regulator Casein Kinase 1α (CK1α). As a result of their stabilized interaction with $CRL4^{CRBN}$, IKZF1, IKZF3, and CK1α are polyubiquitinated and degraded. Degradation of IKZF1 and IKZF3 underlies the anti-myeloma properties of lenalidomide, while degradation of casein kinase explains the therapeutic benefit in the context of del(5q) MDS[15-17,20].

Ubiquitin ligases are extensively regulated, however our knowledge of the molecular and cellular machinery that is required for lenalidomide-mediated degradation of $CRL4^{CRBN}$ targets is largely limited to $CRL4^{CRBN}$ itself. To elaborate on our understanding of the factors required for lenalidomide's therapeutic modulation of $CRL4^{CRBN}$ we have performed a positive selection, genome-scale CRISPR-Cas9 screen in a lenalidomide-sensitive myeloma cell line. *CRBN* was the top-ranking gene, and an IKZF3-EGFP counterscreen of the top 30 genes identified the regulators of CRL neddylation UBE2M and the COP9 signalosome as well as the ubiquitin donating E2 enzymes UBE2D3 and UBE2G1 as necessary for efficient IKZF3 degradation. We demonstrate that loss of UBE2M or the COP9 signalosome alters $CRL4^{CRBN}$ function via hypo- or hyperneddylation of CUL4, and we also establish that the ubiquitin donors UBE2D3 and UBE2G1 play distinct "priming" and "chain-extending" roles, respectively, in order to yield productive K48-linked polyubiquitination of lenalidomide-induced targets of $CRL4^{CRBN}$.

These findings extend our knowledge of the factors required for lenalidomide's therapeutic modulation of $CRL4^{CRBN}$ function and holds implications for what is likely to be a growing class of small molecules targeting ubiquitin ligase-substrate interactions.

RESULTS

Genome-wide CRISPR-Cas9 screen in lenalidomide-sensitive cell line

Positive selection CRISPR-Cas9 screens have proven to be a robust tool for the unbiased identification of genes involved in specific biological pathways[46]. To gain insight into the molecular machinery which is required for lenalidomide-mediated modulation of $CRL4^{CRBN}$ function we therefore performed a positive selection, genome-scale CRISPR-Cas9 screen in the lenalidomide sensitive myeloma cell line, MM1.S.

MM1.S Cas9-expressing cells were infected with the human GeCKOv2 lentiguide-puro gRNA library and eight days after infection, the cells were grown in the presence of either DMSO or lenalidomide for 20 days **(Appendix Figure 16)**. DMSO-treated cells continued to proliferate while the 1uM lenalidomide-treated cells depleted to ~50% of their original cell number by day 20 **(Figure 7a)**. At the conclusion of the screen, the gRNA library was PCR amplified from genomic DNA isolated from the cells, and we then used next-generation sequencing to calculate the fold-change in the frequency of read counts for a given gRNA in the DMSO and lenalidomide experimental conditions. Gene targeting-gRNAs possessed a wider distribution of average fold-changes and an increase in the number of outliers in comparison to the 1,000 control gRNAs in the library **(Figure 7b)**.

CRBN and CRL regulators are the top-ranking genes

Of the ~120,000 gRNAs in the library, all six gRNAs targeting *CRBN*, the substrate receptor of the $CRL4^{CRBN}$ ubiquitin ligase, ranked as the top six most-enriched gRNAs in the lenalidomide replicates at day 20 **(Figure 7c, Appendix Figure 17a-b)**. To identify genes with multiple gRNAs demonstrating enrichment in the presence of lenalidomide we used the STARS

algorithm, which employs a probability mass function based on a binomial distribution to rank genes on the basis of the least probable ranking of their respective gRNAs in the screen **(Figure 7d, Appendix Figure 17c-d)**. Cereblon was the top-ranking gene and was followed by 29 genes exhibiting false discovery rates (FDRs) <0.05.

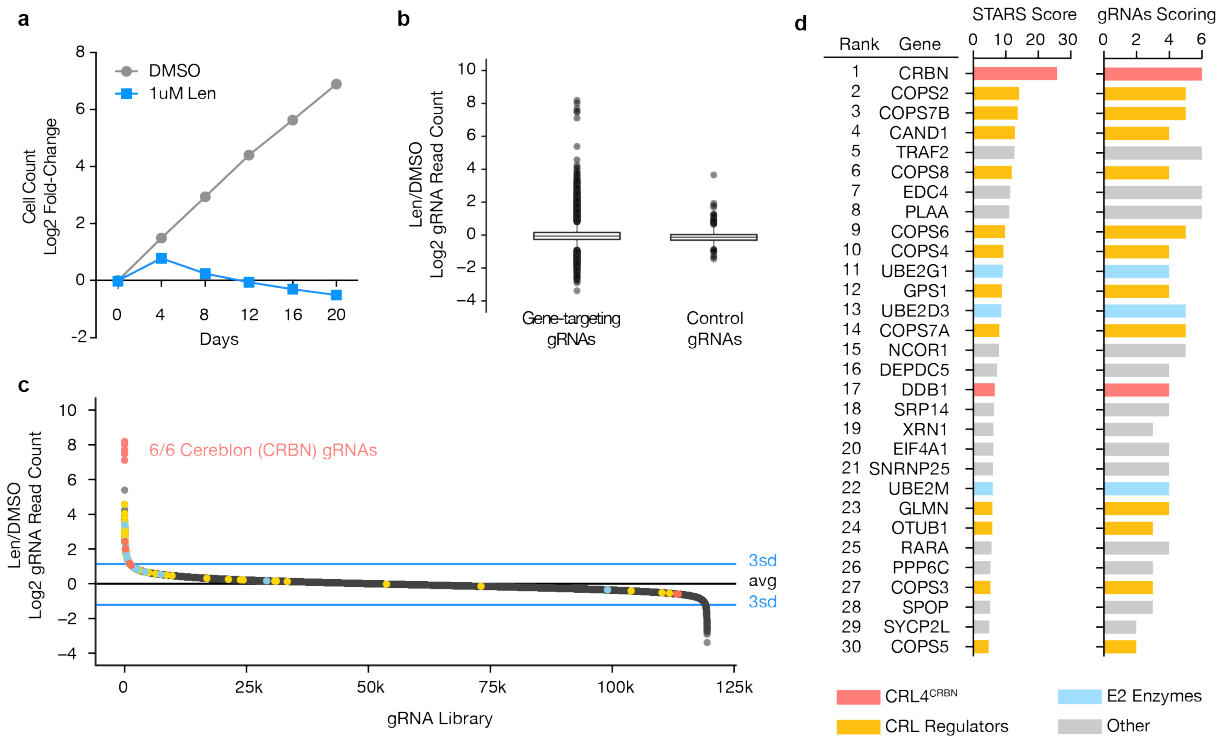


Figure 7 | Genome-scale CRISPR-Cas9 screen in the lenalidomide-sensitive myeloma cell line, MM1.S. a, Cell number over the 20-day assay (DMSO; 1 technical replicate, 1uM Len; average of three technical replicates). **b,** Normalized log2 fold enrichments (Len/DMSO) for gene-targeting gRNAs and 1,000 control gRNAs. **c,** Ranking of gRNAs according to their respective log2 Len/DMSO fold-change in read count (datapoints are average of 3 technical replicates). **d,** Top 30 ranking genes according to STARS algorithm, and number of gRNAs considered to have scored.

17 of the 30 highest-ranking genes could be sub-divided into three categories; members of the CRL4^{CRBN} ubiquitin ligase (*CRBN*, *DDB1*), cullin-ring ligase regulators (9 members of the COP9 signalosome, *CAND1*, *UBE2M*, *GLMN*, *OTUB1*), and E2 ubiquitin-conjugating enzymes

(*UBE2G1*, *UBE2D3*). The remaining 13 genes, which we grouped into a category labeled "other", included members of the NFKB pathway (*TRAF2*), the 5' mRNA decapping complex (*XRN1*, *EDC4*), nuclear hormone receptor signaling (*NCOR1*, *RARA*), MTOR signaling (*DEPDC5*) and genes identified as tumor-suppressors in other malignancies (*PPP6C*, *SPOP*).

Counterscreen identifies genes whose inactivation impairs IKZF3 reporter degradation

We next sought to identify genes that scored in the genome-wide screen due to their regulation of $CRL4^{CRBN}$. Towards this end MM1.S and HEK293T cells expressing an IKZF3 aa130-189-EGFP reporter **(Figure 8a)** were lentivirally transduced with a pooled library of ~117 gRNAs targeting the top 30 genes identified in the genome scale screen, with 3 gRNAs per gene and 12 control gRNAs. Twelve days after infection the cells were treated with DMSO or lenalidomide for 20 hours, after which the lenalidomide-treated samples underwent fluorescence activated cell sorting (FACS) to isolate cells which remained EGFP+ despite treatment with lenalidomide. At the conclusion of the assay, we PCR amplified the gRNAs from genomic DNA isolated from the cells and used next-generation sequencing to calculate the fold-change in the frequency of read counts for a given gRNA in the DMSO and the lenalidomide EGFP+ experimental conditions **(Figure 8b, Appendix Figure 18)**.

gRNAs targeting cereblon exhibited the greatest enrichment in the lenalidomide-treated EGFP+ group in both MM1S and HEK293T cells. The remaining genes segregated into two groups, with the majority of $CRL4^{CRBN}$ members, CRL regulatory factors, and E2 enzymes demonstrating an enrichment in gRNA frequency consistent with impaired degradation of the reporter, while the genes classified in the “other” category performed similar to the control gRNAs **(Figure 8c)**. The results were largely consistent across MM1S and HEK239T cells despite their different tissue of origins; exceptions included RARA, EDC4, and the functionally redundant COPS7A/B isoforms.

The necessity of the neddylation cycle enzymes UBE2M, members of the COP9 signalosome, and CAND1 for lenalidomide-induced IKZF3-EGFP reporter degradation implicates the dynamic regulation of CRL complex assembly and disassembly in lenalidomide’s therapeutic mechanism of action **(Figure 8d)**. CRLs rely on a cycle of neddylation to regulate CRL complex formation and dissociation, a process which allows this family of modular E3 ligases to dynamically

respond to shifting demands for a given E3. The neddylation cycle and its effects have been established largely in the context of the Skip-cullin1-fbox (SCF) cullin ring ligases[47].

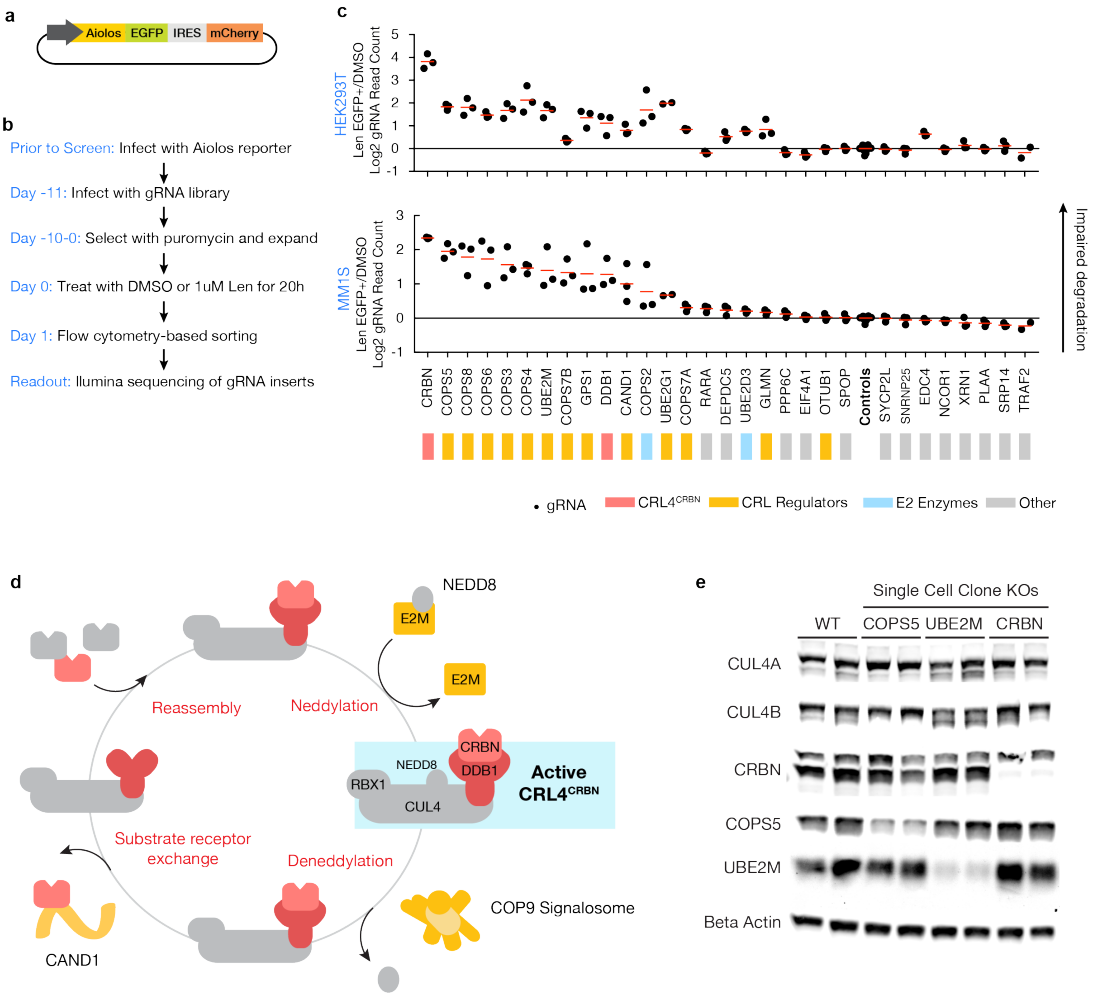


Figure 8 | Counterscreen of top 30 hits using an IKZF3-EGFP reporter identifies genes required for lenalidomide-induced degradation of IKZF3. a, Schematic of the IKZF3 aa130-189 reporter vector. **b,** Experimental overview of the counterscreen. **c,** Genes from the reporter screen ranked according to the normalized average fold-change in the log2-transformed gRNA sequencing read counts (Len-treated EGFP+/DMSO). Each point represents an individual gRNA, and each point is the average of three infection replicates. Red lines represent the average of three gRNAs. **d,** Schematic of cullin-ring ligase neddylation cycle. **e,** HEK293T single cell KO clones were generated via CRISPR-Cas9 introduction of frame-shifting edits. Lysates were harvested from clones and immunobloted for neddylation cycle proteins (representative of three experimental replicates).

We next sought to understand how *UBE2M* and *COPS5* inactivation affected CUL4 neddylation status and CRBN levels. To do so we used CRISPR-Cas9 mediated introduction of frameshifting insertions and deletions to generate UBE2M+/- and COPS5+/- HEK293T single cell clones and immunoblotted for CUL4A, CUL4B, and CRBN **(Figure 8e)**. UBE2M+/- cells demonstrated an increase in the lower molecular weight band of CUL4A and CUL4B, consistent with hyponeddylation of CUL4, which would in turn be expected to render a greater percentage of the CUL4-containing CRLs inactive [48]. Conversely, COPS5+/- cells exhibited an increase in the higher molecular weight band of CUL4A and CUL4B consistent with hyperneddylation, which in turn would be expected to render a greater percentage of CUL4-containing CRLs active. As observed in the context of SCF ligases, hyperneddylation was associated with a concomitant reduction in CRBN levels in one of the clones[49,50].

The E2 enzyme UBE2G1 is required for lenalidomide-mediated degradation of IKZF1, IKZF3, and CK1α

Despite the important role that E2 ubiquitin-conjugating enzymes play in ubiquitin signaling, the micromolar affinity of the E2-E3 interaction has made it difficult to identify which of the ~35 E2 enzymes are utilized by a given E3 ligase[51]. Our initial genome-scale screen and subsequent IKZF3 reporter assay identified UBE2D3 and UBE2G1 as E2 enzymes which are required for lenalidomide's anti-myeloma effects as well as its ability to degrade the IKZF3-EGFP reporter.

Confirming their necessity for lenalidomide's anti-myeloma effects, MM1.S cells infected with gRNAs targeting either *UBE2D3* or *UBE2G1* exhibited a competitive advantage over control gRNA-infected cells in the presence of 1uM lenalidomide **(Figure 9a)**.

To determine the effect of UBE2D3 and UBE2G1 loss in a genotypically homogenous population, we used CRISPR-Cas9 mediated introduction of frameshifting insertions and deletions to generate *UBE2D3*-/- and *UBE2G1*-/- HEK293T single cell clones expressing the IKZF3 aa130-189-EGFP reporter. These cells were then treated with a titration of lenalidomide for 20 hours, and the EGFP/mCherry ratio was assayed using flow cytometry **(Figure 9b)**. *UBE2D3*-/- clones exhibited a minor decrease in their sensitivity to lenalidomide. This phenotype was more prominent in the *UBE2G1*-/- clones, and furthermore, HEK293T cells rendered genetically null for both *UBE2D3* and *UBE2G1* demonstrated the greatest reduction in sensitivity to lenalidomide.

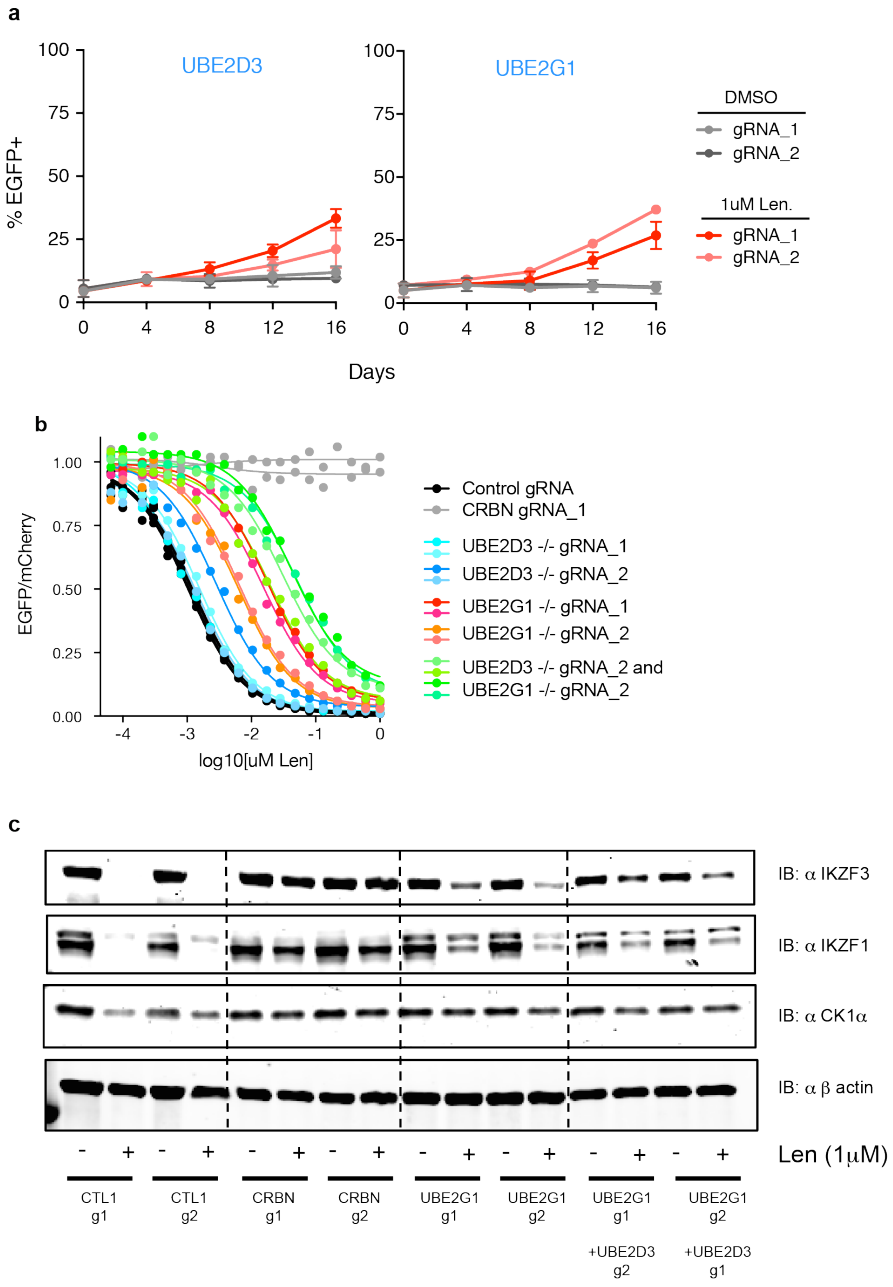


Figure 9 | UBE2G1 is required for lenalidomide-induced degradation of IKZF1, IKZF3, and CK1α. a, MM1.S were infected with vectors expressing EGFP and gRNAs targeting *UBE2D3* or *UBE2G1* and then mixed in a 5:95 ratio with control gRNA infected cells. These mixed populations were then grown in the presence of DMSO or 1uM lenalidomide for 16 days. **b,** HEK293T single cell clones null for either *UBE2D3*, *UBE2G1*, or *UBE2D3* and *UBE2G1* were infected with the IKZF3 aa130-189 EGFP reporter construct and then treated with a titration of lenalidomide for 20h, after which we assayed the EGFP/mCherry ratio via flow cytometry. **c,** MM1.S cells were infected with gRNAs targeting CRBN, UBE2G1, or both UBE2G1 and UBE2D3, treated with 1uM lenalidomide for 20 hours, then protein lysates were harvested and immunoblotted for IKZF1, IKZF3, and CK1α.

Having established that UBE2G1 in particular is required for degradation of the IKZF3 reporter, we asked if UBE2G1 is required for lenalidomide-induced degradation of endogenous IKZF1, IKZF3, and CK1α. MM1.S cells infected with gRNAs targeting *CRBN*, *UBE2D3*, and *UBE2G1* were treated for 20 hours with lenalidomide and then immunoblotted for IKZF1, IKZF3, and CK1α. gRNAs targeting *UBE2G1* resulted in a partial stabilization of IKZF1, IKZF3, and CK1α that was accentuated in the context of simultaneous infection with gRNAs targeting *UBE2D3* and *UBE2G1* **(Figure 9c)**.

UBE2D3 is a monoubiquitinating E2 which primes K48-linked polyubiquitination by UBE2G1

Having demonstrated the necessity of UBE2G1 for efficient degradation of $CRL4^{CRBN}$ target substrates in the presence of lenalidomide, we next sought to use in vitro ubiquitination studies to understand what the individual roles of UBE2D3 and UBE2G1 were in the process of the ubiquitination of IKZF3.

To this end we performed *in vitro* ubiquitination reactions combining recombinant E1, E2, and $CRL4^{CRBN}$ with IKZF3 aa130-189 which had been immunoprecipitated from HEK239T cells **(Figure 10a)**. In the presence of UBE2D3 alone we observed the appearance of prominent, higher molecular weight bands consistent with mono- or multi-mono ubiquitination of the IKZF3 substrate, while in the presence of UBE2G1 alone we failed to observe any change. However, in combination, UBE2D3 and UBE2G1 resulted in “laddering” of higher molecular weight forms of the substrate, consistent with polyubiquitination. The laddering pattern was abrogated with the addition of either a K0 or K48 ubiquitin mutant or in the absence of lenalidomide. These results suggest that UBE2D3 is responsible for “priming” the substrate via monoubiquitination, and that UBE2G1 requires this priming in order to polyubiquitinate a substrate with lys48 linked ubiquitin chains.

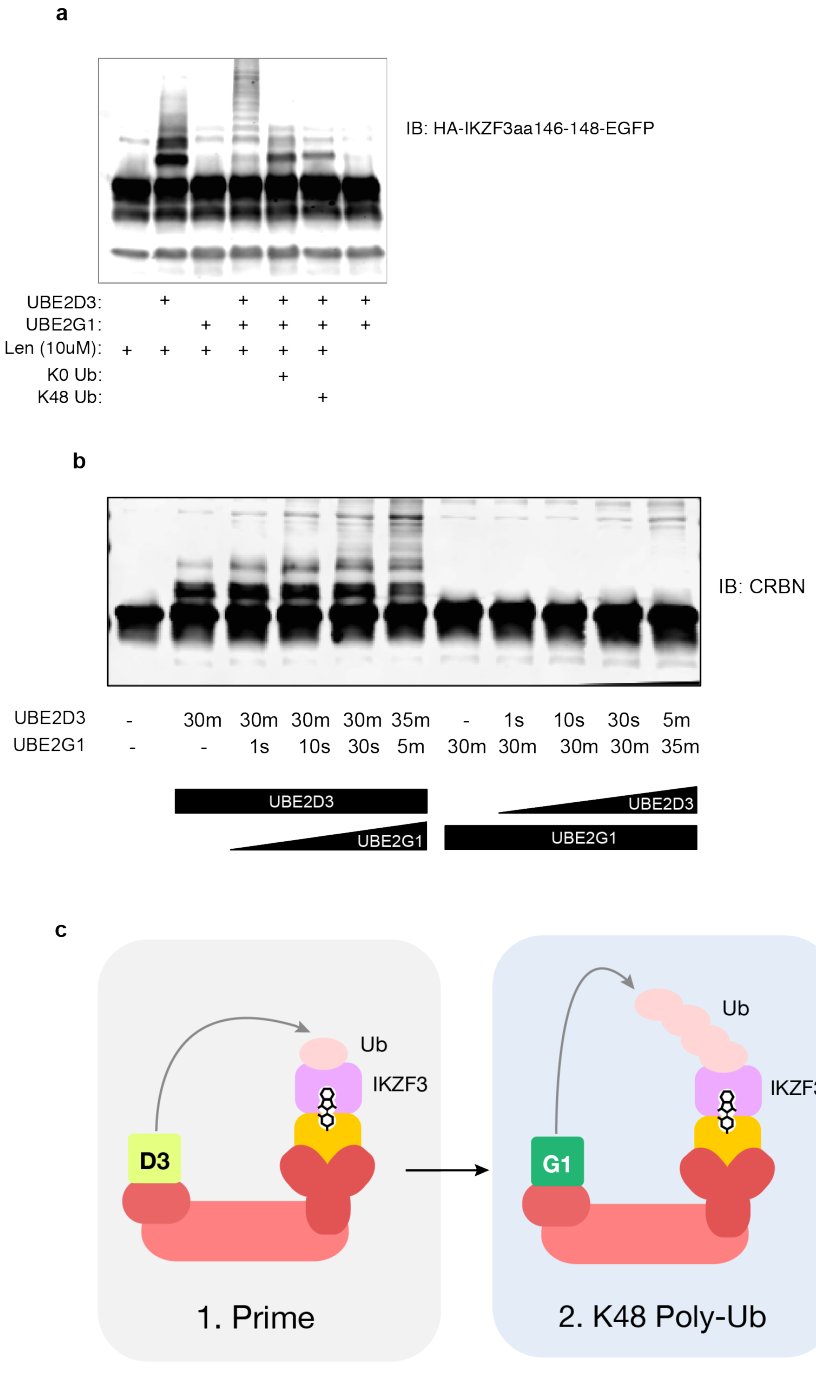


Figure 10 | UBE2D3 monoubiquitinates IKZF3 and primes K48-linked polyubiquitination by UBE2G1. a, In vitro ubiquitination reaction consisting of recombinant E1, E2, $CRL4^{CRBN}$, and HA-IKZF3aa130-189-EGFP which has been immunoprecipitated via the HA tag from HEK293T cells. Reaction was run at 30C for 30 minutes then immunoblotted for IKZF3 and CRBN. **b,** In vitro reaction was run as described with E2 enzymes added at short intervals following a 30min incubation with the reciprocal enzyme. After running the reaction we immunoblotted for CRBN. **c,** Schematic of ubiquitination of lenalidomide-induced $CRL4^{CRBN}$ substrates by UBE2D3 and UBE2G1.

UBE2D3-mediated monoubiquitination is a rate-limiting step in auto-ubiquitination of CRBN

To understand the kinetics of this cooperative process we incubated the reaction for 30 minutes in either the presence of UBE2D3 or UBE2G1 alone, then added the second E2 enzyme and assayed ubiquitin chain formation on CRBN at short intervals **(Figure 10b)**. After 30 minutes of incubation with UBE2D3 we again observed a prominent mono-ubiquitination band, which, following the addition of UBE2G1 resulted in the rapid polyubiquitination of the IKZF3 substrate. Conversely, a 30 minute incubation in the presence of UBE2G1 resulted in no ubiquitination, and the subsequent addition of UBE2D3 failed to produce the rapid polyubiquitination observed in the reciprocal assay. We therefore concluded that monoubiquitination of IKZF3 by UBE2D3 is a rate-limiting step in the reaction.

DISCUSSION

Lenalidomide's ability to induce $CRL4^{CRBN}$-mediated ubiquitination and proteasomal degradation of IKZF1, IKZF3, and CK1α represents a novel therapeutic paradigm. The work presented herein utilizes a genome-scale CRISPR-Cas9 screen to elaborate on our understanding of the cellular machinery which is required for lenalidomide's therapeutic modulation of the $CRL4^{CRBN}$ ubiquitin ligase.

Our unbiased identification of the cellular components required for lenalidomide-induced degradation of $CRL4^{CRBN}$ substrates specifically highlights the importance of UBE2M and the COP9 signalosome in regulating the CRL neddylation cycle, as well as the role that the E2 enzymes UBE2D3 and UBE2G1 play in facilitating K48-linked polyubiquitination of IKZF3. These findings complement a fluorescence-based, arrayed, genome-scale shRNA screen for factors regulating $CRL4^{CDT2}$ degradation of the replication origin licensing factor CDT1 in UV-irradiated HeLa cells[52]. Similarities include the identification of components of the ubiquitin ligase (*CDT2*, *DDB1*, *RBX1*), *UBE2M* (UBC12), and subunits of the COP9 signalosome (*COPS2*, *COPS6*, *COPS5*, *COPS8*, *COPS7B*). Differences, likely due to the use of shRNA-based screening and a shorter assay time point, include the identification proteasomal subunits, *p97*, *UFD1*, *NAE1*, factors which did not score in our assay likely due to their effects on viability.

Unique factors which scored in our assay were the E2 enzymes UBE2D3 and UBE2G1. UBE2G1 in particular does not have extensive literature describing its function or which E3 ligases it pairs with. While G1 appears to be required for all known lenalidomide-induced substrates of $CRL4^{CRBN}$, it remains to be seen whether there are other contexts in which $CRL4^{CRBN}$ might rely on a distinct set of E2 enzymes, whether other CRL4-containing E3 ligases utilize UBE2G1, and what structural motifs on $CRL4^{CRBN}$ or its substrates dictate the use of UBE2G1.

In addition to deepening our understanding of CRL regulation and E2-E3 pairings, the discovery of the factors regulating $CRL4^{CRBN}$ function holds implications for the clinical use of lenalidomide as well as what is likely to become a growing class of drugs that modulate the function of ubiquitin ligases. Indeed, the results of the screen serve as an unbiased and orthogonal confirmation of lenalidomide's mechanism of action; recent work has suggested an E3-ligase independent role for cereblon is responsible for its therapeutic properties[53], however the predominance of *CRBN* and regulators of cullin-ring ligase function in the screen suggests that lenalidomide mediates its anti-myeloma functions principally via modulation of $CRL4^{CRBN}$ function.

The predominance of *CRBN* as the top hit in a genome-scale screen additionally suggests that *CRBN* is likely to be a central node of acquired lenalidomide resistance. Interestingly, preliminary sequencing of CRBN and DDB1 coding exons in 36 patients with lenalidomide-resistant disease did not yield loss-of-function mutations[54]. Instead, several studies have noted that cereblon expression levels and isoform usage are altered at relapse [55-57]. A combination of expression profiling and sequencing of *CRBN* and its surrounding regulatory regions in a large number of patients pre- and post-relapse will clarify the extent to which loss-of-functions mutation or epigenetic regulation of cereblon is indeed the main mechanism of lenalidomide resistance in patients.

In summary, we have utilized an unbiased screen to discover cellular machinery required for lenalidomide's modulation of the $CRL4^{CRBN}$ ubiquitin ligase, a first-in class therapeutic mechanism of action. The recent discovery of the indisulam family of drugs which induce the degradation of RBM39 by $CRL4^{DCAF15}$ [58] suggests that this paradigm will lead to the generation of a broader class of ubiquitin-ligase modifiers. Additionally, lenalidomide's mechanism of action

has also inspired the generation of “degrader compounds” in which $CRL4^{CRBN}$ can be redirected to degrade targets of interest by conjugating a small molecule with binding specificity to lenalidomide or its analogs[41,43]. Owing to their reliance on $CRL4^{CRBN}$, this class of drugs will likely be dependent on the same suite of cellular machinery identified in our work, such as neddylation enzymes, E2 enzymes, and substrate receptors.

METHODS

Cell Culture

MM1S (MM.1s, multiple myeloma, ATCC #2974) were grown in RPMI supplemented with 10% Fetal Bovine Serum, Penicillin, and Strepomycin. HEK293T (Human Embryonic Kidney, Broad Institute Cell Line Repository) were grown in DMEM supplemented with Fetal Bovine Serum, Penicillin, and Strepomycin. Insect cells. Cell line identity was not verified.

Genome-Scale CRISPR-Cas9 Screen

To optimize the dose and time points for the screen, MM1.S cells expressing Cas9 were infected with a non-targeting control gRNA or CRBN-targeting gRNA + EGFP, then mixed at a 95:5 ratio and cultured in a titration of lenalidomide doses for 20 days. 1uM lenalidomide was the minimum dose required to elicit >4 fold enrichment by day 12, therefore we chose 1uM as the screen dose and day 12 and day 20 as the time points for the screen.

Prior to the screen, MM1.S cells were lentivirally transduced with the pLentiCas9-Blast construct (Addgene #52962) and selected with blasticidin. On day -8 MM1S-Cas9 cells were then infected with the human GeCKOv2 LentiGuide-Puro library (Addgene #52962), and on day -7 the cells were selected with 1ug/mL puromycin. On day -4 the cells were passaged into puromycin-free media, and on day 0 we collected a baseline control (120 million cells) and the rest of the cells were divided into the DMSO (60 million) and 1uM lenalidomide (6x replicates of 120 million cells) treatment groups. For the next 20 days, cells were dosed every two days with either DMSO or 1uM lenalidomide and passaged every four, maintaining a cell density of 1 million/mL. On days 12 and 20 we harvested 60 million DMSO cells from the DMSO arm, and all the cells in either replicates 1-3 (D12) or 4-6 (D20) in the 1uM lenalidomide treatment group.

Genomic DNA was harvested using the Qiagen DNA Blood Maxi kit (Qiagen #51192), after which we PCR amplified the gRNA sequences using primers tailed with barcodes and Illumina sequencing adaptors. The resulting amplicons were then pooled and then sequenced across four lanes of the Illumina HiSeq. Reads counts were normalized between samples and log2 transformed. gRNA enrichment was calculated by averaging the replicate read counts for a given gRNA and then subtracting the corresponding DMSO read count from the lenalidomide read count. To collapse the gRNA enrichment results by gene and generate statistics we employed the use of the STARS algorithm (Mudra Henge, Broad Genetic Perturbations Platform, http://www.broadinstitute.org/rnai/public/software/stars).

IKZF3-EGFP Reporter Counterscreen

gRNAs were designed for the top 30 scoring genes in the Primary Screen using the Broad Genetic Perturbations Platform's gRNA prediction algorithm (https://iwww.broadinstitute.org/rnai/db/analysis-tools/sgrna-design). Three gRNAs per gene were chosen with priority given to gRNAs with no perfect matches elsewhere in the genome, location in the first 30% of the protein, and three gRNAs all located in different exons. 12 gRNAs targeting no known sequences were also included as negative controls. gRNAs were ordered from IDT as paired, single-stranded oligos, after which they were annealed and cloned in an arrayed format into pXPR003 (gRNA expression vector with puromycin resistance). Plasmids were then pooled, and used to generate a lentiviral library.

HEK293T cells were transduced with pLentiCas9-Blast construct (Addgene #52962) and selected with blasticidin. The MM1S-Cas9and HEK293T-Cas9 cells were then transduced with the IKZF3 reporter (IKZF3 aa30-189 in Pea, Addgene #74459) and infected cells were isolated using fluorescence activated cell sorting (BD FACS Aria) . On day 0 of the screen three replicates of each reporter cell were infected with the lentiviral library at an efficiency below

50%, and on day 1 the cells were selected with puromycin. On day 4 cells were passaged into puromycin-free media, and on day 11 the cells were dosed with DMSO or 1uM lenalidomide. On day 12, baseline controls were harvested from both the DMSO and 1uM lenalidomide treatment groups, and 1uM Lenalidomide treated cells were then sorted for cells remaining EGFP+ despite treatment with lenalidomide.

Genomic DNA was isolated from the resulting cell pellets using the QiaAmp DNA Blood Midi and Micro Kits (Qiagen #51183, #56304), after which we PCR amplified the gRNA sequences using primers tailed with barcodes and Illumina sequencing adaptors. The resulting amplicons were then pooled and then sequenced across a single lane of the Illumina NextSeq. Reads counts were normalized between samples and log2 transformed. gRNA enrichment in the EGFP+ sort was calculated by subtracting the DMSO gRNA read count from its paired 1uM lenalidomide read count, and then averaging the replicate values of the resulting log2 fold-enrichment. The averaged, log2 fold enrichment score was then normalized to the average enrichment of the 12 non-targeting control gRNAs.

Generation of Single Cell Clone KOs

HEK293T cells expressing Cas9 were transduced with lentiviral vectors expressing gRNA. After ~6-9 days, single cells were plated via limiting dilution in 96-well plates and grown to confluence. Genomic DNA was isolated from the cells, then we PCR amplified the gRNA cut-site and sent the amplicons for MGH Next-gen CRISPR sequencing.

HEK293T Immunoblots

Protein lysates were harvested and electrophoretically run on a polyacrylamide gel, transferred to nitrocellulose membranes, and immunoblotted for the following proteins: CUL4A, CUL4B,

COPS5, UBE2M, CRBN, HA, IKZF1, IKZF3, and CK1α. Imaging was performed on the Licor Odyssey.

IKZF1 (Abcam, ab180713)

IKZF3 (Novus Biologicals, NBP2-16938)

CRBN (Novus Biologicals, Rabbit anti-CRBN, NBP1-91810)

CUL4A (Abcam, ab72548)

CUL4B (Sigma, HPA011880)

CK1α (Santa Cruz Biotech, sc-6477)

UBE2M (Ubc12) (Cell Signaling Technology, 4913S)

COPS5 (Cell Signaling Technology, 9444S)

Competition Assays

MM1.S cells were infected with lentivirus expressing either gRNA and EGFP or control gRNA without a fluorescent marker. After at least 8 days, these cells were then mixed at a 5:95 ratio respectively and cultured in three technical replicates in either DMSO or 1uM lenalidomide. Cells were dosed every two days and passaged every four. With every passage, cells were taken and we used Cytoflex flow cytometers to assay for the percentage of EGFP positive cells.

Lenalidomide Titration in IKZF3-EGFP

HEK293T cells expressing the IKZF3-EGFP reporter were plated in 96-well plates and dosed using the D3000 drug printer with a titration of doses of lenalidomide for 20 hours, after which the cells were harvested and the EGFP/mCherry ratio was calculated on the Cytoflex flow cytometer.

MM1.S Immunoblots

MM1.S cells were infected with gRNA expressing lentivirus and culture for at least 8 days, after which they were plated at 10 million cells per 20mL for 24 hours, then treated with either DMSO or 1uM lenalidomide for 20 hours. Protein lysates were harvested, run on polyacrylamide gels, transferred to nitrocellulose membranes, and immunoblotted for CRBN, IKZF1, IKZF3, and CK1α. Imaging was performed on the Licor Odyssey.

In vitro *ubiquitination assays*

CRL4-CRBN was generated in insect cells as previously described (Fischer, E. S. et al. The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. *Cell* **147**, 1024–1039 (2011)). HEK293T were transfected with IKZF3 degron-HA-GFP and cells were lysed after 48hr. HA-sepharose beads were used to pulldown IKZF3 degron-HA-GFP followed by subsequent elution with 100ug/mL HA peptide at 4C for 30 min. Substrate eluate was added to ubiquitination reaction mixture containing 2uM CRL4-CRBN, 200 nM UBE1 (Boston Biochem), 1uM UBE2G1 (Boston Biochem), 1 uM UBE2D3 (Boston Biochem), 1 uM ubiquitin aldehyde (Boston Biochem), 1x Mg-ATP (Boston Biochem), 1x E3 Ligase Reaction Buffer (Boston Biochem), and 1 uM lenalidomide, or DMSO. *In vitro* ubiquitination assay was run for 30 min at 30C and run on Tris-Glycine SDS-PAGE gel and immunoblotted as indicated.

Chapter 3: Functional and structural definition of C2H2 zinc finger degrons targeted by thalidomide analogs

Quinlan L. Sievers[1,2,3], Georg Petzold[6,7], Brian Liddicoat[1,2], Jenny Chen[4], Tarjei Mikkelsen[1,a], Nicolas H. Thoma[6,7], Benjamin L. Ebert[1,2,5]

[1]Broad Institute of Harvard and MIT, Cambridge, MA
[2]Division of Hematology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA
[3] MD/PhD program, Harvard Medical School, Boston, MA
[4]Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge MA
[5]Dana Farber Cancer Institute, Department of Medical Oncology, Boston, MA
[6]Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland
[7]University of Basel, Basel, Switzerland
[a]Present address: 10X Genomics, Inc., Pleasanton, CA

Corresponding Authors:
Benjamin L. Ebert, MD/PhD
77 Avenue Louis Pasteur, HIM 743
Boston, MA 02115
bebert@partners.org

Nicholas H. Thoma, PhD
Friedrich Miescher Institute for Biomedical Research
Maulbeerstrasse 66
4058 Basel, Switzerland
nicolas.thoma@fmi.ch

CONTRIBUTIONS

Quinlan Sievers and Benjamin Ebert worked collaboratively to form the concept and design of the saturation mutagenesis screen and C2H2 zinc finger library screen. The saturation mutagenesis screen and C2H2 zinc finger library screen were performed by Quinlan Sievers, with help from Tarjei Mikkelssen (saturation mutagenesis library synthesis and analysis) and Jennifer Chen (zinc finger library screen analysis). Validation of the screens and the subsequent experiments were performed by Quinlan Sievers with help from Brian Liddicoat (validation of ZFP91 and ZNF692). Georg Petzold performed the crystallography and TR-FRET assays. Quinlan Sievers drafted the manuscript. All work was conducted under the direction of Nicholas Thoma and Benjamin Ebert.

ACKNOWLEDGEMENTS

Thank you to Donna Neuberg and Robert Redd of the Dana Farber Cancer Institute Biostatistics core for the help analyzing the saturation mutagenesis screen, Cong Zhu for advice for cloning the saturation mutagenesis library, and members of the Ebert lab, specifically Jessica Gasser, for helpful discussions about this work.

ABSTRACT

Thalidomide, lenalidomide, and pomalidomide induce the ubiquitination and proteasomal degradation of the transcription factors Ikaros (IKZF1) and Aiolos (IKZF3) by mediating their interaction with Cereblon (CRBN), the substrate receptor for the $CRL4^{CRBN}$ E3 ubiquitin ligase. Here we have used a combination of saturation mutagenesis, crystallography, and motif-based screening to characterize the IKZF1/IKZF3 degron and identify novel targets of thalidomide analogs. Saturation mutagenesis identified the second C2H2 zinc finger in IKZF1 and IKZF3 as necessary for drug-induced degradation, and specifically highlighted Q147 and A153 within the β-hairpin of zinc finger 2 as residues whose amino acid identity is required for degradation. A 3.5Å crystal structure of zinc finger 2 bound to pomalidomide and cereblon reveals that the β-hairpin of zinc finger 2 forms a ternary degron which makes contacts with both pomalidomide and cereblon, and we note that the side chain of Q147 forms a crucial interaction with the phthaloyl group of pomalidomide. To broaden our understanding of C2H2 zinc finger degrons and identify novel targets we screened ~6,500 unique C2H2 zinc fingers found in the human proteome and identified 10 additional sequences targeted for degradation by either thalidomide, lenalidomide, or pomalidomide. Lastly, we confirm that ZFP91 and ZNF692 are C2H2 degron-containing proteins which are degraded by pomalidomide. In aggregate, our work demonstrates that thalidomide analogs induce $CRL4^{CRBN}$-dependent degradation of proteins via C2H2 zinc finger degrons, a motif which is enriched in transcription factors.

INTRODUCTION

The majority of small molecule therapeutics target proteins with discrete catalytic sites or binding pockets, including ion channels, G-protein coupled receptors, nuclear receptors, and kinases[30]. While this strategy has yielded clinical successes in the treatment of cancer, whole exome sequencing of tumors and cell line-based functional genetic screens have identified therapeutic targets which lack the structural features amenable to targeting with conventional small molecules. Transcription factors in particular are a desirable class of therapeutic targets in cancer because they are often the endpoint of oncogenic signaling pathways controlling cellular identity and proliferation. However, because transcription factors' effector functions are mediated by protein-protein interactions rather than catalytic domains, they remain poor candidates for pharmacologic inhibition.

Thalidomide and its analogs, lenalidomide and pomalidomide, are effective treatments for the hematopoietic malignancy multiple myeloma[44,45,59,60]. Their therapeutic properties are derived from the ability to induce the ubiquitination and proteasomal degradation of the transcription factors IKZF1 (Ikaros) and IKZF3 (Aiolos) via the $CRL4^{CRBN}$ E3 ubiquitin ligase[15-17]. E3 ubiquitin ligases are a class of cellular enzymes which covalently mark their target substrates with the 8kD protein ubiquitin, a signal which, in the form of polyubiquitin chains, targets substrates for rapid degradation by the 26s proteasome. Thalidomide analogs augment this process by stabilizing the interaction of IKZF1/IKZF3 with cereblon (CRBN), the substrate receptor for the $CRL4^{CRBN}$ ubiquitin ligase.

The crystal structure of lenalidomide in complex with CRBN and casein kinase 1α (CK1α), another substrate degraded by these compounds, has revealed that lenalidomide binds at the CK1α-CRBN interface, forming interactions with a tri-tryptophan binding pocket on CRBN and a

β-hairpin in CK1α. Given the inherent structural limitations of targeting transcription factors it is of interest to identify and characterize the structural motif, or degron, which mediates the interaction of IKZF1/IKZF3 with the drug-cereblon complex.

Here we have utilized a combination of saturation mutagenesis and crystallography to identify the β-hairpin of the second C2H2 zinc finger in IKZF1 and IKZF3 as the degron motif, specifically highlighting the drug-interacting residue Q147. To gain a greater understanding of the spectrum of C2H2 zinc fingers capable of being degraded by thalidomide analogs we also have screened ~6,500 unique C2H2 zinc fingers in the human proteome and identified 10 novel C2H2 zinc fingers which are degraded in the presence of thalidomide analogs. Lastly we confirm that ZFP91 and ZNF692 are C2H2 degron-containing proteins which are also degraded at the full-length level.

These results indicate that thalidomide analogs target a greater than previously appreciated number of substrates via C2H2 zinc finger degrons, a motif enriched in transcription factors.

RESULTS

Saturation mutagenesis identifies residues in IKZF3 C2H2 zinc finger 2 required for degradation

We previously established that a 60 amino acid (aa) region of IKZF3 (aa130-189) is necessary and sufficient for thalidomide analog-induced degradation by the $CRL4^{CRBN}$ ubiquitin ligase [15]. To understand more specifically which residues within this region are required for degradation we synthesized a mutagenesis library of IKZF3 aa130-189 such that at each of the 60 amino acids, all 19 possible mutations were represented. The library of ~1200 constructs was cloned in a pooled format into a reporter vector which allowed us to use flow cytometry to monitor post-translational degradation of library constructs via a reduction in the EGFP/mCherry ratio **(Figure 11a)**.

We transduced HEK293T cells with the lentiviral library and after treating the cells for 20 hours with DMSO, thalidomide, lenalidomide, or pomalidomide, we used fluorescence activated cell sorting (FACS) to collect EGFP-/mCherry+ cells. We then PCR-amplified the mutagenesis library from genomic DNA isolated from the sorted cell populations and performed next-generation sequencing of the amplicons **(Figure 11a)**. Amino acids whose identity was crucial for degradation were expected to exhibit a reduction in the frequency of mutant read counts in the drug-treated experimental conditions relative to DMSO-treated controls.

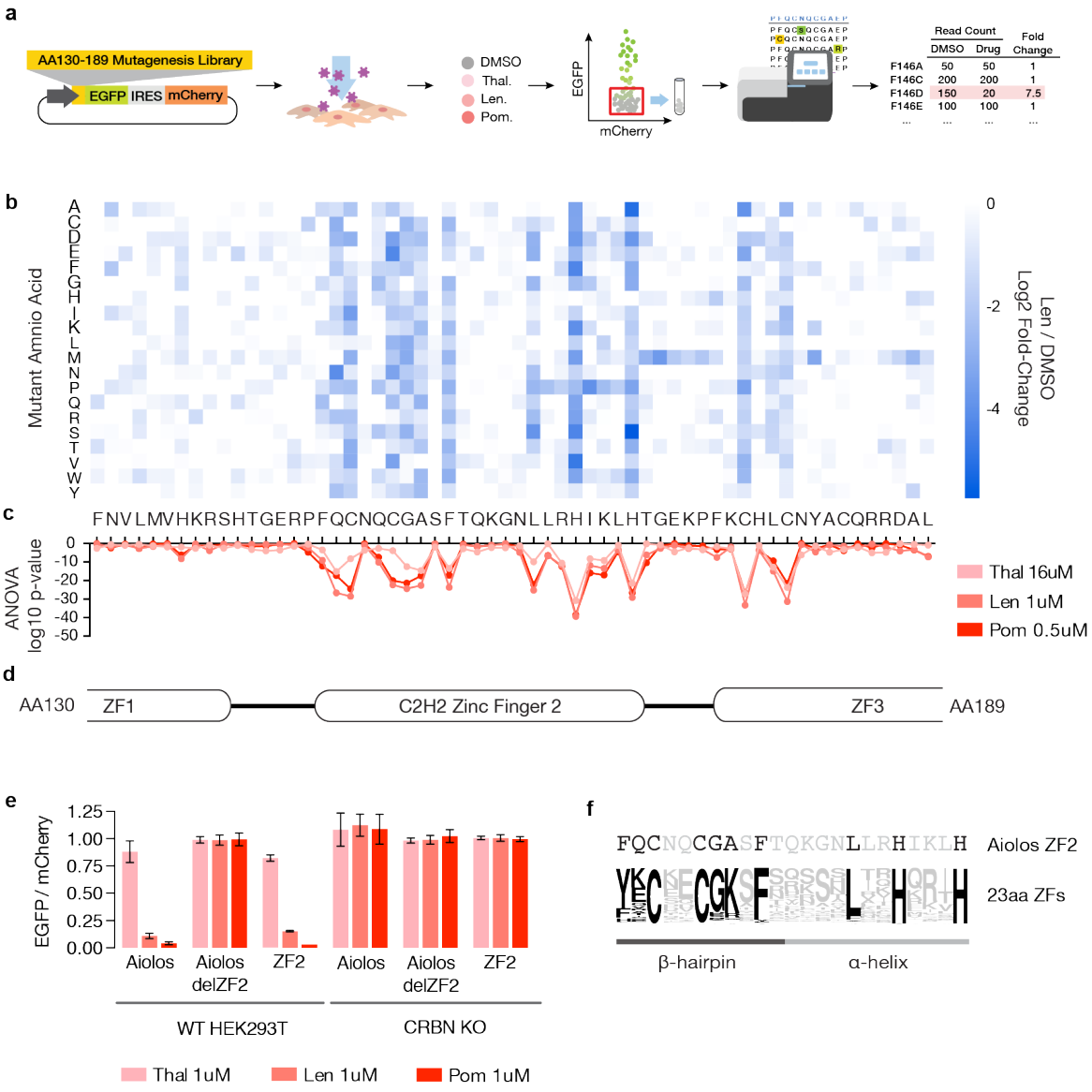


Figure 11 | Saturation mutagenesis of IKZF3 aa130-189 identifies residues required for thalidomide-analog induced degradation. a, Overview of the saturation mutagenesis screen. **b,** Results for lenalidomide, heat map displaying the average log2 fold-change in frequency of sequencing reads encoding mutant amino acids (3 technical replicates). **c,** ANOVA p values for difference in frequency of mutant amino acids at each position in IKZF3 aa130-189, for thalidomide, lenalidomide, and pomalidomide. **d,** Annotated structural features in IKZF3 aa130-189 (Uniprot, Q9UKT9). **e,** Aiolos constructs were cloned into the EGFP/mCherry protein degradation reporter vector and were lentivirally expressed HEK293T cells. After 20 hours of drug treatment the DMSO-normalized EGFP/mCherry ratio was measured via flow cytometry (experimental replicates = 3, technical replicates = 3, bar heights indicate mean of experimental replicates, error bars indicate 95% CI). **f,** Comparison of Aiolos ZF2 amino acid sequence to the frequency of amino acids found in all 5,994 23aa C2H2 zinc fingers in the human proteome.

Lenalidomide treatment resulted in a reduction in the frequency of mutant amino acid reads at 12 non-contiguous amino acids **(Figure 11b, Appendix Figure 22)**. Q147, an amino acid whose mutation to histidine was previously known to prevent the degradation of the homologous proteins Helios (IKZF2) and Eos (IKZF4), was among these conserved residues[15]. We also noted the depletion of proline along aa161-168, a region known to contain an alpha helix, and the depletion of methionine across the C-terminal residues aa169-189, perhaps due to the generation of alternative start sites which bypass the degron. To integrate the data from the 19 mutations at a given locus and compare the three drugs we calculated an ANOVA p-value for the difference in the frequency of mutant read counts observed in the DMSO and drug-treated replicates **(Figure 11c)**. Thalidomide, lenalidomide, and pomalidomide exhibited similar profiles of amino acids whose identity was necessary for degradation.

IKZF1/IKZF3 zinc finger 2 is necessary and sufficient for degradation

Amino acids whose mutation impaired degradation were located primarily within IKZF3's second C2H2 zinc finger, a region which is identical in IKZF1 **(Figure 11d)**. C2H2 zinc fingers are ~23aa motifs whose β-hairpin and α-helix structure is maintained by the chelation of a zinc ion by two N-terminal cysteines and two C-terminal histidines. IKZF1 and IKZF3 each possess four N-terminal C2H2 zinc fingers which mediate DNA binding and two C-terminal C2H2 zinc fingers which facilitate homo- and hetero-dimerization with Ikaros-family transcription factors[61,62] **(Appendix Figure 23)**.

To determine if zinc finger 2 was necessary and sufficient for thalidomide analog-induced degradation we cloned IKZF3, IKZF3 del(ZF2), and ZF2 into the EGFP/mCherry reporter vector used for the initial screen. HEK293T cells transduced with the constructs were treated for 20 hours with DMSO, thalidomide, lenalidomide, or pomalidomide, after which we used flow

cytometry to measure the EGFP/mCherry ratio. Degradation of IKZF3 was blocked by deletion of zinc finger 2, and zinc finger 2 alone conferred drug-induced degradation upon EGFP commensurate with the full-length protein **(Figure 11e)**.

Q147 and A153 are unique residues within the β–hairpin of zinc finger 2 which are required for degradation

To identify residues in zinc finger 2 which explain the specificity of the interaction with the drug-cereblon complex we compared the amino acid sequence of zinc finger 2 with the frequency of amino acids found in the 5,994 23aa C2H2 zinc fingers in the human proteome **(Figure 11f)**. 8 of the 10 residues highlighted by the saturation mutagenesis screen were common amongst C2H2 zinc fingers. The majority of these positions play known roles in maintaining C2H2 tertiary structure; C148, C151, H165, and H168 chelate the zinc ion, and F146, F155, and L161, form a hydrophobic core[63].

The remaining residues, Q147 and A153, are unique: zinc finger 2 in IKZF1 and IKZF3 are the only C2H2 zinc fingers in the proteome to simultaneously possess a Q at position 2 and an A at position 8. In keeping with these observations, mutation of Q147 and A153 in our reporter construct impaired degradation, while mutation of N149, which was not conserved in the screen, did not alter degradation **(Appendix Figure 24)**.

From these experiments we concluded that zinc finger 2 is both necessary and sufficient for thalidomide-analog induced degradation, and that, of the residues highlighted in the screen, Q147 and A153 within the β–hairpin likely drive the specificity of the interaction with the drug-cereblon complex.

Crystal structure of the ZF2-pomalidomide-CRBN complex

To understand the structural correlates of the saturation mutagenesis screen we sought to crystallize the tertiary complex of IKZF1/IKZF3 zinc finger 2 bound to the drug-cereblon complex. Crystals were obtained containing zinc finger 2, pomalidomide, CRBN, and $DDB1^{\Delta BPB}$, and were resolved to a resolution of ~3.5Å **(Figure 12a)**.

In agreement with prior structures, pomalidomide is bound via its glutarimide ring to CRBN's C-terminal domain (CRBN-CTD) with its phthaloyl moiety exposed. Zinc finger 2 sits within a narrow groove on the surface of CRBN-CTD and makes contacts with both the phthaloyl ring of pomalidomide and cereblon via its β–hairpin **(Figure 12b)**. While the shape of the β-hairpin is analogous the β-hairpin of CK1α in complex with lenalidomide-cereblon[29] **(Figure 12c)** the hairpins lack sequence homology save for a glycine residue located where the hairpin passes through a narrowing created by the drug and cereblon **(Figure 12d)**.

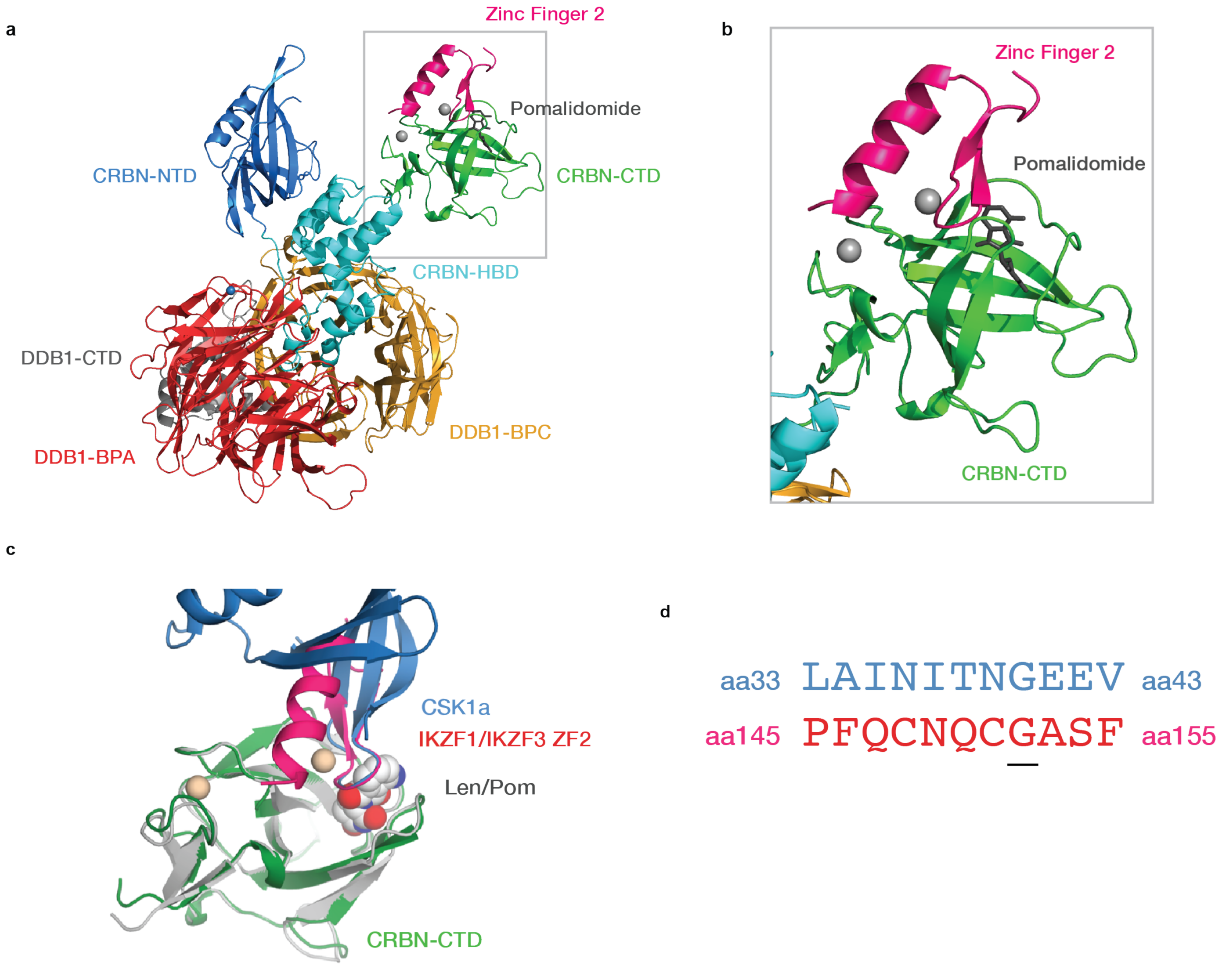


Figure 12 | 3.5Å Crystal structure of the ZF2-pomalidomide-CRBN-DDB1$^{\Delta BPB}$ complex. a, Overview of the complex. **b,** Magnified view of zinc finger 2 bound to pomalidomide-CRBN. **c,** Overlay of ZF2 and CK1α (PDB 4TZ4) in complex with pomalidomide, lenalidomide, and CRBN. **d,** alignment of amino acids comprising the β–hairpins of CK1α (blue) and IKZF1/IKZF3 C2H2 zinc finger 2 (pink). Underlined residue is conserved glycine.

Q147 interacts with the phthaloyl group of pomalidomide

We next sought to determine the structural role of amino acids identified in the saturation mutagenesis screen. The majority of the interactions between the β–hairpin and the pomalidomide-cereblon complex are mediated by the peptide backbone of residues 2-7 in the hairpin and pomalidomide's phthaloyl ring **(Figure 13a)**. Residues which were required for degradation but common to C2H2 zinc fingers act to stabilize the tertiary structure of the zinc finger, facing inward and either chelating the zinc ion or packing the hydrophobic center **(Figure 13b)**. G152 was common amongst C2H2 zinc fingers but does not appear to play a structural role; instead, as noted earlier, G152 sits at position that is sterically limited by pomalidomide and CRBN and cannot permit an amino acid larger than glycine (data not shown).

Q147 and A153 were unique residues conserved in the saturation mutagenesis screen. A153 faces outward towards CRBN and as a result mutation would be predicted to distort the beta loop. Q147 is particularly notable in that it forms a side chain contact with the top face of pomalidomide's phthalimide ring and approaches the aniline group to potentially form a hydrogen bond **(Figure 13b)**. Pomalidomide is identical to thalidomide save for this aniline group, a chemical moiety which renders pomalidomide more potent in both cellular degradation assays and in vitro binding assays **(Figure 13c)**.

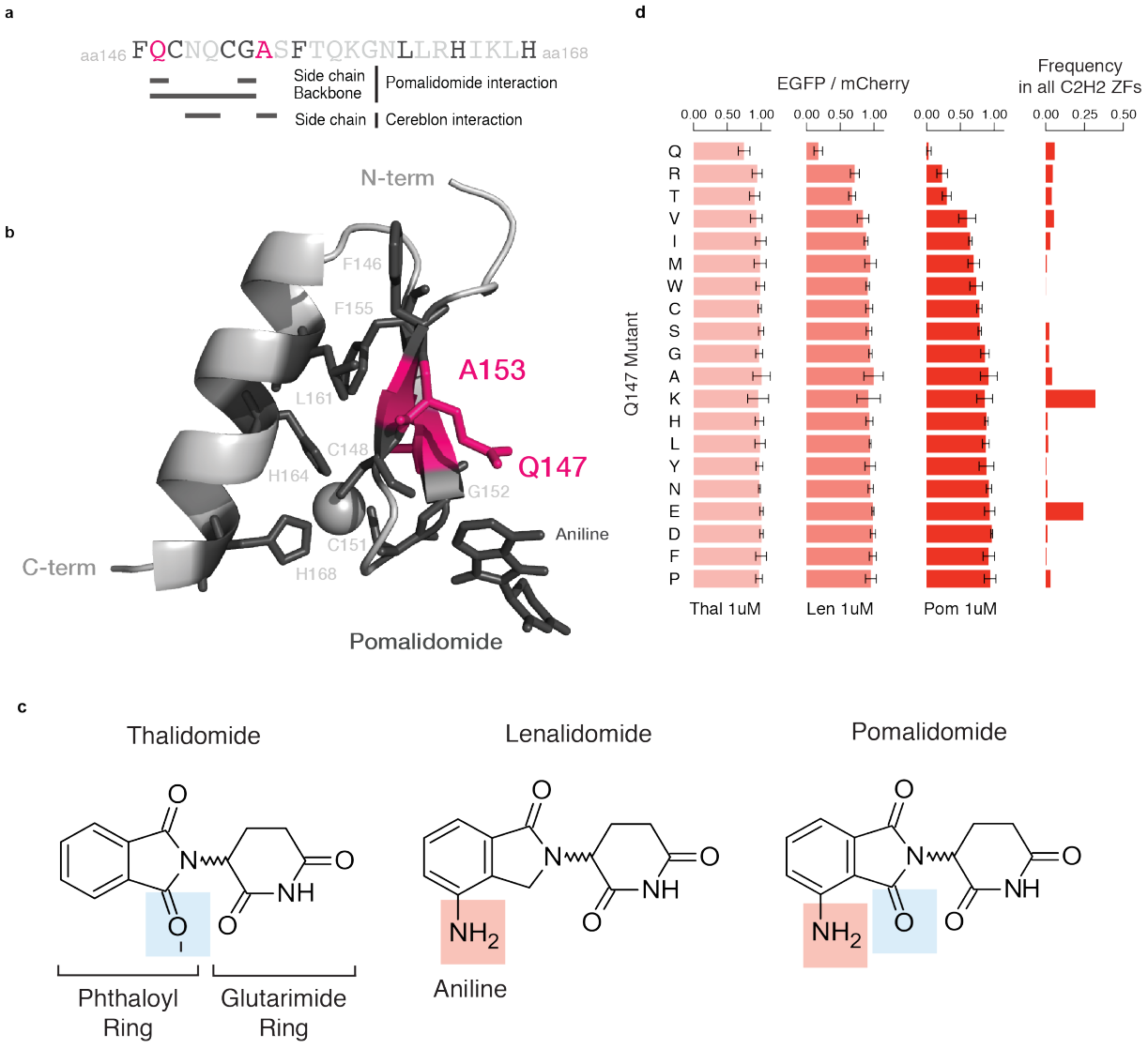


Figure 13 | Q147 interacts with the phthaloyl group of pomalidomide. a, Amino acid sequence of IKZF1/IKZF3 zinc finger 2, with residues identified by the saturation mutagenesis assay highlighted (common residues in black, unique residues in pink) and structure annotations based on the crystal structure. **b,** IKZF1/IKZF3 zinc finger two with amino acids identified in the saturation mutagenesis screen highlighted similar to “c”. **c,** Chemical structures of thalidomide, lenalidomide, and pomalidomide. **d,** HEK293T cells were transduced with Q147 mutants in an EGFP/mCherry reporter vector, treated for 20h with 1uM thalidomide, lenalidomide, or pomalidomide, then the EGFP/mCherry ratio was measured using flow cytometry (experimental replicates =3, technical replicates =3, bar height indicates average of experimental replicates, error bars indicate 95% CI).

To determine the specificity of the Q147-pomalidomide interaction, we generated constructs containing all 19 possible amino acid mutations in our EGFP/mCherry reporter vector,

expressed them in HEK293T cells, treated the cells for 20h with thalidomide, lenalidomide, and pomalidomide, and used flow cytometry to monitor the EGFP/mCherry ratio **(Figure 13d)**. The wild-type Q147 construct exhibited the greatest degradation, and notably, the most common residues at this position in all 23aa zinc fingers, lysine (K) and glutamate (E), completely stabilized the degron in the presence of drug. Arginine (R), Threonine (T), and Valine (V) still allowed for degradation, albeit diminished, to occur.

From these experiments we concluded that zinc finger 2 interacts with the pomalidomide-cereblon complex via its β-hairpin, and that Q147, which was highlighted as a critical residue by our saturation mutagenesis screen, forms a side chain interaction with the phthaloyl group of pomalidomide.

Screening ~6,500 C2H2 zinc finger sequences for thalidomide analog-induced degradation

C2H2 zinc-finger containing proteins are the largest class of putative transcription factors[64]. We therefore next asked if additional C2H2 zinc fingers are degraded in in the presence of thalidomide analogs. To this end we used a C2H2 motif (X-X-C-x(2,4)-C-x(3)-[LIVMFYWC]-x(8)-H-x(3,5)-H) to curate and synthesize a library of 6,819 human C2H2 zinc finger sequences (6,572 unique) originating from 811 genes. We then cloned the corresponding cDNA in a pooled format into the lentiviral EGFP/mCherry reporter vector. The library was expressed in HEK293T cells which were treated for 20 hours with DMSO, thalidomide, lenalidomide, or pomalidomide. We then isolated EGFP+/mCherry+ cells using FACS, PCR amplified the zinc finger library from genomic DNA isolated from the sorted cells, and used next-generation sequencing of the amplicons to quantify the relative frequency of read counts for each zinc finger **(Figure 14a)**. Zinc fingers which were degraded in the presence of drug were expected to exhibit a depletion in the frequency of read counts in the drug-treated condition relative to the DMSO control.

10 novel C2H2 zinc fingers are targeted for degradation by thalidomide analogs

11 zinc fingers scored with a false discovery rate (FDR) of <0.01 in at least one of the three drug treatment conditions (empirical rank-sum p-value with FDR correction) **(Figure 14b)**. Zinc finger 2 from IKZF1 and IKZF3 was among these hits in the lenalidomide and pomalidomide treatment conditions, lending confidence to the sensitivity of the screen.

To validate the results of the screen we cloned the 11 zinc fingers into our EGFP/mCherry reporter vector, expressed them in HEK293T cells, and then treated the cells for 20 hours with a titration of thalidomide, lenalidomide, or pomalidomide **(Figure 14c, Appendix Figure 27)**. All 11 zinc fingers exhibited degradation in the presence of each of the three drugs to varying

degrees. Pomalidomide was the most potent and efficacious. The zinc fingers segregated into two groups on the basis of their response to lenalidomide and thalidomide, possessing greater sensitivity to either one or the other. We additionally tested the next 13 zinc fingers from the screen, corresponding to an FDR threshold of <0.1; none showed evidence of degradation (data not shown).

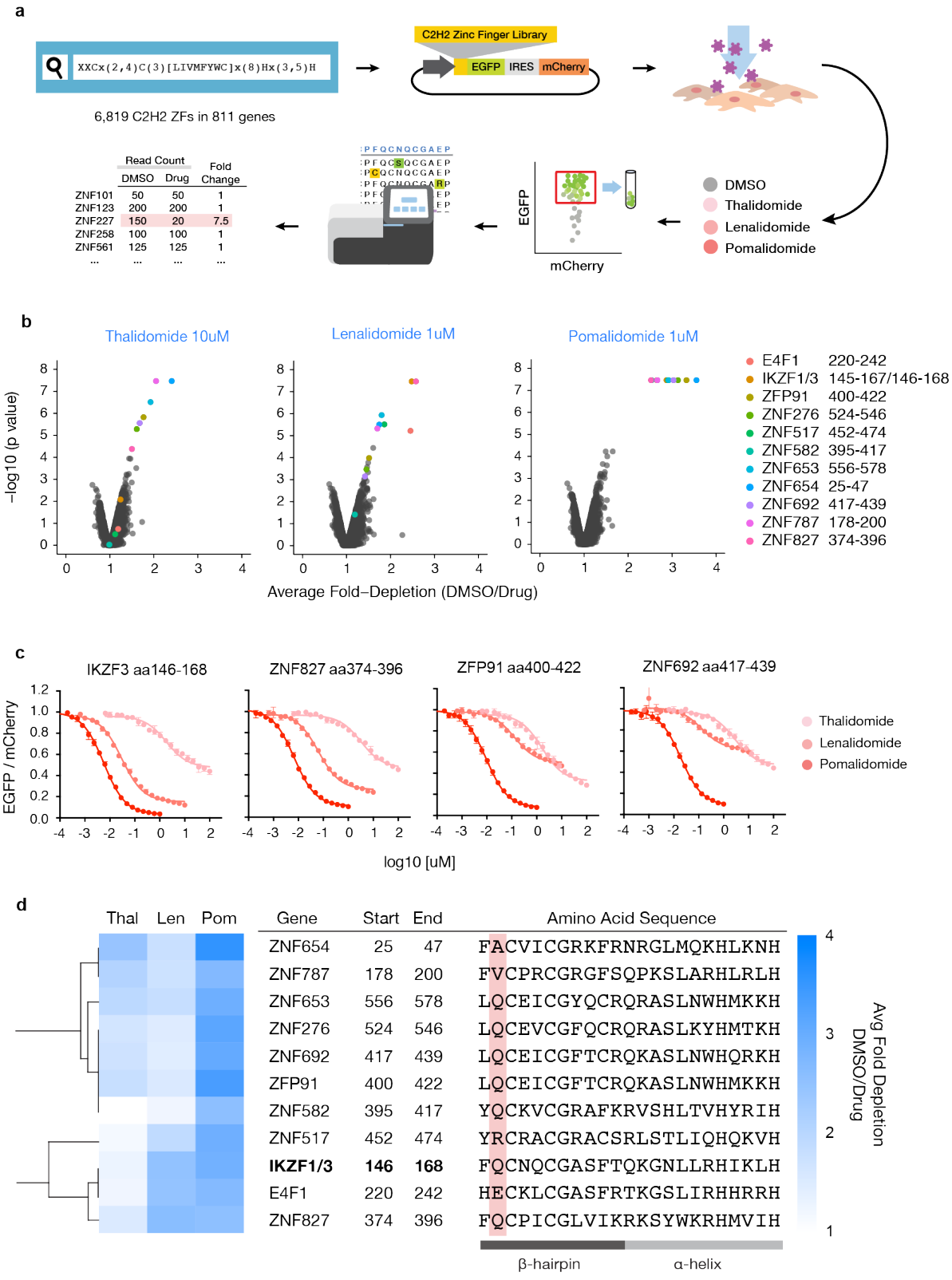
Novel amino acids are found at the drug-interaction position

To examine trends in drug response and amino acid sequence we hierarchically clustered (one-minus pearson correlation) the 11 validated zinc fingers on the basis of their average fold-depletion in the screen and aligned their amino acid sequences **(Figure 14d)**. In keeping with the validation assay, the zinc fingers were clustered into two groups on the basis of their response to thalidomide and lenalidomide. Due to the limited number of hits, we lacked the statistical power to identify amino acid residues differentiating these two categories.

We instead focused on position 2, the amino acid whose side chain in IKZF1/IKZF3 zinc finger 2 was found to interact with pomalidomide. Like IKZF1/IKZF3, 6 of the 10 new zinc fingers possessed a glutamine (Q) at this position, however we also observed zinc fingers with valine (V), arginine (R), alanine (A), and glutamate (E). In the arrayed mutagenesis of Q147 in zinc finger 2 **(Figure 13d)**, Q147A and Q147E completely stabilized the degron, suggesting that epistatic relationships between surrounding residues could render these amino acids permissive for degradation.

Figure 14 | A screen of ~6,500 unique C2H2 zinc fingers identifies novel sequences targeted for degradation by thalidomide, lenalidomide, and pomalidomide. a, Overview of the C2H2 zinc finger library screen. **b,** Fold-depletion of sequencing readcounts (DMSO/drug) and corresponding p values (empirical rank-sum test-statistic) for each of the 5,609 unique C2H2 zinc finger sequences with sufficient representation (raw read count >200 in all three replicates of infected cells) (experimental replicates = 3, labeled data points possess FDR<0.01 in at least one of the three drugs). **c,** HEK239T cells were infected in arrayed format with the top 24 zinc fingers in the EGFP/mCherry reporter vector and treated for 20h with a titration of thalidomide, lenalidomide, or pomalidomide. Displayed are representative graphs of the EGFP/mCherry ratios as assessed via flow cytometry. **d,** Hierarchical clustering (one minus Pearson correlation) of a heat map depicting the average fold depletion in sequencing readcounts (DMSO/drug) for zinc fingers with FDR<0.01 in any one of the three drugs (rank sum test statistic, empirical p-value). The residue predicted to interact with the drug is highlighted in red.

Figure 14 (Continued)



ZFP91 and ZNF692 are thalidomide analog-induced targets of the CRL4^{CRBN} ubiquitin ligase

We next sought to determine whether the C2H2 zinc finger degron-containing proteins identified in the screen were drug-induced targets of the CRL4^{CRBN} ubiquitin ligase. We first examined prior proteomic datasets from MM1S[15], or KG1 cells[20] treated with thalidomide and lenalidomide; of the ten genes possessing novel zinc finger degrons, five were detected, and of these, ZFP91 and ZNF692 demonstrated depletion in the presence of drug **(Appendix Figure 28)**

ZFP91 and ZNF692 each possess an array of five C2H2 zinc finger motifs. To confirm if ZFP91 and ZNF692 are degraded conditional on the presence of the identified C2H2 zinc finger degron we cloned wild-type and zinc finger deletion mutant constructs into the EGFP/mCherry reporter vector, transduced HEK293T cells, treated the cells with thalidomide, lenalidomide, or pomalidomide, and used flow cytometry to quantify the EGFP/mCherry ratio **(Figure 15a)**. The wild-type constructs were degraded in the presence of all three compounds, and this degradation was abrogated upon deletion of the fourth zinc finger, the degron identified in the screen. Also of note, the deletion of the zinc fingers 3 and 5 also appeared to stabilize the construct, suggesting that the context in which the degron is presented can affect degradation.

We next asked whether thalidomide analogs induced degradation of endogenous ZFP91 and ZNF692. Immunoblotting of KG1 cell lysates 20h after treatment with DMSO, thalidomide, lenalidomide, or pomalidomide revealed a reduction in the protein levels of ZFP91 and ZNF692 **(Figure 15b)**. In keeping with the response of their respective degrons to the three drugs, pomalidomide elicited the greatest degradation of ZFP91 and ZNF692, and both thalidomide and lenalidomide had almost no effect on ZFP91.

To determine if ZFP91 and ZNF692 bind CRBN in the presence of thalidomide analogs we immunoprecipitated HA-tagged ZFP91 and immunoblotted for CRBN. Again, in keeping with the response of the individual degrons, all three drugs increased ZFP91 and ZNF692 binding to cereblon, with pomalidomide exhibiting the greatest potency **(Figure 15c)**.

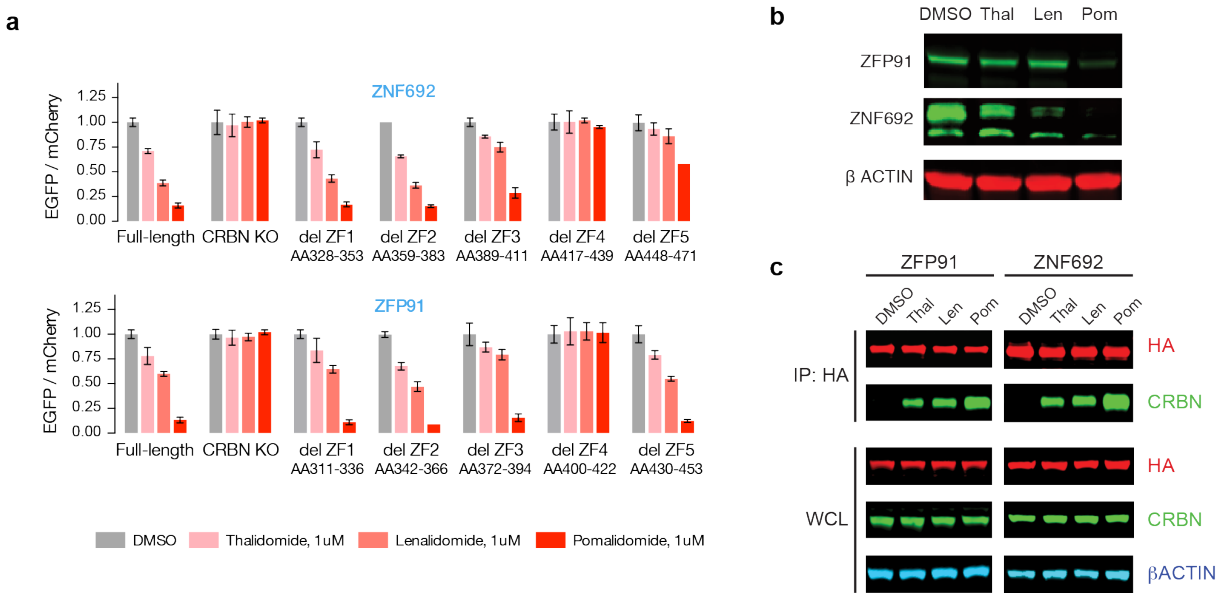


Figure 15 | ZFP91 and ZNF692 are degraded in the presence of thalidomide analogs conditional on the presence of CRBN and a C2H2 zinc finger degron. a, Wild-type and zinc finger deletion mutant constructs of ZFP91 and ZNF692 were cloned into the EGFP/mCherry reporter vector and expressed in wild-type or CRBN-KO HEK293T cells. Following treatment for 20h with DMSO, thalidomide, lenalidomide, or pomalidomide, flow cytometry was used to quantify the EGFP/mCherry ratio. **b,** KG1 cells were treated for 20h with DMSO, thalidomide, lenalidomide, or pomalidomide, after which their lysates were harvested and blotted for ZFP91 and ZNF692. **c,** HA-tagged ZFP91 or ZNF692 was overexpressed in HEK293T cells. Lysates were harvested and immunoprecipitation of the HA tag was carried out in the presence of DMSO or drug, after which we carried out immunoblotting for CRBN.

DISCUSSION

Thalidomide analogs modulate the activity of the $CRL4^{CRBN}$ E3 ubiquitin ligase to ubiquitinate and degrade two disease-relevant transcription factors, IKZF1 and IKZF3. Using a combination of saturation mutagenesis and crystallography, we have found that the β-hairpin of the second C2H2 zinc finger in IKZF1 and IKZF3 interacts with both cereblon and thalidomide analogs. Specifically, we identify Q147 as a residue which interacts with the phthaloyl group of pomalidomide. We utilize a screen of ~6,500 C2H2 zinc finger sequences to identify 10 novel targets of thalidomide analogs.

Traditionally degrons, the amino acid sequence required for protein degradation, have been identified by generating deletion mutants or alanine scanning. Saturation mutagenesis, a technique made possible by advances in oligonucleotide synthesis and next-generation sequencing, is a robust and informative approach for identifying the individual amino acids necessary for degradation and can be applied in other experimental contexts.

We present the crystal structure of IKZF1/IKZF3 zinc finger two in complex with pomalidomide and CRBN. Our identification of a C2H2 zinc finger degron for thalidomide analogs builds on two recent crystal structures of the drug-cereblon complex bound to two non-C2H2 zinc finger-containing proteins; lenalidomide-cereblon in complex with casein kinase 1 alpha [65] and CC-885-cereblon in complex with GSPT1 [66]. While there is limited sequence homology between the β-hairpin degrons, the shape of the loops are strikingly similar and the majority of the drug contacts are mediated by the peptide backbone rather than side chains. We do, however, note that the Q147 position forms a side-chain interaction with the phthaloyl group on pomalidomide, providing a structure-function correlate with implications for the design of novel analogs.

We identified 10 new C2H2 zinc fingers that are degraded by lenalidomide or pomalidomide. Among these, E4F1 and ZNF692 are annotated as transcription factors or as having sequence specific DNA recognition. Our results therefore indicate that thalidomide analogs have the ability to degrade a greater number of zinc finger transcription factors than previously appreciated. This finding suggests the possibility that both the therapeutic and undesirable effects of these compounds may be a result of the degradation of the proteins identified in our screen. Pomalidomide is the most potent and comprehensive drug in this regard, and its ability to appreciably degrade targets which lenalidomide only weakly influences may in part explain why patients who have relapsed on lenalidomide will respond to pomalidomide.

C2H2 zinc finger-containing proteins are the largest class of putative transcription factors, a group of proteins which have been difficult to inhibit with conventional small molecules. Screening diverse structural analogs could illuminate structure-function properties of cereblon modulators as well as the full extent of the “targetable proteome” which these drugs are capable of occupying.

METHODS

Cell Culture

HEK293T (Human Embryonic Kidney, Broad Institute Cell Line Repository) were grown in DMEM supplemented with Fetal Bovine Serum, Penicillin, and Streptomycin. KG1 cells (AML, Broad Institute Cell Line Repository) were grown in RPMI supplemented with Fetal Bovine Serum, Penicillin, and Streptomycin. Cell line identity was not verified.

Compounds

Thalidomide (SelleckChem, S1193)

Lenalidomide (SelleckChem, not current being sold)

Pomalidomide (SelleckChem, S1567)

Vectors

Saturation mutagenesis screen, full length ORFs - Artichoke (Addgene #73320)

C2H2 zinc finger library screen, validation of zinc fingers - Cilantro 2 (Addgene #74450)

Antibodies

ZFP91 (Bethyl Laboratories, Rabbit anti-ZFP91, A303-245A)

ZNF692 (Abcam, ab204595)

CRBN (Novus Biologicals, Rabbit anti-CRBN, NBP1-91810)

HA-affinity gel (Sigma, E6779)

Saturation Mutagenesis

A ssDNA oligonucleotide saturation mutagenesis library for IKZF3 aa130-189 was synthesized on an Agilent array, PCR amplified to form dsDNA, and then cloned into a bacterial vector in-

frame with puromycin resistance gene to select for library elements without frameshifting mutations. The resulting plasmid library was digested with restriction enzymes and cloned into the lentiviral “Artichoke” reporter vector (Addgene 73320).

We performed three technical replicates of the following:
On D0 HEK293T cells were infected at an infection efficiency of ~50%, selected in puromycin from D1-D4, plated for dosing on D7, then dosed for 20h with DMSO, 16uM thalidomide, 1uM lenalidomide, and 0.5uM pomalidomide. Cells then underwent FACS of the EGPF+ (DMSO) or EGFP- (thal, len, pom) cell population (BD FACS Aria).

Genomic DNA was harvested from the sorted cell populations. The IKZF3 mutagenesis library was PCR amplified from the genomic DNA, and in a two-step PCR reaction Illumina primers, adaptors, and sample barcodes were appended to the sequence. The library was then pooled and sequenced on the Illumina MisSeq.

Reads were aligned to the library to obtain read counts for each mutation and wild-type sequences. Only reads which contained a single mutation were analyzed. Read counts within each experimental condition were normalized and log2 transformed. To identify residues which did not tolerate mutations, we used an ANOVA statistical test to compare the relative frequency of reads containing mutations in the DMSO and drug treated conditions.

EGFP/mCherry Ratio Profiling

HEK293T cells were transduced with the EGFP/mCherry reporter vector. Following a 20h drug treatment, the EGFP and mCherry fluorescence was analyzed via flow cytometry. In Flojo (a flow cytometry analysis software), a parameter was derived which calculated the

EGFP/mCherry ratio on a single cell basis. The ratio for a given drug treated sample was normalized to the average of three DMSO-treated controls.

Purification of CRBN and DDB1$^{\Delta BPB}$

CRBN and DDB1$^{\Delta BPB}$ were co-purified as described in Fischer 2014 and Petzold 2016 [27,29]

Purification of IKZF1/IKZF3 constructs

2-4 L of Hi5 cells expressing truncated versions of StrepII-Avi-IKZF1 (Δ256–519, Δ197–238/Δ256–519 and Δ1–82/Δ197–238/Δ256–519; full-length IKZF1 forms aggregates during purification) were lysed in 120-240 ml of lysis buffer containing 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.25 mM TCEP, 1 mM PMSF and 1x protease inhibitor cocktail (Sigma-Aldrich). Following ultracentrifugation, the soluble fraction was passed over ~25 ml Strep-Tactin Sepharose (IBA) affinity resin (Sigma-Aldrich) and eluted in~150 ml of 50 mM Tris-HCl pH 8.0, 200 mM NaCl, 0.25 mM TCEP containing 2.5 mM D-desthiobiotin. Affinity tags were removed by overnight TEV protease treatment. The affinity-purified and TEV cleaved protein was further purified via ion exchange chromatography (low salt buffer: 50 mM Tris-HCl pH 6.0, 10 mM NaCl, 0.25 mM TCEP; high salt buffer: 50 mM Tris-HCl pH 6.0, 1 M NaCl, 0.25 mM TCEP; gradient to 70% high salt over 15 column volumes; 1CV=8ml Poros 50HQ or 50HS): cleaved protein was diluted 1:1 with low salt buffer and passed over Poros 50HQ. The flow through was passed over an 8 ml Poros 50HS column and eluted as described above. The Poros 50HS peak (Ikaros contructs only absorb poorly at 280 nm) was concentrated and subjected to size-exclusion chromatography in 50 mM HEPES pH 7.4, 200 mM NaCl and 0.25 mM TCEP.

Crystallization

70 µM DDB1dB-CRBNdN40 were mixed with 80 µM pomalidomide and ~350 µM cleaved Ikaros-ZF2 (141-173; GP163) in size-exclusion buffer. The complex crystallizes in an optimized

Morpheus HT condition (C2; mixed 1:1) containing 100 mM Morpheus buffer system 1 pH 6.5, 10 % NPS, 15% ethylene glycol, PEG 5K MME between 6.5 – 10 % (dependent on plates [96/24-well] and other unknown factors). Plates are kept at 19°C and crystals appear after 1-2 days. Crystals were flash frozen in liquid nitrogen in the presence of 20 % ethylene glycol.

C2H2 Zinc Finger Screen

To identify instances of C2H2 zinc fingers in the proteome we used the ScanProsite tool (http://prosite.expasy.org/scanprosite/) to search the proteome (UniprotKB, SwissProt Database, no alternative splice variants) for occurrences of the C2H2 motif (X-X-C-x(2,4)-C-x(3)-[LIVMFYWC]-x(8)-H-x(3,5)-H). 6819 instances of C2H2 ZFs were found (6572 unique) in 811 genes.

The amino acid sequences were reverse translated into codon-optimized DNA sequences, which were then appended with restriction enzyme sites and synthesized in a pooled format via array-based oligonucleotide synthesis (CustomArray, http://customarrayinc.com/). The library was converted to dsDNA via low cycle PCR amplification, digested, and cloned into the “Cilantro 2” lentiviral EGFP/mCherry reporter vector (Addgene 74450). Stbl4 electrocompetant bacteria were transformed with the library and plated onto large format LB carbepencillin plates. Resulting colonies were scraped, pooled, and maxiprepped (Qiagen). Lentivirus was generated from the C2H2 zinc finger library using HEK293T cells.

We performed three experimental replicates of the following:
HEK293T cells were infected with the library on D0, puro selected on D1-2, plated in the absence of puromycin for 24h, then were exposed for 20h to either DMSO, thalidomide (10uM), lenalidomide (1uM), or pomalidomide (1uM). Cells were harvested and then underwent flow

cytometry based sorting for EGFP+ and EGFP- cells (SONY cell sorter). Genomic DNA was harvested from the sorted cells as well as a baseline control of unsorted DMSO-treated cells.

The zinc finger sequences were amplified from the genomic DNA with appended Illumina primers and barcodes. As a control, the original plasmid DNA library was also amplified. The resulting Illumina library was sequenced on the Illumina NextSeq.

Analysis of Sequencing Data

To reduce the number of false positives caused by library elements with low representation, zinc fingers with less than 200 raw reads in the DMSO-treated unsorted control were removed from the library. This threshold reduced the library size by ~15%, from 6819 to 5856.

Reads were aligned to the library using bowtie and read counts for each experimental group were normalized (reads per million) and log2 transformed. The fold depletion (log2 normalized read counts) of a given zinc finger from the EGFP+ population upon drug treatment was calculated for each replicate. ZFs were ranked on basis of fold-depletion, and the three experimental replicates were integrated by determining rank-sum empirical p-values, which were in turn used to calculate a false-discovery rate for each zinc finger. An FDR cutoff of 0.01 was used to identify “hits”.

Immunoblots

KG1 cells were treated with either DMSO or 1uM of thalidomide, lenalidomide, or pomalidomide, after which protein lysates were harvested and run on a polyacrylamide gel, transferred to nitrocellulose membrane, and blotted for ZFP91 and ZNF692. Immunoprecipitation was achieved by overexpressing HA-tagged ZFP91 and ZNF692 in KG1 cells. Protein lysates were collected and then incubated with anti-HA sepharose beads in the presence of 1uM of

thalidomide, lenalidomide, or pomalidomide overnight at 4C. Proteins were eluted using HA-peptide and incubating at 30C for 5 minutes.

Chapter 4: Discussion

The remarkable historical arc of thalidomide and its derivatives has continued with the recent discovery that these drugs induce the ubiquitination and proteasomal degradation of targets of the $CRL4^{CRBN}$ ubiquitin ligase. Indeed, their molecular mechanism of action has gained interest as a novel pharmacologic paradigm which makes use of ubiquitin ligases to degrade targets which would otherwise be difficult to inhibit with conventional small molecules. The goal of this thesis has therefore been to leverage screening-based technologies to better understand how thalidomide analogs modulate $CRL4^{CRBN}$ function.

Implications of $CRL4^{CRBN}$ regulation for drug development and resistance

Chapter 2 detailed the use of a genome-scale CRISPR-Cas9 screen to identify the molecular machinery which is required for lenalidomide-induced degradation of IKZF3. The screen identified the neddylation enzymes UBE2M and the COP9 signalosome and the E2 enzymes UBE2D3 and UBE2G1, and we go on to describe the molecular mechanisms by which these proteins influence $CRL4^{CRBN}$.

As discussed in chapter 2, the finding that $CRL4^{CRBN}$ requires additional factors for its function holds implications for our understanding of how cullin-ring ligases are regulated as well as acquired resistance to lenalidomide. However, this work holds broader relevance for a growing class of drugs, including novel thalidomide analogs, thalidomide-based heterobifunctional compounds, and novel small molecules such as indisulam which rely on different CRLs [41,43,58,66,67]. Specifically, our results suggest the possibility of a common set of resistance mechanisms, either in the form of mutations in substrate receptors to which multiple drugs bind,

or mutations in genes of the neddylation cycle which regulate cullin-ring ligases. Additionally, our discovery that more proteins than previously known regulate $CRL4^{CRBN}$ function could facilitate the development of panels of genes whose expression levels may be predictive of response to therapy or the development of acquired resistance.

Genome-scale CRISPR-Cas9 screens to discover drug mechanism of action

Our use of a genome scale CRISPR-Cas9 screen to successfully recapitulate the mechanism of action of lenalidomide additionally highlights the potential of positive selection genome-scale functional genetic screens to define the molecular mechanism of compounds. Indeed, many bioactive compounds lack a defined molecular mechanism of action, knowledge which can benefit efforts to optimize potency and selectivity, appropriate the drug to other disease contexts, as well as an understanding of the pathobiology underlying a malignancy. Large-scale, positive selection CRISPR-Cas9 screening of drugs with unknown function in sensitive cancer cell lines would be a viable approach to illuminate novel pharmacology and biology, in particular in the case of drugs with gain-of-function mechanisms.

Novel targets of thalidomide analogs may play role in therapeutic efficacy

In chapter 3 we used saturation mutagenesis and crystallography to characterize the C2H2 zinc finger degron of IKZF1 and IKZF3, a discovery which in turn facilitated a screen of ~6,500 C2H2 zinc fingers and the subsequent identification of novel targets of thalidomide, lenalidomide, and pomalidomide. The identification of novel targets such as ZFP91 and ZNF692 is of significance for these results establish that thalidomide analogs target a greater-than-previously-appreciated number of proteins for degradation.

In one aspect, this raises the question of whether the therapeutic properties of these compounds is analogous to kinases, whose clinical effects are derived in part from their lack of specificity. Indeed, our assays demonstrate that pomalidomide is a more potent drug with regards to C2H2 zinc-finger-containing proteins and that it facilitates the degradation of a number of targets which pomalidomide and thalidomide are unlikely to degrade to a level that is biologically relevant. Multiple myeloma patients who have relapsed on regimens containing lenalidomide are known to respond to pomalidomide, perhaps as a result of its ability to target a greater number of degrons with increased potency. The therapeutic relevance of these novel targets can be addressed by cell viability assays in which you either introduce gRNAs targeting known degraded targets or by editing the endogenous degron of a protein of interest to block its degradation. Similarly, it is possible that these novel targets could be responsible for the numerous side effects of thalidomide analogs.

Relatedly, it will be necessary to understand the biological role of novel thalidomide analog targets. Many of the 10 genes identified by the zinc finger library screen lack well-annotated functions. Two of the targets identified in the screen have promising, if limited biological annotation: ZFP91 is described as an atypical E3 ligase which ubiquitinates the NFKB pathway member NIK in order to promote NFKB signaling[68,69]. If verified, the degradation of ZFP91 could explain the ability of thalidomide analogs to regulate NFKB signaling and TNFα secretion. Also notable is ZNF827, which is an essential factor for alternative lengthening of telomeres[70], a molecular process which is thought to be exclusively utilized by a subset of cancers and thus is an attractive therapeutic target.

Novel structural analogs of thalidomide may target different repertoires of targets

The full repertoire of substrates which can be targeted for $CRL4^{CRBN}$ degradation by structural analogs of thalidomide has yet to be identified. Indeed, it is unclear to what extent CRBN's binding surface will limit the spectrum of substrates targeted by glutarimide-based compounds. One of the principal goals of future work with thalidomide analogs will therefore be to use mass spectrometry-based proteomics in combination with the C2H2 zinc finger library to systematically test analogs representative of diverse structural classes. Such an approach will define the targetable "space" for this family of cereblon modulators, and in combination with crystallography, will also facilitate the development of structure-function annotation of these drugs. Eventually we hope to enable rational engineering of cereblon modulators with selectivity and specificity for targets of interest.

The notion that structural analogs may target different repertoires of targets has borne out with the novel cereblon modulator CC-885. CC-885 is a novel thalidomide analog which is analogous to lenalidomide, however instead of an amide group it possess a linker connecting to a third phenyl ring. As a result of these additional functional groups, CC-885 uniquely recruits the translation termination factors GSPT1 (eRF3a) to CRBN, mediating $CRL4^{CRBN}$-dependent ubiquitination of GSPT1[66]. Crystallography of the tertiary structure demonstrates that GSPT1, similar to casein kinase and IKZF1/3, interacts directly with both CC-885 and CRBN via a glycine containing beta-beta loop. In keeping with its differential ability to target GSPT1, CC-885 exhibits a differential profile of tumor cell line killing in comparison to its more structurally homologous predecessors, thalidomide, lenalidomide, and pomalidomide.

It is therefore clear from the example of CC-885 that a more comprehensive exam of diverse structural analogs may yield compounds with the ability to degrade novel targets, with a concomitant diversification of therapeutic applications.

Systematic approaches to identifying novel modulators of ubiquitin ligase function

The question of whether other drugs function in an analogous fashion to thalidomide analogs has in part been answered by the recent discovery that the anti-cancer aryl sulfonamides indisulam, CQS, and tasisulam induce the $CRL4^{DCAF15}$-dependent ubiquitination and degradation of the pre-mRNA splicing factor RBM39[58]. Currently indisulam has undergone phase I and II clinical trials in patients with advanced solid malignancies and has been found to induce clinical responses and stabilize disease in 17-36% of patients. This discovery lends itself to a number of the approaches which we have undertaken: mass-spectrometry proteomics, genome-scale resistance screens, saturation mutagenesis and degron library screening, crystallography, and screening of structural analogs to being to ascertain the structure-function relationships.

The identification of a second family of drugs with an analogous mechanism of action justifies a systematic research program towards the identification of additional compounds with similar abilities to modulate the function of one of the ~600 ubiquitin ligases encoded in the genome.

Such an initiative will require a diverse set of approaches which can be divided into two categories:

High-throughput screens of existing bioactive small-molecules

1. Mass spectrometry-based proteomic analysis of existing bio-active small molecules on cell lines representative of canonical tissue types.

2. Genome-scale CRISPR-Cas9 positive selection screens to identify small molecules which rely on substrate receptors and associated ubiquitin ligase machinery for their anti-proliferative effects

3. Genotype-phenotype correlations for bioactive compounds across a diverse set of cancer cell lines to identify small molecules whose cell killing is correlated with the expression levels of substrate receptors and associated ubiquitin ligase machinery.

Identification of E3-substrates pairs and screening for small molecule stabilizers

1. In vitro FRET-based screening of small molecule libraries to identify compounds which stabilize the interaction between substrates of interest and their known E3 ligases (ex. MYC and FBWX7).

2. Assemble reporters for targets of interest and screen a gRNA library of ubiquitin ligase substrate receptors to identify E3 ligases controlling target degradation.

3. Engineer cells in which ubiquitin ligases are either genetically inactivated or overexpressed, and use mass-spectrometry proteomics to measure changes in the proteome of these cells in order to identify candidate substrates.

A coordinated approach across academic and industry to answering the above questions stands to return a number of small molecules with the ability to expand the druggable genome.

Regulation of ubiquitin ligases by endogenous small molecules

Lastly, a question raised by the discovery of drugs such as thalidomide analogs and the indisulam family of compounds is whether these drugs are co-opting an endogenous mechanism by which ubiquitin ligases are regulated by small molecules. Indeed, such a mechanism exists in plants: the small molecule auxin binds the substrate receptor of the SCF^{TIR1} E3 ubiquitin ligase, and like thalidomide analogs, mediates a tertiary interface with both the TIR1 substrate receptor and the transcriptional repressors which it targets for ubiquitination and degradation.

The CRBN-IKZF1/3 interaction has been observed in the absence of drug in immunoprecipitation of cereblon from cell lysates[15], however in vitro cereblon's interaction with IKZF1/3 requires thalidomide or its analogs. This discrepancy could point to the existence of a peptide or post-translational modification which could stabilize the interaction, however steric constraints and the identity of the amino acids within the degron and cereblon suggest that this is not the case. The discovery of endogenous small molecules controlling ubiquitin ligase function and their biosynthetic pathways would be a first-in-class mechanism in mammals and would open a window to drug discovery.

Conclusion

Thalidomide analogs are an promising class of drugs with the ability to mediate $CRL4^{CRBN}$-dependent degradation of the transcription factors IKZF1 and IKZF3, as well as the Wnt pathway regulator casein kinase 1 alpha. Their mechanism of action heralds what is likely to be a growing class of small molecules of significant therapeutic consequence which modulate ubiquitin ligase function. The thesis presented herein not only further characterizes thalidomide-analog induction of $CRL4^{CRBN}$ activity, but also, as part of a larger body of literature on this

family of compounds, presents a road map of principles and techniques which can be used on novel modulators of ubiquitin ligase function. Research programs which pursue these classes of drugs are likely to yield both significant biological insight into E3 ligase function and regulation, as well as clinically relevant compounds addressing the pressing need for small molecules with novel properties.

Appendix

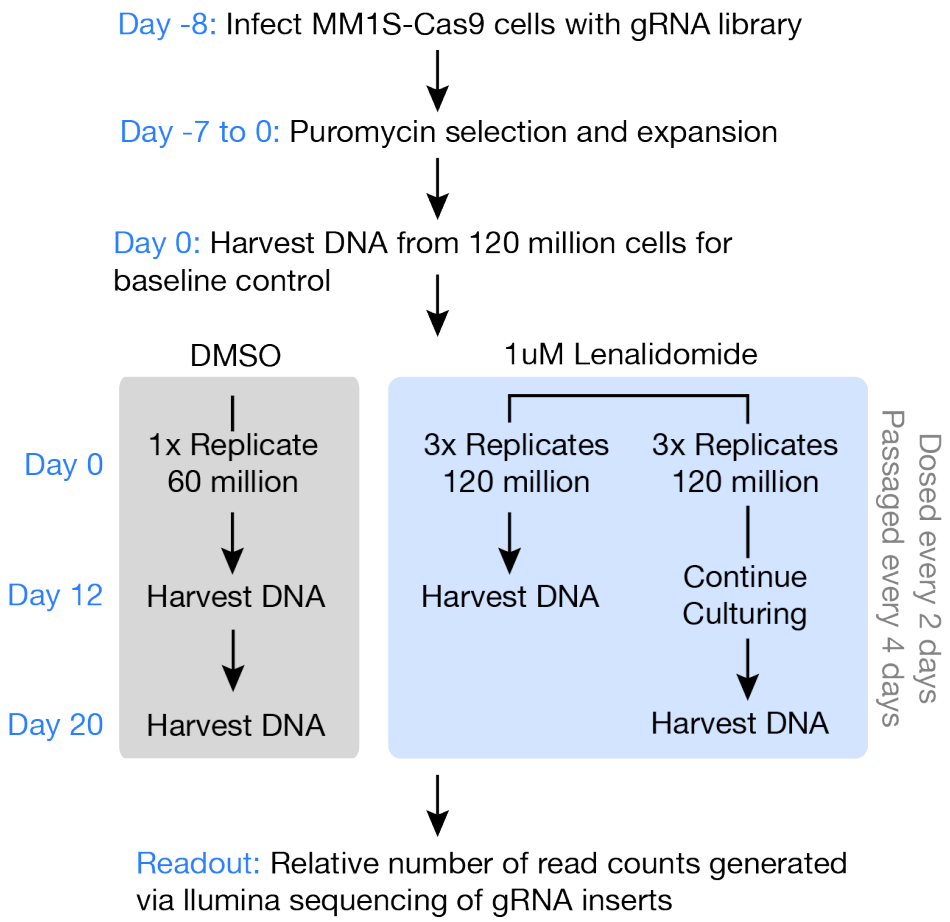


Figure 16 | Detailed workflow of the genome-scale positive selection CRISPR-Cas9 screen.

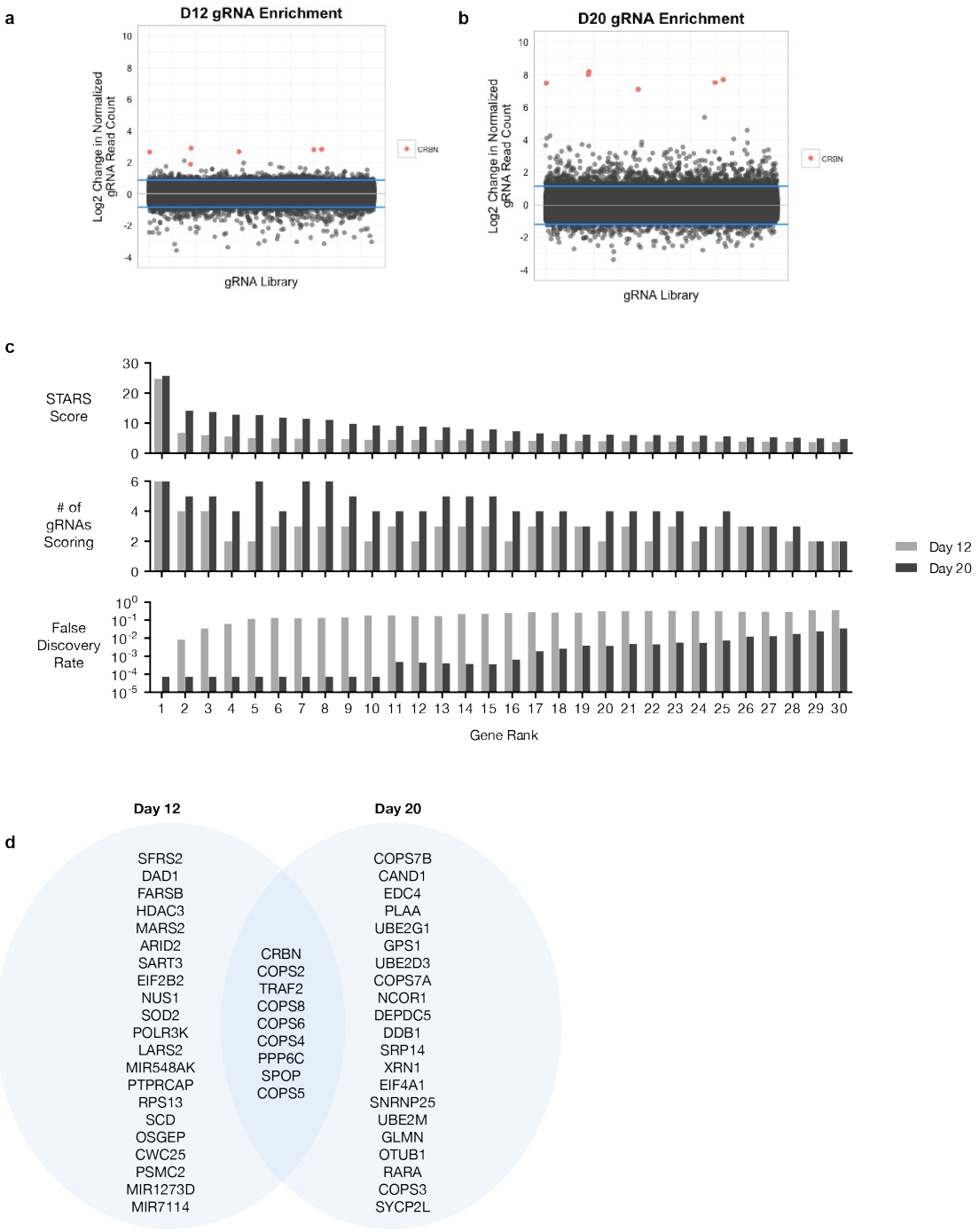


Figure 17 | Comparison of STARS algorithm gene rankings at day 12 and day 20 of the screen. a, Day 12 gRNA average log2 normalized fold-changes (Len/DMSO). **b,** Day 20 gRNA average log2 normalized fold-changes (Len/DMSO). **c,** Day 12 and day 20 values for the STARS score, number of gRNAs considered to have scored, and the FDR (FDR for day 20 hits 1-10 is <7.15e-5). **d,** Comparison of genes which scored at day 12 and day 20.

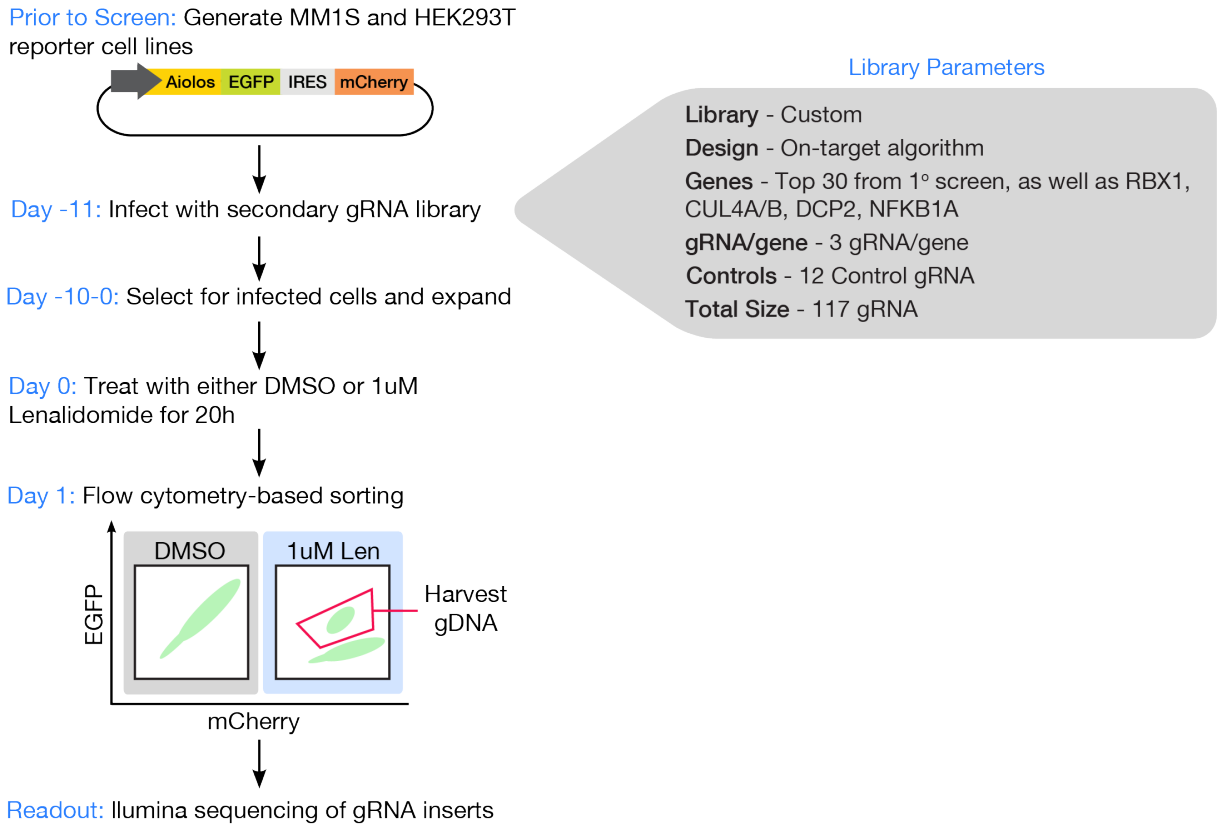


Figure 18 | Workflow of the IKZF3-EGFP Reporter Screen.

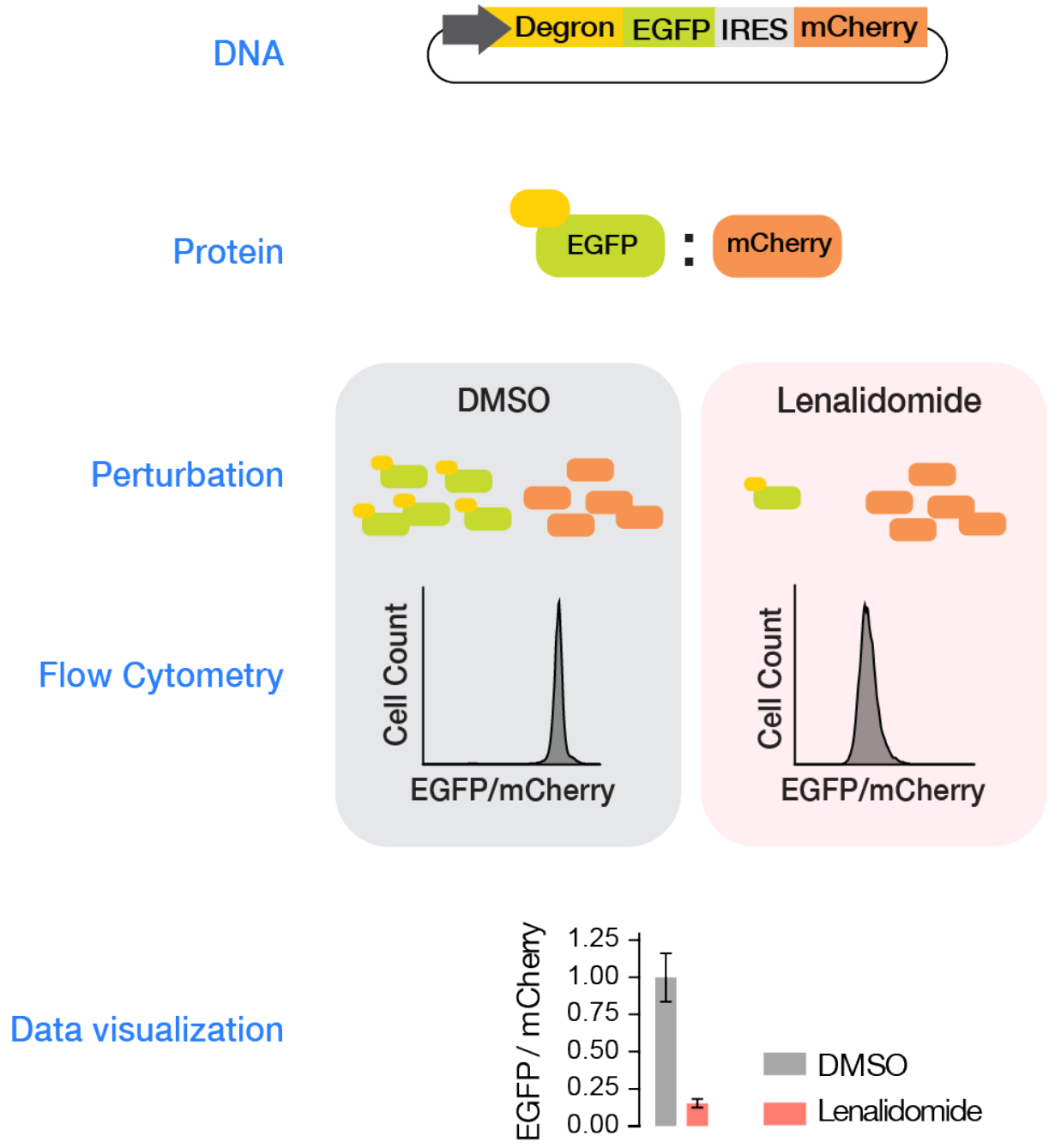


Figure 19 | How the IKZF3-EGFP/mCherry reporter works.

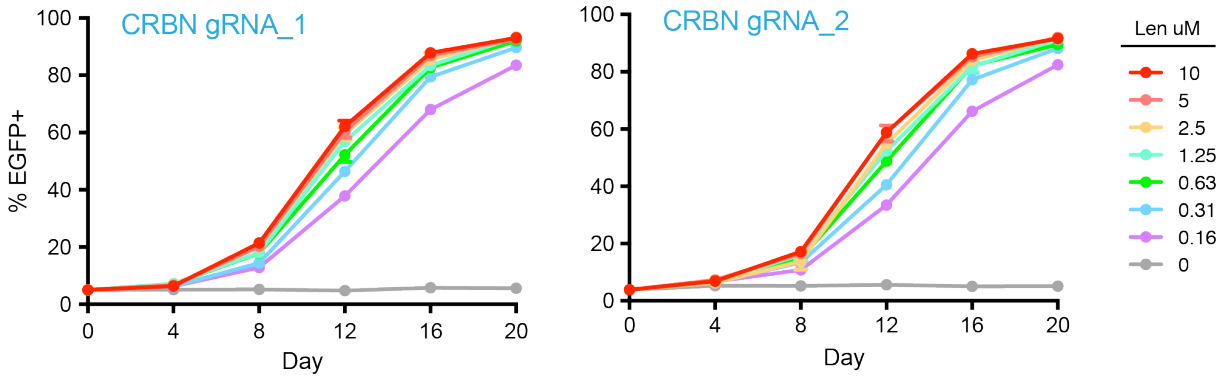


Figure 20 | MM1.S cells infected with gRNAs targeting CRBN exhibit a competitive advantage in a range of doses of lenalidomide. MM1.S cells were transduced with lentiviral vectors either CRBN gRNA and EGFP or a non-targeting control gRNA. After >8 days the cells were mixed at a 5:95 ratio respectively and grown in the presence of a range of concentrations of lenalidomide. The percentage of cells which were EGFP+ was assessed every four days via flow cytometry.

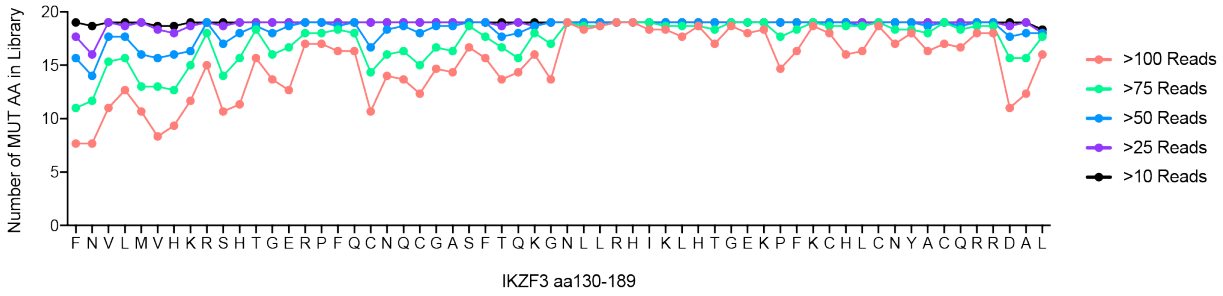


Figure 21 | Number of mutant amino acids represented in the saturation mutagenesis library.

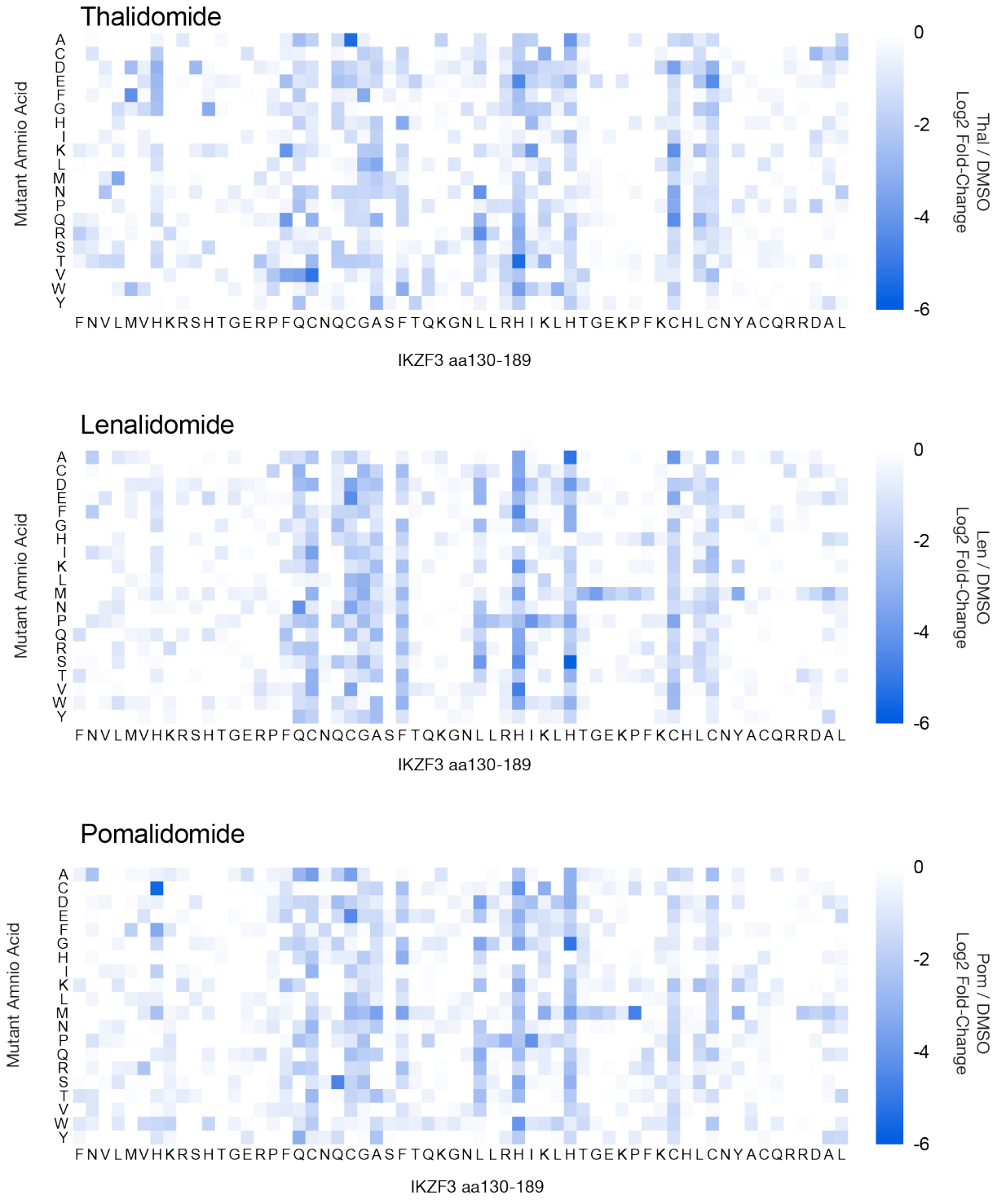


Figure 22 | Log2 fold change in mutant read counts for thalidomide, lenalidomide, and pomalidomide from the saturation mutagenesis screen.

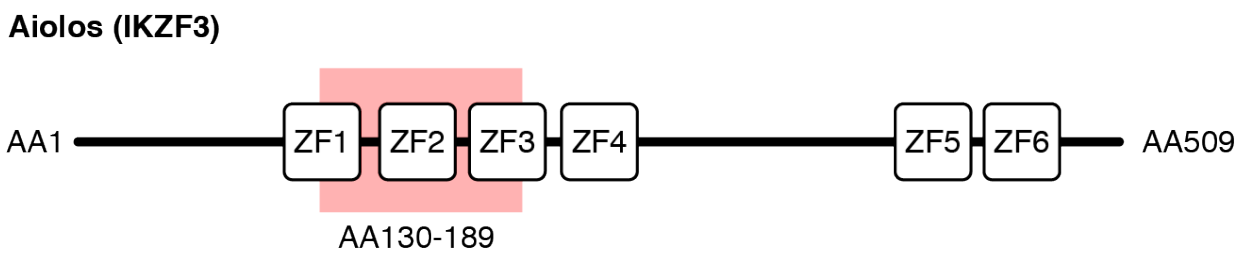


Figure 23 | Location of C2H2 zinc finger sequences in IKZF3.

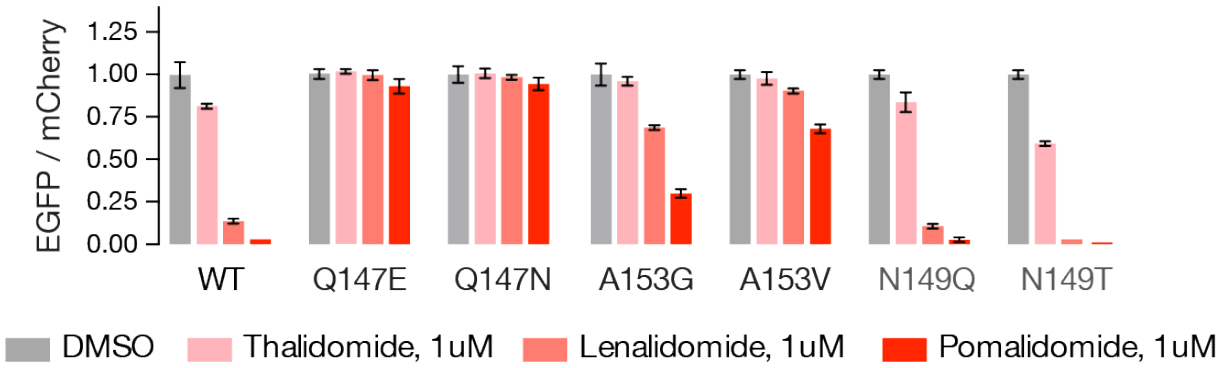


Figure 24 | Mutations at IKZF3 Q147 and A153 impair thalidomide analog-induced degradation. HEK293T cells expressing variants of IKZF3 ZF2 in the EGFP/mCherry reporter vector were treated for 20 hours with DMSO, thalidomide, lenalidomide, or pomalidomide, after which we assessed the EGFP/mCherry ratio using flow cytometry. Values are normalized to the average of three DMSO-treated replicates.

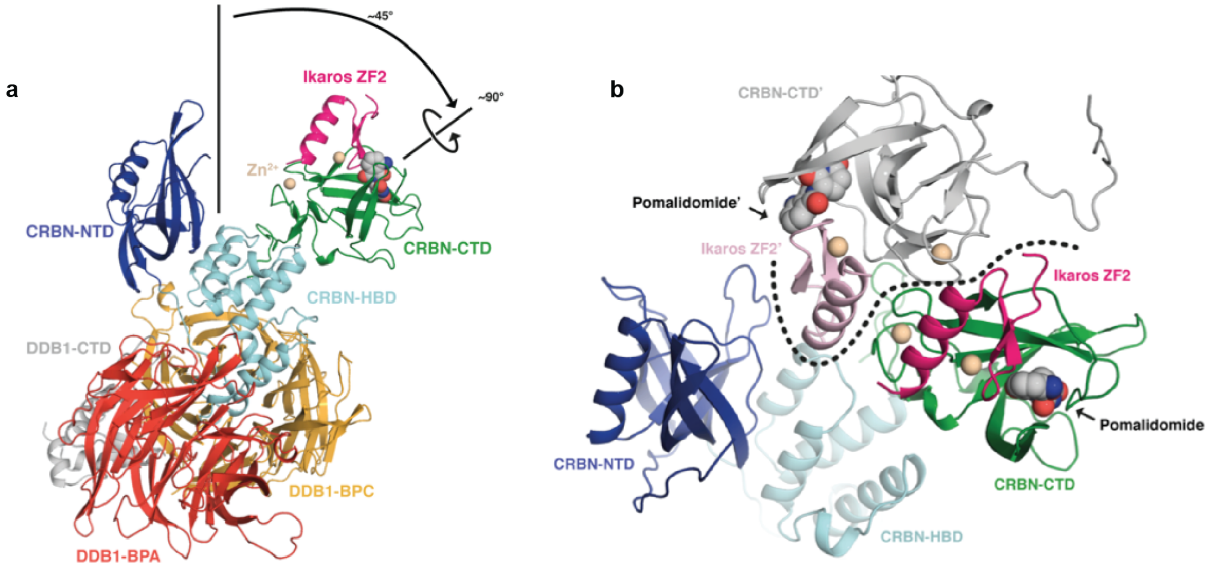


Figure 25 | Dislocation of the CRBN NTD and CTD due to crystal packing. The crystal structure revealed a novel orientation to CRBN; In contrast to prior structures[27-29,66], CRBN's C-terminal domain (CRBN-CTD) has undergone a rigid body motion, dislocating from the N-terminal domain (CRBN-NTD) by approximately 45^{o} and undergoing a 90^{o} rotation. The CTD remains loosely connected to the helical bundle domain (CRBN-HBD) by a short, flexible linker**.** There was no evidence that zinc finger 2 binding induced this conformational change; instead, an examination of the crystal lattice revealed that the zinc finger from an adjacent complex is inserted between the CRBN-NTD and CRBN-CTD, suggesting that the dislocation of the NTD and CTD may be an artifact of crystal packing. However, this finding does not preclude the possibility that this is a biologically plausible confirmation with relevance in the context of other substrates.

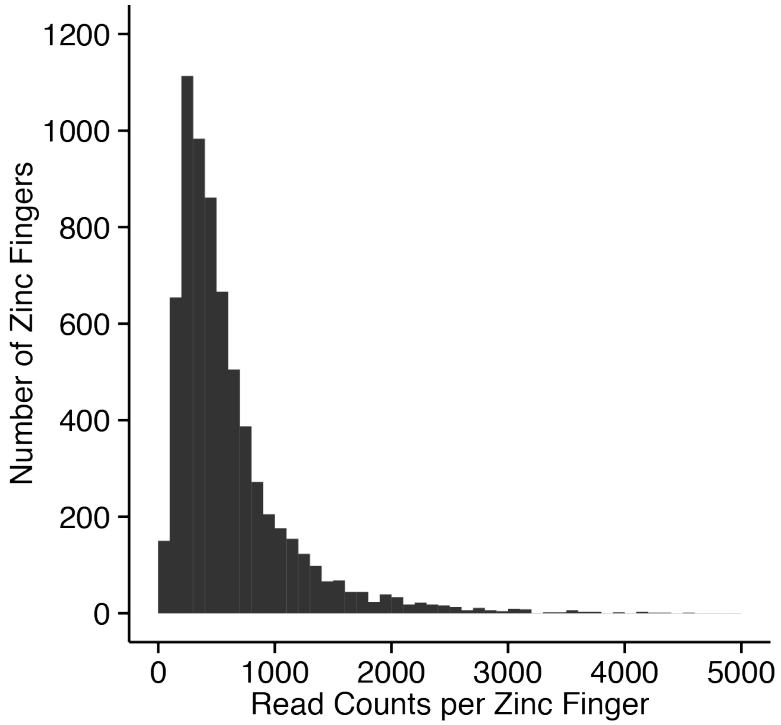


Figure 26 | Quality control of the C2H2 zinc finger library. After being cloned into the EGFP/mCherry reporter vector the C2H2 zinc finger sequences were PCR amplified with Illumina adaptors and sequenced on the Illumina MiSeq.

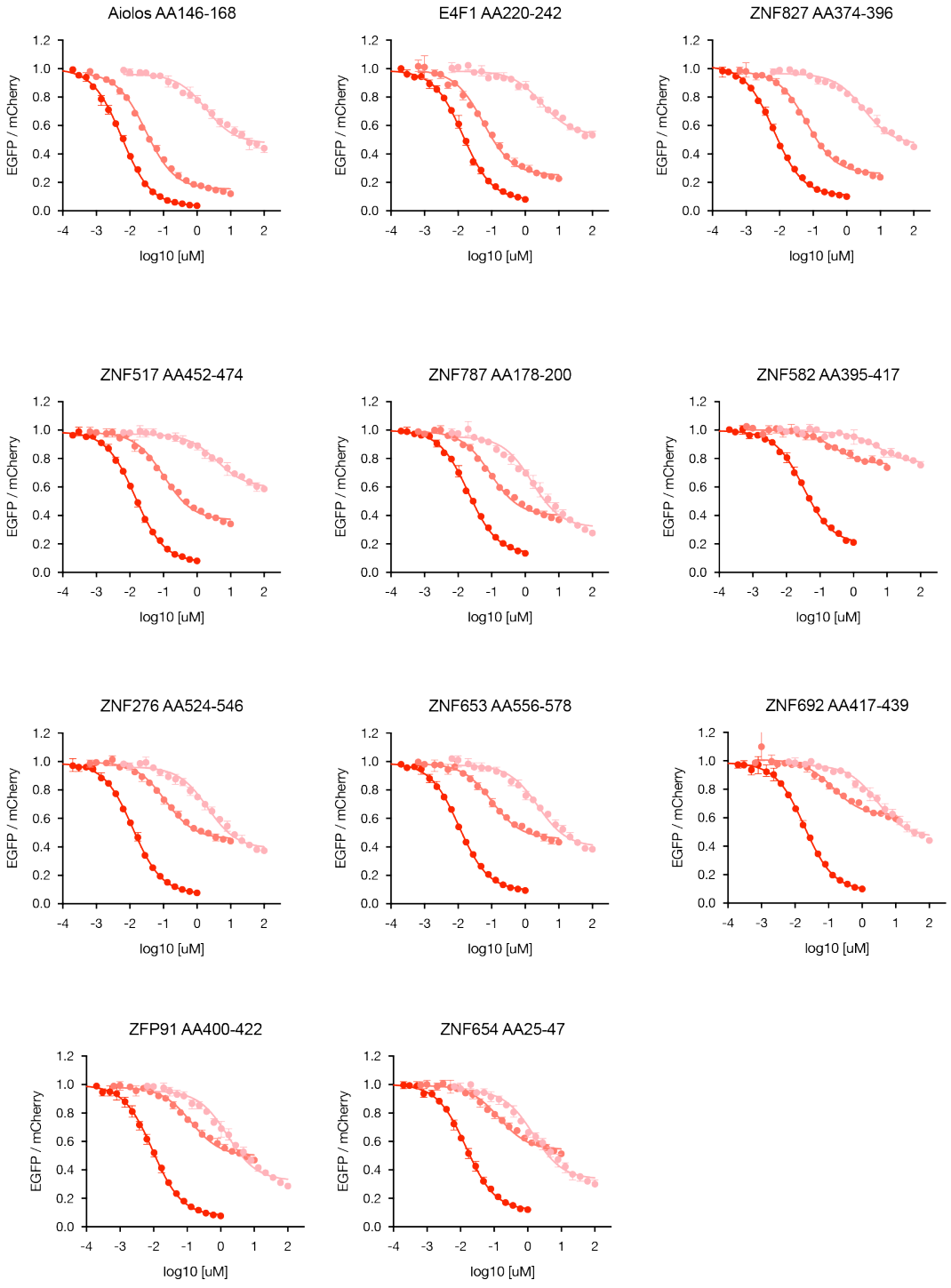


Figure 27 | 11 C2H2 zinc fingers identified in the screen are degraded by thalidomide, lenalidomide, and pomalidomide. HEK293T cells were transduced with the 11 C2H2 zinc fingers identified in the screen in the EGFP/mCherry reporter vector. The cells were treated with a titration of thalidomide, lenalidomide, and pomalidomide for 20h, then the EGFP/mCherry ratio was analyzed using flow cytometry. (experimental replicates=3, technical replicates=3, points are average of experimental replicates, error bars indicate range).

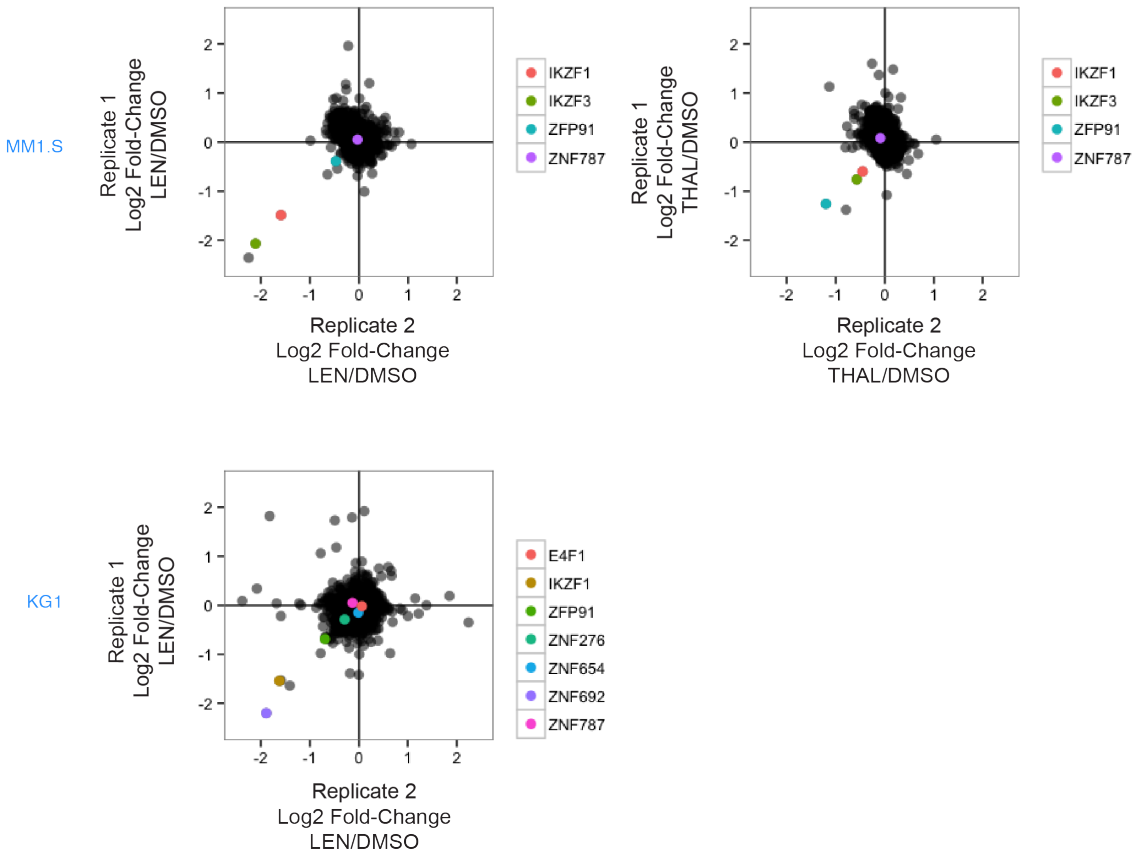


Figure 28 | ZFP91 and ZNF692 show evidence of degradation in prior proteomic datasets. MM1.S and KG1 cells were treated for 12h with 1uM lenalidomide or 10uM thalidomide after which protein lysates were harvested for SILAC-based quantitative mass spectrometry

References

1. Stephens, T. & Brynner, R. *Dark Remedy*. (Basic Books, 2001).

2. Vargesson, N. Thalidomide-induced teratogenesis: History and mechanisms. *Birth Defect Res C* **105,** 140–156 (2015).

3. Mcbride, W. G. THALIDOMIDE AND CONGENITAL ABNORMALITIES. *The Lancet* **278,** 1358 (1961).

4. Lenz, W. & Knapp, K. Foetal malformations due to thalidomide. *Problems of Birth Defects* (1962). doi:10.1007/978-94-011-6621-8_29

5. Sheskin, J. Thalidomide in the treatment of lepra reactions. *Clin Pharmacol Ther* **6,** 303–306 (1965).

6. Iyer, C. G. S. *et al.* WHO co-ordinated short-term double-blind trial with thalidomide in the treatment of acute lepra reactions in male lepromatous patients. *Bulletin of the World Health Organization* **45,** 719 (1971).

7. Sampaio, E. P., Sarno, E. N., Galilly, R., Cohn, Z. A. & Kaplan, G. Thalidomide selectively inhibits tumor necrosis factor alpha production by stimulated human monocytes. *J Exp Med* **173,** 699–703 (1991).

8. Muller, G. W. *et al.* Amino-substituted thalidomide analogs: Potent inhibitors of TNF-α production. *Bioorganic & Medicinal Chemistry Letters* **9,** 1625–1630 (1999).

9. D'Amato, R. J., Loughnan, M. S., Flynn, E. & Folkman, J. Thalidomide is an inhibitor of angiogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **91,** 4082–4085 (1994).

10. Singhal, S. *et al.* Antitumor Activity of Thalidomide in Refractory Multiple Myeloma. *N. Engl. J. Med.* **341,** 1565–1571 (1999).

11. Raza, A. *et al.* Thalidomide produces transfusion independence in long-standing refractory anemias of patients with myelodysplastic syndromes. *Blood* **98,** 958–965 (2001).

12. List, A. *et al.* Efficacy of Lenalidomide in Myelodysplastic Syndromes. *N. Engl. J. Med.* **352,** 549–557 (2005).

13. List, A. *et al.* Lenalidomide in the Myelodysplastic Syndrome with Chromosome 5q Deletion. *N. Engl. J. Med.* **355,** 1456–1465 (2006).

14. Ito, T. *et al.* Identification of a Primary Target of Thalidomide Teratogenicity. *Science* **327,** 1345–1350 (2010).

15. Krönke, J. *et al.* Lenalidomide Causes Selective Degradation of IKZF1 and IKZF3 in Multiple Myeloma Cells. *Science* **343,** 301–305 (2014).

16. Lu, G. *et al.* The Myeloma Drug Lenalidomide Promotes the Cereblon-Dependent Destruction of Ikaros Proteins. *Science* **343,** 305–309 (2014).

17. Gandhi, A. K. *et al.* Immunomodulatory agents lenalidomide and pomalidomide co-stimulate T cells by inducing degradation of T cell repressors Ikaros and Aiolos via modulation of the E3 ubiquitin ligase complex CRL4(CRBN.). *British Journal of Haematology* (2013). doi:10.1111/bjh.12708

18. Guharoy, M., Bhowmick, P., Sallam, M. & Tompa, P. Tripartite degrons confer diversity and specificity on regulated protein degradation in the ubiquitin-proteasome system. *Nature Communications* **7,** 1–13 (1AD).

19. Petroski, M. D. & Deshaies, R. J. Function and regulation of cullin–RING ubiquitin ligases. *Nat. Rev. Mol. Cell Biol.* **6,** 9–20 (2005).

20. Krönke, J. *et al.* Lenalidomide induces ubiquitination and degradation of CK1α in del(5q) MDS. *Nature* **523,** 183–188 (2015).

21. Georgopoulos, K. *et al.* The Ikaros gene is required for the development of all lymphoid lineages. *Cell* **79,** 143–156 (1994).

22. Wang, J. H. *et al.* Aiolos regulates B cell activation and maturation to effector state. *Immunity* **9,** 543–553 (1998).

23. Rebollo, A. & Schmitt, C. Ikaros, Aiolos and Helios: transcription regulators and lymphoid malignancies. *Immunol. Cell Biol.* **81,** 171–175 (2003).

24. Liu, C. *et al.* Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. *Cell* **108,** 837–847 (2002).

25. Luis, T. C. *et al.* Canonical Wnt Signaling Regulates Hematopoiesis in a Dosage-Dependent Fashion. *Cell Stem Cell* **9,** 345–356 (2011).

26. Schneider, R. K. *et al.* Role of Casein Kinase 1A1 in the Biology and Targeted Therapy of del(5q) MDS. *Cancer Cell* **26,** 509–520 (2014).

27. Fischer, E. S. *et al.* Structure of the DDB1-CRBN E3 ubiquitin ligase in complex with thalidomide. *Nature* 1–16 (2014). doi:10.1038/nature13527

28. Chamberlain, P. P. *et al.* Structure of the human Cereblon-DDB1-lenalidomide complex reveals basis for responsiveness to thalidomide analogs. *Nat. Struct. Mol. Biol.* **21,** 803–809 (2014).

29. Petzold, G., Fischer, E. S. & Thomä, N. H. Structural basis of lenalidomide-induced CK1α degradation by the CRL4CRBN ubiquitin ligase. *Nature* **532,** 127–130 (2016).

30. Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How many drug targets are there? *Nature Reviews Drug Discovery* **5,** 993–996 (2006).

31. Nowell, P. & Hungerford, D. A minute chromosome in human chronic granulocytic leukemia. *Landmarks in Medical Genetics: …* (2004).

32. Rowley, J. D. A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature* (1973).

33. Daley, G. Q., Van Ettan, R. A. & Baltimore, D. Induction of chronic myelogenous leukemia in mice. *Science* (1990).

34. Hughes, T. P. & Ross, D. M. Targeted therapies: Remembrance of things past—discontinuation of second-generation TKI therapy for CML. *Nature Reviews Clinical Oncology* **14,** 201–202 (2017).

35. Wells, J. A. & McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein—protein interfaces. *Nature* **450,** 1001–1009 (2007).

36. Beroukhim, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463,** 899–905 (2010).

37. Sakamoto, K. M. *et al.* Protacs: chimeric molecules that target proteins to the Skp1-Cullin-F box complex for ubiquitination and degradation. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 8554–8559 (2001).

38. Sakamoto, K. M. *et al.* Development of Protacs to target cancer-promoting proteins for ubiquitination and degradation. *Mol. Cell Proteomics* **2,** 1350–1358 (2003).

39. Schneekloth, J. S. *et al.* Chemical Genetic Control of Protein Levels: Selective in Vivo Targeted Degradation. *J. Am. Chem. Soc.* **126,** 3748–3754 (2004).

40. Schneekloth, A. R., Pucheault, M., Tae, H. S. & Crews, C. M. Targeted intracellular protein degradation induced by a small molecule: En route to chemical proteomics. *Bioorganic & Medicinal Chemistry Letters* **18,** 5904–5908 (2008).

41. Lu, J. *et al.* Hijacking the E3 Ubiquitin Ligase Cereblon to Efficiently Target BRD4. *Chemistry & Biology* **22,** 755–763 (2015).

42. Zengerle, M., Chan, K.-H. & Ciulli, A. Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. *ACS Chem. Biol.* **10,** 1770–1777 (2015).

43. Winter, G. E. *et al.* Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science* **348,** 1376–1381 (2015).

44. Weber, D. M. *et al.* Lenalidomide plus Dexamethasone for Relapsed Multiple Myeloma in North America. *N. Engl. J. Med.* **357,** 2133–2142 (2007).

45. Dimopoulos, M. *et al.* Lenalidomide plus Dexamethasone for Relapsed or Refractory Multiple Myeloma. *N. Engl. J. Med.* **357,** 2123–2132 (2007).

46. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343,** 84–87 (2014).

47. Duda, D. M. *et al.* Structural regulation of cullin-RING ubiquitin ligase complexes. *Current Opinion in Structural Biology* **21,** 257–264 (2011).

48. Duda, D. M. *et al.* Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation. *Cell* **134,** 995–1006 (2008).

49. Cope, G. A. & Deshaies, R. J. Targeted silencing of Jab1/Csn5 in human cells downregulates SCF activity through reduction of F-box protein levels. *BMC Biochem.* **7,** 1 (2006).

50. Denti, S., Fernandez-Sanchez, M. E., Rogge, L. & Bianchi, E. The COP9 signalosome regulates Skp2 levels and proliferation of human cells. *J. Biol. Chem.* **281,** 32188–32196 (2006).

51. Ye, Y. & Rape, M. Building ubiquitin chains: E2 enzymes at work. *Nat. Rev. Mol. Cell Biol.* **10,** 755–764 (2009).

52. Raman, M., Havens, C. G., Walter, J. C. & Harper, J. W. A Genome-wide Screen Identifies p97 as an Essential Regulator of DNA Damage-Dependent CDT1 Destruction. *Molecular Cell* **44,** 72–84 (2011).

53. Eichner, R. *et al.* Immunomodulatory drugs disrupt the cereblon-CD147-MCT1 axis to exert antitumor activity and teratogenicity. *Nature Medicine* 1–13 (2016). doi:10.1038/nm.4128

54. Thakurta, A. *et al.* Absence of mutations in cereblon (CRBN) and DNA damage-binding protein 1 (DDB1) genes and significance for IMiD therapy. *Leukemia* (2013). doi:10.1038/leu.2013.315

55. Jonasova, A. *et al.* High level of full length cereblon mRNA in lower risk myelodysplastic syndrome with isolated 5q deletion is implicated in the efficacy of lenalidomide. *Eur. J. Haematol.* (2014). doi:10.1111/ejh.12457

56. Broyl, A. *et al.* High cereblon expression is associated with better survival in patients with newly diagnosed multiple myeloma treated with thalidomide maintenance. *Blood* **121,** 624–627 (2013).

57. Heintel, D. *et al.* High expression of cereblon (CRBN) is associated with improved clinical response in patients with multiple myeloma treated with lenalidomide and dexamethasone. *British Journal of Haematology* **161,** 695–700 (2013).

58. Han, T. *et al.* Anticancer sulfonamides target splicing by inducing RBM39 degradation via recruitment to DCAF15. *Science* eaal3755–21 (2017). doi:10.1126/science.aal3755

59. Lacy, M. Q. *et al.* Pomalidomide (CC4047) plus low-dose dexamethasone as therapy for relapsed multiple myeloma. *J. Clin. Oncol.* **27,** 5008–5014 (2009).

60. Rajkumar, S. V., Blood, E., Vesole, D., Fonseca, R. & Greipp, P. R. Phase III Clinical Trial of Thalidomide Plus Dexamethasone Compared With Dexamethasone Alone in Newly Diagnosed Multiple Myeloma: A Clinical Trial Coordinated by the Eastern

Cooperative Oncology Group. *J. Clin. Oncol.* **24,** 431–436 (2006).

61. McCarty, A. S., Kleiger, G., Eisenberg, D. & Smale, S. T. Selective dimerization of a C2H2 zinc finger subfamily. *Molecular Cell* **11,** 459–470 (2003).

62. Cobb, B. S. *et al.* Targeting of Ikaros to pericentromeric heterochromatin by direct DNA binding. *Genes & Development* **14,** 2146–2160 (2000).

63. Wolfe, S. A., Nekludova, L. & Pabo, C. O. DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struct* **29,** 183–212 (2000).

64. Najafabadi, H. S. *et al.* C2H2 zinc finger proteins greatly expand the human regulatory lexicon. *Nat. Biotechnol.* **33,** 555–562 (2015).

65. Petzold, G., Fischer, E. S. & Thomä, N. H. Structural basis of lenalidomide-induced CK1α degradation by the CRL4CRBN ubiquitin ligase. *Nature* 1–16 (2016). doi:10.1038/nature16979

66. Matyskiela, M. E. *et al.* A novel cereblon modulator recruits GSPT1 to the CRL4CRBN ubiquitin ligase. *Nature* 1–24 (2016). doi:10.1038/nature18611

67. Matyskiela, M. E. *et al.* A Cereblon Modulator (CC-220) with Improved Degradation of Ikaros and Aiolos. *J. Med. Chem.* acs.jmedchem.6b01921 (2017). doi:10.1021/acs.jmedchem.6b01921

68. Jin, X. *et al.* An atypical E3 ligase zinc finger protein 91 stabilizes and activates NF-kappaB-inducing kinase via Lys63-linked ubiquitination. *Journal of Biological Chemistry* **285,** 30539–30547 (2010).

69. Jin, H. R., Jin, X. & Lee, J. J. Zinc-finger protein 91 plays a key role in LIGHT-induced activation of non-canonical NF-κB pathway. *Biochem. Biophys. Res. Commun.* **400,** 581–586 (2010).

70. Conomos, D., Reddel, R. R. & Pickett, H. A. NuRD-ZNF827 recruitment to telomeres creates a molecular scaffold for homologous recombination. *Nat. Struct. Mol. Biol.* **21,** 760–770 (2014).