Cereblon modulators: Low molecular weight inducers of protein degradation

Philip P. Chamberlain*, Brian E. Cathers

Celgene Corporation, 10300 Campus Point Drive, Suite 100, San Diego, CA, United States

Targeted protein degradation has become an exciting new paradigm in drug discovery with the potential to target new protein families for therapeutic intervention. In 2010, Hiroshi Handa and colleagues discovered that the drug thalidomide binds to the protein cereblon, a component of the CRL4CRBN E3 ubiquitin ligase. In contrast to the heterobifunctional small molecule degraders reported in the literature, thalidomide is of very low molecular weight (~258Da) with molecular properties (solubility, metabolic stability, permeability etc) that readily support pharmaceutical dosing. It was subsequently shown that thalidomide and the analogues lenalidomide and pomalidomide are able to degrade the transcription factors Ikaros and Aiolos. CK1α and GSPT1 were subsequently identified as substrates for specific ligands, indicating that this molecular class could be tuned for selective protein degradation. Structural studies showed that the thalidomide analogues bind to a shallow hydrophobic pocket on the surface of cereblon, and scaffold a protein–protein interaction with target proteins. Target proteins do not need any affinity for the cereblon modulators, and as such undruggable, or even unligandable, proteins can be targeted for degradation. A similar mechanism of action was subsequently identified for the clinical molecule indisulam, indicating that low molecular weight degraders are not unique to cereblon. The groundbreaking work on cereblon represents a case study for the discovery and characterization of low molecular weight protein degraders for other ligases.

Section editors:
Alessio Ciulli, FRSC – Professor of Chemical & Structural Biology, School of Life Sciences, University of Dundee, Division of Biological Chemistry and Drug Discovery, James Black Centre, Dow Street, Dundee DD1 5EH, United Kingdom.
William Farnaby – Professor of Chemical & Structural Biology, School of Life Sciences, University of Dundee, Division of Biological Chemistry and Drug Discovery, University of Dundee, James Black Centre, Dow Street, Dundee DD1 5EH, United Kingdom.

Introduction

The ubiquitin proteasome system (UPS) has proven to be a challenging but extremely promising area of biology for drug discovery and development. There are several drugs targeting the proteasome itself, but considering that there are ~700 potential drug targets across the combined ubiquitin ligase and deubiquitinating enzyme space, there remains relatively few agents in clinical development and even fewer approved drugs that target these enzymes (Fig. 1). The −E3 ligases are the most numerous members of the UPS and have proven to be very challenging as drug targets. Developing complex enzyme screening cascades is difficult and discovering bona fide inhibitors of E3 ligases even more so. Only inhibitors of the p53mdm2 interaction [1] or SMAC mimetics targeting the IAPs [2] have reached the clinical stage. As the field endeavors to find
Inhibitors of E3 ligases, an alternate approach has sparked enormous interest in recent years. In 2001 it was shown that an E3 ubiquitin ligase could be redirected from endogenous substrates to selectively ubiquitinate a protein of interest resulting in proteasomal degradation [3]. Crews and Deshaies had pioneered a technique whereby an E3 ligase is co-opted to ubiquitinate a non-natural protein substrate and coined the term PROTACs (Proteolysis Targeting Chimeras). In this approach, discrete substrate binding moieties are connected to an E3 ligase recognition element via a linker. Early examples of proteins to be targeted using this approach include the androgen receptor, the estrogen receptor-alpha, and further work has successfully targeted BRD4, and identified a broad spectrum of kinases vulnerable to this strategy [4–8]. The requirement for two independent binding moieties has thus far generated rather large molecules from a drug discovery perspective, but the approach holds considerable promise.

Recently, the discovery that the molecular mechanism of action of thalidomide analogs, namely the IMiD drugs lenalidomide and pomalidomide, is by induced substrate recognition by the CRL4<sub>CRBN</sub> E3 ligase has provided clinical validation for targeted protein degradation as a new paradigm in drug discovery [9]. These low molecular weight degraders remain the only known target protein degraders that have made it through clinical development (thalidomide, lenalidomide, and pomalidomide, Fig. 2). The surprising versatility and selectivity of the thalidomide analogs was demonstrated when it was shown that lenalidomide can induce the degradation of CK1α while the structurally similar pomalidomide cannot [10]. More recently, related molecules referred to as Cereblon E3 Ligand Modulating Drugs (CELMoDs) have expanded the range of potential target substrates even further when the translation termination factor GSPT1 was shown to be degraded by CC-885 [11]. While these types of molecules are much smaller than the chimeric PROTACs, the breadth or scope of CELMoDs may be limited to a smaller set of target proteins based on the requirement for highly compatible protein surfaces and a short glycin-containing hairpin loop [11,12]. However, the exciting possibility that other ligases beyond CRL4<sub>CRBN</sub> are capable of a similar ligand-induced or molecular glue substrate recognition is supported by the discovery that indisulam, an investigational anti-cancer drug, also operates via redirection of an E3 ubiquitin ligase [13,14].

**Discovery of the E3 ligase CRL4<sub>CRBN</sub> as the target of the IMiD drugs**

Nothing was known about the mechanism of action of thalidomide when it was brought to market in the 1950’s. Despite the lack of knowledge thalidomide was shown to have numerous cellular and clinical effects. Its sedative and anti-nausea activity [15], its remarkable activity in treating ENL associated with leprosy [16], its anti-wasting activity and
effects on Kaposi sarcoma associated with AIDS [17–19], and ultimately its activity in multiple myeloma were all discovered without understanding the underlying molecular mechanism.

In seemingly unrelated research, Higgins et al. published a paper describing the association of mild mental retardation in children with a truncating mutation in a gene of unknown function [20]. Because of the cognitive effects associated with this mutation and the sequence similarity with a structural domain from bacterial ion protease, the new gene was called cereblon. In March of 2010, over fifty years after thalidomide was first developed, cereblon was discovered by the laboratory of Hiroshi Handa to be involved in the mechanism of thalidomide teratogenicity and was proposed to be part of a CullE3 ligase [9]. Subsequently, cereblon was demonstrated to be necessary for the immunomodulatory and anti-myeloma effects of lenalidomide and pomalidomide [21,22]. In 2014, the molecular and structural mechanism of the IMiD drugs was revealed. Remarkably, it was shown that the thalidomide analogs imparted a gain of function or ‘neomorphic’ activity to the CRL4CRBN ligase. Interaction of lenalidomide with cereblon enhanced the binding and subsequent ubiquitination of the zinc finger transcription factors, Ikaros and Aiolos [23–25]. Proteasome-dependent destruction of these transcription factors was shown to have a direct anti-proliferative effect on myeloma cell lines while providing a stimulatory effect to T cells and offered a mechanistic rationale for this class of drug that had evaded researchers for over fifty years.

It was subsequently shown that the protein kinase CK1α was a substrate for lenalidomide induced proteasomal degradation [10]. This finding was of clinical significance as it provided an explanation for the efficacy of lenalidomide in del(5q) myelodysplastic syndrome, as the gene for CK1α resides in the 5q chromosomal region, and haploinsufficiency potentiates susceptibility to further degradation [26]. The discovery of CK1α provided evidence that further protein substrates could be found for next-generation cereblon modulators. Furthermore, the discovery of CK1α provided a demonstration that small differences in ligand chemical structure could drive substantial differences in substrate ubiquitination: pomalidomide, thalidomide and CC-122 do not induce the recruitment of CK1α. It also provided clinical validation for targeted protein degradation as an important new mechanism for drug discovery and demonstrated the potential for this approach to target proteins otherwise deemed ‘undruggable’. However, the mechanism for how the IMiD drugs were directing the CRL4CRBN ligase to target Ikaros, Aiolos, and CK1α was not clear.

**Molecular glue as the mechanism of action of IMiD drugs**

Later in 2014, the first crystal structures of IMiD drugs bound to cereblon in complex with DDB1 were published [27,28]. The glutarimide ring common to this class of compound bound in a shallow pocket made up of 3 tryptophan residues (tri-Trp pocket) with a phenylalanine side chain as the base. The variable part of the clinical compounds (Fig. 2), in these cases the phthalimide or isoindolinone groups, protruded out of the tri-Trp pocket and were exposed on the surface of cereblon. The IMiD drugs are very small molecules with modest affinity for cereblon and no measurable affinity for
the target proteins Ikaros and Aiolos, so a direct and specific high affinity interaction between the substrate protein and the ligand was unlikely to drive the observed ubiquitination and degradation. Chamberlain et al. proposed that this binding mode created a ‘hot-spot’ for protein-protein interactions, with unsatisfied hydrogen bonds available in a local hydrophobic environment composed of both the cereblon surface and the exposed portion of the drug molecules [27].

There are several noted examples of small molecules that bridge or induce a protein-protein interaction, and several natural products including FK506 and forskolin work in part by inducing or complementing a protein-protein interaction [29–31]. In the world of E3 ligases, the plant hormones auxin and jasmonate scaffold protein-protein interactions with substrate proteins prompting the team of Ning Zheng to coin the term ‘molecular glue’ [32,33]. These low molecular weight plant hormones display only modest binding affinity for either of their respective ligases or cognate substrates alone, but the ternary complex is formed with sufficiently high affinity to allow for successive rounds of ubiquitination and subsequent proteolytic degradation. Indeed, in 2016 the molecular glue model was confirmed from the first structures of the ternary complexes of two different substrates bound to cereblon: Petzold et al. solved the structure of cereblon in complex with lenalidomide and CK1α [12], and Matyskiela et al. solved the structure of cereblon in complex with a newly described substrate (GSPT1) in complex with a newly described degrader, CC-885 [11]. The discovery of CC-885 and GSPT1 represents a prototype for the expanded potential of this new ligand class, the CELMoDs.

A novel CELMoD targeting Ikaros and Aiolos for degradation was reported in 2017, and show-cases the ability for the molecular understanding of the mechanism to assist in differentiating clinical candidates. Iberdomide (CC-220, Fig. 2) exhibits more than an order of magnitude improved affinity towards cereblon than thalidomide, lenalidomide, or pomalidomide, and consequently exhibits far more potent degradation of Ikaros and Aiolos [34]. The crystal structure of iberdomide revealed that increased potency correlates with more extensive interactions with the surface of cereblon. Iberdomide is currently in phase 2 clinical development for lupus and phase 1b/2a clinical trials for relapsed/refractory multiple myeloma [35].

**Neosubstrate degradation in safety considerations**

Thalidomide is a highly potent teratogen and strict risk management programs have been implemented to minimize the possibility of fetal exposure to molecules in this class. Tragically, thalidomide teratogenicity was established after widespread use of the drug to treat morning sickness in the 1950s, and thousands of babies were affected worldwide. Two studies published in 2018 have now established a plausible explanation for the potent teratogenicity by demonstrating that SALL4 is a robust substrate for thalidomide mediated degradation [37,38]. Critically, human mutations that reduce SALL4 activity have a clinical presentation that resembles thalidomide embryopathy so closely that it has led to misdiagnosis [39–41]. Like Ikaros and Aiolos, SALL4 is a C2H2 zinc finger transcription factor and based on mutagenesis appears to contain the critical glycine-containing neosubstrate degron found in the other validated neosubstrates. Critically, it is not only CELMoDs that exhibit degradation of neosubstrates, heterobifunctional molecules co-opting cereblon have also been commonly reported: SALL4 degradation has been shown to be a property of some heterobifunctional ligands [37,38]. SALL4 degradation is not the only neosubstrate activity to have been identified in heterobifunctional molecules, as Ikaros/Aiolos degradation has been reported to occur with several molecules, and GSPT1 degradation has been discovered as an off-target able to drive antiproliferative effects [42].

Efforts across multiple research teams have reported an expanding number of cereblon neosubstrates including many further members of the zinc finger transcription factor family [10,36–38,43]. Based on homology and mutagenesis, it appears that most of these proteins will be found to exhibit the neosubstrate degron, although cereblon-substrate complexes have so far have only been reported for the substrates IKZF1, ZNF692,
CK1α, and GSPT1 [11,12,36]. The discovery of novel neoubiquitin ligases provides considerable promise for the discovery of differentiated therapeutics across multiple disease areas. However, it is going to be critical to monitor off-target activities in the discovery and development of both new CELMoDs and cereblon dependent heterobifunctional degraders.

**Conclusions**

The discovery of the CELMoD mechanism of action has provided a revolution in drug discovery, providing both clinical validation for the new paradigm of targeted protein degradation, but also rewriting the rules of druggability: Researchers now have chemical tools and candidate therapeutics to drive the destruction of disease causing proteins from families that do not even contain small molecule binding sites. Beyond the discovery that the Ikaros family of zinc finger transcription factors are CELMoD substrates, research teams have been able to substantially expand the known breadth of activity into diverse protein families including translation factors and kinases and further substrates are greatly anticipated.

The discovery that indisulam also acts as a molecular glue implies that the principles uncovered in the first case studies may find broader utility in the discovery of new therapeutics. As the field of drug discovery is now aware of targeted protein degradation as a valid clinical mechanism, we can be vigilant for this modality in existing drugs of unknown mechanism as well as to prospectively engineer molecules to take advantage of the powerful ubiquitin proteasome system to remove disease causing proteins from the cell. Whether via chimeric molecules or molecular glue mechanisms, targeted protein degradation is an important new paradigm in drug discovery.

**Conflict of interest**

Authors are, or have been, employees and shareholders at Celgene.

**References**


Fig. 3. The cereblon neoubiquitin ligase cereblon with CC-885 and GSPT1 with the key glycine shown as a sphere (a). The degron sequence shows no conservation apart from the key glycine as shown in (b), and the occurrence of the degron on validated substrates with different folds (c).
Drug Discovery Today: Technologies | Protein degradation for drug discovery
Vol. 31, 2019


