Targeted protein degradation: expanding the toolbox

Matthieu Schapira[1,2,*], Matthew F. Calabrese[3], Alex N. Bullock[4] and Craig M. Crews[5,6,7]

[1] Structural Genomics Consortium, University of Toronto, Toronto, ON M5G 1L7, Canada

[2] Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON M5S 1A8, Canada

[3] Discovery Sciences, Pfizer Worldwide Research and Development, Groton, CT 06340

[4] Structural Genomics Consortium, University of Oxford, Old Road Campus Research Building, Roosevelt Drive, Oxford OX3 7DQ, United Kingdom

[5] Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, Connecticut 06511, United States.

[6] Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107, United States.

[7] Department of Pharmacology, Yale University, New Haven, Connecticut 06520-8066, United States.

* correspondence should be addressed to matthieu.schapira@utoronto.ca

Abstract

Proteolysis-targeting chimeras (PROTACs) and related molecules that induce targeted protein degradation by the ubiquitin–proteasome system represent a new therapeutic modality and are the focus of great interest, owing to potential advantages over traditional occupancy-based inhibitors with respect to dosing, side effects, drug resistance and modulating 'undruggable' targets. However, the technology is still maturing, and the design elements for successful PROTAC-based drugs are currently being elucidated. Importantly, fewer than 10 of the more than 600 E3 ubiquitin ligases have so far been exploited for targeted protein degradation, and expansion of knowledge in this area is a key opportunity. Here, we briefly discuss lessons learned about targeted protein degradation in chemical biology and drug discovery and systematically review the expression profile, domain architecture and chemical tractability of human E3 ligases that could expand the toolbox for PROTAC discovery.

Introduction

Despite continuous progress in the development of potent and selective small-molecule inhibitors of protein function, multiple targets of high biomedical relevance are still highly challenging for typical small-molecule drugs. In addition, although biologic modalities such as monoclonal antibodies and oligonucleotide therapies can provide opportunities to address such targets, these have limitations such as restricted delivery. Consequently, recent advances indicating that targeted protein degradation with small-molecule drugs could become a new therapeutic modality have attracted substantial interest.

Protein degradation is a normal process of protein turnover within the cell. It provides a mechanism of quality control during protein folding, an ability to rapidly respond to changing cellular signals, and a mechanism to modulate the pool of available amino acids[1]. The majority of proteins will undergo degradation through the ubiquitin–proteasome system (UPS). While this process is broad enough to encompass the vast diversity of the proteome, it nevertheless operates through a collection of regulated and orchestrated steps in which proteins are marked for degradation by covalent post-translational modification with the protein ubiquitin. Ubiquitylation of proteins is carried out by a cascade of three enzymes. In the first step, ATP is consumed by an E1 ubiquitin-activating enzyme to produce an activated ubiquitin–adenylate, which is converted to a thioester intermediate via covalent attachment to a catalytic cysteine in the E1 active site. This is followed by a transthiolation reaction in which ubiquitin is transferred from the catalytic cysteine of the E1 enzyme to the catalytic cysteine of an E2 (ubiquitin conjugation) enzyme. Finally, ubiquitin is transferred to the substrate protein by the action of a bridging E3 ubiquitin ligase, where it forms an isopeptide bond between the carboxy terminus of ubiquitin and a lysine side chain of the target. This cycle can be repeated to generate a poly-ubiquitin chain that directs a substrate for degradation at the proteasome. Within this cascade, E3 ligases are unique in their role of dictating target specificity. Representing a large gene family with ~600 predicted members, E3 ligases typically function as adaptor molecules that recognize substrates through protein–protein interactions and promote ubiquitylation by holding those targets proximal to the associated ubiquitylation machinery.

The potential to redirect protein degradation by artificially recruiting an E3 ligase was demonstrated nearly 20 years ago with the E3 ligase BTRC[2]. In this work, a chimeric molecule comprising a covalent ligand of methionine aminopeptidase 2 (MetAP2) linked to a peptidic ligand of BTRC (a phosphopeptide derived from IκBα) was generated, and this molecule served as a non-native 'bridge' between the target and the E3 ligase (Fig. 1). Excitingly, this compound-dependent recruitment was sufficient to promote degradation of MetAP2 in Xenopus extracts. While this study established a compelling proof of principle, interest in these 'proteolysis targeting chimeras' (PROTACs) as potential drugs remained limited until such molecules progressed beyond peptidic ligands to fully synthetic compounds. This was first demonstrated for recruitment of the androgen receptor to the E3 ligase MDM2 in 2008 (ref.[3]), with extension to the E3 ligase cIAP1 shortly thereafter[4]. At a similar time, independent work from Hiroshi Handa and others demonstrated that the immunomodulatory drug (IMiD) thalidomide directly binds to the E3 ligase cereblon (CRBN)[5], and this binding event was later shown to mediate ligand-dependent degradation of a collection of targets (including IKZF1, IKZF3 and CK1α) by a set of IMiD analogues, including lenalidomide and pomalidomide[6,7,8,9]. Similar to the PROTAC molecules described above, IMiDs are able to promote non-native target degradation via non-native E3 ligase recruitment (Fig. 1).

In the past 5 years, the field of targeted protein degradation has expanded dramatically, with dozens of exemplified substrates being amenable to this mechanism. Studies have progressed beyond cellular assays to in vivo studies in model organisms, and results from the first phase I clinical trials are on the horizon. However, considerable challenges remain. While recruitment to an E3 ligase is a necessary step in targeted protein degradation, it appears insufficient, as subtle features in all segments of PROTAC design can have vast effects on degradation potency[10,11]. In addition, emerging data from proteomics studies using PROTACs based on promiscuous kinase binders suggest that certain targets are more responsive to this mechanism of action than are others[12,13]. As such, the field of protein degradation at present requires

substantial empirical screening. Indeed, even among degradable targets, the choice of E3 pairing appears critical[10].

We propose that a key enabler for the field is a broader exploration of the E3 ligase gene family as a whole. Indeed, only ~1% of the 600 family members have been explored in targeted protein degradation to date. In this article, we discuss opportunities and challenges for the field of targeted protein degradation and review the major classes of E3 ligases that may be exploited to expand the PROTAC toolbox. We address features of E3 ligase ligandability, but note that PROTACs with even moderate E3 ligase affinity have proven capable of directing efficient degradation[14]. If fully realized, this modality holds tremendous potential to pursue established targets in new ways and expand the druggable genome by targeting clinically relevant proteins that are currently not amenable to inhibition.

Progress and lessons learned

Recent observations indicate that IMiDs and PROTACs are poised to have a strong impact on drug discovery in years to come (Table 1). The clinical efficacy of IMiDs is driven at least in part by their capacity to induce CRBN-mediated degradation of neo-substrates: cancer cells become resistant to the multiple myeloma IMiD drug lenalidomide upon mutation of a single amino acid in IKFZ3 that rescues this transcription factor from proteolytic degradation[7]; haploinsufficient expression of CK1α sensitizes myelodysplastic syndrome (MDS) cells with deletion of chromosome 5q (del(5q)) to the CRBN-mediated degradation of CK1α by lenalidomide in patients with MDS (5q)15; and treatment with C-220, a CRBN modulator that induces a more potent degradation of IKFZ1 and IKFZ3 than previous IMiDs[16], elicits positive response in patients with systemic lupus erythematosus in a phase 2a study[17]. Preclinical and clinical studies have demonstrated multiple routes of administration, including oral[18]. Wide distribution of PROTACs to organs and tissues has been demonstrated, including recently to the brain (Arvinas press release; see Related links), and PROTACs directed to oncology targets can induce durable responses in xenograft models[19]. Cancer cells resistant to the kinase inhibitor ibrutinib are responsive to treatment with an ibrutinib-derived PROTAC, suggesting that PROTACs can be

used to address resistance mechanisms affecting parent inhibitors[20]. Finally, a first hetero-bifunctional PROTAC that recruits the androgen receptor to an E3 ligase recently entered the clinic[21].

In addition to clinical or preclinical data, chemical biology studies have highlighted properties intrinsic to the mechanism of action of PROTACs that positively differentiate them from conventional inhibitors (Table 1). PROTACs can achieve high cellular potency due to their catalytic rather than occupancy-based mechanism of action[11,14,22] and can have a duration of action that extends beyond clearance and depends on the turnover rate of the protein target rather than residence time[23,24,25]. Protein domains that feature ligandable binding pockets but are not involved in the pathogenic function of a gene can be favourably targeted by these molecules[26,27], and several PROTACs have been found to be more selective than the inhibitor from which they were derived[12,28,29,30], probably because not all targets brought in proximity to a given ligase are productively ubiquitylated, due to the lack of accessible lysine residues and/or the action of deubiquitinases. The expression profile of the chosen E3 ligase can also be exploited to degrade a target in specific tissues or cellular compartments[31]. Finally, target degradation rather than inhibition is a promising therapeutic modality for diseases driven by the accumulation of aberrant forms of proteins, such as tauopathies[32].

A number of challenges accompanying the development of PROTACs should also be noted. Regarding clinical applications, appropriate dosage may be an issue, as saturating doses of free PROTAC molecules can antagonize the binding of binary PROTAC–protein complexes to their ternary partner and abrogate catalytic degradation, a well-documented phenomenon known as the hook effect in cell assays[23,33,34]. Acquired resistance to PROTAC treatment can be driven by genomic alterations targeting core components of E3 ligase complexes[35]. Mutations affecting PROTAC binding but not protein function should also be expected, maybe more so when PROTACs are exploiting functionally neutral sites. Finally, the PROTACs reported so far have been larger than typical orally available small-molecule drugs. However, available data indicate

that the pharmacokinetics of PROTACs might be less prohibitive than would be expected given their physicochemical properties[36].

The development of PROTACs as chemical biology tools is also fraught with challenges. There is currently little rationale guiding the pairing of a specific E3 ligase with a given protein target: it is unclear at the outset of a PROTAC discovery effort whether fastidious combinatorial sampling of linker chemistry will reveal a suitable combination of linker length and attachment points necessary to the formation of a productive complex, or whether resources will be wasted in attempting to match proteins that cannot be paired[37,38]. Even when a PROTAC with exquisite selectivity is developed, the determinants driving its specificity profile often remain complex, obscure or puzzling[29,30]. PROTACs can also degrade off-targets that were untouched by their parent inhibitors[39], IMiDs degrade diverse arrays of ZnF transcription factors[9], and some CRBN-recruiting PROTACs have been shown to catalyse the degradation of the IMiD targets IKFZ1 and IKFZ3 (refs[40,41]). These molecules are generally larger and more flexible than typical drug-like compounds, which can translate into poor membrane permeability and liability to efflux pumps[39,42,43]. While covalent binding to E3 ligases is acceptable[31,44], PROTACs binding covalently to the protein target probably lose the substoichiometric nature of their mechanism[45]. Finally, what percentage of a target protein should be degraded to trigger a phenotypic response probably depends on the target and the readout, but needs to be systematically investigated.

Recent crystal structures of ternary complexes have advanced the understanding of the structural mechanism of PROTACs. Structural studies on VHL and CRBN show that these Cullin–RING E3 ligases (CRLs) form large, modular, U-shaped complexes in which adaptor proteins mediate the interaction between the substrate-binding element (VHL or CRBN) and Cullin scaffolds that bind the RING-domain protein RBX1, leading to the recruitment of a ubiquitin-conjugated E2 (refs[9,46,47,48,49,50,51,52]) (Fig. 2a,b). The U shape leads to proximal positioning of the E2 and substrate proteins, allowing targeted ubiquitin transfer. The architecture of these large complexes is expected to provide an extended ubiquitylation radius that can accommodate multiple ubiquitylation sites on substrates with diverse sizes and shapes[47,52]. In particular, given

the flexible nature of the adaptor protein DNA damage-binding protein 1 (DDB1), the spectrum of substrates degraded by PROTACs recruiting the CRBN complex (and probably other DDB1–CUL4-associated factor (DCAF) E3 ligases) is expected to be dictated less by the accessibility of the ubiquitylation site and more by the protein synthesis rate as well as the affinity and kinetics of the ligase–PROTAC–target binding event[47,50,52]. Formation of the ternary complex is driven by protein–protein and protein–PROTAC interactions, sometimes including stabilizing contacts between the PROTAC linker and the recruited proteins[51] (Fig. 2a).

The structures of the ternary complex between CRBN, the first bromodomain of the target protein bromodomain-containing protein 4 (BRD4) and diverse PROTACs reveal the plasticity of the interaction, in which different linkers can lead to distinct, target-specific arrangements of the ligase–target interface[46] (Fig. 2c). Similarly, restriction on the interfaces accessible to VHL and p38 isoforms imposed by PROTACs with reduced linker length or specific attachment points can lead to isoform-selective degradation of p38α[29]. Attempts to predict protein interfaces via docking simulations were recently reported[46,53], but the discovery of PROTACs remains empirical at this time.

Unlike PROTACs, CRBN-binding IMiDs lack a substrate-targeting chemical moiety. Rather, the target engages in direct interaction with the CRBN-bound phthalimide group via a β-hairpin that is structurally conserved in all available complex structures, where a critically positioned glycine abuts against the phthalimide in a binding pose that would be incompatible with any other amino acid (Fig. 2d). This unique structural arrangement is preserved in unrelated IMiD targets such as the zinc-finger proteins IKFZ1 and ZNF692 (ref.[9]), the kinase CK1α[47] and the GTP-binding protein GSPT1 (ref.[48]), and was used to identify novel zinc-finger proteins degraded by thalidomide analogues[9]. Importantly, this interaction seems to also be preserved in the context of hetero-bifunctional molecules. For example, a CRBN-recruiting, IMiD-based PROTAC targeting the BTK tyrosine kinase also degrades the phthalimide-binding neo-substrates IKFZ1 and IKFZ3, leading to a synergistic and beneficial effect on mantle cell lymphoma[40,54]. This dual activity raises the possibility that some IMiD-based PROTACs may be prone to degrade off-

targets that feature a glycine-containing β-hairpin degron, including more than 150 zinc-finger proteins[9], which could obscure the interpretation of phenotypic response upon PROTAC treatment or affect the toxicity profiles of drug candidates.

At present, fewer than 10 E3 ligases (CRBN, VHL, IAPs, MDM2, DCAF15, DCAF16, RNF114), out of over 600 in the human proteome[55], have been exploited by degradation-inducing small molecules. Extending the repertoire of ligands to E3 ligases with a variety of structural properties as well as diverse temporal and spatial expression profiles should considerably expand the potential applications of PROTACs for chemical biology and broaden the horizon for future drug discovery efforts. With this opportunity in mind, we now summarize the classification of human E3 ligases, their expression profiles and essentiality in cancer, and then systematically analyse their ligandability.

Expanding the toolbox

Classification of human E3 ligases.

E3 ligases can tag substrate proteins with monoubiquitin or polymeric ubiquitin chains in which successive ubiquitin molecules are connected through distinct isopeptide bonds. Only some of these distinct polyubiquitin chains, such as the ones branched via lysine 48, lead to proteasomal degradation of the tagged protein; non-proteolytic functions of ubiquitylation include DNA repair or subcellular localization[56,57]. The type of polyubiquitin chain installed remains unknown for a large number of E3 ligases, and so is the association of these enzymes with the UPS and their fitness for the development of PROTACs. Based on the annotations of E3 ligases and their binding partners available from UniProt and the Reactome Pathway database, about 270 of the 632 or more human E3 ligases are currently believed to be involved in the UPS[55].

E3 ligases are generally classified in two main categories, based on their mechanism of action: HECT-domain enzymes form a thioester bond with ubiquitin before transferring it to its substrate, whereas RING E3 ligases recruit E2–ubiquitin conjugates via their RING domain and

catalyse the direct transfer of ubiquitin from the E2 enzyme to the substrate. The substrate-binding and RING domains either can belong to the same protein or can be distinct components of multisubunit complexes, such as CRLs, where a Cullin acts as a protein linker between a substrate-targeting subunit and an E2-binding protein58 (Fig. 2a,b). For instance, DCAF E3 ligases generally contain a substrate-binding WD40 repeat (WDR) module and a distinct domain that mediates attachment to Cullin 4 via the adaptor protein DDB1. Cullin 4 simultaneously binds the RING domain protein RBX1, which in turn recruits E2–ubiquitin conjugates for subsequent substrate ubiquitylation. A variation on the Cullin theme is observed in anaphase-promoting complex (APC) ligases, where distinct entities of a multisubunit complex bring an E2–ubiquitin conjugate and a protein substrate into close proximity. RING-between-RING (RBR) E3 ligases are mechanistic hybrids that bind E2s via a RING domain but form an intermediate thioester bond, like HECT enzymes, before transferring ubiquitin to the substrate.

Expression of E3 ligases

PROTACs are only active if the E3 ligase they recruit is available in the cells or tissue of interest. PROTACs relying on ubiquitously expressed E3 ligases could therefore be used as chemical biology tools in a broad range of cellular systems. Based on proteomics data across 81 cell and tissue types available from the Human Protein Atlas[59] (see Related links), 24 E3 ligases are present in at least 90% of cell and tissue types tested (Fig. 3a). Among these, ten (RBBP7, MDM2, TOPORS, TRIM35, TRIM28, FBXW7, UBE3B, PPIL2, UBE3A and RNF20) are known to be involved in the UPS. MDM2, an E3 ligase exploited by existing PROTACs, is expressed in 99% of samples tested, indicating that MDM2-recruiting PROTACs should be valid chemical biology tools in a variety of cellular contexts. It was recently shown that a MDM2-based PROTAC could synergistically combine a catalytic mechanism of action, via degradation of a neo-substrate, and an occupancy-based effect, via competition with, and stabilization of, the endogenous substrate p53 (ref.[60]). This type of synergy may be desirable for a drug, but could obfuscate the phenotype associated with degradation of the neo-substrate. Such an effect is expected to vary on the basis of the stoichiometric balance between E3, PROTAC and substrates in cells, and could be reduced either by decreasing the PROTAC concentration to sub-stoichiometric levels

where it is poorly competitive but still catalytically active, or by exploiting non-functional domains of the E3 ligase. The inhibitor of apoptosis (IAP) proteins BIRC2 and XIAP are hijacked by current PROTACs (also known as specific and non-genetic IAP-dependent protein erasers (SNIPERs)) but do not reach detectable levels in >50% cell and tissue types from the Human Protein Atlas (Fig. 3a), which could be a liability for chemical genomic approaches but is an advantage for translational studies.

PROTACs that recruit E3 ligases with a tissue-selective expression profile are expected to present unique opportunities for therapeutic applications, as they should not degrade the targeted protein in tissues where the E3 ligase is not expressed. Twenty E3 ligases have a narrow window of expression across human tissues, according to the Human Protein Atlas (Fig. 3b). Of these, four are known to induce proteasomal degradation of their substrates (ASB9, KLHL10, KLHL41 and TRIM69). For instance, ASB9, a SOCS box E3 ligase, is exclusively expressed in pancreas and testis, while the F-box E3 ligase FBXL16 is specifically found in caudate and cerebral cortex. Chemical handles binding with sufficient potency and specificity to one of these E3 ligases could be linked to a variety of substrate-targeting ligands for tissue-specific silencing of disease-associated genes.

The expression profile of the currently exploited E3 ligases is comparatively ubiquitous, which may translate to instances of undesired effects of PROTACs recruiting these enzymes (Fig. 3b). These proteomics data rely on the quality and selectivity of the antibodies used for protein detection, and therefore need to be further validated, but they nevertheless illustrate the value of exploiting E3 ligases with diverse tissue expression profiles. In an interesting variation on this theme, PROTACs may be used to induce substrate degradation in specific cellular compartments: a PROTAC covalently recruiting the nuclear E3 ligase DCAF16 was recently shown to degrade exclusively nuclear targets[31].

Essentiality of E3 ligases in cancer

According to the cancer dependency map (depmap.org), which provides gene essentiality derived from CRISPR-knockout studies across >340 cancer cell lines and multiple cancer types[61], a number of E3 ligases are essential — and therefore available for proteasome-targeting applications — in most cancer types (Fig. 3c). Of particular interest, CDC20, the substrate-binding subunit of the APC, can induce degradation of target proteins and is essential in all cancer types tested, but it has low to undetected protein levels in 70% of non-cancer cells according to the Human Protein Atlas. Additionally, a weak small-molecule ligand that binds the substrate-binding domain of CDC20 was previously reported, suggesting that this domain is chemically tractable[62]. More potent chemical handles targeting CDC20 would be attractive tools for the development of PROTACs targeting oncogenes in a diverse array of cancer types. Substrate competitors with CDC20 block mitotic exit and induce tumour cell death[62], which could synergize with PROTAC-driven degradation of oncogenic neo-substrates. Genomic alteration of E3 ligase complexes is a resistance mechanism used by cancer cells in response to chronic treatment with VHL or CRBN-recruiting PROTACs[35]. Exploiting E3 ligases that are essential to the survival of cancer cells is a promising strategy to avoid this resistance mechanism.

Ligandability of E3 ligases

Extending the repertoire of E3 ligases exploited by PROTACs is an engaging prospect and is supported by the observation that five out of six E3 ligases representing diverse enzymatic and structural classes were amenable to recruitment for target degradation when fused to an artificial ligand-binding domain[63]. But this vision can only be realized if the structure of the targeted E3 ligases features pockets or crevices with geometrical and physicochemical properties that allow the binding of a small-molecule ligand. The remainder of the Perspective therefore focuses on the ligandability of the major classes of E3 ligases, based on available structural and chemical data.

DCAF E3 ligases.

DCAF E3 ligases are a subfamily of about 60 enzymes, 52 of which contain a WDR domain[64] (Fig. 4a). The DCAF protein CRBN, which does not feature a WDR domain, is one of the most characterized E3 ligases for PROTAC discovery, and a crystal structure of CRBN was solved in complex with a PROTAC and the non-native substrate BRD4, where the compound and target protein bind the atypical CULT and Lon domains of the ligase, respectively[46]. A class of sulfonamide drugs was recently shown to exploit the WDR-containing DCAF15 to chemically induce the ubiquitylation and proteasomal degradation of the cancer target RBM39 (refs[65,66]). The region of DCAF15 targeted by sulfonamides was not determined, but the WDR domain, a common protein-interaction scaffold that is also the only predicted domain of the protein, is an obvious candidate. The crystal structure of the homologous DCAF1 in complex with VPR indicates that substrate recognition relies on the WDR domain, a doughnut-like protein-interaction module that, in the context of other proteins, has been successfully targeted by small-molecule ligands[67] (Fig. 4a): EED, a critical component of the polycomb repressive complex 2 (PRC2), behaves as a DCAF E3 ligase that recruits PRC1 for subsequent ubiquitylation of histone substrates[68], and nanomolar ligands now in clinical development bind the central cavity of the EED WDR domain[69,70]. Potent and selective compounds are also targeting the central cavity of WDR5, another WDR DCAF E3 ligase[71]. More recently, a PROTAC was shown to induce the degradation of non-native substrates via cysteine-directed covalent recruitment of DCAF16 (ref.[31]). Indeed, covalent binding to E3 ligases should not prevent consecutive ubiquitylation of multiple substrate molecules by a single E3–PROTAC entity and is therefore expected to preserve the sub-stoichiometric catalytic nature of PROTAC-mediated substrate degradation.

The central cavity of WDR domains is generally deep and enclosed, two important properties for potent binding of chemical handles, but this site can sometimes be positively or negatively charged, which reduces its ligandability[72]. For instance, the central cavity of DDB2 and RBBP4 are highly acidic and basic, respectively, compared to those of WDR5, EED, DCAF1, PAFAH1B1, ATG16L1 or ERCC8. Side-chain plasticity can also greatly affect the ligandability of the central cavity. For instance, this site is shallow and looks unligandable in the apo-structure of EED, and

conformational remodelling is necessary for ligand binding[67]. WDR5 and EED are not believed to induce proteasomal degradation and are therefore not suitable for PROTAC discovery, but out of the 52 DCAF E3 ligases containing a WDR domain, it is expected that some are associated with the UPS and are ligandable. The external wall of the WDR domain can also be used as a protein interaction interface. For instance, CDC20, a non-DCAF E3 that acts as the substrate-binding subunit of the APC, uses the side surface of its WDR domain to recruit APC substrates via their D-box motif[73], and apcin, a small molecule that binds at the same CDC20 site, inhibits the ubiquitylation of D-box-containing substrates[62]. As will be seen below, the WDR domain and other structurally related β-propeller structures are found in multiple subfamilies of E3 ligases, indicating that these doughnut-like domains are efficient modules for substrate recruitment, and possibly for PROTAC discovery.

BTB E3 ligases.

Approximately 90 of the 220 human BTB-containing proteins are thought to function as Cullin 3-dependent E3 ligases and uniquely combine the BTB Cullin adaptor and substrate-recognition domains into a single protein (Fig. 4b). BTB-containing E3 ligases are typically distinguished by the presence of a 3-box motif that enables high-affinity binding of Cullin 3 (refs[74,75]). Most BTB domains also homodimerize, affording these E3 ligases with two substrate-recognition centres able to engage multiple degrons within a single substrate, as exemplified by the interactions of KEAP1 and SPOP with the substrates Nrf2 (refs[76,77]) and Ci/Gli[78], respectively. Proof of concept for hijacking the BTB E3 ligases has been provided by a peptide PROTAC targeting Tau for degradation by KEAP1 (ref.[79]). Importantly, this peptide was based on a single degron site within Nrf2, suggesting that multivalency is not likely to be required in equivalent chemical PROTACs.

BTB-Kelch proteins form the largest subfamily within this E3 class and also appear to be the most tractable for drug development. The Kelch domain folds as a six-bladed β-propeller with a central pocket for the binding of substrate degrons or small molecules. Crystal structures of peptide substrate complexes have been reported for four BTB-Kelch family members, including

KEAP1, KLHL2, KLHL3 and KLHL20 (refs[80,81,82]). However, to date, small-molecule development has been restricted to KEAP1, which represents the best-characterized family member, due to its therapeutic potential in chronic inflammatory and neurodegenerative diseases[83]. Importantly, low-nanomolar inhibitors of the Kelch domain of KEAP1 have been reported that demonstrate the ligandability of this target class[84]. Structural comparisons of human Kelch domains reveal significant variation in their pocket shapes and surface charges that may influence how favourable each member is for the development of PROTAC handles (Fig. 4b). A distinct β-propeller structure is also formed in some KCTD family members through oligomerization of their subunits into pentamers, as exemplified by KCTD5, which regulates GPCR signalling through ubiquitin-mediated degradation of Gβγ subunits[85,86]. A WDR β-propeller domain is also predicted in SHKBP1.

BTB-containing E3 ligases have been linked to a variety of proteolytic and non-proteolytic ubiquitin signals that may limit or complicate their utility for PROTACs[87,88]. For example, KLHL12 can induce degradative polyubiquitylation of dishevelled[89], but it can also assemble with specific co-adaptors to monoubiquitylate the COPII component SEC31 for collagen trafficking[90,91,92] or to induce non-lysine ubiquitylation of the dopamine receptors D4.2 and D4.4 (ref.[93]). In addition, some clades of the KCTD family lack Cullin 3 binding[94], while the BTB-Kelch family member KLHL39 appears to function as an antagonist that blocks the ubiquitylation and degradation of PML and DAPK1 by KLHL20 (ref.[95]).

VHL-box and SOCS-box E3 ligases.

VHL-box and SOCS-box proteins contain a BC box for binding to the adaptor proteins Elongin B and C, as well as a Cullin 2 or Cullin 5 box for their assembly into specific CRL2 or CRL5 complexes, respectively[96]. Unfortunately, in the context of PROTAC development, VHL represents a singleton E3 ligase, as the only homologue of the substrate-binding domain, VHL-like protein (VLP), acts as a dominant negative protein that lacks the C-terminal VHL-box required for Elongin B/C and Cullin 2 interaction[97]. Nonetheless, a further 12 diverse proteins contain a VHL-box with confirmed binding to Cullin 2 (Fig. 5a). Many of these mediate protein

destruction by binding to newly described C-terminal degrons[98]. These include two Kelch-domain proteins (KLHDC2 and KLHDC3) that are potentially ligandable. Indeed, crystal structures of KLHDC2 bound to C-terminal diglycine degrons have revealed a deep pocket shaped by three tryptophan and three tyrosine residues[99]. Their low-nanomolar substrate-binding affinities are probably dependent on the buried C-terminal carboxyl group, which establishes a salt bridge and two hydrogen bonds. Thus, like the BTB–Kelch protein KEAP1, these E3 pockets may favour compounds containing acidic moieties that present challenges for cellular permeability. The remaining VHL-box proteins contain either leucine-rich repeats (LRR1, PRAME, ZYG11B and ZER1), tetratricopeptide repeats (APPBP2) or ankyrin repeats (FEM1A–C) that are less characterized, and in the absence of structural information are probably less favourable for small-molecule development.

Another 37 E3 ligases use a SOCS-box domain to form Elongin B/C-containing complexes with Cullin 5 (ref.[100]). The Cullin 5 complex is well known for being hijacked by the HIV viral protein Vif[101], but the CRL5 class of E3 ligases has yet to be targeted by small molecules. Over 20 of the human SOCS-box proteins are believed to induce the degradation of their substrates (Fig. 5a). A potentially attractive WDR domain for PROTAC development is found in three members, although no structural data exist, and WSB1 has been reported to form multiple ubiquitin chain types, including K27 and K29-linked ubiquitylation of LRRK2 (ref.[102]), in addition to its degradation of VHL103 (Fig. 5a).

The ankyrin repeat and SOCS-box family (ASB1–18) is noteworthy for the selective tissue expression profile of some of its members. For instance, the ASB9 protein is specifically found in the pancreas and testis (Fig. 3), and ASB11 in muscles (according to RNA levels, which we did not account for in our expression profile in Fig. 3), while ASB4 is overexpressed in adrenal glands (according to the Human Protein Atlas) and in adrenocortical carcinoma (according to The Cancer Genome Atlas)[104]. PROTACs that recruit ASBs are therefore an attractive prospect to induce tissue-selective degradation of a target. While small-molecule ligands have yet to be

targeted to an ankyrin fold, structural data for ASB9 reveal juxtaposed hydrophobic cavities in the substrate-binding domain that may at least offer some hope for future work[105,106] (Fig. 5a).

Perhaps the best-characterized SOCS-box proteins are the SH2 family of CISH and SOCS1–7. Chemical tractability for SH2 domains has been poor historically, due to difficulties in designing cell-permeable phosphotyrosine mimetics. However, the available peptide co-structures for SOCS3 (refs[107,108,109]) show an expanded hydrophobic pocket that may be more amenable to targeting, as was found for the STAT SH2 family transcription factors. Peptide co-structures are also available for the SPRY domain-containing group of SPSB1–4, including examples with inhibitory cyclic peptides that may enable PROTAC proof-of-principle studies[110,111]. Finally, the RAB40-family proteins contain a poorly characterized GTPase domain that warrants further study for ligandability, as at least one member (RAB40C) has a reported link to the UPS[112].

F-box E3 ligases.

F-box E3 ligases are a subfamily of about 75 proteins that use a canonical F-box domain to interact with the adaptor protein SKP1, mediating Cullin 1 binding for recruitment of E2–ubiquitin conjugates. F-box ligases can be divided into three distinct classes, based on the nature of their substrate-binding domain. Eleven FBXW E3 ligases use a WDR domain for substrate recruitment, eight of which are known to be involved in the UPS (Fig. 5b). Among these, BTRC was efficiently recruited by PROTACs composed of a BTRC-interacting peptide and small-molecule ligands for METAP2, the oestrogen or androgen receptors, leading to degradation of their respective targets[2,113]. Phosphodegrons (phosphorylated peptides) of endogenous targets bind the central cavity of FBXW E3 ligases such as BTRC and FBXW7. The corresponding binding pockets are therefore basic (Fig. 5b), and chemical handles targeting these sites will probably be highly polar, which could be an insurmountable challenge in a typical drug optimization programme but may be overcome in the context of PROTACs. Indeed, PROTACS do not need to bind potently to E3 ligases (a PROTAC recruiting the E3 ligase VHL with a Kd of 320 nM affinity degrades its target RIPK2 with IC50 of 1 nM in cells)[14]; additionally,

previous work has shown that creative linker chemistry can positively affect the physicochemical properties of PROTACs[114].

FBXLs are another class of F-box E3 ligases, relying on a leucine-rich repeat (LRR) for substrate binding (Fig. 5b). There are 17 FBXLs, 15 of which are known to mediate protein degradation. The crystal structure of FBXL3 bound to a substrate does not reveal any well-defined pocket along the LRR domain[115], and the chemical tractability of FBXLs for PROTAC discovery is unclear at this time. Forty-two FBXO ligases, of which 20 at least are associated with the UPS, form the last class of F-box E3 ligases. A variety of substrate-binding domains can be used by FBXO ligases, including an F-box-associated (FBA) domain, found in six of them (Fig. 5b). Structural studies of FBXO44 indicate the presence of cavities at the surface of the FBA domain, but no ligand has been described so far[116].

IAP E3 ligases.

IAP proteins constitute a small class of five E3 ligases that bind substrate proteins via their baculoviral IAP repeat (BIR) domains (Fig. 5c). Unlike the E3 ligase families discussed above, IAP E3 ligases interact directly with E2 proteins via a RING domain. Due to their anti-apoptotic function, IAPs are targets for cancer therapy, and small-molecule antagonists exploiting the substrate-binding site, including non-peptido-mimetic compounds, have been developed with low-nanomolar potency[117,118,119] (Fig. 5c). The binding pocket is structurally conserved, but significant side-chain diversity exists between homologues (Fig. 5c), and ligands with narrow selectivity profiles were recently reported[117]. Small molecules recruiting IAPs for target degradation (also known as SNIPERs) were among the first PROTACs described. Initial compounds relied on the IAP ligand bestatin, which binds BIRC2 with moderate affinity and induces its auto-ubiquitylation and degradation, thereby limiting the effect on targeted substrates[120,121]. Next-generation PROTACs later derived from more potent peptido-mimetic ligands of IAPs were shown to efficiently induce the knockdown of diverse proteins such as ERα, BCR–ABL, BRD4 or PDE4 (ref.[122]). While an ERα-degrading PROTAC bound more potently to BIRC2, silencing XIAP had a more pronounced effect on the activity of the compound, indicating

that XIAP played a preponderant role in mediating target degradation, possibly due to the relative amount or subcellular location of the E3 ligases and substrate. The tissue expression profile of IAP E3 ligases is diverse, the BIR domain is chemically tractable, and future IAP-selective PROTACs should be valuable tools for chemical biology or therapeutic applications (as is indicated by patents WO2017182418 and WO2017211924).

APC E3 ligases.

The APC is a large, multisubunit E3 ligase that induces exit from mitosis by targeting cyclin B and securin for proteasomal degradation[123]. Substrate recognition is carried out by the WDR domain of CDC20 or the close homologue FZR1/Cdh1. As was discussed above, CDC20 is essential in all cancer types, but its expression level is low in the majority of normal tissues, making it an attractive E3 for PROTAC discovery. The structure of the CDC20 WDR domain was solved in complex with a substrate peptide in which a canonical RxxL D-box motif is inserted into a hydrophobic cavity on the side surface of the WDR domain[73]. Structural studies have shown that a D-box peptide can bind a similar pocket in the WDR domain of FZR1 (ref.[124]). A small-molecule ligand, apcin, that binds CDC20 with low-micromolar affinity occupies the D-box binding site of the ligase, suggesting that PROTACs exploiting this site could be developed[62] (Fig. 5d). These data position CDC20 at an interesting intersection of favourable tissue expression profile and promising chemical tractability, which could be exploited to chemically induce the degradation of oncogenes.

***HECT E3 ligases*.**

The ubiquitin ligase activity of the 29 human HECT E3 ligases, most of them involved in the UPS, relies on a reaction intermediate in which ubiquitin chains form a thioester bond with the catalytic HECT domain, followed by transfer to the substrate. It is therefore expected that compounds binding the HECT domain would act as catalytic inhibitors, and future PROTACs should instead exploit other domains. For instance, six HECT E3 ligases (HERC1–HERC6) contain a β-propeller RCC1-like domain (RLD) with a toroidal shape structurally related to the WDR and Kelch domains (Protein Data Bank codes: 3kci, 4o2w, 4l1m) (Fig. 6a). No small-molecule ligand

has been reported so far for RLDs, but the central cavity is deep and probably amenable to PROTAC discovery. Nine HECT E3 ligases feature a WW substrate-binding domain (Fig. 6a). The crystal structures of ITCH and NEDD4 WW domains in complex with peptide substrates reveal a shallow but hydrophobic binding site that accommodates proline-rich motifs, but with unclear ligandability[125].

TRIM E3 ligases.

TRIM proteins are a family of about 73 E3 ligases that directly interact with ubiquitin-conjugated E2 proteins via a canonical RING domain. Of these, 31 are known to be involved in the UPS, but this number will probably grow, as many TRIM E3 ligases are not functionally characterized. These proteins typically homodimerize via a central coiled-coil domain and bind their substrate via a C-terminal module, generally a SPRY domain (Fig. 6b). The SPRY domain of TRIM21, a major autoantigen in autoimmune diseases, was solved in complex with the Fc region of the immunoglobulin IgG[126], revealing a well-defined pocket at the IgG binding site that could be targeted by PROTACs (Fig. 6b). The corresponding site is shallow in the homologue TRIM25 (ref.[127]), and the chemical tractability of the SPRY domain — for which no ligand has been reported so far — is unclear.

A clearly ligandable domain found at the C terminus of TRIM24, TRIM28 and TRIM33 is the bromodomain. This structural module recognizes acetylated lysines and has emerged in recent years as a promising target class in oncology and inflammation[128]. In fact, a small-molecule ligand can bind with low-nanomolar affinity to the bromodomain of TRIM24, and a crystal structure shows that the inhibitor is deeply inserted into the acetyl-lysine binding pocket of the bromodomain[129] (Fig. 6b). All three ubiquitin ligases are involved in the UPS, and TRIM24 can bind an acetylated peptide of the tumour suppressor p53, leading to ubiquitylation and degradation of the target[130]. In this regard, PROTACs derived from the existing TRIM24 ligand may simultaneously induce the degradation of novel substrates via a catalytic mechanism and stabilize the endogenous substrate p53 by a classical occupancy-based competition mechanism. Conversely, a bromodomain ligand was recently linked to a VHL-recruiting chemical handle to

degrade TRIM24 (ref.[27]). TRIM28 is ubiquitously expressed (Fig. 3) and is a good candidate for developing proteasome-targeting chemical biology tools applicable across a diverse array of cellular systems.

A last substrate-binding domain of interest, located at the C terminus of three TRIM E3 ligases —TRIM2, TRIM3 and TRIM32 — is the NHL domain. No NHL structure is available for these proteins, but structural studies of an unrelated protein indicate that the NHL domain adopts a β-propeller topology (PDB code 5ex7) very similar to that of the WDR, Kelch and RLD domains found in other ligases, suggesting that it could be amenable to the development of future PROTACs recruiting UPS-involved TRIM32 (Fig. 6b).

While we highlight here the major classes of E3 ligases, we expect that atypical proteins that are not part of a clearly defined group will also prove to be amenable to PROTAC discovery. For example, GID4, the subunit of the GID E3 ligase complex, features a substrate-binding domain with a deep, enclosed and ligandable pocket that recognizes the amino-terminal proline residue of protein substrates, leading to their proteasomal degradation[131,132].

Outlook

Looking to the future, important questions remain. What fraction of the 'undruggable' proteome will prove amenable to targeted degradation, and which targets will lead to truly transformational therapies? While important progress has been made, considerable work remains. We propose that developing additional E3 ligase 'tools' will be an enabling accelerant for the field as we begin to more fully establish factors that govern efficient target–ligase pairings. In addition, as different E3 ligases are expressed at different times and in different tissues, an additional layer of opportunity (and complexity) may be realized through tissue-specific target knockdown. Because E3 ligase 'tools' primarily represent ligands, we propose direct binding assays as attractive and appropriate hit identification strategies for E3 ligases. One opportunity lies in DNA-encoded libraries (DELs), which have the potential to sample vast chemical space in the span of one or more closely related experiments (>109 compounds)[133].

Alternatively, fragment-based approaches provide an opportunity to screen a more limited compound set chosen from chemically diverse and attractive lower-molecular-weight pharmacophores. In either scenario, we believe progression from initial hits to optimized tools will be greatly aided by the application of structure-based design and molecular modelling.

While the size, breadth and diversity of E3 ligases present considerable challenges, they present similarly substantial opportunities. Indeed, reflecting on the progress made with compounds exploiting only a handful of ligases, it will be interesting to see what the future of protein degradation holds as we scratch below the surface of this gene family.

References

1. Ravid, T. & Hochstrasser, M. Diversity of degradation signals in the ubiquitin–proteasome system. *Nature Reviews Molecular Cell Biology* **9**, 679–689 (2008).
2. Sakamoto, K. M. *et al.* Protacs: Chimeric molecules that target proteins to the Skp1–Cullin–F box complex for ubiquitination and degradation. *PNAS* **98**, 8554–8559 (2001).
3. 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).
4. Itoh, Y., Ishikawa, M., Naito, M. & Hashimoto, Y. Protein knockdown using methyl bestatin-ligand hybrid molecules: design and synthesis of inducers of ubiquitination-mediated degradation of cellular retinoic acid-binding proteins. *J. Am. Chem. Soc.* **132**, 5820–5826 (2010).
5. Ito, T. *et al.* Identification of a Primary Target of Thalidomide Teratogenicity. *Science* **327**, 1345–1350 (2010).

6. 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 CRL4CRBN. *British Journal of Haematology* **164**, 811–821 (2014).
7. Krönke, J. *et al.* Lenalidomide Causes Selective Degradation of IKZF1 and IKZF3 in Multiple Myeloma Cells. *Science* **343**, 301–305 (2014).
8. Lu, G. *et al.* The Myeloma Drug Lenalidomide Promotes the Cereblon-Dependent Destruction of Ikaros Proteins. *Science* **343**, 305–309 (2014).
9. Sievers, Q. L. *et al.* Defining the human C2H2 zinc finger degrome targeted by thalidomide analogs through CRBN. *Science* **362**, eaat0572 (2018).
10. Lai, A. C. *et al.* Modular PROTAC Design for the Degradation of Oncogenic BCR-ABL. *Angewandte Chemie International Edition* **55**, 807–810 (2016).
11. Fisher, S. L. & Phillips, A. J. Targeted protein degradation and the enzymology of degraders. *Curr Opin Chem Biol* **44**, 47–55 (2018).
12. Bondeson, D. P. *et al.* Lessons in PROTAC Design from Selective Degradation with a Promiscuous Warhead. *Cell Chemical Biology* **25**, 78-87.e5 (2018).
13. Huang, H.-T. *et al.* A Chemoproteomic Approach to Query the Degradable Kinome Using a Multi-kinase Degrader. *Cell Chemical Biology* **25**, 88-99.e6 (2018).
14. Bondeson, D. P. *et al.* Catalytic *in vivo* protein knockdown by small-molecule PROTACs. *Nature Chemical Biology* **11**, 611–617 (2015).
15. Krönke, J. *et al.* Lenalidomide induces ubiquitination and degradation of CK1α in del(5q) MDS. *Nature* **523**, 183–188 (2015).

16. Matyskiela, M. E. *et al.* A Cereblon Modulator (CC-220) with Improved Degradation of Ikaros and Aiolos. *J. Med. Chem.* **61**, 535–542 (2018).

17. Gaudy, A. *et al.* SAT0225 Cereblon modulator CC-220 decreases naÏve and memory B cells and plasmacytoid dendritic cells in systemic lupus erythematosus (SLE) patients: exposure-response results from a phase 2A proof of concept study. *Annals of the Rheumatic Diseases* **76**, 858–859 (2017).

18. Sun, X. *et al.* A chemical approach for global protein knockdown from mice to non-human primates. *Cell Discovery* **5**, 10 (2019).

19. Li, Y. *et al.* Discovery of MD-224 as a First-in-Class, Highly Potent, and Efficacious Proteolysis Targeting Chimera Murine Double Minute 2 Degrader Capable of Achieving Complete and Durable Tumor Regression. *J. Med. Chem.* **62**, 448–466 (2019).

20. Buhimschi, A. D. *et al.* Targeting the C481S Ibrutinib-Resistance Mutation in Bruton's Tyrosine Kinase Using PROTAC-Mediated Degradation. *Biochemistry* **57**, 3564–3575 (2018).

21. Mullard, A. First targeted protein degrader hits the clinic. *Nature Reviews Drug Discovery* **18**, 237 (2019).

22. Bondeson, D. P. & Crews, C. M. Targeted Protein Degradation by Small Molecules. *Annu. Rev. Pharmacol. Toxicol.* **57**, 107–123 (2017).

23. Olson, C. M. *et al.* Pharmacological perturbation of CDK9 using selective CDK9 inhibition or degradation. *Nature Chemical Biology* **14**, 163–170 (2018).

24. Churcher, I. Protac-Induced Protein Degradation in Drug Discovery: Breaking the Rules or Just Making New Ones? *J. Med. Chem.* **61**, 444–452 (2018).

25. Burslem, G. M. *et al.* The Advantages of Targeted Protein Degradation Over Inhibition: An RTK Case Study. *Cell Chemical Biology* **25**, 67-77.e3 (2018).

26. Bassi, Z. I. *et al.* Modulating PCAF/GCN5 Immune Cell Function through a PROTAC Approach. *ACS Chem. Biol.* **13**, 2862–2867 (2018).

27. Gechijian, L. N. *et al.* Functional TRIM24 degrader via conjugation of ineffectual bromodomain and VHL ligands. *Nature Chemical Biology* **14**, 405 (2018).

28. Cromm, P. M., Samarasinghe, K. T. G., Hines, J. & Crews, C. M. Addressing Kinase-Independent Functions of Fak via PROTAC-Mediated Degradation. *J. Am. Chem. Soc.* **140**, 17019–17026 (2018).

29. Smith, B. E. *et al.* Differential PROTAC substrate specificity dictated by orientation of recruited E3 ligase. *Nature Communications* **10**, 131 (2019).

30. Brand, M. *et al.* Homolog-Selective Degradation as a Strategy to Probe the Function of CDK6 in AML. *Cell Chemical Biology* **26**, 300-306.e9 (2019).

31. Zhang, X., Crowley, V. M., Wucherpfennig, T. G., Dix, M. M. & Cravatt, B. F. Electrophilic PROTACs that degrade nuclear proteins by engaging DCAF16. *Nat Chem Biol.* **15**, 737–746 (2019).

32. Silva, M. C. *et al.* Targeted degradation of aberrant tau in frontotemporal dementia patient-derived neuronal cell models. *eLife* **8**, e45457 (2019).

33. Douglass, E. F., Miller, C. J., Sparer, G., Shapiro, H. & Spiegel, D. A. A Comprehensive Mathematical Model for Three-Body Binding Equilibria. *J. Am. Chem. Soc.* **135**, 6092–6099 (2013).

34. Buckley, D. L. *et al.* HaloPROTACS: Use of Small Molecule PROTACs to Induce Degradation of HaloTag Fusion Proteins. *ACS Chem. Biol.* **10**, 1831–1837 (2015).

35. Zhang, L., Riley-Gillis, B., Vijay, P. & Shen, Y. Acquired Resistance to BET-PROTACs(Proteolysis Targeting Chimeras) Caused by Genomic Alterations in Core Components of E3 ligase Complexes. *Mol Cancer Ther* molcanther.1129.2018 (2019). doi:10.1158/1535-7163.MCT-18-1129

36. Edmondson, S. D., Yang, B. & Fallan, C. Proteolysis targeting chimeras (PROTACs) in 'beyond rule-of-five' chemical space: Recent progress and future challenges. *Bioorganic & Medicinal Chemistry Letters* **29**, 1555–1564 (2019).

37. Crew, A. P. *et al.* Identification and Characterization of Von Hippel-Lindau-Recruiting Proteolysis Targeting Chimeras (PROTACs) of TANK-Binding Kinase 1. *J. Med. Chem.* **61**, 583–598 (2018).

38. Zoppi, V. *et al.* Iterative Design and Optimization of Initially Inactive Proteolysis Targeting Chimeras (PROTACs) Identify VZ185 as a Potent, Fast, and Selective von Hippel–Lindau (VHL) Based Dual Degrader Probe of BRD9 and BRD7. *J. Med. Chem.* **62**, 699–726 (2019).

39. Popow, J. *et al.* Highly Selective PTK2 Proteolysis Targeting Chimeras to Probe Focal Adhesion Kinase Scaffolding Functions. *J. Med. Chem.* **62**, 2508–2520 (2019).

40. Dobrovolsky, D. *et al.* Bruton tyrosine kinase degradation as a therapeutic strategy for cancer. *Blood* **133**, 952–961 (2019).

41. Jiang, B. *et al.* Development of Dual and Selective Degraders of Cyclin-Dependent Kinases 4 and 6. *Angewandte Chemie International Edition* **58**, 6321–6326 (2019).

42. Powell, C. E. *et al.* Chemically Induced Degradation of Anaplastic Lymphoma Kinase (ALK). *J. Med. Chem.* **61**, 4249–4255 (2018).

43. McCoull, W. *et al.* Development of a Novel B-Cell Lymphoma 6 (BCL6) PROTAC To Provide Insight into Small Molecule Targeting of BCL6. *ACS Chem. Biol.* **13**, 3131–3141 (2018).

44. Ward, C. C. *et al.* Covalent Ligand Screening Uncovers a RNF4 E3 Ligase Recruiter for Targeted Protein Degradation Applications. *ACS Chem. Biol.* (2019). doi:10.1021/acschembio.8b01083

45. Tinworth, C. P. *et al.* PROTAC-Mediated Degradation of Bruton's Tyrosine Kinase Is Inhibited by Covalent Binding. *ACS Chem. Biol.* **14**, 342–347 (2019).

46. Nowak, R. P. *et al.* Plasticity in binding confers selectivity in ligand-induced protein degradation. *Nature Chemical Biology* **14**, 706 (2018).

47. Petzold, G., Fischer, E. S. & Thomä, N. H. Structural basis of lenalidomide-induced CK1α degradation by the $CRL4^{CRBN}$ ubiquitin ligase. *Nature* **532**, 127–130 (2016).

48. Matyskiela, M. E. *et al.* A novel cereblon modulator recruits GSPT1 to the $CRL4^{CRBN}$ ubiquitin ligase. *Nature* **535**, 252–257 (2016).

49. Cardote, T. A. F., Gadd, M. S. & Ciulli, A. Crystal Structure of the Cul2-Rbx1-EloBC-VHL Ubiquitin Ligase Complex. *Structure* **25**, 901-911.e3 (2017).

50. Angers, S. *et al.* Molecular architecture and assembly of the DDB1–CUL4A ubiquitin ligase machinery. *Nature* **443**, 590 (2006).

51. Gadd, M. S. *et al.* Structural basis of PROTAC cooperative recognition for selective protein degradation. *Nature Chemical Biology* **13**, 514–521 (2017).

52. Fischer, E. S. *et al.* The Molecular Basis of CRL4DDB2/CSA Ubiquitin Ligase Architecture, Targeting, and Activation. *Cell* **147**, 1024–1039 (2011).

53. Drummond, M. L. & Williams, C. I. In Silico Modeling of PROTAC-Mediated Ternary Complexes: Validation and Application. *J. Chem. Inf. Model.* **59**, 1634–1644 (2019).

54. Zorba, A. *et al.* Delineating the role of cooperativity in the design of potent PROTACs for BTK. *PNAS* **115**, E7285–E7292 (2018).

55. Liu, L. *et al.* UbiHub: a data hub for the explorers of ubiquitination pathways. *Bioinformatics* doi:10.1093/bioinformatics/bty1067

56. Komander, D. & Rape, M. The Ubiquitin Code. *Annual Review of Biochemistry* **81**, 203–229 (2012).

57. Chen, Z. J. & Sun, L. J. Nonproteolytic functions of ubiquitin in cell signaling. *Mol. Cell* **33**, 275–286 (2009).

58. Mészáros, B., Kumar, M., Gibson, T. J., Uyar, B. & Dosztányi, Z. Degrons in cancer. *Sci. Signal.* **10**, eaak9982 (2017).

59. Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).

60. Hines, J., Lartigue, S., Dong, H., Qian, Y. & Crews, C. M. MDM2-Recruiting PROTAC Offers Superior, Synergistic Antiproliferative Activity via Simultaneous Degradation of BRD4 and Stabilization of p53. *Cancer Res* **79**, 251–262 (2019).

61. Meyers, R. M. *et al.* Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. *Nature Genetics* **49**, 1779–1784 (2017).

62. Sackton, K. L. *et al.* Synergistic blockade of mitotic exit by two chemical inhibitors of the APC/C. *Nature* **514**, 646–649 (2014).

63. Ottis, P. *et al.* Assessing Different E3 Ligases for Small Molecule Induced Protein Ubiquitination and Degradation. *ACS Chem. Biol.* **12**, 2570–2578 (2017).

64. Lee, J. & Zhou, P. DCAFs, the Missing Link of the CUL4-DDB1 Ubiquitin Ligase. *Molecular Cell* **26**, 775–780 (2007).

65. Uehara, T. *et al.* Selective degradation of splicing factor CAPERα by anticancer sulfonamides. *Nature Chemical Biology* **13**, 675–680 (2017).

66. Han, T. *et al.* Anticancer sulfonamides target splicing by inducing RBM39 degradation via recruitment to DCAF15. *Science* **356**, eaal3755 (2017).

67. Schapira, M., Tyers, M., Torrent, M. & Arrowsmith, C. H. WD40 repeat domain proteins: a novel target class? *Nature Reviews Drug Discovery* **16**, 773–786 (2017).

68. Cao, Q. *et al.* The central role of EED in the orchestration of polycomb group complexes. *Nature Communications* **5**, 3127 (2014).

69. He, Y. *et al.* The EED protein–protein interaction inhibitor A-395 inactivates the PRC2 complex. *Nature Chemical Biology* **13**, 389–395 (2017).

70. Qi, W. *et al.* An allosteric PRC2 inhibitor targeting the H3K27me3 binding pocket of EED. *Nature Chemical Biology* **13**, 381–388 (2017).

71. Grebien, F. *et al.* Pharmacological targeting of the Wdr5-MLL interaction in C/EBPα N-terminal leukemia. *Nat. Chem. Biol.* **11**, 571–578 (2015).

72. Song, R., Wang, Z.-D. & Schapira, M. Disease Association and Druggability of WD40 Repeat Proteins. *J. Proteome Res.* **16**, 3766–3773 (2017).

73. Zhang, S. *et al.* Molecular mechanism of APC/C activation by mitotic phosphorylation. *Nature* **533**, 260–264 (2016).

74. Canning, P. *et al.* Structural Basis for Cul3 Protein Assembly with the BTB-Kelch Family of E3 Ubiquitin Ligases. *J. Biol. Chem.* **288**, 7803–7814 (2013).

75. Zhuang, M. *et al.* Structures of SPOP-Substrate Complexes: Insights into Molecular Architectures of BTB-Cul3 Ubiquitin Ligases. *Molecular Cell* **36**, 39–50 (2009).

76. McMahon, M., Thomas, N., Itoh, K., Yamamoto, M. & Hayes, J. D. Dimerization of Substrate Adaptors Can Facilitate Cullin-mediated Ubiquitylation of Proteins by a "Tethering" Mechanism A TWO-SITE INTERACTION MODEL FOR THE Nrf2-Keap1 COMPLEX. *J. Biol. Chem.* **281**, 24756–24768 (2006).

77. Tong, K. I. *et al.* Keap1 Recruits Neh2 through Binding to ETGE and DLG Motifs: Characterization of the Two-Site Molecular Recognition Model. *Molecular and Cellular Biology* **26**, 2887–2900 (2006).

78. Zhang, Q. *et al.* Multiple Ser/Thr-rich degrons mediate the degradation of Ci/Gli by the Cul3-HIB/SPOP E3 ubiquitin ligase. *PNAS* **106**, 21191–21196 (2009).

79. Lu, M. *et al.* Discovery of a Keap1-dependent peptide PROTAC to knockdown Tau by ubiquitination-proteasome degradation pathway. *European Journal of Medicinal Chemistry* **146**, 251–259 (2018).

80. Lo, S.-C., Li, X., Henzl, M. T., Beamer, L. J. & Hannink, M. Structure of the Keap1:Nrf2 interface provides mechanistic insight into Nrf2 signaling. *The EMBO Journal* **25**, 3605–3617 (2006).

81. Schumacher, F.-R., Sorrell, F. J., Alessi, D. R., Bullock, A. N. & Kurz, T. Structural and biochemical characterization of the KLHL3–WNK kinase interaction important in blood pressure regulation. *Biochemical Journal* **460**, 237–246 (2014).

82. Chen, Z., Picaud, S., Filippakopoulos, P., D'Angiolella, V. & Bullock, A. N. Structural basis for recruitment of DAPK1 to the KLHL20 E3 ligase. *Structure* **27**, 1–10 (2019).

83. Cuadrado, A. *et al.* Therapeutic targeting of the NRF2 and KEAP1 partnership in chronic diseases. *Nature Reviews Drug Discovery* **18**, 295 (2019).

84. Davies, T. G. *et al.* Monoacidic Inhibitors of the Kelch-like ECH-Associated Protein 1: Nuclear Factor Erythroid 2-Related Factor 2 (KEAP1:NRF2) Protein–Protein Interaction with High Cell Potency Identified by Fragment-Based Discovery. *J. Med. Chem.* **59**, 3991–4006 (2016).

85. Brockmann, M. *et al.* Genetic wiring maps of single-cell protein states reveal an off-switch for GPCR signalling. *Nature* **546**, 307–311 (2017).

86. Dementieva, I. S. *et al.* Pentameric assembly of potassium channel tetramerization domain-containing protein 5. *J. Mol. Biol.* **387**, 175–191 (2009).

87. Chen, H.-Y., Liu, C.-C. & Chen, R.-H. Cul3-KLHL20 ubiquitin ligase: physiological functions, stress responses, and disease implications. *Cell Div* **11**, 5 (2016).

88. Jerabkova, K. & Sumara, I. Cullin 3, a cellular scripter of the non-proteolytic ubiquitin code. *Semin. Cell Dev. Biol.* (2018). doi:10.1016/j.semcdb.2018.12.007

89. Angers, S. *et al.* The KLHL12-Cullin-3 ubiquitin ligase negatively regulates the Wnt-beta-catenin pathway by targeting Dishevelled for degradation. *Nat. Cell Biol.* **8**, 348–357 (2006).

90. McGourty, C. A. *et al.* Regulation of the CUL3 Ubiquitin Ligase by a Calcium-Dependent Co-adaptor. *Cell* **167**, 525-538.e14 (2016).

91. Scott, D. C. *et al.* Two Distinct Types of E3 Ligases Work in Unison to Regulate Substrate Ubiquitylation. *Cell* **166**, 1198-1214.e24 (2016).

92. Jin, L. *et al.* Ubiquitin-dependent regulation of COPII coat size and function. *Nature* **482**, 495–500 (2012).

93. Skieterska, K., Rondou, P., Lintermans, B. & Van Craenenbroeck, K. KLHL12 Promotes Non-Lysine Ubiquitination of the Dopamine Receptors D4.2 and D4.4, but Not of the ADHD-Associated D4.7 Variant. *PLoS ONE* **10**, e0145654 (2015).

94. Smaldone, G. *et al.* Cullin 3 Recognition Is Not a Universal Property among KCTD Proteins. *PLoS ONE* **10**, e0126808 (2015).

95. Chen, H. Y. *et al.* KLHL39 suppresses colon cancer metastasis by blocking KLHL20-mediated PML and DAPK ubiquitination. *Oncogene* **34**, 5141–5151 (2015).

96. Mahrour, N. *et al.* Characterization of Cullin-box sequences that direct recruitment of Cul2-Rbx1 and Cul5-Rbx2 modules to Elongin BC-based ubiquitin ligases. *J. Biol. Chem.* **283**, 8005–8013 (2008).

97. Qi, H. *et al.* Molecular cloning and characterization of the von Hippel-Lindau-like protein. *Mol. Cancer Res.* **2**, 43–52 (2004).

98. Koren, I. *et al.* The Eukaryotic Proteome Is Shaped by E3 Ubiquitin Ligases Targeting C-Terminal Degrons. *Cell* **173**, 1622-1635.e14 (2018).

99. Rusnac, D.-V. *et al.* Recognition of the Diglycine C-End Degron by CRL2KLHDC2 Ubiquitin Ligase. *Mol. Cell* **72**, 813-822.e4 (2018).

100. Linossi, E. M. & Nicholson, S. E. The SOCS box-adapting proteins for ubiquitination and proteasomal degradation. *IUBMB Life* **64**, 316–323 (2012).

101. Guo, Y. *et al.* Structural basis for hijacking CBF-β and CUL5 E3 ligase complex by HIV-1 Vif. *Nature* **505**, 229–233 (2014).

102. Nucifora, F. C. *et al.* Ubiqutination via K27 and K29 chains signals aggregation and neuronal protection of LRRK2 by WSB1. *Nat Commun* **7**, 11792 (2016).

103. Kim, J. J. *et al.* WSB1 promotes tumor metastasis by inducing pVHL degradation. *Genes Dev.* **29**, 2244–2257 (2015).

104. Zheng, S. *et al.* Comprehensive Pan-Genomic Characterization of Adrenocortical Carcinoma. *Cancer Cell* **29**, 723–736 (2016).

105. Muniz, J. R. C. *et al.* Molecular architecture of the ankyrin SOCS box family of Cul5-dependent E3 ubiquitin ligases. *J. Mol. Biol.* **425**, 3166–3177 (2013).

106. Fei, X. *et al.* Crystal structure of Human ASB9-2 and substrate-recognition of CKB. *Protein J.* **31**, 275–284 (2012).

107. Bergamin, E., Wu, J. & Hubbard, S. R. Structural basis for phosphotyrosine recognition by suppressor of cytokine signaling-3. *Structure* **14**, 1285–1292 (2006).

108. Kershaw, N. J. *et al.* SOCS3 binds specific receptor-JAK complexes to control cytokine signaling by direct kinase inhibition. *Nat. Struct. Mol. Biol.* **20**, 469–476 (2013).

109. Babon, J. J. *et al.* The structure of SOCS3 reveals the basis of the extended SH2 domain function and identifies an unstructured insertion that regulates stability. *Mol. Cell* **22**, 205–216 (2006).

110. Filippakopoulos, P. *et al.* Structural basis for Par-4 recognition by the SPRY domain- and SOCS box-containing proteins SPSB1, SPSB2, and SPSB4. *J. Mol. Biol.* **401**, 389–402 (2010).

111. Sadek, M. M. *et al.* A Cyclic Peptide Inhibitor of the iNOS-SPSB Protein-Protein Interaction as a Potential Anti-Infective Agent. *ACS Chem. Biol.* **13**, 2930–2938 (2018).

112. Yatsu, A., Shimada, H., Ohbayashi, N. & Fukuda, M. Rab40C is a novel Varp-binding protein that promotes proteasomal degradation of Varp in melanocytes. *Biol Open* **4**, 267–275 (2015).

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

114. Qin, C. *et al.* Discovery of QCA570 as an Exceptionally Potent and Efficacious Proteolysis Targeting Chimera (PROTAC) Degrader of the Bromodomain and Extra-Terminal (BET) Proteins Capable of Inducing Complete and Durable Tumor Regression. *J. Med. Chem.* **61**, 6685–6704 (2018).

115. Xing, W. *et al.* SCF(FBXL3) ubiquitin ligase targets cryptochromes at their cofactor pocket. *Nature* **496**, 64–68 (2013).

116. Kumanomidou, T. *et al.* The Structural Differences between a Glycoprotein Specific F-Box Protein Fbs1 and Its Homologous Protein FBG3. *PLoS ONE* **10**, e0140366 (2015).

117. Tamanini, E. *et al.* Discovery of a Potent Nonpeptidomimetic, Small-Molecule Antagonist of Cellular Inhibitor of Apoptosis Protein 1 (cIAP1) and X-Linked Inhibitor of Apoptosis Protein (XIAP). *J. Med. Chem.* **60**, 4611–4625 (2017).

118. Chessari, G. *et al.* Fragment-Based Drug Discovery Targeting Inhibitor of Apoptosis Proteins: Discovery of a Non-Alanine Lead Series with Dual Activity Against cIAP1 and XIAP. *J. Med. Chem.* **58**, 6574–6588 (2015).

119. Fulda, S. & Vucic, D. Targeting IAP proteins for therapeutic intervention in cancer. *Nat Rev Drug Discov* **11**, 109–124 (2012).

120. Okuhira, K. *et al.* Specific degradation of CRABP-II via cIAP1-mediated ubiquitylation induced by hybrid molecules that crosslink cIAP1 and the target protein. *FEBS Lett.* **585**, 1147–1152 (2011).

121. Sekine, K. *et al.* Small molecules destabilize cIAP1 by activating auto-ubiquitylation. *J. Biol. Chem.* **283**, 8961–8968 (2008).

122. Ohoka, N. *et al.* In Vivo Knockdown of Pathogenic Proteins via Specific and Nongenetic Inhibitor of Apoptosis Protein (IAP)-dependent Protein Erasers (SNIPERs). *J. Biol. Chem.* **292**, 4556–4570 (2017).

123. Peters, J.-M. The anaphase promoting complex/cyclosome: a machine designed to destroy. *Nat. Rev. Mol. Cell Biol.* **7**, 644–656 (2006).

124. Chang, L., Zhang, Z., Yang, J., McLaughlin, S. H. & Barford, D. Atomic structure of the APC/C and its mechanism of protein ubiquitination. *Nature* **522**, 450–454 (2015).

125. Qi, S., O'Hayre, M., Gutkind, J. S. & Hurley, J. H. Structural and biochemical basis for ubiquitin ligase recruitment by arrestin-related domain-containing protein-3 (ARRDC3). *J. Biol. Chem.* **289**, 4743–4752 (2014).

126. James, L. C., Keeble, A. H., Khan, Z., Rhodes, D. A. & Trowsdale, J. Structural basis for PRYSPRY-mediated tripartite motif (TRIM) protein function. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 6200–6205 (2007).

127. Koliopoulos, M. G. *et al.* Molecular mechanism of influenza A NS1-mediated TRIM25 recognition and inhibition. *Nat Commun* **9**, 1820 (2018).

128. Filippakopoulos, P. & Knapp, S. Targeting bromodomains: epigenetic readers of lysine acetylation. *Nature Reviews Drug Discovery* **13**, 337–356 (2014).

129. Palmer, W. S. *et al.* Structure-Guided Design of IACS-9571, a Selective High-Affinity Dual TRIM24-BRPF1 Bromodomain Inhibitor. *J. Med. Chem.* **59**, 1440–1454 (2016).

130. Allton, K. *et al.* Trim24 targets endogenous p53 for degradation. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 11612–11616 (2009).

131. Dong, C. *et al.* Molecular basis of GID4-mediated recognition of degrons for the Pro/N-end rule pathway. *Nat. Chem. Biol.* **14**, 466–473 (2018).

132. Chen, S.-J., Wu, X., Wadas, B., Oh, J.-H. & Varshavsky, A. An N-end rule pathway that recognizes proline and destroys gluconeogenic enzymes. *Science* **355**, (2017).

133. Neri, D. & Lerner, R. A. DNA-Encoded Chemical Libraries: A Selection System Based on Endowing Organic Compounds with Amplifiable Information. *Annu. Rev. Biochem.* **87**, 479–502 (2018).

134. You, T. *et al.* Crystal structure of SPSB2 in complex with a rational designed RGD-containing cyclic peptide inhibitor of SPSB2-iNOS interaction. *Biochem. Biophys. Res. Commun.* **489**, 346–352 (2017).

Further information

Acknowledgements

The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada through Ontario Genomics Institute (OGI-055, Innovative Medicines Initiative (EU/EFPIA, ULTRA-DD grant no. 115766), Janssen, Merck KGaA, Darmstadt, Germany, MSD, Novartis Pharma AG, Innovation and Science (MRIS), Pfizer, São Paulo Research Foundation-FAPESP, Takeda and Wellcome (grant 106169/ZZ14/Z). M.S. gratefully acknowledges support from NSERC (grant RGPIN-2019-04416). Research in the C.M.C. lab is supported by grant NIH R35CA197589 and by Arvinas.

Competing interests

M.F.C is an employee of Pfizer. C.M.C. is a consultant and shareholder in Arvinas, which provides research support to his lab.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Related links

Arvinas press release: http://ir.arvinas.com/news-releases/news-release-details/arvinas-present-preclinical-tau-directed-protacr-protein

The Human Protein Atlas: https://www.proteinatlas.org/

Display items

Figure 1 | Substrate recruitment in targeted protein degradation. Schematic representations (left) and crystal structures (right) for substrate recruitment to the E3 ligase cereblon (CRBN) by a hetero-bifunctional proteolysis-targeting chimera (PROTAC; part **a**) or an immunomodulatory drug (IMiD; part **b**)[46,47]. Ligands are shown as stick representations in isolation, in the centre (Protein Data Bank codes: 6boy and 5fqd), or as space-filling models bound in the E3–target complex, to the right. DDB1, DNA damage-binding protein 1.

Figure 2 | Structures of ternary complexes formed during targeted protein degradation. a, b | The VHL (part **a**) and cereblon (CRBN; part **b**) Cullin–RING E3 ligases form a large U-shaped multiprotein complex that positions the ubiquitin-conjugated E2 enzymes in physical proximity to the target protein bromodomain-containing protein 4 (BRD4). The expanded images show how the ligase-binding and target-binding moieties (orange and cyan, respectively), as well as the linker region (white) of the proteolysis-targeting chimera (PROTAC), all contribute to the formation of the ternary complex. The Cullin–RING ligase complexes were each composed by combining two crystal structures. VHL complex: BRD4–PROTAC–VHL–Elongin B/C (ELOB/C) (Protein Data Bank (PDB) code: 5t35) and VHL–ELOB/C–CUL2–RBX1 (PDB code: 5n4w). CRBN complex: BRD4–PROTAC–CRBN–DDB1 (PDB code: 6bn7) and DDB1–CUL4–RBX1 (PDB code: 2hye). **c** | Different PROTACs can induce different binding interfaces between the ligase and the target protein (here, CRBN and BRD4; PDB codes: 6bn7, 6boy and 6bnb, from left to right). Only the carboxy-terminal domain (CTD) of CRBN is shown. The resolution of the BRD4–dBET57–CRBN structure was not sufficient to properly model the PROTAC dBET57, which is symbolized with a green oval. **d** | A glycine-containing beta-hairpin makes a conserved interaction with the immunomodulatory drug (IMiD) scaffold in all CRBN–IMiD–target structures available to date (PDB codes: CK1α, 5fqd; GSPT1, 5hxb; IKFZ1, 6h0f; ZNF692, 6h0g).

Figure 3 | Tissue expression of E3 ligases. a | Ubiquitous E3 ligases. The figure shows ligases with medium to high protein levels in at least 90% of tissues or cell types tested. **b** | Tissue-selective E3 ligases. The protein level of tissue-selective E3 ligases is shown. The data in parts **a** and **b** are based on the protein expression levels in 81 tissues or cell types from the Human Protein Atlas (proteinatlas.org)[59]. Protein levels classified as 'uncertain' by the Human Protein Atlas were ignored. Throughout the figure, E3 ligases associated with the ubiquitin–proteasome system (UPS) in the literature are in bold, and E3 ligases exploited by current PROTACs are indicated with an asterisk. **c** | E3 ligases that are essential in cancer. Colour-coding indicates the median vulnerability score to CRISPR knockout across multiple cell lines for a given tissue type. The number of cell lines is indicated in square brackets next to each tissue type. Only E3 ligases with at least one median dependency score <−1.0 are shown. The last column shows expression data for non-cancer cells from the Human Protein Atlas (data classified as 'uncertain' are not shown). CERES dependency scores corrected for copy number variations are taken from the 30 May 2018 version of the dependency portal at depmap.org. The recommended threshold below which a gene is considered essential is −1 (ref.[61]).

Figure 4 | Ligandability of E3 ligases: DCAF and BTB E3 ligases . a | Almost all DDB1–CUL4-associated factor (DCAF) E3 ligases have a WDR domain. The WDR domain of DCAF1 contributes to substrate binding, and the WDR domain of EED is druggable. Cereblon (CRBN) uses an atypical domain to bind its substrate and is exploited by thalidomide and its analogues. **b** | The BTB domain-containing proteins are a large subfamily of E3 ligases, many of which are known to be involved in the ubiquitin–proteasome system (UPS). Small-molecule ligands can bind to the Kelch domain of KEAP1 with nanomolar affinity. The geometries and electrostatic potentials of the Kelch and WDR domains are diverse, indicating variable ligandability. Blue indicates electropositive, and red indicates electronegative. E3 ligases reported in the published literature to signal for substrate degradation are indicated in bold. E3 ligases for which proteolysis targeting chimeras (PROTACs) have been reported are indicated with an asterisk (in black for chemical or white for peptidic compounds). Protein Data Bank codes: CRBN, 6h0g;

DCAF1, 5jk7; EED, 5k0m; KCTD5, 3drx; KEAP1, 5fnu; KLHL3, 4ch9; KLHL20, 6gy5; KLHL40, 4asc. IMiD, immunomodulatory drug.

Figure 5 | Ligandability of E3 ligases: BC-box, F-box, IAP and APC E3 ligases. **a** | BC-box E3 ligases feature a variety of substrate-binding domains. Ankyrin repeats are the most represented and include two juxtaposed hydrophobic cavities in the structure of ASB9. SH2 domains recruit phosphodegrons and are generally too polar for the development of drug-like inhibitors. The carboxy terminus of USP1 binds the Kelch domain of KLHDC2 at a polar but well-defined pocket. A substrate-derived cyclic peptide targets a shallow site at the surface of the SPSB2 SPRY domain[134]. The E3 ligases listed in italic bind to Cullin 2, and the others bind to Cullin 5. **b** | F-box E3 ligases can use diverse domains for substrate recruitment. Phosphodegron motifs (yellow) bind the central cavity of the BTRC WDR domain, and a phosphopeptide linked to oestradiol recruits oestrogen receptor (ER) to BTRC, leading to ER ubiquitylation and degradation[113]. Other structural modules repeatedly used for substrate recognition, but with unexplored chemical tractability, are the leucine-rich repeat (LRR) domain and the F-box-associated (FBA) domain. **c** | Inhibitor of apoptosis (IAP) E3 ligases feature a baculoviral IAP repeat (BIR) domain that can be exploited for proteolysis targeting chimera (PROTAC) discovery[122]. The ligand-binding pocket is not deep but is hydrophobic (green, hydrophobic; blue/red, hydrogen-bond donors/acceptors, respectively). The BIR3 domain is structurally conserved (bottom left) but features significant side-chain diversity (bottom centre and right). Residues lining the pocket that are not conserved in the multiple-sequence alignment (bottom centre) are shown on the crystal structure. **d** | Anaphase-promoting complex (APC) E3 ligase subunits. A pocket at the side of the WDR domain of the APC-co-activating E3 ligases CDC20 and FZR1 is exploited by the chemical inhibitor apcin (yellow) and by substrate peptides (red)[62]. E3 ligases reported in the published literature to signal for substrate degradation are indicated in bold. E3 ligases for which PROTACs have been reported are indicated with an asterisk (in black for chemical or white for peptidic compounds). Protein Data Bank codes: SPSB2, 5xn3; ASB9, 4uuc; SOCS6, 2vif; KLHDC2, 6do5; FBXL3, 4i6j; BTRC, 1p22; FBXO44, 3wso; XIAP, 5M6L; CDC20, 4n14; FZR1, 4ui9; XIAP, 5m6l.

Figure 6 | Ligandability of E3 ligases: HECT and TRIM E3 ligases. a | The RCC1 repeat is a structural module found in multiple HECT E3 ligases, structurally related to WDR and Kelch domains, with a deep central cavity that may be chemically tractable. The WW domain is also recurrently used for substrate recruitment. The substrate-binding site is hydrophobic (green patches) but may be too shallow for the development of chemical E3 handles. **b** | TRIM proteins are a large subfamily of standalone E3 ligases. Substrate recruitment is generally achieved via a SPRY domain that may be chemically tractable. Ligandable pockets are found on the bromodomain of a few TRIM ligases. Another rare alternative for substrate recruitment is the NHL repeat. Although no structure is available for human proteins, the NHL repeat of the Drosophila melanogaster (Dm) protein Brat reveals a ligandable structure related to WDRs. E3 ligases reported in the published literature to signal for substrate degradation are indicated in bold. Protein Data Bank codes: HERC2, 3kci; ITCH, 4rof; NEDD4, 4n7h; DmBrat, 5ex7; TRIM21, 2iwg; TRIM24, 4yc9.

Table 1 | Selected recent lessons learned on opportunities and challenges for targeted protein degradation.

E3 ligase	Target	Learning point	mechanism of action	selectivity	permeability	druggability	design	structure	in vivo / clinical	Date	Ref
CRBN, VHL	PTK2	Inhibiting PGP efflux pump increases potency; >80% target depletion not sufficient to replicate effect of genetic knockdown; off target PDE6 not hit by parent inhibitor								2019	Popow et al. JMC
VHL	BRD7/9	Empirical nature of E3 ligase selection: VHL originally thought incompatible with BRD7/9; Linker chemistry is fastidious and empirical								2019	Zoppi et al.JMC
CRBN,VHL	BET BRDs	Acquired reistance of cancer cells to PROTACs is driven by genomic alteration of Cullin or chromosomal deletion of CRBN								2019	Zhang et al. MCT
IAP	BTK	Covalent binding to target inhibits degradation								2019	Tinworth et al. ACB
CRBN	Tau	Selective degradation of aberrant/pathogenic form of Tau								2019	Silva eLIFE
VHL	p38	PROTAC is more selective than parent inhibitor; linker length and attachment point dictate isoform selectivity								2019	Smith Nat.Comm
CRBN	CDK6	CRBN-recruiting PROTACs is more selective than parent inhibitor and degrades IMiD targets								2019	Jiang Angew. Chem.
CRBN	CDK6	PROTAC is more selective than parent inhibitor								2019	Brand Cell Chem Bio
CRBN, VHL	BRD4	Predicting complex structure								2019	Drummond JCIM
CRBN	FKBP12	Global target knock-down in most tissues, except in brain								2019	Sun Cell Disc.
CRBN	MDM2	Tumor regression in mouse model								2019	Li JMC
DCAF16	FKBP12	Covalent binding to E3 ligase is compatible with PROTAC modality; harnessing DCAF16 leads to degradation of nuclear proteins only								2018	Zhang biorXiv
CRBN, VHL	Kinases	A promiscuous ligand can be incorporated into a more selective PROTAC								2018	Bondeson Cell Chem Bio
CRBN	CDK9	Prolonged effect of degradation over inhibition; narrow dosage window due to hook effect								2018	Olson NCB
CRBN	BTK	Addressing a resistance mechanism to a covalent inhibitor using a non-covalent PROTAC								2018	Buhimschi Biochem.
MDM2	BRD4	Synergy between degradation of PROTAC target and stabiization of endogenous E3 substrate								2018	Hines Cancer Res.
CRBN	ZnF	IMiDs can degrade undruggable ZnF transcription factors; ternary complex structure								2018	Sievers Science
CRBN	BTK	CRBN-recruiting PROTACs degrade IMiD targets; change in a single linker atom can abolish off-target degradation								2018	Dobrovolsky Blood
VHL	FAK	PROTAC is more selective than parent inhibitor								2018	Cromm JACS
CRBN	BCL6	Permeability 400x lower than parent BCL6 ligand								2018	McCoull ACB
CRBN	ALK	Potency compromised in cells expressing drug transporter ABCB1								2018	Powell JMC
VHL	TRIM24	Targeting functionally neutral domain to degrade undruggable protein								2018	Gechijian NCB
CRBN	PCAF/GCN5	Targeting functionally neutral domain to degrade undruggable protein								2018	Bassi ACB
CRBN	BRD4	Ternary complex structure								2018	Nowak NCB
VHL	c-Met	Prolonged effect of degradation over inhibition								2017	Burslem Cell Chem Bio
VHL	TBK1	Target engagement does not necessarily lead to degradation								2017	Crew JMC
VHL	BRD4	Ternary complex structure								2017	Gadd NCB
CRBN	IKFZ1/3	Positive Phase 2a clinical response to C-220 in SLE patients								2017	Gaudy ARD
CRBN	GSPT1	Ternary complex structure								2016	Matyskiela Nature
CRBN	CK1α	Ternary complex structure								2016	Petzold Nature
VHL	ERRα, RIPK2	Catalytic nature drives potency								2015	Bondeson NCB
VHL	HaloTag:eGFP	Positive ternary complex cooperativity mitigates the Hook Effect								2015	Buckley ABC
CRBN	CK1α	Degradarion of CK1α contributes to efficacy of lenalidomide in del(5q) myelodysplastic syndrome.								2015	Kronke Nature
CRBN	IKFZ1/3	Efficacy of lenalidomide in multiple myeloma driven by induced degradation of neosubstrates								2014	Kronke Science

BRD, bromodomain; CRBN, cereblon; IAP, inhibitor of apoptosis; IMiD, immunomodulatory drug; MoA, mechanism of action; PGP, P-glycoprotein; PROTAC, proteolysis-targeting chimera; SLE, systemic lupus erythematosus.

Figure 1

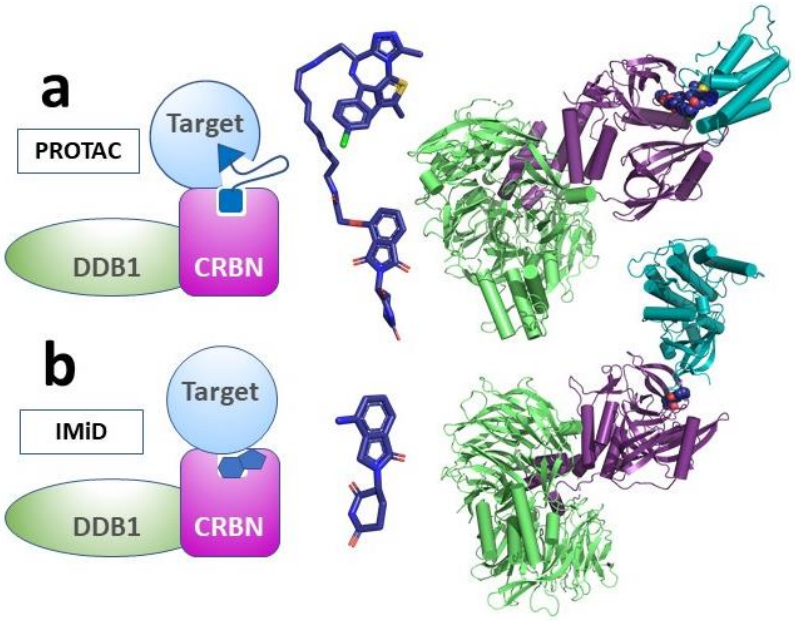


Figure 2

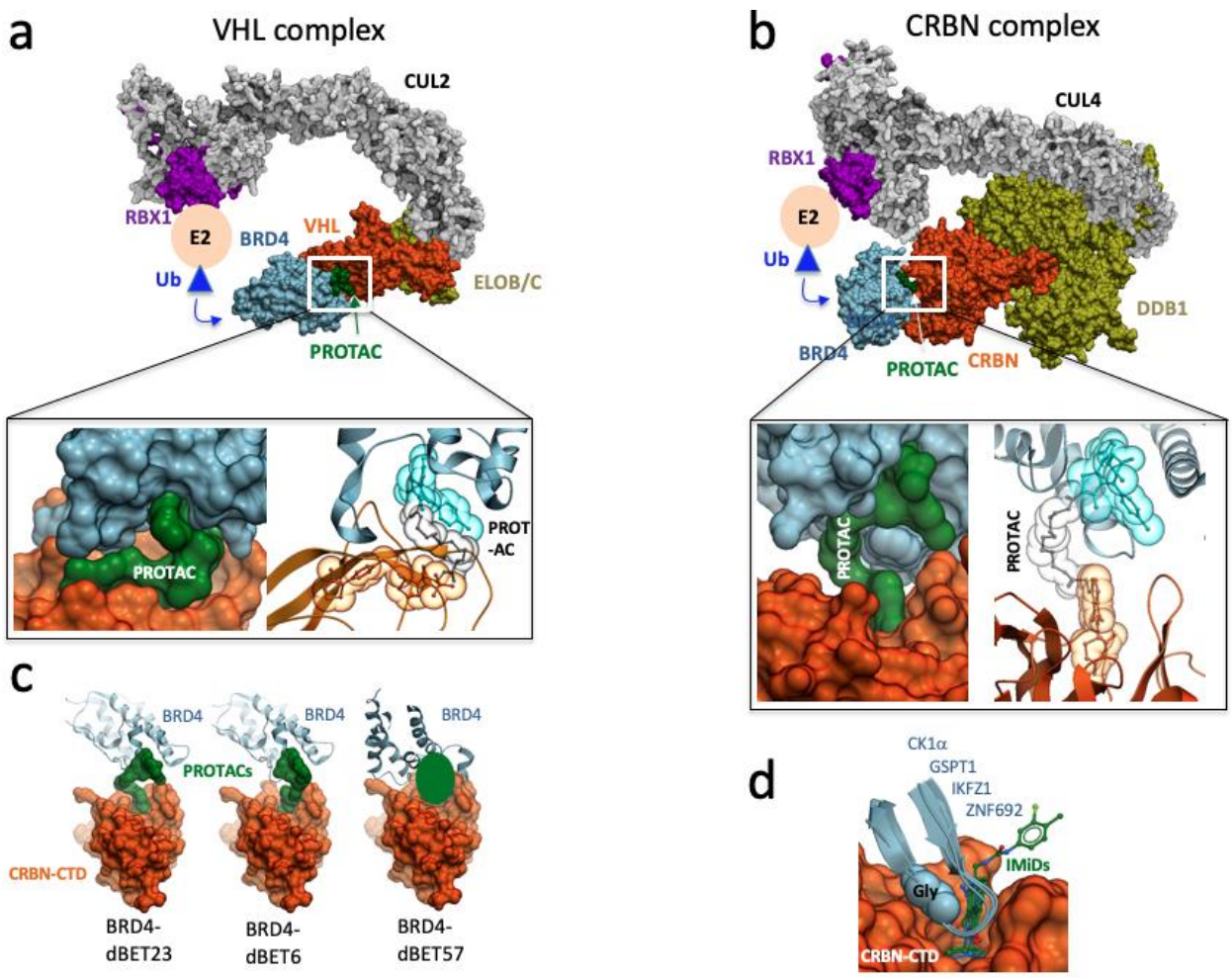
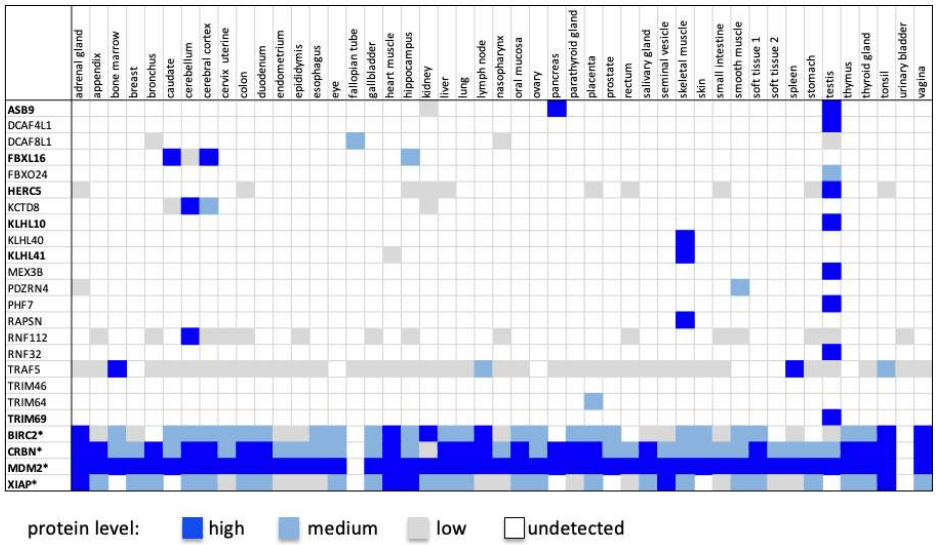

a
VHL complex
CUL2
RBX1
E2
BRD4
VHL
Ub
ELOB/C
PROTAC
PROTAC
PROT
-AC
b
CRBN complex
CUL4
RBX1
E2
Ub
DDB1
BRD4
PROTAC
CRBN
PROTAC
PROTAC
c
BRD4
BRD4
BRD4
PROTACs
CRBN-CTD
BRD4-
dBET23
BRD4-
dBET6
BRD4-
dBET57
d
CK1α
GSPT1
IKFZ1
ZNF692
IMiDs
Gly
CRBN-CTD

Figure 3

A

Protein	% samples where protein levels are medium to high
RBBP7	98.8
MDM2*	98.8
TOPORS	98.7
SNRNP40	97.6
KLHL5	97.5
KDM2A	97.4
TRIM35	96.3
TRIM28	96.3
RBBP4	96.3
RNF214	96.1
PRPF19	95.2
TRIM7	95.0
FBXW7	94.9
TLE1	94.9
RHOBTB1	94.8
CBLL1	93.8
SHPRH	93.6
CNOT4	93.6
UBE3B	92.4
KLHL36	92.1
PPIL2	91.6
UBE3A	91.4
FBXO18	90.9
RNF20	90.1
CRBN*	76.5
BIRC2*	48.8
XIAP*	41.5

B



C

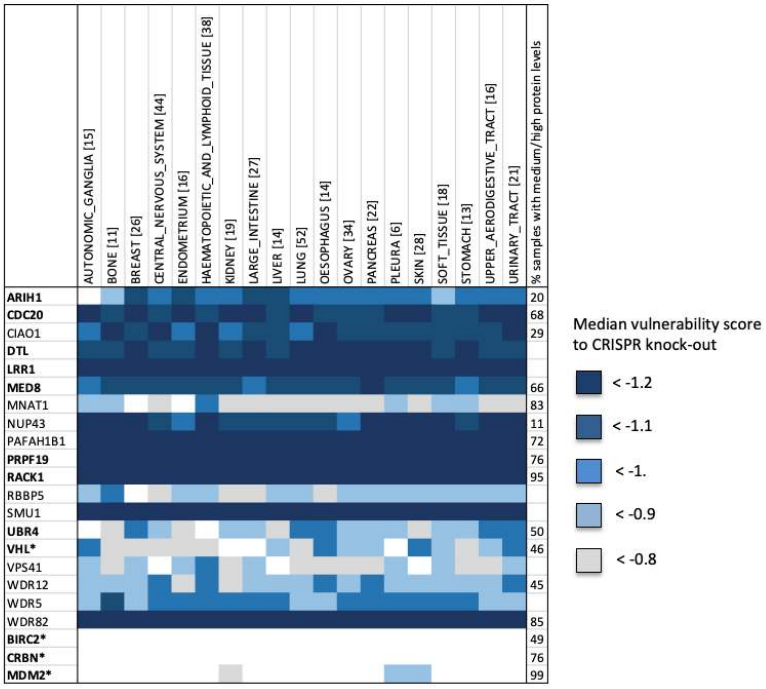


Figure 4

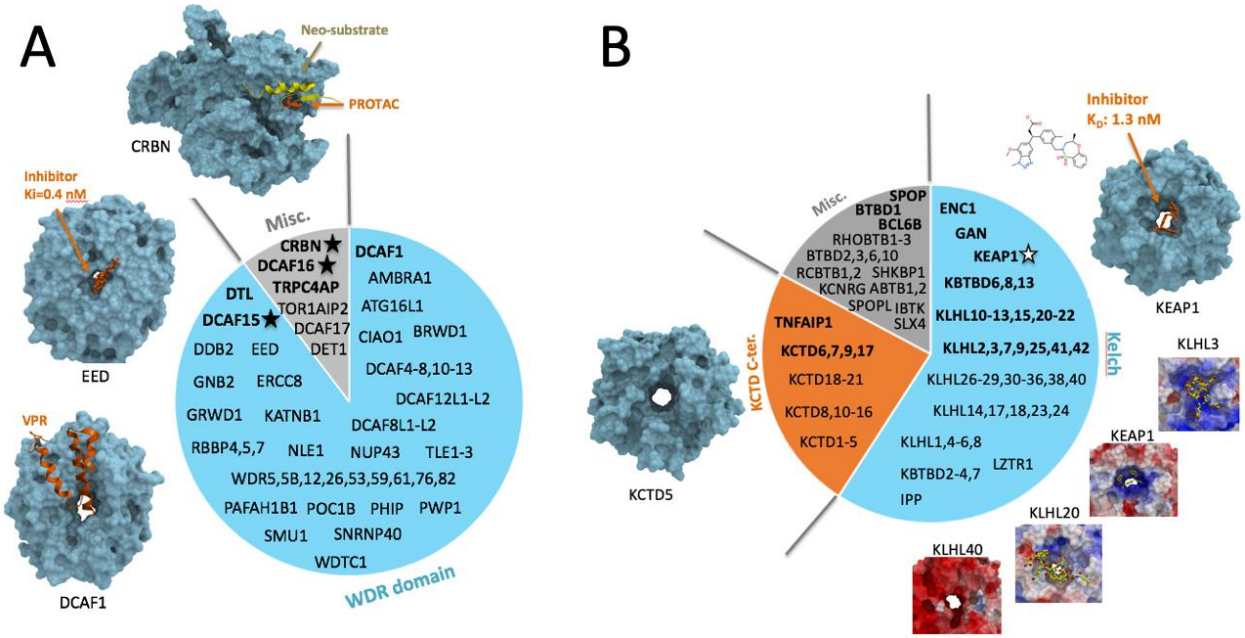

A
Neo-substrate
PROTAC
CRBN
Inhibitor
Ki=0.4 nM
EED
VPR
DCAF1
Misc.
CRBN
DCAF16
TRPC4AP
DTL
DCAF15
DCAF1
AMBRA1
ATG16L1
TOR1AIP2
DCAF17
CIAO1
BRWD1
DDB2
EED
DET1
DCAF4-8,10-13
GNB2
ERCC8
DCAF12L1-L2
GRWD1
KATNB1
DCAF8L1-L2
RBBP4,5,7
NLE1
NUP43
TLE1-3
WDR5,5B,12,26,53,59,61,76,82
PAFAH1B1
POC1B
PHIP
PWP1
SMU1
SNRNP40
WDTC1
WDR domain
B
Inhibitor
Kd: 1.3 nM
KEAP1
KLHL3
KEAP1
KLHL20
KLHL40
KCTD5
Misc.
SPOP
BTBD1
BCL6B
RHOBTB1-3
BTBD2,3,6,10
RCBTB1,2
SHKBP1
KCNRG
ABTB1,2
SPOPL
IBTK
SLX4
KCTD C-ter.
TNFAIP1
KCTD6,7,9,17
KCTD18-21
KCTD8,10-16
KCTD1-5
Kelch
ENC1
GAN
KEAP1
KBTBD6,8,13
KLHL10-13,15,20-22
KLHL2,3,7,9,25,41,42
KLHL26-29,30-36,38,40
KLHL14,17,18,23,24
KLHL1,4-6,8
KBTBD2-4,7
LZTR1
IPP

Figure 5

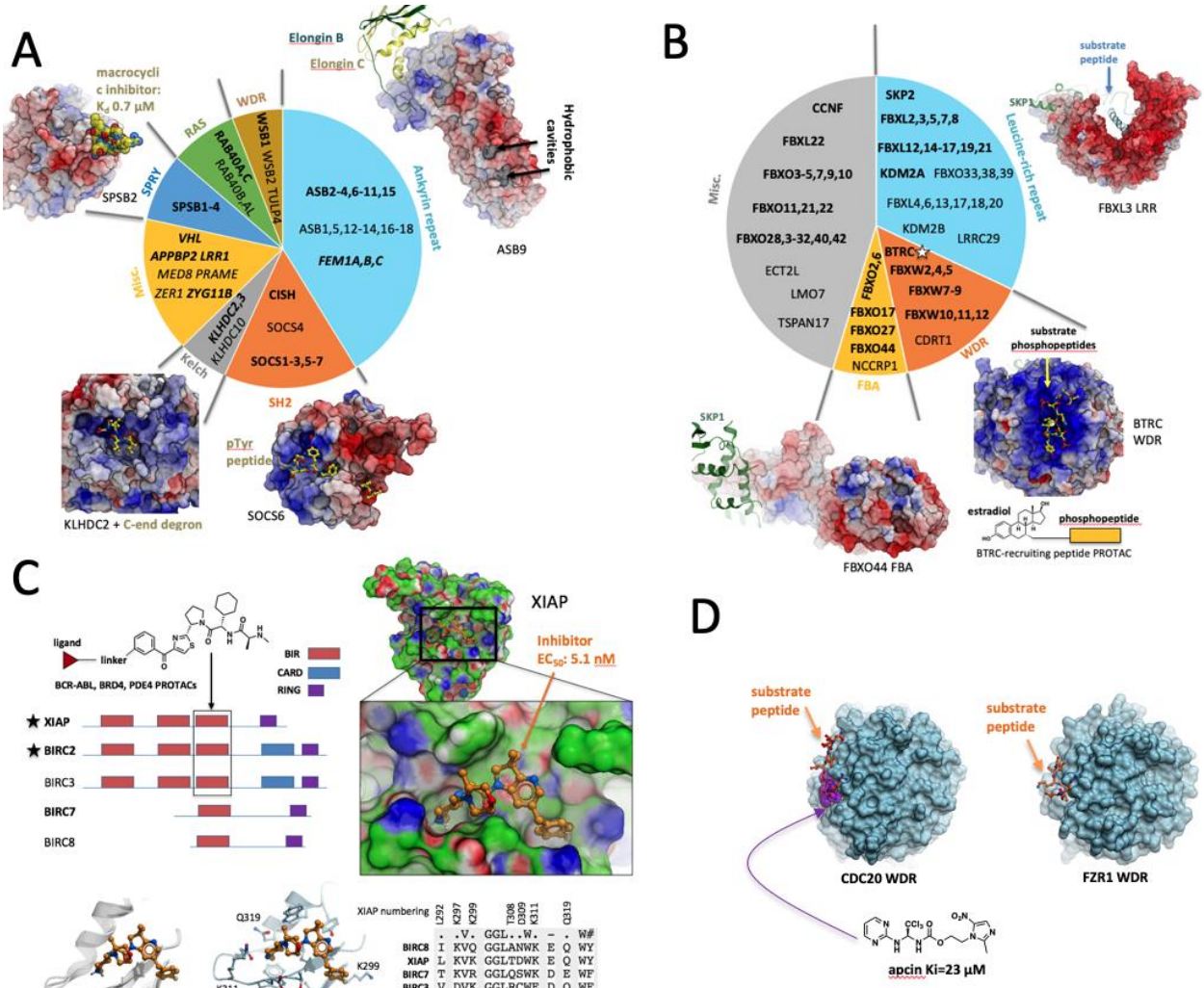
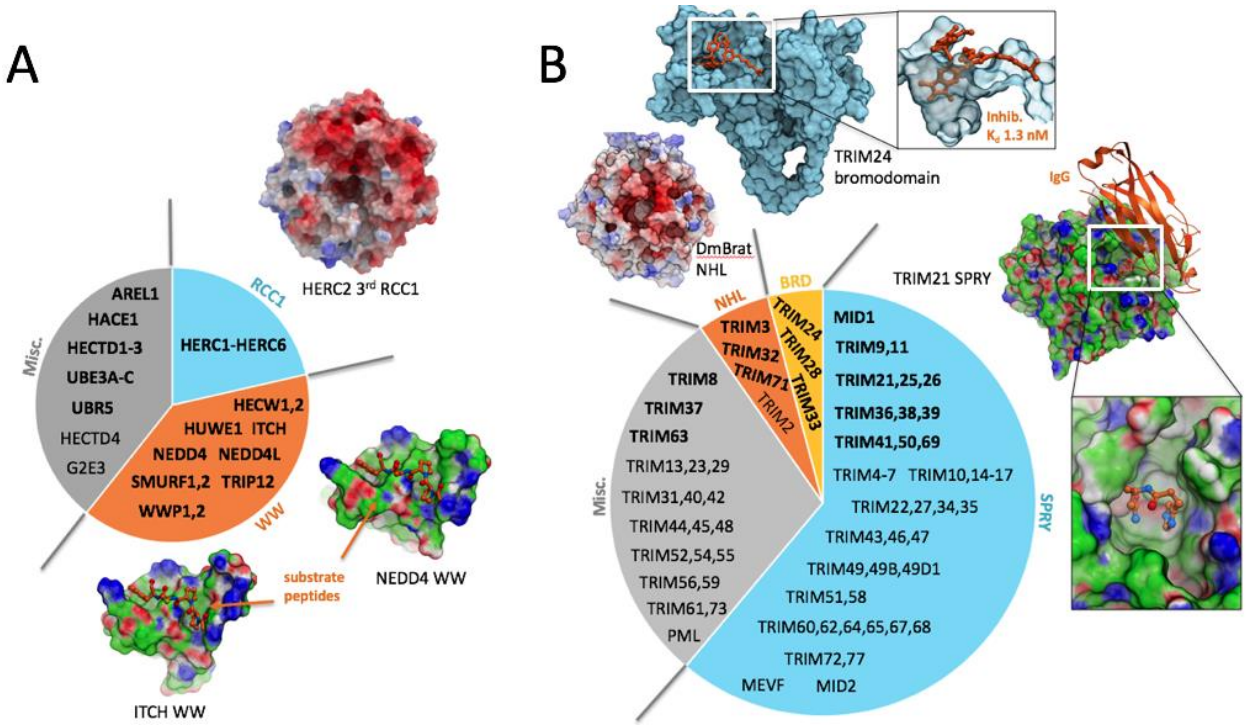

A
macrocyclic inhibitor: K_d 0.7 µM
SPSB2
RAS
SPRY
WDR
RAB40A,C
RAB40B,AL
WSB1 WSB2 TULP4
SPSB1-4
VHL
APPBP2 LRR1
MED8 PRAME
ZER1 ZYG11B
Misc.
KLHDC2,3
KLHDC10
Kelch
CISH
SOCS4
SOCS1-3,5-7
SH2
ASB2-4,6-11,15
ASB1,5,12-14,16-18
FEM1A,B,C
Ankyrin repeat
Elongin B
Elongin C
Hydrophobic cavities
ASB9
pTyr peptide
SOCS6
KLHDC2 + C-end degron
B
CCNF
FBXL22
FBXO3-5,7,9,10
FBXO11,21,22
FBXO28,3-32,40,42
ECT2L
LMO7
TSPAN17
Misc.
SKP2
FBXL2,3,5,7,8
FBXL12,14-17,19,21
KDM2A FBXO33,38,39
FBXL4,6,13,17,18,20
KDM2B
LRRC29
Leucine-rich repeat
FBXO2,6
BTRC
FBXW2,4,5
FBXW7-9
FBXO17
FBXO27
FBXO44
NCCRP1
FBA
FBXW10,11,12
CDRT1
WDR
substrate peptide
SKP1
FBXL3 LRR
substrate phosphopeptides
BTRC WDR
SKP1
FBXO44 FBA
estradiol
phosphopeptide
BTRC-recruiting peptide PROTAC
C
ligand
linker
BCR-ABL, BRD4, PDE4 PROTACs
BIR
CARD
RING
XIAP
BIRC2
BIRC3
BIRC7
BIRC8
XIAP
Inhibitor EC_{50}: 5.1 nM
Q319
K299
K311
D309
T308
K297
L292
XIAP numbering
L292 K297 K299 T308 D309 K311 Q319
BIRC8 I KVQ GGLANWK E Q WY
XIAP L KVK GGLTDWK E Q WY
BIRC7 T KVR GGLQSWK D E WF
BIRC3 V DVK GGLRCWE D Q WF
BIRC2 V DVK GGLRCWE D E WF
D
substrate peptide
substrate peptide
CDC20 WDR
FZR1 WDR
apcin Ki=23 µM

Figure 6


A
HERC2 3rd RCC1
RCC1
HERC1-HERC6
Misc.
AREL1
HACE1
HECTD1-3
UBE3A-C
UBR5
HECTD4
G2E3
HECW1,2
HUWE1 ITCH
NEDD4 NEDD4L
SMURF1,2 TRIP12
WWP1,2
WW
substrate peptides
NEDD4 WW
ITCH WW
B
TRIM24 bromodomain
Inhib. Kd 1.3 nM
DmBrat NHL
IgG
TRIM21 SPRY
NHL
BRD
TRIM3
TRIM32
TRIM71
TRIM2
TRIM24
TRIM28
TRIM33
MID1
TRIM9,11
TRIM21,25,26
TRIM36,38,39
TRIM41,50,69
TRIM4-7 TRIM10,14-17
TRIM22,27,34,35
TRIM43,46,47
TRIM49,49B,49D1
TRIM51,58
TRIM60,62,64,65,67,68
TRIM72,77
MID2
SPRY
Misc.
TRIM8
TRIM37
TRIM63
TRIM13,23,29
TRIM31,40,42
TRIM44,45,48
TRIM52,54,55
TRIM56,59
TRIM61,73
PML
MEVF