Degrons in cancer

Bálint Mészáros,1* Manjeet Kumar,2* Toby J. Gibson,2† Bora Uyar,3 Zsuzsanna Dosztányi4†

Degrons are the elements that are used by E3 ubiquitin ligases to target proteins for degradation. Most degrons are short linear motifs embedded within the sequences of modular proteins. As regulatory sites for protein abundance, they are important for many different cellular processes, such as progression through the cell cycle and monitoring cellular hypoxia. Degrons enable the elimination of proteins that are no longer required, preventing their possible dysfunction. Although the human genome encodes ~600 E3 ubiquitin ligases, only a fraction of these enzymes have well-defined target degrons. Thus, for most cellular proteins, the destruction mechanisms are poorly understood. This is important for many diseases, especially for cancer, a disease that involves the enhanced expression of oncogenes and the persistence of encoded oncoproteins coupled with reduced abundance of tumor suppressors. Loss-of-function mutations occur in the degrons of several oncoproteins, such as the transcription factors MYC and NRF2, and in various mitogenic receptors, such as NOTCH1 and several receptor tyrosine kinases. Mutations eliminating the function of the β-catenin degron are found in many cancers and are considered one of the most abundant mutations driving carcinogenesis. In this Review, we describe the current knowledge of degrons in cancer and suggest that increased research on the “dark degrome” (unknown degron-E3 relationships) would enhance progress in cancer research.

Cancer as a Disease of aberrant Protein Degradation

Cancer is well known to be a disease of gene mutation. More than half of cancers have point mutations in the TP53 gene, compromising the role of the encoded protein p53 (also known as TP53) as the guardian of the genome (1–3). However, cancer is also a disease of aberrant expression of genes. For example, gene duplication, aneuploidy (an imbalance in the number of chromosomes), or remodeling of chromatin state can cause either increases in expression of oncogenes encoding tumor-promoting proteins (oncoproteins) or decreases in expression of genes encoding tumor suppressors (4–7). Gene mutations and gene expression changes are often tightly connected. For example, a mutation that inactivates a tumor suppressor that functions as a negative regulator of gene expression can lead to increased expression of an oncogene. Thus, as a prerequisite to understand any given cancer, both the set of mutations and the gene expression changes need to be fully cataloged.

In this Review, we focus on the importance of degrons in oncogenesis. As the destruction signals in cellular proteins, degrons serve as a nexus between gene mutation and protein abundance in cancer. If degron function is impeded by any form of mutation, the result is that a protein will have a longer half-life, enhancing and prolonging its function in the cell. Nevertheless, with the exception of the frequently mutated degron in the multifunctional protein β-catenin (also known as CTNNB1) (8, 9), mutations that affect degron function have a low profile in the cancer literature. To demonstrate the importance of degron mutations in cancer, we provide an overview of the many ways that degrons, their mutations, and interactions in the E3 ligase systems that govern their recognition and degradation of the degron-containing proteins influence cellular transformation and oncogenesis.

Details of the Ubiquitin-Proteasome System

How the ubiquitin-proteasome degradation pathway works

Intracellular protein degradation in eukaryotes is mainly achieved by the ubiquitin-proteasome system (UPS) (10). The UPS plays an essential role in maintaining cellular homeostasis by participating in protein quality control, for instance, by recognizing and rapidly degrading incorrectly folded or assembled proteins. In addition to the elimination of aberrant proteins, this system is responsible for the dynamic protein turnover involved in cell regulation. The spatially and temporally regulated degradation of proteins is critical for diverse cellular processes, including cell cycle progression, signaling, differentiation, and growth (10).

The UPS has a sophisticated network of enzymes that targets proteins for degradation. The core degradation machine of this system is the 26S proteasome, which consists of two main structural components: the 20S core particle, which functions as a gate through which substrates enter and are degraded, and the 19S regulatory particle, which controls the opening of this gate to selectively let substrate proteins with the proper degradation signals into the core particle (11, 12).

Although various posttranslational modifications, such as poly-adenosine 5′-diphosphate-riboseylation (PARylation) (13) and poly- sumoylation (14), can target substrates to the 26S proteasome, the canonical signal for 26S-mediated proteasomal degradation are the Lys48 (K48)-linked polyubiquitin chains, which are generated by attachment of ubiquitin to a lysine residue of the substrate protein and then repetitive attachment of ubiquitin to Lys48 of the last-attached ubiquitin. A chain of at least four ubiquitin moieties targets the substrate protein to the 26S proteasome (15, 16).

The machinery of ubiquitination

The covalent transfer of ubiquitin, a 76–amino acid protein with a mass of 8 kDa, to substrate proteins is mediated by three types of enzymes: the E1 ubiquitin-activating enzyme, the E2 ubiquitin-conjugating enzyme, and the E3 ubiquitin ligase enzyme (Fig. 1) (17). The human genome contains 2 genes encoding E1 enzymes, 41 genes encoding E2 enzymes, and more than 600 genes encoding E3 enzymes (18), which reflects the importance of the variation of E3 enzymes in the complex spatiotemporal regulation of posttranslational protein stability in the cells. Ubiquitin is activated by the E1 enzyme in an adenosine

1Institute of Enzymology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, 2 Magyar Tudósok krt, Budapest H-1117, Hungary. 2Structural and Computational Biology Unit, European Molecular Biology Laboratory, Meyerhofstraße 1, 69117 Heidelberg, Germany. 3Bioinformatics Platform, Berlin Institute for Medical Systems Biology, Max-Delbrück Center for Molecular Medicine, Robert-Rössle-Strasse 10, 13125 Berlin, Germany. 4MTA-ELTE Lendület Bioinformatics Research Group, Department of Biochemistry, Eötvös Loránd University, Pázmány Péter sétány 1/c, Budapest H-1117, Hungary. †These authors contributed equally to this work. *Corresponding author. Email: dosztanyi@caesar.elte.hu (Z.D.); toby.gibson@embl.de (T.J.G.)
such as adaptor proteins, scaffold proteins, substrate receptor proteins, and accessory proteins. Apart from the RING domain function as part of protein complexes, such as Cullin-RING E3 ligase (CRL) complexes or anaphase-promoting complex (APC/C), some RING-type E3 ligases can function individually; some ligases have E3 ligase activity are divided into two main families on the basis of the core domain that facilitates the transfer of ubiquitin to the substrates. The first family of E3 ligases contain the HECT (homologous to E6-AP C terminus) domain, and the second family of E3 ligases contain the RING (really interesting new gene) domain or variants of it. For HECT E3 ligases, ubiquitin is sequentially transferred from the E2 to the HECT domain and then to the substrate, whereas for the RING E3 ligases, the ubiquitin is directly transferred to the substrate (Fig. 1) (19, 20). HECT E3 ligases comprise ~5% of the known E3 ligases; the rest are RING E3 ligases.

Some E3 ligases function as part of a complex and others function individually. The additional subunits that may be part of the complex include adaptors, scaffolds, and substrate receptors (Fig. 2). The class of RING E3 ligase complexes with the most members and diversity of subunits is the CRL complexes (21), which are further classified according to the scaffold protein Cullin in the complex (22–26). Most of these complexes are named for the adaptor subunit, Cullin scaffold subunit, and the substrate receptor subunit (Fig. 2). Different substrate receptor subunits enable distinct substrate recognition. The ECV and ECS complexes are different from the others in that the ECV and ECS complexes have two adaptor subunits: Elongin B and Elongin C. The ECV complex also only forms with a single receptor subunit, von Hippel-Lindau (VHL), whereas the others can have various receptor subunits. CRL3 is different in that it lacks an adaptor subunit and only contains the E3 ligase, the Cullin scaffold, and one of five substrate receptors. Another E3 ligase complex with RING-type E3 ligase activity is the APC/C, which controls cell cycle progression (27). APC2 is the scaffold with the Cullin domain, and APC11 is a RING box 1 (RBX1)–like RING finger protein that is the catalytic E3 subunit.

5′-triphosphate–dependent manner. The activated ubiquitin is, in turn, transferred to an E2 conjugating enzyme. The interaction of the E2 enzyme with a substrate-specific E3 enzyme leads to the formation of an isopeptide bond between the C-terminal glycine residue of ubiquitin and a lysine residue of the substrate protein (19). The attached ubiquitin proteins can be K48-ubiquitinated to form polyubiquitin chains attached to the substrate protein, thus targeting the substrate protein for 26S proteasome–mediated degradation (16). Proteins and protein complexes that have E3 ligase activity are divided into two main families on the basis of the core domain that facilitates the transfer of ubiquitin to the substrates. The first family of E3 ligases contain the HECT (homologous to E6-AP C terminus) domain (20), and the second family of E3 ligases contain the RING (really interesting new gene) domain or variants of it (19). For HECT E3 ligases, ubiquitin is sequentially transferred from the E2 to the HECT domain and then to the substrate, whereas for the RING E3 ligases, the ubiquitin is directly transferred to the substrate (Fig. 1) (19, 20). HECT E3 ligases comprise ~5% of the known E3 ligases; the rest are RING E3 ligases.

Some E3 ligases function as part of a complex and others function individually. The additional subunits that may be part of the complex include adaptors, scaffolds, and substrate receptors (Fig. 2). The class of RING E3 ligase complexes with the most members and diversity of subunits is the CRL complexes (21), which are further classified according to the scaffold protein Cullin in the complex (22–26). Most of these complexes are named for the adaptor subunit, Cullin scaffold subunit, and the substrate receptor subunit (Fig. 2). Different substrate receptor subunits enable distinct substrate recognition. The ECV and ECS complexes are different from the others in that the ECV and ECS complexes have two adaptor subunits: Elongin B and Elongin C. The ECV complex also only forms with a single receptor subunit, von Hippel-Lindau (VHL), whereas the others can have various receptor subunits. CRL3 is different in that it lacks an adaptor subunit and only contains the E3 ligase, the Cullin scaffold, and one of five substrate receptors. Another E3 ligase complex with RING-type E3 ligase activity is the APC/C, which controls cell cycle progression (27). APC2 is the scaffold with the Cullin domain, and APC11 is a RING box 1 (RBX1)–like RING finger protein that is the catalytic E3 subunit. Although
domain, which consists of an N-terminal RING-like domain, a C-terminal RING-like domain, and an IBR (in-between RING fingers) domain in the middle of the two RING-like domains (31).

**Alternative consequences of ubiquitination**
The fate of ubiquitinated proteins does not have to be degradation by the 26S proteasome (32). The activity of deubiquitinating enzymes (DUBs) that remove ubiquitins from the substrates enables this posttranslational modification to be reversible. Even in the absence of DUB activity, the type of ubiquitin signal dictates the fate of the ubiquitinated protein. Substrate proteins can be modified by a single ubiquitin (monoubiquitination), at multiple sites by single ubiquitins (monomoubiquitination), or by polyubiquitin chains. Polyubiquitin chains can be grown by successive addition of ubiquitin to one of the seven lysine residues of ubiquitin (K6, K11, K27, K29, K33, K48, and K63) or to the N-terminal methionine (M1). Depending on the combination of which of these residues are modified at each round of ubiquitin addition during chain formation, the signal (“ubiquitin code”) on the substrate protein can be interpreted differently in the cell (33, 34).

Whereas K48- and K11-linked polyubiquitin chains usually serve as signals for proteasomal degradation (16, 35), M1- and K63-linked chains may signal other processes. Monoubiquitination, in contrast to K48-linked polyubiquitination, targets transmembrane proteins at the plasma membrane for internalization and delivery to lysosomes, rather than to the 26S proteasome, for degradation. For instance, monoubiquitination of Notch1 by the RING domain–containing individual E3 ligase casitas B-lineage lymphoma (CBL) targets Notch1 to the lysosome for degradation (36). Ubiquitination may also serve completely different functions other than as signals for protein degradation (32, 37–39). For instance, monoubiquitination may serve as a signal for intracellular protein transport in the secretory and endocytic pathways (40). Moreover, K63-linked polyubiquitination of proteins contributes to DNA repair (41), regulation of protein translation (42), and regulation of cell signaling through proteasome-independent activation of protein kinases, such as IκB kinase (43). Other sites of polyubiquitin chain formation have also been associated with various nondegradation-related cellular processes such as DNA repair (K6, K27, and K33), nuclear factor κB (NF-κB) signaling (M1), and post-Golgi trafficking of proteins (K33) (44).

**Definition of degrons and the structural basis of degron recognition**
Fine-tuning the abundance of intracellular proteins usually involves E3 ligases that specifically recognize localized sequence elements, called degrons, in their substrates. An important property of degrons
is that they are transferable, and the attachment of such sequence elements by means of genetic engineering confers instability on otherwise long-lived proteins (45). The degron itself is usually a short linear motif (46, 47) characterized by a specific sequence pattern. The key amino acid residues that are critical for the interaction show strong evolutionary conservation. The largest collection of linear motifs is collected in the Eukaryotic Linear Motif (ELM) database (http://elm.eu.org), which currently contains 15 different classes of degrons with 93 experimentally verified instances in 83 unique protein sequences in human (48). We extended the list with additional examples collected from the literature (table S1; http://dosztanyi.web.elte.hu/CANCER/DEGRON/TP.html).

A well-characterized example of an E3 ligase–degron pair is the degron in p53 and the E3 ligase MDM2 (murine double minute 2), which is a RING domain–containing individual E3 ligase (49). In the absence of DNA damage or other stress signals, MDM2 targets the constantly produced p53 for degradation. The structure formed between MDM2 and p53 (Fig. 3A) shows that a short segment on the N-terminal region of p53, corresponding to the degron motif, forms an α-helical stretch that binds to the SWIB domain of MDM2 (50). The three hydrophobic amino acids of p53 (Phe19, Trp23, and Leu26) form critical contacts with MDM2 and occupy a relatively deep hydrophobic cleft on the surface of MDM2 (Fig. 3A). Other substrates of MDM2 and the structurally similar MDM4 exhibit a similar method of degron binding and share a common motif (Table 1). This sequence motif highlights three key hydrophobic amino acids (a Phe, a Trp, and a Val, Ile, or Leu) that must be spaced such that they are located on the same side of a helix and incorporates the incompatibility of proline residues as part of this recognition motif.

In many cases, the interaction with an E3 ligase requires post-translational modification of the degron motif. For example, a phosphodegron motif is a degron that contains a phosphorylation site, and phosphorylation of this site enables the phosphorylated degron (phosphodegron) to interact with the E3 ligase. The cell cycle inhibitor p27KIP1 (also known as CDKN1B) contains a phosphodegron (51). Substrate recognition of this protein requires SKP2, a receptor subunit of the SCF family of E3 ligases, and the accessory protein CKS1 (CDK regulatory subunit 1). The degron is characterized by its corresponding motif (Table 1) centered around a phosphorylated threonine that binds into a positively charged pocket on CKS1 with high specificity (Fig. 3B) (52). Several other members of the SCF family, such as SCFβ-TrCP and SCFPHBW7, recognize doubly phosphorylated motifs through a WD40 β-propeller structure (Fig. 3C) (53). β-TrCP recognizes a degron motif in its targets with two phosphorylated residues (54), and FBXW7 (F-box and WD repeat domain containing 7) targets contain a similar yet distinct motif with either two phosphorylated residues or a glutamic acid substituting for the second phosphorylated residue (55).

The COP1 E3 ligase is an individual RING domain E3 ligase and also uses the conserved surface of its WD40 domain to interact with its degron, which is usually unmodified (Table 1), such as in the solved structure of COP1 with the TRB1 (Tribbles homolog 1) degron (Fig. 3D). The CBL E3 ligase targets degron motifs containing a phosphotyrosine residue (56). Substrate recognition by the ECV complex, which contains the receptor subunit VHL, critically depends on another posttranslational modification, prolyl hydroxylation (57). The substrates of VHL include hypoxia-inducible factor 1α (HIF-1α), which regulates gene expression in response to oxygen availability (58). Whereas HIF-1α is the most important cancer-linked target of VHL, other targets (such as hsbPB7, β2-adrenergic receptor, or SPRY2) also depend on hydroxylation to exhibit a functional degron. Other posttranslational modifications, such as methylation, can promote the interaction with E3 ligases to mediate degradation, and methylation-dependent degradation is especially important for proteins involved in DNA repair (59, 60).

The complexity of degron regulation

Various mechanisms control degron-dependent degradation. One key regulatory element is the binding strength of the interaction between the E3 ligase

Fig. 3. E3 ligase–degron complexes. (A) p53–MDM2, the p53 degron peptide enters a deep hydrophobic pocket on MDM2. The three key hydrophobic residues from p53 peptide are shown [Protein Data Bank (PDB): 1YCR]. (B) Structure of SKP1-SKP2-CKS1 in complex with the p27KIP1 phosphodegron peptide. SKP2 is rendered in gray surface, whereas CKS1 is represented in dark gray surface, with red (oxygen) and blue (nitrogen) polar functional groups. The small peptide from the p27KIP1 is rendered in ribbon representation with purple color: Its phosphothreonine interacts with positively charged CKS1 surface (blue region), which provides the phosphospecificity. SKP1 is not shown for clarity (PDB: 2AST). (C) Doubly phosphorylated β-catenin degron motif (see Table 1) in complex with β-TrCP1 and SKP1. Both of these molecules are shown in gray surface representation, whereas the β-catenin fragment is rendered as purple ribbon (PDB: 1P22). (D) COP1 E3 ligase with TRB1 degron. TRB1 binds on the conserved surface of COP1 in an extended manner. COP1 with WD40 repeats (β-propellers) is rendered as ribbon, whereas TRB1 peptide is rendered in stick representation and colored purple. (PDB: 5IGQ). Interchain H bonds are represented in magenta. Motif-defining positions have been rendered as yellow sticks for all the cases.
and its substrate. The small interface formed between the protein segments that contain the degron motif and their partner E3 ligase usually enables only a low affinity for the interaction. However, this interface is not the only interaction that occurs between some E3 ligases and their substrates. For example, ERG [erythroblast transformation–specific (ETS)–related gene] is recognized and ubiquitinated by the CRL3 E3 ligase with the receptor subunit Speckle-type POZ protein (SPOP). The binding region of ERG for SPOP

Table 1. Degron motifs and cancer-associated dysfunction.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Protein name</th>
<th>E3 ligase or subunit-recognizing degron</th>
<th>Instance</th>
<th>Starting position</th>
<th>Motif*</th>
<th>Source†</th>
<th>Associated mechanism in cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP53</td>
<td>p53</td>
<td>MDM2</td>
<td>FSDLWKLL</td>
<td>19</td>
<td>F[^P][W][^P][2,3][VIL]</td>
<td>DEG_MDM2_1</td>
<td>Increased degradation by increased E3 ligase activity</td>
</tr>
<tr>
<td>CDKN1B</td>
<td>p230KIP1</td>
<td>SKP2-CKS1</td>
<td>SVEQTPKK</td>
<td>183</td>
<td>[DE],(T),P,K</td>
<td>DEG_SCF_SKP2-CKS1_1</td>
<td>Increased degradation by increased E3 ligase activity</td>
</tr>
<tr>
<td>CTNNB1</td>
<td>β-catenin</td>
<td>β-TrCP1</td>
<td>DSGHS</td>
<td>32</td>
<td>D(S,G)(2,3)(I,T)</td>
<td>DEG_SCF_TRCP1_1</td>
<td>Missense mutations in degron</td>
</tr>
<tr>
<td>MYC</td>
<td>MYC</td>
<td>FBXW7</td>
<td>LLPTPLS</td>
<td>55</td>
<td><a href="0,2">LIVMP</a>(T),P..(ST)</td>
<td>DEG_SCF_FBW7_1</td>
<td>Translocation, mutation of the posttranslational modification site needed for E3 ligase binding, increased gene expression, and stabilization by MCV-mediated inhibition of FBXW7</td>
</tr>
<tr>
<td>ERG</td>
<td>ERG</td>
<td>SPOP</td>
<td>ASSSS</td>
<td>42</td>
<td>[AVP],(ST)(ST)(ST)</td>
<td>(61)</td>
<td>Gene deletion by chromosomal translocation</td>
</tr>
<tr>
<td>HIF1A</td>
<td>HIF-1α</td>
<td>VHL</td>
<td>LAPAAGDTIILDF</td>
<td>400</td>
<td><a href="P">ILA</a>,(6,8) [FLV],(FLVM)</td>
<td>DEG_ODPH_VHL_1</td>
<td>Stabilization by E3 ligase inactivation</td>
</tr>
<tr>
<td>NOTCH1</td>
<td>NOTCH1</td>
<td>FBXW7</td>
<td>PFLTSPES</td>
<td>2508</td>
<td><a href="0,2">LVMP</a>(T),P..(ST)</td>
<td>DEG_SCF_FBW7_2</td>
<td>Truncating mutations in C-terminal PEST region</td>
</tr>
<tr>
<td>TP63</td>
<td>p63</td>
<td>ITCH</td>
<td>PPPY</td>
<td>540</td>
<td>PP,Y</td>
<td>(342)</td>
<td>Reduced gene expression</td>
</tr>
<tr>
<td>NFE2L2</td>
<td>NRF2</td>
<td>KEAP1</td>
<td>DEETGE</td>
<td>77</td>
<td>[DNS],[DES],[TNS]</td>
<td>DEG_Kelch_Keap1_1</td>
<td>Missense mutations in both degrons, NRF2 deletion induces tumor formation</td>
</tr>
<tr>
<td>ETV1</td>
<td>ETV1</td>
<td>COP1</td>
<td>DEQFVPDY</td>
<td>67</td>
<td><a href="1,3">STDE</a>,(0,2) <a href="2,3">TSDE</a> VP<a href="0,1">TSTDEG</a> [FLMVPRA]</td>
<td>DEG_COP1</td>
<td>Translocation deleting the N-terminal degron</td>
</tr>
<tr>
<td>CSF1R</td>
<td>CSF-1R</td>
<td>CBL</td>
<td>LLQPNNYQC†</td>
<td>963</td>
<td>[DN],(Y)(ST),P</td>
<td>(157)</td>
<td>Mutations abolishing posttranslational modification site needed for E3 ligase binding</td>
</tr>
<tr>
<td>MET</td>
<td>c-Met (HGFR)</td>
<td>CBL</td>
<td>DYR</td>
<td>1002</td>
<td>D(Y)R</td>
<td>(157)</td>
<td>Mutations in posttranslational modification site needed for E3 ligase binding/exon skipping/translocation deleting the degron</td>
</tr>
<tr>
<td>SH2B</td>
<td>APS</td>
<td>CBL</td>
<td>RAVENQYSFY</td>
<td>623</td>
<td>RA(VI),NQ(Y)(ST)</td>
<td>(157)</td>
<td>—</td>
</tr>
<tr>
<td>AURKB</td>
<td>Aurora kinase B</td>
<td>CDH1</td>
<td>QKENS</td>
<td>3</td>
<td>.KEN.</td>
<td>DEG_APCC_KENBOX_2</td>
<td>—</td>
</tr>
<tr>
<td>CCNA1</td>
<td>Cyclin A1</td>
<td>CDC20/CDH1</td>
<td>FDIYMD</td>
<td>135</td>
<td>[ILVMP],[ILMPV]</td>
<td>LIG_APCC_ABB8_1</td>
<td>—</td>
</tr>
<tr>
<td>CCNB1</td>
<td>Cyclin B1</td>
<td>CDC20/CDH1</td>
<td>PRTALGDI</td>
<td>41</td>
<td>.R..L,[ILVM]</td>
<td>DEG_APCC_DBOX_1</td>
<td>—</td>
</tr>
<tr>
<td>PTTG1</td>
<td>Securin</td>
<td>CDC20</td>
<td>DKENG</td>
<td>8</td>
<td>.KEN.</td>
<td>DEG_APCC_KENBOX_2</td>
<td>—</td>
</tr>
</tbody>
</table>

*Definition of regular expressions describing linear motif sequence conservation: [ILV], [·] refers to group of amino acids, in this case, Leu, Ile, and Val are all allowed; [^P], anything but Pro is allowed; a period (·) denotes any kind of residue; (T), modified residue (that is, phosphorylated Thr); (0,2), variable length position, in this case, 0, 1, and 2 positions of any kind of amino acid. †Known motifs are represented as standard regular expressions and were taken from the ELM database (48) or from the references cited. ‡Does not meet standard motif definition.
extends N-terminal of the phosphodegron motif, which would likely confer stronger binding (61, 62). Several E3 ligases recognize their substrates through a multisite substrate recognition mechanism. An example of multisite recognition is presented by the interaction between the substrate NRF2 (nuclear factor erythroid 2–related factor 2) and KEAP1 (Kelch-like erythroid cell–derived protein with CNC homology–associated protein 2), a different receptor subunit of the CRL3 E3 ligase (63). NRF2 contains two degrons in close proximity, which act in synergy, mediating binding to a KEAP1 dimer. Although both degrons bind to Kelch domains on respective subunits of the partner KEAP1 dimer, their affinities are different. The N-terminal DLG degron motif has a 200-fold lower affinity for KEAP1 than the high-affinity C-terminal ETGE degron motif (64). The finely tuned affinity of these two sites is key to the stabilization and activation of NRF2 in response to electrophilic stress. Cyclin E also contains two degrons, which are recognized by the SCF family E3 ligase containing the FBXW7 receptor subunit. FBXW7 can exist both as a monomer and as a dimer in normal cells. Like the pair of NRF2 degrons, the ones in cyclin E also have very different affinities, with the N-terminal one forming a much weaker interaction. The two degrons in cyclin E are involved in cooperative binding that is coupled to the dimerization of FBXW7 (65). As with NRF2, the multisite degrons enable tight control for cyclin E turnover (66).

Cooperative interactions likely mediate E3 ligase–mediated degradation of members of the Gli family of transcription factors. Gli2 and Gli3 are critical players in the transcriptional regulation by the Sonic hedgehog pathway (67, 68) and have multiple degron motifs. The degron motifs in these proteins are phosphodegrons that are rich in serine and threonine residues that, when phosphorylated, are recognized by the SPOP-containing CRL3 E3 ligase. Mutation of a subset of these motifs generates resistance to SPOP-mediated protein degradation in Gli2, which suggests that their binding with SPOP happens in a cooperative fashion and uses multiple sites for interaction (69). In contrast to Gli2 and Gli3, Gli1 has fewer Ser- and Thr-rich patches, has no canonical SPOP-recognized degron, and is poorly targeted for degradation by SPOP (69). Instead, Gli1 contains two degrons that are recognized by β-TrCP, a receptor subunit of the E3 ligase CRL1 (also known as SCF). Gli1 lacking both of these degrons is resistant to β-TrCP–dependent degradation (70). β-TrCP–mediated degradation is not the only E3 ligase–dependent degradation mechanism for Gli1. Gli1 has two PPxY motifs and a C-terminal pSP motif (71), which mediate the ubiquitination by HECT-type E3 ligase ITCH in a NUMB (adaptor protein for ITCH)–associated manner, and this mechanism controls Gli1 stability in the nucleus (72).

Regulation of degron recognition often involves molecular switches that can turn on or turn off interactions in response to environmental and cellular cues (73). Posttranslational modifications, including phosphorylation, represent the most common mechanism for either promoting or inhibiting substrate recognition by E3 ligases. Substrate phosphorylation can control E3 ligase–mediated degradation. For example, phosphorylation of Thr18 of p53 blocks its interactions with MDM2, resulting in stabilization of p53 (Fig. 4) (74). Other forms of posttranslational modifications, including acetylation, sumoylation, or glycosylation (59), can also regulate substrate recognition by E3 ligases.

Combinations of posttranslational modifications enable fine-grained control of protein abundance in response to various signals. For example, the formation of phosphorydegrons often involves “priming” phosphorylation events that are necessary before an adjacent residue in the degron can be phosphorylated, and these priming events are often carried out by kinases that are different from the ones that target the residue in the phosphodegron. Such a scenario occurs for the β-TrCP–mediated degradation of β-catenin, for which casein kinase 1 (CK1) is the priming kinase and glycogen synthase kinase 3β (GSK-3β) is the secondary kinase, with phosphorylation of both sites required for the recognition of the phosphodegron (75). In contrast, the binding of the individual E3 ligase CBL to the phosphorylated tyrosine in the degron of epidermal growth factor receptor (EGFR) is blocked by the phosphorylation of neighboring serine residues (76). The HECT-type E3 ubiquitin ligase ITCH (the homolog of mouse Itchy) recognizes tandem PPxY motifs. For the substrate thioredoxin-interacting protein (TXNIP), phosphorylation of the tyrosine residue in the motif abolishes binding to ITCH, instead promoting the interaction of TXNIP with a p-Tyr–binding SH2 domain–containing protein (77). The FBXW7–recognized degron of JUN (Fig. 4) is primed at Ser43 (possibly by DYRK2) for the GSK-3β–activating phosphorylation of Thr239 (78). Conversely, the intrinsically active ITCH degron in JUN (Fig. 4) can be inhibited by ABL phosphorylation of Tyr170 in the PPxY motif (79). For polo-like kinase 4 (PLK4), the biological activity of this enzyme is coupled to its own destruction. PLK4 controls the number of centrioles in the cell. After completion of PLK4–mediated centriole biogenesis (80), these kinases trans-autophosphorylate themselves in a region located in the vicinity of the kinase domain. This creates a phosphodegron that is recognized by β-TrCP, leading to the degradation of PLK4 (81).

Subcellular localization and alternative splicing provide additional mechanisms to regulate protein recognition by E3 ligases and control protein stability. For p27Kip1, degradation is controlled by phosphorylation-dependent changes in subcellular localization. Unphosphorylated p27Kip1 resides in the nucleus and is bound to cyclin E–CDK2 or cyclin D–CDK4 complexes, halting cell cycle progression. Phosphorylation of Ser10 triggers the nuclear export of p27Kip1 to the cytoplasm where it is recognized and ubiquitinated by cytoplasmic CRL3 E3 ligases with the receptor subunit SKP2 and its accessory protein CKS1 (Fig. 4) (82). As exemplified by the transcription factors p53 and HIF-2α (83, 84), alternative splicing can also remove degrons and produce isoforms that have similar domain structure and functionality but differ in stability. Therefore, the relative expression of these different isoforms can also influence the half-lives of the encoded proteins. Chromosomal deletion or fusion events can also influence stability of the encoded proteins, as exemplified by ETV1, which can either undergo a truncation event that removes the degron in the N-terminal region or undergo a chromosomal translocation event leading to the replacement of the N-terminal part of the protein with transmembrane protease, serine 2 (TMPRSS2) (Fig. 4). Either of these events increases the stability of the resulting protein by elimination of the degron recognized by the individual E3 ligase COP1.

Molecular switches can also regulate the E3 ligase by altering its assembly, activity, or subcellular localization. E3 ligases can also be controlled by UPS-mediated degradation, for example, by autoubiquitination, as observed for CBL (85). Another layer of regulatory complexity exists in which one E3 ligase targets another E3 ligase for destruction. TRUSS (tumor necrosis factor receptor–associated ubiquitous scaffolding and signaling protein), a subunit of the CRL4 E3 ligase complex, is one such example. TRUSS contains a phosphodegron that is phosphorylated by GSK-3β, creating a binding site for SKP2 and targeting the CRL1 E3 ligase to tag TRUSS for proteasomal degradation. Thus, the abundance of TRUSS inversely correlates with the abundance of SKP2 (86). Similarly, SKP2 is a substrate for APC/C (87).
proteins, like p27KIP1, are completely disordered, others have a mod-
states instead of a single well-defined conformation. Although some
proteins and protein regions relies on highly flexible conformational
of the UPS involves intrinsic protein disorder. The function of many
protein sequence (Fig. 4). However, within the human proteome, degrons
tend to be less common in the middle of a protein
sequence (Fig. 5).

Many individual proteins are regulated
by multiple E3 ligases that are activated
under different conditions. A well-studied
example is JUN, which contains degrons
for at least three different E3 ligases (Fig.
4). Among these, the binding of COP1
does not depend on phosphorylations.
However, the interaction with the CRL1
receptor subunit FBXW7 requires double
phosphorylation of JUN on Thr239 and
Ser243. The activity of ITCH, a HECT
domain–containing E3 ligase, depends
on phosphorylation by the kinase JUN
N-terminal kinase (JNK) (90). This creates
a negative feedback loop: JNK both acti-
vates and enables degradation of JUN
(91). In contrast, ABL-mediated phos- 
phorylation of the JUN degron recognized
by ITCH protects JUN from interacting 
with ITCH and thus stabilizes JUN (79).
Despite knowing that multiple E3 ligases
regulate some proteins, including MYC,
p53, and NOTCH1, for most cases, the
degr and and E3 ligase binding site are un-
known. Without knowledge of the degron,
it is difficult to establish direct or indi-
rect mechanisms of E3 ligase–dependent
degradation.

Given the large number of E3 ligases
and the combinatorial possibilities to use
them, E3 ligase–degron pairs can poten-
tially provide exquisite specificity for the
UPS. However, substrates and the corre-
sponding degron motifs have been de-
scribed only for a few. The direct targets
of most of the nearly 600 E3 ligases are
unknown. Even for well-characterized
E3 ligases, additional unknown partners
may exist. Moreover, there could be ad-
ditional signals that tag proteins for deg-
radation, such as the N-degron, which is an N-terminal signal that
reduces protein half-life (92). Efficient degradation may also depend
on additional sequence cues in the target proteins, including the presence
of disordered segments. The degradation signal may have three parts
(93): (i) the primary degron that is recognized by the E3 ligase, (ii) sec-
ondary sites of single or multiple lysine residues that can be efficiently
ubiquitinated, and (iii) a structurally disordered segment that initiates
substrate unfolding at the 26S. Although much of the “degrone” remains
to be explored, understanding the details of the controlled degradation
of individual proteins has provided insights into the regulation of cellular
processes, transcription factors, and signaling pathways.

A key structural feature that contributes to the successful operation
of the UPS involves intrinsic protein disorder. The function of many
proteins and protein regions relies on highly flexible conformational
states instead of a single well-defined conformation. Although some
proteins, like p27KIP1, are completely disordered, others have a mod-
ular architecture that involves a combination of ordered globular do-
mains and intrinsically disordered regions. Most known degron
motifs, like linear motifs in general, are located within intrinsically dis-
ordered regions and undergo disorder-to-order transition upon binding
to their partner proteins (88, 89). Disorder can also be important for
the regulation of the degron motifs by ensuring that they are accessible for
posttranslational modifications. The position of the degron motif(s) varies
with some degrons located near the N-terminal end, near the C-terminal end, or in the
middle of the sequence (Fig. 4). However, within the human proteome, degrons
tend to be less common in the middle of a protein
sequence (Fig. 5).
Processes and Pathways Regulated by Degrons

The cell cycle

Degrons can be found in a diverse collection of cell signaling and regulatory pathways. Defects in degron-mediated degradation by the UPS that are particularly important in cancer include the dysfunctional degrons that impair regulation of the cell cycle. The cell cycle consists of four main phases: G1 phase (growth), S phase (DNA replication), G2 phase (more growth and preparation for mitosis), and mitosis (also known as M phase, during which segregation of chromosomes and cytokinesis occur). The cycle is promoted by the activity of cyclins and CDKs and is controlled at many checkpoints by cell cycle regulatory proteins, such as CDK inhibitors (94). The accurate progression of the cell depends on a fine-tuned and tightly controlled complex network of protein interactions, for which the exact amount of each protein is often critical. Therefore, one of the most important mechanisms for the accurate progression of the cell cycle is the ubiquitin-mediated degradation of cell cycle regulatory proteins (27, 95, 96).

Understanding the substrate-specific activity of E3 ligases in the regulation of the cell cycle involves characterization of the degrons that serve as targets for specific E3 ligases. Although exceptions exist (97), most of the well-characterized degrons for the cell cycle are recognized by either the APC/C or the SCF complexes. The substrate receptor subunit of the APC/C complex is either CDC20 or CDH1 (Fig. 2). The APC/C complex can bind to three types of degrons important for the cell cycle: the D-box, the KEN box, and the ABBA motif (Table 1). In contrast, the substrate receptor subunit of the SCF complex can be one of more than 60 different F-box proteins (18). However, so far, only three of these have dedicated degrons that function in the cell cycle: FBXW7, SKP2, and β-TrCP (95). Although there is some overlap between the sets of cell cycle regulatory proteins targeted for degradation by APC/C and SCF complexes, the degrons recognized by APC/C or SCF differ in terms of the cell cycle phases during which they are used to regulate the stability of substrate proteins (Fig. 6). Most APC/

Degron position

1: N terminus; 100: C terminus

Fig. 5. Positional density of degrons in protein sequences. The plot shows the relative density of degron locations in protein sequences. All protein sequences that contain at least one degron (166 proteins from table S1 and http://dosztanyi.web.elte.hu/CANCER/DEGRON/TP.html) were split into 100 equal bins. The bins were numbered from 1 to 100 in the direction of N terminus to C terminus of the protein sequence. The number of degrons overlapping each bin was counted, and a density diagram was generated using ggplot2 geom_density function (360).

C-targeted degrons are critical for the normal mitosis, especially the latter parts of chromosomal segregation and cytokinesis and early G1 phases, whereas most SCF-targeted degrons are critical for G1-S progression, S phase, and early parts of mitosis (27, 96). During the mitosis, APC/C regulates the stability of PLK1 (98), Geminin (99), Cyclin B (100), Securin (101), Bub1 (102), and Cyclin A (103). G1 phase targets of APC/C are SKP2 (87), mE2-C (104), CDC6 (105), CDC20 (106), Nek2A (107), Aurora A (108), and AXIN-2 (109).

With three different receptor subunits that are important for targeting cell cycle regulatory proteins, the SCF complex functions in all phases of the cell cycle (Fig. 6) and controls the stability of both proteins that promote the cell cycle and proteins that inhibit the cell cycle. The F-box proteins FBXW7 and SKP2 have antagonistic effects on cellular proliferation (27). FBXW7 mainly acts as a tumor suppressor because it targets growth-promoting factors, such as MYC (110) and JUN (78), for degradation at late G1 or the G1-S transition. The result of SKP2-mediated targeting of substrates is generally the promotion of the cell cycle. SKP2-targeted degrons occur in tumor suppressors, such as CDK inhibitors p21\(^{kip1}\) (111), p27\(^{kip1}\) (112), and p57\(^{kip2}\) (113). Other SKP2-targeted degrons in proteins relevant to the cell cycle are found in cyclin D (114), ORC1 (115), CDT1 (116), cyclin E (112), E2F1 (117), and p130 (118). Degrons recognized by β-TrCP are found in a diverse set of oncoproteins and tumor suppressors. Most of the characterized β-TrCP degrons function in mitotic onset and mediate the degradation of BORA (119), FBXO31 (120), WEE1A (121), and CDC25B (122). Other cell cycle–relevant β-TrCP degrons occur in proteins distributed throughout the cell cycle: β-catenin (123), CDC25A (124), PLK4 (125), CREB-2 (also known as ATF4) (126), and TPAF4 (127).

The final E3 ligase with a regulatory role in the cell cycle is CRL3. To the best of our knowledge, the only known cell cycle–relevant subunit of CRL3 is p60\(^{Katanin}\), which is recognized by the KLHL42 (Kelch-like protein 42) that forms the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3 and mediate the degradation of BORA (119), FBXO31 (120), WEE1A (121), and CDC25B (122). Other cell cycle–relevant β-TrCP degrons are found in proteins distributed throughout the cell cycle: β-catenin (123), CDC25A (124), PLK4 (125), CREB-2 (also known as ATF4) (126), and TPAF4 (127).

The final E3 ligase with a regulatory role in the cell cycle is CRL3. To the best of our knowledge, the only known cell cycle–relevant subunit of CRL3 is p60\(^{Katanin}\), which is recognized by the KLHL42 (Kelch-like protein 42) that forms the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mitosis, the Kelch domains of KLHL42 form the CRL3 complex with CULLIN3, and binds p60\(^{Katanin}\). During mi...
of p53 from the nucleus to the cytoplasm (quintinate p53, targeting it for degradation, it also promotes the export of p53 (between residues 17 and 29) (Fig. 4). Not only does MDM2 ubiquitinate p53, which recognizes a short degron in the N-terminal part of the gene, but it is present in only low amounts because of constant degradation. This contrasts with the spatially and temporally distributed nature of cell cycle control. The activity of the pathway depends on the concentration of p53. In healthy cells, p53 is constantly produced, but by controlling its accessibility to different E3 ligases (Fig. 4). These ubiquitin ligases is reversible by DUBs. The primary DUB targeting ubiquitin is USP7, which also deubiquitinates and stabilizes MDM2 (137).

Although MDM2 is the main E3 ligase for p53 that mediates its destruction (49), there are several other UPS components that fine-tune p53 abundance under various cellular conditions. The individual RING domain–containing E3 ligases COP1, PIRH2 (p53-induced protein with an RING-H2 domain, and synoviolin; the RING-type E3 ligase MKRN1; the HECT domain–containing E3 ligase ARF-BP1 (HECT); the U-box–type E3 ligase CHIP; and the F-box–containing E3 ligase complexes with JFK are all reported to target p53 for proteasomal degradation, albeit through mechanisms that may be distinct from those used by MDM2 (137). For example, as opposed to MDM2 and COP1, which cannot recognize p53 phosphorylated at the N-terminal region, PIRH2 recognizes N-terminally phosphorylated p53, and thus, PIRH2 is the primary E3 ligase for p53 degradation under conditions of irreparable DNA damage (138). Furthermore, some E3 ligases bind other sites in p53, suggesting that there are other degrons within the central portion of the protein, such as residues 149 to 155, which are recognized by the F-box receptor subunit JFK (139).

Degradation processes in the p53 pathway add additional layers of complexity because several p53-targeting E3 ligases, including MDM2 and COP1, also have autoubiquitinination activity (140, 141). Furthermore, genes encoding the E3 ligases COP1, MDM2, PIRH2, and JFK, all of which target p53, are induced by the transcriptional activity of p53 (138, 139, 142–144). In addition, p53 ubiquitination by these E3 ligases is reversible by DUBs. The primary DUB targeting ubiquitinated p53 is ubiquitin–specific–processing protease 7 (USP7), which stabilizes p53 even in the presence of high amounts of MDM2. However, USP7 also deubiquitinates and stabilizes MDM2 (145, 146). The inherent complexity in p53 degradation is also related to the existence of multiple isoforms (nine isoforms exist in UniProt database for p53; accession no. P04637), some of which lack the MDM2 binding motif in the N-terminal region and thus exhibit prolonged half-lives (147). The subcellular localization of p53 is also regulated by nuclear localization and export signals, which also influence the degradation of p53 by controlling its accessibility to different E3 ligases (Fig. 4). These multilayered and redundant degradation processes ensure the precise control of p53 that is necessary for efficient tumor suppression and maintenance of genome integrity.

Fig. 6. Cell cycle phase diagram showing where degrons function in the cycle. Degron-containing protein names are categorized adjacent to the cell cycle phase during which the degradation of the protein enables progression of the cell cycle.
Wnt/β-catenin pathway

Canonical Wnt signaling regulates various key biological processes during embryogenesis, including those that specify cell fate and contribute to organogenesis (148). A key downstream component of this pathway is β-catenin, which controls gene transcription in response to the glycoproteins of the Wnt family of ligands. In the absence of Wnt, cytosolic β-catenin is degraded by a multisubunit destruction complex. This destruction complex consists of AXIN, adenomatous polyposis coli (APC), the Ser/Thr kinases GSK-3β and CK1, and the protein phosphatase PP2A. The destruction complex marks β-catenin for degradation in a phosphorylation-dependent manner (149). At the destruction complex, CK1 phosphorylates Ser45 of β-catenin, which enables the phosphorylation of Thr37 by GSK-3β. Once Ser45 and Thr37 are phosphorylated, GSK-3β phosphorylates Ser33 and Ser37. These last two sites mark the boundaries of the degron, which, in its phosphorylated state, is recognized by the SCFβ-TRCP E3 ligase. SCFβ-TRCP attaches K48-linked polyubiquitin chains to the Lys19 and Lys49 residues of β-catenin (54), targeting it for proteasomal degradation.

Once the Wnt signaling pathway is activated, the destruction complex translocates to the cytoplasmic side of the transmembrane Frizzled receptors, leading to a series of posttranslational modifications on the adaptor proteins Dishevelled and AXIN. Among these modifications, dephosphorylation of AXIN reduces its binding affinity for β-catenin (150). Moreover, Tankyrases target AXIN for PARsylation, which is critical for the ubiquitination of AXIN. Once PARsylated, AXIN is recognized by E3 ligase RNF146, thereby targeting AXIN for degradation in the proteasomal machinery (151). Consequently, β-catenin accumulates in the cytoplasm, translocates into the nucleus, and recruits cofactors of the TCF/LEF family to regulate transcription of its target genes (152). Many of the activated genes encode proteins connected to cell growth, proliferation, or survival, such as MYC, cyclin D, and JUN (153). Therefore, SCFβ-TRCP-dependent degradation of β-catenin is a critical factor maintaining the “off” state of the canonical Wnt pathway, and RNF146-dependent degradation of AXIN is critical for turning on this pathway.

Receptor tyrosine kinase pathways

Protein ubiquitination plays an important role in regulating the activities of receptor tyrosine kinases (RTKs). Protein tyrosine kinases, both RTKs and non-RTKs, are primary mediators of cell-to-cell communication in multicellular organisms, playing essential roles in most developmental and homeostatic processes, including the regulation of cell growth, proliferation, motility, survival, and differentiation (154). EGFR, colony-stimulating factor-1R (CSF-1R), and Met [the hepatocyte growth factor (HGF), also known as scatter factor (SF) receptor] are members of the RTK family that are particularly important for cancer and exhibit E3 ligase–dependent regulation (155). As single-transmembrane proteins, the RTK transmembrane helix connects the extracellular ligand binding domain to the cytoplasmic region, which contains a conserved tyrosine kinase domain and additional regulatory sequences. RTKs are activated by the binding of growth factors or other polypeptide ligands, which triggers receptor dimerization, oligomerization, or conformational change and activation of the kinase domain. The activated RTK phosphorylates several tyrosine residues on the oligomerized receptors, which form docking sites for cytoplasmic proteins that relay the signal. Simultaneously, RTK activation sets in motion various mechanisms, such as receptor-mediated endocytosis and either receptor degradation or recycling, that will ultimately terminate signaling. One of the major mechanisms of RTK down-regulation involves the E3 ligase family called CBL (156). The CBL family is evolutionarily conserved and contains three members in mammals: CBL, CBL-B, and CBL-C (56, 156). These proteins are characterized by a common N terminus that includes a phosphotyrosine binding (TKB) module connected by a linker helix to a RING finger domain. The TKB module contains an embedded SH2 domain that selectively recruits substrates having phosphorylated tyrosine–containing motifs (Table 1) (157). The RING domain binds to E2 proteins and mediates the transfer of ubiquitin to the substrates. In addition to the conserved N terminus, CBL and CBL-B also have additional C-terminal modules that can mediate multiple protein–protein interactions, as well as an ubiquitin–associated domain, which mediates homodimerization of these two members of the CBL family. This C-terminal region confers adaptor-like functions on CBL and CBL-B. It is ensured at two levels that only activated RTKs are targeted by CBL. First, substrate recognition by CBL requires the phosphorylation of tyrosine residues on the receptors, which is carried out by the activated receptors themselves. Second, the ligase function of CBL only becomes active as a result of the phosphorylation of a conserved tyrosine (Tyr571) in CBL, which is located within the linker helix region connecting the TKB module and the RING domain (158, 159).

The recruitment of CBL proteins to specific phosphotyrosine residues of the activated RTKs leads to ubiquitination of RTKs at the plasma membrane and induces their internalization through the endocytic pathway (160, 161). Through several sorting steps, RTKs can be degraded or recycled back to the cell surface depending on the type and concentration of the ligands. In contrast to the more common mechanisms that target proteins for proteasomal degradation by tagging them with K48-linked ubiquitination, RTKs are associated with a distinct pattern of modifications, featuring multiple monoubiquitination events and K63-linked polyubiquitination and are targeted to the lysosome for degradation (162–164). CBL also has intracellular substrates, such as the kinases ZAP70 and SRC family kinases, which follow the more canonical degradation pathway through the proteasome (165). To fine-tune receptor signaling, there are several mechanisms that regulate CBL-dependent ubiquitination. These include the autoubiquitination of CBLs, the interaction with adaptor molecules that recruit CBL to targets, and proteins that inhibit CBL function. For example, EGFR down-regulation by CBL involves the adaptor protein GRB2 (166), and a similar mechanism may function in CBL-mediated down-regulation of KIT (167). Sprout2 (SPRY2) is a target and an inhibitor of CBL. Activated RTKs induce transcriptional up-regulation and the phosphorylation of SPRY2, thereby creating a binding site for CBL, which targets it for degradation. However, the interaction between SPRY2 and CBL can also attenuate the ubiquitination and endocytosis of some RTKs (168).

Modulation of Protein Degradation in Cancer

Degradation processes can be altered in and contribute to cancer (Fig. 7). In some cancers, the alteration is in the substrate: Specific degron motifs are affected by genetic alterations. The accumulation of these mutations across various types of cancer is consistent with these mutations having a driver role in oncogenesis (169). Altered E3 ligase activity can also contribute to cancer. Furthermore, these basic mechanisms are intertwined in increasingly complex regulatory subystems, presenting multiple mechanisms by which altered function of the UPS affects cancer progression.

Impaired degron functionality by mutation, gene fusion, or truncation

One of the most studied degrons that is modified by cancer mutations is the degron in β-catenin (8, 9). The N-terminal phosphodegron in β-catenin...
SETBP1 contains a phosphodegron that is recognized by SCF
mutations primarily occur in various forms of leukemia and myelo-
tications. Although the exact biological role of SETBP1 is poorly understood,
SET-binding protein (SETBP1), which is another target of cancer muta-
primarily achieved through the inhibition of degradation.

sites. The COSMIC database—a collection of mutations sequenced from
tumor samples (170)—lists more than 5000 missense mutations for β-catenin, 70% of which affect one of the four phosphorylation sites (Ser33, Ser37, Thr41, and Ser45) that are directly or indirectly responsible for SCF
mediated degradation. In addition, most of the identified deletions affect the degron region as well, showing that, for mutations falling within the CTNNB1 gene, the increased β-catenin activity that leads to cancer is primarily achieved through the inhibition of degradation.

A similar mechanism is responsible for the oncogenicity of the nuclear
SET-binding protein (SETBP1), which is another target of cancer mutations. Although the exact biological role of SETBP1 is poorly understood, its involvement in myeloid malignancies is well documented (171). SETBP1 contains a phosphodegron that is recognized by SCF
(172). The degron motif contains almost half of known SETBP1 cancer
mutations despite representing only 0.4% of the protein sequence. Al-
though the same ubiquitin ligase regulates SETBP1 and β-catenin, SETBP1 mutations primarily occur in various forms of leukemia and myelo-
proliferative diseases (172), a more restricted range of cancers than those
that contain β-catenin mutations.

The transcription factor NRF2 lies at the core of the primary pathway
for cellular defense against oxidative stress (173). NRF2 binds to the anti-
oxidant response elements (AREs) in the genome and stimulates the
expression of target genes encoding cytoprotective enzymes, such as
GSTA2 (glutathione S-transferase A2) and NQO1 [NADPH (reduced form of NADP+)] quinone oxidoreductase 1]. In the absence of oxidative
stress, NRF2 is recognized and ubiquitinated by the CRL3 E3 ligase with
the BTB receptor subunit named KEAP1. KEAP1 sequesters NRF2 in the
cytoplasm and ubiquitinates NRF2, targeting it for proteasomal
degradation. NRF2 has two degrons, which bind KEAP1 dimers synergisti-
cally but with different affinities. Despite this asymmetry in the binding
strengths, ubiquitination and subsequent degradation of NRF2 require
binding of KEAP1 to both degrons. Thus, missense point mutations in
either of the two degrons result in excess NRF2 activity. Not surprisingly,
increased NRF2 abundance is present in a wide range of cancers, most
often of the lung, breast, head and neck, ovaries, and the endometrium
(174, 175). The mutations found in NRF2 are spread through the residues
in the degrons without forming apparent mutational “hotspots.” This
pattern contrasts with phosphodegrons, in which the phosphorylated
residues are predominantly mutated, and is consistent with KEAP1 re-
cognition of NRF2 not involving postranslational modifications.

ERG belongs to the ETS transcription factor family and functions as a transcrip-
tional regulator, affecting differentiation, embryonic development, and cell death
(176). In normal cells, this protein is tar-
ged for degradation by CRL3 with the
SPOP receptor subunit, which recognizes
the N-terminal degron motif of ERG. Sim-
ilar to β-catenin and NRF2, disruption of
ERG degradation is connected to cancer,
in particular, prostate cancer (177) and
human myeloid leukemia (178). This de-
gradation process is impaired in a substantial
fraction (~65%) of prostate cancer cases.
However, in contrast to the previous ex-
amples, the ERG degron is not affected by
point mutations. Instead, gene fusion
events, most commonly with the TMPRSS2
protein, disrupt or eliminate the degron.
Transcripts for the TMPRSS2:ERG fusion
protein are associated with androgen-
independent prostate cancer (177, 179),
and this fusion protein is resistant to pro-
tein degradation by CRL3. The two most
common prostate cancer–associated fusions
contain the first exon of TMPRSS2 and
either the first four or five exons of ERG
(TMPS5:ERG exon 1–ERG exon 4 (T1-E4)
and TMPS5:ERG exon 1–T1-E5)). In T1-E4 and T1-E5, the first 39 and 99
residues of ERG are lost, respectively. Al-
though there are multiple putative degrons
in ERG, the N-terminal degron (residues
42 to 46) has critical residues for the SPOP-
ERG interaction. In the case of the T1-E4
fusion protein, the core degron is present,
but the N-terminal flanking region is not. Because the N-terminal flanking region substantially adds to the binding strength, the fusion event leads to an unstable interaction with SPOP and hence impaired ubiquitination of the fusion protein. In the T1-E5 fusion protein, the critical degron is completely lost owing to the absence of the first 99 residues, which makes ERG degradation-resistant. The diminished or abolished degradation eventually leads to the accumulation of these fusion proteins, which drives prostate cancer metastasis (61, 62).

TMPRSS2 also forms genetic fusions through chromosomal translocation with the gene encoding ETV1, ETV4, and ETV5 represent the PEA3 subfamily of the E26 transformation-specific transcription factors (180, 181). The N-terminal regions of the PEA3 subfamily members contain degron motifs recognized by the E3 ligase COP1 (Fig. 4). In addition to the TMPRSS2 fusion, other fusions arising through chromosomal translocations of ETV1 (and, less commonly, ETV4 and ETV5) often produce N-terminal–truncated mutants in human prostate cancers. These degron-lacking ETV proteins evade COP1-mediated degradation (182).

The degradation of NOTCH1 is also often misregulated in cancer, including head and neck squamous cell carcinoma, breast cancer, CLL (chronic lymphocytic leukemia), and lymphoma (183–186), with truncating mutations as the most commonly occurring genetic alterations. NOTCH1 regulates cell fate decisions and differentiation. NOTCH1 is a single-pass transmembrane receptor that is activated by ligands present on the surface of adjacent cells. Ligand binding triggers NOTCH1 cleavage, releasing the Notch intracellular domain (NICD), which translocates to the nucleus and promotes transcriptional programs that stimulate cell division and prevent differentiation. NOTCH1 has been implicated in various cancer types. Multiple ubiquitin ligases, including ITCH, CBL-B, and CBL-C (187), may target NOTCH1. However, the CRL1/SCF E3 ligase with the receptor subunit FBXW7 plays a crucial role in regulating the transcriptional activity of the NICD. FBXW7 recognizes the natively unstructured C-terminal region, which is rich in proline, aspartic acid, serine, and threonine residues (a so-called PEST region). PEST regions are associated with short molecular half-life (188). The NICD PEST region harbors a phosphodegron motif anchored by Thr512, the recognition of which depends on phosphorylation by CDK8 (189). A major mutational hotspot zone in NOTCH1 is in the PEST region, and most of these cancer mutations are frameshift or nonsense truncating mutations that eliminate the degron recognized by FBXW7 and increase NICD stability.

**Degradation impaired through decreased E3 ligase activity**

As the other side of the coin, degradation can also be impaired through various mechanisms that decrease the functional activity of E3 ligases. Although truncating mutations in NOTCH1 frequently eliminate the degron recognized by FBXW7, in T cell acute lymphoblastic leukemia (T-ALL), mutations in FBXW7 are also common. However, cancers with increased NOTCH1 abundance have either the NOTCH1 PEST mutations or the FBXW7 mutations (190–192). This mutual exclusivity is similar to that seen for other substrates and E3 ligases or receptor subunits, such as the substrate ERG and the receptor subunit SPOP. Impaired E3 ligase activity of CBL and TRIM33 has also been reported in various cancers.

Like ERG translocation events, genetic alterations of the SPOP-encoding gene are also connected to prostate cancer (193). Most of SPOP mutations are concentrated in the MATH domain (for example, Y87C and F133V mutations), which is crucial for substrate recognition (194). Recognition of the ERG degron is facilitated by a shallow groove on the MATH domain where the ERG degron motif interacts through an extended conformation. Mutations in this evolutionarily conserved and structurally important region lead to the loss of E3 ligase function toward SPOP-dependent targets (194, 195). As with translocations that eliminate the function of the ERG degron, mutations in SPOP that compromise substrate binding result in increased ERG in prostate cancer cells. There is a mutual exclusivity in the TMPRSS2-E3 fusions (194) and SPOP mutations, consistent with both genetic alterations being able to promote the cancer-associated increase in ERG abundance at an early stage of tumorigenesis (62, 194).

As a phosphotyrosine-directed E3 ligase that controls the degradation of a range of target proteins connected to transducing growth factor signals, many of which are oncoproteins, CBL is also linked to various forms of cancer. CBL mutations are frequently observed in acute myeloid leukemia (AML) and other myeloproliferative diseases, as well as in non–small cell lung cancer (196, 197). Surprisingly, the mutation hotspots largely spare the TK8 domain, which is responsible for substrate recognition. Instead, mutations accumulate in the RING domain and the linker helix region, therefore interfering with E2 binding. The highest frequency of missense mutations occurs at Tyr312 within the linker region, which has a critical role in ligase activity by properly positioning and orienting the TK8-bound substrate relative to the catalytic RING domain (158, 159). CBL mutants that bind RTKs but not the E2 may protect the RTKs against degradation, even in the presence of a second functional CBL allele by preventing access to the substrate.

Tripartite motif containing 13 (TRIM13) is a B-box ubiquitin ligase that is involved in the endoplasmic reticulum (ER)–associated degradation, a process that degrades secretory and membrane proteins synthesized, folded, and assembled in the ER (198). TRIM13 also has cytosolic substrates, and these are the substrates relevant to cancer. The oncogenic relevance of TRIM13 comes from its ability to target MDM2 (the main E3 ligase for p53) and AKT1 (a kinase that activates MDM2) for degradation, thus increasing the abundance of p53 and promoting apoptosis. Physical mapping of the regions in chromosome 13 that are frequently deleted in B cell CLL (B-CLL) has pinpointed TRIM13 to be a core tumor suppressor cancer gene in B-CLL development (199). In contrast to E3 ligases SPOP and CBL, TRIM13 functionality is not typically impaired by mutation; instead, the whole gene is lost.

TRIM33 (also known as transcription intermediary factor 1gamma) plays a key role in erythropoiesis (production of red blood cells) by promoting the differentiation of the red cell precursors (200). TRIM33 inactivation has been shown to lead to chromosomal defects, and reduced TRIM33 expression correlated with an increased rate of genomic rearrangements (201). Accordingly, TRIM33 down-regulation is connected to various forms of cancer, most notably chronic myelomonocytic leukemia (202) and hepatocellular carcinoma (203). In ~35% of leukemic cells, TRIM33 expression was undetectable as a consequence of various epigenetic factors, including CpG hypermethylation and specific histone modifications in the gene promoter (202). At the protein level, the effect of epigenetic gene repression is functionally equivalent to that of gene deletion, similarly to TRIM13. However, expression of TRIM33 can be at least partially restored by demethylating agents, representing—in contrast to previously described mechanisms—a reversible modification of a compromised degradation process.

**Increased degradation through up-regulated E3 ligase activity**

In the alterations of degrons and UPS components described so far, the effects of these modulations are stabilization and increased abundance
of the degron-containing proteins. Thus, impaired ubiquitination promotes tumorigenesis, which means that the target protein has an oncogenic role and the E3 ligase has a tumor suppressor role. However, target proteins that are tumor suppressors can also be involved in cancer-promoting alterations of the UPS (Fig. 7). In this case, the net effect of the modulations of the degradation process is a reduction in the abundance of the target protein, meaning that the activity of certain E3 ligases is increased. At the level of degrons, enhanced degradation cannot be achieved by a simple modification of an already existing degron. Although, in theory, a de novo cancer-promoting degron could be formed by a simple mutation or a fusion event, no such scenario has been documented. However, a gene fusion event could introduce additional lysines that could be ubiquitinated, leading to enhanced degradation, an event that could contribute to cancer (204).

The dominant oncogenic mechanism that leads to the increased degradation of tumor suppressor proteins involves an increase in the abundance of their cognate E3 ligase. A well-known example is MDM2, the primary E3 ligase for p53, which is increased in activity through a range of different molecular mechanisms in various types of cancer. First, transcription of MDM2 can be increased by several mechanisms. A single-nucleotide polymorphism in the promoter of MDM2 increases the binding affinity of the transcription activator SP1, resulting in an increase in the mRNA and of the translated protein (205). This increase in MDM2 attenuates activity of the p53 pathway both in vivo and in vitro and is associated with accelerated development of three different human cancers (205, 206). Gene amplification of MDM2 also increases the activity of this E3 ligase and is common in osteosarcomas, esophageal carcinomas, and soft tissue tumors (207). Incidentally, gene amplification in MDM4, encoding a paralog of MDM2 that also targets p53 for degradation, is also common in retinoblastoma (Rb) (4).

MDM2 abundance and activity can also be increased without direct mutations or changes to the MDM2 gene. Increased transcription of MDM2 is due to increased growth factor signaling by transforming growth factor–β1 (TGF-β1), which activates the heterodimeric transcription factor complex SMAD3/4. Many late-stage carcinomas exhibit excess SMAD3/4 activation, which is associated with high MDM2 abundance (208). Increased MDM2, but without an increase in its transcript, contributes to the development or the severity of several types of cancer (209). The abundance of p53 is indirectly enhanced by p14ARF, which sequesters MDM2 in the nucleolus. Expression of the gene encoding p14ARF is stimulated by the RB/E2F pathway, a pathway that is frequently down-regulated in cancer. Without this negative regulation, MDM2 activity increases and p53 abundance decreases. In general, cancers with increased MDM2 activity have normal p53, indicating that MDM2 up-regulation is sufficient for tumorigenesis in the absence of genetic aberration of TP53 (210).

Similar observations linking cancer to overexpression at the transcript level exist for other E3 ligases that function within the p53 pathway, including COP1, PHR2 (also known as RCHY1), ARF-BP1, and several others (211), as well as for E3 ligases that target other tumor suppressors for degradation (212). The CDK inhibitor p27KIP1 represents an interesting example because at least three E3 ligases control its degradation. The SCF^SKP2^ E3 ligase complex, the RING domain–containing individual PHR2, and the dimeric RING domain–containing E3 ligase complex KPC1 (KIP1 ubiquitination-promoting complex 1) act sequentially, depending on the phase of the cell cycle and cellular location. The E3 ligase with the SKP2 receptor subunit and the accessory subunit CKS1 recognizes the phosphodegron in the C terminus of p27KIP1 (Fig. 4) and targets p27KIP1 for degradation during the S and G2 phases in the nucleus. Although p27KIP1 is seldom mutated in cancer, the abundance of p27KIP1 is frequently reduced because of an increased abundance of SKP2 or CKS1. Increased abundance of PHR2, which is mainly active from late G1 to S phase, may also trigger enhanced degradation of p27KIP1 in malignant cancers (213). In contrast, KPC1, which targets p27KIP1 exported from the nucleus by the nuclear exporter Chromosomal Maintenance 1 protein in G2–G1 transition phases, has not been implicated in cancer. The importance of proper regulation of p27KIP1 in cancer is evidenced by decreased abundance of p27KIP1 serving as a prognostic factor predictive of a poor prognosis in many malignancies (214). There are many E3 ligases that are overexpressed, leading to increased transcript and protein abundance in various cancers (table S2; http://dosztanyi.web.elte.hu/CANCER/DEGRONT/E3_TRS.html). However, for many of these E3 ligases, the targets and the exact mechanisms by which they contribute to cancer have not been explored.

**Disruption of one degron: β-catenin/β-TrCP**

Although we have presented relatively simple examples of cancer-associated changes in degrons and E3 ligase activities, the affected proteins participate in complex and intertwined networks within cells. Consequently, the effects of the degradation—and in many cases the degradation process itself—of certain proteins are interconnected and cannot be separated from one another. We present a set of increasingly complex scenarios showing how cancer and protein degradation are linked in the context of regulatory subnetworks.

The properties of the canonical Wnt/β-catenin pathway enable cancer to arise from simple disruption of a single degron: (i) β-catenin target genes include oncopogenes (215); (ii) β-catenin is constitutively produced and degraded in the absence of Wnt signaling through a mechanism dependent on a single degron (216); (iii) β-catenin appears to be the main or even sole transducer of the Wnt signal into the nucleus. Thus, the pathway lacks redundancies or feedback loops that could overcome degron-disrupting mutations. Degradation of β-catenin requires not only a degron motif but also a functional ubiquitin ligase—in this case, SCF with the receptor subunit β-TrCP. Theoretically, cancer could depend on impaired β-catenin degradation due to disrupted β-TrCP function. However, mutations of β-TrCP are relatively rare events. We predict that the reason for this asymmetry between degron and ligase modulation in cancer is that β-TrCP has a wide range of targets, some with oncogenic and some with tumor suppressor functions. Thus, the effect of β-TrCP inactivation may not be tumorigenic. However, some of its oncogenic targets exhibit degron-disrupting mutations similar to those observed for β-catenin.

**Disruption of synergistic degrons and the corresponding E3 ligase: NRF2/KEAP1**

Although the NRF2/KEAP1 system is similar to the β-catenin/β-TrCP system [constitutive production and degradation of the substrate with signal-mediated stabilization of the substrate and its translocation to the nucleus (217)], a key difference is that the CRL3 E3 ligase with the KEAP1 receptor subunit has few substrates (218). Like β-catenin, the degron mutations of NRF2 promote resistance to degradation, resulting in constitutively active NRF2. Because the primary role of KEAP1 is to limit NRF2 abundance, KEAP1 mutations that disrupt the interaction with NRF2 or epigenetic changes that suppress the expression of the gene are common in cancers (219, 220). The mutational pattern of KEAP1 is broadly distributed with many missense mutations.
scattered along the substrate-binding Kelch domain. Whereas monoallellic mutations in NRF2 are enough to increase NRF2 abundance, usually both alleles of KEAP1 are mutated in cancer (220, 221) because even one functioning copy of the KEAP1 gene can produce the functional E3 ligase. Mutational binding studies showed that KEAP1 mutations that either attenuate or enhance the interaction with NRF2 both decreased NRF2 degradation (222), showing that the interaction is optimized in terms of affinity and thus mutations that alter the affinity in either direction could contribute to tumorigenesis.

**Disruption of multiple E3 ligase components for one target: Cyclin D1/SCF E3 ligases**

In contrast to the degradation of β-catenin or NRF2, the machinery involved in regulating cyclin D1 abundance is complex and intertwined, involving multiple E3 ligases that act at different cellular locations and different phases of the cell cycle. Cyclin D1 is a promoter of the G1–S transition in cell cycle and contains a phosphodegron recognized by SCF E3 ligase with the FBXO4 or the CRL7 E3 ligase with the FBXW8 receptor subunits (223–225). Cyclin D1 functions in the G1 phase when bound to CDK4. In response to mitogenic signals, the cyclin D1–CDK4 complex phosphorylates and inhibits RB, releasing transcription factors of the E2F family and thus enabling gene expression that aids progression through the G1 phase. In its active state during G1 and early S phases, cyclin D1 is in the nucleus. However, at late S phase, cyclin D1 is phosphorylated at the C-terminal Thr286 residue, translocates to the cytoplasm, and is targeted for degradation by CRL complexes SCF<sub>FBXO4</sub> and CRL7<sub>FBXW8</sub> (223, 224). Under normal circumstances, cyclin D1 has a short half-life of less than 30 min (114), which enables the reduction in abundance that occurs at the end of the S phase and is a prerequisite for cell cycle progression. After nuclear export, cytoplasmic cyclin D1 is bound by one of three CRL ubiquitin ligase complexes with different receptor subunits: FBXO4 (226), FBXW8 (224), and FBXO31 (227). The ubiquitination by SCF<sub>FBXO4</sub> and CRL7<sub>FBXW8</sub> is required for normal cell cycle control, with SCF<sub>FBXO4</sub> representing the main pathway. SCF<sub>FBXO4</sub>-mediated cyclin D1 degradation requires the cofactor αB crystallin (223). Degradation of cyclin D1 through CRL7<sub>FBXW8</sub>-mediated ubiquitination is more restricted; FBXW8 and cyclin D1 are not usually present together in most cells. Instead, this E3 ligase complex is activated in response to genotoxic stress and is not required for normal cell cycle control (224). In addition to SCF E3 ligase–mediated degradation, cyclin D1 is also degraded through a phosphorylation-independent process in which an N-terminal degron is recognized by APC/C. Similar to CRL7<sub>FBXW8</sub>-mediated ubiquitination, the APC/C-mediated degradation pathway is activated in response to ionizing radiation and other conditions that cause genotoxic stress (228).

Increased abundance of cyclin D1 occurs in breast and esophageal cancers, but direct gene amplification is responsible for a few of these events, suggesting that disruption of cyclin D1 degradation is the primary mechanism underlying this protein's contribution to tumorigenesis (229). The complex nature of the regulation of cyclin D1 degradation offers multiple target sites for tumorigenic alterations that could promote cyclin D1 stability. Direct degron mutations are relatively rare, although degron mutations are connected to endometrial cancer (230, 231). Instead, the primary mechanism for E3 ligase–degron binding disruption is mutations in the SCF and CRL7 target recognition components. For FBXO4, most mutations affect the dimerization site (232), diminishing the ubiquitin ligase activity, which depends on dimerization for activity. Reduced abundance of αB crystallin, which is associated with breast cancer (229), impairs the interaction between cyclin D1 and FBXO4, thereby stabilizing cyclin D1 and enhancing the progression of cells into S phase.

Cyclin D1 and cyclin D3 are closely related, with nearly identical tissue expression and cellular localization patterns. Furthermore, both exert their main function by inhibiting RB, are linked to mitogen-stimulated activation of the cell cycle, and are controlled through SCF- and CRL7-mediated degradation. However, the receptor subunit that recognizes cyclin D3 is primarily FBXL2 (F-box and leucine-rich repeat protein 2) (233). Despite the similarities between the structure, function, and control of both cyclins, their mutation patterns associated with cancer show striking differences. Instead of mutations that primarily affect the receptor subunit, as occurs in cancers with increased cyclin D1, for cancers associated with increased cyclin D3, the C-terminal degron is often mutated at several sites. In contrast to cyclin D1, the degron in cyclin D3 does not require phosphorylation to be functional, and its recognition by FBXL2 can be disrupted equally well by mutations at many of the degron residues. Furthermore, the cancers associated with these different mutational events are different: Increased cyclin D1, due to mutations or changes in abundance of substrate recognition subunits, is associated with breast (234) and esophageal cancer (235), whereas increased cyclin D3, due to degron mutations, is associated with Burkitt lymphoma (236). This shows how the molecular mechanisms that contribute to tumorigenesis are connected to the precise molecular context of the protein with altered abundance. Seemingly marginal differences, such as differences in substrate recognition components, can shift the balance between the prevalence of tumorigenic mutations affecting degrons or mutations affecting E3 ligase subunits and accessory factors.

**Disruption of one E3 ligase with multiple functionally related targets: RTKs/CBL**

In contrast with E3 ligases with one or a few targets (such as KEAP1 or the SCF and CRL7 target recognition components for cyclin D1), the three members of the CBL E3 ligase family (consisting of CBL, CBL-B, and CBL-C) have several targets with oncogenic potential. Consequently, deregulation of CBL-mediated ubiquitination provides a potential mechanism that can lead to cancer through altered abundance of multiple different oncogenic targets (237). We describe four examples of particular importance in stomach, lung, blood cell, and brain cancers, among others.

Multiple mechanisms of genetic alteration that disrupt CBL recognition of the receptor MET have been observed in several different types of cancer. Some gastric carcinomas are associated with mutations around Tyr<sup>1003</sup> in the juxtamembrane region of MET, which is the site of CBL binding (238). Disruption of the CBL binding site leads to prolonged half-life and increased transforming activity. Exon 14 skipping in the MET gene commonly occurs in lung cancers and also disrupts CBL recognition, thereby stabilizing MET (239, 240). Another oncogenic form of MET that lacks the CBL binding site is TPR (translocated promoter region)–MET, which is generated through a chromosomal translocation that results in combining the dimerization domain of TPR (a scaffolding element of the nuclear pore complex) with the kinase domain of MET. The resulting fusion protein, which is detected in some gastric cancers, is constitutively active and lacks the binding site for CBL and thus is not targeted for degradation (238, 241, 242).

The juxtamembrane region of the receptor KIT (mast/stem cell growth factor receptor) also contains a binding site to which CBL is recruited, either directly or through adaptor proteins. This region is frequently mutated or deleted most frequently in melanomas and gastrointestinal stromal tumors and enhances the transforming activity of...
KIT (237). Mutation in this region impairs not only degradation but also an autoinhibitory function of the receptor (243).

The oncogenic potential of CSF-1R results from the loss or mutation of the CBL binding site (244). The CBL binding site is a phosphodegron at the C-terminal region containing Tyr969, which is autophosphorylated upon activation of the receptor. Because CSF-1R plays an essential role in the regulation of survival, proliferation, and differentiation of macrophage and monocyte precursors, it is not surprising that mutations of Tyr969 are frequently observed in human myelodysplasia and AML (245).

Some glioblastomas have mutations that ablate regions in EGFR that are involved in CBL binding. Although mutation of the degron recognized by CBL is rare, several types of genetic alterations can interfere with CBL-dependent down-regulation of EGFR. Truncated versions of EGFR lacking the binding site for TKB in the EGFR C-terminal region (246) and deletion of exons 2 to 7, which encode parts of the extracellular domain of EGFR, are observed in some glioblastomas. In both cases, the mutant protein does not effectively trigger CBL-dependent down-regulation. For the receptor lacking part of the extracellular domain, although the receptor has less autophosphorylation, which is the mechanism that compromises CBL recognition, the receptor is still tumorigenic despite its decreased activity (246). CBL-mediated destruction of EGFR can also be impaired by increased abundance of other members of the EGFR family. When human epidermal growth factor receptor 2 (HER2) is in excess, as occurs frequently in many cancers, including breast, ovary, prostate, and brain tumors, the increased HER2 shifts the balance between EGFR homodimers and HER2:EGFR heterodimers toward heterodimers, which are not recognized as efficiently by CBL as EGFR homodimers, leading to increased EGFR signaling (237).

Disruption of a single E3 ligase with functionally diverse targets: SCFFBXW7

Although SCF with the FBXW7 receptor subunit is the primary E3 ligase of NOTCH1, this E3 ligase has many other targets. Whereas NOTCH1 and FBXW7 alterations represent two sides of the same tumorigenic story, dysfunction of FBXW7 has far-reaching effect. Other cancer-associated proteins are targeted for degradation by SCF-FBXW7, including the transcription factors MYC and KLF5 (Kruppel-like factor 5). Chromosomal translocation of MYC to one of the immunoglobulin or T cell receptor loci may be one of the primary events in malignant transformation in Burkitt lymphomas (247). However, this event is often accompanied by additional hotspot mutations, including mutations at Thr38 in the phosphodegron of the MYC part of the fusion protein (248, 249). Phosphorylation of this residue is required for FBXW7-dependent ubiquitination and degradation of MYC and the fusion protein (250). Thus, mutation of this residue stabilizes the protein and contributes to tumorigenesis, a similar effect of MYC amplification and overexpression that is frequently observed in other cancer types (251). KLF5 is another transcription factor that is regulated by GSK-3β-mediated phosphorylation of a phosphodegron and FBXW7-mediated degradation. Once again, mutations of the critical residues of the phosphodegron motif were observed in colon cancer samples, as well as other carcinomas, and these mutations increased the KLF5 half-life and transcriptional activity (252, 253).

Large-scale cancer genome studies have revealed that FBXW7 is one of the most commonly mutated cancer genes. In addition to T-ALL, FBXW7 mutations were observed in diverse human cancer types, including colorectal adenocarcinoma, uterine carcinosarcoma, uterine endometrial carcinoma, and bladder carcinoma (190). The oncogenic potential of FBXW7 substrates and frequent allelic loss in human cancers, which is supported by mouse models, indicate that FBXW7 can function as a tumor suppressor in many human cancers (190, 254). However, the mutational spectra also revealed “hotspots” in FBXW7 mutations. A large number of these variations caused amino acid substitutions in three arginine residues that form critical contacts with the phosphodegron motif and are required for high-affinity interactions with the substrates (65, 190, 255). Most tumors with these mutations retain a normal second FBXW7 allele. This is uncommon for most tumor suppressors, for which both alleles typically need to be inactivated or produce dysfunctional protein. The presence of one mutant and one normal FBXW7 allele may suggest a dominant-negative mechanism of action of the mutant protein or may indicate that some residual FBXW7 function is necessary for cell survival. Other mutational hotspots outside the region involved in degron recognition are isoform-specific and likely influence specific subcellular localization (256). FBXW7 function can also be impaired by oncogenic microRNAs or by various mechanisms that reduce FBXW7 expression, for example, through promotor hypermethylation (190). Regulation of FBXW7 function can also be perturbed in cancer by changes that affect its degradation or phosphorylation-dependent heterodimerization of various FBXW7 isoforms. Overall, FBXW7 has a complex cancer-associated pattern of genetic alterations affecting the FBXW7 gene and those encoding other partners or regulators. This pattern depends on the tissue where the cancer originated and the mutational mechanisms. This way tumorigenesis occurs through various molecular events that depend on the exact alterations of FBXW7, presumably through differential effects on specific substrates (190).

Disruption of UPS function by pathogenic interference

In all examples discussed so far, the mechanistic background of tumorigenic processes has been the accumulation of mutational or epigenetic changes to the cellular genome that affect the degradation process. However, these alterations in UPS function and protein stability are often modulated in a way that confers advantage on cancer formation or growth in the absence of genetic or epigenetic alterations. These modulations involve interactions between elements of the UPS and either pathogenic proteins or pathogen-induced alteration of UPS function. Although a connection between bacterial infection and altered UPS function has not yet been reported, many viruses associated with cancer interfere with the UPS. The relevance of this to cancer is evidenced by viruses being collectively responsible for ~16% of worldwide cancer incidence (257).

To date, seven human oncoviruses have been identified: Epstein-Barr virus (EBV), hepatitis B virus (HBV), human T-lymphotropic virus 1 (HTLV I), human papillomaviruses (HPV16 and HPV18), hepatitis C virus (HCV), Kaposi sarcoma-associated herpesvirus (KSHV), and Merkel cell polyomavirus (MCV) (258). All seven of them use various strategies of interfering with certain degradation events of the host cell (258, 259). Similar to other viruses, oncoviruses modulate the host UPS (258) during viral entry and early after viral entry into the cell (260), as well as during other steps of viral life cycle (261, 262). Although the molecular details of the host-pathogen interactions vary among the various viruses, the main human pathways important for cancer that are virally targeted are the p53 pathway (to suppress apoptosis by down-regulating p53), the Rb/E2F pathway (to induce entry into the cell cycle by activating E2F...
proteins), NF-kB pathway (to modulate the immune response), and the Wnt pathway (to enhance proliferation and tissue regeneration through β-catenin activation). Viruses modulate the host cell signaling at these points to overcome cellular senescence and to force the cell to express and activate genes required for progression through the cell cycle so that the virus can exploit the host DNA replication machinery.

Although the main target host pathways are similar for different viruses, the oncoviruses use divergent molecular mechanisms to interfere with these common pathways. Both HPV and EBV interfere with p53 activity by promoting its degradation, but the detailed mechanisms differ. The HPV protein E6 interacts with the HECT-type host E3 ligase E6AP, thereby promoting p53 degradation (263, 264). EBV targets a different element of the host UPS to achieve the same effect. USP7 is a deubiquitinase that stabilizes p53 by removing ubiquitin from p53, counteracting the MDM2-mediated degradation. The EBV nuclear antigen EBNA1 binds USP7, inhibiting its deubiquitinase activity and preventing it from binding to p53 (265), therefore enhancing p53 degradation.

Viruses typically modulate the RB/E2F pathway by disrupting the interaction between RB and the transcription factors of the E2F family; RB maintains E2Fs in an inactive state. Many viruses produce a protein that binds RB, leaving the E2Fs free to drive transcription of genes needed during the cell cycle. The HPV E7 protein binds not only RB but also the CULLIN2 E3 ligase, promoting RB degradation (266).

Like cancer-causing genetic mutations that disrupt the β-catenin degron, viral oncogenesis also targets Wnt signaling through β-catenin. Increased β-catenin stability is achieved by HBV through the viral X protein (Hbx), which activates the kinase Src, which, in turn, suppresses GSK-3β activity (267) and thus prevents β-catenin phosphorylation and subsequent E3 ligase recognition and targeting for degradation. KSHV also targets GSK-3β through the viral latency-associated nuclear antigen, which sequesters GSK-3β in the nucleus (268). HCV also targets β-catenin activity by activating the kinase AKT, which phosphorylates β-catenin on Ser33 and enhances transcriptional activity. Furthermore, patients with liver cancer and underlying HCV infection exhibit a higher rate of mutation of β-catenin. Activation of AKT by the HCV non-structural protein NS5A increased the transcriptional activity of β-catenin and also stabilized the protein through unknown mechanisms (269).

The main oncogenic effect of β-catenin stabilization is the up-regulation of the expression of β-catenin target genes, most notably the gene encoding MYC. MCV enhances MYC abundance through a β-catenin–independent mechanism by impairing its degradation. MCV has two main oncoproteins: the large T antigen (LT) and the small T antigen (st). LT targets the RB/E2F interactions to promote the cell cycle (270). However, this effect is counteracted by the host by targeting LT for degradation by the SCFβTrCP to target IκBα for degradation. SCFβTrCP targets a different element of the host UPS to achieve the same effect. LT also targets for degradation by SCFβTrCP, st also triggers a pronounced increase in MYC stability (271).

In addition to pathways common to most oncoviruses, some viruses alter other aspects of UPS function, producing additional oncogenic consequences. HPV E5 protein blocks the interaction between EGFR and CBL, suppressing EGFR ubiquitination (272). The increased EGFR activity contributes to uncontrolled cell division. Apart from interfering with the host UPS machinery, some viruses also encode proteins with ubiquitin ligase or deubiquitinase function. These include the K3 and K5 E3 ligases of KSHV, which interfere with antigen presentation (273), and the BPLF1 deubiquitinase of EBV, which enhances the production of infectious particles (274). Viruses may also alter APC/C function through as yet undefined mechanisms with unknown physiological consequences (275).

Multiple degron mutations are found in viral and cellular proteins in EBV-induced cancers. Under sustained exposure to pathogen or exposure to environmental carcinogens, EBV-infected cells can become cancerous. EBV is a causative agent for two cancers: Burkitt lymphoma in endemic malaria zones and nasopharyngeal carcinoma (NPC), mainly in southern parts of China (276). In addition, EBV is often a cofactor in other leukemias, such as Hodgkin’s lymphoma. MYC is frequently translocated in Burkitt lymphoma, and the point mutations resulting in T58N and S62P substitutions were identified in the sequences of translocated MYC of cultured Burkitt lymphoma cells (277). These mutated residues are key determinants of the phosphodegron recognized by FBXW7 and contribute to the stability of aberrant MYC in the lymphoma cells (278). The EBV latent membrane protein 1 (LMP1) also has phosphodegron mutations that are important for driving cancer. In this case, a pair of phosphodegrons recognized by β-TrCP act as pseudosubstrates because LMP1 lacks Lys residues that can be ubiquitinated, resulting in sequestration of E3 ligases with the β-TrCP receptor subunit. Both degrons are frequently mutated in NPC, enabling SCFβ-TrCP to target IκBα for degradation, which then stimulates NF-κB proliferative signaling (279). Thus, cancer-driving mutations in degrons are common features of EBV-induced cancers.

Tumorigenesis: Somewhere Between Neutrality and Lethality

The cancer-related changes in protein degradation are between neutral aberrations (a change with little or no phenotypic effect) and lethal aberration (a change that causes death). The effect of the enhancing or impairing the degradation of a gene product depends on the gene’s function and the function of the genes downstream of it. Neither ubiquitin itself nor either of the two known human E1 ubiquitin–activating enzymes are mutated in cancer. This is not surprising because those mutations would impair the whole UPS and are likely lethal aberrations. For similar reasons, with the exception of UB2E1Q1 (E2 ubiquitin–conjugating enzyme Q1), E2 enzymes are generally not implicated in cancer either because their inactivation would affect hundreds or even thousands of proteins. Differential expression of UB2E1Q1 has been linked to cancer, with some colorectal cancer cells showing overexpression in both the transcript and the protein (280). This is likely possible because UB2E1Q1 presumably functions with only two E3 ligases, STUB1 and RNF5 (as assessed by interaction data in the STRING database). In contrast to the limited links between cancer and E2 enzymes, many E3 ligases and their targets are encoded by cancer-driving genes.

Analyzing the link between cancer mutations and various degradation processes uncovers a wide range of possible effects for tumorigenic alterations. In general, germline mutations have a lower impact because these mutations are present in all cells and the organism must be able to adapt to the mutations to develop and survive. Because somatic mutations occur in a limited number of cells, they can cause a greater rewiring or disruption of signaling mechanisms. An example of germline modulation of the degradation of a cancer-driving gene is the gene encoding VHL protein (VHL). VHL is the receptor subunit of an E3 ligase complex that contains Elongin B, Elongin C, and CULLIN2 (named ECV) (Fig. 2). ECV targets a relatively small number of proteins for degradation,
most notably the hypoxia-inducible factor HIF-1α. Patients suffering from the VHL syndrome carry inherited mutations on one allele of VHL and thus have a predisposition for developing benign or cancerous tumors, including hemangioblastoma, clear-cell renal carcinoma (ccRCC), pancreatic tumors, and pheochromocytoma (PCC) (281). However, monoallelic VHL inactivation has a mild effect, and both benign tumorigenesis and cancerous tumorigenesis require the inactivation of the other allele or a secondary mutation in another gene. Biallelic VHL inactivation impairs HIF-1α degradation, enabling HIF-1α to induce angiogenesis, which confers a distinct advantage for tumor growth. However, cancer does not always result from biallelic VHL inactivation and increased HIF-1α activity. Often, additional mutations of other genes are required (282), showing that cancer development connected to VHL inactivation is a multistep process involving the additive effect of mutations that each carry a small advantage toward tumorigenesis (58).

In contrast to germline mutations, somatic mutations only exert their effects in the cells forming the tumor and can have more aggressive phenotypic results. This is the domain of most known cancer genes. Similar to cancer-associated germline mutations, cooperativity among genetic and epigenetic aberrations is a hallmark of cancer-related modulations across genes harboring somatic mutations. However, somatically targeted genes reveal the balance between tumorigenesis and lethality. One such example is EGFR, which is activated by various growth factors (most notably EGF and TGF-α) that induce EGFR dimerization, activation, and downstream signaling that promotes cell proliferation. Accordingly, increased EGFR abundance and activity are linked to various forms of cancer, including squamous cell carcinoma of the lungs, glioblastoma, and head and neck carcinoma. Because the result of increased abundance due to gene amplification or impaired destruction is similar, EGFR degron mutations linked to cancer would come as no surprise; however, they are not common. EGFR is targeted by two isoforms of the E3 ligase CBL through distinct mechanisms (283). Although activating mutations in the intracellular kinase domain of EGFR are very common in cancer, the degrons are almost never mutated, suggesting that the effect of gene amplification and the inhibition of degradation are different for the cell. EGFR is recycled through internalization, and the fate of activated EGFR is decided at the early endosomal stage: Ubiquitination by CBL targets the receptor for lysosomal degradation; otherwise, EGFR recycles to the plasma membrane. EGFR activity resulting from gene amplification is at least partially counteracted by internalization and subsequent degradation. There is no apparent balancing mechanism for impaired degradation. The lack of degron mutations suggest that EGFR overactivation balances on the verge of cancer and lethality, with gene amplification and kinase-activating mutations falling on the viable side and degron mutations on the lethal side.

Oncogenes and tumor suppressors in the UPS

Genes implicated in cancer are usually classified either as (proto) oncogenes or as tumor suppressors. Oncogenes are genes with a potential to promote tumorigenesis such that their (over)activation confers advantages for cancer formation, growth, metastasis, or all three. Tumor suppressors are genes that protect the cell from tumorigenesis, typically by counteracting the functions and activity of oncogenes. Thus, oncogenes are typically overexpressed or their encoded products are overactivated in cancer, whereas either tumor suppressors are less expressed or the encoded products are inactivated. Accordingly, proteins harboring degron mutations have disrupted degradation and are generally expected to be oncogenic, and their corresponding E3 ligases are presumably tumor suppressors. Whereas there are some relatively clear-cut examples of oncogenes and tumor suppressors among E3 ligases and their targets, in certain cases, single proteins cannot be assigned to either category. Although the hundreds of known substrate recognition modules of E3 ligases are highly specific, most of them recognize more than one target protein. The biological effect resulting from the decrease or loss of function of these ligases emerges through the functions of its substrates. VHL can be categorized clearly as a tumor suppressor because it has a very limited number of targets, most of which perform functions that confer advantages to tumorigenesis. Many other ligases have targets that are either oncogenic or tumor suppressors. The two-hit model for tumor suppressors is inadequate, and many other mechanistic models have been proposed (284, 285) that move away from considering separate allelic aberrations as distinct steps toward a continuum model of tumor suppression. Although these models cover a considerably larger spectrum of gene perturbations, they are unable to correctly describe genes with dual oncogenic and tumor suppressor functions. For many elements of the UPS, the classical roles and categories used for cancer genes need to be reassessed.

Defying classification as an oncogene or tumor suppressor

For typical oncogenes, such as EGFR, enhanced activity promotes cancer, but reduced activity does not cause cancer (286) and, to a certain extent, does not threaten viability, making EGFR a typical oncogene. However, this asymmetry in the physiological response to abnormal amounts of a protein is not universal. For many proteins, especially E3 ligases and their targets that control the expression of a large number of genes (usually through transactivation), tipping the delicate scale of normal abundance in either direction can have detrimental consequences, although the change may not result in cancer. For example, by inhibiting several cyclin/CDK systems, cyclin-dependent kinase inhibitor 1C (p57KIP2) inhibits many hallmarks of cancer, including cell invasion, metastasis, tumor differentiation, and angiogenesis. This classifies p57KIP2 as a tumor suppressor (287). Accordingly, the p57KIP2 degron targeted by the SCFSPRO3 E3 ligase (113) is generally not mutated in cancer; rather, expression of the gene encoding p57KIP2 is down-regulated through epigenetic changes, such as DNA methylation and repressive histone marks at the promoter (288). However, mutations in the degron and the corresponding increase in p57KIP2 result in other diseases, collectively called IMAGe syndrome, which is characterized by intrauterine growth restriction, metaphyseal dysplasia, congenital adrenal hypoplasia, and genital anomalies (289). Thus, deviation from the normal p57KIP2 abundance in either direction promotes different disease states.

In certain other components of the UPS, however, any deviation in abundance or function can promote tumorigenesis. Exhibiting either oncogenic or tumor suppressor effects in response to up- or down-regulation, respectively, is characteristic of many UPS components at the E3 level. Examples include COP1 and E3 ligase complexes with SPOP receptor subunits. For example, although the E3 ligase COP1 mainly targets oncogenic substrates, such as JUN and members of the ETV family, COP1 also targets p53 in an MDM2-independent manner (290). Consequently, both the increase and the decrease of COP1 abundance are observed in cancer. The possibility of a clear categorization of COP1 is hindered not only by its mixed targets but also by the complexity of its mechanisms of action as well. Oncogenic targets are recognized by COP1 as part of the CULLIN4-DDB1-RBX1-DET1 E3 complex (291). COP1 mediates p53 ubiquitination as an individual E3 ligase and also acts as an MDM2 cofactor, thereby enhancing p53 degradation (292). To add an additional layer of complexity,
COP1 also promotes its own degradation in response to UV radiation (293). The different substrates (oncogene or tumor suppressor) are also reflected in whether a cancer exhibits an increase (breast, ovarian, liver, and pancreatic cancers) or a decrease (prostate cancers, AML, and melanoma) in COP1 (294, 295).

The CRL3 E3 ligase containing SPOP acts as a tumor suppressor by promoting the degradation of several oncoproteins, including ERG and CDC20 (296), an activator of the APC/C E3 ligase. Consequently, mutations that inactivate SPOP function are present in 15% of prostate cancers (193). Increased SPOP is associated with other forms of cancer, most notably ccRCC. The mechanism by which an increase in SPOP contributes to tumorigenesis is different from the mechanism by which SPOP inactivation leads to cancer (297). The gene encoding SPOP is stimulated by HIF-1α. Thus, hypoxia or mutations that inactivate ECV, the E3 ligase that targets HIF-1α (mutations in the receptor subunit VHL are common in ccRCC), increase the abundance of SPOP. By targeting several tumor suppressors, such as the phosphatase PTEN, ERK phosphatase, and transcription factors such as DAXX and GLI2, the increase in SPOP impairs apoptosis and promotes cell proliferation (297). Thus, E3 ligases with both oncogenic and tumor suppressor targets can promote cancer development when the abundance of the E3 ligase deviates from normal, independently from the direction of the deviation.

The ambiguous nature of many E3 ligases manifests in their activity as oncogenes or tumor suppressors in response to different alterations in their abundance. However, in other cases, the discriminating feature between oncogenic and tumor suppressor behavior for an E3 ligase is not the direction of shift in protein abundance but the cellular context. Genes encoding proteins related to certain biological processes can switch between the two states based on factors other than abundance, such as stage of the cancer or the cell type from which the cancer arises.

A prime example of such a system is NRF2 and its corresponding E3 ligase, which contains the receptor subunit KEAP1. Under normal conditions, NRF2 is kept inactive and targeted for degradation by KEAP1. Exposure or the accumulation of highly oxidative molecules, such as reactive oxygen species (ROS), deactivates KEAP1, inhibiting NRF2 degradation (298). NRF2 induces gene expression through AREs or electrophile response elements, which encode proteins that counteract cellular oxidative stress. NRF2-induced gene products not only detoxify endogenous oxidative agents but also prevent oxidative carcinogenesis from interacting with vital biomolecules, such as DNA, RNA, and proteins, and prevent the accumulation of reactive species that result from cellular metabolism. Accordingly, NRF2-null mice are prone to tumor formation, showing that, in normal and premalignant cells, NRF2 activity has an antitumorigenic effect. The concept of many chemopreventive agents is enhancement of this oxidative stress pathway (299, 300). On the other hand, NRF2 and KEAP1 mutations that keep NRF2 constitutively active are often found in various forms of cancer (175, 301). The reason behind this seemingly paradoxical behavior is that increased resistance to ROS offers a survival advantage for tumor cells and confers resistance to some forms of chemotherapy, decreasing the efficacy of chemotherapy in NRF2–up-regulated cancers (302). A protein’s oncogenic or tumor suppressor status not only can change with the progression of malignancy but also can depend on cell or tissue type. This dependence is evident with NOTCH1, which is up-regulated through impaired degradation in CLL (185). As opposed to this oncogenic role, in certain tissues (and, correspondingly, certain cancer types), NOTCH1 acts as a tumor suppressor by promoting cell cycle arrest through stimulation of expression of the cell cycle inhibitor p21 (303). Although less frequent than the occurrence of degron-eliminating truncation mutations in CLL, tumor suppressor roles of NOTCH1 have been reported in basal cell carcinoma (304, 305), lung squamous cell carcinoma (306), and small cell lung cancer (307).

This type of context dependence is not only characteristic of NOTCH1, but the transcription factor Krüppel-like factor 4 (KLF4) also exhibits cell type–dependent oncogenic or tumor suppressor roles. KLF4 is a key regulator of the cell cycle, apoptosis, differentiation, and stem cell renewal (308). KLF4 is one of few transcription factors that orchestrate the reprogramming of somatic cells into induced pluripotent stem cells (309). KLF4 plays both tumor suppressor and tumor-promoting roles in cancer as a result of its wide range of targets and its function in several key pathways important for maintaining control over cell proliferative potential. Tumor suppressor functions of KLF4 relate to its ability to induce expression of genes encoding the CDK inhibitors p21/WAF1 and p27/KIP1 and to reduce expression of the genes encoding cyclin D1, cyclin B, and FOXM1 (310–312), negatively regulating proliferation in gastric cancer, bladder cancer, lymphoma, and lung carcinomas, among others (313–316). Tumor-promoting activity arises from the ability of KLF4 to repress the expression of the gene encoding p53 (317) and, consequently, reduce expression of p53 target genes encoding proapoptotic proteins, such as BAX (318). The oncogenic activity of KLF4 has been reported in breast cancer and squamous cell carcinoma (317, 319). Because KLF4 and its regulatory factors are distributed unevenly across various cell types, the dominant function of KLF4 is tissue-specific. Consequently, aberrant alterations in KLF4 abundance can either lead to tumorigenesis or tumor suppression, depending on the tissue and the direction in the shift in abundance (320). One of the main mechanisms through which KLF4 abundance is altered in cancer is through the modulation of KLF4 degradation by one of its degrons. KLF4 is targeted by at least three different E3 components of the UPS: VHL, APC/CDH1, and SCIScF–TrCP. Degradation by VHL plays a role in colorectal cancer: Mutations that increase VHL abundance promote KLF4 degradation, indicating that KLF4 functions as a tumor suppressor role in this context (321). This highly specific scenario linked to one cancer type results from not only the tumor suppressor role of KLF4 but also a switch by VHL from its classical tumor suppressor role to an oncogenic role. This oncogenic role of VHL is opposite its classical tumor suppressor function observed in most other cancers associated with altered VHL function, such as renal cell carcinoma (RCC) (322).

The oncogenic role of VHL in certain forms of colorectal cancer is an exceptional deviation from its otherwise universal tumor suppressor function. However, there are other less obvious context-dependent features of VHL mutations. Cancerous manifestations of the VHL syndrome often present as multiple tumors in various tissues, including RCC, retinal hemangioblastoma, pancreatic neuroendocrine tumors, and PCC. On the basis of the presence or absence of these tumor types, VHL diseases are clustered into phenotypic categories (323). The VHL syndrome subtypes reflecting tissues of tumor occurrence correlate with certain genotypic features. For example, most patients with truncating VHL mutations or exon deletions do not develop PCCs (and are classified as type 1), likely because complete VHL inactivation is lethal for PCC cells. In contrast, germline mutations in VHL generally lead to PCC development. These characteristic mutation patterns correlate not only with cancer and tissue type but also with the molecular details that result from the impaired degradation, most notably the balance between the impairment of HIF-dependent and HIF-independent processes.
In general, both oncogenes and tumor suppressors in the UPS can exhibit context dependence at various levels. However, there is another group of proteins linked to cancer development that does not fit into either category. These proteins, which are encoded by genes collectively termed nononcogene addiction (NOA) genes (324), are not tumorigenic themselves but are necessary for sustaining the tumorigenic state. The function of proteins encoded by NOA genes is usually tissue type– and cancer type–specific. A NOA gene encoding a component of the UPS is RNF4, a RING-type E3 ubiquitin ligase. In its canonical function, RNF4 targets polyubiquotylated proteins (including PML, PEA3, CENPI, EPAS1, and PARP1) for degradation, controlling chromosome alignment and spindle assembly (325). This function is not directly related to cancer development, and RNF4 is usually not altered (neither genetically nor epigenetically) in cancer. However, RNF4 can also recognize phosphodegrons in proteins such as β-catenin and MYC and polyubiquitinates these substrates (326). RNF4-mediated polyubiquitination does not target the proteins for degradation; instead, the substrates are stabilized. In contrast to the regular K48-linked ubiquitination, RNF4 catalyzes the formation of K11- and K33-linked polyubiquitin chains, which are not compatible with the UPS degradation pathway (327). Through this mechanism, RNF4 stabilizes a number of oncogenes by translating a transient signal (phosphorylation) into long-term stability, enhancing the tumor phenotype, and in certain tissues, this effect is essential for cancer cell survival (326). Thus, RNF4 represents a distinct class of UPS components that contribute to tumorigenesis solely through the change in their context (in this case, the targeted proteins) without involving a mutation, epigenetic alteration, or change in abundance. This also shows that even the most basic elements of the UPS, such as degrons or polyubiquitination, can have context-dependent meanings and functions in cancer, in this case, changing from the canonical “degradation” role to a detrimental “stabilization” role.

**Implications for Drug Design**

Because the UPS regulates the stability and functions of several cancer-associated proteins, this system provides attractive opportunities for therapeutic interventions. The potential success of this approach in the treatment of cancer was first highlighted by the proteasome inhibitor, bortezomib, which has developed into an important treatment for multiple myeloma and mantle cell lymphoma (328). The proteasome features three dominant proteolytic activities, the trypsin-like, the chymotrypsin-like, and the peptidyl-glutamyl-peptide-cleaving (acidic) activity. Bortezomib inhibits the chymotrypsin-like activity by reversibly binding to the β5 subunit of the 20S proteasome (Fig. 8A). Surprisingly, inhibition of the proteasomal activity showed selective cytotoxicity to certain myeloma and lymphoma cells. This is largely attributed to the increased proteotoxic stress that characterizes tumor cells because of their higher proliferation and rapid growth rates. Many patients can benefit from treatment with bortezomib, but the side effects can be dose- or treatment-limiting (329). Inhibiting the proteasome is hitting the most central and least specific target in the protein degradation machinery; thus, it is perhaps not surprising that treatment is not feasible for most other tumor types, notably solid tumors (329).

Specific inhibitors have also been developed to inhibit the UPS at the level of E1 enzymes. These include PYR-41 and PYZD-4409, the first cell-permeable compounds that are currently in preclinical development stage. Their therapeutic potential originates from their ability to inhibit the catalytic activity of UBE1, the main ubiquitin-activating enzyme in humans (330). However, similar to proteasome inhibitors, these drugs affect many substrates, which could limit their therapeutic value. Neddlyation is a largely similar process to ubiquitination, with NEDD8-activating enzyme (NAE) playing a largely equivalent role to E1 activating enzymes (331). The small molecule, MLN4924, a potent and selective inhibitor of NAE, acts as an indirect inhibitor of Cullin-RING ligases by blocking neddylation that is essential for their activation. MLN4924 was shown to have antitumor activities and is currently undergoing phase 2 clinical trials (332).

Because of their abundance, diversity, and potential for selectivity, the E3 ligase–substrate pairs offer the most targeted approach to therapeutically altering protein abundance at the levels of single or sets of protein substrates. However, current drug design strategies target relatively small, deep pockets, such as the active site of enzymes. This is a practical approach to drug design of enzymatically active proteins because the active site can be readily targeted by small molecules. In contrast, the action of E3 ligases primarily involves protein–protein interactions that, for many years, have been regarded as practically undruggable (333, 334). Results that indicate that protein–protein interactions can be effectively targeted and have shown promise for direct regulation of protein abundance by targeting the activity of substrate recognition of E3 ligases have started to emerge.

A prime example of drugs that directly target an E3 ligase–substrate interaction involves MDM2 and p53. The detailed knowledge of this interface (Fig. 3A), together with high-throughput small-molecule screening, led to the discovery of nutlins, a group of small molecules that can block the MDM2 binding site by mimicking the peptide motif of p53 (Fig. 8B) (335, 336). Nutlins stabilize p53, leading to the activation of cell cycle arrest and apoptosis pathways. Although nutlins and related molecules do not fit into the binding pocket as well as natural ligands, they still were effective competitive inhibitors for p53 (335). Their competitive binding arises because interactions of native degron motifs involve a disorder-to-order transition that entails a large entropic cost, resulting in a relatively weak overall binding (337). This concept is exploited in the design of stapled peptides, in which synthetic α-helical peptides corresponding to the native binding motif are stabilized through covalent linkage (338). Stapled helices not only ensure that the structure and biological activity are maintained but also enable the engineering of lead molecules with optimized pharmacokinetic properties. In addition, stapled helices also bind MDM4, an additional E3 ligase that targets p53, which could be important for treating a broader range of cancers than those currently treated with nutlins (339). These and several other new classes of small inhibitors targeting the MDM2-p53 binding interface or interfering with this interaction in other ways are undergoing clinical trials for the treatment of various solid and hematological cancers (340, 341).

The other two members of the p53 family, p63 and p73, are targeted for degradation by the E3 ligase ITCH (330). Because p63 and p73 share some of the functions of p53, blocking their degradation may be a promising therapeutic strategy, and cyclic peptides that interfere with the binding of ITCH to these targets have been identified (342).

Small molecules or bicyclic peptides that inhibit the activity of members of the HECT family of E3 ligases have also been identified or developed. Rather than interfering with substrate recognition, these inhibitors obstructed the E2 binding site specifically for individual HECT ligases or worked as allosteric inhibitors with a broader specificity for multiple members of the HECT family (343). Another attractive E3 ligase component that could be exploited for cancer therapy is SKP2. Among its substrates, p27KIP1 is the most relevant to the oncogenic function of SKP2.
involves chimeric molecules that combine a degron-mimicking compound to recruit an E3 ligase coupled to another compound to recruit a specific target. These constructs effectively tether the ligase to the substrate, inducing its ubiquitination and degradation. The viability of this concept was demonstrated by combining the degron of IkBα (nuclear factor of κ light polypeptide gene enhancer in B cells inhibitor α) with ovacillin, a drug that covalently binds to methionine aminopeptidase 2 (METAP2). The degron was recognized by β-TrCP, which targeted METAP2 for proteasomal degradation. PROTAC technology has already been successfully applied to various E3 ligases in preclinical experiments (347, 348).

A remarkable story of retargeting an E3 ligase combined with drug repurposing is provided by thalidomide derivatives binding to the protein cereblon (CRBN). CRBN is an unusual receptor subunit of a DCX (CUL4-DDB1-RBX1-CRBN) E3 ligase complex. CRBN has a deep hydrophobic pocket that is the ligand binding site for thalidomide (Fig. 8C), the teratogenic molecule that has been repurposed in lymphoma and leprosy treatments (349, 350). Binding of thalidomide and more recently derived CRBN-binding “immunomodulatory” molecules, such as lenalidomide and pomalidomide, switches the ligand binding preference of CRBN from substrates such as MEIS2 to substrates such as the hematopoietic transcription factors Ikaros and Aiolos, thereby reducing cell division (351). Lenalidomide-bound CRBN also targets the E3 ligase to CK1α, which contributes to the clinical efficacy of this drug in 5q-deletion associated myelodysplastic syndrome [del(5q) MDS] (352). An additional immunomodulatory molecule, CC-885, was shown to have an antitumor activity by inducing the CRBN-dependent degradation of the translation termination factor GSTP1 (353). Structural analysis revealed degrons with a β-hairpin loop in each of GSPT1 and CK1α (Fig. 8D) (352, 353), and degradation of these proteins depends on the binding interface provided by lenalidomide bound to CRBN.

The activity of the CRBN-containing E3 ligase is likely modulated by at least one natural cellular molecule that thalidomide mimics. Uridine binds in the pocket and may be one natural modulator, although to what targets uridine directs CRBN or how uridine affects CRBN function is not known (354). In plants, the concept of endogenous small-molecule regulation of substrate receptor subunits of E3 ligases is well established. The plant hormone auxin induces the degradation of the Aux/IAA family of transcriptional repressors by binding to the F-box component TIR1 of a multisubunit E3 ubiquitin ligase (355). Currently, there are no solved structures for CRBN in complex with a linear motif degron, with or without any small-molecule modulators.
Whether CRBN recognizes both β-hairpin loop degrons and linear motif degrons is unknown but represents an example of the challenge of thinking only in "canonical" terms about substrate recognition by E3 ligases and the potential for the identification of additional types of degron structures in the proteome.

These examples of CRBN-targeted molecules highlight the possibility of identifying or developing small molecules to restore proper degradation by acting as a molecular "glue." Similar approaches have identified small molecules that restore the function of mutant p53 by interacting with its DNA binding domain and stabilizing native folding impaired by point mutations (356). FBXW7, which contains hotspot mutations in its substrate recognition domain in various cancers, is one of the E3 ligase substrate receptor subunits that may be a good candidate for this approach. Specific compounds could stabilize the binding interface with the substrate and reactivate FBXW7 mutants.

The proper functioning of ubiquitin-mediated degradation depends not only on the active E3 ligase but also on the active (recognized) form of the degron. Recognition of degron motifs can be regulated by multiple mechanisms, involving posttranslational modifications or by competing interactions that can enhance or mask the degron motif. These regulatory mechanisms also offer potential strategies for drug design, as demonstrated for n-MYC. n-MYC is a neuronal member of the family of MYC proteins and is specifically amplified in neuroblastoma (357). The degradation of n-MYC involves a phosphodegron motif that is recognized by FBXW7. However, in neuroblastoma cells, Aurora kinase (AURKA) binds to the n-MYC:FBXW7 complex and prevents access of FBXW7 to the degron, thus blocking ubiquitination and stabilizing n-MYC. n-MYC degradation can be restored by small molecules that induce a distortion in the conformation of Aurora A, preventing its interaction with n-MYC (358).

Besides the clinically approved bortezomib and thalidomide derivatives, there are several small molecules targeting various components of the UPS that are under investigation in preclinical studies or are in clinical trials. The most specific effects are expected from targeting those E3 ligases that have the fewest substrates. However, because the sets of targeted proteins vary for different cell types, the effect of E3 ligase–degron inhibitors is also expected to vary for different cancers. This is true for the E3 ligase modulator thalidomide, which has efficacy for multiple myeloma but not for solid tumors (349). Inhibiting the E1–E2–E3 interactions is intermediate in complexity between specific degron blocking and nonspecific inhibition of the proteasome itself. When choosing novel drug targets, the ability to predict the therapeutic effect of any novel inhibitor molecule is extremely limited. Therefore, it makes sense to produce inhibitors for any practical target in the cell and then test them against a range of cell types and disease models for useful effects. The other key use of small-molecule inhibitors is as reagents in experimental research; even inhibitors that are too toxic or otherwise inappropriate for medical use can be very useful research tools. These considerations suggest that all E3 ligase–degron interactions are valid targets for small-molecule inhibition.

**Future Directions**

The relative and absolute amounts of many proteins are critical factors in the establishment and maintenance of cancerous cells. Degrons on oncogenes and tumor suppressors are the targets of E3 ligases that direct the proteins with the degrons for destruction through the UPS. If the degron is not known, it is difficult to evaluate whether posttranslational modifications modulate the E3 ligase–degron interaction and whether there is a direct relationship between an E3 ligase and a putative substrate protein. Cancer-associated mutations compromise degrons in several key oncogenes. By contrast, a lack of mutations in degron sequences in a protein with a proposed role in cancer biology may indicate that the protein is a tumor suppressor. We recommend that scoring degron mutational status should be a routine part of cancer genome screening workflows. The availability of defined E3 ligase–degron complexes enables the screening of compound libraries for inhibitors of these interactions that may have antitumor activity. In summary, knowing the degrons enhances understanding of and therapeutic interference with the destruction of systems operating in different cancers. Yet, compared to the ~600 E3 ligases in the human proteome, only about 25 classes of degron motifs have been defined. Have degrons been neglected? We think so and we believe that putting more effort into degron discovery will benefit biomedical research, providing payoffs in improved cancer treatments.

**SUPPLEMENTARY MATERIALS**

www.sciencesignaling.org/cgi/content/full/10/470/eaak9982/DC1

Table S1. List of known degron motifs.

Table S2. List of E3 target recognition subunits.

Additional information can be found at http://dosztanyi.webelt.hu/CANCER/DEGRON/.

**REFERENCES AND NOTES**


T. Yano, S. Kozuma, Cyclin D1 harboring the T286I mutation promotes oncogenic Oncogene endometrial cancer.

A. Rustgi, S. Y. Fuchs, J. A. Diehl, Phosphorylation-dependent ubiquitination of cyclin D1 with increased risk of adenocarcinoma.

S. B. Baylin, D. Sidransky, E. Gabrielson, M. V. Brock, S. Biswal, Dysfunctional KEAP1-NRF2 tumor suppressor protein.

M. Yamamoto, Oxidative stress sensor Keap1 functions as an adaptor for Cul3-based E3 ligase to regulate proteasomal degradation of Nrf2.

3919 (2008).


T. Yano, S. Kozuma, Cyclin D1 harboring the T286I mutation promotes oncogenic Oncogene endometrial cancer.


Acknowledgments: We thank the researchers who have helped build up the ELM resource and provide insights into linear motif functions within the cell. B.M. and Z.D. thank A. Reményi and A. Zeke (Institute of Enzymology, Research Centre for Natural Sciences, Hungarian Academy of Sciences) for many helpful discussions. Funding: B.U. acknowledges funding received as part of the RNA Bioinformatics Center of the German Network for Bioinformatics Infrastructure (031 AS38C RBC (de:NBII)) from the German Federal Ministry of Education and Research. M.K. is thankful for the postdoctoral fellowship from the Humboldt Foundation. M.B. was funded by the postdoctoral fellowship of the Hungarian Academy of Sciences and the OTKA grant (K115698). Z.D. acknowledges the support of the “Lendület” grant from the Hungarian Academy of Sciences (LP2014-18) and OTKA grant (K108798). Competing interests: The authors declare that they have no competing interests.