Cell, Vol. 97, 427–430, May 14, 1999, Copyright ©1999 by Cell Press

Substrate Targeting in the Ubiquitin System

Minireview

Jeffrey D. Laney and Mark Hochstrasser*
University of Chicago
Department of Biochemistry and
Molecular Biology
Chicago, Illinois 60637

Ubiquitin, named for its expression in a wide variety of tissues and organisms, is extraordinarily well conserved in creatures as diverse as yeast and man. Perhaps it is not surprising then that the proteolytic system involving ubiquitin plays a central role in the control of many basic cellular processes. In addition to degrading damaged, misfolded, or misassembled proteins, the ubiquitin system targets many naturally short-lived proteins, including transcription factors, cell growth modulators, signal transducers, and cell cycle proteins. A requisite step for the degradation of these substrates is the covalent attachment of ubiquitin. Targeted proteins are often modified by polymers of ubiquitin, which triggers their rapid destruction by a large, complex protease, the 26S proteasome (Hershko and Ciechanover, 1998).

The 76–amino acid ubiquitin polypeptide is conjugated to proteins via a reversible isopeptide linkage between the carboxy-terminus of ubiquitin and lysine side chains in the target protein. Attachment of ubiquitin to its substrates is mediated by a conserved cascade of enzymatic reactions. The pathway is initiated by the ATP-dependent activation of ubiquitin's C terminus by ubiquitin-activating enzyme (E1). A concerted, two-step reaction results in a high-energy thioester linkage between ubiquitin and E1. Through a thioester transfer reaction, the ubiquitin is then passed to a ubiquitin-conjugating enzyme (E2). These enzymes function with the E3 ubiquitin–protein ligases to attach ubiquitin to the ϵ-amino group of lysine residues in substrate proteins. In many cases, this ubiquitination complex catalyzes the formation of a polyubiquitin chain on the substrate.

*To whom correspondence should be addressed (e-mail: hoc1@midway.uchicago.edu).

Despite this fairly well-characterized enzymatic pathway for ubiquitin conjugation, how proteins are selected for their ultimate demise is only beginning to be clarified. This process must be highly specific, since short-lived proteins have to be identified and differentiated from the numerous stable polypeptides that coexist in the cell and sometimes even in the same protein complex. Degradation of substrates by the ubiquitin system appears to be mediated by specific degradation signals, which are sequence or structural features of the substrate that are required for rapid proteolysis. In addition to providing a ubiquitinatable lysine residue(s), these signals contain elements that are targeted directly by the ubiquitination apparatus, which in most cases is expected to be a specific E2/E3 complex. In this minireview, we highlight recent findings on the specific recognition of several physiological substrates of the ubiquitin system, emphasizing and contrasting the substrate characteristics that are recognized by the ubiquitin-conjugating machinery. We also consider how the association of these degradation signals with their respective ubiquitination complexes might be regulated. Ubiquitination of cell cycle regulators is discussed in the accompanying minireview (Koepp et al., 1999).

Phosphorylation-Dependent Degradation Signals

The transcription factor NF-κB, a central player in immune and inflammatory responses, is activated by the ubiquitin-proteasome pathway. In quiescent cells, NF-κB is held in a latent state in the cytoplasm in a complex with a member of the IκB family of inhibitors, such that the NF-κB nuclear localization signal (NLS) is occluded. In response to external stimuli, IκB is degraded in a ubiquitin-dependent manner, allowing nuclear translocation of NF-κB, which leads to a variety of transcriptional responses (reviewed by Ghosh et al., 1998).

The following model has emerged for the signal-

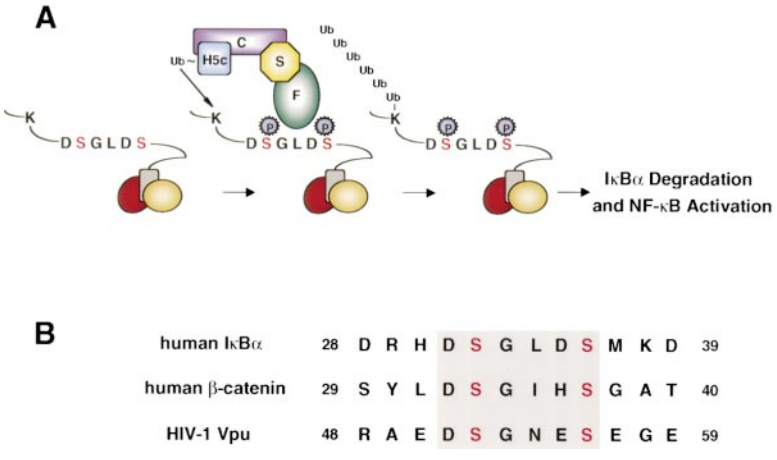

Figure 1. Model for the Phosphorylation-Dependent Degradation of IκBα and Comparison of β-TrCP Recognition Elements

(A) In response to extracellular stimuli, IκBα (in gray), in a complex with NF-κB (p50 subunit in garnet, p65 in gold), is phosphorylated by an IκB kinase on serines 32 and 36 (marked in red). The F box protein β-TrCP (denoted F) recognizes this phosphorylated degradation signal in association with Skp1 (S) and Cul-1 (C). This $SCF^{\beta\text{-TrCP}}$ complex promotes the UbcH5c (H5c)-dependent conjugation of ubiquitin (Ub) to IκBα. For simplicity, only one of the two major sites of ubiquitin conjugation is shown. The multiubiquitinated IκBα is then targeted to the 26S proteasome for degradation.

(B) A comparison of the β-TrCP recognition sequences found in human IκBα, human β-catenin, and HIV-1 Vpu. Similar sequences are shaded and the conserved serines, whose phosphorylation is required for β-TrCP recognition, are shown in red.

induced destruction of IκBα (Figure 1A). Stimulation of cell surface receptors, for example by cytokines, initiates a signal transduction cascade leading to the activation of an IκB kinase complex. This complex phosphorylates NF-κB-bound IκBα on its N-terminal regulatory domain at two specific serine residues, serines 32 and 36. Addition of these phosphates triggers the polyubiquitination of nearby lysines and the subsequent proteasomal degradation of IκBα. Although much is known about the upstream signaling processes that lead to IκBα destruction, it is only recently that experimental results have given insight into the downstream ubiquitination events (Yaron et al., 1997, 1998; Spencer et al., 1999; Winston et al., 1999). With the development of a faithful in vitro IκBα ubiquitination assay, the nature of the degradation signal in IκBα has been characterized and the receptor molecule that recognizes this motif has been defined.

When present in a complex with NF-κB, phosphorylated, but not unphosphorylated IκBα, can be ubiquitinated in cell-free extracts. Using peptides corresponding to the N-terminal regulatory domain of IκBα, Yaron et al. (1997) showed that the targeting component of the E3 ubiquitin ligase can be titrated by peptides that span the phosphorylation sites. IκBα-derived peptides phosphorylated at both serine 32 and 36 strongly inhibit ubiquitin conjugation, whereas unphosphorylated peptides have no effect. The efficacy of singly phosphorylated sequences is reduced at least 20-fold relative to the doubly phosphorylated versions. Peptides that lack the major sites of ubiquitin conjugation, lysines 21 and 22, are still effective inhibitors. Furthermore, immobilizing the inhibitory sequences on a solid support creates an affinity resin that can deplete cell lysates of their IκBα-ubiquitinating activity, without affecting the ubiquitination of other cellular proteins.

Taken together, the results provide compelling evidence for a short, phosphorylation-dependent recognition element in IκBα, DS*GLDS* (where S* represents phosphoserine); this E3-binding segment does not require a ubiquitin conjugation site. Interestingly, a similar sequence is found in the short-lived signal transducer/transcription factor β-catenin (Figure 1B), which is also degraded by the ubiquitin-proteasome system (Aberle et al., 1997). The sequence in β-catenin includes a consensus phosphorylation site for glycogen synthase kinase 3β, and mutation of the serines to residues that cannot be phosphorylated stabilizes the protein (Rubinfeld et al., 1997 and references therein).

Using the unusually well-defined IκBα degradation determinant as an affinity ligand, a host of different researchers have identified an E3 ubiquitin ligase component that binds to this element. The protein isolated by these groups is the human F box and WD domain protein β-TrCP (Yaron et al., 1998; Spencer et al., 1999; Winston et al., 1999). Both the F box and the WD domain are protein–protein interaction motifs present in many proteins, and several F box–containing proteins are involved in the ubiquitin–proteasome pathway (see below). The β-TrCP F box is necessary for IκBα ubiquitination in vitro, and the F box–deleted mutant is a dominant-negative inhibitor of IκBα degradation and NF-κB activation in transfected human cells. These criteria for specificity do not exclude the possibility that a protein related to but distinct from β-TrCP may be the physiologically relevant recognition component of the ubiquitin–IκBα ligase (see Fuchs et al., 1999) or that there are multiple, possibly tissue-specific, recognition factors.

The identification of β-TrCP as a receptor component of an E3 is particularly satisfying since it offers immediate mechanistic insight into a series of genetic observations from different experimental systems. For example, the *Xenopus* version of β-TrCP inhibits the Wnt/β-catenin signaling pathway (Marikawa and Elinson, 1998), and in *Drosophila*, the β-TrCP homolog Slimb regulates that same signaling pathway as well as the Toll/Dorsal pathway (Jiang and Struhl, 1998; Spencer et al., 1999). Inasmuch as β-catenin and its fly homolog Armadillo contain motifs very similar to the E3-binding element in IκBα (Figure 1B), β-TrCP and Slimb became obvious candidates for the recognition component of the ubiquitin ligases for these substrates as well. Recent results confirm this idea: β-TrCP interacts directly with β-catenin in a phosphorylation-dependent manner, and overproduction of an F box–deleted β-TrCP mutant in human cells blocks β-catenin degradation (Hart et al., 1999; Latres et al., 1999; Winston et al., 1999).

In human cells, β-TrCP can also associate with the HIV-1-encoded Vpu protein, a factor that mediates the proteasomal degradation of the HIV receptor CD4. Degradation of CD4, which occurs at the endoplasmic reticulum (ER), may help to prevent viral superinfection and appears to be important for viral envelope protein transport and maturation. In a striking parallel to IκBα and β-catenin, β-TrCP binds to Vpu in a phosphorylation-dependent manner, through a short sequence nearly identical to the IκBα recognition motif (Figure 1B). However, the virus has corrupted the system such that Vpu itself is not targeted for ubiquitination, but instead directs β-TrCP to the host protein CD4 (Margottin et al., 1998 and references therein). These observations suggest that similar targeting signals and recognition mechanisms are used by multiple ubiquitin system substrates.

The finding that an F box protein can function as a receptor for substrates of the ubiquitin pathway is not a novelty: a similar mechanism has been described for the degradation of yeast proteins that regulate the G1 phase and G1–S transition of the cell cycle (reviewed in Patton et al., 1998 and accompanying minireview). Like IκBα, proteolysis of these proteins is phosphorylation dependent, such that the ubiquitination apparatus only recognizes and modifies the phosphorylated substrate. The E3 involved in cell cycle protein recognition is a so-called SCF complex, composed of two evolutionarily conserved factors, *S*KP1 and a member of the *C*ullin family of proteins, in addition to an *F* box protein. β-TrCP functions in the context of a similar SCF assembly (Spencer et al., 1999; Winston et al., 1999), explaining the observation, noted earlier, that β-TrCP mutants lacking the F box, a motif necessary for binding to SKP1, fail to support IκBα ubiquitination in vitro and act in a dominant-negative fashion in vivo. These results, together with the identification of numerous F box proteins from genome sequencing projects, makes it tempting to conclude that the most common way to target a substrate to the ubiquitin-proteasome pathway in a phosphorylation-dependent manner will involve SCF-type E3 complexes. Interestingly, the E2 enzymes responsible for IκBα and cell cycle protein ubiquitination

appear to be different, suggesting an additional combinatorial element to the variability already provided by alternative F box proteins in different SCF assemblies.

The E3 recognition motifs in IκBα and β-catenin are short but specific sequences capable of E3 binding (in the former case at least) even in peptides as short as 7 amino acids. Such short peptides are not expected to exhibit a well-defined structure in solution. Hence, these E3-binding elements may be more akin to the short segments, common in signal transduction proteins, which are recognized via a few key residues in their sequence or by their ability to adopt a particular conformation upon binding to their targets, but not as part of a preformed protein surface. Examples of such motifs include the NLSs that bind nuclear transport factors and the phosphotyrosine-containing protein segments that are targeted by PTB and SH2 domains (reviewed by Harrison, 1996). The rather small size of the E3-recognition motifs raises the important question of how recognition of substrates like IκBα and β-catenin achieves sufficient specificity. Perhaps the interaction of substrate with the E2/E3 complex contributes to more than one step of the ubiquitination process, much like the recognition of appropriate tRNAs by the aminoacyl-tRNA synthetases. Small differences in binding energy can thereby be amplified to improve the fidelity of the ubiquitination reaction. Alternatively, these short protein segments may be the primary E3-binding elements, but a stable association between the E2/E3 complex and substrate may occur only after other weaker contacts are formed. If prolonged binding is necessary for the substrate to become polyubiquitinated, then only substrates that contain these multiple binding elements will acquire ubiquitin chains that reach the length preferred by the proteasome. Substrate deubiquitination might provide an additional proofreading mechanism, further enhancing the specificity of the system.

A Hydrophobic Protein Surface Required for Degradation

In contrast to the short peptide sequences defined for the recognition of IκBα and β-catenin, a surface of a prefolded structure appears to be a critical degradation determinant in the yeast mating type transcription factor α2. Budding yeast cell identity is controlled by the *MAT* mating type locus. The locus exists in haploid cells in two different forms, *MAT***a** and *MAT*α. In certain strains, switching between these two states can occur, with the associated phenotypic changes manifested within a single cell division. For such a rapid change in cellular phenotype, the transcriptional program controlled by the discarded *MAT* locus must be quickly dismantled. Indeed, all three *MAT*-encoded transcription factors (α1, α2, and **a**1) are exceptionally short-lived, and their metabolic stability is determined by the ubiquitin-dependent proteolytic system (Johnson et al., 1998). Maximal degradation of α2 requires four different E2 ubiquitin-conjugating enzymes (Ubc), with Ubc6 and Ubc7 targeting a degradation signal in the N-terminal domain of α2. This proteolytic motif has been termed *Deg1*.

Deletion analysis of the *Deg1* signal defined a region of ~60 amino acids sufficient for targeting to the ubiquitin pathway. Extensive point mutagenesis highlighted a 19-residue element that is critical for the rapid turnover of *Deg1*-containing substrates. This degradation determinant is part of an α2 segment that is predicted to form an amphipathic helix, and the inactivating mutations cluster on the hydrophobic face of this putative structural element (Johnson et al., 1998). The most straightforward interpretation of these results is that the Ubc6/Ubc7-containing ubiquitination apparatus recognizes *Deg1* by binding to the hydrophobic surface of the amphipathic helix in α2. Results from a recent screen in yeast for random protein fragments that act as Ubc6/Ubc7-dependent degradation signals are consistent with this view. Rather than defining specific consensus elements, the screen yielded relatively short peptide sequences with little in common except a strongly hydrophobic character (Gilon et al., 1998). A similar screen for artificial degradation signals in yeast also suggested that a hydrophobic surface can be recognized by a ubiquitination complex. Sadis et al. (1995) identified a peptide signal predicted to form an amphipathic helix, and subsequent mutagenesis experiments supported this structural model. However, unlike *Deg1*, the degradation of proteins containing this signal depends on Ubc4/Ubc5 in addition to Ubc6 and Ubc7. Ubc6 and Ubc7 also participate in the degradation of a number of different ER proteins (Sommer and Wolf, 1997). These substrates appear to be retrotranslocated to the cytosol following (or during) ubiquitination by the ER-localized Ubc6 and Ubc7 enzymes. While being ejected from the ER, these substrates are likely to expose, at least transiently, hydrophobic stretches that are normally in contact with the lipid bilayer or are buried in the protein. Together, these observations suggest that hydrophobic surfaces may be targets of the Ubc6/Ubc7-containing ubiquitination complex. However, since the E3 ubiquitin ligase(s) that functions with these E2s has not been identified and ubiquitination has not been reconstituted in vitro, there is as yet no direct evidence that these hydrophobic segments function in exactly this way.

The structural element critical for *Deg1* function overlaps a region important for binding of α2 to **a**1. This implies that **a**1 may compete with the Ubc6/Ubc7 ubiquitination apparatus for binding to α2. Consistent with this idea of overlapping sites of interaction, coexpression of **a**1 and α2, which occurs naturally in **a**/α diploids, dramatically stabilizes α2 and blocks its ubiquitination. Increasing the concentration of **a**1 in these cells further stabilizes α2. Conversely, mutations that impair the **a**1–α2 coiled-coil interaction interfere with the ability of **a**1 to stabilize α2 (Johnson et al., 1998). Thus, the *Deg1* signal appears to be masked by **a**1 binding.

Coexistence of degradation determinants within protein regions that also help form protein–protein interfaces is likely to be a common phenomenon. Exposure of such segments in components of multiprotein complexes that are synthesized in excess or fail to assemble properly would ensure the rapid elimination of unassembled polypeptides, helping to set the proper stoichiometry of these complexes. More generally, the display of hydrophobic surfaces, normally buried in protein–protein interfaces or found within protein core regions, could serve as a recognition element that helps identify misfolded or otherwise abnormal proteins. Here, the E3 ubiquitin ligases may have a surveillance function similar to that of molecular chaperones, except that the ligases would target these proteins bearing hydrophobic segments to a degradative pathway.

Substrate Recognition: Themes and Variations

As we have seen from the examples discussed above, the recognition mechanisms of many ubiquitin system substrates may be keyed to a fairly limited set of substrate features. Proteins whose degradation rates must be tightly coupled to environmental status, developmental state, or cell cycle stage are often phosphorylated via specific signal transduction cascades. These phosphorylated substrates are targets for particular variants of a large family of multisubunit ubiquitin–protein ligases, the SCF complexes. Interestingly, some of these substrates contain a rather short, closely related phosphopeptide motif.

A very different type of degradation determinant, exemplified by the yeast α2 repressor, is a solvent–exposed hydrophobic protein surface. Such exposure usually only occurs in nonnative proteins or in the absence of a protein partner that contacts and buries the surface. Modulation of protein degradation rates by regulated folding or protein–protein interaction will in general not have the temporal precision of a transiently activated phosphorylation cascade, so the recognition of hydrophobic determinants may be utilized primarily for protein quality control and for bringing about changes in proteolytic rate that occur on somewhat longer time scales. It also seems particularly well-suited for ER-associated proteolysis, for the reasons discussed earlier.

Notwithstanding these common themes, other types of protein motifs are also clearly recognized by the ubiquitin system. The first described E3 recognition determinant is a surprisingly simple one: the amino acid at the protein N terminus (Varshavsky, 1997). In certain proteins, specific N-terminal residues stimulate polyubiquitination and rapid degradation in vivo. A different ubiquitination determinant has been found in the large subunit of yeast RNA polymerase II (Rpb1). Rpb1 is recognized through its conserved carboxy-terminal domain (CTD) by the Rsp5 ubiquitin ligase; this domain is both necessary and sufficient for Rsp5 binding and ubiquitination. The CTD contains the heptapeptide repeat sequence SPTSPSY, which is likely to be recognized by the WW domains of Rsp5. These WW domains, named for their two absolutely conserved tryptophan residues, are protein–protein interaction modules that bind directly to proline-rich sequences or phosphoserine- and phosphothreonine-containing elements in their targets (Lu et al., 1999; Wang et al., 1999 and references therein). Interestingly, a similar interaction occurs between the human epithelial sodium channel (ENaC) and Nedd4, a mammalian homolog of Rsp5. ENaC, whose impaired degradation in the heritable disorder Liddle's syndrome is believed to cause hypertension, bears several proline-rich motifs that bind directly to the WW domains of Nedd4 (Goulet et al., 1998). These examples reinforce the notion that substrate discrimination can rely on very short sequence motifs, like those found in IκBα and β-catenin.

Thus, a few general principles of substrate recognition in the ubiquitin system are at last beginning to emerge, and these should be useful guides for the challenges that lie ahead. These include questions about the details of the E3-substrate interaction and how a substrate can then become polyubiquitinated. Other central issues are how the activity of an E3 is modulated and what controls E2–E3 interactions. Finally, it is known that ubiquitin chains of different lengths and distinct ubiquitin–ubiquitin linkages can form on substrates, so investigation of what regulates the length and topology of these chains will certainly be of interest. There is little doubt that answers to many of these questions will not be long in coming, given the current pace of work in the field.

Selected Reading

Aberle, H., Bauer, A., Stappert, J., Kispert, A., and Kemler, R. (1997). EMBO J. *16*, 3797–3804.

Fuchs, S.Y., Chen, A., Xiong, Y., Pan, Z.-Q., and Ronai, Z. (1999). Oncogene, in press.

Ghosh, S., May, M.J., and Kopp, E.B. (1998). Annu. Rev. Immunol. *16*, 225–260.

Gilon, T., Chomsky, O., and Kulka, R.G. (1998). EMBO J. *17*, 2759–2766.

Goulet, C.C., Volk, K.A., Adams, C.M., Prince, L.S., Stokes, J.B., and Snyder, P.M. (1998). J. Biol. Chem. *273*, 30012–30017.

Harrison, S.C. (1996). Cell *86*, 341–343.

Hart, M., Concordet, J.P., Lassot, I., Albert, I., del los Santos, R., Durand, H., Perret, C., Rubinfeld, B., Margottin, F., Benarous, R., and Polakis, P. (1999). Curr. Biol. *9*, 207–210.

Hershko, A., and Ciechanover, A. (1998). Annu. Rev. Biochem. *67*, 425–479.

Jiang, J., and Struhl, G. (1998). Nature *391*, 493–496.

Johnson, P.R., Swanson, R., Rakhilina, L., and Hochstrasser, M. (1998). Cell *94*, 217–227.

Koepp, D.M., Harper, J.W., and Elledge, S.J. (1999). Cell *97*, this issue, 431–434.

Latres, E., Chiaur, D.S., and Pagano, M. (1999). Oncogene *18*, 849–855.

Lu, P.J., Zhou, X.Z., Shen, M., and Lu, K.P. (1999). Science *283*, 1325–1328.

Margottin, F., Bour, S.P., Durand, H., Selig, L., Benichou, S., Richard, V., Thomas, D., Strebel, K., and Benarous, R. (1998). Mol. Cell *1*, 565–574.

Marikawa, Y., and Elinson, R.P. (1998). Mech. Dev. *77*, 75–80.

Patton, E.E., Willems, A.R., and Tyers, M. (1998). Trends Genet. *14*, 236–243.

Rubinfeld, B., Robbins, P., El-Gamil, M., Albert, I., Porfiri, E., and Polakis, P. (1997). Science *275*, 1790–1792.

Sadis, S., Atienza, C., Jr., and Finley, D. (1995). Mol. Cell Biol. *15*, 4086–4094.

Sommer, T., and Wolf, D.H. (1997). FASEB J. *11*, 1227–1233.

Spencer, E., Jiang, J., and Chen, Z.J. (1999). Genes Dev. *13*, 284–294.

Varshavsky, A. (1997). Genes Cells *2*, 13–28.

Wang, G., Yang, J., and Huibregtse, J.M. (1999). Mol. Cell Biol. *19*, 342–352.

Winston, J.T., Strack, P., Beer-Romero, P., Chu, C.Y., Elledge, S.J., and Harper, J.W. (1999). Genes Dev. *13*, 270–283.

Yaron, A., Gonen, H., Alkalay, I., Hatzubai, A., Jung, S., Beyth, S., Mercurio, F., Manning, A.M., Ciechanover, A., and Ben-Neriah, Y. (1997). EMBO J. *16*, 6486–6494.

Yaron, A., Hatzubai, A., Davis, M., Lavon, I., Amit, S., Manning, A.M., Andersen, J.S., Mann, M., Mercurio, F., and Ben-Neriah, Y. (1998). Nature *396*, 590–594.