Numerous proteins with unique characteristics are degraded by the 26S proteasome following monoubiquitination

Braten, O. and Livneh, I. and Ziv, T. and Admon, A. and Kehat, I. and Caspi, L.H. and Gonen, H. and Bercovich, B. and Godzik, A. and Jahandideh, S. and Jaroszewski, L. and Sommer, T. and Kwon, Y.T. and Guharoy, M. and Tompa, P. and Ciechanover, A.

Proceedings of the National Academy of Sciences of the United States of America
2016 AUG 09 ; 113(32): E4639-E4647
2016 JUL 06 (first published online)
doi: 10.1073/pnas.1608644113

Publisher: National Academy of Science

Numerous Proteins with Unique Characteristics are Degraded by the 26S Proteasome Following Monoubiquitination

Ori Braten[1]*, Ido Livneh[1]*, Tamar Ziv[2], Arie Admon[2], Itzhak Kehat[3], Lilach Caspi[3], Hedva Gonen[1], Beatrice Bercovich[1], Yong-Tae Kwon[4], Adam Godzik[5], Samad Jahandideh[5], Lukasz Jaroszewski[5], Mainak Guharoy[6], Peter Tompa[6,7,8] and Aaron Ciechanover[1+]

[1]Technion Integrative Cancer Center (TICC),
The Rappaport Faculty of Medicine and Research Institute,
[2]Smoler Proteomic Center and Faculty of Biology,
[3]Department of Physiology, The Rappaport Faculty of Medicine and Research Institute,
Technion-Israel Institute of Technology, Haifa, Israel;
[4]Protein Metabolism Medical Research Center and Department of Biomedical Sciences,
College of Medicine, Seoul National University, Seoul, South Korea;
[5]Bioinformatics and Systems Biology Program,
Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA, USA;
[6]VIB Structural Biology Research Center (SBRC); and
[7]Structural Biology Brussels (SBB), Vrije Universiteit Brussel, Brussels, Belgium,
and [8]Institute of Enzymology,
Research Center for Natural Sciences of the Hungarian Academy of Sciences,
Budapest, Hungary

Running title: Monoubiquitination-mediated proteasomal degradation

*Equal contribution

+Correspondence to:
Aaron Ciechanover
Faculty of Medicine
Technion-Israel Institute of Technology,
Efron Street, P.O. Box 9649,
Haifa 31096, ISRAEL

Tel: +972-4-829-5427
Fax: +972-4-852-1193

E-mail: aaroncie@tx.technion.ac.il
Submitted: PNAS

Abstract

The ‘canonical’ proteasomal degradation signal is a substrate-anchored polyubiquitin chain. However, a handful of proteins were shown to be targeted following monoubiquitination. In this study, we established - in both human and yeast cells - a systematic approach for the identification of monoubiquitination-dependent proteasomal substrates. The cellular wild type polymerizable ubiquitin was replaced with ubiquitin that cannot form chains. Using proteomic analysis, we screened for substrates that are nevertheless degraded under these conditions compared to those that are stabilized, and therefore require polyubiquitination for their degradation. For randomly sampled representative substrates, we confirmed that their cellular stability is in agreement with our screening prediction. Importantly, the two groups display unique features: monoubiquitinated substrates are smaller than the polyubiquitinated ones, are enriched in specific pathways, and in humans, are structurally less disordered. We suggest that monoubiquitination-dependent degradation is more widespread than assumed previously, and plays key roles in various cellular processes.

Significance Statement

Unlike the prevailing ‘canon’ that a polyubiquitin chain (with a minimal number of four moieties) serves as the most common proteasomal degradation signal, with only a few exceptional substrates targeted following monoubiquitination, we show here that this mode of proteasomal-targeting modification is much more common. To demonstrate it, we designed a screen that employs a proteomic analysis to identify cellular proteins that are degraded in the exclusive presence of non-polymerizable ubiquitin. The screen revealed numerous such substrates, suggesting that this mode of proteasomal recognition is more widespread than assumed previously. Importantly, the screen revealed also the proteins that require polyubiquitination for their degradation, as they are stabilized in the presence of the non-polymerizable ubiquitin. For randomly selected individual substrates, we confirmed that their mode of modification and subsequent degradation is in agreement with the predictions, corroborating the reliability of the screen. Notably, the two classes of proteins modified by the two different modes of ubiquitination display certain distinct important characteristics. Monoubiquitinated proteins appear to be of lower molecular mass, of lesser structural disorder, and can be assigned to defined cellular pathways/processes. Furthermore, some of the unique characteristics are confined to either human or yeast cells, suggesting that the mechanism of action/recognition of the ubiquitin system in the two organisms must be different somehow.

Introduction

Polymers of ubiquitin (Ub) are formed on proteasomal substrates in eukaryotic cells by the concerted action of three enzymes: The Ub-activating enzyme (E1), a Ub-carrier protein [E2; known also as UBC (Ub-conjugating enzyme)] and a Ub ligase (E3) which is the specific substrate-recognizing element of the system. The Ub chains typically consist of multiple moieties linked to one another via an isopeptide bond between the C-terminal Gly residue of the distal moiety and the ε-NH_2 group of Lys48 of the proximal one (1).

In addition, the system can also catalyze modification by a single Ub moiety (monoubiquitination) or multiple single Ub moieties (multiple monoubiquitinations), each modifying a distinct lysine residue (2). Generally, monoubiquitination has been conceived as a non-destructive signal. Further, it has been suggested that efficient proteasomal targeting requires a chain with a minimal length of four Ub moieties (3). Monoubiquitination is known to be involved in multiple biological processes. For example, monoubiquitination of proteins containing a ubiquitin-binding domain (UBD) often mediates autoinhibition by a UBD-Ub interaction (4, 5). Signal transduction by membrane receptors, such as the EGFR, is attenuated by monoubiquitination-mediated receptor internalization (6). The subcellular localization of small GTPases is controlled, among other post-translational modifications, by monoubiquitination (7, 8), and histone monoubiquitination regulates nucleasomal structure, thus affecting gene expression (9). Emerging reports indicate, however, that several substrates can be degraded following monoubiquitination (10–13). Partial degradation/processing of the p105 precursor of NF-κB, which results in release of the p50 active subunit of the transcription factor, is dependent on multiple monoubiquitinations (14). Importantly, these findings demonstrate that the proteasome can recognize a single Ub moiety(s), and imply the existence of mono- *vs* polyubiquitination "decision" mechanisms. In this context, a previous study has suggested that the chain length required for proteasomal degradation is determined by the size of the substrate, and possibly other characteristics that affect the affinity of the modified substrate to the proteasome. Specifically, it was suggested that substrates smaller than 150 amino acids are degraded following monoubiquitination, whereas longer substrates require longer chains. Thus, a dynamic model was proposed, according to which the chain elongates to a point where the affinity to the proteasome is high enough to secure a stable binding of the conjugated substrate, its concomitant detachment from the E3, and its subsequent degradation (15). Another study demonstrated that restricting the number of ubiquitinatable Lys residues can switch the mode of modification necessary for degradation from multiple monoubiquitinations to polyubiquitination, suggesting that in the cell, the masking of Lys residues by protein-protein interactions or post translational modifications can affect the mode of ubiquitination (11).

However, all these studies have been carried out using specific substrates. Therefore, general conclusions regarding monoubiquitination-dependent degradation mechanisms, the population of

substrates that are degraded following this modification, and importantly, whether they have common distinct characteristics, have remained limited.

In this study, we employed a systematic proteomic approach for the identification and characterization of monoubiquitination-dependent proteasomal substrates. By silencing the endogenous WT Ub followed by expression of non-polymerizable lysine-less Ub, we identified numerous substrates in both mammalian and yeast cells that are targeted by the proteasome following mono- or multiple monoubiquitinations. Interestingly, we confirmed a previous hypothesis (15) that there is a correlation between the length of the substrate and its requirement for either mono- or polyubiquitination, though it is stronger in yeast. Also, the monoubiquitinated substrates are enriched in specific pathways (e.g. oxidative stress, carbohydrate transport and components of the ubiquitin-proteasome system itself), and in humans, are structurally less disordered.

Results

Establishing a system for induction of monoubiquitination in cells

In general, to model monoubiquitination (or multiple monoubiquitinations), we silenced endogenous Ub expression and replaced it with a lysine-less Ub (Ub^{K0}). This non-polymerizable Ub species, in which all seven Lys residues were replaced with Arg, can modify each Lys in the target substrate only once, and cannot be further ubiquitinated (16–19).

To study monoubiquitination in yeast, we used a modification of a previously described Ub replacement method (20). Briefly, all the Ub-coding genes were deleted and replaced by galactose-inducible Ub (ΔUb strain), and either copper-inducible Ub^{WT} or Ub^{K0} (ΔUb^{UbWT} or ΔUb^{UbK0}, respectively). Thus, Ub expression can be silenced by adding glucose, and Ub re-expression can be induced by adding copper to the growth medium (Figure 1A). To validate Ub silencing and re-expression, ΔUb, ΔUb^{UbWT} and ΔUb^{UbK0} yeast strains were treated with glucose and copper. As shown in Figure 1B, Ub expression was efficiently suppressed, and both Ub^{WT} and Ub^{K0} were markedly expressed.

To assess monoubiquitination in human cells, we used a modification of a previously described Ub replacement model in human cultured cells (21). Briefly, endogenous Ub is silenced in U2OS cells by a Ub-specific tetracycline-induced shRNA (shUb), and either HA-Ub^{WT} or HA-Ub^{K0} is expressed following infection with an adenoviral vector (Figure 1A). To evaluate Ub silencing efficiency, we monitored Ub and Ub conjugates level in $U2OS^{shUb}$ cells following tetracycline treatment. As demonstrated in Figure 1Ci, the level of both Ub and Ub-protein conjugates were significantly decreased. In the endogenous Ub-silenced cells, both HA-Ub^{WT} and HA-Ub^{K0} were efficiently expressed and assembled into high molecular mass conjugates following adenoviral expression (Figure 1Cii). It should be noted that the pattern of conjugation appears similar for both Ub^{WT} and Ub^{K0} expression. This is probably due to the numerous substrates with a broad range of molecular mass that are conjugated, and from the possibility that many of them are modified by multiple monoubiquitinations.

To demonstrate Ub replacement using an additional method, we quantified Ub using mass spectrometry. As shown in Figure S1A, tryptic digestion of Ub^{WT} and Ub^{K0} yields both common and differential mass spectrometry (MS)-detectable peptides. To assess Ub replacement in yeast, we treated ΔUb^{UbK0} cells with glucose and copper, and quantified Ub-derived peptides by MS. As illustrated in Figure S1B, Ub^{K0} was markedly more abundant than endogenous Ub. To evaluate Ub^{K0} expression in human cells, we overexpressed HA-Ub^{K0} via adenoviral infection using increasing MOIs. As displayed in Figure S1C, Ub^{K0} expression level was MOI-dependent, and significantly exceeded the level of endogenous Ub.

Taken together, these data demonstrate the effectiveness of our Ub replacement strategies, and suggest that our experimental systems are suitable for studying protein monoubiquitination.

Systematic identification of monoubiquitination-dependent proteasome substrates

To identify substrates that are degraded following monoubiquitination, we replaced Ub with either Ub^{K0} or Ub^{WT} (as a control). We then used anti-K-ε-GG immunoprecipitation (Figure S2) to enrich and quantify by MS GlyGly-modified peptides derived from tryptic digestion of ubiquitinated proteins (Figure S2). This method enabled us also to identify ubiquitination sites (Ubsites). To verify that the ubiquitinated proteins serve indeed as proteolytic substrates, we monitored also the level of non-modified peptides derived from them (Figure 2A). To ascertain reproducibility, we performed several independent biological replicates for each model organism, using both SILAC labeling and label-free quantification (Table S1). As shown in Figure 2B, identification of the proteins was quite reproducible. Similar to previous data (22), identification of Ubsites within proteins was less reproducible.

This established experimental set-up enabled us to discriminate between proteins degraded following modification by mono- (or multiple mono-) ubiquitination, and those that are degraded only following polyubiquitination. We calculated the MS signal intensity ratios following either Ub^{K0} or Ub^{WT} expression for both proteins and immunoprecipitated ubiquitinated sites (denoted 'protein K0/WT intensity ratio' and 'site K0/WT intensity ratio', respectively. See Dataset S1 for raw data). As illustrated in Figure 2C and based on K0/WT intensity ratios, we classified ubiquitinated proteins to putative monoubiquitination- and polyubiquitination-dependent proteasome substrates as following:

Monoubiquitination-dependent proteasomal substrates are expected to be unaffected by Ub^{K0} expression. Alternatively, as Ub^{K0} expression renders proteasomes less occupied by polyubiquitination-dependent substrates, increased proteasome availability may result in accelerated degradation of monoubiquitination-dependent substrates. Thus, we required these substrates to:

(i) have a site K0/WT ratio <1;

(ii) have a detectable MS signal in at least two independent experiments; and

(iii) have a protein K0/WT ratio <1 (if the protein is detectable)

The degradation of a polyubiquitination-dependent substrate is expected to be inhibited upon Ub^{K0} expression. Consequently, we expect their level to increase. Thus, we require these substrates to:

(i) have a site K0/WT ratio >1;

(ii) have a detectable MS signal in at least two independent experiments; and

(iii) have a protein K0/WT ratio >1 (if the protein is detectable)

A small fraction (<3%) of proteins were identified as belonging to the two groups. These proteins were excluded from the survey.

Applying these criteria in both yeast and human cells, we identified 90 and 255 monoubiquitination- and 449 and 328 polyubiquitination-dependent putative proteasomal substrates, respectively (Dataset S2). Samples of each group are presented in Table 1, describing gene names and

ubiquitinated Lys positions. As expected, the polyubiquitination-dependent substrate group included several previously suggested proteasomal substrates, e.g. PDC1, OLE1 and ENO1 in yeast, and HIF1A, POLD2 and IER3 in human cells (23, 24).

Candidate substrate validation

To validate the results of our algorithm and experimental setup, we monitored the cellular stability (using cycloheximide chase) of randomly sampled representative candidate substrates following Ub replacement. As demonstrated in Figure 3A, replacing Ub^{WT} with Ub^{K0} stabilized ARD1 in yeast and CDC20 in human cells (polyubiquitination-dependent substrates; Table 1). In contrast, the predicted monoubiquitination-dependent substrates, GRE1 in yeast and GOT1 in human cells, remained unstable. All four substrates were clearly degraded by the 26S proteasome, as they were stabilized following treatment with a proteasome inhibitor (Figure 3B). Taken together, these results strongly suggest that our experimental setup is suitable for the systematic identification of monoubiquitination-dependent proteasomal substrates.

Physical characteristics of protein substrates play a role in their mode of ubiquitination:

(i) Structural disorder

Evidently, a significant number of proteins are degraded following monoubiquitination (and probably also multiple monoubiquitinations) in both yeast and human cells. This observation challenges the prevailing paradigm of polyubiquitination being the prerequisite for protein degradation. Intriguingly, bioinformatics analysis of the data reveals that yeast and human cells significantly differ in their preferences for using the two types of signals. In yeast, the ratio of proteins degraded by polyubiquitination vs. monoubiquitination is 5.0 (449/90), whereas in human cells it is only 1.3 (328/255). Since the identification method was unbiased, the significant difference in the ratios likely points to important differences in the mode of recognition of the ubiquitin-proteasome system (UPS) in the two species. Yeast apparently operates much more by the traditional signal, polyubiquitin, whereas in humans monoubiquitination is used almost as frequently as polyubiquitination to mark proteins for degradation. The difference may arise from a combination of factors such as using different sets of conjugating enzymes (E2s and E3s), different preferences for local sequence and structural features of substrates, and the recognition elements of the proteasome. Some of these issues are addressed below.

In previous studies, it has been shown that protein ubiquitination and degradation are intimately linked with structural disorder. Intrinsically disordered protein regions lack a well-defined tertiary structure, yet they fulfil important functional roles linked with their highly flexible and adaptable structure (25–27). Structural disorder correlates with all three elements of degradation signals: location of the ubiquitin ligase recognition motif on substrates, the Lys residue(s) to which ubiquitin is attached, and a nearby long disordered region (LDR, a region of at least 30 consecutive disordered residues) that initiates the unfolding of the substrate engaged with the proteasome (15, 28–30).

Structural disorder may also be required for ubiquitin conjugation itself, in two different ways. It is repeatedly reported that the mutation of the Lys residue that is the site of modification does not usually abrogate sensitivity to UPS degradation (11, 31), because flexibility of the substrate enables multiple modifications on neighboring Lys resides. By a similar logic, the buildup of a polyubiquitin chain can also benefit from local structural disorder, because it enables the processive addition of subsequent ubiquitin moieties to the end of the growing polyubiquitin chain.

To test if these disorder features may be related to the use of mono- vs. polyubiquitination, we tested if predicted structural disorder of proteins differs in the different datasets (Figure 4A). We found significant differences between yeast and human proteins: in yeast, the occurrence of structural disorder does not differ between mono- and polyubiquitinated substrates, whereas in human cells, structural disorder prevails in polyubiquitinated substrates. Given that polyubiquitin is a stronger signal, we presume that structural disorder is primarily used for the buildup of the polyubiquitin chain. Since yeast relies more heavily on polyubiquitination, yet its polyubiquitinated substrates do not have more disorder, it is possible that their E3 ligases differ from humans in some critical features. These can be their number/redundancy, structural disorder, or binding heterogeneity, the structural disorder of which ensures processive addition of ubiquitin moieties in the ubiquitin chain.

These differences also have to manifest themselves in the local disorder of the protein chain around the ubiquitinated Lys residues. In general, the sites of post-translational modifications (PTMs) in proteins tend to exhibit local disorder, as studied in detail for phosphorylation (32) and also ubiquitination (28). In our entire dataset, the lysines that are the sites of ubiquitination tend to be locally disordered, but even more importantly, they show a highly characteristic difference between the two species. In yeast, monoubiquitination sites, whereas in human cells polyubiquitination sites, are significantly more disordered (Figure 4B). Our interpretation, again, is that these differences point to likely differences in the UPS in the two species. In yeast, polyubiquitination is robust, and it is the rare monoubiquitination sites that have to be supported by local disorder, probably more for initiation of degradation than modification (29). In human cells, the sites of polyubiquitination are significantly more disordered, probably as much for the processive build-up of the polyubiquitin chain as for initiation of degradation (as suggested above). These are genuine differences and do not result from the natural tendency of lysines to be locally disordered, which, as disorder-promoting amino acids, tend to be located in disordered regions of proteins (33) (Figure S3A).

The distinction between the signaling strength and functionality of mono- and polyubiquitin chains in yeast and human cells is also reflected in characteristic differences in the proximity of LDRs to Ubsites, which might be the sites of initiation of proteasomal degradation (28–30). The need of such assistance for monoubiquitination sites in yeast is apparent from the larger proportion of such sites that are close to an LDR (Figure 4C), and probably also by a larger proportion of such sites that are close to the termini of the proteins, which are generally flexible (Figure S3B).

Conservation in evolution can be an important indicator of the functionality of PTM sites. For example, functional phosphorylation sites (i.e. were shown to have a functional role by direct or indirect evidence - e.g. mutagenesis and/or functional assays) evolve significantly slower than those without evidence for a functional role (34). Interestingly, Ubsites are significantly more conserved in human cells than in yeast (Figure S3C) (these differences are significant because the phylogenetic coverage of the respective multiple alignments is comparable). These results infer that yeast sites are under a significantly lower evolutionary pressure, which may point to the fact that they are functionally more promiscuous.

Physical characteristics of the proteins play a role in their mode of ubiquitination:
(ii) size of the protein

In a previous study, we have shown that substrates degraded by the attachment of a single ubiquitin chain are usually shorter than 150 amino acids (15), which suggests that monoubiquitination is a weaker signal. In accord, there is significant enrichment of shorter proteins among yeast monoubiquitinated substrates, and also a similar trend in humans (Figure S3D). This is particularly true, if we take only average-length proteins (below 600 amino acids) into consideration. This difference is not reflected in a preference of multiple modifications in longer proteins: there are no significant differences in the length of substrates that are degraded after the attachment of a single vs. multiple mono- or polyubiquitin chains (Figure S3E).

Monoubiquitination- and polyubiquitination-dependent substrates are differentially enriched in specific biological processes

To characterize the two different modes of ubiquitination in a cellular function perspective, we searched for enriched Gene Ontology (GO) terms, using the GOrilla tool (35, 36). As illustrated in Figure 5A, it appears that monoubiquitination-dependent proteasomal substrates are enriched in oxidative stress response and carbohydrate transport pathways. As presented in Figure 5B, molecular function-specific GO terms were also identified. For example, monoubiquitination-dependent substrates were found enriched among ribosomal and proteasomal subunits. Similarities between the mono- and polyubiquitinated substrates were also observed, as Ub system components were enriched in both groups. Interestingly however, monoubiquitinated substrates consisted of mostly Ub ligases, E3s (*e.g.* HERC3, ITCH, XIAP), whereas polyubiquitinated substrates were enriched with Ub-conjugating enzymes, E2s (*e.g.* UBE2J1, UBE2T, UBE2L6).

Ubiquitination site sequence analysis reveals unique patterns

Previous studies have demonstrated motifs and sequence patterns specific for protein post-translational modifications such as phosphorylation, acetylation and SUMOylation (29, 30, 37, 38). To identify Ubsite motifs, we analyzed our data using the Motif-X algorithm (39). Consistent with previous findings (24), no specific motif was found when analyzing all identified Ubsites in either yeast or human cells. To analyze Ubsite amino acid composition, all identified Ubsites were examined. We constructed an alignment of peptides that are surrounding each modified Lys residue. Residue- and position-specific amino acid occurrences were calculated and were compared to the corresponding proteomic occurrence in a lysine-centered peptide. As depicted in Figure S4A, ubiquitination sites demonstrated residues-specific enrichment (*e.g.* Ala, Gly, Gln) and under-representation (*e.g.* Cys, His, Lys, Met, Trp) for both yeast and human cells. Enrichment of Glu, and to a lesser extent Asp, was specific to yeast sites. Importantly, for several amino acids, the enrichment factor depended on the proximity to the modified Lys, suggesting that these residues can affect conjugation mechanisms.

To compare the ubiquitination site composition of monoubiquitination- vs. polyubiquitination-dependent substrates, we performed the above analysis for each group separately. As shown in Figures S4B and S4C, each group displays a unique pattern of enrichment factors, and several differences can be observed between the patterns of monoubiquitination- and polyubiquitination-dependent substrates.

Discussion

In this study we identified 90 yeast and 255 mammalian proteins that are degraded by the proteasome following monoubiquitination. These significant numbers point to a much broader phenomenon to what was considered until recently as an exception. Therefore, it allows for analysis that sheds light on the mechanisms that underlie the different modes of ubiquitination. Being aware of the limitations of a proteomic screen, we assume the numbers are larger.

Experimentally, we replaced Ub^{WT} with Ub^{K0} in cells to enforce monoubiquitination. This strategy has been used successfully to inhibit polyubiquitination-dependent proteasomal degradation (14, 17–19). Other methods to study monoubiquitination have been reported, such as inhibition of polyubiquitin chain formation by methylated Ub (40), or detection by Western Blotting of specific substrates that appear to be monoubiquitinated (10). However, using methylated Ub is limited to cell free systems, and using endogenous Ub can identify only individual substrates. Therefore, these methods limit the ability to identify and characterize the broad population of target substrates degraded by the proteasome following mono- and polyubiquitination. Thus, Ub^{WT}-to-Ub^{K0} replacement in cells seemed to be the most suitable strategy for our objective of carrying out a proteome-wide screen to identify these two distinct populations.

Importantly, although efficient Ub^{K0} expression inhibits polyubiquitination, it may still support the conjugation of several Ub^{K0} molecules to a protein substrate, resulting in multiple-monoubiquitinations. In our study, we included mono- and multiple-monoubiquitinated substrates in the same group, as discriminating between the two is complicated experimentally.

Importantly, in our survey we decided not to use proteasome inhibitors as a tool to identify substrates that are nevertheless degraded following Ub replacement. The reason being that in human cells infected with adenoviral HA-Ub^{WT}, we could not observe up-regulation of Ub conjugates following treatment with a proteasome inhibitor (Figure S5A). It should be noted that the proteasome in the cells was active (Figure S5B), and its inhibition was efficient (Figure S5C). This finding is consistent with partial proteasome inhibition by free Ub chains that are accumulated due to the high level of HA-Ub^{WT} (41). Also, the use of proteasome inhibitors is challenging in yeast cells, as low permeability results in low cellular concentration of the drugs (42). Therefore, we have adapted alternative classification criteria to identify mono- vs. polyubiquitinated proteasomal substrates based on MS analyses of the proteome and ubiquitome under normal conditions and under conditions where monoUb is predominant (see under *Results* and in Figure 2B). To confirm our results and to nevertheless relate them to proteasomal degradation, we integrated data from previous studies (23, 24, 43) and constructed a reference list of known ubiquitin-proteasome substrates (Dataset S3). Compared to this list, our polyubiquitination-dependent substrates were highly enriched with known proteasomal substrates in both yeast and human cells (P-value = 6.90×10^{-23} and P-value = 5.06×10^{-5}, respectively, by hypergeometric test). This strongly suggests that our experimental model is faithful and offers a

reliable method for the identification of UPS substrates. Furthermore, we validated biochemically that several mono- and polyubiquitination-dependent candidates emerged from the screen do indeed belong to their respective expected categories (Figure 3).

It is interesting to refer to specific proteins that were identified in previous studies as targeted by monoubiquitination. Thus, Syndecan-4 (12) did exhibit site K0/WT ratio <1 (for Lys105), but was detected in one replicate only. Cks-2 (15) displayed protein K0/WT ratio <1 at the protein level, but had conflicting ratios at site level. Other substrates including Pax-3 (13) and α-synuclein (10) could not be detected, probably due to low abundance in the bone-derived U2OS cells (44). Taken together, it seems that our results are in agreement with random previous data, but should be further substantiated by experiments in cells from different tissues.

Notably, since we did not use proteasome inhibitors, some of our monoubiquitination-dependent candidates may actually be degraded via other cellular pathways. Accordingly, it was previously shown that receptor tyrosine kinases (RTKs) (45) and other membrane proteins are subjected to monoubiquitination-dependent lysosomal degradation.

From the bioinformatics analyses of the substrates, several important and intriguing conclusions can be drawn. First, in agreement with previous studies in which it was shown that substrates of up to 150 residues can be degraded following monoubiquitination (15), we found that the distribution of monoubiquitination-dependent substrates is shifted towards shorter proteins (Figure S3D).

Further, a difference between yeast and human is also apparent when comparing the ratio of monoubiquitination- and polyubiquitination-dependent substrates. Yeast relies more heavily on polyubiquitination, whereas human cells use both monoubiquitination and polyubiquitination with a similar frequency. This difference can be interpreted if structural disorder is considered as shown by our own data (Figures 4 and S3), and if we assume that a single ubiquitin moiety is a weaker signal for degradation than polyubiquitin. As suggested, local structural disorder is involved in various steps of the UPS cascade, from recognition motifs of E3 ligases through local disorder of ubiquitination sites to an LDR initiation site of substrate unfolding (28–30).

We have screened our monoubiquitination-dependent substrates for enriched biological process-related GO terms, and found a highly significant over-representation of genes associated with carbohydrate transport and oxidative stress response pathways (Figure 5A). Since carbohydrate transporters are plasma-membrane proteins, this finding is consistent with previous studies which demonstrated membrane receptors down-regulation via monoubiquitination-mediated endocytosis (46). Notably, as this pathway results in lysosomal/vacuolar rather than proteasomal degradation, this finding highlights the challenge in distinguishing between these two degradation modes using our experimental system. Oxidative stress was shown to activate cellular signal transduction cascades, and to result in gene expression modulation (47). The enrichment of oxidative stress response proteins in our monoubiquitination-dependent substrates may suggest that they are regulated by a common monoubiquitinating E3 ligase(s). Thus, oxidative stress-mediated down regulation of this putative E3

may result in up regulation of oxidative stress pathway components and activate the respective cellular response. Consistently, microarray experiments have shown that the expression of the E3s UBR1 and HUL4, and the E2s CDC43, RAD6 and UBC11, is decreased following exposure to oxidative stress (48).

UPS components were enriched in both groups (Figure 5B). This finding most likely represents the previously reported autoubiquitination of E2s (49–51) and E3s (52–54). Our findings show that E3s are preferably classified as monoubiquitination-dependent substrates.

In this study, we have determined the sequence positions of thousands of ubiquitinated lysines. The unsuccessful attempts to identify a ubiquitination-site motif in this study and in others (24, 55) reflect site-level promiscuity which is supported by low ubiquitination sites conservation across eukaryotic species (55), and by the flexible selection of the ubiquitinated lysines within a given substrate (11, 31). Additionally, we used position-specific analysis of relative amino acid abundance to characterize ubiquitination sites (Figure S4). This method yields a more thorough representation of the amino acid composition, and reflects both enrichment and under-representation trends. According to the notion of promiscuity, a ubiquitination site should merely provide a sterically available ε-amino group of a Lys residue. Consistently, we found enrichment of small residues (e.g. Ala and Gly) and decrease in bulky (e.g. Trp) residues in the proximity of the ubiquitinated Lys. The depletion of Pro at position -1 also supports this concept, since Pro disrupts the protein's secondary structure and might impair the solvent accessibility of a following Lys residue.

Experimental Procedures

Adenovirus-mediated Ub replacement in human cells

For Ub silencing, U2OS^{shUb} [described in (56)] cells were treated with 1 µg/ml tetracycline for 24 hr. Fresh tetracycline and adenoviruses encoding either HA-Ub^{WT} or HA-Ub^{K0} were added, and cells were incubated for additional 24 hr.

Ub replacement in yeast

The construction of ΔUb strain was described previously [SUB328; (57)]. Briefly, endogenous Ub genes were deleted and replaced with a Ub gene expressed under a Gal promoter. To construct ΔUb^{UbWT} and ΔUb^{UbK0} strains, ΔUb yeast cells were transformed with pUb39 Ub^{WT} and pUb39 Ub^{K0}, respectively (both genes are under the Cup1 promoter). To replace the Gal-induced Ub, ΔUb^{UbWT} and ΔUb^{UbK0} yeast cells were grown to 1.0 OD_{600nm} in standard Hartwell's complete medium (HC) without glucose, supplemented with 2% galactose and 2% raffinose. Cells were then washed in DDW to remove galactose and raffinose, and re-suspended in HC medium with glucose and 50 µm $CuSO_4$. Cells were incubated for 16 hr at 30^{o}C, collected by centrifugation, washed in DDW and frozen in liquid N_2.

Ubiquitination sites detection using GG-modified peptide enrichment

Dried peptides (see under *Sample preparation for mass spectrometry* in *Supplemental Information*) were re-suspended in immunoaffinity purification (IAP) buffer (50 mM MOPS/NaOH pH 7.2, 10 mM Na_2HPO_4 and 50 mM NaCl), and cleared by centrifugation. Supernatants were adjusted to pH 7.0 with NaOH and incubated with immobilized anti-K-ε-GG antibody (Cell Signaling Technology) at 4^{o}C for 3 hr. Beads were washed with IAP buffer and then with a wash buffer (500 mM NaCl, 3 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, 0.1% octylglucoside, pH 7.4). GG-modified peptides were eluted with 0.2% trifluoroacetic acid (TFA), desalted on C18 tips, and eluted in two fractions of 20% and 80% acetonitrile. Peptides were analyzed as described under *Mass spectrometry* in *Supplemental Information*.

For additional experimental procedures, see *Supplemental Information*.

Author contributions

O.B. and I.L. designed and performed the experiments and contributed to the writing of the manuscript. T.Z. and A.A. carried out the proteomic analyses and assisted in analyzing its results. I.K. and L.C. helped in developing the adenoviral vectors and in testing their efficiency. H.G. and B.B. helped in carrying out the experiments. Y.T.K. helped with novel ideas regarding the design of experiments and the interpretation of their results. A.G., S.J., L.J. M.G. and P.T. carried out the bioinformatics structural analyses. A.C. conceptualized the project, helped in designing the experiments and in the interpretation of their results, and in writing the manuscript.

Acknowledgements

Research in the laboratory of A.C. is supported by grants from the Dr. Miriam and Sheldon G. Adelson Medical Research Foundation (AMRF), the Israel Science Foundation (ISF), the I-CORE Program of the Planning and Budgeting Committee and the ISF (Grant1775/12), and the Deutsch-Israelische Projektkooperation (DIP). I.L. is supported by the Foulkes Fellowship. We thank Drs. Zhijian James Chen (UT Southwestern) and Daniel Finley (Harvard Medical School) for providing us with the mammalian and yeast systems for deleting endogenous Ub genes, respectively. A.C. is an Israel Cancer Research Fund (ICRF) USA Professor. A.G. is supported by NIH R01 GM101457 grant. This work was supported by the Odysseus grant G.0029.12 from Research Foundation Flanders (FWO, to P.T.) and a VIB/Marie Curie COFUND Postdoctoral (omics@VIB) fellowship (to M.G.).

The authors declare that they do not have any financial conflict of interest that may be construed to the results or interpretation of the manuscript.

References

1. Glickman MH, Ciechanover A (2002) The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol Rev* 82(2):373–428.
2. Komander D (2009) The emerging complexity of protein ubiquitination. *Biochem Soc Trans* 37(Pt 5):937–953.
3. Thrower JS, Hoffman L, Rechsteiner M, Pickart CM (2000) Recognition of the polyubiquitin proteolytic signal. *EMBO J* 19(1):94–102.
4. Hicke L, Schubert HL, Hill CP (2005) Ubiquitin-binding domains. *Nat Rev Mol Cell Biol* 6(8):610–621.
5. Hoeller D, et al. (2006) Regulation of ubiquitin-binding proteins by monoubiquitination. *Nat Cell Biol* 8(2):163–169.
6. Goh LK, Sorkin A (2013) Endocytosis of receptor tyrosine kinases. *Cold Spring Harb Perspect Biol* 5(5):a017459.
7. Ahearn IM, Haigis K, Bar-Sagi D, Philips MR (2012) Regulating the regulator: post-translational modification of RAS. *Nat Rev Mol Cell Biol* 13(1):39–51.
8. Jura N, Scotto-Lavino E, Sobczyk A, Bar-Sagi D (2006) Differential modification of Ras proteins by ubiquitination. *Mol Cell* 21(5):679–687.
9. Chandrasekharan MB, Huang F, Sun ZW (2009) Ubiquitination of histone H2B regulates chromatin dynamics by enhancing nucleosome stability. *Proc Natl Acad Sci U S A* 106(39):16686–16691.
10. Rott R, et al. (2011) alpha-Synuclein fate is determined by USP9X-regulated monoubiquitination. *Proc Natl Acad Sci U S A* 108(46):18666–18671.
11. Dimova N V, et al. (2012) APC/C-mediated multiple monoubiquitylation provides an alternative degradation signal for cyclin B1. *Nat Cell Biol* 14(2):168–176.
12. Carvallo L, et al. (2010) Non-canonical Wnt signaling induces ubiquitination and degradation of Syndecan4. *J Biol Chem* 285(38):29546–29555.
13. Boutet SC, Disatnik MH, Chan LS, Iori K, Rando TA (2007) Regulation of Pax3 by proteasomal degradation of monoubiquitinated protein in skeletal muscle progenitors. *Cell* 130(2):349–362.
14. Kravtsova-Ivantsiv Y, Cohen S, Ciechanover A (2009) Modification by single ubiquitin moieties rather than polyubiquitination is sufficient for proteasomal processing of the p105 NF-kappaB precursor. *Mol Cell* 33(4):496–504.
15. Shabek N, et al. (2012) The size of the proteasomal substrate determines whether its degradation will be mediated by mono- or polyubiquitylation. *Mol Cell* 48(1):87–97.

16. Braten O, Shabek N, Kravtsova-Ivantsiv Y, Ciechanover A (2012) Generation of free ubiquitin chains is up-regulated in stress and facilitated by the HECT domain ubiquitin ligases UFD4 and HUL5. *Biochem J* 444(3):611–617.

17. Hospenthal MK, Freund SM, Komander D (2013) Assembly, analysis and architecture of atypical ubiquitin chains. *Nat Struct Mol Biol* 20(5):555–565.

18. Sun T, et al. (2011) The role of monoubiquitination in endocytic degradation of human ether-a-go-go-related gene (hERG) channels under low K+ conditions. *J Biol Chem* 286(8):6751–6759.

19. Ward CL, Omura S, Kopito RR (1995) Degradation of CFTR by the ubiquitin-proteasome pathway. *Cell* 83(1):121–127.

20. Finley D, et al. (1994) Inhibition of proteolysis and cell cycle progression in a multiubiquitination-deficient yeast mutant. *Mol Cell Biol* 14(8):5501–5509.

21. Xu M, Skaug B, Zeng W, Chen ZJ (2009) A ubiquitin replacement strategy in human cells reveals distinct mechanisms of IKK activation by TNFalpha and IL-1beta. *Mol Cell* 36(2):302–14.

22. Emanuele MJ, et al. (2011) Global identification of modular cullin-RING ligase substrates. *Cell* 147(2):459–474.

23. Mayor T, Graumann J, Bryan J, MacCoss MJ, Deshaies RJ (2007) Quantitative profiling of ubiquitylated proteins reveals proteasome substrates and the substrate repertoire influenced by the Rpn10 receptor pathway. *Mol Cell Proteomics* 6(11):1885–1895.

24. Kim W, et al. (2011) Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol Cell* 44(2):325–340.

25. Tompa P (2011) Unstructural biology coming of age. *Curr Opin Struct Biol* 21(3):419–425.

26. van der Lee R, et al. (2014) Classification of intrinsically disordered regions and proteins. *Chem Rev* 114(13):6589–6631.

27. Guharoy M, Bhowmick P, Tompa P (2016) Design principles involving protein disorder facilitate specific substrate selection and degradation by the ubiquitin-proteasome system. *J Biol Chem*. doi:jbc.R115.692665 [pii]R115.692665 [pii]10.1074/jbc.R115.692665.

28. Guharoy M, Bhowmick P, Sallam M, Tompa P (2016) Tripartite degrons confer diversity and specificity on regulated protein degradation in the ubiquitin-proteasome system. *Nat Commun* 7:10239.

29. Inobe T, Fishbain S, Prakash S, Matouschek A (2011) Defining the geometry of the two-component proteasome degron. *Nat Chem Biol* 7(3):161–167.

30. Ravid T, Hochstrasser M (2008) Diversity of degradation signals in the ubiquitin-proteasome system. *Nat Rev Mol Cell Biol* 9(9):679–690.

31. King RW, Glotzer M, Kirschner MW (1996) Mutagenic analysis of the destruction signal of mitotic cyclins and structural characterization of ubiquitinated intermediates. *Mol Biol Cell* 7(9):1343–1357.

32. Iakoucheva LM, et al. (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res* 32(3):1037–1049.

33. Dunker AK, et al. (2001) Intrinsically disordered protein. *J Mol Graph Model* 19(1):26–59.

34. Landry CR, Levy ED, Michnick SW (2009) Weak functional constraints on phosphoproteomes. *Trends Genet* 25(5):193–197.

35. Eden E, Lipson D, Yogev S, Yakhini Z (2007) Discovering motifs in ranked lists of DNA sequences. *PLoS Comput Biol* 3(3):e39.

36. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z (2009) GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10:48.

37. Amanchy R, et al. (2007) A curated compendium of phosphorylation motifs. *Nat Biotechnol* 25(3):285–286.

38. Basu A, et al. (2009) Proteome-wide prediction of acetylation substrates. *Proc Natl Acad Sci U S A* 106(33):13785–13790.

39. Chou MF, Schwartz D (2011) Biological sequence motif discovery using motif-x. *Curr Protoc Bioinforma* Chapter 13:Unit 13 15–24.

40. Hershko A, Ganoth D, Pehrson J, Palazzo RE, Cohen LH (1991) Methylated ubiquitin inhibits cyclin degradation in clam embryo extracts. *J Biol Chem* 266(25):16376–16379.

41. Amerik Ay, Swaminathan S, Krantz BA, Wilkinson KD, Hochstrasser M (1997) In vivo disassembly of free polyubiquitin chains by yeast Ubp14 modulates rates of protein degradation by the proteasome. *EMBO J* 16(16):4826–4838.

42. Lee DH, Goldberg AL (1996) Selective inhibitors of the proteasome-dependent and vacuolar pathways of protein degradation in Saccharomyces cerevisiae. *J Biol Chem* 271(44):27280–27284.

43. Udeshi ND, et al. (2013) Refined preparation and use of anti-diglycine remnant (K-epsilon-GG) antibody enables routine quantification of 10,000s of ubiquitination sites in single proteomics experiments. *Mol Cell Proteomics* 12(3):825–831.

44. Liu X, Yu X, Zack DJ, Zhu H, Qian J (2008) TiGER: a database for tissue-specific gene expression and regulation. *BMC Bioinformatics* 9:271.

45. Haglund K, et al. (2003) Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation. *Nat Cell Biol* 5(5):461–466.

46. Hicke L, Dunn R (2003) Regulation of membrane protein transport by ubiquitin and ubiquitin-

binding proteins. *Annu Rev Cell Dev Biol* 19:141–172.

47. Ikner A, Shiozaki K (2005) Yeast signaling pathways in the oxidative stress response. *Mutat Res* 569(1-2):13–27.

48. Gasch AP, et al. (2000) Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell* 11(12):4241–4257.

49. David Y, Ziv T, Admon A, Navon A (2010) The E2 ubiquitin-conjugating enzymes direct polyubiquitination to preferred lysines. *J Biol Chem* 285(12):8595–8604.

50. Rape M, Kirschner MW (2004) Autonomous regulation of the anaphase-promoting complex couples mitosis to S-phase entry. *Nature* 432(7017):588–595.

51. Ravid T, Hochstrasser M (2007) Autoregulation of an E2 enzyme by ubiquitin-chain assembly on its catalytic residue. *Nat Cell Biol* 9(4):422–427.

52. Amemiya Y, Azmi P, Seth A (2008) Autoubiquitination of BCA2 RING E3 ligase regulates its own stability and affects cell migration. *Mol Cancer Res* 6(9):1385–1396.

53. Lorick KL, et al. (1999) RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proc Natl Acad Sci U S A* 96(20):11364–11369.

54. Noels H, et al. (2009) Auto-ubiquitination-induced degradation of MALT1-API2 prevents BCL10 destabilization in t(11;18)(q21;q21)-positive MALT lymphoma. *PLoS One* 4(3):e4822.

55. Danielsen JM, et al. (2011) Mass spectrometric analysis of lysine ubiquitylation reveals promiscuity at site level. *Mol Cell Proteomics* 10(3):M110 003590.

56. Xu M, Skaug B, Zeng W, Chen ZJ (2009) A ubiquitin replacement strategy in human cells reveals distinct mechanisms of IKK activation by TNFalpha and IL-1beta. *Mol Cell* 36(2):302–314.

57. Spence J, Sadis S, Haas AL, Finley D (1995) A ubiquitin mutant with specific defects in DNA repair and multiubiquitination. *Mol Cell Biol* 15(3):1265–1273.

Figure Legends

Figure 1. Replacement of endogenous Ub by Ub^{K0} in yeast and mammalian cells

(A) Workflow describing Ub silencing and re-expression (see a detailed description under *Experimental Procedures*). (B) Ub replacement in yeast cells. ΔUb, ΔUb^{UbWT} and ΔUb^{UbK0} yeast cells were treated for Ub silencing and Ub re-expression as indicated. Yeast cells were analyzed by trichloroacetic acid (TCA) lysis followed by SDS-PAGE and WB using the indicated antibodies. (C) Ub replacement in human cells. <i> Ub silencing. To silence Ub, $U2OS^{shUb}$ cells were treated with tetracycline (1 μg/ml) for the indicated times. <ii> Ub re-expression. Following Ub silencing, cells were infected with viral vectors expressing Ub^{WT} or Ub^{K0}. In both panels, lysates were analyzed via SDS-PAGE followed by WB using the indicated antibodies.

Figure 2. Identification of monoubiquitination- and polyubiquitination-dependent proteasome substrates

(A) Experimental workflow. See a detailed description under *Results*. (B) Ubiquitination sites and ubiquitinated proteins identified in independent experiments. See Table S1 for experiment names. (C) The algorithm used for classification to monoubiquitination- and polyubiquitination-dependent substrates.

Figure 3. Validation of representative candidate substrates degraded by the proteasome following mono- and polyubiquitination

(A) Upper panels: MCF7 cells were co-transfected with plasmids coding for GOT1-HA or CDC20-HA along with plasmids coding for Ub^{WT} or Ub^{K0} as indicated. Lower panels: ΔUb^{UbWT} and ΔUb^{UbK0} yeast cells were transformed with plasmids coding for GRE1-HA or ARD1-HA, and Ub replacement was carried out as described under Figure 1. In all experiments, substrate stability was monitored as described under *Experimental Procedures*. (B) Upper panels: MCF7 cells were transfected with either GOT1-HA or CDC20-HA, followed by the treatment with epoxomicin (2 μM, 24 hr) as indicated. Lower panels: ΔPDR5 yeast cells (strain Y12409 from the EUROSCARF collection) were transformed with plasmids coding for either ARD1-HA or GRE1-HA. Strains were treated with cycloheximide (CHX) and bortezomib (100 μM each) as indicated. Samples were collected at the indicated time points. Cell lysates were analyzed via SDS-PAGE followed by WB using the indicated antibodies.

Figure 4. Yeast and human substrates demonstrate distinct patterns of structural disorder

(A) Human polyubiquitinated substrates are more disordered than monoubiquitinated substrates. The distribution of overall disorder content in mono- and polyubiquitinated substrates is plotted for yeast and human. Overall disorder content is the fraction of predicted disordered residues in a given protein sequence. The distribution for the reference proteome corresponding to each species is shown for comparison. Dotted vertical lines correspond to the average values for the distribution. (B) Local

disorder at the ubiquitination sites is more prominent in yeast monoubiquitination substrates and in human polyubiquitination substrates. The distribution of local disorder score is plotted for mono- and polyubiquitinated sites from yeast and human. The disorder profile for each sequence is first calculated using IUPred and then the average disorder score of a 21-residue sequence window centered on each Ubsite is calculated. The distribution for the reference proteome corresponding to each species is shown for comparison. Dotted vertical lines correspond to the average values for the distribution. (C) The distance of Ubsites from their nearest LDR (see *Experimental Procedures* for definition) is plotted for yeast and human sites. The averages of the distributions are not shown here for the sake of clarity.

Figure 5. Analysis of molecular functions and involvement in different cellular processes for mono- and polyubiquitination-dependent substrates.

Monoubiquitination- and polyubiquitination-dependent substrate groups were analyzed for Gene Ontology (GO) term enrichment using the GOrilla tool, as described under *Experimental Procedures*. (A) Biological Process GO term analysis for yeast substrates. (B) Molecular Function GO term analysis for human substrates.

Figures

Figure 1

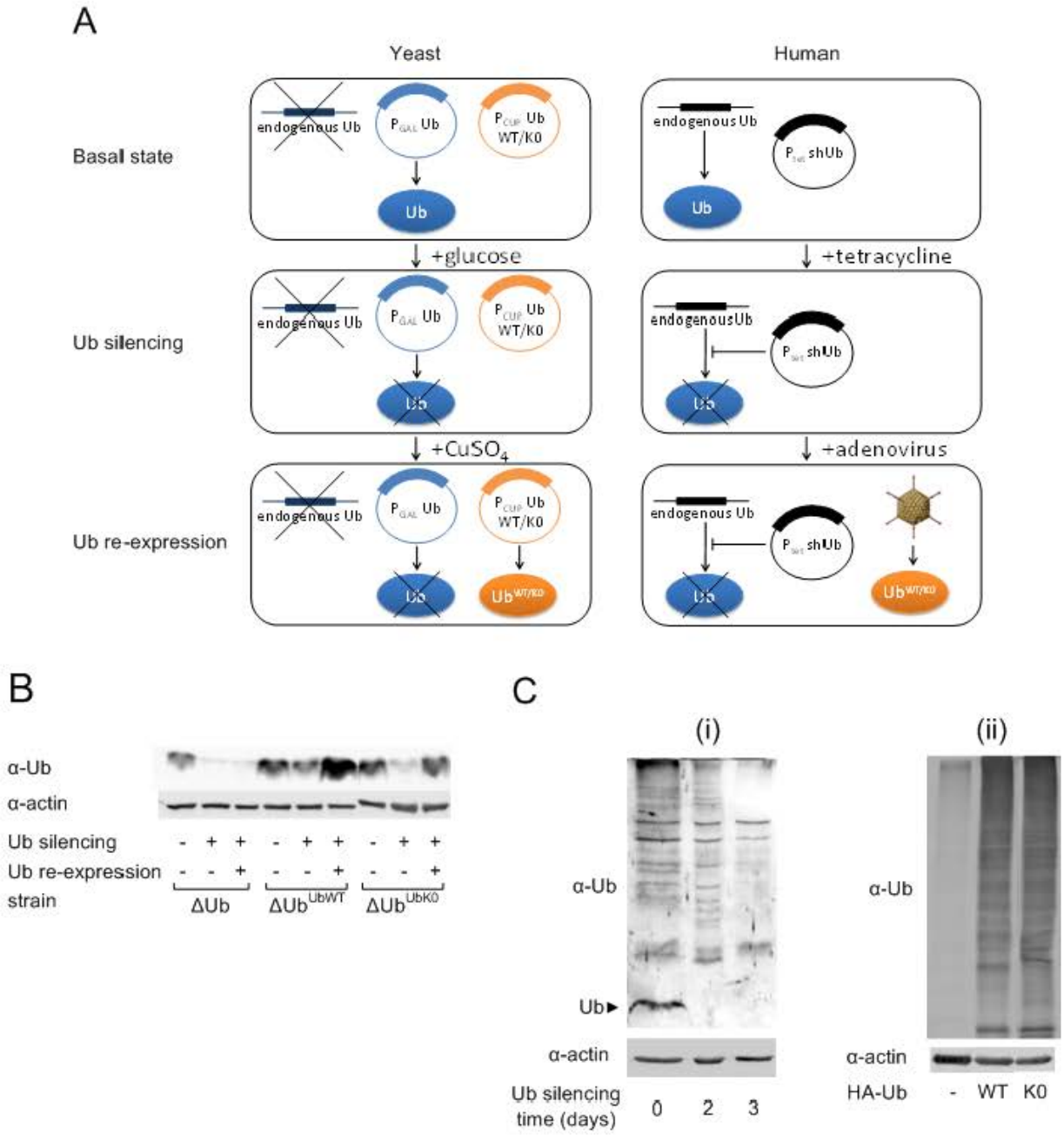


Figure 2

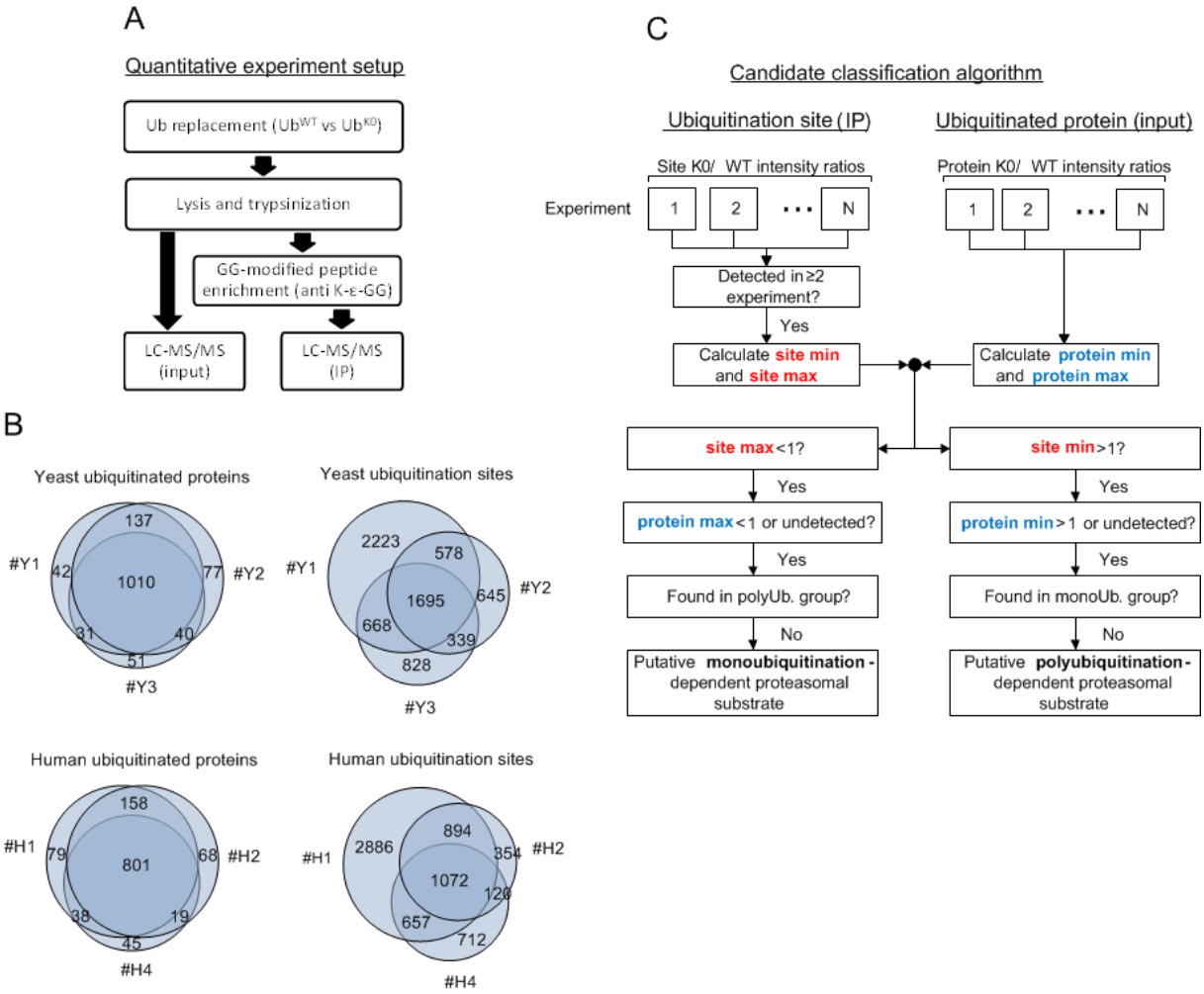

A
Quantitative experiment setup
Ub replacement (Ub^WT vs Ub^K0)
Lysis and trypsinization
GG-modified peptide enrichment (anti K-ε-GG)
LC-MS/MS (input)
LC-MS/MS (IP)
B
Yeast ubiquitinated proteins
#Y1
#Y2
#Y3
137
42
1010
77
31
40
51
Yeast ubiquitination sites
2223
578
1695
645
668
339
828
Human ubiquitinated proteins
#H1
#H2
#H4
158
79
801
68
38
19
45
Human ubiquitination sites
2886
894
354
1072
120
657
712
C
Candidate classification algorithm
Ubiquitination site (IP)
Ubiquitinated protein (input)
Site K0/ WT intensity ratios
Protein K0/ WT intensity ratios
Experiment
1
2
N
Detected in ≥2 experiment?
Yes
Calculate site min and site max
Calculate protein min and protein max
site max<1?
site min>1?
protein max<1 or undetected?
protein min>1 or undetected?
Found in polyUb. group?
Found in monoUb. group?
No
Putative monoubiquitination-dependent proteasomal substrate
Putative polyubiquitination-dependent proteasomal substrate

Figure 3

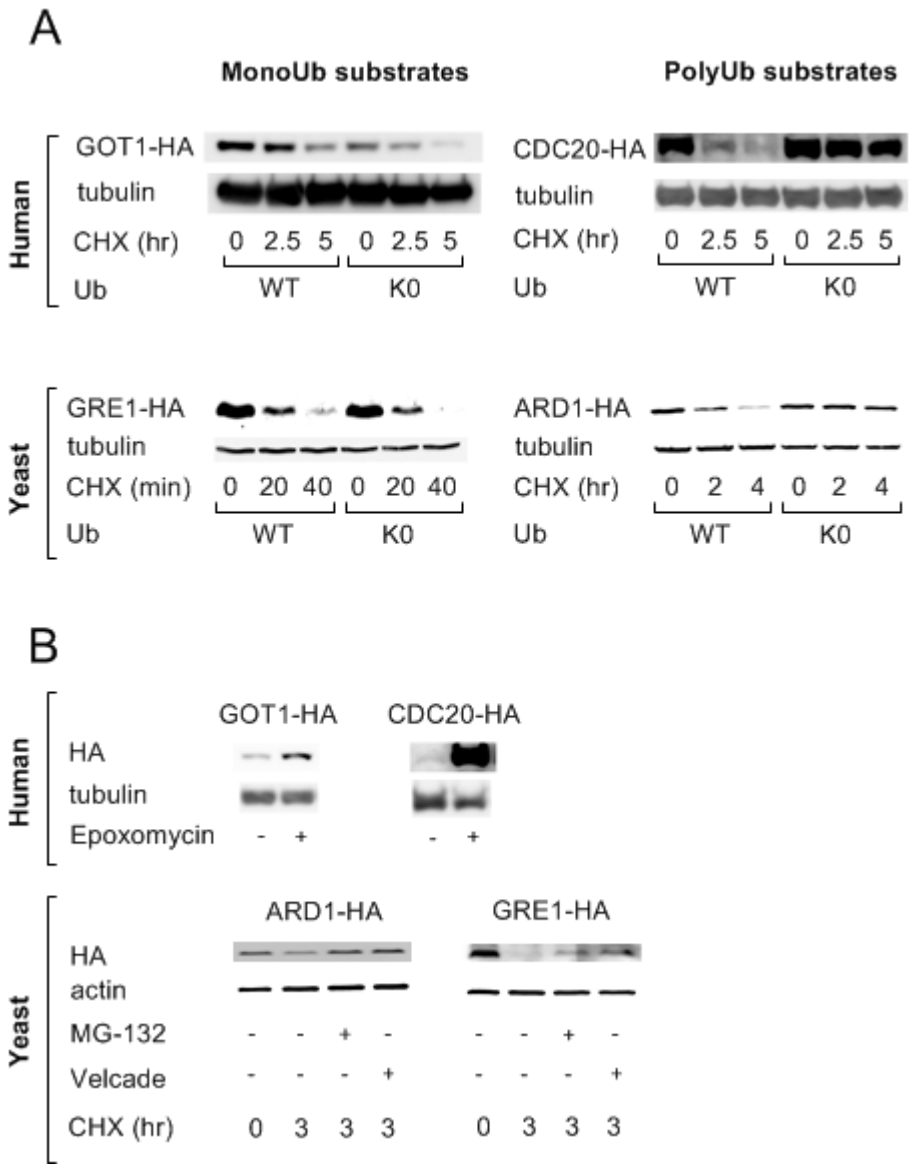

A
MonoUb substrates
PolyUb substrates
Human
GOT1-HA
tubulin
CHX (hr)
0 2.5 5 0 2.5 5
Ub
WT
K0
CDC20-HA
tubulin
CHX (hr)
0 2.5 5 0 2.5 5
Ub
WT
K0
Yeast
GRE1-HA
tubulin
CHX (min)
0 20 40 0 20 40
Ub
WT
K0
ARD1-HA
tubulin
CHX (hr)
0 2 4 0 2 4
Ub
WT
K0
B
Human
GOT1-HA
CDC20-HA
HA
tubulin
Epoxomycin
- + - +
Yeast
ARD1-HA
GRE1-HA
HA
actin
MG-132
- - + - - - + -
Velcade
- - - + - - - +
CHX (hr)
0 3 3 3 0 3 3 3

Figure 4

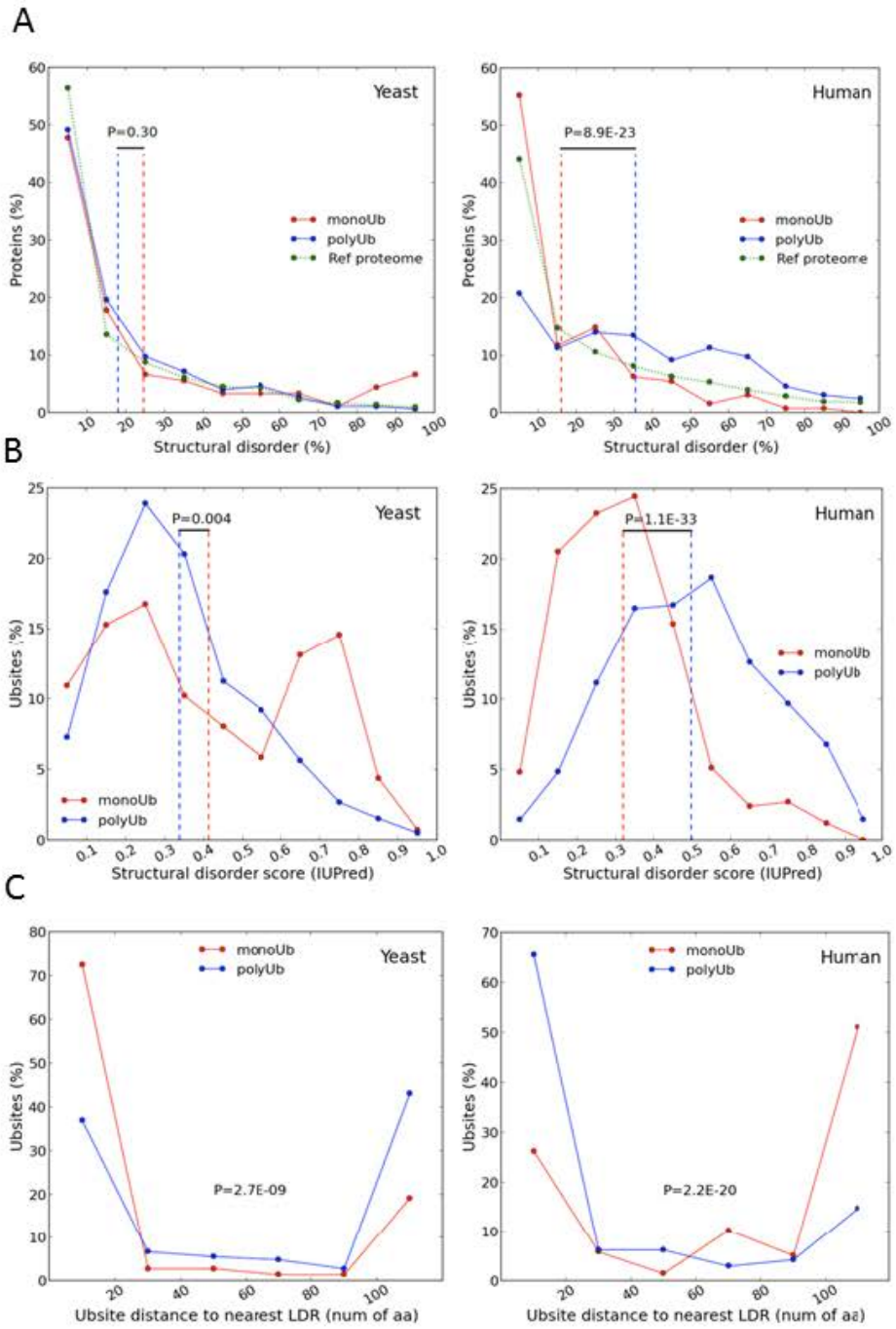
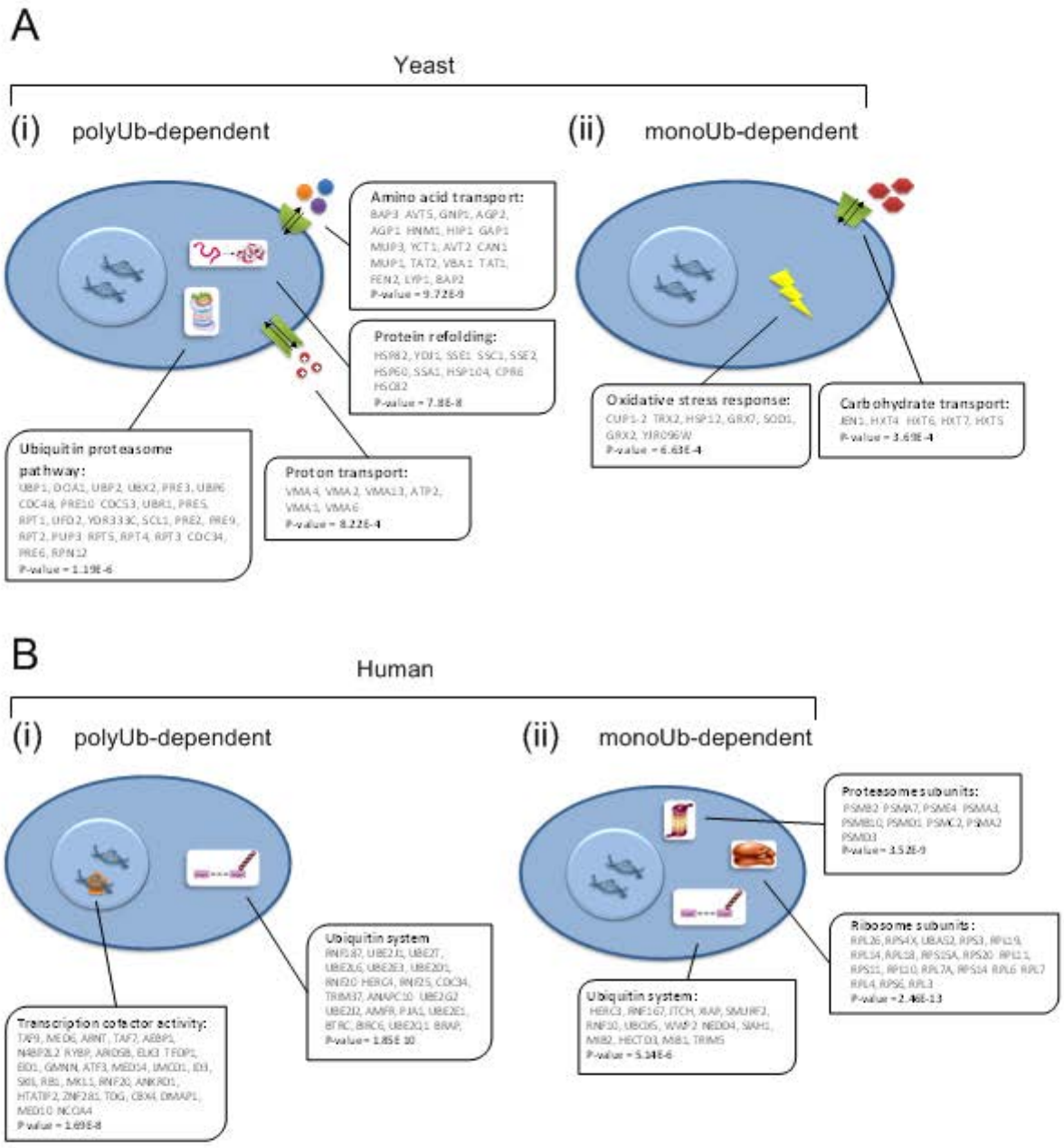

A
Yeast
P=0.30
monoUb
polyUb
Ref proteome
Proteins (%)
Structural disorder (%)
Human
P=8.9E-23
B
Yeast
P=0.004
Ubsites (%)
Structural disorder score (IUPred)
Human
P=1.1E-33
C
Yeast
P=2.7E-09
Ubsite distance to nearest LDR (num of aa)
Human
P=2.2E-20

Figure 5


A
Yeast
(i) polyUb-dependent
(ii) monoUb-dependent
Amino acid transport:
BAP3, AVT5, GNP1, AGP2, AGP1, HNM1, HIP1, GAP1, MUP3, YCT1, AVT2, CAN1, MUP1, TAT2, VBA1, TAT1, FEN2, LYP1, BAP2
P-value = 9.72E-9
Protein refolding:
HSP82, YDJ1, SSE1, SSC1, SSE2, HSP60, SSA1, HSP104, CPR6, HSC82
P-value = 7.8E-8
Ubiquitin proteasome pathway:
UBP1, DOA1, UBP2, UBX2, PRE3, UBP6, CDC48, PRE10, CDC53, UBR1, PRE5, RPT1, UFD2, YDR333C, SCL1, PRE2, PRE9, RPT2, PUP3, RPT5, RPT4, RPT3, CDC34, PRE6, RPN12
P-value = 1.19E-6
Proton transport:
VMA4, VMA2, VMA13, ATP2, VMA5, VMA6
P-value = 8.22E-4
Oxidative stress response:
CUP1-2, TRX2, HSP12, GRX7, SOD1, GRX2, YIR096W
P-value = 6.63E-4
Carbohydrate transport:
JEN1, HXT4, HXT6, HXT7, HXT5
P-value = 3.69E-4
B
Human
(i) polyUb-dependent
(ii) monoUb-dependent
Ubiquitin system
RNF187, UBE2J1, UBE2T, UBE2L6, UBE2E3, UBE2D1, RNF10, HERC4, RNF25, CDC34, TRIM37, ANAPC10, UBE2G2, UBE2J2, AMFR, PJA1, UBE2E1, BTRC, BIRC6, UBE2C11, BIRA
P-value = 1.85E-10
Transcription cofactor activity:
TAF9, MED6, ARNT, TAF7, ARBP5, N4BP2L2, RYBP, ARID5B, ELK3, TFDP1, E2D1, GMNN, ATF3, MED14, LMCD1, ID3, SKI, RB1, MKL1, RNF20, ANKRD1, HTATIP2, ZNF281, TDG, CBX4, DMAP1, MED10, NCOA4
P-value = 1.69E-8
Proteasome subunits:
PSMB2, PSMA7, PSME4, PSMA3, PSMB10, PSMD1, PSMC2, PSMA2, PSMD3
P-value = 3.52E-9
Ribosome subunits:
RPL26, RPS4X, UBA52, RPS3, RPL19, RPL34, RPL18, RPS25A, RPS20, RPL11, RPS11, RPL10, RPL7A, RPS14, RPL6, RPL7, RPL4, RPS6, RPL3
P-value = 2.46E-13
Ubiquitin system:
HERC3, RNF167, ITCH, XIAP, SMURF2, RNF10, UBOX5, WWP2, NEDD4, SIAH1, MIB2, HECTD3, MIB1, TRIM5
P-value = 5.34E-6

Supplemental Figures and Legends

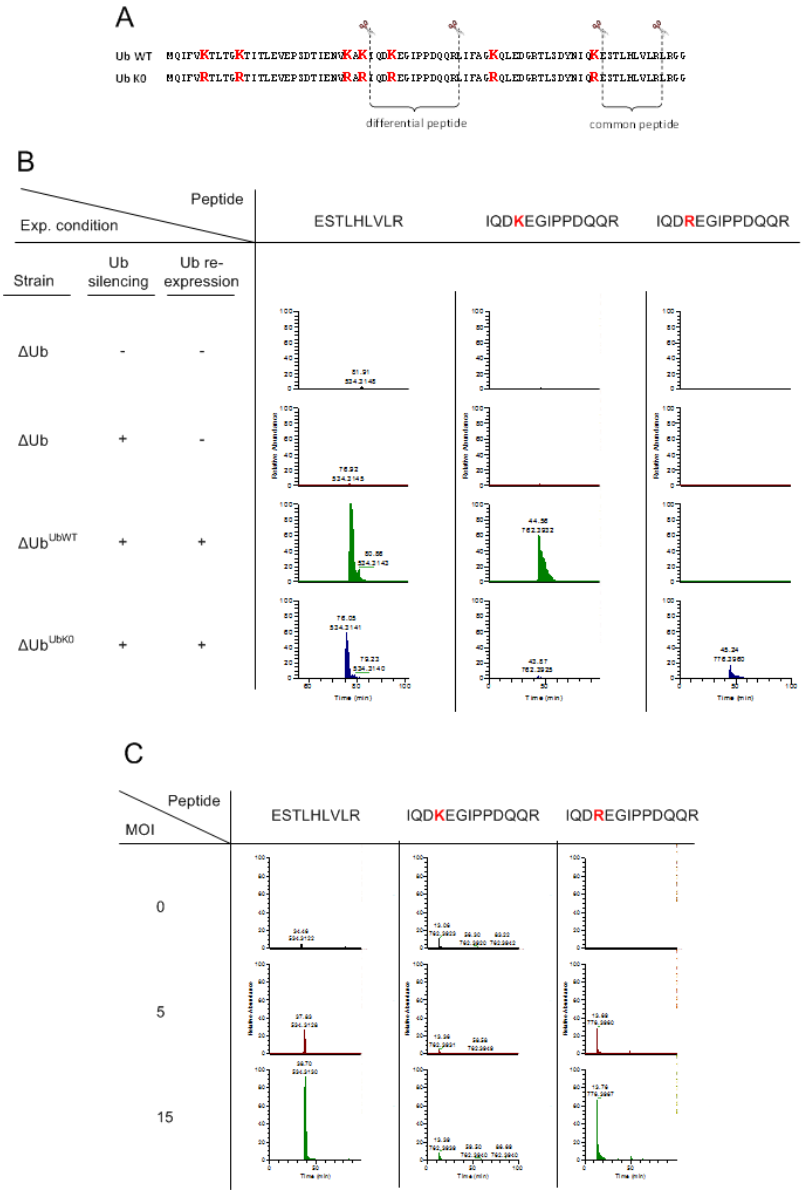


Figure S1. Mass spectrometric verification of replacement of endogenous Ub by Ub^{K0} in yeast and mammalian cells.

(A) Ub^{wt}- and Ub^{K0}-derived tryptic peptides used for mass spectrometric quantification of Ub^{K0} expression. Human Ub sequence is shown. The indicated differential and common peptides correspond also to the yeast Ub. (B) ΔUb, ΔUb^{UbWT} and ΔUb^{UbK0} yeast cells were treated for Ub replacement as indicated. In all experiments, total Ub abundance (K0 and WT species) was assessed based on the relative MS intensity of the common peptide and Ub^{WT} and Ub^{K0} abundances were assessed based on the relevant differential peptide. Ub replacement and MS analysis were performed as described under *Experimental Procedures*. MS peaks representing the indicated peptides are shown with values representing column retention times (upper value; in minutes) and m/z ratio (lower value). (C) U2OS^{shUb} cells were either left untreated or treated for Ub replacement with Ub^{K0} using the indicated MOI. Data were processed as described under (B).

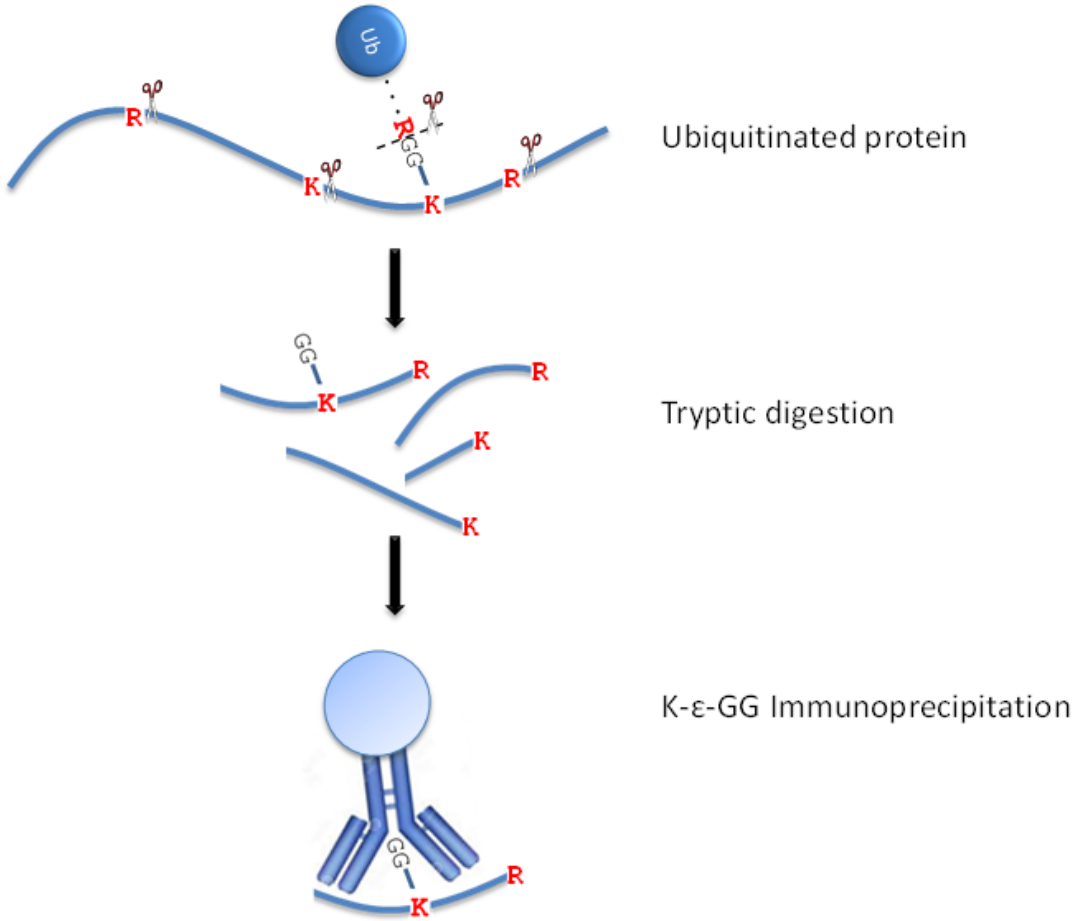


Figure S2. Ubiquitination sites enrichment.

Schematic illustration of isolation of GG-modified peptides using anti-K-ε-GG immunoprecipitation. For detailed description, see under *Experimental Procedures*.

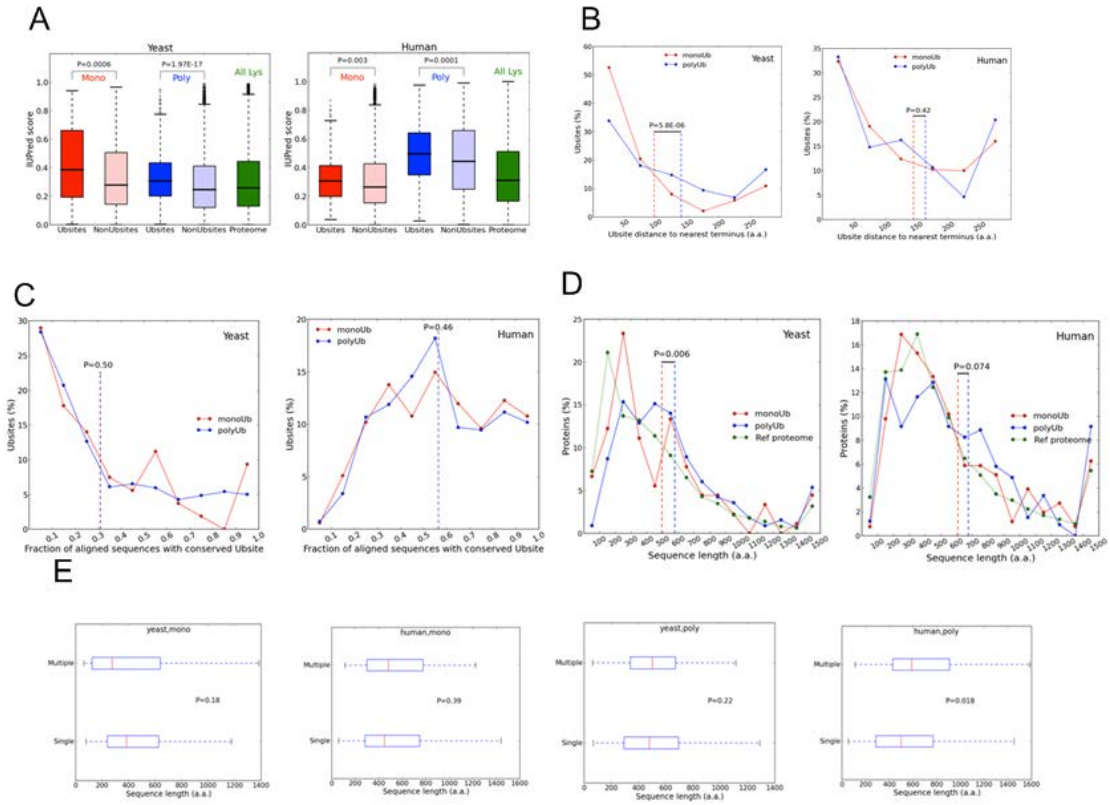


Figure S3. Analyses of different characteristics of ubiquitinated substrates.

(A) Boxplots showing the local disorder scores for Ubsites, NonUbsites and all lysines in the reference proteome. Ubsites correspond to the identified ubiquitination sites in each protein and NonUbsites are the remaining lysines from the proteolytic substrates. The Y-axis corresponds to the average IUPred score of 21-residue sequence windows flanking each lysine. (B) Plot showing the distance of yeast (left) and human (right) Ubsites from their nearer protein chain terminus. Dotted lines represent the average values of each distribution. (C) Plot showing Ubsite residue conservation across orthologous proteins (see *Experimental Procedures* for details). On the X-axis, 0.0 corresponds to a completely variable Ubsite, whereas 1.0 indicates a fully conserved Ubsite. Dotted lines represent the average values of each distribution. (D) Length of substrates that undergo monoubiquitination- and polyubiquitination dependent degradation. The plot is showing the length distribution of mono- and polyubiquitinated substrates from yeast (left) and human (right). The reference proteome corresponding to each species is shown for comparison. Dotted vertical lines correspond to the average values for the distribution. (E) Boxplots of the distribution of sequence lengths for proteins containing a single or multiple Ubsites. The plots show data from yeast and human cells and from mono- and polyubiquitinated proteins. Outliers are not shown.

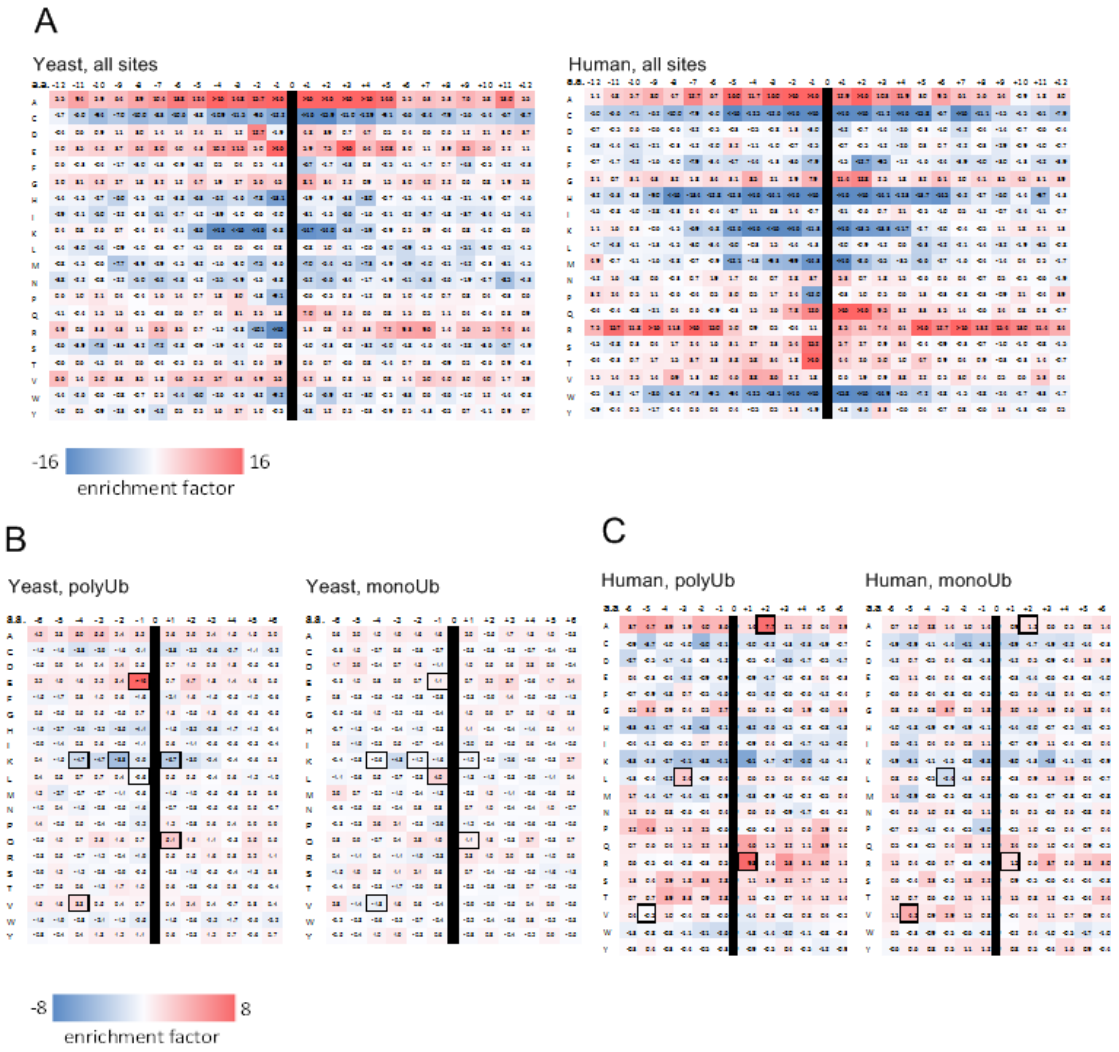


Figure S4. Analysis of amino acids occurrences around ubiquitination sites.

Analysis of sequence windows with ubiquitinated lysines in their center. Residue- and position-specific enrichment factors (EF) were calculated as described under *Experimental Procedures*. (A) All identified ubiquitination sites in yeast (left panel) and in human (right panel). (B) Yeast ubiquitination sites of polyubiquitination- (left panel), and monoubiquitination- (right panel) dependent substrates. The highlighted coordinates represent major pattern differences between mono- and polyubiquitinated substrates in which: $\left|EF_{poly} - EF_{mono}\right| > 4$. (C) Same as in (B) but for human substrates.

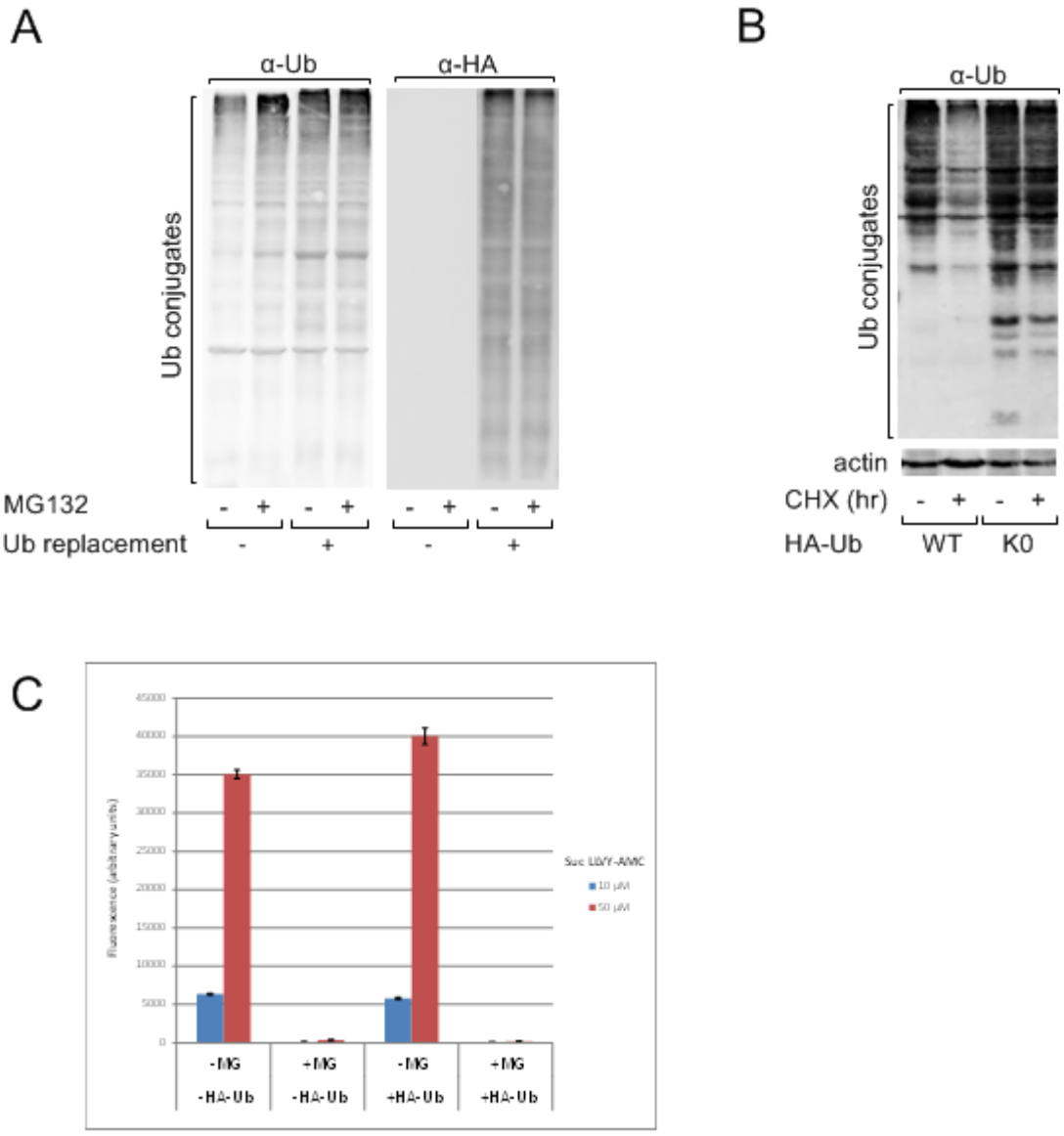


Figure S5. Measurement of ubiquitin-proteasome system activity following Ub replacement in U2OS^{shUb} cells.

(A) Where indicated, endogenous Ub was replaced by HA-Ub^{WT} in U2OS^{shUb} cells as described under *Experimental Procedures*. Cells were further treated as indicated with MG132 at 20 μM for 4 hr. Lysates were analyzed by SDS-PAGE followed by WB using the indicated antibodies. (B) Endogenous Ub was replaced by HA-Ub^{WT} in U2OS^{shUb} cells. The stability of Ub-conjugates was monitored following addition of cycloheximide for 6 hr as described under *Experimental Procedures*. (C) U2OS^{shUb} cells were treated for Ub replacement by HA-Ub^{WT} and for proteasome inhibition with MG132 as indicated. Proteasome activity was measured by a fluorescence-based assay as described under *Experimental Procedures*.

Table S1. Quantitative experiments for the identification of monoubiquitination-dependent proteasomal substrates

# Experiment	Model organism	Experimental setup	Number of identified peptides
Y1	yeast	label free	5164
Y2	yeast	label free	3257
Y3	yeast	label free	3530
H1	human	label free	5509
H2	human	label free	2440
H3	human	SILAC	1156
H4	human	label free	2561

Supplemental Experimental Procedures

Materials

Anti-Ub was described previously (1), anti-HA (16B12) was from Covance, and anti-actin and anti-tubulin were from Millipore. Anti K-ε-GG Immunoaffinity Beads were from Cell Signaling Technology. Suc-Leu-Leu-Val-Tyr-AMC (Suc-LLVY-AMC) was from Boston Biochem. Materials for SDS-PAGE and Bradford reagent were from Bio-Rad. Tissue culture sera, media, and supplements were from Biological Industries (Bet HaEmek, Israel). Peroxidase-conjugated secondary antibodies were from Jackson ImmunoResearch Laboratories. N-carbobenzoxy-L-leucyl-L-leucyl-leucinal (MG132) and epoxomicin were from Calbiochem. Reagents for enhanced chemiluminescence (ECL) were from Pierce. JetPEI^{TM} cell transfection reagent was from Polyplus. Restriction and modifying enzymes were from New England Biolabs. Oligonucleotides were synthesized by Sigma. All other reagents were of high analytical grade.

Plasmids

All of the plasmids were constructed and manipulated using standard molecular biological techniques. For liposome-mediated transient cell transfection, Flag-Ub^{WT} and Flag-Ub^{K6,11,27,29,33,48,63R} (Flag-Ub^{K0}) were subcloned into pCAGGS (2), and GOT1-HA and CDC20-HA were subcloned into pCS2. For adenovirus preparation, HA-Ub^{WT} and HA-Ub^{K0} were subcloned into pAd/CMV/V5-DEST. For yeast transformation, Ub^{WT} and Ub^{K0} were subcloned into pUb39 (3), and ARD1-HA and GRE1-HA were subcloned into pCM190.

Cultured cells

U2OS^{shUb} cells (4) were grown at 37^{o}C in DMEM supplemented with 10% fetal calf serum and antibiotics (penicillin–streptomycin). MCF7 cells were grown 37^{o}C in the same medium supplemented with insulin (100 units/ml).

Protein concentration measurement

Protein concentration was measured by the Bradford method (5) using BSA as the standard.

Protein detection

For Western blotting (WB), proteins were resolved via SDS-PAGE, transferred to nitrocellulose (or PVDF) membrane, and incubated with the appropriate antibodies.

Yeast transformation

The different strains of *Saccharomyces cerevisiae* were transformed with plasmids coding for the proteins of interest using the PEG/LiAc [poly(ethylene glycol)/lithium acetate] method (6), followed by appropriate selection.

Mammalian cell transfection

Transient transfections of mammalian cells with plasmids containing cDNAs coding for proteins of interest were carried out using the jetPEI™ transfection reagent according to manufacturer's instructions.

Monitoring protein stability

Protein synthesis was inhibited by adding 100 µg/ml cycloheximide. Samples were collected at the indicated times, and proteins of interest analyzed by SDS-PAGE followed by WB using the indicated antibodies.

Preparation of adenoviral vectors

HEK293 cells were transfected with pAd/CMV/V5-DEST (AdEasy™ Vector System) encoding the genes of interest. Small and large scale virus amplifications were performed according to the manufacturer's instructions, followed by purification of the viral particles on iodixanol gradient as described previously (7). Viral titer was determined by MOI test according to the manufacturer's instructions.

Fluorescence-based proteasome activity assay

Cells were lysed in 0.3% CHAPS buffer (20 mM HEPES, 100 mM NaCl, 1 mM EDTA, 1.5 mM $MgCl_2$, 0.3% CHAPS, 1 mM *DTT*, 2.5 mM ATP, and protease inhibitor cocktail), and proteasomes were immobilized onto agarose to which anti-α6 was bound. Beads to which the proteasomes were not immobilized were used as control. Beads were washed twice with 0.03% CHAPS buffer. Where indicated, MG132 (100 µM) was added to the immobilized proteasomes and the mixture was incubated at 37°C for 10 min. Next, proteasomes were incubated at 37°C for 30 min with the indicated concentrations of Suc-LLVY-AMC in a reaction buffer (40 mM Tris-HCl pH 7.2, 2 mM DTT, 5 mM $MgCl_2$, 10 mM creatine phosphate, 0.1 mg/ml creatine phosphate kinase, 5 mM ATP). Reaction was stopped by adding 1% SDS, and fluorescence was measured at Ex. 360nm/Em. 460nm.

Cell lysis

Yeast cell lysates for WB were prepared by Trichloroacetic acid (TCA) precipitation as described (8). Yeast lysates for MS analysis and anti-K-ε-GG IP were prepared by suspending cell pellets in urea buffer (8 M urea, 100 mM Tris-HCl pH 8.0, 10 mM iodoacetamide), followed by shaking with 0.5 mm glass beads for 25 min at room temperature. The resulting lysates were cleared by centrifugation. Human cells were lysed by adding urea buffer, followed by brief sonication and clearing by centrifugation.

Sample preparation for mass spectrometry

Cells were harvested and lysed as described under *Cell lysis*. 2-3 mg of protein in 8 M Urea and 100 mM ammonium bicarbonate, were incubated with DTT (2.8 mM; 30 min at 60°C), modified with iodoacetamide (8.8 mM; 30 min at room temperature in the dark), and digested (overnight at 37°C)

with modified trypsin (Promega; 1:50 enzyme-to-substrate ratio) in 2 M Urea and 25 mM ammonium bicarbonate. Additional second trypsinization was carried out for 4 hours. The tryptic peptides were desalted using Sep-Pak C18 (Waters) and dried. 10 µg of the total protein were used for proteome analysis as described under *Mass spectrometry*, and the remaining material was used for ubiquitination sites analysis as described under *Ubiquitination sites detection using GG-modified peptide enrichment*.

Mass spectrometry

Tryptic peptides were analyzed by LC-MS/MS using a Q Exactive plus mass spectrometer (Thermo Fisher Scientific) fitted with a capillary HPLC (easy nLC 1000, Thermo). The peptides were loaded onto a C18 trap column (0.3 x 5 mm, LC-Packings) connected on-line to a home-made capillary column (20 cm, internal diameter 75 micron) packed with Reprosil C18-Aqua (Dr. Maisch GmbH, Germany) in solvent A (0.1% formic acid in water). The peptides mixture was resolved with a 5-28% linear gradient of solvent B (95% acetonitrile with 0.1% formic acid) for 180 min followed by a 5 min gradient of 28-95% and 25 min at 95% acetonitrile with 0.1% formic acid in water at a flow rate of 0.15 µl/min. Mass spectrometry was performed in a positive mode (m/z 350–1800, resolution 70,000) using repetitively full MS scan followed by collision-induced dissociation (HCD at 35 normalized collision energy) of the 10 most dominant ions (>1 charges) selected from the first MS scan. A dynamic exclusion list was enabled with exclusion duration of 20 sec.

Data analysis

The mass spectrometry raw data were analyzed by the MaxQuant software (version 1.4.1.2, http://www.maxquant.org) for peak picking and quantification. This was followed by identification of the proteins using the Andromeda engine, searching against the human or the yeast UniProt database with mass tolerance of 20 ppm for the precursor masses and for the fragment ions. Met oxidation, N-terminal acetylation, N-ethylmaleimide and carbamidomethyl on Cys, GlyGly on Lys, and phosphorylation on Ser, Thr and Tyr residues, were set as variable post-translational modifications. Minimal peptide length was set to six amino acids and a maximum of two mis-cleavages was allowed. Peptide- and protein-level false discovery rates (FDRs) were filtered to 1% using the target-decoy strategy. Protein tables were filtered to eliminate the identifications from the reverse database and from common contaminants. The MaxQuant software was used for label-free semi-quantitative analysis [based on extracted ion currents (XICs) of peptides], enabling quantification from each LC/MS run for each peptide identified in any of the experiments. In samples that were SILAC-labeled, quantification was also done using the MaxQuant 1.4.1.2 software. Data merging and statistical tests were done by the Perseus 1.4 software.

Gene Ontology term enrichment analysis (GOrilla)

GO term analysis was performed by the Gorilla web tool (9, 10), using the *two unranked lists* running mode. We used the monoubiquitination- or polyubiquitination-dependent proteasome substrate groups as target lists, and the complete proteome as a background list. Both yeast and human data were analyzed.

Statistical analysis

To compare protein sequence lengths of monoubiquitination- and polyubiquitination-dependent proteasomal substrates, length distribution histograms were generated. Statistical significance of differences was determined by the Mann-Whitney U-test. To analyze ubiquitination sites amino acid composition, peptides containing GG-modified lysines in their center were aligned, and residue- and position-specific amino acid occurrences (OC) were calculated. As a control, the corresponding proteome's amino acid occurrences in lysine-centered peptides were estimated ($OC_{control}$), using lysines from 1,000 random proteins. The statistical significance of enrichment or depletion for each amino acid in every position was determined by the binomial test, and the enrichment factor (EF) was calculated as following:

$$EF = \begin{cases} -\log(P_{value}); & \frac{OC}{OC_{control}} > 1 \\ \log(P_{value}); & \frac{OC}{OC_{control}} < 1 \end{cases}$$

Thus, increasing positive EF values represent a more significant enrichment, and increasing negative EF values represent a more significant depletion.

Prediction of structural disorder

IUPred (11) software was used to predict structural disorder using amino acid sequences as input. IUPred output scores were determined for each residue (ranging between 0 – 1); scores greater than 0.5 indicate disordered residues (the 'long' default mode of IUPred was used). Using the IUPred scores, long disordered regions (LDRs) were defined as consecutive stretches of at least 30 disordered residues (breaks of up to 3 consecutive ordered residues within an LDR were permitted). Statistical tests for calculating P values were carried out using the Mann-Whitney U test.

Ortholog alignments and calculation of ubiquitination site conservation

Pre-computed multiple sequence alignments of orthologs were obtained from Discovery@Bioware (http://bioware.ucd.ie/~compass/biowareweb/) and used to calculate the fraction of aligned sequences in which lysines were retained in orthologous sequences for identified Ubsites. Statistical tests for calculating P values were carried out using the Mann-Whitney U test.

Supplemental References

1. Hershko A, Eytan E, Ciechanover A, Haas AL (1982) Immunochemical analysis of the turnover of ubiquitin-protein conjugates in intact cells. Relationship to the breakdown of abnormal proteins. *J Biol Chem* 257(23):13964–13970.

2. Niwa H, Yamamura K, Miyazaki J (1991) Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* 108(2):193–199.

3. Finley D, et al. (1994) Inhibition of proteolysis and cell cycle progression in a multiubiquitination-deficient yeast mutant. *Mol Cell Biol* 14(8):5501–5509.

4. Xu M, Skaug B, Zeng W, Chen ZJ (2009) A ubiquitin replacement strategy in human cells reveals distinct mechanisms of IKK activation by TNFalpha and IL-1beta. *Mol Cell* 36(2):302–314.

5. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254.

6. Gietz RD, Woods RA (2006) Yeast transformation by the LiAc/SS Carrier DNA/PEG method. *Methods Mol Biol* 313:107–120.

7. Dormond E, et al. (2010) An efficient process for the purification of helper-dependent adenoviral vector and removal of helper virus by iodixanol ultracentrifugation. *J Virol Methods* 165(1):83–89.

8. Kravtsova-Ivantsiv Y, Cohen S, Ciechanover A (2009) Modification by single ubiquitin moieties rather than polyubiquitination is sufficient for proteasomal processing of the p105 NF-kappaB precursor. *Mol Cell* 33(4):496–504.

9. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z (2009) GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10:48.

10. Eden E, Lipson D, Yogev S, Yakhini Z (2007) Discovering motifs in ranked lists of DNA sequences. *PLoS Comput Biol* 3(3):e39.

11. Dosztanyi Z, Csizmok V, Tompa P, Simon I (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 21(16):3433–3434.