PROTEOMICS

# A systems-wide screen identifies substrates of the $SCF^{\beta TrCP}$ ubiquitin ligase

Teck Yew Low,[1,2] Mao Peng,[1,2]* Roberto Magliozzi,[3]* Shabaz Mohammed,[1,2]† Daniele Guardavaccaro,[3] Albert J. R. Heck[1,2]‡

**Cellular proteins are degraded by the ubiquitin-proteasome system (UPS) in a precise and timely fashion. Such precision is conferred by the high substrate specificity of ubiquitin ligases. Identification of substrates of ubiquitin ligases is crucial not only to unravel the molecular mechanisms by which the UPS controls protein degradation but also for drug discovery purposes because many established UPS substrates are implicated in disease. We developed a combined bioinformatics and affinity purification–mass spectrometry (AP-MS) workflow for the system-wide identification of substrates of $SCF^{\beta TrCP}$, a member of the SCF family of ubiquitin ligases. These ubiquitin ligases are characterized by a multisubunit architecture typically consisting of the invariable subunits Rbx1, Cul1, and Skp1, and one of 69 F-box proteins. The F-box protein of this member of the family is βTrCP. $SCF^{\beta TrCP}$ binds, through the WD40 repeats of βTrCP, to the DpSGXX(X)pS diphosphorylated motif in its substrates. We recovered 27 previously reported $SCF^{\beta TrCP}$ substrates, of which 22 were verified by two independent statistical protocols, thereby confirming the reliability of this approach. In addition to known substrates, we identified 221 proteins that contained the DpSGXX(X)pS motif and also interacted specifically with the WD40 repeats of βTrCP. Thus, with $SCF^{\beta TrCP}$, as the example, we showed that integration of structural information, AP-MS, and degron motif mining constitutes an effective method to screen for substrates of ubiquitin ligases.**

## INTRODUCTION

Relative to most other protein posttranslational modifications (PTMs), ubiquitylation is unique because ubiquitin can be conjugated to a substrate protein in various chain linkages through which the modified substrates are sorted to different cellular fates (*1*). Substrates polyubiquitylated with $Lys^{48}$ linkages are recognized and degraded by the 26*S* proteasome (*2*). Target substrates are modified by the sequential actions of a ubiquitin-activating enzyme (E1), a ubiquitin-conjugating enzyme (E2), and a ubiquitin ligase (E3). The huge repertoire of more than 700 E3 ubiquitin ligases in mammals provides the specificity required for the recognition of diverse substrates (*3*, *4*). Because numerous substrates targeted by the ubiquitin-proteasome system (UPS) are oncoproteins or tumor suppressors, which are potential targets for cancer therapy (*5*, *6*), it is crucial to develop high-throughput assays to identify the complete repertoire of substrates of specific E3 ubiquitin ligases.

Because E3 ubiquitin ligases, as well as other modifying enzymes, are believed to interact transiently with their substrates in a "kiss-and-run" manner (*7*), researchers tend to assay the products of the enzymatic activity rather than the enzyme-substrate interactions. To this end, quantitative proteomics methods are widely used because they allow the identification and quantitation of protein substrates with relatively high throughput. For example, an E3 ubiquitin ligase can be perturbed by RNAi (RNA interference), chemical inhibitors, or genetic manipulations, whereafter the perturbed samples and controls may be analyzed by quantitative mass spectrometry (MS) for any change in protein abundance or the amount of PTM-modified proteins (*8*–*10*). The availability of an antibody that recognizes diglycine residues, derived from ubiquitin after trypsin cleavage (*11*), enables the isolation and identification by MS of ubiquitylated peptides, thereby providing information about the site of ubiquitylation to complement quantitative assays (*12*, *13*).

Alternative approaches based on affinity purification followed by mass spectrometry (AP-MS) (*14*–*16*) using epitope tagging are also feasible for such screens. In AP-MS methods, a protein of interest (bait) is fused in frame with an epitope tag, which enables the isolation of the protein complexes (preys). The affinity-purified complexes are then identified by MS. Although AP-MS experiments have contributed to the identification of substrates of monomeric E3s, identification of substrates of multisubunit E3s is hampered by the weak affinity of substrates for the subunit of the E3 that recognizes the substrate and because the complex between the substrate and the recognition subunit is transient (*17*).

Here, we used a stringent AP-MS approach to systematically identify substrates of $SCF^{\beta TrCP}$, a ubiquitin ligase that targets key regulators of cell proliferation, cell growth (*15*), and apoptosis (*18*) for proteasomal degradation. $SCF^{\beta TrCP}$ consists of Skp1, Cul1, Rbx1, and the F-box protein βTrCP that acts as a substrate recognition subunit (Fig. 1). Mammals have two paralogs of βTrCP: βTrCP1 (also known as FBXW1) and βTrCP2 (also known as FBXW11); however, their biochemical properties appear indistinguishable. Although more than 50 substrates of $SCF^{\beta TrCP}$ have been identified since 1999, the list is undoubtedly incomplete. Certain E3 ligases of the SCF family recognize their substrates through a phosphorylated motif called a phosphodegron. Most known $SCF^{\beta TrCP}$ substrates share the canonical phosphodegron DpSGXX(X)pS, which is a diphosphorylated motif that mediates the interaction of the substrate with

[1]Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Padualaan 8, 3584 CH Utrecht, Netherlands. [2]Netherlands Proteomics Center, Padualaan 8, 3584 CH Utrecht, Netherlands. [3]Hubrecht Institute–KNAW (Royal Netherlands Academy of Arts and Sciences) and University Medical Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, Netherlands.
*These authors contributed equally to this work.
†Present address: Departments of Chemistry and Biochemistry, University of Oxford, New Biochemistry Building, South Parks Road, Oxford OX1 3QU, UK.
‡Corresponding author. E-mail: a.j.r.heck@uu.nl

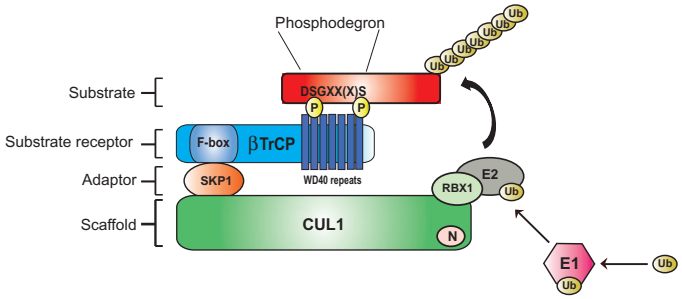


Fig. 1. The architecture of wild-type $SCF^{\beta TrCP}$. $SCF^{\beta TrCP}$ consists of the core subunits Cul1, Rbx1, and Skp1, and the F-box protein βTrCP that determines substrate specificity. Cul1 is a scaffold protein that interacts through its N terminus with the adaptor protein Skp1 and through its C terminus with the RING finger protein Rbx1, which in turn binds to a specific ubiquitin-conjugating enzyme (E2). Skp1 interacts with the F-box domain of βTrCP, which, in turn, recruits substrates through its WD40 repeats. These repeats form a β-propeller structure, which specifically recognizes a diphosphorylated motif with the consensus DpSGXX(X)pS, known as phosphodegron.

the WD40 repeats of βTrCP (Fig. 1) (*19*, *20*). Mutation of either the phosphodegron or the WD40 repeats abolishes substrate binding to βTrCP (*21*). The phosphodegrons of all reported $SCF^{\beta TrCP}$ substrates are conserved in vertebrates (*22*).

A challenge in AP-MS analyses is discriminating bona fide interacting partners from nonspecific protein contaminants (*23*). To resolve this problem, we used label-free quantitative proteomics in combination with two different background-filtering procedures. First, we used the CRAPome (*24*) software suite, which not only includes a large list of common nonspecific contaminants but also provides, due to the embedded SAINT algorithm, a standardized scoring scheme for ranking each prey candidate against these contaminants (*23*, *25*, *26*). We also analyzed the data with a second strategy based on precursor ion (MS1) intensity using the MaxQuant-Perseus software suite (*27*–*30*), which is another method that enriches for high-confidence interactors using label-free quantitative proteomics. By analyzing the AP-MS data with these software tools, we discriminated specific protein interactors from presumed background contaminants (*31*) and studied dynamically regulated interactions (*32*–*34*). Using this workflow, we identified most of the known substrates of $SCF^{\beta TrCP}$ as well as many putative previously unknown substrates of this ubiquitin ligase. We validated this approach by examining several putative $SCF^{\beta TrCP}$ substrates by immunoprecipitation followed by immunoblotting and in vitro ubiquitylation assay. We envisage that this approach may be useful for identifying substrates of other E3 ubiquitin ligases.

## RESULTS

### Mining βTrCP substrate phosphodegron motifs in the human proteome

We first extracted 47 phosphodegrons from reported substrates of βTrCP (table S1) and used the MEME program to formulate a stringent consensus motif (Fig. 2A). We then probed the human proteome (Swiss-Prot, human, July 2013, 20,277 entries) using this consensus motif with the FIMO program, which generates a *P* value for every motif matched (the smaller the *P* value, the more closely the motif matches the consensus). Overall, we obtained 17,779 short sequences that met the requirement of a $P \leq 1 \times 10^{-3}$ and represented 7623 proteins (table S2 and Fig. 2B). These sequences included 45 of the 47 reported phosphodegrons used for generating the consensus sequence; the two that were not recovered were viral protein U (Vpu) and the RNA binding protein ELAV1. Vpu is a viral protein encoded by HIV-1. Because the human Swiss-Prot database that we used for FIMO analysis does not contain this viral protein, the phosphodegron for Vpu (RAEDSGNESE) was not recovered. Our recalculation shows that the FIMO *P* value for this phosphodegron is $1.9 \times 10^{-6}$. The phosphodegron sequence (TNYEEAAMAI) of ELAV1 is so deviated from the consensus that the FIMO *P* value is 0.00124, lower than the $P \leq 1 \times 10^{-3}$ cutoff, and hence is also not included.

Because FIMO *P* values can be used to estimate how similar each putative phosphodegron is to the consensus motif, we ranked each matched motif according to its respective FIMO *P* values (lowest first, representing the closest match to the consensus motif). As expected, the top-ranking reported phosphodegrons generally contained motifs that matched more closely to the DpSGXX(X)pS motif, for example, eEF2K (ranked 1), CTNNB1 (ranked 11), DLG1 (ranked 246), and SIPA1L1 (ranked 404). For ease of reference, we named them category I motifs. Following the SIPA1L1, we observed a sharp drop in the ranking for motifs derived from substrates such as PLK4 (ranked 1214), MYC (ranked 1426), CDC25B (ranked 2931), TWST1 (ranked 3778), and MCL1 (ranked 5218), which have less canonical motifs. The second group is termed category II motifs.

To further refine and corroborate our motif analysis, we queried the PhosphoSitePlus repository (*35*), which contains phosphoproteomics and ubiquitomics data. In total, we found 3881 phosphorylation sites reported for our putative degrons. This analysis revealed that 46% (3507 of 7623) of the proteins with putative degrons have at least one ubiquitylation site (table S2). However, because of the low abundance of some modifications at specific sites and limitation of proteomics technology, we failed to find matching phosphoproteomics data for some well-characterized $SCF^{\beta TrCP}$ substrates, such as PER1, WEE1, EPOR, CPEB-1, PLK4, and FANCM.

Evolutionary conservation can be used to distinguish between functional and nonfunctional motifs (*36*–*38*), and thus, we subjected each βTrCP degron to a conservation analysis. Among the 45 known degrons identified in our experiment, 42 degrons had a conservation score ≥0.7 between human and mouse orthologs (table S2). However, only 23 degrons had a conservation score ≥0.7 between human and fish, and only 1 degron (SYLDSGIHSG) for CTNNB1 was conserved between human and *Drosophila*. This analysis indicated that βTrCP mainly functions in vertebrates, although its ortholog (called Slimb) also exists in flies. By setting a conservation score of 0.7 as observed in human and mouse orthologs, we reduced the total 17,779 putative degrons to 13,109 (74%).

### Identification of known and candidate substrates of βTrCP2

For the AP-MS experiments, we designed three different constructs encoding βTrCP2 carrying N-terminal 2xFLAG and 2xHA epitope tags (Fig. 3, A to C). The rationale of using two sequential immunoprecipitations with two different epitope tags is to reduce the amount of background proteins that copurify, albeit at the expense of weaker protein-protein interactions (PPIs), which may be lost due to the high stringency inherent to this sequential method. We used the epitope tags without any βTrCP2-based fusion protein (EV) as one negative control. The three baits consist of the wild-type βTrCP2 (Fig. 3A), βTrCP2 (R447A) (Fig. 3B), and βTrCP2-ΔF (Fig. 3C). The Arg-to-Ala mutation in βTrCP2 (R447A) is within the WD40 β-propeller domain and abrogates the interaction with and consequently ubiquitylation of $SCF^{\beta TrCP}$ substrates (*19*, *21*); however, like the corresponding βTrCP1 (R474A) mutant, βTrCP2 (R447A) retains the ability to interact with other components and regulators of the SCF complex (*21*, *39*, *40*). We assumed that such a single amino acid mutant serves

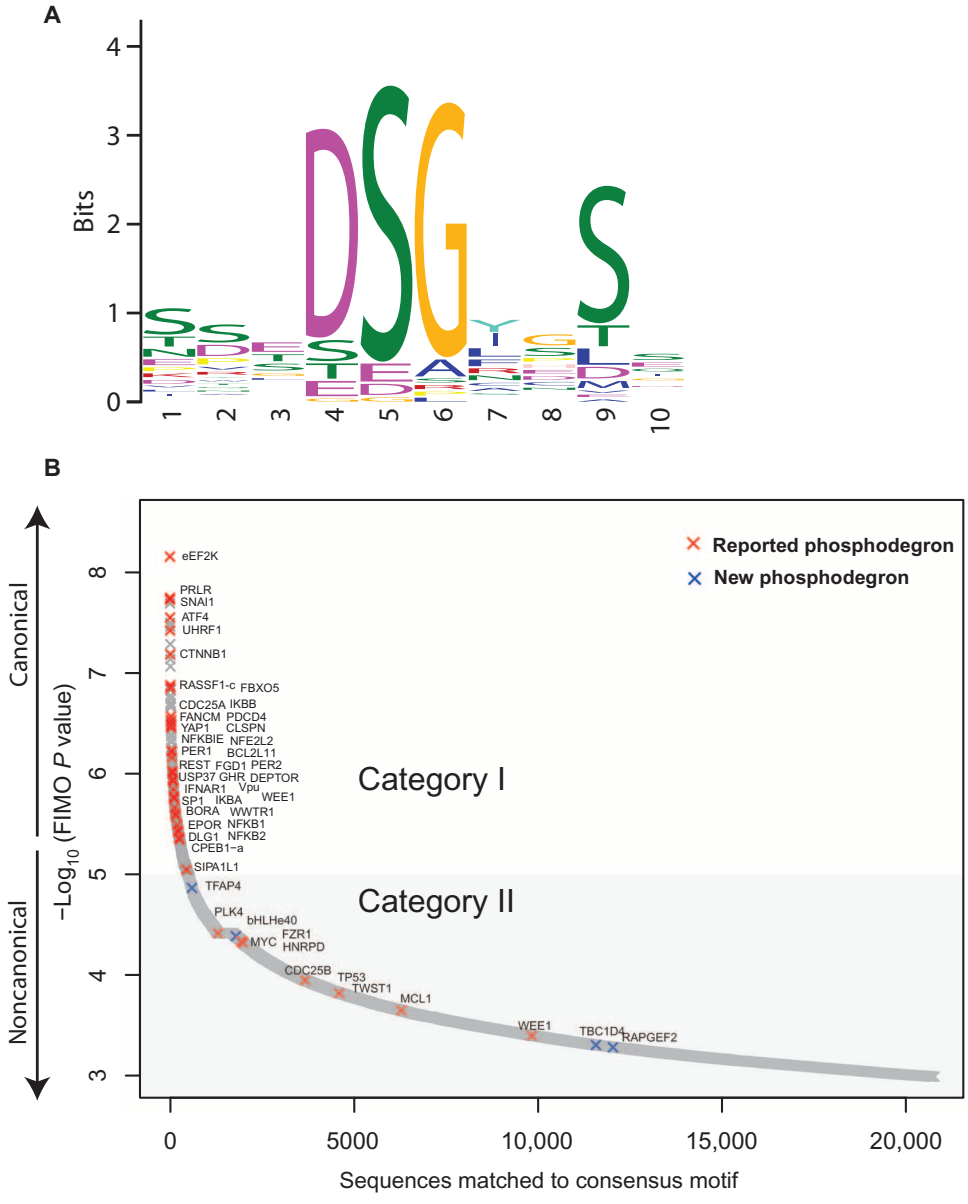


Fig. 2. Mining the human genome for the DpSGXX(X)pS motifs. (A) Consensus motif derived from the phosphodegrons from reported βTrCP substrates: 47 well-characterized phosphodegrons were used to build a consensus motif using MEME (http://meme.nbcr.net), highlighting the already known canonical DpSGXX(X)pS sequence motif where D, S, and G are the best-conserved amino acid residues. (B) Distribution of FIMO –log *P* values against all possible sequence motifs in the human proteome: all 17,779 sequence motifs obtained from FIMO analysis of the human proteome are each plotted against their respective –log *P* values. Some proteins have more than one motif, for example, WEE1 has both a category I and a category II motif.

as a critical negative control in the context of substrate screening, not only because it allows selection against background contaminants but also because it enables the identification of interactors that are specific for the WD40 domain. The βTrCP2-ΔF mutant carries a deletion in the F-box motif and, when overexpressed in cells, acts as a dominant-negative mutant (*41*, *42*) because it binds substrates but cannot interact with Skp1, Cul1, Rbx1, and the E2 enzyme.

We identified 1626 proteins in the AP-MS experiments, 27 of which were previously reported substrates of $SCF^{\beta TrCP}$ (tables S3 and S4). Because the bottleneck for AP-MS experiments is the selection of bona fide protein binders from an abundance of background contaminants, we predicted that most of the 1626 proteins were nonspecific or secondary binders. To differentiate genuine PPIs from background contaminants, we used two distinctive quantitative proteomics filtering tools—the CRAPome and Perseus software suites—and compared immunopurified samples with negative controls that did not contain any bait (Fig. 3D).

The CRAPome software enables background filtering using spectra count data acquired by AP-MS (*24*). It consists of a database of categorized background contaminants with two scoring schemes to independently rank each pair of PPI identified by immunoprecipitation. Here, we ranked wild-type βTrCP2, βTrCP2 (R447A), and βTrCP2-ΔF individually against the negative control (EV). For each bait-prey interaction, CRAPome computed a probability-based SAINT score (*23*) and a separate fold change (FC) score (*24*) (table S3A). For the wild-type βTrCP2 immunoprecipitation, 20 of the 27 reported substrates had high SAINT scores of ≥0.90 (fig. S1 and table S3B), suggesting that the SAINT score alone may be sufficient to identify genuine binders. Therefore, we filtered all PPIs from wild-type βTrCP2, βTrCP2 (R447A), and βTrCP2-ΔF immunoprecipitations against EV with a SAINT score ≥0.90 (fig. S1). However, this analysis revealed that even specific binders of βTrCP2 are not always bona fide substrates, as exemplified by the immunoprecipitation of SCF subunits Skp1, Cul1, and Rbx1, which all had SAINT scores ≥0.9.

The FC score not only estimates the relative abundance (FC) of proteins between two samples but also has a wider dynamic range than the SAINT score (fig. S1). Therefore, we computed a second value (FC ratio) from the FC scores of wild-type βTrCP2 and βTrCP2-ΔF against those obtained for βTrCP2 (R447A), which should not bind any substrates. The FC ratio provides a degree of confidence that the bound proteins are substrates, as well as providing information that can be used to estimate binding stoichiometry. By setting a SAINT score cutoff of 0.90 and an FC ratio cutoff of 10.00, the SCF subunits βTrCP2, Skp1, Cul1, and Rbx1 (SAINT score ≥0.90) were classified as nonsubstrates, whereas 18 of the 27 reported substrates were retained (Fig. 4A). Collectively, from the data, we defined 338 specific interactors of the WD40 repeat region of βTrCP2 [154 for wild-type βTrCP2 and 231 for βTrCP2-ΔF against βTrCP2 (R447A), with 47 in common] on the

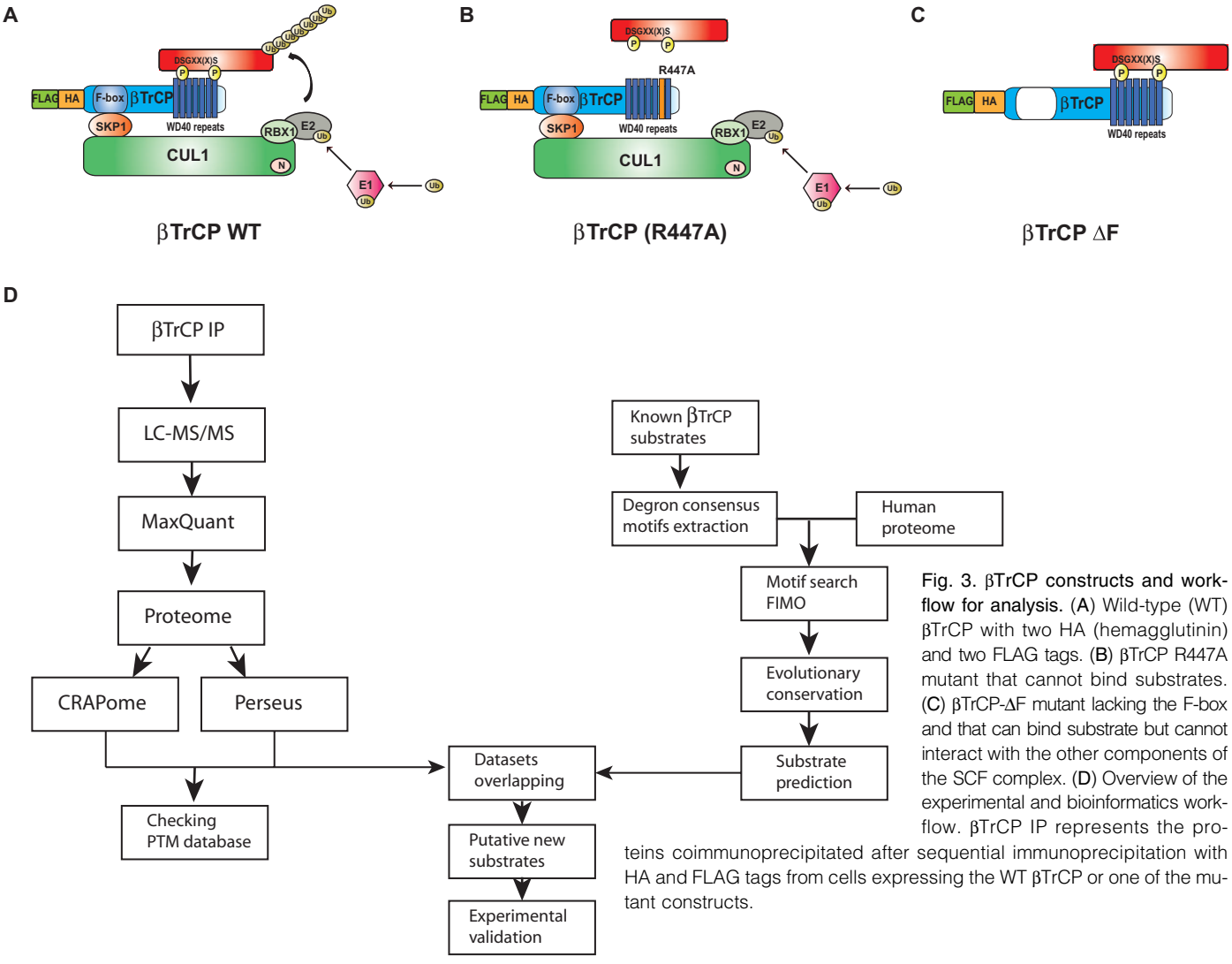


Fig. 3. βTrCP constructs and workflow for analysis. (A) Wild-type (WT) βTrCP with two HA (hemagglutinin) and two FLAG tags. (B) βTrCP R447A mutant that cannot bind substrates. (C) βTrCP-ΔF mutant lacking the F-box and that can bind substrate but cannot interact with the other components of the SCF complex. (D) Overview of the experimental and bioinformatics workflow. βTrCP IP represents the proteins coimmunoprecipitated after sequential immunoprecipitation with HA and FLAG tags from cells expressing the WT βTrCP or one of the mutant constructs.

basis of the CRAPome analysis (table S3C), including 20 (18 from wild type and 3 from ΔF) of 27 previously reported substrates.

To increase confidence in the identified substrates, we performed a second filtering analysis by Perseus (*43*) using the normalized MS1 [label-free quantification (LFQ)] intensity (*44*) obtained from MaxQuant. For each pair of interactions, we performed Student's *t* tests (table S4A and Fig. 4B). With the Perseus workflow, we statistically confirmed 20 of the 27 reported substrates (table S4B), of which 18 were significantly enriched in the wild-type βTrCP2 immunoprecipitation data and 7 in the βTrCP2-ΔF immunoprecipitation data. Together, we obtained 317 significant putative hits (163 for wild-type βTrCP2 and 221 for βTrCP2-ΔF, with 61 in common between the two) using label-free quantitation (table S4C).

Overall, we identified 338 βTrCP2 substrate candidates that were validated with the CRAPome and 317 validated with Perseus; the candidates validated by both methods represented 207 proteins (Fig. 4C). Among the 27 previously reported substrates, we confirmed 22 in the CRAPome analysis (20 substrates) and Perseus (20 substrates), and the overlap between them is 18 (Fig. 4D).

## Reproducibility of AP-MS experiments

We have been working toward the goal of identifying substrates of $SCF^{\beta TrCP}$ for several years and earlier gathered data using an identical sequential immunoprecipitation AP protocol in conjunction with older versions of the separation, MS, and bioinformatics technologies. These older data led to the identification, validation, and characterization of four $SCF^{\beta TrCP}$ substrates: eEF2K (*21*), RAPGEF2 (*39*), TFAP4 (*40*), and bHLH40e (*45*).

To evaluate the reproducibility of our AP-MS workflow, we reanalyzed these older data (data set 2) using CRAPome (table S5). Because these older experiments were performed with MS instruments that are at least two generations older (Orbitraps Classic and Discovery) and are less sensitive and lower in sequencing speed compared to the Orbitrap Elite, fewer proteins were identified (725 in data set 2 versus 1626 in data set 1) (table S6). Of the 725 proteins identified, 531 overlap with data set 1 (73%), indicating that the AP-MS experiments are consistent and reproducible. Upon the CRAPome analysis, we assigned 47 proteins from data set 2 as putative βTrCP substrates, of which 36 overlapped with the 154 putative substrates in data set 1. The lower number of putative substrates validated by the CRAPome analysis in data set 2 reflected the lower spectra

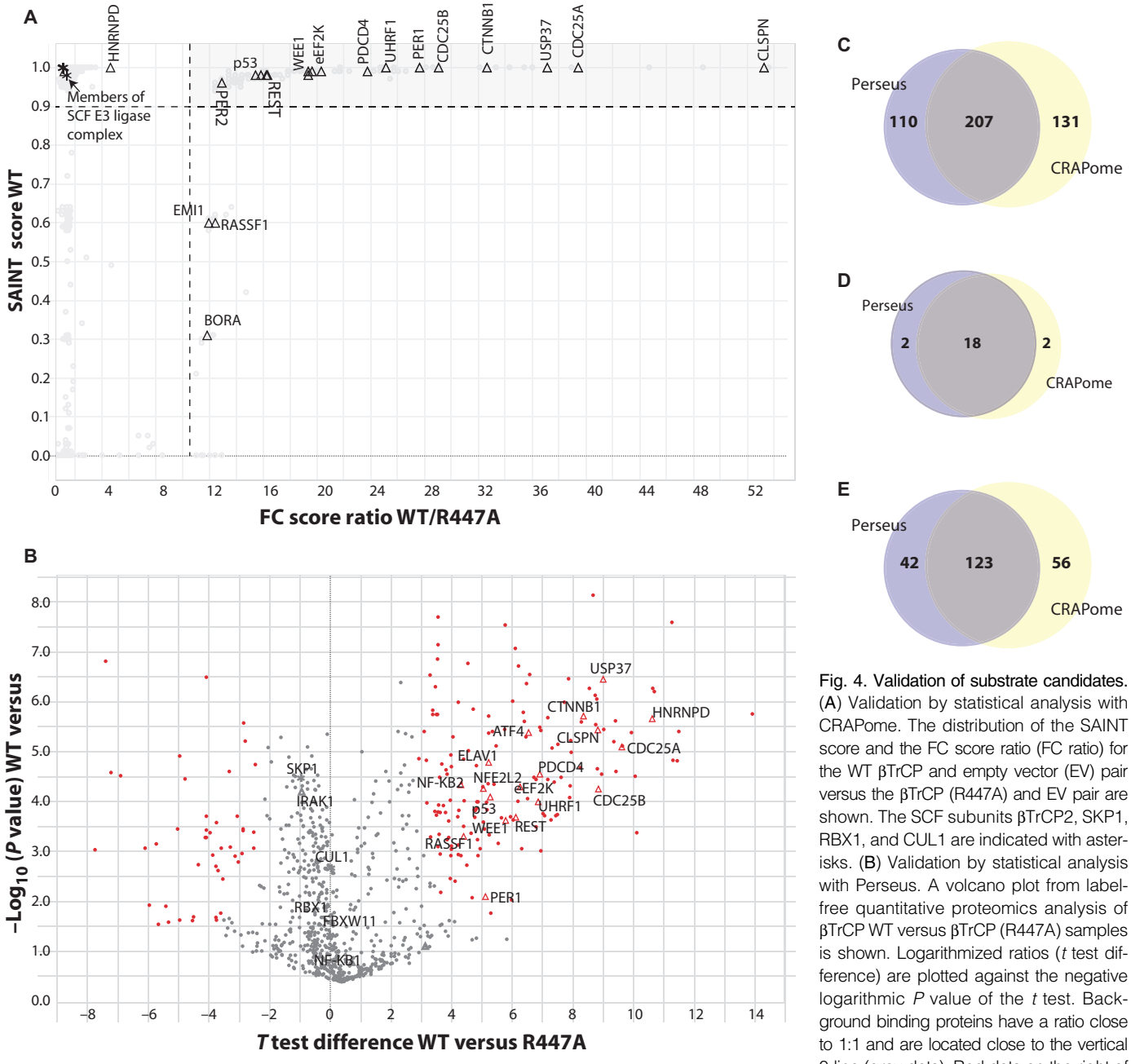


Fig. 4. Validation of substrate candidates. (**A**) Validation by statistical analysis with CRAPome. The distribution of the SAINT score and the FC score ratio (FC ratio) for the WT βTrCP and empty vector (EV) pair versus the βTrCP (R447A) and EV pair are shown. The SCF subunits βTrCP2, SKP1, RBX1, and CUL1 are indicated with asterisks. (**B**) Validation by statistical analysis with Perseus. A volcano plot from label-free quantitative proteomics analysis of βTrCP WT versus βTrCP (R447A) samples is shown. Logarithmized ratios (*t* test difference) are plotted against the negative logarithmic *P* value of the *t* test. Background binding proteins have a ratio close to 1:1 and are located close to the vertical 0-line (gray dots). Red dots on the right of vertical 0-line indicate significant hits for substrates because they are more enriched in the WT sample compared to the R447A mutant. Red triangles represent reported substrates; gray triangles represent previously reported substrates that are not statistically significant in our AP-MS experiments; asterisks represent SCF subunits. (**C**) Venn diagram of the 338 candidates verified by the CRAPome analysis and 317 verified by Perseus analysis. (**D**) Venn diagram of the reported substrates verified by CRAPome and Perseus analysis methods. (**E**) Venn diagram of the substrates analyzed with CRAPome or Perseus but limited to those from the AP-MS data that also contained a degron motif (FIMO $P \leq 1 \times 10^{-3}$) and a conservation score of at least 0.7.

counts (due to the lower sensitivity and speed of older MS), which affected calculation of the multiple CRAPome scores.

We also examined the number of reported substrates recovered. We identified 25 previously reported βTrCP substrates from data set 2 (table S6), of which 21 were also present in the 27 identified in data set 1. However, when we applied the stringent CRAPome filters (SAINT score ≥0.9, FC ratio ≥10), only 13 from data set 2 met the criteria for putative substrates, of which 12 overlapped with those in data set 1. Nevertheless, eEF2K (*21*), RAPGEF2 (*39*), TFAP4 (*40*), and bHLH40e (*45*) met the stringent criteria for putative substrates in both data set 1 and data set 2 after reanalysis. Therefore, the AP-MS experiments were reproducible, but illustrate how the continuous advancement in sensitivity and sequencing speed

of MS also leads to a parallel and inevitable improvement in MS data in terms of both quality and quantity, which allows more stringent data filtering and the recovery of more confident hits.

## Integration of AP-MS and bioinformatics prediction data

Taking the 448 specific interactors classified by each CRAPome and Perseus analyses as input, we filtered these proteins by the presence of motifs matching the phosphodegron consensus sequence and by the evolutionary conservation of the motifs. We combined the AP-MS data with the bioinformatics predictions by first requiring all putative substrates to have a degron motif (FIMO $P \leq 1 \times 10^{-3}$) and a conservation score of at least 0.7 (table S7A). These additional criteria reduced the number of putative substrates in the combined CRAPome and Perseus data set from 448 to 221 (Fig. 4E and table S7A).

Next, we divided these 221 putative substrates into six bins according to their FIMO *P* values (Fig. 5 and table S7B). The number of putative substrates increased with the FIMO *P* value bins. Second, although consisting of smaller numbers, bins with lower *P* values are more enriched in reported substrates. For example, in the first (lowest) bin, eEF2K is the only candidate in the bin and is a reported substrate (enrichment of reported substrates is 100%). This is followed by the second and third bins, both displaying 67% enrichment (two of three and six of nine candidates in the bin represent reported substrates, respectively). In the final bin containing proteins with category I motifs, there is 35% enrichment. Arranging these candidates according to the ranks of their predicted degron motifs (table S5B) placed DOCK7 as the last candidate in the fourth bin, and this protein had a phosphodegron motif that was ranked 432, which is at the borderline of the category I motifs (Fig. 2B). Therefore, this analysis suggested that the putative substrates within bins 1 to 4 are high-confidence $SCF^{\beta TrCP}$ substrates, which is consistent with the recent characterization of eEF2K as an $SCF^{\beta TrCP}$ (*21*).

The fifth and sixth bins contain the highest number of candidates (54 and 134, respectively) but also the lowest number of established substrates (1 and 3, respectively). Most candidates in these bins have motifs (category II motifs) that diverge from the canonical phosphodegron motif of the known substrates. Although we considered the candidates with category II motifs less confident, some candidates from these bins have been identified as $SCF^{\beta TrCP}$ substrates, such as TFAP4 (*40*) (fifth bin), bHLHe40 (*45*) (fifth bin), and RAPGEF2 (*39*) (sixth bin).

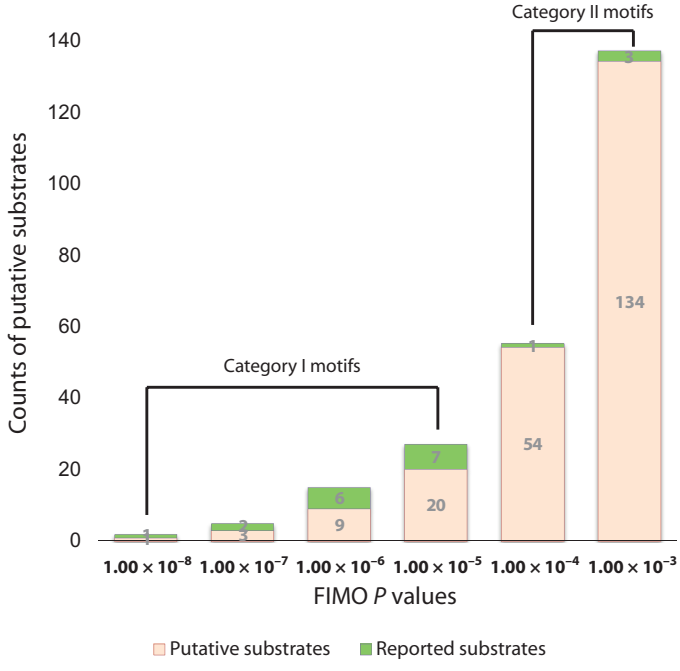


Fig. 5. Distribution of putative substrates by binned FIMO *P* values. Dividing the 221 candidates into six bins according to their FIMO *P* values. All candidates have a least one putative degron (FIMO score $P \leq 1 \times 10^{-3}$) with a conservation score ≥0.7. Green represents number of candidates within a bin that are already reported. The lower number in each bin represents the total number of candidates in the bin.

## Additional biochemical analysis of previously unknown βTrCP substrates

To provide additional support that the previously unknown interactors of βTrCP that we identified are substrates, we performed additional biochemical analysis of two of the putative substrates of $SCF^{\beta TrCP}$. We selected for additional biochemical analysis RAPGEF2, a guanine nucleotide exchange factor for the guanosine triphosphatase (GTPase) RAP, which is involved in the development and maintenance of epithelia (*46*–*48*), and TBC1D4, a Rab GTPase-activating protein, which is involved in insulin-regulated trafficking of the glucose transporter GLUT4 (*49*, *50*).

To test the interaction of βTrCP with RAPGEF2 and TBC1D4, we immunoprecipitated from human embryonic kidney (HEK) 293T cells FLAG-tagged βTrCP1, FLAG-tagged βTrCP2, several other FLAG-tagged F-box proteins, as well as CDH1 and CDC20, which are two WD40-containing substrate recognition subunits of the APC/C (anaphase-promoting complex/cyclosome) ubiquitin ligase. After immunoblotting, we observed that βTrCP1 and βTrCP2, and not the other F-box proteins, coimmunoprecipitated endogenous RAPGEF2 and TBC1D4 (Fig. 6A), as well as other previously reported substrates, REST and β-catenin (also known as CTNNB1), which served as positive controls.

The $Arg^{447}$ residue located within the WD40 repeats of βTrCP2 ($Arg^{474}$ in βTrCP1) is essential for the βTrCP-substrate interaction (*21*, *39*). To assess whether the observed interactions depended on the WD40 β-propeller structure, we immunoprecipitated the FLAG-tagged wild-type βTrCP2 and the βTrCP2 (R447A) mutant and examined their binding to these substrates. As expected, wild-type βTrCP2, but not the βTrCP2 (R447A) mutant, immunoprecipitated endogenous RAPGEF2, bHLH40e (also known as DEC1), and TBC1D4 (Fig. 6B), as well as the known $SCF^{\beta TrCP}$ substrates PDCD4 and β-catenin.

If RAPGEF2 and TBC1D4 are bona fide substrates of $SCF^{\beta TrCP}$, they should be ubiquitylated by this ubiquitin ligase. Purified wild-type βTrCP, but not the inactive βTrCP-ΔF mutant, induced the ubiquitylation of TBC1D4 (Fig. 6C) and RAPGEF2 (*39*) when assayed in vitro with other subunits of the SCF complex. Thus, these data indicated that RAPGEF2 and TBC1D4 can be ubiquitylated by $SCF^{\beta TrCP}$.

## Identification of putative regulators or nonsubstrate interactors of $SCF^{\beta TrCP}$

In addition to identifying putative substrates of βTrCP, our CRAPome and Perseus analyses also indicated that proteins that are not substrates of βTrCP but instead interact with this F-box protein were also enriched in the AP-MS analysis. For example, the SCF core subunits Skp1 and Cul1 were enriched, but these are not substrates of βTrCP. The interaction of Skp1 and Cul1 with βTrCP was readily detected by coimmunoprecipitation from HEK293T cells expressing FLAG-tagged wild-type βTrCP2 or βTrCP2 (R447A). We found that HUWE1 and UBR5 (also known as

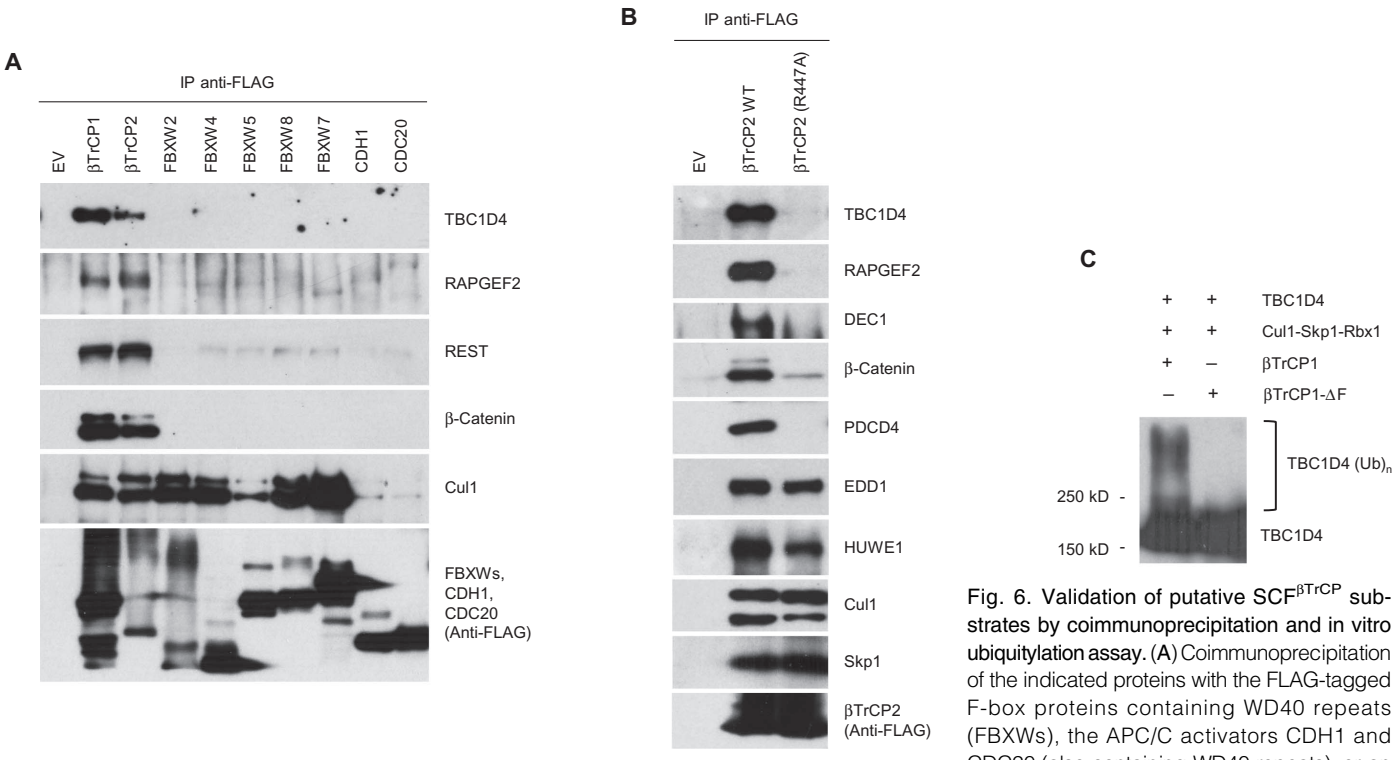


Fig. 6. Validation of putative SCF$^{\beta TrCP}$ substrates by coimmunoprecipitation and in vitro ubiquitylation assay. (**A**) Coimmunoprecipitation of the indicated proteins with the FLAG-tagged F-box proteins containing WD40 repeats (FBXWs), the APC/C activators CDH1 and CDC20 (also containing WD40 repeats), or an empty vector (EV) from HEK293T cells. Immunoprecipitation was performed with the antibody recognizing FLAG, and proteins were detected by immunoblotting with antibodies for the indicated proteins. (**B**) Analysis of the dependence of βTrCP-substrate interaction on the WD40 repeat of βTrCP. Proteins that coimmunoprecipated with FLAG-tagged βTrCP2 WT or βTrCP2 (R447A) were analyzed by immunoblotting. (**C**) In vitro ubiquitylation of TBC1D4 with SCF$^{\beta TrCP1}$ immunoprecipitated from HEK293T cells transfected with TBC1D4, Skp1, Cul1, and Rbx1 in the presence of FLAG-tagged βTrCP1 or FLAG-tagged βTrCP1-ΔF. The bracket indicates a ladder of bands corresponding to polyubiquitylated TBC1D4. Data are representative of three experiments.

EDD1), both of which are components of other E3 ubiquitin ligases, had AP-MS profiles similar to Skp1 and Cul1 (tables S3A and S4A). Therefore, we tested if HUWE1 and UBR5 coimmunoprecipitated with both wild-type βTrCP2 and βTrCP2 (R447A) (Fig. 6B). These results confirmed that both HUWE1 and UBR5 are part of the βTrCP interactome, and indicated that the interaction is independent of the substrate recognition WD40 domain. Thus, this approach not only identified substrates of an E3 ubiquitin ligase but, depending on the mutants used, also revealed potential regulators or other complexes in which the E3 ubiquitin ligase may function.

## DISCUSSION

Here, we showed that searching biological sequence databases for sequences using the MEME-FIMO software tool provides a useful initial data set of putative E3 ubiquitin ligase substrates by identifying, scoring, and ranking sequences that match closely to the conserved phosphodegron motif. We used the DpSGXX(X)pS phosphodegron motif extracted from known SCF$^{\beta TrCP}$ substrates as proof of principle. This bioinformatics tool also captures proteins with phosphodegron that diverged in sequence from the canonical SCF$^{\beta TrCP}$ phosphodegron; these category II motifs had higher FIMO *P* values. To minimize the number of possible candidates, we attempted to match the short sequences to PhosphoSitePlus; however, because of the incomplete repository, the phosphorylation status of some of these motifs could not be confirmed by this method. Therefore, we applied evolutionary conservation to reduce the 17,779 predicted candidates to 13,109 sequences (6091 proteins).

To further limit the substrate repertoire, we used AP-MS to identify proteins that interacted with βTrCP2. To confidently discern genuine interactions from background contaminants, we analyzed the AP-MS data using the spectra count–based CRAPome and MS1 intensity–based Perseus software tools to validate the quantitative immunoprecipitation results. We found that, although there was only a modest 46% overall overlap between the CRAPome and Perseus results (Fig. 4C), the agreement was 82% for the established substrates (Fig. 4D). One possible explanation for why the ability of both CRAPome and Perseus analyses performed similarly, with the highest overlap, for known substrates is the use of the βTrCP2 (R447A) mutant as a negative control. The interactome for the βTrCP2 (R447A) mutant theoretically resembles that of the wild-type βTrCP2, except that it specifically lacks substrates. For these substrates, the AP-MS result should represent an "on" (wild-type βTrCP2 or βTrCP2-ΔF) and "off" [βTrCP2 (R447A)] situation regardless of whether spectra counts or MS1 intensities are used to analyze the data. A newer version of the SAINT algorithm may solve the moderate overlap between the CRAPome and Perseus results for the interactome, because it is also based on MS1 quantitation (*25*). It will be interesting to reanalyze the data once this SAINT algorithm has been incorporated into CRAPome.

We further showed that only 221 of the 448 specific βTrCP interactors contain sequences similar to the consensus motif and at the same time were highly conserved (Fig. 4E). Therefore, not all the specific interactors

that depended on the WD40 repeats are likely substrates; instead, they may represent secondary or tertiary binders.

Many proteomics screens for substrates of E3 ubiquitin ligases are mostly conducted by means of monitoring protein abundance (*8*–*10*, *13*), which lack evidence pertaining to ligase-substrate interaction, because the substrates can be indirectly regulated, as well as regulated by ubiquitylation. Here, we developed an interaction proteomics approach that effectively identifies most of the reported substrates of $SCF^{\beta TrCP}$ in a single screen. In addition, using bioinformatics, we ranked each AP-MS candidate with confidence based on the similarity of their putative degrons to the consensus motif. At least four of the substrates that we identified from this screen—eEF2K (*21*), RAPGEF2 (*39*), TFAP4 (*40*), and bHLHe40 (*45*)—have been validated by various biochemical methods.

We expect that other baits can be designed on the basis of the substrate recognition substrates for other E3 ubiquitin ligases and that this combined AP-MS, bioinformatic, and stringent filtering approach will lead to specific and effective screens for the identification of substrates of ubiquitin ligases, especially those of the F-box family.

## MATERIALS AND METHODS

### Cell culture, transfection, and drug treatment

HEK293T cells were cultured in Dulbecco's modified Eagle's medium (Invitrogen) with 10% fetal calf serum and penicillin-streptomycin (100 U/ml). HEK293T cells were transfected using the polyethylenimine (PEI) method and, 48 hours after transfection, treated with the proteasome inhibitor MG132 (10 μM; Peptide Institute) for 5 hours.

### DNA constructs

Full-length, human wild-type βTrCP2 (isoform C, NP_036432.2) carrying HA and FLAG tags at the N terminus was cloned into the Eco RI and Not I sites of pcDNA3 by polymerase chain reaction. The full-length construct was then used as template to generate the βTrCP2 (R447A) mutant using the QuikChange Site-Directed Mutagenesis kit (Stratagene) according to the manufacturer's directions.

The following oligonucleotides were used: forward primer, 5′-GGGACATGAAGAATTGGTCGCATGCATCCGGTTTGATAACAAG-3′; reverse primer, 5′-CTTGTTATCAAACCGGATGCATGCGACCAATTC-TTCATGTCCC-3′. All constructs were verified by sequencing.

### Biochemical methods

Cells were harvested and lysed with lysis buffer [50 mM tris-HCl at pH 7.4, 1 mM EDTA, 250 mM NaCl, 0.1% Triton X-100, 1 mM DTT (dithiothreitol)] containing protease inhibitors [leupeptin (1 μg/ml), aprotinin (1 μg/ml), soybean trypsin inhibitor (10 μg/ml), 0.1 mM phenylmethylsulfonyl fluoride, tosyl lysyl chloromethyl ketone (10 μg/ml), tosylphenylalanyl chloromethyl ketone (10 μg/ml); Sigma-Aldrich] and phosphatase inhibitors (50 mM sodium fluoride, 100 μM sodium orthovanadate; Sigma-Aldrich) and kept on ice for 30 min. After centrifugation at 18,500 RCF (relative centrifugal force) in a microcentrifuge for 20 min, the supernatant was recovered and protein concentration was determined. Laemmli sample buffer was added to 30 μg of protein extract, which was loaded and separated by SDS–polyacrylamide gel electrophoresis.

### Antibodies

The following monoclonal antibodies were used: antibody recognizing Cul1 (Invitrogen), antibody recognizing FLAG (Sigma-Aldrich), antibody recognizing actin (Santa Cruz Biotechnology), antibody recognizing FLAG M2 (Sigma-Aldrich), antibody recognizing HA 12CA5 (Covance), and antibody recognizing β-catenin (BD Biosciences). Polyclonal antibodies were those recognizing PDCD4 (from Bethyl Laboratories), REST (Millipore), FLAG (Sigma-Aldrich), and TBC1D4 (Cell Signaling Technology).

### βTrCP2 immunopurification

HEK293T cells grown in 15-cm dishes were transfected with pcDNA3-2xFLAG-2xHA-βTrCP2 and treated with 10 μM MG132 for 5 hours. Cells were harvested and subsequently lysed in lysis buffer [50 mM tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, and 0.5% NP-40]. βTrCP2 was immunopurified with the mouse antibody against FLAG M2 coupled to an agarose resin (Sigma-Aldrich). After washing, proteins were eluted by competition with FLAG peptide (Sigma-Aldrich). The eluate was then subjected to a second immunopurification with resin linked to the antibody against HA (12CA5 monoclonal antibody cross-linked to protein G–Sepharose; Invitrogen).

### Recovery of immunoprecipitated proteins for liquid chromatography–MS/MS

Immunoprecipitated proteins were resuspended in a spin filter column (Bio-Rad) and washed three times with 200 μl of phosphate-buffered saline (pH 7.4) to remove residual detergent from the lysis buffer. Proteins bound to the protein G beads were then eluted off with 100 μl of 0.5% RapiGest reagents in 50 mM ammonium bicarbonate, followed by 100 μl of 8 M urea in 50 mM ammonium bicarbonate (pH 8.0). Eluted proteins were reduced with 1 mM DTT and alkylated with 5.5 mM iodoacetamide. For tryptic digestion, proteins were first digested with endoproteinase Lys-C (Wako Chemicals) in room temperature for 4 hours, followed by sequencing grade–modified trypsin (Promega) overnight, after fourfold dilution with 50 mM ammonium bicarbonate. Protease digestion was stopped by addition of trifluoroacetic acid (TFA), and precipitates were removed after centrifugation. Peptides were desalted using reversed-phase Sep-Pak C18 cartridges (Waters), then dried and stored at −20°C.

### Liquid chromatography–MS/MS

For MS analysis, peptides were first separated with a C18 column (Zorbax) and introduced by nanoelectrospray into the LTQ Orbitrap Elite (Thermo Fisher) and MS/MS in data-dependent decision tree mode (collision-induced dissociation/electron transfer dissociation) as described previously (*51*, *52*).

### MS data analysis

MS raw files were analyzed using MaxQuant (version 1.4.0.8) with the match between runs and LFQ options selected. MS/MS spectra, top 8 selected in 100-dalton bin, were searched against the UniProt Human database (version 2013-07, 20,277 entries). Trypsin/P was chosen as the protease, cysteine carbamidomethylation was set as fixed modification, and oxidation of methionine and acetylation of the N terminus of the protein were set as variable modifications. Peptide tolerance was initially set to 20 ppm for first search and 4.5 ppm after recalibration; MS/MS tolerance was set to 0.5 dalton. All peptide-spectrum matches (PSMs) and proteins were validated with 1% false discovery rate (FDR). Only PSMs with a minimum length of 7 amino acids were kept.

### Post-acquisition data analysis: CRAPome and Perseus analyses

Before analysis using the CRAPome or Perseus software suite, the "proteingroups.txt" table generated by MaxQuant was filtered for contaminants and highly abundant proteins (such as keratins, tubulins, and ribosomal proteins), reverse hits, number of unique peptides (>0), and number of peptides (>*1*). Subsequently, analysis and filtering using the CRAPome

software suite were performed essentially as described (*24*). To discriminate bona fide protein interactors of βTrCP2 from the background, we set a SAINT score threshold of 0.9. Subsequently, to identify the putative substrates from this pool of specific interactors, we computed the ratios (FC ratios) for the FC-B scores between the wild-type βTrCP2 or βTrCP2-ΔF against the βTrCP2 (R447A) mutant. For the convenience of FC-ratio estimation, missing values in FC-B scores were imputed with a minimal value of 0.1.

For analysis using the Perseus software suite for identifying specific interactors of wild-type βTrCP2 or βTrCP2-ΔF against that of the βTrCP2 (R447A) mutant, *t* test–based statistics were applied on LFQ (*53*). First, the LFQ values were transformed to logarithm ($log_2$), and the resulting Gaussian distribution of the data was used for imputation of missing values by normal distribution (width = 0.3, shift = 2.5). Statistical outliers were then determined using a two-tailed *t* test followed by multiple testing corrections with a permutation-based FDR method.

### In vitro ubiquitylation assay

HEK293T cells were transfected with HA-tagged TBC1D4 along with His-Skp1, Myc-Rbx1, Cul1, and either FLAG-βTrCP1 or FLAG–βTrCP1-ΔF using the PEI transfection method. Forty-eight hours after transfection, MG132 (10 μM) was added and cells were harvested after 5 hours. Immunoprecipitation using FLAG resin was carried out as described above. Two final washes with QA buffer [10 mM tris-Cl (pH 7.5), 100 mM KCl, 1 mM $MgCl_2$, 0.1 mM $CaCl_2$, 1 mM DTT] were performed in addition to the last washes with lysis buffer. After the last wash, 10 μl of the in vitro ubiquitylation mix containing 50 mM tris (pH 7.6), 5 mM $MgCl_2$, 0.6 mM DTT, 2 mM ATP (adenosine 5′-triphosphate), E1 (1.5 ng/μl) (Boston Biochem), Ubc3 (10 ng/μl), ubiquitin (2.5 μg/μl) (Sigma-Aldrich), and 1 μM ubiquitin aldehyde was added. The reactions were incubated at 30°C for 90 min and analyzed by immunoblotting.

### Motif searching

The βTrCP-binding motif analysis was performed with the MEME and FIMO tools from the MEME software suite (*54*) (http://meme.nbcr.net). We first identified the proper binding motif and then scanned the full human proteome for its presence. The motif identification was performed with the MEME (*55*) program on the set of 47 βTrCP-binding phosphorylated peptides publicly deposited. The motif width was set to 10 residues and specified the occurrences to one per sequence for the MEME search parameters. We then used the motif and position-specific probability matrix generated by MEME to rescan the Swiss-Prot (human, July 2013, 20,277 entries) with the FIMO (*56*) program. All potential binding sites with $P \leq 1 \times 10^{-3}$ were reported.

### Phosphorylation and ubiquitination sites matching

The phosphorylation sites and ubiquitylation sites in predicted βTrCP-binding sites and corresponding proteins were mapped by searching the PhosphoSitePlus (*35*) database (www.phosphosite.org; version of 2013.12.06).

### βTrCP-binding site conservation analysis

Orthologous protein sequences and their alignments among human, mouse, zebrafish, and fruit fly were downloaded from the EggNOG database's metazoan orthologous groups (*57*) (http://eggnog.embl.de/version_4.0.beta/). The EggNOG database uses nonsupervised clustering methods to assign 1133 species to more than 700,000 orthologous protein groups. We scanned the regions of the alignment that correspond to the βTrCP-binding motif (XXXDSGXXSX) in the corresponding human protein. We then assessed whether the motif observed in a human protein was also located at the corresponding position in any of the other three species by calculating the percentage identity of the aligned protein sequences that corresponded to the positions of motifs.

## SUPPLEMENTARY MATERIALS

www.sciencesignaling.org/cgi/content/full/7/356/rs8/DC1
Fig. S1. Statistical validation using the CRAPome.
Table S1. Reported phosphodegron motifs for βTrCP2.
Table S2. List of 17,779 predicted phosphodegron motifs in human proteome.
Table S3. List of proteins isolated by immunoprecipitation, categorized by CRAPome values and sorted into the reported substrates and putative substrates (data set 2).
Table S4. List of proteins isolated by immunoprecipitation, categorized by Perseus values and sorted into the reported substrates and putative substrates.
Table S5. List of proteins isolated by immunoprecipitation comparing wild-type βTrCP2 with R447A mutant, categorized by CRAPome values (data set 1 or old data set).
Table S6. Performance metrics of reproducibility testing.
Table S7. List of putative substrates validated by both CRAPome and Perseus analyses combined with bioinformatics analyses.
References (*58*–*98*)

## REFERENCES AND NOTES


1. I. Dikic, S. Wakatsuki, K. J. Walters, Ubiquitin-binding domains—From structures to functions. *Nat. Rev. Mol. Cell Biol.* **10**, 659–671 (2009).
2. M. H. Glickman, A. Ciechanover, The ubiquitin-proteasome proteolytic pathway: Destruction for the sake of construction. *Physiol. Rev.* **82**, 373–428 (2002).
3. C. M. Pickart, M. J. Eddins, Ubiquitin: Structures, functions, mechanisms. *Biochim. Biophys. Acta* **1695**, 55–72 (2004).
4. C. A. M. Semple; RIKEN GER Group; GSL Members, The comparative proteomics of ubiquitination in mouse. *Genome Res.* **13**, 1389–1394 (2003).
5. A. W. Lau, H. Fukushima, W. Wei, The Fbw7 and beta-TRCP E3 ubiquitin ligases and their roles in tumorigenesis. *Front. Biosci.* **17**, 2197–2212 (2012).
6. S. Lipkowitz, A. M. Weissman, RINGs of good and evil: RING finger ubiquitin ligases at the crossroads of tumour suppression and oncogenesis. *Nat. Rev. Cancer* **11**, 629–643 (2011).
7. R. Fukunaga, T. Hunter, MNK1, a new MAP kinase-activated protein kinase, isolated by a novel expression screening method for identifying protein kinase substrates. *EMBO J.* **16**, 1921–1933 (1997).
8. C. F. Burande, M. L. Heuzé, I. Lamsoul, B. Monsarrat, S. Uttenweiler-Joseph, P. G. Lutz, A label-free quantitative proteomics strategy to identify E3 ubiquitin ligase substrates targeted to proteasome degradation. *Mol. Cell. Proteomics* **8**, 1719–1727 (2009).
9. A. Arabi, K. Ullah, R. M. M. Branca, J. Johansson, D. Bandarra, M. Haneklaus, J. Fu, I. Ariës, P. Nilsson, M. L. Den Boer, K. Pokrovskaja, D. Grandér, G. Xiao, S. Rocha, J. Lehtiö, O. Sangfelt, Proteomic screen reveals Fbw7 as a modulator of the NF-κB pathway. *Nat. Commun.* **3**, 976 (2012).
10. B. K. Koo, M. Spit, I. Jordens, T. Y. Low, D. E. Stange, M. van de Wetering, J. H. van Es, S. Mohammed, A. J. R. Heck, M. M. Maurice, H. Clevers, Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. *Nature* **488**, 665–669 (2012).
11. G. Xu, J. S. Paige, S. R. Jaffrey, Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nat. Biotechnol.* **28**, 868–873 (2010).
12. W. Kim, E. J. Bennett, E. L. Huttlin, A. Guo, J. Li, A. Possemato, M. E. Sowa, R. Rad, J. Rush, M. J. Comb, J. W. Harper, S. P. Gygi, Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol. Cell* **44**, 325–340 (2011).
13. K. A. Lee, L. P. Hammerle, P. S. Andrews, M. P. Stokes, T. Mustelin, J. C. Silva, R. A. Black, J. R. Doedens, Ubiquitin ligase substrate identification through quantitative proteomics at both the protein and peptide levels. *J. Biol. Chem.* **286**, 41530–41538 (2011).
14. A. Peschiaroli, N. V. Dorrello, D. Guardavaccaro, M. Venere, T. Halazonetis, N. E. Sherman, M. Pagano, $SCF^{βTrCP}$-mediated degradation of Claspin regulates recovery from the DNA replication checkpoint response. *Mol. Cell* **23**, 319–329 (2006).
15. N. V. Dorrello, A. Peschiaroli, D. Guardavaccaro, N. H. Colburn, N. E. Sherman, M. Pagano, S6K1- and βTRCP-mediated degradation of PDCD4 promotes protein translation and cell growth. *Science* **314**, 467–471 (2006).
16. D. Guardavaccaro, D. Frescas, N. V. Dorrello, A. Peschiaroli, A. S. Multani, T. Cardozo, A. Lasorella, A. Iavarone, S. Chang, E. Hernando, M. Pagano, Control of chromosome stability by the β-TrCP–REST–Mad2 axis. *Nature* **452**, 365–369 (2008).
17. J. W. Harper, M. K. M. Tan, Understanding cullin-RING E3 biology through proteomics-based substrate identification. *Mol. Cell. Proteomics* **11**, 1541–1550 (2012).
18. E. Dehan, F. Bassermann, D. Guardavaccaro, G. Vasiliver-Shamis, M. Cohen, K. N. Lowes, M. Dustin, D. C. S. Huang, J. Taunton, M. Pagano, βTrCP- and Rsk1/2-mediated degradation of BimEL inhibits apoptosis. *Mol. Cell* **33**, 109–116 (2009).
19. G. Wu, G. Xu, B. A. Schulman, P. D. Jeffrey, J. W. Harper, N. P. Pavletich, Structure of a β-TrCP1-Skp1-β-catenin complex: Destruction motif binding and lysine specificity of the $SCF^{β\text{-}TrCP1}$ ubiquitin ligase. *Mol. Cell* **11**, 1445–1456 (2003).
20. J. T. Winston, P. Strack, P. Beer-Romero, C. Y. Chu, S. J. Elledge, J. W. Harper, The $SCF^{β\text{-}TRCP}$–ubiquitin ligase complex associates specifically with phosphorylated destruction motifs in IκBα and β-catenin and stimulates IκBα ubiquitination in vitro. *Genes Dev.* **13**, 270–283 (1999).

21. F. Kruiswijk, L. Yuniati, R. Magliozzi, T. Y. Low, R. Lim, R. Bolder, S. Mohammed, C. G. Proud, A. J. R. Heck, M. Pagano, D. Guardavaccaro, Coupled activation and degradation of eEF2K regulates protein synthesis in response to genotoxic stress. *Sci. Signal.* **5**, ra40 (2012).
22. D. Frescas, M. Pagano, Deregulated proteolysis by the F-box proteins SKP2 and β-TrCP: Tipping the scales of cancer. *Nat. Rev. Cancer* **8**, 438–449 (2008).
23. H. Choi, B. Larsen, Z. Y. Lin, A. Breitkreutz, D. Mellacheruvu, D. Fermin, Z. S. Qin, M. Tyers, A. C. Gingras, A. I. Nesvizhskii, SAINT: Probabilistic scoring of affinity purification–mass spectrometry data. *Nat. Methods* **8**, 70–73 (2011).
24. D. Mellacheruvu, Z. Wright, A. L. Couzens, J. P. Lambert, N. A. St-Denis, T. Li, Y. V. Miteva, S. Hauri, M. E. Sardiu, T. Y. Low, V. A. Halim, R. D. Bagshaw, N. C. Hubner, A. Al-Hakim, A. Bouchard, D. Faubert, D. Fermin, W. H. Dunham, M. Goudreault, Z. Y. Lin, B. G. Badillo, T. Pawson, D. Durocher, B. Coulombe, R. Aebersold, G. Superti-Furga, J. Colinge, A. J. Heck, H. Choi, M. Gstaiger, S. Mohammed, I. M. Cristea, K. L. Bennett, M. P. Washburn, B. Raught, R. M. Ewing, A. C. Gingras, A. I. Nesvizhskii, The CRAPome: A contaminant repository for affinity purification–mass spectrometry data. *Nat. Methods* **10**, 730–736 (2013).
25. H. Choi, T. Glatter, M. Gstaiger, A. I. Nesvizhskii, SAINT-MS1: Protein-protein interaction scoring using label-free intensity data in affinity purification-mass spectrometry experiments. *J. Proteome Res.* **11**, 2619–2624 (2012).
26. H. Choi, G. Liu, D. Mellacheruvu, M. Tyers, A. C. Gingras, A. I. Nesvizhskii, Analyzing protein-protein interactions from affinity purification-mass spectrometry data with SAINT. *Curr. Protoc. Bioinformatics* **Chapter 8**, Unit 8.15 (2012).
27. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
28. J. Cox, I. Matic, M. Hilger, N. Nagaraj, M. Selbach, J. V. Olsen, M. Mann, A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat. Protoc.* **4**, 698–705 (2009).
29. J. Cox, N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen, M. Mann, Andromeda: A peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).
30. N. C. Hubner, M. Mann, Extracting gene function from protein–protein interactions using Quantitative BAC InteraCtomics (QUBIC). *Methods* **53**, 453–459 (2011).
31. L. Trinkle-Mulcahy, J. Andersen, Y. W. Lam, G. Moorhead, M. Mann, A. I. Lamond, Repo-Man recruits PP1γ to chromatin and is essential for cell viability. *J. Cell Biol.* **172**, 679–692 (2006).
32. B. Blagoev, I. Kratchmarova, S. E. Ong, M. Nielsen, L. J. Foster, M. Mann, A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat. Biotechnol.* **21**, 315–318 (2003).
33. I. Kratchmarova, B. Blagoev, M. Haack-Sorensen, M. Kassem, M. Mann, Mechanism of divergent growth factor effects in mesenchymal stem cell differentiation. *Science* **308**, 1472–1477 (2005).
34. J. A. Ranish, E. C. Yi, D. M. Leslie, S. O. Purvine, D. R. Goodlett, J. Eng, R. Aebersold, The study of macromolecular complexes by quantitative proteomics. *Nat. Genet.* **33**, 349–355 (2003).
35. P. V. Hornbeck, J. M. Kornhauser, S. Tkachev, B. Zhang, E. Skrzypek, B. Murray, V. Latham, M. Sullivan, PhosphoSitePlus: A comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* **40**, D261–D270 (2012).
36. R. Gutman, C. Berezin, R. Wollman, Y. Rosenberg, N. Ben-Tal, QuasiMotiFinder: Protein annotation by searching for evolutionarily conserved motif-like patterns. *Nucleic Acids Res.* **33**, W255–W261 (2005).
37. N. E. Davey, D. C. Shields, R. J. Edwards, Masking residues using context-specific evolutionary conservation significantly improves short linear motif discovery. *Bioinformatics* **25**, 443–450 (2009).
38. M. Peng, A. Scholten, A. J. R. Heck, B. van Breukelen, Identification of enriched PTM cross-talk motifs from large-scale experimental data sets. *J. Proteome Res.* **13**, 249–259 (2014).
39. R. Magliozzi, T. Y. Low, B. G. M. W. Weijts, T. Cheng, E. Spanjaard, S. Mohammed, A. van Veen, H. Ovaa, J. de Rooij, F. J. T. Zwartkruis, J. L. Bos, A. de Bruin, A. J. R. Heck, D. Guardavaccaro, Control of epithelial cell migration and invasion by the IKKβ- and CK1α-mediated degradation of RAPGEF2. *Dev. Cell* **27**, 574–585 (2013).
40. S. D'Annibale, J. Kim, R. Magliozzi, T. Y. Low, S. Mohammed, A. J. R. Heck, D. Guardavaccaro, Proteasome-dependent degradation of transcription factor activating enhancer-binding protein 4 (TFAP4) controls mitotic division. *J. Biol. Chem.* **289**, 7730–7737 (2014).
41. A. Yaron, A. Hatzubai, M. Davis, I. Lavon, S. Amit, A. M. Manning, J. S. Andersen, M. Mann, F. Mercurio, Y. Ben-Neriah, Identification of the receptor component of the IκBα-ubiquitin ligase. *Nature* **396**, 590–594 (1998).
42. F. Margottin, S. P. Bour, H. Durand, L. Selig, S. Benichou, V. Richard, D. Thomas, K. Strebel, R. Benarous, A novel human WD protein, h-βTrCP, that interacts with HIV-1 Vpu connects CD4 to the ER degradation pathway through an F-box motif. *Mol. Cell* **1**, 565–574 (1998).
43. J. Cox, M. Mann, 1D and 2D annotation enrichment: A statistical method integrating quantitative proteomics with complementary high-throughput data. *BMC Bioinformatics* **13** (Suppl. 1), S12 (2012).
44. J. Cox, M. Y. Hein, C. A. Luber, I. Paron, N. Nagaraj, M. Mann, MaxLFQ allows accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction. *Mol. Cell. Proteomics* **13**, 2513–2526 (2014).
45. J. Kim, S. D'Annibale, R. Magliozzi, T. Y. Low, P. Jansen, I. A. Shaltiel, S. Mohammed, A. J. R. Heck, R. H. Medema, D. Guardavaccaro, USP17- and $SCF^{\beta TrCP}$-regulated degradation of DEC1 controls the DNA damage response. *Mol. Cell. Biol.* **34**, 4177–4185 (2014).
46. J. de Rooij, N. M. Boenink, M. van Triest, R. H. Cool, A. Wittinghofer, J. L. Bos, PDZ-GEF1, a guanine nucleotide exchange factor specific for Rap1 and Rap2. *J. Biol. Chem.* **274**, 38125–38130 (1999).
47. Y. Liao, K. Kariya, C. D. Hu, M. Shibatohge, M. Goshima, T. Okada, Y. Watari, X. Gao, T. G. Jin, Y. Yamawaki-Kataoka, T. Kataoka, RA-GEF, a novel Rap1A guanine nucleotide exchange factor containing a Ras/Rap1A-associating domain, is conserved between nematode and humans. *J. Biol. Chem.* **274**, 37815–37820 (1999).
48. T. Ohtsuka, Y. Hata, N. Ide, T. Yasuda, E. Inoue, T. Inoue, A. Mizoguchi, Y. Takai, nRap GEP: A novel neural GDP/GTP exchange protein for Rap1 small G protein that interacts with synaptic scaffolding molecule (S-SCAM). *Biochem. Biophys. Res. Commun.* **265**, 38–44 (1999).
49. M. Larance, G. Ramm, J. Stöckli, E. M. van Dam, S. Winata, V. Wasinger, F. Simpson, M. Graham, J. R. Junutula, M. Guilhaus, D. E. James, Characterization of the role of the Rab GTPase-activating protein AS160 in insulin-regulated GLUT4 trafficking. *J. Biol. Chem.* **280**, 37803–37813 (2005).
50. C. P. Miinea, H. Sano, S. Kane, E. Sano, M. Fukuda, J. Peränen, W. S. Lane, G. E. Lienhard, AS160, the Akt substrate regulating GLUT4 translocation, has a functional Rab GTPase-activating protein domain. *Biochem. J.* **391**, 87–93 (2005).
51. D. L. Swaney, G. C. McAlister, J. J. Coon, Decision tree–driven tandem mass spectrometry for shotgun proteomics. *Nat. Methods* **5**, 959–964 (2008).
52. C. K. Frese, A. F. M. Altelaar, M. L. Hennrich, D. Nolting, M. Zeller, J. Griep-Raming, A. J. R. Heck, S. Mohammed, Improved peptide identification by targeted fragmentation using CID, HCD and ETD on an LTQ-Orbitrap Velos. *J. Proteome Res.* **10**, 2377–2388 (2011).
53. N. C. Hubner, A. W. Bird, J. Cox, B. Splettstoesser, P. Bandilla, I. Poser, A. Hyman, M. Mann, Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. *J. Cell Biol.* **189**, 739–754 (2010).
54. T. L. Bailey, M. Boden, F. A. Buske, M. Frith, C. E. Grant, L. Clementi, J. Ren, W. W. Li, W. S. Noble, MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–W208 (2009).
55. T. L. Bailey, C. Elkan, Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 28–36 (1994).
56. C. E. Grant, T. L. Bailey, W. S. Noble, FIMO: Scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018 (2011).
57. L. J. Jensen, P. Julien, M. Kuhn, C. von Mering, J. Muller, T. Doerks, P. Bork, eggNOG: Automated construction and annotation of orthologous groups of genes. *Nucleic Acids Res.* **36**, D250–D254 (2008).
58. Y. Li, K. G. Suresh Kumar, W. Tang, V. S. Spiegelman, S. Y. Fuchs, Negative regulation of prolactin receptor stability and signaling mediated by $SCF^{\beta\text{-}TrCP}$ E3 ubiquitin ligase. *Mol. Cell. Biol.* **24**, 4038–4048 (2004).
59. B. P. Zhou, J. Deng, W. Xia, J. Xu, Y. M. Li, M. Gunduz, M. C. Hung, Dual regulation of Snail by GSK-3β-mediated phosphorylation in control of epithelial–mesenchymal transition. *Nat. Cell Biol.* **6**, 931–940 (2004).
60. I. Lassot, E. Ségéral, C. Berlioz-Torrent, H. Durand, L. Groussin, T. Hai, R. Benarous, F. Margottin-Goguet, ATF4 degradation relies on a phosphorylation-dependent interaction with the $SCF^{\beta TrCP}$ ubiquitin ligase. *Mol. Cell. Biol.* **21**, 2192–2202 (2001).
61. J. Pons, N. Evrard-Todeschi, G. Bertho, J. Gharbi-Benarous, V. Tanchou, R. Benarous, J. P. Girault, Transfer-NMR and docking studies identify the binding of the peptide derived from activating transcription factor 4 to protein ubiquitin ligase β-TrCP. Competition STD-NMR with β-catenin. *Biochemistry* **47**, 14–29 (2008).
62. H. Chen, H. Ma, H. Inuzuka, J. Diao, F. Lan, Y. G. Shi, W. Wei, Y. Shi, DNA damage regulates UHRF1 stability via the $SCF^{\beta\text{-}TrCP}$ E3 ligase. *Mol. Cell. Biol.* **33**, 1139–1148 (2013).
63. D. V. Hansen, A. V. Loktev, K. H. Ban, P. K. Jackson, Plk1 regulates activation of the anaphase promoting complex by phosphorylating and triggering $SCF^{\beta TrCP}$-dependent destruction of the APC Inhibitor Emi1. *Mol. Biol. Cell* **15**, 5623–5634 (2004).
64. J. Jin, T. Shirogane, L. Xu, G. Nalepa, J. Qin, S. J. Elledge, J. W. Harper, $SCF^{\beta\text{-}TRCP}$ links Chk1 signaling to degradation of the Cdc25A protein phosphatase. *Genes Dev.* **17**, 3062–3074 (2003).
65. R. E. Amir, H. Haecker, M. Karin, A. Ciechanover, Mechanism of processing of the NF-κB2 p100 precursor: Identification of the specific polyubiquitin chain-anchoring lysine residue and analysis of the role of NEDD8-modification on the $SCF^{\beta\text{-}TrCP}$ ubiquitin ligase. *Oncogene* **23**, 2540–2547 (2004).
66. Y. Kee, J. M. Kim, A. D. D'Andrea, Regulated degradation of FANCM in the Fanconi anemia pathway during mitosis. *Genes Dev.* **23**, 555–560 (2009).
67. B. Zhao, L. Li, K. Tumaneng, C. Y. Wang, K. L. Guan, A coordinated phosphorylation by Lats and CK1 regulates YAP stability through $SCF^{\beta\text{-}TRCP}$. *Genes Dev.* **24**, 72–85 (2010).

68. T. Shirogane, J. Jin, X. L. Ang, J. W. Harper, SCF$^{\beta\text{-TRCP}}$ controls clock-dependent transcription via casein kinase 1-dependent degradation of the mammalian period-1 (Per1) protein. *J. Biol. Chem.* **280**, 26863–26872 (2005).
69. S. Chowdhry, Y. Zhang, M. McMahon, C. Sutherland, A. Cuadrado, J. D. Hayes, Nrf2 is controlled by two distinct β-TrCP recognition motifs in its Neh6 domain, one of which can be modulated by GSK-3 activity. *Oncogene* **32**, 3765–3781 (2013).
70. M. Hayakawa, H. Kitagawa, K. Miyazawa, M. Kitagawa, K. Kikugawa, The FWD1/β-TrCP-mediated degradation pathway establishes a "turning off switch" of a Cdc42 guanine nucleotide exchange factor, FGD1. *Genes Cells* **10**, 241–251 (2005).
71. K. Ohsaki, K. Oishi, Y. Kozono, K. Nakayama, K. I. Nakayama, N. Ishida, The role of β-TrCP1 and β-TrCP2 in circadian rhythm generation by mediating degradation of clock protein PER2. *J. Biochem.* **144**, 609–618 (2008).
72. S. Reischl, K. Vanselow, P. O. Westermark, N. Thierfelder, B. Maier, H. Herzel, A. Kramer, β-TrCP1-mediated degradation of PERIOD2 is essential for circadian dynamics. *J. Biol. Rhythms* **22**, 375–386 (2007).
73. A. C. da Silva Almeida, G. J. Strous, A. G. S. H. van Rossum, βTrCP controls GH receptor degradation via two different motifs. *Mol. Endocrinol.* **26**, 165–177 (2012).
74. Y. Zhao, X. Xiong, Y. Sun, DEPTOR, an mTOR inhibitor, is a physiological substrate of SCF$^{\beta\text{TrCP}}$ E3 ubiquitin ligase and regulates survival and autophagy. *Mol. Cell* **44**, 304–316 (2011).
75. A. C. Burrows, J. Prokop, M. K. Summers, Skp1-Cul1-F-box ubiquitin ligase (SCF$^{\beta\text{TrCP}}$)-mediated destruction of the ubiquitin-specific protease USP37 during $G_2$-phase promotes mitotic entry. *J. Biol. Chem.* **287**, 39021–39029 (2012).
76. K. G. S. Kumar, W. Tang, A. K. Ravindranath, W. A. Clark, E. Croze, S. Y. Fuchs, SCF$^{\text{HOS}}$ ubiquitin ligase mediates the ligand-induced down-regulation of the interferon-α receptor. *EMBO J.* **22**, 5480–5490 (2003).
77. N. Watanabe, H. Arai, Y. Nishihara, M. Taniguchi, N. Watanabe, T. Hunter, H. Osada, M-phase kinases induce phospho-dependent ubiquitination of somatic Wee1 by SCF$^{\beta\text{-TrCP}}$. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 4419–4424 (2004).
78. S. Wei, H. C. Chuang, W. C. Tsai, H. C. Yang, S. R. Ho, A. J. Paterson, S. K. Kulp, C. S. Chen, Thiazolidinediones mimic glucose starvation in facilitating Sp1 degradation through the up-regulation of b-transducin repeat-containing protein. *Mol. Pharmacol.* **76**, 47–57 (2009).
79. S. Hatakeyama, M. Kitagawa, K. Nakayama, M. Shirane, M. Matsumoto, K. Hattori, H. Higashi, H. Nakano, K. Okumura, K. Onoé, R. A. Good, Ubiquitin-dependent degradation of IκBα is mediated by a ubiquitin ligase Skp1/Cul 1/F-box protein FWD1. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 3859–3863 (1999).
80. A. Seki, J. A. Coppinger, H. Du, C. Y. Jang, J. R. Yates III, G. Fang, Plk1- and β-TrCP–dependent degradation of Bora controls mitotic progression. *J. Cell Biol.* **181**, 65–78 (2008).
81. W. Huang, X. Lv, C. Liu, Z. Zha, H. Zhang, Y. Jiang, Y. Xiong, Q. Y. Lei, K. L. Guan, The N-terminal phosphodegron targets TAZ/WWTR1 protein for SCF$^{\beta\text{-TrCP}}$-dependent degradation in response to phosphatidylinositol 3-kinase inhibition. *J. Biol. Chem.* **287**, 26245–26253 (2012).
82. C. Y. Liu, Z. Y. Zha, X. Zhou, H. Zhang, W. Huang, D. Zhao, T. Li, S. W. Chan, C. J. Lim, W. Hong, S. Zhao, Y. Xiong, Q. Y. Lei, K. L. Guan, The Hippo tumor pathway promotes TAZ degradation by phosphorylating a phosphodegron and recruiting the SCF$^{\beta\text{-TrCP}}$ E3 ligase. *J. Biol. Chem.* **285**, 37159–37169 (2010).
83. L. Meyer, B. Deau, H. Forejtníková, D. Duménil, F. Margottin-Goguet, C. Lacombe, P. Mayeux, F. Verdier, β-Trcp mediates ubiquitination and degradation of the erythropoietin receptor and controls cell proliferation. *Blood* **109**, 5215–5222 (2007).
84. V. Lang, J. Janzen, G. Z. Fischer, Y. Soneji, S. Beinke, A. Salmeron, H. Allen, R. T. Hay, Y. Ben-Neriah, S. C. Ley, βTrCP-mediated proteolysis of NF-κB1 p105 requires phosphorylation of p105 serines 927 and 932. *Mol. Cell. Biol.* **23**, 402–413 (2003).
85. F. Mantovani, L. Banks, Regulation of the discs large tumor suppressor by a phosphorylation-dependent interaction with the β-TrCP ubiquitin ligase receptor. *J. Biol. Chem.* **278**, 42477–42486 (2003).
86. D. Setoyama, M. Yamashita, N. Sagata, Mechanism of degradation of CPEB during *Xenopus* oocyte maturation. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 18001–18006 (2007).
87. X. L. Ang, D. P. Seeburg, M. Sheng, J. W. Harper, Regulation of postsynaptic RapGAP SPAR by Polo-like kinase 2 and the SCF$^{\beta\text{-TRCP}}$ ubiquitin ligase in hippocampal neurons. *J. Biol. Chem.* **283**, 29424–29432 (2008).
88. G. Guderian, J. Westendorf, A. Uldschmid, E. A. Nigg, Plk4 *trans*-autophosphorylation regulates centriole number by controlling βTrCP-mediated degradation. *J. Cell Sci.* **123**, 2163–2169 (2010).
89. H. Fukushima, K. Ogura, L. Wan, Y. Lu, V. Li, D. Gao, P. Liu, A. W. Lau, T. Wu, M. W. Kirschner, H. Inuzuka, W. Wei, SCF-mediated Cdh1 degradation defines a negative feedback system that coordinates cell-cycle progression. *Cell Rep.* **4**, 803–816 (2013).
90. N. Popov, C. Schülein, L. A. Jaenicke, M. Eilers, Ubiquitylation of the amino terminus of Myc by SCF$^{\beta\text{-TrCP}}$ antagonizes SCF$^{\text{Fbw7}}$-mediated turnover. *Nat. Cell Biol.* **12**, 973–981 (2010).
91. M. L. Li, J. Defren, G. Brewer, Hsp27 and F-box protein β-TrCP promote degradation of mRNA decay factor AUF1. *Mol. Cell. Biol.* **33**, 2315–2326 (2013).
92. Y. Kanemori, K. Uto, N. Sagata, β-TrCP recognizes a previously undescribed nonphosphorylated destruction motif in Cdc25A and Cdc25B phosphatases. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 6279–6284 (2005).
93. H. Inuzuka, A. Tseng, D. Gao, B. Zhai, Q. Zhang, S. Shaik, L. Wan, X. L. Ang, C. Mock, H. Yin, J. M. Stommel, S. Gygi, G. Lahav, J. Asara, Z. X. J. Xiao, W. G. Kaelin, J. W. Harper, W. Wei, Phosphorylation by casein kinase I promotes the turnover of the Mdm2 oncoprotein via the SCF$^{\beta\text{-TRCP}}$ ubiquitin ligase. *Cancer Cell* **18**, 147–159 (2010).
94. J. Zhong, K. Ogura, Z. Wang, H. Inuzuka, Degradation of the transcription factor Twist, an oncoprotein that promotes cancer metastasis. *Discov. Med.* **15**, 7–15 (2013).
95. Q. Ding, X. He, J. M. Hsu, W. Xia, C. T. Chen, L. Y. Li, D. F. Lee, J. C. Liu, Q. Zhong, X. Wang, M. C. Hung, Degradation of Mcl-1 by β-TrCP mediates glycogen synthase kinase 3-induced tumor suppression and chemosensitization. *Mol. Cell. Biol.* **27**, 4006–4017 (2007).
96. P. C. Chu, H. C. Chuang, S. K. Kulp, C. S. Chen, The mRNA-stabilizing factor HuR protein is targeted by β-TrCP protein for degradation in response to glycolysis inhibition. *J. Biol. Chem.* **287**, 43639–43650 (2012).
97. E. Estrabaud, I. Lassot, G. Blot, E. Le Rouzic, V. Tanchou, E. Quemeneur, L. Daviet, F. Margottin-Goguet, R. Benarous, RASSF1C, an isoform of the tumor suppressor RASSF1A, promotes the accumulation of β-catenin by interacting with βTrCP. *Cancer Res.* **67**, 1054–1061 (2007).
98. G. Coadou, J. Gharbi-Benarous, S. Megy, G. Bertho, N. Evrard-Todeschi, E. Segeral, R. Benarous, J. P. Girault, NMR studies of the phosphorylation motif of the HIV-1 protein Vpu bound to the F-box protein β-TrCP. *Biochemistry* **42**, 14741–14751 (2003).


**Funding:** This work has been supported by the Netherlands Proteomics Centre, the Netherlands Organization for Scientific Research (NWO) supporting the Roadmap-embedded large-scale proteomics facility Proteins@Work (project 184.032.201), and the PRIME-XS project grant agreement number 262067 supported by the European Community's Seventh Framework Programme (FP7/2007-2013) to A.J.R.H. **Author contributions:** T.Y.L. designed, performed, and analyzed the proteomics experiments; M.P. devised and optimized the bioinformatics workflow and analyzed the proteomics data; R.M. created and optimized the different bait constructs and transfected cell lines, and performed coimmunoprecipitation and all biochemical validations. T.Y.L., D.G., and A.J.R.H. conceived the idea for this study. S.M., D.G., and A.J.R.H. supervised the study and co-wrote the manuscript with T.Y.L. **Competing interests:** The authors declare that they have no competing financial interests. **Data and materials availability:** The MS proteomics data (data set 1) have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the data set identifier PXD001088. The reanalyzed GeLC-MS data set (data set 2) for reproducibility evaluation is available with the data set identifier PXD001224.

Submitted 8 September 2014
Accepted 21 November 2014
Final Publication 16 December 2014
10.1126/scisignal.2005882
**Citation:** T. Y. Low, M. Peng, R. Magliozzi, S. Mohammed, D. Guardavaccaro, A. J. R. Heck, A systems-wide screen identifies substrates of the SCF$^{\beta\text{TrCP}}$ ubiquitin ligase. *Sci. Signal.* **7**, rs8 (2014).

**A systems-wide screen identifies substrates of the SCF$^{\beta TrCP}$ ubiquitin ligase**

Teck Yew Low, Mao Peng, Roberto Magliozzi, Shabaz Mohammed, Daniele Guardavaccaro and Albert J. R. Heck (December 16, 2014)

*Science Signaling* **7** (356), rs8. [doi: 10.1126/scisignal.2005882]

Science Signaling
AAAS

www.sciencesignaling.org/cgi/content/full/7/356/rs8/DC1

# Supplementary Materials for

## A systems-wide screen identifies substrates of the SCF$^{\beta TrCP}$ ubiquitin ligase

Teck Yew Low, Mao Peng, Roberto Magliozzi, Shabaz Mohammed,
Daniele Guardavaccaro, Albert J. R. Heck*

*Corresponding author. E-mail: a.j.r.heck@uu.nl

Published 16 December 2014, *Sci. Signal.* **7**, rs8 (2014)
DOI: 10.1126/scisignal.2005882

**This PDF file includes:**



**Other Supplementary Material for this manuscript includes the following:**
(available at www.sciencesignaling.org/cgi/content/full/7/356/rs8/DC1)

Table S2 (Microsoft Excel format). List of 17,779 predicted phosphodegron motifs in human proteome.
Table S3 (Microsoft Excel format). List of proteins isolated by immunoprecipitation, categorized by CRAPome values and sorted into the reported substrates and putative substrates (data set 2).
Table S4 (Microsoft Excel format). List of proteins isolated by immunoprecipitation, categorized by Perseus values and sorted into the reported substrates and putative substrates.
Table S5 (Microsoft Excel format). List of proteins isolated by immunoprecipitation comparing wild-type βTrCP2 with R447A mutant, categorized by CRAPome values (data set 1 or old data set).
Table S7 (Microsoft Excel format). List of putative substrates validated by both CRAPome and Perseus analyses combined with bioinformatics analyses.

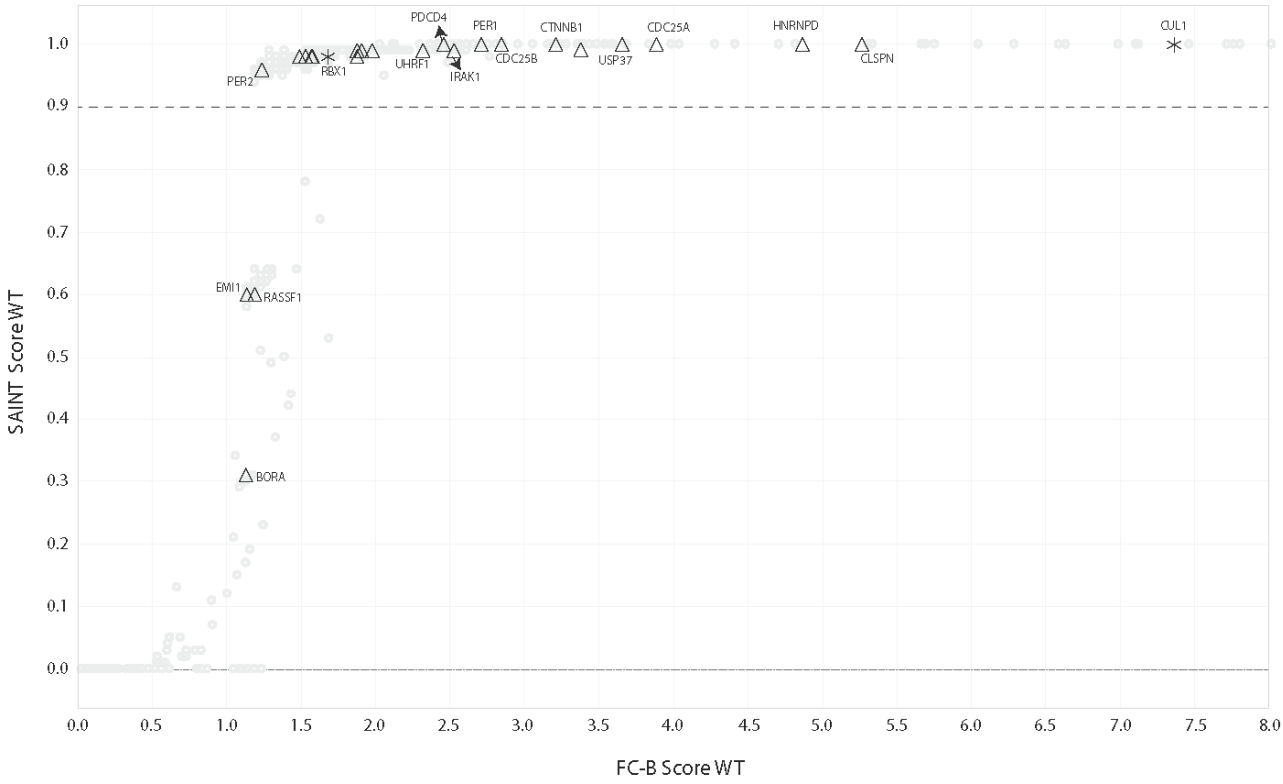


**Fig. S1. Statistical validation using the CRAPome.** The distribution of the SAINT score and the fold-change (FC) score for wild-type βTrCP (WT) versus empty vector immunoprecipitation. A total of 19 (Δ) out of the 27 reported substrates are identified and each carries a SAINT score of ≥ 0.90. SCF subunits (*), such as βTrCP2, Skp1, Rbx1 and Cul1, also have SAINT scores of ≥ 0.90.

**Table S1. Reported phosphodegron motifs for βTrCP2.** Forty-seven phosphodegron motifs were extracted from previously reported βTrCP2 substrates and used for building a consensus motif with MEME. The resulting consensus motif was subsequently used for FIMO analysis of the human SwissProt database to identify similar motifs in the human proteome. The corresponding FIMO p-value and the ranking, taking into account all the putative degrons generated from the human proteome (FIMO p ≤ 1e-3), are also included. (* - degron is not WD40-binding; NA- not identified in FIMO analysis).

| | UniProt Acc. No. | Gene Symbols | Degron Motifs | Degron Range | FIMO p-value | -Log10 (p-value) | Ranking based on p-value |
|---|---|---|---|---|---|---|---|
| *1* | O00418 | eEF2K *(21)* | NSGDSGYPSE | 437-446 | 7.07E-09 | 8.151 | 1 |
| *2* | P16471 | PRLR *(58)* | PDTDSGRGSC | 345-354 | 1.80E-08 | 7.745 | 2 |
| *3* | O95863 | SNAI1/SNAIL *(59)* | SDEDSGKGSQ | 92-101 | 1.89E-08 | 7.724 | 3 |
| *4* | P18848 | ATF4 *(60, 61)* | SDNDSGICMS | 215-224 | 2.83E-08 | 7.548 | 5 |
| *5* | Q96T88 | UHRF1 *(62)* | SDTDSGCCLG | 91-100 | 3.80E-08 | 7.420 | 9 |
| *6* | P35222 | CTNNB1 *(19)* | SYLDSGIHSG | 29-38 | 6.55E-08 | 7.184 | 11 |
| *7* | Q9UKT4 | FBXO5/EMI1 *(63)* | LYEDSGYSSF | 141-150 | 1.42E-07 | 6.848 | 16 |
| *8* | P30304 | CDC25A *(64)* | ESTDSGFCLD | 78-87 | 2.69E-07 | 6.570 | 26 |
| *9* | Q15653 | NFKBIB/IKBB *(65)* | EWCDSGLGSL | 15-24 | 2.93E-07 | 6.533 | 28 |
| *10* | Q8IYD8 | FANCM *(66)* | NVLDSGYNSF | 940-949 | 2.96E-07 | 6.529 | 29 |
| *11* | Q53EL6 | PDCD4 *(15)* | SSRDSGRGDS | 67-76 | 3.18E-07 | 6.498 | 31 |
| *12* | E3WEB5 | YAP1 *(67)* | DESTDSGLSMS | 396-405 | 3.33E-07 | 6.478 | 32 |
| *13* | Q9HAW4 | CLSPN *(14)* | SPSDSGQGSY | 26-35 | 3.49E-07 | 6.457 | 34 |
| *14* | O00221 | NFKBIE *(65)* | SQYDSGIESL | 153-162 | 5.90E-07 | 6.229 | 50 |
| *15* | O15534 | PER1 *(68)* | NPSTSGCSSE | 118-127 | 6.10E-07 | 6.215 | 52 |
| *16* | Q16236 | NFE2L2/NRF2 *(69)* | NDSDSGISLN | 343-348 | 6.88E-07 | 6.162 | 55 |
| *17* | Q13127 | REST *(16)* | IDEDEGIHSH | 1005-1014 | 9.28E-07 | 6.032 | 68 |
| *18* | P98174 | FGD1 *(70)* | PNRDSGIDSI | 279-288 | 9.38E-07 | 6.028 | 69 |
| *19* | O43521 | BCL2L11 *(18)* | SRSSSGYFSF | 90-99 | 9.99E-07 | 6.000 | 75 |
| *20* | O15055 | PER2 *(71, 72)* | NPSTSGCSSD | 91-100 | 1.12E-06 | 5.951 | 80 |
| *21* | P10912 | GHR *(73)* | KDGDSGRTSC | 380-389 | 1.14E-06 | 5.943 | 82 |
| *22* | Q8TB45 | DEPTOR *(74)* | SCGSSGYFSS | 283-292 | 1.20E-06 | 5.921 | 92 |
| *23* | Q86T82 | USP37 *(75)* | SDEDSGNEDV | 854-863 | 1.44E-06 | 5.842 | 100 |
| *24* | P17181 | IFNAR1 *(76)* | TSQDSGNYSN | 531-540 | 1.68E-06 | 5.775 | 110 |
| *25* | P30291 | WEE1_Site2 *(77)* | SWEEEGFGSS | 113-122 | 1.78E-06 | 5.750 | 113 |
| *26* | P08047 | SP1 *(78)* | LPLDSGAGSE | 724-733 | 2.18E-06 | 5.662 | 132 |
| *27* | P25963 | NFKBIA/IKBA *(65, 79)* | DRHDSGLDSM | 28-37 | 2.44E-06 | 5.613 | 148 |
| *28* | Q6PGQ7 | BORA *(80)* | IQMDSGYNTQ | 493-502 | 2.61E-06 | 5.583 | 159 |
| *29* | Q9GZV5 | WWTR1/TAZ *(81, 82)* | QSTDSGLGLG | 310-319 | 3.21E-06 | 5.493 | 192 |
| *30* | P19235 | EPOR *(83)* | VVSDSGISTD | 458-467 | 3.61E-06 | 5.442 | 211 |
| *31* | P19838 | NFKB1 *(84)* | SVCDSGVETS | 923-932 | 3.64E-06 | 5.439 | 213 |
| *32* | Q00653 | NFKB2 *(65)* | VKEDSAYGSQ | 862-871 | 3.78E-06 | 5.423 | 222 |
| *33* | Q12959 | DLG1 *(85)* | KTKDSGLPSQ | 594-603 | 4.28E-06 | 5.369 | 245 |
| *34* | Q91572 | CPEB1-a *(86)* | DSDTSGFSSG | 187-196 | 4.48E-06 | 5.349 | 249 |
| *35* | Q8C0T5 | SIPA1L1/SPIN *(87)* | TSADSGIDTA | 1301-1310 | 8.92E-06 | 5.049 | 402 |
| *36* | O00444 | PLK4 *(88)* | DSIDSGHATI | 281-290 | 3.79E-05 | 4.421 | 1213 |
| *37* | K7EQT1 | FZR1/CDH1 *(89)* | SSPDDGNDVS | 137-146 | 4.51E-05 | 4.346 | 1401 |
| *38* | P01106 | MYC *(90)* | KRSESGSPSA | 275-284 | 4.63E-05 | 4.334 | 1425 |
| *39* | Q14103 | HNRPD/AUF1 *(91)* | HSNSSPRHSE | 79-88 | 4.72E-05 | 4.326 | 1450 |
| *40* | P30305 | CDC25B *(92)* | TEEDDGFVDI | 265-274 | 1.11E-04 | 3.955 | 2930 |
| *41* | P04637 | TP53 *(93)* | EPGGSRAHSS | 358-367 | 1.49E-04 | 3.827 | 3776 |
| *42* | Q15672 | TWST1 *(94)* | PADDSLSNSE | 21-30 | 1.49E-04 | 3.827 | 3777 |
| *43* | Q07820 | MCL1 *(95)* | TSTDGSLPST | 154-163 | 2.21E-04 | 3.656 | 5217 |
| *44* | P30291 | WEE1 *(77)* | TGEDSAFQEP | 49-58 | 3.93E-04 | 3.406 | 8299 |
| *45* | Q15717 | ELAVL1/huR *(96)* | TNYEEAAMAI | 293-302 | 8.69E-04 | 3.06 | 15928 |
| *46* | Q9NS23 | RASSF1-C *(97)* | *STTSSGYCSQ | 15-24 | NA | NA | NA |
| *47* | P05923 | Vpu *(98)* | RAEDSGNESE | 48-57 | NA | NA | NA |

**Table S6. Performance metrics of reproducibility testing.** A comparison of performance and reproducibility metrics of the current AP-MS (dataset 1) and those obtained from past experiments (dataset 2). Dataset 2 was acquired using identical AP methods, but was further separated using SDS-PAGE, followed by LC-MS/MS analysis with an older version Orbitrap mass analyzer. Both datasets were analyzed with the same software and settings.

| | AP-MS Dataset 1 | AP-MS Dataset 2 | Overlapping Proteins | Percentage Overlapping (%) |
|---|---|---|---|---|
| **Number Proteins Identified** | 1627 | 725 | 531 | 73.24 |
| **Number of Proteins passing CRAPome filters (putative substrates)** | 221 | 47 | 31 | 65.96 |
| **Validated putative substrates (eEF2K, RAPGEF2 and TFAP4) passing CRAPome filters** | 3 | 3 | 3 | 100.00 |
| **Reported Substrates Identified** | 27 | 25 | 21 | 84.00 |
| **Reported Substrates Identified and passing CRAPome filters** | 20 | 13 | 12 | 92.31 |

**Supplementary tables supplied as Excel files**

**Table S2. List of 17,779 predicted phosphodegron motifs in human proteome.** A total of 17,779 putative degron sequences were extracted from the human proteome with FIMO ($p \leq 1e\text{-}3$) on the basis of their similarity to the phosphodegron consensus motif obtained from MEME. Data include the FIMO p-values, FIMO q-values, evolutionary conservation scores, and reported phosphorylation and ubiquitylation site information derived from PhosphoSitePlus. (2005882_tableS2.xlsx)

**Table S3. List of the proteins isolated by immunoprecipitation, categorized by CRAPome values and sorted into the reported substrates and putative substrates (data set 2). (A)** A total of 1,626 proteins identified in the AP-MS experiments from three different βTrCP2 constructs, together with their individual SAINT probability scores and also FC-A and FC-B scores computed with the CRAPome software suite. (**B**) A subtable exported from Table S3A of the 27 reported substrates identified in the AP-MS experiments. (**C**) A subtable exported from Table S3A of the 338 putative substrates (after validation by the CRAPome workflow) from the proteins identified in the AP-MS experiments. (2005882_table S3.xlsx)

**Table S4. List of the proteins isolated by immunoprecipitation, categorized by Perseus values and sorted into the reported substrates and putative substrates.** (**A**) A total of 1,626 proteins identified in the AP-MS experiments from three different βTrCP2 constructs together with their t-test significance, t-test difference, and p-values after validation with the Perseus workflow. (**B**) A subtable exported from Table S4A of the 27 reported substrates identified in the AP-MS experiments. (**C**) A sub-table exported from Table S3A of the 317 putative substrates (after validation by the Perseus workflow) from the proteins identified in the AP-MS experiments. 2005882_table S4.xlsx

**Table S5. List of proteins isolated by immunoprecipitation comparing wild-type βTrCP2 with R447A mutant, categorized by CRAPome values (data set 1 or old data set).** A total of 725 proteins were identified in the AP-MS experiments of βTrCP2 (wildtype and R447A mutant) and collated with their individual SAINT probability scores and FC-A and FC-B scores computed with the CRAPome software suite. 2005882_table S5.xlsx

**Table S7. List of putative substrates validated by both CRAPome and Perseus analyses combined with bioinformatics analyses.** (**A**) The 221 putative substrates validated by both CRAPome and Perseus analyses after combining with bioinformatics prediction in which each candidate must contain a putative degron motif (FIMO $p \leq 1E\text{-}3$) with a conservation score $\geq 0.7$. (**B**) The 221 putative substrates divided into 6 bins according to their FIMO p-values. 2005882_table S7.xlsx