Molecular Cell, Volume *83*

Supplemental information

Ubiquitin-independent proteasomal degradation driven by C-degron pathways

Yaara Makaros, Anat Raiff, Richard T. Timms, Ajay R. Wagh, Mor Israel Gueta, Aizat Bekturova, Julia Guez-Haddad, Sagie Brodsky, Yarden Opatowsky, Michael H. Glickman, Stephen J. Elledge, and Itay Koren

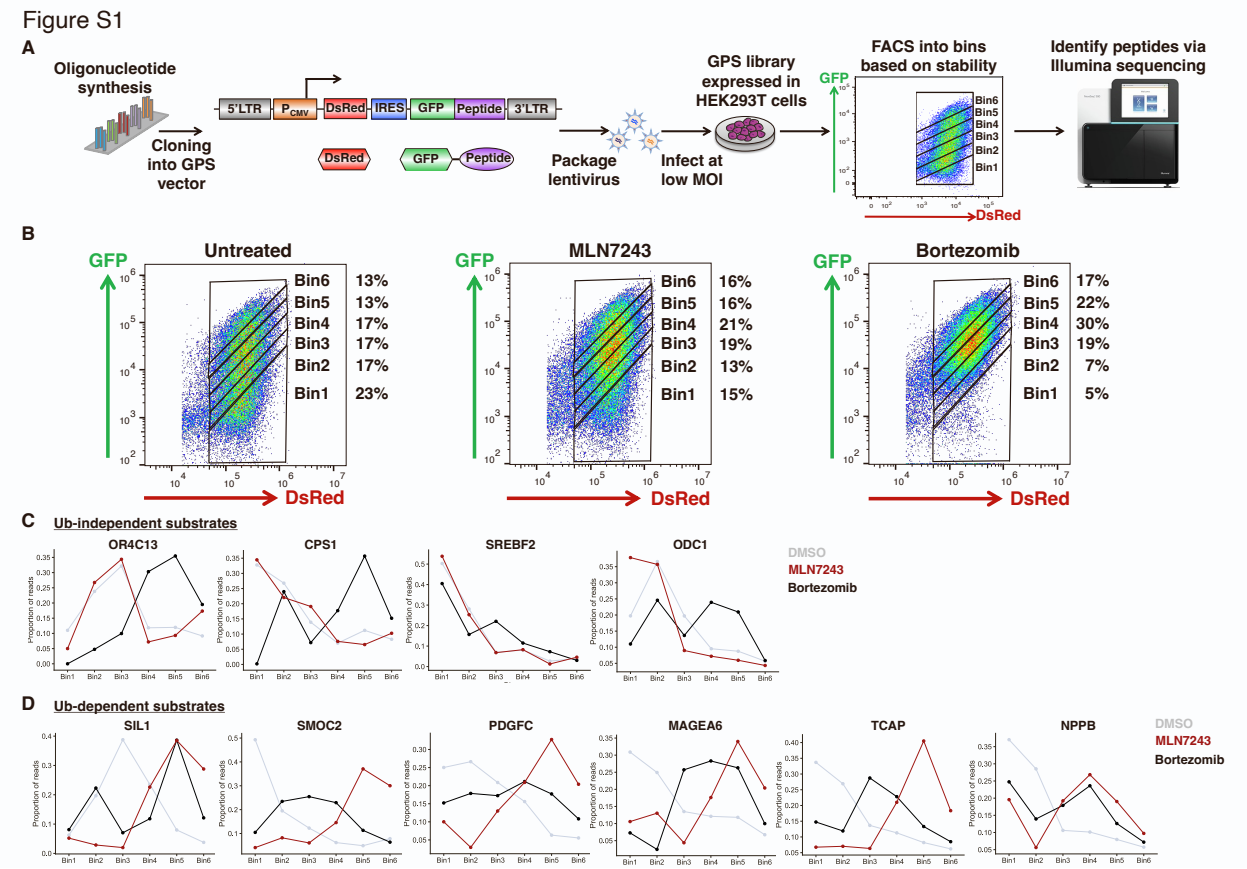

Figure S1. **GPS-peptidome screening pipeline and data analysis. Related to Figure 1.**

(A) Schematic representation of the GPS lentiviral vector and GPS-C23mer screen. GPS is based on a lentiviral construct encoding two fluorescent proteins: DsRed, which serves as an internal reference, and a GFP fusion protein that is translated from an internal ribosome entry site (IRES). As both DsRed and the GFP fusion protein are expressed from the same transcript, the GFP/DsRed ratio can be used to measure the effect of the fusion protein on the stability of GFP. Oligonucleotides encoding the last 23 residues of human proteins were synthesized on a high-density chip and cloned downstream of GFP in the GPS vector. The lentiviral library was introduced into HEK293T cells. Using fluorescence-activated cell sorting (FACS), live cells are separated in real time into sub-populations (bins) based on stability of GFP-fusions. Illumina sequencing is then used to identify and quantify peptides across the stability bins. The stability of each fusion was subsequently quantified by allowing each peptide to be assigned a protein stability index (PSI) score according to the proportion of sequencing reads in each bin (see STAR methods).

(B) Representation of FACS sort into 6 bins done on the GPS-C23mer library across the indicated conditions (untreated, MLN7243, bortezomib). Percentage of cells collected from each bin is indicated.

(C-D) The distribution of Illumina sequencing reads across the six bins in control cells (gray line) versus MLN7243 (red lines) or bortezomib (black line) treatments. The previously identified *UbInPD* substrates validated in the screen, OR4C13, CPS1 and SREBF2, and the positive control *UbInPD* substrate, ODC1 are shown in (C), and the indicated CRL substrates representing ubiquitin-dependent substrates in (D).

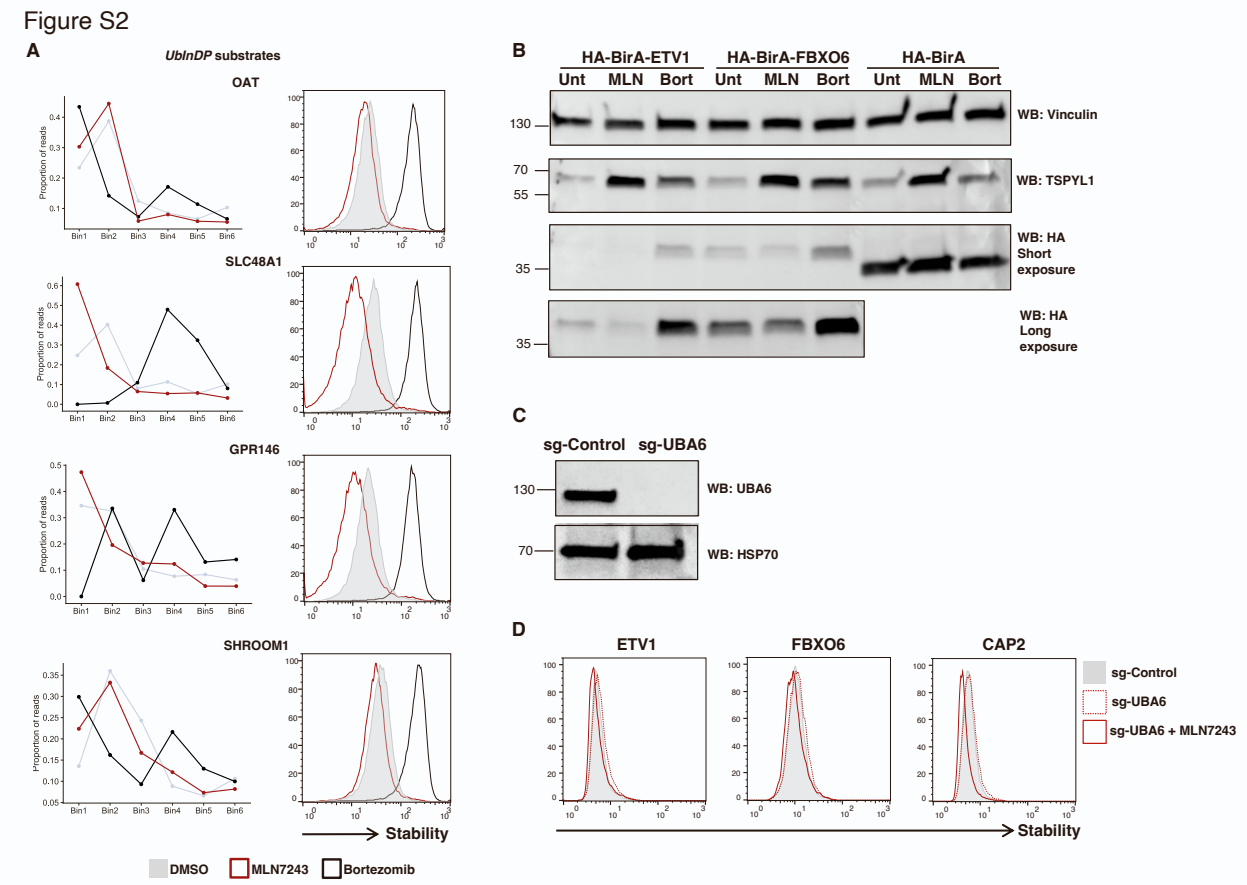

Figure S2. ***UbInPD* substrates turnover is transferable and not dependent on UBA6. Related to Figure 2.**

(A) For each of the indicated GFP-fused C23mer peptides, the GPS screen profiles (representing the distribution of sequencing reads across the six stability bins in control cells (gray line) versus MLN7243 (red lines) or bortezomib (black line) treatments) are shown on the left, and individual validation experiments as analyzed by flow cytometry following treatment with the indicated inhibitors for 7 h are shown on the right.

(B) HEK293T cells were transduced with HA-tagged BirA alone or fused to the C23mer *UbInPD* substrates FBXO6 or ETV1. Following 6 h treatment with MLN7243 (MLN) or bortezomib (Bort) cells were harvested for western blot. BirA abundance was assessed by immunoblot with HA antibody while vinculin antibody was used as a loading control. The protein levels of TSPYL1, a ubiquitin-dependent substrate are monitored with endogenous TSPYL1 antibody.

(C-D) Stability of the indicated *UbInPD* GPS-C23mer substrates is not dependent on UBA6. (C) Western blot of UBA6 knock out (KO) cells (sg-UBA6) vs. AAVSI sgRNA expressing cells (sg-

Control). UBA6 abundance was assessed by immunoblot with UBA6 antibody while HSP70 antibody was used as a loading control. (D) Flow cytometry analysis of UBA6 or control KO cells expressing the indicated GPS-C23mer in MLN7243 treated or untreated cells.

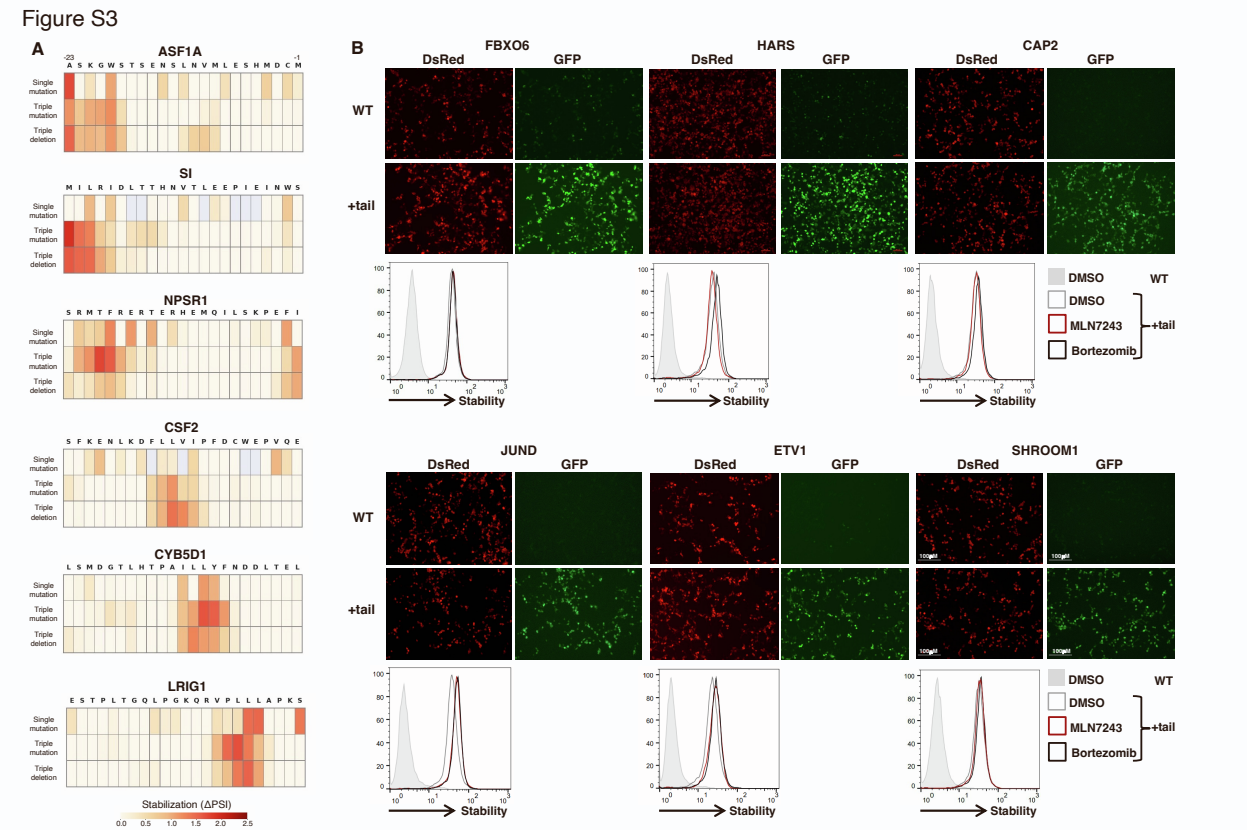

Figure S3. **Specific C-terminal motifs in *UbInPD* substrates control their instability. Related to Figure 3.**

(A) Example of minimal *UbInPD* motifs emerged from scanning mutagenesis experiment. Motifs in that case are not located at the extreme C termini. Scanning mutagenesis was performed for the indicated examples of *UbInPD* substrates. For each gene, three independent mutagenesis experiments were performed and presented in three rows. Top: single mutagenesis ('single mutation'); middle: mutagenesis of three consecutive residues ('Triple mutation'); bottom: deletion of three consecutive residues ('Triple deletion'). In each case, darker colors represent a greater degree of stabilization conferred by the mutation/deletion. Name of the gene is indicated. A universal scale of stabilization is shown at the bottom.

(B) C-terminal capping by adding extra amino acids (-DNYNEPKANQ*) to the C-terminal of the indicated *UbInPD* substrates (+tail) interferes with their proteasomal degradation. For each gene, fluorescence microscope images were taken for wild type (WT) GFP-C23mer peptide versus C23mer+tail of extra residues. DsRed is expressed uniformly and serves as reference expression control. GFP levels are low in WT C23mer but higher in C23mer+tail indicating on increased

protein stabilization of the mutant. In addition, flow cytometry used to quantify and analyze the stability based on the GFP/DsRed ratio for each gene. In that case C23mer+tail was analyzed also in response to 1 μM MLN7243 or bortezomib for 7 h. Stability levels of C23mer+tail is substantially higher than WT peptide with no further stabilization in response to the inhibitors indicating that it is no longer a proteasome substrate. Scale bar, 100 μm

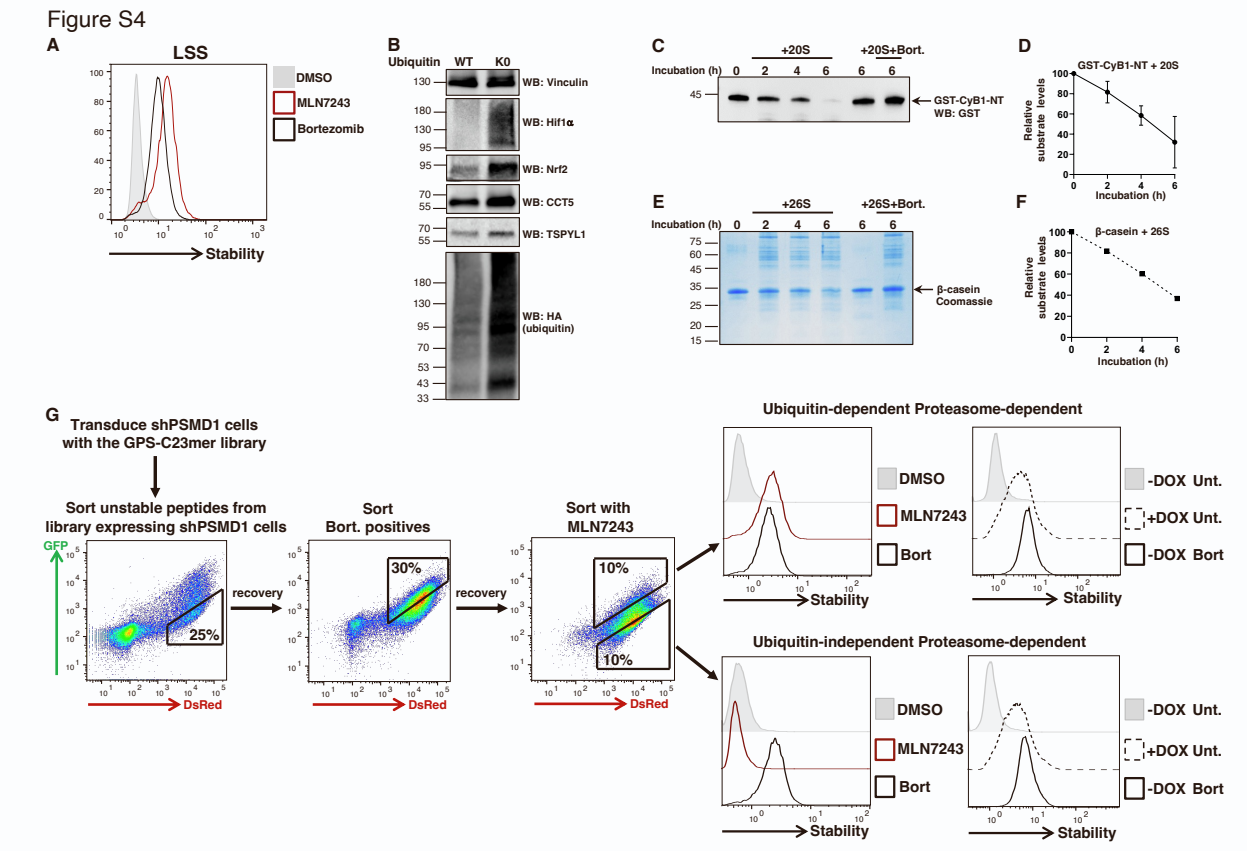

Figure S4. ***UbInPD* is mainly dependent on 26S proteasomes *in vivo*. Related to Figure 4 and Figure 5.**

(A) Cells expressing GPS-LSS, a ubiquitin-dependent proteasome-dependent peptide substrate, were treated with 1 μM MLN7243 or bortezomib for 7 h followed by flow cytometry analysis.

(B) HEK293T cells were transfected with HA-tagged wild type (WT) or lysine-less ubiquitin (K0). 24 h post-transfection cells were harvested and cell extracts analyzed in immunoblot using antibodies against ubiquitin-dependent substrates, Hif1α, Nrf2, CCT5 and TSPYL1. HA antibody is used to detect HA-ubiquitin conjugates. Vinculin served as a loading control.

(C-F) GST-fused CyB1-NT and β-casein serve as positive control for *in vitro* degradation assay with purified 20S and 26S proteasomes, respectively. Recombinant proteins were incubated for the indicated times followed by analysis by western blot with anti-GST antibody (GST-CyB1-NT) (C) or Coomassie stain (β-casein) (E). Bortezomib was added to longest incubation sample to show that degradation is dependent on proteasomes catalytic activity. Quantification of the rate of recombinant proteins proteolysis is presented for GST-CyB1-NT (D) and β-casein (F).

(G) *UbInPD* substrates are mostly degraded by 26S *in vivo*. GPS-C23mer library was transduced into shPSMD1 HEK293T cells followed by FACS separation of 25% most unstable GFP-fusions in the absence of dox. After 7 days of recovery, cells were treated with bortezomib for 6 h followed by sorting of top 30% bortezomib responsive cells. A second sort was done 7 days post sort 1 recovery. The top and bottom 10% MLN7243 positive ("Ubiquitin-dependent Proteasome-dependent") and negative ("ubiquitin-independent Proteasome-dependent") populations were separated by FACS. Following a further 7 days of recovery, the isolated populations were analyzed by flow cytometry following treatment with 1 μM MLN7243 or bortezomib for 7 h (left histogram). Finally, stability was analyzed by flow cytometry following 3 days of dox treatment (right histogram).

Figure S5

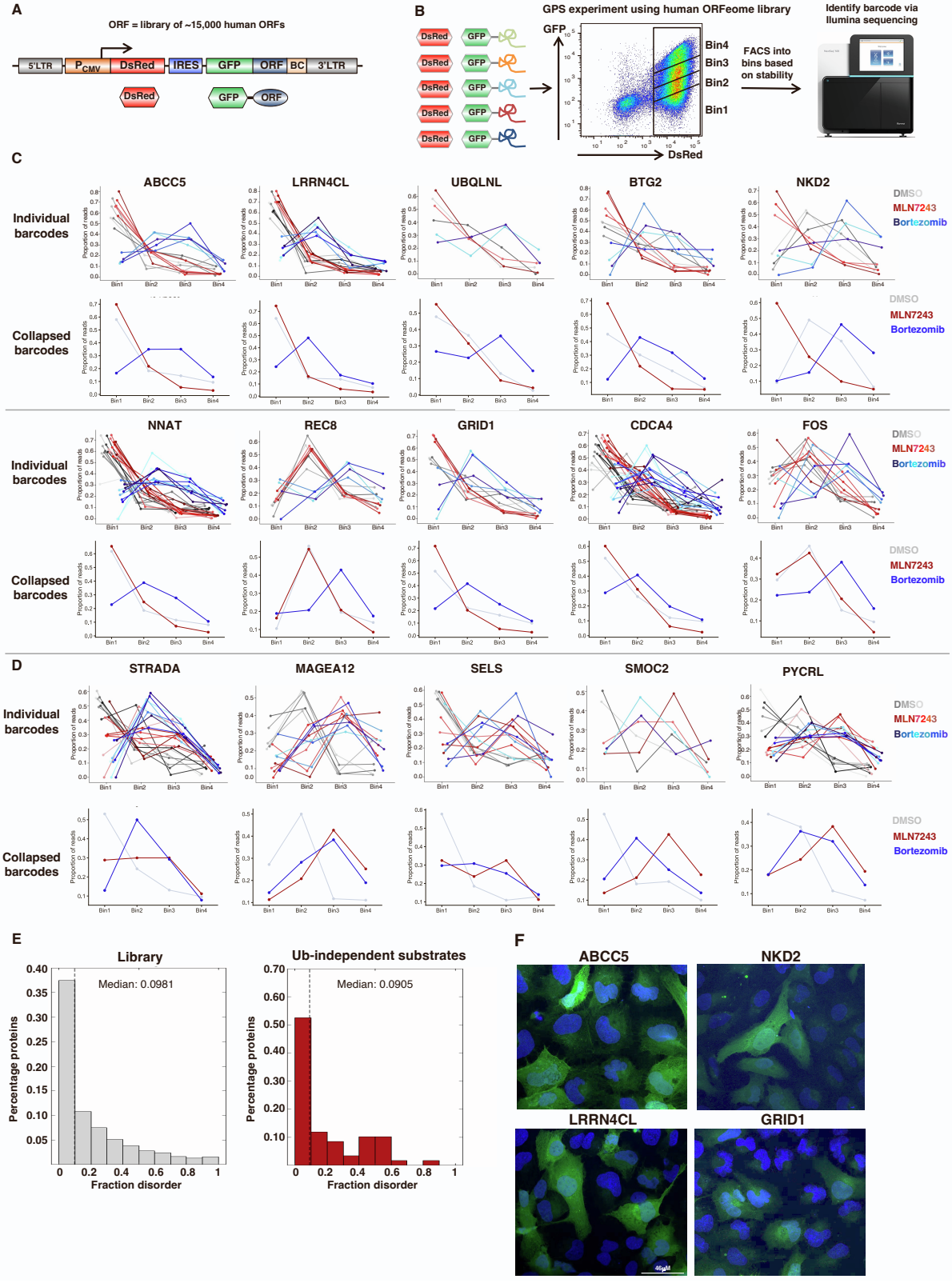
A
ORF = library of ~15,000 human ORFs
5'LTR
PCMV
DsRed
IRES
GFP
ORF
BC
3'LTR
B
GPS experiment using human ORFeome library
GFP
DsRed
Bin4
Bin3
Bin2
Bin1
FACS into bins based on stability
Identify barcode via Illumina sequencing
C
Individual barcodes
Collapsed barcodes
ABCC5
LRRN4CL
UBQLNL
BTG2
NKD2
NNAT
REC8
GRID1
CDCA4
FOS
DMSO
MLN7243
Bortezomib
Proportion of reads
Bin1
Bin2
Bin3
Bin4
D
STRADA
MAGEA12
SELS
SMOC2
PYCRL
E
Library
Median: 0.0981
Ub-independent substrates
Median: 0.0905
Percentage proteins
Fraction disorder
F
ABCC5
NKD2
LRRN4CL
GRID1

Figure S5. **GPS-ORFeome screening pipeline and data analysis. Related to Figure 6.**

(A) Schematic representation of the GPS lentiviral vector. BC, barcode.

(B) Schematic representation of the GPS-ORFeome screen.

(C) Profiles for the nine indicated candidate *UbInPD* substrates. FOS- a known *UbInPD* substrate. Upper histograms: The distribution of sequencing reads for individual barcodes attached to the same GFP-ORF fusion is presented. Each color series represents the different conditions of the experiment. Gray: DMSO (no drug control), Red: MLN7243, Blue: Bortezomib. For each of the different conditions, each color shade represents a different barcode for the same gene. Lower histograms: collapsed average of the individual barcodes in the indicated conditions.

(D) Profiles for known ubiquitin-dependent CRL substrates as in (C).

(E) Disorder tendency distribution plot of ubiquitin-independent substrates versus ORF library. Fraction of disorder was calculated as the fraction of residues predicted to be disordered (disorder tendency calculated by IUPred2>0.5) for each ORF. Median is indicated with dashed line.

(F) U2OS stably expressing the indicated GFP-fused ORFs were fixed and visualized under confocal microscopy. Images show merged channels of GFP (green) and the nuclear staining DAPI (blue). Scale bar, 46 μm

Figure S6

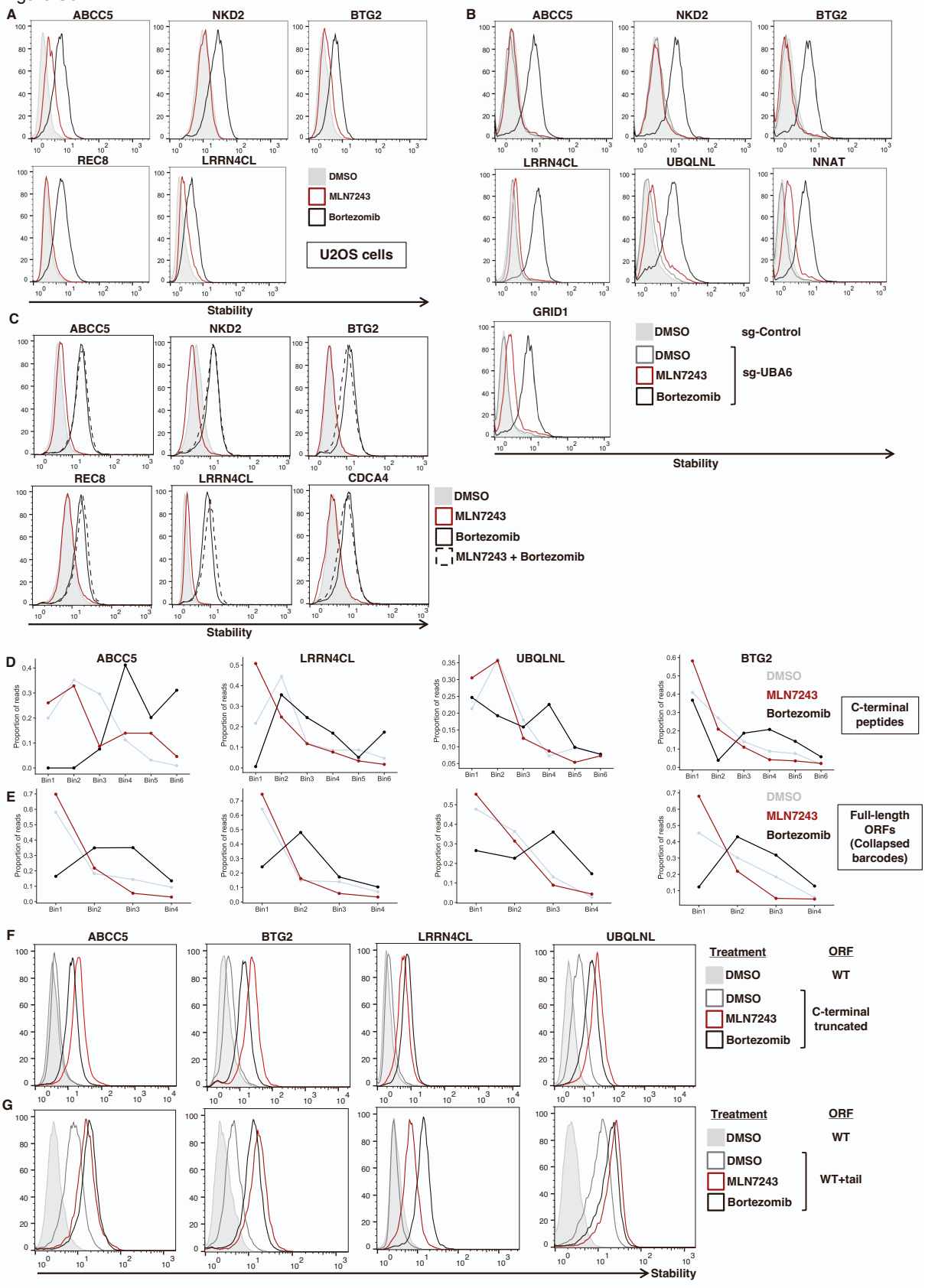
A
ABCC5
NKD2
BTG2
REC8
LRRN4CL
DMSO
MLN7243
Bortezomib
U2OS cells
Stability
B
ABCC5
NKD2
BTG2
LRRN4CL
UBQLNL
NNAT
GRID1
DMSO
sg-Control
DMSO
MLN7243
Bortezomib
sg-UBA6
Stability
C
ABCC5
NKD2
BTG2
REC8
LRRN4CL
CDCA4
DMSO
MLN7243
Bortezomib
MLN7243 + Bortezomib
Stability
D
ABCC5
LRRN4CL
UBQLNL
BTG2
Proportion of reads
Bin1
Bin2
Bin3
Bin4
Bin5
Bin6
DMSO
MLN7243
Bortezomib
C-terminal peptides
E
Full-length ORFs (Collapsed barcodes)
F
ABCC5
BTG2
LRRN4CL
UBQLNL
Treatment
ORF
DMSO
WT
DMSO
MLN7243
Bortezomib
C-terminal truncated
G
WT+tail
Stability

Figure S6. **Full-length *UbInPD* substrates instability is UBA6-independent and depends on their C-terminal. Related to Figure 6.**

(A) U2OS cells expressing the indicated GPS-ORF constructs were treated with 1 μM MLN7243 or bortezomib for 7 h and analyzed by flow cytometry.

(B) Stability of the indicated GPS-ORFs was analyzed in UBA6 KO cells (sg-UBA6) or control cells (sg-Control, expressing sg-AAVSI) in response to 1 μM MLN7243 or bortezomib for 7 h as analyzed by flow cytometry.

(C) HEK293T cells expressing the indicated GPS-ORFs were treated with 1 μM of MLN7243, 10 μM of bortezomib or combination of 1 μM MLN7243 together with 10 μM bortezomib for 6 h followed by flow cytometry analysis.

(D) GPS screen profiles for each of representative GFP-C23mer substrates showing the distribution of Illumina sequencing reads across the six bins in control cells (gray line) versus MLN7243 (red lines) or bortezomib (black line) treatments.

(E) GPS-ORFeome profiles as in (A) for collapsed barcodes across 4 bins of the same indicated genes.

(F-G) Truncation or capping of the C-terminal converts *UbInPD* substrates into ubiquitin-dependent substrates. (F) The last 23 residues of each indicated ORF were deleted. GPS construct encoding wild type (WT) or C-terminal truncated mutant ORF was expressed in cells treated with MLN7243 or bortezomib and stability was analyzed by flow cytometry. (G) Capping of the C termini was done by the addition of extra residues (-DNYNEPKANQ*) to the ORF (WT+tail). Stability was analyzed as in (F).

Figure S7


A
Ubiquilins+RAD23A/B Penta KO
UBQLN1
KO score: 84%
UBQLN2
KO score: 70%
UBQLN4
KO score: 36%
RAD23A
KO score: 51%
RAD23B
KO score: 52%
B
Ubiquilins Triple KO
UBQLN1
KO score: 80%
UBQLN2
KO score: 74%
UBQLN4
KO score: 67%
C
RAD23A/B Double KO
RAD23A
KO score: 80%
RAD23B
KO score: 91%
D
Ubiquilin single KO
UBQLN1
KO score: 94%
UBQLN2
KO score: 76%
UBQLN4
KO score: 88%
E
ABCC5
REC8
LRRN4CL
Control KO
UBQLN1 KO
UBQLN1 KO+ MLN7243
UBQLN2 KO
UBQLN2 KO+ MLN7243
UBQLN4 KO
UBQLN4 KO+ MLN7243
Stability
F
CAP2
JUND
SHROOM1
Control KO
RAD23A/B+ Ubiquilins KO
RAD23A/B+ Ubiquilins KO+ MLN7243
Stability

Figure S7. ***UbInPD*** **substrates stability analysis in shuttling factors KO cells. Related to Figure 7.**

(A-D) Shuttling factors KO efficiency analysis. Genomic DNA sequencing of the various KO cell lines was performed, and KO efficiency is analyzed by Inference of CRISPR Edits (ICE) CRISPR analysis tool. KO score is calculated by out-of-frame insertion and deletion (indel)% compared to wild type sequences. (A) Penta KO cells for UBQLN1, UBQLN2 or UBQLN4 and RAD23A/B. (B) Triple KO cells for UBQLN1, UBQLN2 or UBQLN4. (C) Double KO cells for RAD23A/B. (D) Single KO cells for UBQLN1, UBQLN2 or UBQLN4.

(E) Stability analysis of the indicated GPS-ORFs in control KO cells (sg-AAVSI) or single KO cells for UBQLN1, UBQLN2 or UBQLN4 in the presence or absence of 1 μM MLN7243.

(F) Stability analysis of the indicated *UbInPD* peptides as analyzed by flow cytometry. GPS-C23mers of the indicated genes were expressed in control KO cells (sg-AAVSI) or penta KO cells for RAD23A, RAD23B, UBQLN1, UBQLN2 and UBQLN4 in the presence or absence of 1 μM MLN7243.