

## Supplemental Information

### **Triubiquitin probes for identification of reader and eraser proteins of branched polyubiquitin chains**

Prajwal Paudel, Christine M. Banos, Yujue Liu and Zhihao Zhuang\*

Department of Chemistry and Biochemistry, University of Delaware, 214A Drake Hall, Newark, Delaware, USA, 19716

\* To whom correspondence should be addressed. Email: [zzhuang@udel.edu](mailto:zzhuang@udel.edu), Tel: +1 302-831-8940

## Supplemental Methods

**Purification of Ub species.** HA-Ub<sub>1-75</sub>-MESNA was generated and purified as previously described<sup>1</sup>. A HA-Ub<sub>1-75</sub> intein fusion gene in pTYB1 plasmid was transformed into *E. coli* BL21(DE3) cells. LB media supplemented with 100 µg/mL ampicillin was used to culture the cells at 37 °C. At OD<sub>600</sub> of 0.6, cell culture was induced with 0.4 mM IPTG and cultured further for 16–18 hrs at 16 °C. Harvested cells by centrifugation were resuspended in a 20 mM Tris (pH 7.5) buffer containing 200 mM NaCl, 1 mM EDTA, 5% glycerol and lysed on ice via sonication. The cell supernatant was incubated with chitin resin (NEB) at 4 °C overnight or r.t for 2 hrs. Following wash with 20 mM Tris (pH 7.5), 1 M NaCl, 1 mM EDTA, and 5% glycerol, the chitin resin was equilibrated with 20 mM MES (pH 6.5), 100 mM NaCl. The resin was then incubated with 50 mL cleavage buffer containing 20 mM MES (pH 6.5), 100 mM NaCl, 75 mM MESNA for 12 hrs at r.t. and eluted to yield Ub<sub>1-75</sub>-MESNA. Further purification was performed on a HiLoad Superdex 200 pg 16/60 column (GE Healthcare). The double Lys-to-Cys mutant Ub<sub>1-75</sub>-MESNA was also prepared following a similar procedure.

K63R Ub<sub>1-76</sub>/pET3a, K48R Ub<sub>1-76</sub>/pET3a, K11C/D77 Ub/pET3a, and K63C/D77 Ub/pET3a in BL21(DE3) *E. coli* cells were grown in LB media in the presence of 100 µg/mL ampicillin at 37 °C until OD<sub>600</sub> reached 0.6, then induced with 0.4 mM IPTG and grown at 17 °C for 16 hrs. The cells were harvested by centrifugation at 4,000 rpm for 10 minutes at 4 °C, and resuspended in a lysis buffer (50 mM Tris, pH 7.5, 200 mM NaCl, 1 mM PMSF, 5 % Glycerol, 1 mM EDTA). The cell lysate was centrifuged at 14,000 rpm for 30 minutes at 4 °C to remove the cellular debris. The supernatant was subjected to heat inactivation at 95 °C for 15 minutes, then centrifuged at 14,000 rpm for 30 minutes. The insoluble fraction was discarded, and the heat treatment was repeated. To the final supernatant, 80% ammonium sulfate was added at 4 °C followed by stirring for 10 minutes before centrifugation at 14,000 rpm for 20 minutes to collect the pellet containing the precipitated ubiquitin. The pellet was resuspended in a gel filtration buffer (50 mM Tris, pH 7.5, 150 mM NaCl), syringe filtered, and then purified using a Superdex 75 10/300 gel filtration column. The resulting protein was concentrated and confirmed by mass spectrometry.

**Purification of tUIM domains.** Genes for UIMC1 (a.a. 76-128), USP37 (a.a. 700-851) and AN13A (a.a. 479-594) were synthesized and cloned into pET28a-HIS-TEV vector by Epoch Life Sciences, Inc. The proteins containing UIM domains were expressed and purified from transformed BL21(DE3) cells. The cells were grown at 37 °C until OD<sub>600</sub> of 0.6, then induced with 0.4 mM IPTG and grown at 17 °C for 16 hrs. The cells were harvested by centrifugation at 4,000 rpm for 10 minutes at 4 °C. The cell pellet was resuspended in a lysis buffer containing 25 mM Tris (pH 8.0), 250 mM NaCl, 2 mM β-ME, 1 mM PMSF for 30 minutes at 4 °C. The cells were lysed and the lysate was centrifuged at 14,000 rpm for 30 minutes at 4 °C to remove the cellular debris. The cell lysate containing the His-tagged UIM was incubated with Ni-NTA resin for 2 hrs rotating at 4 °C, then washed with 20 CV of buffer containing 10 mM imidazole and eluted from the resin with 2 CV buffer containing 250 mM imidazole. The UIM-containing protein was then purified with a HiTrap Q FF anion exchange column (Buffer A: 50 mM Tris, pH 8.0, 1 mM DTT; Buffer B: 50 mM Tris, pH 8.0, 1 M NaCl, 1 mM DTT). The protein was further purified on a Superdex 75 10/300 size exclusion chromatography column in a buffer containing 50 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM DTT.

**Purification of deubiquitinases.** The pOPINK-OTUD2 (full-length, a.a. 1-348, Addgene plasmid # 61409) and pOPINK-OTUD3 (OTU+UBA, a.a. 52-275, Addgene plasmid # 61411) plasmids were gifts from David Komander<sup>2</sup>. They were used for expression of His-GST-OTUD2-FL and His-GST-OTUD3 (OTU+UBA domains). Cells were

cultured, harvested and lysed in lysis buffer (25 mM Tris, pH 8.0, 500 mM NaCl, 5% glycerol, 1 mM  $\beta$ -mercaptoethanol) cooled to 4 °C. Lysate was then centrifuged at 36,000  $\times g$  for 30 min at 4 °C. The protein was batch absorbed using Ni-NTA resin (Qiagen), washed with 25 mM Tris (pH 8.0), 500 mM NaCl, 25 mM imidazole and eluted with an elution buffer (25 mM Tris, pH 8.0, 250 mM NaCl, 250 mM imidazole, 1 mM  $\beta$ -mercaptoethanol). The eluted fractions were combined, concentrated and buffer exchanged to a buffer containing 25 mM Tris (pH 8.0), 50 mM NaCl, 1 mM DTT and further purified by an anion exchange chromatography using a HiTrap Q FF column at 4 °C. Pure fractions were collected and concentrated. The protein purity was analyzed by SDS-PAGE.

**Generation of diUb with native linkages.** 5 mg/mL of monoubiquitin species were used in conjunction with 0.1  $\mu$ M E1 and 20  $\mu$ M of the respective E2 ubiquitin-conjugating enzymes overnight at 37 °C in a buffer containing 50 mM Tris pH 8.0, 5 mM  $MgCl_2$ , 10 mM creatine phosphate, 0.6 U/mL inorganic pyrophosphatase, 0.6 U/mL creatine phosphokinase, 0.1 mM DTT and 0.5 mM ATP to generate native diubiquitin. K63R Ub was reacted with D77 or K11C/D77 Ub using mouse E1 (mE1) and Mms2/Ubc13 to make K63 diubiquitin (K63-diUb) or K63 diUb containing a free Cys at K11 of the proximal ubiquitin. K48R Ub was reacted with D77 Ub, K11C/D77 Ub or K63C/D77 Ub using mE1 and E2-25K to generate K48-linked diUb or diUbs containing a free Cys replacing either K11 or K63 in the proximal ubiquitin. K11C Ub was reacted with D77 Ub using mE1 and UBE2S-UBD followed by AMSH treatment to generate K11-linked diubiquitin (K11-diUb)<sup>3</sup>. The diubiquitins were purified using a cation exchange HiTrap SP FF column (Buffer A: 50 mM ammonium acetate, pH 4.5, 1 mM DTT; Buffer B: 50 mM ammonium acetate, pH 4.5, 1 M NaCl, 1 mM DTT).

**Generation of diUb-NC1-PA2 and diUb-NC1.** DiUb-NC1-PA2 and diUb-NC1 were prepared following a previously published method<sup>4</sup>. To make diUb-NC1-PA2 probes, HA-Ub<sub>1-75</sub>-MESNA was reacted with the noncleavable linker (NC) molecule and deprotected as previously described to obtain HA-Ub<sub>1-75</sub>-NC<sup>5</sup>. In parallel, a mutant Ub<sub>1-75</sub>-MESNA with a Lys-to-Cys mutation at the selected lysine site was reacted with propargylamine (PA) as described above to generate Cys-Ub-PA followed by a cation-exchange HiTrap SP FF column purification. Equal molar of HA-Ub-NC and Cys-Ub-PA were incubated at 0.5 mg/mL each in a 20 mM MES (pH 6.5) buffer containing 0.03 mM EDTA for 1 h at room temperature. Subsequently, the reaction mixture was purified using a cation exchange HiTrap SP FF column followed by purification using a Superdex 75 10/300 size exclusion column in 20 mM MES (pH 6.5) buffer to yield the diUb-NC1-PA2 probes. K63-diUb-NC1 was prepared similarly by reacting equal molar Ub<sub>1-75</sub>-NC with K63C Ub, further purified using a cation exchange HiTrap SP FF column.

**Culture and generation of HEK-293T cell lysates.** Human HEK-293T cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) at 37 °C and 5% CO<sub>2</sub> in a 75 mm<sup>3</sup> T-flask in a humidified incubator. To generate cell lysates, adhered cells were washed with cold DPBS buffer and harvested using 0.25% trypsin. The cells were pelleted and washed three times with additional DPBS buffer and resuspended in lysis buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 5 mM  $MgCl_2$ , 0.5 mM EDTA, 5 mM DTT, 1 mM PMSF, 2 mM ATP, 0.5% Triton X-100 and 10% glycerol). The resuspension was sonicated to lyse the cells and the cell lysate was centrifuged for 15 min at 14,000 rpm at 4 °C and supernatant stored at -80 °C until further use.

**Proteomic data analysis using MaxQuant and Perseus.** Raw files (three replicates for each probe and bead control pulldown) were analyzed together using MaxQuant software version 1.4.0.6 as previously described<sup>6</sup>. In summary, peak list was searched with built-in Andromeda search engine against the UniProt human database downloaded from Uniprot website. MaxQuant parameters were set as follow unless otherwise indicated.

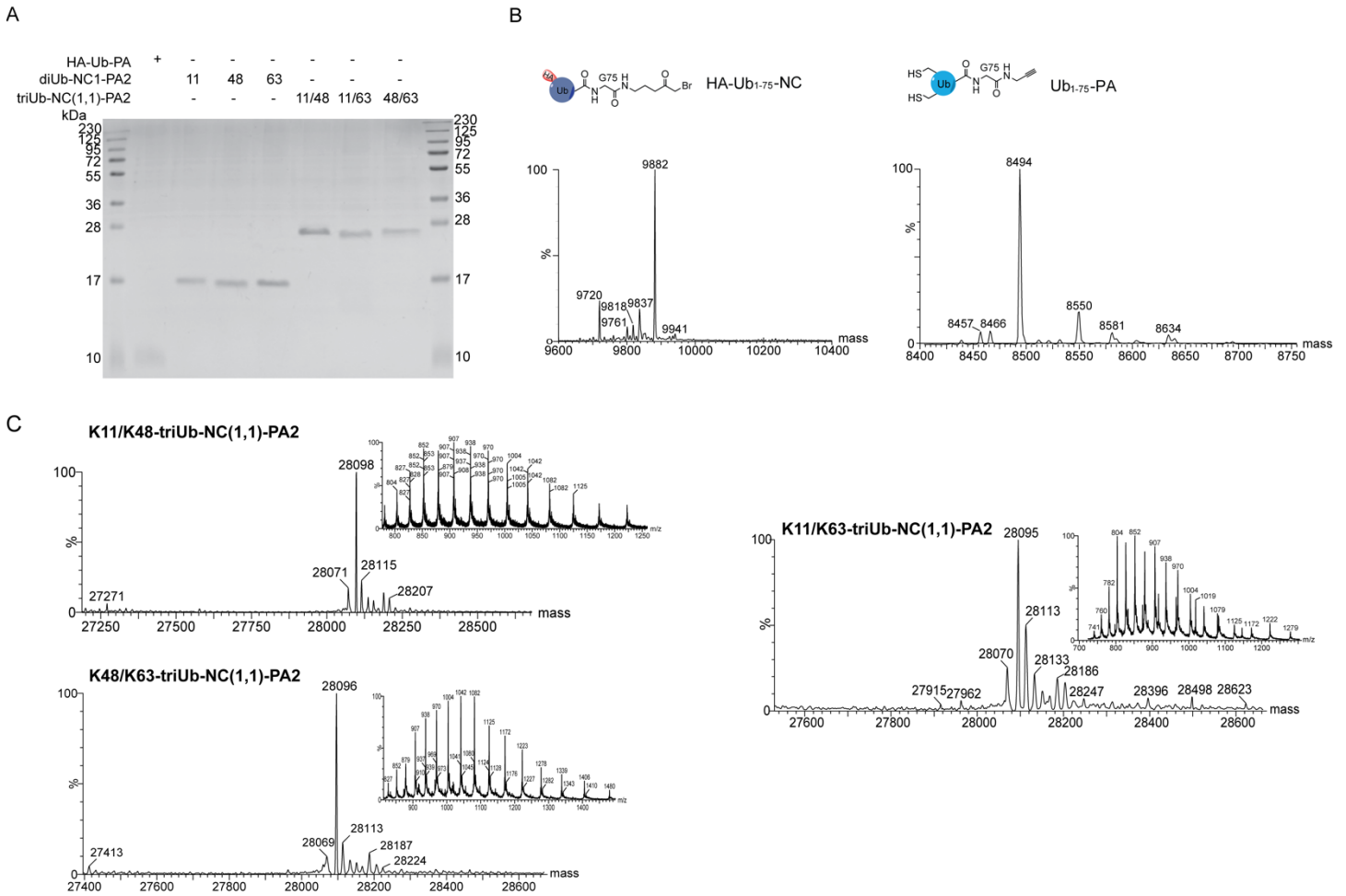
Multiplicity was set to one (no isotope labelling), and trypsin was set to cleave after lysine and arginine, unless followed by a proline. Maximum two missed cleavages were allowed, and maximum charge of peptide was set to 7. Seven amino acids were set to be the minimum peptide length. Carbamidomethylation of cysteine was a fixed modification. N-terminal acetylation, methionine oxidation, phosphorylation (serine, threonine, and tyrosine), and ubiquitin diglycine remnant were set as variable modifications. Label-free Quantification (LFQ) module was enabled with a minimum ratio count of 1, and second peptide search and re-quantify functions were enabled. The maximum number of modifications was set to 5 per peptide. The option of requiring MS/MS for LFQ comparisons was enabled. Target decoy approach was used to filter peptide spectrum matches (PSM), and a false discovery rate (FDR) of 1% was used for protein identification. The output text file "protein groups" was imported and analyzed using Perseus (software version 1.6.1.3). Proteins categorized as "only identified by modified peptides" were eliminated. Proteins identified with reversed decoy mode and the common contaminant proteins were excluded. Experimental replicates (3 total) were grouped. Proteins with at least one valid LFQ values (LFQ intensity >0) in at least one group were kept. Histogram of pre-imputed LFQ values showed relatively normal distribution which was used to impute missing values. Missing values were imputed using the normal distribution of the total matrix using the default parameters (0.3 width and 1.8 down shift) in Perseus.

Using Perseus multi-sample ANOVA test was performed with FDR of 0.01 and  $S_0 = 1$ . ANOVA enriched  $\text{Log}_2(\text{LFQ})$  values were transformed using Z-score normalization for each protein group and averaged for control and each probe pulldown. Heatmap was generated using hierarchical clustering with complete linkage method. Spiderplots were generated using Excel from the bead subtracted  $\text{log}_2(\text{LFQ})$  values for the probes. Gene ontology (GO) analysis was performed on clusters using online ([www.geneontology.org](http://www.geneontology.org)) and domain batch analysis was performed using SMART database ([www.smart.embl.de](http://www.smart.embl.de)) and visualized using iTol ([www.itol.embl.de](http://www.itol.embl.de)). Sequence alignments of relevant proteins were performed with Clustal Omega ([www.ebi.ac.uk/Tools/msa/clustalo/](http://www.ebi.ac.uk/Tools/msa/clustalo/)).

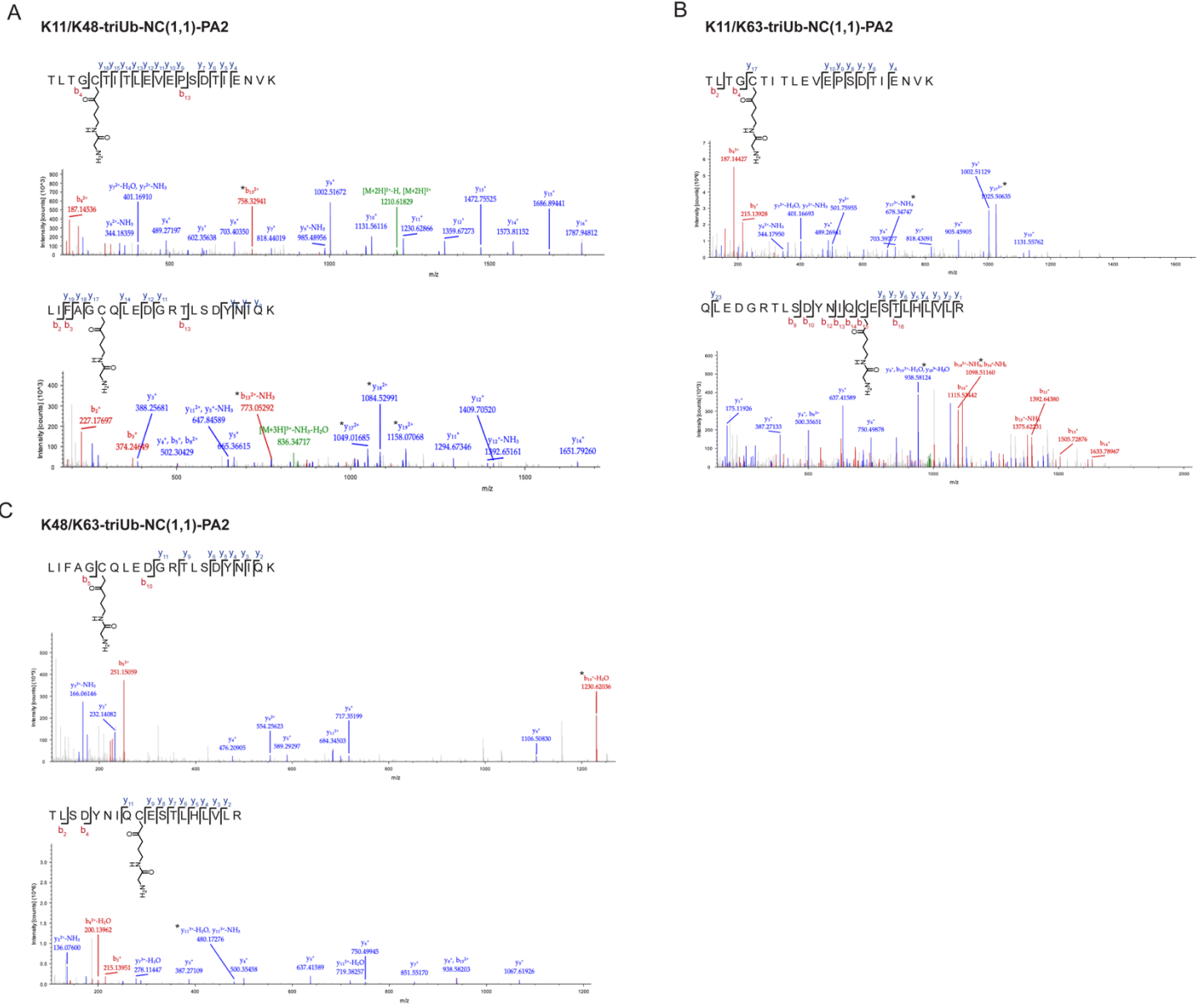
**Labelling of recombinant DUBs using mono-, di- and triUb-PA probes.** Labelling of DUBs with the activity-based Ub probes (ABPs) was performed in 20 mM Tris (pH 8.0), 50 mM NaCl and 1 mM DTT at r.t. for 2 hrs with 1  $\mu\text{M}$  probes and 2  $\mu\text{M}$  enzymes. The reaction was quenched with 6X loading dye with 10 mM DTT and loaded onto an SDS-PAGE gel. The gels were fixed and stained with Coomassie brilliant blue, destained and visualized.

**Western blotting analysis of lysate labelling by di- and tri-Ub-PA probes.** 50  $\mu\text{g}$  HEK-293T cell lysate and 1  $\mu\text{M}$  HA-tagged monoUb, diUb and triUb probes were incubated in the lysis buffer to a final volume of 10  $\mu\text{L}$  for 2 hrs at room temperature. The labelling reaction was quenched using 6X gel loading dye (35 mM Tris, 10% SDS, 30% glycerol and 0.81 M DTT, pH 6.8) and heated on a heat block at 95  $^{\circ}\text{C}$  for 10 min. Following centrifugation, sample was loaded and resolved on a 12% SDS-PAGE gel followed by transfer onto a PDVF membrane (Thermo Fisher Scientific Inc). The membrane was cut below the 50 kDa ladder band and blocked with 5% skim milk. The top half membrane was blotted with mouse anti-HA antibody in 5% BSA followed by a secondary HRP-conjugated anti-mouse antibody. The bottom half membrane was immunoblotted for GAPDH loading control with a mouse anti-GAPDH antibody in 5% BSA followed by blotting with a secondary HRP-conjugated anti-mouse antibody. Membranes were incubated with ECL Western blotting substrate (Thermo Fisher Scientific Inc., Waltham, MA) and signals were imaged.

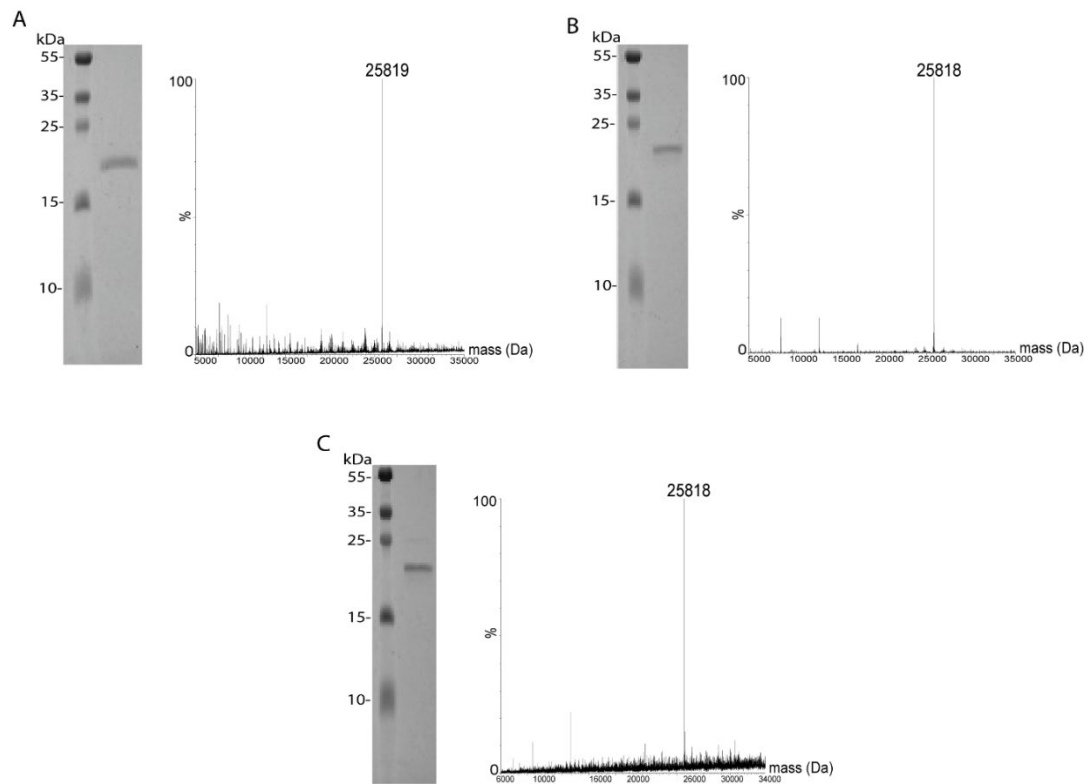
## Supplemental Figures



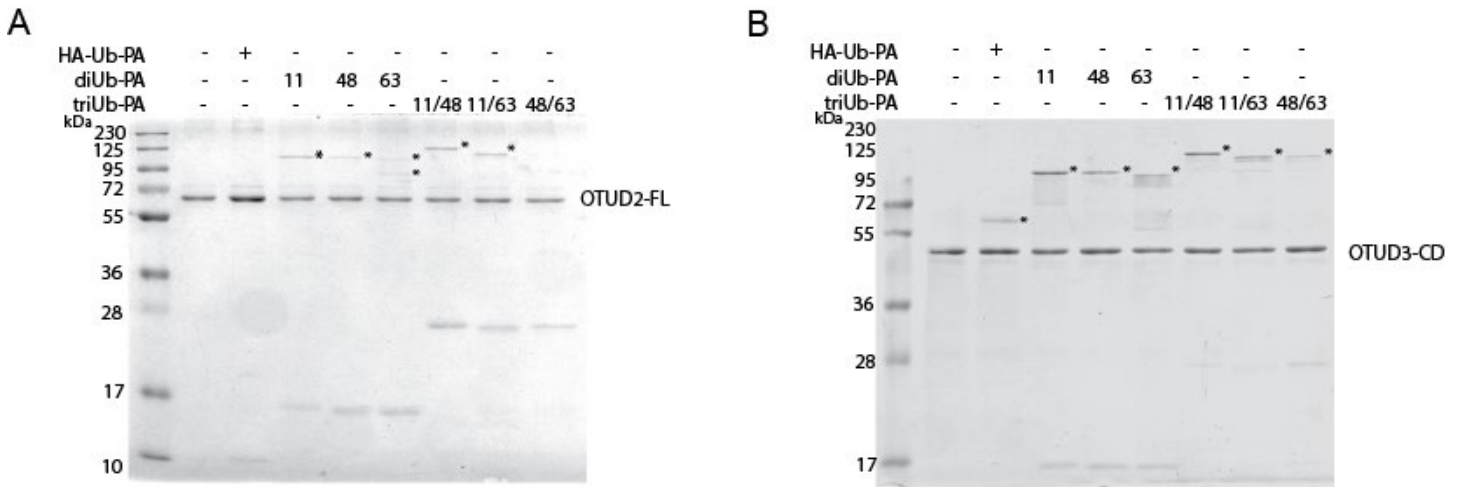
**Figure S1.** (A) SDS-PAGE gel analysis of diUb-PA and triUb-PA probes as indicated. (B) MS analysis of intermediates of branched triUb probe generation (calculated mass for HA-Ub<sub>1-75</sub>-NC: 9882 Da; mutant Ub<sub>1-75</sub>-PA: 8495 Da). (C) MS analysis of the three branched triUb probes (calculated mass: 28,097 Da).



**Figure S2.** Mass spectrometry analysis (MS/MS) of branched chain triUb-PA. (A) K11/K48 triUb-NC(1,1)-PA2, (B) K11/K63 triUb-NC(1,1)-PA2, (C) K48/K63 triUb-NC(1,1)-PA2. The selected peptides were obtained from MS/MS spectra of pull-down experiments using the corresponding branched triUb probes. Asterisk indicates fragments containing the corresponding branching sites.



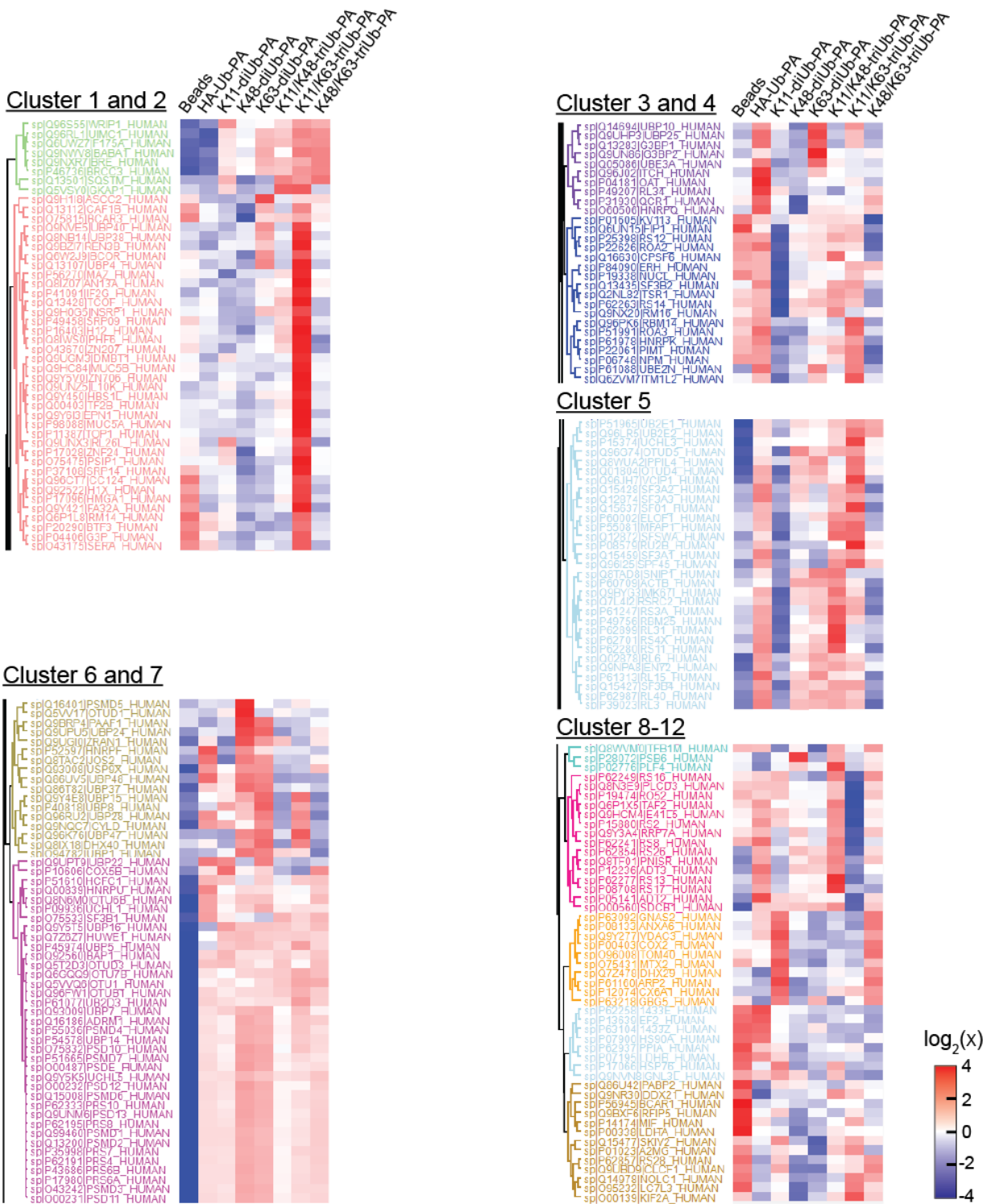
**Figure S3.** SDS-PAGE and mass spectrometry analysis (intact mass) of branched triUb-(1,0) species with (A) K11/K63, (B) K11/K48 and (C) K63/K48 linkages respectively. The calculated molecular weight of the triUb-(1,0) is 25,819 Da.



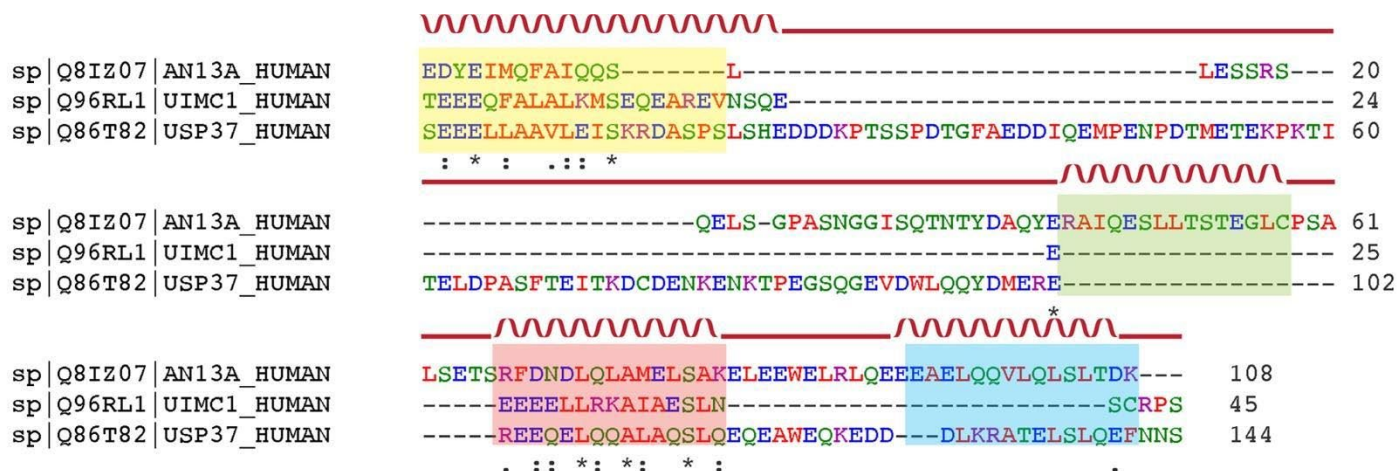
**Figure S4.** *In-vitro* labeling of purified DUBs (A) OTUD2-FL and (B) OTUD3-CD with mono-, di- and branched triUb-PA probes as indicated. SDS-PAGE gels were stained with Coomassie brilliant blue dye. Asterisks indicate detected labeling bands.



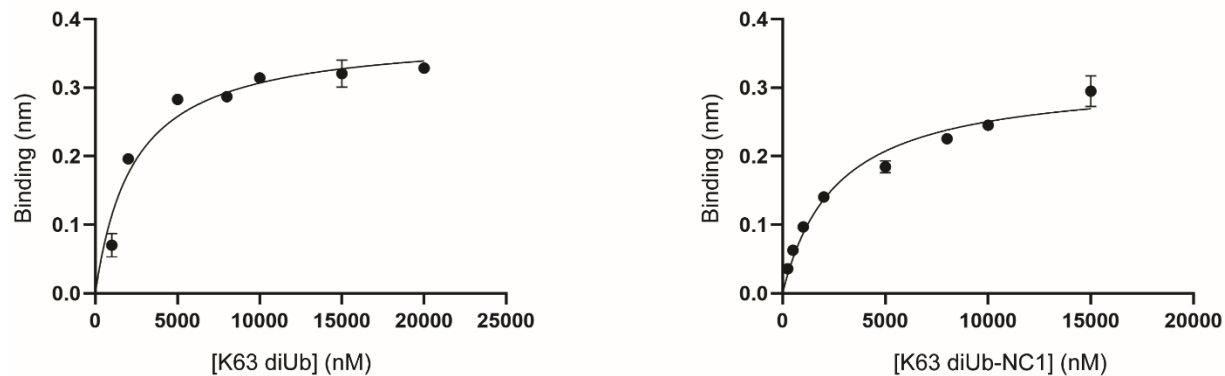




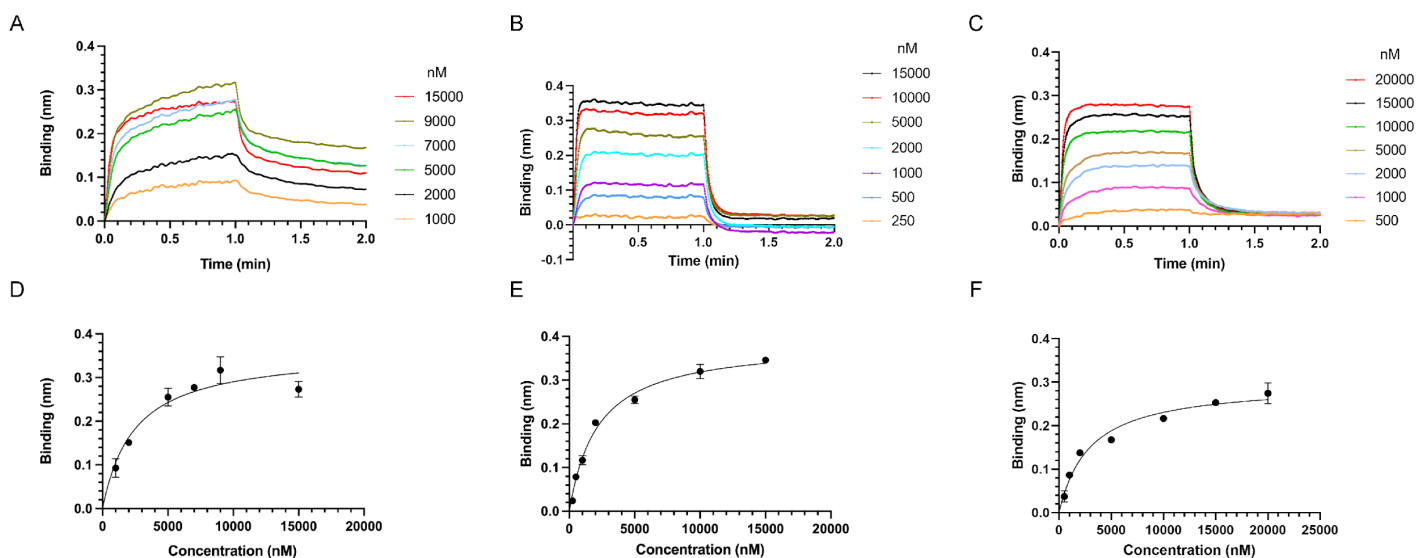
**Figure S6.** Hierarchical clustering of enriched proteins using branched triUb-PA probes in comparison to the monoUb-PA and diUb-PA probes. The protein names are given in each of the clusters. The heatmap reflects a Z-score normalization of  $\log_2(\text{LFQ})$  value of a captured protein by the ubiquitin probes used in the parallel pulldown.



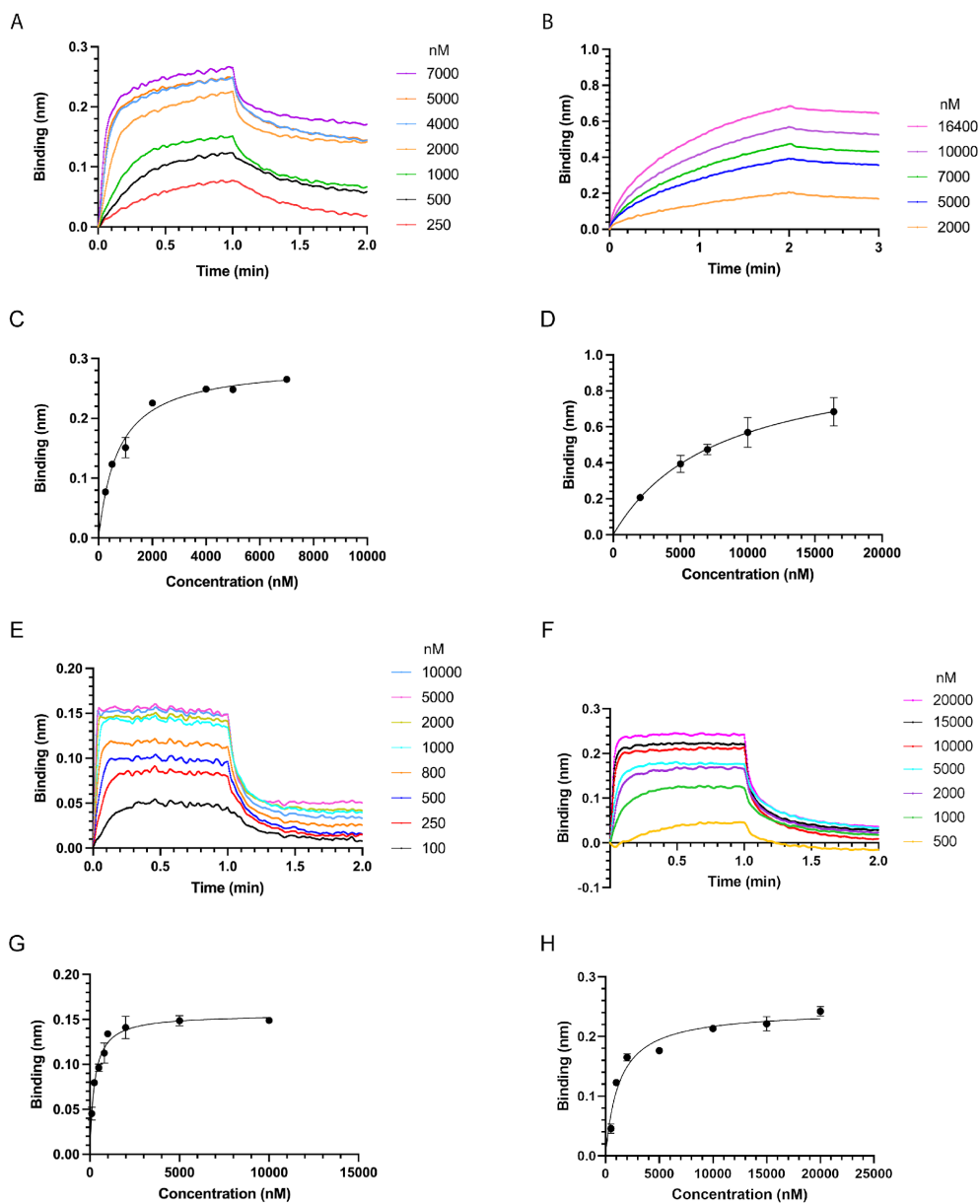
**Figure S7.** Clustal alignment of the identified UIM-containing sequences in UIMC1/Rap80, AN13A and USP37. Colored shading represents predicted UIM sequences. Secondary structure was predicted using AlphaFold modeling<sup>7,8</sup>.



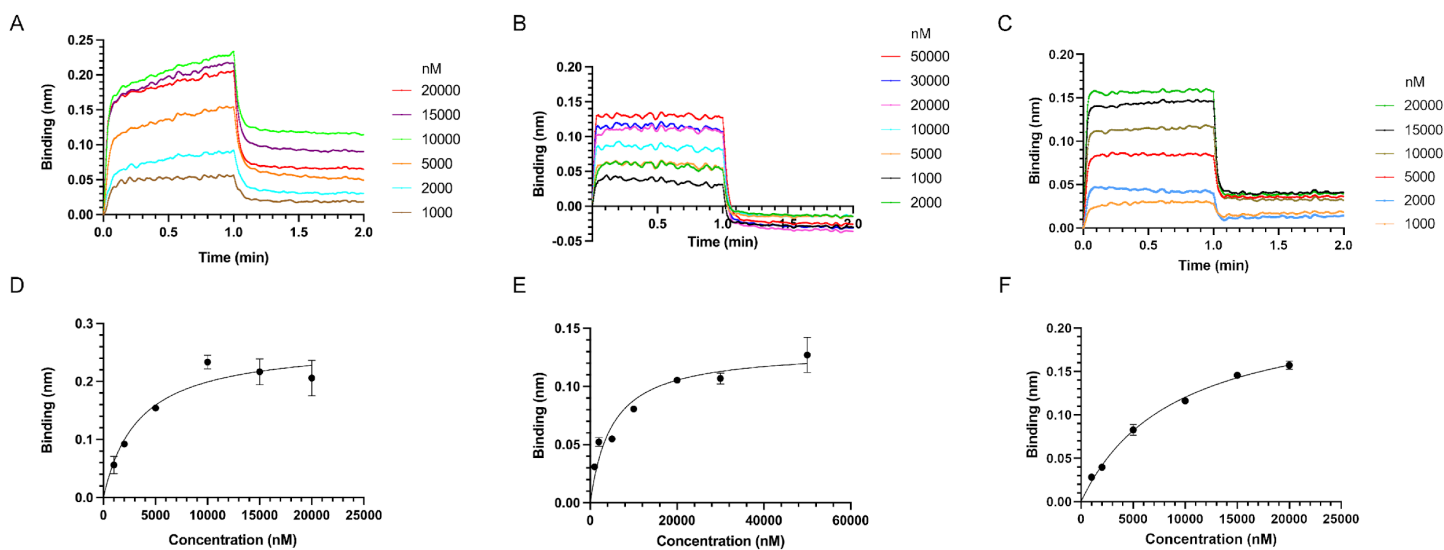
**Figure S8.** Comparison of K63-diUb with a native linkage (K63 diUb) to the K63-diUb containing the thioether linkage (K63 diUb-NC1) in binding to UIMC176-128. The  $K_d$  of UIMC176-128 binding to the Ub species was determined by fitting the maximum BLI binding observed versus the concentration of Ub species used in the BLI binding experiment. UIMC176-128 binds K63-diUb with a  $K_d$  of  $2.34 \pm 0.42 \mu\text{M}$ , in comparison to a  $K_d$  of  $2.63 \pm 0.41 \mu\text{M}$  for K63-diUb-NC1.



**Figure S9.** BLI traces of diUb and branched triUb(1,0) of different linkages binding to UIMC<sub>176-128</sub> immobilized onto the BLI sensor tip surface. BLI traces averaged from two repeats are shown for concentration series of (A) K48/K63-triUb-(1,0); (B) K63-diUb; (C) K11-diUb. The  $K_d$  of binding was determined by fitting the maximum BLI binding observed versus the concentration of Ub species used in the BLI binding experiment for (D) K48/K63-triUb-(1,0), (E) K63-diUb and (F) K11-diUb.

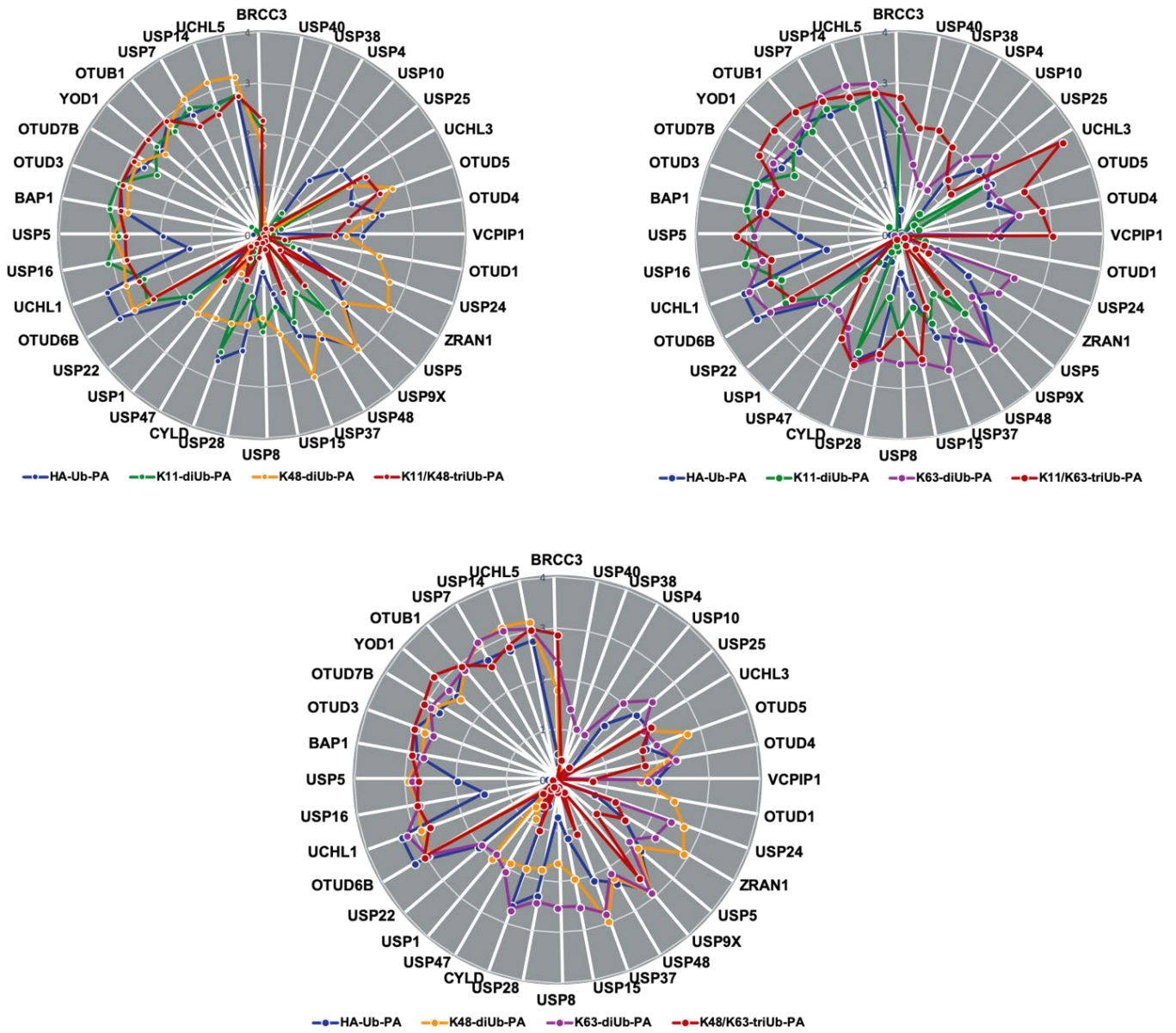


**Figure S10.** AN13A<sub>479-594</sub> binding to diUb and triUb(1,0) of indicated linkages. BLI traces averaged from two repeats are shown for concentration series of (A) K48/K63-triUb-(1,0); (B) K48/K11-triUb-(1,0); (E) K63-diUb; (F) K11-diUb. The  $K_d$  of AN13A<sub>479-594</sub> binding to the different Ub species was determined by fitting the maximum BLI binding observed versus the concentration of Ub species used in the BLI binding experiment for (C) K48/K63-triUb-(1,0), (D) K48/K11-triUb-(1,0); (G) K63-diUb and (H) K11-diUb.



**Figure S11.** USP37<sub>700-851</sub> binding to diUb and triUb(1,0) of indicated linkages. BLI traces averaged from two repeats are shown for concentration series of (A) K63/K11-triUb-(1,0); (B) K63-diUb; (C) K11-diUb. The  $K_d$  of binding was determined by fitting the maximum BLI binding observed versus the concentration of Ub species used in the BLI binding experiment for (D) K63/K11-triUb-(1,0), (E) K63-diUb and (F) K11-diUb.





**Figure S12.** Spiderplot analysis of DUBs enriched by monoUb-PA, di-Ub-PA and branched triUb-PA probes of different linkages.



- (1) Li, G.; Liang, Q.; Gong, P.; Tencer, A. H.; Zhuang, Z. Activity-based diubiquitin probes for elucidating the linkage specificity of deubiquitinating enzymes. *Chemical Comm* **2014**, *50* (2), 216-218.
- (2) Mevissen, T. E.; Hospenthal, M. K.; Geurink, P. P.; Elliott, P. R.; Akutsu, M.; Arnaudo, N.; Ekkebus, R.; Kulathu, Y.; Wauer, T.; El Oualid, F.; et al. OTU deubiquitinases reveal mechanisms of linkage specificity and enable ubiquitin chain restriction analysis. *Cell* **2013**, *154* (1), 169-184.
- (3) Bremm, A.; Freund, S.; Komander, D. Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne. *Nat Struct Mol Biol* **2010**, *17*, 939–947.
- (4) Paudel, P.; Zhang, Q.; Leung, C.; Greenberg, H.; Guo, Y.; Chern, Y.; Dong, A.; Li, Y.; Vedadi, M.; Zhuang, Z.; et al. Crystal Structure and Activity-Based Labeling Reveal the Mechanisms for Linkage-Specific Substrate Recognition by Deubiquitinase USP9X. *Proceedings of the National Academy of Sciences of the United States of America* **2019**, *116* (15), 7288-7297.
- (5) Yang, K.; Li, G.; Gong, P.; Gui, W.; Yuan, L.; Zhuang, Z. Chemical Protein Ubiquitylation with Preservation of the Native Cysteine Residues. *ChemBiochem* **2016**, *17* (11), 995-998.
- (6) Gui, W.; Ott, C. A.; Yang, K.; Chung, J. S.; Shen, S.; Zhuang, Z. Cell-Permeable Activity-Based Ubiquitin Probes Enable Intracellular Profiling of Human Deubiquitinases. *J Am Chem Soc* **2018**, *140* (39), 12424-12433.
- (7) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583-589.
- (8) Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **2022**, *50* (D1), D439-D444.