Figure S1, related to Figure 1

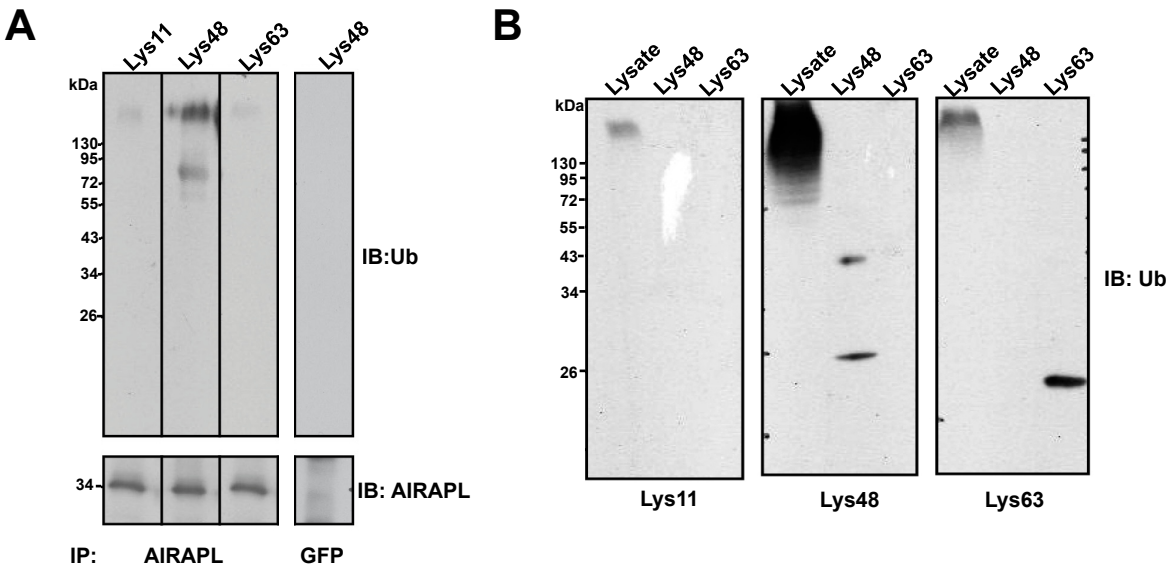


Fig. S1 AIRAPL is selective for Lys48-linked ubiquitin chains (**A**) Endogenous AIRAPL or GFP (serving as a non-relevant Ab.-lane 4) were immunoprecipitated from HEK293T cells and immunobloted against AIRAPL and poly-ubiquitin chains using ubiquitin chain linkage-specific antibodies. Anti-GFP antibody served as a specificity control for AIRAPL IP and Lys48 IB. (**B**) Control for the poly-ubiquitin chain specific antibodies was performed by running total cell lysates as well as commercially available tri-ubiquitin Lys48- and Lys63-linked poly-ubiquitin chains.

Figure S2, related to Figure 1D

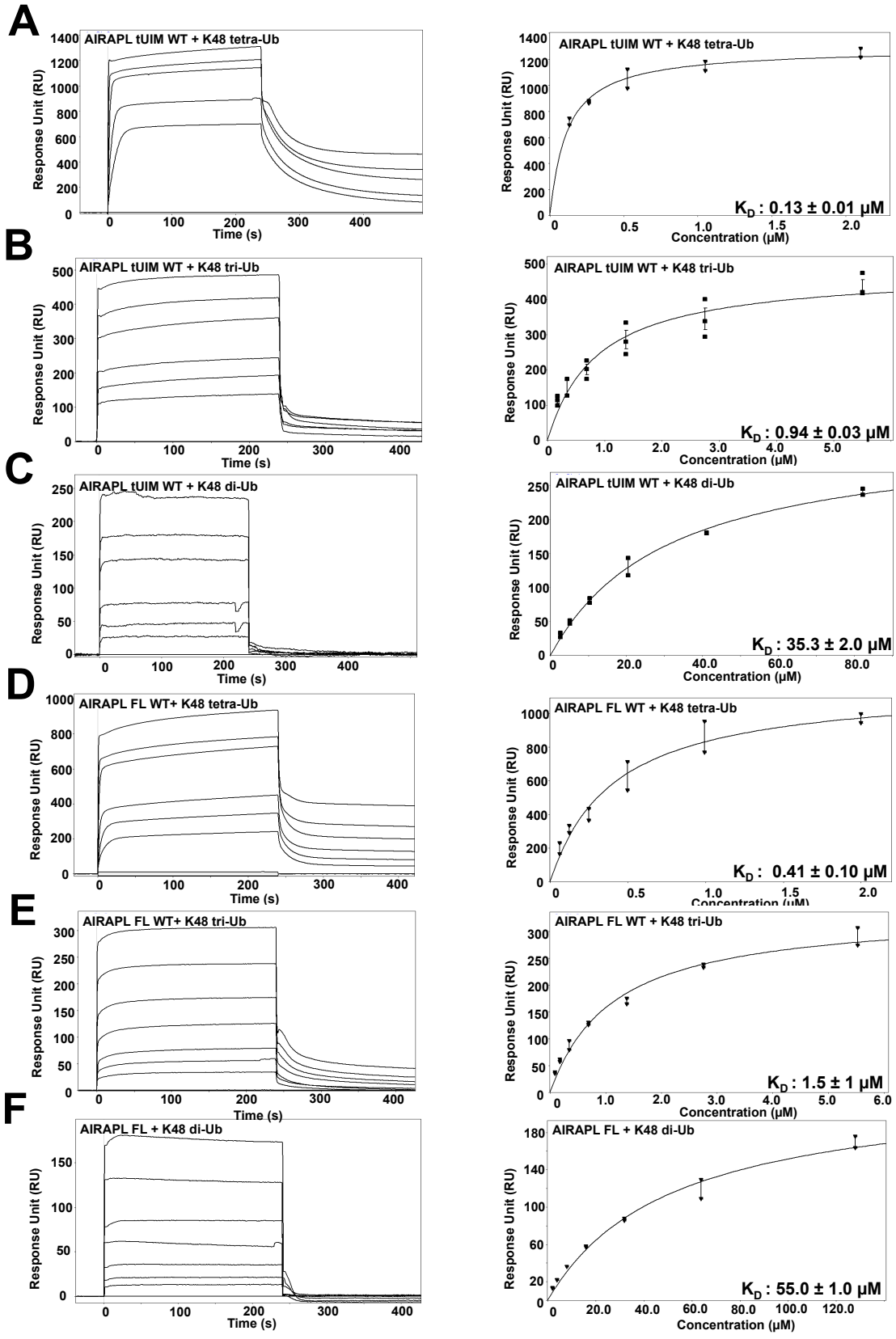


Figure S2, related to Figure 1D

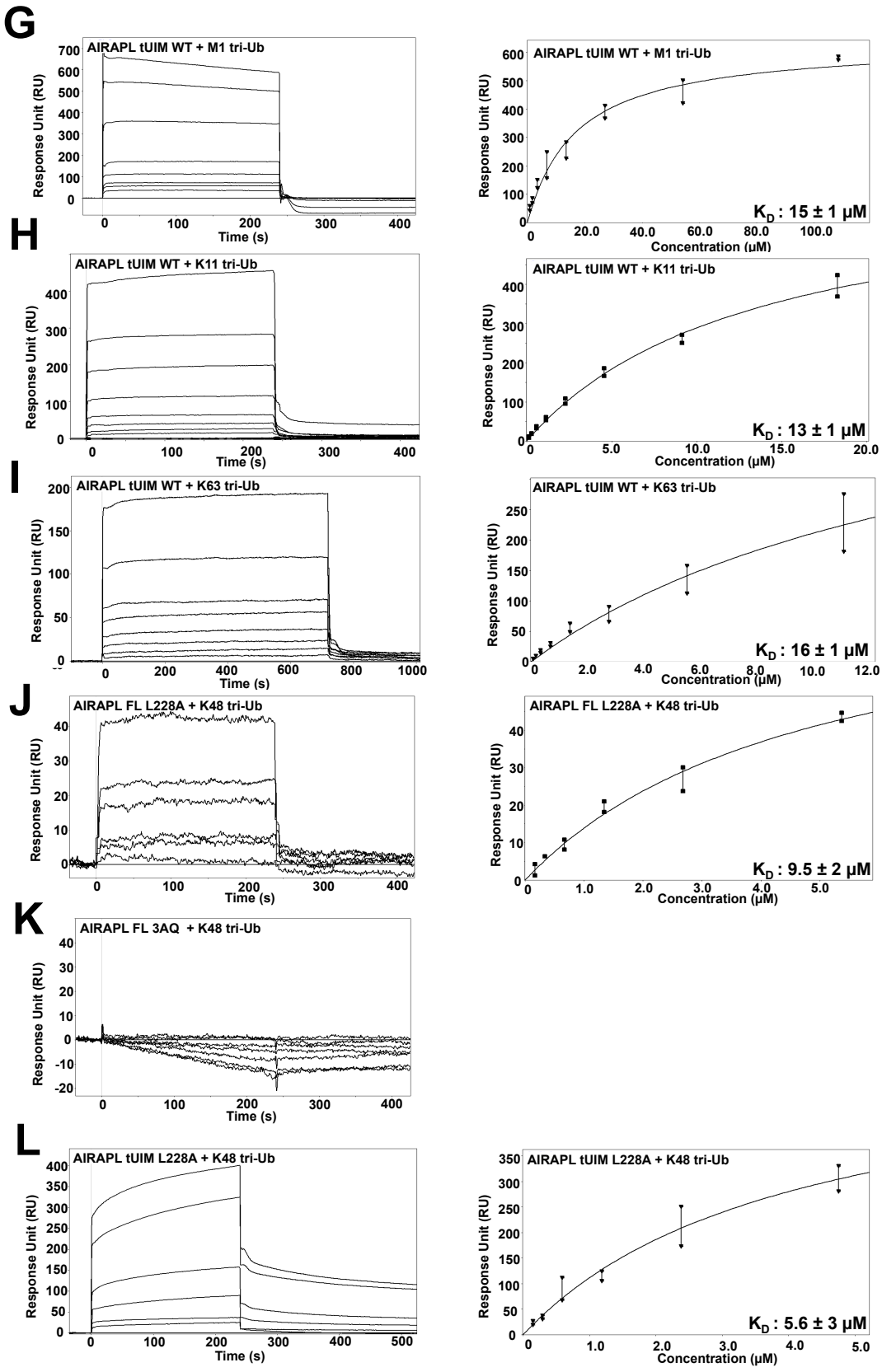


Figure S2, related to Figure 1D

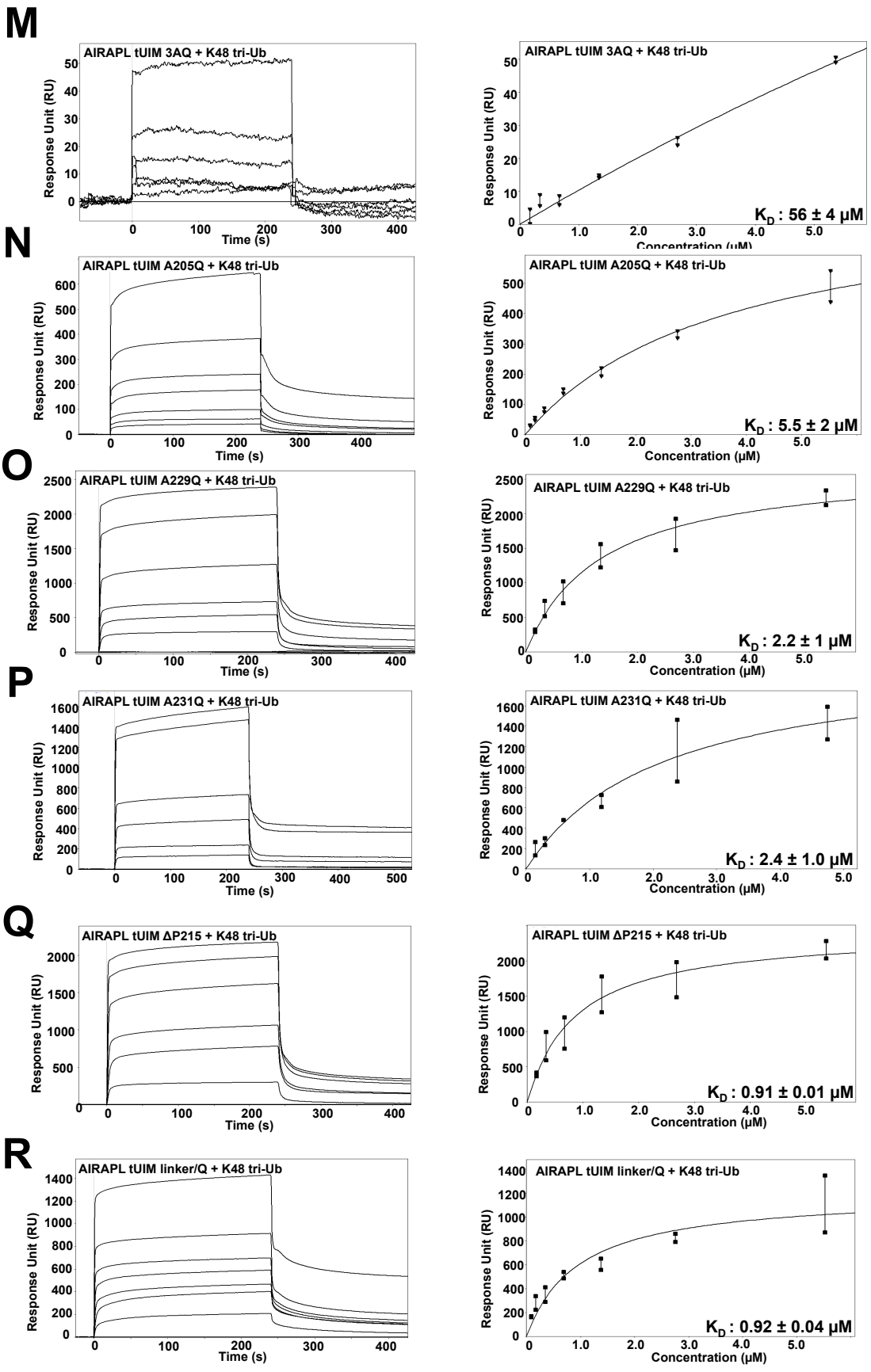


Figure S2, related to Figure 1D

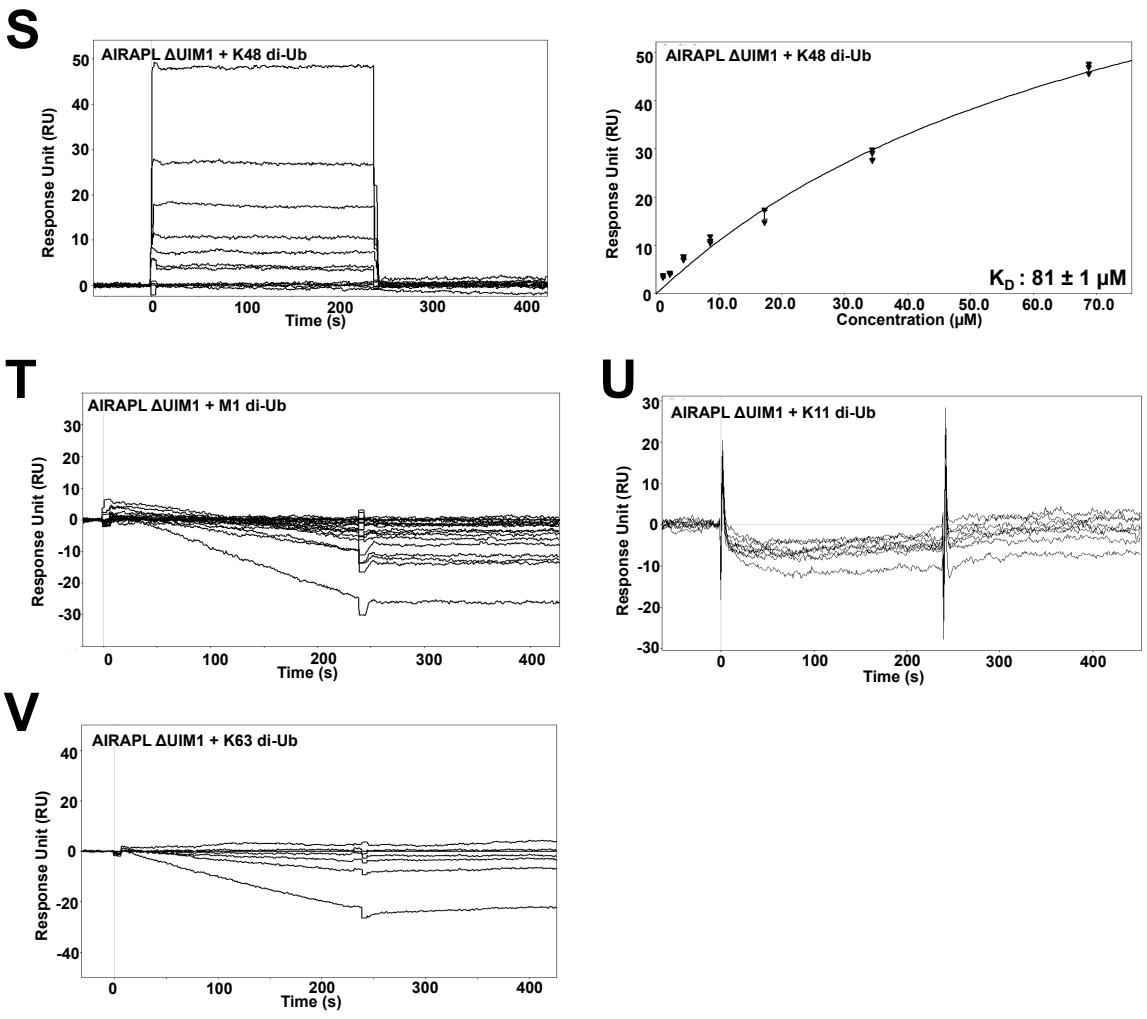
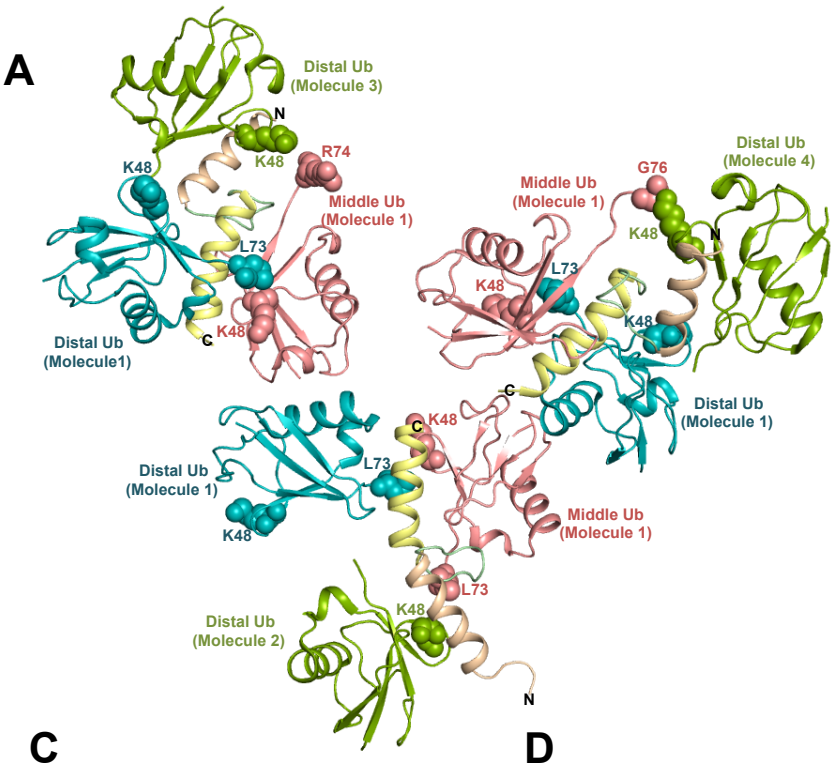
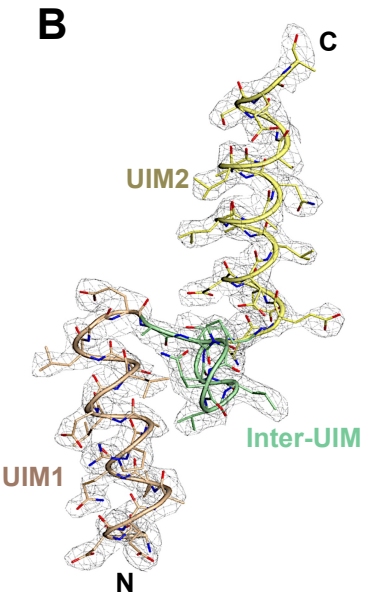
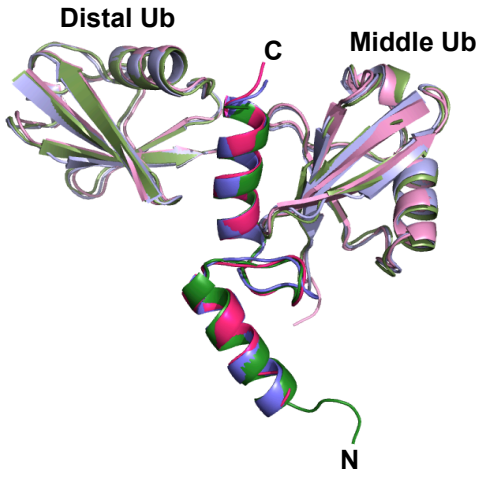


Fig. S2 Surface plasmon resonance (SPR) (left) sensograms and (right) equilibrium fitting for (**A-F**) AIRAPL FL (full-length) or tUIM (residues 188-240) binding to Lys48-linked di-, tri-, and tetra-ubiquitin, (**G-I**) tUIM binding to Met1-, Lys11-, and Ly63-linked tri-ubiquitin, (**J-K**) FL mutants including 3AQ (A205/229/231Q) and L228A binding to Lys48-linked tri-ubiquitin, (**L-R**) tUIM mutants including 3AQ, L228A, A205Q, A229Q, A231Q, ΔP215 (deletion of P215), and linker/Q (all inter-UIM linker residues mutated to Gln) binding to Lys48-linked tri-ubiquitin, (**S-V**) ΔUIM1 (tUIM construct lacking UIM1) binding to Lys48-, Met1-, Lys11-, and Lys63-linked di-ubiquitin chains. Each measurement was repeated at least two times. In (**K, T, U, V**) only SPR sensograms are shown as not binding was detected between the indicated AIRAPL and ubiquitin chain molecules.

Figure S3, related to Figure 2





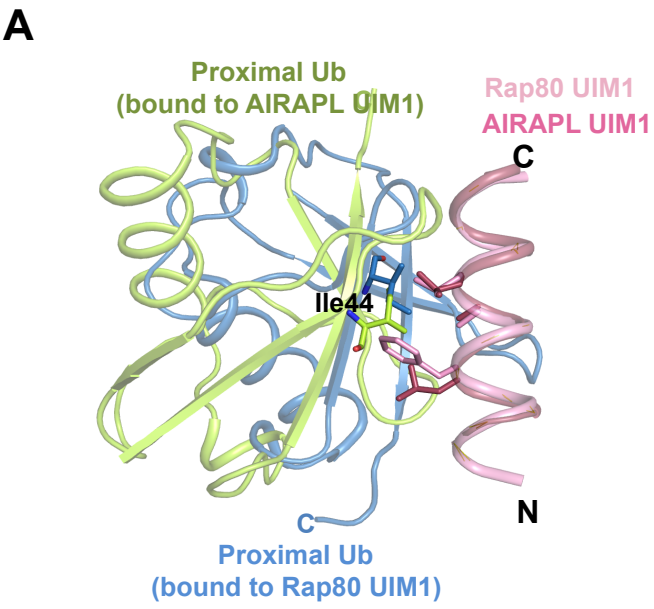


D

	Lys48 tri-Ub	AIRAPL tUIM/Lys48 tri-Ub
Rg (reciprocal space)	24.0	25.3
Rg (real space)	24.0	25.3
Dmax	81.5	81.5
Porod volume	35700	48200
MW estimation by SAXS (kDa)	22.3	30.1
MW from sequence (kDa)	25.5	31.5

Fig. S3 Structure of the AIRAPL tUIM in complex with a Lys48-linked tri-ubiquitin chain. (**A**) The six ubiquitin moieties in an asymmetric unit (AU) of the complex crystal can be related as three di-ubiquitins (distal-middle). However, C-term of each middle ubiquitin is adjacent to the Lys48 of the distal ubiquitin from a neighboring AU. Molecules 1 to 4 represent 4 neighboring asymmetric units. (**B**) Final $2F_o$-F_c electron density map contoured at 1σ for tUIM from tUIM/Lys48-linked tri-ubiquitin complex. (**C**) Superimposition of the three complex molecules in the asymmetric unit of AIRAPL tUIMs/Lys48-linked tri-ubiquitin indicate highly similar overall conformations apart from some flexibility in inter-UIM regions. (**D**) SAXS analysis of Lys48-linked tri-ubiquitins alone (black) and in complex with AIRAPL tUIM. Molecular weights (MW) were estimated by Porod Volume (ref. 28).

Figure S4, related to Figure 3

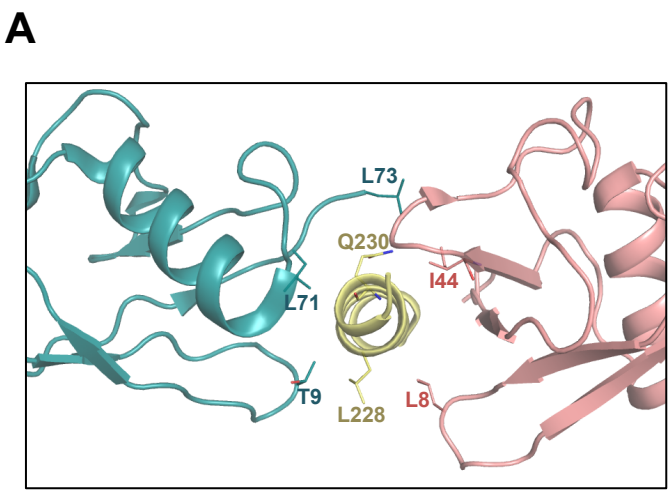


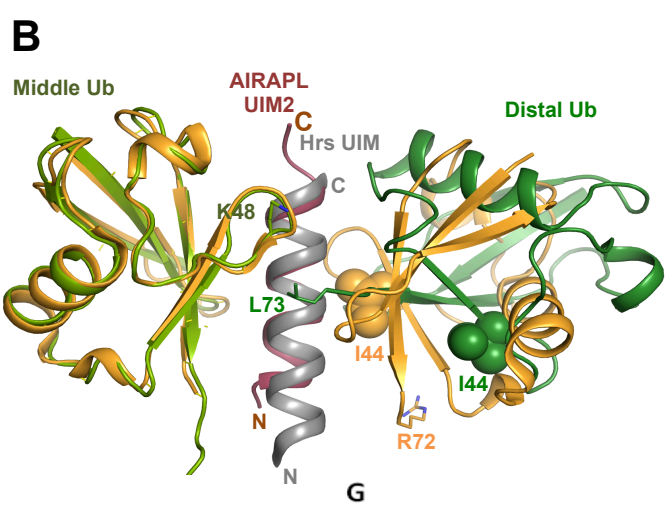
B

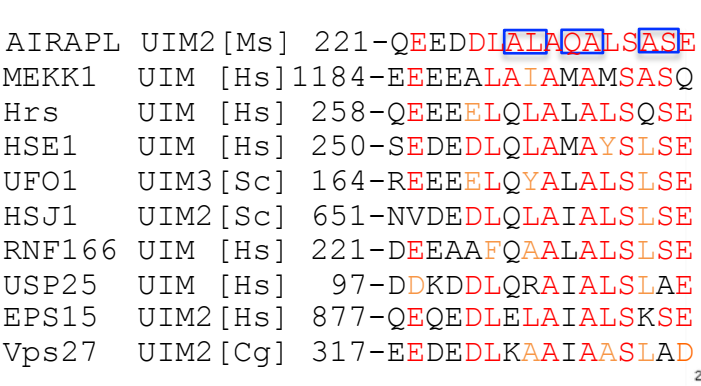

```
Mus musculus       188-PVIALQNGLSEDEALQRALELSLAEAKPQVLSSQEEDDLALAQALSASEAEY-240
Homo sapiens       188-PVIALQNGLSEDEALQRALELSLAETKPQVPSSQEEDDLALAQALSASEAEY-240
Bos taurus         189-PVITLQSGLSEDEALQRALELSLAETKPQVPSSQEEDDLALAQALSASEAEY-241
Rattus norvegicus  188-PVIALQNGLSEDEALQRALELSLAEAKPQIPSSQEEDDLALAQALSASEAEY-240
Macaca mulatta     188-PVIALQNGLSEDEALQRALEMSLAETKPQVPSCQEEDDLALAQALSASEAEY-240
Pan troglodytes    188-PVIALQNGLSEDEALQRALEMSLAETKPQVPSCQEEDDLALAQALSASEAEY-240
Sus scrofa         191-PVIALQNGLSEDEALQRALELSLAETKPQVPSSQEEDDLALAQALSASEAEY-243
Canis familiaris   150-PVIALQNGLSEDEALQRALELSLAETKPQVPSSQEEDDLALAQALSASEAEY-202
Felis Catus        187-PVIALQNGLSEDEALQRALELSLAETKPQVPSSQEEDDLALAQALSASEAEY-239
```

Fig. S4 Characterization of the interactions between AIRAPL tUIM and ubiquitin moieties. (**A**) Comparison of the AIRAPL UIM1/proximal ubiquitin-binding mode with the conventional UIM/ubiquitin binding mode by superimposition on to the Rap80 UIM1/ubiquitin structure (pdb: 3A1Q). Although, proximal ubiquitin bound to AIRAPL UIM1 is inverted compared with the one bound to Rap80 UIM1, both ubiquitins employ hydrophobic patch centered on Ile44 for interactions with UIM domains (**B**) The AIRAPL tUIM region is well conserved among various species. Sequence alignment of the AIRAPL tUIMs illustrates the conservation of a proline residue in the inter-UIM region. Conserved residues in the UIM1, inter-UIM and UIM2 are colored in orange, blue and red, respectively.

Figure S5, related to Figure 4







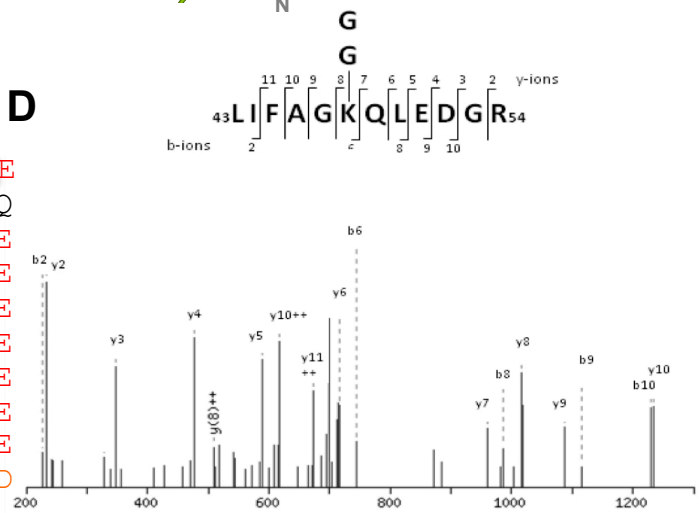


Fig. S5 AIRAPL UIM2 defines selectivity towards Lys48-linked ubiquitin chains. (**A**) Two residues, L228 and Q230, from UIM2 interact with both middle and distal ubiquitins. (**B**) Superimposition of the UIM2/ middle, distal ubiquitins and Hrs double-sided UIM/ mono-ubiquitin complex (pdb: 2D3G) structures. Hrs UIM and ubiquitins bound to it are colored in gray and orange, respectively. Residue Ile44, the central residue in the hydrophobic surface of the middle ubiquitin is shown as a sphere. The last C-terminal residues of the middle ubiquitin and Lys48 residue of the distal ubiquitin are shown as sticks. (**C**) Sequence alignment of the putative double-sided UIMs, found by sequence similarity search using AIRAPL UIM2 as a template in Blast/Uniprot. The fully conserved and homologous residues on the distal and middle Ub-binding surfaces are colored in red and orange, respectively. Blue boxes indicate residues from AIRAPL UIM2 interacting with the distal Ub. Purple boxes illustrate the residues from Hrs dUIM that contact the two Ub moieties. (**D**) MEKK1 (MAPK/ERK kinase kinase 1) was found as a possible Lys48 linkage-selective protein based on conservation of residues on the distal ubiquitin-binding site. LC-MS/MS analysis of HMW poly-ubiquitin chains co-purified with a MEKK1 UIM domain identified only Lys48-linked chains, suggesting the Lys48-linkage specificity of the MEKK1 double-sided UIM. The HMW MEKK1 affinity purified section from a Coomassie gel was subject to MS analysis. Shown is the MS/MS spectrum of the K48-linkage G-G containing peptide 43-54 derived from ubiquitin with a precursor ion mass of 731.0 Da [M+2H]2+. The b/y fragment ions are indicated.

Figure S6, related to Figure 5

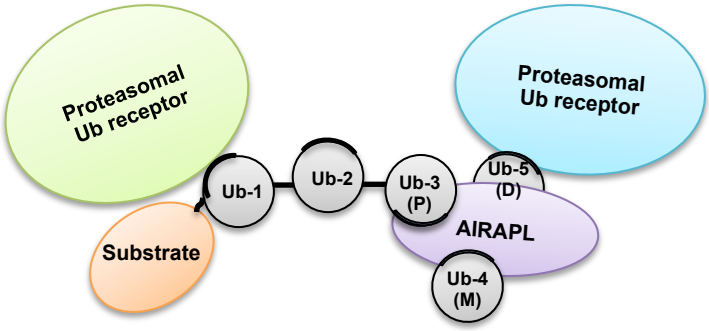


Fig. S6 A model for recruitment of a Lys48-linked ubiquitin-charged substrates to the proteasomal ubiquitin receptors (Rpn10 and Rpn13) by AIRAPL. In this model which is based on the crystal structure of AIRAPL tUIM/Lys48-linked tri-ubiquitin, one of the ubiquitin receptors (Rpn10 or Rpn13) at the proteasome 19S regulatory particle recognizes distal ubiquitin and the chain is elongated by two ubiquitins from proximal site to bind the second ubiquitin receptor. Ub-3, Ub-4, and Ub-5 represent proximal (P), middle (M), and distal (D) ubiquitins, respectively.

Supplemental Experimental Procedures

Protein expression and purification

The GST-tagged proteins were prepared by cloning AIRAPL FL and tUIM construct (188-240) into pGEX-6p1 vector. Protein expression was induced by addition of 0.5 mM IPTG and overnight incubation at 25°C. Bacterial cells were lysed by sonication in phosphate buffer saline (PBS) buffer. GST-fused proteins were purified from the cell lysate using glutathione sepharose 4B column (GE Healthcare). For co-crystallization, SAXS, and CD experiments GST tag was later cleaved using PreScission protease and proteins were eluted from the column with PBS. The proteins were further purified by size exclusion chromatography (Superdex 75 column, GE Healthcare) in a buffer containing 50 mM Tris-HCl, pH 8.0, and 150 mM NaCl. Met1-linked di- and tri-ubiquitins were cloned into pGEX-4T1 and expressed and purified as described above, only the GST tag was cleaved using thrombin protease. For SPR assays, AIRAPL FL and tUIM (wild type or mutants) were cloned into pET-28a vector and expressed in *E. coli* BL21 (DE3) by addition of 0.5 mM IPTG and incubation at 25°C, overnight. Cells were re-suspended in PBS and lysed by sonication. His-tagged proteins were purified by binding to Ni-NTA affinity resin (QIAGEN) and the resin was washed with 10 mM imidazole to get rid of non-specific interactions. The His-tagged proteins were eluted from the Ni-NTA resin with 200 mM imidazole. Lys48-linked di-, tri- and tetra-ubiquitins were synthesized as described by Komander et al. (Komander et al., 2008). Lys11- and Lys63-linked di- and tri-ubiquitins, and Lys48-linked 2-7 ubiquitins were purchased from LifeSensors.

References

KOMANDER, D., LORD, C. J., SCHEEL, H., SWIFT, S., HOFMANN, K., ASHWORTH, A. & BARFORD, D. 2008. The structure of the CYLD USP domain explains its specificity for Lys63-linked polyubiquitin and reveals a B box module. *Mol Cell,* 29**,** 451-64.