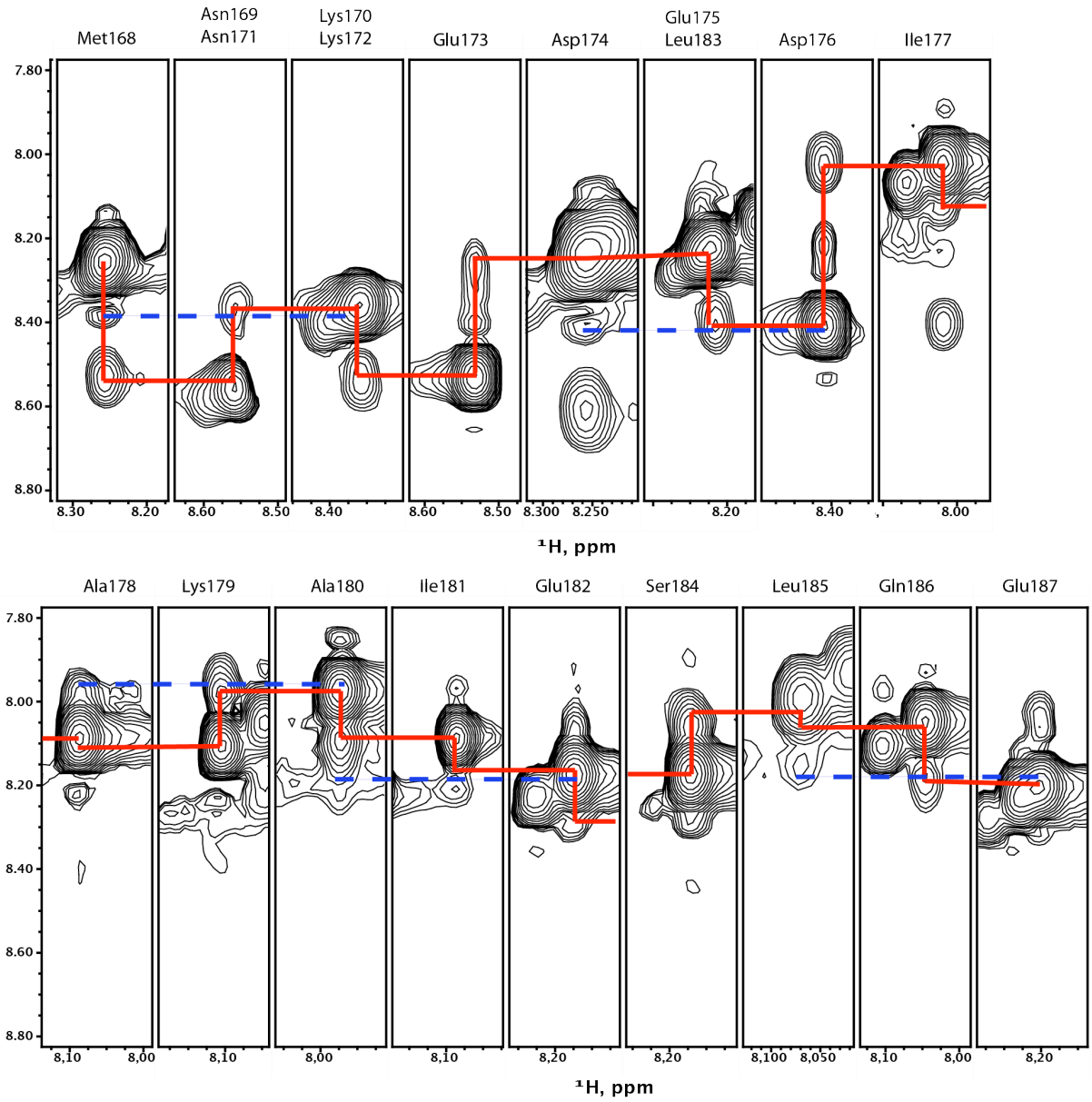
Supplementary Materials

Supplementary Figure S1:
Structural characterization of the STAM2 UIM domain by NMR. Strips from the 3D ^{15}N-dispersed NOESY spectrum of UIM showing cross-peaks corresponding to close NOE contacts between amide protons of residues *i* and *i*±1 (red lines) which are the hallmark of an α-helix. Also shown are detectable, less intense cross-peaks corresponding to NOE contacts between amide protons of residues *i* and *i+2* (blue dashed lines).

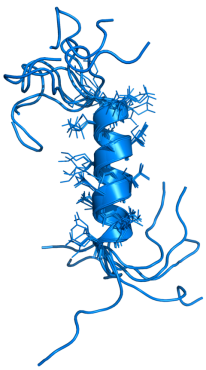


Supplementary Figure S2:
(A) Pairwise alignment of STAM2-UIM with STAM1-UIM, Vps27-UIM1 and Vps27-UIM2. Identical residues between sequences are colored red, whereas orange residues indicate similar residues. STAM1 UIM and STAM2 UIM share 80% identity and 100% similarity. The identity between STAM2 UIM and Vps27 UIM-1 is 55%, with a similarity of 70%. Alignment of STAM2 UIM with Vps27 UIM-2 revealed 40% identity and 70% similarity.
(B) STAM2 UIM structural models generated with "Modeller" homology modeling (1) based on the structure of Vps27 UIM1(2) (PDB code 1Q0V). (C) $^{3}J_{HN\text{-}H\alpha}$ coupling constants derived from ^{1}H,^{15}N Transverse Relaxation Optimized Spectroscopy (TROSY) experiments as a function of residue number. Residues that belong to the UIM helical region do not show any visible splitting due to the small coupling constant whereas residues located in the random coil region exhibit $^{3}J_{HN\text{-}H\alpha}$ coupling constants ranging from 5 to 8 Hz. (D) Expanded representative region of the TROSY spectrum of VHS-UIM. The spectrum clearly shows the presence of a $^{3}J_{HN\text{-}H\alpha}$ coupling constant for residues Thr193 and Thr157 whereas Ser184 does not show any visible splitting.

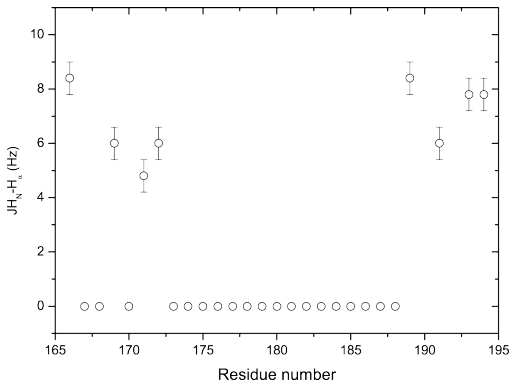
A

```
STAM2-UIM (human)              172 KEDEDIAKAIELSLQE
                                   ||:||:||||||||:||
STAM1-UIM (human)              171 KEEEDLAKAIELSLKE

STAM2-UIM (human)              172 KEDEDIAKAIELSLQE
                                   .|:|.|.|||||||:|.
Vps27-UIM1 (S. cerevisiae)     258 DEEELIRKAIELSLKE

STAM2-UIM (human)              172 KEDEDIAKAIELSLQE
                                   :||.|:..||:.||:|.
Vps27-UIM2 (S.cerevisiae )     301 EEDPDLKAAIQESLRE
```

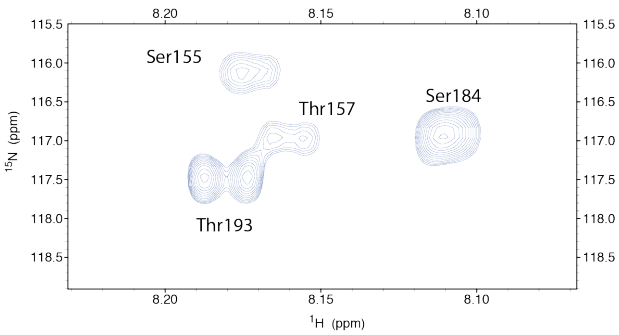
B



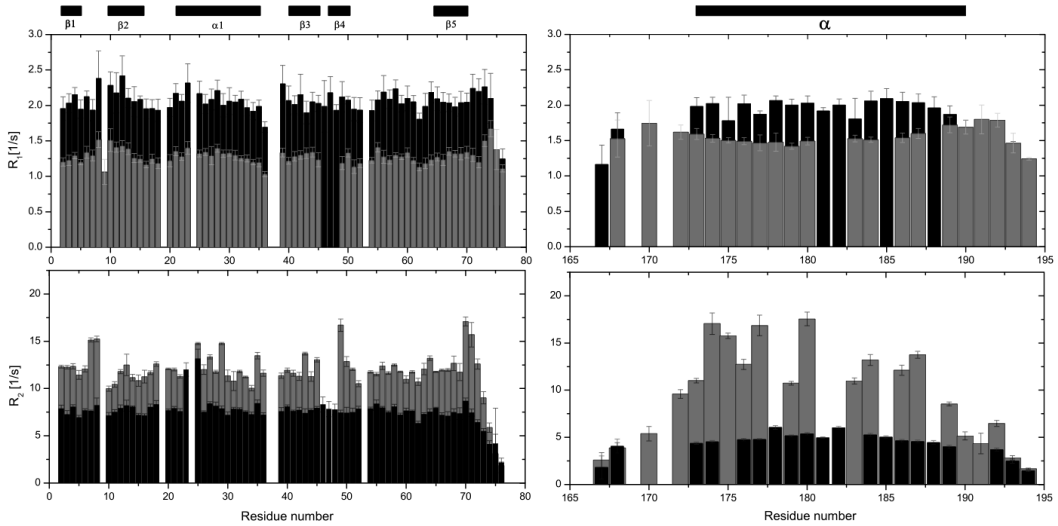
C



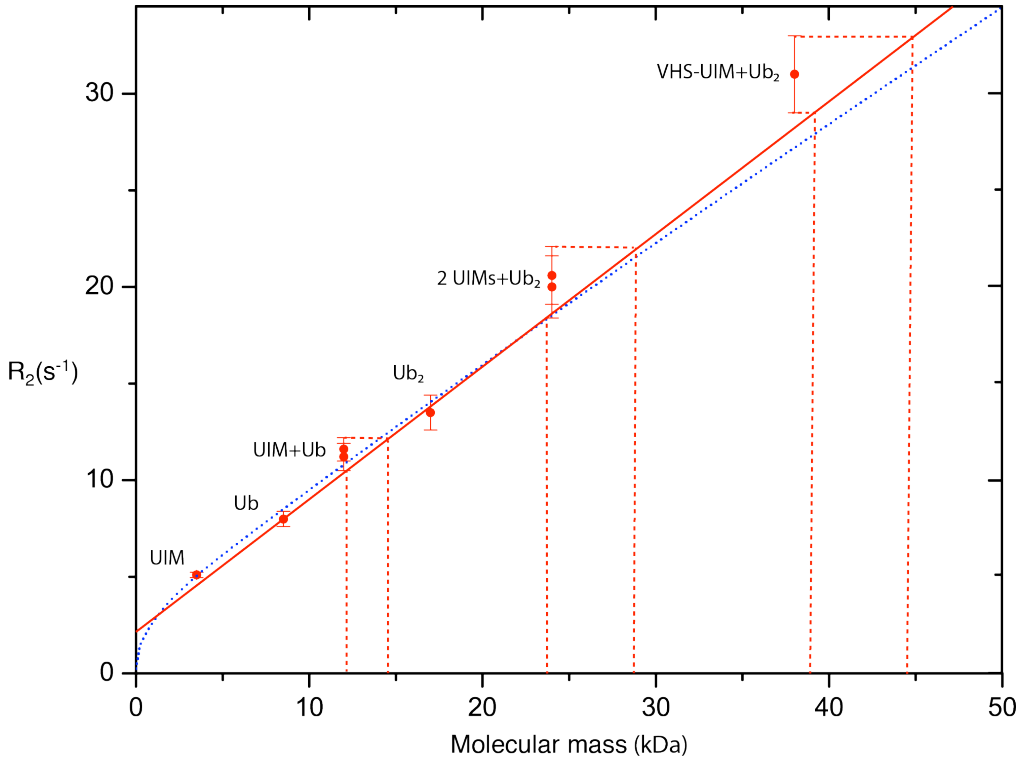
D



Supplementary Figure S3:
(A) ^{15}N relaxation rates, R_1 and R_2, for both monoUb (left panel) and UIM (right panel) in the free (black bars) and bound form (gray bars). The relative shift of the different levels reflects the change in the rate of the overall molecular tumbling for each partner (reflecting increased size/ molecular mass) upon complexation.

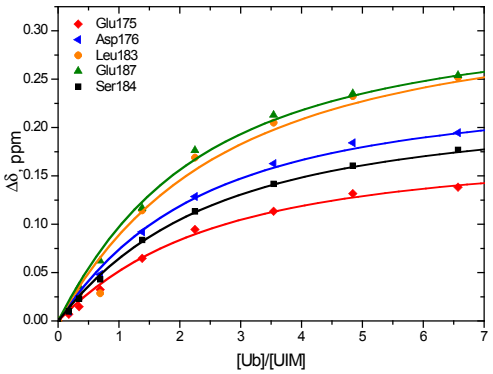


(B) ^{15}N transverse relaxation rate, R_2, as a function of molecular mass. The calibration red line was obtained by considering the experimental R_2 values of Ub and Lys63-Ub_2 distal Ub unit. This line can be used as a "molecular mass ruler" to determine the mass (hence stoichiometry) of molecular complexes of interest. The dashed red lines mark the molecular mass range (based on standard deviation for a given R_2 value) for each of the complexes. The slope of this line is 0.65±0.04 $s^{-1}.kDa^{-1}$, in good agreement with the slope derived by using the Stokes-Einstein equation (0.62±0.02 $s^{-1}.kDa^{-1}$). The latter one, represented by a blue doted line, was obtained by assuming isotropic overall tumbling of the proteins, a squared order parameter S^2 of 0.87 and a local correlation time of 50 ps. We assumed a protein specific density of 0.76 cm^3/g, a hydration shell of 3.2 Å, and solvent viscosity of 0.00114 Pa.s at 288°K (3). The equation used to derive the molecular mass as a function of R_2 (s^{-1}) is the following: $Mw=(R_2-2.5)/0.65$ and is associated with the red line.

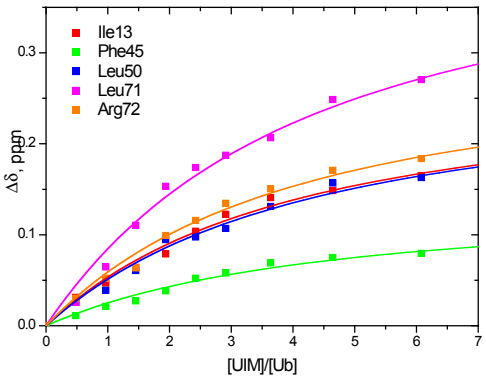


Supplementary Figure S4:
Titration curves on ^{15}N-UIM (A) and ^{15}N-monoUb (B) as a function of the molar ratio of the ligand and the ^{15}N-protein for the monoUb/UIM complex.

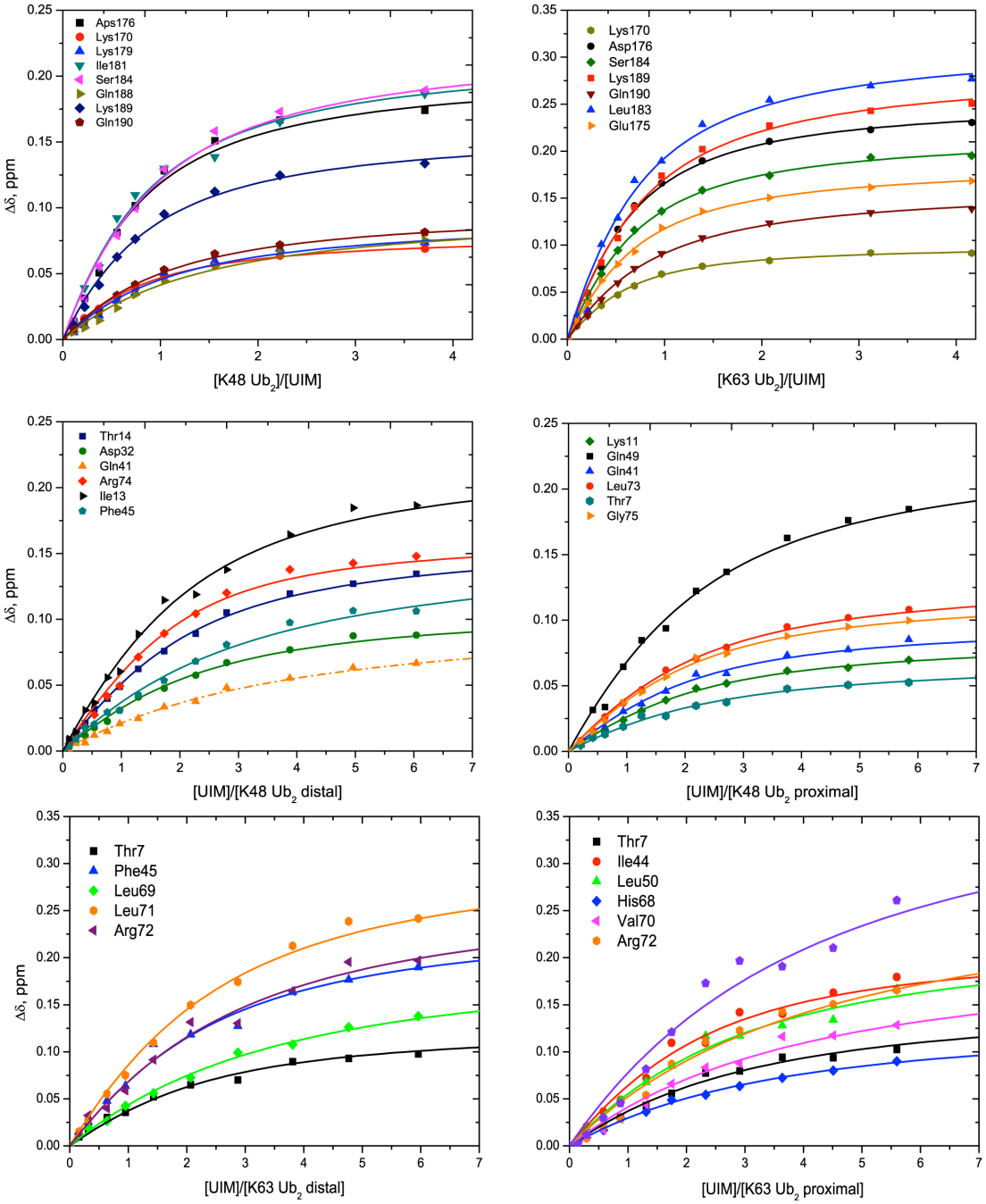
A



B



Supplementary Figure S5:
Titration curves for the UIM/Lys48-Ub_2 and UIM/Lys63-Ub_2 equilibrium as a function of the molar ratio of the ligand and ^{15}N-protein.



Supplementary Figure S6:
Comparison of the trajectories of the peak shifts along the titration of ^{15}N-Lys48- and ^{15}N-Lys63-Ub_2 with UIM, VHS and VHS-UIM. Residues that display the same trajectories are colored yellow whereas green indicates an intermediate trajectory between UIM and VHS.

On Lys63-Ub_2 distal

Residue	UIM	VHS	VHS-UIM
Thr7	↗	→	↗
Leu15	→	↓	→
Ile30	↙	↗	↙
Ile44	↙	↘	↙
Gly47	↖	↗	↖
Leu50	↖	↙	↖
Ile61	↙	→	↙
Glu64	↙	↗	↙
Ser65	←	→	←
His68	↘	→	↘
Leu69	↙	↖	↙
Leu71	↑	←	↑

On Lys63-Ub_2 proximal

Residue	UIM	VHS	VHS-UIM
Thr7	↗	→	→
Thr9	→	↙	↙
Ile30	↙	→	→
Ile44	↙	↘	↓
Leu50	↖	↙	←
Ile61	↙	→	→
Ser65	↙	→	→
His68	↘	↗	↗
Leu69	↙	←	←
Leu73	↖	←	↖
Gly75	↗	↗	↗

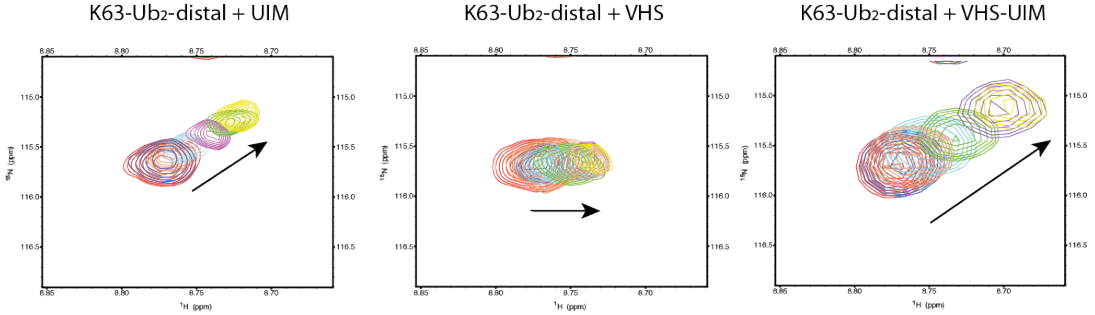
On Lys48-Ub_2 distal

Residue	UIM	VHS	VHS-UIM
Leu8	↙	↘	↓
Lys11	↙	→	↙
Leu15	→	↙	→
Lys29	↙	x	↙
Ile30	↙	→	↙
Glu32	↗	↑	↗
Gln41	↗	x	↑
Leu43	↗	↑	↑
Ile44	↙	↘	↓
Gln49	↑	↙	↙
Leu50	↑	↙	↙
Glu51	↙	←	←
Ile61	←	→	↘
His68	↘	→	↘
Leu69	↙	↑	↙
Arg72	↗	↙	↗

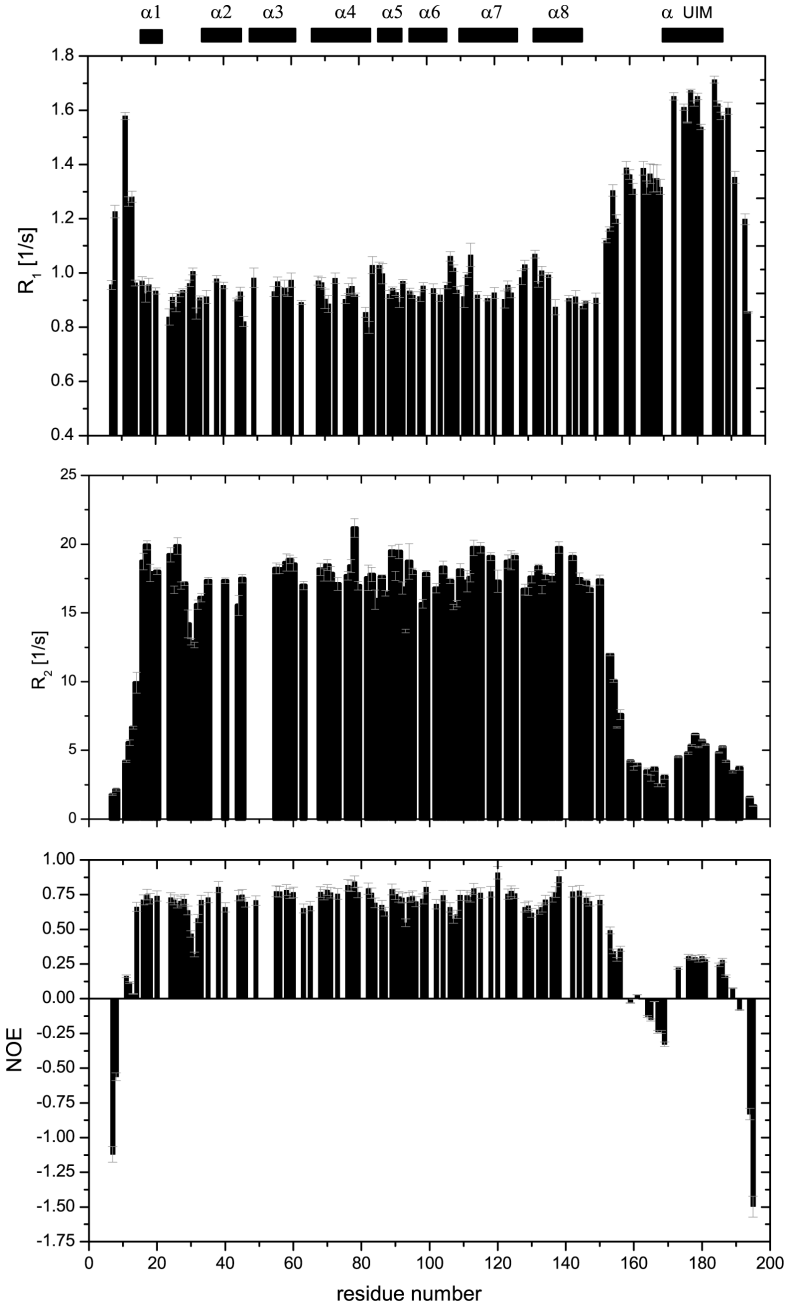
On Lys48-Ub_2 proximal

Residue	UIM	VHS	VHS-UIM
Thr9	↓	↙	↙
Lys11	↙	x	↙
Ile30	↙	↘	↙
Arg42	↖	↗	↖
Leu43	↑	→	↑
Gly47	↖	→	↑
Lys48	↑	↗	↑
Gln49	↑	←	↑
Leu50	↖	→	↑
Leu69	↙	←	↙
Leu73	↖	↑	↑
Gly76	←	↗	↖

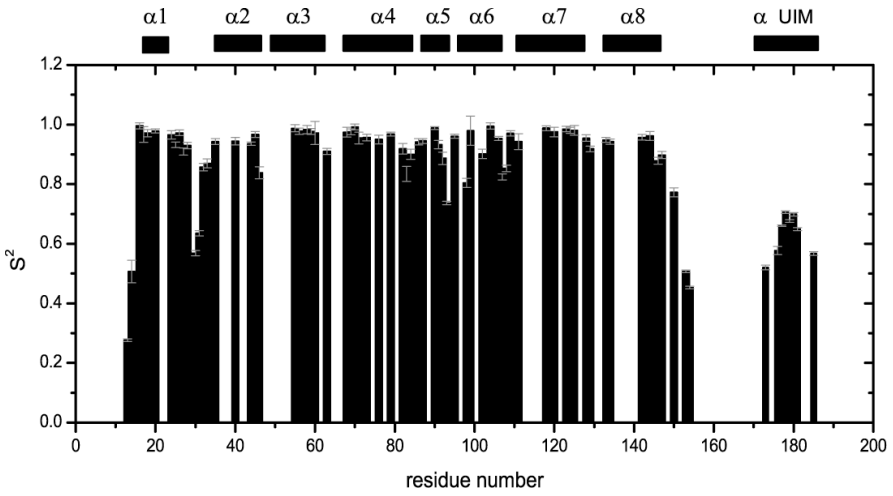
Example of peak shift encountered by Thr7 of Lys63-Ub_2 distal upon titration with UIM, VHS and VHS-UIM



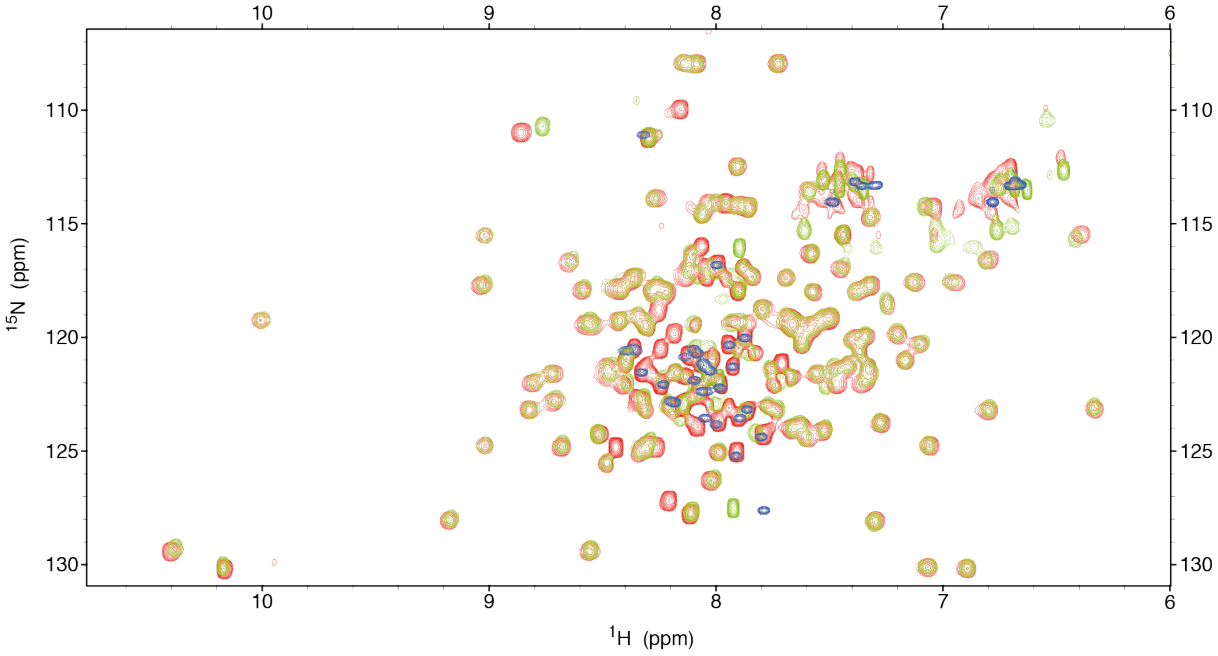
Supplementary Figure S7:
(A) Relaxation parameters R_1, R_2 and heteronuclear NOE for VHS-UIM in the free form. The differences in the R_1 and R_2 levels for VHS and UIM indicate that they tumble independently from each other. Negative and close to zero heteronuclear NOE values also indicate a high degree of backbone flexibility in the intervening linker as well as in the N and C termini of the VHS-UIM construct.



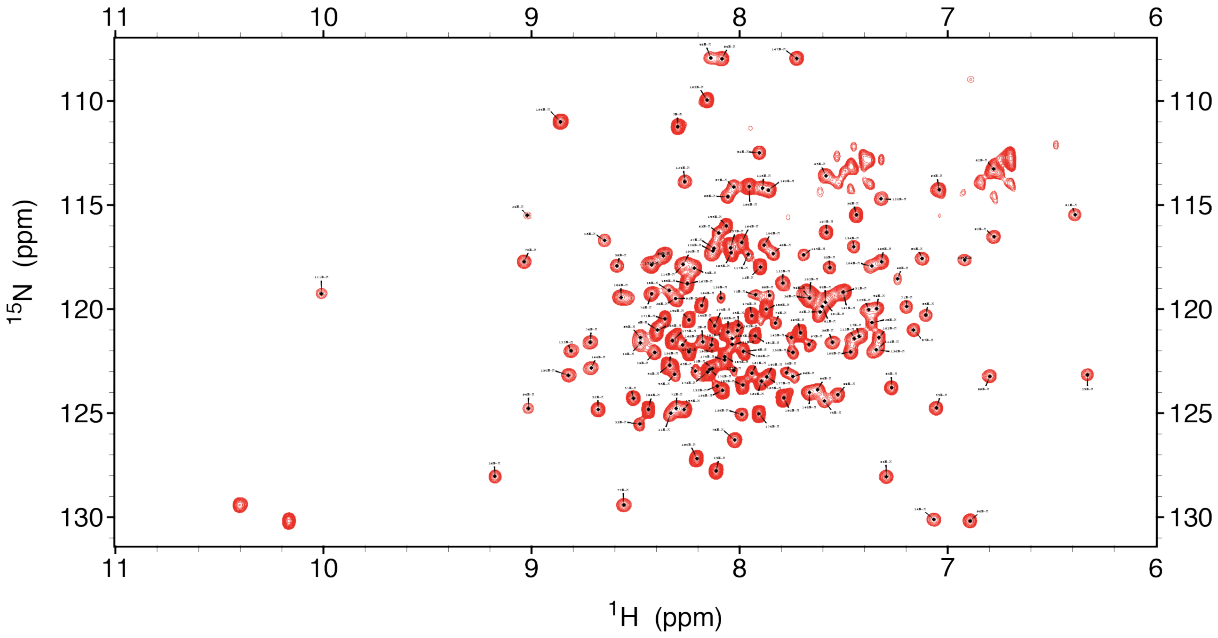
(B) Model-free analysis of the VHS-UIM domain in the free state showing the order parameter S^2 as a function of residue number. The VHS and UIM domains were treated independently. The analysis was performed by using the program DYNAMICS (4,5). Prior to the model-free analysis, the overall rotational diffusion tensor was determined by using the program ROTDIF (6). For the VHS domain, the fully anisotropic model gives the best fit and the following diffusion tensor parameters were used as input in DYNAMICS: D_x= 0.89±0.04 * 10^7 s^{-1}, D_y=0.98±0.05 * 10^7 s^{-1}, D_z=1.37±0.04 * 10^7 s^{-1}, α= 29°, β=87°, γ=165° and the overall rotational correlation time of 15.6±0.4 ns. For the UIM domain, the isotropic model was used with the overall rotational correlation time of 5.8±0.1 ns, as no structural information is available. The ^{15}N CSA value was set to -160 ppm for the model-free analysis.



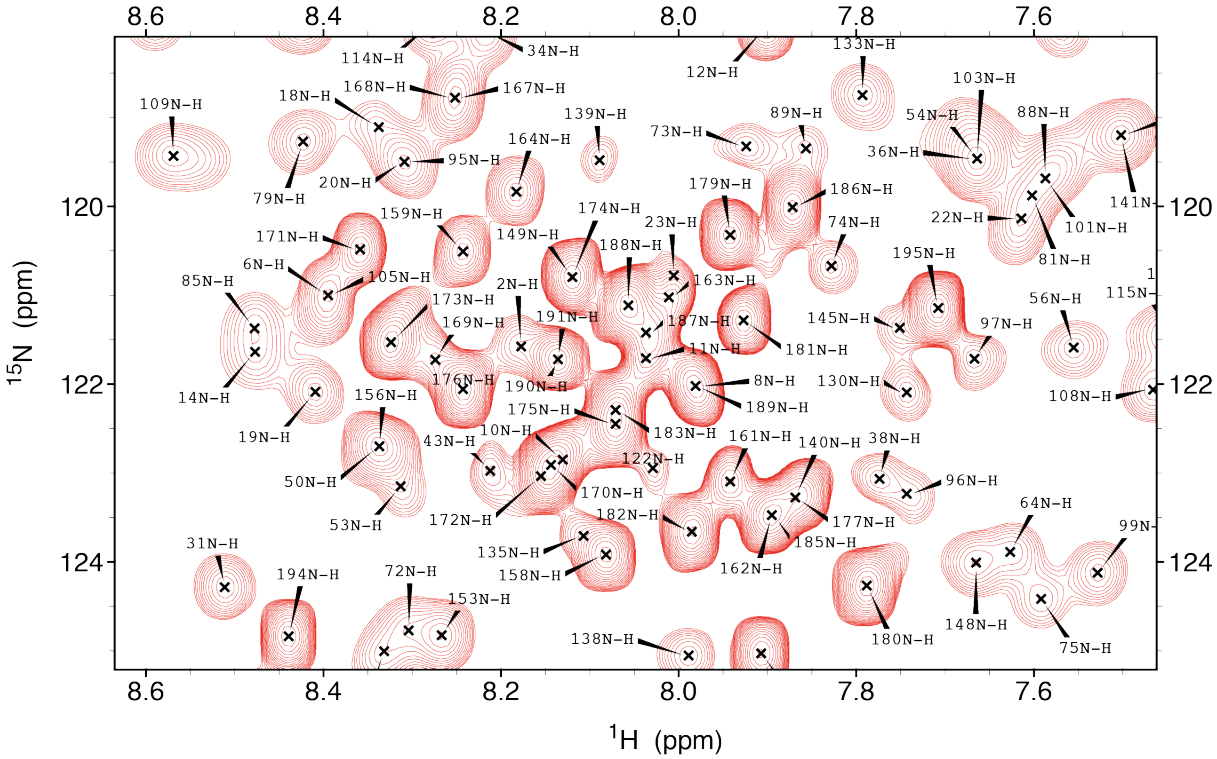
Supplementary Figure S8:
Overlay of ^{1}H,^{15}N-HSQC spectra of VHS-UIM(red), VHS (green) and UIM (blue).



Annotated ^{1}H,^{15}N-HSQC spectrum of the VHS-UIM construct.

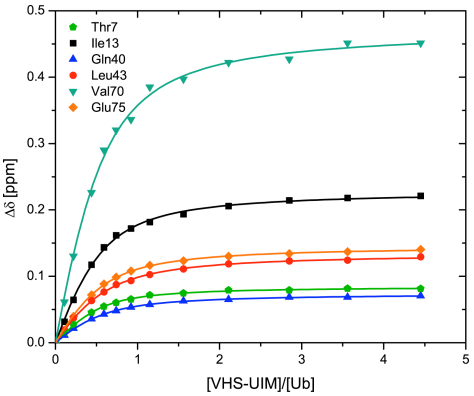


Expanded region of the ^{1}H,^{15}N-HSQC spectrum of the VHS-UIM construct

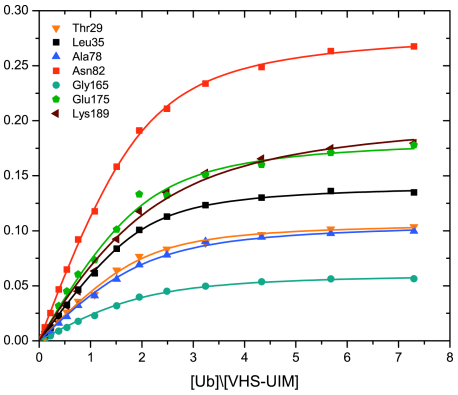


Supplementary Figure S9:
Titration curves for ^{15}N-monoUb (A) and ^{15}N-VHS-UIM (B) for the VHS-UIM/monoUb complex

A

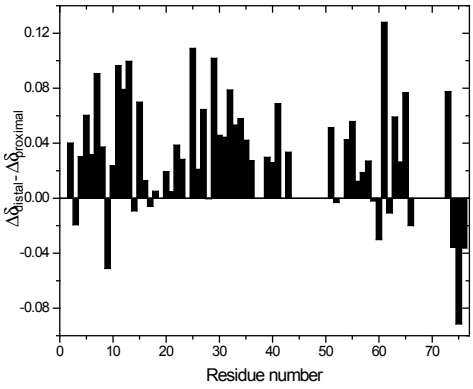


B

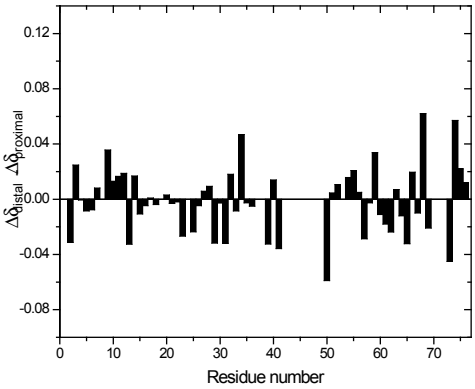


Supplementary Figure S10:
Differences in CSPs between the distal and the proximal domain of ^{15}N-Lys48- and ^{15}N-Lys63-Ub_2 under complexation with the VHS-UIM domain. Differences are taken under saturating conditions and for a similar [ligand]/[protein] ratio. This clearly indicates a difference in binding between the VHS-UIM/Lys48-Ub_2 and VHS-UIM/Lys63-Ub_2 complex.

Lys63-Ub_2

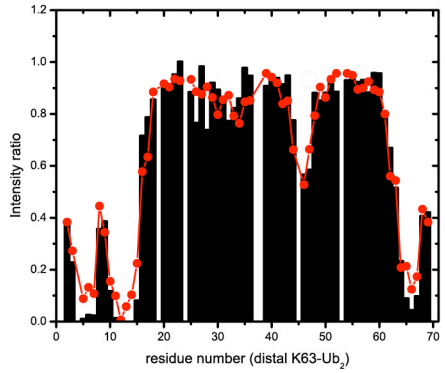


Lys48-Ub_2

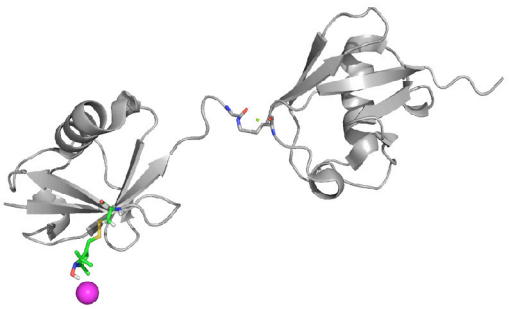


Supplementary Figure S11:
Reconstruction of the unpaired electron's position of MTSL in the case of MTSL attached to the distal domain of ^{15}N-K63-Ub$_2$. (A) The experimental intensity ratio is represented by black bars whereas the red symbols (connected by a line) represent back-calculated PREs for the fitted MTSL position. (B) The reconstructed position of the unpaired electron is shown as a sphere colored magenta while MTSL added *in silico* is shown in sticks representation.
Distances from MTSL derived from PREs measured in the VHS domain of ^{15}N-VHS-UIM + MTSL attached to the proximal Ub of Lys63-Ub$_2$ (C) and in the UIM domain of ^{15}N-VHS-UIM + MTSL attached to the distal Ub of Lys63-Ub$_2$ (D). The black line represents the distances derived from PRE data while the dotted lines describe an upper and lower distance of ±4 Å for the MTSL position fitted from PRE data and likely reflects the flexibility of MTSL. The red symbols and red lines represent the back-calculated distances from the best structure of the modeled VHS-UIM/Lys63-Ub$_2$ complex.

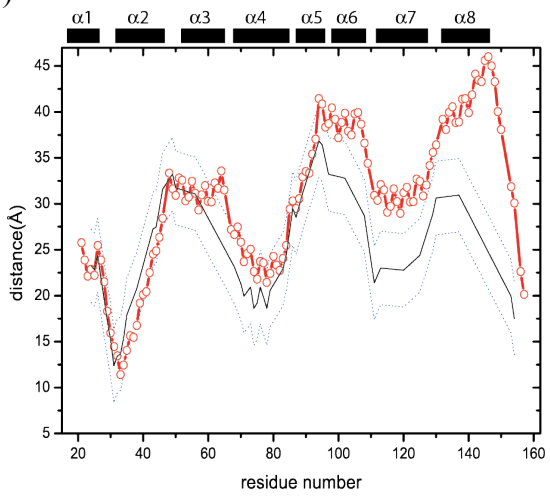
A)



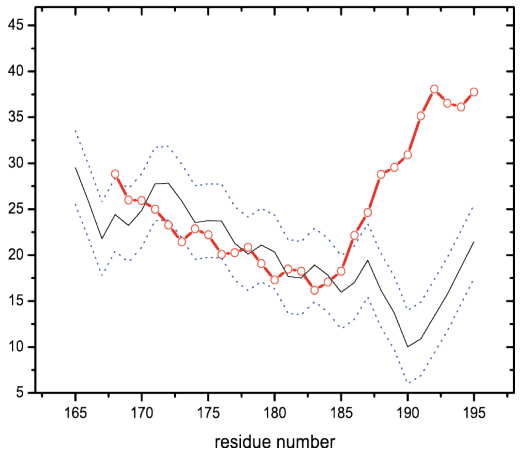
B)



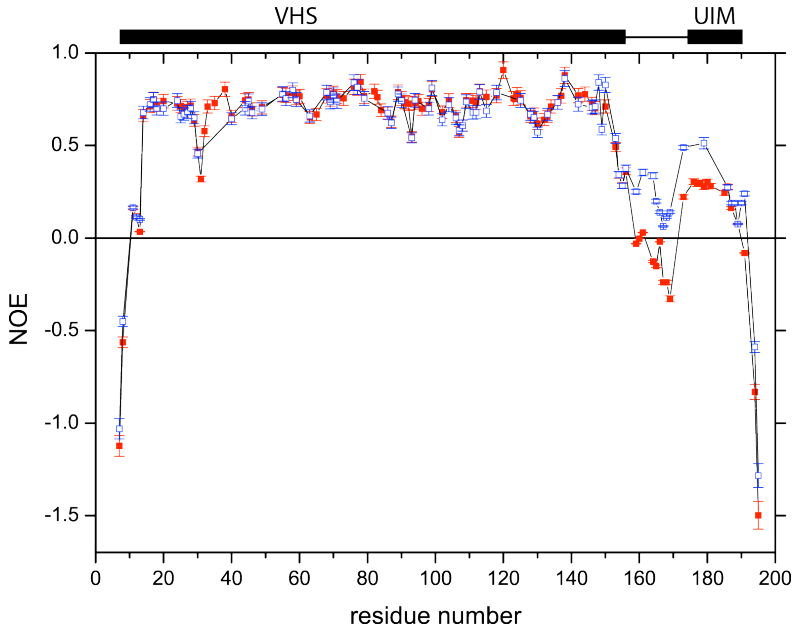
C)



D)



Supplementary Figure S12:
Heteronuclear $^{15}N\{^{1}H\}$NOE for VHS-UIM in the free state (red squares) and in complex with Lys63-Ub_2 at the end of titration (open blue squares). For VHS-UIM in the bound state, residues showing overlapped signal or low signal-to-noise ratio were excluded from the plot.

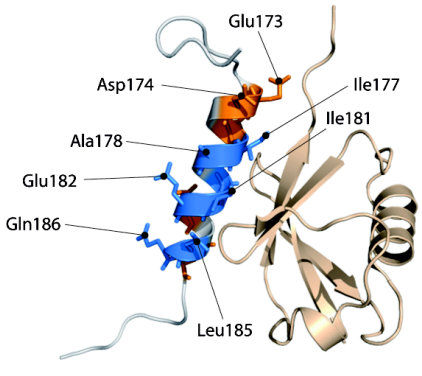


Supplementary Figure S13:
Sequence alignment of several UIM domains (A). Open circles represent UIM residues that experience intermediate exchange upon binding to monoUb. (B) Representation of the UIM/monoUb complex based on the alignment of the modeled STAM2 UIM domain on the Vps27 UIM1/monoUb complex (PDB code 1Q0W). Residues experiencing intermediate exchange upon binding to monoUb are colored blue. Extra residues experiencing intermediate exchange upon binding to Ub_2 are colored orange.

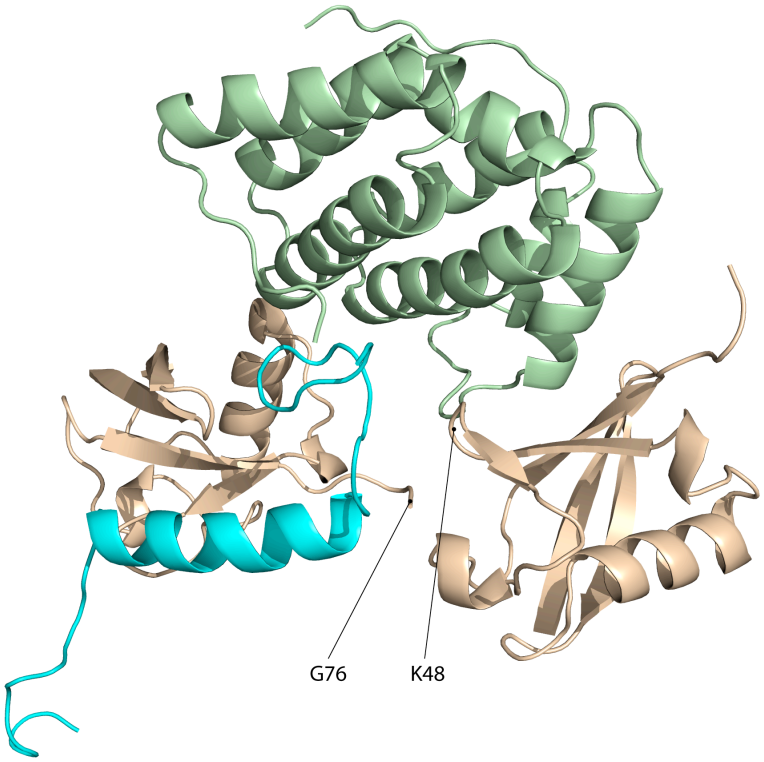
A

```
                          ○○  ○○○   ○○○
STAM2 UIM          172  -KEDEDIAKAIELSLQEQKQQ  191
STAM1 UIM          171  -KEEEDLAKAIELSLKEQRQQ  190
Vps27 UIM-1        258  -DEEELIRKAIELSLKESRNS  277
Vps27 UIM-2        301  -EEDPDLKAAIQESLREAEEA  320
Hse1 UIM           262  -SDDEELQKALKMSLFEYEKQ  181
Ataxin3 UIM-1      224  -EDEEDLQRALALSRQEIDME  243
Ataxin3 UIM-2      244  -DEEADLRRAIQLSMQGSSRN  263
Eps15 UIM-2        878  -QEQEDLELAIALSKSEISEA  897
Eps15 UIM-1        852  -SEEDMIEWAKRESEREEEQR  871
USP25 UIM-1         97  -DDKDDLQRAIALSLAESNRA  116
USP25 UIM-2        123  TDEEQAISRVLEASIAEN---  140
Epsin UIM-1        852  -EEELQLQLALAMSKEEADQE  871
Epsin UIM-2        208  -GDDLRLQMAIEESKRETGGK  227
Rap80 UIM-2        104  EEEEELLRKAIAESLNSCRPS  124
Rap80 UIM-1         79  MTEEEQFALALKMSEQEA---   96
S5A UIM-2          211  -TEEEQIAYAMQMSLQGAEFG  230
```

B



Supplementary Figure S14:
Modeling of the VHS-UIM/Lys48-Ub_2 complex. The resulting complex would prevent the binding of VHS and UIM at the same time due to sterical hindrance of the VHS domain.



Deriving the MTSL position from PRE data

The PRE data were converted to distances between the unpaired electron and the amide proton spins according to the equations previously established (7-10):

$$\Delta R_{2\,para} = \frac{1}{20}\gamma_H^2 g_e^2 \beta_e^2 \left(4\tau_c + \frac{3\tau_c}{1+\omega_H^2 \tau_c^2} \right) \frac{1}{r^6} \; ; \; I_{ox}/I_{red} = \exp\left(-t\Delta R_{2\,para}\right) \text{ and}$$

$$\Delta R_{2\,para} = R_{2ox} - R_{2red} \qquad [1]$$

where τ_c is the rotational correlation time of the molecule. γ_H and ω_H are the ^{1}H gyromagnetic ratio and resonance frequency, g_e is the electronic g-factor, β_e is the Bohr magneton and *r* is the distance between the nucleus under observation and the unpaired electron. Here *t* is the time the ^{1}H magnetization spends in the transverse plane during the pulse sequence.
τ_c was estimated by assuming a linear dependence of τ_c with the molecular weight. R_{2dia} was taken from the transverse relaxation rate of Lys63-Ub_2 bound to VHS-UIM. The position of the unpaired electron was derived after fitting the observed PREs for selected residues in structured parts of the protein. The fit was carried out using the SLfit program (10). Distances were calculated between the position of the paramagnetic center and the position of a given amide proton.

The sequential model for NMR

The two site sequential model can be summarized as:

$$[P] + [L] \underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}} [PL_1] + [L] \underset{k_{-2}}{\overset{k_{+2}}{\rightleftharpoons}} [PL_2] \qquad [2]$$

where P stands for the free protein, PL1 the partly bound protein and PL2 the totally bound protein. The two dissociation constants can be written as:

$$Kd_1 = \frac{k_{-1}}{k_{+1}} = \frac{[P][L]}{[PL_1]} \quad ; \quad Kd_2 = \frac{k_{-2}}{k_{+2}} = \frac{[PL_1][L]}{[PL_2]} \qquad [3]$$

Assuming fast exchange, the chemical shift perturbation can be rewritten as:

$$\Delta\delta = p_P \Delta\delta_P + p_{PL_1} \Delta\delta_{PL_1} + p_{PL_2} \Delta\delta_{PL_2} \qquad [4]$$

where p_i is the corresponding populations of the different complexes. The concentrations of the different complexes are:

$$[P] = \frac{[P_T] \cdot Kd_1 \cdot Kd_2}{Kd_1 \cdot Kd_2 + [L] \cdot Kd_2 + [L]^2}$$

$$[PL_1] = \frac{[L] \cdot [P_T] \cdot Kd_2}{Kd_1 \cdot Kd_2 + [L] \cdot Kd_2 + [L]^2}$$

$$[PL_2] = \frac{[L]^2 \cdot [P_T]}{Kd_1 \cdot Kd_2 + [L] \cdot Kd_2 + [L]^2} \qquad [5]$$

where P_T stands for the total protein concentration. $\Delta\delta$ can be recast as:

$$\Delta\delta = \frac{Kd_1 \cdot Kd_2 \cdot \Delta\delta_P + [L] \cdot Kd_2 \cdot \Delta\delta_{PL_1} + [L]^2 \cdot \Delta\delta_{PL_2}}{Kd_1 \cdot Kd_2 + [L] \cdot Kd_2 + [L]^2} \quad [6]$$

where the two extreme values correspond to the free ($\Delta\delta_P$=0) and totally bound protein ($\Delta\delta=\Delta\delta_{PL2}$) and [L] is obtained by solving the following cubic equation:

$$[L]^3 + [L]^2\left(Kd_2 + 2[P_T] - [L_T]\right) + [L]\left(Kd_1.Kd_2 + Kd_2.[P_T] - Kd_2.[L_T]\right) - [L_T].Kd_1.Kd_2 = 0 \quad [7]$$

The three parameters comprising Kd_1, Kd_2 and $\Delta\delta_{PL1}$ were fitted with non-linear regression by using an in-house Matlab (The Mathworks, Inc) based program.

Using the worm-like chain model for NMR

Following the work of Zhou(11,12), the overall dissociation constant is given by:

$$Kd_a = Kd_1 \cdot Kd_2 \,/\, p(d_0) \quad [8]$$

where $p(d_0)$ is the probability density for the end-to-end vector of the flexible linker with L residues to have a distance r and d_0 is the actual end-to-end distance when the linked VHS and UIM are bound to Lys63-Ub_2. Thus

$$p(r) = \left(3/4\pi l_p l_c\right)^{3/2} \exp\left(-3r^2/4l_p l_c\right)\begin{pmatrix} 1 - 5l_p/4l_c + 2r^2/l_c^2 - 33r^4/80l_p l_c^3 - 79l_p^2/160l_c^2 - 329r^2 l_p/120l_c^3 \\ +6799r^4/1600l_c^4 - 3441r^6/2800l_p l_c^5 + 1089r^8/12800l_p^2 l_c^6 \end{pmatrix} \quad [9]$$

where b=3.8 Å is the nearest C_α-C_α distance, and l_c=bL and lp=3 Å. L is taken equal to 20 as VHS and UIM are separated by a 20 amino acid linker and the distance r has been set to 23Å, which corresponds to the end-to-end length of the VHS-UIM linker for the best structure of the modeled VHS-UIM/Lys63-Ub_2 complex.

Docking of the Lys63-Ub_2/VHS-UIM complex

The model of the Lys63-Ub_2/VHS-UIM complex has been calculated with the HADDOCK2.0 program (13,14), starting with four individual structures: the UIM (modeled by homology), VHS (PDB code 1X5B) and two Ub subunits (PDB code 1D3Z). To ensure for the right connection between VHS and UIM as well as the two Ub subunits, distance constraints were introduced to drive the docking. To account for the linker flexibility, the 20 amino-acids separating VHS and UIM were considered fully flexible.

To take into account PRE data, a cysteine residue and MTSL have been introduced in silico at position 12 on the two Ub subunits. The PRE intermolecular distances were introduced as distances separating the oxygen atom of MTSL and a given amide proton of VHS with an additional upper and lower bound distance of 4Å to reflect the inherent flexibility of MTSL. In addition, CSPs were introduced to define ambiguous interaction restraints (AIRs). Active residues were defined as those having CSPs above 0.05 ppm and 0.1 ppm for VHS and

monoUb respectively and a relative residue accessible surface area larger than 50% for side chains or backbone atoms. Passive residues were identified as residues close to active residues. Flexible segments were defined as stretches of active and passive residues plus one sequential residue. To ensure contact between the different partners, a force constant for the center of mass was turned on. Also used were distance constraints arising from the VHS/monoUb complex(15) as well as homology modeling with respect to the Vps27/monoUb structure(2) (PDB code 1Q0W). Finally, the structure calculation was performed with the standard three step refinement found in HADDOCK with i) a rigid body energy minimization generating 4000 structures ii) semi-flexible refinement using the best 300 structures iii) final refinement in water. The 200 final solutions were ranked according to the Haddock score, defined as $1.0*E_{vdw}+0.2*E_{elec}+0.1*E_{AIR}+1.0*E_{desolv}$. Thus, they were clustered based on a backbone RMSD of 5.5Å and the 10 best structures of the best cluster in terms of Haddock score were retained for analysis. All docking calculations have been carried out on the e-NMR web portal.

Supplementary Table S1: Haddock restraints as well as the results of the structure calculation for the ten best structures of the best cluster for the VHS-UIM/Lys63-Ub_2 interaction.

Haddock restraints

Ambiguous restraints from CSPs	Active	Passive
	4 (Lys63-Ub_2 distal) 4 (Lys63-Ub_2 proximal) 6 (VHS) 3(UIM)	3 (Lys63-Ub_2 distal) 3 (Lys63-Ub_2 proximal) 4 (VHS) 3(UIM)
Unambiguous distance restraints from PREs	16 (9 from Lys63-Ub_2 proximal ; 7 from Lys63-Ub_2 distal)	
from homology modelling	10	
Distal-proximal isopeptide bond	4	
VHS-UIM linker	4	

Lys63-Ub_2/VHS-UIM complex structure calculation

E_{inter}[a]	E_{vdW}	E_{elec}	E_{NOE}[b]	BSA[d]	H_{bond}[e]	Hydro[f]	Score[g]	Rmsd[h]
-676.1	-59.1	-617.0	52.0	2977	17.8	27	-87.0	2.05
(64.0)	(7.5)	(68.9)	(3.4)	(165)	(1.5)	(2)	(8.0)	(0.95)

Procheck analysis	
Residues in most favoured region (%)	88.8
Residues in additional allowed region (%)	10.6
Residues in generously allowed regions (%)	0.4
Residues in disallowed regions (%)	0.2

[a] Intermolecular energy: sum of the van der Waals and electrostatic energies ($kcal.mol^{-1}$)
[b] NOE energy: sum of the ambiguous and unambiguous energies ($kcal.mol^{-1}$)
[d] Total buried surface area for Ub_2 and VHS-UIM ($Å^2$)

[e] Number of hydrogen bonds
[f] number of hydrophobic contact
[g] Haddock score (arbitrary units)
[h] rmsd calculated with respect to the lowest Haddock score structure

References:

1. Sali, A., and Blundell, T. L. (1993) *J Mol Biol* **234**, 779-815
2. Swanson, K. A., Kang, R. S., Stamenova, S. D., Hicke, L., and Radhakrishnan, I. (2003) *EMBO J* **22**, 4597-4606
3. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., Skelton, N. J., and Rance, M. (2006) *Protein NMR Spectroscopy*, 2e edition ed., Academic Press
4. Fushman, D. (2012) *Methods Mol Biol* **831**, 485-511
5. Fushman, D., Cahill, S., and Cowburn, D. (1997) *J Mol Biol* **266**, 173-194
6. Walker, O., Varadan, R., and Fushman, D. (2004) *J Magn Reson* **168**, 336-345
7. Battiste, J. L., and Wagner, G. (2000) *Biochemistry* **39**, 5355-5365
8. Jain, N. U., Venot, A., Umemoto, K., Leffler, H., and Prestegard, J. H. (2001) *Protein Sci* **10**, 2393-2400
9. Fushman, D., Varadan, R., Assfalg, M., and Walker, O. (2004) *Prog Magnetic Reson Spectr* **44**, 189-214
10. Ryabov, Y., and Fushman, D. (2006) *Proteins* **63**, 787-796
11. Zhou, H. X. (2010) *Q Rev Biophys* **43**, 219-293
12. Zhou, H. X. (2003) *J Mol Biol* **329**, 1-8
13. Dominguez, C., Boelens, R., and Bonvin, A. (2003) *J Am Chem Soc* **125**, 1731-1737
14. Vries, S., Dijk, A., Krzeminski, M., Dijk, M., Thureau, A., Hsu, V., Wassenaar, T., and Bonvin, A. M. (2007) *Proteins* **69**, 726-733
15. Lange, A., Hoeller, D., Wienk, H., Marcillat, O., Lancelin, J.-M., and Walker, O. (2011) *Biochemistry* **50**, 48-62