Supplementary Methods & References

Plasmids and cloning
cDNA for human USP4 and USP8 was a gift from Hidde Ploegh and cDNA for USP39 was a gift from R. Medema. USP4CD (aa 296-954) and the D1 fragment (aa 296-490) of human USP4, USP39CD (aa 219-565) and USP8CD (aa 771-1118) were cloned using ligation independent cloning into pET-46 Ek/LIC vector (Novagen). The D2 fragment (aa 766-932) of USP4 was cloned into the pET-NKI b/3C (Luna-Vargas, in preparation). The fused USP4-D1D2 was created by inserting aa 353-359 of USP7 (SIKGKNN) between residues Leu479 and Leu777. The USP4 insert (aa 483-765), Ubl domain (aa 483-571) and Ub were cloned into pGEX-6P-1 (GE Healthcare). The USP8CD mutant was generated by site-directed mutagenesis of the catalytic cysteine (C786A).

Protein preparation
Purification of E2-25K (Pichler et al, 2005), Ubc13/Mms2 (Marteijn et al, 2009) was as described. GST-tagged proteins were overexpressed in *Escherichia coli* strain Rosetta2(DE3)-T1R using IPTG (200μM) induction overnight at 15°C. Cells were lysed by microfluidizer into buffer A (50mM Hepes pH7.5, 150mM NaCl, 5mM β-mercaptoethanol, 1mM PMSF). The fusion protein was purified using glutathione sepharose resin, eluted, followed by removal of the GST-tag with 3C protease and size-exclusion using HiLoad 16/60 Superdex 200 (GE Healthcare). Peak fractions were concentrated to 10mg/ml in 25mM Hepes (pH7.5), 200mM NaCl and 5mM β-mercaptoethanol.
D1 and D2 co-expression, other USP4 variants, USP39CD and USP8CD-mut were overexpressed as above, with 200μM $ZnCl_2$ during induction and lysed in buffer A supplemented with 1mM $ZnCl_2$ and 10mM Imidazole. These His-tagged proteins were purified by a Co^{2+}-affinity (Talon resin) step. Upon Imidazole elution the His-tag was removed by TEV cleavage at 4°C overnight during dialysis in buffer B (25mM Hepes pH7.5, 150mM NaCl, 5mM β-mercaptoethanol). This was followed by POROS Q affinity chromatography and size-exclusion using HiLoad 16/60 Superdex 200 (GE Healthcare), where the protein eluted as a monomer. The peak fractions were concentrated to 5mg/ml in buffer B.

Limited Proteolysis and protein identification
Purified USP4CD (9mg/ml) was incubated with Thermolysin (0.8units) for 1,5hr at room temperature and subjected to size exclusion chromatography using Superdex75 16/60. Fractions containing USP4-D1 and –D2 were subjected to LC-MS analysis. LC-MS measurements were performed on a system equipped with a Waters 2795 Seperation Module (Alliance HT), Waters 2996 Photodiode Array Detector (190-750nm), Waters Alltime C18 (2.1x100mm, 3μm), Waters Symmetry300$_{TM}$ C4 (2.1x100mm, 3.5μm) and LCT$_{TM}$ Orthogonal Acceleration Time of Flight Mass Spectrometer. Data processing was performed using Waters MassLynx Mass Spectrometry Software 4.1 (deconvolution with Maxent 1 function).
N-terminal sequencing of USP4-D1 and –D2 were performed by AltaBioscience in Birmingham, England.

Crystallization and structure determination of the USP4-D1D2
Crystals were grown overnight in sitting-drops mixing 200nl USP4-D1D2 (~3.5mg/ml) with 200 nl 100mM Bis-Tris propane [pH8.5], 25mM Na_2SO_4 and 18% PEG3350 (w/v) at 19°C. Crystals were cryoprotected in mother liquor with 25% ethyleneglycol. The crystals belong to the space group $P2_12_12_1$ with six molecules per asymmetric unit (supplementary Table S1). Diffraction data were collected at the ESRF (Grenoble, France) beamline ID14-2 and processed with MOSFLM (Leslie, 2006) and SCALA (Evans, 2006). The structure was solved by molecular replacement with PHASER (McCoy et al, 2007) using USP8CD (PDB:2GFO) as search model. Iterative rebuilding and refinement were done with Coot (Emsley & Cowtan, 2004) and PHENIX (Terwilliger et al, 2008) and BUSTER (Blanc et al, 2004). The structure was validated with MOLPROBITY (Davis et al, 2007) and WHAT-CHECK (Hooft et al, 1996) and structure figures were generated using PYMOL (Delano, 2002). Cysteine residue 311 in all chains have been chemically modified by β-mercaptoethanol.

Ub-AMC assay
UbAMC assays were done in 50mM Hepes [pH7.5], 100mM NaCl, 5mM DTT, 0,05% Tween-20 and 1mM EDTA and reaction progress was monitored with a Fluostar Optima plate-reader (BMG Tech) by the increase in fluorescence emission at 460nm (λ_{ex} = 355nm) generated by Ub-AMC cleavage.

Di-ubiquitin assay
Di-Ub assays were performed in similar buffer as in UbAMC assays at 37°C in 75µl reaction volume. Aliquots (5µl) were stopped by addition of 4x SDS-sample loading buffer and subjected to SDS-PAGE analysis on a 4-12% coomassie stained gel (Invitrogen).

Surface plasmon resonance
Quantitative binding analysis was done in duplicate at 25°C on a Biacore T-100 instrument (GE Healthcare). GST fused Ub, insert and Ubl domain were immobilized on α-GST antibodies lysine-coupled to a CM5 chip. USPs were injected in varying concentrations over the sensor chip at 30µl/min with a 120s association phase followed by a 10min dissociation phase. For the binding inhibition assay Ub was added in varying concentrations to USP4-D1D2. Standard double referencing data subtraction methods were used before and equilibrium curve fitting with BiaEvaluation (GE Healthcare) and GraphPad software (GraphPad Software Inc).

Isothermal Titration Calorimetry
Isothermal titration calorimetry (ITC) experiments were performed with the VP-ITC Micro Calorimeter (MicroCal, Inc.) at 25°C. Stock solutions of USP4CD, USP4-D1D2 and Ub were prepared by dialysis of the purified proteins against a buffer containing 25mM Hepes pH8.0, 150mM NaCl and 5mM β-mercaptoethanol at 4°C and were degassed before use. The sample cell (1.8ml) contained USP4-D1D2 (10µM) or USP4CD (20µM) which was titrated with 100µM Ub or 200µM Ub respectively using 16 injections. The injections after saturation were used to determine the background signal. Corrected data were analyzed using software supplied by the ITC manufacturer to calculate the dissociation constant K_d and fitted with a one to one binding model.

References

Blanc E, Roversi P, Vonrhein C, Flensburg C, Lea SM, Bricogne G (2004) Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr D Biol Crystallogr* 60: 2210-2221

Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, 3rd, Snoeyink J, Richardson JS, Richardson DC (2007) Molprobity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**: W375-383

Delano WL (2002) The Pymol Molecular Graphics System. San Carlos, CA, USA: DeLano Scientific

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**: 2126-2132

Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**: 72-82

Hooft RWW, Vriend G, Sander C, Abola EE (1996) Errors in protein structures. *Nature* **381**: 272-272

Leslie AG (2006) The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* **62**: 48-57

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**: 658-674

Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, Read RJ, Adams PD (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* **64**:61-69

Supplemental Table S1
Crystallographic parameters.

Table SI Data collection and refinement statistics.

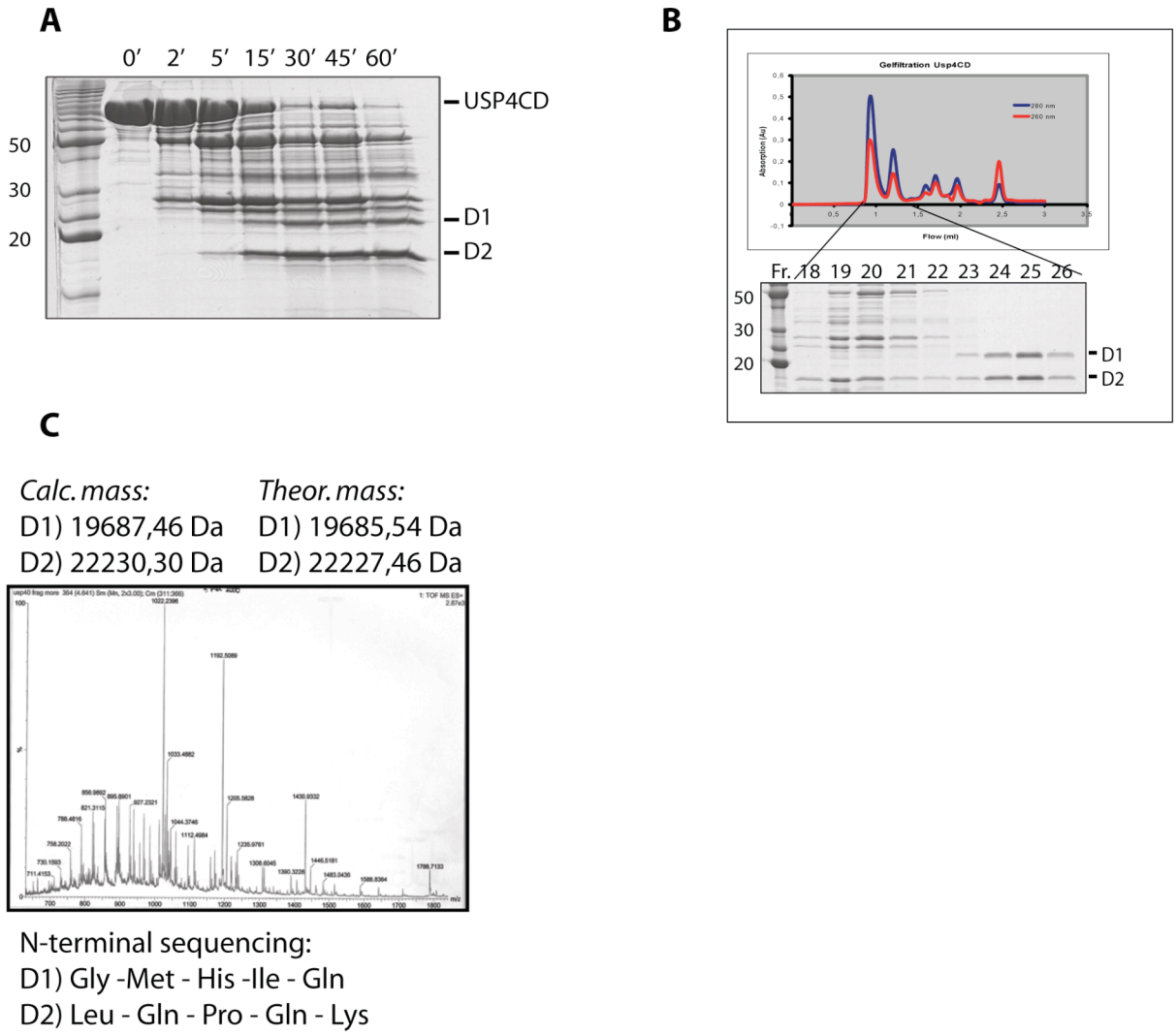
Data collection statistics	**Native**
Wavelength (Å)	0.993
Space group	$P2_12_12_1$
Unit cell (Å)	110.5, 151.0, 178.7
Molecules per asymmetric unit	6
Resolution (Å)[a]	44.6 - 2.4 (2.53 - 2.4)
R_{merge} (%)	8.8 (66.1)
$\langle I/\sigma(I) \rangle$	11.0 (1.1)
Completeness (%)	94.9 (74.0)
Redundancy	3.4 (2.4)
Refinement	
R_{work}/R_{free} (%)	17.8/21.0
Number of reflections	111095
Number of protein atoms	15884
Number of zinc ions	6
Number of waters	1249
RMSD from ideal geometry	
Bond lengths (Å)	0.010
Bond angles (°)	1.02
Ramachandran statistic[b] (prefered/allowed/outliers)	1923 / 34 / 0

[a] Numbers in parentheses are for the highest-resolution shell

[b] Calculated using Molprobity

Supplemental Figures

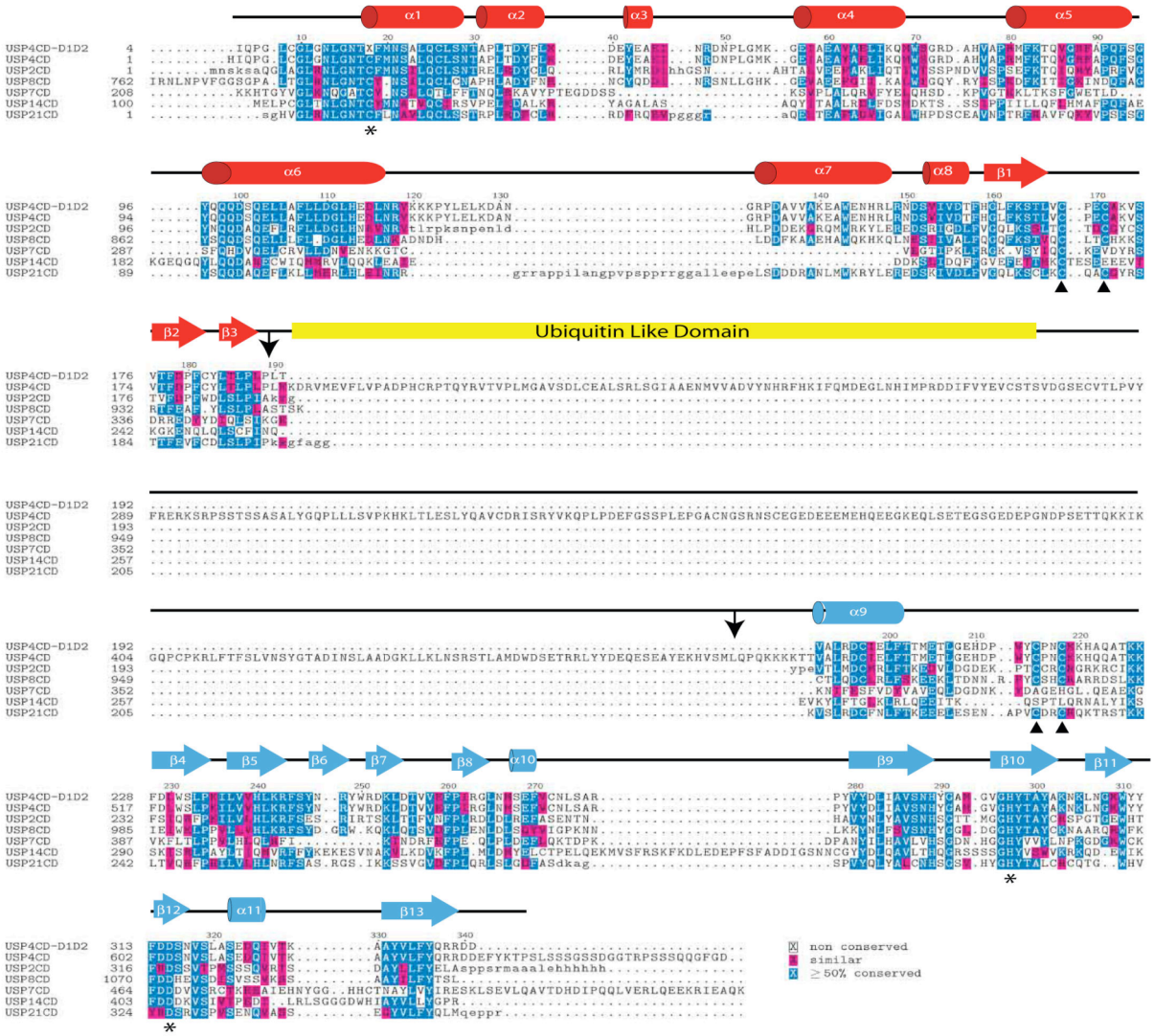
Luna-Vargas et al, Supplemental Figure S1



Supplemental Figure S1

Identification of USP4-D1D2. (**A**) Limited proteolysis analysis of thermolysin cleavage on USP4CD at 37°C. Samples at different time-points were taken and analyzed on a SDS-PAGE gel. (**B**) Proteolytic sample of USP4CD was subjected to a size exclusion chromatography and fraction samples were analyzed on a SDS-PAGE gel. (**C**) Mass Spectrometry analysis and N-terminal sequencing determined the identity of the two fragments, D1 and D2.

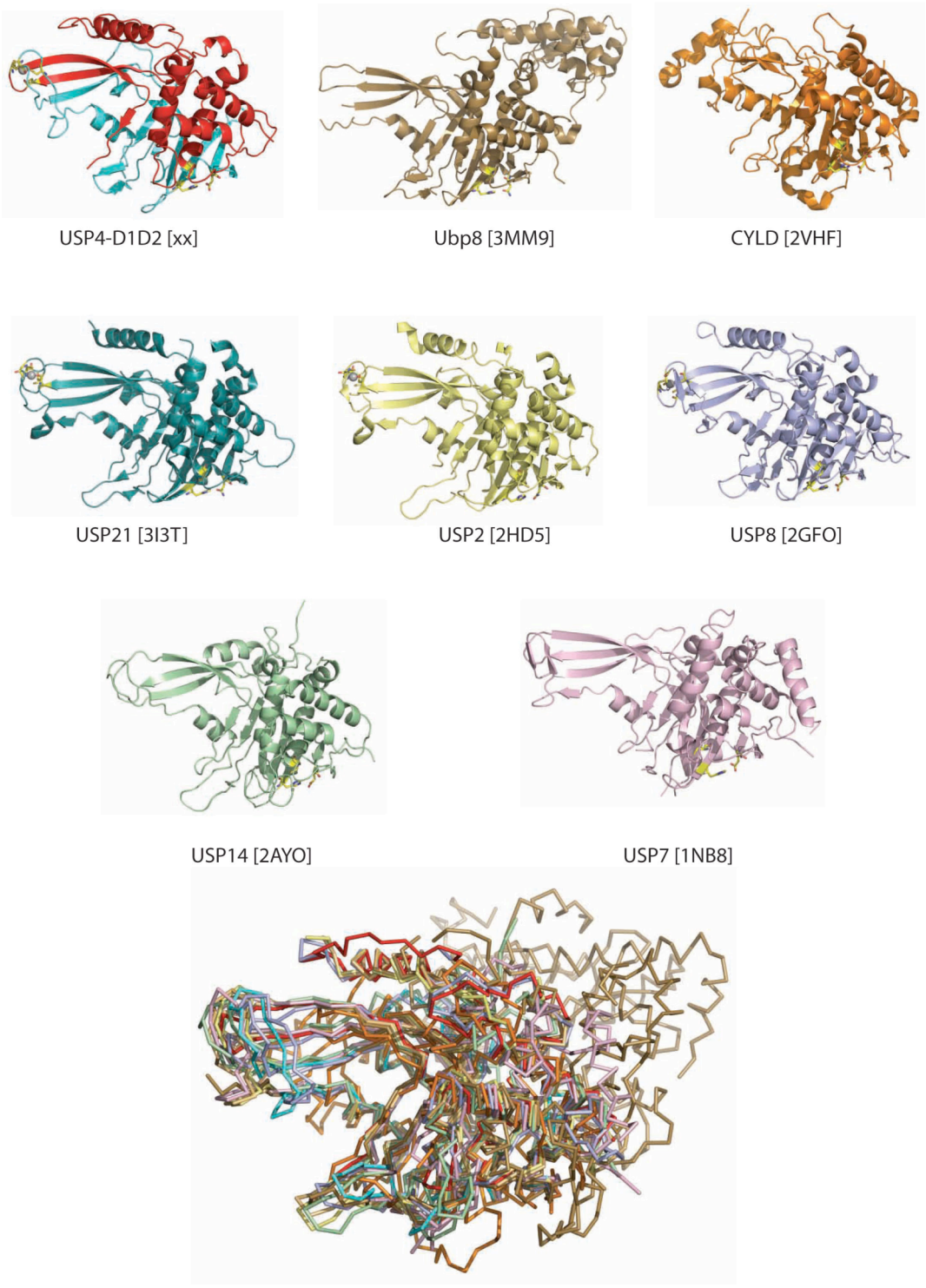
Luna-Vargas et al, Supplemental Figure S2



Supplemental Figure S2

Structure based multiple sequence alignment. Secondary structure elements are colored and labeled according to structure of USP4-D1D2 in Figure 1. The internal Ubl is depicted as a yellow bar and the two black arrows indicate where the protease thermolysin cleaved in the catalytic domain of USP4. The catalytic triad residues are indicated with an asterisk. The four black triangles indicate the positions of the Cys residues coordinating the zinc ion.

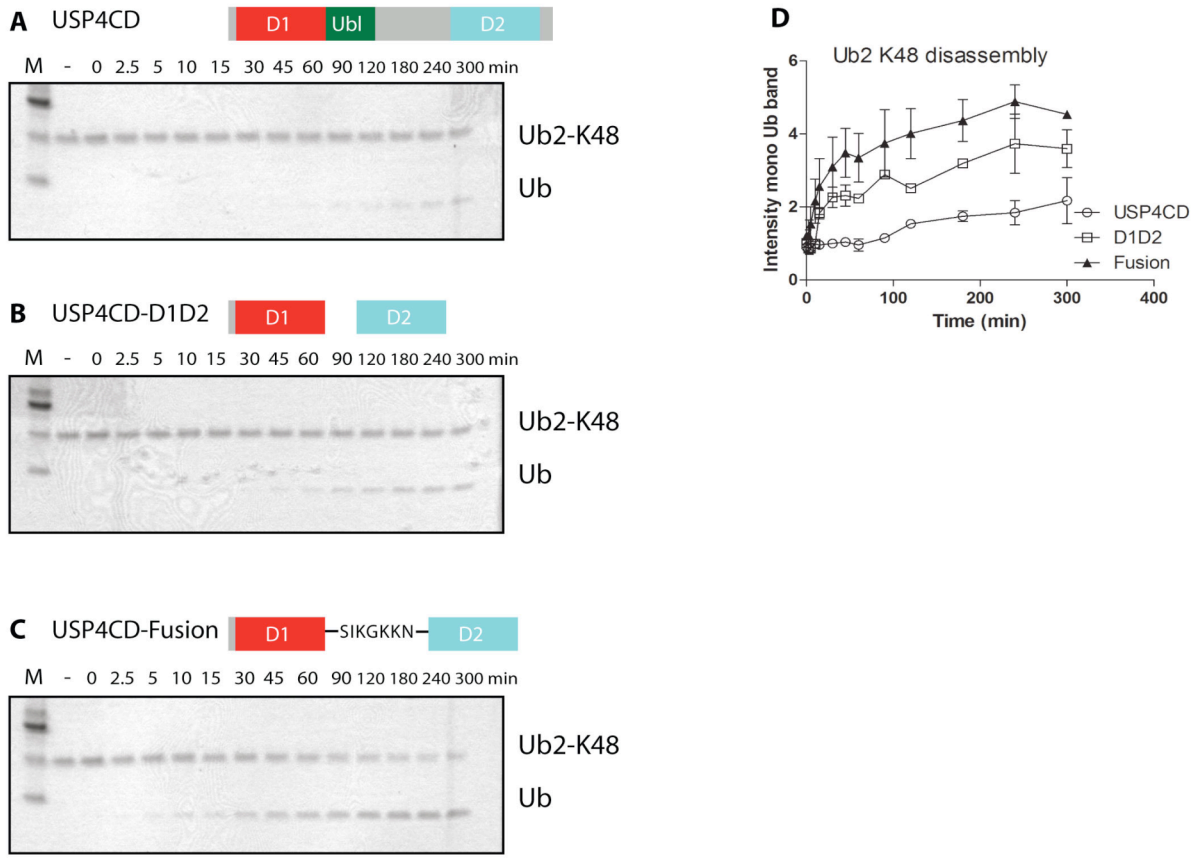
Luna-Vargas et al, Supplemental Figure S3



Supplemental Figure S3

Overview and superposition of USP catalytic domain structures with USP4-D1D2. Comparison between catalytic domains depicted in cartoon representation of USP4-D1D2 (red-cyan), Ubp8 (brown, PDB: 3MM9), CYLD (orange, PDB: 2VHF), USP21 (marine blue, PDB: 3I3T), USP2 (yellow, PDB: 2HD5), USP8 (purple, PDB:2GFO), USP14 (green, PDB: 2AYO) and USP7 (light pink, PDB:1NB8). The structures depicted in ribbon representation were superposed in Coot (RMSD of 2.1Å over 323 residues).

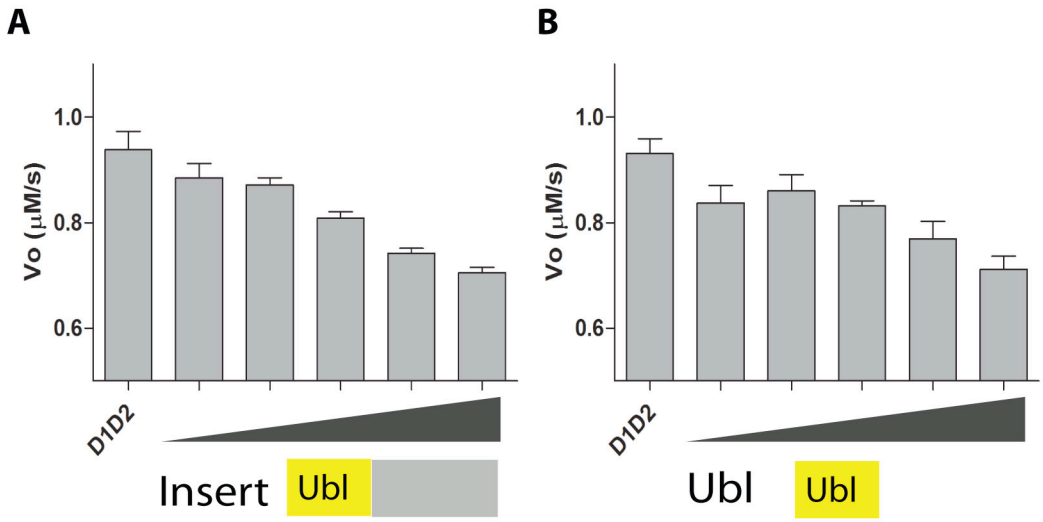
Luna-Vargas et al, Supplemental Figure S4



Supplemental Figure S4

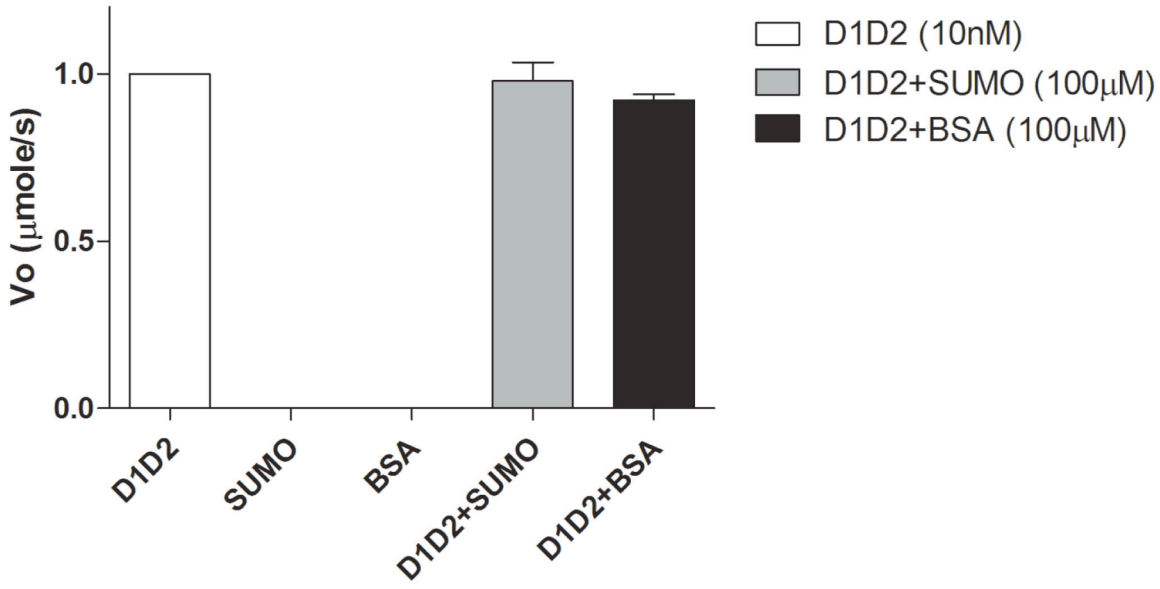
Deubiquitinating assay with K48 di-Ub as substrate (A-C) The full-length USP4 catalytic domain (**A**) is much less active than USP4-D1D2 (**B**) or USP4-fusion **(C)** in a deubiquitinating assay using K48 di-Ub as substrate on coomassie-stained SDS-PAGE gels. (**D**) Quantification of mono-Ub in K48 di-Ub cleavage assays. The intensity of the mono Ub band is plotted against time.

Luna-Vargas et al, Supplemental Figure S5



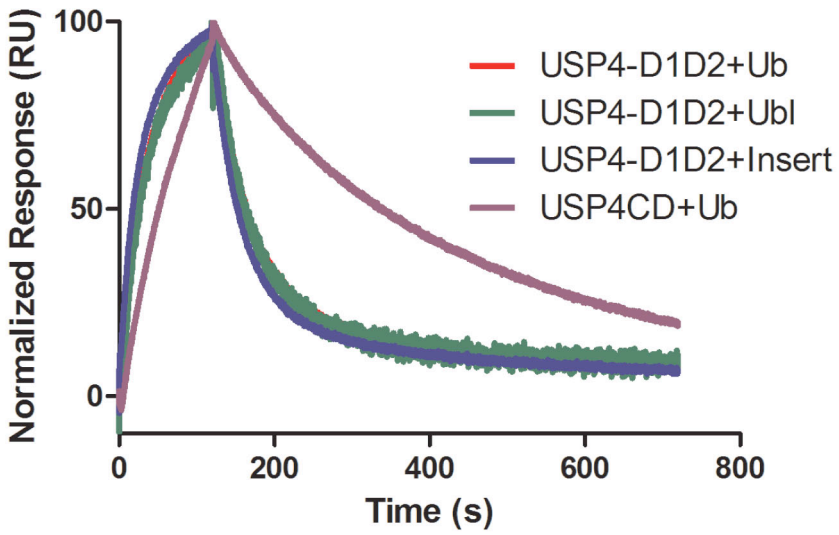
Supplemental Figure S5
***In trans* inhibition of USP4-D1D2 DUB activity** (**A**)The inhibitory effect of the insert is observed in a Ub-AMC assay with increasing amounts of insert (5, 10, 25, 50 and 75μM). **(B)** The Ubl domain (5, 10, 25, 50 and 100μM) is sufficient to show this *in trans* inhibition.

Luna-Vargas et al, Supplemental Figure S6



Supplemental Figure S6
The molecular crowding of high concentrations of SUMO or BSA (100μM) does not have an effect on USP4-D1D2 DUB activity.

Luna-Vargas et al, Supplemental Figure S7



Supplemental Figure S7
Kinetic comparison of USP4-D1D2 and USP4CD binding to Ub, Ubl and insert on SPR. Binding curve of 0.8 μM of USP4-D1D2 and of 1 μM of USP4CD were normalized for maximum binding in order to compare off-rates.