Supplementary Information for

Crystal structure and activity-based labeling reveal the mechanisms for linkage-specific substrate recognition by deubiquitinase USP9X

Prajwal Paudel[1]*, Qi Zhang[2]*, Charles Leung[2]*, Harrison C. Greenberg[1], Yusong Guo[3], Yi-Hsuan Chern[4], Aiping Dong[2], Yanjun Li[2], Masoud Vedadi[2,4], Zhihao Zhuang[1,$], Yufeng Tong[2,4,5$]

[1] Department of Chemistry and Biochemistry, University of Delaware, 214A Drake Hall, Newark, DE 19716, USA
[2] Structural Genomics Consortium, 101 College Street, MaRS Centre, South Tower, Toronto, ON M5G 1L7, Canada
[3] Fisheries College, Guangdong Ocean University, Zhanjiang, Guangdong 524025, China
[4] Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON, Canada
[5] Department of Chemistry and Biochemistry, University of Windsor, Windsor, ON, Canada
[*] Equal contribution
[$] Corresponding authors: ytong@uwindsor.ca, zzhuang@udel.edu

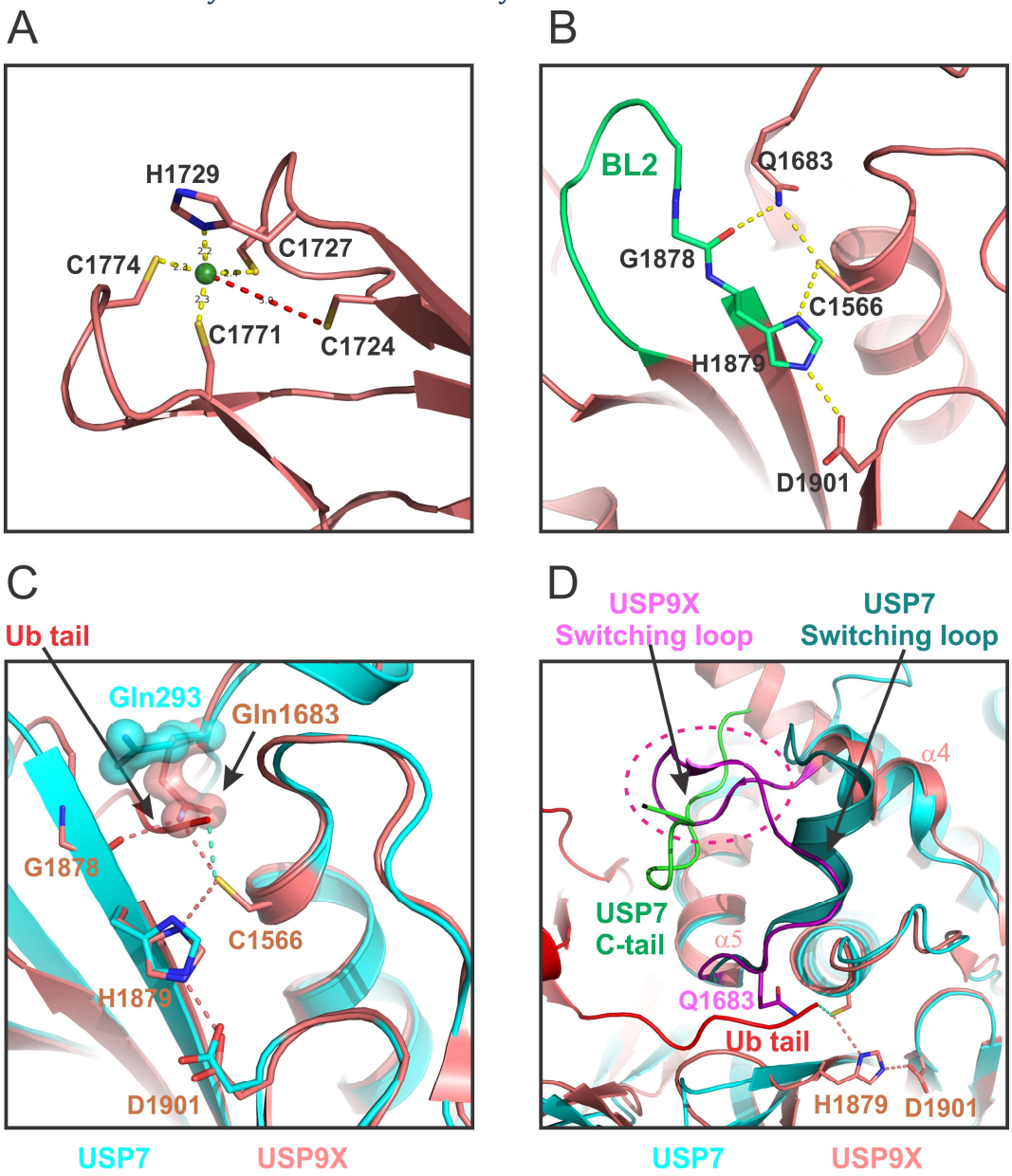
Corresponding authors:
Yufeng Tong
E-mail: ytong@uwindsor.ca

Zhihao Zhuang
E-mail: zzhuang@udel.edu

This PDF file includes:

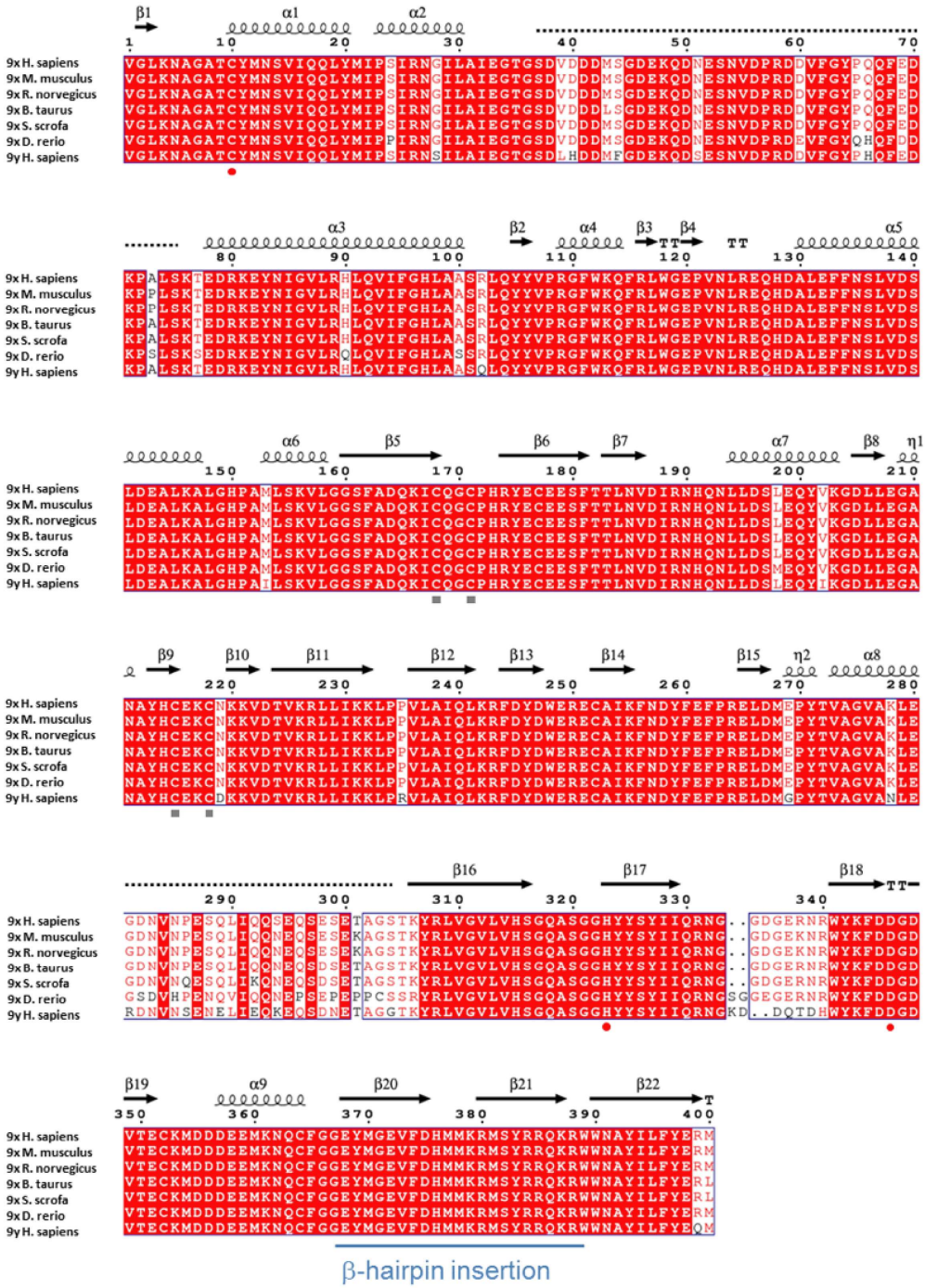
Figs: S1 to S16
Tables S1 to S2
SI reference citations

Fig. S1. Structural analysis of the USP9X catalytic domain.



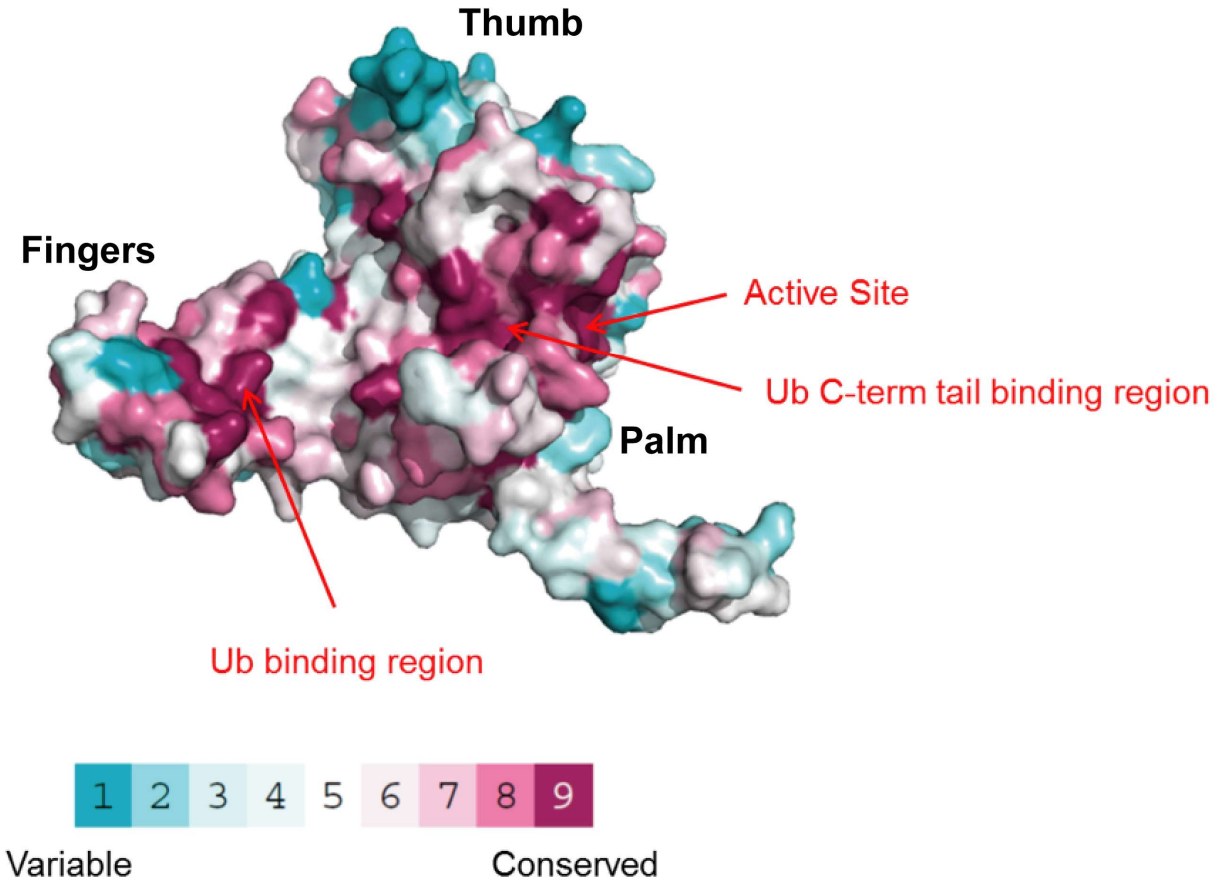
(A) A zinc ion in the tip of fingers subdomain of USP9X is coordinated with Cys1727, His1729, Cys1771, and Cys1774, but not with C1724, as seen in other USPs. **(B)** The USP9X active site catalytic triad with occlusion of the active site by Gln1683 hydrogen bonded to the backbone of blocking loop 2 (BL2, green) and the Cys1566. **(C)** USP9X Gln1683 equivalent Gln293 in USP7 rotates to accommodate the Ub C-terminal tail (red) into the catalytic groove. **(D)** A switching loop (purple) of USP9X between helices α4 and α5 containing Gln1683 is two residues longer compared with the equivalent switching loop of USP7 (cyan for Ub-bound USP7). The USP9X switching loop protrudes into a space otherwise occupied by the C-terminal tail (green) of USP7, suggesting USP9X does not have a similar regulatory mechanism as that of USP7.

Fig. S2. Sequence alignment of USP9X from different species and human USP9Y.



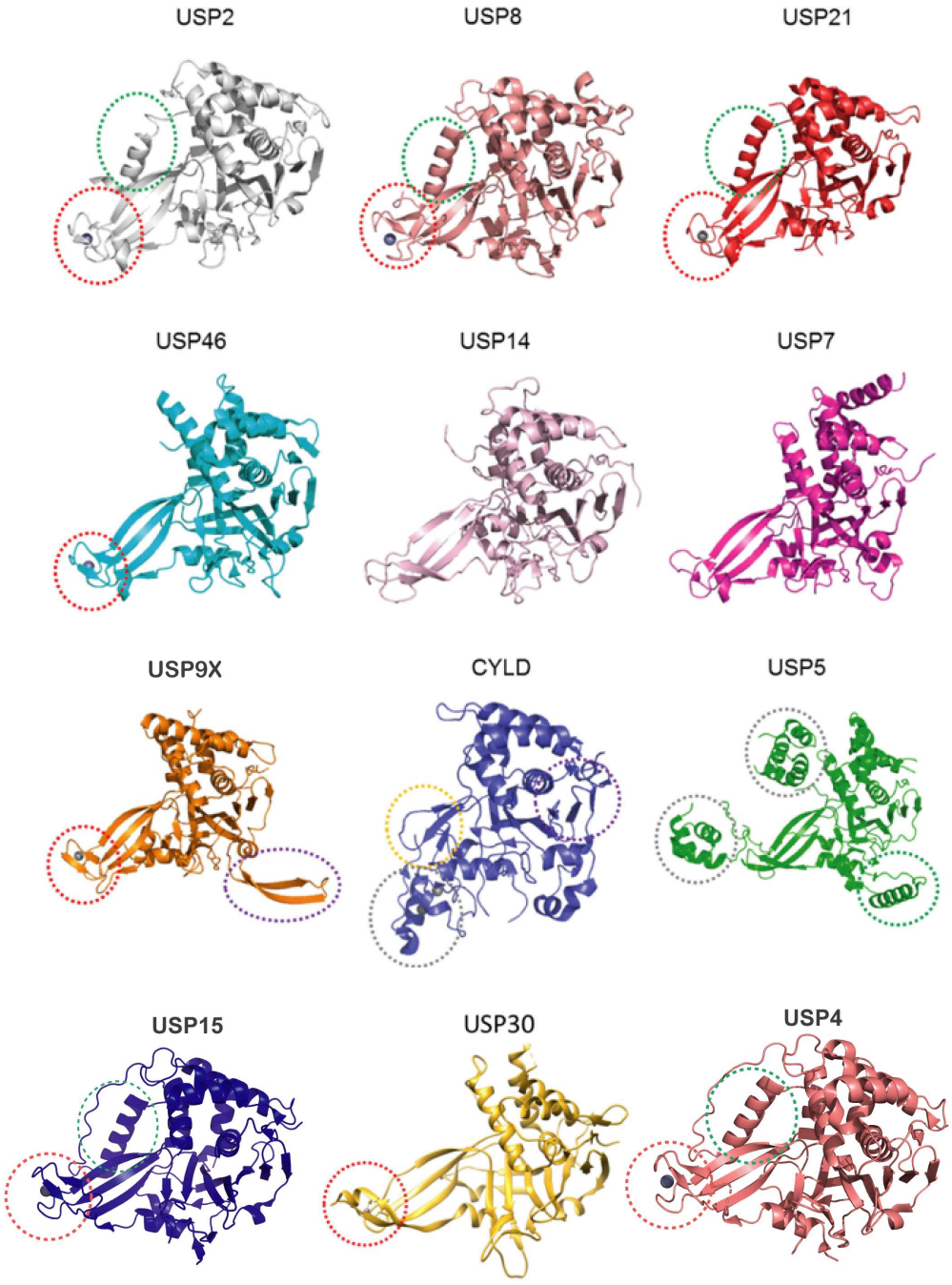
The secondary structure of USP9X is shown on the top of the line. The catalytic triad is indicated with red circles and zinc finger cysteines with grey boxes. The two large disordered insertions and β-hairpin insertion are shown as dotted lines and underlined, respectively. Sequence alignment was generated using Clustal Omega(1) and ESPript 3.0(2).

Fig. S3. Conservation of USP9X CD residues.



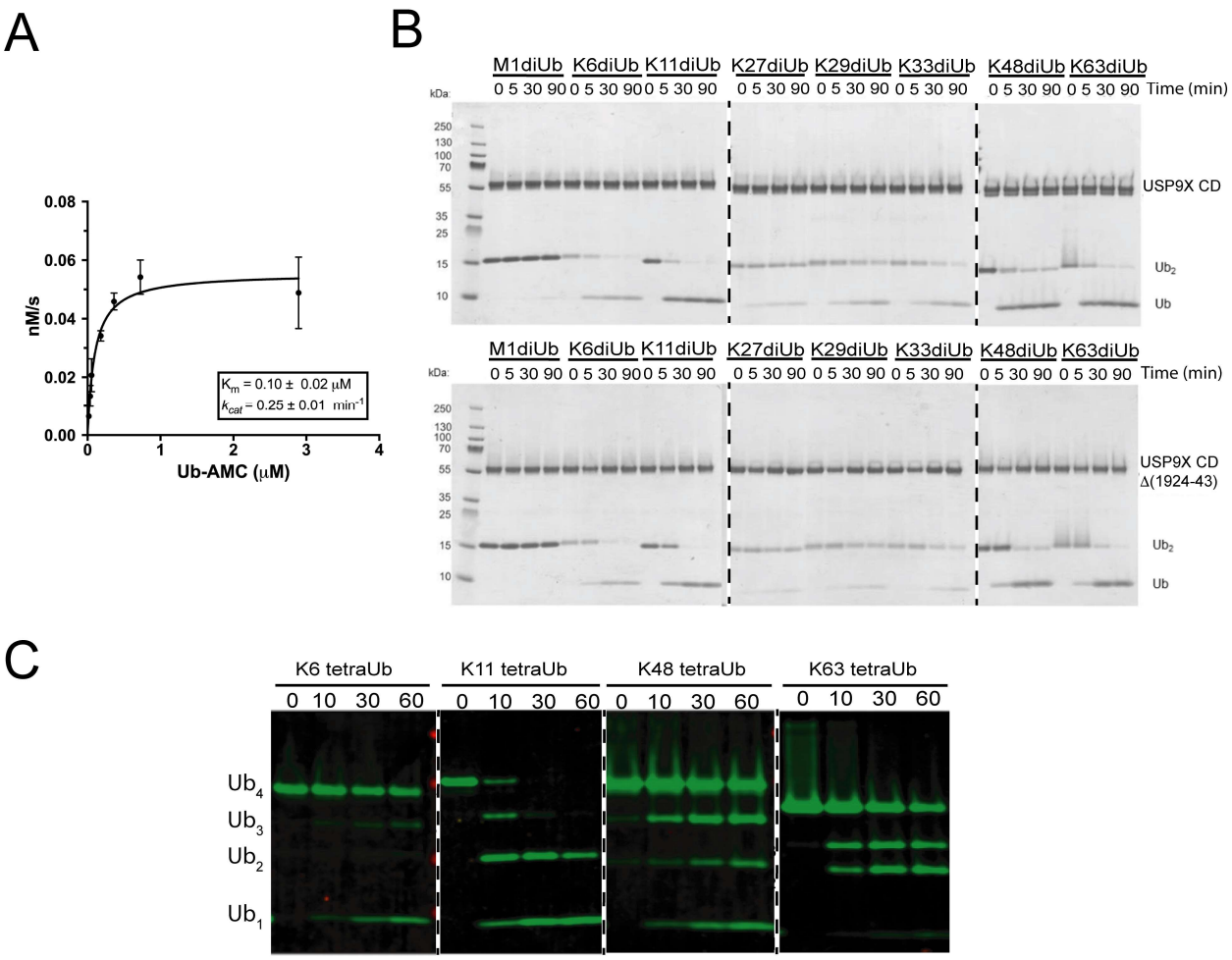
Conservation of USP9X residues was calculated using the ConSurf server(3). The most conserved regions are indicated by arrows.

Fig. S4. Comparison of the structural features of different USP CDs.



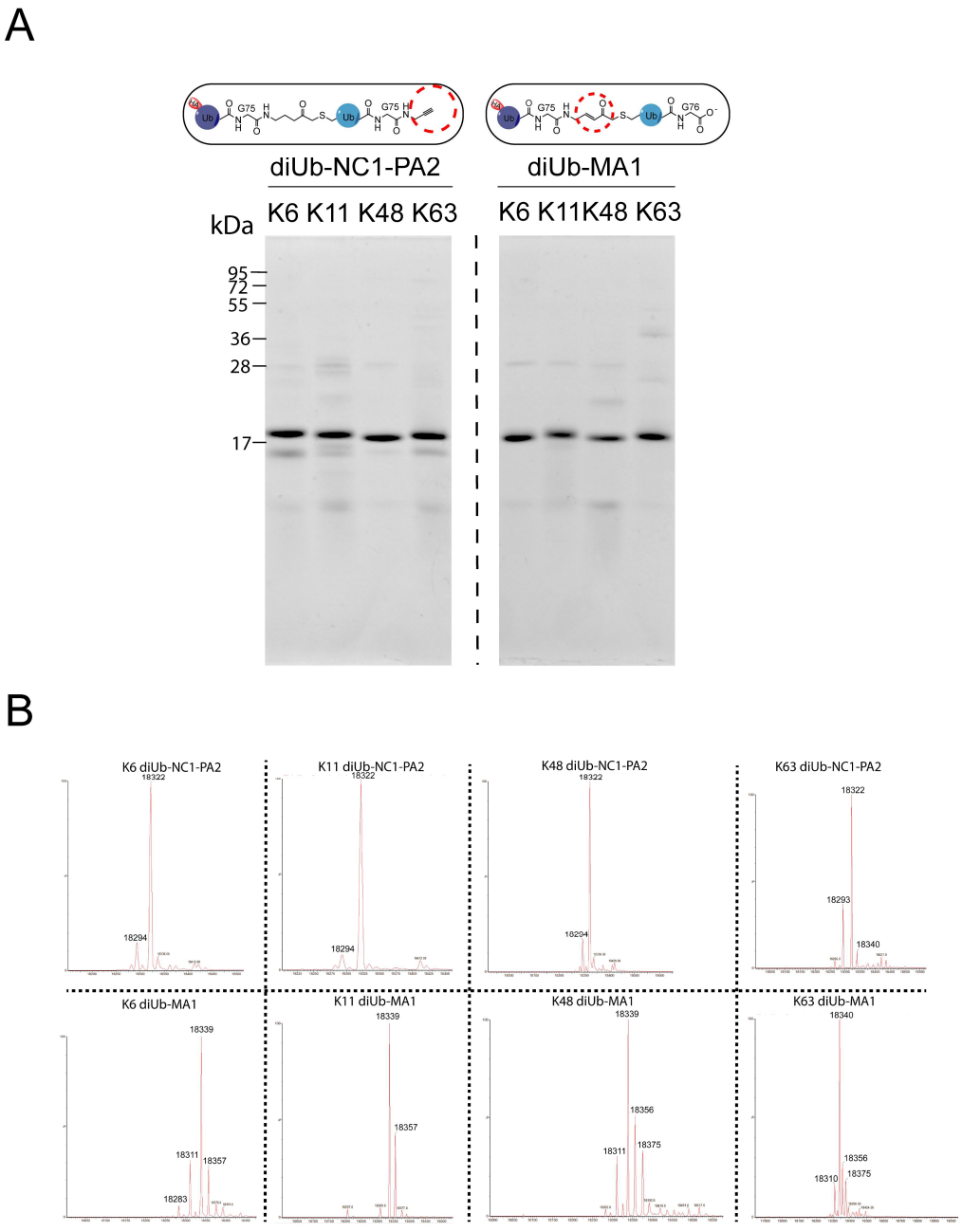
Features are circled for zinc finger (red), truncated fingers subdomain (yellow), β-hairpin insertion (purple), helical insertion (green) and large domain insertion (grey). PDB coordinates used are 2HD5 (USP2), 2GFO (USP8), 3I3T (USP21), 5CVM (USP46), 2AYO (USP14), 1NBF (USP7), 2VFH (CYLD), 3IHP (USP5), 6GHA (USP15), 5OHP (USP30), 2Y6E (USP4). Only the CD structures are shown.

Fig. S5. Activity assays of WT USP9X CD and mutants using Ub-AMC, diUb and tetraUb substrates.



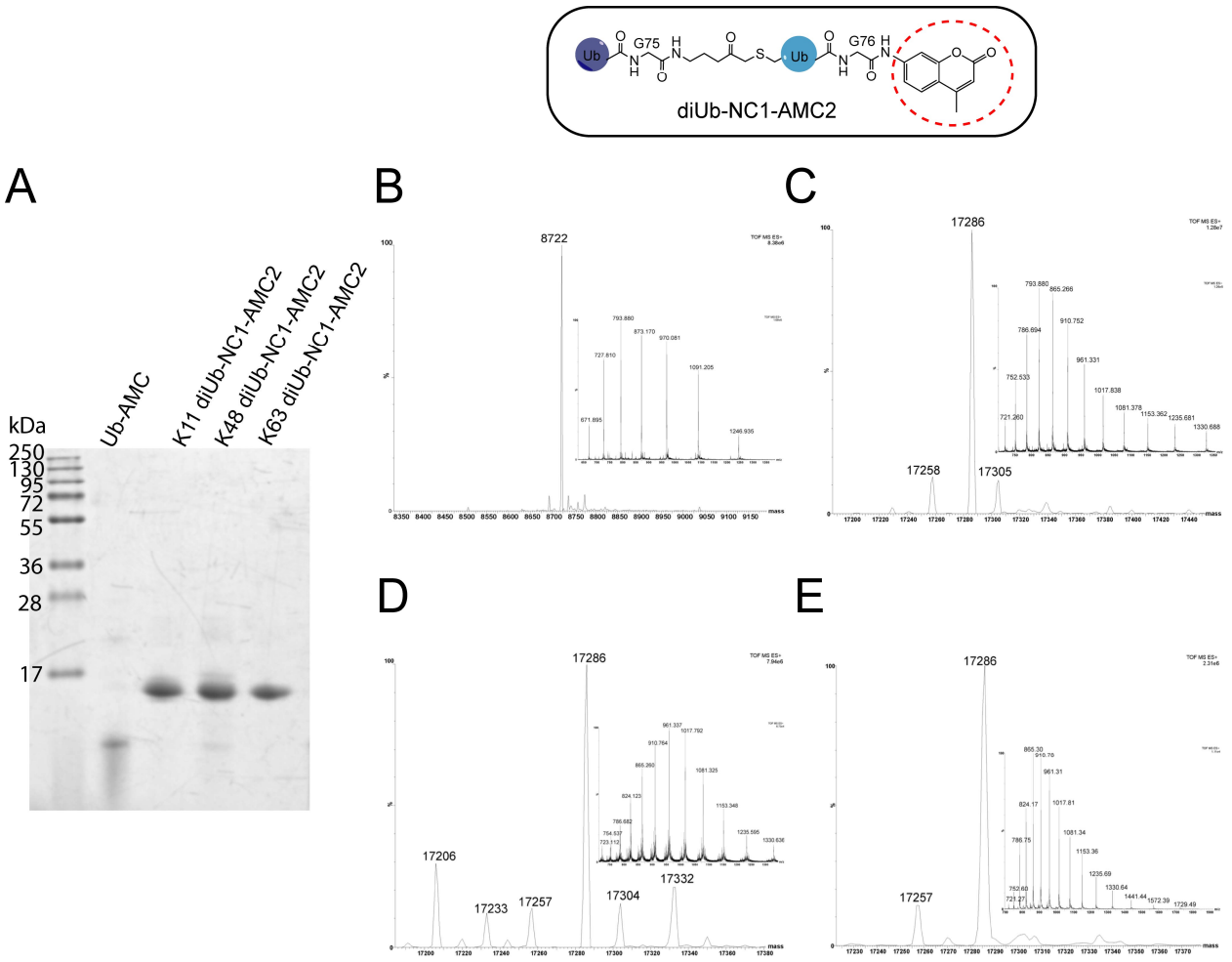
(A) Steady-state kinetics analysis of WT USP9X CD using Ub-AMC as a substrate. Kinetic parameters were obtained by non-linear regression fitting to the Michaelis-Menten equation. Data points are represented as mean ± SD (n = 3). **(B)** Full-size gels of the cleavage of native diUb substrates of different linkages by WT USP9X CD (cf. Fig. 1*E*) and β-hairpin deletion mutant Δ(1924–1943). **(C)** Cleavage of native tetraubiquitin of K6-, K11-, K48- and K63- linkages by WT USP9X CD.

Fig. S6. Generation of linkage-specific diUb probes with terminally (diUb-NC1-PA2) or internally (diUb-MA1) placed warhead.



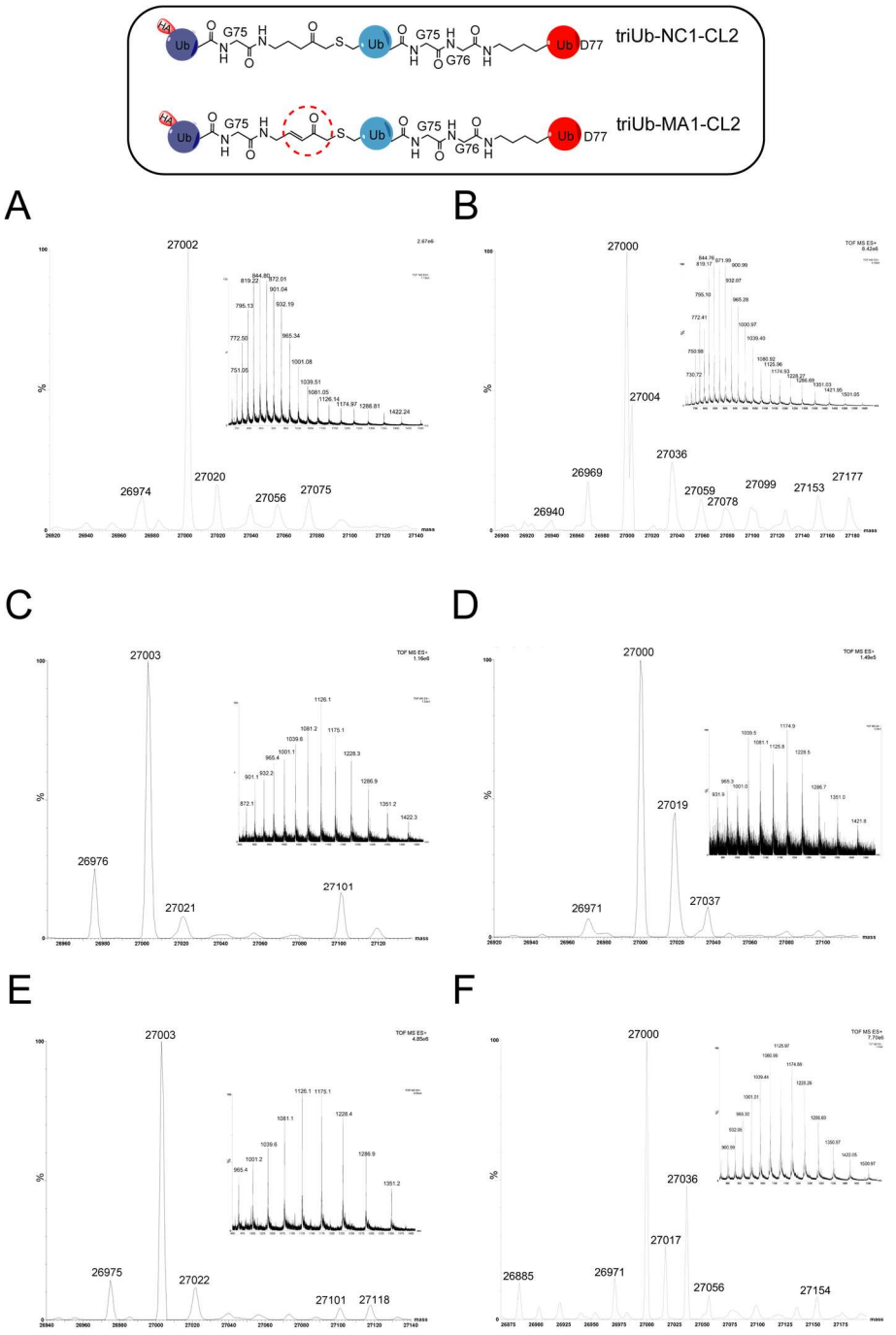
(**A**) SDS-PAGE gel analysis of K6, K11, K48, and K63 diUb-NC1-PA2 probes and K6, K11, K48 and K63 diUb-MA1 probes. (**B**) Xevo Q-TOF MS characterization of the above listed diUb probes. Expected mass of diUb-MA1 is 18339 kDa and of diUb-NC1-PA2 is 18322 kDa.

Fig. S7. Characterization of diUb-NC1-AMC2 DUB substrates.



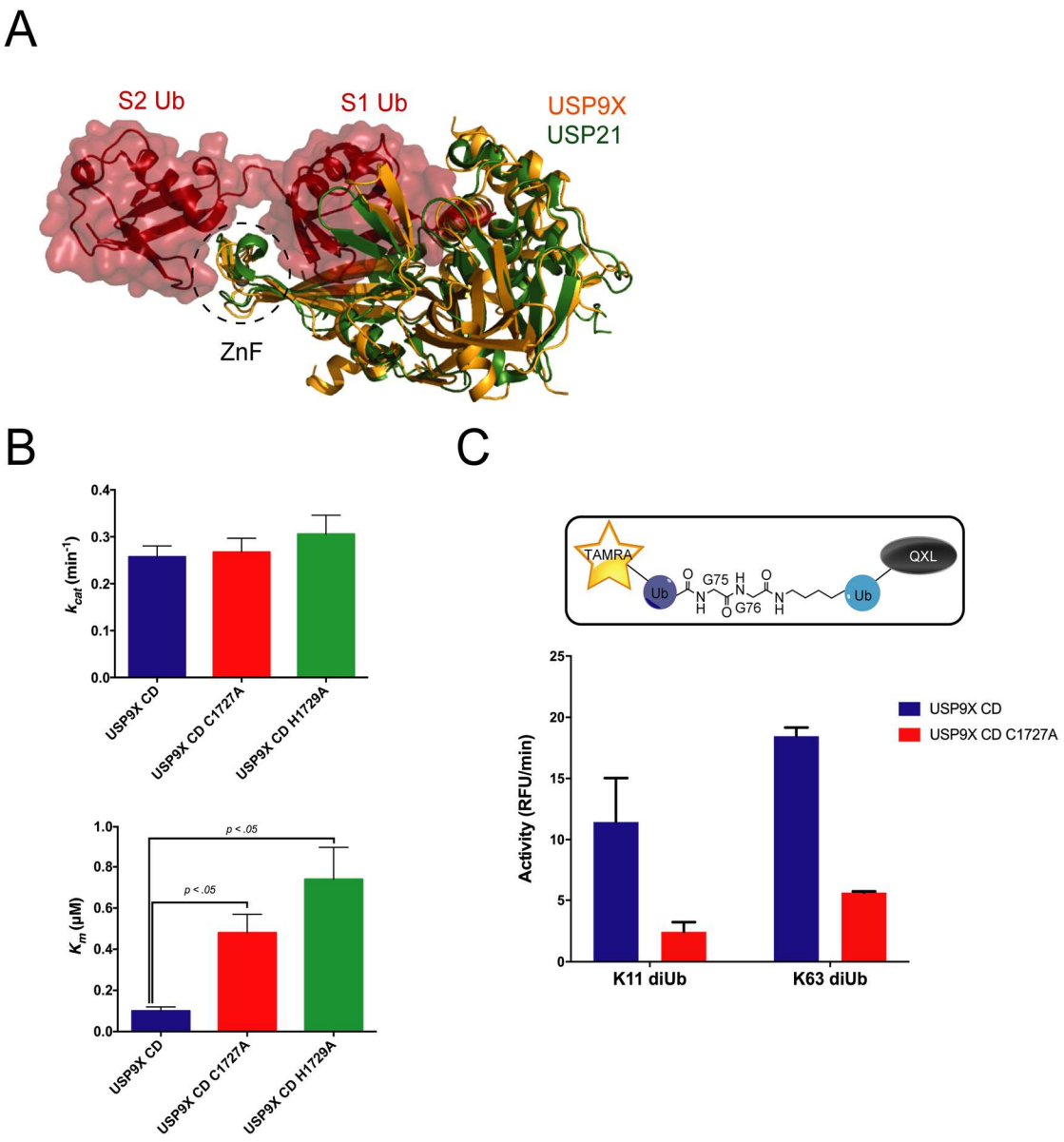
(**A**) The SDS-PAGE analysis of K11, K48, and K63 diUb-NC1-AMC2 species. (**B–E**) Xevo Q-TOF mass spectrometry analysis of Ub-AMC (**B**), K11 diUb-NC1-AMC2 (**C**), K48 diUb-NC1-AMC2 (**D**) and K63 diUb-NC1-AMC2 (**E**). Expected mass for diUb-NC1-AMC2 is 17285 kDa.

Fig. S8. Xevo Q-TOF MS analysis of triUb-NC1-CL2 and triUb-MA1-CL2 probes of different linkages.



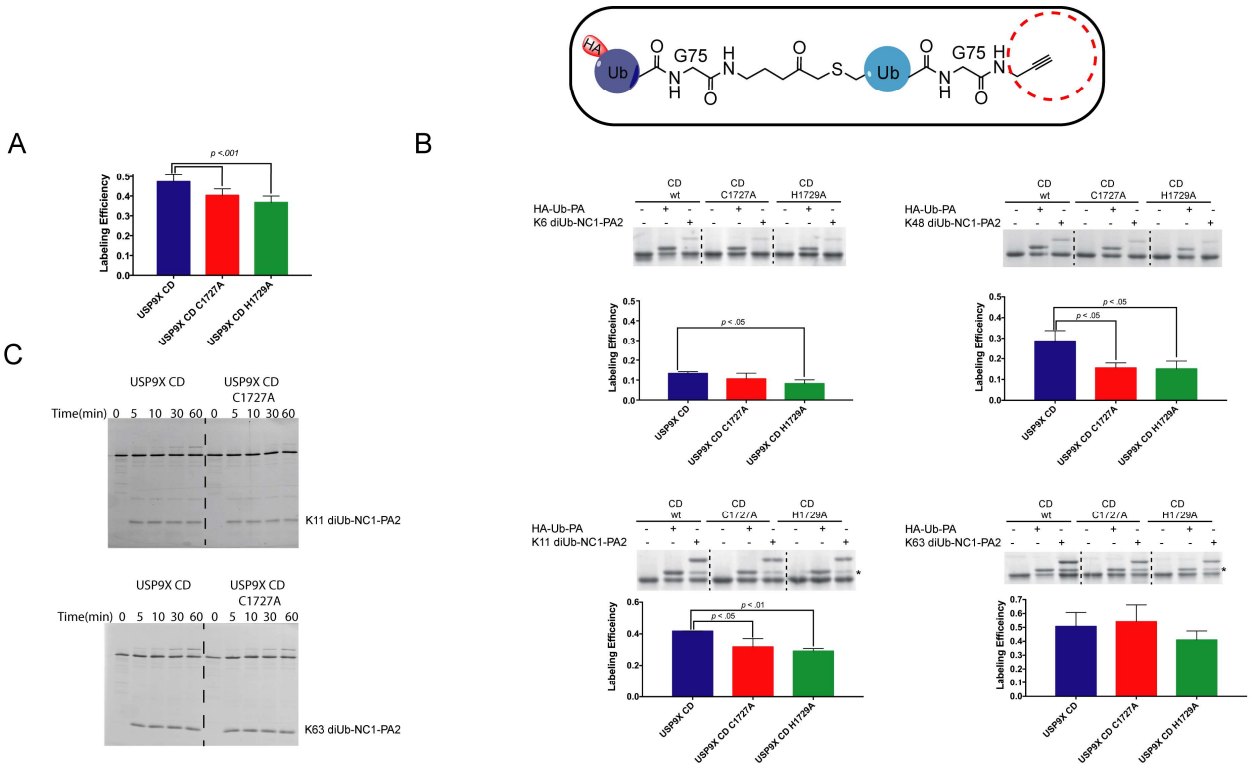
Deconvoluted and multiple-charged state spectra of **(A)** K11 triUb-NC1-CL2, **(B)** K11 triUb-MA1-CL2, **(C)** K48 triUb-NC1-CL2, **(D)** K48 triUb-MA1-CL2, **(E)** K63 triUb-NC1-CL2, and **(F)** K63 triUb-MA1-CL2. Expected mass for triUb-MA1-CL2 is 27001 kDa, and triUb-NC1-CL2 is 27003 kDa.

Fig. S9. Analysis of the ZnF region of USP9X.



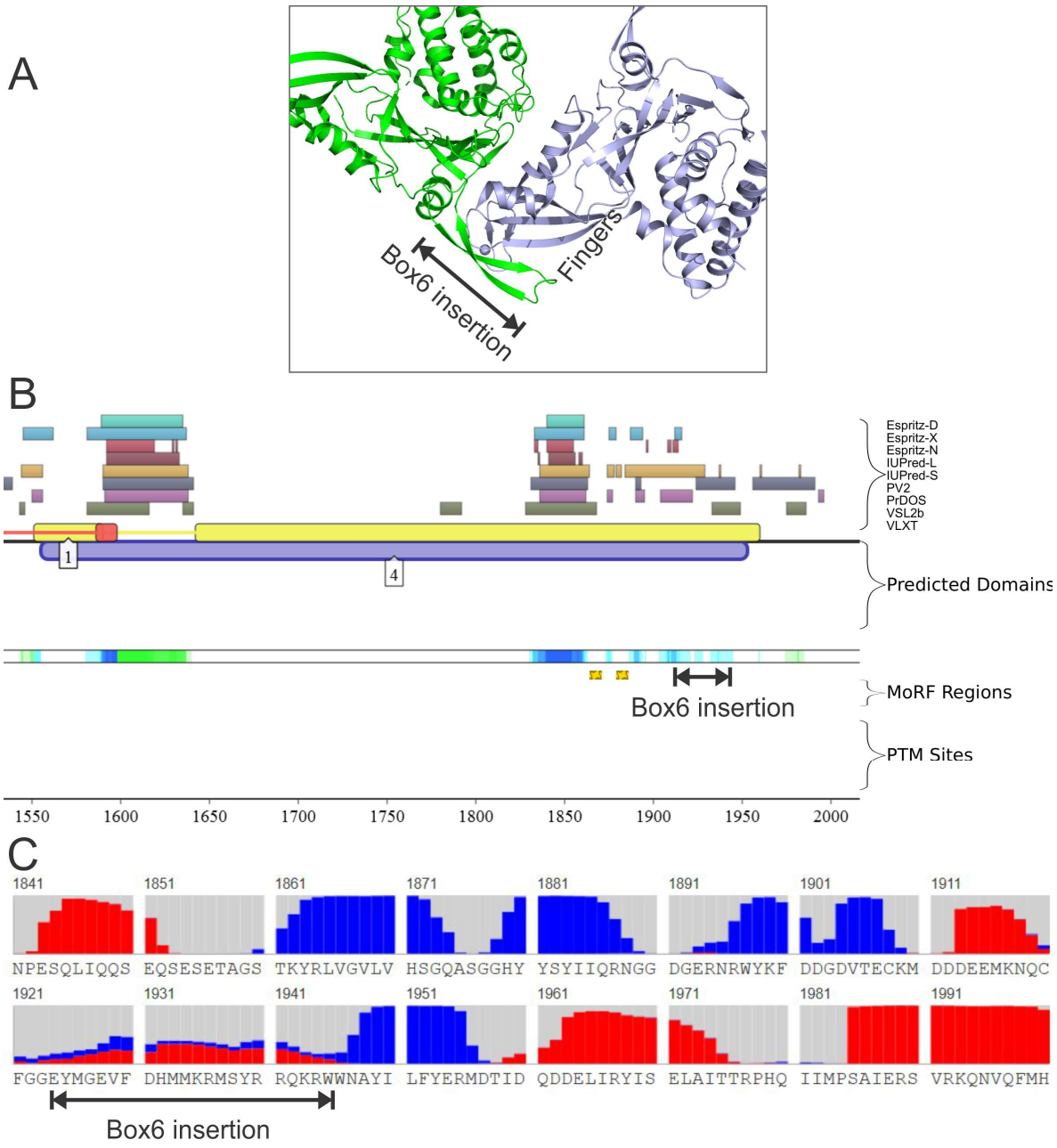
(A) Superposition of the USP9X CD apo structure with the USP21/Met1-diUb co-crystal structure (PDB:2Y5B) suggests a possible S2 site near the USP9X zinc finger (ZnF). **(B)** Kinetic parameters of WT and mutant USP9X CD hydrolyzing Ub-AMC substrate and statistical analysis comparing the K_m and k_{cat} of WT and mutant USP9X CD hydrolyzing Ub-AMC. **(C)** Internally quenched fluorescence (IQF)-diUb substrate cleavage analysis of ZnF mutant C1727A USP9X CD at 50 nM enzyme and 500 nM K11 diUb and 200 nM K63 diUb IQF substrates.

Fig. S10. Labeling of WT and mutant USP9X CD by Ub-PA and diUb-NC1-PA2 probes.



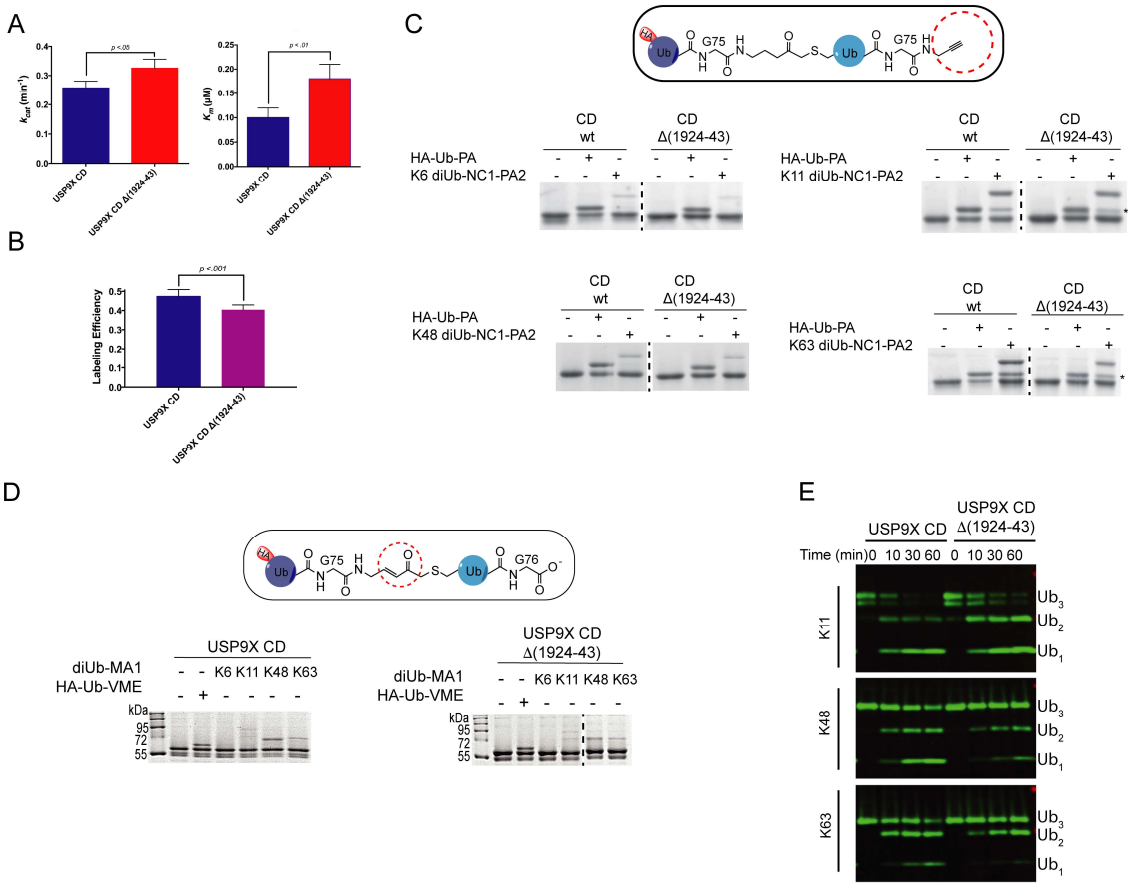
(A) Comparison of the labeling efficiency of WT USP9X CD and the ZnF mutants (C1727A and H1729A) by the HA-tagged Ub-PA probe (n = 12), which has the same warhead as that of diUb-NC1-PA2. **(B)** Labeling of WT and ZnF mutant USP9X CD by K6-, K48-, K11-, and K63-linked diUb-NC1-PA2 probes (n=3), respectively. * indicates labeling band caused by a residual HA-Ub-PA impurity in the DiUb probes. **(C)** Time course of labeling of WT and C1727A USP9X CD by K11- (top panel) and K63- (bottom panel) diUb-NC1-PA2 probes.

Fig. S11. Analysis of the β-hairpin insertion in sequence box 6.



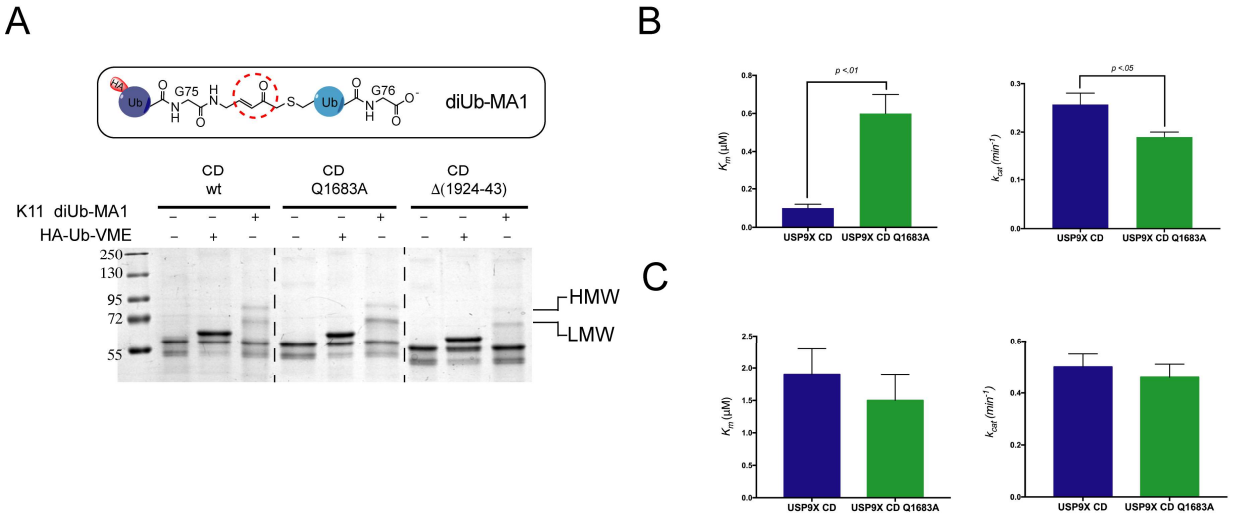
(A) The insertion in USP9X CD box 6 adopts a β-hairpin structure (green) and pairs with the β-strands in the fingers subdomain of a symmetry-related USP9X molecule (light purple) in the crystal lattice to form an extended β-sheet. The Zn ion in the fingers subdomain is shown as a sphere. (**B**) Meta-analysis of the propensity of the box 6 insertion to be disordered in the Database of Disordered Protein Prediction (D2P2). The box 6 insertion has low propensity to be disordered. (**C**) Secondary structure prediction using RaptorX suggests the box 6 insertion has low propensity to be α-helix (red) or β-strand (blue).

Fig. S12. Probing the USP9X CD β-hairpin using diUb probes.



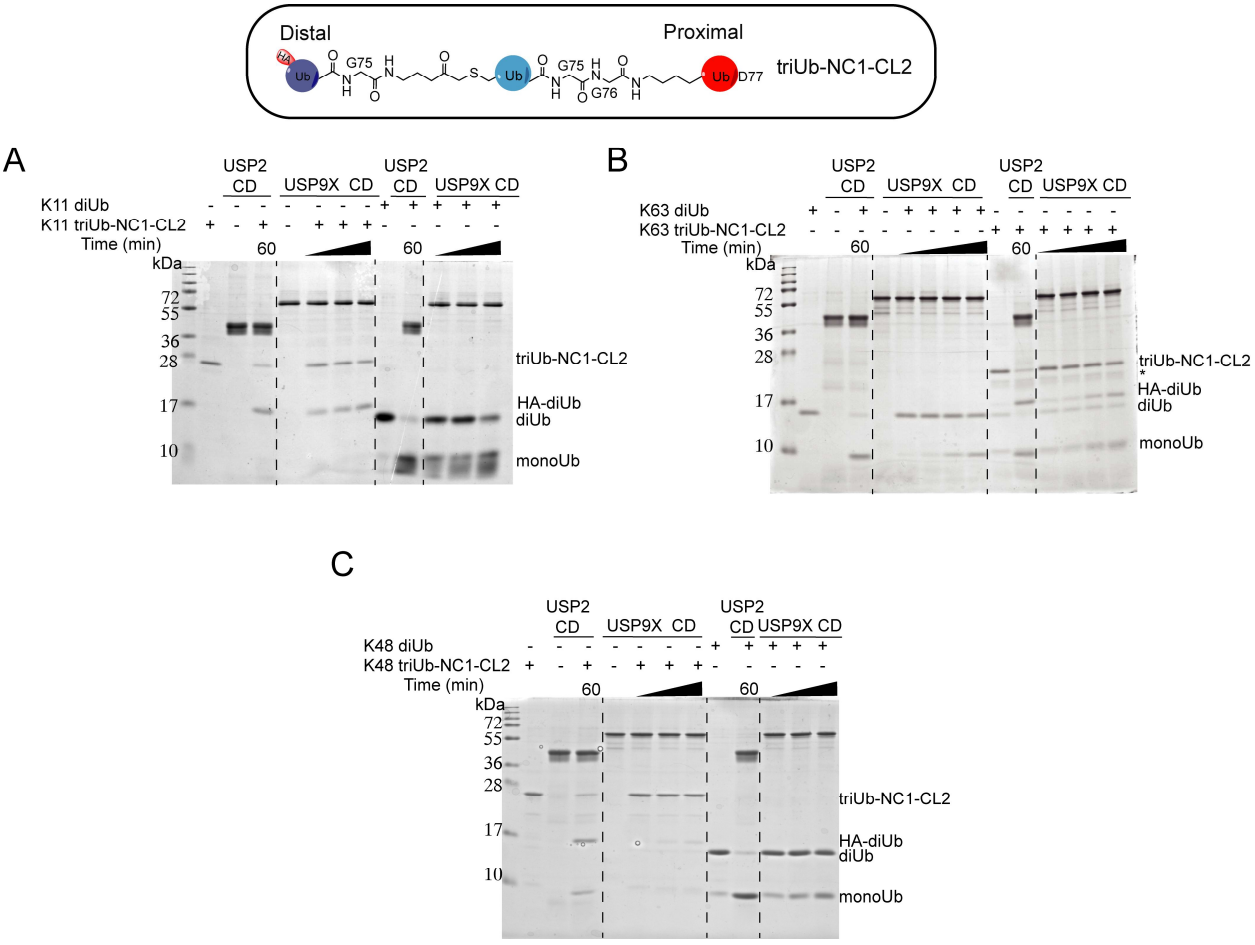
(A) Kinetics of USP9X WT and Δ(1924–1943) against Ub-AMC substrate (n = 3). **(B)** Labeling by the HA-tagged Ub-PA probe shows the effect of β-hairpin deletion (n = 12). **(C)** Labeling of WT and Δ(1924–1943) USP9X CD by K6-, K48-, K11-, and K63-diUb-NC1-PA2 probes. * indicates labeling bands caused by a residual Ub-PA impurity in the DiUb probes. **(D)** The labeling of WT and Δ(1924–1943) USP9X CD by HA-Ub-VME and diUb-MA1 probes of K6-, K11-, K48- and K63-linkages. **(E)** Gel-based triUb cleavage by WT and Δ(1924–1943) USP9X CD.

Fig. S13. Probe the effect of Gln1683 in USP9X CD catalysis and labeling by the K11 diUb-MA1 probe.



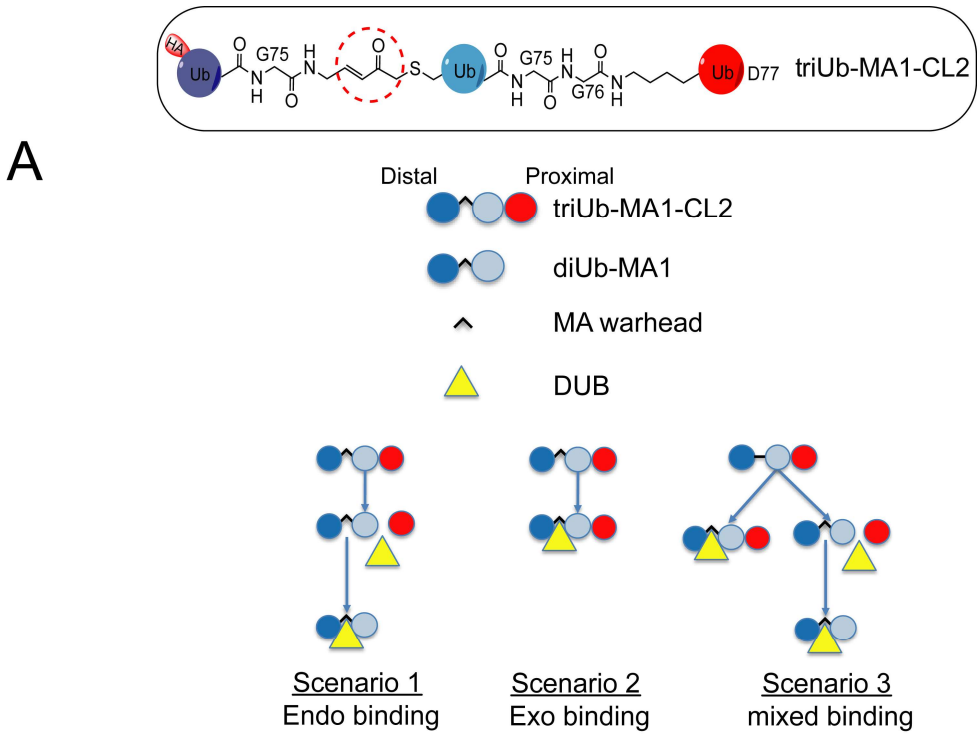
(A) Labeling pattern of WT, Q1683A, and Δ(1924-1943) USP9X CD by K11 diUb-MA1 at r.t. for 1 hr. **(B)** Compare K_m and k_{cat} values of WT and Q1683A USP9X CD using Ub-AMC as a substrate. **(C)** Compare K_m and k_{cat} values of WT and Q1683A USP9X CD using IQF-K11-diUb as a substrate.

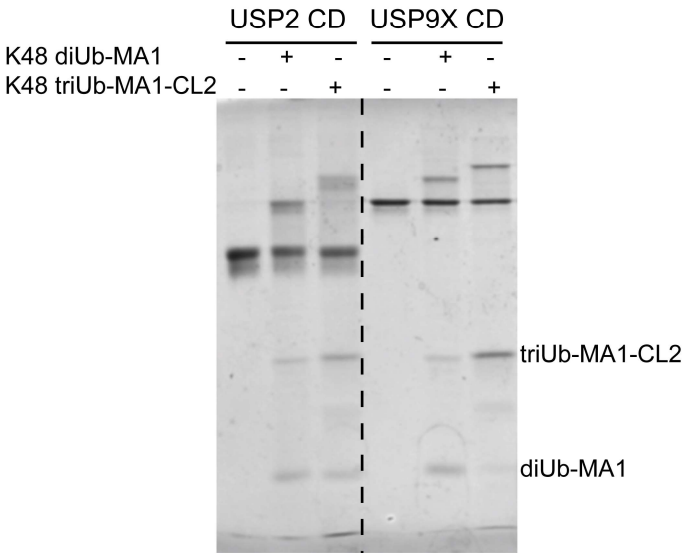
Fig. S14. Cleavage assay of WT USP9X CD using triUb-NC1-CL2.



Cleavage assay using triUb-NC1-CL2 of **(A)** K11-, **(B)** K63- and **(C)** K48-linkage. DiUb substrates of the corresponding linkage are included as a comparison. Time points for K63-linkage are 5, 10, 30, and 60 min while K11- and K48-linkages are 15, 30, 60 min. USP2 CD-catalyzed cleavage is also included in each assay to show that the di- and tri-Ub substrates can be cleaved. Cleavage of triUb-NC1-CL2 results in an HA-tagged diUb product which is 1 kDa larger than the native diUb as indicated on the gel.

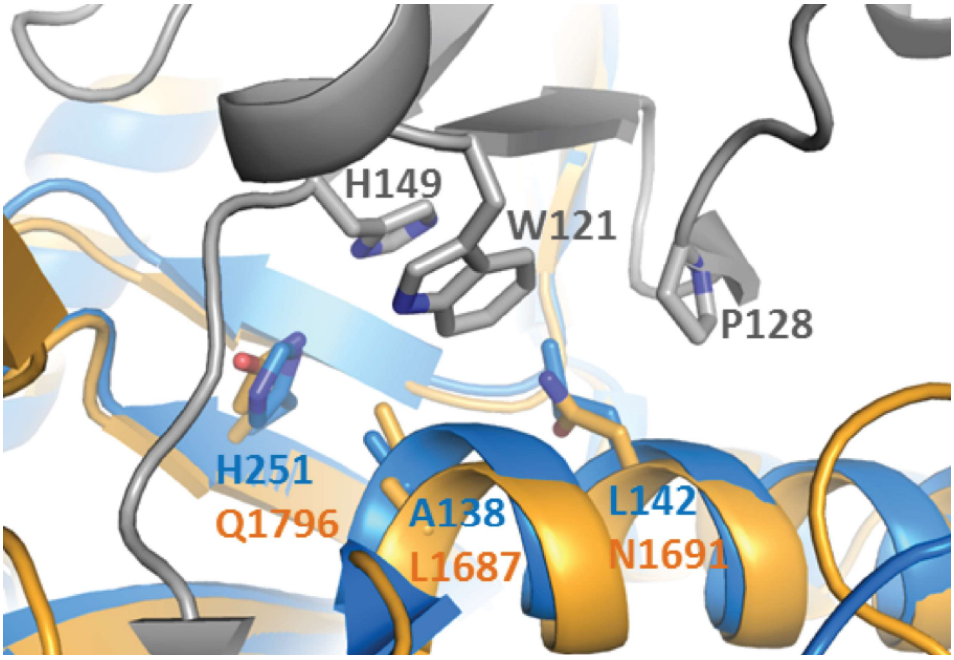
Fig. S15. Cleavage and labeling patterns using the hybrid triUb probe triUb-MA1-CL2.





(A) Three possible scenarios based on the binding mode of DUBs: 1, exclusive endo-cleavage produces mono- and di-Ub; 2, exclusive exo-cleavage leads to labeling of the DUB by the triUb probe; 3, simultaneous endo- and exo-cleavages produce monoUb as well as di- and tri-Ub labeling of the DUB. **(B)** K48-linked triUb-MA1-CL2 and K48-linked diUb-MA1 labeling of WT USP9X CD at r.t. for 1 hr with USP2 CD as a comparison.

Fig. S16. Superposition of the structure of USP9X CD and that of the USP18 CD in complex with ISG15.



USP9X, USP18 and ISG15 are coloured in yellow, blue and grey, respectively. Residues involved in interactions are shown as sticks and labelled.

Table S1. Kinetic parameters of WT USP9X CD and mutants using Ub-AMC

USP9X CD	k_{cat} (min^{-1})	K_m (μM)	*Specificity constant* (min^{-1} μM^{-1})
WT	0.25 ± 0.01	0.10 ± 0.02	2.5
C1727A	0.26 ± 0.02	0.47 ± 0.09	0.56
H1729A	0.31 ± 0.02	0.74 ± 0.16	0.41
Δ(1924–1943)	0.32 ± 0.02	0.18 ± 0.03	1.8
Q1683A	0.19 ± 0.01	0.64 ± 0.10	0.32

Table S2. Kinetic parameters of WT USP9X CD using diUb-NC1-AMC2

Linkage	k_{cat} (min^{-1})	K_m (μM)	*Specificity constant* (min^{-1} μM^{-1})
K11	0.35 ± 0.03	1.0 ± 0.3	0.34
K48	0.025 ± 0.003	0.6 ± 0.2	0.04
K63	0.42 ± 0.05	2.5 ± 0.6	0.17

SI Appendix References

1. Sievers F*, et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539.
2. Robert X & Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42(Web Server issue):W320-324.
3. Ashkenazy H*, et al.* (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* 44(W1):W344-350.