Supporting Information

Ubiquitin specific protease 11 structure in complex with an engineered substrate mimetic reveals a molecular feature for deubiquitination selectivity

Sigrun K. Maurer[1], Matthias P. Mayer[1], Stephanie J. Ward[1], Sana Boudjema[1], Mohamed Halawa[1], Jiatong Zhang[1], Simon G. Caulton[1,2], Jonas Emsley[1], and Ingrid Dreveny[1]*

[1]Biodiscovery Institute, School of Pharmacy, University of Nottingham, Nottingham NG7 2RD, United Kingdom.

[2]Present address: Institute for Microbiology and Infection, School of Biosciences, University of Birmingham, Birmingham B15 2TT, UK.

*To whom correspondence should be addressed: Ingrid Dreveny: Biodiscovery Institute, School of Pharmacy, University of Nottingham, Nottingham NG7 2RD, UK; ingrid.dreveny@nottingham.ac.uk; Tel.+44 1158468015

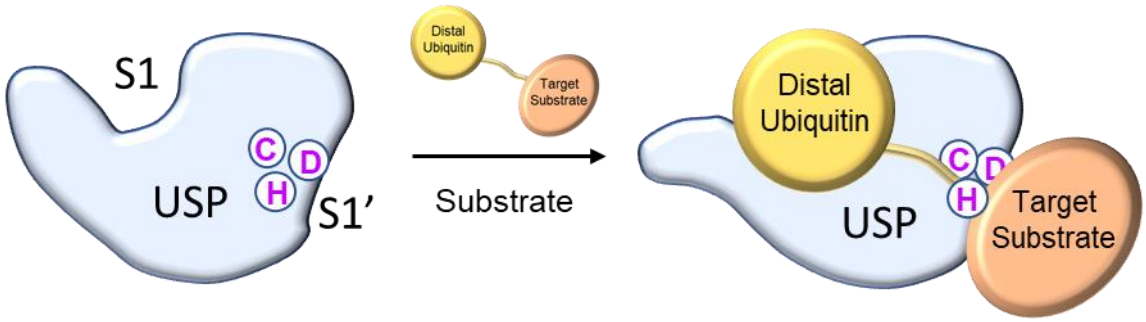


Fig. S1. Schematic representation of substrate binding to a ubiquitin specific protease catalytic domain. The USP catalytic domain composed of D1 and D2 sub-domains is depicted in light blue with the large S1 binding pocket for the distal ubiquitin moiety and much shallower S1' binding site for the conjugated target substrate (which can be a second ubiquitin moiety and would in this case be termed proximal ubiquitin) labelled. Catalytic triad residues are represented in circles.

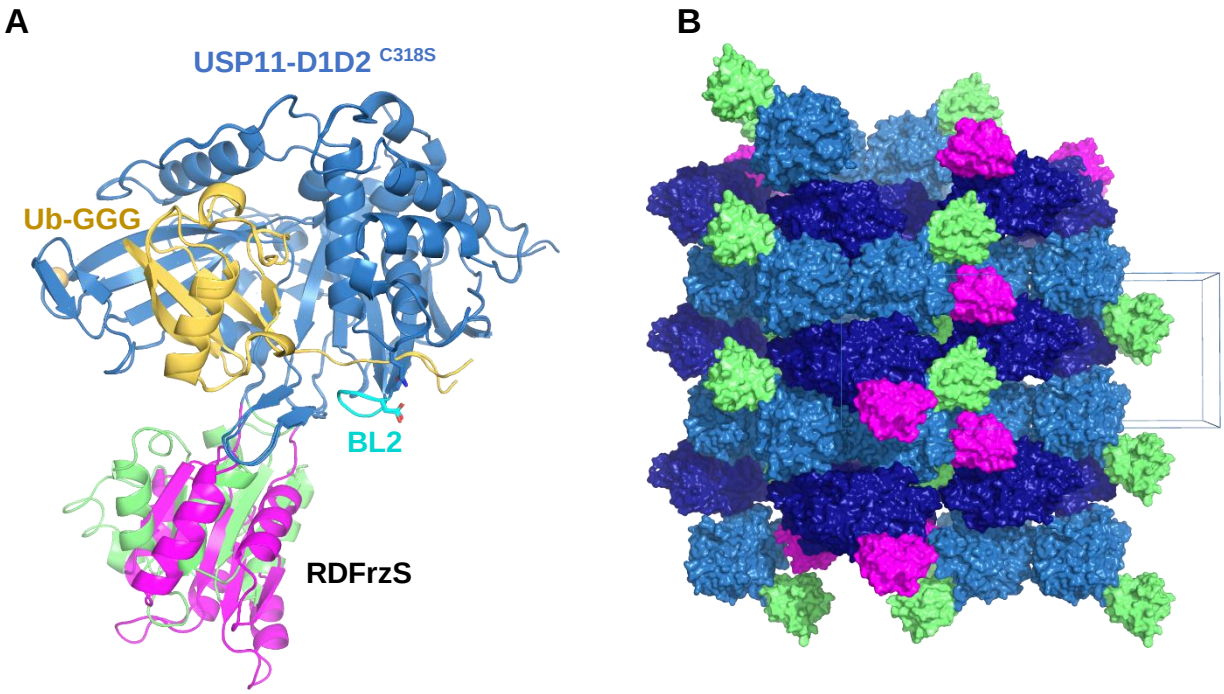


Fig. S2. Crystal packing and copies in asymmetric unit. ***A*** Superposition of the two copies of USP11-D1D2$_{ins}$ ^{C318S} structure in complex with Ub-GGG in the asymmetric unit as cartoon representation with USP11-D1D2 shown in blue and the blocking loop 2 region highlighted in cyan. Ub-GGG molecules are depicted in yellow and the RDFrzS insertion tags are highlighted in magenta and lime green, respectively. ***B*** Crystal packing of USP11-D1D2$_{ins}$ ^{C318S} structure in complex with Ub-GGG with molecules in surface representation. The two copies of the USP11-D1D2 ^{C318S} - Ub-GGG complex in the ASU are shown in light and dark blue, respectively. The RDFrzS insertion tags are colour coded as in ***A***. A unit cell is shown in blue.

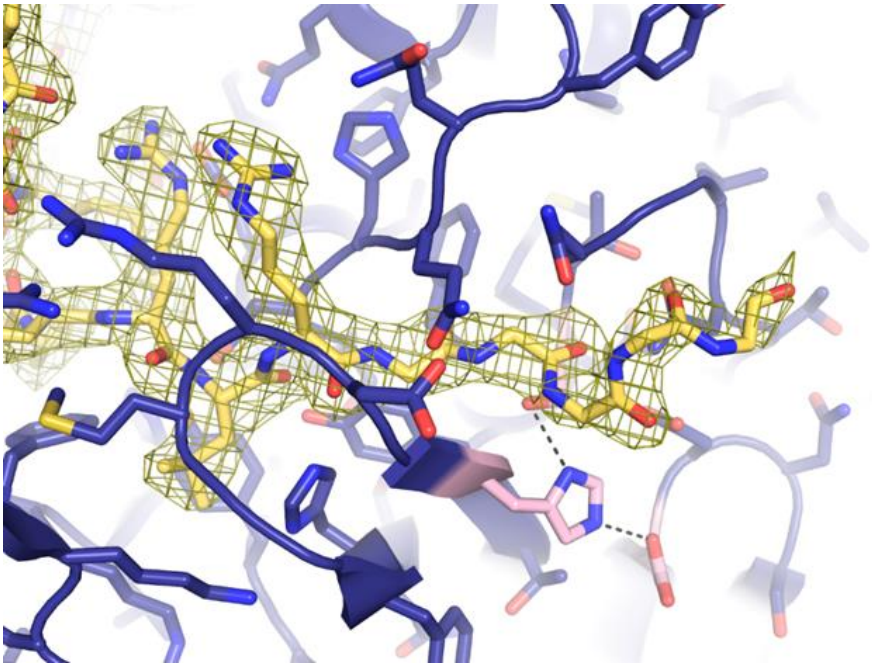


Fig. S3. Electron density of the extended ubiquitin tail region. Cartoon representation depicting the C-terminal residues of the minimal substrate Ub-GGG. USP11 is shown in blue, Ub-GGG in light orange and the catalytic triad residues with the C318S mutation are shown in light pink. Side chains are displayed as sticks and the sigmaA weighted 2Fobs − Fcalc electron density map for Ub-GGG is contoured at 1.05 σ.

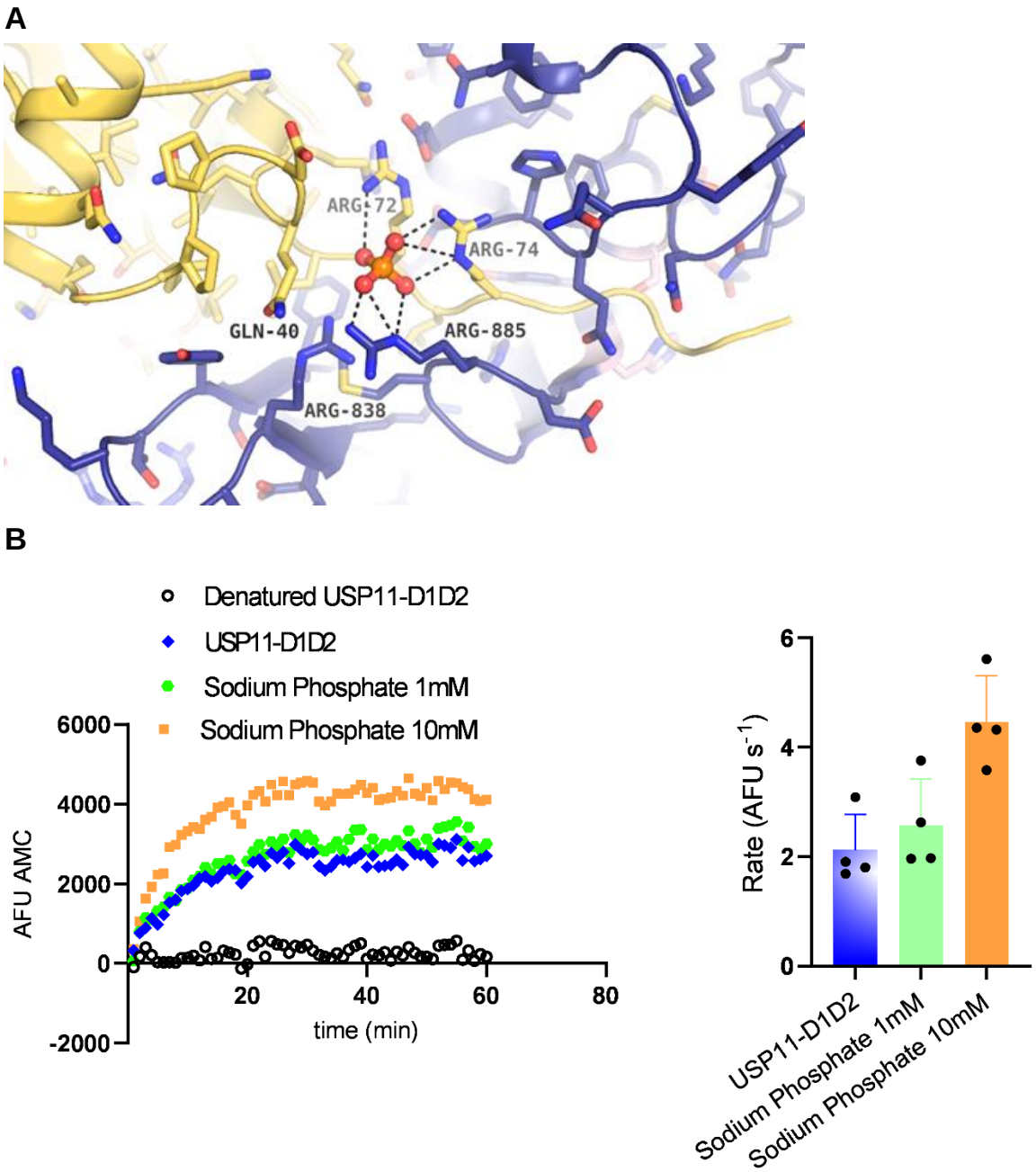


Fig. S4. Phosphate binding site in the structure of USP11-D1D2$_{ins}$ ^{C318S} in complex with Ub-GGG. ***A*** Detail of phosphate ion binding in the USP11 Ub-tail binding channel with engagement of the BL1 and BL2 regions (USP11 green, Ub-GGG yellow). ***B*** Progress curves of USP11-D1D2 (40 nM) - Ub-AMC (750 nM) cleavage reactions in the presence and absence of 1 mM (p-value 0.425) or 10 mM (p-value 0.0045) sodium phosphate with a bar chart (unpaired two-tailed t-test; n=4 independent experiments; error=SD) highlighting that at 10 mM sodium phosphate the rate is slightly enhanced.

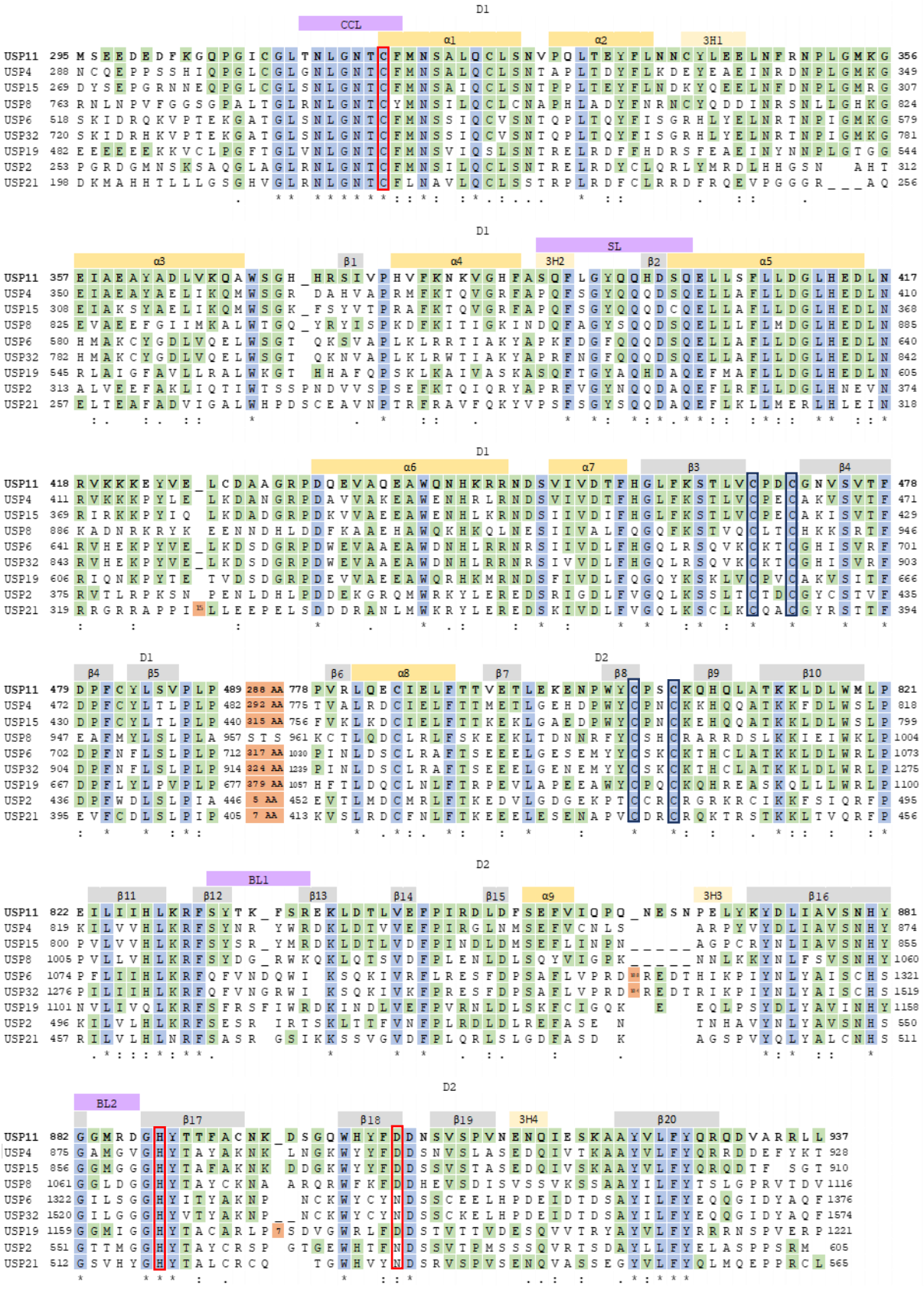
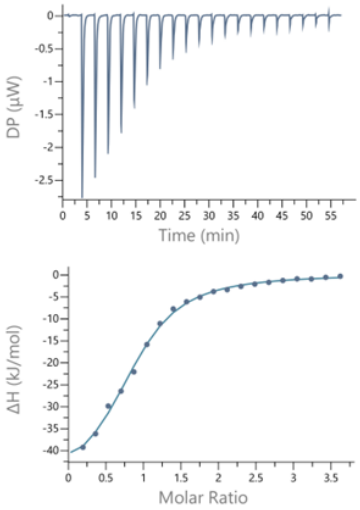
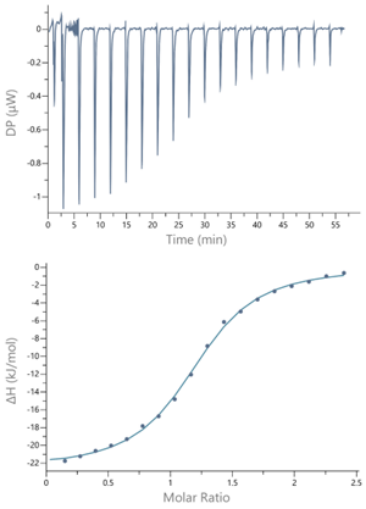

D1
CCL
α1
α2
3H1
USP11 295 M S E E D E D F K G Q P G I C G L T N L G N T C F M N S A L Q C L S N V P Q L T E Y F L N N C Y L E E L N F R N P L G M K G 356
USP4 288 N C Q E P P S S H I Q P G L C G L G N L G N T C F M N S A L Q C L S N T A P L T D Y F L K D E Y E A E I N R D N P L G M K G 349
USP15 269 D Y S E P G R N N E Q P G L C G L S N L G N T C F M N S A I Q C L S N T P P L T E Y F L N D K Y Q E E L N F D N P L G M R G 307
USP8 763 R N L N P V F G G S G P A L T G L R N L G N T C Y M N S I L Q C L C N A P H L A D Y F N R N C Y Q D D I N R S N L L G H K G 824
USP6 518 S K I D R Q K V P T E K G A T G L S N L G N T C F M N S S I Q C V S N T Q P L T Q Y F I S G R H L Y E L N R T N P I G M K G 579
USP32 720 S K I D R H K V P T E K G A T G L S N L G N T C F M N S S I Q C V S N T Q P L T Q Y F I S G R H L Y E L N R T N P I G M K G 781
USP19 482 E E E E E K K V C L P G F T G L V N L G N T C F M N S V I Q S L S N T R E L R D F F H D R S F E A E I N Y N N P L G T G G 544
USP2 253 P G R D G M N S K S A Q G L A G L R N L G N T C F M N S I L Q C L S N T R E L R D Y C L Q R L Y M R D L H H G S N A H T 312
USP21 198 D K M A H H T L L L G S G H V G L R N L G N T C F L N A V L Q C L S S T R P L R D F C L R R D F R Q E V P G G G R _ _ _ A Q 256
D1
SL
α3
β1
α4
3H2
β2
α5
USP11 357 E I A E A Y A D L V K Q A W S G H _ H R S I V P H V F K N K V G H F A S Q F L G Y Q Q H D S Q E L L S F L L D G L H E D L N 417
USP4 350 E I A E A Y A E L I K Q M W S G R D A H V A P R M F K T Q V G R F A P Q F S G Y Q Q Q D S Q E L L A F L L D G L H E D L N 410
USP15 308 E I A K S Y A E L I K Q M W S G K _ F S Y V T P R A F K T Q V G R F A P Q F S G Y Q Q Q D C Q E L L A F L L D G L H E D L N 368
USP8 825 E V A E E F G I I M K A L W T G Q _ Y R Y I S P K D F K I T I G K I N D Q F A G Y S Q Q D S Q E L L L F L M D G L H E D L N 885
USP6 580 H M A K C Y G D L V Q E L W S G T Q K S V A P L K L R R T I A K Y A P K F D G F Q Q Q D S Q E L L A F L L D G L H E D L N 640
USP32 782 H M A K C Y G D L V Q E L W S G T Q K N V A P L K L R W T I A K Y A P R F N G F Q Q Q D S Q E L L A F L L D G L H E D L N 842
USP19 545 R L A I G F A V L L R A L W K G T H H A F Q P S K L K A I V A S K A S Q F T G Y A Q H D A Q E F M A F L L D G L H E D L N 605
USP2 313 A L V E E F A K L I Q T I W T S S P N D V V S P S E F K T Q I Q R Y A P R F V G Y N Q Q D A Q E F L R F L L D G L H N E V N 374
USP21 257 E L T E A F A D V I G A L W H P D S C E A V N P T R F R A V F Q K Y V P S F S G Y S Q Q D A Q E F L K L L M E R L H L E I N 318
D1
α6
α7
β3
β4
USP11 418 R V K K K E Y V E _ L C D A A G R P D Q E V A Q E A W Q N H K R R N D S V I V D T F H G L F K S T L V C P D C G N V S V T F 478
USP4 411 R V K K K P Y L E L K D A N G R P D A V V A K E A W E N H R L R N D S V I V D T F H G L F K S T L V C P E C A K V S V T F 471
USP15 369 R I R K K P Y I Q L K D A D G R P D K V V A E E A W E N H L K R N D S I I V D I F H G L F K S T L V C P E C A K I S V T F 429
USP8 886 K A D N R K R Y K E E N N D H L D D F K A A E H A W Q K H K Q L N E S I I V A L F Q G Q F K S T V Q C L T C H K K S R T F 946
USP6 641 R V H E K P Y V E _ L K D S D G R P D W E V A A E A W D N H L R R N R S I I V D L F H G Q L R S Q V K C K T C G H I S V R F 701
USP32 843 R V H E K P Y V E _ L K D S D G R P D W E V A A E A W D N H L R R N R S I V V D L F H G Q L R S Q V K C K T C G H I S V R F 903
USP19 606 R I Q N K P Y T E T V D S D G R P D E V V A E E A W Q R H K M R N D S F I V D L F Q G Q Y K S K L V C P V C A K V S I T F 666
USP2 375 R V T L R P K S N P E N L D H L P D D E K G R Q M W R K Y L E R E D S R I G D L F V G Q L K S S L T C T D C G Y C S T V F 435
USP21 319 R R G R R A P P I L L E E P E L S D D D R A N L M W K R Y L E R E D S K I V D L F V G Q L K S C L K C Q A G G Y R S T T F 394
D1
D2
β4
β5
β6
α8
β7
β8
β9
β10
USP11 479 D P F C Y L S V P L P 489 288 AA 778 P V R L Q E C I E L F T T V E T L E K E N P W Y C P S C K Q H Q L A T K K L D L W M L P 821
USP4 472 D P F C Y L T L P L P 482 292 AA 775 T V A L R D C I E L F T T M E T L G E H D P W Y C P N C K K H Q Q A T K K F D L W S L P 818
USP15 430 D P F C Y L T L P L P 440 315 AA 756 F V K L K D C I E L F T T K E K L G A E D P W Y C P N C K E H Q Q A T K K L D L W S L P 799
USP8 947 E A F M Y L S L P L A 957 S T S 961 K C T L Q D C L R L F S K E E K L T D N N R F Y C S H C R A R R D S L K K I E I W K L P 1004
USP6 702 D P F N F L S L P L P 712 317 AA 1030 P I N L D S C L R A F T S E E E L G E S E M Y Y C S K C K T H C L A T K K L D L W R L P 1073
USP32 904 D P F N F L S L P L P 914 324 AA 1239 P I N L D S C L R A F T S E E E L G E N E M Y Y C S K C K T H C L A T K K L D L W R L P 1275
USP19 667 D P F L Y L P V P L P 677 379 AA 1057 H F T L D Q C L N L F T R P E V L A P E E A W Y C P Q C K Q H R E A S K Q L L L W R L P 1100
USP2 436 D P F W D L S L P I A 446 5 AA 452 E V T L M D C M R L F T K E D V L G D G E K P T C C R C R G R K R C I K K F S I Q R F P 495
USP21 395 E V F C D L S L P I P 405 7 AA 413 K V S L R D C F N L F T K E E L E S E N A P V C D R C R Q K T R S T K K L T V Q R F P 456
D2
BL1
β11
β12
β13
β14
β15
α9
3H3
β16
USP11 822 E I L I I H L K R F S Y T K _ F S R E K L D T L V E F P I R D L D F S E F V I Q P Q _ N E S N P E L Y K Y D L I A V S N H Y 881
USP4 819 K I L V V H L K R F S Y N R Y W R D K L D T V V E F P I R G L N M S E F V C N L S A R P Y V Y D L I A V S N H Y 874
USP15 800 P V L V V H L K R F S Y S R _ Y M R D K L D T L V D F P I N D L D M S E F L I N P N _ _ _ _ _ A G P C R Y N L I A V S N H Y 855
USP8 1005 P V L L V H L K R F S Y D G _ R W K Q K L Q T S V D F P L E N L D L S Q Y V I G P K _ _ _ _ _ N N L K K Y N L F S V S N H Y 1060
USP6 1074 P F L I I H L K R F Q F V N D Q W I K S Q K I V R F L R E S F D P S A F L V P R D R E D T H I K P I Y N L Y A I S C H S 1321
USP32 1276 P I L I I H L K R F Q F V N G R W I K S Q K I V K F P R E S F D P S A F L V P R D R E D T R I K P I Y N L Y A I S C H S 1519
USP19 1101 N V L I V Q L K R F S F R S F I W R D K I N D L V E F P V R N L D L S K F C I G Q K E E Q L P S Y D L Y A V I N H Y 1158
USP2 496 K I L V L H L K R F S E S R I R T S K L T T F V N F P L R D L D L R E F A S E N T N H A V Y N L Y A V S N H S 550
USP21 457 R I L V L H L N R F S A S R G S I K K S S V G V D F P L Q R L S L G D F A S D K A G S P V Y Q L Y A L C N H S 511
D2
BL2
β17
β18
β19
3H4
β20
USP11 882 G G M R D G H Y T T F A C N K _ D S G Q W H Y F D D N S V S P V N E N Q I E S K A A Y V L F Y Q R Q D V A R R L L 937
USP4 875 G A M G V G H Y T A Y A K N K L N G K W Y Y F D D S N V S L A S E D Q I V T K A A Y V L F Y Q R R D D E F Y K T 928
USP15 856 G G M G G G H Y T A F A K N K D D G K W Y Y F D D S S V S T A S E D Q I V S K A A Y V L F Y Q R Q D T F S G T 910
USP8 1061 G G L D G G H Y T A Y C K N A A R Q R W F K F D D H E V S D I S V S S V K S S A A Y I L F Y T S L G P R V T D V 1116
USP6 1322 G I L S G G H Y I T Y A K N P N C K W Y C Y N D S S C E E L H P D E I D T D S A Y I L F Y E Q Q G I D Y A Q F 1376
USP32 1520 G I L G G G H Y V T Y A K N P _ _ N C K W Y C Y N D S S C K E L H P D E I D T D S A Y I L F Y E Q Q G I D Y A Q F 1574
USP19 1159 G G M I G G H Y T A C A R L P 7 S D V G W R L F D D S T V T T V D E S Q V V T R Y A Y V L F Y R R R N S P V E R P 1221
USP2 551 G T T M G G H Y T A Y C R S P G T G E W H F N D S S V T P M S S S Q V R T S D A Y L L F Y E L A S P P S R M 605
USP21 512 G S V H Y G H Y T A L C R C Q T G W H V Y N D S R V S P V S E N Q V A S S E G Y V L F Y Q L M Q E P P R C L 565

Fig. S5. Sequence alignment of selected USPs. Sequence alignment of the USP11 catalytic core with 8 closely related human USPs. The USP11 catalytic core sequence (295-489 = D1, and 778-937 = D2, UniProt P51784) is shown in alignment with the catalytic core sequences of USP4 (UniProt Q13107), USP15 (UniProt Q9Y4E8-2), USP8 UniProt (P40818), USP6 (UniProt P35125), USP32 (UniProt Q8NFA0), USP19 (UniProt O94966), USP2 (UniProt O75604) and USP21 (UniProt Q9UK80), alignment performed using Clustal Omega. Full sequence conservation between the proteins is highlighted in blue, partial sequence conservation compared to USP11 is marked in green. The following conservation markers of the Clustal Omega output are also shown: * = fully conserved; : = similar properties; . = weakly similar properties; blank = no conservation. The CCL, SL, BL1 and BL2 loop regions are indicated in purple. The catalytic triad and zinc finger residues are boxed in red and blue, respectively. Also indicated are the secondary structure elements (α-helix (yellow), 310-helix (light yellow), β-strand (grey)) of USP11 as seen in the USP11-D1D2 ^{C318S} substrate complex structure, assigned using the DSSP server (https://swift.cmbi.umcn.nl/gv/dssp/).

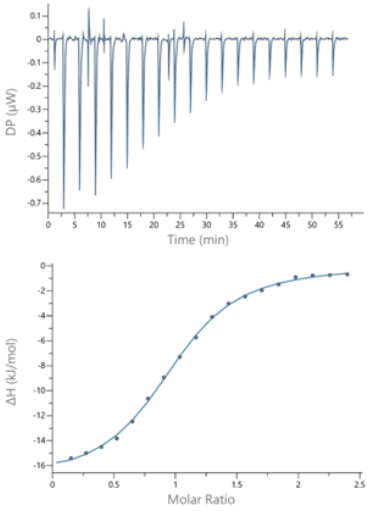
USP11-D1D2
+ Ub



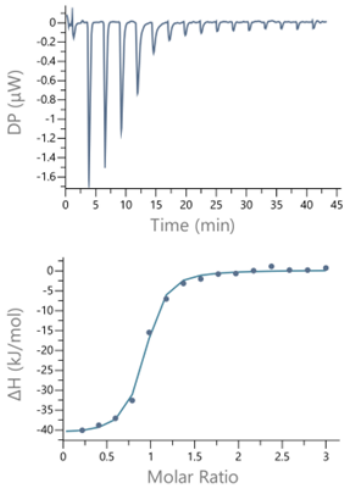
USP11-D1D2 R885G
+ Ub



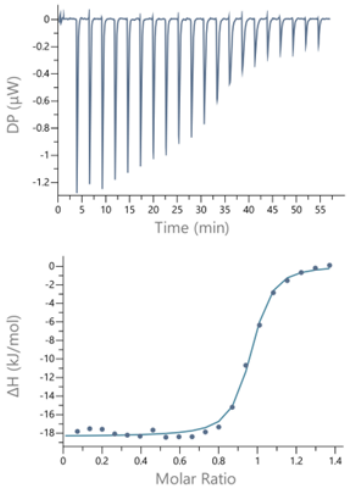
USP11-D1D2 H399Q
+ Ub



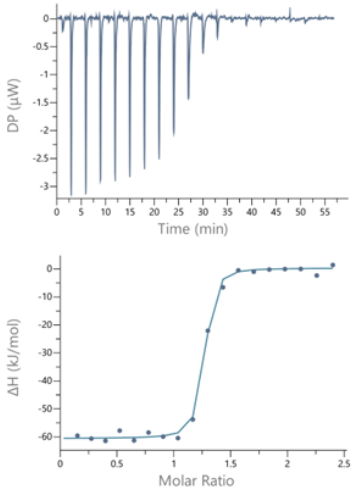
USP11-D1D2 C318S
+ Ub



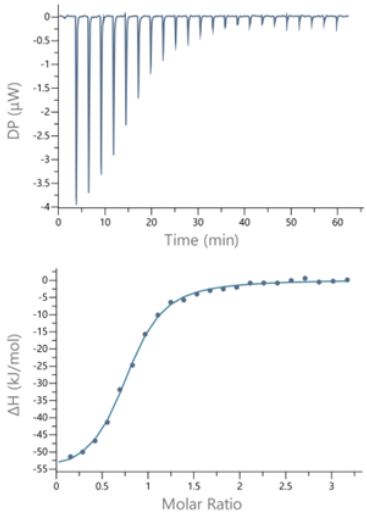
USP11-D1D2 C318S
+ Ub-G



USP11-D1D2 C318S
+ Ub-GGG

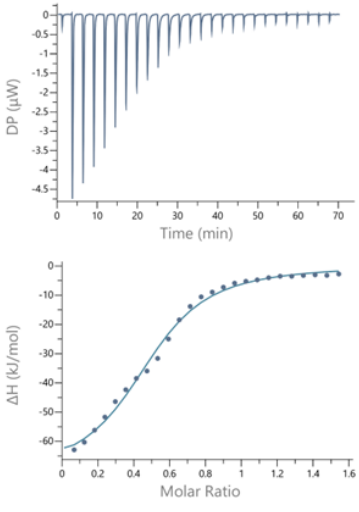


USP11-D1D2 C318S
+ Ub-M



*

USP11-D1D2 C318S
+ linear di-Ub



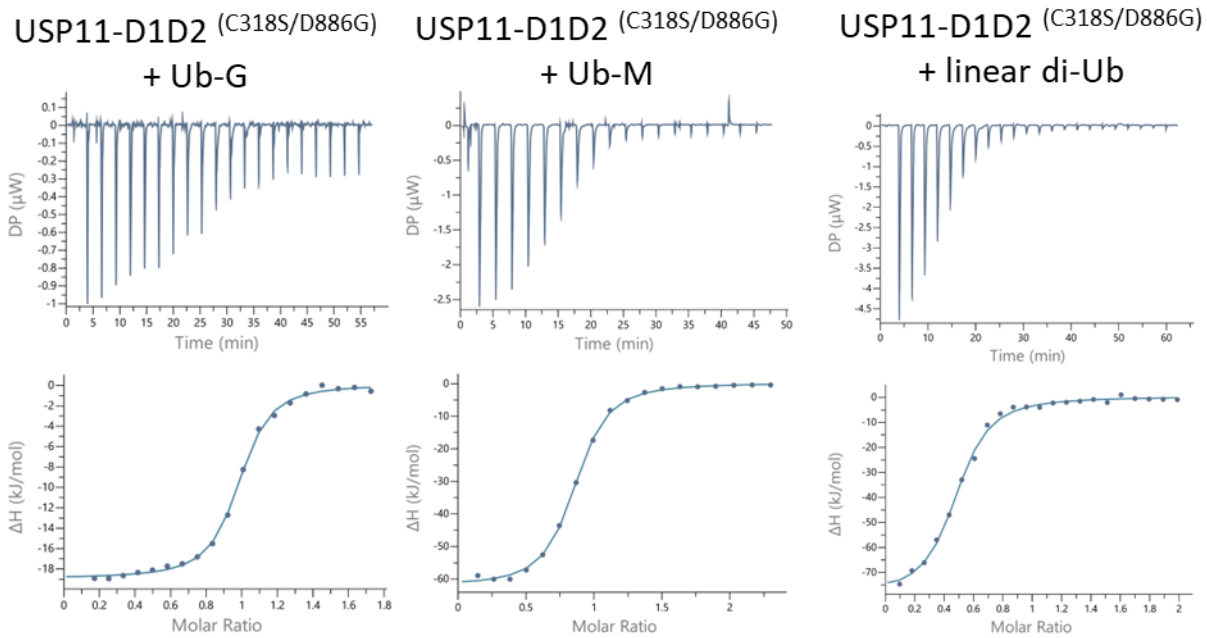


Fig. S6. Isothermal Titration Calorimetry data. Representative examples of ITC raw data and binding isotherms fitted to a one site binding model for USP11 catalytic cores and different substrates, or the product mono-ubiquitin measured at 25°C. * for the USP11-D1D2 ^{C318S} - linear di-Ub titration curve there are indications for two binding sites, but fitting does not converge using a two binding site model.

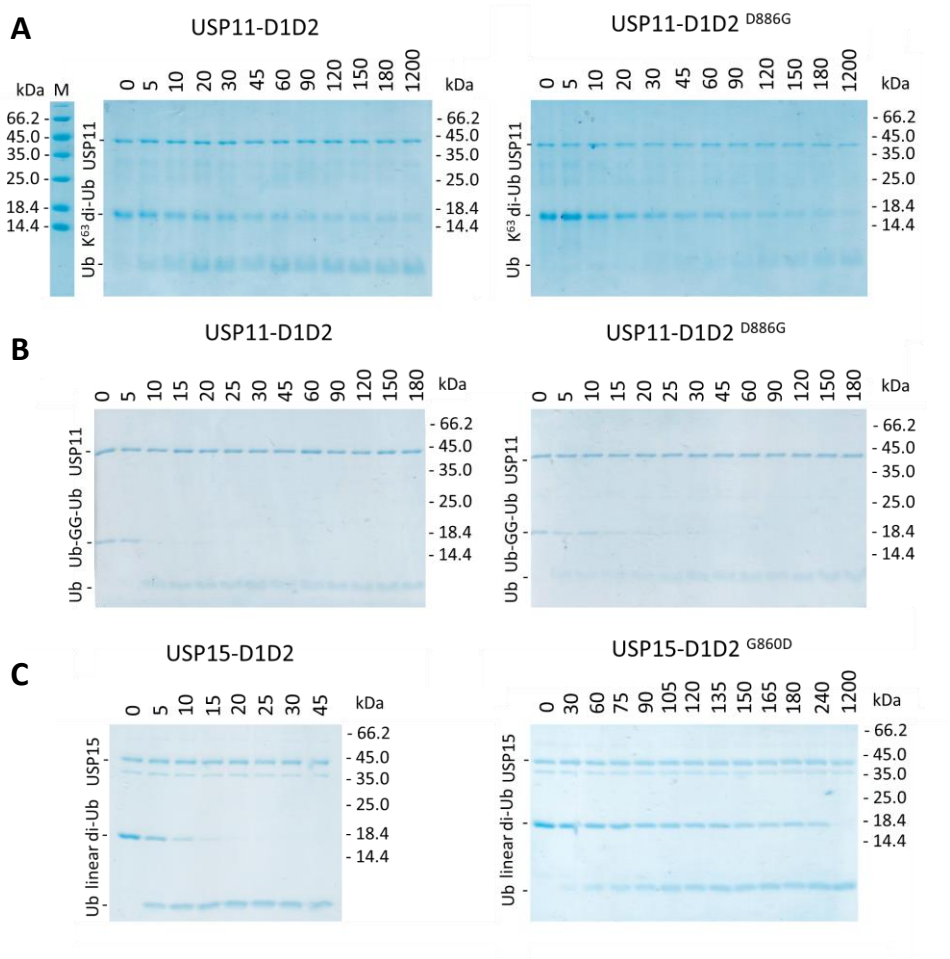


Fig. S7. **Substrate cleavage assays**. Representative examples of SDS-PAGE gels of cleavage assays for ***A, B*** USP11-D1D2 vs USP11-D1D2 ^{D886G} and ***C*** USP15-D1D2 vs USP15-D1D2 ^{G860D} catalytic cores with different substrates including Lys^{63}-linked di-Ub (***A***), Ub-GG-Ub (***B***) or linear di-Ub (***C***).

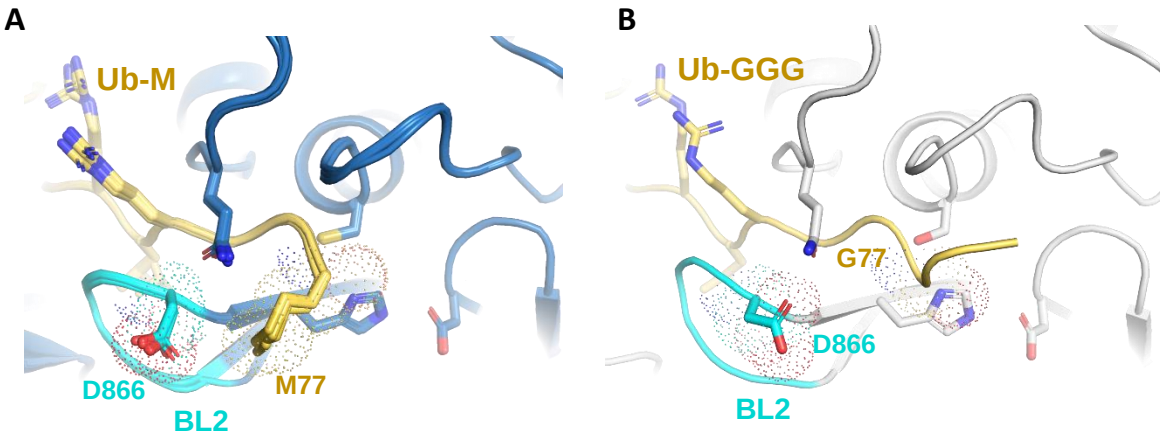


Fig. S8. **Comparison of the predicted USP11-D1D2 structural model in complex with Ub-M with the USP11-D1D2 - Ub-GGG crystal structure.** ***A*** Cartoon representation of the top five USP11-D1D2 models (blue) in complex with Ub-M (yellow) computed by AlphaFold (1). The BL2 region is highlighted in cyan and USP11 Asp^{866}, Gln^{398} and the catalytic triad residues are shown in stick representation as well as the C-terminal residues of Ub-M. In addition, USP11 Asp^{866} and Ub-M Met^{77} are shown in dots representation. USP11 Asp^{866} faces away from the Ub-M C-terminal methionine. ***B*** Crystal structure of USP11-D1D2 (grey) in complex with Ub-GGG (yellow) as reported here. Colour coding as well as residues shown in sticks are equivalent to (***A***). USP11 Asp^{866} and Ub-GGG Gly^{77} are shown as dots.

References

1. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P., and Hassabis, D. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589