

Supplementary Information

Dual domain recognition determines SARS-CoV-2 PLpro selectivity for human ISG15 and K48-linked di-ubiquitin

Jerzy Osipiuk, Pawel M. Wydorski, Benjamin T. Lanham, Christine Tesar, Michael Endres, Elizabeth Engle, Robert Jedrzejczak, Vishruth Mullapudi, Karolina Michalska, Krzysztof Fidelis, David Fushman, Andrzej Joachimiak and Lukasz A. Joachimiak

Source Data 1. Summary of raw crosslinking mass spectrometry data for PLpro^{CoV-2}:substrate complexes.

Supplementary Table 1. Data processing and refinement statistics.

Structure	PLpro:K48-Ub ₂	PLpro:hISG15	K48-Ub ₂	hISG15
Data processing				
Space group	C2	C222 ₁	P2 ₁	P2 ₁
Cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	148, 50, 72	154, 221, 233	24, 57, 46	52, 135, 78
<i>α</i> , <i>β</i> , <i>γ</i> (°)	90, 112, 90	90, 90, 90	90, 93, 90	90, 108, 90
Resolution range (Å) ^a	47.1-1.88 (1.91–1.88)	49.0-2.98 (3.08-2.98)	46.1-1.25 (1.27–1.25)	46.5-2.15 (2.19-2.15)
Unique reflections ^a	38,678 (1,498)	78,254 (3,893)	28,558 (576)	54,460 (2,545)
R-merge ^b	0.106 (0.929)	0.214 (2.18)	0.072 (0.50)	0.070 (0.946)
Mean I/sigma(I)	13.3 (1.10)	12.6 (1.23)	21.2 (1.79)	21.2 (1.12)
CC1/2 ^c	0.986 (0.520)	0.985 (0.364)	0.994 (0.624)	0.991 (0.511)
Completeness (%)	96.1 (74.2)	99.9 (100)	84.7 (34.9)	98.3 (91.6)
Redundancy	3.2 (2.2)	6.9 (6.7)	3.2 (2.1)	3.4 (2.8)
Wilson B-factor (Å ²)	31.8	54.3	11.3	43.5
Refinement				
Resolution range (Å)	47.1-1.88	49.0-2.98	46.1-1.25	46.5-2.15
Reflections work/test	36,778 (1,900)	74,428 (3,786)	27,176 (1,363)	51,806 (2,625)
R _{work} /R _{free}	0.187 / 0.228	0.195 / 0.236	0.129 / 0.177	0.221 / 0.268
RMSD (bonds) (Å)	0.010	0.004	0.009	0.010
RMSD (angles) (°)	1.59	1.28	1.58	1.61
Number of TLS groups	2	10	- ^e	6
Total protein residues	471	2,375	153	942
Number of atoms				
Protein	3122	18423	1286	7000
Ligand / ion	11	5	8	-
Water	169	54	141	119
B-factors				
Average B-factor (Å ²)	46.8	81.9	19.2	58.0
Protein	46.8	81.9	17.6	58.3

Ligand / ion	53.8	142.7	31.8	-
Water	46.8	52.8	33.3	44.8
MolProbity validation				
Rotamer outliers (%) ^d	0.87	1.62	0.65	5.67
Clashscore ^d	3.68	1.83	2.21	2.47
Ramachandran outliers (%) ^d	0.00	0.04	0.00	0.00
MolProbity score ^d	1.40	1.48	1.00	1.60
PDB ID	7RBR	7RBS	7S6O	7S6P

^a Values in parentheses correspond to the highest resolution shell.

^b $R_{\text{merge}} = \frac{\sum_h \sum_j |I_{hj} - \langle I_h \rangle|}{\sum_h \sum_j I_{hj}}$, where I_{hj} is the intensity of observation j of reflection h .

^c As defined by Karplus and Diederichs¹.

^d As defined by Molprobity².

^e Anisotropic refinement of temperature factors (no TLS groups)

Supplementary Figure 1

a

Distal/Proximal domains

miSG15	100%		
hISG15	63.35%	100%	
Ub2	33.55%	32.89%	100%
miSG15		hISG15	Ub2

Distal domains

miSG15	100%		
hISG15	65.33%	100%	
Ub2	34.66%	28.94%	100%
miSG15		hISG15	Ub2

Proximal domains

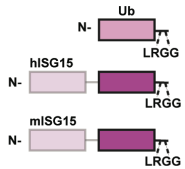
miSG15	100%		
hISG15	61.17%	100%	
Ub2	32.89%	36.84%	100%
miSG15		hISG15	Ub2

b

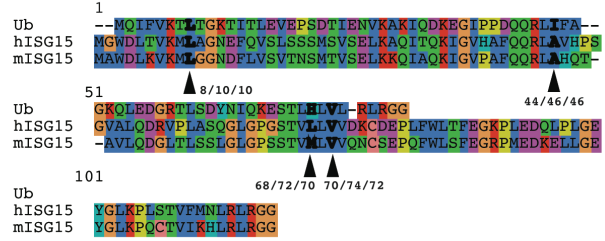
distal



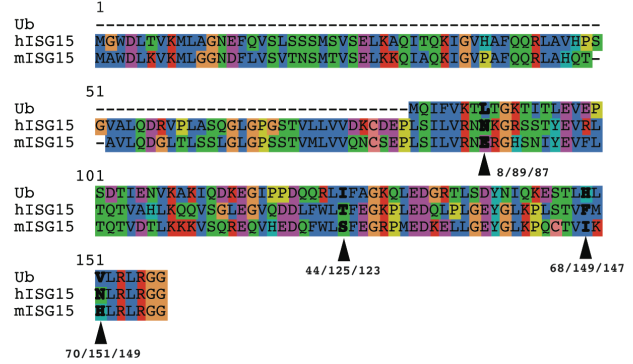
proximal



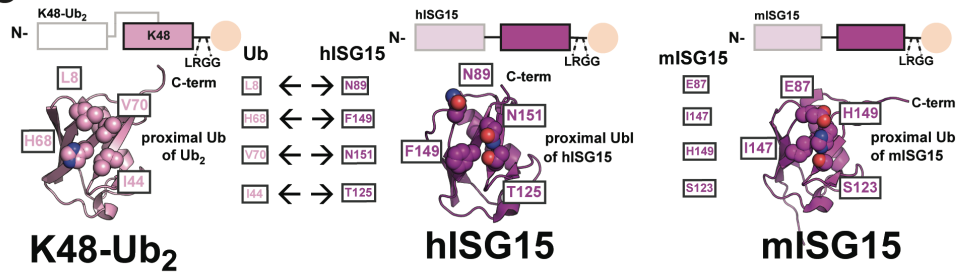
distal UBL of hISG15 to Ub



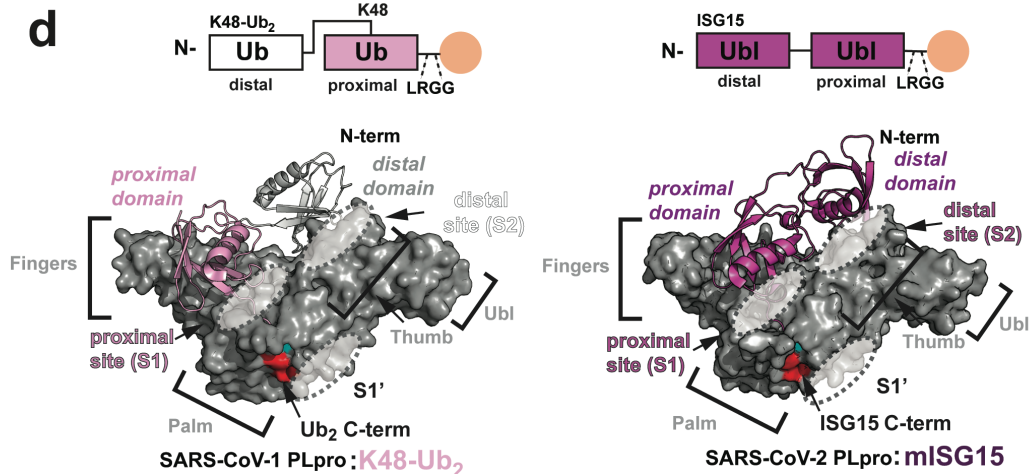
proximal UBL of hISG15 to Ub



c



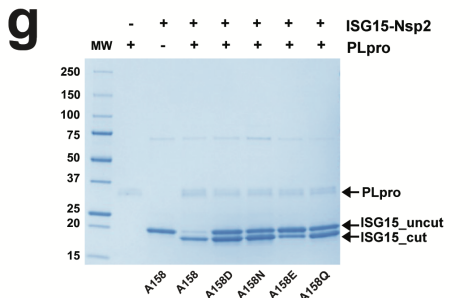
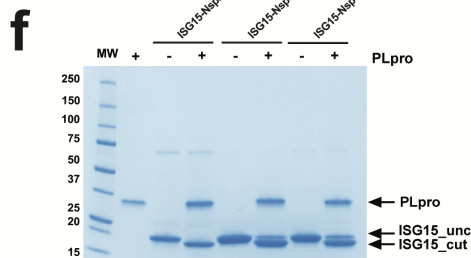
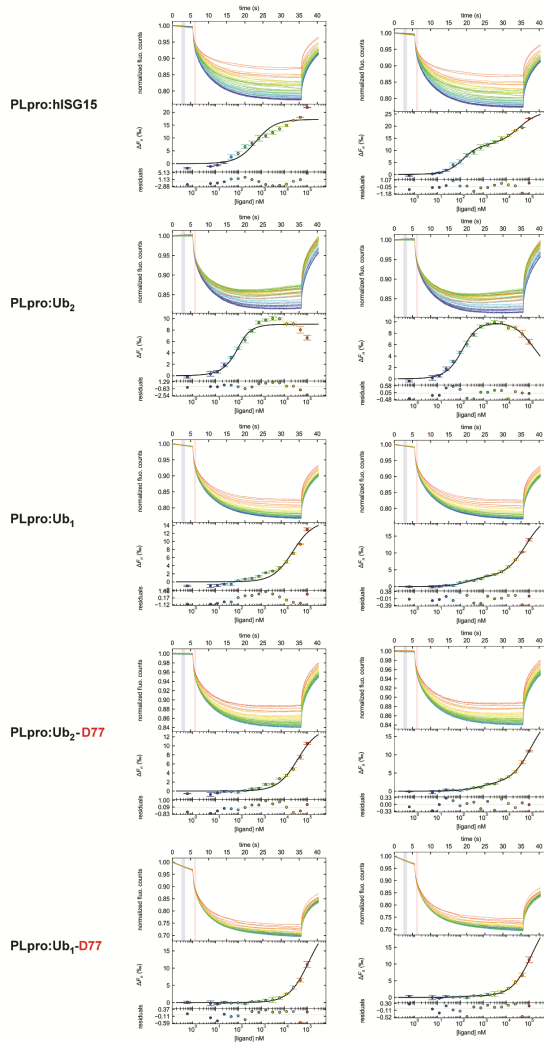
d



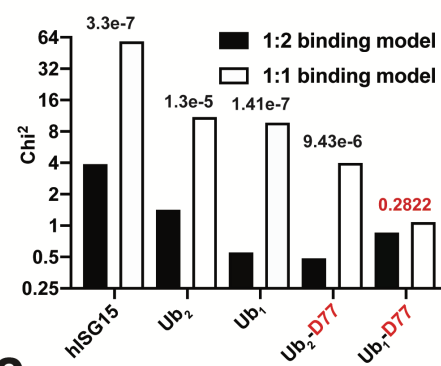
Supplementary Figure 1. PLpro architecture and substrate recognition. (a) Sequence identities

between Ub₂, mISG15 and hISG15 for full sequences and individual domains distal (N-terminal) and proximal (C-terminal) alone. **(b)** Sequence alignments of distal and proximal Ub and UBLs of hISG15 and mISG15. Ub and ISG15 are shown in cartoon representation and are colored pink and magenta. Ub nonpolar patch residues important for Ub recognition is highlighted by arrows on the proximal and distal domains hISG15 and mISG15. **(c)** Key residues that define the nonpolar patch on Ub and ISG15 (proximal domain) are shown in spacefill and highlighted with arrows. **(d)** Binding mode of K48-Ub₂ to PLpro^{CoV-1} and mISG15 to PLpro^{CoV-2}. PLpro is shown in surface representation and colored grey. K48-Ub₂ is shown in cartoon representation and colored white and pink for the proximal and distal Ubs, respectively. We call ‘distal domain’ the Ub unit connected through its C-terminus to K48 of the other Ub unit, called ‘proximal domain’. We define the ‘proximal domain’ and ‘distal domain’ binding sites on PLpro as S1 and S2, respectively, while S1’ hangs off the active site. mISG15 is shown in cartoon representation and colored magenta. In analogy with the two Ubs in Ub₂, the N- and C-terminal UBL (ubiquitin like) domains of ISG15 will be referred to as the distal and proximal UBLs, respectively.

a 1:1 fitting model 1:2 fitting model



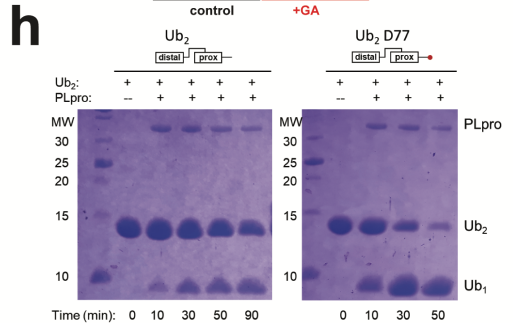
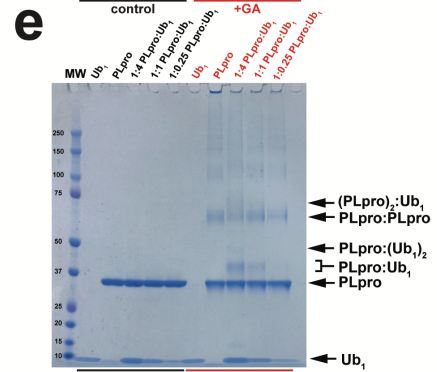
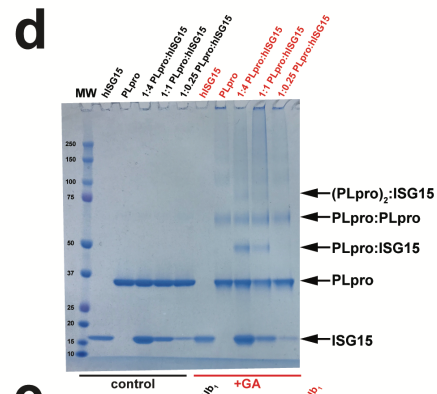
b Supplementary Figure 2



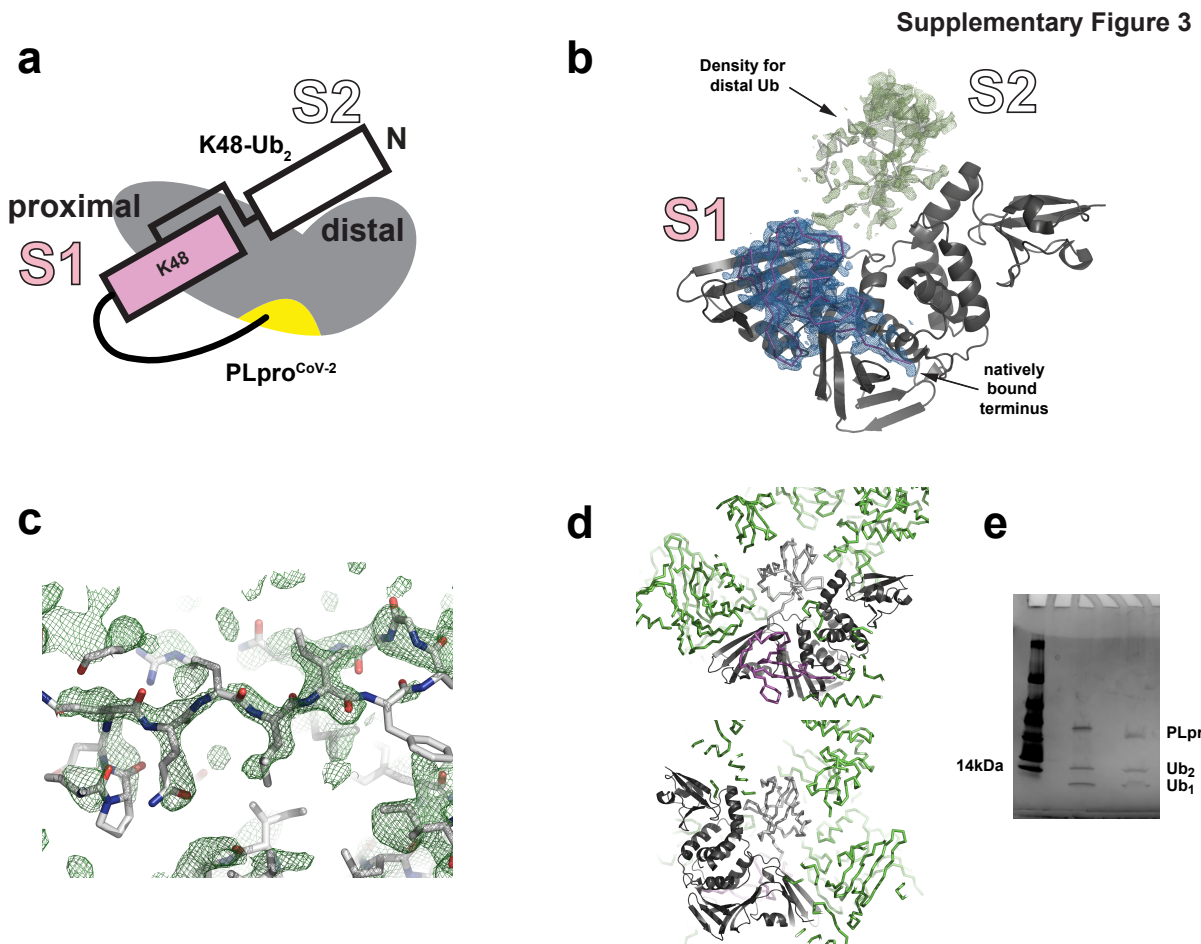
c

	K_{D1} (μ M)				
	ISG15	Ub ₂	Ub ₁	Ub ₂ -D77	Ub ₁ -D77
AVG	0.09	0.16	0.16	0.79	146.67
STDEV	0.03	0.03	0.07	0.54	90.74

	K_{D2} (μ M)			
	ISG15	Ub ₂	Ub ₁	Ub ₂ -D77
AVG	42.67	460.00	75.67	1406.67
STDEV	16.17	557.49	31.88	2246.69



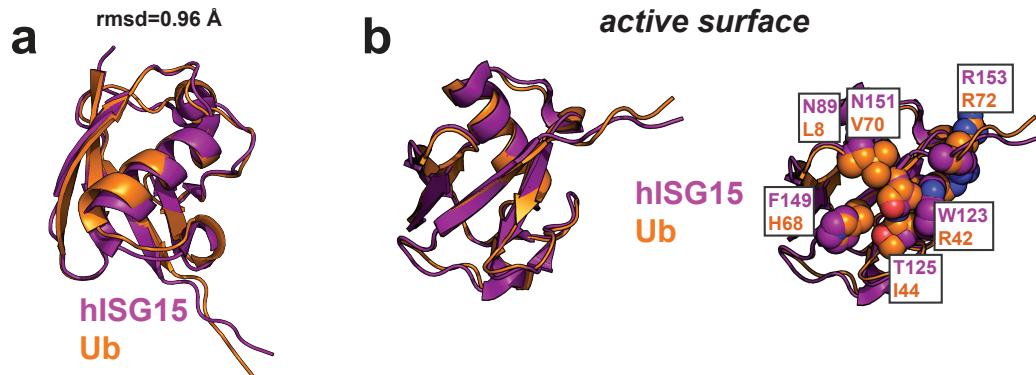
Supplementary Figure 2. 2:1 PLpro^{CoV-2}:substrate binding model explains binding data. (a) kinetic fitting of 1:1 (left) and 1:2 (right) stoichiometry binding models for PLpro^{CoV-2} binding to hISG15, Ub₂, Ub₁, Ub₂ (D77) and Ub₁ (D77). Cumulative raw thermophoresis data for three independent titration experiments. The curves are colored by concentration of PLpro from blue (low) to red (high) and plotted as function of time and normalized fluorescence counts. Change fluorescence (ΔF) is plotted as a function of substrate concentration colored as above. Data is shown as averages with standard deviation across three independent experiments. Best fit is shown for 1:1 and 1:2 binding models. The variance between the average ΔF value and the fit is plotted below. The residuals are shown as average deviation from the fit across three experiments and is colored as above. (b) Comparison of χ^2 derived from the 1:1 and 1:2 fitting models for the PLpro^{CoV-2} binding to hISG15, Ub₂, Ub₁, Ub₂ (D77) and Ub₁ (D77) binding data. Only for the PLpro^{CoV-2}:Ub₁ (D77) did the 1:2 model not statistically improve (Pval=0.2822). All other binding data were better explained by the 1:2 binding model with statistical significance. (c) Summary of K_d calculated from the 1:2 binding model for all except PLpro^{CoV-2}:Ub₁ (D77). K_d s are reported as averages from three independent experiments with shown with standard deviation. (d,e) Glutaraldehyde cross-linking (0.05%) of PLpro^{CoV-2}:hISG15 (d) or PLpro^{CoV-2}:Ub₁ (e) using 1:4, 1:1 and 1:0.25 ratios and resolved by SDS-PAGE. (f) PLpro^{CoV-2} cleavage of hISG15-Nsp C-terminal fusions. SDS-PAGE gel reveals that PLpro cleaves off fragments of Nsp2 (LRGG[^]AYTRYVDNNF), Nsp3 (LRGG[^]APTKVTFGDD), and Nsp4 (LRGG[^]KIVNNWLKQL) from hISG15-Nsp fusions linked to the recognition motif “LRGG”. (g) PLpro^{CoV-2} cleavage of hISG15-Nsp2 C-terminal fusions containing mutations at the first residue (Asp, Gln and Glu) (Nsp2, LRGG[^]AYTRYVDNNF) after the “LRGG” motif and shown in bold. SDS-PAGE gel reveals that PLpro cleaves off fragments of Nsp2 (**XY**TRYVDNNF), (h) PLpro^{CoV-2} cleavage of Ub₂ and Ub₂ (D77). SDS-PAGE gel reveals that PLpro cleaves Ub₂ (D77) more quickly than Ub₂ into Ub.



Supplementary Figure 3. Strong electron density in the proximal domain for PLpro^{CoV-2}: K48-Ub₂ complex. (a) Schematic illustration for PLpro^{CoV-2}:K48-Ub₂ complex. (b) Structure of PLpro^{CoV-2} in complex with K48-Ub₂ overlaid with obtained electron density maps for K48-Ub₂ substrate. PLpro is shown in cartoon representation and colored dark grey. K48-Ub₂ is shown as ribbons and colored magenta and light grey for the proximal and distal Ubs, respectively. The distal Ub model was obtained from the structure of SARS-CoV-1 PLpro bound to a K48-linked Ub₂ (PDB id: 5E6J) superposed on our SARS-CoV-2 PLpro in complex with a K48-linked Ub₂. Arrows show strong electron density map fragment for C-terminus of proximal domain of K48-Ub₂ and observed indication of helical density in its distal domain. The 2F_o – F_c electron density weighted map for the proximal Ub is shown as a blue mesh, contoured at 1.4σ. The F_o – F_c electron density weighted difference map for the distal Ub area is shown as a green mesh, contoured at 2σ. (c) Close-up of area corresponding to the distal Ub not modelled in the structure showing presence of electron density resembling amino acid chain with side chains corresponding to Ub sequence. The sequence in center

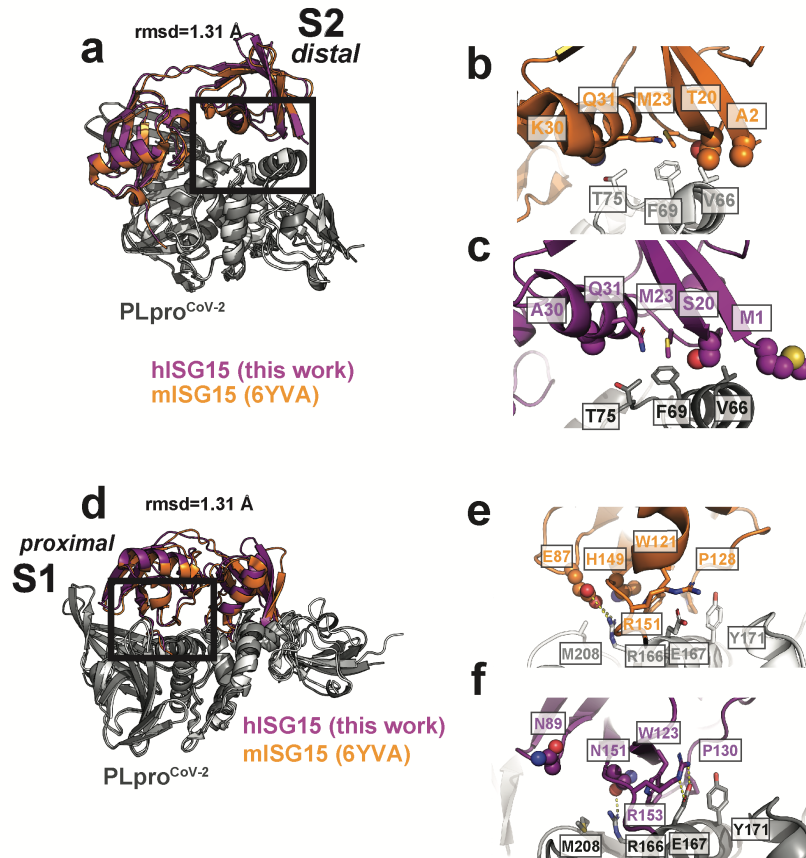
corresponding to Ub residues 41-44 (PDB id: QRLI) is shown in grey with oxygens and nitrogen atoms in red and blue, respectively. The $F_o - F_c$ electron density weighted difference map is shown as a green mesh, contoured at 1.6σ . The 40-46 Ub fragment was moved to the density by Rigid Body Fit Zone in Coot program without any side-chain adjustments. The shift of $C\alpha$ atoms was in 1.3-2.3 Å range. **(d)** The distal Ub not modelled in the structure occupies the "empty" space in PLpro^{CoV-2}-Ub₂ crystal model. PLpro is shown in cartoon representation and colored dark grey. K48-Ub₂ is shown as ribbons and colored magenta and light grey for the proximal and distal Ubs, respectively. The neighbor molecules of PLpro - proximal Ub are shown as green ribbons. **(e)** SDS-PAGE gel of PLpro^{CoV-2}(C111S):Ub₂ crystals used for data collection. PLpro, Ub₂ and Ub₁ are highlighted.

Supplementary Figure 4

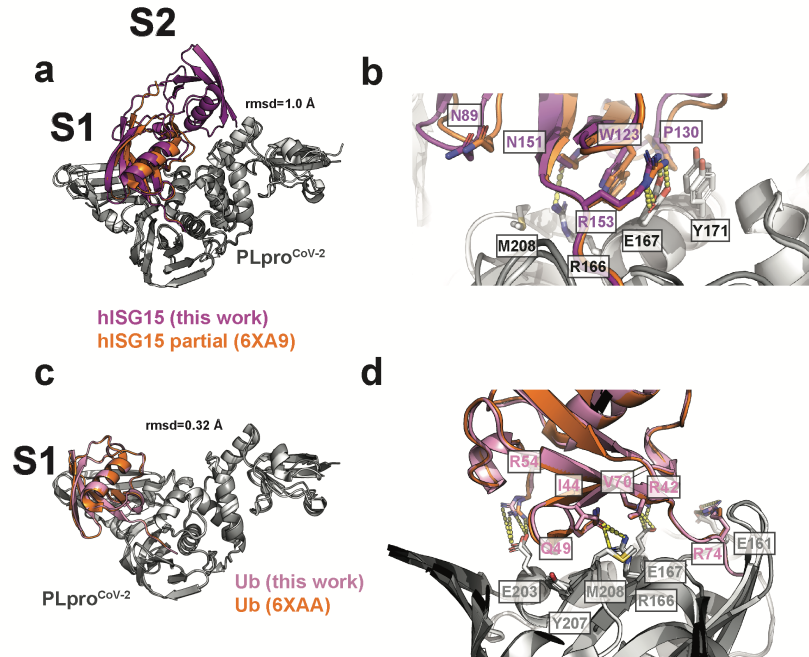


Supplementary Figure 4. Comparison of hISG15 and Ub₁ bound to PLpro^{CoV-2}. (a) Structural alignment of proximal domain of hISG15 to Ub shows high similarity with a Cα rmsd of 0.95 Å. Both, hISG15 and Ub, are shown in cartoon representation. hISG15 is colored magenta and Ub is orange. (b) Comparison of active surface for proximal hISG15 and Ub plus additional contacts involved in binding to PLpro. Key residues are shown in spacefill and are colored as in (a).

Supplementary Figure 5



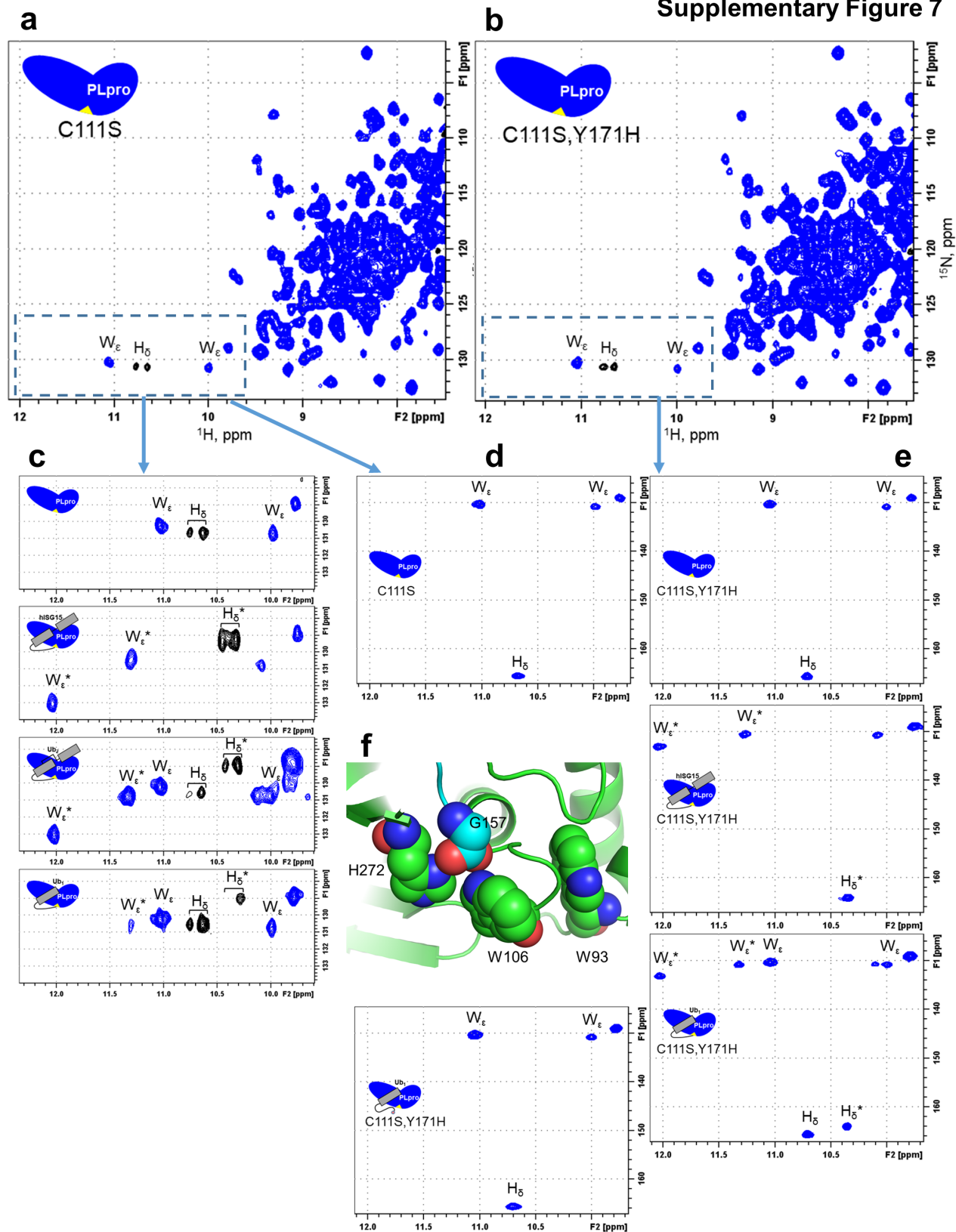
Supplementary Figure 5. Comparison of contacts between hISG15 and mISG15 bound to PLpro^{CoV-2}. (a) Structural overlay of PLpro^{CoV-2}:hISG15 (grey and magenta) from this study (PDB id: 7RBS) and PLpro^{CoV-2}:mISG15 (white and orange) from previously published studies (PDB id: 6VYA) focusing on the distal site (S2). (b,c) Zoom in of boxed contacts from (a) from PLpro to the distal (S2) domain of (b) mISG15 or (c) hISG15. Interface residues are shown in sticks and colored as in (a). Amino acids that vary between mISG15 and hISG15 are shown in spheres. Interface amino acids are labeled. (d) Structural overlay of PLpro^{CoV-2}:hISG15 and PLpro^{CoV-2}:mISG15 focusing on the proximal site (S1) and colored as in (a). (e,f) Zoom in of the boxed contacts from (d) from PLpro to the proximal domain of (e) mISG15 or (f) hISG15. Interface residues are shown in sticks and colored as in (d). Amino acids that vary between mISG15 and hISG15 are shown in spheres. Interface amino acids are labeled.



Supplementary Figure 6. Comparison of partial hISG15 and full hISG15 bound to PLpro^{CoV-2}.

(a) Structural overlay of PLpro^{CoV-2} in complex with hISG15 from this work (PDB id: 7RBS) and from previously published studies with a partial hISG15 (PDB id: 6XA9) shows a C α rmsd of 1.0 Å. PLpro and hISG15 are shown in cartoon representation. PLpro and hISG15 from this paper are colored grey and magenta, respectively. PLpro and partial hISG15 from previous studies are colored white and orange, respectively. (b) Zoom in of the interface comparing the two structures and colored as in (a). Interface residues are labeled and are shown as sticks. Hydrogen bonds are shown as dashed yellow lines. (c) Overlay of PLpro^{CoV-2} in complex with proximal domain of K48-Ub₂ from this paper (PDB id: 7RBR) and from previously published studies (PDB id: 6XAA) shows a C α rmsd of 0.32 Å. PLpro and Ub are shown in cartoon representation. PLpro and proximal domain of Ub from this study is colored grey and pink. PLpro and proximal domain of Ub from previous study is colored in white and orange. (d) Zoom in of the interface comparing the two structures and colored as in (c). Interface residues are labeled and are shown as sticks. Hydrogen bonds are shown as dashed yellow lines.

Supplementary Figure 7



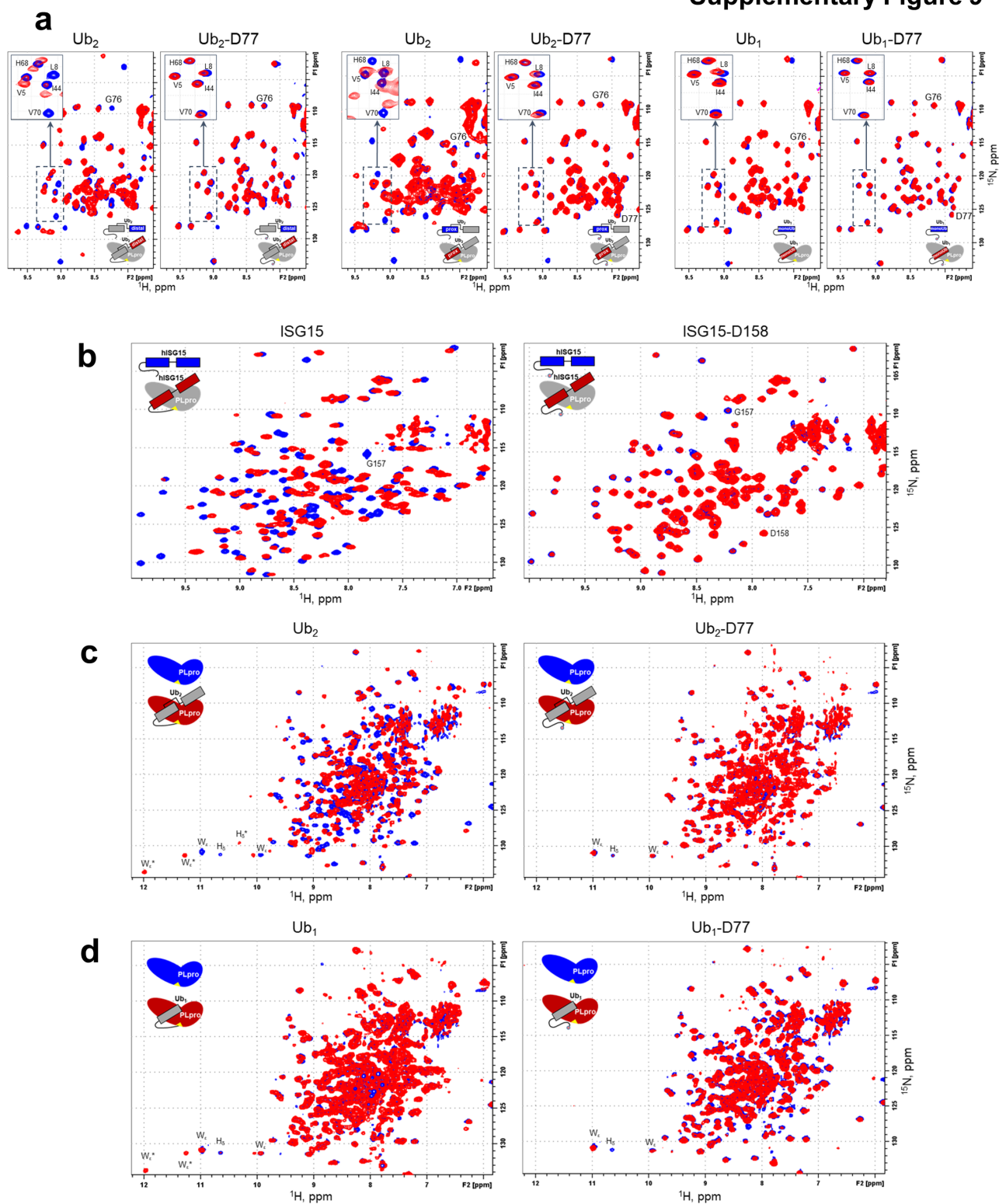
Supplementary Figure 7. NMR characterization and comparison of substrate binding to PLpro^{CoV-2} C111S and C111S,Y171H variants. (a-b) Comparison of ¹H-¹⁵N SOFAST-HMQC spectra of ¹⁵N-labeled PLpro (C111S) (a) and PLpro (C111S,Y171H) (b). The region containing side chain signals of W93, W106, and H272 is marked by a dashed rectangle. Positive signals are blue, negative are black. The histidine imidazole NH group has ¹⁵N chemical shift of 165.5 ppm. Therefore its signal is aliased (folded in) in this spectrum, resulting in negative intensity. It appears as a doublet (consistent with ¹J ~ 94 Hz) because of the inefficient decoupling of ¹⁵N during ¹H acquisition time due to off-resonance effects. See also panels (d and e) where these effects were removed. The crosspeaks corresponding to tryptophan indole and histidine imidazole NH groups are indicated as W_ε and H_δ, respectively. (c) Zoom on the spectral region containing W/H signals (rectangle in panel (a)) for ¹⁵N-PLpro (C111S) alone and upon addition of 1.5 molar equivalents of unlabeled ISG15 or K48-Ub₂ or of 7 molar equivalents of WT Ub. The unbound (W_ε and H_δ) and shifted (W_ε^{*} and H_δ^{*}) side chain tryptophan and histidine signals are indicated. The coloring of spectra is the same as in (a,b). (d) Zoom on the region of ¹H-¹⁵N SOFAST-HMQC spectrum of ¹⁵N-PLpro C111S acquired with the ¹⁵N spectral width of 70 ppm that covers the NH signals from both amides and the side chains of Trp and His, eliminating aliasing and splitting of the His signal. The coloring of signals is the same as in (a,b). (e) Zoom on the region of ¹H-¹⁵N SOFAST-HMQC spectra of ¹⁵N-PLpro (C111S/Y171H) alone and in the presence of the indicated substrates, acquired with the ¹⁵N spectral width of 90 ppm that covers the NH signals from both amides and from the side chains of Trp and His, thus eliminating aliasing and splitting of the His signal. The unbound (W_ε and H_δ) and shifted (W_ε^{*} and H_δ^{*}) side chain tryptophan and histidine signals are indicated. The coloring of signals is the same as in (a,b). (f) Fragment of the PLpro^{CoV-2}:hISG15 structure (PDB id: 7RBS) zooming on the active site of the enzyme to illustrate a close location of the C-terminus of ISG15 (G157) to PLpro^{CoV-2} residues W93, W106, and H272 (shown as spheres). PLpro is colored green, hISG15 is cyan. Carboxylic oxygens of the C-terminal G157 form polar contacts with N_δ of H272 and N_ε of W106. The contacts between G76 of the proximal Ub of K48-Ub₂ and PLpro^{CoV-2} residues W93, W106, and H272 are very similar (PDB id: 7RBS).

b



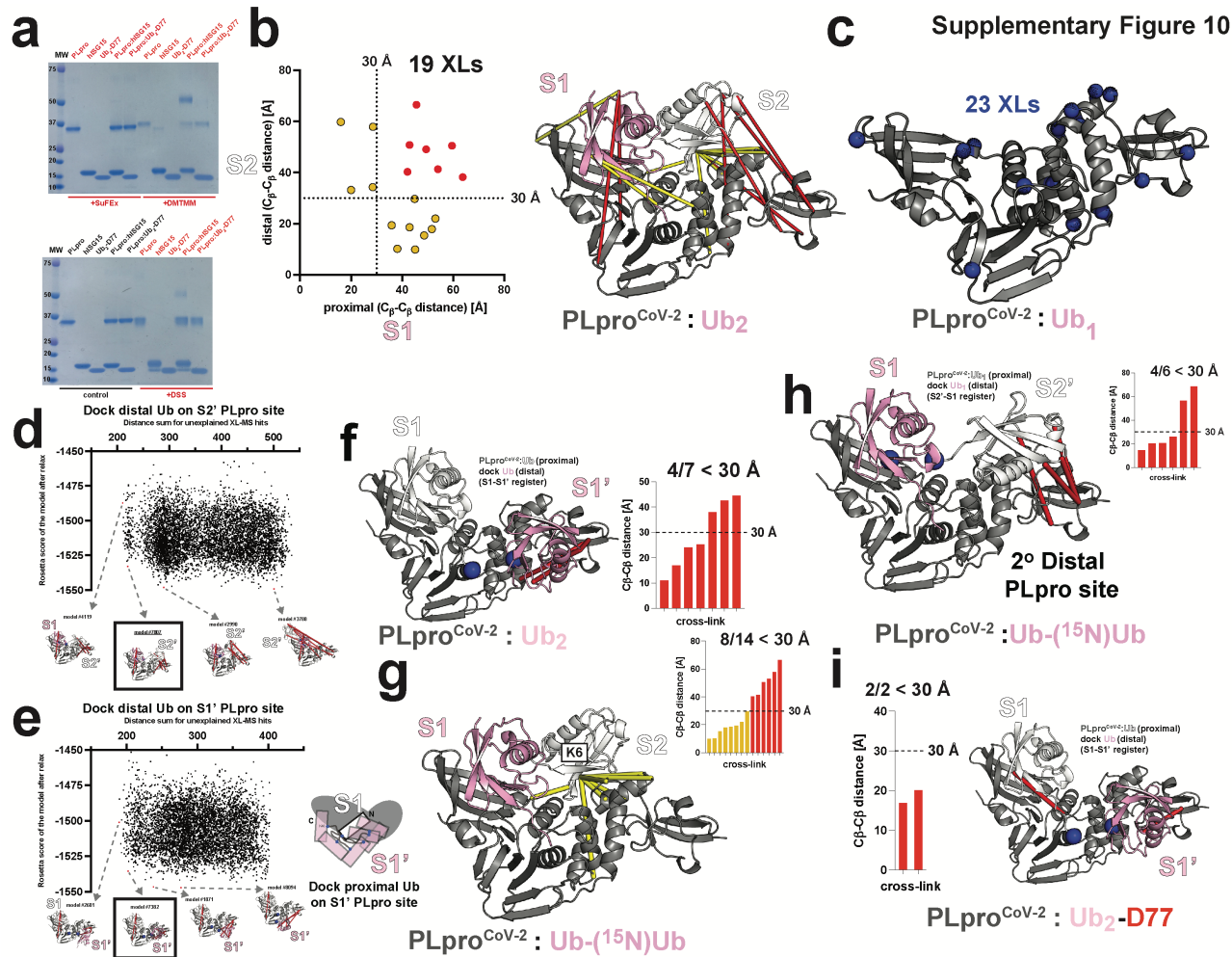
Supplementary Figure 8. Characterization of substrate binding to PLpro^{CoV-2} (C111S) and of substrate competition for binding and cleavage. (a-b) overlay of ¹H-¹⁵N SOFAST-HMQC spectra of Ub₁ (blue) and of the proximal (red) and distal (green) Ubs in K48-Ub₂, free in solution (a) and in complex with PLpro (C111S) (b), see also Figure 3. Examples illustrating the striking difference in the directions of signal shifts in the distal and proximal Ubs of several residues are circled. (c) NMR analysis of the competition between K48-Ub₂ and ISG15 for PLpro binding. Shown is an overlay of ¹H-¹⁵N SOFAST-HMQC spectra of the ¹⁵N-labeled proximal Ub in K48-Ub₂ alone in solution (green), in complex with 1.5 molar equivalents of unlabeled PLpro (C111S) (blue) and upon subsequent addition of 2 molar equivalents of unlabeled ISG15 (red). The spectra show pair-wise comparison. The red contour levels are selected to emphasize the strong signals, for better comparison with the spectra of Ub₂ alone. (d) Intensity of the PLpro-bound signal of G47 (from the binding competition assay as a function of [ISG15]:[Ub₂] molar ratio (dots, normalized to the initial intensity) and the predicted molar fraction of bound Ub₂ (black line) using the K_{d1} values obtained in this work (see Fig. 3j). The dashed red line shows the predicted molar fraction of bound Ub₂ using the K_d values reported by Shin et al. for hISG15 and triazole-linked Ub₂³. These predictions used equations describing competitive binding of two different ligands to the same binding site⁴. (e) SDS-PAGE gels showing the inhibitory effect of ISG15 (middle lanes) and Ub₁ (right lanes) on disassembly of K48-linked Ub₂ by PLpro^{CoV-2}. See also Fig. 3k.

Supplementary Figure 9



Supplementary Figure 9. NMR characterization of the effect of addition of a C-terminal aspartate on interactions of ISG15, K48-Ub₂, and Ub₁ with PLpro^{CoV-2}. (a) Overlay of ¹H-¹⁵N

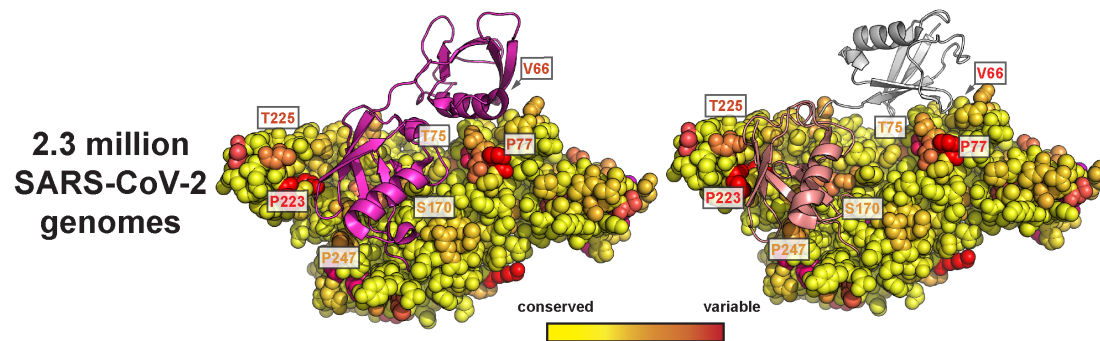
spectra of ^{15}N -labeled distal and proximal Ubs in Ub_2 and Ub_1 alone (blue) and upon addition of $\text{PLpro}^{\text{CoV-2}}$ (C111S). In each pair, the spectrum on the right corresponds to the presence of D77 at the C-terminus of Ub_2 or Ub_1 . PLpro was added in 1:1 (for Ub_2) or 2:1 (for Ub_1) molar ratio for “wild type” proteins and in 2:1 for the D77-extended constructs. Signals of select residues are indicated. Insets zoom on the spectral region containing signals of the hydrophobic patch residues L8, I44, H68, V70, as well as V5. **(b)** Overlay of ^1H - ^{15}N spectra of ^{15}N -labeled ISG15 alone (blue) and upon addition of $\text{PLpro}^{\text{CoV-2}}$ (C111S) (left) or $\text{PLpro}^{\text{CoV-2}}$ (C111S,Y171H) (right). The spectra on the left (TROSY) correspond to WT ISG15, those on the right (SOFAST-HMQC) are for ISG15-D158. The signals corresponding to C-terminal G157 (ISG15) and G157 and D158 (ISG15-D158) are indicated. PLpro was added in ~ 1 molar equivalent to ISG15 and 1.5 molar equivalents to ISG15-D158. **(c)** Overlay of ^1H - ^{15}N TROSY spectra of ^{15}N -labeled $\text{PLpro}^{\text{CoV-2}}$ (C111S,Y171H) alone (blue) and upon addition of K48- Ub_2 (left) or K48- Ub_2 -D77 (right). The signals corresponding to side chain NHs of W93, W106, and H272 are marked as W_ϵ and H_δ for the unbound signals and W_ϵ^* and H_δ^* for the (shifted) bound signals. **(d)** Overlay of ^1H - ^{15}N TROSY spectra of ^{15}N -labeled $\text{PLpro}^{\text{CoV-2}}$ (C111S,Y171H) alone (blue) and upon addition of 8 molar equivalents of WT Ub (left) or Ub-D77 (right). The signals corresponding to side chain NHs of W93, W106, and (likely) H272 are marked as W_ϵ and H_δ for the unbound signals and W_ϵ^* and H_δ^* for the (shifted) bound signals.



Supplementary Figure 10. Intermolecular cross-links from XL-MS analysis of PLpro^{CoV-2} complexes. (a) Cross-linked PLpro^{CoV-2}:hISG15 and PLpro^{CoV-2}:Ub₂-D77 complexes reveal formation of covalent heterodimer complex bands (red) compared to control reactions (untreated and PLpro^{CoV-2} or substrate alone) by SDS-PAGE. (b) Intermolecular cross-links sites (19 pairs) mapped onto proximal or distal Ub binding modes on PLpro^{CoV-2}. Cross-link pair distances between C_β-C_β below 30 Å or above 30 Å are colored in yellow and red, respectively. Calculated proximal and distal distances for cross-link pairs are plotted on a scatter plot. Cross-links are mapped onto structure of PLpro:Ub₂ and colored yellow or red based on distance. Structures are shown in cartoon. Cross-links are shown as lines. (c) Cross-links derived from PLpro^{CoV-2}:Ub₁ complexes (23 pairs). Mapping Ub contact points on PLpro^{CoV-2} derived from heterotypic cross-links pairs. PLpro^{CoV-2} is shown as cartoon (grey), cross-link sites are shown as spheres (blue). (d,e) Docking Ub onto PLpro^{CoV-2}:Ub_{proximal} (S1) to identify alternate S2 site, termed S2' (d) and S1' (e) sites employing a geometric restraint that links K48 from proximal Ub to C-terminus of the distal Ub but with two different binding registers: (1)

proximal Ub in S1 binding site and distal sampling alternative S2' sites (d) or (2) distal Ub in S1 with the proximal Ub sampling alternative S1' sites (e). Energy of complex for over 5000 models was compared to the sum of distances for 7 unexplained cross-links from (b) (in red). Representative structures that sample diverse geometries of the docked Ub that are compatible (left) or incompatible (right) with cross-link distances. Structures are shown in cartoon and colored as before. Ub domains are shown as cartoon on S2'/S1 (d) or S1/S1' binding sites. Cross-links are shown as red lines. (e, right) Schematic for illustrating docking scheme, Ub docked onto PLpro^{CoV-2}:Ub into the S1' site. (f) Seven previously unexplained heterotypic cross-links identified from a PLpro^{CoV-2}:Ub complex. Cross-links between Ub and PLpro^{CoV-2}'s thumb and UBL domains are mapped onto a docked model of PLpro^{CoV-2}:Ub with Ub placed in the S1' site based on the K48-Cterm linkage. Structures are shown as cartoons and colored as previously. The cross-links pairs shown as red lines. Bar plot of calculated distances for 7 cross-links and colored in red. A dashed line denotes the 30 Å threshold. (g) Fourteen heterotypic cross-links identified from a PLpro^{CoV-2}:Ub-(¹⁵N)Ub complex between the protease and the light Ub peptides. Cross-links are mapped onto a model of PLpro^{CoV-2} with the canonical (i.e., as found in the crystal structure) Ub₂ binding mode in S1/S2 binding sites. Structures are shown as cartoons and colored as previously. The cross-links pairs shown as yellow lines. Bar plot of calculated distances for 14 cross-links and colored in yellow (< 30 Å threshold) red (> 30 Å threshold). A dashed line denotes the 30 Å threshold. (h) Six previously unexplained heterotypic cross-links identified from (g, red). Cross-links are mapped onto a docked model of PLpro^{CoV-2}:Ub with Ub placed in the S2' based K48-Cterm linkage geometry. Structures are shown as cartoons and colored as previously. The cross-links pairs shown as red lines. Bar plot of calculated distances for 6 cross-links and colored in red. A dashed line denotes the 30 Å threshold. (i) Two heterotypic cross-links identified from the PLpro^{CoV-2}:Ub₂-D77 complex. Cross-links are mapped onto a docked model of PLpro^{CoV-2}:Ub with Ub placed in the S1' based K48-Cterm linkage geometry. Structures are shown as cartoons and colored as previously. The cross-links pairs shown as red lines. Bar plot of calculated distances for 2 cross-links and colored in red. A dashed line denotes the 30 Å threshold.

Supplementary Figure 11



Supplementary Figure 11. Comparison of ISG15 and PLpro across zoonotic species and humans. Sequence variation across 2.3 million PLpro^{CoV-2} isolates reveals variation in both the proximal and distal binding sites. PLpro is shown in space fill representation and is color-coded according to conservation from yellow-red. Sites of variation are labeled. K48-linked Ub₂ and human ISG15 binding modes are shown for comparison and colored as in Supplementary Fig. 1.

SUPPLEMENTARY MATERIALS AND METHODS

Gene cloning, protein expression and purification of WT and C111S mutant of PLpro

The gene cloning, protein expression and purification were performed using protocols published previously³⁴. Briefly, the Nsp3 DNA sequence corresponding to PLpro protease of SARS-CoV-2 was optimized for *E. coli* expression using the OptimumGene codon optimization algorithm followed by manual editing and then cloned directly into pMCSG53 vector (Twist Bioscience). Human ISG15 gene was also synthesised and cloned directly into pMCSG53 vector. The plasmids were transformed into the *E. coli* BL21(DE3)-Gold strain (Stratagene). *E. coli* cells harboring plasmids for SARS-CoV-2 PLpro WT and C111S mutant and ISG15 expression were cultured in LB medium supplemented with ampicillin (150 µg/ml).

For large-scale purification of PLpro^{CoV-2}-C111S, 4 L cultures of LB Lennox medium were grown at 37 °C (200 rpm) in the presence of ampicillin 150 µg/ml. Once the cultures reached OD₆₀₀ ~1.0, the temperature setting was changed to 4 °C. When the bacterial suspensions cooled down to 18 °C they were supplemented with 0.5 mM IPTG and 40 mM K₂HPO₄ (final concentration). The temperature was set to 18 °C for 20 hours incubation. Bacterial cells were harvested by centrifugation at 7,000g and cell pellets were resuspended in a 12.5 ml lysis buffer (500 mM NaCl, 5% (v/v) glycerol, 50 mM HEPES pH 8.0, 20 mM imidazole pH 8.0, 10 mM β-mercaptoethanol, 1 µM ZnCl₂) per liter culture and sonicated at 120W for 5 minutes (4 sec ON, 20 sec OFF). The cellular debris was removed by centrifugation at 30,000g for 90 minutes at 4 °C. The supernatant was mixed with 3 ml of Ni²⁺ Sepharose (GE Healthcare Life Sciences) which had been equilibrated with lysis buffer supplemented to 50 mM imidazole pH 8.0, and the suspension was applied on Flex-Column (420400-2510) connected to Vac-Man vacuum manifold (Promega). Unbound proteins were washed out via controlled suction with 160 ml of lysis buffer (with 50 mM imidazole pH 8.0). Bound proteins were eluted with 15 ml of lysis buffer supplemented to 500 mM imidazole pH 8.0, followed by Tobacco Etch Virus (TEV) protease treatment at 1:25 protease:protein ratio. The solutions were left at 4 °C overnight. Size exclusion chromatography was performed on a Superdex 75 column equilibrated in lysis buffer. Fractions containing cut protein were collected and applied on a Flex-Column with 3 ml of Ni²⁺ Sepharose which had been equilibrated with lysis buffer. The flow through and a 7 ml lysis buffer rinse were collected. Lysis buffer was replaced using a 30 kDa MWCO filter (Amicon-Millipore) via 10X concentration/dilution repeated 3 times to crystallization buffer (20 mM HEPES

pH 7.5, 150 mM NaCl, 1 μ M ZnCl₂, 10 mM DTT). The final concentration of WT PLpro^{CoV-2} was 25 mg/ml and C111S mutant was 30 mg/ml. For some NMR studies requiring longer measurement times we also utilized C111S/Y171H PLpro^{CoV-2} variant which showed similar binding properties to C111S variant but was more stable in the NMR buffer.

Expression and purification of unlabeled and isotope labeled ISG15 and Ub

ISG15 was purified following the same protocol as for PLpro except that the buffers did not contain ZnCl₂ and a 10 kDa MWCO filter was used for buffer exchange and concentration. The N-terminal polyhistidine tag was removed using TEV protease which left an additional serine/alanine at the N-terminus followed by an additional Ni-NTA step. The final concentration of ISG15 was 40 mg/ml. Unlabeled Ub variants were expressed in BL12(DE3) *E.coli* cells containing a helper pJY2 plasmid and purified as described elsewhere⁵. For expression of uniformly ¹⁵N-labeled Ubs, ISG15, as well as PLpro^{CoV-2} variants the cells were grown in minimal media containing ¹⁵NH₄Cl as the sole source of nitrogen using methods previously described in Varadan et al⁵⁻⁷.

Synthesis of K48-polyUb chains

Ub₂ and Ub₃ chains were assembled from the respective recombinant Ub monomers using controlled chain synthesis catalyzed by Ub-activating E1 enzyme UBA1 and K48-specific Ub-conjugating E2 enzyme UBE2K (aka E2-25K) as detailed elsewhere⁵⁻⁷. Specifically, Ub variants bearing chain-terminating mutations, Ub K48R and Ub D77, were used to ensure that only Ub dimers are produced and to enable incorporation of ¹⁵N labeled Ub units at the desired distal (¹⁵N-Ub K48R) or proximal (¹⁵N-Ub D77) position in the resulting chain for NMR and MS studies. D77 was subsequently removed from the proximal Ub by Ub C-terminal hydrolase YUH1. Ub₃ chains were made in a stepwise manner, by first removing D77 from the proximal Ub in Ub₂ by YUH1 and subsequently conjugating this Ub₂ to Ub D77 using E1 and E2-25K to produce Ub₃. The Ub₃ chain for MS-based cleavage assays had Ub K48R at the distal unit, ¹⁵N-labeled Ub at the endo unit, and Ub D77 at the proximal unit, in order to allow unambiguous identification of the possible cleavage products by mass spectrometry. The correct masses of the synthesized Ub₂ and Ub₃ chains were confirmed using ESI-MS and SDS-PAGE.

Microscale Thermophoresis binding assay

MST experiments were performed using NanoTemper Monolith NT.115 available in the Macromolecular Biophysics Resource core at UTSW and standard protocol was employed during analysis⁸. Ub₁, K48-Ub₂ or hISG15 were labeled with Cyanine5 NHS ester dye (Cy5) and titrated by a 1:1 serial dilution of PLpro^{CoV-2} (C111S mutant). Obtained data were fit and analyzed in PALMIST 1.5.8 using 1:1 and 1:2 binding models and visualized in GUSI 1.4.2^{8,9}. All measurements were done in technical triplicates. To determine whether the 1:1 or 1:2 binding models were more suitable we calculated the probability of getting the χ^2 improvement by chance using a t-test. The binding stoichiometry of the complexes was verified using cross-linking of complexes and visualized by SDS-PAGE (see below). PALMIST and GUSI software is freely available for academic users on UTSW MBR core's website (<https://www.utsouthwestern.edu/labs/mbr/software/>).

PLpro cleavage assay

K48-Ub₃, K48-Ub₂, K48-Ub₂-D77 and ISG15-Nsp2, ISG15-Nsp3 and ISG15-Nsp4 cleavage reactions were performed at 20 °C in 20 mM Tris buffer (pH 7.52) containing 100 mM NaCl, 10 mM DTT, 1 μ M ZnCl₂. The initial volume was 350 μ l and contained 20 μ M of Ub₃. Upon addition of 0.5 μ M PLpro, equal amounts (10 μ l) of reaction samples were aliquoted out at given time points, mixed with equal volume of SDS load buffer and immediately placed in a water bath at 70 °C for 5 minutes to stop the reaction. The samples were then loaded onto 15% urea polyacrylamide gel and resolved using SDS-PAGE. For mass spectrometry analyses cleavage reactions were performed on ice, the samples were buffer exchanged into autoclaved RO water and concentrated to 50 μ l volume prior to analysis. The data shown here were obtained on Bruker Maxis-II ultra-high resolution Q-TOF mass spectrometer available at University of Maryland. Peaks were isolated during separation and analyzed using MagTran 1.03 b2.

Protein crystallization

Crystallizations were performed with the protein-to-matrix ratio of 1:1 using the sitting drop vapor-diffusion method with the help of the Mosquito liquid dispenser (TTP LabTech) in 96-well CrystalQuick plates (Greiner Bio-One). MCSG1, MCSG2, MCSG3, and MCSG4 (Anatrace), Index (Hampton Research) and Wizard 1&2 (Jena Bioscience) screens were used at 16 °C. The PLpro^{CoV-2}-C111S/ISG15 complex (13 mg/ml) crystallized in Index E11 (0.02 M MgCl₂, 0.1 M HEPES pH 7.5, 22% (w/v) polyacrylic acid sodium salt 5,100). For the PLpro^{CoV-2}-C111S/Ub₂-K48R,D77

complex (11 mg/ml), crystals appeared in MCSG-2 F11 and were improved in hanging drops with a protein-to-matrix ratio of 2:1 in 0.2 M disodium tartrate, 15 % (w/v) PEG3350 after seeding with 1/10 volume of PLpro^{CoV-2}-C111S/Ub₂-K48R, D77 microcrystals. The Ub₂ protein crystallized in Pi-PEG D1 (50 mM acetate buffer pH 4.8, 8.6% PEG 2000 MME, 17.1% PEG 400). The ISG15 protein crystallized in MCSG1 G2 (40 mM potassium phosphate, 16% PEG 8000, 20% glycerol). Crystals selected for data collection were washed in the crystallization buffer supplemented with 25% glycerol and flash-cooled in liquid nitrogen.

Data collection, structure determination, and refinement

Single-wavelength X-ray diffraction data were collected at 100 K temperature at the 19-ID beamline of the Structural Biology Center at the Advanced Photon Source at Argonne National Laboratory using the program SBCcollect. The diffraction images were recorded on the PILATUS3 X 6M detector at 12.662 keV energy (0.9792 Å wavelength) using 0.3° rotation and 0.3 sec exposure. The intensities were integrated and scaled with the HKL3000 suite¹⁰. Intensities were converted to structure factor amplitudes in the truncate program¹ from the CCP4 package¹¹. The structures were determined by molecular replacement using HKL3000 suite incorporating the program MOLREP¹²⁻¹⁴. The coordinates of PLpro^{CoV-2} in complex with ubiquitin propargylamide (PDB id: 6XAA) and PLpro^{CoV-2}-C111S with mISG15 (PDB id: 6YVA) were used as the starting models for PLpro^{CoV-2}-K48-Ub₂ and PLpro-ISG15 structure solutions, respectively. For Ub₂ and ISG15 proteins, the structures of PLpro^{CoV-1} bound to a K48-Ub₂ activity based probe (PDB id: 5E6J) and ISG15 (PDB id: 1Z2M) were used as the starting models. The initial solutions were refined, both rigid-body refinement and regular restrained refinement by REFMAC program¹⁵ as a part of HKL3000. Several rounds of manual adjustments of structure models using COOT¹⁶ and refinements with REFMAC program¹⁵ from CCP4 suite¹¹ were done. The stereochemistry of the structure was validated with PHENIX suite¹⁷ incorporating MOLPROBITY² tools. A summary of data collection and refinement statistics is given in Supplementary Table 1.

NMR data collection and analysis

NMR measurements were performed at 25 °C on Avance III 600 MHz and 800 MHz Bruker NMR spectrometers equipped with cryoprobes. The data were processed using Topspin (Bruker) and analyzed using Sparky 3.190¹⁸. NMR signal assignments for hISG15 were obtained from Biological

Magnetic Resonance Data Bank (BMRB Entry ID 5658) and adjusted to the temperature and buffer conditions used in our studies.

The protein samples for NMR measurements were prepared in 50 mM HEPES buffer or in 20 mM Tris buffer, both at pH 7.42 and containing 100 mM NaCl, 1 mM TCEP, 1 μ M ZnCl₂, 0.2% (w/v) NaN₃, and 10% (v/v) D₂O. Binding studies by NMR were carried out by adding pre-calculated amounts of unlabeled PLpro^{CoV-2} (C111S) to ¹⁵N-labeled hISG15 or K48-Ub₂ (with either distal or proximal Ub ¹⁵N labeled) up to ~1:1 molar ratio, or 2:1 molar ratio to Ub₁ and monitoring changes in 2D ¹H-¹⁵N SOFAST-HMQC and/or ¹H-¹⁵N TROSY as well as 1D ¹H spectra. The initial binding studies were performed in HEPES buffer and subsequently repeated in Tris buffer, both produced similar results. Reciprocal titrations were performed by adding unlabeled Ub₂, Ub₁, or ISG15 (or D158-extended) to ¹⁵N-labeled PLpro^{CoV-2} (C111S) or PLpro^{CoV-2} (C111S,Y171H) in a 1.5:1 (Ub₂, ISG15:PLpro) or up to 8:1 (Ub₁:PLpro) molar ratio. ¹⁵N-labeled PLpro^{CoV-2} (C111S,Y171H) was primarily used for lengthy TROSY experiments as this variant proved to be more stable than PLpro^{CoV-2} (C111S,Y171H) in the NMR buffer. Both PLpro^{CoV-2} variants had very similar NMR spectra and essentially identical signal perturbations upon binding of all the substrates tested (see Supplementary Fig. 7).

The protein concentrations in NMR studies in Tris buffer were as follows: 150:150 μ M for ISG15:PLpro^{CoV-2}, 180:270 μ M for ISG15-D158:PLpro^{CoV-2}, 83:104 μ M for ¹⁵N-distal Ub₂:PLpro, 200:200 μ M for ¹⁵N-proximal Ub₂:PLpro^{CoV-2}, 152.8:305.9 μ M for Ub₁:PLpro^{CoV-2}, 125:250 μ M for Ub₁-D77:PLpro^{CoV-2}. Experiments with ¹⁵N-PLpro^{CoV-2} used 301.5:201.5 μ M ISG15:PLpro, 225:150 μ M ISG15-D158:PLpro, 137.5:110 μ M Ub₂:PLpro, 404:202 μ M Ub₂-D77:PLpro, up to 1140:141 μ M Ub₁:PLpro, 1000:125 μ M Ub₁-D77:PLpro. For NMR measurements in HEPES buffer, the concentrations were 115:115 μ M for ISG15: PLpro^{CoV-2}; 71:142 μ M for ¹⁵N-distal Ub₂: PLpro^{CoV-2}; 75:150 μ M for ¹⁵N-proximal Ub₂: PLpro^{CoV-2}; and 81:151 μ M for Ub₁:PLpro^{CoV-2}.

Cross-linking mass spectrometry analysis

Our group has developed standardized protocols for cross-linking and data analysis of samples. For complexes between PLpro^{CoV-2} and Ub₁, Ub₁-D77, K48-Ub₂, K48-Ub₂-D77 or hISG15, we incubated the protease with the substrate at a 1:4 molar ratio for 1 hour at 25 °C. For disuccinimidyl suberate (DSS) reactions, samples were cross-linked with a final 1 mM DSS (DSS-d₀ and -d₁₂, Creative Molecules) for three minutes at 37 °C while shaking at 350 rpm. For sulfonyl fluoride

(SuFEx) reactions, samples were cross-linked with a final 0.658 mM SuFEx (a kind gift from Dr. William DeGrado, UCSF) for one hour at 37 °C while shaking at 350 rpm. For DMTMM reactions, samples were cross-linked with a final 43 mM DMTMM (Sigma-Aldrich) for 15 minutes at 37 °C while shaking at 600 rpm. All cross-linking reactions were quenched with 172 mM (4 times excess) ammonium bicarbonate for 30 minutes at 37 °C while shaking at 350 rpm. Samples were resolved on SDS-PAGE gels (NuPAGE™, 4 to 12%, Bis-Tris, 1.5 mm) and, for DMTMM cross-linker set of samples, bands corresponding to PLpro^{CoV-2}:substrate heterodimers were gel-extracted following standard protocols¹⁹. For glutaraldehyde cross-linking, we used 1:4, 1:1 and 1:0.25 ratios of PLpro:substrate using 6 µM protease in all reactions. Samples were preincubated at 25 °C for 15 minutes followed by addition of 0.05% glutaraldehyde (Sigma) for 1 minute and quenched with Tris pH 8 to a final concentration of 0.2 M. Cross-linked species and control reactions were resolved by SDS-PAGE. Cross-linked samples for XL-MS were digested by 1:50 (m/m) trypsin (Promega) overnight at 37 °C while shaking at 600 rpm. 2% (v/v) formic acid was added to acidify the reaction and further purified by reversed-phase Sep-Pak tC18 cartridges (Waters), next flash frozen in liquid nitrogen and lyophilized. The dried samples were resuspended in water/acetonitrile/formic acid (95:5:0.1, v/v/v) to a final concentration of approximately 0.5 µg/µl. 2 µl of each was injected into Eksigent 1D-NanoLC-Ultra HPLC system coupled to a Thermo Orbitrap Fusion Tribrid system at the UTSW Proteomics core.

The mass spectrometry data were analyzed by in-house version of xQuest 2.1.5 pipeline²⁰. Thermo RAW data files were first converted to open .mzXML format using msconvert (proteowizard.sourceforge.net). The mass spectra across replicates yielded similar intensities. Search parameters were set based on DMTMM as the cross-link reagent as follows: maximum number of missed cleavages = 2, peptide length = 5-50 residues, fixed modifications = carbamidomethyl-Cys (mass shift = 57.02146 Da), variable modification = oxidation of methionine (mass shift = 15.99491 Da), mass shift of cross-linker = -18.010595 Da, no monolink mass specified, MS1 tolerance = 15 ppm, and MS2 tolerance = 0.2 Da for common ions and 0.3 Da for cross-link ions; search in enumeration mode. Next, in-house shell script was employed to identify cross-links between lysines and acidic residues. FDRs were estimated by xProphet²¹ to be 9.8% - 77.8%. For each experiment, five replicate data sets were compared and only cross-link pairs that appeared in all data sets (PLpro^{CoV-2}:hISG15, PLpro^{CoV-2}:K48-Ub₂ and PLpro^{CoV-2}:Ub₁) or at least in four data sets (PLpro^{CoV-2}:K48-Ub₂-D77) were used to generate a consensus data set (Source Data 1).

Modeling of alternate Ub binding sites on PLpro^{CoV-2}

To build an ensemble of alternate Ub binding sites on PLpro^{CoV-2} outside of the canonical proximal domain on the S1 site, we employed docking procedure that combined a geometric restraint between K48 of the immobile proximal Ub to the C-terminus of the mobile Ub to sample alternate S2 binding sites. In the alternative scenario, we employed a geometric restraint between C-terminus of the immobile proximal Ub to K48 of the mobile Ub to sample alternate S1' binding sites below the active sites. An initial conformation of the PLpro^{CoV-2} bound to Ub in the proximal site was built from our structure (PDB id: 7RBR) and converted into a single chain. We next added a mobile Ub as a second chain and produced over 5000 low-resolution centroid mode models employing two different geometric restraints that sampled alternate S2 and S1' binding sites. Each structure was minimized and the total energy of the PLpro^{CoV-2}:Ub₂ complexes was plotted as a function of a sum distances of C_β-C_β from the experimental crosslinks using an in-house script. All simulations were performed with RosettaDock protocol ²² as a part of Rosetta 3.13 suite and ran on UTSW's BioHPC computing cluster. All plots were generated with GraphPad Prism 9. Images were created using PyMOL 2.5.1.

Energetic analysis of PLpro^{CoV-2} in complex with hISG15 and K48-Ub₂

Models of PLpro^{CoV-2} bound two different substrates, hISG15 and K48-Ub₂, that were used in the subsequent *in silico* alanine scan were prepared as follows. For the PLpro^{CoV-2}:hISG15 complex, we used a heterodimer conformation derived from our crystal structure (PDB id: 7RBS). To create the complex between PLpro^{CoV-2} and K48-Ub₂, we used the conformation and binding mode of K48-Ub₂ bound to PLpro^{CoV-1} (PDB id: 5E6J) as a template²³. Briefly, PLpro^{CoV-2} was aligned to PLpro^{CoV-1} bound to K48-Ub₂ to produce a tentative model of PLpro^{CoV-2}:K48-Ub₂. Next, we applied a relax protocol in Rosetta for both complexes: PLpro^{CoV-2}: K48-Ub₂ and PLpro^{CoV-2}:hISG15. To guarantee that each instance of relax is being run with different randomizations, groups of nstruct were run with different, randomly generated seeds using random.org. From 100 total structures (4x25 nstruct for computational efficiency) for each heterocomplex, the lowest energy structure was identified and used in further steps. The list of PLpro residues that may be engaged in interacting with its substrate was created by identifying PLpro residues within 4.0 Å of either hISG15 or K48-Ub₂. The union list of interacting residues identified with heterocomplexes PLpro^{CoV-2}:hISG15 and PLpro^{CoV-2}: K48-Ub₂ was used in the next step. For PLpro^{CoV-2} in complex with hISG15 or K48-Ub₂ 49 positions were

used to describe the combined interface. Selection of common interface residues was carried out in PyMOL 2.5.1. Flex $\Delta\Delta G$ protocol was used as described previously²⁴. The code is available on GitHub: <https://github.com/Kortemme-Lab>. Briefly, selected interacting residues were mutated to alanines. Parameters were used (all default settings): nstruct = 35, max_minimization_iter = 5000, abs_score_convergence_thresh = 1.0, number_backrub_trials = 35000, and to enable earlier time points backrub_trajectory_stride was set to 7000. The $\Delta\Delta G_{\text{binding}}$ score for the last iteration is shown in the Results. These simulations were performed using BioHPC computing cluster at UT Southwestern Medical Center. The results, in raw REU (Rosetta energy units), are shown as a heatmap with $\Delta\Delta G_{\text{binding}}$ values but also as the difference between the $\Delta\Delta G_{\text{binding}}$ from hISG15 compared to K48-Ub₂. The plots were made using GraphPad Prism 9 and mapped onto the protease structure using PyMOL 2.5.1. The relax protocol and Flex ddG used Rosetta v3.13 and v3.12, respectively.

Sequence comparison of Ub, ISG15 and sequence variation across PLpro^{CoV-2} in SARS-CoV-2

Alignments were produced in Clustalo²⁵ and visualized in Seaview²⁶. Sequence identity between Ub₂, hISG15 and mISG15 was calculated using Blast²⁷. PLpro^{CoV-2} sequence variation from 2.3 million sequences (as of October 18, 2021) was derived from the coronavirus3D database²⁸. The per residue mutational frequencies were mapped onto a PLpro^{CoV-2} structure in the context of a bound K48-Ub₂ or hISG15.

Supplementary References

1. Padilla, J. E. & Yeates, T. O. A statistic for local intensity differences: robustness to anisotropy and pseudo-centering and utility for detecting twinning. *Acta Crystallogr D Biol Crystallogr* **59**, 1124–1130 (2003).
2. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12–21 (2010).
3. Shin, D. *et al.* Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature* **587**, 657–662 (2020).
4. Wang, Z. X. An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule. *FEBS Letters* **360**, 111–114 (1995).
5. Pickart, C. M. & Raasi, S. Controlled synthesis of polyubiquitin chains. *Meth. Enzymol.* **399**, 21–36 (2005).
6. Varadan, R., Assfalg, M., Raasi, S., Pickart, C. & Fushman, D. Structural determinants for selective recognition of a Lys48-linked polyubiquitin chain by a UBA domain. *Molecular Cell* **18**, 687–698 (2005).
7. Varadan, R., Walker, O., Pickart, C. & Fushman, D. Structural properties of polyubiquitin chains in solution. *Journal of Molecular Biology* **324**, 637–647 (2002).
8. Scheuermann, T. H., Padrick, S. B., Gardner, K. H. & Brautigam, C. A. On the acquisition and analysis of microscale thermophoresis data. *Analytical Biochemistry* **496**, 79–93 (2016).
9. Tso, S.-C. *et al.* Using two-site binding models to analyze microscale thermophoresis data. *Analytical Biochemistry* **540-541**, 64–75 (2018).
10. Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr* **62**, 859–866 (2006).
11. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235–242 (2011).
12. Morris, R. J., Perrakis, A. & Lamzin, V. S. ARP/wARP and automatic interpretation of protein electron density maps. *Meth. Enzymol.* **374**, 229–244 (2003).
13. Terwilliger, T. SOLVE and RESOLVE: automated structure solution, density modification and model building. *J Synchrotron Radiat* **11**, 49–52 (2004).
14. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**, 22–25 (2010).
15. Joosten, R. P., Long, F., Murshudov, G. N. & Perrakis, A. The PDB_REDO server for macromolecular structure model optimization. *IUCrJ* **1**, 213–220 (2014).
16. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132 (2004).
17. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–221 (2010).
18. Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).
19. Chen, D. *et al.* Tau local structure shields an amyloid-forming motif and controls aggregation propensity. *Nature Communications* **10**, 2493–14 (2019).
20. Rinner, O. *et al.* Identification of cross-linked peptides from large sequence databases. *Nat Meth* **5**, 315–318 (2008).

21. Walzthoeni, T. *et al.* False discovery rate estimation for cross-linked peptides identified by mass spectrometry. *Nat Meth* **9**, 901–903 (2012).
22. Wang, C., Bradley, P. & Baker, D. Protein-protein docking with backbone flexibility. *Journal of Molecular Biology* **373**, 503–519 (2007).
23. Békés, M. *et al.* Recognition of Lys48-Linked Di-ubiquitin and Deubiquitinating Activities of the SARS Coronavirus Papain-like Protease. *Mol Cell* **62**, 572–585 (2016).
24. Barlow, K. A. *et al.* Flex ddG: Rosetta Ensemble-Based Estimation of Changes in Protein-Protein Binding Affinity upon Mutation. *J Phys Chem B* **122**, 5389–5399 (2018).
25. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).
26. Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**, 221–224 (2010).
27. Tatusova, T. A. & Madden, T. L. BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol Lett* **174**, 247–250 (1999).
28. Sedova, M., Jaroszewski, L., Alisoltani, A. & Godzik, A. Coronavirus3D: 3D structural visualization of COVID-19 genomic divergence. *Bioinformatics* **36**, 4360–4362 (2020).