SUPPORTING INFORMATION

Identification of human host substrates of the SARS-CoV-2 M^{pro} and PL^{pro} using subtiligase N-terminomics

Shu Y. Luo[1], Eman W. Moussa[1], Joaquin Lopez-Orozco[2], Alberto Felix-Lopez[3], Ray Ishida[3], Nawell Fayad[2], Erik Gomez-Cardona[1], Henry Wang[1], Joyce A Wilson[4], Anil Kumar[4], Tom C. Hobman[2,3,5], Olivier Julien[1,5]*

[1]Department of Biochemistry, University of Alberta, Edmonton, AB, T6G 2H7, Canada

[2]Department of Cell Biology, University of Alberta, Edmonton, AB, T6G 2H7, Canada

[3]Department of Medical Microbiology & Immunology, University of Alberta, Edmonton, AB, T6G 2H7, Canada

[4]Department of Biochemistry, Microbiology & Immunology, University of Saskatchewan, Saskatoon, SK, S7N 5E5, Canada

[5]Li Ka Shing Institute of Virology, Edmonton, AB, T6G 2E1, Canada

Please address correspondence to: *ojulien@ualberta.ca

Table of Content

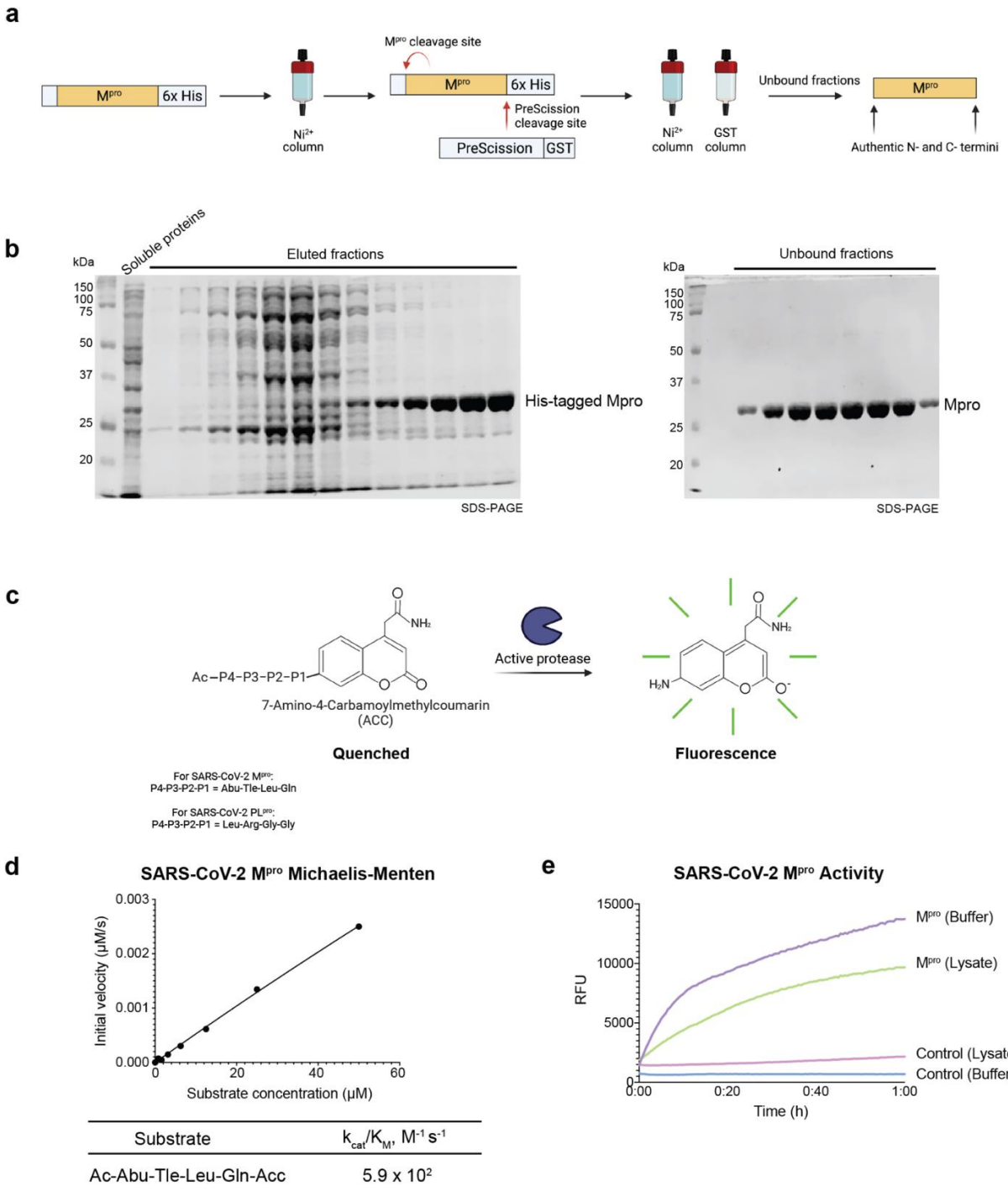

Figure S1. Plasmid construct, protein expression and purification of SARS-CoV-2 M^{pro}. a,b) SARS-CoV-2 M^{pro} was expressed and purified by affinity purification with authentic N- and C-termini. **c**) The fluorescence activity assay was carried out using the optimal coumarin substrate Ac-Abu-Tle-Leu-Gln-ACC. **d**) The enzyme kinetics assay was performed using 0.09 μM M^{pro}, and 0.78 to 50 μM coumarin substrate in 100 μL total assay volume. The k_{cat}/K_M was calculated using the linear region of the Michaelis-Menten curve k_{cat}/K_M = slope / [E], (1) and is consistent with previously reported value. (2) **e**) The protease activity was monitored in parallel with the reverse N-terminomics, in cell-free conditions and cell lysates with 0.5 μM M^{pro} and 2 μM coumarin substrate, showing that it was proteolytically active.

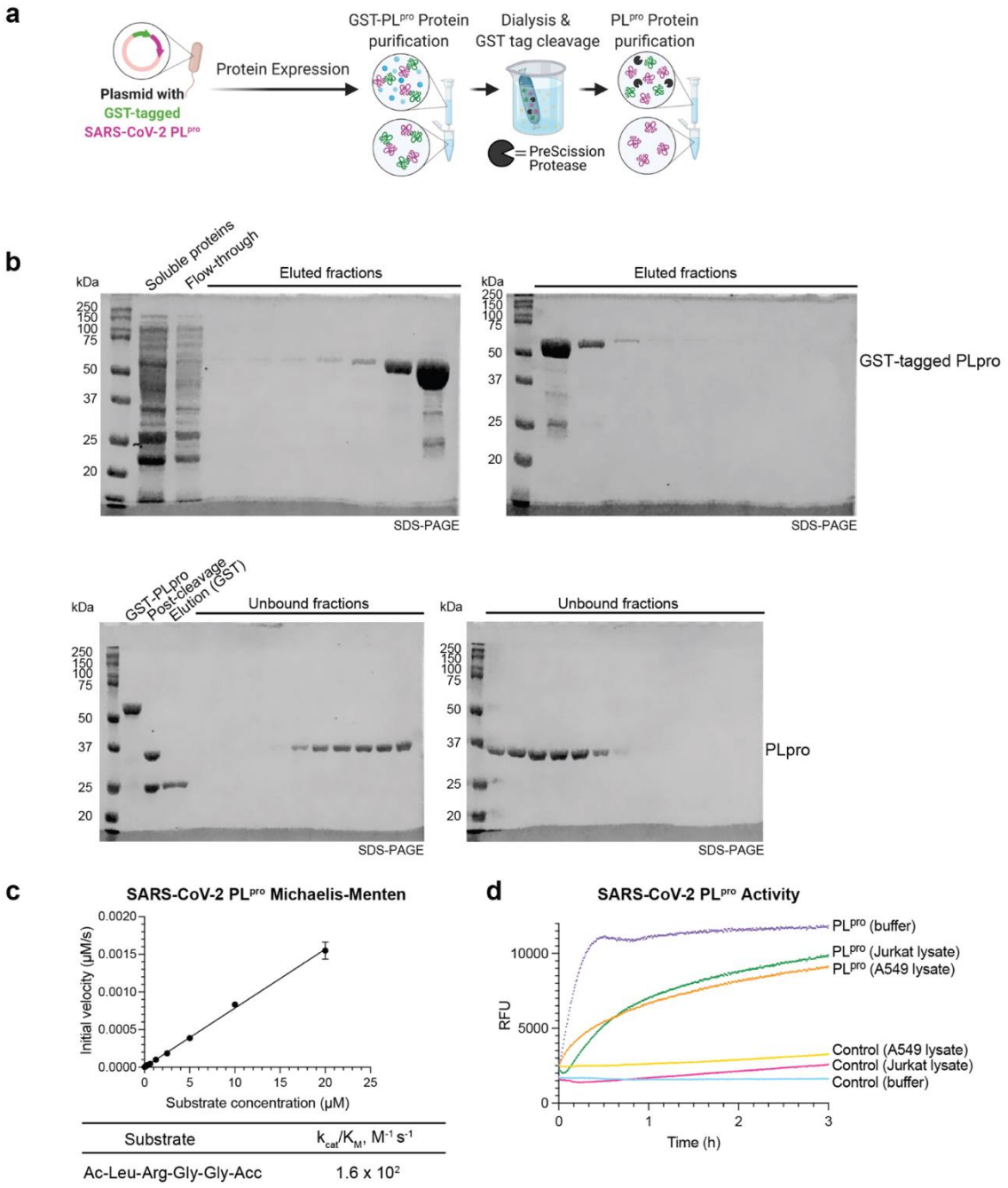

Figure S2. Protein expression, purification and activity of PL^{pro}. a) PL^{pro} expression and purification was conducted using a plasmid encoding for the protease domain of Nsp3 with an N-terminal GST tag. GST-PL^{pro} is purified from *E. coli* lysates using a glutathione sepharose column. The GST tag is removed in an overnight dialysis using a PreScission protease. **b)** Prior to removal of the GST tag, GST-PL^{pro} is collected in the elution fractions. Following GST tag removal, the PreScission protease and GST-tag remain bound to the column and collected in the second elution while PL^{pro} is obtained in the unbound fractions. **c**) The fluorescence activity assay was carried out using the optimal coumarin substrate Ac-Leu-Arg-Gly-Gly-ACC (see **Fig. S1c**). The enzyme kinetics assay was performed using 0.5 μM PL^{pro}, and 0.78 to 20 μM coumarin substrate in 100 μL total assay volume. The k_{cat}/K_M was calculated using the linear region of the Michaelis-Menten curve k_{cat}/K_M = slope / [E]. (1) **d)** The activity of 5 μM PL^{pro} was measured using 10 μM of the coumarin substrate in buffer, A549 and Jurkat cell lysates.

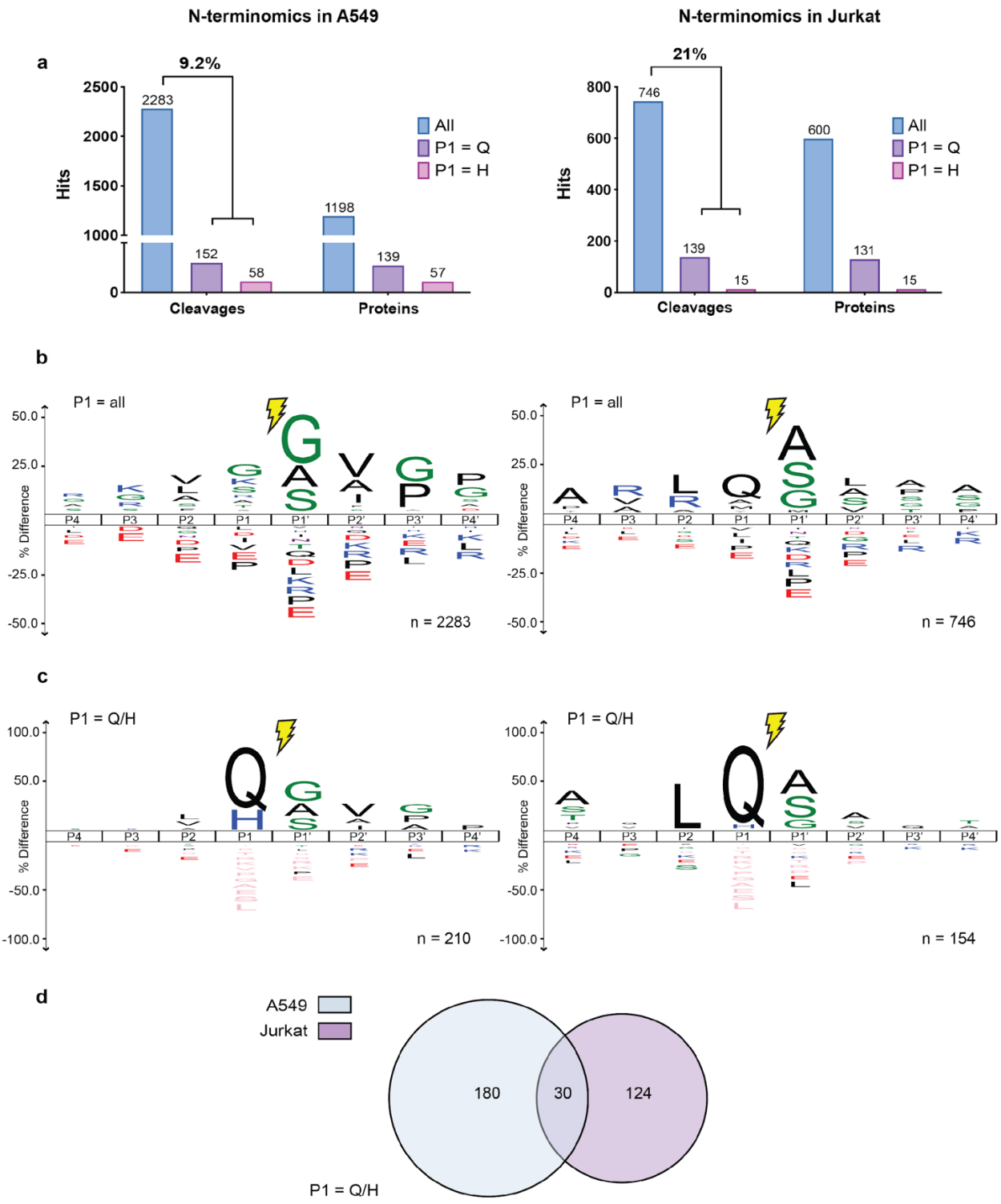

Figure S3. Identification of SARS-CoV-2 M^{pro} substrates. a) N-terminomics statistics of two A549 replicates (left) and of two Jurkat replicates (right). In A549, 2283 unique cleavages were labeled, and 210 sites in 196 host proteins correspond to SARS-CoV-2 M^{pro} specificity with Gln or His at P1 residue (P1=Q/H) at 9.2% enrichment rate. In Jurkat, 746 unique labeled cleavages were identified with 154 sites at P1=Q/H in 146 proteins, showing an enhanced enrichment at 21%. **b**) IceLogo showing P4-P4' residue enrichment in all labeled cleavage sites in A549 (left) and Jurkat (right), and **c**) in sites where P1=Q/H only. **d**) Venn diagram showing the overlap in P1=Q/H cleavages between the A549 and Jurkat proteomes.

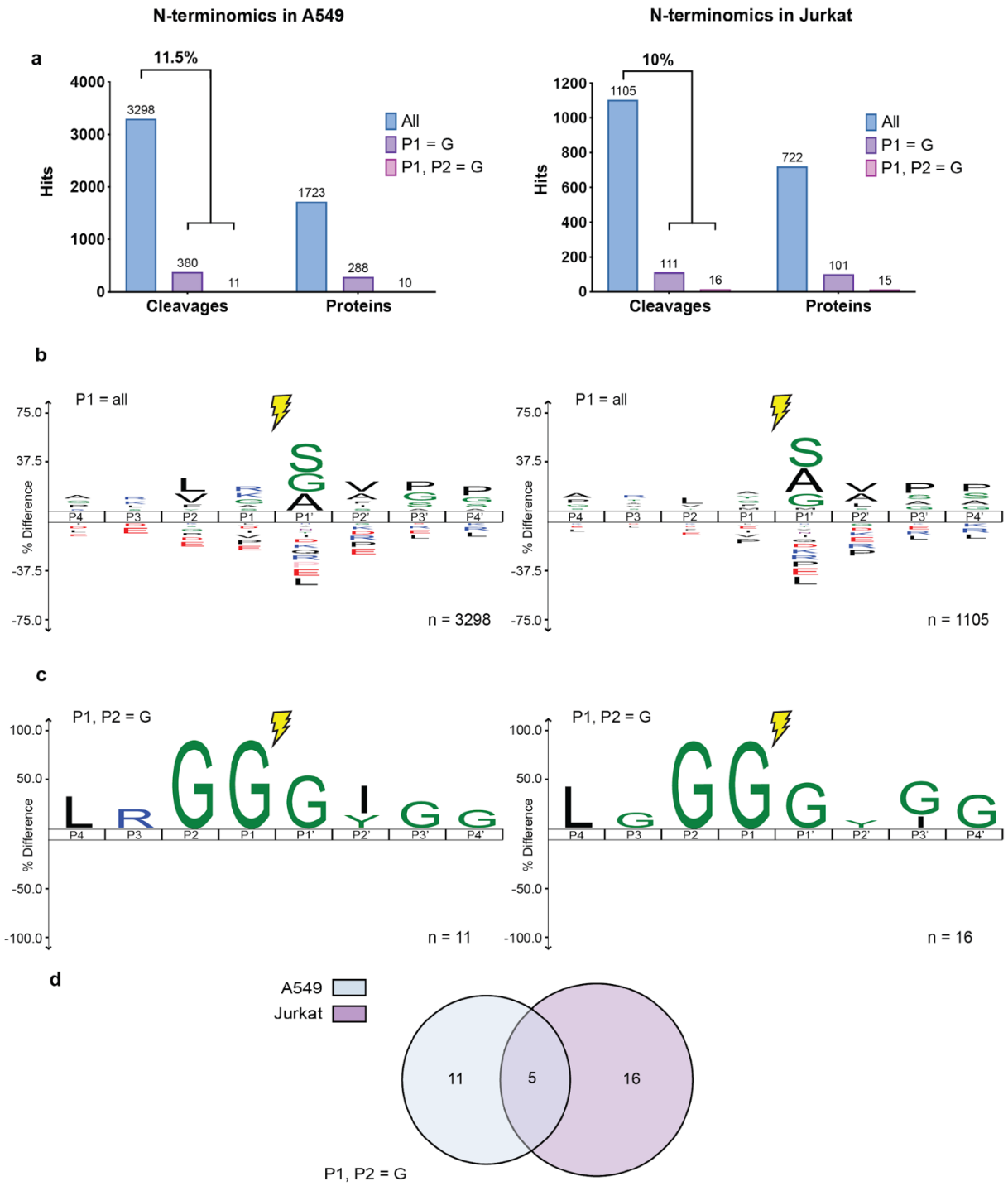

Figure S4. Identification of SARS-CoV-2 PL^{pro} substrates. a) N-terminomics statistics of two A549 (left) and two Jurkat (right) replicates for PL^{pro}. In A549, 3298 unique labeled cleavages were identified with 380 sites at P1=G and 11 sites at P1,P2=G in 288 and 10 proteins, respectively, showing an enhanced enrichment for P1=G at 11.5%. In Jurkat, 1105 unique labeled cleavages were identified with 111 sites at P1=G and 16 sites at P1,P2=G in 101 and 15 proteins, respectively, showing an enhanced enrichment for P1=G at 10%. **b)** IceLogo showing P4-P4' residue enrichment in all labeled cleavage sites in A549 (left) and Jurkat (right), and **c)** in sites where P1,P2=G only. **d)** Venn diagram depicting the overlap in cleavage sites identified in A549 and Jurkat with P1,P2=G.

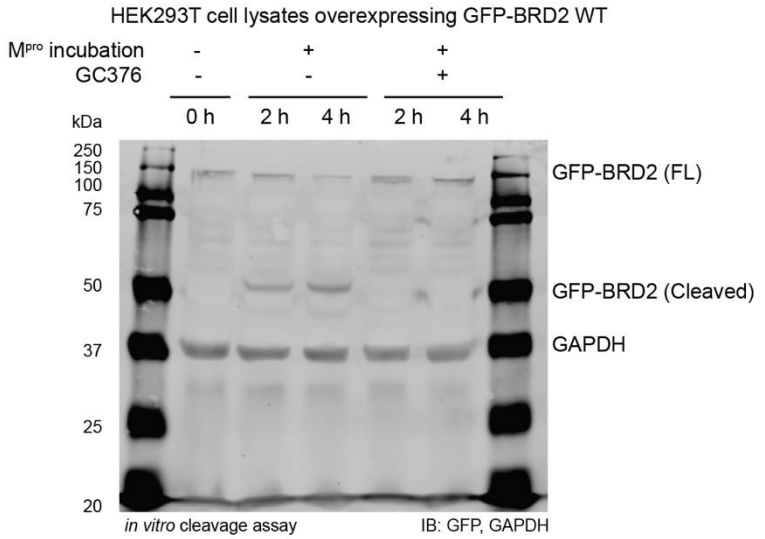

Figure S5. BRD2 is cleaved by SARS-CoV-2 M^{pro} (n=2, biological replicates)**.** HEK293T cell lysates overexpressing GFP-BRD2 were incubated with 0.5 µM M^{pro} for 0, 2, and 4 h in the absence or presence of 8 µM M^{pro} inhibitor GC376 (Selleck Chemicals, #S0475, dissolved in DMSO). GFP-BRD2 cleavage product was only observed without GC376 in the assay.

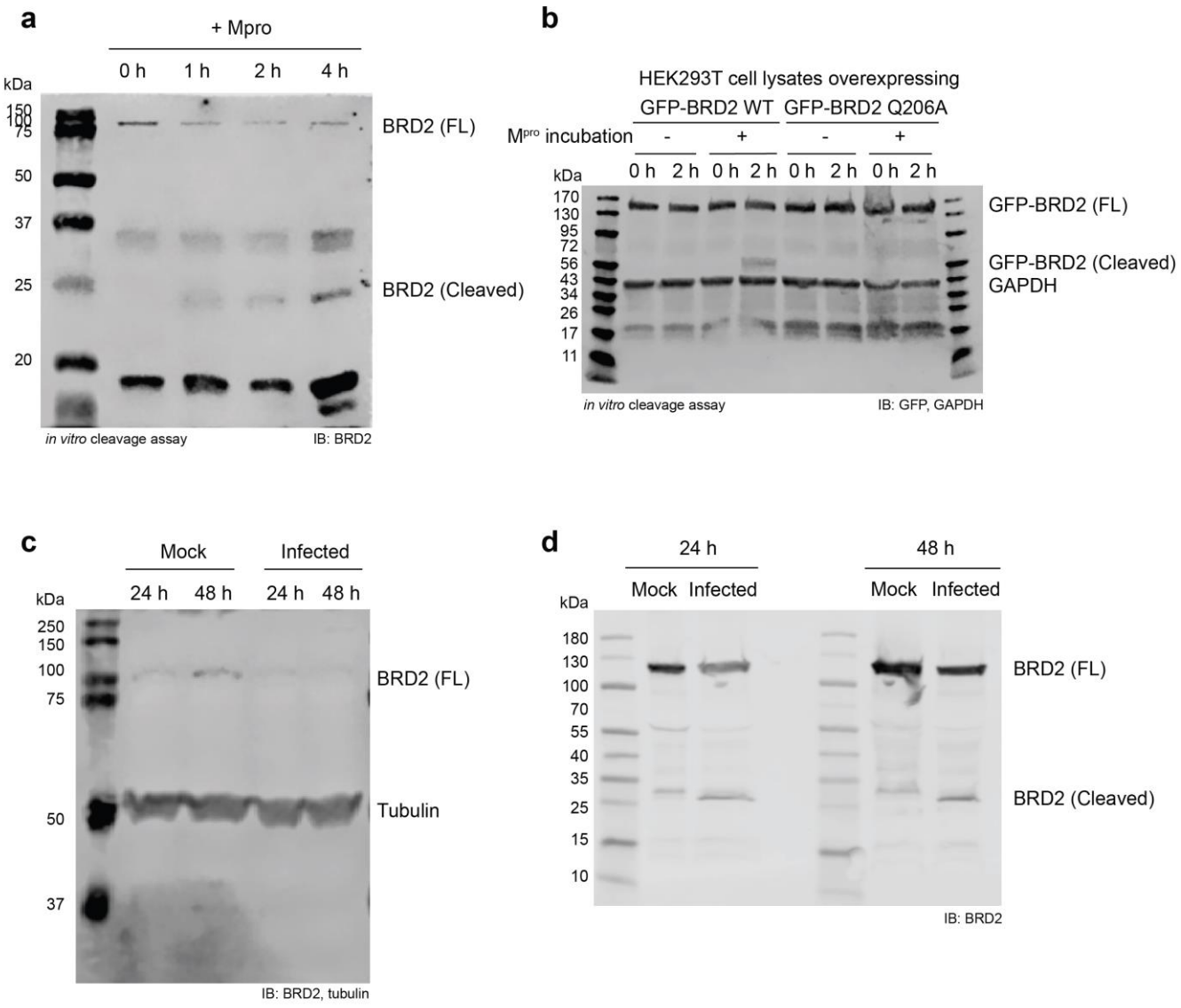

Figure S6. **Full immunoblot images of Fig. 3. Proteolysis of BRD2 by M^{pro} *in vitro* and in SARS-CoV-2 infected cells. a**) BRD2 was cleaved by recombinant M^{pro} in Jurkat cell lysates. Jurkat cell lysates were incubated with recombinant M^{pro} for 0-4 hours, and immunoblotted against BRD2. A cleavage product at 23 kDa appeared with incubation time as the full length BRD2 level decreased. **b**) GFP-BRD2 WT and mutant Q206A overexpression in HEK293T-ACE2 and *in vitro* cleavage by recombinant SARS-CoV-2 M^{pro}. HEK293T-ACE2 cells overexpressing GFP-BRD2 were lysed, and the cell lysates were incubated with M^{pro} for 2 hours and immunoblotted against GFP. Cleavage was only observed with GFP-BRD2 WT. Depletion of full-length BRD2 was also observed in SARS-CoV-2 infected **c**) A549-ACE2 and **d**) HEK293T-ACE2 at 24 and 48 h.p.i.

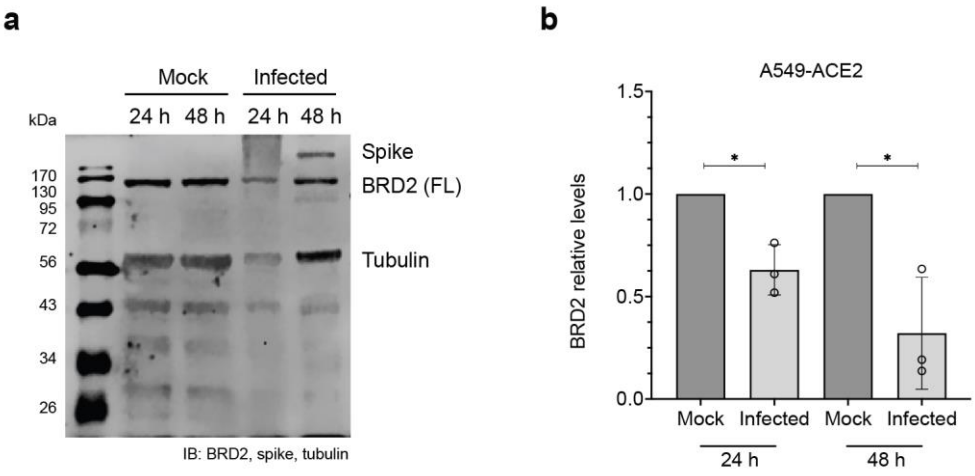

Figure S7. Additional immunoblot of SARS-CoV-2 infected A549-ACE2 cell lysates and BRD2 quantification. a) Another representative replicate was performed in addition to the blot presented in Figure 3 of the main manuscript, for a total of n=3, biological replicates. **b**) Quantification of BRD2 levels at 24 h and 48 h after infection with SARS-CoV-2 compared to mock. BRD2 levels measured were 0.6 ± 0.1 (24 h) and 0.3 ± 0.3 (48 h), with a $*p < 0.05$ using Student's t-test.

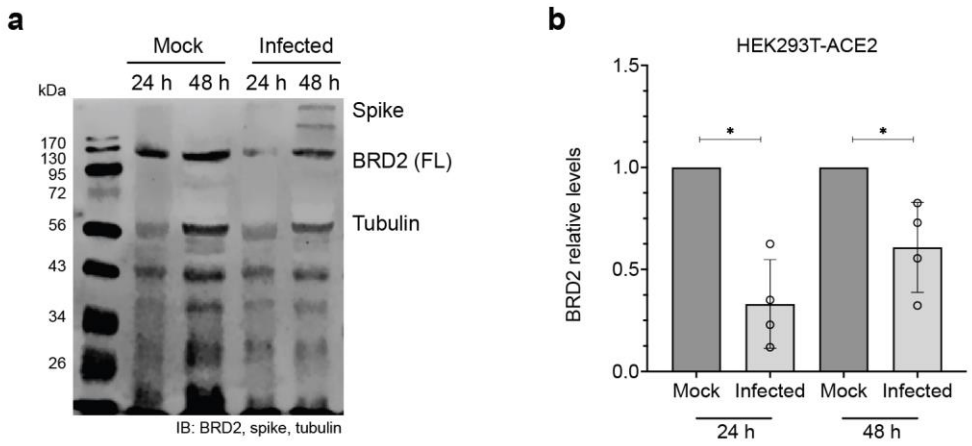

Figure S8. **Additional immunoblot of SARS-CoV-2 infected HEK293T-ACE2 cell lysates and BRD2 quantification. a**) Another representative replicate was performed in addition to the blot presented in Figure 3 of the main manuscript, for a total of n=4, biological replicates. **b**) Quantification of BRD2 levels at 24 h and 48 h after infection with SARS-CoV-2 compared to mock. BRD2 levels measured were 0.3 ± 0.2 (24 h) and 0.6 ± 0.2 (48 h), with a $*p < 0.05$ using Student's t-test.

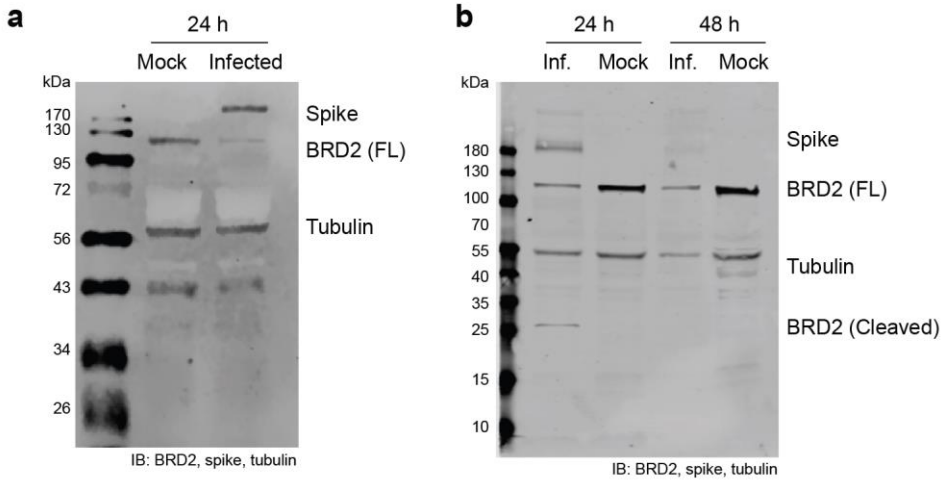

Figure S9. Immunoblots of SARS-CoV-2 infected H23-ACE2 cell lysates in two biological replicates. a,b) A decrease in full-length BRD2 level was observed at 24 h.p.i., as most Infected cells underwent apoptosis at 48 h. A band corresponding to the apparent cleavage product of BRD2 at ~23 kDa was present in **b**).

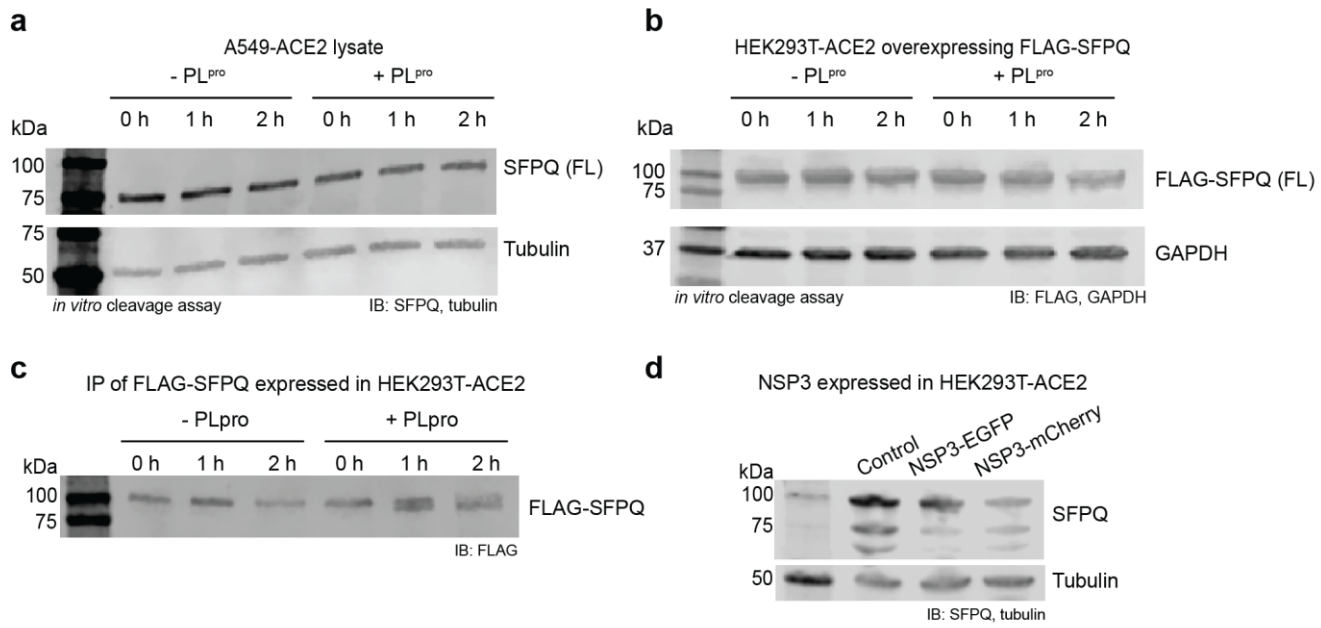

Figure S10. Additional investigation of cleavage of SFPQ by PL^{pro} *in vitro* (n=2, biological replicates)**. a**) Uninfected A549-ACE2 cell lysates were incubated with PL^{pro} and SFPQ cleavage by PL^{pro} could not be detected using immunoblotting. **b**) Cleavage of overexpressed FLAG-tagged SFPQ in HEK293T-ACE2 cells by PL^{pro} was also not detected on immunoblot. **c**) A potential cleavage product was observed when incubating immunoprecipitated FLAG-SFPQ with SARS-CoV-2 PL^{pro}. **d**) Expression of full length NSP3 in HEK293T-ACE2 did not show distinct cleavage of SFPQ compared to the control.

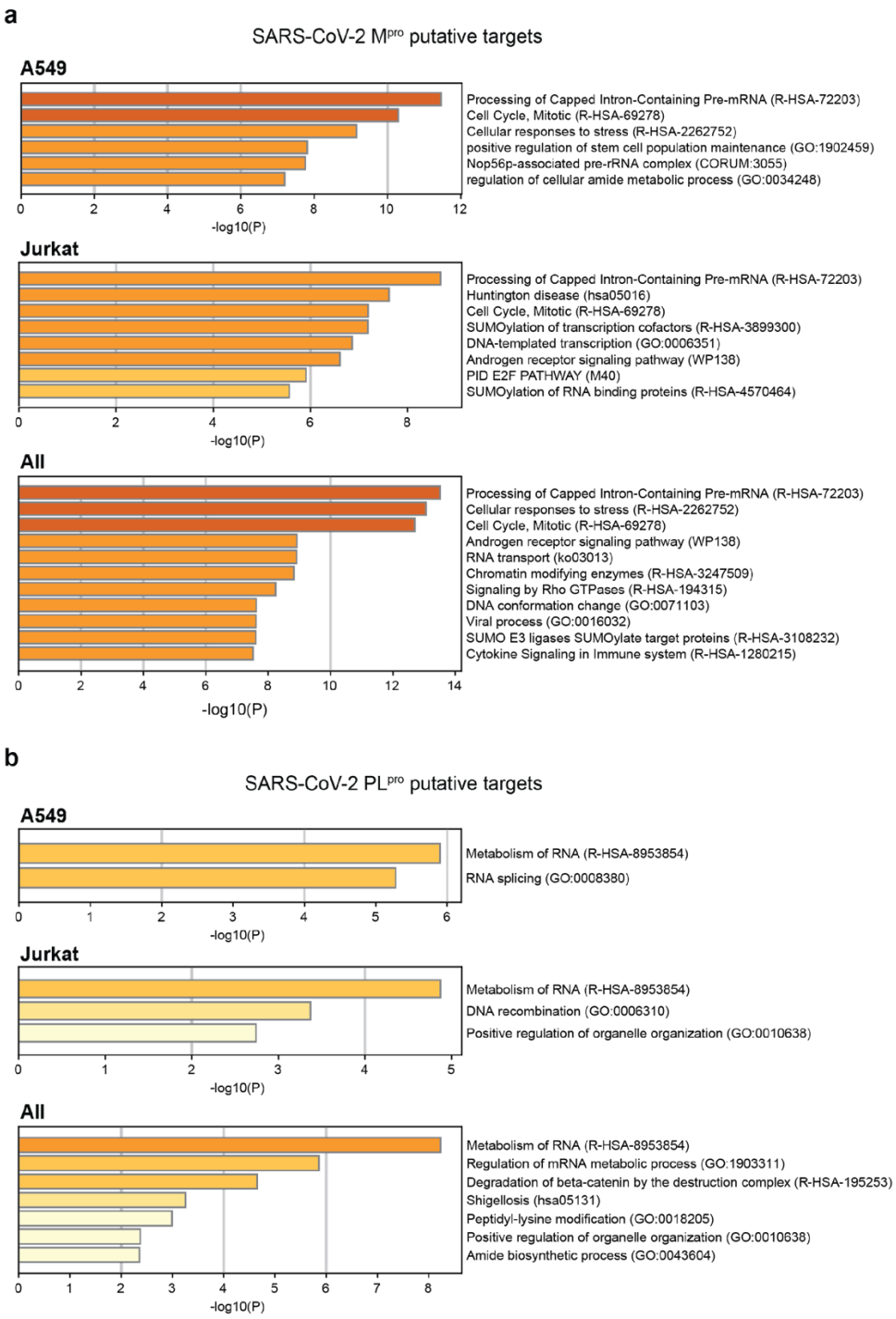

Figure S11. Gene Ontology analysis of a) M^{pro} and b) PL^{pro} putative substrates in A549, Jurkat, and both A549 and Jurkat cell lysates using Metascape. (3)

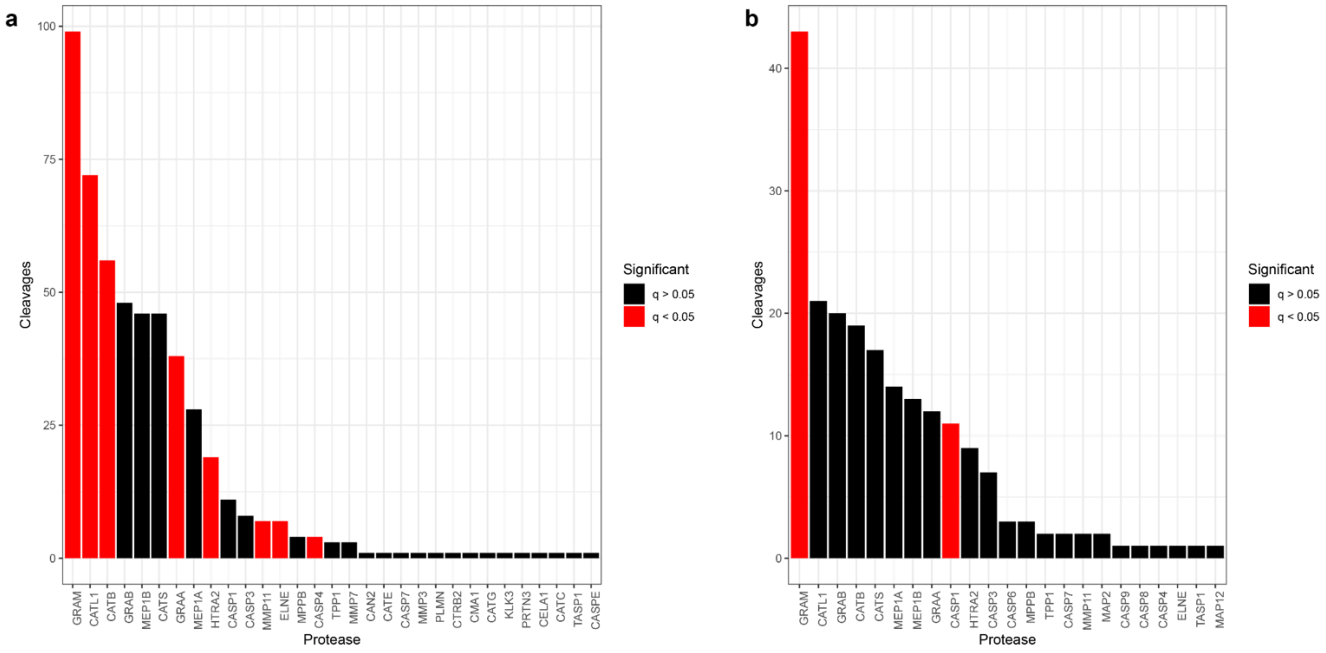

Figure S12. **TopFind analysis of all labeled cleavage sites** in **a**) M^{pro} and **b**) PL^{pro} subtiligase N-terminomics experiments.

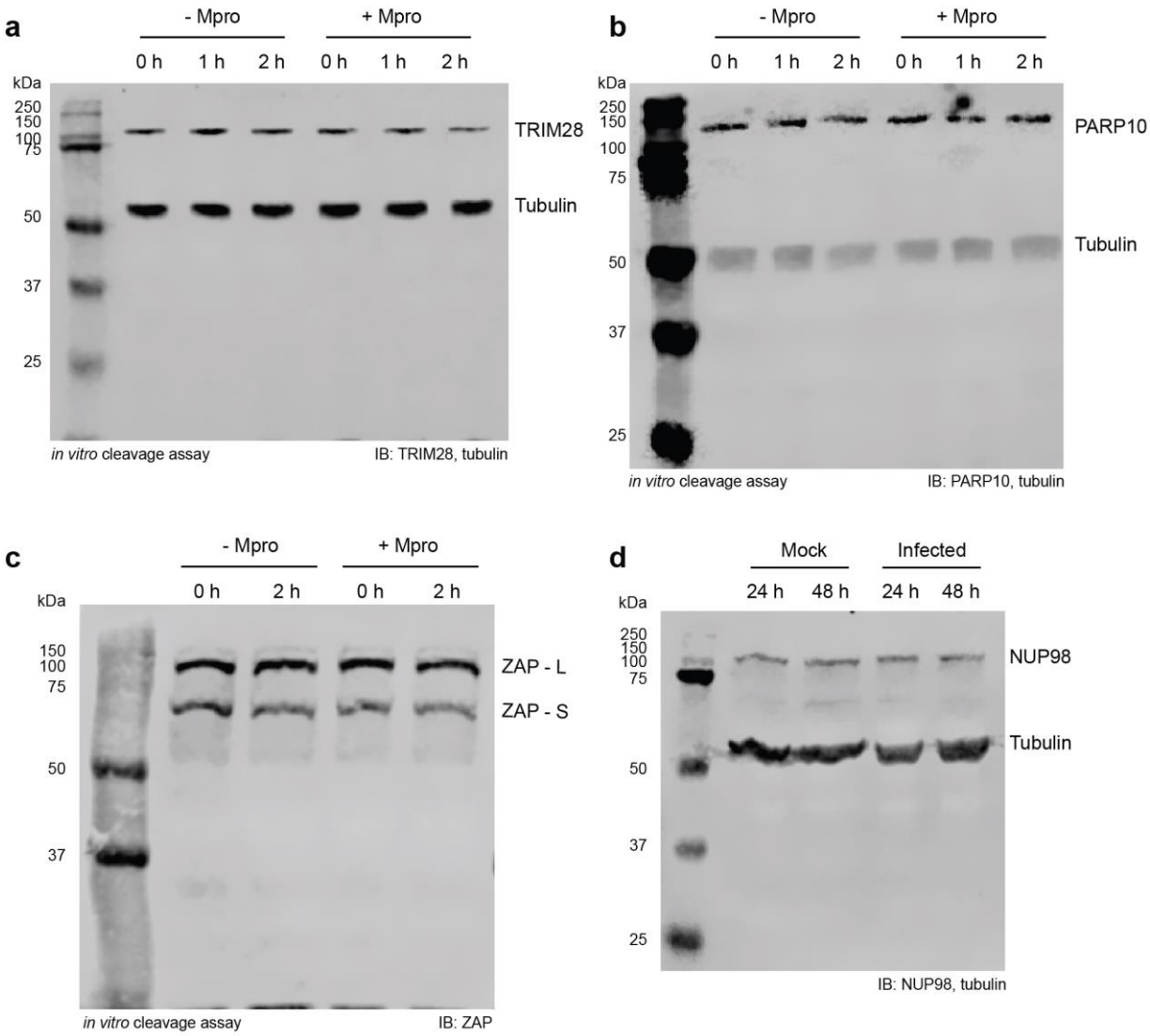

Figure S13. Substrate proteolysis by M^{pro} was not detectable by immunoblot for **a**) *in vitro* cleavage assays of TRIM28 in Jurkat lysates; **b**) *in vitro* cleavage assays of PARP10 in Jurkat lysates; **c**) *in vitro* cleavage assays of ZAP in A549 lysates; and **d**) endogenous NUP98 level in infected A549-ACE2 cells. This suggests that these targets can be cleaved by M^{pro}, but only at a low level detectable only by mass spectrometry.

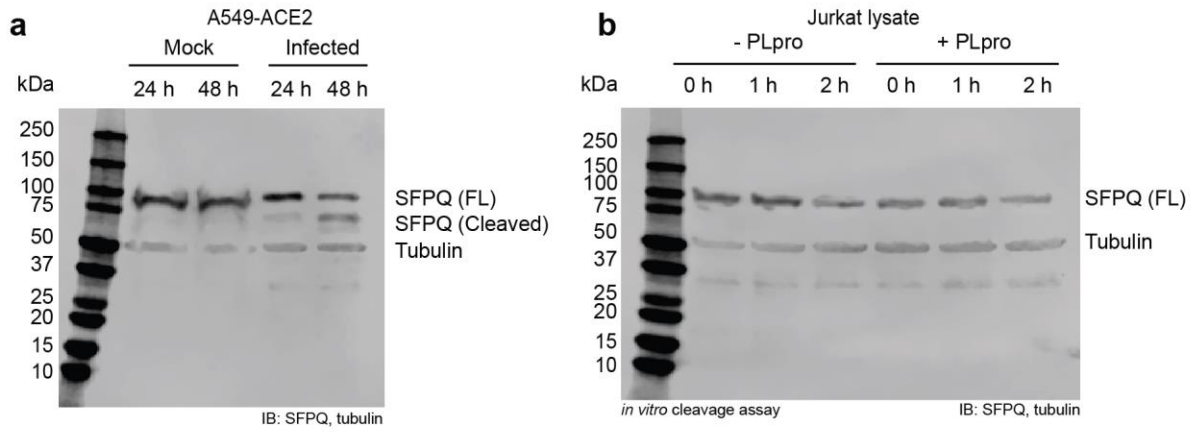

Figure S14. **Full immunoblot images of Fig 4. Proteolysis of SFPQ in SARS-CoV-2 infected cells** (n=2, biological replicates)**. a**) SFPQ was cleaved in A549-ACE2 cells infected with SARS-CoV-2. **b**) SFPQ cleavage by PL^{pro} could not be detected using immunoblotting for Jurkat cell lysates.

Table S1. Antibodies and plasmids used in the study.

Acc #	Gene Name	Sources	Catalog #	Plasmids for overexpression studies
Q13263	TRIM28	R&D Systems	MAB7785	https://www.addgene.org/45569/
Q7Z2W4	ZC3HAV1	Proteintech	16820-1-AP	https://www.addgene.org/45907/
		GeneTex	GTX120134	https://www.addgene.org/45906/
P52948	NUP98	Wozinak Lab		-
P25440	BRD2	Abcam	ab139690	https://www.addgene.org/65376/
Q53GL7	PARP10	LSBio	LS-C747885	-
P23246	SFPQ	Thermo Fisher	PA519663	https://www.addgene.org/166960/
P0DTD1	NSP3	-	-	https://www.addgene.org/165108/ https://www.addgene.org/165131/
P42212	GFP	Abcam	ab6673	-
P04406	GAPDH	Cell Signaling	2118	-
P68363	α-tubulin	Cell Signaling	3873	-
Q93H4B7	β-tubulin	Sigma	T5293	-
P60709	β-actin	Abcam	ab8224	-

References

1. Rut, W.; Lv, Z.; Zmudzinski, M.; Patchett, S.; Nayak, D.; Snipas, S. J.; El Oualid, F.; Huang, T. T.; Bekes, M.; Drag, M.; Olsen, S. K. Activity profiling and crystal structures of inhibitor-bound SARS-CoV-2 papain-like protease: A framework for anti-COVID-19 drug design. *Sci Adv* **2020**, *6* (42): eabd4596.
2. Rut, W.; Groborz, K.; Zhang, L.; Sun, X.; Zmudzinski, M.; Pawlik, B.; Wang, X.; Jochmans, D.; Neyts, J.; Młynarski, W.; Hilgenfeld, R.; Drag, M. SARS-CoV-2 M^{pro} inhibitors and activity-based probes for patient-sample imaging. *Nat Chem Biol* **2021**, *17* (2): 222-228.
3. Zhou, Y.; Zhou, B.; Pache, L.; Chang, M.; Khodabakhshi, A. H.; Tanaseichuk, O.; Benner, C.; Chanda, S. K. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **2019**, *10* (1): 1523.