Cell Reports, Volume *37*

Supplemental information

Mechanistic insights into COVID-19 by global analysis of the SARS-CoV-2 3CL^{pro} substrate degradome

Isabel Pablos, Yoan Machado, Hugo C. Ramos de Jesus, Yasir Mohamud, Reinhild Kappelhoff, Cecilia Lindskog, Marli Vlok, Peter A. Bell, Georgina S. Butler, Peter M. Grin, Quynh T. Cao, Jenny P. Nguyen, Nestor Solis, Srinivas Abbina, Wioletta Rut, John C. Vederas, Laszlo Szekely, Attila Szakos, Marcin Drag, Jayachandran N. Kizhakkedathu, Karen Mossman, Jeremy A. Hirota, Eric Jan, Honglin Luo, Arinjay Banerjee, and Christopher M. Overall

Table S1 related to Figures 1 and 6, and Table 1. Candidate 3CL^{pro} substrates (*n* = 58) and their cleavage sites (*n* = 65) detected in human embryonic kidney (HEK-293) and human lung epithelial (BEAS-2B) cells that require further validation.

Gene Name	UniProt	TAILS neo-N-Terminal P' Peptide	Byonic Score	P4 – P4'*	HEK-293	BEAS-2B		
						Control	IFN-α	IFN-β
PAIRB†	SERBP1	AAAQTNSNAAGKQLR	809	SAAQ↓^{57}AAAQ	●●●	●●●	●●●	●●●
PAIRB†	SERBP1	GEGKIIDR	549	QQLQ↓^{119}GEGK	●●●	●●●	●●●	●●●
K2C8	KRT8	AEIEGLKGQR	644	SRLQ↓^{319}AEIE	●●●	●●●	●●●	●●●
PTBP1	PTBP1	GALAPLAIPSAAAAAAAAGR	641	PNVH↓^{306}GALA	●●●	●●●	●●●	●●●
PININ†	PNN	SQPPSQPEDLSLAVLQPTPQVTQEQGHLLPER	615	PVLQ↓^{496}SQPP	●●●	●●●	●●●	●●●
PININ†	PNN	SQSQPVLQSQPPSQPEDLSLAVLQPTPQVTQEQGHLLPER	571	LQLQ↓^{488}SQSQ	●●●	●●●	●●●	●●●
CHSP1	CARHSP1	ASVGLLDTPR	503	PTHQ↓^{16}ASVG	●●●	●●●	●●●	●●●
ATLA2	ATL2	ATAEANNLAAVAGAR	502	SMLQ↓^{377}ATAE	●●●	●●●	●●●	●●●
LMNA	LMNA	ASSTPLSPTR	463	SGAQ↓^{16}ASST	●●●	●●●	●●●	●●●
ALDOA	ALDOA	SIGTENTEENRR	456	KRLQ↓^{46}SIGT	●●●	●●●	●●●	●●●
STMN1	STMN1	AFELILSPR	440	ASGQ↓^{19}AFEL	●●●	●●●	●●●	●●●
HSP7C	HSPA8	HGKVEIIANDQGNR	541	GVFQ↓^{23}HGKV	●●●	●●	●●●	●●●
TCPZ	CCT6A	AALAVNISAAR	521	ARAQ↓^{18}AALA	●●●	●●	●●●	●●●
FUS	FUS	SYNPPQGYGQQNQYNSSSGGGGGGGGGGNYGQDQSSMSSGG GSGGGYGNQDQSGGGGSGGYGQQDR	265	GQQQ↓^{148}SYNP	●●●	●●●	●●●	●●
ZC3H13	ZC3H14	SSWVYETGR	271	LDMQ↓^{251}SSWV	●●	●	●●●	
PRP4B	PRPF4B	GYESGSEEEGEIHEKAR	593	LILQ↓^{139}GYES	●●●			
HNRPD	HNRNPD	GAAAAAGSGAGTGGGTASGGTEGGSAESEGAKIDASKNEEDEG HSNSSPR	478	AATQ↓^{36}GAAA	●●●			
ZFR	ZFR	SDVQPVGHDYVEEVRNDEGKVIR	469	AALQ↓^{559}SDVQ	●●●			
SFPQ	SFPQ	HHQGPPPGGPGGR	436	YHQQ↓^{255}HHQG	●●●			
GRAP1	GRIPAP1	AENTALQKNVAALQER	426	LRLQ↓^{142}AENT	●●●			
CSTF2†	CSTF2	ASMQGGVPAPGQMPAAVTGPGPGSLAPGGGMQAQVGMPGSGP VSMER	414	PLMQ↓^{248}ASMQ	●●●			
CSTF2†	CSTF2	VGMPGSGPVSMER	342	MQAQ↓^{282}VGMP		●●●	●●●	●●●
RBP56†	TAF15	SQSGYSQSYGGYENQKQSSYSQQPYNNQGQQQNMESSGSQGG	412	SYGQ↓^{59}SQSG	●●●			
RBP56†	TAF15	SGYSQSYGGYENQKQSSYSQQPYNNQGQQQNMESSGSQGGR	390	GQSQ↓^{61}SGYS	●●			
ATX2	ATXN2	AGIIPTEAVAMPIPAASPTPASPASNR	386	ASPQ↓^{712}AGII	●●●			
DAXX	DAXX	GTSSHSADTPEASLDSGEGPSGMASQGCPSASR	369	ARLQ↓^{400}GTSS	●●●			
PRRC1	PRRC1	GTGTTSAITFPEEQEDPR	338	SLAQ↓^{183}GTGT	●●●			
EPIPL	EPPK1	TSGILGPETLR	338	SELH↓^{2724}TSGI	●●●			
TOX4	TOX4	AAAAAAAAASMQLPPPR	334	SVLQ↓^{426}AAAA	●●●			
EWS	EWSR1	SSSYGQQSSFR	309	YSQQ↓^{258}SSSY	●●●			
AKP8L	AKAP8L	STYSDTSAQPTCDYGYGTWNSGTNR	308	TTLQ↓^{16}STYS	●●●			
IF4G1†	EIF4G1	GAYYIPGQGR	304	PASQ↓^{101}GAYY	●●●			
IF4G1†	EIF4G1	MSVEESTPISR	322	TTIQ↓^{295}MSVE		●●●	●●●	●●●
IF4G1†	EIF4G1	QAVPTESTDNRR	263	SALQ↓^{1128}QAVP		●●	●●●	●
ARPIN	ARPIN	SSYKVEAKGDTDR	286	GFLM↓^{98}SSYK	●●●			
FLNA	FLNA	SGTTNKPNKFTVETR	273	PGIQ↓^{1367}SGTT	●●●			
ENOA	ENO1	VVGDDLTVTNPKR	273	AGIQ↓^{315}VVGD	●●●			
SF3B2	SF3B2	GVEVALAPEELELDPMAMTQKYEEHVR	268	PELQ↓^{823}GVEV	●●●			
GOGA3	GOLGA3	AQVECSHSSQQR	256	TKLQ↓^{451}AQVE	●●●			
TAB1	TAB1	SEQQPSWTDDLPLCHLSGVGSASNR	240	SLLQ↓^{11}SEQQ	●●●			
SETD2	SETD2	AAPVPLPVDVAVR	237	TVLM↓^{224}AAPV	●●●			
TMED8	TMED8	SEHTGAIDVLSADLESADLLGDHR	203	VMIQ↓^{127}SEHT	●●●			
ROCK2	ROCK2	ALHIGLDSSSIGSGPGDAEADDGFPESR	406	SQLQ↓^{1125}ALHI	●●			
CNN3	CNN3	MGTNKGASQAGMLAPGTR	386	ISLQ↓^{208}MGTN	●●			
PUR9†	ATIC	SSESKDTSLETR	357	TEMQ↓^{160}SSES	●●			
PUR9†	ATIC	MDQSYKPDENEVR	272	CVLQ↓^{367}MDQS		●●	●●●	●●●
PTBP3	PTBP3	SGSLALSGGPSNEGTVLPGQSPVLR	353	SAVQ↓^{159}SGSL	●●			
NEB2	PPP1R9B	MGTTAGPSGEAGGGAGLAEAPR	312	MFLQ↓^{64}MGTT	●●			
ZKSC8	ZKSCAN8	SEATQHKSPVPQESQER	234	TQLQ↓^{170}SEAT	●●			
PRUN2	PRUNE2	GTQLASFPDTCQPASLNER	378	TGLQ↓^{2588}GTQL		●●●	●●●	●●●
TRIR	TRIR	GREAPGPAGGGGGGSR	350	AEPQ↓^{12}GREA		●●●	●●●	●●●
ANLN	ANLN	AADTISDSVAVPASLLGMR	311	VQPQ↓^{109}AADT		●●●	●●●	●●●
ELYS	AHCTF1	AEFDMSAIPR	238	LTVQ↓^{1963}AEFD		●●●	●●●	●●●
RABE1	RABEP1	SREQVSEELVR	342	VLMQ↓^{634}SREQ	●	●●●	●●●	●
FA98A	FAM98A	TSSSYTGSGYQGGGYQQDNR	326	GGFQ↓^{428}TSSS		●●	●●●	●●●
VATH	ATP6V1H	GQMISAEDCEFIQR	259	SYLQ↓^{39}GQMI		●●	●●●	●●●
GOGB1	GOLGB1	LSQTQAEQAAQQVVR	237	LQAQ↓^{206}LSQT		●●	●●●	●●●
DLGP4	DLGAP4	SIGVQVEDDWR	231	TKFQ↓^{690}SIGV		●●●	●●●	●●
DREB	DBN1	AWAGPMEEPPQAQAPPR	431	AAPQ↓^{396}AWAG		●	●●●	●●●

DEN5B	DENND5B	SDVLATGPTSNNR	294	PKLQ↓[660]SDVL	000	●●●	●●●	●
SRRM2	SRRM2	SGSDSSPEPKAPAPR	222	PRAQ↓[1616]SGSD	000	●	●●●	●●●
LRRF1	LRRFIP1	SREIDCLSPEAQKLAEAR	267	AAAQ↓[9]SREI	000	●●●	●●	●
SPT6H	SUPT6H	SAQAQPQPSSSSR	357	TTPQ↓[1646]SAQA	000	●	●	● ●
K1C18	KRT18	SVENDIHGLR	341	AMRQ↓[177]SVEN	000	●	000	●●●
LRC59	LRRC59	HMKAVQADQER	243	KVLQ↓[147]HMKA	000	000	000	●●●

Candidate substrates are those that did not meet high stringency manual inspection of the MS/MS spectra of the filtered neo-N-terminal P' peptides. Fields marked as "●" or "●" indicate in which of the *N* = 12 independent cell experiments the neo-N-terminal P' peptide was found by TAILS LC-MS/MS with an FDR ≤ 0.01 at the peptide level. *, Amino acid sequence of the cleavage site and P1' amino acid position identified from the neo-N-terminal peptide. ↓, scissile bond. †, Proteins with multiple cut sites are shown with each other according to the highest-ranked peptide.

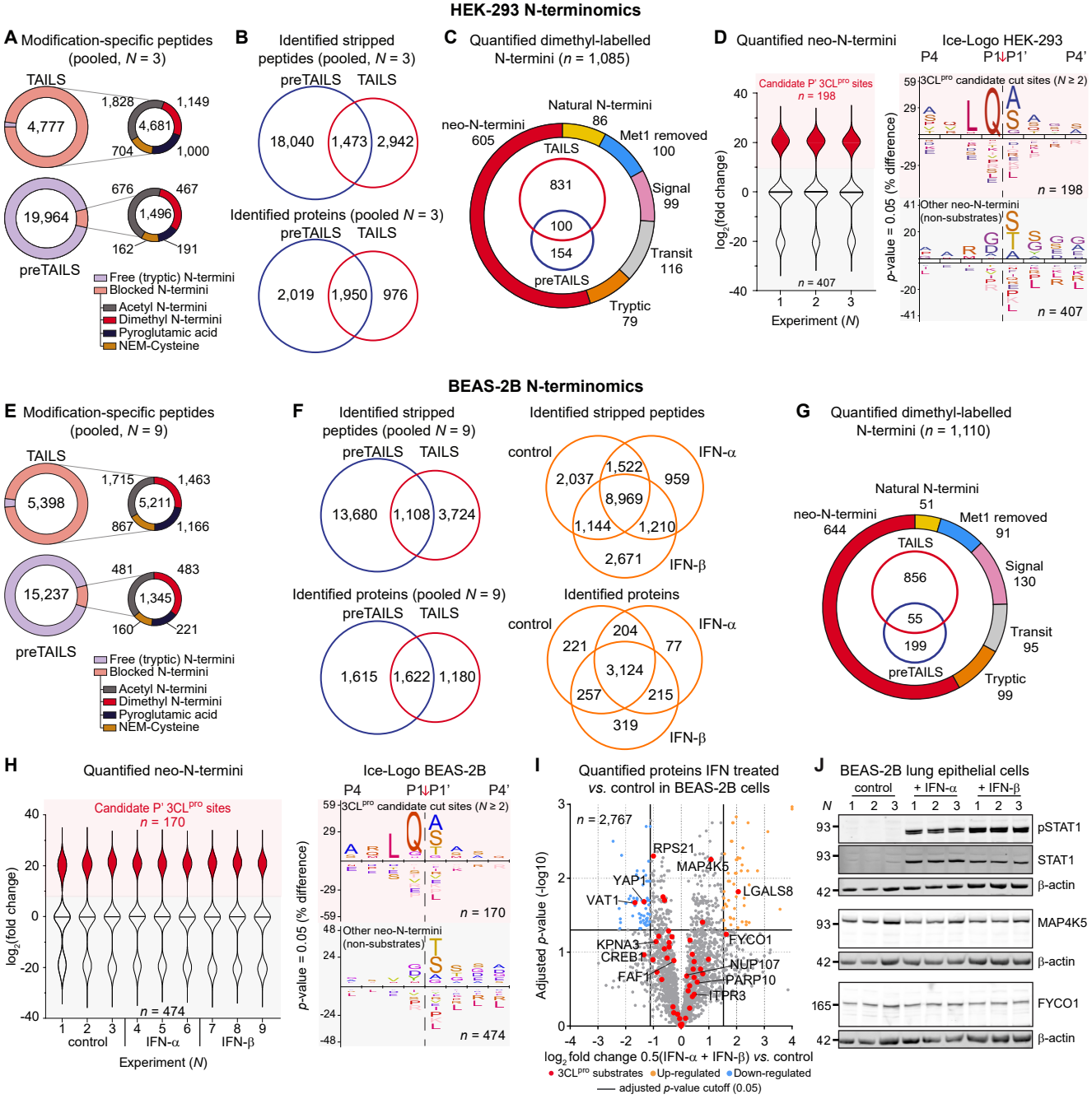

HEK-293 N-terminomics
A Modification-specific peptides (pooled, N = 3)
TAILS
4,777
1,828
4,681
1,149
704
1,000
19,964
676
1,496
467
162
191
preTAILS
Free (tryptic) N-termini
Blocked N-termini
Acetyl N-termini
Dimethyl N-termini
Pyroglutamic acid
NEM-Cysteine
B Identified stripped peptides (pooled, N = 3)
preTAILS TAILS
18,040 1,473 2,942
Identified proteins (pooled N = 3)
preTAILS TAILS
2,019 1,950 976
C Quantified dimethyl-labelled N-termini (n = 1,085)
Natural N-termini 86
neo-N-termini 605
Met1 removed 100
TAILS
831
Signal 99
100
154
Transit 116
preTAILS
Tryptic 79
D Quantified neo-N-termini
Candidate P' 3CLpro sites n = 198
log2(fold change)
n = 407
Experiment (N)
Ice-Logo HEK-293
P4 P1↓P1' P4'
3CLpro candidate cut sites (N ≥ 2)
n = 198
Other neo-N-termini (non-substrates)
n = 407
p-value = 0.05 (% difference)
BEAS-2B N-terminomics
E Modification-specific peptides (pooled, N = 9)
TAILS
5,398
1,715
5,211
1,463
867
1,166
15,237
481
1,345
483
160
221
preTAILS
Free (tryptic) N-termini
Blocked N-termini
Acetyl N-termini
Dimethyl N-termini
Pyroglutamic acid
NEM-Cysteine
F Identified stripped peptides (pooled N = 9)
preTAILS TAILS
13,680 1,108 3,724
Identified stripped peptides
control IFN-α IFN-β
2,037 1,522 959
8,969
1,144 1,210
2,671
Identified proteins (pooled N = 9)
preTAILS TAILS
1,615 1,622 1,180
Identified proteins
control IFN-α IFN-β
221 204 77
3,124
257 215
319
G Quantified dimethyl-labelled N-termini (n = 1,110)
Natural N-termini 51
neo-N-termini 644
Met1 removed 91
TAILS
856
Signal 130
55
199
Transit 95
preTAILS
Tryptic 99
H Quantified neo-N-termini
Candidate P' 3CLpro sites n = 170
log2(fold change)
n = 474
control IFN-α IFN-β
Experiment (N)
Ice-Logo BEAS-2B
P4 P1↓P1' P4'
3CLpro candidate cut sites (N ≥ 2)
n = 170
Other neo-N-termini (non-substrates)
n = 474
p-value = 0.05 (% difference)
I Quantified proteins IFN treated vs. control in BEAS-2B cells
n = 2,767
RPS21
MAP4K5
YAP1
VAT1
LGALS8
KPNA3
CREB1
FAF1
FYCO1
NUP107
PARP10
ITPR3
Adjusted p-value (-log10)
log2 fold change 0.5(IFN-α + IFN-β) vs. control
3CLpro substrates
Up-regulated
Down-regulated
adjusted p-value cutoff (0.05)
J BEAS-2B lung epithelial cells
control + IFN-α + IFN-β
N 1 2 3 1 2 3 1 2 3
pSTAT1
STAT1
β-actin
MAP4K5
β-actin
FYCO1
β-actin

Figure S1 related to Figure 1. Quantitative TAILS and shotgun proteomics of HEK-293 and BEAS-2B cells.

(A) Distribution of free and N-terminal-modified peptides identified by TAILS and PreTAILS in $N = 3$ independent HEK-293 cell experiments.

(B) Venn diagrams of the total numbers of peptides and proteins identified, FDR ≤ 0.01.

(C) Classification of 1,085 quantified isotope-labelled, dimethylated N-termini by TopFinder 4.1 ($N = 3$).

(D) Log_2(fold change) distribution of all quantified neo-N-termini ($n = 605$) used to calculate the native protein P4–P4' cleavage site specificity of $3CL^{pro}$ ($n = 198$) *versus* other cell protease activity before cell harvesting and addition of protease inhibitors ($n = 407$).

(E) Distribution of free and N-terminal-modified peptides identified by TAILS and PreTAILS in $N = 9$ independent BEAS-2B cell experiments.

(F) Venn diagrams of the total numbers of peptides and proteins identified (left), and after 24-h treatment by IFN-α ($N = 3$), IFN-β ($N = 3$), and vehicle ($N = 3$), FDR ≤ 0.01.

(G) Classification of 1,110 quantified isotope-labelled, dimethylated N-termini by TopFinder 4.1 ($N = 9$).

(H) Log_2(fold change) distribution of quantified neo-N-termini ($n = 644$) used to calculate the native protein P4–P4' cleavage site specificity of $3CL^{pro}$ ($n = 170$) *versus* other cell protease activity before cell harvesting and addition of protease inhibitors ($n = 474$).

(I) Volcano plot of log_2-fold change in expression of 2,767 proteins quantified in type I interferon-treated BEAS-2B cells ($N = 6$) relative to controls ($N = 3$). Red-filled circles, $3CL^{pro}$ substrates ($n = 45$). Yellow and blue-filled circles, IRG proteins with increased ($n = 49$) or decreased ($n = 64$) expression, respectively. A multiple sample t-test assessed statistical significance with Benjamini Hochberg correction. To focus on the $3CL^{pro}$ substrates, 113 IRG proteins with an adjusted p-value > 0.05 and log_2-fold change in expression ≥ 1.53 or ≤ –1.11 are not plotted and are listed in table S6.

(J) Immunoblot analysis of elevated phospho-STAT1-Tyr^{701} protein, and unaltered FYCO1 and MAP4K5 protein levels in response to type I interferon signalling. b-actin loading controls. Protein levels of FYCO1 showed no apparent change, and although exhibiting a large relative fold-change at the peptide level by LC-MS/MS, this was not statistically significant. MAP4K5 peptides exhibited a small fold-change upon interferon treatment by LC-MS/MS that was significant. However, this small change was not readily apparent by immunoblotting and is unlikely to be biologically significant.

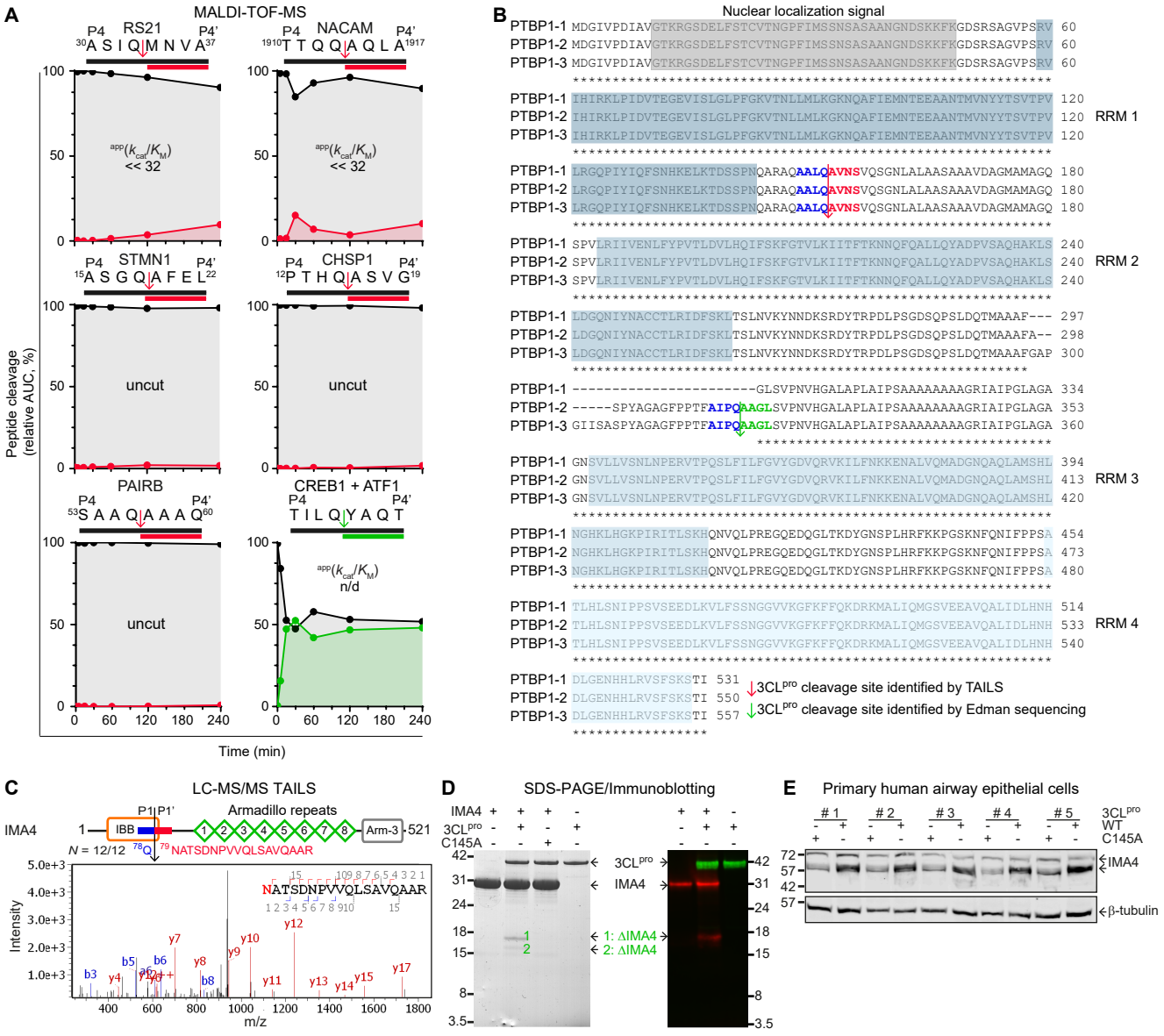

A
MALDI-TOF-MS
RS21
NACAM
STMN1
CHSP1
PAIRB
CREB1 + ATF1
Peptide cleavage (relative AUC, %)
uncut
n/d
Time (min)
B
Nuclear localization signal
PTBP1-1
PTBP1-2
PTBP1-3
RRM 1
RRM 2
RRM 3
RRM 4
↓3CLpro cleavage site identified by TAILS
↓3CLpro cleavage site identified by Edman sequencing
C
LC-MS/MS TAILS
IMA4
IBB
Armadillo repeats
Arm-3
N = 12/12
NATSDNPVVQLSAVQAAR
Intensity
m/z
D
SDS-PAGE/Immunoblotting
3CLpro
C145A
IMA4
1: ΔIMA4
2: ΔIMA4
E
Primary human airway epithelial cells
WT
β-tubulin

Figure S2 related to Figures 2, 3 and 4. 3CL^{pro} substrate validation by peptide and protein cleavage assays.

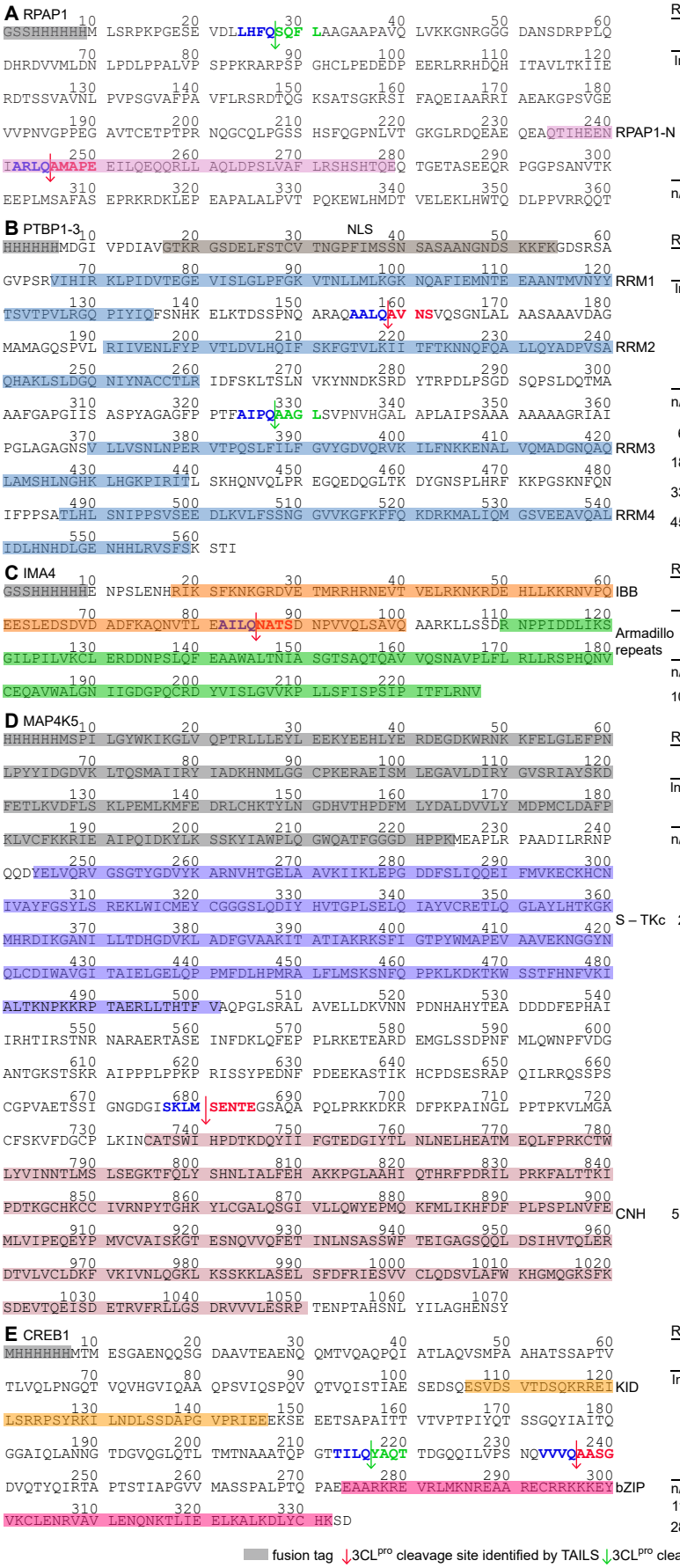
(A) MALDI-TOF-MS analyses of peptides assay incubated with 3CL^{pro} (1:20 molar ratio, E:S). Peptides from STMN1, CHSP1 and PAIRB, were not cut; thus, they were not included in our final substrate list (Table 1). The CREB1 cleavage site (TILQ^{223}↓YAQT) was identified by Edman sequencing of recombinant human CREB1 cleavage products (Figure 4A) and is identical in ATF1. Product generation (red, site identified by TAILS; green, site identified by Edman sequencing) and substrate consumption (black) were calculated as the peak area normalized to the total peak area in the spectrum.

(B) PTBP1 isoforms 1, 2 and 3 sequence alignments by Clustal Omega. P and P' sites identified by TAILS (AALQ^{152}↓AVNS) and Edman sequencing (AIPQ^{322}↓AAGL, isoform-3 numbering), blue/red and blue/green, respectively.

(C) Relative location of the 3CL^{pro} cleavage site in the importin beta binding domain (IBB) of IMA4 identified by the TAILS neo-N-terminal peptide (red). Representative MS/MS spectrum of neo-N-terminal peptides identifying the IMA4-cleavage site in every TAILS experiment (*N* = 12/12).

(D) Validation of 3CL^{pro} cleavage of recombinant human IMA4 by SDS-PAGE and immunoblotting for IMA4 and 3CL^{pro} (mouse anti-FLAG antibody). IMA4 was incubated (1:5 mol/mol, E:S) with 3CL^{pro} or inactive 3CL^{pro}-C145A, 18 h. ΔIMA4, no sequence of cleavage product was obtained. Figure S3C, Edman sequence data for IMA4 cleavage products 1 and 2.

(E) IMA4 immunoblot of cell lysates from HAEC from five donors. Lysates were incubated with 3CL^{pro} or 3CL^{pro}-C145A (1:200 w/w, E:S) for 18 h, 37°C. β-tubulin, loading control.



Recombinant partial RPAP1 analysed in Fig. 3B

Bands	ID	Edman sequencing	MW_{Theor} (kDa)	MW_{Exp} (kDa)
Intact RPAP1 (1 – 351)	RPAP1	GSSHH	39.4	~52.0
G1 – Q27	-	-	3.1	-
S28 – T360	1	n/a	36.2	~44.9
G1 – Q245	2	n/a	26.4	~33.0
S28 – Q245	3	SQFLA	23.2	~29.0
A246 – T360	4	AMAPE	13.0	~17.0
Truncated RPAP1	5	GSSHH	unknown	~15.0
S28 – unknown	6	SQFLA	unknown	~8.0

n/a: no amino acids identified 225 RPAP1-N 270 RPAP1-like N-terminal 356 RPAP1-C 422 RPAP1-like C-terminal

Recombinant PTBP1-3 analysed in Fig. 3B

Bands	ID	Edman sequencing	MW_{Theor} (kDa)	MW_{Exp} (kDa)
Intact PTBP1 (1 – 557)	PTBP1	n/a	60.5	~64.0
H1 – Q158	-	-	17.2	-
A159 – I563	-	-	43.3	-
H1 – Q327	-	-	35.1	-
A328 – I563	1	AAGLS	25.3	~27.0
A159 – unknown	2	AVNSV	-	~24.0
A159 – Q327	3	AVNSV	18.0	~23.0

n/a: no amino acids identified 11 NLS 48 Nuclear localization signal

60 RRM1 129
185 RRM2 254
338 RRM3 407
455 RRM4 525
RNA recognition motif

Recombinant partial IMA4 analysed in Fig. S2D

Bands	ID	Edman sequencing	MW_{Theor} (kDa)	MW_{Exp} (kDa)
Intact IMA4 (3 – 220)	IMA4	GSSHH	25.4	~29.0
N86 – V227	1	n/a	15.3	~16.0
G1 – Q85	2	n/a	10.01	~15.0

n/a: no amino acids identified 11 IBB 93 Importin beta binding domain

103 440 Armadillo repeats (1 to 8) 447 Arm-3 449 Atypical armadillo repeat

Recombinant MAP4K5 analysed in Fig. 4C

Bands	ID	Edman sequencing	MW_{Theor} (kDa)	MW_{Exp} (kDa)
Intact MAP4K5 (1 – 846)	MAP4K5	n/a	121.3	~119.0
H1 – M680	1	n/a	77.2	~75.0
S681 – Y1070	2	SENTE	44.1	~43.0

n/a: no amino acids identified

20 S – TKc 277 Serine/Threonine protein kinase

511 CNH 826 Citron homology domain

Recombinant CREB1 analysed in Fig. 4C

Bands	ID	Edman sequencing	MW_{Theor} (kDa)	MW_{Exp} (kDa)
Intact CREB1 (1 – 327)	CREB1	MHHHH	36.1	~46.0
M1 – Q216	1	n/a	22.9	~31.0
Contaminant band	§	MKVAK	unknown	~25.0
Y217 – D334	2	YAQTT	13.2	~17.0
M1 – Q236	-	-	25.1	-
A237 – D334	3	AASGD	11.0	~16.0
Y217 – Q236	-	-	2.2	-

n/a: no amino acids identified

113 pKID 153 Phosphorylated kinase-inducible domain

281 bZIP 339 Basic region leucine zipper

Figure S3 related to Figures 3, 4, and S2. Recombinant protein substrate details.
Protein sequence and Edman sequencing of $3CL^{pro}$-cut sites in the recombinant proteins shown in Figures 3, 4 and S2. Blue, P-side. Red, P'-side (identified by TAILS). Green, P'-side (identified by Edman sequencing). Theoretical molecular weight (ProtParam, ExPASy) and experimental molecular weight (GelAnalyzer version 19.1) of intact proteins and the proteolytic fragments generated by $3CL^{pro}$. Protein domain information was accessed from Smart (https://smart.embl.de).
(A) Partial recombinant RPAP1 (1 – 351 aa, BC000246).
(B) PTBP1 (1 – 557 aa, NP_002810.1).
(C) Partial recombinant IMA4 (3 – 220 aa, NP_002258.2).
(D) MAP4K5 (1 – 846 aa, NP_006566.2).
(E) CREB1 (1 – 327 aa, NM_004379).

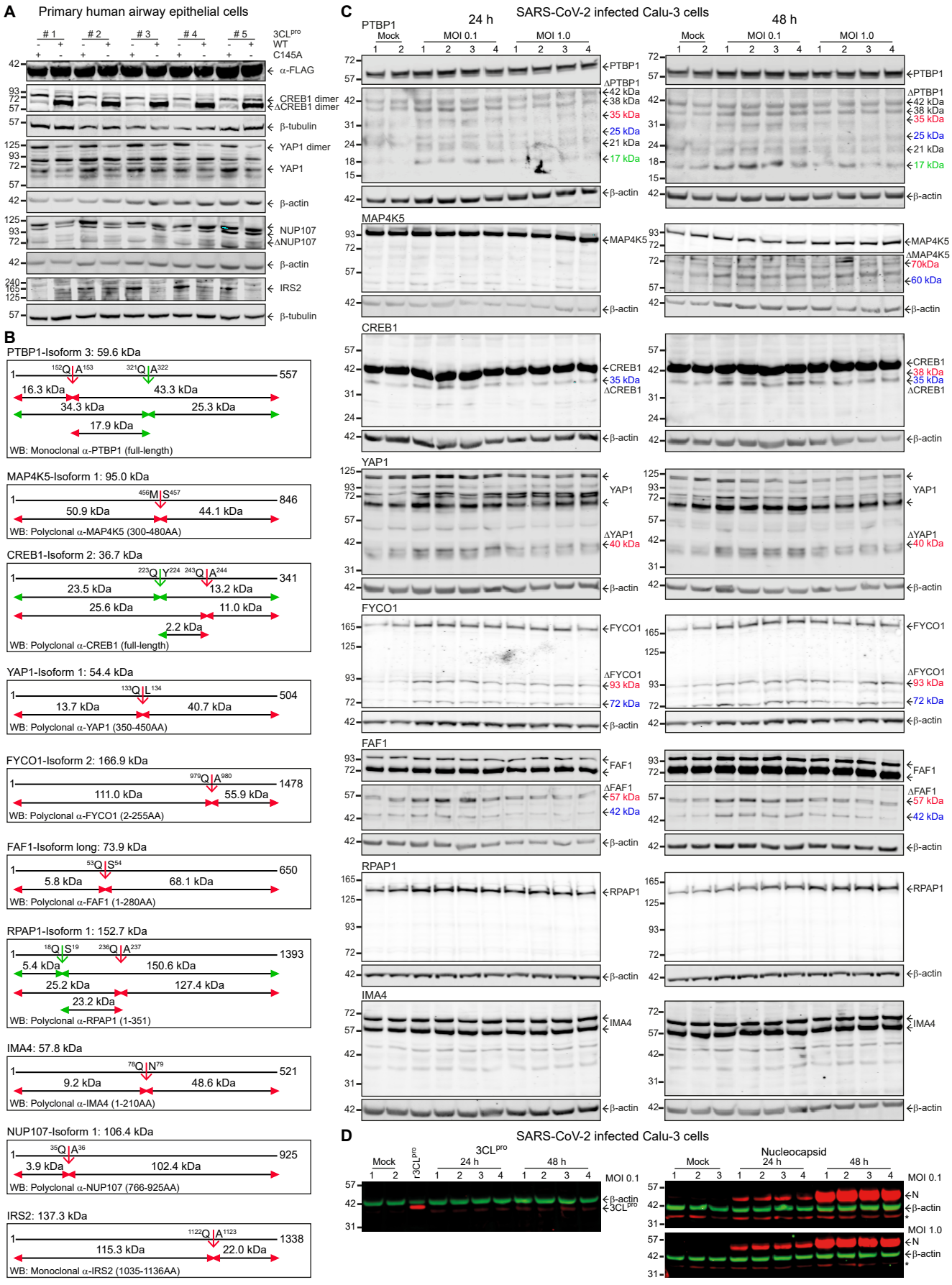

A
Primary human airway epithelial cells
1
2
3
4
5
3CLpro
WT
C145A
α-FLAG
CREB1 dimer
ΔCREB1 dimer
β-tubulin
YAP1 dimer
YAP1
β-actin
NUP107
ΔNUP107
β-actin
IRS2
β-tubulin
B
PTBP1-Isoform 3: 59.6 kDa
152Q|A153
321Q|A322
557
16.3 kDa
43.3 kDa
34.3 kDa
25.3 kDa
17.9 kDa
WB: Monoclonal α-PTBP1 (full-length)
MAP4K5-Isoform 1: 95.0 kDa
456M|S457
846
50.9 kDa
44.1 kDa
WB: Polyclonal α-MAP4K5 (300-480AA)
CREB1-Isoform 2: 36.7 kDa
223Q|Y224
243Q|A244
341
23.5 kDa
13.2 kDa
25.6 kDa
11.0 kDa
2.2 kDa
WB: Polyclonal α-CREB1 (full-length)
YAP1-Isoform 1: 54.4 kDa
133Q|L134
504
13.7 kDa
40.7 kDa
WB: Polyclonal α-YAP1 (350-450AA)
FYCO1-Isoform 2: 166.9 kDa
979Q|A980
1478
111.0 kDa
55.9 kDa
WB: Polyclonal α-FYCO1 (2-255AA)
FAF1-Isoform long: 73.9 kDa
53Q|S54
650
5.8 kDa
68.1 kDa
WB: Polyclonal α-FAF1 (1-280AA)
RPAP1-Isoform 1: 152.7 kDa
18Q|S19
236Q|A237
1393
5.4 kDa
150.6 kDa
25.2 kDa
127.4 kDa
23.2 kDa
WB: Polyclonal α-RPAP1 (1-351)
IMA4: 57.8 kDa
78Q|N79
521
9.2 kDa
48.6 kDa
WB: Polyclonal α-IMA4 (1-210AA)
NUP107-Isoform 1: 106.4 kDa
35Q|A36
925
3.9 kDa
102.4 kDa
WB: Polyclonal α-NUP107 (766-925AA)
IRS2: 137.3 kDa
1122Q|A1123
1338
115.3 kDa
22.0 kDa
WB: Monoclonal α-IRS2 (1035-1136AA)
C
SARS-CoV-2 infected Calu-3 cells
24 h
48 h
Mock
MOI 0.1
MOI 1.0
PTBP1
ΔPTBP1
42 kDa
38 kDa
35 kDa
25 kDa
21 kDa
17 kDa
β-actin
MAP4K5
ΔMAP4K5
70kDa
60 kDa
CREB1
ΔCREB1
YAP1
ΔYAP1
40 kDa
FYCO1
ΔFYCO1
93 kDa
72 kDa
FAF1
ΔFAF1
57 kDa
RPAP1
IMA4
D
SARS-CoV-2 infected Calu-3 cells
3CLpro
Nucleocapsid
Mock
24 h
48 h
MOI 0.1
MOI 1.0
β-actin
N

Figure S4 related to Figures 3 and 4. 3CL^{pro} substrates in bronchial or SARS-CoV-2 infected human lung cells.

(A) Immunoblots for the proteins indicated of lysates from HAECs (N = 5 donors) incubated with 3CL^{pro} or 3CL^{pro}-C145A (1:200 w/w, E:S) for 18 h, 37°C. Anti-FLAG antibody verified that equivalent amounts of 3CL^{pro} and inactive 3CL^{pro}-C145A were incubated.

(B) Schematics of all substrates validated in Figures 3 and 4. 3CL^{pro} cleavage sites identified by TAILS (red arrow) and Edman sequencing (green arrow).

(C) Uncropped immunoblots of substrates in Calu-3 lung epithelial cells infected with SARS-CoV-2 for 24 and 48 h (MOI 0.1 and 1.0, n = 4, n = 2 mock). Proteolytic fragments indicated by apparent molecular weights or Δ.

(D) Immunoblots of 3CL^{pro} and nucleocapsid protein (N) in Calu-3 cells infected with SARS-CoV-2 at a MOI 0.1 and 1.0 for 24 (n = 4) and 48 (n = 4) hpi, or mock-treated (n = 3). Positive control, recombinant (r)3CL^{pro} in control cell lysate. b-actin and b-tubulin loading controls.

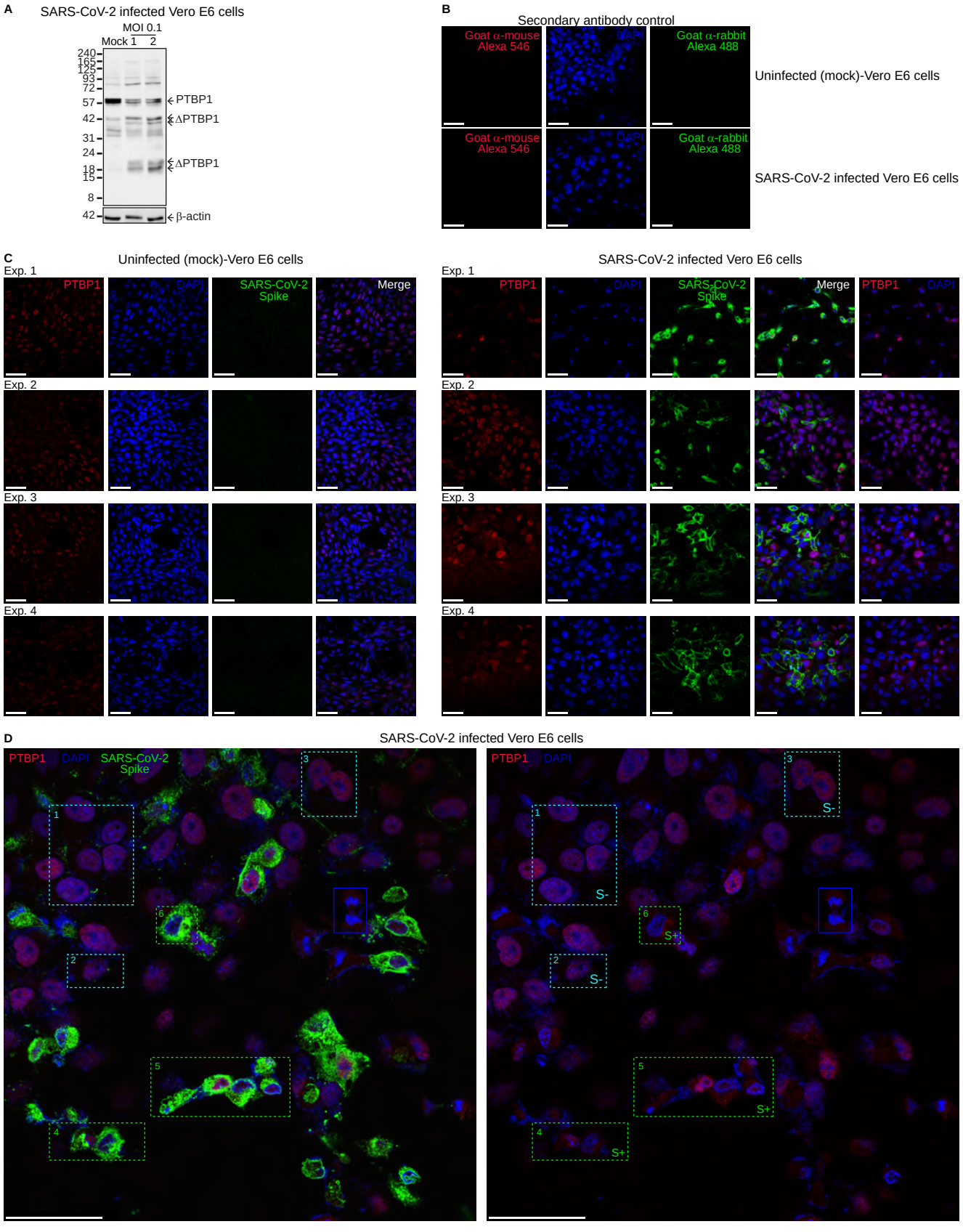

A
SARS-CoV-2 infected Vero E6 cells
MOI 0.1
Mock 1 2
240
165
125
93
72
57
42
31
24
18
15
8
42
PTBP1
ΔPTBP1
ΔPTBP1
β-actin
B
Secondary antibody control
Goat α-mouse Alexa 546
DAPI
Goat α-rabbit Alexa 488
Uninfected (mock)-Vero E6 cells
SARS-CoV-2 infected Vero E6 cells
C
Uninfected (mock)-Vero E6 cells
SARS-CoV-2 infected Vero E6 cells
Exp. 1
Exp. 2
Exp. 3
Exp. 4
PTBP1
DAPI
SARS-CoV-2 Spike
Merge
D
SARS-CoV-2 infected Vero E6 cells
PTBP1
DAPI
SARS-CoV-2 Spike
1
2
3
4
5
6
S-
S+

Figure S5 related to Figure 3. Subcellular localization of PTBP1 in Vero E6 cells infected with SARS-CoV-2.

(A) Immunoblot of PTBP1 and cleaved (Δ) PTBP1 in the SARS-CoV-2-infected Vero E6 cells (MOI 0.1, 48 hpi) used for subcellular localization in B–D. b-actin loading controls.

(B) Secondary antibody controls.

(C) Subcellular localization of PTBP1 in uninfected ($N = 4$, mock) and SARS-CoV-2-infected ($N = 4$) Vero E6 cells. Confocal images of immunofluorescent localization of PTBP1 and SARS-CoV-2 Spike S1 glycoprotein. DAPI, nuclear stain. Scale bar, 50 μm. Far-right panels, merge of PTBP1 and DAPI channels to visualize cytoplasmic *versus* nuclear PTBP1.

(D) Enlarged and uncropped confocal imaging of Figure 3G, Spike-positive (S+) and Spike-negative (S-) (scale bar, 50 μm). Cyan boxes, S- uninfected cells. Green boxes, S+ infected cells. Dark blue-box, cytoplasmic PTBP1 in anaphase cells.

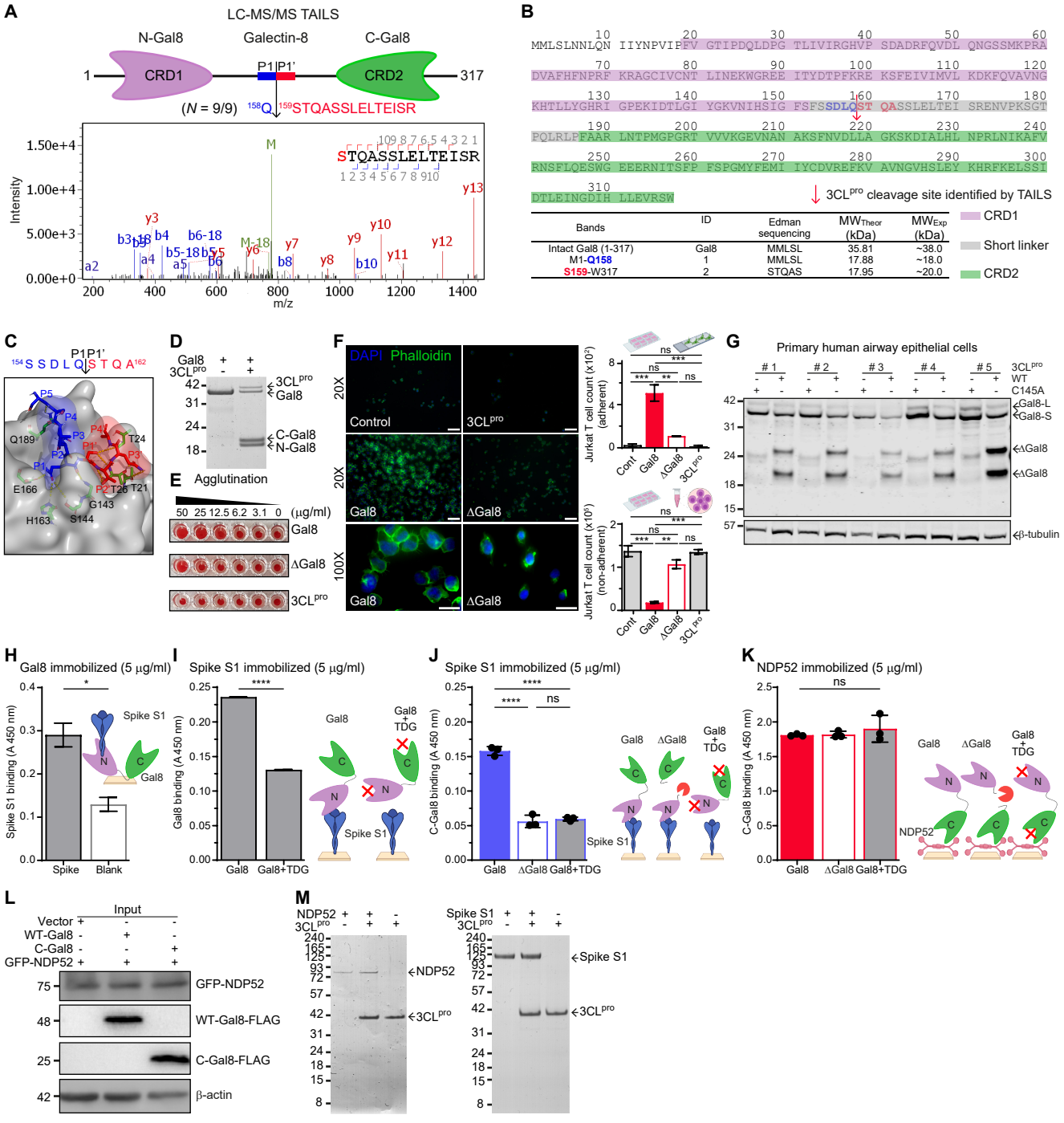

A
LC-MS/MS TAILS
N-Gal8
Galectin-8
C-Gal8
CRD1
CRD2
P1 P1'
(N = 9/9) 158Q 159STQASSLELTEISR
STQASSLELTEISR
Intensity
m/z
B
MMLSLNNLQN IIYNPVIPFV GTIPDQLDPG TLIVIRGHVP SDADRFQVDL QNGSSMKPRA
DVAFHFNPRF KRAGCIVCNT LINEKWGREE ITYDTPFKRE KSFEIVIMVL KDKFQVAVNG
KHTLLYGHRI GPEKIDTLGI YGKVNIHSIG FSFSSDIQST QASSLELTEI SRENVPKSGT
PQLRLPFAAR LNTPMGPGRT VVVKGEVNAN AKSFNVDLLA GKSKDIALHL NPRLNIKAFV
RNSFLQESWG EEERNITSFP FSPGMYFEMI IYCDVREFKV AVNGVHSLEY KHRFKELSSI
DTLEINGDIH LLEVRSW
3CLpro cleavage site identified by TAILS
Bands
ID
Edman sequencing
MWTheor (kDa)
MWExp (kDa)
Intact Gal8 (1-317)
Gal8
MMLSL
35.81
~38.0
M1-Q158
1
MMLSL
17.88
~18.0
S159-W317
2
STQAS
17.95
~20.0
CRD1
Short linker
CRD2
C
P1 P1'
154S S D L Q S T Q A162
Q189
T24
E166
T21
G143
H163
S144
D
Gal8
3CLpro
3CLpro
C-Gal8
N-Gal8
E
Agglutination
50 25 12.5 6.2 3.1 0 (μg/ml)
Gal8
ΔGal8
3CLpro
F
DAPI Phalloidin
Control
3CLpro
Gal8
ΔGal8
20X
100X
Jurkat T cell count (x10²) (adherent)
Jurkat T cell count (x10²) (non-adherent)
Cont
Gal8
ΔGal8
3CLpro
G
Primary human airway epithelial cells
1
2
3
4
5
3CLpro
WT
C145A
Gal8-L
Gal8-S
ΔGal8
β-tubulin
H
Gal8 immobilized (5 μg/ml)
Spike S1 binding (A 450 nm)
Spike S1
Gal8
Spike
Blank
I
Spike S1 immobilized (5 μg/ml)
Gal8 binding (A 450 nm)
Gal8
Gal8+TDG
Spike S1
J
Spike S1 immobilized (5 μg/ml)
C-Gal8 binding (A 450 nm)
Gal8
ΔGal8
Gal8+TDG
Spike S1
K
NDP52 immobilized (5 μg/ml)
C-Gal8 binding (A 450 nm)
Gal8
ΔGal8
Gal8+TDG
NDP52
L
Input
Vector
WT-Gal8
C-Gal8
GFP-NDP52
GFP-NDP52
WT-Gal8-FLAG
C-Gal8-FLAG
β-actin
M
NDP52
3CLpro
Spike S1
3CLpro

Figure S6 related to Figure 5. Loss of galectin-8 biological function and binding to SARS-CoV-2 Spike S1 glycoprotein after cleavage by 3CL^{pro}.
(A) MS/MS spectrum of TAILS neo-N-terminal peptide (red) identifying 3CL^{pro} cleavage site in galectin-8 (Gal8) in $N = 9/9$ BEAS-2B cell experiments.
(B) Reference sequence of recombinant Gal8 (short linker) (1 – 317) #AAF 19370.1. Edman sequencing of 3CL^{pro}-cleaved Gal8 in Figure 5C.
(C) Structural model of Gal8 P5–P1 (blue sticks) and P1'–P4' (red sticks) amino acid residues interacting with the 3CL^{pro} substrate-binding S and S'-subsites, respectively (grey surface).
(D) Cleaved recombinant human Gal8 (ΔGal8) prepared for agglutination (S6E), adhesion (S6F), and binding (Figures 5G and 5J) assays.
(E) Human erythrocyte agglutination assay with intact or ΔGal8, or 3CL^{pro} (control).
(F) Adherent *versus* non-adherent Jurkat T cells in serum-free RPMI containing intact or ΔGal8, or 3CL^{pro} (control) were counted after DAPI/phalloidin staining. Scale bars, 20x: 50 µm; 100x: 20 µm. Mean ± SD, $n = 2$, *** $P \leq 0.001$, ** $P \leq 0.01$, ns $P > 0.05$, one-way ANOVA with Tukey's multiple comparisons test.
(G) Gal8 immunoblots of HAECs lysates from $N = 5$ donors incubated with 3CL^{pro} or 3CL^{pro}-C145A (1:200 w/w, E:S) for 18 h, 37°C. b-tubulin loading control.
(H) ELISA of SARS-CoV-2 Spike S1 glycoprotein binding to immobilized Gal8 (mean ± SD, $n = 2$, $N = 2$, * $P \leq 0.05$, Student's t-test).
(I) Anti-Gal8 antibody ELISA of Gal8 binding to immobilized SARS-CoV-2 Spike S1 glycoprotein in the presence of thiodigalactoside (TDG) (mean ± SD, $n = 2$, $N = 2$, **** $P \leq 0.0001$, Student's t-test).
(J) Anti-C-Gal8 antibody ELISA of intact Gal8, cleaved ΔGal8, or intact Gal8 in the presence of TDG binding to immobilized Spike S1.
(K) Anti-C-Gal8 antibody ELISA of intact Gal8, cleaved ΔGal8, or TDG-inhibited Gal8 binding to immobilized NDP52. One-way ANOVA with Tukey's multiple comparisons test (mean ± SD, $n = 3$, ns $P > 0.05$).
(L) Transfected GFP-NDP52, WT-Gal8-FLAG or C-Gal8 (159-317)-FLAG 3CL^{pro}-cleavage analogue detected in the Hela-cell lysates before immunoprecipitation experiments shown in Figure 5H. Immunoblots were performed with specific anti-FLAG antibody, anti-NDP52 antibody, and with β-actin used as the loading control.
(M) SDS-PAGE of NDP52 and SARS-CoV-2 Spike S1 incubated with recombinant 3CL^{pro}.