The SARS-CoV-2 RNA–protein interactome in infected human cells

Nora Schmidt, Caleb A. Lareau, Hasmik Keshishian, Sabina Ganskih, Cornelius Schneider, Thomas Hennig, Randy Melanson, Simone Werner, Yuanjie Wei, Matthias Zimmer, Jens Ade, Luisa Kirschner, Sebastian Zielinski, Lars Dölken, Eric S. Lander, Neva Caliskan, Utz Fischer, Jörg Vogel, Steven A. Carr, Jochen Bodem and Mathias Munschauer

Characterizing the interactions that SARS-CoV-2 viral RNAs make with host cell proteins during infection can improve our understanding of viral RNA functions and the host innate immune response. Using RNA antisense purification and mass spectrometry, we identified up to 104 human proteins that directly and specifically bind to SARS-CoV-2 RNAs in infected human cells. We integrated the SARS-CoV-2 RNA interactome with changes in proteome abundance induced by viral infection and linked interactome proteins to cellular pathways relevant to SARS-CoV-2 infections. We demonstrated by genetic perturbation that cellular nucleic acid-binding protein (CNBP) and La-related protein 1 (LARP1), two of the most strongly enriched viral RNA binders, restrict SARS-CoV-2 replication in infected cells and provide a global map of their direct RNA contact sites. Pharmacological inhibition of three other RNA interactome members, PPIA, ATP1A1, and the ARP2/3 complex, reduced viral replication in two human cell lines. The identification of host dependency factors and defence strategies as presented in this work will improve the design of targeted therapeutics against SARS-CoV-2.

The rapid spread of a new severe acute respiratory syndrome-related coronavirus (SARS-CoV-2) around the globe has led to a worldwide spike in SARS-like respiratory illness termed coronavirus disease 2019 (COVID-19). To date, more than one million lives have been lost due to COVID-19. A detailed understanding of the molecular interactions and perturbations occurring during SARS-CoV-2 infection is required to understand the biology of SARS-CoV-2 and design therapeutic strategies. SARS-CoV-2 is an enveloped, positive-sense, single-stranded RNA virus that, upon infection of a host cell, deploys a ‘translation-ready’ RNA molecule, which uses the protein synthetic machinery of the host to express a set of viral proteins crucial for replication. Replication of the full-length viral genome and transcription of subgenomic RNAs both involve the synthesis of highly denaturing purification procedure and is ideally suited to the rapid capture and identification of those proteins that bind directly to SARS-CoV-2 RNAs.

To purify SARS-CoV-2 RNAs and the complement of directly crosslinked cellular proteins from infected human cells, we designed a pool of biotinylated DNA oligonucleotides antisense to the positive-sense viral RNA and designed therapeutics against SARS-CoV-2.

Results
Capturing SARS-CoV-2 RNAs in infected human cells. To purify SARS-CoV-2 RNAs and the complement of directly crosslinked cellular proteins from infected human cells, we designed a pool of biotinylated DNA oligonucleotides antisense to the positive-sense SARS-CoV-2 RNA and its subgenomic messenger RNAs. As a cellular system, we selected the human liver cell line HuH7, which is naturally permissive to both SARS-CoV-1 and SARS-CoV-2.

1Helmholtz Institute for RNA-based Infection Research, Helmholtz-Center for Infection Research, Würzburg, Germany. 2School of Medicine, Stanford University, Palo Alto, CA, USA. 3Broad Institute of MIT and Harvard, Cambridge, MA, USA. 4Institute for Molecular Infection Biology, University of Würzburg, Würzburg, Germany. 5Department of Biochemistry, University of Würzburg, Würzburg, Germany. 6Institute for Virology and Immunobiology, Julius-Maximilians-University Würzburg, Würzburg, Germany. 7Department of Biology, MIT, Cambridge, MA, USA. 8Department of Systems Biology, Harvard Medical School, Boston, MA, USA. 9Faculty of Medicine, University of Würzburg, Würzburg, Germany. 10These authors contributed equally: Nora Schmidt, Caleb A. Lareau, Hasmik Keshishian. ✉ e-mail: jochen.bodem@vim.uni-wuerzburg.de; mathias.munschauer@helmholtz-hiri.de
replication\cite{20,21}. SARS-CoV-2 preferentially infects cells in the respiratory tract, but infection of multiple organs, including the liver, has been reported\cite{22}.

To test if our pool of antisense capture probes was suitable for the purification of SARS-CoV-2 RNAs from infected Huh7 cells, we performed RAP–MS 24 h after infection when viral replication levels were high\cite{21}. We implemented a covalent protein capture step after the purification of SARS-CoV-2 RNAs from infected Huh7 cells, we compared the protein content of SARS-CoV-2 RNA purifications to that of an unrelated control ribonucleoprotein complex of known composition. As the control, we used the endogenously expressed human ribonuclease mitochondrial RNA processing (RMRP) RNA and purified both SARS-CoV-2 RNA and RMRP from infected Huh7 cells. RMRP was selected for several reasons: (1) RMRP interacts with approximately ten well-known proteins that serve as an internal control\cite{15,23}; (2) RMRP is not translated; and (3) RMRP does not originate from SARS-CoV-2 RNA made up 93 and 92% of all mapped reads in 2 highly correlated replicate experiments ($r = 0.994$; Extended Data Fig. 1b,c).

To identify proteins that specifically interact with SARS-CoV-2 RNAs as opposed to non-specific background proteins, we compared the protein content of SARS-CoV-2 RNA purifications to that of an unrelated control ribonucleoprotein complex of known composition. As the control, we used the endogenously expressed human ribonuclease mitochondrial RNA processing (RMRP) RNA and purified both SARS-CoV-2 RNA and RMRP from infected Huh7 cells. RMRP was selected for several reasons: (1) RMRP interacts with approximately ten well-known proteins that serve as an internal control\cite{15,23}; (2) RMRP is not translated; and (3) RMRP does not originate from SARS-CoV-2 RNA made up 93 and 92% of all mapped reads in 2 highly correlated replicate experiments ($r = 0.994$; Extended Data Fig. 1b,c).
not globally bind to mRNA. Hence, RMRP-binding proteins are distinct from the group of proteins expected to bind to SARS-CoV-2 RNAs, making it an ideal control for the discovery of unknown interactors. Further, the purification of SARS-CoV-2 RNA and RMRP from infected cells avoids biases resulting from widespread changes in the host cell proteome induced by viral infection.

On average, approximately 90% of all crosslinked RNA fragments originated from the SARS-CoV-2 genome in SARS-CoV-2 RNA purifications, while more than 99% of crosslinked RNA fragments from RMRP purifications mapped to the human genome (Extended Data Fig. 1d). Western blot analysis confirmed the specific capture of SARS-CoV-2 nucleocapsid protein only in SARS-CoV-2-purified samples (Extended Data Fig. 1e). The RMRP component POP1 was detected only in RMRP purifications. Together, these experiments verify the high specificity of our approach for capturing the desired RNAs and the proteins that directly bind to them.

An atlas of SARS-CoV-2 RNA–protein interactions in human cells. Next, we subjected proteins purified with RMRP and SARS-CoV-2 RNAs to tandem mass tag (TMT) labelling and relative quantification by liquid chromatography coupled with tandem mass spectrometry (LC–MS/MS). In two replicate experiments, we identified 699 proteins, of which 583 were detected with 2 or more unique peptides (Supplementary Table 1 and Methods). As shown in Fig. 1b, we found five known RMRP components among the ten most significantly enriched proteins in RMRP purifications.

Next, we analysed proteins enriched in SARS-CoV-2 RNA purifications and found 15 SARS-CoV-2 proteins, 6 of which were among the 20 most significantly enriched proteins (Fig. 1b,c). In addition to 5 viral proteins translated from distinct open reading frames (ORFs), 10 of the 16 non-structural proteins (NSPs), which are derived from a precursor polyprotein, were detected by RAP–MS. As expected, the SARS-CoV-2 nucleocapsid protein, which binds the viral RNA, was one of the two most significantly enriched viral proteins, followed by several known viral RNA binders, such as the endoribonuclease NSP15 (ref. 23), the RNA-dependent RNA polymerase (RdRP) NSP12 (ref. 24), the methyltransferase NSP16 (ref. 25), the RNA-binding protein NSP9 (ref. 28), the capping factor NSP10 (ref. 27), the primase NSP8 (ref. 23), the 5′-UTR binder NSP1 (ref. 24) and the multifunctional protein NSP3 (ref. 23). Remarkably, NSP3 and the most strongly enriched protein in our data, NSP6, were required for the formation of double-membrane vesicles and both proteins are candidate constituents of a molecular pore complex involved in the export of RNA from coronavirus double-membrane vesicles. We also found ORF3a, which binds the 5′-end of the SARS-CoV-1 genome, as well as ORF9b and the S and M proteins among strongly enriched candidates. While M is known to interact with the nucleocapsid protein, a model for genomic RNA packaging further suggests a possible RNA-binding function for M. An RNA-binding activity of S was not previously reported. While S covers the surface of the viral envelope, it has a transmembrane domain and an intracellular tail, making it conceivable that S may indeed contact viral RNA.

Discovery of 104 human proteins that bind SARS-CoV-2 RNA. We next focused on the human proteins enriched in SARS-CoV-2 RNA purifications. We identified 276 proteins with a positive log, fold change. Of these, 57 were significantly enriched (adjusted P < 0.05, two-tailed t-test), which we subsequently defined as the set of core SARS-CoV-2 RNA interacting proteins (Fig. 1c). Additionally, we also defined an expanded SARS-CoV-2 RNA interactome using a relaxed false discovery rate (FDR) of less than 20% (Fig. 1c).

The expanded SARS-CoV-2 RNA interactome encompassed 104 human proteins and included 13 SARS-CoV-2-encoded proteins. The vast majority of the human RNA interactome proteins (100 proteins, 96%) have been identified previously in system-wide studies aimed at capturing proteins that crosslink to RNA (Supplementary Table 2). Comparing this expanded SARS-CoV-2 RNA interactome with the poly(A)-RNA interactome in Huh7 cells, revealed high overlap between both datasets (69 proteins, 66%) (Fig. 2a). Next, we compared our direct SARS-CoV-2 RNA interactome with proteins that directly or indirectly associate with the RNA genomes of Dengue and Zika viruses in Huh7 cells. Sixty-six proteins (63%) of the expanded SARS-CoV-2 RNA interactome also associated with the Dengue and Zika virus RNAs, while 38 proteins (36.5%) were unique SARS-CoV-2 RNA binders (Fig. 2a). Since coronaviruses form replication/transcription complexes (RTCs), we also compared the expanded SARS-CoV-2 RNA interactome to the protein content of murine coronavirus RTCs and found 64 shared proteins (Supplementary Table 2).

Finally, only 10 of the 332 human proteins that bind recombinant SARS-CoV-2 proteins in uninfected cells also bound directly to viral RNA in infected cells (Supplementary Table 2). These results highlight the importance of discriminating between protein–protein and RNA–protein interactions when dissecting the biology of SARS-CoV-2.

Biological functions of SARS-CoV-2 RNA-binding proteins. To analyse the biological functions of SARS-CoV-2 RNA binders, we performed a hypergeometric gene ontology (GO) enrichment analysis on the expanded SARS-CoV-2 RNA interactome. We observed strong enrichment for GO terms linked to translational initiation (GO:0006413), nonsense-mediated decay (GO:0000184), signal-recognition particle-dependent cotranslational protein targeting to the membrane (GO:0006614) and viral transcription (GO:0019083) (Fig. 2b and Supplementary Table 3). Consistent with the enrichment of these GO terms, the importance of subgenomic mRNA translation at the endoplasmic reticulum membrane is well established for coronaviruses. Further, nonsense-mediated mRNA decay was recently described as an antiviral mechanism targeting coronavirus RNAs.

In agreement with the crucial role of mRNA translation, the expanded SARS-CoV-2 RNA interactome included 19 ribosomal proteins and 12 translation factors. Among the translation factors, the eukaryotic translation initiation factor 4F (EIF4F) components EIF4G1 and EIF4B are regulated by mammalian target of rapamycin (mTOR) signalling. EIF4B is important for recruiting the 40S subunit to mRNA and both the phosphatidylinositol-3-kinase (PI3K)/mTOR and mitogen-activated protein kinase (MAPK) pathways target EIF4B to control its activity. Inhibition of PI3K/Akt/mTOR signalling has been demonstrated to suppress SARS-CoV-2 replication in Caco2 cells.

To examine the connectivity of the identified SARS-CoV-2 RNA-binding proteins and their relationship to virus-associated biological processes systematically, we constructed a protein–protein association network using our expanded RNA interactome (Fig. 2c and Supplementary Table 4). We observed a striking enrichment for physical interactions when comparing the total connectivity among RNA interactome proteins to the connectivity of equally sized networks sampled from expressed proteins (Extended Data Fig. 2 and Methods; permutation test P < 2.2 × 10^-16). In addition to ribosomal proteins and translation factors, many virus-associated RNA-binding proteins are prominently represented in this network (Fig. 2c). Since RNA-binding proteins can be regulated by phosphorylation, we intersected our expanded SARS-CoV-2 RNA interactome with a recent phosphoproteomic dissection of SARS-CoV-2-infected cells, highlighting 30 proteins that might be dynamically phosphorylated in response to SARS-CoV-2 infection (Fig. 2d).

We next integrated known drug–target interactions within this network and identified 23 SARS-CoV-2 RNA interactome proteins that can be targeted with existing compounds, including...
peptidyl-prolyl cis-trans isomerase A (PPIA), actin-related protein 2 (ACTR2; henceforth ARP2), sodium/potassium -transporting ATPase subunit alpha-1 (ATP1A1), coflin-1 (CFL1) and epidermal growth factor receptor (EGFR) (Fig. 2e). Notably, EGFR is a known target of compounds that inhibit SARS-CoV-2 replication7,8,10.

**Identification of activated host response pathways.** To gain deeper insight into host response pathways activated upon SARS-CoV-2 infection, we globally measured protein abundance changes in infected cells. We performed triplicate MS experiments on SARS-CoV-2-infected and uninfected Huh7 cells and identified 10,956 proteins with 2 or more unique peptides (Fig. 3a and Supplementary Table 5). Among the detected proteins, 4,578 proteins were regulated (adjusted \(P < 0.05\), two-tailed \(t\)-test) after 24 h of SARS-CoV-2 infection, which is consistent with widespread proteome regulation and agrees well with previously published data (Extended Data Fig. 3a)8,9. As expected, proteome samples clustered according to their infection status in a principal component analysis (Extended Data Fig. 3b). Among differentially expressed proteins, we detected 13 viral proteins and 56 proteins from our expanded SARS-CoV-2 RNA interactome (Fig. 3a).

We next performed gene set enrichment analysis (GSEA) using our proteome abundance measurements. Among the most
**Fig. 3 | Connecting the SARS-CoV-2 RNA interactome to perturbations in host cells.**

**a,** Volcano plot of proteome abundance measurements in SARS-CoV-2-infected and uninfected Huh7 cells 24 h post-infection (n = 3) (Supplementary Table 5). Adjusted P value: two-tailed moderated t-test. SARS-CoV-2-encoded proteins are shown in magenta; human SARS-CoV-2 RNA interactome proteins are shown in teal; interferon response-related proteins are shown in purple. **b,** GSEA for the global proteome abundance measurements shown in **a.** Selected gene sets are shown; the full table displaying additional enriched gene sets is provided in Extended Data Fig. 3c. Statistical test: Kolmogorov–Smirnov test with Benjamini–Hochberg adjustment. NES, normalized enrichment score. **c,** Protein–protein association network of core SARS-CoV-2 RNA interactome proteins and their connections to differentially regulated proteins in SARS-CoV-2-infected cells based on curated interactions in STRING v.11 (ref. 96). Upregulated proteins are shown in light grey; downregulated proteins are shown in dark grey. Circle sizes scale to the number of connections of each interactome protein. Selected GO enrichments for network communities are shown in the transparent circles (Methods). Full GO term analysis is provided in Supplementary Table 8.
significantly enriched hallmark gene sets were ‘TGFB signalling’, ‘TNF-α signalling via NF-kB’, ‘interferon (IFN)-γ response’ and ‘IL-6 JAK STAT3 signalling’ (Fig. 3b and Extended Data Fig. 3c), which is consistent with the induction of broad pro-inflammatory and antiviral responses in infected cells. Further, we observed significant enrichment of the gene sets ‘GO regulation of MAPK cascade’, ‘GO positive regulation of MAPK activity’ and ‘GO response to type I interferon’ (Supplementary Table 6). Recent evidence indicates that these pathways are indeed highly relevant in the context of SARS-CoV-2 infections55,26. Inhibition of growth factor signalling through the MAPK pathway, which responds to and controls the production of pro-inflammatory cytokines, including TNFα and IL-6, was shown to modulate SARS-CoV-2 replication26.

In agreement with recent transcriptome studies6,7, our proteome data suggest activation of interferon signalling upon SARS-CoV-2 infection. Among interferon-related genes, we observed significant upregulation of several major components of IFN signalling pathways, including STAT1 and IRF9, which together with STAT2 make up the interferon stimulated gene factor 3 (ISGF3) complex, their upstream components TYK2 and JAK1, as well as their downstream targets IFIT1, IFIT3, IFITM3, OAS2 and ISG15 (Fig. 3a).

Other strongly upregulated IFN-related genes include BST2, SP110, UBE2L6, ADAR and TGF1 (Supplementary Table 5). Notably, many SARS-CoV-2 RNA interactome members are linked to the IFN response. These include the strongly enriched PUM1 (ref. 49), YBX1 (ref. 50), SYNCRIP51, G3BP1 (refs. 43,44), G3BP2 (refs. 8,45), EIF4B52, MOV10 (ref. 53), CAPRIN1 (ref. 54), DDX3X55, LSM14A56, RyDEN57, STRAP58, ANXA1 (ref. 59), DDX1 (ref. 60), PCBP2 (ref. 61), CAPRIN1 (ref. 62), DDX3X52, LSM14A53, YWHAZ63, N haunt targets, pathways, including TGF-β signalling64, IFIT3, IRF9, and STAT2, which together with IFITM3, OAS2 and IFITM3 were highly susceptible to infections with different pathogens64,65. Notably, CNBP-deficient animals are highly susceptible to virus-induced cell death, which suggests that CNBP may act as an antiviral regulator. To corroborate the functional importance of CNBP in SARS-CoV-2 infections, we generated polyclonal Hu7 CNBP knockout cell lines using CRISPR-Cas9 (Fig. 4b). We infected CNBP knockout cells with SARS-CoV-2 and noted significantly elevated levels of intracellular viral RNA compared to matched Hu7 control cells (Fig. 4b).

To confirm the direct physical engagement of SARS-CoV-2 RNAs by CNBP, we performed enhanced crosslinking and immunoprecipitation (eCLIP) in SARS-CoV-2-infected Hu7 cells and quantified the enrichment of CNBP peaks relative to size-matched input libraries7. First, we analysed CNBP binding to the human transcriptome. Consistent with earlier reports7, CNBP bound to protein-coding transcripts and displayed a preference for binding within the coding sequence (CDS) of mRNAs (Fig. 4c,d). A large number of transcripts bound by CNBP in SARS-CoV-2-infected cells were previously reported as CNBP targets (approximately 46%; Supplementary Table 9). We next analysed CNBP binding to SARS-CoV-2 RNA and observed several strongly enriched binding sites in the viral genome (Fig. 4e). These data provide strong evidence for a direct interaction between CNBP and SARS-CoV-2 RNAs in infected cells and validate that RAP–MS indeed identifies direct RNA binders. Further, the finding that CNBP preferentially associates with the CDS of mature mRNAs lends credibility to its previously proposed role as a translational regulator2 in addition to its function in regulating pro-inflammatory cytokines.

**LARP1 binds genomic and subgenomic SARS-CoV-2 RNAs.** Other than CNBP, two members of the L-related protein (LARP) family, namely LARP1 and LARP4, were strongly enriched in SARS-CoV-2 RNA purifications. While LARP1 did not quite meet our significance cut-off, both LARP1 and LARP4 were among the 15 host proteins with the strongest enrichment based on overall effect size, indicating that LARP1 is very likely a SARS-CoV-2 RNA binder. Additionally, LARP1 was detected among protein–protein interactors of the nucleocapsid protein in uninfected cells16.

Given that LARP1 is a major downstream target of mammalian target of rapamycin complex 1 (mTORC1) (refs. 74,75) and inhibition of PI3K/Akt/mTOR was recently shown to inhibit SARS-CoV-2 replication in Caco2 cells16, we sought to characterize the LARP1-SARS-CoV-2 axis in greater detail. We performed eCLIP7 to map direct physical interactions between LARP1 and its RNA targets. LARP1 predominantly bound protein-coding transcripts and we observed most of the enriched peaks in the CDS, followed by 5'-UTR and 3'-UTR sequences (Fig. 5a). Previous work suggested
that LARP1 binds the 7-methylguanosine triphosphate (m7Gppp) moiety of the cap and the adjacent 5′-terminal oligopyrimidine (5′ TOP) motif of mRNAs to regulate their translation. Consistent with this finding, our eCLIP data revealed a strong enrichment of 5′-proximal nucleotides in 5′-UTR sequences and we recovered an oligopyrimidine motif reminiscent of TOP-like sequences in approximately 30% of all bound 5′-UTRs (Fig. 5b,c). Out of 112 mRNAs that are regulated by LARP1 downstream of mTOR, we observed LARP1 binding to 84 mRNAs (75%; Supplementary Table 10). In line with the known regulatory functions of LARP1 (ref. 76), LARP1 target transcripts were most strongly enriched for GO terms linked to translational regulation (Supplementary Table 11). Together, these data demonstrate that our eCLIP experiments recovered known regulatory interactions of LARP1.

Having confirmed the quality of our eCLIP experiment on host RNAs, we next characterized LARP1 binding to SARS-CoV-2 RNAs and found several regions of enrichment that coincided with oligopyrimidine sequences (Fig. 5d). Notably, we observed LARP1 binding to the first 70 nucleotides at the 5′-end of the SARS-CoV-2 genome, which corresponds to the viral 5′-leader sequence and contains a TOP-like motif instance (Fig. 5d). Binding to the 5′-leader, which is present in all viral subgenomic mRNAs, suggests a direct association of LARP1 with subgenomic mRNAs.

LARP1 represses SARS-CoV-2 replication. To determine the impact of LARP1 depletion on SARS-CoV-2 replication, we generated four clonal LARP1 knockout cell lines using CRISPR–Cas9 in HEK293 cells (Extended Data Fig. 4a). We infected cells with...
SARS-CoV-2 and measured intracellular viral RNA levels and the production of infectious virus. Compared to wild-type (WT) cells, LARP1 knockout cells displayed approximately fivefold higher levels of intracellular viral RNA and a similar increase in the production of infectious virus (Fig. 5c). Conversely, transient overexpression of LARP1 fused to green fluorescent protein (GFP) in WT cells led to a significant reduction of viral RNA and infectious virus when compared to GFP expression alone (Fig. 5f and Extended Data Fig. 4b). Next, we complemented LARP1 knockout cells with transiently expressed LARP1–GFP proteins (Fig. 5g, Extended Data Fig. 4c). In all knockout cell lines, we observed a clear reduction in intracellular viral RNA that approached WT levels when compared to cells transfected with GFP alone. These experiments established that LARP1 functions as a repressor of SARS-CoV-2 replication in infected human cells.

RyDEN suppresses ribosomal frameshifting during SARS-CoV-2 RNA translation. LARP1 interacts with PABPC1 and both LARP1 and PABPC1 have been proposed to reside in the same ribonucleoprotein complex with RyDEN, all of which were enriched in RAP–MS experiments. In addition to being an IFN-induced protein, RyDEN suppresses Dengue virus production in infected cells and inhibits programmed -1 ribosomal frameshifting (-1FS) in human immunodeficiency virus type 1 (HIV-1) infections.

In coronaviruses, production of RdRP requires translation of the ORF1b gene, which is controlled by -1FS. For SARS-CoV-2, it is presently unknown if the efficiency of -1FS is important for the viral life cycle. To dissect if RyDEN can modulate the frequency of -1FS in SARS-CoV-2, we generated a dual-colour fluorescence reporter system to quantify frameshifting efficiency in response to RyDEN induction, as seen upon SARS-CoV-2 infection (Extended Data Fig. 4d and Methods). Using a reporter containing the HIV-1 frameshift element as a positive control, we confirmed that overexpression of RyDEN fused to enhanced cyan fluorescent protein (eCFP) suppressed -1FS when compared to eCFP expression alone (Fig. 5h). Importantly, overexpression of RyDEN also led to a significant reduction of -1FS during translation of the SARS-CoV-2 frameshift element (Fig. 5h). Together, our results show that RyDEN is induced upon SARS-CoV-2 infection, associates with the SARS-CoV-2 RNA in infected cells and modulates the efficiency of SARS-CoV-2 -1FS.

Pharmacological inhibition of interactome proteins restricts viral replication. Next, we tested if targeting the SARS-CoV-2 RNA interactome and its associated pathways with known inhibitors is effective in restricting viral replication. We selected four inhibitors that target component of our expanded RNA interactome: PPIA; ARP2; ATP1A1 and DDX3X. While DDX3X is a DEAD-box RNA helicase and canonical RNA-binding protein, PPIA, ARP2 and ATP1A1 are non-classical RNA binders that are nonetheless robustly detected among RNA-binding proteins in Huh7 cells. In addition to Huh7 cells, we evaluated all inhibitors in Calu3 cells, a human lung epithelial cell line that is naturally susceptible to SARS-CoV-2 infection.

We observed a dose-dependent inhibition of intracellular viral RNA expression accompanied by a reduction in the production of infectious virus for the PP1A inhibitor cyclosorin A (Extended Data Fig. 5a,b), the ARP2/3 complex inhibitor CK-548 and the ATP1A1 inhibitor ouabain (Fig. 6a,b). The observed effect was highly consistent between Calu3 and Huh7 cells (Fig. 6a,b). While CK-548 treatment reduced cell viability at the highest concentration in Huh7 cells, we did not observe such effects at identical concentrations in Calu3 cells. All other efficacious inhibitors had no apparent effect on cell viability (Extended Data Fig. 5c,d). Unlike the three aforementioned compounds, inhibition of DDX3X only led to a moderate reduction of intracellular viral RNA and infectious virus in Calu3 cells at the highest concentration (Fig. 6a,b).

Beyond inhibiting direct RNA binders, we also targeted mTORC1, the upstream regulatory complex that controls LARP1 activity. Consistent with LARP1 restricting SARS-CoV-2 replication, we observed that inhibiting mTORC1/2 resulted in reduced viral replication in Huh7 and Calu3 cells (Fig. 6a,b). These findings agree well with previous results showing that mTORC1 phosphorylates LARP1, which leads to a translational de-repression of LARP1 target mRNAs. Indeed, recent phosphoproteomic surveys demonstrate that LARP1 undergoes dynamic phosphorylation in response to SARS-CoV-2 infection. Inhibition of another upstream regulator, TANK-binding kinase 1, which interacts with the SARS-CoV-2 RNA binders DDX3X and ANXA1, increased the levels of viral RNA and infectious virus in A549-ACE2 cells, but did not show a consistent effect in Huh7 or Calu3 cells (Extended Data Fig. 5a,b). Together, our experiments demonstrate that RNA interactome proteins represent viable targets for inhibiting SARS-CoV-2 replication. The SARS-CoV-2 RNA interactome provides valuable starting points for future mechanistic studies and may help developing new antiviral approaches for COVID-19.

Discussion

Decoding how the RNA genomes of pathogenic RNA viruses interface with the host cell proteome has been a long-standing challenge. In this study, we provide detailed molecular insights into the identity of host factors and cellular machinery that directly and specifically bind SARS-CoV-2 RNAs during infection of human cells. We integrate CRISPR perturbation data and perform genetic and pharmacological validation experiments that together suggest functional roles for 18 RNA interactome proteins in SARS-CoV-2 infections.
Beyond identifying proteins that bind SARS-CoV-2 RNAs, we globally mapped where CNBP and LARP1 contact viral and human RNA and report binding preferences that are consistent with previously described regulatory functions of both proteins. While we show that CNBP acts as an antiviral factor, it remains to be determined if its role as a regulator of mRNA translation or its effect
on cytokine expression is critical for this function. We provide strong genetic evidence for a functional role of LARP1 in restricting SARS-CoV-2 replication. Remarkably, the SARS-CoV-2 5' leader contains a TOP-like sequence motif that is bound by LARP1 in infected cells. While the TOP-like sequence is still several nucleotides away from the 5' end of the SARS-CoV-2 leader, it is tempting to speculate that binding of LARP1 would negatively influence translation of SARS-CoV-2 RNAs similar to LARP1-mediated translational repression of host-encoded 5' TOP mRNAs.

In addition to genetic perturbation, we inhibited SARS-CoV-2 RNA binders pharmacologically. Notably, all host proteins and complexes that are effectively targeted by these inhibitors have previously been linked to viral diseases: (1) PPIA is involved in protein folding and has a well-documented impact on the replication of viruses\(^{60,61}\). Its direct interaction with SARS-CoV-2 RNA expand these previously described functions. While the PPIA inhibitor cyclosporin A has immunosuppressive properties, the non-immunosuppressive cyclosporin A analogue alisporivir may offer greater translational potential\(^{62}\); (2) a role for the RNA-binding metabolic enzyme ATP1A1\(^{63}\) in coronavirus and respiratory syncytial virus infections has been reported previously\(^{64,65}\). ATP1A1 had a significant effect on virus-induced cell death in a SARS-CoV-2 CRISPR perturbation screen\(^{66}\). Hence, both genetic and pharmacological evidence point to ATP1A1 as an important SARS-CoV-2 host factor; (3) ARP2 is part of the actin-related protein 2/3 complex and contributes to regulating cell shape and motility, which can affect intracellular pathogens\(^{67}\). ARP2 has been identified as a respiratory syncytial virus host factor and is involved in filopodia formation\(^{68}\). Recent work demonstrated that SARS-CoV-2 infection induced a dramatic increase in filopodia and viral particles localized to these actin-rich protrusions\(^{69}\).

In addition to the aforementioned factors, we observed various other notable proteins among SARS-CoV-2 RNA binders. These include vesicle trafficking proteins (SCFD1, USO1, RAB1A, RAB6D, RAB6A, RAB7A, GDI2), cytoskeleton regulators (ARP2, CFL1, PFN1, ACTA1), RNA editing cofactors (RBM47, A1CF) and subunits of a transfer RNA-splicing ligase complex (DDX1, RTCB). Our work highlights opportunities for targeting proteins or pathways linked to the SARS-CoV-2 RNA interactome to interfere with viral infection. We believe that our approach provides a general roadmap for dissecting the biology of RNA viruses and the interactions between hosts and pathogens at the molecular level.

**Methods**

**Tissue culture.** We maintained HuH7, Calu3, HEK293, ACE2-A549 (a generous gift from A. Pichlmaier) and TMPRSS2-Vero E6 cells (a generous gift from S. Pöhlmann) in DMEM medium (Thermo Fisher Scientific) supplemented with 10% heat-inactivated FCS (Thermo Fisher Scientific) and 100 U ml\(^{-1}\) penicillin. Cells were grown at 37 °C and 5% CO\(_2\). Heat-inactivated FCS (Thermo Fisher Scientific) was used in Calu3 and ACE2-A549 cells.

**Generation of LARP1 knockout cell lines using CRISPR–Cas9.** To generate the LARP1 CRISPR knockout cells, we used the pX335-U6-chimeric-BB-CBh-hSpCas9 9n(D10A) nickase (a generous gift from F. Zhang) together with GTTGGGTGGCGGATTTACGGGT and GCCACCCAAAAGCACATGA as guide sequences. HEK293 cells were transfected with TransIT-T2 (Mirus Bio) and selected with 2 µg ml\(^{-1}\) of puromycin in DMEM for 48h and with 1 µg ml\(^{-1}\) for another 48h. We picked single colonies and screened for LARP1 deletion by western blotting.

Plasmids for LARP1 overexpression were generated using C-terminal Myc-DDK-tagged human LARP1 (NM_015315), which was purchased from OriGene and subcloned into pEGFP-C1 retaining the C-terminal Myc-DDK tag.
Generation of CNBP knockout cell lines using CRISPR-Cas9. A total of 2.5×10^5 Huh7 cells per well were seeded in a 6-well plate and transfected the next day with 2.5 μg of a commercially available CNBP CRISPR-Cas9 knockout plasmid (catalogue no. sc-404990; Santa Cruz Biotechnology) and Lipofectamine 2000 (Thermo Fisher Scientific) per 1 μg of DNA. A plasmid containing a puromycin resistance gene was cotransfected as the selection marker. Control cells were transfected only with the puromycin resistance plasmid. Successfully transfected cells were selected with puromycin (5 μg/ml) starting at 24 h post-transfection for 2 d. CNBP expression in polyclonal cell populations was analysed by western blot.

**Virus production.** We used previously described patient-derived SARS-CoV-2 of streptomycin and 100 mg ml^-1^ of penicillin. Control samples that were used for RMRP purifications. SARS-CoV-2 infection was carried out as described previously with the following modifications: to capture endogenous SARS-CoV-2 RNAs, we designed and synthesized 5’ biotinylated 90-mer DNA oligonucleotides (Integrated DNA Technologies) antisense to the complementary strands of 67 probes such that one probe binding site occurred roughly every 400 bases in the approximately 30-kilobase (kb) SARS-CoV-2 genome and excluded regions of 0.005 plaque-forming units (p.f.u.) per cell for virus propagation. After 1 h of incubation at 37 °C, the inoculum was removed and fresh DMEM supplemented with 5% FCS; 100 U ml^-1^ of penicillin were added to the cells. Cells were washed with PBS and incubated with the inoculum for 1 h at 37 °C. The inoculum was removed and fresh DMEM supplemented with 5% FCS, 100 μg ml^-1^ of streptomycin and 100 mg ml^-1^ of penicillin were added to the cells.

**RAP-MS.** Rap-MS was carried out as described previously with the following modifications: to capture endogenous SARS-CoV-2 RNAs, we designed and synthesized 5’ biotinylated 90-mer DNA oligonucleotides (Integrated DNA Technologies) antisense to the complementary strands of 67 probes such that one probe binding site occurred roughly every 400 bases in the approximately 30-kilobase (kb) SARS-CoV-2 genome and excluded regions of 0.005 plaque-forming units (p.f.u.) per cell for virus propagation. After 1 h of incubation at 37 °C, the inoculum was removed and fresh DMEM supplemented with 5% FCS; 100 U ml^-1^ of penicillin were added to the cells. Cells were washed with PBS and incubated with the inoculum for 1 h at 37 °C. The inoculum was removed and fresh DMEM supplemented with 5% FCS, 100 μg ml^-1^ of streptomycin and 100 mg ml^-1^ of penicillin were added to the cells.

**Proteome analyses of SARS-CoV-2-infected cells.** For the proteome measurements, we expanded the Huh7 cells to two 10-cm tissue culture plates per replicate. Cells were infected with a previously described SARS-CoV-2 isolate at an MOI of 10 and incubated for 24 h before being collected. Three process isolates covalently crosslinked to proteins purified with RAP–MS, we carried out as described previously with the following modifications: to capture endogenous SARS-CoV-2 RNAs, we designed and synthesized 5’ biotinylated 90-mer DNA oligonucleotides (Integrated DNA Technologies) antisense to the complementary strands of 67 probes such that one probe binding site occurred roughly every 400 bases in the approximately 30-kilobase (kb) SARS-CoV-2 genome and excluded regions of 0.005 plaque-forming units (p.f.u.) per cell for virus propagation. After 1 h of incubation at 37 °C, the inoculum was removed and fresh DMEM supplemented with 5% FCS; 100 U ml^-1^ of penicillin were added to the cells. Cells were washed with PBS and incubated with the inoculum for 1 h at 37 °C. The inoculum was removed and fresh DMEM supplemented with 5% FCS, 100 μg ml^-1^ of streptomycin and 100 mg ml^-1^ of penicillin were added to the cells.

**Quantification and identification of peptides and proteins.** For the proteome analysis, we expanded the Huh7 cells to two 10-cm tissue culture plates per replicate. Cells were infected with a previously described SARS-CoV-2 isolate at an MOI of 10 and incubated for 24 h before being collected. Three process isolates covalently crosslinked to proteins purified with RAP–MS, we carried out as described previously with the following modifications: to capture endogenous SARS-CoV-2 RNAs, we designed and synthesized 5’ biotinylated 90-mer DNA oligonucleotides (Integrated DNA Technologies) antisense to the complementary strands of 67 probes such that one probe binding site occurred roughly every 400 bases in the approximately 30-kilobase (kb) SARS-CoV-2 genome and excluded regions of 0.005 plaque-forming units (p.f.u.) per cell for virus propagation. After 1 h of incubation at 37 °C, the inoculum was removed and fresh DMEM supplemented with 5% FCS; 100 U ml^-1^ of penicillin were added to the cells. Cells were washed with PBS and incubated with the inoculum for 1 h at 37 °C. The inoculum was removed and fresh DMEM supplemented with 5% FCS, 100 μg ml^-1^ of streptomycin and 100 mg ml^-1^ of penicillin were added to the cells.

**LC–MS/MS analysis (RAP–MS and proteome).** For the proteome analysis, we expanded the Huh7 cells to two 10-cm tissue culture plates per replicate. Cells were infected with a previously described SARS-CoV-2 isolate at an MOI of 10 and incubated for 24 h before being collected. Three process isolates covalently crosslinked to proteins purified with RAP–MS, we carried out as described previously with the following modifications: to capture endogenous SARS-CoV-2 RNAs, we designed and synthesized 5’ biotinylated 90-mer DNA oligonucleotides (Integrated DNA Technologies) antisense to the complementary strands of 67 probes such that one probe binding site occurred roughly every 400 bases in the approximately 30-kilobase (kb) SARS-CoV-2 genome and excluded regions of 0.005 plaque-forming units (p.f.u.) per cell for virus propagation. After 1 h of incubation at 37 °C, the inoculum was removed and fresh DMEM supplemented with 5% FCS; 100 U ml^-1^ of penicillin were added to the cells. Cells were washed with PBS and incubated with the inoculum for 1 h at 37 °C. The inoculum was removed and fresh DMEM supplemented with 5% FCS, 100 μg ml^-1^ of streptomycin and 100 mg ml^-1^ of penicillin were added to the cells.
The eCLIP experiments, as described in the eCLIP library preparation protocol\(^\text{71}\), starting with the reverse phenol-chloroform extraction. All subsequent manipulation steps were carried out with 2.5 mM of tris(2-carboxyethyl)phosphine (TCEP) and sodium deoxycholate. We then washed the beads twice in 1 ml of T4 FastAP buffer (10 mM of Tris pH 7.5, 10 mM of NaCl, 2 mM of EDTA, 2% (v/v) Triton X-100, 0.2% NP-40, 0.2% sodium deoxycholate and incubated each washing step for 5 min at 37 °C. These heated washing steps were followed by two additional washes in the same buffer at room temperature. Subsequently, beads were rinsed on the magnet in 1 ml of FastAP buffer. Next, end repair was carried out by resuspending the beads in 50 μl of FastAP mix (39 μl of H₂O, 5 μl of FastAP buffer (Thermo Fisher Scientific), 1 μl of murine RNase inhibitor, 5 μl of FastAP enzyme (Thermo Fisher Scientific) and incubating for 20 min at 37 °C. In the meantime, we prepared 150 μl of T4 polyuridylate dilution (PUD) mix (120 μl of H₂O, 10 μl of T4 PUD buffer (New England Biolabs), 1 μl of murine RNase inhibitor, 7 μl of T4 PNK, 1 μl of TURBO DNase), which was added to the FastAP reaction and incubated for another 20 min at 37 °C. After end repair, we washed the beads once in modified RTL buffer and twice in detergent wash buffer (20 mM of Tris pH 7.5, 50 mM of NaCl, 0.2% Triton X-100, 0.2% NP-40, 0.2% sodium deoxycholate). Then, we resuspended the beads in the magnetic beads lysis buffer (50 μl of Tris-HCl pH 7.5, 10 μl of MgCl₂, 0.1 mM of EDTA, 1% (v/v) NP40, 0.5% sodium deoxycholate, 0.25 mM of DTT), and incubating for 20 min at 37 °C. After end repair, we washed the beads once in modified RTL buffer and twice in detergent wash buffer (20 mM of Tris pH 7.5, 50 mM of NaCl, 0.2% Triton X-100, 0.2% NP-40, 0.2% sodium deoxycholate) and incubated each washing step for 5 min at 37 °C. After proteinase K digestion, we performed four technical qPCR replicates from the reverse transcription of recovered RNA fragments. 

**Infection of HuH CNBP knockout cells.** A total of 1×10⁵ cells were seeded per well of a 24-well plate. The next day, cells were infected with SARS-CoV-2 at an MOI of 0.5 PFU per cell as described above. At the indicated time points post-infection, supernatants were collected for plaque assay analyses and cells were lysed by reverse transcription (RT-qPCR) analyses. 

**Infection of HEK293 LARP1 knockout cells.** A total of 1.5×10⁵ cells were seeded per well of a poly-L-lysine-coated 24-well plate. For the rescue experiments, cells were treated with the indicated inhibitors as described for the infection assays. After 24 h post-transfection, cells were lysed and analysed by Western blot. Supernatants were collected for the plaque assay analyses and cells were lysed by reverse transcription (RT-qPCR) analyses. 

**RNA extraction and RT–qPCR.** Cells were lysed in 300 μl of TRIzol per well and RNA was extracted using the Direct-zol RNA MicroPrep Kit (Zymo Research). RNA was reverse-transcribed into complementary DNA using the AffinityScript cDNA synthesis kit (Stratagene) according to the manufacturer's instructions. Viral RNA was quantified by qPCR using the PowerUp SYBR Green Master Mix (Thermo Fisher Scientific) and primers specific to the SARS-CoV-2 RdRP gene (forward: GTGARATGGTCAATGTCG, reverse: CARATGTTAAASACACATTATGCATA) and 18S ribosomal RNA (forward: ATGGCGCTTCTTGTTG, reverse: GATAGTAAAGCCTACA). We carried out qPCR reactions in the presence of the ΔΔ⁻¹ method versus 18S. To achieve power to detect small effects in gene expression, we performed four technical qPCR replicates from the same cDNA and took the median value for further analysis. 

**Western blot.** In general, we added NuPAGE LDS Sample Buffer (Thermo Fisher Scientific) to a 1× concentration and incubated samples for 3 min at 95 °C. Proteins were resolved by SDS–polyacrylamide gel electrophoresis using NuPAGE 4 to 12% Bis-Tris-1Gels (Thermo Fisher Scientific) at 200 V for 1 h, followed by transfer to a nitrocellulose membrane using the iBlot dry blotting system (Thermo Fisher Scientific). Western blots were developed using the BiBlot Automated Western System (Thermo Fisher Scientific). For protein detection, we used the following primary antibodies: nucleoplasid protein (catalogue no. ab272882; Abcam); POP1 (catalogue no. 12029-1-AP; Proteintech); LARP1; CNBP; α-Tubulin (catalogue no. 2144; Cell Signaling Technology); β-Actin (catalogue no. sc-7778; Santa Cruz Biotechnology). We used the following secondary antibodies: IRDye 800CW goat anti-rabbit IgG (LI-COR); IRDye 680CW goat anti-mouse IgG (LI-COR). For the visualization of bands, we used the Odyssey Clx Infrared Imaging System (LI-COR). 

**Plaque assay.** TPRM5S2-Vero E6 cells were infected with 10-fold serial dilutions of the virus-containing sample in DEMEM with 1% FCS. After a 1-h incubation, the inoculum was removed and cells were overlaid with 0.6% (w/v) methylcellulose (Carl Roth) in MEM (Gibco) supplemented with 25 mM of HEPES, 0.44% NaHCO₃, 2 mM of GlutaMAX (Gibco), 100 μM of streptomycin, 100 μM of penicillin and 5% FCS. At 4d post-infection, cells were fixed and stained by adding 2x fixation solution (0.23% crystal violet, 8% formaldehyde, 10% ethanol) directly to the medium for 2h. Cells were washed with H₂O and plaques were counted to determine viral titres. 

**Cell viability assay.** For the cell viability assays, cells were seeded in 96-well plates (2×10⁴ cells per well for HuH and A549-ACE2 cells, 6×10⁴ cells per well for Calu3 cells) and treated with inhibitors as described for the infection assays. After 24 h post-infection, cell viability was assessed using the CellTiter Glo reagent (Promega Corporation) according to the manufacturer's instructions. 

**Quantification of ribosomal frameshifting.** HEK293 cells were transiently transfected with either the control or frameshifting construct of our dual-colour enhanced GFP (eGFP)-mCherry translation reporter outlined in Extended Data Fig. 4d. Briefly, cells transfected with this reporter express a single fluorescent protein (eGFP) when the 0 reading frame is translated (Extended Data Fig. 4d). Expression of a second fluorescent protein (mCherry) downstream of eGFP is dependent on -1FS, which prevents translation of an inframe stop codon. Thus, the ratio between mCherry and eGFP directly correlates to -1FS efficiency. As a normalization control, we used a construct lacking a stop codon in the 0 reading frame, leading to the expression of eGFP and mCherry in equal ratios.
providing a direct readout of ribosomal frameshifting efficiency. Accordingly, frameshifting efficiency was calculated using the ratio of mCherry to eGFP observed with the frameshifting reporter construct relative to the mCherry/eGFP ratio observed with the control construct (Extended Data Fig. 4d).

**Computational analyses.** *Protein–protein interaction network.* To establish protein–protein interactions for the proteins identified from the MS experiments, we utilized STRING v11 (ref. 38). For all network and interaction inferences, we used the ‘combined score’ from STRING, which utilizes both physical and functional interactions. Specifically, for the RAP–MS network (Fig. 2), we seeded all proteins detected with an adjusted P < 0.02 and positive log fold change from the moderated t-test between SARS-CoV-2 RNA and RMRP purifications. The edges between interacting proteins were included for those above a combined interaction score of 550. To generate the combined RAP–MS and proteome MS network, we seeded nodes where the adjusted P < 0.05 for either of the assays. Edges between RAP–MS and proteome MS nodes were included for combined interaction scores exceeding 700.

**Gene set and pathway enrichment analysis.** First, we performed a hypergeometric GO enrichment analysis for the expanded MS-Cov-2 RNA interactome proteins using the DAVID tool (v6.8, https://david.ncifcrf.gov/tools.jsp) and applying default settings (Fig. 2b). Additionally, we performed GSEA for the proteome experiments with the clusterProfiler R package (v3.18) utilizing the Hallmark and C5 biological processes gene sets available through Molecular Signatures Database (v7.2) (Fig. 3a). Genes were ranked based on the product of the log fold change and the log_{10} moderated t-test P value between SARS-CoV-2 and mock treatments. To establish enriched terms for communities within the interactome network (Fig. 3c), we considered all regulated genes in the proteome measurements interacting with a specific direct binder and computed enrichments using the C5 biological processes gene sets.

eCLIP and RNA sequencing analysis.** Paired-end sequencing reads from (1) eCLIP experiments or (2) sequencing of crosslinked RNA fragments after RAP–MS, were aligned to hg38 using STAR v2.7.1 (Fig. 3b). Genes were ranked based on the product of the log fold change and the log_{10} moderated t-test P value between SARS-CoV-2 and mock treatments. To establish enriched terms for communities within the interactome network (Fig. 3c), we considered all regulated genes in the proteome measurements interacting with a specific direct binder and computed enrichments using the C5 biological processes gene sets.

**Deposition of the data.** The high-throughput sequencing data have been deposited with the Gene Expression Omnibus under the accession no. GSE154430. Source data are provided with this paper.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability.** The original mass spectra for all experiments and the protein sequence databases used for the searches have been deposited with the MassIVE repository (https://massive.ucsd.edu) and can be accessed at ftp://massive.ucsd.edu/MSV000085734/. The high-throughput sequencing data have been deposited with the Gene Expression Omnibus under the accession no. GSE154430. Source data are provided with this paper.

**Code availability.** The computer code for the custom analyses is publicly available at https://munschauerlab.github.io/SCoV2-proteome-atlas/.

Received: 16 July 2020; Accepted: 3 December 2020; Published online: 21 December 2020

**References**

1. Guan, W.-J. et al. Clinical characteristics of coronavirus disease 2019 in China. *N. Engl. J. Med.* **382**, 1708–1720 (2020).

2. Tay, M. Z., Poh, C. M., Réna, L., MacArty, P. A. & Ng, L. F. P. The trinity of COVID-19: immunity, inflammation and intervention. *Nat. Rev. Immunol.* **20**, 363–374 (2020).

3. Sola, I., Almazán, F., Zúñiga, S. & Enjuanes, L. Continuous and discontinuous RNA synthesis in coronaviruses. *Annu. Rev. Virol.* **2**, 265–288 (2015).

4. Chan, Y. K. & Gack, M. U. Viral evasion of intracellular DNA and RNA sensing. *Nat. Rev. Microbiol.* **14**, 360–373 (2016).

5. Blanco-Melo, D. et al. Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell Rep.* **31**, 1036–1045 (2020).

6. Emanuel, W. et al. Bulk and single-cell gene expression profiling of SARS-CoV-2 infected human cell lines identifies molecular targets for therapeutic intervention. Preprint at bioRxiv https://www.biorxiv.org/content/10.1101/2020.05.05.079194v1 (2020).

7. Bouhaddou, M. et al. The global phosphorylation landscape of SARS-CoV-2 infection. *Cell* **182**, 685–712 (2020).

8. Klann, K. et al. Growth factor receptor signaling inhibition prevents SARS-CoV-2 replication. *Mol. Cell* **80**, 164–174 (2020).

9. Bojkova, D. et al. Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature* **583**, 469–472 (2020).

10. Gordon, D. E. et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).

11. Gordon, D. E. et al. Comparative host–coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science* **370**, eabe9403 (2020).

12. Vanhavre, H., Huulquist, J. F., Kaake, R. M., Hüttenhain, R. & Krogan, N. J. A systems approach to infectious disease. *Nat. Rev. Genet.* **21**, 339–354 (2020).

13. Maranon, D. G., Anderson, J. R., Maranon, A. G. & Wilusz, J. The interface between coronaviruses and host cell RNA biology: novel potential insights for future therapeutic intervention. *Wiley Interdiscip. Rev. RNA* **11**, e16140 (2020).

14. McHugh, C. A. et al. The Xist IncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521**, 232–236 (2015).

15. Munschauer, M. et al. The NORAD IncRNA assembles a topoisomerase complex critical for genome stability. *Nature* **561**, 132–136 (2018).

16. Chu, C. et al. Systematic discovery of X1 RNA binding proteins. *Cell* **161**, 404–416 (2015).

17. Minajigi, A. et al. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* **349**, aab2276 (2015).

18. Ramanathan, M., Porter, D. F. & Khavarri, P. A. Methods to study RNA–protein interactions. *Nat. Methods* **16**, 225–234 (2019).

19. Lee, F. C. Y. & Ule, J. Advances in CLIP technologies for studies of protein–RNA interactions. *Mol. Cell* **69**, 354–369 (2018).

20. Nie, Y. et al. Highly infectious SARS-CoV-ß pseudotyped virus reveals the cell tropism and its correlation with receptor expression. *Biochem. Biophys. Res. Commun.* **531**, 994–1000 (2004).

21. Harcourt, J. et al. Severe acute respiratory syndrome coronavirus 2 from patient with coronavirus disease, United States. *Emerg. Infect. Dis.* **26**, 1266–1273 (2020).

22. Wang, Y. et al. SARS-CoV-2 infection of the liver directly contributes to hepatic impairment in patients with COVID-19. *J. Hepatol.* **73**, 807–816 (2020).

23. Perederina, A. et al. Cryo-EM structure of catalytic ribonucleoprotein complex RNase MRP. *Nat. Commun.* **11**, 3474 (2020).

24. Snijder, E. J., Decroly, E. & Ziebuhr, J. The nonstructural proteins directing coronavirus RNA synthesis and processing. *Adv. Virus Res.* **96**, 59–126 (2016).

25. Kim, Y. et al. Crystal structure of Nsp15 endoribonuclease NendoU from SARS-CoV-2. *Protein Sci.* **29**, 1596–1605 (2020).

26. Hillen, H. S. et al. Structure of replicating SARS-CoV-2 polymerase. *Nature* **584**, 154–156 (2020).

27. Vwlanathan, T. et al. Structural basis of RNA cap modification by SARS-CoV-2. *Nat. Commun.* **11**, 3718 (2020).

28. Sutton, G. et al. The nsp9 replicate protein of SARS-coronavirus, structure and functional insights. *Structure* **12**, 341–353 (2004).

29. Tanaka, T., Kamitani, W., DeDeo, M. L., Enjuanes, L. & Matsuura, Y. Severe acute respiratory syndrome coronavirus nsp5 facilitates efficient propagation in cells through a specific translational shutoff of host mRNA. *J. Virol.* **86**, 11128–11137 (2012).

30. Neuman, B. W. et al. Proteomics analysis unravels the functional repertoire of coronavirus nonstructural protein 3. *J. Virol.* **82**, 5279–5294 (2008).

31. Angelini, M. M., Akhlaghpour, M., Neuman, B. W. & Buchmeier, M. J. Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double membrane vesicles. *mbio* **4**, e00524-13 (2013).

32. Wolff, G. et al. A molecular pore spans the double membrane of the coronavirus replication organelle. *Science* **369**, 1395–1398 (2020).

33. Sharma, K. et al. The 3a accessory protein of SARS coronavirus specifically interacts with the 5′UTR of its genomic RNA, using a unique 75 amino acid interaction domain. *Biochemistry* **46**, 6488–6499 (2007).

34. Masters, P. S. Coronavirus genomic RNA packaging. *Virology* **357**, 198–207 (2019).

35. Shang, J. et al. Cell entry mechanisms of SARS-CoV-2. *Proc. Natl Acad. Sci. USA* **117**, 11727–11734 (2020).
Articles

NATURE MICROBIOLOGY

61. Han, X., Han, Y., Jiao, H. & Jie, Y. 14-3-3
60. Wang, L., Wen, M. & Cao, X. Nuclear hnRNPA2B1 initiates and amplifies
59. Xin, Z. et al. PCBP2 enhances the antiviral activity of IFN-α
54. Suzuki, Y. et al. Characterization of RyDEN (C19orf66) as an
55. Wang, X. et al. Regulation of HIV-1 Gag-Pol expression by Shiftless, an
53. Li, Y. et al. LSm14A is a processing body-associated sensor of viral nucleic
51. Cuevas, R. A. et al. MOV10 provides antiviral activity against RNA viruses
49. Bidet, K., Dadlani, D. & Garcia-Blanco, M. A. G3BP1, G3BP2 and
48. Alam, U. & Kennedy, D. G3BP1 and G3BP2 regulate translation of
47. Arif, A., Chatterjee, P., Moodt, R. A. & Fox, P. L. Heterotrimeric GAIT
46. Takeuchi, A. et al. YB-1 suppression induces STAT3 proteolysis and
45. Narita, R. et al. A novel function of human Pumilio proteins in cytoplasmic
43. Shahbazian, D. et al. The mTOR/PI3K and MAPK pathways converge
38. Ooi, Y. S. et al. An RNA-centric dissection of host complexes controlling
23. Hassan, R. A. & Brown, A. J. 2'-5′A synthesis is required for macropinocytic entry of respiratory syncytial virus (RSV) in human respiratory epithelial cells. PLoS Pathog. 15, e1007963 (2019).
22. Backlund, M. et al. Plasticity of nuclear and cytoplasmic stress responses of RNA-binding proteins. Nucleic Acids Res. 48, 4725–4740 (2020).
21. Zhou, D., Mei, Q., Li, J. & He, H. Cyclophilin A and viral infections. Proc. Natl Acad. Sci. USA 109, 15160–15165 (2012).
20. Kamata, M. et al. Stress protein inhibitor alisporivir. Clin. Infect. Dis. 62, 1349–1360 (2020).
19. Lee, E. et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat. Methods 13, 508–514 (2016).
18. Banerjee, A. K. et al. SARS-CoV-2 disrupts splicing, translation, and protein trafficking to suppress host defenses. Cell 183, 1325–1339 (2020).
17. Wang, L., Wang, N., Han, Y., Chen, J., Li, Y., Li, X., Tang, J., Wang, Y. & He, C. A 3D human brain organotypic model for SARS-CoV-2 infection. Cell 182, 36–41 (2020).
16. Morikawa, T. et al. SARS-CoV-2 infection induces IkB kinase activity and NFκB signaling via caveolin 1. EMBO J. 40, e1001017 (2021).
15. Li, J. et al. Cell-surface proteomic profiling in the fly brain uncovers wiring diagrams. eLife 9, e53357 (2020).
14. Pettersson, J. S. COVID-19 pandemic: time to revive the cyclophilin inhibitor alisporivir. Clin. Infect. Dis. 71, 2191–2194 (2020).
13. Li, J. et al. cyclophilins as target for pan-coronavirus inhibitors. Nature Rev. Drug Discov. 19, 363–374 (2020).
12. Liu, J. et al. Efficient RNA interference in mammalian cells using short interfering RNAs with 20-nucleotide stems. Proc. Natl Acad. Sci. USA 104, 11370–11375 (2007).
11. Moscat, J. et al. Autophagy in health and disease. Cell 165, 272–290 (2016).
10. Li, J. et al. A novel transmission model for SARS-CoV-2 infection in human respiratory epithelial cells. PLoS Pathog. 17, e1008992 (2021).
9. Peng, X. et al. Identification of putative SARS-CoV-2 target proteins of cyclosporin A. Acta Pharmacol. Sin. 41, 234–249 (2020).
8. Peng, X. S. et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat. Methods 13, 508–514 (2016).
7. Lee, E. et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat. Methods 13, 508–514 (2016).
6. Wang, L., Wen, M. & Cao, X. Nuclear hnRNPA2B1 initiates and amplifies
5. Wang, L. et al. Nuclear localization of the RNA-binding protein hnRNPA2B1 is required for macropinocytic entry of respiratory syncytial virus (RSV) in human respiratory epithelial cells. PLoS Pathog. 15, e1007963 (2019).
4. Alam, U. & Kennedy, D. G3BP1 and G3BP2 regulate translation of interferon-stimulated genes: IFITM1, IFITM2 and IFITM3 in the cancer cell line MCF7. Mol. Cell. Biochem. 459, 189–204 (2019).
3. Liu, Y. et al. Lsm14A is a processing body-associated sensor of viral nucleic acids that initiates cellular antiviral response in the early phase of viral infection. Proc. Natl Acad. Sci. USA 115, E10157–E10166 (2018).
2. Kurokawa, K. et al. Interferon-dependent engagement of cytoplasmic initiation factor 4B via S6 kinase (S6K)- and ribosomal protein S6K-mediated signals. Mol. Cell. Biol. 29, 2865–2875 (2009).
1. Cuevas, R. A. et al. MOV10 provides antiviral activity against RNA viruses by enhancing RIG-1-MAVS-independent IFN induction. J. Immunol. 196, 3877–3886 (2016).

NATURE MICROBIOLOGY | VOL 6 | MARCH 2021 | 339–353 | www.nature.com/naturemicrobiology

352
Author contributions
M.M. conceived, designed and supervised the study. J.B. supervised the virus work and performed the infection experiments with assistance from L.K. N.S. performed most of the experiments, analysed the data and provided analytical ideas. C.A.L. performed all computational analyses and developed the analytical concepts and ideas. S.A.C. and H.K. supervised the mass spectrometry work executed by R.M., helped with experimental design and performed the data analyses. S.G. performed the CLIP experiments. C.S. generated the knockout cell lines under the supervision of U.F. and provided analytical ideas. T.H. established the protocols, helped with the infection experiments and generated the cell lines under the supervision of L.D. S.W. performed the qPCR and western blot experiments with help from J.A. and S.Z. Y.W. performed additional computational analyses. M.Z. performed the frameshifting assay and analysis under the supervision of N.C. E.S.L. and J.V. provided support and contributed to study design. M.M. wrote the manuscript with input from all authors.

Competing interests
The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41564-020-00846-z.
Supplementary information is available for this paper at https://doi.org/10.1038/s41564-020-00846-z.

Correspondence and requests for materials should be addressed to J.B. or M.M.

Peer review information Nature Microbiology thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.
© The Author(s) 2020
Extended Data Fig. 1 | Capturing SARS-CoV-2 RNAs and bound proteins with RAP-MS. a, Alignment of protein-crosslinked RNA fragments to the SARS-CoV-2 genome following RNA antisense purification of SARS-CoV-2 RNAs from infected Huh7 cells. Two replicate experiments are shown. b, Fraction of crosslinked RNA fragments mapping to the human or SARS-CoV-2 genomes in pilot RAP-MS experiments. c, Correlation plot for two replicate RAP experiments. CPM values for SARS-CoV-2 genes are shown. CPM: counts per million. d, As in b, but for full-scale SARS-CoV-2 RNA RAP-MS and RMRP RAP-MS experiments. e, Western blot of two replicate SARS-CoV-2 RNA and RMRP RAP-MS experiments. Indicated antibodies were used for protein detection.
Extended Data Fig. 2 | Connectivity in RAP-MS protein-protein association network. Total number of connections observed in protein-protein association network constructed based on expanded SARS-CoV-2 RNA interactome (red bar, 1,534 connections), compared to number of connections observed in random networks of equal size (grey bars, mean 60 connections, z-score 76) using random sampling of proteins detected in proteome measurements.
Extended Data Fig. 3 | See next page for caption.
Extended Data Fig. 3 | Proteome abundance changes in SARS-CoV-2 infected cells. **a,** Correlation of protein abundance measurements reported in Klann et al. and this study ($r = 0.411$). Proteins displayed are significant at an adjusted $P$ value threshold of 0.01 in both studies ($n=712$). **b,** Principle component analysis for proteome measurements of SARS-CoV-2 (SCoV2) infected or mock infected Huh7 cells. **c,** GSEA for proteins significantly regulated in global proteome measurements. Gene sets enriched in addition to those shown in Figure 3b are presented. Statistical test: Kolmogorov-Smirnov test with Benjamini-Hochberg adjustment. **d,** Protein-protein association network of expanded SARS-CoV-2 RNA interactome proteins (blue: interactome protein, not regulated; red: interactome protein, regulated) and their connections to differentially regulated proteins upon SARS-CoV-2 infection. Upregulated proteins are shown in light grey; downregulated proteins are shown in dark grey. Circle sizes scale to the number of connections of each interactome protein.
Extended Data Fig. 4 | Functional validation of SARS-CoV-2 RNA binders. a, Western blot of WT HEK293 cells and four different HEK293 LARP1 knockout (KO) cell lines generated with CRISPR-Cas9 (see Methods). Expression of LARP1 was evaluated relative to Tubulin. b, Western blot of HEK293 cells transiently overexpressing (OE) GFP or LARP1-GFP proteins at 48 h post transfection. Arrows indicate endogenous LARP1 proteins and GFP-tagged LARP1. c, Western blot of four different HEK293 LARP1 knockout cell lines transiently transfected with plasmids encoding GFP or LARP1-GFP proteins at 48 h post transfection. Experiments were repeated at least two times. d, Schematic of dual-fluorescence translation reporter to quantify ribosomal frameshifting efficiency. The depicted control construct contains enhanced GFP (eGFP) and mCherry in an in-frame orientation, leading to the production of both fluorescent proteins separated by a self-cleaving 2A peptide when the 0 reading frame is translated. In the frameshift construct depicted below, eGFP and mCherry are separated by an in-frame stop codon, preventing the production of mCherry when the 0 reading frame is translated. −1FS leads to the production of eGFP and mCherry and the ratio between both fluorescent proteins is a direct measure of frameshifting efficacy. −1FS: -1 ribosomal frameshifting.
Extended Data Fig. 5 | See next page for caption.
Extended Data Fig. 5 | Pharmacological inhibition of SARS-CoV-2 RNA interactome proteins. a, RT-qPCR measurements of intracellular SARS-CoV-2 RNA (RdRP gene) in infected Calu3, Huh7 and A549-ACE2 cells after inhibitor treatment. Inhibitors were used at indicated concentration (left to right). Calu3 cells were assayed 24 h post-infection, Huh7 and A549-ACE2 cells were assayed 48 h post-infection. Values are normalized to 18S rRNA measurements and compared to untreated or DMSO treated cells. b, Infectious viral titers in the supernatants of infected Calu3, Huh7 and A549-ACE2 cells after inhibitor treatment. Inhibitors were used at indicated concentration (left to right). Calu3 cells were assayed 24 h post-infection, Huh7 and A549-ACE2 cells were assayed 48 h post-infection. All values in a–b are mean ± s.d. (n = 3 independent infections) c–d, Cell viability assay in inhibitor-treated and untreated cells. Values are the mean ± s.d. (n = 3 independent treatments). P values determined in unpaired two-tailed t-test. ***P < 0.001; **P < 0.01; *P < 0.05; ns, not significant.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a

☑️ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement

☑️ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly

☑️ The statistical test(s) used AND whether they are one- or two-sided

Only common tests should be described solely by name; describe more complex techniques in the Methods section.

☐ A description of all covariates tested

☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons

☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)

☐ For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted

Give P values as exact values whenever suitable.

☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings

☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes

☐ Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Spectrum Mill MS Proteomics Workbench, Illumina NextSeq 500 system and software

Date analysis

Graphpad Prism 9.0, Microsoft Excel(v.16.16.27), R (v.4.0.2), Python (v2.7; v3.8), MACS2 (v2.2.7), PicardTools (v2.22.0) bwa (v0.7.17), clusterProfiler (v3.18), MEME Suite (v5.2), deepTools (v2.0), Integrative Genomics Viewer (v2.6.0), Database for Annotation Visualization and Integrated Discovery (DAVID v6.8), MSigDB (v7.2) database, STRING (v11) database, The Drug Gene Interaction Database (DGIdb) (v3.0).

Computer code for custom analyses is publicly available at: https://munschauerlab.github.io/SCoV2-proteome-atlas/.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The original mass spectra for all experiments, and the protein sequence databases used for searches have been deposited in the public proteomics repository MassiVE (https://massive.ucsd.edu) and are accessible at ftp://massive.ucsd.edu/MSV000085734/.
High-throughput sequencing data are in Gene Expression Omnibus (GEO) and are available under the accession number GSE154430.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

- Sample size: No sample size calculations were performed a priori. Sample sizes were determined following best practices in the field. We performed at least 3 independent experiments for all assays unless otherwise noted.
- Data exclusions: Mass spectrometry data were filtered for common laboratory contaminants and keratins. Otherwise no data were excluded.
- Replication: As reported in the figure legends, main text and Method section, the findings were reliably reproduced.
- Randomization: There were no variables or interventions to randomize in this study.
- Blinding: Blinding is not relevant to our study, as our tools are not dependent on blinding. Investigators could not be blinded during data collection or analysis. Analyses were performed in an exploratory manner where blinding is not possible.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
| ✗   | Antibodies            |
| ✗   | Eukaryotic cell lines |
| ✗   | Palaeontology and archaeology |
| ✗   | Animals and other organisms |
| ✗   | Human research participants |
| ✗   | Clinical data         |
| ✗   | Dual use research of concern |

### Methods

| n/a | Involved in the study |
|-----|-----------------------|
| ✗   | ChIP-seq              |
| ✗   | Flow cytometry        |
| ✗   | MRI-based neuroimaging |

### Antibodies

- **Antibodies used**
  - CNBP - Proteintech #67109-1-ig (IP, Western Blot)
  - SARS-CoV-2 Nucleoprotein - Abcam #ab272852 (Western blot)
  - POP1 - Proteintech #12029-1-AP (Western blot)
  - LARP1 - Bethyl #A302-087A (IP, Western blot)
  - TUBULIN - Cell Signaling Technologies #12144 (Western blot)
  - ACTIN - Santa Cruz #sc-47778 (Western blot)
  - IRDye 800CW Goat anti-Rabbit IgG - LI-COR #926-3221 (Western blot)
  - IRDye 800CW Goat anti-Mouse IgG - LI-COR #926-32210 (Western blot)

- **Validation**
  - Antibodies were validated by the manufacturer and relevant data is available at the manufacturer’s website.

### Eukaryotic cell lines

- **Cell line source(s)**
  - HuH7 and Calu3 (provided by the Virology Diagnostics Unit at Institute of Virology and Immunobiology, University of Würzburg), ACE2-AS49 (a generous gift from Andreas Pichlmair), Vero-E6-TMPRSS2 cells (a generous gift from Stefan Pöhlmann), HEK293 (provided by the Utz Fischer laboratory).
| Authentication | Cell lines were authenticated by the provider. |
|---------------|-----------------------------------------------|
| Mycoplasma contamination | Cell lines regularly tested negative for mycoplasma contamination. |
| Commonly misidentified lines (See ICLAC register) | No commonly misidentified cell lines were used. |