A newly described coronavirus named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which is the causative agent of coronavirus disease 2019 (COVID-19), has infected over 2.3 million people, led to the death of more than 160,000 individuals and caused widespread social and economic disruption. There are no antiviral drugs with proven clinical efficacy for the treatment of COVID-19, nor are there any vaccines that prevent infection with SARS-CoV-2, and efforts to develop drugs and vaccines are hampered by the limited knowledge of the molecular details of how SARS-CoV-2 infects cells. Here we cloned, tagged and expressed 26 of the 29 SARS-CoV-2 proteins in human cells and identified the human proteins that physically associated with each of the SARS-CoV-2 proteins using affinity-purification mass spectrometry, identifying 332 high-confidence protein–protein interactions between SARS-CoV-2 and human proteins. Among these, we identify 66 druggable human proteins or host factors targeted by 69 compounds (of which, 29 drugs are approved by the US Food and Drug Administration, 12 are in clinical trials and 28 are preclinical compounds). We screened a subset of these in multiple viral assays and found two sets of pharmacological agents that displayed antiviral activity: inhibitors of mRNA translation and predicted regulators of the sigma-1 and sigma-2 receptors. Further studies of these host-factor-targeting agents, including their combination with drugs that directly target viral enzymes, could lead to a therapeutic regimen to treat COVID-19.
Analyses of SARS-CoV-2–host protein interactions

Our AP-MS analysis identified 332 high-confidence protein interactions between SARS-CoV-2 proteins and human proteins, observing correlations between replicate experiments of each viral bait (Pearson’s R = 0.46–0.72) (Extended Data Fig. 2 and Supplementary Tables 1, 2). We studied the interacting human proteins with regards to their biological functions, anatomical expression patterns, expression changes during SARS-CoV-2 infection5,8 and in relation to other maps of host–pathogen interacting proteins43,52 (Fig. 2a). We analysed each viral protein for Gene Ontology enrichment (Fig. 2b and Extended Data Fig. 3) and identified the major cell processes of the interacting proteins, including lipoprotein metabolism (S), nuclear transport (NSP7) and ribonucleoprotein complex biogenesis (NSP8). To discover potential binding interfaces, we enriched for domain families within the interacting proteins of each viral bait (Extended Data Fig. 4). For instance, DNA polymerase domains are enriched among proteins that interact with NSP1, and bromodomains and extra-terminal domain (BET) family domains are enriched among proteins that interact with E (Supplementary Discussion and Supplementary Methods).

Although the cell line used for these AP-MS experiments, HEK-293T/17, can be infected with the SARS-CoV-2 virus11, it does not represent the primary physiological site of infection—lung tissue. From 29 human tissues12, we identified the lung as the tissue with the highest expression of the prey proteins relative to the average proteome (Fig. 2c). Consistent with this, the interacting proteins were enriched in the lung relative to other tissues (Extended Data Fig. 5a), and compared to overall RefSeq gene expression in the lung (median transcripts per million (TPM) = 3.198), proteins that interacted with SARS-CoV-2 proteins were expressed at a higher level (median TPM = 25.52, P = 0.0007; Student’s t-test) (Extended Data Fig. 5b), supporting the hypothesis that SARS-CoV-2 preferentially hijacks proteins that are expressed in lung tissue.

We also studied the evolutionary properties of the host proteins bound by SARS-CoV-2 (Supplementary Table 3, Supplementary Methods and Supplementary Discussion). In addition, we analysed changes in protein abundance during SARS-CoV-2 infection21. When possible, the correlation between changes in the abundance of viral proteins and their human interaction partners across four time points. Interacting pairs typically had stronger correlated changes than other pairs of viral–human proteins (Fig. 2d) (Kolmogorov–Smirnov test P = 4.8 × 10−5), indicating that the AP-MS-derived interactions are relevant for the target tissue and the infection context. We compared our SARS-CoV-2 interaction map with those of ten other pathogens (Fig. 2e) and found that West Nile virus34 and Mycobacterium tuberculosis35 had the most similar host-protein interaction partners. The association with M. tuberculosis is of particular interest as it also infects lung tissue.

The interactome reveals SARS-CoV-2 biology

Our study highlights interactions between SARS-CoV-2 proteins and human proteins that are involved in several complex and biological processes (Fig. 3). These included DNA replication (NSP1), epigenetic and gene-expression regulators (NSP5, NSP8, NSP13 and E), vesicle trafficking (NSP2, NSP6, NSP7, NSP10, NSP13, NSP15, ORF3a, E, M and ORF8), lipid modification (S), RNA processing and regulation (NSP8 and N), ubiquitin ligases (ORF10), signalling (NSP7, NSP8, NSP13, N and ORF9b), nuclear transport machinery (NSP9, NSP15 and ORF6), and identified the major cell processes of the interacting proteins, (when homologues exist). n = 4 structural proteins; n = 16 NSPs; n = 9 accessory factors.
level changes during SARS-CoV-2 infection for pairs of viral–human proteins through multiple hypothesis testing. A false-discovery rate was used to account for the overlap of human interacting proteins between SARS-CoV-2 and other pathogens using a hypergeometric test (unadjusted for multiple testing). The background gene set for the test consisted of all unique proteins detected by mass spectrometry across all pathogens (n = 10,181 proteins). HCV, hepatitis C virus; HIV, human immunodeficiency virus; HPV, human papillomavirus; KSHV, Kaposi’s sarcoma-associated herpesvirus; Mtb, M. tuberculosis; WNV, West Nile virus.

Fig. 2 | Global analysis of SARS-CoV-2 protein interactions. a, Overview of global analyses performed. b, Gene Ontology (GO) enrichment analysis was performed on the human interacting proteins of each viral protein. P values were calculated by hypergeometric test and a false-discovery rate was used to account for multiple hypothesis testing (Methods). The top GO term of each viral protein was selected for visualization. c, Degree of differential protein expression for the human interacting proteins (n = 332) across human tissues. We obtained protein abundance values for the proteome in 29 human tissues and calculated the median level of abundance for the human interacting proteins (top 16 tissues shown). This was then compared with the abundance values for the full proteome in each tissue and summarized as a Z score from which a P value was calculated. A false-discovery rate was used to account for multiple hypothesis testing. d, The distribution of the correlation of protein level changes during SARS-CoV-2 infection for pairs of virus–human proteins (median, white circles; interquartile range, black bars) is higher than non-interacting pairs of viral–human proteins (P = 4.8 × 10−5; Kolmogorov–Smirnov test). The violin plots show each viral–human protein correlation for preys (n = 210, minimum = −0.986, maximum = 0.999, quartile (Q)1 = −0.468, Q2 = 0.396, Q3 = 0.850) and non-preys (n = 54765, minimum = −0.999, maximum = 0.999, Q1 = −0.599, Q2 = 0.006, Q3 = 0.700). e, Significance of the overlap of human interacting proteins between SARS-CoV-2 and other pathogens using a hypergeometric test (unadjusted for multiple testing). The background gene set for the test consisted of all unique proteins detected by mass spectrometry across all pathogens (n = 10,181 proteins). HCV, hepatitis C virus; HIV, human immunodeficiency virus; HPV, human papillomavirus; KSHV, Kaposi’s sarcoma-associated herpesvirus; Mtb, M. tuberculosis; WNV, West Nile virus.

SARS-CoV-2 interacts with innate immune pathways

Several innate immune signalling proteins are targeted by SARS-CoV-2 viral proteins. The interferon pathway is targeted by NSP13 (TBK1 and TBKBP1), NSP15 (RNFI41 (also known as NRD1P)) and ORF9b (TOPM70); and the NF-κB pathway is targeted by NSP13 (TLE1, TLE3 and TLE5) and ORF9c (NLX1, F2RL1 and NDFIP2). Furthermore, two other E3 ubiquitin ligases that regulate antiviral innate immune signalling, TRIM59 and MIB1, are bound by ORF3a and NSP9, respectively.

We also identified interactions between SARS-CoV-2 ORF6 and the NUP98–RAGE complex (Fig. 4a), an interferon-inducible mRNA sequence of HDAC2 (Extended Data Fig. 6a–d), suggesting that NSP5 may inhibit the transport of HDAC2 into the nucleus and could potentially affect the ability of HDAC2 to mediate the inflammation and interferon response. MIB1 interacts with Sigma receptors that have been implicated in lipid remodelling and the stress response of the endoplasmic reticulum; these proteins interact with many human drugs (see ‘Antiviral activity of host-directed compounds’).
Fig. 3 | SARS-CoV-2 protein–protein interaction network. 332 high-confidence interactions between 26 SARS-CoV-2 proteins (red diamonds) and human proteins (circles; drug targets: orange; protein complexes: yellow; proteins in the same biological process: blue). Edge colour proportional to MiST score; edge thickness proportional to spectral counts. Physical interactions among host proteins (thin black lines) were curated from CORUM, IntAct, and Reactome. An interactive protein–protein interaction map can be found at kroganlab.ucsf.edu/network-maps. ECM, extracellular matrix; ER, endoplasmic reticulum; snRNP, small nuclear ribonucleoprotein. n = 3 biologically independent samples.
nuclear export complex that is hijacked or degraded by multiple viruses including vesicular stomatitis virus (VSV), influenza A, Kaposi’s sarcoma-associated herpesvirus and poliovirus, and is a restriction factor for influenza A infection. The X-ray structure of the VSV M protein complexed with NUP98–RAE1 reveals key binding interactions, including a buried methionine residue on the M protein that packs into a hydrophobic pocket in RAE1, and neighbouring acidic residues that interact with a basic patch on the NUP98–RAE1 complex. These features are also present in a conserved motif in the C-terminal region of SARS-CoV-2 ORF6 (Fig. 4b–d and Extended Data Fig. 7a, b), providing a structural hypothesis for the observed interaction. ORF6 of SARS-CoV antagonizes host interferon signalling by perturbing nuclear transport, and the NUP98–RAE1 interaction with ORF6 may perform the same function for SARS-CoV-2.

**SARS-CoV-2 interacts with host translation machinery**

Nucleocapsid (N) of SARS-CoV-2 binds to the stress granule proteins G3BP1 and G3BP2, and to other host mRNA-binding proteins including the mTOR-regulated translational repressor LARP1, two subunits of casein kinase 2 (CK2), and mRNA decay factors UPF1 and MOV10 (Fig. 4e). Manipulation of the stress granule and related RNA biology is common among Coronavirusae and stress granule formation is thought to be a primarily antiviral response. The promotion of G3BP aggregation by eIF4A inhibitors may partially explain their antiviral activity (see ‘Antiviral activity of host-directed compounds’).

All coronavirus mRNAs rely on cap-dependent translation to produce their proteins, a process enhanced in trans by the SARS-CoV N protein. Key eIF4F–cap binding complex constituents—the cap binding protein eIF4E, scaffold protein eIF4G and the DEAD-box helicase eIF4A—are candidates for therapeutic targeting of coronaviruses. Therapeutic targeting (Fig. 4f, g) of viral translation by interfering with the eIF4F complex formation or the interactions between viral proteins N, NSP2 or NSP8 and the translational machinery may have therapeutic benefits (see ‘Antiviral activity of host-directed compounds’).

**Fig. 4** The SARS-CoV-2 interactome reveals novel aspects of SARS-CoV-2 biology and pharmacological targets. a, ORF6 interacts with an mRNA nuclear export complex that can be targeted by selinexor. b, The C-terminal peptide of SARS-CoV-2 ORF6 (dark purple) is modelled into the binding site of the VSV M protein (yellow)–NUP98 (green)–RAE1 (light purple) complex (Protein Data Bank (PDB) ID: 4OWR). ORF6 and M protein residues are labelled. c, The C-terminal sequence of SARS-CoV-2 ORF6, highlighting the described trafficking motifs and putative NUP98–RAE1 binding sequence. The chemical properties of amino acids are shown as follows: polar, green; neutral, purple; basic, blue; acidic, red; and hydrophobic, black. d, Putative NUP98–RAE1 interaction motifs. Negatively charged residues (red) surround a conserved methionine (yellow) in several virus species. e, N targets stress granule proteins. f, Inhibition of casein kinase II (by silmitasertib or TMCB) disrupts phosphorylation of eIF4E. fER1Cat blocks the interaction of eIF4E with eIF4G. Inhibition of eIF4A (zotatifin) may prevent the unwinding of the viral 5’ untranslated region to prevent its translation. h, Targeting the translation elongation factor-1A ternary complex (ternatin-4) or the Sec61 translocon (PS3061) may prevent viral protein production and membrane insertion, respectively. i, ORF10 interacts with the CUL2 complex for the ubiquitination (Ub) of host proteins, which can be inhibited by pevonedistat. j, The protein E interacts with bromodomain proteins. j, Alignment of E proteins from SARS-CoV-2, SARS-CoV and bat-CoV with histone H3 and NS1 protein of influenza A H3N2. Identical and similar amino acids are highlighted. k, Bromodomain inhibitors (iBETs) may disrupt the interaction between E and BRDs. l, FDA-approved drugs are shown in green, clinical candidates are shown in orange and preclinical candidates are shown in purple.

nuclear export complex that is hijacked or degraded by multiple viruses including vesicular stomatitis virus (VSV), influenza A, Kaposi’s sarcoma-associated herpesvirus and poliovirus, and is a restriction factor for influenza A infection.

Key eIF4F–cap binding complex constituents—the cap binding protein eIF4E, scaffold protein eIF4G and the DEAD-box helicase eIF4A—are candidates for therapeutic targeting of coronaviruses. Therapeutic targeting (Fig. 4f, g) of viral translation by interfering with the eIF4F complex formation or the interactions between viral proteins N, NSP2 or NSP8 and the translational machinery may have therapeutic benefits (see ‘Antiviral activity of host-directed compounds’).
NSP8 and three SRP components suggest that the virus hijacks the Sec61-mediated protein translocation pathway for entry into the endoplasmic reticulum. Sec61 inhibitors of protein biogenesis such as PS3061 (Fig. 4h), which has previously been shown to inhibit other enveloped RNA viruses, may also block SARS-CoV-2 replication and assembly.

**Fig. 5 | Drug–human target network**. Protein–protein interactions of SARS-CoV-2 baits with approved drugs (green), clinical candidates (orange) and preclinical candidates (purple) with experimental activities against the host proteins (white background) or previously known host factors (grey background) are shown.
SARS-CoV-2 interacts with a Cullin ubiquitin ligase

Viruses commonly hijack ubiquitination pathways for replication and pathogenesis\(^4\). The ORF10 of SARS-CoV-2 interacts with members of a cullin-2 (CUL2) RING E3 ligase complex (Fig. 4i), specifically the CUL2\(^{ZYG11B}\) complex. ZYG11B is the highest scoring protein in the ORF10 interactome, suggesting that there is a direct interaction between ORF10 and ZYG11B. Despite its small size (38 amino acids), ORF10 appears to contain an α-helical region (Fig. 4j) that may be adopted in complex with CUL2\(^{ZYG11B}\). The ubiquitin transfer to a substrate requires neddylation of CUL2 by NEDD8-activating enzyme (NAE), which is a druggable target\(^5\) (Fig. 4k). ORF10 may bind to the CUL2\(^{ZYG11B}\) complex.
and hijack it for ubiquitination and degradation of restriction factors, or alternatively, ZYG18 may bind to the N-terminal glycine in ORF10 to target it for degradation.

**SARS-CoV-2 interacts with bromodomain proteins**

We found that the transmembrane E protein, which is probably resident on the endoplasmic reticulum–Golgi intermediate compartment and Golgi membranes, binds to BRD2 and BRD4 (Fig. 4I), members of the bromodomain and extra-terminal (BET) domain family of epigenetic readers that bind to acetylated histones to regulate gene transcription. The C-terminal region of E mimics the N-terminal segment of histone H3, which is a known interacting partner of bromodomains. Notably, this region of E is highly conserved in SARS and bat coronaviruses, which suggests that it has a conserved function (Fig. 4M). A similar short peptide motif has also been identified in the N51S1 protein of the influenza A H3N2 strain, in which it interferes with transcriptional processes that support an antiviral response. Bromodomain inhibitors might disrupt the interaction between protein E and BRDs (Fig. 4N).

For a more comprehensive overview of virus–host interactions, see Supplementary Discussion and Supplementary Methods.

**Identification of drugs that target host factors**

To disrupt the SARS-CoV-2 interactome, we sought ligands of human proteins that interact with viral proteins (Methods). Molecules were prioritized by the MIST score of the interaction between the human and viral proteins; by their status as approved drugs, investigational drugs (drugs in clinical trials) or as preclinical candidates; by their selectivity; and by their availability (Supplementary Tables 4, 5). Chemoinformatic searches from the IUPHAR/BPS Guide to Pharmacology (2020-3-12) and the ChEMBL database on the human interactors yielded 16 approved drugs, 3 investigational drugs and 18 preclinical candidates (Supplementary Table 4); and target- and pathway-specific literature search revealed 13 approved drugs, 9 investigational drugs and 10 preclinical candidates (Supplementary Table 5). Of the 332 human targets that interact with the viral bait proteins with a high-confidence score (Fig. 3), 62 have 69 drugs, investigational drugs or preclinical molecules that modulate them and can be overlaid on our protein-interaction network (Fig. 5).

**Antiviral activity of host-directed compounds**

We next investigated the antiviral activity of these drugs and compounds, using two viral assays (Fig. 6A). First, at Mount Sinai Hospital in New York, we developed a medium-throughput immunofluorescence-based assay (which detects the viral NP protein) to screen 37 compounds for inhibition of SARS-CoV-2 infection in the Vero E6 cell line. Second, at the Institut Pasteur in Paris, viral RNA was monitored using quantitative PCR with reverse transcription (RT–qPCR) after treatment with 44 drugs and compounds. Together, both locations tested 47 of the 69 compounds that we identified, plus 13 to expand testing of the sigma-1 and sigma-2 receptors and mRNA translation targets, and 15 additional molecules that had been prioritized by other methods (Methods and Supplementary Table 6). Viral growth and cytotoxicity were monitored at both institutions (Extended Data Figs. 8, 9 and Supplementary Table 6). Two classes of molecules emerged as effectively reducing viral infectivity: protein biogenesis inhibitors (zotatifin, ternat-4 and PS3061) (Fig. 6B and Extended Data Fig. 9) and ligands of the sigma-1 and sigma-2 receptors (haloperidol, PB28, PD-144418 and hydroxychloroquine, which is undergoing clinical trials in patients with COVID-19 (ClinicalTrials.gov, trial number NCT04332991)). We also subsequently found that the sigma-1- and sigma-2-receptor active drugs clemastine, cloperastine and progesterone, and the clinical molecule siramesine, were antiviral drugs (Fig. 6C and Extended Data Fig. 9). Median tissue culture infectious dose (TCID₅₀) assays on supernatants from infected cells treated with PB28 (90% inhibitory concentration (IC₉₀) = 0.278 μM) and zotatifin (IC₉₀ = 0.037 μM) revealed a more potent inhibition than was observed in the NP-staining assay (Fig. 6D). Notably, in this assay, PB28 was around 20 times more potent than hydroxychloroquine (IC₉₀ = 5.78 μM).

To better understand the mechanism by which these inhibitors exert their antiviral effects, we performed a time course assay in which the drugs were added at different times before or after infection (Fig. 6E). Cells were infected during a single cycle of infection at high multiplicity of infection (MOI = 2) over the course of 8 h, and the drugs were added either 2 h before infection or 0, 2 or 4 h after infection. PB28, zotatifin and hydroxychloroquine all decreased the detection of the viral NP protein even in this single cycle assay, indicating that the antiviral effect occurs before viral egress from the cell (Fig. 6E). Furthermore, all three molecules inhibited NP expression when added up to 4 h after infection, after viral entry has occurred. Thus, these molecules seem to exert their antiviral effect during viral replication.

Coronaviruses rely on cap-dependent mRNA translation through the translation machinery of the host. eIF4H, which interacts with NSP9, is a partner of eIF4A, and we observed a strong antiviral effect after treatment with the eIF4A inhibitor zotatifin (Fig. 6B), which is currently in a phase-I clinical trial for the treatment of cancer. We also observed potent antiviral effects of the elongation factor-1A (eEF1A) inhibitor ternat-4 (Fig. 6B), which may suggest that the rate of translation elongation is critical for obtaining optimal levels of viral proteins. Of note, the eEF1A inhibitor aplidin/plitidepsin is used clinically in patients with multiple myeloma. Multiple SARS-CoV-2 proteins are predicted to undergo SRP- and Sec61-mediated co-translational insertion into the endoplasmic reticulum, and SRP19, SRP54 and SRP72 were identified as NSP8-interacting proteins (Fig. 3). Consistent with previous studies of flaviviruses, the Sec61 inhibitor PS3061 also blocked SARS-CoV-2 replication (Extended Data Fig. 9). The two translation inhibitors had cytostatic effects in uninfected Vero cells, which are immortalized cell lines with indefinite proliferative capacity that have mutations in key cell cycle inhibitors. These cells are more sensitive to antiviral compounds, which affect the cell cycle state of immortalized cells more strongly than non-immortalized cells. A critical question going forward is whether these or related inhibitors of viral protein biogenesis will show therapeutic benefits in patients with COVID-19. Plitidepsin is currently under consideration by the Spanish Medicines Agency for a phase-II trial in hospitalized patients with COVID-19.

Molecules that target the sigma-1 and sigma-2 receptors perturb the virus through different mechanisms than the translation inhibitors, which could include the cell stress response. These molecules are also active against other aminergic receptors; however, the only targets shared among all of the tested molecules are the sigma-1 and sigma-2 receptors (Fig. 6F), into which these drugs can be readily modelled (Extended Data Fig. 10A). For instance, the antipsychotic drug haloperidol inhibits the dopamine D2 and histamine H1 receptors, whereas clemastine and cloperastine are antihistamines; each of these drugs is a Sigma receptor ligand with antiviral activity (Fig. 6C). Conversely, the antipsychotic drug olanzapine, which also inhibits histamine H1 and dopamine D2 receptors, has little Sigma receptor activity and does not show antiviral activity (Extended Data Fig. 10B). Which of the Sigma receptors is responsible for the activity remains uncertain, as does the role of pharmacologically related targets, such as EBP and related sterol isomerases, the ligand recognition of which resembles those of the Sigma receptors. Notably, the sigma-1-receptor benzomorphan agonist dextromethorphan has proviral activity (Fig. 6G), further supporting the role of these receptors in viral infection. Overall, two features should be emphasized. First, several of the molecules that target the Sigma receptors, such as clemastine, cloperastine and progesterone, are approved drugs with a long history in human therapy. Many other widely used drugs, which show activity against the Sigma receptors,
remain to be tested; and indeed, several drugs such as astemizole, which we show is a sigma-2 receptor ligand (with an \( K_s \) of 95 nM) (Extended Data Fig. 11), verapamil and amiodarone, have been reported by others to be active in viral replication assays, although this has not been linked to their Sigma receptor activity\(^{12,13}\). Second, the Sigma receptor ligands have a clear separation between antiviral and cytotoxic effects (Fig. 6b, c), and ligands such as PB28 have substantial selectivity for the Sigma receptors compared with side-effect targets, such as the hERG ion channel. Indeed, the lack of selectivity of chloroquine and hydroxychloroquine for hERG (Fig. 6b) and other off-targets (Extended Data Fig. 12) may be related to the adverse cardiac drug reactions\(^{19}\) that have limited their use.

Discussion

In this study, we have identified 332 high-confidence SARS-CoV-2 protein–human protein interactions that are connected with multiple biological processes, including protein trafficking, translation, transcription and regulation of ubiquitination. We found 69 ligands, including FDA-approved drugs, compounds in clinical trials and preclinical compounds, that target these interactions. Antiviral tests in two different laboratories reveal two broad sets of active drugs and compounds; those that affect translation and those modulate the sigma-1 and sigma-2 receptors. Within these sets are at least five targets and more than ten different chemotypes, providing a rich landscape for optimization.

The chemo-proteomic analysis that emerges from this study not only highlights clinically actionable drugs that target human proteins in the interactome, but also provides a context for interpreting their mechanism of action. The potent efficacy of the translation inhibitors on viral infectivity—in the 10 to 100 nM range—makes these molecules attractive as candidate antiviral agents, and also highlights this pathway as a point of intervention. Although the mechanism of action of the drugs that target the sigma-1 and sigma-2 receptors remains less defined, their activity as both anti- and pro-inflammatory agents is mechanistically suggestive. The relatively strong efficacy of PB28, at an IC\(_{90}\) of 280 nM in the viral titre assay, and its high selectivity against off-target proteins, suggests that molecules of this class may be optimized as therapeutic agents. Although it is unclear whether approved drugs such as clemastine and cloperastine, which are used as antihistamines and antitussive drugs, have pharmacokinetics that are suitable for antiviral therapy, and although they are not free of binding to targets that cause side effects (Fig. 6f and Extended Data Fig. 12), these drugs have been used for decades. We caution against their use outside of controlled studies, because of their side-effect liabilities. By the same standard, we find that the widely used antitussive drug dextromethorphan has proviral activity and that therefore its use should merit caution and further study in the context of the treatment of COVID-19. More positively, there are dozens of approved drugs that show activity against the Sigma receptors, not all of which are generally recognized as Sigma receptor ligands. Many of these drugs remain to be tested as a treatment for COVID-19; although some have begun to appear in other studies\(^{21,22}\). This area of pharmacology has great promise for the repurposing and optimization of new agents in the fight against COVID-19.

Our approach of host-directed intervention as an antiviral strategy overcomes problems associated with drug resistance and may also provide panviral therapies as we prepare for the next pandemic. Furthermore, the possibilities for cotherapies are expanded—for example, with drugs that directly target the virus, including remdesivir—and, as we demonstrate in this study, there are numerous opportunities for the repurposing of FDA-approved drugs. More broadly, the pipeline described here represents an approach for drug discovery not only for panviral strategies, but also for the research of many diseases, and illustrates the speed with which science can be moved forward using a multi-disciplinary and collaborative approach.
Methods

Genome annotation
The GenBank sequence for SARS-CoV-2 isolate 2019-nCoV/USA-WA1/2020, accession MN985325, was downloaded on 24 January 2020. In total, we annotated 29 possible ORFs and proteolytically mature proteins encoded by SARS-CoV-2.2 Proteolytic products that result from NPS3 and NPS5-mediated cleavage of the ORF1a/ORF1ab polyprotein were predicted on the basis of the protease specificity of SARS-CoV proteases, and 16 predicted nonstructural proteins (NSPs) were subsequently cloned (NSP1–NSP16). For the NSPS protease (3Clike/3CLpro), we also designed the catalytically dead mutant NPS5 (C145A)23,28. ORFs at the 3′ end of the viral genome annotated in the original GenBank file included 4 structural proteins: S, E, M, and N and the additional ORFs ORF3a, ORF6, ORF7a, ORFs and ORF10. On the basis of the analysis of ORFs in the genome and comparisons with other annotated SARS-CoV ORFs, we annotated a further four ORFs: ORF3b, ORF7b, ORF9b, and ORF9c.

Cloning
ORFs and proteolytically mature NSPs annotated in the SARS-CoV-2 genome were human codon-optimized using the IDT codon-optimization tool (https://www.idtdna.com/codonopt) and internal EcoRI and BamHI sites were eliminated. Start and stop codons were added as necessary to NSPs 1–16, a Kozak sequence was added before each start codon, and a 2×Strep tag with linker was added to either the N or C terminus. To guide our tagging strategy, we used GPS-Lipid each start codon, and a 2×Strep tag with linker was added to either the N or C terminus. T otal plasmid was normalized to 15 μg with the N or C terminus. T o guide our tagging strategy, we used GPS-Lipid each start codon, and a 2×Strep tag with linker was added to either

Transfection
For each affinity purification (26 wild-type baits and one catalytically dead SARS-CoV-2 bait, one GFP control and one empty vector control), ten million HEK-293T/17 cells were plated per 15-cm dish and transfected with up to 15 μg of individual Strep-tagged expression constructs after 20–24 h. Total plasmid was normalized to 15 μg with empty vector and complexed with PolyJet Transfection Reagent (SignaGen Laboratories) at a 1:3 μg:μl ratio of plasmid:transfection reagent based on the manufacturer’s recommendations. After more than 38 h, cells were dissociated at room temperature using 10 ml Dulbecco’s phosphate-buffered saline without calcium and magnesium (DPBS) supplemented with 10 mM EDTA for at least 5 min and subsequently washed with 10 ml DPBS. Each step was followed by centrifugation at 200g, 4 °C for 5 min. Cell pellets were frozen on dry ice and stored at −80 °C. For each bait, n = 3 independent biological replicates were prepared for affinity purification.

Affinity purification
Frozen cell pellets were thawed on ice for 15–20 min and resuspended in 1 ml lysis buffer (IP buffer (50 mM Tris-HCl, pH 7.4 at 4 °C, 150 mM NaCl, 1 mM EDTA) supplemented with 0.5% Non-Idet P40 substitute (NP40; Fluka Analytical) and Complete mini EDTA-free protease and PhosSTOP phosphatase inhibitor cocktails (Roche)). Samples were then frozen on dry ice for 10–20 min and partially thawed at 37 °C before incubation on a tube rotator for 30 min at 4 °C and centrifugation at 13,000g, 4 °C for 15 min to pellet debris. After reserving 50 μl lysate, up to 48 samples were arrayed into a 96-well Deepwell plate for affinity purification on the KingFisher Flex Purification System (Thermo Scientific) as follows: MagStrep ‘type3’ beads (30 μl; IBA Lifesciences) were equilibrated twice with 1 ml wash buffer (IP buffer supplemented with 0.05% NP40) and incubated with 0.95 ml lysate for 2 h. Beads were washed three times with 1 ml wash buffer and then once with 1 ml IP buffer. To directly digest bead-bound proteins as well as elute proteins with biotin, beads were manually suspended in IP buffer and divided in half before transferring to 50 μl denaturation–reduction buffer (2 M urea, 50 mM Tris-HCl pH 8.0, 1 mM DTT) and 50 μl×1 buffer BXT (IBA Lifesciences) dispensed into a single 96-well KF microtitre plate. Purified proteins were first eluted at room temperature for 30 min with constant shaking at 1,100 rpm on a ThermoMixer C incubator. After removing eluates, on-bead digestion proceeded (see ‘On-bead digestion’). Strep-tagged protein expression in lysates and enrichment in eluates were assessed by western blot and silver stain, respectively.

On-bead digestion
Bead-bound proteins were denatured and reduced at 37 °C for 30 min and after being brought to room temperature, alkylated in the dark with 3 μM iodoacetamide for 45 min and quenched with 3 μM DTT for 10 min. Proteins were then incubated at 37 °C, initially for 4 h with 1.5 μl trypsin (0.5 μg/μl; Promega) and then another 1–2 h with 0.5 μl additional trypsin. To offset evaporation, 15 μl 50 mM Tris-HCl, pH 8.0 were added before trypsin digestion. All steps were performed with constant shaking at 1,100 rpm on a ThermoMixer C incubator. Resulting peptides were combined with 50 μl 50 mM Tris-HCl, pH 8.0 used to rinse beads and acidified with trifluoroacetic acid (0.5% final, pH < 2.0). Acidified peptides were desalted for MS analysis using a BioPureSPE Mini 96-Well Plate (20 mg PROTO 300 C18; The Nest Group) according to standard protocols.

MS data acquisition and analysis
Samples were re-resuspended in 4% formic acid, 2% acetonitrile solution, and separated by a reversed-phase gradient over a Nanoflow C18 column (Dr Maisch). Each sample was directly injected via an Easy-nLC 1200 (Thermo Fisher Scientific) into a Q-Exactive Plus mass spectrometer (Thermo Fisher Scientific) and analysed with a 75 min acquisition, with all MS1 and MS2 spectra collected in the orbitrap; data were
acquired using the Thermo software Xcalibur (4.2.47) and Tune (2.11 QFI Build 3006). For all acquisitions, QCloud was used to control instrumental longitudinal performance during the project. All proteomic data were searched against the human proteome (Uniprot-reviewed sequences downloaded 28 February 2020), the eGFp sequence and the SARS-CoV-2 protein sequences using the default settings for MaxQuant (v.1.6.11.0)\(^{44,45}\). Detected peptides and proteins were filtered to 1% false-discovery rate in MaxQuant, and identified proteins were subsequently subjected to protein–protein interaction scoring with both SAINtExpress (v.3.6.3)\(^{44}\) and MiST (https://github.com/kroganlab/mist)\(^{46,47}\). We applied a two-step filtering strategy to determine the final list of reported interactors, which relied on two different scoring stringency cut-offs. In the first step, we chose all protein interactions that had a MIST score \(\geq 0.7\), a SAINtExpress Bayesian false-discovery rate (BFDR) \(\leq 0.05\) and an average spectral count \(\geq 2\). For all proteins that fulfilled these criteria, we extracted information about the stable protein complexes that they participated in from the CORUM database of known protein complexes. In the second step, we then relaxed the stringency and recovered additional interactors that (1) formed complexes with interactors determined in filtering step 1 and (2) fulfilled the following criteria: MIST score \(\geq 0.6\), SAINtExpress BFDR \(\leq 0.05\) and average spectral counts \(\geq 2\). Proteins that fulfilled filtering criteria in either step 1 or step 2 were considered to be high-confidence protein–protein interactions (HC-PPIs) and visualized with Cytoscape (v.3.7.1)\(^{48}\). Using this filtering criteria, nearly all of our bait datasets recovered a number of HC-PPIs in close alignment with previous datasets reporting an average of around 6 PPIs per bait\(^{49}\). However, for a subset of baits (ORF8, NSP8, NSP13 and ORF9c), we observed a much higher number of PPIs that passed these filtering criteria. For these four baits, the MIST scoring was instead performed using a larger in-house database of 87 baits that were prepared and processed in an analogous manner to this SARS-CoV-2 dataset. This was done to provide a more comprehensive collection of baits for comparison, to minimize the classification of non-specifically binding background proteins as HC-PPIs. All MS raw data and search results files have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD018117\(^{70,71}\). PPI networks have also been uploaded to NDEX.

**GO overrepresentation analysis**

The targets of each bait were tested for enrichment of GO biological process terms. The overrepresentation analysis was based on the hypergeometric distribution and performed using the enricher function of clusterProfiler package in R with default parameters. The GO terms were obtained from the C5 category of Molecular Signature Database (MSigDB v6.1). Significant GO terms (1% false-discovery rate) were identified and further refined to select non-redundant terms. To select non-redundant gene sets, we first constructed a GO term tree based on distances (1 – Jaccard similarity coefficients of shared genes) between the significant terms. The GO term tree was cut at a specific level (\(h = 0.99\)) to identify clusters of non-redundant gene sets. For a bait with multiple significant terms belonging to the same cluster, we selected the broadest term that is, largest gene set size.

**Virus interactome similarity analysis**

Interactome similarity was assessed by comparing the number of shared human-interacting proteins between pathogen pairs, using a hypergeometric test to calculate significance. The background gene set for the test consisted of all unique proteins detected by MS across all pathogens \(n = 10,181\) genes).

**ORF6 peptide modelling**

The proposed interaction between ORF6 and the NUP98–RAE1 complex was modelled in PyRosetta 4 (release v.2020.02-dev61090)\(^{72}\) using the crystal structure of VSV matrix (M) protein bound to NUP98–RAE1 as a template\(^73\) (PDB 4OWR; downloaded from the PDB-REDO server\(^{74}\)). The M protein chain (C) was truncated after residue 54 to restrict the model to the putative interaction motif in ORF6 (M protein residues 49–54, sequence DEMTD). These residues were mutated to the ORF6 sequence, QPMEID, using the mutate_residue function in the module pyrosetta.toolbox, without repacking at this initial step. After all six residues were mutated, the full model was relaxed to a low-energy conformation using the FastRelax protocol in the module pyrosetta.rosetta.protocols.relax. FastRelax was run with constraints to starting coordinates and scored with the ref2015 score function. The resulting model was inspected for any large energetic penalties associated with the modelled peptide residues or those NUP98 and RAE1 residues interacting with the peptide, and was found to have none. The model was visualized in PyMOL (The PyMOL Molecular Graphics System, v.3.4, Schrödinger).

**ORF10 secondary structure prediction**

The secondary structure of ORF10 was predicted using JPRED (https://www.compbio.dundee.ac.uk/jpred/index.html)\(^{75}\).

**Protein E alignment**

Protein E sequences from SARS-CoV-2 (YP_009724392.1), SARS-CoV (NP_828854.1) and bat SARS-like CoV (AGZ48809.1) were aligned using Clustal Omega\(^{76}\) and then manually aligned to the sequences of histone H3 (P68431) and influenza A H3N2 NS1 (YP_308845.1).

**Cheminformatic analysis of SARS-CoV-2-interacting partners**

To identify drugs and reagents that modulate the 332 host factors that interact with SARS-CoV-2 and HEK-293T/17 cells (MIST \(\geq 0.7\)), we used two approaches: (1) a chemoinformatic analysis of open-source chemical databases and (2) a target- and pathway-specific literature search, drawing on specialist knowledge within our group. Chemoinformatically, we retrieved 2,472 molecules from the UPHAR/BPS Guide to Pharmacology (2020-3-12)\(^{48}\) (Supplementary Table 7) that interacted with 30 human ‘prey’ proteins (38 approved, 71 in clinical trials), and found 10,883 molecules (95 approved, 369 in clinical trials) from the ChEMBL database\(^{48}\) (Supplementary Table 8). For both approaches, molecules were prioritized on their FDA approval status, activity at the target of interest better than 1 \(\mu\)M and commercial availability, drawing on the ZINC database\(^{48}\). FDA-approved molecules were prioritized except when clinical candidates or preclinical research molecules had substantially better selectivity or potency on-target. In some cases, we considered molecules with indirect mechanisms of action on the general pathway of interest based solely on literature evidence (for example, captopril modulates ACE2 indirectly via its direct interaction with angiotensin-converting enzyme, ACE). Finally, we predicted 6 additional molecules (2 approved, 1 in clinical trials) for proteins with MIST scores between 0.7 and 0.6 to viral baits (Supplementary Tables 4, 5). Complete methods can be found at https://github.com/momeara/BioChemPantry/tree/master/vignette/COVID19.

**Molecular docking**

After their chemoinformatic assignment to the sigma-1 receptor, cloperastine and clemastine were docked into the agonist-bound state structure of the receptor (6DKI)\(^{79}\) using DOCK3.7\(^{80}\). The best scoring configurations that ion pair with Glu172 are shown; both l-cloperastine and clemastine were docked into the agonist-bound state configuration that ion pair with Glu172 are shown; both l-cloperastine and clemastine were docked into the agonist-bound state configuration. Molecular docking

**Viral growth and cytotoxicity assays in the presence of inhibitors**

For studies carried out at Mount Sinai, SARS-CoV-2 (isolate USA-WA1/2020 from BEI RESOURCES NR-52281) was propagated in Vero E6 cells. Two thousand Vero E6 cells were seeded into 96-well plates in DMEM (10% FBS) and incubated for 24 h at 37 °C, 5% CO\(_2\).
Vero E6 cells were used at passage number 20 derived from the Institut Pasteur in compliance with the guidelines of the Institut Pasteur following BSL3 containment procedures in approved laboratories. All experiments were performed in at least three biologically independent samples.

Antiviral activity assays

Vero E6 cells were seeded at 1.5 × 10^5 cells per well in 96-well plates 18 h before the experiment. Then, 2 h before infection, the cell-culture supernatant of triplicate wells was replaced with medium containing 10 μM, 2 μM, 500 nM, 200 nM, 100 nM or 10 nM of each compound or the equivalent volume of maximum DMSO vehicle used as a control. At the time of infection, the drug-containing medium was removed, and replaced with virus inoculum (MOI of 0.1 PFU per cell) containing TPCK-trypsin (Sigma Aldrich). Following a 1 h adsorption incubation at 37 °C, the virus inoculum was removed and 200 μl of drug- or vehicle-containing medium was added. Then, 48 h after infection, the cell-culture supernatant was used to extract RNA using the Direct-zol-96 RNA extraction kit (Zymo) following the manufacturer’s instructions. Detection of viral genomes in the extracted RNA was performed by RT-qPCR, using previously published SARS-CoV-2-specific primers. Specifically, the primers target the N gene region: 5′-TAATCAGACAAGGAACTGATTA-3′ (forward) and 5′-CAGAGGTGACTCCTCATG-3′ (reverse). RT-qPCR was performed using the Luna Universal One-Step RT–qPCR Kit (NEB) in an Applied Biosystems QuantStudio 6 thermocycler, using the following cycling conditions: 55 °C for 10 min, 95 °C for 1 min, and 40 cycles of 95 °C for 15 s, followed by 60 °C for 1 min. The quantity of viral genomes is expressed as PFU equivalents, and was calculated by performing a standard curve with RNA derived from a viral stock with a known viral titre. In addition to measuring viral RNA in the supernatant derived from drug-treated cells, infectious virus was quantified by plaque-forming assay.

Cell viability assays

Cell viability in drug-treated cells was measured using Alamar blue reagent (ThermoFisher). In brief, 48 h after treatment, the drug-containing medium was removed and replaced with Alamar blue and incubated for 1 h at 37 °C for fluorescence measurement in a Tecxan Infinity 2000 plate reader. Percentage viability was calculated relative to untreated cells (100% viability) and cells lysed with 20% ethanol (0% viability), included in each plate.

Plaque-forming assays

Viruses were quantified by plaque-forming assays. For this, Vero E6 cells were seeded in 24-well plates at a concentration of 7.5 × 10^4 cells per well. The following day, tenfold serial dilutions of individual virus samples in serum-free MEM medium were added to infect the cells at 37 °C for 1 h. After the adsorption time, the overlay medium was added at final concentration of 2% FBS/MEM medium and 0.05% agarose to achieve a semi-solid overlay. Plaque-forming assays were incubated at 37 °C for 3 days before fixation with 4% formalin and visualization using crystal violet solution.

Off-target assays for Sigma receptor drugs and ligands

hERG binding assays were carried out as previously described. In brief, compounds were incubated with hERG membranes, prepared from HEK-293 cells stably expressing hERG channels, and [3H]dofetilide (5 nM final) in a total of 150 μl for 90 min at room temperature in the dark. Reactions were stopped by filtering the mixture onto a glass fibre and were quickly washed three times to remove unbound [3H]dofetilide. The filter was dried in a microwave, melted with a scintillant cocktail and wrapped in a plastic film. Radioactivity was counted on a MicroBeta counter and results were analysed in Prism by fitting the built-in one binding function to obtain affinity Kᵢ. Radioligand binding assays for the muscarinic and alpha-adrenergic receptors were performed as previously described. Detailed protocols are available.
on the NIMH PDSP website at https://pspdb.unc.edu/html/tutorials/UNC-CH%20Protocol%20Book.pdf.

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

Data availability
The AP-MS raw data and search results files generated during the current study are available in the ProteomeXchange Consortium via the PRIDE partner repository with dataset identifier PXD018117 (https://www.ebi.ac.uk/pride/archive/projects/PXD018117) and PPI networks have also been uploaded to NDEX (https://public.ndexbio.org/#/network/43803262-6d99-11ea-bfbdc-9ac135e8b9ac). An interactive version of these networks, including relevant drug and functional information, can be found at http://kroganlab.ucsf.edu/network-maps. All data generated or analysed during this study are included in the article and its Supplementary Information. Expression vectors used in this study are readily available from the authors to biomedical researchers and educators in the non-profit sector. Source data are provided with this paper.

Code availability
Complete methods for chemoinformatics analysis can be found on Github (https://github.com/momeara/BioChemPantry/tree/master/vignette/COV19D19); details on MiST scoring can be found on Github (https://github.com/kroganlab/mist/).

Acknowledgements
We thank J. Sarlo for his artistic contributions to the manuscript; A. Kruse for his advice on Sigma receptor functional activity and pharmacologically related enzymes; S. Worland and eFFECTOR Therapeutics for providing zotatifin; T. Haley for his effort helping us to ship out drugs and SARS-CoV-2 expression constructs across the globe; R. Albrecht for support with BSL3 procedures and their partners and educators in the non-profit sector. Source data are provided with this paper.

Author contributions
The study was conceived by N.J.K. and D.E.G. Genome annotation was carried out by D.E.G., G.M.J. and T.L.R. and maximal peptide ratio extraction, termed MaxLFQ.

75. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Biol. Evol. 32, 549–555 (2020).

76. Huang, X.-P., Mangano, T., Hufeisen, S., Setola, V. & Roth, B. L. Identification of human Ether-a-go-go related gene modulators by three screening platforms in an academic drug discovery setting. Assay Drug. Dev. Technol. 8, 727–736 (2010).

77. Sterling, T. & Irwin, J. J. ZINC15—ligand discovery for everyone. J. Chem. Inf. Model. 55, 2324–2333 (2015).

78. Schmidt, H. R., Betz, R. M., Dorr, R. O. & Kruse, A. C. Structural basis for α, receptor ligand recognition. Nat. Struct. Mol. Biol. 25, 981–987 (2018).

79. Mysinger, M. M. & Shoichet, B. K. Rapid context-dependent ligand desolvation in molecular docking. J. Chem. Inf. Model. 50, 1561–1573 (2010).

80. Reed, L. J. & Muench, H. A simple method of estimating fifty percent endpoints. Am. J. Hyg. 27, 489–497 (1938).

81. Chu, D. K. W. et al. Molecular diagnosis of a novel coronavirus (2019-nCoV) causing an outbreak of pneumonia. Clin. Chem. 66, 549–555 (2020).

82. Jimenez-Morales, D., Rosa Campos, A., Von Dollen, J., Krogan, N. J. & Swaney, D. L. artMS: analytical R tools for mass spectrometry. R package version 1.6.5 http://bioconductor.org/ packages/release/bioc/html/artMS.html (2020).

83. El-Gebali, S. et al. The Pfam protein families database in 2019. Nucleic Acids Res. 47, D427–D432 (2019).
ShangPharma/ChemPartners. A.A. is a co-founder of Tango Therapeutics, Azkarra Therapeutics, Ovibio Corporation, a consultant for SPARC, Bluestar, ProLymx, Earli, Cura, GenVivo and GSK; a member of the SAB of Genentech and GLAdiator; receives grant/research support from SPARC and AstraZeneca; holds patents on the use of PARP inhibitors held jointly with AstraZeneca, from which he has benefitted financially (and may do so in the future). The authors have not filed for patent protection on the SARS-CoV-2 host interactions or the use of predicted drugs for the treatment of COVID-19 to ensure all of the information is freely available to accelerate the discovery of a treatment.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2286-9.
Correspondence and requests for materials should be addressed to M.V., A.G.-S., K.M.S., B.K.S. or N.J.K.
Peer review information Nature thanks Michael P. Weekes and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.
Reprints and permissions information is available at http://www.nature.com/reprints.
Extended Data Fig. 1 | Mutations in overlapping coding regions result in premature termination of ORF3a and ORF9c. a, Table of the SARS-CoV-2 proteins, including molecular mass, sequence similarity with the SARS-CoV homologue and inferred function based on the SARS-CoV homologue. b, Immunoblot detection of 2×Strep tag demonstrates expression of each bait in input samples, as indicated by the red arrowheads. For each bait, input from one of the three replicates prepared and affinity purified for mass spectrometry was used for western blot (n = 1). For gel source data, see Supplementary Fig. 1. c, Schematic of ORF3a (light green) and ORF3b (dark green) overlapping regions. A premature stop codon in ORF3b at position 14 (E14*) corresponds to a Q57H mutation in ORF3a. d, Schematic of the N (red), ORF9b (green) and ORF9c (green) overlapping regions. Two mutations in the N protein (S194L and S197L) correspond to premature stop codons at positions 41 and 44 in ORF9c. The analysis is based on 2,784 sequences obtained from GISAID on 4 April 2020.
Extended Data Fig. 2 | Clustering analysis of the AP-MS dataset reveals good correlation between biological replicates of individual baits. All AP-MS runs ($n = 3$ biologically independent samples) were compared and clustered using artMS$^{45}$. All Pearson’s pairwise correlations between MS runs are shown and are clustered according to similar correlation patterns.

Correlation between replicates for individual baits ranges from 0.46 to 0.72, and in most cases the experiments corresponding to each bait cluster together, with the exception of a few baits with lower numbers of specific host interactions (for example, E, NSP2, ORF6, ORF3a and ORF3b).
Extended Data Fig. 3 | Enrichment of GO biological processes for SARS-CoV-2 host factors. We performed GO biological process enrichment analyses (Methods) for the host factors identified as binding to each SARS-CoV-2 viral protein and the top five most significant terms for each viral protein are shown. The P-values were calculated using a hypergeometric test and a false-discovery rate was used to account for multiple-hypothesis testing.
Extended Data Fig. 4 | Enrichment of Pfam protein families for SARS-CoV-2 host factors. The enrichments of individual protein family domains from the Pfam database was calculated using a hypergeometric test, for which success is defined as the number of domains, and the number of trials is the number of individual preys that were affinity purified with each viral bait. The population values were the numbers of individual Pfam domains in the human proteome. The P values were not adjusted for multiple testing. To make sure that the P values that indicated enrichment were meaningful, we only considered Pfam domains that have been affinity purified at least three times with any SARS-CoV-2 protein and that occurred in the human proteome at least five times. Here, we show Pfam domains with the lowest P value for a given viral bait protein.
Extended Data Fig. 5 | Lung mRNA expression and specificity of SARS-CoV-2-interacting human proteins relative to other proteins.  

**a**, Scatterplot of the lung mRNA expression (TPM) versus enrichment of lung mRNA expression (lung TPM/median of all tissue TPM) for human-interacting proteins. Red points denote drug targets that are labelled with their gene names. Points above the horizontal blue line represent interacting proteins that are enriched in lung expression and show how most SARS-CoV-2-interacting proteins tend to be enriched in the lung. **b**, Gene expression in the lung of the high-confidence human-interacting proteins was observed to be higher compared to all other proteins. Blue, interacting proteins \((n = 332, \text{median} = 25.52 \text{ TPM})\); grey, all other proteins \((n = 13,583, \text{median} = 3.198 \text{ TPM})\). \(P = 0.0007\) using a Student’s \(t\)-test.
Extended Data Fig. 6 | Candidate targets for the viral NSP5 protease.

a, Wild-type NSP5 and NSP5(C145A) (catalytic dead mutant) interactome.
b, Domain maps of HDAC2 and TRMT1, illustrating predicted cleavage sites (using NetCorona 1.0). HDAC, histone deacetylase domain; NLS, nuclear localization sequence; MTS, mitochondrial targeting sequence; SAM-MT, S-adenosylmethionine-dependent methyltransferase domain.
c, Peptide docking of predicted cleavage peptides into the crystal structure of SARS-CoV NSP5. d, NSP5-cleavage consensus site for SARS-CoV (left) and SARS-CoV-2 (right).
Extended Data Fig. 7 | Consensus analysis of SARS-CoV-2 ORF6 homologues.

**a**, Sequence logo of SARS-CoV-2 ORF6 homologues, showing sequence conservation at each position computed from a multiple-sequence alignment of 35 sequences. The key methionine M58, and the acidic residues E55, E59 and D61 of the putative NUP98–RAE1-binding motif are shown to be highly conserved. Homology was determined from alignments to full-length sequences. Colours indicated chemical properties of amino acids: polar (G, S, T, Y, C; green), neutral (Q, N; purple), basic (K, R, H; blue), acidic (D, E; red) and hydrophobic (A, V, L, I, P, W, F, M; black).

**b**, Multiple-sequence alignment of SARS-CoV-2 ORF6 homologues. The query sequence is shown at the top (sequence 1, ref|YP_009724394.1). Sequence coverage (cov) and percentage identity (pid) are shown for each homologous sequence.
Extended Data Fig. 8 | Viral growth and cytotoxicity for compounds tested in New York. Viral growth (percentage infection; red) and cytotoxicity (black) results for compounds tested at Mount Sinai in New York. TCID$_{50}$ assay results (green) for zotatifin, hydroxychloroquine and PB28 are also shown. Zotatifin and midostaurin were tested in two independent experiments and data are shown in two individual panels. Data are mean ± s.d.; n = 3 biologically independent samples. The full dataset is available in Supplementary Table 6.
Extended Data Fig. 9 | Virus plaque assays, qRT–PCR and cell viability for compounds tested in Paris. Plaque assay (viral titre; red), qRT–PCR (viral RNA; blue) and cell viability (Alamar blue; black) results for compounds tested at the Pasteur Institute in Paris. PF-846 was tested in two independent experiments and data are shown in two individual panels. Data are mean ± s.d.; for virus plaque assay and RT- qPCR, n = 3 biologically independent samples for drug-treated cells, n = 5 for PS3061, n = 6 for DMSO controls; for cell viability, n = 6 biologically independent samples for drug-treated cells and DMSO controls. The full dataset is available in Supplementary Table 6.
Extended Data Fig. 10 | Activity of sigma ligands. a, The drugs cloperastine and clemastine can be readily fit into the agonist-bound structure of the sigma-1 receptor. b, Compounds tested for antiviral activity with annotated sigma-1 receptor and/or sigma-2 receptor (also known as TMEM97) activity are shown. Inhibition pIC$_{50}$ values of SARS-CoV-2 infection are shown from blue to yellow, mode of functional activity at the sigma-1 receptor is shown by mark shape (upwards triangle, agonist; downwards triangle, antagonist; circle, binding), and pKi values for the sigma-1 receptor and sigma-2 receptor are shown along the x and y axes. We have not yet tested chloroquine for antiviral activity. Binding of E-52862 at the sigma-2 receptor is reported to be greater than 1 μM. Activities of pimozide and olanzapine at the sigma-2 receptor have not been reported. Activity of olanzapine at the sigma-1 receptor is reported to be greater than 5 μM.
Extended Data Fig. 11 | Astemizole is a potent sigma-2 receptor ligand. \(a, b\). Concentration–response curves of astemizole from radioligand displacement assays for the sigma-2 (\(K_i = 95\) nM) and the sigma-1 (\(K_i = 1.3\) μM) receptors are shown. Data are mean ± s.e.m.; \(n = 4\) independent assays for each receptor.
Extended Data Fig. 12 | Off-target activities for characteristic Sigma receptor ligands. Dose–response curves against a panel of eight targets that can confer adverse cardiac responses, respiratory difficulties and dry-mouth effects for chloroquine, hydroxychloroquine, PB28, PD-144418 and clemastine. These results are not meant to represent or replace a comprehensive test against off-target panels, as might commonly be assayed in drug progression for clinical use. The eight targets include the alpha-2A adrenergic receptors: alpha 2A (encoded by \textit{ADRA2A}), alpha 2B (encoded by \textit{ADRA2B}), and alpha 2C (encoded by \textit{ADRA2C}); as well as the muscarinic acetylcholine receptors: M1 (encoded by \textit{CHRM1}), M2 (encoded by \textit{CHRM2}), M3 (encoded by \textit{CHRM3}), M4 (encoded by \textit{CHRM4}) and M5 (encoded by \textit{CHRM5}). Data are mean ± s.d.; \( n = 3 \) biologically independent samples. The table summarizes the fitted pKi values for the five ligands at the eight targets.
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 Authors & Referees 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

- Confirm 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

All MS data was acquired on a Thermo Fisher Scientific Q-Exactive Plus mass spectrometer using the Thermo software Xcalibur (4.2.47) and Tune (2.11 Q1 Build 3006).

Data analysis

Raw mass spectrometry data were searched using MaxQuant (version 1.6.11.0) and scored using MIST (available at https://github.com/kroganlab/mist) and SAINT (version 3.6.3). Custom scripts were designed to map interacting proteins to drugs and compounds (available at https://github.com/momeara/BioChemPantry/tree/master/vignette/COVID19). Data on virus assays were analyzed using GraphPad Prism version 7.00 for Mac (GraphPad Software, La Jolla California USA, www.graphpad.com). High-confidence protein-protein interactions were visualized using Cytoscape version 3.7.1. The over-representation analysis (ORA) was performed using the enricher function of clusterProfiler package in R with default parameters. Protein E sequences were aligned using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/).

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/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

Data availability: The mass spectrometry raw data and search results files generated during the current study are available in the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD018117 (https://www.ebi.ac.uk/pride/archive/projects/PXD018117) and PPI networks have also been uploaded to NDEX (https://public.ndexbio.org/#/network/43803262-5d69-11ea-bf6c-0ac135ebacdf). All data generated or analyzed during this study are included...
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
It is an accepted practice in the field of large-scale interactomics (via AP-MS), that biological triplicate measurements are sufficient for measuring high confidence interactions using the methods and software performed in this study. At least three biological replicates were independently prepared for affinity purification.

All antiviral experiments at Mount Sinai were performed in triplicate. For experiments performed at the Institut Pasteur: A sample size of n=3 was chosen for each treatment assessing the effect of drug treatment on viral RNA and infectious units following infection in pre-treated cells. In control-treated cells, n=6. This sample size was chosen as it is sufficient in medium-throughput screening in order to identify differences in viral replication in vitro following drug treatment compared to control conditions. A sample size of n=6 was chosen for assessing the effect of drug treatment on cell viability (alamar blue). This was because this larger sample size enabled us to identify any outliers.

Data exclusions
On cell viability assays performed at the Institut Pasteur, replicates were excluded if they were considered outliers (if cell viability was significantly different compared to other replicates). No other data were excluded from the study.

Replication
Reproducibility between bioreplicates can be measured by the degree of variance explained by matching LC-MS feature identifications (peptide and charge) between replicates. We used standard artMS procedures. First, LC-MS features were identified and quantified by MaxQuant in each LC-MS run. Next, the strength of effect was measured as a correlation coefficient (Pearson’s r) between each pair of LC-MS runs, pairing individual feature intensities between runs by their peptide and charge identifications. Correlation patterns between LC-MS runs from biological replicates are clustered here along the x and y axes, showing both high correlation coefficients (near 1.0) as well as a trend for most same-bait replicates to cluster by similarity with each other, indicating consistent and bait-specific results.

Randomization
Sample randomization is not relevant to our study because experimental groups do not exist. Moreover, AP-MS samples were processed and collected on the same instruments in a short time frame (roughly 3 weeks time). Therefore instrument performance did not have time to drift. QCloud was used to control instrument longitudinal performance during the project.

Blinding
Blinding is not relevant to the AP-MS data because our data are acquired and processed systematically with established scoring algorithms, excluding human bias. For viral assays at Institut Pasteur, Investigators were blinded to group allocation by defining each drug with a number. The name of each drug (numbered 1 to 66) were not revealed to investigators during the screening. Additionally, different investigators were involved at different stages of the process (pre-treatment, infection, data collection, analysis).

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 | Methods |
|---------------------------------|---------|
| n/a | Involved in the study | n/a | Involved in the study |
| ☑ Antibodies | ☑ ChIP-seq |
| ☑ Eukaryotic cell lines | ☑ Flow cytometry |
| ☑ Palaeontology | ☑ MRI-based neuroimaging |
| ☑ Animals and other organisms | | |
| ☑ Human research participants | | |
| ☑ Clinical data | | |

Antibodies

Antibodies used: Anti-strep antibody, Qiagen # 34850 (1:2,500). Anti-mouse-HRP conjugate, BioRad # 1706516 (1:20,000)
Poly-clonal anti-SARS-CoV-NP antisera produced in a single rabbit, Garcia-Sastre lab at Mount Sinai (1:10,000). Due to the current shelter-in-place order we were unable to identify the lot numbers of commercially available antibodies.

**Validation**

Use of the anti-strep antibody and anti-mouse HRP conjugate by western blot only detects signal in cell lysate from cells expressing Strep-tagged fusion proteins.

### Eukaryotic cell lines

**Policy information about cell lines**

**Cell line source(s)**

HEK-293T/17 cells were procured from the UCSF Cell Culture Facility, now available through UCSF’s Cell and Genome Engineering Core (https://cgec.ucsf.edu/cell-culture-and-banking-services); cell line collection listed here: https://ucsf.app.box.com/s/6xydeqhr8a2xes0mb023333i3k1ndqv (CCLZR076). Vero E6 cells used at Mount Sinai and Institut Pasteur were purchased from ATCC (VERO C1008 [Vero 76, clone E6, Vero E6] (ATCC® CRL-1586™)).

**Authentication**

STR analysis by the Berkeley Cell Culture Facility on August 8, 2017 authenticates our HEK-293T/17 cells with 94% probability. The African green monkey kidney epithelial Vero E6 (ATCC CRL-1586) is derived from ATCC, and thus is already authenticated.

**Mycoplasma contamination**

Cells were tested on July 3, 2019 using the MycoAlert™ Mycoplasma Detection Kit (Lonza LT07-318) and were negative: B/A ratio < 1 (no detected mycoplasma).

Vero E6 cells: The cell line was tested for mycoplasma contamination prior to commencement of experiments and was negative.

**Commonly misidentified lines**

No commonly misidentified cell lines were used in this study.