

# Proteomics-Based Insights Into the SARS-CoV-2-Mediated COVID-19 Pandemic: A Review of the First Year of Research

## Authors

Jeremy L. Praissman, and Lance Wells

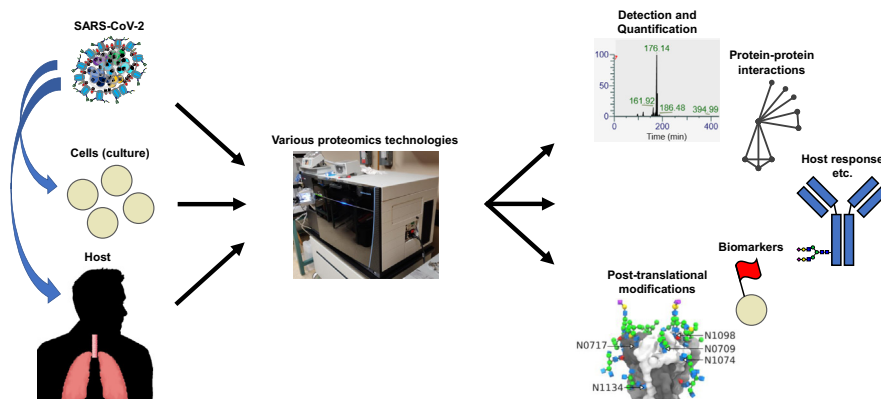
## Correspondence

[lwells@ccrc.uga.edu](mailto:lwells@ccrc.uga.edu)

## Graphical Abstract

### In Brief

SARS-CoV-2, the betacoronavirus that caused the COVID-19 pandemic, became a major source of human disease and death in 2020. The fundamental constituents of a virus being its genome and proteome, characterizing the proteome is essential to understanding its biology. In this review article, we survey the proteomics literature from the first year of the COVID-19 pandemic, including protein-protein interaction studies, post-translational modification studies, and work using proteomics technologies to probe host response, which collectively inform efforts to ameliorate the pandemic.



### Highlights

- SARS-CoV-2, which caused the COVID-19 pandemic, depends on features of its proteome and host proteomes for transmissibility and virulence.
- Qualitative and quantitative proteomic assay development underlies attempts to further understand the virus.
- Protein-protein interaction and post-translational modification studies inform efforts to develop therapies.
- In sum, the range of proteomics technologies provides valuable information in the fight against COVID-19.

# Proteomics-Based Insights Into the SARS-CoV-2–Mediated COVID-19 Pandemic: A Review of the First Year of Research

Jeremy L. Praissman and Lance Wells\*

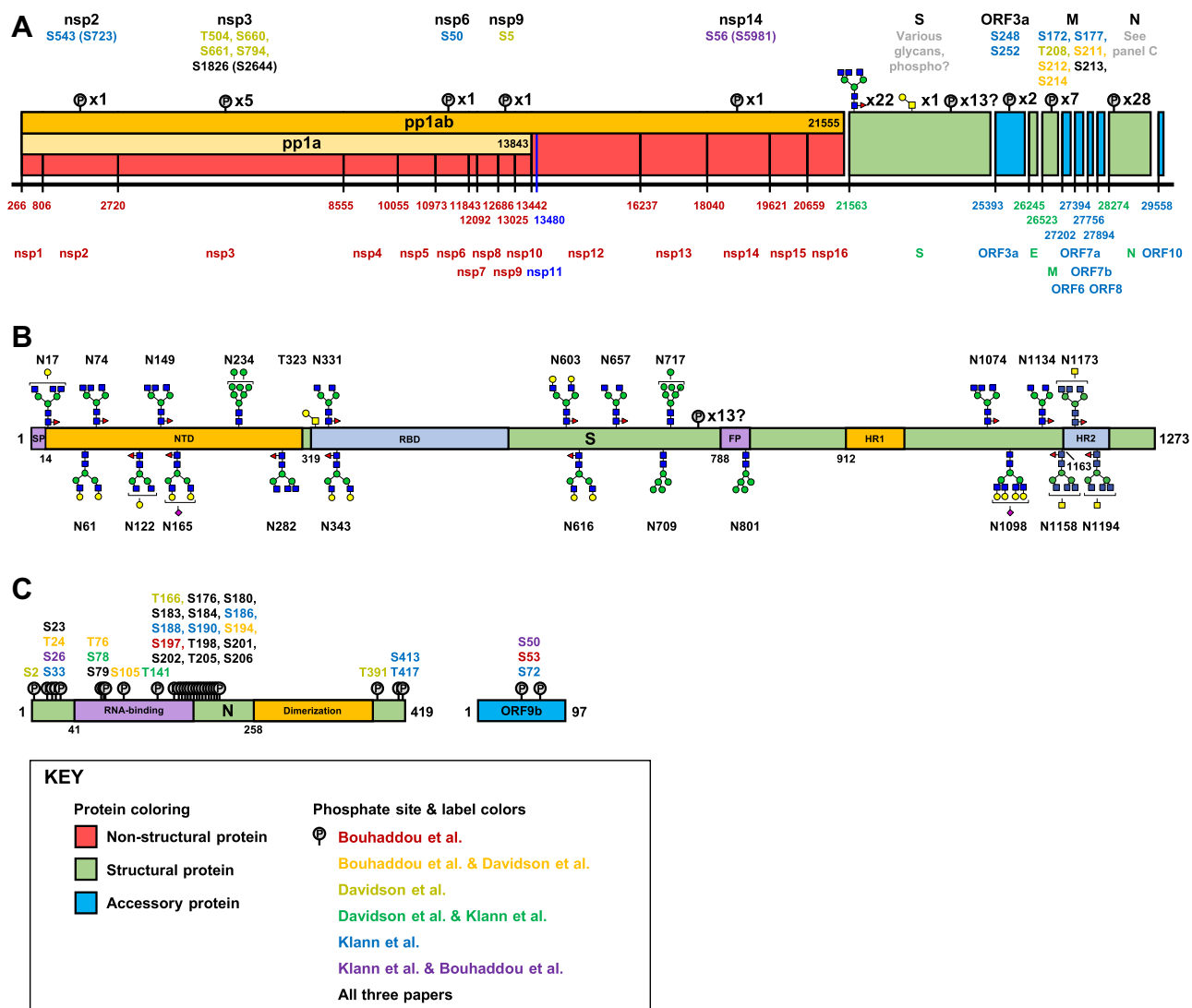
In late 2019, a virus subsequently named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged in China and led to a worldwide pandemic of the disease termed coronavirus disease 2019. The global health threat posed by this pandemic led to an extremely rapid and robust mobilization of the scientific and medical communities as evidenced by the publication of more than 10,000 peer-reviewed articles and thousands of preprints in the first year of the pandemic alone. With the publication of the initial genome sequence of SARS-CoV-2, the proteomics community immediately joined this effort publishing, to date, more than 100 peer-reviewed proteomics studies and submitting many more preprints to preprint servers. In this review, we focus on peer-reviewed articles published on the proteome, glycoproteome, and glycome of SARS-CoV-2. At a basic level, proteomic studies provide valuable information on quantitative aspects of viral infection course; information on the identities, sites, and microheterogeneity of post-translational modifications; and, information on protein–protein interactions. At a biological systems level, these studies elucidate host cell and tissue responses, characterize antibodies and other immune system factors in infection, suggest biomarkers that may be useful for diagnosis and disease-course monitoring, and help in the development or repurposing of potential therapeutics. Here, we summarize results from selected early studies to provide a perspective on the current rapidly evolving literature.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a betacoronavirus that began infecting people in 2019 with the index case identified as a hospitalized patient who initially became ill on December 1, 2019 (1). The genome of the then unknown etiological agent was rapidly sequenced and made available to other researchers in early January (2–4). The SARS-CoV-2 genome reported (29,903 bases, single-stranded RNA) was annotated as encoding 26 or more proteins and has high sequence similarity (~80% identity at the nucleotide level) to the extensively studied SARS-CoV-1 responsible for SARS outbreaks in 2002 and 2003 as well as to multiple animal coronaviruses (2, 4–6).

The SARS-CoV-2 National Center for Biotechnology Information reference genome was released shortly after publication of the initial genome sequences and contains annotations for 28 encoded proteins. Starting from the 5′-end, the annotated proteins consist of 16 nonstructural proteins (denoted as nsp1–nsp16), translated as components of large polyproteins and then separated by viral proteases, followed by structural proteins and additional ORFs at the 3′-end (Fig. 1). Putative functions of the encoded proteins were initially inferred by sequence homology to previously studied coronaviruses (2, 4, 5). The nsp proteins include two with protease functions essential for polyprotein processing—Mpro, also called 3CLpro (3C-like protease, nsp5) and PLpro (papain-like protease, nsp3)—as well as the viral replication–transcription complex subunits. Structural proteins encoded include spike (S), envelope (E), membrane (M), and nucleocapsid (N). A number of accessory proteins that have (partially) determined roles in host defense interference, intracellular trafficking, transcription, and replication in related coronaviruses are also encoded (4, 7, 8). In the context of proteomics, it is important to note that such homology-based annotations provide useful initial models but are necessarily subject to more definitive empirical characterization. This is particularly essential for most RNA viruses because of their comparatively compact genomes and the resulting multiform and multifunction nature of their encoded proteins driven by evolutionary constraints (9). Consequently, different groups have used a variety of marginally different annotations in their studies. Researchers have, therefore, continued to work on refining annotations of regulatory elements and encoded polypeptides that may not have been completely characterized by homology-based methods (10–14). For example, Finkel *et al.* (10) reported 23 unannotated ORFs in their study using ribosome profiling, and Davidson *et al.* (14) reported that 14% of the transcripts detected in their study do not code for a known ORF and subsequently identified peptides from these transcripts. To our knowledge, there is no current definitive database compiling this information, and interested readers are directed to the original research articles.

From the Complex Carbohydrate Research Center, University of Georgia, Athens, Georgia, USA

\*For correspondence: Lance Wells, [lwells@ccrc.uga.edu](mailto:lwells@ccrc.uga.edu).



**FIG. 1. The SARS-CoV-2 proteome and its post-translational modifications (PTMs).** The SARS-CoV-2 NCBI reference sequence proteome delineated along its genome (A). The 28 proteins annotated in the NCBI reference sequence are represented as boxes with the starting base corresponding to each protein in the genome listed later along with most protein names (pp1ab and pp1a are labeled inside boxes). Note that the nsp proteins are expressed as parts of large polyproteins (pp1ab and pp1a), which are subsequently cleaved by proteases contained in the polyproteins themselves. A summary of PTMs detected in proteomics studies is listed above each protein except for N and S, which are shown in detail in panels B and C. Numbers in parentheses indicate the residue number in pp1ab as given in the study by Klann *et al.* (102). The PTMs of S. A partial domain structure is shown for orientation with coloring for contrast and start residue numbers. The most abundant N-glycans from the most abundant Oxford class at each site are shown as reported by Zhao *et al.* (86). The class abundances at each site reported by Watanabe *et al.* (83) are similar although the protein they analyzed showed a small but clear tendency toward slightly less processed glycoforms. Articles have reported varying amounts of O-glycosylation on S almost exclusively at T323, occupancy generally ~10% or less. Note also that Davidson *et al.* (14) identified 13 sites of phosphorylation on S; however, most were not cytoplasmic. Secretory pathway kinases have been confirmed (e.g., FAM20C), but it is not clear that these sites fit with known specificity determinants. The PTMs of N and ORF9b. Domain structure shown with coloring for contrast and start residue numbers. ORF9b is an alternative ORF in the N coding sequence that is not annotated in the NCBI reference sequence. FP, fusion peptide; HR1, heptad repeat 1; HR2, heptad repeat 2; NCBI, National Center for Biotechnology Information; nsp, nonstructural protein; RBD, receptor binding domain; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

While the genome and preliminary information on the proteome of the SARS-CoV-2 itself were being defined, researchers also began determining the host cell proteins required for or facilitative of infection. Angiotensin-converting enzyme 2 (ACE2) was known to be the host cell surface

receptor for several other coronaviruses, including SARS-CoV-1, and its identity as the host cell surface receptor for SARS-CoV-2 was quickly confirmed (15–18). Mature membrane-bound ACE2 is a 788-amino acid single-pass type I membrane protein (~91 kDa without post-translational

modifications [PTMs]) consisting of an N-terminal peptidase domain and a C-terminal collectrin-like domain that causes homodimerization at the cell surface and contains the transmembrane helix (19–21). Many viruses in addition utilize host cell proteases and other host cell machinery to enable and facilitate initial cell infection. The protease transmembrane serine protease 2 (TMPRSS2) was confirmed as a key factor in SARS-CoV-2 infection in one of the early articles confirming the identity of the cell surface receptor as ACE2 (18). Researchers have since continued to pursue potential alternative host cell surface receptors, related or alternative proteases (TMPRSS4, cathepsin B and L), and proteins involved in processes such as endosome maturation (a FYVE finger-containing phosphoinositide kinase [PIKFyve], two-pore channel 2) that are critical for infection (18, 22–26). However, many questions still remain regarding the role and behavior of host proteases and proteins during infection.

While experimental studies have continued, furthering understanding of the biology of SARS-CoV-2, important parallel efforts have focused on cataloging and increasing the accessibility of this information. Of particular note to the proteomics community are resources compiling genome sequences, annotations, protein–protein interactions (PPIs), PTMs, and proteomics datasets. Genome sequences for SARS-CoV-2 from which protein sequences may be derived are available from the Global Initiative on Sharing Avian Influenza Data (27), National Center for Biotechnology Information (28), European Molecular Biology Laboratory's European Bioinformatics Institute (29), and other organizations. The coronavirus disease 2019 (COVID-19) data portal (<https://www.covid19dataportal.org/>) from European Molecular Biology Laboratory's European Bioinformatics Institute maintains an updated curated collection of SARS-CoV-2 and host proteins and their relationships and information on pathways and from gene expression studies (29). A database specific for PPIs is available from the Biological General Repository for Interaction Datasets curation project (<https://thebiogrid.org/project/3>) and may be consulted for continuously updated information (30). Many proteomics datasets from relevant studies are available through ProteomeXchange and its subsidiary databases (<http://www.proteomexchange.org/>) (31). A number of additional databases contain proteomics data particularly useful for analyzing (or developing assays to analyze) host cell factors, for example, proteomicsDB (<https://www.proteomicsdb.org/>) (32), PAXdb (<https://pax-db.org/>) (33), the Clinical Proteomic Tumor Analysis Consortium Data Portal (<https://cptac-data-portal.georgetown.edu/cptacPublic/>) (34), Human Proteome Map (<http://www.humanproteomemap.org/>) (35), and The Human Protein Atlas (<https://www.proteinatlas.org/>) (36). Glycan information from various studies is compiled in GlyGen (<https://www.glygen.org/>) (37). In addition, glycan and other PTM information is available through PhosphoSitePlus (<https://www.phosphosite.org/>) (38).

In the first half of this article, we focus on the virus and its cell entry factors including the host cell receptor ACE2. The literature in this area may be further divided among (1) studies examining the basic qualitative behavior of viral and host cell entry factor peptides in mass spectrometric experiments; (2) quantitative proteomics studies that either detail viral protein expression over time or examine the distribution of host cell entry factors in human tissues and cells; and (3) studies of viral and ACE2 structure and PTMs. In the second half of this review, we provide an overview of studies focused on the proteomes of host (primarily human) cells and tissues and their responses and interactions with the SARS-CoV-2 virus. These studies encompass PPI mapping experiments, the quantitative proteomics of host cell protein expression during infection, determination of putative biomarkers, and characterization of immune system responses and SARS-CoV-2–directed antibodies during infection. Another review of SARS-CoV-2 proteomics was published during preparation of this article and may be of interest for further reading (39).

#### VIRUS AND HOST CELL ENTRY FACTOR STUDIES

##### *Basic Qualitative Proteomics and Potential Clinical Diagnostics*

A number of studies have been published containing information on the basic qualitative proteomics of the virus and the potential of proteomics technology–based assays for clinical diagnostics development. Important results include the determination of peptides suitable for targeted method development in LC–MS experiments in terms of level of detection and quantification, specificity and stability of amino acid sequences in reported genomes, and the presence or absence of PTMs. Testing of developed methods with relevant clinical samples has also been reported by several groups.

In an early study, Gouveia *et al.* (40) detected 101 (tryptic) peptides across six viral proteins (N, S, M, ORF1ab, ORF3a, and ORF8) from virus-infected Vero cells and further recommended 14 peptides for targeted assays (Table 1 in the article of Gouveia *et al.*). Zecha *et al.* (41) characterized both tryptic viral peptides (from cells or cell culture supernatant independently) and tryptic host cell peptides from four relevant cell line models (discussed further later in this review). Parallel reaction monitoring (PRM) assays for viral proteins were developed for 23 peptides with favorable properties in nanoflow PRM and 21 peptides in microflow PRM. For each peptide, the top six transitions are also detailed (supplemental Table S3 in the article by Zecha *et al.* (41), spectral libraries are available at Panorama Public (42)). Gouveia *et al.* and Zecha *et al.* recommended similar lists of peptides for targeted method development, sharing two peptides for M, five peptides for N, and one peptide for S. Additional articles have been published more recently and may be consulted for further reference (43–45) (Table 1).

TABLE 1  
Selected peptides for SARS-CoV-2 detection and quantification reported in at least two publications

Accession	Protein	Peptide sequence	Theoretical unmodified precursor [M + H] <sup>+</sup>	Observed precursor z	Common mods	References
VME1_SARS2	M	EITVATSR	876.48	2	None	(40, 41)
		VAGDSGFAAYSR	1200.57	2	None	(40, 41)
NCAP_SARS2	N	ADETQALPQR	1128.57	2	Deamidation (NQ)	(40, 41, 43, 44, 50)
		AYNVTQAFGR	1126.57	2	Deamidation (NQ)	(40, 41)
		GFYAEGRS	886.41	2	None	(40, 41, 50)
		IGMEVTPSGTWLTYTGAIK	2025.04	2, 3	Oxidation (M)	(40, 41)
SPIKE_SARS2	S	NPANNAIVLQLPQGTTLPK	2060.15	2, 3	Deamidation (NQ)	(40, 41)
		FQTLALHR	1098.64	3	None	(40, 44)
		LQSLQTYVTQQLIR	1690.95	2, 3	None	(40, 41)

Host cell entry factor peptides have also been qualitatively characterized. Zecha *et al.* (41) developed PRM assays for ACE2 and TMPRSS2 (human and, partly nonshared, monkey) targeting 16 tryptic human ACE2 sequences and six tryptic human TMPRSS2 sequences (supplemental Table S1 in the article by Zecha *et al.*, spectral libraries are available at Panorama Public (42)). They found that ACE2 could be detected in all four of their model cell lines (ACE2-A549, an ACE2 overexpressor, Vero E6, Calu-3, and Caco-2) using PRM methods but only in two using a data-dependent acquisition (DDA)-based method. TMPRSS2 was only detectable in two of the cell lines tested (Calu-3 and Caco-2). Other known and potential viral entry factors including TMPRSS4, CTSS, cathepsin L, BSG (CD147), and FURIN variously appeared across cell lines in DDA data (41). In addition, numerous articles have reported analyses of *de novo* or publicly available proteomics datasets aimed at characterizing the cell, tissue, and bodily fluid distribution of relevant proteins and may be useful as further references (46–49).

One longstanding goal in the proteomics field is the development of clinical diagnostics utilizing proteomics, and particularly MS-based, methods. This goal is of particular interest during a time of supply-chain disruptions and shortages of necessary reagents for PCR-based assays and other frequently used clinical laboratory methods. However, MS-based proteomics often suffers from sensitivity, specificity, and throughput issues. Zecha *et al.* (41) concluded that the PRM methods developed by their group were inadequate to serve as a reasonable clinical diagnostic platform. Furthermore, considering the literature more broadly (including pre-prints), they found wide variability among studies suggesting caution in attempting to apply these methods in a clinical setting. Additional studies may be consulted for further information on current progress in developing SARS-CoV-2 diagnostics from nasopharyngeal swabs, gargle solutions, other human samples, and simulated (mock) samples (from *in vitro*-derived mucus and inactivated virus) (43–45, 50–53).

Additional studies not substantially focused on characterization of viral or host cell entry factor peptides but containing lists of detected peptides and further relevant information

(often with deposited datasets available in various proteomics databases) have also been published (14, 51, 54–58).

### Quantitative Proteomics

Quantitative proteomics experiments involving SARS-CoV-2 and SARS-CoV-2 host cell entry factors generally fall into three categories: experiments quantifying virus (proteins), experiments examining the cell type and tissue distribution of host cell entry factors, and experiments quantifying changes in host cell (and other) factors during infection. Many of these experiments are natural extensions of the qualitative experiments outlined in the previous section.

Quantification of SARS-CoV-2 proteins has been carried out to understand the kinetics of viral infection, examine the effect of administration of potential therapeutics, and assess viral abundance in infected patients. Grenga *et al.* (58) characterized viral kinetics by assaying viral protein expression at several time points (day 1, 2, 3, 4, or 7) for cultures infected either at multiplicity of infection 0.01 or 0.001 on day 0. These results were compared with quantitative PCR (qPCR) measurements of viral RNA and found to be consistent validating their methods in this context (Grenga *et al.*; Fig. 2). Appelberg *et al.* (59) also looked at viral RNA levels by qPCR as compared with viral protein abundance over time (24, 48, or 72 h after infection) from SARS-CoV-2-infected Huh7 cells using a tandem mass tag (TMT)-based method (TMT-MS) and found similar trend agreement between qPCR results and their developed TMT-MS method (Appelberg *et al.*; Fig. 1). Gordon *et al.* (56) quantified virus protein expression at 8 h after infection in cell culture in relation to treatment (singly) with three potential therapeutic compounds previously identified in their study through other methods (two ligands of sigma-1 and sigma-1 receptors and one protein biogenesis inhibitor), confirming the effectiveness of these compounds in putatively disrupting viral replication. Zecha *et al.* applied their PRM-based methods to quantify viral proteins in patient samples, although their assays were designed for repeatability rather than for accurate and precise quantification. For samples in which SARS-CoV-2 peptides were detected, peptide intensities were generally in good correspondence with PCR

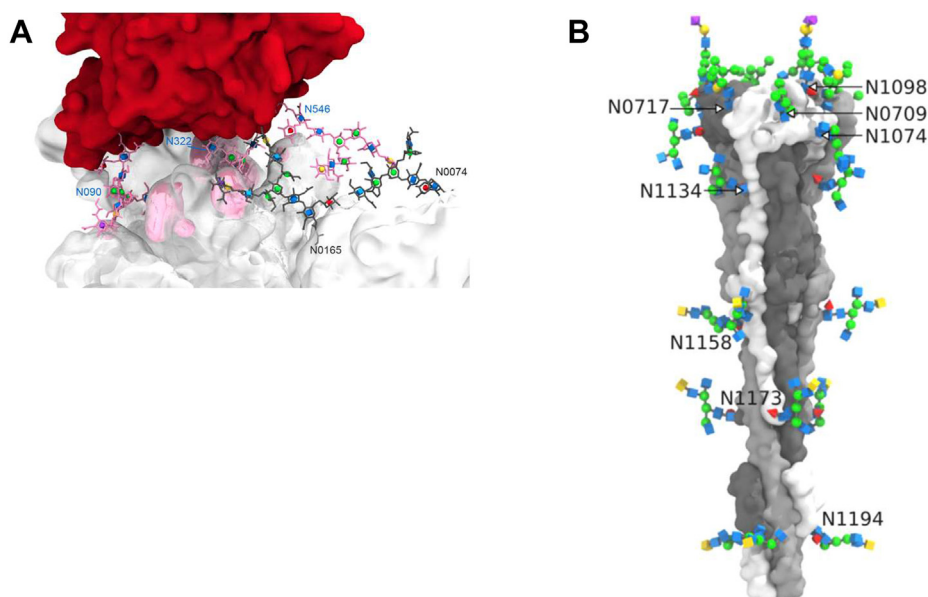


FIG. 2. **Views of the SARS-CoV-2 spike protein and its glycosylation.** Images courtesy of Oliver C. Grant (unused graphics from Zhao *et al.* (86)). Protein models courtesy of Professor Bing Chen. *A*, the interface of SARS-CoV-2 S (white) bound to ACE2 (red) showing glycans involved in glycan-peptide and glycan-glycan interactions. *B*, the postfusion structure of SARS-CoV-2 S showing its distinctive columnar structure and regular spacing of N-glycans. ACE2, angiotensin-converting enzyme 2; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

results. However, the PRM assay had a prohibitive rate of false negatives in patient samples (43 of 54 or approximately 80% false negative), a difficulty encountered with MS-based assays in general (41). Gouveia *et al.* further developed and tested a method based on their first article (on SARS-CoV-2 peptide analytical characteristics) establishing a lower limit of detection and concluding that two tryptic peptides from the nucleocapsid protein provide the best basis for a DDA (with inclusion list) reversed-phase LC-MS/MS-based diagnostic platform of the type described. However, these experiments similarly achieved a low rate of detection with diagnosed patient samples (two of nine patients or ~22% from a PCR validated cohort) (50). Relevant peptides from these studies are summarized in Table 1.

Several studies have been published to date examining the tissue and cell-type distribution of host cell entry factors. One of the classic studies is by Hamming *et al.* (60), published in 2004 shortly after the original SARS outbreaks and still a very relevant resource to consider newer research against, although not MS based. In more recent research, Zecha *et al.* (41) were able to relatively quantify ACE2 in their four cell line models, ACE2-A549, Vero E6, Calu-3, and Caco-2, using a PRM-based method they developed (discussed in more detail previously), finding expression to be more than 1000 times lower in Calu-3 and Caco-2 cells compared with ACE2-A549 cells. TMPRSS2 and other factors involved or putatively involved in host cell entry also varied widely in expression (Fig. 1, supplemental Fig. S1C, and supplemental Table S1 in the article by Zecha *et al.*). Hikmet *et al.* (46) and Aguiar *et al.*

(61) carried out immunohistochemical analysis of tissue and protein profiling by Western blot and from two publicly available MS datasets (one in common to both articles, Kim *et al.*, 2014 (35)). Both articles extensively analyzed tissue and cell distribution finding a similar pattern of ACE2 expression (high abundance in kidney and testis, lower abundance in gallbladder, and so on; Figs. 4–6 in the article by Aguiar *et al.* (61), Fig. 6 in the article by Hikmet *et al.* (46)). Aguiar *et al.* (61) also profiled TMPRSS2, BSG (CD147), and HSPA5 in detail (Figs. 4–6 in the article by Aguiar *et al.*). Wang *et al.* (47) carried out a similar analysis entirely using previously published data and generating similar results (Fig. 2 in the article by Wang *et al.*). Finally, Stanley *et al.* looked specifically at reproductive cells using data from the Human Protein Atlas (36) and the Human Proteome Map (35) and concluded that reproductive consequences of SARS-CoV-2 infection are low given a lack of detectable coexpression of ACE2 and TMPRSS2 at the protein level (49). One of the most interesting results from these and prior studies, considering the primary respiratory route of viral transmission, has been the difficulty of detecting ACE2 and TMPRSS2 in upper airway samples (61) and other lung tissue samples (46, 47, 61). Researchers have typically explained this by suggesting that alternative entry factors may exist, noting that ACE2 expression appears to be restricted to a subset of (generally) epithelial cells, and by noting that interferon can upregulate ACE2 expression once infection is established (60–62). Among other tissues notable in pathogenesis and symptom presentation, ACE2 and TMPRSS2 were codetected in multiple intestinal samples (46, 47, 61),

and intestine has long been known to be particularly enriched in ACE2 expression (60).

Studies have also examined changes in host cell entry factor abundance during infection. Understanding the role of ACE2 modulation during infection is clinically significant because ACE2 is both the viral receptor as well as a lung protective factor (notably in SARS-CoV-1 infection) (18, 62–68). However, characterizing this modulation has proven challenging as ACE2 (cell surface) levels are variously upregulated or downregulated by different factors in infection including interferon signaling and proteases (e.g., ADAM17) (69, 70). In SARS-CoV-2 infection, Bojkova *et al.* (and Bock *et al.* (71) based on the same dataset) have now reported reduction of ACE2 abundance in the Caco-2 cell line (a human colon carcinoma line) (57). In contrast to these data, however, Zecha *et al.* (41) did not find ACE2 decrease in their cell line models (ACE2-A549, Vero E6, Calu-3, and notably also Caco-2), although they found a decrease in abundance of cathepsin L over time. Continued work will further refine our understanding of host cell entry factor changes during the course of infection.

### PTMs

Coronavirus proteins, receptors, and other relevant host cell proteins are often post-translationally modified (72, 73). Review of the SARS-CoV-2 proteome literature revealed studies of glycosylation, phosphorylation, and at least one report detailing (lysine, arginine, and glutamic acid) methylation and proline oxidation on SARS-CoV-2 S and human ACE2 produced in insect cells (74). The methylation results are somewhat surprising since most methyltransferases are localized to the nucleus although there are reports of aspartic and glutamic acid methylation in the secretory pathway (75). No studies have yet confirmed other modifications typically observed with coronaviruses such as ADP ribosylation, sumoylation, palmitoylation, or ubiquitination although sites have been predicted by bioinformatics (73, 76, 77), and relevant protein interactions for such modifications have been demonstrated (56, 71, 78).

**Glycoproteomics and Glycomics**—Several SARS-CoV-2 proteins (S, M, E, and certain “orf” proteins such as Orf8 (79, 80)—a viroporin—and likely Orf7 (81), as with related coronaviruses), as well as host cell factors important in infection, transit the secretory pathway during expression and thus may be glycosylated by secretory pathway glycosyltransferases. The “S” protein (also called “spike” or “surface glycoprotein”) assembles as homotrimers and coats SARS-CoV-2 virions (16, 17). The primary binding partner of spike for host cell entry, ACE2, is also a glycoprotein (15, 18, 82). Representing natural targets for both antibodies and inhibitors, there has been substantial interest in both these proteins and their protein-linked carbohydrate moieties that may shield or otherwise alter PPIs and protein accessibility. In reviewing glycoproteomic and glycomic studies (particularly preprints), it is important to note that recombinant protein

design may lead to non-native modifications (e.g., reports of secretory pathway glycosylation of N).

To date, there have been six glycoproteomics and glycomics studies published on the carbohydrates covalently attached to SARS-CoV-2 spike, two of which also characterized the glycosylation of ACE2 (74, 83–87). All results published to date confirm that the spike protein is predominantly modified by N-glycans (at 22 sites) and that there may be varying amounts of O-glycans present at one site (T323). Distinguishing parameters of primary importance among these results are the cell model, recombinant protein design, and purification strategy used. Two of the studies utilized experimental designs (human cell line expression and trimer purification) that have previously been widely shown with viruses in general to produce proteins very close in character to those derived from actual viral infections (83, 86). These studies, by Watanabe *et al.* (83) and Zhao *et al.* (86), were in substantial agreement regarding the identities of glycans present and the occupancy of each glycosite and in addition demonstrate the importance of multiple protease digestion and the use of different types of fragmentation activation for comprehensive glycan and glycosite characterization. Zhao *et al.* (86) were also one of only two studies to date in which glycomics was carried out to refine the topologies of the glycans present (87). Other studies cited previously either utilized proteins not produced in human cell lines (74, 85) or protein other than full-length trimer purified spike (84, 87), raising additional questions as to the biological relevance of the glycosylation results obtained with respect to actual SARS-CoV-2 virions in human hosts. The apparent resulting differences provide valuable information to researchers considering antibody or vaccine candidate production in nonhuman cell lines or using non-trimer purified protein. Finally, it is worth noting that SARS-CoV-2 glycosylation is significantly more host like than the glycosylation found on many other viruses such as HIV when considering N-glycan processing and density, although high mannose glycans still occur with greater prevalence than on most host proteins (83, 86, 88).

The two recent studies characterizing the glycosylation of ACE2 form a subset of the SARS-CoV-2 spike articles (74, 86) and provide a much more complete picture than earlier articles (89–92). In particular, Zhao *et al.* (86) carried out comprehensive glycomics-informed glycoproteomic analysis on a purified soluble version of the protein. ACE2 has seven “canonical” N-glycosylation sequons, six of which were included in the expression construct used in the article and analyzed in depth. The N-glycosylation of ACE2 was found to be broadly similar to that seen with other human proteins that traffic through the secretory pathway. Only a small amount of O-glycosylation was detected. This detailed analysis of glycosylation also allowed the authors to carry out molecular dynamics simulations of SARS-CoV-2 S glycoprotein bound to ACE2 glycoprotein suggesting that several N-linked glycans on each protein are intimately involved in glycoprotein–glycoprotein

interaction (Fig. 2). In addition, a model of the postfusion SARS-CoV-2 S glycoprotein was developed showing its distinctive columnar structure and even N-glycan spacing (Fig. 2). A number of other articles containing molecular dynamics simulation results and molecular modeling work have been published based on different glycoproteomics studies and are valuable additional references (93–100). Sun *et al.* (74) reported very similar results on the N-glycosylation of ACE2. Although they were unable to detect N-glycans at N053 and N322, by using a construct including the N-glycosylation site N690, Sun *et al.* (74) were able to characterize glycosylation at this seventh site in contrast with Zhao *et al.* (86). In total, the glycosylation patterns of the SARS-CoV-2 spike protein and its host cell surface receptor have been characterized in detail by multiple groups using different biological models providing important information for future research and particularly informing modeling that may be crucial in understanding and addressing the emergence of potential vaccine and antibody escape variants.

**Phosphorylation**—Currently, three published studies have specifically examined the phosphorylation of SARS-CoV-2 proteins (Fig. 1) and host cell entry factors (14, 101, 102). Davidson *et al.* (14) used TiO<sub>2</sub> followed by ferric nitrilotriacetate phosphopeptide enrichment on in-gel trypsin-digested peptides from whole cell lysate enabling the mapping of 44 phosphorylation sites among five viral proteins (Figs. 3–5 and Table 4 in the article by Davidson *et al.*: nsp3, nsp9, M, N, and S—nsp12 and ORF3a also produced phosphorylated peptides, but sites could not be confidently assigned) in a typical DDA LC–MS/MS higher-energy collisional dissociation experiment (14). The report of phosphorylation sites distributed across spike is surprising, as noted by Davidson *et al.* in their article, given its secretory pathway expression, and it will be interesting to see if future studies confirm this result. Bouhaddou *et al.* (101) comprehensively examined phosphorylation of both viral proteins and host proteins, enriching from cell lysate tryptic peptides *via* iron affinity and analyzing *via* LC–MS/MS (DDA and data-independent acquisition) with higher-energy collisional dissociation. Using this approach, they identified 25 phosphorylation sites (omitting results from Davidson *et al.* (14) that they also presented) on SARS-CoV-2 viral proteins (Fig. 2 and supplemental Table S2 in the article by Bouhaddou *et al.* (101): on nsp3, nsp14, orf9b, M, and N). The phosphorylation of N has since been shown to be functionally important in nucleocapsid assembly and viral replication and transcription (103–105) as was previously observed with other coronaviruses. Several sites in the C-terminal tail of M are noted to be present in other viruses suggesting potential functional importance although it does not appear that this has been functionally verified for SARS-CoV-2 yet. Among host cell entry factors, PIKfyve and cathepsin L are phosphorylated (Table S1 in the article by Bouhaddou *et al.* (101)), and regulation of the phosphorylation of PIKfyve was further confirmed to be involved in SARS-CoV-2 infection (101). Klann

*et al.* (102) used ferric nitrilotriacetate enrichment after in-solution digest and TMT labeling of proteins from whole cell lysate finding 33 sites on six viral proteins (Fig. 1, E–J in the study by Klann *et al.* (102): pp1ab, nsp6, ORF3a, ORF9b, M, and N). Additional datasets available in the ProteomeXchange (31) contain various sets of relevant phosphorylation hits and may be found by searching this database.

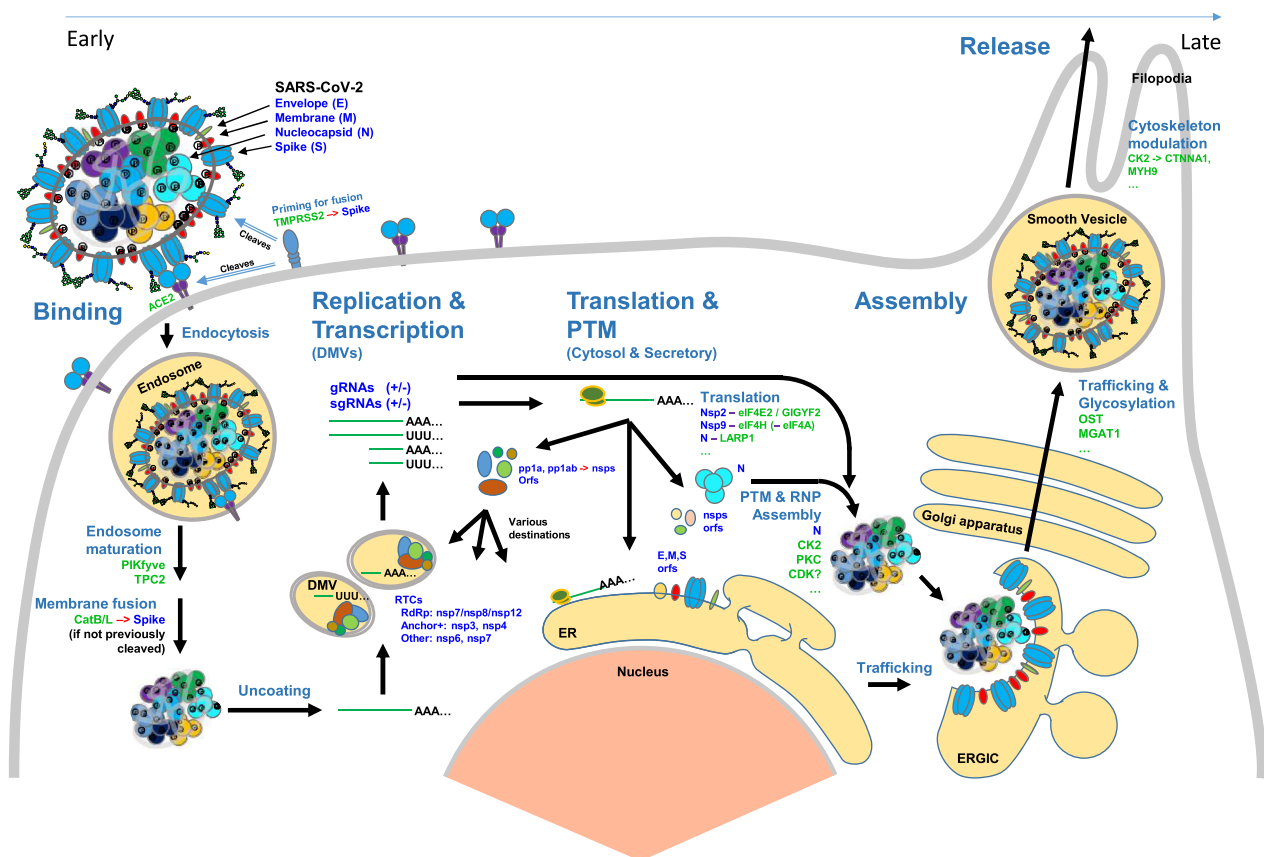
## HOST CELL PROTEOMICS

### PPI

PPIs play a primary role in the life cycle of animal viruses, from attachment to cells through endosomal compartment escape (for most viruses) and ultimately reorganization of cellular machinery to support viral reproduction and diminish host defense (106, 107). Consequently, enumerating host–virus PPIs is crucial to understanding the biology of viruses and developing a starting point for investigation of potential therapeutics. To date, several comprehensive studies of SARS-CoV-2 host–virus PPIs using proteomics technologies have been published, setting the stage for additional studies expanding on this work—studies that may definitively validate or invalidate proposed interactions. An extremely useful resource for SARS-CoV-2 PPIs was brought to our attention during review of this article, and interested readers are directed to it for a more comprehensive up-to-date view of published interactions: <https://thebiogrid.org/project/3> (30). The interactions reported in this database and in the research cited later may also be considered in the context of more recently published functional genomics studies for additional perspective (25, 26, 108–110).

A comprehensive interactome was published by Gordon *et al.* who carried out affinity purification, using transfected viral proteins (human embryonic kidney 293 cells) as baits against human cell proteins in the cells used for viral protein expression, followed by tryptic digestion and LC–MS/MS (56). Through this workflow, they were able to map 332 high-confident PPIs and identify 66 druggable human proteins concluding, through additional experiments with compounds targeting interactors, that inhibitors of mRNA translation and regulators of sigma-1 and sigma-2 receptors show potential for SARS-CoV-2 treatment. Noteworthy PPIs involve factors in host mRNA nuclear export and overall mRNA regulation/translation, phosphorylation, secretory pathway targeting of proteins, and protein degradation (Fig. 3, Table 2, and supplemental Table S1). This group has since published a newer study comparing interactors across related coronaviruses SARS-CoV-1 and Middle East respiratory syndrome-CoV and validating or further validating the clinical relevance of three host factor interactors reported in their original study (Tom70 or TOMM70, ILR17RA, and SigmaR1) (111). Li *et al.* (78) more recently published another comprehensive interactome using similar affinity-purification LC–MS/MS methods and found 45 targets shared with the article by Gordon *et al.*





**FIG. 3. The SARS-CoV-2 viral life cycle and selected host proteins involved.** The viral life cycle is displayed proceeding from host cell entry through new virion synthesis, packaging, and export. Host cell proteins are labeled in *green*, and SARS-CoV-2 proteins are labeled in *blue*. *Red arrows* (→) indicate protease cleavage. The representation of virus shows ribonucleoproteins (RNPs) (consisting of five dimers of N) in the tetrahedral geometry recently reported (Yao *et al.* (99)). This article reported an average of  $26 \pm 15$  copies of prefusion S per virion and  $26 \pm 15$  RNP per virion. The life cycle in a given cell begins with host cell entry mediated by ACE2 (the receptor), TMPRSS2 (or alternatively CatB/L—CSTB/CTSL—fusion priming enzymes), and proceeds with trafficking through endosomes. Endosomal maturation required for viral–host–cell membrane fusion involves the proteins PIKfyve and TPC2. After fusion and uncoating of the viral RNA, the replication–transcription complex is expressed, and new viral genomic RNAs (gRNAs, + and – sense) and subgenomic RNAs (sgRNAs, + and – sense) are produced. The translation of viral proteins and modulation of host protein translation is affected by protein–protein interactions (Nsp2–eIF4E2/GIGYF2, Nsp9–eIF4H, and N–LARP1 are shown) and signaling. New virion structural protein N is phosphorylated (CK2, PKC, and CDK), forms RNPs, winds gRNAs, and collects at the ERGIC membrane for envelopment. Viral proteins E, M, and S traffic through the secretory pathway for further processing including addition of glycans. Filopodia formation is enhanced (proposed to be CK2 driven by Bouhaddou *et al.* (101)) and may improve transmission of egressing virus between cells. ACE2, angiotensin-converting enzyme 2; CTSL, cathepsin L; ERGIC, endoplasmic reticulum golgi intermediate compartment; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; TMPRSS2, transmembrane serine protease 2.

(56) (~16% of their interactome, including, *e.g.*, SigmaR1) while uncovering many unreported PPIs. In particular, Li *et al.* found many immune system–related interactions (78) (Table 2 and supplemental Table S1).

Two studies have used interactome data previously published on other coronaviruses to predict the SARS-CoV-2 interactome, in one case thoroughly integrating this analysis with the data by Gordon *et al.* (56). Perrin-Cocon *et al.* (112) integrated the Gordon *et al.* (56) interactome with coronavirus–host interactions from literature on 12 other coronaviruses. Many of the interactions found were already widely reported although several previously relatively overlooked (for SARS-CoV-2) interactions were noted including with protein translation repressors, mitochondrial homeostasis regulators, and

an S-adenosylmethionine synthase crucial to the DNA methylation pathway (Table 2 and supplemental Table S1) (112). Messina *et al.* (113) carried out a more restricted analysis based only on the spike protein and without integrating the findings of Gordon *et al.* (56), highlighting several noteworthy pathways (113) (Table 2 and supplemental Table S1). It is important to note that functional genomic screens that are now available have not directly confirmed the biological importance of many of these interactions, in some cases, potentially because of their design and limitations.

Several studies have integrated results from Gordon *et al.* (56) with other types of host proteome data. Feng *et al.* (48) integrated the interactome with proteomics datasets from human lung, colon, kidney, liver, and heart, to piece together a

TABLE 2  
Selected host proteins in infection

Primary process	Gene/complex/family	Protein name	PPI?	Abundance?	Phosphorylation?	Act.?	Function in infection (known and/or hypothesized)	Cell location	Selected SARS-CoV-2 proteomics references
Host cell entry	ACE2	Angiotensin-converting enzyme 2	S	+/-			Virus receptor	PM	(41, 46–49, 57, 61, 74, 86)
	TMPRSS2	Transmembrane protease serine 2	S				Cleaves S (“priming”), especially at S2' site	PM	(41, 48, 49, 61)
	CTSB	Cathepsin B	S?				Cleaves S, alternative to TMPRSS2	EN	(41, 46, 78, 112)
	CTSL	Cathepsin L	S				Cleaves S, alternative to TMPRSS2	EN	(41, 46, 49, 61, 112)
Endosomal release	PIKFYVE	1-phosphatidylinositol 3-phosphate 5-kinase				Endosome maturation, with TPC2	EN	(101)	
Protein expression	NUP98	Nuclear pore complex protein Nup98	Orf6		+		Prevent host nuclear mRNA export	NM	(56, 101)
	LARP1	La-related protein 1	N		-		Prioritize virus protein expression	CP, NU	(56, 101, 112)
	UPF1	Regulator of nonsense transcripts 1	N				Binding by N represses NMD?	CP, NU	(56, 112)
	EIF4H	Eukaryotic translation initiation factor 4H	Nsp9				Cap-dependent mRNA translation	PN	(56)
	EIF4E2	Eukaryotic translation initiation factor 4E type 2	Nsp2				Represses cap-dependent translation	CP	(56, 112)
	Sec61 complex	SEC61 channel-forming translocon complex	Nsp8				Protein entry into endoplasmic reticulum	CP, ERM	(56, 57)
	BRD4	Bromodomain-containing protein 4	E	+			Interference with antiviral response?	NU	(48, 56)
	FURIN	Furin	S				Cleaves S (“priming”), especially at S1/S2 site	Golgi	
Protein processing	CUL2	Cullin 2	Orf10	+			Increase degradation of restriction factors?	CP, NU	(56, 59)
Protein degradation	CDK	Cyclin-dependent kinase				-	Cell cycle arrest, S/G2	NU, MT, CP	(101)
Cell signaling	MAPK	Mitogen-activated protein kinase		+		+	Viral replication+, stress response	NU, CP, MT	(59, 101, 102, 122)

TABLE 2—Continued

Primary process	Gene/complex/family	Protein name	PPI?	Abundance?	Phosphorylation?	Act.?	Function in infection (known and/or hypothesized)	Cell location	Selected SARS-CoV-2 proteomics references
	AKT	RAC-alpha serine/threonine-protein kinase		+	+	-/+	Viral replication+, cell proliferation & apoptosis regulation	NU, CP (PM)	(59, 101)
Cell structure	PHB complex	Prohibitin complex	nsp2				Signaling interference, mitochondrial antiviral signaling, apoptosis–	MT, NU, CP, PM	(56, 112, 129)
	CK2 complex	Casein kinase II	N			+	Cytoskeleton changes, filopodia+	CP, NU	(56, 101, 102)
Stress, immunity	HSPA5	Endoplasmic reticulum chaperone BiP					Unfolded protein response, virus receptor?	CP, PM	(61)
	NKRF	NF-kappaB-repressing factor	(nsp10)				IL-8 induction	NO, NU, CP	(78)
	CFB	Complement factor B		-/+		+	Alternative complement pathway factor	Secreted	(120–122)
	CFD	Complement factor D					Activates complement-dependent killing	Secreted	(120, 123)
	CFI	Complement factor I		+		+	Prevented from modulating complement	Secreted	(120, 123)
	CFH	Complement factor H		+/-		+	Prevented from modulating complement	Secreted	(120, 122, 123)

Abbreviations: Cell location—CP, cytoplasm; EN, endosome; ER, endoplasmic reticulum; ERM, ER membrane; LY, lysosome; MT, mitochondria; NM, nuclear membrane; NO, nucleolus; NU, nucleus; PM, plasma membrane; PN, perinuclear; other—HS, heparan sulfate; MAVS, mitochondrial antiviral signaling; NMD, nonsense mediated decay.

See [supplemental Table S1](#) for more information on these proteins, complexes, and families. NKFR = nsp10 PPI may not be direct.

more tissue-centric view of the results from Gordon *et al.* (56). They also extended this analysis with differential quantitative “infectome” data from Bojkova *et al.* (57) (discussed in detail in the next section) carrying out further PPI network analysis and identifying critical “hubs” and hub proteins, grouped by tissue, and involved in mRNA processing (HNRNPC, SRSF1, and HNRNPA3), gene expression (BRD4), the tumor necrosis factor signaling cascade (RIPK1), and the olfactory signaling pathway (REEP5) among others (48). Zecha *et al.* (41) integrated the results from Gordon *et al.* (56) with their own quantitative data on the host proteomes of four relevant cell lines (human ACE2-A549, Caco-2, Calu-3, and African green monkey Vero E6) finding 56 interacting host proteins differing quantitatively during infection of Vero E6 cells (Zecha *et al.*; Figure 4C, supplemental Fig. S8D) (41). Appelberg *et al.* (59) also analyzed interactions in Gordon *et al.* (56) in terms of quantitative transcriptomic and proteomic data (TMT based) that they generated (Huh7 cells) (59). Finally, Bouhaddou *et al.* (101) integrated information from the Gordon *et al.* (56) interactome with quantitative proteomics and phosphoproteomics as discussed in more detail in next section.

#### Host Cell and Tissue Proteome Changes in Infection

Related to direct PPIs is the set of proteins regulated or otherwise modulated during infection. This set includes many proteins involved in cell survival and the cell cycle, signaling, host defense (often directly inhibited by pathogenic viruses), cell trafficking, and other critical cellular and host processes (114, 115). In contrast with studies of PPIs using techniques like affinity purification, differential quantitative proteomics is subject to several additional confounding factors. First—viral protein expression increases in particularly permissive host cells dramatically over the course of even relatively short time intervals (8 h–2 days), which can negatively impact sensitivity and dynamic range in addition to making accurate comparative quantification challenging (58, 101). Second—in cell culture, different populations of cells will exist concurrently at any given time point (uninfected cells through to apoptotic, necrotic, and fully lysed cells). Third—in patient samples, quantified proteins must be compared against some standard usually chosen to be “uninfected controls.” This raises questions about how closely the proteomes of small sets of people (in the studies carried out to date) match under ordinary conditions. In this section, we will discuss articles on differential quantitative proteomics of host cell lines in SARS-CoV-2 infection and briefly mention extensions of those studies. We will follow this with discussions of differential quantitative proteomics (and potential biomarkers) from patient samples.

Beginning with the publication of Bojkova *et al.* (57), initially available as a preprint in March 2020, seven articles have reported host primary cell or host cell line quantitative differential proteomics in the context of SARS-CoV-2 infection. Six of these articles carried out extensive differential expression,

network, pathway, and gene ontology analysis (57–59, 78, 101, 102) (Bouhaddou *et al.* (101) and Klann *et al.* (102) primarily in terms of the phosphoproteome), whereas one examined differential expression in a more cursory fashion (41). Studies differed in cells or cell line models used (patient peripheral blood mononuclear cells (78), Caco-2 from human colon carcinoma (57, 102), Vero E6 from African green monkey kidney epithelial cells (58, 101), and Huh7 from human hepatocyte carcinoma (59)), time points examined, and analytical techniques and technologies. Three groups utilized TMT labeling for quantification (41, 59, 78), one group used multiplexed enhanced protein dynamics proteomics based on pulsed stable isotope labeling by amino acids in cell culture and TMT (57) for their first article and TMT alone for their follow-up article (102), and the other two groups used forms of label-free quantification (58, 101). Pathways identified in common among various articles include spliceosomal (57, 58, 101, 102), hypoxia-inducible factor 1 signaling (41, 57, 59), innate immune system (41, 59, 78, 101), and pathways involved in carbon metabolism (57–59, 78, 102). Many of the host proteins shown to directly interact with viral proteins do not appear centrally in these results presumably since interaction is distinct from regulation. However, for purposes of comparison, HMOX1 (heme degradation) (41, 56, 59, 78), RIPK1 (NF- $\kappa$ B, tumor necrosis factor signaling) (56, 59), CUL2 and RBX1 (ubiquitination) (57, 59), and La-related protein 1 (an mechanistic target of rapamycin–regulated translation repressor) (56, 59, 101) are some proteins shared between PPI studies and these studies (Fig. 3, Table 2, and supplemental Table S1). Also of interest, certain pathways such as complement and innate immune activation were detected as differentially regulated only at low multiplicity of infections and/or in certain cell lines (41, 59).

Two of the differential quantitative proteomics articles examined the phosphoproteome in depth and found significant phosphoproteome changes consistent with the central role of phosphorylation in cell signal transduction. Bouhaddou *et al.* (101) compared the proteome and phosphoproteome of infected cells at six time points (0, 2, 4, 8, 12, or 24 h) and noninfected cells at two time points (0 or 24 h). This analysis revealed changes in phosphorylation of proteins involved in RNA processing (including La-related protein 1 and ribosomal RNA processing 1), nuclear export (NUP98; Table 2, supplemental Table S1), cytoskeleton organization and filopodia formation (VIM, STMN1, CTNNA1, MYH9; Fig. 3, Table 2, and supplemental Table S1—see casein kinase II [CK2]), the p38/mitogen-activated protein kinase pathway (NELFE, HSPB1, and STAT1), and cell cycle arrest (cyclin-dependent kinase 2; Fig. 3, Table 2, and supplemental Table S1). Based on the amino acid sequences of mapped sites, kinases and kinase families involved were also predicted (CK2, cyclin-dependent kinase, PKC, mitogen-activated protein kinase 12, CAMK2G, and AKT1/2 among others; Fig. 3,

Table 2, and supplemental Table S1). In addition, Bouhaddou *et al.* were able to classify phosphoproteome changes into sets based on their appearance throughout the viral life cycle (early, replication, and egress). Finally, Bouhaddou *et al.* (101) also analyzed their results in the context of the Gordon *et al.* interactome (56) generating additional hypotheses regarding the mechanisms of a subset of the regulation they observed (e.g., allosteric regulation of CK2 by N). Klann *et al.* (102) carried out similar analyses finding significant regulation of proteins involved in numerous pathways, including receptor signaling, endocytosis, the cell cycle, translation initiation, and splicing. Although there is some overlap between proteins revealed by these two studies, there are clear differences in the set of proteins and pathways reported that will not be covered further in this review.

Several additional articles consisting of reanalyses or extensions of data from the articles discussed previously have also been published. Feng *et al.* (48), as discussed earlier, reanalyzed and extended the data from Bojkova *et al.* (57) and Gordon *et al.* (56) to the tissue level. Bock *et al.* (71) reanalyzed the data from Bojkova *et al.* (57) using different methods highlighting, for example, the complement cascade in contrast with the original analysis. Many additional articles utilizing these differential host proteomics datasets and results have also been published (10, 102, 116, 117).

### *Host Biofluid Proteome Changes, Biomarkers, and Immune Response*

A number of extensive reviews have recently been published containing information on host biofluid proteome changes, potential biomarkers, and immune system response. Therefore, we will not cover the primary literature in these areas in detail here but will briefly summarize several reviews pointing the reader to them for more extensive discussion.

One recent review focused substantially on proteomics was published by Whetton *et al.* Whetton *et al.* (118) discussed the complexity of COVID-19, covered relevant proteomics techniques, compiled a table of selected publications detailing proteome changes in disease, presented network analyses of coronavirus- and SARS-related PubMed abstracts (in terms of cytokines, cells, and connected diseases), and concluded that future pandemic response would greatly benefit from more integrated OMICS and informatics pipelines. Ultimately, in terms of characterized proteome changes, they highlight the widely reported cytokine storm profile present in patients. They also point out markers associated with disease severity, including IL-6, troponin I, and B-type natriuretic peptide. Of particular interest, in light of a recent report on the importance of the alternative pathway of the complement system in COVID-19 disease pathogenesis (119), Whetton *et al.* (118) called attention to the appearance of alternative pathway-relevant proteins in a couple of proteomics articles (120, 121). Additional proteomics articles examining complement

system proteins in SARS-CoV-2 infection have since been published (122, 123) (see also Table 2 and supplemental Table S1). Another recent review of note but less focused on proteomics in isolation (among others (124, 125)), focusing rather on immune response and immunological pathways is by Vabret *et al.* and expands on the immunology and biochemistry of much of this information (126). This review also includes information on current clinical trials and therapeutic options. A number of other articles with a proteomic character, including many previously cited in our review (several only in tables), also point to various immune pathways and biofluid proteins and may be consulted for further information (41, 57, 59, 78, 101, 102, 113, 120–123, 127–129).

### CONCLUSION

In less than a year, there are now more than 100 directly proteomic-relevant articles already published in peer-reviewed journals. In this short period, scientific and medical research progress on the COVID-19 pandemic virus, SARS-CoV-2, has been prodigious, also representing a testament to the (funding of) decades worth of basic research required for such rapid progress when confronting new challenges. While the constraints of this review have prevented a comprehensive presentation of all relevant literature (and led to the exclusion of several topics such as the use of native MS in pharmacological studies), each subsection will hopefully facilitate the readers ability to further investigate specific SARS-CoV-2-related topics of interest. From basic proteomics and potential diagnostics through to a deeper understanding of host-wide changes that occur during infection that could inform potential therapeutic strategies, it is clear that the contribution of the proteomics community to addressing the COVID-19 pandemic is already substantial. Future research will continue to build on existing findings and is highly likely to generate novel insight into the COVID-19 pandemic.

*Supplemental data*—This article contains [supplemental data](#) (130–137).

*Acknowledgments*—We thank Oliver C. Grant and Rob Woods for the images in [Figure 2](#) and Professor Bing Chen for the underlying models used.

*Author contributions*—J. L. P. drafting and revising the article; L. W. revising and approval of the final version.

*Funding and additional information*—This work was supported in part by a grant from the National Institutes of Health/National Institute of General Medical Sciences (R01GM130915 to L. W.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**Conflict of interest**—The authors declare no competing interests.

**Abbreviations**—The abbreviations used are: ACE2, angiotensin-converting enzyme 2; CK2, casein kinase II; COVID-19, coronavirus disease 2019; DDA, data-dependent acquisition; nsp, nonstructural protein; PIKFyve, a FYVE finger-containing phosphoinositide kinase; PPI, protein-protein interaction; PRM, parallel reaction monitoring; PTM, post-translational modification; qPCR, quantitative PCR; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; TMPRSS2, transmembrane serine protease 2; TMT, tandem mass tag.

Received May 4, 2021 Published, MCPRO Papers in Press, June 4, 2021, <https://doi.org/10.1016/j.mcpro.2021.100103>

## REFERENCES

- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., et al. (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497–506
- Wu, F., Zhao, S., Yu, B., Chen, Y. M., Wang, W., Song, Z. G., Hu, Y., Tao, Z. W., Tian, J. H., Pei, Y. Y., Yuan, M. L., Zhang, Y. L., Dai, F. H., Liu, Y., Wang, Q. M., et al. (2020) A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269
- Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., et al. (2020) A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733
- Chan, J. F., Kok, K. H., Zhu, Z., Chu, H., To, K. K., Yuan, S., and Yuen, K. Y. (2020) Genomic characterisation of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerg. Microbes Infect.* **9**, 221–236
- Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P., Meng, J., Zhu, Z., Zhang, Z., Wang, J., Sheng, J., Quan, L., Xia, Z., Tan, W., Cheng, G., et al. (2020) Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host Microbe* **27**, 325–328
- Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N., Bi, Y., Ma, X., Zhan, F., Wang, L., Hu, T., et al. (2020) Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding. *Lancet* **395**, 565–574
- Wu, C., Liu, Y., Yang, Y., Zhang, P., Zhong, W., Wang, Y., Wang, Q., Xu, Y., Li, M., Li, X., Zheng, M., Chen, L., and Li, H. (2020) Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharm. Sin. B* **10**, 766–788
- Yoshimoto, F. K. (2020) The proteins of severe acute respiratory syndrome coronavirus-2 (SARS CoV-2 or n-COV19), the cause of COVID-19. *Protein J.* **39**, 198–216
- Holmes, E. C. (2003) Error thresholds and the constraints to RNA virus evolution. *Trends Microbiol.* **11**, 543–546
- Finkel, Y., Mizrahi, O., Nachshon, A., Weingarten-Gabbay, S., Morgenstern, D., Yahalom-Ronen, Y., Tamir, H., Achdout, H., Stein, D., Israeli, O., Beth-Din, A., Melamed, S., Weiss, S., Israely, T., Paran, N., et al. (2021) The coding capacity of SARS-CoV-2. *Nature* **589**, 125–130
- Kim, D., Lee, J. Y., Yang, J. S., Kim, J. W., Kim, V. N., and Chang, H. (2020) The architecture of SARS-CoV-2 transcriptome. *Cell* **181**, 914–921.e10
- Nomburg, J., Meyerson, M., and DeCaprio, J. A. (2020) Pervasive generation of non-canonical subgenomic RNAs by SARS-CoV-2. *Genome Med.* **12**, 108
- Pancer, K., Milewska, A., Owczarek, K., Dabrowska, A., Kowalski, M., Łabaj, P. P., Branicki, W., Sanak, M., and Pyrc, K. (2020) The SARS-CoV-2 ORF10 is not essential in vitro or in vivo in humans. *PLoS Pathog.* **16**, e1008959
- Davidson, A. D., Williamson, M. K., Lewis, S., Shoemark, D., Carroll, M. W., Heesom, K. J., Zambon, M., Ellis, J., Lewis, P. A., Hiscox, J. A., and Matthews, D. A. (2020) Characterisation of the transcriptome and proteome of SARS-CoV-2 reveals a cell passage induced in-frame deletion of the furin-like cleavage site from the spike glycoprotein. *Genome Med.* **12**, 68
- Zhou, P., Yang, X. L., Wang, X. G., Hu, B., Zhang, L., Zhang, W., Si, H. R., Zhu, Y., Li, B., Huang, C. L., Chen, H. D., Chen, J., Luo, Y., Guo, H., Jiang, R. D., et al. (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273
- Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C. L., Abiona, O., Graham, B. S., and McLellan, J. S. (2020) Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263
- Walls, A. C., Park, Y. J., Tortorici, M. A., Wall, A., McGuire, A. T., and Velesler, D. (2020) Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **181**, 281–292.e6
- Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N. H., Nitsche, A., Müller, M. A., Drosten, C., and Pöhlmann, S. (2020) SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280.e8
- Heurich, A., Hofmann-Winkler, H., Gierer, S., Liepold, T., Jahn, O., and Pöhlmann, S. (2014) TMPRSS2 and ADAM17 cleave ACE2 differentially and only proteolysis by TMPRSS2 augments entry driven by the severe acute respiratory syndrome coronavirus spike protein. *J. Virol.* **88**, 1293–1307
- Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., and Zhou, Q. (2020) Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **367**, 1444–1448
- Donoghue, M., Hsieh, F., Baronas, E., Godbout, K., Gosselin, M., Stagliano, N., Donovan, M., Woolf, B., Robison, K., Jeyaseelan, R., Breitbart, R. E., and Acton, S. (2000) A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. *Circ. Res.* **87**, E1–E9
- Ou, X., Liu, Y., Lei, X., Li, P., Mi, D., Ren, L., Guo, L., Guo, R., Chen, T., Hu, J., Xiang, Z., Mu, Z., Chen, X., Chen, J., Hu, K., et al. (2020) Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat. Commun.* **11**, 1620
- Zang, R., Gomez Castro, M. F., McCune, B. T., Zeng, Q., Rothlauf, P. W., Sonnek, N. M., Liu, Z., Brulois, K. F., Wang, X., Greenberg, H. B., Diamond, M. S., Giorba, M. A., Whelan, S. P. J., and Ding, S. (2020) TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Sci. Immunol.* **5**, 47
- Matsuyama, S., Nao, N., Shirato, K., Kawase, M., Saito, S., Takayama, I., Nagata, N., Sekizuka, T., Katoh, H., Kato, F., Sakata, M., Tahara, M., Kutsuna, S., Ohmagari, N., Kuroda, M., et al. (2020) Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 7001–7003
- Danilowski, Z., Jordan, T. X., Wessels, H. H., Hoagland, D. A., Kasela, S., Legut, M., Maniatis, S., Mimitou, E. P., Lu, L., Geller, E., Danziger, O., Rosenberg, B. R., Phatnani, H., Smibert, P., Lappalainen, T., et al. (2021) Identification of required host factors for SARS-CoV-2 infection in human cells. *Cell* **184**, 92–105.e16
- Wei, J., Alfajaro, M. M., DeWeirdt, P. C., Hanna, R. E., Lu-Culligan, W. J., Cai, W. L., Strine, M. S., Zhang, S. M., Graziano, V. R., Schmitz, C. O., Chen, J. S., Mankowski, M. C., Filler, R. B., Ravindra, N. G., Gasque, V., et al. (2021) Genome-wide CRISPR screens reveal host factors critical for SARS-CoV-2 infection. *Cell* **184**, 76–91.e13
- Shu, Y., and McCauley, J. (2017) GISAID: Global initiative on sharing all influenza data - from vision to reality. *Euro Surveill.* **22**, 30494
- Coordinators, N. R. (2018) Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **46**, D8–D13
- Cantelli, G., Cochrane, G., Brooksbank, C., McDonagh, E., Flicek, P., McEntyre, J., Birney, E., and Apweiler, R. (2021) The European bioinformatics institute: Empowering cooperation in response to a global health crisis. *Nucleic Acids Res.* **49**, D29–D37
- Oughtred, R., Rust, J., Chang, C., Breitkreutz, B. J., Stark, C., Willems, A., Boucher, L., Leung, G., Kolas, N., Zhang, F., Dolma, S., Coulombe-Huntington, J., Chatr-Aryamontri, A., Dolinski, K., and Tyers, M. (2021) The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci.* **30**, 187–200
- Deutsch, E. W., Csordas, A., Sun, Z., Jarnuczak, A., Perez-Riverol, Y., Ternent, T., Campbell, D. S., Bernal-Llinares, M., Okuda, S., Kawano, S.,

- Moritz, R. L., Carver, J. J., Wang, M., Ishihama, Y., Bandeira, N., *et al.* (2017) The ProteomeXchange consortium in 2017: Supporting the cultural change in proteomics public data deposition. *Nucleic Acids Res.* **45**, D1100–D1106
32. Samaras, P., Schmidt, T., Frejno, M., Gessulat, S., Reinecke, M., Jarzab, A., Zecha, J., Mergner, J., Giansanti, P., Ehrlich, H. C., Aiche, S., Rank, J., Kienegger, H., Krcmar, H., Kuster, B., *et al.* (2020) ProteomicsDB: A multi-omics and multi-organism resource for life science research. *Nucleic Acids Res.* **48**, D1153–D1163
33. Wang, M., Herrmann, C. J., Simonovic, M., Szklarczyk, D., and von Mering, C. (2015) Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. *Proteomics* **15**, 3163–3168
34. Edwards, N. J., Oberti, M., Thangudu, R. R., Cai, S., McGarvey, P. B., Jacob, S., Madhavan, S., and Ketchum, K. A. (2015) The CPTAC data portal: A resource for cancer proteomics research. *J. Proteome Res.* **14**, 2707–2713
35. Kim, M. S., Pinto, S. M., Getnet, D., Nirujogi, R. S., Manda, S. S., Chaerkady, R., Madugundu, A. K., Kelkar, D. S., Isserlin, R., Jain, S., Thomas, J. K., Muthusamy, B., Leal-Rojas, P., Kumar, P., Sahasrabudhe, N. A., *et al.* (2014) A draft map of the human proteome. *Nature* **509**, 575–581
36. Uhlen, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szgyarto, C. A., *et al.* (2015) Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419
37. York, W. S., Mazumder, R., Ranzinger, R., Edwards, N., Kahsay, R., Aoki-Kinoshita, K. F., Campbell, M. P., Cummings, R. D., Feizi, T., Martin, M., Natale, D. A., Packer, N. H., Woods, R. J., Agarwal, G., Arpinar, S., *et al.* (2020) GlyGen: Computational and informatics resources for glycoscience. *Glycobiology* **30**, 72–73
38. Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., and Skrzypek, E. (2015) PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43**, D512–D520
39. Mahmud, I., and Garrett, T. J. (2020) Mass spectrometry techniques in emerging pathogens studies: COVID-19 perspectives. *J. Am. Soc. Mass Spectrom.* **31**, 2013–2024
40. Gouveia, D., Grenga, L., Gaillard, J. C., Gallais, F., Bellanger, L., Pible, O., and Armengaud, J. (2020) Shortlisting SARS-CoV-2 peptides for targeted studies from experimental data-dependent acquisition tandem mass spectrometry data. *Proteomics* **20**, e2000107
41. Zecha, J., Lee, C.-Y., Bayer, F. P., Meng, C., Grass, V., Zerweck, J., Schnatbaum, K., Michler, T., Pichlmair, A., Ludwig, C., and Kuster, B. (2020) Data, reagents, assays and merits of proteomics for SARS-CoV-2 research and testing. *Mol. Cell. Proteomics* **19**, 1503–1522
42. Sharma, V., Eckels, J., Schilling, B., Ludwig, C., Jaffe, J. D., MacCoss, M. J., and MacLean, B. (2018) Panorama public: A public repository for quantitative data sets processed in skyline. *Mol. Cell. Proteomics* **17**, 1239–1244
43. Singh, P., Chakraborty, R., Marwal, R., Radhakrishnan, V. S., Bhaskar, A. K., Vashisht, H., Dhar, M. S., Pradhan, S., Ranjan, G., Imran, M., Raj, A., Sharma, U., Singh, P., Lall, H., Dutta, M., *et al.* (2020) A rapid and sensitive method to detect SARS-CoV-2 virus using targeted-mass spectrometry. *J. Proteins Proteom.* **11**, 159–165
44. Cazares, L. H., Chaerkady, R., Samuel Weng, S. H., Boo, C. C., Cimbri, R., Hsu, H. E., Rajan, S., Dall'Acqua, W., Clarke, L., Ren, K., McTamney, P., Kallewaard-LeLay, N., Ghaedi, M., Ikeda, Y., and Hess, S. (2020) Development of a parallel reaction monitoring mass spectrometry assay for the detection of SARS-CoV-2 spike glycoprotein and nucleoprotein. *Anal. Chem.* **92**, 13813–13821
45. Iles, R. K., Zmuidinaite, R., Iles, J. K., Carnell, G., Sampson, A., and Heeney, J. L. (2020) Development of a clinical MALDI-ToF mass spectrometry assay for SARS-CoV-2: Rational design and multi-disciplinary team work. *Diagnostics (Basel)* **10**, 746
46. Hikmet, F., Méar, L., Edvinsson, Å., Micke, P., Uhlén, M., and Lindskog, C. (2020) The protein expression profile of ACE2 in human tissues. *Mol. Syst. Biol.* **16**, e9610
47. Wang, Y., Wang, Y., Luo, W., Huang, L., Xiao, J., Li, F., Qin, S., Song, X., Wu, Y., Zeng, Q., Jin, F., and Wang, Y. (2020) A comprehensive investigation of the mRNA and protein level of ACE2, the putative receptor of SARS-CoV-2, in human tissues and blood cells. *Int. J. Med. Sci.* **17**, 1522–1531
48. Feng, L., Yin, Y. Y., Liu, C. H., Xu, K. R., Li, Q. R., Wu, J. R., and Zeng, R. (2021) Proteome-wide data analysis reveals tissue-specific network associated with SARS-CoV-2 infection. *J. Mol. Cell Biol.* **12**, 946–957
49. Stanley, K. E., Thomas, E., Leaver, M., and Wells, D. (2020) Coronavirus disease-19 and fertility: Viral host entry protein expression in male and female reproductive tissues. *Fertil. Steril.* **114**, 33–43
50. Gouveia, D., Miotello, G., Gallais, F., Gaillard, J.-C., Debroas, S., Bellanger, L., Lavigne, J.-P., Sotto, A., Grenga, L., Pible, O., and Armengaud, J. (2020) Proteotyping SARS-CoV-2 virus from nasopharyngeal swabs: A proof-of-concept focused on a 3 min mass spectrometry window. *J. Proteome Res.* **19**, 4407–4416
51. Ihling, C., Tänzler, D., Hagemann, S., Kehlen, A., Hüttelmaier, S., Arlt, C., and Sinz, A. (2020) Mass spectrometric identification of SARS-CoV-2 proteins from gargle solution samples of COVID-19 patients. *J. Proteome Res.* **19**, 4389–4392
52. Rivera, B., Leyva, A., Portela, M. M., Moratorio, G., Moreno, P., Durán, R., and Lima, A. (2020) Quantitative proteomic dataset from oro- and nasopharyngeal swabs used for COVID-19 diagnosis: Detection of viral proteins and host's biological processes altered by the infection. *Data Brief* **32**, 106121
53. Nachtigall, F. M., Pereira, A., Trofymchuk, O. S., and Santos, L. S. (2020) Detection of SARS-CoV-2 in nasal swabs using MALDI-MS. *Nat. Biotechnol.* **38**, 1168–1173
54. Nikolaev, E. N., Indeykina, M. I., Brzhozovskiy, A. G., Bugrova, A. E., Kononikhin, A. S., Starodubtseva, N. L., Petrotchenko, E. V., Kovalev, G. I., Borchers, C. H., and Sukhikh, G. T. (2020) Mass-spectrometric detection of SARS-CoV-2 virus in scrapings of the epithelium of the nasopharynx of infected patients via nucleocapsid N protein. *J. Proteome Res.* **19**, 4393–4397
55. Villar, M., Fernández de Mera, I. G., Artigas-Jerónimo, S., Contreras, M., Gortázar, C., and de la Fuente, J. (2020) Coronavirus in cat flea: Findings and questions regarding COVID-19. *Parasit. Vectors* **13**, 409
56. Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., O'Meara, M. J., Rezelj, V. V., Guo, J. Z., Swaney, D. L., Tummino, T. A., Hüttenhain, R., Kaake, R. M., Richards, A. L., Tutuncuoglu, B., *et al.* (2020) A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468
57. Bojkova, D., Klann, K., Koch, B., Widera, M., Krause, D., Ciesek, S., Cinatl, J., and Münch, C. (2020) Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature* **583**, 469–472
58. Grenga, L., Gallais, F., Pible, O., Gaillard, J. C., Gouveia, D., Batina, H., Bazaline, N., Ruat, S., Culotta, K., Miotello, G., Debroas, S., Roncato, M. A., Steinmetz, G., Foissard, C., Desplan, A., *et al.* (2020) Shotgun proteomics analysis of SARS-CoV-2-infected cells and how it can optimize whole viral particle antigen production for vaccines. *Emerg. Microbes Infect.* **9**, 1712–1721
59. Appelberg, S., Gupta, S., Svensson Akusjärvi, S., Ambikan, A. T., Mikaeloff, F., Saccon, E., Végvári, Á., Benfeitas, R., Sperk, M., Ståhlberg, M., Krishnan, S., Singh, K., Penninger, J. M., Mirazimi, A., and Neogi, U. (2020) Dysregulation in Akt/mTOR/HIF-1 signaling identified by proteo-transcriptomics of SARS-CoV-2 infected cells. *Emerg. Microbes Infect.* **9**, 1748–1760
60. Hamming, I., Timens, W., Bulthuis, M. L., Lely, A. T., Navis, G., and van Goor, H. (2004) Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J. Pathol.* **203**, 631–637
61. Aguiar, J. A., Tremblay, B. J., Mansfield, M. J., Woody, O., Lobb, B., Banerjee, A., Chandiramohan, A., Tiessen, N., Cao, Q., Dvorkin-Gheva, A., Revill, S., Miller, M. S., Carlsten, C., Organ, L., Joseph, C., *et al.* (2020) Gene expression and in situ protein profiling of candidate SARS-CoV-2 receptors in human airway epithelial cells and lung tissue. *Eur. Respir. J.* **56**, 2001123
62. Ziegler, C. G. K., Allon, S. J., Nyquist, S. K., Mbano, I. M., Miao, V. N., Tzouanas, C. N., Cao, Y., Yousif, A. S., Bals, J., Hauser, B. M., Feldman, J., Muus, C., Wadsworth, M. H., Kazer, S. W., Hughes, T. K., *et al.* (2020) SARS-CoV-2 receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is detected in specific cell subsets across tissues. *Cell* **181**, 1016–1035.e19
63. Ingraham, N. E., Barakat, A. G., Reilkoff, R., Bezdicek, T., Schacker, T., Chipman, J. G., Tignanelli, C. J., and Puskarich, M. A. (2020)

- Understanding the renin-angiotensin-aldosterone-SARS-CoV axis: A comprehensive review. *Eur. Respir. J.* **56**, 2000912
64. Liu, M. Y., Zheng, B., Zhang, Y., and Li, J. P. (2020) Role and mechanism of angiotensin-converting enzyme 2 in acute lung injury in coronavirus disease 2019. *Chronic Dis. Transl. Med.* **6**, 98–105
  65. Kai, H., and Kai, M. (2020) Interactions of coronaviruses with ACE2, angiotensin II, and RAS inhibitors—lessons from available evidence and insights into COVID-19. *Hypertens. Res.* **43**, 648–654
  66. Samavati, L., and Uhal, B. D. (2020) ACE2, much more than just a receptor for SARS-CoV-2. *Front. Cell Infect. Microbiol.* **10**, 317
  67. Lumbers, E. R., Delforce, S. J., Pringle, K. G., and Smith, G. R. (2020) The lung, the heart, the novel coronavirus, and the renin-angiotensin system; the need for clinical trials. *Front. Med. (Lausanne)* **7**, 248
  68. Li, Y., Zeng, Z., Cao, Y., Liu, Y., Ping, F., Liang, M., Xue, Y., Xi, C., Zhou, M., and Jiang, W. (2016) Angiotensin-converting enzyme 2 prevents lipopolysaccharide-induced rat acute lung injury via suppressing the ERK1/2 and NF-kappaB signaling pathways. *Sci. Rep.* **6**, 27911
  69. Verdecchia, P., Cavallini, C., Spanevello, A., and Angeli, F. (2020) The pivotal link between ACE2 deficiency and SARS-CoV-2 infection. *Eur. J. Intern. Med.* **76**, 14–20
  70. Glowacka, I., Bertram, S., Herzog, P., Pfefferle, S., Steffen, I., Muench, M. O., Simmons, G., Hofmann, H., Kuri, T., Weber, F., Eichler, J., Drosten, C., and Pöhlmann, S. (2010) Differential downregulation of ACE2 by the spike proteins of severe acute respiratory syndrome coronavirus and human coronavirus NL63. *J. Virol.* **84**, 1198–1205
  71. Bock, J. O., and Ortea, I. (2020) Re-analysis of SARS-CoV-2-infected host cell proteomics time-course data by impact pathway analysis and network analysis: A potential link with inflammatory response. *Aging (Albany NY)* **12**, 11277–11286
  72. Fung, T. S., and Liu, D. X. (2018) Post-translational modifications of coronavirus proteins: Roles and function. *Future Virol.* **13**, 405–430
  73. Gupta, R., Charron, J., Stenger, C. L., Painter, J., Steward, H., Cook, T. W., Faber, W., Frisch, A., Lind, E., Bauss, J., Li, X., Sirpilla, O., Soehnen, X., Underwood, A., Hinds, D., et al. (2020) SARS-CoV-2 (COVID-19) structural and evolutionary dynamicome: Insights into functional evolution and human genomics. *J. Biol. Chem.* **295**, 11742–11753
  74. Sun, Z., Ren, K., Zhang, X., Chen, J., Jiang, Z., Jiang, J., Ji, F., Ouyang, X., and Li, L. (2020) Mass spectrometry analysis of newly emerging coronavirus HCoV-19 spike protein and human ACE2 reveals camouflaging glycans and unique post-translational modifications. *Engineering (Beijing)*. <https://doi.org/10.1016/j.eng.2020.07.014>
  75. Sprung, R., Chen, Y., Zhang, K., Cheng, D., Zhang, T., Peng, J., and Zhao, Y. (2008) Identification and validation of eukaryotic aspartate and glutamate methylation in proteins. *J. Proteome Res.* **7**, 1001–1006
  76. Requena, D., Médico, A., Chacón, R. D., Ramírez, M., and Marín-Sánchez, O. (2020) Identification of novel candidate epitopes on SARS-CoV-2 proteins for South America: A review of HLA frequencies by country. *Front. Immunol.* **11**, 2008
  77. Sardar, R., Satish, D., Birla, S., and Gupta, D. (2020) Integrative analyses of SARS-CoV-2 genomes from different geographical locations reveal unique features potentially consequential to host-virus interaction, pathogenesis and clues for novel therapies. *Heliyon* **6**, e04658
  78. Li, J., Guo, M., Tian, X., Wang, X., Yang, X., Wu, P., Liu, C., Xiao, Z., Qu, Y., Yin, Y., Wang, C., Zhang, Y., Zhu, Z., Liu, Z., Peng, C., et al. (2020) Virus-host interactome and proteomic survey reveal potential virulence factors influencing SARS-CoV-2 pathogenesis. *Med (N. Y.)* **2**, 99–112.e7
  79. Wang, X., Lam, J. Y., Wong, W. M., Yuen, C. K., Cai, J. P., Au, S. W., Chan, J. F., To, K. K. W., Kok, K. H., and Yuen, K. Y. (2020) Accurate diagnosis of COVID-19 by a novel immunogenic secreted SARS-CoV-2 orf8 protein. *mBio* **11**, e02431–20
  80. Sicari, D., Chatziioannou, A., Koutsandreas, T., Sitia, R., and Chevet, E. (2020) Role of the early secretory pathway in SARS-CoV-2 infection. *J. Cell Biol.* **219**, e202006005
  81. Zhang, J., Cruz-Cosme, R., Zhuang, M. W., Liu, D., Liu, Y., Teng, S., Wang, P. H., and Tang, Q. (2020) A systemic and molecular study of subcellular localization of SARS-CoV-2 proteins. *Signal Transduct. Target. Ther.* **5**, 269
  82. Letko, M., Marzi, A., and Munster, V. (2020) Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B beta-coronaviruses. *Nat. Microbiol.* **5**, 562–569
  83. Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S., and Crispin, M. (2020) Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **369**, 330–333
  84. Shajahan, A., Supekar, N. T., Gleinich, A. S., and Azadi, P. (2020) Deducing the N- and O-glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. *Glycobiology* **30**, 981–988
  85. Zhou, D., Tian, X., Qi, R., Peng, C., and Zhang, W. (2021) Identification of 22 N-glycosites on spike glycoprotein of SARS-CoV-2 and accessible surface glycopeptide motifs: Implications for vaccination and antibody therapeutics. *Glycobiology* **31**, 69–80
  86. Zhao, P., Praissman, J. L., Grant, O. C., Cai, Y., Xiao, T., Rosenbalm, K. E., Aoki, K., Kellman, B. P., Bridger, R., Barouch, D. H., Brindley, M. A., Lewis, N. E., Tiemeyer, M., Chen, B., Woods, R. J., et al. (2020) Virus-receptor interactions of glycosylated SARS-CoV-2 spike and human ACE2 receptor. *Cell Host Microbe* **28**, 586–601.e6
  87. Antonopoulos, A., Broome, S., Sharov, V., Ziegenfuss, C., Easton, R. L., Panico, M., Dell, A., Morris, H. R., and Haslam, S. M. (2021) Site-specific characterisation of SARS-CoV-2 spike glycoprotein receptor binding domain. *Glycobiology* **31**, 181–187
  88. Watanabe, Y., Berndsen, Z. T., Raghvani, J., Seabright, G. E., Allen, J. D., Pybus, O. G., McLellan, J. S., Wilson, I. A., Bowden, T. A., Ward, A. B., and Crispin, M. (2020) Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. *Nat. Commun.* **11**, 2688
  89. Towler, P., Staker, B., Prasad, S. G., Menon, S., Tang, J., Parsons, T., Ryan, D., Fisher, M., Williams, D., Dales, N. A., Patane, M. A., and Pantoliano, M. W. (2004) ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. *J. Biol. Chem.* **279**, 17996–18007
  90. Kristiansen, T. Z., Bunkenborg, J., Gronborg, M., Molina, H., Thuluvath, P. J., Argani, P., Goggins, M. G., Maitra, A., and Pandey, A. (2004) A proteomic analysis of human bile. *Mol. Cell. Proteomics* **3**, 715–728
  91. Chen, R., Jiang, X., Sun, D., Han, G., Wang, F., Ye, M., Wang, L., and Zou, H. (2009) Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry. *J. Proteome Res.* **8**, 651–661
  92. Wu, K., Li, W., Peng, G., and Li, F. (2009) Crystal structure of NL63 respiratory coronavirus receptor-binding domain complexed with its human receptor. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19970–19974
  93. Woo, H., Park, S. J., Choi, Y. K., Park, T., Tanveer, M., Cao, Y., Kern, N. R., Lee, J., Yeom, M. S., Croll, T. I., Seok, C., and Im, W. (2020) Developing a fully glycosylated full-length SARS-CoV-2 spike protein model in a viral membrane. *J. Phys. Chem. B* **124**, 7128–7137
  94. Ke, Z., Otonari, J., Qu, K., Cortese, M., Zila, V., McKeane, L., Nakane, T., Zivanov, J., Neufeldt, C. J., Cerikan, B., Lu, J. M., Peukes, J., Xiong, X., Kräusslich, H. G., Scheres, S. H. W., et al. (2020) Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature* **588**, 498–502
  95. Turonova, B., Sikora, M., Schurmann, C., Hagen, W., Welsch, S., Blanc, F., Bulow, S., Gecht, M., Bagola, K., Horner, C., Zandbergen, G., Landry, J., Azevedo, N., Mosalaganti, S., Schwarz, A., et al. (2020) In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. *Science* **370**, 203–208
  96. Grant, O. C., Montgomery, D., Ito, K., and Woods, R. J. (2020) Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition. *Sci. Rep.* **10**, 14991
  97. Prates, E. T., Garvin, M. R., Pavicic, M., Jones, P., Shah, M., Demerdash, O., Amos, B. K., Geiger, A., and Jacobson, D. (2021) Potential pathogenicity determinants identified from structural proteomics of SARS-CoV and SARS-CoV-2. *Mol. Biol. Evol.* **38**, 702–715
  98. Wintjens, R., Bifani, A. M., and Bifani, P. (2020) Impact of glycan cloud on the B-cell epitope prediction of SARS-CoV-2 Spike protein. *NPJ Vaccines* **5**, 81
  99. Yao, H., Song, Y., Chen, Y., Wu, N., Xu, J., Sun, C., Zhang, J., Weng, T., Zhang, Z., Wu, Z., Cheng, L., Shi, D., Lu, X., Lei, J., Crispin, M., et al. (2020) Molecular architecture of the SARS-CoV-2 virus. *Cell* **183**, 730–738.e13
  100. Casalino, L., Gaieb, Z., Goldsmith, J. A., Hjorth, C. K., Dommer, A. C., Harbison, A. M., Fogarty, C. A., Barros, E. P., Taylor, B. C., McLellan, J. S., Fadda, E., and Amaro, R. E. (2020) Beyond shielding: The roles of glycans in the SARS-CoV-2 spike protein. *ACS Cent. Sci.* **6**, 1722–1734



101. Bouhaddou, M., Memon, D., Meyer, B., White, K. M., Rezelj, V. V., Correa Marrero, M., Polacco, B. J., Melnyk, J. E., Ulferts, S., Kaake, R. M., Batra, J., Richards, A. L., Stevenson, E., Gordon, D. E., Rojic, A., *et al.* (2020) The global phosphorylation landscape of SARS-CoV-2 infection. *Cell* **182**, 685–712.e19
102. Klann, K., Bojkova, D., Tascher, G., Ciesek, S., Münch, C., and Cinatl, J. (2020) Growth factor receptor signaling inhibition prevents SARS-CoV-2 replication. *Mol. Cell* **80**, 164–174.e4
103. Carlson, C. R., Asfaha, J. B., Ghent, C. M., Howard, C. J., Hartooni, N., Safari, M., Frankel, A. D., and Morgan, D. O. (2020) Phosphoregulation of phase separation by the SARS-CoV-2 N protein suggests a biophysical basis for its dual functions. *Mol. Cell* **80**, 1092–1103.e4
104. Lu, S., Ye, Q., Singh, D., Cao, Y., Diedrich, J. K., Yates, J. R., Villa, E., Cleveland, D. W., and Corbett, K. D. (2021) The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. *Nat. Commun.* **12**, 502
105. Yang, M., He, S., Chen, X., Huang, Z., Zhou, Z., Zhou, Z., Chen, Q., Chen, S., and Kang, S. (2020) Structural insight into the SARS-CoV-2 nucleocapsid protein C-terminal domain reveals a novel recognition mechanism for viral transcriptional regulatory sequences. *Front. Chem.* **8**, 624765
106. Marsh, M., and Helenius, A. (2006) Virus entry: Open sesame. *Cell* **124**, 729–740
107. Mercer, J., Schelhaas, M., and Helenius, A. (2010) Virus entry by endocytosis. *Annu. Rev. Biochem.* **79**, 803–833
108. Wang, R., Simoneau, C. R., Kulsuptrakul, J., Bouhaddou, M., Trivisano, K. A., Hayashi, J. M., Carlson-Stevermer, J., Zengel, J. R., Richards, C. M., Fozouni, P., Oki, J., Rodriguez, L., Joehnk, B., Walcott, K., Holden, K., *et al.* (2021) Genetic screens identify host factors for SARS-CoV-2 and common cold coronaviruses. *Cell* **184**, 106–119.e14
109. Schneider, W. M., Luna, J. M., Hoffmann, H. H., Sánchez-Rivera, F. J., Leal, A. A., Ashbrook, A. W., Le Pen, J., Ricardo-Lax, I., Michailidis, E., Peace, A., Stenzel, A. F., Lowe, S. W., MacDonald, M. R., Rice, C. M., and Poirier, J. T. (2021) Genome-scale identification of SARS-CoV-2 and pan-coronavirus host factor networks. *Cell* **184**, 120–132.e14
110. Zhu, Y., Feng, F., Hu, G., Wang, Y., Yu, Y., Zhu, Y., Xu, W., Cai, X., Sun, Z., Han, W., Ye, R., Qu, D., Ding, Q., Huang, X., Chen, H., *et al.* (2021) A genome-wide CRISPR screen identifies host factors that regulate SARS-CoV-2 entry. *Nat. Commun.* **12**, 961
111. Gordon, D. E., Hiatt, J., Bouhaddou, M., Rezelj, V. V., Ulferts, S., Braberg, H., Jureka, A. S., Obernier, K., Guo, J. Z., Batra, J., Kaake, R. M., Weckstein, A. R., Owens, T. W., Gupta, M., Pourmal, S., *et al.* (2020) Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science* **370**, eabe9403
112. Perrin-Cocon, L., Diaz, O., Jacquemin, C., Barthel, V., Ogire, E., Ramière, C., André, P., Lotteau, V., and Vidalain, P. O. (2020) The current landscape of coronavirus-host protein-protein interactions. *J. Transl. Med.* **18**, 319
113. Messina, F., Giombini, E., Agrati, C., Vairo, F., Ascoli Bartoli, T., Al Moghazi, S., Piacentini, M., Locatelli, F., Kobinger, G., Maeurer, M., Zumla, A., Capobianchi, M. R., Lauria, F. N., and Ippolito, G. (2020) COVID-19: Viral-host interactome analyzed by network based-approach model to study pathogenesis of SARS-CoV-2 infection. *J. Transl. Med.* **18**, 233
114. Fung, T. S., and Liu, D. X. (2019) Human coronavirus: Host-pathogen interaction. *Annu. Rev. Microbiol.* **73**, 529–557
115. Frieman, M., and Baric, R. (2008) Mechanisms of severe acute respiratory syndrome pathogenesis and innate immunomodulation. *Microbiol. Mol. Biol. Rev.* **72**, 672–685. Table of Contents
116. Poran, A., Harjanto, D., Malloy, M., Arieta, C. M., Rothenberg, D. A., Lenkala, D., van Buuren, M. M., Addona, T. A., Rooney, M. S., Srinivasan, L., and Gaynor, R. B. (2020) Sequence-based prediction of SARS-CoV-2 vaccine targets using a mass spectrometry-based bioinformatics predictor identifies immunogenic T cell epitopes. *Genome Med.* **12**, 70
117. Zeng, X., Song, X., Ma, T., Pan, X., Zhou, Y., Hou, Y., Zhang, Z., Li, K., Karypis, G., and Cheng, F. (2020) Repurpose open data to discover therapeutics for COVID-19 using deep learning. *J. Proteome Res.* **19**, 4624–4636
118. Whetton, A. D., Preston, G. W., Abubeker, S., and Geifman, N. (2020) Proteomics and informatics for understanding phases and identifying biomarkers in COVID-19 disease. *J. Proteome Res.* **19**, 4219–4232
119. Yu, J., Yuan, X., Chen, H., Chaturvedi, S., Braunstein, E. M., and Brodsky, R. A. (2020) Direct activation of the alternative complement pathway by SARS-CoV-2 spike proteins is blocked by factor D inhibition. *Blood* **136**, 2080–2089
120. Messner, C. B., Demichev, V., Wendisch, D., Michalick, L., White, M., Freiwald, A., Textoris-Taube, K., Vernardis, S. I., Egger, A. S., Kreidl, M., Ludwig, D., Kilian, C., Agostini, F., Zelezniak, A., Thibeault, C., *et al.* (2020) Ultra-high-throughput clinical proteomics reveals classifiers of COVID-19 infection. *Cell Syst.* **11**, 11–24.e4
121. Shen, B., Yi, X., Sun, Y., Bi, X., Du, J., Zhang, C., Quan, S., Zhang, F., Sun, R., Qian, L., Ge, W., Liu, W., Liang, S., Chen, H., Zhang, Y., *et al.* (2020) Proteomic and metabolomic characterization of COVID-19 patient sera. *Cell* **182**, 59–72.e15
122. Hou, X., Zhang, X., Wu, X., Lu, M., Wang, D., Xu, M., Wang, H., Liang, T., Dai, J., Duan, H., Xu, Y., Yu, X., and Li, Y. (2020) Serum protein profiling reveals a landscape of inflammation and immune signaling in early-stage COVID-19 infection. *Mol. Cell. Proteomics* **19**, 1749–1759
123. D'Alessandro, A., Thomas, T., Dzieciatkowska, M., Hill, R., Francis, R., Hudson, K., Zimring, J., Hod, E., Spitalnik, S., and Hansen, K. (2020) Serum proteomics in COVID-19 patients: Altered coagulation and complement status as a function of IL-6 level. *J. Proteome Res.* **19**, 4417–4427
124. Yang, L., Liu, S., Liu, J., Zhang, Z., Wan, X., Huang, B., Chen, Y., and Zhang, Y. (2020) COVID-19: Immunopathogenesis and immunotherapeutics. *Signal Transduct. Target. Ther.* **5**, 128
125. Zhong, J., Tang, J., Ye, C., and Dong, L. (2020) The immunology of COVID-19: Is immune modulation an option for treatment? *Lancet Rheumatol.* **2**, e428–e436
126. Vabret, N., Britton, G. J., Gruber, C., Hegde, S., Kim, J., Kuksin, M., Levantovsky, R., Malle, L., Moreira, A., Park, M. D., Pia, L., Risson, E., Saffern, M., Salomé, B., Esai Selvan, M., *et al.* (2020) Immunology of COVID-19: Current state of the science. *Immunity* **52**, 910–941
127. Schulte-Schrepping, J., Reusch, N., Paclik, D., Baßler, K., Schlickeiser, S., Zhang, B., Krämer, B., Krammer, T., Brumhard, S., Bonaguro, L., De Domenico, E., Wendisch, D., Grasshoff, M., Kapellos, T. S., Beckstette, M., *et al.* (2020) Severe COVID-19 is marked by a dysregulated myeloid cell compartment. *Cell* **182**, 1419–1440.e23
128. Kimhofer, T., Lodge, S., Whitley, L., Gray, N., Loo, R. L., Lawler, N. G., Nitschke, P., Bong, S.-H., Morrison, D. L., Begum, S., Richards, T., Yeap, B. B., Smith, C., Smith, K. G. C., Holmes, E., *et al.* (2020) Integrative modeling of quantitative plasma lipoprotein, metabolic, and amino acid data reveals a multiorgan pathological signature of SARS-CoV-2 infection. *J. Proteome Res.* **19**, 4442–4454
129. Shin, D., Mukherjee, R., Grewe, D., Bojkova, D., Baek, K., Bhattacharya, A., Schulz, L., Widera, M., Mehdipour, A. R., Tascher, G., Geurink, P. P., Wilhelm, A., van der Heden van Noort, G. J., Ova, H., Müller, S., *et al.* (2020) Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature* **587**, 657–662
130. Bestle, D., Heindl, M. R., Limburg, H., Van Lam van, T., Pilgram, O., Moulton, H., Stein, D. A., Harges, K., Eickmann, M., Dolnik, O., Rohde, C., Klenk, H. D., Garten, W., Steinmetzer, T., and Böttcher-Friebertshäuser, E. (2020) TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Sci. Alliance* **3**, e202000786
131. Leon, K., and Ott, M. (2021) An 'arms race' between the nonsense-mediated mRNA decay pathway and viral infections. *Semin. Cell Dev. Biol.* **111**, 101–107
132. Beacon, T. H., Su, R. C., Lakowski, T. M., Delcuve, G. P., and Davie, J. R. (2020) SARS-CoV-2 multifaceted interaction with the human host. Part II: Innate immunity response, immunopathology, and epigenetics. *IUBMB Life* **72**, 2331–2354
133. Surjit, M., and Lal, S. K. (2009) The nucleocapsid protein of the SARS coronavirus: Structure, function and therapeutic potential. *Mol. Biol. SARS-Coronavirus*, 129–151
134. Schoeman, D., and Fielding, B. C. (2019) Coronavirus envelope protein: Current knowledge. *Virology* **16**, 69
135. Asghari, A., Naseri, M., Safari, H., Saboory, E., and Parsamanesh, N. (2020) The novel insight of SARS-CoV-2 molecular biology and pathogenesis and therapeutic options. *DNA Cell Biol.* **39**, 1741–1753
136. Burtscher, J., Cappellano, G., Omori, A., Koshiba, T., and Millet, G. P. (2020) Mitochondria: In the cross fire of SARS-CoV-2 and immunity. *iScience* **23**, 101631
137. [preprint] Ramlall, V., Thangaraj, P., Meydan, C., Foox, J., Butler, D., May, B., Freitas, J., Glicksberg, B., Mason, C., Tatonetti, N., and Shapira, S. (2020) Identification of immune complement function as a determinant of adverse SARS-CoV-2 infection outcome. *medRxiv*. <https://doi.org/10.1101/2020.05.05.20092452>