Open Modification Searching of SARS-CoV-2–Human Protein Interaction Data Reveals Novel Viral Modification Sites

Authors
Charlotte Adams, Kurt Boonen, Kris Laukens, and Wout Bittremieux

Correspondence
wbittremieux@health.ucsd.edu

In Brief
Public MS data were reanalyzed using open modification searching to investigate the presence of post-translational modifications during interaction of viral severe acute respiratory syndrome coronavirus 2 proteins and human host proteins. We examined several post-translational modifications, including phosphorylation, ubiquitination, and S-nitrosylation, in the context of the virus–host protein–protein interaction network to obtain novel hypotheses about the functional roles of these modifications during viral infection by severe acute respiratory syndrome coronavirus 2.

Highlights
- Identification of protein interactions between SARS-CoV-2 and its human host.
- Insights into the role of post-translational modifications during viral infection.
- Human kinases interact with phosphorylated viral proteins.
- Indications that SARS-CoV-2 hijacks the host ubiquitination system.
- Hypothesis of antiviral mechanism related to S-nitrosylation of SARS-CoV-2 proteins.

© 2022 THE AUTHORS. Published by Elsevier Inc on behalf of American Society for Biochemistry and Molecular Biology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
https://doi.org/10.1016/j.mcpro.2022.100425
Open Modification Searching of SARS-CoV-2–Human Protein Interaction Data Reveals Novel Viral Modification Sites

Charlotte Adams¹,², Kurt Boonen²,³, Kris Laukens¹, and Wout Bittremieux⁴,*

The outbreak of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of the coronavirus 2019 disease, has led to an ongoing global pandemic since 2019. Mass spectrometry can be used to understand the molecular mechanisms of viral infection by SARS-CoV-2, for example, by determining virus–host protein–protein interactions through which SARS-CoV-2 hijacks its human hosts during infection, and to study the role of post-translational modifications. We have reanalyzed public affinity purification–mass spectrometry data using open modification searching to investigate the presence of post-translational modifications in the context of the SARS-CoV-2 virus–host protein–protein interaction network. Based on an over twofold increase in identified spectra, our detected protein interactions show a high overlap with independent mass spectrometry-based SARS-CoV-2 studies and virus–host interactions for alternative viruses, as well as previously unknown protein interactions. In addition, we identified several novel modification sites on SARS-CoV-2 proteins that we investigated in relation to their interactions with host proteins. A detailed analysis of relevant modifications, including phosphorylation, ubiquitination, and S-nitrosylation, provides important hypotheses about the functional role of these modifications during viral infection by SARS-CoV-2.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a highly transmissible and pathogenic coronavirus that causes coronavirus disease 2019 (COVID-19). Since late 2019, it has caused an ongoing global pandemic and socio-economic crisis, with currently over 610 million people infected and over 6.5 million deaths globally (World Health Organization, September 2022) (1). Despite intensive research by the scientific community, many important questions regarding its molecular mechanisms remain unanswered.

Viruses are opportunistic intracellular pathogens that depend on their interactions with host proteins to ensure their survival and propagation. The study of protein–protein interactions (PPIs) between SARS-CoV-2 and human proteins is important to understand the mechanisms of viral infection by SARS-CoV-2 (2) and develop therapeutic treatments (3). Proteins and their function can be altered by post-translational modifications (PTMs) to increase the functional diversity of the limited number of proteins that are encoded. Interestingly, PTMs have been characterized on several coronavirus proteins, in spite of the fact that coronaviruses lack enzymes capable of introducing such modifications (4). Elucidating the roles of PTMs in a mechanistic context is important to understand viral infection, as they are crucial for viral protein function and promote viral replication, assembly, and release. For example, phosphorylation of the SARS-CoV-2 nucleocapsid protein allows its shuttling between cellular compartments (5). In addition, insights in viral and host PTM dynamics offer a potential avenue toward the development of antiviral therapies. Removing PTMs that play a role in the enzymatic activity of viral proteins can aid the host in overcoming viral infection. Alternatively, viral proteins can be modified leading to their inactivation and/or proteasomal degradation, for example, by attaching ubiquitin (6).

Affinity purification–mass spectrometry (AP–MS)–based proteomics was used to compile the first SARS-CoV-2–human PPI map, revealing interactions with proteins involved in major cellular processes, including DNA replication, RNA processing, and vesicle trafficking to give insights into SARS-CoV-2 infection (2). In addition, the interaction map revealed several human proteins that are targeted by existing drugs approved by the US Food and Drug Administration, which may be potential targets for drug repurposing. The majority of previously conducted studies use standard sequence database searching to process AP–MS data (2, 7–10). Most search engines require the user to explicitly specify potential protein modifications to be considered during searching. Unfortunately, taking into account multiple
were denatured, reduced, carbamidomethylated, and enzymatically 
cut (biological replicates) in HEK-293T cells. Bead-bound proteins 
CoV-2 proteins that were individually tagged and expressed in tripli-

Spectrum identification Using Open Modification Searching

First, the downloaded raw files were converted to MGF files using 
ThermoRawFileParser (version 1.2.3) (17). Next, OMS was performed 
using the ANN-SoLo spectral library search engine (version 0.2.4) 
(13, 14). A combined human–SARS-CoV-2 spectral library was used for 
searching. The MassIVE-KB library (version June 15, 2018) was used as 
human spectral library. This is a comprehensive human higher energy 
collisional dissociation spectral library containing 2,154,269 unique 
precursors corresponding to 1,114,503 unique peptides, derived from 
publicly available MS data in the MassIVE repository (18). SARS-CoV-2 
spectra were simulated by generating all possible tryptic peptide se-
quences from the SARS-CoV-2 protein sequences downloaded from 
UniProt (version March 05, 2020) using Pyteomics (version 4.3.2) (19) 
and predicting the corresponding spectra using ProSigt (version prosi-
t intensity_2020_hcd; collision energy 33 as determined by ProSigt 
collision energy calibration) (20). A simulated spectral library for the 
green fluorescent protein was generated in a similar fashion. A final 
spectral library was compiled by merging all spectra using SpectraST 
(version 5.0) (21) and adding decoy spectra in a 1:1 ratio using the 
shuffle-and-reposition method (22). ANN-SoLo was configured to use a 
20 ppm precursor mass tolerance during the first step of its cascade 
search and a 500 Da precursor mass tolerance during its open search. 
Other search settings were to filter peaks below 101 m/z, above 1500 m/ 
z, and in a 0.5 m/z window around the precursor mass; a 0.02 m/z 
fragment mass tolerance; and a bin size of 0.05 m/z. The remaining 
settings were kept at their default values. Peptide-spectrum matches 
(PSMs) were filtered at 1% FDR using ANN-SoLo’s built-in subgroup 
FDR procedure (supplemental Table S1). 

In addition, OMS was performed using MStagger (version 3.5) (15) 
and FragPipe (version 18.0) against a concatenated FASTA file con-
taining human protein sequences (UniProt reviewed sequences 
downloaded on February 28, 2020) (23), the SARS-CoV-2 protein 
sequences (version March 05, 2020), and the green fluorescent protein 
sequence. An equal number of decoy protein sequences was gener-
ated using FragPipe. The MSStagger search settings included a pre-
cursor mass tolerance between −150 Da and 500 Da, a fragment mass 
tolerance of 0.02 Da, and trypsin cleavage with up to two missed 
cleavages. Cysteine carbamidomethylation was used as a fixed 
modification, and oxidation of methionine and N-terminal acetylation 
were used as variable modifications. Other search settings were kept 
at their default values. PSMs were processed using PeptideProphet 
(version 4.4.0) (24) with the FragPipe default settings for open 
searches and filtered at 1% FDR.

Experimental Procedures

Experimental Design and Statistical Rationale

A previously generated AP–MS dataset to study the SARS-CoV-2 virus–host interactome (2) was retrieved from the PRIDE repository 
(PXD018117) (16). For the full experimental details, see the original study by Gordon et al. (2). In brief, AP was performed using 27 SARS-
CoV-2 proteins that were individually tagged and expressed in tripli-
cate (biological replicates) in HEK-293T cells. Bead-bound proteins were 
denatured, reduced, carbamidomethylated, and enzymatically 
digested using trypsin, and each sample was injected via an Easy-nLC 
1200 (Thermo Fisher Scientific) into a Q-Exactive Plus mass spec-
trometer (Thermo Fisher Scientific). The SARS-CoV-2 proteins that 
were included are all mature nonstructural proteins (Nsp5), except for 
Nsp5 and Nsp16; a mutated version of Nsp5 to disable its proteolytic 
activity (Nsp5, C145A); and all predicted SARS-CoV-2 ORFs, including 
the spike (S), membrane (M), nucleocapsid (N), and envelope (E) 
protein. Spectrum identifications were filtered at 1% FDR, and PPIs 
were filtered using a SAINTexpress Bayesian false discovery rate 
(BFDR) ≤ 0.05, an average spectral count ≥ 2, and a MiST score ≥ 0.7 
(see further).

Protein Inference and PPI Filtering

Protein inference was performed using the Protein Inference Algo-
rithms tool (version 1.3.13) (25) based on Occam’s razor. Only proteins 
with minimum two unique peptides were retained, whereas other 
settings were kept at their default values (supplemental Table S2). 
Combined scoring of interacting proteins using SAINTexpress (version 3.6.3) (26) and MiST ([https://modbase.compbio.ucsf.edu/mist/], 
version main.e2da2b0) (27) was used to filter high-confidence PPIs 
(supplemental Table S3). Scoring thresholds were a SAINTexpress 
BFDR ≥ 0.05, an average spectral count ≥ 2, and a MiST score ≥ 0.7 
(supplemental Table S4). 

To validate the filtered PPIs, a list of SARS-CoV-2–human interactions reported in seven alternative MS-based SARS-CoV-2 virus–host interactome studies (7–10, 28–30) was obtained from the BioGRID repository (31). In addition, PPIs from a previous reanalysis of the original AP–MS study were included (12). Human interaction
partners were queried in the VirHostNet database (version 3.0) (32) to investigate whether these proteins are also targeted by other viruses.

**Gene Ontology Enrichment Analysis**

Gene Ontology (GO) enrichment analysis was performed for the human proteins that interact with each viral protein, using the enrichGO function of the clusterProfiler package (version 4.0.5) in R. Significant GO terms corresponding to the biological process category (1% Benjamini–Hochberg FDR) were extracted and further refined to select nonredundant terms using the rvgo package (version 1.4.4) with default parameters (https://ssayols.github.io/rvgo).

**Investigation of PTMs**

The precursor mass differences observed in the ANN-SoLo results were referenced against the Unimod database (33) to determine the modifications that were present. PSMs that included specific PTMs (phosphorylation, ubiquitination, and S-nitrosylation) were manually investigated in more detail to disambiguate between alternative PTM assignments with near-identical mass and determine the modification site by visual inspection using the spectrum_utils Python package (version 0.3.3) (34).

Identified PTM sites were verified using PTM prediction tools and through literature study. Phosphorylation results were compared with three independent SARS-CoV-2 phosphoproteomic studies (35–37). Phosphorylation site prediction was performed using NetPhos (version 3.1) (38, 39). In addition, GPS (version 5.9) (40) was used to predict kinase-specific phosphorylation sites (with a “high” threshold) for each of the SARS-CoV-2 proteins that were found to interact with known human kinases, retrieved from the KinHub database (41) (supplemental Tables S5 and S6). Ubiquitination results were compared with two independent SARS-CoV-2 ubiquitination studies (7, 42) and ubiquitination sites predicted by BDM-PUB (version 1.0) (http://bdmpub.biocuckoo.org/) (supplemental Table S7). To our knowledge, to date, there have been no S-nitrosylation sites reported on SARS-CoV-2. S-nitrosylation site prediction was performed with iSNO-PseAAC (version 1.0) (43) to validate observed S-nitrosylation sites (supplemental Tables S8 and S9).

**RESULTS**

**Increased Spectrum Identification Rate Using Open Modification Searching Boosts PPI Confidence**

Open modification searching using ANN-SoLo succeeded in identifying 830,743 of 2,503,010 total MS/MS spectra at 1% FDR (33% identification rate). This represents a 214% increase in identified spectra compared with the originally reported results (2) obtained by standard searching using MaxQuant (44) (Fig. 1A). Notably, 402,586 PSMs correspond to modified peptides with nonzero precursor mass differences from open modification searching (Fig. 1C and supplemental Table S10). Besides the increase in identified spectra, an important advantage of open modification searching is its ability to identify any type of PTM in an unbiased fashion, without the need to explicitly specify a limited number of variable modifications. This makes it possible to explore the general presence of PTMs in the context of the SARS-CoV-2 virus–host interactome. Frequently observed PTMs include modifications that were likely artificially introduced during sample processing (45), such as oxidation, dioxygenation, and acetylation. Such ubiquitous modifications are typically included as variable modifications during standard searching. The OMS results also include unique observations of biologically relevant modifications at lower abundances, such as phosphorylation, ubiquitination, and S-nitrosylation.

**Similar beneficial results can also be achieved with alternative OMS tools, such as MSFragger (15). Performing an open search using MSFragger instead of ANN-SoLo succeeded in identifying 574,394 of 2,503,010 total MS/MS spectra at 1% FDR (23% identification rate). This represents a 148% increase in identified spectra compared with the originally reported results (2) (supplemental Fig. S1A). Furthermore, there is a strong correspondence in identification results between ANN-SoLo and MSFragger (supplemental Fig. S1, B and C), which indicates the robustness of open modification searching, irrespective of the search engine that is employed.**

For simplicity of the downstream analyses and because ANN-SoLo was more sensitive than MSFragger, the ANN-SoLo results were used to investigate protein interactions during viral infection by SARS-CoV-2.

**PPI filtering is crucial to separate true interactors from nonspecific binders and contaminants. A combination of SAINTextpress and MiST filtering of the open modification searching results produced 375 high-confidence PPIs (supplemental Table S4). These results contain 164 PPIs that overlap with the previously reported results (2) and 211 novel PPIs (supplemental Fig. S2). Notably, a previous reanalysis of these AP–MS data using alternative bioinformatics tools reported a similar overlap with the original PPI results (12). The difference in detected PPIs is partly because of the difference in spectrum identification and PPI filtering strategies. In the original analysis, a two-step filtering strategy was used. In the first step, the PPIs were filtered by a SAINTextpress BFDR ≤ 0.05, an average spectral count ≥ 2, and a MiST score ≥ 0.7. All proteins that fulfilled the first filtering step were searched in the CORUM (46) database of known protein complexes, and information was extracted about the stable protein complexes that they participated in. In the second step, all the proteins that formed known complexes with interactors identified in the first step were subjected to a filtering step with a lower stringency (MiST score ≥ 0.6) (2). In contrast, in the reanalysis, only a single filtering step was performed using a SAINTextpress BFDR ≤ 0.05, an average spectral count ≥ 2, and a MiST score ≥ 0.7.**

When compared with the unfiltered PPI data of the original analysis (2), 199 of the 211 novel PPIs were previously detected as well but failed the original PPI filtering thresholds (supplemental Fig. S3). Most discarded PPIs (~88%) did not pass the minimum threshold for the MiST score, which is a linear combination of the prey abundance, the prey reproducibility across repeated runs, and the specificity of the prey relative to other baits (47). About 28% of the discarded PPIs...
did not pass the SAINTexpress BFDR filter, which is based on the abundance of the preys and control proteins (26). The difference in filtered PPIs in the reanalysis is driven by the increased number of PSMs per identified protein from the open modification results, which resulted in a more reliable identification of both background proteins and potential interaction partners. This influences the PPI filtering and results in a more robust identification of true interactors.

The PPIs were validated against seven alternative MS-based SARS-CoV-2 virus–host interactome studies (7–10, 28–30) and the results from a previous reanalysis (12). The largest overlap was found with the previous reanalysis and the original analysis, as these studies make use of the same experimental data and only differ in the computational tools that were employed (Fig. 2A). Although most independent studies contributed unique PPIs, there is a strong correspondence in the superset of detected PPIs across all independent studies, with 317 of the 375 PPIs that were detected in this study previously reported in one or more of the other studies (Figs. 2A and S2). In addition, investigating the overlap with the VirHostNet database (32) shows that 328 of the human interaction partners are also targeted by other viruses (Fig. 2B). These results are significantly enriched in the VirHostNet database (Fisher’s exact test, \( p \ll 0.001 \)), which is expected as different viruses have similar strategies to hijack the host, for example, through functional conservation or viral recombination (32).

GO enrichment analysis based on the human interacting proteins for each SARS-CoV-2 protein indicates how viral infection might hijack major cellular processes, including metabolic processes involving noncoding RNA (ncRNA; Nsp8) and glycoproteins (Orf8), and RNA export from the nucleus (Orf6) (Fig. 3). Both microRNA and long ncRNA are ncRNA that can regulate gene expression and exhibit different expression profiles in COVID-19 patients compared with healthy people (48, 49). In addition, they have recently been found to act as viral modulators, regulating viral infection and host defense (50, 51). Glycoprotein metabolic processes are expected to be relevant to Orf8, as induction of proinflammatory cytokine production by secreted Orf8 is glycosylation dependent (52).
RNA export from the nucleus is also relevant, as SARS-CoV-2-infected cells have an increased level of nuclear mRNA accumulation. Confirming our detected interactions between Orf6 and the mRNA export factors Rae1 and Nup98, a recent study found that Orf6 uses this process to trap host mRNA (53). In addition to these major cellular processes, other biological processes relevant during viral infection were found, including viral transcription and viral gene expression (supplemental Fig. S4).

**Interaction Between Human Kinases And Phosphorylated Viral Proteins**

Phosphorylation is a common PTM that impacts many basic cellular processes (54). It usually results in a functional change of the target protein, interfering with its enzymatic activity, cellular location, and/or association with other proteins. Several studies have indicated that also the function of viral proteins can be affected by their phosphorylation status. For example, association of the SARS-CoV-2 nucleocapsid protein with the 14-3-3 host proteins depends on its phosphorylation status (55), and phosphorylation of the SARS-CoV-1 nucleocapsid protein has been proven to be important for the regulation of the viral life cycle (56).

Twenty phosphorylated viral peptides were identified using open modification searching, and phosphosites were localized to specific residues using visual inspection where possible (supplemental Table S5 and supplemental Figs. S5–S24). Notably, despite the fact that no phosphopeptide enrichment protocol was used during the original AP–MS study, we identified several phosphorylation sites that match results from specialized phosphoproteomics studies performed on SARS-CoV-2 infected cells (35–37) (Fig. 4A).

Phosphorylation status is determined by the interplay between kinases and phosphatases. Kinases catalyze the attachment of phosphate groups to target proteins, and phosphatases remove phosphate groups from target proteins (57). SARS-CoV-2 does not encode any kinases or
phosphatases and thus relies on host proteins for its phosphorylation status. Several kinases and phosphatases have been identified as potential drug targets, and a few kinase inhibitors are currently being used to treat COVID-19. For example, baricitinib is recommended by the World Health Organization for patients with severe or critical COVID-19. It is a Janus kinase inhibitor that suppresses overstimulation of the immune system by preventing phosphorylation of key proteins involved in the signal transduction that leads to immune activation and inflammation (58), and it is able to prevent SARS-CoV-2 from entering the cell by inhibiting clathrin-mediated endocytosis (59).

We detected multiple interactions between SARS-CoV-2 proteins and host kinases and phosphatases (supplemental Tables S11 and S12), including receptor interaction protein kinase 1 (RIPK1; Nsp12) and Tank-binding kinase 1 (TBK1; Nsp13), both of which have previously been linked to SARS-CoV-2. Active phosphorylated RIPK1 was found in epithelial cell samples from COVID-19 patients (60), and a recent study discovered that Nsp12 promotes activation of RIPK1, which in turn enhances viral replication by stimulating the expression of viral receptors, such as angiotensin-converting enzyme 2 and epidermal growth factor receptor (61). Furthermore, Nsp13 is found to limit the activation of TBK1 by directly binding to TBK1 (62). In addition, we uncovered previously unknown interactions with host kinases, such as the interaction between Nsp12 and TLK1 (tousled-like kinase 1; Fig. 4B). This interaction is particularly interesting as kinase-specific phosphorylation site prediction revealed that an observed phosphorylation site on Nsp12 (T21) is most likely caused by TLK1.

Ubiquitination of SARS-CoV-2 Proteins

Ubiquitination is a common PTM that can affect the localization, stability, and function of proteins (63). It regulates a variety of cellular processes, including protein degradation, protein trafficking, transcription, cell-cycle control, and cell signaling (64). Depending on the cellular context, ubiquitin attachment may either promote or inhibit the viral life cycle. Viruses have developed means to exploit protein ubiquitination by enhancing or inhibiting ubiquitination of specific substrates depending on their needs. They can redirect protein degradation toward proteins with antiviral activity (65) or use ubiquitination to regulate viral proteins. For example, transcriptional function of the HIV type-1 Tat protein is increased by the addition of a single ubiquitin molecule (66). More recently, it has been discovered that the inhibition of interferon α signaling by SARS-CoV-2 Orf7a depends on the ubiquitination of K119 (67).

Ten ubiquitinated viral peptides were identified, and, where possible, ubiquitination sites were localized to specific residues using visual inspection (supplemental Table S7 and supplemental Figs. S25–S34). Notably, all ubiquitination sites that could be confidently localized have been previously reported in independent studies (7, 42) (Fig. 5A) or could be predicted. Interestingly, although ubiquitination of K16 on Orf9c has not been reported in SARS-CoV-2 yet, this site has been observed to be ubiquitinated in SARS-CoV-1 (7).

The final step of ubiquitination is performed by a ubiquitin (E3) ligase. This enzyme is particularly important as it determines to which substrate protein the ubiquitin is attached. A ubiquitin can be removed by deubiquitinating enzymes (DUBs). Both host E3 ligases and DUBs were found to interact with SARS-CoV-2 proteins (supplemental Tables S13 and S14), suggesting that SARS-CoV-2 might hijack the host ubiquitination system. For example, Nsp13 was found to interact with the DUB USP13 (Fig. 5B). According to a recent study, Nsp13 likely hijacks USP13 to prevent itself from...
degradation. Both knockdown of USP13 and its inhibition through spautin-1 resulted in decreased levels of Nsp13, suggesting that USP13 deubiquitinates and consequently stabilizes Nsp13 (68).

S-nitrosylation of SARS-CoV-2 Proteins

A promising compound currently undergoing clinical trials for COVID-19 is nitric oxide (NO) (69). Besides its role as an important vasodilator, to prevent blood clot formation, it functions as a vital immune mediator, exerting broad-spectrum antiviral effects (70). NO potentially prevents infection by SARS-CoV-2, as it was suggested for SARS-CoV (71). The surprisingly low prevalence of smokers among hospitalized COVID-19 patients (72) could be attributed to the intermittent burst of high NO concentration in cigarette smoke (73). Also nitrate-rich nutrition, exercise,
and breathing through your nose is hypothesized to prevent a SARS-CoV-2 infection, as they all increase the NO concentration (71, 74).

The general antiviral mechanism appears to be the NO-mediated S-nitrosylation of viral and host proteins (75). S-nitrosylation is the reversible covalent attachment of NO to the thiol side chain of cysteine. It is one of the most important and universal PTMs, and it can act as a global regulator of protein function akin to phosphorylation and ubiquitination (76). Interestingly, reactive cysteine residues are present in many viral and host proteins, representing possible targets for NO (75). Results of an in vitro study suggest that S-nitrosylation of the SARS-CoV-2 3CL protease (Nsp5) can directly inhibit its protease activity and reduce viral replication (77).

To the best of our knowledge, there are currently no reported S-nitrosylation sites on SARS-CoV-2 proteins. Because cysteine residues are typically considered with a fixed carbamidomethylation modification (+57.021464 Da) introduced during sample processing by alkylation with iodoacetamide (45), Prosit (20)—which was used to simulate SARS-CoV-2 spectra—always considers cysteine residues to be carbamidomethylated. However, the presence of a prior modification can block reduction and alkylation of cysteine residues (78). Therefore, rather than identifying S-nitrosylation directly based on its mass difference of +28.990164 Da, this modification is represented by a mass difference of −28.031300 Da (28.990164 Da – 57.021464) in the open modification searching results. Because the corresponding S-nitrosylation mass difference is identical to the mass difference observed from the valine to alanine and methionine to cysteine amino acid substitutions, careful visual inspection to localize the modification to specific amino acid residues was performed to confirm three S-nitrosylation sites (supplemental Table S8 and supplemental Figs. S35–S52).

Interestingly, one of these S-nitrosylation sites (C45) is located on the main protease Nsp5, specifically, near the catalytic site. Nsp5 is one of two cysteine proteases necessary for viral replication and assembly, and its protease activity is suggested to be directly inhibited by S-nitrosylation (77). A recent study found that C45 is a hyper-reactive cysteine with a higher nucleophilicity than C145—the catalytic cysteine of Nsp5—and identified C45 as an attractive binding site for the development of a covalent inhibitor (79).

**DISCUSSION**

Mass spectrometry research can help to provide insights into the etiology of SARS-CoV-2 infection and identify potential therapeutic targets by investigating the roles of viral and host proteins during infection, their protein interactions, and PTMs (12). In addition, it can be used to develop diagnostic strategies, such as through efforts of the CoV-MS consortium (80)—a partnership by multiple academic and industrial groups to increase applicability, accessibility, sensitivity, and robustness of an MS-based diagnostic test that detects proteolytically digested SARS-CoV-2 proteins.

Here, we have used open modification searching to reprocess SARS-CoV-2–human AP–MS data to investigate PTMs in the context of PPIs between SARS-CoV-2 and its human host. Studying these two aspects in tandem is especially relevant to understand the mechanisms of viral infection, because although PTMs are essential for viral replication, coronaviruses lack the enzymes to introduce these themselves. Whereas most SARS-CoV-2 studies so far have only considered a limited number of modifications commonly introduced during sample processing, our open modification searching strategy enabled the unbiased investigation of any PTM. This increased the number of identified spectra by more than twofold, corresponding to newly identified modified peptides, and enabled us to put these PTMs in the context of the virus–host PPI network.

We discovered several combinations of relevant protein interactions and PTMs that hint at functional roles during viral infection. Specifically, we investigated phosphorylation, ubiquitination, and S-nitrosylation in more detail. These are reversible PTMs that are essential for a variety of cellular processes and could be of importance during viral infection. We found interactions between phosphorylated SARS-CoV-2 proteins and host kinases, between ubiquitinated SARS-CoV-2 proteins and host E3 ligases, and novel S-nitrosylation sites on SARS-CoV-2 proteins. Notably, even though no specialized modification enrichment was performed, we were able to confidently detect and localize multiple PTM sites that have been previously reported in independent enrichment studies, as well as novel PTM sites on SARS-CoV-2 proteins.

An important consideration in this study is the overlap in PPI results with the original study (2), which used standard searching instead of open modification searching. Although the vast majority of original spectrum identifications (79%) could be replicated using open modification searching, the overlap in PPIs was smaller (49%). Notably, another recent reanalysis of these data using alternative bioinformatics software tools showed a similarly limited overlap with the original PPI results (12). Besides using stringent FDR control and other PPI filtering settings during all data processing steps, we validated the detected protein interactions through comparison with alternative MS-based SARS-CoV-2 studies and general virus–host interactions in the ViHostNet database. This showed a highly significant overlap between our PPI results and these independent sources, reinforcing the validity of the newly detected interactions.

The partial mismatch in PPI results can be largely explained by more complete identification of both background proteins and potential interaction partners using OMS. As such, the current study highlights the dependence of PPI filtering on the preceding protein identification results. First, open modification searching can be used to obtain high-quality identification data from which comprehensive PPI results can be obtained.
Second, PPI filtering algorithms have to be able to robustly deal with uncertainty and missingness in the identification results when determining true protein interactions. Because OMS increased both the robustness of the PPI results when determining true protein interactions. Because OMS increased both the robustness of the PPI results when determining true protein interactions. Because OMS increased both the robustness of the PPI results when determining true protein interactions. Because OMS increased both the robustness of the PPI results when determining true protein interactions.

In conclusion, we have used open modification searching to reanalyze open SARS-CoV-2 virus–host protein interaction data. The presented results enrich our knowledge of viral infection by SARS-CoV-2 by putting PTMs in the context of the virus–host PPI network, which provides important hypotheses on the functional roles of PTMs during SARS-CoV-2 infection.

DATA AVAILABILITY

All reanalysis results have been deposited to the ProteomeXchange Consortium (81) via the MassIVE repository with dataset identifiers RMSV00000359.1 and RMSV00000309.4. Annotated spectra can be inspected in MS-Viewer (https://msviewer.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msviewer) using the key mcorgdnzfy. All data processing codes are freely available on GitHub as open source under the Apache 2.0 license at https://github.com/adamscharlotte/SARS-Cov-2-analysis.

Supplemental data—This article contains supplemental data.

Acknowledgments—K. L. acknowledges support of the University of Antwerp Core Facility BIOMINA. This research was supported by the University of Antwerp BOF DOCPRO, FWO G070722N, and IBOF VLIR MIMICRY. The computational resources and services used in this work were provided by the HPC core facility CalCUA of the University of Antwerp, and VSC ( Flemish Supercomputer Center), funded by the Research Foundation—Flanders (FWO) and the Flemish Government. The graphical abstract was created with BioRender.com.

Author contributions—W. B. conceptualization; K. B., K. L., and W. B. methodology; C. A. and W. B. software; C. A. and W. B. validation; C. A. and K. B. formal analysis; C. A., K. B., K. L., and W. B. writing—original draft; C. A., K. B., K. L., and W. B. writing—review & editing; C. A. visualization; K. B., K. L., and W. B. supervision; K. B., K. L., and W. B. project administration; K. B. and K. L. funding acquisition.

Conflict of interest—The authors declare no competing interests.

Abbreviations—The abbreviations used are: AP–MS, affinity purification–mass spectrometry; BFDR, Bayesian false discovery rate; COVID-19, coronavirus disease 2019; DUB, deubiquitinating enzyme; FDR, false discovery rate; GO, Gene Ontology; HCD, higher energy collisional dissociation; MS, mass spectrometry; ncRNA, noncoding RNA; NO, nitric oxide; Nsp, nonstructural protein; OMS, open modification searching; PPI, protein–protein interaction; PSM, peptide-spectrum match; PTM, post-translational modification; RIPK1, receptor interaction protein kinase 1; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; TBK1, Tank-binding kinase 1.

Received March 16, 2022, and in revised form, September 18, 2022 Published, MCPRO Papers in Press, October 12, 2022, https://doi.org/10.1016/j.mcpro.2022.100425

REFERENCES

1. Dong, E., Du, H., and Gardner, L. (2020) An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect. Dis. 20, 533–534
2. Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obemier, K., White, K. M., et al. (2020) A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature 583, 459–468
3. Valiente, P. A., Wen, H., Nim, S., Lee, J., Kim, H. J., Kim, J., et al. (2021) Computational design of potent D-peptide inhibitors of SARS-CoV-2. J. Med. Chem. 64, 14955–14967
4. Mishra, S., Bassi, G., and Nyomba, B. G. (2021) Inter-proteomic post-translational modifications of the SARS-CoV-2 and the host proteins A new frontier. Exp. Biol. Med. (Maywood) 246, 749–757
5. Tung, H. Y. L., and Limtung, P. (2020) Mutations in the phosphorylation sites of SARS-CoV-2 encoded nucleocapsid protein and structure model of sequestration by protein 14-3-3. Biochem. Biophys. Res. Commun. 532, 134–138
6. Kumar, R., Mehta, D., Mishra, N., Nayak, D., and Sunil, S. (2020) Role of host-mediated post-translational modifications (PTMs) in RNA virus pathogenesis. Int. J. Mol. Sci. 22, 323
7. Stukalov, A., Girault, V., Grass, V., Karayel, O., Bergant, V., Urban, C., et al. (2021) Multi-level proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV. Nature 594, 246–252
8. Li, J., Guo, M., Tian, X., Wang, X., Yang, X., Wu, P., et al. (2020) Virus-host interactome and proteomic survey reveal potential virulence factors influencing SARS-CoV-2 pathogenesis. Med (NY) 2, 99–112.e7
9. [preprint] Chen, Z., Wang, C., Feng, X., Nie, L., Tang, M., Zhang, H., et al. (2021) Comprehensive analysis of the host-virus interactome of SARS-CoV-2. bioRxiv, https://doi.org/10.1101/2020.12.31.429661
10. Liu, X., Huuskonen, S., Laitinen, T., Redchuk, T., Bogacheva, M., Salokas, K., et al. (2021) SARS-CoV-2-host proteome interactions for antiviral drug discovery. Mol. Syst. Biol. 17, e10396
11. Elias, J. E., and Gygi, S. P. (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat. Met. 4, 207–214
12. Bittremieux, W., Adams, C., Laukens, K., Donestein, P. C., and Bandeira, N. (2021) Open science resources for the mass spectrometry-based analysis of SARS-CoV-2. J. Proteome Res. 20, 1464–1475
13. Bittremieux, W., Meysman, P., Noble, W. S., and Laukens, K. (2018) Fast open modification spectral library searching through approximate nearest neighbor indexing. J. Proteome Res. 17, 3463–3474
14. Bittremieux, W., Laukens, K., and Noble, W. S. (2019) Extremely fast and accurate open modification spectral library searching of high-resolution mass spectra using feature hashing and graphics processing units. J. Proteome Res. 18, 3792–3799
15. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellaucherveu, D., and Nesvijzhitski, A. I. (2017) MSFagger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. Nat. Met. 14, 513–520
16. Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J., et al. (2019) The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucl. Acids Res 47, D442–D450
17. Hulstaert, N., Shostahl, J., Sachsenberg, T., Walzer, M., Barnea, H., Martens, L., et al. (2020) ThermoRawFileParser: modular, scalable, and cross-platform RAW file conversion. J. Proteome Res. 19, 537–542
18. Wang, M., Wang, J., Carver, J., Pullman, B. S., Cha, S. W., and Bandeira, N. (2018) Assembling the community-scale discoverable human proteome. *Cell Syst.* 7, 412–421.e5

19. Levitsky, L. I., Klein, J. A., Ivanov, M. V., and Gorkhkov, M. V. (2019) Pyleomics 4.0: five years of development of a Python proteomics framework. *J. Proteome Res.* 18, 709–714

20. Gessulat, S., Schmidt, T., Zolg, D. P., Samaras, P., Schnatbaum, K., Zerweck, J., et al. (2019) Prolong: proteome-wide prediction of peptide tandem mass spectra by deep learning. *Nat. Met.* 16, 509–518

21. Lam, H., Deutsch, E. W., Eddes, J. S., Eng, J. K., King, N., Stein, S. E., et al. (2007) Development and validation of a spectral library searching method for peptide identification from MS/MS Proteome. *7, 655–667*

22. Lam, H., Deutsch, E. W., and Aebersold, R. (2010) Artificial decoy spectral libraries for false discovery rate estimation in spectral library searching in proteomics. *J. Proteome Res.* 9, 605–610

23. Breuza, L., Poux, S., Estreichet, A., Famiglietti, M. L., Magrane, M., Tognolli, M., et al. (2016) The UniProtKB guide to the human proteome. *Database (Oxford)* 2016, bav120

24. Keller, A., Nesvizhskii, A. I., Kleiner, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *J. Proteome Res.* 1, 2988–2997

25. Teo, G., Liu, G., Zhang, J., Nesvizhskii, A. I., Gingras, A. C., and Choi, H. (2014) SambaExpress: improvements and additional features in significance analysis of InteRactome software. *J. Proteomics* 100, 37–43

26. Jäger, S., Cimermancic, P., Gulbahce, N., Johnson, J. R., Mcgovern, K. E., Clarke, S. C., et al. (2011) Global landscape of human proteome complexes. *Nature 481*, 365–370

27. [preprint] Samavarchi-Tehrani, P., Abdouni, H., Knight, J. D. R., Astori, A., et al. (2021) The global phosphorylation landscape of SARS-CoV-2 reveals a cell passage induced in-frame viral replication. *bioRxiv*, https://doi.org/10.1101/2021.08.28.269175

28. [preprint] Samavarchi-Tehrani, P., Abdouni, H., Knight, J. D. R., Astori, A., et al. (2021) A SARS-CoV-2 BioID-based virus-host membrane protein interactome and virus peptide compartmentalisation: new proteomics resources for COVID-19 research. *bioRxiv*. https://doi.org/10.1101.2020.03.28.99175

29. [preprint] Laurent, E. M. N., Sojifantos, Y., Komarova, A., Gimeno, J.-P., Tehrani, P. S., Kim, D.-K. et al. (2020) Global BioID-based SARS-CoV-2 proteins proximal interactome unveils novel ties between viral poly-peptides and host factors involved in multiple COVID19-associated mechanisms. *bioRxiv*. https://doi.org/10.1101.2020.08.27.272955

30. Oughtred, R., Rust, J., Chang, C., Breitkreutz, B. J., Stark, C., Willems, A., et al. (2021) The BioGRID database: a comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci. 30*, 187–200

31. Guirimard, T., Delmotte, S., and Navratil, V. (2015) VirusHostNet 2.0: surfing on the web of virus/host molecular interactions data. *Nucl. Acids Res* 43, D983–D987. Database issue

32. Creasy, D. M., and Cottrell, J. S. (2004) Unimod: protein modifications for mass spectrometry. *Proteomics* 4, 1534–1536

33. Bittremieux, W. (2020) Spectrum-utils: a Python package for mass spectrometry data processing and visualization. *Anal. Chem. 92*, 659–661

34. Klann, K., Bojkova, D., Tascher, G., Ciesek, S., Münch, C., and Cinat, J. (2020) Growth factor receptor signaling inhibition prevents SARS-CoV-2 replication. *Mol. Cell 80*, 164–174.e4

35. Davidson, A. D., Williamson, M. K., Lewis, S., Shoemark, D., Carroll, M. W., Hesom, K. J., et al. (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

36. Bouhaddou, M., Memon, D., Meyer, B., White, K. M., Rezvi, J. E., Correa Marrero, M., et al. (2020) The global phosphorylation landscape of SARS-CoV-2 infection. *Cell 182*, 685–712.e19

37. Blom, N., Gammeltoft, S., and Brunak, S. (1999) Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol. 294*, 1351–1362

38. Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S., and Brunak, S. (2004) Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics 4*, 1633–1649

39. Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S., and Brunak, S. (2004) Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics 4*, 1633–1649
62. Chathuranga, K., Weerawardhana, A., Dodantenna, N., and Lee, J. S. (2021) Regulation of antiviral innate immune signaling and viral evasion following viral genome sensing. *Exp. Mol. Med.* 53, 1647–1668

63. Hershko, A., and Ciechanover, A. (1998) The ubiquitin system. *Annu. Rev. Biochem.* 67, 425–479

64. Pickart, C. M. (2004) Back to the future with ubiquitin. *Cell* 116, 181–190

65. Lindner, H. A. (2007) Deubiquitination in virus infection. *Virology* 362, 245–256

66. Bres, V., Kiernan, R. E., Linares, L. K., Chable-Bessia, C., Plechakova, O., Tréand, C., *et al.* (2003) A non-proteolytic role for ubiquitin in Tat-mediated transactivation of the HIV-1 promoter. *Nat. Cell Biol.* 5, 754–761

67. Cao, Z., Xia, H., Rajsbaum, R., Xia, X., Wang, H., and Shi, P. Y. (2021) Ubiquitination of SARS-CoV-2 ORF7a promotes antagonism of interferon response. *Cell Mol. Immunol.* 18, 746–748

68. Guo, G., Gao, M., Gao, X., Zhu, B., Huang, J., Luo, K., *et al.* (2021) SARS-CoV-2 non-structural protein 13 (nsp13) hijacks host deubiquitinase USP13 and counteracts host antiviral immune response. *Signal. Transduct. Target. Ther.* 6, 119

69. Bhatta, V., Elsagary, L., and Dakshinamurti, S. (2021) Tracing the path of inhaled nitric oxide: biological consequences of protein nitrosylation. *Pediatr. Pulmonol.* 56, 525–538

70. Akaike, T. (2001) Role of free radicals in viral pathogenesis and mutation. *Rev. Med. Virol.* 11, 87–101

71. Ignarro, L. J. (2020) Inhaled NO and COVID-19. *Br. J. Pharmacol.* 177, 3848–3849

72. Meini, S., Fortini, A., Andreini, R., Sechi, L. A., and Tascini, C. (2021) The paradox of the low prevalence of current smokers among COVID-19 patients hospitalized in nonintensive care wards: results from an Italian multicenter case-control study. *Nicotine Tob. Res.* 23, 1436–1440

73. Hedenstierna, G., Chen, L., Hedenstierna, M., Lieberman, R., and Fine, D. H. (2020) Nitric oxide dosed in short bursts at high concentrations may protect against Covid 19. *Nitric Oxide* 103, 1–3

74. Swathi Krishna, S., Thennavan, A., and Kanthial, S. K. (2021) Dietary foods containing nitric oxide donors can be early curators of SARS-CoV-2 infection: a possible role in the immune system. *J. Food Biochem.* 46, e13884

75. Colasanti, M., Persichini, T., Venturini, G., and Ascenzi, P. (1999) S-Nitrosylation of viral proteins: molecular bases for antiviral effect of nitric oxide. *IUBMB Life* 48, 25–31

76. Mnatsakanyan, R., Markoutsa, S., Walbrunn, K., Roos, A., Verhelst, S. H. L., and Zahedi, R. P. (2019) Proteome-wide detection of S-nitrosylation targets and motifs using bioorthogonal cleavable-linker-based enrichment and switch technique. *Nat. Commun.* 10, 2195

77. Akaberi, D., Krambrich, J., Ling, J., Luni, C., Hedenstierna, G., Järhult, J. D., *et al.* (2020) Mitigation of the replication of SARS-CoV-2 by nitric oxide in vitro. *Redox Biol.* 37, 101734

78. Liu, M., Hou, J., Huang, L., Huang, X., Heibeck, T. H., Zhao, R., *et al.* (2010) Site-specific proteomics approach for study protein S-nitrosylation. *Anal. Chem.* 82, 7160–7168

79. Verma, N., Henderson, J. A., and Shen, J. (2020) Proton-coupled conformational activation of SARS coronavirus main proteases and opportunity for designing small-molecule broad-spectrum targeted covalent inhibitors. *J. Am. Chem. Soc.* 142, 21883–21890

80. Van Puyvelde, B., Van Uytvanck, K., Tytgat, O., Van Oudenhove, L., Gabriels, R., Bouwmeester, R., *et al.* (2021) Cov-MS: a community-based template Assay for mass-spectrometry-based protein detection in SARS-CoV-2 patients. *JACS Au* 1, 750–765

81. Deutsch, E. W., Csordas, A., Sun, Z., Jarnuczak, A., Perez-Riverol, Y., Ternent, T., *et al.* (2017) The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. *Nucl. Acids Res.* 45, D1100–D1106