Journal of proteome research

pubs.acs.org/jpr

Communication

Mass Spectrometric Identification of SARS-CoV-2 Proteins from Gargle Solution Samples of COVID-19 Patients

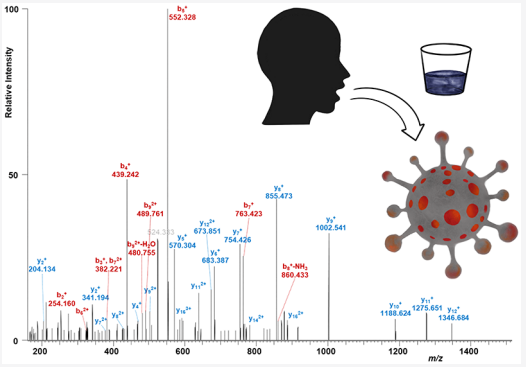
Christian Ihling, Dirk Tänzler, Sven Hagemann, Astrid Kehlen, Stefan Hüttelmaier, Christian Arlt, and Andrea Sinz*

Cite This: https://dx.doi.org/10.1021/acs.jproteome.0c00280

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Mass spectrometry (MS) can deliver valuable diagnostic data that complement genomic information and allow us to increase our current knowledge of the COVID-19 disease caused by the SARS-CoV-2 virus. We developed a simple, MS-based method to specifically detect SARS-CoV-2 proteins from gargle solution samples of COVID-19 patients. The protocol consists of an acetone precipitation and tryptic digestion of proteins contained within the gargle solution, followed by a targeted MS analysis. Our methodology identifies unique peptides originating from SARS-CoV-2 nucleoprotein. Building on these promising initial results, faster MS protocols can now be developed as routine diagnostic tools for COVID-19 patients. Data are available via ProteomeXchange with identifier PXD019423.



KEYWORDS: *COVID-19, mass spectrometry, nucleoprotein, SARS-CoV-2*

INTRODUCTION

The disease COVID-19, caused by the SARS-CoV-2 virus, was declared a pandemic by the World Health Organization on March 12, 2020[1] and has caused more than 348 000 fatalities worldwide (date: May 26, 2020). Providing novel mass-spectrometry (MS)-based diagnostic tools that complement genomic approaches is one of the major goals of the recently formed COVID-19 mass spectrometry coalition (www.covid19-msc.org).[2] On the basis of the prospective goals of this coalition, the aim of this "proof-of-principle" study is to highlight the potential of MS in identifying SARS-CoV-2 proteins, even from highly diluted samples, such as gargle solutions of COVID-19 patients.

EXPERIMENTAL SECTION

Protein Precipitation and Digestion

Gargle solutions (20 mL in 0.9% NaCl) were obtained from three patients with confirmed COVID-19 infection. All samples had been classified as SARS-CoV-2-positive by three reverse-transcription and quantitative polymerase chain reaction (RT-qPCR) analyses identifying E-, S-, and N-gene RNAs. For protein precipitation, 1 mL of acetone (−20 °C) was added to 750 μL of gargle solution and stored overnight at −20 °C. After centrifugation, the supernatant was discarded, and the pellet was dissolved in 10 μL of SMART Digest buffer (Thermo Fisher Scientific). The solution was kept at 75 °C for 10 min. Afterward, 1 μL of PNGase F (nonreducing, NEB) was added to remove protein *N*-glycosylations, and the sample was incubated at 50 °C for 10 min. One μg of trypsin (SMART Digest, bulk resin option, Thermo Fisher Scientific) was added to 50 μL of SMART Digest buffer (Thermo Fisher Scientific), and the solution was incubated at 70 °C for 30 min. Reduction and alkylation were performed with dithiothreitol (DTT) and iodoacetamide, respectively. Before LC/MS analysis, trifluoroacetic acid (TFA) was added to a final concentration of 0.5% (v/v).

Nano-HPLC/Nano-ESI-Orbitrap-MS/MS

LC−/MS analysis was performed on a nano-HPLC system (UltiMate RSLC 3000) coupled to an Orbitrap Fusion Tribrid mass spectrometer with nano-ESI source (Thermo Fisher Scientific). The samples were loaded onto a trapping column (Acclaim PepMap C18, 300 μm × 5 mm, 5 μm, 100 Å, Thermo Fisher Scientific) and washed for 15 min with 0.1% (v/v) TFA at a flow rate of 30 μL/min. Trapped peptides were eluted using a separation column (200 cm μPAC, C18, PharmaFluidics) that had been equilibrated with 3% B (A: 0.1% (v/v) formic acid in water; B: 0.08% (v/v) formic acid in acetonitrile). Peptides were separated with linear gradients from 3 to 10% B over 15 min, followed by 10% B to 30% B

Special Issue: Proteomics in Pandemic Disease

Received: April 29, 2020
Published: June 22, 2020

A

```
1   MSDNGPQNQR NAPRITFGGP SDSTGSNQNG ERSGARSKQR RPQGLPNNTA SWFTALTQHG KEDLKFPRGQ
71  GVPINTNSSP DDQIGYYRRA TRRIRGGDGK MKDLSPRWYF YYLGTGPEAG LPYGANKDGI IWVATEGALN
141 TPKDHIGTRN PANNAAIVLQ LPQGTTLPKG FYAEGSRGGS QASSRSSSRS RNSSRNSTPG SSRGTSPARM
211 AGNGGDAALA LLLLDRLNQL ESKMSGKGQQ QQGQTVTKKS AAEASKKPRQ KRTATKAYNV TQAFGRRGPE
281 QTQGNFGDQE LIRQGTDYKH WPQIAQFAPS ASAFFGMSRI GMEVTPSGTW LTYTGAIKLD DKDPNFKDQV
351 ILLNKHIDAY KTFPPTEPKK DKKKKADETQ ALPQRQKKQQ TVTLLPAADL DDFSKQLQQS MSSADSTQA
```

B

```
1   MSDNGPQNQR NAPRITFGGP SDSTGSNQNG ERSGARSKQR RPQGLPNNTA SWFTALTQHG KEDLKFPRGQ
71  GVPINTNSSP DDQIGYYRRA TRRIRGGDGK MKDLSPRWYF YYLGTGPEAG LPYGANKDGI IWVATEGALN
141 TPKDHIGTRN PANNAAIVLQ LPQGTTLPKG FYAEGSRGGS QASSRSSSRS RNSSRNSTPG SSRGTSPARM
211 AGNGGDAALA LLLLDRLNQL ESKMSGKGQQ QQGQTVTKKS AAEASKKPRQ KRTATKAYNV TQAFGRRGPE
281 QTQGNFGDQE LIRQGTDYKH WPQIAQFAPS ASAFFGMSRI GMEVTPSGTW LTYTGAIKLD DKDPNFKDQV
351 ILLNKHIDAY KTFPPTEPKK DKKKKADETQ ALPQRQKKQQ TVTLLPAADL DDFSKQLQQS MSSADSTQA
```

Figure 1. Sequence coverage of SARS-CoV-2 nucleoprotein in gargle samples (A, B) of two COVID-19 patients. Identified peptides are shown in green. The peptide (aa 41−61) identified in both samples is highlighted in gray.

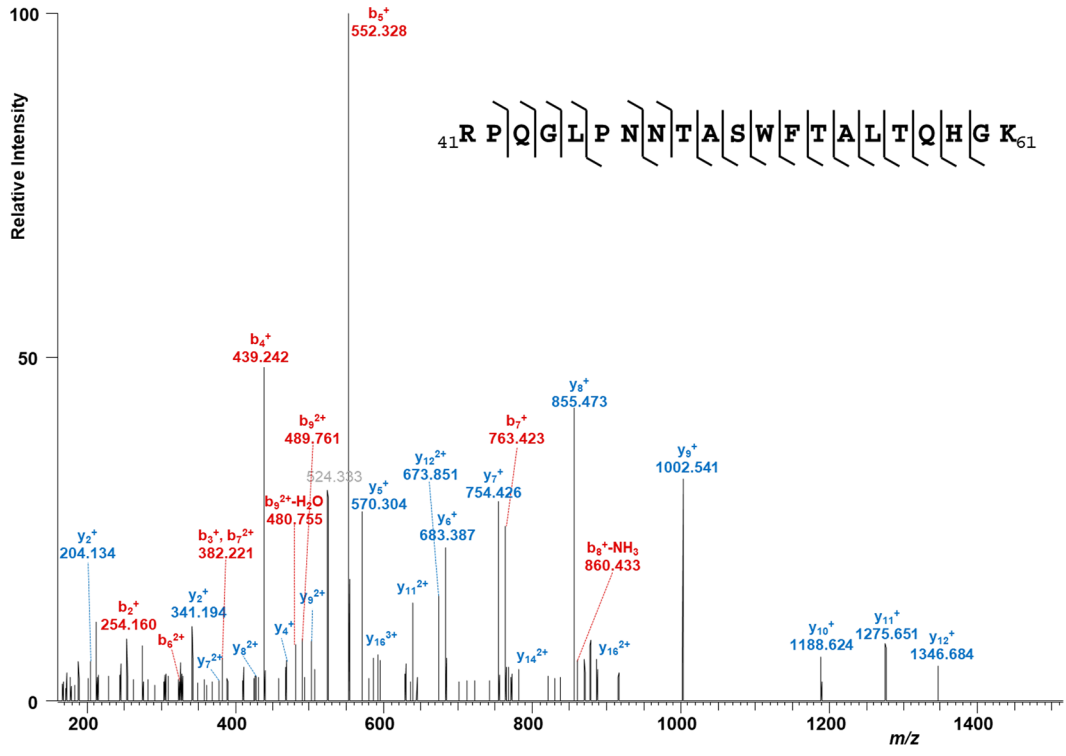


Figure 2. Fragment-ion mass spectrum (HCD-MS/MS) of the 4+ charged signal of peptide RPQGLPNNTASWFTALTQHGK (amino acids 41−61) from SARS-CoV-2 nucleoprotein.

https://dx.doi.org/10.1021/acs.jproteome.0c00280
J. Proteome Res. XXXX, XXX, XXX−XXX

over 165 min. The column was kept at 30 °C, and the flow rate was set to 300 nL/min. Targeted MS analysis was performed based on an inclusion list of tryptic peptides from SARS-CoV-2 proteins (Table S1).[3] The inclusion list was created with MeroX using the following settings: Tryptic cleavage at Lys and Arg, maximum of two missed cleavage sites, 5−30 amino acids per peptide, fixed modification: Cys carbamidomethylated, variable modifications: Met oxidized, Asn deamidated. Charge states 2−4 were considered for masses up to 3000, whereas charge states 2−5 were considered for masses between 3000 and 4000. In case none of the target masses from the inclusion list were detected, a TOP-5s method was applied. The resolving power in MS mode was set to 120 000 at m/z 200. A parallel data acquisition strategy was employed, relying on Orbitrap-HCD (higher energy collision-induced dissociation, R = 15 000) and linear ion trap-CID (collision-induced dissociation)-MS/MS experiments. A database search (UniProt human, dated November 2019 with 20 315 entries, and SARS-CoV-2[3]) was performed with Proteome Discoverer version 2.4 (Thermo Fisher Scientific). The search was restricted to human and SARS-COV-2 proteins. Two missed tryptic cleavages were allowed, and carbamidomethylation of Cys was set as a fixed modification, whereas Met oxidation and deamidation of Asp, due to PNGase F treatment, were set as variable modifications.

The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[4] partner repository with the data set identifier PXD019423 and 10.6019/PXD019423.

RESULTS AND DISCUSSION

Protein MS analysis of the three gargle samples from COVID-19 patients revealed a vast abundance of human proteins (Tables S2, S3, and S4). Strikingly, our protocol allowed us to identify unique peptides originating from SARS-CoV-2 nucleoprotein that were identified in two out of three samples (Figure 1). These results are in agreement with preceding polymerase chain reaction (PCR) analyses, exhibiting a higher viral load in the respective samples. We estimate the viral load to be on the order of 10^5 to 10^6 genome equivalents/μL gargle solution for the two samples, in which we detected SARS-CoV-2 nucleoprotein, whereas the MS-negative sample contains only ~10^3 genome equivalents/μL gargle solution.

A peptide originating from SARS-CoV-2 nucleoprotein, comprising the sequence RPQGLPNNTASWFTALTQHGK (amino acids 41−61, Figure 2), was found in both samples, giving the first hints regarding its value for establishing an MS-based COVID-19 diagnostic method. SARS-CoV-2 nucleoprotein packages the positive-strand viral genome RNA into a helical ribonucleocapsid and plays a fundamental role during virion assembly through its interactions with the viral genome and membrane protein M.[3] It is also important for enhancing the efficiency of subgenomic viral RNA transcription as well as viral replication.[3] The N-gene was found to be the most expressed viral gene,[5] which is in agreement with peptides of SARS-CoV-2 nucleoprotein being detected in our MS analysis. Clearly, further experiments are needed to test larger patient cohorts, but our findings show that SARS-CoV-2 proteins can be identified from highly diluted solutions. Our current method has the drawback of long analysis times (3 h per sample), requiring significant optimization for future high-throughput measurements of patient samples. Also, alternative MS instrumentation, such as triple−quadrupole instruments, will be required to allow the detection and quantification of a larger number of virus proteins.

CONCLUSIONS

We present a protein MS-based method to specifically detect SARS-CoV-2 virus proteins from highly diluted gargle solutions of COVID-19 patients. This straightforward method relies on an acetone precipitation of proteins, followed by a tryptic digestion and a targeted mass spectrometric analysis of SARS-CoV-2 proteins. Using this approach, we were able to identify peptides originating from SARS-CoV-2 nucleoprotein in gargle solution samples. We are planning to apply our protocol to additional sample material, such as bronchoalveolar lavage. The future goal is to provide robust, sensitive, and reliable MS-based methods as a routine diagnostic of COVID-19 patients that complement PCR-based methods.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00280.

Table S1: Inclusion list of tryptic peptides from SARS-CoV-2 proteins (XLSX)

Table S2: List of proteins identified in gargle solution sample 1 (XLSX)

Table S3: List of proteins identified in gargle solution sample 2 (XLSX)

Table S4: List of proteins identified in gargle solution sample 3 (XLSX)

AUTHOR INFORMATION

Corresponding Author

Andrea Sinz − *Department of Pharmaceutical Chemistry & Bioanalytics, Institute of Pharmacy, Martin Luther University Halle-Wittenberg, 06120 Halle, Germany;* orcid.org/0000-0003-1521-4899; Phone: +49-345-5525-170; Email: andrea.sinz@pharmazie.uni-halle.de

Authors

Christian Ihling − *Department of Pharmaceutical Chemistry & Bioanalytics, Institute of Pharmacy, Martin Luther University Halle-Wittenberg, 06120 Halle, Germany*

Dirk Tänzler − *Department of Pharmaceutical Chemistry & Bioanalytics, Institute of Pharmacy, Martin Luther University Halle-Wittenberg, 06120 Halle, Germany*

Sven Hagemann − *Department of Molecular Cell Biology, Institute for Molecular Medicine, Charles Tanford Center, Medical Faculty, Martin Luther University Halle-Wittenberg, 06120 Halle, Germany*

Astrid Kehlen − *Institute of Medicinal Microbiology, Biocenter, Medical Faculty, Martin Luther University Halle-Wittenberg, 06120 Halle, Germany*

Stefan Hüttelmaier − *Department of Molecular Cell Biology, Institute for Molecular Medicine, Charles Tanford Center, Medical Faculty, Martin Luther University Halle-Wittenberg, 06120 Halle, Germany*

Christian Arlt − *Department of Pharmaceutical Chemistry & Bioanalytics, Institute of Pharmacy, Martin Luther University Halle-Wittenberg, 06120 Halle, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.0c00280

https://dx.doi.org/10.1021/acs.jproteome.0c00280
J. Proteome Res. XXXX, XXX, XXX−XXX

Notes

The authors declare no competing financial interest.
The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[4] partner repository with the data set identifier PXD019423 and 10.6019/PXD019423.

■ ACKNOWLEDGMENTS

A.S. acknowledges financial support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), RTG 2467, project number 391498659 "Intrinsically Disordered Proteins − Molecular Principles, Cellular Functions, and Diseases". Prof. Gary Sawers is acknowledged for critical reading of the manuscript.

■ REFERENCES

(1) Yang, X.; Yu, Y.; Xu, J.; Shu, H.; Xia, J.; Liu, H.; Wu, Y.; Zhang, L.; Yu, Z.; Fang, M.; Yu, T.; Wang, Y.; Pan, S.; Zou, X.; Yuan, S.; Shang, Y. Clinical Course and Outcomes of Critically Ill Patients With SARS-CoV-2 Pneumonia in Wuhan, China: A Single-Centered, Retrospective, Observational Study. *Lancet Respir. Med.* **2020**, *8* (5), 475−481.

(2) Struwe, W.; Emmott, E.; Bailey, M.; Sharon, M.; Sinz, A.; Corrales, F. J.; Thalassinos, K.; Braybrook, J.; Mills, C.; Barran, P. COVID-19 MS Coalition (2020). The COVID-19 MS Coalition-accelerating Diagnostics, Prognostics, and Treatment. *Lancet* **2020**, *395* (10239), 1761−1762.

(3) SARS Coronavirus 2 - ViralZone Page. https://viralzone.expasy.org/8996, accessed April 6, 2020.

(4) Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; Pérez, E.; Uszkoreit, J.; Pfeuffer, J.; Sachsenberg, T.; Yilmaz, S.; Tiwary, S.; Cox, J.; Audain, E.; Walzer, M.; Jarnuczak, A. F.; Ternent, T.; Brazma, A.; Vizcaíno, J. A. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **2019**, *47* (D1), D442−D450 (PubMed ID: 30395289).

(5) Kim, D.; Lee, J. Y.; Yang, J. S.; Kim, J. W.; Kim, V. N.; Chang, H. Architecture of SARS-CoV-2 transcriptome. *BioRxiv* **2020**, DOI: 10.1101/2020.03.12.988865.

https://dx.doi.org/10.1021/acs.jproteome.0c00280
J. Proteome Res. XXXX, XXX, XXX−XXX