GDCh

**Research Articles**

*Angewandte*
*Chemie*
International Edition
www.angewandte.org

# Comprehensive Fragment Screening of the SARS-CoV-2 Proteome Explores Novel Chemical Space for Drug Development

*Hannes Berg[+], Maria A. Wirtz Martin[+], Nadide Altincekic[+], Islam Alshamleh,*
*Jasleen Kaur Bains, Julius Blechar, Betül Ceylan, Vanessa de Jesus,*
*Karthikeyan Dhamotharan, Christin Fuks, Santosh L. Gande, Bruno Hargittay,*
*Katharina F. Hohmann, Marie T. Hutchison, Sophie Marianne Korn, Robin Krishnathas,*
*Felicitas Kutz, Verena Linhard, Tobias Matzel, Nathalie Meiser, Anna Niesteruk,*
*Dennis J. Pyper, Linda Schulte, Sven Trucks, Kamal Azzaoui, Marcel J. J. Blommers,*
*Yojana Gadiya, Reagon Karki, Andrea Zaliani, Philip Gribbon, Marcius da Silva Almeida,*
*Cristiane Dinis Anobom, Anna L. Bula, Matthias Bütikofer, Ícaro Putinhon Caruso,*
*Isabella Caterina Felli, Andrea T. Da Poian, Gisele Cardoso de Amorim,*
*Nikolaos K. Fourkiotis, Angelo Gallo, Dhiman Ghosh, Francisco Gomes-Neto,*
*Oksana Gorbatyuk, Bing Hao, Vilius Kurauskas, Lauriane Lecoq, Yunfeng Li,*
*Nathane Cunha Mebus-Antunes, Miguel Mompeán, Thais Cristtina Neves-Martins,*
*Martí Ninot-Pedrosa, Anderson S. Pinheiro, Letizia Pontoriero, Yulia Pustovalova,*
*Roland Riek, Angus J. Robertson, Marie Jose Abi Saad, Miguel Á. Treviño,*
*Aikaterini C. Tsika, Fabio C. L. Almeida, Ad Bax, Katherine Henzler-Wildman,*
*Jeffrey C. Hoch, Kristaps Jaudzems, Douglas V. Laurents, Julien Orts, Roberta Pierattelli,*
*Georgios A. Spyroulias, Elke Duchardt-Ferner, Jan Ferner, Boris Fürtig, Martin Hengesbach,*
*Frank Löhr, Nusrat Qureshi, Christian Richter, Krishna Saxena, Andreas Schlundt,*
*Sridhar Sreeramulu, Anna Wacker, Julia E. Weigand, Julia Wirmer-Bartoschek,*
*Jens Wöhnert, and Harald Schwalbe\**

**Abstract:** SARS-CoV-2 (SCoV2) and its variants of concern pose serious challenges to the public health. The variants increased challenges to vaccines, thus necessitating for development of new intervention strategies including anti-virals. Within the international Covid19-NMR consortium, we have identified binders targeting the RNA genome of SCoV2. We established protocols for the production and NMR characterization of more than 80 % of all SCoV2 proteins. Here, we performed an NMR screening using a fragment library for binding to 25 SCoV2 proteins and identified hits also against previously unexplored SCoV2 proteins. Computational mapping was used to predict binding sites and identify functional moieties (chemotypes) of the ligands occupying these pockets. Striking consensus was observed between NMR-detected binding sites of the main protease and the computational procedure. Our investigation provides novel structural and chemical space for structure-based drug design against the SCoV2 proteome.

## Introduction

SARS-CoV-2 (SCoV2) is the cause for the COVID-19 pandemic resulting in more than 5 million deaths across the world and continues to pose serious challenges to public health and safety.[1] Countering the continuously evolving virus has not only seen an unprecedented success in the vaccine development but also given birth to several novel campaigns for anti-viral drug discovery,[2,3] including the recently approved oral antivirals paxlovid (Pfizer) and molnupiravir (Merck & Co.).[4–6]

The extensively mutated and highly infective variant of SCoV2, Omicron,[7] is resistant to several therapeutic antibodies,[8,9] evades double immunization,[8,10] and dominates the pandemic in 2022, calling for the development of new therapeutic strategies in combating the virus, specifically, by exploiting the conserved features.[11,12]

The SCoV2 genome consists of an ≈29.9 kb long positive-sense single-stranded RNA,[13] two-thirds of which comprises the open-reading frames (ORF) 1a and 1ab. Both ORFs encode polyproteins, which are proteolytically processed into 16 different non-structural proteins (nsp1-

[*] H. Berg,[+] M. A. Wirtz Martin,[+] N. Altincekic,[+] Dr. I. Alshamleh,
J. Kaur Bains, J. Blechar, B. Ceylan, V. de Jesus, C. Fuks,
Dr. S. L. Gande, B. Hargittay, K. F. Hohmann, M. T. Hutchison,
R. Krishnathas, F. Kutz, V. Linhard, T. Matzel, N. Meiser,
Dr. A. Niesteruk, Dr. D. J. Pyper, Dr. L. Schulte, Dr. S. Trucks,
Dr. J. Ferner, Dr. B. Fürtig, Dr. M. Hengesbach, Dr. N. Qureshi,
Dr. C. Richter, Dr. K. Saxena, Dr. S. Sreeramulu, Dr. A. Wacker,
Dr. J. Wirmer-Bartoschek, Dr. Prof. Dr. H. Schwalbe
Institute for Organic Chemistry and Chemical Biology, Goethe
University Frankfurt, 60438Frankfurt am Main (Germany)

H. Berg,[+] M. A. Wirtz Martin,[+] N. Altincekic,[+] Dr. I. Alshamleh,
J. Kaur Bains, B. Ceylan, V. de Jesus, K. Dhamotharan,
Dr. S. L. Gande, B. Hargittay, K. F. Hohmann, M. T. Hutchison,
Dr. S. Marianne Korn, R. Krishnathas, F. Kutz, V. Linhard, T. Matzel,
Dr. A. Niesteruk, Dr. D. J. Pyper, Dr. L. Schulte, E. Duchardt-Ferner,
Dr. J. Ferner, Dr. B. Fürtig, Dr. F. Löhr, Dr. N. Qureshi, Dr. C. Richter,
Dr. K. Saxena, Dr. A. Schlundt, Dr. S. Sreeramulu, Dr. A. Wacker,
Dr. J. Wirmer-Bartoschek, Prof. Dr. J. Wöhnert,
Dr. Prof. Dr. H. Schwalbe
Center of Biomolecular Magnetic Resonance (BMRZ), Goethe
University Frankfurt, Frankfurt am Main (Germany)
Max-von-LaueStrasse 7+9, 60438 Frankfurt am Main (Germany)
E-mail: schwalbe@nmr.uni-frankfurt.de

K. Dhamotharan, Dr. S. Marianne Korn, E. Duchardt-Ferner,
Dr. F. Löhr, Dr. A. Schlundt, Prof. Dr. J. Wöhnert
Institute for Molecular Biosciences, Goethe University Frankfurt,
Max-von-Laue Strasse 7+9, 60438 Frankfurt am Main (Germany)

Dr. K. Azzaoui, Dr. M. J. J. Blommers
Saverna Therapeutics, Pumpmattenweg 3, 4105 Biel-Benken (Switzerland)

Y. Gadiya, R. Karki, A. Zaliani, Dr. P. Gribbon
Fraunhofer Institute for Translational Medicine and Pharmacology
(ITMP), Screening Port, Schnackenburgallee 114, 22525 Hamburg
(Germany)

Y. Gadiya, R. Karki, A. Zaliani, Dr. P. Gribbon
Fraunhofer Cluster of Excellence for Immune-Mediated Diseases
(CIMD), Theodor-Stern-Kai 7, 60596 Frankfurt am Main (Germany)

M. da Silva Almeida, Í. Putinhon Caruso, A. T. Da Poian,
G. Cardoso de Amorim, N. Cunha Mebus-Antunes,
T. Cristtina Neves-Martins, F. C. L. Almeida
Institute of Medical Biochemistry, Federal University of Rio de
Janeiro, 21941-902 Rio de Janeiro (Brazil)

M. da Silva Almeida
Fraunhofer Cluster of Excellence for Immune-Mediated Diseases
(CIMD), Theodor-Stern-Kai 7, 60596 Frankfurt am Main (Germany)

C. Dinis Anobom, F. Gomes-Neto*, F. C. L. Almeida
National Center of Nuclear Magnetic Resonance (CNRMN),
CENABIO, Federal University of Rio de Janeiro, 21941-902Rio de
Janeiro (Brazil)

C. Dinis Anobom, A. S. Pinheiro
Department of Biochemistry, Institute of Chemistry, Federal
University of Rio de Janeiro, 21941-902Rio de Janeiro (Brazil)

A. L. Bula, K. Jaudzems
Latvian Institute of Organic Synthesis, Aizkraukles 21, LV-1006 Riga
(Latvia)

M. Bütikofer, D. Ghosh, R. Riek
ETH, Swiss Federal Institute of Technology, Laboratory of Physical
Chemistry, HCI F217, Vladimir-Prelog-Weg 2, 8093 Zürich (Switzerland)

Í. Putinhon Caruso
Multiuser Center for Biomolecular Innovation (CMIB), Department
of Physics, São Paulo State University (UNESP), 01049-010São
José do Rio Preto (Brazil)

I. Caterina Felli, L. Pontoriero, R. Pierattelli
Magnetic Resonance Center (CERM), University of Florence, Via
Luigi Sacconi 6, Sesto Fiorentino, 50019, Florence (Italy)

I. Caterina Felli, L. Pontoriero, R. Pierattelli
Department of Chemistry "Ugo Schiff", University of Florence, Via
della Lastruccia 3-13, Sesto Fiorentino, 50019, Florence (Italy)

G. Cardoso de Amorim
Multidisciplinary Center for Research in Biology (NUMPEX),
Campus Duque de Caxias Federal University of Rio de Janeiro,
25.250-470 Duque de Caxias (Brazil)

N. K. Fourkiotis, A. Gallo, A. C. Tsika, G. A. Spyroulias
Department of Pharmacy, University of Patras, 26504Patras
(Greece)

A. Gallo
Department of Chemistry, University of Torino IT−10126 Torino
(Italy)

F. Gomes-Neto*
Laboratory of Toxinology, Oswaldo Cruz Foundation (FIOCRUZ),
21040-900Rio de Janeiro (Brazil)

O. Gorbatyuk, B. Hao, Y. Li, Y. Pustovalova, J. C. Hoch
Department of Molecular Biology and Biophysics UConn Health
263 Farmington Ave., Farmington, CT 06030-3305 (USA)

V. Kurauskas, K. Henzler-Wildman
Department of Biochemistry, University of Wisconsin-Madison,
Madison, WI 53706 (USA)

L. Lecoq, M. Ninot-Pedrosa
Molecular Microbiology and Structural Biochemistry, UMR5086
CNRS/Université Lyon 1, 7, passage du Vercors, 69367 Lyon
(France)

M. Mompeán, M. Á. Treviño, D. V. Laurents
"Rocasolano" Institute for Physical Chemistry, CSIC 28006Madrid-
(Spain)

A. J. Robertson, A. Bax
Laboratory of Chemical Physics, National Institute of Diabetes and
Digestive and Kidney
20892-0520 Bethesda (USA)

M. Jose Abi Saad, J. Orts
University of Vienna, Department of Pharmaceutical Sciences,
Josef-Holaubek-Platz 2, A-1090 Vienna (Austria)

Dr. J. E. Weigand
Department of Biology, Technical University of Darmstadt, 64289
Darmstadt (Germany)

[+] These authors contributed equally to this work.

nsp16).[14,15] Four structural proteins: spike (S), envelope (E), membrane (M) and nucleocapsid (N) and nine additional accessory factors are expressed from the 13 ORFs located at the 3′ end of the viral genome. In total, the viral genome encodes for at least 28 peptides or proteins.[16–18] Repurposing of (approved) drugs has been actively pursued as a strategy to counter SCoV2 infections,[19–22] however, with little clinical success.[23] Most of repurposed drugs were primarily an outcome of structure-based virtual screening campaigns and solely focused on a small fraction of the proteome, namely proteases (nsp3d, nsp5) or polymerase (nsp12) as targets.[24–30] Within the viral life cycle, the enzymes nsp3 (papain-like protease), nsp5 (main protease), nsp7·nsp8 (primase complex), nsp12 (primary RNA-dependent RNA polymerase (RdRp)), nsp13 (helicase), nsp14 (exoribonuclease) and the methyltransferases nsp14/nsp16 are important components of the replicase-transcriptase complex and hence are also listed as attractive drug targets.[16,31] X-ray crystallography and NMR have been successfully used to screen either fragments, approved drugs, or drugs in clinical trials, against a subset of key SCoV2 protein drug targets like nsp5, nsp3b, nsp13 and nsp14.[32–40]

The current drug development has typically focused its efforts around the two key viral proteins, a protease (nsp5) and a polymerase (nsp12, RdRp), and soon such a monotherapy can result in the virus developing resistance against the first-generation antivirals, thus warranting us to develop new antivirals involving different targets.[41] Recently, using a range of biochemical assays, several drugs were identified as inhibitors against a total of seven enzymes of SCoV2.[42–49] Therefore, developing drugs or synergistic combinations involving multiple viral targets appears as a viable therapeutic strategy for the treatment of COVID-19.[2,3,50]

Within the Covid19-NMR consortium, we undertook a massive NMR-based ligand screening with the aim of identifying fragments as new chemical entities targeting SCoV2 proteins. Previously, via consorted efforts between NMR groups worldwide we have successfully developed protocols for large-scale production of more than 80 % of all SCoV2 proteins.[51] Soon, the availability of proteins and the experience gained from the completion of >20 screens with the DSI-PL fragment library for binding against the viral RNA[52] positioned us to embark on this massive screening campaign. For this purpose, >20 SCoV2 proteins (nsp1, nsp2 (CtDR), nsp3a, nsp3b, nsp3b·GS-441524, nsp3c (SUD-N), nsp3c (SUD-MC), nsp3d, nsp3e, nsp3Y, nsp5, $_{GHM}$nsp5, $_{GS}$nsp5, nsp7, nsp8, nsp9, nsp10, nsp10·nsp14, $_{His6}$nsp15, nsp10·nsp16, ORF9a (IDR1-NTD-IDR2), ORF9a (NTD), ORF9a (NTD-SR), ORF9a (CTD), ORF9b; (for definitions see Supporting Information Table 1) were produced in NMR groups at sites all over the world and subsequently shipped to the Frankfurt NMR center (BMRZ) for conducting the NMR screening. We applied ligand-observed $^1$H NMR experiments and identified 311 binders across the 25 screened SCoV2 proteins. Further, we used FTMap,[53] a computational mapping server which has been proven to be more accurate than the conventional GRID and MCSS methods to identify binding sites (or hot spots) on macromolecules (protein, DNA or RNA). Active sites in enzymes are usually concave surfaces that are suitable for ligand binding and therefore, in our study, binding site, hot spot, and active site are used interchangeably. FTMap predicts chemical scaffolds and functional units occupying these binding pockets. A comparison of the predicted scaffolds and functional units with the constitution of the experimental fragment hits for which we detected binding in our experimental screens showed striking correlation, as exemplified by comparing predicted and experimentally determined binding pockets for the main protease nsp5, the latter obtained both from crystallographic screens[54] as well as NMR protein-based screens conducted here. We thus propose this novel methodology for the analysis of ligand binding capability across multiple protein targets as provided in this work. Such methodology bears excellent potential to act as a unique resource for developing novel inhibitors.

## Results and Discussion

We conducted fragment-based screenings for a large number of SCoV2 viral proteins (Table 1 and Supporting Information Table 1). The viral proteins can be classified broadly into three different classes, namely, (i) proteases, (ii) replicase-transcriptase (RT) complex proteins and (iii) other accessory proteins. The main protease (nsp5, Mpro, CLpro) and the Papain-like protease (nsp3d, PLpro) are two important viral proteases that play a functionally important role in viral maturation.[55,56] Nsp5 is responsible for the cleavage of 12 nsps (nsp4-nsp16) and therefore represents one of the most attractive drug targets. We screened three different constructs (nsp5, $_{GS}$nsp5 and $_{GHM}$nsp5) of nsp5. The two ($_{GS}$nsp5) or three ($_{GHM}$nsp5) additional amino acids in the N-terminus resulted from cloning. SEC-MALS analysis of these two proteins revealed that they are monomeric in solution compared to the dimeric wildtype nsp5.[51] Recently, it has been shown that the monomer-dimer equilibrium is coupled to the catalytic activity of nsp5, with maximum activity associated with the dimeric state.[57] Therefore, identifying small molecules that interfere with the dimer formation is considered as an alternative strategy to impair catalytic activity[58] and so screening of both monomeric and dimeric states of the proteins may act as a valuable tool in identifying and developing allosteric ligands. Nsp3d is responsible for the cleavage of the N-terminus of the polyprotein, releasing nsp1, nsp2 and nsp3 and is therefore also a potential drug target. The RT-complex is composed of multiple enzymes, and we screened the SCoV2 putative primases (nsp7 and nsp8) and the methyltransferases (nsp14 and nsp16) in complex with its co-factor nsp10 (nsp10·nsp14, nsp10·nsp16). The other screened set of proteins included several nsps, various domain constructs of nsp3 and structural and accessory proteins (ORF9a (N-protein) and ORF9b). The molecular weight of the screened proteins ranged between 5 kDa (nsp2 (CtDR)) to 78 kDa (nsp10·nsp14). Further, the 25 screened proteins also included intrinsically disordered proteins (nsp2 (CtDR)), proteins with intrinsically disordered regions (N-protein),

***Table 1:*** SCoV2 protein constructs screened by NMR.

| Protein genome position (nt)[a] | Trivial name Construct expressed | Size (aa)[b] | Boundaries | MW [kDa] | PDB code used for FTMap | Number of binders identified | Crossclusters in Cleft1 | Crossclusters in Cleft2 |
|---|---|---|---|---|---|---|---|---|
| nsp1 266–805 | *Leader* | 180 | | 19.8 | | | | |
| | Globular Domain (GD) | 116 | 13–127 | 12.7 | 7k7p | 5 | 0, 7 | 1, 2, 3 |
| nsp2 806–2,719 | | 638 | | 70.5 | | | | |
| | C-terminal IDR (CtDR) | 45 | 557–601 | 4.9 | - | 19 | – | – |
| nsp3 2,720–8,554 | | 1,945 | | 217.3 | | | | |
| a | Ub-like (UBl) domain | 111 | 1–111 | 12.4 | 7kag | 14 | 3, 6 | 0, 5, 8, 10 |
| b | nsp3b (Macro domain) | 170 | 207–376 | 18.3 | 6vxs | 10 | 0, 1, 2, 3, 5 | – |
| b | nsp3b·GS-441524 | 170 | 207–376 | 18.3 | 6vxs | 5 | – | – |
| c | SUD-N | 140 | 409–548 | 15.4 | 2w2 g | 10 | 0, 2, 4, 5, 6, 9 | – |
| c | SUD-MC | 193 | 551–743 | 21.5 | 2kqv | 154 | 1, 2, 3, 4, 6 | 0 |
| d | Papain-like protease PL^pro | 318 | 743–1,060 | 36 | 6w9c | 150 | 5, 7 | 1, 2, 4 |
| e | NAB | 116 | 1,088–1,203 | 13.4 | 2k87 | 21 | 1, 4 (Cleft 3) | - |
| Y | | 286 | | 31.5 | | 81 | – | – |
| nsp5 10,055–10,972 | *Main protease (M^pro)* | 306 | | 33.8 | | | | |
| | GS nsp5 | 306 | 1–306 | 33.8 | – | 12 | – | – |
| | GHM nsp5 | 306 | 1–306 | 33.8 | - | 38 | – | – |
| | Full-length | 306 | 1–306 | 33.8 | 5r83 | 78 | 3, 4, 6 | 1, 2, 7, 8 |
| nsp7 11,843–12,091 | | 83 | | 9.2 | | | | |
| | Full-length | 83 | 1–83 | 9.2 | 2kys | 92 | 0, 1, 3, 6 | - |
| nsp8 12,092–12,685 | | 198 | | 21.9 | | | | |
| | Full-length | 198 | 1–198 | 21.9 | 6wiq | 35 | 1, 3, 4, 5, 6 | – |
| nsp9 12,686–13,024 | | 113 | | 12.4 | | | | |
| | Full-length | 113 | 1–113 | 12.4 | 6w4b | 2 | 1, 3 | 0, 2, 4 |
| nsp10 13,025–13,441 | | 139 | | 14.8 | | | | |
| | Full-length | 139 | 1–139 | 14.8 | 6zpe | 38 | 0, 3, 5, 6 | – |
| nsp15 19,621–20,658 | *Endonuclease* | 346 | | 38.8 | | | | |
| | His6 nsp15 | 346 | 1–346 | 38.8 | 6w01 | 42 | 1, 2 | 4 |
| nsp10·nsp16 20,659–21,552 | *Methyltransferase* | 298 | | 33.3 | | | | |
| | nsp10·nsp16 | 298 | 1–298 (nsp16) | 33.3 | 6w4 h | 92 | 3, 4, 5, 7, 8, 10 | 0 |
| nsp10·nsp14 18,040–19,620 | *Exoribonuclease* | 527 | | 61.4 | | | | |
| | nsp10·nsp14 | 527 | 7–527 (nsp14) | 61.4 | modelled | 44 | 2, 5, 9 (Cleft 3) | – |
| ORF9a 28,274–29,533 | *Nucleocapsid (N)* | 419 | | 45.6 | | | | |
| | IDR1-NTD-IDR2 | 248 | 1–248 | 26.5 | 6yi3 | 7 | – | – |

***Table 1:*** (Continued)

|  |  |  |  |  |  |  |  |  |
|---|---|---|---|---|---|---|---|---|
|  | NTD-SR | 169 | 44–212 | 18.1 | 6yi3 | 5 | – | – |
|  | NTD | 136 | 44–180 | 14.9 | 6yi3 | 32 | 0, 1, 3, 5, 6, 7 | 2 |
|  | CTD | 118 | 247–364 | 13.3 | 7c22 | 9 | 1, 2, 6, 8 | – |
| ORF9b *28,284–28,574* |  | 97 |  | 10.8 |  |  |  |  |
|  | Full-length | 97 | 1–97 | 10.8 | 6z4u | 8 | 0, 3, 5 (Cleft 3) | – |

[a] Genome position in nucleotide (nt) corresponding to SCoV2 NCBI reference genome entry NC_045512.2, identical to GenBank entry MN908947.3. [b] number of amino acids excluding the additional residues due to cloning

and even a protein-inhibitor complex (nsp3b·GS-441524) with the quest to identify ligands binding in close proximity to the nucleotide binding pocket as starting point for fragment growth medicinal chemistry.

The DSI-poised library (DSI-PL, Supporting Information, excel sheet 1 DSI PL Poised Library.xlsx)[59–61] has already been successfully used to screen the druggability of the RNA regulatory elements and the main protease nsp5 from SCoV2.[52,54] This library is composed of 768 highly diverse and poised fragments specifically designed to



***Figure 1.*** NMR based identification of binding fragments. A) Schematic representation of all NMR experiments used in the screening that show exemplary effects indicating binding events in the presence of ligand compared to ligand free spectra. B) NMR spectra (1D $^1$H, wLOGSY, STD, and $T_2$-CPMG (5 ms and 100 ms) and chemical structure (binder 20 and binder 3) of two binding fragments identified for nsp7. Single fragment spectra (top) are used for chemical shift deconvolution in the mixture. Binder 20 shows clear sign changes in the STD and wLOGSY in presence of nsp7 protein. Binder 3 also shows signal in the STD and a sign change in the wLOGSY, as well as a $T_2$ reduction of approximately 50% in presence of nsp7 protein.

facilitate easy downstream synthesis. We applied ligand-observed $^1$H NMR experiments and performed the screening with 64 mixes containing 12 fragments each as described previously.[52] In these screening experiments, changes in the $^1$H signals of the ligand in the presence and absence of the protein served as readout for binding.

For identifying binders within the mixtures, we first compared spectra from four different NMR experiments and analyzed differences by visual inspection. As criteria, chemical shift perturbations (CSPs) or severe line broadening, sign change in the waterLOGSY (wLOGSY), STD signal or significant decrease of signal intensity in a $T_2$-relaxation experiment were used to identify binders (Figure 1A). Ligands were assigned as a binder if one of the four criteria was satisfied. For example, binder 20 qualifies as a binder, showing changes in wLOGSY and STD, while only minor CSP and change in $T_2$ (Figure 1B, left). Similarly, binder 3 qualifies as a hit, displaying changes in wLOGSY, STD and $T_2$, but no CSP (Figure 1B, right).

NMR-based screening resulted in 311 binders across the 25 screened SCoV2 proteins (Figure 2). Our results show that the overall binders identified against a target ranged from 2 (nsp9) to 154 (nsp3c (SUD-MC)). No correlation was observed between the molecular weight of the target and the number of binders (Supporting Information Figure 1). Strikingly, the intrinsically disordered domain of nsp2 (CtDR) shows 19 binders. By contrast, the well folded protein nsp3b has only 3 binders. The protease nsp3d and the nsp3c (SUD) as a didomain with its middle and C-terminus (MC), are amongst those with the largest number of binders (Supporting Information Table 2). The nsp3b (macro domain) is evolutionarily conserved and regarded as
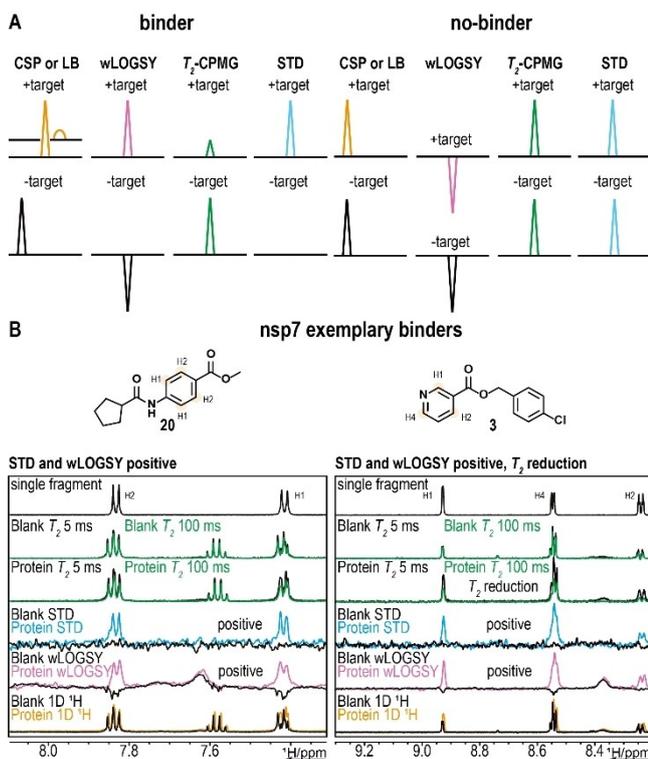
***Table 2:*** Affinities of the SCoV2 protein binders.

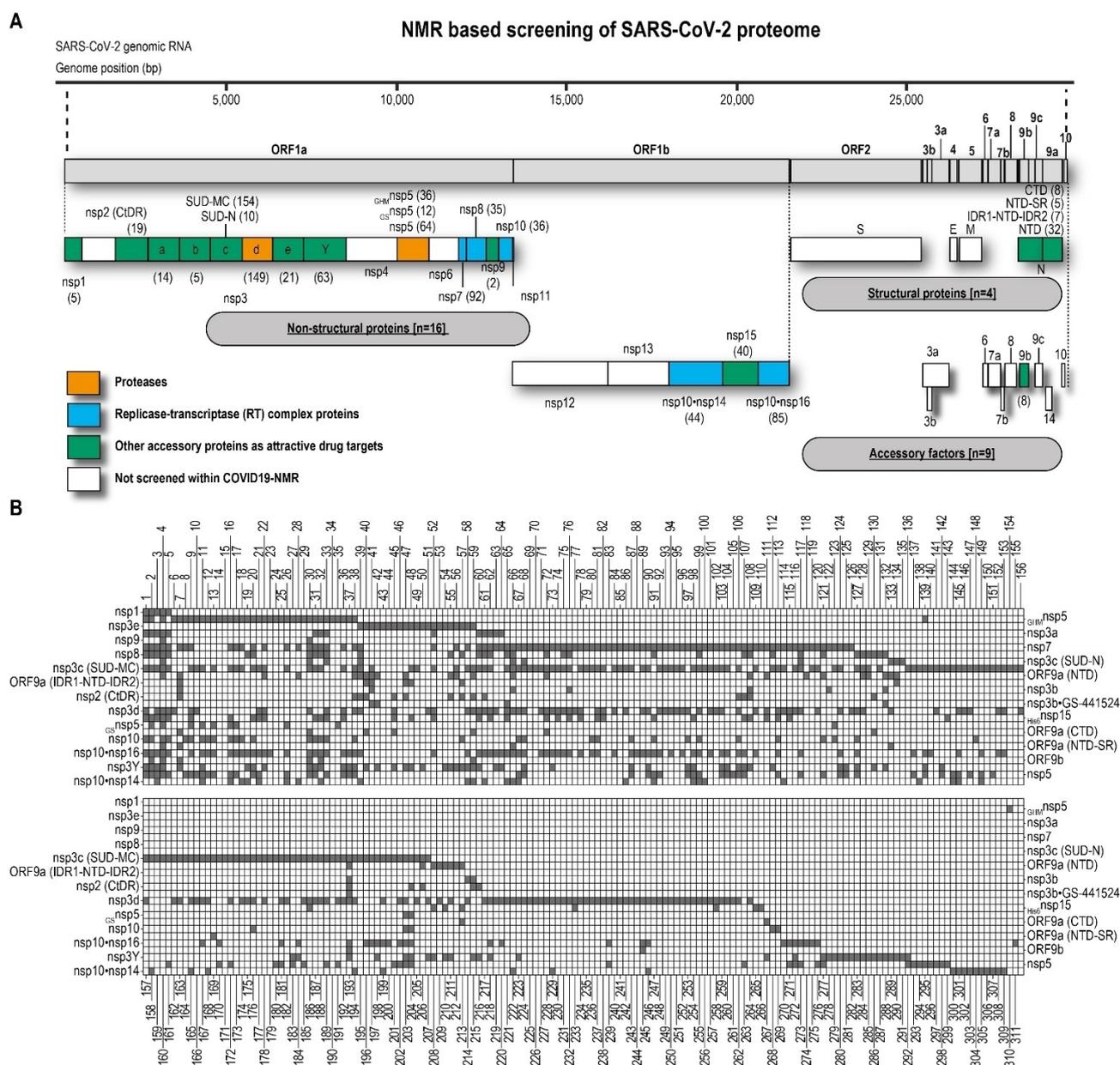| Ligand observed | | | | Protein Observed | |
|---|---|---|---|---|---|
| Binder | ORF9a (NTD)[a] | nsp3c (SUD-MC)[a] | nsp5[a] | Binder nsp5[a] | nsp10[a] |
| 40 | >5 | – | – | 21 | 0.46 ± 0.04 |
| 129 | >5 | – | – | 32 | – 0.44 ± 0.05 |
| 209 | >5 | – | – | 2 | – 1.70 ± 0.54 |
| 68 | – | 0.45 ± 0.71 | – | | |
| 30 | – | >5 | – | | |
| 13 | – | – | 0.02 ± 0.007 | | |
| 26 | – | – | >5 | | |

[a] $K_D$ in millimolar.

**Figure 2.** 311 binding fragments identified for SCoV2 proteins from NMR based fragment screening. A) Schematic representation of the SCoV2 genome (adapted from[16]). B) The two tables summarize all binding fragments identified in the NMR screening for their corresponding protein (grey). The first table shows binder 1 to 156 (columns) and the corresponding bound proteins (right and left rows). The second table shows binder 157 to 311 and the corresponding proteins (left and right rows).

a potential drug target. We conducted screening in its apo/free state and in the presence of GS-441524, the active drug and metabolite of remdesivir. We observed one common binder (binder 41) and two and four unique binders, respectively (Supporting Information Table 2). The main protease nsp5 is a dimeric cysteine protease and its N-terminus forms a part of the dimer interface. Subtle changes in the amino acid sequence at the N-terminus influence the oligomeric state ($_{GS}$nsp5 and $_{GHM}$nsp5, monomeric; nsp5, dimeric) of the protein.[51] For the three (nsp5, $_{GS}$nsp5 and $_{GHM}$nsp5) screened constructs we identified 78, 12, and 38 binders, respectively. Only 8 binders overlapped (Supporting Information Figure 2) between the three constructs,

suggesting that indeed there are differential surfaces exposed for ligand binding, which in turn stems from the monomer/dimer state of the protein constructs.[51] Previously, using the DSI-PL, nsp5 and nsp14 have been screened by crystallography identifying 39[54] and 41[38] binders, respectively. In contrast, 78 binders were identified by NMR for the identical construct of nsp5, and for a subset of these identified binders crystallization could be reproduced in house.

A comparison of the binders revealed 6 common binders including two 3-aminopyrimidine-like compounds (21 and 26) that form the chemical starting points within the COVID moonshot initiative.[40] The twice as large number of binders

identified by NMR is potentially attributed either to the presence of multiple stable conformations of nsp5 in solution[62] or to the fact that the different NMR-based screening experiments can identify binders within different affinity regimes (low micromolar to high millimolar). For nsp10·nsp14, we identified 44 binders with only one binder (binder 168) overlapping with the X-ray hits, wherein the screening was performed in the absence of nsp10. Further, 7 overlapping binders were found between nsp10·nsp14 and nsp10 NMR screens (Supporting Information Figure 2). Given the fact that significant conformational differences exist between nsp14 and nsp10·nsp14 structures,[38] it is not surprising that different sets of binders are identified in X-ray and NMR screens. Further, NMR competition experiments with sinefungin, a methyltransferase inhibitor and structural analog of s-adenosyl methionine (SAM), identified that binder 141 and 146 bind to the SAM binding site.

The relatively diverse and varying number of binders across the screened SCoV2 proteins in this work is likely correlated to the accessible surface of a given protein. In general, proteins that routinely bind to either small molecules or substrates to perform their function have well-defined cavities and pockets. For example, the cysteine protease nsp5 and the nsp3b (macro domain) each have a substrate or endogenous ligand binding cleft that both are currently exploited for designing functional inhibitors. Traditionally, ligand binding pockets in proteins are deter-

mined experimentally either by X-ray crystallography or NMR. Such experimental identification of binding pockets for large sets of binders across several targets of SCoV2 reported here would be very time-consuming and sample intensive. Thus, we deduced the ligand binding sites of the SCoV2 proteins using FTMap.[53] FTMap uses 16 small organic molecules (Supporting Information Figure 3) as probes to scan the surface of the protein target and to identify regions that bind multiple of these probes, thus forming a probecluster. Several probeclusters which are in close proximity on the protein surface form one crosscluster, thus defining a consensus site or hot spot. We performed the FTMap analysis for the 18 of the 25 screened proteins for which structural coordinates were available (Supporting Information Table 3). Except for nsp3e, the pdb structures for all proteins were from SCoV2. Further, for structures with multiple chains but with the same sequence (for example: dimer) the FTMap protocol recommends each chain to be independently mapped and therefore a single monomer unit was used for all the proteins except for nsp5, ORF9a (CTD) and ORF9b that is known to exist as a stable dimer in solution and both monomeric and dimeric state were analyzed. Typically, one to three binding sites (Supporting Information Figure 4 to 21) were identified for each of the proteins. For example, the binding sites in monomeric nsp5 clustered mainly around three distinct
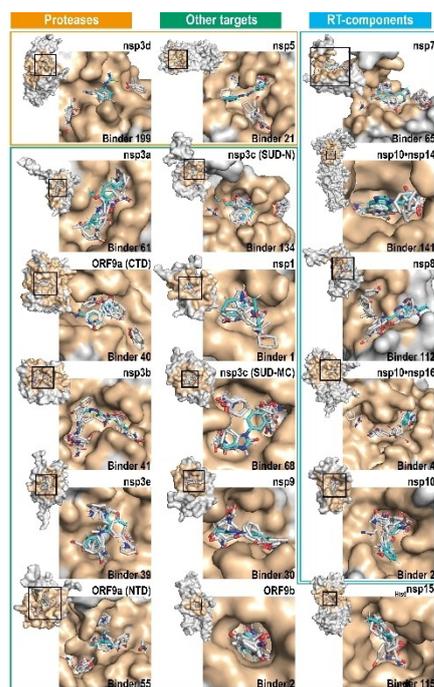


**Figure 3.** Hot spots identified using FTMap along with the docking of the NMR identified binders for 18 SCoV2 proteins. Proteases are highlighted with an orange box, RT-components with a blue box, and other targets with a green box. Zoom-ins show one of the identified clefts (beige colored) from PDBSum with its corresponding hot spots (and probes in grey sticks) from the FTMap analysis. For each of the targets, one of the binders was docked using SwissDock, shown in cyan.
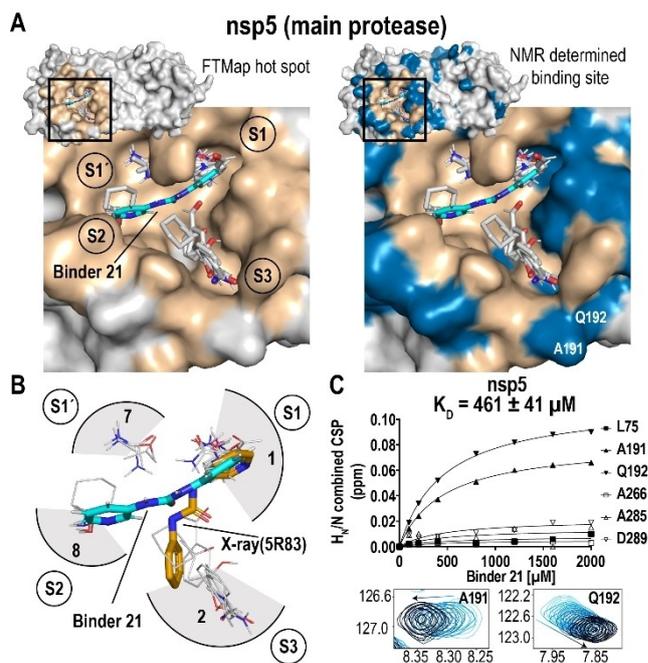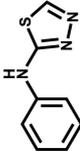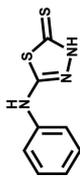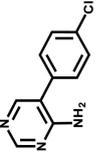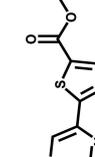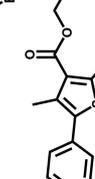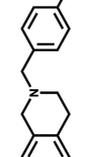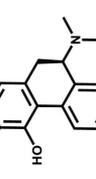


**Figure 4.** Agreement between bioinformatic and experimental mapping of the binding site. A) The FTMap identified hot spot for nsp5. The subsites of the active site are labeled as S1, S1′, S2 and S3. The crossclusters (1, 2, 7, and 8) occupying the binding site are shown in grey sticks. The docked pose of binder 21 is shown in cyan. Mapping of the CSPs (in blue) on to the structure of nsp5. B) Active site of nsp5 with an overlay of a docking (cyan) and X-ray determined (orange) structure of binder 21. C) The interaction of binder 21 and nsp5 was monitored via NMR titration. Binder 21 binds to nsp5 with a $K_D$ of 461 μM. The inset shows two shifting peaks (A191 and Q192) with increasing concentration of binder 21 (light blue-low to black-high).

**Table 3:** Fragment hits from NMR-based screening and related analogues identified as biologically active compounds in SCoV2 related assays in public databases.

| Binder | Structure | Binding targets detected by NMR | Bioactive analogue | CHEMBL Compound ID | Name | Tanimoto score | IC$_{50}$ in SCoV2 related assay (nM) | SCoV2 related BioAssay-ID | Assay ChEMBL ID | Assay Description |
|---|---|---|---|---|---|---|---|---|---|---|
| 67 | | nsp7, nsp3c (SUD-MC), nsp3d, nsp10, nsp10·nsp16, nsp5 | | CHEMBL289356 | CL-17107 | 0.88 | 390 | | CHEMBL4495583 | Biochemical, nsp5 (SCoV2 3CL-Pro protease inhibition) IC$_{50}$ FRET format with a peptide substrate |
| 74 | | nsp7, nsp3c (SUD-MC), nsp3d, nsp3d_His6, nsp15, nsp10·nsp16 | | CHEMBL264373 | Metoprine | 0.85 | 2340 | | CHEMBL4513083 | Cell based, SCoV2 induced cytotoxicity of VERO-E6 cells after 48 hours exposure to 0.01 MOI SCoV2 virus by high content imaging |
| 79 | | nsp7, nsp3c (SUD-MC) | | CHEMBL210545 | Oxyclozanide | 0.82 | 3710 | | CHEMBL4303812 | Cell based, Antiviral activity against SCoV2 (viral titer) measured by plaque assay in Vero cells at MOI 0.0125 after 24 hr |
| 150 | | nsp3c (SUD-MC) | | CHEMBL1382627 | silver-sulfadiazine | 0.81 | 750 | | CHEMBL4495583 | Biochemical, nsp5 (SCoV2 3CL-Pro protease inhibition) IC$_{50}$ FRET format with a peptide substrate |
| 50 | | nsp3e, nsp3y | | CHEMBL1380480 | VANITIOLIDE | 0.74 | 1320 | | CHEMBL4495583 | Biochemical, nsp5 (SCoV2 3CL-Pro protease inhibition) IC$_{50}$ FRET format with a peptide substrate |
| 6 | | nsp5, nsp3d | | CHEMBL226652 | 4-DAMP | 0.72 | 2360 | | CHEMBL4495583 | Biochemical, nsp5 (SCoV2 3CL-Pro protease inhibition) IC$_{50}$ FRET format with a peptide substrate |
| 157 | | nsp3c (SUD-MC), nsp3d | | CHEMBL243652 | PD096194 | 0.72 | 2040 | | CHEMBL4495583 | Biochemical, nsp5 (SCoV2 3CL-Pro protease inhibition) IC$_{50}$ FRET format with a peptide substrate |
| 37 | | nsp5, nsp10·nsp16, nsp3y, nsp5 | | CHEMBL566136 | PD121351 | 0.71 | 3190 | | CHEMBL4495583 | Biochemical, nsp5 (SCoV2 3CL-Pro protease inhibition) IC$_{50}$ FRET format with a peptide substrate |
| 56 | | nsp3e, nsp7, nsp3c (SUD-MC), ORF9a (NTD), nsp3y | | CHEMBL1616 | APOMORPHINE HYDROCHLORIDE | 0.71 | 520 | | CHEMBL4495583 | Biochemical, nsp5 (SCoV2 3CL-Pro protease inhibition) IC$_{50}$ FRET format with a peptide substrate |
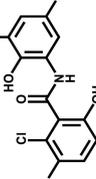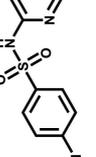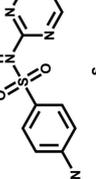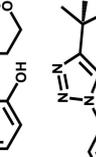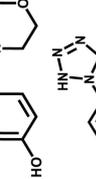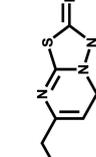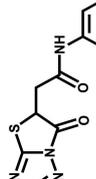
regions of the protein, including the already known catalytic active site (Supporting Information Figure 20). However, FTMap analysis performed on the dimeric nsp5 does not identify the catalytic site (Supporting Information Figure 22 and Supporting Information Figure 23), which is in line with one of the limitations of FTMap that it works best for single domains. Therefore, monomeric form of nsp5 was utilized for the analysis of druggability. For nsp3b, hot spots clustered mainly in the ADPr binding site (Supporting Information Figure 13). Similarly, we observed the same (previously known and additional binding pockets) trend of hot spot clustering in the other proteins of SCoV2, which facilitated the definition of the relevant clefts on the protein. We used PDBsum[63] to calculate the cleft regions and ranked the clefts according to their volume. Integration of the PDBsum derived cleft information and the FTMap-identified binding sites strikingly revealed that for 13 out of 18 proteins, the hot spots identified by FTMap overlapped with cleft 1, for three proteins with cleft 2 and for three proteins with cleft 3 as identified by PDBsum (Figure 3 and Supporting Information Table 4). Importantly, FTMap analysis together with the cleft analysis for each of the SCoV2 proteins investigated here revealed that indeed, the 18 proteins contain defined potential ligand binding sites and are thus druggable. As a next step, for a given hot spot, we compared and correlated the types of FTMap probes predicted to bind in the binding sites with the chemical substructures present in the experimentally identified fragments in the DSI-PL. For this purpose, we scanned and extracted the number of occurrences of the 16 FTMap probes for all the 768 compounds from the DSI-PL using cheminformatic tools (Supporting Information excel sheet 2 DSI PL Poised Library Characterized into the 16 Probes of FTMap.xlsx). As a next step, for each of the identified binder for a given target, we quantified the overlap of probes between the hits and FTMap probes (Supporting Information excel sheets). We then selected one binder for each target, for which binding effects were observed in one or more NMR experiment. Mapping of the ligand-derived functional units revealed that for 14 out of 18 of these ligands, a 100% correlation was observed with the probes found within one or more of the crossclusters spanning the predicted cleft (Figure 3 and Supporting Information Table 4). For example, binder 21 showed positive binding effects in both wLOGSY and STD NMR experiments for nsp5 and was hence chosen as ligand of choice for this target. FTMap and cleft analysis of nsp5 suggested that crossclusters 1, 2, 7, and 8 were situated within the known active site (cleft 2) of the protease. Binder 21 is composed of mainly three (methanamine, benzene and urea) FTMap probes, and all of them are present in the crosscluster 1 (100%). The crossclusters 2, 7 and 8 each consist of one of the three probes (33%). These observations show that there is a good overlap between the chemical substructures of the FTMap ligands and those experimental fragments that occupy the hot spots, suggesting a likely binding site for this ligand. Further, in order to gain insight into the binding site of the ligand, we performed molecular docking using the Swissdock web server.[64,65] For 50% of the targets, we observed that the top-ranked pose (i.e., the ligand with the lowest binding free energy) of the ligand docks onto the binding site (Figure 3, docked ligand shown in cyan).

In order to test the validity of our predicted ligand binding sites, we performed ligand-observed (ORF9a (NTD), nsp3 (SUD-MC) and nsp5) and/or protein observed (nsp5 and nsp10) titrations and determined the dissociation constants for a subset of targets by NMR. In general, the dissociation constants $K_D$ for the fragments ranged from 50 to 2000 μM (Table 2 and Supporting Information Figure 24). Binder 13 (Z979145504) bound to nsp5 with the highest affinity. In addition, we also performed protein-observed titrations for ligands that bind to nsp5 and nsp10. An advantage of protein-observed NMR titrations is that apart from obtaining information on the dissociation constants, it is also possible to visualize the binding site of the ligand by mapping the CSPs, provided the backbone amides are assigned. Previously, within the Covid19-NMR consortium we have achieved the near-to-complete backbone assignments of nsp10 and nsp5.[36,62,66] Binder 21 was titrated to nsp5 and bound with a $K_D$ of ≈500 μM (Figure 4, bottom right). Mapping of the CSPs revealed that apart from remote CSP effects, the residues involved in the binding mainly clustered around the active site (Figure 4, top right, blue regions), which was in good agreement with the binding cleft identified by FTMap. Moreover, FTMap and cleft analysis of nsp5 not only identified the same two sites (S1 and S3) in line with the crystal structure of binder 21 in complex with nsp5 (Figure 4, lower left, orange stick), but also reveals two additional sites (S1' and S2). A similar analysis performed for a weak binder (binder 2, $K_D$ of ≈2000 μM) of nsp10 (Supporting Information Figure 25) reveals a striking correlation between the binding site mapped based on NMR CSPs and the FTMap-detected hot spot, thus supporting the robustness and validity of our analysis. Further, FTMap analysis of the 6 and 8 overlapping binders for X-ray/NMR screening and three nsp5 constructs, respectively, suggests, that the active site (cleft 2) is their putative binding site (Supporting Information Table 5 and Supporting Information Table 6). Moreover, the 6 X-ray/NMR overlapping binders revealed identical docking poses for single chains of either monomeric (5r83) or dimeric (7khp) structures as documented in Supporting Information Figure 26.

The NMR-based fragment hit structures were compared to >2 million molecules contained in the ChEMBL,[67] PubChem[68] and NCATS (https://opendata.ncats.nih.gov/covid19/) associated data resources of bioactive compounds. 2D Tanimoto scoring[69] was used to identify analogues annotated as active in SCoV2 bioassays. To capture "weak associations" between hits and bioactive analogues, a cut-off of 0.65 was set, which revealed 35 hit fragments associated with 50 analogues identified as active in 16 different SCoV2 assays, representing a total of 154 distinct bioactivities (Supporting Information excel sheet 3 Hits to Bioactives.xlsx). A knowledge graph additionally annotated with links to public SCoV2 assay information and relevant metadata on the bioactivities and primary targets of the 154 compounds can be accessed at https://github.com/Fraunhofer-ITMP/COVID NMR -KG. At a more stringent

Tanimoto cut-off of 0.70, a group of 9 hit fragments representing 9 analogues were identified (Table 3). Seven of the analogues, with $IC_{50}$ values between 390 nM and 3190 nM, were identified as inhibitors of protease activity, in the study by Kuzikov et al.,[70] who screened a compound repurposing collection in a FRET-based biochemical assay against full-length nsp5. Although the fragment hits binding to nsp5 also binds to at least one additional protein, three (binder 6, 37 and 67) have analogues that inhibit nsp5 activity. Two analogue compounds were also active in phenotypic assays monitoring the anti-cytopathic effect of SCoV2 in Vero E6 cell models (Metoprine, $IC_{50} = 2340$ nM and Oxyclozanide, $IC_{50} = 3710$ nM).[71] The NMR hit (binder 74) related to Metroprine, binds multiple proteins (nsp7, nsp3c (SUD-MC), nsp3d, $_{His6}$nsp15, nsp10 and nsp16) whilst the Oxyclozanide related compound (binder 79) targets a smaller group of viral proteins, namely nsp7 and nsp3c (SUD-MC).

## Conclusion

Covid19 has triggered enormous research efforts. For the less than 30 viral proteins and 15 conserved RNA regulatory elements, holistic approaches screening almost all viral components can be pursued. X-ray crystallography with recently introduced automatization of fragment screening approaches[33,54] has spearheaded medicinal chemistry approaches focusing on a subset of the viral protein targets. Previously (Sreeramulu, Richter et al.) and here, we exploit the unique advances of NMR spectroscopy for screening of structured elements of the RNA genome as well as the soluble parts of the proteome. The work described thus provides information for $25*768 = 19200$ possible protein-ligand interactions monitored by 4 different ligand-based NMR experiments. The 768 ligands come from a highly privileged fragment library. They have been assembled previously and validated by NMR for their chemical purity and solubility.[59,60]

The screening identifies 311 hits (1.5% overall hit rate). The work goes, however, beyond reporting these screening results. We delineate a procedure to combine computational methods to validate binding site prediction from FTMap and PDBsum with the experimentally detected binding ligands. This procedure relies on the prediction of chemical sub-moieties essential for binding and the similarity of these substructures in the set of experimental binders. The thus identified and prioritized binding sites allow application of focused docking protocols and further, the experimental cross-validation by protein-based NMR experiments. From these protein-based NMR experiments, we show that dissociation constants of these fragments with proteins range from 80 μM to several millimolar. The determination of binding affinities can be used to prioritize medicinal chemistry campaigns. Using bioinformatics, identification of fragment binders also serves as starting point for database searches of known binders, using chemical similarity scores between fragments and known inhibitors as selection criterion. Thus, the herein developed workflow allows for

holistic screening of the majority of the viral proteome. It provides highly valuable data for the day-to-day support of medicinal chemistry campaigns aiming at developing novel drugs applying fragment-based drug discovery. These data will also serve development of artificial intelligence (AI) based algorithms to inform hit-to-lead campaigns.

## Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

The data that support the findings of this study are openly available in BMRbig (bmrbig45 to bmrbig69) at https://bmrbig.bmrb.io/released/.[72–96]

**Keywords:** COVID19-NMR · Drug Discovery · Fragment Screening · NMR Spectroscopy · Protein · SARS-CoV-2

[1] D. Adam, *Nature* **2022**, *601*, 312–315.
[2] G. J. Kontoghiorghes, S. Fetta, C. N. Kontoghiorghe, *Front. Biosci.* **2021**, *26*, 1723–1736.
[3] Z. A. Shyr, K. Gorshkov, C. Z. Chen, W. Zheng, *J. Pharmacol. Exp. Ther.* **2020**, *375*, 127–138.
[4] T. T. Le, J. P. Cramer, R. Chen, S. Mayhew, *Nat. Rev. Drug Discovery* **2020**, *19*, 667–668.
[5] M. Mei, X. Tan, *Front. Mol. Biosci.* **2021**, *8*, https://doi.org/10.3389/fmolb.2021.671263.
[6] M. Cully, *Nat. Rev. Drug Discovery* **2022**, *21*, 3–5.
[7] E. Petersen, F. Ntoumi, D. S. Hui, A. Abubakar, L. D. Kramer, et al., *Int. J. Infectious Diseases* **2022**, *114*, 268–272.
[8] M. Hoffmann, N. Krüger, S. Schulz, A. Cossmann, C. Rocha, et al., *Cell* **2022**, *185*, 447–456.
[9] L. A. VanBlargan, J. M. Errico, P. J. Halfmann, S. J. Zost, J. E. Crowe, et al., *Nat. Med.* **2022**, *28*, 490–495.
[10] W. F. Garcia-Beltran, K. J. st Denis, A. Hoelzemer, E. C. Lam, A. D. Nitido, et al., *Cell* **2022**, *185*, 457–466.
[11] Y.-W. Zhou, Y. Xie, L.-S. Tang, D. Pu, Y.-J. Zhu, et al., *Signal Transduction Targeted Ther.* **2021**, *6*, 317.
[12] B. Malone, N. Urakova, E. J. Snijder, E. A. Campbell, *Nat. Rev. Mol. Cell Biol.* **2022**, *23*, 21–39.
[13] F. Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, et al., *Nature* **2020**, *579*, 265–269.
[14] A. R. Fehr, S. Perlman, *Methods Mol. Biol.* **2015**, *1282*, 1–23.
[15] E. J. Snijder, P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, et al., *J. Mol. Biol.* **2003**, *331*, 991–1004.
[16] D. E. Gordon, G. M. Jang, M. Bouhaddou, J. Xu, K. Obernier, et al., *Nature* **2020**, *583*, 459–468.
[17] C. W. Nelson, Z. Ardern, T. L. Goldberg, C. Meng, C. H. Kuo, et al., *eLife* **2020**, *9*, e59633.
[18] A. Pavesi, *Virology* **2020**, *546*, 51–66.
[19] P. Venkatesan, *Lancet Respir. Med.* **2021**, *9*, e63.
[20] B. Cao, Y. Wang, D. Wen, W. Liu, J. Wang, et al., *N. Engl. J. Med.* **2020**, *382*, 1787–1799.
[21] M. Wang, R. Cao, L. Zhang, X. Yang, J. Liu, et al., *Cell Research* **2020**, *30*, 269–271.
[22] W.-C. Ko, J.-M. Rolain, N.-Y. Lee, P.-L. Chen, C.-T. Huang, et al., *Int. J. Antimicrob. Agents* **2020**, *55*, 105933.
[23] M. A. Martinez, *Front. Immunol.* **2022**, *12*, https://doi.org/10.3389/fimmu.2021.635371.
[24] J. O. Ogidigo, E. A. Iwuchukwu, C. U. Ibeji, O. Okpalefe, M. E. S. Soliman, *J. Biomol. Struct. Dyn.* **2022**, *40*, 2284–2301.
[25] C. Liu, X. Zhu, X. Lu, X. Zhang, X. Jia, T. Yang, *J. Pharm. Anal.* **2021**, *11*, 272–277.
[26] M. A. White, W. Lin, X. Cheng, *J. Phys. Chem. Lett.* **2020**, *11*, 9144–9151.
[27] M. T. J. Quimque, K. I. R. Notarte, R. A. T. Fernandez, M. A. O. Mendoza, R. A. D. Liman, et al., *J. Biomol. Struct. Dyn.* **2021**, *39*, 4316–4333.
[28] A. Carino, F. Moraca, B. Fiorillo, S. Marchianò, V. Sepe, et al., *Front. Chem.* **2020**, *8*, https://doi.org/10.3389/fchem.2020.572885.
[29] S. Barage, A. Karthic, R. Bavi, N. Desai, R. Kumar, et al., *J. Biomol. Struct. Dyn.* **2022**, *40*, 2557–2574.
[30] M. Macchiagodena, M. Pagliai, P. Procacci, *Chem. Phys. Lett.* **2020**, *750*, 137489.
[31] S. Yazdani, N. de Maio, Y. Ding, V. Shahani, N. Goldman, M. Schapira, *J. Proteome Res.* **2021**, *20*, 4212–4215.
[32] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, et al., *Science* **2020**, *368*, 409–412.
[33] M. Schuller, G. J. Correy, S. Gahbauer, D. Fearon, T. Wu, et al., *Sci. Adv.* **2021**, *7*, eabf8711.
[34] S. Günther, P. Y. A. Reinke, Y. Fernández-García, J. Lieske, T. J. Lane, et al., *Science* **2021**, *372*, 642–646.
[35] J. A. Newman, A. Douangamath, S. Yadzani, Y. Yosaatmadja, A. Aimon, et al., *Nat. Commun.* **2021**, *12*, 4848.
[36] F. Cantrelle, E. Boll, L. Brier, D. Moschidi, S. Belouzard, et al., *Angew. Chem. Int. Ed.* **2021**, *60*, 25428–25435; *Angew. Chem.* **2021**, *133*, 25632–25639.
[37] A. L. Kantsadi, E. Cattermole, M.-T. Matsoukas, G. A. Spyroulias, I. Vakonakis, *J. Biomol. NMR* **2021**, *75*, 167–178.
[38] N. Imprachim, Y. Yosaatmadja, J. A. Newman, *bioRxiv* **2022**, https://doi.org/10.1101/2022.03.11.483836.
[39] V. Napolitano, A. Dabrowska, K. Schorpp, A. Mourão, E. Barreto-Duran, et al., *Cell Chem. Biol.* **2022**, *29*, 774–784.
[40] T. C. M. Consortium, H. Achdout, A. Aimon, E. Bar-David, H. Barr, et al., *bioRxiv* **2022**, https://doi.org/10.1101/2020.10.29.339317.
[41] M. Kozlov, *Nature* **2022**, *601*, 496.
[42] J. F. X. Diffley, *Biochem. J.* **2021**, *478*, 2533–2535.
[43] C. T. Lim, K. W. Tan, M. Wu, R. Ulferts, L. A. Armstrong, et al., *Biochem. J.* **2021**, *478*, 2517–2531.
[44] J. C. Milligan, T. U. Zeisner, G. Papageorgiou, D. Joshi, C. Soudy, et al., *Biochem. J.* **2021**, *478*, 2499–2515.
[45] S. Basu, T. Mak, R. Ulferts, M. Wu, T. Deegan, et al., *Biochem. J.* **2021**, *478*, 2481–2497.
[46] B. Canal, R. Fujisawa, A. W. McClure, T. D. Deegan, M. Wu, et al., *Biochem. J.* **2021**, *478*, 2465–2479.
[47] B. Canal, A. W. McClure, J. F. Curran, M. Wu, R. Ulferts, et al., *Biochem. J.* **2021**, *478*, 2445–2464.
[48] A. P. Bertolin, F. Weissmann, J. Zeng, V. Posse, J. C. Milligan, et al., *Biochem. J.* **2021**, *478*, 2425–2443.
[49] J. Zeng, F. Weissmann, A. P. Bertolin, V. Posse, B. Canal, et al., *Biochem. J.* **2021**, *478*, 2405–2423.
[50] P. Ren, W. Shang, W. Yin, H. Ge, L. Wang, et al., *Acta Pharm. Sin.* **2022**, *43*, 483–493.
[51] N. Altincekic, S. M. Korn, N. S. Qureshi, M. Dujardin, M. Ninot-Pedrosa, et al., *Front. Mol. Biosci.* **2021**, *8*, 89.
[52] S. Sreeramulu, C. Richter, H. Berg, M. A. Wirtz Martin, B. Ceylan, et al., *Angew. Chem. Int. Ed.* **2021**, *60*, 19191–19200; *Angew. Chem.* **2021**, *133*, 19340–19349.
[53] D. Kozakov, L. E. Grove, D. R. Hall, T. Bohnuud, S. E. Mottarella, et al., *Nature Protocols* **2015**, *10*, 733–755.
[54] A. Douangamath, D. Fearon, P. Gehrtz, T. Krojer, P. Lukacik, et al., *Nat. Commun.* **2020**, *11*, 5047.
[55] D. Shin, R. Mukherjee, D. Grewe, D. Bojkova, K. Baek, et al., *Nature* **2020**, *587*, 657–662.
[56] K. Anand, J. Ziebuhr, P. Wadhwani, J. R. Mesters, R. Hilgenfeld, *Science* **2003**, *300*, 1763–1767.
[57] N. T. Nashed, A. Aniana, R. Ghirlando, S. C. Chiliveri, J. M. Louis, *Commun. Biol.* **2022**, *5*, 160.
[58] B. Goyal, D. Goyal, *ACS Comb. Sci.* **2020**, *22*, 297–305.
[59] O. B. Cox, T. Krojer, P. Collins, O. Monteiro, R. Talon, A. Bradley, O. Fedorov, J. Amin, B. D. Marsden, J. Spencer, F. von Delft, P. E. Brennan, *Chem. Sci.* **2016**, *7*, 2322–2330.

[60] S. Sreeramulu, C. Richter, T. Kuehn, K. Azzaoui, M. J. J. Blommers, et al., *J. Biomol. NMR* **2020**, *74*, 555–563.

[61] H. Berg, M. A. Wirtz Martin, A. Niesteruk, C. Richter, S. Sreeramulu, H. Schwalbe, *J. Visualized Exp.* **2021**, *172*, e62262.

[62] A. J. Robertson, J. Ying, A. Bax, *Magn. Reson.* **2021**, *2*, 129–138.

[63] R. A. Laskowski, J. Jabłońska, L. Pravda, R. S. Vařeková, J. M. Thornton, *Protein Sci.* **2018**, *27*, 129–134.

[64] A. Grosdidier, V. Zoete, O. Michielin, *J. Comput. Chem.* **2011**, *32*, 2149–2159.

[65] A. Grosdidier, V. Zoete, O. Michielin, *Nucleic Acids Res.* **2011**, *39*, W270–W277.

[66] N. Kubatova, N. S. Qureshi, N. Altincekic, R. Abele, J. K. Bains, et al., *Biomol. NMR Assign.* **2021**, *15*, 65–71.

[67] D. Mendez, A. Gaulton, A. P. Bento, J. Chambers, M. De Veij, et al., *Nucleic Acids Res.* **2019**, *47*, D930–D940.

[68] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, et al., *Nucleic Acids Res.* **2021**, *49*, D1388–D1395.

[69] X. Chen, C. H. Reynolds, *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1407–1414.

[70] M. Kuzikov, E. Costanzi, J. Reinshagen, F. Esposito, L. Vangeel, et al., *ACS Pharmacol. Transl. Sci.* **2021**, *4*, 1096–1110.

[71] S. Jeon, M. Ko, J. Lee, I. Choi, S. Y. Byun, S. Park, D. Shum, S. Kim, *Antimicrob. Agents Chemother.* **2020**, *64*, https://doi.org/10.1128/AAC.00819-20.

[72] nsp1, H. Berg, 2022, DSI-PL_COVID19-NMR _nsp1, BMRbig, bmrbig48, https://bmrbig.bmrb.io/released/bmrbig48.

[73] _GHM_nsp5, H. Berg, 2022, DSI-PL_COVID19-NMR _GHM_nsp5, BMRbig, bmrbig45, https://bmrbig.bmrb.io/released/bmrbig45.

[74] nsp3e, H. Berg, 2022, DSI-PL_COVID19-NMR _nsp3e, BMRbig, bmrbig56, https://bmrbig.bmrb.io/released/bmrbig56.

[75] nsp3a, H. Berg, 2022, DSI-PL_COVID19-NMR _nsp3a, BMRbig, bmrbig50, https://bmrbig.bmrb.io/released/bmrbig50.

[76] nsp9, H. Berg, 2022, DSI-PL_COVID19-NMR _nsp9, BMRbig, bmrbig61, https://bmrbig.bmrb.io/released/bmrbig61.

[77] nsp7, H. Berg, 2022, DSI-PL_COVID19-NMR _nsp7, BMRbig, bmrbig59, https://bmrbig.bmrb.io/released/bmrbig59.

[78] nsp8, H. Berg, 2022, DSI-PL_COVID19-NMR _nsp8, BMRbig, bmrbig60, https://bmrbig.bmrb.io/released/bmrbig60.

[79] nsp3c (SUD-N), H. Berg, 2022, DSI-PL_COVID19-NMR _nsp3c_SUD-N, BMRbig, bmrbig54, https://bmrbig.bmrb.io/released/bmrbig54.

[80] nsp3c (SUD-MC), H. Berg, 2022, DSI-PL_COVID19-NMR _nsp3c_SUD-MC, BMRbig, bmrbig53, https://bmrbig.bmrb.io/released/bmrbig53.

[81] ORF9a (NTD), H. Berg, 2022, DSI-PL_COVID19-NMR _ORF9a_NTD, BMRbig, bmrbig67, https://bmrbig.bmrb.io/released/bmrbig67.

[82] ORF9a (IDR1-NTD-IDR2), H. Berg, 2022, DSI-PL_COVID19-NMR _ORF9a_IDR1-NTD-IDR2, BMRbig, bmrbig66, https://bmrbig.bmrb.io/released/bmrbig66.

[83] nsp3b, H. Berg, 2022, DSI-PL_COVID19-NMR _nsp3b, BMRbig, bmrbig51, https://bmrbig.bmrb.io/released/bmrbig51.

[84] nsp2 (CtDR), H. Berg, 2022, DSI-PL_COVID19-NMR _nsp2_CtDR, BMRbig, bmrbig49, https://bmrbig.bmrb.io/released/bmrbig49.

[85] nsp3b·GS-441524, H. Berg, 2022, DSI-PL_COVID19-NMR _nsp3b_GS-441524, BMRbig, bmrbig52, https://bmrbig.bmrb.io/released/bmrbig52.

[86] nsp3d, H. Berg, 2022, DSI-PL_COVID19-NMR _nsp3d, BMRbig, bmrbig55, https://bmrbig.bmrb.io/released/bmrbig55.

[87] _His6_nsp15, H. Berg, 2022, DSI-PL_COVID19-NMR _His6_nsp15, BMRbig, bmrbig46, https://bmrbig.bmrb.io/released/bmrbig46.

[88] _GS_nsp5, H. Berg, 2022, DSI-PL_COVID19-NMR _GS_nsp5, BMRbig, bmrbig47, https://bmrbig.bmrb.io/released/bmrbig47.

[89] ORF9a (CTD), H. Berg, 2022, DSI-PL_COVID19-NMR _ORF9a_CTD, BMRbig, bmrbig65, https://bmrbig.bmrb.io/released/bmrbig65.

[90] nsp10, H. Berg, 2022, DSI-PL_COVID19-NMR _nsp10, BMRbig, bmrbig62, https://bmrbig.bmrb.io/released/bmrbig62.

[91] ORF9a (NTD-SR), H. Berg, 2022, DSI-PL_COVID19-NMR _ORF9a_NTD-SR, BMRbig, bmrbig68, https://bmrbig.bmrb.io/released/bmrbig68.

[92] nsp10·nsp16, H. Berg, 2022, DSI-PL_COVID19-NMR _nsp10_nsp16, BMRbig, bmrbig63, https://bmrbig.bmrb.io/released/bmrbig63.

[93] ORF9b, H. Berg, 2022, DSI-PL_COVID19-NMR _ORF9b, BMRbig, bmrbig69, https://bmrbig.bmrb.io/released/bmrbig69.

[94] nsp3y, H. Berg, 2022, DSI-PL_COVID19-NMR _nsp3y, BMRbig, bmrbig57, https://bmrbig.bmrb.io/released/bmrbig57.

[95] nsp5, H. Berg, 2022, DSI-PL_COVID19-NMR _nsp5, BMRbig, bmrbig58, https://bmrbig.bmrb.io/released/bmrbig58.

[96] nsp10·nsp14, H. Berg, 2022, DSI-PL_COVID19-NMR _nsp10_nsp14, BMRbig, bmrbig64, https://bmrbig.bmrb.io/released/bmrbig64.