A Journal of the Gesellschaft Deutscher Chemiker

Angewandte

GDCh

International Edition

Chemie

www.angewandte.org

Accepted Article

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

Authors: 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. Hutchinson, 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 Lina Bula, Matthias Buetikofer, Ícaro Putinhon Caruso, Isabella Caterina Felli, Andrea T Da Poian, Gisele Cardoso de Amorim, Nikolaos K Fourkiotis, Angelo Gallo, Dhiman Ghosh, Francesco Gomes-Neto, Oksana Gorbatyuk, Bing Hao, Vilius Kurauskas, Lauriane Lecoq, Yunfeng Li, Nathane Cunha Mebus-Antunes, Miguel Mompean, Thais Cristtina Neves-Martins, Marti Ninot-Pedrosa, Anderson S Pinheiro, Letizia Pontoriero, Yulia Pustovalova, Roland Riek, Angus Robertson, Marie Jose Abi Saad, Miguel A Treviño, Aikaterini C Tsika, Fabio C.L. Almeida, Ad Bax, Katherine Henzler-Wildman, Jeffrey C Hoch, Kristaps Jaudzems, Douglas V Laurents, Julien Orts, Roberta Pieratelli, Georgios A Spyroulias, Elke Duchardt-Ferner, Jan Ferner, Boris Fuertig, Martin Hengesbach, Frank Löhr, Nusrat Qureshi, Christian Richter, Krishna Saxena, Andreas Schlundt, Sridhar Sreeramulu, Anna Wacker, Julia E Weigand, Julia Wirmer-Bartoschek, Jens Woehnert, and Harald Schwalbe

To be cited as: *Angew. Chem. Int. Ed.* **2022**, e202205858

WILEY-VCH

Link to VoR: https://doi.org/10.1002/anie.202205858

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

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

[a,b] H. Berg[ÿ], M. A. Wirtz Martin[ÿ], N. Altincekic[ÿ], Dr. I. Alshamleh, J. K. 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, 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
[a] Institute for Organic Chemistry and Chemical Biology, Goethe University Frankfurt, Frankfurt am Main, Germany
[b] 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

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

[d] Dr. K. Azzaoui, Dr. M. J. J. Blommers
Saverna Therapeutics, Pumpmattenweg 3, 4105 Biel-Benken, Switzerland

[e,f] Y. Gadiya, R. Karki, A. Zaliani, Dr. P. Gribbon
[e] Fraunhofer Institute for Translational Medicine and Pharmacology (ITMP), Screening Port, Schnackenburgallee 114, 22525 Hamburg, Germany
[f] Fraunhofer Cluster of Excellence for Immune-Mediated Diseases (CIMD), Theodor-Stern-Kai 7, 60596 Frankfurt am Main, Germany

[g,h] M. S. Almeida, A. T. Da Poian, N. C. Mebus-Antunes, T. C. Neves-Martins, F. C. L. Almeida
[g] Institute of Medical Biochemistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
[h] Fraunhofer Cluster of Excellence for Immune-Mediated Diseases (CIMD), Theodor-Stern-Kai 7, 60596 Frankfurt am Main, Germany

[i,j] C. D Anobom, A. S. Pinheiro
[i] National Center of Nuclear Magnetic Resonance (CNRMN), CENABIO, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
[j] Department of Biochemistry, Institute of Chemistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

[k] A. L. Bula, K. Jaudzems
[k] Latvian Institute of Organic Synthesis, Aizkraukles 21, LV-1006 Riga, Latvia

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

[g,m] Í. P. Caruso
[m] Multiuser Center for Biomolecular Innovation (CMIB), Department of Physics, São Paulo State University (UNESP), São José do Rio Preto, Brazil

[n,o] I. C. Felli, L. Pontoriero, R. Pierattelli
[n] Magnetic Resonance Center (CERM), University of Florence, Via Luigi Sacconi 6, Sesto Fiorentino, 50019, Florence, Italy
[o] Department of Chemistry "Ugo Schiff", University of Florence, Via della Lastruccia 3-13, Sesto Fiorentino, 50019, Florence, Italy

[g,p] G. C. de Amorim
[p] Multidisciplinary Center for Research in Biology (NUMPEX), Campus Duque de Caxias Federal University of Rio de Janeiro, Duque de Caxias, Brazil

[q] N. K. Fourkiotis, A. C. Tsika, G. A. Spyroulias
[q] Department of Pharmacy, University of Patras, Patras, Greece

[q,r] A. Gallo,
[r] Department of Chemistry, University of Torino IT -10126 Torino. Italy

[i,s] F. Gomes-Neto
[s] Laboratory of Toxinology, Oswaldo Cruz Foundation (FIOCRUZ), Rio de Janeiro, Brazil

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

[u] V. Kurauskas, K. Henzler-Wildman
[u] Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706

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

[w] M. Mompeán, M. Á. Treviño, D. V. Laurents
[w] "Rocasolano" Institute for Physical Chemistry, CSIC
[x] A. J. Robertson, A. Bax
[x] Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases,
[y] M. J. Abi Saad, J. Orts
[y] University of Vienna, Department of Pharmaceutical Sciences, Josef-Holaubek-Platz 2, A-1090 Vienna, Austria
[z] Dr. J. E. Weigand
[z] Department of Biology, Technical University of Darmstadt, Darmstadt, Germany
[ÿ] These authors contributed equally to this work.

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.) [4ï 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-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 [16ï 18]. Repurposing of (approved) drugs has been actively pursued as a strategy to counter SCoV2 infections [19ï 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 [24ï 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 [32ï 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 [42ï 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), 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) [59ï 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 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.

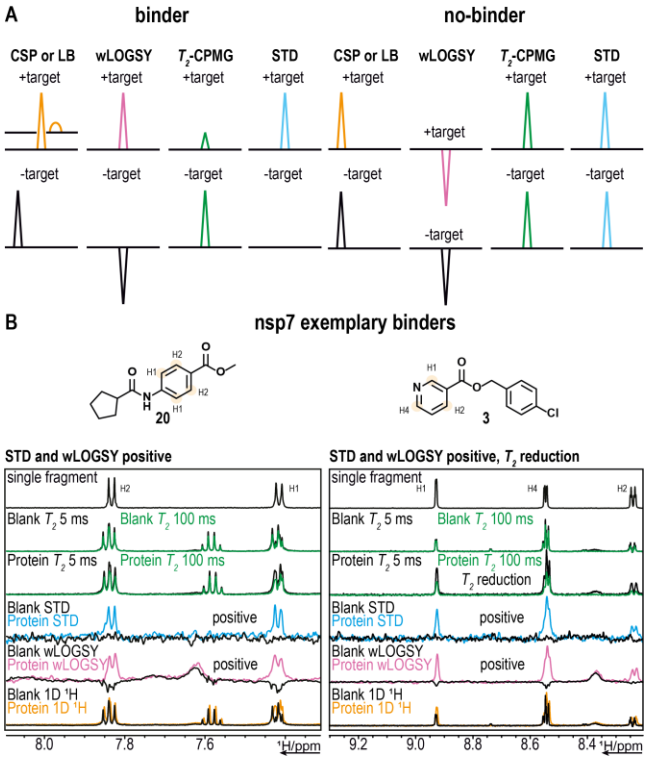


Figure 1. NMR based identification of binding fragments. (A) Schematic representation of all NMR experiments used in the screening that show exemplary effects indicating binding events in the presence of ligand compared to ligand free spectra. (B) NMR spectra (1D ^{1}H, wLOGSY, STD, and T_2-CPMG (5 ms 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.

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 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.

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	2w2g	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				

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
	$_{GS}$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	6w4h	92	3, 4, 5, 7, 8, 10	0
nsp10¶nsp14 *18,040-19,620*	*Exoribonuclease*	527		61.4				
	nsp10Ånsp14	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	-	-
	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	-

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
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

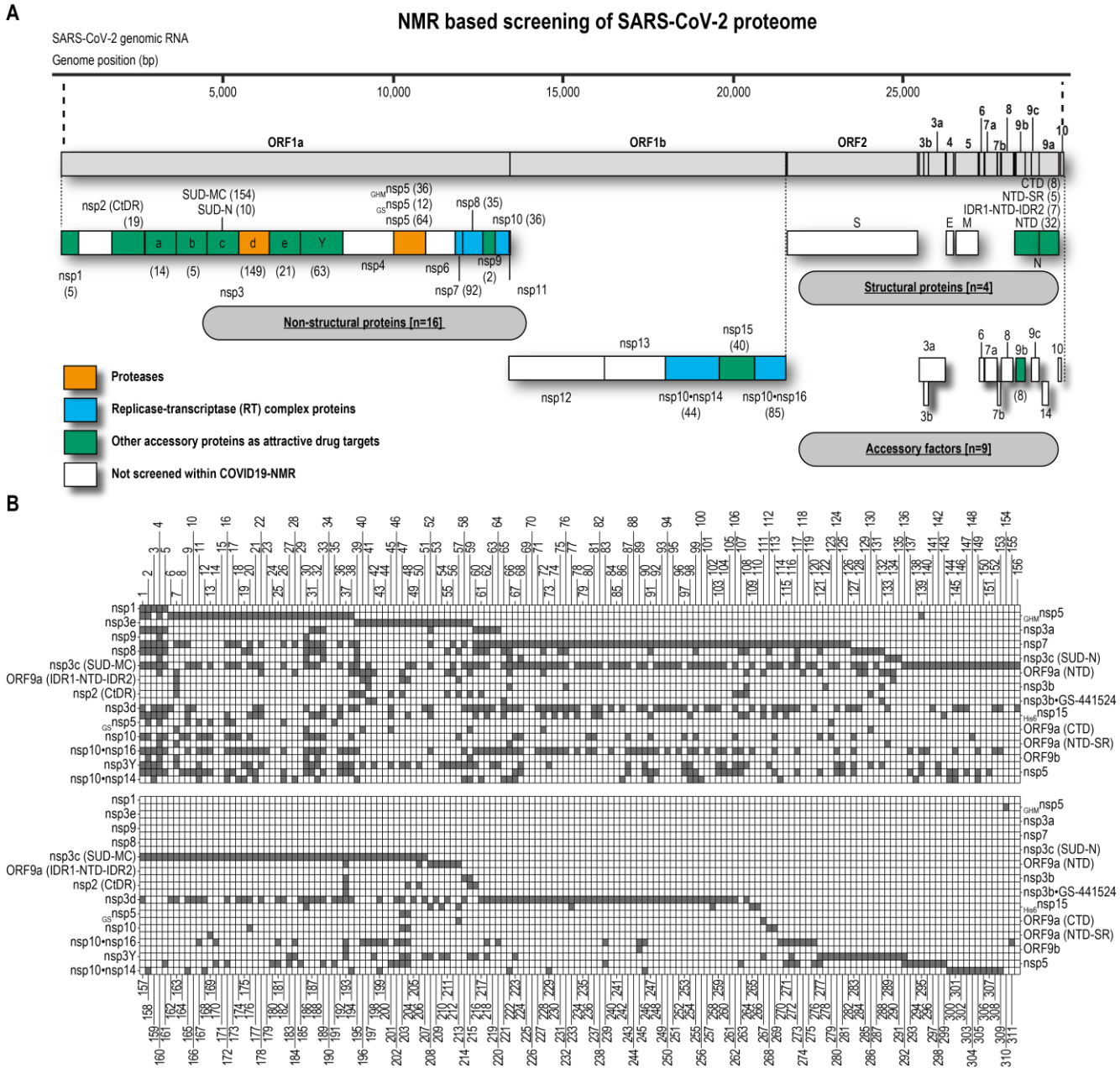


Figure 2. 311 binding fragments identified for SCoV2 proteins from NMR based fragment screening. (A) Schematic representation of the SCoV2 genome (adapted from [16]). (B) The two tables summarize all binding fragments identified in the NMR screening for their corresponding protein (grey). The first table 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).

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 determined 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 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).

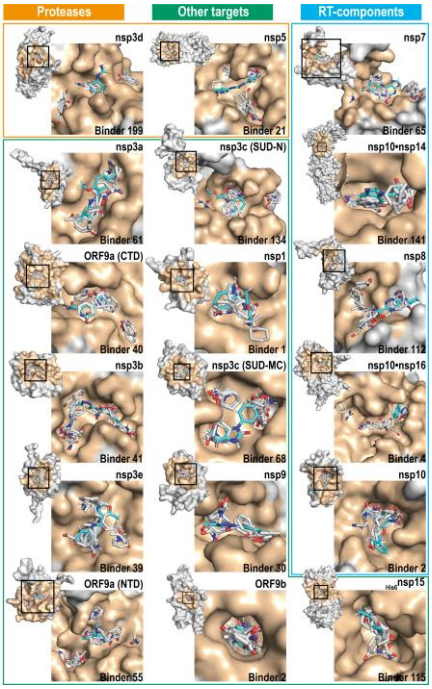


Figure 3. Hot spots identified using FTMap along with the docking of the NMR identified binders for 18 SCoV2 proteins. Proteases are highlighted with an orange box, RT-components with a blue box, and other targets with a green box. Zoom-ins show one of the identified clefts (beige colored) from PDBSum 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.

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 Supplementary 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 ñweak associationsò 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).

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.

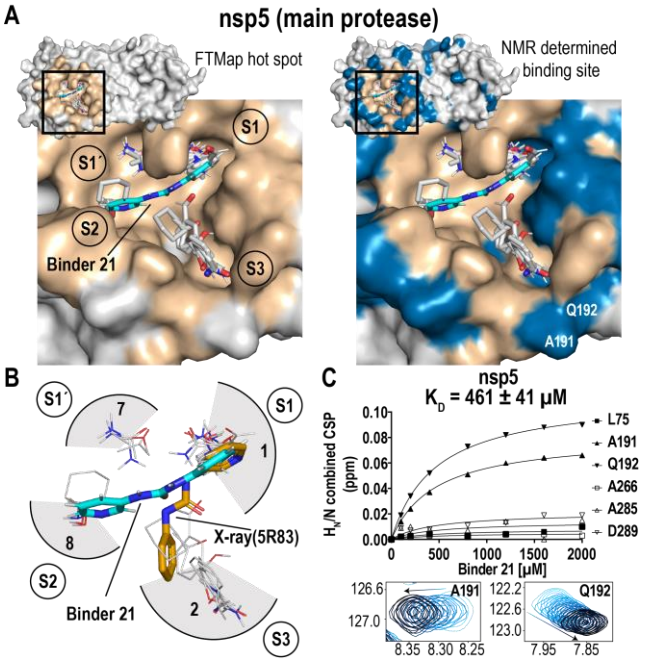
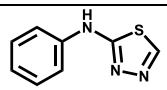
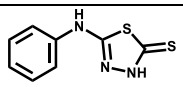
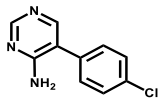
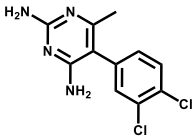
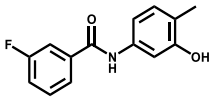
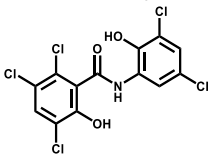
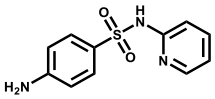
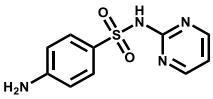
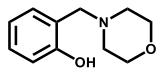
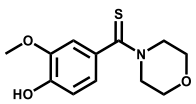
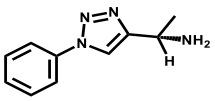
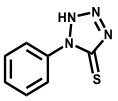
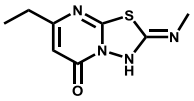
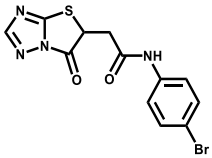
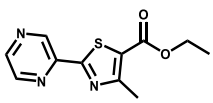
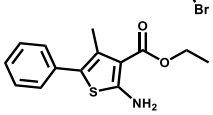
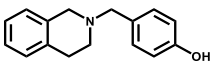
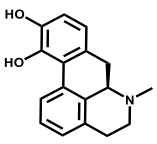


Figure 4. Agreement between bioinformatic and experimental mapping of the binding site. (A) The FTMap identified hot spot for nsp5. The subsites of the active site are labeled as S1, S1´, S2 and S3. The crossclusters (1, 2, 7, and 8) occupying the binding site are shown in grey sticks. The docked pose 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).

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].

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 BioAssay (nM)	Assay ChEMBL ID	Assay Description
67		nsp7, nsp3c (SUD-MC), nsp3d, nsp10¶nsp16, nsp5		CHEMBL2893356	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, $_{His6}$nsp15, nsp10¶nsp16		CHEMBL2643713	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)		CHEMBL2105450	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		$_{GHM}$nsp5, nsp3d		CHEMBL2266502	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		CHEMBL2436542	PD096194	0.72	2040	CHEMBL4495583	Biochemical, nsp5 (SCoV2 3CL-Pro protease inhibition) IC_{50} FRET format with a peptide substrate
37		$_{GHM}$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

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.

Acknowledgements

The work has been conducted in the international consortium of Covid19-NMR (covid19-nmr.de). We would like to thank Roderick Lambertz and Katharina Targaczewski for technical assistance. We thank Peter Maas and his team at SPECS for assembling the fragment library. Work at BMRZ is supported by the state of Hesse. Work in Covid19-nmr was supported by the Goethe Corona Funds, by the IWB-EFRE-program 20007375 of state of Hesse, the DFG through CRC902: ñMolecular Principles of RNA-based regulation.ò and through infrastructure funds (project numbers: 277478796, 277479031, 392682309, 452632086, 70653611) and by European Union's Horizon 2020 research and innovation program iNEXT-discovery under grant agreement No 871037. BY-COVID receives funding from the European Unionôs Horizon Europe Research and Innovation Programme under grant agreement number 101046203. "INSPIRED" (MIS 5002550) project, implemented under the Action ôReinforcement of the Research and Innovation Infrastructure,ô funded by the Operational Program ôCompetitiveness, Entrepreneurship and Innovationô (NSRF 2014ï 2020) and co-financed by Greece and the EU (European Regional Development Fund) and the FP7 REGPOT CT-2011-285950 ï ñSEE-DRUGò project (purchase of UPATôs 700 MHz NMR equipment). The support of the CERM/CIRMMP center of Instruct-ERIC is gratefully acknowledged. This work has been funded in part by a grant of the Italian Ministry of University and Research (FISR2020IP_02112, ID-COVID) and by Fondazione CR Firenze. A.S. is supported by the Deutsche Forschungsgemeinschaft [SFB902/B16, SCHL2062/2-1] and the Johanna Quandt Young Academy at Goethe [2019/AS01]. M.H. and C.F. thank SFB902 and the Stiftung Polytechnische Gesellschaft for the Scholarship. L.L. work was supported by the French National Research Agency (ANR, NMR-SCoV2-ORF8), the Fondation de la Recherche Médicale (FRM, NMR-SCoV2-ORF8), FINOVI and the IR-RMN-THC Fr3050 CNRS. Work at UConn Health was supported by grants from the US National Institutes of Health (R01GM135592 to B.H., P41GM111135 and R01GM123249 to J.C.H.) and the US National Science Foundation (DBI 2030601 to J.C.H.). Latvian Council of Science Grant No. VPP-COVID-2020/1-0014. National Science Foundation EAGER MCB-2031269. This work was supported by the grant Krebsliga KFS-4903-08-2019 and SNF-311030_192646 to J.O. P.G. (ITMP) The EOSC Future project is co-funded by the European Union Horizon Programme call INFRAEOSC-03-2020 - Grant Agreement Number 101017536.

Keywords: COVID19-nmr Ådrug discovery Åfragment screening ÅNMR spectroscopy ÅSARS-CoV-2 Åprotein

References

[1] D. Adam, *Nature* **2022**, *601*, 312ï 315.
[2] G. J. Kontoghiorghes, S. Fetta, C. N. Kontoghiorghe, *Front Biosci (Landmark Ed)* **2021**, *26*, 1723ï 1736.
[3] Z. A. Shyr, K. Gorshkov, C. Z. Chen, W. Zheng, *Journal of Pharmacology and Experimental Therapeutics* **2020**, *375*, 127ï 138.
[4] T. T. Le, J. P. Cramer, R. Chen, S. Mayhew, *Evolution of the COVID-19 Vaccine Development Landscape*, Nature Reviews **2020,** *19*, 667-668.
[5] M. Mei, X. Tan, *Frontiers in Molecular Biosciences* **2021**, *8*, DOI 10.3389/fmolb.2021.671263.
[6] M. Cully, *Nature Reviews Drug Discovery* **2022**, *21*, 3ï 5.
[7] E. Petersen, F. Ntoumi, D. S. Hui, A. Abubakar, L. D. Kramer, *et. al*, *International Journal of Infectious Diseases* **2022**, *114*, 268ï 272.
[8] M. Hoffmann, N. Krüger, S. Schulz, A. Cossmann, C. Rocha, *et. al*, *Cell* **2021**, *185*, 447-456, DOI 10.1016/j.cell.2021.12.032.
[9] L. A. VanBlargan, J. M. Errico, P. J. Halfmann, S. J. Zost, J. E. Crowe, *et. al*, *Nature Medicine* **2022**, *28*, 490-495, DOI 10.1038/s41591-021-01678-y.
[10] W. F. Garcia-Beltran, K. J. st. Denis, A. Hoelzemer, E. C. Lam, A. D. Nitido, *et. al*, *Cell* **2022**, *185*, 457-466, DOI 10.1016/j.cell.2021.12.033.
[11] Y.-W. Zhou, Y. Xie, L.-S. Tang, D. Pu, Y.-J. Zhu, *et. al*, *Signal Transduction and Targeted Therapy* **2021**, *6*, 317.
[12] B. Malone, N. Urakova, E. J. Snijder, E. A. Campbell, *Nature Reviews Molecular Cell Biology* **2022**, *23*, 21ï 39.
[13] F. Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, *et. al*, *Nature* **2020**, *579*, 265ï 269.
[14] A. R. Fehr, S. Perlman, *Methods Mol Biol* **2015**, *1282*, 1ï 23.
[15] E. J. Snijder, P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, *et. al*, *J Mol Biol* **2003**, *331*, 991ï 1004.
[16] D. E. Gordon, G. M. Jang, M. Bouhaddou, J. Xu, K. Obernier, *et. al*, *Nature* **2020**, *583*, 459ï 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*, 51ï 66.
[19] P. Venkatesan, *The Lancet Respiratory Medicine* **2021**, *9*, e63.
[20] B. Cao, Y. Wang, D. Wen, W. Liu, J. Wang, *et. al*, *New England Journal of Medicine* **2020**, *382*, 1787ï 1799.
[21] M. Wang, R. Cao, L. Zhang, X. Yang, J. Liu, *et. al*, *Cell Research* **2020**, *30*, 269ï 271.
[22] W.-C. Ko, J.-M. Rolain, N.-Y. Lee, P.-L. Chen, C.-T. Huang, *et. al*, *International Journal of Antimicrobial Agents* **2020**, *55*, 105933.
[23] M. A. Martinez, *Frontiers in Immunology* **2021**, *12*, DOI 10.3389/fimmu.2021.635371.

[24] J. O. Ogidigo, E. A. Iwuchukwu, C. U. Ibeji, O. Okpalefe, M. E. S. Soliman, *Journal of Biomolecular Structure and Dynamics* **2020**, *40*, 2284ï 2301.

[25] C. Liu, X. Zhu, Y. Lu, X. Zhang, X. Jia, T. Yang, *Journal of Pharmaceutical Analysis* **2021**, *11*, 272ï 277.

[26] M. A. White, W. Lin, X. Cheng, *The Journal of Physical Chemistry Letters* **2020**, *11*, 9144ï 9151.

[27] M. T. J. Quimque, K. I. R. Notarte, R. A. T. Fernandez, M. A. O. Mendoza, R. A. D. Liman, *et. al*, *Journal of Biomolecular Structure and Dynamics* **2021**, *39*, 4316ï 4333.

[28] A. Carino, F. Moraca, B. Fiorillo, S. Marchianò, V. Sepe, *et. al*, *Frontiers in Chemistry* **2020**, *8*, DOI 10.3389/fchem.2020.572885.

[29] S. Barage, A. Karthic, R. Bavi, N. Desai, R. Kumar, *et. al*, *Journal of Biomolecular Structure and Dynamics* **2020**, *40*, 2557-2574.

[30] M. Macchiagodena, M. Pagliai, P. Procacci, *Chemical Physics Letters* **2020**, *750*, 137489.

[31] S. Yazdani, N. de Maio, Y. Ding, V. Shahani, N. Goldman, M. Schapira, *Journal of Proteome Research* **2021**, *20*, 4212ï 4215.

[32] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, *et. al*, *Science (1979)* **2020**, *368*, 409ï 412.

[33] M. Schuller, G. J. Correy, S. Gahbauer, D. Fearon, T. Wu, *et. al*, *Science Advances* **2021**, *7*, eabf8711, DOI 10.1126/sciadv.abf8711.

[34] S. Günther, P. Y. A. Reinke, Y. Fernández-García, J. Lieske, T. J. Lane, *et. al*, *Science (1979)* **2021**, *372*, 642ï 646.

[35] J. A. Newman, A. Douangamath, S. Yadzani, Y. Yosaatmadja, A. Aimon, *et. al*, *Nature Communications* **2021**, *12*, 4848.

[36] F. Cantrelle, E. Boll, L. Brier, D. Moschidi, S. Belouzard, *et. al*, *Angewandte Chemie International Edition* **2021**, *60*, 25428ï 25435.

[37] A. L. Kantsadi, E. Cattermole, M.-T. Matsoukas, G. A. Spyroulias, I. Vakonakis, *Journal of Biomolecular NMR* **2021**, *75*, 167ï 178.

[38] N. Imprachim, Y. Yosaatmadja, J. A. Newman, *bioRxiv* **2022**, 2022.03.11.483836.

[39] V. Napolitano, A. Dabrowska, K. Schorpp, A. Mourão, E. Barreto-Duran, *et. al*, *Cell Chemical Biology* **2022**, *29*, 774-784, DOI 10.1016/j.chembiol.2021.11.006.

[40] T. C. M. Consortium, H. Achdout, A. Aimon, E. Bar-David, H. Barr, *et. al*, *bioRxiv* **2022**, 2020.10.29.339317.

[41] M. Kozlov, *Nature* **2022**, *601*, 496, DOI 10.1038/d41586-022-00112-8.

[42] J. F. X. Diffley, *Biochemical Journal* **2021**, *478*, 2533ï 2535.

[43] C. T. Lim, K. W. Tan, M. Wu, R. Ulferts, L. A. Armstrong, *et. al*, *Biochemical Journal* **2021**, *478*, 2517ï 2531.

[44] J. C. Milligan, T. U. Zeisner, G. Papageorgiou, D. Joshi, C. Soudy, *et. al*, *Biochemical Journal* **2021**, *478*, 2499ï 2515.

[45] S. Basu, T. Mak, R. Ulferts, M. Wu, T. Deegan, *et. al*, *Biochemical Journal* **2021**, *478*, 2481ï 2497.

[46] B. Canal, R. Fujisawa, A. W. McClure, T. D. Deegan, M. Wu, *et. al*, *Biochemical Journal* **2021**, *478*, 2465ï 2479.

[47] B. Canal, A. W. McClure, J. F. Curran, M. Wu, R. Ulferts, *et. al*, *Biochemical Journal* **2021**, *478*, 2445ï 2464.

[48] A. P. Bertolin, F. Weissmann, J. Zeng, V. Posse, J. C. Milligan, *et. al*, *Biochemical Journal* **2021**, *478*, 2425ï 2443.

[49] J. Zeng, F. Weissmann, A. P. Bertolin, V. Posse, B. Canal, *et. al*, *Biochemical Journal* **2021**, *478*, 2405ï 2423.

[50] P. Ren, W. Shang, W. Yin, H. Ge, L. Wang, *et. al*, *Acta Pharmacologica Sinica* **2022**, *43*, 483ï 493.

[51] N. Altincekic, S. M. Korn, N. S. Qureshi, M. Dujardin, M. Ninot-Pedrosa, *et. al*, *Frontiers in Molecular Biosciences* **2021**, *8*, 89.

[52] S. Sreeramulu, C. Richter, H. Berg, M. A. Wirtz Martin, B. Ceylan, *et. al*, *Angewandte Chemie - International Edition* **2021**, *60*, 19191-19200, DOI 10.1002/anie.202103693.

[53] D. Kozakov, L. E. Grove, D. R. Hall, T. Bohnuud, S. E. Mottarella, *et. al*, *Nature Protocols* **2015**, *10*, 733ï 755.

[54] A. Douangamath, D. Fearon, P. Gehrtz, T. Krojer, P. Lukacik, *et. al*, *Nature Communications* **2020**, *11*, DOI 10.1038/s41467-020-18709-w.

[55] D. Shin, R. Mukherjee, D. Grewe, D. Bojkova, K. Baek, *et. al*, *Nature* **2020**, *587*, 657ï 662.

[56] K. Anand, J. Ziebuhr, P. Wadhwani, J. R. Mesters, R. Hilgenfeld, *Science (1979)* **2003**, *300*, 1763ï 1767.

[57] N. T. Nashed, A. Aniana, R. Ghirlando, S. C. Chiliveri, J. M. Louis, *Communications Biology* **2022**, *5*, 160.

[58] B. Goyal, D. Goyal, *ACS Combinatorial Science* **2020**, *22*, 297ï 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, *Chemical Science* **2016**, *7*, 2322-2330, DOI 10.1039/c5sc03115j.

[60] S. Sreeramulu, C. Richter, T. Kuehn, K. Azzaoui, M. J. J. Blommers, *et. al*, *Journal of Biomolecular NMR* **2020**, *74*, 555-563, DOI 10.1007/s10858-020-00327-9.

[61] H. Berg, M. A. Wirtz Martin, A. Niesteruk, C. Richter, S. Sreeramulu, H. Schwalbe, *Journal of Visualized Experiments* **2021**, *172*, e62262, DOI 10.3791/62262.

[62] A. J. Robertson, J. Ying, A. Bax, *Magnetic Resonance* **2021**, *2*, 129ï 138.

[63] R. A. Laskowski, J. Jabůrek, L. Pravda, R. S. Vaškov§, J. M. Thornton, *Protein Science* **2018**, *27*, 129ï 134.

[64] A. Grosdidier, V. Zoete, O. Michielin, *Journal of Computational Chemistry* **2011**, *32*, 2149ï 2159.

[65] A. Grosdidier, V. Zoete, O. Michielin, *Nucleic Acids Research* **2011**, *39*, W270ï W277.

[66] N. Kubatova, N. S. Qureshi, N. Altincekic, R. Abele, J. K. Bains, *et. al*, *Biomolecular NMR Assignments* **2021**, *15*, 65ï 71.

[67] D. Mendez, A. Gaulton, A. P. Bento, J. Chambers, M. De Veij, *et. al*, *Nucleic Acids Research* **2019**, *47*, D930ï D940.

[68] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, *et. al*, *Nucleic Acids Research* **2021**, *49*, D1388ï D1395.

[69] X. Chen, C. H. Reynolds, *Journal of Chemical Information and Computer Sciences* **2002**, *42*, 1407ï 1414.

[70] M. Kuzikov, E. Costanzi, J. Reinshagen, F. Esposito, L. Vangeel, *et. al*, *ACS Pharmacology & Translational Science* **2021**, *4*, 1096ï 1110.

[71] S. Jeon, M. Ko, J. Lee, I. Choi, S. Y. Byun, S. Park, D. Shum, S. Kim, *Antimicrobial Agents and Chemotherapy* **2020**, *64*, DOI 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.

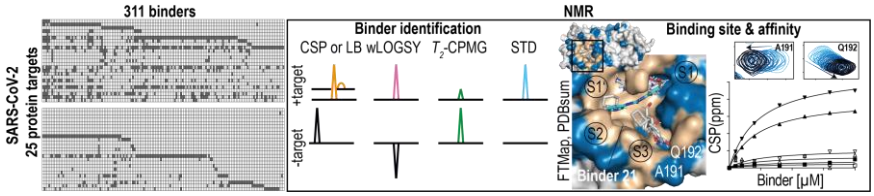
[85] nsp3b·GS-441524, H. Berg, 2022, DSI-PL_COVID19-NMR_nsp3b_GS-441524, BMRbig, bmrbig52, https://bmrbig.bmrb.io/released/bmrbig52.

[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.

Entry for the Table of Contents



Using a fragment-based screening strategy by NMR, we identified 311 small molecule binders of 25 SARS-CoV-2 proteins, thus expanding the previously unexplored chemical and target space. Further, using experimental and bioinformatic analysis we identify potential binding sites. This comprehensive data would greatly assist medicinal chemistry efforts even beyond COVID-19.

Institute and/or researcher Twitter usernames: ((optional))