

## Covid Virus

How to cite: *Angew. Chem. Int. Ed.* **2021**, *60*, 19191–19200

International Edition: doi.org/10.1002/anie.202103693

German Edition: doi.org/10.1002/ange.202103693

## Exploring the Druggability of Conserved RNA Regulatory Elements in the SARS-CoV-2 Genome

Sridhar Sreeramulu<sup>+</sup>, Christian Richter<sup>+</sup>, Hannes Berg, Maria A. Wirtz Martin, Betül Ceylan, Tobias Matzel, Jennifer Adam, Nadide Altincekic, Kamal Azzaoui, Jasleen Kaur Bains, Marcel J. J. Blommers, Jan Ferner, Boris Fürtig, Michael Göbel, J. Tassilo Grün, Martin Hengesbach, Katharina F. Hohmann, Daniel Hyman, Bozana Knezic, Jason N. Martins, Klara R. Mertinkus, Anna Niesteruk, Stephen A. Peter, Dennis J. Pyper, Nusrat S. Qureshi, Ute Scheffer, Andreas Schlundt, Robbin Schnieders, Elke Stirnal, Alexey Sudakov, Alix Tröster, Jennifer Vögele, Anna Wacker, Julia E. Weigand, Julia Wirmer-Bartoschek, Jens Wöhnert, and Harald Schwalbe\*

**Abstract:** SARS-CoV-2 contains a positive single-stranded RNA genome of approximately 30000 nucleotides. Within this genome, 15 RNA elements were identified as conserved between SARS-CoV and SARS-CoV-2. By nuclear magnetic resonance (NMR) spectroscopy, we previously determined that these elements fold independently, in line with data from in vivo and ex-vivo structural probing experiments. These elements contain non-base-paired regions that potentially harbor ligand-binding pockets. Here, we performed an NMR-based screening of a poised fragment library of 768 compounds for binding to these RNAs, employing three different <sup>1</sup>H-based 1D NMR binding assays. The screening identified common as well as RNA-element specific hits. The results allow selection of the most promising of the 15 RNA elements as putative drug targets. Based on the identified hits, we derive key functional units and groups in ligands for effective targeting of the RNA of SARS-CoV-2.

## Introduction

Since early 2020, enormous scientific efforts are geared towards antiviral treatment for SARS-CoV-2 (SCoV-2).

Impressive advancements are reported in vaccine development. Furthermore, experimental approaches are being actively pursued in drug repurposing and in design and synthesis of new drugs. Such studies are supported by virtual screening and molecular docking campaigns.<sup>[1–7]</sup> Compound libraries with more than 10<sup>6</sup> molecules have been screened virtually. So far, most efforts focus on targeting proteins to inhibit viral propagation, while few attempts have been reported for direct targeting of the large viral RNA genome.<sup>[8]</sup> This focus on proteins as drug targets comes despite the fact that structured RNA elements have been increasingly recognized as drug targets in recent years,<sup>[9–16]</sup> including SARS-CoV (SCoV). In SCoV, for example, the pseudoknot (PK) element involved in ribosomal frameshifting was identified as a potent drug target.<sup>[17]</sup> The abundance of viral RNA in infected host cells opens a further opportunity for pharmaceutical intervention: for coronaviruses in general and for SCoV-2 infected cells in particular, viral RNA accounts for approximately two thirds of the total RNA.<sup>[18]</sup> In fact, the transcriptome of SCoV-2 is discussed to contain numerous potential druggable sites.<sup>[19,20]</sup> Currently, however, only few experimental screenings for ligands directly targeting the

[\*] Dr. S. Sreeramulu,<sup>[+]</sup> Dr. C. Richter,<sup>[+]</sup> H. Berg, M. A. Wirtz Martin, B. Ceylan, T. Matzel, J. Adam, N. Altincekic, J. K. Bains, Dr. J. Ferner, Dr. B. Fürtig, Prof. Dr. M. Göbel, J. T. Grün, Dr. M. Hengesbach, K. F. Hohmann, D. Hyman, B. Knezic, J. N. Martins, K. R. Mertinkus, Dr. A. Niesteruk, D. J. Pyper, Dr. N. S. Qureshi, Dr. U. Scheffer, Dr. R. Schnieders, E. Stirnal, A. Sudakov, A. Tröster, Dr. A. Wacker, Dr. J. Wirmer-Bartoschek, Prof. Dr. H. Schwalbe  
Institute for Organic Chemistry and Chemical Biology, Center for Biomolecular Magnetic Resonance (BMRZ), Johann Wolfgang Goethe-University  
Max-von-Laue-Str. 7 + 9, 60438 Frankfurt/M. (Germany)  
E-mail: schwalbe@nmr.uni-frankfurt.de

Dr. A. Schlundt, J. Vögele, Prof. Dr. J. Wöhnert  
Institute for Molecular Biosciences, Center for Biomolecular Magnetic Resonance (BMRZ), Johann Wolfgang Goethe-University  
Max-von-Laue-Str. 7 + 9, 60438 Frankfurt/M. (Germany)

S. A. Peter, Dr. J. E. Weigand  
Department of Biology, Technical University of Darmstadt  
Schnittspahnstr. 10, 64287 Darmstadt (Germany)

Dr. K. Azzaoui, Dr. M. J. J. Blommers  
Saverna Therapeutics  
Pumpmattenweg 3, 4105 Biel-Benken (Switzerland)  
Dr. N. S. Qureshi  
Present address: EMBL Heidelberg  
Meyerhofstr. 1, 69117 Heidelberg (Germany)

[\*] These authors contributed equally to this work.

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:  
https://doi.org/10.1002/anie.202103693.

© 2021 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

structured RNA elements in SCoV-2 have been reported.<sup>[8,21,22]</sup>

Here, we report a holistic NMR-based screening campaign investigating previously identified structured RNA elements.<sup>[23]</sup> We determine the druggability of these regulatory RNA elements by NMR-based screening,<sup>[24]</sup> using a previously implemented workflow.<sup>[25]</sup> We use a well-characterized,<sup>[26]</sup> highly diverse fragment library poised for follow-up chemistry. The screen of 768 fragments conducted here yielded a high hit rate, despite the stringent hit definition from three independent NMR experiments, demonstrating the general druggability of the viral RNA targets.

Previously, we identified secondary structures in these RNA targets (Supporting Information, Figure S1) by NMR spectroscopy.<sup>[23]</sup> Our selection of 15 SCoV-2 RNA targets from the genome was based on functional and structural conservation among Betacoronaviruses.

In the context of screening the viral RNA targets, we briefly describe the functional roles of the 15 RNA elements. In the 5'-genomic end, stem loop 1 (5\_SL1) is associated with viral mRNA escape from Nsp1-mediated translational shutdown.<sup>[27]</sup> 5\_SL2 + 3 harbors a stringently conserved stem-loop (Supporting Information, Figure S1)<sup>[28]</sup> and the transcription regulatory sequence regulating the production of subgenomic RNAs that is embedded within a second, more labile stem-loop. It further comprises a nucleotide stretch involved in long-range interactions.<sup>[29]</sup> The stem-loop 5\_SL4 is indispensable for replication and contains a short upstream open reading frame (uORF).<sup>[30]</sup> We dissected the large SL5 element into the three sub-elements 5\_SL5a, 5\_SL5b + c and 5\_SLstem, the latter of which contains the ORF1 start codon. The functions of SL6, 7 and 8 are less well defined, but their high degree of structure conservation between SCoV and SCoV-2 suggests regulatory roles as well.

The frameshifting region is a well-established target for binding of ligands of low molecular weight.<sup>[17,31,32]</sup> We thus screened the attenuator hairpin (att HP) and PK. Within the 3'-UTR, a putative structural switch is supposed to play a role in the initiation events of (-)-strand synthesis.<sup>[33]</sup> The individual elements forming this elaborate RNA architecture are 3\_SL1, 3\_SL2 and 3\_SL3base. Furthermore, the hyper-variable region of the 3'-UTR harbours the 3\_s2m motif, which is recurring among a variety of RNA viruses, and seems to be involved in coronavirus pathogenicity.<sup>[34]</sup> For each of these 15 RNA elements chosen for fragment-based screening, their existence and structural integrity within the SCoV-2 genome was demonstrated by us<sup>[23]</sup> and others.<sup>[35-37]</sup> Correct folding prior to screening experiments was ensured as detailed in the Material and Methods section.

## Results and Discussion

We analyzed the non-canonical RNA elements (Supporting Information, Table S1), revealing four different classes of non-canonical segments:

(i) Capping loops (blue and grey in Supporting Information, Figure S1, Table S1): In general, loops are ca. 6 nts long and uridine enriched, while adenosine depleted, with

respect to the genomic nucleobase composition (Supporting Information, Table S3). We here distinguish the capping loops by their different degrees of structure, with the loops of 5\_SL2 and 5\_SL5c adopting typical tetraloop conformations.<sup>[28]</sup> A hexaloop 5'-UUUCGU-3' found in 5\_SL5a also caps 5\_SL5b. It is interesting to note that one of the few viral mutations described since early 2020 includes a C to U mutation at position 4 in almost all SCoV-2 strains.<sup>[38]</sup> The loops of the PK, 3\_s2m and especially 3\_SL2 are exceptionally large (9, 9 and 13 nts, respectively). The 3\_SL2 loop has been proposed to engage in a 3'-pseudoknot structure.<sup>[33,39]</sup>

- ii) Bulges/internal loops (green in Supporting Information, Figure S1, Table S1). Bulges and internal loops dictate the relative orientations of helical segments. They are key elements in RNA tertiary architecture, and are often interaction sites for proteins, RNAs, metal ions or small molecules. Frequently, the helical segments of the SCoV-2 RNA elements are interrupted either by a single unpaired nucleotide or by stretches of unpaired residues. We classify single and tandem mismatched residues separately (under iii). With this classification, the remaining internal loops are asymmetric. The average size of internal loops is ca. 7 nts, and only 5\_SL6, 3\_SL1 and 3\_s2m contain larger internal loops. In contrast to the capping loops, the internal loops and bulges are uridine depleted and adenosine enriched with respect to the genomic base composition (Supporting Information, Table S3). Most bulges consist of only one nucleotide, but up to four bulged-out nucleotides are found.
- (iii) Non-canonical base-pairs/mismatches (orange in Supporting Information, Figure S1). Next to the well-known non-canonical G-U/U-G base-pair, which disrupts A-helical conformation only marginally,<sup>[40]</sup> the RNA elements show a variety of mismatched residues. Interestingly, tandem pyrimidine-pyrimidine (Y-Y) base pairs are identified in 3\_SL1 and 3\_SL3base by their base-pair connectivities, whereas isolated Y-Y mismatches in 5\_SL4, 5\_SL5a, 5\_SL5b and 5\_SL8 do not form base pairs with detectable H-bonds.
- (iv) Three-helix junctions (pink in Supporting Information, Figure S1). Single-stranded regions connecting three helices are found in the PK and 3\_SL3base and are structurally flexible. Helix junctions are generally considered high-potential binding pockets for small molecules, reminiscent to bacterial riboswitches.<sup>[41]</sup>

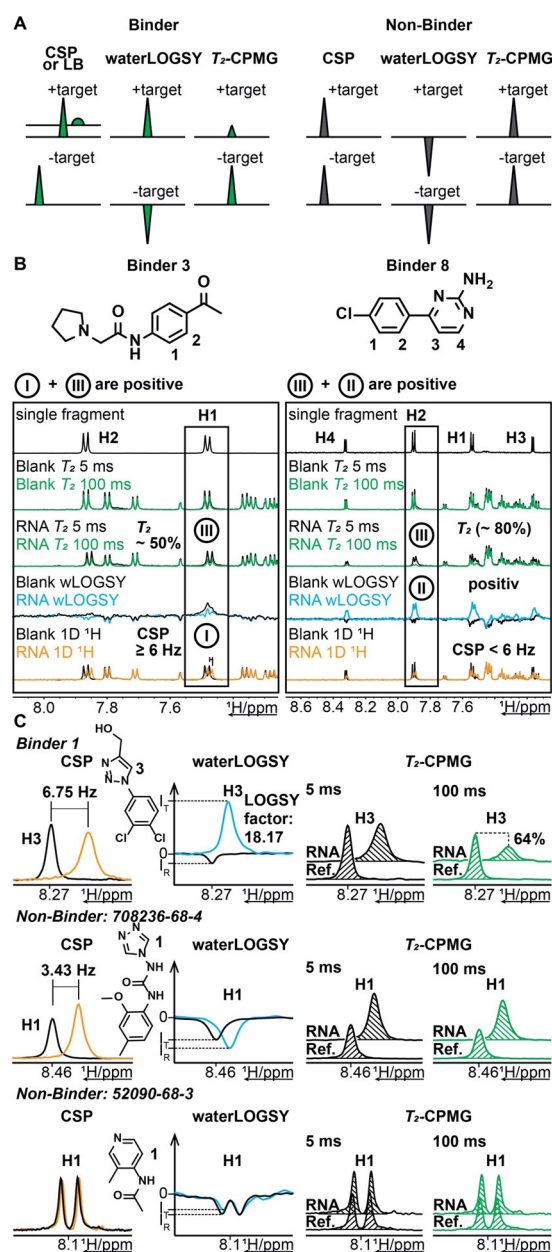
Supporting Information, Table S1 gives an overview of loop and bulge sequences for the 15 RNAs. Column 3 shows the frequency of a sequence in the SCoV-2 genome<sup>[42]</sup> compared to its abundance in predicted structured regions therein.<sup>[20]</sup> Column 4 correlates each sequence to the number of actual loops or bulges predicted (or, if deviating, experimentally determined<sup>[23]</sup>) for this sequence. To estimate the relevance of a given loop or bulge sequence, its conservation (resistance to mutation) is stated in column 5.<sup>[43]</sup> Supporting Information, Table S1 also provides information on unique loop or bulge motifs within the structured RNA genome as defined by Rangan et al.<sup>[20]</sup>

The DSI-poised library (DSI-PL, Supporting Information, Table S2. DSI-PL-768\_Ligands) is the second generation of the initially developed poised library,<sup>[44]</sup> composed of 768 highly diverse and poised fragments specifically designed to facilitate easy downstream synthesis.

This library has already been successfully used to screen the main protease nsp5 from SCoV-2.<sup>[45]</sup> Ligand-observed NMR-based screening is a versatile method routinely adapted in the discovery of fragments binding to a target. Screening individual fragments for individual targets is, however, time-consuming from a library with hundreds of fragments. Instead, evaluation of binding in mixtures ranging from 5 to 20 compounds is highly advantageous. One of the prerequisites in designing an ideal mixture is minimal signal overlap between the <sup>1</sup>H-NMR signals of the individual compounds so that spectral changes in the presence of the RNA target can be associated to a single compound. Previously, we performed molecular clustering analysis of the DSI-PL fragment library. We found a total of ca. 400 distinguished chemical clusters, from which ca. 200 clusters contained singletons, suggesting a high chemical diversity of the library.<sup>[26]</sup> We hypothesized that the chemical diversity of ligands might also result in significantly diverse <sup>1</sup>H-NMR spectra of the fragments in each sub-library mix. In total, we generated 64 mixtures with 12 fragments (Supporting Information, Figure S2C) each from the 768 compounds of the DSI-PL. Indeed, the <sup>1</sup>H-NMR spectra of each of the randomly chosen 12 fragments is significantly different from one another (Supporting Information, Figure S2A and B; black spectra), so that the signals of the individual compounds could be traced back in the <sup>1</sup>H-NMR spectrum (Supporting Information, Figure S2A and B; blue spectra).

We applied ligand-observed <sup>1</sup>H-NMR experiments against the 15 RNAs ranging in size from 29–90 nts depicted in Supporting Information, Figure S1 and five additional larger RNAs with sizes between 118–472 nts that comprise several RNA elements (Supporting Information, Table S1 and Files “1–21\_Hits” detailing all experimental hits). In screening experiments, changes in the <sup>1</sup>H signals of the ligand in the presence and absence of the RNA served as readout for binding. After optimization (Supporting Information, Figure S3), we recorded <sup>1</sup>H 1D spectra at *T* = 293 K to detect chemical shift perturbations (CSPs) or line broadening (LB) (Figure 1 A), waterLOGSY<sup>[46,47]</sup> and *T*<sub>2</sub>-relaxation<sup>[48]</sup> experiments (Supporting Information, Table S4). Previously, these experiments have successfully been utilized to identify small molecule binders for RNA.<sup>[49,50]</sup> NMR signals of each of the 64 fragment mixtures were monitored at a ratio of each ligand to one RNA target of 20:1. 190 μl of a 10 μM RNA in screening buffer (25 mM KPi, 50 mM KCl, pH 6.2) was manually pipetted into 3 mm NMR tubes. 10 μl of the fragment mixture was added using a pipetting robot to a final concentration of 200 μM for each fragment. In waterLOGSY, the signal from a bound fragment is positive while the signal of a free ligand is negative. *T*<sub>2</sub>-relaxation-based experiments take advantage of the fact that large biomolecular targets including all RNAs studied here relax faster than small fragments (Figure 1 A).

For identifying binders within the mixtures, we first compared spectra from all three of the above experiments

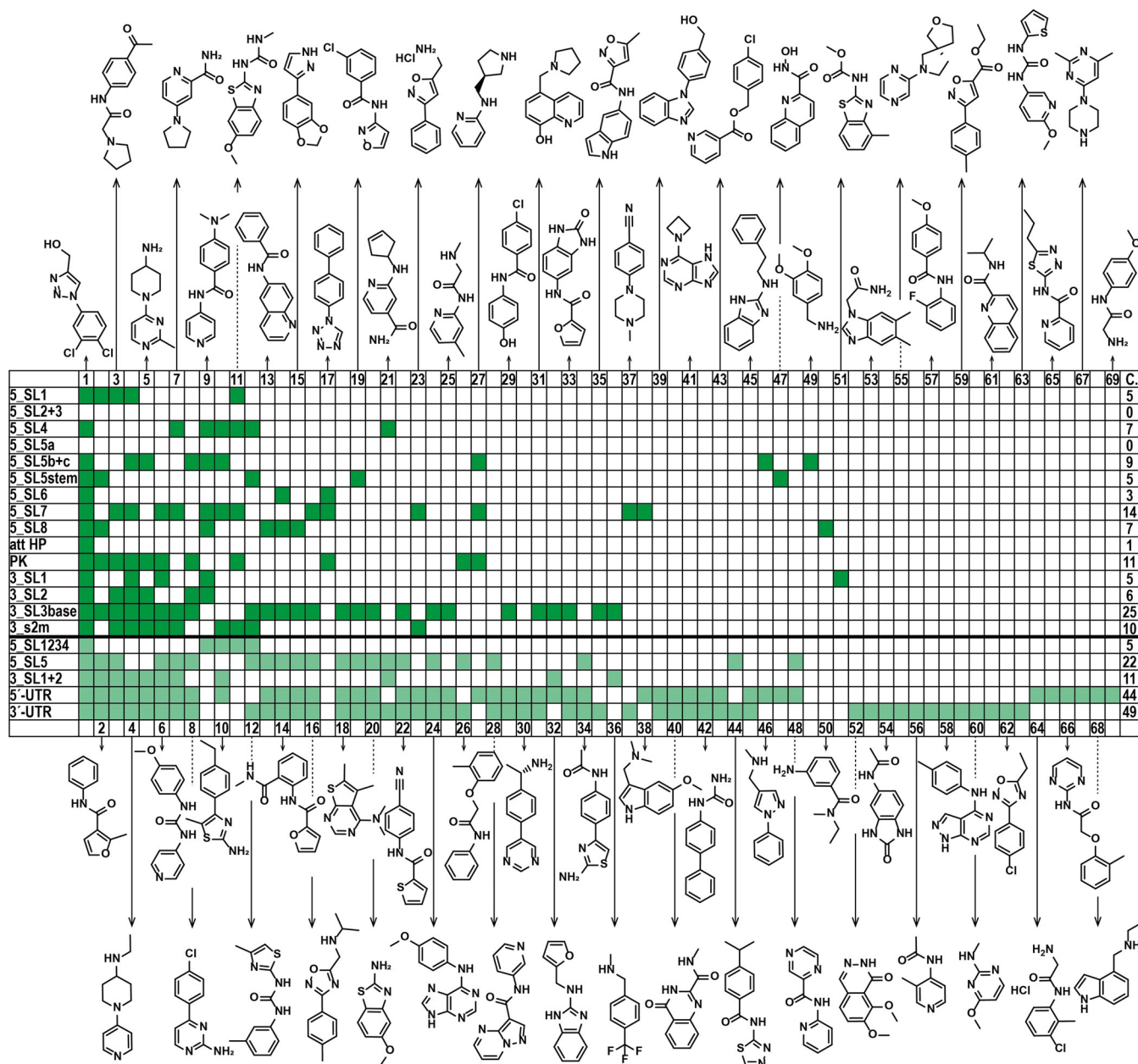


**Figure 1.** NMR-based fragment screening and hit identification: A, Schematic representation of the experiments and criteria used to identify binders/ non-binders. B, Chemical structures of Binder 3 and Binder 8 along with an overlay of their corresponding NMR spectra [<sup>1</sup>D <sup>1</sup>H (i), waterLOGSY (ii) and *T*<sub>2</sub>-CPMG (iii); (5 ms and 100 ms)] of the mixture in absence (blank) and presence of RNA. The top <sup>1</sup>D <sup>1</sup>H-spectrum corresponds to the single fragment used for chemical shift deconvolution in the mixture. The inset highlights the signal corresponding to the hit used for analysis. Visual inspection of the overlay indicates that Binder 3 displays CSP (≥ 6 Hz) and *T*<sub>2</sub> reduction of ca. 50%. Binder 8 shows no CSP, but a clear sign change in waterLOGSY and *T*<sub>2</sub>-reduction of ca. 80%. C, <sup>1</sup>D <sup>1</sup>H-NMR spectral regions of fragments with strong (top row), weak (middle row) or no binding (lower row) to 3<sub>SL2</sub>. Binder 1 shows a CSP of 6.75 Hz, a positive waterLOGSY signal and a *T*<sub>2</sub>-reduction of 64% suggesting that it binds to 3<sub>SL2</sub> (top row). 708236–68-4 shows a minor CSP of 3.43 Hz, but neither a positive waterLOGSY signal nor a *T*<sub>2</sub>-reduction, indicating either very weak or no binding to 3<sub>SL2</sub> (middle row). 52090–68-3 does not bind to 3<sub>SL2</sub> (lower row of the mixture (Supporting Information, Figure S2, blue spectrum)).



and analysed differences by visual inspection (Figure 1 B). As stringent criteria, CSPs  $\geq 6$  Hz or severe line broadening, sign change in the waterLOGSY or  $\geq 20\%$  decrease of signal intensity in the  $T_2$ -relaxation experiment was used to identify binders. Ligands were assigned as a hit if two of the three criteria were satisfied. For example, binder **3** was defined as a hit by a CSP  $\geq 6$  Hz, reduced signal intensity of  $T_2$  (ca. 50%), while the waterLOGSY effect cannot unambiguously identified (Figure 1 B, left). Similarly, binder **8** qualifies as a hit, showing changes in waterLOGSY and  $T_2$ , but not CSP (Figure 1 B, right). In the second stage of analysis, for the hits thus identified, we quantified CSP, LOGSY effect<sup>[51,52]</sup> and  $T_2$ -reduction ( $Q^{\text{bind}}$ ) in intensity.<sup>[25]</sup>

Previously, several reports attempted to delineate common scaffolds binding to RNA.<sup>[9,53,54]</sup> Our results here show that 40 different fragments are found to bind to the 15 RNA elements and an additional 29 fragments bind to the five multi-element RNAs. In total, we observed 108 binding events. For 5\_SL2+3 and 5\_SL5a, no fragment binding was detected. For three RNAs, we identified hits that only bind to this specific RNA element (5\_SL5b+c, 5\_SL8, 3\_SL1). Furthermore, we identified 48 fragments that recognize more than one target RNA. The range of binding promiscuity ranges from binding of two RNA targets (ligands 35–48, Figure 2) up to 18 out of the 20 screened RNA targets (ligand 1, Figure 2).



**Figure 2.** NMR-based fragment screening identifies 69 fragment hits across the SCoV-2 RNAs. Tabular column summarizing screened SCoV-2 RNA elements and their corresponding hits. Dark green indicates hits for individual RNA elements; light green indicates hits for larger RNAs. The last column (C.) lists the number of binders to the investigated RNA target. Chemical structures of respective fragment hits are shown.

We observed no clear correlation between the number of non-helical residues (those forming loops, bulges and mismatches) in an RNA and the number of hits, although the non-targeted RNAs 5\_SL2+3 and 5\_SL5a belong to the smallest RNAs investigated. Strikingly, 5\_SL7 appears to harbor little non-helical structure, but shows 14 hits. By contrast, 3\_SL1 contains roughly three times as much non-helical space and has only four hits. However, the RNA elements with the most complex structures, 3\_SL3base and PK, are also among those with the highest hit rates (Supporting Information, Table S5).

Inspection of all hits reveals the most significant molecular descriptors distinguishing between binders and non-binders are the number of aromatic rings and the  $sp^3$ -character of the compounds. In general, binders have one to three aromatic rings and less  $sp^3$ -carbon atoms than non-binders (see Supporting Information, Table S6).

We noticed that 30 out of 40 fragments have a modular architecture: They consist of two functional units, which are mainly non-, mono- or di-substituted rings of the size 5 and 6 (Scheme 1) and connected by four different types of linker units. The mostly aromatic or heteroaromatic core motifs are often modified with a small set of substituents, which are key elements for generating affinity towards the RNA target. The remaining 10 fragments contain a single functional unit (class 1), often substituted with a functional group. Analysis of the additional 29 hits for which binding is detected only to the larger RNA elements are given in the Supporting Information, Figure S5.

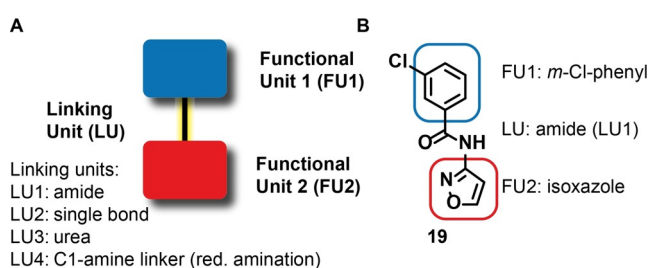
Within the functional units, six functional groups are frequently found (Schemes 1 and 2). In the majority of hits, the two functional units are connected either by an amide bond or by a single bond. We further find urea and carbamates as linkers. They contain hydrogen bond donors and acceptors. Additionally, functional groups are often linked by a C1-linker resulting from reductive amination. Furthermore, similar simultaneous donor and acceptor properties are found in the six-membered (pyridine-acetamide, pyrimidine-2-amine) as well as five membered heterocyclic aromatic rings (2-amine-thiazole, 2-amine-imidazole). Primary, secondary and tertiary amine groups are often found and their categorization into functional group or linker unit is somewhat arbitrary. Some of these amines have  $pK_a$  values that allow for a change in protonation state under physio-

logical conditions, such as imidazole with a  $pK_a$  of 6.95. The classification of the fragments into functional units (see Scheme 2) leads to 26 different functional units (Supporting Information, Figure S6).

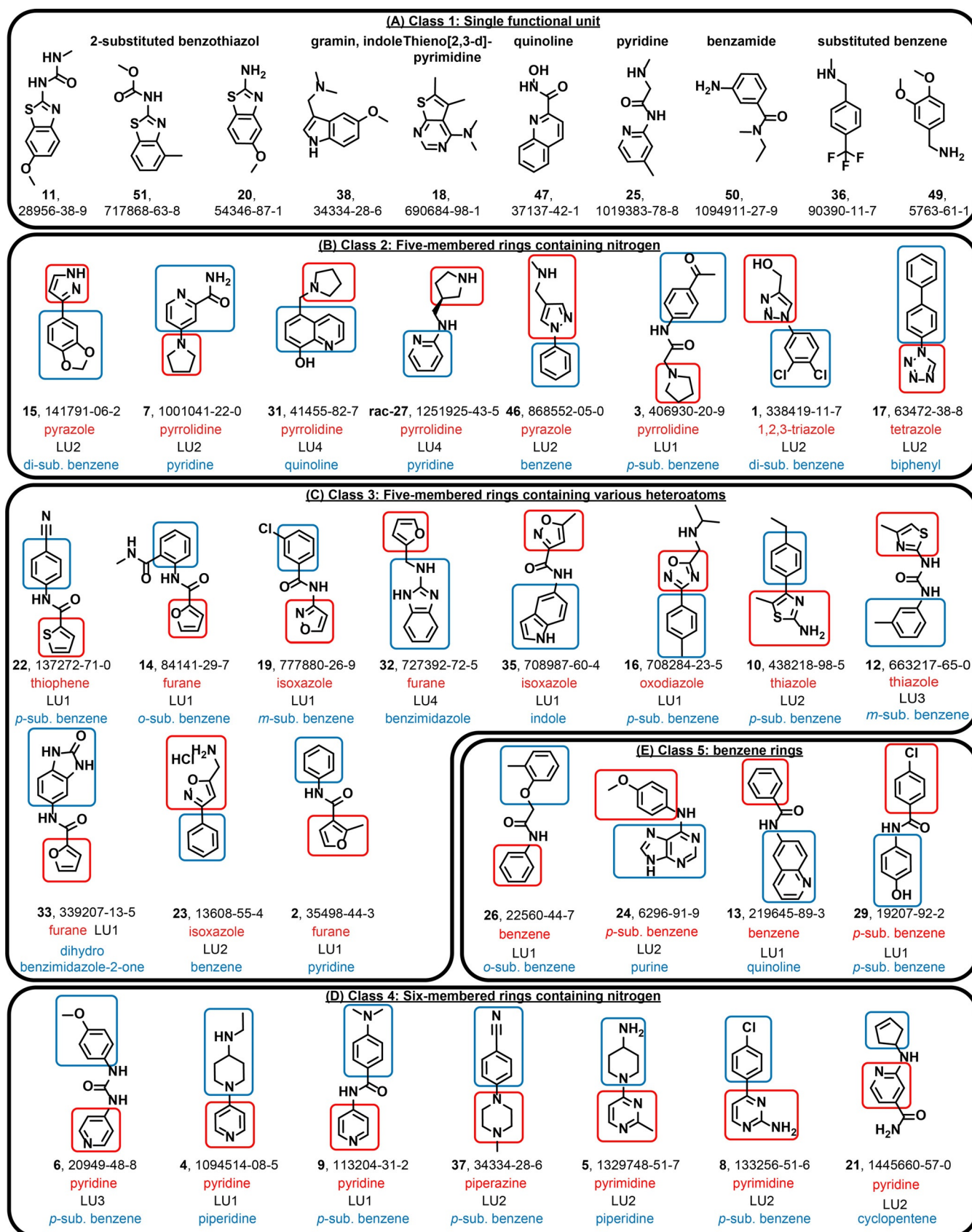
Next to the analysis of binders, comparison to functionally closely related fragments that do not bind is insightful (Scheme 3). Neither fragments containing carboxylic (842971-05-5, for compound identity, Supporting Information, Table S2. DSI-PL-768\_Ligands) nor benzoic acid (3303-18-2, 1152510-62-7, 693776-70-4) functional groups, nor sulfone (24092-75-9) or sulfamide (1423029-76-8, 1388691-56-2), nor fragments with hydroxyl groups (74548-62-2, 1849283-80-2), and none of the 17 acetamides within DSI-PL are found among the hits. The library contains not only the para-substituted pyridine derivative 6, but also ortho- and meta-substituted pyridines 2000 55-7 and 313386-33-3 (Scheme 3A), but only para-substituted pyridines bind. Only thiophene-containing fragment 22, but not methyl-substituted thiophenes (717873-31-9 and 445007-73-8 shows binding to three RNA targets. A particularly intriguing example of surprisingly high selectivity for binding over not binding are four triazole-containing ligands (Scheme 3B). While triazole 1 is the most promiscuous binder, the closely related ligands 133902-66-6, 1099631-80-7, and 2322927-70-6 do not bind at all.

Hits containing two functional units can be classified into four additional classes. Class 2: five-membered heterocyclic rings containing a single nitrogen heteroatom, Class 3: five-membered rings with several heteroatoms, in particular nitrogen, oxygen and sulfur, Class 4: six-membered heterocyclic rings containing nitrogen atom(s), and Class 5: substituted and unsubstituted benzenes. In class 1, we find three differently 2-substituted benzo-thiazoles with either an amino group (compound 20) or two different linker units attached (methyl urea linker (11) and methyl carbamate (51)). Related to the benzothiazoles are the fragment hits thieno[2,3-*d*]pyrimidine (18) and the C7-methoxy-substituted gramin (37). A last single functional unit-containing fragment is quinoline, substituted at C2 with a hydroxamic acid functional group (47). Furthermore, we find two methylene-amino-substituted benzenes (36 and 49).

For a subset of hits (Table 1) that bound PK and 3\_s2m, we determined the dissociation constants using ligand-observed  $^1\text{H}$  NMR-based titrations. In general, the determined dissociation constants  $K_D$  for the fragment hits ranged between 64 and 1318  $\mu\text{M}$  (Table 1 and Supporting Information, Figure S7). Fragment 4 bound with highest affinity both towards PK and 3\_s2m (Figure 3A). In addition to the quantification of the  $T_2$ -reduction obtained from the primary screen in mixtures at 20-fold excess of ligand over RNA ( $[\text{RNA}] = 10 \mu\text{M}$ ) (Supporting Information, Figure S4), we determined the transverse relaxation time  $T_2$  for the binding ligands at 100  $\mu\text{M}$  concentration at an RNA concentration of 35  $\mu\text{M}$ . Within the subset of the fragment hits, the ligands with the highest degree of  $T_2$ -reduction also showed the strongest affinity. For ligands with a  $K_D < 100 \mu\text{M}$ , binding affinity correlated with the  $T_2$ -reduction ( $Q^{\text{bind}}$ ). No correlation was, however, observed for  $K_D > 100 \mu\text{M}$ . This finding suggests that  $T_2$  is a qualitative indicator for binding strength of

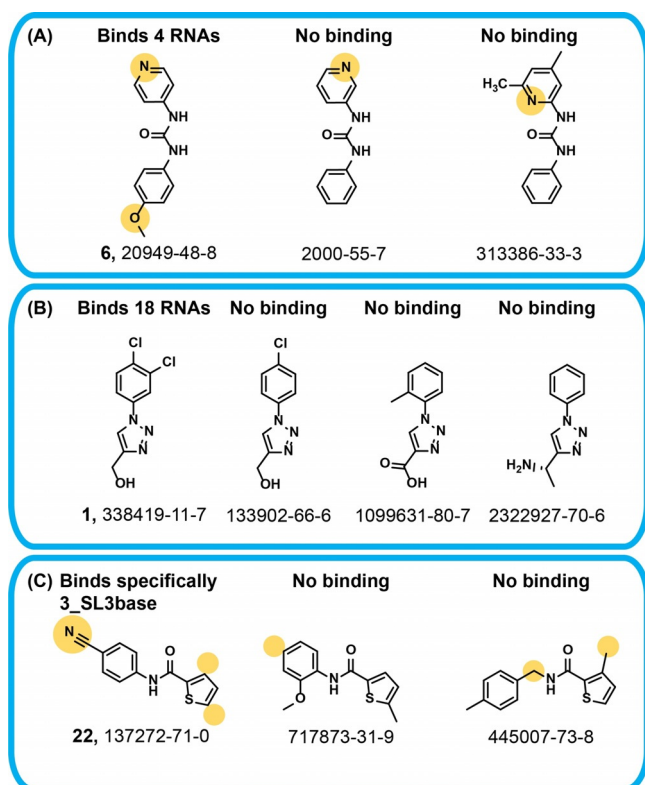


**Scheme 1.** Molecular descriptors and chemical diversity of RNA binders in the DSI-poised library. About 75 % of the fragments showing binding to SCoV-2 RNA contain two functional units, connected by four different linking units. In case of LU3, the urea functional group likely contributes to binding affinity.



Scheme 2. Classification into five classes of the 40 ligands that bind to the 15 regulatory RNAs. LU = linking unit.





**Scheme 3.** Comparison of binding and non-binding fragments with closely related chemical structure.

fragments to one specific target for binding in fast exchange and with a low micromolar  $K_D$ .

As a follow-up and in the quest for finding molecules with improved affinity, we used the fragment hits obtained for five RNAs (3\_s2m, 3\_SL1, 5\_SL4, PK and 5'-UTR) to identify commercially available potential binders and obtained one compound (D01) that bound to both PK and 3\_s2m in single-digit  $\mu\text{M}$  affinity (Figure 3B). At the same time, D01 showed no saturable binding towards control RNAs including the most stable RNA UUCG-tetraloop (Supporting Information, Figure S8). For fragments binding with  $K_D$  above 200  $\mu\text{M}$ , it is our experience<sup>[55,56]</sup> that ligand binding sites can often not be detected even at maximum possible excess of ligand (limited by solubility) over RNA for a number of reasons. In fact, we

show this lack of binding site detection also for one of the fragments that binds with a  $K_D$  of 46  $\mu\text{M}$  (fragment 8, 133256-51-6) in the Supporting Information, Figure S9A. By contrast, for the high-affinity binder D01, we provide experimental mapping of its binding site to PK (Figure 3), which is in line with the binding site determined by in-line probing (Supporting Information, Figure S9C). NMR titration experiments were performed using  $^1\text{H}$ ,  $^{15}\text{N}/^{13}\text{C}$ -HSQC of PK. In these experiments, many of the signals showed either severe line-broadening or CSPs strongly indicating binding. Mapping of these signals onto the cryo-electron microscopy structure of PK shows that D01 mainly binds to the stem 3 (Figure 3C and Supporting Information, Figure S9B).

## Conclusion

We conducted here an NMR-based fragment screening against all regulatory RNA elements of the viral genome of SCoV-2. Our results show that the fragments of the DSI-PL library can cover almost the entire structural space of the viral RNA genome. Thus, the SCoV-2-RNA genome can be targeted by ligands of low molecular weight. We detect that several fragments already show specificity at this very first stage of drug development (Supporting Information, Figure S6). Assuming that increasing structural complexity enhances the potential to achieve specificity for small molecules,<sup>[32,57]</sup> among the 15 RNAs, PK and 3\_SL3base represent the most promising targets for follow-up development of lead molecules.

The current state-of-the-art discovery of small molecules targeting RNAs (RNA drug discovery) and the current status of this field have been described by Juru and Hargrove,<sup>[53]</sup> by Meyer et al<sup>[9]</sup> and by Warner et al.<sup>[57]</sup> The screening campaign conducted here supports previous reports that identified the existence of a privileged chemical space for targeting RNA.<sup>[58]</sup> Delineating RNA target space into secondary structure motifs, definable from studies of isolated RNA elements, finds further support by a recent paper from the Das group that provides FARFAR models for all conserved viral RNA elements.<sup>[59]</sup>

Here, we use a fragment library that has previously been used to screen proteins,<sup>[44]</sup> but also 14 different RNAs, and five DNAs.<sup>[55]</sup> We use NMR as primary screen because it not

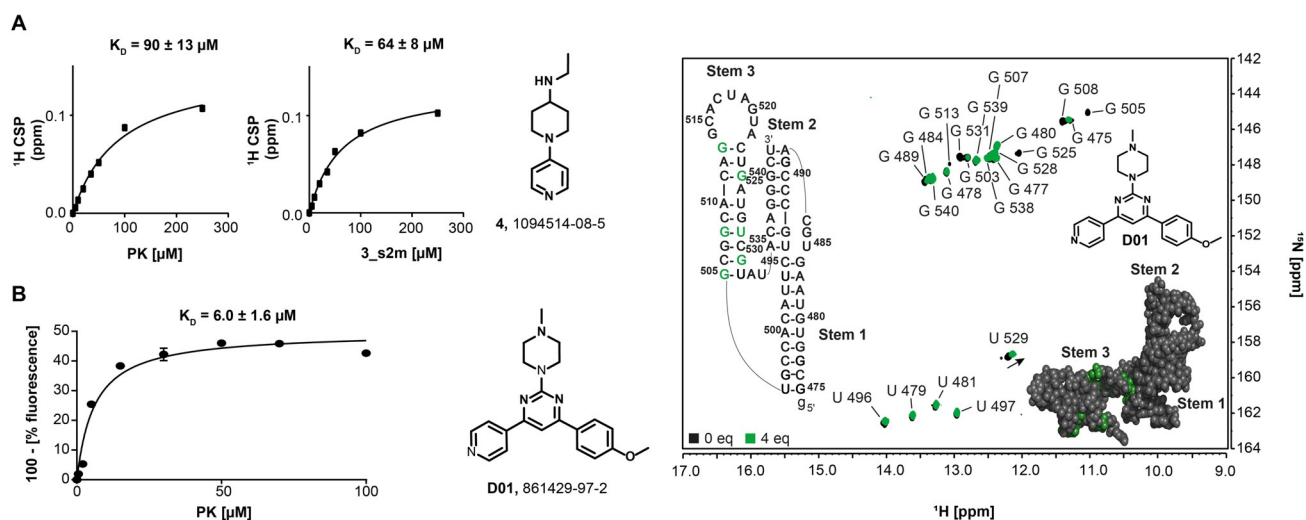
only provides information on chemical purity of both ligand and RNA target, but also on affinity and binding site. Identifying RNA fragments with affinities between 60–400  $\mu\text{M}$  allows us to identify binders containing similar structural motifs. A thus identified binder, D01, binds with 6  $\mu\text{M}$  affinity to PK and 3\_s2m, allowing fragment linking approaches.<sup>[32,60]</sup>

It is apparent that all ligands contain at least one, but most often two aromatic or heteroaromatic

**Table 1:** Affinities and  $T_2$  relaxation times of the RNA binders.

ID	PK		3_s2m			
	$T_2$ [Q <sup>bind</sup> ] primary screen <sup>[a]</sup>	$T_2$ [ms] determined <sup>[b]</sup>	$K_D$ [ $\mu\text{M}$ ] <sup>[c]</sup>	$T_2$ [Q <sup>bind</sup> ] primary screen <sup>[a]</sup>	$T_2$ [ms] determined <sup>[b]</sup>	$K_D$ [ $\mu\text{M}$ ] <sup>[c]</sup>
4	54	83 ± 6	90 ± 13	48	125 ± 12	64 ± 8
5	38	68 ± 12	160 ± 6	15	98 ± 15	206 ± 5
3	39	84 ± 6	336 ± 18	39	106 ± 7	212 ± 30
23	n.a.	n.a.	n.a.	26	131 ± 10	678 ± 25
2	31	87	456 ± 33	n.a.	n.a.	n.a.
D01	n.a.	n.a.	6 ± 1.6	n.a.	n.a.	4.4 ± 2.3

[a] Primary screen; ligand in mixtures ([RNA]:[Ligand] = 1:20), [RNA] = 10  $\mu\text{M}$ . [b] Single ligand measurements ([RNA]:[Ligand] = 1:3), [RNA] = 35  $\mu\text{M}$ . [c]  $^1\text{H}$  NMR-based ligand observed titrations [Ligand 100  $\mu\text{M}$ ; 0 to 250  $\mu\text{M}$ ].



**Figure 3.** Affinities of binders and mapping of binding sites in SCoV-2 RNAs. A) NMR-based (ligand-observed; protons chosen to follow CSP are indicated) titration curves for the interaction of fragment **4** with PK and 3\_s2m. B) Fluorescence-based titration curve and dissociation constant of **D01** with PK. C) Interaction of **D01** with PK monitored by RNA-observed NMR.  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC overlay of PK alone and in presence of four equivalents of **D01** (green). Addition of **D01** to PK results in CSPs or line broadening beyond detection. Affected nucleotides are mapped onto the PK secondary and 3D structures (PDB:7O7Z<sup>[61]</sup>).

rings (Supporting Information, Figure S6). Heteroaromatic rings with two condensed rings including indoles, purines, quinolones (class I, Scheme 2) are found but not particularly enriched. Many of the hits do not have large planar moieties nor very basic sites. For example, **1** as tertiary amine is strongly basic, 2-aminopyridine has a  $\text{p}K_{\text{a}}$  of around 3.5. By contrast, 1,2,3-triazole is already not very basic but shows very specific binding (Scheme 3). Further, comprehensively comparing the distribution of functional units between the hits and the non-hits across the library suggests that the hits are enriched with pyrimidines and benzimidazoles (Supporting Information, Table S7). Thus, the absence of trivial functional units in the DSI-PL appears particularly striking: fusion of monocyclic aromatic or heteroaromatic ring systems with H-bond donors or -acceptors seems sufficient to reach binding.

The future analysis of our screening results might open new routes towards binding selectivity. The screening campaign was conducted for 20 different RNA elements. It also included RNAs that combined two or more of the individually screened elements (e.g. 5\_SL1234) and even the entire 5'- and 3'-untranslated regions of SCoV-2 comprising more than 337 nts for each of the constructs. The size range of these targets is markedly different. At this point, exact quantification of the NMR screening experiments relies on approximations including effective overall rotational correlation times of the RNA targets, relative population of free vs. bound ligand, and their on- and off-rates. These assumptions are, however, no longer fulfilled when one compares pools of RNAs with significantly different molecular weight and potential anisotropic tumbling. Within a given target, however, the  $T_2$ -reduction recorded within the NMR screening of all 768 fragments provides a qualitative indication for binding with  $K_{\text{D}}$  values below 100  $\mu\text{M}$  (Table 1).

We analyzed the investigated RNA target space towards sequence-derived properties and derived unique target sites within the SCoV-2 genome that are conserved among

coronaviruses. Exceptionally rare mutations in SCoV-2 during 2020 have maintained the target space for ligands up to now. By analysis of the experimental hits, we can derive privileged RNA target space from a medicinal chemistry perspective. Further, our approach relies on the hypothesis that the viral genome can be specifically targeted even in the presence of human cellular RNA. The herein provided analysis of secondary structure abundance (Supporting Information, Table S1) is restricted to the viral RNA genome and will have to be extended to the entire cell-specific human transcriptome. However, given all these theoretical considerations it is striking to recognize that small molecules have been developed that target splicing<sup>[11,56]</sup> in a highly specific manner.

The results as well as the methodological approach presented here will impact medicinal chemistry approaches but also cellular targeting of SCoV-2 RNA. For example, as immediate follow-up, we currently conduct secondary NMR screens with commercially available compounds using the identified ligands as guide. We consider our broad, systematic and coherent identification of binders to be a significant step forward. The focus of our study is to experimentally validate the druggability of the SCoV-2 genome. In fact, this potential had so far only been predicted from theoretical studies.<sup>[59]</sup> Our efforts extend on-going studies targeting the viral proteome. In conclusion, we establish here the conserved RNA elements as potential space for small molecule targeting towards Coronavirus-specific medication, even beyond SCoV-2.

## Acknowledgements

The work has been conducted in the international consortium of Covid19-NMR (covid19-nmr.de). 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: “Molecular Principles of RNA-based regulation.” 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. We thank Peter Maas (SPECS) for providing the compound D01 as a kind gift. Open access funding enabled and organized by Projekt DEAL.

### Conflict of Interest

The authors declare no conflict of interest.

**Keywords:** Covid19-nmr · fragment screening · NMR spectroscopy · RNA · SARS-CoV-2

- [1] M. Macchiagodena, M. Pagliai, P. Procacci, *Chem. Phys. Lett.* **2020**, *750*, 137489.
- [2] S. Barage, A. Karthic, R. Bavi, N. Desai, R. Kumar, V. Kumar, K. W. Lee, *J. Biomol. Struct. Dyn.* **2020**, 1–18.
- [3] A. Carino, F. Moraca, B. Fiorillo, S. Marchianò, V. Sepe, M. Biagioli, C. Finamore, S. Bozza, D. Francisci, E. Distrutti, et al., *Front. Chem.* **2020**, *8*, 572885.
- [4] M. T. J. Quimque, K. I. R. Notarte, R. A. T. Fernandez, M. A. O. Mendoza, R. A. D. Liman, J. A. K. Lim, L. A. E. Pilapil, J. K. H. Ong, A. M. Pastrana, A. Khan, et al., *J. Biomol. Struct. Dyn.* **2020**, 1–18.
- [5] M. A. White, W. Lin, X. Cheng, *J. Phys. Chem. Lett.* **2020**, *11*, 9144–9151.
- [6] C. Liu, X. Zhu, Y. Lu, X. Zhang, X. Jia, T. Yang, *J. Pharm. Anal.* **2020**, <https://doi.org/10.1016/j.jpha.2020.08.002>.
- [7] J. O. Ogidigo, E. A. Iwuchukwu, C. U. Ibeji, O. Okpalefe, M. E. S. Soliman, *J. Biomol. Struct. Dyn.* **2020**, 1–18.
- [8] H. S. Haniff, Y. Tong, X. Liu, J. L. Chen, B. M. Suresh, R. J. Andrews, J. M. Peterson, C. A. O’Leary, R. I. Benhamou, W. N. Moss, et al., *ACS Cent. Sci.* **2020**, *6*, 1713–1721.
- [9] S. M. Meyer, C. C. Williams, Y. Akahori, T. Tanaka, H. Aikawa, Y. Tong, J. L. Childs-Disney, M. D. Disney, *Chem. Soc. Rev.* **2020**, *49*, 7167–7199.
- [10] M. L. Kelly, C.-C. Chu, H. Shi, L. R. Ganser, H. P. Bogerd, K. Huynh, Y. Hou, B. R. Cullen, H. M. Al-Hashimi, *RNA* **2020**, *rna.076257*.120.
- [11] S. Campagne, S. Boigner, S. Rüdiger, A. Moursy, L. Gillioz, A. Knörlein, J. Hall, H. Ratni, A. Cléry, F. H.-T. Allain, *Nat. Chem. Biol.* **2019**, *15*, 1191–1198.
- [12] Y. Wang, R. R. Rando, *Chem. Biol.* **1995**, *2*, 281–290.
- [13] M. F. Bardaro, Z. Shajani, K. Patora-Komisarska, J. A. Robinson, G. Varani, *Nucleic Acids Res.* **2009**, *37*, 1529–1540.
- [14] D. Raghunathan, V. M. Sánchez-Pedregal, J. Junker, C. Schwegk, M. Kalesse, A. Kirschning, T. Carlomagno, *Nucleic Acids Res.* **2006**, *34*, 3599–3608.
- [15] F. A. Abulwerdi, M. D. Shortridge, J. Sztuba-Solinska, R. Wilson, S. F. J. Le Grice, G. Varani, J. S. J. Schneekloth, *J. Med. Chem.* **2016**, *59*, 11148–11160.
- [16] M. Zeiger, S. Stark, E. Kalden, B. Ackermann, J. Ferner, U. Scheffer, F. Shoja-Bazargani, V. Erdel, H. Schwalbe, M. W. Göbel, *Bioorg. Med. Chem. Lett.* **2014**, *24*, 5576–5580.
- [17] S. J. Park, Y. G. Kim, H. J. Park, *J. Am. Chem. Soc.* **2011**, *133*, 10094–10100.
- [18] D. Kim, J.-Y. Lee, J.-S. Yang, J. W. Kim, V. N. Kim, H. Chang, *Cell* **2020**, *181*, 914–921.
- [19] I. Manfredonia, C. Nithin, A. Ponce-Salvatierra, P. Ghosh, T. K. Wirecki, T. Marinus, N. S. Ogando, E. J. Snijder, M. J. van Hemert, J. M. Bujnicki, et al., *Nucleic Acids Res.* **2020**, *48*, 12436–12452.
- [20] R. Rangan, I. N. Zheludev, R. Das, *RNA* **2020**, *26*, 937–959.
- [21] Z. Martina, H. Christina, L. Le, D.-C. Jesse, Y.-C. Liang, S. M. Christian, A. G. Monaghan, A. A. Kennedy, J. D. Yesselman, R. R. Gifford, et al., *bioRxiv* **2020**, DOI: <https://doi.org/10.1101/2020.12.05.409821>.
- [22] J. A. Kelly, A. N. Olson, K. Neupane, S. Munshi, J. San Emeterio, L. Pollack, M. T. Woodside, J. D. Dinman, *J. Biol. Chem.* **2020**, *295*, 10741–10748.
- [23] A. Wacker, J. E. Weigand, S. R. Akabayov, N. Altincekic, J. K. Bains, E. Banijamali, O. Binas, J. Castillo-Martinez, E. Cetiner, B. Ceylan, et al., *Nucleic Acids Res.* **2020**, *48*, 12415–12435.
- [24] M. Pellecchia, I. Bertini, D. Cowburn, C. Dalvit, E. Giral, W. Jahnke, T. L. James, S. W. Homans, H. Kessler, C. Luchinat, et al., *Nat. Rev. Drug Discovery* **2008**, *7*, 738–745.
- [25] O. Binas, V. de Jesus, T. Landgraf, A. E. Völklein, J. Martins, D. Hymon, H. Berg, J. K. Bains, T. Biedenbänder, B. Fürtig, et al., *ChemBioChem* **2020**, *22*, 423–433.
- [26] S. Sreeramulu, C. Richter, T. Kuehn, K. Azaoui, M. J. J. Blommers, R. Del Conte, M. Fragai, N. Trieloff, P. Schmieder, M. Nazare, et al., *J. Biomol. NMR* **2020**, *74*, 555–563.
- [27] C. Huang, K. G. Lokugamage, J. M. Rozovics, K. Narayanan, B. L. Semler, S. Makino, *PLoS Pathog.* **2011**, *7*, e1002433–e1002433.
- [28] C. W. Lee, L. Li, D. P. Giedroc, *FEBS Lett.* **2011**, *585*, 1049–1053.
- [29] O. Ziv, J. Price, L. Shalamova, T. Kamenova, I. Goodfellow, F. Weber, E. A. Miska, *Mol. Cell* **2020**, *80*, 1067–1077.
- [30] D. Yang, J. L. Leibowitz, *Virus Res.* **2015**, *206*, 120–133.
- [31] K. Neupane, S. Munshi, M. Zhao, D. B. Ritchie, S. M. Ieperuma, M. T. Woodside, *J. Mol. Biol.* **2020**, *432*, 5843–5847.
- [32] T. Hermann, *Wiley Interdiscip. Rev. RNA* **2016**, *7*, 726–743.
- [33] R. Zust, T. B. Miller, S. J. Goebel, V. Thiel, P. S. Masters, *J. Virol.* **2008**, *82*, 1214–1228.
- [34] S. J. Goebel, T. B. Miller, C. J. Bennett, K. A. Bernard, P. S. Masters, *J. Virol.* **2007**, *81*, 1274–1287.
- [35] I. Manfredonia, D. Incarnato, *Biochem. Soc. Trans.* **2020**, *48*, 341–352.
- [36] Z. Miao, A. Tidu, G. Eriani, F. Martin, *RNA Biol.* **2020**, *17*, 1–10.
- [37] J. Zhao, J. Qiu, S. Aryal, J. L. Hackett, J. Wang, *Viruses* **2020**, *12*, 1473.
- [38] J. Navarro Gonzalez, A. S. Zweig, M. L. Speir, D. Schmelter, K. R. Rosenbloom, B. J. Raney, C. C. Powell, L. R. Nassar, N. D. Maulding, C. M. Lee, et al., *Nucleic Acids Res.* **2021**, *49*, D1046–D1057.
- [39] S. J. Goebel, B. Hsue, T. F. Dombrowski, P. S. Masters, *J. Virol.* **2004**, *78*, 669–682.
- [40] N. B. Leontis, J. Stombaugh, E. Westhof, *Nucleic Acids Res.* **2002**, *30*, 3497–3531.
- [41] J. E. Barrick, R. R. Breaker, *Genome Biol.* **2007**, *8*, R239.
- [42] F. Wu, S. Zhao, B. Yu, Y. M. Chen, W. Wang, Z. G. Song, Y. Hu, Z. W. Tao, J. H. Tian, Y. Y. Pei, et al., *Nature* **2020**, *579*, 265–269.
- [43] J. Hadfield, C. Megill, S. M. Bell, J. Huddleston, B. Potter, C. Callender, P. Sagulenko, T. Bedford, R. A. Neher, *Bioinformatics* **2018**, *34*, 4121–4123.
- [44] O. B. Cox, T. Krojer, P. Collins, O. Monteiro, R. Talon, A. Bradley, O. Fedorov, J. Amin, B. D. Marsden, J. Spencer, et al., *Chem. Sci.* **2016**, *7*, 2322–2330.
- [45] A. Douangamath, D. Fearon, P. Gehrtz, T. Krojer, P. Lukacik, C. D. Owen, E. Resnick, C. Strain-Damerell, A. Aimon, P. Ábrányi-Balogh, et al., *Nat. Commun.* **2020**, *11*, 5047.
- [46] C. Dalvit, G. Fogliatto, A. Stewart, M. Veronesi, B. Stockman, *J. Biomol. NMR* **2001**, *21*, 349–359.
- [47] C. Dalvit, *J. Biomol. NMR* **1998**, *11*, 437–444.

- [48] P. J. Hajduk, E. T. Olejniczak, S. W. Fesik, *J. Am. Chem. Soc.* **1997**, *119*, 12257–12261.
- [49] D. R. Calabrese, C. M. Connelly, J. S. Schneekloth, in *Methods Enzymol.* **2019**, *623*, 131–149, DOI: <https://doi.org/10.1016/bs.mie.2019.05.030>.
- [50] J. H. Davis, B. F. Dunican, S. A. Strobel, *Biochemistry* **2011**, *50*, 7236–7242.
- [51] A. Ciulli, G. Williams, A. G. Smith, T. L. Blundell, C. Abell, *J. Med. Chem.* **2006**, *49*, 4992–5000.
- [52] A.-S. Bretonnet, A. Jochum, O. Walker, I. Krimm, P. Goekjian, O. Marcillat, J.-M. Lancelin, *J. Med. Chem.* **2007**, *50*, 1865–1875.
- [53] A. Umuhire Juru, A. E. Hargrove, *J. Biol. Chem.* **2021**, *296*, 100191.
- [54] N. F. Rizvi, J. P. Santa Maria, A. Nahvi, J. Klappenbach, D. J. Klein, P. J. Curran, M. P. Richards, C. Chamberlin, P. Saradjian, J. Burchard, et al., *SLAS Discovery* **2020**, *25*, 384–396.
- [55] O. Binas, V. Jesus, T. Landgraf, A. E. Völklein, J. Martins, D. Hymon, J. Kaur Bains, H. Berg, T. Biedenbänder, B. Fürtig, et al., *ChemBioChem* **2020**, *22*, 423–433.
- [56] A. Garcia-Lopez, F. Tessaro, H. R. A. Jonker, A. Wacker, C. Richter, A. Comte, N. Berntenis, R. Schmucki, K. Hatje, O. Petermann, et al, *Nat. Commun.* **2018**, *9*, 2032.
- [57] K. D. Warner, C. E. Hajdin, K. M. Weeks, *Nat. Rev. Drug Discovery* **2018**, *17*, 547–558.
- [58] B. S. Morgan, J. E. Forte, R. N. Culver, Y. Zhang, A. E. Hargrove, *Angew. Chem. Int. Ed.* **2017**, *56*, 13498–13502; *Angew. Chem.* **2017**, *129*, 13683–13687.
- [59] R. Rangan, A. M. Watkins, J. Chacon, R. Kretsch, W. Kladwang, I. N. Zheludev, J. Townley, G. Thain, R. Das, *Nucleic Acids Res.* **2021**, *49*, 3092–3108.
- [60] M. D. Disney, A. J. Angelbello, *Acc. Chem. Res.* **2016**, *49*, 2698–2704.
- [61] P. R. Bhatt, A. Scaiola, G. Loughran, M. Leibundgut, A. Kratzel, R. Meurs, R. Dreos, K. M. O'Connor, A. McMillan, J. W. Bode, V. Thiel, D. Gatfield, J. F. Atkins, N. Ban, *Science* **2021**, *372*, 1306–1313.

Manuscript received: March 18, 2021

Revised manuscript received: June 3, 2021

Accepted manuscript online: June 23, 2021

Version of record online: August 6, 2021