Chemical-guided SHAPE sequencing (cgSHAPE-seq) informs the binding site of RNA-degrading chimeras targeting SARS-CoV-2 5' untranslated region

Zhichao Tang[1,3], Shalakha Hegde[1,3], Siyuan Hao[2], Manikandan Selvaraju[1], Jianming Qiu[2], Jingxin Wang*[1]

[1]Department of Medicinal Chemistry, University of Kansas, Lawrence, KS, USA.

[2]Department of Microbiology, Molecular Genetics and Immunology, University of Kansas Medical Center, Kansas City, KS, USA.

[3]These authors contributed equally.

*Email: wang.jingxin@ku.edu

One of the hallmarks of RNA viruses is highly structured untranslated regions (UTRs) in their genomes. These conserved RNA structures are often essential for viral replication, transcription, or translation. In this report, we discovered and optimized a new type of coumarin derivatives, such as **C30** and **C34**, which bind to a four-way RNA helix called SL5 in the 5' UTR of the SARS-CoV-2 RNA genome. To locate the binding site, we developed a novel sequencing-based method namely cgSHAPE-seq, in which the acylating chemical probe was directed to crosslink with the 2'-OH groups of ribose at the ligand binding site. This crosslinked RNA could then create read-through mutations during reverse transcription (i.e., primer extension) at single-nucleotide resolution to uncover the acylation locations. cgSHAPE-seq unambiguously determined that a bulged G in SL5 was the primary binding site of **C30** in the SARS-CoV-2 5' UTR, which was validated through mutagenesis and in vitro binding experiments. **C30** was further used as a warhead in RNA-degrading chimeras to reduce viral RNA expression levels. We demonstrated that replacing the acylating moiety in the cgSHAPE probe with ribonuclease L recruiter (RLR) moieties yielded RNA degraders active in the in vitro RNase L degradation assay and SARS-CoV-2 5’ UTR expressing cells. We further explored another RLR conjugation site on the E ring of **C30/C34** and discovered improved RNA degradation activities in vitro and in cells. The optimized RNA-degrading chimera **C64** inhibited live virus replication in lung epithelial carcinoma cells.

RNA viruses usually have highly structured 5’ and 3’ UTRs in their RNA genome, which can potentially serve as therapeutic targets[1]. In this report, we used SARS-CoV-2 as a specific test-case example and explored to use RNA-degrading chimeras to inhibit virus replication. SARS-CoV-2 is an enveloped ssRNA(+) virus. The whole genome of SARS-CoV-2 (~30,000 nucleotides) is encoded in a single RNA molecule[2]. The viral RNA in transmitted virions is 5' capped and 3' polyadenylated, and, therefore, it is first recognized and treated as mRNA[3]. In this step, the 5’ UTR is used to hijack the host ribosome to translate viral proteins[4]. Furthermore,

the 5’ UTR plays an essential role in RNA transcription for each coronavirus structural protein, which is accomplished through a "discontinuous” transcription mechanism. Specifically, the replication transcription complex binds to the 5’ UTR leader transcriptional regulatory sequences (TRS-L), and then “hops” onto the body TRS (TRS-B) sequence located at the 5’-end of each structural gene[5,6]. That said, all SARS-CoV-2 transcripts share the same 5’ UTR leader sequence. In addition, the SARS-CoV-2 5’ UTR was reported essential for viral RNA packaging[7].

Given the importance of the UTRs in SARS-CoV-2, we and others elucidated the RNA structures in SARS-CoV-2 UTRs[8–12]. The 5’ UTR RNA structures in cell-free buffers, virus-infected cells, and our reporter cell model are highly consistent [8–13], suggesting superior stability and suitability serving as drug targets. The 5’ UTR of SARS-CoV-2 contains five stem-loops, namely SL1–5. The start codon resides in SL5, a unique four-way helix[8–12] (Fig. 1a). SL5 exists in all betacoronovirus species, including MERS and SARS-CoV, and the shapes of this RNA structure are similar[9]. We aligned the SARS-CoV-2 RefSeq and different lineages and demonstrated that the SL5 is highly conserved among all strains[1]. Although a predominant mutation was found in SL5B loop region from recent viral lineages (C241T), this mutation is unlikely to change the overall structure of SL5[1]. In SARS-CoV-2, several structures in the 5’ UTR, including SL4, SL5A, and SL6, were found binding to amilorides[14] (Fig. 1a). Amilorides demonstrated antiviral activity in SARS-CoV-2 infected cells.

Here, we report a pipeline in antiviral discovery and optimization of RNA-degrading chimeras targeting SL5, highlighting a novel sequencing-based method namely chemical-guided (cg) selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) sequencing (seq), or cgSHAPE-seq, to rapidly locate the RNA ligand binding site (Fig. 1a). First, we screened a small coumarin derivative library in SL5 RNA binding assay and optimized the hit through structure-activity-relationship studies. To elucidate the RNA ligand binding site, we synthesized and applied a new type of chemical probe that can selectively acylate the 2’-OH on the ribose at the location of binding (Fig. 1b)[26,27]. The 2’-OH acylation locations were “recorded” onto RNA molecules by reverse transcriptase as single-point mutations at the modification sites during primer extension. The mutation sites were then captured and deconvoluted by next-generation sequencing[28,29]. Mutational profiling analysis in cgSHAPE-seq unambiguously identified a bulged G in SL5 as the primary binding site in the SARS-CoV-2 5’ UTR. In the literature, other sequencing-based methods were reported using affinity probes bearing nitrogen mustard or diazirine moiety (e.g., ChemCLIP-seq[15–19] and PEARL-seq[20]). However, a major limitation of these methods is a strong labeling bias toward guanosines[21]. Similar to the SHAPE, cgSHAPE reacts with the 2’-OH group of the ribose in all nucleotides, A, U, G, or C. This can potentially increase the scope and accuracy of proximity-induced chemical reactions on RNAs for mapping purposes. While we were preparing for the manuscript, the Kool lab also reported a proximity-induced acylation approach and determined RNA-binding sites for several FDA-approved drugs[22].

We further developed RNA-degrading chimeras by replacing the 2’-OH acylating moiety with RNase L recruiter (RLR) moieties on the cgSHAPE probe, as well as by conjugating the RLR moieties on other putative solvent-accessible sites on the RNA ligand. RNA-degrading chimera

utilizing endogenous ribonuclease (RNase) L was first reported by the Silverman group in 1993[23]. Recently, the modality of RNA-degrading chimera was further developed by the Disney group and was demonstrated to be active using small-molecule RNA ligands[24–31] (Fig. 1c). The Disney group also coined the name "ribonuclease targeting chimera or RIBOTAC for this type of RNA degraders. RIBOTACs were shown efficacious to degrade microRNA in cells[25,26] and mouse models[27]. Importantly, a small-molecule RIBOTAC was recently used to degrade the SARS-CoV-2 RNA genome by targeting an RNA structure named attenuator hairpin near the programmed frameshift (PFS) regulatory element[24]. The viral RNA transcript level was shown to be reduced in a model cell system by ~50% at 8 μM RIBOTAC[24]. Antisense-based RNA-degrading chimeras targeting spike-protein or envelope-protein encoding RNAs also demonstrated efficacy in SARS-CoV-2 infected cells[29]. Our optimized RIBOTAC robustly degraded SARS-CoV-2 RNA in cellular models at 1 μM and inhibited virus replication at 20 μM in lung epithelial cells. No significant toxicity was observed. Interestingly, we discovered the natural RNase L binding moiety is similar to or less active than the synthetic RLR in the RIBOTAC modality.

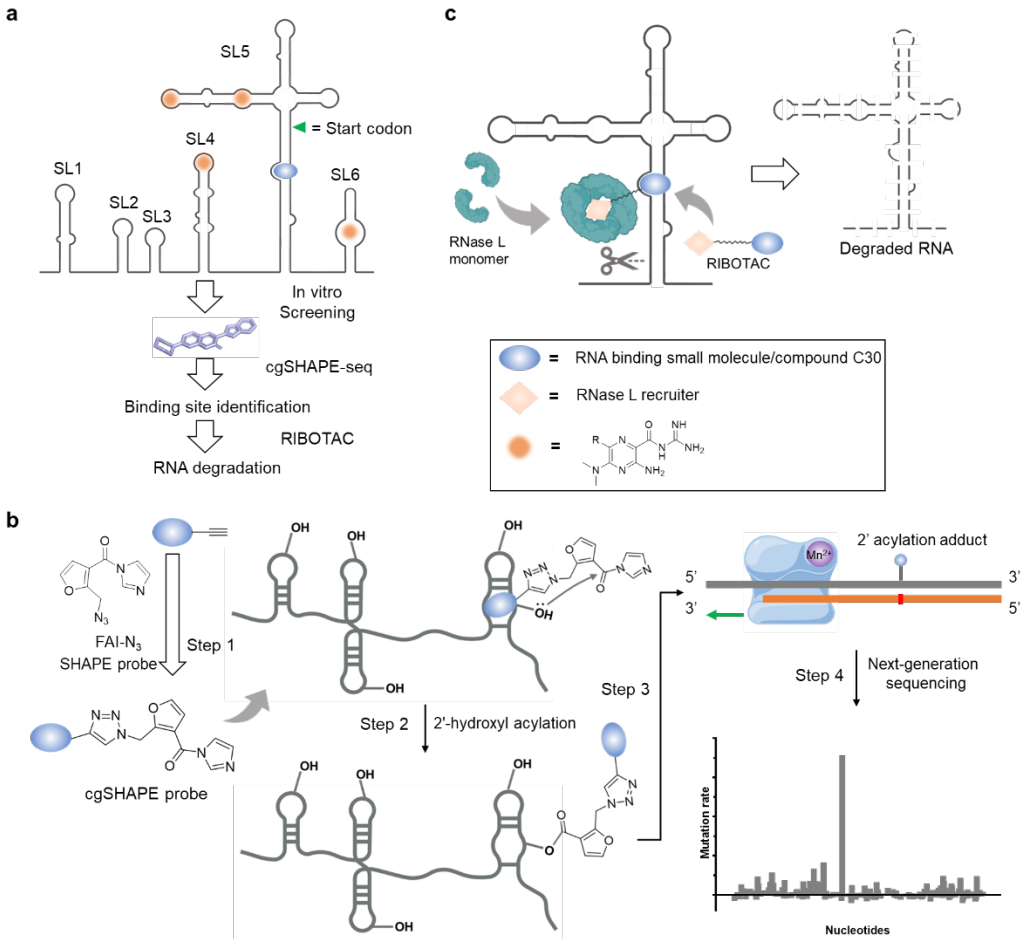


Fig. 1 | Development of cgSHAPE-seq and anti-viral RIBOTAC. a, RNA secondary structures in SARS-CoV-2 5' UTR and the pipeline in identification of the ligand binding site and the development of RNA-degrading chimeras. **b**, Principle of cgSHAPE-seq for identifying small molecule binding sites in four steps. Step 1: Synthesis of FAI conjugated chemical probe. Step 2: Chemical-guided acylation at the 2'-OH of ribose at the binding site. Step 3: Reverse transcription in the presence of Mn^{2+} creates single-point mutations at the acylation site. Step 4: Mutational profiling

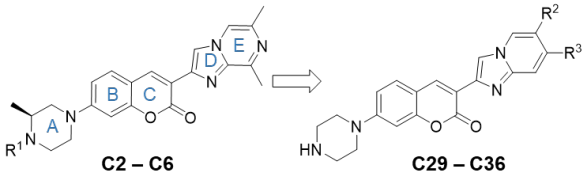
and quantification identify the putative binding sites (figure adapted from Figdraw). **c**, RNA-degrading chimeras (RIBOTACs) recruit RNase L at the target RNA to degrade viral RNAs.

Results

Chemical optimization of the coumarin derivatives for SL5 RNA binding

We previously synthesized a collection of coumarin derivatives that are known to bind to RNAs[32]. Each of these coumarin derivatives is fluorescent (excitation/emission ~400/480 nm), enabling us to use fluorescence polarization (FP) to determine the in vitro binding affinity with the RNA receptors. We in vitro transcribed SL5 RNA (144–303, RefSeq NC_045512) and screened the compound library using the FP assay (Extended Data Fig. 1). **C2** binds to SL5 at a dissociation constant (K_d) of 1.45 μM (Table 1). Elimination of the ethyl substituent on the A ring (**C4**) did not improve the binding affinity to SL5 RNA (Table 1). In contrast, a bulky substituent on the A ring (**C6**) impeded the interaction with SL5 (Table 1). This indicated that a non-sterically hindered linker might be suitable for conjugation on ring A without affecting the ligand binding affinity to SL5 RNA. We found **C29** the best ligand in our compound collection, which binds to SL5 RNA at a K_d of 0.47 μM. Compared to **C4**, **C29** has a different set of substituents on ring E, and therefore, we further investigated ring E while keeping ring A as unsubstituted piperazine (Table 1). In vitro FP assay showed that a fluorinated analog, **C30**, further improves the binding affinity (K_d = 0.22 μM). Changing the F group into Cl (**C32**) or CF_3 (**C36**) groups or alternating the fluorinated site (**C31**) all reduced the binding affinities (Table 1). In contrast, replacing the fluorine group in **C30** with methoxy group maintained the binding affinity (**C34**, K_d = 0.14 μM). To our knowledge, **C30 and C34** are the strongest small-molecule ligand to the SARS-CoV-2 SL5 RNA.

Table 1. Structure-activity-relationship study of coumarin derivatives for in vitro binding affinities with SL5 RNA.



Name	R^1	R^2	R^3	SL5 K_d (μM)
C2	CH_2CH_3	NA	NA	1.45 ± 0.38
C4	H	NA	NA	1.24 ± 0.29
C6	CH_2CH_2NHBoc	NA	NA	5.67 ± 4.78
C29	NA	H	H	0.47 ± 0.09
C30	NA	H	F	0.22 ± 0.09
C31	NA	F	H	0.69 ± 0.07
C32	NA	H	Cl	0.42 ± 0.09
C36	NA	H	CF_3	1.41 ± 0.25
C34	NA	H	OMe	0.14 ± 0.03

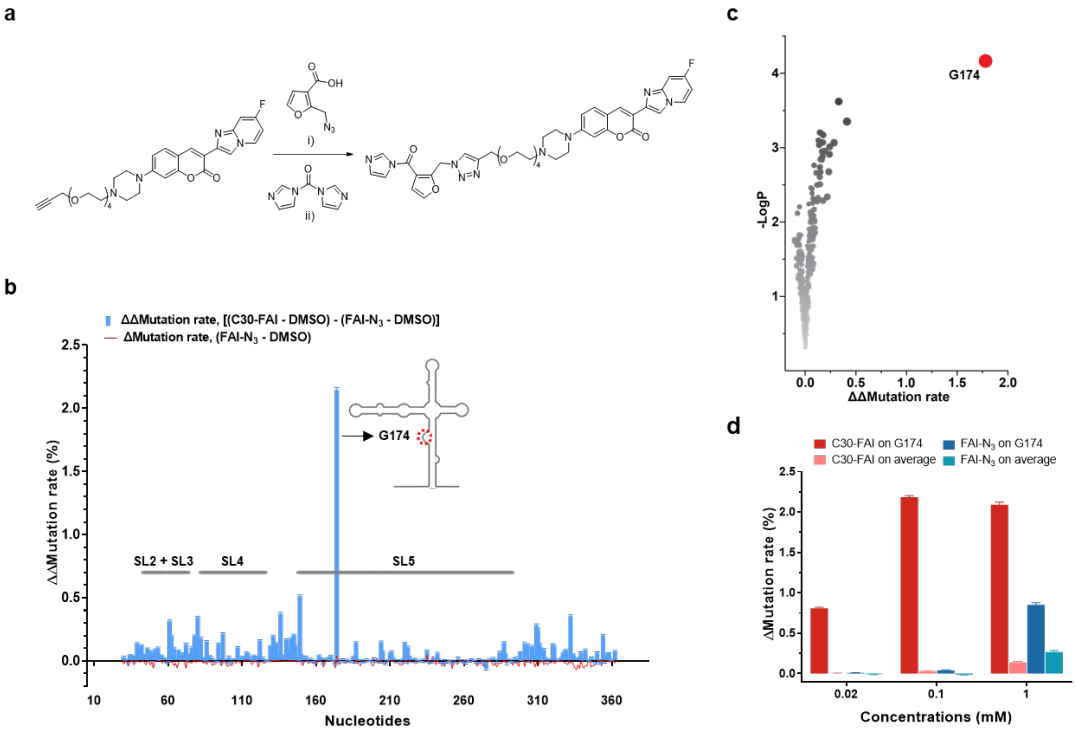
127 **cgSHAPE-seq uncovered the bulged G in SL5 as the C30 binding site**

128 We were inspired by a commonly used method for RNA structure elucidation called SHAPE and
129 sought to develop a new chemical-guided sequencing-based method to identify the binding site
130 of **C30/C34** in SL5 RNA. We named the new chemical probing method chemical-guided SHAPE
131 sequencing, or cgSHAPE-seq. Conventional SHAPE uses electrophilic reagents that can form
132 ester adducts on the 2'-OH of the ribose. The unpaired nucleotides have higher accessibility for
133 acylation reactions, which is the basis of structure-based differential acylation activity.
134 Therefore, identification of the acylation site would provide information on RNA base-pairing in
135 the conventional SHAPE. Importantly, the electrophile-ribose adduct can create a mutation
136 during reverse transcription (i.e., primer extension). As a result, conventional SHAPE coupled
137 with quantitative mutational profiling has become a gold standard method to explore RNA
138 topology in vitro and in vivo[33–37]. The key advantage of SHAPE is that the electrophile can
139 usually react with all four nucleotides (A, U, G, or C). We wondered if the ribose acylation could
140 be repurposed for identification of small-molecule binding sites by covalently linking electrophile
141 moieties to RNA-binding chemical ligands.

142 First, we selected furoyl acylimidazole (FAI) as the acylating electrophile for synthesizing the
143 chemical probe. Compared to other electrophile moieties such as anhydride and acyl cyanide,
144 FAI is more resistant to hydrolysis with a reported half-life of ~73 min in aqueous solutions[38–40].
145 The hydrolysis-resistant probe design renders the synthesis and storage less demanding for
146 anhydrous experimental facilities. An azide group on the furoyl moiety was used to provide a
147 click-chemistry handle to conjugate with alkyne-modified **C30** (Fig. 2a). We confirmed that this
148 alkyne group does not eliminate the binding of **C30** to SL5 RNA, even though the binding affinity
149 reduced by 7-fold (Supplementary Information, Fig. S1). After the [3+2] cycloaddition reaction,
150 the carboxylic acid was then converted into an acyl imidazole moiety under a mild condition,
151 finalizing the synthesis of **C30-FAI** probe (Fig. 2a). The acyl imidazole probe was prepared
152 freshly and used directly without further purification (see Method).

153 Next, we applied **C30-FAI** in SARS-CoV-2 5' UTR RNA to elucidate the binding site of **C30**. It is
154 known that conventional SHAPE experiments usually require high concentrations (10–100 mM)
155 of acyl imidazole (e.g., FAI) for ribose acylation. To avoid obtaining structure-based SHAPE
156 activity caused by FAI moiety alone, we chose to use a much lower dose of the chemical probe
157 for cgSHAPE-seq. We reasoned that at low concentrations of the probe (0.02–1 mM), the
158 differential acylation activity would be predominantly caused by ligand binding (i.e., proximity-
159 promoted acylation). Briefly, the total RNA was extracted from SARS-CoV-2 5' UTR expressing
160 cells (see Method) and refolded in buffer. **C30-FAI**, FAI-N_3, or DMSO was individually reacted
161 with the folded RNA for 15 min at 37 °C. After the reaction, we used ProtoScript II reverse
162 transcriptase (New England Biolabs) in the presence of $MnCl_2$ (3 mM) for primer extension
163 using a protocol modified from the literature report (see Method)[33]. In this step, we screened
164 several commercially available reverse transcriptases and found ProtoScript II one of the best
165 enzymes that can tolerate Mn^{2+} in the reaction (Supplementary Information, Fig. S2). The cDNA
166 was then amplified by PCR in the SARS-CoV-2 5' UTR region and the resulting amplicon was
167 subsequently sequenced. We applied an existing software package ShapeMapper2 developed
168 by the Weeks group for mutational profiling analysis[33,39]. We calculated the background

169 Δmutation rate (FAI-N_3 – DMSO) for each nucleotide and pleasingly observed a very low
170 background signal at 0.02 and 0.1 mM of FAI-N_3, indicating that the structure-based differential
171 acylation activity is negligible in cgSHAPE-seq at these concentrations (Extended Data Fig. 2).
172 In contrast, FAI-N_3 at 1 mM significantly increased the Δmutation rate, implying the contribution
173 of structure-based SHAPE activities started to emerge at high concentrations, confirming the
174 suitable probe concentration for cgSHAPE is in the range 0.02−0.1 mM (Fig. 2d). We then
175 calculated the RNA ligand-induced ΔΔmutation rate [(C30-FAI – DMSO) – (FAI-N_3 – DMSO)] for
176 each nucleotide and identified G174 as the only significantly mutated nucleotides in 0.02–0.1
177 mM probe-treated samples (Fig. 2b and 2c, Extended Data Fig. 3). In 1 mM probe-treated
178 samples, the cgSHAPE signal of G174 is comparable with 0.1 mM probe-treated ones, but a
179 high signal-to-noise ratio was observed due to the structure induced SHAPE background (Fig.
180 2d). Mapping the nucleotide with previously identified secondary structures uncovered that
181 G174 is a single-nucleotide bulge in the SL5 stem region[8–12]. It is worth noting that a previous
182 study showed that G174 was highly reactive in a regular SHAPE experiment using another
183 acylation agent, NAI, at 100 mM[11]. This concentration is at least 1,000 times higher than the
184 optimal cgSHAPE probe concentration determined in our study (0.02−0.1 mM), and therefore,
185 the cgSHAPE signal is unlikely attributed to the structure-induced SHAPE background at this
186 position (Fig. 2d).

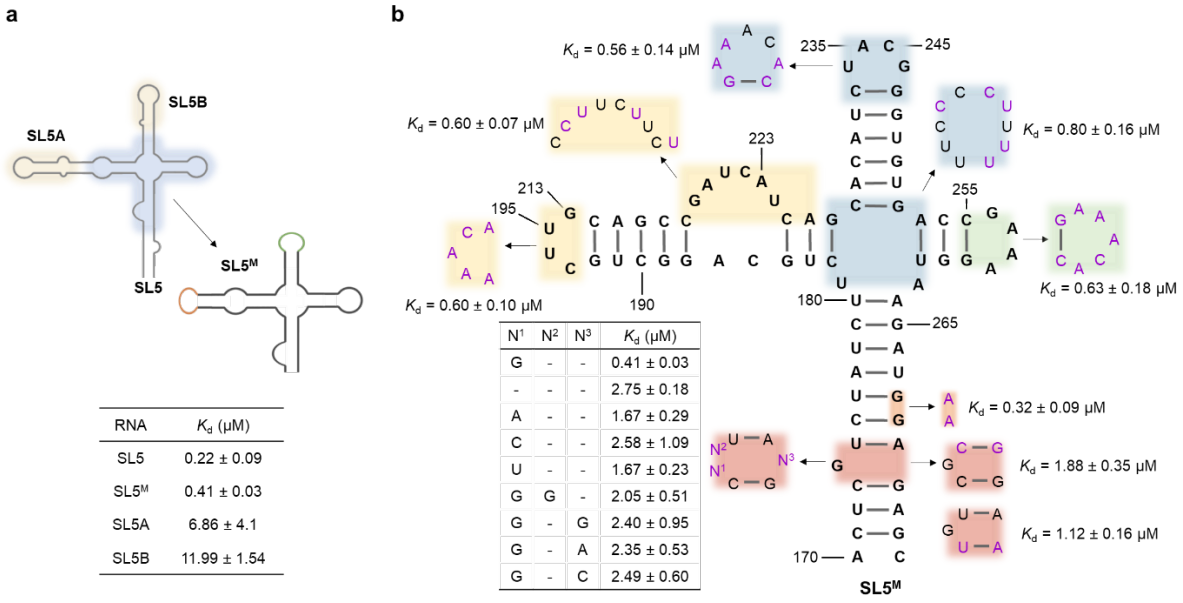


187

188 **Fig. 2 | Identification of the binding site by cgSHAPE-seq. a**, cgSHAPE-seq probe (**C30-FAI**) synthetic route.
189 Reaction conditions: i) tris(hydroxypropyltriazolylmethyl)amine, $CuSO_4$, sodium ascorbate, DMSO, room temperature;
190 ii) anhydrous DMSO, room temperature. **b**, cgSHAPE-seq mutational profiling analysis of the SL5 sequence in total
191 RNA extract treated with **C30-FAI** (0.1 mM). Δmutation rate (FAI-N_3 – DMSO) indicates the background structure-
192 based differential acylation. ΔΔmutation rate [(C30-FAI – DMSO) – (FAI-N_3 – DMSO)] indicates the proximity-based
193 differential acylation. The cgSHAPE-seq experiments were performed with three replicates (N = 3). **c**, Scatter plot of -
194 LogP vs ΔΔmutation rate. **d**, Comparison of the Δmutation rates of G174 and on average in RNAs treated with

different concentrations of **C30-FAI** or FAI-N_3. Three data points (C75, G216, G285) were removed as outliers as they had an abnormally high mutation rate (*Z*-score > 4.0) in DMSO-treated samples.

We then validated the **C30** binding site in SL5 by testing individual substructures of the SL5 RNA. The loop region of SL5A, SL5B, and a minimized four-helix junction, named $SL5^M$ (containing shorter stems) were synthesized chemically or enzymatically (Fig. 3a). The in vitro binding results demonstrated that only $SL5^M$ retained similar binding affinity to **C30** (Fig. 3a, Extended Data Fig. 4). To further validate the putative binding site G174 in $SL5^M$, we designed and synthesized $SL5^M$ RNAs with different mutations that disrupt the bulged G or other RNA structures (Fig. 3b). As expected, deletion of G174 or base-pairing G174 with an additional C both resulted in a 7-fold decrease in binding affinity to **C30**. Replacement of G174 with A, C, or U also significantly reduced **C30** binding. We also expanded the bulged G by inserting different nucleotides between C173 and U175 or between A270 and G271, all these mutated RNAs demonstrated 5–6-fold reduced binding affinity to **C30**. These results suggested the importance of a single bulged G in accommodating C30's binding. Changing the closing U-A base pair into C-G (3'-end of G174) or C-G base pair into U-A (5'-end of G174) also resulted in a 4-fold decreased binding. Mutations on other parts of the RNA have less impact (i.e., within 2-fold) in changing the binding affinity to **C30** (Fig. 3b). Altogether, these observations validated that the bulged G region is the primary binding site in SL5 RNA for **C30**. We concluded that cgSHAPE-seq is a validated method for identifying the binding site of RNA-binding small molecules.



RNA	K_d (μM)
SL5	0.22 ± 0.09
$SL5^M$	0.41 ± 0.03
SL5A	6.86 ± 4.1
SL5B	11.99 ± 1.54

N^1	N^2	N^3	K_d (μM)
G	-	-	0.41 ± 0.03
-	-	-	2.75 ± 0.18
A	-	-	1.67 ± 0.29
C	-	-	2.58 ± 1.09
U	-	-	1.67 ± 0.23
G	G	-	2.05 ± 0.51
G	-	G	2.40 ± 0.95
G	-	A	2.35 ± 0.53
G	-	C	2.49 ± 0.60

Fig. 3 | Validation of the binding site. a, Structural fragments of SL5 and their binding affinities to **C30**. **b**, $SL5^M$ mutants and their binding affinities to **C30**.

Comparison of cgSHAPE-seq and other sequencing-based RNA ligand localization methods

We compared cgSHAPE-seq with two existing sequencing-based methods for determining RNA ligand binding sites, PEARL-seq[20] and Chem-CLIP[15–19]. PEARL-seq used photocrosslinking

221 moieties such as diazirine to covalently link the bound RNA nucleotides. On the other hand,
222 Chem-CLIP uses a nitrogen mustard moiety to alkylate nucleobases by nucleophilic reactions.
223 To compare these two approaches with cgSHAPE-seq, we synthesized three new C30 probes:
224 both C30-D and C30-BD contain a diazirine group for photocrosslinking, and C30-NM contains
225 a nitrogen mustard group for alkylation. We also confirmed that these probes have similar
226 binding affinity to SL5 RNA compared to the cgSHAPE probe (Supplementary Information, Fig.
227 S1 and S3). These three chemical probes were used to treat the total RNA extracted from
228 SARS-CoV-2 5' UTR expressing cells at 0.1 mM under appropriate crosslinking conditions (see
229 Supplementary Information). We used the same reverse transcription and analysis pipeline as
230 those used in our cgSHAPE-seq experiments. We observed that C30-D in PEARL-seq can also
231 identify G174 as the binding site albeit with a higher background than cgSHAPE-seq, while the
232 other two probes didn't significantly differentiate G174 from other modified nucleotides
233 (Extended Data Fig. 5). Taken together, we concluded that the FAI-based acylation probe is a
234 more sensitive chemical probe to identify RNA binding sites of small molecules.

235 **Comparison of two RNase L recruiting moieties in RIBOTACs**

236 We then conjugated **C30/C34** with RLRs to synthesize SL5-targeting RNA degraders
237 (RIBOTACs). For RLR conjugation on the A ring, the acylating moiety in the above cgSHAPE
238 probe (**C30-FAI**) was replaced with RLR moieties. We noticed that in most reported co-NMR
239 structures of RNA bulges and small molecules that are similar to **C30/C34**, both termini of the
240 small molecules are solvent accessible[41–43], and, therefore, ring E was also explored for
241 conjugation to RLRs, where the methoxy group of **C34** was substituted for RLR attachment (Fig.
242 4a).

243 For RLR moieties, the natural RNase L ligand 2'-5'-lined oligoadenylate (2-5A) and its synthetic
244 mimic **D1** were both previously reported to be used in RNA-degrading chimeras (Extended Data
245 Fig. 6)[24,25]. Combining the two conjugation sites and two RLR structures, we obtained four
246 RIBOTAC candidates, **C47**, **C48**, **C64**, and **C65**, for SL5 RNA degradation (Fig. 4a). We
247 validated that the polyethylene glycol (PEG) linker on 2-5A does not affect the activity in a
248 reported RNase L degradation assay with a 5' 6-fluorescein-tagged model RNA containing
249 multiple RNase L cleavage sites[44,45] (Extended Data Fig. 6a). It was demonstrated that the
250 binding affinity between RNase L and **D1** ($K_d \approx 18$ μM) is 80,000-fold weaker than that observed
251 for 2-5A[45]. Consistent with this reported in vitro binding data, the synthetic **D1** alone is > 10,000
252 times weaker than 2-5A in the in vitro RNase L degradation assay (Extended Data Fig. 6b and
253 6c).

254 Next, we tested the four RIBOTACs in the in vitro RNase L degradation assay with purified SL5
255 RNA and observed their activities in order: **C64** > **C47** ≈ **C48** > **C65** (Fig. 4b, Extended Data Fig.
256 7a). To our surprise, the RIBOTAC **C64** with **D1** as the RLR moiety is much stronger than **C65**
257 with 2-5A at 50 μM (Fig. 4b). This result is contrary to what we would have predicted based on
258 the activities of the RLR moieties per se. We validated these in vitro findings in SARS-CoV-2 5'
259 UTR expressing 293T cells. In this cell model, the SARS-CoV-2 5' UTR sequence was fused to
260 a CMV promoter-controlled Gaussia luciferase expression cassette (Fig. 4c; for sequences, see
261 Method). Consistent with the RNase L degradation assay result, the maximum potency of **C64**

(i.e., RNA reduction level) was significantly better than **C65** (Fig. 4c). The activities of **C47** and **C48** in this cell model are similar to those of **C64** and **C65** (Extended Data Fig. 7b).

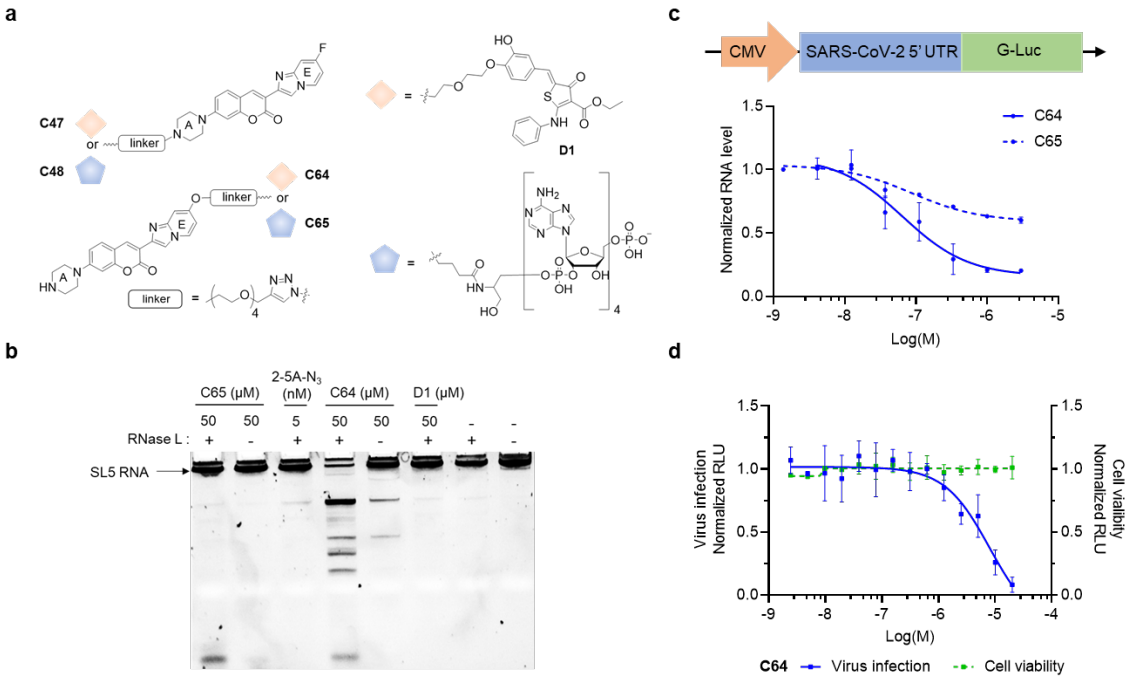


Fig. 4 | RNA degrading activity and anti-viral activity of C30-based RIBOTACs. a, Synthesis of **C30**-based RIBOTACs using conjugation sites on rings A or E of **C30**. **b**, Comparison of two RLR moieties in the RIBOTAC modality using the in vitro RNase L degradation assay with purified SL5 RNA. **c**, Cellular activity of RIBOTACs in SARS-CoV-2 5' UTR expressing cells. **d**, Inhibitory effect of RIBOTAC **C64** in SARS-CoV-2 infected A549 cells. The cytotoxicity of the compound was also evaluated. The dose-response curves are representative of three independent measurements (N = 3).

Efficacy of RIBOTAC in live virus infection assay

Finally, we tested the activity of **C64** in SARS-CoV-2 infected cells. The SARS-CoV-2 virus was engineered to include a Nano Luciferase (NLuc) reporter by fusing NLuc onto ORF7 of the SARS-CoV-2 genome[46]. In this way, the NLuc signal is proportional to the viral protein copy number in cells. We applied a human lung epithelial carcinoma cell line A549 that expresses high level of ACE2 as the host cell[46]. The cells were infected with the SARS-CoV-2-NLuc virus at a multiplicity of infection (MOI) of 2.0 at 1 h before the treatment with RIBOTACs **C64** for 3 d. To our satisfaction, **C64** showed > 95% inhibition at 20 μM (Fig. 4d). At the same concentration, no major toxicity is observed in A549 cells (Fig. 4d).

Discussion

In cgSHAPE-seq, it is critical to employ a concentration at which the predominant source of SHAPE activity is the proximity-enhanced activity. Specifically, a standard SHAPE probing should be performed parallelly with an identical probe concentration as that of the cgSHAPE probe to determine the full set of SHAPE reactivity in the folded RNA (Fig. 2d). At the optimal cgSHAPE probe concentration (i.e., 0.02−0.1 mM in this study), the standard SHAPE activity

286 should be minimal (typically < 0.1% as demonstrated in Fig. 2d). In addition, the putative binding
287 sites determined by cgSHAPE-seq require extensive validation. In the current study, we used
288 mutagenesis and in vitro binding assay to validate the binding site (Fig. 3). We envision that
289 reverse genetics with a mutated binding site coupled with relevant functional assays can also be
290 used for validation purposes if the RNA binder has well-defined biological consequences.

291 Multiple crucial factors merit attention when designing cgSHAPE probes: (1) the cgSHAPE
292 probe should be considerably stable in water solution to allow sufficient target engagement and
293 to avoid hydrolysis by moisture in the air during the experiment preparation steps. In this current
294 study, we chose FAI moiety with a reported half-life of ~73 min. We attempted to use C30-NAI
295 conjugate but observed that the probe is considerably less stable than **C30-FAI** (data not
296 shown). (2) The conjugation site of acylation moiety should be solvent accessible and not
297 impede ligand binding. In this current study, the **C30-FAI** precursor for cgSHAPE and **C30-D**
298 probe for PEARL-seq both showed significant binding to the RNA target, albeit 3−7 fold less
299 binding affinity (Supplementary Information, Fig. S1 and S3). (3) The functional group should be
300 compatible with acylating moiety. For example, the FAI-based probes used in this report would
301 not be compatible with nucleophilic RNA ligands[14] due to self-reaction. cgSHAPE-seq potentially
302 has a limitation with bias towards certain types of small molecule-RNA interactions. As shown in
303 conventional SHAPE, FAI moiety has a higher reactivity towards unpaired RNA nucleotides.
304 Although most of the reported RNA ligands target the unpaired region[47,48], cgSHAPE-seq may
305 be less reactive for ligands that bind to the double-stranded RNA grooves.

306 In our cgSHAPE-seq result, apart from G174, we also observed a cluster of nucleotides from
307 A131 to G149 showing slightly higher mutation rates than others (Fig. 2b and 2c). This can be
308 potentially caused by the nonspecific binding of **C30** with flexible sequence[32]. RNA targeting
309 strategies are known to have off-target effects due to shallow binding sites on RNAs and
310 relatively weak binding affinity for small molecules. **C64** at 3 μM can cause >100 genes up- or
311 down-regulated (|Log2FoldChange| > 2) in the transcriptome (Extended Data Fig. 8,
312 Supplementary Table S1). The activity of the RIBOTAC might be improved if a more potent and
313 selective RNase L recruiter is used[49]. Specifically, we showed that the natural RNase L
314 recruiter/activator 2-5A, which is negatively charged, is sometimes not compatible with the
315 positively charged RNA binder **C30** (Fig. 4b). For this reason, new synthetic RNase L recruiter
316 should probably be considered to be neutral or positively charged as most of the reported RNA
317 ligands are also positively charged[50–52].

318 In summary, we developed a new generalizable chemical probing method called cgSHAPE-seq
319 for quickly identifying small molecule-RNA binding sites by sequencing. cgSHAPE probes react
320 with the 2'-OH groups on the ribose close to the binding sites with a mitigated dependency of
321 the nucleobase identity observed in other reported methods. We used cgSHAPE-seq to identify
322 a bulged G on SL5 as the primary binding site on the SARS-CoV-2 5' UTR targeted by the
323 newly discovered coumarin derivatives, such as **C30** and **C34**. Finally, we developed a novel
324 **C30/C34**-based RNA degrader (RIBOTAC) capable of degrading viral RNA transcripts in cells
325 and inhibiting virus replication in SARS-CoV-2 infected cells, offering crucial insights into RNA
326 degrading chimeras' design.

References

1. Hegde, S., Tang, Z., Zhao, J. & Wang, J. Inhibition of SARS-CoV-2 by Targeting Conserved Viral RNA Structures and Sequences. *Front. Chem.* **9**, 802766 (2021).

2. Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 (2020).

3. Malone, B., Urakova, N., Snijder, E. J. & Campbell, E. A. Structures and functions of coronavirus replication-transcription complexes and their relevance for SARS-CoV-2 drug design. *Nat. Rev. Mol. Cell Biol.* **23**, 21–39 (2022).

4. de Breyne, S. *et al.* Translational control of coronaviruses. *Nucleic Acids Research* vol. 48 12502–12522 (2020).

5. Zúñiga, S., Sola, I., Alonso, S. & Enjuanes, L. Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. *J. Virol.* **78**, 980–94 (2004).

6. Sola, I., Almazán, F., Zúñiga, S. & Enjuanes, L. Continuous and Discontinuous RNA Synthesis in Coronaviruses. *http://dx.doi.org/10.1146/annurev-virology-100114-055218* **2**, 265–288 (2015).

7. Escors, D., Izeta, A., Capiscol, C. & Enjuanes, L. Transmissible Gastroenteritis Coronavirus Packaging Signal Is Located at the 5′ End of the Virus Genome. *J. Virol.* **77**, 7890–7902 (2003).

8. Zhao, J., Qiu, J., Aryal, S., Hackett, J. L. & Wang, J. The RNA Architecture of the SARS-CoV-2 3′-Untranslated Region. *Viruses* **12**, 1473 (2020).

9. Sun, L. *et al.* In vivo structural characterization of the SARS-CoV-2 RNA genome identifies host proteins vulnerable to repurposed drugs. *Cell* **184**, 1865-1883.e20 (2021).

10. Huston, N. C. *et al.* Comprehensive in vivo secondary structure of the SARS-CoV-2 genome reveals novel regulatory motifs and mechanisms. *Mol. Cell* **81**, 584-598.e5 (2021).

11. Manfredonia, I. *et al.* Genome-wide mapping of SARS-CoV-2 RNA structures identifies therapeutically-relevant elements. *Nucleic Acids Res.* (2020) doi:10.1093/nar/gkaa1053.

12. Sanders, W. *et al.* Comparative analysis of coronavirus genomic RNA structure reveals conservation in SARS-like coronaviruses. *bioRxiv Prepr. Serv. Biol.* (2020) doi:10.1101/2020.06.15.153197.

13. Lan, T. C. T. *et al.* Structure of the full SARS-CoV-2 RNA genome in infected cells. *bioRxiv* (2020) doi:10.1101/2020.06.29.178343.

14. Zafferani, M. *et al.* Amilorides inhibit SARS-CoV-2 replication in vitro by targeting RNA structures. *Sci. Adv.* **7**, eabl6096 (2021).

15. Velagapudi, S. P. *et al.* Defining RNA-Small Molecule Affinity Landscapes Enables Design of a Small Molecule Inhibitor of an Oncogenic Noncoding RNA. *ACS Cent. Sci.* **3**, 205–216 (2017).

16. Suresh, B. M. *et al.* A general fragment-based approach to identify and optimize bioactive

365 ligands targeting RNA. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 33197–33203 (2020).

366 17. Velagapudi, S. P. *et al.* Design of a small molecule against an oncogenic noncoding
367 RNA. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 5898–903 (2016).

368 18. Velagapudi, S. P., Li, Y. & Disney, M. D. A cross-linking approach to map small molecule-
369 RNA binding sites in cells. *Bioorganic Med. Chem. Lett.* **29**, 1532–1536 (2019).

370 19. Guan, L. & Disney, M. D. Covalent Small-Molecule-RNA Complex Formation Enables
371 Cellular Profiling of Small-Molecule-RNA Interactions. *Angew. Chemie Int. Ed.* **52**,
372 10010–10013 (2013).

373 20. Mukherjee, H. *et al.* PEARL-seq: A Photoaffinity Platform for the Analysis of Small
374 Molecule-RNA Interactions. *ACS Chem. Biol.* **15**, 2374–2381 (2020).

375 21. Balaratnam, S. *et al.* A chemical probe based on the PreQ1 metabolite enables
376 transcriptome-wide mapping of binding sites. *Nat. Commun.* **12**, 5856 (2021).

377 22. Fang, L. *et al.* Pervasive transcriptome interactions of protein-targeted drugs. *Nat. Chem.*
378 *2023 1510* **15**, 1374–1383 (2023).

379 23. Torrence, P. F. *et al.* Targeting RNA for degradation with a (2'-5')oligoadenylate-
380 antisense chimera. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 1300–4 (1993).

381 24. Haniff, H. S. *et al.* Targeting the SARS-CoV-2 RNA Genome with Small Molecule Binders
382 and Ribonuclease Targeting Chimera (RIBOTAC) Degraders. *ACS Cent. Sci.* **6**, 1713–
383 1721 (2020).

384 25. Costales, M. G., Matsumoto, Y., Velagapudi, S. P. & Disney, M. D. Small Molecule
385 Targeted Recruitment of a Nuclease to RNA. *J. Am. Chem. Soc.* **140**, 6741–6744 (2018).

386 26. Costales, M. G., Suresh, B., Vishnu, K. & Disney, M. D. Targeted Degradation of a
387 Hypoxia-Associated Non-coding RNA Enhances the Selectivity of a Small Molecule
388 Interacting with RNA. *Cell Chem. Biol.* **26**, 1180-1186.e5 (2019).

389 27. Costales, M. G. *et al.* Small-molecule targeted recruitment of a nuclease to cleave an
390 oncogenic RNA in a mouse model of metastatic cancer. *Proc. Natl. Acad. Sci.* **117**, 2406–
391 2411 (2020).

392 28. Liu, X. *et al.* Targeted Degradation of the Oncogenic MicroRNA 17-92 Cluster by
393 Structure-Targeting Ligands. *J. Am. Chem. Soc.* **142**, 6970–6982 (2020).

394 29. Su, X. *et al.* Efficient Inhibition of SARS-CoV-2 Using Chimeric Antisense
395 Oligonucleotides through RNase L Activation*. *Angew. Chem. Int. Ed. Engl.* **60**, 21662–
396 21667 (2021).

397 30. Zhang, P. *et al.* Reprogramming of Protein-Targeted Small-Molecule Medicines to RNA
398 by Ribonuclease Recruitment. *J. Am. Chem. Soc.* **143**, 13044–13055 (2021).

399 31. Bush, J. A. *et al.* Ribonuclease recruitment using a small molecule reduced c9ALS/FTD
400 r(G4C2) repeat expansion in vitro and in vivo ALS models. *Sci. Transl. Med.* **13**,
401 eabd5991 (2021).

402 32. Tang, Z. *et al.* Recognition of single-stranded nucleic acids by small-molecule splicing

403 modulators. *Nucleic Acids Res.* **49**, 7870-7883. PMC8373063 (2021).

404 33. Smola, M. J., Rice, G. M., Busan, S., Siegfried, N. A. & Weeks, K. M. Selective 2 ′ -
405 hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP)
406 for direct , versatile and accurate RNA structure analysis. *Nat. Protoc.* **10**, 1643–1669
407 (2015).

408 34. Weeks, K. M. & Mauger, D. M. Exploring RNA Structural Codes with SHAPE Chemistry.
409 *Acc. Chem. Res.* **44**, 1280–1291 (2011).

410 35. Siegfried, N. A., Busan, S., Rice, G. M., Nelson, J. A. E. & Weeks, K. M. RNA motif
411 discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods* **11**, 959
412 (2014).

413 36. Spitale, R. C. *et al.* Structural imprints in vivo decode RNA regulatory mechanisms.
414 *Nature* **519**, 486–490 (2015).

415 37. Flynn, R. A. *et al.* Transcriptome-wide interrogation of RNA secondary structure in living
416 cells with icSHAPE. *Nat. Protoc.* **11**, 273–290 (2016).

417 38. Chan, D., Beasley, S., Zhen, Y. & Spitale, R. C. Facile synthesis and evaluation of a dual-
418 functioning furoyl probe for in-cell SHAPE. *Bioorg. Med. Chem. Lett.* **28**, 601–605 (2018).

419 39. Busan, S. & Weeks, K. M. Accurate detection of chemical modifications in RNA by
420 mutational profiling (MaP) with ShapeMapper 2. *RNA* **24**, 143–148 (2018).

421 40. Chan, D., Feng, C., Zhen, Y., Flynn, R. A. & Spitale, R. C. Comparative Analysis Reveals
422 Furoyl in Vivo Selective Hydroxyl Acylation Analyzed by Primer Extension Reagents
423 Form Stable Ribosyl Ester Adducts. *Biochemistry* **56**, 1811–1814 (2017).

424 41. Campagne, S. *et al.* Structural basis of a small molecule targeting RNA for a specific
425 splicing correction. *Nat. Chem. Biol. 2019 1512* **15**, 1191–1198 (2019).

426 42. Nagano, K., Kamimura, T. & Kawai, G. Interaction between a fluoroquinolone derivative
427 and RNAs with a single bulge. *J. Biochem.* **171**, 239–244 (2022).

428 43. Ichijo, R., Kamimura, T. & Kawai, G. Interaction between a fluoroquinolone derivative
429 KG022 and RNAs: Effect of base pairs 3′ adjacent to the bulged residues. *Front. Mol.*
430 *Biosci.* **10**, 168 (2023).

431 44. Thakur, C. S., Xu, Z., Wang, Z., Novince, Z. & Silverman, R. H. A convenient and
432 sensitive fluorescence resonance energy transfer assay for RNase L and 2',5'
433 oligoadenylates. *Methods Mol. Med.* **116**, 103–13 (2005).

434 45. Thakur, C. S. *et al.* Small-molecule activators of RNase L with broad-spectrum antiviral
435 activity. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 9585–90 (2007).

436 46. Adhikary, P. *et al.* Discovery of Small Anti-ACE2 Peptides to Inhibit SARS-CoV-2
437 Infectivity. *Adv. Ther.* **4**, 2100087 (2021).

438 47. Hargrove, A. E. Small molecule–RNA targeting: starting with the fundamentals. *Chem.*
439 *Commun.* **56**, 14744–14756 (2020).

440 48. Childs-Disney, J. L. *et al.* Targeting RNA structures with small molecules. *Nat. Rev. Drug*

441 *Discov.* **21**, 736–762 (2022).

442 49. Meyer, S. M. *et al.* DNA-Encoded Library Screening To Inform Design of a Ribonuclease
443 Targeting Chimera (RiboTAC). *J. Am. Chem. Soc.* **144**, 21096–21102 (2022).

444 50. Morgan, B. S., Forte, J. E., Culver, R. N., Zhang, Y. & Hargrove, A. E. Discovery of Key
445 Physicochemical, Structural, and Spatial Properties of RNA-Targeted Bioactive Ligands.
446 *Angew. Chemie - Int. Ed.* **56**, 13498–13502 (2017).

447 51. Velagapudi, S. P., Gallo, S. M. & Disney, M. D. Sequence-based design of bioactive
448 small molecules that target precursor microRNAs. *Nat. Chem. Biol.* **10**, 291–297 (2014).

449 52. Disney, M. D. *et al.* Inforna 2.0: A Platform for the Sequence-Based Design of Small
450 Molecules Targeting Structured RNAs. *ACS Chem. Biol.* **11**, 1720–1728 (2016).

451

452 **Methods**
453 **Synthesis of C30-FAI (cgSHAPE probe)**

454 Compound **C30-alkyne** (50 mg, 0.08 mmol) in DMSO (1 mL) was added 2-(azidomethyl)furan-
455 3-carboxylic acid (15 mg, 0.09 mmol), THPTA (9 mg, 0.02 mmol), sodium ascorbate (8 mg, 0.04
456 mmol) and $CuSO_4$ (3 mg, 0.02 mmol). The reaction vial was sealed, evacuated, and refilled with
457 N_2 three times and stirred at room temperature overnight. DMSO was removed under vacuum
458 and the residue was purified by silica gel column chromatography (0 – 10% CH_3OH in CH_2Cl_2)
459 to afford **C30-FCA** as a yellow solid (45 mg, 70%). MS-ESI (*m*/*z*) $[M+1]^+$ 746.28.

460 ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.71 – 8.68 (m, 2H), 8.52 (s, 1H), 8.09 (s, 1H), 7.71 – 7.68 (m,
461 2H), 7.38 (dd, *J* = 10.1, 2.6 Hz, 1H), 7.02 (dd, *J* = 9.0, 2.3 Hz, 1H), 6.97 (td, *J* = 7.6, 2.6 Hz, 1H),
462 6.88 (d, *J* = 2.4 Hz, 1H), 6.72 (d, *J* = 1.9 Hz, 1H), 5.92 (s, 2H), 4.52 (s, 2H), 3.58 – 3.51 (m,
463 14H), 3.38 (t, *J* = 5.1 Hz, 4H), 2.59 (t, *J* = 5.1 Hz, 4H), 2.56 (t, *J* = 5.8 Hz, 2H).

464 ^{13}C NMR (126 MHz, DMSO-d_6) δ 160.2 (d, *J* = 252 Hz), 159.5, 154.9, 153.3, 152.2, 144.4,
465 144.3, 143.4, 139.4, 138.6, 129.6, 129.2 (d, *J* = 11.8 Hz), 124.2, 114.5, 112.1, 111.7, 111.4,
466 110.2, 104.2 (d, *J* = 29.8 Hz), 99.6, 99.4, 69.8, 69.7, 69.1, 63.4, 57.0, 54.9, 52.6, 46.7, 44.8.

467 Compound **C30-FCA** (45 mg, 0.06 mmol) in anhydrous DMSO (0.6 mL) was added
468 carbonyldiimidazole (CDI, 10 mg, 0.06 mmol) and the reaction mixture was stirred at room
469 temperature for 1 h. The reaction mixture contains ~75% **C30-FAI** and ~25% unreacted **C30-**
470 **FCA** (see Supplementary Information) and was used directly in RNA modification. The stock
471 solution was used as **75 mM** and can be stored at –80 °C for long-term storage.

472 **cgSHAPE-seq using Total RNA Extract from Cells**

473 SARS-CoV-2 5' UTR expressing cells were harvested and pelleted. Total RNA was extracted
474 using TRIzol Reagent (Invitrogen) per the user's manual. An on-column DNA digestion was
475 performed to remove the residual genomic DNA in total RNA using DNase I (10 U/μL, Roche)
476 and RDD buffer (Qiagen). Purified total RNA was dissolved in water and stored at –80 °C before
477 use. For RNA modification, 5 μg total RNA was used for each reaction. **C30-FAI** and FAI-N_3

478 were prepared at 20 mM, 2 mM, and 0.4 mM in DMSO as 20× working solution. Briefly, total
479 RNA was added water and 5× folding/reaction buffer (500 mM HEPES pH 7.4, 500 mM KCl, 30
480 mM $MgCl_2$) to make a 47.5 μL solution. The solution was incubated at 37 °C for 30 min to refold.
481 2.5 μL **C30-FAI** (cgSHAPE probe), FAI-N_3 (background control) or DMSO was added to the
482 total RNA and mixed well by pipetting. The mixture was incubated at 37 °C for 15 min and then
483 quenched by adding RLT buffer (Qiagen). The RNA was then extracted using RNeasy kit
484 (Qiagen). 500 ng total RNA was used for reverse transcription and then PCR as described
485 below. All reactions were performed in triplicates.

486 For reverse transcription (10 μL reaction), probe or DMSO treated RNA and reverse
487 transcription primer (0.5 μM in final reaction buffer) were heated at 70 °C for 5 min and snap-
488 cooled on ice for 1min. 5× reaction buffer (375 mM Tris-HCl, 500 mM KCl, 15 mM $MnCl_2$, pH
489 7.4, 2 μL), DTT (100 mM, 1 μL), dNTP (10 mM, 0.5 μL), ProtoScript II (0.5 μL, New England
490 Biolabs, M0368L) and RNase inhibitor (0.2 μL, ApexBio, K1046) were added. The reaction was
491 incubated at 42 °C for 1 h and deactivated at 70 °C for 15 min. In each PCR reaction (50 μL),
492 cDNA (2.5 μL) was mixed with Phire Hot Start II DNA Polymerase (Thermo Fisher, 1 μL), dNTP
493 (10 mM, 1 μL), 5× Phire Green reaction buffer (10 μL), primers (0.5 μM in final reaction buffer)
494 and water (35.5 μL). After reaction, the amplicon was purified using a DNA Clean &
495 Concentrator kit (Zymo Research) following the user's manual. The purified DNA was submitted
496 for next-generation sequencing (Amplicon-EZ, Azenta Life Sciences).

497 An integrated software package developed by Busan and Weeks, ShapeMapper2 was used to
498 analyze the fastq files for mutational profiling and the rresult was used to generate Figures 2b,
499 2c, and Extended Data Fig. 3[39]. The reference sequence (SARS-CoV-2_5_UTR.fa) required for
500 ShapeMapper2 is listed below.

501 >SARS-CoV-2_5_UTR

502 aggtttataccttcccaggtaacAAACCAACCAACTTTCGATCTCTTGTAGATCTGTTCTCTAAACGAAC
503 TTTAAAATCTGTGTGGCTGTCACTCGGCTGCGTGCTTAGTGCACTCACGCAGTATAATTAAT
504 AACTAATTACTGTCGTTGACAGGACACGAGTAACTCGTCTATCTTCTGCAGGCTGCTTACG
505 GTTTCGTCCGTGTTGCAGCCGATCATCAGCACATCTAGGTTTCGTCCGGGTGTGACCGAAA
506 GGTAAGATGGAGAGCCTTGTCCCTGGTTTCAACGAGGGAGTCAAAGTTCTGTTTGCCCTGA
507 TCTGCATCGCTGTGGCCGAGGCCAAGCCCACCGAGAACAACGAagacttcaacatcgtggccg.
508 (lowercase = primer binding sequences).

509

510 ***In Vitro* RNase L Degradation Assay**

511 Purified recombinant GST-tagged RNase L was purchased from MyBioSource (MBS1041064).
512 The buffer of RNase L was exchanged into a buffer containing 50 mM Tris-HCl (pH 7.4) and 100
513 mM NaCl using Zeba Desalting Column (Thermo Fisher, 8766) using the manufacturer's
514 protocol. For RNase L degradation of SL5 RNA, T7 transcribed SL5 RNA was first purified by
515 polyacrylamide gel electrophoresis (PAGE) and recovered using small-RNA PAGE Recovery Kit
516 (Zymo Research, R1070). RNase L (1.3 μg in 5 μL) was incubated in the presence of **C47**, **C48**,

517 **C64**, **C65**, or DMSO control in the cleavage buffer (final reaction volume is 8 μL) at 4 °C for 12
518 h. The 1X cleavage buffer contains 25 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 100 mM KCl, 50
519 μM ATP, and 7 mM β-mercaptoethanol. The SL5 RNA (120 ng in 2 μl H_2O) was then added into
520 the reaction mixture and incubated for another 2 h at 22 °C. The reaction was stopped by
521 adding RNA Gel Loading Dye (Thermo Fisher, R0641) at 1:1 ratio. The samples (4 μl) were
522 then loaded on TBE-urea polyacrylamide gel (20%) for electrophoresis (180 V for 85 min). The
523 gel was stained with SYBR safe (1/50,000, ApexBio, A8743) in TBE buffer for 1 min and
524 visualized on gel imager (Thermo Fisher, iBright FL1500).

525 The RNA sequence of SL5 used in this assay is: 5'-
526 UCGUUGACAGGACACGAGUAACUCGUCU
527 AUCUUCUGCAGGCUGCUUACGGUUUCGUCCGUGUUGCAGCCGAUCAUCAGCACAUCUA
528 GGUUUCGUCCGGGUGUGACCGAAAGGUAAGAUGGAGAGCCUUGUCCCUGGUUUCAACG
529 A.

530 For RNase L degradation of a model 6-FAM-tagged RNA, RIBOTACs in the above protocol
531 were replaced with **D1** (0.37 μg in 0.75 μl DMSO) or **2-5A-N_3** (0.12 ng in 0.75 μl H_2O)[44,45]. After
532 electrophoresis, the gel was not stained and was directly visualized on the gel imager at the 6-
533 FAM fluorescence channel. The 6-FAM RNA (5'-6-FAM-
534 UUAUCAAAUUCUUAUUUGCCCCAUU
535 UUUUUGGUUUA-BHQ) was purchased from IDT.

536

537 **SARS-CoV-2 5' UTR Expressing Stable Cell Line**

538 293T cells (Thermo Fisher, R70007) were cultured in DMEM growth medium (Gibco, 11995040)
539 supplemented with 10% FBS (Cytiva, SH30910.03) and 1% Antibiotic-Antimycotic (Gibco,
540 15240062) at 37 °C in 5% CO_2 atmosphere. For producing the lentivirus, 293T cells were
541 seeded in a 6-well plate (Fisher, FBO12927) at 3 x 10^5 cells per well and transfected with 1 μg
542 of SARS-CoV-2 5' UTR expressing lentivirus vector (pLV-SARS-CoV-2-5'UTR-GLuc) along with
543 the packaging plasmids pMD2.G (0.4 μg) and psPAX2 (0.6 μg) using Lipofectamine 2000
544 (Invitrogen, 11668019). At 24 h post-transfection, the cell medium was replaced with fresh
545 growth medium. 48 h after the change of media, the supernatant containing the lentivirus
546 particles was siphoned and centrifuged at 500 g for 10 min at 4 °C to remove the cell debris.
547 The virus particles were further concentrated at 10X in volume using Lenti-X-concentrator
548 (Clontech, PT4421-2) according to the manufacturer's protocol. The lentivirus can be quantified
549 using literature method[53]. Usually, 10^7–10^8 plaque forming units (pfu)/mL lentivirus was obtained
550 after the concentrator treatment. For lentiviral transduction, 293T cells were inoculated with the
551 concentrated viral suspension (multiplicity of infection ~10) using polybrene (Sigma-Aldrich, TR-
552 1003-G) at a final concentration of 8 μg/mL. At 24 h post-transduction, the culture medium was
553 replaced with fresh growth media. After recovery for 24 h, the transduced cells were then
554 selected in blasticidin (10 μg/mL, Invivogen, ant-bl) for 2 weeks. For stable single clone
555 selection, the cells were diluted in the growth medium containing blasticidin (10 μg/mL) to a final
556 density of 1 cell per 100 μL. The diluted cell suspension was then dispensed to a 96-well plate

557 (100 μL per well). The plate was incubated at 37 °C for 4 weeks. A single cell colony from one
558 of the wells was then selected for experiments.

559 pLV-SARS-CoV-2-5UTR-Luc was constructed by inserting the SARS-CoV-2 5' UTR and
560 Gaussian luciferase into the pLV vector, under the control of the CMV promoter. The insert
561 sequence is as follows:

562 attaaaggtttataccttcccaggtaacaaaccaaccaactttcgatctcttgtagatctgttctctaaacgaactttaaaatctgtgtggctg
563 tcactcggctgcgtgcttagtgcactcacgcagtataattaataactaattactgtcgttgacaggacacgagtaactcgtctatcttctgc
564 aggctgcttacggtttcgtccgtgttgcagccgatcatcagcacatctaggtttcgtccgggtgtgaccgaaaggtaagatggagagcct
565 tgtccctggtttcaacgagggagtcaaagttctgtttgccctgatctgcatcgctgtggccgaggccaagcccaccgagaacaacgaa
566 gacttcaacatcgtggccgtggccagcaacttcgcgaccacggatctcgatgctgaccgcgggaagttgcccggcaagaagctgcc
567 gctggaggtgctcaaagagatggaagccaatgcccggaaagctggctgcaccaggggctgtctgatctgcctgtcccacatcaagt
568 gcacgcccaagatgaagaagttcatcccaggacgctgccacacctacgaaggcgacaaagagtccgcacagggcggcataggc
569 gaggcgatcgtcgacattcctgagattcctgggttcaaggacttggagcccatggagcagttcatcgcacaggtcgatctgtgtgtgga
570 ctgcacaactggctgcctcaaagggcttgccaacgtgcagtgttctgacctgctcaagaagtggctgccgcaacgctgtgcgacctttg
571 ccagcaagatccagggccaggtggacaagatcaagggggccggtggtgactaa (lowercase = SARS-CoV-2 5'
572 UTR; uppercase = Gaussia luciferase; underline = start codon).

573

574 **Quantitative Reverse Transcription PCR (RT-qPCR) Assay**

575 The SARS-CoV-2 5' UTR expressing cells were seeded at 3 x 10^5 cells per well in 12-well
576 plates in 1 mL growth medium at 37 °C for 3 h. The cells were then treated with the compounds
577 (**C47**, **C48**, **C64**, or **C65**) at various concentrations (1.3 nM–3 μM) for 48 h. After treatment, the
578 supernatant was aspirated from each well and the total RNA was then extracted from the cells
579 using RNeasy mini kit (Qiagen, 74104). The total RNAs were quantified by ultraviolet absorption
580 at 260 nm (Thermo Fisher, NanoDrop 1000). Usually 10–20 μg total RNA was obtained from
581 each well. cDNAs were synthesized from 500 ng of total RNA for each sample using M-MLV
582 reverse transcriptase (Promega, M1701) and $(dT)_{25}$ according to the manufacturer's protocol. 1
583 μl of cDNA mixture was used in a 15 μl RT-qPCR reaction (Apex-Bio, K1070). The human
584 GAPDH RNA level was used as the reference for normalization. The RT-qPCR primer
585 sequences used for the PCR are shown below:

586 SL5-SYBR-FW: 5'-CGTTGACAGGACACGAGTAA

587 SL5-SYBR-RV: 5'-TTGAAACCAGGGACAAGGCTC

588 GAPDH-FW: 5'-GACAAGGCTGGGGCTCATTT

589 GAPDH-RV: 5'-CAGGACGCATTGCTGATGAT

590

591 **SARS-CoV-2 Inhibition Assay**

592 Vero-E6 cells (ATCC® CRL-1586™) and A549 cells (ATCC® CCL-185) were cultured in
593 Dulbecco's modified Eagle's medium (DMEM, Cytiva Life Science, SH30022) with addition of
594 10% fetal bovine serum (FBS, Millipore Sigma, F0926) at 37°C under 5% CO_2 atmosphere.
595 A549 cells were transduced with a human ACE2-expressing lentivirus vector, and the
596 transduced were cultured in the DMEM plus 2 μg/μL puromycin[46].

597 *Virus and titration*

598 SARS-CoV-2-Nluc was created by engineering the nanoluciferase (Nluc) gene into the OFR7 of
599 the SARS-CoV-2 genome. The insertion site of Nluc at ORF7 was based on previous
600 mNeonGreen reporter SARS-CoV-2[54]. The virus was propagated in Vero-E6 cells once,
601 aliquoted in DMEM, and stored at –80˚C. A biosafety protocol to work on SARS-CoV-2 in the
602 BSL3 Lab was approved by the Institutional Biosafety Committee of the University of Kansas
603 Medical Center.

604 *Plaque assay*[55,56]

605 Vero-E6 cells were seeded in 24-well plates at a density of 0.5 ×10^6 cells per well. A virus stock
606 was serially diluted at 10-fold in Dulbecco's phosphate-buffered saline, pH7.4 (DPBS). 200 μL of
607 the diluent were added to each well and incubated for 1 h on a rocking rotator. After removing
608 the virus diluent, 0.5 mL of overlay media (1% methylcellulose in DMEM with 5% FBS) were
609 added to each well. The plates were incubated at 37 °C under 5% CO_2 for 4 days. The
610 methylcellulose overlays were aspirated, and the cells were fixed with 10% formaldehyde
611 solution for 30 min and stained with 1% crystal violet solution followed by extensive washing.
612 Plaques in each well were counted and multiplied by the dilution factor to determine the virus
613 titer at pfu/mL.

614 *Determination of half-maximal inhibitory concentration (IC_{50})*[46,57]

615 ACE2-A549 cells were seeded into 96 well plates. When the cells were confluent, SARS-CoV-2-
616 NLuc viruses were diluted with cold PBS and added into each well at a multiplicity of infection
617 (MOI) of 2 (2 pfu/cell). The plates were kept in the CO_2 incubator for 1 hour. Compound **C64**
618 was diluted at 2 x serials from 20 μM to 0.002 μM. The virus-PBS solution was aspirated. Each
619 well was washed with cold PBS three times, and was loaded with the diluted compounds. Each
620 concentration was loaded in triple wells in the plates, and the total volume of each well was 0.2
621 mL. The plates were kept in the incubator. After 3 days post-infection, the culture media were
622 aspirated from each well and the wells were washed with PBS three times. The nano-luciferase
623 activity assay (Promega, N1110) was carried out following the manufacturer's instructions.
624 Briefly, 100 ul of cell lysis buffer were added to each well for 10 minutes to completely lyse the
625 cells. Then 100 ul of nano-luciferase reaction reagent were added to each well and the
626 luminescent signal was determined at A490 absorbance on a plate reader (Bio-Tek, Synergy).
627 The IC_{50} was calculated using GraphPad Prism 8.0 software.

628

629 **Statistical Analysis**

All data shown as means ± s.d. with sample size (N) listed for each experiment. Statistical analysis was carried out with Prism GraphPad 8.0. Unpaired two-sample t-tests were used to analyze significant differences between the group means. The P values were calculated by Prism GraphPad 8.0 or R. For data generated from ShapeMapper2, the standard error (stderr) associated with the mutation rate at a given nucleotide in the S (probe treated) or U (DMSO treated) samples was calculated as: stderr = $\sqrt{mutation\ rate}/\sqrt{reads}$. The standard error of the Δmutation rate at a given nucleotide is: $\sqrt{{stderr_S}^2 + {stderr_U}^2}$.

Data Availability

The cgSHAPE-seq data for **C30-FAI**, FAI-N_3 or DMSO-treated total RNAs, and RNA-seq data for **C64**-treated cells were deposited in NCBI SRA with accession numbers PRJNA1029650 and PRJNA947619, respectively.

References

53. Barczak, W., Suchorska, W., Rubiś, B. & Kulcenty, K. Universal Real-Time PCR-Based Assay for Lentiviral Titration. *Mol. Biotechnol.* **57**, 195–200 (2015).

54. Xie, X. *et al.* A nanoluciferase SARS-CoV-2 for rapid neutralization testing and screening of anti-infective drugs for COVID-19. *Nat. Commun.* **11**, 5214 (2020).

55. Hao, S. *et al.* Long-term modeling of SARS-CoV-2 infection of in vitro cultured polarized human airway epithelium. *MBio* **11**, 1–17 (2020).

56. Zou, W. *et al.* The sars-cov-2 transcriptome and the dynamics of the s gene furin cleavage site in primary human airway epithelia. *MBio* **12**, (2021).

57. Hao, S. *et al.* Establishment of a Replicon Reporter of the Emerging Tick-Borne Bourbon Virus and Use It for Evaluation of Antivirals. *Front. Microbiol.* **11**, 2229 (2020).

Acknowledgments

Research reported in this article was supported by the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health (NIH) under award numbers R35GM147498 and P20GM113117, W. M. Keck Foundation, and the University of Kansas General Research Funds. We also thank Dr. Robert Silverman at Cleveland Clinic for providing the authentic 2-5A samples for comparison. We have obtained SARS-CoV-2-Nluc from Drs. Shi and Menachery through The University of Texas Medical Branch (UTMB)'s World Reference Center for Emerging Viruses and Arboviruses. We thank the University of Kansas Medical Center Genomics Core for doing RNA-seq experiment. (The KUMC Genomics Core was supported by Kansas Intellectual and Developmental Disabilities Research Center (NIH U54 HD

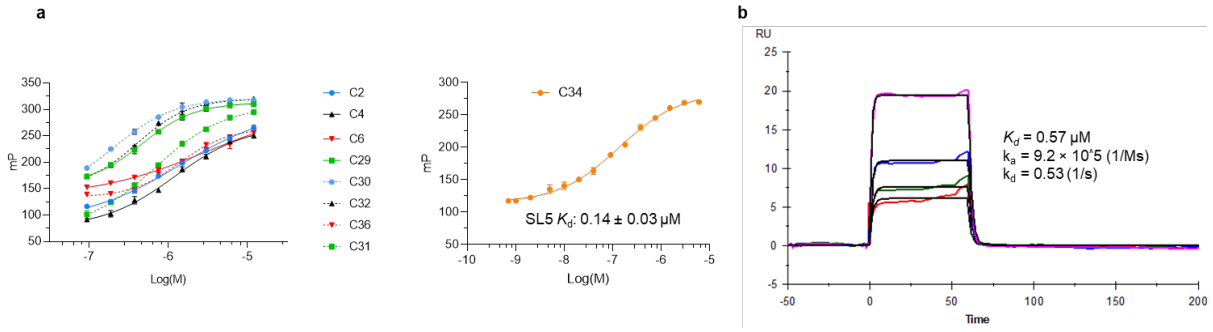
090216), the Molecular Regulation of Cell Development and Differentiation COBRE (P30 GM122731-03), the NIH S10 High-End Instrumentation Grant (NIH S10OD021743) and the Frontiers CTSA grant (UL1TR002366)).

Author contributions

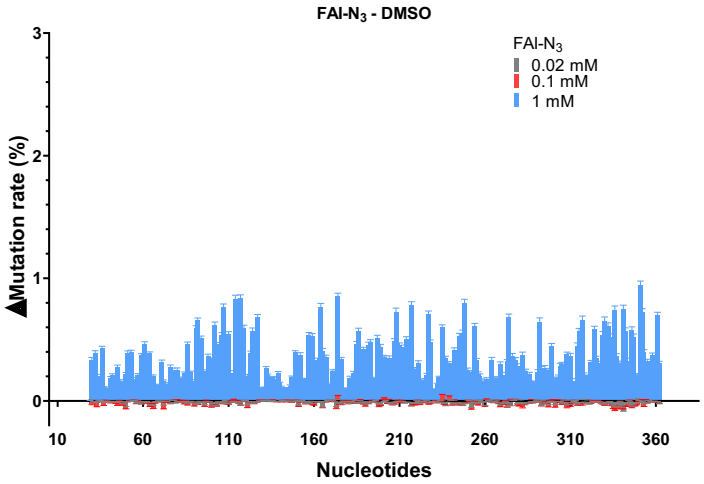
J.W. conceived the work. J.W., Z.T., S. Hegde and J.Q. wrote the paper. Z.T and M.S. performed chemical synthesis. Z.T. and J.W. conducted the RNA sequencing and analysis. S. Hegde collected RNA degradation data in vitro and in cells. S. Hao and J.Q. performed the live virus assay.

Competing interests

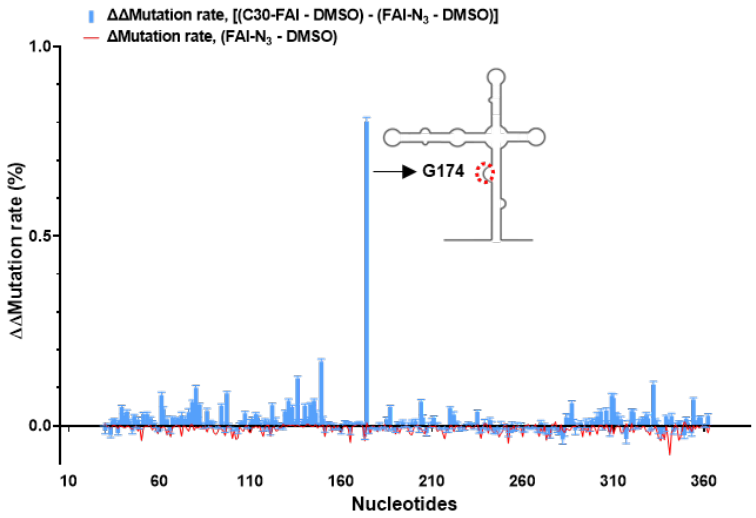
The authors declare no competing interests.



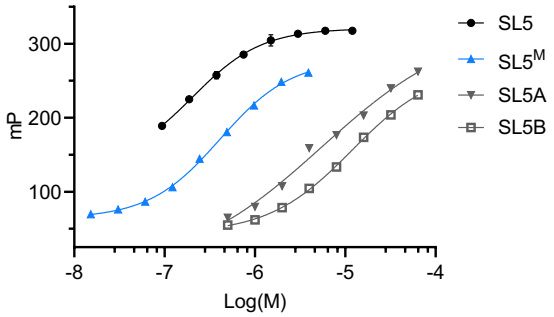
Extended Data Fig. 1 | Binding affinity of coumarin derivatives to SL5 RNA. a, Dose-response curves of coumarin derivatives in fluorescence polarization assay with in vitro transcribed SL5 RNA. All compounds were used at a concentration of 80 nM. Each data point represents the mean fluorescence polarization value of two independent replicates (N = 2). **b**, Surface plasmon resonance (SPR) binding analysis of **C30** with SL5 RNA. (Curve fitting is shown as black line. [SL5 RNA] = 0.075, 0.15, 0.3, and 0.6 μM).



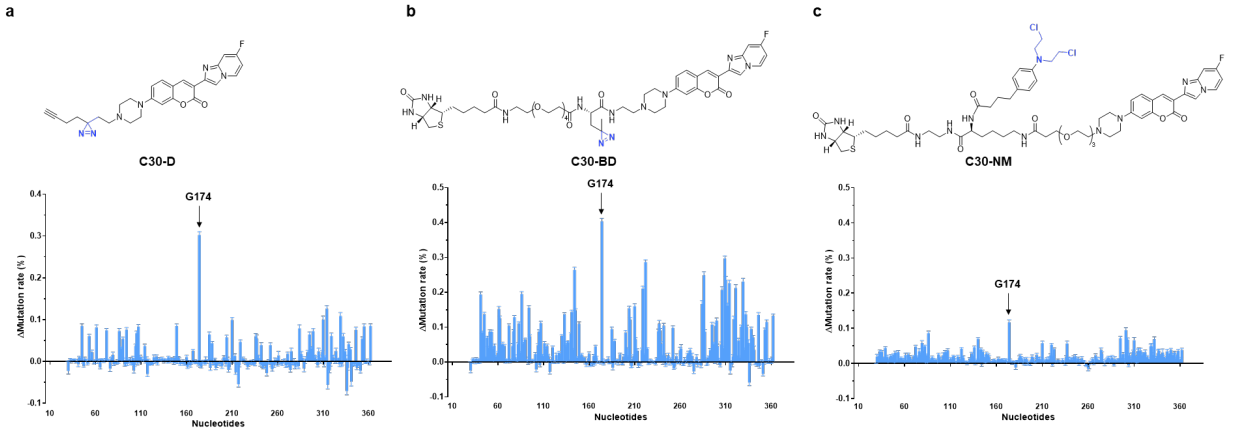
Extended Data Fig. 2 | Background Δmutation rate (FAI-N_3 – DMSO) of the SL5 sequence in total RNA extract treated with different concentrations of FAI-N_3. Three data points (C75, G216, G285) were removed as outliers as they had an abnormally high mutation rate (*Z*-score > 4.0) in DMSO-treated samples.



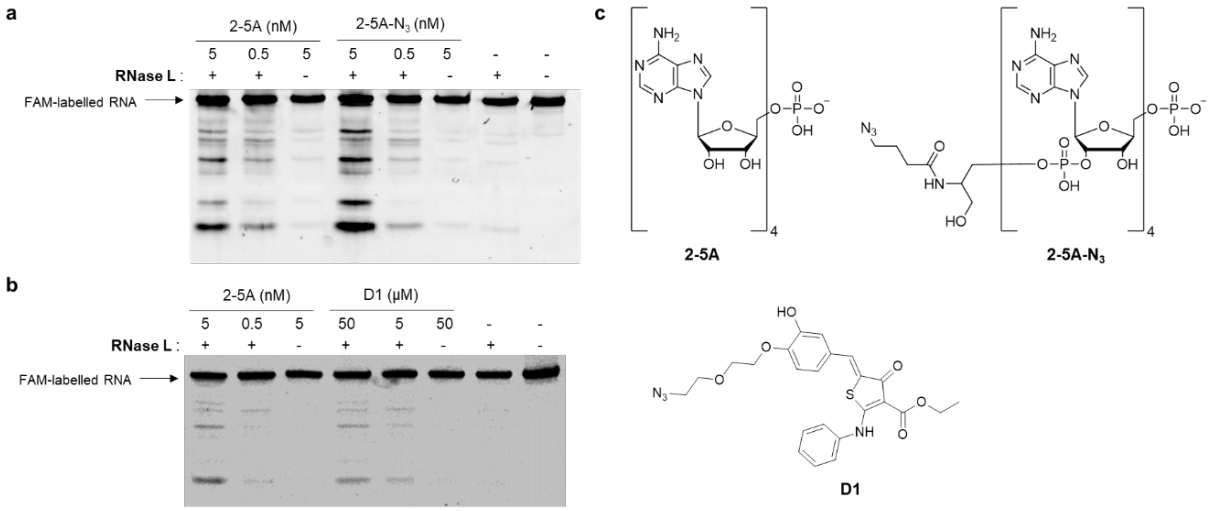
Extended Data Fig. 3 | cgSHAPE-seq mutational profiling analysis using C30-FAI (0.02mM). Δmutation rate (FAI-N_3 – DMSO) indicates the background structure-based differential acylation. ΔΔmutation rate [(C30-FAI – DMSO) – (FAI-N_3 – DMSO)] indicates the proximity-based differential acylation. The cgSHAPE-seq experiments were performed with three replicates (N=3). Three data points (C75, G216, G285) were removed as outliers as they had an abnormally high mutation rate (*Z*-score > 4.0) in DMSO-treated samples.



Extended Data Fig. 4 | Dose-response curves of **C30** (80 nM) in fluorescence polarization assay with SL5 RNA and its substructures. Each data point represents the mean fluorescence polarization value of two independent replicates (N = 2).



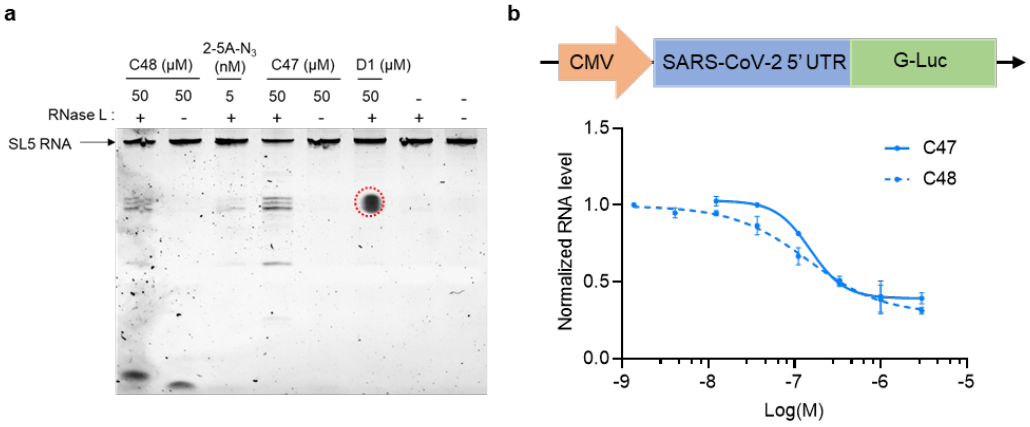
701 **Extended Data Fig. 5 | Mutational profiling analysis of crosslinking probes. a,** Structure of C30-D and
702 mutational profiling analysis. **b,** Structure of C30-BD and mutational profiling analysis. **c,** Structure of C30-NM and
703 mutational profiling analysis. All experiments were performed with three replicates (N = 3). Three data points (C75,
704 G216, G285) were removed as outliers as they had an abnormally high mutation rate (*Z*-score > 4.0) in DMSO-
705 treated samples.



706

707 **Extended Data Fig. 6 | In vitro RNase L degradation assay with a 5' 6-fluorescein-tagged model RNA**
708 **containing multiple RNase L cleavage sites**. **a**, Comparison of RNA degradation activity of RNase L in the
709 presence of **2-5A** and **2-5A-N$_3$** with no significant differential activity observed. **b**, Comparison of synthetic RNase L
710 recruiter (D1) and 2-5A. The activity of **D1** in RNase L activation is ~ 10,000 times weaker than **2-5A**. **c**, Chemical
711 structures of **2-5A**, **2-5A-N$_3$**, and **D1**.

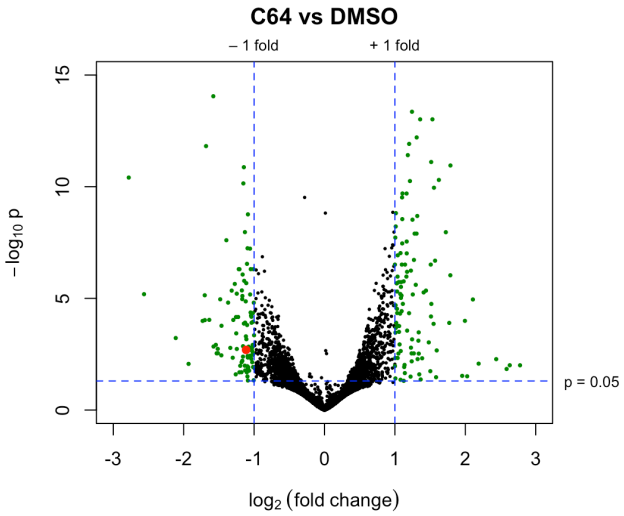
712



713

714 **Extended Data Fig. 7 | RNA degrading activity of C47 and C48**. **a**, Comparison of two RLR moieties in the
715 RIBOTAC modality using the in vitro RNase L degradation assay with purified SL5 RNA (red circle is a staining
716 artifact). **b**, Cellular activity of **C47** and **C48** in SARS-CoV-2 5' UTR expressing cells.

717



Extended Data Fig. 8 | Volcano plot of differential gene expression in SARS-CoV-2 5' UTR expressing cells treated with C64 (3 μM). DMSO-treated cells were used as a control. Red spot = SARS-CoV-2 5' UTR transcript. The RNA-seq analysis was performed with three biological replicates (N = 3).