



HAL
open science

Small molecule approaches to targeting RNA

Sandra Kovachka, Marc Panosetti, Benedetto Grimaldi, Stéphane Azoulay,
Audrey Di Giorgio, Maria Duca

► **To cite this version:**

Sandra Kovachka, Marc Panosetti, Benedetto Grimaldi, Stéphane Azoulay, Audrey Di Giorgio, et al..
Small molecule approaches to targeting RNA. *Nature Reviews Chemistry*, 2024, 8 (2), pp.120-135.
10.1038/s41570-023-00569-9 . hal-04801128

HAL Id: hal-04801128

<https://hal.science/hal-04801128v1>

Submitted on 25 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

Small molecule approaches to targeting RNA

Sandra Kovachka,^{1,‡} Marc Panosetti,^{1,2,‡} Benedetto Grimaldi,² Stéphane Azoulay,¹ Audrey Di Giorgio,¹ Maria Duca^{1,*}

¹ Université Côte d'Azur, CNRS, Institute of Chemistry of Nice, 28 avenue Valrose, Nice, France

² Molecular Medicine Research Line, Istituto Italiano di Tecnologia (IIT), Via Morego 30, 16163, Genoa, Italy

[‡] These authors contributed equally to the manuscript

* Corresponding author: maria.duca@univ-cotedazur.fr

Abstract

The development of innovative methodologies to identify RNA binders has attracted enormous attention in chemical biology and drug discovery. Although antibiotics targeting bacterial ribosomal RNA have been on the market for decades, the renewed interest in RNA targeting reflects the need to better understand complex intracellular processes involving RNA. In this context, small molecules are privileged tools to explore the biological functions of RNA, to validate RNA as therapeutic targets and, eventually, to become new drugs. Despite recent progresses, the rational design of specific RNA binders requires a better understanding of the interactions which occur with the RNA target to reach the desired biological response. In this review, we discuss the challenges to approaching this underexplored chemical space, together with recent strategies to bind, interact and affect biologically relevant RNAs.

[H1] Introduction

The paradigm of RNA targeting with both oligonucleotides and small molecules has existed in chemical and biological sciences for many years.¹ The recognition of the essential roles played by RNAs in living systems has resulted in the discovery of many new classes of RNAs involved not only in transcription and translation, but also in the regulation of gene expression. Oligonucleotides recognize RNAs by sequence complementarity, thanks to the formation of Watson-Crick-Franklin hydrogen bonds with single-stranded RNAs. In contrast, small molecules target specific sites in structured RNAs where the association of single-stranded and double-stranded regions leads to the formation of specific tridimensional structures.¹ This structure-recognition depends on the formation of binding pockets similar to those found in proteins. While oligonucleotides are important tools both for the fundamental study of RNA functions and for their potential therapeutic applications, their use remains limited owing to pharmacokinetic and pharmacodynamic limitations, such as stability, selectivity and biodelivery.² Furthermore, since RNAs, especially non-coding RNAs, are highly structured, oligonucleotides cannot access or interact efficiently with these structures. Some of the limitations of oligonucleotides have been addressed, such as stability by using modified nucleotides resistant to nucleases. Some others have still to be solved, such as biodelivery that asks for complex route of administration and thus impairing the applicability of such an approach.³ An approach complementary to oligonucleotides is the use of small molecules and holds great potential for chemical intervention on RNA functions.^{1,4-6}

The first RNA binders were discovered in the early forties when various natural products, such as the first aminoglycoside streptomycin (FIG1), were identified as antibiotics (FIG. 1).⁷ The antimicrobial activity was a result of binding to the prokaryotic ribosomal RNA which induced the impairment of protein synthesis in bacteria.⁸ Aminoglycosides consist of positively charged aminooligosaccharides that interact with the negatively charged RNA backbone.⁹

Subsequently, other natural compounds were identified as ribosome binders with antimicrobial properties. These include tetracyclines (such as doxycycline), macrolides (such as erythromycin), and oxazolidinones (such as linezolid).⁸ The therapeutic potential of RNA-targeting was also recently highlighted by the FDA approval of risdiplam for spinal muscular atrophy (SMA) treatment.¹⁰ Additionally, numerous compounds have been documented as RNA binders, to further enhance the understanding of small-molecule RNA targeting strategies.¹

In this review, we illustrate the structures of small molecule RNA binders together with the interactions they form with the RNA target and discuss the important molecular features for binding, selectivity and biological activity. The recent modalities adding new functions to specific RNA binders are also detailed. Finally, the knowledge gaps remaining for the field to progress for new

59 applications in chemical biology and therapeutics will be presented. Since the screening
60 methodologies and the biophysical approaches employed to identify and study RNA binders have been
61 reviewed and detailed elsewhere, they will be described only briefly.¹ This manuscript will focus on
62 the challenges that have been recently overcome and the ones that remain to design effective chemical
63 tools for RNA targeting. This review aims to promote medicinal chemists and chemical biologists to
64 enter the field of small molecule RNA targeting with the insight and understanding required for future
65 developments and innovations.

66 67 **[H1] Design of small molecules for RNA binding**

68 RNA's inherent structure means identifying RNA binders can be challenging (See BOX 1). A first
69 obstacle is the highly negatively charged backbone restricting compatible structures and favoring
70 interactions with positively charged species.¹¹ Also, RNA has only four bases compared to proteins' 20
71 amino acids, meaning selecting appropriate chemical scaffolds is key to designing ligands suitable for
72 RNA binding.⁴ Finally, unlike most protein targets, biologically relevant RNAs exist in a dynamic
73 ensemble of structures depending on base pairing.¹² Factors like salt concentration, intracellular
74 interactions or mutations influence the prevalence of certain sub-structures. Altering this dynamic
75 ensemble may impact the RNA's function but determining which RNA conformation is biologically
76 active *in vivo* remains challenging. The lack of knowledge about the isoform the compound will bind
77 to and how this will affect RNA function represents the third major drawback. Fortunately, structural
78 biologists now have a range of techniques in addition to X-ray, such as advanced NMR and cryo-EM,
79 that should provide the tools for important developments in the future.¹

80
81 Given these premises, the design of selective RNA binders as chemical probes or potential drugs
82 cannot be performed in the same way as protein targeting. Indeed, small molecule/RNA interactions
83 are dominated by hydrogen bonding and π -stacking interactions that constitute more than half of the
84 interactions underlying small molecule-RNA recognition in reported binders.¹³ Hydrophobic
85 interactions and weak hydrogen bonds (e.g. CH-O interactions) are also privileged interactions but
86 occur less frequently than hydrogen bonding and π -stacking. Conversely, the systematic analysis of
87 the interactions involved in small molecule-protein complexes are mainly represented by hydrophobic
88 contacts that constitute almost half of the studied interactions followed by hydrogen bonding and π -
89 stacking. Since the discovery of the first RNA binding small molecules, until the recent discovery of
90 specific and drug-like molecules, such as the FDA approved Risdiplam modifying RNA splicing in
91 spinal muscular atrophy and described later in this manuscript,¹⁰ show that there may well be
92 techniques to overcome the above-mentioned challenges. However, it is still difficult to perform a
93 rational design of ligands for a specific RNA target.

95 In the upcoming sections, we provide an exploration of the challenges and limitations that have been
96 tackled in the quest to efficiently target RNA with small molecules from the chemistry perspective.
97 We will not describe all the details about the biological consequences induced by RNA binding, but
98 instead detail how the compounds act at the molecular level and how they interact with the target. We
99 indeed believe that a better understanding of the formed interactions upon the study of RNA binders is
100 key to the development of improved ligands.

101
102 **[H2] Improving selectivity with cationic ligands.** The highly negatively charged RNA backbone
103 makes ligand design particularly challenging. Furthermore, the presence of paired and unpaired
104 regions, results in a 3D structure that differs from double-helical DNA and can be considered more
105 similar to proteins structure. The phosphate backbone promotes the use of binding compounds
106 containing chemical groups protonated at physiological pH, such as aminoglycosides, that can provide
107 electrostatic interactions with phosphates and hydrogen bonds with unpaired bases.⁷

108 To illustrate this concept, one can look at aminoglycosides that predominantly bind to RNA regions
109 with internal loops and bulges that distort the double helix.¹⁵ For example, plazomicin (compound **1**,
110 FIG2a right), a recent aminoglycoside approved for the treatment of infections caused by multidrug
111 resistant organisms, interacts with the ribosomal A-site through π -stacking and hydrogen bond
112 interactions with nucleobases as revealed by X-ray (FIG. 2a left).¹⁶ A-site region of the ribosome is the
113 one binding to the incoming tRNA bringing a new amino acid to be introduced in the growing peptidic
114 chain during protein synthesis. Additionally, plazomicin's hydroxy-aminobutyric acid (HABA) tail
115 (highlighted in blue in FIG2a) forms a H-bond with an uracil base. Together, this leads to a strong and
116 selective interaction even if electrostatic interactions are also formed, which reinforce binding but
117 hamper aminoglycoside specificity. To further increase selectivity, aminoglycosides have been
118 conjugated with moieties able to establish specific interactions with the target.¹⁷ One such strategy to
119 improve aminoglycoside selectivity is through conjugation with nucleobases,¹⁸ for example, a
120 modified neomycin designed to selectively target the HIV-1 dimerization initiation site (DIS). This is
121 a highly conserved RNA stem-loop structure located in the 5'-non-coding region of the viral genome.¹⁹
122 Its dimerization is a key step in HIV replication and its inhibition has been considered for potential
123 antiviral therapies. The structural study of neomycin in a complex with DIS highlighted that an
124 adenine base was involved in a key interaction with ring I of neomycin. Comparison with A-site, that
125 shows substantial similarity to DIS, illustrated that this interaction was absent in this case.¹⁹ Indeed,
126 ring I was replaced by a thymine leading to neomycin derivative **2** (FIG. 2b) able to bind DIS with an
127 improved specificity compared to neomycin. This example illustrates how the appropriate addition of
128 H-bonds donors and acceptors complementary to the structure of the RNA target can increase
129 selectivity for the desired target. In a similar way, neomycin and paromomycin were conjugated with
130 nucleobases as in compound **3** (FIG. 2c) to obtain ligands with an improved affinity for HIV-1 TAR
131 RNA over A-site.²⁰ This result was difficult to predict, as rational drug design remains challenging

132 even in the presence of structural information, most likely owing to the intrinsically dynamic RNA
133 systems. Furthermore, no biological activity was reported, suggesting issues such as poor permeability
134 or lack of specificity in cells.

135 Other types of heteroaromatic compounds have been applied to prepare aminoglycoside conjugates
136 targeting the biogenesis of oncogenic microRNAs (miRNAs or miR) that are essential regulators of
137 gene expression.²¹ The biogenesis of miRNAs goes through two highly structured precursors: a
138 primary miRNA (pri-miRNA) of about 1 kbases and a precursor of miRNA (pre-miRNA) that are
139 sequentially cleaved by ribonucleases Drosha in the nucleus and Dicer in the cytoplasm to lead to a
140 mature miRNA of about 22-24 nucleotides. This latter is able to bind to mRNAs and inhibit their
141 translation into proteins. Overexpression of some miRNAs appears to be responsible for the
142 development and progression of various cancers and these oncogenic miRNAs became not only
143 biomarkers for the different kinds of cancers but also potential anticancer targets. To design selective
144 ligands of stem-loop-structured precursors of oncogenic miRNAs and inhibit miRNAs biogenesis,
145 neomycin was conjugated to artificial nucleobases,²²⁻²⁵ heteroaromatic compounds able to form
146 specific hydrogen bonds with DNA and RNA base pairs, to form base triplets.²⁶ This led to
147 compounds, such as conjugate **4** (FIG. 2d), that selectively inhibit the biogenesis of a small set of
148 overexpressed oncogenic miRNAs in cancer cells.²⁵ Notably, incorporation of the phenyl-thiazole
149 heteroaromatic component directed neomycin to the stem-loop precursor of oncogenic miR-372, pre-
150 miR-372, inhibiting Dicer cleavage and miRNA maturation. This resulted in effective miR-372
151 inhibition *in vitro* and specific proliferation inhibition in cancer cells overexpressing miR-372 without
152 toxicity. Enzymatic footprinting (analysis of which residues of the RNA sequence are protected by
153 ribonuclease cleavage in the presence of a ligand) and molecular docking analyses suggested that
154 incorporation of a third RNA binding domain could enhance affinity and selectivity. Consequently,
155 amino acids, natural RNA binders, were conjugated to neomycin-nucleobase derivatives leading to
156 improved binding and inhibitory activity.²³ Notably, despite their selectivity for pre-miRNAs over
157 tRNA and DNA, these ligands also bind to a small set of other pre-miRNAs. The specificity of
158 biological activity for cells overexpressing miR-372, corroborated by other studies,²⁷ indicates that
159 selectivity strongly depends on intracellular target expression levels as well as on binding RNA in a
160 functional site to effectively inhibit its function or processing. Strong binding to non-functional off-
161 target sites may not result in any biological activity nor toxicity. Therefore, associating binding with
162 function is crucial in designing biologically active, selective RNA binders and evaluating potential off-
163 target implications.

164
165 In a complementary approach, two-dimensional combinatorial screening (2-DCS) involves a large set
166 of compounds immobilized on microarrays and screened against thousands of RNA motifs, including
167 bulges or internal loops.²⁸ 2-DCS was used to identify guanidilated aminoglycosides as specific
168 binders of RNA internal loops. Guanidilated neomycin (compound **5**, FIG. 2e) was found to bind a 5'-

169 AUACC/3'-UAAGG internal loop that is present in the precursor of oncogenic miR-10b and inhibit its
170 production in cells with micromolar activity. Compound **5** was thus a specific ligand of pre-miR10b,
171 with the guanidinium groups probably playing an important role in RNA binding.²⁹

172
173 Positively charged compounds interact with RNA in a selective manner providing electrostatic
174 interactions that are important for binding.³⁰ The association of these kind of compounds with moieties
175 able to form specific hydrogen bonds and hydrophobic interactions increases selectivity. Therefore, it
176 may be possible to use positively charged compounds to approach an RNA phosphate backbone, while
177 coupled to structures able to establish specific hydrogen bonds or hydrophobic interactions for
178 increased selectivity. The search for ligands whose main binding mode is not driven by the formation
179 of electrostatic interactions led to the discovery of more drug-like and specific binders as it will be
180 highlighted in the following sections.

181
182 **[H2] Targeting secondary and tertiary RNA structures.** The second challenge in the discovery of
183 RNA binders is the particular chemical space of the RNA targets that is different to the one of
184 proteins. This calls for a new paradigm to design selective ligands using strategies that can be broadly
185 divided into two categories: (i) the rational design of binders based on the identification of specific
186 interactions that could be formed with the RNA structure and (ii) the use of high-throughput
187 methodologies combining screening with bioinformatics and biophysics to identify new scaffolds. The
188 success of these strategies relies heavily on a thorough understanding of the secondary and tertiary
189 structures of the target, which play a critical role in determining potential binding pockets for ligands.

190 **[H3] RNA binders by design**

191 For the first strategy, a relevant example is represented by dimethylamiloride (DMA, compound **6** in
192 FIG. 3a) that, compared with aminoglycosides, showed lower affinity but promising selectivity.³¹
193 Iterative modifications at C(5)- and C(6)-positions allowed the establishment of strong interactions
194 with the bulge regions of HIV-1 TAR RNA. Derivative **7** (FIG. 3a) demonstrated a 100-fold increase
195 in activity over the parent compound **6** in Tat peptide displacement assays against TAR.³¹ This
196 suggests a stronger binding in a site that is particularly important for Tat interaction. The same
197 scaffold was modified to direct the ligands against other viral RNAs.³²⁻³⁴ Some DMA analogs, such as
198 compound **8** (FIG. 3a) showed promising antiviral activity by reducing SARS-CoV-2 virus load in
199 infected cells. These amiloride compounds bind to bulges and/or loop residues of 5'-UTR domains of
200 SARS-CoV-2 RNA genome owing to the interaction with specific binding sites.³⁴ Amilorides could
201 thus represent promising chemical probes for understanding SARS-CoV-2 pathology, even if they
202 generally show modest selectivity as they are also able to bind other RNAs or viral proteins.³⁵

203
204 Similar interactions were found for the triaminotriazine moiety that has been employed in the design
205 of various DNA and RNA ligands and is responsible for the recognition of T-T and U-U mismatches.

206 The two edges of the triazine heterocycle can form a full set of Janus-wedge hydrogen bonds with
207 poorly paired uracil and thymine residues (FIG. 3b) or can bind one of the two nucleobases and
208 displace the other one out of the helix. The addition of bisamidinium groove-binding linkers, as in
209 compounds **9-11** enabled preparation of multimeric compounds to target DNA and RNA triplet repeat
210 expansions, that are represented by abnormal repetition of three or more nucleotides in DNA and
211 consequently in RNA. These have been linked to a number of genetic diseases since they induce the
212 production of aberrant proteins. The presence of amidinium groups also added electrostatic
213 interactions with the phosphate backbone in the major groove of double helix.^{36,37} These compounds
214 were specific for the CUG expanded triplet repeats, r(CUG)^{exp}, found in the 3'-untranslated region of
215 the dystrophin myotonic protein kinase (DMPK) gene causing the myotonic dystrophy type 1
216 (DM1).³⁸ The pathology of DM1 is a result of splicing errors caused by these repeated RNA triplets,
217 thereby their targeting has been proposed for the discovery of treatments.³⁹ The biological evaluation
218 of compound **9** resulted in efficient inhibition of MBNL1-r(CUG)^{exp} interaction with micromolar
219 activities *in vitro* and in cells and paved the way to the synthesis of oligomers of compound **9** to
220 increase the potency by binding to multiple repeats.^{40,41}

221
222 Conjugation of an acridine intercalator with a triaminotriazine scaffold led to compound **12** (FIG.
223 3d).⁴² While the triazine heterocycle forms hydrogen bonds with poorly paired uracil residues, both
224 acridine and triaminotriazine moieties form important π -stacking interactions with paired bases,
225 eventually increasing affinity and selectivity, as illustrated by the crystal structure of CTG DNA repeat
226 expansion in complex with **12** (FIG. 3d).⁴³ While, it is clear that these compounds are selective for
227 r(CUG)^{exp} and r(CTG)^{exp}, they do not distinguish between DNA and RNA which limits their biological
228 applications.^{36,37}

229
230 Similarly, the naphthyridine scaffold has been employed in compounds with the aim of producing
231 DNA and RNA molecular glues.⁴⁴ This compound can form specific hydrogen bonds with unpaired
232 guanine and cytosine bases (FIG. 3e) and dimers have been prepared to control hybridization events
233 between two single-stranded DNAs and higher order RNA structures.^{45,46} For instance, compounds **13**
234 and **14** (FIG. 3e) are able to bind to the loop residues inducing the dimerization of two pre-miRNAs
235 structures. This tridimensional assembly prevents Dicer processing and represents an original
236 mechanism of inhibition.⁴⁷ Other types of naphthyridine derivatives, such as cyclic mismatch binding
237 ligands (CMBLs), were recently reported as rationally designed ligands to target CGG repeats in
238 RNA.⁴⁸

239 **[H3] RNA binders by high-throughput approaches**

240 In the context of the second type of targeting approaches, based on high-throughput techniques, a
241 remarkable study reported the integration of high-throughput sequencing with conventional SELEX
242 (Systematic Evolution of Ligands by Exponential enrichment) followed by bioinformatic analysis was

243 employed to identify hairpin loop sequences of pre-miRNA showing high binding affinity and
244 specificity for a water-soluble cyclic mismatch binding ligand **15** (FIG. 3f).⁴⁹ Binding was observed
245 for both endogenous pre-miR-33a and pre-miR-24-2 comprising characteristic binding motifs. Dose-
246 dependent decrease in the cleaved products formation was also observed in the presence of **15** for the
247 *in vitro* Dicer cleavage reaction of pre-miR33a thus confirming the potential of such an approach,
248 although no intracellular activity was reported.

249
250 While chemically distinct from aminoglycosides, guanidilated and protonated compounds maintain
251 interactions based on electrostatic forces, which can compromise selectivity. In this context, ribocil
252 (compound **16**, FIG. 3g) marks a substantial shift towards specific RNA binders with drug-like
253 properties.¹⁴ Identified as an inhibitor of the flavin mononucleotide (FMN) riboswitch, ribocil targets
254 these conserved prokaryotic RNAs that adjust gene expression upon binding to specific endogenous
255 ligands. The FMN riboswitch regulates riboflavin concentrations and controls genes related to its
256 biosynthesis and transport. A high-throughput screening highlighted the *S* enantiomer of ribocil as an
257 effective inhibitor of FMN, reducing riboflavin levels in *E. coli*. Various analogs, including ribocil D
258 (compound **17**, FIG. 3g), have been explored, revealing π -stacking interactions and crucial H-bonds
259 between the carbonyl, the 2'-OH of adenosine 48, and the exocyclic NH₂ of adenosine 99.⁵⁰ This
260 suggests that RNA binding does not strictly require protonated groups if hydrogen bonds are
261 established. Additional drug-like compounds, such as those binding to Xist, which plays a role in X
262 chromosome inactivation, further exemplify this trend.^{51,52}

263 **[H3] Integration of high-throughput and rational strategies**

264 From the examples above, it is clear that both the design based on the potential formed interactions
265 and high-throughput approaches are efficient strategies to discover selective ligands. The more is
266 known about RNA structure, the better the design and understanding of the interactions. A
267 combination of all these approaches would thus be ideal for future discovery of efficient RNA binders.

268
269 The closest approach to the rational design of RNA ligands based on the targeted structure is
270 represented by InfoRNA methodology that belongs to a series of recently developed discovery tools.⁵³
271 While most of these approaches, such as dynamic combinatorial chemistry (DCC) or DNA-encoded
272 libraries (DEL) are still being explored (see BOX 2), InfoRNA played a crucial role in the discovery
273 of new binders that have been validated *in vitro*, in cells and *in vivo*. InfoRNA methodology is based
274 on the 2-DCS approach described above, combined with a structure–activity relationships through
275 sequencing (StARTS) process – a statistical method for anticipating affinity and selectivity of RNA
276 library's compounds and so scores binding interactions.⁵⁴ This leads to the identification of the RNA
277 motifs that are the most suitable targets of a particular small-molecule binder. Finally, the obtained
278 database is mined against folded RNA structures within the human transcriptome to identify
279 potentially druggable RNA targets from sequence and the corresponding specific ligand.⁵⁵

280 The benzimidazole scaffold at the core of Hoechst derivatives that are very common fluorophores
281 known to interact with nucleic acids by intercalation and it was found in many of the RNA binders
282 identified by InfoRNA.⁵⁶ A typical example is compound **18** (FIG. 3h) that was identified as a specific
283 binder of miR-210 hairpin precursor. MiR-210 regulates hypoxia inducible factors (HIFs) that play
284 critical roles in cancer maintenance since a hypoxic state is critical to the metastatic and invasive
285 features of most solid cancers.²⁷ Compound **18** was able to bind to pre-miR-210 and inhibit its
286 processing toward mature miRNA through a specific interaction with 5'-ACU-3'/3'-UCA-5' internal
287 loop that represents its Dicer-cleavage site. This also induced a decrease of HIF-1 α and apoptosis of
288 triple negative breast cancer.²⁷ This highly specific effect shows that Hoechst intercalators can be
289 modified to specifically recognize a RNA region (sequence and structure) and to avoid off-target
290 effects.

291 In another remarkable example, compound **19** (FIG. 3i) was discovered with InfoRNA, to specifically
292 interact with a 1 x 1 UU internal loop located in the cleavage site of Dicer on pre-miR-200c, while
293 compound **20** was discovered as it interacts with adjacent AU base pair.⁵⁷ MiR-200c belongs to the
294 miR-200 family including five different members associated with type 2 diabetes (T2D).
295 Overexpression of miR-200c alone is sufficient to induce β cell apoptosis *in vivo*. Assembling
296 compounds **19** and **20**, led to compound **21** that is remarkably specific for miR-200c with respect to
297 the other members of the miR-200 family since it binds two different sites that are found only on pre-
298 miR-200c. Binding of this small molecule inhibited the biogenesis of miR-200c in a cellular model of
299 T2D thus reducing the amount of the miRNA as well as β cell apoptosis. Compound **21** has no other
300 targets across the miRnome, the transcriptome or the proteome, showing that conjugation of two
301 ligands specific for two different RNA sites could direct binding on a unique region where these two
302 sites are close and distinguish between similar pre-miRNAs sequences. Importantly, this compound is
303 specifically targeting pre-miR-200c that could not be targeted using an oligonucleotide, this latter
304 needing more extended single-stranded regions to bind. Using multiple methodologies, specific
305 compounds were identified to correct Tau pre-mRNA splicing linked to neurodegenerative diseases.⁵⁸
306 Use of InfoRNA applied to a library of RNA binders led to the identification of compound **22** (FIG.
307 3j). This was followed by a chemical similarity search for compounds bearing a chemical structure
308 close to **22** that led to analog **23**. NMR and fluorescence assays built a pharmacophore model, leading
309 to hit expansion and optimized compounds **24** and **25**. Compound **26** was finally obtained through
310 molecular docking, and bears micromolar activity *in vitro*, in cells, and *in vivo*. The comparison of the
311 formed interaction between the three compounds and the targets, as studied by NMR and molecular
312 docking, revealed that concomitant to the stacking interactions formed by the aromatic rings and the
313 RNA nucleobases, The *N*-methylcyclohexylamine group of **26** appears to form van der Waals
314 interactions with the wall of the major groove, probably increase affinity, specificity and eventually
315 the biological activity. This example highlights the importance of combining different techniques to
316 optimize RNA ligands and understand the underlying molecular mechanism. InfoRNA is a suitable

317 approach to identify RNA binders bearing a high and specific biological activity. The intracellular
318 activity has also been directly related to the RNA target upon target engagement experiments that
319 confirmed which intracellular RNA is responsible for the activity. In vivo confirmation of the activity
320 further highlighted the efficacy of this methodology.

321
322 In summary, the design of ligands tailored for the formation of strong interactions with the RNA target
323 can be very efficient but can also lead to poorly specific binders. Integrated approaches, such as
324 SELEX or InfoRNA as well as other methodologies described in Box 2, despite possible increased
325 failure rates, present a higher potential for the discovery of specific ligands. Nucleobase targeting by
326 small molecules predominantly relies on a structure-based approach over a sequence-based one.
327 Successful interactions depend on the target's three-dimensional environment and the potential for
328 RNA to adapt conformationally, creating specific binding sites.⁵⁹ Therefore, understanding the
329 relevant 3D structure is essential for the discovery of new ligands.

330
331 **[H2] Targeting dynamic 3D RNA structure.** The secondary and tertiary RNA structures of the target
332 and the pairing interactions involved in the formation of binding pockets, are strongly connected. A
333 further level of concept complexity is added by the fact that in most cases the RNA target bears more
334 than one possible conformation and that the conformations adopted in solution or in complex with the
335 intracellular partners are of major importance for its function.¹² While the identification of strong (at
336 least nanomolar) RNA binders is accessible, the determination of the biologically relevant RNA
337 structure is essential for the design of selective binders. Notably, ligand binding to an RNA target can
338 induce a conformational change, as demonstrated by the benzimidazole ligands of the IRES Iia
339 domain in the HCV RNA genome or by naphthyridine carbamate dimers (NCD) against
340 r(UGGAA)exp repeats in spinocerebellar ataxia type 31.^{59,60} Binding could thus involve induced fit
341 but also conformational selection mechanisms and this renders the understanding of the interactions
342 more challenging as well as the design of ligands appropriate for a specific target.⁶¹⁻⁶³ Structural
343 biology, biochemistry, biophysics and molecular modeling are thus essential fields for the design of
344 efficient binders. Even if algorithms are available to predict the RNA base pairing and the resulting
345 structure, experimental validation is essential to assess their reliability.

346
347 NMR remains a privileged technique to assess RNA structure and ligand binding.⁶⁴ These approaches
348 are mainly based on the observation of either the target or ligand ¹H nuclei, and have been developed
349 to study RNA dynamics (detection of dynamic sub-states) as well as RNA–ligand interactions.^{65,66} As
350 such, NMR studies allowed for the elucidation of risdiplam's mode of action, the first splicing
351 modifier approved for clinical use in 2020 for the treatment of SMA and initially identified through
352 phenotypic screening.⁶⁷ The binding mode of SMN-C5, risdiplam's analog, is shown in FIG. 4a, and
353 confirms that the compound acts by stabilizing a bulged adenine (A1). This induce a modified splicing

354 of Survival of Motor Neuron-2 protein (SMN2) that takes the place of the SMN1 defective protein.
355 Further to its ability to bind RNA through H-bonding and π -stacking interactions, risdiplam and its
356 analogs stabilize the interaction between the 5' splice site and the U1 small nuclear ribonucleoprotein
357 (snRNP) of the spliceosome, acting as a molecular glue between protein and RNA. ^1H NMR
358 experiments can be difficult to interpret when studying larger strands of RNA and in this case, it is
359 possible to use ^{19}F NMR to follow conformational changes and/or binding events. As an example,
360 chiral fluorinated diaminocyclopentanes (such as **27**, FIG. 4b) could bind to different RNA structures
361 leading to distinct ^{19}F NMR signal splittings.⁶⁸ Therefore, fluorinated ligands can be exploited as
362 sensitive NMR probes to investigate dynamic RNA structures, delivering a topological footprint of the
363 RNA. These small fluorinated reporters were also used for monitoring the binding of unlabeled RNA
364 with unlabeled ligands and to qualitatively rank different binders for a structured RNA.⁴⁵

365
366 Among the techniques available to study RNA secondary and tertiary structures, selective 2'-hydroxyl
367 acylation analyzed by primer extension (SHAPE) is largely applied to accurately identify paired and
368 unpaired bases on RNA sequences as well as transcriptome-wide.^{69,70} Mutational profiling (SHAPE-
369 MaP) makes the *de novo* and large-scale identification of RNA functional motifs possible.⁷¹ Indeed,
370 mutational profiling allows for the identification of sites prone to mutation that give a measure of local
371 flexibility and dynamics at the nucleotide level. This helped in the understanding of actual structures
372 that can be found *in vivo* and are ligands targets even if the dynamics of RNA and of ligand binding
373 events make the design of compounds appropriate for the desired biological activity very difficult.⁷²
374 Recently, SHAPE chemical probing of RNA and fragment-based ligand discovery were combined to
375 discover small fragments binding an RNA structure of interest, the thiamine pyrophosphate (TPP)
376 riboswitch, and to identify the individual sites of binding at roughly nucleotide resolution.⁷³ The
377 fragments were identified through a particularly clever workflow including a modular RNA screening
378 construct containing two target motifs that allowed the assessment of affinity and specificity during
379 the screening. The authors then linked fragments interacting with two different sites on the targeted
380 RNA to efficiently obtain optimized compounds, such as compounds **28** and **29** (FIG. 4c). Without
381 resemblance to the native ligand, they are not only druglike and have high ligand efficiency, but also
382 able to modulate RNA conformation during co-transcriptional folding.

383
384 A combination of artificial intelligence (AI)-augmented molecular simulations with experimental
385 techniques such as SHAPE-MaP was also recently employed to evaluate the predictability of RNA
386 dynamics and ligand binding events.^{74,12} For cognate and synthetic ligands of PreQ1 riboswitch,
387 unbiased molecular dynamics (MD) simulations with classical all-atom force-fields for RNA and
388 water can provide the same flexibility profile as measured in SHAPE experiments. Since the time
389 scales for ligand dissociation are far slower than MD, an AI-augmented MD method was developed
390 and could successfully reproduce relative binding affinity measurements obtained experimentally.

391 Furthermore, this methodology was able to highlight which nucleotides are essential for ligand
392 dissociation in a reliable way. These interactions would have been difficult to identify otherwise as
393 they are distant from the binding site and do not interact directly with the studied ligands.

394

395 Other structural biology approaches, especially crystallography and cryo-EM, can be employed to
396 better understand the three-dimensional structure of the target together with its dynamics, but can
397 rarely account for the entire conformational landscape of a particular RNA alone. It is thus important
398 that a combination of approaches are used to achieve the goal of identifying specific ligands for
399 functional structures towards a target biological output.

400

401 **[H1] New modalities for RNA binding**

402 Recent strategies have enabled the design of ligands with enhanced features, such as covalent binding
403 or cleavage properties, primarily driven by the need for target identification and validation. An
404 efficient method in this regard is chemical cross-linking and isolation by pull-down (Chem-CLIP).⁷⁵ In
405 this technique, an RNA-binding molecule is equipped with a chemical cross-linker, like nitrogen
406 mustard or diazirine, and a biotin purification tag (FIG. 5a).⁷⁶ Upon RNA binding, the cross linker is
407 brought into high concentration and undergoes a proximity-based reaction. Resulting covalent adducts,
408 purified *via* biotin, reveal the molecule's cellular targets. Chem-CLIP was employed to map binding
409 sites against oncogenic pre-miR-21, leading to the identification of fragments from 460 diazirine
410 compounds that bind RNA upon irradiation. Notably, combining promising fragments with a pre-miR-
411 21 binder produced compound **30** (FIG. 5b), which inhibited pre-miR-21 processing in breast cancer
412 cells. Covalent RNA ligand binding offers a promising approach for ligand discovery and intracellular
413 RNA target validation. Combined transcriptome-wide analyses can further elucidate intracellular
414 targets, as evidenced by studies on PreQ1 riboswitch and QSOX1 (quiescin sulfhydryl oxidase 1)
415 mRNA 5'-UTR binders.^{77,78} Such research shows the feasibility of studying small molecule-RNA
416 interactions in intact biological systems. and in an unbiased way.

417

418 Combining cleavage activity with RNA binding is another strategy to increase the biological activity
419 by permanently damage the RNA target. Inspiration for this came from RNase enzymes that cleave
420 RNA owing to the presence of one or two histidine residues in their catalytic site. These residues act
421 as acidic and basic imidazolium and imidazole units coordinating the phosphorous intermediate during
422 the transition state together with other protonated side chains such as lysine (FIG. 5c) inducing the
423 cleavage of the phosphodiester bond. Adding one or more histidine or imidazole residues to RNA
424 ligands such as intercalators,⁷⁹ peptides,⁸⁰ or aminoglycosides⁸¹ has also been applied as a strategy to
425 induce RNA cleavage, with the specificity dependent on the selectivity of the RNA binder.

426

427 Bleomycin is a known compound able to induce DNA and RNA cleavage. It is a natural product
428 extracted from *Streptomyces* and used in clinics as an anticancer agent with a complex mechanism of
429 action which includes topoisomerase inhibition.⁸² Bleomycin was also applied to RNA cleavage by
430 studying its preferential cleavage site and identifying it in the precursors of miR-10b (FIG. 5d).⁸³
431 Cleavage occurred close to the Drosha and Dicer cleavage site, essential for the production of mature
432 miRNA, and it was demonstrated *in vitro* and in cells to reduce the production of oncogenic miR-10b.
433 Bleomycin was then conjugated to binders specific for a particular RNA target such as targaprimiR-96
434 (compound **31**, FIG. 5b)⁸⁴ or ligands of r(CUG)^{exp} to induce a specific cleavage and degradation of the
435 target.⁸⁵

436
437 Altogether, these examples based on the conjugation of a covalent or cleavage moiety to an RNA
438 ligand illustrate the value of these methodologies not only in target validation studies, but also to
439 increase the actual biological effect when selectivity is not affected.

440
441 Starting from these findings and observations, a recent innovation in RNA targeting is the
442 development of ribonuclease targeting chimeras, or RIBOTACs.⁸⁶ These involve covalently linking a
443 specific RNA ligand to a molecule that activates RNase L, an RNA-degrading enzyme (FIG. 5e). This
444 approach has been applied to various RNA binders, including targapremir-210 (**18**, FIG. 3h) and has
445 been explored for targeting the RNA genome of SARS-CoV2, highlighting its potential in RNA
446 therapeutics.⁸⁷ Using InfoRNA, a binder was identified for the SARS-CoV-2 frameshifting element
447 RNA, that bears a 1x1 nucleotide UU internal loop crucial for frameshifting function.⁸⁷ Subsequent
448 detailed *in vitro* and *in cellulo* analysis led to the discovery of compound **32** (FIG. 5e), a ligand with
449 nanomolar affinity that stabilized the RNA hairpin structure and inhibited frameshifting. Chem-CLIP
450 was employed to ascertain the intracellular target of **32**, upon conjugation to a nitrogen mustard.
451 Additionally, **32** was modified to create RIBOTAC **33** (FIG. 5b), which induces targeted RNA
452 cleavage, in a similar manner to how PROTACs target proteins. This compound effectively degraded
453 the entire SARS-CoV construct. Optimization of the RIBOTAC strategy enhanced the compound's
454 bioactivity, as evidenced by intracellular luciferase reporter assays. This methodological combination
455 of chemical modification and use of various assays facilitated the identification of specific ligands that
456 effectively target cells without inducing toxicity. The question of the applicability of these complex
457 tools *in vivo* remains open, but the numerous examples reported so far hold the promise for further
458 developments in an analogous fashion to what happened for PROTACs.⁸⁸ The combination of RNA
459 binding and selectivity properties of a ligand with a chemical tool with covalent and/or cleavage
460 properties certainly represents an interesting approach for chemical biology and medicinal chemistry
461 applications. This leads to compounds that move away from the commonly accepted drug-like
462 properties⁵ but recent developments show that these criteria need to remain flexible to allow for new
463 modalities that may be beneficial for therapeutic discoveries.

464 [H1] Conclusion and outlook

465 Here, we have explored design strategies and chemical approaches that have been used in recent years
466 to approach RNA as a biological target. The understanding of RNA binders is continuously growing
467 and major improvements are in progress thanks to innovation in this fascinating field. RNA that was
468 once considered undruggable and almost impossible to be targeted specifically. However, it is now
469 considered a valuable target for the discovery of chemical probes to better understand its numerous
470 and elusive functions. Many efforts by the academic and private sector have aimed to expand the field
471 of RNA targeting by small molecules. Despite the numerous developments that have been performed
472 during the last decades, and even if the contribution for a better understanding of complex biological
473 processes is undeniable, it remains hard to predict if RNA binders will fulfill all expectations and open
474 new avenues for therapy. In this context, the future of the field is still uncertain, and more efforts are
475 needed to assess whether RNA targeting with small molecules will actually open new avenues to
476 answer important biological questions and discover new drugs.

477 The analysis of the RNA binders reported so far enables the identification of some common features,
478 such as the presence of protonated groups, the introduction of nitrogen-containing heteroaromatic
479 moieties or nucleobases and often the use of sp³-rich compounds. These latter are particularly
480 interesting in medicinal chemistry because they lead to better pharmacological profiles compared to
481 planar aromatic compounds, a concept that is called escape from flatland.⁸⁹ Drug-likeness has been
482 achieved successfully, but it is essential to allow for new modalities outside the common drug-like
483 landscape.

484 The most important open question concerns the lack of a standard methodology to obtain the desired
485 and specific biological effect with RNA-targeting small molecules. Despite relevant *in vivo* studies
486 and a few clinical successes, many reported ligands lack proof of specificity in cells and *in vivo* and
487 correlation between the observed biological effect and the specific targeting of the desired RNA.
488 Further studies are therefore needed, especially target validation experiments, structural biology to
489 determine the actual conformation and structure of the target and *in vivo* validation of the biological
490 activity. Also, it is hard to distinguish between biological activity induced by pure RNA binders, by
491 RNA-protein inhibitors or RNA-protein glues in a biological context where off target interactions
492 could reduce the biological activity even without inducing toxic effects. This is of course true for all
493 bioactive molecules, but is especially evident in the RNA-targeting field. A deep study of the site(s) of
494 interaction and the accurate choice of these sites that must be functional to lead to activity, is an
495 essential parameter to reach the desired activity. The first examples of the use of AI in the discovery of
496 RNA binders have already been reported and it is likely that it will continue to play a role in the field.

497 The tools presented here, including chemical strategies for the targeting of RNA primary, secondary
498 and tertiary structures and innovative approaches to understand RNA binding, function and biological

499 outcome represent the main basis for what should be the next step: rational design of ligands for
500 specific targets with potent biological activity. This will help to create new approaches to answer
501 complex biological questions and open new avenues for therapy. The upcoming years will be decisive
502 to confirm whether or not this optimistic view of the field is accurate.

503

505 **References**

- 506 1 Childs-Disney, J. L. *et al.* Targeting RNA structures with small molecules. *Nat. Rev. Drug*
507 *Discov.* **21**, 736-762 (2022).
- 508 2 Crooke, S. T., Baker, B. F., Crooke, R. M. & Liang, X. H. Antisense technology: an overview and
509 prospectus. *Nat. Rev. Drug Discov.* **20**, 427-453 (2021).
- 510 3 Roberts, T. C., Langer, R. & Wood, M. J. A. Advances in oligonucleotide drug delivery. *Nat.*
511 *Rev. Drug Discov.* **19**, 673-694 (2020).
- 512 4 Falese, J. P., Donlic, A. & Hargrove, A. E. Targeting RNA with small molecules: from
513 fundamental principles towards the clinic. *Chem. Soc. Rev.* **50**, 2224-2243 (2021).
- 514 5 Warner, K. D., Hajdin, C. E. & Weeks, K. M. Principles for targeting RNA with drug-like small
515 molecules. *Nat. Rev. Drug Discov.* **17**, 547-558 (2018).
- 516 6 Zamani, F. & Suzuki, T. Synthetic RNA Modulators in Drug Discovery. *J. Med. Chem.* **64**, 7110-
517 7155 (2021).
- 518 7 Magnet, S. & Blanchard, J. S. Molecular insights into aminoglycoside action and resistance.
519 *Chem. Rev.* **105**, 477-498 (2005).
- 520 8 Wilson, D. N. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nat.*
521 *Rev. Microbiol.* **12**, 35-48 (2014).
- 522 9 Thomas, J. R. & Hergenrother, P. J. Targeting RNA with small molecules. *Chem. Rev.* **108**,
523 1171-1224 (2008).
- 524 10 Ratni, H., Scalco, R. S. & Stephan, A. H. Risdiplam, the First Approved Small Molecule Splicing
525 Modifier Drug as a Blueprint for Future Transformative Medicines. *ACS Med. Chem. Lett.* **12**,
526 874-877 (2021).
- 527 11 Di Giorgio, A. & Duca, M. Synthetic small-molecule RNA ligands: future prospects as
528 therapeutic agents. *MedChemComm* **10**, 1242-1255 (2019).
- 529 12 Ganser, L. R. *et al.* High-performance virtual screening by targeting a high-resolution RNA
530 dynamic ensemble. *Nat. Struct. Mol. Biol.* **25**, 425-434 (2018).
- 531 13 Padroni, G., Patwardhan, N. N., Schapira, M. & Hargrove, A. E. Systematic analysis of the
532 interactions driving small molecule-RNA recognition. *RSC Med. Chem.* **11**, 802-813 (2020).
- 533 14 Howe, J. A. *et al.* Selective small-molecule inhibition of an RNA structural element. *Nature*
534 **526**, 672-677 (2015).
- 535 15 Wang, S., Huber, P. W., Cui, M., Czarnik, A. W. & Mei, H. Y. Binding of neomycin to the TAR
536 element of HIV-1 RNA induces dissociation of Tat protein by an allosteric mechanism.
537 *Biochemistry* **37**, 5549-5557 (1998).
- 538 16 Golkar, T. *et al.* Structural basis for plazomicin antibiotic action and resistance. *Commun. Biol.*
539 **4**, 729 (2021).
- 540 17 Aradi, K., Di Giorgio, A. & Duca, M. Aminoglycoside Conjugation for RNA Targeting:
541 Antimicrobials and Beyond. *Chem. Eur. J.* **26**, 12273-12309 (2020).
- 542 18 Bera, S., Mondal, D., Palit, S. & Schweizer, F. Structural modifications of the neomycin class of
543 aminoglycosides. *MedChemComm*, 1499 (2016).
- 544 19 Ennifar, E. *et al.* Structure-guided discovery of a novel aminoglycoside conjugate targeting
545 HIV-1 RNA viral genome. *ACS Chem. Biol.* **8**, 2509-2517 (2013).
- 546 20 Blount, K. F. & Tor, Y. A tale of two targets: differential RNA selectivity of nucleobase-
547 aminoglycoside conjugates. *ChemBioChem* **7**, 1612-1621 (2006).
- 548 21 Gebert, L. F. R. & MacRae, I. J. Regulation of microRNA function in animals. *Nat. Rev. Mol.*
549 *Cell. Biol.* **20**, 21-37 (2019).
- 550 22 Maucort, C. *et al.* Design and Implementation of Synthetic RNA Binders for the Inhibition of
551 miR-21 Biogenesis. *ACS Med. Chem. Lett.* **12**, 899-906 (2021).
- 552 23 Vo, D. D. *et al.* Building of neomycin-nucleobase-amino acid conjugates for the inhibition of
553 oncogenic miRNAs biogenesis. *Org. Biomol. Chem.* **16**, 6262-6274 (2018).

- 554 24 Vo, D. D. *et al.* Targeting the production of oncogenic microRNAs with multimodal synthetic
555 small molecules. *ACS Chem. Biol.* **9**, 711-721 (2014).
- 556 25 Vo, D. D. *et al.* Oncogenic MicroRNAs Biogenesis as a Drug Target: Structure-Activity
557 Relationship Studies on New Aminoglycoside Conjugates. *Chem. Eur. J.* **22**, 5350-5362 (2016).
- 558 26 Malnuit, V., Duca, M. & Benhida, R. Targeting DNA base pair mismatch with artificial
559 nucleobases. Advances and perspectives in triple helix strategy. *Org. Biomol. Chem.* **9**, 326-
560 336 (2011).
- 561 27 Costales, M. G. *et al.* Small Molecule Inhibition of microRNA-210 Reprograms an Oncogenic
562 Hypoxic Circuit. *J Am Chem Soc* **139**, 3446-3455 (2017).
- 563 28 Disney, M. D. *et al.* Two-dimensional combinatorial screening identifies specific
564 aminoglycoside-RNA internal loop partners. *J. Am. Chem. Soc.* **130**, 11185-11194 (2008).
- 565 29 Velagapudi, S. P. & Disney, M. D. Two-dimensional combinatorial screening enables the
566 bottom-up design of a microRNA-10b inhibitor. *Chem. Commun. (Camb)* **50**, 3027-3029
567 (2014).
- 568 30 Davis, B. *et al.* Rational design of inhibitors of HIV-1 TAR RNA through the stabilisation of
569 electrostatic "hot spots". *J. Mol. Biol.* **336**, 343-356 (2004).
- 570 31 Patwardhan, N. N. *et al.* Amiloride as a new RNA-binding scaffold with activity against HIV-1
571 TAR. *MedChemComm* **8**, 1022-1036 (2017).
- 572 32 Davila-Calderon, J. *et al.* IRES-targeting small molecule inhibits enterovirus 71 replication via
573 allosteric stabilization of a ternary complex. *Nat. Commun.* **11**, 4775 (2020).
- 574 33 Patwardhan, N. N., Cai, Z., Newson, C. N. & Hargrove, A. E. Fluorescent peptide displacement
575 as a general assay for screening small molecule libraries against RNA. *Org. Biomol. Chem.* **17**,
576 1778-1786 (2019).
- 577 34 Zafferani, M. *et al.* Amilorides inhibit SARS-CoV-2 replication in vitro by targeting RNA
578 structures. *Sci. Adv.* **7**, eabl6096 (2021).
- 579 35 Wilson, L., Gage, P. & Ewart, G. Hexamethylene amiloride blocks E protein ion channels and
580 inhibits coronavirus replication. *Virology* **353**, 294-306 (2006).
- 581 36 Hagler, L. D. *et al.* Versatile Target-Guided Screen for Discovering Bidirectional Transcription
582 Inhibitors of a Trinucleotide Repeat Disease. *ACS Med. Chem. Lett.* **12**, 935-940 (2021).
- 583 37 Hagler, L. D. *et al.* Expanded DNA and RNA Trinucleotide Repeats in Myotonic Dystrophy
584 Type 1 Select Their Own Multitarget, Sequence-Selective Inhibitors. *Biochemistry* **59**, 3463-
585 3472 (2020).
- 586 38 Mirkin, S. M. Expandable DNA repeats and human disease. *Nature* **447**, 932-940 (2007).
- 587 39 Wong, C. H. *et al.* Targeting toxic RNAs that cause myotonic dystrophy type 1 (DM1) with a
588 bisamidinium inhibitor. *J. Am. Chem. Soc.* **136**, 6355-6361 (2014).
- 589 40 Krueger, S. B., Lanzendorf, A. N., Jeon, H. H. & Zimmerman, S. C. Selective and Reversible
590 Ligand Assembly on the DNA and RNA Repeat Sequences in Myotonic Dystrophy.
591 *ChemBioChem* **23**, e202200260 (2022).
- 592 41 Pushechnikov, A. *et al.* Rational design of ligands targeting triplet repeating transcripts that
593 cause RNA dominant disease: application to myotonic muscular dystrophy type 1 and
594 spinocerebellar ataxia type 3. *J. Am. Chem. Soc.* **131**, 9767-9779 (2009).
- 595 42 Jahromi, A. H. *et al.* A novel CUG(exp).MBNL1 inhibitor with therapeutic potential for
596 myotonic dystrophy type 1. *ACS Chem. Biol.* **8**, 1037-1043 (2013).
- 597 43 Chien, C. M. *et al.* Structural Basis for Targeting T:T Mismatch with Triaminotriazine-Acridine
598 Conjugate Induces a U-Shaped Head-to-Head Four-Way Junction in CTG Repeat DNA. *J. Am.*
599 *Chem. Soc.* **142**, 11165-11172 (2020).
- 600 44 Murata, A., Otabe, T., Zhang, J. & Nakatani, K. BzDANP, a Small-Molecule Modulator of Pre-
601 miR-29a Maturation by Dicer. *ACS Chem. Biol.* **11**, 2790-2796 (2016).
- 602 45 Lombes, T. *et al.* Investigation of RNA-ligand interactions by 19F NMR spectroscopy using
603 fluorinated probes. *Angew. Chem. Int. Ed. Engl.* **51**, 9530-9534 (2012).
- 604 46 Murata, A., Nakamori, M. & Nakatani, K. Modulating RNA secondary and tertiary structures
605 by mismatch binding ligands. *Methods* **167**, 78-91 (2019).

606 47 Murata, A. *et al.* Small Molecule-Induced Dimerization of Hairpin RNA Interfered with the
607 Dicer Cleavage Reaction. *Biochemistry* **60**, 245-249 (2021).

608 48 Konieczny, P. *et al.* Cyclic mismatch binding ligands interact with disease-associated CGG
609 trinucleotide repeats in RNA and suppress their translation. *Nucleic Acids Res.* **49**, 9479-9495
610 (2021).

611 49 Mukherjee, S. *et al.* HT-SELEX-based identification of binding pre-miRNA hairpin-motif for
612 small molecules. *Mol. Ther. Nucleic Acids* **27**, 165-174 (2022).

613 50 Howe, J. A. *et al.* Atomic resolution mechanistic studies of ribocil: A highly selective unnatural
614 ligand mimic of the E. coli FMN riboswitch. *RNA Biol.* **13**, 946-954 (2016).

615 51 Aguilar, R. *et al.* Targeting Xist with compounds that disrupt RNA structure and X inactivation.
616 *Nature* **604**, 160-166 (2022).

617 52 Carrette, L. L. G. *et al.* A mixed modality approach towards Xi reactivation for Rett syndrome
618 and other X-linked disorders. *Proc. Natl. Acad. Sci. USA* **115**, E668-E675 (2018).

619 53 Ursu, A. *et al.* Design of small molecules targeting RNA structure from sequence. *Chem Soc*
620 *Rev* **49**, 7252-7270 (2020).

621 54 Velagapudi, S. P., Seedhouse, S. J. & Disney, M. D. Structure-activity relationships through
622 sequencing (StARTS) defines optimal and suboptimal RNA motif targets for small molecules.
623 *Angew. Chem. Int. Ed. Engl.* **49**, 3816-3818 (2010).

624 55 Disney, M. D. *et al.* Inforna 2.0: A Platform for the Sequence-Based Design of Small Molecules
625 Targeting Structured RNAs. *ACS Chem. Biol.* **11**, 1720-1728 (2016).

626 56 Velagapudi, S. P., Gallo, S. M. & Disney, M. D. Sequence-based design of bioactive small
627 molecules that target precursor microRNAs. *Nat. Chem. Biol.* **10**, 291-297 (2014).

628 57 Haniff, H. S. *et al.* A structure-specific small molecule inhibits a miRNA-200 family member
629 precursor and reverses a type 2 diabetes phenotype. *Cell Chem. Biol.* **29**, 300-311 e310
630 (2022).

631 58 Chen, J. L. *et al.* Design, Optimization, and Study of Small Molecules That Target Tau Pre-
632 mRNA and Affect Splicing. *J. Am. Chem. Soc.* **142**, 8706-8727 (2020).

633 59 Dibrov, S. M. *et al.* Structure of a hepatitis C virus RNA domain in complex with a translation
634 inhibitor reveals a binding mode reminiscent of riboswitches. *Proc. Natl. Acad. Sci. USA* **109**,
635 5223-5228 (2012).

636 60 Shibata, T. *et al.* Small molecule targeting r(UGGAA)(n) disrupts RNA foci and alleviates
637 disease phenotype in Drosophila model. *Nat. Commun.* **12**, 236 (2021).

638 61 Boehr, D. D., Nussinov, R. & Wright, P. E. The role of dynamic conformational ensembles in
639 biomolecular recognition. *Nat. Chem. Biol.* **5**, 789-796 (2009).

640 62 Nakatani, K. Possibilities and challenges of small molecule organic compounds for the
641 treatment of repeat diseases. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **98**, 30-48 (2022).

642 63 Vogt, A. D. & Di Cera, E. Conformational selection or induced fit? A critical appraisal of the
643 kinetic mechanism. *Biochemistry* **51**, 5894-5902 (2012).

644 64 Bailor, M. H., Mustoe, A. M., Brooks, C. L., 3rd & Al-Hashimi, H. M. Topological constraints:
645 using RNA secondary structure to model 3D conformation, folding pathways, and dynamic
646 adaptation. *Curr. Opin. Struct. Biol.* **21**, 296-305 (2011).

647 65 Marusic, M., Toplishek, M. & Plavec, J. NMR of RNA - Structure and interactions. *Curr. Opin.*
648 *Struct. Biol.* **79**, 102532 (2023).

649 66 Shortridge, M. D. *et al.* Drug-Like Small Molecules That Inhibit Expression of the Oncogenic
650 MicroRNA-21. *ACS Chem Biol* **18**, 237-250 (2023).

651 67 Campagne, S. *et al.* Structural basis of a small molecule targeting RNA for a specific splicing
652 correction. *Nat. Chem. Biol.* **15**, 1191-1198 (2019).

653 68 Moumne, R. *et al.* Fluorinated diaminocyclopentanes as chiral sensitive NMR probes of RNA
654 structure. *J. Am. Chem. Soc.* **132**, 13111-13113 (2010).

655 69 Bevilacqua, P. C., Ritchey, L. E., Su, Z. & Assmann, S. M. Genome-Wide Analysis of RNA
656 Secondary Structure. *Annu. Rev. Genet.* **50**, 235-266 (2016).

657 70 Deigan, K. E., Li, T. W., Mathews, D. H. & Weeks, K. M. Accurate SHAPE-directed RNA
658 structure determination. *Proc. Natl. Acad. Sci. USA* **106**, 97-102 (2009).

659 71 Siegfried, N. A., Busan, S., Rice, G. M., Nelson, J. A. & Weeks, K. M. RNA motif discovery by
660 SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods* **11**, 959-965 (2014).

661 72 Weeks, K. M. SHAPE Directed Discovery of New Functions in Large RNAs. *Acc. Chem. Res.* **54**,
662 2502-2517 (2021).

663 73 Zeller, M. J. *et al.* SHAPE-enabled fragment-based ligand discovery for RNA. *Proc. Natl. Acad.*
664 *Sci. USA* **119**, e2122660119 (2022).

665 74 Wang, Y., Parmar, S., Schneekloth, J. S. & Tiwary, P. Interrogating RNA-Small Molecule
666 Interactions with Structure Probing and Artificial Intelligence-Augmented Molecular
667 Simulations. *ACS Cent. Sci.* **8**, 741-748 (2022).

668 75 Velagapudi, S. P., Li, Y. & Disney, M. D. A cross-linking approach to map small molecule-RNA
669 binding sites in cells. *Bioorg Med Chem Lett* **29**, 1532-1536 (2019).

670 76 Suresh, B. M. *et al.* A general fragment-based approach to identify and optimize bioactive
671 ligands targeting RNA. *Proc. Natl. Acad. Sci. USA* **117**, 33197-33203 (2020).

672 77 Balaratnam, S. *et al.* A chemical probe based on the PreQ(1) metabolite enables
673 transcriptome-wide mapping of binding sites. *Nat Commun* **12**, 5856 (2021).

674 78 Tong, Y. *et al.* Transcriptome-Wide Mapping of Small-Molecule RNA-Binding Sites in Cells
675 Informs an Isoform-Specific Degradator of QSOX1 mRNA. *J Am Chem Soc* **144**, 11620-11625
676 (2022).

677 79 Vlassov, V. V., Zuber, G., Felden, B., Behr, J. P. & Giege, R. Cleavage of tRNA with imidazole
678 and spermine imidazole constructs: a new approach for probing RNA structure. *Nucleic Acids*
679 *Res.* **23**, 3161-3167 (1995).

680 80 Tamkovich, N. *et al.* Design, RNA cleavage and antiviral activity of new artificial ribonucleases
681 derived from mono-, di- and tripeptides connected by linkers of different hydrophobicity.
682 *Bioorg. Med. Chem.* **24**, 1346-1355 (2016).

683 81 Martin, C. *et al.* Design, Synthesis, and Evaluation of Neomycin-Imidazole Conjugates for RNA
684 Cleavage. *ChemPlusChem* **87**, e202200250 (2022).

685 82 Hecht, S. M. Bleomycin: new perspectives on the mechanism of action. *J Nat Prod* **63**, 158-
686 168 (2000).

687 83 Angelbello, A. J. & Disney, M. D. Bleomycin Can Cleave an Oncogenic Noncoding RNA.
688 *ChemBioChem* **19**, 43-47 (2018).

689 84 Li, Y. & Disney, M. D. Precise Small Molecule Degradation of a Noncoding RNA Identifies
690 Cellular Binding Sites and Modulates an Oncogenic Phenotype. *ACS Chem. Biol.* **13**, 3065-
691 3071 (2018).

692 85 Angelbello, A. J. *et al.* Precise small-molecule cleavage of an r(CUG) repeat expansion in a
693 myotonic dystrophy mouse model. *Proc. Natl. Acad. Sci. USA* **116**, 7799-7804 (2019).

694 86 Liu, X. *et al.* Targeted Degradation of the Oncogenic MicroRNA 17-92 Cluster by Structure-
695 Targeting Ligands. *J. Am. Chem. Soc.* **142**, 6970-6982 (2020).

696 87 Haniff, H. S. *et al.* Targeting the SARS-CoV-2 RNA Genome with Small Molecule Binders and
697 Ribonuclease Targeting Chimera (RIBOTAC) Degradators. *ACS Cent. Sci.* **6**, 1713-1721 (2020).

698 88 Li, K. & Crews, C. M. PROTACs: past, present and future. *Chem Soc Rev* **51**, 5214-5236 (2022).

699 89 Lovering, F., Bikker, J. & Humblet, C. Escape from flatland: increasing saturation as an
700 approach to improving clinical success. *J Med Chem* **52**, 6752-6756 (2009).

701 90 Morgan, B. S. & Hargrove, A. E. Synthetic Receptors for Biomolecules: Design Principles and
702 Applications. *The Royal Society of Chemistry*, 253-325 (2015).

703 91 Kenyon, J., Prestwood, L. & Lever, A. Current perspectives on RNA secondary structure
704 probing. *Biochem Soc Trans* **42**, 1251-1255 (2014).

705 92 Batey, R. T., Rambo, R. P. & Doudna, J. A. Tertiary Motifs in RNA Structure and Folding.
706 *Angew Chem Int Ed Engl* **38**, 2326-2343 (1999).

707 93 Butcher, S. E. & Pyle, A. M. The molecular interactions that stabilize RNA tertiary structure:
708 RNA motifs, patterns, and networks. *Acc. Chem. Res.* **44**, 1302-1311 (2011).

- 709 94 Jones, C. P. & R., F.-D. A. A. RNA quaternary structure and global symmetry. *Trends in*
710 *Biochemical Sciences* **40**, 211-220 (2015).
- 711 95 Donlic, A. *et al.* R-BIND 2.0: An Updated Database of Bioactive RNA-Targeting Small
712 Molecules and Associated RNA Secondary Structures. *ACS Chem. Biol.* **17**, 1556-1566 (2022).
- 713 96 Morgan, B. S., Forte, J. E., Culver, R. N., Zhang, Y. & Hargrove, A. E. Discovery of Key
714 Physicochemical, Structural, and Spatial Properties of RNA-Targeted Bioactive Ligands.
715 *Angew. Chem. Int. Ed. Engl.* **56**, 13498-13502 (2017).
- 716 97 Xie, J. & Frank, A. T. Mining for Ligandable Cavities in RNA. *ACS Med. Chem. Lett.* **12**, 928-934
717 (2021).
- 718 98 Murray, C. W. & Rees, D. C. The rise of fragment-based drug discovery. *Nat. Chem.* **1**, 187-
719 192 (2009).
- 720 99 Umuhire Juru, A., Cai, Z., Jan, A. & Hargrove, A. E. Template-guided selection of RNA ligands
721 using imine-based dynamic combinatorial chemistry. *Chem. Commun. (Camb)* **56**, 3555-3558
722 (2020).
- 723 100 Rzuczek, S. G., Park, H. & Disney, M. D. A toxic RNA catalyzes the in cellulo synthesis of its
724 own inhibitor. *Angew. Chem. Int. Ed. Engl.* **53**, 10956-10959 (2014).
- 725 101 Benhamou, R. I. *et al.* DNA-encoded library versus RNA-encoded library selection enables
726 design of an oncogenic noncoding RNA inhibitor. *Proc. Natl. Acad. Sci. USA* **119** (2022).
- 727 102 Rizvi, N. F. *et al.* Discovery of Selective RNA-Binding Small Molecules by Affinity-Selection
728 Mass Spectrometry. *ACS Chem. Biol.* **13**, 820-831 (2018).

729

730

731 **Acknowledgements**

732 The authors thank the Italian Institute of Technology (IIT) and Université Côte d'Azur for financial
733 support to M.P. as PhD fellowship, respectively. The authors also thank BoostUrCareer COFUND
734 project with Région PACA and European Union's Horizon 2020 research and innovation program
735 under grant agreement no. 847581 for PhD fellowship to S.K.

736 **Author contributions**

737 All authors researched data for the article. S.K., M.P. contributed substantially to the illustrations describing the reported examples. M.D., S.A., A.D.G contributed substantially to
738 discussion of the content. M.D. wrote the article. All authors reviewed and/or edited the manuscript before submission.

739 **Competing interests**

740 The authors declare no competing interests.

741 **Peer review information**

742 *Nature Reviews Chemistry* thanks Matthew Disney, Jessica Bush, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
743
744

745 Fig. 1. | **Historical view of RNA ligands as marketed drugs.** Timeline illustrating the main landmarks in the field of RNA
746 targeting with small molecules and examples of marketed drugs.

747

748 Fig. 2 | **RNA binders based on aminoglycoside structure.** **a** | On the left, crystal structure of plazomicin **1** within its A-site
749 binding site. The ligand is shown in orange sticks while contact residues are shown in light blue sticks, H-bonds are shown as
750 black dashes (PDB ID: 71H5). On the right, 2D representation of plazomicin and the contact nucleotides in the binding site. **b**
751 | On the left, crystal structure of neomycin within the DIS kissing complex, target residue for specific interaction A16 is
752 shown as blue sticks (PDB ID: 2FCY). Comparison between the H-bond A16-Neomycin-ring I and A16-Neomycin-thymine
753 is shown below. On the right 2D representation of Neomycin-Thymine conjugate **2** and the contact nucleotides in the binding
754 site. **c** | Neomycin-nucleobase conjugates **3**; **d** | Neomycin-artificial nucleobase conjugate **4** and its binding to stem-loop
755 structured precursor of oncogenic miR-372 as confirmed by footprinting and molecular docking. **4** interacts predominantly at
756 the stem region (U12-G14 and A40-G42) close to the site of cleavage by Dicer. **e** | Guanidilated neomycin **5** identified upon
757 2-DCS as a specific binder of the internal loop of pre-miR-10b.

758

759 Fig. 3. | **RNA binders targeting secondary structures.** **a** | Compound **6** optimization led to **7** and **8**; **b** | Janus-wedge
760 hydrogen bonds formed by the triaminotriazine moiety; **c** | Cartoon representation of a CUG repeat (PDB ID: 3GM7) with
761 highlighted the binding residues for compounds **9-11** in sticks; **d** | Compound **12** binding to DNA CTG repeats: the
762 triaminotriazine forms H-bonds with T10 inducing a displacement of T4 (represented in dark blue sticks) while the acridine
763 intercalates between T10 and C3-G11 pair (PDB ID: 6M4T); **e** | Naphthyridines **13-15** forming H-bonds with C and G; **f** |
764 Ribocil (**16**) and Ribocil-D (**17**) chemical structures and binding mode in FMN riboswitch (PDB ID: 5KX9). **g** | InfoRNA
765 pipeline led to Targapremir-210 (**18**); **h** | Discovery of compound **21** starting from **19** binding to the U-U internal loop, and
766 **20** binding to the adjacent U-A pair; **i** | Structure-based drug design of **26**. The apo NMR structure bearing the A bulge
767 (highlighted in light orange) is shown on the left (PDB ID: 6VA1). The complex between **22** and the RNA results in A6
768 displacement and in the design of **26** whose binding mode is shown on the left (PDB ID 6VA4).

769

770 Fig. 4. | **NMR and SHAPE methodologies for RNA structure determination.** **a** | SMN **C5** molecular glue between U1
771 snRNA (represented in light blue) and E7 5' SS (represented in white) by stabilizing a bulged A (white sticks) inside the
772 duplex. Binding mode determined by NMR (PDB ID:6HMO); **b** | Fluorinated probe **27** binds to different RNA
773 conformations existing in solution and thus the conformational ensemble can be studied by ¹⁹F NMR. Upon binding of the
774 unmarked ligand, the selection of the bound conformation and the perturbation of the equilibrium can be monitored by ¹⁹F
775 NMR; **c** | Cartoon representation of SHAPE-MaP methodology. Binding events perturb SHAPE signal and the comparison
776 between bound and unbound structure leads to identification of the binding nucleotides (indicated in green). Binding
777 fragments **1** and **2** are next linked together to improve binding affinity, leading to compound **28**. Binding mode elucidation of
778 **28** by X-ray inside the TPP riboswitch (PDB ID: 7TZU, ligand **28** represented with green sticks, key contact nucleotides are
779 shown) and subsequent structure based molecular modeling led to the design of optimized compound **29**.

780

781 Fig. 5 | **Achieving RNA binding with additional functionalities.** **a** | Chem-CLIP methodology for target engagement or
782 binding site mapping through covalent binding; **b** | RNA ligands conjugated with imidazole are able to induce RNA cleavage,
783 the mechanism is shown in the grey box; **c** | Bleomycin conjugated to efficient RNA ligands induces cleavage of the RNA
784 target; **d** | Ribonuclease targeting chimeras (RIBOTACs) are bivalent compounds able to recruit RNases and induce RNA
785 cleavage; **e** | Chemical structures of pre-miR-21 binder **30** identified via Chem-CLIP-Frag-Map, Bleomycin conjugate of
786 targapri-miR96 **31** and RIBOTAC **32** against the SARS-CoV-2 RNA genome.

788 **Box 1 – Primary, secondary and tertiary structures of RNA.**

789 Ribonucleic acids (RNAs) are multifaceted molecules that accomplish various tasks in living cells by
790 acquiring complex and unique 3D structures. They are formed through different levels of organization,
791 including primary, secondary, tertiary, and quaternary structures and provide useful sites for selective
792 identification by proteins or small molecules. The primary structure of RNA consists of a linear
793 sequence of nucleotides, each comprising a nucleobase, a ribose sugar, and a phosphate group that
794 links the sugars through phosphodiester bonds. The four nucleobases: guanine (G), adenine (A),
795 cytosine (C), and uracil (U), attach to the backbone with, respectively, their N9 for purines (A and G)
796 and their N1 for pyrimidines (C and U) and the ribose sugars.⁹⁰

797 RNA secondary structures result from hydrogen bonds between nucleobases, forming various folding
798 patterns. The most common secondary structure is the Watson-Crick-Franklin canonical base pairing
799 between A-U and C-G base pairs, resulting in RNA double helices with an A-form, instead of the B-
800 form found in DNA due to a hydroxyl group at the 2' position of the sugar. Other secondary structures
801 such as internal loops, hairpin loops, and bulges are formed by single-stranded sections of
802 nucleotides.⁹¹ Overall, secondary structures are behind the formation of 3D arrangements of RNA,
803 called tertiary structures.^{92,93}

804 RNA tertiary structures arise from connections between secondary structures, leading to specific
805 tridimensional arrangements. The stacking of bases increases the likelihood of these interactions,
806 leading to complex structures such as pseudoknots and kissing loop complexes.⁹⁴ G-quadruplexes,
807 riboswitches and triple helices are also part of RNA tertiary structures. RNA quaternary structures
808 occur when RNA interacts with other biomolecules to carry out its biological function.

809

810

811 **Box 2 - Integrated technologies for the discovery of RNA ligands**

812 Alongside the development of predictive approaches and chemoinformatic tools to describe the main
813 features for RNA targeting, such as R-BIND (RNA-targeted Bioactive Ligand Database), innovative
814 screening methodologies have been developed for the discovery of RNA binders and have been
815 integrated to the methodologies described in this manuscript.⁹⁵⁻⁹⁷ Fragment-based drug discovery
816 (FBDD),⁹⁸ coupled to Chem-CLIP or SHAPE, has been particularly useful for the discovery of
817 original chemical scaffolds.^{73,76}

818 Dynamic combinatorial chemistry (DCC, a) is a powerful strategy to identify specific bioactive
819 compounds. Compounds able to assemble and form a dynamic library *via* a reversible reaction are put
820 in the presence of the biological target to favor the formation of the strongest binder(s).⁹⁹ *In situ*
821 approaches have also provided informative results. Potent modulators of RNA function can be
822 assembled *in cellulo* by using cells as a reaction vessel and a disease-causing RNA as a catalyst.¹⁰⁰

823 The use of DNA-encoded libraries (DEL, b) has been applied in the search for RNA binders. For
824 example, a DNA-encoded library was screened against a library of RNA structures to evaluate 300
825 million interactions in total, allowed for the identification of numerous ligand/target pairs.¹⁰¹

826 High-throughput screenings have been also been applied to the discovery of RNA binders. Alongside
827 phenotypic screenings, various recent innovative methodologies have been reported. Among them,
828 Automated Ligand Detection System (ALIS, c), an indirect affinity-selection mass spectrometry (AS-
829 MS) has been employed for the high-throughput screening of unbiased libraries of small molecules for
830 ncRNAs interactions.^{102, 52} Catalytic enzyme-linked click chemistry assays (cat-ELCCA) have also
831 been applied to the screening for compounds able to inhibit pre-let7/Lin28 interactions. The main
832 advantages of these assays are the robustness, the high sensitivity owing to catalytic signal
833 amplification and the negligible compound interference compared to other fluorescence-based assays.

834

835

836 **Short summary**

837 This review highlights the strategies and challenges for targeting RNA with small molecules in
838 medicinal chemistry. It emphasizes their potential as drugs and tools for understanding complex
839 biological processes while encouraging chemists to contribute to this field for future advances.

840

841

842