Nucleic Acids Research, 2025, **53**, gkaf559
https://doi.org/10.1093/nar/gkaf559
Chemical Biology and Nucleic Acid Chemistry

Mechanistic studies of small molecule ligands selective to RNA single G bulges

Shalakha Hegde[1,†], Sana Akhter[2,†], Zhichao Tang[1,†], Chang Qi[3], Chenguang Yu[4], Anna Lewicka[5], Yu Liu[6], Kushal Koirala[2], Mikhail Reibarkh[3], Kevin P. Battaile[7], Anne Cooper[8], Scott Lovell[8], Erik D. Holmstrom[9], Xiao Wang[3], Joseph A. Piccirilli[5,10], Qi Gao[3,*], Yinglong Miao[2,*], Jingxin Wang[1,*]

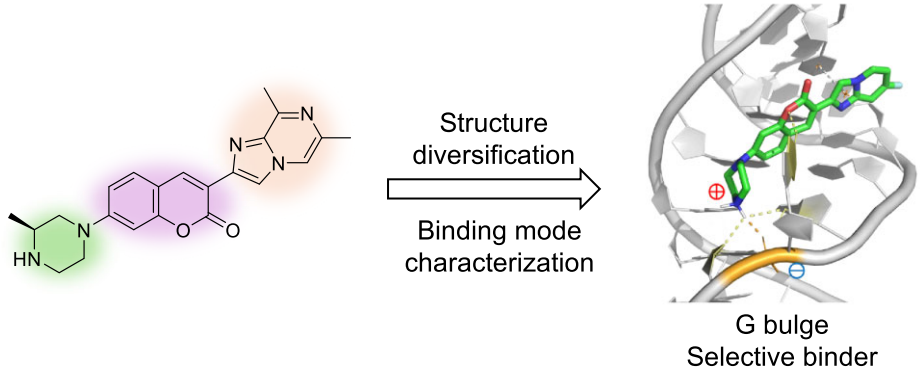
[1]Section of Genetic Medicine, Department of Medicine, Biological Sciences Division, University of Chicago, Chicago, IL 60637, United States
[2]Department of Pharmacology and Computational Medicine Program, University of North Carolina, Chapel Hill, NC 27599, United States
[3]Analytical Research & Development, Merck & Co., Inc, Rahway, NJ 07065, United States
[4]Calibr–Skaggs Institute for Innovative Medicines, The Scripps Research Institute, La Jolla, CA 92037, United States
[5]Department of Biochemistry and Molecular Biology, Biological Sciences Division, University of Chicago, Chicago, IL 60637, United States
[6]Department of Chemistry, Rockhurst University, Kansas City, MO 64110, United States
[7]New York Structural Biology Center, Upton, NY 11973, United States
[8]Protein Structure and X-ray Crystallography Laboratory, University of Kansas, Lawrence, KS 66047, United States
[9]Department of Molecular Biosciences, University of Kansas, Lawrence, KS 66045, United States
[10]Department of Chemistry, Physical Sciences Division, University of Chicago, Chicago, IL 60637, United States

[*]To whom correspondence should be addressed. Email: wang.jingxin@uchicago.edu
Correspondence may also be addressed to Yinglong Miao. Email: yinglong_miao@med.unc.edu
Correspondence may also be addressed to Qi Gao. Email: qi.gao1@merck.com
[†]The first three authors should be regarded as Joint First Authors.

Abstract

Small-molecule RNA binders have emerged as an important pharmacological modality. A profound understanding of the ligand selectivity, binding mode, and influential factors governing ligand engagement with RNA targets is the foundation for rational ligand design. Here, we report a novel class of coumarin derivatives exhibiting selective binding affinity towards single G RNA bulges. Harnessing the computational power of all-atom Gaussian accelerated molecular dynamics simulations, we unveiled a rare minor groove binding mode of the ligand with a key interaction between the coumarin moiety and the G bulge. This predicted binding mode is consistent with results obtained from structure-activity relationship studies and transverse relaxation measurements by nuclear magnetic resonance spectroscopy. We further generated 444 molecular descriptors from 69 coumarin derivatives and identified key contributors to the binding events, such as charge state and planarity, by lasso (least absolute shrinkage and selection operator) regression. Our work deepened the understanding of RNA-small molecule interactions and integrated a new framework for the rational design of selective small-molecule RNA binders.

Graphical abstract



Introduction

RNA plays critical roles in gene regulation and various cellular processes in almost all life forms, including transcription, translation, splicing, and epigenetic modifications [1, 2]. Selective targeting of RNA structures using small molecules is an important pharmacological modality that complements

Received: November 6, 2024. **Revised:** April 22, 2025. **Editorial Decision:** May 10, 2025. **Accepted:** June 4, 2025

traditional protein targeting approaches [3–9]. For example, bacteria ribosomal RNA (rRNA) is an important antibiotic target with numerous clinically validated drug classes, such as aminoglycoside, tetracycline, macrolide, lincosamide, and oxazolidinone [10]. Recently, two synthetic compounds, risdiplam and branaplam, both targeting precursor messenger RNA (pre-mRNA)-U1 small nuclear ribonucleoprotein (snRNP) complex, attracted tremendous attention as RNA splicing modulators to treat genetic diseases, including spinal muscular atrophy [11–17] and Huntington's disease [18–20]. We previously demonstrated that a class of coumarin analogues of risdiplam can induce GA-rich sequences to form loop-like structures using molecular dynamics (MD) simulations and proposed that this interaction in cells provided additional selectivity of the coumarin derivatives to the GA-rich SMN2 gene [21].

In addition to rRNA in bacteria and pre-mRNA in humans, several other classes of RNA have been targeted by chemical probes and drug candidates, including bacteria riboswitches [22–24], yeast self-splicing introns [25], microRNAs [26–31], untranslated regions (UTR) of mRNAs [32–34], and long non-coding RNAs [35–37]. In viruses, highly structured RNA regions have also been explored as targets for small molecules [38], such as an internal ribosome entry site in the 5′ UTR of the hepatitis C virus [39–42] and a transactivation response hairpin in human immunodeficiency virus 1 [43–45]. After the outbreak of SARS-CoV-2, we and others illustrated that the structural elements in the SARS-CoV-2 genome can also be targeted to suppress virus replication [46–51]. Specifically, we discovered that some coumarin derivatives (Fig. 1) can be repurposed to selectively bind to a single G bulge in 5′ UTR of SARS-CoV-2 without retaining splicing modulatory activities or binding to GA-rich loops [47]. We further demonstrated that covalently linking a ribonuclease (RNase) L recruiter and the coumarin-based G bulge binder yielded an active ribonuclease targeting chimera, which is effective in targeting SARS-CoV-2-infected epithelial cells [47].

In general, attaining selective and effective targeting of RNA using small molecules is a challenging endeavour due to several factors, including its conformational flexibility and heterogeneity and its polyanionic backbone, which all prevents the formation of deep hydrophobic binding pockets for small-molecule binding [8, 52]. Cheminformatics works have uncovered key factors governing the activity and selectivity of RNA binders [53–56] and have recently been further advanced by machine learning approaches [57]. However, the lack of methods for mechanistic studies of flexible RNA–small molecule ligand interactions critically limits further optimization of RNA ligands. A powerful approach to studying RNA-small molecule interactions is to use molecular dynamics (MD) simulations [21, 58, 59], which are able to fully account for the RNA flexibility on an atomistic level. Here, we present an integrated framework that combines all-atom Gaussian accelerated MD (GaMD) simulations, which can rapidly predict ligand binding modes, with nuclear magnetic resonance (NMR) transverse relaxation (R_2) measurements and structure-activity relationship (SAR) studies, which can experimentally probe RNA–ligand interactions [21, 60]. We envision that our mechanistic studies of RNA ligands could serve as a model for understanding interactions between RNA and small molecules and for enhancing binding affinity across various RNA targets.

Materials and methods

Fluorescence polarization assay

Synthetic RNA oligonucleotides were procured from GenScript (Piscataway, NJ, USA) and reconstituted in nuclease-free water. Compounds were prepared at a concentration of 50 μM in dimethyl sulfoxide (DMSO) and were diluted in 2× assay buffer (50 mM 2-(*N*-morpholino)ethanesulfonic acid [MES], 100 mM NaCl, 0.004% TritonX, pH 6.1) to a concentration of 0.2 μM. A 1:3 dilution series of six points was prepared for each RNA in 20 μl water, resulting in concentrations ranging from 0.1 to 20 μM. Subsequently, 20 μl of 2× working solution containing the assay buffer and small-molecule ligand was added to each RNA sample in a 1:1 (v/v) ratio and mixed by pipetting. To measure fluorescence polarization (FP), 8 μl of the sample solution was transferred into a 384-well, black, flat-bottom microplate (Greiner, #784076) in triplicate. The plate was equilibrated at room temperature for 5 min prior to data acquisition using BioTek Synergy H1 (Winooski, VT, USA) with a blue FP filter cube (excitation/emission = 360/460 nm; BioTek, #8040563) at 25°C. Experimental data were analysed using the Prism 8.0 software package (GraphPad, San Diego, USA). A nonlinear curve fitting was employed to calculate the dissociation constant (K_d), reported with a 95% confidence interval.

Lasso regression

The structure of each of the 69 compounds was individually optimized using density-functional theory (DFT) calculation performed with Gaussian 09 software, employing the B3LYP functional and 6-31G(d) basis set in the ground state with default convergence criteria (see Supplementary data). The protonation state was predicted by Molecular Operating Environment (MOE) 2022 software (Chemical Computing Group, Montreal, Canada) at pH 7.0. The structures of all compounds were loaded on MOE 2022, and 443 molecular descriptors were generated using MOE 2022. Two molecular descriptors, h_pKa and h_pKb, were excluded because of the differential protonation states within the compound library. The dihedrals in degree units (°) between BC and DE rings within the DFT-optimized 3-dimensional (3D) structures were added as a new molecular descriptor. In the lasso regression analysis, the natural logarithm of the dissociation constant [Ln(K_d), K_d in molar unit] was utilized as the dependent variable (y-value) (for the molecular descriptor values used for lasso regression, see MolecularDescriptors.csv). Lasso regression (L1 regularization) was performed in R (4.2.2) according to the reported protocol [61] (see LassoRegression.rmd for the R code). The non-zero coefficients were determined as described in Supplementary Table S1.

GaMD simulations

GaMD is a computational enhanced sampling technique in which a harmonic boost potential is added to smooth the potential energy surface and reduce the system energy barriers [60]. GaMD does not require predefined collective variables or reaction coordinates and is thus advantageous for unconstrained enhanced sampling of complex biological systems such as ligand binding to highly flexible RNA. Since the GaMD boost potential exhibits a Gaussian distribution, the original biomolecular free energy profiles can be properly recovered through cumulant expansion to the second order [60].

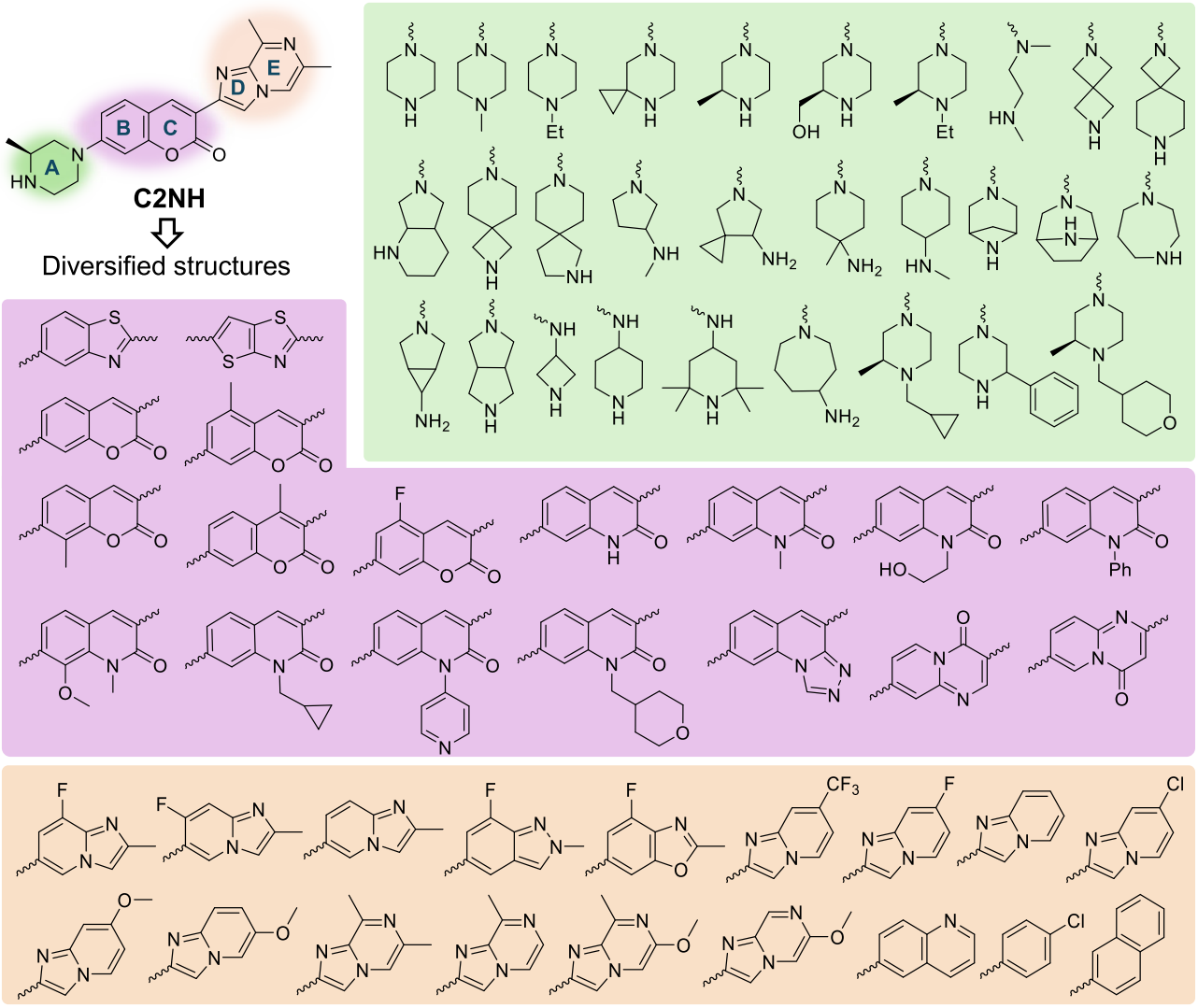


Figure 1. Molecular diversity of coumarin derivative analogues designed to bind bulged G RNA. The A, BC, and DE rings were replaced by the shaded green, purple, and orange structures, respectively (wavy lines = connecting bonds).

GaMD has been successfully demonstrated on biomolecular simulations and revealed physical pathways and mechanisms of protein folding and ligand binding, which were consistent with experiments and long-timescale conventional MD simulations [62–66]. Recently, it was also applied to successfully capture ligand binding to single-stranded RNA [21]. Therefore, we adopt GaMD for RNA-ligand binding simulations in this study.

The RNA composer (available at https://rnacomposer.cs.put.poznan.pl) [67] was applied to build the simulation structure of target hairpin RNA with a G 1 × 0 bulge (RNA5: 5′-AAGAUGGAGAGCGAAACACACUCGUCUAUCUU). 3-Dimensional (3D) structures of three coumarin derivatives (C30, C30-Me, and SMSM64) were prepared using ChemBio Tool. All small molecules were protonated with a + 1 charge at the nitrogen in the piperazine ring. Parameterization of the small molecules was performed using antechamber with atomic charges calculated using the AM1-BCC option [68]. Each small molecule was initially placed at a distance of >15 Å away from the nucleic acid surface. Each simulation system was then prepared using the *tleap* module in AMBER 22 [69]. The system charges were neutralized in 0.15 M NaCl [70] and 0.01 M $MgCl_2$ [71] solution. The AMBER OL3 force field [72, 73] was employed for RNA, the latest GAFF2 [74] for small molecules, and TIP3P [75] for water in the system. The AMBER OL3 force field was chosen for RNA because it has been optimized with specific parameters to model the intrinsic structural features of RNA, such as base pairing, stacking interactions, and high flexibility. These features are critical for accurately predicting how small molecule ligands interact with RNA. AMBER OL3 excels in capturing the delicate balance of interactions between RNA and ligands. The usage of AMBER OL3 and GAFF2 force fields has already been demonstrated to successfully capture ligand binding to single-stranded RNA [21] and is thus adopted in this study.

GaMD simulations were conducted using the *pmemd.cuda*, GPU-accelerating program in AMBER22 [69]. In all simulations, the hydrogen-heavy atom bonds were constrained using the SHAKE [76] algorithm, and the simulation time step was set to 2.0 fs. The particle mesh Ewald (PME) [77] method was employed to compute the long-range electrostatic interactions, and a cutoff value of 9.0 Å was applied to treat the non-

bonded atomic interactions. The temperature was controlled using the Langevin thermostat with a collision frequency of 1.0 ps^{-1}. Each simulation system was energy minimized using the steepest descent algorithm. This was followed by subjecting the system to heating from 0 K to 300 K for 200 ps. The system was further equilibrated using the constant number, volume, and temperature run for 800 ps at 300 K and constant number, pressure, and temperature ensemble at 300 K and 1 bar for 1 ns with 1 kcal/mol/$Å^2$ constraints on the heavy atoms of the RNA and ligand. A short conventional MD simulation for 10 ns without any constraint was performed to collect initial potential statistics, including the maximum, minimum, average, and standard deviation (SD).

Dual-boost GaMD simulations were performed on binding of the three coumarin derivatives to target RNA. One boost potential was applied to the dihedral energetic term and the other to the total potential energetic term. The reference energy E for applying the boost potentials was set to the lower bound, i.e. $E = V_{max}$. The average and SD of the system potential energies were calculated every 300 000 simulation steps (600 ps). The upper limit of the boost potential SD, σ0, was set to 6.0 kcal/mol for both the dihedral and the total potential energetic terms. The GaMD equilibration was carried out for 63 ns after adding the boost potential. Finally, three independent 1500 ns GaMD production simulations were conducted for each system with randomized initial atomic velocities.

Simulation analysis

VMD [78] and CPPTRAJ [79] were used to perform trajectory analysis. The hierarchical agglomerative clustering algorithm [80] available in CPPTRAJ was used to conduct clustering of the ligand snapshots to identify low-energy conformations. The frames were sieved at a stride of 500 for clustering. The top 10 structural clusters were analysed to identify the representative conformations of each system. The center-of-mass distances were calculated between heavy atoms of the BC ring of the ligand and the RNA nucleotide G24 and between heavy atoms of the DE ring of the ligand and RNA nucleotide C12 using CPPTRAJ [79]. Furthermore, these distances were used as reaction coordinates for calculating a 2D potential of mean force free energy profile by reweighting all three GaMD simulations combined. The PyReweighting toolkit was used for reweighting the GaMD simulations [81]. A bin size of 1 Å was used for the ligand distances and the cutoff was set to 500 frames in a bin or cluster for reweighting.

NMR experiments

An 84.1 nmol unlabelled RNA5 sample (Sigma–Aldrich) was dissolved in 135 µl of potassium buffer (25 mM potassium phosphate buffer, 50 mM potassium chloride, pH 6.2, 10% D_2O) to prepare an ∼600 µM NMR sample. A similar sample condition was used in a previous report [82], where NMR assignment was determined for a ^{15}N-labeled RNA containing the segment of RNA5.

All NMR spectra were acquired using a Bruker 800 MHz Ascend spectrometer equipped with a TCI cryoprobe at 298 K (Supplementary Fig. S1). RNA proton peak assignment was performed by comparing the measured ^{1}H chemical shifts with literature values (Supplementary Table S2) [82]. Proton peaks were assigned if the chemical shift difference was <0.15 ppm. All NH protons of U and G residues that showed peaks in the 10−14 ppm region were assigned accordingly, except for the one from G24, where the literature assignment was missing. Two proton peaks observed in this chemical shift region that were not previously assigned should correspond to the NH of G24 as well as that of G13. G13 is a part of the linker that differs from the RNA sequence used in the literature. The proton peaks of G7-NH and U25-NH did not appear until the addition of a 70 µM ligand. All the unambiguously assigned resonances are summarized in Supplementary Table S2.

The ligand stock solution prepared for the titration experiment contained 10 mM C30 in DMSO-d6. 0.47, 0.47, 0.94, 1.88, 0.94, 1.88, and 2.82 µl of the stock solution were titrated into the RNA NMR sample to achieve final ligand concentrations of 35, 70, 140, 280, 350, 490, and 700 µM, respectively. Proton transverse relaxation rate R_2 was measured from SOFAIR (band-selective optimized flip-angle internally encoded relaxation) (Supplementary Fig. S2) [83]. The band-selective excitation pulse p39 was centred at 12.1 ppm with a bandwidth of 5.2 ppm for NH region, and was centred at 7.6 ppm with a bandwidth of 3 ppm for NH_2/aromatic region. Transverse relaxation was encoded through the incrementation of a delay t flanking the refocusing pulse p40. The delay time was set to between 0 and 0.4 s with a total of 12 increments (0, 0.002, 0.005, 0.010, 0.015, 0.020, 0.025, 0.030, 0.050, 0.1, 0.2, and 0.4 s). The duration of each experiment is ∼18.5 min. Data were processed and analysed using MestReNova. R_2 of each resonance was determined through area integration and fitting the integrals to the following equation: $I = I_0 e^{-tR_2} + B$, where t is the delay time, R_2 is the transverse relaxation rate, and B is a constant to account for any baseline differences between experiments.

Results

Coumarin derivatives selectively bind to RNA G (1 × 0) bulge

The prototype coumarin derivative C2NH, which binds to RNA single G bulges (denoted as G 1 × 0 bulge) at a moderate binding affinity, contains five heterocyclic rings: piperazine (A), coumarin (BC), and a [5,6]-fused ring (DE) (Fig. 1). We previously reported C2NH as an active splicing modulator that can bind to a GA-rich loop within the SMN2 gene [21]. We modified the E ring to remove the splicing modulatory activity and repurposed the scaffold to other RNA targets, resulting in a potent G 1 × 0 bulge binder that strongly associates with a structural motif in the RNA genome of SARS-CoV-2 [21, 47]. To further probe the mechanism of the coumarin derivative in RNA binding interaction, we synthesized a collection of 69 analogues of C2NH (Fig. 1). Each compound in this collection comprises at least one ring distinct from the parent compound. For instance, in Ring A, the piperazine was replaced by cyclic amines of varying sizes. In Ring BC, the coumarin was substituted by other heterocycles with various substituents. Similarly, the [5,6]-fused Ring DE was replaced by [6,5]- or [6,6]-fused rings (Fig. 1).

All compounds in this collection are fluorescent with an excitation/emission wavelength at ∼400/480 nm, which allowed us to use FP assay to rapidly determine their binding affinity to the bulge G RNA. Using a bulged G RNA segment from SARS-CoV-2 SL5 RNA as a model, we extensively profiled this 69-compound library against all four 1 × 0 RNA bulge variants (RNA1–4) for binding affinities

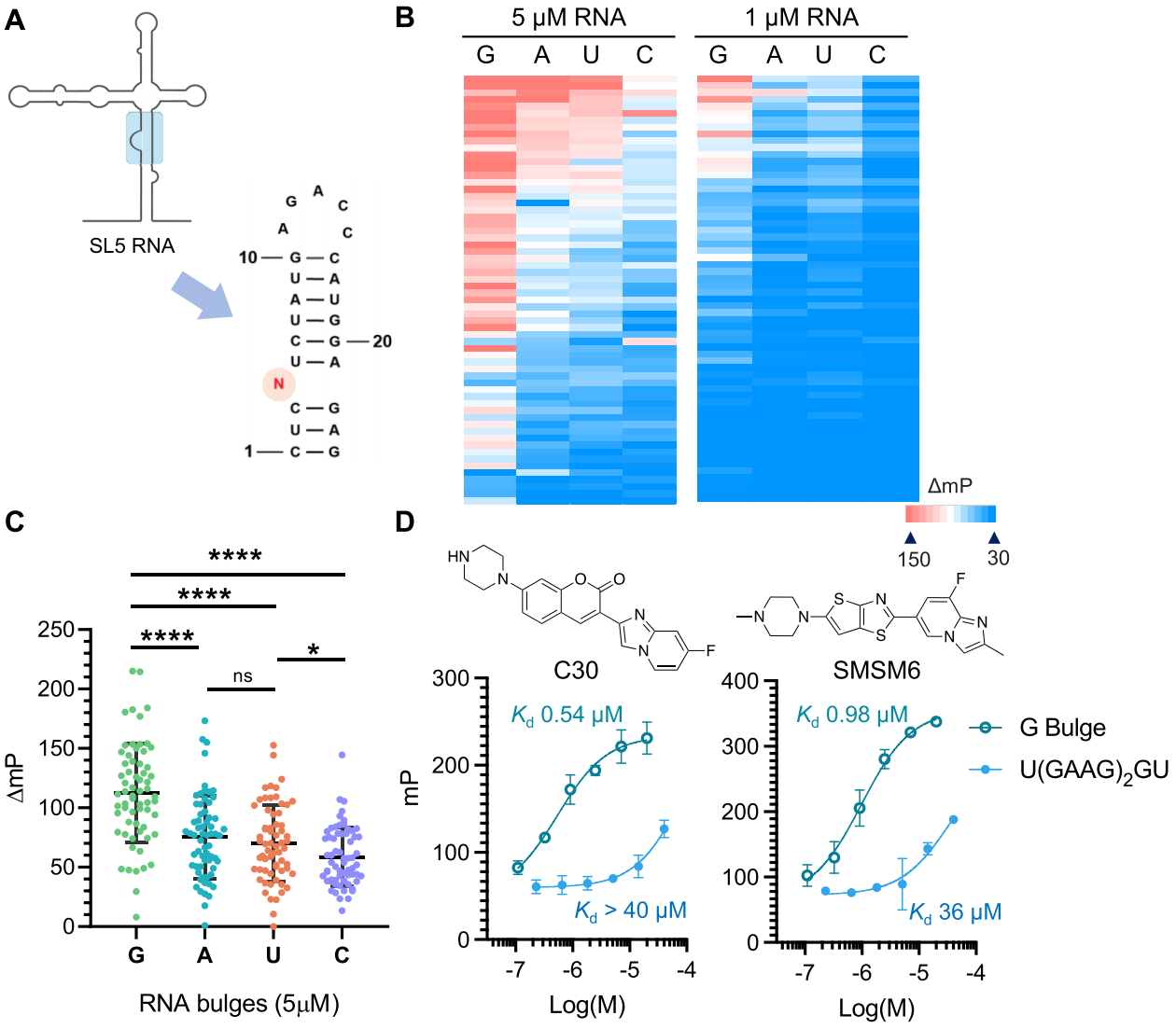


Figure 2. Coumarin derivatives can selectively bind to RNA G 1 × 0 bulges: (**A**) RNA structures of the 1 × 0 RNA bulges used for *in vitro* binding profiling using the FP assay. *N* = *G*, *A*, *U*, or *C* (RNA1-4). (**B**) Heatmap profile of the ΔmP = ($FP^{RNA\text{-}ligand}$ – FP^{ligand}) × 1000 for RNA binders in the presence of [RNA] = 5 or 1 μM (red = high polarization, blue = low polarization). (**C**) ΔmP of RNA-ligand complex for RNA ligands at 5 μM. Each data point represents a measurement of a ligand in the 69-compound collection. **** indicates $P < .0001$. (**D**) Dose-response curves for compounds (SMSM6, C30) selectively binding to the bulged G RNA (RNA1) compared to an 11-nucleotide GA-rich sequence that would form a double loop-like RNA structure measured by the FP assay.

(Fig. 2A, Supplementary Figs S3 and S4, and Supplementary Table S3). Almost all binding molecules showed superior selectivity for the G bulge compared to other RNA bulges (bulged A, U, and C), as judged by the polarization change (ΔmP) at two concentrations (1 and 5 μM) (Fig. 2B).

Statistical analysis of all 69 compounds revealed that the binding affinity for different RNA bulges followed the following trend: G >> A ≈ U > C (Fig. 2C). Since C2NH can also bind to GA-rich RNA loops [21] via an induced-fit mechanism, resulting in the formation of a double loop-like structure, we also tested the binding of coumarin derivatives to a flexible GA-rich RNA (5′-U(GAAG)$_2$GU) (Supplementary Fig. S5 and Supplementary Table S4). Interestingly, certain compounds, such as C30 and SMSM6, demonstrated >35-fold selectivity towards bulged G over GA-rich RNA (Fig. 2D), while a few compounds bind to both RNAs with comparable binding affinities (Supplementary Fig. S6). This suggests that a portion of coumarin derivatives may employ a distinct binding mechanism to target the RNA 1 × 0 G bulge selectively. The binding affinity of selected G bulge binders to RNA1 was validated using the microscale thermophoresis assay, which provided consistent results. (Supplementary Fig. S7).

GaMD simulations captured spontaneous binding of coumarin derivatives to RNA G (1 × 0) bulges in the minor groove

To explore the binding of specific RNA G bulge ligands, we performed all-atom simulations using the GaMD method [60] on binding of three coumarin derivatives to the model hairpin RNA5 with a G bulge (Supplementary Fig. S8A; see simulation details in "Materials and methods" section).

We found that C30 bound spontaneously to the G bulge and minor groove of RNA5 during the GaMD equilibration

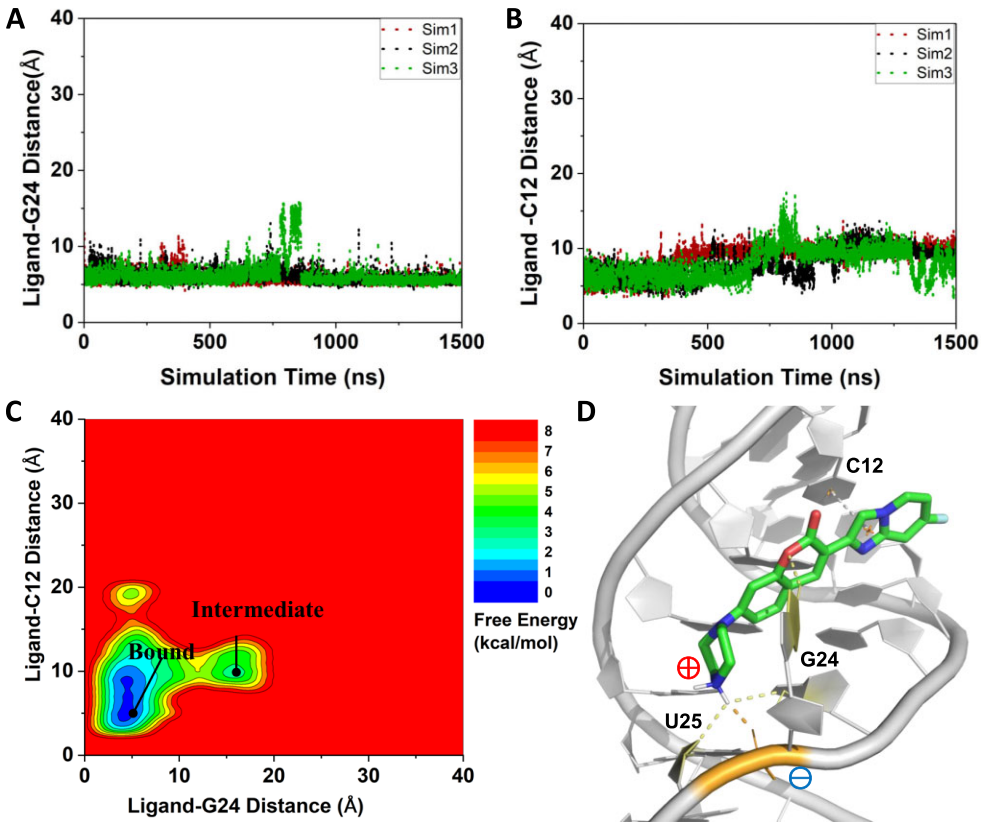


Figure 3. GaMD simulations captured stable binding of coumarin derivative C30 to RNA5: (**A**) Time courses of the center-of-mass distance between heavy atoms of the BC ring of ligand C30 and the RNA bulge G24 were calculated from three 1500 ns GaMD production simulations. (**B**) The center-of-mass distance between heavy atoms of the DE ring of ligand C30 and RNA nucleotide C12 is plotted as a function of simulation time. (**C**) 2D free energy profile calculated with all three GaMD simulations combined, showing two distinct low-energy states, namely the 'Bound' and 'Intermediate'. (**D**) Representative conformation of RNA–C30 complex in the bound state (grey dash line = π–π stacking, yellow dash line = hydrogen bonding, orange dash line = ionic interaction). The 'Intermediate' conformation is shown in Supplementary Fig. S9.

simulation (Supplementary Fig. S8B). It then maintained the bound conformation in three independent 1500 ns GaMD production simulations (Fig. 3). C30 remained bound to the RNA5 minor groove throughout nearly the entire three 1500 ns GaMD production simulations, despite small fluctuations around ~800 ns in Sim3. Upon binding to the minor groove of RNA5, a short distance was observed between the coumarin core (BC ring) of the ligand and the bulged G24 at 3.5–5 Å (Fig. 3A). Moreover, a π–π stacking interaction was observed between C12 and the fused D/E ring of the C30 ligand in simulations (Fig. 3B). We used these distances as reaction coordinates to further calculate a 2D free energy profile of C30 binding to RNA5, which showed two low-energy states, designated as 'Bound' (more stable) and 'Intermediate' states (Fig. 3C; bound state structure was deposited in Model Archive Project ma-q6hl4). To experimentally probe the minor groove binding mechanism that we observed in the GaMD simulations, we conducted additional FP binding assays using C30 and various DNA versions of the RNA G bulge sequences (same sequences as RNA1 and RNA5). Our results show that the deoxyribose modification gives rise to a >13-fold decrease in binding affinity (Supplementary Fig. S10). This result differed from what we observed with GA-rich loop binders, where the DNA aptamers bind to the ligands with a higher binding affinity than the RNA aptamers with the same sequences [21]. Given that double-stranded (ds) DNA typically adopts a different groove geometry than dsRNA [84], the reduced binding of C30 to DNA supports a groove-binding mechanism and highlights the RNA selectivity in C30 ligand recognition.

In the 'Bound' state, C30 formed three primary interactions within the minor groove of RNA5 (Fig. 3D): (i) The bulged G (G24) contributed to a hydrogen bond via its N1 position to the coumarin lactone moiety in C30. (ii) A phosphate group in the RNA backbone was involved in an ionic interaction with the protonated NH_2^+ group in the piperazine ring of C30. (iii) Nucleotide C12 formed π–π stacking interactions with the ligand C30 in the RNA minor groove. In the transient 'Intermediate' state, C30 was located at a much larger distance from the G24 nucleotide and did not insert into the RNA minor groove (Supplementary Fig. S9).

We further performed GaMD simulations on two inactive analogues of C30, namely C30-Me (Fig. 4A) and SMSM64 (Supplementary Fig. S11). C30-Me merely has an additional methyl group on the C ring compared to C30, which would break the planarity of Rings BC and DE within the compound (see discussions below). On the other hand, SMSM64 has an N-pyridinyl quinolone replacing the coumarin core, whose bulkiness might block the polar interaction with the RNA G bulge. In experiments, both compounds exhibited >100-fold reduced binding affinities towards RNA5, with SMSM64 displaying a dissociation constant (K_d) of >50 μM, in compar-

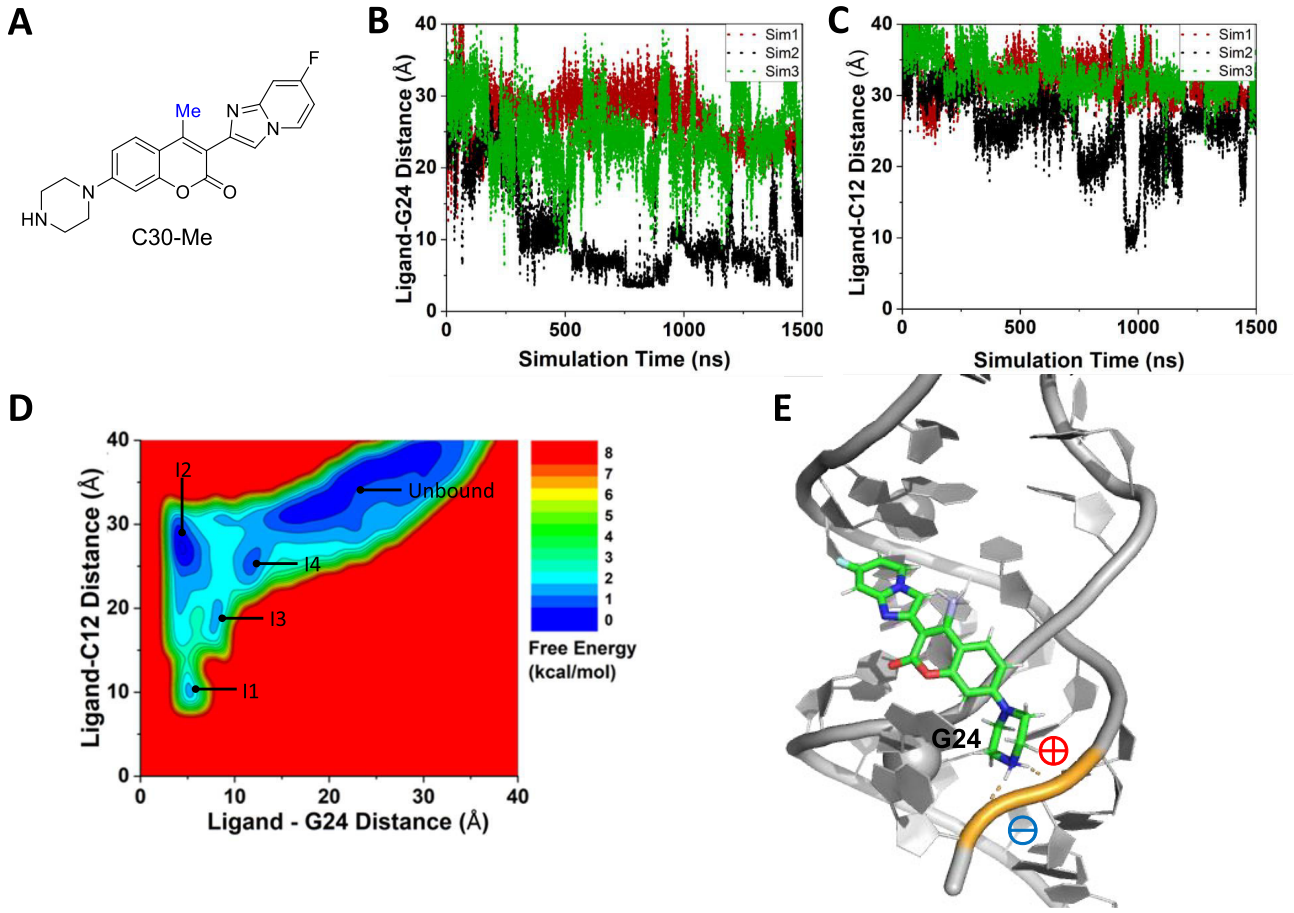


Figure 4. GaMD simulations captured transient binding of ligand C30-Me to RNA5: (**A**) Chemical structure of C30-Me. (**B**) Time courses of the center-of-mass distance between heavy atoms of the coumarin core (BC ring) in the ligand and the RNA bulged G24 calculated from three independent 1500 ns GaMD simulations. (**C**) The center-of-mass distance between the heavy atoms of the DE ring of C30-Me and RNA nucleotide C12 plotted as a function of simulation time. (**D**) 2D free energy profile calculated with all three GaMD simulations combined, showing five low-energy states, namely the 'I1', 'I2', 'I3', 'I4', and 'Unbound'. (**E**) Representative conformation of C30-Me–RNA5 complex in the I1 state (orange dash line = ionic interaction).

ison to C30, which has a K_d of 0.27 ± 0.01 μM to RNA5 (Supplementary Fig. S12). Similar binding affinities were observed for these compounds when binding to RNA1 (see Supplementary Table S3).

In all three 1500 ns GaMD simulations for C30-Me, the ligand seldom reached the target site in the minor groove of RNA5 (Fig. 4). In the situation where C30-Me transiently interacted with G24 nucleotide ('Sim2' in Fig. 4B), the ligand remained out of the RNA minor groove with a distance >10 Å from nucleotide C12 (Fig. 4C). Altogether, four transient binding states were identified from the free energy profile, designated as 'Intermediate' states I1–I4, as well as the 'Unbound' state, where the ligand dissociated from the RNA (Fig. 4D and Supplementary Fig. S13). The presence of multiple intermediate states suggested that the ligand explored various binding positions but was unable to achieve stable insertion into the minor groove. These intermediate conformations all maintained ionic interactions between the positively charged piperazine ring on the ligand and at least one phosphate group on the RNA backbone (Fig. 4E). For SMSM64, the ligand remained mostly >∼15 Å away from key nucleotides G24 and C12 throughout the 1500 ns GaMD simulations (Supplementary Fig. S11B and C). The resulting free energy profile showed only an 'Unbound' state (Supplementary Fig. S11D and E).

During the GaMD simulations, we observed high fluctuations of G24 in RNA5 in the absence of stable ligand binding as for the C30-Me and SMSM64 coumarin derivatives. In comparison, stronger binding of C30 to RNA5 led to significantly reduced root-mean-square fluctuations of the RNA nucleotides, especially G24 (Supplementary Fig. S14). Experimentally, we screened ∼20 crystal structures of RNA1 obtained using fragment antigen-binding region (Fab) chaperon-assisted crystallography (for a representative structure, see Protein Data Bank with accession code 9DN4) and observed dynamic conformations of the G bulge nucleotide [85], whereas other nucleotides remained relatively static (Supplementary Fig. S15 and Supplementary Table S5). This result is also consistent with our chemical probing results in SARS-CoV-2 RNA, where a high SHAPE (selective 2′ hydroxyl acylation analysed by primer extension) signal was observed with high concentrations of acylation agents (e.g. 10 mM FAI-N3) [47].

NMR validation of the minor groove binding mode

Next, we used NMR experiments to validate the predicted binding mode between coumarin analogues and the RNA with a G 1 × 0 bulge. First, we assigned imino protons and some other protons on the nucleobases in ^{1}H NMR using a reported

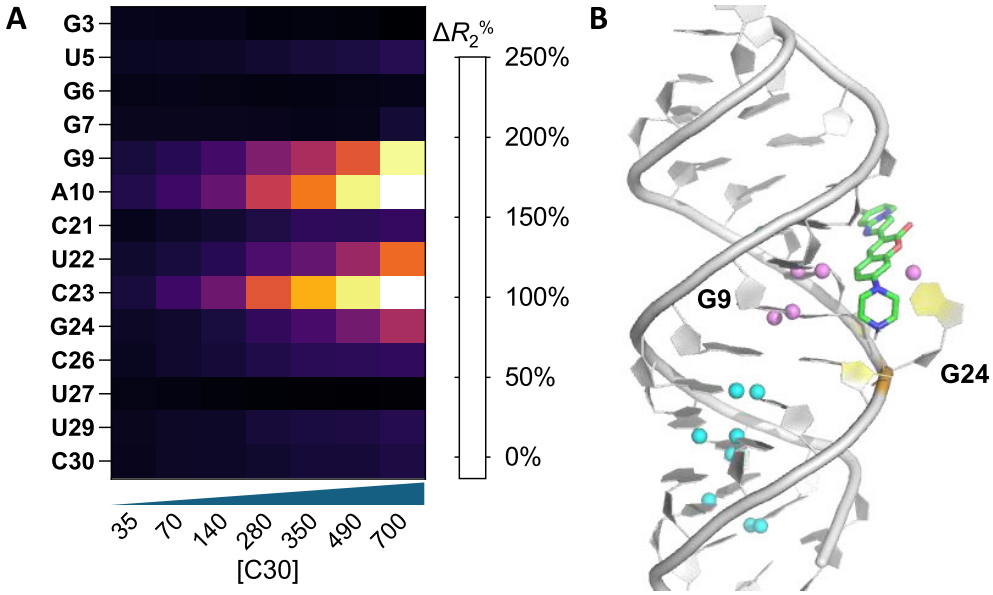


Figure 5. NMR relaxometry validation of the minor groove binding mode: (**A**) Normalized R_2 relaxation rate percentage changes obtained from each assigned proton of RNA nucleotide in the absence and presence of the ligand C30. The coloured columns in the bar plot represented R_2 measured at different ligand concentrations. Normalized R_2 values were calculated using the equation: $\Delta R_2^{\%} = \frac{R_2^{ligand+} - R_2^{ligand-}}{R_2^{ligand-}} \times 100\%$. (**B**) Overlay of the simulation-predicted 'Bound' state of C30-RNA5 and all identifiable protons in ^{1}H SOFAIR (pink = increased $\Delta R_2{}^{\%}$; blue = unchanged or decreased $\Delta R_2{}^{\%}$).

assignment that contains the segment of RNA5 [82]. The assigned peaks were distinguishable ones within 0.15 ppm from the reported ^{1}H NMR chemical shifts (Supplementary Table S2). Next, we applied a recently published NMR method, ^{1}H SOFAIR (band-Selective Optimized Flip-Angle Internally-encoded Relaxation) [86], to quantify R_2 relaxation rate of the receptor signals in order to characterize RNA-ligand interactions. R_2 relaxation reflects on dynamics and motion changes of molecules, which is sensitive to weak binding (K_d ~μM), and has been widely utilized as an NMR approach for identifying the binding sites of biomolecules [87, 88]. Here, the SOFAIR pulse sequence [86] was utilized to facilitate signal acquisition with high sensitivity of an RNA sample at mM concentration. Notably, SOFAIR was specifically designed to speed up data acquisition, and in this instance, led to a reduction in acquisition time from several hours, characteristic of conventional proton R_2 measurements using Carr-Purcell-Meiboom-Gill type of methods [89, 90] to ~20 min.

The R_2 relaxation rates were obtained from RNA nucleobases during the titration of the ligand C30 (Supplementary Table S6). As shown in Fig. 5A, the titration of C30 induced an overall R_2 change, indicating binding between C30 and the bulged G RNA. The most pronounced increase in R_2 was observed in G9, A10, U22, C23, and G24, implying the direct involvement of these nucleotides in binding. In contrast, the relaxation rates observed from G3 to G7 and C26 to C30 exhibited much smaller increases or even negative changes upon ligand addition, suggesting that these regions of the RNA are not directly involved in binding. These findings from NMR experiments regarding the bound and unbound RNA nucleotides are consistent with those obtained from the GaMD simulations (Fig. 5B). Interestingly, the putative binding location is selective to one side of the G bulge (U22–G24), implying sequence selectivity in the minor groove. It is worth noting that only a few small molecular ligands have been reported as minor groove binders (e.g. PDB 1QD3) [91], likely because the minor groove is wide and shallow in A-form dsRNAs. In summary, our NMR data strongly supported a minor groove binding mechanism for C30, as the ligand is unlikely to bind to the major groove of the RNA given the observed R_2 relaxation changes. This finding further highlights the critical role of the bulged G in ligand interactions within this unusual minor groove binding mechanism.

Molecular features on the ligands for RNA binding

We also determined how the molecular characteristics of these ligands contribute to their efficacy in RNA G bulge binding. Our approach involved an SAR analysis using regularized regression based on *in vitro* binding affinity data. Given the similarities in shape, size, and molecular scaffolds of our 69-compound library, we expected the SAR analysis to offer detailed molecular insight into the specific structural and electronic properties responsible for their potency.

We used MOE software to individually predict the most likely protonation state based on the 2D structures of each molecule. The majority of molecules were found to be monoprotonated at the aliphatic cyclic amine (A ring), with a few exceptions that carried two positive charges (Supplementary data). We then optimized the 3D structure of each molecule using *ab initio* DFT calculation with B3LYP 6–31G(d) basis set (Supplementary data). Using these 3D structures as input, we generated 443 molecular descriptors using MOE software (Supplementary data). To account for the planarity of the coumarin derivatives, we introduced a new dihedral descriptor between the aromatic rings BC and DE based on the most stable conformer predicted by DFT calculations. We then used the least absolute shrinkage and selection operator (Lasso) regression technique to identify the important elec-

Table 1. Molecular descriptors selected by lasso regression for RNA binding[a]

Molecular descriptor	Description[a]	Class/impact[b]
FCharge	Total formal charge of the molecule.	Charge/+
a_base	Number of basic atoms.	Charge/+
PEOE_VSA_FPOS	Fractional positive VDWSA.	Charge/+
PEOE_VSA_FNEG[c]	Fractional negative VDWSA.	Charge/−
PEOE_VSA_NEG	Total negative VDWSA.	Charge/−
NPR1	PMI[d] ratio: PMI1/PMI3	Shape/−
std_dim1	The square root of the largest eigenvalue of the covariance matrix of atomic coordinates.	Shape/+
ω(BC–DE)	Dihedral between BC and DE rings in the optimal structure.	Shape/−

[a]VDWSA = van der Waals surface area.
[b]+ and − signs indicate variables positively or negatively correlated to the binding affinity. [c]PEOE_VSA_FNEG = –PEOE_VSA_FPOS. [d]PMI = Normalized principal moments of inertia.

tronic and structural features among these molecular descriptors using a modified analytical pipeline [57, 61]. Lasso regression is a linear regression approach used for feature selection, which effectively eliminates unimportant variables. This process resulted in 16 molecular descriptors that significantly contributed to the binding affinity (Supplementary Table S1), of which eight molecular features are related to the charge and shape of the RNA ligands (Table 1).

The five charge-related molecular descriptors were based on the total formal charge of the molecule (FCharge), the number of basic atoms that can be potentially protonated in physiological pH (A_base), the fractional positive (PEOE_VSA_FPOS) and negative (PEOE_VSA_FNEG) charges per unit area, and the total negative charge per unit area (PEOE_VSA_NEG). Since RNA is densely negatively charged, it is reasonable that positive charges would significantly contribute to RNA binding due to charge attraction. In the GaMD simulations with C30, intermolecular ionic interactions between the positive charge on the piperazine ring of C30 and phosphate groups on the RNA backbone were critical in maintaining the stability of the RNA-ligand complex. When we acetylated the piperazine ring of C30 at the N4 position (C30-Ac) to prevent protonation, the binding affinity decreased by a factor of >5, highlighting the importance of electrostatic interaction between the ligand and RNA (Supplementary Fig. S16). We further hypothesized that the ligand used the positive charge on Ring A to explore suitable binding pockets at the early stage of the binding process. This hypothesis was supported by GaMD simulations, in which we observed all identifiable transient binding states ('Intermediate' states) of C30–RNA5 and C30-Me–RNA5 complexes retained an ionic interaction with RNA backbone phosphates (Supplementary Fig. S9 and S13; Supplementary Movie S1).

We also verified the impact of local positive charges on coumarin derivatives on *in vitro* binding by selecting four compounds, C29, C36, C34, and C34b, which only differ in the structures of the E ring. These compounds have two potential protonation sites: a piperazine A ring and an imidazole D ring. The second protonation site on the D ring can be partially stabilized by the coumarin moiety by forming an internal hydrogen bond. We speculated that the propensity of imidazolium formation significantly depends on the substituents on the E ring (Fig. 6A). For example, substituting the E ring with a trifluoromethyl group makes the molecule less amenable to protonation due to the electron-withdrawing effects. In contrast, the presence of an electron-donating methoxy group in compound C34b enhances the favourability of imidazolium formation. When the methoxy group is positioned at the 4′ location (C34), the existence of a resonance structure further contributes to stabilizing the positive charge (Fig. 6A). We verified the protonation energy of the four compounds relative to C29 using DFT calculations and compared it with the *in vitro* binding data (Fig. 6B and C). The dissociation constants for these four compounds exhibit a consistent trend concerning protonation energy, providing compelling evidence that local positive charges on ligands significantly contribute to RNA binding.

The 3D shape descriptors also strongly correlate with the binding affinity (Table 1). For example, NPR1 and NPR2 are numeric shape descriptors with values between 0 and 1 that characterize the general 3D geometries of molecules [92]. All compounds in our compound collection exhibit a small NPR1 value (<0.2) and a large NPR2 value (>0.85), indicating rod-like molecular structures. This observation is consistent with a prior cheminformatic analysis of diverse RNA-binding molecules [54]. In addition, the positive contribution of the shape descriptor std_dim1 indicates that a longer molecule makes the ligand more favorable for binding, which is consistent with our expectations for groove binders.

Finally, we observed a positive correlation between planarity and binding affinity, as indicated by the inverse relationship between the dihedral angle of Rings BC and DE [ω(BC–DE)] and the natural logarithm of the binding constant (LnK_d). In C30, the dihedral angle between the BC–DE ring is ∼0, making it a planar molecule (Supplementary Fig. S17), which facilitates groove binding. However, adding a methyl group on ring C of C30 (C30-Me) causes steric hindrance between the methyl group and the lone pair electron of the imidazole nitrogen, disrupting the planarity of the molecule, rending C30-Me a poor binder (Fig. 7A and Supplementary Fig. S17). We also tested the role of this methyl group on ring B (C30-Me^{RingB}), where the methyl group no longer sterically clashes with the imidazole ring. As expected, C30-Me^{RingB} is planar in its most favourable conformation (Supplementary Fig. S17), and the binding affinity was comparable to that of C30 (Fig. 7A). Planarity might also contribute to the high binding affinity of C34 ($K_d = 0.10 \pm 0.01$ μM to RNA1). In the second protonation site of C34, the imidazole ring can form an internal hydrogen bonding with the coumarin lactone, further stabilizing the planar conformation (Fig. 7B and Supplementary Fig. S17).

Discussion

In this study, we have reported a new group of coumarin derivatives that exhibit selective binding to bulge G RNA. Using all-atom GaMD simulations, NMR, and SAR studies, we have identified critical interactions that permit minor groove binding as well as crucial molecular properties of the ligands that significantly contribute to their binding affinity to bulged G RNAs. The ligand-RNA minor groove binding interface was validated by ^{1}H SOFAIR NMR experiments that can rapidly characterize the ligand binding behaviour on RNAs. Our research establishes a new example for understanding RNA-small molecule interactions with nanomolar binding affinity.

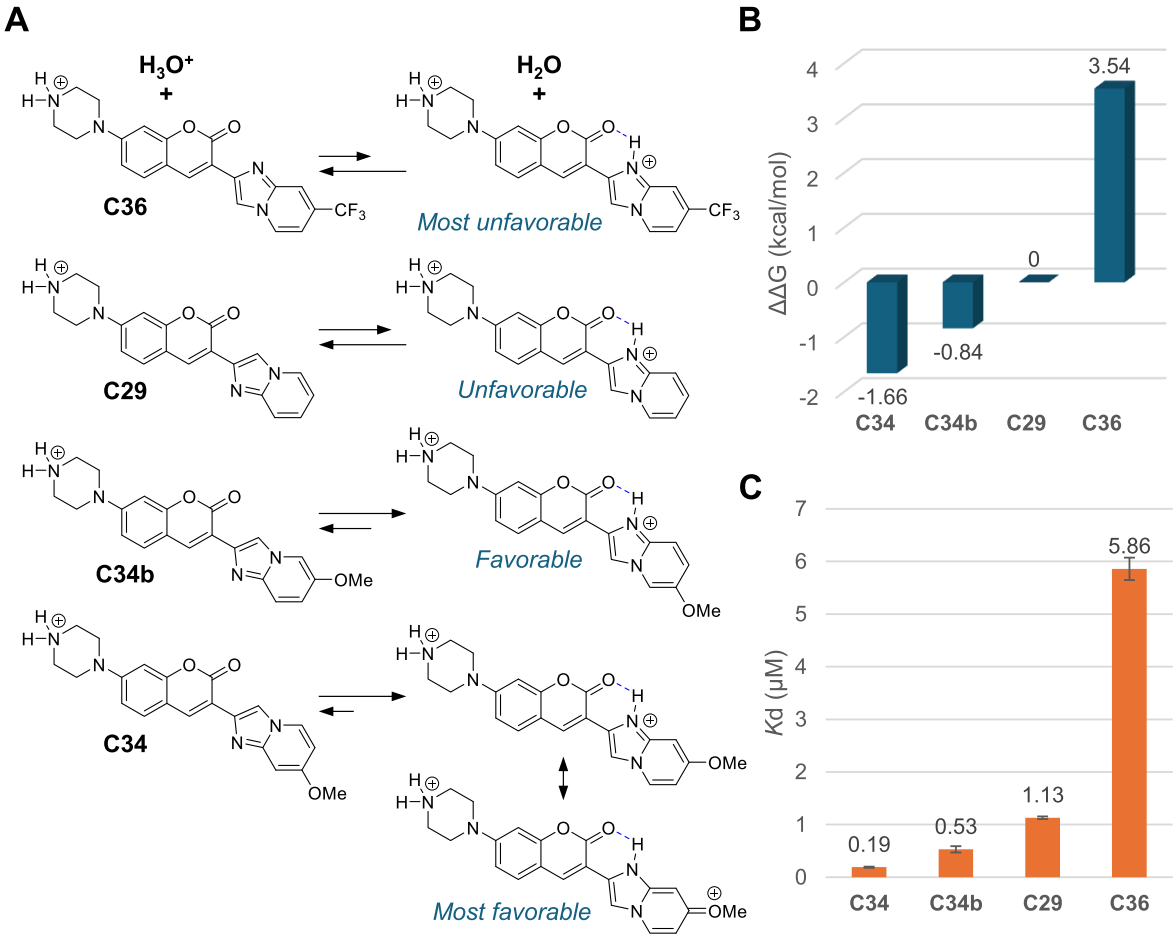


Figure 6. Protonation state of the ligands contributes to the RNA binding. (**A**) Equilibria for the protonation reactions of four coumarin derivatives. (**B**) Protonation energy (relative to C29) was calculated using DFT with B3LYP 6–31G(d) basis set. (**C**) Observed binding affinity of the four compounds.

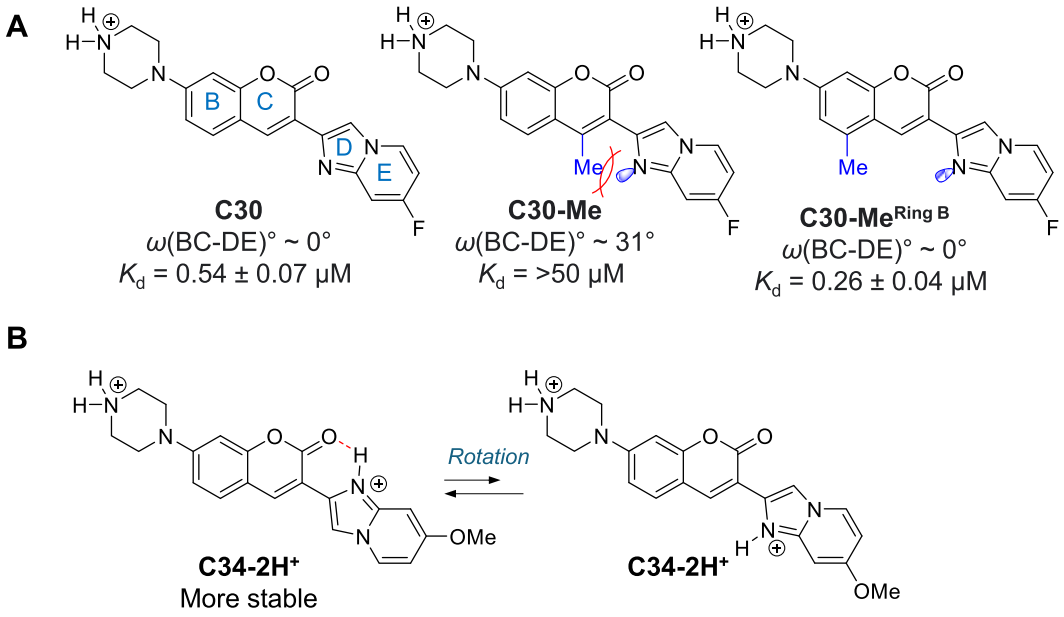


Figure 7. Planarity of the ligands contributes to the RNA binding. (**A**) The dihedral angle of Rings BC and DE [ω(BC–DE)] of the ligands in their ground states demonstrates a positive correlation with the dissociation constant (K_d). Methyl group on the Ring C can introduce steric clashes between Rings BC and DE, reducing their co-planarity. (**B**) Di-protonated C34 maintains planarity in both conformations due to the flipping of the DE ring. The left conformation is stabilized by internal hydrogen bonding at the imidazolium ring.

Acknowledgements

The authors acknowledge support from the MRL Postdoctoral Research Program and thank Xingjian Xu (Merck & Co., Inc.) for discussions regarding NMR measurements. The authors also thank Amoldeep S. Kainth and Wenzhe Chen (University of Chicago) for discussions during manuscript revision. This research used the NYX beamline 19-ID, supported by the New York Structural Biology Center, at the National Synchrotron Light Source II, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE-SC0012704.

Author contributions: Shalakha Hegde (Investigation, Visualization, Formal analysis, Writing—original draft, Writing—review & editing), Sana Akhter (Investigation, Data curation, Visualization, Software, Writing—review & editing), Zhichao Tang (Investigation, Visualization, Writing—review & editing), Chang Qi (Investigation, Data curation), Chenguang Yu (Software, Data curation, Methodology), Anna Lewicka (Investigation, Visualization), Yu Liu (Investigation), Kushal Koirala (Investigation), Mikhail Reibarkh (Supervision, Resources), Kevin P. Battaile (Investigation), Anne Cooper (Investigation), Scott Lovell (Supervision, Data curation), Erik D. Holmstrom (Investigation), Xiao Wang (Supervision, Investigation, Methodology), Joseph A. Piccirilli (Supervision, Investigation), Qi Gao (Supervision, Project administration, Methodology, Investigation, Resources, Writing—review & editing), Yinglong Miao (Conceptualization, Funding acquisition, Project administration, Supervision, Resources, Data curation, Visualization, Writing—original draft, Writing—review & editing), Jingxin Wang (Conceptualization, Funding acquisition, Project administration, Supervision, Resources, Visualization, Writing—original draft, Writing—review & editing).

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

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 (to J.W.) and start-up project 27110 at the University of North Carolina–Chapel Hill (to Y.M.). This research used the NYX beamline 19-ID, supported by the New York Structural Biology Center, at the National Synchrotron Light Source II, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE-SC0012704. The NYX detector instrumentation was supported by grant S10OD030394 through the Office of the Director of the National Institutes of Health. The Center for Bio-Molecular Structure (CBMS) is primarily supported by the NIH-NIGMS through a Center Core P30 Grant (P30GM133893) and by the DOE Office of Biological and Environmental Research (KP1607011). NSLS2 is a U.S. DOE Office of Science User Facility operated under Contract No. DE-SC0012704. This publication also resulted from the data collected using the beamtime obtained through NECAT BAG proposal # 311950.

Data availability

The representative C30-RNA5 bound conformation generated by GaMD simulations is available in PDB format in the Model Archive repository (https://modelarchive.org) under project ma-q6hl4. The representative apo RNA crystal structure of RNA1 is available in PDB under accession number 9DN4. NMR data is available in BMRB repository (https://bmrb.io/): https://deposit.bmrb.io/entry/load/93501976-1149-4710-989d-f797df292649.

References

1. Butler AA, Webb WM, Lubin FD. Regulatory RNAs and control of epigenetic mechanisms: expectations for cognition and cognitive dysfunction. *Epigenomics* 2016;**8**:135–51. https://doi.org/10.2217/epi.15.79
2. Wilkinson E, Cui YH, He YY. Roles of RNA modifications in diverse cellular functions. *Front Cell Dev Biol* 2022;**10**:828683. https://doi.org/10.3389/fcell.2022.828683
3. Meyer SM, Williams CC, Akahori Y *et al.* Small molecule recognition of disease-relevant RNA structures. *Chem Soc Rev* 2020;**49**:7167–99. https://doi.org/10.1039/D0CS00560F
4. Ursu A, Childs-Disney JL, Andrews RJ *et al.* Design of small molecules targeting RNA structure from sequence. *Chem Soc Rev* 2020;**49**:7252–70. https://doi.org/10.1039/D0CS00455C
5. Childs-Disney JL, Yang X, Gibaut QMR *et al.* Targeting RNA structures with small molecules. *Nat Rev Drug Discov* 2022;**21**:736–62. https://doi.org/10.1038/s41573-022-00521-4
6. Donlic A, Hargrove AE. Targeting RNA in mammalian systems with small molecules. *Wiley Interdiscip Rev RNA* 2018;**9**:e1477. https://doi.org/10.1002/wrna.1477
7. Hargrove AE. Small molecule-RNA targeting: starting with the fundamentals. *Chem Commun* 2020;**56**:14744–56. https://doi.org/10.1039/D0CC06796B
8. Connelly CM, Moon MH, Schneekloth JS. The emerging role of RNA as a therapeutic target for small molecules. *Cell Chem Biol* 2016;**23**:1077–90. https://doi.org/10.1016/j.chembiol.2016.05.021
9. Warner KD, Hajdin CE, Weeks KM. Principles for targeting RNA with drug-like small molecules. *Nat Rev Drug Discov* 2018;**17**:547–58. https://doi.org/10.1038/nrd.2018.93
10. Wilson DN. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nat Rev Microbiol* 2014;**12**:35–48. https://doi.org/10.1038/nrmicro3155
11. Naryshkin NA, Weetall M, Dakka A *et al.* SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. *Science* 2014;**345**:688–93. https://doi.org/10.1126/science.1250127
12. Palacino J, Swalley SE, Song C *et al.* SMN2 splice modulators enhance U1-pre-mRNA association and rescue SMA mice. *Nat Chem Biol* 2015;**11**:511–7. https://doi.org/10.1038/nchembio.1837
13. Woll MG, Qi H, Turpoff A *et al.* Discovery and optimization of small molecule splicing modifiers of survival motor neuron 2 as a treatment for spinal muscular atrophy. *J Med Chem* 2016;**59**:6070–85. https://doi.org/10.1021/acs.jmedchem.6b00460
14. Ratni H, Karp GM, Weetall M *et al.* Specific correction of alternative survival motor neuron 2 splicing by small molecules: discovery of a potential novel medicine to treat spinal muscular atrophy. *J Med Chem* 2016;**59**:6086–100. https://doi.org/10.1021/acs.jmedchem.6b00459

15. Pinard E, Green L, Reutlinger M *et al.* Discovery of a novel class of survival motor neuron 2 splicing modifiers for the treatment of spinal muscular atrophy. *J Med Chem* 2017;**60**:4444–57. https://doi.org/10.1021/acs.jmedchem.7b00406
16. Ratni H, Ebeling M, Baird J *et al.* Discovery of risdiplam, a selective survival of motor neuron-2 (SMN2) gene splicing modifier for the treatment of spinal muscular atrophy (SMA). *J Med Chem* 2018;**61**:6501–17. https://doi.org/10.1021/acs.jmedchem.8b00741
17. Cheung AK, Hurley B, Kerrigan R *et al.* Discovery of small molecule splicing modulators of survival motor neuron-2 (SMN2) for the treatment of spinal muscular atrophy (SMA). *J Med Chem* 2018;**61**:11021–36. https://doi.org/10.1021/acs.jmedchem.8b01291
18. Krach F, Stemick J, Boerstler T *et al.* An alternative splicing modulator decreases mutant HTT and improves the molecular fingerprint in Huntington's disease patient neurons. *Nat Commun* 2022;**13**:6797. https://doi.org/10.1038/s41467-022-34419-x
19. Bhattacharyya A, Trotta CR, Narasimhan J *et al.* Small molecule splicing modifiers with systemic HTT-lowering activity. *Nat Commun* 2021;**12**:7299. https://doi.org/10.1038/s41467-021-27157-z
20. Keller CG, Shin Y, Monteys AM *et al.* An orally available, brain penetrant, small molecule lowers huntingtin levels by enhancing pseudoexon inclusion. *Nat Commun* 2022;**13**:1150. https://doi.org/10.1038/s41467-022-28653-6
21. Tang Z, Akhter S, Ramprasad A *et al.* Recognition of single-stranded nucleic acids by small-molecule splicing modulators. *Nucleic Acids Res* 2021;**49**:7870–83. https://doi.org/10.1093/nar/gkab602
22. Blount KF, Breaker RR. Riboswitches as antibacterial drug targets. *Nat Biotechnol* 2006;**24**:1558–64. https://doi.org/10.1038/nbt1268
23. Howe JA, Wang H, Fischmann TO *et al.* Selective small-molecule inhibition of an RNA structural element. *Nature* 2015;**526**:672–7. https://doi.org/10.1038/nature15542
24. Tran B, Pichling P, Tenney L *et al.* Parallel discovery strategies provide a basis for riboswitch ligand design. *Cell Chem Biol* 2020;**27**:1241–9. https://doi.org/10.1016/j.chembiol.2020.07.021
25. Fedorova O, Jagdmann GE, Adams RL *et al.* Small molecules that target group II introns are potent antifungal agents. *Nat Chem Biol* 2018;**14**:1073–8. https://doi.org/10.1038/s41589-018-0142-0
26. Gumireddy K, Young DD, Xiong X *et al.* Small-molecule inhibitors of microrna miR-21 function. *Angew Chem Int Ed* 2008;**47**:7482–4. https://doi.org/10.1002/anie.200801555
27. Suresh BM, Li W, Zhang P *et al.* A general fragment-based approach to identify and optimize bioactive ligands targeting RNA. *Proc Natl Acad Sci USA* 2020;**117**:33197–203. https://doi.org/10.1073/pnas.2012217117
28. Li Y, Disney MD. Precise small molecule degradation of a noncoding RNA identifies cellular binding sites and modulates an oncogenic phenotype. *ACS Chem Biol* 2018;**13**:3065–71. https://doi.org/10.1021/acschembio.8b00827
29. Tong Y, Lee Y, Liu X *et al.* Programming inactive RNA-binding small molecules into bioactive degraders. *Nature* 2023;**618**:169–79. https://doi.org/10.1038/s41586-023-06091-8
30. Lorenz DA, Kaur T, Kerk SA *et al.* Expansion of cat-ELCCA for the discovery of small molecule inhibitors of the pre-let-7-Lin28 RNA-protein interaction. *ACS Med Chem Lett* 2018;**9**:517–21. https://doi.org/10.1021/acsmedchemlett.8b00126
31. Borgelt L, Li F, Hommen P *et al.* Trisubstituted pyrrolinones as small-molecule inhibitors disrupting the protein-RNA interaction of LIN28 and Let-7. *ACS Med Chem Lett* 2021;**12**:893–8. https://doi.org/10.1021/acsmedchemlett.0c00546
32. Zhang P, Park H-J, Zhang J *et al.* Translation of the intrinsically disordered protein α-synuclein is inhibited by a small molecule targeting its structured mRNA. *Proc Natl Acad Sci USA* 2020;**117**:1457–67. https://doi.org/10.1073/pnas.1905057117
33. Angelbello AJ, Rzuczek SG, Mckee KK *et al.* Precise small-molecule cleavage of an r(CUG) repeat expansion in a myotonic dystrophy mouse model. *Proc Natl Acad Sci USA* 2019;**116**:7799–804. https://doi.org/10.1073/pnas.1901484116
34. Bush JA, Aikawa H, Fuerst R *et al.* Ribonuclease recruitment using a small molecule reduced c9ALS/FTD r(G4C2) repeat expansion *in vitro* and *in vivo* ALS models. *Sci Transl Med* 2021;**13**:eabd5991. https://doi.org/10.1126/scitranslmed.abd5991
35. Donlic A, Morgan BS, Xu JL *et al.* Discovery of small molecule ligands for MALAT1 by tuning an RNA-binding scaffold. *Angew Chem Int Ed* 2018;**57**:13242–7. https://doi.org/10.1002/anie.201808823
36. Donlic A, Zafferani M, Padroni G *et al.* Regulation of MALAT1 triple helix stability and *in vitro* degradation by diphenylfurans. *Nucleic Acids Res* 2020;**48**:7653–64. https://doi.org/10.1093/nar/gkaa585
37. Rakheja I, Ansari AH, Ray A *et al.* Small molecule quercetin binds MALAT1 triplex and modulates its cellular function. *Mol Ther Nucleic Acids* 2022;**30**:241–56. https://doi.org/10.1016/j.omtn.2022.09.016
38. Wang J. Genome-wide identification of stable RNA secondary structures across multiple organisms using chemical probing data: insights into short structural motifs and RNA-targeting therapeutics. *Biochemistry* 2025;**64**:1817–27. https://doi.org/10.1101/2024.10.08.617329
39. Seth PP, Miyaji A, Jefferson EA *et al.* SAR by MS: discovery of a new class of RNA-binding small molecules for the hepatitis C virus: internal ribosome entry site IIA subdomain. *J Med Chem* 2005;**48**:7099–102. https://doi.org/10.1021/jm050815o
40. Parsons J, Castaldi MP, Dutta S *et al.* Conformational inhibition of the hepatitis C virus internal ribosome entry site RNA. *Nat Chem Biol* 2009;**5**:823–5. https://doi.org/10.1038/nchembio.217
41. Wang W, Préville P, Morin N *et al.* Hepatitis C viral IRES inhibition by phenazine and phenazine-like molecules. *Bioorg Med Chem Lett* 2000;**10**:1151–4. https://doi.org/10.1016/S0960-894X(00)00217-1
42. Dibrov SM, Parsons J, Carnevali M *et al.* Hepatitis C virus translation inhibitors targeting the internal ribosomal entry site. *J Med Chem* 2014;**57**:1694–707. https://doi.org/10.1021/jm401312n
43. Sztuba-Solinska J, Shenoy SR, Gareiss P *et al.* Identification of biologically active, HIV TAR RNA-binding small molecules using small molecule microarrays. *J Am Chem Soc* 2014;**136**:8402–10. https://doi.org/10.1021/ja502754f
44. Abulwerdi FA, Shortridge MD, Sztuba-Solinska J *et al.* Development of small molecules with a noncanonical binding mode to HIV-1 trans activation response (TAR) RNA. *J Med Chem* 2016;**59**:11148–60. https://doi.org/10.1021/acs.jmedchem.6b01450
45. Davidson A, Leeper TC, Athanassiou Z *et al.* Simultaneous recognition of HIV-1 TAR RNA bulge and loop sequences by cyclic peptide mimics of Tat protein. *Proc Natl Acad Sci USA* 2009;**106**:11931–6. https://doi.org/10.1073/pnas.0900629106
46. Zhao J, Qiu J, Aryal S *et al.* The RNA architecture of the SARS-CoV-2 3′-untranslated region. *Viruses* 2020;**12**:1473. https://doi.org/10.3390/v12121473
47. Tang Z, Hegde S, Hao S *et al.* Chemical-guided SHAPE sequencing (cgSHAPE-seq) informs the binding site of RNA-degrading chimeras targeting SARS-CoV-2 5' untranslated region. *Nat Commun* 2025;**16**:483. https://doi.org/10.1038/s41467-024-55608-w
48. Haniff HS, Tong Y, Liu X *et al.* Targeting the SARS-CoV-2 RNA genome with small molecule binders and ribonuclease targeting chimera (RIBOTAC) degraders. *ACS Cent Sci* 2020;**6**:1713–21. https://doi.org/10.1021/acscentsci.0c00984
49. Sun Y, Abriola L, Niederer RO *et al.* Restriction of SARS-CoV-2 replication by targeting programmed -1 ribosomal frameshifting. *Proc Natl Acad Sci USA* 2021;**118**:e2023051118. https://doi.org/10.1073/pnas.2023051118

50. Zafferani M, Haddad C, Luo L *et al.* Amilorides inhibit SARS-CoV-2 replication *in vitro* by targeting RNA structures. *Sci Adv* 2021;**7**:eabl6096. https://doi.org/10.1126/sciadv.abl6096
51. Mikutis S, Rebelo M, Yankova E *et al.* Proximity-induced nucleic acid degrader (PINAD) approach to targeted RNA degradation using small molecules. *ACS Cent Sci* 2023;**9**:892–904. https://doi.org/10.1021/acscentsci.3c00015
52. Bell 2018_Pnas_Si_Spe. *Proc Natl Acad Sci USA* 2017;**120**:2017.
53. Velagapudi SP, Seedhouse SJ, French J *et al.* Defining the RNA internal loops preferred by benzimidazole derivatives via 2D combinatorial screening and computational analysis. *J Am Chem Soc* 2011;**133**:10111–8. https://doi.org/10.1021/ja200212b
54. Morgan BS, Forte JE, Culver RN *et al.* Discovery of key physicochemical, structural, and spatial properties of RNA-targeted bioactive ligands. *Angew Chem Int Ed* 2017;**56**:13498–502. https://doi.org/10.1002/anie.201707641
55. Eubanks CS, Forte JE, Kapral GJ *et al.* Small molecule-based pattern recognition to classify RNA structure. *J Am Chem Soc* 2017;**139**:409–16. https://doi.org/10.1021/jacs.6b11087
56. Allen TEH, Mcdonagh JL, Broncel M *et al.* Physicochemical principles driving small molecule binding to RNA. bioRxiv, https://doi.org/10.1101/2024.01.31.578268, 2 February 2024, preprint: not peer reviewed.
57. Yazdani K, Jordan D, Yang M *et al.* Machine learning informs RNA-binding chemical space. *Angew Chem Int Ed* 2023;**62**:e202211358. https://doi.org/10.1002/anie.202211358
58. Wang Y, Parmar S, Schneekloth JS *et al.* Interrogating RNA-small molecule interactions with structure probing and artificial intelligence-augmented molecular simulations. *ACS Cent Sci* 2022;**8**:741–8. https://doi.org/10.1021/acscentsci.2c00149
59. Akhter S, Tang Z, Wang J *et al.* Mechanism of ligand binding to theophylline RNA aptamer. *J Chem Inf Model* 2024;**64**:1017–29. https://doi.org/10.1021/acs.jcim.3c01454
60. Miao Y, Feher VA, McCammon JA. Gaussian accelerated molecular dynamics: unconstrained enhanced sampling and free energy calculation. *J Chem Theory Comput* 2015;**11**:3584–95. https://doi.org/10.1021/acs.jctc.5b00436
61. Cai Z, Zafferani M, Akande OM *et al.* Quantitative structure-activity relationship (QSAR) study predicts small-molecule binding to RNA structure. *J Med Chem* 2022;**65**:7262–77. https://doi.org/10.1021/acs.jmedchem.2c00254
62. Miao Y, McCammon JA. Graded activation and free energy landscapes of a muscarinic G-protein-coupled receptor. *Proc Natl Acad Sci USA* 2016;**113**:12162–7. https://doi.org/10.1073/pnas.1614538113
63. Palermo G, Miao Y, Walker RC *et al.* CRISPR-Cas9 conformational activation as elucidated from enhanced molecular simulations. *Proc Natl Acad Sci USA* 2017;**114**:7260–5. https://doi.org/10.1073/pnas.1707645114
64. Miao Y, McCammon JA. Gaussian accelerated molecular dynamics: theory, implementation, and applications. *Annu Rep Comput Chem* 2017;**13**:231–78.
65. Miao Y, Huang Y-MM, Walker RC *et al.* Ligand binding pathways and conformational transitions of the HIV protease. *Biochemistry* 2018;**57**:1533–41. https://doi.org/10.1021/acs.biochem.7b01248
66. Miao Y, McCammon JA. Mechanism of the G-protein mimetic nanobody binding to a muscarinic G-protein-coupled receptor. *Proc Natl Acad Sci USA* 2018;**115**:3036–41. https://doi.org/10.1073/pnas.1800756115
67. Popenda M, Szachniuk M, Antczak M *et al.* Automated 3D structure composition for large RNAs. *Nucleic Acids Res* 2012;**40**:e112. https://doi.org/10.1093/nar/gks339
68. Salomon-Ferrer R, Case DA, Walker RC. An overview of the Amber biomolecular simulation package. *WIREs Comput Mol Sci* 2013;**3**:198–210. https://doi.org/10.1002/wcms.1121
69. Case DA, Aktulga HM, Belfon K *et al.* AmberTools. *J Chem Inf Model* 2023;**63**:6183–91. https://doi.org/10.1021/acs.jcim.3c01153
70. Joung IS, Cheatham TE. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J Phys Chem B* 2008;**112**:9020–41. https://doi.org/10.1021/jp8001614
71. Li P, Song LF, Merz KM. Systematic parameterization of monovalent ions employing the nonbonded model. *J Chem Theory Comput* 2015;**11**:1645–57. https://doi.org/10.1021/ct500918t
72. Fadrná E, Špačková N, Sarzyńska J *et al.* Single stranded loops of quadruplex DNA as key benchmark for testing nucleic acids force fields. *J Chem Theory Comput* 2009;**5**:2514–30. https://doi.org/10.1021/ct900200k
73. Zgarbová M, Otyepka M, Sponer J *et al.* Refinement of the *Cornell et al.* nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J Chem Theory Comput* 2011;**7**:2886–902. https://doi.org/10.1021/ct200162x
74. He X, Man VH, Yang W *et al.* A fast and high-quality charge model for the next generation general AMBER force field. *J Chem Phys* 2020;**153**:114502. https://doi.org/10.1063/5.0019056
75. Jorgensen WL, Chandrasekhar J, Madura JD *et al.* Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 1983;**79**:926–35. https://doi.org/10.1063/1.445869
76. Ryckaert J-P, Ciccotti G, Berendsen HJ. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* 1977;**23**:327–41. https://doi.org/10.1016/0021-9991(77)90098-5
77. Essmann U, Perera L, Berkowitz ML *et al.* A smooth particle mesh Ewald method. *J Chem Phys* 1995;**103**:8577–93. https://doi.org/10.1063/1.470117
78. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graphics* 1996;**14**:33–8. https://doi.org/10.1016/0263-7855(96)00018-5
79. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput* 2013;**9**:3084–95. https://doi.org/10.1021/ct400341p
80. Louhichi S, Gzara M, Ben Abdallah H. A density based algorithm for discovering clusters with varied density. In: *2014 World Congress on Computer Applications and Information Systems (WCCAIS)*. 2014, Hammamet, Tunisia, 1–6. https://doi.org/10.1109/WCCAIS.2014.6916622
81. Miao Y, Sinko W, Pierce L *et al.* Improved reweighting of accelerated molecular dynamics simulations for free energy calculation. *J Chem Theory Comput* 2014;**10**:2677–89. https://doi.org/10.1021/ct500090q
82. Wacker A, Weigand JE, Akabayov SR *et al.* Secondary structure determination of conserved SARS-CoV-2 RNA elements by NMR spectroscopy. *Nucleic Acids Res* 2020;**48**:12415–35. https://doi.org/10.1093/nar/gkaa1013
83. Xu X, Poggetto GD, Mccoy M *et al.* Rapid characterization of structural and behavioral changes of therapeutic proteins by relaxation and diffusion 1 H-SOFAST NMR experiments. 2024;**96**:16322–9.
84. Vuković LV, Koh HR, Myong S *et al.* Substrate recognition and specificity of double-stranded RNA binding proteins. 2014;**53**:3457–66. https://doi.org/10.1021/bi500352s
85. Koirala D, Shelke SA, Dupont M *et al.* Affinity maturation of a portable Fab-RNA module for chaperone-assisted RNA crystallography. *Biochemistry* 2018;**46**:2624–35. https://doi.org/10.1093/nar/gkx1292
86. Xu X, Dal Poggetto G, McCoy M *et al.* Rapid characterization of structural and behavioral changes of therapeutic proteins by relaxation and diffusion 1H-SOFAST NMR experiments. *Anal Chem* 2024;**96**:16322–9. https://doi.org/10.1021/acs.analchem.4c03479
87. Vise PD, Baral B, Latos AJ *et al.* NMR chemical shift and relaxation measurements provide evidence for the coupled folding and binding of the p53 transactivation domain. *Nucleic Acids Res* 2005;**33**:2061–77. https://doi.org/10.1093/nar/gki336

88. Walinda E, Morimoto D, Sugase K. Resolving biomolecular motion and interactions by R2 and R1ρ relaxation dispersion NMR. *Methods* 2018;**148**:28–38. https://doi.org/10.1016/j.ymeth.2018.04.026
89. Carr HY, Purcell EM. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys Rev* 1954;**94**:630–8. https://doi.org/10.1103/PhysRev.94.630
90. Meiboom S, Gill D. Modified spin-echo method for measuring nuclear relaxation times. *Rev Sci Instrum* 1958;**29**:688–91. https://doi.org/10.1063/1.1716296
91. Faber C, Sticht H, Schweimer K *et al.* Structural rearrangements of HIV-1 Tat-responsive RNA upon binding of neomycin B. *J Biol Chem* 2000;**275**:20660–6. https://doi.org/10.1074/jbc.M000920200
92. Sauer WHB, Schwarz MK. Molecular shape diversity of combinatorial libraries: a prerequisite for broad bioactivity. *J Chem Inf Comput Sci* 2003;**43**:987–1003. https://doi.org/10.1021/ci025599w

Received: November 6, 2024. **Revised:** April 22, 2025. **Editorial Decision:** May 10, 2025. **Accepted:** June 4, 2025