Driving Factors in Amiloride Recognition of HIV RNA Targets

Journal:	*Organic & Biomolecular Chemistry*
Manuscript ID	OB-ART-08-2019-001702.R1
Article Type:	Paper
Date Submitted by the Author:	21-Sep-2019
Complete List of Authors:	Patwardhan, Neeraj; Duke University, Chemistry Cai, Zhengguo; Duke University, Chemistry Umuhire Juru, Aline; Duke University Hargrove, Amanda; Duke University, Chemistry

SCHOLARONE™
Manuscripts

Driving Factors in Amiloride Recognition of HIV RNA Targets

Neeraj N. Patwardhan; Zhengguo Cai; Aline Umuhire Juru; Amanda E. Hargrove*

Department of Chemistry, Duke University, 124 Science Drive, Durham, NC 27705, USA

*Corresponding author / Lead Contact: amanda.hargrove@duke.edu

Abstract

Noncoding RNAs are increasingly promising drug targets yet ligand design is hindered by a paucity of methods that reveal driving factors in selective small molecule:RNA interactions, particularly given the difficulties of high-resolution structural characterization. HIV RNAs are excellent model systems for method development given their targeting history, known structure-function relationships, and the unmet need for more effective treatments. Herein we report a strategy combining synthetic diversification, profiling against multiple RNA targets, and predictive cheminformatic analysis to identify driving factors for selectivity and affinity of small molecules for distinct HIV RNA targets. Using this strategy, we discovered improved ligands for multiple targets and the first ligands for ESSV, an exonic splicing silencer critical to replication. Computational analysis revealed guiding principles for future designs and a predictive cheminformatics model of small molecule:RNA binding. These methods are expected to facilitate progress toward selective targeting of disease-causing RNAs.

Keywords: cheminformatics; HIV RNA; molecular recognition; QSAR, RNA recognition

Introduction:

The recent resurgence in targeting RNA with small molecules has been driven in part by a renewed appreciation for regulatory RNAs in mammalian systems as well as several successful examples of biologically active RNA ligands.[1-8] At the same time, these examples are currently limited to a handful of different RNAs, and there are few examples of success in animal models or clinical trials. To further advance the field, continued effort is needed in several areas including the identification of new chemical scaffolds capable of targeting RNA, the development of robust and general high-throughput screening assays, and the

development of a detailed understanding on the best pathways available for optimizing RNA-targeted chemical probes both for potency and selectivity.

We and others have used scaffold-based design and/or well-studied RNA model systems, such as HIV-1-TAR, to test methods and approaches for RNA-targeted probe development.[9-21] For example, we developed amiloride as an RNA-targeted scaffold based on initial reports of dimethylamiloride (**DMA-001**) binding to HIV-1-TAR[12] and the scaffold's promising pharmacological properties (amiloride itself is an FDA-approved drug). In our initial work, we evolved the lead molecule (**DMA-001**) to **DMA-169**, which demonstrated 100-fold increased activity *in vitro* and high selectivity against DNA and tRNA controls (**Figure 1**).[12, 19, 22] SOFAST ^{1}H-^{13}C [HMQC] NMR studies revealed site-specific binding of **DMA-169** to the bulge region of HIV-1-TAR. In addition, the range of activities observed for different amiloride derivatives allowed for the prediction of HIV-1-TAR selectivity based on cheminformatic parameters, suggesting that this scaffold might be used to reveal fundamental insights into small molecule:RNA binding. We thus chose amiloride as the small molecule scaffold for this study and synthesized a library of new derivatives based on insights from our previous work targeting HIV-1-TAR RNA,[19] selectivity profiling,[22] and RNA-privileged chemical space.[23]

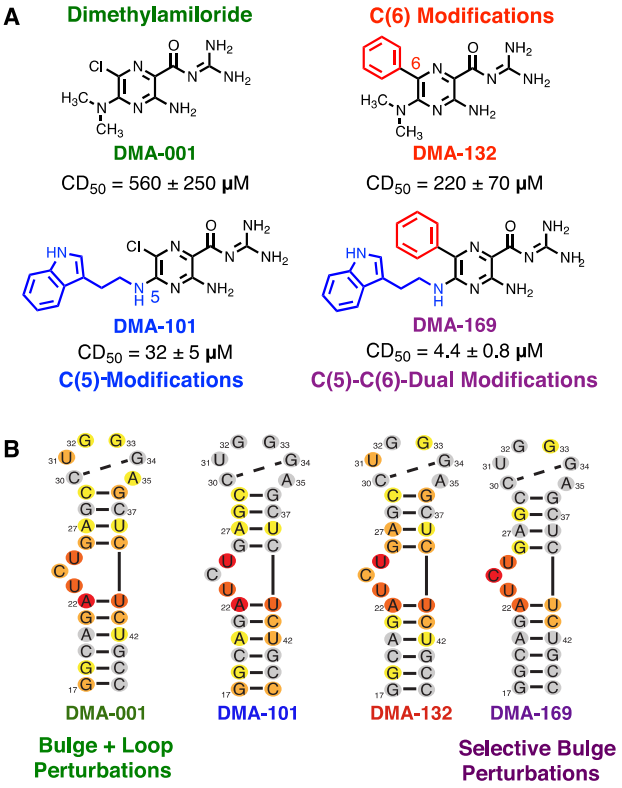


Figure 1. A) Stepwise modification at the C(5) and C(6) positions of amiloride scaffold. B) Heat maps of ^{1}H-^{13}C [HMQC] SOFAST NMR experiments with amiloride and HIV-1-TAR RNA. Reproduced/adapted from reference [19] with permission from the Royal Society of Chemistry.

HIV RNAs were chosen as the ideal RNA model system given that the structure and function of much of the RNA genome is well characterized and HIV/AIDS remains an important area of clinical need.[24, 25] Five HIV RNAs, HIV-1-TAR, HIV-2-TAR, RRE-IIB, HIV-1-FSS and ESSV RNAs (**Figure 2**) were chosen for their structural diversity and known roles in HIV replication and virulence.[16] Moreover, HIV-1-TAR, HIV-2-TAR, RRE-IIB and HIV-1-FSS are well explored small molecule targets.[18] ESSV RNA, while not explored for small molecule targeting, has been identified as critical for viral mRNA splicing and viral production,[26-28] with mutations to ESSV leading to a 95% reduction in HIV-1 replication.[29] These five RNA targets thus provide ideal model systems for assessing the structural selectivity of our small molecule library with an added benefit of uncovering new lead compounds.

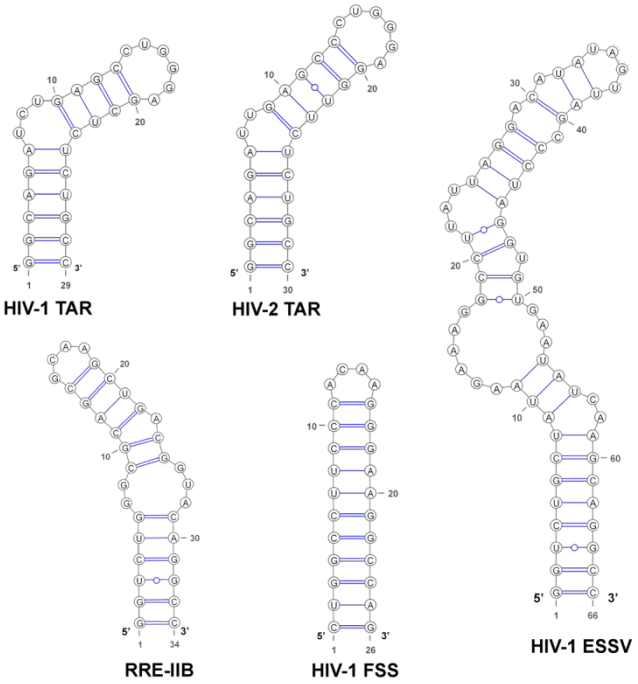


Figure 2. Secondary structures of RNAs screened. HIV-1 and HIV-2-TAR RNA contain a hexanucleotide hairpin loop with a trinucleotide bulge, RRE-IIB contains an 1X2 internal loop with a smaller hairpin loop. HIV-1-FSS contains only a smaller hairpin loop while the HIV-1-ESSV structure is larger with several potential binding pockets. TAR = transactivation response element; RRE = Rev response element; FSS = frameshift-stimulating; ESSV = exonic splicing silencer of Vpr.

We herein report a new strategy, developed using these model systems, that elucidates the critical factors necessary for designing high-affinity and selective amiloride derivatives toward HIV RNA constructs. This

strategy relies on systematic synthetic modifications of the scaffold, profiling of affinity against multiple RNA targets, and quantitative cheminformatic analysis. Using this strategy, we report the first small molecule ligands for ESSV RNA, a promising new target in HIV, along with derivatives with enhanced activity towards HIV-1-TAR. Furthermore, we report quantitative and predictive cheminformatics studies for small molecule binding of specific RNAs,[30-32] which highlights chemical parameters that can be exploited for the future design of selective derivatives for ESSV and HIV-1-TAR. These methods lay out a pathway toward rational targeting of biologically relevant RNAs without requiring high-resolution structure.

Results and Discussion:

Synthesis: We began by replacing the phenyl substituent at the C(6) position of **DMA-169** with other aryl groups since substituted aryl and naphthyl substituents at C(6) previously showed improved activity against HIV-1-TAR (See Scheme S1 and S2 of the ESI).[19] Additionally, derivatives with nitrogen containing heterocycles pyridine and pyrimidine were synthesized (**DMA-190 – DMA-191**) to explore the effect of adding hydrogen bond forming heteroatoms at the C(6) position. The effect of other amine subunits at the C(5) positions, while keeping the C(6) phenyl substituent intact, was then explored, where subunits were chosen from the C(5) monosubstituted derivatives previously described.[33, 19] Notably, **DMA-188** features a C(5) subunit with a shorter methylene linker between the pyrazine and indole rings relative to **DMA-169**. As the C(5) pyrrolidine-containing derivative **DMA-19** showed strong binding to HIV-1-TAR,[19] C(6) aryl derivatives with a pyrrolidine ring at C(5) were synthesized. Since the first study only included 2 cyclic secondary amine subunits at C5 (**DMA-19**, and **DMA-20**), three new monosubstituted derivatives **DMA-201-203** were synthesized, followed by their C(6) phenyl derivatives. Finally, we included further amine subunits at the C(6) position of the amiloride core pyrazine scaffold via the addition of a 2-chloro pyrimidine subunit following the strategy recently reported by Kelso and coworkers.[34] These derivatives were synthesized to explore whether addition of further diversity at the C(6) position can increase the RNA affinity and selectivity. Synthetic details are available in the Supporting Information (pS21).

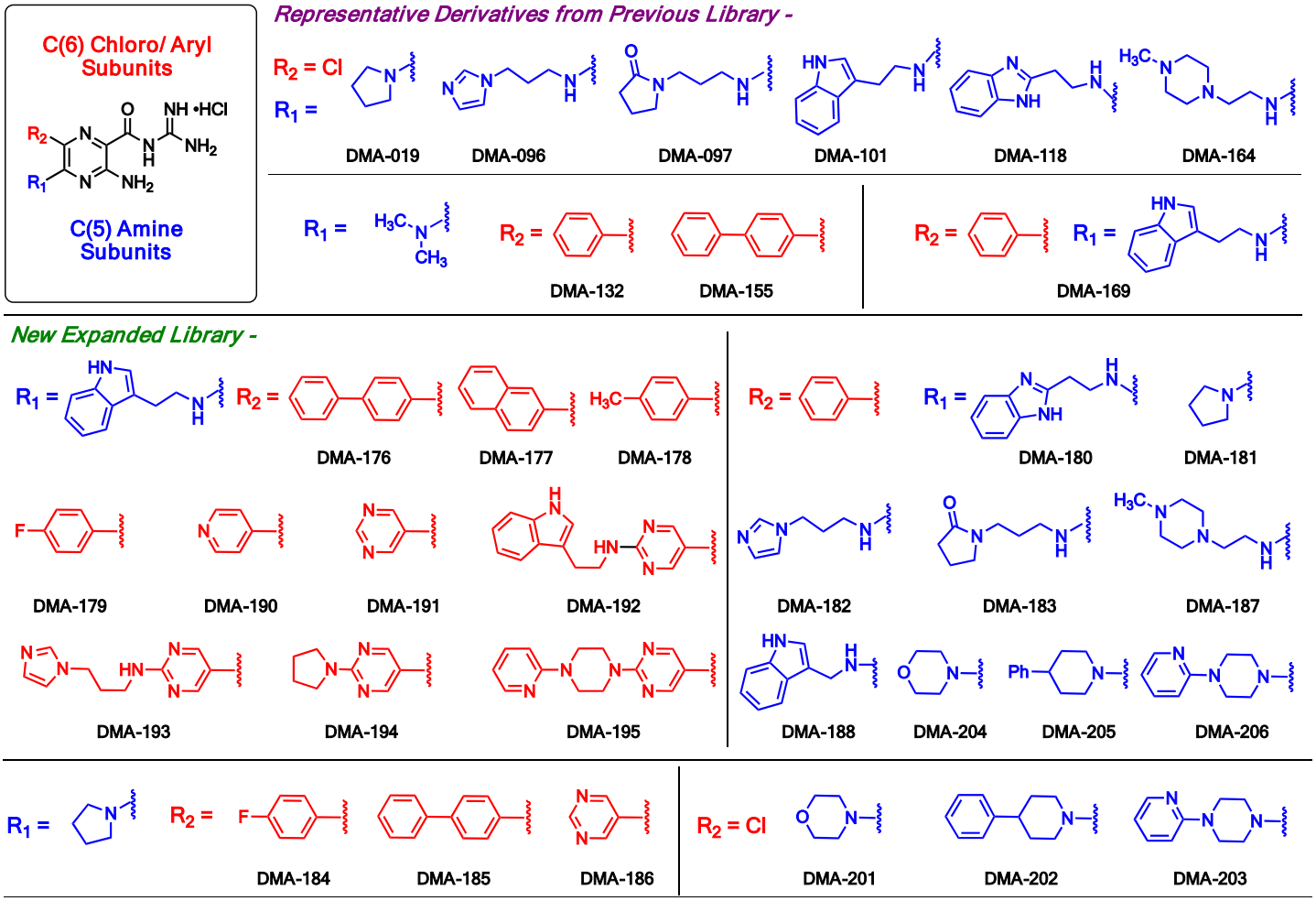


Figure 3: Structures of amiloride derivatives tested. (See supporting information for synthetic details of expanded library).

Profiling of Affinity and Selectivity against 5 HIV RNA Targets. The five HIV RNA targets (HIV-1-TAR, HIV-2-TAR, RRE-IIB, HIV-1-FSS and ESSV) were screened against 25 newly synthesized C(5), C(6) derivatives, along with 10 related previously prepared amiloride derivatives (**Figures 1** and **3**). We employed a simple and fast displacement assay utilizing a FRET pair labeled Tat peptide as recently reported for HIV-1-TAR, HIV-2-TAR, and RRE.[22] To expand this screening method to HIV-1-FSS and ESSV, we first measured the binding affinity of the indicator peptide to the RNAs (Table S1), adjusted the RNA assay concentrations for fraction bound, and measured the Z'-score to evaluate assay quality (Table S2). While HIV-1-FSS had a weaker Tat binding affinity and thus required a higher RNA concentration, both assays were within the acceptable range. The small molecules were screened at concentrations of 5, 10, and 25 μM to identify both strong and weak binders. The percent fluorescent indicator displacement (%FID) was calculated as reported (See supporting

information), and compounds that produced >25% FID were designated as hits based on previous work.[22] In all assays, neomycin was included as a positive control (data shown in SI). While more hits were identified at the higher concentrations such as 25 μM (Figure S1 and S2), screening at the lower concentration of 10 μM better highlighted potential selectivity profiles (**Figure 4**). For example, **DMA-169** and **DMA-188** were hits for only the two TAR RNAs, **DMA-180 and DMA-182** are only hits for HIV-1-TAR, **DMA-194** and **DMA-195** are only hits for ESSV, and **DMA-193** binds all RNAs. These selectivity profiles likely reflect differences in binding pockets between the RNAs studied. The higher numbers of hits for ESSV and TAR relative to RRE-IIB and HIV-1-FSS were consistent with the previously observed bulge selectivity for amiloride derivatives.[19, 22] The increase in hits for HIV-1-TAR RNA vs. HIV-2-TAR suggested that the larger bulge of HIV-1-TAR may be more suitable for amiloride binding. The large number of hits for ESSV may indicate that the presence of multiple secondary structures in close proximity allows the RNA to accommodate a wide range of ligands. Finally, HIV-1-FSS has far more limited structural features relative to the other constructs (a single 4 n.t. hairpin, **Figure 2**), which may be responsible for the low number of hits, though it should be noted that the higher RNA concentrations required in this assay may be leading to false negatives.

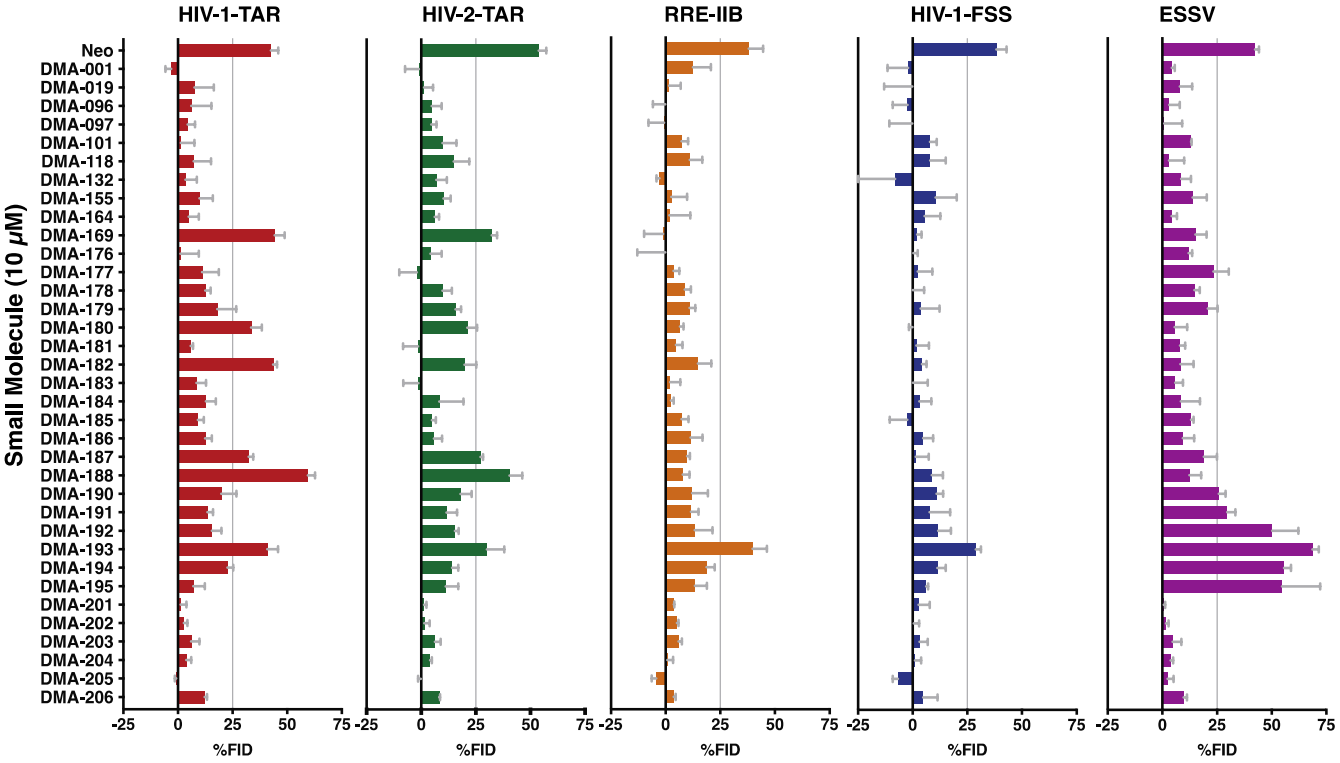


Figure 4: Screening data for 35 amiloride derivatives at 10 μM small molecule concentration against HIV-1-TAR, HIV-2-TAR, RRE-IIB, HIV-1-FSS, and ESSV RNA targets. Neomycin (Neo) was included as a positive control. **Buffer:** 50 mM Tris, 50 mM KCl, 0.01% Triton-X-100, 5% DMSO, pH = 7.4; **Tat peptide**: 50 nM; **RNA:** HIV-1-TAR = 40 nM, HIV-2-TAR =

40 nM, RRE-IIB = 75 nM, HIV-1-FSS = 250 nM, ESSV = 70 nM; **Small molecule:** 10 μM; Excitation λ = 485 nm, Emission λ = 590 nm; Incubation time = 30 min.

Trends in affinity and selectivity were further visualized through agglomerative hierarchical clustering (AHC) using the %FID values at 5, 10, and 25 μM small molecule concentrations, which allowed unbiased categorization of the small molecules based on binding profiles (**Figure 5**). Importantly, ESSV and TAR selective ligands are readily distinguished. In addition, the more promiscuous ligands are generally clustered together, with **DMA-193** in a single group due to its much higher overall binding. Compounds with weak or no binding are also grouped together, and the AHC analysis further highlights the reduced number of hits identified for RRE-IIB and HIV-1 FSS.

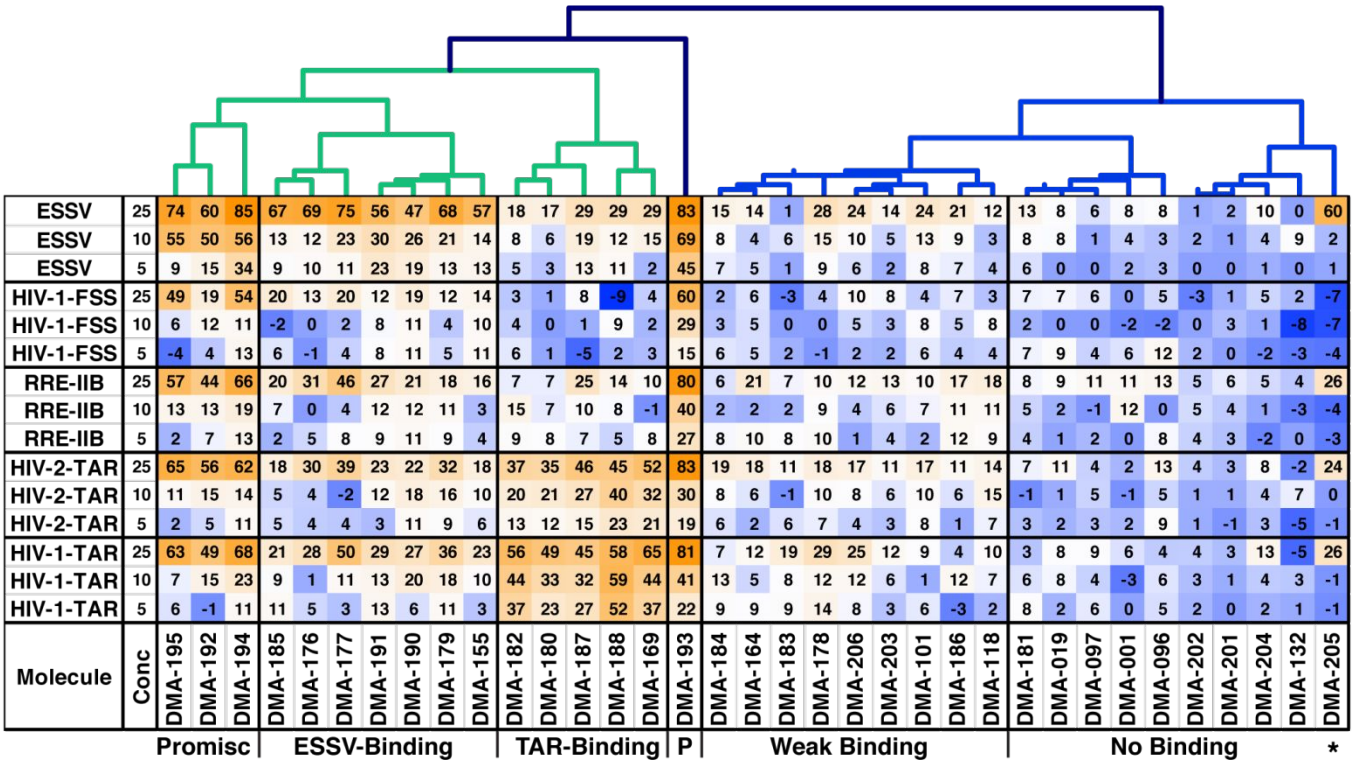


Figure 5. Heat map and agglomerative hierarchical clustering analysis of screening data at 3 different small molecule concentrations. Numbers reflect %FID at listed concentration. Clusters are labeled to indicate ESSV-binding, TAR-binding, promiscuous binding (Promis, P), no binding, and weak binding. ***DMA-205** was incorrectly clustered as a non-binder due to very low %FID at lower concentrations and was moved to the Weak Binding group in further analyses.

Preliminary cheminformatic analysis. To further understand the molecular basis for the observed RNA-binding patterns, we calculated 20 cheminformatic parameters for the amiloride library.[35, 36, 23] These parameters were then used in linear discriminant analysis (LDA) as previously described[19] using the clusters

formed in AHC as groups **(Figure 5**). **(Figure 6A**) Groups were well predicted by the cheminformatic parameters during training (100%) but less well in the cross-validation (37%) (Table S6-S7). Qualitative trends could be inferred from the loading plots (**Figure 6B**). For example, non- or weak binding ligands tended to have more oxygens, more sp3 centers and higher relative polar surface area (PSA) while promiscuous binders had high nitrogen counts, heteroaromatic rings, hydrogen bond acceptors, and topological polar surface area (tPSA). Interesting, ESSV ligands were less hydrophilic than other classes (further analysis and data in Supporting Information, pS15). This analysis was limited by the use of only three concentration points as well as the number of molecules and parameters used. We thus conducted more detailed analyses of binding, which allowed for in-depth structure-activity/selectivity analysis followed by quantitative analysis with a larger set of parameters.

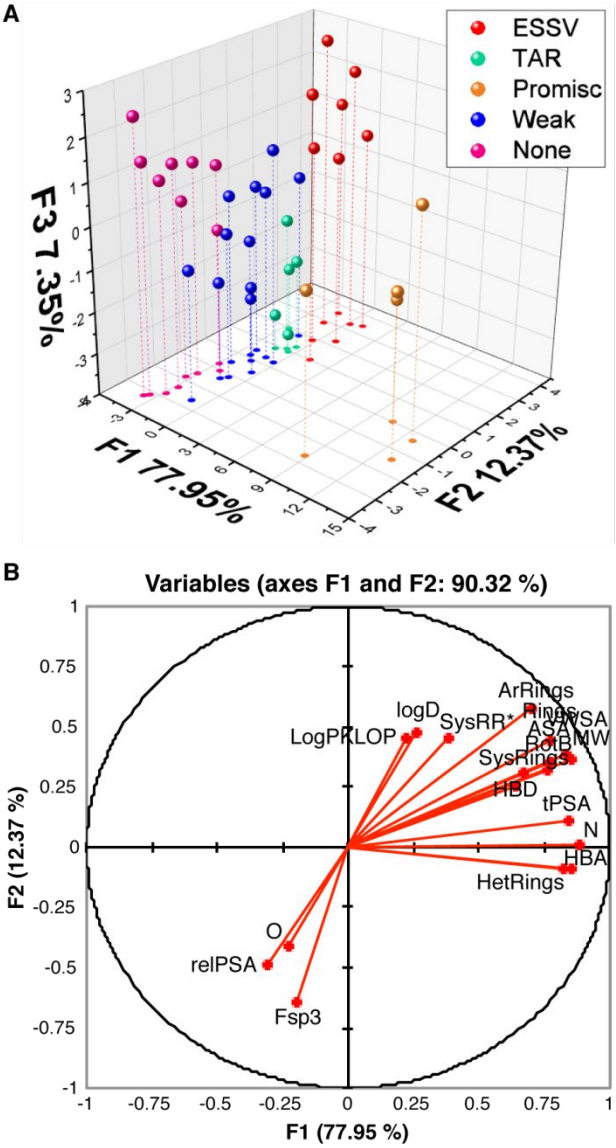


Figure 6. A. LDA plot for the first three axes (factors) depicting clustering for the five groups of RNA ligands using cheminformatic parameters. **B**. Loading plot of the contributions of each cheminformatic parameter on F1 vs F2.

Quantitative assessment of amiloride binding and trends. To better assay selectivity patterns, the CD_{50} values were determined for "hit" amiloride derivatives against the corresponding RNA targets by performing small molecule titrations with the Tat peptide displacement assay (**Table 1**). As 18 of the 35 DMAs tested were a hit for at least one RNA at the 25 μM concentration point, several interesting trends can be noted.

For HIV-1 and HIV-2-TAR binders, a clear preference emerged for the C6 phenyl substituted derivatives versus other aryl subunits for both binding and selectivity (e.g. **DMA-169** > **DMA-176-179**, **DMA-190-195),** possibly due to steric hinderance (**Table 1**, **Figure 3**). Additionally, substituting a C(6) phenyl group for the

C(6) chloro rescued HIV-1-TAR binding activity of several derivatives that showed low %FID values or interference in previous work. For example, **DMA-180**, **182, 187** all showed higher % FID and improved CD_{50} to HIV-1-TAR (**Table 1**) as compared to their respective C(6)-chloro counterparts,[19] with **DMA-180** and **182** showing remarkable selectivity towards HIV-1 and HIV-2-TAR RNA. These results could be due to a number of factors, including better complementarity, stacking and/or electronic properties conferred by the C(6) phenyl group. However, no improvement in affinity was found for **DMA-183**, a C(6) phenyl derivative of the previously described non-binder **DMA-097**, highlighting the impact of both C(5) and C(6) substituents towards observed RNA-binding activity.

Closer evaluation of TAR binding compounds relative to **DMA-169** revealed additional trends in selectivity and structural properties. First, reducing the flexibility and sp3 carbon count at the C(5) position (**DMA-188** vs. **DMA-169**) led to a 2-3-fold increase in affinity for HIV-1-TAR as well as enhanced selectivity relative to HIV-2-TAR (from 2 to 5-fold) and ESSV (from 8 to 22 fold) (Table S3). This result is consistent with increased binding affinity and selectivity relating to reduced entropy and/or pre-organization, in line with the current (**Figure 6**) and previous cheminformatic analyses of RNA ligands.[23] By contrast, the replacement of a tryptamine subunit (**DMA-169**) by an aminoethyl benzimidazole subunit (**DMA-180**), a switch in aromatic heterocycle with the same sized linker, reduces the affinity for HIV-1-TAR and leaves the selectivity over HIV-2-TAR unimproved, despite increased selectivity against ESSV.

Table 1: CD_{50} values for DMA molecules identified as hits in the 25 μM conc. screening experiment.

Small Molecule	CD_{50} (μM)				
	HIV-1-TAR	HIV-2-TAR	RRE-IIB	HIV-1-FSS	ESSV
DMA-155	[b]	-	-	-	20.1 ± 3.1
DMA-169	3.9 ± 0.7 [a]	9.3 ± 1.3 [a]	29.0 ± 7.1[a]	-	30.3 ± 4.5
DMA-176	18.3 ± 1.0	17.2 ± 0.24	16.2 ± 0.45	-	11.3 ± 0.6
DMA-177	16.4 ± 0.57	18.8 ± 0.4	18.0 ± 0.35	-	11.9 ± 1.1
DMA-178	46.0 ± 21.7	-	-	-	17.2 ± 1.3
DMA-179	48.1 ± 39.0	20.5 ± 1.5	-	-	15.2 ± 1.1
DMA-180	7.73 ± 0.92	18.5 ± 1.2	-	-	-
DMA-182	7.47 ± 0.80	21.8 ± 2.5	-	-	-
DMA-185	-	-	-	-	19.5 ± 0.8
DMA-187	13.3 ± 0.87	12.3 ± 1.1	71.8 ± 30.3	-	46.0 ± 26.7
DMA-188	1.36 ± 0.15	6.2 ± 0.6	-	-	30.5 ± 14.7
DMA-190	24.2 ± 5.57	-	-	-	48.9 ± 23.1
DMA-191	12.5 ± 3.25	39.3 ± 10.2	29.8 ± 6.8	-	12.6 ± 2.6
DMA-192	7.21 ± 1.30	9.68 ± 0.23	8.51 ± 0.52	[b]	5.0 ± 1.1
DMA-193	6.79 ± 0.91	12.5 ± 1.1	14.6 ± 2.7	28.5 ± 7.2	4.4 ± 0.6
DMA-194	15.9 ± 2.53	22.9 ± 2.9	25.7 ± 3.9	35.3 ± 13.7	8.02 ± 0.82
DMA-195	13.1 ± 1.07	11.3 ± 0.98	11.9 ± 0.66	18.2 ± 0.7	10.7 ± 0.5
DMA-205	33.6 ± 0.51	30.1 ± 3.28	19.0 ± 3.9	-	34.8 ± 3.16

[a] Previously reported CD_{50} values; [b] Interference observed at high concentrations, likely due to aggregation. (-) Not determined. **Buffer:** 50 mM Tris, 50 mM KCl, 0.01% Triton-X-100, 5% DMSO, pH = 7.4; **Tat peptide conc**: 50 nM; **RNA conc:** HIV-1-TAR = 40 nM, HIV-2-TAR = 40 nM, RRE-IIB = 75 nM, HIV-1-FSS = 250 nM, ESSV = 70 nM; **Small molecule conc:** 0-100 μM; Excitation λ = 485 nm, Emission λ = 590 nm; Incubation time = 30 min. CD_{50} values below 10 μM are highlighted in blue.

In contrast to HIV-1 and HIV-2-TAR, strong ESSV binding was observed by the derivatives with increased bulk at the C6 position (**DMA-176-179**, and **DMA-192-195**), nearly irrespective of the subunit at the C5 position (e.g. **DMA-155** vs **DMA-185**). Selectivity ratios for strong ESSV ligands were significant for **DMA-194**, **DMA-178** and **DMA-179**, each with a 2-3-fold selectivity over HIV-1-TAR.

The more promiscuous ligands (**DMA-191-196 in Table 1**) had significantly increased number of nitrogen atoms at the C(6) position, consistent with the LDA analysis above. In addition, the tryptamine subunit at C(5) appeared to confer affinity, as all 10 derivatives with this subunit and an aryl C(6) group were hits for both HIV-1-TAR and ESSV. In contrast, decreased RNA binding was observed with substitutions of secondary amines at the C(5) position (**DMA-181**, **184-186**, and **201-206**) (Figure 3). Only **DMA-205**, with a C(6)-phenyl substitution and C(5)-phenyl piperidine, was found to be a "hit" for any RNA and demonstrated weak binding and little selectivity.

Quantitative Modeling of RNA Affinity Profiles. Given the preliminary cheminformatic trends observed via LDA and the observed trends in CD_{50} value, we performed a more detailed cheminformatics analysis. We used the CD_{50} values as a proxy for binding affinity, which were modeled with ligand molecular descriptors in order to build a quantitative structure-activity relationship (QSAR).[37, 38, 30, 32, 39] This was performed for HIV-1-TAR, HIV-2-TAR and ESSV as they had the greatest number of hits and thus data points for analysis. For these RNA targets, the best two-parameter linear models ($y \sim 1 + x_1 + x_2$) were constructed after model searching using the method of exhaustion and evaluated by the R^2 value and leave-one-out cross validation (LOOCV).

For HIV-1-TAR, multivariate regression produced models with moderate R^2 values. The best model ($R^2 = 0.71$) contains two descriptors, *PEOE_VSA-0* (#179, *sum of van der Waals surface area of atoms whose partial charge is in the range [-0.05,0.00*) and *SMR_VSA2* (#331, *sum of van der Waals surface area of atoms whose contribution to molar refractivity (MR) is in [0.26,0.35]*) (**Figure 7A**). The QSAR model constructed by these two descriptors indicates that DMAs having reduced surface area that is weakly basic and more aromatic carbons/heteroatoms[40, 41] will lead to lower CD_{50} values. These factors might account for the relatively low CD_{50} value of **DMA-188**, where both C(5) and C(6) positions feature aromatic rings (reflected by high MR) but the C(5) position contains a shorter linker between the pyrazine and indole rings relative to other DMAs (reflected by reduced weakly basic surface area). Analysis with HIV-2-TAR produced a less robust model (SI, pS17), possibly due to the reduced number of data points used.

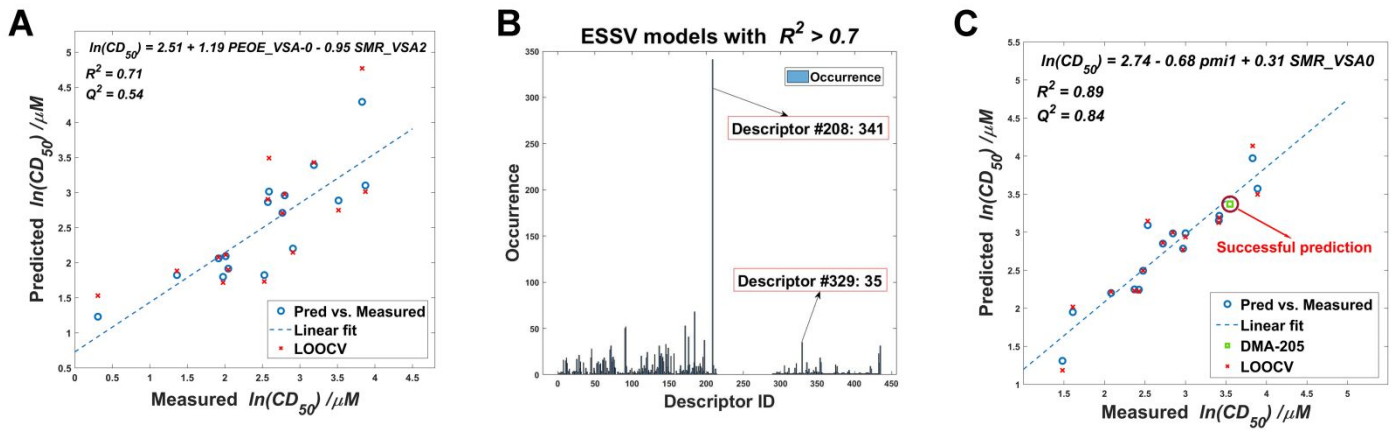


Figure 7: A. HIV-1-TAR model summary: model equation, linear regression and LOOCV efficiency. **B.** Occurrence histogram of descriptors appearing in the two-parameter linear models of ESSV with R^2 >0.7, the descriptor *pmi1* dominates, another descriptor used in the best two-parameter linear model of ESSV is *SMR_VSA0*. **C.** ESSV model summary: model equation, linear regression, LOOCV efficiency and DMA-205 prediction.

Modeling for ESSV was more predictive and identified different descriptors. More than 1100 models have R^2 >0.7 and thus the number of top models is suitable for a statistical occurrence analysis. We counted the number of occurrences of each descriptor that appeared in the above top models and summarized them into a histogram (**Figure 7B**). Occurrence analysis showed that a topological descriptor *pmi1* (#208, *first diagonal element of diagonalized moment of inertia tensor*) dominated the top models with R^2 > 0.7, indicating the shape of the molecules strongly affect the binding. The best model includes another descriptor: *SMR_VSA0* (#329, *sum of van der Waals surface area of atoms whose contribution to molar refractivity is in [0,0.11]*), suggesting that polarizability of the molecule also contributes to ESSV binding. The coefficients of the two parameters in the model suggested that the topological term contributes more than molecular size and polarizability (*SMR_VSA* term[40]). The linear combination of these two descriptors produced a QSAR model with R^2 = 0.89 in regression analysis and Q^2 = 0.84 in the leave-one-out cross-validation analysis (**Figure 7C**), indicating that individual CD_{50} values are well predicted. Because we trained the QSAR model using the dataset without **DMA-205**, we could further test the model by predicting the CD_{50} value of **DMA-205** (**Figure 7C**) using the *pmi1* and *SMR_VSA0* values. The predicted CD_{50} (29.1 μM) is close to the measured one (34.8 μM), suggesting that this QSAR model has predictive power and may be used to guide future modifications of this specific scaffold for tighter binding to ESSV. To begin validation of this model, we built a virtual library consisting of 15 DMAs with synthetic feasibility (Figure S9). The CD_{50} values of these molecules were

predicted by substituting their *pmi1* and *SMR_VSA0* values into the QSAR model (Figure S9). The predicted CD_{50} values indicate that one compound (**DMA-207**, Scheme S1) would have tighter binding ($CD_{50,\ pred}$ =2.3 μM) than the strongest ESSV ligand in the initial screen (**DMA-193**, CD_{50} = 4.4 μM). Therefore, we synthesized this compound and its measured CD_{50} value ($CD_{50,\ measured}$ = 6.5 μM) fits reasonably with the prediction, supporting the use of QSAR modeling in rational design. Future refinements of the model will include additional molecules, especially in this lower CD_{50} value range, as well as examine selectivity.

Conclusions

In summary, we demonstrated that harnessing a combination of organic, biochemical, and computational techniques can identify key factors contributing to affinity and selectivity in small molecule:RNA interactions in vitro. Using these methods, we discovered the first reported ligands for ESSV as well as improved amiloride ligands for HIV-1-TAR and developed cheminformatic models for these small molecule:RNA interactions. These outcomes were achieved by expanding the synthetic methods for diversification of amilorides, screening against 5 distinct HIV RNA targets, profiling structure-activity and selectivity relationships, and performing quantitative cheminformatic analysis to identify driving factors for both HIV-1-TAR and ESSV RNA affinity. In addition to traditional SAR results, such as position-specific steric requirements, flexibility, and functional group type, the developed quantitative models correlated specific chemical properties with RNA binding for both HIV-1-TAR and ESSV and found that the most predictive descriptors were different for the two RNA targets. For example, these results suggested that tuning amilorides for HIV-1-TAR binding should focus on aromatic surface area and electrostatic properties while tuning for ESSV should focus on molecular shape. The differential binding of ligands from the same scaffold to different RNA structures suggests that the unique features of the RNA determine the molecular properties critical for binding and that screening of larger, more diverse libraries could yield additional insights. In progress characterizations of ESSV:amiloride interactions through NMR experiments will lend further insight as will evaluation of specific ESSV-targeted biological activity. Moving forward, it will be critical to refine and implement these quantitative models in biological systems. The methods used here outline a pathway toward the rational design of RNA-specific small

molecules, without the incorporation of difficult or time-consuming structural studies, that should be applicable to other RNAs of interest, including those without high-resolution structures.

Experimental Procedures

Readers are referred to the Supplementary Information for detailed experimental procedures.

Conflicts of Interest: The authors declare no conflict of interest.

Acknowledgments

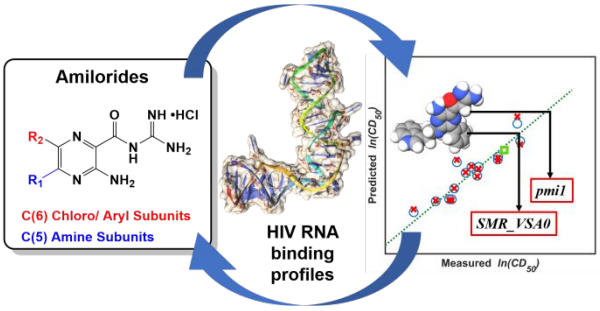
We would like to thank Professor Blanton Tolbert and Gaius Takor (Case Western Reserve University) for providing the ESSV RNA used for this study and members of the Hargrove Lab for valuable insight. The authors acknowledge financial support for this work from Duke University, the US National Institute of Health Grant/Award Number: U54AI150470, and the National Science Foundation (CAREER Award 1750375). We would also like to thank Dr. Peter Silinski, Director of Analytical Services, Duke University for performing the high-resolution mass spectroscopy analyses, and the Duke University NMR Center for use of the 500 MHz and 400 MHz spectrometers for analysis of small molecules.

References

1. T. R. Cech and J. A. Steitz, *Cell*, 2014, **157**, 77-94.
2. K. V. Morris and J. S. Mattick, *Nat. Rev. Genet.*, 2014, **15**, 423.
3. A. J. Angelbello, J. L. Chen, J. L. Childs-Disney, P. Zhang, Z.-F. Wang and M. D. Disney, *Chem. Rev.*, 2018, **118**, 1599-1663.
4. A. Donlic and A. E. Hargrove, *Wiley Inter. Rev. RNA*, 2018, **9**, e1477.
5. B. S. Morgan, J. E. Forte and A. E. Hargrove, *Nucleic Acids Res.*, 2018, **46**, 8025-8037.
6. H. Ratni, M. Ebeling, J. Baird, S. Bendels, J. Bylund, K. S. Chen, N. Denk, Z. Feng, L. Green, M. Guerard, P. Jablonski, B. Jacobsen, O. Khwaja, H. Kletzl, C.-P. Ko, S. Kustermann, A. Marquet, F. Metzger, B. Mueller, N. A. Naryshkin, S. V. Paushkin, E. Pinard, A. Poirier, M. Reutlinger, M. Weetall, A. Zeller, X. Zhao and L. Mueller, *J. Med. Chem.*, 2018, **61**, 6501-6517.
7. A. J. Angelbello, S. G. Rzuczek, K. K. Mckee, J. L. Chen, H. Olafson, M. D. Cameron, W. N. Moss, E. T. Wang and M. D. Disney, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 7799-7804.
8. J. Lee, Y. Bai, U. V. Chembazhi, S. Peng, K. Yum, L. M. Luu, L. D. Hagler, J. F. Serrano, H. Y. E. Chan, A. Kalsotra and S. C. Zimmerman, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 8709-8714.
9. L. Ratmeyer, M. Zapp, M. Green, R. Vinayak, A. Kumar, D. W. Boykin and W. D. Wilson, *Biochemistry*, 1996, **35**, 13689-13696.
10. M. L. Zapp, D. W. Young, A. Kumar, R. Singh, D. W. Boykin, W. D. Wilson and M. R. Green, *Bioorg. Med. Chem.*, 1997, **5**, 1149-1155.
11. N. Gelus, C. Bailly, F. Hamy, T. Klimkait, W. D. Wilson and D. W. Boykin, *Bioorg. Med. Chem.*, 1999, **7**, 1089-1096.

12. A. C. Stelzer, A. T. Frank, J. D. Kratz, M. D. Swanson, M. J. Gonzalez-Hernandez, J. Lee, I. Andricioaei, D. M. Markovitz and H. M. Al-Hashimi, *Nat. Chem. Biol.*, 2011, **7**, 553-559.
13. N. M. Bell, A. L'Hernault, P. Murat, J. E. Richards, A. M. Lever and S. Balasubramanian, *Biochemistry*, 2013, **52**, 9269-9274.
14. A. Blond, E. Ennifar, C. Tisne and L. Micouin, *ChemMedChem*, 2014, **9**, 1982-1996.
15. J. Sztuba-Solinska, S. R. Shenoy, P. Gareiss, L. R. Krumpe, S. F. Le Grice, B. R. O'Keefe and J. S. Schneekloth, Jr., *J. Am. Chem. Soc.*, 2014, **136**, 8402-8410.
16. S. F. J. Le Grice, *Curr. Top. Microbiol. Immunol.*, 2015, **389**, 147-169.
17. F. A. Abulwerdi, M. D. Shortridge, J. Sztuba-Solinska, R. Wilson, S. F. Le Grice, G. Varani and J. S. Schneekloth, Jr., *J. Med. Chem.*, 2016, **59**, 11148-11160.
18. T. Hermann, *Wiley Interdiscip Rev RNA*, 2016, **7**, 726-743.
19. N. N. Patwardhan, L. R. Ganser, G. J. Kapral, C. S. Eubanks, J. Lee, B. Sathyamoorthy, H. Al-Hashimi and A. E. Hargrove, *MedChemComm*, 2017, **8**, 1022-1036.
20. L. R. Ganser, J. Lee, A. Rangadurai, D. K. Merriman, M. L. Kelly, A. D. Kansal, B. Sathyamoorthy and H. M. Al-Hashimi, *Nat. Struct. Mol. Biol.*, 2018, **25**, 425-434.
21. S. S. Chavali, R. Bonn-Breach and J. E. Wedekind, *J. Biol. Chem.*, 2019, DOI: 10.1074/jbc.REV119.006860.
22. N. N. Patwardhan, Z. Cai, C. N. Newson and A. E. Hargrove, *Org. Biomol. Chem.*, 2019, **17**, 1778-1786.
23. B. S. Morgan, J. E. Forte, R. N. Culver, Y. Zhang and A. E. Hargrove, *Angew. Chem., Int. Ed.*, 2017, **56**, 13498-13502.
24. R. W. Shafer, *Clin. Infect. Dis.*, 2015, **62**, 136-137.
25. S. Sengupta and R. F. Siliciano, *Immunity*, 2018, **48**, 872-895.
26. J. M. Madsen and C. M. Stoltzfus, *Retrovirology*, 2006, **3**, 10.
27. C. M. Stoltzfus and M. M. Joshua, *Current HIV Research*, 2006, **4**, 43-55.
28. J. Tazi, N. Bakkour, V. Marchand, L. Ayadi, A. Aboufirassi and C. Branlant, *The FEBS Journal*, 2010, **277**, 867-876.
29. J. M. Madsen and C. M. Stoltzfus, *J. Virol.*, 2005, **79**, 10478-10486.
30. Y. N. Sekhar, M. R. S. Nayana, N. Sivakumari, M. Ravikumar and S. K. Mahmood, *J. Mol. Graphics Modell.*, 2008, **26**, 1338-1352.
31. P. Setny and J. Trylska, *J. Chem. Inf. Model.*, 2009, **49**, 390-400.
32. I. Maciagiewicz, S. Zhou, S. C. Bergmeier and J. V. Hines, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 4524-4527.
33. S. P. Velagapudi, A. Pushechnikov, L. P. Labuda, J. M. French and M. D. Disney, *ACS Chem. Biol.*, 2012, **7**, 1902-1909.
34. B. J. Buckley, A. Aboelela, E. Minaei, L. X. Jiang, Z. Xu, U. Ali, K. Fildes, C.-Y. Cheung, S. M. Cook, D. C. Johnson, D. A. Bachovchin, G. M. Cook, M. Apte, M. Huang, M. Ranson and M. J. Kelso, *J. Med. Chem.*, 2018, **61**, 8299-8320.
35. C. F. Stratton, D. J. Newman and D. S. Tan, *Bioorg. Med. Chem. Lett.*, 2015, **25**, 4802-4807.
36. T. A. Wenderski, C. F. Stratton, R. A. Bauer, F. Kopp and D. S. Tan, *Methods Mol. Biol. (N. Y., NY, U. S.)*, 2015, **1263**, 225-242.
37. Y. Guang-Fu and H. Xiaoqin, *Curr. Pharm. Des.*, 2006, **12**, 4601-4611.
38. P. Gramatica, *QSAR Comb. Sci.*, 2007, **26**, 694-701.
39. S. Zhao, T. Gensch, B. Murray, Z. L. Niemeyer, M. S. Sigman and M. R. Biscoe, *Science*, 2018, **362**, 670-674.
40. S. A. Wildman and G. M. Crippen, *J. Chem. Inf. Comput. Sci.*, 1999, **39**, 868-873.
41. P. Labute, *J. Mol. Graphics Modell.*, 2000, **18**, 464-477.

Table of Contents Graphic and Text



RNA profiling reveals key contributors to affinity and selectivity in amiloride:RNA interactions, including a predictive model for ESSV recognition