MedChemComm

RESEARCH ARTICLE

View Article Online
View Journal

Cite this: DOI: 10.1039/c6md00729e

Amiloride as a new RNA-binding scaffold with activity against HIV-1 TAR†‡

Neeraj N. Patwardhan,[a] Laura R. Ganser,[b] Gary J. Kapral,[a] Christopher S. Eubanks,[a] Janghyun Lee,§[b] Bharathwaj Sathyamoorthy,¶[b] Hashim M. Al-Hashimi[ab] and Amanda E. Hargrove *[ab]

Received 31st December 2016,
Accepted 14th March 2017

DOI: 10.1039/c6md00729e

rsc.li/medchemcomm

Diversification of RNA-targeted scaffolds offers great promise in the search for selective ligands of therapeutically relevant RNA such as HIV-1 TAR. We herein report the establishment of amiloride as a novel RNA-binding scaffold along with synthetic routes for combinatorial C(5)- and C(6)-diversification. Iterative modifications at the C(5)- and C(6)-positions yielded derivative **24**, which demonstrated a 100-fold increase in activity over the parent dimethylamiloride in peptide displacement assays. NMR chemical shift mapping was performed using the 2D SOFAST-[^{1}H-^{13}C] HMQC NMR method, which allowed for facile and rapid evaluation of binding modes for all library members. Cheminformatic analysis revealed distinct differences between selective and non-selective ligands. In this study, we evolved dimethylamiloride from a weak TAR ligand to one of the tightest binding selective TAR ligands reported to date through a novel combination of synthetic methods and analytical techniques. We expect these methods to allow for rapid library expansion and tuning of the amiloride scaffold for a range of RNA targets and for SOFAST NMR to allow unprecedented evaluation of small molecule:RNA interactions.

Introduction

Although often regarded as an intercessor in the transfer of genetic information, RNA is presently identified as a key regulatory element in several cellular processes and has been implicated in a range of disease states.[1,2] Furthermore, several regulatory RNA molecules play significant roles in infectious diseases.[3] HIV infection, for example, relies on a single-stranded RNA genome in which specific RNA structures and RNA:protein interactions are recognized as crucial to HIV replication and as potential therapeutic targets.[4–7] In line with this growing appreciation for the regulatory power of RNA comes a need for the continued advancement of methods and tools for the elucidation of RNA structure and cellular function as well as for the therapeutic targeting of RNA.

Small molecule chemical probes offer a unique opportunity to study both the structure and function of biomacromolecules, including RNA.[8–10] The small size and extensive tunability of small molecules have led to highly selective chemical probes for many biological targets, particularly proteins, and have been invaluable to the field of chemical biology. This potential has yet to be fully realized for RNA, possibly due to the limited chemical functionality of RNA that can be exploited for specific molecular recognition and by the highly dynamic structure of RNA. Screening of general small molecule libraries for selective ligands has been successful for some RNA targets; however, this approach has generally yielded low hit-rates.[8–20] One plausible explanation is that current high-throughput screening libraries are largely designed and optimized for protein targets. Moreover, screens for RNA targets often identify promiscuous ligands such as large polycationic aminoglycosides,[21,22] which are well known to possess poor pharmacological properties and interact with several other cellular targets.[8–10] It is worth noting, however, that aminoglycosides and their multivalent analogs have featured prominently in several studies elucidating high affinity interactions with RNA and have demonstrated utility in studying RNA biology.[23–30] Although significant advances have recently been made towards rationally designing small molecule ligands for RNA,[9,14,31–37] the limited available knowledge of the guiding principles of small molecule:RNA recognition has remained a significant

[a] *Department of Chemistry, Duke University, Durham, North Carolina 27708, USA. E-mail: amanda.hargrove@duke.edu; Tel: +1 919 660 1522*
[b] *Department of Biochemistry, Duke University Medical Center, Durham, North Carolina 27708, USA*

† The authors declare no competing interests.

‡ Electronic supplementary information (ESI) available: Synthetic procedures, compound characterizations, biological assays and data, procedures and parameters for docking, and cheminformatic analysis. See DOI: 10.1039/c6md00729e

§ Current Address: Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, 34141 South Korea.
¶ Current Address: Department of Chemistry, Indian Institute of Science Education and Research Bhopal, Bhopal 462066, India.

hurdle. At the same time, some of the most successful screens have been conducted on RNA-targeted libraries generated by the diversification of molecular scaffolds known to interact with RNA, such as phenylbenzamidazoles,[38–40] oxazolidinones,[41–44] and diphenylfurans[45–47] though the library sizes and number of scaffolds tested has been fairly limited.[40,48] Based on these results, we see an urgent need for the identification and development of new RNA-targeted scaffolds that can expand scientists' repertoire for probing RNA structure and function, particularly RNA structures critical to infectious agents such as HIV.

Specifically, small molecule probes targeted to discrete HIV RNA structures would both facilitate the elucidation of critical pathways in HIV biology and open novel therapeutic opportunities.[5,7] One of the most studied HIV RNA targets is the HIV-1 transactivation response (TAR) element, which binds to the Tat protein and is necessary for efficient HIV replication.[49–52] High-throughput and fragment-based screens have yielded small molecules that bind TAR in the micromolar range and inhibit Tat interactions.[53–57] Recently, small molecule microarrays have been employed by Schneekloth and coworkers to discover a TAR-binding probe that inhibits the HIV-induced cell death of T-lymphocytes at 28 μM without measured toxicity.[17] Importantly, this thienopyridine-based probe was charge neutral and bound TAR with a dissociation constant of 2.4 μM. Finally, branched and cyclic peptides[58–60] as well as small synthetic proteins[61] have also shown promise as potent inhibitors of the TAR:Tat interaction. In order to expedite screening and increase hit rates, the Al-Hashimi lab recently pioneered a virtual docking method for RNA by first generating dynamic ensembles of TAR RNA through molecular dynamics (MD) simulations and NMR spectroscopy refinements.[62] Virtual screening of a ~51 000 molecule library followed by experimental binding studies led to identification of six TAR ligands, which included aminoglycosides and intercalators as well as dimethylamiloride **1** (DMA).[14] Amiloride had not been previously identified among RNA ligands, though DNA duplex binding was reported by Waring and co-workers at pH below the pK_a of amiloride (~8.7).[63,64]

Although **1** bound TAR with a modest affinity (K_d = 120 μM), it demonstrated a unique predicted binding profile in its preference for the apical loop of TAR over the more commonly bound bulge region in docking.[14,63] Furthermore, amiloride derivatives are readily accessed through common synthetic transformations on a commercially available scaffold and offer a straightforward route to combinatorial library design.[65,66] In this study we describe the development of an amiloride-based small molecule library and analysis of the suitability of this library for targeting HIV-1 TAR. We evaluated the activity of these derivatives against the TAR:Tat interaction and explored their binding properties *via* 2D NMR and docking experiments. The identification of multiple amiloride-based ligands that selectively interact with TAR RNA supports the approach of identifying RNA chemical probes from targeted small molecule libraries generated through rapid diversification of RNA-binding scaffolds.

Results and discussion

DMA binds both the apical loop and bulge of TAR

We began our study by characterizing the activity and TAR-binding properties of the parent molecule dimethylamiloride **1** (Fig. 1A). Activity screening was performed by measuring displacement of a fluorescently labeled Tat-derived peptide (N-AAARKKRRQRRRAAAK-C) from a truncated TAR sequence featuring the stem-bulge-loop region.[67] The competitive dosage for 50% displacement (CD_{50}) was calculated for

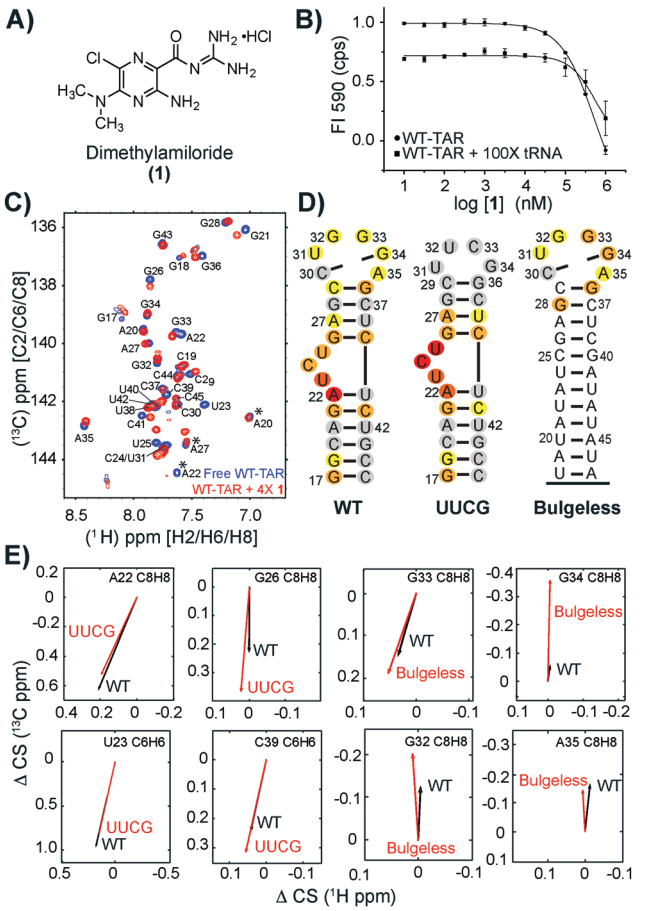


Fig. 1 A) Structure of dimethylamiloride **1** (DMA). B) Displacement assay binding curve with **1** (50 mM Tris-HCl, 50 mM KCl, 0.01% Triton-X-100 at pH ~ 7.5). Error bars represent standard deviation for three independent replicates. C) SOFAST-HMQC spectra of TAR in the presence and absence of **1** (50 μM TAR, 200 μM DMA, 25 °C, buffer: 15 mM Na_2HPO_4/NaH_2PO_4, 15 mM NaCl, 0.1 mM EDTA at pH ~ 6.4). *C2H2 peaks were folded into the spectral window. D) Heat maps for WT-, UUCG- and Bulgeless-TAR. Colors represent the magnitude of chemical shift changes for spins C2H2, C6H6, or C8H8 on the given residue upon addition of 4× compound **1**: <0.025 ppm (gray); 0.025–0.05 (yellow); 0.05–0.1 ppm (light orange); >0.1 ppm (dark orange); and peak unassignable due to broadening or large chemical shift changes (red). E) Vector diagrams for select residues upon DMA binding to WT-TAR compared to DMA binding to UUCG-TAR (left) and Bulgeless-TAR (right).

displacement of Tat peptide induced by **1**, and control experiments included repeating the assay in the presence of 100-fold excess tRNA and measuring fluorescence background in the absence of TAR (Fig. 1B, page S107 of ESI‡). While **1** demonstrated moderate inhibition with a CD_{50} of 540 ± 171 μM, activity was not diminished in the presence of the tRNA control (CD_{50} = 542 ± 47 μM), which supported a selective **1**:TAR interaction as previously reported.[14] Furthermore, the activity of **1** was not diminished in the presence of a 30-mer (dA-dT) alternating DNA duplex (see page S112 and S113 of the ESI‡), which is consistent with the lack of intercalation observed by Waring and co-workers for C(5)-substituted amilorides[63] and highlights the potential selectivity of this scaffold.

The binding mode of **1** to TAR was then studied using the previously described 2D selective optimized flip angle short transient (SOFAST) [^{1}H-^{13}C] HMQC NMR method,[68] which allowed for the rapid assessment of small molecule-induced structural perturbations as observed by changes in ^{13}C and ^{1}H chemical shifts of signals from the aromatic residues on TAR (Fig. 1C). To allow ready visualization, we depicted the magnitude of each perturbation as a heat map overlay on the TAR secondary structure (Fig. 1D). Whereas low range chemical shift perturbations <0.025 ppm (grey) were considered insignificant, the measurably larger shifts ranging from 0.025–0.050 ppm (yellow), 0.05–0.1 ppm (light orange) and >0.1 ppm (dark orange) were surmised to originate from interaction of the small molecule with the RNA, reflecting direct ligand–RNA interactions and/or ligand-induced conformational changes. Lastly, peaks that could not be conclusively identified in the bound spectrum, either due to broadening or the magnitude of the shift, were labeled red. In the presence of **1**, significant perturbations were observed in residues 21–26, as well as G17, G36, C39, and C41, which we hypothesize to originate from the binding of **1** to both the bulge and apical loop regions of TAR.

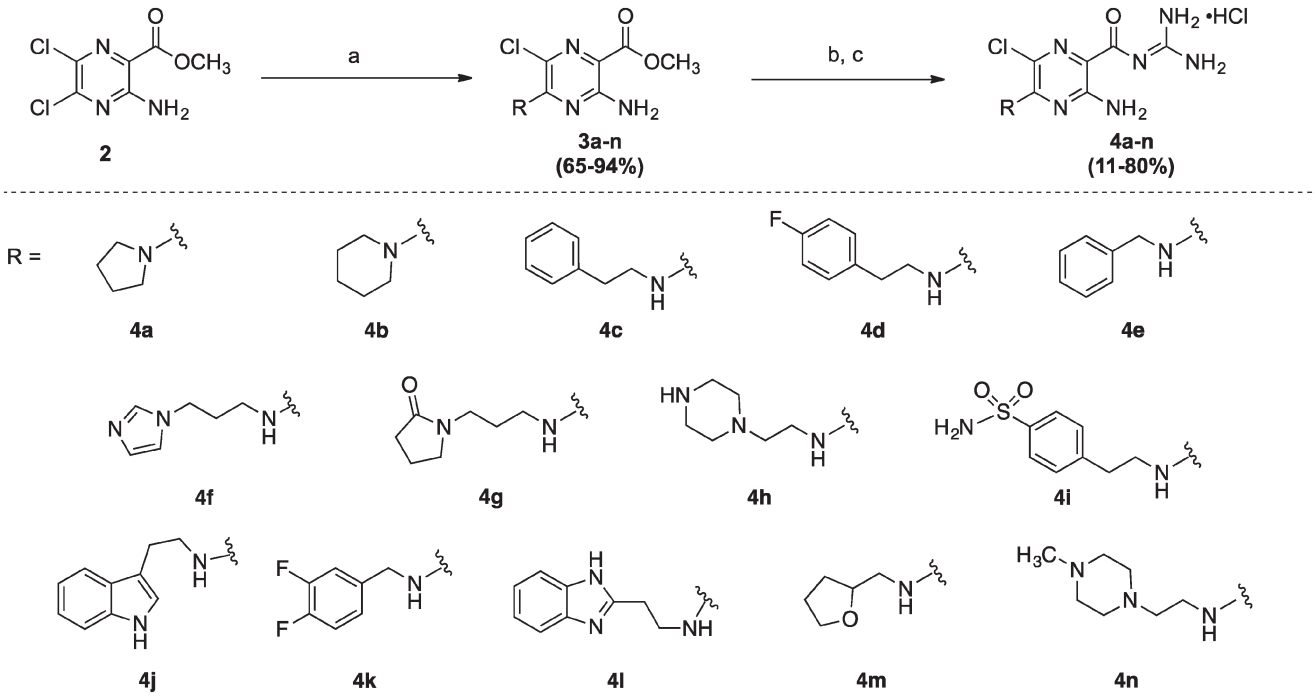
In the previous studies, traditional HSQC NMR chemical shift mapping experiments revealed that **1** induced significant perturbations to resonances belonging to the apical loop as well as in and around bulge residues while also demonstrating a non-linear dependence on concentration under these conditions (50 μM TAR, 15 mM Na_2HPO_4/NaH_2PO_4, 0.1 mM EDTA, 25 mM NaCl, 10% D_2O at pH ~ 6.4 and 25 °C).[14] In principle, the perturbations observed at the apical loop could have arisen due to conformational changes that are induced due to binding of **1** at the bulge or due to the presence of two discrete binding sites. To determine whether **1** does indeed bind the TAR apical loop as indicated by docking,[14] control experiments were performed with TAR mutants where residues of the apical loop or bulge were selectively modified. The UUCG-TAR RNA mutant featured a truncated and compact 4 nucleotide (UUCG) apical loop while maintaining the remaining bulge and stem structural features of WT-TAR, whereas the Bulgeless-TAR mutant replaced the three nucleotide bulge residues with a fully base paired stem while retaining the apical loop of the WT-TAR structure (Fig. 1D). In the compact loop UUCG-TAR mutant, significant perturbations were seen in the bulge residues, which are very similar in magnitude and direction to the perturbations observed in WT-TAR (Fig. 1D). Likewise, in the Bulgeless-TAR mutant, significant perturbations were seen in the apical loop residues, which were very similar in magnitude and direction to the perturbations observed in WT-TAR (Fig. 1D). The similarity of the chemical shift changes in the WT and mutant sequences were clearly seen when plotting the 2D chemical shift changes as vectors (Fig. 1E) and were consistent with two independent binding sites in WT-TAR, one localized in the apical loop and another in the bulge.[69] The small differences in direction or magnitude of the perturbations, particularly for G34(C8H8), may suggest that having both the bulge and apical loop has a subtle effect on the binding of **1** to either site. Accurate stoichiometry of binding, however, could not be gleaned from these experiments. Low solubility of **1** prevented the use of traditional techniques such as isothermal titration calorimetry, and, due to the low affinity of DMA, it was not possible to carry out NMR NOE experiments to more directly assess ligand binding sites.

The selective binding of **1** to TAR in the presence of tRNA and DNA confirmed the potential for amiloride as a promising RNA scaffold while the ability of **1** to bind both the bulge and apical loop implies that a range of motif specificities may be possible with further scaffold diversification. To further explore the utility of amiloride as an RNA selective scaffold, we pursued synthetic routes that allowed for diversification of the C(5)- and C(6)-positions. We chose a variety of motifs, including several found in reported RNA-binding ligands[38,39] and others enriched in aromatic, nitrogen-rich, and hydrogen bonding groups. We then synthesized 28 amiloride derivatives, purchased 6 commercial amilorides, and characterized the TAR binding properties of the combined library using the Tat displacement assay and NMR chemical shift mapping.

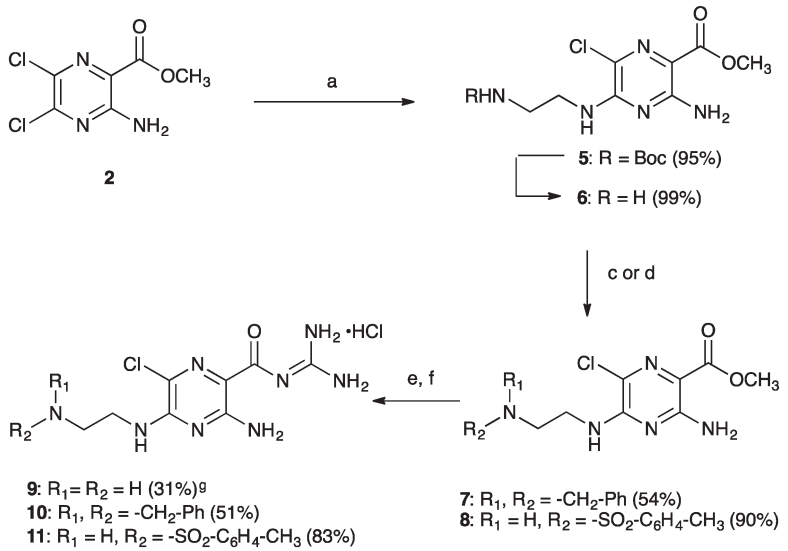
Synthetic routes toward a library of C(5)- and C(6)-substituted amiloride derivatives.

Amiloride derivatives have been previously synthesized and evaluated against a variety of biological targets.[65,66,70–73] Based on reported synthetic procedures for amilorides and other analogous compounds,[74–77] we designed and tested several synthetic routes to achieve diversity at the C(5)- and C(6)-positions (Schemes 1–5) of the core pyrazine ring. C(5)-derivatives were achieved through a concise two-step pathway (Scheme 1). Commercially available methyl 3-amino-5,6-dichloropyrazine-2-carboxylate **2** was substituted with various primary and secondary amines to obtain C(5)-substituted intermediates **3a–n** in moderate to high yields. Reactions of the methyl esters **3a–n** with guanidine at 65 °C yielded derivatives **4a–n**, which were isolated as the HCl salts. The acyl guanidine was installed at the final step to avoid challenges in purification.

To allow rapid expansion of the C(5)-derivative library outside of commercially available primary amines, we installed



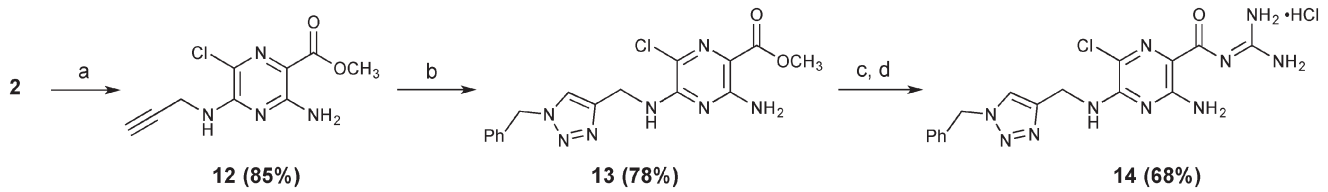
Scheme 1 Synthesis of C(5)-derivatives *via* designed two-step protocol. Reagents and conditions: a: amine, diisopropyl ethylamine (DIEA), DMF, rt, 18 h; b: guanidine (2 M solution in methanol), THF, 65 °C, 18 h; c: HCl (2 M in Et_2O), Et_2O.



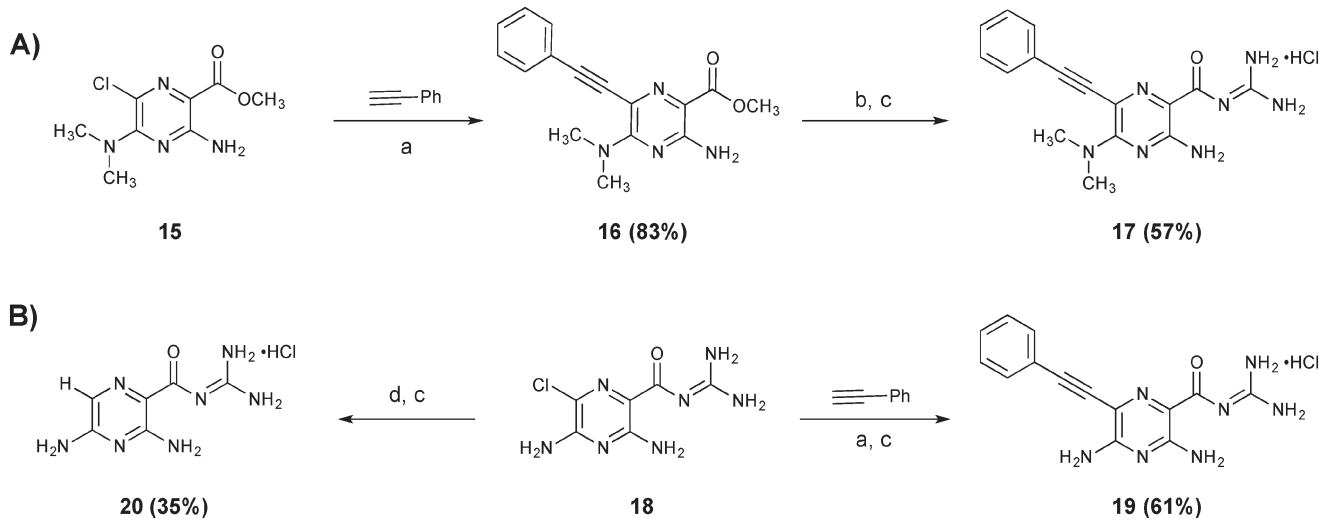
Scheme 2 Synthesis of ethylenediamine containing amiloride derivatives. Reagents and conditions: a: *N*-Boc-1,2-diaminoethane, DIEA, DMF, rt, 18 h; b: HCl (1 M) in MeOH, rt, 12 h; c: BnBr, DIEA, DMF, rt, 18 h; d: TsCl, DIEA, THF, rt, 18 h; e: guanidine 2 M solution in methanol, THF, 65 °C, 18 h; f: HCl (2 M in Et_2O), Et_2O; g: synthesized by guanidinylation of **6**.

an ethyleneamine linker (Scheme 2). Intermediate **6** can undergo a wide range of reactions, including functionalization with benzyl or tosyl groups as achieved here to yield intermediates **7** and **8**, respectively. In addition to the three derivatives **9**–**11** tested in this study, it is expected that ethylenediamine derivative **6** will open several possible routes to future diversification.

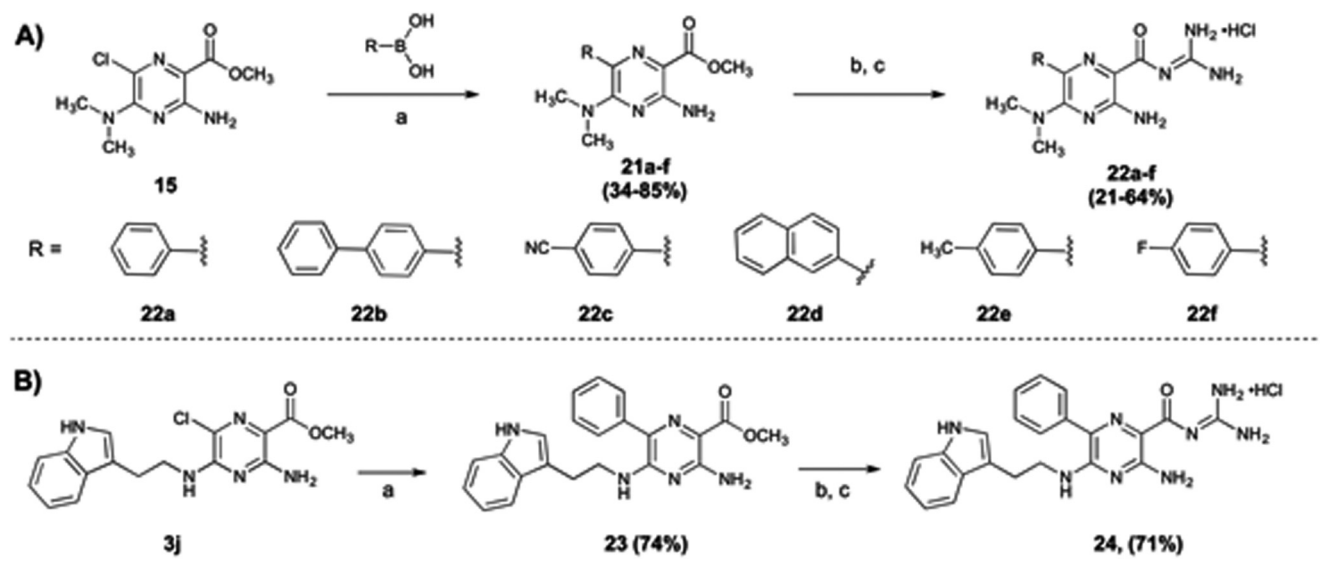
To test the possibility of using click-type reactions for further diversification, we installed an alkyne through reaction with propargyl amine (Scheme 3). Reaction of intermediate **12** with benzyl azide yielded the 1,2,3-triazole containing methyl ester derivative **13**, which was subsequently guanidinylated to yield product **14**. As with derivative **6**, terminal alkyne derivative **12** opens up a variety of possibilities



Scheme 3 Synthesis of 1,2,3-triazole containing derivative **14**. Reagents and conditions: a: propargylamine, DIEA, DMF, rt, 18 h; b: benzyl azide, $CuSO_4$·HCl, sodium ascorbate, iPrOH : water : CH_2Cl_2 (1 : 1 : 1), rt, 18 h; c: guanidine 2 M solution in methanol, THF, 65 °C, 18 h; d: HCl (2 M in Et_2O), Et_2O.



Scheme 4 (A) and (B): Sonogashira cross-coupling reactions and palladium catalyzed reduction for synthesis of C(6)-amiloride derivatives **17**, **19**, and **20**. Reagents and conditions: a: phenylacetylene, CuI (1 mol%), $Pd(PPh_3)_2Cl_2$ (2.5 mol%), triethylamine (5 eq.), DMF, 60 °C, 18 h; b: guanidine (2 M in methanol), THF, 65 °C, 18 h; c: HCl (2 M in Et_2O), Et_2O; d: H_2, 10% Pd–C (20 mol%), methanol, rt, 48 h.



Scheme 5 (A) Suzuki cross-coupling reactions of intermediate **15** for synthesis of C(6)-derivatives **22a–f**. (B): Synthesis of C(5)–C(6) dual substituted derivative **24**. Reagents and conditions: a: boronic acid (1.25 eq.), $Pd(PPh_3)_4$ (5 mol%), Na_2CO_3 (5 eq.), THF : water (1 : 1), 60 °C, 18 h; b: guanidine (2 M in methanol), THF, 65 °C, 18 h; c: HCl (2 M in Et_2O), Et_2O.

for diversification through coupling to various commercially available azides, in addition to the derivative evaluated here.

While synthesis of C(6)-amiloride derivatives had been previously described through a chloride reduction –

bromination/iodination procedure,[65] we achieved direct palladium-catalyzed cross coupling reactions on C(6)-chloroamiloride derivatives (Schemes 4 and 5). Sonogashira cross-coupling reaction was first performed on intermediate **15** (Scheme 4A), which was synthesized by the reaction of *N*,*N*-dimethylamine with intermediate **2**. Guanidinylation of intermediate **16** yielded derivative **17** (Scheme 4A). Commercially available amiloride **18**, where the acyl guanidine residue is already present, was directly converted to two derivatives (Scheme 4B): derivative **19** (right) through cross-coupling as above, and C(6)-reduced derivative **20** (left) *via* a palladium-catalyzed hydrogenation.

Suzuki cross-coupling reactions were next performed with intermediate **15** and a series of commercially available boronic acids to generate derivatives **22a–f** (Scheme 5A). Finally, in order to test if functionalization of both C(5)- and C(6)-positions yielded more potent binders, intermediate **3j**, featuring a tryptamine side-chain at the C5 position, was subject to the Suzuki cross-coupling reaction with benzeneboronic acid to yield bis-functionalized amiloride derivative **24** (Scheme 5B). The combination of these functionalization methods not only allowed for the generation of 28 diverse amiloride derivatives but also established a combination of efficient synthetic pathways that will allow rapid expansion of amiloride libraries in future studies.

Screening of amiloride derivatives using a displacement assay with fluorescently labeled Tat peptide

The 28-member synthetic amiloride library along with the 6 commercially available derivatives were screened for activity using the Tat displacement assay previously described.[67] While all of the commercially available derivatives (**S1–S6**, Fig. S1‡) were found to be insoluble or interfere with the assay signal, several of the synthetic derivatives yielded promising results (Table 1).

The observed displacement activities amongst synthetic amiloride derivatives can be separated into four groups: A) clear displacement of the Tat peptide in the presence of tRNA, *i.e.* selective binding; B) clear displacement of the Tat peptide only in the *absence* of tRNA, *i.e.* non-selective binding; C) significant signal interference, presumably caused by the fluorescence emission of the small molecule or aggregation; or D) no displacement. Group A is composed of 10 derivatives that showed mild to good displacement ($CD_{50} \sim 4$–200 μM) and were selective for TAR in presence of tRNA as seen from their minimally changed CD_{50} values in the presence of 100-fold excess tRNA (Fig. 2, left). Notably, each of the selective ligands features aromatic substituents. While only one of the C(5)-derivatives, **4j**, fell into this category, 9 of the 10 synthesized C(6)-derivatives, including **22a**, displayed selective interactions. These results imply that C(6)-functionalized amilorides represent a privileged class of ligands for targeting TAR. The differences in C(5)- and C(6)-derivatives likely stem from increased rigidity of the ligand and/or the presence of extended conjugation. Gratifyingly, combining the C(5)-indole motif of **4j** and C(6)-phenyl motif of **22a** to design **24** led to the most active of all small molecules tested ($CD_{50} = 4.4 \pm 1.4$ μM). The 7 non-selective ligands (Group B, Fig. 2, right) include most of the amilorides with substituents that are expected to be positively charged at pH 7.4 (**4h**, **9**, **4n**), suggesting that their binding events involve nonspecific electrostatic interactions with the RNA backbone. Several ligands with neutral alkyl and aromatic substituents, however, are also included in this non-selective group. Unfortunately, a subset of derivatives (Group C) displayed strong background fluorescence in the presence of the Tat peptide alone, presumably due to fluorescence of the small molecule itself or through aggregation. Finally, **4c**, **4g**, **4i**, and **14** showed no displacement of the Tat peptide (Group D). Though no single trend was identified for this small group, it is worth noting that **4g** and **4i** contain electronegative functional groups, specifically an amide and sulfonamide, which may be repulsed by the RNA phosphate backbone. The very strong interaction of the Tat peptide with TAR RNA under these conditions prevents calculation of a dissociation constant for the TAR:Tat interaction, which in turn prohibits the determination of discrete dissociation constants for amiloride derivatives.

Although the displacement assay results did not point to definitive correlations between the structural properties of amiloride derivatives and their observed displacement activity, possibly due to the many accessible binding conformations of TAR, some broad conclusions could be derived. To begin, all synthetic derivatives that demonstrated measurable activity in the displacement assay were more active than the parent compound **1**. As expected, polar positively charged subunits at the C(5)-position improved the displacement activity but did so nonspecifically as binding was partially or completely abrogated in the presence of the tRNA control. Both aryl and alkynyl groups at the C(6)-position showed improved affinity and selectivity compared to derivatives bearing a C(6)-chloro or C(6)-hydrogen substituent. The high hit rate of Tat peptide displacement observed with amiloride derivatives (18 out of 34, 53%) supports amiloride as an effective RNA-binding scaffold while the decreased activity with tRNA observed for 8 of the 18 active ligands highlights the inherent difficulties of developing *selective* ligands for RNA structures. We further tested the tRNA-selective amiloride analogues against DNA and found no loss in activity, similar to that observed with analogue **1** (see pages S112 and S113 of the ESI‡).

Evaluation of binding preferences *via* SOFAST-[^{1}H-^{13}C] HMQC NMR and docking

To evaluate the individual binding modes of the synthesized amiloride derivatives to HIV-1 TAR, we employed the SOFAST-[^{1}H-^{13}C] HMQC NMR experiment. In general, the perturbations induced by the derivatives were unique when compared to other known TAR binders.[14] This result suggests a unique overall mode of binding, though there were significant

Table 1 Tat peptide displacement screening and NMR assay results for synthetic amiloride derivatives

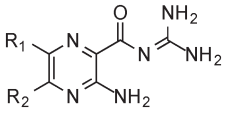
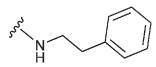
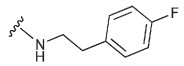
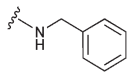
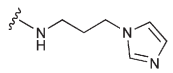
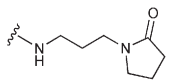
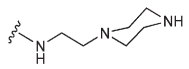
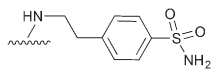
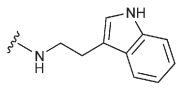
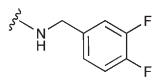
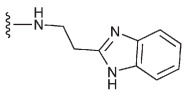
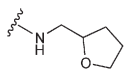
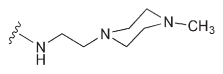
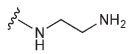
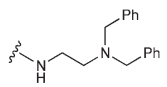
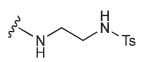
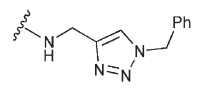
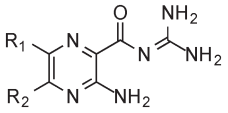
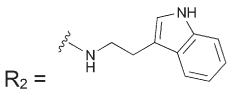
Amiloride		Group[a]	CD_{50}[b] Tat disp. (μM)	CD_{50} Tat disp. w/ tRNA[c] (μM)	NMR assay results[i]
C(5)-substitutions (R_1 = Cl)					
4a	R_2 = 	B	12 ± 4.5	No disp.[e]	Binds
4b	R_2 = 	C	Interference[d]	N.A[f]	Binds
4c	R_2 = 	D	No disp.[d]	N.A[f]	Binds
4d	R_2 = 	B	74 ± 26	No disp.[e]	Binds
4e	R_2 = 	C	Interference[d]	N.A[f]	Binds
4f	R_2 = 	C	Interference[d]	N.A[f]	Binds
4g	R_2 = 	D	No disp.[e]	No disp.[e]	No binding
4h	R_2 = 	B	120 ± 30	No disp.[e]	Binds
4i	R_2 = 	D	No disp.[e]	No disp.[e]	No binding
4j	R_2 = 	A	33 ± 12	42 ± 17	Binds
4k	R_2 = 	C	Interference[d]	N.A[f]	Binds
4l	R_2 = 	B	47 ± 23	No disp.[e]	Binds
4m	R_2 = 	B	143 ± 19	No disp.[e]	Binds
4n	R_2 = 	C	Interference[d]	No disp.[e]	Binds[h]
9	R_2 = 	B	33 ± 5.3	No disp.[e]	Binds
10	R_2 = 	C	Interference[d]	N.A[f]	Binds
11	R_2 = 	C	Interference[d]	N.A[f]	Binds
14	R_2 = 	D	No disp.[e]	No disp.[e]	No binding

Table 1 *(continued)*

Amiloride		Group[a]	CD_{50}[b] Tat disp. (μM)	CD_{50} Tat disp. w/ tRNA[c] (μM)	NMR assay results[i]
C(6)-substitutions					
17	R_1 = C≡C-Ph R_2 = $N(CH_3)_2$	A	54 ± 17	86 ± 46	Binds
19	R_1 = C≡C-Ph R_2 = NH_2	A	40 ± 19	48 ± 18	Binds
20	R_1 = H R_2 = NH_2	B	730 ± 223[g]	No disp.[e]	Binds
22a	R_1 = Ph R_2 = $N(CH_3)_2$	A	192 ± 163	124 ± 43	Binds
22b	R_1 = (*p*-C_6H_4)-Ph R_2 = $N(CH_3)_2$	A	—[h]	18 ± 3.6	Binds
22c	R_1 = (*p*-C_6H_4)-CN R_2 = $N(CH_3)_2$	A	101 ± 40	31 ± 4.6	Binds
22d	R_1 = β-napthyl R_2 = $N(CH_3)_2$	A	32 ± 12	29 ± 8.2	Binds
22e	R_1 = (*p*-C_6H_4)-CH_3 R_2 = $N(CH_3)_2$	A	—[h]	57 ± 8.9	Binds
22f	R_1 = (*p*-C_6H_4)-F R_2 = $N(CH_3)_2$	A	60 ± 18	64 ± 8.7	Binds
C(5)- and C(6)-substitutions					
24	R_1 = Ph R_2 = 	A	4.4 ± 1.4	13 ± 2.5	Binds

[a] See text for description of groups. [b] CD_{50}: competitive dosage required for displacement of 50% of Tat peptide from preformed TAR:Tat complex, as measured by fluorescence of FAM- and TAMRA-labeled Tat peptide at 590 nm, the emission λ_{max} of TAMRA. Errors are standard deviation for three independent replicates. [c] 100× tRNA (brewer's yeast, Roche) used as control to test selectivity of amiloride derivatives towards TAR. [d] Interference originating from the fluorescent small molecule or from aggregation of the assay mixture at 590 nm. [e] No displacement of Tat peptide observed. [f] Not applicable. [g] Curve fitting is poor due to weak binding of 20. Mean and standard deviation calculated from two CD_{50} values. [h] **22b** and **22e** curves could only be accurately fit in the presence of tRNA, possibly due to differences in the background (no TAR added) signal. [i] As seen by SOFAST NMR experiments explained below. Some derivatives that did not measurably displace Tat did demonstrate binding by NMR.

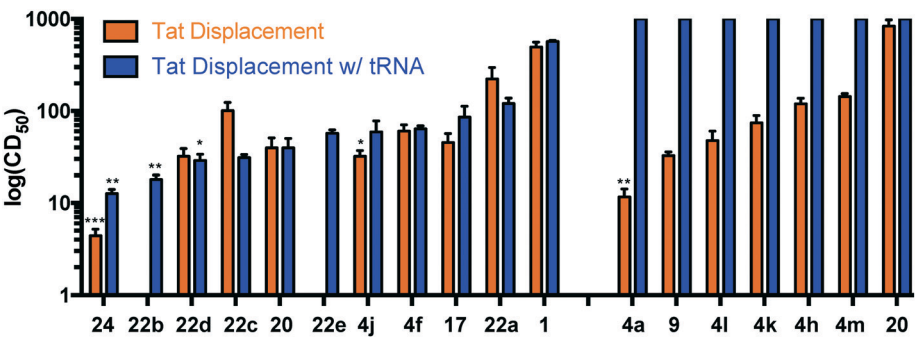


Fig. 2 CD_{50} values determined for all active amiloride derivatives in the Tat-displacement assay (50 mM Tris-HCl, 50 mM KCl, 0.01% Triton-X-100 at pH ~ 7.5) using WT-TAR alone (orange) and with 100-fold excess tRNA (blue). Error bars represent standard deviation for three independent replicates. Ligands on the right did not demonstrate measurable activity in the presence of tRNA. Note: CD_{50} values could not be determined for **22b** and **22e** in the absence of tRNA (see Table 1).

variations among the different amiloride derivatives as well. Similar to the study of **1** shown in Fig. 1, Fig. 3 shows residue-specific heat maps for NMR chemical shift perturbations caused by our synthetic amiloride derivatives.

As previously discussed, the perturbations in the NMR chemical shifts caused by amiloride derivatives measured *via* SOFAST NMR provided a fast and direct method to lend insight into the site-selectivity of binding interactions.

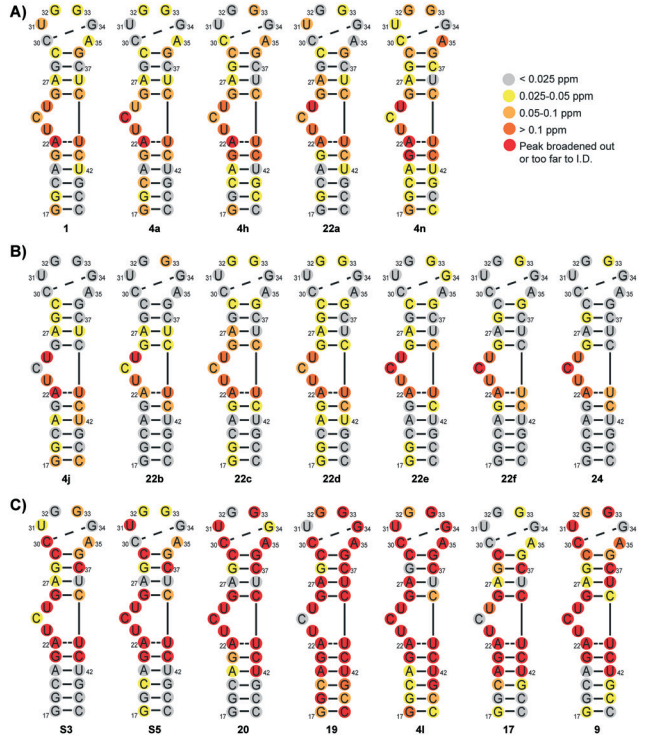


Fig. 3 Representative chemical shift maps reflecting the interaction of synthetic amiloride derivatives with HIV-1 TAR using the SOFAST-[^{1}H-^{13}C] HMQC NMR experiment. Results represented as heat maps showing the extent of perturbations to the TAR C8H8/C6H6/C2H2 chemical shifts caused by amiloride derivatives. A) Perturbations to the apical loop and bulge residues; B) perturbations primarily to the bulge residues; C) broadening of signals. Spectra and remaining chemical shift maps can be found on pages S114–S122 of the ESI.‡

Compared to standard NMR HSQC experiments, the SOFAST method allowed for rapid (<1.5 h per ligand) screening of all derivatives using low (50 μM) concentrations of RNA.[68] As seen from Fig. 3 (also see Fig. S3 in ESI‡), the amiloride derivatives show significant variation in the degree and sites of signal perturbation, which is represented with respective colors in the heat maps. Initial comparison of the heat maps revealed that a small subset of derivatives (Fig. 3A) perturbed signals in both the apical loop and bulge region, similar to those observed with **1**. A much larger subset of the library members, however, perturbed signals almost exclusively in the bulge region (Fig. 3B, Fig. S3‡). Notably, no perturbations in any nucleotides of TAR were observed upon addition of derivatives **4g**, **4i**, and **14**, which was in alignment with the negative results in the Tat displacement assay. On the other hand, **4c** displayed no significant activity in the displacement assay but caused significant NMR chemical shift perturbations, presumably due to the high sensitivity of the NMR experiment, which can reveal weak binding interactions. A binding mode that does not displace the Tat peptide, however, cannot be ruled out.

Importantly, synthetic derivatives **4b**, **4e**, **4k**, and **10** along with commercial derivatives **S1–S6**, which all demonstrated interference in the displacement assay, are identified as TAR binding derivatives by NMR, confirming the importance of multiple analytical methods. **S1**, **S3**, **S6**, **20**, **19**, **4l**, **17**, and **9** caused significant broadening of resonances and disappearance of several peaks (Fig. S2‡). This widespread broadening may originate from strong interactions of the small molecule with the RNA that lead to exchange at the micro-to-millisecond timescale and/or from other bulk events such as aggregation. If broadening was due to aggregation and consequently lower overall tumbling rates, we would expect all residues to experience similar extents of broadening. However, we found that the loss of intensity was much more significant for specific residues, largely in and around the bulge, compared to residues in the helical regions (Fig. S2‡). The non-uniform nature of the broadening suggests that it arose due to intermediate exchange at the micro-to-millisecond timescale between free and small molecule bound TAR states, possibly due to tighter binding and slower off rates as compared to the parent compound **1**. However, we could not rule out that the broadening was partially due to small molecule induced aggregation. To examine this latter possibility, we measured the dynamic light scattering (DLS) of the amiloride derivatives alone as well as in the presence of TAR and found that only **9** demonstrated a significant increase in particle size over background (see Fig. S8‡). Still we cannot rule out transient formation of larger aggregates that are invisible to these DLS experiments and that contribute to the observed broadening in the NMR experiments.

For a more in-depth analysis of the ligand-induced spectral changes, vector diagrams were generated that depict both the magnitude and direction of the chemical shift perturbations (Fig. 4, Fig. S4–S7‡). To begin, amiloride derivatives that caused strong perturbations in the apical loop residues (*e.g.* **4a**, **4h**, **22a**, and **4n**) generally caused similar chemical shift changes as observed with **1** (Fig. 4). On the other hand, most derivatives perturbed the bulge residue signals with the greatest magnitude and a subset of derivatives exhibit striking differences in direction, particularly the C(6)-derivatives (Fig. 4). For example, both **22f** and **24** caused perturbations at bulge residues U25(C6H6) and G26(C8H8) in directions similar to each other but distinct relative to other ligands, consistent with a specific interaction in this region. **22d**, which contains a large naphthyl group at the C(6)-position, also caused perturbations at bulge residues A22(C2H2), U25(C6H6), and G26(C8H8) but in unique directions. Interestingly, the perturbations in the helical regions of TAR, such as A27(C8H8), displayed lower magnitudes but the greatest variation in direction among the ligands. This latter effect likely reflects the subtle differences in the TAR conformations stabilized by each amiloride derivative, rather than specific interactions with the ligands. Taken together, the differences in perturbations among amiloride derivatives, particularly those closely related such as **22a**, **22e** and **22f**, which differ by a Ph-H, Ph-CH_3, and Ph-F substitution, respectively, suggest that subtle changes in the shape of the small molecule may significantly impact the bound conformation of TAR.

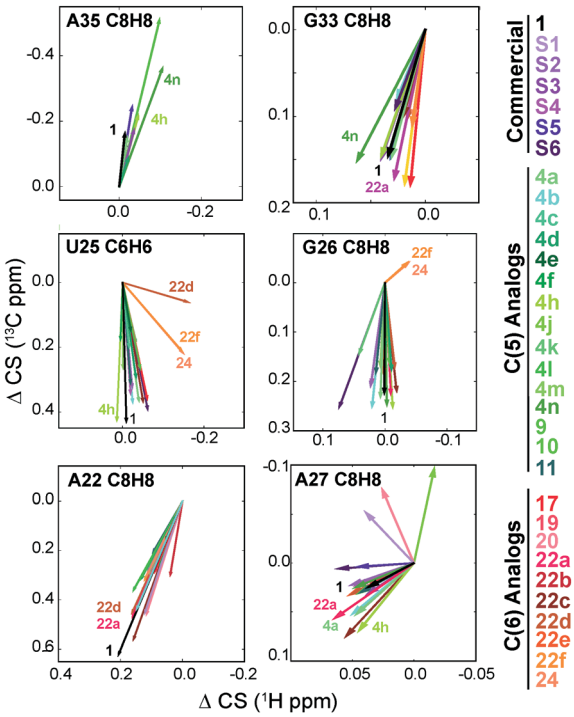


Fig. 4 Representative vector diagrams indicating the direction and magnitude of chemical shift perturbations induced by each synthetic amiloride derivative binding to HIV-1 TAR measured using the SOFAST-[^{1}H-^{13}C] HMQC NMR experiment. Perturbations less than 0.025 ppm were considered negligible and excluded. Perturbations that could not be assigned due to large chemical shifts and/or broadening are listed in Fig. S4–S7,‡ as are the vector diagrams for all other residues.

The ability of the amiloride derivatives to adopt multiple binding modes was reflected in docking experiments of the synthesized library against the TAR ensemble previously reported by the Al-Hashimi lab,[78] and using the docking program, ICM (Molsoft).[79] Each derivative was docked against each of the 20 TAR conformers (see page S124 of ESI‡). The docking energies for each complex, as measured by ICM's interaction score, were used to predict the population of each TAR conformer when bound to the amiloride derivatives assuming a Boltzmann distribution (see page S124 of ESI‡). We examined the population of each TAR conformer over all amiloride derivatives and found that four of the 20 conformers were strongly preferred for amiloride: TAR interactions (Fig. 5). These docking results confirm the NMR-based observation that both the bulge and loop are potential ligand binding sites for different amiloride derivatives (Fig. 3 and 4). Lead derivative **24** was the strongest predicted binder with an interaction score (−51 kcal mol^{-1}) three standard deviations above the mean (mean score = −34 ± 5 kcal mol^{-1}, see Table S2‡), and exclusively demonstrated bulge binding interactions (Fig. 5). This result is consistent with the bulge-localized NMR chemical shift perturbations measured for **24**. Derivative **22d** yielded distinct NMR perturbations (Fig. 4) relative to the other derivatives and is the only derivative predicted to have strong binding to Conformer 14. It is important to note, however, that the predicted binding analysis did not show other trends that could be clearly correlated with the NMR or CD_{50} data. Remarkably, the fact that docking often resulted in more than one favorable binding mode for a single molecule as well as strikingly different binding modes for very similar molecules (*i.e.* **22a**, **22e**, and **22f**) reinforces the previously described lack of clearly discernible structure–activity relationships with amiloride derivatives. Interestingly, a deeper analysis of the terms in the scoring function revealed that the selective binders (Group A) predicted higher contributions from van der Waals forces while the non-selective binders (Group B) predicted higher contributions from hydrogen bonding (Fig. S9‡), suggesting a possible variation in the binding modes of the two sets of ligands.

Cheminformatic analysis of amiloride library

In order to determine chemical features associated with each of the three binding classes, (*i.e.* specific binding, non-specific binding, and non-binding), we evaluated our amiloride library using cheminformatic analyses. We first performed a preliminary evaluation of library diversity[80–84] by calculating the Tanimoto coefficient (T_c) values for all possible pairs of library members (see page S126 of ESI‡).[85,86] Interestingly, only two library members (**4a** and **9**) were predicted to be similar to **1** ($T_c > 0.85$) (Fig. 6A).[87] Despite the presence of the amiloride scaffold in every molecule, only 8% of the total possible pairings of synthetic derivatives were found to be similar, and nearly half of these were within the aryl-C(6)-derivative class.

We further sought to evaluate whether standard cheminformatic parameters[83,84] may be able to explain the binding patterns within our synthetic derivative library. Given the difference in size between the group that bound by NMR (N = 25) and the group that showed no binding (N = 3), statistical comparisons between these groups were not possible. Qualitatively, however, the binding group displayed a mean decrease in the hydrogen bond acceptor number (HBA), total polarizable surface area (tPSA) and rotatable bonds (RotB) relative to the non-binding ligands, along with a decreased number of oxygens and nitrogens (Fig. 6B, Table S4‡). Statistical analysis was performed between the groups whose CD_{50} values were consistent with selective (N = 10) or non-selective (N = 7) binding interactions. Group analysis, rather than direct correlation of cheminformatic parameters to CD_{50} values, was performed due to the narrow range of CD_{50} values recorded. The library members that displayed interference in the displacement assay were removed from this analysis. As previously discussed, of the four library members with positively charged side chains at pH 7.4, two were in the non-selective binding group and two interfered with the assay. All selective binders in this library had neutral substituents. The most strikingly different cheminformatic parameter between the two groups was the predicted distribution coefficient

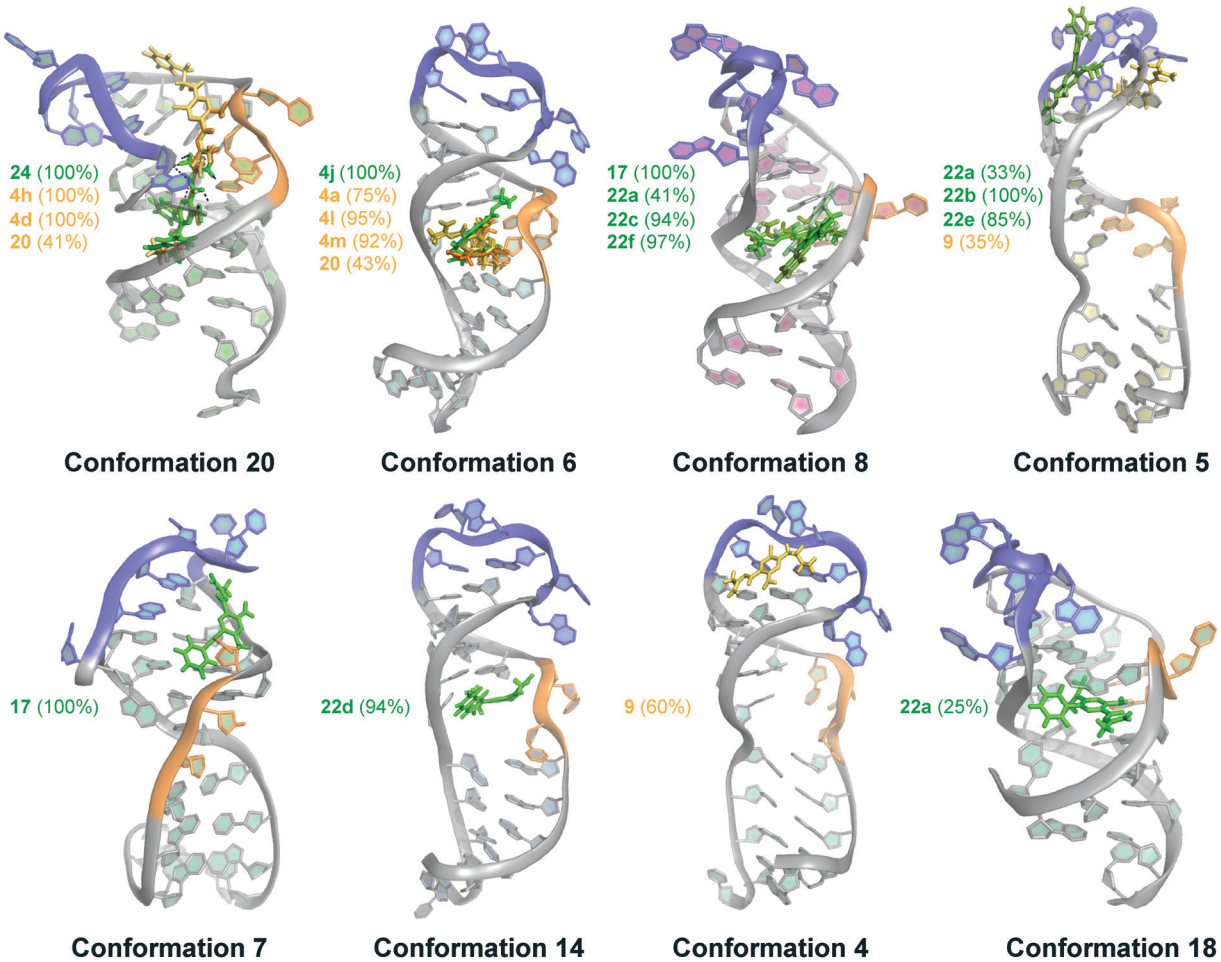
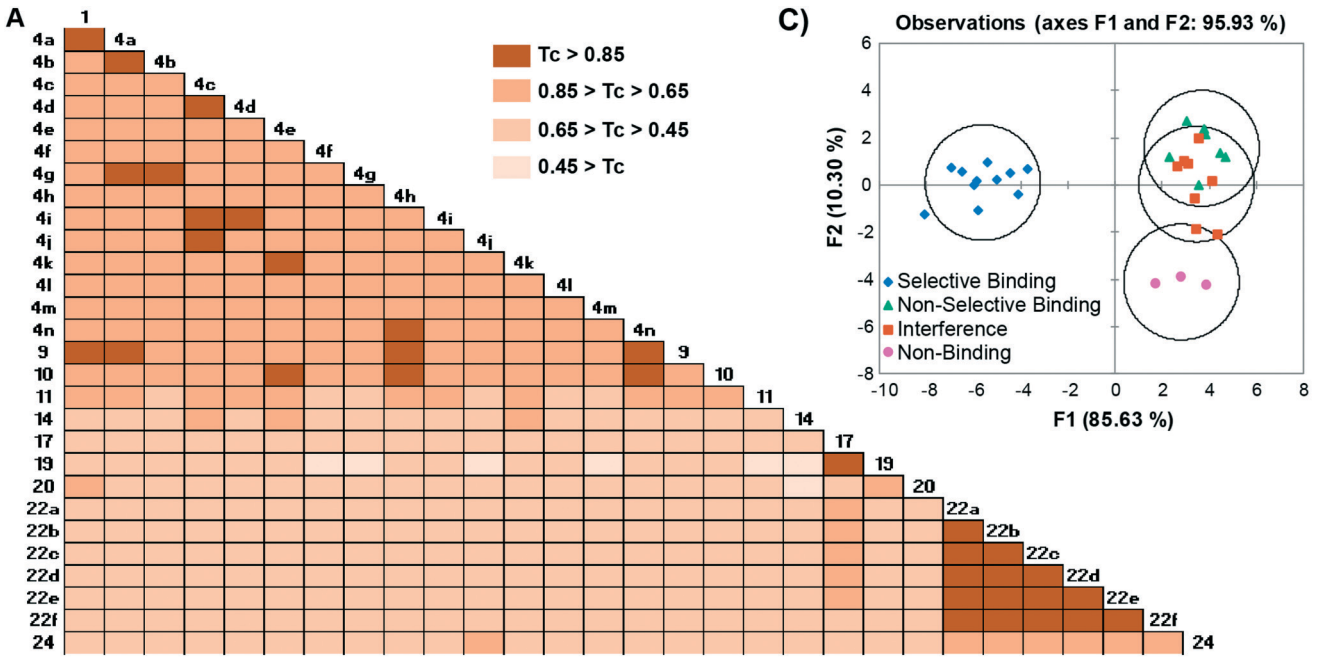


Fig. 5 Docking predicted bound poses of amiloride derivatives to their preferred conformers of the HIV-1 TAR ensemble. The TAR loop is in blue, the bulge in orange, and the stem in gray, while the ligands are shown in green (Group A) or yellow (Group B). Conformers are ordered based on the docking predicted population of each, averaged over all amilorides, with the most strongly preferred conformer (20) listed first. The listed percentages indicate the docking predicted population of each ligand:TAR complex for the *individual* ligands.

(log *D*), a pH-dependent partition coefficient, where the non-selective binders are predicted to be more hydrophilic (mean log *D* = −0.78) while the selective binders more lipophilic (mean log *D* = 1.75). The selective binders also displayed the lowest fraction of sp3 hybridized carbons (Fsp3), hydrogen bond acceptor number (HBA), hydrogen bond donor number (HBD), and relative polar surface area (relPSA) of the three classes. These analyses, along with the docking results, are consistent with stacking and other van der Waals interactions, rather than electrostatics and hydrogen bonding, as the driving forces behind selective TAR binding among amiloride derivatives. Interestingly, all library members were found to follow Lipinski's rules of five[88] except for **4i**, **11**, and **14**, none of which displaced the Tat peptide, though it is not clear whether Lipinski's rules of five are applicable to RNA-targeted ligands. No library members were identified as PAINS compounds.[89]

To determine whether a combination of parameters might be used to predict selectivity among the amiloride derivatives, we performed linear discriminate analysis (LDA) of the three binding groups and found that 99.9% of the variance could be explained with a single linear discriminant that weighted HBA and tPSA the most heavily, followed by O, N, RotB, log *D*, and HBD (see pages S129–S130 of the ESI‡). While excellent separation was observed (estimation sample 100%), leave-one-out cross-validation analysis returned correct group assignments with 65% accuracy, suggesting that larger sample sizes will be needed for these calculations to be predictive. When the fourth group, *i.e.* those derivatives that interfered with the displacement assay but bound by NMR, was added to the LDA analysis, this group was found to overlap with both the non-selective and non-binding groups (Fig. 6C). This result may imply that these derivatives were largely weak binders.

Finally, we further sought to compare the diversified scaffold design to fragment-based screening by examining the difference between **1**, **4j**, and **24**. **4j** demonstrated an order of magnitude decrease in CD_{50} value as well as more resolved NMR chemical shifts as compared to **1** and differs in the addition of an indole at the C(5)-position. *N*-Acetyltryptamine,



B)

	MW	logD	logP	HBA	HBD	RotB	tPSA	Fsp3	rel PSA	N	O	Rings	Ar Rings	Sys Rings	VWSA
All DMA derivatives (N = 28)	**341.82**	**0.64**	**0.99**	**8.46**	**4.00**	**4.36**	**155.19**	**0.20**	**0.36**	**7.68**	**1.21**	**2.21**	**2.00**	**2.07**	**446.18**
Std Dev	53.16	1.56	1.16	0.69	0.82	1.66	17.81	0.15	0.08	0.86	0.57	0.63	0.82	0.47	74.79
Subclasses															
No Binding (N = 3)	389.49	-0.01	0.40	9.67	4.33	6.00	182.36	0.24	0.37	8.67	2.00	2.33	2.00	2.33	493.06
Total Binding (N = 25)	336.09	0.72	1.06	8.32	3.96	4.16	151.93	0.20	0.36	7.56	1.12	2.20	2.00	2.04	440.56
Non-Selective Binding (N = 7)	304.96	-0.78	1.18	8.43	4.29	3.71	158.24	0.27	0.43	7.71	1.14	1.86	1.43	1.71	386.84
Selective Binding (N = 10)	338.52	1.75	0.78	8.10	3.50	3.70	146.08	0.12	0.32	7.30	1.00	2.50	2.50	2.20	457.55
		**						*	*				*	*	*

Fig. 6 A) Heat map of Tanimoto coefficients for synthesized derivatives; B) representative cheminformatic parameters among binding classes (**p-value < 0.01, *p-value < 0.05 between non-selective and selective binding classes); C) linear discriminant analysis (LDA) plot of cheminformatic-based clusters of binding classes.

which mimics the C(5)-side chain in **4j**, was tested in the displacement assay as well as the SOFAST NMR experiments and showed no displacement of the Tat peptide or perturbations to the TAR NMR signals, suggesting that the subunit may not have been identified as a hit in a fragment based screen. The simultaneous combination of the amiloride and tryptamine fragment can thus be concluded to be important for binding activity of **4j**. Expansion of **4j** in **24** led to an additional order of magnitude decrease in CD_{50} value, consistent with increased molecular complexity,[90] which may prove an important feature in small molecule:RNA recognition.

Conclusions and future directions

In summary, we have used scaffold diversification to generate a small RNA-targeted library that establishes amiloride as a new RNA-binding scaffold capable of highly specific interactions. Expansion of this scaffold led to the discovery of 10 novel and selective ligands for HIV-1 TAR RNA that demonstrated increased or similar affinity relative to other small molecule ligands.[9,17] Furthermore, use of the recently developed SOFAST NMR method allowed us to directly evaluate the binding mode of amiloride derivatives with TAR at low concentration in a time efficient assay. Chemical shift mapping demonstrates that while the parent ligand **1** bound both the apical loop and bulge, modifications at the C(5)- and particularly at the C(6)-positions allowed this binding to be tuned to tight and specific binding in the bulge region. More detailed docking studies are underway to pursue correlations with the NMR perturbation data, which would prove a powerfully predictive docking technique and, along with the established synthetic methods, allow for the tailoring of ligands to bind specific TAR conformations, including those that favor the apical loop. Cheminformatic analysis suggested that stacking and other van der Waals interactions are the most critical for selective binding of amilorides to TAR RNA and that it may be possible to predict this selectivity in future library designs. Finally, the success of the combined C(5)- and C(6)-modifications in **24** to yield a ligand with 100× improved activity relative to the parent molecule and a unique binding mode strongly supports the pursuit of combinatorial library designs on modifiable RNA scaffolds. Indeed, **24** is one of the tightest non-aminoglycoside TAR ligands reported to date, and future work will investigate the combination of

even more potent C(6)-derivatives. Furthermore, the use of multiple experimental and computational methods allowed specific insight into the molecular recognition of RNA by these small molecules without the determination of a high-resolution structure. Future studies will include explorations of multivalent amiloride-based probes such as dimers and conjugates with other RNA active molecules, as these strategies have been shown to enhance activity in some cases.[24,27,91] We hope that the tools and insights developed in this study will allow the development of much-needed guiding principles for small molecule:RNA interactions.

Notes

The authors declare no competing financial interests.

Abbreviations used

TAR	Trans-activation response element
DMA	Dimethylamiloride
CD_{50}	Competitive displacement of 50% Tat peptide
tRNA	Transfer RNA
SOFAST	Selective optimized flip angle short transient
HMQC	Heteronuclear multiple-bond quantum correlation
HSQC	Heteronuclear single-bond quantum correlation
ITC	Isothermal titration calorimetry
DLS	Dynamic light scattering
MW	Molecular weight
HBA	Hydrogen bond acceptor
HBD	Hydrogen bond donor
RotB	Rotatable bonds
tPSA	Topical polar surface area
relPSA	Relative polar surface area
LDA	Linear discriminant analysis

Acknowledgements

We thank the members of the Hargrove and Al-Hashimi labs for stimulating discussions. AEH and HMA wish to acknowledge financial support for this work from Duke University and the US National Institute of Health (P50GM103297). We also wish to thank: Dr. George R. Dubay, Director of Analytical Services, Duke University for performing the high resolution mass spectroscopy analyses; the Duke University NMR Center for use of the 800 MHz and 600 MHz spectrometers for 2D SOFAST NMR experiments and the 500 MHz and 400 MHz spectrometers for 1D analysis of small molecules; Prof. Michael Therien and Dr. Melanie O'Sullivan for assistance and access to dynamic light scattering instruments; Mr. David Jones, Ms. Dillan Casanova, and Ms. Aline Umuhire-Juru for assistance with synthesis of amiloride derivatives; and Mr. Raymond Li for help with the fluorescence assays.

References

1 E. P. Consortium, *Nature*, 2012, **489**, 57.
2 T. R. Cech and J. A. Steitz, *Cell*, 2014, **157**, 77.
3 I. Caldelari, Y. Chao, P. Romby and J. Vogel, *Cold Spring Harbor Perspect. Med.*, 2013, **3**, a010298.
4 A. Blond, E. Ennifar, C. Tisné and L. Micouin, *ChemMedChem*, 2014, **9**, 1982.
5 S. F. J. Le Grice, *Curr. Top. Microbiol. Immunol.*, 2015, 147.
6 C. M. Connelly, M. H. Moon and J. S. Schneekloth Jr, *Cell Chem. Biol.*, 2016, **23**, 1077.
7 T. Hermann, *Wiley Interdiscip. Rev.: RNA*, 2016, **7**, 726.
8 L. Guan and M. D. Disney, *ACS Chem. Biol.*, 2012, **7**, 73.
9 B. S. Morgan and A. E. Hargrove, in *Synthetic Receptors for Biomolecules: Design Principles and Applications*, The Royal Society of Chemistry, 2015, p. 253.
10 J. R. Thomas and P. J. Hergenrother, *Chem. Rev.*, 2008, **108**, 1171.
11 T. Hermann, *Curr. Opin. Struct. Biol.*, 2005, **15**, 355.
12 J. Poehlsgaard and S. Douthwaite, *Nat. Rev. Microbiol.*, 2005, **3**, 870.
13 K. F. Blount and R. R. Breaker, *Nat. Biotechnol.*, 2006, **24**, 1558.
14 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.
15 N. M. Bell, A. L'Hernault, P. Murat, J. E. Richards, A. M. L. Lever and S. Balasubramanian, *Biochemistry*, 2013, **52**, 9269.
16 N. A. Naryshkin, M. Weetall, A. Dakka, J. Narasimhan, X. Zhao, Z. Feng, K. K. Y. Ling, G. M. Karp, H. Qi, M. G. Woll, G. Chen, N. Zhang, V. Gabbeta, P. Vazirani, A. Bhattacharyya, B. Furia, N. Risher, J. Sheedy, R. Kong, J. Ma, A. Turpoff, C.-S. Lee, X. Zhang, Y.-C. Moon, P. Trifillis, E. M. Welch, J. M. Colacino, J. Babiak, N. G. Almstead, S. W. Peltz, L. A. Eng, K. S. Chen, J. L. Mull, M. S. Lynes, L. L. Rubin, P. Fontoura, L. Santarelli, D. Haehnke, K. D. McCarthy, R. Schmucki, M. Ebeling, M. Sivaramakrishnan, C.-P. Ko, S. V. Paushkin, H. Ratni, I. Gerlach, A. Ghosh and F. Metzger, *Science*, 2014, **345**, 688.
17 J. Sztuba-Solinska, S. R. Shenoy, P. Gareiss, L. R. H. Krumpe, S. F. J. Le Grice, B. R. O'Keefe and J. S. Schneekloth, *J. Am. Chem. Soc.*, 2014, **136**, 8402.
18 J. A. Howe, H. Wang, T. O. Fischmann, C. J. Balibar, L. Xiao, A. M. Galgoci, J. C. Malinverni, T. Mayhood, A. Villafania, A. Nahvi, N. Murgolo, C. M. Barbieri, P. A. Mann, D. Carr, E. Xia, P. Zuck, D. Riley, R. E. Painter, S. S. Walker, B. Sherborne, R. de Jesus, W. Pan, M. A. Plotkin, J. Wu, D. Rindgen, J. Cummings, C. G. Garlisi, R. Zhang, P. R. Sheth, C. J. Gill, H. Tang and T. Roemer, *Nature*, 2015, **526**, 672.
19 J. Palacino, S. E. Swalley, C. Song, A. K. Cheung, L. Shu, X. Zhang, M. Van Hoosear, Y. Shin, D. N. Chin, C. G. Keller, M. Beibel, N. A. Renaud, T. M. Smith, M. Salcius, X. Shi, M. Hild, R. Servais, M. Jain, L. Deng, C. Bullock, M. McLellan, S. Schuierer, L. Murphy, M. J. J. Blommers, C. Blaustein, F. Berenshteyn, A. Lacoste, J. R. Thomas, G. Roma, G. A. Michaud, B. S. Tseng, J. A. Porter, V. E. Myer, J. A. Tallarico, L. G. Hamann, D. Curtis, M. C. Fishman, W. F. Dietrich, N. A. Dales and R. Sivasankaran, *Nat. Chem. Biol.*, 2015, **11**, 511.
20 K. M. Felsenstein, L. B. Saunders, J. K. Simmons, E. Leon, D. R. Calabrese, S. Zhang, A. Michalowski, P. Gareiss, B. A. Mock and J. S. Schneekloth, *ACS Chem. Biol.*, 2016, **11**, 139.

21 H. Wang and Y. Tor, *J. Am. Chem. Soc.*, 1997, **119**, 8734.
22 K. F. Blount, F. Zhao, T. Hermann and Y. Tor, *J. Am. Chem. Soc.*, 2005, **127**, 9818.
23 S. R. Kirk and Y. Tor, *Bioorg. Med. Chem.*, 1999, **7**, 1979.
24 K. Michael, H. Wang and Y. Tor, *Bioorg. Med. Chem.*, 1999, **7**, 1361.
25 Y. Tor, T. Hermann and E. Westhof, *Chem. Biol.*, 1998, **5**, R277.
26 D. P. Arya, R. L. Coffee, B. Willis and A. I. Abramovitch, *J. Am. Chem. Soc.*, 2001, **123**, 5385.
27 S. Kumar and D. P. Arya, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 4788.
28 N. Ranjan, S. Kumar, D. Watkins, D. Wang, D. H. Appella and D. P. Arya, *Bioorg. Med. Chem. Lett.*, 2013, **23**, 5689.
29 S. Kumar, N. Ranjan, P. Kellish, C. Gong, D. Watkins and D. P. Arya, *Org. Biomol. Chem.*, 2016, **14**, 2052.
30 C. S. Eubanks, J. E. Forte, G. J. Kapral and A. E. Hargrove, *J. Am. Chem. Soc.*, 2017, **139**, 409.
31 T. Hermann, *Biochimie*, 2002, **84**, 869.
32 B. Davis, M. Afshar, G. Varani, A. I. H. Murchie, J. Karn, G. Lentzen, M. Drysdale, J. Bower, A. J. Potter, I. D. Starkey, T. Swarbrick and F. Aboul-ela, *J. Mol. Biol.*, 2004, **336**, 343.
33 A. I. H. Murchie, B. Davis, C. Isel, M. Afshar, M. J. Drysdale, J. Bower, A. J. Potter, I. D. Starkey, T. M. Swarbrick, S. Mirza, C. D. Prescott, P. Vaglio, F. Aboul-ela and J. Karn, *J. Mol. Biol.*, 2004, **336**, 625.
34 M. D. Disney, I. Yildirim and J. L. Childs-Disney, *Org. Biomol. Chem.*, 2014, **12**, 1029.
35 L. Nguyen, L. M. Luu, S. Peng, J. F. Serrano, H. Y. E. Chan and S. C. Zimmerman, *J. Am. Chem. Soc.*, 2015, **137**, 14180.
36 M. D. Shortridge and G. Varani, *Curr. Opin. Struct. Biol.*, 2015, **30**, 79.
37 F. A. Abulwerdi, M. D. Shortridge, J. Sztuba-Solinska, R. Wilson, S. F. J. Le Grice, G. Varani and J. S. Schneekloth, *J. Med. Chem.*, 2016, **59**, 11148.
38 J. L. Childs-Disney, J. Hoskins, S. G. Rzuczek, C. A. Thornton and M. D. Disney, *ACS Chem. Biol.*, 2012, **7**, 856.
39 S. P. Velagapudi, A. Pushechnikov, L. P. Labuda, J. M. French and M. D. Disney, *ACS Chem. Biol.*, 2012, **7**, 1902.
40 S. P. Velagapudi, S. M. Gallo and M. D. Disney, *Nat. Chem. Biol.*, 2014, **10**, 291.
41 J. Means, S. Katz, A. Nayek, R. Anupam, J. V. Hines and S. C. Bergmeier, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 3600.
42 R. Anupam, A. Nayek, N. J. Green, F. J. Grundy, T. M. Henkin, J. A. Means, S. C. Bergmeier and J. V. Hines, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 3541.
43 I. Maciagiewicz, S. Zhou, S. C. Bergmeier and J. V. Hines, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 4524.
44 C. M. Orac, S. Zhou, J. A. Means, D. Boehm, S. C. Bergmeier and J. V. Hines, *J. Med. Chem.*, 2011, **54**, 6786.
45 N. Gelus, C. Bailly, F. Hamy, T. Klimkait, W. D. Wilson and D. W. Boykin, *Bioorg. Med. Chem.*, 1999, **7**, 1089.
46 G. Xiao, A. Kumar, K. Li, C. T. Rigl, M. Bajic, T. M. Davis, D. W. Boykin and W. D. Wilson, *Bioorg. Med. Chem.*, 2001, **9**, 1097.
47 W.-Y. Yang, R. Gao, M. Southern, P. S. Sarkar and M. D. Disney, *Nat. Commun.*, 2016, **7**, 11647.
48 M. D. Disney, A. M. Winkelsas, S. P. Velagapudi, M. Southern, M. Fallahi and J. L. Childs-Disney, *ACS Chem. Biol.*, 2016, **11**, 1720.
49 K. Weeks, C. Ampe, S. Schultz, T. Steitz and D. Crothers, *Science*, 1990, **249**, 1281.
50 F. Aboul-ela, J. Karn and G. Varani, *J. Mol. Biol.*, 1995, **253**, 313.
51 N. J. Keen, M. J. Gait and J. Karn, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 2505.
52 M. Stevens, E. De Clercq and J. Balzarini, *Med. Res. Rev.*, 2006, **26**, 595.
53 K. S. Long and D. M. Crothers, *Biochemistry*, 1995, **34**, 8885.
54 C. Parolin, B. Gatto, C. Del Vecchio, T. Pecere, E. Tramontano, V. Cecchetti, A. Fravolini, S. Masiero, M. Palumbo and G. Palù, *Antimicrob. Agents Chemother.*, 2003, **47**, 889.
55 Y. Ming, *Curr. Drug Targets: Infect. Disord.*, 2005, **5**, 433.
56 A. Davidson, D. W. Begley, C. Lau and G. Varani, *J. Mol. Biol.*, 2011, **410**, 984.
57 M. Zeiger, S. Stark, E. Kalden, B. Ackermann, J. Ferner, U. Scheffer, F. Shoja-Bazargani, V. Erdel, H. Schwalbe and M. W. Göbel, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 5576.
58 A. Davidson, T. C. Leeper, Z. Athanassiou, K. Patora-Komisarska, J. Karn, J. A. Robinson and G. Varani, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 11931.
59 D. I. Bryson, W. Zhang, P. M. McLendon, T. M. Reineke and W. L. Santos, *ACS Chem. Biol.*, 2012, **7**, 210.
60 J. Lu, L. Nguyen, L. Zhao, T. Xia and X. Qi, *Biochemistry*, 2015, **54**, 3687.
61 D. W. Crawford, B. D. Blakeley, P.-H. Chen, C. Sherpa, S. F. J. Le Grice, I. A. Laird-Offringa and B. R. McNaughton, *ACS Chem. Biol.*, 2016, **11**, 2206.
62 A. T. Frank, A. C. Stelzer, H. M. Al-Hashimi and I. Andricioaei, *Nucleic Acids Res.*, 2009, **37**, 3670.
63 C. Bailly, A. W. Cuthbert, D. Gentle, M. R. Knowles and M. Waring, *J. Biochem.*, 1993, **32**, 2514.
64 A. W. Cuthbert, *Mol. Pharmacol.*, 1976, **12**, 945.
65 H. Matthews, M. Ranson, J. D. A. Tyndall and M. J. Kelso, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 6760.
66 M. Murai, S. Habu, S. Murakami, T. Ito and H. Miyoshi, *Biosci., Biotechnol., Biochem.*, 2015, **79**, 1061.
67 C. Matsumoto, K. Hamasaki, H. Mihara and A. Ueno, *Bioorg. Med. Chem. Lett.*, 2000, **10**, 1857.
68 B. Sathyamoorthy, J. Lee, I. Kimsey, L. Ganser and H. Al-Hashimi, *J. Biomol. NMR*, 2014, **60**, 77.
69 E. A. Dethoff, A. L. Hansen, C. Musselman, E. D. Watt, I. Andricioaei and H. M. Al-Hashimi, *Biophys. J.*, 2008, **95**, 3906.
70 E. J. Cragoe, O. W. Woltersdorf, J. B. Bicking, S. F. Kwong and J. H. Jones, *J. Med. Chem.*, 1967, **10**, 66.
71 H. Palandoken, K. By, M. Hegde, W. R. Harley, F. A. Gorin and M. H. Nantz, *J. Pharmacol. Exp. Ther.*, 2005, **312**, 961.
72 S. D. Kuduk, C. N. Di Marco, R. K. Chang, R. M. DiPardo, S. P. Cook, M. J. Cato, A. Jovanovska, M. O. Urban, M. Leitl, R. H. Spencer, S. A. Kane, M. T. Bilodeau, G. D. Hartman and M. G. Bock, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 2514.

73 A. P. Massey, W. R. Harley, N. Pasupuleti, F. A. Gorin and M. H. Nantz, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 2635.
74 W. J. Thompson, J. H. Jones, P. A. Lyle and J. E. Thies, *J. Org. Chem.*, 1988, **53**, 2052.
75 A. K. Ghosh, W. J. Thompson, M. K. Holloway, S. P. McKee, T. T. Duong, H. Y. Lee, P. M. Munson, A. M. Smith and J. M. Wai, *J. Med. Chem.*, 1993, **36**, 2300.
76 P.-P. Kung, M. D. Casper, K. L. Cook, L. Wilson-Lingardo, L. M. Risen, T. A. Vickers, R. Ranken, L. B. Blyn, J. R. Wyatt, P. D. Cook and D. J. Ecker, *J. Med. Chem.*, 1999, **42**, 4705.
77 S. H. Kim, G. N. Anilkumar, L. G. Zawacki, Q. Zeng, D.-Y. Yang, Y. Shao, G. Dong, X. Xu, W. Yu, Y. Jiang, C.-H. Jenh, J. W. Hall Iii, C. D. Carroll, D. W. Hobbs, J. J. Baldwin, B. F. McGuinness, S. B. Rosenblum, J. A. Kozlowski, B. B. Shankar and N.-Y. Shih, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 6982.
78 L. Salmon, G. Bascom, I. Andricioaei and H. M. Al-Hashimi, *J. Am. Chem. Soc.*, 2013, **135**, 5457.
79 R. Abagyan, M. Totrov and D. Kuznetsov, *J. Comput. Chem.*, 1994, **15**, 488.
80 E. J. Martin, J. M. Blaney, M. A. Siani, D. C. Spellmeyer, A. K. Wong and W. H. Moos, *J. Med. Chem.*, 1995, **38**, 1431.
81 A. Bender and R. C. Glen, *Org. Biomol. Chem.*, 2004, **2**, 3204.
82 H. Eckert and J. Bajorath, *Drug Discovery Today*, 2007, **12**, 225.
83 C. F. Stratton, D. J. Newman and D. S. Tan, *Bioorg. Med. Chem. Lett.*, 2015, **25**, 4802.
84 T. A. Wenderski, C. F. Stratton, R. A. Bauer, F. Kopp and D. S. Tan, in *Chemical Biology: Methods and Protocols*, ed. E. J. Hempel, H. C. Williams and C. C. Hong, Springer New York, New York, NY, 2015, p. 225.
85 T. T. Tanimoto, *IBM Internal Report*, IBM Corporation, Armonk NY, 1957.
86 G. Maggiora, M. Vogt, D. Stumpfe and J. Bajorath, *J. Med. Chem.*, 2014, **57**, 3186.
87 D. E. Patterson, R. D. Cramer, A. M. Ferguson, R. D. Clark and L. E. Weinberger, *J. Med. Chem.*, 1996, **39**, 3049.
88 C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Delivery Rev.*, 2001, **46**, 3.
89 J. B. Baell and G. A. Holloway, *J. Med. Chem.*, 2010, **53**, 2719.
90 P. Selzer, H.-J. Roth, P. Ertl and A. Schuffenhauer, *Curr. Opin. Chem. Biol.*, 2005, **9**, 310.
91 S. R. Kirk, N. W. Luedtke and Y. Tor, *J. Am. Chem. Soc.*, 2000, **122**, 980.