Published in final edited form as:

Bioorg Med Chem. 2019 April 15; 27(8): 1759–1765. doi:10.1016/j.bmc.2019.03.016.

Molecular Recognition of a Branched Peptide with HIV-1 Rev Response Element (RRE) RNA

Yumin Dai^{a,†}, Ashley N. Peralta^{a,†}, Jessica E. Wynn^{a,†}, Chringma Sherpa^b, Hao Li^a, Astha Verma^a, Stuart F. J. Le Grice^b, and Webster L. Santos^a

^aDepartment of Chemistry and Virginia Tech Center for Drug Discovery, Virginia Tech, Blacksburg, VA, 24061, United States

^bBasic Research Laboratory, National Cancer Institute, Frederick, MD, 21702, United States

Abstract

Interaction of HIV-1 rev response element (RRE) RNA with its cognate protein. Rev, is critical for HIV-1 replication. Understanding the mode of interaction between RRE RNA and ligands at the binding site can facilitate RNA molecular recognition as well as provide a strategy for developing anti-HIV therapeutics. Our approach utilizes branched peptides as a scaffold for multivalent binding to RRE IIB (high affinity rev binding site) with incorporation of unnatural amino acids to increase affinity via non-canonical interactions with the RNA. Previous high throughput screening of a 46,656-member library revealed several hits that bound RRE IIB RNA in the sub-micromolar range. In particular, the lead compound, 4B3, displayed a K_d value of 410 nM and demonstrated selectivity towards RRE. A ribonuclease protection assay revealed that 4B3 binds to the stem-loop structure of RRE IIB RNA, which was confirmed by SHAPE analysis with 234 nt long NL4-3 RRE RNA. Our studies further indicated interaction of 4B3 with both primary and secondary Rev binding sites.

Graphical Abstract

[†]denotes co-first authorship

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

HIV-1; RRE RNA; Branched peptides; Boronic acids; RNA targeting

1. Introduction

RNA plays vital roles in multiple biological processes such as transcription, translation, catalysis, and splicing.^{1–4} RNA's secondary structure including stems, bulges, and hairpin loops, affords a unique three-dimensional architecture that presents a challenge when developing selective binding of various ligands.⁵ Due to its critical role and unique structure, RNA thus provides an ideal target for the development of therapeutics.⁶ The human immunodeficiency virus (HIV-1) RNA genome has been investigated with a view to therapeutic interventions. Among several *cis*-acting sequences, the highly conserved HIV-1 Rev Response Element (RRE) RNA is essential to RNA replication. RRE stem-loop IIB has been recognized as the primary binding site of Rev protein, with secondary binding extended to stem IA.⁷ Once Rev interacts with RRE, the Rev-RRE ribonucleoprotein complex facilitates the transport of unspliced and singly spliced transcripts to the cytoplasm, which contain *gag*, *pol*, and *env* genes as well as the genome for a budding virion.⁸ In the cytoplasm, translation of these genes provide genomic RNA and the proteins necessary for subsequent replication. Targeting the Rev/RRE axis, therefore, provides a strategy to inhibit HIV-1 replication.

To date, several ligands including small molecules, aminoglycosides, and antisense oligonucleotides have been designed to target the Rev-RRE complex.^{9,10,11} In particular, several peptide-based molecules or peptidomimetics displayed high RRE binding affinity. Inspired by the Rev peptide, an α -helical fragment that recognizes the major groove of RRE tertiary structure,¹² efforts have been made to introduce α -helicity into peptide ligands. For example, RSG1.2 was selected from a combinatorial library and shown to bind to RRE SL-IIB hairpin with 7-fold higher affinity and 15-fold higher specificity than the arginine-rich domain of Rev itself.¹³ To target the RRE major groove, α -helicity was induced in R₆QR₇, a RRE specific peptide, through macrolactam constraints. The best peptidomimetic displayed a K_d value of 40 nM, which was 4-fold higher compared to that of the original peptide.¹⁴ Other strategies incorporate metal binding linkers in the ligands. For example, work by

Frankel and coworkers reported a 14 amino acid sequence derived from the Rev peptide with an engineered zinc finger framework. Through the histidine residue, the peptides were shown to bind specifically to RRE with 7-fold improved affinity in a zinc-dependent manner.¹⁵ Moreover, Cowan and co-workers reported a series of metal-chelate Rev conjugates that displayed 10-fold higher binding affinity for the RRE compared to Rev. These chelate-Rev conjugates were used to form coordination complexes with metals and the arginine-rich Rev peptide could mediate localization towards the RRE SL-IIB site¹⁶.

Our laboratories have developed medium size branched peptides (BPs) that target RNA structures.¹⁷ Their medium size (1000-2000 Da) is small enough to facilitate cell permeability but large enough to allow for increased surface area contact, making them more competitive against other RNA-binding ligands.¹⁸ With one bead one compound (OBOC) peptide synthesis, diverse peptide libraries with a wide range of amino acids can be accessed, affording potential macromolecular interactions to increase selectivity and affinity. For example, exotic monomers with boronic acid appendage can be introduced.^{19,20} The empty p orbital in boron facilitates Lewis acid-base complexation likely through the 2'-hydroxyl group of the ribose ring. Such an interaction results in a selective molecular recognition of HIV-1 Rev-RRE RNA versus the corresponding identical DNA sequence.¹⁹ Furthermore, our investigations demonstrated that the binding affinity of branched peptides can be significantly increased with an intercalating moiety such as the fused aromatic ring system in acridinyl lysine, generating peptides in the low double digit nanomolar level.²¹ We recently reported the high-throughput screening and characterization of a fourth generation branched peptide library that consisted solely of unnatural amino acids at all variable positions.²² Each unnatural amino acid offers a unique mode of binding (electrostatic interaction, hydrogen bonding, pi stacking, etc) and include L-Guanidinoproline (L), D-aminoproline (D), guanine peptide nucleic acid (G), lysine boronic acid (B) and 1-naphthylalanine (N) (Fig. 1A). As part of our investigations that focus on understanding the intermolecular interactions that govern binding of ligands to folded RNA structures, we report the characterization of 4B3 against HIV-1 RRE UB and a much larger 234 nt RRE RNA construct using a combination of biochemical and biophysical methods. These studies suggest that 4B3 binds to the primary and secondary binding site of Rev on RRE and causes local structural reorganization.

2. Results and Discussion

2.1 Selectivity of hit branched peptide 4B3 toward RRE IIB tertiary structure

As listed in Table 1, biophysical characterization of the top hit compounds revealed binding affinities in the sub-micromolar range. Among the branched peptides investigated, 4B3 (Fig. 1B), which consisted of L-guanidinoproline, D-aminoproline, guanine peptide nucleic acid, lysine boronic acid, lysine and tyrosine, displayed the highest binding affinity with a K_d value of 410 nM. We selected 4B3 to characterize the RNA-peptide interactions that result in improved RNA binding. Selectivity was determined with a fluorescence competition assay using 2-AP-labeled RRE SL-IIB RNA in presence of 10-fold excess unlabeled nucleic acid competitors: LoopA/B/BulgeA Deleted RRE SL-IIB RNA (No Bulge RRE), homologous sequence RRE SL-IIB DNA, and TAR RNA (Fig. 2). The HIV-1 transactivation response

element region (TAR) RNA is a highly conserved sequence containing a hexanucleotide loop and a 3-nt bulge located on the 5' end of all nascent mRNA²³ TAR interacts with the arginine rich motif (ARM) of the transcriptional activator protein, Tat, which promotes efficient transcription that ultimately leads to efficient viral replication. As TAR has a short stem-loop and a bulge, it was used as competitor RNA. As shown in Fig. 3, the fluorescence of 2-AP-labeled RRE SL-IIB RNA decreased as a function of increasing peptide concentration suggesting a change in the environment around 2-amino purine. In the presence of competitor nucleic acids, increasing K_d values indicate interaction of the peptide with the competing agent. As shown in Table 2, 4B3 showed good selectivity for RRE IIB in the presence of No Bulge RRE RNA and RRE DNA, with approximately a 3-fold loss of binding affinity observed with both competitor nucleic acids. These results were consistent from our previous work, indicating the loop and bulge in the RRE SL-IIB construct, especially the upper stem region, are essential for peptide binding.^{22, 24} In addition, 4B3 showed lower selectivity for RRE SL-IIB in the presence of TAR RNA as the K_d against RRE IIB increased approximately 10-fold. Since both TAR and RRE RNAs share a small stem-loop with a consensus sequence UGGG, this unique structure of the RNA may be necessary for 4B3 binding.

2.2 Determination of 4B3 RNA binding site by RNase footprinting

RRE binding site of lead compound 4B3 was investigated *via* footprinting assays using a variety of RNases. The RNA construct was designed from the NL43 sequence and structure of RRE SL-IIB RNA is part of SL-II, which is the primary Rev binding site. Nucleotides outside RRE SL-IIB that have been shown to affect Rev-RRE function were also included. In this assay, increasing concentrations of BP (0-20 μ M) were incubated with 5'-³²P labeled RRE SL-IIB, which was then subjected to enzymatic cleavage by RNases A and VI (Fig. 4). Protection of the RNA from enzymatic cleavage is indicative of potential contact points with the branched peptide. For 4B3, the most noticeable cleavage was caused by RNase VI in the upper stem between bases A52 and G55, as this enzyme cleaves double-stranded RNA. This cleavage pattern revealed similar binding sites compared to previously characterized BPs from other 3.3.4 BP libraries.²⁴ A concentration-dependent protection from RNase VI in the upper stem region of the RNA was observed as evidenced by reduced cleavage at G53, C54, and C65. Protection of RRE SL-IIB from RNase A, which cleaves the 3' end of pyrimidines C and U, was also observed in the internal loop region of the RNA (U43 and U72). Interactions of the branched peptide with this region of the RNA were not surprising, as this is also the region where Rev binds.^{25, 26} Overall, **4B3** interacts with both the internal loop and upper stem/apical loop regions of RRE SL-IIB, spanning a significant portion of the entire stem-loop structure.

2.3 SHAPE analysis of 4B3:RNA interactions

The full length RRE is highly structured wherein long-range contacts between SL-I and SL-II occurs to orchestrate cooperative binding of multiple Rev molecules.²⁷ Such long range interactions have the potential to influence the binding of 4B3. To investigate the effect of 4B3, instead of focusing on the SL-IIB arm, we expanded 4B3 footprint study to the 234 nt RRE, the minimal sized functional RRE,²⁸ using SHAPE-MaP (selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling) technology. SHAPE-MaP

interrogates RNA flexibility using SHAPE chemistry (2'-hydroxyl nucleobase accessibility to an electrophile) combined with real time PCR and high-throughput sequencing.²⁹ SHAPE based footprinting has also been previously employed successfully to study ligand binding to RNA.^{22, 30–32} Free RNA and RNA incubated with various concentrations of 4B3 were treated with electrophile 1-methyl-7-nitroisatoic anhydride (IM7), and reactivity differences were plotted as a function of nucleotide position. A cutoff of ± 0.3 was arbitrarily chosen to identify nucleotides that were rendered more flexible or constrained in the presence of 4B3.

As shown in Fig. 5, protection in the secondary Rev binding site loop (Stem IA) was observed.³³ This protection is enhanced when the peptide level is increased. At 1:1 RNA:BP ratio, decreased reactivity was observed at nt 28, suggesting protection by 4B3 at this position. At a 1:10 RNA:BP ratio, this protection extended to nt 27, as well as to nucleotides on the opposite strand of the loop (nt 206 and 207). Additionally, since the 3-dimensional architecture of RRE²⁷ indicates an interaction of SL-I with SL-II, reactivity changes in the SL-I region might indicate a binding event in SL-II that affects flexibility in the secondary Rev binding site via long range interactions. We, however, did not observe 4B3-induced decrease in IM7 reactivity in SL-II suggesting the peptide might predominantly bind to an already constrained base paired region. This observation is consistent with our RNase protection assay data in the previous section that showed protection effect of 4B3 in the upper stem region of SL-IIB. Nevertheless, nt 46, which forms non-canonical base pairing upon Rev,¹² exhibited increased reactivity in the presence of 4B3. This reactivity change at nt 46 might reflect local structural changes imparted by 4B3 binding in the neighboring SL-II sequence. Changes in IM7 reactivity were also observed at nt 105 and nt 134, which might be due to 4B3 binding or 4B3 induced local and long-range structural changes on RRE.

Overall, SHAPE MaP data indicated our best binder 4B3 may interact with both the SL-I and SL-II of NL43 RRE, which are two important Rev binding sites on the RRE.

3. Conclusion

In summary, we have investigated the interaction of a branched peptide (4B3) with the stem loop IIB and larger RRE RNA construct. The studies were focused on 4B3 because it has the highest binding affinity ($K_d \sim 400$ nM) with SL IIB of RRE. In the presence of competitor RNAs as well as DNA, 4B3 exhibited good selectivity with the RNA target. Indeed, detailed structural analysis of its binding site via RNase footprinting and SHAPE-MaP reactivity assays suggested interaction not only on stem loop IIB but also in the bulge region of stem loop I, which are the primary and secondary binding sites for Rev, respectively. Our studies demonstrate the recognition of RRE RNA by the branched peptide that include long range intermolecular interactions as well as local structural rearrangements. Our work can inform future strategies towards the molecular recognition of RNA.

4. Experimental section

4.1 Peptide synthesis, purification, and characterization

Synthesis of the branched peptides was achieved by solid phase peptide synthesis using N- α -Fmoc protected L- and D-amino acids (3 equiv.), Pyoxim (Novabiochem) (3 equiv.) in DMF as coupling reagent, and DIE A (Aldrich) (6 equiv.) on Rink amide MBF1A resin (100-200 mesh) (Novabiochem) with 0.6 mmol/g loading. The Fmoc group was deprotected with 20% piperidine in DMF. Fmoc-Lys(Fmoc)-OFI was used as a branching unit, and molar equivalencies of reagents were doubled in coupling reactions after installation of the branching unit. Solid phase synthesis was performed on a vacuum manifold (Qiagen) outfitted with 3-way Luer lock stopcocks (Sigma) in either Poly-Prep columns or Econo-Pac polypropylene columns (Bio-Rad). The resin was mixed in solution by bubbling argon during all coupling and washing steps. After Fmoc deprotection of the N-terminal amino acids, the resin was bubbled in a phenylboronic acid solution (0.2 g/mL, 1.6 M) in DMF overnight to remove the pinacol groups of boron-containing side chains. Finally, the resin was treated with 95:2.5:2.5 TFA (Trifluoroacetic acid, Acros/H₂O/TIPS (Triisopropylsilane, Acros) (v/v/v) for 3.5 h. The supernatant was dried under reduced pressure, and the crude peptide triturated from cold diethyl ether. Peptides were purified on a Jupiter 4 μ m Proteo 90 Å semiprep column (Phenomenex) using a solvent gradient composed of 0.1% TFA in Milli-Q water and FPLC grade acetonitrile. Peptide purity was determined using a Jupiter 4 μ m Proteo 90 Å analytical column (Phenomenex), and identity confirmed by MALDI-TOF analysis. Peptide concentrations were measured in nuclease free water at 280 nm using their calculated extinction coefficient. The extinction coefficient for lysine pyrazine was experimentally determined (3374 M⁻¹cm⁻¹) by monitoring the absorbance of pyranic acid in nuclease free water at 280 nm. Previously reported extinction coefficients were used for guanine PNA (7765 M⁻¹cm⁻¹) and the naphthalene derivative (3374 M⁻¹cm⁻¹) at 280 nm in aqueous medium.^{34,35}

4.2 Transcription of HIV-1 SL-IIB RNA and SL-IIB RRE variants

Wild-type and mutant RRE SL-IIB RNAs were transcribed *in vitro* by T7 polymerase with the Ribomax T7 Express System (Promega) using previously reported techniques.^{24, 36} The antisense template, sense complementary strand (5'-ATGTAATACGACTCACTATAGG-3') and RRE SL-IIB reverse PCR primer (5'-GGCTGGCCTGTAC-3') were purchased from Integrated DNA Technologies (Coralville, Iowa). Antisense templates were: RRE SL-IIB RNA 5'-GGCTGGCCTGTACCGTCAGCGTCATTGACGCTGCGCCCATACCAGCCCTATAGTGAGTTCGTATTACAT-3'; Loop A/B/Bulge A Deleted 5'-GGCTGGCAGCGTCATTGACGCTGCCAGCCCTATAGTGAGTCGTATTACAT-3'. Wild type HIV-1 RRE SL-IIB was PCR amplified using HotstarTaq DNA polymerase (Qiagen) followed by a clean-up procedure using a spin column kit (Qiagen). For preparation of all other sequences, the antisense DNA template was annealed with the sense DNA complementary strand in reaction buffer at 95 °C for 2 min, then cooled on ice for 4 min. T7 transcription proceeded at 42 °C for 1.5 h. Subsequently, DNA templates were degraded with DNase at 37 °C for 45 min and the RNA was purified by a 12% polyacrylamide gel containing 7.5 M urea. The band corresponding to the RNA of interest was excised from the

gel and eluted overnight in 1× TBE buffer at 4 °C. The sample was desalted using a Sep-Pak syringe cartridge (Waters Corporation) and lyophilized. Purified RNA was stored as a pellet at –80 °C. RRE SL-IIB DNA (5'-GGCTGGTATGGGCGCAGCGTCAATGACGCTGACGGTACAGGCCAGCC-3') was purchased from Integrated DNA Technologies and stored at –20 °C.

4.3 2-Aminopurine (AP) assay

All fluorescence spectra were measured on a Varian Cary Eclipse fluorescence spectrophotometer using a xenon flash lamp with a thermoelectrically controlled cell holder. The excitation slit width and the emission slit width were set to 10 nm. Excitation of the sample was performed at 310 nm and fluorescence spectra were collected from 340 nm to 450 nm. A quartz cell of 1 cm path length transparent on three sides (Starna Cells, Inc.) was used.

The dissociation constants for the branched peptides were determined by following the decrease in fluorescence at 372 nm of 2-AP labeled RRE SL-IIB RNA (5'-CUGGUAUGGGCGCAGCGUCAAUGACGCGACGG-2AP-ACAGGCCAGCC-3', Integrated DNA Technologies) as a function of increase in branched peptide concentration.³⁷ 2-AP labeled RRE SL-IIB was refolded by heating at 95 °C for 3 min and snap-cooling on ice. RNA concentration was fixed at 0.1 μM during the titration and the peptide concentration was varied from 0-20 μM. Both peptides and RNA were prepared with 0.2 μM sterile-filtered 1× phosphate buffer (10 mM potassium phosphate, 100 mM KCl, 1 mM MgCl₂, 20 mM NaCl, pH 7.0). Binding data were analyzed using a Hill equation (Eq. 1) with Kaleidagraph (Synergy Software).³⁸ In this equation, *b* and *y* are the fluorescence emission intensities of the RNA in the absence and presence of peptide; *m* is the fluorescence emission intensity of the RNA in the presence of an infinite drug concentration; *R* and *x* are the total concentrations of the RNA and peptide; *K* is the *K_d* value of the peptide binding to the RNA; *n* is the apparent cooperativity. Each experiment was performed at least in duplicate and error bars represent the standard deviation calculated over the replicates.

$$y = b + (m - b) * \left(\frac{1}{1 + \left(\frac{K}{x} \right)^n} \right) \quad \text{Eq. 1}$$

4.4 Preparation of 32-P labeled RNA

HIV-1 RRE SL-IIB RNA was dephosphorylated with calf intestinal phosphatase (CIP) in NEBuffer 3 (New England Biolabs) according to the manufacturer's protocol. The product was recovered by a standard phenol extraction followed by ethanol precipitation. RNA was labeled at the 5'-end by treating 10 pmol of dephosphorylated RNA with 20 pmol of [γ-³²P] ATP (111 TBq mol⁻¹) and 20 units of T4 polynucleotide kinase in 70 mM Tris•HCl, 10 mM MgCl₂, and 5 mM dithiothreitol, pH 7.6. The mixture was incubated at 37 °C for 30 min, and then at Room temperature for 20 min. Kinase was heat-inactivated at 65 °C for 10 min, and RNA recovered by ethanol precipitation. RNA purity was examined using 12% denaturing PAGE followed by autoradiography.

4.5 Nuclease protection assays

RNA was first refolded by heating a solution of 5'-³²P-labeled RRE-IIB (10 nM) and excess unlabeled RRE SL-IIB (200 nM) at 95 °C for 3 min and then snap cooling on ice. Refolded RNA was incubated on ice for 4 h in a solution containing the BPBA and buffer composed of 10 mM Tris, pH 7, 100 mM KCl, and 10 mM MgCl₂. RNase (Ambion) was then added to the solution, which was further incubated on ice for 10 min (0.002 Units RNase V1), or 1 h (1 Unit RNase T1; 20 ng RNase A). Inactivation/precipitation buffer (Ambion) was added to halt digestion, and the RNA was pelleted by centrifugation at 13,200 rpm for 15 min. Pelleted RNA was redissolved into tracking dye and fractionated through a 12 % polyacrylamide gel containing 7.5 M urea. The gel was dried at 80 °C for 1 h and imaged by autoradiography.

4.6 SHAPE analysis of 4B3

RNA preparation for SHAPE: 234 nt NL4-3 5'-RRE RNA was prepared by *in vitro* transcription using the MegaShortScript kit (Ambion/Life Technologies) per manufacturers' recommendations. DNA template used in the transcription reaction was generated by PCR from a pro viral pNL4-3 plasmid using high fidelity platinum Taq DNA polymerase (Invitrogen) using forward oligo 234RREf (5' AGCGTACTTAATACGACTCACTATAGGGAGGAGCTTTGTTTCCTTG) and reverse oligo 234RREr (5' AGGAGCTGTTGATCCTTTAG). The forward primer was designed to introduce T7 promoter sequence at the 5' end of the RRE. RNA was then treated with Turbo DNase I for 1 h at 37 °C, heated at 85 °C for 2 min and fractionated on a 5% denaturing polyacrylamide gel (1× TBE, 7 M urea) at constant temperature (45 °C, 30 W max). The SL-II containing gel slice was excised, and RNA electroeluted at 200 V for 2 h at 4 °C, ethanol precipitated and stored at -20 °C in TE buffer (10 mM Tris, pH 7.6; 0.1 mM EDTA) prior to use.

RNA folding and modification: 5 pmol of RNA, treated with renaturation buffer (final concentration 10 mM Tris/HCl pH 8.0, 100mM KCl, 0.1mM EDTA), in a volume of 5 µL was heated to 95°C for 2 mins and snap cooled on ice. Renatured RNA was incubated with RNA folding buffer (final concentration: 70 mM Tris/HCl pH 8.0, 180 mM KCl, 0.3 mM EDTA, 8 mM MgCl₂, 5% of glycerol) in a total volume of 8 µL at 37°C for 15 min and cooled to 4°C. During this period 5 pmoles/µL and 50 pmoles/µL of peptide dilutions were made in 1 × PBS. 1 µL of 1 × PBS or the above two peptide dilutions were then added to different tubes containing folded RNA to created peptide negative control reaction, 1:1, and 1:10 RNA:peptide ratio test reactions. Reactions were incubated at 37°C for 15 min and the RNA probed using 2.5 nM (final concentration) 1M7 in DMSO at 37°C for 5 min. 1M7 negative control reactions were generated as above with the exception that 1 µL of DMSO was added instead of 1M7. All the 1M7 (+) and (-) reactions were placed on ice and the denaturing control reaction was generated as previously described.²⁹ All RNAs were purified using RNeasy mini kit RNA cleanup protocol.

Mutational profiling: SHAPE-MaP experiments were largely performed as previously described.²⁹ Approximately 0.5 pmoles of the purified RNA was used per reaction to generate cDNA, by mutational profiling (MaP) reverse transcription using RRE specific

oligo 234-RRE (5' AGGAGCTGTTGATCCTTTA G) as primers. RNA in the cDNA reaction was hydrolyzed by adding 1 μ L 2 N NaOH to each reaction and the reaction neutralized by adding 1 μ L 2N HCl, cDNA was purified by passing each sample through G50 spin columns. The entirety of each resulting cDNA was used as template in a 100 μ L PCR (PCR1) reaction (1.1 μ L each of 50 pmoles of forward and reverse oligo, 2 μ L of 10mM dNTPs, 20 μ L of 5 $\times$ Q5 reaction buffer, 1 μ L of hot Start High-Fidelity DNA polymerase). Cycling conditions comprised: 98°C for 30 secs, 15 cycles of [98°C for 10 secs, 50°C for 30 secs, 72°C for 30 sec], 72°C for 2 mins. to generate ds DNA with half of the Illumina adapters on each end. The resulting PCR product was purified using a gel purification kit (Qiagen) and the entire PCR product used in subsequent PCR reactions that added the rest half of the illumina adapters with appropriate indices. PCR2 reaction and cycling conditions were same as those for PCR1. The resulting sequencing amplicon library was fractionated on a 2% agarose gel and the amplicons recovered and purified from the gel slices by electro-elution at room temperature for 2 hours followed by ethanol precipitation. Each library was quantitated by real time PCR using KAPA Universal Library Quantitation kit (cat# KK4824) per manufacturer's protocol. Sequencing libraries were pooled and mixed with 20% phiX and sequenced using MiniSeq Mid Output Kit (300-cycles) on MiniSeq sequencer following the manufacturer's protocol to generate 2 $\times$ 150 paired-end reads. SHAPE reactivity profiles were created by aligning reads to the 234-nt NL4-3 RRE sequence using ShapeMapper (v1.2, <http://chem.unc.edu/rna/software.html>) with default settings. Reactivity values corresponding to PCR1 primer binding regions (nt 1–nt 17 and nt 219 – nt 234) were excluded and the rest renormalized to generate the final reactivity profiles.

Acknowledgments

We thank K. Ray and R. Helm in the VT mass spectrometry incubator for their assistance with MALDI-TOF analysis, and K. Decourcy at the Fralin Life Science Institute for her assistance with confocal microscopy. SL and CS were supported by the Intramural Research Program of the National Cancer Institute of Health, Department of Health and Human Services. This work was supported by National Institutes of Health grant RO1 GM093834.

References

1. Wanowska E, Kubiak MR, Rosikiewicz W, Makałowska I, Szczepaniak MW, Natural antisense transcripts in diseases: From modes of action to targeted therapies, Wiley Interdisciplinary Reviews: RNA e1461–n/a.
2. Ren A, Micura R, Patel DJ, Structure-based mechanistic insights into catalysis by small self-cleaving ribozymes, *Curr Opin Chem Biol* 2017; 41: 71–83. [PubMed: 29107885]
3. Larochelle S, Making the cut, *Nat Chem Biol* 2018; 14: 107.
4. Harvey RF, Smith TS, Mulroney T, Queiroz RML, Pizzinga M, Dezi V, Villeneuve E, Ramakrishna M, Lilley KS, Willis AE, Trans-acting translational regulatory RNA binding proteins, *WIREs RNA* e1465–n/a.
5. Disney MD, Dwyer BG, Childs-Disney JL, Drugging the RNA World, *Cold Spring Harb Perspect Biol* 2018; 10: doi: 10.1101/cshperspect.a034769.
6. Zaman GJ, Michiels PJ, van Boeckel CA, Targeting RNA: new opportunities to address drugless targets, *Drug Discov Today* 2003; 8: 297–306. [PubMed: 12654542]
7. Jayaraman B, Mavor D, Gross JD, Frankel AD, Thermodynamics of Rev–RNA Interactions in HIV-1 Rev–RRE Assembly, *Biochemistry* 2015; 54: 6545–6554. [PubMed: 26422686]
8. Rausch JW, Le Grice SF, HIV Rev Assembly on the Rev Response Element (RRE): A Structural Perspective, *Viruses* 2015; 7: 3053–3075. [PubMed: 26075509]

9. Nakaya T, Iwai S, Fujinaga K, Sato Y, Otsuka E, Ikuta K, Decoy approach using RNA-DNA chimera oligonucleotides to inhibit the regulatory function of human immunodeficiency virus type 1 Rev protein, *Antimicrob Agents CH* 1997; 41: 319–325.
10. Prado S, Beltrán M, Moreno Á, Bedoya LM, Alcamí J, Gallego J, A small-molecule inhibitor of HIV-1 Rev function detected by a diversity screen based on RRE-Rev interference, *Biochem Pharmacol* 2018; 156: 68–77. [PubMed: 30071201]
11. Cho J, Rando RR, Specificity in the Binding of Aminoglycosides to HIV-RRE RNA *Biochemistry* 1999; 38: 8548–8554. [PubMed: 10387102]
12. Battiste JL, Mao H, Rao NS, Tan R, Muhandiram DR, Kay LE, Frankel AD, Williamson JR, Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex, *Science* 1996; 273: 1547–1551. [PubMed: 8703216]
13. Zhang Q, Harada K, Cho HS, Frankel AD, Wemmer DE, Structural characterization of the complex of the Rev response element RNA with a selected peptide, *Chem Biol* 2001; 8: 511–520. [PubMed: 11358697]
14. Mills NL, Daugherty MD, Frankel AD, Guy RK, An alpha-helical peptidomimetic inhibitor of the HIV-1 Rev-RRE interaction, *J Am Chem Soc* 2006; 128: 3496–3497. [PubMed: 16536504]
15. McColl DJ, Honchell CD, Frankel AD, Structure-based design of an RNA-binding zinc finger, *Proc Natl Acad Sci U S A* 1999; 96: 9521–9526. [PubMed: 10449725]
16. Joyner JC, Cowan JA, Targeted Cleavage of HIV RRE RNA by Rev-Coupled Transition Metal Chelates, *J Am Chem Soc* 2011; 133: 9912–9922. [PubMed: 21585196]
17. Wynn JE, Santos WL, HIV-1 drug discovery: targeting folded RNA structures with branched peptides, *Org. Biomol. Chem* 2015; 13: 5848–5858. [PubMed: 25958855]
18. Bryson DI, Zhang W, McLendon PM, Reineke TM, Santos WL, Toward Targeting RNA Structure: Branched Peptides as Cell-Permeable Ligands to TAR RNA, *ACS Chem. Biol* 2012; 7: 210–217. [PubMed: 22003984]
19. Zhang W, Bryson DI, Crumpton JB, Wynn J, Santos WL, Targeting folded RNA: a branched peptide boronic acid that binds to a large surface area of HIV-1 RRE RNA, *Org. Biomol. Chem* 2013; 11: 6263–6271. [PubMed: 23925474]
20. Zhang W, Bryson DI, Crumpton JB, Wynn J, Santos WL, Branched peptide boronic acids (BPBAs): a novel mode of binding towards RNA, *Chem. Commun. (Cambridge, U. K.)* 2013; 49: 2436–2438.
21. Wynn JE, Zhang W, Tebit DM, Gray LR, Hammarskjöld ML, Rekosh D, Santos WL, Effect of intercalator and Lewis acid-base branched peptide complex formation: boosting affinity towards HIV-1 RRE RNA, *Medchemcomm* 2016; 7: 1436–1440. [PubMed: 27453773]
22. Dai Y, Wynn JE, Peralta AN, Sherpa C, Jayaraman B, Li H, Verma A, Frankel AD, Le Grice SF, Santos WL, Discovery of a Branched Peptide That Recognizes the Rev Response Element (RRE) RNA and Blocks HIV-1 Replication, *J Med Chem* 2018; 61: 9611–9620. [PubMed: 30289719]
23. Rana TM, Jeang K-T, Biochemical and Functional Interactions between HIV-1 Tat Protein and TAR RNA, *Arch Biochem Biophys* 1999; 365: 175–185. [PubMed: 10328810]
24. Zhang W, Bryson DI, Crumpton JB, Wynn J, Santos WL, Targeting folded RNA: a branched peptideboronic acid that binds to a large surface area of HIV-1 RRE RNA, *Org. Biomol. Chem* 2013; 11: 6263–6271. [PubMed: 23925474]
25. Kjems J, Calnan BJ, Frankel AD, Sharp PA, Specific binding of a basic peptide from HIV-1 Rev, *EMBO J* 1992; 11: 1119–1129. [PubMed: 1547776]
26. Kjems J, Brown M, Chang DD, Sharp PA, Structural analysis of the interaction between the human immunodeficiency virus Rev protein and the Rev response element, *Proc. Natl. Acad. Sci. U S A* 1991; 88: 683–687. [PubMed: 1992459]
27. Bai Y, Tambe A, Zhou K, Doudna JA, RNA-guided assembly of Rev-RRE nuclear export complexes, *ELife* 2014; 3: e03656. [PubMed: 25163983]
28. Malim MH, Hauber J, Le S-Y, Maizel JV, Cullen BR, The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA, *Nature* 1989; 338: 254–257. [PubMed: 2784194]

29. Smola MJ, Rice GM, Busan S, Siegfried NA, Weeks KM, Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis, *Nat. Protoc* 2015; 10: 1643–1669. [PubMed: 26426499]
30. Siegfried NA, Busan S, Rice GM, Nelson JA, Weeks KM, RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP), *Nat Methods* 2014; 11: 959–965. [PubMed: 25028896]
31. Sztuba-Solinska J, Shenoy SR, Gareiss P, Krumpe LRH, Le Grice SFJ, O'Keefe BR, Schneekloth JS, Identification of Biologically Active, HIV TAR RNA-Binding Small Molecules Using Small Molecule Microarrays, *J Am Chem Soc* 2014; 136: 8402–8410. [PubMed: 24820959]
32. Crawford DW, Blakeley BD, Chen P-H, Sherpa C, Le Grice SFJ, Laird-Offringa IA, McNaughton BR, An Evolved RNA Recognition Motif That Suppresses HIV-1 Tat/TAR-Dependent Transcription, *ACS Chem Biol* 2016; 11: 2206–2215. [PubMed: 27253715]
33. Daugherty MD, D'Orso I, Frankel AD, A Solution to Limited Genomic Capacity: Using Adaptable Binding Surfaces to Assemble the Functional HIV Rev Oligomer on RNA, *Mol Cell* 31: 824–834. [PubMed: 18922466]
34. Mallik B, Katragadda M, Spruce LA, Carafides C, Tsokos CG, Morikis D, Lambris JD, Design and NMR characterization of active analogues of compstatin containing non-natural amino acids, *J Med Chem* 2005; 48: 274–286. [PubMed: 15634022]
35. Smith SJ, Rittinger K, Preparation of GTPases for structural and biophysical analysis, *Methods Mol Biol* 2002; 189: 13–24. [PubMed: 12094582]
36. Milligan JF, Groebe DR, Withered GW, Uhlenbeck OC, Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates, *Nucleic Acids Res* 1987; 15: 8783–8798. [PubMed: 3684574]
37. Lacourciere KA, Stivers JT, Marino JP, Mechanism of neomycin and Rev peptide binding to the Rev responsive element of HIV-1 as determined by fluorescence and NMR spectroscopy, *Biochemistry* 2000; 39: 5630–5641. [PubMed: 10801313]
38. Pagano JM, Clingman CC, Ryder SP, Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes, *RNA* 2011; 17: 14–20. [PubMed: 21098142]

Fig. 1.
(A) Structures of unnatural amino acid monomers used in the library. (B) Structure of peptide 4B3.

Fig. 2.
 (A) Structure of 2-AP-RRE-IIB RNA, X = 2-aminopurine. (B) TAR RNA. (C) LoopA/B/
 BulgeA Deleted RRE IIB RNA (RRE No Bulge)

Fig. 3.
Titration curve of 4B3 against competitor nucleic acids.

Fig. 4. RNase protection assay of RRE SL-IIB. The gel depicts the autoradiogram of alkaline hydrolysis (AH) and RNase protection experiments using RNases A, and V1 with increasing concentration of 4B3. Colored triangles highlight bases protected from cleavage by RNase A (blue) and RNase V1 (red).

Fig. 5. SHAPE-MaP reactivity profiles of 234 nt RRE RNA in the presence of peptide 4B3. **Upper panel:** Reactivity values of the RRE superimposed on the 5-SL structure of NL43 RRE (A) in the absence of 4B3. (B) in the presence of equal molar concentration of 4B3. and (C) in the presence of 10-fold molar excess of 4B3. Green arrows indicate nucleotides whose reactivity changes significantly in the presence of 4B3. **Lower panel:** Reactivity difference profile generated for (D) equimolar RRE:peptide ratio and (E) 1:10 molar ratio of RRE and peptide by subtracting the reactivity values of the RRE in the absence of the 4B3 from the reactivity values of the RRE in its presence. Here ± 0.3 difference is used as cutoff value to identify nucleotides that are more relaxed or more constrained in the presence of 4B3.

Table 1.

Dissociation constant of selected hit compounds

No.	Peptides	Sequences	K_d (μM) ^a
1	4B3	(DLL) ₂ KLGBY	0.41 ± 0.05
2	13B8	(LDL) ₂ KLGBY	0.63 ± 0.03
3	T1	(LDN) ₂ KGGNY	0.66 ± 0.02
4	11F6	(DLD) ₂ KDLDY	0.72 ± 0.03
5	4A5	(GGD) ₂ KPGLY	0.88 ± 0.02
6	10B10	(LDD) ₂ KPGDY	1.02 ± 0.07
7	16D9	(PDL) ₂ KGLDY	1.11 ± 0.05
8	RA7	(LDL) ₂ KPPGY	1.12 ± 0.08
9	19B3	(LDG) ₂ KGPBY	1.18 ± 0.05
10	RA8	(LDP) ₂ KLLPY	1.30 ± 0.03

^aEach experiment was performed at least in duplicate and error bars represent the standard deviation calculated over the replicates.

Table 2.

Dissociation constant of 4B3 with RRE IIB in the presence of indicated RNAs.

Nucleic Acids	K_d (μM) ^a
Wild Type RRE IIB	0.40 ± 0.05
10× RRE DNA	1.48 ± 0.03
10× No Bulge RRE RNA	1.14 ± 0.10
10× TAR RNA	3.52 ± 0.52

^aEach experiment was performed at least in duplicate and error bars represent the standard deviation calculated over the replicates.