Targeting HIV-1 RNAs with Medium Sized Branched Peptides Featuring Boron and Acridine—Branched Peptide Library Design, Synthesis, High-Throughput Screening and Validation

Wenyu Zhang

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Chemistry

Webster L. Santos, Chair
David G. I. Kingston
Richard D. Gandour
Daniel G. S. Capelluto

(Date of Defense)
Blacksburg, VA

Keywords: HIV-1, TAR RNA, RRE RNA, Branched Peptide Library, High-throughput Screening, Branched Peptide Boronic Acids, Acridine Branched Peptides, Unnatural Amino Acids.

Targeting HIV-1 RNAs with Medium Sized Branched Peptides Featuring Boron and Acridine—Branched Peptide Library Design, Synthesis, High-Throughput Screening and Validation

Wenyu Zhang

Abstract

RNAs have gained significant attention in recent years because they can fold into well-defined secondary or tertiary structures. These three dimensional architectures provide interfaces for specific RNA–RNA or RNA–protein interactions that are essential for biological processes in a living system. These discoveries greatly increased interest in RNA as a potential drug target for the treatment of diseases. Two of the most studied RNA based regulatory systems are HIV-1 *trans*-activating response element (TAR)/Tat replication pathway and Rev response element (RRE)/Rev export pathway. To efficiently target TAR and RRE RNA, we designed and synthesized three generations of branched peptide libraries that resulted in medium sized molecules.

The first generation of BPs were discovered from screening a one-bead one-compound library (4,096 compounds) against HIV-1 TAR RNA. One peptide FL4 displayed a binding affinity of 600 nM to TAR RNA, which is tighter than its native protein counterpart, Tat. Biophysical characterization of these BP demonstrated that 'branches' in BPs impart multivalency, and they are cell permeable and non-toxic.

The second generation peptides were discovered from an on-bead high-throughput screening of a 3.3.4 branched peptide boronic acids (BPBAs) library that bind selectively to the tertiary structure of RRE IIB. The library comprised of 46,656 unique sequences. We demonstrate that our highest affinity BPBA (**BPBA1**) selectively binds RRE IIB in the

presence of competitor tRNAs as well as against six RRE IIB structural variants. Further, we show that the boronic acid moieties afford a novel binding mode towards RNA that is tunable; their Lewis acidity has critical effects on binding affinity. In addition, biophysical characterizations provide evidence that "branching" in these peptides is a key structural motif for multivalent interactions with the target RNA. Finally, RNA footprinting studies revealed that the **BPBA1** binding site encompasses a large surface area that spans both the upper stem as well as the internal loop regions of RRE IIB. BPBA1 is cell permeable and non-toxic.

In the next generation of branched peptides, a 3.3.4 branched peptide library composed of 4,096 unique sequences that featured boronic acid and acridine moieties was designed. We chose acridine as the amino acid side chain due to its potential for π-stacking interaction that provides high binding affinity to RNA target. The library was screened against HIV-1 RRE IIB RNA. Fifteen peptides were sequenced and four contained acridine alone and/or in conjunction with boronic acid moieties displayed dissociation constants lower than 100 nM. The ribonuclease protection assays of **A7**, a sequence that contains both boronic acid and acridine residues, showed a similar protection pattern compared to previous peptide **BPBA1**, suggesting that the 3.3.4 branched peptides shared similar structural elements and contacted comparable regions of the RRE IIB RNA.

The results from this research indicated that "branching" in peptides imparts multivalent interactions to the RNA, and that functional groups such as boronic acid and acridine are key structural features for efficient binding and selectivity for the folded RNA target. We demonstrated that the branched peptides are cell permeable and non-toxic.

Acknowledgements

First and foremost, I would like to thank my doctoral advisor and friend Dr. Webster L. Santos. His guidance throughout my studies at Virginia Tech has been indispensable. I could not have asked for a better mentor. He has done everything possible to provide me with the tools needed to be a successful scientist in the future, and his tireless work has provided my research with the funding required to keep it moving forward. I am constantly amazed and appreciative of his availability to provide answers to my questions or concerns. He has always made me feel welcome to knock on his office door for a brief chat. This thesis would not have been possible without his inspiration and support. His profound knowledge and great personality deeply influenced me in my research and life.

I must also thank my current and previous Ph. D. committee members Drs. Kingston, Gandour, Capelluto and Reineke. Your positive support and critiques throughout my six years at Virginia Tech have helped fuel my desire for research and bolstered my confidence as a scientist.

I would also like to thank my friends and colleagues in the Santos Group—Dr. Ming Gao, Dr. Mithun Raje, Dr. David Bryson, Dr. Jason Crumpton, Dr. Brandon Thorpe, Dr. Neeraj Patwardhan, Ken Knott, Michael Perfetti, Jing Sun, Xi Guo, Jessica Wynn, Amanda Nelson, Molly Congdon, Joe Calderone, Emily Morris, Cheryl Peck, Beth Childress, Hao Li, Russell Snead, Analyn Carreon, Marietou Paye, Ryan Stephens, Leah Heist, Wes Morris, and Matt Nguyen. I would also like to thank our collaborators on this project—Dr. W. Keith Ray, Dr. Patrick M. McLendon, Dr. David M. Rekosh, Dr. Joseph Falkinham, and Dr. Takashi Morii.

Finally, I would like to thank my Mother, Father, and family. Your support and encouragement has been unwavering throughout my life, and you have always believed in me even when I didn't believe in myself. I love you all so much.

List of Figures

List of Tables

Chapter 1. HIV-1 RNAs as Therapeutic Targets for Drug Development

1.1 Introduction

The human immunodeficiency virus type 1 (HIV-1) is the causative agent of the worldwide health problem, Acquired Immune Deficiency Syndrome (AIDS). AIDS has been under significant investigation for over 30 years, and more than 20 drugs have been approved for treatment of AIDS.[1] The majority of these drugs can be divided into four groups: nucleoside reverse transcriptase inhibitors (NRTIs) / nucleotide reverse transcriptase inhibitors (NtRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs), protease inhibitors (PIs), and fusion inhibitors (FIs).[2] The initial treatment for patients featured the use of one or two drugs from a single group of inhibitors, which showed moderate effects on reducing the plasma HIV-1 RNA. This led to the rapid emergence of drug resistance. Consequently, three or more drugs from two or more different groups of inhibitors were used for treatment, now known as highly active antiretroviral therapy (HAART). HAART showed notable success in reducing plasma viral loads to undetectable levels during HIV infection.[3] However, later studies found that the virus displayed drug resistance and cross-drug resistance as a result of using HAART. In the next viral colony, they even showed increased resistance. The observation of replication from some residual virus during HAART revealed that HAART has failed to completely eliminate the virus from the body due to the remaining chronically HIV-infected $CD4^+$ T cells, which contain integrated but transcriptionally dormant HIV provirus.[3-4] In addition, the toxicity of HAART, including lipodystrophy, dyslipidaemia, hypersensitivity, liver

dysfunction and mitochondrial toxicity that results from inhibition of mitochondrial polymerase-γ, has been reviewed and is considered an important issue in the treatment of AIDS patients.[5] Drug resistance and toxicity, as well as other side effects of current HIV therapies make an increasing urgency to develop new anti-HIV treatment strategies and novel HIV targets with reasonable toxicity and resistance profiles.

1.2 HIV Life Cycle

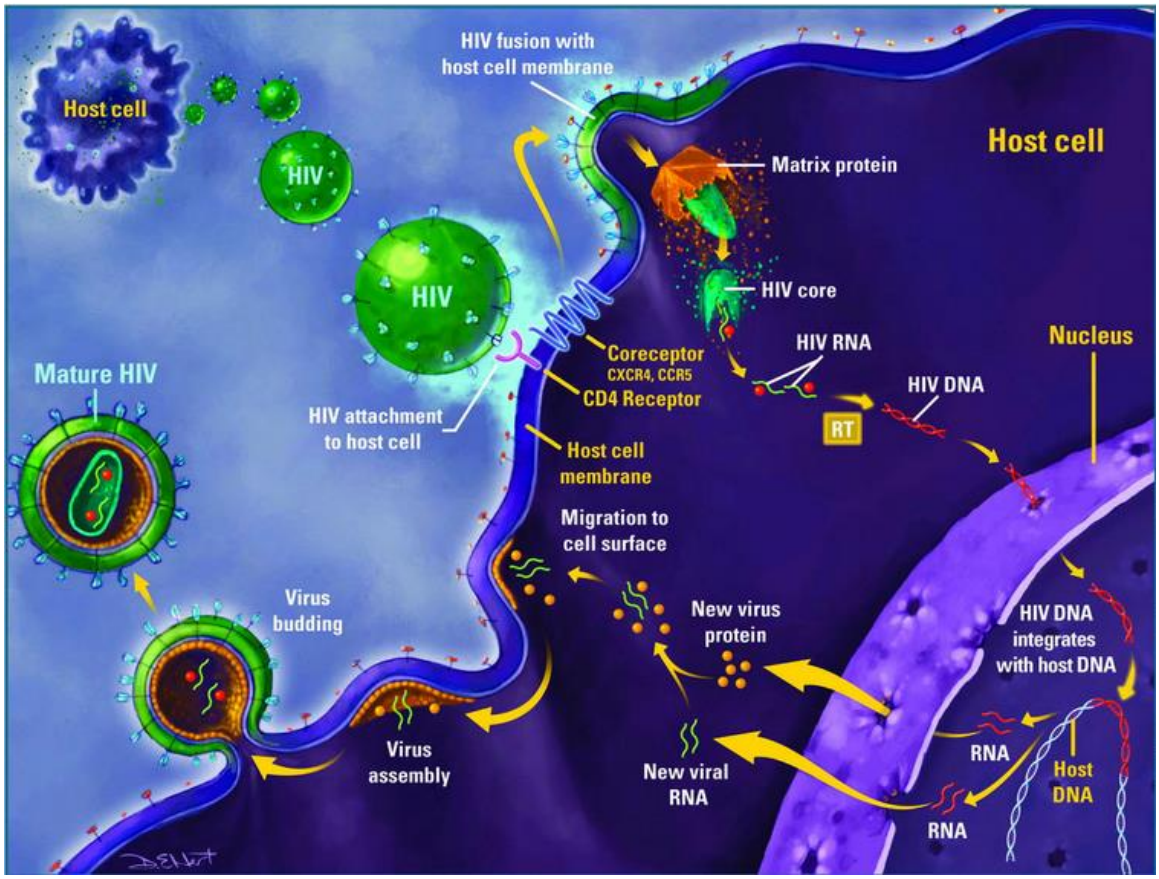


Figure 1.1 HIV-1 life cycle (www.HIVwebstudy.org).

HIV-1, like most retroviruses, has a replication cycle that is comprised of four central steps: infection, reverse transcription and integration, viral-gene expression, and virus assembly and maturation (Figure 1.1). Infectious viruses initially bind to the cellular receptor CD4 on the surface of T cells via its envelope (Env) glycoprotein gp120. The

binding of CD4 and gp120 results in a conformational change in gp120, allowing it to bind to one of the CD4 chemokine co-receptors, CCR5 and CXCR4, and cause the exposition of gp41 (a glycoprotein non-covalently bound to gp120).[6] Protein gp41 assists the fusion of the outer membrane of the virus with the host cell membrane, and the viral genome information is released to the host cytoplasm. The single stranded RNA genome is reverse transcribed into double stranded DNA by viral reverse transcriptase (RT). At this point, the DNA enters the nucleus as a nucleic acid–protein complex and is integrated at random sites of the host cell's genome by integrase (IN). This covalently integrated form of viral DNA is defined as a provirus. Once integrated, the proviral genome serves as the template of viral transcription which is regulated by the host cell as well as its own viral transcriptional control mechanisms.[7] This is then followed by viral RNA synthesis, processing, and transport to the cytoplasm for the translation of viral proteins. The *gag* and *pol* core proteins are assembled with two copies of full length viral RNA, budding through the plasma membrane. In the meantime, a third viral enzyme protease (PR) cleaves the core proteins and releases mature virus particles.[8-9]

In principle, any suitable target involved in an essential step during the replication of the virus could be developed as a drug target. The initial step of viral infection is a complex process. The interaction of the gp120 and CD4 receptor, as well as their co-receptor CCR5 or CXCR4 and a membrane protein gp41, offers a potential therapeutic target. While tremendous effort has been made to develop entry inhibitors, only one drug has been approved by FDA: enfuvirtide (Figure 1.2), which is a fusion inhibitor. Enfuvirtide contains 36 amino acids that is derived from a prototypic gp41 subunits region (Heptad repeats 2, HR2). It blocks HIV-1 entry by binding to another gp41 subunit (HR1).

Though highly potent, its low oral bioavailability and huge production cost limits its application. Notwithstanding, enfuvirtide is the only FDA-approved fusion inhibitor; many others are in development or in preclinical/clinical trials.[10] Two piperazine derivatives (**1** and **2**) led by structure–activity relationship (SAR) studies have been identified as potent inhibitors that bind directly to the viral protein gp120 and stop the interaction with CD4 receptors on host T cells. Further pharmacokinetic studies have shown that both compounds have a moderate bioavailability and excellent anti-viral activity.[11] Another entry inhibitor, maraviroc acts as an antagonist of two major co-receptors: CCR5 and CXCR4 and is the only FDA-approved co-receptor inhibitor. It was first discovered from a high-throughput screen of the Pfizer compound file.[12] Maraviroc demonstrated potent antiviral activity against CCR5 by blocking gp120 binding to co-receptor CCR5.[12] Further clinical research showed that 100 mg of maraviroc twice daily could reduce the viral load for at least 10 days.[13] All of these compounds have shown great success inhibiting HIV-1 replication by targeting virion structural and accessory proteins.

HIV enzymes, like reverse transcriptase, protease, and integrase, are another kind of potential target. The most investigated reverse transcriptase inhibitors are NRTIs/NtRTIs and NNRTIs.[14] One clinically approved NRTI, AZT, was discovered as the nucleoside analogue lacking a 3'-OH group; once this compound is phosphorylated by cellular nucleoside kinase (CNK) and integrated into a host DNA strand, the polymerization process is terminated.[2] Different from NRTIs, NNRTIs, such as nevirapine, were designed to target reverse transcriptase by binding noncompetitively to a hydrophobic pocket in the palm subdomain.[2] Indinavir is a peptidomimetic protease inhibitor. It showed

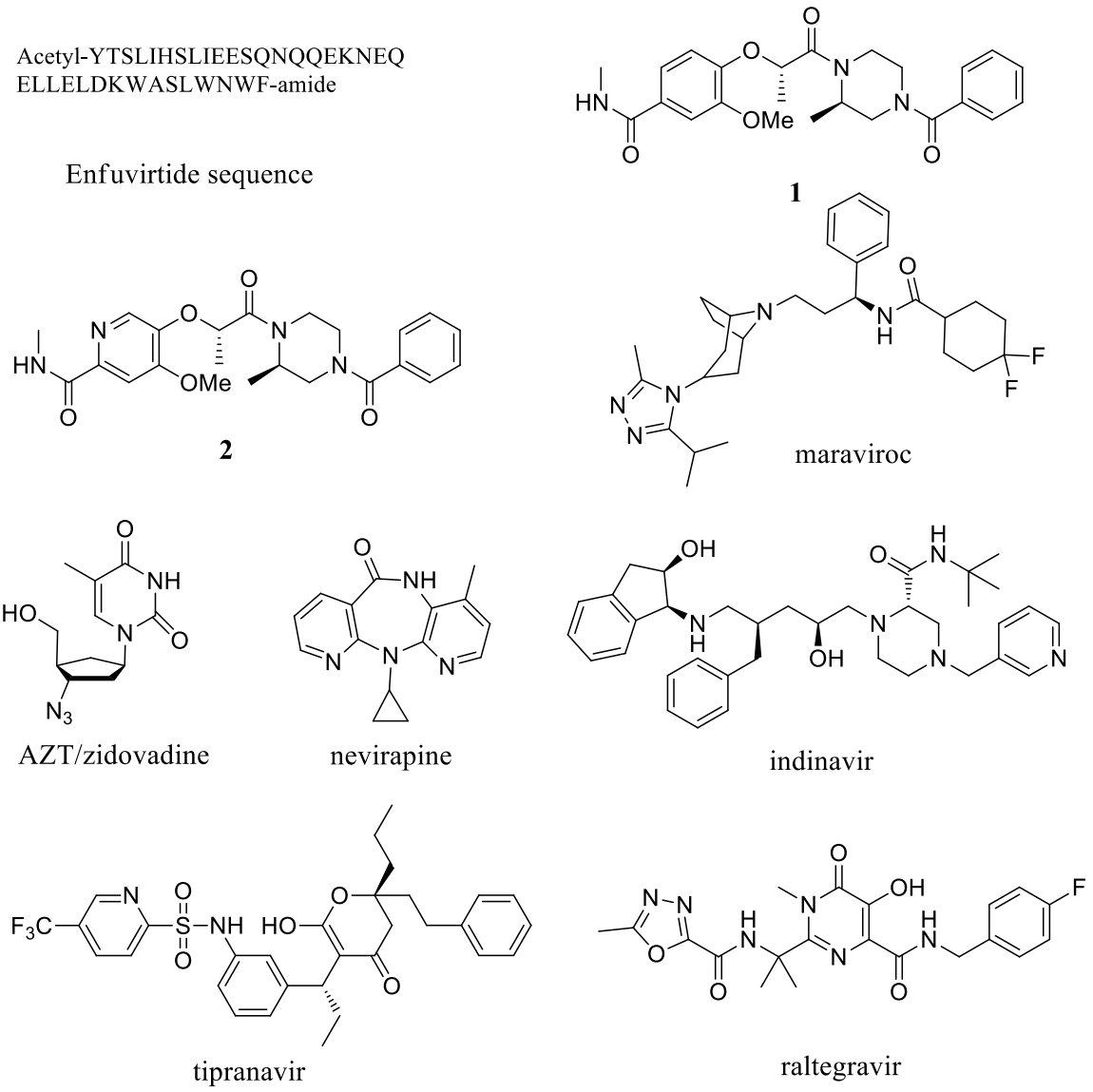


Figure 1.2 Viral entry inhibitors and enzyme inhibitors.

a high binding affinity to protease for both HIV-1 and HIV-2. However, it also showed cross-resistance to other peptidomimetic PIs as well as side effects that may be due to the lack of selectivity.[15] Tipranavir was the first non-peptidomimetic PI to lead a new class of protease inhibitors. It exhibited picomolar range binding affinity to protease and reduced cross-resistance.[2] Integrase is another favored target because there are no counterparts

present in mammals. Raltegravir was the first FDA approved integrase inhibitor and was discovered from a modification to a compound containing β diketo acid (DKA). Raltegravir is known for its inhibitory activity at the strand transfer step in integration. It binds to IN and stops the incorporation of the full length HIV DNA copy into the host cell DNA.[16] Although hundreds of enzyme inhibitors have been developed, they still face drug-resistance, cross-resistance and toxicity issues. Especially for RT inhibitors, a single point mutation in the RT results in viral resistance to other NNRTIs.[17] Although enormous advances have been achieved, the rate and number of HIV infections continue to rise. It is clear that the investigation of new drugs and the development of new strategies is of utmost importance.

1.3 Introduction of HIV-1 Tat–TAR and RRE–Rev Interactions

Two viral proteins control the most essential viral replication process of transcription: *trans*-activator (Tat) protein and the regulator of virion expression (Rev). Tat and Rev are RNA-binding proteins. Tat functions as an adaptor protein to stimulate transcriptional elongation from the viral long terminal repeat (LTR) via binding to *trans*-activating response element (TAR) RNA; Rev is required as a shuttle to export viral RNAs from the nucleus to the cytoplasm via binding to Rev response element (RRE) RNA. Both TAR and RRE are conserved regions that are not susceptible to mutation during replication. Therefore, a novel approach to inhibit Tat–TAR and RRE–Rev interactions and stop the transcription processes is to target these regulatory proteins or their binding RNAs.

1.3.1 HIV-1 Tat and TAR

The Tat protein consists of 101 amino acids for all natural viral isolates. An 86 amino acid version of Tat has also been isolated from HIV-1 Zaire 2, which does not exist

in natural viral isolates.[18] NMR studies indicate that there exists a hydrophobic core region of 16 amino acids and a basic region containing six arginines and two lysines in the Tat protein. The core region and basic region across all known Tat proteins are highly conserved.[19] The core region is required for sequence-specific RNA binding, and the basic region is involved in RNA binding, as well as nuclear localization of Tat.[20-21] A cysteine-rich region and a glutamine-rich region of Tat are also essential for transactivation; however, they may not be directly involved in RNA recognition and binding. Tat protein belongs to the arginine-rich motif (ARM) RNA binding family.[19]

Tat is believed to initiate the process of transcribing polymerases, and under some conditions it may also enhance the rate of this process. In the presence of Tat, the production of viral mRNAs increases by 100-fold, which is essential for viral replication, while without Tat, transcription generally stops beyond a few nucleotides.[22] However, Tat stimulates the transcription process by binding with the *trans*-activating response element (TAR) RNA. TAR RNA is a 59-residue hairpin loop structure located at the 5′ end of the nascent viral transcripts.[23] Deletion studies have demonstrated that the region from 19 to 42 spans the minimum nucleotides that are required for recognition and sufficient binding of Tat. This short RNA forms a stem-loop structure, which contains a hexa-loop, and a three-pyrimidine rich bulge (Figure 1.3).[24] The arginine-rich domain of Tat binds the U-rich bulge region of TAR RNA, and another arginine residue helps to recognize it.[25-26] After forming the Tat–TAR complex, a cofactor complex of cyclin T1 (CycT1) and cyclin-dependent kinase 9 (CDK9) induces a conformational change in Tat and binds to the hairpin loop region of TAR RNA (Figure 1.3).[27] CycT1 and Tat binding to TAR RNA are highly cooperative. CycT1 does not bind to TAR RNA without Tat, and Tat binds to TAR

RNA inefficiently in the absence of CycT1.[28] Therefore, CycT1 enhances the interaction of Tat protein and TAR RNA, and the TAR RNA further increases the affinity between Tat and CycT1.[1]

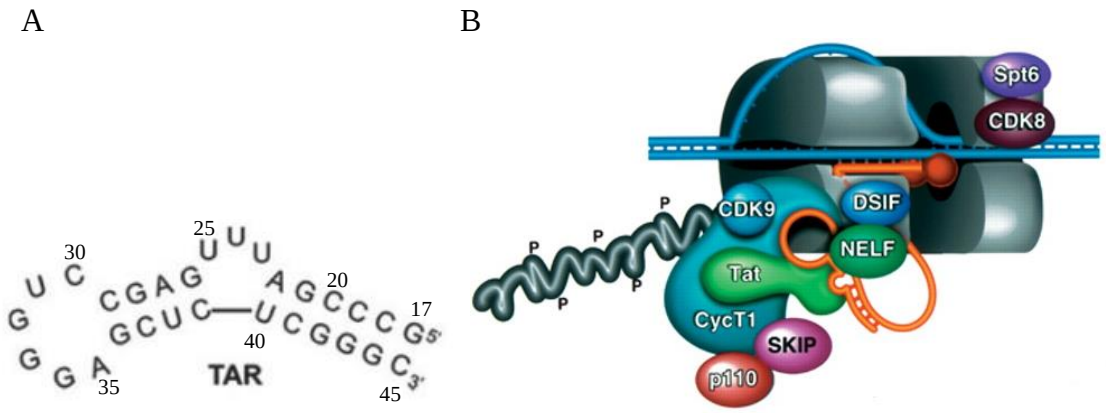


Figure 1.3 (A) minimum TAR RNA sequence required for binding and b) mechanism of HIV-1 transcription stimulated by Tat. [Brès, V.; Gomes, N.; Pickle, L.; Jones, K. A., A human splicing factor, SKIP, associates with P-TEFb and enhances transcription elongation by HIV-1 Tat. *Genes Dev.* **2005,** *19* (10), 1211-1226.], Copyright 2005; [used with Creative commons attribution 4.0 international public license].

A polypeptide ADP-1 was used to determine the interaction of the Tat–TAR complex. ADP-1 is a 37-mer that contains the minimal RNA recognition region of the Tat protein and mimics the Tat binding region. When ADP-1 binds, the bulge region of TAR RNA undergoes a local rearrangement so that the U40, A22 and U23 residues at the 5′ end of the bulge are positioned in the major groove (Figure 1.4). ADP-1 then forms a stacking interaction with U23, and the binding is further enhanced by contacting the critical phosphate groups P22, P23 and P40.[29] The binding is strongly inhibited by modification of these phosphates, which suggests that the interaction between the phosphates and Tat may account for the specific and selective binding to TAR RNA rather than other bulged RNAs.[18, 30] Other studies found that the sequences containing flanking residues from the core region along with basic and glutamine regions showed higher specificity than the

sequences containing only basic and glutamine regions.[20] This suggests that amino acids from core regions may provide additional interaction with TAR or introduce a conformational change of the peptide–RNA complex.[18]

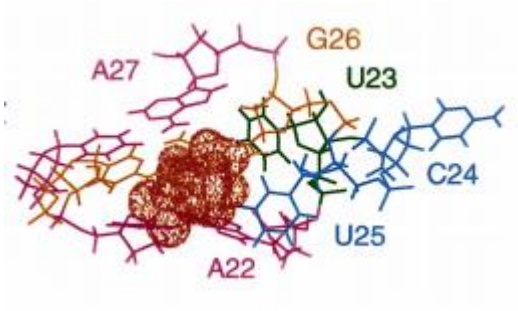


Figure 1.4 Close- up view of the bulge region with bound arginine residue shown by van der Waals surface. [Aboulela, F.; Karn, J.; Varani, G., The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J. Mol. Biol.* **1995,** *253* (2), 313-332.], Copyright 1995; [used with permission from Elsevier].

Taken together, TAR RNA plays an essential role in stimulating transcriptional elongation by interacting with Tat protein. Disruption of this interaction provides a potential anti-HIV strategy. To support this hypothesis, overexpression of TAR RNA "decoy" sequences also inhibits viral replication by acting as competitors.[31] The observation of the conserved TAR RNA structure, which is necessary during the replication, suggests that ligands can be used to explore development of high affinity inhibitors.[32]

1.3.2 HIV-1 Rev and RRE

Another essential HIV replication pathway is regulated by RRE–Rev interaction. Rev consists of 116-amino acids that are virally encoded, sequence specific RNA binding protein. Rev mainly exists in the nucleus/nucleolus.[33-34] It is believed to activate nuclear export of unspliced and singly spliced HIV-1 mRNA to the cytoplasm by interacting with its cis-acting RNA target sequence, the Rev response element (RRE).[35] Mutation studies have identified several essential regions in Rev.[36-38] An arginine-rich motif (position 35-

50, RQARRNRRRRWRERQR) near the N-terminal region serves as a nuclear export signal that directs nuclear location of Rev.[34, 37] A synthetic peptide that contains this region has been shown to interact with the same set of nucleotides as the Rev protein, which suggests that this arginine-rich motif is responsible for the specific binding to RRE.[37, 39-40] In fact, mutation experiments have demonstrated that each of these arginines as well as Thr34 and Asn40 are critical for specific RNA interaction, both *in vivo* and *in vitro*.[41] Previous circular dichroism studies also indicated this short peptide binds to RRE in an α-helical conformation and the α-helical conformation is further stabilized upon binding with RRE.[41-42] In addition, a relatively conserved hydrophilic α-helix region near the N-terminus has been identified as a nuclear inhibitory signal (NIS). This small domain contains 15 amino acids (position 10–24), and maintains the distribution of dominant subcellular proteins in nuclear, as well as the function of Rev itself.[38]

In order to utilize the limited single viral genome to express multiple genes that are required for mature viral package, the mRNA needs to be alternatively spliced into three major classes: ~2-kb fully spliced mRNAs, ~4-kb singly spliced mRNAs and ~9-kb unspliced full length transcripts.[8, 43] Due to the limited size of nuclear pores, only small proteins or RNAs with a molecular weight lower than 60 kD can pass freely.[38] At this point, only multiply spliced transcripts can be exported to the cytoplasm and translated into their corresponding proteins including Tat, Rev and Nef.[8] These proteins are also small enough to diffuse back into the nucleus. This stage has been described as early phase, where Rev and Tat have accumulated in the nucleus and the mRNA production has dramatically increased (Figure 1.5). In addition, Rev binds cooperatively to RRE, which is a span of ~240 nucleotides (351 nucleotides in more recent studies) located in the *env* gene of all

singly spliced and unspliced HIV-1 transcripts (Figure 1.6).[8, 44] RRE has been found to be an extremely well-conserved sequence of RNA across different HIV-1isolates.[45] In the late phase, the resulting RRE–Rev ribonucleoprotein complex binds the host Crm1 and is then shuttled out of the nucleus through the nuclear pore after the larger complex binds to Ran-GTP.[46-47] It has been demonstrated that proviral colonies without the rev gene have no replicative abilities and in the absence of rev protein, the stability of unspliced mRNA is decreased.[48] Since this cooperative binding allows for the export of full-length and singly spliced transcripts, the Rev/RRE export pathway has become a high profile drug target for its critical role in proliferation of HIV-1.

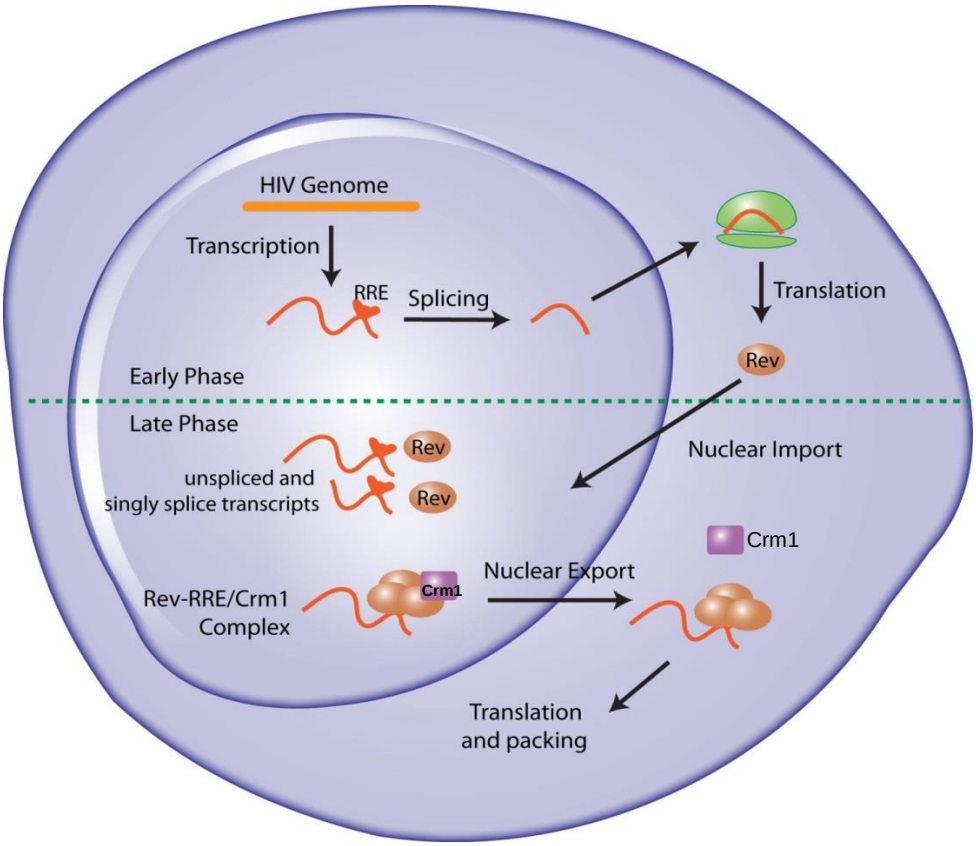


Figure 1.5 RRE–Rev pathway.

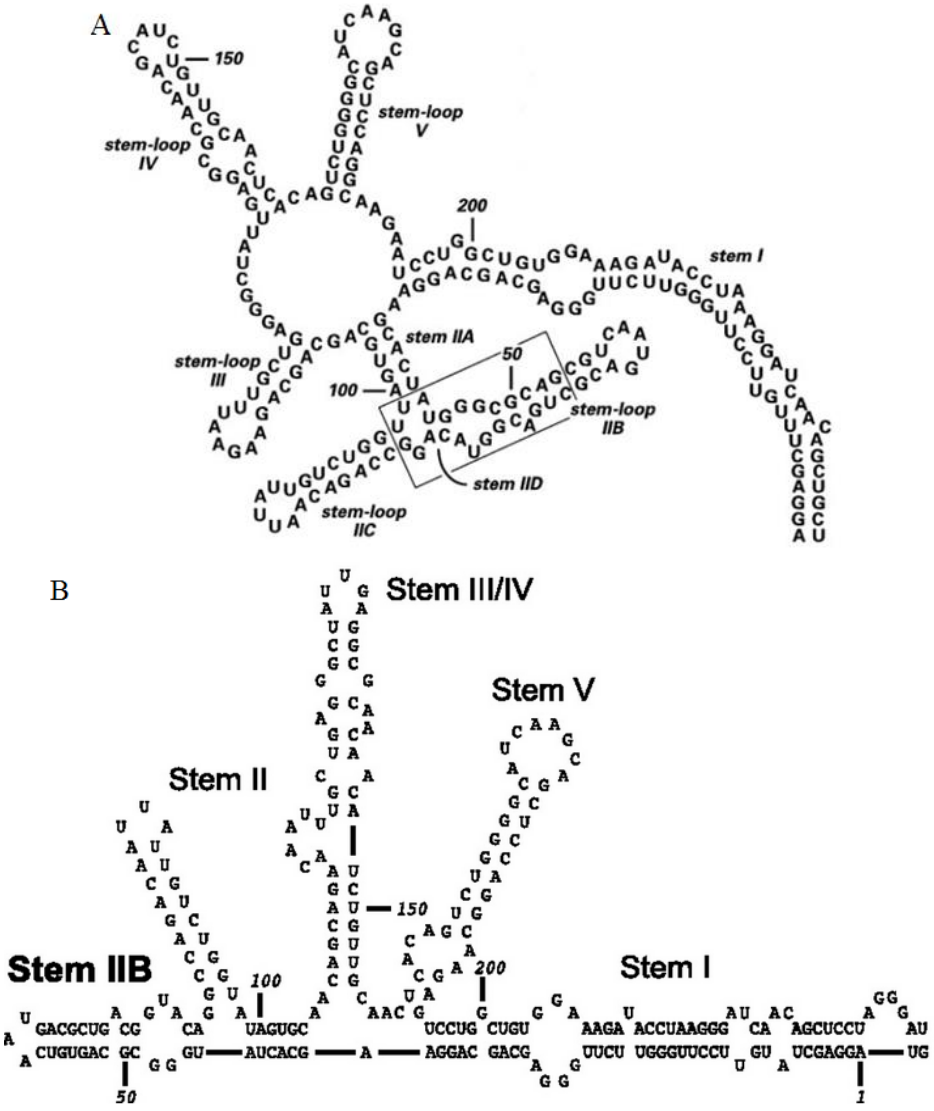


Figure 1.6 Two possible secondary structures of RRE (A) and (B). The squire box indicated sites for Rev recognization. [Mann, D. A.; Mikaélian, I.; Zemmel, R. W.; Gait, M. J.; Karn, J., A molecular rheostat: Co-operative Rev binding to stem I of the Rev-response element modulates human immunodeficiency virus type-1 late gene expression. *J. Mol. Biol.* **1994,** *241* (2), 193-207.], Copyright 1994; [used with permission from Elsevier].

1.4 Structural Basis for HIV-1 RRE–Rev Interaction

The specific sequences within RRE that are recognized by Rev are surprisingly limited. In fact, a stem-loop structure, which is termed RRE IIB, has been recognized as the high affinity site, where Rev initially binds.[49-50] Mutagenesis, chemical interference,

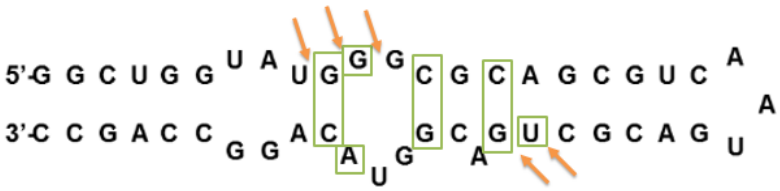


Figure 1.7 Structure of RRE IIB and important nucleotides.

and *in vitro* RNA selection experiments showed that several nucleotides (indicated by box) and phosphates on the RNA back bone (indicated by arrows) are necessary for Rev binding (Figure 1.7).[40, 51-52] NMR studies of a short Rev peptide derivative and RRE IIB have revealed several details of this specific interaction. As an A-form helix RNA, RRE carries a deep and narrow major groove that can barely adopt an α-helix peptide. However, two

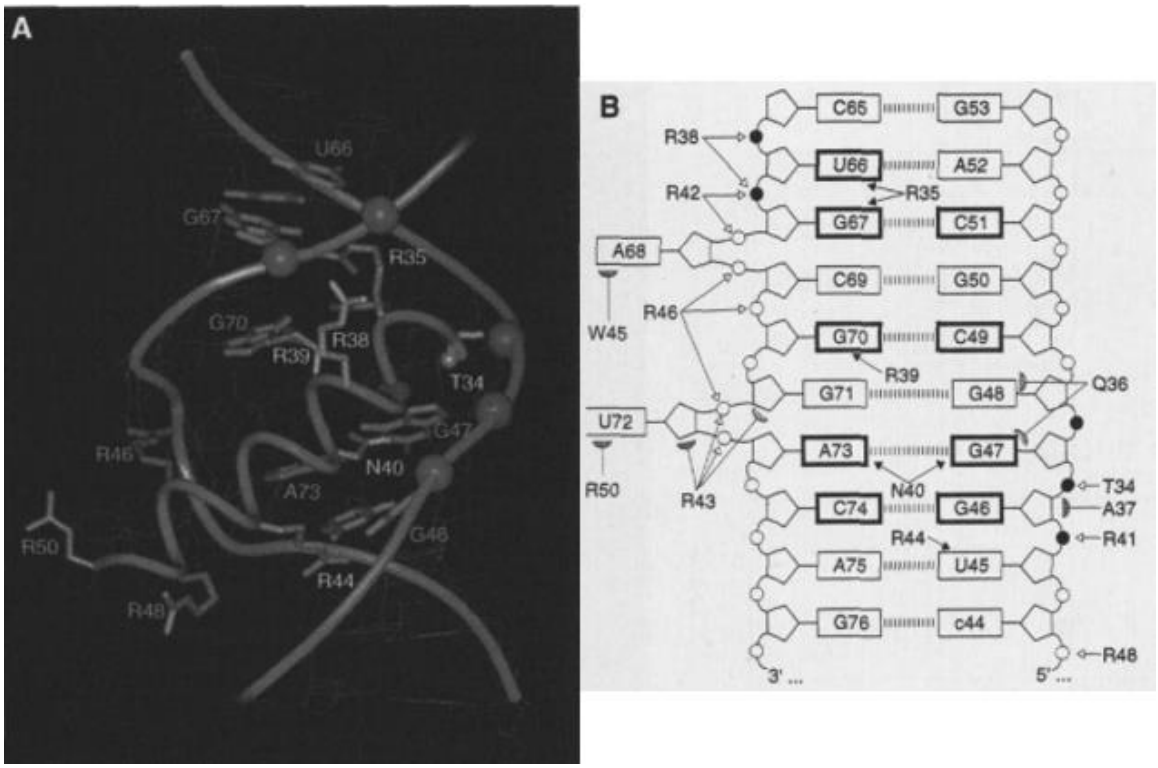


Figure 1.8 RNA-peptide interaction: A) View of core binding site; B) Schematic of specific RNA-peptide contacts. (Black arrows, open arrows and hatched arcs indicate base-specific, phosphate backbone and van der Waals contacts, respectively. Thick boxes and black circles indicate important nucleotides and phosphates, respectively.) [F Battiste, J. L.; Mao, H.; Frankel, A. D.; Williamson, J. R., α Helix-RNA major groove recognition in an HIV-1 Rev peptide–RRE RNA complex. *Science* **1996,** *273* (5281), 1547-1551.], Copyright 1996; [used with permission from AAAS].

non Watson–Crick base pairs (G48–G71 and G47–A73) in RRE IIB distort the sugar-phosphate backbone. [53] As a result, an S-shaped architecture to the backbone of nucleotides G70 and A73 is formed and opens the major groove by ~5 Å. This allows a deeper penetration of the Rev α-helix peptide compared to regular DNA binding proteins.[54] More detailed interaction of Rev α-helix peptide and RRE IIB has been clarified in Figure 1.8. Arg35, Arg39, Asn40 and Arg44 interact with nucleotides on the two opposite sides of the groove and likely form hydrogen bonds rather than electrostatic contacts.[41, 54] Thr34 and other six arginines (38, 41, 42, 43, 46 and 48) interact with the phosphate backbone via electrostatic interactions or hydrogen bonding. Thr34 also appear to interact with both peptide and RNA. Arg46, Arg48, as well as Arg50 at the C-terminus may have van der Waals contacts that help to orient the peptide.[54]

Later studies have determined an *in vivo* selected arginine rich peptide RSG1-2 (DRRRRGSRPSGAERRRRRAAAA) that has higher binding affinity and specificity to RRE compared with Rev peptide. An NMR study using a stable oligonucleotide structure derived from RRE IIB and peptide RSG-1.2 revealed that the major region in RRE IIB for peptide recognition is similar to Rev; RSG-1.2 penetrated deeper towards the RRE major groove and induced a conformational switch of the RRE U72 base.[55] This research suggested that the high binding affinity and specificity came from the deep interaction of peptide and RNA, possibly providing a new strategy for HIV drug development.

The stem-loop RRE IIB is the only high affinity binding site for Rev and initial studies indicate that Rev binds to RRE IIB in a one to one stoichiometry; other observations suggest that achieving full Rev function requires multiple Rev molecules binding to the RRE full sequence.[56-58] In fact, Rev could bind to RRE to form a hexameric complex that

is 500-fold tighter than any of the single interactions.[59] A correlation between the oligomer affinity and RNA export further support that Rev oligomerization is required for the Rev-RNA complex to be recognized by cellular components, which direct the export of the complexes.[59] Early studies have demonstrated that the high affinity binding site RRE IIB, RNA-binding ARM, and flanking oligomerization domains of Rev are essential for Rev oligomer assembly.[8] In addition to RRE IIB, Rev also utilizes different surfaces of the binding domain to interact with other low binding sites, such as stem IA.[59] These findings provide an initial understanding of Rev oligomer assembly, and a complete study of Rev oligomerization mechanism is needed. However, this task is a challenge due to limited Rev solubility.[60] A recent study points out that the oligomeric state and solubility of Rev are controlled by RRE.[61] In the presence of RNA or a RNA surrogate, Rev is highly soluble and allows the structure of the complex at desired concentrations to be examined. Furthermore, a Rev mutant, L60R, generates dimeric complexes that cooperatively bind to the extended RRE IIB RNA and increase specificity.[61] They form a distinct lobed structure with the disordered C terminus.[61]

To fully understand how Rev oligomerization facilitates the specific RRE RNA recognition and viral complex export, a crystal structure is indeed helpful. Recently, the Rev crystal structure with either Fab or RRE was elucidated in terms of explaining the structural basis of multimerization.[62-63] A modified Rev structure termed Rev_{70}-dimer was used in his study. Each monomer forms an antiparallel helix-loop-helix structure while the second monomer assembles face to face in a V shape fashion relative to the first monomer with a superimposable hydrophobic core region and an accessible RNA-binding ARM. Three amino acids, Leu18, Phe21 and IIe55, are identified as the critical residues, which

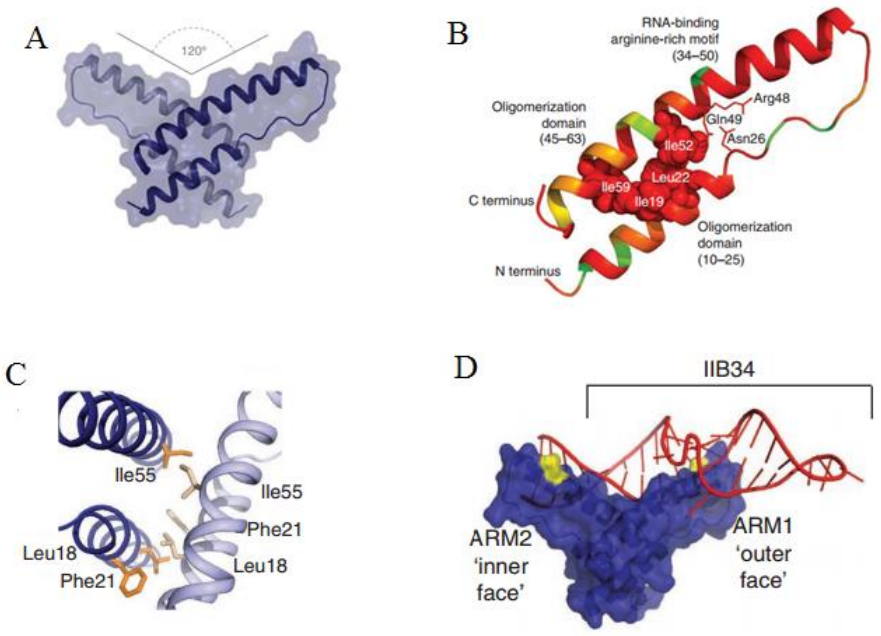


Figure 1.9 (A).Side view of Rev dimer and cross angle; (B). Each of the four Rev monomers has a folded core from residues 9–63, as shown, with structural and functional regions and amino acid numberings indicated; (C). Three hydrophobic residues mediate symmetric interactions at the dimerization interface; (D). The Rev dimer (blue) modeled with the 34-nt stem IIB (labeled IIB34; red) from the NMR structure of an ARM peptide–RNA complex. [Daugherty, M. D.; Liu, B.; Frankel, A. D., Structural basis for cooperative RNA binding and export complex assembly by HIV Rev. *Nat. Struct. Mol. Biol.* **2010,** *17* (11), 1337-42], Copyright 2010; [used with permission from Nature Publishing Group].

are buried in the hydrophobic core during the dimerization and stabilize the dimer by over 22 kcal/mol (Figure 1.9).[63-64] This dimer binds to RRE IIB RNA cooperatively with Asn40 in ARM of monomer 1 contacting RRE IIB and the other Asn40 in ARM of monomer 2 contacting the adjacent RNA binding site.[63] A Rev hexamer model is also generated using an arrangement of three dimers.[61, 63] The hexamer, along with the RRE full sequence, forms a jellyfish-like complex. On one side, the oriented ARMs are responsible for specific and cooperative binding to RRE; and on the other side, the disordered C terminus containing NES that are responsible for launching up to two nuclear export factor Crm1.[63]

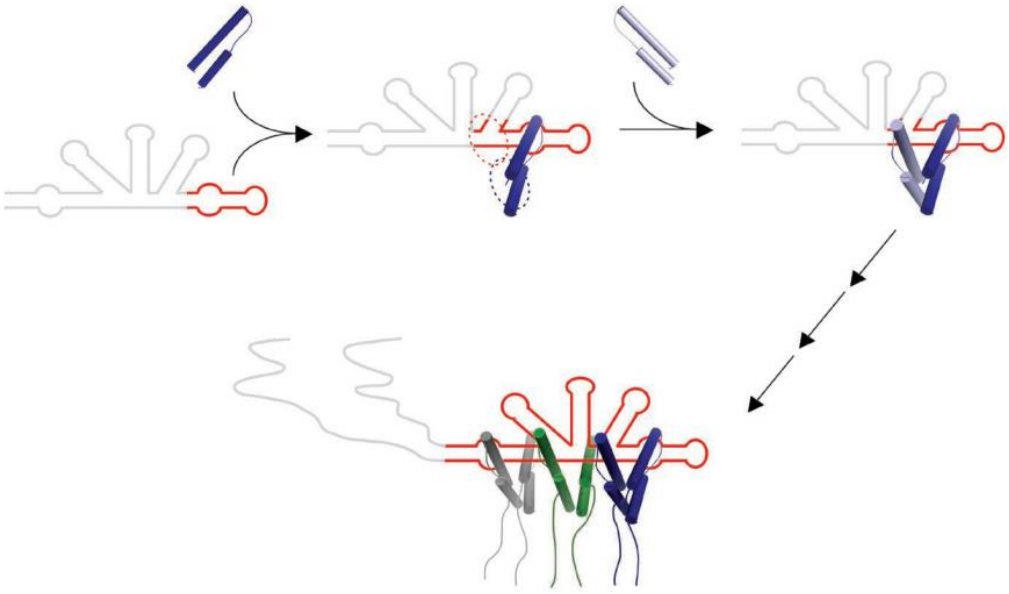


Figure 1.10 Models for oligomerization-mediated cooperative assembly. [Daugherty, M. D.; Liu, B.; Frankel, A. D., Structural basis for cooperative RNA binding and export complex assembly by HIV Rev. *Nat. Struct. Mol. Biol.* **2010,** *17* (11), 1337-42], Copyright 2010; [used with permission from Nature Publishing Group].

Binding of Rev to the RRE is believed to be initiated at the RRE IIB site. Upon binding, the second Rev binding site needs to be well oriented relative toward the RRE IIB site, which allows the second Rev to cooperatively bind at the initial complex through the exposed dimerization interface.[63] This requirement allows the Rev oligomer to distinguish the RRE from other RNAs. Further oligomerization and protein–RNA interactions complete the Rev hexamer assembly (Figure 1.10).[63] The model explains the specific recognition, cooperative RNA binding, and export complex formation very well; however, a complete crystal structure that includes the full length RRE is needed. Even though it is a challenge to obtain a crystal structure of the Rev–RRE complex, alternative methods are available to investigate the details inside this complex. A recent study utilizing small angle X-ray scattering (SAXS) revealed the fine structural features of the solution RRE structure.[65] A novel 'A' form RRE topology is proposed where the two major binding

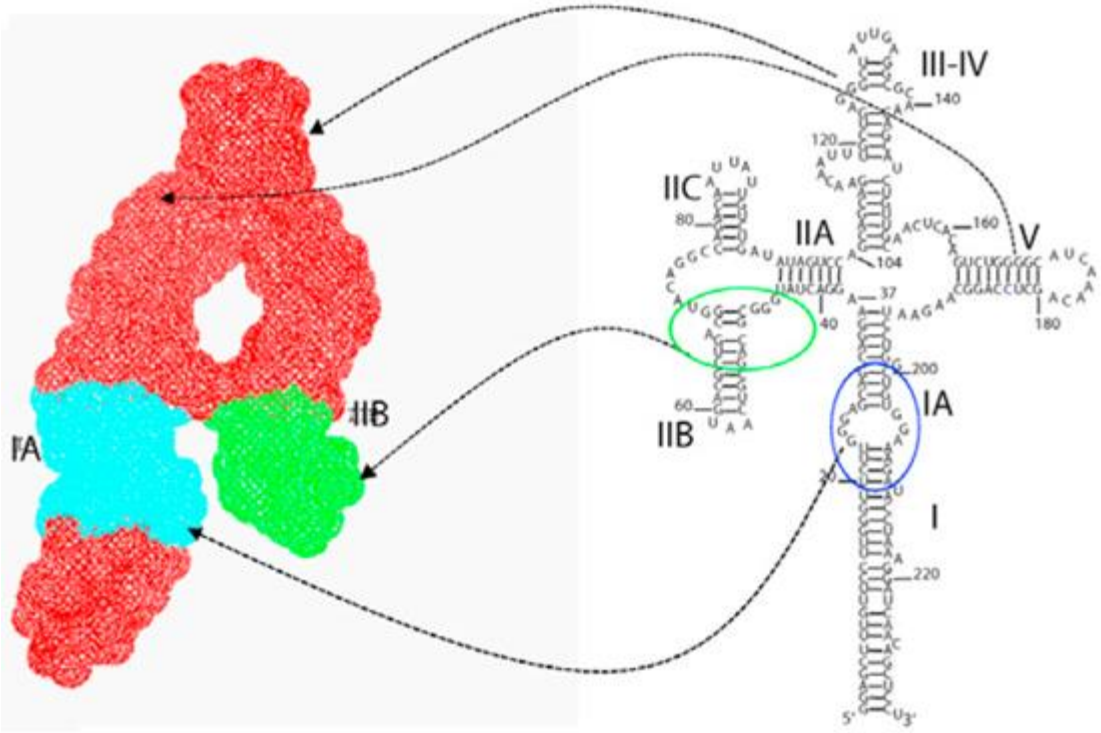


Figure 1.11 Summary of RRE domain locations in topological structures [Fang, X.; Wang, J.; Rein, A.; Wang, Y.-X., An unusual topological structure of the HIV-1 Rev response element. *Cell* **2013,** *155* (3), 594-605], Copyright 2013; [used with permission from Elsevier].

sites RRE IIB and IA are located in the separated legs.[65] The distance between these two sites is ~55 Å, similar to the span between two N-termini in the Rev dimer (Figure 1.11).[63] Deletion of either legs or increasing the distance between the two legs diminishes the ability of RRE to adopt the higher order Rev oligomer, which sequentially results in the loss of activity both *in vitro* and *in vivo*.[65] These observations are in agreement with previous findings that the RRE full sequence is required for export of viral genome.[57]

With this novel 'A' form structure of RRE elucidated, a new Rev oligomerization and binding mechanism has been proposed. Different from Frankel's model, the two legs span the space that allows the Rev to come into contact with both legs. They form a

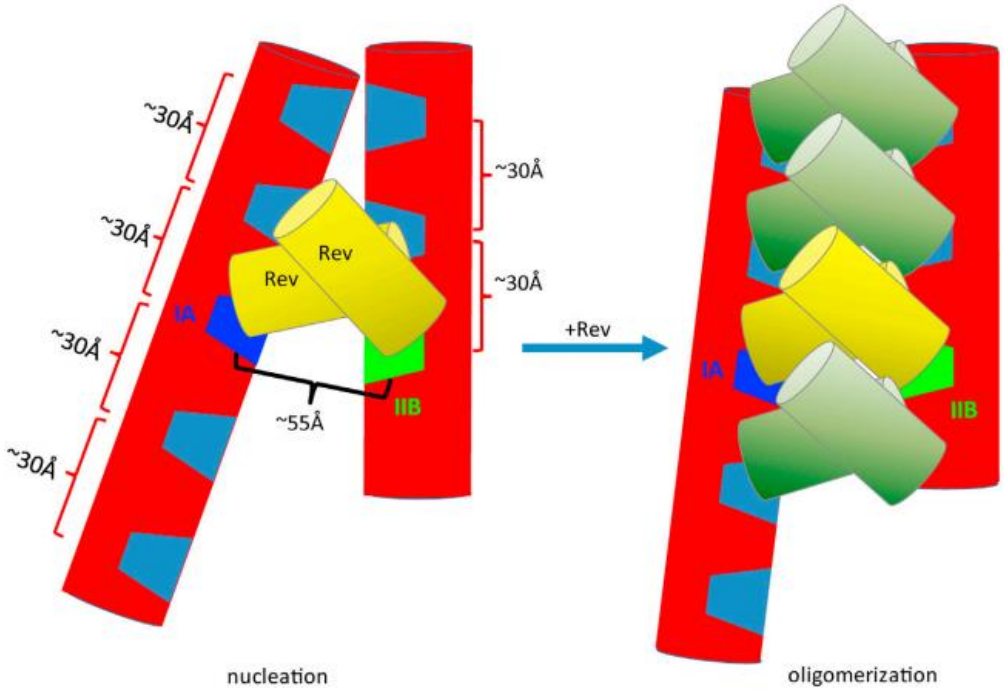


Figure 1.12 Models for initial Rev binding and Rev oligomerization on the RRE RNA [Fang, X.; Wang, J.; Rein, A.; Wang, Y.-X., An unusual topological structure of the HIV-1 Rev response element. *Cell* **2013,** *155* (3), 594-605], Copyright 2013; [used with permission from Elsevier].

sandwich-like structure with the Rev multimer (Figure 1.12). It is believed that the initial binding of the first two Rev monomers (yellow cylinders) to RRE IIB (green) and IA (dark blue) results in a nucleation site for subsequent Rev oligomerization on the RRE RNA. This oligomerization is partially driven by hydrophobic interaction between proteins and is constrained by the major groove (light blue) spacing and the arrangement of the two legs.[65] Overall, this global structural constraint explains the specificity and cooperative binding of multiple Revs that are required for full RRE function. The understanding of the Rev–RRE interaction at a single residue level may facilitate the discovery of new drug targets or strategies for a rapidly mutating virus like HIV-1. In fact, the three dimensional structures could provide detailed information not only on molecular interaction, but also on biological mechanisms. In this case, the overall architecture of the RRE–Rev

ribonucleoprotein has a critical role in the exposure of the NES for Crm1, and in organization of the complex to bind to other cellular proteins. As a result, the final complex could export the viral genome that is essential for viral replication. Thus, studies of the structural basis for Rev–RRE interaction benefit the exploration of new anti-HIV drugs, as well as understanding the basic knowledge of how the virus utilizes the oligomerization of proteins to recognize essential RNA.

1.5 Strategies for Targeting HIV-1 RNAs

The Tat–TAR and RRE–Rev interactions are essential for viral replication. Trans-activation inhibition of the promoter processed by Tat may force the virus to remain in dormant stages, where acute and chronic HIV infections may be inhibited. This may also result in less resistance.[66] In addition, disruption of RRE–Rev interaction either by binding to the Rev protein, RRE RNA, or RRE–Rev complex may inhibit the export of full length mRNA and singly spliced mRNA, which halts the replication during the translocation step. Both TAR and RRE have become high profile targets. In recent years, different approaches have been applied through interference with Tat–TAR and RRE–Rev interactions. There are several classes of inhibitors: intercalators, hybrid small molecules, modified macromolecules like PNAs, ribozymes, RNA decoys, siRNA, cationic peptides, and newly developed medium sized molecules.[32, 66-67]

1.5.1 Small Molecules That Bind to HIV-1 TAR

High throughput screening *in vitro* is used to select lead compounds that are inhibitors of the Tat–TAR interaction. Recent research has shown that small molecules such as aminoglycosides, aminoquinolones, and benzodiazepines can bind to TAR RNA

with nanomolar EC_{50} values. They show great potential as Tat–TAR interaction inhibitors *in vitro*.[68-70]

One series of compounds act as intercalators. Previously, ethidium has been shown to intercalate with both RNA and DNA containing G–C base pairs to a similar degree, which causes poor selectivity (Figure 1.13). A carboxylated ethidium was shown to strongly bind to the bulge region of TAR RNA rather than other similar RNA or DNA lacking the bulge region.[71] This binding could selectively disrupt the replication process of the virus with less cellular toxicity.[71] In order to enhance or direct the specificity of general intercalators for TAR RNA, an arginine was conjugated to ethidium via various linker lengths. Due to the different binding sites, these kind of molecules were considered to be bifunctional ligands. Indeed, one of these compounds, compound **3** was well studied through NMR, molecular modeling, and RNase footprinting assays. The results demonstrated that ethidium intercalates between G43–G44 and C18–C19 while the arginine residue interacts with the UCU bulge region (Figure 1.14).[72] It has been shown that the linker length is important for compounds to place the arginine residue into the correct position.[72]

Beside regular intercalators, a new type of inhibitor, In-PRiNts (inhibitor of protein-ribonucleotide sequences) that utilize the intercalator compounds has been developed. In-PRiNts have three components: a π-stacking system for interacting with the bases, a positively charged residue for interacting with the phosphate backbone, and a property linker to connect these two regions. One compound, CGP40336A, which features 6-chloro-2-methoxy acridine and spermidine, proved to be superior at inhibiting the Tat–TAR interaction with an apparent K_d of 150 nM for the wild type TAR sequence.[73] Gel shift

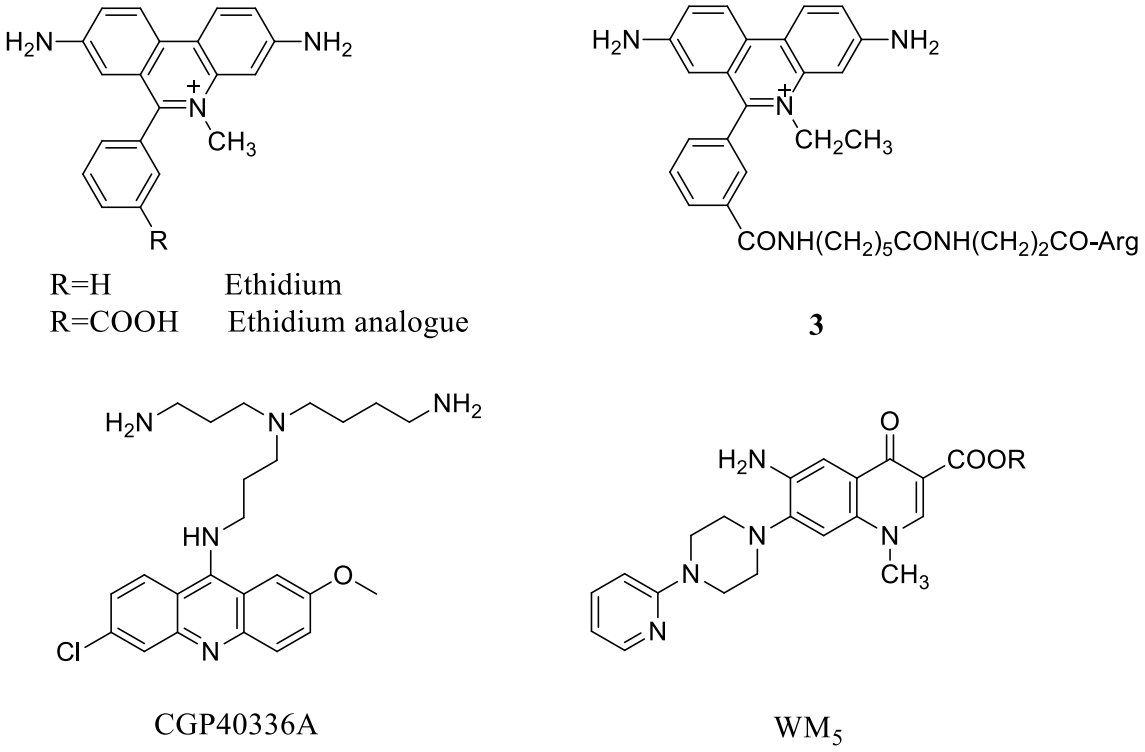


Figure 1.13 Inhibitors of Tat–TAR interaction.

assays, RNase foot printing, and NMR experiments revealed that CGP40336A binds to the bulge region of TAR, where acridine stacks between A22 and U23 while the secondary amine at the 9-position and the methoxy group at the 2-position form hydrogen bonds with the G26–C39 base pair.[73] Further mutation experiments on G26–C39 showed a decreased binding affinities, confirming that CPG40336A interacts with this region.[74]

As part of the development of anti-HIV drugs, a series of 6-aminoquinolone compounds were explored with an HIV-1 *in vitro* cell culture assay. Compound WM5, bearing a methyl substituent at the N-1 position and a 4-(2-pyridyl)-1-piperazine moiety at the C-7 position, was found to actively inhibit HIV-1 replication.[69] Later research investigated the mechanism of this inhibition and found that WM5 inhibited HIV-1 replication both in acutely infected cells and in chronically infected cells. WM5 showed great affinity to TAR RNA with a dissociation constant in the nanomolar range. Mutation

experiments of TAR RNA showed that WM5 mainly interacted with TAR RNA through the bulge region. However, it also showed an apparent affinity for protease and nucleic acids, which may explain its cytotoxicity.[75]

Several aminoglycosides have been shown to bind RNA motifs with high affinities. Among the aminoglycoside antibiotics, neomycin B is the most efficient inhibitor for disrupting the Tat–TAR RNA interaction, which is believed to occur through an allosteric mechanism (Figure 1.14). The IC_{50} value of neomycin is less than 1μM.[68] The subsequent work has found that it acts as a non-competitive inhibitor that directly binds to the Tat–TAR RNA complex. Circular dichroism (CD) spectroscopy and NMR

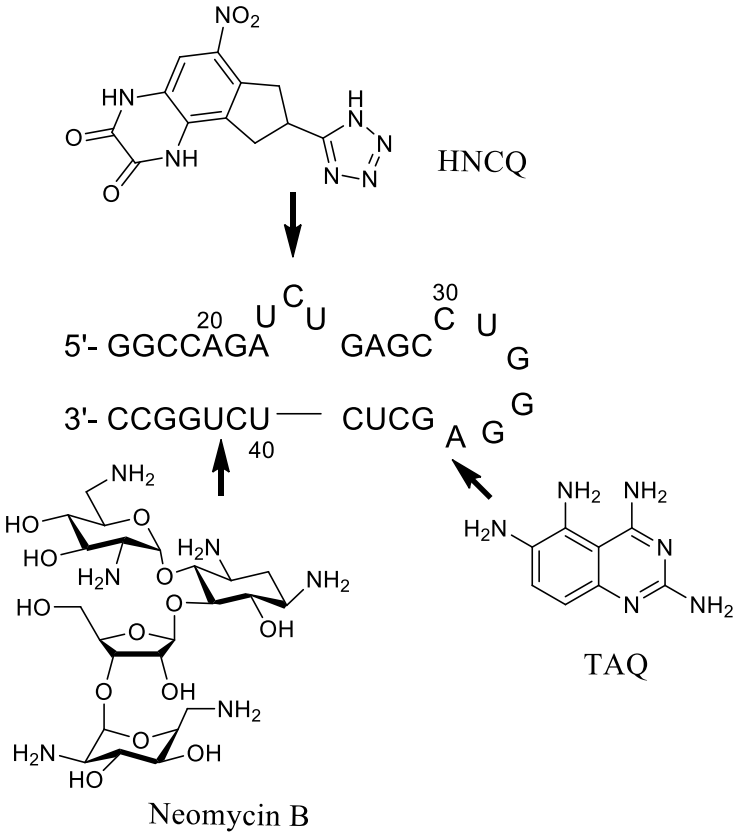


Figure 1.14 Primary TAR RNA binding sites of the three small molecules.

data show that neomycin binds to the lower stem immediately next to the three-nucleotide bulge region and induces a structural change in TAR. Binding to the Tat–TAR RNA complex increases the dissociation constant of the Tat–TAR complex.[68] This is considered to be valuable because the lead compound can associate with both TAR-RNA and the Tat–TAR complex and increase the efficiency of the inhibitors. Another two compounds, 2,3-dioxo-8-[2-(5-tetrazolyl)]-2,3,4,7,8,9-hexahydro-1H-6-nitrocyclopenta-[f]quinoxaline (HNCQ) and 2,4,5,6-tetraamino-quinozaline (TAQ) were compared with neomycin using a footprinting assay. They were found to inhibit HIV-1 by binding to two distinct sites of TAR rather than the lower stem region. HNCQ directly binds to the bulge region of TAR and acts as a competitive inhibitor of Tat. It also inhibits the Tat binding of HIV-2. Further assays support the conclusion that HNCQ preferentially binds to TAR as compared to calf thymus DNA and tRNA. TAQ is found to interact with the loop region of TAR. A Tat-activated reporter gene assay found that TAQ inhibits HIV-1 replication with an EC_{50} value of 19 μM, and no obvious cellular toxicity was observed.[76] These three molecules provide reasonable templates for further development of inhibitors for the Tat–TAR interaction.

The research on neomycin B brings great opportunity to study other aminoglycosides as well as their derivatives. Since the structure basis of the HIV-1 TAR–Tat interaction involves the ARM, other studies aim to conjugate various numbers of arginine into aminoglycoside to improve their affinity. The three compounds, R4GC1a (arginine-pure C1a gentamicin isomer conjugate), R3G (arginine-gentamicin conjugate) and R4K (arginine-kanamycin conjugate), stand out for their improved binding affinities (Figure 1.15). Gratefully, these small molecules display nearly no toxicity. The NMR studies and foot-printing assay indicate that they interact with both the upper stem next to

the bulge region of TAR RNA and the lower stem-bulge region of TAR, as well as the wide major groove of TAR, which is formed by the bulge region. [77,78]

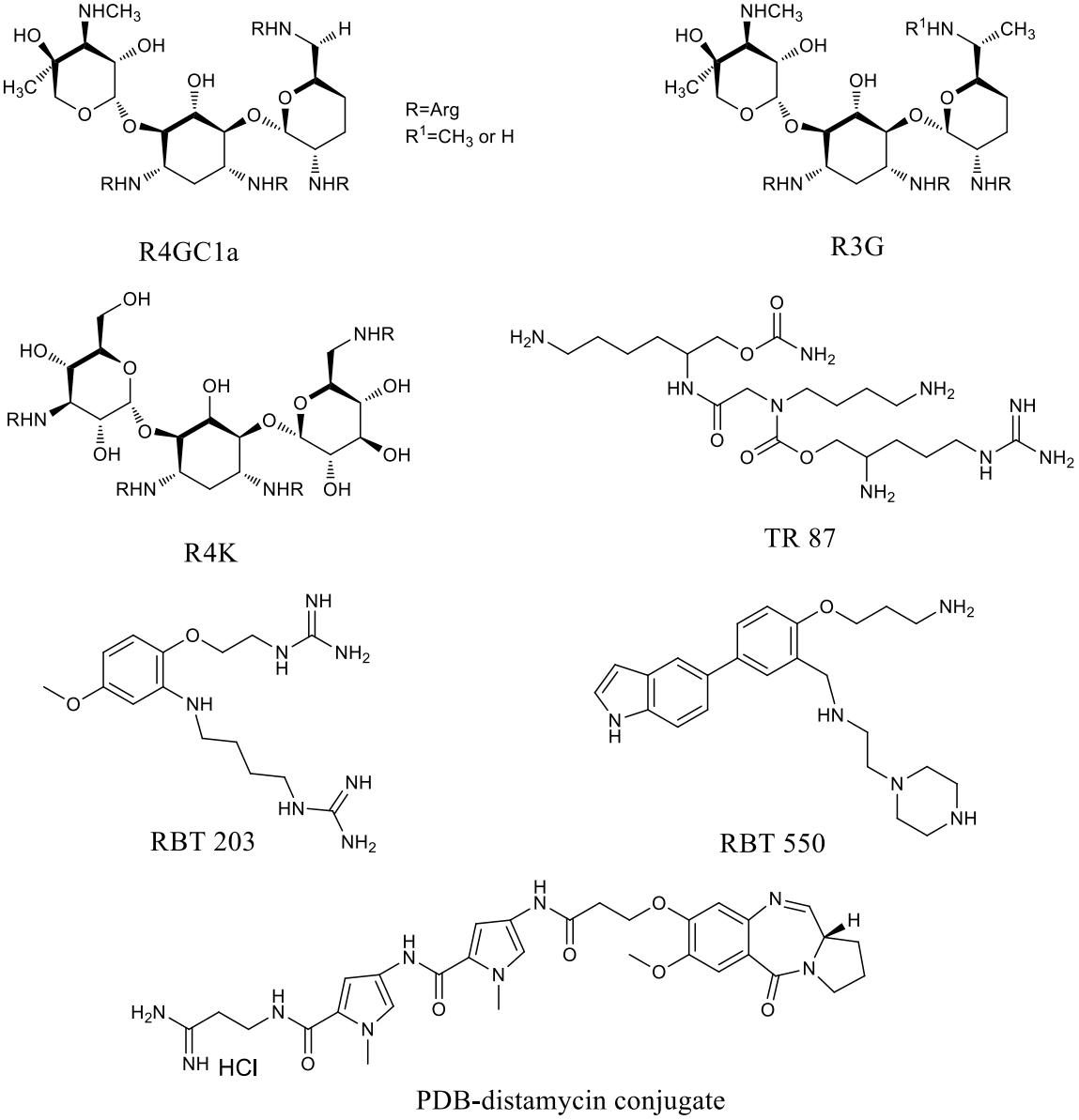


Figure 1.15 Inhibitors of Tat–TAR interaction.

Other research on exploring TAR inhibitors was based on high-throughput screen assays. Rana and co-workers screened a library containing 39,304 small molecules. They found 19 compounds that bound well to TAR RNA.[79] The *in vitro* assay showed that several compounds reduced the viral loads in infected cells at least by 5–20 fold in one or

two weeks. One lead compound, TR87, showed great potential in inhibiting Tat–TAR interactions (Figure 1.15). An *in vitro* study with Jurkat cells observed no cytotoxicity, and TR87 inhibited HIV-1 replication for over 24 days. An *in vivo* toxicity study with Taconic outbred albino mice showed that in five weeks all the mice remained healthy, and apparent weight change and toxicity were not observed. Fluorescence resonance energy transfer (FRET) was used to confirm that TR87 specifically interrupts the Tat–TAR interaction. Compared with a large number of potential inhibitors, its extremely low toxicity and high affinity for TAR RNA allowed it to be used in clinical trials.

In order to guide the scientists to develop TAR inhibitors, a Poisson-Boltzmann electrostatic surface potential was calculated based on previous research. The results showed that two electronegative spots exist within the bulge region.[80] With this in mind, Karn and co-workers developed a series of compounds containing a biaryl heterocycle on a benzylic scaffold.[80] The initial lead compound RBT203 was studied by NMR, and the results showed that RBT203 covers the major groove of the upper stem of TAR, and at higher concentrations of RNA, the binding becomes non-specific (Figure 1.15). Due to the different binding mode of biaryl residues, it was believed that guanidinium substituents may not be necessary for binding affinity. By mimicking the Tat-derived peptides and optimizing the previous library, another compound RBT550 was designed. It was demonstrated that the indole ring of RBT550 intercalated between the A22–U40 and G26–C39 base pairs while the amines contacted the phosphate back-bone. However, a cytotoxicity assay and a Tat–TAR reporter assay showed that they only exhibit two- to five-fold selectivity.[81]

Pyrrolo [2,1-c] [1,4] benzodiazepine (PBD) and distamycin are well known DNA binding drugs that have no affinity for structured RNA. It is surprising that the hybridized PBD-distamycin showed a preference for binding TAR RNA rather than A-T rich or G-C rich RNA structures. The results from *in vitro* experiment by EMSA and *ex vivo* experiments in the HL3T1 cell line showed that longer polypyrrole backbone results in higher affinity to TAR RNA. The inhibitor with four pyrroles is the most efficient one. However, a reverse effect was observed in the cytotoxicity assay. One compound, which contains 2 pyrroles was considered to be the most interesting as it exhibited the lowest toxicity (Figure 1.15). Further development of these hybrid compounds is necessary to produce clinical candidates.[70]

Although thousands of small molecules are studied for their inhibition of the Tat–TAR interaction, few of them are used in clinical trials due to their toxicity and lack of selectivity. Macromolecules provide us more options as Tat – TAR RNA inhibitors.

1.5.2 Macromolecules That Bind to HIV-1 TAR

Macromolecules such as oligonucleotides, ribozymes, polyamide nucleotide analogs (PNA) and short interfering RNA (siRNA) have shown potential inhibition of the Tat–TAR interaction with high efficiency and selectivity. Antisense oligonucleotide

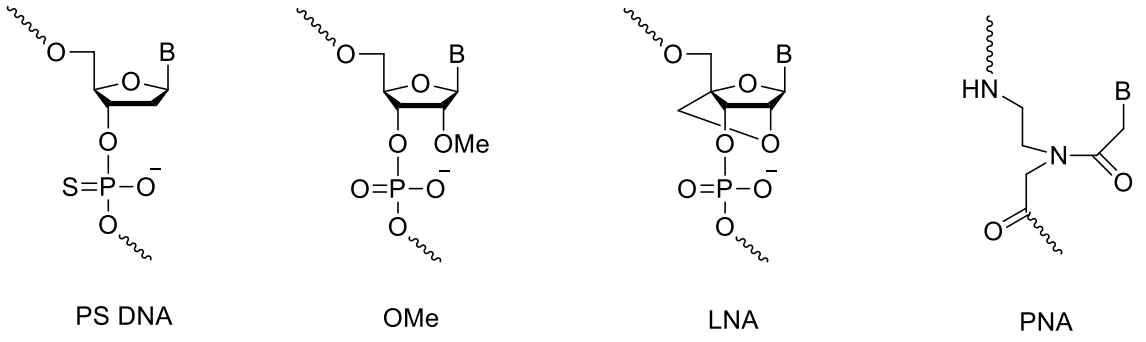


Figure 1.16 Oligonucleotides used as anti-RNA molecules

analogues inhibit the TAR–Tat interaction by strand invasion to form a pseudo-half knot structure and induce RNase H dependent destruction, which cleaves the DNA-RNA hybrid. The first experiment using antisense oligonucleotides as HIV-1 inhibitors was designed to form complexes with the viral primer binding site (PBS) by 12-26 long unmodified oligodeoxyribonucleotides. While it showed inhibitory activity at high concentrations, it was unstable to serum and cellular nucleases. Further improvement was made by replacing the oligonucleotide phosphate backbone with phosphorothioate (PS). This simple change increased its stability, as well as the sequence specificity and affinity while maintaining its recognition by RNase H. Although the oligonucleotides that induce RNase H cleavage inhibit gene expression efficiently, they are possible to miss bind to a similar sequence and result in cleavage. This will lead to off target mRNA inhibition effects.[82] In contrast, other approaches were achieved by modifying oligonucleotides that could inhibit translation by steric blockade of the ribosome. For example, several modifications were made by building 2′-O-methyl oligoribonucleotide analogues (OMe), locked nucleic acids (LNAs), and peptide nucleic acids (PNAs) (Figure 1.16). Their potent effects on inhibition of the Tat–TAR interaction were reported with HeLa cell nuclear extract. In this research, three types of the same length (12-mer) oligonucleotides, an OMe, a chimeric oligonucleotide containing 7xOMe and 5x5-methyl C LNA residues, and a PNA oligonucleotides were evaluated. They showed strong binding affinity to TAR and nearly equal inhibition *in vitro* (Table 1.1).[83] The affinity for TAR RNA was in the nanomolar range. A mutant sequence containing 1xOMe and 11x5-methyl C LNA residues was designed to increase the affinity and specificity according to previous research; however, it showed less binding affinity towards TAR and was less potent in the *in vitro* experiment. By using a cationic lipid

transporter, the oligonucleotides were delivered into HeLa cells, and only the 12-mer chimeric 7xOMe/5xLNA showed substantial, dose-dependent, and selective inhibition of firefly luciferase. This may be due to the strong uptake of the compound by HeLa cells.[83]

Table 1.1 Modified oligonucleotides and sequences.

Name	**Sequence 5'-3'**	**K_d (nM)***
12 TAR OMe	CUC CCA GGC UCA	60.7
12 TAR 7xOMe/5xLNA	CUC CCA GGC UCA	25.4
12 TAR 1xOMe/11xLNA	CTC CCA GGC TCA	89
12 TAR PNA	NH_2-CTC CCA GGC TCA-Lys-Lys-Lys-COOH	1.2

*Buffer condition: 50 mM Tris HCl, pH 7.4, 20 mM KCl

The major advantage of oligonucleotides is selectivity, and the introduction of LNA units enhances nuclear delivery and resistance to serum and nucleases. However, cell permeability is a major issue for oligonucleotides, which severely limits clinical studies.

PNA is an oligonucleotide that contains peptide bonds instead of the sugar phosphate backbone. As a result, PNA molecules are very stable in human serum and cell extracts, as well as proteases. Different from other oligonucleotides, the unrestrained polyamide backbones of PNA favor binding their targets in an orientation-independent manner.[84] PNA forms a more stable complex with DNA and RNA compared to a DNA-DNA or a DNA-RNA complex.[85] Even a single base pair mismatch will reduce the binding affinity which suggests that it recognizes the target sequence specifically. In previous research, a 15-mer (5′-TCCCAGGCTCAGATC-3′) PNA corresponding to the hairpin loop and the bulge region of TAR RNA was designed to inhibit Tat–TAR interaction (Figure 1.14). The 15-mer PNA showed great affinity for TAR RNA. Mutation of either bulge region, loop region, or upper stem will decrease binding ability or prevent binding to TAR. Further experiments revealed that the PNA inhibits Tat-mediated transcription both *in vitro*

and *in vivo* as well as in the presence of Tat protein. This suggests that PNA acts as a competitive inhibitor of TAR, preventing the binding of Tat. In the culture medium, it can block the Tat – TAR dependent expression of chloroamphenicol acetyltransferase (CAT) at 100 nM. Because PNAs are neutral, they have the potential to be cell permeable.[86] In order to increase cell permeability, a transporter conjugated PNA was developed. The membrane-permeating peptide vector was conjugated to a PNA to increase its cellular uptake. In the meantime, the affinity for TAR RNA *in vivo* remained similar to its value *in vitro*. These studies provided a cell permeable TAR–Tat inhibitor that is still efficient and specific.[87, 88]

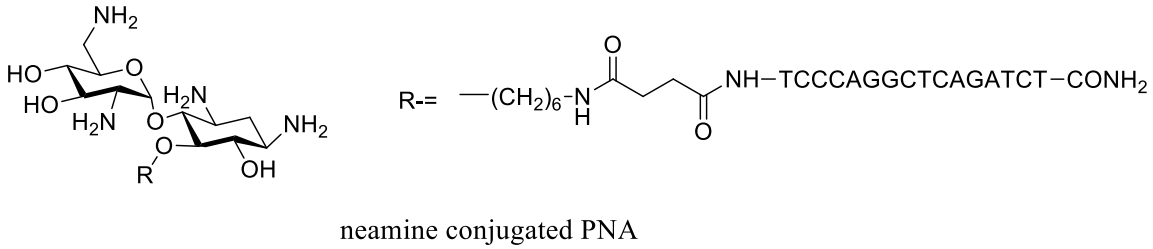


Figure 1.17 Structure of neamine conjugated PNA.

Although PNA had great success in inhibiting transcription selectively and efficiently, the solubility and cellular uptake was still a major issue. It must be conjugated with cell membrane penetrated peptides to approach its target. In order to optimize the design, a neamine conjugated PNA was developed to target TAR RNA (Figure 1.17). The neamine, which is a part of aminoglycoside neomycin B, has been shown to be an essential structural element that is involved in specific recognition of target RNA. The conjugation of neamine to the 5′ end of a previously reported PNA (5′-TCCCAGGCTCAGATC-3′) allowed the cellular uptake of the complex while maintaining the same binding affinity and inhibition efficiency. In addition, a unique RNA cleavage pattern was discovered that was

specific to its target site and functional at physiological concentrations of Mg^{2+}.[89] This new aminoglycoside conjugated PNA could be explored as a potential antiviral drug due to its high solubility and unique RNA cleavage properties.

The RNA–protein hybrid ribozyme is another approach to inhibit the Tat–TAR interaction. The activity of ribozymes depends on the expression, intracellular stability, target co-localization, and cleavage site access. When targeting a longer RNA, the efficiency is heavily dependent upon the secondary and tertiary structure of RNA because a majority of target sites may be inaccessible. To overcome this problem, Taira and co-workers designed a ribozyme linked to an RNA helicase that could relieve any secondary structure. They further introduced an RNA motif, a constitutive transport element (CTE), to the ribozyme that was demonstrated to interact with RNA helicase both *in vivo* and *in vitro* (Figure 1.18).[90] Among all hybrid ribozymes, the TAR Rz4 and TAR Rz5 were conformed to target the TAR region and inhibit the expression of the luciferase gene under

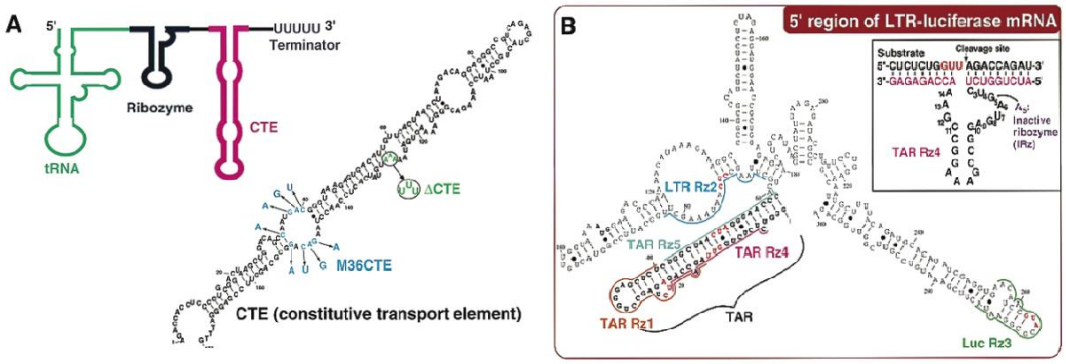


Figure 1.18 CTE-Rz design. (A)Schematic representation of a CTE-Rz (Left) and the secondary structure of the constitutive transport element (CTE) (Right). The sites of mutation that generate M36CTE and CTE are indicated by blue and green letters, respectively. (B)The secondary structure of the 5' region of long terminal repeat (LTR)-luciferase mRNA targeted by the indicated Rz. The sequence of the TAR Rz 4 is shown as a typical example. [Warashina, M.; Kuwabara, T.; Kato, Y.; Sano, M.; Taira, K., RNA-protein hybrid ribozymes that efficiently cleave any mRNA independently of the structure of the target RNA. *Proc. Natl. Acad. Sci.* **2001,** *98* (10), 5572-5577], Copyright (copyright 2001) National Academy of Sciences, USA].

the control of CTE. CTE improved the cleavage activity for the ribozyme and was also able to cleave TAR even when Tat was present. This was due to the recognition of CTE by RNA helicase.[90] Compared with previous ribozymes, the RNA-protein hybrid ribozyme can cleave the target RNA at any site.

RNA aptamers are oligonucleotides that can bind specifically to their target. In the case of TAR, the aptamers can strongly bind with the apical loop of TAR RNA through a loop-loop interaction to form a kissing complex. The aptamers can be selected by SELEX (Systematic Evolution of Ligands by Exponential Enrichment). A human RNA library was screened against TAR using genomic SELEX to see if a human transcript could interact with the retroviral genome. A genomic aptamer, a1, was found to interact with TAR through an apical loop complementary to five nucleotides of the hairpin loop of TAR (Figure 1.19). The affinity of a1 was four times higher than the previous aptamer R06.[92] NMR studies of the TAR–aptamer complex in solution demonstrated that this structure was highly stable, possibly due to the GA base pair in the loop closing position .[91] However,

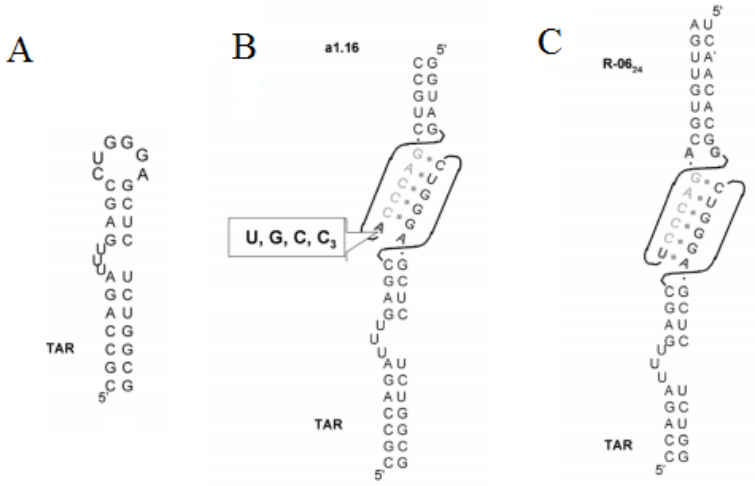


Figure 1.19 (A) TAR structures used for in vitro genomic selection; B) Kissing complex between TAR and truncated a1 genomic aptamers where mutations were introduced; C) Previously studied kissing anti-TAR hairpin R06.

in *in vivo* studies RNA aptamers were unstable to nucleases and were degraded rapidly. To overcome this issue, Toulme's group developed a series of 2′-OMe derivatives and found that these aptamers still recognize TAR RNA with similar properties, and selectively inhibit Tat-mediated transcription.[93] The modification was believed to increase the nuclease resistance of the aptamers and as a result, they showed higher efficiency in an *in vitro* experiment compared to previous aptamers. Further improvement is still needed to achieve higher efficiency, selectivity and stability.

Despite a significant amount of work demonstrating that macromolecules exhibit great affinity and high selectivity for inhibiting Tat–TAR interactions, little cellular data has been obtained. This is due to the lack of cell permeability of macromolecules. In addition, toxicity has been reported that might arise from the off-target effects.[94]

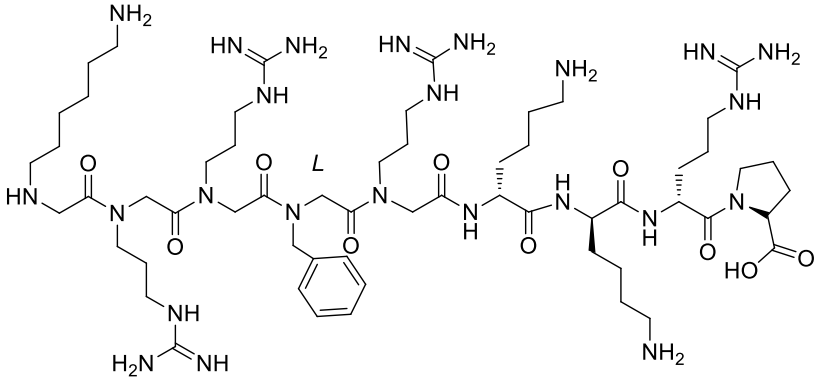
1.5.3 Medium Sized Molecules That Bind to HIV-1 TAR

Clearly, the design of TAR RNA inhibitors is far from straightforward. Both small and large molecules have beneficial properties. Small molecules can simultaneously achieve high affinity and cell permeability, yet they are not selective towards their target. Due to the limited surface area, they can insert into major groves of RNA and DNA easily, which may be the main reason for the lower selectivity and higher toxicity. Large molecules, on the other hand, heavily rely on base pairing so they do not distinguish the three-dimensional architectures. Additionally, they have poor cell permeability, and in some cases, they exhibit cell toxicity and poor stability profiles. Therefore, a series of compounds with molecular weight between small molecules and macromolecules (~1100 Da-2500 Da), so called "medium-sized" molecules were developed. These kinds of

molecules exhibit properties that allow them to interact with the RNA target via multiple contacts while also retaining cell permeability.

One major class of medium-sized Tat–TAR inhibitors is linear peptidomimetics. Most peptide inhibitors are designed based on mimicking the core region and basic domain of the Tat protein. Previously, a Tat-antagonistic compound (RKKRRQRRRK, Tat9-K-biotin), which contains the nine amino acid sequence of the basic domain of Tat protein, was reported to compete with Tat for binding to TAR. In Jurkat cells, Tat9-K-biotin reduced the expression of CAT, which is encoded by a Tat-dependent gene and is expressed from the HIV-1 long terminal repeat (LTR). Although Tat9-K-biotin successfully inhibited the expression of Tat *in vivo*, it showed little effect on CAT mRNA levels, which suggested that this peptide works at the post-transcriptional level.[95-96] Further investigation indicated that Tat9-K-biotin may prevent the association of mRNA from polysomes. As a result, it reduced gene expression of Tat-dependent genes.[96] This research shows that peptide analogues may inhibit HIV-1 replication at different steps.

Another approach to generate peptide inhibitors is with a combinatorial library. One hybrid peptoid/peptide oligomer of nine residues (CGP64222) (Figure 1.20) was



CGP 64222

Figure 1.20 Structure of CGP 64222.

selected from a pool of different chemical entities.[97] Peptoid and D-amino acids were used because of their flexibility and potential to form a unique secondary structure. A mobility shift assay showed that CGP64222 could inhibit the formation of the Tat–TAR complex at nanomolar concentration *in vitro*. NMR studies demonstrated that CGP64222 binds to TAR in a similar way to Tat and induces a conformational change at the binding site. Additionally, in a cellular trans-activation assay, inhibition was observed at 3-5 μM. Delightfully, no cytotoxicity was observed for CGP64222 up to 100 μM. It exhibited great biological stability.[97] Further investigation found that CGP64222 also inhibited the early step of the viral replication by selectively interacting with the CXC-chemokine receptor four coreceptor, and therefore, blocked viral entry into cells.[98] CGP64222 is the first example of a peptidomimetic compound that selectively inhibits Tat–TAR interaction. It appears to be a potential inhibitor of HIV-1 by inhibiting more than one step in viral replication.

Besides homochiral peptides, other studies have investigated the effect of peptide sequence chirality. A small family of mostly heterochiral tripeptides was selected from an on-bead screening assay.[99] Two peptides (L-Lys-D-Lys-L-Asn and D-Thr-D-Lys-L-Asn) exhibited dissociation constants of 420 nM and 564 nM, respectively. They were found to specifically bind to the bulge region of TAR RNA. Two diastereomers of the peptide L-Lys-D-Lys-L-Asn were also found in the pool. Their dissociation constants were shown to be 7 to 10 times higher than the peptide L-Lys-D-Lys-L-Asn, which may indicate that the tripeptide bound to TAR RNA is highly stereospecific. The best tripeptide L-Lys-D-Lys-L-Asn completely suppressed the transcription process in human cells with an IC_{50} of about 50 nM. *In vivo* studies demonstrated that these tripeptides were cell permeable, nontoxic

to cells, and may be able to interfere with gene expression.[99] Use of D- and L-amino acids together yields a rich stereo-chemical variety of ligands that induce the stereospecific interaction with its target RNA.

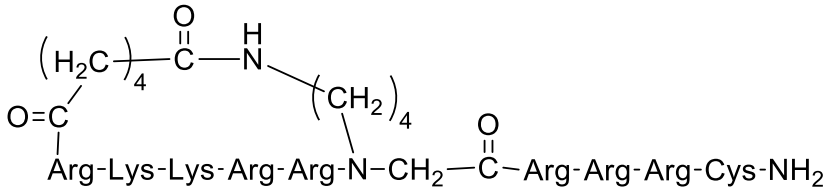


Figure 1.21 Structure of cyclic peptide Tat 11.

Cyclic peptides have been investigated to a greater extent. The initial approach used backbone cyclic ARM mimetic peptides. They were screened for their ability to mediate nuclear import, and from that, a peptide Tat11 (Figure 1.21**)**, was selected. Experimental data demonstrate that it can inhibit HIV-1 replication through inhibition of nuclear import and protein–RNA binding.[100] Later, a 3D structure of bovine immunodeficiency virus (BIV) TAR in a complex with a Tat-derived peptide was solved.[101-102] Although different from HIV-1 TAR, they share a highly similar sequence

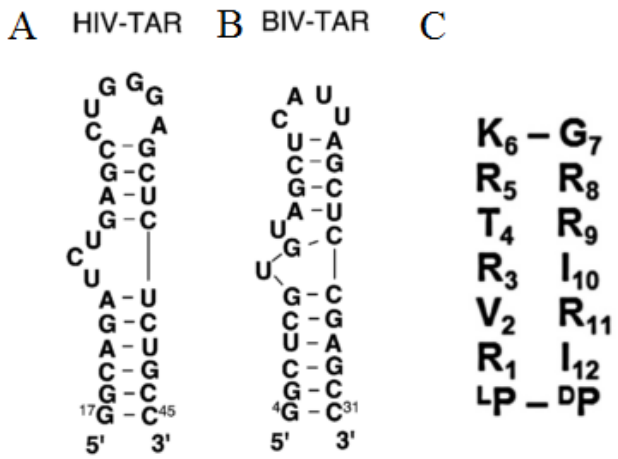


Figure 1.22 Structure of (A) HIV TAR RNA and (B) BIV TAR RNA; (C) template of cyclic peptides.

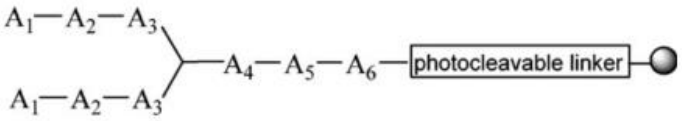
(Figure 1.22). In the BIV system, the structural study revealed a well-defined β-hairpin conformation of the bound Tat peptide, which aided the rational design of conformationally restrained β-hairpin peptidomimetics. Therefore, a small eight-membered peptide library of hairpin mimetics was designed. It contained 12 amino acids from Tat essential residues and mounted on the D-Pro-L-Pro template. The sequence BIV2 (RVRTRGKRRIRV) was found to successfully adopt a β-hairpin structure and had a K_d of 150 nM.[103-104] Remarkably, the interaction between BIV2 and BIV TAR was highly specific. This finding greatly influenced the design of HIV-1 TAR inhibitors. Further modification of the library led to the discovery of three β-hairpin peptide mimetics (L-22, L-50 and L-51, Table 1.2) that specifically inhibited the Tat–TAR interaction at the nM range.[105] Surprisingly, this interaction could not be disrupted by adding a 10,000-fold excess tRNA. Some of these peptides, for example L-22 and L-51, could even discriminate between HIV-1 TAR and BIV-TAR. NMR studies of the peptide L22-RNA structures indicated that both peptides interacted with the hairpin loop region and the bulge region. Binding of the peptide induced the formation of unusual deep groove that further stabilized the complex. Furthermore, multiple interactions such as hydrophobic, polar, and electrostatic interactions were observed in the binding complex. Competition experiments were performed to demonstrate that the cyclic peptides were specific and selective.[105] Indeed, the structure-related study provided an insight towards optimization of their antiviral activities. These peptides were shown to be potent HIV-1 replication

Table 1.2 Sequences of cyclic peptides and their dissociation constants to HIV TAR

Mimetic	1	2	3	4	5	6	7	8	9	10	11	12	K_d (nM)
L-22	R	V	R	T	R	K	G	R	R	I	R	I	30
L-50	R	V	R	T	R	G	K	R	R	I	R	R	1
L-51	R	T	R	T	R	G	K	R	R	I	R	V	5

inhibitors. Further characterization of L50 demonstrated that this peptide inhibits a wide range of HIV-1 isolate strains in different cell lines with an IC_{50} of 250 nM.[106] It exhibited cell permeability and low cytotoxicity. The inhibition mechanism studies revealed that L50 inhibits both HIV-1 reverse transcription and HIV-1 Tat-dependent mRNA transcription.[106] This is one of first examples that a drug inhibit HIV-1 replication with dual mechanisms.

The natural way that most biological systems form interactions between ligand and substrate is through polyvalent connections, which are defined as the simultaneous binding of multiple ligands to multiple receptors of one biological target.[107] These interactions were found to highly increase the affinity of ligands to receptors. Recently, a series of branched peptides were discovered by Santos and his colleagues from a branched peptide library that could bind to TAR RNA in a nanomolar range. The library was designed such that there were three variable amino acid positions at both the N- and C-termini (A1–A3 and A4–A6, respectively), and each variable position was composed of four possible side chains (Figure 1.23). Each of the four possible side chains was chosen for its potential to interact with the TAR RNA through different binding modes.[108] One of the peptides FL4 even



A_1	A_2	A_3	A_4	A_5	A_6
Asp	Arg	Tyr	His	Tyr	Phe
Arg	Asn	Leu	Val	Arg	Leu
Tyr	His	Trp	Asn	Ala	Asp
His	Val	Ala	Trp	Gln	Ser

Figure 1.23 Structure of branched peptide library.

bound to TAR much tighter than its native protein counterpart, Tat. Further mutation experiments demonstrated that branching in peptide ligands provides strong multivalent interactions with TAR RNA. These branched peptides displayed no cytotoxicity, and provided excellent cell permeability. A foot-printing assay revealed that FL4 may interact with the apical loop and bulge region.[109] This work provides a general platform to generating the selective RNA-binding ligands with branched arms, which utilize multivalent interactions to improve affinity and specificity.

Despite the great success that has been obtained from developing small molecules, large molecules, as well as medium sized molecules, none of the TAR RNA binding inhibitors has made it to clinic trail. One possible reason may be due to the lack of a HIV-1 Tat–TAR crystal structure. This further demonstrates the inherent difficulty in designing RNA binding ligands. Although the task is a daunting one, the exploration of new inhibitors is still ongoing.

1.5.4 Small Molecules That Bind to HIV-1 RRE

The RRE–Rev interaction is completely viral in nature, providing a high-valued therapeutic target completely independent from the natural cellular processes of the host. This is a huge advantage that would allow the interaction to be targeted selectively with minimal risk of side effects. Owing to the therapeutic potential of the Rev/RRE export pathway, many ligands have been designed to interrupt the RRE–Rev interaction, but with limited clinical success. The molecules that inhibit the RRE–Rev interaction can be divided into three major classes like TAR inhibitors. They are small molecules, macromolecules, and medium sized molecules.

Previous studies have suggested that several aminoglycoside antibiotics could interact with RNA in a sequence specific fashion.[110] They have been shown to act on ribosomal RNA.[110-112] Inspired from these studies, Green and co-workers screened 32 aminoglycoside and non-aminoglycoside compounds in the presence of Rev and RRE RNA to test their abilities to compete with Rev for binding with RRE.[113] Not surprisingly, none of non-aminoglycosides showed inhibitory activity at the tested concentration. Only three aminoglycosides, neomycin B, tobramycin, and lividomycin A, significantly inhibited RRE–Rev binding (Figure 1.24). The initial investigation showed the inhibition to be highly RRE–Rev specific. Adding unrelated sequence-specific RNA binding proteins or DNA binding proteins, even with similar arginine rich motif peptide

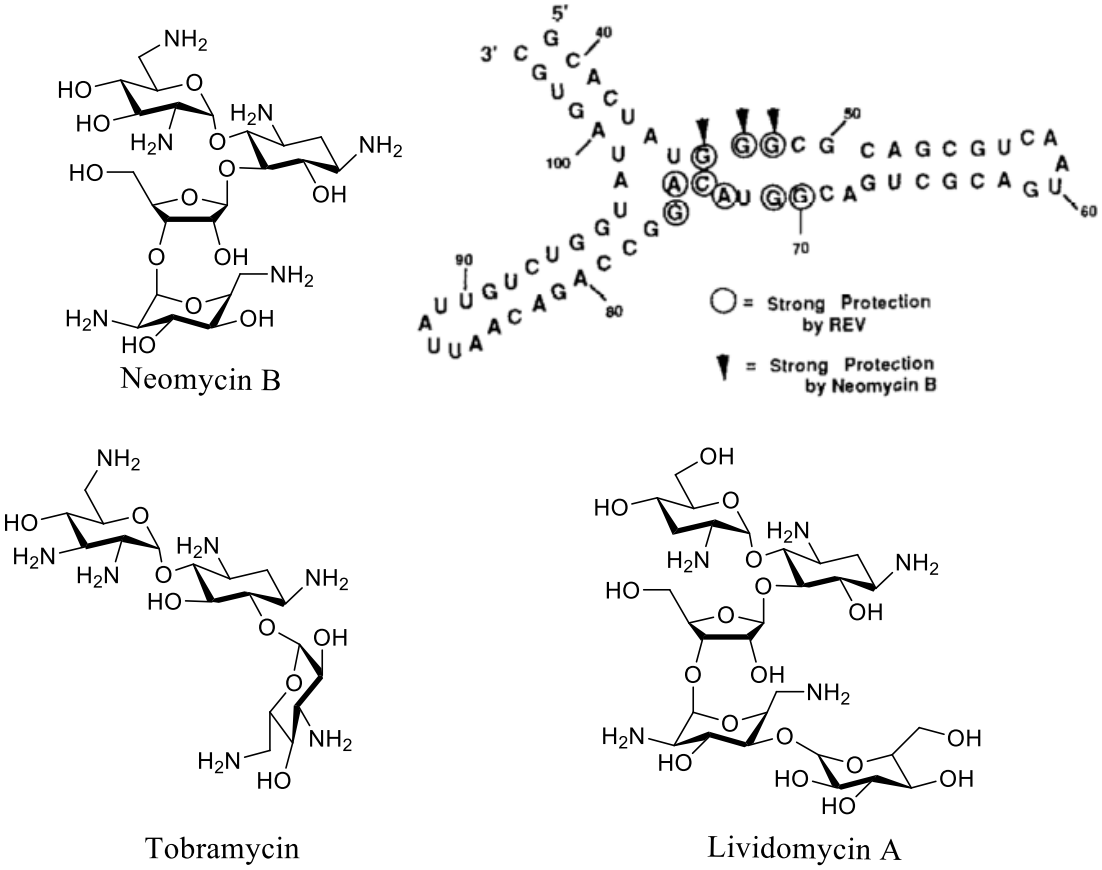
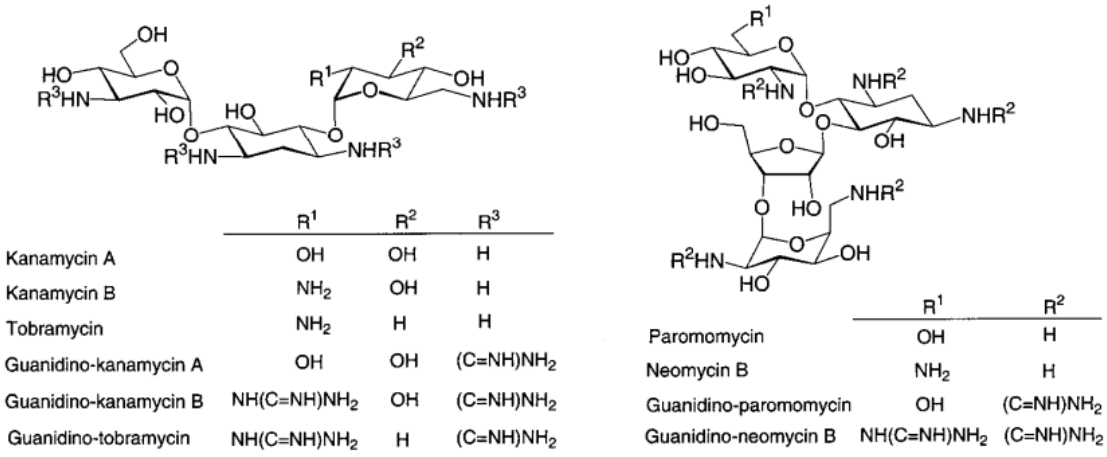


Figure 1.24 Structures of aminoglycosides and RRE RNA used in assay.

Tat did not disrupt the inhibition. Further experiments showed that neomycin B did not inhibit mammalian pre-mRNA splicing, which indicated that inhibition of RRE–Rev binding by neomycin B is highly specific. However, later research showed that neomycin B also interacts with the Tat–TAR complex by binding to the lower stem of TAR RNA, suggesting that it could not discriminate the secondary structure of RNA like the bulge or loop.[68] A foot-printing assay revealed that both neomycin B and tobramycin had a strong interaction with bases 46 to 48. Additional weak protections were observed within the IIA region.[113] These two aminoglycosides share similar protection patterns, especially within the core region. This indicated that they interact with core elements of RRE similarly to Rev. Fortunately, neomycin B and tobramycin showed specific inhibitory activities in both *in vitro* and *in vivo* experiments. Furthermore, neomycin B successfully inhibited the HIV-1 production in a cell line that was Rev-dependent in the activation of viral production. Although aminoglycosides are specific RRE–Rev inhibitors, their poor cellular permeability and potential cytotoxicity limit their use as antiviral drugs.[114]

It was believed that the general affinity for RNA is related to the ability of aminoglycosides to bind RNA through electrostatic interactions mediated by ammonium groups. To approach the same function, guanidinium groups are potential groups due to their highly basic, planar character and their directionality in their H-bonding interactions. The guanidinylation of polyfunctional amines of kanamycin A, kanamycin B, tobramycin, paromomycin, and neomycin B resulted in a 5- to 10-fold increase in inhibitory activity relative to the parent compounds (Figure 1.25).[115] A solid phase displacement assay indicated that kanamycin A, kanamycin B, tobramycin, and paromomycin had a binding preference for RRE over poly r[A]-r[U]. Only guanidinylated neomycin B was less

selective for RRE in the presence of poly r[A]-r[U]. Overall, the guanidinoglycosides showed very little affinity for double-stranded DNA.[115] The ability of guanidinoneomycin B and guanidinotobramycin to inhibit viral replication in HIV-1-infected $CD4^+$ HeLa cells was measured.[116] They inhibited HIV replication of activities approximately 100 times higher than their parent compounds. The antiviral activity of the guanidinoglycosides may be related to their ability to bind the RRE with high affinity and specificity, thus preventing the HIV-1 RRE–Rev interaction and viral replication. The transformation of aminoglycosides into guanidinoglycosides influences the biological activity of glycosides and implicates their therapeutic potential as antiviral drugs.



	R^1	R^2	R^3
Kanamycin A	OH	OH	H
Kanamycin B	NH_2	OH	H
Tobramycin	NH_2	H	H
Guanidino-kanamycin A	OH	OH	$(C{=}NH)NH_2$
Guanidino-kanamycin B	$NH(C{=}NH)NH_2$	OH	$(C{=}NH)NH_2$
Guanidino-tobramycin	$NH(C{=}NH)NH_2$	H	$(C{=}NH)NH_2$

	R^1	R^2
Paromomycin	OH	H
Neomycin B	NH_2	H
Guanidino-paromomycin	OH	$(C{=}NH)NH_2$
Guanidino-neomycin B	$NH(C{=}NH)NH_2$	$(C{=}NH)NH_2$

Figure 1.25 Structures of aminoglycosides and guanidinylated aminoglycosides.

Earlier studies have evaluated the binding affinity of several heterocyclic aromatic compounds to RNA. In particularly, diphenylfuran derivatives have been shown to bind RNA in a structure dependent manner.[117-118] A series of heterocyclic compounds have been tested for their ability to inhibit RRE–Rev interactions. The majority of these compounds contained a furan core, and all of them contained two or more benzene rings attached with alkylamine substituents. Their inhibitory ability was evaluated by gel shift assay. One compound, DB182, was identified as a potential inhibitor with an IC_{50} around 0.1 μM,

which is 10 times lower than neomycin B (Figure 1.26). Replacing the central furan ring with a pyrrole or increasing the number of substituents failed to optimize the inhibitory activity. Further SAR investigations revealed that both a heterocyclic core and aromaticity are critical for binding. More detailed interactions of DB182 with RRE were identified with a foot-printing assay. Different from aminoglycoside compounds, DB182 strongly protected nucleotides A44, U99, and U102, which are located within stem IIA of the RRE. Binding of DB182 may disrupt these base pairs, preventing recognition and binding of Rev. Surprisingly, these nucleotides were not protected by Rev and neomycin B, suggesting that DB182 may interact with RRE in a different manner.

To further improve DB182, four different parts of the scaffold were synthesized.[119] These groups included cationic substituents, phenyl rings, central furan heterocycles, and the position of the cationic side chain (Figure 1.26). It is clear that the inhibition activities decreased when the chain length is shortened, the terminal amino group is removed, or the diethylamino group is replaced by a cyclic amino group. Hence, the position of the charge and the size of the alkylamine group, as well as the number of charges affect the activity of inhibitors. In this experiment, they used a more stringent binding condition which yielded a relatively higher IC_{50} (5.1 μM). Further modification of the furan ring to the corresponding thiophene DB316 and oxazole DB449 improved activities. When they replaced one phenylfuran with a benzimidazole ring, a dramatic increase in inhibitor activity was observed, which yielded the best compound, DB340. DB340 bound strongly to RRE and increased the T_m dramatically. Other modifications were made on the position of the alkylamine substituents. Even though there was no apparent trend with this modification, changing the *para* cationic substituents in DB182 to the *meta* position

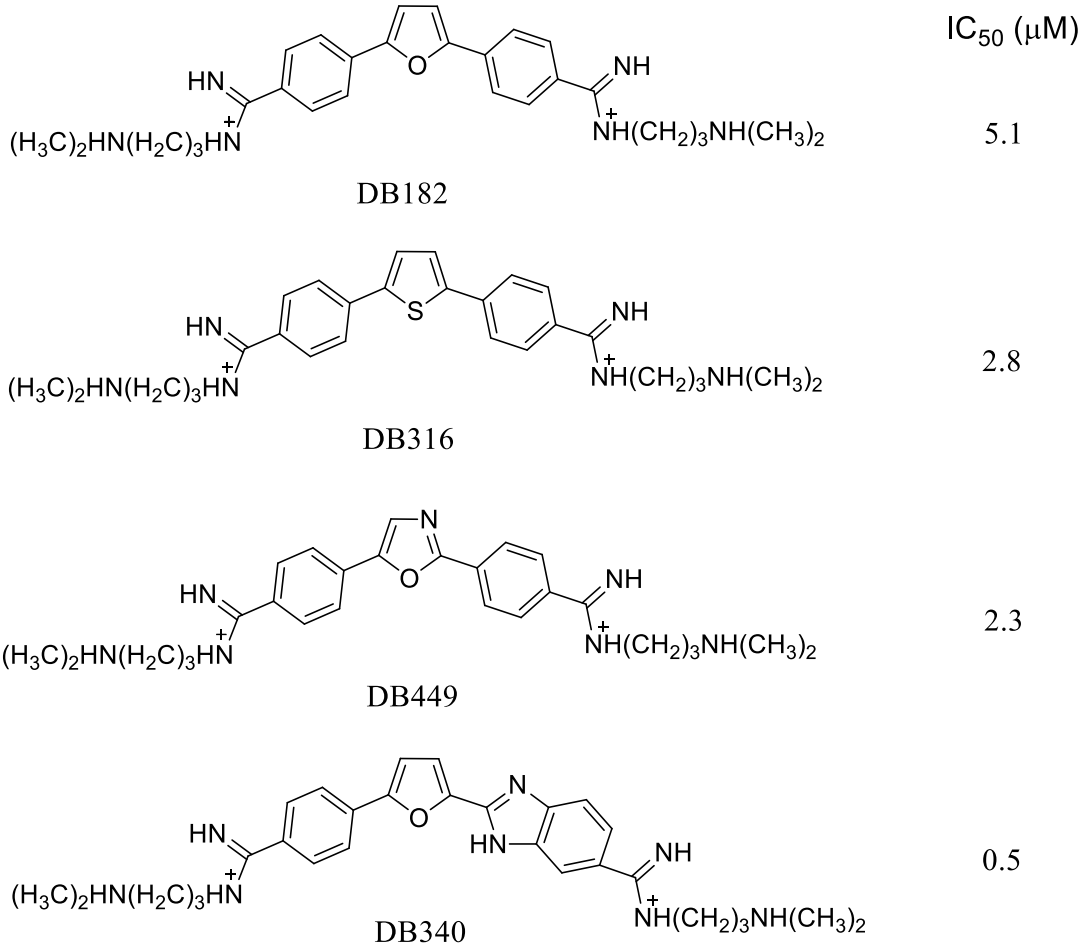
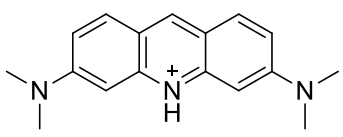


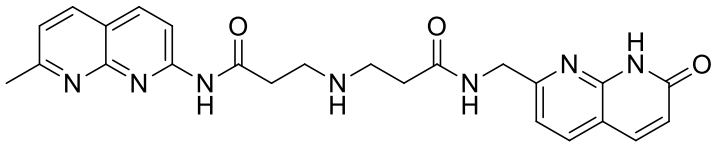
Figure 1.26 Structures of diphenylfuran derivatives.

significantly improved activity. Melting experiments showed that there was a non-linear relationship between T_m and activity. In general, the higher the T_m is, the better the compound activity. Overall, DB340 had promising binding affinity to RRE and was the most potent compound among the investigated compounds. Subsequent NMR studies indicated that DB340 binds to RRE as a dimer.[120] The dimer is placed in the minor groove of RRE near G–A base pair at the end of the internal loop. A series of biological assays found that DB340 binds the internal loop, inducing a conformational change in the upper stem and hairpin loop.[120] DB340 binds to RRE as a dimer in a highly structured and cooperative complex in the RRE minor groove.

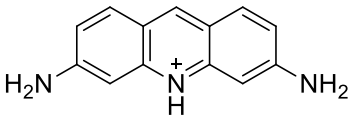
In addition to the focused studies on aminoglycosides and diphenylfuran cations, other small molecules like proflavine were investigated. Initially, twelve acridine-like compounds were screened due to their potential binding to RNA, as well as their drug-like chemical properties.[121] Among these compounds, proflavine and acridine orange showed the most significant decrease in fluorescence in a 2-aminopurine based assay (Figure 1.27). The change in fluorescence usually indicated a direct or indirect interaction with 2-aminopurine-labeled base.[122] The NMR studies revealed that proflavine binds tightly to a single binding site on RRE-IIB as a dimer in a stacked manner. It was believed that proflavine first occupies the bulge region with high affinity followed by a second proflavine binding to the same region with low affinity. The binding of the dimer completely inhibits the formation of the short rev peptide-RRE IIB complex, which suggests that proflavine induces a conformational change in RRE IIB that diminishes the binding site where its native protein binds. Two competition experiments revealed that proflavine bound to RRE IIB competitively with an IC_{50} of 0.1 μM.[122] Although proflavine was considered to belong to a class of nonspecific nucleic acid intercalators, these binding activity suggests that they can also be used as RRE inhibitors.



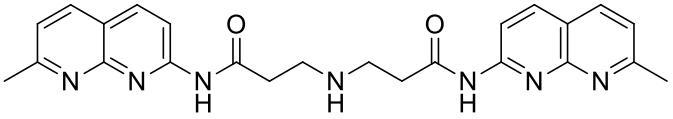
Acridine orange



Naphthyridine-Azaquinolone (NA)



Proflavine



Naphthyridine dimer (ND)

Figure 1.27 Structures of small molecules that target RRE RNA.

Due to the inherent difficulty of targeting complicated three dimensional RNA structures, other researchers have first applied their strategies using a DNA model. Nakatani and co-workers discovered a naphthyridine compound that selectively recognizes a single guanine bulge of duplex DNA.[123] In order to improve their specificity, two other naphthyridine hybrid compounds have been discovered. The naphthyridine dimer (ND) can selectively recognize a G-G mismatch via stacking with one guanine and complementary hydrogen bonding with the other guanine (Figure 1.27). [124] Based on the same concept, another hybrid compound, naphthyridine–azaquinolone (NA) was designed to recognize G-A mismatches.[125] Azaquinolone adopted complementary hydrogen bonding to adenine that might account for the high binding affinity. The melting temperature of a single G–A mismatch DNA duplex was significantly increased in the presence of the NA compound. Because the core recognition region of RRE adopted a G–G and a G–A mismatch, ND and NA can be potentially used as specific ligands that bind to the RRE IIB internal loop. In fact, the evaluation of these compounds on binding to a hairpin RRE IIB RNA, a completely matched RNA sequence, a bulge r(UCU) containing TAR RNA, and a single strand RNA through surface plasmon resonance (SPR) revealed that ND and NA do not recognize completely matched RNA duplexes, a r(UCU) bulge, as well as the single strand RNA.[126] However, further experiments with various RNA mutants suggested that ND also favored the G–G mismatch flanked by a single nucleotide bulge. It was believed that the single nucleotide bulge may loosen the major groove of the G-G mismatch RNA and provide an ideal pocket. A fluorescence anisotropy displacement assay revealed that ND could competitively bind with RRE IIB in the presence of a short Rev peptide. It is more efficient to induce RRE–Rev dissociation than NA and neomycin B. It is clear that the

naphthyridine hybrid compounds show potential inhibitory activity towards HIV-1 RRE. However, none of their K_d or IC_{50} values have been reported.

1.5.5 Macromolecules That Bind to HIV-1 RRE

PNA molecules have been demonstrated to bind oligonucleotides in an efficient and sequence-specific manner.[84-85] Due to the highly flexible backbone of PNA, it is hard to form conformational structures like peptides and proteins. In order to enhance its ability to recognize an RNA specifically, Mihara and co-workers designed a novel RNA-binding molecule using a short rev peptide that adopted an α-helix structure combined with PNAs that can specifically interact with the bases in RRE IIB.[127] Because the length of the PNA monomer equals two amino acids, an extra glycine was attached to the C-terminus for reference compounds. They replaced Arg50, which was shown to interact with U72, to PNA monomers bearing different nucleic acids (Figure 1.28). It is clear that only the adenine PNA monomer showed comparable binding affinity towards RRE RNA as the positive reference compound. Further placement of the PNA monomer in the middle of the Rev sequence showed a significant decrease in the K_d values, as well as the α-helix content. These results suggest that the adenine PNA monomer interacts with

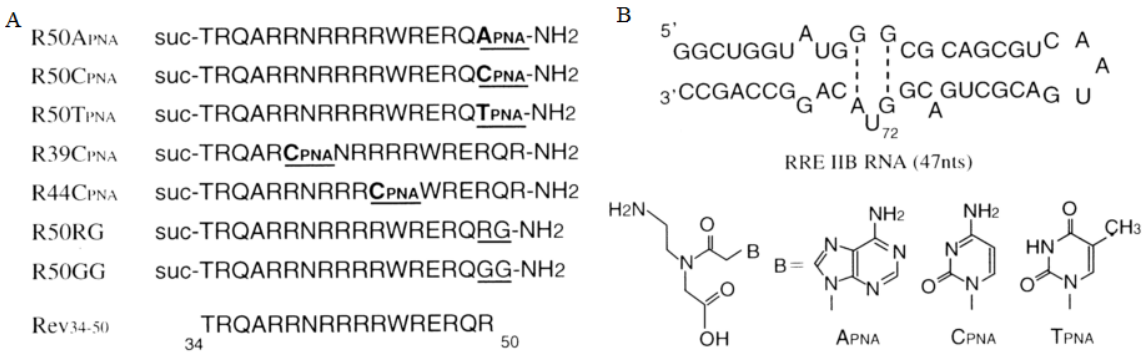


Figure 1.28 (A) The sequence of PNA peptides and (B) structures of RRE IIB and PNA.

U72 through specific contacts rather than hydrophobic interactions. In addition, the position of the PNA monomers is important since PNA can disrupt the structure that the peptide adopts. This strategy provides a possible solution to the development of artificial molecules that recognize RNA with high specificity and affinity.

Encouraged by previous results, the same group developed a series of N-terminal conjugated PNA–peptide hybrids.[128] Starting with the model sequence, $Rev_{34\text{-}50}$, one residue glutamine Q36 was modified in this research due to its role in contacting the base G48.[54] Initially, four PNA monomers (A, T, G and C), as well as a reference compound, acetyl(2-aminoethyl)glycine (agg), were introduced at the N-terminus of the $Rev_{37\text{-}50}$ peptide that had three deleted residues (Figure 1.29). Each of these sequences displayed increased binding affinity compared to the $Rev_{37\text{-}50}$ peptide. Furthermore, their K_ds were comparable with each other, which suggested that the elongation of chain length was important. Addition of a second and third monomer further improved the binding affinity towards RRE via a sequence-specific manner. Only the sequence of CGC PNA conjugated peptides showed the highest binding affinity toward RRE; this may indicate

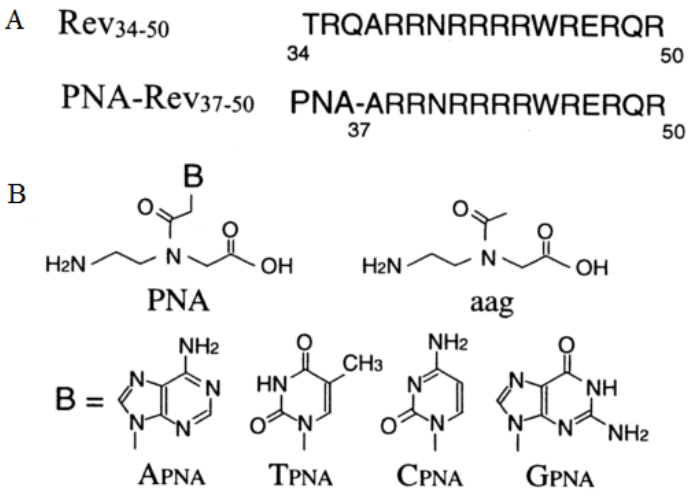


Figure 1.29 (A) Sequences of $Rev_{34\text{-}50}$ and PNA $Rev_{37\text{-}50}$; (B) PNA structures.

that the CGC PNA interacts with the G48, C59 and G50 sequences of RRE specifically. Finally, the sequence GCGC PNA and TGCGC PNA showed a 15-fold tighter binding affinity for RNA. Deleting either the peptide sequence or the PNA sequence resulted in decreasing binding affinities. CD characterization revealed that the PNA conjugated peptides displayed a strong α-helix content and validated that conjugating PNA at the terminus of the peptide will not disrupt the secondary structure of the peptide.

Nature always uses chiral molecules to help them be recognized specifically in their complicated system. For example, most enzymes are composed of L-amino acids. They are highly specific and selective to recognize and bind to their target. Compared with achiral PNA units, introducing chiral nucleobase amino acids has more advantage in recognizing target conformations.[127] Based on Mihara and co-workers' previous work, they continued exploring compounds that can specifically bind to HIV-1 RRE RNA with chiral nucleobase amino acids (NBA) (Figure 1.30).[129] The L-α-amino-γ-(nucleobase)-butyric acid was chosen to carry the nucleobase due to its flexible side chain and ease in operation during solid phase synthesis. With the well characterized $Rev_{34\text{-}50}$ peptide as the template, Gln36 and Asn40, which were reported to interact with G47-A73 specifically, were replaced with different nucleobases, respectively. It is not surprising that the

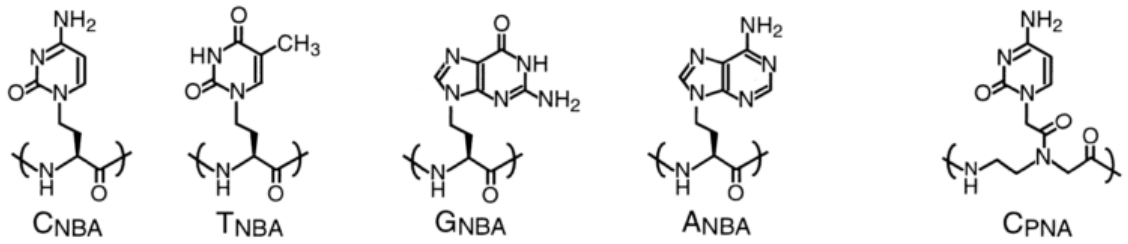


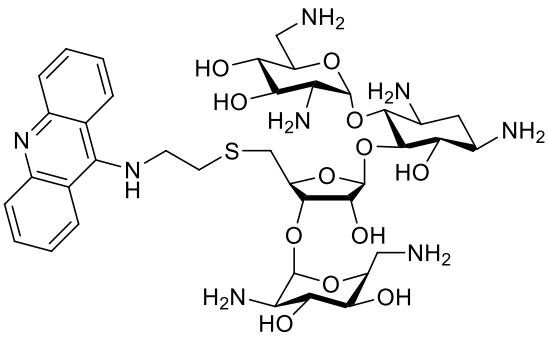
Figure 1.30 Structures of nucleobase amino acids and traditional PNAs

sequence N40C_{NBA} exhibited a similar K_d as Rev_{34-50}, while sequence N40C_{PNA} showed 50-fold higher K_d when compared to the sequence N40C_{NBA}. Furthermore, replacing C_{NBA} with other nucleobases resulted in decreasing binding affinities. CD spectroscopy showed that introducing one nucleobase increased α-helical content. Although introducing the second nucleobase slightly decreased the α-helical content, it still maintained a comparable α-helical content with Rev_{34-50} peptide. It is gratifying that replacing both functional amino acids with proper nucleobases results in ligands with comparable binding constant. However, none of these nucleobase amino acid peptides exhibited significant increases in binding affinities towards RRE IIB. However, this study led to a new strategy applicable to the construction of molecules that specifically recognize structured RNAs with various NBA unites on peptide structures.

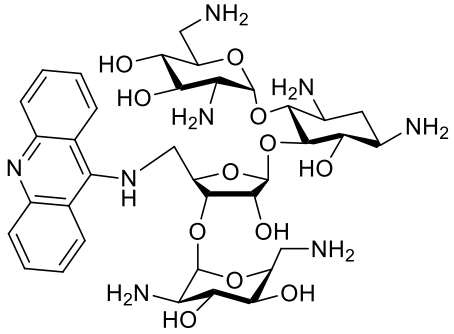
1.5.6 Medium Sized Molecules That Bind to HIV-1 RRE

Over the last few decades, medium sized molecules have gained sufficient attention due to their high potency on recognizing RNA structures and cellular permeability, as well as low toxicity.[130-132] As previously discussed, aminoglycoside antibiotics are highly selectively for RNA over DNA but bind to a wide range of unrelated RNAs.[111-112] In an attempt to improve the binding affinity and selectivity of the binding of aminoglycosides to RRE RNA, a series of neomycin B derivatives and dimers were synthesized and evaluated. Tor and co-workers first conjugated an acridine monomer to neomycine B via a thioether linker (Figure 1.31).[133] The new compound, neo-S-acridine, showed a 10 times stronger binding affinity compared to neomycin B in a displacement assay through a gel shift mobility assay. Further fluorescence anisotropy competition experiments also confirmed that neo-S-acridine exhibited a higher binding affinity towards RRE than both

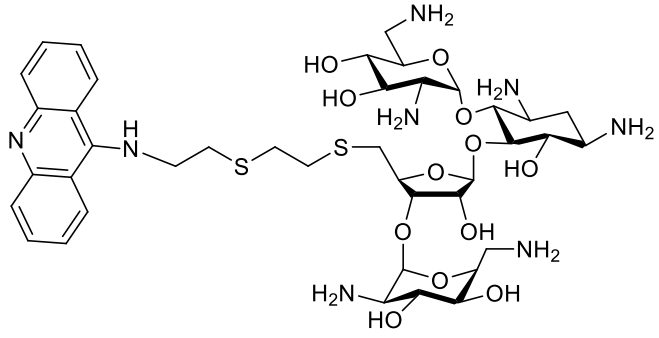
neomycin B and 9-aminoacridine, and the binding affinity is slightly lower than Rev peptide. To locate the binding site of neo-S-acridine on the RRE, enzymatic protection experiments were conducted. Both Rev peptide and neo-S-acridine protected the G46-G48 bulge region as well as U72. This suggested that neo-S-acridine competes for the same binding site that Rev binds and also disrupts the G48-G71 base pair. In addition, protection of the nucleotides A44 and C74 was not observed with the Rev peptide.[133] However, later



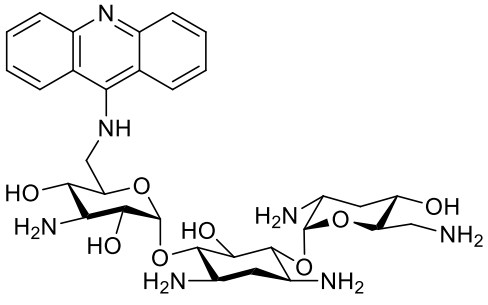
Neo-S-acridine



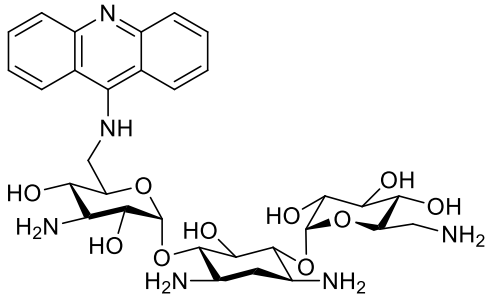
Neo-N-acridine



Neo-C-acridine



Tobra-N-acridine



KanaA-N-acridine

Figure 1.31 Structures of acridine-aminoglycoside conjugates

exploration of neo-S-acridine indicated that it maintained the selectivity for RNA over DNA but had a very low specificity for the RRE compared to the Rev peptide. It also only exhibited a 2-fold higher affinity for RRE compared to poly r[A]-r[U]. Therefore, the linker was modified to improve the RRE specificity.[134] Two new neomycin-acridine conjugates with different linker lengths were synthesized and evaluated. It is interesting that all three compounds showed similar affinity for RRE; however, they had different RRE specificity ratios. Neo-N-acridine had the shortest linker length and the best RRE specificity. It maintained most of its activity in the presence of excess *calf thymus (*CT*)* DNA and $tRNA^{mix}$. Furthermore, it showed a 100-fold higher affinity for RRE compared to poly r[A]-r[U]. In contrast, neo-C-acridine had the longest linker length and the worst specificity. These results indicate that the linker length has significant effects on RRE specificity of neomycin-acridine conjugates.[134] Tor and co-workers also conjugated aminoacridine to the 6' position of both tobramycin and kanamycin A to generate two new aminoglycoside-acridine conjugates tobra-N-acridine and kanaA-N-acridine. A solid phase displacement assay indicated that these two compounds have approximately 400- and 2500-fold higher RRE affinities, respectively, compared to their parent molecules. However, there was no

Table 1.3 Calculated affinities (K_i) for various RNAs and DNAs

Aminoglycoside derivatives	**K_i /RRE66 [nM]**	**K_i /r[A]-r[U] [nM]**	**K_i /CT DNA [nM]**	**RRE specific ratio[a]**
Neo-N-acridine	2.4	200	510	17
Neo-S-acridine	2.6	6.5	69	26
Neo-C-acridine	2.6	32	31	50
Tobra-N-acridine	5.9	410	290	14
KanaA-N-acridine	9.6	1400	1300	3.9
Neo-Neo	2.6	12	390	49
Neo-N-Neo	2.1	1.6	260	–
Tobra-Tobra	8.4	190	610	11
KanaA-KanaA	11.0	1900	7500	6.5

[a] Specificity ratio= (average IC_{50} in the presence of DNA and $tRNA^{mix}$)/(IC_{50} in the absence of other nucleic acids).

selectivity for RNA over DNA, which indicated that high RRE affinity is not a prerequisite for high RRE selectivity. Among the aminoglycoside–acridine conjugates tested, kanaA-N-acridine exhibited the best RRE specificity ratio due to its relatively low affinity for both CT DNA and poly r[A]-r[U].[134]

Earlier research has shown that dimerization can dramatically increase the affinity of aminoglycosides for many different RNAs, including the hammerhead ribozyme (HH16), a dimerized A-site, and $tRNA^{Phe}$.[135-137] The same approach has been applied to RRE-binding compounds. It has been reported that RRE contained multiple binding sites for aminoglycosides. The dimerization of these compounds could potentially increase their binding affinity to RRE and possibly, selectivity as well as specificity.[122, 138] Unlike the acridine conjugates, the selectivity of RNA over DNA is largely maintained upon dimerization. However, the specificity varies across different aminoglycoside dimers.[134] The dimeric neomycin B derivatives neo-neo and neo-N-neo (Figure 1.32) have similar affinities to both poly r[A]-r[U] and the RRE66 (Figure 1.24), while neo-neo has a higher affinity to a mixture of tRNAs. Tobra-tobra has about a 3-fold lower RRE affinity, but it retains most of its RRE–Rev inhibition activity in the presence of excess CT DNA and loses partial activity in the presence of $tRNA^{mix}$. The dimerization of kanamycin A increases its affinity for the RRE by 2200-fold, and only 80-fold to poly r[A]-r[U]. It has the lowest affinity for simple duplex DNA and RNA. Compared to all the aminoglycoside dimers evaluated, the kanaA-kanaA dimer has the lowest RRE affinity but the best RRE specificity ratio and neo-neo has the highest RRE affinity but the worst RRE specific ratio. These trends are very similar to those observed for aminoglycoside–acridine conjugates.

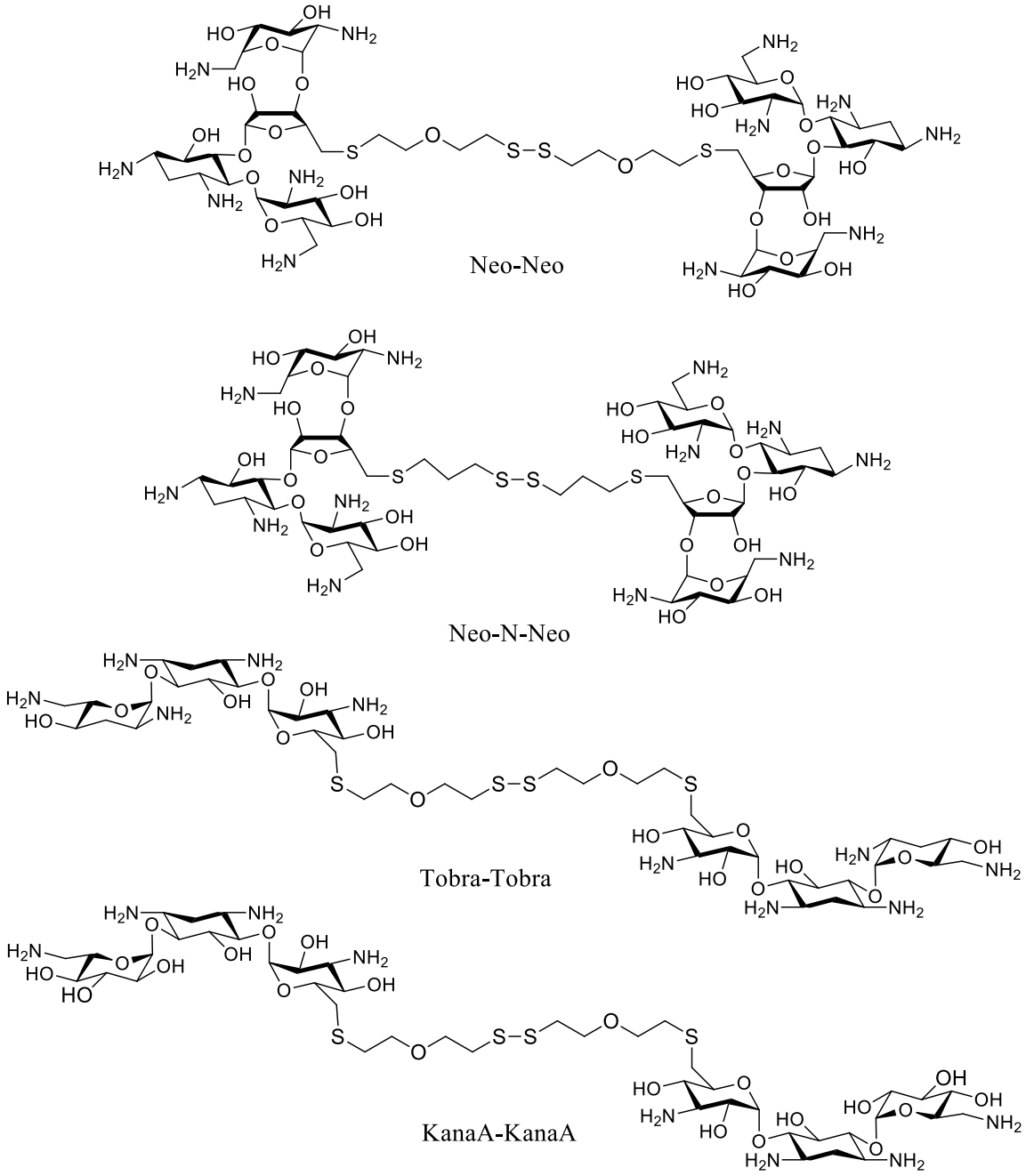


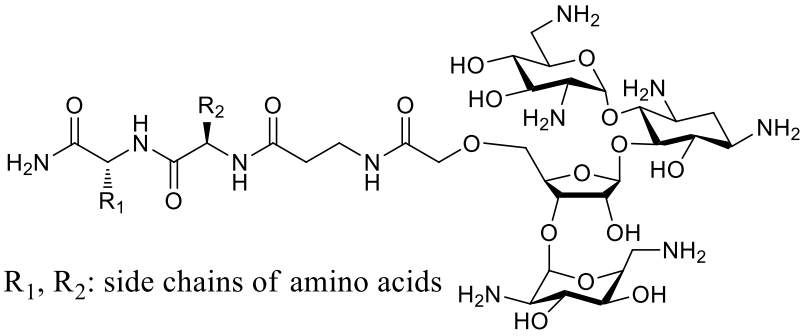
Figure 1.32 Structures of aminoglycoside dimers.

Further modification of linker length of dimers did not affect either affinity or specificity.[134] Taken together, it appears that there is an inverse relationship between the RNA affinity and RRE specificity for aminoglycoside derivatives. This may represent a

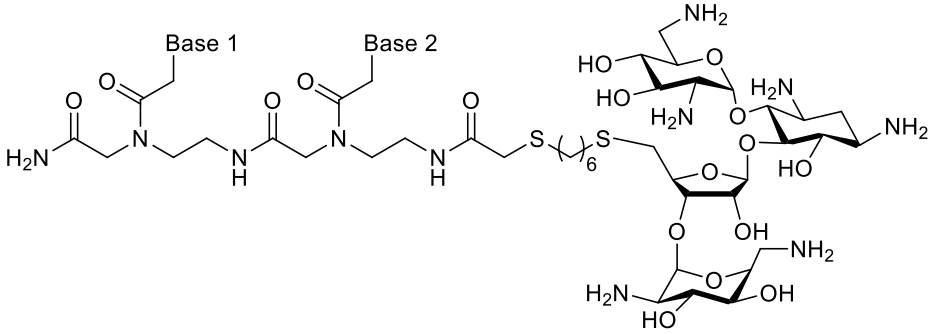
general principle for molecule-RNA interactions. It also highlights the challenge of targeting a structured RNA by using ligands with both high affinity and high specificity. Indeed, among most molecules that have been designed, few of them exhibit a high RRE specificity even though high RRE affinity can be easily achieved.

To make more diverse pharmacophores, the strategy involves addition of a simple peptide or peptide nucleic acid to the aminoglycoside pharmacophore in order to make new heteroconjugates.[139-140] The added moiety should carry a site for additional interactions between drugs and RNA. For example, peptides possessing both hydrogen bonding donors and acceptors could potentially form hydrogen bonds with bases in a loop/stem region of RNA. Further, peptides can be easily prepared and diversified by conventional solid phase synthesis. Yu and co-workers developed a neomycin-dipeptide library based on this purpose (Figure 1.33). Sixteen amino acids were chosen in each position to give a total of 256 different dipeptides. The acetic-acid linked neomycin was coupled after a β-Ala space-linker. Only two amino acid positions were used for peptides because they believed that dipeptide would be more stable than longer peptides. Furthermore, they may possess the correct distance to make hydrogen bonding.[134] Five peptides were isolated through solid phase screening against fluorescein-labeled RRE RNA. The surface plasmon resonance (SPR) and fluorescence anisotropy (FA) were used to evaluate their activity for biotinylated RRE. Even though the binding affinities slightly improved in comparison to neomycin, one conjugate, Neo-β-Ala-Arg-His, exhibited the best specificity for RRE, which was about 5-fold more specific then neomycin itself.[139] Although Neo-β-Ala-Arg-His showed only sub-micromolar affinity for RRE, the data suggested that there was no direct relationship between binding affinity and specificity. The conjugation of dipeptide

to neomycin apparently improved the specificity of the compounds and further provided a new strategy to design and synthesize more specific RNA binding molecules.



Neomycin B-peptide conjugates



Neomycin B-PNA conjugates

Figure 1.33 Structures of peptide-neomycin B and PNA-neomycin B conjugates

In order to further improve the activity of neomycin-dipeptide conjugates, Yu and co-workers designed a neomycin-PNA conjugate library due to the ubiquitous binding properties of PNA (Figure 1.33). Four monomeric PNAs and sixteen dimeric PNAs were linked to neomycin B through a six carbon spacer-linker. The molecular weight range of these molecules was from 1000 to 1500. This property allows neomycin-PNAs to penetrate the cells without cellular uptake issues. Six conjugates were found to have dissociation constants lower than 100 nM, which were as much as five times lower than that of neomycin B.[140] This demonstrated that PNAs aided the interactions between conjugates and RNAs. In addition, the sequences of PNA played a role in the determination of the

magnitude of interactions. Further comparison of binding affinities of these conjugates for RRE RNA and tRNA showed that the two compounds, 4-AA and 4GG, have the highest preference for RRE RNA. This result, along with the specificity experiment, suggested that these two molecules were more specific for RRE RNA. RNase protection assays showed that 4-AA and 4-GG bound the region where neomycin binds. They also spanned the loop region and possibly formed base pairs with the loop.[140] The combination of a small molecule with the moieties from a medium-sized molecule or a macromolecule gave a new strategy to generate multifunctional medium sized molecules that possess both a high binding affinity site as well as a specific interaction site. This strategy may be generally applicable to targeting any other RNA molecules.

As previously described, HIV-1 RRE functions through its interactions with the Rev protein.[43] Furthermore, it has been well demonstrated that only short fragments of the Rev protein are required for binding to the RRE RNA structure.[86] Since nature has provided a blue print for RNA ligand design, it is not surprising that peptidic compounds and peptidomimetics are the most well explored medium-sized molecular species designed for RNA binding. Peptidic ligands are also particularly attractive due to their synthetic accessibility, functional diversity, and amenability to the rapid preparation of combinatorial libraries.

Previous work has shown that binding specificity for RRE correlates with peptide helicity.[42] Kiplin and co-workers designed and evaluated the first α-helical peptidomimetics targeted for HIV-1 RRE RNA using macrolactam constraints between amino acid side chains (i, i+4), a known strategy to induce helicity in small peptides.[141] They generated the constrained peptidomimetics based upon a highly specific RRE

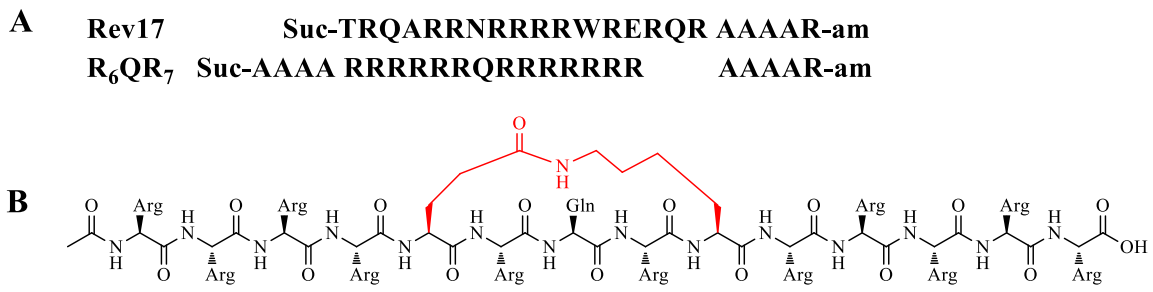


Figure 1.34 (A) Sequence of Rev17 and R_6QR_7; (B) Structure of α-helical peptidomimetic.

binding peptide R_6QR_7.[132] It binds the RRE with similar affinity compared to Rev17 (Figure 1.34). The lactam rings were generated using Lys and Asp or Glu. They replaced Arg on different positions along the chain to explore the effects of length, location, and orientation of the lactam ring. EMSA was used with wild-type RRE IIB and a C46–G74 mutant RRE to determine the dissociation constant, as well as the specificity. One peptidomimetic, **11**, stood out for its high binding affinity (45 nM) as well as its excellent specificity (26-fold). Reversing the direction of the lactam ring, removing the lactam ring, or replacing the central glutamine with asparagine (peptidomimetics **9**, **16** and **18**, Table 1.4) resulted in a loss of both binding affinity and specificity. Further CD experiments confirmed that only the peptidomimetics containing a Glu–Lys linkage exhibited a classical α-helical curve shape while other peptidomimetics and Rev17 showed an unstructured peptide signal.[132] These results illustrated that the linkage had to be positioned correctly to allow the essential glutamine to be viable for binding. Furthermore, α-helicity is required for high binding affinity. In a fluorescence polarization competition assay, peptidomimetic **11** was shown to effectively compete with the fluorescein-labeled Rev17 with an IC_{50} of 150 nM. This conformationally constrained α-helical peptidomimetic can bind to RRE with high affinity and specificity, as well as compete for the Rev binding site. This

suggested that the approach may be useful for generating other molecules that specifically target the RRE RNA for anti-viral therapy.

Table 1.4 Composition and binding affinities of R_6QR_7 peptidomimetic.

Peptide	Sequence [a]	K_{sp} [nM]	K_{nonsp} [nM]	Specificity
4	Ac-RRRKRRQDRRRRRR-OH	800	1200	1.5
5	Ac-RRRKRRQERRRRRR-OH	1200	1600	1.3
6	Ac-RRRDRRQKRRRRRR-OH	1200	1600	1.3
7	Ac-RRRERRQKRRRRRR-OH	700	1200	1.7
8	Ac-RRRRKRQRDRRRRR-OH	1200	1600	1.3
9	Ac-RRRRKRQRERRRRR-OH	200	1200	6
10	Ac-RRRRDRQRKRRRRR-OH	700	1200	1.7
11	Ac-RRRRERQRKRRRRR-OH	45	1200	26
12	Ac-RRRRRKQRRDRRRR-OH	1200	1600	1.3
13	Ac-RRRRRKQRRERRRR-OH	1200	1200	1
14	Ac-RRRRRDQRRKRRRR-OH	1200	1200	1
15	Ac-RRRRREQRRKRRRR-OH	700	1200	1.7
16	Ac-RRRRERQRKRRRRR-OH	1200	1200	1
17	Ac-RRRRRRQRRRRRRR-OH	1600	1600	1
18	Ac-RRRRERNRKRRRRR-OH	800	1200	1
Rev17	Suc-TRQARRNRRRRWRER QRAAAAR	100	1600	16

[a]Red residues are involved in the macrolactam constraint. Peptide **17** is a general linear R_6QR_7 control peptide. RRE binding K_d's were determined by electrophoretic mobility shift assays, with specificity defined as the ratio of dissociation constants of wild-type RRE IIB and a C46-G74 mutant RRE, shown not to specifically recognize Rev or R_6QR_7 peptides.

Earlier studies have identified an amphiphilic peptide originally aimed at calmodulin (peptide **19**, Table 1.5).[142] BLAST (Basic Local Alignment Search Tool) search of this peptide showed high homology with several typical RNA-binding proteins.[143] Yu and co-workers utilized this peptide and designed a series *N*-methylated peptides to explore the effect of methylation on RNA binding affinities.[144] Peptides bearing one or two *N*,*N*-dimethyl-Lys groups were synthesized and evaluated (Table 1.5).[144] CD experiments showed that all the peptides had similar α-helicities even though the position and number of methylated Lys are different. However, only peptide **25** containing two *N*,*N*-dimethyl-Lys groups showed high binding affinity comparable to Rev peptide in the fluorescence

Table 1.5 Sequences and affinities of α-helical peptides containing *N*,*N*-dimethyl-Lys

Peptide	**Sequence (position(s) of K*)[a]**	**α-Helicity (%)[b]**	**K_d vs RRE [nM]**	**K_d vs TAR [nM][c]**	**K_d vs $tRNA^{mix}$ [nM][c]**
19	LKKLLKLLKKLLKLKG	26/57	22	62 (2.8)	55 (2.5)
20	LKKLLKLLKKLLK*LKG	5/45	79	–	–
21	LKKLLKLLK*KLLKLKG	9/48	74	–	–
22	LKK*LLKLLKKLLKLKG	8/50	75	–	–
23	LKKLLKLLK*KLLK*LKG	7/49	30	–	–
24	LKK*LLKLLKKLLK*LKG	6/43	69	–	–
25	LKK*LLKLLK*KLLKLKG	8/52	9.1	53 (5.8)	42 (4.6)
26	LKK*LLKLLK*KLLK*LK G	6/47	87	–	–
Rev17	Suc-RQARRNRRRRWRER QRAAAAR	33/73	8.5	21 (2.5)	20 (2.4)

[a] K* = N^{ε},N^{ε}-dimethyl Lys. [b] In 10 mM H_3PO_4 / 50% 2,2,2-trifluoroethanol (TFE) in 10 mM H_3PO_4, pH 7.4. [c] Discrimination ratios (K_d against other RNA / K_d against RRE) are given in parenthesis.

anisotropy assay (9.1 nM and 8.5 nM, respectively). Other modified peptides with one, two, or even three *N*,*N*-dimethyl-Lys groups displaced weaker binding affinities for RRE RNA though all of them share similar α-helicity. This result suggested that the position of the *N*,*N*-dimethyl-Lys rather than α-helicity played significant roles in the RNA binding affinities for the peptides. Furthermore, specificity was evaluated against $tRNA^{mix}$ and TAR RNA. Peptide **25** gave an average discrimination ratio of 5.2, while both unmodified peptide **19** and Rev17 exhibited a discrimination ratio of only 2.5. The specificity and affinity of peptide **25** was believed to come from an introduction of conformational change when the peptide bound to the RNA. This was confirmed by CD experiments; upon addition of peptide **25** to 0.1 and 0.3 equivalent of RRE, the α-helicity increased from 7.5% to 11% and 18%. RNA foot-printing studies also suggested that both peptide **25** and the Rev17 peptide share a similar binding site on RRE, which is located at the internal loop and nearby stem regions.

In order to further explore the strategies to increase the binding affinity and specificity, Yu and co-workers hypothesized that the installation of a larger substitute than a methyl group might give rise to a diverse set of α-helical peptides that recognize RNA specifically. Acridine was introduced into the peptide through the ε-amino group of Lys as an intercalator due to its potential to form additional π-π interactions with bases in the target RNA. Peptide **19** acted as prototype, each Lys was replaced by *N*-acridinyl-Lys individually to generate 6 sequences. The α-helicity was evaluated through CD experiments and the binding affinities against RRE, TAR and $tRNA^{mix}$ were obtained through FA experiments. Three peptides (peptide **27**, **30** and **31**, Table 1.6) bound strongly to RRE and TAR; however, they were barely able to discriminate the difference between RRE and TAR RNA. The position of *N*-acridinyl-Lys directly related to the binding affinity of the peptides. To further confirm this, three bis-acridinylated peptides (peptide **33**, **34** and **35**, Table 1.6) were designed by selecting two out of the three positions that led to the highest binding affinities of the mono-acridinylated peptides. The binding affinity significantly improved when using bis-acridinylated peptides. Peptide **33** displayed a binding affinity of 610 pM to RRE, which is about 36-fold tighter than the unmodified peptide. It also maintained a similar α-helicity to the mono-acridinylated substance. Even though the peptides could not discriminate between the structure of RRE and TAR RNA, peptide **33** and peptide **35** exhibited about 8-fold and 18-fold discrimination ratios against $tRNA^{mix}$, which suggested that the two acridines could increase both the binding affinity and selectivity. Another tris-acridinylated peptide **36** showed the strongest binding affinity to all RNA targets and displayed poor specificity. Further analysis of different hairpin RNAs demonstrated that the peptides preferred the stem-bulge region of the hairpin RNAs.

It is clear that the addition of multiple acridines significantly increased binding affinity and specificity, which led to one of the best binding pharmacophores against RNA targets.

Table 1.6 Sequences and binding affinity of peptides featuring *N*-acridinyl Lys

Peptide	Sequence[a]	α-Helicity (%)[b]	K_d vs RRE [nM][c]	K_d vs TAR [nM][c]	K_d vs $tRNA^{mix}$ [nM]
19	LKKLLKLLKKLLKLKG	26/57	22 (2.5)	62 (0.89)	55
27	LK*KLLKLLKKLLKLKG	14/52	3.2 (4.4)	2.5 (5.6)	14
28	LKK*LLKLLKKLLKLKG	18/57	11	17	16
29	LKKLLK*LLKKLLKLKG	18/55	14	8.3	19
30	LKKLLKLLK*KLLKLKG	36/62	4.3 (3.0)	1.3 (10)	13
31	LKKLLKLLKK*LLKLKG	25/63	6.8 (2.1)	1.5 (9.3)	14
32	LKKLLKLLKKLLK*LKG	15/65	13	16	9.3
33	LK*KLLKLLK*KLLKLKG	35/57	0.61 (8.0)	0.55 (8.9)	4.9
34	LK*KLLKLLKK*LLKLKG	15/48	0.72 (8.5)	0.64 (9.5)	6.1
35	LKKLLKLLK*K*LLKLKG	12/54	0.92 (7.5)	0.37 (18)	6.9
36	LK*KLLKLLK*K*LLKLKG	15/49	0.25 (0.9)	0.20 (1.1)	0.23

[a] K* = N^{ε}-acridinyl-Lys. [b] In 10 mM H_3PO_4 / 50% TFE in 10 mM H_3PO_4, pH 7.4. [c] Discrimination ratios (K_d against $tRNA^{mix}$/ K_d against RRE or TAR) are given in parenthesis.

Even though these Lys- and Leu-rich peptides displayed excellent binding affinity and selectivity towards hairpin RNAs, it is difficult to find one peptide that can target a specific RNA.[131, 144] Indeed, hairpin RNAs share great similarities within the stem, bulge, and loop structures. Furthermore, most hairpin RNAs are flexible and can easily undergo conformational changes.[81] To overcome this problem, Yu and co-workers performed an alanine scan using peptide **19**, which is a Lys- and Leu-rich α-helical peptide that has been chosen to bind hairpin RNA structures.[144-145] The Lys residue was replaced by Ala individually to generate 7 mutant peptides. Their binding affinities towards RRE, TAR, 16S rRNA A-site from *E. coli* and IRES domain IV of HCV were tested and evaluated (Figure 1.35). For most of mutant peptides, they showed similar binding affinities for

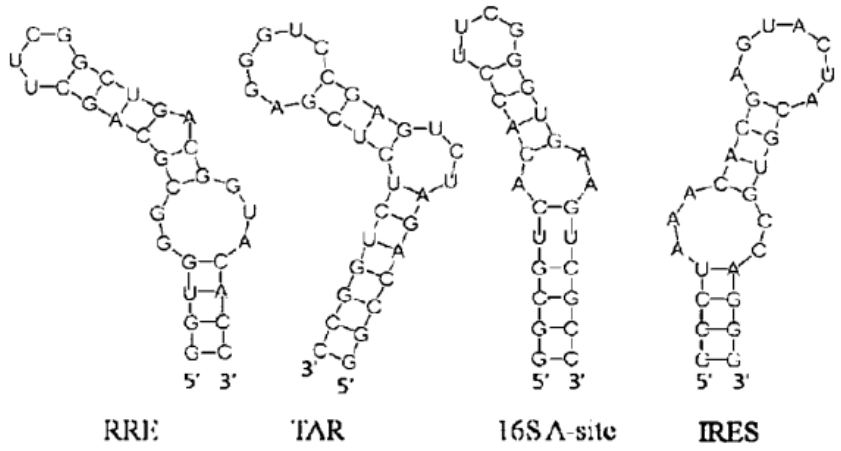


Figure 1.35 Secondary structures of hairpin RNA targets used in study

TAR RNA compared to the unsubstituted peptide **19** (Table 1.7). This may imply that TAR RNA adopts a more flexible structure, which could accommodate several similar peptides. Within these four RNAs tested, HCV IRES had the highest standard deviation (70 nM) of all the binding affinities across the seven Ala substituted peptides and only one peptide exhibited a high binding affinity for IRES. This indicated that IRES was the most specific hairpin and therefore, was chosen to perform further research. During the alanine scan, three Lys positions (6, 9 and 13) were shown to be important for RNA binding. They were systematically replaced by ornithine (Orn), 1,4-diaminobutyric acid (Dab), and 1,3-diaminopropionic acid (Dap). Three mutations, peptide **47**, **48**, and **49** (Table 1.8), have the greatest improved affinities against the IRES hairpin. Further modifications were generated by replacing two or three Lys with acridine residues from these three positions. Only one peptide (**50**, Table 1.8) exhibited a significantly improved binding affinity (K_d=680 pM). Fortunately, this sequence could discriminate IRES from RRE, TAR, and the 16S A-site by 25-fold. They performed further mutations on the hydrophobic region by replacing Leu with Trp due to the possible interaction between the aromatic rings with the bases of RNAs. The results indicated that the position of Trp significantly affected the

Table 1.7 Sequences and affinity of Ala and Trp substituted peptides

Peptide	Sequence	α-Helicity (%)[a]	K_d vs RRE [nM]	K_d vs TAR [nM]	K_d vs 16S A-Site [nM]	K_d vs IRES [nM]
19	LKKLLKLLKKLLKL KG	26/57	22 ± 2	66 ± 5	100 ± 4	45 ± 3
37	LAKLLKLLKKLLKL KG	9/44	140 ± 10	16 ± 3	160 ± 10	53 ± 9
38	LKALLKLLKKLLKL KG	10/42	56 ± 6	87 ± 10	92 ± 6	62 ± 8
39	LKKLLALLKKLLKL KG	12/47	150 ± 10	150 ± 9	160 ± 12	150 ± 10
40	LKKLLKLLAKLLKL KG	12/43	130 ± 10	74 ± 10	180 ± 15	120 ± 20
41	LKKLLKLLKALLKL KG	17/45	92 ± 10	25 ± 5	64 ± 3	33 ± 4
42	LKKLLKLLKKLLAL KG	8/44	90 ± 8	39 ± 5	200 ± 9	210 ± 10
43	LKKLLKLLKKLLKL AG	9/43	25 ± 2	63 ± 9	53 ± 2	49 ± 2
44	WKKLLKLLKKLLKL AG	48/65	2.5 ± 0.3	2.0 ± 0.3	0.76 ± 0.1	0.74 ± 0.1
45	LKKLLKWLKKLLKL AG	19/71	1.2 ± 0.1	21 ± 1	17 ± 2	0.87 ± 0.1
46	LKKLLKLLKKLLKW AG	20/52	18 ± 2	7.8 ± 0.6	3.3 ± 0.3	0.69 ± 0.1

[a] In 10 mM H_3PO_4 / 50% TFE in 10 mM H_3PO_4, pH 7.4.

binding affinity, as well as the α-helicity. Mutations on position 1, 7, and 14 increased the binding affinity to subnanomolar K_ds (**44**, **45**, and **46**, Table 1.7). Therefore, the third generation mutant peptides were generated by combining the mutants of both the hydrophilic regions and hydrophobic regions of the mother peptide. Only two peptides (**51** and **52**), significantly improved the binding affinities with **51** having a discrimination ratio of 45. Replacing Trp with Phe or Tyr decreased both binding affinities for IRES, suggesting that the indole moiety had a specific interaction with bases in the IRES hairpin.

Table 1.8 Sequences and affinity of Orn, Dab, Dap, and Trp substituted peptides

Peptide	Sequence	α-Helicity (%)[a]	K_d vs IRES [nM][b]
47	LKKLLDabLLKKLLKLAG	23/63	1.6 ± 0.2 (6.3)
48	LAKLLKLLOrnKLLKLAG	28/57	1.3 ± 0.2 (9.1)
49	LKALLKLLKKLLDapLAG	39/54	1.4 ± 0.2 (12)
50	LKALLKLLOrnKLLDapLAG	48/61	0.68 ± 0.06 (25)
51	LKALLKWLOrnKLLDapLAG	32/38	0.55 ± 0.05 (45)
52	LKKLLDabLLKKLLKWAG	17/45	0.62 ± 0.05 (28)

[a] In 10 mM H_3PO_4 / 50% TFE in 10 mM H_3PO_4, pH 7.4. [b] Discrimination ratios (K_d against other hairpins/ K_d against IRES) are given in parenthesis.

As previously described, constrained peptidomimetics at proper positions could increase the α-helicity. Later, Yu and co-workers designed a series of constrained peptides from the previously studied amphiphilic peptide **19**.[146] A link was built between the fifth and twelfth Leu residues to give a seven amino acid gap, which was reported to increase α-helicity propensities.[147] They chose to incorporate linkers on the hydrophobic side due to the theory that there were less interactions possible with the RNAs. The fifth and twelfth Leu were replaced by Cys, and cross-linked by 1,4-(bismaleimido)butane (BMB), bis(maleimido)ethane (BMOE), 1,4-bismaleimidyl-2,3-dihydroxybutane (BMDB), or bis(maleimido)hexane (BMH) to give four different sequences (Table 1.9). These peptides, along with their parent peptides, were tested for their binding affinity for RRE and TAR of HIV-1, and IRES of HCV. Surprisingly, none of these peptides displayed improved binding affinities compared to their parent peptide. Even though the α-helicity increased to 50% for the BMB linker, the binding affinities towards all three hairpin RNAs decreased. These results suggested that the hydrophobic phase may play a significant role in RNA recognition. Earlier reports have indicated that α-helicity could maintain high values even in aqueous media upon dimerization. Here, they used Cys to replace Leu to form intermolecular disulfide bonds. According to the different positions of the Cys sets, five dimer peptides were generated and evaluated (Table 1.10).

Table 1.9 The sequences, percent helicity, and K_d values of covalently cross-linked peptides against hairpin RNA structures.

Peptide	Sequence	α-Helicity (%)[a]	K_d vs RRE [nM]	K_d vs TAR [nM]	K_d vs IRES [nM]
99	LKKLLKLLKKLLKLAG	9/43	25	63	49
109	LKKLCKLLKKLCKLAG	18/73	13	9.6	34
110(BMB)[b]	LKKLCKLLKKLCKLAG Linker	50/61	90	67	47
110(BMOE)[b]	LKKLCKLLKKLCKLAG Linker	13/23	46	43	31
110(BMDB)[b]	LKKLCKLLKKLCKLAG Linker	15/53	56	8.4	25
110(BMH)[b]	LKKLCKLLKKLCKLAG Linker	24/44	85	52	27

[a] In 10 mM H_3PO_4 / 50% TFE in 10 mM H_3PO_4, pH 7.4. [b] 1,4-(Bismaleimido)butane (BMB), bis(maleimido)ethane (BMOE), 1,4-bismaleimidyl-2,3-dihydroxybutane (BMDB), or bis(maleimido)hexane (BMH).

Table 1.10 The sequences, percent helicity, and K_d values of covalently cross-linked peptide dimers against hairpin RNA structures.

Peptide	Sequence	α-Helicity (%)[a]	K_d vs RRE [nM]	K_d vs TAR [nM]	K_d vs IRES [nM]
111	CKKLLKLCKKLLKLAG CKKLLKLCKKLLKLAG	77/80	0.21	0.17	0.14
112	LKKCLKLLKKCLKLAG LKKCLKLLKKCLKLAG	79/88	0.18	0.16	0.068
113	LKKLCKLLKKLCKLAG LKKLCKLLKKLCKLAG	74/79	0.18	0.061	0.060
114	LKKLLKCLKKLLKCAG LKKLLKCLKKLLKCAG	80/83	0.040	0.021	0.024
115	LKKLLKLCKKLLKLCG LKKLLKLCKKLLKLCG	73/82	0.24	0.12	0.093

[a] In 10 mM H_3PO_4 / 50% TFE in 10 mM H_3PO_4, pH 7.4.

As expected, all five peptides exhibited a high α-helicity (>70%). Although they showed different orientations of the hydrophilic phase, all dimers displayed picomolar K_ds. Clearly that the position of cross-linking defined the structure of two α-helicity peptides as a dimer. FA experiments determined the 1:1 stoichiometry of the peptide-RNA complex, which suggested that the RRE RNA groove, as well as other hairpin RNA grooves have sufficient room to accommodate a dimeric peptide. Furthermore, Dimer **57** is much more stable than the reference peptides both in reductive cytoplasmic conditions and in human serum. However, none of these peptides have selectivity for the RNAs tested. Further fine tuning is needed to increase the selectivity by decreasing the non-specific interactions between peptides and RNAs through positive charges.

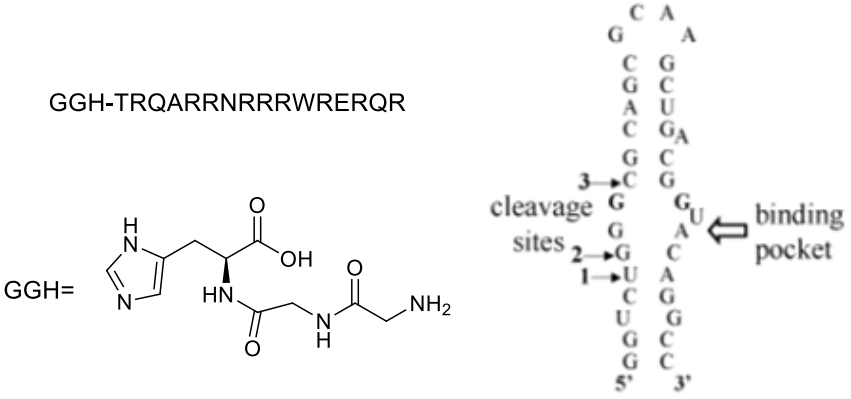


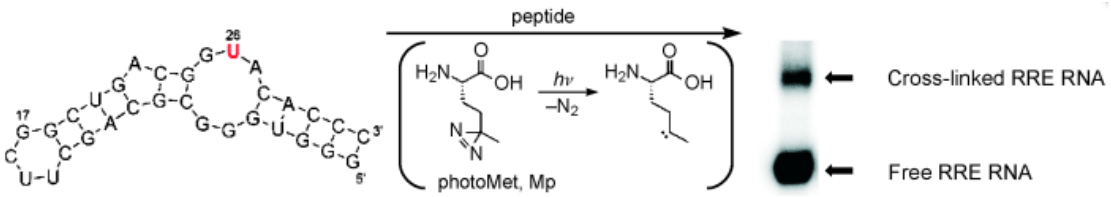
Figure 1.36 Sequence of peptide and structure of GGH and RRE RNA IIB, the cleavage sites that are identified by mass spectrometry studies are highlighted by arrows.[149]

The most popular strategy used for targeting RNA is to design molecules that bind to the target RNA through electrostatic interactions, hydrogen bonding, π-π interactions, Van der Waals contacts, and Watson-Crick base pairs interactions.[93, 148] Even though many molecules have been developed based on these principles, few of them are selective due to non-specific interactions. Furthermore, most of these interactions are reversible, which could potentially compete with their natural counterpart. To overcome this problem, other

strategies were developed from either forming a covalent bond between drug and RNA target or cleaving the RNA target.[149-150] Cowan and co-workers used the minimized sequence of Rev protein and an N-terminal metal binding ATCUN (amino-terminal copper and nickel binding) motif, GGH, to recognize and cleave the RRE RNA (Figure 1.36).[149] The copper-Rev peptide complex was studied *in vitro* using 5'-fluorescein end-labeled RNA in the presence and absence of the mild reducing agent ascorbate under physiological conditions. The cleavage products were displayed as three bands under stoichiometric conditions. Both mass spectrometric studies and PAGE data confirmed that the cleavage is site specific. Other control experiments illustrated that both copper and RRE specific binding peptide Rev are important for the site specific cleavage, and incorporation of the metal binding motif at the N-terminal of Rev peptide did not disrupt the binding affinity. This was the first example of metal-peptide complex that stoichiometrically targeted HIV RRE RNA for site-specific cleavage with a biological co-reactant under mild conditions.

In recent research, photocrosslinking has been widely used for exploring natural RNA-protein interactions.[151] Substances that do not alternative interactions could serve as reactants to generate carbenes. Therefore, reactions of photogenerated carbenes with nucleophiles in RNA might form the covalent adducts specifically.[152] To this end, Yu and co-workers used an amphiphilic Arg-rich peptide to generate photo-controlled crosslinking peptides. Each Leu and Arg was replaced by a photoactive diazirine-bearing Met analogue (photoMet) to avoid diminishing the interactions with RRE RNA (Figure 1.37). The binding affinities for RRE RNA were obtained through FA, and the numbers varied between 0.29 nM to 14 nM. However, the modification at the hydrophilic regions were generally low binding, which suggested that Arg residues contribute more greatly to

binding with RNA. At micromolar concentration of RRE, the crosslinked adduct was observed only when R10Mp was irradiated at 302 nm. Foot-printing experiments determined the crosslink site. The nucleophile in uridine 26, located in the bulge region, was found to be responsible for covalent adduct formation. Exposure of uridine 26 increased the possibility that the electron rich carbon-carbon double bond in this base could approach closely to the carbene of the photo reactant, and therefore, promote the reaction. The covalent adducts between RRE RNA and peptides can be generally used for developing site-specific RNA binders.



R10Mp: Ac-LRRLLRLLRMpLLRLAG

Figure 1.37 Structures of RRE, photoMet and sequence of R10Mp.[150]

1.6 Other RNA Targets

RNA is a unique macromolecular entity that plays key roles in carrying genetic information, catalyzing protein synthesis, and regulating gene expression in a living cell. Along with protein-protein interactions, RNA-protein interactions are the gateway to the diversity of function that mediates a variety of biological effects. One attractive approach to perturbing the system is to inhibit RNA–protein interactions by disassembling the construct either through binding to the protein or RNA portion. Targeting RNA as a macromolecular entity is a herculean task. That is likely due to the fact that RNA is single stranded and folds into unique structures like a stem, bulge, or loop to minimize its energy. The major groove of an A-form RNA is deeper and narrower than a B-form DNA and the

minor groove is shallower. As a result, the most specific ligands for DNA binding were not specific for RNA binding. Furthermore, some RNA structures are flexible and change conformation upon binding; therefore, specific binding to a single conformation is difficult.[153] In addition to the HIV-1 TAR RNA and RRE RNA discussed above, other RNAs like HIV-1 dimerization initiation site (HIV-1 DIS), expanded nucleotide repeats of myotonic dystrophy (DM), and HCV internal ribosome entry site (IRES) have been well studied. Indeed, the strategies described in this review could be applied to most RNA targets. Here, we will simply introduce the recent exploration of molecules that could target DIS and expanded nucleotide repeats of DM.

All retroviruses, including HIV-1, contain two identical or nearly identical segments of the RNA genome that are non-covalently linked near their 5′-ends.[154] This process, known as genomic RNA dimerization, proceeds through the HIV-1 dimerization initiation site (DIS), which is a highly conserved sequence in the 5′-noncoding region of the viral genomic RNA.[155] It has been demonstrated that DIS is associated with RNA dimerization, packaging, and reverse transcription, as well as viral infectivity.[154] The presence of a self-complementary sequence in the DIS loop suggests that dimerization starts with the formation of a kissing loop complex, and upon binding with the viral NGp7 nucleocapsid protein, the complex stabilizes into an extended duplex form (Figure 1.38A).[154]

Earlier research has identified the X-ray structures of the DIS kissing-loop complex and extended duplex forms.[156-157] It is surprising that they share similar sequence and structure with the bacterial 16S rRNA aminoacyl-tRNA decoding site (A-site).[158] Therefore, the well-known A-site ligands, like aminoglycosides, could be used to design

and explore DIS ligands. In fact, examination of all DIS/aminoglycosides revealed that the conserved base A280 in DIS directly interacts with ring I of aminoglycosides (Figure 1.38B). Based on this, Floris and co-workers designed a new aminoglycoside by replacing the ring I of neomycin with a thymine, possibly leading to Watson-Crick base-pairing interactions (Figure 1.38C).[159] The binding affinity was evaluated *in vitro* using isothermal titration calorimetry (ITC) with 23-nucleotides containing the minimal DIS hairpin loop. As expected, neomycin-thymine conjugate bound the DIS kissing loop in a 2:1 stoichiometry, which is lower than the original neomycin. This is possibly due to the loss

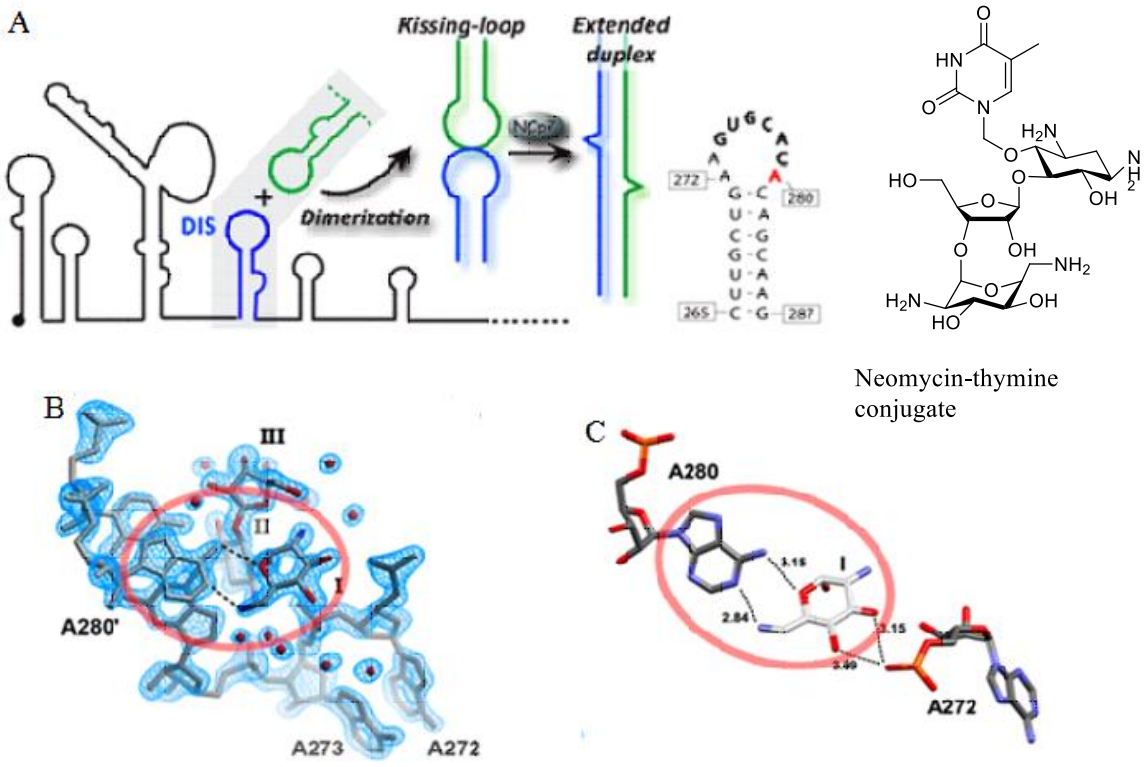


Figure 1.38 (A) HIV-1 genomic RNA dimerization mechanism and RNA sequence corresponding to the HIV-1 subtypes A and F used in this study. Structure of neomycin-thymine conjugates; (B) DIS extended duplex crystal structure (PDB ID 3C3Z) superimposed with the electron density map; (C) DIS kissing-loop complex crystal structure (PDB ID 2FCZ) [Ennifar, E.; Aslam, M. W.; Strasser, P.; Hoffmann, G.; Dumas, P.; van Delft, F. L., Structure-guided discovery of a novel aminoglycoside conjugate targeting HIV-1 RNA viral genome. *ACS Chem. Biol.* **2013,** *8* (11), 2509-2517.], Copyright 2013; [used with permission from American Chemical Society].

of four amine groups and optimal stacking interactions between ring I of neomycin and G271. Specificity of binding was also performed by ITC experiments. As expected, neomycin-thymine conjugate did not show any preference for binding to various RNA sequences, as well as a DIS A280U mutant sequence. These results suggested that the specificity of neomycin–thymine conjugate may come from the Watson–Crick hydrogen bonds between residue 280 and the thymine. Further analysis also demonstrated that neomycin-thymine conjugate can bind both HIV subtype B and F, which were unable to interact with aminoglycosides due to steric hindrance. This was the first ligand to recognize any HIV DIS RNA dimers, thus overcoming the limitation of natural aminoglycosides that could not bind DIS subtype RNAs. This work heavily relied on the crystal structure of the DIS complex, as well as molecular modeling of the drug-RNA interactions, suggesting a new strategy to develop specific RNA binders.

DM is an autosomal dominant genetic neuromuscular disease caused by either a trinucleotide repeat (CTG) in the 3' UTR of the dystrophia myotonica protein kinase (DMPK) gene (DM1) or a tetranucleotide repeat (CCTG) in intron 1 of the zinc finger 9 protein (ZNF9) gene (DM2).[153] The RNA displays a unique flexible 1x1 nucleotide internal loop repeat, which is the binding site for muscle blind-like 1 protein (MBNL1).[160] The binding of MBNL1 to the repeats can cause various diseases, such as the disregulation of pre-mRNA splicing.[161] In an effort to exploit potential ligands that can disrupt the interaction between MBNL1 and the RNA repeats, Disney and co-workers investigated a series of small molecules. In their first generation of ligands, the compounds were composed of a peptoid backbone attached to bisbenzimidazole Ht, called 2H-4. The compound 2H-4 displayed high binding affinity and bioactivity (Figure 1.39).[162] In the

second generation of ligands, they studied the linker between two binding motifs by using various backbones like polyamines, α-peptide, β-peptide, peptoid and peptide tertiary amide (PTAs).[163] These molecules were designed as dimers, which have been shown to have a slower off rate. Biological studies demonstrated that PTA is the most optimal scaffold, displaying high potency in both *in vitro* and *in vivo* experiments, cellular permeability, stability as well as toxicity experiments. Furthermore, the optimal compound 2H-K4NMe also improved DM1 associated alternative splicing defects in a mouse model. Therefore, it has a high clinic potential.

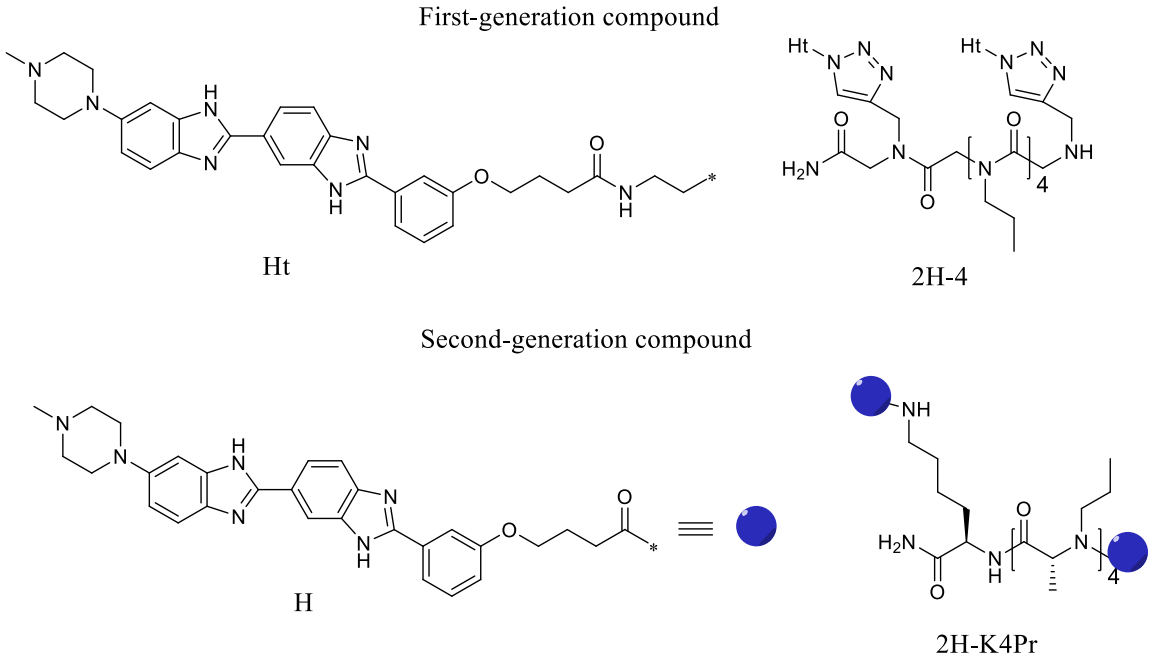


Figure 1.39 Structures of 2H-4 and 2H-K4Pr.

1.7 Conclusion

The development of ligands to bind functional RNA is a challenging task due to the fact that most RNAs possess flexible structures and undergo conformational changes upon binding. Furthermore, few RNA crystal structures have been elucidated. HIV-1 TAR RNA and RRE RNA are two of these RNAs. During the last few decades, more and more

molecules and strategies have been developed to explore clinical drugs. Due to the limited sampling area, small molecules usually are non-selective. Large molecules usually bind to RNA selectively, primarily through Watson-Crick base pairing rules, and they are large enough to compete with the endogenous protein ligands. However, due to the highly negatively charged backbone of large molecules, they are not cell permeable. In contrast, medium sized molecules have been demonstrated to be cell permeable. Several medium-sized compounds have displayed sub-nanomolar binding affinities to RRE RNA and have been shown to inhibit RNA controlled processes in cell-based assays. Although extensive research has been done to exploit HIV-1 TAR and RRE ligands, none of the molecules have made it to the clinic. This further creates the urgency to develop new anti-HIV treatment strategies and novel HIV targets with drug-like properties.

1.8 References

1. Baba, M., Recent status of HIV-1 gene expression inhibitors. *Antiviral Res.* **2006,** *71* (2-3), 301-306.

2. Meadows, D. C.; Gervay-Hague, J., Current developments in HIV chemotherapy. *ChemMedChem.* **2006,** *1* (1), 16-29.

3. Stevens, M.; Clercq, E. D.; Balzarini, J., The regulation of HIV-1 transcription: Molecular targets for chemotherapeutic intervention. *Med. Res. Rev.* **2006,** *26* (5), 595-625.

4. Martinez-Picado, J.; DePasquale, M. P.; Kartsonis, N.; Hanna, G. J.; Wong, J.; Finzi, D.; Rosenberg, E.; Günthard, H. F.; Sutton, L.; Savara, A.; Petropoulos, C. J.; Hellmann, N.; Walker, B. D.; Richman, D. D.; Siliciano, R.; D'Aquila, R. T., Antiretroviral resistance during successful therapy of HIV type 1 infection. *Proc. Natl. Acad. Sci. U. S. A.* **2000,** *97* (20), 10948-10953.

5. Carr, A., Toxicity of antiretroviral therapy and implications for drug development. *Nat. Rev. Drug Discov.* **2003,** *2* (8), 624-634.

6. Zhou, T.; Xu, L.; Dey, B.; Hessell, A. J.; Van Ryk, D.; Xiang, S.-H.; Yang, X.; Zhang, M.-Y.; Zwick, M. B.; Arthos, J.; Burton, D. R.; Dimitrov, D. S.; Sodroski, J.; Wyatt, R.; Nabel, G. J.; Kwong, P. D., Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature* **2007,** *445* (7129), 732-737.

7. Mehellou, Y.; De Clercq, E., Twenty-six years of anti-HIV drug discovery: Where do we stand and where do we go? *J. Med. Chem.* **2009,** *53* (2), 521-538.

8. Pollard, V. W.; Malim, M. H., The HIV-1 Rev protein. *Annu. Rev. Microbiol.* **1998,** *52* (1), 491-532.

9. Gait, M. J.; Karn, J., Progress in anti-HIV structure-based drug design. *Trends in Biotechnology* **1995,** *13* (10), 430-438.

10. O'Hara, B. M.; Olson, W. C., HIV entry inhibitors in clinical development. *Curr. Opin. Pharmacol.* **2002,** *2* (5), 523-528.

11. Williams, D. H.; Adam, F.; Fenwick, D. R.; Fok-Seang, J.; Gardner, I.; Hay, D.; Jaiessh, R.; Middleton, D. S.; Mowbray, C. E.; Parkinson, T.; Perros, M.; Pickford, C.; Platts, M.; Randall, A.; Siddle, D.; Stephenson, P. T.; Tran, T.-D.; Vuong, H., Discovery of a small molecule inhibitor through interference with the gp120-CD4 interaction. *Bioorg. Med. Chem. Lett.* **2009,** *19* (17), 5246-5249.

12. Dorr, P.; Westby, M.; Dobbs, S.; Griffin, P.; Irvine, B.; Macartney, M.; Mori, J.; Rickett, G.; Smith-Burchnell, C.; Napier, C.; Webster, R.; Armour, D.; Price, D.;

Stammen, B.; Wood, A.; Perros, M., Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity. *Antimicrob. Agents Chemother.* **2005,** *49* (11), 4721-4732.

13. Fatkenheuer, G.; Pozniak, A. L.; Johnson, M. A.; Plettenberg, A.; Staszewski, S.; Hoepelman, A. I. M.; Saag, M. S.; Goebel, F. D.; Rockstroh, J. K.; Dezube, B. J.; Jenkins, T. M.; Medhurst, C.; Sullivan, J. F.; Ridgway, C.; Abel, S.; James, I. T.; Youle, M.; van der Ryst, E., Efficacy of short-term monotherapy with maraviroc, a new CCR5 antagonist, in patients infected with HIV-1. *Nat. Med.* **2005,** *11* (11), 1170-1172.

14. Hughes, A.; Barber, T.; Nelson, M., New treatment options for HIV salvage patients: an overview of second generation PIs, NNRTIs, integrase inhibitors and CCR5 antagonists. *J. Infect.* **2008,** *57* (1), 1-10.

15. Elston, R.; Kuritzkes, D.; Bethell, R., An investigation into the influence of the tipranavir-asssociated V82L/T mutations on the susceptibility to Darunavir and Brecanavir. *Fourteenth CROI.* **2007**.

16. Hicks, C.; Gulick, R. M., Raltegravir: The first HIV type 1 integrase inhibitor. *Clin. Infect. Dis.* **2009,** *48* (7), 931-939.

17. Domaoal, R. A.; Demeter, L. M., Structural and biochemical effects of human immunodeficiency virus mutants resistant to non-nucleoside reverse transcriptase inhibitors. *Int. J. Biochem. Cell Biol.* **2004,** *36* (9), 1735-1751.

18. Rana, T. M.; Jeang, K.-T., Biochemical and functional interactions between HIV-1 Tat protein and TAR RNA. *Arch. Biochem. Biophys.* **1999,** *365* (2), 175-185.

19. Bayer, P.; Kraft, M.; Ejchart, A.; Westendorp, M.; Frank, R.; Rosch, P., Structural studies of HIV-1 tat protein. *J. Mol. Biol.* **1995,** *247* (4), 529-535.

20. Churcher, M. J.; Lamont, C.; Hamy, F.; Dingwall, C.; Green, S. M.; Lowe, A. D.; Butler, P. J. G.; Gait, M. J.; Karn, J., High affinity binding of TAR RNA by the human immunodeficiency virus type-1 Tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region. *J. Mol. Biol.* **1993,** *230* (1), 90-110.

21. Karn, J.; Graeble, M. A., New insights into the mechanism of HIV-1 trans-activation. *Trends Genet.* **1992,** *8* (11), 365-368.

22. Frankel, A. D.; Young, J. A. T., HIV-1: Fifteen proteins and an RNA. *Annu. Rev. Biochem.* **1998,** *67* (1), 1-25.

23. Jakobovits, A.; Smith, D. H.; Jakobovits, E. B.; Capon, D. J., A discrete element 3' of human immunodeficiency virus 1 (HIV-1) and HIV-2 mRNA initiation sites mediates transcriptional activation by an HIV trans activator. *Mol. Cell. Biol.* **1988,** *8* (6), 2555-2561.

24. Berkhout, B.; Silverman, R. H.; Jeang, K.-T., Tat trans-activates the human immunodeficiency virus through a nascent RNA target. *Cell* **1989,** *59* (2), 273-282.

25. Puglisi, J. D.; Tan, R.; Calnan, B. J.; Frankel, A. D.; Williamson, J. R., Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science* **1992,** *257* (5066), 76-80.

26. Delling, U.; Roy, S.; Sumner-Smith, M.; Barnett, R.; Reid, L.; Rosen, C. A.; Sonenberg, N., The number of positively charged amino acids in the basic domain of Tat is critical for trans-activation and complex formation with TAR RNA. *Proc. Natl. Acad. Sci. U. S. A.* **1991,** *88* (14), 6234-6238.

27. Marcello, A.; Zoppe, M.; Giacca, M., Multiple modes of transcriptional regulation by the HIV-1 Tat transactivator. *IUBMB Life* **2001,** *51* (3), 175-181.

28. Cao, H.; Rana, T. M., Specific HIV-1 TAR RNA Loop Sequence and functional groups are required for human cyclin T1-Tat–TAR ternary complex formation. *Biochemistry* **2002,** *41* (20), 6391-6397.

29. Aboul-ela, F.; Karn, J.; Varani, G., The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J. Mol. Biol.* **1995,** *253* (2), 313-332.

30. Churcher, M. J.; Lamont, C.; Hamy, F.; Dingwall, C.; Green, S. M.; Lowe, A. D.; Butler, P. J. G.; Gait, M. J.; Karn, J., High affinity binding of TAR RNA by the human immunodeficiency virus type-1 Tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region. *J. Mol. Biol.* **1993,** *230* (1), 90-110.

31. Graham, G. J.; Maio, J. J., RNA transcripts of the human immunodeficiency virus transactivation response element can inhibit action of the viral transactivator. *Proc. Natl. Acad. Sci. U. S. A.* **1990,** *87* (15), 5817-5821.

32. Bannwarth, S.; Gatignol, A., HIV-1 TAR RNA: The target of molecular interactions between the virus and its host. *Current HIV Research*, Bentham Science Publishers Ltd.: 2005; Vol. 3, pp 61-71.

33. Cochrane, A. W.; Perkins, A.; Rosen, C. A., Identification of sequences important in the nucleolar localization of human immunodeficiency virus Rev: relevance of nucleolar localization to function. *J. Virol.* **1990,** *64* (2), 881-885.

34. Malim, M. H.; Böhnlein, S.; Hauber, J.; Cullen, B. R., Functional dissection of the HIV-1 Rev trans-activator—Derivation of a trans-dominant repressor of Rev function. *Cell* **1989,** *58* (1), 205-214.

35. Malim, M. H.; Tiley, L. S.; McCarn, D. F.; Rusche, J. R.; Hauber, J.; Cullen, B. R., HIV-1 structural gene expression requires binding of the rev trans-activator to its RNA target sequence. *Cell* **1990,** *60* (4), 675-683.

36. Hope, T. J.; McDonald, D.; Huang, X. J.; Low, J.; Parslow, T. G., Mutational analysis of the human immunodeficiency virus type 1 Rev transactivator: essential residues near the amino terminus. *J. Virol.* **1990,** *64* (11), 5360-5366.

37. Cochrane, A. W.; Chen, C. H.; Rosen, C. A., Specific interaction of the human immunodeficiency virus Rev protein with a structured region in the env mRNA. *Proc. Natl. Acad. Sci. U. S. A.* **1990,** *87* (3), 1198-1202.

38. Satoshi Kubota; Pomerantz, R. J., A cis-acting peptide signal in human immunodeficiency virus type I Rev which inhibits nuclear entry of small proteins. *Oncogene* **1998,** *16,* 10.

39. Malim, M. H.; Hauber, J.; Le, S.-Y.; Maizel, J. V.; Cullen, B. R., The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* **1989,** *338* (6212), 254-257.

40. Kjems, J.; Calnan, B. J.; Frankel, A. D.; Sharp, P. A., Specific binding of a basic peptide from HIV-1 Rev. *Embo J.* **1992,** *11* (3), 1119-29.

41. Tan, R.; Frankel, A. D., Costabilization of peptide and RNA structure in an HIV Rev peptide-RRE complex. *Biochemistry* **1994,** *33* (48), 14579-14585.

42. Tan, R.; Chen, L.; Buettner, J. A.; Hudson, D.; Frankel, A. D., RNA recognition by an isolated alpha-helix. *Cell* **1993,** *73* (5), 1031-1040.

43. Cullen, B. R.; Malim, M. H., The HIV-1 Rev protein: prototype of a novel class of eukaryotic post-transcriptional regulators. *Trends Biochem. Sci.* **1991,** *16* (0), 346-350.

44. Mann, D. A.; Mikaélian, I.; Zemmel, R. W.; Green, S. M.; Lowe, A. D.; Kimura, T.; Singh, M.; Jonathan, P.; Butler, G.; Gait, M. J.; Karn, J., A molecular rheostat: Co-operative Rev binding to stem I of the Rev-response element modulates human immunodeficiency virus type-1 late gene expression. *J. Mol. Biol.* **1994,** *241* (2), 193-207.

45. Malim, M. H.; Hauber, J.; Le, S. Y.; Maizel, J. V.; Cullen, B. R., The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* **1989,** *338* (6212), 254-7.

46. Askjaer, P.; Jensen, T. H.; Nilsson, J.; Englmeier, L.; Kjems, J., The specificity of the CRM1-Rev nuclear export signal interaction is mediated by RanGTP. *J. Biol. Chem.* **1998,** *273* (50), 33414-22.

47. Fornerod, M.; Ohno, M.; Yoshida, M.; Mattaj, I. W., CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* **1997,** *90* (6), 1051-60.

48. Felber, B. K.; Hadzopoulou-Cladaras, M.; Cladaras, C.; Copeland, T.; Pavlakis, G. N., rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. *Proc. Natl. Acad. Sci. U. S. A.* **1989,** *86* (5), 1495-1499.

49. Malim, M. H.; Tiley, L. S.; McCarn, D. F.; Rusche, J. R.; Hauber, J.; Cullen, B. R., HIV-1 structural gene expression requires binding of the Rev trans-activator to its RNA target sequence. *Cell* **1990,** *60* (4), 675-83.

50. Huang, X. J.; Hope, T. J.; Bond, B. L.; McDonald, D.; Grahl, K.; Parslow, T. G., Minimal Rev-response element for type 1 human immunodeficiency virus. *J. Virol.* **1991,** *65* (4), 2131-4.

51. Kjems, J.; Brown, M.; Chang, D. D.; Sharp, P. A., 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* (3), 683-7.

52. Tiley, L. S.; Malim, M. H.; Tewary, H. K.; Stockley, P. G.; Cullen, B. R., Identification of a high-affinity RNA-binding site for the human immunodeficiency virus type 1 Rev protein. *Proc. Natl. Acad. Sci. U. S. A.* **1992,** *89* (2), 758-762.

53. Bartel, D. P.; Zapp, M. L.; Green, M. R.; Szostak, J. W., HIV-1 rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* **1991,** *67* (3), 529-536.

54. Battiste, J. L.; Mao, H.; Rao, N. S.; Tan, R.; Muhandiram, D. R.; Kay, L. E.; Frankel, A. D.; Williamson, J. R., α Helix-RNA major groove recognition in an HIV-1 Rev peptide-RRE RNA complex. *Science* **1996,** *273* (5281), 1547-1551.

55. Yuying Gosser; Thomas Hermann; Ananya Majumdar; Weidong Hu; Ronnie Frederick; Feng Jiang, W. X.; Patel, D. J., Peptide-triggered conformational switch in HIV-1 RRE RNA complexes. *Nat. Struc. Biol.* **2001,** *8*, 5.

56. Cook, K. S.; Fisk, G. J.; Hauber, J.; Usman, N.; Daly, T. J.; Rusche, J. R., Characterization of HIV-1 Rev protein: binding stoichiometry and minimal RNA substrate. *Nucleic Acids Res* **1991,** *19* (7), 1577-1583.

57. Malim, M. H.; Cullen, B. R., HIV-1 structural gene expression requires the binding of multiple Rev monomers to the viral RRE: Implications for HIV-1 latency. *Cell* **1991,** *65* (2), 241-248.

58. Olsen, H. S.; Cochrane, A. W.; Dillon, P. J.; Nalin, C. M.; Rosen, C. A., Interaction of the human immunodeficiency virus type 1 Rev protein with a structured region in env mRNA is dependent on multimer formation mediated through a basic stretch of amino acids. *Genes Dev.* **1990,** *4* (8), 1357-1364.

59. Daugherty, M. D.; D'Orso, I.; Frankel, A. D., A solution to limited genomic vapacity: Using adaptable binding surfaces to assemble the functional HIV Rev oligomer on RNA. *Mol. Cell.* **2008,** *31* (6), 824-834.

60. Heaphy, S.; Finch, J. T.; Gait, M. J.; Karn, J.; Singh, M., Human immunodeficiency virus type 1 regulator of virion expression, rev, forms nucleoprotein filaments after binding to a purine-rich "bubble" located within the rev-responsive region of viral mRNAs. *Proc. Natl. Acad. Sci. U. S. A.* **1991,** *88* (16), 7366-7370.

61. Daugherty, M. D.; Booth, D. S.; Jayaraman, B.; Cheng, Y.; Frankel, A. D., HIV Rev response element (RRE) directs assembly of the Rev homooligomer into discrete asymmetric complexes. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107* (28), 12481-12486.

62. DiMattia, M. A.; Watts, N. R.; Stahl, S. J.; Rader, C.; Wingfield, P. T.; Stuart, D. I.; Steven, A. C.; Grimes, J. M., Implications of the HIV-1 Rev dimer structure at 3.2 Å resolution for multimeric binding to the Rev response element. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107* (13), 5810-5814.

63. Daugherty, M. D.; Liu, B.; Frankel, A. D., Structural basis for cooperative RNA binding and export complex assembly by HIV Rev. *Nat. Struct. Mol. Biol.* **2010,** *17* (11), 1337-42.

64. Vallone, B.; Miele, A. E.; Vecchini, P.; Chiancone, E.; Brunori, M., Free energy of burying hydrophobic residues in the interface between protein subunits. *Proc. Natl. Acad. Sci. U. S. A.* **1998,** *95* (11), 6103-6107.

65. Fang, X.; Wang, J.; O'Carroll, Ina, P.; Mitchell, M.; Zuo, X.; Wang, Y.; Yu, P.; Liu, Y.; Rausch, Jason, W.; Dyba, Marzena, A.; Kjems, J.; Schwieters, Charles, D.; Seifert, S.; Winans, Randall, E.; Watts, Norman, R.; Stahl, Stephen, J.; Wingfield, Paul, T.; Byrd, R. A.; Le, Grice, Stuart, F. J.; Rein, A.; Wang, Y.-X., An unusual topological structure of the HIV-1 Rev response element. *Cell* **2013,** *155* (3), 594-605.

66. Froeyen, M.; Herdewijn, P., RNA as a target for drug design, the example of Tat–TAR interaction. *Curr. Top. Med. Chem.*, **2002**, 10 (2), 1123-1145.

67. Fulcher, A. J.; Jans, D. A., The HIV-1 Tat transactivator protein: a therapeutic target? *IUBMB Life* **2003,** *55* (12), 669-680.

68. Huber, P. W.; Cui, M.; Czarnik, A. W.; Mei, H.-Y., Binding of neomycin to the TAR element of HIV-1 RNA induces dissociation of Tat protein by an allosteric mechanism. *Biochemistry* **1998,** *37* (16), 5549-5557.

69. Cecchetti, V.; Parolin, C.; Moro, S.; Pecere, T.; Filipponi, E.; Calistri, A.; Tabarrini, O.; Gatto, B.; Palumbo, M.; Fravolini, A., 6-Aminoquinolones as new potential anti-HIV agents. *J. Med. Chem.* **2000,** *43* (20), 3799-3802.

70. Mischiati, C.; Finotti, A.; Sereni, A.; Boschetti, S.; Baraldi, P. G.; Romagnoli, R.; Feriotto, G.; Jeang, K.-T.; Bianchi, N.; Borgatti, M.; Gambari, R., Binding of hybrid molecules containing pyrrolo [2,1-c][1,4]benzodiazepine (PBD) and oligopyrrole carriers to the human immunodeficiency type 1 virus TAR-RNA. *Biochem. Pharmacol.* **2004,** *67* (3), 401-410.

71. Ratmeyer, L. S.; Vinayak, R.; Zon, G.; Wilson, W. D., An ethidium analog that binds with high specificity to a base-bulged duplex from the TAR RNA region of the HIV-1 genome. *J. Med. Chem.* **2002,** *35* (5), 966-968.

72. Peytou, V.; Condom, R.; Patino, N.; Guedj, R.; Aubertin, A.-M.; Gelus, N.; Bailly, C.; Terreux, R.; Cabrol-Bass, D., Synthesis and antiviral activity of ethidium−arginine conjugates directed against the TAR RNA of HIV-1. *J. Med. Chem.* **1999,** *42* (20), 4042-4053.

73. Hamy, F.; Brondani, V.; Flörsheimer, A.; Stark, W.; Blommers, M. J. J.; Klimkait, T., A new class of HIV-1 Tat antagonist acting through Tat−TAR inhibition. *Biochemistry* **1998,** *37* (15), 5086-5095.

74. Gelus, N.; Hamy, F.; Bailly, C., Molecular basis of HIV-1 TAR RNA specific recognition by an acridine tat-antagonist. *Bioorg. Med. Chem.* **1999,** *7* (6), 1075-1079.

75. Parolin, C.; Gatto, B.; Del Vecchio, C.; Pecere, T.; Tramontano, E.; Cecchetti, V.; Fravolini, A.; Masiero, S.; Palumbo, M.; Palu, G., New anti-human immunodeficiency virus type 1 6-aminoquinolones: Mechanism of action. *Antimicrob. Agents Chemother.* **2003,** *47* (3), 889-896.

76. Mei, H.-Y.; Cui, M.; Heldsinger, A.; Lemrow, S. M.; Loo, J. A.; Sannes-Lowery, K. A.; Sharmeen, L.; Czarnik, A. W., Inhibitors of protein-RNA complexation that target the RNA: Specific recognition of human immunodeficiency virus type 1 TAR RNA by small organic molecules. *Biochemistry* **1998,** *37* (40), 14204-14212.

77. Litovchick, A.; Evdokimov, A. G.; Lapidot, A., Arginine-aminoglycoside conjugates that bind to HIV transactivation responsive element RNA in vitro. *FEBS Lett.* **1999,** *445* (1), 73-79.

78. Evdokimov, A. G.; Lapidot, A., Aminoglycoside arginine conjugates that bind TAR RNA: Synthesis, characterization, and antiviral activity. *Biochemistry* **2000,** *39* (11), 2838-2852.

79. Seongwoo H., Natarajan. T., Karen K., Hong C., Akbar A., Yueh-Hsin P., Kuan-Teh J., and Tariq M. R., Discovery of a small molecule Tat-trans-activation-responsive RNA antagonist that potently inhibits human immunodeficiency virus-1 replication *J. Biol. Chem.* **2003,** *278*.

80. Davis, B.; Afshar, M.; Varani, G.; Murchie, A. I. H.; Karn, J.; Lentzen, G.; Drysdale, M.; Bower, J.; Potter, A. J.; Starkey, I. D.; Swarbrick, T.; Aboul-ela, F., Rational design of inhibitors of HIV-1 TAR RNA through the stabilisation of electrostatic "Hot Spots". *J. Mol. Biol.* **2004,** *336* (2), 343-356.

81. Murchie, A. I. H.; Davis, B.; Isel, C.; Afshar, M.; Drysdale, M. J.; Bower, J.; Potter, A. J.; Starkey, I. D.; Swarbrick, T. M.; Mirza, S.; Prescott, C. D.; Vaglio, P.; Aboul-ela, F.; Karn, J., Structure-based drug design targeting an inactive RNA conformation: exploiting the flexibility of HIV-1 TAR RNA. *J. Mol. Biol.* **2004,** *336* (3), 625-638.

82. Turner, J. J.; Fabani, M.; Arzumanov, A. A.; Ivanova, G.; Gait, M. J., Targeting the HIV-1 RNA leader sequence with synthetic oligonucleotides and siRNA: Chemistry and cell delivery. *BBA-Biomembranes* **2006,** *1758* (3), 290-300.

83. Walsh, A. P.; Rajwanshi, V. K.; Kumar, R.; Wengel, J.; Gait, M. J., Inhibition of HIV-1 Tat-dependent trans activation by steric block chimeric O-Methyl/LNA oligoribonucleotides. *Biochemistry* **2001,** *40* (48), 14645-14654.

84. Peffer, N. J.; Hanvey, J. C.; Bisi, J. E.; Thomson, S. A.; Hassman, C. F.; Noble, S. A.; Babiss, L. E., Strand-invasion of duplex DNA by peptide nucleic acid oligomers. *Proc. Natl. Acad. Sci. U. S. A.* **1993,** *90* (22), 10648-10652.

85. Møllegaard, N. E.; Buchardt, O.; Egholm, M.; Nielsen, P. E., Peptide nucleic acid. DNA strand displacement loops as artificial transcription promoters. *Proc. Natl. Acad. Sci. U. S. A.* **1994,** *91* (9), 3892-3895.

86. Kaushik, N.; Pandey, P. K.; Kashanchi, F.; Deng, L.; Pandey, V. N., Inhibition of Tat-mediated transactivation of HIV-1 LTR transcription by polyamide nucleic acid targeted to TAR hairpin element. *Biochemistry* **2000,** *39* (38), 11532-11539.

87. Kaushik, N.; Basu, A.; Palumbo, P.; Myers, R. L.; Pandey, V. N., Anti-TAR polyamide nucleotide analog conjugated with a membrane-permeating peptide inhibits human immunodeficiency virus type 1 production. *J. Virol.* **2002,** *76* (8), 3881-3891.

88. Turner, J. J.; Ivanova, G. D.; Verbeure, B.; Williams, D.; Arzumanov, A. A.; Abes, S.; Lebleu, B.; Gait, M. J., Cell-penetrating peptide conjugates of peptide nucleic acids (PNA) as inhibitors of HIV-1 Tat-dependent trans-activation in cells. *Nucl. Acids Res.* **2005,** *33* (21), 6837-6849.

89. Riguet, E.; Tripathi, S.; Chaubey, B.; Désiré, J.; Pandey, V. N.; Décout, J.-L., A peptide nucleic acid−neamine conjugate that targets and cleaves HIV-1 TAR RNA inhibits viral replication. *J. Med. Chem.* **2004,** *47* (20), 4806-4809.

90. Warashina, M.; Kuwabara, T.; Kato, Y.; Sano, M.; Taira, K., RNA-protein hybrid ribozymes that efficiently cleave any mRNA independently of the structure of the target RNA. *Proc. Natl. Acad. Sci. U. S. A.* **2001,** *98* (10), 5572-5577.

91. Van Melckebeke, H.; Devany, M.; Di Primo, C.; Beaurain, F.; Toulme, J.-J.; Bryce, D. L.; Boisbouvier, J., Liquid-crystal NMR structure of HIV TAR RNA bound to its SELEX RNA aptamer reveals the origins of the high stability of the complex. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** *105* (27), 9210-9215.

92. Watrin, M.; Von Pelchrzim, F.; Dausse, E.; Schroeder, R.; Toulme, J.-J., In vitro selection of RNA aptamers derived from a genomic human library against the TAR RNA element of HIV-1. *Biochemistry* **2009,** *48* (26), 6278-6284.

93. Arzumanov, A.; Gait, M. J.; Di Primo, C.; Toulme, J. J., 2'-O-Methyl-RNA hairpins generate loop-loop complexes and selectively inhibit HIV-1 Tat-mediated transcription. *Biochemistry* **2002,** *41* (40), 12186-12192.

94. Fedorov, Y.; Anderson, E. M.; Birmingham, A.; Reynolds, A.; Karpilow, J.; Robinson, K.; Leake, D.; Marshall, W. S.; Khvorova, A., Off-target effects by siRNA can induce toxic phenotype. *RNA* **2006,** *12* (7), 1188-1196.

95. Choudhury I; Wang J, R. A., Stein S, Pooyan S, Stein S, Leibowitz MJ., Inhibition of HIV-1 replication by a Tat RNA-binding domain peptide analog. *JAIDS* **1998,** *17*, 8.

96. Choudhury, I.; Wang, J.; Stein, S.; Rabson, A.; Leibowitz, M., Translational effects of peptide antagonists of Tat protein of human immunodeficiency virus type 1. *J. Gen. Virol.* **1999,** *80* (3), 777-782.

97. Hamy, F.; Felder, E. R.; Heizmann, G.; Lazdins, J.; Aboul-ela, F.; Varani, G.; Karn, J.; Klimkait, T., An inhibitor of the Tat/TAR RNA interaction that effectively suppresses HIV-1 replication. *Proc. Natl. Acad. Sci. U. S. A.* **1997,** *94* (8), 3548-3553.

98. Daelemans, D.; Schols, D.; Witvrouw, M.; Pannecouque, C.; Hatse, S.; van Dooren, S.; Hamy, F.; Klimkait, T.; de Clercq, E.; VanDamme, A.-M., A second target for the peptoid Tat/transactivation response element inhibitor CGP64222: Inhibition of

human immunodeficiency virus replication by blocking CXC-chemokine receptor 4-mediated virus entry. *Mol. Pharmacol.* **2000,** *57* (1), 116-124.

99. Hwang, S.; Tamilarasu, N.; Ryan, K.; Huq, I.; Richter, S.; Still, W. C.; Rana, T. M., Inhibition of gene expression in human cells through small molecule-RNA interactions. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** *96* (23), 12997-13002.

100. Friedler A, F. D., Luedtke NW, Tor Y, Loyter A, Gilon C, Development of a functional backbone cyclic mimetic of the HIV-1 Tat arginine-rich motif. *J. Biol. Chem.* **2000,** *275*, 23783-23789.

101. Ye, X.; Kumar, R. A.; Patel, D. J., Molecular recognition in the bovine immunodeficiency virus Tat peptide-TAR RNA complex. *Chem. Biol.* **1995,** *2* (12), 827-840.

102. Puglisi, J. D.; Chen, L.; Blanchard, S.; Frankel, A. D., Solution structure of a bovine immunodeficiency virus Tat–TAR peptide-RNA complex. *Science* **1995,** *270* (5239), 1200-1203.

103. Athanassiou, Z.; Dias, R. L. A.; Moehle, K.; Dobson, N.; Varani, G.; Robinson, J. A., Structural Mimicry of Retroviral Tat Proteins by Constrained β-hairpin peptidomimetics: Ligands with high affinity and selectivity for viral TAR RNA regulatory elements. *J. Am. Chem. Soc.* **2004,** *126* (22), 6906-6913.

104. Leeper, T. C.; Athanassiou, Z.; Dias, R. L. A.; Robinson, J. A.; Varani, G., TAR RNA recognition by a cyclic peptidomimetic of Tat protein. *Biochemistry* **2005,** *44* (37), 12362-12372.

105. Davidson, A.; Leeper, T. C.; Athanassiou, Z.; Patora-Komisarska, K.; Karn, J.; Robinson, J. A.; Varani, G., Simultaneous recognition of HIV-1 TAR RNA bulge and loop sequences by cyclic peptide mimics of Tat protein. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106* (29), 11931-11936.

106. Lalonde, M. S.; Lobritz, M. A.; Ratcliff, A.; Chamanian, M.; Athanassiou, Z.; Tyagi, M.; Wong, J.; Robinson, J. A.; Karn, J.; Varani, G.; Arts, E. J., Inhibition of both HIV-1 reverse transcription and gene expression by a cyclic peptide that binds the Tat-transactivating response element (TAR) RNA. *PLoS Pathog.* **2011,** *7* (5), e1002038.

107. Mammen, M.; Choi, S.-K.; Whitesides, G. M., Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors. *Angew. Chem. Int. Ed.* **1998,** *37* (20), 2754-2794.

108. Bryson, D. I.; Zhang, W.; Ray, W. K.; Santos, W. L., Screening of a branched peptide library with HIV-1 TAR RNA. *Mol. BioSyst.* **2009,** *5*, 1070-1073.

109. Bryson, D. I.; Zhang, W.; McLendon, P. M.; Reineke, T. M.; Santos, W. L., Toward targeting RNA structure: branched peptides as cell-permeable ligands to TAR RNA. *ACS Chem. Biol.* **2012,** *7* (1), 210-7.

110. Moazed, D.; Noller, H. F., Interaction of antibiotics with functional sites in 16S ribosomal RNA. *Nature* **1987,** *327* (6121), 389-394.

111. Cundliffe, E., On the nature of antibiotic binding sites in ribosomes. *Biochemistry* **1987,** *69* (8), 863-869.

112. Moazed, D.; Noller, H. F., Chloramphenicol, erythromycin, carbomycin and vernamycin B protect overlapping sites in the peptidyl transferase region of 23S ribosomal RNA. *Biochemistry* **1987,** *69* (8), 879-884.

113. Zapp, M. L.; Stern, S.; Green, M. R., Small molecules that selectively block RNA binding of HIV-1 rev protein inhibit rev function and viral production. *Cell* **1993,** *74* (6), 969-978.

114. Umezawa, S.; Tsuchiya, T., Total synthesis and chemical modification of the aminoglycoside antibiotics. *Aminoglycoside Antibiotics*, Umezawa, H.; Hooper, I., Eds. Springer Berlin Heidelberg: 1982; Vol. 62, pp 37-110.

115. Luedtke, N. W.; Baker, T. J.; Goodman, M.; Tor, Y., Guanidinoglycosides: A novel family of RNA ligands. *J. Am. Chem. Soc.* **2000,** *122* (48), 12035-12036.

116. Baker, T. J.; Luedtke, N. W.; Tor, Y.; Goodman, M., Synthesis and anti-HIV activity of guanidinoglycosides. *J. Org. Chem.* **2000,** *65* (26), 9054-9058.

117. Fernandez-Saiz, K. L. M.; Rigl, C. T.; Kumar, A.; Ragunathan, K. G.; McConnaughie, A. W.; Boykin, D. W.; Schneider, H.-J.; Wilson, W. D., Design and analysis of molecular motifs for specific recognition of RNA. *Bioorg. Med. Chem.* **1997,** *5* (6), 1157-1172.

118. Zhao, M.; Ratmeyer, L.; Peloquin, R. G.; Yao, S.; Kumar, A.; Spychala, J.; Boykin, D. W.; David Wilson, W., Small changes in cationic substituents of diphenylfuran derivatives have major effects on the binding affinity and the binding mode with RNA helical duplexes. *Bioorg. Med. Chem.* **1995,** *3* (6), 785-794.

119. Xiao, G.; Kumar, A.; Li, K.; Rigl, C. T.; Bajic, M.; Davis, T. M.; Boykin, D. W.; Wilson, W. D., Inhibition of the HIV-1 RRE–Rev complex formation by unfused aromatic cations. *Bioorg. Med. Chem.* **2001,** *9* (5), 1097-113.

120. Li, K.; Davis, T. M.; Bailly, C.; Kumar, A.; Boykin, D. W.; Wilson, W. D., A heterocyclic inhibitor of the RRE–Rev complex binds to RRE as a dimer. *Biochemistry* **2001,** *40* (5), 1150-1158.

121. DeJong, E. S.; Chang, C. E.; Gilson, M. K.; Marino, J. P., Proflavine acts as a Rev inhibitor by targeting the high-affinity Rev binding site of the Rev responsive element of HIV-1. *Biochemistry* **2003,** *42* (26), 8035-46.

122. Lacourciere, K. A.; Stivers, J. T.; Marino, J. P., 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* (19), 5630-5641.

123. Nakatani, K.; Sando, S.; Saito, I., Recognition of a single guanine bulge by 2-acylamino-1,8-naphthyridine. *J. Am. Chem. Soc.* **2000,** *122* (10), 2172-2177.

124. Nakatani, K.; Sando, S.; Kumasawa, H.; Kikuchi, J.; Saito, I., Recognition of guanine−guanine mismatches by the dimeric form of 2-amino-1,8-naphthyridine. *J. Am. Chem. Soc.* **2001,** *123* (50), 12650-12657.

125. Hagihara, S.; Kumasawa, H.; Goto, Y.; Hayashi, G.; Kobori, A.; Saito, I.; Nakatani, K., Detection of guanine–adenine mismatches by surface plasmon resonance sensor carrying naphthyridine–azaquinolone hybrid on the surface. *Nucleic Acids Res.* **2004,** *32* (1), 278-286.

126. Nakatani, K.; Horie, S.; Goto, Y.; Kobori, A.; Hagihara, S., Evaluation of mismatch-binding ligands as inhibitors for Rev–RRE interaction. *Bioorg. Med. Chem.* **2006,** *14* (15), 5384-5388.

127. Kumagai, I.; Takahashi, T.; Hamasaki, K.; Ueno, A.; Mihara, H., Construction of HIV Rev peptides containing peptide nucleic acid that bind HIV RRE IIB RNA. *Bioorg. Med. Chem. Lett.* **2000,** *10* (4), 377-379.

128. Kumagai, I.; Takahashi, T.; Hamasaki, K.; Ueno, A.; Mihara, H., HIV Rev peptides conjugated with peptide nucleic acids and their efficient binding to RRE RNA. *Bioorg. Med. Chem. Lett.* **2001,** *11* (9), 1169-1172.

129. Takahashi, T.; Hamasaki, K.; Ueno, A.; Mihara, H., Construction of peptides with nucleobase amino acids: design and synthesis of the nucleobase-conjugated peptides derived from HIV-1 rev and their binding properties to HIV-1 RRE RNA. *Bioorg. Med. Chem.* **2001,** *9* (4), 991-1000.

130. Belousoff, M.; Gasser, G.; Graham, B.; Tor, Y.; Spiccia, L., Binding of HIV-1 TAR mRNA to a peptide nucleic acid oligomer and its conjugates with metal-ion-binding multidentate ligands. *J. Biol. Inorg. Chem.* **2009,** *14* (2), 287-300.

131. Lee, Y.; Hyun, S.; Kim, H.; Yu, J., Amphiphilic helical peptides containing two acridine moieties display picomolar affinity toward HIV-1 RRE and TAR. *Angew. Chem. Int. Ed.* **2008,** *47* (1), 134-137.

132. Mills, N. L.; Daugherty, M. D.; Frankel, A. D.; Guy, R. K., An α-helical peptidomimetic inhibitor of the HIV-1 Rev–RRE interaction. *J. Am. Chem. Soc*. **2006,** *128* (11), 3496-3497.

133. Kirk, S. R.; Luedtke, N. W.; Tor, Y., Neomycin−acridine conjugate: A potent inhibitor of RRE–Rev binding. *J. Am. Chem. Soc.* **2000,** *122* (5), 980-981.

134. Luedtke, N. W.; Liu, Q.; Tor, Y., RNA−ligand interactions: Affinity and specificity of aminoglycoside dimers and acridine conjugates to the HIV-1 Rev response element. *Biochemistry* **2003,** *42* (39), 11391-11403.

135. Wang, H.; Tor, Y., Dimeric aminoglycosides: Design, synthesis and RNA binding. *Bioorg. Med. Chem. Lett.* **1997,** *7* (14), 1951-1956.

136. Tok, J. B. H.; Huffman, G. R., Enhanced binding of aminoglycoside dimers to a "dimerized" A-site 16S rRNA construct. *Bioorg. Med. Chem. Lett.* **2000,** *10* (14), 1593-1595.

137. Kirk, S. R.; Tor, Y., $tRNA^{Phe}$ binds aminoglycoside antibiotics. *Bioorg. Med. Chem.* **1999,** *7* (9), 1979-1991.

138. Hendrix, M.; Priestley, E. S.; Joyce, G. F.; Wong, C. H., Direct observation of aminoglycoside−RNA interactions by surface plasmon resonance. *J. Am. Chem. Soc.* **1997,** *119* (16), 3641-3648.

139. Ahn, D. R.; Yu, J., Library construction of neomycin–dipeptide heteroconjugates and selection against RRE RNA. *Bioorg. Med. Chem.* **2005,** *13* (4), 1177-1183.

140. Hyun, S.; Lee, K. H.; Yu, J., A strategy for the design of selective RNA binding agents. Preparation and RRE RNA binding affinities of a neomycin-peptide nucleic acid heteroconjugate library. *Bioorg. Med. Chem. Lett.* **2006,** *16* (18), 4757-4759.

141. Geistlinger, T. R.; Guy, R. K., An inhibitor of the interaction of thyroid hormone receptor β and glucocorticoid interacting protein 1. *J. Am. Chem. Soc.* **2001,** *123* (7), 1525-1526.

142. Cox, J. A.; Comte, M.; Fitton, J. E.; DeGrado, W. F., The interaction of calmodulin with amphiphilic peptides. *J. Biol. Chem.* **1985,** *260* (4), 2527-2534.

143. Pain, A.; Renauld, H.; Berriman, M.; Murphy, L.; Yeats, C. A.; Weir, W.; Kerhornou, A.; Aslett, M.; Bishop, R.; Bouchier, C.; Cochet, M.; Coulson, R. M. R.; Cronin, A.; de Villiers, E. P.; Fraser, A.; Fosker, N.; Gardner, M.; Goble, A.; Griffiths-Jones, S.; Harris, D. E.; Katzer, F.; Larke, N.; Lord, A.; Maser, P.; McKellar, S.; Mooney, P.; Morton, F.; Nene, V.; O'Neil, S.; Price, C.; Quail, M. A.; Rabbinowitsch, E.; Rawlings, N. D.; Rutter, S.; Saunders, D.; Seeger, K.; Shah, T.; Squares, R.; Squares, S.; Tivey, A.; Walker, A. R.; Woodward, J.; Dobbelaere, D. A. E.; Langsley, G.;

Rajandream, M.-A.; McKeever, D.; Shiels, B.; Tait, A.; Barrell, B.; Hall, N., Genome of the host-cell transforming parasite theileria annulata compared with T. parva. *Science* **2005,** *309* (5731), 131-133.

144. Hyun, S.; Kim, H. J.; Lee, N. J.; Lee, K. H.; Lee, Y.; Ahn, D. R.; Kim, K.; Jeong, S.; Yu, J., α-Helical peptide containing N,N-dimethyl lysine residues displays low-nanomolar and highly specific binding to RRE RNA. *J. Am. Chem. Soc.* **2007,** *129* (15), 4514-4515.

145. Lee, S. J.; Hyun, S.; Kieft, J. S.; Yu, J., An approach to the construction of tailor-made amphiphilic peptides that strongly and selectively bind to hairpin RNA targets. *J. Am. Chem. Soc.* **2009,** *131* (6), 2224-2230.

146. Hyun, S.; Na, J.; Lee, S. J.; Park, S.; Yu, J., RNA Grooves can accommodate disulfide-bridged bundles of α-helical peptides. *ChemBioChem.* **2010,** *11* (6), 767-770.

147. Schafmeister, C. E.; Po, J.; Verdine, G. L., An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides. *J. Am. Chem. Soc.* **2000,** *122* (24), 5891-5892.

148. Wilson, W. D.; Li, K., Targeting RNA with small molecules. *Curr. Med. Chem.* **2000,** *7* (1), 73-98.

149. Jin, Y.; Cowan, J. A., Targeted cleavage of HIV Rev response element RNA by metallopeptide complexes. *J. Am. Chem. Soc.* **2005,** *128* (2), 410-411.

150. Hyun, S.; Han, A.; Yu, J., Photocrosslinking of RNA and photoMet-containing amphiphilic α-helical peptides. *ChemBioChem.* **2009,** *10* (6), 987-989.

151. Tanaka, Y.; Kohler, J. J., Photoactivatable crosslinking sugars for capturing glycoprotein interactions. *J. Am. Chem. Soc.* **2008,** *130* (11), 3278-3279.

152. Nakashima, H.; Hashimoto, M.; Sadakane, Y.; Tomohiro, T.; Hatanaka, Y., Simple and versatile method for tagging phenyldiazirine photophores. *J. Am. Chem. Soc.* **2006,** *128* (47), 15092-15093.

153. Guan, L.; Disney, M. D., Recent advances in developing small molecules targeting RNA. *ACS Chem. Biol.* **2011,** *7* (1), 73-86.

154. Paillart, J.-C.; Shehu-Xhilaga, M.; Marquet, R.; Mak, J., Dimerization of retroviral RNA genomes: an inseparable pair. *Nat. Rev. Micro.* **2004,** *2* (6), 461-472.

155. Skripkin, E.; Paillart, J. C.; Marquet, R.; Ehresmann, B.; Ehresmann, C., Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **1994,** *91* (11), 4945-4949.

156. Ennifar, E.; Yusupov, M.; Walter, P.; Marquet, R.; Ehresmann, B.; Ehresmann, C.; Dumas, P., The crystal structure of the dimerization initiation site of genomic HIV-1 RNA reveals an extended duplex with two adenine bulges. *Structure* **1999,** *7* (11), 1439-1449.

157. Ennifar, E.; Walter, P.; Ehresmann, B.; Ehresmann, C.; Dumas, P., Crystal structures of coaxially stacked kissing complexes of the HIV-1 RNA dimerization initiation site. *Nat. Struct. Mol. Biol.* **2001,** *8* (12), 1064-1068.

158. Vicens, Q.; Westhof, E., Crystal structure of paromomycin docked into the eubacterial ribosomal decoding a site. *Structure* **2001,** *9* (8), 647-658.

159. Ennifar, E.; Aslam, M. W.; Strasser, P.; Hoffmann, G.; Dumas, P.; van Delft, F. L., Structure-guided discovery of a novel aminoglycoside conjugate targeting HIV-1 RNA Viral Genome. *ACS Chem. Biol.* **2013,** *8* (11), 2509-2517.

160. Sobczak, K.; Michlewski, G.; de Mezer, M.; Kierzek, E.; Krol, J.; Olejniczak, M.; Kierzek, R.; Krzyzosiak, W. J., Structural diversity of triplet repeat RNAs. *J. Biol. Chem.* **2010,** *285* (17), 12755-12764.

161. Paul, S.; Dansithong, W.; Kim, D.; Rossi, J.; Webster, N. J.; Comai, L.; Reddy, S., Interaction of musleblind, CUG‐BP1 and hnRNP H proteins in DM1-associated aberrant IR splicing. *EMBO J.* **2006,** *25* (18), 4271-4283.

162. Childs-Disney, J. L.; Hoskins, J.; Rzuczek, S. G.; Thornton, C. A.; Disney, M. D., Rationally designed small molecules targeting the RNA that causes myotonic dystrophy type 1 are potently bioactive. *ACS Chem. Biol.* **2012,** *7* (5), 856-862.

163. Rzuczek, S. G.; Gao, Y.; Tang, Z.-Z.; Thornton, C. A.; Kodadek, T.; Disney, M. D., Features of modularly assembled compounds that impart bioactivity against an RNA target. *ACS Chem. Biol.* **2013,** *8* (10), 2312-2321.

Chapter 2. Development of First Generation Branched Peptides Ligand for HIV-1 TAR RNA

Attributions

This chapter was taken from Bryson, D. I.; Zhang, W. Y.; Ray, W. K.; Santos, W. L. Screening of a branched peptide library with HIV-1 TAR RNA. *Mol. BioSyst.* **2009**, 5, 1070-1073, and Bryson, D. I.; Zhang, W.; McLendon, P. M.; Reineke, T. M.; Santos, W. L. Toward targeting RNA structure: Branched peptides as cell-permeable ligands to TAR RNA. *ACS Chem. Biol*. **2012**, 7, 210-217. The author of this dissertation was responsible for re-synthesis, purification, and characterization of most hit peptide compounds. Interpretation of most binding affinity data, and synthesis of several RNAs. She also performed a significant number of dot blot assays and contributed to the editing of the manuscript. David I. Bryson was responsible for the synthesis of the peptide library, high throughput screening, RNA design and preparation, electrophoretic mobility shift experiments, interpretation of binding affinity data, Hill plot analyses, and liquid chromatography coupled with electrospray ionization mass spectrometry. He also contributed significantly to the writing and editing of the manuscript. Dr. W. Keith Ray of the VT mass spectrometry incubator performed MALDI analysis of hit compounds and performed the majority of sequence deconvolutions. Patrick M. McLendon of the Reineke Group performed all cellular assays. The final manuscript was prepared by Dr. Webster L. Santos. [Reproduced by permission of the Royal Society of Chemistry and *ACS Chem. Biol.* 2012, 7, 210-221. Copyright 2012.]

Abstract

The *trans*-activating response element (TAR) RNA is a 59-residue hairpin loop structure located at the 5′ end of the nascent viral transcripts and is recognized by the HIV-1 Tat protein. It is responsible for the efficient generation of full RNA transcripts and can force the virus to remain in dormant stages. Therefore, TAR RNA is an attractive target for drug design due to its potential to inhibit acute and chronic HIV infections. Herein, we report the discovery of non-toxic and cell permeable branched peptide ligands that bind to TAR RNA in the submicromolar range from an on-bead high throughput screening of a 3.3.3-library comprised of 4,096 unique compounds. We found that one of the peptides, FL4, displayed a tighter binding affinity than its native protein counterpart, Tat (600 nM vs 780 nM, respectively). We demonstrate that "branching" in peptides provides multivalent interactions, which increase binding affinities for RNA. Mutation experiments indicated that FL4 interacts with both the loop region and the bulge region.

2.1 Introduction

For decades, ribonucleic acid (RNA) was recognized for playing a key role in carrying genetic information. The recently discovered functions of RNA have expanded its cellular roles. For example, RNAs, like ribozymes, are regarded as functional biomolecules that have inherent enzyme-like properties for catalyzing chemical reactions in the absence of proteins.[1] MicroRNAs (miRNAs) are non-coding RNAs that can negatively regulate gene expression.[2] RNAs also mediate a variety of biological effects such as transcription, splicing, replication, and transport.[3] This exploration in RNA biology marks the important roles of RNA in a living cell and highlights the potential of targeting a functional RNA for the treatment of diseases. Despite the considerable effort in utilizing RNA as a drug target,

there are comparatively few compounds that target RNA with high affinity and specificity. Therefore, targeting RNA still remains a challenge and continues to be a subject of intense investigation.[4] One of the main issues for targeting RNA is the involvement of a large surface area for recognition. Furthermore, the conformational dynamics of RNA often results in structural differences making it difficult for the *de novo* design of RNA ligands.[5]

Although RNA is chemically similar to DNA, the three dimensional structures are quite different. DNA is typically helical because it is double stranded. RNA is single stranded and folds into a variety of secondary structures, including hairpins, bulges, stems, loops, pseudoknots, and turns, which are similar to protein pockets. Further folding of these local structures can give rise to tertiary structures that are unique to specific RNA constructs and potentially allow the RNA to be targeted selectively.[6] The major groove of an A-form RNA is deeper and narrower than a B-form DNA, and the minor groove of the A-form RNA is shallower than the B-form DNA. As a result, the most specific ligands for DNA binding are usually not specific for RNA binding and the RNA architecture is recognized in a manner that is not solely dependent on Watson–Crick base pairing. For example, the ability of small molecules to target the tertiary structures of ribosomal RNA (rRNA) has been well demonstrated by aminoglycoside, macrolide, oxazolidinone, and tetracycline antibiotics.[7-12] However, the development of RNA-binding small molecules is far from a straightforward process. Small molecules can simultaneously achieve high affinity and cell permeability and yet not be selective towards their target. Due to their limited surface area, the small molecules can insert into the major groves of RNA and DNA easily, which may be the main reason for the lower selectivity and higher toxicity.[4] RNA-targeted gene repression can be attained using antisense or RNA interference

technologies.[13] However, they heavily rely on base pairing, so they do not distinguish the three-dimensional architectures. Additionally, their poor cell permeability, cytotoxicity and stability profiles still remain a challenge, although several siRNAs are in clinical trials.[13-14] The most desired approach for rational, structure-based RNA-targeted drug discovery is still in its infancy, making the design of RNA ligands difficult.[15]

The human immunodeficiency virus type 1 (HIV-1) is a global issue despite significant advancements in its research, treatment, and prevention. One viral RNA, *trans*-activating response element (TAR) RNA, controls the most essential viral replication process of transcription via stimulating transcriptional elongation from the viral long terminal repeat (LTR) by binding to its target Tat protein.[16] TAR RNA is a 59-residue hairpin loop structure located at the 5′ end of the nascent viral transcripts and is a conserved region that is not susceptible to mutation during replication.[17-19] Deletion studies have demonstrated that the region between 19 to 42 spans the minimum number of nucleotides that are required for recognition and sufficient binding of Tat. This short RNA forms a stem-loop structure, which contains a hexa-loop and a three-pyrimidine rich bulge (Figure 2.1).[20] When the Tat protein is not associated with HIV-1 TAR, basal viral transcription is very low, and short RNA transcripts are generated.[21] When the arginine-rich domain of Tat binds the U-rich bulge region of TAR RNA, another arginine residue helps to recognize it and form the Tat – TAR complex, a cofactor complex of cyclin T1 (CycT1) and cyclin-dependent kinase 9 (CDK9) that induces a conformational change in Tat. This complex also binds to the hairpin loop region of TAR RNA.[22-23] This process stimulates the efficient transcription of the long terminal repeat (LTR).[19] Therefore,

blocking this Tat–TAR interaction is a potent strategy for controlling the proliferation of the virus and provides a potential anti-HIV therapy.

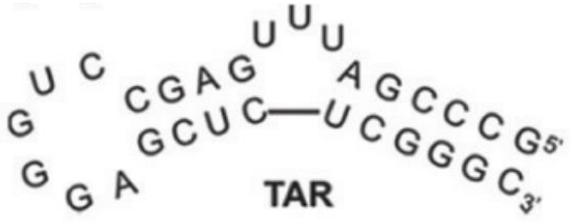


Figure 2.1 The sequence and secondary structure of HIV-1 TAR RNA.

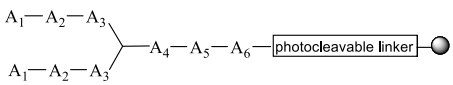
The Tat/TAR regulation pathway has become a high profile drug target precisely for its critical role in the proliferation of HIV-1.[24] Indeed, targeting RRE may provide an important opportunity for developing drugs that do not induce drug resistance in the virus because of the pressure to preserve this element from an RNA-protein binding standpoint.[25] Owing to the therapeutic potential of the Tat/TAR regulation pathway, many ligands have been designed to interrupt the Tat/TAR interaction with limited clinical success. Small molecules such as neomycin B, as well as other aminoglycoside derivatives, have demonstrated submicromolar binding ligands of TAR.[26-28] Other molecules like quinoxaline, argininamide, purine analogs, bis-guanidine compounds, tripeptides and pyrrole derivatives have been shown to interact with the bulge region of RNA.[4, 29-34] However, their lack of binding specificity, poor cell permeability, and toxicity make them therapeutically undesirable. In addition, large macromolecules like TAR RNA decoys are currently being investigated.[35-36] Furthermore, compounds of intermediate size such as oligomeric amines and β-hairpin peptidomimetics are subject to increased attention in part because of better overlap with the large surface area of RNA.[37-39] In particular, a β-hairpin cyclic peptide L50 has recently been shown to inhibit the tat dependent transcription process and the reverse transcription step by binding to TAR RNA.[40]

While extensive efforts have been made to develop ligands that can recognize RNA structures specifically and inhibit HIV-1 replication, none of them have made it to the clinic. In an effort to provide a general platform for targeting RNA structures, we focused on inhibiting RNA-protein interactions involving Tat and TAR RNA with medium-sized molecules that may not suffer cell permeability and cytotoxicity issues. Herein, we report an approach that uses branched peptides (BPs) to target RNA. We believe that the branches in peptides have the potential for multivalent interactions with various regions of TAR—a desirable property that can increase selectivity and affinity.[41] The combination of various peptide architectures allows the highly populated solution conformations of TAR RNA to bind the best complementary compounds. Furthermore, different amino acids can be incorporated in the library to achieve functional diversity, and their synthetic accessibility is straightforward. Indeed, rapid generation of a large number of peptides on beads with the split and pool technique can provide access to libraries of varying molecular weight. In this way, we were able to generate a 3.3.3-branched peptide library that was comprised of 4,096 unique compounds. We found that one of the peptides (FL4), displayed a tighter binding affinity than its native protein counterpart, Tat (600 nM vs 780 nM, respectively). We have demonstrated that 'branching' in peptides provides multivalent interactions, which increases binding affinities for RNA. Mutation experiments indicated that FL4 interacts with both the loop region and the bulge region. Further, we demonstrate that BPs are cell permeable and non-toxic, making them excellent chemical biological tools for targeting well-defined RNA structures and useful in anti-HIV drug discovery.

2.2 Results and Discussion

2.2.1 Library Design and On-Bead High-Throughput Screening

The BP library was synthesized in triplicate on Tentagel beads via split and pool synthesis. The library was prepared such that there were three variable amino acid positions at both the N- and C-termini (A_1-A_3 and A_4-A_6, respectively). Each variable position contained one of four possible amino acid monomers (Figure 2.2). The branches A_1-A_3 had identical sequences in order to simplify the *de novo* sequencing by MALDI/mass spectrometry (MS) and were branched through the α- and ε-amino groups on lysine. Each of the four possible sidechains at variable positions A_1-A_6 were chosen for its potential to interact with the TAR RNA. In position A_2, for example, we selected amino acids with functional groups that can interact with the RNA through hydrophobic interaction (Val), electrostatic attraction (Arg), hydrogen bonding (Asn), and π stacking (His). The goal was to present varying functional groups for the RNA target. Using standard solid phase peptide synthesis, 4,096 compounds were generated by the split and pool technique. It was important that the peptides were linked to the bead by a photocleavable linker (3-amino-3-(2-nitrophenyl)propionic acid, ANP), which allowed an on-bead high throughput screen and the efficient release of hit molecules by UV-irradiation(Figure 2.2).[42]



A_1	A_2	A_3	A_4	A_5	A_6
Asp	Arg	Tyr	His	Tyr	Phe
Arg	Asn	Leu	Val	Arg	Leu
Tyr	His	Trp	Asn	Ala	Asp
His	Val	Ala	Trp	Gln	Ser

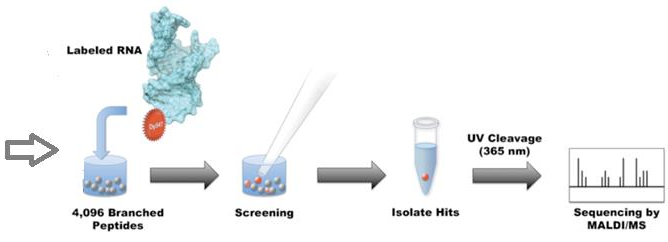


Figure 2.2 High throughput screening of a branched peptide library with TAR RNA reveals peptide hits that are sequenced by MALDI-TOF.

The BP library was subjected to an on-bead high throughput screen against DY547 labeled HIV-1 TAR RNA.[41] We initially transferred the library containing 3 copies of each member in an Eppendorff tube and washed them with water (5x) and buffer (3x) containing 50 mM Tris·HCl, 20 mM KCl, and 0.1% Triton X-100 at pH 7.4.[43] In order to address the issue of selectivity, unlabeled α-synuclein mRNA, a 340-nucleotide long RNA, was added in the incubation mixture to minimize non-specific binding. We hypothesized that longer RNA sequences were likely to bind non-specifically to our library, and abolishing promiscuous binders could decrease the potential for off-target binding. After washing, 3.4 nM DY547 labeled HIV-1 TAR RNA was added. Specific binding of the target RNA to the peptide resulted in increased fluorescence of the bead, as indicated by fluorescence microscopy. Seventeen beads were selected as possible hits. As expected, rescreening of the same beads with DY547-TAR RNA under a more stringent condition still yielded fluorescence, confirming RNA binding. These peptides were photocleaved via UV irradiation and then sequenced by MALDI MS-MS analysis.[44] Hit compounds were resynthesized so that a single N-terminus was labeled with fluorescein with fluorescein isothiocyanate (FITC) for further biophysical characterizations.

2.2.2 Binding Affinities of Hit Branched Peptides to HIV-1 TAR

Following a standard dot blot assay, ^{32}P-labeled TAR RNA was titrated with increasing concentration of BPs. The sequence and dissociation constants of the hit peptides are shown in Table 2.1. The BP hits contained significant numbers of Arg moieties, which was not surprising due to the strong electrostatic attraction between the positively charged side chain of the peptides and the negatively charged phosphate backbone of the RNA target. One peptide, **FL10**, lacked a positive charge and was determined to be a false

Table 2.1. Binding constants and molecular weights of hit BPs.

Peptide	Sequence	K_d (μM)	MW (g/mol)
FL4	$(RRW)_2$*HAL	0.6 ± 0.1	1463.74
FL6	$(RRY)_2$*VRL	1.2 ± 0.2	1464.77
FL7	$(RRW)_2$*HYD	1.2 ± 0.2	1557.77
FL12	$(RRW)_2$*HAS	1.9 ± 0.3	1437.66
FL5	$(RRL)_2$*NRF	2.0 ± 0.3	1413.72
FL1	$(RRL)_2$*WYL	2.5 ± 0.3[a]	1458.80
FL17	$(RRL)_2$*HRF	3.1 ± 0.3	1436.76
FL9	$(RRA)_2$*NYF	3.7 ± 1.9	1336.55
FL2	$(YRA)_2$*HRF	7.5 ± 0.9	1366.58
FL16	$(RRL)_2$*HYL	7.7 ± 1.1	1409.73
FL8	$(YRL)_2$*WRL	7.8 ± 1.9[a]	1465.79
FL11	$(RRA)_2$*VYF	8.1 ± 1.3	1321.58
FL3	$(HRW)_2$*WAS	8.7 ± 3.1[a]	1448.64
FL14	$(RRY)_2$*VQL	20.2 ± 2.9	1436.71
FL10	$(DNL)_2$*HYF	NB[a]	1277.38
FL13	$(RRY)_2$*NQD	> 75	1453.61
FL15	$(RRA)_2$*VRD	> 75	1282.51
T4-1	RRWGHAL	> 75	
Tat_{47-57}	GRKKRRQRRR	0.78	

[a] Performed with a final concentration of 5% DMSO. No binding observed (NB). * = Lysine branching unit. Each value is an average of at least three experiments.

positive hit. Our results indicated fourteen peptides had binding affinities in the low micromolar range. In particular, **FL4**, **FL6,** and **FL7** contained four positively charged Arg residues in the N-termini and had K_d values of 600 nM, 1.2 μM, and 1.2 μM, respectively. Gratifyingly, **FL4** bound much tighter than its native protein counterpart, Tat (K_d = 780 nM). This tight binding was not directly associated with numbers of Arg; for example, **FL6** contained six Arg residues exhibited 2-fold lower binding affinities for TAR. Further sequence analysis indicated that amino acids in the C-terminus affected the binding affinity. Replacing AL to YD (**FL4** to **FL7**) or replacing L to S (**FL4** to **FL12**) resulted in decreased binding affinities (2-fold and 3-fold, respectively). This indicates that the hydrophobic residues at the C-terminus have specific interactions with TAR RNA. Therefore,

substituting these groups with hydrophilic groups diminished the interaction. In general, however, a hydrophobic moiety was desired in the C-terminus and a hydrophilic moiety was desired in the N-terminus (Table 2.1**)**.

From the pool of seventeen BPs, three peptides (**FL1**, **FL3** & **FL8**) did not contain the consensus Arg–Arg in the N-termini. The dissociation constants of these compounds were evaluated by a dot blot assay with 5% DMSO in the buffer. Lacking positively charged residues may decrease the solubility, and therefore, decrease the binding affinities. Finally, it is noteworthy that the electrostatic interactions mediated by the two Arg moieties in the N-termini did not contribute to the majority of the binding affinity with TAR since **FL13** and **FL15** had K_d values > 75 μM.

2.2.3 The Branched N-terminus Imparts Multivalency

To determine whether the additional "branch" in our peptides results in a complementary increase in binding affinity as a consequence of multivalency, we synthesized **T4-1** as truncated variants of the strongest binder, **FL4**. **T4-1** (RRWGHAL) featured a single N-terminus and a C-terminus that was identical to **FL4**. We substituted the branching Lys unit with Gly in **T4-1** to preserve the spacing between the N- and C-termini and to avoid incorporating any additional functionality compared to **FL4**. To our delight, we observed a > 125-fold decrease in binding affinity for **T4-1** (K_d > 75 μM) compared to **FL4** when measured by dot blot assay (Table 2.1). Compare to the original sequence of **FL4** two positively charged side-chains were omitted in **T4-1**. Thus, the decreased binding affinity may also due to the smaller number of positive charges. However, from the sequence analysis, the presence of Arg–Arg in the N-termini did not necessarily result in effective RNA binding, and there was no direct relationship between

Arg residue numbers and binding affinity. Furthermore, mutation of only one amino acid in the C-terminus (**FL4** to **FL12**) resulted in a decreased binding affinity. Taken together, these results suggest that electrostatic interactions are not solely responsible for high affinity binders but that all branches of the 3.3.3 peptides are responsible for tight binding with the RNA.

2.2.4 Selectivity of FL4 Toward TAR RNA Tertiary Structure and Binding Site Determination

To confirm that the dissociation constants determined from the dot blot assay are reliable, an electrophoretic mobility shift assay (EMSA) was performed with **FL4**. As shown in Figure 2.3A, the dissociation constants determined with the dot blot and EMSA were in excellent agreement (600 nM and 500 nM, respectively), consistent with previous reports using EMSA. To characterize the selectivity against other RNA structures, binding

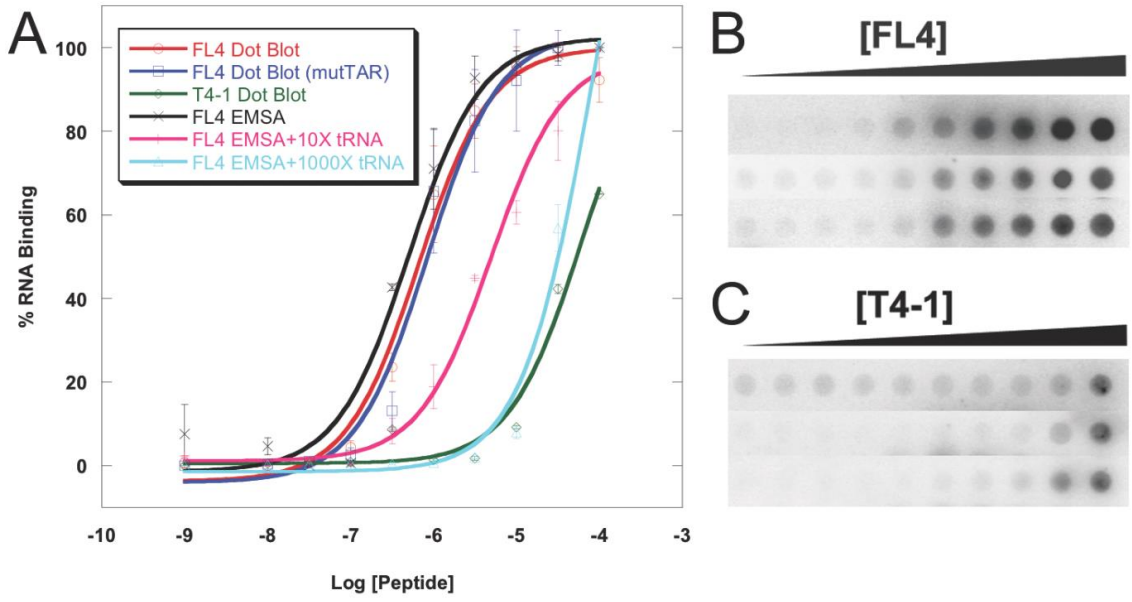


Figure 2.3 Titration curves for RNA binding. (A) Binding curves for **FL4** and **T4-1** with or without competitor tRNA using dot blot and EMSA. Image of dot blot filter-binding assay of (B) **FL4** and (C) **T4-1**.

affinities between **FL4** and TAR RNA were measured in the presence of excess bacterial tRNA.[37, 46] Addition of 10- and 1000-fold excess of tRNA caused a shift in the binding affinity for ^{32}P-labeled TAR RNA, suggesting **FL4** was partially selective in the presence of competitor tRNAs (Figure 2.3A).

To further determine the selectivity of **FL4** for TAR RNA, we measured the K_d for TAR containing a single mutation (24C>U) at the 3-nt bulge, the site where Tat is known to bind, and the K_d was found to be 0.8 μM. The slight change in K_d indicates the peptide may not directly interact with U or C in the bulge region. Indeed, there is the evidence that TAR (UCU in bulge region) presents in several clades of HIV-1 and is expected to maintain the native structure of TAR.[47] Several mutants were designed to determine whether **FL4** could discriminate between the native tertiary structure of TAR and related analogs with missing structural elements. As a result, a more dramatic increase in K_d was observed when the 3-nt bulge region of the TAR RNA tertiary structure was removed (bulgeless TAR, 6.6 ± 1.3 μM) or the size of the apical loop region was decreased by 2-nt (tetraloop TAR 5.5 ±2.0 μM) (Figure 2.4). The approximately ten-fold decrease in binding affinity is exciting

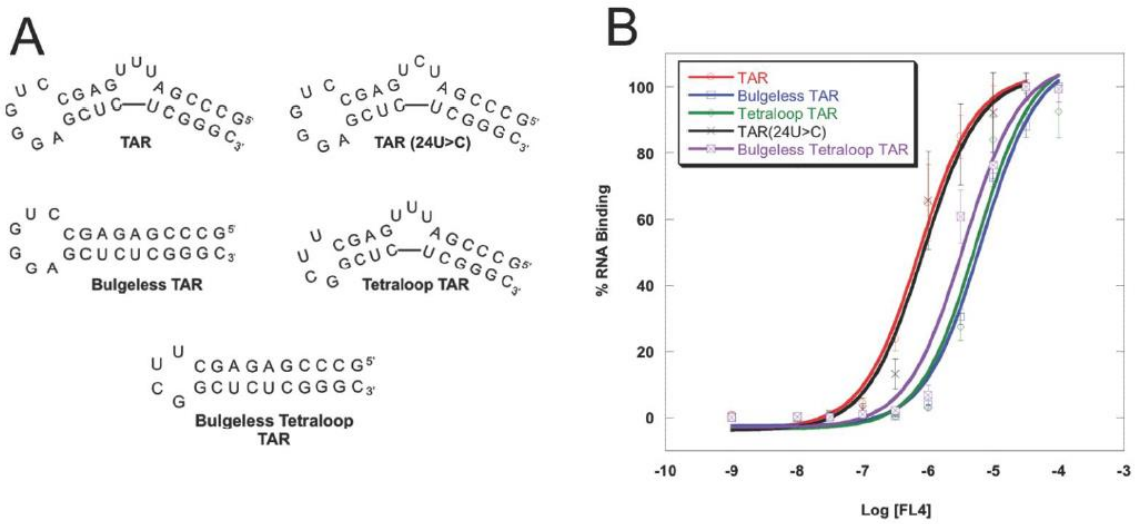


Figure 2.4 (A) Sequence and secondary structure of TAR RNA and variants; (B) Titration curves of **FL4** with these RNAs.

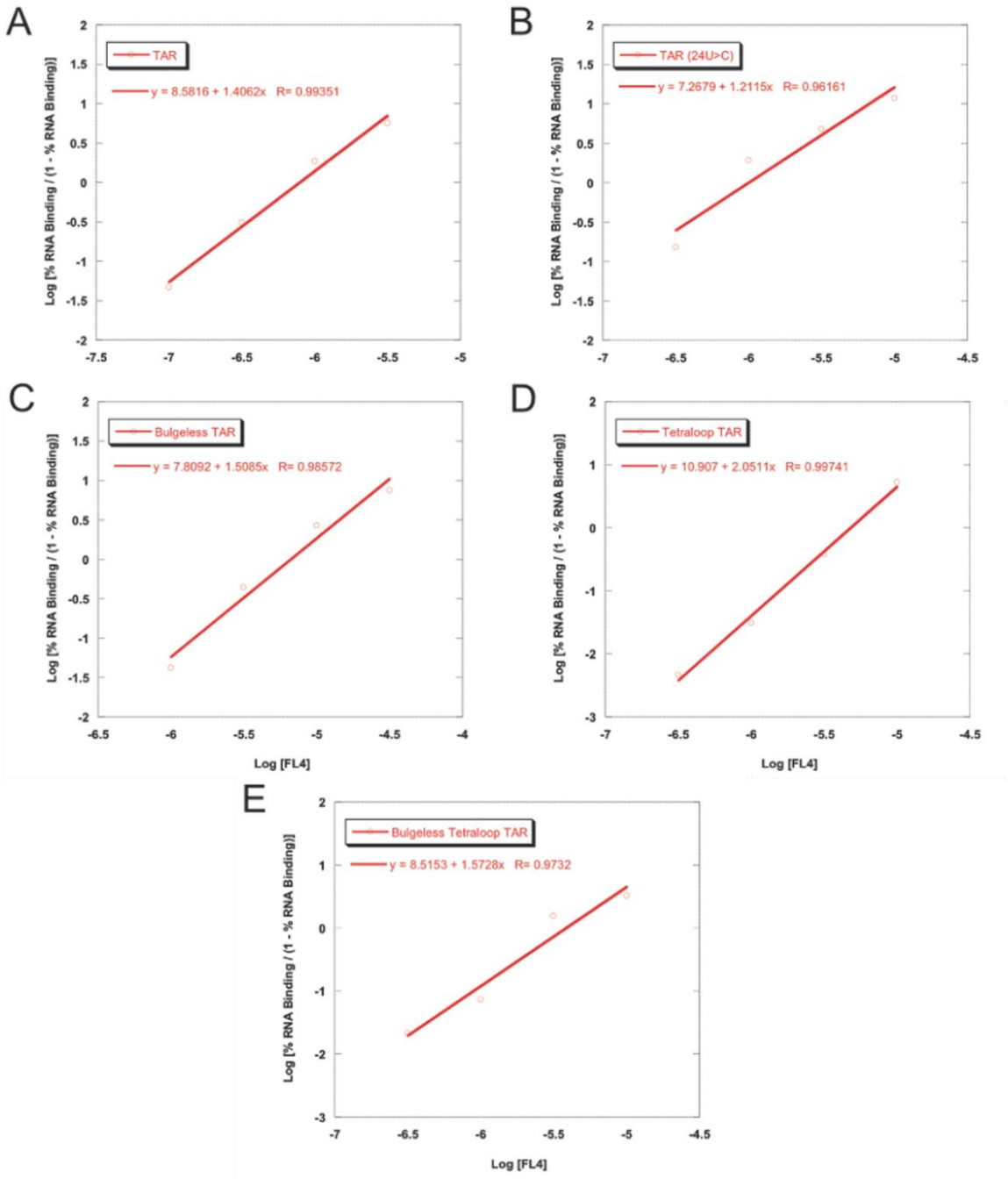


Figure 2.5 Hill plots of **FL4** dot blot data with native TAR RNA and TAR RNA variants. (A) TAR, (B) TAR(24U>C), (C) Bulgeless TAR, (D) Tetraloop TAR, and (E) Bulgeless Tetraloop TAR.

because it suggests that **FL4** is selective for the three dimensional structure of native TAR RNA and that **FL4** may interact with both of these structural elements when bound to the native TAR RNA. Therefore, we anticipated that modifying both structural elements

simultaneously (bulgeless tetraloop TAR) would further decrease the binding affinity if **FL4** indeed spans the bulge and apical loop of TAR RNA. However, implementing both modifications resulted in a K_d value of 3.5 ± 1.0 μM, which was within error of tetraloop TAR. It is possible that the additional decrease in affinity is not observed with bulgeless tetraloop TAR because the tertiary structure of these RNAs is sufficiently altered to preclude **FL4** from binding specifically. Hill analyses of our dot blot data for native TAR RNA and TAR (24U>C) yielded Hill coefficients (n) of 1.4 and 1.2, respectively (Figure 2.5), suggesting non-cooperative binding of **FL4** when n is near 1. This supports our hypothesis of a single binding site that spans the bulge and apical loop, as the binding affinity is clearly decreased when these structural elements are individually removed. Furthermore, cooperative binding ($n \geq 1.5$) is observed for bulgeless TAR, tetraloop TAR, and bulgeless tetraloop TAR, which presented Hill coefficients of 1.5, 2.1, and 1.6 respectively. Taken together, these results suggest that modifications on TAR result in multiple, low-affinity binding sites for **FL4**.

2.2.5 Branched Peptides are Cell Permeable and Exhibit no Cytotoxicity

An attractive feature for developing BPs as RNA ligands is the ability to control the molecular weight, a property that can have significant influence on cellular uptake. Although Arg-rich BPs (Arg > 8) have been shown to be cell permeable,[48] the spacing and number of Arg residues present in our BP hits are significantly different than previously reported examples. We biased our investigation to 3.3.3-BP in part because of their medium molecular weight, which is around 1400 Da. Based on the hit sequences revealed through HTS, we predicted that our BPs would be cell permeable due to the high content of basic residues and because they were less than 20-amino acids in size.[49-52]

Cellular uptake in HeLa cells was determined upon incubation with FITC-labeled BPs (1 μM) in the culture medium for 4 hr at 37 ℃. After washing, cells were fixed with 4% paraformaldehyde and imaged by confocal microscope. Cells incubated with **FL3** contained fluorescence that was diffuse throughout the nucleus and cytoplasm (Figure 2.6A). Punctate structures were also observed in cells incubated with other BPs, supporting evidence that BPs were cell permeable (Figure 2.6B). Further, the HeLa cell line was viable in an MTT assay, suggesting that BPs that penetrate cells were nontoxic (Figure 2.6C).

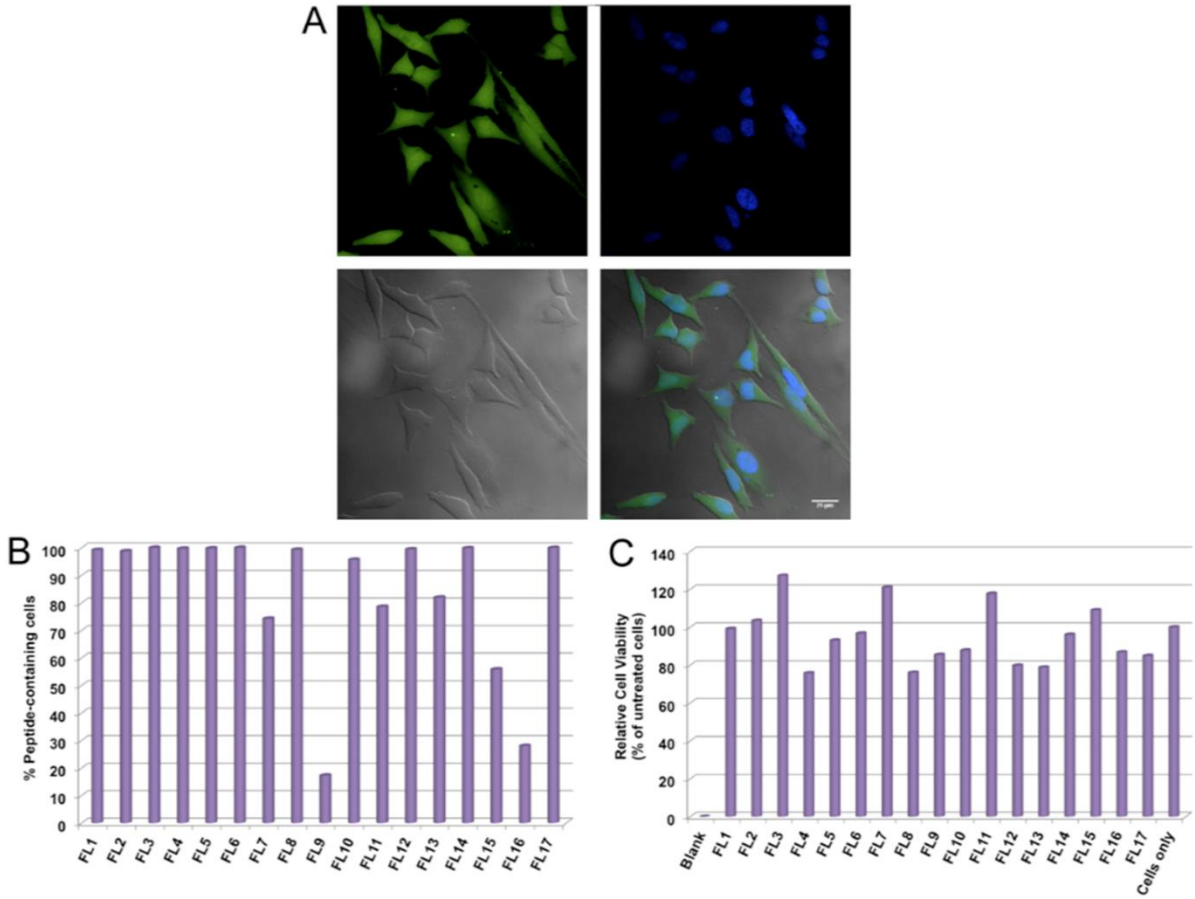


Figure 2.6 (A) Cellular uptake of branched peptides into HeLa cells, top left: fluorescence image of cells; top right: DAPI staining of the nucleus; bottom left: phase contrast image; bottom right: overlay of all images. White scale bar is 25μm. (B) Cell permeability of FITC-labeled branched peptides by flow cytometry in HeLa cells. (C) MTT cell toxicity assay.

2.3 Conclusion

In this study, we generated a general platform for development of selective RNA-binding ligands based on multivalent branched peptides. A BP library was designed to bind selectively to the tertiary structure of HIV-1 TAR RNA. High throughput screening and bio-physical characterization of hit compounds resulted in peptides with binding affinities in the submicromolar range. Characterization of **FL4** demonstrated that 'branching' in peptides provides multivalent interactions, which increase binding affinities for RNA. Mutation experiments indicated that **FL4** interacts with both the loop region and the bulge region. Furthermore, the branched peptides are cell permeable and nontoxic. The results from this study highlight the therapeutic potential of these RNA binding compounds, as well as the utility of this method for RNA ligand discovery. Current efforts are aimed to further improve the binding affinity and selectivity of BPs and demonstrate their therapeutic potential in cell-based assays. Taken together, the results from this study will inform our ongoing efforts to target highly structured RNAs with high affinity and selectivity.

2.4 Materials and Methods

2.4.1 Synthesis of Branched Peptide Library

We used standard solid phase peptide synthesis techniques to generate the 3.3.3 library via the split and pool method. *N*-α-Fmoc protected L-amino acids (Novabiochem), HCTU (Novabiochem) and *N*,*N*-Diisopropylethylamine (DIEA, Aldrich) were used in coupling reactions. Three copies of library were prepared simultaneously by using a three-fold excess of Tentagel Macrobead-NH_2 resin (0.37 mmol/g) (Peptides International). The resin was swollen in DCM (20 mL, 2 x 15 min) followed by DMF (20 mL, 15 min). Three

equivalents photocleavable linker Fmoc-ANP-OH were first coupled to the resin in DMF for 3 hr in the presence of three equivalent HCTU and five equivalent DIEA. After coupling, the resin was washed with DMF (20 mL, 1 min), DCM (20 mL, 1 min) and DMF (20 mL, 1 min). The same washing procedure was applied after every step. Then, 20% piperidine in DMF (20 mL, 2 x 10 min) was used for Fmoc deprotection. A Kaiser test was used after each coupling and deprotection step to confirm reaction completion. N-Fmoc amino acids (3 equiv), HCTU (3 equiv), and DIEA (5 equiv) were added to each reaction vessels in DMF and coupled for 30 min. Fmoc-Lys(Fmoc)-OH was used as a branching unit, and molar equivalencies of reagents were doubled in coupling reactions after installation of the branching unit. Finally, the resin was treated with 95:2.5:2.5 TFA (trifluoroacetic acid, Acros)/H_2O/TIS (triisopropylsilane, Acros) (v/v/v) for 3 hr. After deprotection, the resin was washed extensively with DMF, DCM, and MeOH before drying and storing at −20 °C.

2.4.2 On-bead Screening Assay

DY547 labeled HIV-1 TAR RNA (5′-DY547-GCC-CGA-UUU-GAG-CCU-GGG-AGC-UCU-CGG-GC-3′) was purchased from Dharmacon and prepared according to the manufacturer's protocol. We performed a preliminary incubation of a small subset of beads (~2000) in BSA (1mg/ml) and 500 nM unlabeled α-synuclein mRNA at 4 ℃ for 3 hr and washed (3x) the beads to remove unbound BSA and RNA in solution. Finally, TAR RNA screening was effected by DY547-labeled TAR RNA at 3.4 nM for 3 hr at 4 ℃. The solution was then filtered, washed successively with buffer, placed in a 96-well plate and analyzed by using confocal microscopy, 6 beads were selected (Figure 2.7**)**. It was evident from the initial experiment that the stringency of the assay was inadequate because out of ~2000 beads, six beads were sufficiently fluorescent to isolate them (0.3% hit rate).

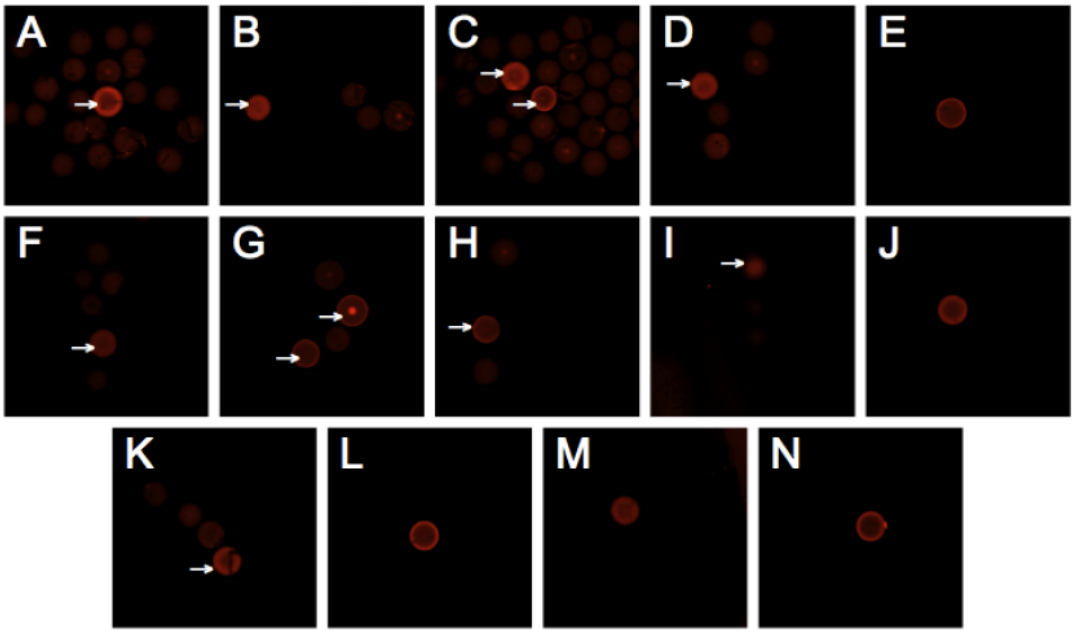


Figure 2.7 The 16 hits screened from the 3.3.3 library. A–E: Initial 6 hits. F–N: Additional 10 hits found with more stringent screening conditions.

The results of the preliminary incubation studies suggested that a more stringent protocol is necessary to decrease the number of hit compounds that will have to be analyzed. To achieve this goal, the second incubation time was reduced to 1 hr. Following this protocol for the remaining 10,000 library members, 10 additional beads were selected (Figure 2.7) (0.1% hit rate). These beads were washed with buffer and organic solvents to remove the bound RNA until fluorescence was no longer visible under the confocal microscope. Finally, labeled RNA was removed with copious buffer and organic solvent washes, which was followed by photolytic release of branched peptides from the resin by exposure to UV irradiation (365 nm, handheld UV lamp, 4 W) for 1 hr.

2.4.3 Synthesis of Fluorescein-labeled and Non-labeled Branched Peptides

The BPs were synthesized following standard solid-phase peptide synthesis techniques using Rink Amide MBHA resin. Each BP was prepared such that a single N-terminus was labeled with fluorescein using fluorescein isothiocyanate (FITC) (Figure 2.8).

The branching unit was attached to the ε-nitrogen of Lys that was orthogonally protected with ivDde; therefore, although the two branches were similar, the spacing was different. Over the course of this study, we found that directly coupling the N-terminal amino acid to FITC resulted in poor isolated yields after HPLC purification. Upon further analysis of the crude peptides, we observed a strong signal in the MALDI-TOF mass spectrum corresponding to truncated peptide, where fluorescein and the adjacent amino acid were cleaved. This result suggested that acid mediated formation of fluorescein thiohydantoins was occurring upon cleavage from solid-support resin with TFA.[53] Installation of aminohexanoic acid (Ahx) as spacer between the N-terminal amino acid and fluorescein resolved the problematic autocleavage and resulted in increased overall yield. Further, one particular peptide, **FL15**, epimerized resulting in two diastereomers of equal intensity that were separated by HPLC and confirmed by LC-ESI/MS. **FL15** was finally prepared as the single diastereomer by using [ethyl cyano(hydroxyimino)acetato-O^2]tri-1-pyrrolidinylphosphonium hexafluorophosphate (PyOxim) as coupling reagent.

Unlabeled and FITC-labeled peptides were synthesized on Rink amide MBHA resin (100-200 mesh) (Novabiochem). For FITC-labeled peptides, Fmoc-Lys(ivDde)-OH was used as branching units. Acetic anhydride in DMF (1:1 v/v) with 10 equivalents of DIEA was used to cap the first N-terminus. Then, ivDde was removed by treatment with 2% hydrazine in DMF for 1 hr, and the second N-terminus was synthesized through the ε-N of the Lys side chain. Fmoc-6-Ahx-OH (AnaSpec) was coupled to the N-terminal amino acidfollowed by reacting with FITC (5 equiv) and DIEA DIEA (14 equiv) for 6 hr. All subsequent steps were protected from light. The supernatant was dried under reduced pressure, and the crude peptide was triturated from cold diethyl ether. The peptides were

purified by using a Jupiter 4 µm Proteo 90 Å semiprep column (Phenomenex) with a solvent gradient composed of 0.1% TFA in Milli-Q water and HPLC grade acetonitrile. Peptide purity was determined by using a Jupiter 4 µm Proteo 90 Å analytical column (Phenomenex), and peptide identity was confirmed by MALDI-TOF analysis. Unlabeled peptide concentrations were measured in nuclease free water at 280 nm using their calculated extinction coefficients. FITC-labeled peptide were monitored at 495 nm using the extinction coefficient of FITC at 77,000 mol^{-1} cm^{-1} in 100 mM glycine, pH 9.0.

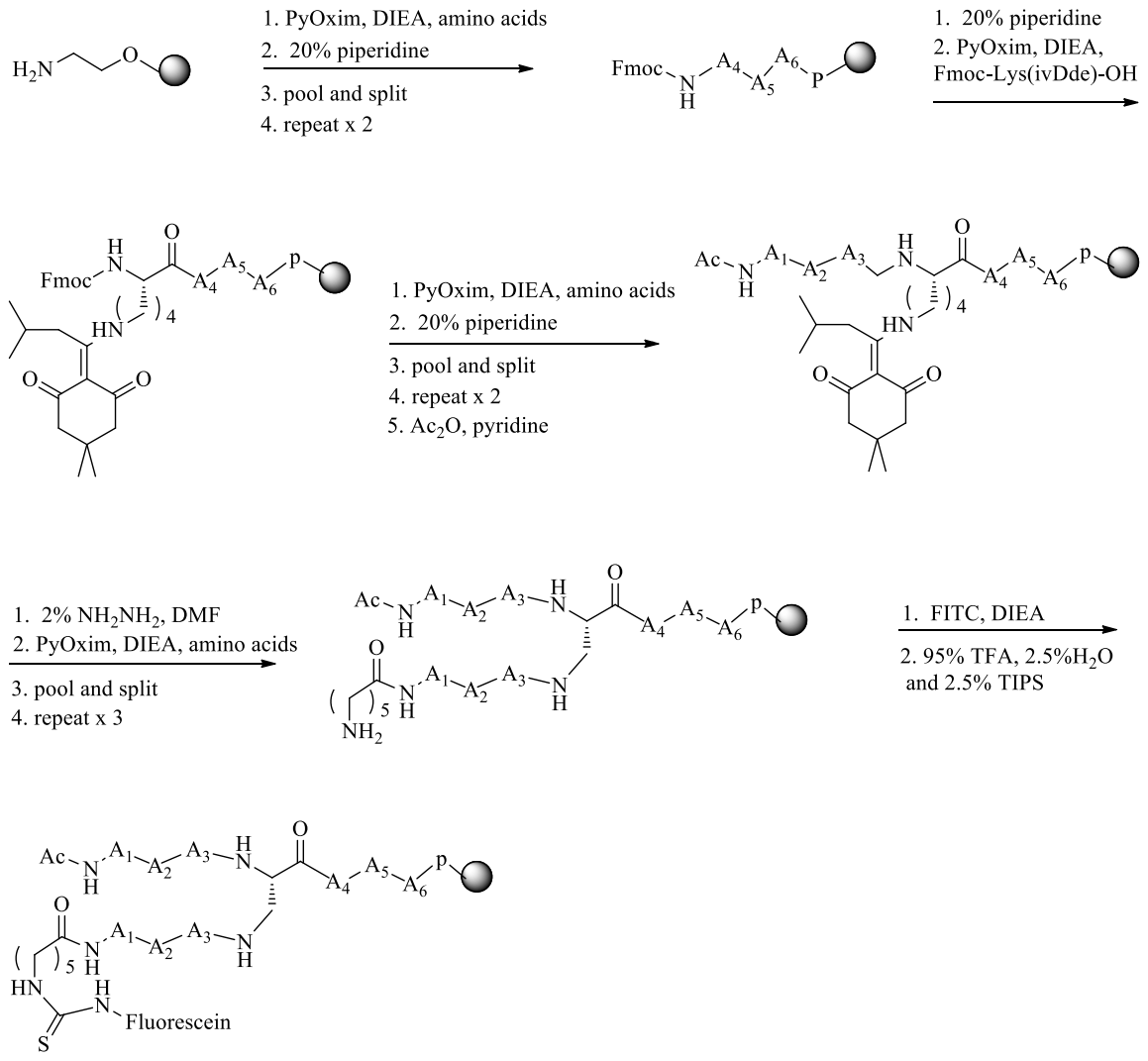


Figure 2.8 Synthesis of fluorescein-labeled branched peptides by solid-phase peptide synthesis (SPPS).

2.4.4 Preparation of ^{32}P-Labeled HIV-1 TAR RNA

HIV-1 TAR RNA was prepared by *in vitro* transcription with the Ribomax T7 Express System (Promega) using previously reported techniques.[54] All steps were done by using RNase free conditions. HIV-1 TAR, 5'-GCCCGAGAGCTCCCAGGCTCAAATC GGGCCTATAGTGAGTCGTATTACAT; TAR(24U>C), 5'-GCCCGAGAGCTCCCA GGCTCATATCGGGCCTATAGTGAGTCGTATTACAT; bulgeless TAR, 5'-GCCCGA GAGCTCCCAGGCTCTCGGGCCTATAGTGAGTCGTATTACAT; tetraloop TAR 5'-GCCCGAGAGCCGAAGCTCAAATCGGGCCTATAGTGAGTCGTATTACAT; bulgeless tetraloop TAR, 5'-GCCCGAGAGCCGAAGCTCTCGGGCCTATAGT GAGTCGTATTACAT and complementary strands, 5'-ATGTAATACGACTCACTA TAGG (Integrated DNA Technologies), were annealed prior to transcription by heating an equimolar mixture of the ssDNA strands in water at 65 ^{o}C for 2 minutes followed by a 2 minute incubation in an ice bath. After transcription, the DNA was degraded by DNase. The newly transcribed HIV-1 TAR RNA was purified by 20% denaturing PAGE. The topmost band was excised and eluted from the gel overnight in 300 mM sodium acetate, 10 mM Tris•HCl, pH 7.4, and 10 mM EDTA. The sample was desalted by using a NAP-25 column (GE Healthcare) and lyophilized to dryness before treating with calf intestinal phosphatase in NEBuffer 3 (New England BioLabs). The dephosphorylated RNA was recovered by phenol extraction followed by ethanol precipitation. Preparation of the full-length transcript was confirmed by MALDI-TOF. Dephosphorylated RNA was stored as a pellet at −80 ^{o}C.

HIV-1 TAR RNA was labeled at the 5'-end by treating 100 pmoles of dephosphorylated RNA with 20 nmoles of [γ-^{32}P]ATP (111 TBq mol^{-1}) and 20 units of T4

polynucleotide kinase in 70 mM Tris•HCl, pH 7.6, 10 mM $MgCl_2$, and 5 mM dithiothreitol. The mixture was incubated at 37 ºC for 1 hr followed by a 10 minute incubation at 65 ºC. The labeled RNA was recovered by ethanol precipitation. A 20% denaturing PAGE run at 25W for 2.25 hr followed by autoradiography was used to ensure that the RNA sample was ≥95% pure prior to use in EMSA and dot blot assays. The ^{32}P-labeled RNA was stored at a concentration of 500 nM in water at -20 ºC for up to two months without measurable degradation.

2.4.5 Dot Blot Assay

Dot blot assays were performed in triplicate using a Whatman Minifold I 96 well Dot Blot system and Whatman 0.45 μm pore size Protran nitrocellulose membranes. Assays were carried out by first refolding 800 pM ^{32}P-labeled HIV-1 TAR RNA in 2X TK buffer (100 mM Tris•HCl, pH 7.4, and 200 mM KCl) by heating the sample at 95 ºC for three minutes and then allowing it to cool slowly at room temperature for 20 minutes. Peptides were diluted with water or 10% DMSO from 200 μM stocks to prepare each series at 2X concentrations ranging from 0.002 μM to 200 mM at half-log intervals. 25 μL aliquots of refolded ^{32}P-labeled HIV-1 TAR RNA in 2X TK buffer were added to 25 μL aliquots of the peptides at 2X concentration to give the desired final concentration of each component. The mixtures incubated at room temperature for 20 minutes prior to filtering each 50 μL sample through the nitrocellulose membrane, which had been pre-equilibrated in 1X TK buffer. Two consecutive 50 μL washes with 1X TK buffer followed each filtration. Peptide binding was measured by autoradiography by using a storage phosphor screen (GE Healthcare), which was imaged on a Typhoon Trio (GE Health Care). Densitometry measurements were taken in ImageQuant TL (Amersham Biosciences).

Binding curves were generated by using Kaleidagraph (Synergy Software). Error bars represent the standard deviation calculated for three replicates.

2.4.6 EMSA

EMSA assays were performed in duplicate by first refolding 4 nM ^{32}P-labeled HIV-1 TAR RNA in 2X TK buffer (100 mM Tris•HCl, pH 7.4, and 200 mM KCl) by using the previously described method (*vide supra*). Peptides were diluted with water from 200 μM stocks to prepare each series at 2X concentration. 10 μL aliquots of refolded ^{32}P-labeled HIV-1 TAR RNA in 2X TK buffer were added to 10 μL aliquots of the peptides at 2X concentration to give the desired final concentration of each component. The mixtures incubated at room temperature for 20 minutes followed by the addition of 2 μL of 30% glycerol for loading. 15 μL of each sample was loaded on to a 10% native PAGE, which had pre-run for 1 hr at 300 V. The samples electrophoresed for 1 hr at 300 V. Gels were dried to filter paper prior to autoradiography. Data were measured as the percentage of bound RNA in each lane and error bars represent the standard deviation calculated for two replicates. Selectivity studies were performed in the presence of 10-fold excess by mass of tRNA from E.Coli MRE 600 (Roche). The tRNA was refolded separately from HIV-1 TAR RNA prior to use.

2.4.7 Cellular Internalization of Peptides

HeLa cells were plated at 1 x 10^5 cells/well in DMEM containing 10% FBS in 12-well tissue culture plates (Corning) and allowed to attach at 37 ℃ in a humid 5% CO_2 atmosphere for 24 hr. After removing media and washing cells with PBS, 600 μl of FITC-labeled peptide in Opti-MEM (1 μM) was added to each well. Cells were incubated with peptides for 4 hr, then 1.5 ml DMEM was added and incubated for 30 minutes. Cells were

detached with 500 μL of trypsin-EDTA, quenched with internalization media (1 mL), and the contents of each well were collected into Falcon Tubes (BD Biosciences). Cells were centrifuged at 4 ℃ and 1250 rpm for 10 minutes. The supernatant was removed and the cell pellets rinsed with PBS, and centrifuged again at identical conditions. Supernatant was removed and cell pellet was again suspended in 2% FBS in PBS. Propidium iodide (PI; 5 μg ml^{-1}, Molecular probes) was added to each tube 2–5 minutes prior to analysis. Cellular uptake of FITC-labeled peptides was measured on a FACS Canto II flow cytometer (BD Biosciences). FITC was excited by using a 488 nm solid state laser, and detected at 530 ± 30 nm bandpass filter, and PI was excited by using a 488 nm solid state laser and detected with a 670 nm longpass filter. Appropriate gating was done against the untreated cells control to ensure that autofluorescence was not measured as cellular uptake, and only live (PI-negative) cells were included in subsequent analysis. 10,000-20,000 gated events were collected for each sample, and experiments were done in duplicate unless noted otherwise.

2.4.8 MTT Toxicity Assay

HeLa cells were plated at 10,000 cells/well in 10% FBS-containing DMEM in 48-well tissue culture plates and incubated for 24 hr at 37 ℃ in a humid 5% CO_2 atmosphere. Culture medium was removed and cells rinsed with PBS. Peptides (175 μL in Opti-MEM, 1 μM) were added to each well and incubated for 4 hr. Peptides were removed and cells were rinsed with PBS. DMEM (300 μL), containing 10% FBS and 0.5 mg ml^{-1} MTT (Invitrogen) was added to each well. Cells were incubated for one additional hr before removing MTT-containing media, rinsing with PBS, and dissolving cells and formazan product in 300 μl DMSO. Absorbance was measured at 570 nm for each sample, and toxicity was normalized against DMSO only and untreated cells.

2.5 References

1. Doudna, J. A.; Cech, T. R., The chemical repertoire of natural ribozymes. *Nature* **2002,** *418* (6894), 222-228.

2. Jovanovic, M.; Hengartner, M. O., miRNAs and apoptosis: RNAs to die for. *Oncogene* **2006,** *25* (46), 6176-6187.

3. Draper, D. E., Protein-RNA recognition. *Annu. Rev. Biochem.* **1995,** *64*, 593-620.

4. Thomas, J. R.; Hergenrother, P. J., Targeting RNA with small molecules. *Chem. Rev.* **2008,** *108* (4), 1171-224.

5. Lu, J.; Kadakkuzha, B. M.; Zhao, L.; Fan, M.; Qi, X.; Xia, T., Dynamic ensemble view of the conformational landscape of HIV-1 TAR RNA and allosteric recognition. *Biochemistry* **2011,** *50* (22), 5042-57.

6. Zaman, G. J. R.; Michiels, P. J. A.; van Boeckel, C. A. A., Targeting RNA: new opportunities to address drugless targets. *Drug Discov. Today* **2003,** *8* (7), 297-306.

7. Carter, A. P.; Clemons, W. M.; Brodersen, D. E.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V., Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* **2000,** *407* (6802), 340-348.

8. Tor, Y., The ribosomal A-site as an inspiration for the design of RNA binders. *Biochemistry* **2006,** *88* (8), 1045-1051.

9. Brodersen, D. E.; Clemons Jr, W. M.; Carter, A. P.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V., The Structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. *Cell* **2000,** *103* (7), 1143-1154.

10. Wilson, D. N.; Harms, J. M.; Nierhaus, K. H.; Schlünzen, F.; Fucini, P., Species-specific antibiotic-ribosome interactions: implications for drug development. *J. Biol. Chem.* **2005,** *386* (12), 1239-1252.

11. Hermann, T., Drugs targeting the ribosome. *Curr. Opin. Struct. Biol.* **2005,** *15* (3), 355-366.

12. Leach, K. L.; Swaney, S. M.; Colca, J. R.; McDonald, W. G.; Blinn, J. R.; Thomasco, Lisa, M.; Gadwood, R. C.; Shinabarger, D.; Xiong, L.; Mankin, A. S., The site of action of oxazolidinone antibiotics in living bacteria and in human mitochondria. *Mol. Cell.* **2007,** *26* (3), 393-402.

13. Davidson, B. L.; McCray, P. B., Current prospects for RNA interference-based therapies. *Nat. Rev. Genet.* **2011,** *12* (5), 329-340.

14. Watts, J. K.; Corey, D. R., Clinical status of duplex RNA. *Bioorg. Med. Chem. Lett.* **2010,** *20* (11), 3203-3207.

15. Lee, M. M.; Pushechnikov, A.; Disney, M. D., Rational and modular design of potent ligands targeting the RNA that causes myotonic dystrophy 2. *ACS Chem. Biol.* **2009,** *4* (5), 345-355.

16. Stevens, M.; De Clercq, E.; Balzarini, J., The regulation of HIV-1 transcription: molecular targets for chemotherapeutic intervention. *Med. Res. Rev.* **2006,** *26* (5), 595-625.

17. Jakobovits, A.; Smith, D. H.; Jakobovits, E. B.; Capon, D. J., A discrete element 3' of human immunodeficiency virus 1 (HIV-1) and HIV-2 mRNA initiation sites mediates transcriptional activation by an HIV trans activator. *Mol. Cell. Biol.* **1988,** *8* (6), 2555-2561.

18. Rana, T. M.; Jeang, K.-T., Biochemical and functional interactions between HIV-1 Tat protein and TAR RNA. *Arch. Biochem. Biophys.* **1999,** *365* (2), 175-185.

19. Bannwarth, S.; Gatignol, A., HIV-1 TAR RNA: The target of molecular interactions between the virus and its host. *Cur. HIV Res.* **2005,** *3* (1), 61-71.

20. Berkhout, B.; Silverman, R. H.; Jeang, K.-T., Tat trans-activates the human immunodeficiency virus through a nascent RNA target. *Cell* **1989,** *59* (2), 273-282.

21. Keen, N. J.; Gait, M. J.; Karn, J., Human immunodeficiency virus type-1 Tat is an integral component of the activated transcription-elongation complex. *Proc. Natl. Acad. Sci. U. S. A.* **1996,** *93* (6), 2505-10.

22. Puglisi, J. D.; Tan, R.; Calnan, B. J.; Frankel, A. D.; Williamson, J. R., Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science* **1992,** *257* (5066), 76-80.

23. Delling, U.; Roy, S.; Sumner-Smith, M.; Barnett, R.; Reid, L.; Rosen, C. A.; Sonenberg, N., The number of positively charged amino acids in the basic domain of Tat is critical for trans-activation and complex formation with TAR RNA. *Proc. Natl. Acad. Sci. U. S. A.* **1991,** *88* (14), 6234-6238.

24. Dayton, A. I.; Sodroski, J. G.; Rosen, C. A.; Goh, W. C.; Haseltine, W. A., The trans-activator gene of the human T cell lymphotropic virus type III is required for replication. *Cell* **1986,** *44* (6), 941-7.

25. Fernandes, J.; Jayaraman, B.; Frankel, A., The HIV-1 Rev response element: an RNA scaffold that directs the cooperative assembly of a homo-oligomeric ribonucleoprotein complex. *RNA biology* **2012,** *9* (1), 6-11.

26. Huber, P. W.; Cui, M.; Czarnik, A. W.; Mei, H.-Y., Binding of neomycin to the TAR element of HIV-1 RNA induces dissociation of Tat protein by an allosteric mechanism. *Biochemistry* **1998,** *37* (16), 5549-5557.

27. Litovchick, A.; Evdokimov, A. G.; Lapidot, A., Arginine-aminoglycoside conjugates that bind to HIV transactivation responsive element RNA in vitro. *FEBS Lett.* **1999,** *445* (1), 73-79.

28. Evdokimov, A. G.; Lapidot, A., Aminoglycoside arginine conjugates that bind TAR RNA: Synthesis, characterization, and antiviral activity. *Biochemistry* **2000,** *39* (11), 2838-2852.

29. Tao, J.; Frankel, A. D., Specific binding of arginine to TAR RNA. *Proc. Natl. Acad. Sci. U. S. A.* **1992,** *89* (7), 2723-6.

30. Pang, R.; Zhang, C.; Yuan, D.; Yang, M., Design and SAR of new substituted purines bearing aryl groups at N9 position as HIV-1 Tat–TAR interaction inhibitors. *Biorg. Med. Chem.* **2008,** *16* (17), 8178-8186.

31. Hwang, S.; Tamilarasu, N.; Ryan, K.; Huq, I.; Richter, S.; Still, W. C.; Rana, T. M., Inhibition of gene expression in human cells through small molecule-RNA interactions. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** *96* (23), 12997-3002.

32. Mei, H. Y.; Mack, D. P.; Galan, A. A.; Halim, N. S.; Heldsinger, A.; Loo, J. A.; Moreland, D. W.; Sannes-Lowery, K. A.; Sharmeen, L.; Truong, H. N.; Czarnik, A. W., Discovery of selective, small-molecule inhibitors of RNA complexes--I. The Tat protein/TAR RNA complexes required for HIV-1 transcription. *Bioorg. Med. Chem.* **1997,** *5* (6), 1173-84.

33. Davis, B.; Afshar, M.; Varani, G.; Murchie, A. I.; Karn, J.; Lentzen, G.; Drysdale, M.; Bower, J.; Potter, A. J.; Starkey, I. D.; Swarbrick, T.; Aboul-ela, F., Rational design of inhibitors of HIV-1 TAR RNA through the stabilisation of electrostatic "hot spots". *J. Mol. Biol.* **2004,** *336* (2), 343-56.

34. Murchie, A. I.; Davis, B.; Isel, C.; Afshar, M.; Drysdale, M. J.; Bower, J.; Potter, A. J.; Starkey, I. D.; Swarbrick, T. M.; Mirza, S.; Prescott, C. D.; Vaglio, P.; Aboul-ela, F.; Karn, J., Structure-based drug design targeting an inactive RNA conformation: exploiting the flexibility of HIV-1 TAR RNA. *J. Mol. Biol.* **2004,** *336* (3), 625-38.

35. Bohjanen, P. R.; Colvin, R. A.; Puttaraju, M.; Been, M. D.; GarciaBlanco, M. A., A small circular TAR RNA decoy specifically inhibits Tat-activated HIV-1 transcription. *Nucleic Acids Res* **1996,** *24* (19), 3733-3738.

36. Michienzi, A.; Li, S.; Zaia, J. A.; Rossi, J. J., A nucleolar TAR decoy inhibitor of HIV-1 replication. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99* (22), 14047-52.

37. Wang, D.; Iera, J.; Baker, H.; Hogan, P.; Ptak, R.; Yang, L.; Hartman, T.; Buckheit, R. W., Jr.; Desjardins, A.; Yang, A.; Legault, P.; Yedavalli, V.; Jeang, K. T.; Appella, D. H., Multivalent binding oligomers inhibit HIV Tat–TAR interaction critical for viral replication. *Bioorg. Med. Chem. Lett.* **2009,** *19* (24), 6893-7.

38. Davidson, A.; Leeper, T. C.; Athanassiou, Z.; Patora-Komisarska, K.; Karn, J.; Robinson, J. A.; Varani, G., Simultaneous recognition of HIV-1 TAR RNA bulge and loop sequences by cyclic peptide mimics of Tat protein. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106* (29), 11931-6.

39. Athanassiou, Z.; Patora, K.; Dias, R. L.; Moehle, K.; Robinson, J. A.; Varani, G., Structure-guided peptidomimetic design leads to nanomolar beta-hairpin inhibitors of the Tat–TAR interaction of bovine immunodeficiency virus. *Biochemistry* **2007,** *46* (3), 741-51.

40. Lalonde, M. S.; Lobritz, M. A.; Ratcliff, A.; Chamanian, M.; Athanassiou, Z.; Tyagi, M.; Wong, J.; Robinson, J. A.; Karn, J.; Varani, G.; Arts, E. J., Inhibition of both HIV-1 reverse transcription and gene expression by a cyclic peptide that binds the Tat-transactivating response element (TAR) RNA. *PLoS pathog.* **2011,** *7* (5), e1002038.

41. Bryson, D. I.; Zhang, W.; Ray, W. K.; Santos, W. L., Screening of a branched peptide library with HIV-1 TAR RNA. *Mol. BioSyst.* **2009,** *5*, 1070-1073.

42. Brown, B. B.; Wagner, D. S.; Geysen, H. M., A single-bead decode strategy using electrospray ionization mass spectrometry and a new photolabile linker: 3-amino-3-(2-nitrophenyl)propionic acid. *Mol. Divers.* **1995,** *1* (1), 4-12.

43. Churcher, M. J.; Lamont, C.; Hamy, F.; Dingwall, C.; Green, S. M.; Lowe, A. D.; Butler, J. G.; Gait, M. J.; Karn, J., High affinity binding of TAR RNA by the human immunodeficiency virus type-1 tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region. *J. Mol. Biol.* **1993,** *230* (1), 90-110.

44. B. Crumpton, J.; Zhang, W.; L. Santos, W., Facile analysis and sequencing of linear and branched peptide boronic acids by MALDI mass spectrometry. *Anal. Chem.* **2011,** *83* (9), 3548-3554.

45. Wang, X.; Huq, I.; Rana, T. M., HIV-1 TAR RNA recognition by an unnatural biopolymer. *J. Am. Chem. Soc.* **1997,** *119* (27), 6444-6445.

46. Gelman, M. A.; Richter, S.; Cao, H.; Umezawa, N.; Gellman, S. H.; Rana, T. M., Selective binding of TAR RNA by a Tat-derived beta-peptide. *Org. Lett.* **2003,** *5* (20), 3563-5.

47. Wang, D.; Iera, J.; Baker, H.; Hogan, P.; Ptak, R.; Yang, L.; Hartman, T.; Buckheit Jr, R. W.; Desjardins, A.; Yang, A.; Legault, P.; Yedavalli, V.; Jeang, K.-T.; Appella, D.

H., Multivalent binding oligomers inhibit HIV Tat–TAR interaction critical for viral replication. *Bioorg. Med. Chem. Lett.* **2009,** *19* (24), 6893-6897.

48. Futaki, S.; Nakase, I.; Suzuki, T.; Zhang; Sugiura, Y., Translocation of branched-chain arginine peptides through cell membranes: Flexibility in the spatial disposition of positive charges in membrane-permeable peptides. *Biochemistry* **2002,** *41* (25), 7925-7930.

49. Park, S. H.; Doh, J.; Park, S. I.; Lim, J. Y.; Kim, S. M.; Youn, J. I.; Jin, H. T.; Seo, S. H.; Song, M. Y.; Sung, S. Y.; Kim, M.; Hwang, S. J.; Choi, J. M.; Lee, S. K.; Lee, H. Y.; Lim, C. L.; Chung, Y. J.; Yang, D.; Kim, H. N.; Lee, Z. H.; Choi, K. Y.; Jeun, S. S.; Sung, Y. C., Branched oligomerization of cell-permeable peptides markedly enhances the transduction efficiency of adenovirus into mesenchymal stem cells. *Gene Ther.* **2010,** *17* (8), 1052-61.

50. Frankel, A. D.; Pabo, C. O., Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* **1988,** *55* (6), 1189-1193.

51. Green, M.; Loewenstein, P. M., Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein. *Cell* **1988,** *55* (6), 1179-1188.

52. Torchilin, V. P., Tat peptide-mediated intracellular delivery of pharmaceutical nanocarriers. *Adv. Drug Deliv. Rev.* **2008,** *60* (4-5), 548-558.

53. Jullian, M.; Hernandez, A.; Maurras, A.; Puget, K.; Amblard, M.; Martinez, J.; Subra, G., N-terminus FITC labeling of peptides on solid support: the truth behind the spacer. *Tetrahedron Lett.* **2009,** *50* (3), 260-263.

54. Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C., Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* **1987,** *15* (21), 8783-8798.

Chapter 3. Branched Peptide Boronic Acids (BPBAs): A Novel Mode of Binding Towards RNA

Attribution

This chapter was taken from Zhang, W.; Bryson, D. I.; Crumpton, J. B.; Wynn, J.; Santos, W. L.; "Branched Peptide Boronic Acids (BPBAs): A Novel Mode of Binding Towards RNA". *Chem. Commun.* **2013**, 49. 2436. The author of this dissertation performed a major portion of the work described in this chapter. She designed and synthesized the peptide library, performed the majority of the high throughput screening, synthesized the majority of the hit peptides for biophysical characterization, performed a significant amount of dot blot and electrophoretic mobility shift assays, and synthesized the unnatural amino acids containing boron. She was also responsible for all binding data interpretation, RNA synthesis, and hit peptide deconvolution. Design of the peptide library sequence was a collaborative effort between the author of this dissertation, David I. Bryson, and Jason B. Crumpton of the Santos Group. David I. Bryson prepared and designed the majority of the RNA constructs, and performed dot blot and electrophoretic mobility shift assay of several hit peptides. Jason B. Crumpton performed all MALDI analysis of peptides and aided in the synthesis of the unnatural amino acids containing boron. Jessica Wynn aided in the synthesis of several hit peptides. The author also significantly contributed to the writing and editing of the manuscript. The final manuscript was prepared by Dr. Webster L. Santos. [Reproduced by permission of the Royal Society of Chemistry]

Abstract

We report branched peptide boronic acids (BPBAs) that bind to RRE IIB from an on-bead high-throughput screening of a 3.3.4-library. We demonstrate that boronic acids are tunable moieties that afford a novel binding mode towards RNA.

3.1 Introduction

The human immunodeficiency virus type 1 (HIV-1) is regarded as the prototype member of the lentivirus subfamily of retroviruses, and is the causative agent of the worldwide health problem, acquired immune deficiency syndrome (AIDS). In the past few decades, more than 20 drugs have become available for treatment of HIV-1 infection;[1] at present, favored treatment focuses on combination therapy by using HIV-1 reverse transcriptase (RT), protease, and gp41 inhibitors. Despite the fact that highly active antiretroviral therapy (HAART) has obtained notable successes in reducing plasma viral loads to undetectable levels, HAART fails to completely eliminate the virus from the body due to the remaining chronically HIV-infected $CD4^+$ T cells, which contain the integrated but transcriptionally dormant HIV provirus.[2] In addition, the emergence of drug-resistant viruses have been reported in patients receiving HAART, resulting from mutations in the virally encoded enzymes.[1] In order to keep pace with the rapidly evolving HIV-1, there is a need for development of drugs that target novel viral mechanisms that are genetically well-conserved and less prone to mutation under selective pressure.

The highly structured HIV-1 Rev response element (RRE), a span of ~240-nucleotides located in the *env* gene of all singly spliced and unspliced HIV-1 transcripts, RRE is an example of an extremely well-conserved sequence of RNA across different HIV-1 isolates and plays an essential role in RNA replication by interaction with the Rev

protein.[3] It has been demonstrated that proviral colonies without the rev gene have no replicative abilities, and in the absence of rev protein, the stability of unspliced mRNA is decreased.[4] While the details of the RRE–Rev export pathway are being investigated, some steps have been identified. It is known that of the singly spliced transcripts and multiply spliced transcripts, only multiply spliced transcripts can be exported to the cytoplasm and translated to their corresponding proteins including Rev.[5] Once Rev is expressed, it is imported into the nucleus where it binds cooperatively to RRE.[3] In particular, the stem-loop IIB of RRE (RRE IIB) has been recognized as the high affinity site, where Rev initially binds.[6-7] The resulting RRE–Rev ribonucleoprotein complex binds the host Crm1 and is then shuttled out of the nucleus through the nuclear pore after the larger complex binds to Ran- GTP.[8-9] Since this cooperative binding allows for the export of full-length and singly spliced transcripts, the Rev/RRE export pathway has become a high profile drug target for its critical role in proliferation of HIV-1.[10]

In continuation of our effort toward developing molecules that target the tertiary structure of RNA, we focused our attention on RRE and envisioned utilizing RNA–ligand interactions that are outside the typical canonical mode of binding. We previously demonstrated that branching in peptide ligands provides strong multivalent interactions with another HIV-1 related RNA, the transactivation response element (TAR).[11-12] These branched peptides (BPs) displayed no cytotoxicity, provided excellent cell permeability, and bind to TAR in the submicromolar regime. Herein, we report the discovery and biophysical characterization of branched peptide boronic acids (BPBAs) as medium-sized ligands that bind to the tertiary structure of HIV-1 RRE IIB. Our investigations suggest that the boronic acid moiety plays a pivotal role in increasing binding affinity.

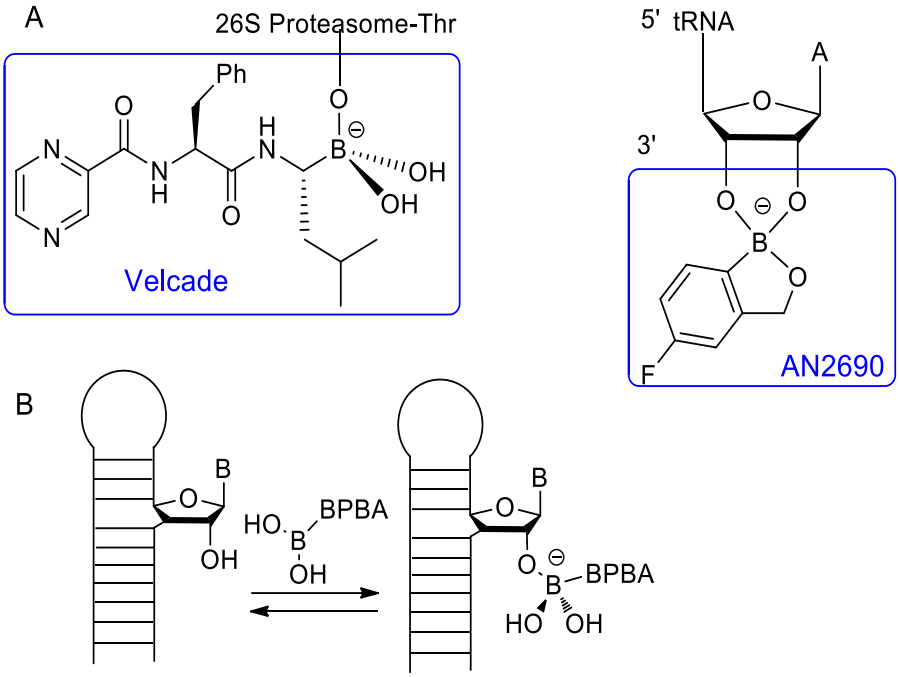


Figure 3.1 (A) Therapeutically relevant boronic acid derivatives that bind to protein or RNA targets. (B) Possible reversible covalent bond formation between Lewis bases in RNA and a boronic acid moiety of branched peptides.

We embarked on a strategy to improve the selectivity and binding affinity to the RNA target through the incorporation of unnatural amino acid side chains featuring the boronic acid functional group. Boronic acids have been used in various biomolecules. For example, boronic acids anchored to a cellulose polymer support was first used to separate and purify RNA.[13] Peptides displaying boronic acid moieties have been demonstrated to form reversible covalent bonds with alizarin and glucose, in addition to being utilized as potent protease inhibitors (Figure 3.1A).[14-15] Furthermore, boron-containing compounds are well tolerated *in vivo* as is evident from the FDA approval of the first boron-containing drug, Bortezomib (Velcade) (Figure 3.1A).[16-17] Another boron-containing small molecule, Tavaborole (AN2690), is currently in phase III clinical trials for treatment of onychomycosis and its mode of action involves trapping the 2′- and 3′-oxygen atoms of the terminal adenosine in leucyl-tRNA synthetase as a boronate adduct (Figure 3.1A).[18] Since

peptidyl boronic acids have not been investigated to target RNA, we hypothesized that we can capitalize on the empty *p*-orbital of boron by forming a reversible covalent bond with the 2′-hydroxyl group of RNA. We reasoned that the empty *p*-orbital of boron is advantageous in ligand binding by acting as an acceptor for the 2′-hydroxyl group and this will not only selective for RNA over DNA, but also boost both affinity and selectivity to the RNA target due to the formation of reversible covalent bonds (Figure 3.1B). Further, the boronic acid moiety is an excellent candidate for probing RNA-ligand binding interactions that are atypical of canonical modes of binding. The boronic acid moiety presents a unique mode of binding that increased the complexity of the library and the incorporation of unnatural side chains would likely impart resistance to enzymatic cleavage.

3.2 Results and Discussion

A BPBA peptide library was synthesized on Tentagel beads by split and pool synthesis. The library was prepared such that there were three variable amino acid positions at both the N- and C-termini (A_1–A_3 and A_4–A_6, respectively), and each variable position was composed of six possible amino acids (Figure 3.2A). Each of the six possible amino acids was chosen for its potential to interact with the RRE IIB target RNA. In position A_1, for example, we selected amino acids with functional groups that can interact with the RNA through hydrophobic interactions (Val), electrostatic attraction (Lys), hydrogen bonding (Ser), π-stacking (Trp), and reversible covalent bonding between boron and a Lewis base presented by the RNA target (K_{BBA}/F_{BPA}) (Figure 3.2B). We chose to incorporate two boron-containing side chains at each variable position in the library to examine whether RRE IIB had a preference in the side chain length or boron Lewis acidity. Hence, the 3.3.4 branched peptide boronic acid library was composed of 46,656 possible amino acid

sequences linked to the bead by a photocleavable linker (3-amino-3-(2-nitrophenyl) propionic acid, ANP). Tyr was included at position A_7 as a convenient spectroscopic handle in quantifying peptide concentrations.

A **3.3.4. Library (6 variable positions)**

A_1	A_2	A_3	A_4	A_5	A_6	A_7
Trp	Tyr	Phe	His	Tyr	Trp	Tyr
Ser	Asn	Thr	Gln	Thr	Ser	Tyr
Val	Pro	Ala	Pro	Leu	Ala	Tyr
Lys	Lys	Lys	Lys	Lys	Lys	Tyr
K_{BBA}	K_{BBA}	K_{BBA}	K_{BBA}	K_{BBA}	K_{BBA}	Tyr
F_{BPA}	F_{BPA}	F_{BPA}	F_{BPA}	F_{BPA}	F_{BPA}	Tyr

6^6 = 46,656 library members

B

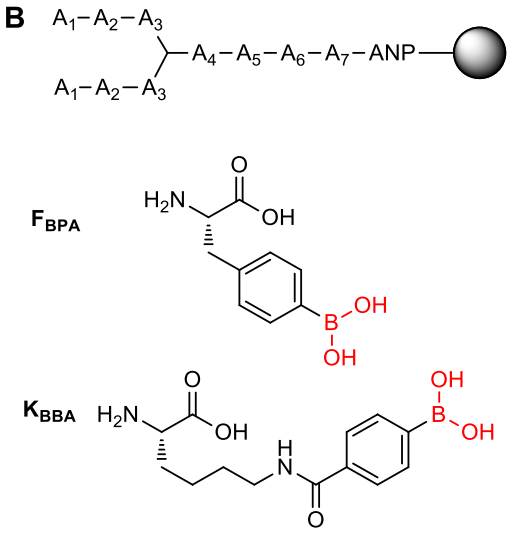


Figure 3.2 (A) 3.3.4-Branched peptide boronic acid library, (B) structure of BPBA library and unnatural amino acids bearing boronic acid groups.

The BPBA library was subjected to on-bead high throughput screening[12] against DY547 labeled HIV-1 RRE IIB RNA. During the screening process, the beads were first pretreated with excess of both bovine serum albumin and competitor tRNA to minimize non-specific binding. Specific binding of the target RNA to peptides resulted in increased fluorescence of the bead, which was monitored by fluorescence microscopy. Eleven beads were selected as possible hits. These peptides were photocleaved via UV irradiation and then sequenced by MALDI MS-MS analysis.[19] Hit compounds were resynthesized for further biophysical characterizations.

The sequence and dissociation constant of the hit BPBAs are shown in Table 3.1 By using standard dot blot assay techniques, ^{32}P-labeled RRE IIB was titrated with increasing concentration of BPBAs.[11] Three peptides showed no binding affinities,

therefore they were identified as false positives. The hits contained varying numbers of boronic acid residues ranging from 0 to 5. In particular, three peptides (**BPBA1, BPBA2 and BPBA3**) had low micromolar binding affinities (1.4, 3.3 and 8.7 μM, respectively) and were investigated further. To confirm that the dissociation constants determined from the dot blot assay were reliable, electrophoretic mobility shift assay (EMSA) was performed with **BPBA1**, **BPBA2** and **BPBA3**. The dissociation constants determined through EMSA (0.3 ± 0.1, 0.6 ± 0.2 and 4.6 ± 2.4 μM, respectively) were comparable to the low micromolar results obtained via dot blot assay. The slight discrepancy with the K_d values may be a result of nonspecific interactions between the peptide and nitrocellulose. Gratifyingly, all three of these sequences contained boronic acid residues with boron containing amino acids present in positions A_1, A_3 and A_4. Interestingly, no boron containing amino acids were found at positions A_5 or A_6 in any hit sequence nor was there any correlation between the number of boronic acid moieties and the resulting binding affinities. Further, it was observed that the longer, more Lewis acidic residue of K_{BBA} was the preferred boron-containing amino acid over the less electron demanding F_{BPA} residue. This result suggests that boron containing side chains can provide non-canonical, alternative, and positive binding interactions with the RNA target.

Further sequence analysis indicated that a Lys residue was preferred at the majority of the positions A_1-A_6. A preference for positively charged side chains, however, is not unexpected because the Rev protein binds RRE IIB through the polyarginine nuclear localization signal (NLS) located in its N-terminal region.[20-21] The positively charged Lys side chains in our hits may function similarly to the Arg residues of the Rev NLS and likely provide necessary electrostatic attraction to the negatively charged phosphate backbone of

RRE IIB. However, it is noteworthy that the number of Lys residues in the hit sequences did not have significant influence in the binding affinity. For example, an increase in the net positive charge did not result in increased binding affinity (compare **BPBA1** to **BPBA7** and **BPBA8**). Taken together, the data suggested that the dissociation constants of the hit peptides for RRE IIB are not entirely due to electrostatic attraction. Additional analysis shows that aliphatic residues were the least preferred side chains in our pool of hit compounds. This suggests that modes of binding other than hydrophobic interactions are generally more beneficial for strong binding between these BPBAs and RRE IIB.

Table 3.1 Dissociation Constant and Molecular Weight of Hit Compounds.

peptide	sequence[a]	K_d (μM)	MW(g mol^{-1})
BPBA1	$(WKK)_2$ * $K_{BBA}YWY$	1.4 ± 0.4	1817.99
BPBA2	$(K_{BBA}KF_{BPA})_2$ * $K_{BBA}KKY$	3.3 ± 1.2	2031.10
BPBA3	$(F_{BPA}YF_{BPA})_2$ * NKSY	8.7 ± 2.3	1727.78
BPBA4	$(KKK_{BBA})_2$ * $F_{BPA}TSY$	26.8 ± 4.4	1751.98
BPBA5	$(KK_{BBA}F)_2$ * KKWY	27.2 ± 6.9	1853.04
BPBA6	$(WYK)_2$ * PTWY	28.5 ± 4.4	1646.34
BPBA7	$(KK_{BBA}K)_2$ * KLKY	58.4 ± 4	1742.09
BPBA8	$(K_{BBA}YK)_2$ * HKKY	86.5 ± 10	1836.04

[a] * = Lysine branching unit. Each value is an average of at least three experiments.

To determine the role of the boronic acid functional group in binding to the RRE IIB target RNA, we designed control peptide variants of **BPBA1** and **BPBA3** wherein the boronic acid moiety was removed. As much as possible, modifications were performed with minimal structural perturbation. First, the boronic acid moiety in the *para* position of **BPBA1** was replaced with a hydrogen atom to afford **BPBA1.1**. The change was accompanied by an approximate 6-fold increase in the observed K_d value (1.4 vs 8.2 μM,

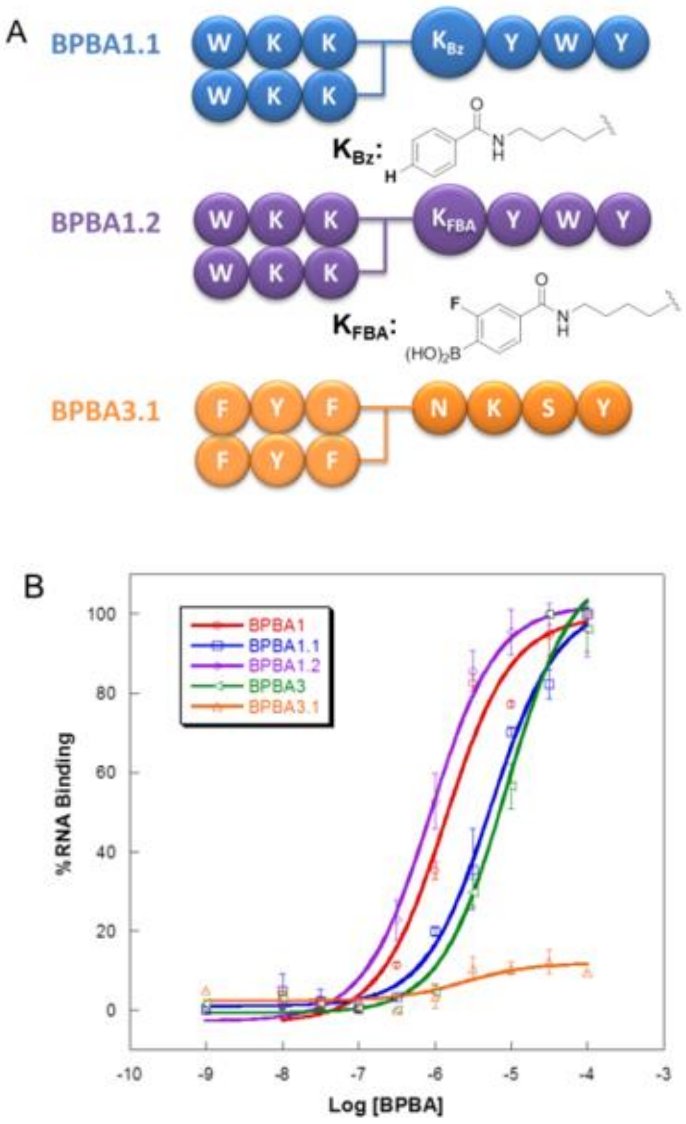


Figure 3.3 (A) Structure and sequences, (B) Binding curves of boron and non-boron containing branched peptides using dot blot assay with RRE IIB RNA.

Figure 3.3). Although the precise role of the boronic acid moiety in binding is currently unknown, the decrease in affinity suggests that the boronic acid functional group interacts with RRE. The single boronic acid moiety in **BPBA1** contributes a modest fraction of binding affinity with RRE IIB. In this case, electrostatic attraction accounts for majority of interactions with the RNA provided by the high density of Lys residues. In order to investigate the hypothesis that structural alterations, such as introduction of electron

withdrawing groups on K_{BBA} which results in increased Lewis acidity, would facilitate increased complexation, we synthesized a fluorinated analog, **BPBA1.2**, wherein fluorine was placed *ortho* to the boronic acid at position A_4. When the K_d was determined by dot blot assay, the binding affinity improved to 0.8 ± 0.1 μM, suggesting an induction of stronger interactions with RRE through the boronic acid moiety. This result supports the notion that the binding affinity of BPBAs can be tuned by manipulating the Lewis acidity of the boronic acid and is consistent with the observation that more electrophilic boronic acids are more acidic as they can form stable Lewis acid-base complexes.[22]

We also investigated the effect of the number of boronic acid moieties in hit peptide **BPBA3**, which contains four boron atoms. When a boronic acid-free analog of **BPBA3** was synthesized (**BPBA3.1**), nearly all binding was abolished (Figure 3.3). The drastic loss of binding with **BPBA3.1** is likely due to the loss of the high number of boron containing side chains because the major source of binding interactions, presumably Lewis acid-base complexation, is lost. These results indicate that boronic acid side chains can be utilized in peptides to boost binding affinity with a highly structured RNA target. Further, such interaction represents an additional and unique mode of binding that increases the repertoire of RNA binding motifs.

3.3 Conclusion

In conclusion, we generated a **BPBA** library that was designed to interrogate the effect of boronic acids when screened against the tertiary structure of a RNA target: RRE IIB. High throughput screening and biophysical characterization of hit compounds resulted in peptides with binding affinities in the low micromolar range, whereas the addition of boronic acid groups introduced a novel, alternative mode of interaction. Characterization

of **BPBA**s demonstrated that the binding affinity can be tuned by changing the electronic property of the boron center. This results highlights that boronic acid moieties can have a significant impact when binding to RNA and they can be easily engineered into peptides and present a new binding mode.

3.4 Materials and Methods

3.4.1 Synthesis of Branched Peptide Boronic Acids Library

We used standard solid phase peptide synthesis techniques to generate the 3.3.4 library via the split and pool method using the previously described apparatus.[12] *N*-α-Fmoc protected L-amino acids (Novabiochem), PyOxim (Novabiochem) and *N*,*N*-Diisopropylethylamine (DIEA, Aldrich) were used in coupling reactions. The synthesis and full characterization of Fmoc-*N*-ε-(4-boronobenzoyl)-L-lysine (K_{BBA}), Fmoc-*N*-ε-benzoyl-L-lysine (K_{Bz}) and Fmoc-*N*-ε-(4-borono-3-fluorobenzoyl)-L-lysine (K_{FBA}) will be described elsewhere. Fmoc-L-4-boronophenylalanine (F_{BPA}) and Fmoc-ANP-OH were synthesized as previously reported.[23-25] Three copies of the library were prepared simultaneously by using a three-fold excess of Tentagel Macrobead-NH_2 resin (2.13 g, 0.57 mmol/g) (Peptides International). The resin was swollen in DCM (20 mL, 2 x 15 min) followed by DMF (20 mL, 15 min). The photocleavable linker Fmoc-ANP-OH (739 mg, 1.71 mmol) was first coupled to the resin in DMF for 3 hrs in the presence of PyOxim (901.53 mg, 1.71 mmol) and DIEA (493 μL, 2.85 mmol). After coupling, the resin was washed with DMF (20 mL, 1 min), DCM (20 mL, 1 min) ,and DMF (20 mL, 1 min). The same washing procedure was applied after every step. Then, 20% piperidine in DMF (20 mL, 2 x 10 min) was used for Fmoc deprotection. A Kaiser test was used after each coupling and deprotection step to confirm reaction completion. *N*-Fmoc amino acids (3

equiv), PyOxim (3 equiv), and DIEA (5 equiv) were added to each reaction vessel in DMF and coupled for 30 min. Fmoc-Lys(Fmoc)-OH was used as a branching unit and molar equivalencies of reagents were doubled in coupling reactions after installation of the branching unit. After Fmoc deprotection of the N-terminal amino acids, the resin was bubbled in a phenylboronic acid solution (0.2 g/mL) 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)/H_2O/TIS (Triisopropylsilane, Acros) (v/v/v) for 3 hrs. After deprotection, the resin was washed extensively with DMF, DCM, and MeOH before drying and storing at −20 °C (Figure 3.4**)**.

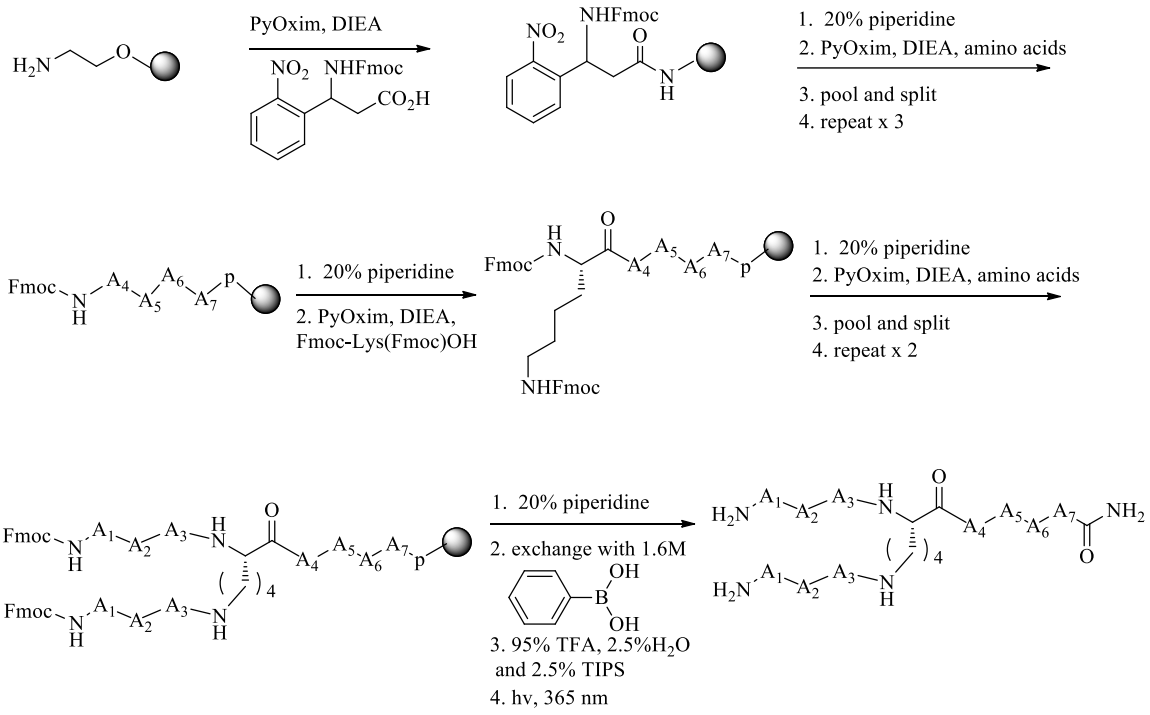


Figure 3.4 Synthesis of branched boronic acid peptide library by solid-phase peptide synthesis (SPPS)

3.4.2 On-bead Screening Assay

DY547 labeled HIV-1 RRE-IIB RNA (5′-DY547-GGC-UGG-UAU-GGG-CGC-AGC-GUC-AAU-GAC-GCU-GAC-GGU-ACA-GGC-CAG-CC-3′) was purchased from

Dharmacon and prepared according to the manufacturer's protocol. To account for the autofluorescence of Tentagel Macrobead-NH_2 resin, a control peptide, $(KYR)_2$*FDS, was incubated in 100 nM DY547 labeled HIV-1 RRE-IIB RNA for 1 hr in phosphate buffer (10 mM potassium phosphate, 100 mM KCl, 1 mM $MgCl_2$, 20 mM NaCl, pH 7.0). These beads were washed extensively and placed into a sterile 96-well plate (Nunc) and imaged by a Zeiss Axiovert 200 fluorescent microscope under a rhodamine filter. The fluorescence intensity was adjusted to remove the background autofluorescence of untreated beads.

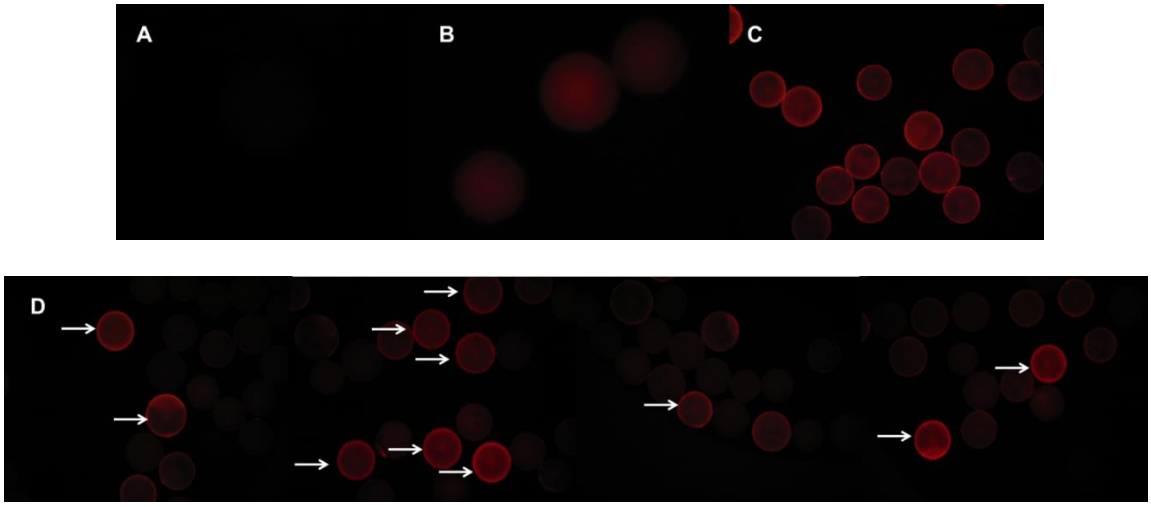


Figure 3.5 Bead images for determining the optimal filter settings. (A) Tentagel Macrobead-NH_2 resin prior to peptide coupling visualized with a fluorescence microscope using rhodamine filter, (B) Branched peptide boronic acid derivatized macrobeads, (C) Positive control peptide $(KYR)_2$*FDS with 100 nM RRE incubated in phosphate buffer for 1 hr, washed, and visualized under fluorescence microscope, and (D) eleven hits from optimized screening of the 3.3.4 branched boronic acid peptide library.

Screening conditions were initially tested using around 2,400 beads from the library. The beads were placed into a 1.5 mL non-stick microfuge tube (Fisher) with a 500 µL final volume of phosphate buffer and mixed by a Barnstead/Thermolyne Labquake rotisserie shaker. The beads were first treated with 1 mg/mL bovine serum albumin (BSA) (New England BioLabs) and 0.8 mg/mL *E. coli* tRNA (Roche) (~3,144-fold molar excess to RRE

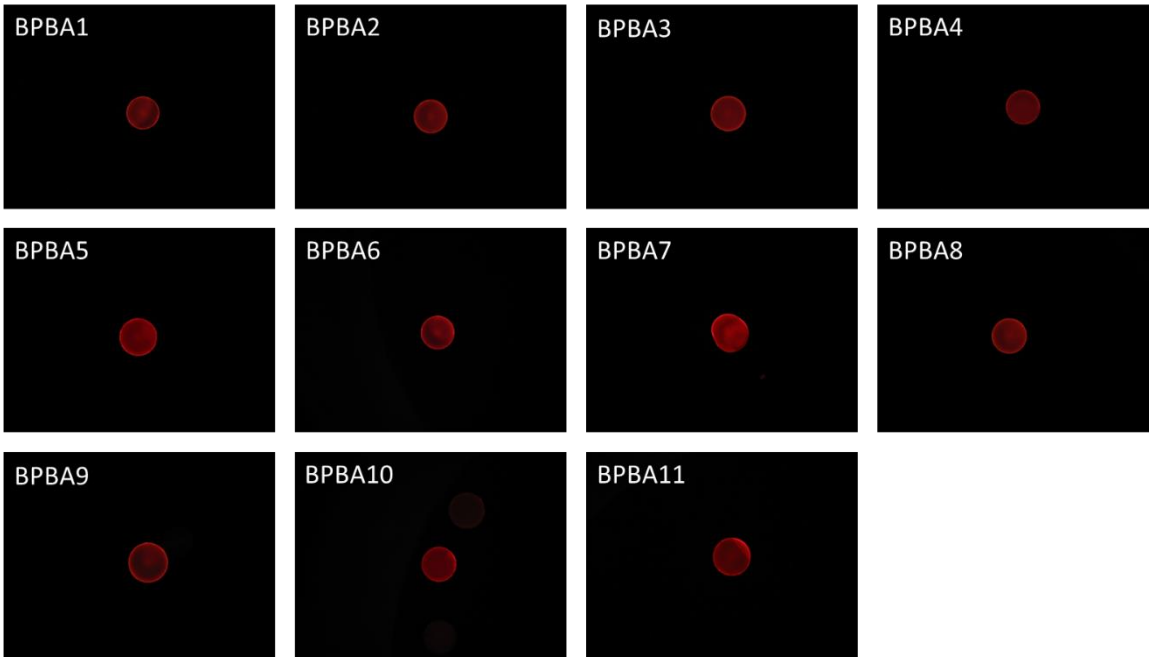


Figure 3.6 Images of isolated hits beads under fluorescence microscope using a rhodamine filter.

IIB RNA) for 3 hrs at rt to block nonspecific binding peptide sequences. Next, the beads were washed 5 times with phosphate buffer and incubated in 500 µL of 10 nM DY547 labeled RRE IIB RNA in phosphate buffer for 5 hrs at rt. After the final incubation, the resin was extensively washed with buffer and the beads were imaged under a fluorescence microscope in a 96-well plate using the previously optimized settings. The initial screening afforded 7 hits which were isolated, rinsed with DMF (5 x 500 µL) and MeOH (5 x 500 µL), and photocleaved in clear non-stick 0.5 mL microfuge tubes in 15 µL of 1:1 MeOH: H_2O (v/v) by irradiation at 365 nm with a 4W handheld UV lamp. The supernatant was retained and subjected to MALDI-TOF analysis. MALDI-MS/MS fragmentation analysis generated 6 sequences (one was deemed a false positive) providing a hit rate of 0.25%. The remaining beads of the 3.3.4 library were screened using more stringent conditions in an attempt to reduce the total number of hit beads. First, the beads were blocked at rt for 3 hrs in phosphate buffer with 1 mg/mL BSA and 1.25 mg/mL tRNA (5000-fold molar excess

to RRE stem IIB RNA). Next, the beads were washed in buffer and incubated with 10 nM DY547 labeled RRE-IIB RNA in phosphate buffer for 5 hrs at RT. The resin was washed extensively after final incubation prior to screening. Another 5 beads (from approximately half of the total compound library) were found with elevated fluorescence by using these more stringent conditions (Figure 3.6**)**. These beads were photocleaved and sequenced by using MALDI-TOF.

3.4.3 Peptide Synthesis, Purification and Characterization

Unlabeled and fluorescein 5-isothiocyanate (FITC) (Sigma) labeled peptides were synthesized on Rink amide MBHA resin (100-200 mesh) (Novabiochem). In the preparation of the FITC-labeled peptides, Fmoc-Lys(ivDde)-OH (Novabiochem) was used as the branching unit.[11] Acetic anhydride in DMF (1:1 v/v) with 10 equivalents of DIEA was used to cap the first N-terminus. Then, ivDde was removed by treatment with 2% hydrazine in DMF for 1 hr, and the second N-terminus was synthesized through theε-N of the Lys side chain. Fmoc-6-Ahx-OH (AnaSpec) was coupled to the N-terminal amino acid to provide a linker for FITC, which prevents autocleavage of FITC under acidic conditions.[11, 26] All subsequent steps were protected from light. FITC (5 equiv) was reacted with the deprotected N-terminus of the peptides for 6 hrs using DIEA (14 equiv). The boronic acid deprotection was performed as before. The supernatant was dried under reduced pressure and the crude peptide was triturated from cold diethyl ether. The peptides were purified by using a Jupiter 4 μm Proteo 90 Å semiprep column (Phenomenex) with a solvent gradient composed of 0.1% TFA in Milli-Q water and HPLC grade acetonitrile. Peptide purity was determined by using a Jupiter 4 μm Proteo 90 Å analytical column (Phenomenex) and peptide identity was confirmed by MALDI-TOF analysis. Unlabeled

peptide concentrations were measured in nuclease free water at 280 nm by using their calculated extinction coefficients. FITC-labeled peptide concentrations were monitored at 495 nm by using the extinction coefficient of FITC at 77,000 mol^{-1} cm^{-1} in 100 mM glycine, pH 9.0.

3.4.4 Dot Blot Assay

Dot blot assays were performed at rt by using a Whatman Minifold I 96 well Dot Blot system and Whatman 0.45 μm pore size Protran nitrocellulose membranes. To determine the binding affinities, 0.4 nM radiolabeled RNA was titrated with the peptide (0.001–100 μM). First, a solution of 0.8 nM ^{32}P-labeled RNA was refolded in 2x phosphate buffer (20 mM potassium phosphate, 200 mM KCl, 1 mM $MgCl_2$, 40 mM NaCl, pH 7.0) by heating at 95 °C for 3 min and then slowly cooling at rt for 20 min. Next, 25 μL of the [^{32}P]-RNA solution was added to 25 μL of peptide in nuclease free water and incubated at rt for 4 hrs. The 50 μL mixtures were filtered through the nitrocellulose membrane which was immediately followed by two consecutive 50 μL washes with 1x phosphate buffer. Peptide binding was visualized by autoradiography using a storage phosphor screen (GE Healthcare) and a Typhoon Trio phosphorimager (GE Healthcare). Densitometry measurements were quantified using ImageQuant TL (Amersham Biosciences). Binding curves were generated using a four parameter logistic equation with Kaleidagraph (Synergy Software): y=m1+(m2-m1)/(1+10^(log(m3)-x); m1=100; m2=1; m3=.000003, where y=percentage of RNA binding, x= log[peptide], m1=percentage of RNA binding affinity at infinite concentration (nonspecific binding), m2= percentage of RNA binding affinity at zero concentration, m3=peptide concentration at 50% binding (K_d). Each experiment was

performed in triplicate and error bars represent the standard deviation calculated over three replicates.

3.4.5 Preparation of ^{32}P-labeled RNA

Wild-type RRE-IIB RNAs was transcribed *in vitro* by T7 polymerase with the Ribomax T7 Express System (Promega) by using previously reported techniques.[27-28] The antisense templates, sense complementary strand (5'-ATGTAATACGACTCACT ATAGG-3'), and RRE IIB reverse PCR primer (5'-GGCTGGCCTGTAC-3') were purchased from Integrated DNA Technologies. Antisense templates were used as follows: RRE IIB RNA 5'-GGCTGGCCTGTACCGTCAGCGTCATTGACGCTGCGCCCATAC CAGCCCTATAGTGAGTCGTATTACAT-3'; RRE IIB was PCR amplified by using HotstarTaq DNA polymerase (Qiagen) followed by a clean-up by procedure using a spin column kit (Qiagen). T7 transcription proceeded at 42 °C for 1.5 hrs. After transcription, 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 1x TBE buffer at 4 °C. The sample was desalted using a Sep-Pak syringe cartridge (Waters Corporation) and lyophilized. The products were dephosphorylated with calf intestinal phosphatase (CIP) in NEBuffer 3 (New England Biolabs) according to manufacturer's protocol. The product was recovered by a standard phenol extraction followed by ethanol precipitation. Purified RNA was stored as a pellet at −80 °C. RRE IIB DNA (5′-GGCTGGTATGGGCGCAGC GTCAATGACGCTGACGGT-ACAGGCCAGCC-3′) was purchased from Integrated DNA Technologies and stored at −20 ˚C.

HIV-1 RRE IIB RNA was labeled at the 5'-end by treating 10 pmol of dephosphorylated RNA / DNA with 20 pmol of [γ-^{32}P] ATP (111 TBq mol^{-1}) and 20 units of T4 polynucleotide kinase in 70 mM Tris•HCl, 10 mM $MgCl_2$, and 5 mM dithiothreitol, pH 7.6. The mixture was incubated at 37 °C for 30 min and then at rt for 20 min. The kinase was heat-inactivated at 65 °C for 10 min. The product was recovered by ethanol precipitation and the purity was examined by using 12% denaturing PAGE followed by autoradiography.

3.4.6 EMSA of BPBA1

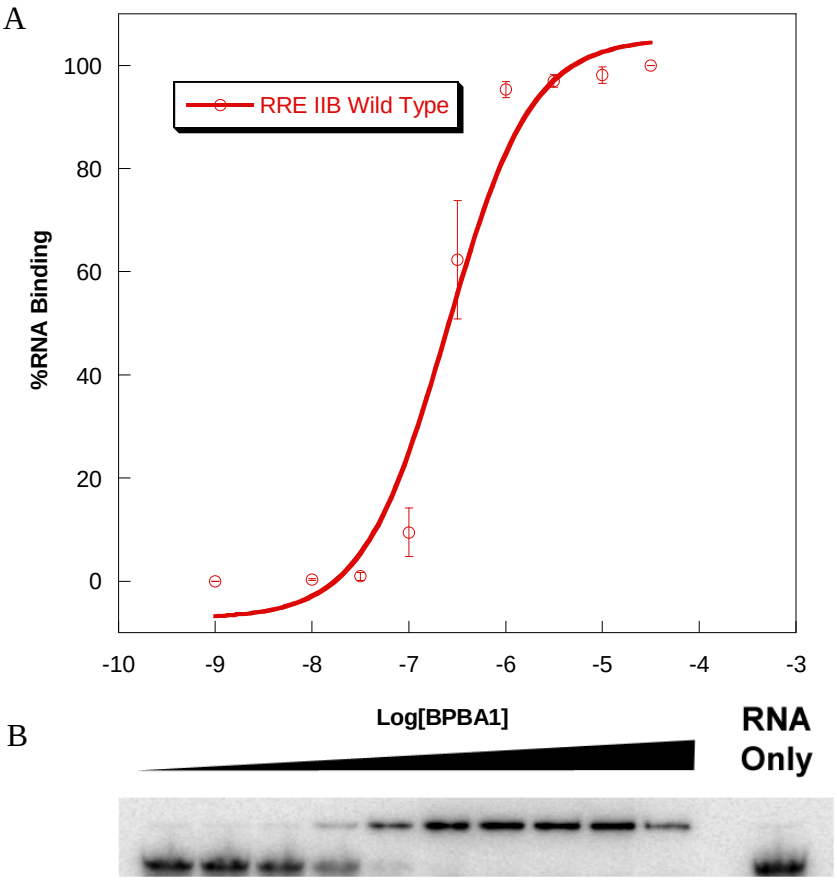


Figure 3.7 Binding curve (A) and EMSA result (K_d =0.25 ± 0.08 μM) of **BPBA1** with RRE IIB RNA (B).

EMSA were performed at rt. First, 4 nM ^{32}P labeled RNA in 2x phosphate buffer (20 mM potassium phosphate, 200 mM KCl, 1 mM $MgCl_2$, 40 mM NaCl, pH 7.0) was refolded by heating at 95 °C for 3 min and cooling at rt for 20 min. A 10 μL solution of the refolded RNA was added to 10 μL of peptide in nuclease free water and incubated at rt for 4 hrs. The final concentration of peptide was varied from 0.001 to 100 μM. After incubation, 3 μL of 30% glycerol was added for loading. Peptide-RNA complexes were resolved on 10% non-denaturing PAGE, which had been pre-run for at least 1 hr. Gels were electrophoresed at 150 V for 35 min at rt. Gels were dried to filter paper and visualized by autoradiography (Figure 3.7). Each experiment was repeated 3 times. Data were measured as the percentage of bound RNA in each lane and error bars represent the standard deviation calculated over three replicates.

3.5 References

1. Baba, M., Recent status of HIV-1 gene expression inhibitors. *Antiviral Res.* **2006,** *71* (2-3), 301-306.

2. Stevens, M.; Clercq, E. D.; Balzarini, J., The regulation of HIV-1 transcription: Molecular targets for chemotherapeutic intervention. *Med. Res. Rev.* **2006,** *26* (5), 595-625.

3. Malim, M. H.; Hauber, J.; Le, S. Y.; Maizel, J. V.; Cullen, B. R., The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* **1989,** *338* (6212), 254-7.

4. Felber, B. K.; Hadzopoulou-Cladaras, M.; Cladaras, C.; Copeland, T.; Pavlakis, G. N., rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. *Proc. Natl. Acad. Sci. U. S. A.* **1989,** *86* (5), 1495-1499.

5. Pollard, V. W.; Malim, M. H., The HIV-1 Rev protein. *Annu. Rev. Microbiol.* **1998,** *52* (1), 491-532.

6. Malim, M. H.; Tiley, L. S.; McCarn, D. F.; Rusche, J. R.; Hauber, J.; Cullen, B. R., HIV-1 structural gene expression requires binding of the Rev trans-activator to its RNA target sequence. *Cell* **1990,** *60* (4), 675-83.

7. Huang, X. J.; Hope, T. J.; Bond, B. L.; McDonald, D.; Grahl, K.; Parslow, T. G., Minimal Rev-response element for type 1 human immunodeficiency virus. *J. Virol.* **1991,** *65* (4), 2131-4.

8. Askjaer, P.; Jensen, T. H.; Nilsson, J.; Englmeier, L.; Kjems, J., The specificity of the CRM1-Rev nuclear export signal interaction is mediated by RanGTP. *J. Biol. Chem.* **1998,** *273* (50), 33414-22.

9. Fornerod, M.; Ohno, M.; Yoshida, M.; Mattaj, I. W., CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* **1997,** *90* (6), 1051-60.

10. Dayton, A. I.; Sodroski, J. G.; Rosen, C. A.; Goh, W. C.; Haseltine, W. A., The trans-activator gene of the human T cell lymphotropic virus type III is required for replication. *Cell* **1986,** *44* (6), 941-7.

11. Bryson, D. I.; Zhang, W.; McLendon, P. M.; Reineke, T. M.; Santos, W. L., Toward targeting RNA structure: Branched peptides as cell-permeable ligands to TAR RNA. *ACS Chem. Biol.* **2012,** *7* (1), 210-7.

12. Bryson, D. I.; Zhang, W.; Ray, W. K.; Santos, W. L., Screening of a branched peptide library with HIV-1 TAR RNA. *Mol. BioSyst.* **2009,** *5*, 1070-1073.

13. Moore, E. C.; Peterson, D.; Yang, L. Y.; Yeung, C. Y.; Neff, N. F., Separation of ribonucleotides and deoxyribonucleotides on columns of borate covalently linked to cellulose. Application to the assay of ribonucleoside diphosphate reductase. *Biochemistry* **1974,** *13* (14), 2904-7.

14. Priestley, E. S.; De Lucca, I.; Ghavimi, B.; Erickson-Viitanen, S.; Decicco, C. P., P1 Phenethyl peptide boronic acid inhibitors of HCV NS3 protease. *Bioorg. Med. Chem. Lett.* **2002,** *12* (21), 3199-3202.

15. Kinder, D. H.; Katzenellenbogen, J. A., Acylamino boronic acids and difluoroborane analogues of amino acids: potent inhibitors of chymotrypsin and elastase. *J. Med. Chem.* **1985,** *28* (12), 1917-25.

16. Baker, S. J.; Ding, C. Z.; Akama, T.; Zhang, Y. K.; Hernandez, V.; Xia, Y., Therapeutic potential of boron-containing compounds. *Future Med. Chem.* **2009,** *1* (7), 1275-88.

17. Kane, R. C.; Bross, P. F.; Farrell, A. T.; Pazdur, R., Velcade: U.S. FDA approval for the treatment of multiple myeloma progressing on prior therapy. *Oncologist* **2003,** *8* (6), 508-13.

18. Rock, F. L.; Mao, W.; Yaremchuk, A.; Tukalo, M.; Crepin, T.; Zhou, H.; Zhang, Y. K.; Hernandez, V.; Akama, T.; Baker, S. J.; Plattner, J. J.; Shapiro, L.; Martinis, S. A.; Benkovic, S. J.; Cusack, S.; Alley, M. R., An antifungal agent inhibits an aminoacyl-tRNA synthetase by trapping tRNA in the editing site. *Science* **2007,** *316* (5832), 1759-61.

19. Crumpton, J. B.; Zhang, W.; Santos, W. L., Facile analysis and sequencing of linear and branched peptide boronic acids by MALDI mass spectrometry. *Anal. Chem.* **2011,** *83* (9), 3548-3554.

20. Battiste, J. L.; Mao, H.; Rao, N. S.; Tan, R.; Muhandiram, D. R.; Kay, L. E.; Frankel, A. D.; Williamson, J. R., Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. *Science* **1996,** *273* (5281), 1547-51.

21. Daugherty, M. D.; Liu, B.; Frankel, A. D., Structural basis for cooperative RNA binding and export complex assembly by HIV Rev. *Nat. Struct. Mol. Biol.* **2010,** *17* (11), 1337-42.

22. Hall, D., *Boronic acids: Preparation, applications in organic synthesis, medicine and materials*. 2 ed.; Wiley-VCH GmbH & Co.: Weinheim, 2011; Vol. 1 and 2.

23. Malan, C.; Morin, C., A Concise preparation of 4-borono-l-phenylalanine (l-BPA) from l-phenylalanine. *J. Org. Chem.* **1998,** *63* (22), 8019-8020.

24. Duggan, P. J.; Offermann, D. A., The preparation of solid-supported peptide boronic acids derived from 4-borono-l-phenylalanine and their affinity for Alizarin. *Aust. J. Chem.* **2007,** *60* (11), 829-834.

25. Tan, D. S.; Foley, M. A.; Stockwell, B. R.; Shair, M. D.; Schreiber, S. L., Synthesis and preliminary evaluation of a library of polycyclic small molecules for use in chemical genetic assays. *J. Am. Chem. Soc.* **1999,** *121* (39), 9073-9087.

26. Jullian, M.; Hernandez, A.; Maurras, A.; Puget, K.; Amblard, M.; Martinez, J.; Subra, G., N-terminus FITC labeling of peptides on solid support: the truth behind the spacer. *Tetrahedron Lett.* **2009,** *50* (3), 260-263.

27. Bryson, D. I.; Zhang, W.; McLendon, P. M.; Reineke, T. M.; Santos, W. L., Toward targeting RNA structure: Branched peptides as cell-permeable ligands to TAR RNA. *ACS Chem. Biol.* **2012,** *7* (1), 210-217.

28. Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C., Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucl. Acids Res.* **1987,** *15* (21), 8783-8798.

Chapter 4. Targeting Folded RNA: A Branched Peptide Boronic Acid That Binds to a Large Surface Area of HIV-1 RRE RNA

Attribution

This chapter was taken from Zhang, W.; Bryson, D. I.; Crumpton, J. B.; Wynn, J.; Santos, W. L.; Targeting folded RNA: A branched peptide boronic acid that binds to large surface area of HIV-1 RRE RNA". *Org. Biomol. Chem.* **2013**, 11, 6263. The author of this dissertation performed a major portion of the work described in this chapter. She synthesized the majority of the hit peptides for biophysical characterization, performed a significant amount of the dot blot and electrophoretic mobility shift assays and synthesized the unnatural amino acids containing boron. She was also responsible for all binding data interpretation, RNA synthesis, and Hill plot analyses. David I. Bryson was responsible for RNA footprint assays and footprint optimization. He also prepared and designed the majority of the RNA constructs. Jason B. Crumpton performed all MALDI analyses of peptides and aided in the synthesis of the unnatural amino acids containing boron. Jessica Wynn performed all cell-based assays, and aided in the synthesis of several hit peptides. The author also significantly contributed to the writing and editing of the manuscript. The final manuscript was prepared by Dr. Webster L. Santos. [Reproduced by permission of the Royal Society of Chemistry]

Abstract

On-bead high throughput screening of a medium sized (1000-2000 Da) branched peptide boronic acid (BPBA) library, consisting of 46,656 unique sequences, against HIV-1 RRE RNA generated peptides with binding affinities in the low micromolar range. In particular, **BPBA1** had a K_d of 1.4 μM with RRE IIB, 27-fold preference for RNA over DNA, and selectivity of up to >75 fold against a panel of RRE IIB variants. Structure–activity studies suggest that the boronic acid moiety and "branching" in peptides are key structural features for efficient binding and selectivity for the folded RNA target. **BPBA1** was efficiently taken up by HeLa and A2780 cells. RNA-footprinting studies revealed that the **BPBA1** binding site encompasses a large surface area that spans the upper stem as well as the internal loop regions of RRE IIB.

4.1 Introduction

Ribonucleic acid (RNA) is a unique macromolecular entity that plays key roles in a living cell, serving in essential functions such as a carrier of genetic information. RNA also catalyzes protein synthesis and regulates gene expression. Along with protein–protein interactions, RNA–protein interactions are the gateway to the diversity of function that mediates a variety of biological effects. One attractive approach to perturbing the system is to inhibit RNA–protein interaction by disassembling the construct by binding to either the protein or RNA. Targeting RNA as a macromolecular entity with small molecules is a herculean task that is complicated by RNA structural dynamics—specific binding to a single conformation is difficult. Despite considerable campaigns, clinically effective small molecule inhibitors of RNA are rare outside of antibiotics that bind to the ribosome.[1-2]

In silico approaches, either from virtual screening methods based on either binding preferences of ligands toward certain RNA motifs[3-4] or from docking molecules onto RNA dynamic ensembles generated by nuclear magnetic resonance and molecular dynamics studies, are powerful methods for the rational design of RNA ligands.[5-6] While these investigations are a significant leap forward, these approaches are still in their infancy. A complementary approach is high throughput screening of chemical libraries against an RNA target.[7-9] Chemical libraries that exploit chemical space outside the region used for protein-targeting small molecules are ideal since structural features present in RNA are vastly different than proteins.[7]

Although chemically similar, the presence of 2'-hydroxyl groups and other nucleotide modifications in RNA generate far more complex tertiary structures than those found in DNA.[10] For example, DNA forms a double stranded helical structure while a single stranded RNA folds onto a variety of secondary structures. Hairpins, bulges, loops, pseudoknots, and turns give rise to three-dimensional architecture akin to targetable regions of proteins. Theoretically, these strategies can create unique binding pockets suitable for intermolecular binding with small molecules. While attractive, the discovery of small molecules that selectively bind to a well-folded RNA has proven difficult.[1-2] New molecular scaffolds that can recognize three-dimensional structures of RNA are needed. Recently, Disney and co-workers used a modular assembly approach to target r(CCUG) repeats that cause myotonic dystrophy type 2.[4] Three molecules of kanamycin A that bound to the internal loop were tethered by a linker and resulted in the multivalent inhibition of the protein-RNA complex with an IC_{50} of 25 nM.

In contrast to molecules that target RNA via Watson-Crick base pairing, we surmise

that an alternative mode of binding the native three-dimensional fold of RNA could be advantageous. Firstly, this would afford a complementary approach to targeting RNA molecules with inaccessible primary sequences as a consequence of RNA folding. Secondly, the tertiary structure of RNA could present multiple crevices or pockets suitable for medium-sized molecules to penetrate and bind favorably. A collection of small binding interactions could accumulate to significant affinity that can also aid in selectivity. We previously developed a first generation branched peptide library (BP) that selectively bound with an HIV-1 related RNA tertiary structure, the transactivation response element (TAR).[11-12] We demonstrated that medium-sized BPs (MW ~1,000-2,000 Da) were cell permeable and displayed minimal to no toxicity. Moreover, our studies revealed that branching in peptides plays a significant role in increasing binding affinity to the target RNA. More recently, we reported the screening of a second generation BP library that was diversified with unnatural amino acids decorated with boronic acid moieties against HIV-1 RRE IIB RNA.[13] These medium-sized branched peptide boronic acids (BPBAs) were capable of binding to the tertiary structure of HIV-1 RRE IIB in the low micromolar regime.

The Rev/RRE export pathway is essential for HIV-1 viral replication.[14] The RRE–Rev interaction is also completely viral in nature, which provides a high-value therapeutic target completely independent from the natural cellular processes of the host. This is a huge advantage that could allow the interaction to be targeted selectively with minimal risk of side effects. Owing to the therapeutic potential of the Rev/RRE export pathway, many ligands have been designed to interrupt the RRE–Rev interaction with limited clinical success. Small molecules such as neomycin B, and other aminoglycosides, have demonstrated submicromolar binding ligands of RRE; however, their lack of binding

specificity, poor cell permeability, and toxicity make them therapeutically undesirable.[2, 15-17] Other inhibitors such as aromatic heterocycles, antisense oligonucleotides, transdominant negative Rev mutant proteins, RRE-based decoys, cyclic peptides, and α-helical peptidomimetics, have also been identified yet none of these have found clinical success. [18-32]

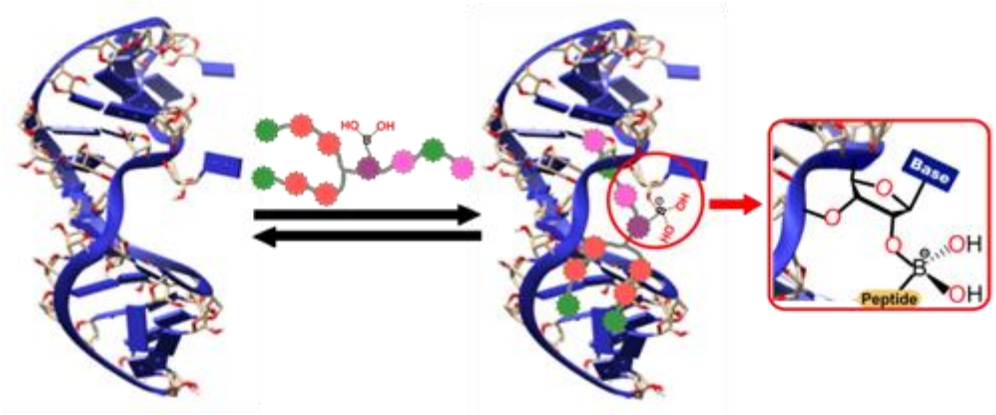


Figure 4.1 RNA-branched peptide boronic acid interaction: formation of a possible reversible covalent bond between the 2'-hydroxyl group of ribose and empty p-orbital of boronic acids.

From an academic standpoint, RNA ligands that minimize nonspecific electrostatic interactions are highly desirable. We hypothesized that the empty *p*-orbital of boron would be a surrogate for a positive charge and act as an acceptor for the 2′-hydroxyl group or other Lewis bases of the RNA (Figure 4.1). Further, we expected that the boronic acid moiety would demonstrate bias against DNA and prefer binding to RNA. In this report, we detail the biophysical characterization of a branched peptide boronic acid (**BPBA1**) against HIV-1 RRE IIB, other variant RNAs, and DNA. Our investigations suggest that **BPBA1** forms complex interactions that span a large surface area of the structured RNA and that its binding affinity is highly dependent on the BP sequence.

4.2 Results and Discussion

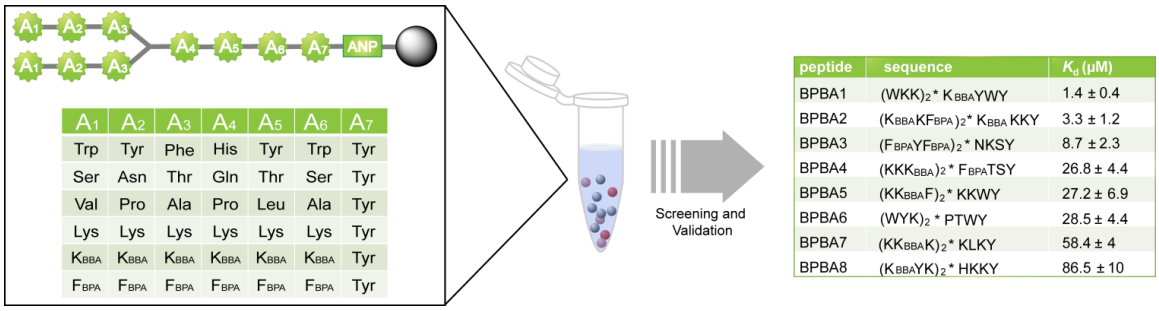
4.2.1 Library Design and On-bead High-throughput Screening

We synthesized a 46,656-membered BPBAs library by split-and-pool synthesis and subjected it to an on-bead high throughput screening assay.[13] Each library member featured two identical N-termini branched by a Lys residue and a single C-terminus. The peptides were linked to beads through the C-termini via a photocleavable linker (3-amino-3-(2-nitrophenyl)propionic acid, ANP). There were six variable amino acid positions at N- and C-termini (A_1–A_3 and A_4–A_6, respectively) and six possible side chains in each position (Figure 4.2). The identity of each side chain was chosen for its potential to interact with the RRE IIB target RNA through hydrophobic interaction, electrostatic attraction, hydrogen bonding, and π-stacking interactions. In addition, two unnatural amino acids containing boron side chains were incorporated at each variable position to provide reversible covalent bond formation between boron and a Lewis base presented by the RNA target, such as the 2' hydroxyl group of the ribose sugars. By varying the side chain length and boron Lewis acidity of these residues (F_{BPA} and K_{BBA}), we were able to probe their effects on binding to our RNA target. Finally, a Tyr was included at position A_7 to quantify peptide concentrations. Presentation of diverse moieties on the bead surface would allow interaction of DY547-labeled RRE IIB with all possible modes of binding.

Upon synthesis, we subjected the library to an on-bead high throughput screening assay. We first minimized non-specific binding by blocking the library with excess bovine serum albumin and competitor tRNA. DY547-labeled RRE-IIB RNA was added during the second incubation. Increased fluorescence of beads, which was monitored by fluorescence microscopy, indicated specific binding to the target RNA. Eleven beads were

selected, photocleaved and sequenced by MALDI MS/MS analysis following a previously published procedure.[33]

Standard dot blot assay was used to verify the dissociation constant of hit BPBAs. In particular, one peptide (**BPBA1**) stood out for its low micromolar binding affinity (1.4 μM).[13] Further modifications were made wherein the boronic acid was removed (BPBA1.1, 8.2 μM) or fluorine was placed *ortho* to the boronic acid at position A_4 (BPBA1.2, 0.8 μM) to probe the role of the boron atom in RNA binding. The results suggested that the boronic acid functional group had involvement in the interaction with RRE IIB, and that the Lewis acidity of the boronic acid had a significant effect on binding affinity. These results demonstrate that the incorporation of boronic acid side chains into peptides can provide increased binding affinities to structured RNA targets and introduce a unique binding mode that increases the gamut of RNA-binding motifs.



peptide	sequence	K_d (μM)
BPBA1	$(WKK)_2$* $K_{BBA}YWY$	1.4 ± 0.4
BPBA2	$(K_{BBA}KF_{BPA})_2$* K_{BBA} KKY	3.3 ± 1.2
BPBA3	$(F_{BPA}YF_{BPA})_2$ * NKSY	8.7 ± 2.3
BPBA4	$(KKK_{BBA})_2$ * F_{BPA}TSY	26.8 ± 4.4
BPBA5	$(KK_{BBA}F)_2$* KKWY	27.2 ± 6.9
BPBA6	$(WYK)_2$ * PTWY	28.5 ± 4.4
BPBA7	$(KK_{BBA}K)_2$* KLKY	58.4 ± 4
BPBA8	$(K_{BBA}YK)_2$ * HKKY	86.5 ± 10

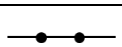
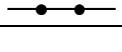
Figure 4.2 High throughput screening of the 3.3.4-branched peptide boronic acid library reveal hit compounds with varying binding affinities.

4.2.2 BPBA1 Binds RRE IIB via All Branches of the Peptide

To determine whether branching in our peptides contributed to the binding affinity toward RRE IIB (Table 4.1), control linear peptide variants of **BPBA1** were synthesized. To minimize perturbations from the parent structure of the peptide, glycine was used in place of the branching lysine residue between positions A_3 and A_4 (**LPBA1**, **LPBA4**, and

LPBA5). Truncation of the N-terminal branching peptide fragment linked to either the α- or ε-nitrogen of lysine resulted in peptides **LPBA3** and **LPBA1**, respectively, with a dramatic loss (> 50 fold) of binding affinity compared to **BPBA1**. Moreover, when the C-terminal fragment was removed to generate **LPBA2**, a similar decrease (> 50 fold) in binding affinity was observed. Collectively, these results suggested that all "branches" of the branched peptide are required for high affinity binding. In addition to probing the effects of branching, modifications to the structure of **BPBA1** simultaneously probed the contribution of electrostatics to the binding of RRE. For example, a marked decrease in affinity was observed for **LPBA2**, which is isoelectronic with **BPBA1**. Peptide derivatives **LPBA1** and **LPBA3** with two fewer lysine residues than **BPBA1** and decreased net positive charge also demonstrated significant loss in binding affinity (K_d > 75 μM). To investigate whether the change in binding affinity could be a consequence of decreased electrostatic interactions, the WKK branching fragment was installed on either the N-terminus or C-terminus of **LPBA1** to provide **LPBA4** and **LPBA5**, respectively. Consistent with previous reports that electrostatic interactions have a significant role in

Table 4.1 Sequence and dissociation constant of **BPBA1** variants

Compound	Scheme	Sequence [a]	K_d (μM)
BPBA1		**$(WKK)_2$*K_{BBA}YWY**	**1.4 ± 0.4**
LPBA1		**WKKGK_{BBA}YWY**	**>75**
LPBA2		**$(WKK)_2$***	**>75**
LPBA3		**Ac(WKK)*K_{BBA}YWY[b]**	**>75**
LPBA4		**WKKWKKGK_{BBA}YWY**	**7 ± 2**
LPBA5		**WKKGK_{BBA}YWYWKK**	**9 ± 2**
RPBA1		**AcYKW*KWKY (KWK_{BBA})[c]**	**>75**

[a] * = Lysine branching unit. [b] α-Nitrogen of lysine is acetylated and KKW is attached to the ε-nitrogen. [c] α-Nitrogen of tyrosine is acetylated and KWK_{BBA} is attached to the ε-nitrogen of branching lysine. Each value is an average of at least three experiments.

boosting binding affinity and not necessarily selectivity, [1, 11]the K_d values of these peptides (~ 8 μM) were improved compared to LPBA2 but remained at least 5-fold weaker than BPBA1. Gratifyingly, when a branched peptide containing a scrambled sequence of BPBA1was tested (**RPBA1**), the K_d value increased to > 75 μM suggesting that the sequence of BPBA1 is essential for tight binding to RRE IIB. Taken together, these results indicated that binding occurs through interactions with the three branches of **BPBA1**, and that the sequence of the branched peptide plays a significant role in binding.

4.2.3 Selectivity of BPBA1 Toward RRE IIB Tertiary Structure

Several mutations were made to RRE IIB to determine whether **BPBA1** could discriminate between the native tertiary structures of the target RNA and closely related structural analogs. RRE IIB RNA is composed of two stems, two internal loops, a single nucleotide bulge and an apical loop, which are expected to contribute to its tertiary structure (Figure 4.3). We synthesized 'hexaloop RNA', where the size of the trinucleotide apical loop (AAU) was increased to a hexa-nucleotide loop (AUGGCC) (Figure 4.3A). The measured K_d for this variant using dot blot assay increased by ~3-fold to 4.4 μM compared to the wild type (Figure 4.3B, Table 4.2). We next removed the 4 base pairs in the upper stem to generate 'stem B deleted RNA' and obtained a similar K_d of 3.7 μM. The small change in K_d with these two mutant RRE RNA structures suggested that **BPBA1** had a minor interaction with the apical loop and stem B region of RRE IIB. Hill analysis of the hexaloop RNA and stem B deleted RNA provided coefficients (n) of 1.7 and 1.6, respectively, and suggested cooperative binding of **BPBA1** to these variant structures. We also probed regions toward the lower stem by generating 'loop B/bulge A deleted RNA',

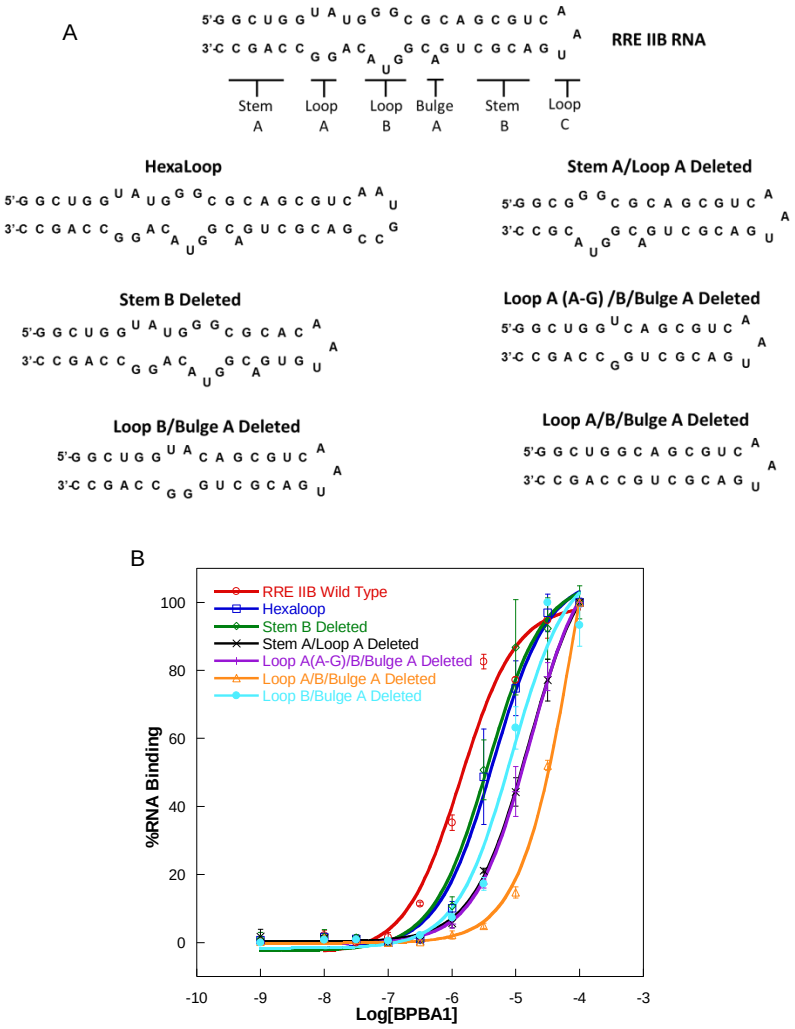


Figure 4.3 (A) Sequence and structure of RRE IIB and variants, (B) Titration curves of **BPBA1** with indicated RNAs.

Table 4.2 Dissociation constants of **BPBA1** with indicated RNAs

RRE IIB Variants	K_d (μM)
RRE IIB wild type	1.4 ± 0.4
HexaLoop	4.4 ± 0.8
Stem B Deleted	3.7 ± 0.9
Loop B/Bulge A Deleted	8.7 ± 2.7
Loop A(A-G)/B/Bulge A Deleted	16.5 ± 0.9
Stem A/Loop A Deleted	15.9 ± 1.0
Loop A/B/Bulge A Deleted	91.7 ± 14.5

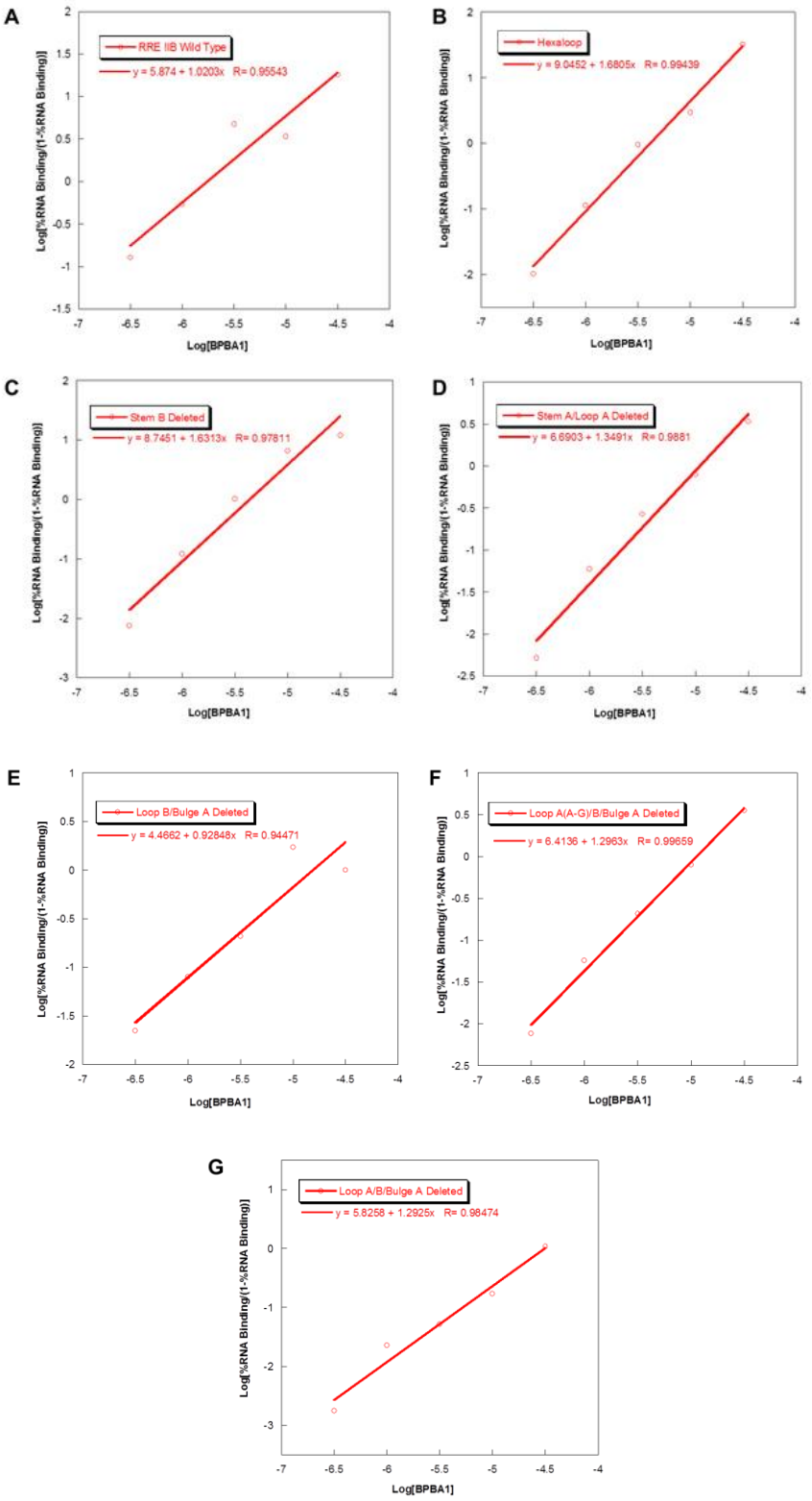


Figure 4.4 Hill plot of **BPBA1** (from dot blot data) against (A) RRE IIB Wild Type, (B) HexaLoop, (C) Stem B Deleted, (D) Stem A/Loop A Deleted, (E) Loop B/Bulge A Deleted, (F) Loop A(A-G)/B/Bulge A Deleted, and (G) Loop A/B/Bulge A Deleted.

and observed 6-fold lower affinity (K_d = 8.7 μM). The RRE IIB variants 'loop A(A-G)/B/bulge A deleted', which contained a smaller loop A from the removal of the A-G mismatch, and 'stem A/loop A deleted RNA' were synthesized to determine whether the affected region of the RNA is the potential site for binding **BPBA1**. These mutant RNAs demonstrated >10 fold lower affinity with **BPBA1** (Figure 4.3B, Table 4.2) and suggested major interaction with the nucleotides deleted from RRE IIB. We suspected that changes or removal of sequences in loop A, loop B, and bulge A would result in an altered tertiary structure that minimizes interaction with **BPBA1**. When these structural elements were eliminated to produce 'loop A/B/bulge A deleted RNA' as a stem loop with a significantly altered tertiary structure, a dramatic loss (> 75 fold) in binding affinity was observed. These results supported our hypothesis that the exact tertiary structure of RRE IIB is required for optimal binding with **BPBA1**, and suggested that **BPBA1** likely makes contacts primarily with nucleotides present in loops A and B. Hill analysis of the dot blot data from 'loop B/bulge A deleted RNA,' 'stem A/loop A deleted RNA,' 'loop A/B/bulge A deleted RNA,' and 'loop A(A-G)/B/bulge A deleted RNA' generated Hill coefficients of 0.9, 1.3, 1.3, and 1.3, respectively, indicating noncooperative binding of **BPBA1** (Figure 4.4).

To further characterize the selectivity against other RNA structures, binding affinities between **BPBA1** and RRE IIB were measured by dot blot in the presence of excess bacterial tRNA. Initially, a 10-fold molar excess of competing tRNA relative to RRE IIB was included during incubation with the peptide. This resulted in an observed K_d (2.1 μM) that was within experimental error with that obtained in the absence of tRNA (Figure 4.5). Encouraged by this result, we increased the amount of competing tRNA to a 1000-fold molar excess. Gratifyingly, the observed K_d was 5.3 μM, representing a minute

shift in RRE IIB binding. These results suggested that peptide **BPBA1** is selective for RRE IIB in the presence of competing complex mixture of tRNA structures. It is also important to note that these results are quite improved from our previous generation of branched peptides that did not feature boronic acids. Peptides from the previous library demonstrated a dramatic shift (10-fold) in binding affinity in the presence of only a 10-fold molar excess of tRNAs.[11]

Next we explored the selectivity of **BPBA1** against DNA, composed of the same RRE IIB sequence, by dot blot assay. The high dissociation constant (31.8 ± 6.6 μM) confirmed our hypothesis that BPBAs confer selectivity toward RNA (23-fold). A change in the tertiary structure, loss of the 2'-OH groups, or a combination of both could account for the dramatic shift in binding affinity.

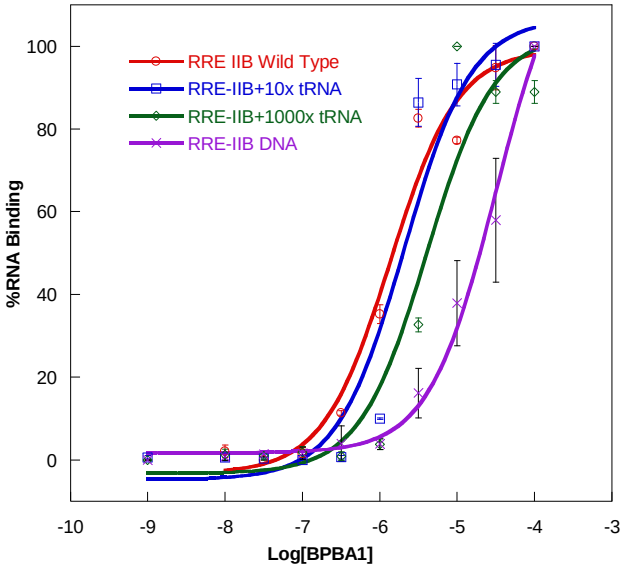


Figure 4.5 Titration curves of **BPBA1** with WT RRE IIB (red), RRE IIB DNA (purple), 10-fold tRNA (blue), and 1000-fold tRNA (green)

4.2.4 Determination of the BPBA1 Binding Site by RNase Protection Assay

The results of our binding selectivity experiments with RRE IIB mutants allowed the evaluation of structural elements required for binding of **BPBA1**. To better determine

regions of specific nucleotide contacts with RRE IIB, we performed ribonuclease protection assays. For the assays, 5'-^{32}P-labeled RRE IIB was incubated with varying concentrations (up to 20 μM) of **BPBA1** in the presence of RNase T1, RNase A, or RNase V1 (Figure 4.6). Potential binding sites were confirmed based on the ability of the peptide to protect the RNA from enzymatic cleavage. While there were no noticeable changes in band intensity with RNase T1, the most prominent cleavage bands over all the experiments

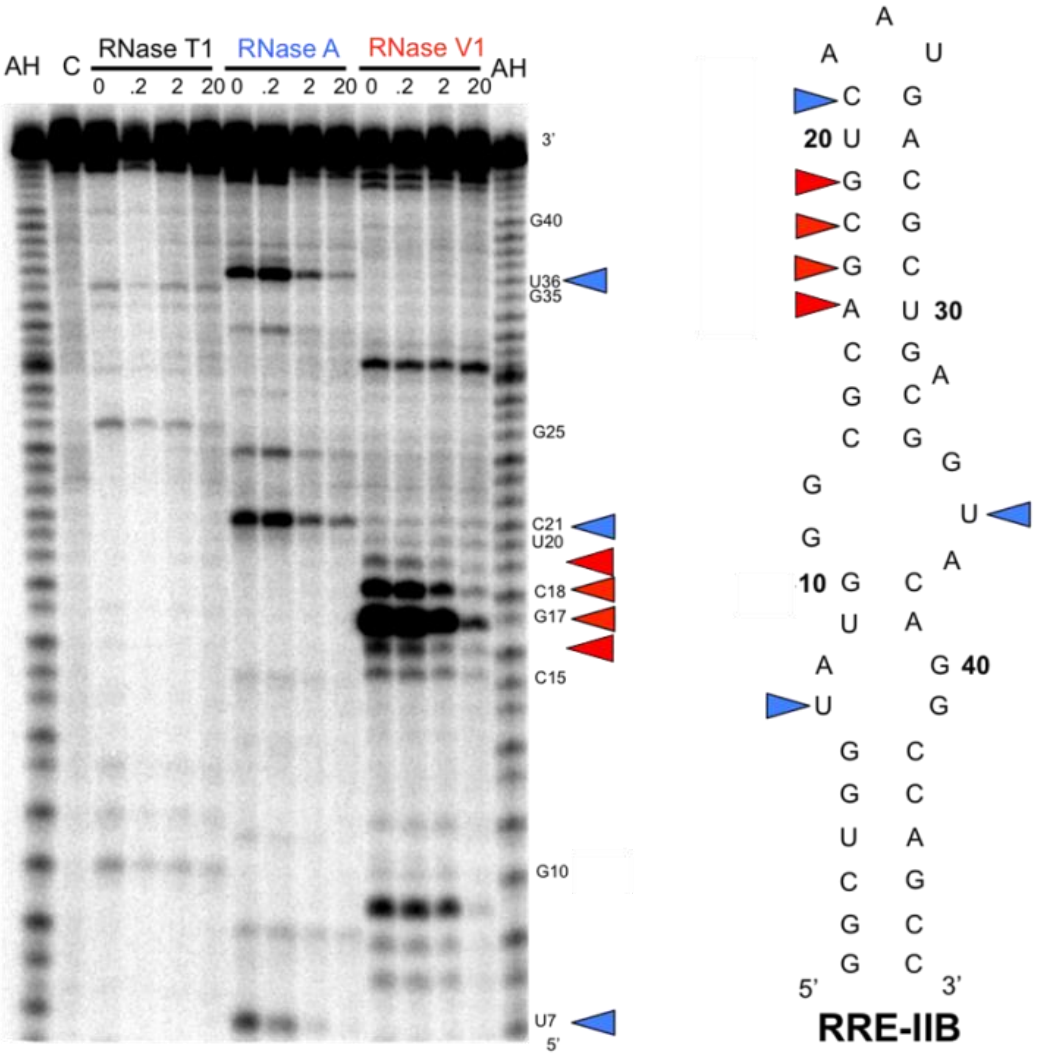


Figure 4.6 RNase protection assay of RRE IIB. The gel depicts the autoradiogram of alkaline hydrolysis (AH) and RNase protection experiments using RNases T1, A, and V1 with increasing concentration of **BPBA1**. Colored triangles highlight bases protected from cleavage by RNase A (blue) and RNase V1 (red).

conducted were observed between A16 through G19 using RNase V1, which preferentially cleaves double stranded regions of RNA. As shown in Figure 4.6, there was a concentration dependent protection (0.2 μM to 20 μM) of this region in the presence of **BPBA1**, which suggested that this portion of RRE IIB served as a potential binding site. Interestingly, this region is also the site where the native protein partner Rev binds.[34-35] **BPBA1** also visibly protected U7, C21, and U36 from cleavage by RNase A, which hydrolyzes at the 3' side of unpaired pyrimidine bases. In particular, U7 and U36 were potential contact points located in the internal loop region of RRE IIB. Indeed, this same internal loop region, when deleted, caused dramatically reduced binding affinities with **BPBA1**. Taken together, the RNA footprinting assay indicated that the binding site for **BPBA1** encompasses a large surface area constituting stem B and loops A and B. Hill analysis (n = 1) supported a noncooperative, possibly 1:1 binding stoichiometry between BPBA1 and RRE IIB. This indicated that BPBA1 may be bound in a folded groove saddled between the internal loop regions and upper stem bases of RRE IIB, since this compound appeared to be interacting with these structural elements. Although the precise nature of the RNA–peptide interaction is currently unknown, protection from enzymatic cleavage of these regions could result either from steric blockade or remodeling of its tertiary structure.

BPBA8 proved to be less apt to protect RRE-IIB from enzymatic hydrolysis compared to **BPBA1** overall. Like **BPBA1**, **BPBA8** was able to protect U7 and G10 from cleavage by RNase A and RNase V1, respectively, but levels of hydrolysis at U36 were unchanged. Curiously, as **BPBA8** concentrations were increased G35 became more prone towards digestion by RNase T1, an enzyme that is selectively cleaves the 3' side of single stranded guanosine bases. Previous reports have demonstrated that a non-canonical G:G

base-pair could be formed between G12 and G35 in the high affinity site of Rev.[36-37] Thus it is likely that this homo-purine base pair was disrupted upon binding **BPBA8**. Although nuclease protection by **BPBA8** provided less information compared to **BPBA8**, we concluded that **BPBA8** was likely binding in the bulge region of RRE-IIB. However, a simultaneous interaction with the upper stem structure was not occurring.

4.2.5 Cell Permeability and Cytotoxicity of Branched Peptide Boronic Acids

Our previous generation of branched peptides that featured Arg-rich N-termini exhibited excellent cell permeability and non-cytotoxic properties.[11] The key feature of the current library was the incorporation of unnatural amino acids with boronic acid functional groups and the elimination of Arg in place of Lys. It was unclear whether these changes would have an impact on cell permeability. Despite these changes, we suspected that BPBAs would be cell permeable in part because they maintained a medium molecular weight (~1000–2000 Da), and contained multiple basic residues.[38] Cellular uptake was assayed in both HeLa and A2780 cells by incubation with the FITC-labeled version of **BPBA1** (**FBPBA1**) by using an established procedure in tissue culture medium for 4 hr at 37 ℃.[11] After washing, cells were fixed with 4% (w/v) paraformaldehyde in PBS followed by incubation with BSA and then DAPI stain. Cells were imaged by using a confocal microscope. Initially, intense fluorescent background prevented the acquisition of a confocal image due to poor cell contrast. Comparison of a previously imaged branched peptide (**FL3**) to **FBPBA1** under identical conditions and acquisition parameters shows a large imaging difference between the two peptides (Figure 4.7). Since the difference between the two peptides was the installation of boronic acid on **FBPBA1**, it was thought that boronic acid moieties on the residual peptide that adhered to the coverslip were

leaching into the mounting media, and thereby increasing the background noise to unmanageable levels when imaged. Once a modified procedure was developed (see materials and methods), both HeLa and A2780 cells incubated with **FBPBA1** showed fluorescence evenly distributed throughout the cytoplasm and nucleus, suggesting that the boronic acid moiety did not have a negative impact on the membrane permeability of BPBAs (Figure 4.8A–B).

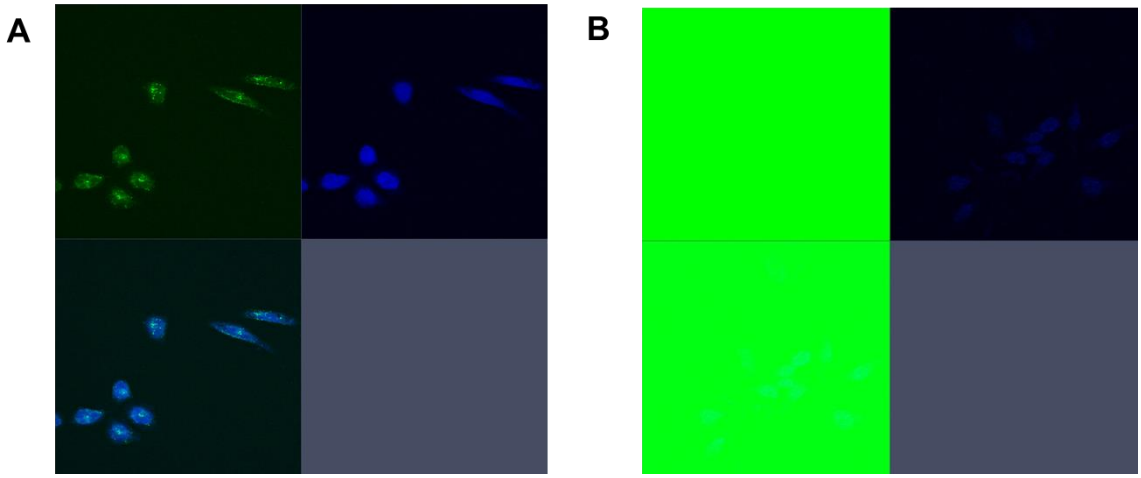


Figure 4.7 Cell microscopy from initial incubation of peptides in the presence of mounting media. Confocal image of peptide **FL3** (A) and **FBPBA1** (B) showed large difference in fluorescent background with FITC channel (top left). Top left: fluorescence image of cells; Top right: DAPI staining of the nucleus; Bottom left: overlay of the two images; Bottom right: Empty

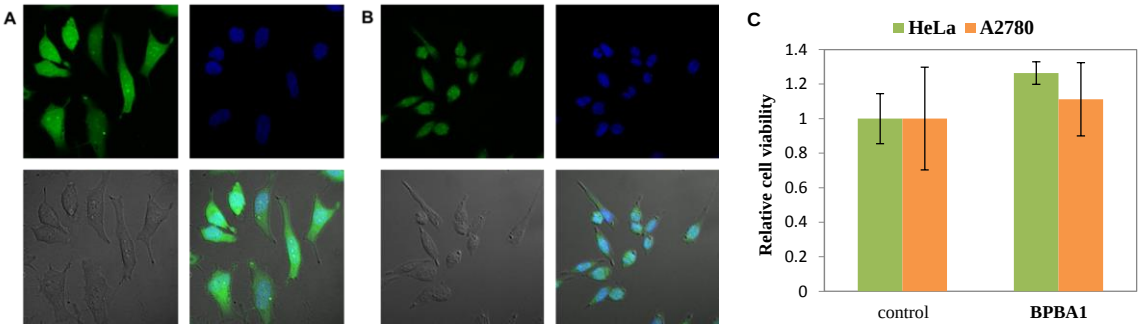


Figure 4.8 Cellular uptake of 1 μM **FBPBA1** into (A) HeLa and (B) A2780 cells. Top left: fluorescence image of cells; top right: DAPI staining of the nucleus; bottom left: DIC image; bottom right: overlay of the three images. (C) MTT cell toxicity assay of 30 μM **BPBA1** at 24 hrs of exposure.

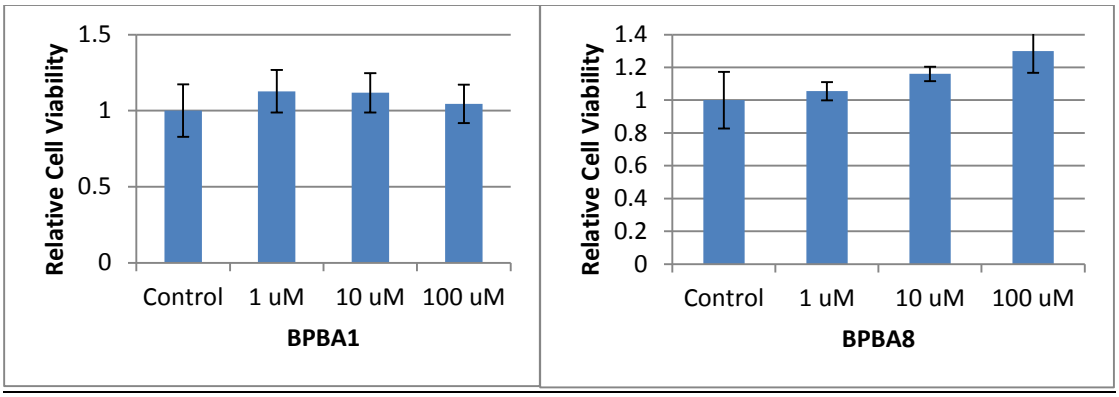


Figure 4.9 MTT cell toxicity assay using A2780 cells for **BPBA1** and **BPBA8** for 4 hr.

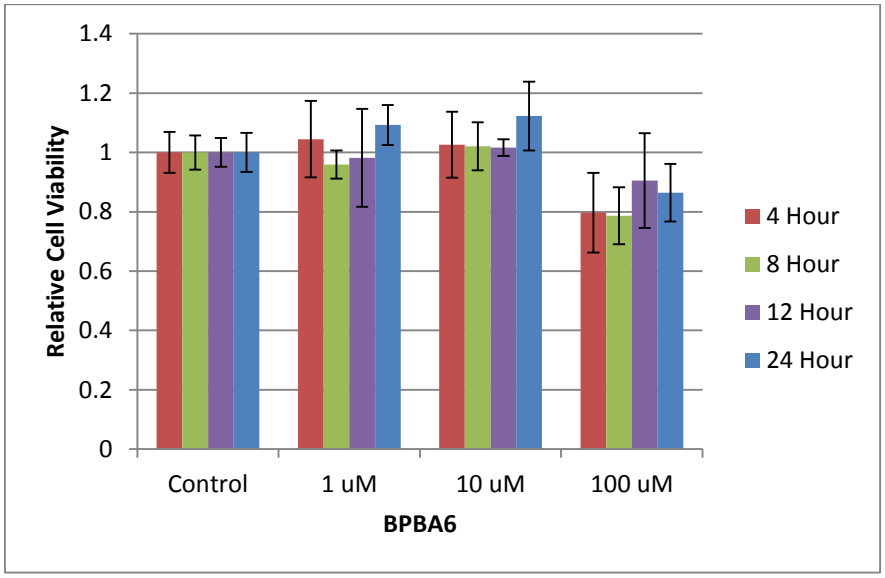


Figure 4.10 A2780 cell MTT assay 24 hr time study as a function of the concentration of **BPBA6**.

Both HeLa and A2780 cells were viable upon incubation with various **BPBAs** when monitored *via* an MTT assay. **BPBA1 and BPBA8** were shown to be non-toxic in concentrations up to 100 μM when incubated with A2780 cells for 4 hr (Figure 4.9). In time studies, **BPBA6** remained non-toxic to A2780 cells even after a 24 hours incubation, with a slight drop-off in viability at 100 μM for all incubation times (Figure 4.10). The concentration of **BPBA6** was then reduced to 30 μM for the time study, and delightfully

the cells treated with peptide showed excellent viability relative to control cells at this concentration for all incubation times (Figure 4.11). Further, **BPBA1** was incubated in both cell lines at viable concentration for 24 hr in an MTT assay suggesting that this branched peptide boronic acid was nontoxic at concentration ranges from 100 nM to 30 μM (Figure 4.8C). It should be noted that increasing the concentration of **BPBA1** to 60 μM and above showed a dramatic reduction in cell viability (Figure 4.12).

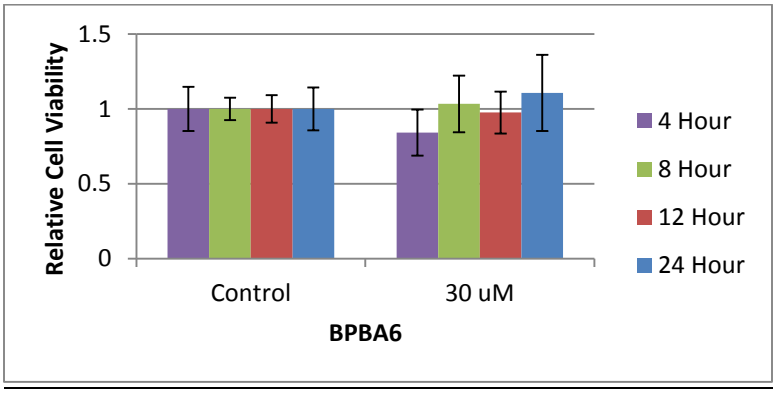


Figure 4.11 A2780 cell MTT assay 24 hr time study of **BPBA6** at 30 μM.

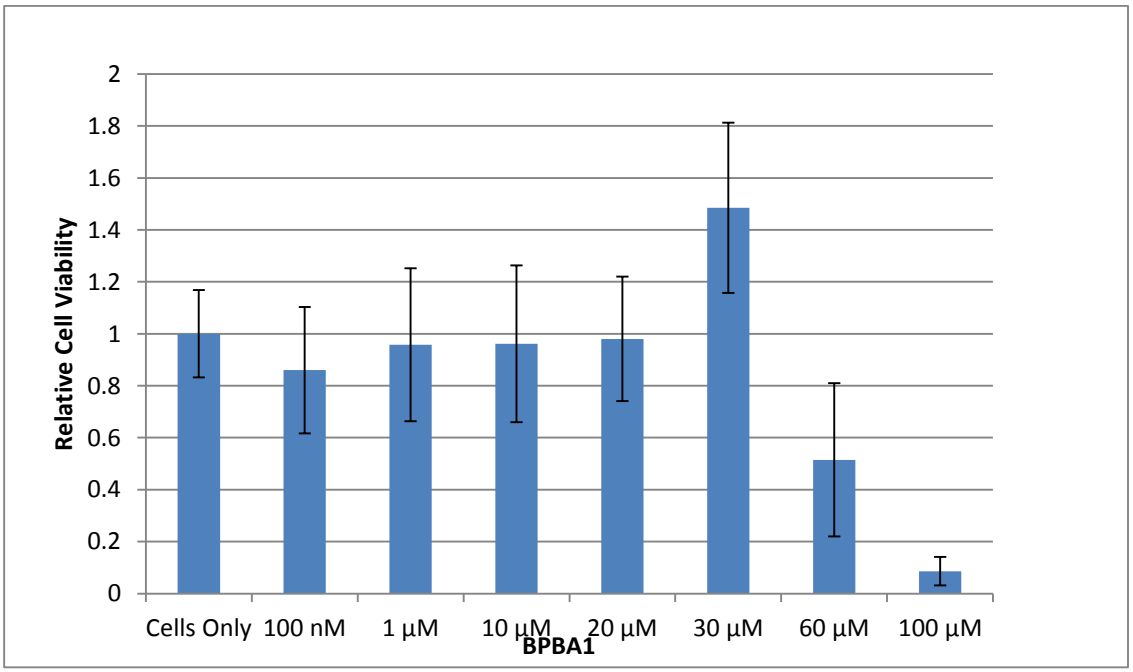


Figure 4.12 MTT toxicity assay of HeLa cells with **BPBA1** at varying concentrations over 24 hours. The relative cell viability data shown is the average of three independent experiments, each run in triplicate.

Taken together, **BPBAs** displayed excellent cell permeability in two cell lines. In addition, these peptides were shown to be non-toxic in concentrations up to 30 μM in both HeLa and A2780 cell lines for 24 hours. It is clear that incorporation of boronic acid in peptides does not affect cell permeability, as well as cell viability. These results encouraged further research on exploring other branched peptides that bind to RNA ligands.

4.3 Conclusion

The unique aspect of our study was the use of boronic acid groups that introduced an alternative mode of binding to a well-folded, therapeutically relevant target RNA. To our knowledge, this is the first instance of boronic acid side chains being utilized in peptidic RNA ligands. In this work, we selected and characterized the peptide **BPBA1** from a 3.3.4-branched peptide boronic acid library. Our results demonstrated that this medium-sized peptide boronic acid selectively binds RRE IIB RNA against competitor tRNAs, six RRE IIB related structural variant RNAs, and an RRE IIB DNA analogue. We also showed that **BPBA1** is cell permeable and poses minimal cytotoxicity in two eukaryotic cell lines. Finally, we provided evidence that **BPBA1** binding to RRE IIB spans a large surface area in the upper stem and internal loops of the RNA. This work demonstrates that complex, multiple intermolecular interactions along a large surface area as a key feature in developing selective binders of unique RNA tertiary structures. These results highlight the potential for this novel class of RNA-binding compounds to be further refined for use in therapeutic development as well as the utility of this method as a general platform for RNA ligand discovery. Current efforts are aimed at improving the binding affinity and selectivity of BPBAs and demonstrating their inhibition activity in cell-based assays. Lessons learned

from this study will inform our ongoing efforts to target highly structured RNAs with high affinity and selectivity.

4.4 Materials and Methods

4.4.1 Peptide Synthesis, Purification and Characterization

Unlabeled and fluorescein-5-isothiocyanate (FITC) (Sigma) labeled peptides were synthesized on Rink amide MBHA resin (100-200 mesh) (Novabiochem). In the preparation of the FITC-labeled peptides, Fmoc-Lys(ivDde)-OH (Novabiochem) was used as the branching unit.[11] Acetic anhydride in DMF (1:1 v/v) with 10 equivalents of DIEA was used to cap the first N-terminus. Then, ivDde was removed by treatment with 2% hydrazine in DMF for 1 hr. The second N-terminus was synthesized through the ε-*N* of the Lys side chain. Fmoc-6-Ahx-OH (AnaSpec) was coupled to the N-terminal amino acid to provide a linker for FITC, to prevent autocleavage of FITC under acidic conditions.[11, 39] All subsequent steps were protected from light. FITC (5 equiv) was reacted with the deprotected N-terminus of the peptides for 6 hr using DIEA (14 equiv). The boronic acid deprotection was performed as previously described.[40] The supernatant was dried under reduced pressure, and the crude peptide was triturated from cold diethyl ether. The peptides were purified by using a Jupiter 4 μm Proteo 90 Å semiprep column (Phenomenex) using a solvent gradient composed of 0.1% TFA in Milli-Q water and HPLC grade acetonitrile. Peptide purity was determined by using a Jupiter 4 μm Proteo 90 Å analytical column (Phenomenex), while peptide identity was confirmed by MALDI-TOF analysis. Unlabeled peptide concentrations were measured in nuclease free water at 280 nm by using their calculated extinction coefficients. FITC-labeled peptide concentrations were monitored at

495 nm by using the extinction coefficient of FITC at 77,000 mol^{-1} cm^{-1} in 100 mM glycine, pH 9.0.

4.4.2 Preparation of ^{32}P-labeled RNA and DNA

Wild-type and mutant RRE-IIB RNAs were transcribed *in vitro* by T7 polymerase with the Ribomax T7 Express System (Promega) by using previously reported techniques.[41] The antisense templates, sense complementary strand (5'-ATGTAATACGACTCACT ATAGG-3') and RRE IIB reverse PCR primer (5'-GGCTGGCCT GTAC-3') were purchased from Integrated DNA Technologies. Antisense templates were used as follows: RRE IIB RNA 5'-GGCTGGCCTGTACCGTCAGCGT CATTGACGCTGCGCCCATACCAGCCCTATAGTGAGTCGTATTACAT-3'; HexaLoop 5'-GGCTGGCCTGTACGTCAGCGTCGGCATTGACGCTGCGCCCATAC CAGCCCTATAGTGAGTCGTATTACAT-3'; Stem A/Loop A Deleted 5'-GGCGTA CCGTCAGCGTCATTGACGCTGCGCCCGCCCTATAGTGAGTCGTATTACAT-3'; Stem B Deleted 5'-GGCTGGCCTGTACCGTCACATTGTGCGCCCATACCAGCCC TATAGTGAGTCGTATTACAT-3'; Loop A(A-G)/B/Bulge A Deleted 5'-GGCTGGCC AGCGTCATTGACGCTGACCAGCCCTATAGTGAGTCGTATTACAT-3'; Loop B/Bulge A Deleted 5'-GGCTGGCCTGC AGCGTCATTGACGC TGCATACCAGCCCT ATAGTGAGTCGTATTACAT-3'; Loop A/B/Bulge A Deleted 5'-GGCTGGCAGC GTCATTGACGCTGCCAGCCCTATAGTG AGTCGTATTACAT-3'. TetraLoop and RRE IIB were both PCR amplified by using HotstarTaq DNA polymerase (Qiagen) followed by a clean-up procedure using a spin column kit (Qiagen). For the 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 hr. After transcription, 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 1x TBE buffer at 4 °C. The sample was desalted by using a Sep-Pak syringe cartridge (Waters Corporation) and lyophilized. The products, along with bacterial tRNA, were dephosphorylated with calf intestinal phosphatase (CIP) in NEBuffer 3 (New England Biolabs) according to manufacturer's protocol. The product was recovered by a standard phenol extraction followed by ethanol precipitation. Purified RNA was stored as a pellet at −80 °C. RRE IIB DNA (5′-GGCTGGTATGGGCGCAGCGTCAATGACGCTGACGG TACAGGCCAGCC-3′) was purchased from Integrated DNA Technologies and stored at −20 ˚C.

HIV-1 RRE IIB RNA as well as the mutant RNA sequences, tRNA, and RRE IIB DNA were labeled at the 5'-end by treating 10 pmol of dephosphorylated RNA / DNA with 20 pmol of [γ-^{32}P] ATP (111 TBq mol^{-1}) and 20 units of T4 polynucleotide kinase in 70 mM Tris•HCl, 10 mM $MgCl_2$, and 5 mM dithiothreitol, pH 7.6. The mixture was incubated at 37 °C for 30 min, and then at rt for 20 min. The kinase was heat-inactivated at 65 °C for 10 min. The product was recovered by ethanol precipitation, and the purity was examined using 12% denaturing PAGE followed by autoradiography.

4.4.3 Dot Blot Assay

Dot blot assays were performed at rt using a Whatman Minifold I 96 well Dot Blot system and Whatman 0.45 μm pore size Protran nitrocellulose membranes. To determine the binding affinities, 0.04 nM radiolabeled RNA was titrated with peptide (0.001–100 μM). First, a solution of 0.08 nM ^{32}P-labeled RNA was refolded in 2x phosphate buffer (20 mM

potassium phosphate, 200 mM KCl, 1 mM $MgCl_2$, 40 mM NaCl, pH 7.0) by heating at 95 °C for 3 min and then slowly cooling at rt for 20 min. Next, 25 μL of the [^{32}P]-RNA solution was added to 25 μL of peptide in nuclease free water and incubated at rt for 4 hr. The 50 μL mixtures were filtered through the nitrocellulose membrane, which was immediately followed by two consecutive 50 μL washes with 1x phosphate buffer. Peptide binding was visualized by autoradiography by using a storage phosphor screen (GE Healthcare) and a Typhoon Trio phosphorimager (GE Healthcare). Densitometry measurements were quantified by using ImageQuant TL (Amersham Biosciences). Binding curves were generated using a four parameter logistic equation with Kaleidagraph (Synergy Software): y = m1+(m2-m1)/(1+10^(log(m3)-x); m1 = 100; m2 = 1; m3 = .000003, where y = percentage of RNA binding, x = log[peptide], m1 = percentage of RNA binding affinity at infinite concentration (nonspecific binding), m2 = percentage of RNA binding affinity at zero concentration, m3 = peptide concentration at 50% binding (K_d). Each experiment was performed in triplicate and error bars represent the standard deviation calculated over three replicates.

4.4.4 Nuclease Protection Assay

RNA was first refolded by heating a solution of 5'-^{32}P-labeled RRE-IIB (10 nM) and excess unlabeled RRE-IIB (200 nM) at 95 °C for 3 min and then snap cooling on ice. The refolded RNA was incubated on ice for 4 hr in a solution containing the BPBAs and buffer composed of 10 mM Tris, pH 7, 100 mM KCl, and 10mM $MgCl_2$. RNase (Ambion) was then added to the solution, which was further incubated on ice for 10 min (0.002 Units RNase V1), or 1 hr (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 run through a 12 % PAGE containing 7.5 M urea and imaged by autoradiography.

4.4.5 Cellular Internalization of Peptides and MTT Toxicity Assay

Cellular uptake of fluorescein-labeled (FITC) branched peptide boronic acid **BPBA1** in HeLa and A2780 cells was determined by modifying a procedure for FITC-labeled branched peptides.[42] Attempts to obtain cell images by using the published method resulted in intense FITC background signal throughout the Prolong Gold antifade mounting media (Invitrogen) and on the surface of the coverslip. Therefore, removal of the mounting media from the procedure was necessary; cells were grown in 35 mm poly-lysine treated glass bottom dishes containing a no. 1.5 coverglass (MatTek) to allow for submersion of fixed cells in PBS for imaging. The dishes also allowed for more vigorous PBS washes after peptide incubation without dislodging large amounts of cells from the coverslip. Bovine serum albumin (BSA) was employed as a blocking agent to help reduce the FITC background on the coverslip.

HeLa cells were grown and plated for peptide internalization by using previously established methods.[11] A2780 cells were grown and plated in RPMI 1640 media supplemented with 10% heat-inactivated fetal bovine serum (FBS), 1% L-glutamine, 100units/mg penicillin, 100 μg/mL streptomycin, and 0.25 μg/mL amphotericin (Invitrogen). Both cell lines were maintained in a 37 °C incubator containing a 5% CO_2 atmosphere and were subcultured once per week. Cell samples were prepared in 35mm poly-lysine treated glass-bottom dishes containing a no. 1.5 coverglass (MatTek) by using the manufacturer's protocol. The dishes were pre-equilibrated with media for 15 min at 37 °C in 5% CO_2 atmosphere. Media was aspirated from the dishes and the cells were plated

at 1.5 x 10^4 cells/dish in a 500 μL media suspension and incubated for 1 hr to allow for initial cell adherence. Additional media (2 mL) was added and the cells were incubated for a total of 48 hr. After the removal of medium and washing of cells with PBS, 600 μL of FITC-labeled branched peptide boronic acid (FBPBA1**,** 1 μM) in Opti-MEM (Invitrogen/Gibco) was added to the dish; cells were incubated with the peptide for 4 hr. Control cell samples were incubated in 600 μL of Opti-MEM for 4 hr. After incubation and following each subsequent step, a 15 min PBS wash was applied to all cell samples at rt. Cells were fixed with 4% (w/v) paraformaldehyde (Acros) in PBS for 15 min, and incubated with 3% (v/v) BSA (New England Biolabs) in PBS for 30 min. Nucleus staining was performed with the addition of 150 μL of 600 nM DAPI to the cell samples for 4 min. Finally, cells were submersed in PBS, and the dishes were sealed with parafilm and stored at 4°C in the dark until imaged. Cells were imaged using a 40X water-immersion objective (N.A. = 1.2) on an LSM 510 confocal system mounted on an Axiovert 100 inverted microscope (Zeiss). For each cell line, identical acquisition parameters were used for both peptide and control samples. Brightness and contrast were adjusted for the image processing of DAPI and DIC channels using AxioVisionLE software (Zeiss).

For the MTT toxicity assays, 96-well plates (Nunc) were incubated with 100 μL of poly-lysine (Sigma) for 24 hr at 4 °C. The poly-lysine was removed and the wells were rinsed ten times with sterile nuclease-free water. HeLa and A2780 cells were plated in the poly-lysine treated wells and analyzed via an MTT toxicity assay by using a previously published procedure.[11]

4.5 References

1. Guan, L.; Disney, M. D., Recent advances in developing small molecules targeting RNA. *ACS Chem. Biol.* **2012,** *7* (1), 73-86.

2. Thomas, J. R.; Hergenrother, P. J., Targeting RNA with small molecules. *Chem. Rev.* **2008,** *108* (4), 1171-1224.

3. Seedhouse, S. J.; Labuda, L. P.; Disney, M. D., The privileged chemical space predictor (PCSP): a computer program that identifies privileged chemical space from screens of modularly assembled chemical libraries. *Bioorg. Med. Chem. Lett.* **2010,** *20* (4), 1338-43.

4. Lee, M. M.; Pushechnikov, A.; Disney, M. D., Rational and modular design of potent ligands targeting the RNA that causes myotonic dystrophy 2. *ACS Chem. Biol.* **2009,** *4* (5), 345-55.

5. Stelzer, A. C.; Frank, A. T.; Kratz, J. D.; Swanson, M. D.; Gonzalez-Hernandez, M. J.; Janghyun, L.; Andricioaei, I.; Markovitz, D. M.; Al-Hashimi, H. M., Discovery of selective bioactive small molecules by targeting an RNA dynamic ensemble. *Nat. Chem. Biol.* **2011,** *7* (8), 553-559.

6. Filikov, A. V.; Mohan, V.; Vickers, T. A.; Griffey, R. H.; Cook, P. D.; Abagyan, R. A.; James, T. L., Identification of ligands for RNA targets via structure-based virtual screening: HIV-1 TAR. *J. Comput. Aided. Mol. Des.* **2000,** *14* (6), 593-610.

7. Aboul-ela, F., Strategies for the design of RNA-binding small molecules. *Future Med. Chem.* **2010,** *2* (1), 93-119.

8. Galicia-Vázquez, G.; Lindqvist, L.; Wang, X.; Harvey, I.; Liu, J.; Pelletier, J., High-throughput assays probing protein–RNA interactions of eukaryotic translation initiation factors. *Anal. Biochem.* **2009,** *384* (1), 180-188.

9. Yen, L.; Magnier, M.; Weissleder, R.; Stockwell, B. R.; Mulligan, R. C., Identification of inhibitors of ribozyme self-cleavage in mammalian cells via high-throughput screening of chemical libraries. *RNA* **2006,** *12* (5), 797-806.

10. Czerwoniec, A.; Dunin-Horkawicz, S.; Purta, E.; Kaminska, K. H.; Kasprzak, J. M.; Bujnicki, J. M.; Grosjean, H.; Rother, K., MODOMICS: a database of RNA modification pathways. 2008 update. *Nucleic Acids Res* **2009,** *37* (Database issue), D118-21.

11. Bryson, D. I.; Zhang, W.; McLendon, P. M.; Reineke, T. M.; Santos, W. L., Toward targeting RNA structure: branched peptides as cell-permeable ligands to TAR RNA. *ACS Chem. Biol.* **2012,** *7* (1), 210-7.

12. Bryson, D. I.; Zhang, W.; Ray, W. K.; Santos, W. L., Screening of a branched peptide library with HIV-1 TAR RNA. *Mol. BioSyst.* **2009,** *5*, 1070-1073.

13. Zhang, W.; Bryson, D. I.; Crumpton, J. B.; Wynn, J.; Santos, W. L., Branched peptide boronic acids (BPBAs): a novel mode of binding towards RNA. *Chem. Commun.* **2013,** *49* (24), 2436-8.

14. Dayton, A. I.; Sodroski, J. G.; Rosen, C. A.; Goh, W. C.; Haseltine, W. A., The trans-activator gene of the human T cell lymphotropic virus type III is required for replication. *Cell* **1986,** *44* (6), 941-7.

15. Hamasaki, K.; Ueno, A., Aminoglycoside antibiotics, neamine and its derivatives as potent inhibitors for the RNA-protein interactions derived from HIV-1 activators. *Bioorg. Med. Chem. Lett.* **2001,** *11* (4), 591-4.

16. Zapp, M. L.; Young, D. W.; Kumar, A.; Singh, R.; Boykin, D. W.; Wilson, W. D.; Green, M. R., Modulation of the RRE–Rev interaction by aromatic heterocyclic compounds. *Bioorg. Med. Chem.* **1997,** *5* (6), 1149-55.

17. Wilson, W. D.; Li, K., Targeting RNA with small molecules. *Curr. Med. Chem.* **2000,** *7* (1), 73-98.

18. Lee, Y.; Hyun, S.; Kim, H. J.; Yu, J., Amphiphilic helical peptides containing two acridine moieties display picomolar affinity toward HIV-1 RRE and TAR. *Angew. Chem. Int. Ed.* **2008,** *47* (1), 134-137.

19. Zapp, M. L.; Young, D. W.; Kumar, A.; Singh, R.; Boykin, D. W.; Wilson, W. D.; Green, M. R., Modulation of the RRE–Rev interaction by aromatic heterocyclic compounds. *Bioorg. Med. Chem.* **1997,** *5* (6), 1149-55.

20. Xiao, G.; Kumar, A.; Li, K.; Rigl, C. T.; Bajic, M.; Davis, T. M.; Boykin, D. W.; Wilson, W. D., Inhibition of the HIV-1 RRE–Rev complex formation by unfused aromatic cations. *Bioorg. Med. Chem.* **2001,** *9* (5), 1097-113.

21. Luedtke, N. W.; Liu, Q.; Tor, Y., RNA-ligand interactions: affinity and specificity of aminoglycoside dimers and acridine conjugates to the HIV-1 Rev response element. *Biochemistry.* **2003,** *42* (39), 11391-403.

22. Li, K.; Davis, T. M.; Bailly, C.; Kumar, A.; Boykin, D. W.; Wilson, W. D., A Heterocyclic inhibitor of the RRE–Rev complex binds to RRE as a dimer. *Biochemistry.* **2001,** *40* (5), 1150-1158.

23. Hart, R. A.; Billaud, J. N.; Choi, S. J.; Phillips, T. R., Effects of 1,8-diaminooctane on the FIV Rev regulatory system. *Virology* **2002,** *304* (1), 97-104.

24. DeJong, E. S.; Chang, C. E.; Gilson, M. K.; Marino, J. P., Proflavine acts as a Rev inhibitor by targeting the high-affinity Rev binding site of the Rev responsive element of HIV-1. *Biochemistry* **2003,** *42* (26), 8035-46.

25. Prater, C. E.; Saleh, A. D.; Wear, M. P.; Miller, P. S., Chimeric RNase H-competent oligonucleotides directed to the HIV-1 Rev response element. *Bioorg. Med. Chem.* **2007,** *15* (16), 5386-95.

26. Legiewicz, M.; Badorrek, C. S.; Turner, K. B.; Fabris, D.; Hamm, T. E.; Rekosh, D.; Hammarskjold, M. L.; Le Grice, S. F., Resistance to RevM10 inhibition reflects a conformational switch in the HIV-1 Rev response element. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** *105* (38), 14365-70.

27. Lee, T. C.; Sullenger, B. A.; Gallardo, H. F.; Ungers, G. E.; Gilboa, E., Overexpression of RRE-derived sequences inhibits HIV-1 replication in CEM cells. *New Biol.* **1992,** *4* (1), 66-74.

28. Symensma, T. L.; Baskerville, S.; Yan, A.; Ellington, A. D., Polyvalent Rev decoys act as artificial Rev-responsive elements. *J. Virol.* **1999,** *73* (5), 4341-9.

29. Giver, L.; Bartel, D. P.; Zapp, M. L.; Green, M. R.; Ellington, A. D., Selection and design of high-affinity RNA ligands for HIV-1 Rev. *Gene* **1993,** *137* (1), 19-24.

30. Chaloin, L.; Smagulova, F.; Hariton-Gazal, E.; Briant, L.; Loyter, A.; Devaux, C., Potent inhibition of HIV-1 replication by backbone cyclic peptides bearing the Rev arginine rich motif. *J Biomed Sci* **2007,** *14* (5), 565-84.

31. Friedler, A.; Friedler, D.; Luedtke, N. W.; Tor, Y.; Loyter, A.; Gilon, C., Development of a functional backbone cyclic mimetic of the HIV-1 Tat arginine-rich motif. *J. Biol. Chem.* **2000,** *275* (31), 23783-9.

32. Mills, N. L.; Daugherty, M. D.; Frankel, A. D.; Guy, R. K., An α-Helical peptidomimetic inhibitor of the HIV-1 Rev–RRE interaction. *J. Am. Chem. Soc.* **2006,** *128* (11), 3496-3497.

33. Crumpton, J. B.; Zhang, W.; Santos, W. L., Facile analysis and sequencing of linear and branched peptide boronic acids by MALDI mass spectrometry. *Anal. Chem.* **2011,** *83* (9), 3548-3554.

34. Kjems, J.; Calnan, B. J.; Frankel, A. D.; Sharp, P. A., Specific binding of a basic peptide from HIV-1 Rev. *Embo J.* **1992,** *11* (3), 1119-29.

35. Kjems, J.; Brown, M.; Chang, D. D.; Sharp, P. A., 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* (3), 683-7.

36. Ippolito, J. A.; Steitz, T. A., The structure of the HIV-1 RRE high affinity rev binding site at 1.6 Å resolution. *J. Mol. Biol.* **2000,** *295* (4), 711-717.

37. Bartel, D. P.; Zapp, M. L.; Green, M. R.; Szostak, J. W., HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* **1991,** *67* (3), 529-36.

38. Park, S. H.; Doh, J.; Park, S. I.; Lim, J. Y.; Kim, S. M.; Youn, J. I.; Jin, H. T.; Seo, S. H.; Song, M. Y.; Sung, S. Y.; Kim, M.; Hwang, S. J.; Choi, J. M.; Lee, S. K.; Lee, H. Y.; Lim, C. L.; Chung, Y. J.; Yang, D.; Kim, H. N.; Lee, Z. H.; Choi, K. Y.; Jeun, S. S.; Sung, Y. C., Branched oligomerization of cell-permeable peptides markedly enhances the transduction efficiency of adenovirus into mesenchymal stem cells. *Gene Ther.* **2010,** *17* (8), 1052-1061.

39. Jullian, M.; Hernandez, A.; Maurras, A.; Puget, K.; Amblard, M.; Martinez, J.; Subra, G., N-terminus FITC labeling of peptides on solid support: the truth behind the spacer. *Tetrahedron Lett.* **2009,** *50* (3), 260-263.

40. Coutts, S. J.; Adams, J.; Krolikowski, D.; Snow, R. J., Two efficient methods for the cleavage of pinanediol boronate esters yielding the free boronic acids. *Tetrahedron Lett.* **1994,** *35* (29), 5109-5112.

41. Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C., Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucl. Acids Res.* **1987,** *15* (21), 8783-8798.

42. Bryson, D. I.; Zhang, W.; McLendon, P. M.; Reineke, T. T.; Santos, W. L., Cell permeable branched peptide RNA ligands to disrupt RNA-protein interactions. *ACS Chem. Biol.* **2011**.

Chapter 5. Targeting HIV-1 RRE with Branched Peptides Featuring Unnatural Amino Acids

Attribution

This chapter is adapted from a manuscript that is currently in preparation for submission to a peer-reviewed journal. The author of this dissertation performed a major portion of the work described in this chapter. She designed and synthesized the peptide library, performed the high throughput screening, synthesized the majority of the hit peptides for biophysical characterization, performed a significant amount of electrophoretic mobility shift assay and fluorescence spectroscopy. She was also responsible for synthesis of the unnatural amino acids, binding data interpretation, RNA synthesis, and hit peptide deconvolution. Jessica Wynn was responsible for RNA footprint assays and MALDI analysis of peptides. She also aided in the synthesis of several peptides. Dr. Neeraj Patwardhan aided in the synthesis of unnatural amino acids. The author also significantly contributed to the writing and editing of the manuscript.

Abstract

The Rev Response Element (RRE) within the *env* gene of the HIV-1 RNA genome is a highly conserved region that is recognized by the HIV-1 Rev protein and activates nuclear export of unspliced and singly spliced HIV-1 mRNA to the cytoplasm. The RRE/Rev export pathway is an attractive target for drug design due to its critical role in allowing the translocation of HIV-1 late transcripts, translation of viral structure proteins, and proliferation of the virus. Early studies have identified the stem-loop structure, known as RRE stem IIB (RRE IIB), as the high affinity binding site of the Rev protein. Herein,

we report a branched peptide library comprised of 4,096 unique sequences featuring boronic acid and acridine moieties. We reveal that four of the hit sequences containing acridine alone or in conjunction with boronic acid improved binding affinity significantly to less than 100 nM. These results indicate that acridine is beneficial for boosting binding affinity towards RRE RNA.

5.1 Introduction

The human immunodeficiency virus type 1 (HIV-1) continues to be a global issue despite significant advancements in research, treatment, and prevention efforts. An estimated 2.7 million new cases of infection occurred in 2013 bringing the total number of people living with HIV-1 to 35.3 million according to the World Health Organization.[1] HIV-1 remains a problem today due in part to the emergence of drug-resistant viral strains driven by selective pressure. Indeed, the high rate of mutation among retroviruses allows rapid evolution of the drug target that enables it to escape purview of the medication being administered.[2-5] Combination therapies are requisite for effective treatment of HIV-1, yet discovery of new drug combinations that reduce the development of drug resistance remains a formidable challenge. The majority of drugs approved for the treatment of HIV-1 target are either reverse transcriptase or HIV protease inhibitors. However, fusion inhibitors, entry inhibitors, and integrase inhibitors have also been approved.[6] Thus, the variety of HIV-1 targets being exploited presently is expanding. Unfortunately, each of the FDA approved drugs is vulnerable to mutations of HIV-1 as they function through inhibition of a specific enzyme, receptor, or co-receptor.[6] Therefore, there remains a great need to continue the search for the alternative therapies that target other essential viral activities.

In recent years, RNAs have gained significant attention because they can fold into well-defined secondary and tertiary structures, providing interfaces for specific RNA–RNA or RNA–protein interactions, which are essential for transcription, translation, splicing, replication, transport and catalysis in a living cell.[7-11] These discoveries greatly increased the interest in RNA as a potential drug target for treatment of diseases. One of the most studied RNA-based regulatory systems is the HIV-1 Rev response element (RRE)/Rev export pathway. Different from most retrovirus, HIV-1 uses a unique pathway to evade the well-controlled and tightly coupled processes of mRNA splicing and nuclear export in eukaryotes. This alternative pathway involves interaction between the virally encoded Rev protein and an extremely well-conserved sequence of RRE RNA, which spans ~240-nucleotides located in the *env* gene of all singly spliced and unspliced HIV-1 transcripts.[12] The RRE–Rev interaction is essential for these RNA transcripts to translocate from the nucleus without first being fully spliced; therefore, this interaction is necessary for the synthesis of viral proteins that are vital to the HIV-1 life cycle such as Gag, Pol, and Env, as well as the accumulation of HIV-1 genomic RNA in the cytoplasm.[12-14] In the absence of this process, the structural proteins necessary for viral particle construction are not produced because the intron containing mRNA transcripts are sequestered in the nucleus.[15] Thus, the life cycle of HIV-1 cannot continue.

The Rev/RRE export pathway is mediated by constant shuttling of Rev in and out of the nucleus, and although many key steps have been extensively investigated, the process has not yet been fully elucidated.[16] Initially, a HIV-1 encoded Rev protein consisting of 116-amino acid residues is imported into the nucleus by binding to multiple importin proteins by using the nuclear localization signal located at the N-terminus of

Rev.[17-19] Once inside the nucleus, Rev binds cooperatively to RRE with high affinity in monomeric increments to form a multimeric complex that ultimately involves up to six Rev monomers.[12, 20-23] The stem-loop IIB of RRE (RRE IIB) has been recognized as the high affinity site, where Rev initially binds.[22-27] The resulting RRE–Rev ribonucleoprotein complex is recognized by the host Crm1 and Ran-GTP, which directs the complex to the nuclear pore and initiates export of the complex.[28-30]

The Rev/RRE export pathway has become a high profile drug target for its critical role in the proliferation of HIV-1.[15] It's important to note that the RRE–Rev interactions are independent from the cellular process, although the cellular factors may have influences in RRE–Rev bindings in the cell.[31-32] Therefore, specific targeting the RRE–Rev interactions by a therapeutic drug may lead to inhibit the viral replication with minimal risk of side effect. Additionally, targeting RRE may provide an important opportunity to develop drugs that do not induce drug resistance in the virus because of a pressure to preserve this element both from an RNA-protein binding standpoint as well as from the overlap with the open reading frame of the *env* gene.[33] Owing to the therapeutic potential of the Rev/RRE export pathway, a variety of small molecules, like aminoglycosides have been designed to interrupt the RRE–Rev interaction with limited clinical success.[34-37] Other inhibitors such as aromatic heterocycles, antisense oligonucleotides, RRE-based decoys, cyclic peptides, α-helical peptidomimetics, and others have also been identified.[38-52] However none of these strategies have found clinical success.

The traditional strategy of targeting RNAs is to design molecules that either binding to RNA with single binding mode or one simple moiety. While when only a limited number of interfaces presented for recognition and interaction, multivalent ligands are preferred to

enhance binding affinity and specificity to their targets. Therefore, conjugating ligands that have affinities for different motifs of RNA (i.e. loop, stem and bulge) or have different binding modes against the same RNA target (i.e. electrostatic interaction, hydrogen bonding and intercalation) is attractive. Indeed, acridine has been conjugated to a variety of ligands, and their bioactivities have been extensively studied. For example, one compound, CGP40336A, which features 6-chloro-2-methoxyacridine and spermidine, proved to be superior at inhibiting the Tat–TAR interaction through π-stacking and electrostatic interaction (Figure 5.1A).[53] An acridine-aminoglycoside conjugate, termed acridine-N-neomycin, and an amphiphilic helical peptide bearing two acridine moieties through the N^{ε} atom of the Lys amino acids, have been demonstrated to bind to RRE with a binding affinity of 2.4 nM and 610 pM, respectively, which is much tighter than the non-

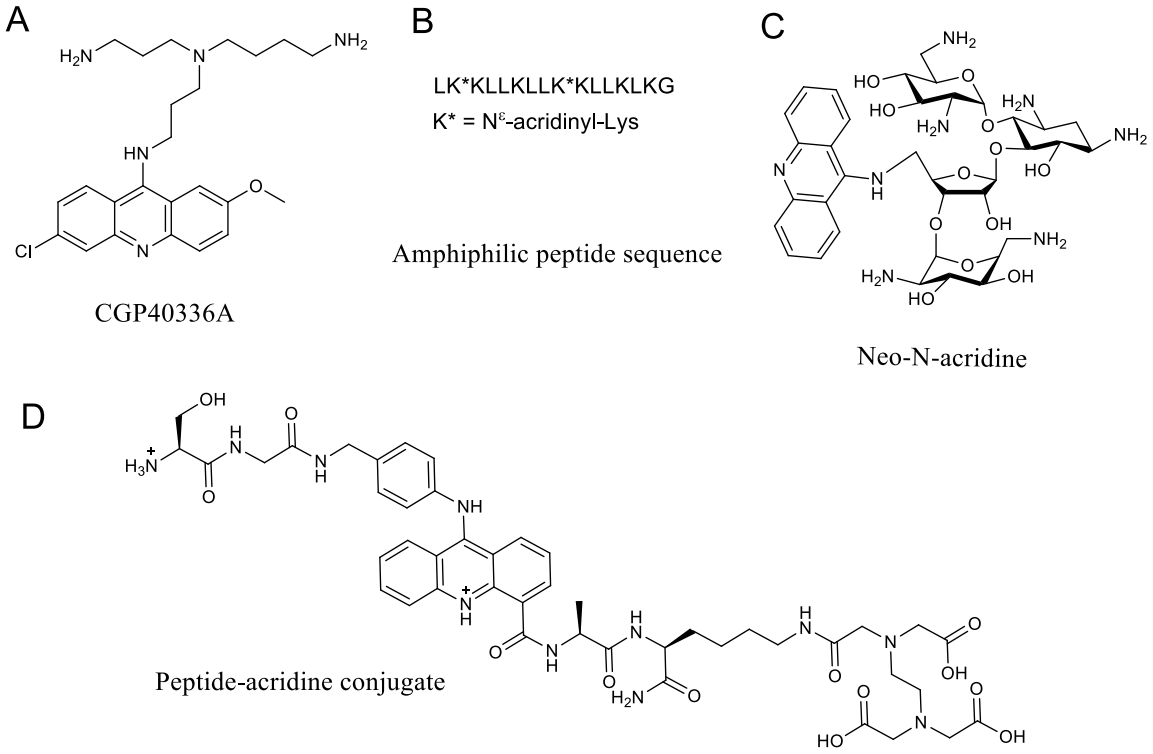


Figure 5.1 Structures of (A) CGP40336A; (B) amphiphilic peptide; (C) Neo-N-acridine; (D) Peptide-acridine conjugate.

acridine containing counterparts (Figure 5.1B-C).[54-55] Further CD experiments have shown that incorporation of two acridine moieties in proper positions do not disrupt α-helicity of the sequence. [55] In addition, a peptide–acridine conjugate has been proven to bind to a duplex RNA, where acridine is shown to interact via a threading intercalation process (Figure 5.1D).[56] Particularly, acridine has been selected to be a potential pharmacophore both for in vitro screening and virtual screening against RNA targets, primarily because of its potential to intercalate into RNA.[44, 57]

In continuation of our effort towards developing molecules that target the tertiary structure of RNA instead of its primary sequence, we focused our attention on RRE and envisioned utilizing RNA–ligand interactions that are outside the typical canonical modes of binding. As discussed in chapter 3, we identified a novel cell-permeable and nontoxic branched peptide boronic acid (**BPBA1**) from an on-bead high-throughput screening of a 3.3.4-library that selectively bind to RRE IIB in the presence of various competitors. We demonstrated that both the boronic acid moieties and the "branching" in these peptides were key structural motifs for multivalent interactions with the target RNA.[58-59] In this study, we performed a high throughput screening of a simple 3.3.4-library comprised of 4,096 unique sequences that feature both boronic acid and acridine moieties. We reveal that four of the hit sequences containing acridine alone or in conjunction with boronic acid significantly improved the binding affinity to less than 100 nM.

5.2 Results and Discussion

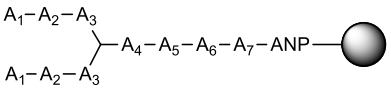
5.2.1 Library Design and On-Bead High-Throughput Screening

While designing our third generation of branched peptide (BP) library, we embarked on a strategy to improve the selectivity and binding affinity to the RNA target

through the incorporation of unnatural amino acid side chains featuring the acridine functional group (Figure 5.2) in addition to the boronic acid moiety. We hypothesized that incorporation of the acridine intercalator into the branched peptides might result in stronger π-π interactions with bases in the target RNA.[55] This strong interaction may direct the whole peptide to get close to the target RNA, therefore, both acridine and peptide side chain could position properly to allow tight binding and specific recognition for the RNA target. Furthermore, we utilized Fmoc-*N*-ε-(4-boronobenzoyl)-L-lysine (K_{BBA}) (Figure 5.2) as boronic acid containing amino acid because the consensus sequence analysis from our previous library revealed that the hit peptides had a preference for the K_{BBA} moiety.[58] We expected that the empty p-orbital of boron could accept electrons from RNA to form reversible covalent bonds, therefore, increased both the selectivity and the affinity for RNA targets. In addition, the amino acid containing boronic acid presented various binding modes (i.e. electrostatic, π-stacking, and covalent binding) that could increase the complexity of the library. Finally, the incorporation of unnatural side chains, as well as the unnatural branch structure of the peptide, would likely impart resistance to enzymatic cleavage.

The BP library was synthesized in triplicate on Tentagel beads by split and pool synthesis. The library was prepared such that there were three variable amino acid positions at both the N- and C-termini (A_1–A_3 and A_4–A_6, respectively), and each variable position was composed of four side chains (Figure 5.2). Each of the four possible side chains at variable positions A_1–A_6 was chosen for its potential to interact with the RRE IIB target RNA. Different from previous library, we only chose four amino acids to represent all possible interactions for the whole library. For example, from position A_1–A_6, we used

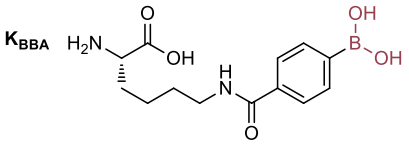
Fmoc-*N*-ε-(9-aminoacridinyl)-L-lysine (K_{ACR}) for interactions with the RNA through π stacking, Lys for electrostatic attraction and hydrogen bonding, Leu for hydrophobic interaction, and K_{BBA} for and reversible covalent bonding in each position. The goal was to present varying binding modes to the RNA target through four simple amino acids. Most importantly, we investigated whether intercalating agents can impart potency in branched peptides. Further modification could be achieved by replacing amino acid in each position with various other amino acids that have similar properties.



3.3.4. Library (4 variable positions)

A_1	A_2	A_3	A_4	A_5	A_6	A_7
Leu	Leu	Leu	Leu	Leu	Leu	Tyr
K_{ACR}	K_{ACR}	K_{ACR}	K_{ACR}	K_{ACR}	K_{ACR}	Tyr
Lys	Lys	Lys	Lys	Lys	Lys	Tyr
K_{BBA}	K_{BBA}	K_{BBA}	K_{BBA}	K_{BBA}	K_{BBA}	Tyr

4^6 = 4,096 library members



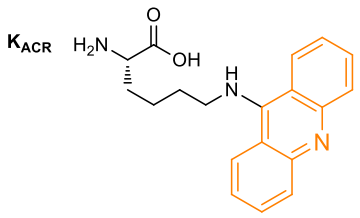


Figure 5.2 Branched peptide library and structure of K_{BBA} and K_{ACR}.

The BP library was subjected to on-bead high throughput screen[60] against DY547 labeled HIV-1 RRE IIB RNA. During the screening process, the beads were first pretreated with bovine serum albumin and competitor tRNA to minimize non-specific binding. Specific binding of the target RNA to peptide resulted in increased fluorescence of the bead, which was monitored by both fluorescence microscopy and confocal microscopy. Twenty beads were selected as possible hits. These beads were photocleaved via UV irradiation and then sequenced by MALDI MS-MS analysis.[61] Hit compounds were resynthesized for further biophysical characterizations.

5.2.2 Sequence and Binding Affinity of BPs Toward HIV-1 RRE IIB

Among the twenty beads that were analyzed, as possible hits, fifteen sequences were determined. The sequence of four of the beads could not be revealed by MALDI-TOF, possibly due to quality of these beads. Three sequences were found in duplicate, which increased confidence on the quality of the library, as well as screening process. During MALDI-TOF analysis, we noticed that the hit compounds containing acridine moiety were not ionizing very well; therefore, one or two MS-MS peaks for sequence identification were missing in the MS spectra. As a result, we chose the sequences that were reliably and confidently sequenced for further biological assay. Following the standard electrophoretic mobility shift assay (EMSA), ^{32}P-labeled RRE IIB was titrated with increasing concentration of BPs. The sequence and dissociation constant of the hit BPs were shown in Table 5.1. The hits contained either 1–3 boronic acid residues or 1–2 acridine residues in addition to both boronic acid and acridine residues. In particular, the peptides containing acridine residues significantly decreased the dissociation constants to lower than 100 nM. In fact, this number is lower than the dissociation constants obtained from the EMSA analysis of Rev17 (Suc-TRQARRNRRRRWRERQRAAAAR-am), the minimum binding elements of Rev protein, with wild-type RRE IIB RNA.[52] Especially, peptides **A9** and **A10** displayed a K_d of 20 nM and 30 nM, respectively, which were the best two sequences in the library. Indeed, peptides **A2** and **A10** only had a slightly difference in sequence order (K_{ACR}LK to LKK_{ACR}); however, **A10** displayed a 7-fold tighter binding affinity for RRE IIB than **A2**. Further, peptides **A2** and **A9** are only different in one position (L to K_{ACR} at position A_5), **A9** bound 5-fold tighter to RRE IIB than **A2**. This may be due to the presence of the second acridine in the sequence, providing extra π-stacking interactions. Consensus

sequence analysis revealed that acridine residues only present at position A_1, A_2, A_4, and A_5. While it is difficult to draw conclusion for the positional preference of acridine moieties because only five sequences out of fifteen contained acridine, it appears that there is no correlation between the number of acridine moieties and the resulting binding affinities. Further sequence analysis suggested that K_{BBA} was heavily preferred at position A_5 and A_6 of the C-terminus. This distribution may suggest that the C-terminus branch could get close to RNA sugar rings or bases where boron acts as a Lewis acid, providing non-canonical, and improved binding interaction with the RNA target. No boron containing amino acids were found at positions A_1 or A_4 in any hit sequence and the boronic acid moiety was rarely preferred at the N-terminus.

Further sequence analysis indicated that Lys was present at most positions (A_1 to A_4), especially at N-terminus. A preference for positively charged side chains, however, is not unexpected because the Rev protein binds RRE IIB through the polyarginine nuclear localization signal (NLS) located in its N-terminal region.[23, 27, 62-63] The positively charged Lys side chains in our hits likely function similarly to the Arg residues of the Rev NLS and may provide the necessary electrostatic attraction with the negatively charged phosphate backbone of RRE IIB. However, it is noteworthy that the number of Lys residues in the hit compounds did not correlate to increasing binding affinity—*i.e*, an increase in the net positive charge did not result in increased binding affinity. For example, **A14** with six Lys residues had a >7 fold higher K_d value (1,430 vs 210 nM) compared to **A3** with four Lys residues. Furthermore, the peptides containing five Lys possessed the binding affinities ranging from 70 nM to 530 nM. Indeed, we have observed this phenomenon in previous generation of branched peptides designed to bind HIV-1 TAR RNA and HIV-1 RRE IIB

Table 5.1 Dissociation constant and molecular weight of hit compounds.

Entry	Peptide	Sequence[a]	K_d(nM)	MW(g mol^{-1})
1	A1	$(KKK)_2$*$LK_{BBA}K_{BBA}Y$	280 ± 80	1742.7
2	A2	$(KKK)_2$*$K_{ACR}LKY$	140 ± 50	1624.3
3	A3	$(LLK)_2$*$KK_{BBA}KY$	210 ± 30	1550.8
4	A4	$(KKK)_2$*$LK_{BBA}LY$	130 ± 30	1579.8
5	A5	$(KLK)_2$*$LKK_{BBA}Y$	300 ± 40	1564.8
6	A6	$(KKL)_2$*$LKK_{BBA}Y$	330 ± 100	1564.8
7	A7	$(K_{ACR}KK)_2$*$KLK_{BBA}Y$	60 ± 10	1949.2
8	A8	$(KK_{BBA}K)_2$*LLKY	530 ± 180	1727.4
9	A9	$(KKK)_2$*$K_{ACR}K_{ACR}KY$	30 ± 4	1816.3
10	A10	$(KKK)_2$*$KLK_{ACR}Y$	20 ± 4	1624.1
11	A11	$(KKK_{BBA})_2$*$KLK_{BBA}Y$	120 ± 40	1890.7
12	A12	$(KKL)_2$*$KK_{BBA}K_{BBA}Y$	170 ± 50	1727.7
13	A13	$(KKK)_2$*$KK_{BBA}LY$	ND	1594.9
14	A14	$(KLK)_2$*$KKK_{BBA}Y$	1430 ± 870	1579.8
15	A15	$(KK_{ACR}K)_2$*$KLK_{BBA}Y$	70 ± 30	1949.2

[a] * = Lysine branching unit. Each value is an average of at least three experiments.

RNA, where higher numbers of Arg/Lys side chains did not generally result in increased binding affinity—an alternative mode of binding was suspected as compensation for the loss of electrostatic interactions.[64-65] Taken together, the acridine moieties contribute significantly to the low dissociation constants of the hit peptides for RRE IIB. Finally, aliphatic residues were the least preferred side chains in our pool of hit compounds. This suggests that modes of binding other than hydrophobic interactions are generally more beneficial for strong binding between these BPBAs and RRE IIB.

5.2.3 Determination of Binding Affinity of BPs Toward HIV-1 RRE IIB via Fluorescence Spectroscopy

While we examined the dissociation constants obtained from the EMSA, we noticed that some peptides processed a huge error bar—about 60% of the calculated value. This was found mainly between the peptides without acridine moieties. Because our peptides were positively charged, they were hard to get into the gel during EMSA. As a result, it

was difficult to calculate the dissociation constant with partial bands disappeared on the gel. To confirm that the dissociation constants determined from EMSA were reliable, a 2-aminopurine (2-AP) based fluorescence spectroscopy was performed with five peptides.

Previous studies have shown that incorporation of 2-AP in the RNA sequence could probe the structure and dynamics of the specific sites.[66] The fluorescence of 2-AP is usually highly quenched when it is stacked with other bases, but increases significantly when it is exposed to solvent.[67] The change of fluorescence is highly sensitive to the microenvironment, which allows the detection on the conformational change when bound to ligand. Furthermore, because the structure similarity of 2-AP with adenine, it is able to form a thermodynamically equivalent base pair with uridine.[67] As a result, the substitution of bases with 2-AP usually will not disrupt the original RNA conformation. Earlier studies of RRE–Rev interaction have revealed that one base, U36 in the internal loop region, was flipped-out from the RNA helix upon Rev peptide binding.[27] Mutagenesis studies have indicated that U36 could be substituted without disrupting Rev peptide binding.[49, 68] Therefore, U36 is an ideal base for positioning a 2-AP probe to monitor the peptide–RNA interactions.[66, 69]

The binding affinities of BPs to HIV-1 RRE IIB was determined by titrating 2-AP labeled RRE IIB RNA with increasing concentration of BPs. Initially, without adding any compounds, the 2-AP labeled RRE IIB RNA displayed the maximum fluorescent intensity, which suggested that 36AP base was less stacked in the absent of ligands. Upon adding peptide, a decrease in fluorescence intensity was observed as a function of increasing concentration of peptides. The results that the environment around 36AP base changed as

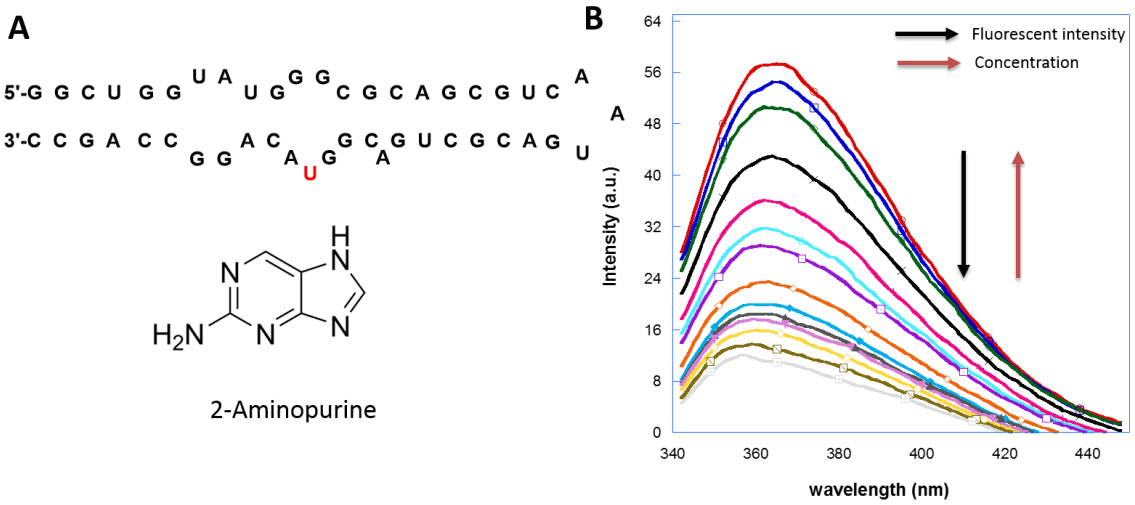


Figure 5.3 (A) Structure of 2-aminopurine and RRE IIB RNA, red indicates U36 is replaced by 2-AP; (B) Fluorescent intensity decreased as the concentration of **A3** increased.

a consequence of stacking or a more hydrophobic environments (Figure 5.3). This observation is quite different from previous research such that binding of small molecules with RRE IIB usually exposed 36AP base more toward the solvent.[66] These results suggested that our BPs interact with RRE IIB differently compared to the small molecules like neomycin. The dissociation constants determined through fluorescence spectroscopy were comparable to the results obtained via EMSA for acridine containing peptides (Table 5.2). However, the dissociation constants of non-acridine containing peptides determined through fluorescence spectroscopy was quite different compared to the result obtained via EMSA. When we compared to the EMSA result of acridine containing peptides with

Table 5.2 Dissociation constants obtained from EMSA and FS.

Compounds	K_d (nM)/EMSA	K_d (nM)/FS
A2	140 ± 50	150 ± 30
A3	210 ± 30	870 ± 20
A7	60± 10	14 ± 1
A10	20± 4	10 ± 8
A15	70 ± 30	120 ± 30

Each value is an average of at least three experiments.

non-aciridine containing peptides, we found that acridine containing peptide traveled into the gel much better than non-acridine containing peptides, which likely resulted in calculation error (Figure 5.4). Job's plot analysis of the fluorescence titration spectra of **A3** with 2-AP labeled RRE IIB indicate a 1:1 stoichiometry-a 1:1 complex between **A3** and RRE IIB RNA (Figure 5.5).

Figure 5.4 EMSA image of acridine containing peptide (**A7**) and non-acirdine containing peptide (**A3**).

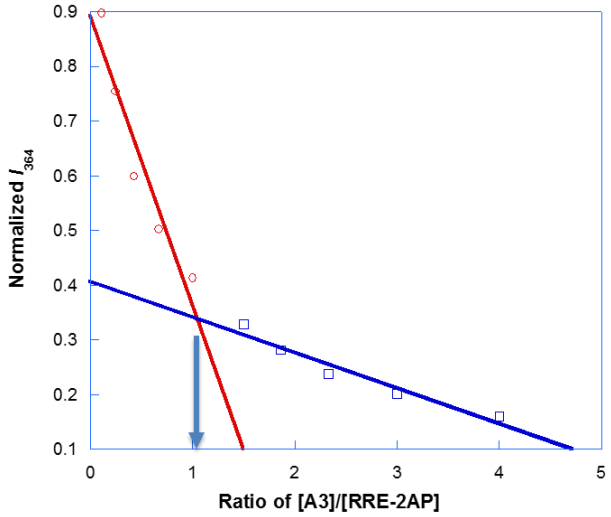


Figure 5.5 Job's plot of **A3** indicated a 1:1 stoichiometry.

5.2.4 Selectivity of BPs Toward RRE IIB Tertiary Structure

In our previous studies, several mutants were designed to determine whether BPs could discriminate between the native tertiary structure of RRE IIB and related analogs wherein structural elements are missing. We found that when loops A-B and bulge A were eliminated to produce loop A/B/bulge A deleted RNA as a stem loop with significantly

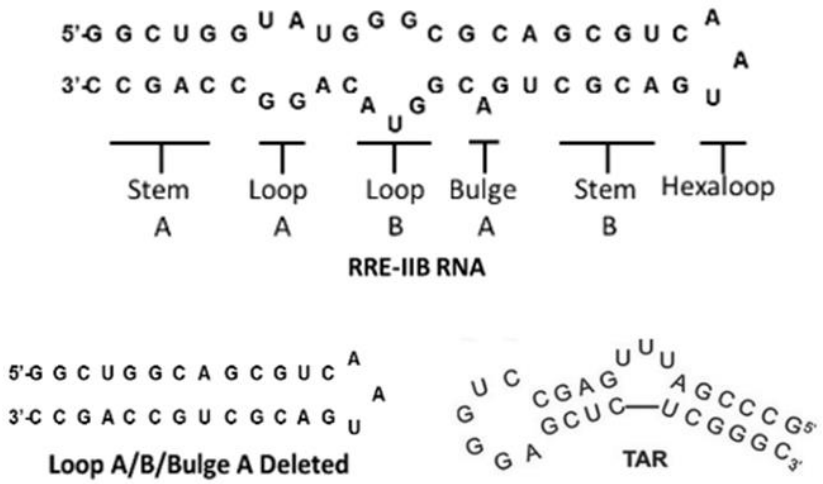


Figure 5.6 Structures of mutant RRE IIB and TAR.

different tertiary structure, a dramatic loss (> 75 fold) in binding affinity (Figure 5.6). Therefore, we evaluated whether BPs containing acridine and boronic acid moieties could discrminate between RRE IIB and the most different structure of RRE IIB variants. The measured K_d of **A3** and **A7** for this variant using EMSA were 190 nM and 70 nM, respectively, which were comparable to the K_ds for the wild type RRE IIB. This suggested that our BPs could not recognize RRE IIB specifically, possibly due to the rich of positive charges that mainly contact RNA phosphate back bone via electrostatic interaction. In addition, we also evaluated the binding affinities of BPs to another well studied viral RNA target, HIV-1 TAR RNA. TAR RNA contains a bulge region and a hexa-loop region that has a simple structure compared to RRE IIB. The measured K_d of **A3** and **A7** for TAR by using EMSA were 170 nM and 80 nM, respectively, which were the same compared to the K_ds for wild type RRE IIB. It is not surprising that BPs can't discriminate between hairpin RNAs because both TAR and RRE IIB shared similar secondary structures such as stem, bulge and loops. Furthermore, the acridine moiety could intercalate with bases tight enough that it is hard to discriminate between hairpin RNAs. Indeed, this observation has been

previously reported that bases in the bulges and kinks of a hairpin RNA were significantly involved in interactions with acridine moieties.[55]

To further characterize the selectivity against other RNA structures, binding affinities between BPs and RRE IIB were measured by EMSA in the presence of excess bacterial tRNA. Initially, a 1000-fold molar excess of competing tRNA was included during incubation with the **A7** and RRE IIB, which resulted in an observed K_d (1.0 μM) that was 16-fold higher than the K_d obtained in the absence of tRNA. This was also observed for other BPs, suggesting BPs were partially selective in the presence of competitor tRNAs. Next we explored the selectivity of **A3** against DNA composed of the same RRE IIB sequence by dot blot assay. Unfortunately, the measured K_d of **A3** was 170 nM, which was the same within experimental error with the K_d of wild type RRE IIB. The reason for BPs lacking selectivity to RRE IIB is unclear, further experiments are needed to make a definitive conducsion.

5.2.5 Determination of Binding Site of BPBA1 by RNase Protection Assay

To better determine regions of specific nucleotide contacts with RRE IIB, we performed ribonuclease protection assays. Hence, 5'-^{32}P-labeled RRE IIB was incubated with varying concentrations (up to 20 μM) of **A7** in the presence of RNase T1, RNase A, or RNase V1 (Figure 5.7). Potential binding sites were confirmed based on the ability of the peptide to protect the RNA from enzymatic cleavage. While there were no noticeable changes in band intensity with RNase T1, the most prominent cleavage bands over all the experiments conducted were observed at G17 and C18 using RNase V1, which preferentially cleaves double stranded regions of RNA. As shown in Figure 5.7, there is a concentration dependent protection (0.2 μM to 20 μM) of this region in the presence of **A7**,

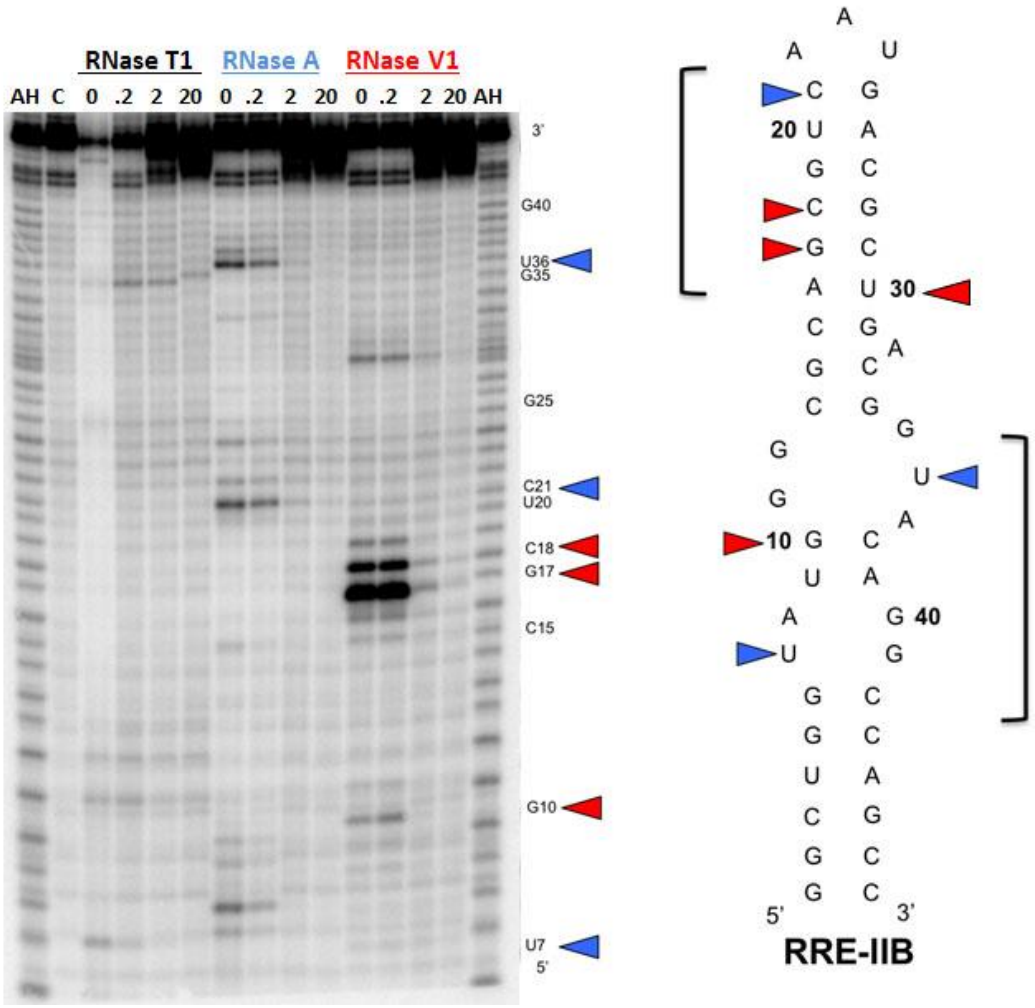


Figure 5.7 RNase protection assay of RRE IIB. The gel depicts the autoradiogram of alkaline hydrolysis (AH) and RNase protection experiments using RNases T1, A, and V1 with increasing concentration of **A7**. Colored triangles highlight bases protected from cleavage by RNase A (blue) and RNase V1 (red).

which suggests that this portion of RRE IIB is a potential binding site. Indeed, this region is also the site wherein the native protein binding partner Rev binds.[70-71] **A7** also visibly protected U7, C21, and U36 from cleavage by RNase A, which hydrolyzes at the 3' side of unpaired pyrimidine bases. In particular, U7 and U36 are potential contact points located in the internal loop region of RRE IIB. Indeed, the same protection pattern was observed with our previous peptide **BPBA1**, suggesting that the interaction of peptide with RRE IIB is highly structure dependent because they shared the same branch peptide structure.[59]

Taken together, the RNA footprinting assay indicated that the binding site for **A7** encompasses a large surface area constituting stem B and loops A and B. Job's plot confirmed a 1:1 binding stoichiometry between **A7** and RRE IIB, and may suggest that **A7** is bound in a folded groove saddled between the internal loop regions and upper stem bases of RRE IIB, as this compound is clearly interacting with these structural elements simultaneously. Although the precise nature of the RNA–peptide interaction is currently unknown, protection from enzymatic cleavage of these regions could result either from steric blockade or remodeling of its tertiary structure.

5.3 Conclusion

In conclusion, the third generation 3.3.4-branched peptide library featuring boronic acid and acridine moieties yielded a series of non-toxic peptides with the dissociation constant less than 100 nM towards HIV-1 RRE IIB RNA. The peptides we tested exhibited no selectivity towards hairpin RNAs and tRNA, which might be the result of strong intercalation of acridine moiety to stem like structures, coupled to the relatively high density of positive charges presented in the peptides. Furthermore, the library was only composed of four amino acids (Lys, Leu, K_{BBA} and K_{ACR}), hence, the library had limited diversity that might account for the low selectivity. RNA-footprinting studies revealed that the BPs binding site encompasses a large surface area that spans the upper stem as well as the internal loop regions of RRE IIB. These results indicated that acridine is beneficial for boosting binding affinity towards RRE RNA. This highlight the potential for the RNA-binding compounds containing acridine and boronic acid groups to be further refined for use in therapeutic development. Further efforts are aimed at improving the binding affinity and selectivity of BPs by increasing the library diversity. Lessons learned from this study

will inform our ongoing efforts to target highly structured RNAs with high affinity and selectivity.

5.4 Materials and Methods

5.4.1 Synthesis of Branched Peptide Library

We used standard solid phase peptide synthesis techniques to generate the 3.3.4 library via the split and pool method using the previously described procedue (Figure 5.8).[65] *N*-α-Fmoc protected L-amino acids (Novabiochem), PyOxim (Novabiochem) and *N*,*N*-Diisopropylethylamine (DIEA, Aldrich) were used in coupling reactions. The synthesis and full characterization of Fmoc-*N*-ε-(4-boronobenzoyl)-L-lysine (K_{BBA}) will be described elsewhere. Fmoc-*N*-ε-(9-aminoacridinyl)-L-lysine (K_{ACR}) and Fmoc-ANP-OH were synthesized as previously reported.[55, 72] Three copies of library were prepared simultaneously by using a three-fold excess of Tentagel Macrobead-NH_2 resin (0.19 g, 0.05

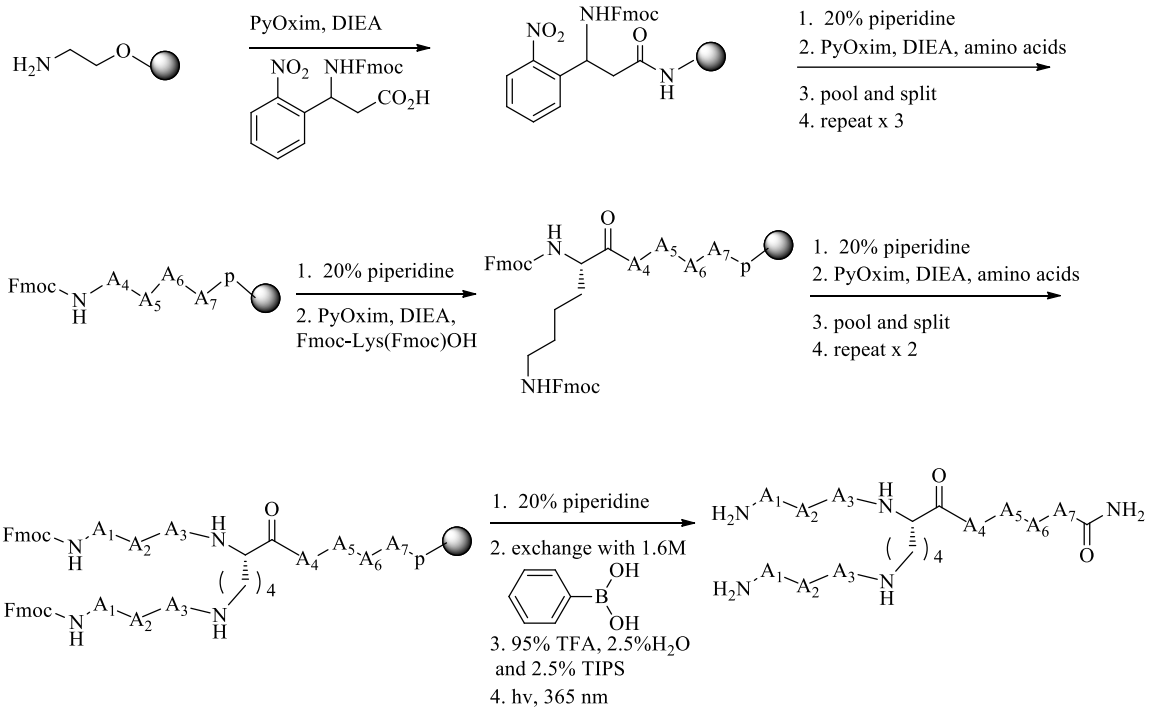


Figure 5.8 Synthesis of branched peptide library.

mmol/g) (Peptides International). The resin was swollen in DCM (20 mL, 2 x 15 min) followed by DMF (20 mL, 15 min). The photocleavable linker Fmoc-ANP-OH (58 mg, 0.14 mmol) was first coupled to the resin in DMF for 3 hr in the presence of PyOxim (70 mg, 0.14 mmol) and DIEA (47 μL, 0.27 mmol). After coupling, the resin was washed with DMF (20 mL, 1 min), DCM (20 mL, 1 min) and DMF (20 mL, 1 min). The same washing procedure was applied after every step. Then, 20% piperidine in DMF (20 mL, 2 x 10 min) was used for Fmoc deprotection. A Kaiser test was used after each coupling and deprotection step to confirm reaction completion. *N*-Fmoc amino acids (3 equiv), PyOxim (3 equiv), and DIEA (5 equiv) were added to each reaction vessels in DMF and coupled for 30 min. Fmoc-Lys(Fmoc)-OH was used as a branching unit, and molar equivalencies of reagents were doubled in coupling reactions after installation of the branching unit. After Fmoc deprotection of the N-terminal amino acids, the resin was bubbled in a phenylboronic acid solution (0.2 g/mL) overnight to remove the pinacol groups of boron-containing side chains. Finally, the resin was treated with 95:2.5:2.5 TFA /H_2O/TIS (v/v/v) for 3 hr. After deprotection, the resin was washed extensively with DMF, DCM, and MeOH before drying and storing at −20 °C.

5.4.2 On-bead Screening Assay

DY547 labeled HIV-1 RRE-IIB RNA (5′-DY547-GGC-UGG-UAU-GGG-CGC-AGC-GUC-AAU-GAC-GCU-GAC-GGU- ACA-GGC-CAG-CC-3′) was purchased from Dharmacon and prepared according to the manufacturer's protocol. To account for the autofluorescence of Tentagel Macrobead-NH_2 resin, control peptide $(LLK)_2$*$K_{BBA}K_{ACR}$LY was incubated in 100 nM DY547 labeled HIV-1 RRE-IIB RNA for 1 hr in phosphate buffer (10 mM potassium phosphate, 100 mM KCl, 1 mM $MgCl_2$, 20 mM NaCl, pH 7.0). These

beads were washed extensively and placed into a sterile 96-well plate (Nunc) and imaged by both a Zeiss Axiovert 200 fluorescent microscope under a rhodamine filter and a Zeiss LSM 510 microscope et to longpass 585. The fluorescence intensity of these RNA treated beads was compared with the auto-fluorescence of untreated control branched peptide resin, and the detector sensitivity was adjusted to remove auto-fluorescence of the resin (Figure 5.9). For fluorescence microscopy, adjusting the fluorescence intensity to a minimum setting would not eliminate the auto-fluorescence of the untreated beads; however, the fluorescence of control beads are higher enough to discriminate from untreated beads. Screening conditions were initially tested by using about 500 beads from the library. The beads were placed into a 1.5 mL non-stick microfuge tube (Fisher) with a 200 μL final volume of phosphate buffer and mixed by a Barnstead/Thermolyne Labquake rotisserie shaker. The beads were first treated with 1 mg/mL bovine serum albumin (BSA) (New England BioLabs) and 6.25 mg/mL tRNA (Roche) (~5,000-fold molar excess to RRE-IIB RNA) for 4 hr at rt to block nonspecific binding peptide sequences. Then, the beads were washed 5 times with phosphate buffer and incubated in 200 μL of 50 nM DY547 labeled RRE-IIB RNA in phosphate buffer for 3 hr at rt. After the final incubation, the resin was extensively washed with buffer, and the beads were imaged under both a confocal microscope and a fluorescence microscope in a 96-well plate by using the previously optimized settings (Figure 5.10). The initial screening afforded 4 hits which were isolated, rinsed with DMF (5 x 500 μL) and MeOH (5 x 500 μL), and photocleaved in clear non-stick 0.5 mL microfuge tubes in 15 μL of 1:1 MeOH: H_2O (v/v) by irradiation at 365 nm with a 4W handheld UV lamp. The supernatant was retained and subjected to MALDI-

TOF analysis. MALDI-MS/MS fragmentation analysis generated 2 sequences (two was deemed a false positive) providing a hit rate of 0.4%.

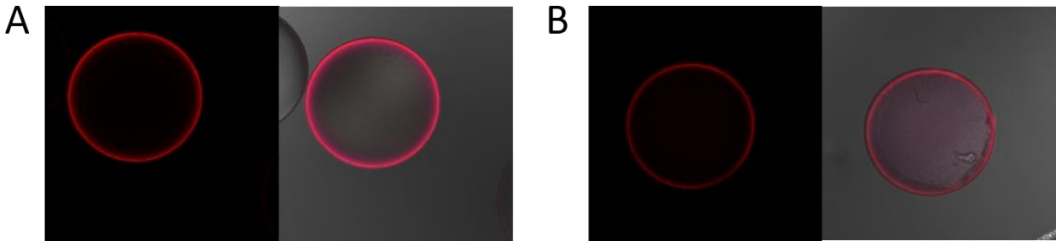


Figure 5.9 (A) Positive control peptide $(LLK)_2$*$K_{BBA}K_{ACR}LY$ and (B) Branched peptide library with 100 nM RRE incubated in phosphate buffer for 1 hr, washed, and visualized under confocal microscope (Left), and overlap of confocal microscopy with transmitted light (right).

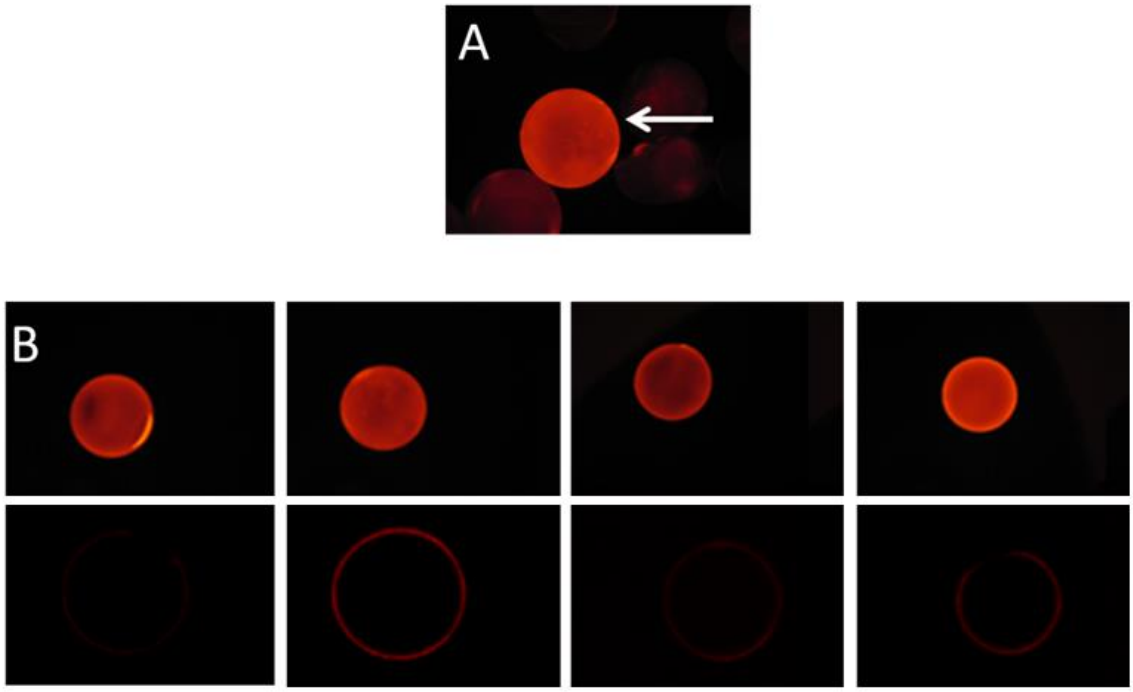


Figure 5.10 (A) Fluorescence image of incubated library, arrow indicate possible hit. (B) Examples of images of hits isolated from the library obtained via fluorescent microscopy (top) and confocal microscopy (bottom).

The remaining beads of the 3.3.4 library were divided into groups, in which contained 50 mg of beads. Each group was screened by using more stringent conditions in an attempt to reduce the total number of hit beads. First the beads were blocked at rt for 4

hr in phosphate buffer with 1 mg/mL BSA, and 12.5 mg/mL tRNA (10,000-fold molar excess to RRE stem IIB RNA). Next, the beads were washed in buffer and incubated with 50 nM DY547 labeled RRE-IIB RNA in phosphate buffer for 3 hr at RT. The resin was washed extensively after final incubation prior to screening. Another 16 beads were found with elevated fluorescence by using these more stringent conditions. These beads were photocleaved and sequenced using MALDI-TOF.

5.4.3 Peptide Synthesis, Purification and Characterization

Synthesis of the branched peptides was achieved by solid phase peptide synthesis using *N*-α-Fmoc protected L-amino acids (Novabiochem) (3 equivalent), Pyoxim (Peptides International) (3 equivalent) in DMF as coupling reagent, and DIEA (Aldrich) (5 equivalent) on Rink amide MBHA resin (100-200 mesh) (Novabiochem) with 0.4 mmol/g loading. The Fmoc group was deprotected with 20% piperidine in DMF. Fmoc-Lys(Fmoc)-OH was used as a branching unit, and molar equivalencies of reagents were doubled in coupling reactions after installation of the branching unit. The solid phase synthesis was done 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) overnight to remove the pinacol groups of boron-containing side chains. Finally, the resin was treated with 95:2.5:2.5 TFA/H_2O/TIS (v/v/v) for 3 hr. The supernatant was dried under reduced pressure, and the crude peptide was triturated from cold diethyl ether. The peptides were purified by using a Jupiter 4 μm Proteo 90 Å semiprep column (Phenomenex) using a solvent gradient composed of 0.1% TFA in Milli-Q water

and HPLC grade acetonitrile. Peptide purity was determined by using a Jupiter 4 μm Proteo 90 Å analytical column (Phenomenex), and peptide identity was confirmed by MALDI-TOF analysis. Non-acridine containing peptide concentrations were measured in nuclease free water at 280 nm by using their calculated extinction coefficients. Acridine containing peptide concentrations were monitored at 412 nm using the extinction coefficient of acridine at 13,200 mol^{-1} cm^{-1} in nuclease free water.

5.4.4 Preparation of ^{32}P-labeled RNA

Wild-type and mutant RRE-IIB RNAs were transcribed *in vitro* by T7 polymerase with the Ribomax T7 Express System (Promega) by using previously reported techniques.[64, 73] The antisense templates, sense complementary strand (5'-ATGTAATACGACTCACT ATAGG-3') and RRE-IIB reverse PCR primer (5'-GGCTGGCCTGTAC-3') were purchased from integrated DNA Technologies. Antisense templates were used as follows: HIV-1 RRE-IIB Wild Type RNA 5'-GGCTGGCCTGTACCGTCAGCGTCATTGACGCTGCGCCCATACCAGCCCTATA GTGAGTCGTATTACAT-3'; RRE-IIB Loop A(A-G)/B/Bulge A Deleted RNA 5'-GGCTGGCCAGCGTCATTGACGCTGACCAGCCCT ATAGTGAGTCGTATTACAT-3'; RRE-IIB Loop A/B/Bulge A Deleted RNA 5'-GGC TGGCAGCGTCATTGACGCTGC CAGCCCTATAGTGAGTCGTATTACAT-3'; HIV-1 TAR, 5'-GCCCGAGAGCTCCC AGGCTCAAATCGGGCCTATAGTGAGTCGTATT AC`AT. HIV-1 RRE-IIB Wild Type was PCR amplified by using HotstarTaq DNA polymerase (Qiagen) followed by a clean-up procedure by using a spin column kit (Qiagen). For the 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 hr. After transcription, 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 1x TBE buffer at 4 °C. The sample was desalted using a Sep-Pak syringe cartridge (Waters Corporation), lyophilized, and dephosphorylated with calf intestinal phosphatase (CIP) in NEBuffer 3 (New England Biolabs) according to manufacturer's protocol. The product was recovered by a standard phenol extraction followed by ethanol precipitation. Purified RNA was stored as a pellet at −80 °C.

HIV-1 RRE-IIB RNA, as well as other RNA sequences, and RRE IIB DNA were labeled at the 5'-end by treating 10 pmol of dephosphorylated RNA with 20 pmol of [γ-^{32}P] ATP (111 TBq mol^{-1}) and 20 units of T4 polynucleotide kinase in 70 mM Tris•HCl, 10 mM $MgCl_2$, and 5 mM dithiothreitol, pH 7.6. The mixture was incubated at 37 °C for 30 min, and then at rt for 20 min. The kinase was heat-inactivated at 65 °C for 10 min. The RNA was recovered by ethanol precipitation, and the purity was examined by using 12% denaturing PAGE followed by autoradiography.

5.4.5 EMSA

EMSA were performed at room temperature. First, 4 nM ^{32}P labeled RNA in 2x phosphate buffer (20 mM potassium phosphate, 200mM KCl, 1mM $MgCl_2$, 40mM NaCl, pH 7.0) was refolded by heating at 95 °C for 3 min and cooling at room temperature for 20 min. A10 μL solution of the refolded RNA was added to 10 μL of peptide in nuclease free water and incubated at room temperature for 4 hours. The final concentration of peptide was varied from 0.001 to 10 μM. After incubation, 3 μL of 30% glycerol was added for loading. Peptide-RNA complexes were resolved on 10% non-denaturing PAGE, which had

been pre-run for at least 1 hour. Gels were electrophoresed at 400 V for 20 min at 4 °C. Gels were dried to filter paper and visualized by autoradiography. Each experiment was repeated 3 times. Data were measured as the percentage of bound RNA in each lane, and error bars represent the standard deviation calculated over three replicates.

5.4.6 Fluorescence Binding Assays

All fluorescence spectra were measured on a Varian Cary Eclipse fluorescence spectrophotometer by using a xenon flash lamp with thermoelectrically controlled cell holder. The excitation slit width and the emission slit width were set to 10 nm. The excitation of the sample was done at 310 nm wavelength and fluorescence spectra were collected from 340 nm to 450 nm. A quartz cell of 1 cm path length transparent from three side was used. All experiments were done at 20 °C. The dissociate constant for BPs binding to 2-AP labeled RRE IIB was determined by following the decrease in fluorescence at 368 nm as a function of increase in BP concentration. The 2-AP labeled RRE IIB was refolded and fixed at 0.1 μM during the titration and the peptide concentration was varied from 0–3 μM. Both peptides and RNA were prepared with phosphate buffer (10 mM potassium phosphate, 100 mM KCl, 1 mM $MgCl_2$, 20 mM NaCl, pH 7.0).

Binding data were analysized by using following equations with Kaleidagraph (Synergy Software). It has been previously discussed that when K_d of proteins binding to RNA less than 20–50 nM, it would be optimized to fit the data to a quadratic equation (**Equation 1**). In other cases, a hill equation would be good (**Equation 2**).[74]

$$y = b + \frac{(m-b)}{2R} * [(x + R + K) - \sqrt{(x + R + K)^2 - 4Rx}\ ]; \quad (1)$$

$$y = b + (m - b) * \left(\frac{1}{1+(K/x)^n}\right); \quad (2)$$

In these equations, 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 values of the peptide binding to the RNA; n is the apparent cooperativity. Each experiment was performed in triplicate and error bars represent the standard deviation calculated over three replicates.

The stoichiometry of binding was determined by Job's plot by using fluorescence microscopy.The total concentration of BPs and RNA was fixed at 0.4 μM, where the mole ratio of peptide vs RNA was varied from 0.25 to 4. The fluorescence intensity was plotted against the ration of [peptide]/[RNA]. The intersection of the two linear portions of the Job's plot yielded the mole ratio corresponding to the approximate binding stoichiometry between peptide and RNA. In our system, we observed a 1:1 stoichiometry for all compounds.

5.4.7 Nuclease Protection

RNA was first refolded by heating a solution of 5'-^{32}P-labeled RRE-IIB (10 nM) and excess unlabeled RRE-IIB (200 nM) at 95 °C for 3 min and then snap cooling on ice. The refolded RNA was incubated on ice for 4 hr in a solution containing the BPs and buffer composed of 10 mM Tris, pH 7, 100 mM KCl, and 10 mM $MgCl_2$. RNase (Ambion) was then added to the solution, which was further incubated on ice for 10 min (0.002 Units RNase V1), or 1 hr (1 Unit RNase T1; 20 ng RNase A). Inactivation/precipitation buffer (Ambion) was added to halt digestion, and then the RNA was pelleted by centrifugation at 13,200 rpm for 15 min. Pelleted RNA was redissolved into tracking dye and run through a 12 % PAGE containing 7.5 M urea and imaged by autoradiography.

5.5 References

1. Organization, W. H.; UNICEF; UNAIDS., *Global HIV/AIDS Response: Construction of Core Indicators for Monitoring the 2011 UN Political Declaration on HIV/AIDS. Progress Report 2013,*. Cambodia, 2013.

2. Balfe, P.; Simmonds, P.; Ludlam, C. A.; Bishop, J. O.; Brown, A. J., Concurrent evolution of human immunodeficiency virus type 1 in patients infected from the same source: rate of sequence change and low frequency of inactivating mutations. *J. Virol.* **1990,** *64* (12), 6221-33.

3. Domaoal, R. A.; Demeter, L. M., Structural and biochemical effects of human immunodeficiency virus mutants resistant to non-nucleoside reverse transcriptase inhibitors. *Int. J. Biochem. Cell Biol.* **2004,** *36* (9), 1735-51.

4. Wei, X.; Decker, J. M.; Liu, H.; Zhang, Z.; Arani, R. B.; Kilby, J. M.; Saag, M. S.; Wu, X.; Shaw, G. M.; Kappes, J. C., Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. *Antimicrob. Agents Chemother.* **2002,** *46* (6), 1896-905.

5. Fikkert, V.; Van Maele, B.; Vercammen, J.; Hantson, A.; Van Remoortel, B.; Michiels, M.; Gurnari, C.; Pannecouque, C.; De Maeyer, M.; Engelborghs, Y.; De Clercq, E.; Debyser, Z.; Witvrouw, M., Development of resistance against diketo derivatives of human immunodeficiency virus type 1 by progressive accumulation of integrase mutations. *J. Virol.* **2003,** *77* (21), 11459-70.

6. Mehellou, Y.; De Clercq, E., Twenty-six years of anti-HIV drug discovery: Where do we stand and where do we go? *J. Med. Chem.* **2009,** *53* (2), 521-538.

7. Manche, L.; Green, S. R.; Schmedt, C.; Mathews, M. B., Interactions between double-stranded RNA regulators and the protein kinase DAI. *Mol. Cell. Biol.* **1992,** *12* (11), 11.

8. Sledz, C. A.; Holko, M.; de Veer, M. J.; Silverman, R. H.; Williams, B. R. G., Activation of the interferon system by short-interfering RNAs. *Nat. Cell. Biol.* **2003,** *5* (9), 834-839.

9. Sen, G. L.; Blau, H. M., A brief history of RNAi: the silence of the genes. *FASEB J.* **2006,** *20* (9), 1293-1299.

10. Bayne, E. H.; Allshire, R. C., RNA-directed transcriptional gene silencing in mammals. *Trends Genet.* **2005,** *21* (7), 370-373.

11. Jovanovic, M.; Hengartner, M. O., miRNAs and apoptosis: RNAs to die for. *Oncogene* **2006,** *25* (46), 6176-6187.

12. Malim, M. H.; Hauber, J.; Le, S. Y.; Maizel, J. V.; Cullen, B. R., The HIV-1 Rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* **1989,** *338* (6212), 254-7.

13. Hammarskjold, M. L.; Heimer, J.; Hammarskjold, B.; Sangwan, I.; Albert, L.; Rekosh, D., Regulation of human immunodeficiency virus env expression by the rev gene product. *J. Virol.* **1989,** *63* (5), 1959-1966.

14. Felber, B. K.; Hadzopoulou-Cladaras, M.; Cladaras, C.; Copeland, T.; Pavlakis, G. N., Rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. *Proc. Natl. Acad. Sci. U.S.A.* **1989,** *86*, 1496-1499.

15. Dayton, A. I.; Sodroski, J. G.; Rosen, C. A.; Goh, W. C.; Haseltine, W. A., The trans-activator gene of the human T cell lymphotropic virus type III is required for replication. *Cell* **1986,** *44* (6), 941-7.

16. Suhasini, M.; Reddy, T. R., Cellular proteins and HIV-1 Rev function. *Curr. HIV Res.* **2009,** *7* (1), 91-100.

17. Henderson, B. R.; Percipalle, P., Interactions between HIV Rev and nuclear import and export factors: the Rev nuclear localisation signal mediates specific binding to human importin-beta. *J. Mol. Biol.* **1997,** *274* (5), 693-707.

18. Truant, R.; Cullen, B. R., The arginine-rich domains present in human immunodeficiency virus type 1 Tat and Rev function as direct importin beta-dependent nuclear localization signals. *Mol. Cell Biol.* **1999,** *19* (2), 1210-7.

19. Arnold, M.; Nath, A.; Hauber, J.; Kehlenbach, R. H., Multiple importins function as nuclear transport receptors for the Rev protein of human immunodeficiency virus type 1. *J. Biol. Chem.* **2006,** *281* (30), 20883-90.

20. Rosen, C. A.; Terwilliger, E.; Dayton, A.; Sodroski, J. G.; Haseltine, W. A., Intragenic cis-acting art gene-responsive sequences of the human immunodeficiency virus. *Proc. Natl. Acad. Sci. U. S. A.* **1988,** *85* (7), 2071-5.

21. Kjems, J.; Brown, M.; Chang, D. D.; Sharp, P. A., 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* (3), 683-7.

22. DiMattia, M. A.; Watts, N. R.; Stahl, S. J.; Rader, C.; Wingfield, P. T.; Stuart, D. I.; Steven, A. C.; Grimes, J. M., Implications of the HIV-1 Rev dimer structure at 3.2 A resolution for multimeric binding to the Rev response element. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107* (13), 5810-4.

23. Daugherty, M. D.; Liu, B.; Frankel, A. D., Structural basis for cooperative RNA binding and export complex assembly by HIV Rev. *Nat. Struct. Mol. Biol.* **2010,** *17* (11), 1337-42.

24. Cook, K. S.; Fisk, G. J.; Hauber, J.; Usman, N.; Daly, T. J.; Rusche, J. R., Characterization of HIV-1 Rev protein: binding stoichiometry and minimal RNA substrate. *Nucleic Acids Res.* **1991,** *19* (7), 1577-83.

25. Malim, M. H.; Tiley, L. S.; McCarn, D. F.; Rusche, J. R.; Hauber, J.; Cullen, B. R., HIV-1 structural gene expression requires binding of the Rev trans-activator to its RNA target sequence. *Cell* **1990,** *60* (4), 675-83.

26. Huang, X. J.; Hope, T. J.; Bond, B. L.; McDonald, D.; Grahl, K.; Parslow, T. G., Minimal Rev-response element for type 1 human immunodeficiency virus. *J. Virol.* **1991,** *65* (4), 2131-4.

27. Battiste, J. L.; Mao, H.; Rao, N. S.; Tan, R.; Muhandiram, D. R.; Kay, L. E.; Frankel, A. D.; Williamson, J. R., Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. *Science* **1996,** *273* (5281), 1547-51.

28. Askjaer, P.; Jensen, T. H.; Nilsson, J.; Englmeier, L.; Kjems, J., The specificity of the CRM1-Rev nuclear export signal interaction is mediated by RanGTP. *J. Biol. Chem.* **1998,** *273* (50), 33414-22.

29. Fornerod, M.; Ohno, M.; Yoshida, M.; Mattaj, I. W., CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* **1997,** *90* (6), 1051-60.

30. Neville, M.; Stutz, F.; Lee, L.; Davis, L. I.; Rosbash, M., The importin-beta family member Crm1 bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Curr. Biol.* **1997,** *7* (10), 767-75.

31. Cochrane, A. W.; Chen, C. H.; Rosen, C. A., Specific interaction of the human immunodeficiency virus Rev protein with a structured region in the env mRNA. *Proc. Natl. Acad. Sci. U. S. A.* **1990,** *87* (3), 1198-1202.

32. Bartel, D. P.; Zapp, M. L.; Green, M. R.; Szostak, J. W., HIV-1 rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* **1991,** *67* (3), 529-536.

33. Fernandes, J.; Jayaraman, B.; Frankel, A., The HIV-1 Rev response element: an RNA scaffold that directs the cooperative assembly of a homo-oligomeric ribonucleoprotein complex. *RNA Biol.* **2012,** *9* (1), 6-11.

34. Hamasaki, K.; Ueno, A., Aminoglycoside antibiotics, neamine and its derivatives as potent inhibitors for the RNA-protein interactions derived from HIV-1 activators. *Bioorg. Med. Chem. Lett.* **2001,** *11* (4), 591-4.

35. Zapp, M. L.; Young, D. W.; Kumar, A.; Singh, R.; Boykin, D. W.; Wilson, W. D.; Green, M. R., Modulation of the RRE–Rev interaction by aromatic heterocyclic compounds. *Bioorg. Med. Chem.* **1997,** *5* (6), 1149-55.

36. Wilson, W. D.; Li, K., Targeting RNA with small molecules. *Curr. Med. Chem.* **2000,** *7* (1), 73-98.

37. Thomas, J. R.; Hergenrother, P. J., Targeting RNA with Small Molecules. *Chem. Rev.* **2008,** *108* (4), 1171-1224.

38. Lee, Y.; Hyun, S.; Kim, H. J.; Yu, J., Amphiphilic helical peptides containing two acridine moieties display picomolar affinity toward HIV-1 RRE and TAR. *Angew. Chem. Int. Ed.* **2008,** *47* (1), 134-137.

39. Zapp, M. L.; Young, D. W.; Kumar, A.; Singh, R.; Boykin, D. W.; Wilson, W. D.; Green, M. R., Modulation of the RRE–Rev interaction by aromatic heterocyclic compounds. *Bioorg. Med. Chem.* **1997,** *5* (6), 1149-55.

40. Xiao, G.; Kumar, A.; Li, K.; Rigl, C. T.; Bajic, M.; Davis, T. M.; Boykin, D. W.; Wilson, W. D., Inhibition of the HIV-1 RRE–Rev complex formation by unfused aromatic cations. *Bioorg. Med. Chem.* **2001,** *9* (5), 1097-113.

41. Luedtke, N. W.; Liu, Q.; Tor, Y., RNA-ligand interactions: affinity and specificity of aminoglycoside dimers and acridine conjugates to the HIV-1 Rev response element. *Biochemistry.* **2003,** *42* (39), 11391-403.

42. Li, K.; Davis, T. M.; Bailly, C.; Kumar, A.; Boykin, D. W.; Wilson, W. D., A heterocyclic inhibitor of the RRE–Rev complex binds to RRE as a dimer. *Biochemistry.* **2001,** *40* (5), 1150-1158.

43. Hart, R. A.; Billaud, J. N.; Choi, S. J.; Phillips, T. R., Effects of 1,8-diaminooctane on the FIV Rev regulatory system. *Virology* **2002,** *304* (1), 97-104.

44. DeJong, E. S.; Chang, C. E.; Gilson, M. K.; Marino, J. P., Proflavine acts as a Rev inhibitor by targeting the high-affinity Rev binding site of the Rev responsive element of HIV-1. *Biochemistry* **2003,** *42* (26), 8035-46.

45. Prater, C. E.; Saleh, A. D.; Wear, M. P.; Miller, P. S., Chimeric RNase H-competent oligonucleotides directed to the HIV-1 Rev response element. *Bioorg. Med. Chem.* **2007,** *15* (16), 5386-95.

46. Legiewicz, M.; Badorrek, C. S.; Turner, K. B.; Fabris, D.; Hamm, T. E.; Rekosh, D.; Hammarskjold, M. L.; Le Grice, S. F., Resistance to RevM10 inhibition reflects a conformational switch in the HIV-1 Rev response element. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** *105* (38), 14365-70.

47. Lee, T. C.; Sullenger, B. A.; Gallardo, H. F.; Ungers, G. E.; Gilboa, E., Overexpression of RRE-derived sequences inhibits HIV-1 replication in CEM cells. *New Biol.* **1992,** *4* (1), 66-74.

48. Symensma, T. L.; Baskerville, S.; Yan, A.; Ellington, A. D., Polyvalent Rev decoys act as artificial Rev-responsive elements. *J. Virol.* **1999,** *73* (5), 4341-9.

49. Giver, L.; Bartel, D. P.; Zapp, M. L.; Green, M. R.; Ellington, A. D., Selection and design of high-affinity RNA ligands for HIV-1 Rev. *Gene* **1993,** *137* (1), 19-24.

50. Chaloin, L.; Smagulova, F.; Hariton-Gazal, E.; Briant, L.; Loyter, A.; Devaux, C., Potent inhibition of HIV-1 replication by backbone cyclic peptides bearing the Rev arginine rich motif. *J. Biomed. Sci.* **2007,** *14* (5), 565-84.

51. Friedler, A.; Friedler, D.; Luedtke, N. W.; Tor, Y.; Loyter, A.; Gilon, C., Development of a functional backbone cyclic mimetic of the HIV-1 Tat arginine-rich motif. *J. Biol. Chem.* **2000,** *275* (31), 23783-9.

52. Mills, N. L.; Daugherty, M. D.; Frankel, A. D.; Guy, R. K., An α-Helical peptidomimetic inhibitor of the HIV-1 Rev–RRE interaction. *J. Am. Chem. Soc.* **2006,** *128* (11), 3496-3497.

53. Hamy, F.; Brondani, V.; Flörsheimer, A.; Stark, W.; Blommers, M. J. J.; Klimkait, T., A new class of HIV-1 Tat antagonist acting through Tat−TAR inhibition. *Biochemistry* **1998,** *37* (15), 5086-5095.

54. Luedtke, N. W.; Liu, Q.; Tor, Y., RNA−Ligand Interactions: Affinity and specificity of aminoglycoside dimers and acridine conjugates to the HIV-1 Rev response element. *Biochemistry* **2003,** *42* (39), 11391-11403.

55. Lee, Y.; Hyun, S.; Kim, H.; Yu, J., Amphiphilic helical peptides containing two acridine moieties display picomolar affinity toward HIV-1 RRE and TAR. *Angew. Chem. Int. Edit.* **2008,** *47* (1), 134-137.

56. Gooch, B. D.; Beal, P. A., Recognition of duplex RNA by helix-threading peptides. *J. Am. Chem. Soc.* **2004,** *126* (34), 10603-10610.

57. Renner, S.; Ludwig, V.; Boden, O.; Scheffer, U.; Göbel, M.; Schneider, G., New inhibitors of the Tat–TAR RNA interaction found with a "Fuzzy" pharmacophore model. *ChemBioChem.* **2005,** *6* (6), 1119-1125.

58. Zhang, W.; Bryson, D. I.; Crumpton, J. B.; Wynn, J.; Santos, W. L., Branched peptide boronic acids (BPBAs): A novel mode of binding towards RNA. *Chem. Commun.* **2013,** *49* (24), 2436-8.

59. Zhang, W.; Bryson, D. I.; Crumpton, J. B.; Wynn, J.; Santos, W. L., 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* (37), 6263-6271.

60. Bryson, D. I.; Zhang, W.; Ray, W. K.; Santos, W. L., Screening of a branched peptide library with HIV-1 TAR RNA. *Mol. BioSyst.* **2009,** *5*, 1070-1073.

61. B. Crumpton, J.; Zhang, W.; L. Santos, W., Facile analysis and sequencing of linear and branched peptide boronic acids by MALDI mass spectrometry. *Anal. Chem.* **2011,** *83* (9), 3548-3554.

62. Daugherty, M. D.; Booth, D. S.; Jayaraman, B.; Cheng, Y.; Frankel, A. D., HIV Rev response element (RRE) directs assembly of the Rev homooligomer into discrete asymmetric complexes. *Proc.Natl. Acad. Sci. U.S.A.* **2010,** *107* (28), 12481-6.

63. DiMattia, M. A.; Watts, N. R.; Stahl, S. J.; Rader, C.; Wingfield, P. T.; Stuart, D. I.; Steven, A. C.; Grimes, J. M., Implications of the HIV-1 Rev dimer structure at 3.2 A resolution for multimeric binding to the Rev response element. *Proc.Natl. Acad. Sci. U.S.A.* **2010,** *107* (13), 5810-4.

64. Bryson, D. I.; Zhang, W.; McLendon, P. M.; Reineke, T. M.; Santos, W. L., Toward targeting RNA structure: branched peptides as cell-permeable ligands to TAR RNA. *ACS Chem. Biol.* **2011**.

65. Bryson, D. I.; Zhang, W. Y.; Ray, W. K.; Santos, W. L., Screening of a branched peptide library with HIV-1 TAR RNA. *Mol. BioSyst.* **2009,** *5* (9), 1070-1073.

66. Lacourciere, K. A.; Stivers, J. T.; Marino, J. P., 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* (19), 5630-5641.

67. Millar, D. P., Fluorescence studies of DNA and RNA structure and dynamics. *Current Opin. Struc. Biol.* **1996,** *6* (3), 322-326.

68. Giver, L.; Bartel, D.; Zapp, M.; Pawul, A.; Green, M.; Ellington, A. D., Selective optimization of the Rev-binding element of HIV-1. *Nucl. Acids Res.* **1993,** *21* (23), 5509-5516.

69. Kumar, S.; Bose, D.; Suryawanshi, H.; Sabharwal, H.; Mapa, K.; Maiti, S., Specificity of RSG-1.2 peptide binding to RRE-IIB RNA element of HIV-1 over Rev peptide is mainly enthalpic in origin. *PLoS ONE* **2011,** *6* (8), e23300.

70. Kjems, J.; Calnan, B. J.; Frankel, A. D.; Sharp, P. A., Specific binding of a basic peptide from HIV-1 Rev. *EMBO J.* **1992,** *11* (3), 1119-29.

71. Kjems, J.; Brown, M.; Chang, D. D.; Sharp, P. A., 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* (3), 683-7.

72. Tan, D. S.; Foley, M. A.; Stockwell, B. R.; Shair, M. D.; Schreiber, S. L., Synthesis and preliminary evaluation of a library of polycyclic small molecules for use in chemical genetic assays. *J. Am. Chem. Soc.* **1999,** *121* (39), 9073-9087.

73. Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C., Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucl. Acids Res.* **1987,** *15* (21), 8783-8798.

74. Pagano, J. M.; Clingman, C. C.; Ryder, S. P., Quantitative approaches to monitor protein–nucleic acid interactions using fluorescent probes. *RNA* **2011,** *17* (1), 14-20.

Chapter 6. Conclusions and Future Directions

We successfully developed three generation of BP libraries and demonstrated that 'branching' in peptides imparted multivalent interactions to the RNA and that the unnatural functional groups such as boronic acid and acridine groups were key structural features for efficient binding and selectivity for the folded RNA target.

In the first generation BP library, we created a 3.3.3. BP library that was designed to target HIV-1 TAR, a structured RNA element that is absolutely vital for the viral life cycle. We identified a BP, **FL4**, that was selective for HIV-1 TAR with a binding constant (K_d) of 600 nM. We discovered this particular compound from screening a one-bead, one-compound library (4,096 compounds) that was prepared by split-and-pool technique. We demonstrated that 'branching' in peptides provides multivalent interactions, which increases binding affinity for RNA. Mutation experiments indicated that **FL4** interacts with both the loop and bulge regions. Further, we demonstrated that BPs are cell permeable and non-toxic, making them excellent chemical biology tools for targeting well-defined RNA structures. However, we tested the binding affinity of **FL4** in the presence of excess tRNA as a competitor, we found that **FL4** started to lose binding affinity with 10-fold excess tRNA, and significantly lost binding affinity with 1000-fold excess tRNA. This suggested that FL4 is partially selective toward RRE IIB RNA.

To further improve the binding properties of BPs, we incorporated unnatural amino acid side chains featuring boronic acid moieties into the new 3.3.4 boronic acid branched peptide (BPBA) library (46,656 unique sequences). The logic behind the inclusion of boron was that the 2'-hydroxyl group of the RNA ribose or Lewis base in RNA could form a reversible covalent bond with boron to generate a stable boronate species. This will not

only allow the ligand to select RNA over DNA, but also boost both affinity and selectivity to bind the RNA target. Furthermore, the boronic acid moiety presented a unique mode of binding that increased the complexity of the library and the incorporation of unnatural side chains that likely imparted resistance to enzymatic cleavage. The newly synthesized library was screened for binding to HIV-1 RRE, an RNA that is required for nuclear export of unspliced and singly spliced viral transcripts and is essential for HIV-1 replication. Through EMSA and dot blot assay, we were able to determine that these compounds bind to RRE IIB in the submicromolar range.

Three peptides containing boronic acid residues stood out for their low binding affinities. A subsequent structure–activity relationship study on the role of the boron moiety from our hit compounds revealed that boronic acid moieties are responsible for boosting the high binding affinities. Further studies of the role of the 'branches' revealed that each branch is important for binding and the high density of Lys residue for electrostatic interaction and H-bonding is not necessary for building the high affinity for RRE. We also performed EMSA of **BPBA1** against six mutant RRE IIB variants as well as RRE IIB DNA and tRNA. We demonstrated that **BPBA1** had strong interactions with the lower bulge region of RRE IIB and could distinguish RRE RNA from RRE DNA and tRNA. This significantly improved the selectivity of the BPs compared to the BPs from the last library. To better determine regions of specific nucleotide contacts with RRE IIB, we performed ribonuclease protection assays. **BPBA1** displayed a strong protection of the stem region between A16 through G19, which are within the site that the native protein partner Rev binds.[1-2] The full protection pattern demonstrated that the binding site encompassed a large surface area that spanned the upper stem as well as the internal loop

regions of RRE IIB. In addition, we also examined if the boronic acid will make the peptide toxic. We treated both HeLa cell line and A2870 cell line with up to 30 μM **BPBA1** for 24 hours, the cell viability remained, suggesting **BPBA1** is not toxic. **BPBA1** also displayed cell permeability that makes it a desirable drug candidate.

We were greatly encouraged by previous results that boronic acid moiety could improve the binding affinity toward RNA. In the next generation of branched peptides, we designed a simple 3.3.4 branched peptide library that featured boronic acid and acridine moieties. The library was only composed of four amino acids, which were Lys, Leu, K_{BBA} and K_{ACR}, hence, the library had limited diversity. We chose acridine as the amino acid side chain due to its high potential for π-stacking interaction with the bases of RNA. Providing high binding affinity to RNA target. The library (4,096 compounds) was screened against HIV-1 RRE IIB RNA in the presence of 10,000-fold tRNA and 1 mg/mL BSA. Fifteen peptide sequences were sequenced and four contained acridine moieties and displayed dissociation constants in the low 100 nM scales. However, the EMSA results of BPs binding to RRE IIB variants, RRE IIB DNA, TAR RNA and tRNA revealed that these peptides could not discriminate between these structures. We are not surprised that BPs did not distinguish between small hairpin RNAs and larger tRNA. This might be the result of strong intercalation of acridine moiety to stem like structures, coupled to the relatively high density of positive charges presented in the peptides. Furthermore, we performed ribonuclease protection assays with **A7**, a sequence that contains both boronic acid and acridine residues. The protecting pattern was similar compared to previous peptide **BPBA1**, suggesting that the 3.3.4 branched peptides shared similar structural elements and contacted comparable regions of the RRE IIB RNA.

In the future, branched peptide libraries will impart unnatural amino acids in every position to achieve high binding affinity and selectivity. The library is designed to possess tighter and more rigid structure by using shorter side chains and stiffer branching unit. Previously, research on an amphiphilic peptide provided information that replacing Lys with amino acids bearing shorter side chain (Orn, Dap and Dab) could boost both the binding affinity and the selectivity to HCV IRES.[3] In addition, unnatural amino acids will impart unique binding modes different from traditional RNA-peptide interactions. This strategy will increase the complexity of the library and impart resistance to enzymatic cleavage. Taken together, research on branched peptide library provides a general platform for targeting therapeutic RNA associated with diseases.

6.1 References

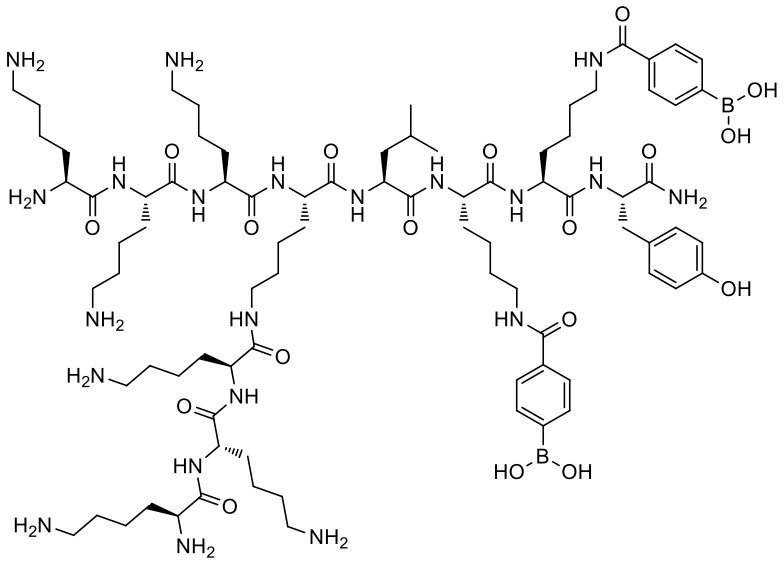
1. Kjems, J.; Calnan, B. J.; Frankel, A. D.; Sharp, P. A., Specific binding of a basic peptide from HIV-1 Rev. *Embo J.* **1992,** *11* (3), 1119-29.

2. Kjems, J.; Brown, M.; Chang, D. D.; Sharp, P. A., 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* (3), 683-7.

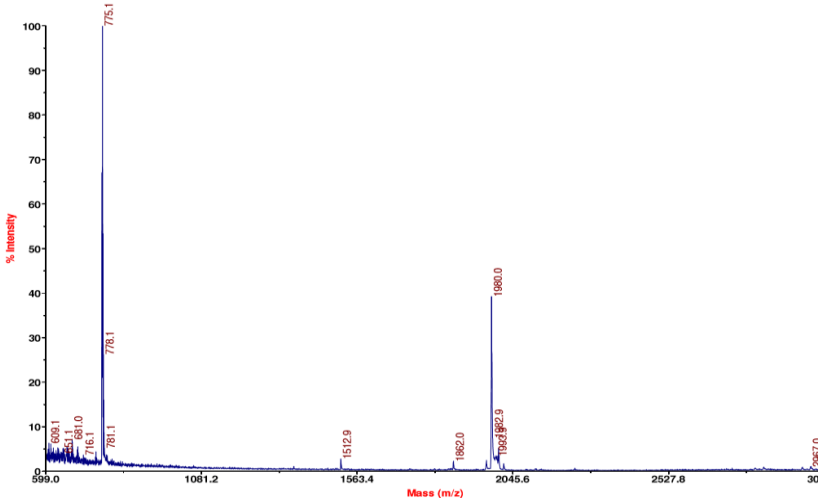
3. Lee, S. J.; Hyun, S.; Kieft, J. S.; Yu, J., An approach to the construction of tailor-made amphiphilic peptides that strongly and selectively bind to hairpin RNA targets. *J. Am. Chem. Soc.* **2009,** *131* (6), 2224-2230.

Appendix A Structures, MALDI-TOF and HPLC

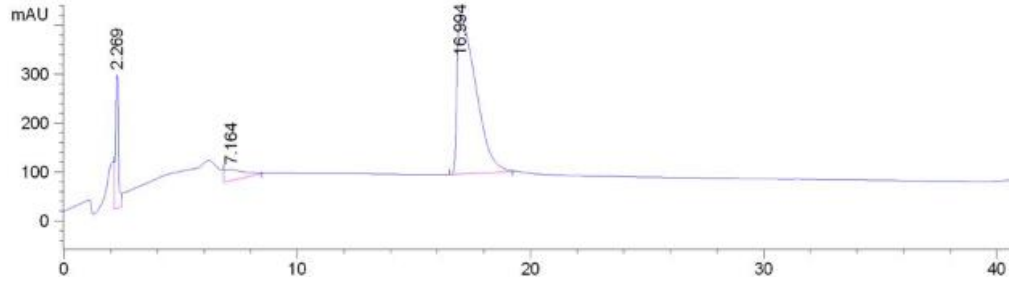
A1: $(KKK)_2$*$LK_{BBA}K_{BBA}Y$; MS(MALDI-TOF): Calcd for $C_{83}H_{141}B_2N_{21}O_{18}$ [M-DHB +H]$^+$: 1979.9, found: 1980.0; Purity (HPLC):93%.



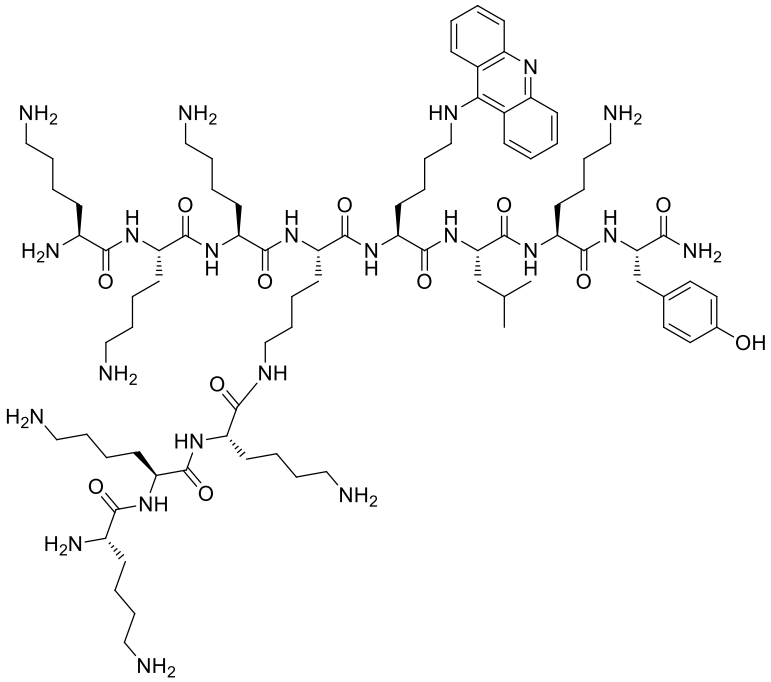
MALDI-TOF Spectrum:



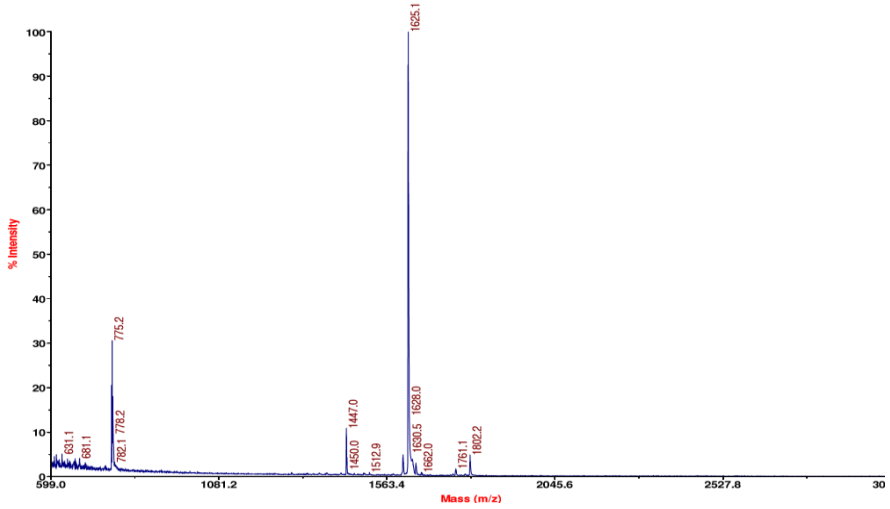
HPLC Spectrum: retention time is labeled on the peak.



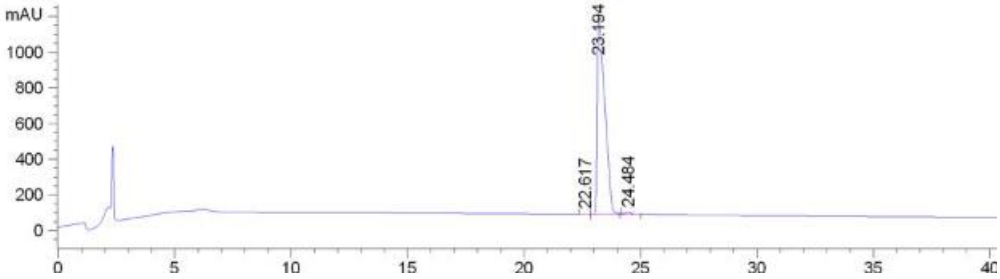
A2: $(KKK)_2$*K_{ACR}LKY; MS(MALDI-TOF): Calcd for $C_{82}H_{138}N_{22}O_{12}$ $[M\text{-}DHB+H]^+$: 1625.3, found: 1625.1; Purity (HPLC):98%.



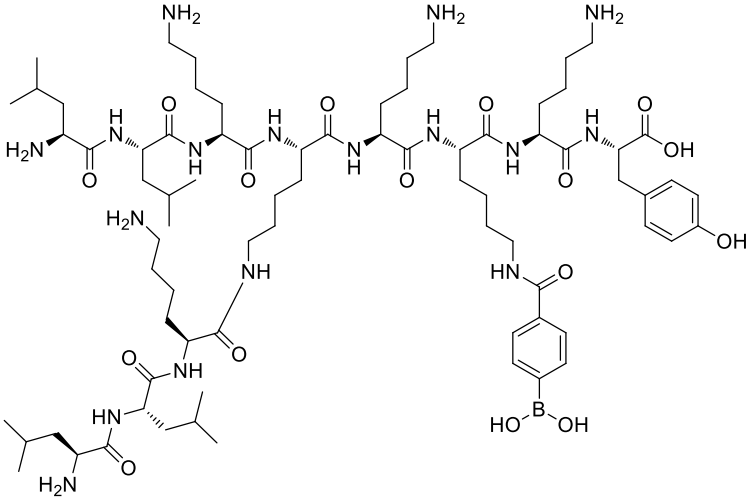
MALDI-TOF Spectrum



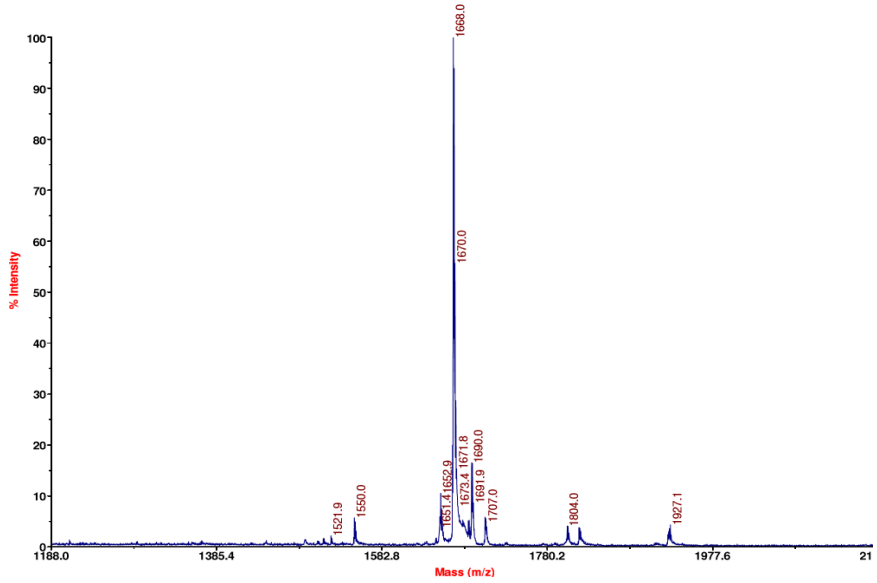
HPLC Spectrum: retention time is labeled on the peak.



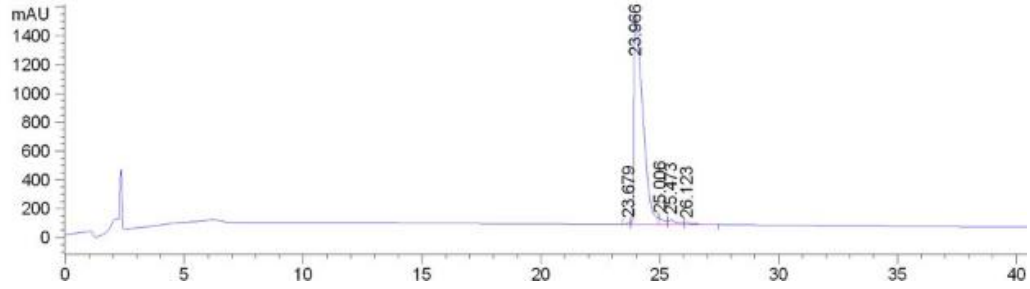
A3: $(LLK)_2$*$KK_{BBA}KY$; MS(MALDI-TOF): Calcd for $C_{76}H_{132}B_1N_{17}O_{16}$ $[M\text{-}DHB+H]^+$: 1669.0, found: 1669.0; Purity (HPLC):93%.



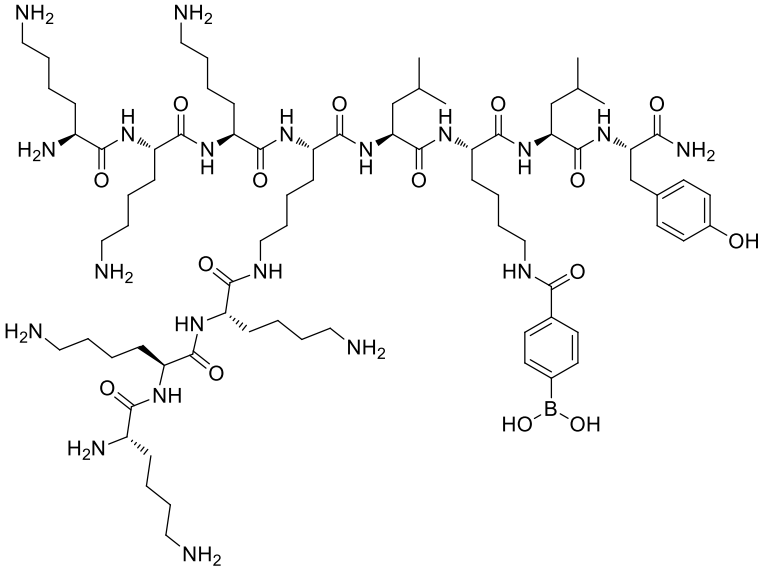
MALDI-TOF Spectrum



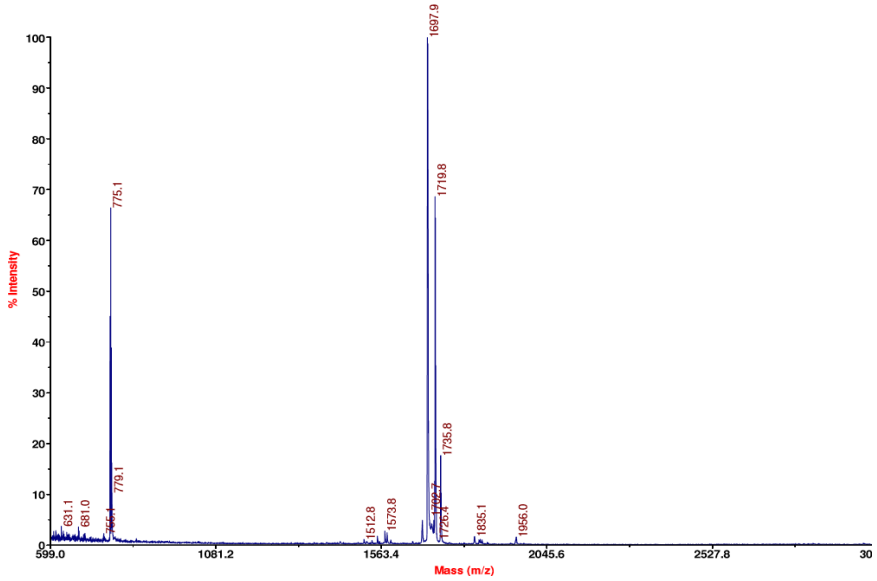
HPLC Spectrum: retention time is labeled on the peak.



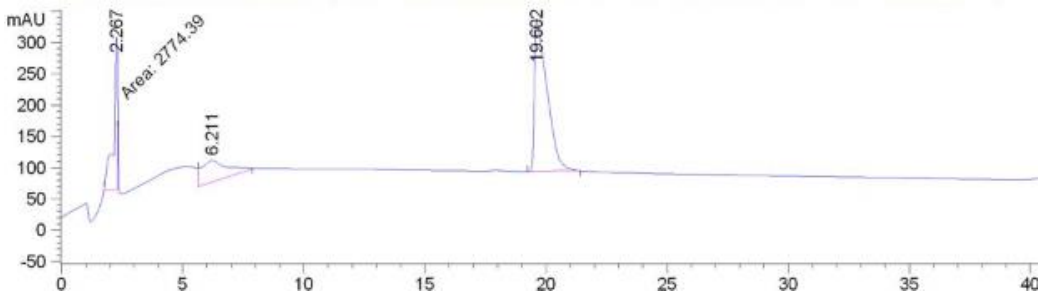
A4: $(KKK)_2$*$LK_{BBA}LY$; MS(MALDI-TOF): Calcd for $C_{76}H_{135}BN_{20}O_{15}$ $[M\text{-}DHB+H]^+$: 1699.0, found: 1698.9; Purity (HPLC):93%.



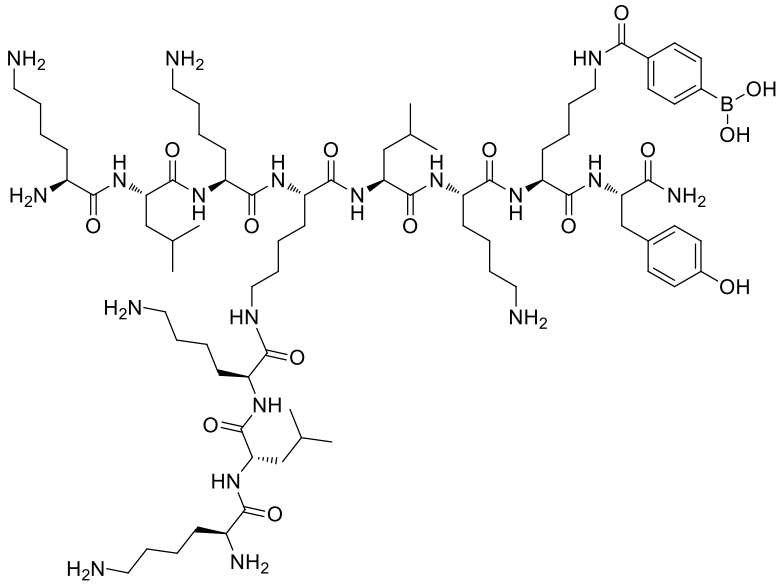
MALDI-TOF Spectrum



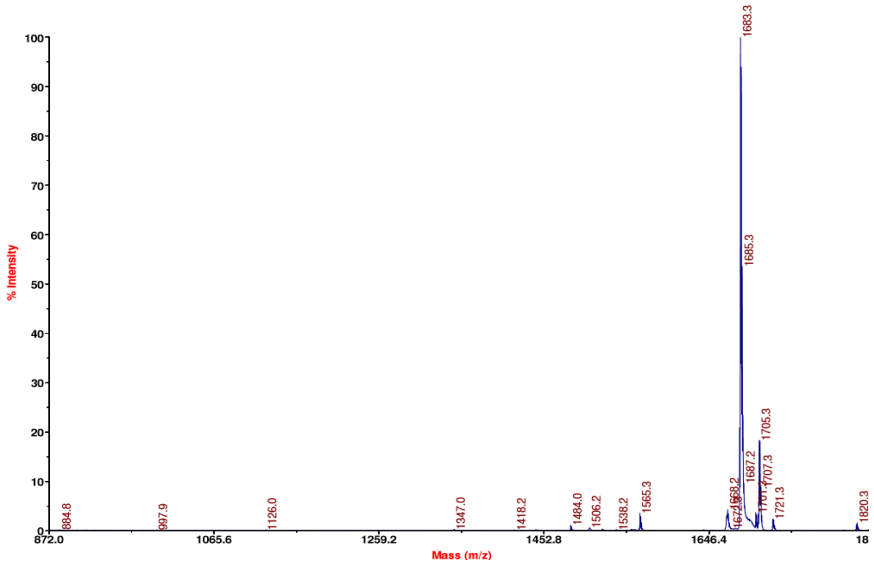
HPLC Spectrum: retention time is labeled on the peak.



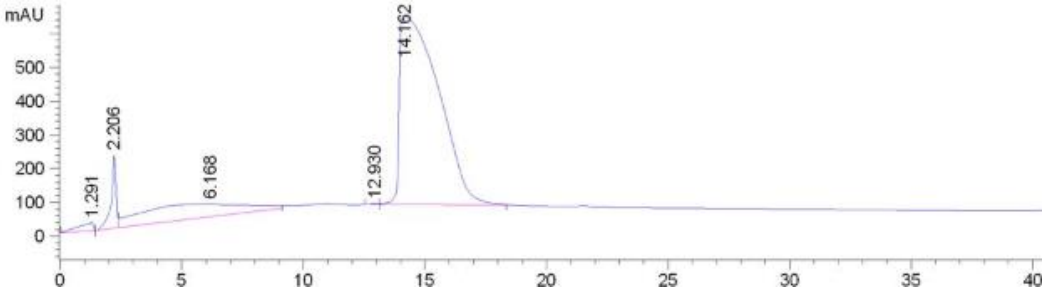
A5: $(KLK)_2$*$LKK_{BBA}Y$; MS(MALDI-TOF): Calcd for $C_{76}H_{134}BN_{19}O_{15}$ $[M\text{-}DHB+H]^+$: 1684.1, found: 1980.0; Purity (HPLC):96%.



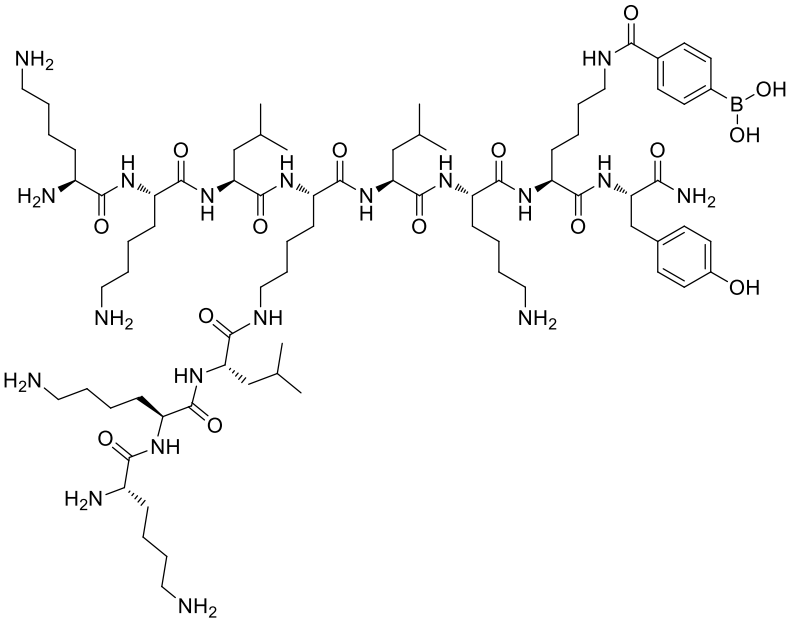
MALDI-TOF Spectrum:



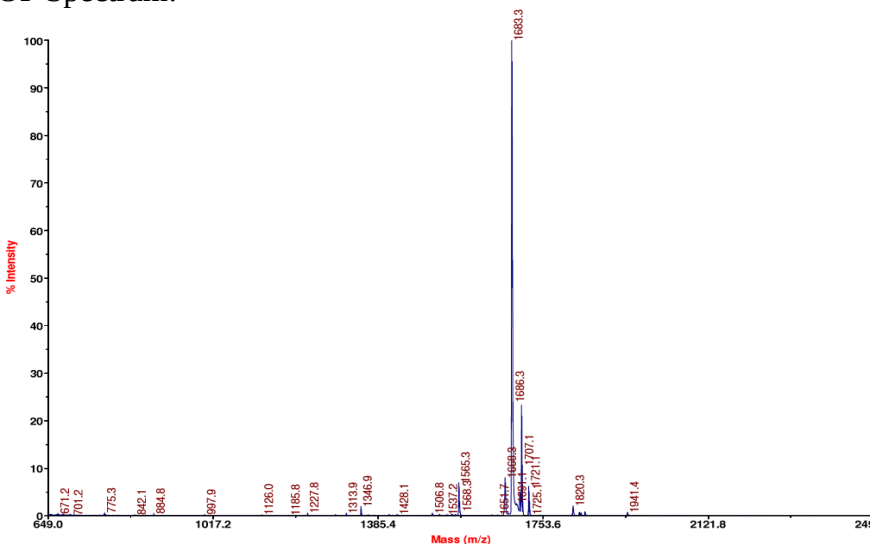
HPLC Spectrum: retention time is labeled on the peak.



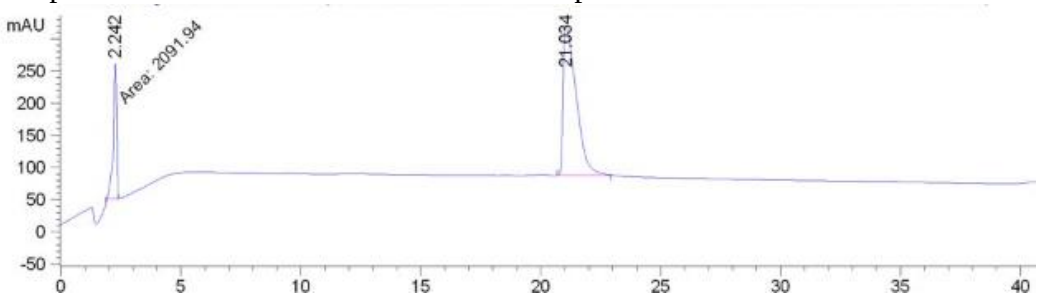
A6: $(KKL)_2$*$LKK_{BBA}Y$; MS(MALDI-TOF): Calcd for $C_{76}H_{134}BN_{19}O_{15}$ $[M\text{-}DHB+H]^+$: 1684.1, found: 1980.0; Purity (HPLC):99%.



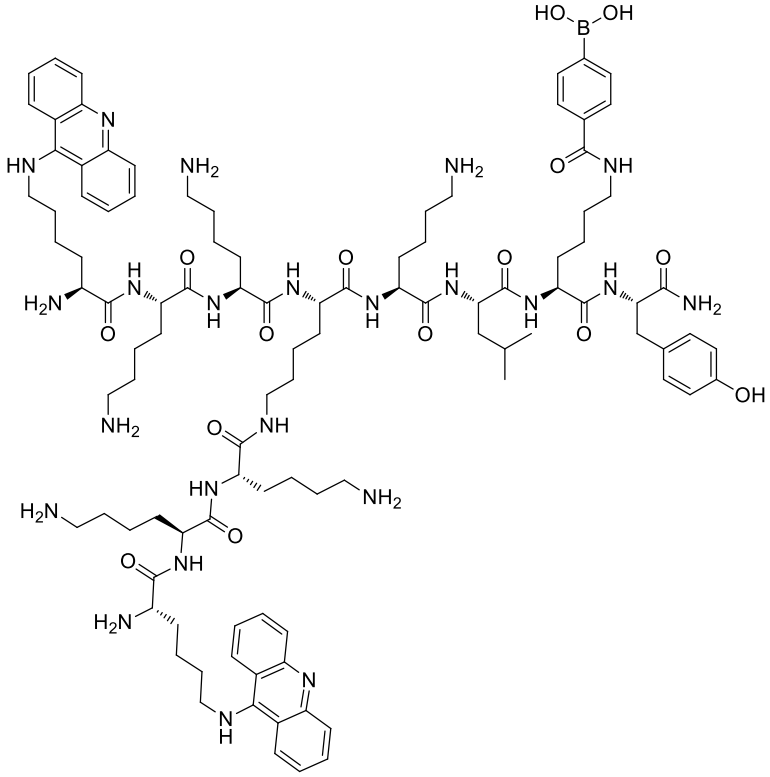
MALDI-TOF Spectrum:



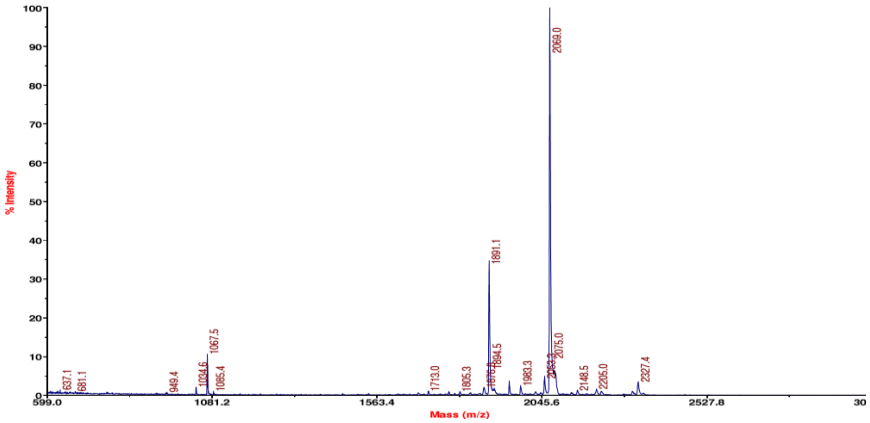
HPLC Spectrum: retention time is labeled on the peak.



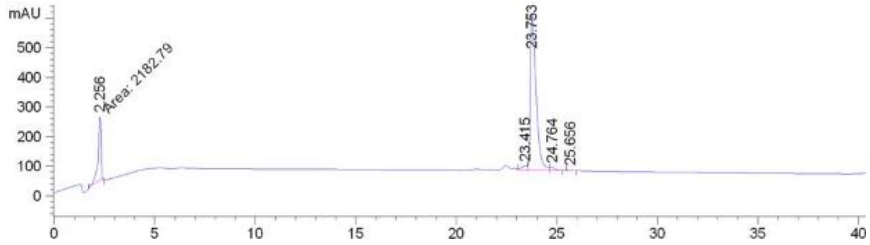
A7: $(K_{ACR}KK)_2$*$KLK_{BBA}Y$; MS(MALDI-TOF): Calcd for $C_{102}H_{150}BN_{23}O_{15}$ [M-DHB +H]$^+$: 2068.5, found: 2069.0; Purity (HPLC):91%.



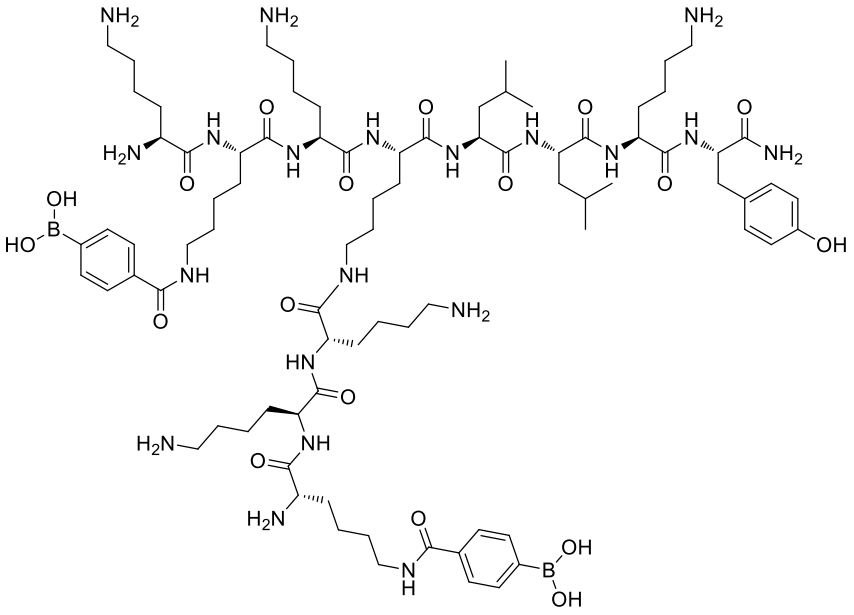
MALDI-TOF Spectrum:



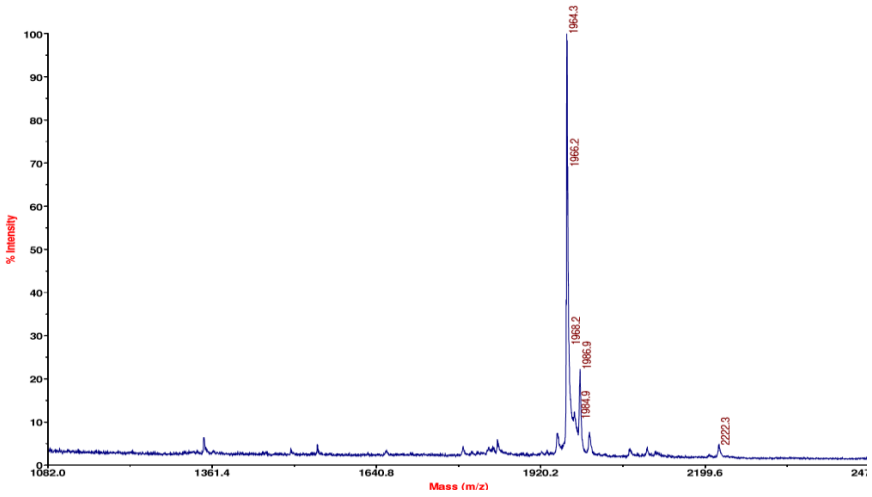
HPLC Spectrum: retention time is labeled on the peak.



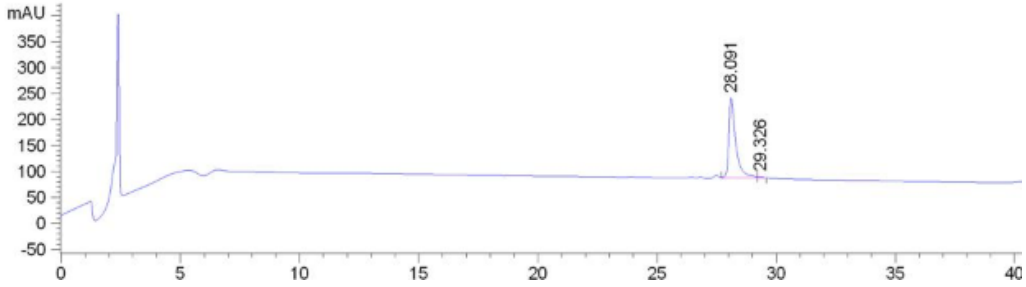
A8: $(KK_{BBA}K)_2$*LLKY; MS(MALDI-TOF): Calcd for $C_{83}H_{140}B_2N_{20}O_{18}$ $[M\text{-}DHB+H]^+$: 1965.1, found: 1965.3; Purity (HPLC):93%.



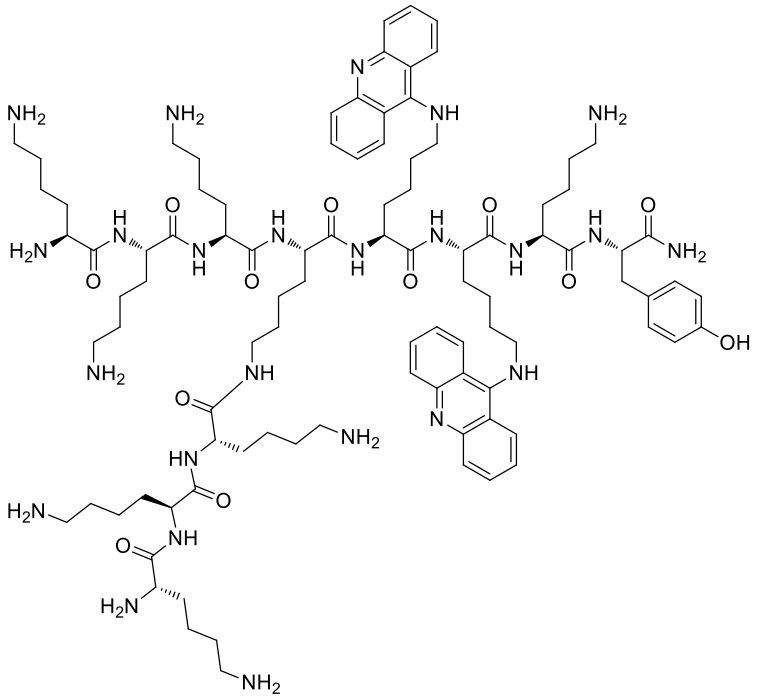
MALDI-TOF Spectrum



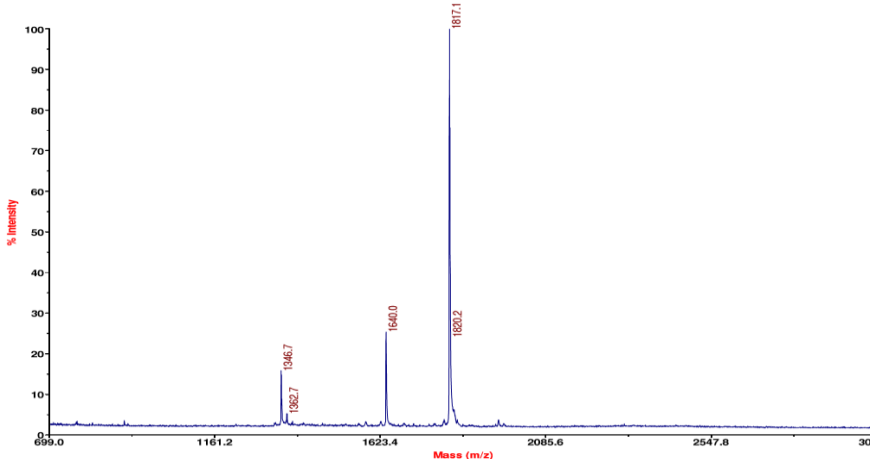
HPLC Spectrum: retention time is labeled on the peak.



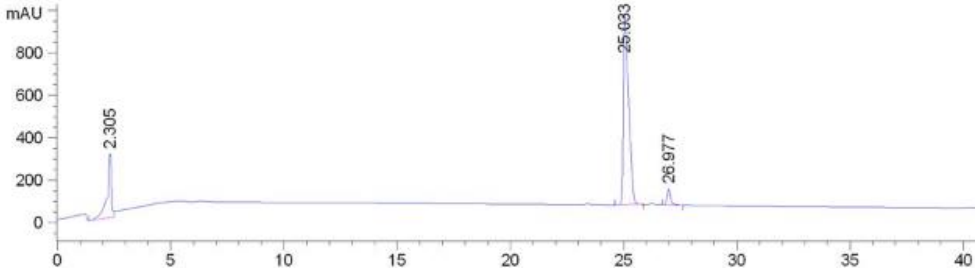
A9: $(KKK)_2$*$K_{ACR}K_{ACR}KY$; MS(MALDI-TOF): Calcd for $C_{95}H_{146}N_{24}O_{12}$ $[M\text{-}DHB+H]^+$: 1817.0, found: 1817.1; Purity (HPLC):95%.



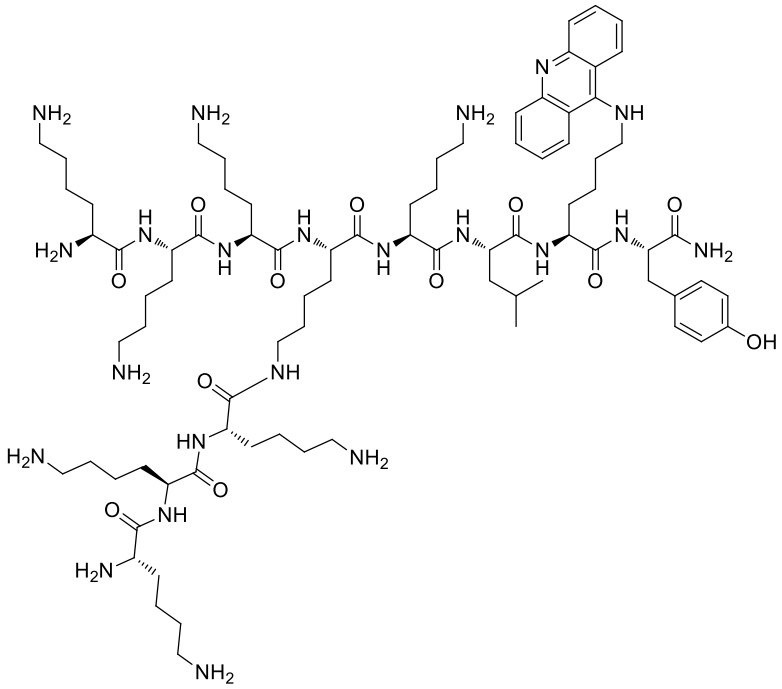
MALDI-TOF Spectrum:



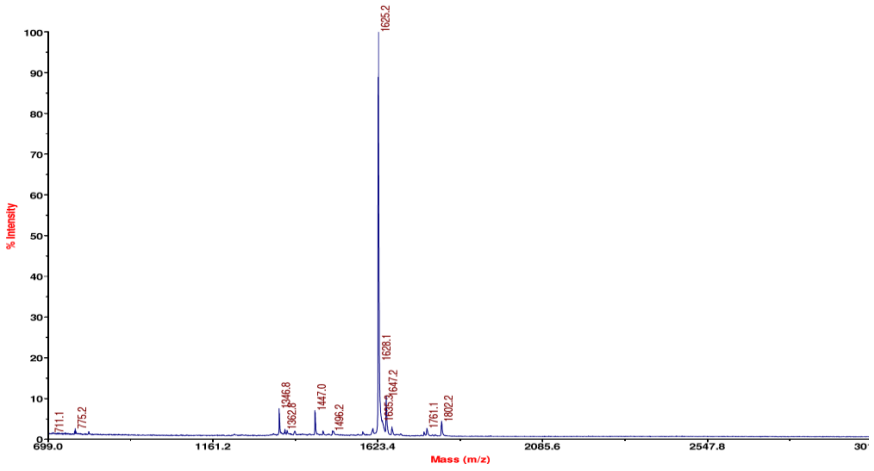
HPLC Spectrum: retention time is labeled on the peak.



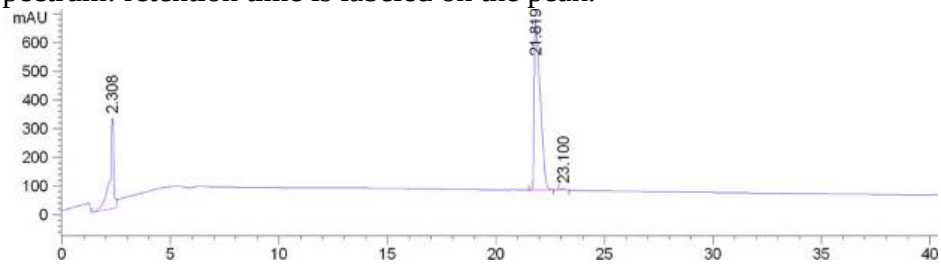
A10: $(KKK)_2$*$KLK_{ACR}Y$; MS(MALDI-TOF): Calcd for $C_{82}H_{138}N_{22}O_{12}$ $[M\text{-}DHB+H]^+$: 1625.1, found: 1625.2; Purity (HPLC):96%.



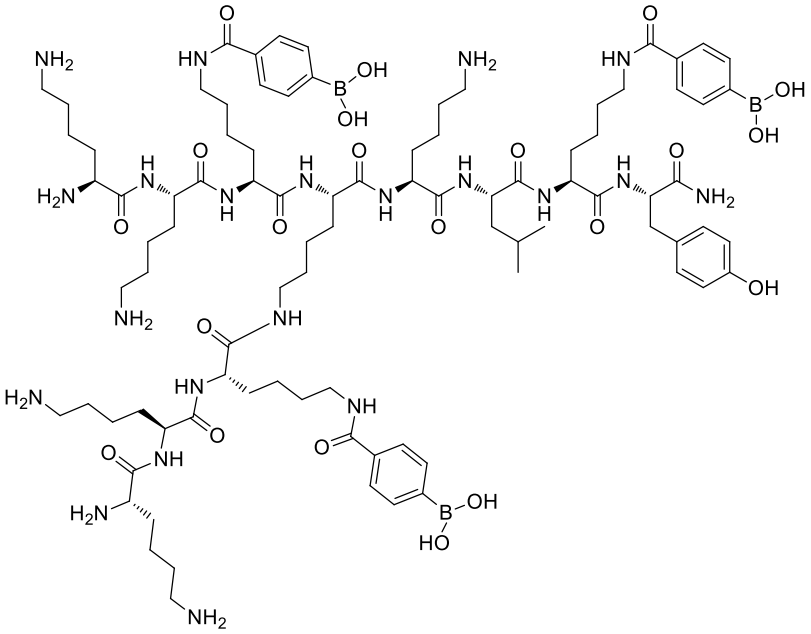
MALDI-TOF Spectrum:



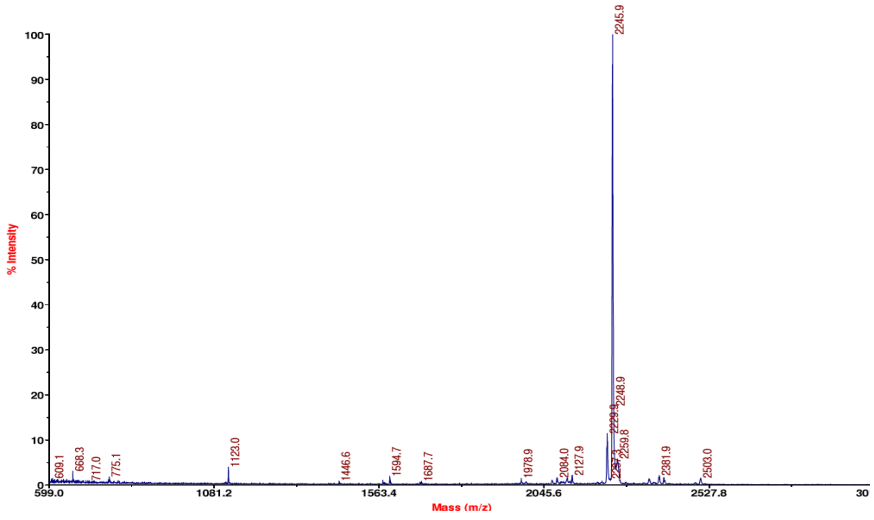
HPLC Spectrum: retention time is labeled on the peak.



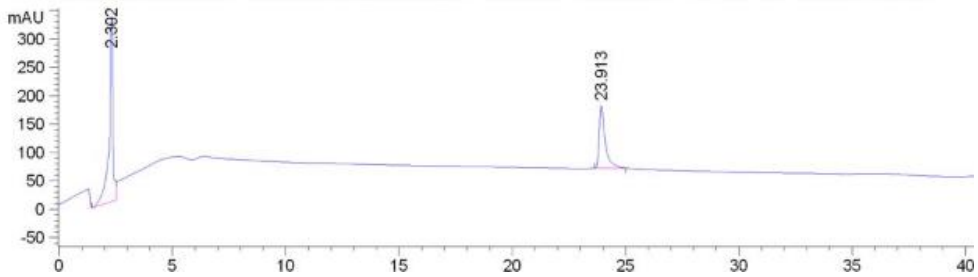
A11: $(KKK_{BBA})_2$*$KLK_{BBA}Y$; MS(MALDI-TOF): Calcd for $C_{90}H_{146}B_3N_{21}O_{21}$ $[M\text{-}DHB+H]^+$: 2246.0, found: 2245.9; Purity (HPLC):99%.



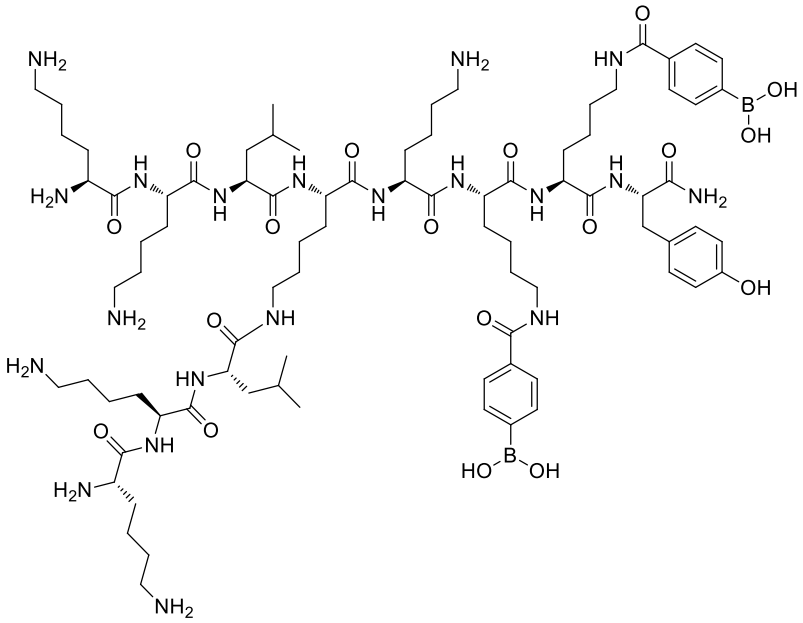
MALDI-TOF Spectrum:



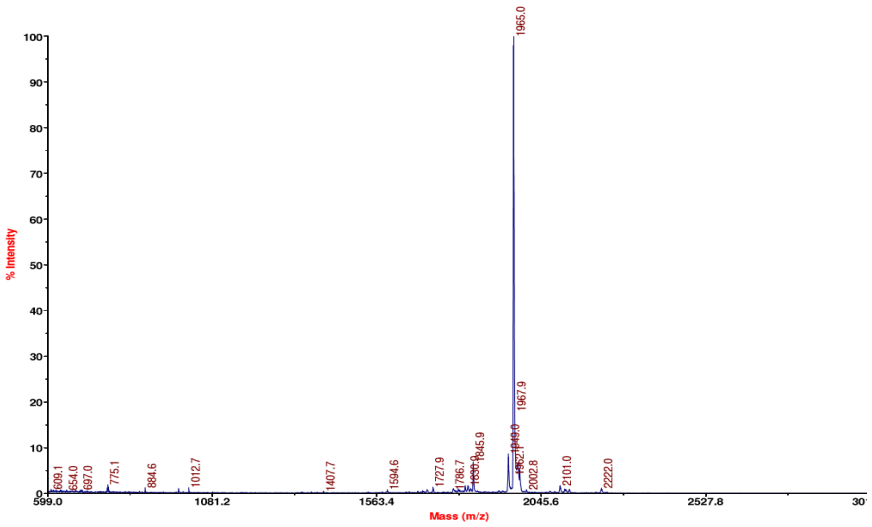
HPLC Spectrum: retention time is labeled on the peak.



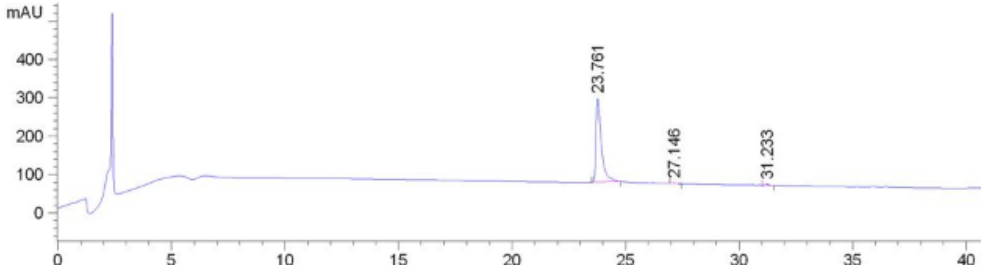
A12: $(KKL)_2$*$KK_{BBA}K_{BBA}Y$; MS(MALDI-TOF): Calcd for $C_{83}H_{140}B_2N_{20}O_{18}$ [M-DHB +H]$^+$: 1965.2, found: 1965.0; Purity (HPLC):98%.



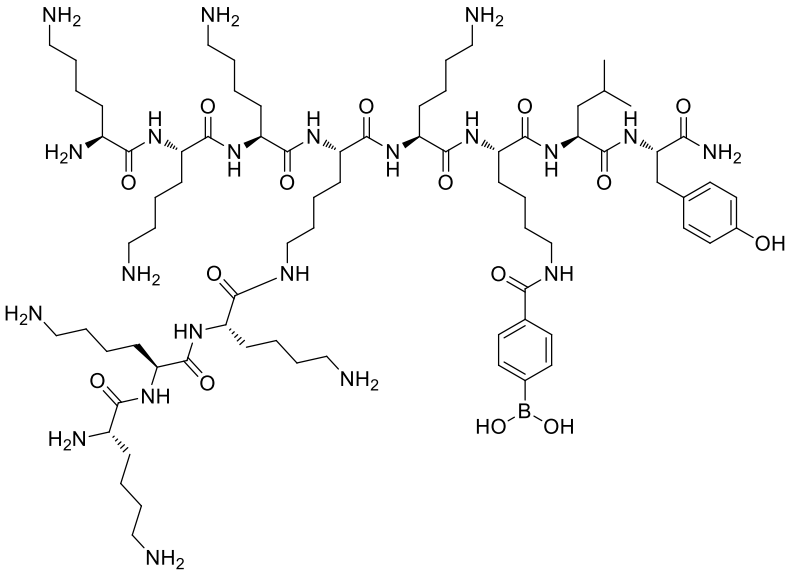
MALDI-TOF Spectrum:



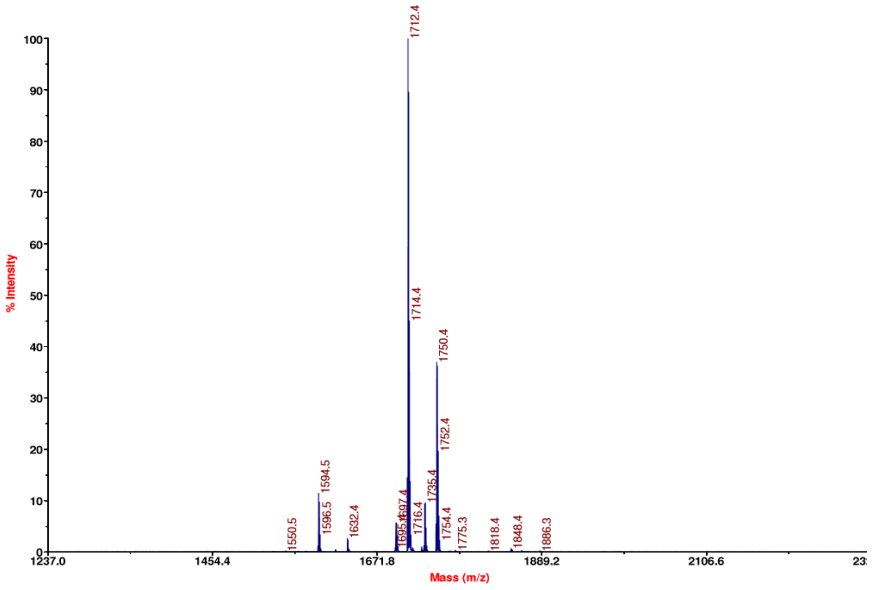
HPLC Spectrum: retention time is labeled on the peak.



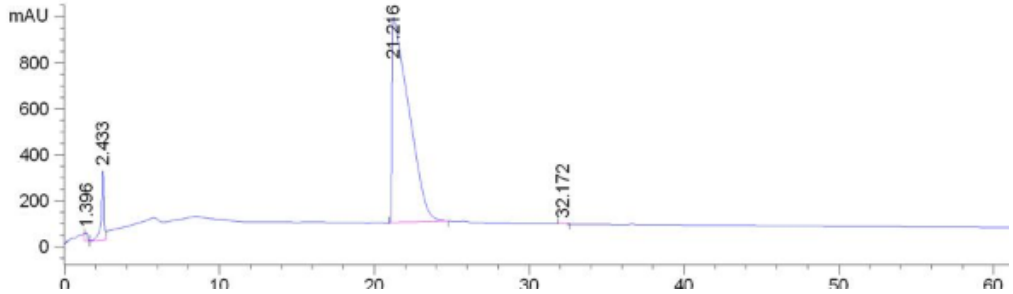
A13: $(KKK)_2$*$KK_{BBA}LY$; MS(MALDI-TOF): Calcd for $C_{76}H_{136}BN_{21}O_{15}$ $[M\text{-}DHB+H]^+$: 1714.1, found: 1714.4; Purity (HPLC):99%.



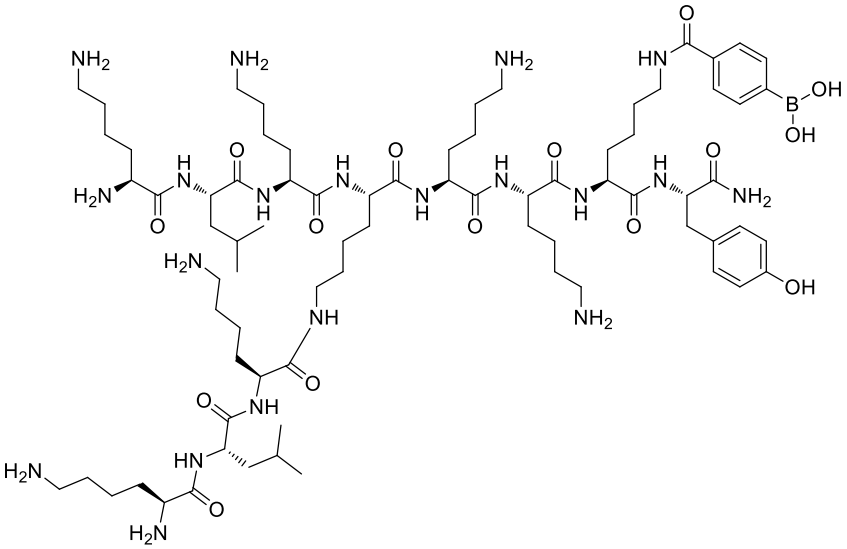
MALDI-TOF Spectrum:



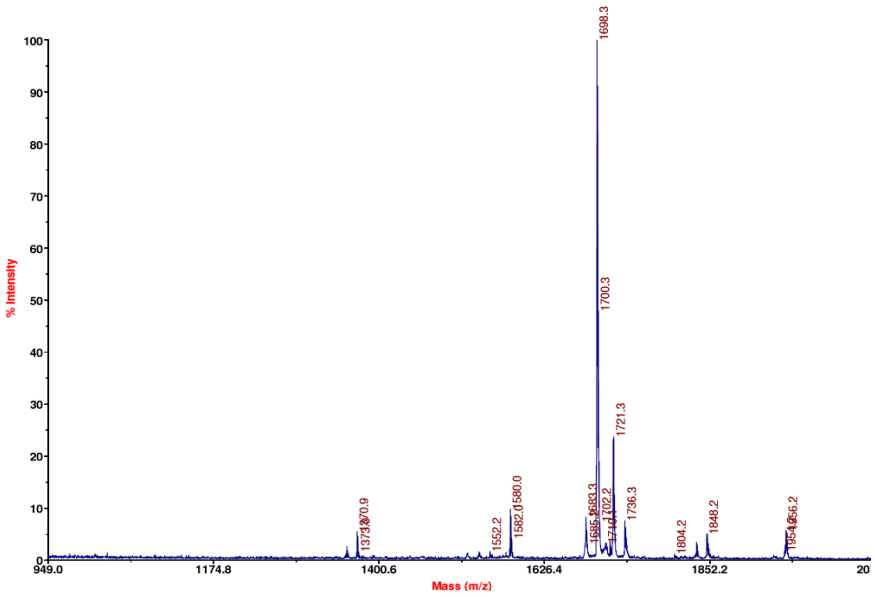
HPLC Spectrum: retention time is labeled on the peak.



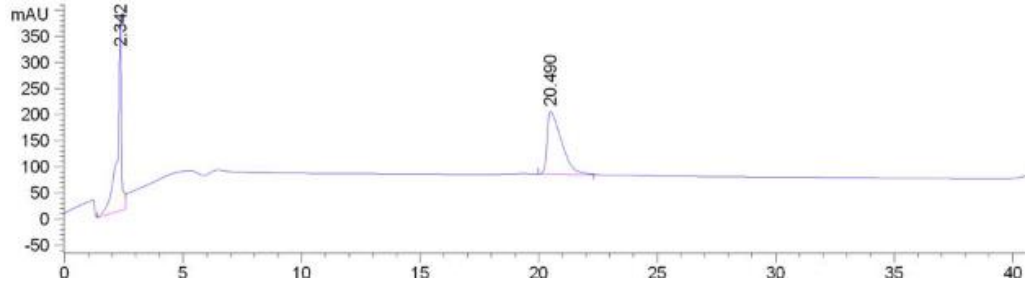
A14: $(KLK)_2$*$KKK_{BBA}Y$; MS(MALDI-TOF): Calcd for $C_{76}H_{135}BN_{20}O_{15}$ $[M\text{-}DHB + H]^+$: 1699.1, found: 1699.3; Purity (HPLC):93%.



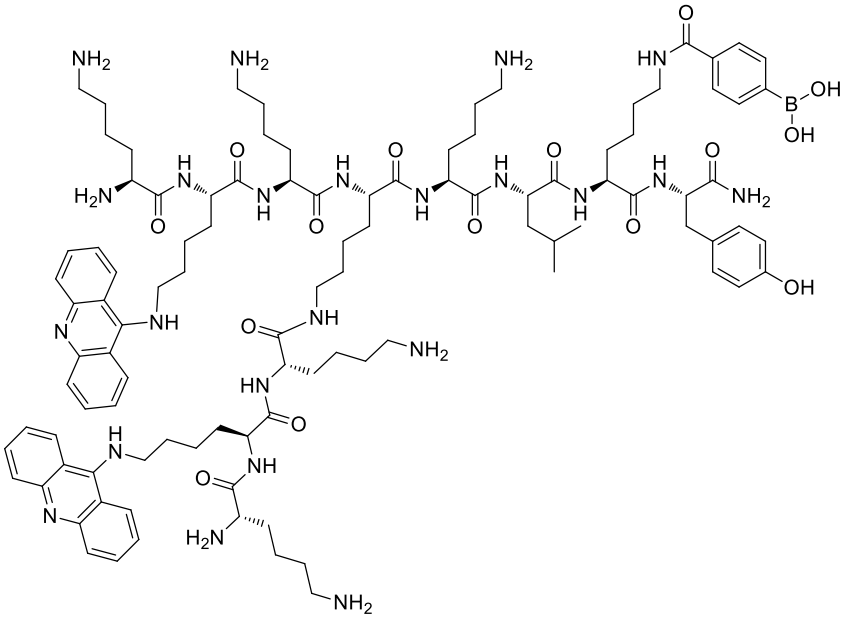
MALDI-TOF Spectrum:



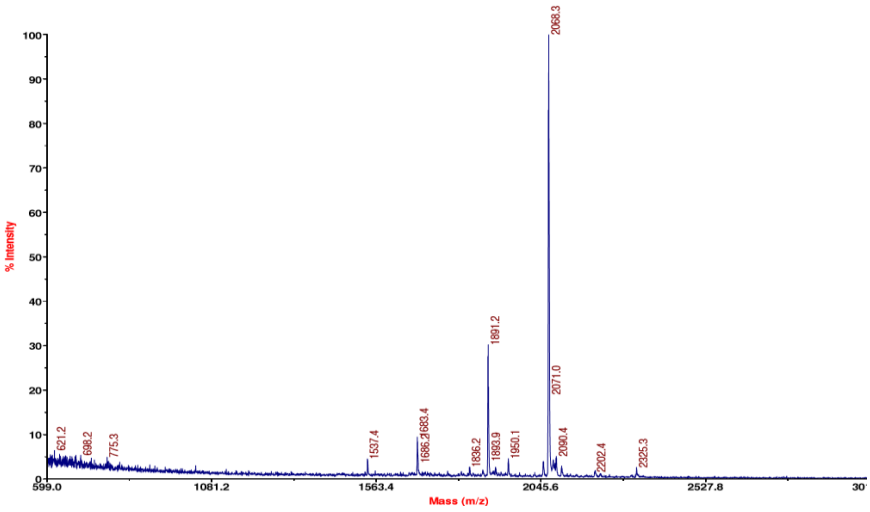
HPLC Spectrum: retention time is labeled on the peak.



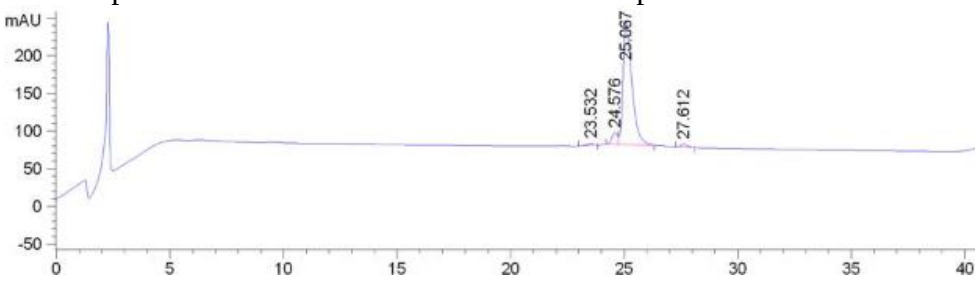
A15: $(KK_{ACR}K)_2$*$KLK_{BBA}Y$; MS(MALDI-TOF): Calcd for $C_{102}H_{150}BN_{23}O_{15}$ [M-DHB +H]$^+$: 2068.5, found: 2068.3; Purity (HPLC):89%.



MALDI-TOF Spectrum:



HPLC Spectrum: retention time is labeled on the peak.



Appendix B Citations of Copyrighted Works

Figure 1.4 Reprinted by permission from Elsevier: [Aboulela, F.; Karn, J.; Varani, G., The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J. Mol. Biol.* **1995,** *253* (2), 313-332.], Copyright 1995.

Figure 1.6 Reprinted by permission from Elsevier: [Mann, D. A.; Mikaélian, I.; Zemmel, R. W.; Gait, M. J.; Karn, J., A molecular rheostat: Co-operative Rev binding to stem I of the Rev-response element modulates human immunodeficiency virus type-1 late gene expression. *J. Mol. Biol.* **1994,** *241* (2), 193-207.], Copyright 1994.

Figure 1.8 Reprinted by permission from AAAS: [F Battiste, J. L.; Mao, H.; Frankel, A. D.; Williamson, J. R., α Helix-RNA major groove recognition in an HIV-1 Rev peptide-RRE RNA complex. *Science* **1996,** *273* (5281), 1547-1551.] , Copyright 1996.

Figure 1.9 Reprinted by permission from Nature Publishing Group: [Daugherty, M. D.; Liu, B.; Frankel, A. D., Structural basis for cooperative RNA binding and export complex assembly by HIV Rev. *Nat. Struct. Mol. Biol.* **2010,** *17* (11), 1337-42], Copyright 2010.

Figure 1.10 Reprinted by permission from Nature Publishing Group: [Daugherty, M. D.; Liu, B.; Frankel, A. D., Structural basis for cooperative RNA binding and export complex assembly by HIV Rev. *Nat. Struct. Mol. Biol.* **2010,** *17* (11), 1337-42], Copyright 2010.

Figure 1.11 Reprinted by permission from Elsevier: [Fang, X.; Wang, J.; Rein, A.; Wang, Y.-X., An Unusual Topological Structure of the HIV-1 Rev Response Element. *Cell* **2013,** *155* (3), 594-605], Copyright 2013.

Figure 1.12 Reprinted by permission from Elsevier: [Fang, X.; Wang, J.; Rein, A.; Wang, Y.-X., An Unusual Topological Structure of the HIV-1 Rev Response Element. *Cell* **2013,** *155* (3), 594-605], Copyright 2013.

Figure 1.38 Reprinted by permission from American Chemical Society: [Ennifar, E.; Aslam, M. W.; Strasser, P.; Hoffmann, G.; Dumas, P.; van Delft, F. L., Structure-guided discovery of a novel aminoglycoside conjugate targeting HIV-1 RNA viral genome. *ACS Chem. Biol.* **2013,** *8* (11), 2509-2517.], Copyright 2013.