DESIGN OF WELL-FOLDED β-HAIRPIN PEPTIDES FOR MOLECULAR RECOGNITION OF RNA AND IMPROVED RESISTANCE TO PROTEOLYSIS

Lauren Latshaw Cline

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry.

Chapel Hill
2009

Approved by:

Professor Marcey Waters

Professor Dorothy Erie

Professor Linda Spremulli

Professor Matthew Redinbo

Professor Nancy Allbritton

ABSTRACT

Lauren Latshaw Cline: Design of Well-Folded β-Hairpin Peptides for Molecular Recognition of RNA and Improved Resistance to Proteolysis
(Under the direction of Marcey L. Waters)

Ribonucleic acid (RNA) plays a vital role in many biological processes of the cell, which makes it an attractive target for drug discovery. In this work, small β-hairpin peptides were used to better understand RNA molecular recognition. Specifically, peptides were designed with the capability to bind via aromatic, electrostatic, and hydrogen bonding interactions to the single stranded RNA.

To investigate RNA recognition using a conjugated system of different binding motifs, a β-hairpin peptide was connected to an RNA intercalator. It was found that the individual parts did not bind well to RNA, but the conjugated system binds RNA with a dissociation constant of 3.7 μM. RNase footprinting experiments showed that the intercalator threads the stem region while the peptide interacts with an internal bulge region, and that the β-hairpin structure of the peptide is important for binding.

Combinatorial chemistry and *de novo* peptide design were used to improve binding of the BIV Tat (bTat) peptide for BIV TAR RNA (bTAR). A combinatorial library was designed which incorporated natural and unnatural residues in the N-terminal region of bTat with the purpose of inducing interactions with the bTAR hairpin loop region. Two peptides were isolated from fluorescent library screens, one which had one less charge than bTat and contained an unnatural homophenylalanine residue. It was determined by gel shift assays

that the two peptides bound equally as well to bTAR as the native sequence. In a separate project, strong cation-π side chain interactions were used to promote β-hairpin structure of a bTAR-binding peptide without the need for cyclization. The side-chain interactions were enough to stabilize the hairpin structure; however some RNA binding affinity was lost.

To successfully develop peptide drugs that bind RNA, the receptors need to be stable to proteolytic degradation. For this purpose, it was investigated whether stable β-hairpin peptides have increased resistance to proteolysis due to their structure. A series of peptides with ranging amount of thermodynamic stability were designed and digested using a variety of specific and non-specific proteases. It was determined that increased thermodynamic stability of the β-hairpin peptides does correlate to an increase in proteolytic stability.

ACKNOWLEDGEMENTS

I want to thank Professor Marcey Waters for being my advisor and my mentor. I admire her never-ending excitement for research and her enthusiasm for training graduate students. I truly appreciate all of the scientific advice and career advice that she has given me over the past five years.

I would also like to thank my committee members, Professor Linda Spremulli, Professor Dorothy Erie, Professor Matt Redinbo, and Professor Nancy Allbritton, and the rest of the UNC chemistry faculty. They have provided me with an excellent education. I appreciate their input on my research and the challenges they have provided in the classroom.

I want to say a big thank you to all the Waters lab members, past and present. I couldn't have asked for better co-workers and friends. Two things I will definitely miss are the Christmas party gift exchange and the weekly supply of delicious home-baked goods. I also want to thank the Thorp lab, especially Patty, Julie, and Dana. They have been great friends and lab-mates, graciously giving up their time and their space so that I could learn about the world of RNA.

I also want to thank Dr. David Klapper, Ashutosh Tripathy, the people at the Michael Hooker Microscopy Facility, and Dr. Sohrab Habibi. In one way or another they have all taught me the skills I needed to make my dissertation work possible. I appreciate their time and patience.

Finally, I want to thank my family, especially my husband Dan, for supporting me on this long, stressful, and ultimately rewarding endeavor. He is always there when I need him, no matter if I want to talk about science or if I want to escape science. Being with him always raises my spirit, makes things seem brighter, and motivates me to continue reaching higher.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SCHEMES

ABBREVIATIONS

A	Adenine
Ala, A	Alanine
Alloc	Allyloxycarbonyl
Arg, R	Arginine
Asn, N	Asparagine
ATP	Adenosine triphosphate
β-Ala	β-Alanine
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
BIV	Bovine immunodeficiency virus
Boc	t-Butoxycarbonyl
C	Cystosine
Cit	Citrulline
Cy-3	Cyanine-3
Cys, C	Cysteine
DCM	Dichloromethane
DEPC	Diethylpyrocarbonate
DIPEA	Disopropylethyl amine
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FMOC	N-9-Fluorenylmethyloxycarbonyl
G	Guanosine
Gln, Q	Glutamine
Glu, E	Glutamic acid
Gly, G	Glycine
HBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate

hF	Homophenylalanine
His, H	Histidine
HOBT	N-hydroxybenzotriazole
HPLC	High pressure liquid chromatography
Ile, I	Isoleucine
ivDde	1-(4,4-Dimethyl-2,6-dioxo-cyclohexylidene)-3-methyl-butyl
K(Me)2	Dimethyllysine
Leu, L	Leucine
Lys, K	Lysine
NBT	Nitro blue tetrazolium chloride
Orn, O	Ornithine
Pbf	2,2,4,6,7-Pentamethyl-dihydrobenzofurane-5-sulfonyl
PBS	Phosphate buffered saline
PDB	Protein data bank
Peg	Polyethylene glycol
pfF	Parafluorophenylalanine
Phe	Phenylalanine
PhG	Phenyl glycine
$_{D}$-Pro, p	$_{D}$-Proline
Pro, P	Proline
R(Me)2	Dimethyl arginine
RNA	Ribonucleic acid
SDS	Sodium Dodecyl Sulfate
Ser	Serine
TAMRA	5-(6)-Carboxytetramethylrhodamine
TAR	Transactivation response RNA
Tat	Transactivation Protein
tBu	t-Butyl
TFA	Trifluoroacetic acid

Thr, T	Threonine
TIPS	Triisopropyl silane
Tris	Tris(hydroxymethyl)aminomethane
Trp, W	Tryptophan
trt	Trityl
Tyr, Y	Tyrosine
U	Uracil
Val, V	Valine

Chapter I

INTRODUCTION

A. Significance of this work

Ribonucleic acid (RNA) plays several vital roles in the life cycle of the cell. The most prominent role is as messenger RNA (mRNA), which carrys genetic information from the DNA to ribosomes which produce proteins. RNA is also found as ribozymes, ribo-protein complexes (rRNA and snRNPs), gene silencers (siRNA), carriers of amino acids (tRNA), the primary genetic material of retroviruses, and possibly more roles that haven't been discovered yet.[1] Because RNA affects so many processes of the cell, it is now believed that targeting it may be an alternative to traditional drug discovery which targets proteins.[2] In order to design effective RNA ligands, it is imperative to understand the structure and dynamics of binding this complex biomolecule.[3] It is only then that we will be able to control cellular events at an RNA level.

---

[1] Campbell, M. K.; Farrell, S. O.; 6th ed.; Thompson Brooks/Cole: Belmont, CA, 2009.

[2] Melnikova, I. *Nat Rev Drug Discovery* **2007**, *6*, 863-864.

[3] (a) Tor, Y. *Angew Chem Int Ed Engl* **1999** *38*, 1579-1582. (b) Tor, Y. *Chembiochem* **2003**, *4*, 998-1007.

B. Targeting RNA

i. DNA vs. RNA

Molecular recognition of double stranded DNA structure has been well-studied and many synthetic molecules and protein motifs have been developed which bind sequence specifically to DNA. These interactions generally occur by base pairing through Hoogsten interactions in the major and minor grooves.[4] On the other hand, double stranded RNA is not as amenable to molecular recognition through hydrogen bonding in the groove. Due to the 2'hydroxyl groups, double stranded RNA exists in an A-form helix. The minor groove of an A-form helix is wide and shallow and cluttered with 2'hydroxyl groups and the major groove is narrow and deep and sterically hindered by phosphate groups (Figure 1.1).[5]

Unlike DNA, RNA often folds into a complex tertiary structure, which commonly contains exposed loops, bulges, and psuedoknots (Figure 1.2). Many structures have been solved in which single stranded regions of RNA are recognized by proteins through a combination of aromatic, electrostatic, and hydrogen bonding interactions.[6] It is through these interactions that proteins can recognize specific RNA sequences. If we can design molecules that target the single stranded regions of RNA, it may also provide selectivity for binding a particular RNA.

---

[4] (a) Wemmer, D. E. *Annu Rev Biophys Biomol Struct* **2000**, 29, 439-461. (b) Segal, D. J.; Barbas, C. F., 3rd *Curr Opin Chem Biol* **2000**, *4*, 34-9.

[5] Carlson, C. B.; Stephens, O. M.; Beal, P. A. *Biopolymers* **2003**, *70*, 86-102.

[6] (a) Nagai, K. *Curr Opin Chem Biol* **1992**, *2*, 131-137. (b) Deo, R. C.; Bonanno, J. B.; Sonenberg, N.; Burley, S. K. *Cell* **1999**, *98*, 835-45. (c) Handa, N.; Nureki, O.; Kurimoto, K.; Kim, I.; Sakamoto, H.; Shimura, Y.; Muto, Y.; Yokoyama, S. *Nature* **1999**, *398*, 579-85. (d) Price, S. R.; Evans, P. R.; Nagai, K. *Nature* **1998**, *394*, 645-50. (e) Legault, P.; Li, J.; Mogridge, J.; Kay, L. E.; Greenblatt, J. *Cell* **1998**, *93*, 289-99. (f) Oubridge, C.; Ito, N.; Evans, P. R.; Teo, C. H.; Nagai, K. *Nature* **1994**, *372*, 432-8.

It has also been noted that upon ligand binding, induced fit interactions are more extreme for RNA rather than DNA.[7] This makes it even more difficult to rationally design an RNA binding ligand. For this reason, scientists have found creative methods, such as combinatorial chemistry, to discover novel compounds. These methods offer a high-throughput way of screening large numbers of compounds, and it allows the system to determine the best ligand.

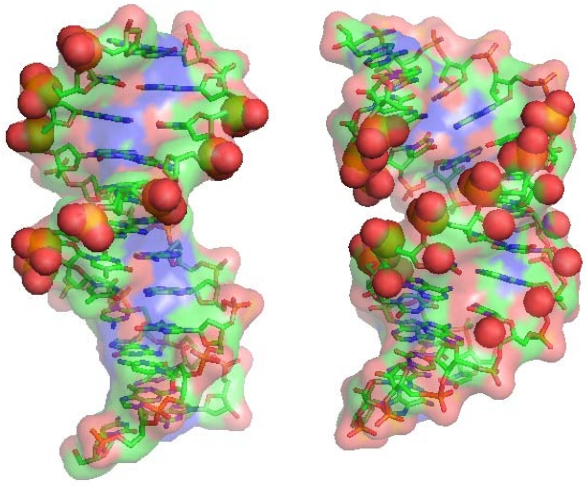


Figure 1.1. Space filling models of the structures of duplex DNA and RNA. Highlighted in red are the phosphate groups on DNA and both the phosphate groups and the 2'hydroxyl groups of RNA.

[7] Cheng, A. C.; Calabro, V.; Frankel, A. D. *Curr Opin Struct Biol* **2001**, *11*, 478-84.

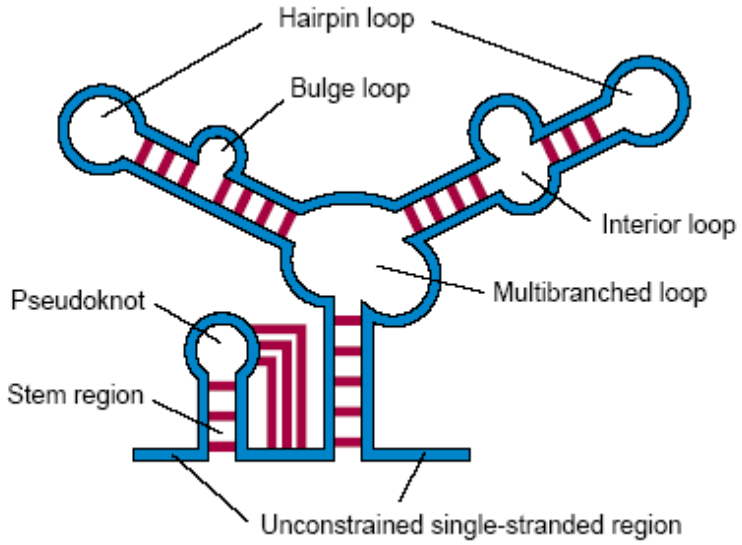


Figure 1.2. Possible types of structural features found in RNA tertiary structure.[8]

ii. Molecular recognition of RNA: An emerging field

Despite the challenges of designing molecules that bind RNA, for the efforts toward drug discovery RNA may be an even better target than DNA or proteins. For molecular recognition of DNA sequence selectivity is imperative because it exists primarily as a double stranded helix. However RNA's tertiary structure offers regions that could be targeted in either a structure or sequence specific manner. RNA also has the advantage of being found both inside and outside the nucleus, making it more accessible to designed ligands. In composition, RNA resembles DNA, but in many ways RNA molecules are more like proteins. They are similar in that they both form complex structures that form pockets and surfaces available for binding.[7] Yet the chemical composition is less complex in RNA than in proteins, because there are only four RNA bases as opposed to 20 natural amino acids.

[8] Reviewed in: Zamen, G. J. R.; Michiels, P. J. A.; van Boeckel, C. A. A. *Drug Discovery Today* **2003**, *8*, 297-306.

This may simplify the design of ligands because there are fewer types of chemical moieties exposed on the surface, making RNA an even better target than proteins.

However, the field of RNA molecular recognition has significantly lagged behind studies of DNA and protein recognition due mainly to the lack of structural information. Before 1992, the only RNA-protein crystal structures were of tRNAs in aminoacyl-tRNA synthetase complex and the 50S subunit.[6a] Since then, there have been an abundance of crystal structure and NMR structures of RNA and RNA-protein complexes published in the PDB. This sudden increase of information has intensified the search for RNA molecules that could be effective medicinal targets. One class of molecules that is actively studied is viral RNA sequences such as the trans-activating response (TAR) RNA and the Rev response element (RRE) both from the HIV-1 genome. Other attractive targets that are currently being studied are ribozymes, mRNAs, microRNAs, and rRNAs.[9]

There are now many different synthetic molecules found that bind well to RNA[9] such as aminoglycosides,[10] polyamines,[11] deoxystreptamine dimers,[12] peptidomimetics,[13] peptides,[14]

---

[9] Reviewed in: Thomas, J. R.; Hergenrother, P. J. *Chem Rev* **2008**, *108*, 1171-224.

[10] Silva, J. G.; Carvalho, I. *Curr Med Chem* **2007**, *14*, 1101-19.

[11] Lawton, G. R.; Appella, D. H. *J Am Chem Soc* **2004**, *126*, 12762-3.

[12] (a) Liu, X.; Thomas, J. R.; Hergenrother, P. J. *J Am Chem Soc* **2004**, *126*, 9196-7. (b) Thomas, J. R.; Liu, X.; Hergenrother, P. J. *J Am Chem Soc* **2005**, *127*, 12434-5. (c) Thomas, J. R.; Liu, X.; Hergenrother, P. J. *Biochemistry* **2006**, *45*, 10928-38.

[13] (a) Athanassiou, Z.; Dias, R. L.; Moehle, K.; Dobson, N.; Varani, G.; Robinson, J. A. *J Am Chem Soc* **2004**, *126*, 6906-13. (b) Leeper, T. C.; Athanassiou, Z.; Dias, R. L.; Robinson, J. A.; Varani, G. *Biochemistry* **2005**, *44*, 12362-72. (c) Athanassiou, Z.; Patora, K.; Dias, R. L.; Moehle, K.; Robinson, J. A.; Varani, G. *Biochemistry* **2007**, *46*, 741-51. (d) Burns, V. A.; Bobay, B. G.; Basso, A.; Cavanagh, J.; Melander, C. *Bioorg Med Chem Lett* **2008**, *18*, 565-7. (e) McNaughton, B. R.; Gareiss, P. C.; Miller, B. L. *J Am Chem Soc* **2007**, *129*, 11306-7.

and small intercalating agents[15]. One of these classes of molecules has already been proven to be an effective drug. Aminoglycosides make up a family of antibiotics that target RNA, specifically the 16S rRNA subunit of the ribosome.[8] Two things to be learned from this application are that (1) effectiveness can be attained by either selectivity against size of binding pocket or the sequence of binding pocket, and (2) even moderate binding affinities in the micromolar range can make effective drugs. Most aminoglycosides work by targeting the bacterial ribosomes using hydrogen bonding and electrostatic interactions with an interior loop of the RNA.[8] The aminoglycosides do not affect 12S human mitochondrial rRNA because the human rRNA has a larger interior loop at the same site and it doesn't bind the 18S human cytoplamic rRNA because of differences in sequence. The binding affinities of the aminoglycosides for the RNA range from 0.1 μM to 10 μM.[16] Despite the moderate binding affinity, this is tight enough to be effective. This offers hope that even with a moderate binding affinities RNA can be a viable therapeutic target.

---

[14] (a) Hyun, S.; Kim, H. J.; Lee, N. J.; Lee, K. H.; Lee, Y.; Ahn, D. R.; Kim, K.; Jeong, S.; Yu, J. *J Am Chem Soc* **2007**, *129*, 4514-5. (b) Pustowka, A.; Dietz, J.; Ferner, J.; Baumann, M.; Landersz, M.; Konigs, C.; Schwalbe, H.; Dietrich, U. *Chembiochem* **2003**, *4*, 1093-7. (c) Raja, C.; Ferner, J.; Dietrich, U.; Avilov, S.; Ficheux, D.; Darlix, J. L.; de Rocquigny, H.; Schwalbe, H.; Mely, Y. *Biochemistry* **2006**, *45*, 9254-65. (d) Kawakami, J.; Sugimoto, N.; Tokitoh, H.; Tanabe, Y. *Nucleosides Nucleotides Nucleic Acids* **2006**, *25*, 397-416. (e) Austin, R. J.; Xia, T.; Ren, J.; Takahashi, T. T.; Roberts, R. W. *J Am Chem Soc* **2002**, *124*, 10966-7.

[15] (a) Carlson, C. B.; Beal, P. A. *Bioorg Med Chem Lett* **2000**, *10*, 1979-82. (b) Carlson, C. B.; Spanggord, R. J.; Beal, P. A. *Chembiochem* **2002**, *3*, 859-65.

[16] (a) Wong, C. H.; Hendrix, M.; Manning, D. D.; Rosenbohm, C.; Greenberg, W. A. *J Am Chem Soc* **1998**, *120*, 8319-8327. (b) Griffey, R. H.; Hofstadler, S. A.; Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T. *Proc Nat Acad Sci USA* **1999**, *96*, 10129-10133. (c) Carter, A. P.; Clemons, W. M.; Brodersen, D. E.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V. *Nature* **2000**, *407*, 340-348. (d) Wang, Y.; Hamasaki, K.; Rando, R. R. *Biochemistry* **1997**, *36*, 768-779.

iii. Current state of research

RNA molecular recognition is still in its early stages of discovery, which makes it an exciting and important field to study. Although many classes of small molecules have been discovered that bind RNA, only moderate binding affinities have been achieved and specificity is a significant challenge. Aminoglycosides are still the only class of RNA drugs, because most other compounds fail to show efficacy in cell culture or animal studies.[9] The "holy grail" of this field would be the discovery of a set of rules for designing ligands that are sequence or structure specific for RNA.

C. RNA-protein interactions

i. Classes of RNA binding proteins

In order to be able to design molecules that bind to RNA, it is imperative to understand how nature has designed RNA-binding proteins. There are several motifs that are commonly found in RNA-binding proteins, some of which will be highlighted below because of their relevance to this work. By studying RNA binding motifs it is possible to learn what types of interactions that are important for affinity and the way specificity is achieved.

The RNA recognition motif, RRM, is a binding domain that is found in a majority of RNA-binding proteins in higher vertebrates.[17] It is characterized by a mixed α/β structure made up of approximately 100 amino acids. The interactions with RNA are primarily located on the β-sheet portion of the protein. The RRM domain has been found to bind anywhere from 2-8 nucleotides simultaneously with 4 being the most common number. The binding occurs when the nucleotides stack with three conversed aromatic residues on the

[17] Reviewed in: (a) Cheng, A. C.; Calabro, V.; Frankel, A. D. *Curr Opin Struct Biol* **2001**, *11*, 478-84. (b) Auweter, S. D.; Oberstrass, F. C.; Allain, F. H. *Nucleic Acids Res* **2006**, *34*, 4943-59. (c) Messias, A. C.; Sattler, M. *Acc Chem Res* **2004**, *37*, 279-87.

protein. One example of this interaction is the small nuclear ribonucleoprotein U1A (U1 snRNP) binding to a hairpin loop region of RNA derived from U1 snRNA (Figure 1.3).[18] It can be noted from the structure that the RNA bases in the loop are flipped out in solution so as to be accessible for protein binding. The protein recognizes the RNA through stacking interactions between aromatic residues in the β-sheet domain and the nucleotide bases.

Zinc finger motifs are frequently found in proteins that bind to DNA sequence selectively, but they have also been found in RNA binding proteins. The zinc ions coordinate to an an α-helix and β-sheet region of the peptide that is made up of approximately 30 residues.[7] Like the RRM motif, if has been found that aromatic residues are important for the binding of zinc finger domains. The zinc ion is required to maintain protein structure but it is the conserved aromatic residues that create a binding pocket for exposed RNA bases. For example in the crystal structure for MMLV nucleocapsid protein, a tyrosine residue (Tyr28) stacks between a uridine (U306) and a cytosine (C307) (Figure 1.4).[19] A second conserved aromatic residue is found nearby, Trp35, and it stacks with a guanosine (G309).

For both RRM and zinc finger domains the pre-formed structure is very important for binding as are aromatic interactions. However in the arginine rich motif (ARM), short, unstructured, arginine-rich peptides (~20 residues) fold into a variety of secondary structures upon binding to RNA.[17a] Arginine is well equipped to bind RNA because of the guandinium group on the side chain.

---

[18] Oubridge, C.; Ito, N.; Evans, P. R.; Teo, C. H.; Nagai, K. *Nature* **1994**, *372*, 432-8.

[19] D'Souza, V.; Summers, M. F. *Nature* **2004**, *431*, 586-90.

One example of an ARM motif is the bovine immunodeficiency virus (BIV) Tat peptide that recognizes BIV TAR RNA (bTAR). Unlike the human version of this complex, HIV-1 bTat/TAR, the NMR structure of bTat/bTAR has been solved (Figure 1.5).[20] The peptide is found to be unstructured in solution, but when bound to bTAR RNA it adopts a β-hairpin structure and lies in the major groove.[21] The specificity in this system originates from hydrogen bonding between the arginine residues and guanine bases in the major groove. Also important for this interaction is the hydrophobic packing of isoleucine and glycine on the peptide. This example provides evidence that a β-hairpin peptide has the ability to fit in the RNA groove.

[20] Ye, X.; Kumar, R. A.; Patel, D. J. *Chem Biol* **1995**, *2*, 827-40.

[21] Reviewed in: Patel, D. J. *Curr Opin Struct Biol* **1999**, *9*, 74-87.

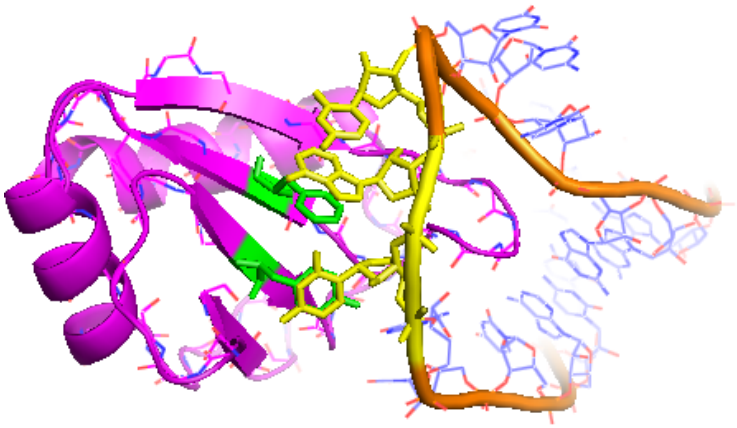


Figure 1.3. Crystal structure of the RRM domain of U1A snRNP binding to an RNA hairpin loop.[18] The RNA phosphate backbone is colored orange with the important hairpin loop residues highlighted in yellow. The protein backbone is colored magenta, and the aromatic residues that participate in binding RNA are highlighted in green.

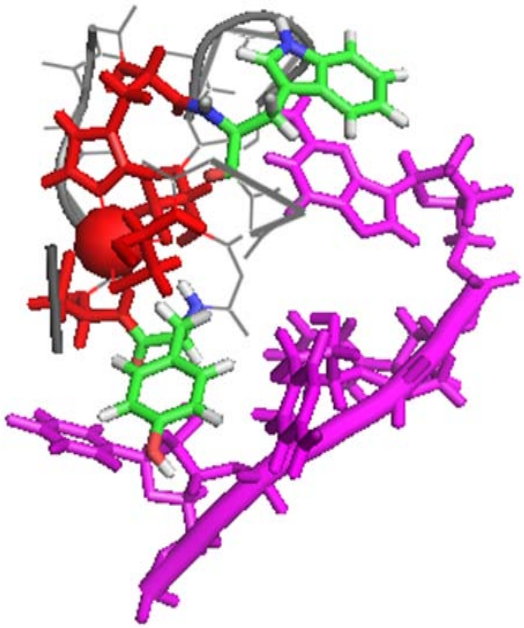


Figure 1.4. Crystal structure of the zinc finger binding domain of MMLV nucleocapsid protein binding to RNA (PDB 1U5P).[19] The zinc coordination site is labeled red, the RNA is colored magenta, and the two conserved aromatic residues, Tyr28 and Trp 35, are highlighted in green.

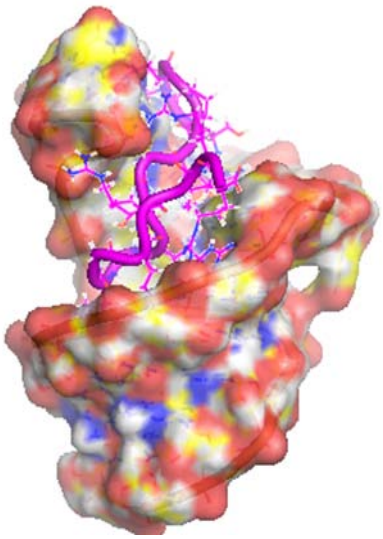


Figure 1.5. NMR structure of BIV TAR RNA stem-loop bound to Tat peptide (PDB 1BIV).[21] The C-terminus of the peptide is shown in the foreground with the β-hairpin extending to the bottom left. The N-terminus loop is near the RNA hairpin loop at the top of the picture.

ii. Types of interactions involved

By examining the specific examples above, it is clear that hydrogen bonding, electrostatics, and aromatic interactions are important in protein-nucleic acid recognition. It is well-known that polar interactions, specifically hydrogen bonding, are responsible for the complementary base pairing in double stranded RNA, but they are also the basis of specificity in protein binding. It is energetically favorable for hydrogen bonding to occur as polar residues are buried in the RNA-protein interface. Using statistical analysis, it has been found that 90% of the time the hydrogen bond donor group is the protein whereas the acceptor is the nucleic acid.[22] The analysis also revealed that the negatively charge phosphate groups make excellent hydrogen bond acceptors, but about 20% of the time the

[22] Nadassy, K.; Wodak, S. J.; Janin, J. *Biochemistry* **1999**, *38*, 1999-2017.

hydrogen bonds occur with the 2'-hydroxyl group. For proteins the most common hydrogen bond donors are lysines, arginines, and amide protons. It has been estimated that the energy for a neutral hydrogen bond is between 2.1-6.3 kJ/mol, but a hydrogen bond where one of the species is charged can yield 15-19 kJ/mol.[23] This amount of energy can add up to be a substantial amount when there is a network of hydrogen bonds occurring at the protein-nucleic acid interface. This is believed to be the source of sequence selectivity.

Electrostatic interactions are equally as important, but they contribute more toward binding affinity rather than sequence selectivity. Statistical analysis shows that the surfaces of proteins that bind RNA are primarily positively charged, and it is extremely rare to see an acidic residue at the RNA interface.[23] In fact an RNA-protein binding event is believed to release into solution a large number of counterions that are associated with the phosphate backbone and other charged moieties. Therefore it is entropically favorable for electrostatic interactions to form between the RNA and the protein. These interactions are believed to contribute greatly to the overall free energy of binding.[24]

The aromatic composition of the RNA bases makes π-stacking interactions possible. This interaction is partially responsible for the stability of double stranded RNA, but in single stranded RNA the unstacked bases are free to interact with proteins. Statistical analysis reveals that all the bases are able to stack with aromatic amino acids, however phenylalanine is most commonly found.[25] The "jack-of-all-trades" amino acid, arginine, can also

---

[23] Fersht, A. R.; Shi, J. P.; Knill-Jones, J.; Lowe, D. M.; Wilkinson, A. J.; Blow, D. M.; Brick, P.; Carter, P.; Waye, M. M.; Winter, G. *Nature* **1985**, *314*, 235-8.

[24] Reviewed in: (a) Bloomfield, V. A.; Crothers, D. M.; Tinoco, J. I. *Nucleic Acids: Structures, Properties, and Functions*; University Science Books: Sausalito, CA, 2000. (b) Auweter, S. D.; Oberstrass, F. C.; Allain, F. H. *Nucleic Acids Res* **2006**, *34*, 4943-59.

[25] Allers, J.; Shamoo, Y. *J Mol Biol* **2001**, *311*, 75-86.

participate in aromatic interactions with RNA. This protonated residue can form a cation-π interaction with the electron-rich π-cloud of the aromatic bases. These interaction have also been observed with lysines and histidines.[25b]

Stacking interactions are particularly common in loop regions where bases are more solvent exposed. As was evident with the RRM motifs and the zinc binding domains, the single stranded regions provide good targets for protein binding. They also may be good targets for designing inhibitors of protein-RNA interactions.

iii. Using β-hairpin peptides to mimic β-sheet structure

We have investigated mimicking β-sheet interactions with single stranded regions of RNA by using small β-hairpin peptides that are well-folded in aqueous solution. The use of β-hairpin peptides as model systems dates back to 1993 when the first monomeric, water-soluble β-hairpin was discovered.[26] This discovery launched a field of research focused on determining the rules for designing small peptide systems that accurately portray the large β-sheet domains of peptides. These model systems can not only be used to study the intricacies of biomolecular interactions, they can also be used to design inhibitors of those interactions.

In these past 15 years, it has been determined that there are three main contributors to the stability of β-hairpin structure, (1) turn sequence, (2) side chain interactions, and (3) amino acid β-sheet propensity. In the following projects, these characteristics were optimized in order to find the best peptides for binding RNA. One way we have optimized the turn sequence was to investigate both Asn-Gly and $_{D}$-Pro-Gly sequences which have different effects when incorporated into a β-hairpin peptide. The Asn-Gly turn is type I' and is

[26] Reviewed in: Hughes, R. M.; Waters, M. L. *Curr Opin Struct Biol* **2006**, *16*, 514-24.

commonly found in the β-hairpin loops of protein crystal structures.[27] The $_{D}$Pro-Gly turn has been found to be an even better promoter of β-hairpin formation which adopts a type II' turn.[28] Due to the unnatural chirality of this turn sequence it is not found in native proteins, but the proline residue provides rigidity to the turn region.

We also optimized the side chain interactions in several of the following projects. Both cation-π and π-π stacking interactions were used to stabilize the peptide structures by incorporating aromatic tryptophan residues and cationic lysine residues into non-hydrogen bonded positions. This design strategy has been extensively studied in the past and is shown to yield some of the most stable β-hairpin peptides published to date.[29] In Chapter IV, we used the two different faces of a β-hairpin peptide for two different roles. The face displaying the hydrogen bonded residues was designed to bind RNA by inserting the amino acids known to bind bTAR RNA. The non-hydrogen bonded face contained residues that were able to make side chain-side chain interactions which stabilized the overall structure.

The rules for designing well-folded β-hairpin peptides are known and it has been shown that β-sheet proteins bind well to RNA, but it is still difficult to design β-hairpin peptides which bind RNA as well as proteins. One interesting observation regarding how peptides and proteins bind RNA is that there is a difference between their modes of binding. Peptides tend to alter their conformation upon binding RNA but the RNA seems to stay rigid, whereas large proteins seem to stay rigid when they bind and the RNA undergoes conformational

---

[27] Hutchinson, E. G.; Thornton, J. M. *Protein Sci* **1994**, *3*, 2207-16.

[28] Stanger, H. E.; Gellman, S. H. *J Am Chem Soc* **1998**, *120*, 4236-4237.

[29] (a) Riemen, A. J.; Waters, M. L. *Biochemistry* **2009**, *48*, 1525-31. (b) Hughes, R. M.; Waters, M. L. *J Am Chem Soc* **2005**, *127*, 6518-9. (c) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc Natl Acad Sci U S A* **2001**, *98*, 5578-83.

changes.[30] If this is true, there are two options for design. Either design a peptide that is so stable the RNA will conform to it, similar to what occurs with proteins, or design a peptide that has the potential to be both well-folded and flexible. In this case, the peptide will have a pre-defined binding pocket, but upon RNA binding will be able to adapt its conformation to fit. We have attempted to do both in the projects herein.

iv. Previously designed peptide-based systems which bind RNA

There have been other published reports of groups developing peptide-based systems to bind to the single stranded regions of RNA by taking advantage of the same interactions used in RNA-binding proteins. Austin et al. designed an arginine-rich α-helical peptide which was modeled after the $P22_{N21}$ and the λ_{N22} proteins that bind to hairpin loop regions of P22boxB and λboxB RNA, respectively.[31] The conjugate peptide combined the native stability of the $P22_{N21}$-P22boxB complex with the specific loop-binding structure of the λ_{N22} protein. The mutant λ_{N22} peptides increased the stability of the peptide-RNA complex and achieved picomolar binding affinity. Kawakami et al. used *in vitro* selection methods to find RNA hairpin loop sequences that bound to a model peptide from lambda N protein using hydrogen bonding and π-π stacking interactions.[32] The consensus sequence for the RNA hairpin loop selected from the library, -GCUAA-, creates a novel peptide-RNA pair which has higher specificity although slightly lower affinity than the native RNA sequence.

Many groups have used combinatorial methods and found that peptides containing aromatic-rich sequences where commonly selected. These residues could potentially make

---

[30] Hermann, T. *Angew Chem Int Ed Engl* **2000**, *39,* 1890-1904.

[31] Austin, R.; Xia, T.; Ren, J.; Takahashi, T.; Roberts, R. *J Am Chem Soc* **2002,** *124,* 10966-10967.

[32] Kawakami, J.; Sugimoto, N.; *Nucleosides, Nucleotides Nucleic Acids* **2006,** *25,* 397-416.

π-π or hydrogen bonding interactions with flexible bases in single stranded regions of RNA. Dietrich and coworkers have used phage display technologies to discover peptide ligands that bind specifically to the RNA packaging structure psi (ψ) in HIV-1.[33] The resulting hits were aromatic-rich peptides which resembled the natural ligand, NCp7, which uses aromatic residues to interact with the hairpin loop RNA. In a related study, Mély and coworkers used one of the selected peptides, Ac-HKWPWW-NH_2, and studied the affinity and interactions involved in binding TAR and PBS RNA sequences.[34] They determined that the peptide uses stacking interactions between the tryptophan residues and the RNA to bind with low micromolar affinity. Burns et al. have found a cyclic peptide system using combinatorial methods that binds to bTAR RNA.[35] Both of the peptides that were further characterized contained a tryptophan residue and one or two histidine residues. Computer modeling of the system suggests that contacts are being with the internal bulge region of bTAR. The authors hypothesize that the peptides are making specific interactions due to fact that the sequences are not polycationic.

D. Purpose of this work

We have investigated the important features of RNA-protein interactions by studying minimized systems involving peptides binding to RNA. We set out to discover how these interactions can be used to design novel substrates for RNA. We proposed to do this using a three-fold approach, (1) designing a conjugated binding system with the potential to bind

[33] Pustowka, A.; Dietz, J.; Ferner, J.; Baumann, M.; Landersz, M.; Konigs, C.; Schwalbe, H.; Dietrich, U. *ChemBioChem*, **2003,** *4,* 1093-1097.

[34] Raja, C.; Ferner, J.; Dietrich, U.; Avilov, S.; Ficheux, D.; Darlix J.; de Rocquigny, H.; Schwalbe, H.; Mély, Y.; *Biochemistry*, **2006,** *45,* 9254-9265.

[35] Burns, V.; Bobay, B.; Basso, A.; Cavanagh, J.; Melander, C. *Bioorg Med Chem Lett* **2008,** *18,* 565-567.

using two different modes, (2) utilizing combinatorial chemistry, and (3) incorporation of side-chain interactions to stabilize the structure of an RNA binding peptide.

The ultimate goal for our studies would be to design an inhibitor for medicinal applications. That is difficult to accomplish when working with peptides, because of how quickly they are degraded in the body. With that in mind, we also investigated whether β-hairpin peptides, similar to the ones used to bind RNA, have increased resistance to proteolysis because they are well-structured.

Chapter II

DESIGN OF A CONJUGATED β-HAIRPIN PEPTIDE-INTERCALATOR PROBE FOR BINDING RNA

A. Background

In our first approach to designing peptides that bind to RNA, we have designed a multivalent system which uses both intercalation and non-covalent interactions for the recognition of multiple regions of a single RNA strand. The advantage of using this system is that the individual parts do not bind with high affinity, but when linked together the affinity is greatly increased as is the specificity.

i. Previous work using conjugated systems

As is often seen in biology, proteins are multivalent receptors making numerous weak interactions to a target which add together to provide overall high binding affinity.[1] Several groups have tried to mimic this by designing molecules that combine several binding motifs in an attempt to improve RNA binding affinity. Kirk et al. covalently linked an aminoglycoside, neomycin, to an intercalating agent, acridine, and found that the

[1] Mammen, M.; Choi, S. K.; Whitesides, G. M. *Angew Chem Int Ed* **1998**, *37*, 2755-2794.

multivalent molecule had a much higher binding affinity to RRE RNA than the individual parts, and the neomycin-acridine conjugate could competitively inhibit the natural receptor, Rev peptide.[2] Kaiser et al. also synthesized an aminoglycoside-quinacridine intercalator conjugate and found improved binding to telomerase RNA.[3] Both of these groups used aminoglycosides and intercalators, which were relatively promiscuous binders. Kirk et al. found that their complex binds near an internal bulge in RRE and could effectively inhibit the native Rev peptide from binding. However, they did not screen against any other RNA sequence containing a bulge to investigate its specificity for that target. Likewise, Kaiser et al. found that their complexes required a loop region to bind to the duplex RNA but they did not test different loop sequences to show selectivity.

Several groups have also found success by attaching peptides to RNA intercalators. In one example, a crystal structure was determined for an ethidium-arginine conjugate bound to TAR RNA of HIV-1.[4] Yu and coworkers have designed several multivalent systems, including a neomycin-peptide nucleic acid conjugate[5], an aminoglycoside-small molecule conjugate[6], and a neomycin-dipeptide conjugate.[7] Another molecule designed by the Yu lab is particularly relevant to this project because it uses an intercalator and a structured peptide.

---

[2] Kirk, S. R.; Luedtke, N. W.; Tor, Y. *J Am Chem Soc* **2000**, *122*, 980-981.

[3] Kaiser, M.; Sainlos, M.; Lehn, J. M.; Bombard, S.; Teulade-Fichou, M. P. *Chembiochem* **2006**, *7*, 321-329.

[4] Peytou, V.; Condom, R.; Patino, N.; Guedj, R.; Aubertin, A. M.; Gelus, N.; Bailly, C.; Terreux, R.; Cabrol-Bass, D. *J Med Chem* **1999**, *42*, 4042-4053.

[5] Hyun, S.; Lee, K. H.; Yu, J. H. *Bioorg Med Chem Lett* **2006**, *16*, 4757-4759.

[6] Lee, J. K.; Kwon, M. Y.; Lee, K. H.; Jeong, S. J.; Hyun, S.; Shin, K. J.; Yu, J. H. *J Am Chem Soc* **2004**, *126*, 1956-1957.

[7] Ahn, D. R.; Yu, J. *Bioorg Med Chem* **2005**, *13*, 1177-83.

They synthesized a helical, polycationic peptide and attached acridine molecules at various locations throughout the peptide.[8] The peptide interacts with the RNA in the stem regions with additional π-π interactions being made by the acridine molecule. The addition of two acridine molecules resulted in peptides with dramatic increases in the binding affinities to TAR RNA (170-fold increase) and RRE RNA (40-fold increase). This group extensively investigated the structure and sequence specificity of their helical peptide-acridine conjugates. They found that the binding of their complexes significantly decreased when they removed the bulge regions, but was not affected by the size and shape of the loop regions.

ii. Goal of this project

We wanted to develop a two-part probe that binds RNA in both the single stranded loop region and the double stranded stem regions simultaneously. By linking the two parts, we wanted to increase the overall affinity and induce structural selectivity. To do this we designed a conjugated system that contains an RNA intercalator connected to a well-folded β-hairpin peptide through a flexible linker. The peptide portion used in these studies was the WKWK peptide, which is named for the Trp-Lys-Trp-Lys residues on one face of the hairpin. This peptide has been shown to bind ATP with a K_D of 170 μM and as a dimer bind single stranded DNA (ssDNA) with low micromolar affinity. This peptide has a high propensity to form a β-hairpin and has been characterized by NMR to be 95% folded in an aqueous environment.[9] From mutational studies, it is hypothesized that ATP interacts with

[8] Lee, Y.; Hyun, S.; Kim, H. J.; Yu, J. *Angew Chem Int Ed Engl* **2007**, *46*, 1-5.

[9] (a) Butterfield, S. M.; Waters, M. L. *J Am Chem Soc* **2003**, *125*, 9580-1. (b) Butterfield, S. M.; Sweeney, M. M.; Waters, M. L. *J Org Chem* **2005**, *70*, 1105-14. (c) Butterfield, S. M.; Cooper, W. J.; Waters, M. L. *J Am Chem Soc* **2005**, *127*, 24-5.

the two tryptophan residues through π-π stacking and also makes electrostatic interactions with the lysine side chains.[9]

The intercalator portion used in these studies was the quinoline-based RNA intercalator developed by Beal et al. to intercalate into double stranded regions of RNA.[10] This intercalator and a related acridine intercalator have been found to preferentially thread between two G-C base pairs which are directly adjacent to bulge sites.[10, 11] The Beal lab synthesized a small five residue peptide, aminobutyric acid-Ser-Val-*Quin*-Arg, which contained this intercalator (*Quin*) as an unnatural amino acid and footprinting experiments found that it bound to a specific RNA sequence with low micromolar affinity. We chose to use this particular intercalator due its site specificity and ease of incorporation into peptide synthesis.

These two pieces utilize two different methods for binding the RNA, and by connecting them together we expected to gain a more selective system with a tighter binding affinity due to cooperatively. The WKWK peptide should target single, unpaired bases in the single stranded regions of RNA, while the intercalator will orient the complex by binding a specific site in the duplex RNA. Our results indicate that the use of a multivalent system greatly improves the binding affinity as compared to the individual parts. We also determined that the peptide makes specific contacts with at least one RNA base and protects it from RNase cleavage.

---

[10] Krishnamurthy, M.; Gooch, B. D.; Beal, P. A. *Org Biomol Chem* **2006**, *4*, 639-45.

[11] (a) Carlson, C. B.; Beal, P. A. *Bioorg Med Chem Lett* **2000**, *10*, 1979-82. (b) Carlson, C. B.; Vuyisich, M.; Gooch, B. D.; Beal, P. A. *Chem Biol* **2003**, *10*, 663-72. (c) Gooch, B. D.; Beal, P. A. *J Am Chem Soc* **2004**, *126*, 10603-10. (d) Krishnamurthy, M.; Gooch, B. D.; Beal, P. A. *Org Lett* **2004**, *6*, 63-6. (e) Gooch, B. D.; Krishnamurthy, M.; Shadid, M.; Beal, P. A. *Chembiochem* **2005**, *6*, 2247-54. (f) Krishnamurthy, M.; Simon, K.; Orendt, A. M.; Beal, P. A. *Angew Chem Int Ed Engl* **2007**, *46*, 7044-7.

B. Results and discussion

i. General design

We have designed a conjugated probe that contains a β-hairpin peptide connected through a β-alanine (βAla) linker to an RNA threading intercalator, **WKWK-Int** (Figure 2.1). We selected βAla for the linker region because of its flexibility and ability to provide additional points of hydrogen bonding interactions. During early studies the probes were synthesized with one, two, or three βAla residues linking the two pieces. The results from initial binding experiments suggested that a linker made with one βAla did not bind as well as the probes made with two or three β-alanines, which both had comparable binding. For ease of synthesis and to eliminate additional entropic costs of using a longer linker all the probes were made with a two residue linker.

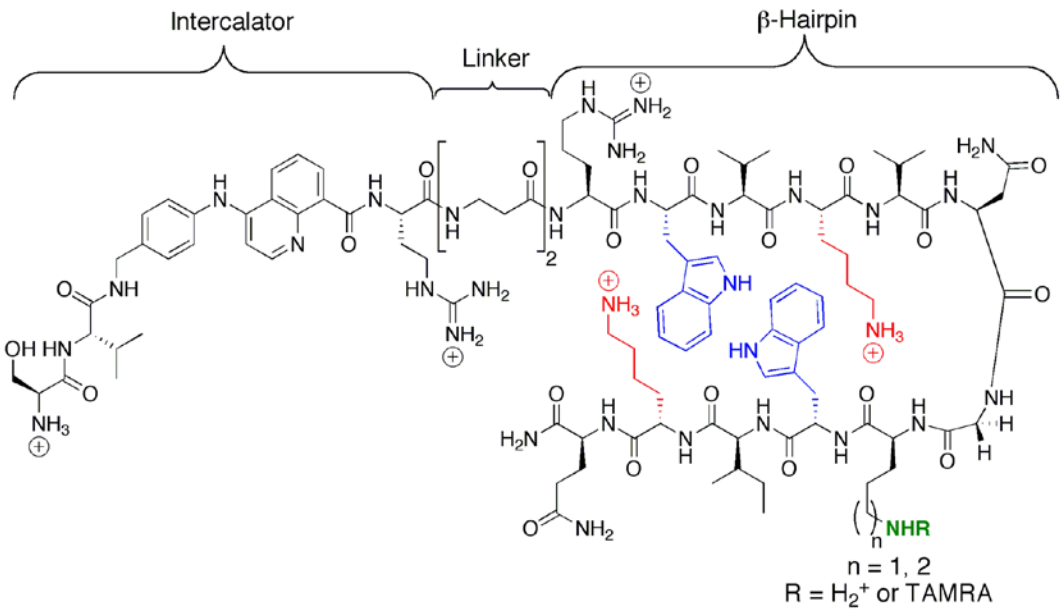


Figure 2.1. Structure and sequence of the **WKWK-Int** probe. The nucleotide binding pocket on the β-hairpin, WKWK, is colored red and blue and the position for attachment of the fluorophore is green.

ii. Design of the control peptides

Control peptides were also designed in order to test what features of the peptide-intercalator probe were contributing to binding. The **WKWK-control** contains the β-hairpin peptide and linker regions alone and the **Int-control** peptide contains the intercalator and linker regions only (Figure 2.2a and b). Additionally, two controls were designed to determine the influence of the peptide structure on RNA binding. The **GKGK-Int** control was designed to have all the amino acids in the β-hairpin portion replaced with glycine except the charged residues (Figure 2.2c). This gives **GKGK-Int** the same charge state as **WKWK-Int**, but it lacks secondary structure. The **GWKWK-Int** control peptide was designed to have glycines replace all the residues except the charged residues and the aromatic tryptophan residues (Figure 2.2d). This control peptide is also unstructured because it lacks an Asn-Gly turn sequence and key side chain-side chain interactions, but it has the same charge state as **WKWK-Int** and has the ability to make aromatic interactions.

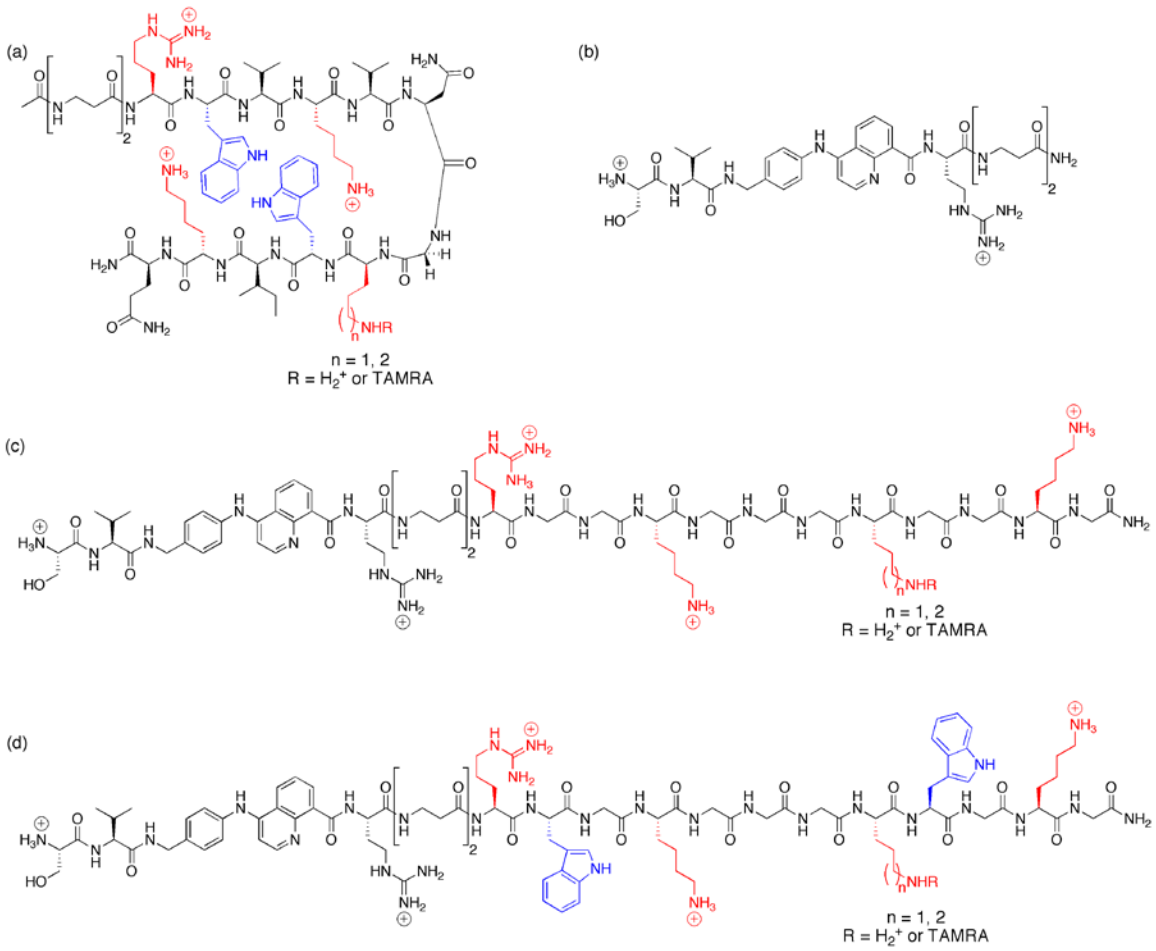


Figure 2.2. Structure and sequences of the control peptides. The cationic residues from the WKWK peptide portion of the probes are highlighted in red and the aromatic residues are highlighted in blue. a) **WKWK-control**. b) **Int-control**. c) **GKGK-Int**. d) **GWKWKG-Int**.

iii. Design of the RNA target

The RNA sequence chosen for these experiments contains regions of double stranded and single stranded bases which are suitable for binding both parts of the **WKWK-Int** probe (Figure 2.3). We wanted to target the hairpin loop region of the RNA, and there was no structural information known about the RNA sequence used by the Beal Lab. We chose to substitute the hairpin loop region for a more biologically relevant target that also has structural information known about it. The selected sequence, **GGAGloop RNA**, contains a single stranded GGAG hairpin loop that which is part of the SL3 Ψ-RNA recognition

element of the HIV-1 nucleocapsid protein and the published NMR structure shows that three of the four residues are unstacked and potentially flipped out from the hairpin loop (Figure 2.4).[12] The remainder of the **GGAGloop RNA** sequence is identical to the sequence used in Beal's studies, including a single stranded internal AAUU bulge that is located near the intercalation site which may also contain bases that are solvent exposed. This bulge also contributes to the flexibility of the intercalation site, which consists of two G-C base pairs surrounded by a 5'U bulge and 3'bulge.[10]

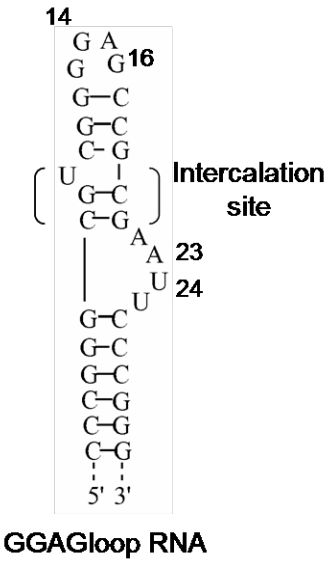


Figure 2.3. Structure and sequences of the **GGAGloop RNA.** The intercalation site is denoted by parentheses, and the significant bases in the hairpin loop and internal bulge are labeled with numbers.

[12] De Guzman, R. N.; Wu, Z. R.; Stalling, C. C.; Pappalardo, L.; Borer, P. N.; Summers, M. F. *Science* **1998**, *279*, 384-388.

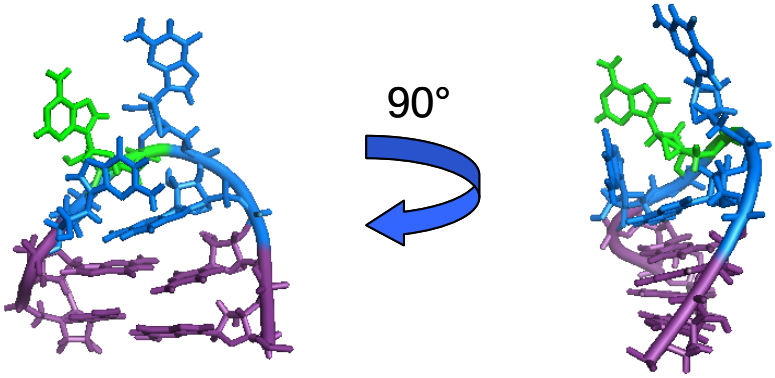


Figure 2.4. Structure of GGAG RNA hairpin loop (PDB 1A1T)[12]. The stem region is highlighted in purple, guanosine residues are blue, and adenosine residues are green.

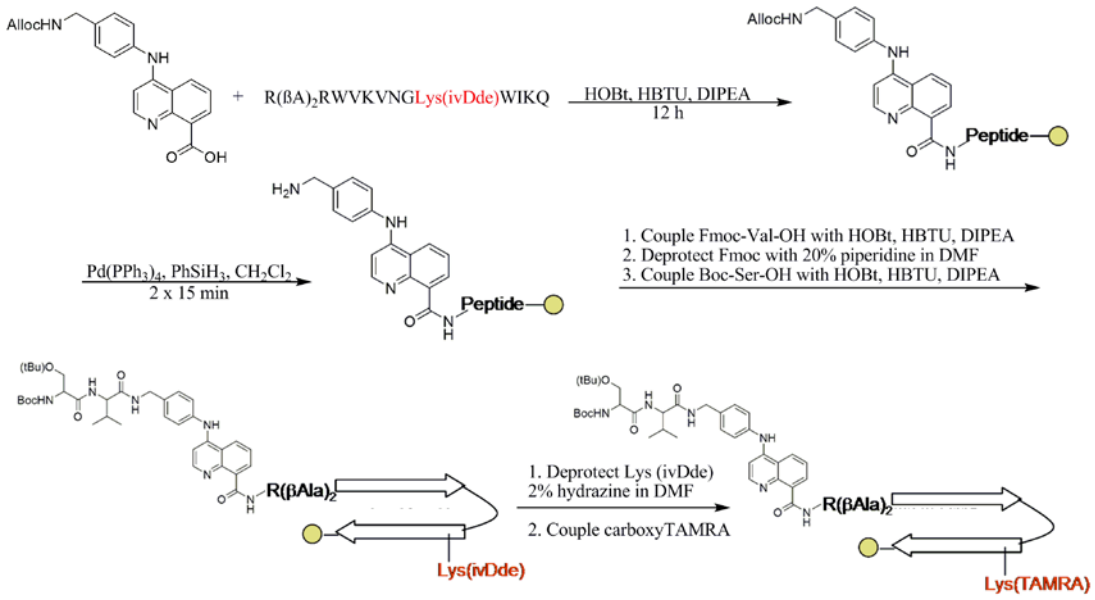
iv. Synthetic scheme

All of the probes were synthesized on bead using standard FMOC solid phase peptide synthesis. The intercalator amino acid was synthesized following the previously published procedure.[10] This unnatural amino acid was then incorporated into the synthesis using an HOBT/HBTU activator solution. Orthogonal deprotection of the N-terminus was then done using $Pd(PPh_3)_4$ and phenylsilane, and the reaction was followed using the Kaiser test.[13] The synthesis of the probe was then completed by adding the final valine and serine residues and deprotecting the N-terminus.

To use fluorescence anisotropy for measuring binding, the probes needed to be labeled with a fluorophore. We chose to attach the 5,6-carboxytetramethylrhodamine (TAMRA) through a lysine side chain that is on the opposite face of the nucleotide binding pocket. The modification of that particular lysine in WKWK was previously shown to not disrupt β-hairpin formation.[14] In order to label the peptide, a slightly different procedure for synthesis

[13] Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. *Anal Biochem* **1970**, *34*, 595-8.
[14] Cooper, W. J.; Waters, M. L. *Org Lett* **2005**, *7*, 3825-3828.

was developed (Scheme 2.1). At the site where the fluorophore label was to be incorporated, an ivDde-protected lysine residue was added instead of the traditional FMOC-Orn(Boc)-OH. The rest of the peptide synthesis was completed as stated above with the exception of the final serine, which was changed to Boc-protected serine in order to protect the N-terminus during the labeling reaction. The Boc group was then easily removed during the acid cleavage from the resin.



Scheme 2.1. Synthesis of the TAMRA-labeled peptide-intercalator probes.

v. Circular dichroism studies

Circular dichroism (CD) studies were done to confirm that **GKGK-Int** and **GWKWK-Int** are unstructured compared to **WKWK-Int** (Figure 2.5). The CD studies help determine the global secondary structure of the peptides. A peptide which folds into a β-sheet type structure will typically show a minimum mean residue ellipticity (MRE, [Θ]) around 215 nm. Whereas a peptide which is unstructured or contains a random coil type structure will show a minimum MRE value around 200 nm. The results show that the **WKWK-Int** probe

resembles a β-sheet structure, but the **GKGK-Int** and **GWKWKG-Int** controls are unstructured.

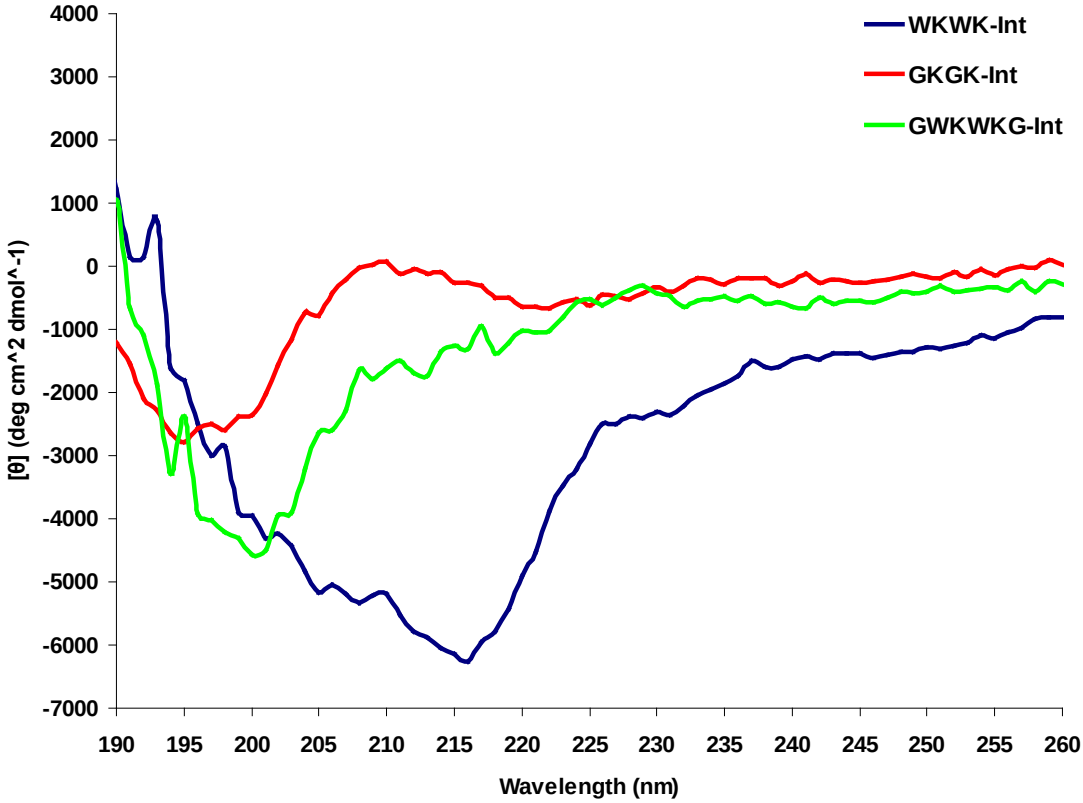


Figure 2.5. Circular dichroism spectrum for the **WKWK-Int** probe (blue) and the unstructured controls, **GKGK-Int** (red) and **GWKWKG-Int** (green). The CD experiments were run using 100 μM peptide in 10 mM sodium phosphate buffer (pH 7.6) at 25°C in a 0.1 cm cell. The spectra that are shown are an average of three scans.

vi. Fluorescence anisotropy binding assays

Fluorescence anisotropy was used to find the binding affinities of each of the peptide-intercalator probes to the RNA. As described above, the peptides were labeled with TAMRA through a lysine side chain on the opposite face of the WKWK binding pocket. The anisotropy experiments showed that **WKWK-control** and **Int-control** peptides did not measurably bind to the **GGAGloop** RNA (Figure 2.6). However, when the two pieces were

connected together, as in **WKWK-Int**, the binding affinity greatly increased (K_d= 3.7 ± 0.7 μM). It was determined that the unstructured **GKGK-Int** bound to the RNA with slightly lower binding affinity (K_d= 9.7 ± 2.2 μM) and the **GWKWKG-Int** bound with a dissociation constant of 31.8 ± 13.2 μM (Figure 2.7). This suggests that although electrostatic interactions contribute significantly to the binding, there is some increased binding due to either the β-hairpin structure or specific contact with non-charged side chains.

To test the specificity of the binding event, we investigated the binding of **WKWK-Int** to BIV TAR RNA, which has similar secondary structure to that of **GGAGloop RNA** but a different sequence. The BIV TAR RNA (bTAR) sequence comes from the bovine immunodeficiency virus and is known to bind with low nanomolar affinity to the β-hairpin peptide, BIV Tat. This RNA sequence contains two uridine bulges and a four member hairpin loop with the sequence 5'-CAUU-3' (Figure 2.8). The fluorescence anisotropy binding experiments show that the **WKWK-Int** probe binds more tightly to the **GGAGloop RNA** than it does to the **bTAR RNA** (Figure 2.8). The specificity is most likely due to the intercalator which has been shown to preferentially insert between G-C base pairs that are surrounded by bulge regions.[10, 11] This type of binding site is not present in the **bTAR RNA**.

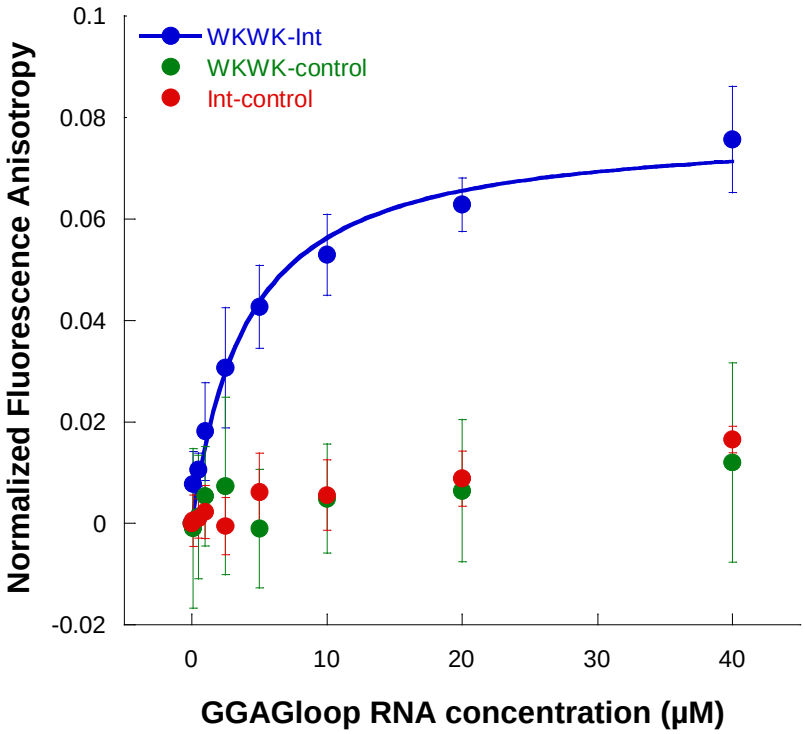


Figure 2.6. Fluorescence anisotropy binding experiments of TAMRA-labeled probes (0.5 μM) with **GGAGloop RNA**. **WKWK-Int** (blue) binds in the low micromolar range (3.7 ± 0.7 μM) whereas **WKWK-control** (green) and **Int-control** (red) do not measureably bind. Experiments were done in PBS buffer (containing 10 mM sodium phosphate, 140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) at 25°C and each curve is an average of 3 or 4 runs.

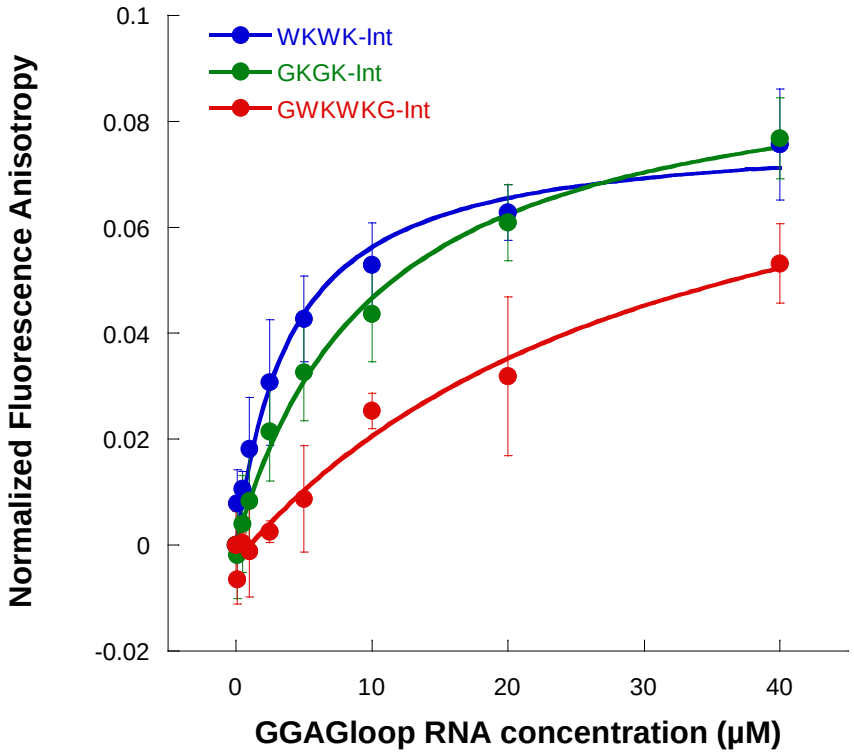


Figure 2.7. Fluorescence anisotropy binding experiments of TAMRA-labeled probes (0.5 μM) with **GGAGloop RNA**. **WKWK-Int** (blue) binds with the highest affinity (3.7 ± 0.7 μM) whereas **GKGK-Int** (green) and **GWKWKG-Int** (red) bind less tightly (9.7 ± 2.2 μM and 31.8 ± 13.2 μM, respectively). Experiments were done in PBS buffer (containing 10 mM sodium phosphate, 140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) at 25°C and each curve is an average of 3 or 4 runs.

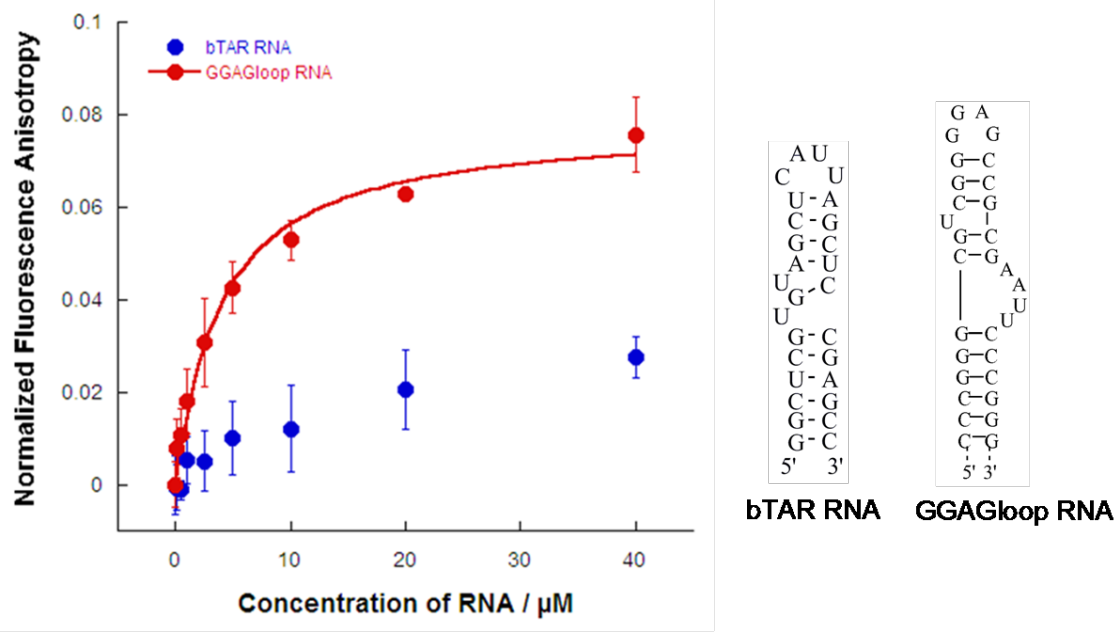


Figure 2.8. Fluorescence anisotropy binding experiments of TAMRA-labeled **WKWK-Int** (0.5 µM) binding to **BIV TAR RNA** or **GGAGloop RNA**. **WKWK-Int** binds more tightly to **GGAGloop RNA** (K_d= 3.7 ± 0.7 µM) than it does to **BIV TAR RNA**. Experiments were done with 0.5 µM **WKWK-Int** in PBS buffer (containing 10 mM sodium phosphate, 140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) at 25°.

vii. RNase footprinting assays with WKWK-Int and GGAGloop RNA

Enzymatic footprinting experiments were done to identify which regions of the RNA were making contacts with the conjugated probes. The **GGAGloop RNA** sequence was 3'-labeled with ^{32}pCp, and enzymatic digestions were preformed with RNases that target single stranded regions of RNA (RNase T1 and RNase 1) (Figure 2.9a) and double stranded regions of RNA (RNase V1) (Figure 2.9b). These experiments helped to identify whether contacts were being made in the hairpin loop region, the bulge region, or the stem region.

The RNase footprinting experiments show that **WKWK-Int** binds to **GGAGloop RNA** with low micromolar affinity, which is consistent with the fluorescence anisotropy results. Unlike the fluorescence experiments, the footprinting experiments were all done in the presence of unlabeled yeast tRNA to slow down the enzymatic degradation of the target

RNA. The fact that the binding affinities are similar for the two methods is further proof that the peptide is specific for this sequence of RNA. The **WKWK-Int** probe specifically protects residues G16, A23, and U24 in the single stranded regions and it seems to also protect much of the double stranded stem region.

The resulting bands from cleavage at bases G16, A23, and U24 were quantified using ImageQuant analysis program and plotted as fraction bound versus the concentration of peptide (Figure 2.10). The results confirm that A23 and U24 are protected from RNase cleavage. There was moderate binding at position G16, which suggests that there are multiple orientations of binding, but binding at the internal bulge is slightly preferred.

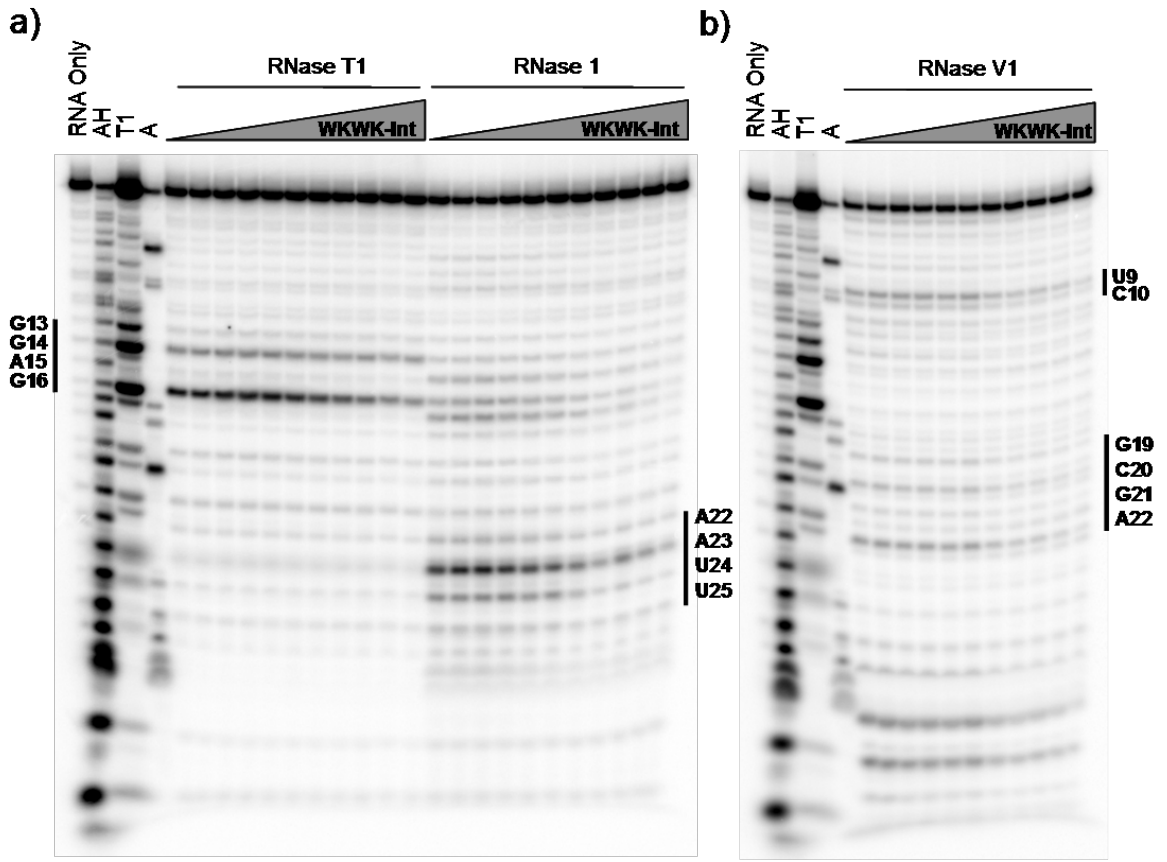


Figure 2.9. RNase footprinting experiments with **WKWK-Int** and **GGAGloop RNA**. Shown is the phosphor image of a 20% denaturing polyacrylamide gel. Structure lanes: RNA only, folded RNA in buffer; AH, alkaline hydrolysis; T1, RNase T1 (G-lane); A, RNase A (C and U lane). Increasing concentrations (0 to 100 μM) of **WKWK-Int** were added to **GGAGloop RNA** (0.18 μM) in 10 mM PBS buffer (140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) and 5 ng/μL yeast tRNA. a) RNase T1 (cleaves G residues) was used to visualize the G-rich loop region and RNase 1 (cleaves ssRNA) was used to visualize both the loop and internal bulge regions. Bases G14 and G16 were the major cleavage sites for RNase T1, whereas bases A23 and U24 were the major cleavage sites for RNase 1. b) RNase V1 (cleaves dsRNA) was used to visualize the stem regions of the RNA.

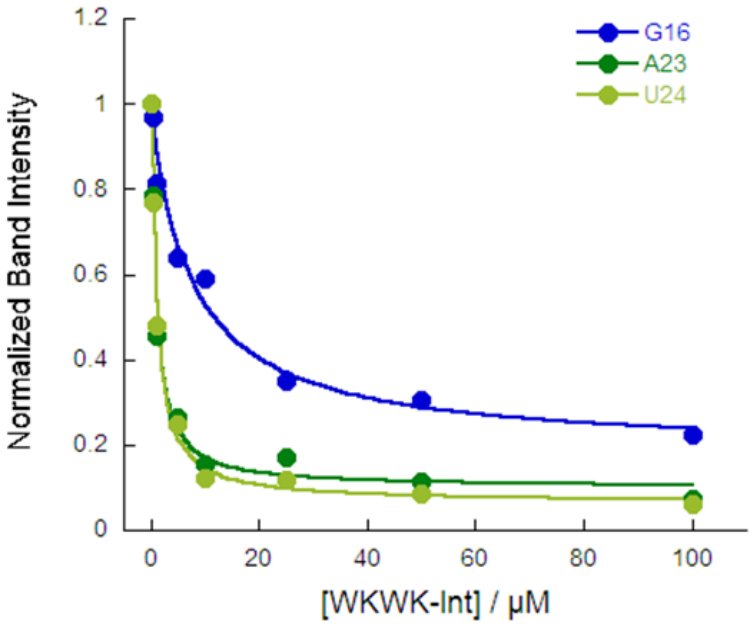


Figure 2.10. ImageQuant analysis of RNase footprinting assay of **WKWK-Int** binding **GGAGloop RNA.** The tightest binding occurs in the internal bulge region (dark and light green, local binding affinity for A23 equals 0.9 ± 0.2 μM and affinity for U24 was 1.0 ± 0.2 μM), and moderate binding occurs in the loop region (dark blue, local binding affinity for G16 =7.8 ± 2.3 μM). The band intensities were the average of at least two RNase footprinting assays.

viii. RNase footprinting assays with WKWK-control and Int-control

The binding of the **WKWK-control** and the **Int-control** peptides to GGAGloop RNA was also investigated using footprinting assays (Figures 2.11 and 2.12). Neither of the control peptides protects the observable regions of the **GGAGloop RNA** from enzyme cleavage as well as **WKWK-Int** protected the RNA. These results are consistent with the fluorescence anisotropy binding data which showed there was no measurable binding affinity for the control peptides. For both control peptides there is no binding detected in the single stranded regions, but the **Int-control** peptide shows minimal binding in the double stranded

regions at the highest concentration (100 μM). This confirms the published results showing that the intercalator binds by threading through the stem regions of the RNA.[10]

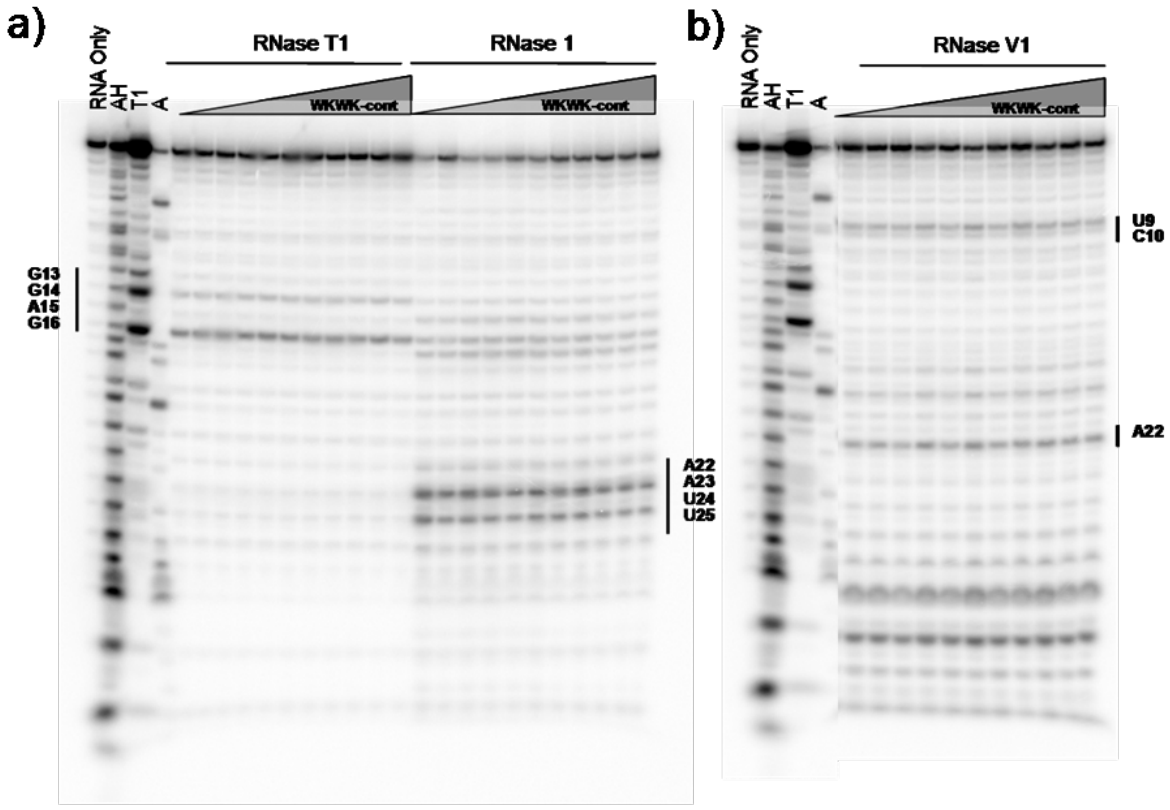


Figure 2.11. RNase footprinting experiments with **WKWK-control** and **GGAGloop RNA**. Shown is the phosphor image of a 20% denaturing polyacrylamide gel. Structure lanes: RNA only, folded RNA in buffer; AH, alkaline hydrolysis; T1, RNase T1 (G-lane); A, RNase A (C and U lane). Increasing concentrations (0 to 100 μM) of **WKWK-control** were added to **GGAGloop RNA** (0.18 μM) in 10 mM PBS buffer (140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) and 5 ng/μL yeast tRNA. a) RNase T1 (cleaves G residues) was used to visualize the G-rich loop region and RNase 1 (cleaves ssRNA) was used to visualize both the loop and internal bulge regions. b) RNase V1 (cleaves dsRNA) was used to visualize the stem regions of the RNA.

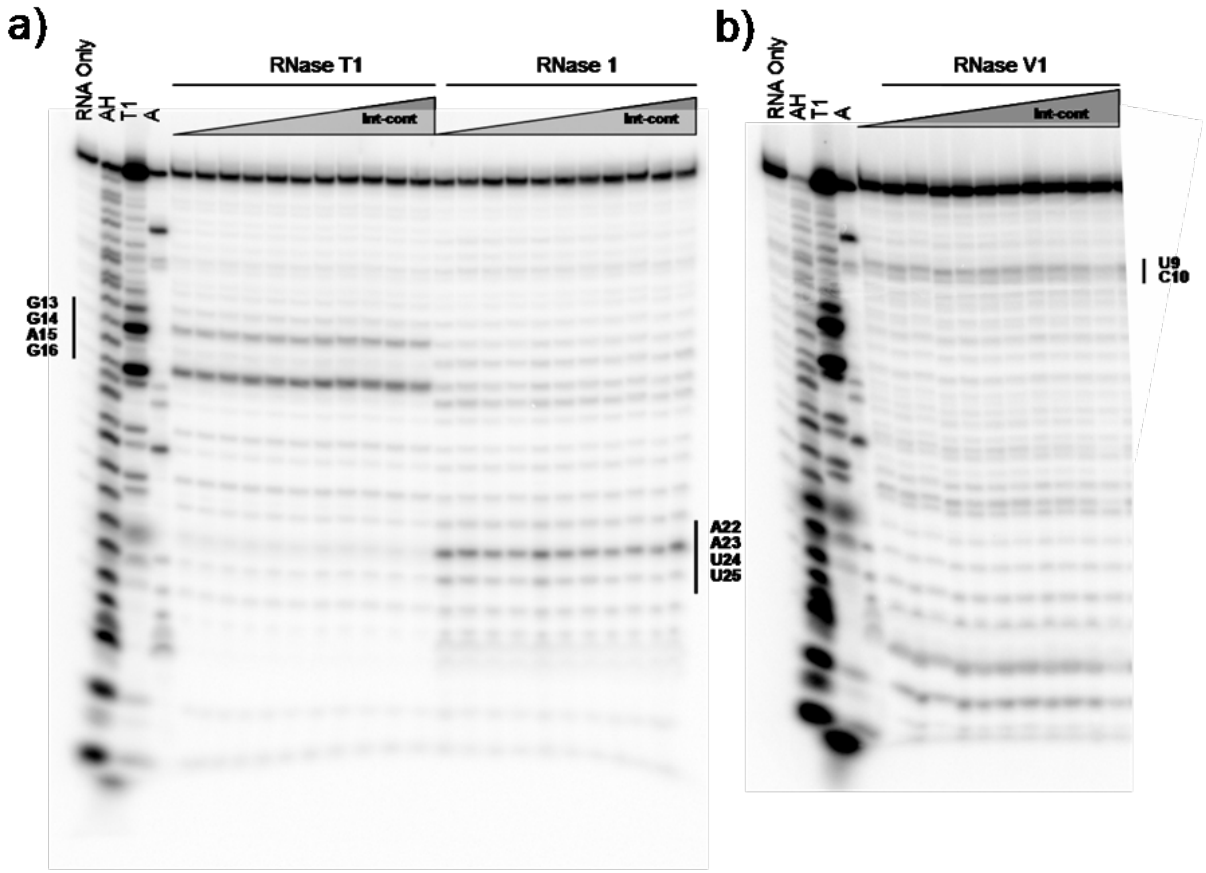


Figure 2.12. RNase footprinting experiments with **Int-control** and **GGAGloop RNA**. Shown is the phosphor image of a 20% denaturing polyacrylamide gel. Structure lanes: RNA only, folded RNA in buffer; AH, alkaline hydrolysis; T1, RNase T1 (G-lane); A, RNase A (C and U lane). Increasing concentrations (0 to 100 μM) of **Int-control** were added to **GGAGloop RNA** (0.18 μM) in 10 mM PBS buffer (140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) and 5 ng/μL yeast tRNA. a) RNase T1 (cleaves G residues) was used to visualize the G-rich loop region and RNase 1 (cleaves ssRNA) was used to visualize both the loop and internal bulge regions. b) RNase V1 (cleaves dsRNA) was used to visualize the stem regions of the RNA.

ix. RNase footprinting assays with GKGK-Int control

The results from the footprinting assay for **GKGK-Int** binding to **GGAGloop RNA** show that this peptide also makes contacts with the loop, stem, and bulge regions of the RNA (Figure 2.13). However, in comparing the cleavage pattern of **GKGK-Int** to that of **WKWK-Int**, we see some differences in the bulge region. The **WKWK-Int** protects residue

A23 from RNase 1 cleavage, whereas **GKGK-Int** does not. This peptide has the same charge but not the same structure as **WKWK-Int** and is lacking aromatic residues. These results suggest that protection of that particular base is due to the structure of the peptide or aromatic interactions and not electrostatic interactions.

The footprinting assays done with **GKGK-Int** were analyzed and plotted in the same way as the **WKWK-Int** results (Figure 2.14). The results confirm that A23 is not protected from RNase cleavage by the **GKGK-Int** peptide, but G16 and U24 are protected. The binding in the loop region resembles that of **WKWK-Int**, and this agrees with the hypothesis that there are multiple orientations of binding, but binding at the internal bulge is preferred.

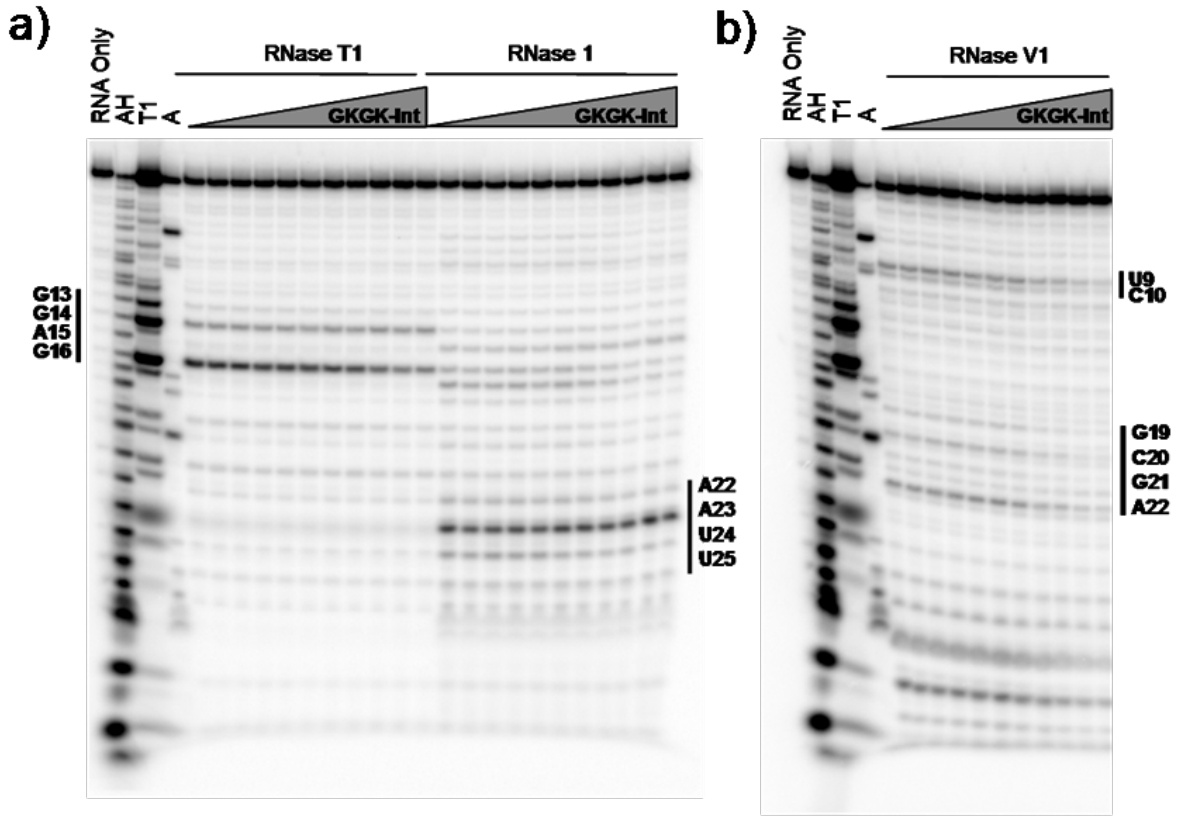


Figure 2.13. RNase footprinting experiments with **GKGK-Int** and **GGAGloop RNA**. Shown is the phosphor image of a 20% denaturing polyacrylamide gel. Structure lanes: RNA only, folded RNA in buffer; AH, alkaline hydrolysis; T1, RNase T1 (G-lane); A, RNase A (C and U lane). Increasing concentrations (0 to 100 μM) of **GKGK-Int** were added to **GGAGloop RNA** (0.18 μM) in 10 mM PBS buffer (140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) and 5 ng/μL yeast tRNA. a) RNase T1 (cleaves G residues) was used to visualize the G-rich loop region and RNase 1 (cleaves ssRNA) was used to visualize both the loop and internal bulge regions. b) RNase V1 (cleaves dsRNA) was used to visualize the stem regions of the RNA.

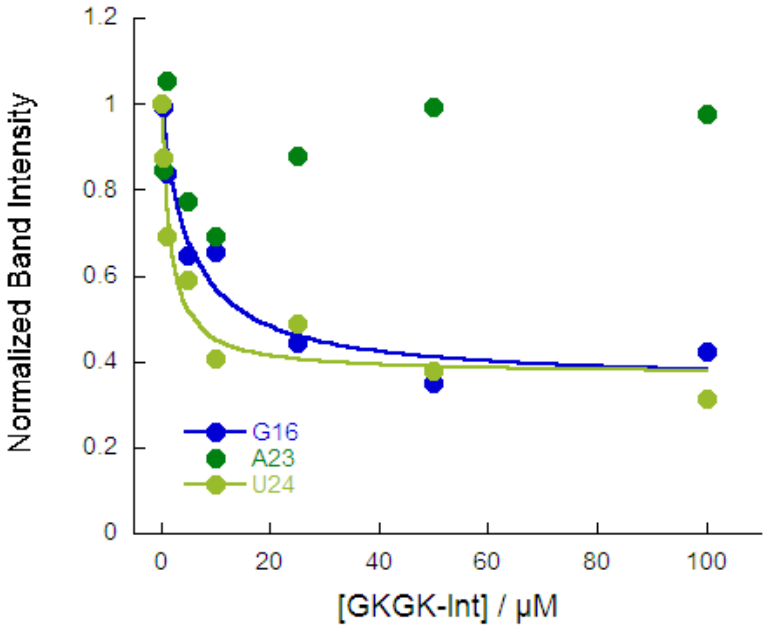


Figure 2.14. ImageQuant analysis of RNase footprinting assay of **GKGK-Int** binding **GGAGloop RNA.** There is no RNase protection observed at bases G14 (dark blue) or A23 (dark green), but binding is detected at bases G16 (light blue) and U24 (light green). The band intensities were the average of at least two RNase footprinting assays.

x. RNase footprinting assays with GWKWKG-Int control

The **GWKWKG-Int** control peptide was designed to test what effect aromatic interactions have on RNA binding. The footprinting assays show that this peptide makes similar contacts with the loop, stem, and bulge regions of the RNA as the other peptides (Figure 2.15). By comparing the cleavage pattern at the A23 site, we saw that it protects this base but it has a weaker affinity than the structured **WKWK-Int** peptide. We can conclude from these results that the aromatic residues are important for binding in the bulge region, but the β-hairpin structure of the WKWK peptide preorganizes it for binding.

The analysis of these footprinting assays confirm that although A23 is protected from RNase cleavage by the **GWKWKG-Int** peptide, it is a weaker interaction (Figure 2.16). Both bases G16 and U24 continue to be protected from RNase cleavage as was shown for the other peptides. The curves for all the peptides were fit using a nonlinear least squares binding equation and local binding affinities were determined for the individual bases (Table 2.1). These values clearly show that the binding at A23 was greatly affected by the lack of structure of the peptide.

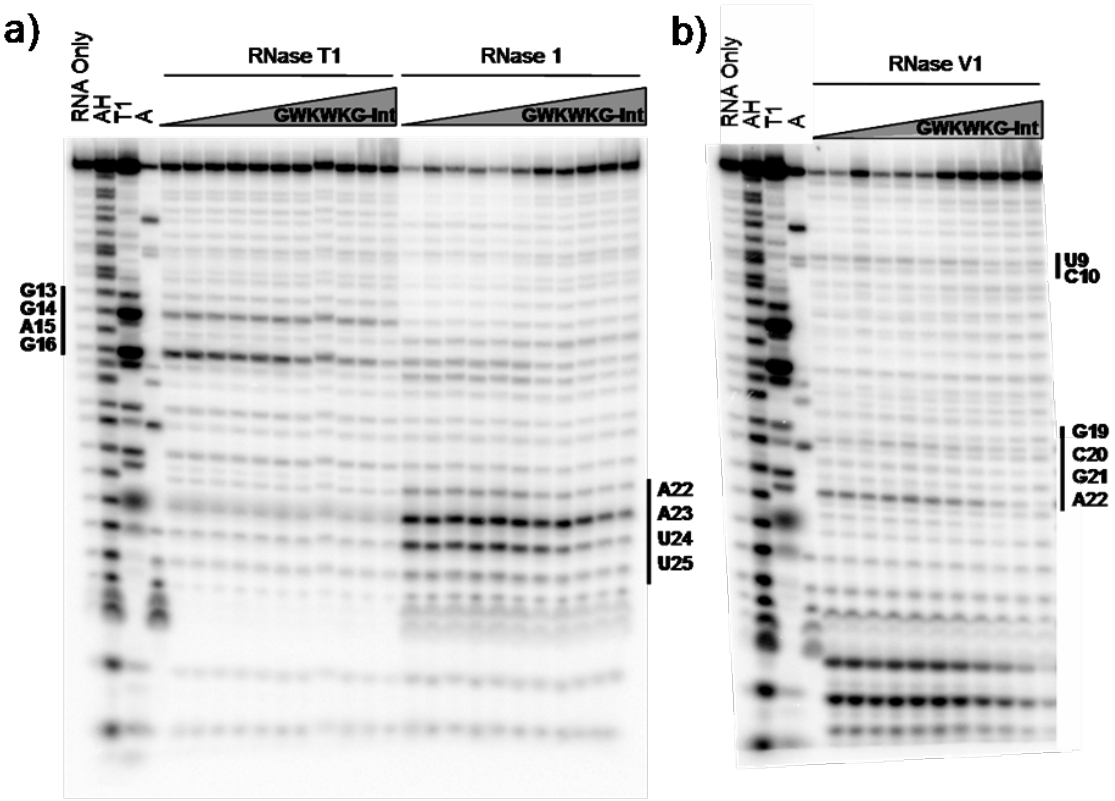


Figure 2.15. RNase footprinting experiments with **GWKWKG-Int** and **GGAGloop RNA**. Shown is the phosphor image of a 20% denaturing polyacrylamide gel. Structure lanes: RNA only, folded RNA in buffer; AH, alkaline hydrolysis; T1, RNase T1 (G-lane); A, RNase A (C and U lane). Increasing concentrations (0 to 100 μM) of **GWKWKG-Int** were added to **GGAGloop RNA** (0.18 μM) in 10 mM PBS buffer (140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) and 5 ng/μL yeast tRNA. a) RNase T1 (cleaves G residues) was used to visualize the G-rich loop region and RNase 1 (cleaves ssRNA) was used to visualize both the loop and internal bulge regions. b) RNase V1 (cleaves dsRNA) was used to visualize the stem regions of the RNA.

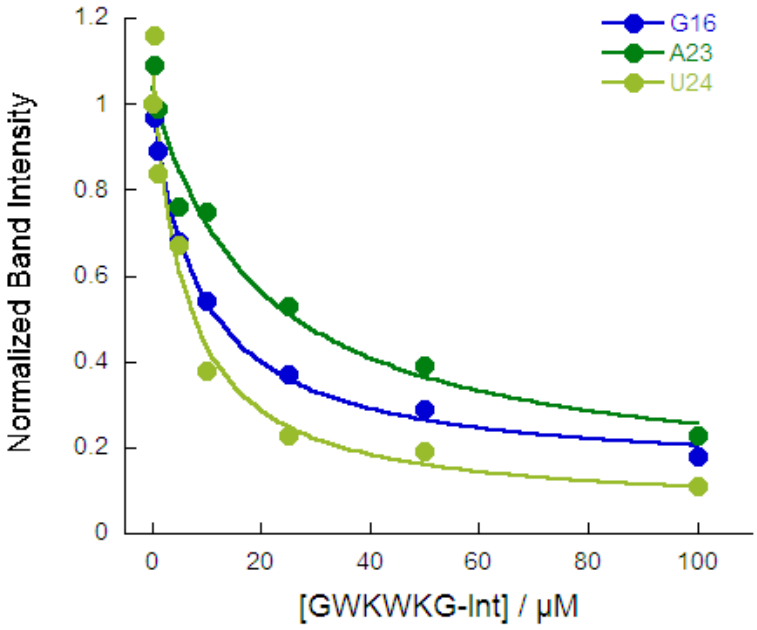


Figure 2.16. ImageQuant analysis of RNase footprinting assay of **GWKWKG-Int** binding **GGAGloop RNA.** Tightest binding was detected at bases G16 (dark blue) and U24 (light green), and moderate binding was detected at A23 (dark green). The band intensities were the average of two RNase footprinting assays.

Table 2.1. Comparison of Binding Affinities[a] of the Peptide-Intercalator Conjugates to **GGAGloop RNA**

	Local binding affinity / μM (± error)		
Peptide	**G16**	**A23**	**U24**
WKWK-Int	**7.8 (2.3)**	**0.9 (0.2)**	**1.0 (0.2)**
GKGK-Int	**5.2 (2.3)**	**n.b.[b]**	**1.6 (0.8)**
GWKWKG-Int	**8.8 (1.0)**	**19.6 (6.9)**	**6.0 (2.6)**

[a] Local binding affinities were determined by plotting the RNase footprinting results as band intensity versus concentration of peptide and fitting the data using equation 1.
[b] no binding detected or binding too weak to determine a dissociation constant

xi. Investigation of cyclic WKWK-Int binding GGAGloop RNA

A cyclic version of the WKWK-Int probe, **cycWKWK-Int**, was synthesized using a disulfide linkage between the N- and C-termini of the β-hairpin portion of the probe (Figure 2.17). This peptide has a pre-formed binding pocket, and does not need to overcome the entropic cost of folding. CD studies show that **cycWKWK-Int** has a large minimum between 200-215 nm (Figure 2.18). This unexpected spectrum may be caused by the mixture of highly structured peptide and unstructured intercalator region. The disulfide bond may also be perturbing the β-hairpin structure. The maximum signal at 225 nm confirms that the peptide is well-folded because the peak is caused by the exciton coupling of the two tryptophan residues when they are oriented in a folded structure.

Fluorescence anisotropy was used to measure the binding affinity of **cycWKWK-Int** to the **GGAGloop RNA (**Figure 2.19). These experiments showed that **cycWKWK-Int** binds with a dissociation constant of 3.0 ± 0.8 μM, which is nearly the same binding affinity that **WKWK-Int** probe showed. These data suggest that the WKWK peptide has a comparable structure to the cyclic version and therefore binds with a similar affinity.

The RNase footprinting assays confirm that both **cycWKWK-Int** and **WKWK-Int** bind in the same orientation (Figure 2.20). There is moderate protection of the bases in the hairpin loop region and there is strong binding observed at bases A23 and U24. The binding in the double stranded region of RNA also matches what was detected for all the other peptides. By quantifying the footprinting results for **cycWKWK-Int**, the local binding affinities were determined and compared to that for the **WKWK-Int** probe (Figure 2.21, Table 2.2). The cyclized version of WKWK seems to bind worse to the hairpin loop region, but with a comparable affinity in the bulge region.

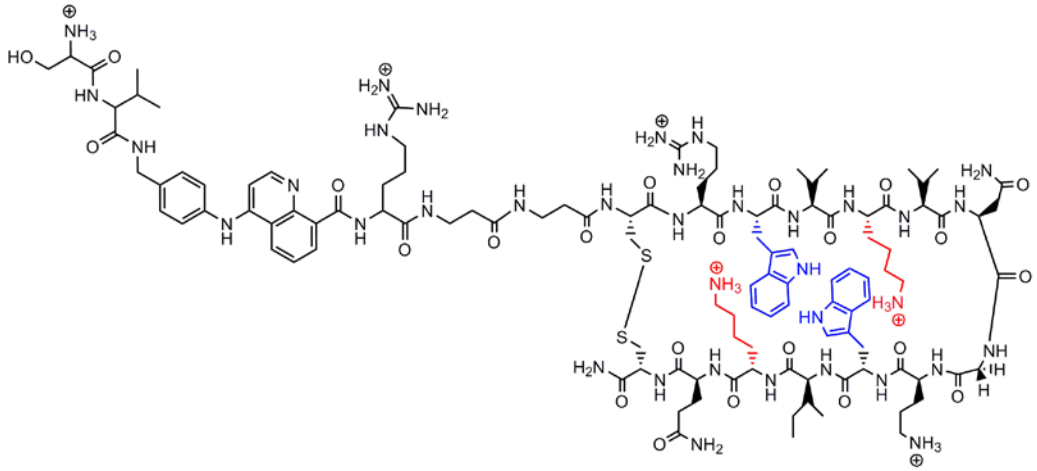


Figure 2.17. Structure and sequence of **cycWKWK-Int**. The cationic residues from the WKWK peptide portion of the probes are highlighted in red and the aromatic residues are highlighted in blue.

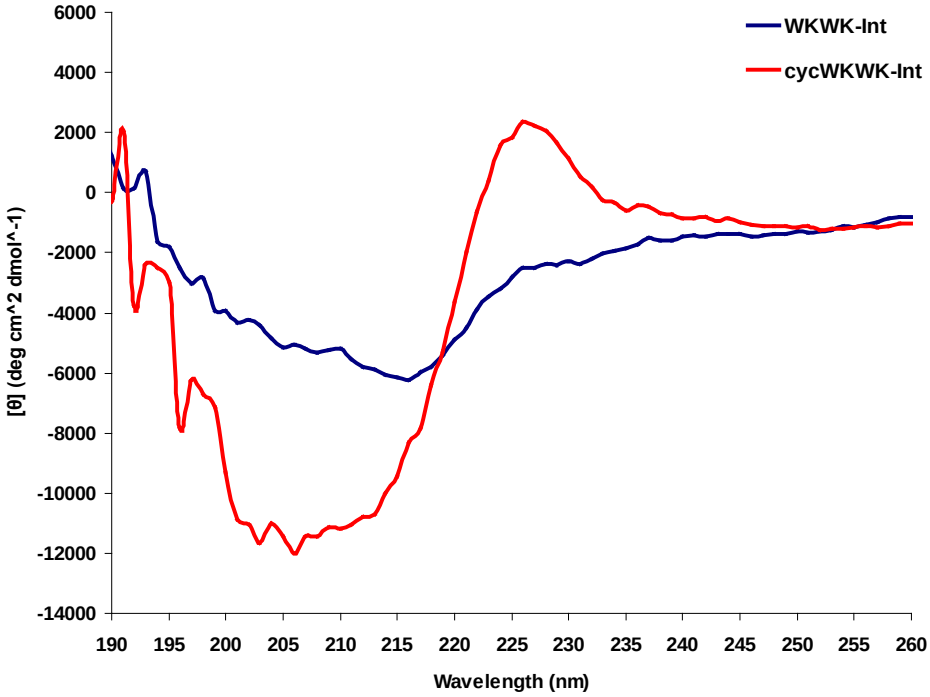


Figure 2.18. CD spectrum for **WKWK-Int** (blue) and **cycWKWK-Int** (red). The CD experiments were run using 100 μM peptide in 10 mM sodium phosphate buffer (pH 7.6) at 25°C in a 0.1 cm cell. The spectra that are shown are an average of three scans.

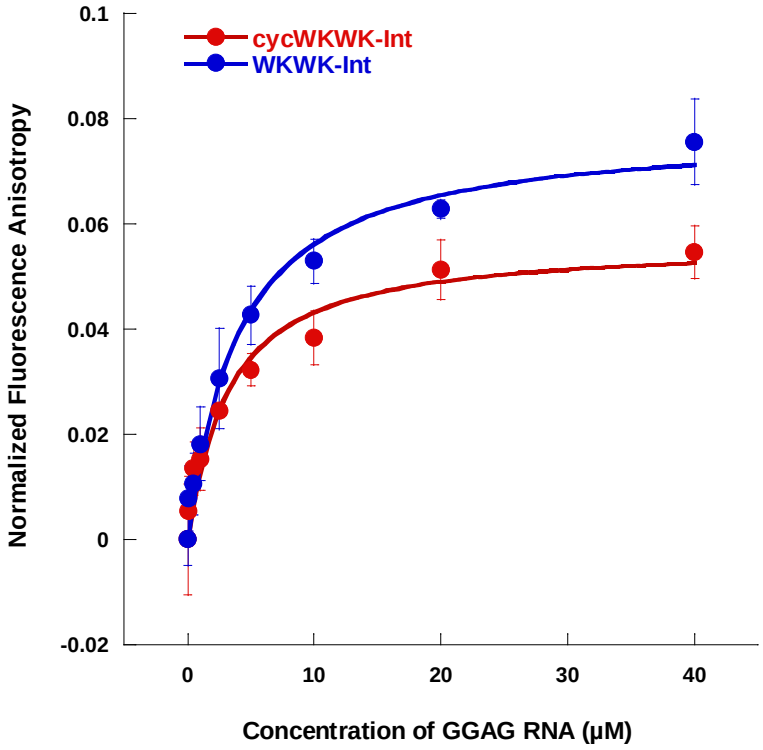


Figure 2.19. Fluorescence anisotropy binding experiments of TAMRA-labeled **cycWKWK-Int** and **WKWK-Int** binding to **GGAGloop RNA**. Experiments were done with 0.5 μM peptide in PBS buffer (containing 10 mM sodium phosphate, 140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) at 25°. Experimental data is an average of at least three runs.

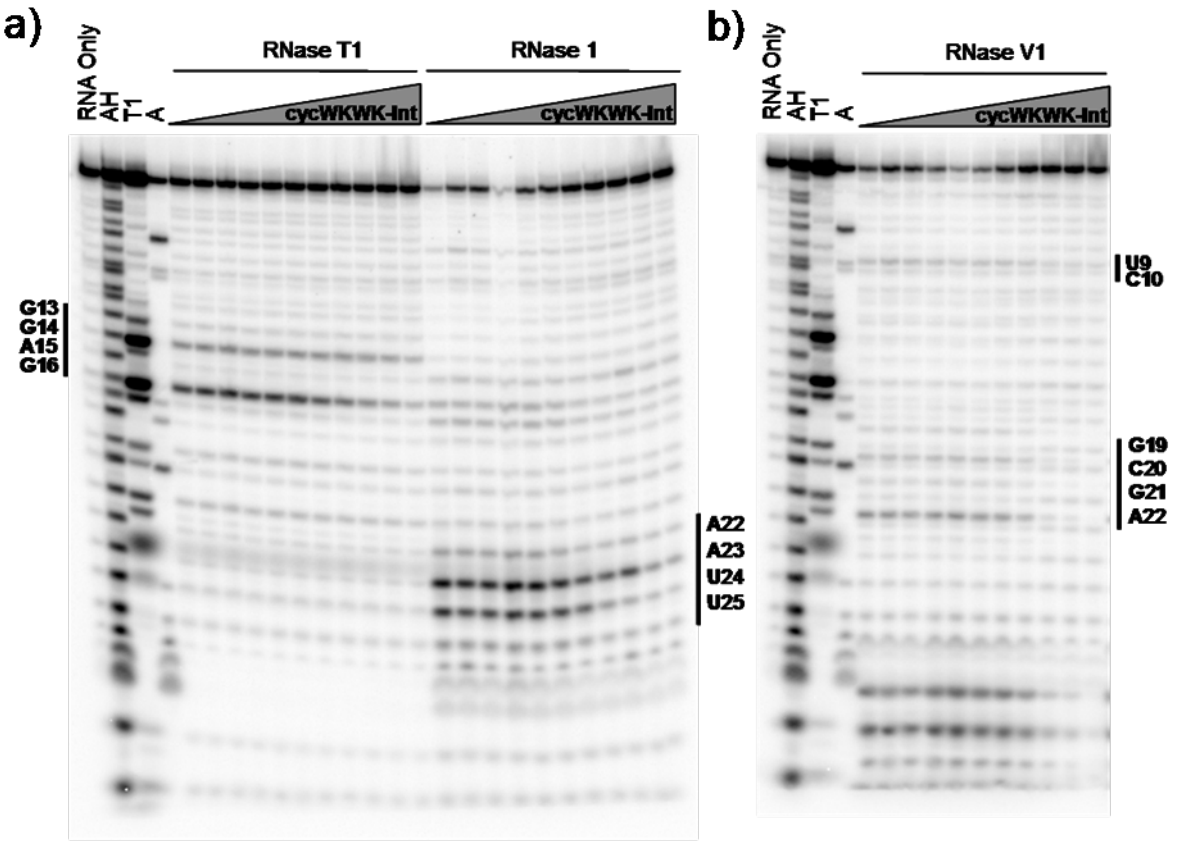


Figure 2.20. RNase footprinting experiments with **cycWKWK-Int** and **GGAGloop RNA**. Shown is the phosphor image of a 20% denaturing polyacrylamide gel. Structure lanes: RNA only, folded RNA in buffer; AH, alkaline hydrolysis; T1, RNase T1 (G-lane); A, RNase A (C and U lane). Increasing concentrations (0 to 100 μM) of **cycWKWK-Int** were added to **GGAGloop RNA** (0.18 μM) in 10 mM PBS buffer (140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) and 5 ng/μL yeast tRNA. a) RNase T1 (cleaves G residues) was used to visualize the G-rich loop region and RNase 1 (cleaves ssRNA) was used to visualize both the loop and internal bulge regions. b) RNase V1 (cleaves dsRNA) was used to visualize the stem regions of the RNA.

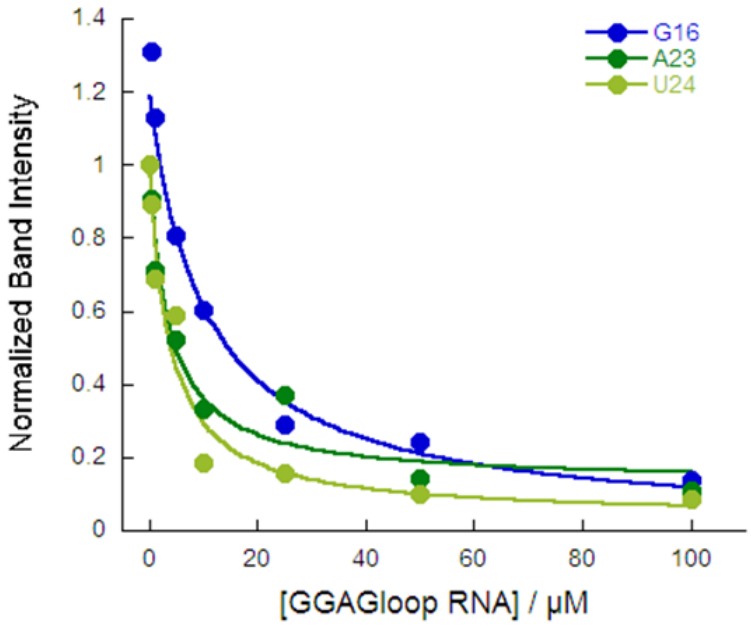


Figure 2.21. ImageQuant analysis of RNase footprinting assay of **cycWKWK-Int** binding **GGAGloop RNA.** Weak binding was observed in the hairpin loop regions (G16, blue), but tighter binding was detected in the bulge region (A23, dark green; U24, light green). The band intensities were the average of two RNase footprinting assays.

Table 2.2. Comparison of Local Binding Affinities[a] for **cycWKWK-Int** and **WKWK-Int**

	Local binding affinity / μM (± error)		
Peptide	**G16**	**A23**	**U24**
WKWK-Int	**7.8 (2.3)**	**0.9 (0.2)**	**1.0 (0.2)**
cycWKWK-Int	**10.4 (5.1)**	**3.8 (1.6)**	**3.9 (1.7)**

[a] Local binding affinities were determined by plotting the RNase footprinting results as band intensity versus concentration of peptide and fitting the data using equation 1.

xii. Investigation of a WKWK-Int probe with reversed orientation

In the design of the **WKWK-Int** probe, the intercalator was placed on the N-terminus of the peptide. The footprinting results have shown that the C-terminal β-hairpin peptide interacts with the both the bulge region and the hairpin loop region. To test whether there is a preference for which terminus of the intercalator threads the RNA, we designed a probe where the intercalator and the peptide have been switched, **revWKWK-Int** (Figure 2.22). The fluorescence anisotropy data show that by reversing the orientation there is a six-fold decrease in binding affinity (Figure 2.23). The dissociation constant for **revWKWK-Int** is 18.2 ± 8.3 μM as compared to 3.7 ± 0.7 μM for **WKWK-Int**.

However, the footprinting experiments show the same regions of the RNA are protected from RNase cleavage for r**evWKWK-Int** as they were for **WKWK-Int** (Figure 2.24). Analysis of the protected regions showed that the local binding affinities for bases A23 and U24 are 1.5 (± 0.5) μM and 0.7 (± 0.1) μM, respectively and the dissociation constant for G16 was 6.8 (± 1.8) μM (Figure 2.25, Table 2.3). These binding affinities are within error of **WKWK-Int** binding **GGAG RNA**. This suggests that the intercalator can thread either direction, but the peptide promotes binding in the bulge region. It also suggests that the TAMRA-fluorophore used in the fluorescence anisotropy binding experiments may be interfering with binding in the case of **revWKWK-Int**. This would explain the discrepancies between the binding affinities determined using the two different methods.

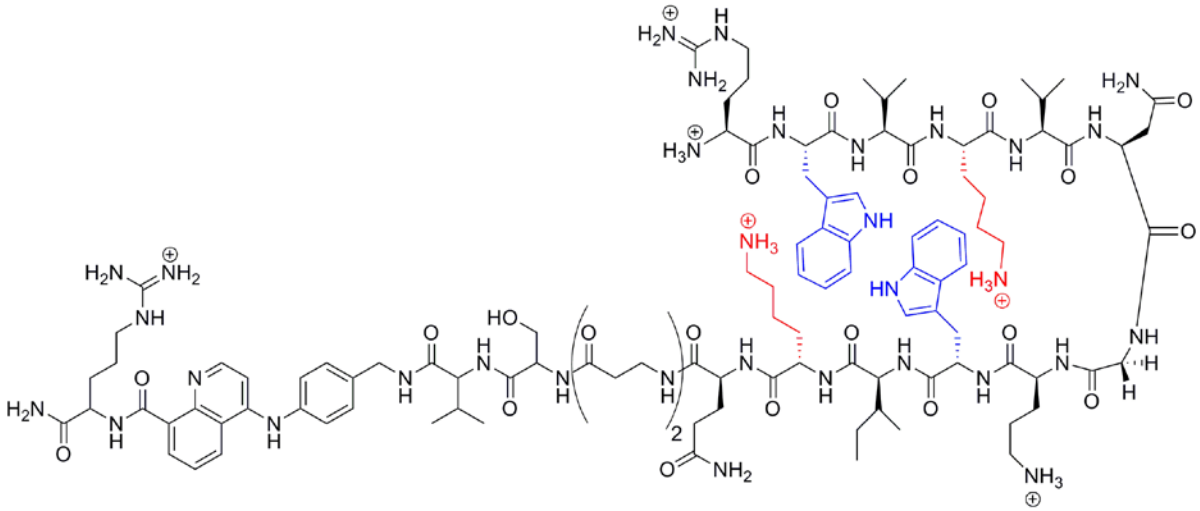


Figure 2.22. Structure and sequence of **revWKWK-Int**. The cationic residues from the WKWK peptide portion of the probes are highlighted in red and the aromatic residues are highlighted in blue.

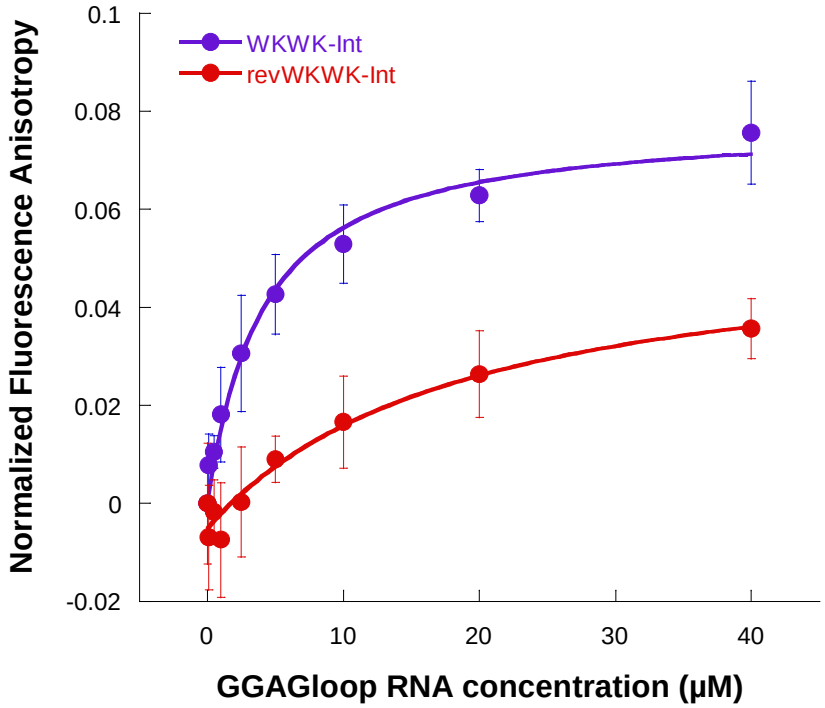


Figure 2.23. Fluorescence anisotropy binding experiments of TAMRA-labeled **revWKWK-Int** (red) and **WKWK-Int** (blue) binding to **GGAGloop RNA**. Peptide concentration was 0.5 μM in all experiments. Experiments were done in 10 mM PBS buffer (140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) at 25°C and the revWKWK-Int data is an average of six runs.

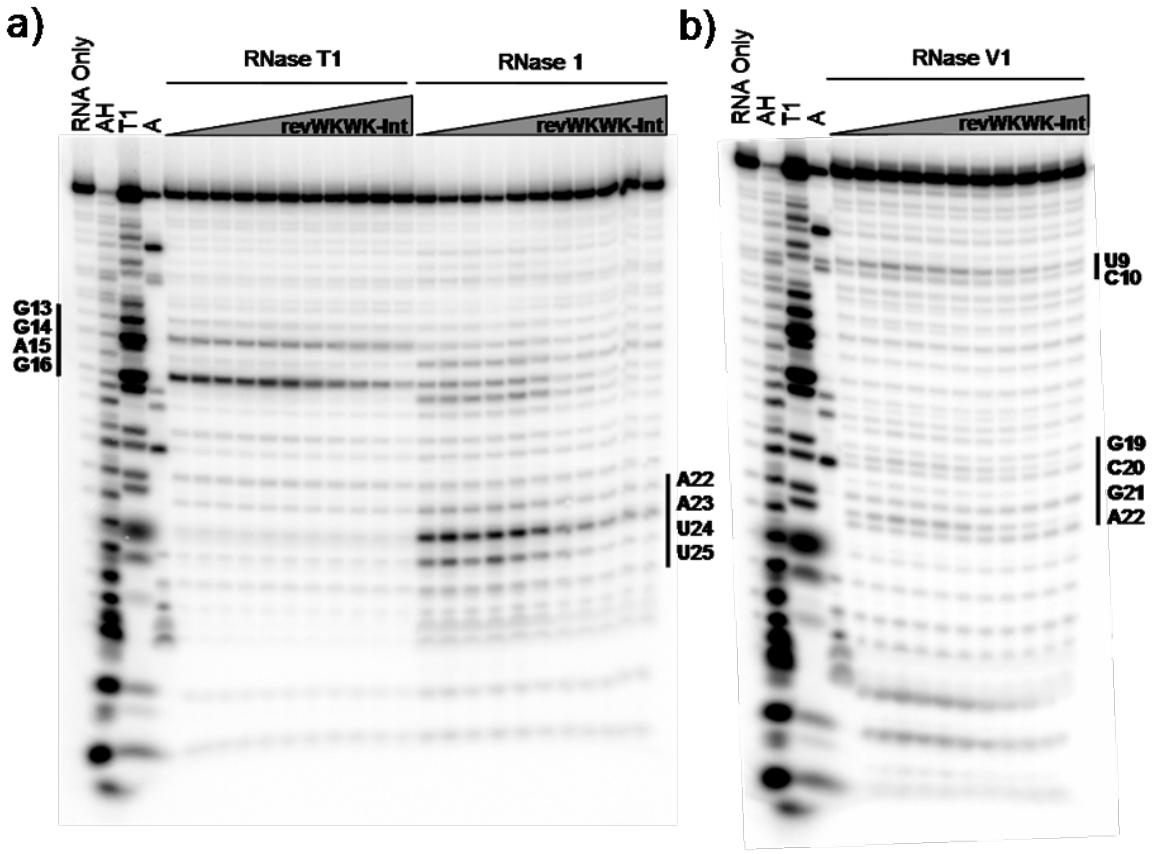


Figure 2.24. RNase footprinting experiments with **revWKWK-Int** and **GGAGloop RNA**. Shown is the phosphor image of a 20% denaturing polyacrylamide gel. Structure lanes: RNA only, folded RNA in buffer; AH, alkaline hydrolysis; T1, RNase T1 (G-lane); A, RNase A (C and U lane). Increasing concentrations (0 to 100 μM) of **revWKWK-Int** were added to **GGAGloop RNA** (0.18 μM) in 10 mM PBS buffer (140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) and 5 ng/μL yeast tRNA. a) RNase T1 (cleaves G residues) was used to visualize the G-rich loop region and RNase 1 (cleaves ssRNA) was used to visualize both the loop and internal bulge regions. b) RNase V1 (cleaves dsRNA) was used to visualize the stem regions of the RNA.

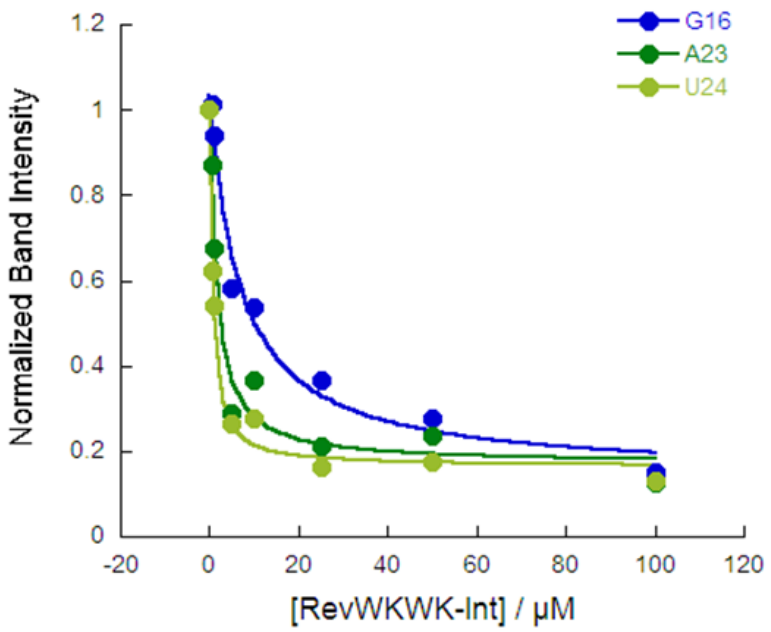


Figure 2.25. ImageQuant analysis of RNase footprinting assay of **revWKWK-Int** binding **GGAGloop RNA.** Only moderate binding was observed at base G16 (dark blue), but tighter binding was detected at bases A23 (dark green) and U24 (light green). The band intensities were calculated from the average of two RNase footprinting assays.

Table 2.3. Comparison of Local Binding Affinities[a] for **revWKWK-Int** and **WKWK-Int**

	Local binding affinity / μM (± error)		
Peptide	**G16**	**A23**	**U24**
WKWK-Int	**7.8 (2.3)**	**0.9 (0.2)**	**1.0 (0.2)**
revWKWK-Int	**6.8 (1.8)**	**1.5 (0.5)**	**0.7 (0.1)**

[a] Local binding affinities were determined by plotting the RNase footprinting results as band intensities versus concentration of peptide and fitting the data using equation 1.

xiii. Studies using an AUCA hairpin loop RNA sequence

In the initial design we had chosen a hairpin loop sequence which is expected to have solvent exposed bases in the hairpin loop available for peptide binding. Our results have showed that the peptide preferentially binds to the internal 5'-AAUU-3' bulge region and minimally binds in the loop region. Therefore we designed a second RNA sequence, **AUCAloop RNA**, which has a 5'-AUCA-3' hairpin loop sequence. The change in sequence could make better contacts with the β-hairpin peptide or it could change the structure of the loop which may facilitate binding (Figure 2.26). This structural data for this sequence shows that the loop residues are also solvent exposed (Figure 2.27).[15] This sequence is a member of a class of RNA tetraloop structures with the consensus sequence of RNYA or ANYA (R = purine, N= any nucleotide, Y = pyrimidine, A = adenine). These loops have been found to be important in recognition of phage coat proteins, and the RNA-protein complexes easily crystallize.[16] This particular sequence is derived from the RNA aptamer-MS2 phage coat protein complex, and it was chosen because it contains an adenine base that is flipped out of the loop and not stacked with any nearby bases. The native RNA sequence is adjacent to 5'GG-CC-3' base paired stem region, and this was maintained in our RNA design so as to most accurately mimic the native loop structure.[12]

Fluorescence anisotropy experiments were run to determine the binding affinity of the TAMRA-labeled **WKWK-Int** probe for the new RNA sequence, **AUCAloop RNA** (Figure 2.28). **WKWK-Int** binds the RNA with a dissociation constant of 7.1 ± 1.3 μM, which is

---

[15] Rowsell, S.; Stonehouse, N. J.; Convery, M. A.; Adams, C. J.; Ellington, A. D.; Hirao, I.; Peabody, D. S.; Stockley, P. G.; Phillips, S. E. *Nat Struct Biol* **1998**, *5*, 970-5.

[16] Klosterman, P. S.; Hendrix, D. K.; Tamura, M.; Holbrook, S. R.; Brenner, S. E. *Nucleic Acids Res* **2004**, *32*, 2342-52.

slightly weaker than its affinity for **GGAGloop RNA** (Table 2.4). The new loop sequence could make changes to the global structure of the RNA, but because binding was nearly the same we can conclude that the structural changes are only minor.

RNase footprinting assays were also done to investigate whether **WKWK-Int** is interacting differently with the **AUCAloop RNA**. To visualize the hairpin loop region, RNase A was used because it cleaves uridine and cytidine residues, which are both present in the loop. As was done in previous studies, RNase 1 was used to visualize the bulge region. The footprinting results using **AUCAloop RNA** were very similar to the results with **GGAGloop RNA** (Figure 2.29). We have maintained low micromolar binding affinity in the bulge region, and there is some minimal binding in the loop region. The local binding affinities were calculated for the A23 and U24 residues in the bulge and were found to be 9.4 ± 3.9 μM and 6.9 ± 2.6 μM, respectively (Figure 2.30, Table 2.5). An average of these values is consistent with the binding affinity determined by fluorescence anisotropy. All further studies were done with the **GGAGloop RNA** because the affinity for the **AUCAloop RNA** was not any better and the peptide still seems to prefer binding the internal loop.

AUCAloop RNA

Figure 2.26. Structure and sequence of the **AUCAloop RNA**. The RNA hairpin loop region, which differs in sequence from **GGAGloop RNA**, is highlighted in red.

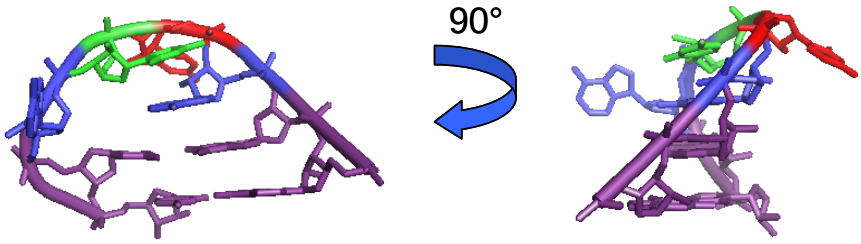


Figure 2.27. Structure of AUCA RNA hairpin loop (5MSF)[14]. The stem region is highlighted in purple, adenosine residues are blue, cytidine residues are green, and uridine residues are red.

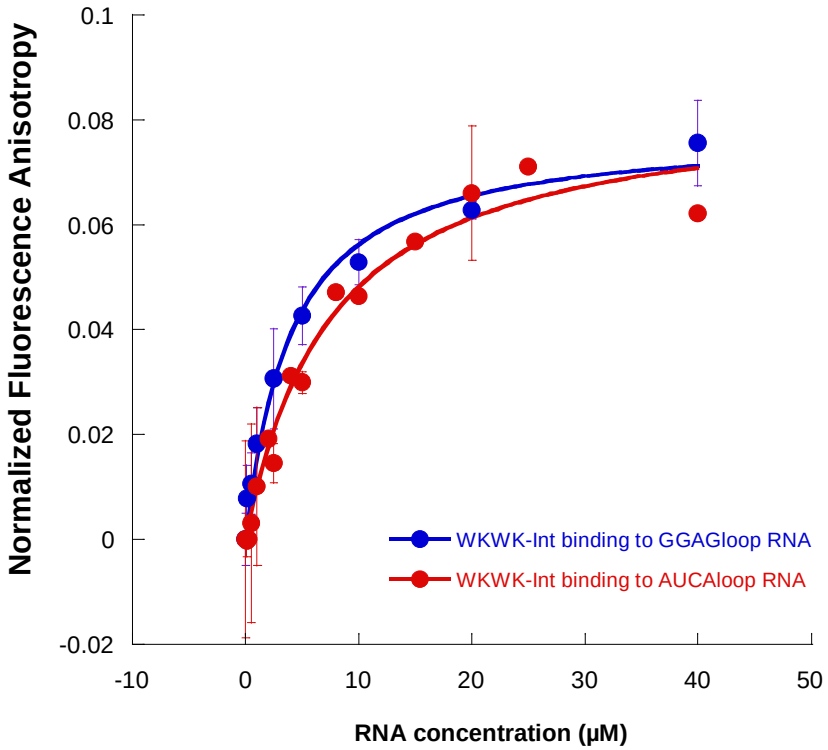


Figure 2.28. Fluorescence anisotropy binding experiments of TAMRA-labeled **WKWK-Int** binding to **GGAGloop RNA** (blue) or **AUCAloop RNA** (red). Peptide concentration was 500 nM in all experiments. Experiments were done in 10 mM PBS buffer (140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) at 25°C and each curve is an average of 3 runs.

Table 2.4. Dissociation Constants for the **WKWK-Int** Binding to the **GGAGloop RNA** and **AUCAloop RNA**.[a]

Peptide-Intercalator Probe	RNA Sequence	K_D (uM) (± error)
WKWK-Int	**GGAGloop RNA**	**3.7 (0.7)**
WKWK-Int	**AUCAloop RNA**	**7.1 (1.3)**

[a] As determined by fluorescence anisotropy experiments. Binding constants were determined by fitting the data to equation 1.

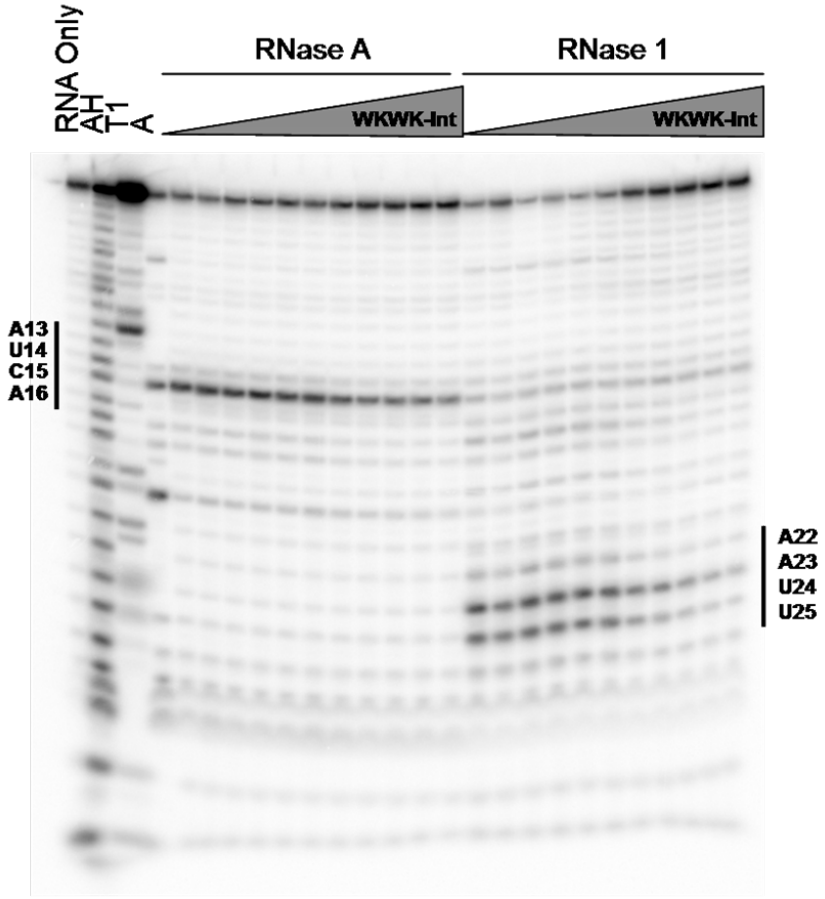


Figure 2.29. RNase footprinting experiments with **WKWK-Int** and **AUCAloop RNA**. Shown is the phosphor image of a 20% denaturing polyacrylamide gel. Structure lanes: RNA only, folded RNA in buffer; AH, alkaline hydrolysis; T1, RNase T1 (G-lane); A, RNase A (C and U lane). Increasing concentrations (0 to 100 μM) of **WKWK-Int** were added to **AUCAloop RNA** (0.18 μM) in 10 mM PBS buffer (140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) and 5 ng/μL yeast tRNA. RNase A (cleaves C and U residues) was used to visualize the loop region and RNase 1 (cleaves ssRNA) was used to visualize both the loop and internal bulge regions.

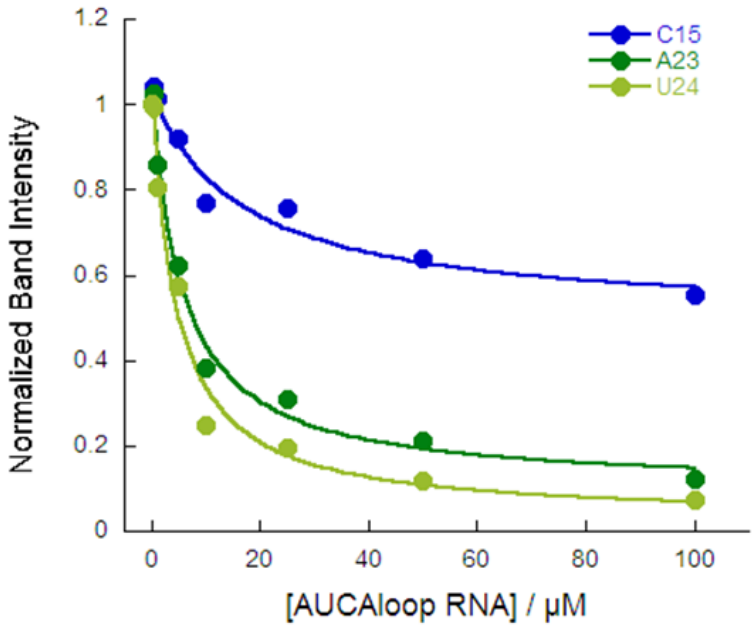


Figure 2.30. ImageQuant analysis of RNase footprinting assay of **WKWK-Int** binding **AUCAloop RNA.** There was weak binding detected in hairpin loop region (blue), but the peptide does bind to the bulge region (dark and light green). The band intensities are the average of three RNase footprinting assays.

Table 2.5. Local Binding Affinities[a] for **WKWK-Int** binding to **AUCAloop RNA**

	Local binding affinity / μM (± error)		
Peptide	**G18**	**A27**	**U29**
WKWK-Int	**16.7 (6.7)**	**5.6 (1.4)**	**4.5 (1.3)**

[a] Local binding affinities were determined by plotting the RNase footprinting results as band intensity versus concentration of peptide and fitting the data using equation 1.

xiv. Studies using an extended length RNA.

The footprinting results showed that the **WKWK-Int** probe bound tightly to the internal AAUU bulge region, but that some weaker binding was occurring in the hairpin loop region. In order to isolate the binding to the bulge region, the **GGAGloop RNA** sequence was extended in length so that the hairpin loop region would be farther away from the intercalation site. The extended RNA, **GGAGext RNA**, made it impossible for the probes to interact with the hairpin loop region, and therefore eliminate any binding in that region (Figure 2.31).

The footprinting experiments with the **GGAGext RNA** showed that **WKWK-Int** no longer binds in the hairpin loop region, but the interaction in the bulge region remained the same (Figure 2.32). Quantitation and analysis of the band densities supports the conclusion that binding in the loop region has been removed, and the binding in the bulge region remains (Figure 2.33). The local dissociation constants for bases A27 and U28 are 2.9 (± 1.7) μM and 2.4 (± 2.2) μM, respectively (Table 2.6). These values are within error of the binding affinities for A23 and U24 in the **GGAGloop RNA**.

The binding experiments show that although we have isolated the peptide binding to the bulge region, the probes have a lower binding affinity. Both **WKWK-Int** and the unstructured **GKGK-Int** have a three-fold lower binding affinity for **GGAGext RNA** than for the shorter RNA sequence (Figure 2.34, Table 2.7). This confirms that the probe has lost the ability to reach the single stranded hairpin loop region, and therefore shows an overall decrease in the binding affinity.

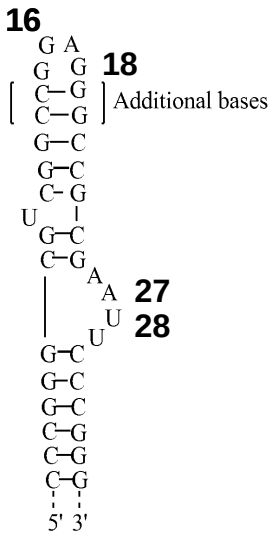


Figure 2.31. Structure and sequence of the **GGAGext RNA**. The additional bases pairs are denoted by parentheses and the key bases in the hairpin loop and internal bulge are labeled with numbers.

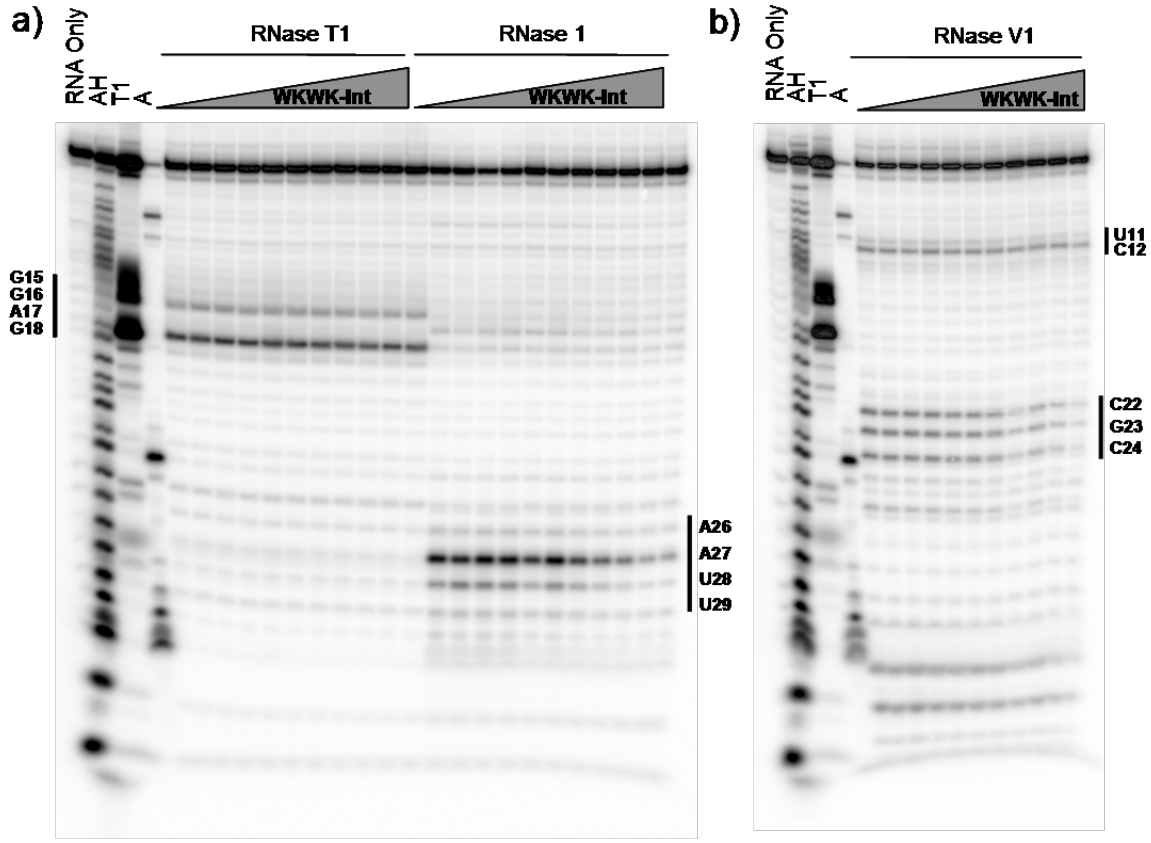


Figure 2.32. RNase footprinting experiments with **WKWK-Int** and **GGAGext RNA**. Shown is the phosphor image of a 20% denaturing polyacrylamide gel. Structure lanes: RNA only, folded RNA in buffer; AH, alkaline hydrolysis; T1, RNase T1 (G-lane); A, RNase A (C and U lane). Increasing concentrations (0 to 100 μM) of **WKWK-Int** were added to **GGAGext RNA** (0.18 μM) in 10 mM PBS buffer (140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) and 5 ng/μL yeast tRNA. a) RNase T1 (cleaves G residues) was used to visualize the G-rich loop region and RNase 1 (cleaves ssRNA) was used to visualize both the loop and internal bulge regions. b) RNase V1 (cleaves dsRNA) was used to visualize the stem regions of the RNA.

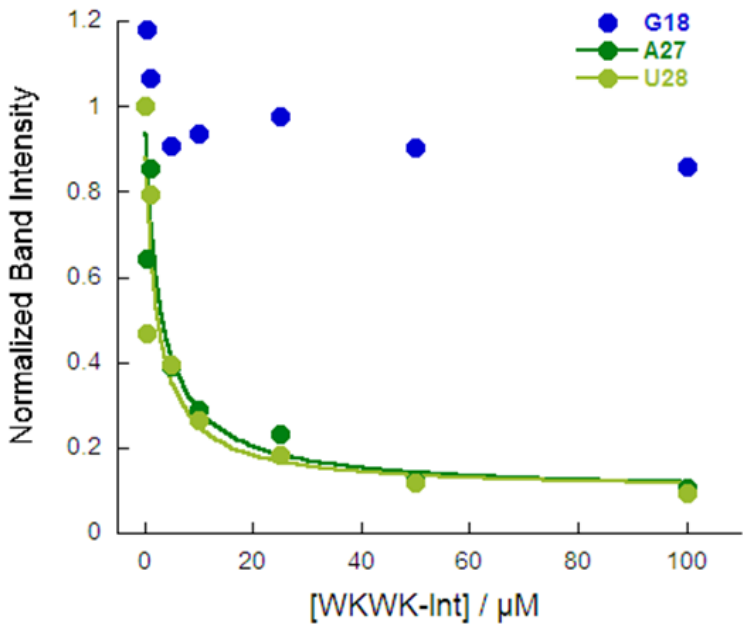


Figure 2.33. ImageQuant analysis of RNase footprinting assay of **WKWK-Int** binding **GGAGext RNA.** There was no binding detected in hairpin loop region (dark and light blue), but the peptide does bind to the bulge region (dark and light green). The band intensity data are the average of two RNase footprinting assays.

Table 2.6. Local Binding Affinities[a] for **WKWK-Int** binding to **GGAGext RNA**

	Local binding affinity / μM (± error)		
Peptide	**G18**	**A27**	**U29**
WKWK-Int	**n.b.**[b]	**2.9 (1.7)**	**2.4 (2.2)**

[a] Local binding affinities determined by plotting the RNase footprinting results as band intensity versus concentration of peptide and fitting the data using equation 1.
[b] no binding detected or binding too weak to determine a dissociation constant

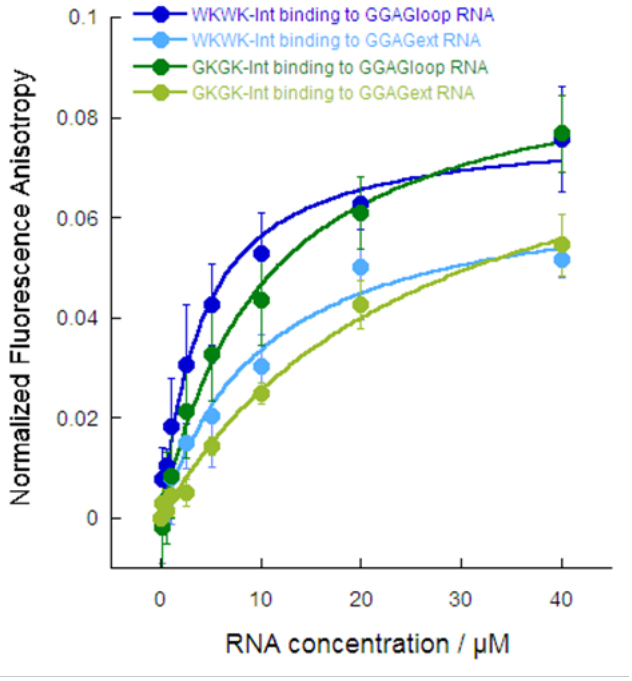


Figure 2.34. Fluorescence anisotropy experiments of TAMRA-labeled probes binding to **GGAGloop RNA** or **GGAGext RNA**. There is a consistent decrease in binding affinity for **WKWK-Int** (dark blue and light blue circles) and **GKGK-Int** (dark green and light green triangles) binding to the extended length RNA. Increasing concentrations of RNA was incubated with the peptide (500 nM) for 30 minutes in 10 mM PBS buffer (140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) at 25°C. Each curve is an average of 3 or 4 runs.

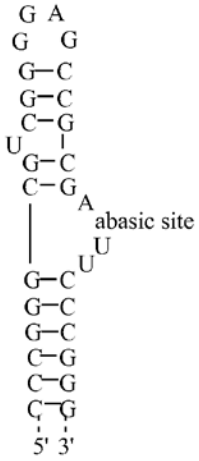
Table 2.7. Dissociation Constants for the Different Peptide Probes Binding to the **GGAGloop RNA** and **GGAGext RNA**.[a]

Peptide-Intercalator Probe	RNA Sequence	K_D (uM) (± error)
WKWK-Int	**GGAGloop RNA**	**3.7 (0.7)**
WKWK-Int	**GGAGext RNA**	**9.7 (2.2)**
GKGK-Int	**GGAGloop RNA**	**9.9 (1.3)**
GKGK-Int	**GGAGext RNA**	**25.6 (5.1)**

[a] As determined by fluorescence anisotropy. Peptide concentration was 500 nM in all experiments. Binding constants were determined by fitting the data to equation 1.

xv. Studies using GGAGloop RNA with an abasic site at A23.

We wanted to investigate further the importance of base A23 in the internal loop of **GGAGloop RNA** for the binding of **WKWK-Int** probe. In order to do this a new RNA sequence was designed, **GGAGloop A23 abasic**, which has an abasic site incorporated at position 23 (Figure 2.35). Fluorescence anisotropy was used to find the dissociation constant of **WKWK-Int** for the abasic RNA sequence. This value was then compared to the K_d found for **GGAGloop RNA**. The anisotropy experiments show that the binding affinity is very similar for both RNA sequences (Figure 2.36). The dissociation constant determined for **WKWK-Int** binding to **GGAG A23 abasic RNA** was 2.6 ± 1.2 μM, which is within error of the value determined for **GGAGloop RNA** (3.7 ± 0.7 μM). These data are somewhat inconclusive, because the effect of eliminating binding at position A23 may be too small to detect when measuring the global binding affinity. Footprinting studies were done, but no relevant information was learned because the RNase 1 enzyme does not cleave at abasic sites. Future work would include incorporation of different RNA bases at that site, and using footprinting to determine whether there is sequence selectivity.



GGAGloop A23
abasic site

Figure 2.35. Structure and sequence of the **GGAG A23 abasic site RNA**. The A23 abasic site is labeled.

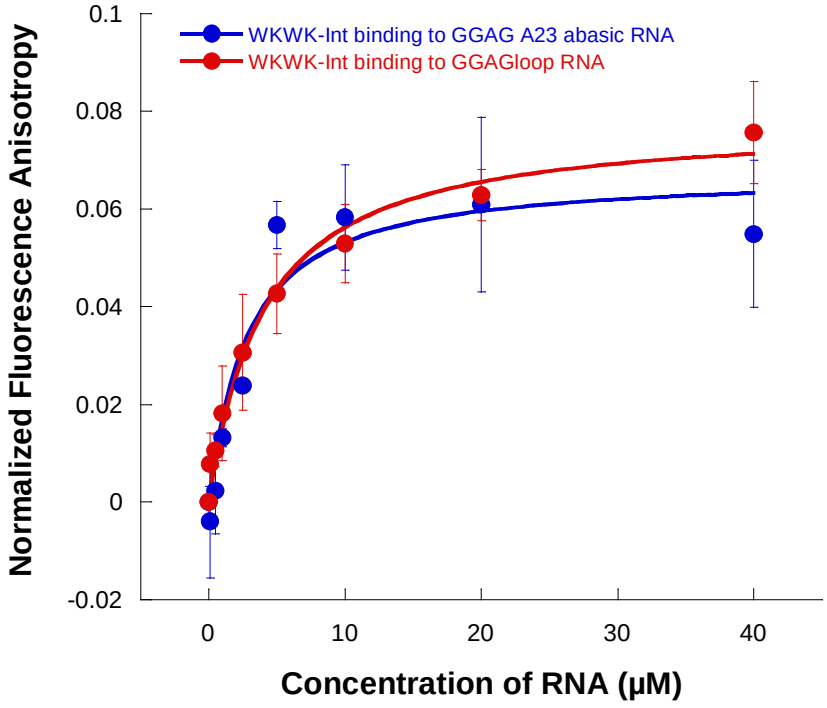


Figure 2.36. Fluorescence anisotropy experiments of TAMRA-labeled probes with **WKWK-Int** binding to **GGAG A23 abasic and GGAGloop RNA** sequences. The binding affinity of WKWK-Int for each sequence is comparable. Increasing concentrations of RNA was incubated with the **WKWK-Int** (500 nM) for 30 minutes in 10 mM PBS buffer (140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) at 25°C. Each curve is an average of 2 or 3 runs.

C. Conclusions

We have designed a peptide-intercalator conjugate that recognizes RNA by interacting with both the double stranded stem region and single stranded bulge region simultaneously. The binding data shows that both the intercalator and the peptide portion of the probe are required for low micromolar binding to occur. This system provides an example of how two ligands with low affinity for a substrate can together create a tighter binding system. This dual-mode method of binding creates a system that is specific for two types of structure, and may be a viable way of providing specificity for RNA binding molecules.

The RNase footprinting experiments showed that **WKWK-Int** interacts strongly with the internal bulge region (residues A23 and U24) and moderately with the hairpin loop region (residue G16). The cleavage patterns for **GKGK-Int** and **Int-control** probe looked similar, but neither of these probes protected the A23 residue. The cleavage pattern for **GWKWK-Int** also looked similar, but it had a weak affinity for the A23 residue. We can conclude that specific interactions are being made between the structured peptide and this residue. Aromatic interactions are important for this interacation to occur, but the β-hairpin structure increases the affinity. Because the residue is located in the single stranded bulge, the base may be exposed and binding into the aromatic pocket of the peptide. This hypothesis is supported by structural data for a similar adenine-rich bulge.[17]

The design of a probe that contained a fully folded version of the WKWK peptide, **cycWKWK-Int**, lowered the entropic cost for folding but also caused the peptide structure to be more constrained. The results showed that this did not seem to help nor hurt the binding

[17] Hermann, T.; Patel, D. J. *Structure* **2000**, *8*, R47-54.

affinities, and both **cycWKWK-Int** and **WKWK-Int** bound the GGAGloop RNA in a similar conformation.

When the directionality of the peptide-intercalator probe was reversed, we found comparable protection from RNase cleavage. Therefore, we conclude that the intercalator is threading in either direction and the peptide is driving the binding to the bulge region. If the intercalator was driving the binding orientation, we would have seen either a loss of protection at the bulge or a different region of the RNA would have been protected, as the **revWKWK-Int** would place the peptide in the opposite groove.

The sequence of the RNA hairpin loop was changed in an attempt to create more favorable interactions with that region. The **WKWK-Int** probe bound nearly the same to AUCAloop RNA as it did to GGAGloop RNA. The results show that the sequence of this region can be changed without a negative affect on binding. However it also further proves that the loop region does not make significant interacations with the peptide portion of **WKWK-Int**.

We have also shown that by extending the length of the RNA, we can terminate binding in the hairpin loop region. The loss of binding results in a three-fold decrease in the binding affinity of both the structured and unstructured peptide intercalators for the RNA. The intermolecular distance between the intercalator and the β-hairpin peptide is such that it is not long enough to reach the hairpin loop once the RNA sequence is extended.

In this study, we have tested a β-hairpin "mini-protein" that has a binding pocket for unpaired nucleotides for binding to RNA loop and bulge regions. We used an RNA intercalator that not only added an additional point of binding, but also helped to orient the peptide near the bulge region. This is one of the only examples of a multivalent RNA

binding motif that interacts with a specific site on the RNA and simultaneously makes multiple interactions.

Further studies using this probe would focus on identifying whether there is sequence specificity in the bulge region. This could be done using an RNA sequence where the sequence of the bulge region is changed, especially the A23 base which seems to be important for binding the β-hairpin peptide. Additionally the tryptophan binding pocket could be mutated to scan which of the residues are making important interactions with the RNA. We suspect that the RNA is interacting with the non-hydrogen bonded face (Trp-Lys-Trp-Lys) of the peptide, but it could be confirmed through peptide mutations.

D. Experimental section

i. Synthesis of the TAMRA-labeled intercalator-peptide probes

The intercalator portion, *4-(4'-Methylaminoallyoxycarbamate)aniline-quinoline-8-carboxylic acid*, was synthesized following the previously published procedure by Krishnamurthy et al. (Scheme 2.2) [10] One-dimensional NMR was used to confirm the identity of the intermediates and final product (Figures 2.37-2.42). The peptide portions were synthesized on FMOC-PAL-PEG-PS resin (0.07 mmol scale, 150 mg resin) using standard FMOC solid phase peptide synthesis methodology on a Creosalus Tetras automated synthesizer until the position where FMOC-Lys(ivDde)-OH (0.14 mmol, 2 eq.) was coupled manually. The peptide was put back onto the synthesizer for coupling of the remainder of the natural amino acids. After the βAla-βAla-Arg linker was added, the intercalator portion was manually coupled overnight using HOBT/HBTU (4 eq. each) and DIPEA (8 eq.). The N-terminal Alloc protecting group was deprotected using $Pd(PPh_3)_4$ (1 eq.) and phenylsilane (24 eq.) in dry

dichloromethane. The FMOC-Val-OH and BOC-Ser-OH amino acids, respectively, were coupled using HOBT/HBTU and DIPEA (0.4 mmol of each) in DMF. The peptides were fluorescently labeled by first deprotecting the ivDde protecting group with 2% hydrazine in DMF three times for 3 minutes each time. The free amine on the lysine side chain was coupled overnight to 5-(and -6)-carboxytetramethylrhodamine (TAMRA) (0.16 mmol). The BOC group on the N-terminus was used in order to keep the N-terminus protected during the hydrazine deprotection and fluorophore coupling steps. All the peptides were cleaved and deprotected by treating the resin with a (95:2.5:2.5) TFA:TIPS:Water mixture and bubbling with N_2 for 3 hours. The peptides were isolated by evaporating the TFA with nitrogen gas, then precipitating the peptide in cold diethyl ether, and then peptide extraction using MilliQ water. The resulting solution was frozen and lyophilized to give a dark pink powder. The peptides were purified using a Waters reverse-phase HPLC system. During purification a C18 column was used, and the peptides were eluted using Standard A (95% water, 5% acetonitrile, 0.1% TFA) and Standard B (95% acetonitrile, 5% water, 0.1% TFA). The masses of the peptides were confirmed using ESI-TOF mass spectrometry (Table 2.8).

ii. Cyclization reaction for cycWKWK-Int

After cleavage from the resin, **cycWKWK-Int** peptide was lyophilized to dryness then redissolved in 25 mL of 10 mM phosphate buffer (pH 7.5) and 250 μL DMSO. The solution was stirred at room temperature overnight and then lyophilized to dryness. The peptide was purified using reverse-phase HPLC and masses were confirmed by mass spectrometry as described above.

iii. Secondary structure determination using circular dichroism

Studies were done using an Aviv Model 62DS Circular Dichroism Spectrometer. A wavelength scan (185nm-260nm) was done of **WKWK-Int, GKGK-Int, GWKWK-Int,** and **cycWKWK-Int** in 10 mM sodium phosphate buffer (pH 7.6) at 37°C in a 0.1 cm cell. The peptide concentration was 100 μM in both cases and the scans were done in triplicate.

iv. Fluorescence anisotropy binding experiments

The RNAs with defined sequences were purchased as a purified, lyophilized powder from Integrated DNA Technologies. The RNA was dissolved in DEPC-treated water to a concentration of 500 μM and annealed by heating to 95°C for 2 min, then slowly cooling to room temperature. The folded RNA solution was separated into smaller aliquots and stored at -20°C. The peptide-intercalator probes were dissolved in DEPC-treated water and the concentration of the solution was determined using UV-Vis (λ_{559}=91000 M^{-1} cm^{-1}) [18]. The binding experiments were done by incubating 0.5 μM TAMRA-labeled peptide-intercalator with increasing concentrations of folded RNA in reaction buffer (10 mM phosphate buffer, 140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) for 30 minutes at room temperature. The anisotropy data were collected at 25°C in a 1 cm pathlength quartz microcuvette (excited at 548 nm and emission monitored at 580 nm). The anisotropy data were plotted as a function of fluorescence anisotropy versus concentration of RNA added. Each of the curves was normalized so that the initial fluorescence anisotropy point equaled zero. The plot was fitted using **equation 1** derived from a previous report:[19]

[18] Extinction coefficient provided by Integrated DNA Technologies, www.idtdna.com.

[19] Wang, Y.; Killian, J.; Hamasaki, K.; Rando, R. R. *Biochemistry* **1996**, *35*, 12338-46.

$$r = \left(\left(\frac{(a + x + K_D) \pm \sqrt{(-a - x - K_D)^2 - 4(a \cdot x)}}{2 \cdot a} \right) \cdot (r_\infty - r_0) \right) + r_0 \qquad (Equation\ 1)$$

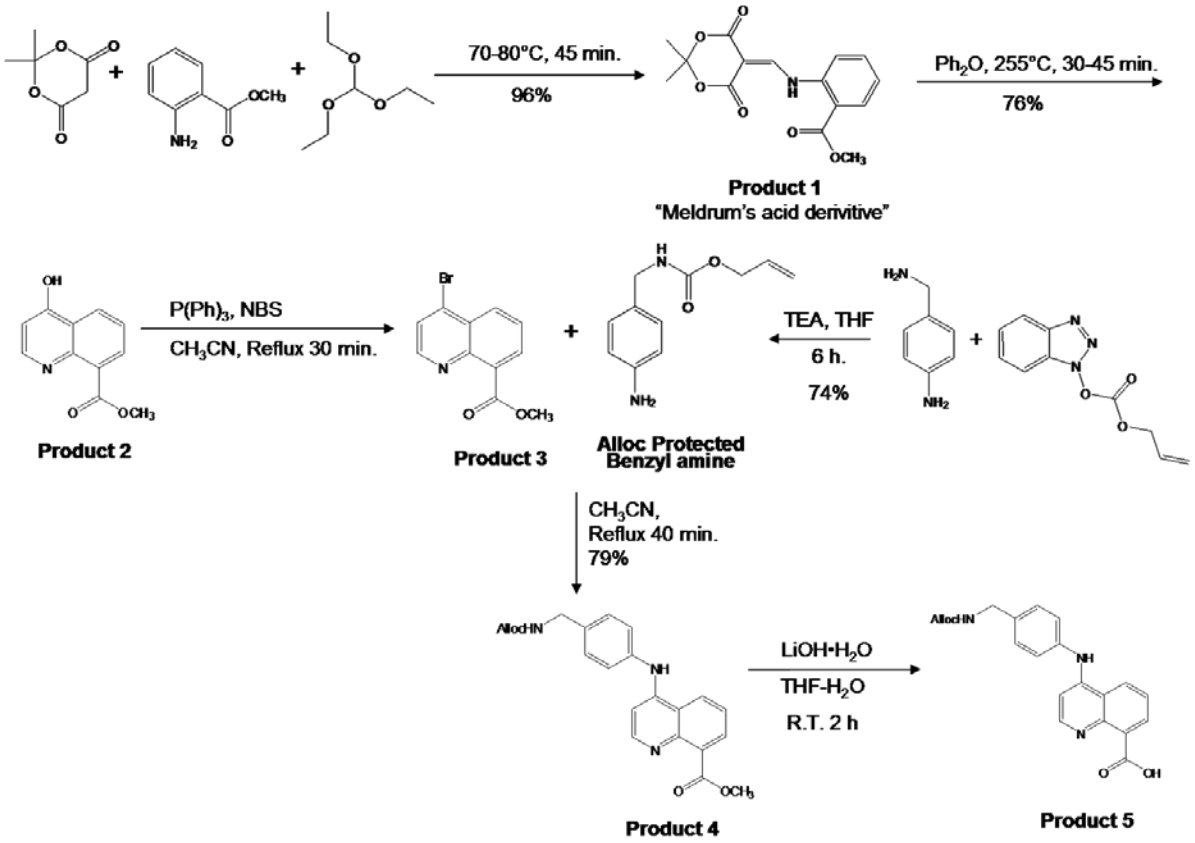
where r is anisotropy, r_0 is initial anisotropy value, r_∞ is the maximum anisotropy, a is the total peptide-intercalator conjugate concentration, x is the concentration of RNA added, K_D is the dissociation constant.

v. Quantitative RNase footprinting experiments

3'-^{32}P RNA labeling. Approximately 10 μg of RNA construct was treated with T4 RNA ligase (2 U/μL), T4 RNA ligase buffer (1X), DMSO (10%), RNase inhibitor (1 U/μL), and ^{32}pCp (0.05 mCi) and incubated at 4°C overnight. The labeled RNA was run through an RNA spin column and loaded onto a 20% denaturing polyacrylamide gel for purification. The labeled RNA was visualized by autoradiography, excised from the gel and eluted overnight with elution buffer (0.5 M NH_4OAc, 1 mM EDTA, 0.1 % SDS, pH 8.0). After filtration through a Millipore centrifugal filter device (PTFE membrane 0.20 μM), the RNA was precipitated with 95% ethanol, sodium acetate buffer, and linear acrylamide overnight at -20°C. The RNA was brought up in DEPC-treated water and the concentration was determined by UV absorbance using the equation: concentration of RNA SAMPLE = 40 × OD_{260} × dilution factor.

RNase footprinting experiments. The labeled RNA was brought up in DEPC-treated water (360 nM) and annealed at 95°C for 2 min and slowly cooled to room temperature. The peptide-intercalator ligands were synthesized using the same procedure as above except without the fluorophore label. Instead of inserting an FMOC-Lys(ivDde)-OH residue, an FMOC-Orn(Boc)-OH residue was used. The unlabeled peptide-intercalator probes were dissolved in DEPC-treated water and the concentration of the solution was determined using

the absorbance of the quinoline molecule (ε_{359}=11700 M^{-1} cm^{-1}).[10] The folded RNA (final concentration =180 nM) was incubated with increasing concentrations of the peptide ligands, 5 μg/mLyeast tRNA, and 10 mM PBS buffer (140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) for 30 minutes at room temperature. RNase digestions were done with RNase T1 (0.1 U/μL), RNase 1 (10 U/μL), and RNase V1 (0.01 U/μL) for 5 min, 30 s, and 1 min, respectively, and quenched with loading dye. Sequencing lanes were done using unfolded RNA and 1X alkaline hydrolysis buffer (purchased from Ambion), RNase T1, and RNase A (1 μg/mL). Digested RNA and sequencing reactions were heat denatured and loaded onto a 20% denaturing polyacrylamide gel electrophoresis (51W, 2.5 h). The gel was visualized by phosphorimaging, and the intensity of the bands was quantified using ImageQuant analysis program. In order to correct the data for differences in loading, the intensity of each band was divided by the intensity of a control lane which is unchanged during the reaction. To normalize the data, the corrected band intensity was then divided by the intensity of the RNA band when no peptide-intercalator probe was added. The data were averaged and plotted on Kaleidograph as normalized band intensitry versus concentration of peptide-intercalator added and fit using **equation 1**. In this case, r is the normalized band intensity, r_0 is initial band intensity, r_∞ is band intensity at the maximum peptide-intercalator conjugate concentration, a is the total RNA concentration, x is the total concentration of peptide-intercalator conjugate added, K_d is the dissociation constant.



Scheme 2.2. Synthesis of the intercalator, *4-(4'-Methylaminoallyoxycarbamate)aniline-quinoline-8-carboxylic acid.* [10]

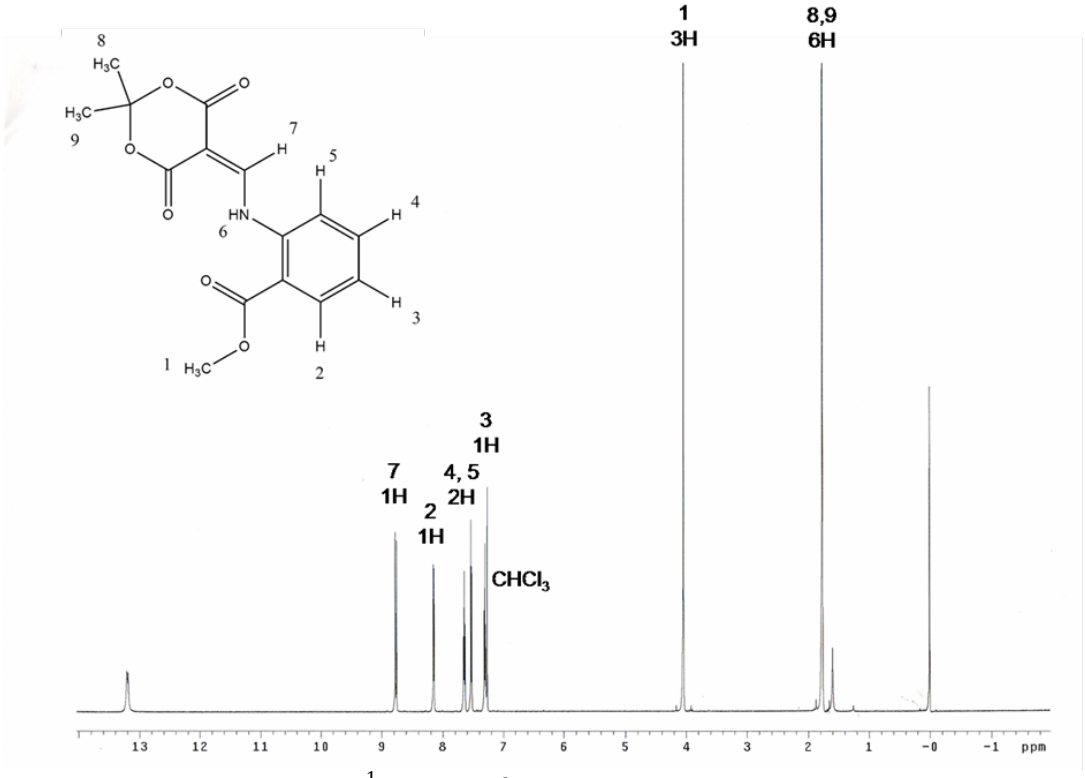


Figure 2.37. ^{1}HNMR of Product 1 dissolved in $CDCl_3$.

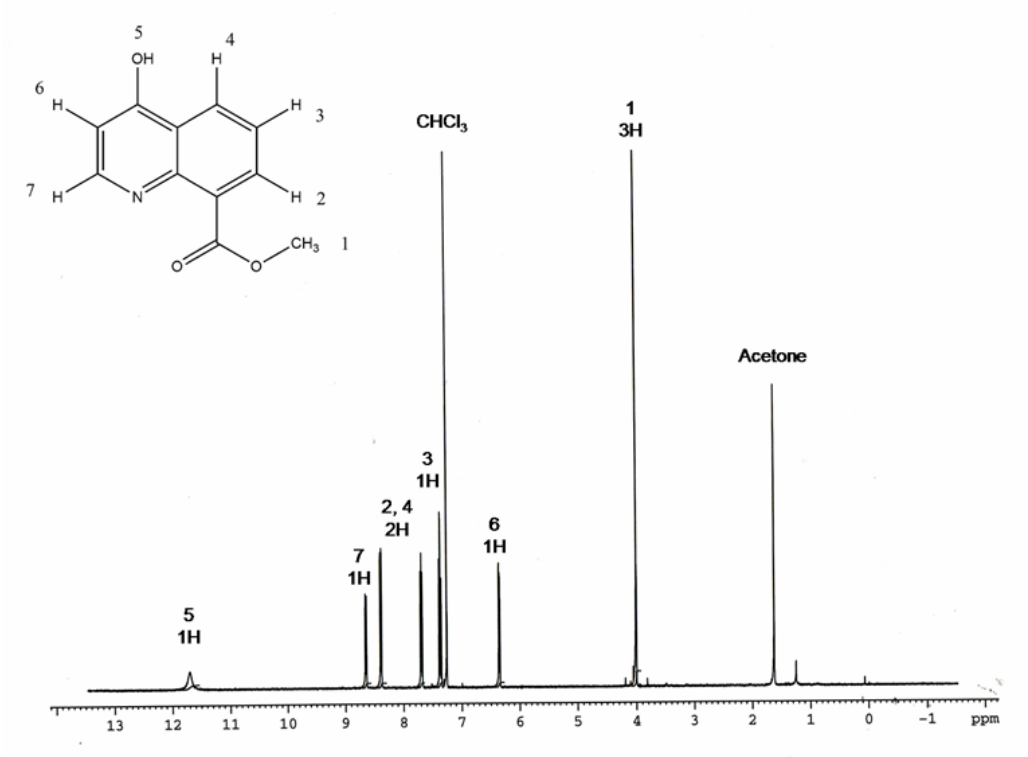


Figure 2.38. ^{1}HNMR of Product 2 dissolved in $CDCl_3$.

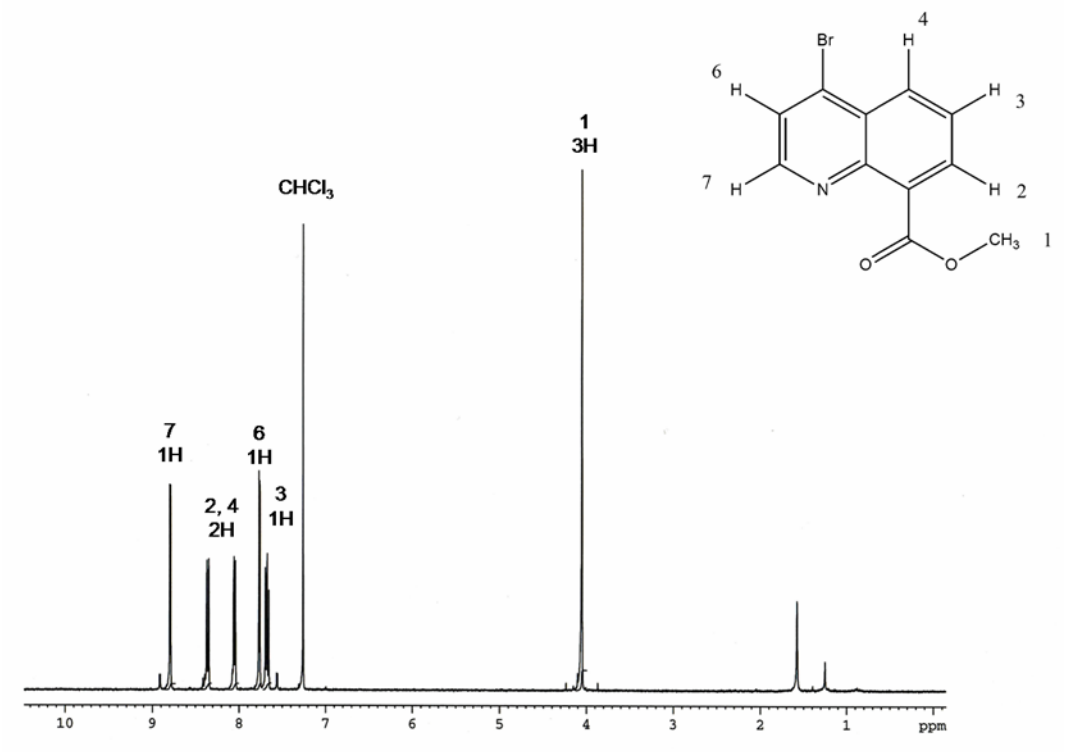


Figure 2.39. [1]HNMR of Product 3 dissolved in $CDCl_3$.

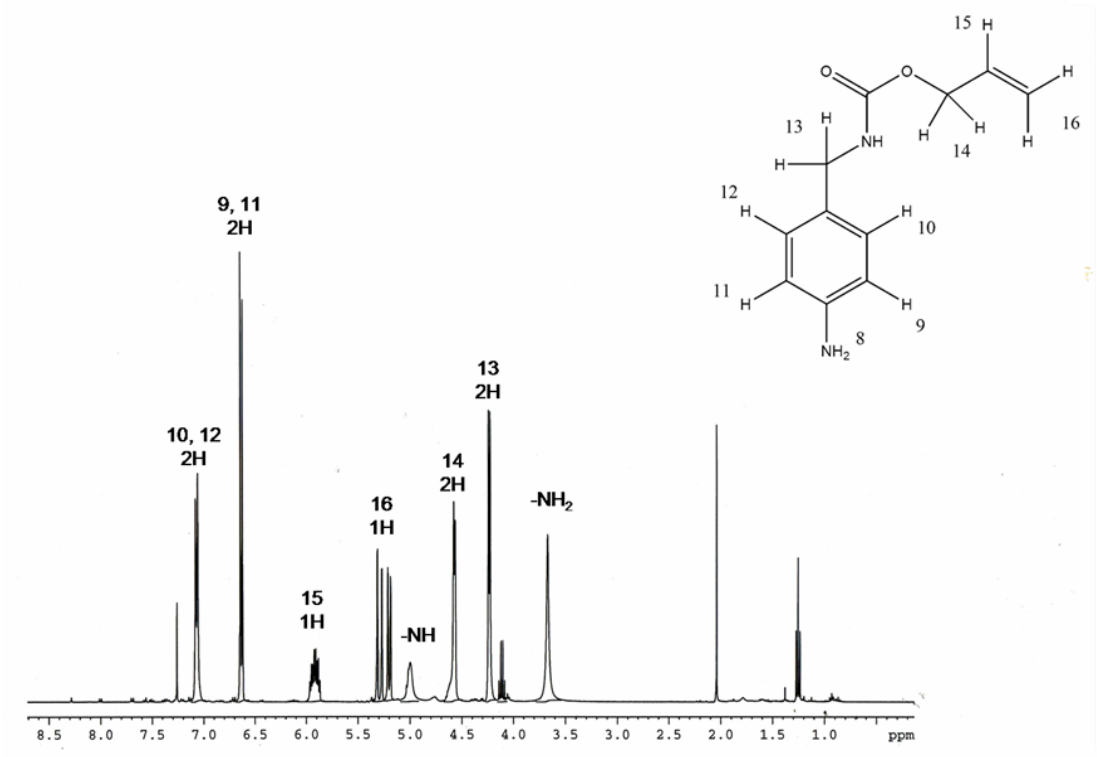


Figure 2.40. [1]HNMR of the Alloc-protected benzyl amine, *allyl-4-aminobenzylaminocarbamate*, dissolved in $CDCl_3$.

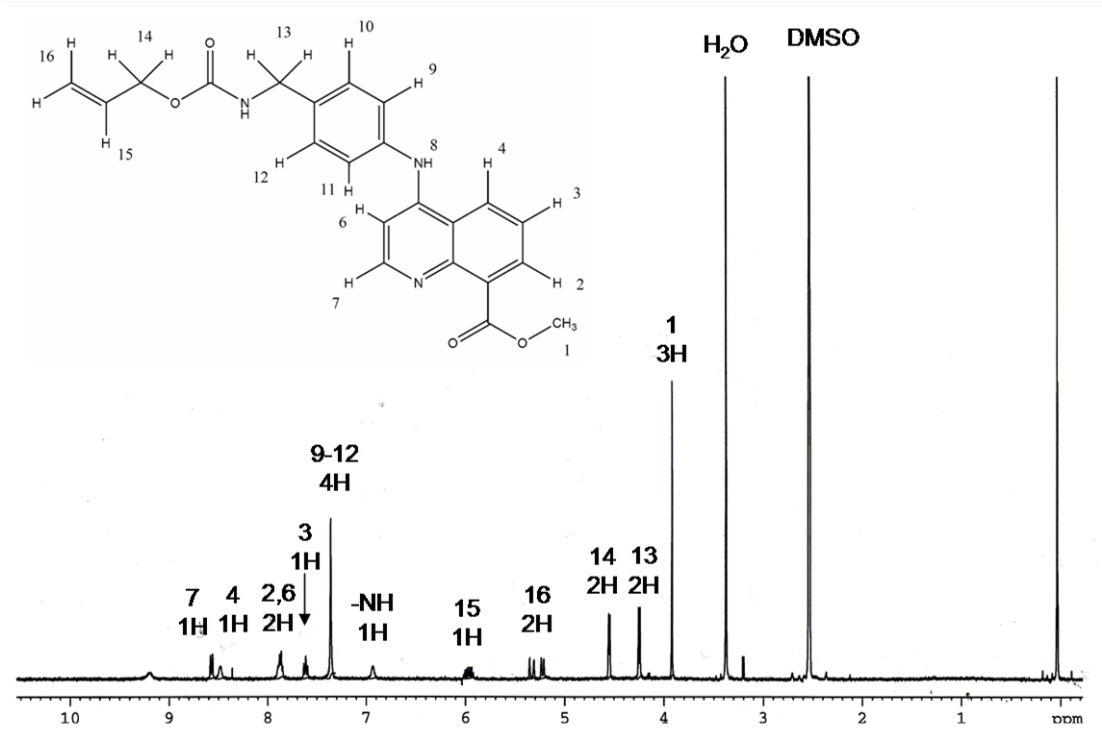


Figure 2.41. ^{1}HNMR of Product 4 dissolved in $(CD_3)_2SO$.

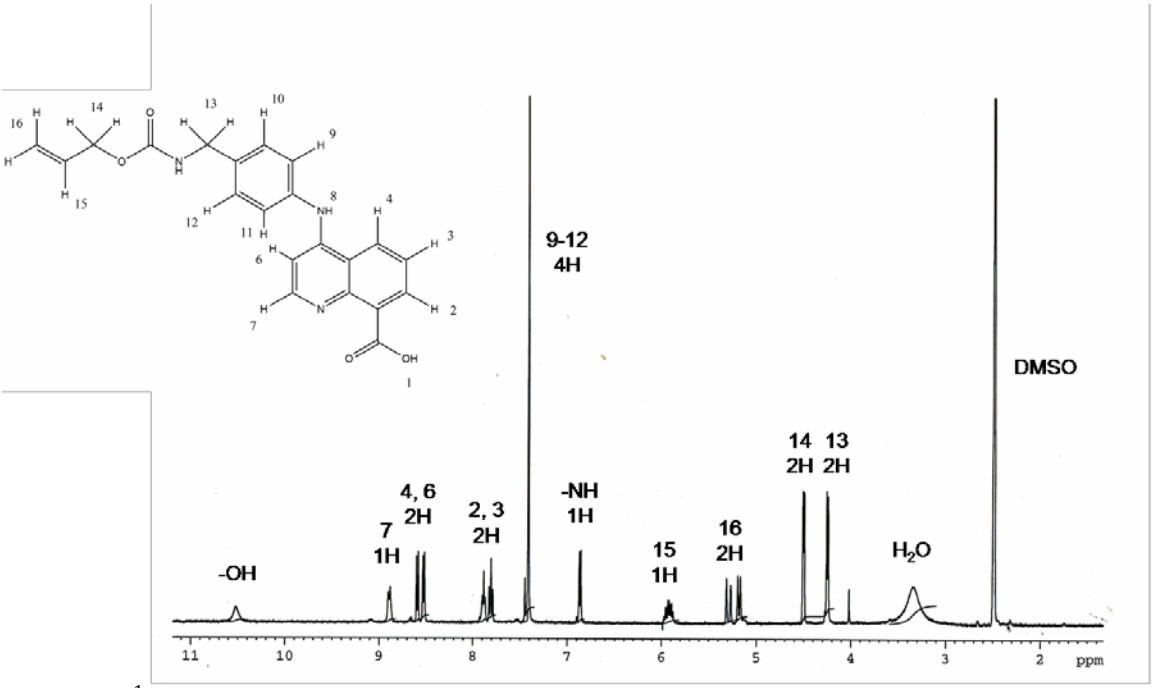


Figure 2.42. ^{1}HNMR of Product 5 dissolved in $(CD_3)_2SO$.

Table 2.8. Mass Spectrometry Data Obtained for Peptides Studied in this Work

Peptide	Expected Mass (Da)	Observed Mass (Da)
WKWKcontrol	1710.0	1710.4
WKWKcontrol(TAMRA)	2136.1	2136.2
Intcontrol	776.4	776.6
Intcontrol(TAMRA)	1316.6	1316.8
WKWK-Int	2285.3	2285.0
WKWK(TAMRA)-Int	2711.5	2711.5
GKGK-Int	1759.0	1758.9
GKGK(TAMRA)-Int	2187.2	2186.1
GWKWKG-Int	2017.1	2017.1
GWKWKG(TAMRA)-Int	2445.3	2444.9
cycWKWK-Int	2489.3	2489.6
cycWKWK(TAMRA)-Int	2915.5	2915.8
revWKWK-Int	2285.3	2286.0
revWKWK(TAMRA)-Int	2710.5	2711.1

Chapter III

DESIGN OF A BIV TAT PEPTIDE LIBRARY FOR INCREASING BINDING AFFINITY TO BIV TAR RNA

A. Background

In a second approach toward investigating peptides that interact with RNA, we have used combinatorial chemistry to improve the affinity of a peptide to an RNA hairpin loop region. RNA is different from DNA because it commonly contains unique structural elements such as loops and bulges. These elements are often recognized by proteins using an induced fit mode of binding. For these reasons the design of RNA binding molecules is challenging because it is difficult to predict how the RNA structure will change to accommodate the ligand. We believe a combinatorial chemistry approach may overcome these design hurdles.

The use of combinatorial chemistry to find different molecules that bind RNA targets has become prevalent in the past decade. Some recent examples use computation methods to design libraries,[1] high throughput fluorescence polarization assays,[2] high throughput cellular

---

[1] (a) Schueller, A.; Suhartono, M.; Fechner, U.; Tanrikulu, Y.; Breitung, S.; Scheffer, U.; Goebel, M. W.; Schneider, G. *J Comput -Aided Mol Des* **2008**, *22*, 59-68. (b) Mayer, M.; James, T. L. *Nuclear Magnetic Resonance of Biological Macromolecules, Part C* **2005**, *394*, 571-587.

[2] Galicia-Vazquez, G.; Lindqvist, L.; Wang, X.; Harvey, I.; Liu, J.; Pelletier, J. *Anal Biochem* **2009**, *384*, 180-8.

assays, [3] and dynamic combinatorial chemistry.[4] Several groups have used combinatorial methods to find peptide-based ligands that bind to RNA.[5] The use of a peptide library has the advantage of being amenable to solid phase synthesis of the library and which also simplifies the screening process. This was exemplified by the work of Hwang et al., who synthesized a tri-peptide library and performed an on-bead screen for ligands that bound the bulge loop of TAR RNA.[6] They found that even though some of the peptide is displayed on the interior of the resin bead, the peptides were still accessible for RNA binding. The on-bead method also made it possible to screen against non-specific binding to undesirable RNA sequences.

We proposed to use a similar on-bead screening methodology, but instead of trying to find novel RNA-binding ligands, we wanted to improve upon a known RNA-peptide interaction. The benefit of this approach was that we were able to use the binding affinity of the native peptide as a guideline to compare to newly discovered molecules. We also used this as a gauge to tell how well our library screens worked. We chose to investigate a well-characterized RNA-protein interaction, the transactivation response (TAR) RNA sequence

---

[3] Jimenez Bueno, G.; Klimkait, T.; Gilbert, I. H.; Simons, C. *Bioorg Med Chem* **2003**, *11*, 87-94.

[4] (a) McNaughton, B. R.; Gareiss, P. C.; Miller, B. L. *J Am Chem Soc* **2007**, *129*, 11306-7. (b) Karan, C.; Miller, B. L. *J Am Chem Soc* **2001**, *123*, 7455-6. (c) Gareiss, P. C.; Sobczak, K.; McNaughton, B. R.; Palde, P. B.; Thornton, C. A.; Miller, B. L. *J Am Chem Soc* **2008**, *130*, 16254-61. (d) Bugaut, A.; Toulme, J. J.; Rayner, B. *Org Biomol Chem* **2006**, *4*, 4082-8.

[5] (a) Tisne, C.; Dardel, F. *Comb Chem High Throughput Screen* **2002**, *5*, 523-9. (b) Tisne, C.; Guilliere, F.; Dardel, F. *Biochimie* **2005**, *87*, 885-8. (c) Hamy, F.; Felder, E. R.; Heizmann, G.; Lazdins, J.; Aboul-ela, F.; Varani, G.; Karn, J.; Klimkait, T. *Proc Natl Acad Sci U S A* **1997**, *94*, 3548-53. (d) Ahn, D. R.; Yu, J. *Bioorg Med Chem* **2005**, *13*, 1177-83.

[6] Hwang, S.; Tamilarasu, N.; Ryan, K.; Huq, I.; Richter, S.; Still, W. C.; Rana, T. M. *Proc Natl Acad Sci U S A* **1999**, *96*, 12997-3002.

from bovine immunodeficiency virus (BIV) bound to the transactivation (bTat) protein (Figure 3.1). This was an attractive system to study because (1) there was extensive structural information known about it, (2) the bTat binding domain is a small peptide that is amenable to combinatorial chemistry, and (3) previous results suggested that one region of the peptide was not making tight contacts with the RNA loop region which allowed us to investigate binding to RNA loops.

The TAR RNA stem-loop-bulge structure and sequence has been shown to be required for Tat binding and that interaction is required for viral replication.[7] The bTat-TAR interaction has high specificity and affinity, and unlike human immunodeficiency virus Tat-TAR, there is a published NMR structure of the TAR RNA sequence bound to the minimal recognition sequence of bTat (Figure 3.1).[8] Mutagenesis studies have shown that the residues important for binding are all located in the C-terminal β-hairpin of the bTat peptide (Figure 3.2).[9] However, the NMR structure shows that residues in the N-terminal loop (residues 65-69) of the peptide are in close proximity to the RNA hairpin loop, although these interactions are not detected in the mutagenesis studies. The TAR RNA hairpin loop has bases 17-20 flipped out and not base-paired to each other, and there have been no reported studies to date that alter the N-terminal loop of bTat peptide in order to better interact with the exposed bases. We proposed to improve binding by using combinatorial

[7] (a) Fisher, A. G.; Feinberg, M. B.; Josephs, S. F.; Harper, M. E.; Marselle, L. M.; Reyes, G.; Gonda, M. A.; Aldovini, A.; Debouk, C.; Gallo, R. C; et al. *Nature* **1986**, *320*, 367-71. (b) Dayton, A. I.; Sodroski, J. G.; Rosen, C. A.; Goh, W. C.; Haseltine, W. A. *Cell* **1986**, *44*, 941-7.

[8] Ye, X.; Kumar, R. A.; Patel, D. J. *Chem Biol* **1995**, *2*, 827-40.

[9] Chen, L.; Frankel, A. D. *Proc Natl Acad Sci U S A* **1995**, *92*, 5077-81.

chemistry to find mutations in that region which improve binding to the RNA hairpin loop which would also improve the overall binding of the system (Figure 3.2).

We also wanted to incorporate some rational design elements in the combinatorial libraries, so we have incorporated both natural and unnatural residues that could make aromatic, cation-π, hydrogen bonding, and hydrophobic interactions. We also aimed to avoid library screens resulting in hits that are dominated by electrostatic interactions, as they typically lead to nonspecific interactions.

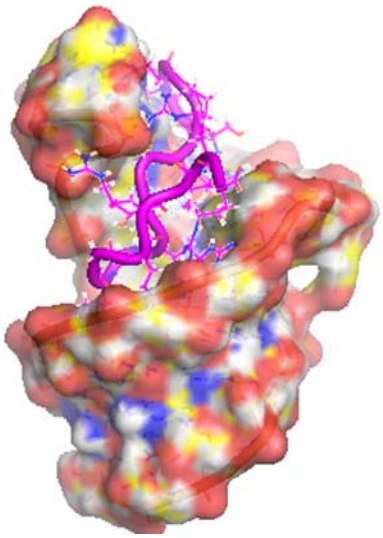


Figure 3.1. NMR structure of BIV Tat-TAR (PDB 1BIV). TAR RNA stem-loop bound to Tat peptide. The C-terminus of the peptide is shown in the foreground with the β-hairpin extending to the bottom left. The N-terminus loop is near the RNA hairpin loop at the top of the picture.

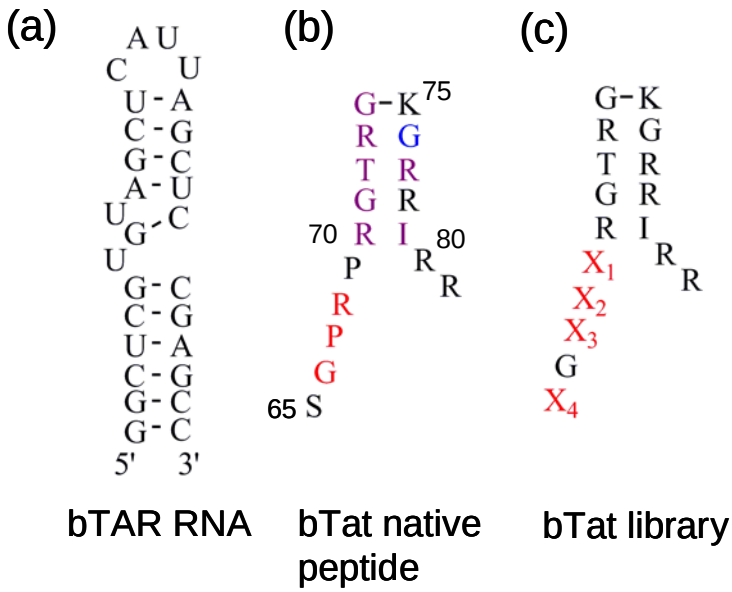


Figure 3.2. Sequence and secondary structure of bTAR RNA, the binding domain of the bTat protein, and the proposed bTat library. (a) bTAR RNA sequence and secondary structure. Expected base pairing is indicated by dashed lines. (b) bTat native peptide sequence and bound structure with β-sheet portion and frayed ends. Mutagenesis studies have isolated the residues important for binding (highlighted in blue and purple), but the NMR structure showed that N-terminal residues also contact the RNA (highlighted in red and purple). (c) Proposed design of a bTat peptide library.

B. Library design and synthesis

The bTat library was designed after studying the NMR structure of the bTat peptide bound to bTAR. The structure showed that there are several possible contacts between Pro69, Arg68, and Pro67, but that the Ser65 did not significantly interact (Figure 3.3). A variety of residues were incorporated at these positions in an attempt to increase binding in that region. We used aromatic residues to promote hydrophobic or π-stacking interactions, polar residues to increase hydrogen bond interactions, basic residues for cation-π or electrostatic interactions, and non-polar residues to find favorable hydrophobic interactions (Figure 3.4). We also included several non-natural amino acids to increase the possibility of finding a novel RNA binding interaction (Figure 3.5). There were 2640 total possible

compounds in the biased library, and two glycine residues were used a spacer between the bead the bTat library. We also synthesized a positive control which was the bTat native peptide synthesized on the Tentagel beads.

Synthesis of the library was done by the split and pool method (Scheme 3.1).[10] The benefits of using split and pool synthesis is that each bead only contains one compound. This simplifies hit identification, because single bead sequencing can be done to determine which peptide bound to the RNA. Tentagel resin was used because it contains a non-cleavable linker, and because the swelling properties of the resin work well in aqueous media.

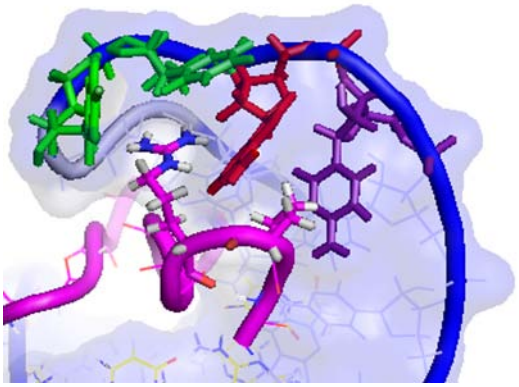


Figure 3.3. RNA-peptide interactions near the N-terminal loop of bTat peptide (PDB BIV1) displayed using PyMol. The peptide is shown in magenta with Arg68 and Pro67 side chains interacting with two uracils (green), an adenine (red), and a cytosine (purple).

[10] Lam, K. S.; Lebl, M.; Krchnak, V. *Chem Rev* **1997**, *97*, 411-448.

●-Non-cleavable linker-GGRRIRRGKGRTGR$X_4X_3X_2GX_1$-NH_3^+

X_1= S, K, R, W, Y, or A

X_2= P, F, W, PhG, or A

X_3= R, $R(Me)_2$, $K(Me)_2$, Cit , F, W, Y, hF, PhG, pfF, or A

X_4= P, G, R, K, Y, W, F, or A

Figure 3.4. BIV Tat library sequence including all the possible amino acid substitutions. One letter abbreviations are used for all the residues except the non-natural amino acids which are labeled as followed: phenylglycine (PhG), dimethyl arginine ($R(Me)_2$), dimethyl lysine ($K(Me)_2$), citrulline (Cit), homophenylalanine (hF), and parafluorophenyalanine (pfF).

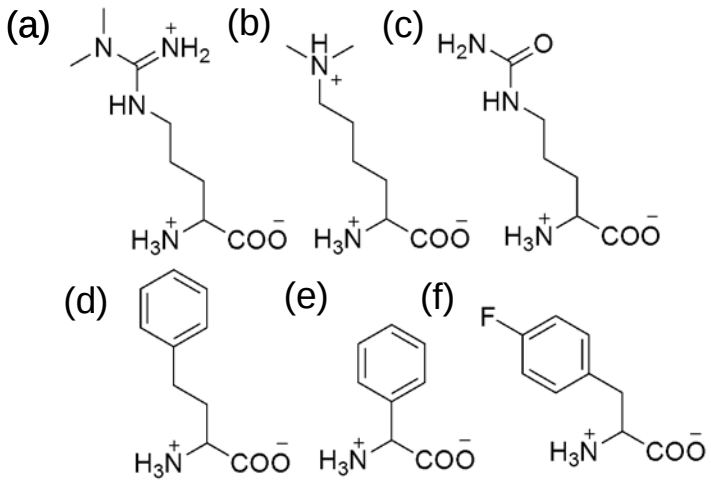
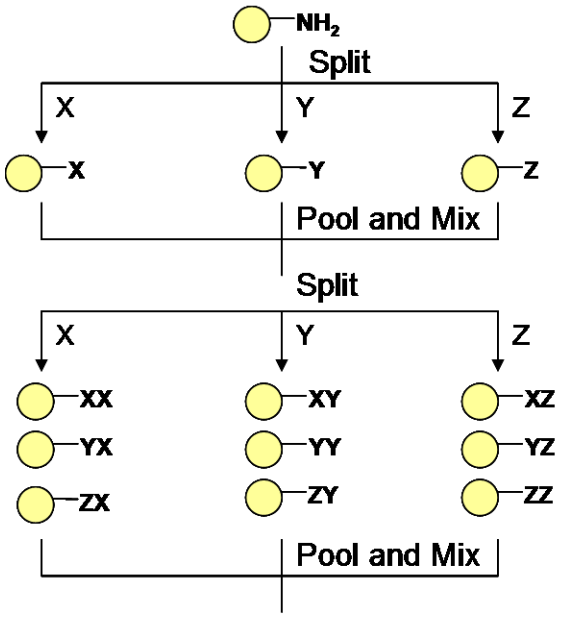


Figure 3.5. Stuctures of the non-natural amino acids used in the synthesis of bTat library. (a) dimethyl arginine, (b) dimethyl lysine, (c) citrulline, (d) homophenylalanine, (e) phenylglycine, and (f) parafluorophenyalanine.



Scheme 3.1. Diagram of split and pool synthesis used to make the bTat peptide library. This method of synthesis results in each bead containing one peptide sequence.

C. Results and discussion

i. Identification of the best screening method

Initially the library screening was done by incubating the biotinylated BIV TAR RNA with streptavidin-labeled alkaline phosphatase in a high salt PBS buffer. This mixture was then added to the peptide library and incubated for 3 h. After extensive washing a mixture of BCIP (5-Bromo-4-Chloro-3'-Indolyphosphate p-Toluidine salt) and NBT (Nitro-Blue Tetrazolium Chloride) was added. It was expected that any of the beads that bound well to the TAR RNA would have alkaline phosphatase attached which would reacts with BCIP/NBT to give a purple precipitate that dyes the bead.[11] However, a negative control reaction was done where no biotinylated TAR RNA was added, and this resulted in all the

[11] Morken, J. P.; Kapoor, T. M.; Feng, S. B.; Shirai, F.; Schreiber, S. L. *J Am Chem Soc* **1998**, *120*, 30-36.

beads turning purple. It was hypothesized that the bTat library binds to one of the other components of the screening method.

To eliminate the need for alkaline phosphatase and the BCIP/NBT, we used an AlexaFluor488-streptavidin conjugate and monitored binding using fluorescence microscopy. The results showed that the beads did fluorescence in the presence of 100 nM each of biotin-bTAR and AlexaFluor488-streptavidin (Figure 3.6). However when the negative control experiment (without biotin-bTAR added) was done, the beads were equally as fluorescent. It was hypothesized that since the streptavidin was the only constant variable between these and the previous experiments, it must be binding to the bTatlibrary. However we were unsure if the streptavidin was binding the bTat peptides or if it was merely adhering to the Tentagel resin. To test this theory, Tentagel beads without any peptide attached were incubated with 100 nM AlexaFluor488-streptavidin. Microscopy showed that the beads were not fluorescent, and so we can conclude that it is the bTatlibrary that binds to streptavidin.

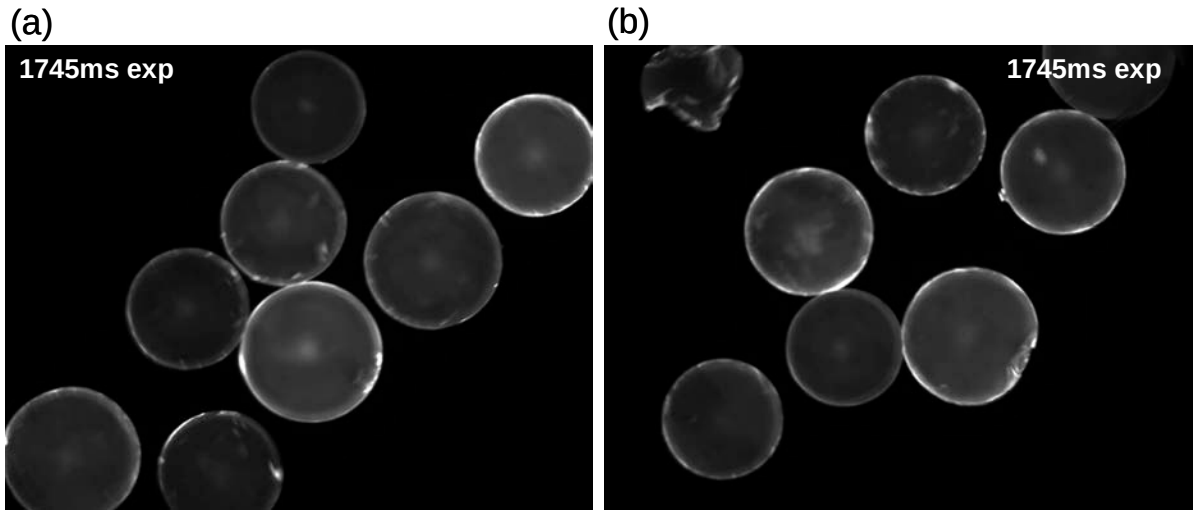


Figure 3.6. Fluorescence microscopy images of bTatlibrary beads with and without bTAR RNA added and incubation with AlexaFluor488-streptavidin conjugate. (a) Reaction with 100 nM biotin-bTAR RNA and 100 nM AlexaFluor488-streptavidin added. (b) Reaction with 100 nM AlexaFluor488-streptavidin added but no RNA.

In order to remove the streptavidin from the screening conditions, we chose to directly label the bTAR RNA with a Cy3 fluorophore and forego a multi-step colorimetric assay. We switched fluorophores because we were able to purchase the RNA with Cy3 already attached, which we were unable to do using AlexaFluor488. The negative control experiment for this screening method was done using Tentagel beads containing peptides (labeled WKWKlib) which do not bind bTAR RNA. The negative control was successful, because none of the resin beads were fluorescent after addition of the Cy3-bTAR (Figure 3.7a). A positive control was also done with Tentagel beads containing the native bTat peptide. This experiment resulted in all fluorescent beads (Figure 3.7b). However, under the same buffer conditions with the bTatlibrary, all of the beads in the library were found to bind Cy3-bTAR (Figure 3.7c). This result suggested that electrostatic interactions dominated and so there was no way to distinguish between peptides with different binding affinities.

To minimize the number of hits detected, the experiments were repeated using more stringent salt conditions in an attempt to eliminate non-specific electrostatic binding. Several rounds of screening were done using varying concentrations of Cy3-bTAR and varying salt conditions until only a few beads were fluorescent (Figure 3.7d). The most effective screening conditions were using 5 nM Cy3-bTAR and 10 mM phosphate buffer with 1 M NaCl and 1 mM $MgCl_2$ at pH 7.4.

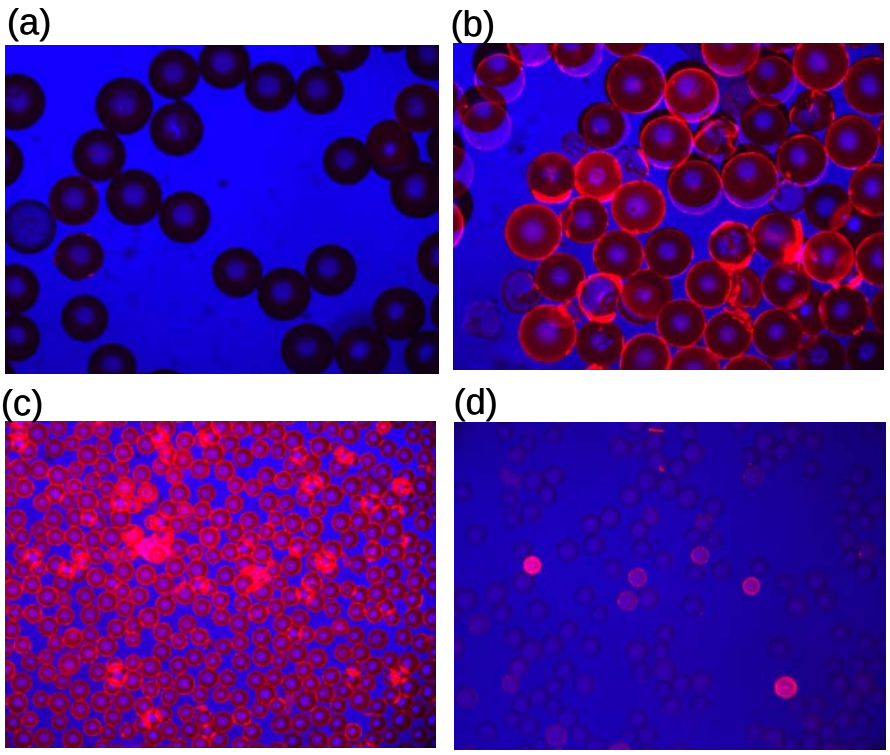


Figure 3.7. Fluorescence microscopy images of affinity screens using Cy3-bTAR RNA. (a) negative control reaction using WKWKlib beads: 10 nM Cy3-bTAR, 10 mM phosphate, 500 mM NaCl, 1 mM MgCl2, pH 7.4 (b) positive control reaction using bTatnative beads: 10 nM Cy3-bTAR, 10 mM phosphate, 500 mM NaCl, 1 mM MgCl2, pH 7.4 (c) bTat library screening: 10 nM Cy3-bTAR, 10 mM phosphate, 500 mM NaCl, 1 mM MgCl2, pH 7.4 (d) bTat library screening: 5 nM bTAR-Cy3 in 10 mM phosphate, 1 M NaCl, 1 mM $MgCl_2$, pH 7.4

ii. Hit identification

The fluorescent beads were picked out individually using a pipette tip and sequenced by single bead Edman degradation. In the first round of sequencing, thirteen of the darkest beads were selected and analyzed. The sequencing results are tabulated in Table 3.1. At position X_1 basic residues were preferred, at X_2 there were both aromatic and non-polar residues selected, and at both positions X_3 and X_4 charged residues were preferred. None of the selected beads contained peptides with non-natural amino acids. All of the peptides that were selected contained either a +1 or +2 charge over the native sequence. These results suggest that even in the high salt (1 M NaCl) screening conditions, electrostatic interactions

were contributing to binding, but in a sequence selective manner. The consensus sequence from the first group of hits was (NH_3-**RGFRK**RGTRGKGRRIRR-GG-resin).

After letting the beads sit in solution for two weeks, some of the beads were still fluorescent, so six beads were chosen for the second round of sequencing (Table 3.1). The second batch of peptide sequences were similar to the original sequences because charged residues were preferred as X_1, X_3, and X_4, and aromatic residues were preferred at X_2. However this round of sequencing did produce several peptides that contained non-natural amino acids. One interesting peptide, NH_3-**YGFhFP**RGTRGKGRRIRR-GG-resin, had one less positive charge than the native sequence and no charged residues at any of the varied positions. Both this sequence, **bTatYGFhFP**, and the consensus sequence, **bTatRGFRK**, were chosen for further investigation.

Table 3.1. Sequencing Results for Hits Selected in Screening the bTatlibrary

$^{+}NH_3$-X_1G$X_2X_3X_4$ RGTRGKGRRIRRGG-●						
	X_1	G	X_2	X_3	X_4	Overall Peptide Charge
Native Seq.	S	G	P	R	P	9
Bead 1	R	G	F	R	R	11
Bead 2	K	G	A	R	K	11
Bead 3	R	G	F	Y	R	10
Bead 4	R	G	A	R	R	11
Bead 5	S	G	W	R	K	10
Bead 6	K	G	F	R	K	11
Bead 7	R	G	P	K	R	11
Bead 8	R	G	P	R	R	11
Bead 9	R	G	A	R	K	11
Bead 10	K	G	F	R	A	10
Bead 11	K	G	W	R	K	11
Bead 12	S	G	P	K	K	10
Bead 13	R	G	A	R	W	10
Hits after 2 weeks						
Bead 1	Y	G	F	hF	P	8
Bead 2	R	G	P	R	R	11
Bead 3	R	G	F	R	R	11
Bead 4	K	G	PhG	R	R	11
Bead 5	R	G	P	R	R	11
Bead 6	R	G	PhG	R	R	11

iii. Binding studies: Fluorescence anisotropy

The full-length **bTatYGFhFP** and **bTatRGFRK** sequences were synthesized along with the native bTat sequence (**bTatnative**) and the binding affinities were characterized using fluorescence anisotropy. To the N-terminus of the peptides was attached 5-(6)-carboxytetramethylrhodamine (TAMRA) with a spacer of two β-alanine residues included in order to avoid the fluorophore interfering with binding. Unlike the library screening, the binding studies were done at physiological salt concentration (10 mM PBS, 137 mM NaCl, 10 mM $MgCl_2$, pH 7.4). The bTAR RNA was unlabeled for the fluorescence experiments.

During the experiment, unlabeled RNA was titrated into a solution of labeled peptide and as binding occurs the mobility of the RNA slows down and changes the amount of parallel and perpendicular polarized light that is emitted. The ratio of emitted parallel and perpendicular polarized light is used to determine the anisotropy value which can then be correlated to binding affinity.[12]

Previously published results have shown that the interaction of the native bTat peptide and bTAR RNA has dissociation constant in the nanomolar range.[13] Our fluorescence anisotropy studies showed that the **bTatnative** peptide binds bTAR with a binding affinity of approximately 110 pM, and the two mutants have very similar binding affinities (Figure 3.8, Table 3.2). The discrepancy in measured binding affinity between our results and previously published values can be attributed to the different methods used to determine binding and slightly different buffer conditions.

For these experiments the ideal concentration of labeled peptide would be $1/10^{th}$ the K_d value. So for a binding affinity of approximately 100 pM, the ideal concentration would be around 10 pM and not 2 nM at which these were done. However the sensitivity of our fluorimeter is too low to measure picomolar concentrations. Therefore the experiments were repeated using a high salt buffer (250 mM NaCl) in an attempt to weaken the binding by screening the electrostatic charges. This still allowed us to compare the binding affinities, and we were able to work at a higher concentration of TAMRA-labeled peptide. The experiments done at a higher salt concentration showed more discrepancy in the binding

[12] Gelman, M. A.; Richter, S.; Cao, H.; Umezawa, N.; Gellman, S. H.; Rana, T. M. *Org Lett* **2003**, *5*, 3563-3565.

[13] (a) Chen, L.; Frankel, A. D. *Biochemistry* **1994**, *33*, 2708-15. (b) Athanassiou, Z.; Dias, R. L.; Moehle, K.; Dobson, N.; Varani, G.; Robinson, J. A. *J Am Chem Soc* **2004**, *126*, 6906-13.

affinity of the three peptides (Figure 3.9). The **bTatnative** bound with a K_d of 1.2 ± 0.4 nM, the **bTatRGFRK** bound three-fold better with a K_d of 0.4 ± 0.2 nM, and **bTatYGFhFP** bound the weakest with a K_d of 4.6 ± 2.3 nM (Table 3.3). The large errors associated with those values suggest that the conditions for the experiment are still not ideal. Therefore, another method of measure binding affinity was used to confirm the results.

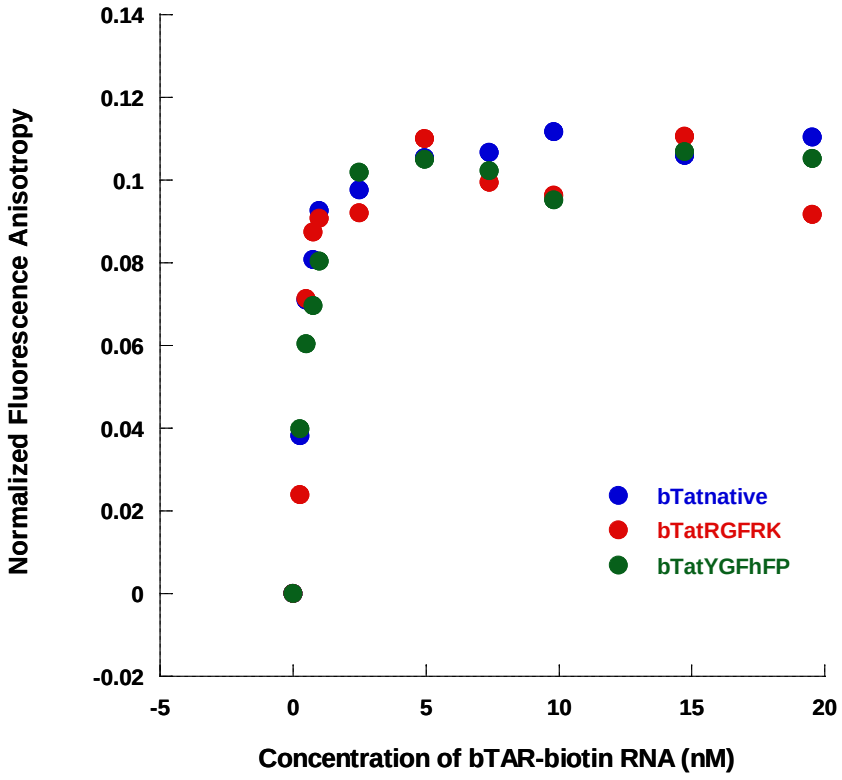


Figure 3.8. Fluorescence anisotropy experiments for measuring the binding affinity of bTAR RNA to the **bTatnative, bTatRGFRK,** and **bTatYGFhFP**. Experiments were done in 10 mM phosphate buffer, 140 mM NaCl, 1 mM $MgCl_2$ pH 7.4. The peptides (2 nM) were labeled with a TAMRA-fluorophore and were excited at 548 nm and emission was collected at 580 nm.

Table 3.2. Dissociation constants determined from the fluorescence anisotropy experiments done at physiological conditions.

Peptide	K_d (nM) (± error)
bTatnative	0.11 (0.02)
bTatRGFRK	0.08 (0.04)
bTatYGFhFP	0.17 (0.03)

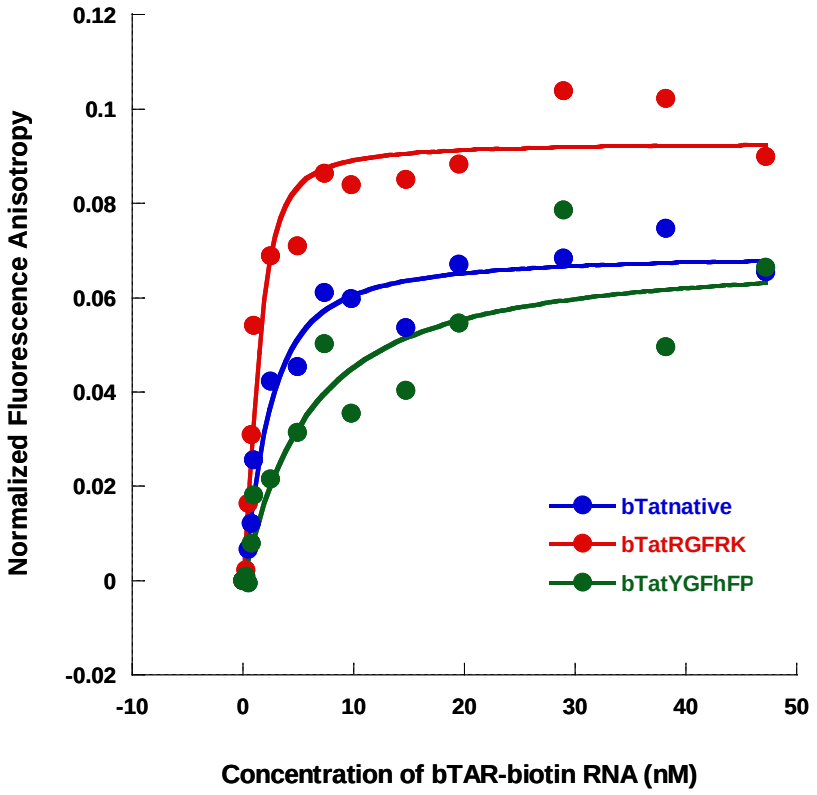


Figure 3.9. Fluorescence anisotropy experiments for measuring the binding affinity of bTAR RNA to the **bTatnative, bTatRGFRK,** and **bTatYGFhFP**. Experiments were done in 10 mM phosphate buffer, 250 mM NaCl, 1 mM $MgCl_2$ pH 7.4. The peptides (5 nM) were labeled with a TAMRA-fluorophore and were excited at 548 nm and emission was collected at 580 nm.

Table 3.3. Dissociation constants determined from the fluorescence anisotropy experiments done with high salt buffer.

Peptide	K_d (nM) (± error)
bTatnative	1.2 (0.4)
bTatRGFRK	0.4 (0.2)
bTatYGFhFP	4.6 (2.3)

iv. Binding studies: Gel shift assays

Gel shift assays were done to supplement the fluorescence anisotropy binding results. Each of the peptides was synthesized without the N-terminal TAMRA label leaving a free amine. This mimics how the original bTatlibrary was synthesized. An additional control peptide was made in which the last five residues of the N-terminus were omitted, $^{+}NH_3$-RGTRGKGRRIRR-NH_2 (**bTatnoloop**). This peptide served as a negative control because it cannot make any contacts in the hairpin loop region.

The procedure for the gel shift assays was a modified version of the one used by Athanassiou et al.[13b] This group also reported a dissociation constant of approximately 50 nM for the native bTat peptide binding to bTAR RNA. Therefore, we were able to directly compare our results with theirs. The binding studies were done using a 50 mM Tris-HCl buffer containing 50 mM KCl, 200 mM DTT, and 0.05% Triton-X at pH 8. The gel shift assays with the positive and negative control peptides were done in duplicate, whereas the binding experiments for the bTat mutants were done in triplicate. We expected the binding affinity to be in the low nanomolar range, so the bTAR RNA concentration was held constant at 0.5 nM. After the gels were phosphorimaged, the band densities were quantified using ImageQuant software. In all cases the lowest and highest concentrations of peptide added

were estimated to be 0% and 100% peptide bound, so the data were normalized to reflect the values as fraction bound. In order to compare the binding affinities, the curves were fit using a 4-parameter logistical model (Equation 2) than analyzed to find the concentration of peptide when 50% of the RNA was bound (K_{eff}).

The positive control peptide (**bTatnative)** and the negative control peptide (**bTatnoloop)** bound bTAR RNA with measurable differences in K_{eff} (Figures 3.10 and 3.11). The **bTatnative** peptide bound with an K_{eff} of 6.3 ± 0.4 nM (Figure 3.12, Table 3.4), which is on the same order of magnitude of those values previously reported.[13a] For the **bTatnoloop** peptide, the K_{eff} was 40.8 ± 3.3 nM (Figure 3.12, Table 3.3). The fact that the binding affinity was about six-fold worse when the N-terminal loop of the peptide is not present confirms that this portion of the peptide contributes to binding with the RNA. However because binding was not completely lost, that portion is not essential for binding to occur.

The two bTat mutant peptides that were identified from the library screens were found to have a similar binding affinity for bTAR as the native peptide (Figures 3.13 and 3.14). The K_{eff} for **bTatRGFRK** peptide was 8.5 ± 0.6 nM (Figure 3.15, Table 3.4), and had a similar shaped binding curve as **bTatnative**. This peptide sequence was derived from the consensus sequence of first round of screening. The sequence has one more positive charge than the native sequence, but the binding affinity is nearly the same. The **bTatYGFhFP** was found to have an K_{eff} of 9.9 ± 0.4 nM (Figure 3.15, Table 3.3). This peptide has one less positive charge than the native sequence, and the binding affinity was only slightly worse.

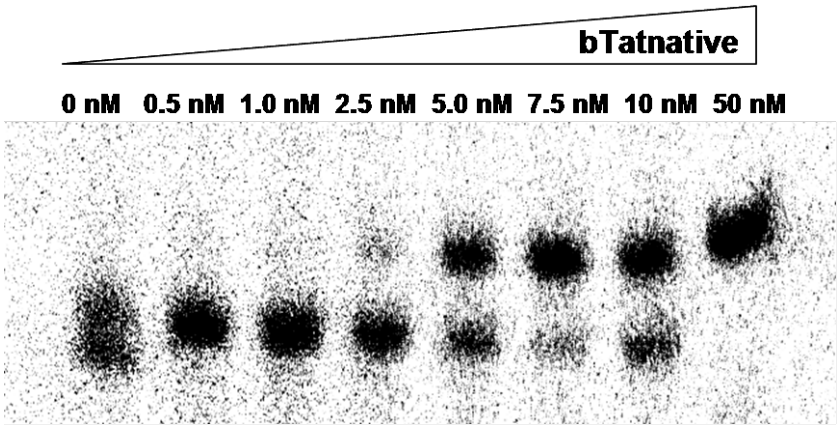


Figure 3.10. Electromobility gel shift assay for **bTatnative** peptide binding to bTAR RNA. The ^{32}pCp-bTAR RNA (0.5 nM) was incubated with increasing concentrations of **bTatnative** peptide in Tris-HCl buffer (50 mM Tris-HCl, 50 mM KCl, 200 mM DTT, 0.05% Triton-X pH 8) for 30 min. The reactions were then run on a native 12% PAGE at 12W and 4°C in 1X TBE buffer.

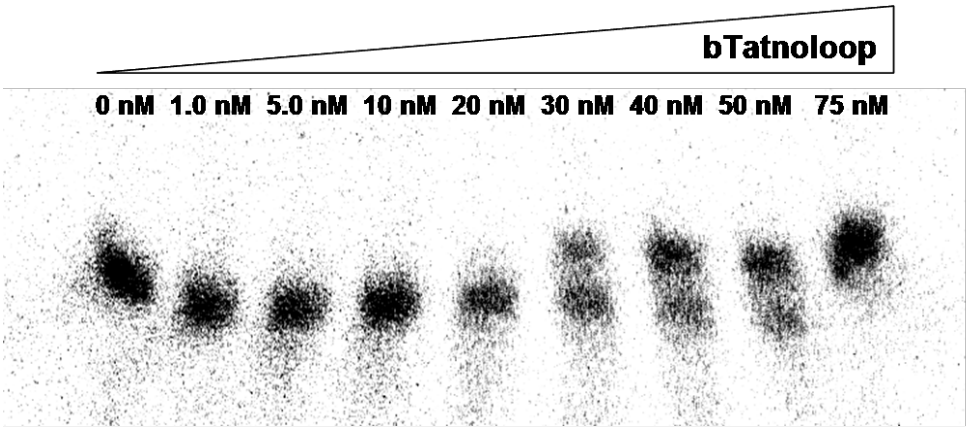


Figure 3.11. Electromobility gel shift assay for **bTatnoloop** peptide binding to bTAR RNA. The ^{32}pCp-bTAR RNA (0.5 nM) was incubated with increasing concentrations of **bTatnoloop** peptide in Tris-HCl buffer (50 mM Tris-HCl, 50 mM KCl, 200 mM DTT, 0.05% Triton-X pH 8) for 30 min. The reactions were then run on a native 12% PAGE at 12W and 4°C in 1X TBE buffer.

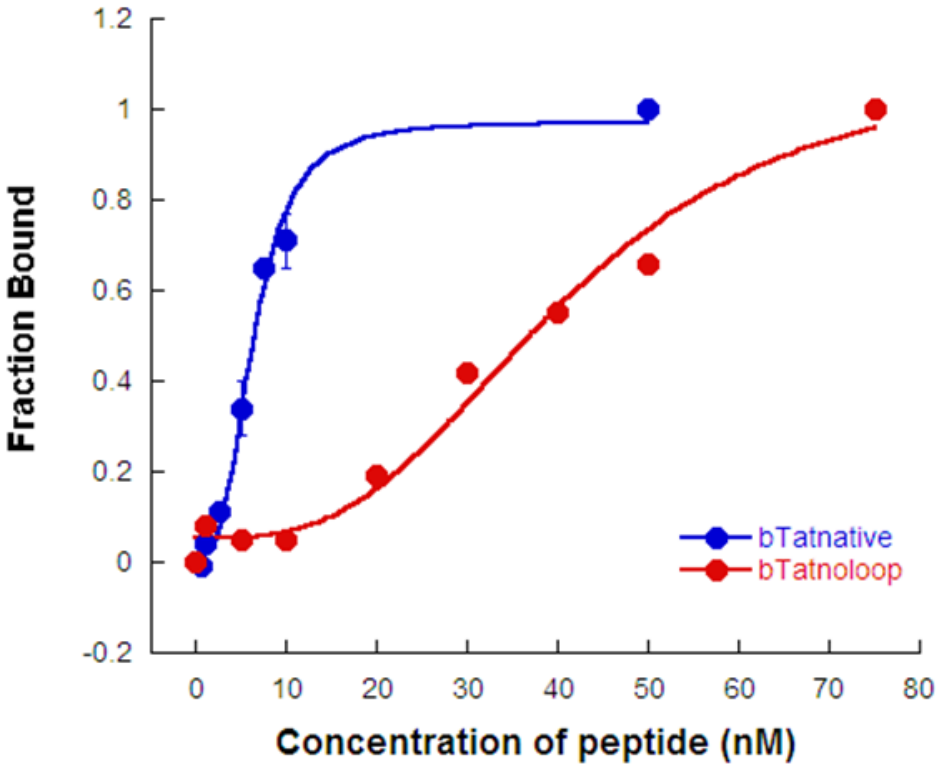


Figure 3.12. Plot of the fraction of bTAR RNA bound by the **bTatnative** (blue) and **bTatnoloop** (red) peptides as determined by electromobility gel shift assays. The data were averaged from two gel shift assays. K_{eff} values were determined by fitting the curve to Equation 2.

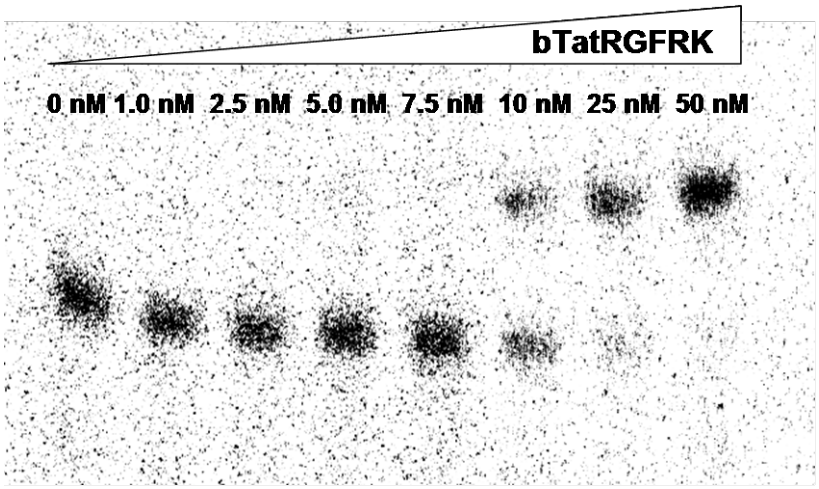


Figure 3.13. Electromobility gel shift assay for **bTatRGFRK** peptide binding to bTAR RNA. The ^{32}pCp-bTAR RNA (0.5 nM) was incubated with increasing concentrations of **bTatRGFRK** peptide in Tris-HCl buffer (50 mM Tris-HCl, 50 mM KCl, 200 mM DTT, 0.05% Triton-X pH 8) for 30 min. The reactions were then run on a native 12% PAGE at 12W and 4°C in 1X TBE buffer.

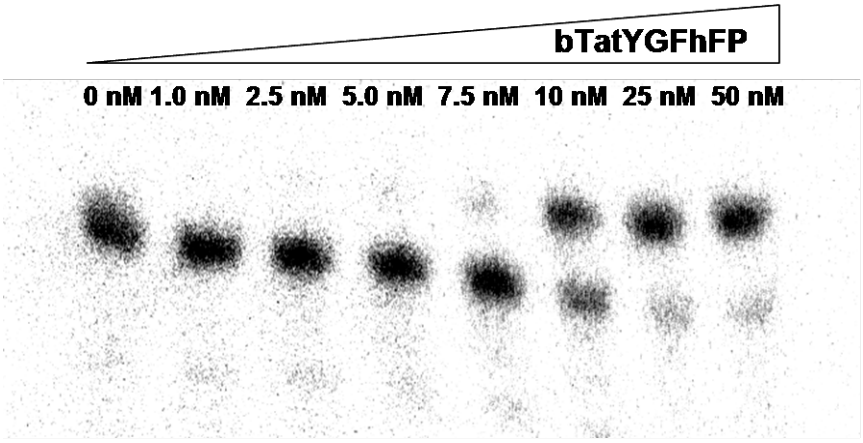


Figure 3.14. Electromobility gel shift assay for **bTatYGFhFP** peptide binding to bTAR RNA. The ^{32}pCp-bTAR RNA (0.5 nM) was incubated with increasing concentrations of **bTatYGFhFP** peptide in Tris-HCl buffer (50 mM Tris-HCl, 50 mM KCl, 200 mM DTT, 0.05% Triton-X pH 8) for 30 min. The reactions were then run on a native 12% PAGE at 12W and 4°C in 1X TBE buffer.

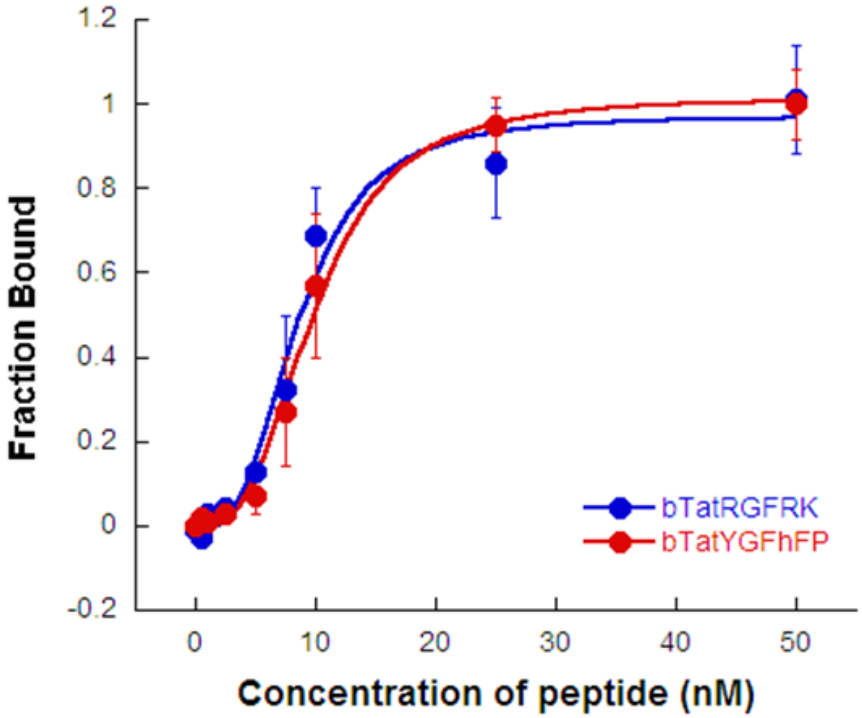


Figure 3.15. Plot of the fraction of bTAR RNA bound by the **bTatRGFRK** (blue) and **bTatYGFhFKP** (red) peptides as determined by electromobility gel shift assays. The data were averaged from two gel shift assays. K_{eff} values were determined by fitting the curves to Equation 2.

Table 3.4. K_{eff} values for bTAR binding to bTatnative and the bTat mutants peptides as determined by electromobility shift assays.

Peptide	K_{eff}, nM (± error)
bTatnative	**6.3 (0.4)**
bTatnoloop	**40.8 (3.3)**
bTatRGFRK	**8.5 (0.6)**
bTatYGFhFP	**9.9 (0.4)**

D. Conclusions

In beginning this project, we set out to improve the bTat/bTAR binding affinity by finding novel substrates with through optimization of hairpin loop binding. We tried to incorporate non-natural amino acids into the peptide substrates which would be able to bind with RNA using interactions other than electrostatics.

After working out the screening methodology, we set out to investigate two of the interesting hits that were chosen, **bTatRGFRK** and **bTatYGFhFP**. The goal was to find a novel bTAR substrate which had a higher binding affinity than the native bTat peptide. More specifically we were searching for peptides which made tighter interactions between the N-terminus of the peptide and the hairpin loop of the RNA. Initial attempts to use fluorescence anisotropy to determine dissociation constants were unsuccessful because of issues with instrument sensitivity at low concentrations of peptide. Gel shift assays were used instead, and the results showed that both of the bTat mutants both had low nM affinity for the RNA, but neither bound tighter than the **bTatnative**. However, all three of those peptides bound better than the negative control, **bTatnoloop**. Despite our efforts, we did not discover a novel substrate that bound better than the native peptide. We were able to find a peptide, **bTatYGFhFP**, which had one less positive charge which did not seem to cause a loss in binding affinity.

The other interesting aspect of the **bTatYGFhFP** peptide was that it contained a non-natural residue, homophenylalanine. During the screening process, high salt buffers were used in order to reduce the amount of electrostatic interactions, and promote peptides that bound with different types of interactions (aromatic, hydrophobic, cation-π, etc). We also included several non-natural amino acids that could participate in these types of interactions.

The initial screen did not yield peptides containing unnatural amino acids, but after the beads had incubated in buffer for two weeks several hits did contain interesting residues. We can conclude from this result that there may have only been a few hits that contained non-natural amino acids, but the peptides that did had undiminished binding affinity over time.

In summary, we have screened an on-bead bTat peptide library for increased binding to bTAR RNA. Despite the fact that we discovered novel substrates for this RNA sequence, there was not a drastic improvement in binding affinity. Future directions for the project would be to investigate whether there were increase contacts of the bTat mutants to the RNA hairpin loop region using RNase footprinting or NMR studies. Additionally the methodology could be used to find novel peptides which bind to other interesting RNA sequences.

E. Experimental

i. Peptide library synthesis

The synthesis of the bTat library and bTat peptide was done manually on non-cleavable Tentagel resin beads (0.21 mmol/g loading) on a 0.16 mmol and 0.1 mmol scales, respectively. The resin was swelled using DMF, and each of the FMOC-protected amino acids (4 eq.) were coupled onto the peptide using an activator solution of HOBT (4 eq.), HBTU (4 eq.), and DIPEA (4 eq.). The resin was agitated for 1.5 h, and then rinsed thoroughly. FMOC deprotection was done with 20% piperidine in DMF (2 x 15min). Coupling reactions and Fmoc deprotections were monitored using the Kaiser test.[14] The side chains were deprotected for 3 hours using a cleavage cocktail of TFA:TIPS:Water (9.5:0.25:0.25), and then subsequently rinsed with TFA and methanol.

[14] Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. *Anal Biochem* **1970**, *34*, 595-8.

ii. Library screening

The 3' labeled Cy3-bTAR RNA was purchased as a purified solid from IDT DNA Technologies. The RNA (5 nM or 10 nM) was redissolved in RNase-free PBS (10 mM phosphate, 140 mM NaCl, pH 7.4) and annealed by heating to 95°C for 2 min and slow cooling. The RNA stock was aliquoted into smaller samples and stored at -20°C. For the library screening reactions, the resin was swelled with DCM, and then dried using an aspirator. Approximately 50-100mg of resin was weighed into a small, fritted reaction vessel. The resin was rinsed with MeOH, 50% MeOH:DEPC-treated water, DEPC-treated water, high salt buffer (10 mM phosphate, 140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4) and high salt screening buffer (10 mM phosphate, 500 or 1000 mM NaCl, 1 mM $MgCl_2$, 0.01% Triton-X, pH 7.4). The resin was incubated in high salt screening buffer for 30 min, then drained and a solution of high salt screening buffer and RNase inhibitor (1 μL) was added. To that solution the folded Cy3-bTAR RNA was added and the resin is shaken for 2 h. The resin is drained and rinsed twice with high salt screening buffer, twice with high salt PBS buffer, and then twice with high salt screening buffer. The beads were plated on cell culture plates and studied using a Leica DMIRB Inverted Fluorescence/DIC microscope. Highly fluorescent beads were picked out using a pipette and sequenced using Edman degradation.

iii. Fluorescence Anisotropy Binding Experiments

The non-labeled bTAR RNA sequences were purchased as a purified, lyophilized powder from IDT DNA. The RNA was dissolved in RNase-free water and annealed by heating to 95°C for 2 min, then slowly cooling to room temperature. The folded RNA solution was separated into smaller aliquots and stored at -20°C. The bTatnative peptide and bTat mutant peptides were synthesized on PAL-PEG-PS resin (0.07 mmol) using an Applied Biosystems

Pioneer synthesizer. The peptides were fluorescently labeled by coupling overnight 5-(and -6)-carboxytetramethylrhodamine (0.16 mmol) with the free N-terminus of the peptide using HOBT, HBTU, and DIPEA (0.4 mmol of each) in DMF. Side-chain deprotection and cleavage from the resin using a TFA cocktail as previously described. The peptides were purified using a Waters reverse-phase HPLC system to give a bright pink powder. The masses of the peptides were confirmed using ESI-TOF mass spectrometry (Table 3.5).

The probes were dissolved in RNase-free PBS buffer (10 mM phosphate, 140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4), and the concentration of the solution was determined using UV-Vis (λ_{559}=91000 M^{-1} cm^{-1}). The binding experiments were done by titration of a concentrated stock of RNA into a 2 nM solution of TAMRA labeled-peptide (excited at 548 nm and emission monitored at 580 nm). The anisotropy experiments were preformed at 25°C in a 1 cm pathlength quartz cuvette that had been rinsed with RNaseZap (Ambion), RNase-free water, and methanol. The anisotropy data was plotted as a function of anisotropy versus concentration of bTAR RNA added. The plot was fitted using equation 1 derived from a previous report: [15]

$$r = \left(\left(\frac{(a + x + K_D) \pm \sqrt{(-a - x - K_D)^2 - 4(a \cdot x)}}{2 \cdot a} \right) \cdot (r_\infty - r_0) \right) + r_0 \quad \textit{Equation 1}$$

where r is anisotropy, r_0 is initial anisotropy value, r_∞ is the maximum anisotropy, a is the total TAMRA-labeled peptide-intercalator concentration, x is the total concentration of RNA added, K_d is the dissociation constant.

[15] Wang, Y.; Killian, J.; Hamasaki, K.; Rando, R. R. *Biochemistry* **1996**, *35*, 12338-46.

iv. Gel shift assays

3'-^{32}P RNA labeling. Approximately 10 μg of bTAR RNA construct was treated with T4 RNA ligase (2 U/μL), T4 RNA ligase buffer (1X), DMSO (10%), RNase inhibitor (1 U/μL), and ^{32}pCp (0.05 mCi) in a total volume of 30 μL and incubated at 4°C overnight. The labeled RNA solution was run through an RNA spin column and loaded onto a 20% denaturing polyacrylamide gel for purification. The labeled RNA was visualized by autoradiography, excised from the gel and eluted overnight with elution buffer (0.5 M NH_4OAc, 1 mM EDTA, 0.1 % SDS, pH 8.0). After filtration the RNA was precipitated with 95% ethanol, sodium acetate buffer, and linear acrylamide overnight at -20°C. The RNA was brought up in DEPC-treated water and the concentration was determined by UV absorbance using the equation: concentration of RNA SAMPLE = 40 × OD_{260} × dilution factor.

Gel Shift Assays. These are modified versions of the procedures used by Athanassiou et al.[13b] The labeled RNA was brought up in DEPC-treated water and annealed at 95°C for 2 minutes and slowly cooled to room temperature. The bTat peptides were synthesized using the same procedure as above except without the fluorophore label. They all contained a free N-terminus so as to mimic the total charge on the bTat library. Stock solutions were made by weighing out unlabeled peptides and dissolving them in DEPC-treated water to give a 1 mM final concentration. Increasing amounts of peptide solution was added to 0.5 nM bTAR RNA and incubated at room temperature in 1X Tris-HCl buffer (50 mM Tris-HCl, 50 mM KCl, 0.05% Triton X, 200 mM DTT, pH 8.0) for 30 min. To each reaction was added 1 uL BPB dye. After the addition of the dye, the total reaction volume was 11 uL, and 8 uL of the sample was loaded onto a 12% native polyacrylamide gel. The gel was electrophoresed in

1X TBE buffer at 12W and 4°C. The gels were exposed to the phosphor imaging screen overnight and scanned using a Storm phosphor imager. The band intensities were analyzed using ImageQuant software, and fraction bound was determined by dividing the number of radioactive counts in the higher band by the total counts in the lane. The fraction bound values were then normalized by assuming the initial point was 0% bound and the final point was fully bound. The binding constant, K_{eff}, was determined using a 4-parameter logistical model which fit the data according to Equation 2:[16]

$$F = F_{\max} + \frac{\left(F_o - F_{\max}\right)}{\left(1 + \left(\frac{x}{K_{eff}}\right)^2\right)} \qquad \textit{Equation 2}$$

where F is the fraction of RNA bound, F_{max} is the fraction bound at infinite peptide concentration, F_0 is the fraction bound at zero peptide concentration, x is the peptide concentration, K_{eff} is the peptide concentration when 50% of the RNA is bound.

[16] Findlay, J. W.; Dillard, R. F. *AAPS J* **2007**, *9*, E260-7.

Table 3.5. Mass Spectrometry Data Obtained for Peptides Studied in this Project

Peptide	Expected Mass (Da)	Observed Mass (Da)
bTatnative	1961.2	1961.2
bTatnative(TAMRA)	2515.5	2515.6
bTatnoloop	1467.0	1467.0
bTatnoloop(TAMRA)	1879.1	1879.0
bTatRGFRK	2111.4	2111.3
bTatRGFRK(TAMRA)	2665.6	2665.6
bTatYGFhFP	2092.3	2092.2
bTatYGFhFP(TAMRA)	2646.5	2647.0

Chapter IV

THE USE OF SIDE-CHAIN INTERACTIONS TO STABILIZE β-HAIRPIN PEPTIDES FOR BINDING TO BIV TAR RNA

A. Background

In a third study, we used strong side-chain interactions to stabilize a mutant BIV Tat (bTat) peptide, BIV2, which was previously designed to bind BIV TAR RNA (bTAR). Using our knowledge of β-hairpin stability, we were able to convert the BIV2 peptide, which required cyclization to maintain the structure, into an acyclic peptide with stable β-hairpin structure. We investigated whether our newly designed peptides could bind RNA with the same binding affinity as the original. The binding affinity of the linear peptides was compared to the cyclic peptide using fluorescence anisotropy and electromobility gel shift assays.

This project was based on work by Athanassiou et al. in which they cyclized the native BIV Tat peptide with a $_{D}$Pro-$_{L}$Pro turn template. [1] The native BIV Tat peptide is unstructured in solution, but upon binding BIV TAR RNA it forms a β-hairpin structure and lies in the

[1] (a) Athanassiou, Z.; Dias, R. L.; Moehle, K.; Dobson, N.; Varani, G.; Robinson, J. A. *J Am Chem Soc* **2004**, *126*, 6906-13. (b) Leeper, T. C.; Athanassiou, Z.; Dias, R. L.; Robinson, J. A.; Varani, G. *Biochemistry* **2005**, *44*, 12362-72. (c) Athanassiou, Z.; Patora, K.; Dias, R. L.; Moehle, K.; Robinson, J. A.; Varani, G. *Biochemistry* **2007**, *46*, 741-51.

major groove.[2] The group made mutations to optimize the cyclized bTat peptide until a new sequence, BIV2, was found which had nanomolar binding affinity to bTAR RNA (Figure 4.1). BIV2 adopts a stable β-hairpin structure and upon binding RNA one face of the peptide makes interactions with the major groove of bTAR while the other face is solvent exposed (Figure 4.2a).[1] The BIV2 peptide mimics the native bTat peptide in the way that it binds RNA, but the designed peptide has a pre-formed structure and eliminates the entropic cost of folding (Figure 4.2b).

One advantage to using side-chain interactions to provide structure to a peptide system instead of cyclization is that it simplifies the synthesis of the peptide. There is no need for solution phase amide coupling which requires additional purification. Another reason for doing this project was to attempt to design a peptide which maintained both its β-sheet secondary structure and displayed an RNA binding interface. For this purpose we designed a β-hairpin peptide where one face of the peptide stabilized the peptide through side-chain interactions, while the other face displayed residues that mimic a desired binding interaction. Using side-chain interactions for stabilization may be better than using cyclization which could potentially perturb the structure or interfere with binding.

In this study we chose to insert a Trp-Trp-Lys-Leu (WWKL) sequence into the non-hydrogen bonded positions of BIV2 peptide. This peptide sequence, known as Trp pocket, stabilizes β-hairpin structure because the two tryptophans form an aromatic cleft where the cationic lysine side chain interacts through cation-π interactions.[3] The leucine also stabilizes the structure through hydrophobic interactions to the N-terminal tryptophan. A small β-

---

[2] (a) Puglisi, J. D.; Chen, L.; Blanchard, S.; Frankel, A. D. *Science* **1995**, *270*, 1200-3. (b) Ye, X.; Kumar, R. A.; Patel, D. J. *Chem Biol* **1995**, *2*, 827-40.

[3] Riemen, A. J.; Waters, M. L. *Biochemistry* **2009**, *48*, 1525-31.

hairpin containing this sequence has been shown to be the most thermally stable designed hairpin reported to date.

For BIV2, it was determined that the important residues for RNA binding were Arg1, Arg3, Arg5, Arg8, and Ile10, which are all located in the hydrogen bonded positions (as determined by the NMR structure, PDB 2A9X).[1] Therefore, we were able to replace residues Val2, Thr4, Arg9, and Arg11 of the non-hydrogen bonded positions with the Trp pocket residues. Using the NMR structure as a guide, we chose to cut BIV2 between residues Gly6 and Lys7 in order to make it a linear peptide. These two residues do not seem to make any significant contribution to binding, and seem to be solvent exposed from the NMR structure. We also replaced the $_{D}$Pro-$_{L}$Pro turn template with either an Asn-Gly or DPro-Gly turn sequence. Not only do these turns promote β-hairpin formation,[4] they are more flexible than the $_{D}$Pro-$_{L}$Pro turn, and will allow the peptide to conform more easily to the RNA.

After synthesizing and characterizing the structure of the newly designed linear peptides, we also investigated whether they still bind to BIV TAR RNA. Initially fluorescence anisotropy was used to characterize the binding, but due to fluorophore quenching issues we switched to electromobility gel shift assays. Although the designed peptides do maintain β-hairpin structure without cyclization, they did not maintain the binding affinity for the RNA target

[4] Stanger, H. E.; Gellman, S. H. *J Am Chem Soc* **1998**, *120*, 4236-4237.

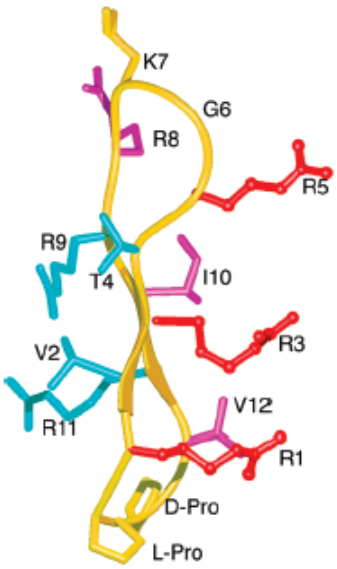


Figure 4.1. Structure of BIV2 as determined by Robinson et al.[1] The residues that interact with bTAR RNA are shown in red and pink and the solvent exposed face of the peptide is colored blue.

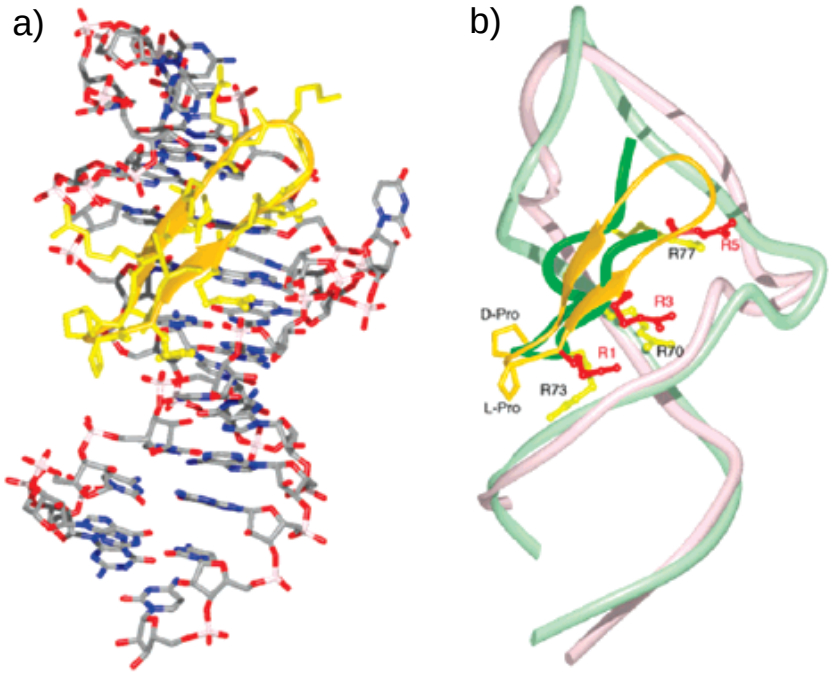
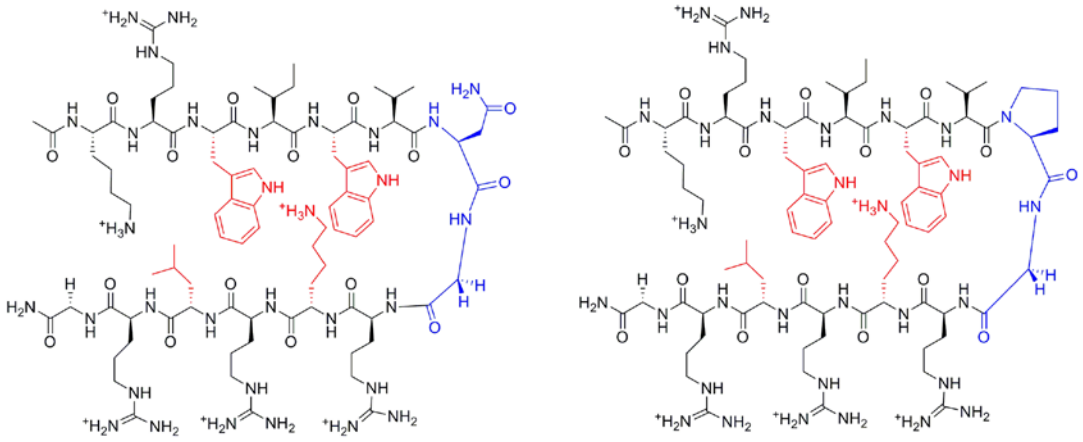


Figure 4.2. Structure of the BIV2-bTAR complex as determined by Robinson et al.[1] a) BIV2 (yellow) lies in the major groove of BIV TAR RNA. The $_{D}$Pro-$_{L}$Pro turn template orients itself in the RNA stem, away from the hairpin loop and bulge regions. b) Superimposed images of BIV2 (yellow, PDB 2A9X) and BIV Tat peptide (green, PDB 1MNB) binding in a similar orientation to BIV TAR RNA. The key arginine residues that are consistent between the two peptides are labeled and highlight in yellow and red.

B. Design and synthesis

Two peptides, **bTatWWKL-NGturn** and **bTatWWKL-pGturn**, were designed based on the BIV2 sequence, but differ from that peptide because they are linear and not cyclized (Figure 4.3). Instead the Trp pocket domain has been incorporated on the face of the β-hairpin which does not interact with bTAR RNA. We have also added either an Asn-Gly (NG) or $_{D}$-Pro-Gly (pG) turn sequence to help promote β-hairpin formation. These turn sequences represent two different turn types, and we tested both to see which would be a better conformation for binding the RNA. We expected that all of these changes would stabilize the β-hairpin structure without the need for a cyclic peptide.

The peptides were synthesized using standard Fmoc-based solid phase peptide synthesis methods. The extent of folding was determined using circular dichroism and NMR. The peptides were also synthesized with a 5-(6)-carboxytetramethylrhodamine (TAMRA) fluorophore on the N-terminus, which allowed for fluorescence anisotropy to be used to measure binding of the peptide to the RNA.



bTatWWKL-NGturn bTatWWKL-pGturn

Figure 4.3. Sequence and structures of **bTatWWKL-NGturn** and **bTatWWKL-pG** turn. The WWKL face of the peptide is shown in red, and the turn sequence is shown in blue.

C. Results and discussion

i. Secondary structure determination: Circular dichroism

Circular dichroism (CD) was used to determine the global secondary structure of the WWKL peptides in solution. The CD spectra for **bTatWWKL-NGturn** and **bTatWWKL-pGturn** were compared to the CD spectrum for a linear version of the BIV2 peptide, Ac-KRRIRVpPRVRTRG-NH_3^+ (Figure 4.4). Both of the bTatWWKL peptides showed a minimum CD signal around 215 nm which is indicative of β-sheet structure. The CD spectrum for the linear version of BIV2 peptide was very different, because it showed a minimum around 200 nm which is indicative of a random coil conformation. This provides evidence that the incorporation of the WWKL pocket and the new turn sequences was enough to promote β-sheet sheet structure without the need for cyclization.

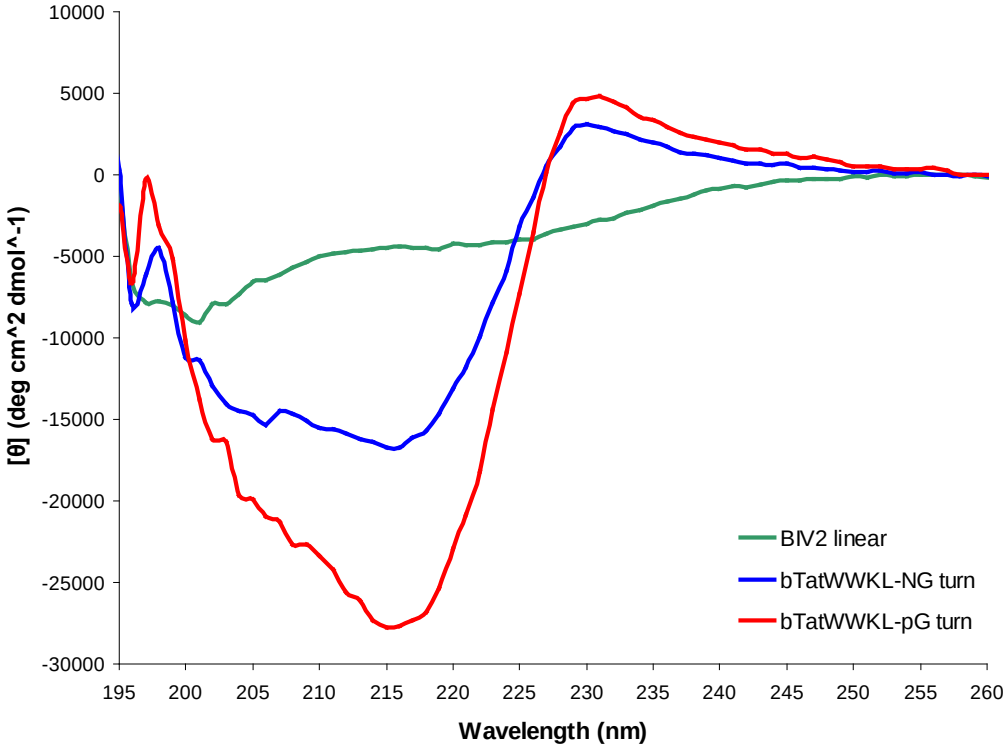


Figure 4.4. Circular Dichroism Spectra for linear BIV2 (green), **bTatWWKL-NGturn** (blue), and **bTatWWKL-pGturn** (red). Studies were done at 100 μM peptide concentration in 10 mM sodium phosphate buffer pH 7.5 at 25°C. Both bTatWWKL peptides fold into a β-sheet conformation with a minimum mean residue ellipticity around 215 nm. Linear BIV2 does not fold into a β-sheet conformation, and instead resembles a random coil conformation with a minimum around 200 nm.

ii. Secondary structure determination: One dimensional NMR

Working in collaboration with Dr. Zach Laughrey, we confirmed the β-hairpin structure of the bTatWWKL peptides using one-dimensional NMR. For well-folded β-hairpin peptides the alpha protons, Hα, are significantly downfield shifted as compared to an unfolded control.[5] The fraction folded can be determined using the equation 1, where the unstructured controls are truncated peptides containing either the N-terminal residues including the turn or the C-terminal residues including the turn. The fully folded control was

[5] Wishart, D. S.; Sykes, B. D.; Richards, F. M. *Biochemistry* **1992**, *31*, 1647-51.

a cyclized version of the bTatWWKL peptides which had cysteine residues on both the N- and C-terminus which formed a terminal disulfide bond. The results confirmed that both peptides are highly folded with the exception of the ends which are commonly frayed (Figure 4.5). Overall the **bTatWWKL-pGturn** peptide was determined to be 88% folded, which was slightly more than the **bTatWWKL-NGturn** peptide which is 82% folded.

$$Fraction\,Folded = \frac{(H_\alpha(hairpin) - H_\alpha(unstructured\ control))}{(H_\alpha(\ fully\ \ folded\ control) - H_\alpha(unstructured\ control))} \qquad (Equation\ 1)$$

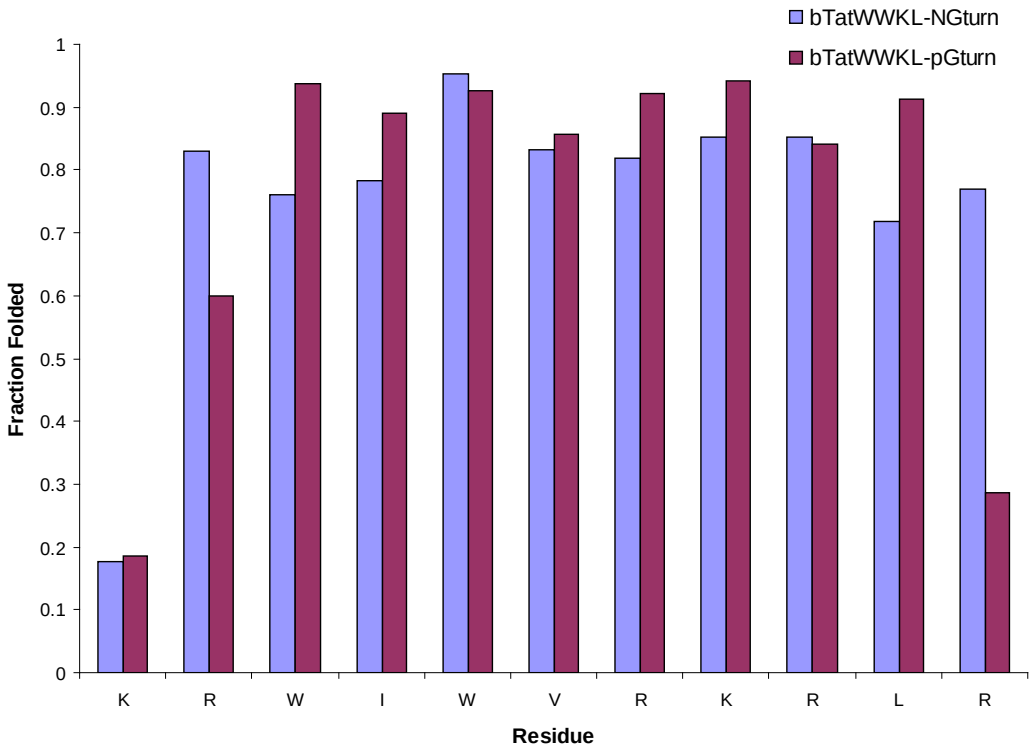


Figure 4.5. Fraction folded derived from the Hα chemical shifts. Both the **bTatWWKL-NGturn** (blue) and the **bTatWWKL-pGturn** (purple) peptides are found to be well-folded β-hairpins. Data was collected and analyzed by Dr. Zach Laughrey.

iii. Fluorescence quenching studies

We had intended to use fluorescence anisotropy to measure the binding of the bTatWWKL peptides to bTAR RNA. However during the initial studies it was observed that the fluorescence intensity of the samples was changing over the course of the binding experiments. Therefore, we set up several experiments to investigate when and why TAMRA fluorescence was being quenched.

First, a solution of 0.1 μM TAMRA-labeled **bTatWWKL-NGturn** peptide was made without any RNA added. The fluorescence intensity of the sample was measured immediately and several more times over the course of 45 min (Figure 4.6). Within 10 minutes the fluorescence intensity of the sample had decreased by half, and continued to decrease until leveling off around 30 min.

To determine if the decrease in signal was a result of photobleaching, three identical samples of **bTatWWKL-NGturn** were made simultaneously. A fluorescence scan of one of the samples was taken immediately, and the other two samples were stored overnight at room temperature. One of them was wrapped in foiled and stored in the dark, and the other was stored on the bench top exposed to light. The following day the fluorescence intensities for both samples were tested using the same fluorimeter settings as the prior day. Additionally a fourth sample, identical to the others, was made and scanned immediately. The results show that no matter the storage conditions, the fluorescence of **bTatWWKL-NGturn** decreases significantly over time (Figure 4.7). When the fourth sample was made and scanned immediately it matched the intensity of the original sample. This proves that the quenching is occurring within the peptide itself and not because of any outside triggers. We hypothesize

that the conformation of the peptide may be changing in solution and during that process the environment around the fluorophore is also changing.

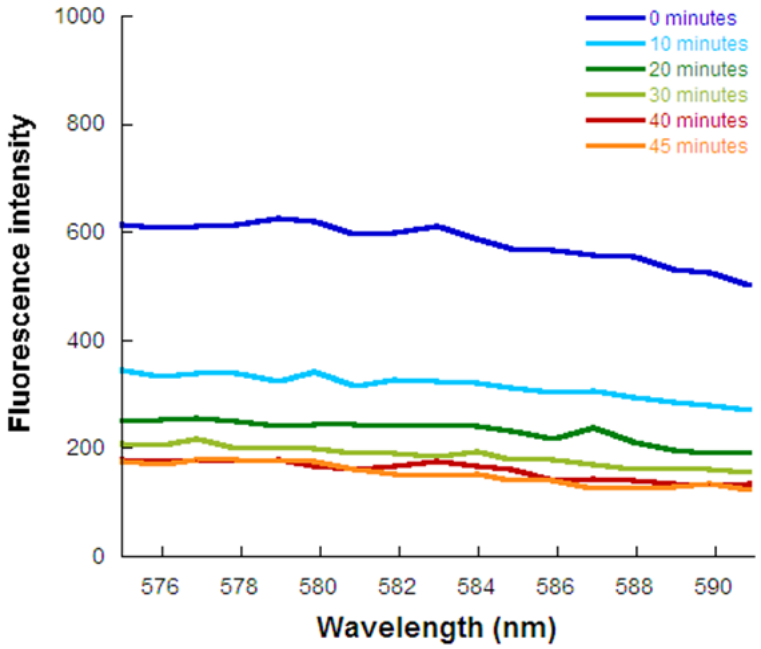


Figure 4.6. Fluorescence intensity of **bTatWWKL-NGturn** peptide measured over time. A 0.1 μM solution of **bTatWWKL-NG** in 10 mM phosphate buffer, 140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$ (pH 7.5) was used. For each scan the excitation and emission slit widths were 10 nm each, 900v light was used, and the instrument was set at 15°C.

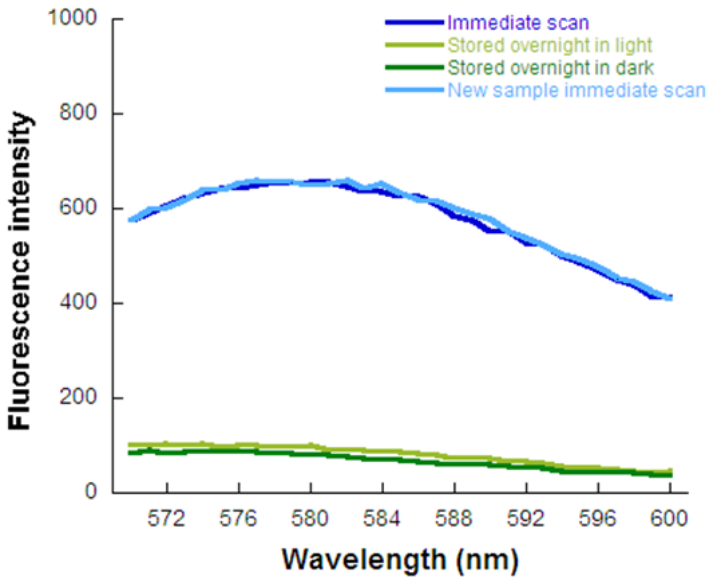


Figure 4.7. Fluorescence intensity of four separate samples of **bTatWWKL-NGturn** peptide measured at different time points. Each sample contains 0.1 μM solution of **bTatWWKL-NG** in 10 mM phosphate buffer, 140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$ (pH 7.5). For each scan the excitation and emission slit widths were 10 nm each, 875v light was used, and the instrument was set at 15°C.

iv. Fluorescence anisotropy studies

We also investigated whether the anisotropy values of TAMRA-labeled **bTatWWKL-NGturn** were affected over time. To do this experiment, a 0.1 μM solution of **bTatWWKL-NGturn** in buffer was made with no RNA added. The fluorescence anisotropy values were measured over time, and no change in the values was seen (Figure 4.8). From this we can assume then that during a binding experiment any change in the fluorescence anisotropy value will be due to binding RNA, and not due to fluorescence quenching.

From our earlier experiments we had determined that the decrease in fluorophore intensity stabilized after 30 min. Even though the fluorescence anisotropy does not seem to be affected over time, the samples were allowed to incubate for 30 min before data was collected. This also allowed time for the system to reach equilibrium. The results showed

that both peptides bind bTAR in the low micromolar range, although the **bTatWWKL-pGturn** peptide binds slightly better (Figure 4.9, Table 4.1). The dissociation constant determined for **bTatWWKL-NGturn** was 4.6 ± 1.2 μM, whereas it was 2.1 ± 0.2 μM for **bTatWWKL-pGturn**. These dissociation constants are about 10-fold higher than what Robinson determined for cyclic BIV2 binding to bTAR (K_d = 0.15 μM).[1]

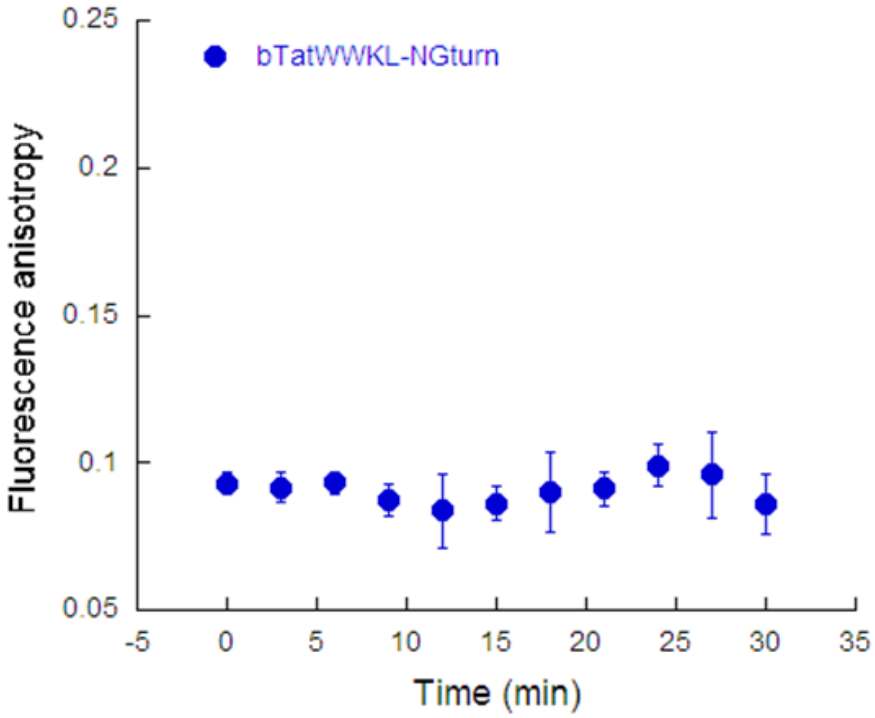


Figure 4.8. Fluorescence anisotropy over time of **bTatWWKL-NGturn** peptide with no RNA added. A solution of 0.1 μM peptide in 10 mM phosphate buffer, 140 mM Na^+/K^+Cl^-, and 1 mM $MgCl_2$ pH 7.5 at 15°C.

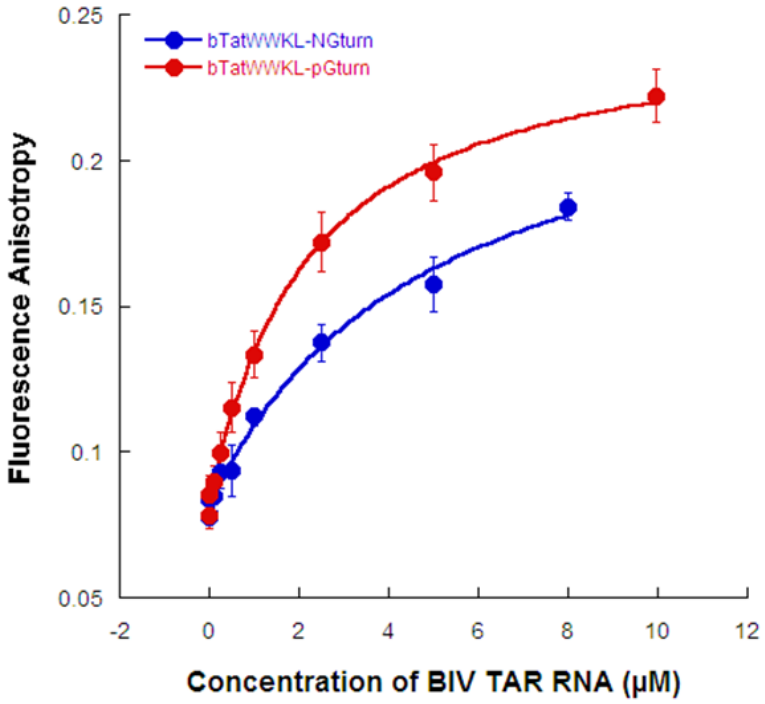


Figure 4.9. Fluorescence anisotropy experiments for the bTatWWKL peptides binding to bTAR RNA. A 0.2 μM solution of either bTatWWKL-NGturn (blue) or bTatWWKL-pGturn (red) peptides in 10 mM phosphate buffer, 140 mM Na^+/K^+Cl^-, and 1 mM $MgCl_2$ pH 7.5 was incubated with increasing concentrations of bTAR RNA for 30 min.

Table 4.1. The dissociation constants for bTatWWKL peptides binding to bTAR RNA as determined by fluorescence anisotropy.

	K_D (μM) (error)
bTatWWKL-NGturn	4.6 (1.2)
bTatWWKL-pGturn	2.1 (0.2)

v. Electromobility gel shift assays

We decided to use electromobility gel shift assays (EMSAs) to measure the binding, as this was the method that Robinson's group used to measure BIV2 binding, so we could directly compare our results to theirs. Another advantage to using EMSAs is that we no longer needed to label the peptides with TAMRA, and we avoided the fluorescence quenching problems.

For the EMSAs, we used the same buffer conditions that Robinson et al. used, which was 50 mM Tris-HCl (pH 8.0), 50 mM KCl, 200 mM DTT, and 0.05% Triton X-100. The only difference was that we did not include yeast tRNA in our studies. We also followed their gel conditions, which were using a 12% native polyacrylamide in 0.5X TBE buffer electrophoresed at 15W in 4°C. Running the gel at a low temperature helps ensure that the complex remains bound on the gel. The bTAR RNA was labeled on the 3' end using ^{32}pCp, and the peptide was synthesized without TAMRA on the N-terminus.

Based on the fluorescence anisotropy results, **bTatWWKL-pGturn** bound the tightest to bTAR RNA, so this peptide used for further studies. It was expected that this peptide would bind the RNA with a K_d of approximately 2 μM, so for the EMSAs we used peptide concentrations ranging from 0.025 μM to 50 μM (Figure 4.10). However, unlike the anisotropy data, no binding was detected by the gel shift assay.

To determine whether there was a problem with the assay conditions, we ran a positive control experiment using the native bTat peptide binding to bTAR RNA (Figure 4.11). This K_d for this interaction is known to be approximately 10 nM.[6] Our results showed that after only 0.025 μM peptide was added, the band corresponding to unbound bTAR was gone and a

[6] a) Chen, L.; Frankel, A. D. *Biochemistry* **1994**, *33*, 2708-15. (b) Athanassiou, Z.; Dias, R. L.; Moehle, K.; Dobson, N.; Varani, G.; Robinson, J. A. *J Am Chem Soc* **2004**, *126*, 6906-13.

new band corresponding to the peptide-RNA complex appeared which ran slower on the gel. A third band appears after more than 1 μM peptide has been added, and it is suspected that this corresponds to large complexes being formed by nonspecific associations. The binding affinity agrees well with the published binding data for the bTat-bTAR interaction, so we can conclude that the assay conditions were fine.

To mimic more exactly the fluorescence anisotropy experiments, we decided to use the TAMRA-labeled **bTatWWKL-pGturn** peptide for the EMSA. As was done with the previous experiments, increasing concentrations of the peptide were incubated with bTAR RNA for 30 min prior to electrophoresis (Figure 4.12). The results show that the band corresponding to unbound bTAR began to disappear after 1 μM peptide was added and is replaced by a higher band near the top of the gel. The data were analyzed using ImageQuant software and the curve was fit using Equation 2 (Figure 4.13). The binding constant determined for the **TAMRA-bTatWWKL-pGturn** and BIV TAR RNA interaction was 0.6 ± 0.2 μM, which is very similar to what was determined by fluorescence anisotropy. We can conclude from these data that the TAMRA fluorophore is responsible for the binding of this peptide to bTAR RNA. When the peptide did not contain the label, there was no measurable binding detected.

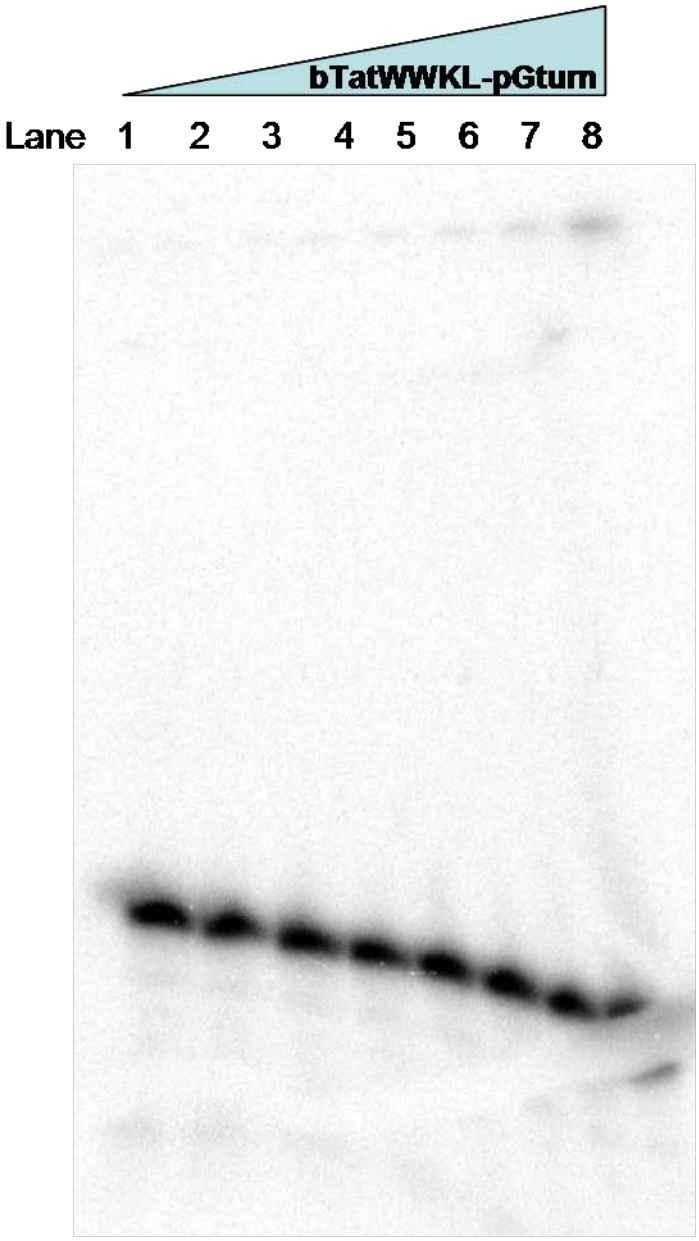


Figure 4.10. Electromobility gel shift assay of **bTatWWKL-pGturn** bound to bTAR RNA (10 nM). Lane 1: RNA only, Lane 2: 0.025 μM peptide, Lane 3: 0.1 μM peptide, Lane 4: 0.5 μM peptide, Lane 5: 1 μM peptide, Lane 6: 5 μM peptide, Lane 7: 10 μM peptide, Lane 8: 50 μM peptide. Buffer conditions were 50 mM Tris-HCl (pH 8.0), 50 mM KCl, 200 mM DTT, and 0.05% Triton X-100. Gel was a 12% native polyacrylamide in 0.5X TBE buffer electrophoresed at 15W in 4°C. The peptide and RNA were incubated at room temperature for 30 min before loading on the gel. The gel was dried for 1 h and phosphor imaged overnight. The distorted band in lane 8 was caused by a tear in the gel.

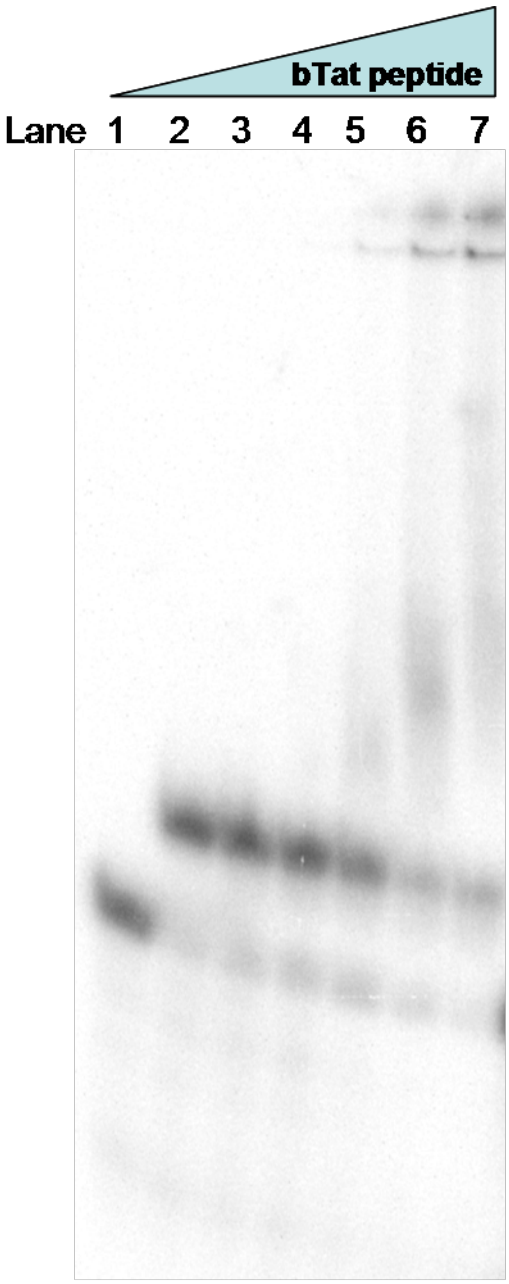


Figure 4.11. Electromobility gel shift assay of the native bTat peptide bound to bTAR RNA (10 nM). Lane 1: RNA only, Lane 2: 0.025 μM peptide, Lane 3: 0.1 μM peptide, Lane 4: 0.5 μM peptide, Lane 5: 1 μM peptide, Lane 6: 5 μM peptide, Lane 7: 10 μM peptide, Lane 8: 50 μM peptide. Buffer conditions were 50 mM Tris-HCl (pH 8.0), 50 mM KCl, 200 mM DTT, and 0.05% Triton X-100. Gel was a 12% native polyacrylamide in 0.5X TBE buffer electrophoresed at 15W in 4°C. The peptide and RNA were incubated at room temperature for 30 min before loading on the gel. The gel was dried for 1 h and phosphor imaged overnight.

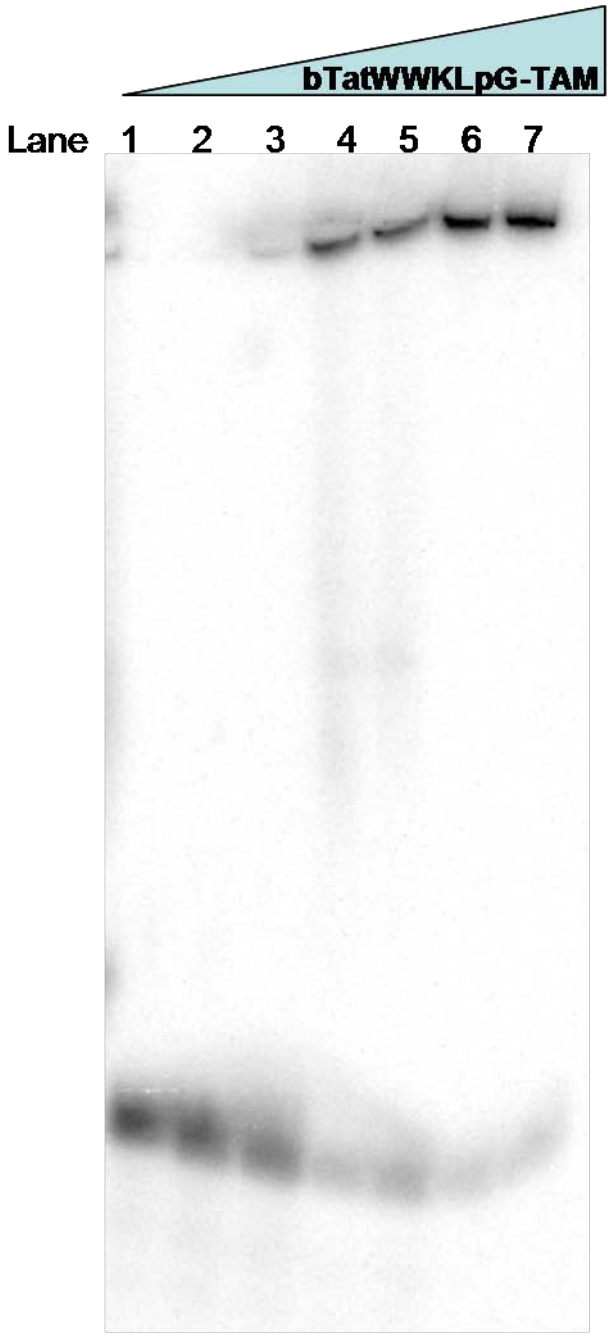


Figure 4.12. Electromobility gel shift assay of the native **TAMRA labeled bTatWWKL-pGturn** peptide bound to bTAR RNA (10 nM). Lane 1: RNA only, Lane 2: 0.025 μM peptide, Lane 3: 0.1 μM peptide, Lane 4: 0.5 μM peptide, Lane 5: 1 μM peptide, Lane 6: 5 μM peptide, Lane 7: 10 μM peptide, Lane 8: 50 μM peptide. Buffer conditions were 50 mM Tris-HCl (pH 8.0), 50 mM KCl, 200 mM DTT, and 0.05% Triton X-100. Gel was a 12% native polyacrylamide in 0.5X TBE buffer electrophoresed at 15W in 4°C. The peptide and RNA were incubated at room temperature for 30 min before loading on the gel. The gel was dried for 1 h and phosphor imaged overnight.

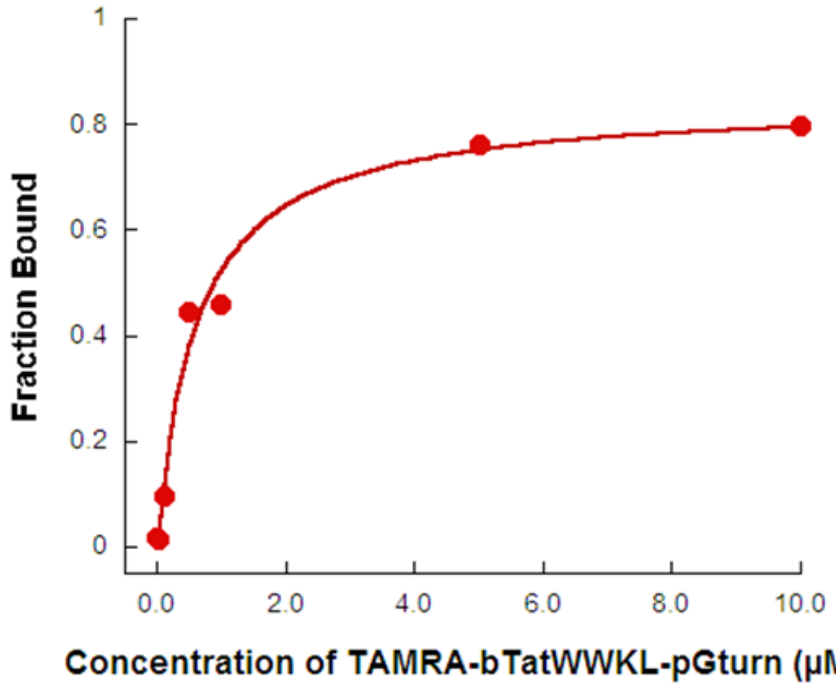


Figure 4.13. Fraction of 3'-p^{32}Cp-BIV TAR RNA that was bound by **TAMRA-bTatWWKL-pGturn** as determined by the electromobility gel shift assay. For each reaction 10 nM RNA was used, and the line was fit using Equation 2 and gives a K_d of 0.6 ± 0.2 μM. The data reflects the results from one gel shift assay.

D. Conclusions

The goal of this project was to mimic a known interaction between RNA and a cyclic peptide by designing a new peptide which was linear but whose structure was stabilized by side chain interactions. This "proof-of-principle" project would provide an example of how a stabilizing sequence of amino acids could be incorporated into a different peptide and still maintain structure and binding capability. We also set out to improve the original system by using a linear peptide that is both easier to synthesize and more accurately emulates the native interaction.

Our structural experiments proved that it was possible to incorporate the Trp pocket domain and promote β-hairpin structure. The CD experiments proved that when it is not cyclized, the BIV2 peptide is unstructured in solution. However with the addition of the Trp

pocket sequence, WWKL, and either an Asn-Gly or $_{D}$Pro-Gly turn sequence the peptide displays β-sheet structure. These results were confirmed using NMR studies.

We went on to investigate whether the bTatWWKL peptides maintained the low binding affinity for BIV TAR RNA that was found for BIV2. Fluorescence studies were hindered by quenching of the TAMRA label on the peptide. We suspect this may have been caused by intramolecular binding of the fluorophore to the rest of peptide, because it occurred even when no RNA was added. As an alternative method to measure binding, we used gel shift assays. Unfortunately no binding was detected for the unlabeled bTatWWKL-pGturn peptide to radiolabeled BIV TAR RNA. Further studies provided evidence that the peptide binds with a K_d of 0.6 μM, but only when the TAMRA-label is attached. We can conclude that by stabilizing the structure with the Trp pocket domain, we have altered the conformation of the peptide and perturbed binding to RNA.

In the future, it may be advantageous to revert back to the $_{D}$Pro-$_{L}$Pro turn template from BIV2, because that may provide an important contact for binding. It was significant that we were able to stabilize an unstructured peptide by incorporating the Trp pocket domain, but the bTat-bTAR interaction may not have be the right system to use for proof-of-principle studies. The BIV2 peptide contained five residues that were important for binding. However, it may be better to repeat these experiments with a system that has fewer contacts that need to be correctly aligned.

E. Experimental section

i. Peptide synthesis and labeling with fluorophore

The bTatWWKL peptides were synthesized on PAL-PEG-PS resin (0.1 mmol scale) using an Applied Biosystems Pioneer synthesizer. The peptides were fluorescently labeled on the free N-terminus by coupling overnight with 5-(-6)-carboxytetramethylrhodamine (0.16 mmol), HOBT (0.4 mmol), HBTU (0.4 mmol), and DIPEA (0.4 mmol) in DMF. Side-chain deprotection and cleavage from the resin was done using a TFA:Triisopropylsilane:Water mixture (95:2.5:2.5) and bubbling with N_2 for 3 hours. The peptides were isolated by first reducing the volume of TFA in a stream of nitrogen, then precipitating with cold diethyl ether, followed by extraction using deionized water. The resulting peptide solution was frozen and lyophilized. The peptides were purified using a Waters semi-preparative HPLC system with a reverse-phase C18 column. The peptides were eluted using Standard A (95% Water, 5% Acetonitrile, 0.1% TFA) and Standard B (95% Acetonitrile, 5% Water, 0.1% TFA). The masses of the peptides were confirmed using ESI-TOF mass spectrometry (Table 4.2).

ii. Fluorescence quenching experiments

Time experiments. The fluorescence of a 100 nM sample of bTatWWKL-NGturn peptide was measured on a Varian Cary Eclipse Fluorescence Spectrophotometer. The sample was excited at 549 nm at 900 volts and 10 nm slit widths and the emission was collected from 578-600 nm. The fluorescence of the same sample was intermittently measured over the time period of 45 min. In a separate experiment, two identical samples were made up and the fluorescence of the first was measured immediately and other was excited 30 min later with the same conditions.

Photobleaching experiments. Three samples each containing 100 nM bTatWWKL-NGturn were made, and one was used immediately to measure the fluorescence of the sample. The instrument conditions were set at 875 volts and 10 nm slit widths with the excitation at 549 nm and the emission collected ranging from 570-600 nm. The other two were stored overnight, one in the dark and the other exposed to light, and the fluorescence the tested the following day using the same conditions. Another 100 nM bTatWWKL-NGturn samples was then made using the same stock as the previous day. The fluorescence of this sample was also measured using the same instrument conditions.

iii. Fluorescence anisotropy binding experiments

The bTAR RNA was purchased as a purified, lyophilized powder from Integrated DNA Technologies. The RNA was dissolved in DEPC-treated water (100 μM) and annealed by heating to 95°C for 2 min, then slowly cooling to room temperature. The TAMRA-labeled bTatWWKL peptides were dissolved in DEPC-treated water and the concentration of the solution was determined using UV-Vis (ε_{559}=91000 M^{-1} cm^{-1}).[7] The binding experiments were done by titrating increasing amounts of folded RNA into 0.2 μM TAMRA-labeled peptide in reaction buffer (10 mM phosphate buffer, 140 mM Na^+/K^+Cl^-, 1 mM $MgCl_2$, pH 7.4). The anisotropy data were collected at 25°C in a 1 cm pathlength quartz microcuvette (excited at 548 nm and emission monitored at 580 nm) on the Cary Eclipse Fluorescence Spectrophotometer. The anisotropy data were plotted as a function of fluorescence anisotropy versus concentration of RNA. The plot was fitted using Equation 2 derived from a previous report:[8]

---

[7] Extinction coefficient provided by Integrated DNA Technologies, www.idtdna.com.

[8] Wang, Y.; Killian, J.; Hamasaki, K.; Rando, R. R. *Biochemistry* **1996**, *35*, 12338-46.

$$r = \left(\left(\frac{(a + x + K_D) \pm \sqrt{(-a - x - K_D)^2 - 4(a \cdot x)}}{2 \cdot a} \right) \cdot (r_\infty - r_0) \right) + r_0 \quad (\textit{Equation } 2)$$

where r is anisotropy, r_0 is initial anisotropy value, r_∞ is the maximum anisotropy, a is the total concentration of TAMRA-labeled peptide-intercalator, x is the total concentration of RNA added, K_D is the dissociation constant.

iv. Gel shift assays

3'-^{32}P RNA labeling. Approximately 10 μg of bTAR RNA construct was treated with T4 RNA ligase (2 U/μL), T4 RNA ligase buffer (1X), DMSO (10%), RNase inhibitor (1 U/μL), and ^{32}pCp (0.05 mCi) and incubated at 4°C overnight. The labeled RNA solution was run through an RNA spin column and loaded onto a 20% denaturing polyacrylamide gel for purification. The labeled RNA was visualized by autoradiography, excised from the gel and eluted overnight with elution buffer (0.5 M NH_4OAc, 1 mM EDTA, 0.1 % SDS). After filtration the RNA was precipitated with 95% ethanol, sodium acetate buffer, and linear acrylamide overnight at -20°C. The RNA was brought up in DEPC-treated water and the concentration was determined by UV absorbance using the equation: concentration of RNA SAMPLE = 40 × OD_{260} × dilution factor.

Gel shift assays. The labeled RNA was brought up in DEPC-treated water and annealed at 95°C for 2 min and slowly cooled to room temperature. The bTatWWKL-pGturn peptide was synthesized using the same procedure as above except without the fluorophore label and instead contained an N-terminal acetyl cap. Stock solutions were made by dissolving the peptides in DEPC-treated water and measuring the tryptophan absorbance (ε_{280}= 11380 M^{-1} cm^{-1}) in the presence of 5M guanidine hydrochloride. Increasing amounts of peptide solution was added to 10 nM bTAR RNA and incubated at room temperature in 0.5X Tris buffer (50

mM Tris-HCl, 50 mM KCl, 0.05% Triton X, 200 mM DTT, pH 8.0) for 30 minutes. To each reaction was added 1 uL 80% glycerol BPB dye. After the addition of the dye, the total reaction volume was 11 uL, and 10 uL of the sample was loaded onto a 12% native polyacrylamide gel. The gel was electrophoresed in 0.5X TBE buffer at 15W and 4°C. The gels were exposed to the phosphor imaging screen overnight and scanned using a Storm phosphor imager.

Table 4.2. Mass spectrometry data for peptides used in this study.

Peptide	Expected mass (Da)	Observed mass (Da)
bTatWWKL-NGturn (TAMRA)	2235.3	2235.6
bTatWWKL-pGturn (TAMRA)	2218.3	2218.2

Chapter V

DESIGN OF HIGHLY STRUCTURED β-HAIRPIN PEPTIDES WITH IMPROVED RESISTENCE TO PROTEOLYSIS

(Reproduced, in part, with permission from Cline, L.L.; Waters, M.L. *accepted for publication in Peptide Science.)*

A. Background

The use of peptides as drugs, especially those taken through oral ingestion, is hindered by the fact that peptides are readily degraded by proteases in vivo. Despite the disadvantages of peptide drugs, the advantages could include higher potency, higher specificity, and fewer problems with toxicity than small organic drugs.[1] Therefore efforts are being made to find ways of making peptides more resistant to proteases.

One approach has been to design peptidomimetics that resemble L-α-amino acids but are more resistant to proteolytic cleavage. Hamamoto et al. have incorporated D-amino acids into antimicrobial peptides and found that the unnatural peptides retained the ability to lyse bacteria but were more stable to trypsin cleavage than the L-amino acid counterparts.[2] The incorporation of $_{D}$-amino acids has also been used to improve peptidomimetic vaccines.

---

[1] Marx, V. *Chem Eng News* **2005**, *83*, 17-24.

[2] Hamamoto, K.; Kida, Y.; Zhang, Y.; Shimizu, T.; Kuwano, K. *Microbiol Immunol* **2002**, *46*, 741-9.

Tugyi et al. have incorporated $_{D}$-Thr and $_{D}$-Pro into the N- and C-terminus of peptide immunogens, which added resistance to enzymatic degradation and still bound to the appropriate antibodies.[3] Additionally Saikumari et al. have reported the design of peptides that are resistant to trypsin cleavage when a $_{D}$-Pro-$_{L}$-Pro or $_{D}$-Pro-Gly sequence is placed near the cleavage site. The use of $_{D}$-Pro also hindered the degradation of the peptides by proteinase K, subtilisin, elastase, and collagenase.[4]

Another example of the use of unnatural amino acids in peptides is the incorporaton of β- and γ-amino acids in place of α-amino acids. The Seebach lab has found that peptides containing all β- and all γ-amino acids were stable to fifteen different peptidases *in vitro*.[5] A similar approach by Porter et al. showed that β-peptides that were designed to mimic antimicrobial α-peptides were also resistant to proteolytic degradation and retained their activity.[6] The advantages to these approaches are that β-peptides form complementary secondary structures to native motifs such as helices, β-sheets, and hairpins. However, many of the unnatural amino acids used to make these peptides are not commercially available and need to be manually synthesized. In addition, there is no way of expressing the unnatural peptides in bacteria.

Previously our group has designed an all α-amino acid peptide, **WKWK**, which has a high propensity to form a β-hairpin and has been characterized by NMR to be 95% folded in

---

[3] Tugyi, R.; Uray, K.; Ivan, D.; Fellinger, E.; Perkins, A.; Hudecz, F. *Proc Natl Acad Sci U S A* **2005**, *102*, 413-8.

[4] Saikumari, Y. K.; Ravindra, G.; Balaram, P. *Protein Pept Lett* **2006**, *13*, 471-6.

[5] (a) Frackenpohl, J.; Arvidsson, P. I.; Schreiber, J. V.; Seebach, D. *Chembiochem* **2001**, *2*, 445-55. (b) Hook, D. F.; Bindschadler, P.; Mahajan, Y. R.; Sebesta, R.; Kast, P.; Seebach, D. *Chem Biodivers* **2005**, *2*, 591-632.

[6] Porter, E. A.; Weisblum, B.; Gellman, S. H. *J Am Chem Soc* **2002**, *124*, 7324-30.

aqueous environment (Figure 5.1).[7] This highly structured peptide provides the ideal system to study whether or not β-hairpin secondary structure in small peptides can provide protection from peptidase degradation. Recent studies in this field have used FRET based assay both in vivo and in vitro to study structure and stability of small peptides.[8] We have used HPLC based experiments which remove the need for bulky GFP-derived fluorophores that may interfere with the enzyme binding and function. We have chosen to test the stability of the peptides against three different types of proteases. α-Chymotrypsin and trypsin are serine endopeptidases found in the pancreas which cleave specifically at bulky aromatic residues and basic residues, respectively. Figure 5.2 shows examples of the possible cut sites for these enzymes degrading the **WKWK** peptide. Pronase E is a zinc-dependent metalloendopeptidase from *Streptomyces griseus* which shows no preference for cleavage site.

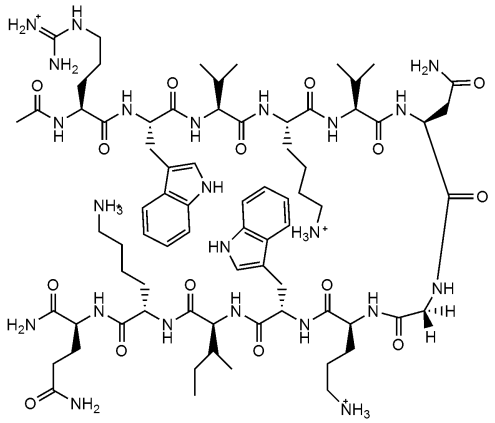


Figure 5.1. Structure and sequence of the **WKWK** peptide.

[7] Butterfield, S. M.; Waters, M. L. *J Am Chem Soc* **2003**, *125*, 9580-1.(b) Butterfield, S. M.; Sweeney, M. M.; Waters, M. L. *J Org Chem* **2005**, *70*, 1105-14.

[8] Cheng, Z.; Campbell, R. E. *Chembiochem* **2006**, *7*, 1147-50.

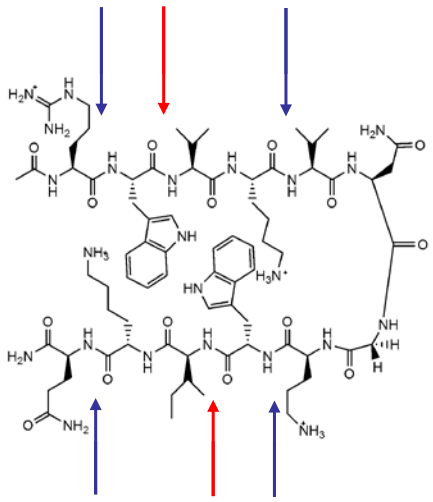


Figure 5.2. Possible cut sites on the **WKWK** peptide for α-chymotrypsin (red) and trypsin (blue) proteases. Both proteases cleave the C-terminal peptide bond.

B. Design and synthesis

To correlate secondary structure and peptidase resistance, peptides were designed with varying propensities to fold into a β-hairpin (Figure 5.3). Each of the peptides was based on the **WKWK** peptide scaffold, which contains all α-amino acids. An unstructured version of this peptide which has all the same residues in a scrambled sequence was synthesized as a control (Figure 5.3a). One of the enzymes used in this study, α-chymotrypsin, is an endopeptidase from bovine pancreas known to readily cleave proteins and peptides after bulky aromatic residues such as Trp, Tyr, and Phe, and more slowly cleave after Asn, His, Met, and Leu.[9] Therefore **WKWK-scrambled** was designed so that the residues flanking the amide bond where cleavage occurs remained consistent to those in the parent peptide.

[9] Voet, D.; Voet, J. *Biochemistry*; 3rd ed.; John Wiley & Sons, Inc: Hoboken, NJ, 2004.

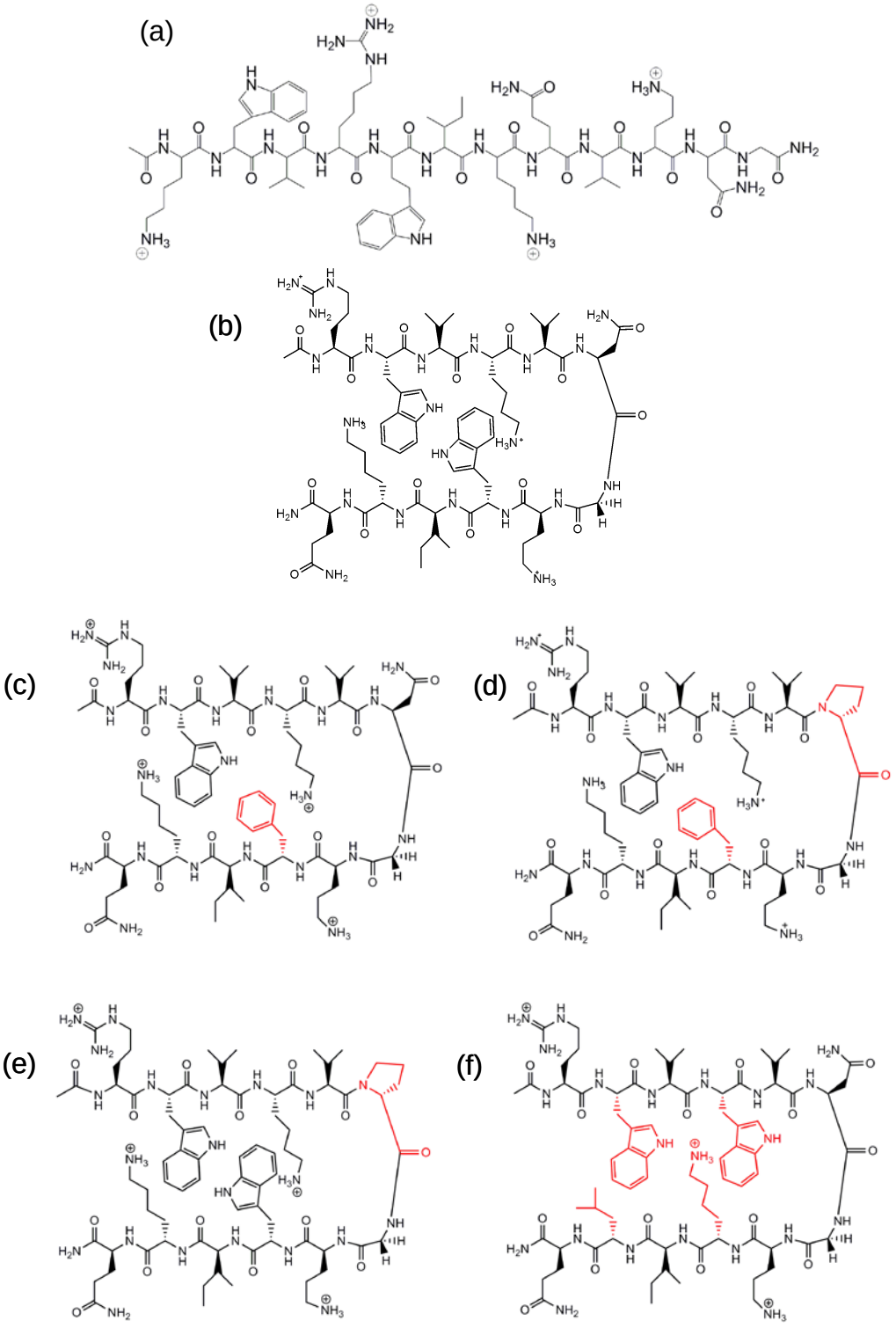


Figure 5.3. Sequence and structures of all the peptides used in the protease studies. (a) **WKWK-scrambled**. (b) **WKWK**. (c) **WKFK**. (d) **WKFK-pGturn**. (e) **WKWK-pGturn**. (f) **Trp Pocket**. Residues that are highlighted in red are different from the WKWK peptide.

C. Results and discussion

i. Structural characterization of the peptides

Circular dichroism studies were performed to confirm that **WKWK** exists as a folded β-hairpin at physiological pH and temperature, whereas WKWK-scrambled is unstructured (Figure 5.4). **WKWK-scrambled** showed a characteristic CD signal for random coil peptides whereas **WKWK** has a minimum signal around 210 nm, which is slightly shifted from the typical minimum of 215 nm for β-sheet peptides. This is likely due to the contribution of Trp-Trp exciton coupling, as indicated by the maximum at 226 nm.

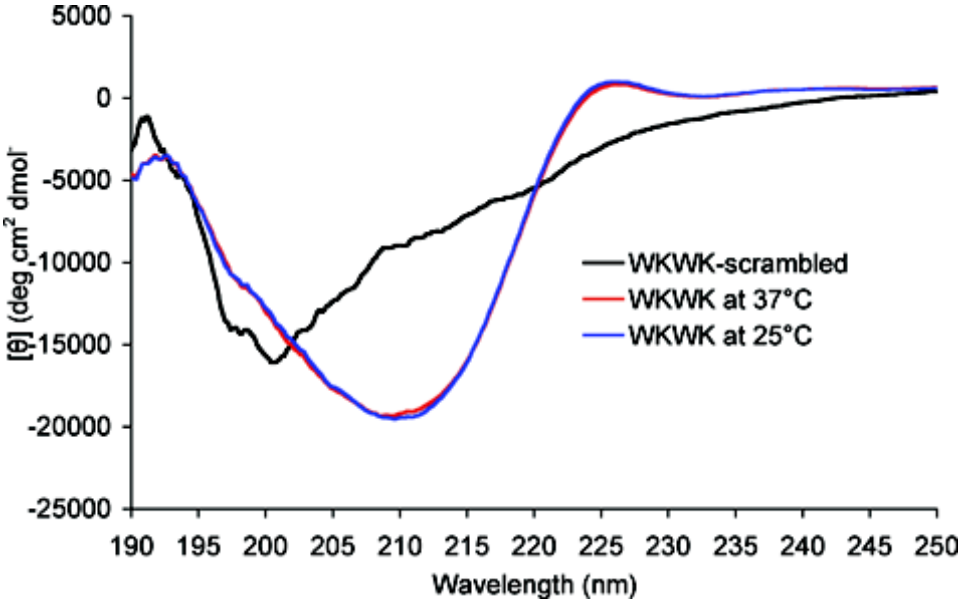


Figure 5.4. Circular Dichroism Spectra for **WKWK** and **WKWK-scrambled** peptides. Studies were done at 100 μM peptide concentration in 10 mM sodium phosphate buffer pH 7.6 at 25°C and 37°C. **WKWK** (red and blue) folds into a β-sheet conformation with a minimum mean residue ellipticity around 212 nm. **WKWK-scrambled** (black) is shown to be in a random coil conformation.

ii. NMR characterization of peptide folding

For all the other peptides used in this study, mutations were made to the **WKWK** peptide to alter the stability of the folded structure (Table 5.1). Folding and stability of the hairpins was investigated by NMR and circular dichroism (CD). One-dimensional NMR studies were preformed to determine the fraction folding for the peptides based on Hα chemical shifts and glycine splitting as described in the Experimental section.[10] Apart from the unfolded control, the **WKFK** peptide was the least folded at 67%. This peptide has Trp9 mutated to Phe, which decreases the strength of its cross-strand interactions with Trp2 and Lys4 and has a lower β-sheet propensity. By incorporating a $_{D}$-proline in the turn region of **WKFK** (**WKFK-pGturn**) the turn becomes more rigid and the fraction folding increases to 86%. To increase the extent of folding, a $_{D}$-proline was also incorporated into the **WKWK** peptide. It was also possible to make a more stable β-hairpin peptide by increasing the cross strand interactions of the peptide. The **Trp pocket** peptide has been shown to create an aromatic pocket with which the lysine side chain can interact and create a cation-π interaction. This peptide which contains all α-amino acids is the most well folded small β-hairpin peptide reported to date.[11]

[10] Griffiths-Jones, S. R.; Maynard, A. J.; Searle, M. S. *J Mol Biol* **1999**, *292*, 1051-69.
[11] Riemen, A. J.; Waters, M. L. *Biochemistry* **2009**, *48*, 1525-31.

Table 5.1. Sequences of the β-hairpin peptides used in this study and the corresponding percent folded values at 25°C.

Peptide name	Sequence	Percent Folded	Source
WKWK-scrambled	Ac-KWVRWIKQVONG-NH_2	0%[a]	This work
WKFK	Ac-RWVKVNGOFIKQ-NH_2	67% [b]	Ref 7b
WKFK-pGturn	Ac-RWVKVpGOFIKQ-NH_2	86%[c]	This work
WKWK	Ac-RWVKVNGOWIKQ-NH_2	96%[b] (96%)[d]	Ref 7a
WKWK-pGturn	Ac-RWVKVpGOWIKQ-NH_2	>98%[c]	This work
Trp pocket	Ac-RWVWVNGOKILQ-NH_2	>99% [b,c] (98%)[d]	Ref 11

(a) The percent folded of WKWK-scrambled was estimated from the CD data. (b) Determined using 1D NMR glycine splitting data. (c) Determined based on Hα chemical shifts. Conditions: 50 mM NaOAc-d3 (pD 4.0), and experiments were done at 25°C. (d) Value in parentheses is the fraction folded at 37°C.

iii. Peptidase studies using α-chymotrypsin

α-Chymotrypsin cleaves on the C-terminal side of aromatic amino acids. All the peptides studied contain two aromatic residues that which are followed by a branched aliphatic amino acid, so differences in the rate of cleavage are not a result of differences in the neighboring residue. The unstructured peptide was used to determine the concentration of enzyme for each reaction. Conditions were optimized to give a $t_{1/2}$ of 7 min for the scrambled peptide as a basis for comparison. The reactions were quenched with glacial acetic acid at various time points up to 5 h, and the decrease of the full-length peptide was monitored with RP-HPLC. An example of the resulting HPLC trace is shown in Figure 5.5. To analyze the data, the parent peak was integrated using Empower software and plotted as a function of time.

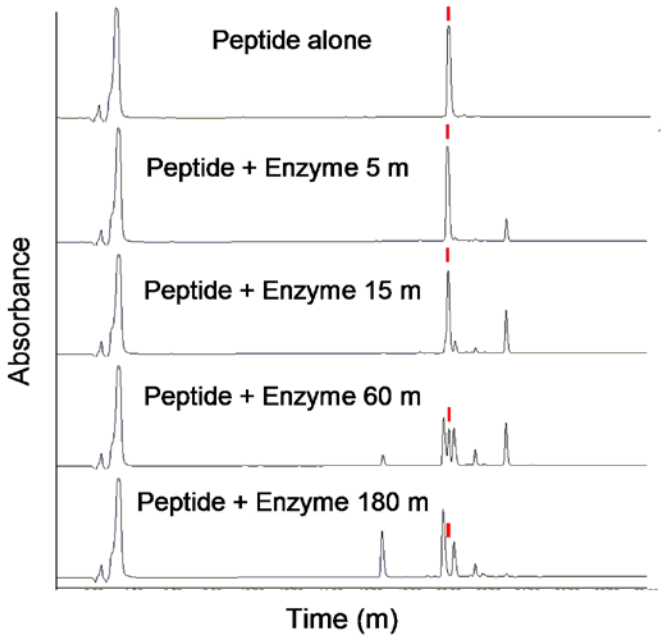


Figure 5.5. Example of how RP-HPLC was used to monitor the degradation of peptide by enzyme. The full length peptide peak is highlighted in red, and disappears at increasing time points. Acetic acid, which was used to both quench the reaction and as an internal control, elutes at 3 minutes.

The structure of the β-hairpin peptides significantly affected the peptide degradation by α-chymotrypsin (Figure 5.6). After 15 minutes there was no full-length **WKWK-scrambled** remaining, but all the structured peptides were still completely intact. The **WKFK** peptide was degraded with a half-life of approximately 36 min, and the **WKWK** peptide was degraded by 50% in approximately 198 min. The other two peptides containing $_{D}$-Pro in the turn were completely stable to the enzyme even after 5 hours. All of the peptides have the same number of cleavage sites for α-chymotrypsin, therefore any difference in the rate of cleavage is due to the accessibility of the aromatic residues to the enzyme. Analysis by mass spectrometry shows that in the case of **WKWK** the enzyme first cleaves either Trp2 or Trp9 and then shortens the peptides further by cutting at the second site (Table 5.2). Since

WKFK-pGturn and **WKWK-pGturn** demonstrated complete resistance over 5 h, **Trp pocket**, which is the most well-folded, was not investigated with this enzyme.

To further demonstrate that the structure of the peptide is the reason for the enzymatic resistance, equal concentrations of **WKWK** and **WKWK-scrambled** were reacted with α-chymotrypsin (Figure 5.7). After 15 min, the parent peak for **WKWK-scrambled** is completely degraded whereas the **WKWK** peptide has been untouched by the enzyme. A similar experiment was also done by incubating **WKWK-pGturn** alone with α-chymotrypsin for 5 h and then adding **WKWK-scrambled** peptide (Figure 5.8). **WKWK-pGturn** was not cleaved during the 5 hours or after **WKWK-scrambled** was added, however the unstructured peptide was completely cleaved within 30 min of incubation. This experiment confirms that the enzyme is still active after 5 h, and that the structure of the peptide is inhibiting the degradation.

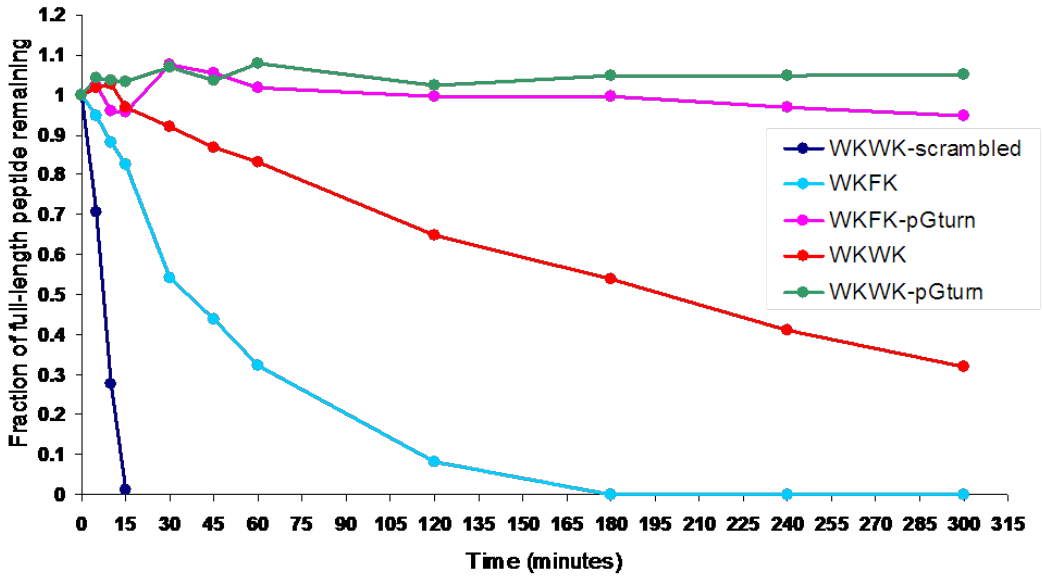


Figure 5.6. HPLC analysis of peptide degradation by α-chymotrypsin. The **WKWK-scrambled** is completely degraded after 15 min. The **WKFK**, **WKFK-pGturn**, **WKWK**, and **WKWK-pGturn** peptides are all substantially more resistant to enzymatic degradation. Experiments were done in duplicate at physiological conditions.

Table 5.2. Percentage of total peak area for the full-length **WKWK** peptide and each of the cleavage products from α-chymotrypsin degradation

	Peak Area (%)[a]				
	Full Length **WKWK** Ac-RWVKVNGOWIKQ-NH_2 Mass= 1567.8 g/mol	Ac-RWVKVNGOW-CO_2^- Mass= 1199.6 g/mol	H_3N^+-VNGOWIKQ-NH_2 Mass= 1183.6 g/mol	Ac-RW-CO_2^- Mass= 402.2 g/mol	H_3N^+-VKVNGOW-CO_2^- Mass= 815.4
5 m	97.6	1.2	0.7	0.5	0.0
15 m	94.2	1.2	2.6	2.0	0.0
30 m	89.1	2.0	4.2	4.0	0.7
60 m	80.5	2.7	6.6	8.3	1.9
120 m	65.1	4.2	9.5	15.1	6.0
180 m	51.3	3.6	11.2	22.7	11.2
240 m	39.1	3.1	11.3	30.5	16.1
300 m	30.3	2.3	10.7	35.0	21.7

[a] Each reaction was analyzed using RP-HPLC and peaks were identified by mass spectrometry. The percent peak area was calculated by dividing the individual peak area by the total area of all peaks.

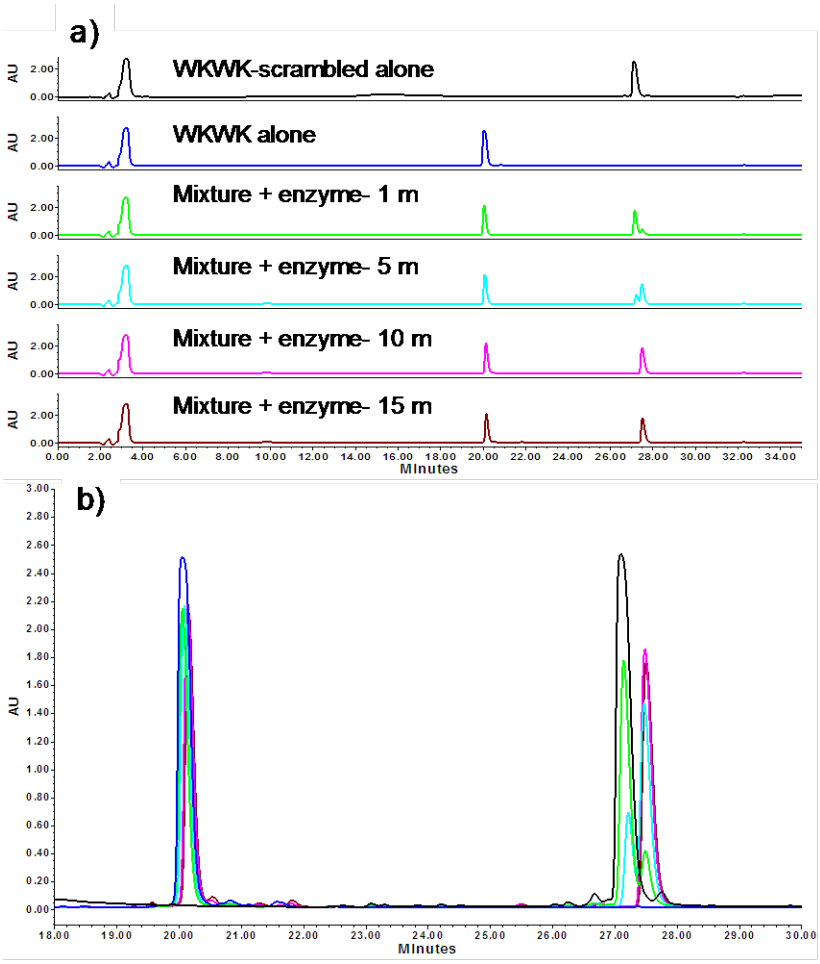


Figure 5.7. HPLC trace of **WKWK-scrambled** and **WKWK** alone and then an equimolar mixture of the two reacted with α-chymotrypsin. (a) Reactions were quenched at varying time points. The **WKWK** peptide elutes at 20 min and **WKWK-scrambled** elutes at 27 minutes. The **WKWK-scrambled** peak decreases within 5 min and a new peak grows in beside it, whereas the full-length **WKWK** peak is still present after 15 min. (b) Overlay view of the relevant portion of the graph.

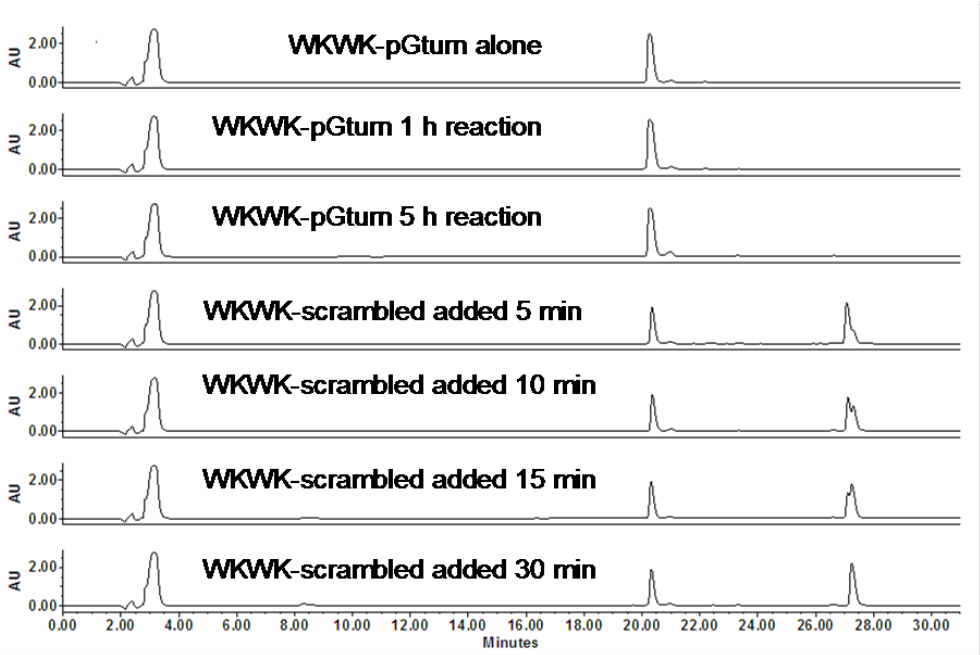


Figure 5.8. HPLC trace of α-chymotrypsin degradation of **WKWK-pGturn** over five hours and then subsequent addition of **WKWK-scrambled** peptide to the solution. The plot shows **WKWK-pGturn** alone and reactions with the enzyme after 1 h and 5 h. **WKWK-scrambled** peptide was then added and aliquots were quenched after 5 m, 10 m, 15 m, and 30 m. The full-length **WKWK-pGturn** elutes at approximately 20.5 min and the full-length **WKWK-scrambled** peptide elutes at approximately 27.0 min. All other peaks correspond to acetic acid (3.0 min) or degradation products.

iv. Peptidase studies using trypsin

Trypsin cleaves after basic amino acids such as Arg and Lys. Degradation studies show that unlike α-chymotrypsin, the peptides were not completely stable to trypsin. However, the structured peptides were still substantially more stable than the unstructured control (Figure 5.9). The degradation times correlated well with the extent of folding. **WKFK** and **WKFK-pGturn** both degraded at the same rate with a half-life of approximately 23 min, but this was still four times slower than the half-life for the unstructured peptide. The similar rate for **WKFK** and **WKFK-pGturn** suggests that the more rigid turn did not deter the enzyme from cleaving the peptides. The **Trp pocket** peptide was the most stable to trypsin. However, this

peptide had one less basic residue than the other peptides, which could be the cause of the increased resistance.

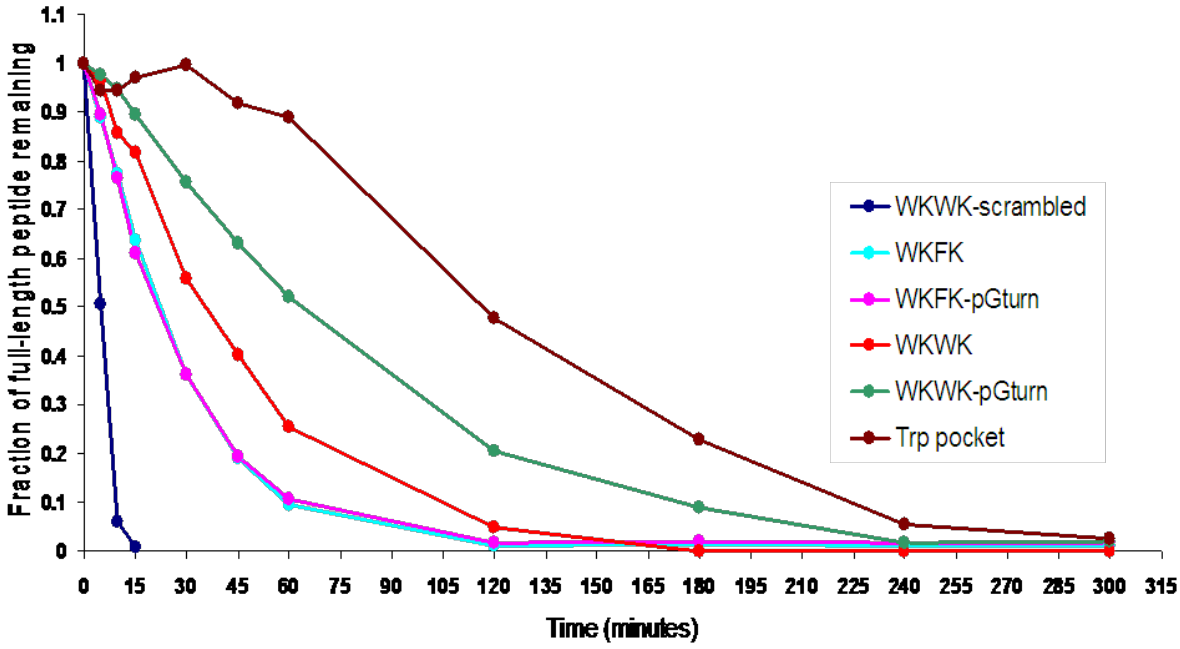


Figure 5.9. HPLC analysis of peptide degradation by trypsin. The **WKWK-scrambled** peptide is completely degraded after 15 minutes. The **WKFK**, **WKFK-pGturn**, **WKWK**, **WKWK-pGturn**, and **Trp pocket** peptides are all substantially more resistant to enzymatic degradation. Experiments were done in duplicate at physiological conditions.

v. Peptidase studies using pronase E

The most aggressive enzyme tested in these experiments was pronase E, which is a non-specific serine protease found in S. griseus. This enzyme degraded the least folded β-hairpin peptide **WKFK** in the same amount of time it degraded the **WKWK-scrambled** peptide, however the other peptides demonstrated resistance (Figure 5.10). The **WKFK** peptide, which is only 67% folded, may be flexible enough for the enzyme to cleave it quickly, but as the peptides become more folded the pronase E is less able to cut the peptides. It is interesting to note that different rates of degradation are observed with pronase E even

though the terminal residues of β-hairpins are frayed. For this enzyme the fraction folded did not directly correlate with the degradation rate. Instead, both **WKFK-pGturn** and **WKWK-pGturn** were much more peptidase-resistant than their analogs with an Asn-Gly turn sequence. The **WKFK-pGturn** peptide is equally resistant to enzyme cleavage as **WKWK** although WKFK-pGturn is less folded. The same trend was seen for **WKWK-pGturn** which is slightly less folded than the **Trp pocket** peptide, but is more resistant to pronase E. The fact that both **WKWK-pGturn** and **WKFK-pGturn** contain D-proline must cause the slower rates of degradation, because of the rigidity of the β-hairpin structure.

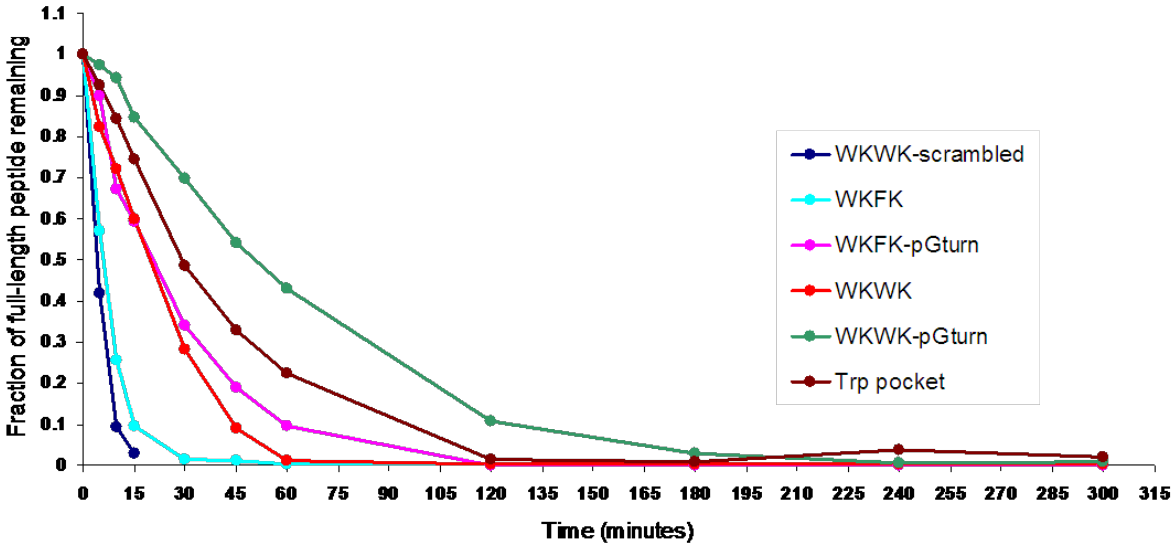


Figure 5.10. HPLC analysis of peptide degradation by pronase E. The **WKWK-scrambled** is completely degraded after 15 min. The **WKFK**, **WKFK-pGturn**, **WKWK**, **WKWK-pGturn**, and **Trp pocket** peptides are all substantially more resistant to enzymatic degradation. Experiments were done in duplicate at physiological conditions.

Table 5.3. Approximate half-life for proteolytic degradation of the β-hairpin peptides

	Approximate $t_{1/2}$ (minutes)[a]		
Peptide name	α-Chymotrypsin	Trypsin	Pronase E
WKWK-scrambled	7	5	5
WKFK	36	23	6
WKFK-pGturn	>300	23	21
WKWK	198	36	20
WKWK-pGturn	>300	64	51
Trp pocket	n/d[b]	117	30

(a) Approximate time when 50% of the uncleaved peptide is remaining
(b) Not determined

D. Conclusions

This study used in vitro experiments to demonstrate that β-hairpin structure, even in a small 12-residue peptide, can add resistance to proteolytic degradation. The structures were stabilized by incorporation of cross-strand cation-π interactions and/or of a more rigid **$_D$-Pro-Gly** turn sequence. The results from all the experiments are summarized in Table 5.3, and it shows that for trypsin there was a direct correlation between fraction folded and stability. For both α-chymotrypsin and pronase E an increase in fraction folded is beneficial, but the use of unnatural $_D$-proline residue in the turn also added to its resistance.

We have shown with the **Trp Pocket** peptide that increasing the side chain-side chain interactions to promote structure added to the proteolytic resistance. Therefore the peptide **WKWK-pGturn** which contained both a rigid β-turn and strong side chain interactions demonstrated the greatest resistance across the set of enzymes tested. Although these peptides did not completely block degradation, this may be an advantage for applications

where the peptide needs to be degradable over time. For the design of peptide based drugs, the use of a highly structured peptide is a viable way to add resistance to proteolysis without using cyclic peptides or unnatural amino acids. Moreover, the combination of peptide structure with unnatural amino acids, such as $_{D}$-proline, can provide even greater resistance in some cases.

E. Ongoing work: protease studies in a biological system

The *in vitro* studies proved that β-hairpin structure can help to add resistance to proteolytic degradation. However, those studies were done in a controlled system and only tested a single enzyme at a time. We wanted to investigate whether our structured peptides, which we've nicknamed "protectides", were still more stable than an unstructured control in a more biologically relevant system.

For this reason we began collaborating with the Allbritton Lab in order to test *ex vivo* the ability of our protectides to block the degradation of a Protein Kinase C (PKC) substrate. The studies will be done with cell lysate and if the results are promising the protectides will be microinjected into whole cells for testing. There are very few peptidases found in the cytoplasm of the cell, and of those that are known to be either aminopeptidases or endopeptidases. The crystal structures of these enzymes show that most of these enzymes have a deep cleft through which the peptide substrate is threaded.[12] To protect the PKC substrate against aminopeptidases, we have attached the protectides to the N-terminus. We

[12] (a) Fulop, V.; Bocskei, Z.; Polgar, L. *Cell* **1998**, *94*, 161-70. (b) Aertgeerts, K.; Ye, S.; Tennant, M. G.; Kraus, M. L.; Rogers, J.; Sang, B. C.; Skene, R. J.; Webb, D. R.; Prasad, G. S. *Protein Sci* **2004**, *13*, 412-21. (c) Brandstetter, H.; Kim, J. S.; Groll, M.; Huber, R. *Nature* **2001**, *414*, 466-70. (d) Guhaniyogi, J.; Sohar, I.; Das, K.; Stock, A. M.; Lobel, P. *J Biol Chem* **2009**, *284*, 3985-97. (e) Pal, A.; Kraetzner, R.; Gruene, T.; Grapp, M.; Schreiber, K.; Gronborg, M.; Urlaub, H.; Becker, S.; Asif, A. R.; Gartner, J.; Sheldrick, G. M.; Steinfeld, R. *J Biol Chem* **2009**, *284*, 3976-84. (f) Ray, K.; Hines, C. S.; Coll-Rodriguez, J.; Rodgers, D. W. *J Biol Chem* **2004**, *279*, 20480-9.

hypothesize that the structural stability of the protectides will also hinder degradation by endopeptidases, because it will be difficult to thread the structured peptide through the enzyme's deep cleft. The Allbritton lab will utilize capillary electrophoresis (CE) to track the activity of the PKC enzyme by monitoring phosphorylation of a fluorophore-labeled substrate. If the PKC substrate is being degraded by endogenous proteases, this can also be detected by this method. Enzyme activity is dependent on the structural stability of the substrate peptide and by attaching our WKWK-derived protectides we hope to increase resistance to proteolysis. Using the developed CE methodology we will be able to test the stability of the PKC substrate covalently linked to the series of **WKWK** protectides in a cellular environment.

The protectides were attached to the N-terminus of a PKC substrate with the sequence RFRRFQTLKIKAKA (Figure 5.11). This sequence was used by Newton et al. and is known to be a good substrate for all isoforms of PKC, but is not efficiently phosphorylated by other kinases.[13] We also chose this particular substrate because it was used by others for in cell studies, and its efficiency and specificity are well-studied in that environment. We have incorporated a two unit polyethylene glycol (PEG_2) linker between the two pieces to create some distance from the protein binding site and the protectides. The PEG_2 linker also helps to increase the solubility of the large complex which aids in the synthesis. As a requirement for the CE method, we have labeled the protectide portion with (6)-carboxyfluorescein (6-FAM) through a lysine side chain. The lysine is positioned near the β-hairpin turn and has

[13] (a) Violin, J. D.; Zhang, J.; Tsien, R. Y.; Newton, A. C. *J Cell Biol* **2003**, *161*, 899-909. (b) Gallegos, L. L.; Kunkel, M. T.; Newton, A. C. *J Biol Chem* **2006**, *281*, 30947-56. (c) Nishikawa, K.; Toker, A.; Johannes, F. J.; Songyang, Z.; Cantley, L. C. *J Biol Chem* **1997**, *272*, 952-60.

been previously shown not to be disruptive to the peptide structure.[14] In total we aim to make five protectide-PKC complexes (Table 5.4). The protectide sequences were the most folded peptides from the previous studies, as well as the unstructured WKWK-scrambled sequence as a negative control and a cyclic WKWK sequence.

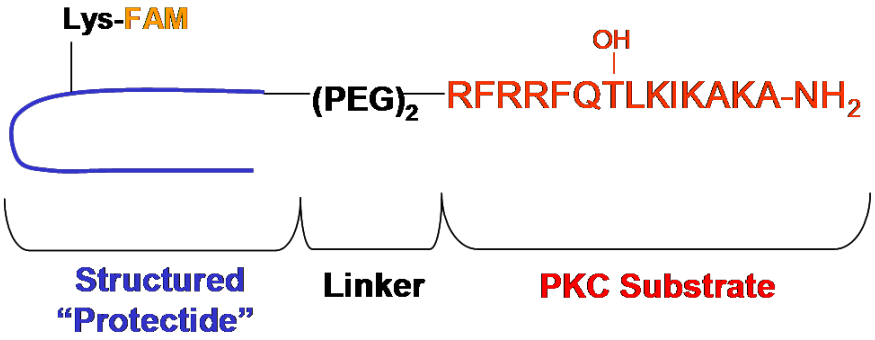


Figure 5.11. Diagram of the **Protectide-Peg$_2$-PKC** complexes. Both the site of PKC phosphorylaion and the site of fluorescein attachment are shown.

Table 5.4. Proposed sequences for the series of **Protectide-Peg$_2$-PKC** complexes

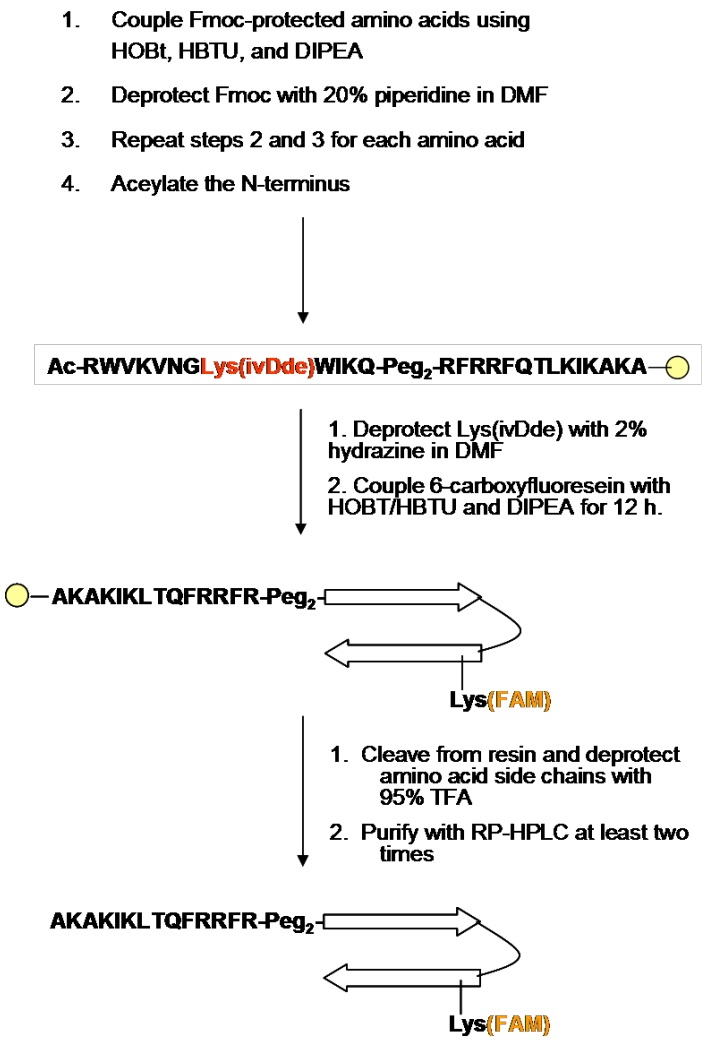
Name	Sequence
WKWKscram-Peg$_2$-PKC	Ac-K(FAM)WVRWIKQVKNG-Peg$_2$-RFRRFQTLKIKAKA-NH$_2$
WKWK-Peg$_2$-PKC	Ac-RWVKVNGK(FAM)WIKQ-Peg$_2$-RFRRFQTLKIKAKA-NH$_2$
WKWKpGturn-Peg$_2$-PKC	Ac-RWVKVpGK(FAM)WIKQ-Peg$_2$-RFRRFQTLKIKAKA-NH$_2$
TrpPkt-Peg$_2$-PKC	Ac-RWVWVNGK(FAM)KILQ-Peg$_2$-RFRRFQTLKIKAKA-NH$_2$
cycWKWK-Peg$_2$-PKC	HN-RWVKVNGK(FAM)WIKQC-Peg$_2$-RFRRFQTLKIKAKA-NH$_2$ 

The synthetic scheme for these large complexes was relatively straightforward using solid phase peptide synthesis, but it was not trivial (Scheme 5.1). One issue was that the overall size of the complex, 28 residues total, is not ideal for an on-bead synthesis. The

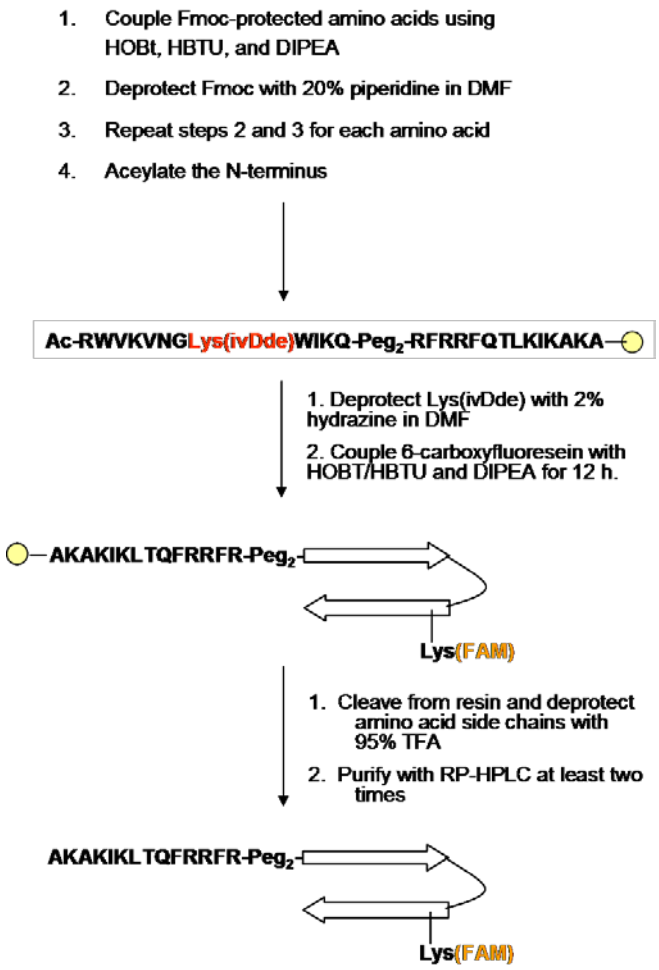
[14] Cooper, W. J.; Waters, M. L. *Org Lett* **2005**, *7*, 3825-3828.

yields tend to be much lower for long peptides and the increase in truncated sequences makes the purification more difficult. The 6-FAM was coupled to the lysine side chain after orthogonal deprotection of the amine. This step is done immediately before cleavage from the bead. The synthesis of the fully folded **cycWKWK-Peg$_2$-PKC** complex will involve bromo-acetylation of the N-terminus on bead and then intramolecular cyclization through a cysteine side chain in solution (Scheme 5.2).[15] This type of cyclization was chosen over the commonly used disulfide bond due to its stability in the reducing cellular environment. Three protectide-PKC complexes have been synthesized thus far, **WKWKscram-Peg$_2$-PKC**, **WKWK- Peg$_2$-PKC**, and **TrpPkt-Peg$_2$-PKC**. The synthesis issues were overcome by increasing the coupling times and purifying the peptides twice. Several other derivatives of the WKWK peptide series are currently being made, and the Allbritton lab is carrying out the CE analysis of peptide degradation.

[15] Ivanov, B.; Grzesik, W.; Robey, F. A. *Bioconjug Chem* **1995**, *6*, 269-77.



Scheme 5.1. Synthesis of the linear **Protectide-Peg2-PKC** complexes



Scheme 5.2. Synthesis of the **cycWKWK-Peg2-PKC** complex

F. Experimental section

i. Peptide synthesis and purification

Peptides were synthesized on either an Applied Biosystems Pioneer or Creosalus Tetras automated peptide synthesizer. All the peptides were synthesized using standard solid phase peptide synthesis protocols with Fmoc-protected amino acids and Fmoc-PAL-PEG-PS resin. Peptides were synthesized on a 0.1 mmol scale and capped at the N-terminus with an acetyl group. All the peptides were cleaved and deprotected by treating the resin with a (95:2.5:2.5) TFA:Triisopropylsilane:Water mixture and bubbling with N_2 for 3 h. The peptides were

isolated by first reducing the volume of TFA in a stream of nitrogen, then precipitating with cold diethyl ether, followed by extraction using deionized water. The resulting peptide solution was frozen and lyophilized. The peptides were purified using a Waters semi-preparative HPLC system with a reverse-phase C18 column. The peptides were eluted using Standard A (95% Water, 5% Acetonitrile, 0.1% TFA) and Standard B (95% Acetonitrile, 5% Water, 0.1% TFA). The masses of the peptides were confirmed using ESI-TOF mass spectrometry (Table 5.5).

ii. Peptide concentration determination

Approximately 15 mg of peptide was brought up in 2-3 mL of 10 mM sodium phosphate, 140 mM Na^+/K^+Cl^- buffer (pH 7.6). UV-Vis was used to determine the concentration of the peptide stocks by dilution into 6 M guanidine hydrochloride and measuring absorbance at 280nm (Trp ε_{280}=5690 $M^{-1}cm^{-1}$).

iii. Peptidase concentration

Enzymes were purchased from Sigma-Aldrich, Pronase E from Streptomyces griseus (EC 3.4.24.31) as a solid, α-chymotrypsin from bovine pancreas (3.4.21.1) as a lyophilized powder, and trypsin from bovine pancreas as a lyophilized powder containing lactose. Pronase E, α-chymotrypsin, and trypsin were brought up in 10 mM sodium phosphate, 140 mM NaCl buffer, pH 7.6, to approximately 0.1 mg/mL, 0.031 mg/mL, and 0.011 mg/mL, respectively. For each enzyme, the volume used was 6 μL, 10 μL, and 5 μL, respectively, and these values were based on the amount of enzyme that completely cleaved **WKWK-scrambled** peptide in 15 min.

iv. Peptidase degradation reactions

The procedure for peptidase studies was adapted from Seebach et al.[5a] For each reaction, 0.5 mM peptide in the phosphate buffer (140 mM Na^+/K^+Cl^- buffer pH 7.6) was reacted with the appropriate concentration of enzyme at 37°C. A 75 μL aliquot was removed after 5, 10, 15, 30, 45, 60, 120, 180, 240, and 300 minute time points, quenched with 10 μL glacial acetic acid and 15 μL of buffer, bringing the final volume to 100 μL. A 40 μL aliquot was directly injected on the analytical RP-HPLC using a gradient of 0% Standard B for 5 min, 36% Standard B in 24 min, 100% Standard B in 11 min. Examples of the HPLC analysis from the chymotrypsin reactions are shown in Figures 5.12-5.16, the trypsin reactions are shown in Figures 5.17-5.22, and the pronase reactions are shown in Figures 5.23-5.28. The acetic acid peak, which elutes from the HPLC column around 3 min, was used as an internal control to account for differences in the injection volumes from the autosampler. The parent peak and the acetic acid peak were integrated using the Empower Pro software. To determine the fraction of full-length peptide remaining at each of the time points, the peak area for the full-length peptide was divided by the peak area at the zero time point. The points were plotted as an average of two runs.

v. Circular dichroism

Studies were done using an Aviv Model 62DS Circular Dichroism Spectrometer. A wavelength scan (185nm-260nm) was done of both **WKWK-scrambled** and **WKWK** in 10 mM sodium phosphate buffer (pH 7.6) at 37°C in a 0.1 cm cell. The peptide concentration was 100 μM in both cases and the scans were done in triplicate.

vi. NMR spectroscopy

NMR experiments were performed at room temperature. Samples were made at a concentration of 1 mM peptide in D_2O buffered to pD 4.0 (uncorrected) with 50 mM NaOAc-d3. Fraction folded determinations were done using the glycine splitting method, which compares the extent of glycine splitting for the desired peptide to the glycine splitting of a fully folded control. In the case of **WKWK-pGturn**, a cyclic version **cycWKWK-pGturn** was synthesized but for the **WKFK-pGturn** the peptide was compared to data that had been previously collected by Butterfield et al.[16] One dimensional NMR spectra were collected using 32K data points and between 8 and 64 scans using 1-3 s presaturation (Figures 5.29, 5.30, 5.31). Two-dimensional TOCSY experiments were used to determine the α-H chemical shifts (Tables 5.6, 5.7, and 5.8) and confirm the results found using the glycine splitting method. The experiments were done with pulse sequences from the Chempack software experiments, and 16-64 scans were done in the first dimension and 64-256 scans in the second dimension. All spectra were analyzed using standard window functions (sinebell and Gaussian). The peak assignments were made using standard methods as described by Wüthrich.[17]

vii. Determination of fraction folded

Fraction folded was determined based on Hα and/or Gly chemical shifts, as compared to random coil and fully folded control peptides. To determine the unfolded chemical shifts, 7-mers were synthesized as unstructured controls and disulfide-bonded cyclic peptides were synthesized for fully folded controls. The chemical shifts of the fully folded state were taken

[16] Butterfield, S. M., The University of North Carolina at Chapel Hill, 2004.
[17] Wüthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons, Inc: New York, 1986.

from the cyclic peptides. The fraction folded at each residue was determined from equation 1,

$$\text{Fraction Folded} = [\delta obs - \delta 0]/[\ \delta 100 - \delta 0], \qquad \text{[Equation 1]}$$

where δobs is the observed Hα chemical shift, δ100 is the Hα chemical shift of the cyclic peptides, and δ0 is the Hα chemical shift of the unfolded 7-mers. The overall fraction folded for the entire peptide was obtained by averaging the fraction folded of resides Val3, Orn8, and Ile10. These residues are in hydrogen bonded positions have been shown to be the most reliable in determining fraction folded.[18] The overall fraction fold was also determined using the extent of Hα glycine splitting observed in the turn residue Gly10 given in equation 2.

$$\text{Fraction Folded} = [\Delta\delta_{\text{Gly Obs}}]/[\ \Delta\delta_{\text{Gly 100}}], \qquad \text{[Equation 2]}$$

where $\Delta\delta_{\text{Gly Obs}}$ is the difference in the glycine Hα chemical shifts of the observed, and $\Delta\delta_{\text{Gly 100}}$ is the difference in the glycine Hα chemical shifts of the cyclic peptides.

viii. Synthesis and purification of PKC-protectides

Peptides were synthesized on either the Creosalus Tetras automated peptide synthesizer. All the peptides were synthesized using standard solid phase peptide synthesis protocols with Fmoc-protected amino acids and Fmoc-PAL-PEG-PS resin. For the addition of the FMOC-PEG_2-OH amino acid (purchased from Novabiochem) only 2 equivalents of the amino acid were used and the stirring time was increased to two hours. The same conditions were used for the addition of FMOC-Lys(ivDde)-OH. Peptides were synthesized on a 0.065 mmol scale and all were capped with an N-terminal acetyl group except for **cycWKWK-PKC**. Before cleaving from the resin, the ivDde protecting group was removed by reacting with 10 mL of 2% hydrazine in DMF (3 x 3 min). The single isomer (6)-carboxyfluorescein was

[18] Syud, F. A., Espinosa, J.F., Gellman, S.H. *J Am Chem Soc* **1999**, *121*, 11577–11578.

then coupled to the free lysine side chain overnight and in the dark using HOBT/HBTU activator solution (4 equivalents of each) and 4 equivalents of DIPEA. All the peptides were cleaved and deprotected by treating the resin with a (95:2.5:2.5) TFA:Triisopropylsilane:Water mixture and bubbling with N_2 for 3 hours. The peptides were isolated by first reducing the volume of TFA in a stream of nitrogen, then precipitating with cold diethyl ether, followed by extraction using deionized water. The resulting peptide solution was frozen and lyophilized. The peptides were twice purified using a Waters semi-preparative HPLC system with a reverse-phase C18 column. The peptides were eluted using Standard A (95% Water, 5% acetonitrile, 0.1% TFA) and Standard B (95% acetonitrile, 5% Water, 0.1% TFA). The masses of the peptides were confirmed using ESI-TOF mass spectrometry.

ix. Cyclization reaction for cyc-WKWK-PKC

The cyclization reaction was a modified procedure from Ivanov et al.[15] Before cleaving the **cyc-WKWK-PKC** peptide from the resin, the N-terminal FMOC group was deprotected and the resin was swelled with DCM. In a separate scintillation vial, bromoacetic acid (8 eq., 0.52 mmol), diisopropyl carbodiimide (4 eq., 0.26 mmol), and DMF (4 mL) were stirred at 0°C for 30 min. This solution was then added to the resin along with DIPEA (4 eq., 0.26 mmol) and the reaction proceeded for 1.5 hours or until the reaction is completed as determined by the Kaiser test. The resin was then thoroughly rinsed with DMF and DCM, dried, and cleaved for 3 hours with the standard TFA cocktail as described above. The peptides were isolated by precipitating with cold diethyl ether and extraction with deionized water. The resulting peptide solution was frozen and lyophilized. The peptide was then redissolved in 10 mL of water (minimal DMSO may be added to help dissolve) and 6

equivalents of triethylamine (0.39 mmol) was added to initiate cyclization. The reaction was stirred overnight and then lyophilized to dryness for subsequent purification using HPLC.

Table 5.5. Mass Spectrometry Data Obtained for Peptides Studied in this Work

Peptide	Expected mass (Da)	Observed mass (Da)
WKWK-scrambled	1567.9	1568.1
WKFK	1528.9	1528.5
WKFK-pGturn	1511.9	1512.2
WKWK	1567.9	1568.0
WKWK-pGturn	1551.0	1551.0
WKWKscram-Peg_2-PKC	3887.2	3887.2
WKWK-Peg_2-PKC	3887.2	3887.9
TrpPkt-Peg_2-PKC	3872.2	3872.3

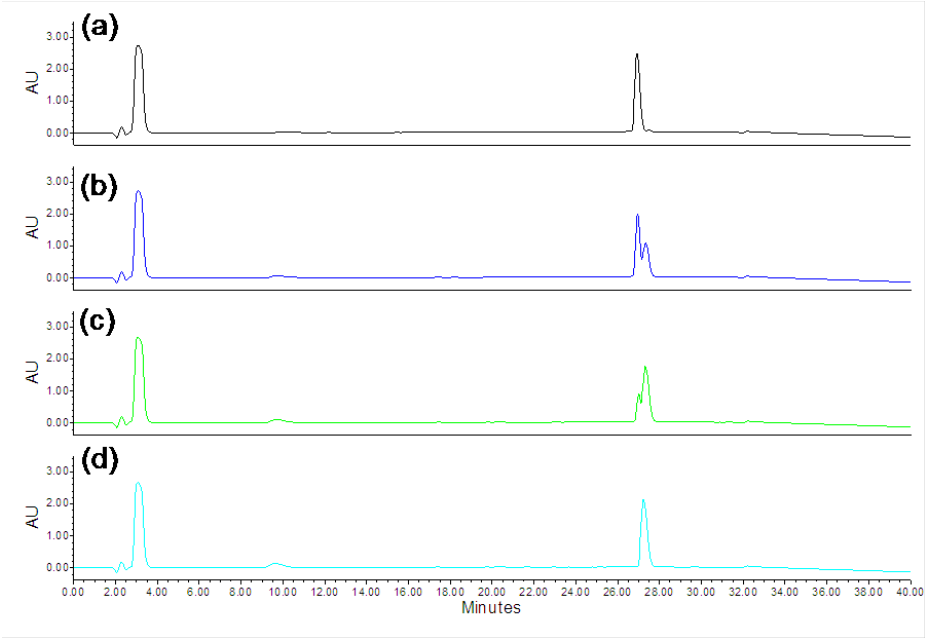


Figure 5.12. RP-HPLC analysis of **WKWK-scrambled** peptide incubated with α-chymotrypsin for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 5 m, (c) 10 m, and (d) 15 m. Absorbance was monitored at 220 nm.

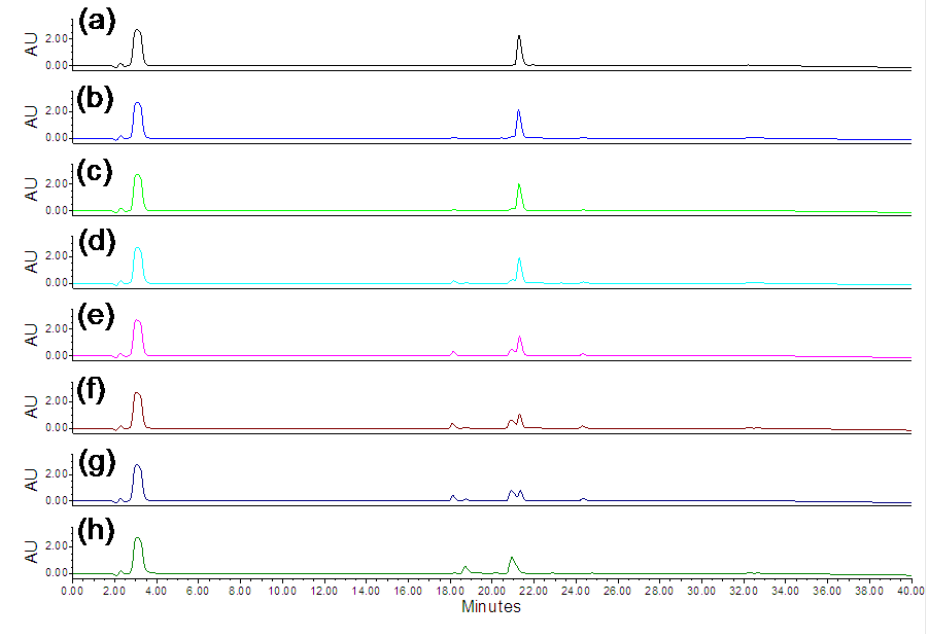


Figure 5.13. RP-HPLC analysis of **WKFK** peptide incubated with α-chymotrypsin for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 5 m, (c) 10 m, (d) 15 m, (e) 30 m, (f) 45 m, (g) 1 h, and (h) 5 h. Absorbance was monitored at 220 nm.

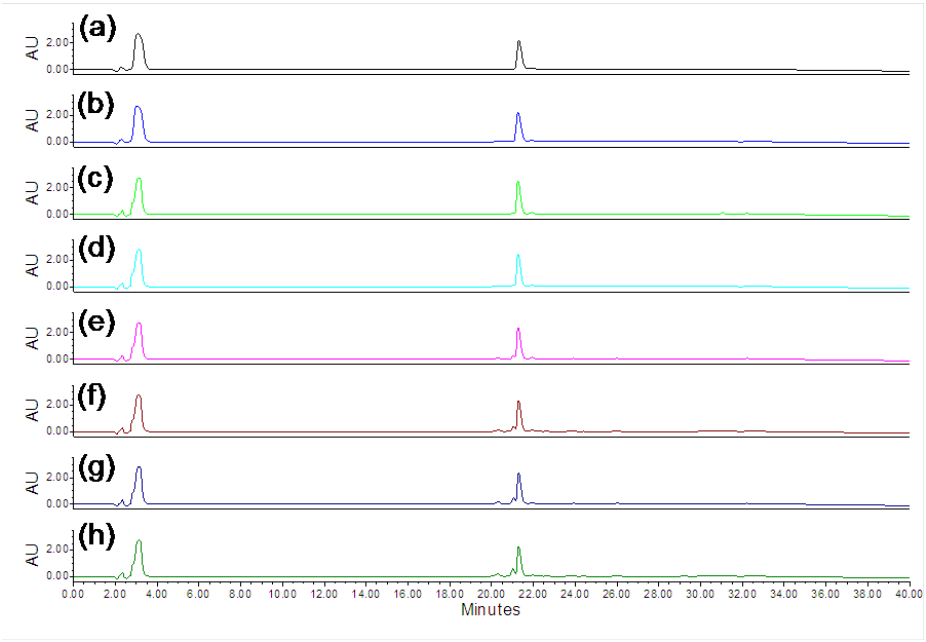


Figure 5.14. RP-HPLC analysis of **WKFK-pGturn** peptide incubated with α-chymotrypsin for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 15 m, (c) 30 m, (d) 1 h, (e) 2 h, (f) 3 h, (g) 4 h, and (h) 5 h. Absorbance was monitored at 220 nm.

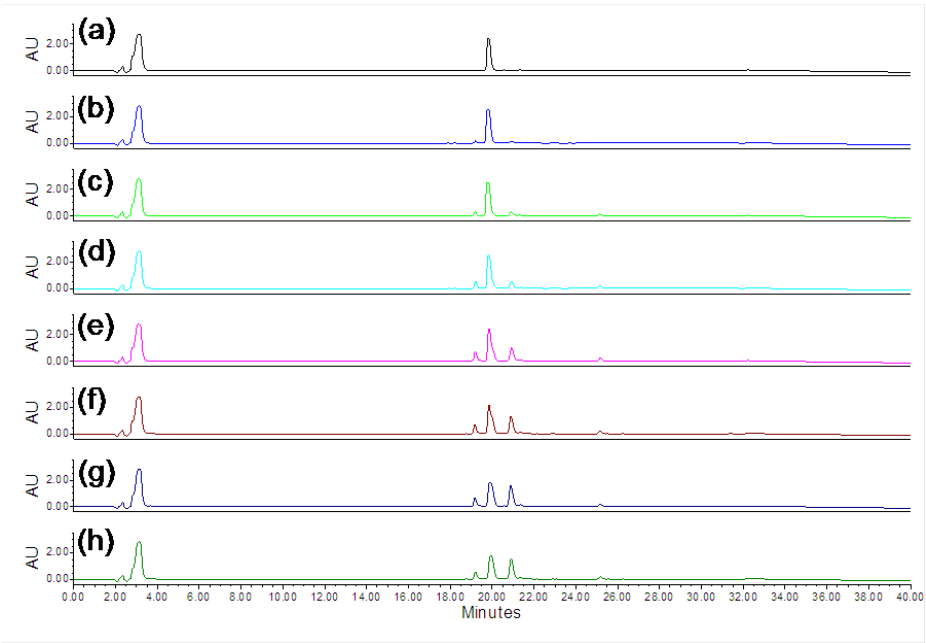


Figure 5.15. RP-HPLC analysis of **WKWK** peptide incubated with α-chymotrypsin for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 15 m, (c) 30 m, (d) 1 h, (e) 2 h, (f) 3 h, (g) 4 h, and (h) 5 h. Absorbance was monitored at 220 nm.

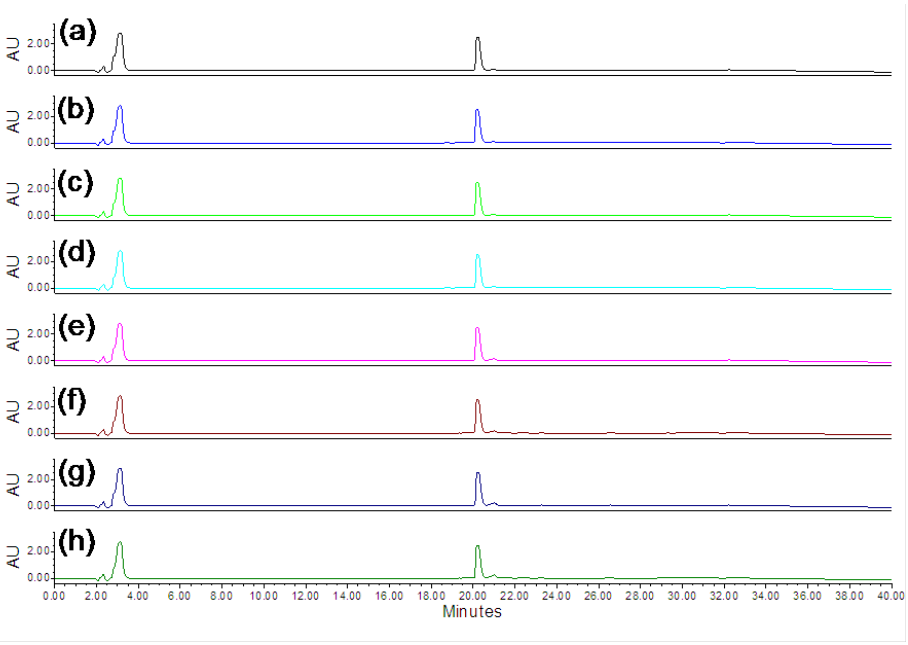


Figure 5.16. RP-HPLC analysis of **WKWK-pGturn** peptide incubated with α-chymotrypsin for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 15 m, (c) 30 m, (d) 1 h, (e) 2 h, (f) 3 h, (g) 4 h, and (h) 5 h. Absorbance was monitored at 220 nm.

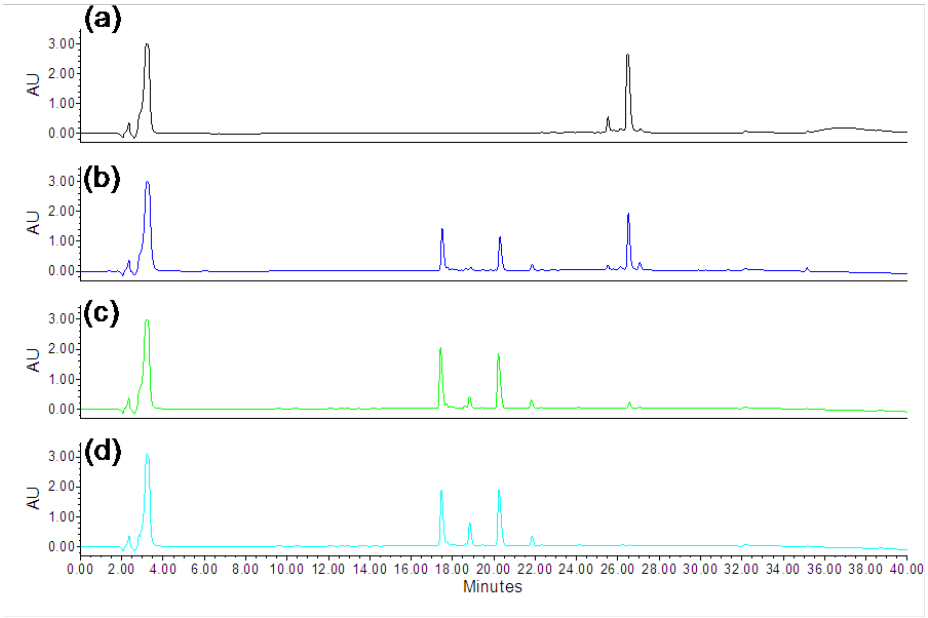


Figure 5.17. RP-HPLC analysis of **WKWK-scrambled** peptide incubated with trypsin for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 5 m, (c) 10 m, and (d) 15 m. Absorbance was monitored at 220 nm.

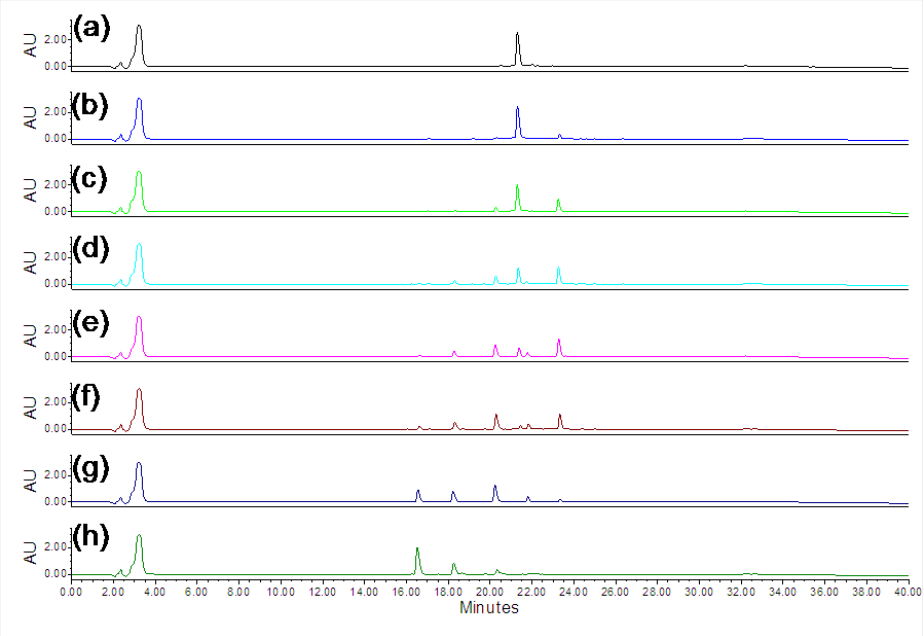


Figure 5.18. RP-HPLC analysis of **WKFK** peptide incubated with trypsin for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 5 m, (c) 15 m, (d) 30 m, (e) 45 m, (f) 1 h, (g) 2 h, and (h) 5 h. Absorbance was monitored at 220 nm.

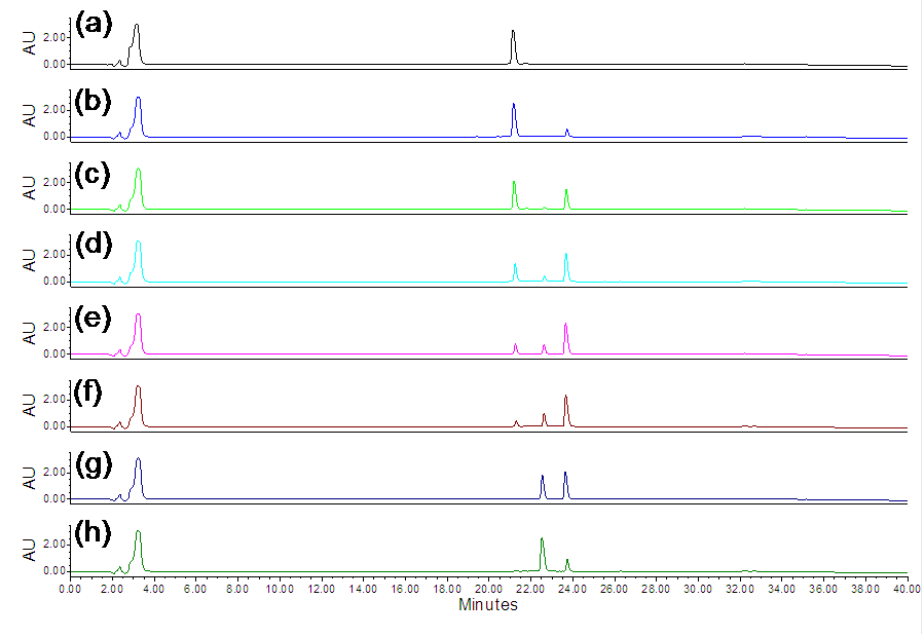


Figure 5.19. RP-HPLC analysis of **WKFK-pGturn** peptide incubated with trypsin for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 5 m, (c) 15 m, (d) 30 m, (e) 45 m, (f) 1 h, (g) 2 h, and (h) 5 h. Absorbance was monitored at 220 nm.

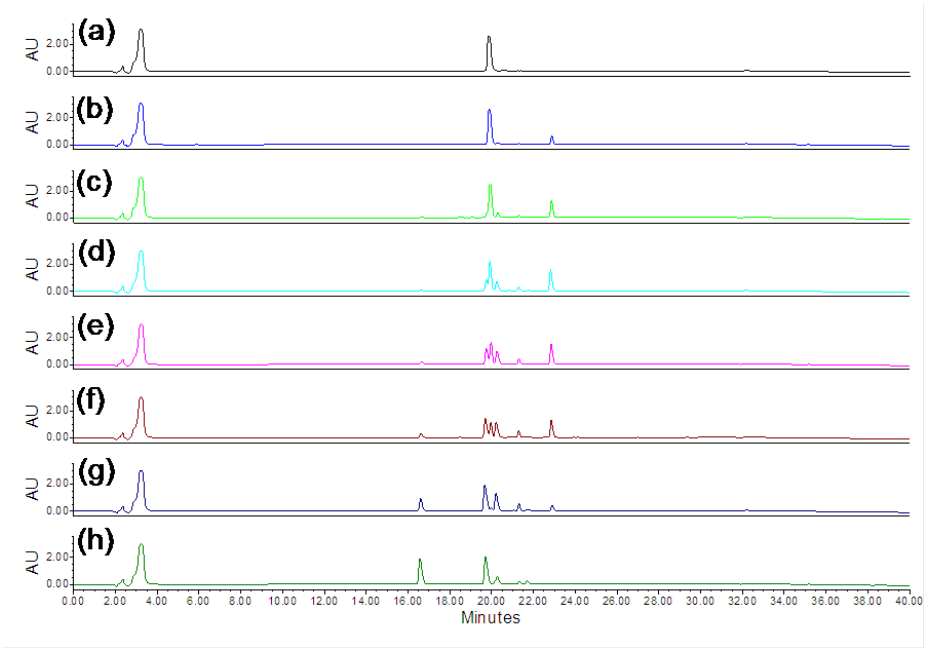


Figure 5.20. RP-HPLC analysis of **WKWK** peptide incubated with trypsin for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 5 m, (c) 15 m, (d) 30 m, (e) 45 m, (f) 1 h, (g) 2 h, and (h) 5 h. Absorbance was monitored at 220 nm.

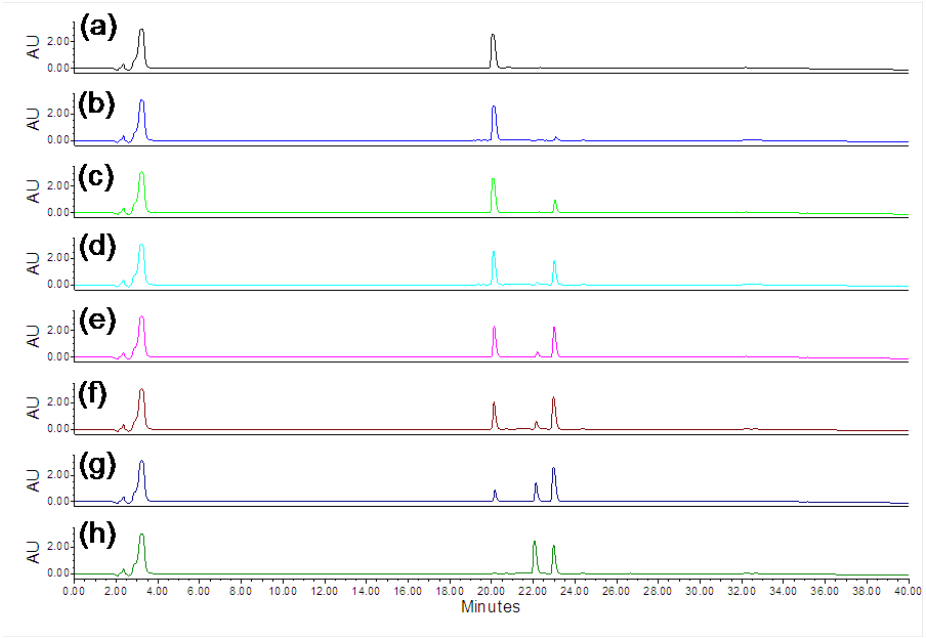


Figure 5.21. RP-HPLC analysis of **WKWK-pGturn** peptide incubated with trypsin for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 5 m, (c) 15 m, (d) 30 m, (e) 45 m, (f) 1 h, (g) 2 h, and (h) 5 h. Absorbance was monitored at 220 nm.

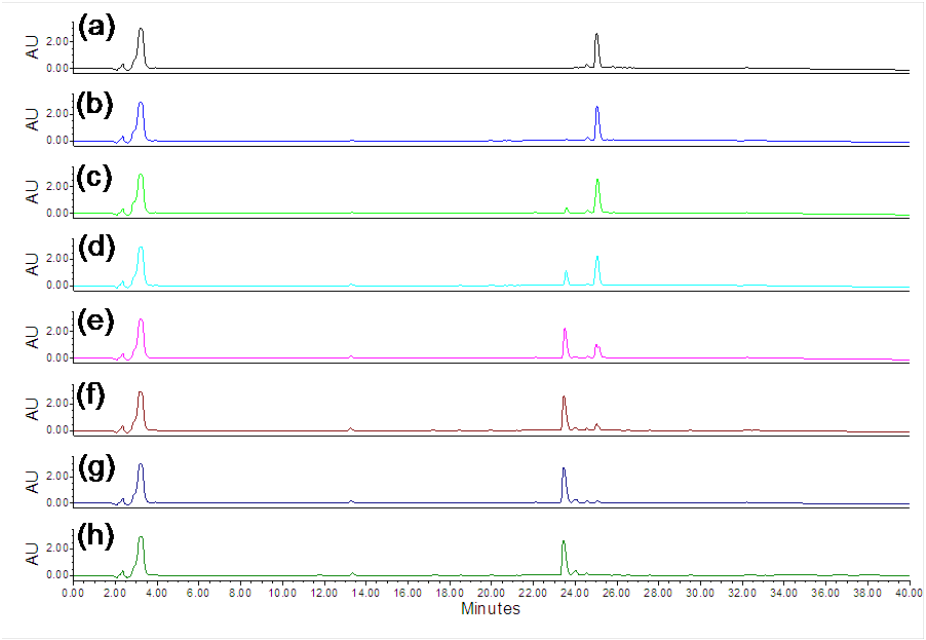


Figure 5.22. RP-HPLC analysis of **Trp pocket** peptide incubated with trypsin for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 15 m, (c) 30 m, (d) 1 h, (e) 2 h, (f) 3 h, (g) 4 h, and (h) 5 h. Absorbance was monitored at 220 nm.

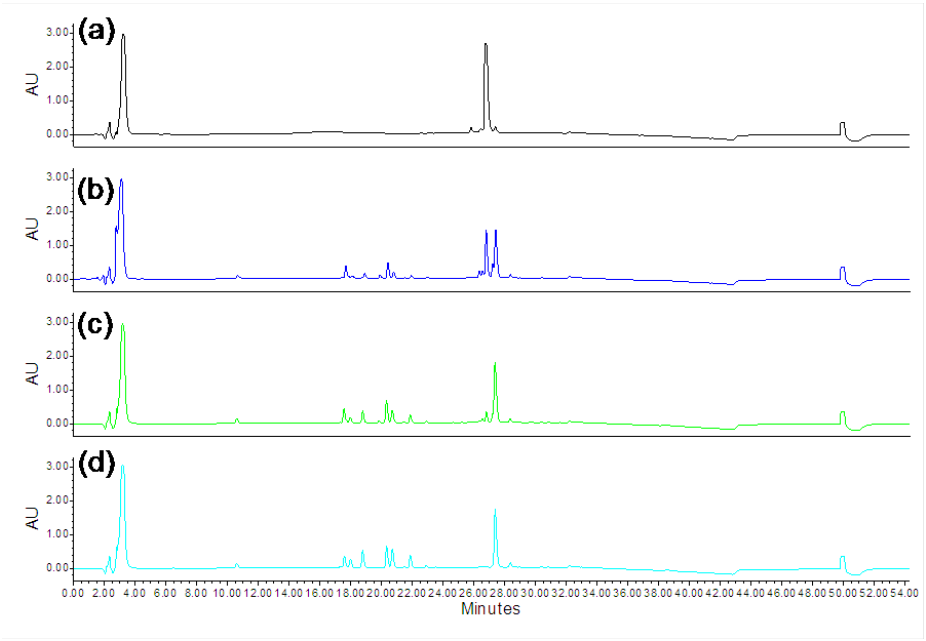


Figure 5.23. RP-HPLC analysis of **WKWK-scrambled** peptide incubated with pronase for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 5 m, (c) 10 m, and (d) 15 m. Absorbance was monitored at 220 nm.

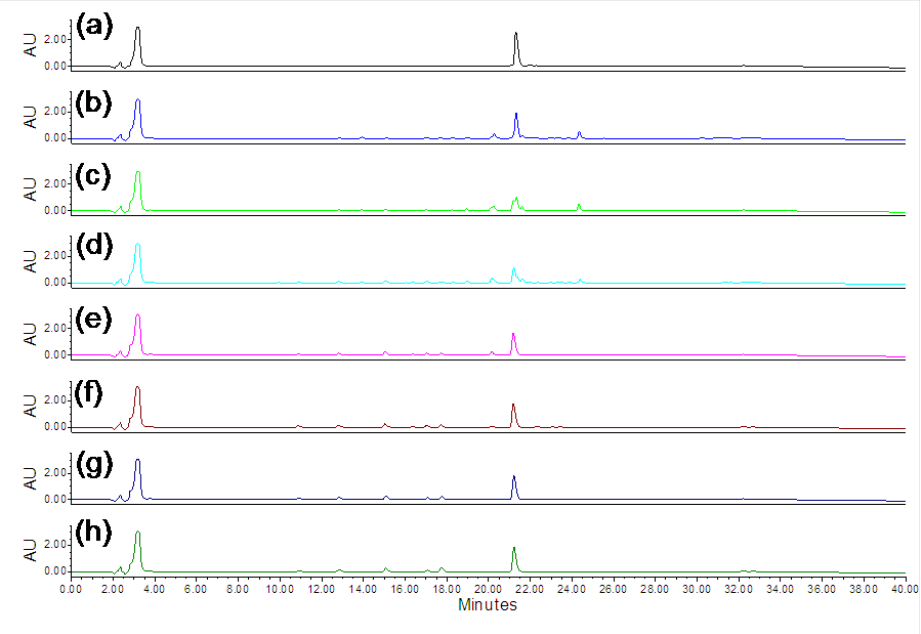


Figure 5.24. RP-HPLC analysis of **WKFK** peptide incubated with pronase for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 5 m, (c) 10 m, (d) 15 m, (e) 30 m, (f) 45 m, (g) 1 h, and (h) 5 h. Absorbance was monitored at 220 nm.

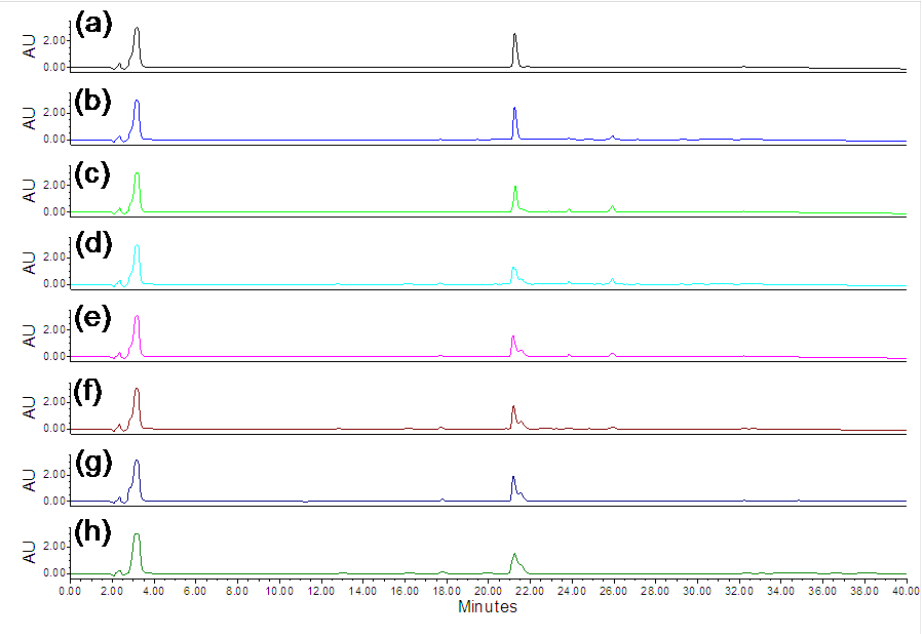


Figure 5.25. RP-HPLC analysis of **WKFK-pGturn** peptide incubated with pronase for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 5 m, (c) 15 m, (d) 30 m, (e) 45 m, (f) 1 h, (g) 2 h, and (h) 5 h. Absorbance was monitored at 220 nm.

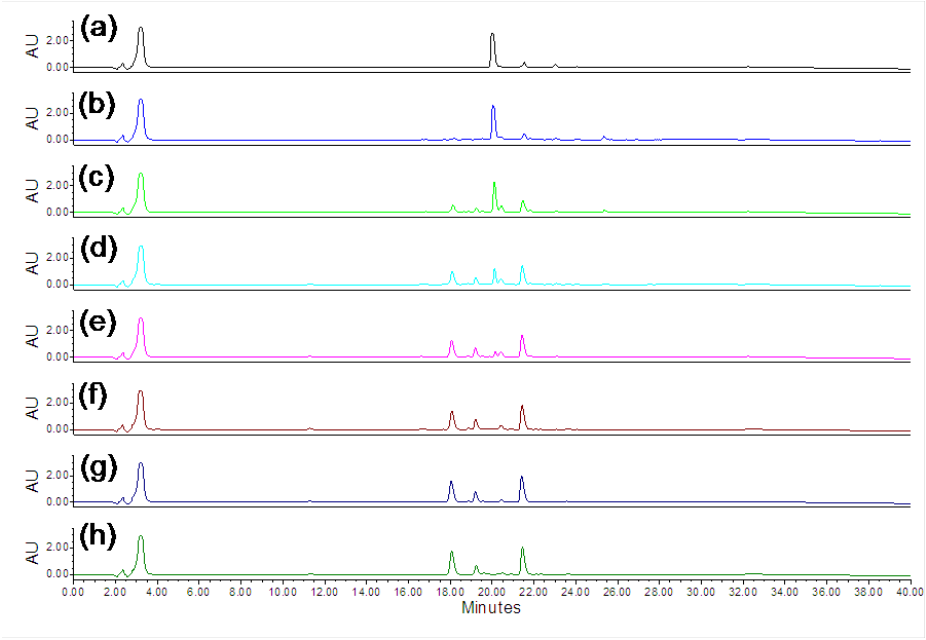


Figure 5.26. RP-HPLC analysis of **WKWK** peptide incubated with pronase for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 5 m, (c) 15 m, (d) 30 m, (e) 45 m, (f) 1 h, (g) 2 h, and (h) 5 h. Absorbance was monitored at 220 nm.

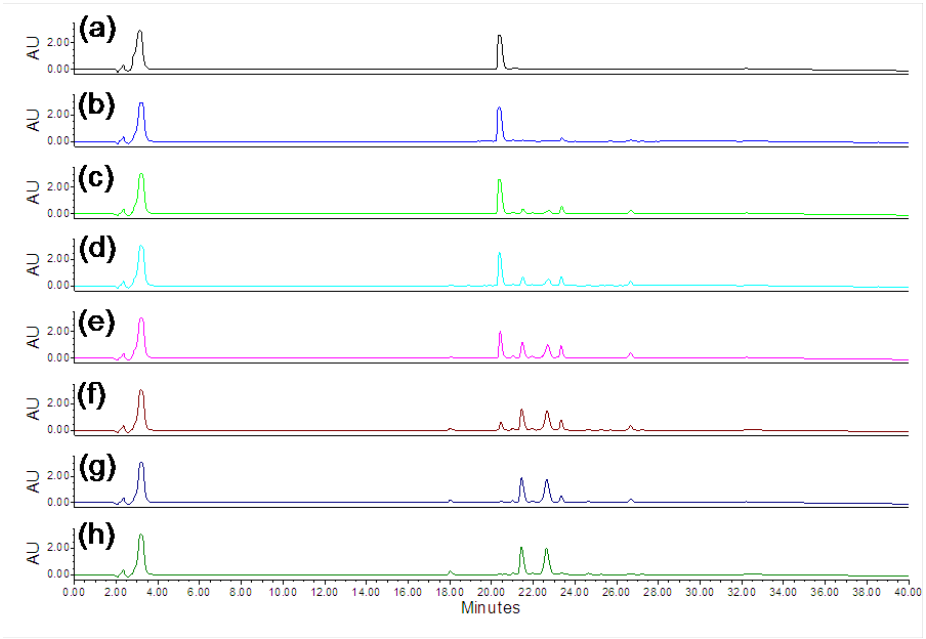


Figure 5.27. RP-HPLC analysis of **WKWK-pGturn** peptide incubated with pronase for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 5 m, (c) 15 m, (d) 30 m, (e) 1 h, (f) 2 h, (g) 3 h, and (h) 5 h. Absorbance was monitored at 220 nm.

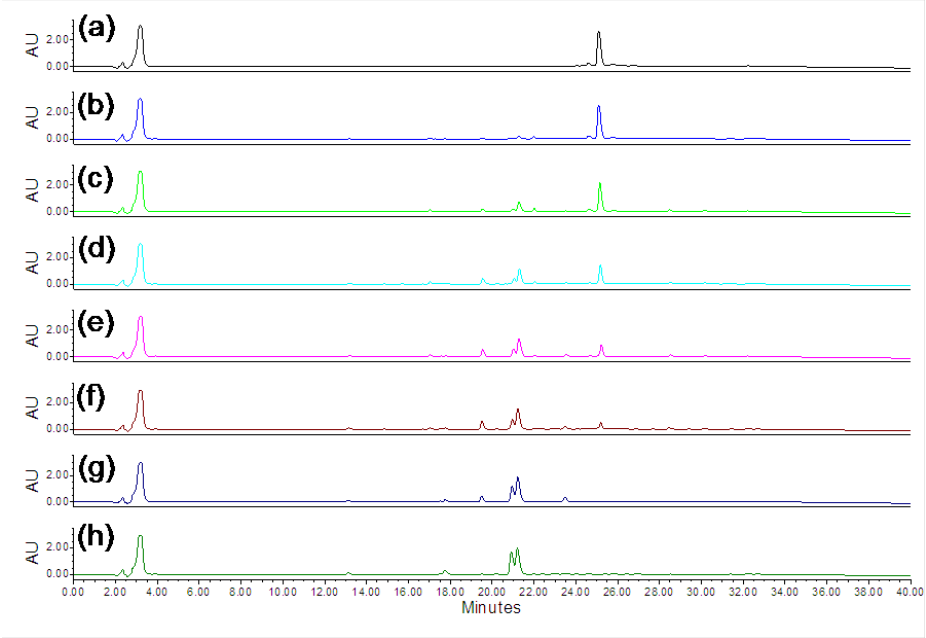


Figure 5.28. RP-HPLC analysis of **Trp pocket** peptide incubated with pronase for various amounts of time. Reactions were quenched with acetic acid at (a) 0 m, (b) 5 m, (c) 15 m, (d) 30 m, (e) 45 m, (f) 1 h, (g) 2 h, and (h) 5 h. Absorbance was monitored at 220 nm.

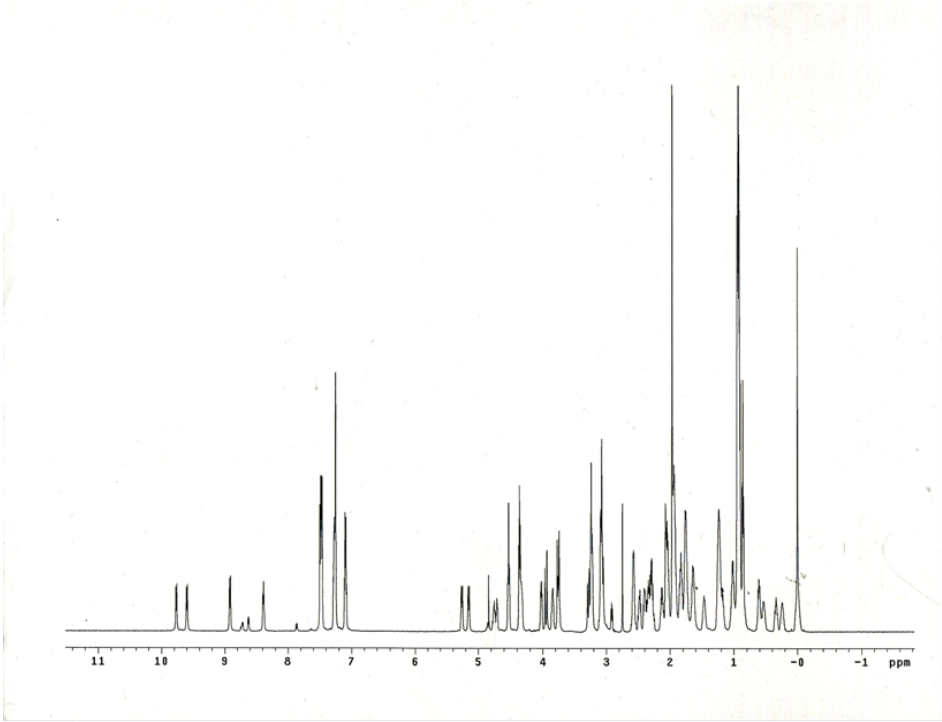


Figure 5.29. ^{1}HNMR of **WKWK-pGturn** peptide: Ac-Arg-Trp-Val-Lys-Val-$_{D}$Pro-Gly-Orn-Trp-Ile-Lys-Gln-NH_2

Table 5.6. Proton Chemical Shift Assignments for **WKWK-pGturn** peptide

Residue	α	β	γ	δ	ε
Arg	4.46	1.75	1.62	3.17	
Trp	5.17	3.2, 3.04			
Val	4.65	1.97	0.86		
Lys	4.27	1.15, 0.14	0.47	0.95	2.31
Val	4.46	1.75	0.83		
$_{D}$Pro	4.30	2.29, 1.92	1.92	3.74	
Gly	4.30, 3.71				
Orn	4.70	1.81	1.72	3.01	
Trp	5.08	3.2, 3.01			
Ile	4.70	1.81	0.85	0.85	
Lys	3.95	1.15, 0.21	0.55	0.55	2.52
Gln	4.30	1.9	2.25		

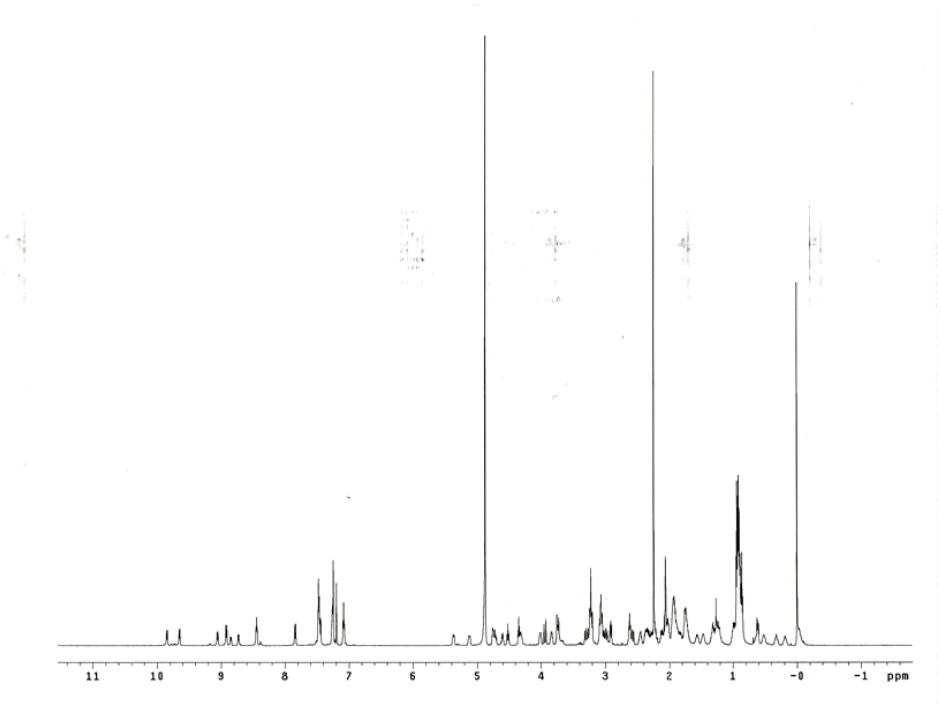


Figure 5.30. ^{1}HNMR of **cycWKWK-pGturn** peptide: Ac-Cys-Arg-Trp-Val-Lys-Val-$_{D}$Pro-Gly-Orn-Phe-Ile-Lys-Gln-Cys-NH_2

Table 5.7. Proton Chemical Shift Assignments for **cycWKWK-pGturn** peptide

Residue	α	β	γ	δ	ε
Cys	5.10	2.99			
Arg	4.68	1.83	1.60	3.18	
Trp	5.33	3.21, 2.99			
Val	4.68	1.83	0.88		
Lys	4.26	1.18	0.48	-0.05	2.29
Val	4.48	1.89	0.85		
$_{D}$Pro	4.36	2.29, 1.92	1.92	3.74	
Gly	4.31, 3.74				
Orn	4.68	1.89	1.7	3.04	
Trp	5.16	3.2, 3.02			
Ile	4.68	1.89	0.88	0.88	
Lys	4.05	1.21	0.59	0.23	2.56
Gln	4.56	1.81	2.20		
Cys	5.20	2.96, 2.36			

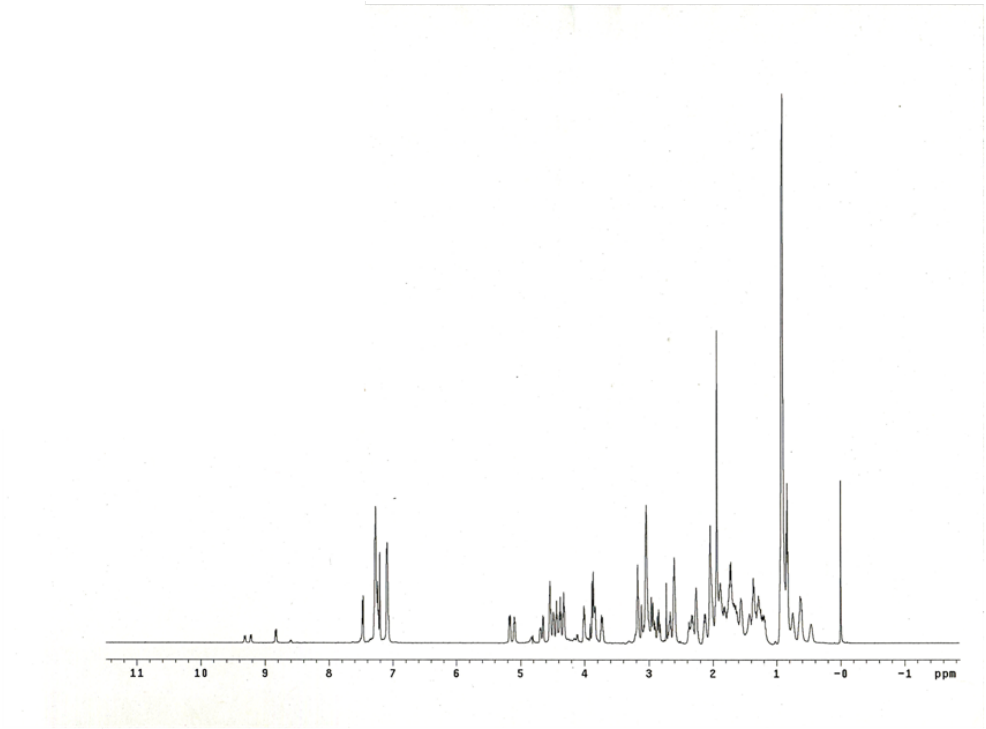


Figure 5.31. ^{1}HNMR of **WKFK-pGturn** peptide: Ac-Arg-Trp-Val-Lys-Val-$_{D}$Pro-Gly-Orn-Phe-Ile-Lys-Gln-NH_2

Table 5.8. Proton Chemical Shift Assignments for **WKFK-pGturn** peptide

Residue	α	β	γ	δ	ε
Arg	4.44	1.7	1.57	3.13	
Trp	5.13	3.01, 2.86			
Val	4.59	1.84	0.86		
Lys	4.44	1.60, 0.58	0.86	1.30	2.59
Val	4.49	1.97	0.87		
$_{D}$Pro	4.32	2.27, 1.95	1.95	3.74	
Gly	4.33, 3.73				
Orn	4.63	1.84	1.72	3.0	
Phe	5.06	3.01			
Ile	4.59	1.84	0.86	0.86	
Lys	3.96	1.27, 0.47	0.66	0.66	2.56
Gln	4.27	1.84	2.23		

BIBLIOGRAPHY

Aertgeerts, K.; Ye, S.; Tennant, M. G.; Kraus, M. L.; Rogers, J.; Sang, B. C.; Skene, R. J.; Webb, D. R.; Prasad, G. S., Crystal structure of human dipeptidyl peptidase IV in complex with a decapeptide reveals details on substrate specificity and tetrahedral intermediate formation. *Protein Sci* **2004,** 13, (2), 412-21.

Ahn, D. R.; Yu, J., Library construction of neomycin-dipeptide heteroconjugates and selection against RRE RNA. *Bioorg Med Chem* **2005,** 13, (4), 1177-83.

Allers, J.; Shamoo, Y., Structure-based analysis of protein-RNA interactions using the program ENTANGLE. *J Mol Biol* **2001,** 311, (1), 75-86.

Athanassiou, Z.; Dias, R. L.; Moehle, K.; Dobson, N.; Varani, G.; Robinson, J. A., Structural mimicry of retroviral tat proteins by constrained beta-hairpin peptidomimetics: ligands with high affinity and selectivity for viral TAR RNA regulatory elements. *J Am Chem Soc* **2004,** 126, (22), 6906-13.

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.

Austin, R. J.; Xia, T.; Ren, J.; Takahashi, T. T.; Roberts, R. W., Designed arginine-rich RNA-binding peptides with picomolar affinity. *J Am Chem Soc* **2002,** 124, (37), 10966-7.

Auweter, S. D.; Oberstrass, F. C.; Allain, F. H., Sequence-specific binding of single-stranded RNA: is there a code for recognition? *Nucleic Acids Res* **2006,** 34, (17), 4943-59.

Bloomfield, V. A.; Crothers, D. M.; Tinoco, J. I., *Nucleic Acids: Structures, Properties, and Functions*. University Science Books: Sausalito, CA, 2000.

Brandstetter, H.; Kim, J. S.; Groll, M.; Huber, R., Crystal structure of the tricorn protease reveals a protein disassembly line. *Nature* **2001,** 414, (6862), 466-70.

Bugaut, A.; Toulme, J. J.; Rayner, B., SELEX and dynamic combinatorial chemistry interplay for the selection of conjugated RNA aptamers. *Org Biomol Chem* **2006,** 4, (22), 4082-8.

Burns, V. A.; Bobay, B. G.; Basso, A.; Cavanagh, J.; Melander, C., Targeting RNA with cysteine-constrained peptides. *Bioorg Med Chem Lett* **2008,** 18, (2), 565-7.

Butterfield, S. M. Aromatic Interactions In Folding And Recognition In Peptide Model Systems. The University of North Carolina at Chapel Hill, Chapel Hill, NC, 2004.

Butterfield, S. M.; Cooper, W. J.; Waters, M. L., Minimalist protein design: a beta-hairpin peptide that binds ssDNA. *J Am Chem Soc* **2005,** 127, (1), 24-5.

Butterfield, S. M.; Sweeney, M. M.; Waters, M. L., The recognition of nucleotides with model beta-hairpin receptors: investigation of critical contacts and nucleotide selectivity. *J Org Chem* **2005,** 70, (4), 1105-14.

Butterfield, S. M.; Sweeney, M. M.; Waters, M. L., The recognition of nucleotides with model beta-hairpin receptors: investigation of critical contacts and nucleotide selectivity. *J Org Chem* **2005,** 70, (4), 1105-14.

Butterfield, S. M.; Waters, M. L., A designed beta-hairpin peptide for molecular recognition of ATP in water. *J Am Chem Soc* **2003,** 125, (32), 9580-1.

Calnan, B. J.; Tidor, B.; Biancalana, S.; Hudson, D.; Frankel, A. D., Arginine-mediated RNA recognition: the arginine fork. *Science* **1991,** 252, (5010), 1167-71.

Campbell, M. K.; Farrell, S. O., Biochemistry. In 6th ed.; Thompson Brooks/Cole: Belmont, CA, 2009.

Carlson, C. B.; Beal, P. A., Solid-phase synthesis of acridine-based threading intercalator peptides. *Bioorg Med Chem Lett* **2000,** 10, (17), 1979-82.

Carlson, C. B.; Spanggord, R. J.; Beal, P. A., Selection of small-molecule mediators of the RNA regulation of PKR, the RNA-dependent protein kinase. *Chembiochem* **2002,** 3, (9), 859-65.

Carlson, C. B.; Stephens, O. M.; Beal, P. A., Recognition of double-stranded RNA by proteins and small molecules. *Biopolymers* **2003,** 70, (1), 86-102.

Carlson, C. B.; Vuyisich, M.; Gooch, B. D.; Beal, P. A., Preferred RNA binding sites for a threading intercalator revealed by in vitro evolution. *Chem Biol* **2003,** 10, (7), 663-72.

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.

Chen, L.; Frankel, A. D., An RNA-binding peptide from bovine immunodeficiency virus Tat protein recognizes an unusual RNA structure. *Biochemistry* **1994,** 33, (9), 2708-15.

Chen, L.; Frankel, A. D., A peptide interaction in the major groove of RNA resembles protein interactions in the minor groove of DNA. *Proc Natl Acad Sci U S A* **1995,** 92, (11), 5077-81.

Cheng, A. C.; Calabro, V.; Frankel, A. D., Design of RNA-binding proteins and ligands. *Curr Opin Struct Biol* **2001,** 11, (4), 478-84.

Cheng, Z.; Campbell, R. E., Assessing the structural stability of designed beta-hairpin peptides in the cytoplasm of live cells. *Chembiochem* **2006,** 7, (8), 1147-50.

Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A., Tryptophan zippers: stable, monomeric beta -hairpins. *Proc Natl Acad Sci U S A* **2001,** 98, (10), 5578-83.

Cooper, W. J.; Waters, M. L., Turn residues in beta-hairpin peptides as points for covalent modification. *Organic Letters* **2005,** 7, (18), 3825-3828.

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.

De Guzman, R. N.; Wu, Z. R.; Stalling, C. C.; Pappalardo, L.; Borer, P. N.; Summers, M. F., Structure of the HIV-1 nucleocapsid protein bound to the SL3 Psi-RNA recognition element. *Science* **1998,** 279, (5349), 384-388.

Deo, R. C.; Bonanno, J. B.; Sonenberg, N.; Burley, S. K., Recognition of polyadenylate RNA by the poly(A)-binding protein. *Cell* **1999,** 98, (6), 835-45.

D'Souza, V.; Summers, M. F., Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus. *Nature* **2004,** 431, (7008), 586-90.

Fersht, A. R.; Shi, J. P.; Knill-Jones, J.; Lowe, D. M.; Wilkinson, A. J.; Blow, D. M.; Brick, P.; Carter, P.; Waye, M. M.; Winter, G., Hydrogen bonding and biological specificity analysed by protein engineering. *Nature* **1985,** 314, (6008), 235-8.

Findlay, J. W.; Dillard, R. F., Appropriate calibration curve fitting in ligand binding assays. *AAPS J* **2007,** 9, (2), E260-7.

Fisher, A. G.; Feinberg, M. B.; Josephs, S. F.; Harper, M. E.; Marselle, L. M.; Reyes, G.; Gonda, M. A.; Aldovini, A.; Debouk, C.; Gallo, R. C.; et al., The trans-activator gene of HTLV-III is essential for virus replication. *Nature* **1986,** 320, (6060), 367-71.

Frackenpohl, J.; Arvidsson, P. I.; Schreiber, J. V.; Seebach, D., The outstanding biological stability of beta- and gamma-peptides toward proteolytic enzymes: an in vitro investigation with fifteen peptidases. *Chembiochem* **2001,** 2, (6), 445-55.

Fulop, V.; Bocskei, Z.; Polgar, L., Prolyl oligopeptidase: an unusual beta-propeller domain regulates proteolysis. *Cell* **1998,** 94, (2), 161-70.

Galicia-Vazquez, 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), 180-8.

Gallegos, L. L.; Kunkel, M. T.; Newton, A. C., Targeting protein kinase C activity reporter to discrete intracellular regions reveals spatiotemporal differences in agonist-dependent signaling. *J Biol Chem* **2006,** 281, (41), 30947-56.

Gareiss, P. C.; Sobczak, K.; McNaughton, B. R.; Palde, P. B.; Thornton, C. A.; Miller, B. L., Dynamic combinatorial selection of molecules capable of inhibiting the (CUG) repeat RNA-MBNL1 interaction in vitro: discovery of lead compounds targeting myotonic dystrophy (DM1). *J Am Chem Soc* **2008,** 130, (48), 16254-61.

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. *Organic Letters* **2003,** 5, (20), 3563-3565.

Gooch, B. D.; Beal, P. A., Recognition of duplex RNA by helix-threading peptides. *J Am Chem Soc* **2004,** 126, (34), 10603-10.

Gooch, B. D.; Krishnamurthy, M.; Shadid, M.; Beal, P. A., Binding of helix-threading peptides to E. coli 16S ribosomal RNA and inhibition of the S15-16S complex. *Chembiochem* **2005,** 6, (12), 2247-54.

Griffey, R. H.; Hofstadler, S. A.; Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T., Determinants of aminoglycoside-binding specificity for rRNA by using mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* **1999,** 96, (18), 10129-10133.

Griffiths-Jones, S. R.; Maynard, A. J.; Searle, M. S., Dissecting the stability of a beta-hairpin peptide that folds in water: NMR and molecular dynamics analysis of the beta-turn and beta-strand contributions to folding. *J Mol Biol* **1999,** 292, (5), 1051-69.

Guhaniyogi, J.; Sohar, I.; Das, K.; Stock, A. M.; Lobel, P., Crystal structure and autoactivation pathway of the precursor form of human tripeptidyl-peptidase 1, the enzyme deficient in late infantile ceroid lipofuscinosis. *J Biol Chem* **2009,** 284, (6), 3985-97.

Hamamoto, K.; Kida, Y.; Zhang, Y.; Shimizu, T.; Kuwano, K., Antimicrobial activity and stability to proteolysis of small linear cationic peptides with D-amino acid substitutions. *Microbiol Immunol* **2002,** 46, (11), 741-9.

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), 3548-53.

Handa, N.; Nureki, O.; Kurimoto, K.; Kim, I.; Sakamoto, H.; Shimura, Y.; Muto, Y.; Yokoyama, S., Structural basis for recognition of the tra mRNA precursor by the Sex-lethal protein. *Nature* **1999,** 398, (6728), 579-85.

Hermann, T., Strategies for the Design of Drugs Targeting RNA and RNA-Protein Complexes. *Angew Chem Int Ed Engl* **2000,** 39, (11), 1890-1904.
Hermann, T.; Patel, D. J., RNA bulges as architectural and recognition motifs. *Structure* **2000,** 8, (3), R47-54.

Hook, D. F.; Bindschadler, P.; Mahajan, Y. R.; Sebesta, R.; Kast, P.; Seebach, D., The proteolytic stability of 'designed' beta-peptides containing alpha-peptide-bond mimics and of mixed alpha,beta-peptides: application to the construction of MHC-binding peptides. *Chem Biodivers* **2005,** 2, (5), 591-632.

Hughes, R. M.; Waters, M. L., Influence of N-methylation on a cation-pi interaction produces a remarkably stable beta-hairpin peptide. *J Am Chem Soc* **2005,** 127, (18), 6518-9.

Hughes, R. M.; Waters, M. L., Model systems for beta-hairpins and beta-sheets. *Curr Opin Struct Biol* **2006,** 16, (4), 514-24.

Hutchinson, E. G.; Thornton, J. M., A revised set of potentials for beta-turn formation in proteins. *Protein Sci* **1994,** 3, (12), 2207-16.

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.

Hyun, S.; Kim, H. J.; Lee, N. J.; Lee, K. H.; Lee, Y.; Ahn, D. R.; Kim, K.; Jeong, S.; Yu, J., Alpha-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), 4514-5.

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

Ivanov, B.; Grzesik, W.; Robey, F. A., Synthesis and use of a new bromoacetyl-derivatized heterotrifunctional amino acid for conjugation of cyclic RGD-containing peptides derived from human bone sialoprotein. *Bioconjug Chem* **1995,** 6, (3), 269-77.

Jimenez Bueno, G.; Klimkait, T.; Gilbert, I. H.; Simons, C., Solid-phase synthesis of diamine and polyamine amino acid derivatives as HIV-1 tat-TAR binding inhibitors. *Bioorg Med Chem* **2003,** 11, (1), 87-94.

Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I., Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal Biochem* **1970,** 34, (2), 595-8.

Kaiser, M.; Sainlos, M.; Lehn, J. M.; Bombard, S.; Teulade-Fichou, M. P., Aminoglycoside-quinacridine conjugates: Towards recognition of the P6.1 element of telomerase RNA. *Chembiochem* **2006,** 7, (2), 321-329.

Karan, C.; Miller, B. L., RNA-selective coordination complexes identified via dynamic combinatorial chemistry. *J Am Chem Soc* **2001,** 123, (30), 7455-6.

Kawakami, J.; Sugimoto, N.; Tokitoh, H.; Tanabe, Y., A novel stable RNA pentaloop that interacts specifically with a motif peptide of lambda-N protein. *Nucleosides Nucleotides Nucleic Acids* **2006,** 25, (4-6), 397-416.

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

Klosterman, P. S.; Hendrix, D. K.; Tamura, M.; Holbrook, S. R.; Brenner, S. E., Three-dimensional motifs from the SCOR, structural classification of RNA database: extruded strands, base triples, tetraloops and U-turns. *Nucleic Acids Res* **2004,** 32, (8), 2342-52.

Krishnamurthy, M.; Gooch, B. D.; Beal, P. A., Peptide quinoline conjugates: a new class of RNA-binding molecules. *Org Lett* **2004,** 6, (1), 63-6.

Krishnamurthy, M.; Gooch, B. D.; Beal, P. A., RNA binding and thiolytic stability of a quinoline-containing helix-threading peptide. *Org Biomol Chem* **2006,** 4, (4), 639-45.

Krishnamurthy, M.; Simon, K.; Orendt, A. M.; Beal, P. A., Macrocyclic helix-threading peptides for targeting RNA. *Angew Chem Int Ed Engl* **2007,** 46, (37), 7044-7.
Lam, K. S.; Lebl, M.; Krchnak, V., The "One-Bead-One-Compound" Combinatorial Library Method. *Chem Rev* **1997,** 97, (2), 411-448.

Lawton, G. R.; Appella, D. H., Nonionic side chains modulate the affinity and specificity of binding between functionalized polyamines and structured RNA. *J Am Chem Soc* **2004,** 126, (40), 12762-3.

Lee, J. K.; Kwon, M. Y.; Lee, K. H.; Jeong, S. J.; Hyun, S.; Shin, K. J.; Yu, J. H., An approach to enhance specificity against RNA targets using heteroconjugates of aminoglycosides and chloramphenicol (or linezolid). *Journal of the American Chemical Society* **2004,** 126, (7), 1956-1957.

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 Engl* **2007,** 46, 1-5.

Leeper, T. C.; Athanassiou, Z.; Dias, R. L.; Robinson, J. A.; Varani, G., TAR RNA recognition by a cyclic peptidomimetic of Tat protein. *Biochemistry* **2005,** 44, (37), 12362-72.

Legault, P.; Li, J.; Mogridge, J.; Kay, L. E.; Greenblatt, J., NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell* **1998,** 93, (2), 289-99.

Lim, W. A.; Fox, R. O.; Richards, F. M., Stability and Peptide Binding-Affinity of an Sh3 Domain from the Caenorhabditis-Elegans Signaling Protein Sem-5. *Protein Science* **1994,** 3, (8), 1261-1266.

Liu, X.; Thomas, J. R.; Hergenrother, P. J., Deoxystreptamine dimers bind to RNA hairpin loops. *J Am Chem Soc* **2004,** 126, (30), 9196-7.

Mammen, M.; Choi, S. K.; Whitesides, G. M., Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. *Angewandte Chemie-International Edition* **1998,** 37, (20), 2755-2794.

Marx, V., *Chem Eng News* **2005,** 83, 17-24.

Mayer, M.; James, T. L., Discovery of ligands by a combination of computational and NMR-based screening: RNA as an example target. *Nuclear Magnetic Resonance of Biological Macromolecules, Part C* **2005,** 394, 571-587.

McNaughton, B. R.; Gareiss, P. C.; Miller, B. L., Identification of a selective small-molecule ligand for HIV-1 frameshift-inducing stem-loop RNA from an 11,325 member resin bound dynamic combinatorial library. *J Am Chem Soc* **2007,** 129, (37), 11306-7.

Melnikova, I., RNA-based therapies. *Nature Reviews Drug Discovery* **2007,** 6, 863-864.
Messias, A. C.; Sattler, M., Structural basis of single-stranded RNA recognition. *Acc Chem Res* **2004,** 37, (5), 279-87.

Morken, J. P.; Kapoor, T. M.; Feng, S. B.; Shirai, F.; Schreiber, S. L., Exploring the leucine-proline binding pocket of the Src SH3 domain using structure-based, split-pool synthesis and affinity-based selection. *Journal of the American Chemical Society* **1998,** 120, (1), 30-36.

Nadassy, K.; Wodak, S. J.; Janin, J., Structural features of protein-nucleic acid recognition sites. *Biochemistry* **1999,** 38, (7), 1999-2017.

Nagai, K., RNA-protein interactions. *Curr Opin Chem Biol* **1992,** 2, 131-137.

Nishikawa, K.; Toker, A.; Johannes, F. J.; Songyang, Z.; Cantley, L. C., Determination of the specific substrate sequence motifs of protein kinase C isozymes. *J Biol Chem* **1997,** 272, (2), 952-60.

Oubridge, C.; Ito, N.; Evans, P. R.; Teo, C. H.; Nagai, K., Crystal structure at 1.92 A resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **1994,** 372, (6505), 432-8.

Pal, A.; Kraetzner, R.; Gruene, T.; Grapp, M.; Schreiber, K.; Gronborg, M.; Urlaub, H.; Becker, S.; Asif, A. R.; Gartner, J.; Sheldrick, G. M.; Steinfeld, R., Structure of tripeptidyl-peptidase I provides insight into the molecular basis of late infantile neuronal ceroid lipofuscinosis. *J Biol Chem* **2009,** 284, (6), 3976-84.

Patel, D. J., Adaptive recognition in RNA complexes with peptides and protein modules. *Curr Opin Struct Biol* **1999,** 9, (1), 74-87.

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. *Journal of Medicinal Chemistry* **1999,** 42, (20), 4042-4053.

Porter, E. A.; Weisblum, B.; Gellman, S. H., Mimicry of host-defense peptides by unnatural oligomers: antimicrobial beta-peptides. *J Am Chem Soc* **2002,** 124, (25), 7324-30.

Price, S. R.; Evans, P. R.; Nagai, K., Crystal structure of the spliceosomal U2B"-U2A' protein complex bound to a fragment of U2 small nuclear RNA. *Nature* **1998,** 394, (6694), 645-50.

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), 1200-3.

Pustowka, A.; Dietz, J.; Ferner, J.; Baumann, M.; Landersz, M.; Konigs, C.; Schwalbe, H.; Dietrich, U., Identification of peptide ligands for target RNA structures derived from the HIV-1 packaging signal psi by screening phage-displayed peptide libraries. *Chembiochem* **2003,** 4, (10), 1093-7.

Raja, C.; Ferner, J.; Dietrich, U.; Avilov, S.; Ficheux, D.; Darlix, J. L.; de Rocquigny, H.; Schwalbe, H.; Mely, Y., A tryptophan-rich hexapeptide inhibits nucleic acid destabilization chaperoned by the HIV-1 nucleocapsid protein. *Biochemistry* **2006,** 45, (30), 9254-65.

Ray, K.; Hines, C. S.; Coll-Rodriguez, J.; Rodgers, D. W., Crystal structure of human thimet oligopeptidase provides insight into substrate recognition, regulation, and localization. *J Biol Chem* **2004,** 279, (19), 20480-9.

Riemen, A. J.; Waters, M. L., Design of highly stabilized beta-hairpin peptides through cation-pi interactions of lysine and n-methyllysine with an aromatic pocket. *Biochemistry* **2009,** 48, (7), 1525-31.

Rowsell, S.; Stonehouse, N. J.; Convery, M. A.; Adams, C. J.; Ellington, A. D.; Hirao, I.; Peabody, D. S.; Stockley, P. G.; Phillips, S. E., Crystal structures of a series of RNA aptamers complexed to the same protein target. *Nat Struct Biol* **1998,** 5, (11), 970-5.

Saikumari, Y. K.; Ravindra, G.; Balaram, P., Structure formation in short designed peptides probed by proteolytic cleavage. *Protein Pept Lett* **2006,** 13, (5), 471-6.

Schueller, A.; Suhartono, M.; Fechner, U.; Tanrikulu, Y.; Breitung, S.; Scheffer, U.; Goebel, M. W.; Schneider, G., The concept of template-based de novo design from drug-derived molecular fragments and its application to TAR RNA. *Journal of Computer-Aided Molecular Design* **2008,** 22, (2), 59-68.

Segal, D. J.; Barbas, C. F., 3rd, Design of novel sequence-specific DNA-binding proteins. *Curr Opin Chem Biol* **2000,** 4, (1), 34-9.

Silva, J. G.; Carvalho, I., New insights into aminoglycoside antibiotics and derivatives. *Curr Med Chem* **2007,** 14, (10), 1101-19.

Stanger, H. E.; Gellman, S. H., Rules for Antiparallel β-Sheet Design: $_{D}$-Pro-Gly Is Superior to L-Asn-Gly for β-Hairpin Nucleation. *Journal of the American Chemical Society* **1998,** 120, 4236-4237.

Syud, F. A., Espinosa, J.F., Gellman, S.H., NMR-Based Quantification of ß-Sheet Populations in Aqueous Solution through Use of Reference Peptides for the Folded and Unfolded States *J Am Chem Soc* **1999,** 121, (49), 11577–11578.

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

Thomas, J. R.; Liu, X.; Hergenrother, P. J., Size-specific ligands for RNA hairpin loops. *J Am Chem Soc* **2005,** 127, (36), 12434-5.

Thomas, J. R.; Liu, X.; Hergenrother, P. J., Biochemical and thermodynamic characterization of compounds that bind to RNA hairpin loops: toward an understanding of selectivity. *Biochemistry* **2006,** 45, (36), 10928-38.

Tisne, C.; Dardel, F., Optimisation of a peptide library for screening specific RNA ligands by flow-injection NMR. *Comb Chem High Throughput Screen* **2002,** 5, (7), 523-9.

Tisne, C.; Guilliere, F.; Dardel, F., NMR-based identification of peptides that specifically recognize the d-arm of tRNA. *Biochimie* **2005,** 87, (9-10), 885-8.

Tor, Y., RNA and the Small Molecule World. *Angew Chem Int Ed Engl* **1999** 38, (11), 1579-1582.

Tor, Y., Targeting RNA with small molecules. *Chembiochem* **2003,** 4, (10), 998-1007.

Tugyi, R.; Uray, K.; Ivan, D.; Fellinger, E.; Perkins, A.; Hudecz, F., Partial D-amino acid substitution: Improved enzymatic stability and preserved Ab recognition of a MUC2 epitope peptide. *Proc Natl Acad Sci U S A* **2005,** 102, (2), 413-8.

Violin, J. D.; Zhang, J.; Tsien, R. Y.; Newton, A. C., A genetically encoded fluorescent reporter reveals oscillatory phosphorylation by protein kinase C. *J Cell Biol* **2003,** 161, (5), 899-909.

Voet, D.; Voet, J., *Biochemistry*. 3rd ed.; John Wiley & Sons, Inc: Hoboken, NJ, 2004.

Wang, Y.; Hamasaki, K.; Rando, R. R., Specificity of aminoglycoside binding to RNA constructs derived from the 16S rRNA decoding region and the HIV-RRE activator region. *Biochemistry* **1997,** 36, (4), 768-779.

Wang, Y.; Killian, J.; Hamasaki, K.; Rando, R. R., RNA molecules that specifically and stoichiometrically bind aminoglycoside antibiotics with high affinities. *Biochemistry* **1996,** 35, (38), 12338-46.

Wemmer, D. E., Designed Sequence-Specific Minor Groove Ligands. *Annual Review of Biophysics and Biomolecular Structure* **2000,** 29, 439-461.

Wong, C. H.; Hendrix, M.; Manning, D. D.; Rosenbohm, C.; Greenberg, W. A., A library approach to the discovery of small molecules that recognize RNA: Use of a 1,3-hydroxyamine motif as core. *Journal of the American Chemical Society* **1998,** 120, (33), 8319-8327.

Wüthrich, K., *NMR of Proteins and Nucleic Acids*. John Wiley & Sons, Inc: New York, 1986.

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), 827-40.

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