Abstract

GRAHAM, WILLIAM D. Isolating HIV NCp7 Peptide Mimics that Target the Anticodon Stem Loop Region of Human Transfer RNA Lysine[3] in a Modification Dependent Manner. (Under the direction of Dr. Paul F. Agris).

Nucleocapsid protein 7 (NCp7) of human immunodeficiency virus type 1 (HIV-I) is a multifunctional nucleic acid chaperone protein responsible for the processing of human tRNA Lysine 3 (htRNA^{Lys3}) into a primer for viral reverse transcription and the packaging of nucleic acid into progeny virion. To investigate the relationship between the post transcriptional modifications 2-methylthio-N6-threonyl carbamoyladenosine at position 37 ($ms^2t^6A_{37}$) and 5-methoxycarbonylmethyl-2-thiouridine at position 34 ($mcm^5s^2U_{34}$) of the anticodon stem loop region (ASL) of htRNA^{Lys3} and NCp7's recognition and recruitment of htRNA^{Lys3}, fluorescence quenching assays were performed. The ASL region of htRNA^{Lys3} was chemically synthesized with both modifications ms^2t^6A and mcm^5s^2U, occurring at positions 37 and 34 respectively (ASL$^{Lys3}mcm^5s^2U_{34}$; $ms^2t^6A_{37}$). NCp7 demonstrated a higher affinity and specificity for the modified ASL^{Lys3} $mcm^5s^2U_{34}$; $ms^2t^6A_{37}$ with a dissociation constant (K_d) of 0.28 µM compared to the protein's interaction with chemically synthesized ASL regions of unmodified htRNA^{Lys3} (ASL^{Lys3}) and unmodified ASL valine[3] (ASL^{Val3}) both demonstrating K_d values in the high µM range. These results suggest that the post-transcriptional modification present on htRNA^{Lys3} can potentially be used as a target for peptide antiviral therapeutics. Phage display was used to screen a library of peptides generated from RNA-binding proteins for their affinity to bind ASL^{Lys3} in a sequence and modification specific manner. Peptide sequences demonstrating the highest affinity for binding ASL$^{Lys3}mcm^5s^2U_{34}$; $ms^2t^6A_{37}$ were selected and synthesized into a library of

fluoescein tagged peptides.   A fluorescein isothiocyanate (FITC) conjugate was attached to the N-terminus of each peptide and assayed for modification specificity using $ASL^{Lys3}mcm^5s^2U_{34}$; $ms^2t^6A_{37}$, singly modified ASL^{Lys1} ($ASL^{Lys1}t^6A_{37}$), ASL^{Lys3} and ASL^{Val3}. Among the 20 peptides synthesized, peptides 6 (P6) and 17 (P17) demonstrated the highest affinity and specificity for $ASL^{Lys3}mcm^5s^2U_{34}$; $ms^2t^6A_{37}$ with K_d values of 0.45 µM and 0.60 µM respectively. A more in-depth analysis of the interaction between P6 and P17 with $ASL^{Lys3}mcm^5s^2U_{34}$; $ms^2t^6A_{37}$ by circular dichroism revealed that the peptides were able to mimic NCp7's ability to restructure the ASL region of $htRNA^{Lys3}$.

Isolating HIV NCp7 Peptide Mimics that Target the Anticodon Stem Loop Region of Human Transfer RNA Lysine[3] in a Modification Dependent Manner

by
William Darnell Graham

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Molecular and Structural Biochemistry

Raleigh, North Carolina

2010

APPROVED BY:

_____   _____
Michael B. Goshe, Ph.D.   William Miller, Ph.D.
Committee Chair

_____
John Cavanagh, Ph.D.

Biography

William Graham was born on a small military base in Havelock, North Carolina. As the son of a Marine Drill Sergeant and a 5th grade teacher the importance of discipline, a good work ethic, and education were instilled in him at a very young age and became the foundation of the man he is today. At the age of 17, he moved to Raleigh, NC where he worked his way through college and obtained a Bachelor of Science degree in Microbiology in 2001. Immediately after graduation, he found work as a Research Technician at his alma mater, where he worked his way through an advanced degree in Biochemistry in 2010.

Table of Contents

List of Tables

List of Figures

Chapter 1
Introduction

The lentivirus, Human Immunodeficiency Virus type 1 (HIV-1), is a highly adaptive virus that uses a complex system of proteolytic cleavages of the viral polyprotein Pr55gag (Gag), to replicate and integrate its viral genome into host cell chromosomes. The Gag polyprotein contains a zinc finger motif that is conserved among most retroviral proteins and is the only component of the HIV-1 virion that is required for virus particle assembly [1-3]. The viral life cycle of HIV-1 includes an assembly and budding step that results in the proteolytic cleavage of the Gag polyprotein by HIV-1 protease (PR) into six mature Gag proteins. Arguably the most vital protein produced from this cleavage is Nuclear Capsid Protein 7 (NCp7) [1,3,4]. NCp7 is a 71 amino acid protein (55 without spacer peptide 2 region) (Figure 1) that is cleaved from the C-terminus of the Gag polyprotein and functions as a nucleic acid chaperone during various stages of viral progression [5-7].

FIGURE 1. NCp7 Sequence[7]

NCp7's chaperone activity is characterized by its ability to destabilize nucleic acid structures and rearrange them into the most thermodynamically favorable conformations. This characteristic is vital during the replication stage of HIV-1 pathogenesis, in which human tRNA Lysine 3 (htRNA^{Lys3}) (Figure 2A) is rearranged to function as a primer for reverse transcription. During the assembly and packaging stage of HIV-1 pathogenesis, NCp7's role as a nucleic acid chaperone involves the binding of host cell tRNAs and packaging them into newly formed virion particles [5,7]. NCp7's ability to bind and manipulate nucleic acids for successful viral propagation, and the mechanism surrounding its interaction with tRNA make it a primary target for the development of antiviral therapeutics.
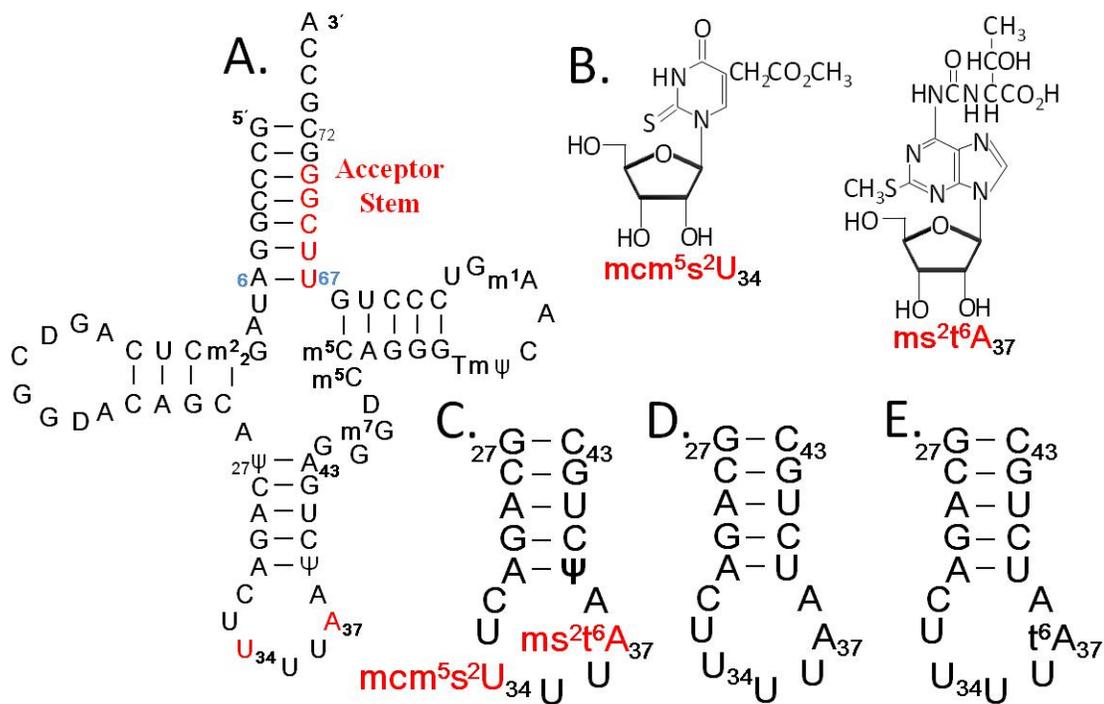
FIGURE 2. Human tRNA Lysine 3 Sequence and Structure: *(A).* Human tRNA Lysine 3 (htRNA^{Lys3}) *(B).* Modified nucleosides: methylthio-N6-threonyl carbamoyladenosine (ms^{2}t^{6}A) and 5-methoxycarbonylmethyl-2-thiouridine (mcm^{5}s^{2}U) *(C).* ASL domain of htRNA^{Lys3} with the modifications (ms^{2}t^{6}A$_{37}$) and 5-methoxycarbonylmethyl-2-thiouridine (mcm^{5}s^{2}U) at position 34 (ASL^{Lys3}ms^{2}t^{6}A$_{37}$: mcm^{5}s^{2}U$_{34}$) *(D).* Anticodon Stem loop (ASL) domain of htRNA^{Lys3} (ASL^{Lys3}) *(E).* ASL domain of htRNA^{Lys1} with the modification N6-threonyl carbamoyladenosine (t^{6}A) at position 37 (ASL^{Lys1}t^{6}A$_{37}$).

NCp7 binds to nucleic acids in a complex ionic strength dependent manner. This is characterized by the proteins promiscuity and ability to bind a variety of nucleic acids with varying affinity. In terms of sequence specificity, NCp7 demonstrates the highest affinity for binding the L-shape of tRNA, on the acceptor stem, rich in G-U and A base pairs, which is indicative of the anticodon stem loop (ASL) region of htRNA^{Lys3} (Figure 2D), which functions as a primer for reverse transcription [5,8]. The ability of NCp7 to melt the nucleic acid into proper conformation suggests that it simultaneously binds the loop region and the stem of the tRNA. Structural studies on NCp7 have identified that the zinc finger domains (Figure 1) from the parental Gag protein are retained by the chaperone protein. This explains how NCp7 can attach to the htRNA^{Lys3} in two different regions [4,9-11]. Mature htRNA^{Lys3} contains the naturally occurring post transcriptional modifications methylthio-N6-threonyl carbamoyladenosine (ms^{2}t^{6}A$_{37}$) at position 37 and 5-methoxycarbonylmethyl-2-thiouridine (mcm^{5}s^{2}U$_{34}$) at position 34 on the ASL region (Figure 2B). In many instances, post transcriptional modifications are essential for protein-RNA recognition [12,13]. Most RNA binding proteins recognize their cognate tRNA's identity through the structure and chemistry influenced by naturally occurring modifications found in the ASL region [13-15]. More specifically post transcriptional modifications on htRNA allow for another mechanism of regulation in terms of protein synthesis. The post transcriptional modifications, ms^{2}t^{6}A$_{37}$ at position 37 and mcm^{5}s^{2}U$_{34}$ at position 34 have demonstrated the ability to increase the affinity of htRNA^{Lys3} to bind mRNA on the ribosome, and increase the efficiency of translation *in vitro* [14,16,17]. Based on the literature it is inferred that naturally occurring post transcriptional modifications present on htRNA^{Lys3} assist in the recognition of tRNA^{Lys3} by NCp7 [18,19]. The high specificity of NCp7 for htRNA^{Lys3} and its

ability to melt tRNA^{Lys3} into the appropriate conformation for reverse transcription could be explained by the structural characteristics of NCp7 and the presence of post transcriptional modifications in the ASL. The electrostatic properties of these modifications destabilize the loop region of htRNA^{Lys3}, making it more dynamic and susceptible to conformational influence by NCp7 [16]. To investigate the relationship between post transcriptional modification and the recognition of htRNA^{Lys3} by NCp7, a fluorescence quenching assay was performed, which identified that the recognition of htRNA^{Lys3} by NCp7 is greatly enhanced when modifications are present in the ASL region.

In addition to translation and viral replication, there are many nucleic acid pathways that require the recognition of RNA by proteins mediated by post transcriptional modifications. Sequence analysis of RNA-binding proteins has led to the identification of specific motifs that are responsible for RNA structure and recognition [20]. These motifs are characterized by the presence of positively charged residues, like arginine, that interact with the negatively charged backbone of RNA. These proteins have been well characterized using a variety of analytical techniques that do not take into account structure and physiological conditions [21]. The recognition of RNA by proteins is heavily dependent on the conformational state of the RNA. Post transcriptional modifications exert dynamic influence on RNA structure, specifically the loop region of tRNA, and dictate folding conformations that are favorable for protein-RNA interaction. Most of the studies performed on nucleic acid recognizing proteins have led to the generation of peptide libraries that do not take into account modifications and structural characteristics of RNA in physiological conditions. Phage display screening is a biophysical technique that has been used in previous studies to effectively generate a library of protein

derivatives, peptides, that bind specifically to modified nucleosides with high affinity (K_d = nM to μM), with regard to structure and proper physiological conditions [22-24]. The generation of a library of peptides that can effectively function as RNA binding protein mimics can potentially lead to the development of a therapeutic that can silence viral gene expression *in vivo*. In this study the initial screen for peptides that bind specifically to the ASL region of htRNA^{Lys3}mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$ (ASL^{Lys3}mcm^{5}s^{2}U$_{34}$, ms^{2}t^{6}A$_{37}$) (Figure 1D) with high affinity was conducted by phage display. The peptide library generated from this screen was further analyzed for its affinity to bind ASL^{Lys3}mcm^{5}s^{2}U$_{34}$, ms^{2}t^{6}A$_{37}$ using fluorescence quenching assays. This procedure produced two peptides that not only bind specifically to ASL^{Lys3}mcm^{5}s^{2}U$_{34}$, ms^{2}t^{6}A$_{37}$ with K_d values comparable to NCp7, but can also mimic NCp7s ability to melt the RNA into specific conformations. The sequences of the peptides are currently under review for a patent application and can not be shown at this time.

Chapter 2
Experimental Methods

RNA Sample Preparation

The unmodified *Escherichia coli* ASL region of tRNA Valine 3 (ASL^{Val3}) and the unmodified human ASL region of $tRNA^{Lys3}$ (ASL^{Lys3}) used in this study were chemically synthesized by Dharmacon (ThermoFisher, Lafayette, CO) using "ACE" chemistry [25]. The modified $ASL^{Val3}m^6A_{37}$ and modified $ASL^{Lys3}ms^2t^6A_{37}$: $mcm^5s^2U_{34}$ was synthesized by Integrated DNA Technologies BVBA (formerly RNA-Tec) [26]. The newly synthesized ASLs were then analyzed by HPLC nucleoside composition analyses[27] and further confirmed by mass spectrometry (MALDI-TOF), and NMR analyses. A previous study performed using $ASL^{Val3}m^6A_{37}$ also confirmed the presence of the modification m^6A_{37} by X-ray crystallography [16]. The modified human ASL region of tRNA Lysine 1 ($htRNA^{Lys1}$) containing the modification t^6A ($ASL^{Lys1}t^6A_{37}$) was synthesized by the North Carolina State University Nucleic Acid Facility with standard ribonucleoside phosphoamidites (ChemGenes) and with little change in standard synthesis protocols [28]. All ASLs were purified by HPLC using a Nucleogen 60-7 DEAE (250 mm × 10 mm) column and desalted (Waters Corporation Sep-pak columns, Milford, MA) [29].

Phage Display Library Selection

Biotinylated ASL^{Lys3} and $ASL^{Lys3}ms^2t^6A_{37}$: $mcm^5s^2U_{34}$ receptors were immobilized on the surface of streptavidin-coated high capacity binding, 96-well microplates (Pierce, Rockford, IL) in 200 μL of TTDBA buffer (1 mg/ml BSA and 0.02% NaN3 in 200/1 TBS/Tween, vol/vol); TBS buffer, (50 mM Tris HCl, pH 7.5, 150 mM NaCl). The streptavidin-coated plates containing

the fixed ASLs were incubated for 4hr at 4°C to ensure attachment of the ASLs to streptavidin surface. Any unbound receptors were removed by washing the plates five separate times with TBS-tween. The plates were then washed with a buffer solution containing 10 mM biotin and incubated at room temperature for 10 minutes to ensure maximum coating of the streptavidin surface [30,31]. The streptavidin plates were washed with input phage and incubated for 4hr at 4°C. After the incubation period, to remove any unbound phage, the plates were washed 5 times with a TBS-tween solution. To elute the bound phage from the plates, using both acidic and basic conditions separately, the plates were rinsed with 200 μL of acid elution buffer (acidic conditions) or alkaline elution buffer (basic conditions) and placed on a shaker for 10 minutes [21]. The elution mixture containing eluted phage was transferred to a microtube containing respective neutralizing buffer [21]. The isolated phage were then amplified for a total of 5 rounds of subsequent screening. The final round of screening was concluded with a serial dilution of the isolated phage and a quantitative assessment performed to determine yield. Clones were then constructed using starved *E. coli* K91BluKan cells. DNA was then isolated from the clones and sequenced [21]. The resulting sequences were used to produce the peptide library.

Peptide Synthesis

Peptides were designed with a Fluorescein isothiocyanate (FITC) tag on the N-terminus for screening by fluorescence spectroscopy. The synthesis of the library was performed by Sigma Aldrich.

Fluorescence Spectroscopy Screening

The NCp7 fluorescence quenching studies were performed using the tryptophan residue found at position 37 on the full length NCp7 protein. The amino acid tryptophan possesses fluorescent properties that can be monitored for molecular interaction by way of fluorescence quenching. The NCp7 was titrated with varying amounts of ASLs and the tryptophan fluorescence signal monitored at an excitation point of 285 nm and an emission point of 360 nm using a Spectramax Gemini XS dual scanning microplate spectrofluoremeter (Sunnyvale, CA). The peptide interaction with ASL targets was performed by monitoring the fluorescence signal from the N-terminus labeled FITC tag. The peptides were titrated with varying concentrations of ASLs in 20mM phosphate buffer and the fluorescence signal monitored at an excitation point of 486 nm and an emission wavelength of 524 nm. The binding constant (K_d) value for each fluorescence experiment was assessed using the one site binding equation $Y=Bmax*X/[K_d+X]$.

Circular Dichroism

CD spectra were recorded using a Jasco J600 spectropolarimeter and an interfaced computer [32,33]. RNA samples were prepared in 20 mM sodium phosphate buffer at a concentration of 1.5 μM. The prepared RNA samples were then placed in a 1 mL Hellma cuvette and read in both near and far U.V. (235 nm to 280 nm). The RNA samples were then titrated with pre-dried pure peptide sample to compensate for buffer and volume change, starting with an aliquot that would equal a concentration of 1.5 μM in a final volume of 1 mL then peaking at a final concentration of 9 μM. The experiments were run at 4°C, at a speed of 10 nm per minute, at a resolution of 1 nm, cell path length of 1cm and a total run accumulation of 12. The raw data was then converted to percent elipticity using RNA sample alone as the baseline.

Chapter 3
Experimental Results

NCp7 RNA Post Transcriptional Modification Recognition

NCp7 is promiscuous and can bind to a variety of nucleic acids with varying affinities [5]. Studies of NCp7 binding and structure suggest that it functions as a chaperone protein and can bind to the stem region of nucleic acid structures with little specificity. In contrast NCp7 has the ability to exhibit great sequence specificity when interacting with human tRNA (htRNA) [5,34,35]. This observation suggests that NCp7 binds htRNA lysine 3 (htRNA^{Lys3}) in a sequence first, modification second manner. When interacting with tRNA^{Lys3}, the NCp7 binds initially to the acceptor stem region then binds to the anticodon loop (ASL) region of the htRNA to introduce structural instability. The change in ASL dynamics makes it energetically favorable for NCp7 to fold the RNA into the correct conformation for reverse transcription [1].

To examine how modifications influence the binding of NCp7 to htRNA^{Lys3} fluorescence quenching assays were performed using chemically synthesized ASL of htRNA^{Lys3} with and without modifications [36,37]. The 55 amino acid form of NCp7 was first titrated with the ASL region of unmodified htRNA lysine 3 (ASL^{Lys3}) (Figure 2D) to obtain a baseline titration curve. The fluorescence signal of NCp7's tryptophan residue was monitored with an emission wavelength of 360 nm and an excitation wavelength of 285 nm. The resulting signal spectrum was normalized into percent quenching and plotted against respective RNA concentrations. A binding constant (K_d) from the interaction of NCp7 and ASL^{Lys3} was derived from the analysis of the quench curve (Figure 3) using a one site binding model equation. The interaction between NCp7 and ASL^{Lys3} produced a weak K_d value in the low μM range (Table 1).
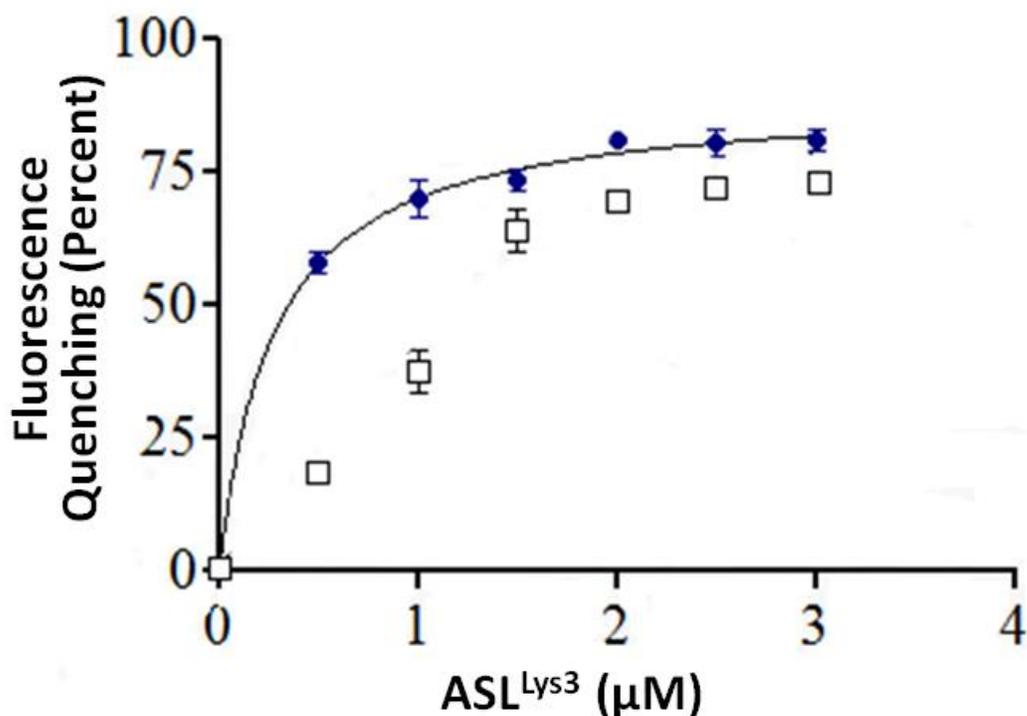
FIGURE 3. Fluorescence Quench Curve of NCp7 and ASL of htRNALys3: NCp7 concentration was kept constant and titrated with increasing amounts of modified and unmodified ASL Lysine 3 (ASL^{Lys3}): □ASL^{Lys3} unmodified (ASL^{Lys3}), ◆ASL^{Lys3} modified ($ASL^{Lys3}mcm^5s^2U_{34}$; $ms^2t^6A_{37}$).

NCp7 was then titrated with $ASL^{Lys3}mcm^5s^2U_{34};ms^2t^6A_{37}$ using the same parameters and conditions as ASL^{Lys3}. A 10 fold decrease in K_d value was observed when modifications are introduced into the ASL region of $htRNA^{Lys3}$. Further examination of Figure 3 reveals that a max quenching of only 68% is achieved when NCp7 is titrated with ASL^{Lys3} compared to the 75% quenching achieved by $ASL^{Lys3}mcm^5s^2U_{34,}ms^2t^6A_{37}$. These results strongly suggest that NCp7 recognition of $htRNA^{Lys3}$ is significantly enhanced when modifications are present. This supports the theory that modifications are not necessary for the initial non-specific binding of

NCp7 to the acceptor stem region of htRNA^{Lys3}, but they are essential for protein recognition of the anticodon stem loop region.

Phage Display Library Selection of Peptides

Peptides that bind selectively to the anticodon stem loop of tRNA Lysine 3 containing the post transcriptional ms^{2}t^{6}A$_{37}$:mcm^{5}s^{2}U$_{34}$ (ASL ^{Lys3}mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$) were selected using a phage display library. The peptides used in the Phage Library (Library 73303) were 15 to 17 amino acids in length and were randomly generated from a sequence homology search of known RNA-chaperone proteins. The target RNAs were 17 nucleotides in length and synthesized with a biotin tag at the 3′ end. Unmodified ASL Lysine 3 (ASL^{Lys3}) and ASL ^{Lys3}mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$ were fixed onto streptavidin-coated, high capacity microplates. The binding and elution conditions were designed to emulate physiological conditions with regard to standard phage conditions. To conserve material the first round of screening was conducted with ASL^{Lys3} as the target nucleoside. Streptavidin coated plates containing biotin conjugated ASL^{Lys3} fixed to the surface, were washed with Phage Library 73303 and rinsed thoroughly with physiologically relevant buffer. The chemistry of nucleic acids would suggest that the proteins that bind the target RNA contain positive residues that may be sensitive to both alkaline and acidic conditions. To prevent a false positive or negative, the peptides were eluted from the plates using both alkaline and acidic conditions [21] and compared. The difference in phage output and phage input, represented by percent yield, was minimal. Peptides that demonstrated an affinity to bind ASL^{Lys3}, after four rounds of selective screening, were catalogued and eliminated from the second step of screening. The second step of screening was conducted on the remaining peptides that did not show specificity for ASL^{Lys3}, using ASL ^{Lys3}mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$. Out of the 200

12

colonies produced from four rounds of phage selection against ASL ^{Lys3}mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$ under basic conditions, 25 peptides were isolated. These peptides were sequenced and synthesized with a Fluorescein Isothiocyanate (FITC) tag at the N-terminus. After analysis of the generated library and the elimination of repeat peptides, 20 genuine RNA binding protein mimics were produced.

**Peptide RNA Post Transcriptional Modification Recognition**

Phage display library screening is an inexpensive and effective way of screening a large number of peptide-RNA interactions under biologically relevant conditions. In order to characterize the molecular chemistry of the interaction between the peptides and target nucleic acids, a more sensitive technique is required. The peptide library was synthesized with a FITC tag at the N-terminus to allow for high throughput screening by fluorescence spectroscopy. The quenching of the fluorophore was monitored at an emission wavelength of 524 nm and an excitation wavelength of 486 nm. Analysis of the peptide library sequences revealed that peptides 6 (P6), 15 (P15), 17 (P17) and 21 (P21) were isolated multiple times from separate colonies produced from phage display selection. This observation suggests that these peptides will bind to modified nucleosides with great affinity. For this reason the first peptide screen by fluorescence spectroscopy was conducted using these peptides. The modified nucleosides used for this study are extremely hard to isolate and must be chemically synthesized. To conserve precious material the initial fluorescence screen of the generated peptide library was performed with ASL^{Lys3}. The preparation of the peptide samples for fluorescence spectroscopy revealed a solubility problem with some of the peptides. To achieve sample homology in solution, the peptide mixtures were placed through several rounds of dilutions and a final concentration

assessed by BCA assay. This eliminated a large portion of peptide candidates including P15.

Once in solution P6, P17 and P21 were titrated with ASL^{Lys3}. The binding characteristics of each

peptide was then assessed by fluorescence quenching of the N-terminus fluorophore.

The resulting spectrum from each peptide was normalized into percent quenching and

plotted against the respective concentration of ASL^{Lys3}. Analysis of the quench curve (Figure 4)

revealed that P21 exhibited non-specific binding of ASL^{Lys3} while P6 and P17 demonstrated no

interaction with ASL^{Lys3}. The remaining two peptides were then screened against ASL

$^{Lys3}mcm^5s^2U_{34}$; $ms^2t^6A_{37}$ (Figure 5). A strong interaction between P6 and ASL $^{Lys3}mcm^5s^2U_{34}$;

$ms^2t^6A_{37}$ was observed, resulting in a K_d value of 0.44 µM. P17 demonstrated an equally strong

interaction with ASL $^{Lys3}mcm^5s^2U_{34}$; $ms^2t^6A_{37}$, with a K_d value of 0.60 µM. As a control a non-

RNA binding peptide (P31) was used as a baseline (data not shown). The data from the P31was

not represented on the graph because of the high error rate that occurs with non-specific

interaction. In an effort to fully characterize the relationship between modifications and the

recognition of $htRNA^{Lys3}$ by these peptide mimics, a second fluorescence screen was performed

on P6 and P17 using the ASL region of lysine 1 containing the modification N6-threonyl

carbamoyladenosine (t^6A) at position 37 ($ASL^{Lys1}t^6A_{37}$). The fluorescence assay was

performed under the same conditions and parameters as ASL^{Lys3} and ASL $^{Lys3}mcm^5s^2U_{34}$;

$ms^2t^6A_{37}$. Analysis of the data in Figure 5 illustrates a clear influence by the modification on the

recognition of the ASL region of $htRNA^{Lys3}$ for both P6 and P17. The calculated K_d values for

P17 and P6 were 0.60 µM and 1.75 µM respectively. In comparison the molecular interaction

between these peptides and $ASL^{Lys1}t^6A_{37}$ is not as strong as the interaction observed when the

peptides are titrated with ASL $^{Lys3}mcm^5s^2U_{34}$; $ms^2t^6A_{37.}$ Despite this observation P6 titrated with

$ASL^{Lys1}t^6A_{37}$ exhibited a max signal quenching of 41% compared to 22% observed with ASL^{Lys3}. These results reiterate the necessity of modifications to achieve energetically favorable interaction between chaperone proteins and nucleic acids.
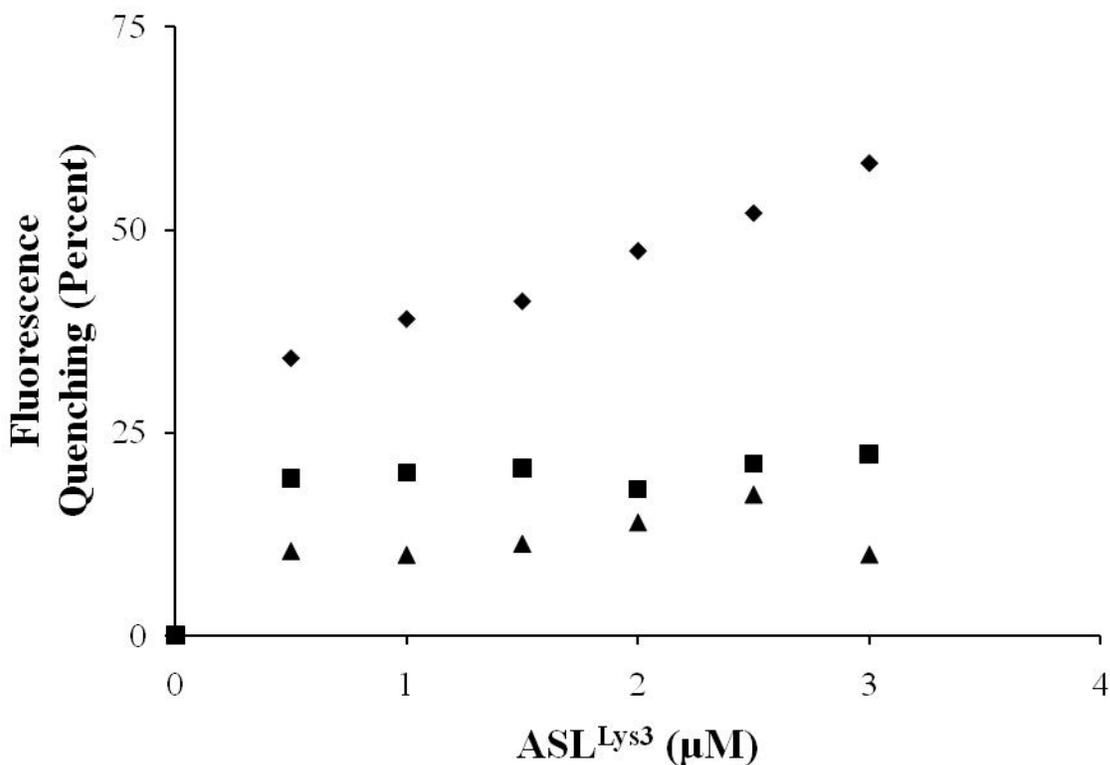
FIGURE 4. Fluorescence Quench Curve of Peptide 6 and Peptide 17 with control ASL^{Lys3}: To ensure integrity of the screen peptides generated from the phage display library were prescreened by fluorescence quenching using the anticodon stem loop (ASL) region of Lysine 3 unmodified (ASL^{Lys3}) to eliminate any nonspecific binding proteins. Peptide specificity control screen against ASL^{Lys3}: ■P6, ◆P21 (Non-specific control peptide), and ▲P17.
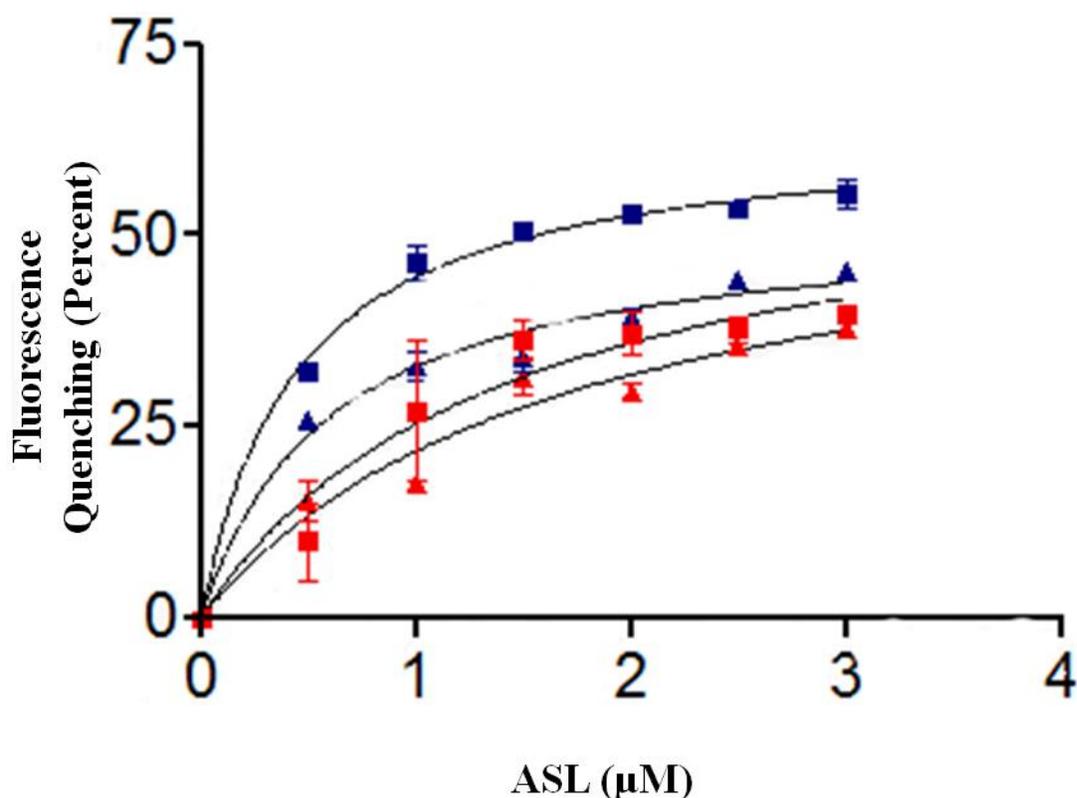
FIGURE 5. Fluorescence Quench Curve of Peptide 6 and Peptide 17 with Modified ASL^{Lys}: Peptides that did not exhibit specificity for ASL^{Lys3} were then screened for modification specificity: ■P6 titrated with $ASL^{Lys3}mcm^5s^2U_{34}$; $ms^2t^6A_{37}$, ▲P17 titrated with $ASL^{Lys3}mcm^5s^2U_{34}$; $ms^2t^6A_{37}$, ■P6 titrated with $ASL^{Lys1}t^6A_{37}$ ▲P17 titrated with $ASL^{Lys1}t^6A_{37}$.

A final set of fluorescence experiments were performed using modified and unmodified valine 3 ASLs. These experiments were designed to identify sequence specificity and modification recognition. The ASL region of valine 3 (ASL^{Val3}) (Figure 6A) was chemically synthesized to contain the modification N6-methyladenosine (m^6A) at position 37 ($ASL^{Val3}m^6A_{37}$) (Figure 6B). This modification is not the same naturally occurring modification that is found on mature $htRNA^{Lys3}$, but structural studies have demonstrated that this

16

modification exhibits the same influence on loop dynamics [14,38] as what is observed with ASL^{Lys3} and post transcriptional modifications.
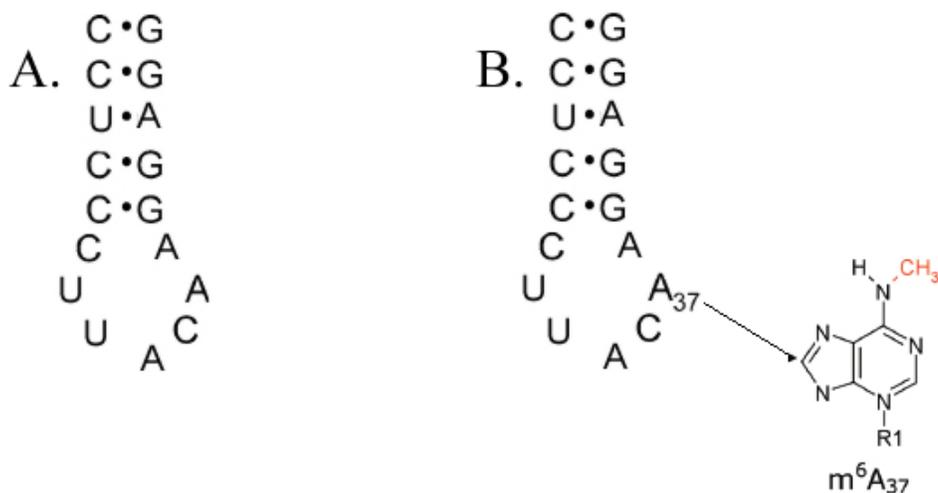
FIGURE 6. Modified and Unmodified *E.coli* ASL Valine 3 Sequence and Structure: *(A).* Anticodon Stem Loop (ASL) region of tRNA Valine 3 (ASL^{Val3}). *(B).* ASL^{Val3} containing the modification N6-methyladenosine (m^{6}A) at position 37 (ASL^{Val3}m^{6}A$_{37}$).

The same conditions for the fluorescence assay were used as with the previous experiments. Figure 7 is a scatter plot of the percent quenching against RNA concentration. The peptides exhibited very little specificity to bind ASL^{Val3} or ASL^{Val3}m^{6}A$_{37}$. A K_d value could not be extrapolated from the resulting quench curves because the interaction between the peptides and both modified and unmodified ASL^{Val3} was not strong enough to generate significant signal quenching. These results identify sequence specificity for ASL^{Lys3} exhibited by both peptides 6 and 17. This correlates with the idea that the interaction between specific proteins and ASL^{Lys3}

occur in a two state manner, where sequence recognition occurs in the acceptor region and is

enhanced by the presence of modifications.
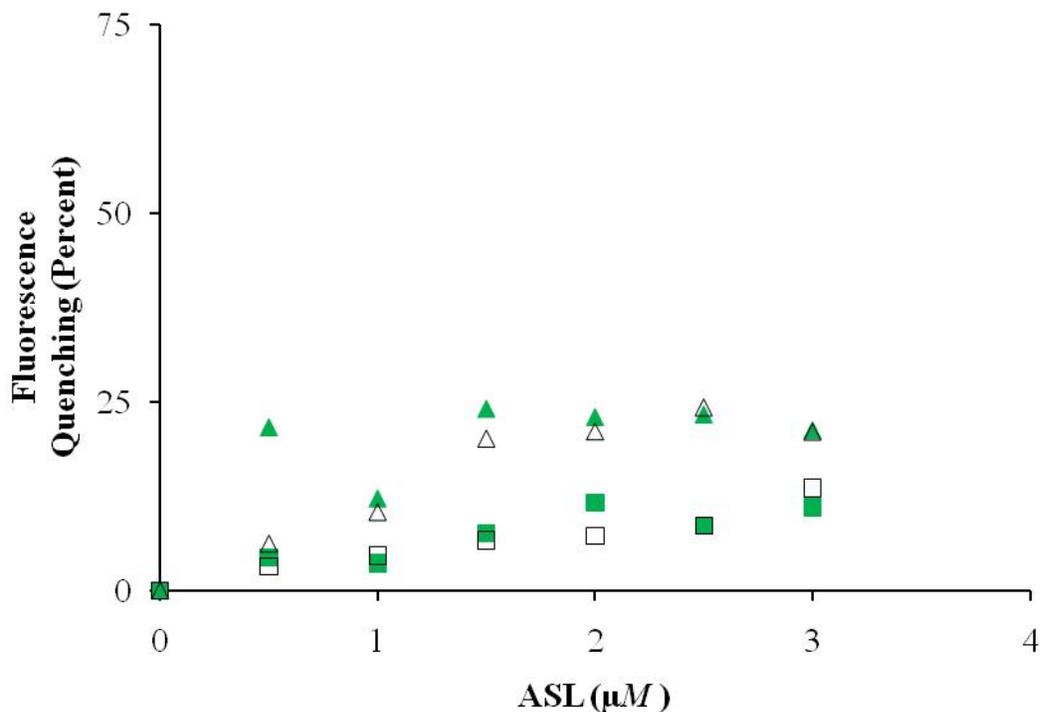
FIGURE 7. Fluorescence Quench Curve of Peptide 6 and Peptide 17 with Modified and Unmodified *E.coli* ASL^{Val3}: Peptide screen for sequence specificity. □P6 titrated with unmodified ASL Valine 3(ASL^{Val3}), △P17 titrated with ASL^{Val3}, ■P6 titrated with modified ASL Valine 3 m^{6}A at position 37 (ASL^{Val3}m^{6}A$_{37}$) and ▲P17 titrated with ASL^{Val3}m^{6}A$_{37}$.

Table 1. Dissociation constant (K$_d$) values for peptide 6, peptide 17 and NCp7 interactions with various ASLs (K$_d$ = µM).

Peptide	ASL^{Lys3}	ASL^{Lys3} t^{6}A$_{37}$	mcm^{5}s^{2}U$_{34}$ms^{2}t^{6}A$_{37}$
Peptide 6	IC	0.60 +/- 0.09	0.44 +/- 0.05
Peptide 17	IC	1.75 +/- 0.23	0.60 +/- 0.09
NCP7	2.29 +/-0.62	ND	0.27 +/-0.03

IC, Incalculable   ND, not determined

18

ASL^{Lys3} Modification Dependent Restructuring by NCp7 and Peptide Mimics

The interaction between NCp7and modified tRNA^{Lys3} is dynamic in nature, resulting in the conformational change of the tRNA into a primer for viral genome reverse transcription [39]. Previous studies have shown that NCp7 binds preferentially to the L-shape of tRNA^{Lys3}, on the acceptor stem, rich in G-U and A base pairs, allowing the protein access to the loop region of the tRNA [39,40]. Studies illustrating this mechanism of interaction were performed in the absence of post transcriptional modifications that are present on the anticodon stem loop (ASL) of the tRNA *in vivo*. Studies analyzing the structural influence of post transcriptional modifications contained within the ASL region of tRNA, have demonstrated that modifications allow for the loop to be more dynamic and susceptible to conformational rearrangement [14]. In this study the ASL region of target tRNA were chemically synthesized with and without modifications and monitored for a change in conformation when titrated with NCp7 using circular dichroism (CD). Figure 8 is the resulting spectra of unmodified human ASL Lysine 3 (ASL^{Lys3}) titrated with NCp7. Measured in both near and far U.V.(230 nm to 285 nm), the CD spectra illustrate very little change in elipticity when the tRNA is titrated with NCp7 in the absence of RNA modifications.
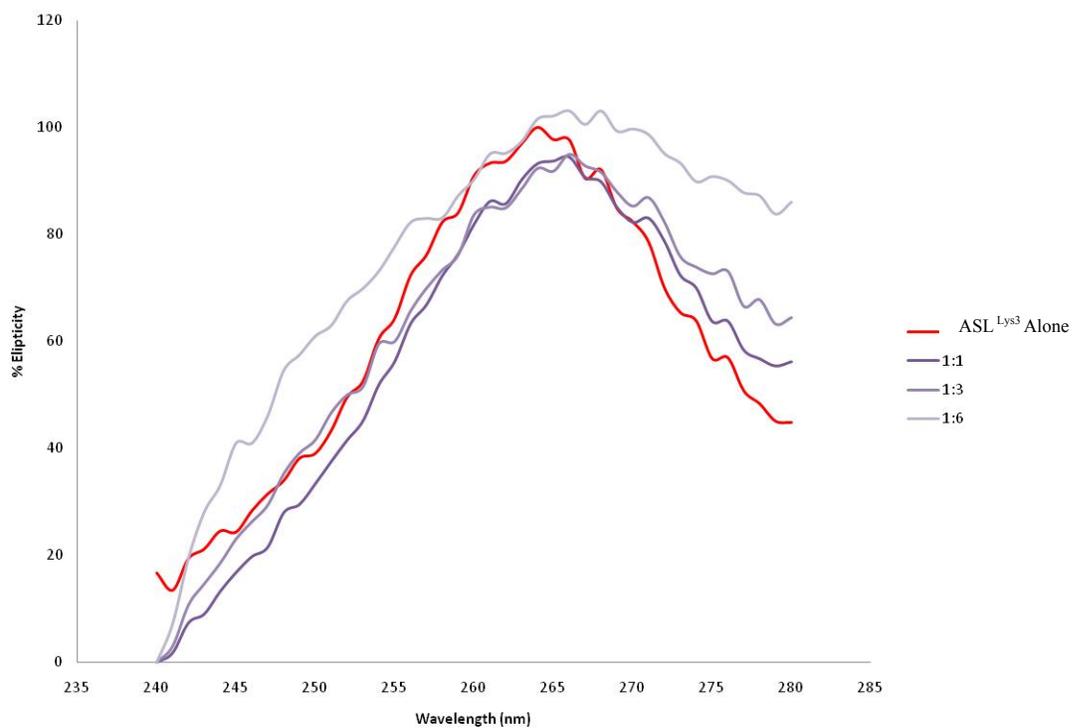
FIGURE 8. Circular Dichroism Spectra of NCp7 with ASL[Lys3] : (ASL[Lys3]) kept at a constant concentration of 1.5 μM then titrated with NCp7 starting at a 1:1 ratio and ranging to a 1:6 ratio of RNA to protein.

Under the same conditions the ASL region of htRNA[Lys3] containing modifications (ASL [Lys3]$mcm^5s^2U_{34}$; $ms^2t^6A_{37}$) was titrated with NCp7. The resulting data is represented in Figure 9. The presence of modifications significantly enhances the ability of NCp7 to melt the ASL region of htRNA[Lys3].
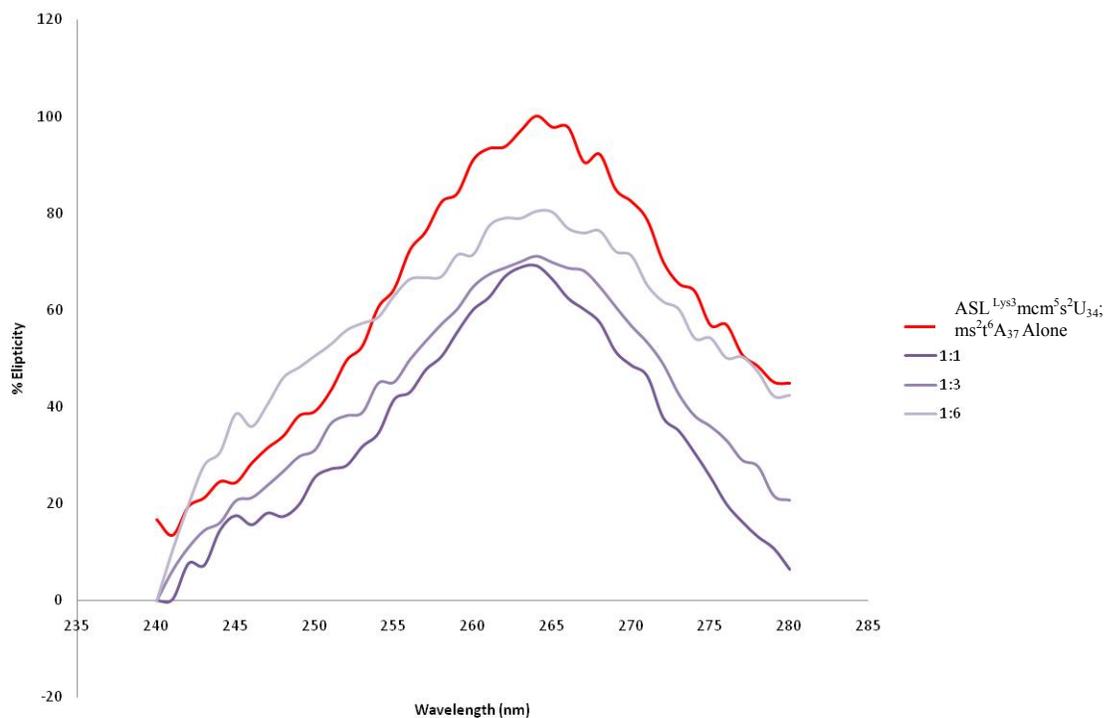
FIGURE 9. Circular Dichroism Spectra of NCp7 with Modified ASL^{Lys3}: kept at a constant concentration of 1.5 µM then titrated with NCp7 starting at a 1:1 ratio and ranging to a 1:6 ratio of RNA to protein.

The spectra produced from the titration of ASL ^{Lys3}mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$ with NCp7 compared to ASL^{Lys3} titrated with NCp7 indicates that the presence of modifications on the ASL region of htRNA^{Lys3} allows NCp7 to recognize the loop region of the tRNA and melt the tRNA. To determine if the two peptides isolated from the phage display library screen are able to mimic NCp7's ability to unfold tRNA, a second set of experiments were performed under the same conditions using the peptides 6 (P6) and 17 (P17). Figures 10 and 11 are scans representing control data of ASL^{Lys3} titrated with peptides 6 and 17. Analysis of the spectra indicates no interaction between peptide and ASL^{Lys3}. When modifications are introduced into the ASL region of the human tRNA^{Lys3}, a decrease in ASL base stacking is observed. Figure 12 is of ASL

^{Lys3}mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$ titrated with P6. Analysis of the spectra reveals that P6 is able to mimic

NCp7's ability to melt the tRNA. More specifically P6, is able to completely denature the ASL,

unlike NCp7, which appears to partially unfold the tRNA when introduced in the same

concentrations. Preliminary results even suggest that NCp7, when introduced in sufficient

concentrations, will begin to reorder the ASL region of the tRNA, resulting in an increase in

elipticity. Analysis of P17's interaction with ASL ^{Lys3}mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$ shows that P17 does

not function in the same manner as NCp7 and P6. In Figure 13 the elipticity of ASL

^{Lys3}mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$ is increased when titrated with P17, suggesting that P17 increases base

stacking in the ASL region of human tRNA^{Lys3} when modifications are present.

FIGURE 10. Circular Dichroism Spectra of Peptide 6 with ASL^{Lys3}:   (ASL^{Lys3}) kept at a
constant concentration of 1.5 µM then titrated with peptide 6 (P6) starting at a 1:1 ratio and
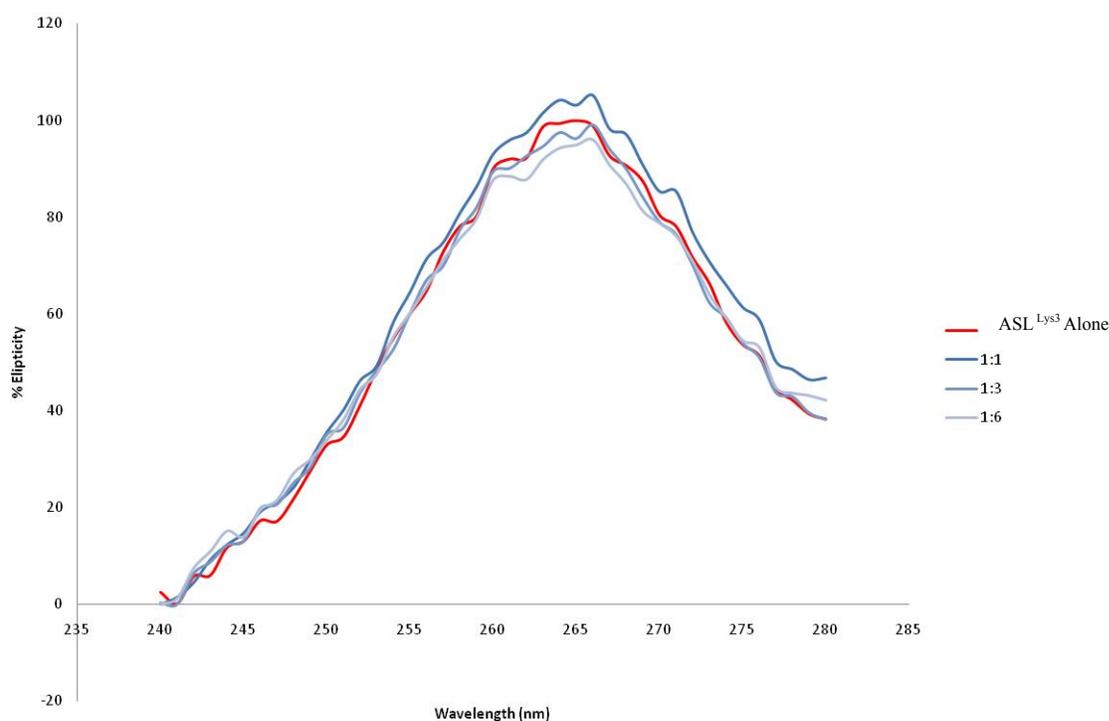ranging to a 1:6 ratio of RNA to protein.
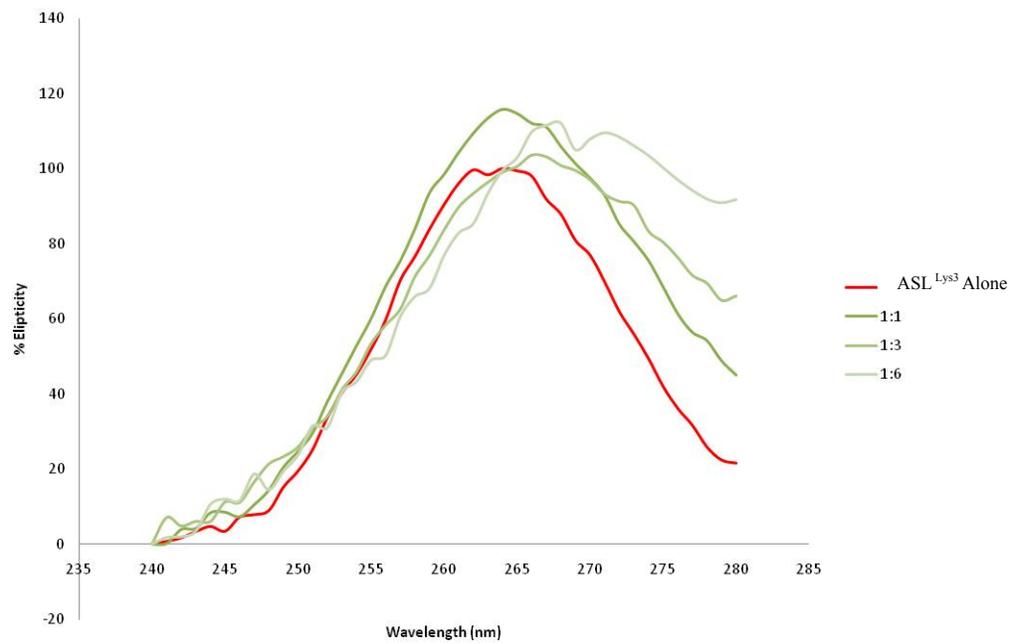
FIGURE 11. Circular Dichroism Spectra of Peptide 17 with ASL^{Lys3}: (ASL^{Lys3}) kept at a constant concentration of 1.5 μM then titrated with peptide 17 (P17) starting at a 1:1 ratio and ranging to a 1:6 ratio of RNA to protein.
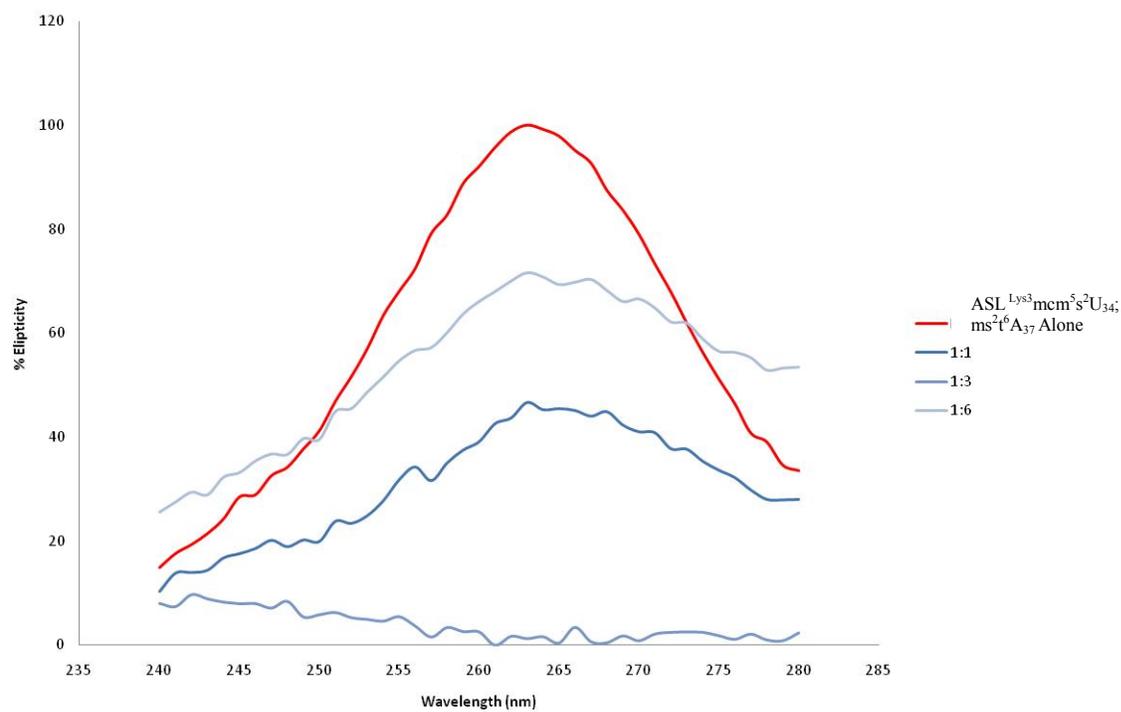
FIGURE 12. Circular Dichroism Spectra of Peptide 6 with Modified ASL^{Lys3}:
$(ASL^{Lys3}mcm^5s^2U_{34}; ms^2t^6A_{37})$ kept at a constant concentration of 1.5 µM then titrated with peptide 6 (P6) starting at a 1:1 ratio and ranging to a 1:6 ratio of RNA to protein.
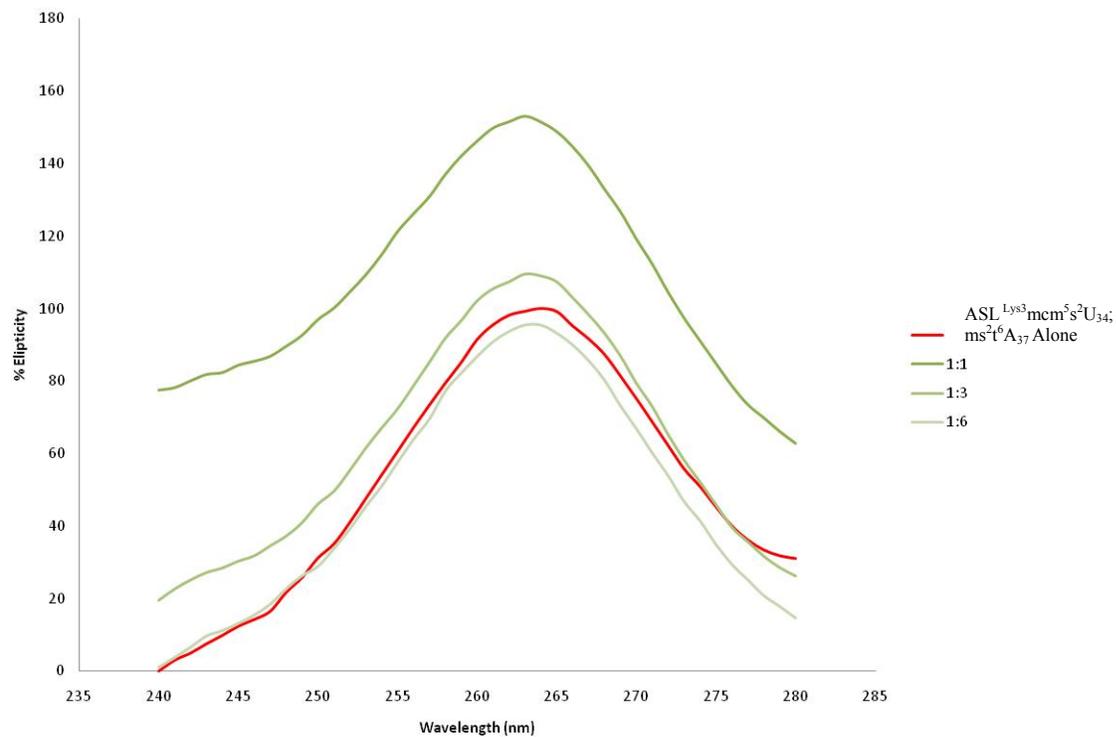
FIGURE 13. Circular Dichroism Spectra of Peptide 17 with Modified ASL^{Lys3}: (ASL ^{Lys3}mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$) kept at a constant concentration of 1.5 µM then titrated with peptide 17 (P17) starting at a 1:1 ratio and ranging to a 1:6 ratio of RNA to protein.

Chapter 4
Discussion

NCp7 Recognition of ASL^{Lys3} Post Transcriptional Modifications

The diverse nature of NCp7 during HIV-1 viral pathogenesis make it a prime target for the development of host discriminatory antiviral therapeutics. NCp7 is composed of two zinc fingers that allow the protein to be promiscuous in nature and bind a variety of nucleic acids with varying affinities [41]. Although this structural characteristic of NCp7 would suggest that its interaction with host cell nucleic acids is non-specific in nature, HIV-1 viral progeny are specifically packaged with viral RNA, which only account for less than 1% of the total nucleic acids *in vivo* during the lytic phase of viral progression. [42] In addition to promoting viral RNA dimerization and encapsidation, NCp7 initiates the annealing of the primer human tRNA lysine 3 (htRNA^{Lys3}) for the process of reverse transcription by recruiting host cell htRNA^{Lys3} and rearranging the tRNA into the best conformation that favors reverse transcription [43-47]. The results of this study suggest that this measure of specificity and recognition is accomplished by sequence specificity and enhanced by the presence of post transcriptional modifications.

Previous studies performed on NCp7 recognition of nucleic acid targets have demonstrated NCp7's ability to discriminate based on sequence but fail to incorporate the influence of post transcriptional modifications [5,43,44,48]. Further analysis by NMR has revealed that NCp7 binds preferentially to the inside of the L-shape of the acceptor stem of tRNA that contain regions of G6-U67 and T54(A58) pairs, which are highly sensitive to protein catalysis [40]. The results from this study support these earlier observations by using the anticodon stem loop (ASL) domains of various tRNAs to determine sequence specificity. This study is also unique in

nature because of its incorporation of the post transcriptional modifications methylthio-N6-threonyl carbamoyladenosine ($ms^2t^6A_{37}$) at position 37 and 5-methoxycarbonylmethyl-2-thiouridine (mcm^5s^2U) at position 34, which are present on mature htRNA^{Lys3} *in vivo*. NCp7 was titrated with the ASL domains for htRNA^{Lys3}, htRNA$^{Lys1}t^6A_{37}$, htRNA^{Lys3} $mcm^5s^2U_{34}$; $ms^2t^6A_{37}$ and *E. coli* tRNA^{Val3} and the level of interaction determined by tryptophan fluorescence quenching. The data generated from the resulting quench analysis identified that NCp7 demonstrates an increased affinity for ASL domains that contain stems that are rich in guanine on the D side of the tRNA (data not shown). When modifications are present on the ASL region of htRNA^{Lys3} (ASL^{Lys3} $mcm^5s^2U_{34}$; $ms^2t^6A_{37}$) the binding between NCp7 and htRNA^{Lys3} is significantly enhanced. This observation is illustrated in Figure 3, where the K_d value for NCp7 and ASL^{Lys3} is 2.3 µM compared to the K_d value of 0.27 µM that is observed during the interaction between NCp7 and ASL^{Lys3} $mcm^5s^2U_{34}$; $ms^2t^6A_{37}$. A mechanism that would explain these observations is that NCp7 binds target nucleic acids in a sequence first, modification second manner. NCp7 initially binds the acceptor stem region of tRNA rich in guanine with great affinity allowing the protein to recognize the loop region and identify the presence of post transcriptional modifications. The presence of modifications allows for the protein to bind with increased affinity to the tRNA and destabilize the loop region, making the tRNA more susceptible to conformational manipulation.[12,49] Recent studies performed on the structural influence of post transcriptional modifications on tRNA also support this observation in which modifications can destabilize the loop region of tRNA and alter secondary structure. This mechanism provides another level of protein-nucleic acid recognition *in vivo* [12,16,49] and provides

the perfect model for NCp7's promiscuous nature when studied *in-vitro* but its high specificity *in vivo.*

Peptide Selection and Screening

The development of potential therapeutics that target the recruitment of nucleic acids by NCp7 requires the selection of peptides that demonstrate great sequence and modification specificity. A peptide library derived from nucleic acid interacting proteins was synthesized and initially screened by phage display. This mode of screening was ideal for modification recognition because of the rich chemistry that is involved with the modifications $mcm^5s^2U_{34}$ and $ms^2t^6A_{37}$ protein interaction. The binding characterization, involving the derivation of dissociation constants, of the 20 peptides isolated from the phage screen was performed using fluorescence quenching. The peptides were monitored for sequence and modification specificity under physiological conditions. Of the library of peptides, peptides 6 (P6) and 17 (P17) demonstrated the highest affinity for binding the anticodon stem loop (ASL) domain of modified htRNA^{Lys3} (ASL^{Lys3} $mcm^5s^2U_{34}$; $ms^2t^6A_{37}$). The peptides were synthesized with a FITC conjugated tag at the N-terminus. The peptides were initially screened for non-specific binding using the ASL of unmodified htRNA^{Lys3} (ASL^{Lys3}). P6 and P17 demonstrated the least amount of interaction with the ASL^{Lys3}. In contrast peptide 21 (P21) demonstrated a high affinity for binding ASL^{Lys3}. These results are illustrated in Figure 4, where the interaction between P21 and ASL^{Lys3} results in over 50% quenching of the fluorescence signal. In contrast only a 25% quench was observed for the interaction between both P6 and P17 with ASL^{Lys3}. Although many peptides were screened using ASL^{Lys3}, chemical synthesis of RNA with post transcriptional modifications is extremely difficult and cost prohibitive. For this reason only the peptides that

28

exhibited the least amount of interaction with ASL^{Lys3} were selected for modification specificity screening. These peptides were P6 and P17. P21 was also selected from this round of screening to serve as a control peptide.

The modification recognition screen was performed using the ASL domains of $htRNA^{Lys1}t^6A_{37}$ ($ASL^{Lys1}t^6A_{37}$) and $htRNA^{Lys3}$ (ASL^{Lys3} $mcm^5s^2U_{34}$; $ms^2t^6A_{37}$). The peptides were individually titrated with each modified tRNA under physiologically relevant conditions and monitored for fluorescence quenching. A correlation between the number of modifications present in the ASL region of the tRNA and the level of binding can clearly be inferred from the resulting quench curves illustrated in Figure 4. P6 exhibited the greatest level of modification recognition with a K_d value of 0.6 µM when titrated with the singly modified $ASL^{Lys1}t^6A_{37}$. When titrated with ASL^{Lys3} $mcm^5s^2U_{34}$; $ms^2t^6A_{37}$, the binding is enhanced further, producing a K_d value of 0.4 µM. P17 also demonstrated a high specificity for modification containing tRNA, with a K_d value of 1.7 µM and 0.6 µM when titrated with $ASL^{Lys1}t^6A_{37}$ and ASL^{Lys3} $mcm^5s^2U_{34}$; $ms^2t^6A_{37}$ respectively. In comparison to NCp7, neither P6 nor P17 bind as tightly to ASL^{Lys3} $mcm^5s^2U_{34}$; $ms^2t^6A_{37}$, but both peptides exhibit K_d values in the low µM –nM range (Table 1). When monitoring biological interactions *in vitro* using a FITC conjugated tag as opposed to a fluorescent residue like tryptophan, the calculated K_d values are not as precise. In the case of P6 and P17, the FITC tag is placed at the N-terminus of the peptide as opposed to the interior of the peptide, where the protein-RNA interaction generally occurs.

To determine the ability of P6 and P17 to exhibit sequence and modification specificity, the peptides were screened against *E.coli* ASL region $tRNA^{Val\ 3}$ unmodified (ASL^{Val3}) (Figure 6A) and the ASL region of *E.coli* $tRNA^{Val3}$ m^6A_{37} ($ASL^{Val3}m^6A_{37}$) (Figure 6B). Figure 7 is the

resulting quench curve from the titration of both $ASL^{Val3}m^6A_{37}$ and ASL^{Val3} with P6 and P17.

Analysis of the curve reveals that a very small amount of binding occurs in the case of ASL^{Val3}

titrated with P6 and P17. The same curve pattern is observed when $ASL^{Val3}m^6A_{37}$ is titrated with

P6 and P17. The results illustrate that P6 and P17 are able to differentiate between sequences as

well as native human tRNA post transcriptional modifications verses naturally occurring

bacterial modifications.

NCp7 and Peptide Functional Analysis by Circular Dichroism

The ability of P6 and P17 to discriminate between modified and unmodified $htRNA^{Lys3}$

make them prime candidates for development into HIV-1 viral therapeutics. The data obtained

from the fluorescence studies suggest that the binding of these peptides to target $htRNA^{Lys3}$ is

comparable to NCp7. NMR analysis of NCp7's interaction with $htRNA^{Lys3}$ suggest that the

positive residues surrounding the zinc knuckle regions of the protein interact with the $htRNA^{Lys3}$

acceptor stem region of the ASL [40]. As discussed previously, the synthesis of modified tRNA is

extremely expensive and cost prohibitive making it difficult to observe the same interaction in

the presence of post transcriptional modifications by NMR. Circular dichroism (CD) is a less

sensitive technique compared to NMR, but is sufficient in observing changes in secondary and

tertiary structures as the result of molecular interactions. The theorized mechanism for the

interaction between NCp7 and $htRNA^{Lys3}$ involves the protein denaturing the tRNA. For this

reason the interaction between NCp7 and $htRNA^{Lys3}$ was monitored in the near U.V. range using

CD where nucleic acid structure can be observed. The ASL domain of $htRNA^{Lys3}$ (ASL^{Lys3}) was

titrated with NCp7, with the CD set to scan from 235 nm to 285 nm encompassing both near and

far U.V.. An initial experiment was performed with the CD set to scan from 190 nm to 285 nm

to observe both NCp7 and ASL^{Lys3} during the interaction study. The experiment resulted in little to no change in the far U.V. region. Figure 8 represents the resulting CD spectra of ASL^{Lys3} titrated with NCp7. Analysis of the spectra reveals no significant structural change in the htRNA^{Lys3} when titrated with NCp7. There was a 15% decrease in elipticity when NCp7 was initially titrated into the sample of RNA, which can be explained by the sequence recognition observed in the fluorescence assay. At a ratio of 1 to 6, RNA to NCp7, the observed elipticity increases to 100%. The elipiticity remained constant at 100% when the ratio of RNA to NCp7 is pushed to 1:10. This observation suggests that the RNA reaches a saturation point at a ratio of 1:6 RNA to protein. The increase in elipticity was most likely the result of protein aggregation from the oversaturation of the ASL^{Lys3}.

Figure 9 represents the titration of htRNA^{Lys3} mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$ (ASL^{Lys3} mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37)}$ with NCp7 under the same conditions as ASL^{Lys3}. An overall 40% decrease in elipticity was observed at a 1:1 ratio of NCp7 to ASL^{Lys3} mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$. Further titration of the RNA with NCP7 resulted in a steady increase in elipticity until the RNA mostly likely reaches a saturation point of 1:6, where the elipticity increases from 60% to 80% and remains constant even at a ratio of 1:10, RNA to protein. These results support the theory that NCp7 initially binds to the RNA in a sequence specific manner, then uses the presence of modifications to alter the structure of the ASL region of the tRNA^{Lys3} into the most favorable conformation that aids in the initiation of reverse transcription. Once the RNA reaches saturation, it is mostly likely in the proper conformation for recognition by reverse transcriptase.

Peptides 6 and 17 were observed by fluorescence quenching assays to exhibit modification specificity and bind effectively to modified htRNA^{Lys3} with similar affinity to

NCp7. A titration experiment monitored by CD under the same conditions as NCp7 and ASL^{Lys3} was conducted using peptide 6 and 17 as titrant. Figures 10 and 11 illustrate that both peptides 6 and 17 do not drastically alter the structure of ASL^{Lys3} when modifications are not present. In Figure 11 it is apparent that P17 increases the elipticity of ASL^{Lys3} by 20% after the initial titration, but then remains constant despite the addition of more peptide. This is significant because the data suggest that P17 is functioning to enhance the stability of the ASL domain as opposed to NCp7's primary function of melting the RNA.

Titration of ASL^{Lys3} mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$ with P6 and P17 revealed that both peptides were able to exhibit the same ability as NCp7 to alter the conformation of htRNA^{Lys3}. P6 demonstrated the ability to completely melt the RNA. Figure 12 illustrates a complete 100% loss in elipticity. This observation indicates that P6 recognizes the modifications in the ASL domain of htRNA^{Lys3} and completely destabilizes the loop. Analysis of Figure 13 reveals that P17 functions in the opposite manner as P6. P17 significantly enhances the structural stability of ASL^{Lys3} mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$ when it is initially titrated into solution resulting in an increase from 100% to 160% elipticity. Further titration of ASL^{Lys3} mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$ by P17 resulted in a drop in elipticity back to 100%. This observation suggests that P17 binds to the RNA at a different site compared to P6 and NCp7. This is supported by the fact that the sequence for P17 is significantly different containing arginine, lysine, and histidine compared to P6 which is composed of only arginine and histidine residues.

NCp7 is a 71 amino acid long protein in comparison to peptides 6 and 17 which are 15 and 16 residues in length. In the NCp7 and htRNA^{Lys3} study by CD, NCp7 demonstrated the ability to recognize the presence of modifications and alter the conformation of ASL^{Lys3}

mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$ in comparison to its interaction with ASL^{Lys3}. Peptides 6 and 17 demonstrated a similar ability to manipulate the structure of the ASL domain of the htRNA^{Lys3} when they were titrated into solution containing ASL^{Lys3} mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$ and monitored by CD.

Chapter 5
Conclusion

Conclusion

The general structure of all tRNAs consist of a D loop, T loop and an anticodon stem loop (ASL) domain. The nucleosides contained within the ASL domain and the conformation of the tRNA, as a result of loop dynamics are how different species of tRNA are identified *in vivo* by RNA binding proteins. A well characterized example of this are aminoacyl-tRNA synthetases, which use tRNA structure and nucleoside sequence to determine proper aminoacylation [50]. In terms of modifications, t^6A_{37} chemistry in combination with its influence over the conformation of the anticodon domain of host tRNA, was determined to be a key recognition factor for yeast tRNA phenylalanine[30].

In general, there are two types of post-transcriptional modifications [51]. The first is a set of modifications that are commonly found and generally located at identical or similar residues in RNAs, such as Ψ, 2′-O-methyl nucleosides. These modifications participate in RNA structure and conformational stability [17,52-54]. Additionally, the modifications, m^2G and m^2_2G, D, m^5C, T, and m^1A are important to tRNA folding and thermal stability [55-57]. A second set of modifications are those that are distinct in nature. These modifications are sometimes found as as a single modification or in combination, allowing for enhanced variability among RNA. An example of this is 1-methyl-3-(3-amino-3-carboxypropyl) pseudouridine, $m^1acp^3\Psi$, of the eukarya rRNAs. More particularly, the modifications of tRNA's anticodon region are distinctive. The derivatives of 2-thiouridines at the wobble position 34 ($*s^2U_{34}$) are limited to certain amino acid accepting

34

species [58] and those tRNAs with an N6-threonyladensine at position 37 (*t^6A_{37}), 3′-adjacent to the anticodon, have a U at position 36. Both *s^2U_{34} and *t^6A_{37} are important for cognate and wobble codon recognition on the ribosome, [24] and for maintaining the translational reading frame [59]. In general modifications have been proven to be essential for protein-RNA recognition.

The introduction of post-transcriptional modifications to the ASL domain of tRNAs in most cases reduces thermal stability by disrupting the hydrogen bonding dynamics within the loop. As a direct result of this, the loop becomes more dynamic and more susceptible to base stacking interactions [12,49]. In the case of NCp7, the presence of post transcriptional modifications lowers the thermal stability of the ASL domain, making it energetically favorable for the protein to alter the ASL domain of htRNA^{Lys3} conformation. Though the peptides are less than a third of the size of NCp7, they may have some ability to unfold or remodel the ASL conformation. NCp7's ability to bind and manipulate tRNA^{Lys3} for successful viral propagation, and the mechanism surrounding its interaction with the tRNA make it a prime target for the development of antiviral therapeutics. In mimicking the binding of NCp7 and perhaps other proteins involved in HIV's recruitment of htRNA^{Lys3} for the priming of reverse transcriptase, the peptides have the potential to be tools in understanding HIV replication and may also provide insight to the design of novel therapeutics.

Preliminary Results and Future Experiments

The dissociation constant (K_d) value derived from the fluorescence quenching assay of ASL^{Lys3} $mcm^5s^2U_{34}$; $ms^2t^6A_{37}$ and peptides 6 and 17, suggest that these peptides could potentially be developed into therapeutics that would compete with NCp7 for interaction with $htRNA^{Lys3}$ in HIV-1 infected cells. To test this theory, a preliminary competition assay was performed by fluorescence quenching of the tryptophan residue of NCp7, in which the NCp7 was saturated with ASL^{Lys3} $mcm^5s^2U_{34}$; $ms^2t^6A_{37}$ reducing the fluorescent signal to its lowest point. The mixture was then titrated with P6, with the idea that P6 would compete off the ASL^{Lys3} $mcm^5s^2U_{34}$; $ms^2t^6A_{37}$, resulting in an increase of fluorescence signal from the tryptophan. When the experiment was conducted, the tryptophan signal actually decreased further. This result suggests that peptide 6 interacts with NCp7 in addition to interacting with ASL^{Lys3} $mcm^5s^2U_{34}$; $ms^2t^6A_{37}$. Examination of the sequence of P6 could explain this observation. P6 contains 3 histidine residues that may bind to the zinc present in the zinc knuckles of NCp7. Material constraints prevented the conduction of a competition experiment using P17. The sequence of P17 only consists of one histidine, which suggests P17 may be able to compete off the RNA without any interaction with NCp7.

The observation that P6 can bind to NCp7 in addition to the structural data generated from the CD experiments suggest that a more sensitive technique should be used to characterize the peptide residues involved in the interaction between the peptides and ASL^{Lys3} $mcm^5s^2U_{34}$; $ms^2t^6A_{37}$ as well as peptide and NCp7. Mass spectrometry experiments were conducted with the specific aim of observing complex formation between the peptides and modified nucleosides.

ESI-MS is an analytical mass spectrometry technique that can be used to observe non-covalent complex formation between nucleic acid and proteins in solution, using nM concentrations [60]. For this reason ESI-MS is the perfect and most cost effective technique to screen peptide 6 for its interaction with $ASL^{Lys3}mcm^5s^2U_{34};ms^2t^6A_{37}$. A similar methodology has been used in which nucleic acid ligands were screened for their interaction with HIV-1 related proteins [61]. As a control experiment unbound ASL^{Lys3} was analyzed by ESI-MS, to produce a baseline of ASL^{Lys3}, which resulted in a spectra indicating a -4 charge state for the nucleic acid. The RNA readily fragmented providing complete sequence coverage. Peptide 6 was analyzed under the same conditions, then titrated with both modified and unmodified ASL^{Lys3}. Analysis and comparison of the resulting spectra led to the observation of a loss of the peptide along with the appearance of the unbound -6 charge state of the RNA, consistent with the peptide having a +2 charge, when peptide 6 interacts with $ASL^{Lys3}mcm^5s^2U_{34};ms^2t^6A_{37}$. The shift in charge state did not occur on the spectra produced from peptide 6 interaction with ASL^{Lys3}, suggesting that shift in charge state is a result of complex formation.

Reference List

1. Freed, E. O. HIV-1 Gag Proteins: Diverse Functions in the Virus Life Cycle. *Virology* **1998,** *251* (1), 1-15.

2. Coffin, J.; S.Hughes; H.Varmus *Retroviruses;* Cold Spring Harbor Press: Plainview, NY, 1997.

3. Coren, L. V.; Thomas, J. A.; Chertova, E.; Sowder, R. C.; Gagliardi, T. D.; Gorelick, R. J.; Ott, D. E. Mutational analysis of the C-terminal gag cleavage sites in human immunodeficiency virus type 1. *J Virol* **2007,** *81* (18), 10047-10054.

4. Thomas, J. A.; Gorelick, R. J. Nucleocapsid protein function in early infection processes. *Virus Res* **2008,** *134* (1-2), 39-63.

5. Levin, J. G.; Guo, J.; Rouzina, I.; Musier-Forsyth, K. Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism. *Prog. Nucleic Acid Res Mol Biol* **2005,** *80*, 217-286.

6. Qualley, D. F.; Stewart-Maynard, K. M.; Wang, F.; Mitra, M.; Gorelick, R. J.; Rouzina, I.; Williams, M. C.; Musier-Forsyth, K. C-terminal domain modulates the nucleic acid chaperone activity of human T-cell leukemia virus type 1 nucleocapsid protein via an electrostatic mechanism. *J Biol Chem* **2010,** *285* (1), 295-307.

7. Williams, M. C.; Rouzina, I.; Wenner, J. R.; Gorelick, R. J.; Musier-Forsyth, K.; Bloomfield, V. A. Mechanism for nucleic acid chaperone activity of HIV-1 nucleocapsid protein revealed by single molecule stretching. *Proc Natl Acad Sci U. S. A* **2001,** *98* (11), 6121-6126.

8. Berkowitz, R.; Fisher, J.; Goff, S. P. RNA packaging. *Curr Top Microbiol Immunol* **1996,** *214*, 177-218.

9. Guo, J.; Wu, T.; Anderson, J.; Kane, B. F.; Johnson, D. G.; Gorelick, R. J.; Henderson, L. E.; Levin, J. G. Zinc Finger Structures in the Human Immunodeficiency Virus Type 1 Nucleocapsid Protein Facilitate Efficient Minus- and Plus-Strand Transfer. *J. Virol.* **2000,** *74* (19), 8980-8988.

10. Morellet, N.; de Rocquigny, H.; Mθly, Y.; Jullian, N.; Dθmθnθ, H.; Ottmann, M.; Gθrard, D.; Darlix, J. L.; Fournie-Zaluski, M. C.; Roques, B. P. Conformational behaviour of the active and inactive forms of the nucleocapsid NCp7 of HIV-1 studied by 1H NMR. *Journal of Molecular Biology* **1994,** *235* (1), 287-301.

11. Poon, D. T.; Wu, J.; Aldovini, A. Charged amino acid residues of human immunodeficiency virus type 1 nucleocapsid p7 protein involved in RNA packaging and infectivity. *J. Virol.* **1996,** *70* (10), 6607-6616.

12. Agris, P. F.; Malkiewicz, A.; Guenther, R.; Basti, M.; Sengupta, R.; Stuart, J. Importance of modified nucleosides to the structure and function of tRNAs. *Nucleic Acids Symp. Ser.* **1995,** (33), 254-255.

13. Nobles, K. N.; Yarian, C. S.; Liu, G.; Guenther, R. H.; Agris, P. F. Highly conserved modified nucleosides influence Mg2+-dependent tRNA folding. *Nucleic Acids Res* **2002,** *30* (21), 4751-4760.

14. Gustilo, E. M.; Vendeix, F. A.; Agris, P. F. tRNA's modifications bring order to gene expression. *Curr Opin Microbiol* **2008,** *11* (2), 134-140.

15. Helm, M. Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic Acids Res* **2006,** *34* (2), 721-733.

16. Vendeix, F. A.; Dziergowska, A.; Gustilo, E. M.; Graham, W. D.; Sproat, B.; Malkiewicz, A.; Agris, P. F. Anticodon domain modifications contribute order to tRNA for ribosome-mediated codon binding. *Biochemistry* **2008,** *47* (23), 6117-6129.

17. Agris, P. F.; Vendeix, F. A. P.; Graham, W. D. tRNA's Wobble Decoding of the Genome: 40 Years of Modification. *Journal of Molecular Biology* **2007,** *366* (1), 1-13.

18. Bilbille, Y.; Vendeix, F. A.; Guenther, R.; Malkiewicz, A.; Ariza, X.; Vilarrasa, J.; Agris, P. F. The structure of the human tRNALys3 anticodon bound to the HIV genome is stabilized by modified nucleosides and adjacent mismatch base pairs. *Nucleic Acids Res.* **2009,** *37* (10), 3342-3353.

19. Auxilien, S.; Keith, G.; Le Grice, S. F.; Darlix, J. L. Role of post-transcriptional modifications of primer tRNALys,3 in the fidelity and efficacy of plus strand DNA transfer during HIV-1 reverse transcription. *J Biol Chem* **1999,** *274* (7), 4412-4420.

20. Cusack, S. RNA-protein complexes. *Curr Opin Struct. Biol* **1999,** *9* (1), 66-73.

21. Eshete, M.; Marchbank, M. T.; Deutscher, S. L.; Sproat, B.; Leszczynska, G.; Malkiewicz, A.; Agris, P. F. Specificity of phage display selected peptides for modified anticodon stem and loop domains of tRNA. *Protein J.* **2007,** *26* (1), 61-73.

22. Agris, P. F.; Marchbank, M. T.; Newman, W.; Guenther, R.; Ingram, P.; Swallow, J.; Mucha, P.; Szyk, A.; Rekowski, P.; Peletskaya, E.; Deutscher, S. L. Experimental models of protein-RNA interaction: isolation and analyses of tRNA(Phe) and U1 snRNA-binding peptides from bacteriophage display libraries. *J Protein Chem* **1999,** *18* (4), 425-435.

23. Mucha, P.; Szyk, A.; Rekowski, P.; Guenther, R.; Agris, P. F. Interaction of RNA with phage display selected peptides analyzed by capillary electrophoresis mobility shift assay. *RNA.* **2002,** *8* (5), 698-704.

24. Murphy, F. V.; Ramakrishnan, V.; Malkiewicz, A.; Agris, P. F. The role of modifications in codon discrimination by tRNA(Lys)UUU. *Nat. Struct. Mol. Biol.* **2004,** *11* (12), 1186-1191.

25. Scaringe, S. A.; Kitchen, D.; Kaiser, R. J.; Marshall, W. S. Preparation of 5'-silyl-2'-orthoester ribonucleosides for use in oligoribonucleotide synthesis. *Curr Protoc. Nucleic Acid Chem* **2004,** *Chapter 2*, Unit.

26. Agris, P. F.; Malkiewicz, A.; Kraszewski, A.; Everett, K.; Nawrot, B.; Sochacka, E.; Jankowska, J.; Guenther, R. Site-selected introduction of modified purine and pyrimidine ribonucleosides into RNA by automated phosphoramidite chemistry. *Biochimie* **1995,** *77* (1-2), 125-134.

27. Gehrke, C. W.; Kuo, K. C.; McCune, R. A.; Gerhardt, K. O.; Agris, P. F. Quantitative enzymatic hydrolysis of tRNAs: reversed-phase high-performance liquid chromatography of tRNA nucleosides. *J Chromatogr.* **1982,** *230* (2), 297-308.

28. Ogilvie, K. K.; Usman, N.; Nicoghosian, K.; Cedergren, R. J. Total chemical synthesis of a 77-nucleotide-long RNA sequence having methionine-acceptance activity. *Proc Natl Acad Sci U. S. A* **1988,** *85* (16), 5764-5768.

29. Agris, P. F.; Malkiewicz, A.; Kraszewski, A.; Everett, K.; Nawrot, B.; Sochacka, E.; Jankowska, J.; Guenther, R. Site-selected introduction of modified purine and pyrimidine ribonucleosides into RNA by automated phosphoramidite chemistry. *Biochimie* **1995,** *77* (1-2), 125-134.

30. Mucha, P.; Szyk, A.; Rekowski, P.; Weiss, P. A.; Agris, P. F. Anticodon domain methylated nucleosides of yeast tRNA(Phe) are significant recognition determinants in the binding of a phage display selected peptide. *Biochemistry* **2001,** *40* (47), 14191-14199.

31. Agris, P. F.; Marchbank, M. T.; Newman, W.; Guenther, R.; Ingram, P.; Swallow, J.; Mucha, P.; Szyk, A.; Rekowski, P.; Peletskaya, E.; Deutscher, S. L. Experimental models of protein-RNA interaction: isolation and analyses of tRNA(Phe) and U1 snRNA-binding peptides from bacteriophage display libraries. *J Protein Chem* **1999,** *18* (4), 425-435.

32. Guenther, R. H.; Hardin, C. C.; Sierzputowska-Gracz, H.; Dao, V.; Agris, P. F. A magnesium-induced conformational transition in the loop of a DNA analog of the yeast tRNA(Phe) anticodon is dependent on RNA-like modifications of the bases of the stem. *Biochemistry* **1992,** *31* (45), 11004-11011.

33. Jones, C. I.; Spencer, A. C.; Hsu, J. L.; Spremulli, L. L.; Martinis, S. A.; DeRider, M.; Agris, P. F. A counterintuitive Mg2+-dependent and modification-assisted functional folding of mitochondrial tRNAs. *J Mol Biol* **2006,** *362* (4), 771-786.

34. Druillennec, S.; Caneparo, A.; de, R. H.; Roques, B. P. Evidence of interactions between the nucleocapsid protein NCp7 and the reverse transcriptase of HIV-1. *J. Biol. Chem.* **1999,** *274* (16), 11283-11288.

35. Abbink, T. E. M.; Berkhout, B. HIV-1 reverse transcription initiation: A potential target for novel antivirals? *Virus Research* **2008,** *134* (1-2), 4-18.

36. Yarian, C.; Marszalek, M.; Sochacka, E.; Malkiewicz, A.; Guenther, R.; Miskiewicz, A.; Agris, P. F. Modified nucleoside dependent Watson-Crick and wobble codon binding by tRNALysUUU species. *Biochemistry* **2000,** *39* (44), 13390-13395.

37. Yarian, C.; Townsend, H.; Czestkowski, W.; Sochacka, E.; Malkiewicz, A. J.; Guenther, R.; Miskiewicz, A.; Agris, P. F. Accurate translation of the genetic code depends on tRNA modified nucleosides. *J Biol Chem* **2002,** *277* (19), 16391-16395.

38. Weixlbaumer, A.; Murphy, F. V.; Dziergowska, A.; Malkiewicz, A.; Vendeix, F. A. P.; Agris, P. F.; Ramakrishnan, V. Mechanism for expanding the decoding capacity of transfer RNAs by modification of uridines. *Nat Struct Mol Biol* **2007,** *14* (6), 498-502.

39. Tisnθ, C.; Roques, B. P.; Dardel, F. Specific recognition of primerátRNALys3 by HIV-1 nucleocapsid protein: involvement of the zinc fingers and the N-terminal basic extension. *Biochimie* **2003,** *85* (5), 557-561.

40. Tisnθ, C.; Roques, B. P.; Dardel, F. Heteronuclear NMR studies of the interaction of tRNA3Lys with HIV-1 nucleocapsid protein. *Journal of Molecular Biology* **2001,** *306* (3), 443-454.

41. Guo, J.; Wu, T.; Anderson, J.; Kane, B. F.; Johnson, D. G.; Gorelick, R. J.; Henderson, L. E.; Levin, J. G. Zinc Finger Structures in the Human Immunodeficiency Virus Type 1 Nucleocapsid Protein Facilitate Efficient Minus- and Plus-Strand Transfer. *J. Virol.* **2000,** *74* (19), 8980-8988.

42. Poon, D. T.; Wu, J.; Aldovini, A. Charged amino acid residues of human immunodeficiency virus type 1 nucleocapsid p7 protein involved in RNA packaging and infectivity. *J. Virol.* **1996,** *70* (10), 6607-6616.

43. Fisher, R. J.; Rein, A.; Fivash, M.; Urbaneja, M. A.; Casas-Finet, J. R.; Medaglia, M.; Henderson, L. E. Sequence-specific binding of human immunodeficiency virus type 1 nucleocapsid protein to short oligonucleotides. *J Virol* **1998,** *72* (3), 1902-1909.

44. Tsuchihashi, Z.; Brown, P. O. DNA strand exchange and selective DNA annealing promoted by the human immunodeficiency virus type 1 nucleocapsid protein. *J Virol* **1994,** *68* (9), 5863-5870.

45. Cristofari, G.; Darlix, J. L. The ubiquitous nature of RNA chaperone proteins. *Prog. Nucleic Acid Res Mol Biol* **2002,** *72,* 223-268.

46. Herschlag, D. RNA chaperones and the RNA folding problem. *J Biol Chem* **1995,** *270* (36), 20871-20874.

47. Bampi, C.; Jacquenet, S.; Lener, D.; Decimo, D.; Darlix, J. L. The chaperoning and assistance roles of the HIV-1 nucleocapsid protein in proviral DNA synthesis and maintenance. *Curr HIV Res* **2004,** *2* (1), 79-92.

48. Vuilleumier, C.; Bombarda, E.; Morellet, N.; Gerard, D.; Roques, B. P.; Mely, Y. Nucleic acid sequence discrimination by the HIV-1 nucleocapsid protein NCp7: a fluorescence study. *Biochemistry* **1999,** *38* (51), 16816-16825.

49. Vendeix, F. A. P.; Dziergowska, A.; Gustilo, E. M.; Graham, W. D.; Sproat, B.; Malkiewicz, A.; Agris, P. F. Anticodon Domain Modifications Contribute Order to tRNA for Ribosome-Mediated Codon BindingГÇáГÇí. *Biochemistry* **2008,** *47* (23), 6117-6129.

50. Beuning, P. J.; Musier-Forsyth, K. Transfer RNA recognition by aminoacyl-tRNA synthetases. *Biopolymers* **1999,** *52* (1), 1-28.

51. Rozenski, J.; Crain, P. F.; McCloskey, J. A. The RNA Modification Database: 1999 update. *Nucl. Acids Res.* **1999,** *27* (1), 196-197.

52. Durant, P. C.; Davis, D. R. Stabilization of the anticodon stem-loop of tRNALys,3 by an A+-C base-pair and by pseudouridine. *Journal of Molecular Biology* **1999,** *285* (1), 115-131.

53. Helm, M. Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucl. Acids Res.* **2006,** *34* (2), 721-733.

54. Ishitani, R.; Nureki, O.; Nameki, N.; Okada, N.; Nishimura, S.; Yokoyama, S. Alternative Tertiary Structure of tRNA for Recognition by a Posttranscriptional Modification Enzyme. *Cell* **2003,** *113* (3), 383-394.

55. Kowalak, J. A.; Dalluge, J. J.; McCloskey, J. A.; Stetter, K. O. The Role of Posttranscriptional Modification in Stabilization of Transfer RNA from Hyperthermophiles. *Biochemistry* **1994,** *33* (25), 7869-7876.

56. Ginell, S. L.; Parthasarathy, R. Conformation of N2-methylguanosine, A modified nucleoside of tRNA. *Biochemical and Biophysical Research Communications* **1978,** *84* (4), 886-894.

57. Stuart, J. W.; Gdaniec, Z.; Guenther, R.; Marszalek, M.; Sochacka, E.; Malkiewicz, A.; Agris, P. F. Functional anticodon architecture of human tRNALys3 includes disruption of intraloop hydrogen bonding by the naturally occurring amino acid modification, t6A. *Biochemistry* **2000,** *39* (44), 13396-13404.

58. Chan, P. P.; Lowe, T. M. GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucl. Acids Res.* **2009,** *37* (suppl_1), D93-D97.

59. Bjork, G. R.; Ericson, J. U.; Gustafsson, C. E.; Hagervall, T. G.; Jonsson, Y. H.; Wikstrom, P. M. Transfer RNA modification. *Annu. Rev Biochem* **1987,** *56*, 263-287.

60. Colgrave, M. L.; Beck, J. L.; Sheil, M. M.; Searle, M. S. Electrospray ionisation mass spectrometric detection of weak non-covalent interactions in nogalamycin-DNA complexes. *Chem Commun (Camb.)* **2002,** (6), 556-557.

61. Turner, K. B.; Brinson, R. G.; Yi-Brunozzi, H. Y.; Rausch, J. W.; Miller, J. T.; Le Grice, S. F.; Marino, J. P.; Fabris, D. Structural probing of the HIV-1 polypurine tract RNA:DNA hybrid using classic nucleic acid ligands. *Nucleic Acids Res* **2008,** *36* (8), 2799-2810.