Published in final edited form as:
J Mol Biol. 2011 July 22; 410(4): 698–715. doi:10.1016/j.jmb.2011.04.025.

Functional Recognition of the Modified Human tRNA$^{Lys3}_{UUU}$ Anticodon Domain by HIV's Nucleocapsid Protein and a Peptide Mimic

William D. Graham[1,†], **Lise Barley-Maloney**[1,†], **Caren J. Stark**[2,†], **Amarpreet Kaur**[1], **Khrystyna Stolyarchuk**[1], **Brian Sproat**[3], **Grazyna Leszczynska**[4], **Andrzej Malkiewicz**[4], **Nedal Safwat**[1], **Piotr Mucha**[5], **Richard Guenther**[1], and **Paul F. Agris**[2,*]

[1] Molecular and Structural Biochemistry, North Carolina State University, Raleigh NC 27695, USA [2] Te RNA Institute, Biological Sciences, University at Albany, Albany, NY 12222, USA [3] Integrated DNA Technologies BVBA, Interleuvenlaan 12A, B-3001 Leuven, Belgium [4] Institute of Organic Chemistry, Technical University, Żeromskiego 116, 90-924, ŁódŸ, Poland [5] Department of Chemistry, University of Gdansk, Sobieskiego 18, 80-952 Gdansk, Poland

Abstract

The HIV-1 nucleocapsid protein, NCp7, facilitates the use of human tRNA$^{Lys3}_{UUU}$ as the primer for reverse transcription. NCp7 also remodels the htRNA's amino acid accepting stem and anticodon domains in preparation for their being annealed to the viral genome. To understand the possible influence of the htRNA's unique composition of post-transcriptional modifications on NCp7 recognition of htRNA$^{Lys3}_{UUU}$, the protein's binding and functional remodeling of the human anticodon stem and loop domain (hASL^{Lys3}) were studied. NCp7 bound the hASL$^{Lys3}_{UUU}$ modified with 5-methoxycarbonyl methyl-2-thiouridine at position-34 (mcm^{5}s^{2}U$_{34}$) and 2-methylthio-N^6-threonylcarbamoyladenosine at position-37 (ms^{2}t^{6}A$_{37}$) with a considerably higher affinity than the unmodified hASL$^{Lys3}_{UUU}$ (K_d = 0.28 ± 0.03 and 2.30 ± 0.62 μM, respectively). NCp7 denatured the structure of the hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$;ψ$_{39}$ more effectively than that of the unmodified hASL$^{Lys3}_{UUU}$. Two 15 amino acid peptides selected from phage display libraries demonstrated a high affinity (average K_d = 0.55 ± 0.10 μM) and specificity for the ASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$ comparable to that of NCp7. The peptides recognized a t^{6}A$_{37}$-modified ASL with an affinity (K_d = 0.60 ± 0.09 μM) comparable to that for hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$, indicating a preference for the t^{6}A$_{37}$ modification. Significantly, one of the peptides was capable of relaxing the hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$;ψ$_{39}$ structure in a manner similar to that of NCp7, and therefore could be used to further study protein recognition of

*Corresponding authors present address: P.F. Agris, The RNA Institute, University at Albany-SUNY, 1400 Washington Ave., Albany, NY 12222, USA; Tel: 518-437-4448; FAX: 518-437-4456; PAgris@albany.edu.
†These authors contributed equally to the research. For W.D.G., the work is in partial fulfillment of the graduate degree.
Present Addresses:
L. Barley-Maloney: BD Diagnostics – Women's Health and Cancer, 21 Davis Dr., RTP, NC 277709, USA; Lise_Barley-Maloney@bd.com
B. Sproat: Chemconsilium GCV, Jaarmarktstraat 48, 2221, Booischot, Belgium; fa205401@skynet.be
N. Safwat: Clinical Marketing – Microbiology, bioMerieux Inc., 100 Rodolphe St., Durham, NC 27712, USA; nedal.safwat@biomerieux.com
R. Guenther: Department of Plant Pathology, North Carolina State University, Raleigh, NC 27695, USA; Dick_Guenther@ncsu.edu

RNA modifications. The post-transcriptional modifications of htRNA$^{Lys3}_{UUU}$ have been found to be important determinants of NCp7's recognition prior to the tRNA$^{Lys3}_{UUU}$ being annealed to the viral genome as the primer of reverse transcription.

Keywords

Modified nucleosides; *N*^{6}*-threonylcarbamoyladenosine*; 5-methoxycarbonylmethyl-2-thiouridine; NCp7; anticodon stem and loop

Introduction

The lentivirus, Human Immunodeficiency Virus type 1 (HIV-1), is a highly adaptive virus that uses a complex system of proteolytic cleavages of its polyprotein precursor Pr55^{Gag} (Gag), to replicate and integrate its genome into host cell chromosomes. The Gag polyprotein contains two zinc finger motifs that are 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 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 one of the most vital proteins produced from this cleavage is the nuclear capsid protein 7 (NC or NCp7) (recently reviewed in reference 4).[1,2,4,5] NCp7 is a strongly basic, 55 amino acid protein containing two shortened zinc fingers (knuckles) (Figure 1a). NCp7 functions as a nucleic acid chaperone during various stages of viral progression.[4,6] Both theoretical and empirical studies indicate that approximately 1400 copies of NCp7 coat each of the two copies of HIV's RNA viral genome in the capsid.[7–9]

NCp7's chaperone activity is characterized by its ability to destabilize RNA structure, facilitating a conformational rearrangement, or remodeling, that is critical to the replication of HIV-1. NCp7 is involved in the specific use of the human host cell lysine tRNA isoacceptor 3, htRNA$^{Lys3}_{UUU}$, as the primer of HIV reverse transcription.[10,11] The htRNA$^{Lys3}_{UUU}$ with the anticodon UUU is one of some 45 tRNAs, and the minor species of three lysine tRNAs in human cells. A conformational alteration of its canonical L-shaped tRNA structure is required for htRNA$^{Lys3}_{UUU}$ to function as the primer for reverse transcription.[12] In facilitating the annealing of the primer to the HIV genomic RNA, NCp7 recognizes, binds and destabilizes the htRNA$^{Lys3}_{UUU}$.[4] The Watson-Crick hydrogen bonding of the 3′-terminal nucleotides and that of the U-rich anticodon stem and loop (ASL) are destabilized and annealed to the complementary 18 nucleotides of the viral RNA's primer binding site (PBS)[13,14] and approximately 12 nucleotides of the A-rich bulge of Loop 1,[15,16] respectively (Figure 1b). The htRNA$^{Lys3}_{UUU}$ is packaged along with tRNA^{Lys1,2},[18,19] the human lysyl-tRNA synthetase (KRS)[20] and the viral genome into the newly formed virions and used as the primer during a subsequent infection.[4] NCp7's ability to bind and manipulate both the viral genomic RNA and host cell's htRNA^{Lys3} for successful viral propagation make it a primary target for the development of antiviral therapeutics.[21]

NCp7 binds to RNAs in an ionic strength dependent manner that is characterized by a range of affinities.[22] The protein has a demonstrated preference for single stranded polyGs, especially in GU (and GT) stretches,[23] and appears to melt G•U wobble base pairs, first.[6,24,25] In its recognition and binding of htRNA$^{Lys3}_{UUU}$, NCp7 exhibits its highest affinity for the amino acid acceptor stem.[4,26] As many as eight to ten copies of the protein are thought to bind to the stem and loop regions of the tRNA in a non-specific manner.[27–29] In addition to the two zinc fingers, NCp7 has a number of positively charged amino acids that are suggestive of how NCp7 could attach to any RNA, including the different regions

and conformations of htRNA$^{Lys3}_{UUU}$.[4,5,30–32] Sequence analysis of RNA-binding proteins has led to the identification of specific motifs that are responsible for protein recognition of RNA structure.[33] Many of these motifs contain positively charged residues, such as arginine, that interact in a non-specific manner with the negatively charged backbone of the RNA. However, any degree of specificity of a protein's recognition of a particular RNA, such as NCp7 for htRNA$^{Lys3}_{UUU}$, would be dependent on the RNA's nucleotide sequence and resulting conformation. Post-transcriptional modifications alter the conformation and dynamics of the RNA structure, specifically the loop regions of tRNAs where most modifications are found, thereby creating conformations that are favorable for a specific protein-tRNA interaction.[34] The modified nucleosides of htRNA$^{Lys3}_{UUU}$ may affect NCp7 recognition of and binding to the tRNA.

Mature htRNA$^{Lys3}_{UUU}$ contains the naturally occurring post-transcriptional modifications 5-methoxycarbonylmethyl-2-thiouridine (mcm^{5}s^{2}U$_{34}$) at wobble position 34 and 2-methylthio-*N*^{6}-threonylcarbamoyladenosine (ms^{2}t^{6}A$_{37}$) at position 37 within the tRNA's anticodon stem and loop domain (hASL$^{Lys3}_{UUU}$; Figure 1c,d). Pseudouridine, ψ_{39} is located at position 39, 3′ to the anticodon loop. Post-transcriptional modifications within the anticodon domain are identity elements that are recognized by the cognate aminoacyl-tRNA synthetases for some tRNAs and are important for effective aminoacylation of the tRNA for many, particularly for those tRNAs having position 34 modified uridines (glutamic acid, glutamine and lysine).[35,36] The post-transcriptional modifications of U$_{34}$ and A$_{37}$ increase the affinity of the hASL$^{Lys3}_{UUU}$ to bind the AAA and AAG codons on the ribosomal-bound mRNA,[37,38] and their bacterial counterparts increase the efficiency of translation and maintain the translational reading frame.[39] Thus, the naturally occurring post-transcriptional modifications present on htRNA$^{Lys3}_{UUU}$ may assist in the recognition of htRNA$^{Lys3}_{UUU}$ by NCp7, and other HIV-1 proteins, though the importance of these modifications to the tRNA's recognition by HIV proteins has not been explored. The ability of NCp7 to disrupt the structure of the htRNA$^{Lys3}_{UUU}$ anticodon domain for subsequent annealing to the A-rich bulge of the viral Loop 1 sequence could also be facilitated by the presence of post-transcriptional modifications in the ASL. Anticodon domain modifications are characterized, generally, by their ability to maintain an open ASL loop and yet enhance base stacking for codon binding.[40,41] As a consequence, the open loop would be more susceptible to conformational change by NCp7.

We have investigated the relationship between post-transcriptional modifications and the recognition of htRNA$^{Lys3}_{UUU}$ by NCp7. Here, we report that the anticodon domain modifications considerably enhanced NCp7 binding affinity for the hASL$^{Lys3}_{UUU}$. Modifications facilitated NCp7 disruption of the ASL's ordered conformation as a prelude to the structural rearrangement that precedes its annealing to the HIV genome. In addition, we describe the selection and characterization of a 15 amino acid peptide that both mimics the NCp7 recognition of the modified hASL$^{Lys3}_{UUU}$, and is also capable of disrupting the RNA structure.

RESULTS

NCp7 recognition of the modified anticodon domain of htRNA$^{Lys3}_{UUU}$

NCp7 binds to different RNAs to varying degrees.[4,6] Studies of NCp7 binding to RNA suggest that it binds to single stranded regions of RNA with little specificity.[4,6] When interacting with the unmodified transcript of the htRNA$^{Lys3}_{UUU}$ and stable isotope labeled tRNA expressed in *Escherichia coli*, NCp7 appears to bind initially to the nucleotides at the position of the tertiary interactions at the center of the L-shaped structure from which the melting of acceptor stem/T-stem helices and anticodon domain would occur.[4,24,42,43] The induced conformational change and dynamics now favor the annealing of the disrupted

acceptor and T stems to the complementary PBS and the tRNA's U-rich anticodon domain to the virus' A-rich bulge of the Loop 1 sequence.[15,16] In order to examine in some detail the extent to which the naturally occurring modifications within the $hASL^{Lys3}_{UUU}$ may influence NCp7 recognition of the tRNA, the protein's binding to various constructs of the $hASL^{Lys3}_{UUU}$ was investigated. The ASL^{Lys3}_{UUU} and its fragments (residues 32–43) were chemically synthesized with and without the complete and incomplete human modifications $mcm^5s^2U_{34}$, $ms^2t^6A_{37}$, 2-thiouridine (s^2U_{34}) and ψ_{39}, and the *E. coli* modification 5-methylaminomethyluridine (mnm^5U_{34}).[38,44] NCp7 (75 nM) was titrated with the unmodified ASL^{Lys3}_{UUU} (0 – 3 μM) and the change in fluorescence, or quenching, of NCp7's single tryptophan residue (W37) was recorded (Figure 2a).[7] NCp7's interaction with the unmodified $hASL^{Lys3}_{UUU}$ yielded a K_d in the low μM range (K_d = 2.3 ± 0.62 μM; Table 1) and as such established a baseline for recognition of the RNA in the absence of modifications.

NCp7 was also titrated with the doubly modified $hASL^{Lys3}_{UUU}$ having the naturally occurring modifications, $mcm^5s^2U_{34}$ at wobble position 34 and $ms^2t^6A_{37}$ at position 37, 3′-adjacent to the anticodon, $hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$ (Figure 2a). Conditions were identical to those of the titration with the unmodified ASL^{Lys3}_{UUU}, described above. NCp7 exhibited a 10-fold increase in affinity for the modified ASL (K_d = 0.28 ± 0.03 μM; Table 1). The titration of NCp7 with the unmodified $hASL^{Lys3}_{UUU}$ resulted in a maximum fluorescence quenching of only 68% as compared to 75% reduction achieved with the doubly modified $hASL^{Lys3}_{UUU}$ (Figure 2a). When the $htRNA^{Lys3}_{UUU}$ is annealed to the viral genome for the priming of reverse transcription, the U-rich ASL^{Lys3}_{UUU} anneals to the A-rich bulge of the viral stem and loop (Figure 1b) which also must be destabilized by NCp7. In order to compare NCp7's binding of the modified and unmodified ASL^{Lys3}_{UUU} to that of the A-rich bulge of the viral RNA, the NCp7 was titrated with a 27 nucleotide, stem and loop construct of the HIV-1, serotype-G RNA (VSL-1) and the fluorescence of tryptophan 37 was monitored (Figure 2b). NCp7 bound the VSL-1 with an affinity comparable to its binding of the $hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$, and to other stem/loop constructs of viral RNA (reviewed in reference 45), but considerably greater than that of its binding to the unmodified ASL^{Lys3}_{UUU} (Figure 2b; Table 1).

We also assessed the binding of the NCp7 to fragments of the $hASL^{Lys3}_{UUU}$ that were composed of the single stranded, dodecamer oligoribonucleotide residues 29–40, and its single stranded complementary sequence of the viral A-rich Loop 1 (Figure 1b). The fragments were chosen to be non-self complementary, and thus could not form a stem and loop structure. Chemical constraints to the syntheses of these fragments in sufficient quantities limited us to the introduction of the incomplete modifications of s^2U_{34} and N^6-threonylcarbamoyladenosine (t^6A_{37}), along with ψ_{39}. In addition, the bacterial analog of mcm^5U_{34}, mnm^5U_{34}, was incorporated into the fragments. Because of its known preference for single stranded RNA, it is not surprising that NCp7 exhibited a higher affinity for the shorter, unstructured fragments of the anticodon and Loop 1 than for the unmodified $hASL^{Lys3}_{UUU}$ hairpin (Table 1). In fact, NCp7 bound with high affinity to the single stranded unmodified anticodon fragment (K_d = 0.62 ± 0.12 μM) and even more so to the A-rich Loop 1 (K_d = 0.06 ± 0.01 μM) than to a double stranded RNA composed of these two fragments (K_d = 1.11 ± 0.16 μM) or the unmodified $hASL^{Lys3}_{UUU}$ (K_d = 2.30 ± 0.62 μM).

Perhaps more interesting is the comparison of binding affinities between modified and unmodified fragments. Introduction of the modifications further enhanced the affinity of NCp7 for these fragments. In particular, the individual introductions of s^2U_{34}, mnm^5U_{34} and t^6A_{37} significantly increased the protein's affinity for the anticodon fragment (Table 1). NCp7 bound the modified dodecamer fragments of the $hASL^{Lys3}_{UUU}$ with an affinity higher than that reported for DNA dodecamers.[46] These results for single stranded modified

anticodon fragments along with those of the ASL study suggest that NCp7's binding of htRNA$^{Lys3}_{UUU}$ is considerably enhanced when modifications are present. The modifications may not be necessary for NCp7's initial non-specific binding of the tRNA's acceptor stem, but they appear to be of considerable importance for the protein's recognition of the anticodon stem and loop domain.

Peptide recognition of the modified anticodon domain of htRNA$^{Lys3}_{UUU}$

Phage Display Library selection of peptides is an effective way of screening a large number of peptide-RNA interactions under biologically relevant conditions to determine the contribution of modifications to protein recognition of tRNAs, and the amino acids required.[47,48] Peptides that bound the hASL$^{Lys3}_{UUU}$-ms^{2}t^{6}A$_{37}$;mcm^{5}s^{2}U$_{34}$;ψ$_{39}$ were selected using two different phage display libraries and two different conditions for elution from microplates.[49] This triply modified hASL$^{Lys3}_{UUU}$ was chemically synthesized with biotin at the 3′-end, and bound to streptavidin-coated, high capacity microplates. The first round of screening was conducted with the unmodified hASL$^{Lys3}_{UUU}$ bound to streptavidin plates, followed by multiple rounds of screening with the triply modified hASL$^{Lys3}_{UUU}$. Phages were eluted from the plates using both alkaline and acid conditions.[49] Phage that demonstrated an affinity to bind the unmodified hASL$^{Lys3}_{UUU}$ after four rounds of selection were catalogued and eliminated from further screening. In the subsequent step, phage that did not bind hASL$^{Lys3}_{UUU}$ were selected using hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$; ms^{2}t^{6}A$_{37}$;ψ$_{39}$. After four iterations of phage selection against the fully modified hASL$^{Lys3}_{UUU}$, 155 distinct colonies were chosen for DNA sequencing. Several sequences were found to occur repeatedly among the 155 colonies. Twenty-five peptides were chosen for further study, however five sequences either failed chemical synthesis or were insoluble due to a high degree of hydrophobicity. Thus, 20 peptides were synthesized with a conjugated fluorescein isothiocyanate (FITC) at the N-terminus for analysis of their interaction with the ASLs using fluorescence spectroscopy. Peptides 6 (P6), 15 (P15), 17 (P17) and 21 (P21) were chosen for further study as they appeared multiple times from separate colonies during the phage display selection (Table 2) suggesting that these four would bind to the hASL$^{Lys3}_{UUU}$-ms^{2}t^{6}A$_{37}$;mcm^{5}s^{2}U$_{34}$ with high affinity. Unfortunately, P15 lacked the solubility required of the assay and was eliminated from the study.

Though the biotin conjugated hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$;ψ$_{39}$ was used for the selection of phage, its synthesis and purification in quantities required for characterizing peptide affinity and specificity was cost prohibitive. We and others have found that ψ$_{39}$ slightly improves the thermal stability of hASLs, but adds little to their chemistry and conformation, and nothing to their affinity for the cognate codon on the ribosome.[50,51] Therefore, the affinity of the fluorescein-conjugated peptides for the doubly modified hASL^{Lys3}-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$ was determined by monitoring the fluorescence quenching upon binding of the ASLs (Figure 3). Each of the peptides exhibited fluorescence quenching with increasing concentrations of doubly modified hASL$^{Lys3}_{UUU}$. Both P6 and P17 bound the hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$ with a high degree of affinity (K_d = 0.5 μM ± 0.1 and 0.6 ± 0.1 μM, respectively; Table 1), and comparable to that of NCp7.

The specificities of P6, P17 and P21 for the unmodified ASL$^{Lys3}_{UUU}$ were assessed and compared. In contrast to their binding of the modified ASL$^{Lys3}_{UUU}$, P6 and P17 demonstrated little to no interaction with this ASL, and a dissociation constant could not be determined (Figure 3a,b; Table 1). However, P21 bound the hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$ and the unmodified hASL$^{Lys3}_{UUU}$ with considerable, and almost equal, affinity (K_d = 0.98 ± 0.03 μM and K_d = 1.30 ± 0.04 μM, respectively) demonstrating that this peptide's binding of the ASLs is relatively non-specific. To determine if the peptides were non-specifically recognizing the five base-paired stem and/or the seven-residue loop of the ASL hairpin structure, we tested their abilities to bind the *E. coli*

$tRNA^{Val3}_{UAC}$. The *E. coli* $tRNA^{Val3}_{UAC}$ has a different stem, and an anticodon loop nucleoside sequence similar to that of $hASL^{Lys3}_{UUU}$ (Figure 3c). The $tRNA^{Val3}_{UAC}$ construct had its A_{37} modified at N^6, but with only a methyl group, i.e. m^6A_{37} (ASL^{Val3}_{UAC}-m^6A_{37}). P6 did not bind the *E. coli* ASL^{Val3}_{UAC}-m^6A_{37} indicating that peptide P6 is specifically recognizing the modifications of the $hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$. However, the fluorescence of peptide P17 was quenched when titrated with the ASL^{Val3}_{UAC}-m^6A_{37} indicating recognition of an N^6-modified A_{37} (Figure 3a,b).

The specificity and affinity of P6 and P17 for $hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$ may be due to only one of the two modifications rather than both. Therefore, we determined their affinities for the singly modified anticodon stem and loop of the human $tRNA^{Lys}$ species 1 ($hASL^{Lys1}_{CUU}$-t^6A_{37}) (Figure 1g). Though $htRNA^{Lys1}_{CUU}$ has an unmodified C_{34}, A_{37} is modified to t^6A_{37} and naturally lacks the ms^2-moiety. Otherwise the stem and loop sequences of the ASLs are the same. Both peptides exhibited significant affinities for $hASL^{Lys1}_{CUU}$-t^6A_{37} (K_d= 0.60 ± 0.09 μM and 1.8 ± 0.20 μM, respectively; Table 1). Not only was the affinity of the peptides for the $hASL^{Lys1}_{CUU}$-t^6A_{37} considerable, but the fluorescence was quenched to over 50% when the peptide was saturated with $hASL^{Lys1}_{CUU}$-t^6A_{37}, as compared to just 12% with the unmodified $hASL^{Lys3}_{UUU}$. Apparently, the single modification t^6A_{37} is an important recognition element.

NCp7 and a peptide mimic alter the $hASL^{Lys3}$ conformation in a modification dependent manner

NCp7 has the ability to disrupt RNA secondary structure to enable the annealing of the primer $htRNA^{Lys3}_{UUU}$ to the HIV genomic RNA which is energetically favored. To test NCp7's ability to disrupt the base stacking and hydrogen bonds of the $hASL^{Lys3}_{UUU}$, we designed an unmodified $hASL^{Lys3}_{UUU}$ to exhibit a fluorescence (Förster) resonance energy transfer (FRET) between the donor, 5′-conjugated fluorescein, and the acceptor, a 3′-conjugated accepting quencher, dabsyl (Figure 1f). Because fluorescence emission from the 5′-terminal fluorescein is transferred to the neighboring 3′-terminal dabsyl, increases in fluorescence would be expected should the NCp7 disrupt the structure of the $hASL^{Lys3}_{UUU}$ and thereby increase the distance between the two termini. In fact, the protein's ability to denature the unmodified $hASL^{Lys3}_{UUU}$ was similar to the thermal denaturation of the same ASL (Figure 4). At a molar ratio of ~2:1 (NCp7 to $hASL^{Lys3}_{UUU}$) and a temperature of ~40 °C, the ASL appears to be half melted. Having established that NCp7 will disrupt the structure of the unmodified $hASL^{Lys3}_{UUU}$, we were interested in knowing if the natural modifications had an effect on this function and whether the peptides were capable of mimicking NCp7's denaturation of the RNA. However, for these experiments we were compelled to use a different approach because the synthesis of the fully modified $hASL^{Lys3}_{UUU}$ with fluorescein and dabsyl was chemically difficult and inordinately expensive.

Changes in nucleic acid base stacking, and indirectly hydrogen bonding, caused by an alteration in the nucleotide sequence, changes in temperature, or by binding of ligands can be monitored by circular dichroism (CD) spectrapolarimetry.[52,53] The titration of an RNA with either a protein or peptide adds ellipticity to the CD spectrum of the nucleic acid, especially in the wavelength range where backbone interactions are observed, 220–250 nm. Fortunately, both proteins and peptides have a spectral null in the wavelength range of interest for observing base stacking interactions, 250–280 nm,[52] enabling the recording of changes in the RNA conformation even in the presence of proteins. NCp7 alone exhibited the expected spectral null in the wavelength range of 260–280 nm (Figure 5a). The unmodified $hASL^{Lys3}_{UUU}$ (1.5 μM) was titrated with NCp7 and its CD ellipticity monitored for a change in RNA conformation (Figure 5a). The CD spectra of the unmodified

hASL$^{Lys3}_{UUU}$ when titrated with increasing concentrations of NCp7 exhibited a decrease in ellipticity, indicating that the protein binds and relaxes the ASL structure (Figure 5a), which is consistent with decreased stacking interactions, as would be expected from the results of the prior FRET experiment.

The triply modified hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$;ψ$_{39}$ was also titrated with NCp7 and the protein again appeared to disrupt the base stacking interactions and to relax the ASL$^{Lys3}_{UUU}$ structure (Figure 5b). We compared the changes in ellipticity caused by NCp7 for unmodified ASL$^{Lys3}_{UUU}$ and modified hASL$^{Lys3}_{UUU}$ at the wavelength maximum and found that hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$;ψ$_{39}$ required less protein to cause denaturation of the ASL structure (Figure 5c). The amount of NCp7 required to achieve the point at which half of the ASL was melted was 0.3 μM for the modified hASL$^{Lys3}_{UUU}$, and 0.46 μM for the unmodified hASL$^{Lys3}_{UUU}$ (Figure 5c). Unlike the binding experiments monitored by fluorescence, the CD experiments are not indicative of the degree to which binding occurs. At concentrations of NCp7 above the ratio of 1, NCp7 to ASL, there was little change in CD ellipticity with addition of more protein. Since we did not perform a kinetic analysis, we cannot say that at any snapshot in time, there would be more NCp7 on the modified ASL versus the unmodified ASL under conditions in which each is saturated with protein. However, these data confirmed that the hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$;ψ$_{39}$ is bound and its secondary structure more effectively disrupted by NCp7 than is that of the unmodified hASL$^{Lys3}_{UUU}$.

The hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$;ψ$_{39}$ and unmodified hASL$^{Lys3}_{UUU}$ constructs were also titrated with the P6 peptide to determine if it was able to mimic the ability of NCp7 to bind and unfold the ASL structure. The spectra indicated that the peptide mimics NCp7 in its ability not only to bind hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$;ψ$_{39}$ better than the unmodified hASL$^{Lys3}_{UUU}$, but also in its ability to relax the ASL^{Lys3} structures (Figure 5d). Substantive differences were evident in the P6-facilitated denaturation of the triply-modified hASL$^{Lys3}_{UUU}$ in comparison to that of the unmodified hASL$^{Lys3}_{UUU}$. When changes in ellipticity were compared for increasing concentrations of P6 at the wavelength maximum, it was apparent that the peptide completely disrupted the structure of the hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$;ψ$_{39}$ at a concentration of 2.15 μM (Figure 5d). In contrast, the P6 denaturation of the unmodified hASL$^{Lys3}_{UUU}$ was incomplete even at a concentration of 5.15 μM (Figure 5d). The concentration at which P6 had disrupted half the hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$;ψ$_{39}$ was ~0.90 μM. Thus, P6 was able to mimic NCp7's ability to melt the RNA, and the presence of modifications were important to how little of the peptide was needed to do so. It is also noteworthy that the modified hASL$^{Lys3}_{UUU}$ was half-melted with an NCp7 concentration of 0.3 μM or by P6 at 0.9 μM, very much like their respective K_ds of 0.28 and 0.5 μM.

Discussion

The results presented here suggest that NCp7 recognition and remodeling of htRNA$^{Lys3}_{UUU}$ *in vivo* may be more specific than originally revealed, since it is enhanced by the presence of post-transcriptional modifications. The recruitment of htRNA$^{Lys3}_{UUU}$ for the priming of HIV-1 reverse transcription involves the Gag and the Gag-Pol polyprotein, and the human KRS, or probably a composite of these proteins.[4] The NCp7 domain of Gag is an RNA chaperone of htRNA$^{Lys3}_{UUU}$ and facilitates its annealing to the viral RNA.[4] The anticodon domain of tRNA^{Lys} species is an identity determinant for the synthetase.[54,55] The synthetase is involved in the unacylated tRNA's recruitment to the viral RNA. It is proposed that the synthetase may be recognizing an htRNA$^{Lys3}_{UUU}$-like element (TLE) of the viral RNA that has sequence homology with the U-rich anticodon region and thereby directs the tRNA to the complementary viral sequences for annealing.[4] The complexity of protein components

and the detailed mechanisms by which they recognize, recruit and anneal $htRNA^{Lys3}_{UUU}$ to the viral genome have yet to be described.

NCp7 facilitates the use of the $htRNA^{Lys3}_{UUU}$ species as the primer for reverse transcription. The protein remodels the tRNA conformation, thus facilitating its annealing to the viral genome for the initiation of reverse transcription during the next round of infection.[4,56] The HIV-1 capsids are packaged with viral RNA and concentrate $htRNA^{Lys3}_{UUU}$ some 60-fold to a ratio of tRNA to viral genome of approximately 6:1 to 8:1.[4,57] The two zinc fingers of NCp7 provide the protein with a highly structure-specific binding of RNA, whereas a large number of basic residues produce a high degree of promiscuity in its binding to a variety of nucleic acids with varying affinities *in vitro*.[4] NCp7 exhibits some sequence discrimination,[6,24,25] however, prior studies did not include the potential influence of the naturally occurring, post-transcriptional modifications that are found in the host cell $tRNA^{Lys3}_{UUU}$. As would be expected for a protein involved in the recognition and remodeling of the $htRNA^{Lys3}_{UUU}$ conformation, NMR spectroscopy revealed that NCp7 bound to the base paired regions inside of the canonical L-shape of the acceptor stem of tRNA wherein lies tertiary structure hydrogen bonding and the G6•U67 and T54•A58 pairs.[24] With ^{15}N-labeling, the imino protons of the base paired stems of a $htRNA^{Lys3}_{UUU}$ product of *E. coli* cloning and thus with bacterial modifications were observable, but those of the anticodon loop were in too fast an exchange to be observed. In experiments with unmodified D- and T-domain fragments of the tRNA, NCp7 exhibited an affinity for the D-domain fragment comparable to the association that we observed with the anticodon domain.[58]

Modifications had a considerable impact upon NCp7 recognition of ASL^{Lys3}_{UUU}. In fact, the affinity of NCp7 for the $hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$ (K_d = 0.28 ± 0.03 μM) was equivalent to the affinity reported for its binding to the native bovine $tRNA^{Lys3}_{UUU}$ (K_d = 0.31 μM) using, as we did, the inherent fluorescence of the protein's W37 and under conditions of comparable ionic strength.[28] NCp7 appears to bind nucleic acids in a salt-effected manner as would be expected of an electrostatic interaction with the phosphodiester backbone contributing significantly to the binding constant.[59] The binding of the stem loop structures of HIV's ψ-domain contribute significantly to our knowledge of its affinities for these structures. NCp7 has a high affinity for model hairpins (20mers) of the ψ-domain stem and loops SL2 and SL3 (K_d ≈ 0.1–0.2 μM) and a lower affinity for SL4 (K_d ≈ 14 μM).[60–62] The affinities of NCp7 that we have observed for the $hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$ was comparable (K_d= 0.28 ± 0.03 μM) and acquired under similar salt conditions. The equilibrium binding constant of the bacterial modified $Fragment_{29–40}$-mnm^5U_{34};t^6A_{37} (0.26 ± 0.06 μM) was comparable to that of the $hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$ (0.28 ± 0.03 μM), and little different than that of the fragment with t^6A_{37} only (0.15 ± 0.06 μM) especially when one takes into consideration the errors of the binding study.

The human $tRNA^{Lys3}_{UUU}$ has 18 nucleotides at the 3′-terminus, most of which are engaged in A-form RNA duplexes of the amino acid acceptor stem and the T-stem. These nucleotides are complementary to the viral PBS and must first be freed from their internal hydrogen bonding to participate in binding to the PBS. Though this sequence is an important distinctive property of the tRNA and critical to its priming function, the combination of modified nucleosides found in the anticodon loop of $htRNA^{Lys3}_{UUU}$ are a chemically rich and exclusive attribute to that tRNA. The modifications s^2U_{34} and t^6A_{37} are found in other tRNAs.[63,64] Therefore, it is interesting to note that $htRNA^{Lys3}$, $htRNA^{Lys1,2}$, $htRNA^{Ile}$, and $htRNA^{Asn}$ are among those tRNAs most frequently found within the HIV viral capsid.[18,19,65] The $htRNA^{Lys1,2}$, $htRNA^{Ile}$, and $htRNA^{Asn}$ have the nucleoside modification t^6A_{37}. We have found that NCp7 binds with high affinity to the $hASL^{Lys1}_{CUU}$ with t^6A_{37}, as well as to the $hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$, but not to the unmodified

hASL$^{Lys3}_{UUU}$. Thus, the modifications are themselves recognition determinants for NCp7. Alternatively, the anticodon domain conformation resulting from the modifications enhances recognition. We hypothesize that this recognition mechanism provides another level of specificity to protein-nucleic acid interactions *in vivo*.

NCp7 disruption of the hydrogen bonding of the tRNA's amino acid accepting stem and the stem of the T-domain is well documented and is an accepted functional mechanism of facilitating the binding of the tRNA to the viral PBS.[4] Presumably, NCp7 is also involved in disrupting the structure of the U-rich anticodon stem and loop domain to facilitate the subsequent binding of this sequence to the A-rich bulge of the Loop 1 region of the viral RNA.[29,42] Using a FRET analysis between the 5′- and 3′-termini of the hASL$^{Lys3}_{UUU}$, we found that when titrated with NCp7, the hASL$^{Lys3}_{UUU}$ structure is disrupted. This denaturation was analogous to the thermal denaturation of the ASL.

The interaction of NCp7 with the full length transcript of tRNA^{Lys3} had been monitored previously by circular dichroism and a change in the protein's spectrum was reported.[62] However, little to no change in the tRNA's spectrum was observed with the addition of three mole-equivalents of NCp7 to one of RNA, ~1/25 of NCp7 to nucleotides.[66] This result contrasts to that of other reports[29,42] and also to our own CD results with the ASLs. The protein's denaturation of the hASL$^{Lys3}_{UUU}$ was observed by monitoring CD spectra while increasing the concentration of NCp7. The presence of modifications in the hASL$^{Lys3}_{UUU}$ facilitated NCp7's denaturation of the hASL$^{Lys3}_{UUU}$. The heptadecamer ASL was half melted at a stoichiometry of approximating 2:1, NCp7 to ASL, consistent with the reported ratio of one NCp7 per 6–8 nucleotides. At first this stoichiometry appears to contrast with the one-site, non-linear curve fitting of the binding curves to the ASL. The protein does have a considerably greater affinity for the single stranded anticodon domain fragments than for the ASL hairpin. Thus, the second site binding of NCp7 to the ASLs could be weaker than that of the first and not observed under the conditions of our fluorescence studies. The 12mer fragments would have one or possibly two NCp7 bound per RNA, the heptadecamer RNA stem and loops would have two molecules of NCp7 bound and the 27mer VSL construct (Figure 2b) being considerably larger could bind more NC proteins.

A stoichiometric binding of NCp7, one per 6–8 nucleotides, would decrease the possibility of annealing the ASL to the viral RNA until the local amount of the protein is reduced, perhaps through the kinetics of on/off rates. In other systems, (e.g., *trans*-activating region RNA interaction, TAR-TAR, and the dimer initiation signal interaction, DIS-DIS), it appears that only a small amount of nucleocapsid protein is sufficient to produce annealing and structural rearrangement (e.g., from kissing dimer to extended duplex).[45,67,68] In fact, our CD experiments that monitored disruption of base stacking and not melting indicated that NCp7 at a ratio a of 0.2:1 (NCp7/ASL) and peptide at 0.5:1 were sufficient to dissolve half the base stacking interactions of the modified hASL$^{Lys3}_{UUU}$ (Figure 5c,d). Though the ratio seems low, CD results average all of the interactions taking place in the experiment even when the amount of protein is below that of the ASL. Thus, we see base stacking being significantly disrupted with as little as 0.2:1 ratio of NCp7/ASL, ie. 20% of the ASL is bound by the NCp7. We are not alone in this observation of NCp7 interactions with htRNA$^{Lys3}_{UUU}$ (personal communication, K. Musier-Forsyth).[45,67,68] We found little difference in the concentrations of NCp7 required for the half-maximal denaturation of the triply modified and unmodified hASL$^{Lys3}_{UUU}$, "destablization K_d" determined by CD (Figure 5c). This is in contrast to the significant difference in the binding K_d (Table 1). NCp7 affinity for the modified hASL$^{Lys3}_{UUU}$ was comparable to that for NCp7 binding of viral RNA and significantly higher than that for the unmodified hASL$^{Lys3}_{UUU}$. Although there are some similarities between the mechanisms of action of DNA single stranded binding protein (SSB) and NCp7, NCp7's faster kinetics and ability to aggregate, as well as

destabilize, RNA and DNA appear to be most important for its functions.[69] Thus, destabilization of RNA occurs at sub-saturating NCp7 concentrations, whereas annealing/aggregation requires saturating levels of NC.[70]

The effects of modifications on the chemistry, structure and conformational dynamics of the anticodon loop of $htRNA^{Lys3}_{UUU}$ increased the binding affinity of NCp7 and of peptide P6, and facilitated disruption of the ASL structure. The major recurring RNA chemistries available as identity determinants for protein recognition are the nucleobase functional groups (NH_2 and CO), the ribose 2′-OH and the phosphate backbone. But, these *per se* do not create a unique identity. Thus, with some notable exceptions among the most studied of RNA-binding proteins,[71] there is little understanding of how proteins specifically interact with their target RNA to form native complexes. Peptide-RNA interactions are more easily studied in detail, such as the thoroughly investigated Tat-TAR[72] and Rev-RRE systems.[73] Previously, we had used Peptide Phage Display libraries to effectively screen and select peptides that bind specifically to modified nucleosides with high affinity and specificity.[47–49] In this study, peptides were selected for their abilities to bind specifically to the modified ASL region of the human $tRNA^{Lys3}$ ($hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$). The peptides selected from this screen were analyzed in detail for their abilities to bind the $hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$. Two peptides were found to have affinities and specificities for ASL^{Lys3}-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$ that were comparable to those of NCp7. In addition, peptide P6 could also mimic NCp7's ability to denature the RNA's hairpin conformation, albeit with somewhat reduced efficiency. Potentially, P6 could facilitate the annealing of the $hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$ to the A-rich bulge of the HIV's stem and Loop 1. In binding to $hASL^{Lys1}_{CUU}$-t^6A_{37}, P6 and P17 appeared to favor the t^6A_{37} over the $mcm^5s^2U_{34}$ as the most important of the modifications in their recognition of the modified $hASL^{Lys3}_{UUU}$. Perhaps in using $mcm^5s^2U_{34}$ and $ms^2t^6A_{37}$ and the t^6-moiety in particular as identity elements, the two peptides bind to the anticodon, the 3′-side of the loop and the 3′-side of the stem.

The short sequence (15 amino acids) of peptide P6 and its lack of structure as determined by CD effectively mimics the modification-facilitated binding and functionality of the larger and structured NCp7 (55 amino acids and two zinc fingers). The binding of P6 to the anticodon region of the modified $hASL^{Lys3}_{UUU}$ was confirmed with preliminary data obtained from mass spectrometry (MS). MS experiments were conducted with the specific aim of observing complex formation between the peptides and modified nucleosides. Electrospray ionization-MS (ESI-MS), an analytical technique that is used to observe non-covalent complex formation between nucleic acids and proteins in solution at nM concentrations,[74] has been used to investigate the specific interactions between NCp7 and the different domains of the HIV-1 packaging signal.[45,75,76] For this reason, we employed ESI-MS as the most cost effective screening technique to observe the interaction between peptide P6 and $hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$; $ms^2t^6A_{37}$. As a control, the modified $hASL^{Lys3}$ was analyzed by ESI-MS in the absence of the peptide. Spectra indicate a −4 charge state for the RNA (Figure 6). Peptide P6 was analyzed alone and under the same conditions. Both the modified and unmodified $hASL^{Lys3}_{UUU}$ were titrated with P6. A comparison of the resulting spectra led to the observation of a decrease in the abundance of free peptide in solution, along with an increase in the abundance of its complex with $hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$. With equimolar amounts of RNA and peptide, ~50% of the $hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$ was in a complex with P6, comparable to the results obtained when the fluorescein-conjugated peptide's binding of the modified $hASL^{Lys3}_{UUU}$ was monitored by fluorescence. At a ratio of 1:5, RNA to peptide, ~100% of the modified $hASL^{Lys3}_{UUU}$ had been bound by P6 and the free ASL was no longer observed (Figure 6b). In these preliminary experiments, small amounts of the unmodified $hASL^{Lys3}_{UUU}$ were bound by P6, and only 50% of *E. coli* ASL^{Val3}_{UAC} was in a complex with P6 when the

peptide was in five-fold excess (data not shown). Thus, peptide P6 bound the modified hASL$^{Lys3}_{UUU}$ preferentially.

Anticodon loop modifications create an architecture optimized for the tRNA's binding to the mRNA codon by reducing thermal stability and conformational dynamics.[40,41,77] The modifications at wobble position 34, mcm^{5}s^{2}U$_{34}$ and 3′-adjacent to the anticodon at position 37, ms^{2}t^{6}A$_{37}$, reduce thermal stability by negating intra-loop hydrogen bonds (Figure 7). They alter the nucleobase's hydrogen acceptors and donors known to engage in non-canonical, as well as canonical base pairing. Their prominent physical volume widens the loop. However, modifications at purine-37 also increase the base stacking of the nucleosides on the 3′-side of the loop and thus increase the order within the loop.[41,77] Therefore, the modifications would contribute a physicochemical advantage, as well as their being chemically distinct identity determinants, for the NCp7 denaturation of the htRNA$^{Lys3}_{UUU}$. In contrast to an unmodified hASL$^{Lys3}_{UUU}$, two NCp7 molecules could ostensibly bind sequentially, perhaps cooperatively, to the hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$ and thus prepare the ASL for its annealing to the viral RNA (Figure 7). The introduction of the natural modifications provides the perfect model for recognition of htRNA$^{Lys3}_{UUU}$ by viral proteins such as NCp7 when studied *in vitro*, and may replicate the higher degree of specificity appreciated *in vivo.* The abilities of peptides such as P6 to discriminate between modified and unmodified htRNA^{Lys3}, and for P6 also to be a functional mimic of NCp7 in its ability to disrupt ASL structure, make these peptides prime candidates for the development of HIV-1 antiviral therapeutics. The selection of peptides that can effectively function as mimics of RNA binding proteins, could potentially lead to the development of tools in drug discovery and therapeutics that would silence viral gene expression *in vivo*.

Materials and Methods

RNA sample preparation

The unmodified, heptadecamer oligoribonucleotide corresponding to the anticodon stem and loop of the human tRNA$^{Lys3}_{UUU}$ (hASL$^{Lys3}_{UUU}$) and that of the *Escherichia coli* tRNA$^{Val}_{UAC}$ isoacceptor 3 (ASL$^{Val3}_{UAC}$) were chemically synthesized by Dharmacon (Thermo Fisher, Lafayette, CO) using "ACE" chemistry.[78] The hASL$^{Lys3}_{UUU}$-ms^{2}t^{6}A$_{37}$; mcm^{5}s^{2}U$_{34}$;ψ$_{39}$ with biotin conjugated at the 3′-end, the hASL$^{Lys3}_{UUU}$ - mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$, and the ASL$^{Val3}_{UAC}$ with the natural modification *N*^{6}-methyladenosine, ASL$^{Val3}_{UAC}$-m^{6}A$_{37,}$ were synthesized by Integrated DNA Technologies BVBA using phosphoramidite chemistries.[79] The t^{6}A$_{37}$-modified ASL of the human tRNA$^{Lys}_{CUU}$ isoacceptor 1 (htRNA$^{Lys1}_{CUU}$-t^{6}A$_{37}$) was synthesized by the North Carolina State University Nucleic Acid Facility with the phosphoramidite of t^{6}A$_{37}$ having the ribose 2′ and 5′ hydroxyls and the threonine carboxyl and hydroxyl suitably protected, and with standard ribonucleoside phosphoramidites and little change in standard synthesis protocols. The unmodified ASL$^{Lys3}_{UUU}$ with a 5′-conjugated fluorescein and a 3′-dabsyl, a 27 nucleotide construct of the HIV A-rich bulge stem and loop, and 12 nucleotide anticodon domain fragments were synthesized by the NCSU Nucleic Acid Facility using standard 2′-*O*-TBDMS protected ribonucleoside phosphoramidites (ChemGenes, Wilmington, MA).[80] The newly synthesized ASLs were purified by anion exchange HPLC (Nucleogen 60–7 DEAE; 250 mm × 10 mm column), and desalted (Waters Corporation Sep-pak columns, Milford, MA).[81] The purified ASLs were analyzed for their nucleoside composition by HPLC,[82] and mass spectrometry, and the modifications observed by NMR,[40] and X-ray crystallography bound to codon on the *E. coli* 30S ribosomal subunit (Vendeix, A.F.P., Murphy, F.V. 4th and Agris, P.F., unpublished data).

Isolation and purification of NCp7

The 55 amino acid NCp7 expression clone, pRD2, and NCp7 protein were the gifts of Dr. Michael Summers (University of Maryland, Baltimore County). Additionally, NCp7 protein was expressed using BL21 (DE3) *E. coli* cells and purified as described previously.[61,83]

Phage display selection of peptides and peptide synthesis

Peptides that bound to the fully modified anticodon stem and loop domain of $htRNA^{Lys3}$ ($hASL^{Lys3}_{UUU}$-$ms^2t^6A_{37}$;$mcm^5s^2U_{34}$;ψ_{39}) were selected using two different phage display libraries, fuse5 and f88-cys6, gifts from Dr. George Smith (University of Missouri-Columbia). The fuse5 library is that of a completely random 15 amino acid sequence containing more than 10^7 different phage.[84] The f88-cys6 is a 16 amino acid sequence randomized for every residue except for the locations of two cysteines and containing more than 2.7×10^8 phage.[85] The 3′-biotin conjugated $hASL^{Lys3}_{UUU}$ -$ms^2t^6A_{37}$;$mcm^5s^2U_{34}$;ψ_{39} was bound to 96-well, streptavidin-coated, high capacity microplates (Pierce) in TTDBA buffer (1 mg/ml BSA and 0.02% NaN_3 in 200/1 TBS/Tween, vol/vol; TBS buffer, 50 mM Tris HCl, pH 7.5, 150 mM NaCl). The plates were incubated with phage for 4 hr at 4°C, and were then washed 5 times with a TBS-tween solution in order to remove any unbound phage. To elute the bound phage from the plates with either acidic or alkaline conditions, the plate wells were rinsed with acid elution buffer (Acidic conditions) or alkaline elution buffer (Basic conditions) on a shaker with gentle agitation for 10 minutes.[49] The elution mixture containing eluted phage was transferred to a microtube containing the respective pH-neutralizing buffer.[49] The selected and isolated phage were amplified and the selection was reiterated for a total of five rounds. A screen was conducted with the unmodified $hASL^{Lys3}_{UUU}$ bound to strepavidin plates. Phage that demonstrated an affinity to bind $hASL^{Lys3}_{UUU}$ were catalogued and eliminated. The final round of selection was concluded with a serial dilution of the isolated phage and quantified to determine yield. Phage clones were then constructed on starved *E. coli* K91BluKan cells. Of the phage that demonstrated an affinity for the fully modified $hASL^{Lys3}_{UUU}$ and not the unmodified $hASL^{Lys3}_{UUU}$, 155 distinct colonies were chosen, DNA was isolated and sequenced.[49] The resulting sequences were used to select the peptides to be synthesized. Some sequences were found to occur repeatedly among the 155 colonies (Table 1). Twenty-five sequences were chosen, but five sequences either failed chemical synthesis or were insoluble due to a high degree of hydrophobicity. Thus, 20 peptides were synthesized with a conjugated fluorescein isothiocyanate (FITC) at the N-terminus (Table 1).

The 20 peptides were chemically synthesized (Sigma-Aldrich) with fluorescein isothiocyanate (FITC) conjugated to the N-terminus for assaying peptide interaction with the ASLs by fluorescence spectroscopy. The final concentration was determined by the Bradford assay using a standard curve generated from bovine serum albumin, and a peptide of known concentration. Peptides of low solubility were eliminated from further consideration.

Fluorescence spectroscopy

The importance of anticodon domain modified nucleosides to NCp7's recognition of $htRNA^{Lys3}_{UUU}$ was characterized by monitoring the changes in the intrinsic fluorescence properties of the protein's one tryptophan residue (position 37, Figure 1) upon the addition of RNA. For binding of the ASL's oligonucleotide fragments, tryptophan fluorescence was monitored with a QuantaMaster™ Model C 61 spectrofluorometer (Photon Technology International Inc.). The titrations were performed at 25 °C with increasing amounts of RNA added to a fixed concentration of 75 nM NCp7 in 400 μL buffer (50 mM HEPES, 50 mM NaCl, 0.04% PEG 8000 buffer, pH 7.5; or with equivalent results 20 mM phosphate buffer, 20 mM Na_2HPO_4 and 20 mM KH_2PO_4 of equal volume, pH 6.8). NCp7 fluorescence was

measured at 340 nm emission (290 nm excitation) in a 1 cm path length cuvette. In monitoring the fluorescence of fluorescein 3′-conjugated hASL$^{Lys3}_{UUU}$ with a 5′-dabsyl, fluorescence was measured at 518 nm emission (490 nm excitation) in a 1 cm path length cuvette. Experiments were repeated three times with fluorescence detection collected over 10 seconds after an initial 0.5 min period of equilibration. In assessing the binding of NCp7 to the heptadecamer ASLs, a microplate spectrofluorometer was used (Molecular Devices Spectramax Gemini XS with dual scanning monochrometers, Sunnyvale, CA). Under the conditions of 20 mM phosphate buffer, the tryptophan fluorescence was observed at 360 nm (285 nm excitation). Experiments were performed in triplicate and repeated at least twice. The resulting spectra were normalized and the percentage change in fluorescence was plotted against respective RNA concentrations. An equilibrium binding constant (K_d) was derived from the analysis of the curve using a one site binding, non-linear regression model (Prism, GraphPad Software, San Diego, CA). To assess and correct for inner-filter effects and background, fluorescence intensities were adjusted for dilution, buffer fluorescence, and screening effects due to the presence of RNA. The RNAs inner-filter effect was minimal, and the buffer chosen had insignificant effects on NCp7 tryptophan fluorescence. Photobleaching was determined to be only 3% of total fluorescence during a period of 10 min. Peptide interaction with the ASLs was monitored through changes in the fluorescence of the N-terminal FITC that was observed at 524 nm (486 nm excitation). The peptides (0.5 μM) were titrated with varying concentrations of ASLs in 20 mM phosphate buffer. The K_ds for each fluorescence experiment were assessed using the one site binding equation $Y = [Bmax][X]/[K_d+X]$. As a control a non-RNA binding peptide (P31; AGPVPLHSLSYYYNQ) was used as a baseline (data not shown). The data from the P31 was not represented on the graph because of the high rate of error that occurred with this non-specific interaction.

UV-monitored, thermal denaturations

Human ASL$^{Lys3}_{UUU}$ was dissolved to a concentration of 2 μM (20 mM phosphate buffer). Thermal denaturations and renaturations, performed in triplicate in cells of 1 cm path length, were monitored by UV absorbance (260 nm) using a Cary 3 spectrophotometer as previously described.[50] Data points were averaged over 20 seconds and collected three times/minute with a temperature ramp of 1 °C/min from 5 – 90 °C. Data from denaturations and renaturations were treated similarly. No hysteresis was observed.

Circular dichroism (CD)

CD experiments, performed in triplicate, were conducted with a Jasco J815 spectropolarimeter and an interfaced computer.[86,87] RNA samples (1.5 μM) were prepared in 10 mM phosphate buffer, pH 6.8, and spectra collected from 225 nm to 325 nm. To maintain RNA and buffer concentrations and to minimize volume changes, the RNA samples were added to lyophillized protein or peptide samples of known quantity resulting in a range of protein concentrations from 0.05 μM to 3.65 μM and peptide concentration from 0.05 μM to 5.15 μM. The CD spectra were collected at 4 °C, at a rate of 10 nm per minute, a resolution of 1 nm, and the spectra were averaged over 6 runs. Normalization of spectra was performed by calculating molar circular dichroism $\Delta\varepsilon$ (cm^2/mmol) using the formula $\Delta\varepsilon = (\theta/32980)(C)(L)(N)$, where θ = raw CD amplitude (mdeg), C = concentration (mol/L), L = path length (cm) and N = number of nucleotides in RNA. Data were plotted and a one-site (NCp7) or two-site (P6), non-linear regression analysis was performed using Prism 5 (GraphPad Software, San Diego, CA).

Acknowledgments

The authors wish to acknowledge the help of Winnell Newman in the synthesis of some of the RNA at the North Carolina State University Nucleic Acid Facility, Michael Summers for NCp7 and a clone of the protein, George Smith for the specific phage display libraries, Karin Musier-Forsyth for her suggestions for the manuscript, Glenn Björk for the modified nucleoside composition analysis by HPLC, and Daniele Fabris for help with mass spectrometry and for manuscript suggestions. This work was supported by grants from the National Institutes of Health, the National Science Foundation, and the North Carolina Biotechnology Center to P.F.A. (2RO1-GM23037, MCB0548602 and MRG 1102/Agris, respectively).

Abbreviations

NCp7	HIV nucleocapsid protein
hASL$^{Lys}_{UUU}$	anticodon stem and loop of human tRNA$^{Lys3}_{UUU}$
mcm^{5}s^{2}U$_{34}$	5-methoxycarbonylmethyl-2-thiouridine at anticodon wobble position 34
ms^{2}t^{6}A$_{37}$	2-methylthio-*N*^{6}-threonylcarbamoyladenosine at position 37
Ψ$_{39}$	pseudouridine at anticodon stem position 39

References

1. Freed EO. HIV-1 Gag proteins: Diverse functions in the virus life cycle. Virology. 1998; 251:1–15. [PubMed: 9813197]
2. Coffin, J.; Hughes, S.; Varmus, H. Retroviruses. Cold Spring Harbor Press; Plainview, NY: 1997.
3. Coren LV, Thomas JA, Chertova E, Sowder RC, Gagliardi TD, Gorelick RJ, Ott DE. Mutational analysis of the C-terminal gag cleavage sites in human immunodeficiency virus type 1. J Virol. 2007; 81:10047–10054. [PubMed: 17634233]
4. Levin JG, Mitra M, Mascarenhas A, Musier-Forsyth K. Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription. RNA Biology. 2010; 7:754–774. [PubMed: 21160280]
5. Thomas JA, Gorelick RJ. Nucleocapsid protein function in early infection processes. Virus Res. 2008; 134:39–63. [PubMed: 18279991]
6. Levin JG, 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. [PubMed: 16164976]
7. You JC, McHenry CS. HIV nucleocapsid protein. Expression in *Escherichia coli*, purification, and characterization. J Biol Chem. 1993; 268:16519–16527. [PubMed: 8344933]
8. Lapadat-Tapolsky M, de Rocquigny H, Van Gent D, Roques B, Plasterk R, Darlix JL. Interactions between HIV-1 nucleocapsid protein and viral DNA may have important functions in the viral life cycle. Nucleic Acids Res. 1993; 21:831–839. [PubMed: 8383840]
9. de Baar MP, van der Horn KHM, Goudsmit J, de Ronde A, de Wolf F. Detection of human Immunodeficiency virus type 1 nucleocapsid protein p7 *in vitro* and *in vivo*. J Clin Microbiol. 1999; 37:63–67. [PubMed: 9854065]
10. Wain-Hobson S, Sonigo P, Danos O, Cole S, Alizon M. Nucleotide sequence of the AIDS virus, LAV. Cell. 1985; 40:9–17. [PubMed: 2981635]
11. Ratner L, Haseltine W, Patarca R, Livak KJ, Starcich B, Josephs SF, Doran ER, Rafalski JA, Whitehorn EA, Baumeister K, Ivanoff L, Petteway SR Jr, Pearson ML, Lautenberger JA, Papas TS, Ghrayeh J, Chang NT, Gallo RC, Wong-Staal F. Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature. 1985; 313:277–284. [PubMed: 2578615]
12. Marquet R, Isel C, Ehresmann C, Ehresmann B. tRNAs as primer of reverse transcriptases. Biochimie. 1995; 77:113–124. [PubMed: 7541250]
13. Mak J, Kleiman L. Primer tRNAs for reverse transcription. J Virol. 1997; 71:8087–8095. [PubMed: 9343157]
14. Isel C, Ehresmann C, Marquet R. Initiation of HIV reverse transcription. Viruses. 2010; 2:213–243. [PubMed: 21994608]

15. Isel C, Marquet R, Keith G, Ehresmann C, Ehresman B. Modified nucleosides of $tRNA^{Lys3}$ modulate primer/template loop-loop interaction in the initiation complex of HIV-1 reverse transcription. J Biol Chem. 1993; 268:25269–25272. [PubMed: 7503978]
16. Bilbille Y, Vendeix FAP, Guenther R, Malkiewicz A, Ariza X, Vilarrasa J, Agris PF. The structure of the human $tRNA^{Lys3}$ anticodon bound to the HIV genome is stabilized by modified nucleosides and adjacent mismatch base pairs. Nucleic Acids Res. 2009; 37:3342–3353. [PubMed: 19324888]
17. Wilkinson KA, Gorelick RJ, Vasa SM, Guex N, Rein A, Mathews DH, Giddings MC, Weeks KM. High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. PLoS Biol. 2008; 6:e96, 0883–0899. [PubMed: 18447581]
18. Jiang M, Mak J, Ladha A, Cohen E, Klein M, Rovinski B, Kleiman L. Identification of tRNAs incorporated into wild-type and mutant human immunodeficiency virus type 1. J Virol. 1993; 67:3246–3253. [PubMed: 8497049]
19. Pavon-Eternod M, Wei M, Pan T, Kleiman L. Profiling non-lysyl tRNAs in HIV-1. RNA. 2010; 16:267–273. [PubMed: 20007329]
20. Cen S, Khorchid A, Javanbakht H, Gabor J, Stello T, Shiba K, Musier-Forsyth K, Kleiman L. Incorporation of lysyl-tRNA synthetase into human immunodeficiency virus type 1. J Virol. 2001; 75:5043–5048. [PubMed: 11333884]
21. Druillennec S, Roques BP. HIV-1 NCp7 as a target for the design of novel antiviral agents. Drug News Perspect. 2000; 13:337–349. [PubMed: 12937655]
22. Athavale SS, Ouyang W, McPike MP, Hudson BS, Borer PN. Effects of the nature and concentration of salt on the interaction of the HIV-1 nucleocapsid protein with SL3 RNA. Biochemistry. 2010; 49:3525–3533. [PubMed: 20359247]
23. Fisher RJ, Rein A, Fivash M, Urbaneja MA, Casas-Finet JR, Medaglia M, Henderson LE. Sequence-specific binding of human immunodeficiency virus type 1 nucleocapsid protein to short oligonucleotides. J Virol. 1998; 72:1902–1909. [PubMed: 9499042]
24. Tisné C, Roques BP, Dardel F. Heteronuclear NMR studies of the interaction of $tRNA^{3Lys}$ with HIV-1 nucleocapsid protein. J Mol Biol. 2001; 306:443–454. [PubMed: 11178904]
25. Tisné C, Roques BP, Dardel F. Specific recognition of primer $tRNA^{Lys3}$ by HIV-1 nucleocapsid protein: involvement of the zinc fingers and the N-terminal basic extension. Biochimie. 2003; 85:557–561. [PubMed: 12763315]
26. Berkowitz R, Fisher J, Goff SP. RNA packaging. Curr Topics Microbiol Immunol. 1996; 214:177–218.
27. Khan R, Giedroc DP. Recombinant human immunodeficiency virus type 1 nucleocapsid (NCp7) protein unwinds tRNA. J Biol Chem. 1992; 267:6689–6695. [PubMed: 1551877]
28. Mély Y, de Rocquigny H, Sorinas-Jimeno M, Keith G, Roques BP, Marquet R, Gérard D. Binding of the HIV-1 nucleocapsid protein to the primer $tRNA^{3Lys}$, *in vitro*, is essentially not specific. J Biol Chem. 1995; 270:1650–1656. [PubMed: 7829498]
29. Hargittai MRS, Mangla AT, Gorelick RJ, Musier-Forsyth K. HIV-1 nucleocapsid protein zinc finger structures induce $tRNA^{Lys3}$ structural changes but are not critical for primer/template annealing. J Mol Biol. 2001; 312:985–997. [PubMed: 11580244]
30. Guo J, Wu T, Anderson J, Kane BF, Johnson DG, Gorelick RJ, Henderson LE, Levin JG. Zinc finger structures in the human immunodeficiency virus type 1 nucleocapsid protein facilitate efficient minus- and plus-strand transfer. J Virol. 2000; 74:8980–8988. [PubMed: 10982342]
31. Morellet N, de Rocquigny H, Mély Y, Jullian N, Déméné H, Ottmann M, Gérard D, Darlix JL, Fournie-Zaluski MC, Roques BP. Conformational behaviour of the active and inactive forms of the nucleocapsid NCp7 of HIV-1 studied by 1H NMR. J Mol Biol. 1994; 235:287–301. [PubMed: 8289249]
32. Poon DT, 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; 70:6607–6616. [PubMed: 8794295]
33. Cusack S. RNA-protein complexes. Curr Opin Struct Biol. 1999; 9:66–73. [PubMed: 10400475]
34. Agris PF. The Importance of being modified: roles of modified nucleosides and Mg^{2+} in RNA structure and function. Progr Nucleic Acid Res Mol Biol. 1996; 53:79–129.

35. Seno T, Agris PF, Söll D. Involvement of the anticodon region of *Escherichia coli* $tRNA^{Gln}$ and $tRNA^{Glu}$ in the specific interaction with cognate aminoacyl-tRNA synthetase. Alteration of the 2-thiouridine derivatives located in the anticodon of the tRNAs by BrCN or sulfur deprivation. Biochim Biophys Acta. 1974; 349:328–338. [PubMed: 4366808]
36. Madore E, Florentz C, Giegé R, Sekine S, Yokoyama S, Lapointe J. Effect of modified nucleotides on *Escherichia coli* $tRNA^{Glu}$ structure and on its aminoacylation by glutamyl-tRNA synthetase. Predominant and distinct roles of the mnm^5 and s^2 modifications of U_{34}. Eur J Biochem. 1999; 266:1128–1135. [PubMed: 10583410]
37. Ashraf SS, Sochacka E, Cain R, Guenther R, Malkiewicz A, Agris PF. Single atom modification (O→S) of tRNA confers ribosome binding. RNA. 1999; 5:188–194. [PubMed: 10024171]
38. Yarian C, Marszalek M, Sochacka E, Malkiewicz A, Guenther R, Miskiewicz A, Agris PF. Modified nucleoside dependent Watson-Crick and wobble codon binding by $tRNA^{Lys}_{UUU}$ species. Biochemistry. 2000; 39:13390–13395. [PubMed: 11063576]
39. Björk GR, Durand JM, Hagervall TG, Leipuviene R, Lundgren HK, Nilsson K, Chen P, Qian Q, Urbonavicius J. Transfer RNA modification: influence on translational frameshifting and metabolism. FEBS Lett. 1999; 452:47–51. [PubMed: 10376676]
40. Stuart JW, Gdaniec Z, Guenther R, Marszalek M, Sochacka E, Malkiewicz A, Agris PF. Functional anticodon architecture of human $tRNA^{Lys3}$ includes disruption of intraloop hydrogen bonding by the naturally occurring amino acid modification, t^6A. Biochemistry. 2000; 39:13396–13404. [PubMed: 11063577]
41. Gustilo EM, Vendeix FA, Agris PF. tRNA's modifications bring order to gene expression. Curr Opin Microbiol. 2008; 11:134–140. [PubMed: 18378185]
42. Chan B, Wiedermaier K, Yip WT, Barbara PF, Musier-Forsyth K. Intra-tRNA distance measurements for nucleocapsid protein-dependent tRNA unwinding during priming of HIV reverse transcription. Proc Natl Acad Sci USA. 1999; 96:459–464. [PubMed: 9892655]
43. Barat C, Schatz O, Le Grice S, Darlix JL. Analysis of the interactions of HIV1 replication primer $tRNA^{Lys,3}$ with nucleocapsid protein and reverse transcriptase. J Mol Biol. 1993; 231:185–190. [PubMed: 7685391]
44. Yarian C, Townsend H, Czestkowski W, Sochacka E, Malkiewicz AJ, Guenther R, Miskiewicz A, Agris PF. Accurate translation of the genetic code depends on tRNA modified nucleosides. J Biol Chem. 2002; 277:16391–16395. [PubMed: 11861649]
45. Hagan N, Fabris D. Direct mass spectrometric determination of the stoichiometry and binding affinity of the complexes between nucleocapsid protein and RNA stem-loop hairpins of the HIV-1 ψ-recognition element. Biochemistry. 2003; 42:10736–10745. [PubMed: 12962498]
46. Avilov SV, Godet J, Piémont E, Mély Y. Site-specific characterization of HIV-1 nucleocapsid protein binding to oligonucleotides with two binding sites. Biochemistry. 2009; 48:2422–2430. [PubMed: 19186983]
47. Agris PF, Marchbank MT, Newman W, Guenther R, Ingram P, Swallow J, Mucha P, Szyk A, Rekowski P, Peletskaya E, Deutscher SL. 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; 18:425–435. [PubMed: 10449040]
48. Mucha P, Szyk A, Rekowski P, Weiss PA, Agris PF. Anticodon domain methylated nucleosides of yeast $tRNA^{Phe}$ are significant recognition determinants in the binding of a phage display selected peptide. Biochemistry. 2001; 40:14191–14199. [PubMed: 11714272]
49. Eshete M, Marchbank MT, Deutscher SL, Sproat B, Leszczynska G, Malkiewicz A, Agris PF. Specificity of phage display selected peptides for modified anticodon stem and loop domains of tRNA. The Protein J. 2007; 26:61–73.
50. Yarian CS, Basti MM, Cain R, Ansari G, Guenther RH, Sochacka E, Czerwinska G, Malkiewicz A, Agris PF. Structural and functional roles of the N1- and N3-protons of ψ at tRNA's position 39. Nucleic Acids Res. 1999; 27:3543–3549. [PubMed: 10446245]
51. Durant PC, Davis DR. Stabilization of the anticodon stem-loop of $tRNA^{Lys,3}$ by an A^+-C base-pair and by pseudouridine. J Mol Biol. 1999; 285:115–131. [PubMed: 9878393]
52. Fasman, GD. Circular dichroism and the conformational analysis of biomolecules. Springer-Verlag; NY: 1996.

53. Martin SR, Schilstra MJ. Circular dichroism and its application to the study of biomolecules. Methods Cell Biol. 2008; 84:263–293. [PubMed: 17964935]
54. Cusack S, Yaremchuk A, Tukalo M. The crystal structures of T. thermophilus lysyl-tRNA synthetase complexed with *E. coli* $tRNA^{Lys}$ and a *T. thermophilus* $tRNA^{Lys}$ transcript: anticodon recognition and conformational changes upon binding of a lysyl-adenylate analogue. EMBO J. 1996; 15:6321–6334. [PubMed: 8947055]
55. Francklyn C, Musier-Forsyth K, Martinis SA. Aminoacyl-tRNA synthetases in biology and disease: new evidence for structural and functional diversity in an ancient family of enzymes. RNA. 1997; 3:954–960. [PubMed: 9292495]
56. Bampi C, Jacquenet S, Lener D, Décimo D, Darlix JL. The chaperoning and assistance roles of the HIV-1 nucleocapsid protein in proviral DNA synthesis and maintenance. Curr HIV Res. 2004; 2:79–92. [PubMed: 15053342]
57. Huang Y, Mak J, Cao Q, Li Z, Wainberg MA, Kleiman L. Incorporation of excess wild type and mutant $tRNA^{Lys3}$ into human immunodeficiency virus type 1. J Virol. 1994; 68:7676–7683. [PubMed: 7966556]
58. Barraud P, Gaudin C, Dardel F, Tisné C. New insights into the formation of HIV-1 reverse transcription initiation complex. Biochimie. 2007; 89:1204–1210. [PubMed: 17383790]
59. Urbaneja MA, Kane BP, Johnson DJ, Gorelick RJ, Henderson LE, Casas-Finet JR. Binding properties of the human immunodeficiency virus type 1 nucleocapside protein p7 to a model RNA: Elucidation of the structural determinants for function. J Mol Biol. 1999; 287:59–75. [PubMed: 10074407]
60. De Guzman RN, Wu ZR, Stalling CC, Pappalardo L, Borer PN, Summers MF. Structure of the HIV-1 nucleocapsid protein bound to the SL3 ψ-RNA recognition element. Science. 1998; 279:384–388. [PubMed: 9430589]
61. Amarasinghe GK, De Guzman RN, Turner RB, Chancellor KJ, Wu ZR, Summers MF. NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. J Mol Biol. 2000; 301:491–511. [PubMed: 10926523]
62. Amarasinghe GK, Zhou J, Miskimon M, Chancellor KJ, McDonald JA, Matthew AG, Miller RR, Rouse MD, Summers MF. Stem-loop SL4 of the HIV-1 ψ RNA packaging signal exhibits weak affinity for the nucleocapsid protein. Structural studies and implications for genome recognition. J Mol Biol. 2001; 314:961–970. [PubMed: 11743714]
63. Cantara WA, Crain PF, Rozenski J, McCloskey JA, Harris KA, Zhang X, Vendeix FAP, Fabris D, Agris PF. The RNA modification database, RNAMDB: 2011 update. Nucleic Acids Res. 2011; 39:D195–D201. [PubMed: 21071406]
64. Jühling F, Mörl M, Hartmann RK, Sprinzl M, Stadler PF, Pütz J. tRNAdb 2009: compilation of tRNA sequences and tRNA genes. Nucleic Acids Res. 2009; 37:D159–D162. [PubMed: 18957446]
65. Wakefield JK, Wolf AG, Morrow CD. Human immunodeficiency virus type 1 can use different tRNAs as primers for reverse transcription but selectively maintains a primer binding site complementary to $tRNA^{3Lys}$. J Virol. 1995; 69:6021–6029. [PubMed: 7545240]
66. Gregoire CJ, Gautheret D, Loret EP. No $tRNA^{3Lys}$ unwinding in a complex with HIV NCp7. J Biol Chem. 1997; 272:25143–25148. [PubMed: 9312125]
67. Huthoff H, Berkhout B. Two alternating structures of the HIV-1 leader RNA. RNA. 2001; 7:143–157. [PubMed: 11214176]
68. Zeng Y, Liu HW, Landes CF, Kim YJ, Ma X, Zhu Y, Musier-Forsyth K, Barbara PF. Probing nucleation, reverse annealing, and chaperone function along the reaction path of HIV-1 single-strand transfer. Proc Natl Acad Sci USA. 2007; 104:12651–12656. [PubMed: 17578926]
69. Wu H, Rouzina I, Williams MC. Single-molecule stretching studies of RNA chaperones. RNA Biol. 2010; 7:73–84.
70. My-Nuong Vo MN, Barany G, Rouzina I, Musier-Forsyth K. Mechanistic studies of mini-TAR RNA/DNA annealing in the absence and presence of HIV-1 nucleocapsid protein. J Mol Biol. 2006; 363:244–261. [PubMed: 16962137]

71. Cléry A, Blatter M, Allain FH. RNA recognition motifs: boring? Not quite Curr Opin Struct Biol. 2008; 18:290–298.
72. Davidson A, Leeper TC, Athanassiou Z, Patora-Komisarska K, Karn J, Robinson JA, Varani G. Simultaneous recognition of HIV-1 TAR RNA bulge and loop sequences by cyclic peptide mimics of Tat protein. Proc Natl Acad Sci USA. 2009; 106:11931–11936. [PubMed: 19584251]
73. Michael LA, Chenault JA, Miller BR 3rd, Knolhoff AM, Nagan MC. Water, shape recognition, salt bridges, and cation-pi interactions differentiate peptide recognition of the HIV rev-responsive element. J Mol Biol. 2009; 392:774–786. [PubMed: 19631217]
74. Hofstadler SA, Sannes-Lowery KA, Hannis JC. Analysis of nucleic acids by FTICR MS. Mass Spectrom Rev. 2005; 24:265–285. [PubMed: 15389854]
75. Turner KB, Brinson RG, Yi-Brunozzi HY, Rausch JW, Miller JT, Le Grice SF, Marino JP, Fabris D. Structural probing of the HIV-1 polypurine tract RNA:DNA hybrid using classic nucleic acid ligands. Nucleic Acids Res. 2008; 36:2799–2810. [PubMed: 18400780]
76. Hagan NA, Fabris D. Dissecting the protein-RNA and RNA-RNA interactions in the nucleocapsid-mediated dimerization and isomerization of HIV-1 stem loop 1. J Mol Biol. 2007; 365:396–410. [PubMed: 17070549]
77. Agris PF, Vendeix FAP, Graham WD. tRNA's wobble decoding of the genome: 40 years of modification. J Mol Biol. 2007; 366:1–13. [PubMed: 17187822]
78. Hartsel SA, Kitchen DE, Scaringe SA, Marshall WS. RNA oligonucleotide synthesis via 5′-silyl-2′-orthoester chemistry. Methods Mol Biol. 2005; 288:33–50. [PubMed: 15333896]
79. Agris PF, 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; 77:125–134. [PubMed: 7599270]
80. Ogilvie KK, Usman N, Nicoghosian K, Cedergren RJ. Total chemical synthesis of a 77-nucleotide-long RNA sequence having methionine-acceptance activity. Proc Natl Acad Sci USA. 1988; 85:5764–5768. [PubMed: 3413059]
81. Guenther RH, Gopal DH, Agris PF. Purification of transfer RNA species by single-step ion-exchange high performance liquid chromatography. J Chromatogr. 1988; 444:79–87. [PubMed: 3204142]
82. Gehrke CW, Kuo KC, McCune RA, Gerhardt KO, Agris PF. Quantitative enzymatic hydrolysis of tRNAs: reversed-phase high-performance liquid chromatography of tRNA nucleosides. J Chromatogr. 1982; 230:297–308. [PubMed: 7050138]
83. Lee BM, De Guzman RN, Turner BG, Tjandra N, Summers MF. Dynamical behavior of the HIV-1 nucleocapsid protein. J Mol Biol. 1998; 279:633–649. [PubMed: 9641983]
84. Cheadle C, Ivashchenko Y, South V, Searfoss GH, French S, Howk R, Ricca GA, Jaye M. Identification of a Src SH3 domain binding motif by screening a random phage display library. J Biol Chem. 1994; 269:24034–24039. [PubMed: 7929055]
85. Chen L, Wang Y, Liu X, Dou S, Liu G, Hnatowich DJ, Rusckowski M. A new TAG-72 cancer marker peptide identified by phage display. Cancer Lett. 2008; 272:122–132. [PubMed: 18723274]
86. Scheunemann AE, Graham WD, Vendeix FAP, Agris PF. Binding of aminoglycoside antibiotics to helix 69 of 23S rRNA. Nucleic Acids Res. 2010; 38:3094–3105. [PubMed: 20110260]
87. Gagnon KT, Zhang X, Agris PF, Maxwell ES. Assembly of the archaeal box C/D sRNP can occur via alternative pathways and requires temperature-facilitated sRNA remodeling. J Mol Biol. 2006; 362:1025–1042. [PubMed: 16949610]

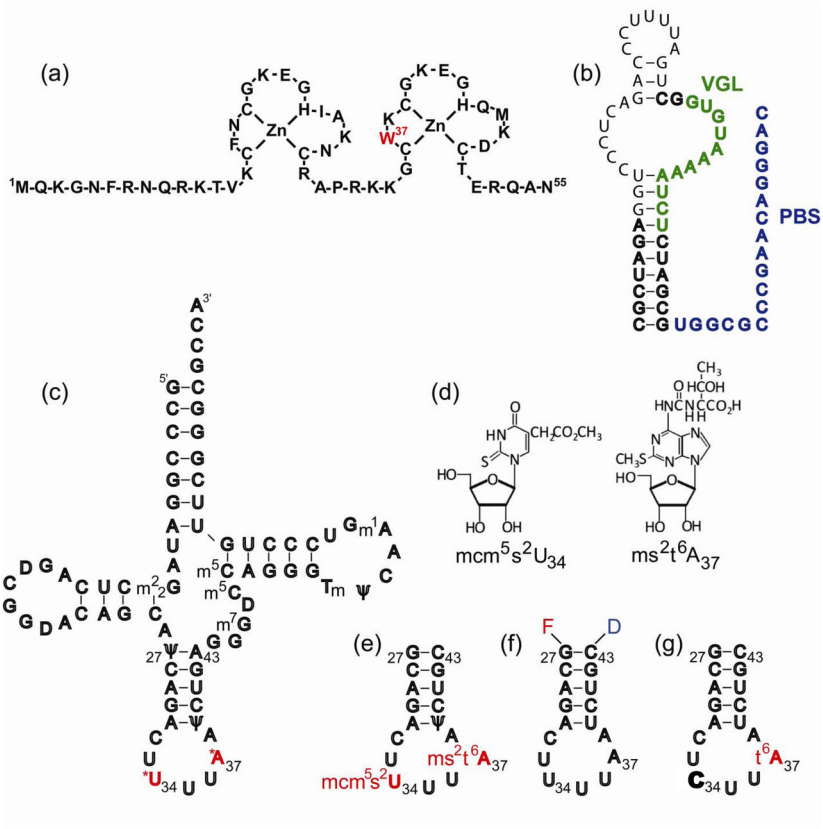

Figure 1.

HIV nucleocapsid protein, NCp7, human $tRNA^{Lys3}_{UUU}$, its anticodon stem and loop domain and its modified nucleosides. **(a)** NCp7 sequence of 55 amino acids and zinc finger structures. The single tryptophan is denoted, W37. **(b)** Sequence and secondary structure of the Primer Binding Site (PBS) and A-rich Stem and Loop 1 of HIV-1, sero(sub)type G, viral RNA. The secondary structure of the HIV-1 sero(sub)type-G, having the highest complementariness of the Loop I sequences, is drawn from that reported for the HIV-1 sero(sub)type B, the most prevalent serotype in the Americas, Europe and Oceania.[17] The A-rich bulge of the viral serotype-G Loop 1 sequence (**VGL; green**) binds the U-rich anticodon domain of the $htRNA^{Lys3}_{UUU}$ and is adjacent to the 18-residue primer binding site (**blue**). The single stranded Loop 1 sequence with significant complementarity to the anticodon domain was used alone and in a truncated construct (bolded) with the anticodon domain fragment for experiments with NCp7. **(c)** Sequence and secondary structure of human $tRNA^{Lys3}$ ($htRNA^{Lys3}_{UUU}$) with all of its known modified nucleosides: N^2,N^2-dimethylguanosine at position 10, $m^2{}_2G_{10}$; dihydrouridine at positions 16, 20 and 48, D; pseudouridine-27, 39 and 55, ψ; 5-methoxycarbonylmethyl-2-thio uridine-34, $mcm^5s^2U_{34}$ (***U**$_{34}$* in red); 2-methylthio-N^6-threonylcarbamoyladenosine-37, $ms^2t^6A_{37}$ (***A**$_{37}$* in red); N^7-methyl-guanosine, m^7G; 5-methylcytidine at positions 48 and 49, m^5C; 2′-O-methylribothymidine-54 (2′-O-methyl-5-methyluridine), Tm; N^1-methyladenosine-58, m^1A. **(d)** Chemical structures of the hypermodified nucleosides within the anticodon stem and loop domain of $htRNA^{Lys3}_{UUU}$: $mcm^5s^2U_{34}$, and $ms^2t^6A_{37}$. **(e)** Modified anticodon stem and loop domain of $htRNA^{Lys3}$ ($hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$;ψ_{39}). The modifications $mcm^5s^2U_{34}$ and $ms^2t^6A_{37}$ are in red. The ASLs were synthesized with a G_{27}–C_{34} terminal base pair instead of the ψ_{27}–A_{43} in order to stabilize the stem. The $hASL^{Lys3}_{UUU}$-$mcm^5s^2U_{34}$;$ms^2t^6A_{37}$;ψ_{39} with a 3′-terminal biotin was used for the peptide

selection, whereas the hASL^{Lys3}-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$ was used for the characterization of the peptides. **(f)** The unmodified hASL$^{Lys3}_{UUU}$ sequence. The unmodified hASL$^{Lys3}_{UUU}$ was synthesized with and without fluorescein conjugated to the 5′-terminus (**F**) and dabsyl at the 3′-terminus (**D**). **(g)** The anticodon domain of htRNA$^{Lys1}_{CUU}$ with the modification *N*^{6}-threonylcarbamoyladenosine (t^{6}A$_{37}$, in red), hASL$^{Lys1}_{CUU}$-t^{6}A$_{37}$. The modified and unmodified dodecamer anticodon domain fragments used in experiments with NCp7 were comprised of residues 29–40 and modifications were introduced at U$_{34}$, A$_{37}$ and U$_{39}$).

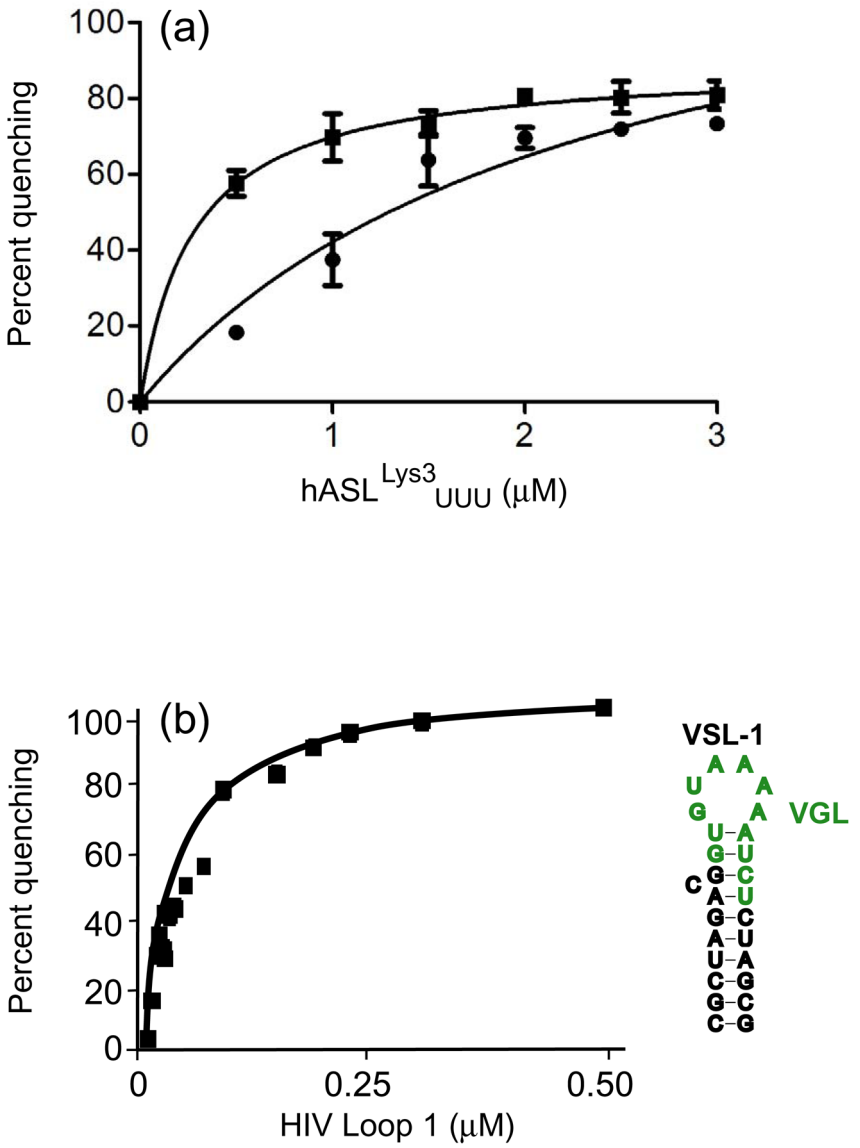

Figure 2.

NCp7 association with the anticodon stem and loop domain of htRNA$^{Lys3}_{UUU}$. The binding of hASL$^{Lys3}_{UUU}$ to NCp7 was monitored by observing the change in fluorescence (quenching) of the protein's one tryptophan (W37). **(a)** NCp7 (75 nM) titrated with increasing concentrations (0 – 3 μM of the hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$ (●) and the unmodified hASL$^{Lys3}_{UUU}$ (■). **(b)** NCp7 binding of the HIV-1, serotype-G, viral Stem and Loop 1. The quenching of the NCp7 W37 was monitored in a single experiment when titrated with a truncated construct of the A-rich HIV-1, serotype-G, viral Stem and Loop 1 (VSL-1 with VGL).

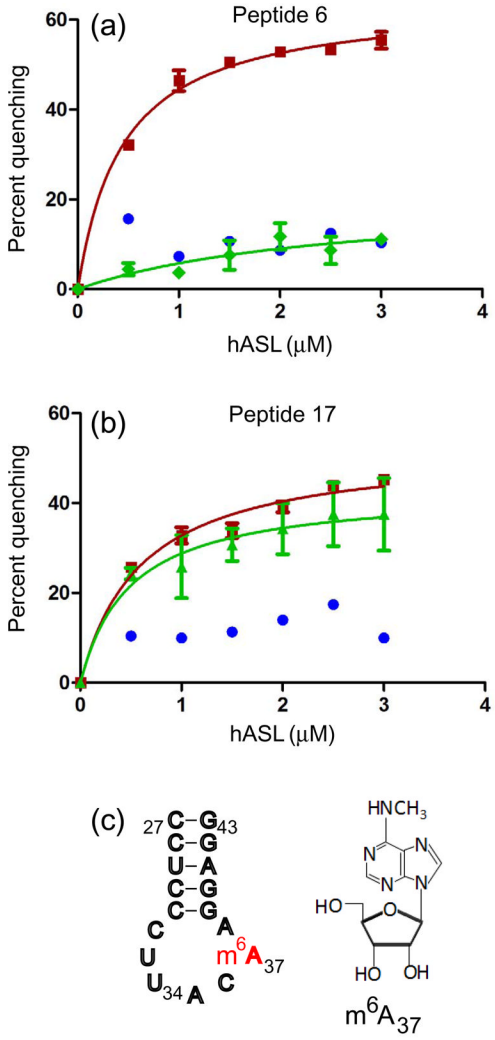

Figure 3.

The association of phage display selected peptides with hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$. **(a)** Peptide P6 binding of hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$. The fluorescence of the FITC-conjugated peptide 6 (P6) (percent quenching) was monitored with addition of the ASL. P6 titrated with hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$ (■); P6 titrated with the unmodified hASL$^{Lys3}_{UUU}$ (●); P6 titrated with *E. coli* ASL$^{Val3}_{UAC}$-m^{6}A$_{37}$ (◆). **(b)** Peptide P17 binding of hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$. P17 titrated with the hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$ (■); P17 titrated with the unmodified hASL$^{Lys3}_{UUU}$ (●); P17 titrated with ASL$^{Val3}_{UAC}$-m^{6}A$_{37}$ (◆). **(c)** The anticodon stem and loop of *E. coli* tRNA$^{Val3}_{UAC}$ with the modification *N*^{6}-methyladenosine at position 37 (m^{6}A$_{37}$) (ASL$^{Val3}_{UAC}$-m^{6}A$_{37}$).

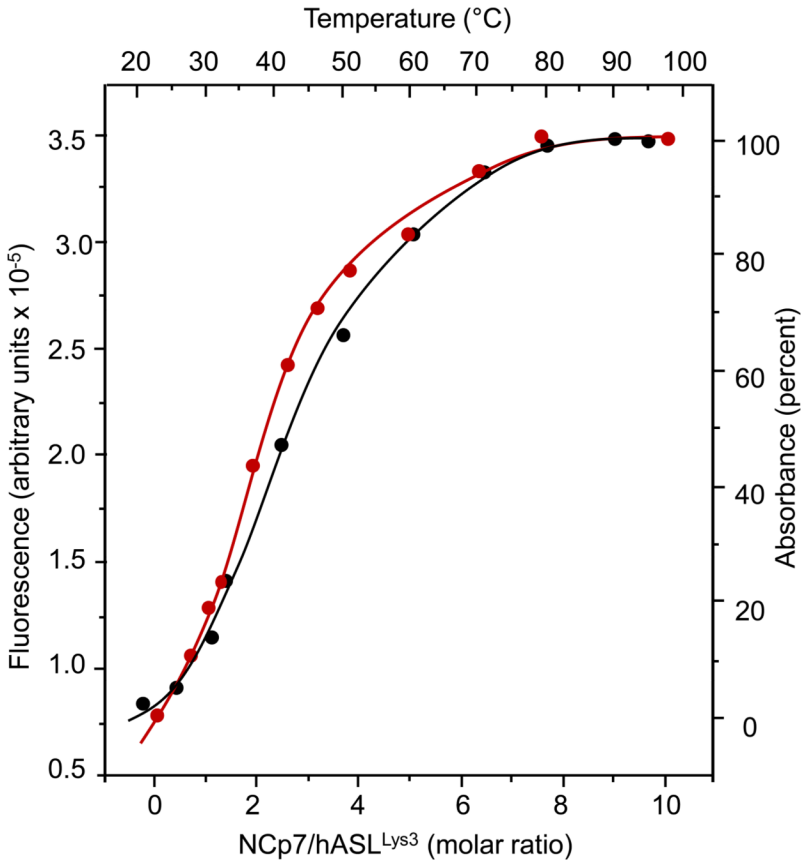

Figure 4.

NCp7 disruption of the htRNA$^{Lys3}_{UUU}$ anticodon stem and loop domain compared to thermal denaturation. An unmodified anticodon stem and loop domain construct (hASL$^{Lys3}_{UUU}$) with a 5′-conjugated fluorescein and a 3′-conjugated dabsyl was titrated with NCp7 and the increase in fluorescence monitored (●). The hASL$^{Lys3}_{UUU}$ was half "melted" at an NCp7 to RNA ratio of ~2. In addition, an unmodified hASL$^{Lys3}_{UUU}$ (without fluorescein or dabsyl) was thermally denatured/renatured from 20 to 95 °C and the UV absorbance at 260 nm was recorded, averaged and normalized (percent absorbance, ●). The temperature at which the sample was half-melted was ~40 °C.

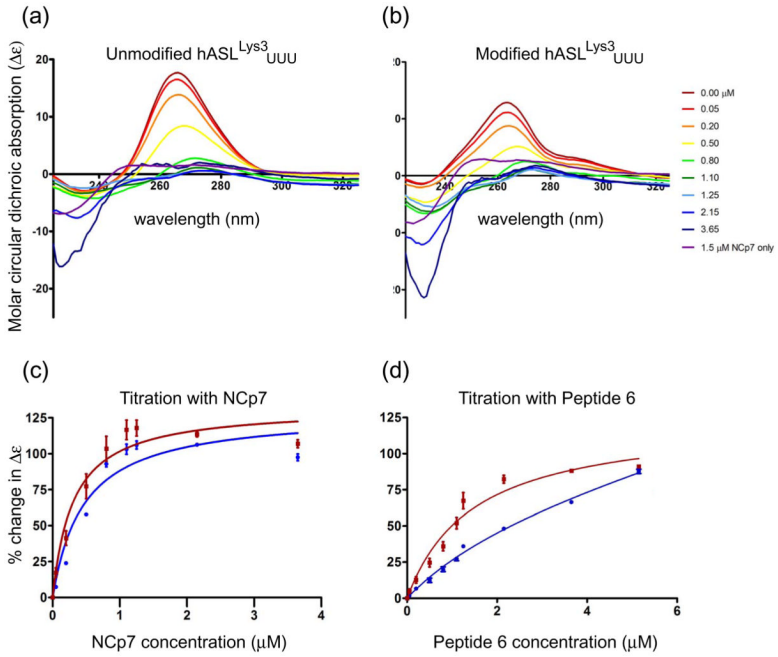

Figure 5.

NCp7 and peptide P6 remodeling of the modified and unmodified hASL$^{Lys3}_{UUU}$ as monitored with circular dichroism. **(a)** The circular dichroism spectra (225 – 325 nm) of the unmodified hASL$^{Lys3}_{UUU}$ alone and with increasing concentrations of NCp7. Increasing concentrations of NCp7 decreased the CD spectral ellipticity (255–275 nm) of the hASL$^{Lys3}_{UUU}$. The hASL$^{Lys3}_{UUU}$ (1.5 μM) was titrated with NCp7 (0.0 – 3.65 μM). A spectrum of the NCp7 alone (1.50 μM; purple) demonstrates that the peptide has approximately a zero ellipticity between 255 and 275 nm. **(b)** The circular dichroism spectra (225 – 325 nm) of the triply modified hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$;ψ$_{39}$ alone and with increasing concentrations of NCp7. Increasing concentrations of NCp7 (0.0 – 3.65 μM) decreased the CD spectral ellipticity (255–275 nm) of the ASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$;ψ$_{39}$ (1.5 μM). A spectrum of the NCp7 alone (1.50 μM; purple) demonstrates a zero ellipticity between 255 and 275 nm. **(c)** Changes in ellipticity of unmodified ASL$^{Lys3}_{UUU}$ and modified hASL$^{Lys3}_{UUU}$ when titrated by NCp7. CD spectral changes were monitored at the wavelength maximum during titration of hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$;ψ$_{39}$ (■) and the unmodified hASL$^{Lys3}_{UUU}$ (●). **(d)** Changes in ellipticity of unmodified ASL$^{Lys3}_{UUU}$ and hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$;ψ$_{39}$ when titrated by peptide P6. CD spectral changes were monitored at the wavelength maximum during titration of hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$;ψ$_{39}$ (■) and the unmodified hASL$^{Lys3}_{UUU}$ (●).

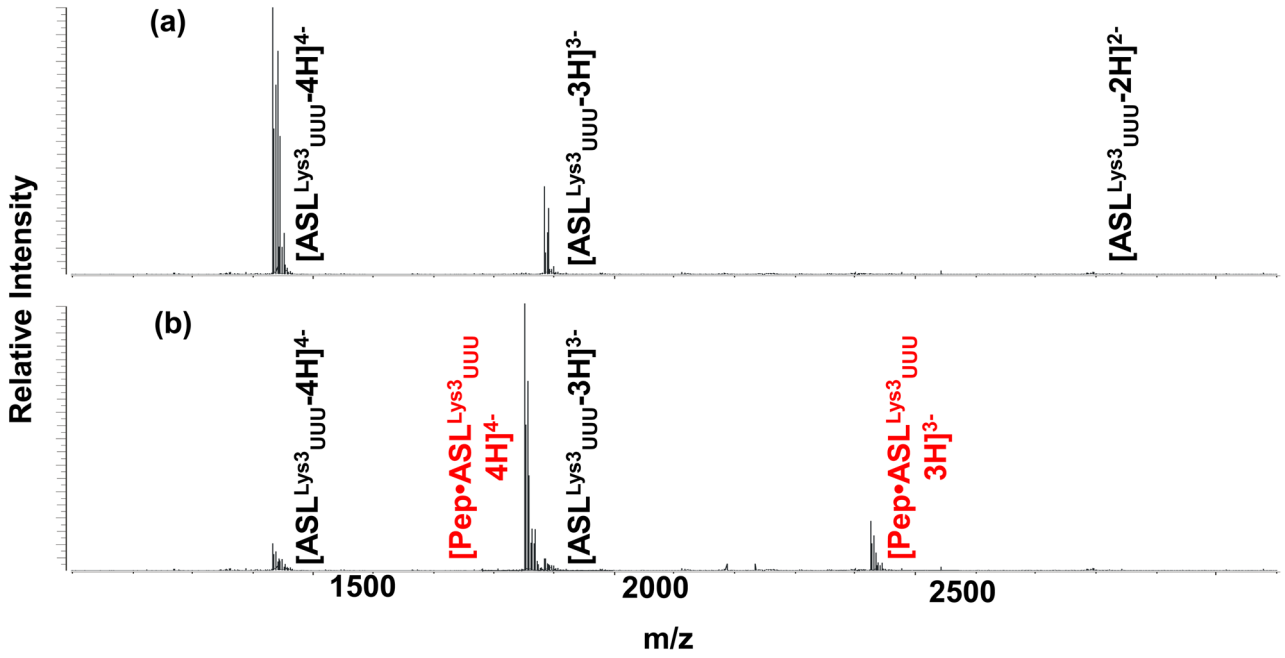

Figure 6.

Mass spectrometry of hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$ bound by P6. **(a)** ESI-MS of hASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$ (relative intensity vs. mass/charge ratio) **(b)** ESI-MS of hASL^{Lys3}-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$ bound to P6. The ratio of P6 to ASL is 5/1.

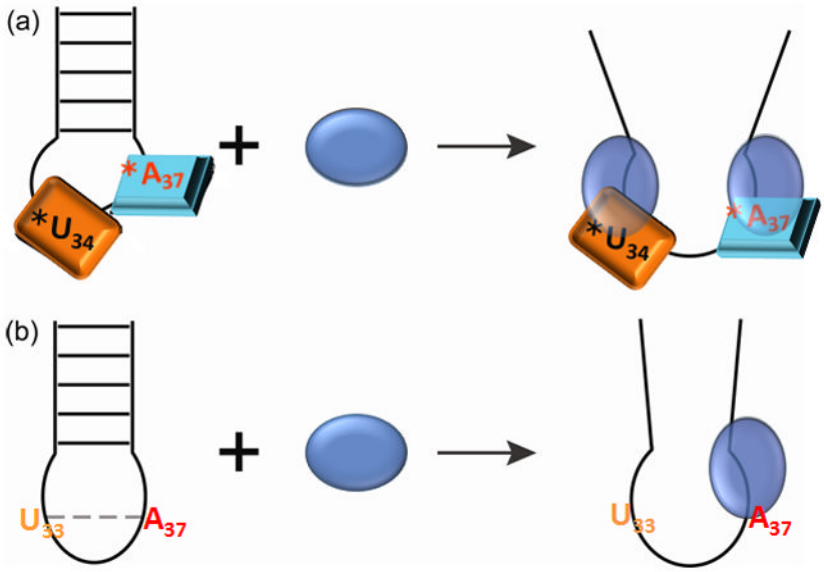

Figure 7.

Natural modifications provide identity determinants and a physicochemical advantage to NCp7 remodeling of htRNA$^{Lys3}_{UUU}$. **(a)** NCp7 (blue oval) interaction with the modified anticodon domain of htRNA$^{Lys3}_{UUU}$. The modifications of the anticodon stem and loop domain of htRNA$^{Lys3}_{UUU}$, mcm^{5}s^{2}U$_{34}$ (*U$_{34}$) and ms^{2}t^{6}A$_{37}$ (*A$_{37}$), provide unique identity determinants for NCp7. **(b)** NCp7 interaction with the unmodified anticodon domain of htRNA$^{Lys3}_{UUU}$. In contrast to the modified anticodon domain of htRNA$^{Lys3}_{UUU}$, the unmodified domain lacks the identity determinants and has intra-loop hydrogen bonds (----) between the invariant U$_{33}$ and A$_{37}$.[40]

Table 1

Affinity of NCp7, and peptides P6 and P17 for hASL^{Lys}

NCp7 or Peptide	ASL^{Lys}, Fragment, or HIV Stem/Loop I	Dissociation Constant K_d (μM)
NCp7	Unmodified ASL$^{Lys3}_{UUU}$	2.30 ± 0.62
	ASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$	0.28 ± 0.03
	ASL$^{Lys1}_{CUU}$-t^{6}A$_{37}$	ND
	HIV Stem/Loop I (VSL-1)	0.3
	Unmodified Fragment$_{29–40}$	0.62 ± 0.12
	Fragment$_{29–40}$-s^{2}U$_{34}$	0.01 ± 0.01
	Fragment$_{29–40}$-ψ$_{39}$	0.48 ± 0.03
	Fragment$_{29–40}$-mnm^{5}U$_{34}$	0.20 ± 0.07
	Fragment$_{29–40}$-t^{6}A$_{37}$	0.15 ± 0.06
	Fragment$_{29–40}$-mnm^{5}U$_{34}$;ψ$_{39}$	0.25 ± 0.05
	Fragment$_{29–40}$-mnm^{5}U$_{34}$;t^{6}A$_{37}$	0.26 ± 0.06
	A-rich Loop 1 dodecamer	0.06 ± 0.01
	Duplex Anticodon/Loop 1	1.11 ± 0.16
P6	Unmodified ASL$^{Lys3}_{UUU}$	ID
	ASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$	0.50 ± 0.10
	ASL$^{Lys1}_{CUU}$-t^{6}A$_{37}$	0.60 ± 0.09
P17	Unmodified ASL^{Lys3}	ID
	ASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$	0.60 ± 0.10
	ASL$^{Val3}_{ACU}$-m^{6}A$_{37}$	0.49 ± 0.31
	ASL$^{Lys1}_{CUU}$-t^{6}A$_{37}$	1.80 ± 0.20
P21	Unmodified ASL$^{Lys3}_{UUU}$	0.98 ± 0.03
	ASL$^{Lys3}_{UUU}$-mcm^{5}s^{2}U$_{34}$;ms^{2}t^{6}A$_{37}$	1.30 ± 0.04

ND = experiment not done

ID = Indeterminable, could not be calculated from the data

Table 2

Peptides Generated from Phage Display Library Screening

Clone No.	15 and 16 Amino acid Peptides	Library/Elution	Times Appearing
1	FSVSFPSLPAPPDRS	Fuse5/basic	18
3	GRVTYYSCGVSLLFQ	Fuse5/basic	4
4	AGPVPLHSLSYYYNQ	Fuse5/basic	1
5	RAVMTVVWPVSFAGF	Fuse5/acidic	5
6	**RVTHHAFLGAHRTVG**	**Fuse5/acidic**	**10**
8	PAVASTSSLIIDGPF	Fuse5/acidic	2
9	PKAFQYGGRAVGGLW	Fuse5/acidic	1
10	AAHVSEHYVSGSLRP	Fuse5/acidic	1
11	ASVGPAPWAMTPPVS	Fuse5/acidic	1
12	APALWYPWRSLLPLY	Fuse5/acidic	1
13	ASLHPVPKTWFFLLS	Fuse5/acidic	1
14	WSHSRNTADVPVSML	Fuse5/acidic	1
15	HRGYCRDRWNCGEYF	F88-cys6/basic	8
17	**PHRQCSAPAKSCKILP**	**F88-cys6/basic**	**8**
19	TLPACHELPKHCKRRG	F88-cys6/basic	4
20	TLPACHELPKHCNEAR	F88-cys6/basic	1
21	NGPECNAYMVRCRGYH	F88-cys6/acidic	4
23	GNSNCPMLNEQCPWQD	F88-cys6/acidic	1
24	HTETCINIRNTCTTVA	F88-cys6/acidic	1
25	LKLPCKITINNCQLAG	F88-cys6/acidic	1