

Article

Mapping m⁶A Sites on HIV-1 RNA Using Oligonucleotide LC-MS/MS

Alice Baek^{1,2,3,4,†} , Asif Rayhan^{5,†}, Ga-Eun Lee^{1,2,3,4} , Sarah Golconda^{1,2,3,4}, Hannah Yu^{1,2,3,4} ,
Shihyoung Kim^{1,2,3,4}, Patrick A. Limbach⁵ , Balasubrahmanyam Addepalli^{5,*}  and Sanggu Kim^{1,2,3,4,*} 

- ¹ Center for Retrovirus Research, The Ohio State University, Columbus, OH 43210, USA; baek.71@buckeyemail.osu.edu (A.B.); lee.8950@buckeyemail.osu.edu (G.-E.L.); golconda.2@buckeyemail.osu.edu (S.G.); yu.2127@osu.edu (H.Y.); kim.6754@osu.edu (S.K.)
² Department of Veterinary Biosciences, The Ohio State University, Columbus, OH 43210, USA
³ Infectious Diseases Institute, The Ohio State University, Columbus, OH 43210, USA
⁴ Translational Data Analytics Institute, The Ohio State University, Columbus, OH 43210, USA
⁵ Rieveschl Laboratories for Mass Spectrometry, Department of Chemistry, University of Cincinnati, Cincinnati, OH 45221, USA; rayhanaf@mail.uc.edu (A.R.); limbacpa@ucmail.uc.edu (P.A.L.)
* Correspondence: bala_addepalli@waters.com (B.A.); kim.6477@osu.edu (S.K.); Tel.: +1-614-292-6524 (S.K.)
† These authors contributed equally to this work.

Abstract: The biological significance of chemical modifications to the ribonucleic acid (RNA) of human immunodeficiency virus type-1 (HIV-1) has been recognized. However, our understanding of the site-specific and context-dependent roles of these chemical modifications remains limited, primarily due to the absence of nucleotide-resolution mapping of modification sites. In this study, we present a method for achieving nucleotide-resolution mapping of chemical modification sites on HIV-1 RNA using liquid chromatography and tandem mass spectrometry (LC-MS/MS). LC-MS/MS, a powerful tool capable of directly analyzing native RNAs, has proven effective for mapping RNA modifications in small RNA molecules, including ribosomal RNA and transfer RNA. However, longer RNAs have posed challenges, such as the 9 Kb HIV-1 virion RNA, due to the complexity of and ambiguity in mass differences among RNase T1-cleaved RNA fragments in LC-MS/MS data. Here, we introduce a new target RNA enrichment method to isolate small local RNA fragments of HIV-1 RNA that potentially harbor site-specific N⁶-methyladenosine (m⁶A) modifications. In our initial trial, we used target-specific DNA probes only and encountered insufficient RNA fragmentation due to inefficient S1 digestion near the target site. Recognizing that inefficient S1 digestion by HIV-1 RNA is likely due to the formation of secondary structures in proximity to the target site, we designed multiple DNA probes annealing to various sites of HIV-1 RNA to better control the structures of RNA substrates for S1 digestion. The use of these non-target DNA probes significantly improved the isolation of more homogeneous target RNA fragments of approximately 50 bases in length. Oligonucleotide LC-MS/MS analysis of these isolated target RNA fragments successfully separated and detected both m⁶A-methylated and non-methylated oligomers at the two m⁶A-predicted sites. The principle of this new target enrichment strategy holds promise and should be broadly applicable to the analysis of any lengthy RNA that was previously deemed infeasible for investigation using oligonucleotide LC-MS/MS.

Keywords: RNA modification; N⁶-methyladenosine (m⁶A); HIV-1; mass spectrometry; oligonucleotide LC-MS/MS; target RNA enrichment



Citation: Baek, A.; Rayhan, A.; Lee, G.-E.; Golconda, S.; Yu, H.; Kim, S.; Limbach, P.A.; Addepalli, B.; Kim, S. Mapping m⁶A Sites on HIV-1 RNA Using Oligonucleotide LC-MS/MS. *Methods Protoc.* **2024**, *7*, 7. <https://doi.org/10.3390/mps7010007>

Academic Editor: Anthony Purcell

Received: 17 November 2023

Revised: 12 December 2023

Accepted: 13 December 2023

Published: 10 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Chemical modifications on RNA, collectively known as epitranscriptomes, have been demonstrated to influence various aspects of RNA, including stability, folding, translocation, alternative splicing, and translation [1]. The dysregulation of epitranscriptomes has been implicated in more than 100 human diseases [2–4]. Since their initial discovery in

the late 1950s, over 300 different types of chemical modifications to RNA have been identified [3,5]. However, we have only recently begun to comprehend the global distribution of these modifications thanks to advances in sequencing technologies [6].

Most studies to date have utilized short-read sequencing methods, involving the reverse transcription of fragmented RNA and PCR amplification of reverse-transcribed cDNA, providing only low-resolution and population-average values of chemical modifications. While useful in identifying chemical modification sites without prior knowledge, short-read data are complicated by potential biases introduced during antibody-mediated target RNA enrichment and cDNA preparation [6,7]. Due to the lack of nucleotide-resolution mapping of site-specific modifications, our understanding of chemical modifications is mostly based on cellular perturbation (overexpression or knockout/knockdown) of host effectors [6,8–15], overlooking the site-specific roles of the epitranscriptome [2,16–21]. These studies often reported disparate effects of epitranscriptome depending on gene type, cellular status and viral species [1,9,22].

Mass spectrometry (MS) is a powerful tool, capable of directly analyzing chemical modifications of native RNA molecules by measuring the mass shift of RNA oligomers [23,24]. To identify an unknown modification (a new type of chemical modification), one can employ complete hydrolysis of RNA before LC-MS [25]; however, this approach does not provide modification-site information. Oligonucleotide LC-MS/MS, however, identifies both the positions and types of chemical modifications at single-nucleotide resolution. In this technology, chemical modification sites can be identified via nucleobase-specific digestion of RNA molecules via guanosine (G)-specific ribonuclease T1, cytidine (C)-specific Cusativin, or uridine (U)-specific MC1 ribonucleases, followed by liquid-chromatography-mediated separation of digested products and tandem mass spectrometry of separated oligonucleotides [23,24,26]. Selective cleavage of RNA with one of these nucleobase-specific RNases generates RNA fragments simple enough for MS to detect. These RNA fragments possess unique sequences and display mass differences between chemically modified and unmodified status. This technology can determine chemical modifications in RNAs with known sequences. However, the positions and types of modifications can be identified without prior information [23,24,26]. This direct RNA analysis is also free from the potential technical biases associated with short-read sequencing methods [6,7].

HIV-1 RNAs have significantly higher epitranscriptomic modifications than cellular RNA [12,27]. However, the precise roles of these modifications in HIV-1 replication and their evolutionary benefits remain unclear, exhibiting both pro- and anti-viral effects depending on the experimental conditions and tested replication stages [9,15,28–31]. Despite frequent cleavage with nucleobase-specific RNases, the analysis of small RNAs, such as transfer RNA (tRNA) and ribosomal RNA (rRNA) with oligonucleotide LC-MS/MS, has proven effective [7,24,32–38]. However, the analysis of larger RNAs, such as the 9 Kb HIV-1 RNA, remains nearly impossible due to the ambiguity in mass differences among RNase-cleaved fragments in LC-MS/MS data. For example, RNase T1 digestion generates many RNA fragments with the same molecular weight, originating from different parts of the HIV-1 RNA. This prevents the unambiguous mapping of RNase T1-cleaved fragments. In this study, we developed a new method that can effectively enrich target RNA fragments within a mixed pool of long RNA strands. With the new method, we fragmented 9 Kb HIV-1 virion RNA via S1 nuclease digestion and selectively enriched and purified small (~50 bases) target RNA fragments with sufficient purity for oligonucleotide LC-MS/MS analysis. The MS data confirmed the precise locations of two dominant m⁶A modifications on HIV-1 RNA.

2. Materials and Methods

2.1. HIV-1 Viral RNA Extraction

HIV-1 virion RNAs were extracted from HEK293T cells transfected with the HIV-1 proviral DNA construct, pNL4-3, following the procedures described previously [39] (Figure 1a). Briefly, we transfected HEK293T cells with pNL4-3 using polyethylenimine

(PEI) following the manufacturer's instructions. The cell culture medium was exchanged with fresh medium at 6 h post-transfection and the supernatant was harvested at 72 h for virion RNA extraction. To purify viral particles, the supernatant was subjected to a 10% sucrose gradient centrifugation at $28,000\times g$ at $4\text{ }^{\circ}\text{C}$ for 1 h 40 min. After discarding the supernatant, the pellet was resolved in $160\text{ }\mu\text{L}$ 1X HBSS and treated with DNase I treatment (12 U/reaction; New England Biolab, Ipswich, MA, USA) for 30 min at $37\text{ }^{\circ}\text{C}$. Virion RNA was extracted using TRIzol™ (Thermo Fisher, Waltham, MA, USA) following the company-provided standard protocol. The RNA pellet after the 70% ethanol precipitation was air-dried in the hood and then resuspended in $30\text{ }\mu\text{L}$ of DEPC-treated water followed by the addition of RNase inhibitor (New England Biolabs; final concentration 1 U/uL).

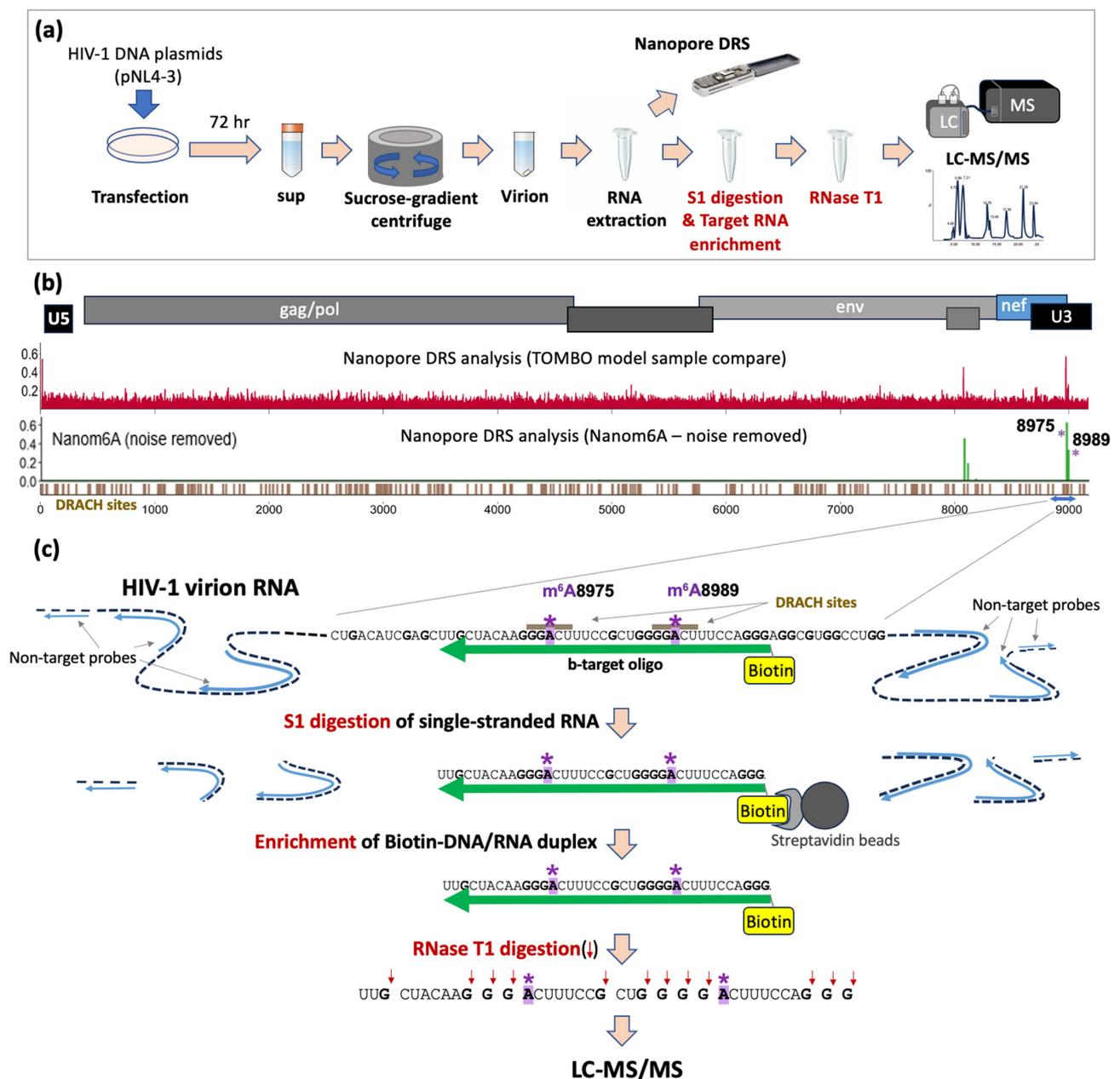


Figure 1. Schematic view of HIV-1 virion RNA analysis. (a) HIV-1 virion RNA samples were prepared by transfection of HEK293T cells with HIV-1 proviral DNA plasmids (pNL4-3), followed by sucrose-gradient of supernatant at 72 h post transfection and RNA extraction from purified virions [39]. Virion RNA was then analyzed both by Nanopore Direct RNA Sequencing [40] and RNase T1-mediated oligonucleotide LC-MS for the detection of site-specific chemical modifications. (b) Nanopore Direct

RNA Sequencing (DRS) of full-length HIV-1 RNA analysis via Tombo–model–sample–compare (Tombo-MS), and Nanom6A are shown. Both revealed distinct site-specific modification signals at A8975 and A8989 (both denoted by *), coinciding with m⁶A DRACH motifs. (c) A schematic view of oligonucleotide LC-MS for the confirmation of m⁶As at A8975 and A8989 sites. The nucleotide sequence of the local area near A8975 and A8989, denoted by the blue arrow in (b), is shown. Target RNA fragments, harboring both A8975 and A8989, are enriched via selective S1 digestion on single-stranded RNA sites, followed by streptavidin bead isolation of biotinylated DNA/RNA (b-DNA/RNA) duplex. Blue lines indicate non-target HIV-1 RNA-specific probes (length 20–25 bp) and green lines indicate biotinylated DNA probe (b-target-oligo) binding to the target RNA site. S1 nuclease digestion removes single-stranded RNA, which results in DNA and RNA duplexes. The Biotin-DNA/RNA duplex recovered using streptavidin magnetic beads was subjected to RNase T1 digestion and LC-MS analysis.

2.2. HIV-1-Specific Oligos

We used two sets of DNA oligonucleotide to enrich the target RNA fragments: (i) biotinylated, target-specific oligomers (b-target: /5'-GCU ACA AGG GAC UUU CCG CUG GGG ACU UUC CAG GGA G-3'; Integrated DNA Technologies) and (ii) non-target oligomers specific to 110 different sites of HIV-1 at least 40–50 bases away from each other (non-target oligo sequences are available in Supplementary Table S1).

2.3. S1 Nuclease Digestion and the Recovery of Biotinylated-DNA/RNA Duplex

To enrich target RNA, 10 µg of HIV-1 virion RNA was initially incubated with either b-target oligos only (HIV-1 RNA:DNA oligo molar ratio of 1:100) or a mixture of b-target oligos and non-target oligos (at an HIV-1 RNA:DNA molar ratio of 1:30) at 65 °C for 5 min in 0.5X nuclease-free duplex buffer (Integrated DNA Technologies; 30 mM HEPES, pH 7.5; 100 mM potassium acetate) in heat block and then slowly cooled to room temperature. It was followed by nuclease S1 digestion (Invitrogen; final concentration 2.5 U/µL). The sample was briefly mixed by tapping and left for 2 h at room temperature for RNA digestion. The S1-digested RNA samples were purified using phenol–chloroform [41] and eluted in 40 µL TE buffer.

The target DNA:RNA duplexes with b-target oligos were recovered using Streptavidin beads (Dynabeads™ MyOne™ C1, Thermo Fisher Scientific) following the manufacturer's instructions. Briefly, the S1-digested RNA sample was mixed with the beads, resuspended in 2X binding and washing buffer (equal volume with the sample,) and incubated for 10 min at room temperature on HulaMixer™ Sample Mixer (Thermo Fisher) in vibration mode (100 rpm). The streptavidin-RNA sample was then put onto a magnetic stand until the solution became clear and supernatant was discarded. The sample was washed twice with 1X binding and washing buffer by repeating this step. The beads were then resuspended with 20 µL of 4 µM biotin solution (in 0.5% dimethylsulfoxide), briefly mixed via tapping, and left at room temperature for 10 min. The sample tube was briefly spun down and put onto a magnetic stand until it became clear (~5 min). The biotinylated-DNA/RNA (b-DNA/RNA) duplex was then recovered by carefully transferring the supernatant to a new tube. The presence and the size of b-DNA/RNA duplex was analyzed on 12% polyacrylamide gel electrophoresis.

2.4. Oligonucleotide LC-MS/MS of HIV-1 RNA Fragment

For RNase T1 digestion, about 200 ng of the purified target b-DNA/RNA duplex (see 2.3 above) was initially denatured at 95 °C for 2 min, snap cooled at 4 °C, and digested with 50 units of RNase T1 (Worthington, Columbus, OH, USA) at 37 °C for 2 h. The digested sample was then dried in a speedvac system (Thermo Scientific). LC-MS/MS analysis was performed using a BEH C18 column (1.7 µm, 0.3 × 150 mm, Waters, Milford, MA, USA) with Ultimate 3000 Ultra-High-Performance Liquid Chromatography (UHPLC; Thermo Scientific) coupled to a Synapt G2-S (Quadrupole Time-of-Flight, Q-TOF, Waters, Milford,

MA, USA) mass spectrometer. Gradient chromatography was conducted with a flow rate of $5 \mu\text{L min}^{-1}$ using mobile phase A (8 mM TEA and 200 mM HFIP, pH = 7.8 in water) and mobile phase B (8 mM TEA and 200 mM HFIP in 50% methanol) at 60°C . The gradient included an initial hold at 3% mobile phase B for sample loading, followed by a ramp to 55% mobile phase B over 70 min, reaching 99% in 2 min with a 5 min hold before re-equilibration (30 min) at 3% mobile phase B for initial conditions. The digestion products in the chromatographic eluent were detected in negative ion mode via electrospray ionization (ESI) on a Synapt G2-S (Quadrupole Time-of-Flight, Q-TOF). The mass spectrometer was operated in V-mode (sensitivity mode). The ESI conditions involved 2.2 kV at capillary, 30 V at sample cone, with source and desolvation temperatures maintained at 120°C and 400°C , respectively, and gas flow rates at 3 and 600 L h^{-1} , respectively. A scan range of 545–2000 m/z (0.5 s) was employed for first-stage (MS) data acquisition and 250–2000 m/z (1 s) for second (MS/MS)-stage data acquisition. The first stage of MS identifies the mass of the oligonucleotide, while the second stage helps to identify the nucleotide sequence based on sequence-informative fragment ion patterns. The top three most abundant ions in the first stage were chosen to face fragmentation for MS/MS using an m/z -dependent collision energy profile (20–23 V at m/z 545; 51–57 V at m/z 2000) before exclusion for 60 s using the dynamic exclusion feature.

2.5. Synthetic RNA Control

As a positive control experiment for oligonucleotide LC-MS/MS, we used two synthetic RNA oligos each harboring $m^6\text{A}$ at A8975 ($m^6\text{A}8975$: GAA CUG CUG ACA UCG AGC UUG CUA CAA GGG $m^6\text{ACU}$ UUC CGC UGG GGA CUU UCC AGG GAG GCG UGG CCU GGG CGG GAC UGG GGA; Dharmacon) and at A8989 ($m^6\text{A}8989$: GAA CUG CUG ACA UCG AGC UUG CUA CAA GGG ACU UUC CGC UGG GGG $m^6\text{A}$ CUU UCC AGG GAG GCG UGG CCU GGG CGG GAC UGG GGA; Dharmacon). These synthetic oligos were used in parallel with the HIV-1 RNA fragments in “Oligonucleotide LC-MS/MS of HIV-1 RNA fragment” above.

2.6. LC-MS/MS Data Processing

Mongo Oligo mass calculator (<http://mass.rega.kuleuven.be/mass/mongo.htm>; accessed on 11 November 2021) was used to predict the m/z values for RNase T1 digestion products (and their fragment ions) of a 40-base-long HIV-1 RNA sequence. To manually identify and assign the $m^6\text{A}$ modification, a ~ 14 Da mass shift was scored for the theoretically expected oligonucleotides, considering cleavage at the 3'-end of guanosine, in the modified RNA compared to the unmodified counterpart. Control experiments were conducted using synthetic oligos, incorporating $m^6\text{As}$ at either A8975 or A8989 to attribute the $m^6\text{A}$ modification to positions 8975 and 8989, respectively.

3. Results

3.1. Identification of Potential Site-Specific $m^6\text{A}$ Modifications on HIV-1 Viral RNA

Several common chemical modifications, including $m^6\text{A}$, 5-methylcytosine ($m^5\text{C}$), 2'-O-methylation (Nm), and N4-acetylcytidine (ac^4C), have been mapped onto the HIV-1 genome using short-read sequencing methods [10,12–14,30,42]. These studies, however, provide only low-resolution, fragmented information, typically spanning approximately 40 to 200 bases, about modification sites. Due to the lack of nucleotide-resolution data, the site-specific and context-dependent roles of RNA modifications remain largely unknown for HIV-1.

We recently developed novel Nanopore Direct RNA Sequencing (DRS) methods for full-length HIV-1 RNA analysis and demonstrated potential chemical modification sites on full-length HIV-1 RNAs [40]. Our DRS data revealed three predominant site-specific modification signals that coincide with $m^6\text{A}$ sequence motifs (DRACH: D = A/G/U, R = A/G, H = A/C/U) at positions A8079, A8975, and A8989 within the 3' untranslated region (U3) of HIV-1 RNA [40]. Using $m^6\text{A}$ -detection tools, such as $m^6\text{Anet}$ [43] and Nanom^6A [44], these three sites were identified as potential $m^6\text{A}$ modification sites (Figure 1b) [40].

3.2. Enriching Target RNA Fragments of HIV-1 RNA for Oligonucleotide LC-MS/MS

To validate the potential site-specific m⁶As identified with Nanopore DRS, we employed oligonucleotide LC-MS/MS. Given the challenges for the analysis of full-length.

Examining HIV-1 RNA, we developed a new method to enrich target RNA fragments for oligonucleotide LC-MS/MS. The local RNA containing the target sites needs to be fragmented and enriched to a nearly homogeneous state, sufficient for RNase T1 treatment. The aim is to yield unique-sized RNA isomers that oligonucleotide LC-MS/MS can distinguish. To achieve this, we designed a 37-base biotinylated DNA probe (b-target oligo) to selectively purify small target RNA fragments that harbor potential m⁶A sites (A8975 and A8989) (Figure 1c). An S1 digestion of b-target-probe-bound HIV-1 RNA, followed by streptavidin-bead isolation, enriched target RNA/DNA duplexes (b-DNA/RNA) suitable for oligonucleotide LC-MS/MS. Initial trials for target enrichment using the b-target-oligo probe alone substantially reduced RNA fragments in general (Figure 2(ai)), but the majority of enriched RNAs were substantially larger than expected, heterogeneous in length distribution, and ranged from 75 to 300 bases, indicating incomplete S1 digestion of RNA near the target site (Figure 2b).

3.3. Enhanced Enrichment of More Homogeneous Target RNA Fragments Using Non-Target Probes

Given the complex secondary and tertiary structures of HIV-1 RNA structures [45,46], we reasoned that the enrichment of the unexpectedly lengthy and heterogeneous RNAs may result from poor S1 digestion near the target sites due to local secondary structures (Figure 2(ai)). S1 digestion is inefficient for double-stranded RNA [47]. To test this, we designed a set of DNA probes (for a total of 110 non-target probes, see Supplementary Table S1) that anneal to different sites of HIV-1 RNA in order to better control the structure of RNA substrates for S1 digestion. These non-target probes were designed to interfere with the formation of local secondary structures based on data from a previous SHAPE assay [46] and leave at least 40–50 bases of single-stranded areas between the probe binding sites (Figure 2a,b). The two non-target probes (NL9360R and NL9530R, see supplemental Table S1) nearest to the target site align approximately 57–58 bases away from the b-target oligo binding sites. These single-stranded regions are favorable targets for S1 digestion [47]. Indeed, when we used a mixture of both b-target oligos and a set of total 110 non-target probes (Supplementary Table S1) to release the local structure of HIV-1 RNA, the S1-digestion and streptavidin-bead purification process resulted in more homogeneous b-DNA/RNA duplex bands near 50-base marker DNA on gel electrophoresis, indicating a substantial improvement in our target enrichment process (Figure 2(bii)). The streptavidin-bead-purified b-DNA/RNA duplexes were then subjected to oligonucleotide LC-MS/MS.

3.4. Oligonucleotide LC-MS/MS of Synthetic RNA Oligos with m⁶As at A8975 or A8989

As a positive control experiment, we used two synthetic RNA oligos (Figure 3a; see the methods section for the sequences). The sequences of the two synthetic RNA oligos match to the 81 bases that cover positions 8945 to 9025 of HIV-1 RNA. These oligos have fully m⁶A-methylated nucleotides at either 8975A or 8989A. RNase T1 digestion of these oligos, followed by ultra-high-performance liquid chromatography (UHPLC) and mass-spectrometry (MS), revealed the presence of various RNase T1-cleaved oligos of expected sizes, including m⁶A-methylated and non-methylated forms of the two target sites (A8975: ACUUUCCGp, m⁶A8975: m⁶ACUUUCCGp, A8989: ACUUUCCAGp, and m⁶A8989: m⁶ACUUUCCAGp) (Figure 3a). Two cyclic phosphate forms were also detected for the A8975 site (ACUUUCCG>p and m⁶ACUUUCCG>p). As expected, only the m⁶A-methylated A8975 form (m⁶ACUUUCCGp as well as m⁶ACUUUCCG>p) and non-methylated A8989 (ACUUUCCAGp) oligo form were detected for the m⁶A8975 synthetic RNA oligos; on the other hand, for the m⁶A8989 synthetic RNA oligos, the corresponding oligo form for m⁶ACUUUCCAGp and non-methylated versions for A8975 (ACUUUCCGp and/or ACUUUCCG>p) were detected (Figure 3a).

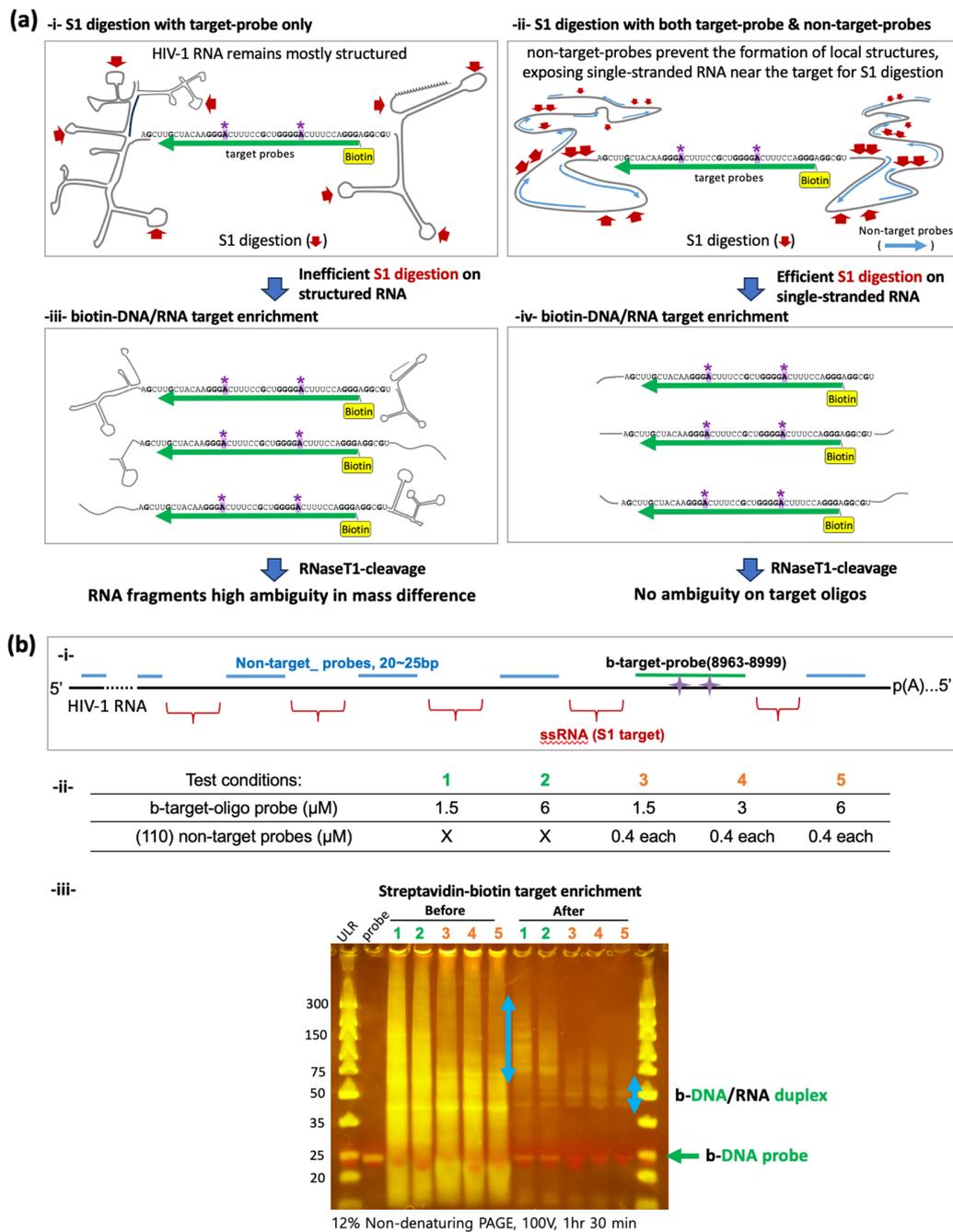


Figure 2. Comparison of S1 digestion with or without non-target probes. (a) A schematic view comparing S1 digestion on HIV-1 RNA. Complex RNA secondary structure formation may result in inefficient S1 digestion, leaving unanticipatedly long RNA strands in the enriched biotin-DNA/RNA pool (i,iii). We used 110 different non-target probes to prevent secondary structure formation, which may enhance S1 digestion (ii) and enrich more homogeneous target RNAs (iv). A8975 and A8989 sites are denoted by (*). (b) (i) a schematic view of S1-digestion on single-stranded RNA (ssRNA). (ii) Different concentrations of target (b-target-oligo) probes and non-target probes were tested. Conditions 1 and 2 were tested without non-target probes, while 3, 4 and 5 tested with 110 different non-target probes (0.4 μM each). (iii) Non-denaturing PAGE gel electrophoresis before and after the streptavidin–biotin target enrichment. Light blue arrow indicates isolated b-DNA/RNA duplexes. The green arrow indicates unbound DNA probes. In all conditions, streptavidin–biotin enrichment step significantly reduced non-target RNA fragments (before vs. after). After the enrichment, conditions 3, 4, and 5 using non-target probes showed b-DNA/RNA duplex near the 50-base marker, whereas conditions 1 and 2 showed much larger fragments ranging from 75 to 300 bases.

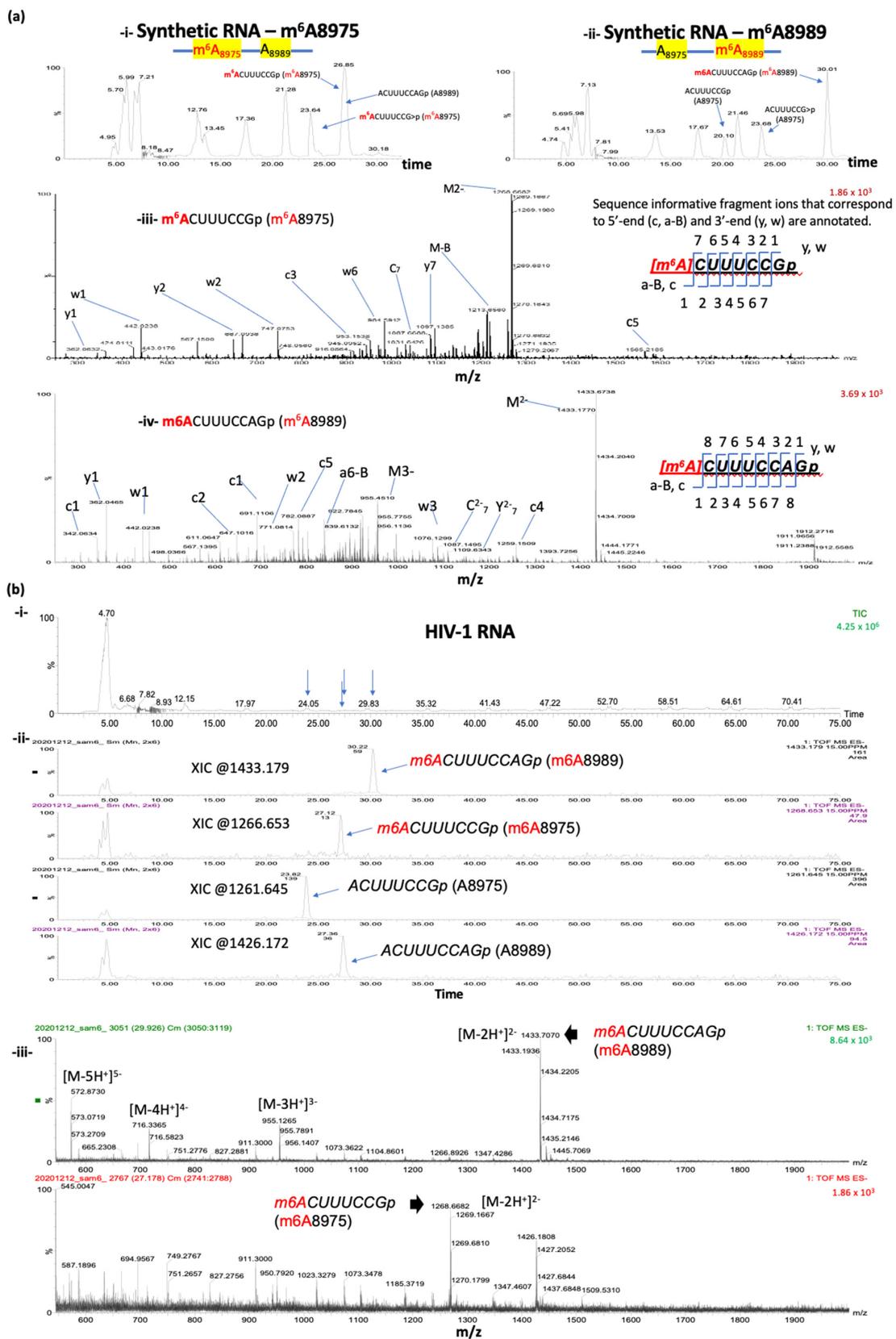


Figure 3. LC-MS/MS based comparison of the RNase T1 digests of target HIV-1 RNA segment with the methylated synthetic oligos of same sequence. (a) Detection of RNase T1 digestion products of synthetic oligonucleotides bearing single methylation at either 8975 (i) or 8989 (ii) position in the total ion chromatogram. Mass spectrum of RNase T1 digestion product, m⁶ACUUUCCGp (*m/z* 1268.653, mass error < 15 PPM) at position 8975 from synthetic RNA (i,iii) and m⁶ACUUUCCAGp (1.81 1433.179, TIC 4.25 × 10⁶). (b) XICs and MS/MS spectra for the identified peaks.

mass error < 15 PPM) at position 8989 (ii,iv). Sequence informative fragment ions that correspond to 5'-end (a-B_n, c_n) and 3'-end (y_n, w_n) for each of the digestion product are annotated. (b) LC-MS analysis of the oligonucleotide segment purified from HIV-1 RNA (i) Total ion chromatogram of the RNase T1 digest. Retention time positions of methylated and unmethylated oligonucleotides are indicated by arrows. (ii) Extracted ion chromatograms for methylated and unmethylated versions of oligomers corresponding to positions at 8975 and 8989. (iii) Mass spectra of the methylated oligonucleotides depicting the multiply charged ions. Note the differences in relative abundance of methylated oligo m⁶ACUUUCCAGp (position 8989, top panel) vs. m⁶ACUUUCCGp (position 8975, bottom panel). Oligonucleotide LC-MS/MS confirmed m⁶A at 8989A and 8975A in HIV-1 viral RNA (iii).

3.5. Oligonucleotide LC-MS/MS Confirms the Presence of m⁶As at A8975 and A8989 of HIV-1 RNA

For virion RNA analyses, we isolated approximately 200 ng of target HIV-1 RNA fragments, harboring both A8975 and A8989 sites, following the S1-digestion and streptavidin-enrichment. These RNA were subjected to subsequent RNase T1 digestion and oligonucleotide analysis via LC-MS/MS (Figures 1 and 2). The analysis identified all the expected methylated and non-methylated versions at the two m⁶A-sites (Figure 3b and Table 1). The detection accuracy was ascertained by the retention time and the mass spectral (*m/z* values) matches between the HIV-1 RNA samples and synthetic RNA segments. These data therefore validate the presence of m⁶A at positions A8975 and A8989 in HIV-1 RNA. These m⁶As were further confirmed by the reduction in modification signals following m⁶A-eraser (ALKBH5) treatment and through site-directed mutagenesis in an independent study [40].

Table 1. The m⁶A modification frequencies at 8975 and 8989 sites of HIV-1 RNA.

m ⁶ A Position	Oligomer	Retention Time	Peak Area Counts	Total Area	Relative Ratio of Modified vs. Total
8975	ACUUUCCGp	23.82	434	500	69/569 = 0.1213
	ACUUUCCG>p	20.12	66		
	m ⁶ ACUUUCCGp	27.12	49	69	
	m ⁶ ACUUUCCG>p	23.91	20		
8989	ACUUUCCAGp	27.36	102	102	174/276 = 0.6304
	ACUUUCCAG>p	0	0		
	m ⁶ ACUUUCCAGp	30.22	174	174	
	m ⁶ ACUUUCCAG>p	0	0		

While LC-MS/MS had the potential to detect any RNA modifications if they were deposited in a site-specific fashion above the detection levels, we did not detect any other modifications in our experiments.

3.6. Detection of Both Methylated and Unmethylated Versions of the Same Oligomer Indicates Partial Methylation

Current analysis of the LC-MS/MS data indicates differential modification stoichiometry based on a comparison of the relative abundance of unmodified and modified oligonucleotides (Table 1). The computation of the peak area ratios following baseline noise subtraction estimates that approximately 12% of HIV-1 RNA is methylated at 8975A, and ~63% is methylated at 8989A. More systematic studies to optimize the quantification via LC-MS/MS-based analyses and additional orthogonal assays are needed, however, if we are to accurately determine the stoichiometry of methylation.

4. Discussion

Developing methods that enable the analysis of the “true” sequences of native RNA molecules has been a major interest [8]. Liquid chromatography coupled with mass spectrometry (LC-MS) is a powerful tool for modification analysis due to its ability to directly measure the mass values of modified and unmodified RNA molecules. However, its application to long RNAs and weakly abundant transcripts has been challenging. Enriching small, highly homogeneous RNA fragments of target sites can be a way forward but has been known to be a significant challenge for HIV-1 RNA. In this study, we present a novel method that significantly improves target RNA enrichment by effectively fragmenting lengthy HIV-1 RNA using the S1 enzyme and multiple non-target probes. The principles of our method have the potential for broad application, enabling the analysis of lengthy, kilobase-scale RNAs that were previously considered impractical for investigation by oligonucleotide LC-MS analysis.

Our target enrichment strategy notably improved the enrichment and purification of small target fragments of HIV-1 RNA containing both A8975 and A8989 sites. However, this process requires a substantial amount (~10 µg) of 9 Kb full-length virion RNAs to obtain an adequate amount of target RNA for LC-MS analysis. The current sensitivity achieved with this RNA (200 ng of target RNA) was adequate for detecting the presence of m⁶As at these two sites. Nevertheless, accurate quantification of m⁶A modifications would necessitate a significantly larger quantity of viral RNA and a well-designed quantification experiment. Method precision can be further improved by using another ribonuclease enzyme with complementary specificity so that overlapping digestion product is generated to confirm the modification location.

One limitation of this approach is its inability to distinguish nitrogenous base methylation from ribose methylation through oligonucleotide sequencing. This limitation may be addressed by incorporating nucleoside analysis, where base methylation and ribose methylation exhibit different retention times during chromatography. The point of methylation can then be understood through MS/MS analysis of nucleosides.

While this method is effective in analyzing one target section, it has limitations as it retains only the specified target, discarding the remaining sections during the process. Looking ahead, developing a multiplex technique capable of isolating multiple diverse targets from the same RNA sample would be valuable. Such an advancement could broaden the applicability of the approach, enabling the mapping of other regions of HIV-1 RNA or small segments within any long RNA molecules with known sequences.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/mps7010007/s1>, Table S1: non-target probes for LC-MS sample preparation.

Author Contributions: Conceptualization, S.K. (Sanggu Kim) and B.A.; methodology, A.B., A.R., G.-E.L., S.G., H.Y. and S.K. (Shihyoung Kim); software, A.B., G.-E.L. and A.R.; validation, A.B. and A.R.; formal analysis, A.B., A.R., B.A. and S.K. (Sanggu Kim); investigation, A.B., A.R., B.A. and S.K. (Sanggu Kim); resources, P.A.L., B.A. and S.K. (Sanggu Kim); data curation, A.R. and B.A.; writing—original draft preparation, A.B., A.R., B.A. and S.K. (Sanggu Kim); writing—review and editing, A.B., A.R., P.A.L., B.A. and S.K. (Sanggu Kim); visualization, A.B., A.R., B.A. and S.K. (Sanggu Kim); supervision, B.A. and S.K. (Sanggu Kim); project administration, P.A.L., B.A. and S.K. (Sanggu Kim); funding acquisition, P.A.L. and S.K. (Sanggu Kim). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Institutes of Health for this work, grant numbers HG010318 [Kim] and GM058843 [Limbach], and Department of Defense, grant number HT9425-23-1-0582. Ga-Eun Lee is a recipient of the graduate fellowship from the C. Glenn Barber Fund Trust.

Institutional Review Board Statement: This work was approved by the OSU Institutional Biosafety Committee protocols (2016R00000009-R1-AM1 and 2020R00000030). This study does not involve animals or humans.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the current study are available from the corresponding authors upon request.

Acknowledgments: We thank the Ohio Supercomputer Center (OSC) for their service; We thank all the research funding agencies.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Murakami, S.; Jaffrey, S.R. Hidden codes in mRNA: Control of gene expression by m⁶A. *Mol. Cell* **2022**, *82*, 2236–2251. [[CrossRef](#)]
2. Jiang, X.; Liu, B.; Nie, Z.; Duan, L.; Xiong, Q.; Jin, Z.; Yang, C.; Chen, Y. The role of m⁶A modification in the biological functions and diseases. *Signal Transduct. Target. Ther.* **2021**, *6*, 74. [[CrossRef](#)]
3. Boccaletto, P.; Stefaniak, F.; Ray, A.; Cappannini, A.; Mukherjee, S.; Purta, E.; Kurkowska, M.; Shirvanizadeh, N.; Destefanis, E.; Groza, P.; et al. MODOMICS: A database of RNA modification pathways. 2021 update. *Nucleic Acids Res.* **2021**, *50*, D231–D235. [[CrossRef](#)] [[PubMed](#)]
4. Abebe, J.S.; Verstraten, R.; Depledge, D.P. Nanopore-Based Detection of Viral RNA Modifications. *mBio* **2022**, *13*, e0370221. [[CrossRef](#)] [[PubMed](#)]
5. Davis, F.F.; Allen, F.W. Ribonucleic acids from yeast which contain a fifth nucleotide. *J. Biol. Chem.* **1957**, *227*, 907–915. [[CrossRef](#)] [[PubMed](#)]
6. Zhang, Y.; Lu, L.; Li, X. Detection technologies for RNA modifications. *Exp. Mol. Med.* **2022**, *54*, 1601–1616. [[CrossRef](#)]
7. Helm, M.; Motorin, Y. Detecting RNA modifications in the epitranscriptome: Predict and validate. *Nat. Rev. Genet.* **2017**, *18*, 275–291. [[CrossRef](#)]
8. Alfonzo, J.D.; Brown, J.A.; Byers, P.H.; Cheung, V.G.; Maraia, R.J.; Ross, R.L. A call for direct sequencing of full-length RNAs to identify all modifications. *Nat. Genet.* **2021**, *53*, 1113–1116. [[CrossRef](#)]
9. Baquero-Perez, B.; Geers, D.; Diez, J. From A to m⁶A: The Emerging Viral Epitranscriptome. *Viruses* **2021**, *13*, 1049. [[CrossRef](#)]
10. Courtney, D.G.; Tsai, K.; Bogerd, H.P.; Kennedy, E.M.; Law, B.A.; Emery, A.; Swanstrom, R.; Holley, C.L.; Cullen, B.R. Epitranscriptomic Addition of m(5)C to HIV-1 Transcripts Regulates Viral Gene Expression. *Cell Host Microbe* **2019**, *26*, 217–227.e6. [[CrossRef](#)]
11. Tsai, K.; Vasudevan, A.A.J.; Campos, C.M.; Emery, A.; Swanstrom, R.; Cullen, B.R. Acetylation of Cytidine Residues Boosts HIV-1 Gene Expression by Increasing Viral RNA Stability. *Cell Host Microbe* **2020**, *28*, 306–312.e6. [[CrossRef](#)] [[PubMed](#)]
12. Kennedy, E.M.; Bogerd, H.P.; Kornepati, A.V.R.; Kang, D.; Ghoshal, D.; Marshall, J.B.; Poling, B.C.; Tsai, K.; Gokhale, N.S.; Horner, S.M.; et al. Posttranscriptional m⁶A Editing of HIV-1 mRNAs Enhances Viral Gene Expression. *Cell Host Microbe* **2016**, *19*, 675–685. [[CrossRef](#)] [[PubMed](#)]
13. Ringear, M.; Marchand, V.; Decroly, E.; Motorin, Y.; Bennasser, Y. FTSJ3 is an RNA 2'-O-methyltransferase recruited by HIV to avoid innate immune sensing. *Nature* **2019**, *565*, 500–504. [[CrossRef](#)] [[PubMed](#)]
14. Lichinchi, G.; Gao, S.; Saletore, Y.; Gonzalez, G.M.; Bansal, V.; Wang, Y.; Mason, C.E.; Rana, T.M. Dynamics of the human and viral m⁶A RNA methylomes during HIV-1 infection of T cells. *Nat. Microbiol.* **2016**, *1*, 16011. [[CrossRef](#)] [[PubMed](#)]
15. Tirumuru, N.; Zhao, B.S.; Lu, W.; Lu, Z.; He, C.; Wu, L. N(6)-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 Gag protein expression. *eLife* **2016**, *5*, e15528. [[CrossRef](#)] [[PubMed](#)]
16. Shi, H.; Wei, J.; He, C. Where, When, and How: Context-Dependent Functions of RNA Methylation Writers, Readers, and Erasers. *Mol. Cell* **2019**, *74*, 640–650. [[CrossRef](#)] [[PubMed](#)]
17. Liu, N.; Dai, Q.; Zheng, G.; He, C.; Parisien, M.; Pan, T. N6-methyladenosine-dependent RNA structural switches regulate RNA–protein interactions. *Nature* **2015**, *518*, 560–564. [[CrossRef](#)]
18. König, J.; Zarnack, K.; Rot, G.; Curk, T.; Kayikci, M.; Zupan, B.; Turner, D.J.; Luscombe, N.M.; Ule, J. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.* **2010**, *17*, 909–915. [[CrossRef](#)]
19. Llibre, J.M.; Pulido, F.; Garcia, F.; Deltoro, M.G.; Blanco, J.L.; Delgado, R. Genetic barrier to resistance for dolutegravir. *AIDS Rev.* **2015**, *17*, 56–64.
20. Sun, L.; Fazal, F.M.; Li, P.; Broughton, J.P.; Lee, B.; Tang, L.; Huang, W.; Kool, E.T.; Chang, H.Y.; Zhang, Q.C. RNA structure maps across mammalian cellular compartments. *Nat. Struct. Mol. Biol.* **2019**, *26*, 322–330. [[CrossRef](#)]
21. Wu, B.; Su, S.; Patil, D.P.; Liu, H.; Gan, J.; Jaffrey, S.R. Molecular basis for the specific and multivalent recognitions of RNA substrates by human hnRNP A2/B1. *Nat. Commun.* **2018**, *9*, 420. [[CrossRef](#)] [[PubMed](#)]
22. Park, O.H.; Ha, H.; Lee, Y.; Boo, S.H.; Kwon, D.H.; Song, H.K.; Kim, Y.K. Endoribonucleolytic Cleavage of m⁶A-Containing RNAs by RNase P/MRP Complex. *Mol. Cell* **2019**, *74*, 494–507.e8. [[CrossRef](#)] [[PubMed](#)]
23. Gaston, K.W.; Limbach, P.A. The identification and characterization of non-coding and coding RNAs and their modified nucleosides by mass spectrometry. *RNA Biol.* **2014**, *11*, 1568–1585. [[CrossRef](#)] [[PubMed](#)]
24. Thakur, P.; Estevez, M.; Lobue, P.A.; Limbach, P.A.; Addepalli, B. Improved RNA modification mapping of cellular non-coding RNAs using C- and U-specific RNases. *Analyst* **2020**, *145*, 816–827. [[CrossRef](#)] [[PubMed](#)]

25. Jora, M.; Borland, K.; Abernathy, S.; Zhao, R.; Kelley, M.; Kellner, S.; Addepalli, B.; Limbach, P.A. Chemical Amination/Imination of Carbonothiolated Nucleosides during RNA Hydrolysis. *Angew. Chem. Int. Ed.* **2020**, *60*, 3961–3966. [CrossRef] [PubMed]
26. Addepalli, B.; Lesner, N.P.; Limbach, P.A. Detection of RNA nucleoside modifications with the uridine-specific ribonuclease MC1 from *Momordica charantia*. *RNA* **2015**, *21*, 1746–1756. [CrossRef] [PubMed]
27. McIntyre, W.; Netzband, R.; Bonenfant, G.; Biegel, J.M.; Miller, C.; Fuchs, G.; Henderson, E.; Arra, M.; Canki, M.; Fabris, D.; et al. Positive-sense RNA viruses reveal the complexity and dynamics of the cellular and viral epitranscriptomes during infection. *Nucleic Acids Res.* **2018**, *46*, 5776–5791. [CrossRef]
28. Pereira-Montecinos, C.; Toro-Ascuy, D.; Ananías-Sáez, C.; Gaete-Argel, A.; Rojas-Fuentes, C.; Riquelme-Barríos, S.; Rojas-Araya, B.; García-De-Gracia, F.; Aguilera-Cortés, P.; Chnaiderman, J.; et al. Epitranscriptomic regulation of HIV-1 full-length RNA packaging. *Nucleic Acids Res.* **2022**, *50*, 2302–2318. [CrossRef]
29. Konan, S.N.; Ségéral, E.; Bejjani, F.; Bendoumou, M.; Said, M.A.; Gallois-Montbrun, S. YTHDC1 regulates distinct post-integration steps of HIV-1 replication and is important for viral infectivity. *Retrovirology* **2022**, *19*, 4. [CrossRef]
30. Tsai, K.; Bogerd, H.P.; Kennedy, E.M.; Emery, A.; Swanstrom, R.; Cullen, B.R. Epitranscriptomic addition of m⁶A regulates HIV-1 RNA stability and alternative splicing. *Minerva Anestesiol.* **2021**, *35*, 992–1004. [CrossRef]
31. Lu, W.; Tirumuru, N.; Gelais, C.S.; Koneru, P.C.; Liu, C.; Kvaratskhelia, M.; He, C.; Wu, L. N⁶-Methyladenosine-binding proteins suppress HIV-1 infectivity and viral production. *J. Biol. Chem.* **2018**, *293*, 12992–13005. [CrossRef] [PubMed]
32. D'ascenzo, L.; Popova, A.M.; Abernathy, S.; Sheng, K.; Limbach, P.A.; Williamson, J.R. Pytheas: A software package for the automated analysis of RNA sequences and modifications via tandem mass spectrometry. *Nat. Commun.* **2022**, *13*, 2424. [CrossRef] [PubMed]
33. Edmonds, C.G.; Crain, P.F.; Gupta, R.; Hashizume, T.; Hocart, C.H.; A Kowalak, J.; Pomerantz, S.C.; Stetter, K.O.; McCloskey, J.A. Posttranscriptional modification of tRNA in thermophilic archaea (Archaeobacteria). *J. Bacteriol.* **1991**, *173*, 3138–3148. [CrossRef] [PubMed]
34. 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*, 7869–7876. [CrossRef] [PubMed]
35. Bakin, A.; Kowalak, J.A.; McCloskey, J.A.; Ofengand, J. The single pseudouridine residue in *Escherichia coli* 16S RNA is located at position 516. *Nucleic Acids Res.* **1994**, *22*, 3681–3684. [CrossRef]
36. A Kowalak, J.; A McCloskey, J.; O Stetter, K.; Crain, P.F.; Bruenger, E.; Kowalak, J.A.; Kuchino, Y.; McCloskey, J.A.; Mizushima, H.; Stetter, K.O.; et al. 5S rRNA modification in the hyperthermophilic archaea *Sulfolobus solfataricus* and *Pyrodicticum occultum*. *FASEB J.* **1993**, *7*, 196–200. [CrossRef]
37. Kowalak, J.A.; Bruenger, E.; Hashizume, T.; Peltier, J.M.; Ofengand, J.; McCloskey, J.A. Structural characterization of U^{*}-1915 in domain IV from *Escherichia coli* 23S ribosomal RNA as 3-methylpseudouridine. *Nucleic Acids Res.* **1996**, *24*, 688–693. [CrossRef]
38. Kowalak, J.A.; Bruenger, E.; McCloskey, J.A. Posttranscriptional modification of the central loop of domain V in *Escherichia coli* 23 S Ribosomal RNA. *J. Biol. Chem.* **1995**, *270*, 17758–17764. [CrossRef]
39. Phillips, S.; Baek, A.; Kim, S.; Chen, S.; Wu, L. Protocol for the generation of HIV-1 genomic RNA with altered levels of N (6)-methyladenosine. *STAR Protoc.* **2022**, *3*, 101616. [CrossRef]
40. Alice Baek, G.-E.L.; Golconda, S.; Rayhan, A.; Manganaris, A.A.; Chen, S.; Tirumuru, N.; Yu, H.; Kim, S.; Kimmel, C.; Zablocki, O.; et al. Single-RNA-Level Analysis of Full-Length HIV-1 RNAs Reveals Functional Redundancy of m⁶As. Preprint (Version 1) Research Square. 2023. Available online: <https://assets.researchsquare.com/files/rs-2679540/v1/47af6f79-e742-4fdc-8ed6-77e46ea8cfd8.pdf?c=1686941532> (accessed on 15 December 2023).
41. Sambrook, J.; Russell, D.W. Purification of nucleic acids by extraction with phenol:chloroform. *Cold Spring Harb. Protoc.* **2006**, *2006*, pdb.prot4455. [CrossRef]
42. Tirumuru, N.; Wu, L. HIV-1 envelope proteins up-regulate N⁶-methyladenosine levels of cellular RNA independently of viral replication. *J. Biol. Chem.* **2019**, *294*, 3249–3260. [CrossRef] [PubMed]
43. Hendra, C.; Pratanwanich, P.N.; Wan, Y.K.; Goh, W.S.S.; Thiery, A.; Göke, J. Detection of m⁶A from direct RNA sequencing using a multiple instance learning framework. *Nat. Methods* **2022**, *19*, 1590–1598. [CrossRef] [PubMed]
44. Gao, Y.; Liu, X.; Wu, B.; Wang, H.; Xi, F.; Kohonen, M.V.; Reddy, A.S.N.; Gu, L. Quantitative profiling of N⁶-methyladenosine at single-base resolution in stem-differentiating xylem of *Populus trichocarpa* using Nanopore direct RNA sequencing. *Genome Biol.* **2021**, *22*, 22. [CrossRef] [PubMed]
45. Tomezsko, P.J.; Corbin, V.D.A.; Gupta, P.; Swaminathan, H.; Glasgow, M.; Persad, S.; Edwards, M.D.; McIntosh, L.; Papenfuss, A.T.; Emery, A.; et al. Determination of RNA structural diversity and its role in HIV-1 RNA splicing. *Nature* **2020**, *582*, 438–442. [CrossRef]
46. Watts, J.M.; Dang, K.K.; Gorelick, R.J.; Leonard, C.W.; Bess, J.W., Jr.; Swanstrom, R.; Burch, C.L.; Weeks, K.M. Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature* **2009**, *460*, 711–716. [CrossRef]
47. Desai, N.A.; Shankar, V. Single-strand-specific nucleases. *FEMS Microbiol. Rev.* **2003**, *26*, 457–491. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.