Waters™

Tunable Digestions of RNA Using RapiZyme™ RNases to Confirm Sequence and Map Modifications

Balasubrahmanyam Addepalli, Tatiana Johnston, Christian Reidy, Matthew A. Lauber

Waters Corporation

Abstract

RNA therapeutics such as sgRNA and mRNA are important modalities for Gene Therapy applications to treat human diseases. Rapid development of these RNA molecules and their deployment require fast tracked development of new analytical methodologies. This methodology aims at determining the identity, purity, and modification(s) of therapeutic RNA through oligo mapping of its sequence by liquid chromatography coupled with mass spectrometry (LC-MS). MS-based sequencing involves direct detection of RNA fragments and localization of the nucleoside modifications in the sequence. Currently available enzymatic tools to produce RNA fragments yield incomplete or partial coverage with a very high ambiguity in sequence interpretation. The availability of enzyme tools with complementary cleavage specificity can overcome this problem and fill sequence gaps while streamlining data interpretation and increasing confidence. This application note evaluates an optimized protocol for RapiZyme MC1 and RapiZyme Cusativin and explores each RNase's reproducibility and complementary specificity with a focus on increased sequence coverage during sgRNA oligo mapping. The new workflows include data processing software, including the MAP Sequence app within waters_connect™ and two supporting MicroApps. These provide a streamlined workflow to facilitate data analysis for routine RNA fragment identification, sequence coverage assessment ,and quantification.

Benefits

- Complete sequence coverage of sgRNA generating more unique masses while maintaining LC-MS compatibility and high signal

- Simple and reproducible protocol that is adaptable for automation without chemical denaturants or enzyme inhibitors

- The ability to perform partial digestions and generate overlapping digestion products

- Automated mRNA digest annotation based on accurate-mass matching as facilitated by *in silico* mRNA digestion calculations by waters_connect applications

- Fully integrated into waters_connect MAP Sequence for streamlined data analysis

Introduction

Background:

Success of mRNA vaccines against SARS-Cov-2 pandemic accelerated the development of RNA therapeutics to treat infectious diseases and provide gene therapies through genome editing involving CRISPR sgRNA. Parallel advancements in analytical technology can accelerate the turnaround of characterization and release testing work being performed in discovery, development ,and QC laboratories. Confirmation of identity, purity ,and modifications in therapeutic mRNA through oligo-mapping and sequencing via LC-MS has been the widely accepted strategy for direct molecular analysis. Similar to a peptide mapping involving bottom-up approaches, RNA is enzymatically digested to generate RNA fragments that can then be separated on a reversed phase (or hydrophilic interaction) column through liquid chromatography (LC). As digestion components are eluted from the column, they can subsequently be ionized and detected by mass spectrometry (MS).[1] To date, RNase T1 has been the most commonly used enzyme for oligo mapping LC-MS. This enzyme cleaves RNA at the 3' end of guanosine residues (G). As a result, RNase T1 often generates smaller digestion products (monomers, dimers, trimers, etc.), especially when it acts on G-rich regions. These products ambiguously match to more than one location in the sequence leading to data interpretation ambiguity. Further, identical small segments of an RNA molecule are not useful for computing the sequence coverage. This leads to coverage gaps and low confidence sequence information.

To address these limitations, Waters™ has developed a new set of recombinant endonucleases for RNA digestion that exhibit cleavage specificities that are complementary to RNase T1. Recombinant RapiZyme MC1 and RapiZyme Cusativin exhibit optimal activity recognizing and cleaving dinucleotide sites as demonstrated in Figure 1.[2] Furthermore, both RapiZyme RNases also have a propensity for producing partial digestions which produce unique, overlapping products amenable to LC-MS analysis and sequence determination. This application note demonstrates the use of RapiZyme MC1 and RapiZyme Cusativin in simple one-pot protocols to achieve reproducible digestions of a CRISPR single guide (sg)RNA. Use of these enzymes along with IonHance™ HFIP (hexafluoroisopropanaol) and RNA data analysis workflows involving micro apps such as mRNA Cleaver, MAP Sequence, Coverage Viewer and UNIFI processing workflows help enable faster LC-MS characterization of RNA therapeutics.[3] With their reproducibility and compliant ready architecture, these tools are suitable for use in both discovery and QC labs in support of regulatory filings and easy-to-deploy future options for release testing.
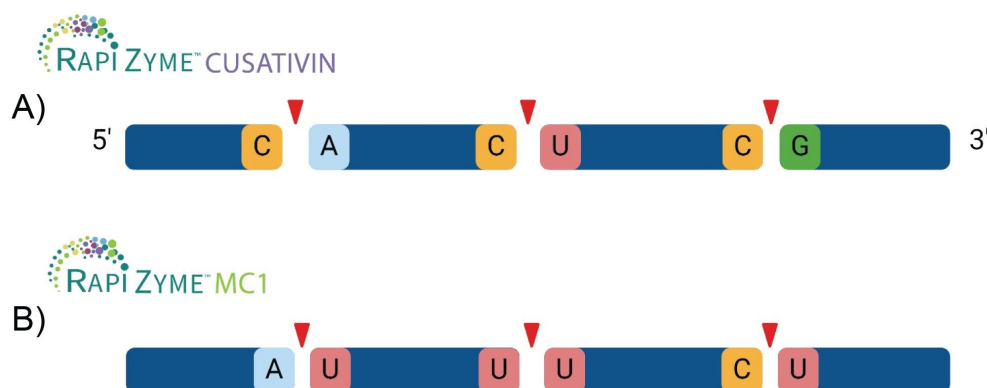

Figure 1. Example illustration showing the primary cleavage location of A) RapiZyme Cusativin Cp[A/U/G] and UpA and B) RapiZyme MC1 at [A/U/C]pU. A list of secondary cleavage sites can be found in the RapiZyme RNases Specificity section under Results and Discussion.

Experimental

Sequence
mG*mA*mU*GAUCUCUCAACUUUAACGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U

Elemental composition	Length (nt)	Theoretical MW (monoisotopic)	Theoretical MW (average)
C957H1187N374O690P99S6	100	32,212.16 Da	32,227.55 Da

An asterisk (*) denotes a PS bond and an="m" denotes a 2'-O-methyl (OMe) modification.

RNase free environment: Prior to RNA work, the workspace was made RNase-free by decontaminating the area with 70% isopropyl alcohol (IPA). All reagents used were RNase-free (certified) and prepared with RNase-free water.

Enzyme preparation: One vial (10,000 Units) of RapiZyme MC1 was dissolved in 200 μL of 200 mM ammonium acetate buffer at pH 8 creating a stock solution of 50 U/μL of RNase. One vial (10,000 Units) of RapiZyme Cusativin was dissolved in 200 μL of 200 mM ammonium acetate buffer at pH 9 creating a stock solution of 50 U/μL of RNase.

sgRNA Digestion with RapiZyme MC1

100 pmol (3.2 μg) of Hs.Cas9.HPRT1.2.AA sgRNA (IDT Technologies) was diluted to 10 μL with 200 mM ammonium acetate pH 8 and denatured at 90 °C for 2 minutes followed by rapid cooling at 4 °C. 1 μL of the RapiZyme RNase stock solution (50 Units total) was diluted to 10 μL of 18.2 MΩ water and added to each sample, gently vortexed and centrifuged to ensure mixing. The digestion mixture was incubated for 1 hour at 30 °C. The enzyme was heat inactivated by incubation at 70 °C for 15 minutes.

sgRNA digestion with RapiZyme Cusativin

100 pmol (3.2 μg) of Hs.Cas9.HPRT1.2.AA sgRNA (IDT Technologies) was diluted to 10 μL with 200 mM ammonium acetate pH 9 and denatured at 90 °C for 2 min followed by cooling at 4 °C. 1 μL of the RapiZyme Cusativin stock solution (50 Units total) was diluted to 10 μL of 18.2 MΩ water and added to each sample and gently vortexed and centrifuged to ensure mixing. The digestion mixture was incubated for 1 hour at 30 °C. The enzyme was heat inactivated by incubation at 75 °C for 15 minutes.

sgRNA digestion with RNase T1

100 pmol (3.2 µg) of Hs. Cas9.HPRT1.2.AA sgRNA (IDT Technologies) was diluted to 10 µL with 200 mM ammonium acetate pH 7 and denatured at 90 °C for 2 minutes followed by cooling at 4 °C. 250 units RNase T1 (Worthington Biochemical Corporation) diluted with 10 µL of 18.2 MΩ water was added, gently vortexed and centrifuged to ensure mixing. The digestion mixture was incubated for 1 hour at 37 °C. The enzyme was heat inactivated by incubation at 80 °C for 15 minutes.

sgRNA digestion with RNase 4

10 µg of Hs. Cas9.HPRT1.2.AA sgRNA (IDT Technologies) in 6.25 µL was mixed with 3.75 µL of 8 M urea to achieve a final concentration of 3M urea as recommended by NEB. The RNA was denatured at 90 °C for 5 minutes followed by cooling at 25 °C. The denatured RNA mixture was diluted into 20 µL of 1.5 X NEBuffer™ r1.1 for a final 30 µl volume as recommended by NEB and 1 µL of RNase 4 (NEB #M1284) stock (50U/µL) was added. The mixture was incubated at 37 °C for 1 hour. The reaction was stopped by adding 1 µL of NEB RNase inhibitor (NEB #M0314) followed by incubation at room temperature for 10 minutes.

Each digestion mixture was subsequently transferred to low bind QuanRecovery™ vials (p/n: 186009186 < https://www.waters.com/nextgen/global/shop/vials-containers--collection-plates/186009186-quanrecovery-with-maxpeak-hps-12-x-32-mm-screw-neck-vial-300--l-.html>) and an equivalent amount of RNA digest (50 pmol or 1.6 µg equivalent) for each enzyme was injected for LC-MS analysis.

LC Conditions

Column:	ACQUITY™ Premier Oligonucleotide BEH™ C$_{18}$, 300 Å, 1.7 µm, 2.1 x 50 mm (p/n: 186010539)
Mobile phase A:	0.1% N,N-diisopropylethylamine (DIPEA) and 1% (v/v) IonHanceHFIP (p/n:186010781) in 18.2 MΩ water
Mobile phase B:	0.0375% DIPEA and 0.075% (v/v) IonHance HFIP (p/n:186010781) in 55:10:35

	Acetonitrile:Methanol: 18.2 MΩ
LC system:	ACQUITY Premier BSM System
Detector:	Xevo™ G3 QTof
Wavelength:	254 nm
Flow rate:	0.4 mL/min
Injection:	10 μL
Column temperature:	70 °C
Sample temperature:	4 °C

Gradient Table

Time	%A	%B	Curve
0.0	99.0	1.0	Initial
2.0	99.0	1.0	6
36.0	80.0	20.0	6
41.0	30.0	70.0	6
42.0	99.0	1.0	6
52.0	99.0	1.0	6
53.0	99.0	1.0	11

MS Conditions

Source type:	ESI
Polarity:	Negative

Analyzer mode:	Sensitivity
Capillary voltage:	1.5 kV
Sample cone voltage:	40V
Source temperature:	100 °C
Desolvation temperature:	550 °C
Cone gas:	50 L/h
Desolvation gas:	650 L/h
Scan mass range:	550–2000
Scan time:	1.0 sec
Collision energy:	Low 6 V
High energy ramp:	10–45 V
Intelligent dat capture - intensity threshold:	Medium (10)
Lock correction:	Automatic (30 sec intervals)
Events: LC, sample:	2.0–40.0 min
Lock mass: leucine enkephalin:	554.26202 $(M-H^+)^{1-}$
Lock mass: combine width and mass window:	3 scans, 0.5 m/z

Results and Discussion

The purpose of this study was to evaluate the batch-to-batch performance of the RapiZyme RNases for RNA digestion and their utility for oligo mapping of therapeutic RNAs. In this study, we investigated the analysis of a CRISPR sgRNA designed for targeting the human HPRT (Hypoxanthine-guanine phosphoribosyltransferase) gene (Hs.Cas9.HPRT1.2.AA).

RapiZyme RNases Specificity

RapiZyme MC1 and Cusativin exhibit a preference for substrate docking and cleavage at specific dinucleotide locations. These preferences are outlined in Table 1. RapiZyme MC1 cleaves RNA at the 5′ end of uridine at ApU, CpU, and UpU sites with a slightly lower preference for CpA and CpG dinucleotide combinations. Incidentally, MC1 does not cleave GpU bonds. RapiZyme Cusativin exhibits slightly broader specificity cleaving RNA at the 3′ end of cytidine at CpA, CpG, CpU sites and with a slightly lower preference at ApU, GpU, UpU, and UpA bonds. Further, Cusativin exhibits little or no cleavage at CpC bonds. Despite having slightly broader specificity, these enzymes can be used to reproducibly achieve partial digestions, and they do not require protective hybridization oligos to achieve such results. This capability provides the opportunity for generating digestion products with overlapping features to increase the sequence coverage and facilitate confident data interpretation. Partial digestion through missed cleavages is a unique feature of both RapiZyme MC1 and Cusativin that creates more diverse digestion components and thus more chances to investigate sequence components and structural details of an RNA analyte. The mass values generated from these longer digestion products reduce the frequency of observed isomers that require more rigorous MS/MS analysis and sequence interpretation.

Informatics for data interpretation: The Waters MicroApp *mRNA Cleaver* predicts the *in silico* mass values of the expected digestion products for common digestion enzymes, including RapiZyme RNases, based on their cleavage specificity. This file is used to match and identify the observed digestion products in the LC-MS analysis through a waters_connect app *MAP Sequence*. The results are used to calculate and visualize the sequence coverage with *Coverage Viewer* for unique or unique and non-unique digestion products. These capabilities are highlighted using HPRT sgRNA as a substrate RNA analyte in the current studies.

Cleavage heat map								
RNase	RapiZyme MC1				RapiZyme Cusativin			
Nucleotide site	A	G	C	U	A	G	C	U
A				>70%				30–70%
G								30–70%
C	30–70%			>70%	>70%	>70%		>70%
U				30–70%	>70%			30–70%

Cleavage under tested conditions		
Not observed	30–70%	>70%

Table 1. Approximation of relative cleavage reaction rates for RapiZyme MC1 and RapiZyme Cusativin at each dinucleotide combination.

RapiZyme RNases Batch to Batch Reproducibility

In oligonucleotide mapping sample preparations, reproducible digestion of the RNA substrate is paramount. Reproducible digestion behavior of RapiZyme MC1 and RapiZyme Cusativin using HPRT sgRNA was evaluated across multiple batches in triplicate. Figure 2 shows LC-UV-MS analysis of the resulting enzyme digests. Both enzymes exhibited excellent reproducibility across UV profiles with minimum differences in peak height, although RapiZyme MC1 showed minor differences in abundance of a subset of digestion products (Figure 2A and 2B). Furthermore, MS signal was evaluated by use of the MAP Sequence app to identify the digestion products and their retention times based on mRNA Cleaver input. *In silico* predicted digestion products were generated with 4 missed cleavages using mRNA Cleaver. The digestion product file so created was used to identify the digestion products in the LC-MS data using the MAP Sequence application. After manual verification of any alternative assignments in the MAP Sequence output, a UNIFI template excel file was prepared denoting the oligonucleotide digestion product, elemental composition, and retention time. To demonstrate and visualize the reproducibility, this file was imported as library components into the waters_connect administration tool so that a UNIFI processing method for data analysis could be applied to all injections. From this MS analysis, extracted ion chromatogram overlays were prepared in the review mode. These are shown in Figure 2C for RapiZyme Cusativin and Figure 2D for RapiZyme MC1.

LC Conditions
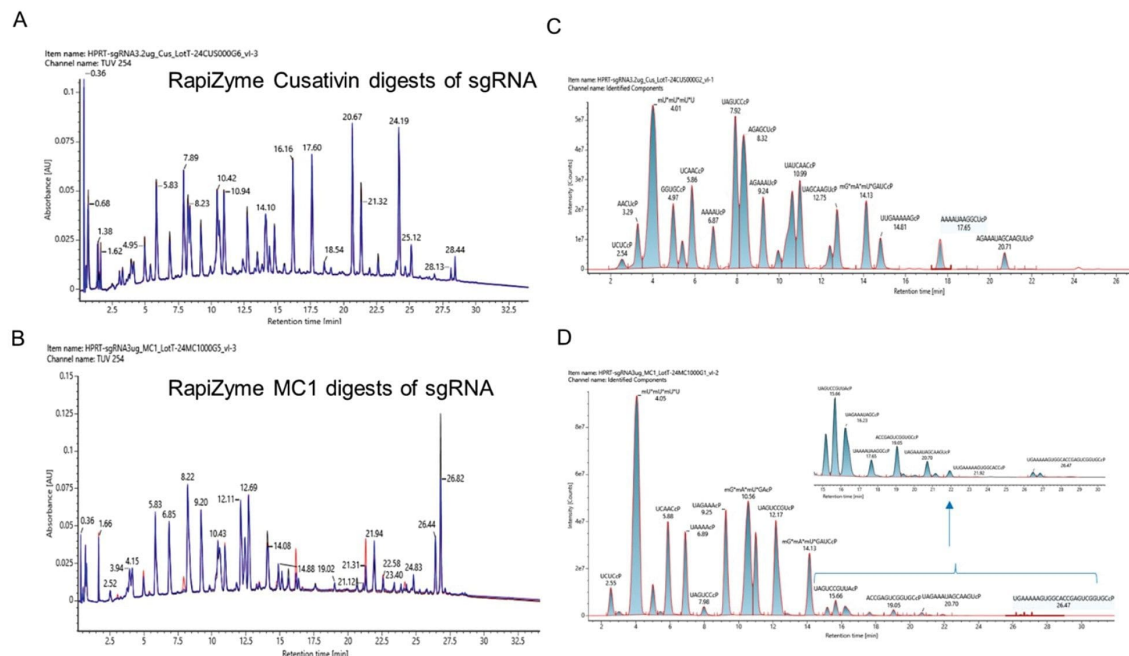
Figure 2. LC-UV-MS analysis of the HPRT sgRNA digested with recombinant RapiZyme Cusativin and MC1. Three independent recombinant enzyme batches of RapiZyme Cusativin (A) and RapiZyme MC1 (B) were used to digest HPRT sgRNA and the three UV chromatograms (black, blue and red traces) are overlayed on top of each other to indicate the reproducible digestion behavior of three preparations. (C) Extracted ion chromatogram overlay of the identified RapiZyme Cusativin digestion products indicating their abundance in the representative TUV trace. (D) Extracted ion chromatogram overlay of the identified RapiZyme MC1 digestion products indicating their abundance in the representative TUV trace. Low abundant digestion product profile of RapiZyme MC1 digestion products are shown in the inset figure.

To compare the digestion product profiles of RapiZyme Cusativin and MC1 to RNase T1 and RNase 4, sgRNA was digested with all four RNases as described in the methods section. The total ion chromatograms (TICs) of the four RNases are shown in Figure 3. A rich diversity of digestion products with abundant MS signal can be easily observed with RapiZyme MC1 and Cusativin digests compared to RNase T1 and RNase 4. Furthermore, the signal intensity was weaker for RNase 4 digest even though a similar mass load (50 pmol equivalent digest) was injected for analysis.

LC-MS data analysis of the RNase T1 and RNase 4 digests revealed the presence of discrete set of digestion products with no overlapping features even though missed cleavages were considered for these enzymes in the mRNA Cleaver Microapp based digestion product predictions. The observed digestion products for each enzyme are separately overlayed onto the HPRT sgRNA sequence as shown in Figure 4 for comparison purposes.
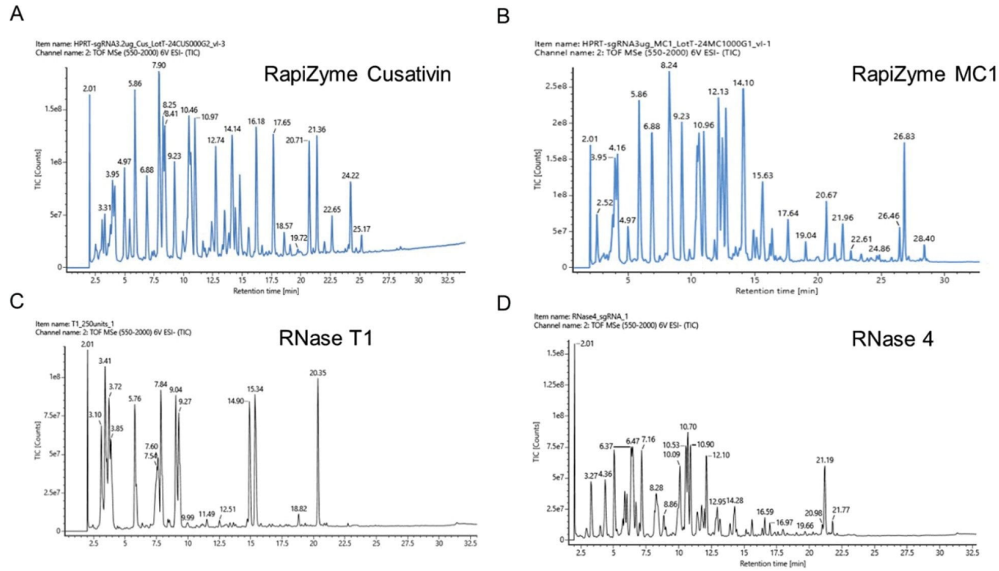
Figure 3. LC-MS analysis of HPRT1 sgRNA digests made with RapiZyme Cusativin (A), RapiZyme MC1 (B), RNase T1 (C) and RNase 4 (D).

Both RapiZyme MC1 and Cusativin digestion products exhibited overlaps which provide extra sequence location and context, thereby increasing the confidence in data interpretation. These sequence overlaps are possible because of the reproducible partial digestion of RNA by RapiZyme MC1 and Cusativin. Moreover, complete (100%) sequence coverage was achieved when observed digestion products were mapped to the sgRNA sequence for both RapiZyme MC1 and Cusativin. In contrast, RNase T1 yielded partial sequence coverage (~78%) with gaps in coverage and ambiguity where a single digestion product can be mapped to at least two regions of the sequence. Although RNase 4 digest provided complete sequence coverage, its digestion products are discrete with no sequence overlaps.

Regarding the terminal phosphate type on the digestion products, RapiZyme MC1 and Cusativin digests exhibited predominantly cyclic phosphates while the RNase 4 exhibited equivalent amounts of cyclic and linear phosphates, which helps explain the lower intensities of signal seen across the example RNase 4 data. On the

other hand, linear phosphate is the predominant form for RNase T1 digestion products (Table 2). The length of digestion products also varied with RapiZyme MC1 exhibiting the longest digestion product (27 nt) compared to other enzymes.
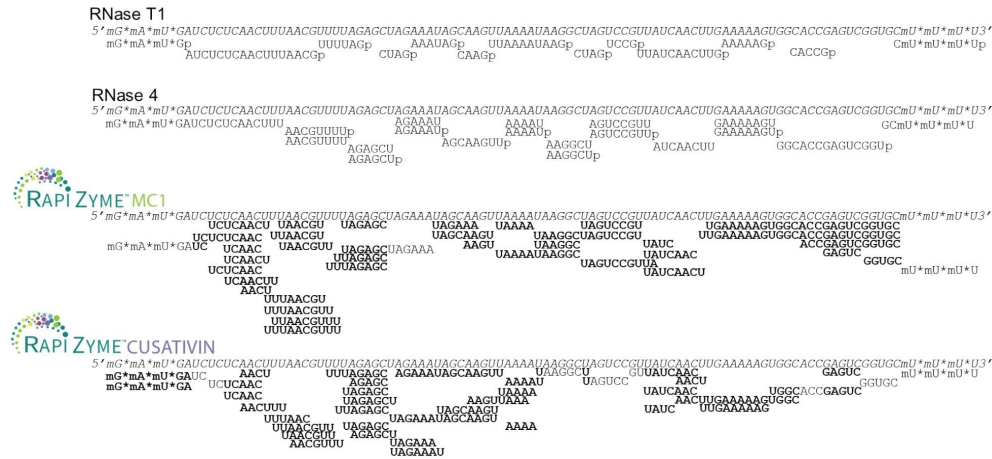

Figure 4. Depiction of identified HPRT sgRNA digestion products for RNase T1, RNase 4, RapiZyme MC1, and RapiZyme Cusativin. Bolded bases signify areas of the sequence where residues were detected in overlapping digestion products. Sequences that end in p (e.g. AACGUUUUp) signifies the presence of linear phosphate product detected in high abundance.

HPRT sgRNA results						
RNase	Total preparation time (min)	Number of digestion products	Sequence coverage	Sequence overlaps	Longest digestion product (nt)	3' Product formed
RapiZyme MC1	90	37	100%	Yes	27	Cyclic phosphate (>95%)
RapiZyme Cusativin	90	51	100%	Yes	16	Cyclic phosphate (>75%)
RNase T1	75	13	72%	No	17	Linear phosphate
RNase 4	90	12	100%	No	17	50% Cyclic 50% Linear

Table 2. HPRT sgRNA Results.

Conclusion

In this work, we have established a robust analytical workflow involving sample preparation and data analysis of synthetic RNA therapeutics using IP-RP-LC-MS.

· Synthetic sgRNA was reproducibly digested using RapiZyme MC1 and RapiZyme Cusativin with as little as 100 pmol or 3.2 μg RNA consumed in the sample preparation protocol and only 50 pmol (1.6 μg) of the RNA digest injected for analysis

· High chromatographic resolution as well as excellent UV and MS signal was achieved for both short and long digestion products using ion-pairing reversed phase chromatography performed with an ACQUITY Premier Binary System, ACQUITY Premier Oligonucleotide BEH C18 Column, and IonHance HFIP mobile phase reagent

· RNA digest fragments were identified and mapped to the sequence based on accurate-mass matching as facilitated by in silico RNA digestion calculations through mRNA Cleaver, MAP Sequence, and Coverage Viewer

· Complete coverage with sequence overlaps was obtained with both RapiZyme RNases to provide confident data interpretation and output

Momentum for newer and more advanced RNA therapeutics, including mRNA vaccines, CRISPR sgRNA, and synthetic oligonucleotides, continues to build. To help address the complexities related to characterization, new solutions for sample preparation, chromatographic separation, informatics for data analysis, and data interpretation are needed. The methodology outlined in this application note with RapiZyme MC1 and RapiZyme Cusativin demonstrates how RNA therapeutics such as sgRNA can be reproducibly characterized and monitored through the generation of unique digestion products and LC-MS analysis for confident and complete sequence coverage. This work demonstrates the generation of orthogonal and additive sequence mapping information for deployment of multiple nucleases with different cleavage preferences.

More information detailing the waters_connect MAP Sequence workflow for RNA characterization can be found in RNA Digestion Product Mapping Using an Integrated UPLC-MS and Informatics Workflow, Waters Application Note - 720008553.

References

1. Gau, B.; Dawdy, A.; Wang, H.L. *et al.* (2023) Oligonucleotide Mapping Via Mass Spectrometry to Enable Comprehensive Primary Structure Characterization of an mRNA Vaccine Against SARS-Cov-2. *Scientific Reports* 13, 9038.

2. Thakur P, Atway J, Limbach PA, Addepalli B. RNA Cleavage Properties of Nucleobase-Specific RNase MC1 and Cusativin are Determined by the Dinucleotide-Binding Interactions in the Enzyme-Active Site. *Int J Mol Sci.* 2022 Jun 24;23(13):7021.

3. Gaye MM, Fox J, Vissers JPC, Reah I, Knowles C, Lauber M (2022) Synthetic mRNA Oligo Mapping Using Ion Pairing Liquid Chromatography and Mass Spectrometry. Waters Application Note – 720007669. June 2022.

Featured Products

ACQUITY Premier System <
https://www.waters.com/nextgen/global/products/chromatography/chromatography-systems/acquity-premier-

system.html>

Xevo G3 QTof <https://www.waters.com/nextgen/global/products/mass-spectrometry/mass-spectrometry-systems/xevo-g3-qtof.html>

RapiZyme RNases <https://www.waters.com/nextgen/global/products/standards-and-reagents/rapizyme-rnases.html>

IonHance MS-Grade Mobile Phase Additives and Buffers < https://www.waters.com/nextgen/global/products/standards-and-reagents/ionhance-ms-grade-mobile-phase-additives-and-buffers.html>

waters_connect Software Solutions <https://www.waters.com/nextgen/global/products/informatics-and-software/waters_connect.html>

UNIFI Scientific Information System <https://www.waters.com/waters/nav.htm?cid=134801359>

720008539, September 2024

∧