Waters™

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

Catalin Doneanu, Alexandre F Gomes, Tatiana Johnston, Chris Preston, Matt Gorton, Bala Addepalli, Scott Berger, Ying Qing Yu

Waters Corporation, Milford, MA, United States

Abstract

Recent advances in liquid chromatography-mass spectrometry (LC-MS) and bioinformatics technologies have enhanced our detailed understanding of therapeutic RNA molecules. This application note presents a compliance-ready workflow for sequence mapping of digested single-guide RNAs (sgRNA). The workflow incorporates sample preparation using a variety of digestion enzymes, including RNase T1 as well as RapiZyme™ MC1 and RapiZyme Cusativin ribonucleases. UPLC™-MS^E data (DIA – data independent acquisition) was acquired on the Xevo MRT Mass Spectrometer and automated data analysis was performed with the waters_connect™ MAP Sequence Application (v2.0). The Xevo MRT Mass Spectrometer combines cutting-edge multi-reflecting time-of-flight technology with hybrid quadrupole time-of-flight (QTOF) capabilities, delivering sub-ppm mass accuracy that ensures confident identification of digested RNA products. In this study, complete (100%) sequence coverage of an sgRNA was achieved after combining the individual digestion results generated

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

1

for all three enzymes investigated. The unique assignments for the accurate mass measurements of the sgRNA digestion products were confirmed by their fragmentation using the ramped elevated energy MS^E fragmentation performed in the collision cell on the Xevo MRT Mass Spectrometer.

Benefits

· Confident characterization of sgRNA ribonuclease digestion products with sub-ppm average root mean square (RMS) mass accuracy for both digestion product precursors and their corresponding fragment ions

· A compliance-ready informatics workflow featuring the waters_connect MAP Sequence Application (v2.0) streamlines oligonucleotide sequence mapping of sgRNA digests

· RapiZyme MC1 and RapiZyme Cusativin, two recently introduced ribonucleases, offer unique cleavage specificity and the opportunity to generate overlapping digestion products, enabling complete (100%) and confident sequence coverage for sgRNAs

Introduction

sgRNAs were first described in 2012,[1] when two RNA molecules essential for CRISPR-Cas9-mediated DNA cleavage— the trans-activating CRISPR RNA (tracrRNA), which serves as a scaffold for the Cas9 nuclease and the CRISPR RNA (crRNA), responsible for DNA target recognition— were fused into a single RNA construct. This fusion created the sgRNA, a critical component of the CRISPR-Cas9 gene editing system.[1,2] The sgRNA molecule directs the Cas9 nuclease to introduce precise double-stranded breaks in DNA, enabling targeted genetic modifications.

Since its discovery, which was honored with the 2020 Nobel Prize in Chemistry, CRISPR technology has transformed gene editing applications. Beyond basic research, it has been adapted for rapid diagnostic tools, such as COVID-19 tests developed during the 2020 global pandemic,[3] and holds promise for therapeutic interventions in genetic diseases, cancer, and infectious diseases. As sgRNAs are typically synthesized via solid-phase oligonucleotide synthesis, their analytical characterization requires confirmation of both the intact molecular weight[4] and the sequence verification to ensure functional accuracy.[5]

Traditional LC-MS workflows for sgRNA digest oligo mapping have been labor-intensive and time-consuming, relying heavily on manual data analysis and interpretation. Recently,[6] a UPLC-MS^E (DIA) and informatics workflow

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

2

was presented that was designed for automated sgRNA sequence mapping following enzymatic digestion with a panel of ribonucleases, including RNase T1, hRNAse 4, and two RNase T2 enzymes: RapiZyme MC1 and Cusativin. This application note extends the capability of this workflow (Figure 1) when implemented on a Xevo MRT Mass Spectrometer (Figure 1), focusing on the multi-enzyme digestion of the Waters sgRNA LC-MS standard, where complete sequence coverage for the digested products and sequence-confirming fragmentation was obtained.
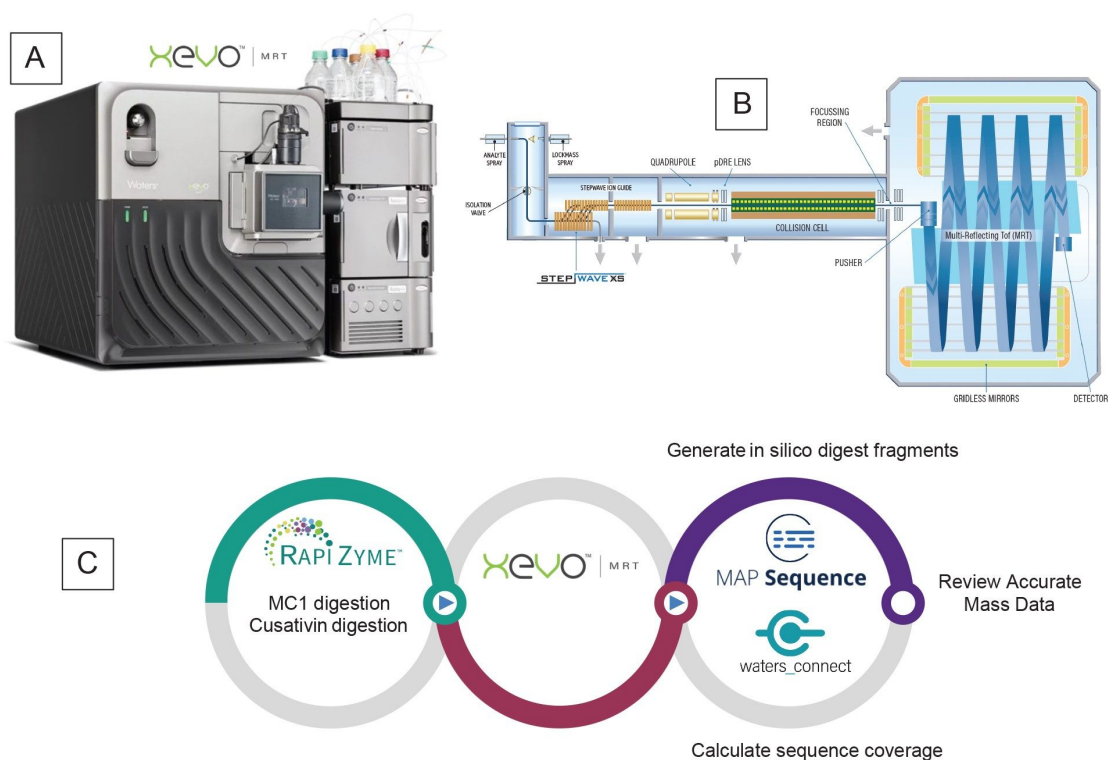

Figure 1. (A) Xevo MRT (Multi-Reflecting Time-of-flight) QTOF Mass Spectrometer with the ACQUITY Premier System; (B) Schematic diagram of the Xevo MRT QTOF Mass Spectrometer; (C) Workflow diagram for UPLC-MS^E processing of sgRNA digests using the MAP Sequence App v2.0.

Experimental

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

3

Reagents and Sample Preparation

N,N-diisopropylethylamine (DIPEA, 99.5% purity, catalog number 387649-100 ML) was purchased from MilliporeSigma (St Louis, MO) and 1,1,1,3,3,3-hexafluoro-2-propanol (IonHance™ HFIP, p/n:186010781) was obtained from Waters (Milford, MA). Methanol (LC-MS grade, catalog number 34966-1L) and acetonitrile (LC-MS grade, catalog number 34967-6XL) were obtained from Honeywell (Charlotte, NC). HPLC grade Type I deionized (DI) water was purified using a Milli-Q® system (Millipore, Bedford, MA). Mobile phases were prepared fresh daily. Ultrapure nuclease-free water (catalog number J71786.AE) for sgRNA digestions was purchased from Thermo Fisher Scientific (Waltham, MA).

The Waters sgRNA LC-MS Standard (p/n: 186011357) is a 100-mer sgRNA, encoding specifically for the mouse GATA2 transcription factor. Its sequence is: 5' - C*U*U* CAA CCA UCU CGA CUC GCG UUU UAG AGC UAG AAA UAG CAA GUU AAA AUA AGG CUA GUC CGU UAU CAA CUU GAA AAA GUG GCA CCG AGU CGG UGC U*U*U* U-3'. It contains a 2'-OMe modification on its first three 5' nucleotides (C*U*U*, denoted as the XYY sequence as shown in Figure 9), as well as on its last three 3' nucleotides (U*U*U*, sequence YYY in Figure 9). The asterisk indicates that all these six nucleotides are phosphorothioated. A 50 µM stock solution of the Waters sgRNA LC-MS Standard was prepared using nuclease-free water.

Fungus-derived animal free purified ribonuclease T1 (catalogue no IFGRNASET1AFLY500KU) was ordered from Innovative Research (Novi, MI) and the lyophilized enzyme was dissolved in 5 mL of 100 mM ammonium bicarbonate (catalogue no 5.33005-50G, MilliporeSigma) to prepare a solution containing 100 units/µL. For sgRNA digestion with RNase T1, 2 µL of 5 µM sgRNA were mixed with 28 µL of nuclease-free water and denatured at 90 ºC for 2 minutes. Next, 10 µL of RNase T1 enzyme (1000 units) were added and the digestion was allowed to proceed at 37 ºC for 15 minutes. The digestion mixture was prepared in a QuanRecovery™ MaxPeak™ 300 µL Vial. The digest was analyzed immediately by LC-MS using 5 µL injections.

The digestion protocols used for RapiZyme MC1 (Waters p/n: 186011190 < https://www.waters.com/nextgen/global/shop/standards--reagents/186011190-rapizyme-mc1-10000-units.html> , 10000 units/tube) and RapiZyme Cusativin (p/n: 186011192 < https://www.waters.com/nextgen/global/shop/standards--reagents/186011192-rapizyme-cusativin-10000-units.html> , 10000 units/tube) are very similar. For RapiZyme MC1, the sgRNA (10µL, 5 µM solution) was denatured at 90 ºC for 2 minutes in a buffer containing 200 mM ammonium acetate (catalog no AM9071-500ML, Thermo Fisher, Waltham, MA), pH 8.0. For RapiZyme Cusativin, the sgRNA (10µL, 5 µM solution) was denatured at 90 ºC for 2 minutes in a buffer containing 200 mM ammonium acetate, pH 9.0. Both samples were cooled on ice and spun in a microcentrifuge to collect the sample droplets. After adding 50 units of digestion enzyme (1 µL of

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

4

either RapiZyme MC1 or Cusativin) and 9 μL of nuclease-free water to obtain a final volume of ~ 20 μL, the sgRNA was digested at 30 °C for 30 minutes in an Eppendorf® thermomixer. The enzymatic digestion was stopped by exposure to 70 °C for 15 minutes for MC1 digesions, and to 75 °C (15 minutes) for Cusativin, in order to inactivate the enzymes. The digest was analyzed immediately by LC-MS using 5 μL injections.

All datasets were acquired with waters_connect Informatics Platform (v4.1.0.17) and subsequently processed using the SYNTHETIC Library App (v 2.0) and the MAP Sequence App (v2.0).

LC Conditions

LC-MS system:	Xevo MRT (multi-reflecting time-of-flight) QTOF Mass Spectrometer coupled with ACQUITY™ Premier UPLC (Binary) System
Column:	ACQUITY Premier Oligonucleotide BEH™ C18 Column 300 Å, 1.7 μm, 2.1 x 150 mm, (p/n: 186010541)
Column temperature:	70 °C
Flow rate:	400 μL/min
Mobile phases:	Solvent A: 0.1% DIPEA (N,N-diisopropylethylamine), 1% HFIP (1,1,1,3,3,3-hexafluoroisopropanol) in DI water, pH 8.5 Solvent B: 0.0375% DIPEA, 0.075% HFIP in 65% acetonitrile
Sample temperature:	8 °C
Sample vials:	QuanRecovery MaxPeak HPS Vials (p/n: 186009186)

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

5

| Injection volume: | 5 µL |

Wash solvents:	Purge solvent: 10% methanol in DI water
	Sample manager wash solvent: 50% MeOH
	Seal wash: 10% methanol in DI water

Gradient Table

Time (min)	Flow rate (mL/min)	Solvent A composition (%)	Solvent A composition (%)	Curve profile
0.00	0.4	97	3	Initial
5.00	0.4	97	3	6
60.00	0.4	70	30	6
61.00	0.4	15	85	6
62.00	0.4	15	85	6
63.00	0.4	97	3	6
70.00	0.4	97	0	6

MS Conditions

| MS system: | Xevo MRT QTOF Mass Spectrometer |

| Ionization mode: | ESI(-) |

| Acquisition mode: | MS^E |

| Acquisition rate: | 2 Hz |

| Capillary voltage: | 1.5 kV |

| Cone voltage: | 40 V |

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

6

Source offset:	10 V
Source temperature:	120 °C
Desolvation temperature:	550 °C
Cone gas flow:	0 L/h
Desolvation gas flow:	1000 L/hr
TOF mass range:	50–4000 (MS^E acquisition)
Low energy CE:	6 V
High energy CE ramp:	15 to 25 V
Lock-mass:	50 pg/µL Leu Enk in 0.1% formic acid, 50% ACN
Data acquisition:	waters_connect Informatics Platform v4.1.0.17
Data processing:	SYNTHETIC Library v2.0
	MAP Sequence v2.0

Results and Discussion

An informatics workflow (Figure 1) featuring the waters_connect MAP Sequence App (v 2.0) was used to facilitate automated data processing of UPLC-MS^E datasets acquired following the enzymatic digestion (in separate reactions) of an sgRNA standard with three different enzymes: RNase T1, RapiZyme MC1, and RapiZyme Cusativin.[6,7] The informatics processing workflow consisted of three steps which performed within the waters_connect MAP Sequence v 2.0 App:

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

7

1. *In-Silico* Digestion: The app generates predicted, *in-silico* digested oligonucleotide digestion products and sequence coverage, based on selecting the target sgRNA sequence from the SYNTHETIC Library App v 2.0 and the digestion enzyme of choice. Combined sequence coverage from multiple enzyme digests can also be displayed.

2. Data Processing and Mass-based Peak Assignments: The App processes the UPLC-MS^E data, matching the in-silico predicted neutral monoisotopic masses of digested oligonucleotides to the experimentally acquired MS1 data. In addition, the elevated energy MS^E data channel is used for assigning the corresponding fragment ions to each digestion product, therefore increasing the confidence for all digestion product assignments and enabling the resolution of ambiguous isomeric and isobaric nucleotides.

3. Sequence Coverage Visualization: The resulting sequence coverage for each sgRNA digest is summarized and visualized in the Results section. The combined sequence coverage obtained from multiple digest runs can also be displayed.

All data acquisition and processing were conducted within the waters_connect Informatics Platform, designed for users requiring software tools that maintain data integrity and support organizations subjected to regulatory scrutiny.

The data processing workflow begins with entering the sgRNA sequence into the SYNTHETIC Library App (v2.0), as illustrated in Figure 2. The Waters sgRNA standard includes two modified nucleotides: a 2'-OMe phosphorothioated cytidine, designated as nucleotide X, and a 2'-OMe phosphorothioated uridine, designated as nucleotide Y (see Figures 2 and 8 for the complete sgRNA sequence). Based on this sequence, the App automatically calculated the elemental composition as well as the monoisotopic and average masses of the 100-mer oligonucleotide. In addition, the SYNTHETIC Library App predicted sequence coverage for the same sgRNA digested with various enzymes. The predicted sequence coverages (Figure 3) for the Waters sgRNA Standard individually digested with three enzymes— RNase T1, RapiZyme MC1, and RapiZyme Cusativin— is displayed with the assumption of no missed cleavages. This predictive capability highlights how combining the individual sequence coverage results from multiple enzyme mapping analyses can potentially increase the overall sequence coverage of the RNA substrate.

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

8

Figure 2. SYNTHETIC Library entry for the Waters sgRNA Oligonucleotide Standard. This molecule contains two modified nucleotides, a 2'-OMe phosphorothioated cytidine defined as nucleotide X, and a 2'-OMe phosphorothioated uridine defined as nucleotide Y. The app automatically calculates the elemental composition and the monoisotopic mass of the 100-mer nucleotide.

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0
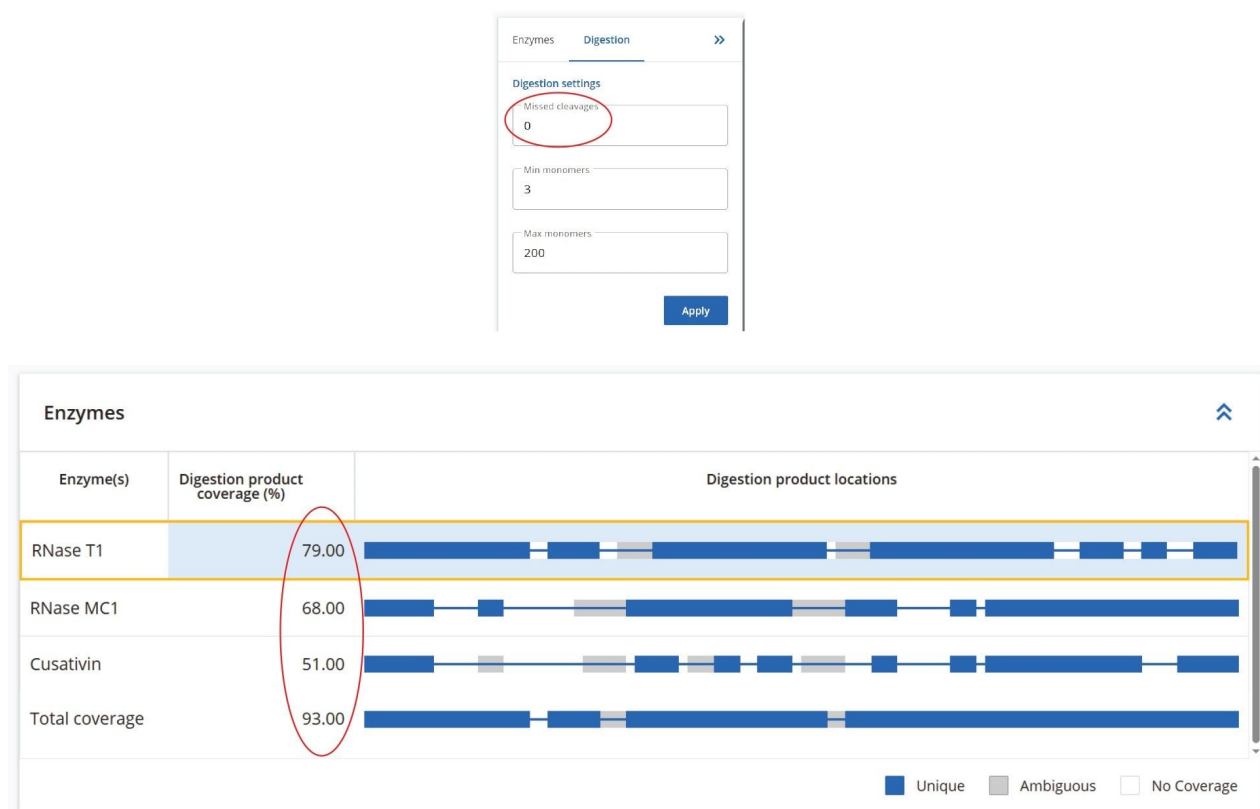
9

Figure 3. Predicted sequence coverages for digestion of the Waters sgRNA Standard, calculated using the SYNTHETIC Library App (v2.0). The figure compares the individual sequence coverages expected for three enzymes - RNase T1, RapiZyme MC1, and RapiZyme Cusativin - assuming no missed cleavages for each enzyme. The combined sequence coverage illustrates the effectiveness of a multi-enzyme digestion strategy in enhancing the overall sequence coverage of sgRNAs.

Three UPLC-MS[E] datasets were acquired for the Waters sgRNA LC-MS oligo Standard digested with RNase T1, RapiZyme MC1 and RapiZyme Cusativin. These datasets were acquired on the Xevo MRT Mass Spectrometer (Figure 1A,B), which uses MRT to achieve exceptional MS resolution, speed, sensitivity with sub-ppm mass accuracy for precursors and fragment ions.

Automated processing of all datasets was carried out using the waters_connect MAP Sequence App (ver 2.0), applying the processing parameters as shown in Figure 4. Both data channels— low-energy MS[E] for digested oligonucleotide precursors and elevated-energy MS[E] for their corresponding fragment ions— were analyzed with

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

10

a maximum mass error tolerance of 20 ppm. The initial search allowed for up to 10 missed cleavages, although filtering options can be applied to display results with fewer missed cleavages. Isotopic distributions of both precursor and fragment ions were assessed using an isotope similarity score to ensure high-quality matches between theoretical and experimental patterns, typically requiring a score of at least 70 out of 100. For fragment ions, additional criteria included a minimum sequence coverage of 20% per digested oligonucleotide product and detection of at least 20% of the theoretical fragment ions for that oligonucleotide.

The TIC chromatogram resulting from the digestion of Waters sgRNA with RapiZyme MC1 is displayed within the MAP Sequence App v2.0, as shown in Figure 5. The blue trace indicates that the majority of the detected chromatographic signal corresponds to theoretically predicted MC1 digestion products. A portion of the results table, generated using the processing parameters outlined in Figure 4, is shown in Figure 6. Green circles next to each digestion product denote unique identifications based on both low-energy MS^E data (precursor ions) and elevated-energy MS^E data (fragment ions). All digestion products were confidently identified by matching experimentally observed monoisotopic oligonucleotide precursor masses to their in-silico predicted counterparts. The average RMS mass accuracy across all assigned digestion products was within the sub-ppm range (<1 ppm). Additional details for each assigned precursor are accessible by clicking the corresponding sequence, highlighted in blue in the left column of the table.

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

11

Figure 4. MAP Sequence App (v2.0) processing parameters typically used for the precursors and fragments of oligonucleotide digested products.

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

12

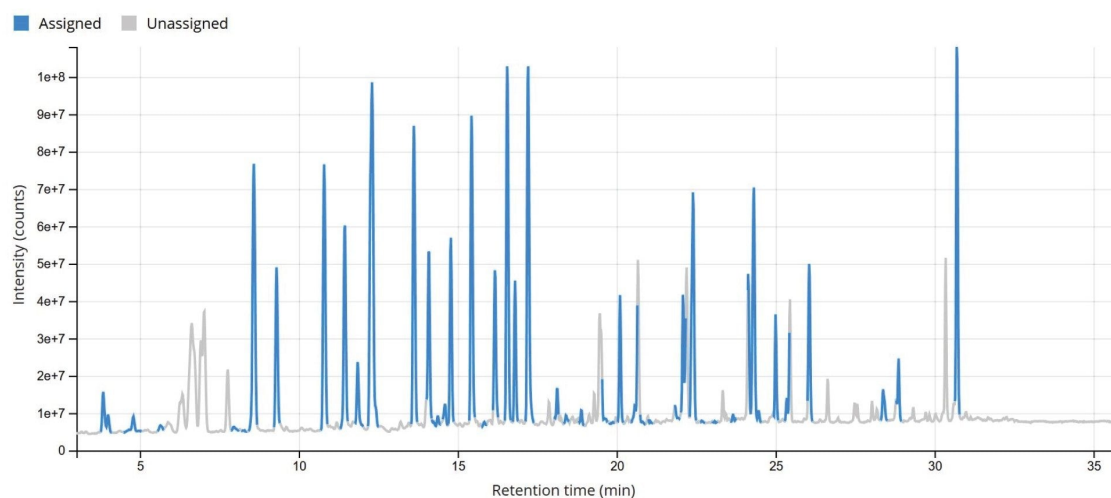*Figure 5. TIC chromatograms displayed by MAP Sequence v 2.0 for UPLC-MS^E analysis of the Waters sgRNA LC-MS Standard digested with RapiZyme MC1. The blue trace indicates that most of the chromatographic peaks detected in this chromatogram were assigned to predicted digestion products.*

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

13

Digestion product	Sequence	Length	Expected mass (Da)	Observed mass (Da)	Response	Mass error (ppm)	Observed RT (min)	Results ▲
X1:C4	XYYC-cP	4	1312.1116	1312.1111	215,847	-0.39	11.38	● Found
U17:C20	UCGC-cP	4	1261.1553	1261.1565	1,119,572	0.94	3.85	● Found
U17:U22	UCGCGU-cP	6	1912.2280	1912.2301	5,493,785	1.06	11.45	● Found
U17:U23	UCGCGUU-cP	7	2218.2533	2218.2546	4,533,617	0.58	14.10	● Found
U17:U24	UCGCGUUU-cP	8	2524.2786	2524.2819	3,055,158	1.29	16.18	● Found
G21:U24	GUUU-cP	4	1263.1233	1263.1238	531,616	0.40	4.01	● Found
U23:C30	UUUAGAGC-cP	8	2571.3171	2571.3190	3,400,808	0.75	16.81	● Found
U24:C30	UUAGAGC-cP	7	2265.2918	2265.2945	4,822,204	1.19	14.79	● Found
U31:A36	UAGAAA-cP	6	1967.2828	1967.2829	9,070,282	0.05	13.63	● Found
U31:U44	UAGAAAUAGCAAGU-cP	14	4561.6271	4561.6272	1,850,577	0.02	25.01	● Found
U37:U44	UAGCAAGU-cP	8	2594.3443	2594.3458	9,309,290	0.56	17.22	● Found
U45:A49	UAAAA-cP	5	1622.2354	1622.2355	9,238,532	0.10	10.80	● Found

Figure 6. Results table excerpt from the MAP Sequence App (v2.0). Unique digestion products (indicated by green circles) were confidently identified by matching the experimentally measured monoisotopic oligonucleotide precursors (MS1) with the in-silico predicted monoisotopic masses. The sequences of these unique digestion products were further validated through fragment ion assignments, using the processing parameters as described in Figure 4. The average mass accuracy for all assigned MC1 digestion products is <1 ppm. Additional details for the highlighted 14-mer digestion product (shown in blue) are provided in the following figure.

As an example, the U31:U44 digestion product—a 14-mer shown at the bottom of Figure 6—is further detailed in Figure 7. All isotopic peaks corresponding to the seven charge states (ranging from -2 to -8) detected for U31:U44 are shown in panel 7A, while the enlarged view for the M–8H]$^{8-}$ charge state is displayed in panel 7B. The high-resolution capability of the Xevo MRT Mass Spectrometer instrument (100,000 resolving power) enabled full isotopic resolution of this charge state. The overlaid extracted ion chromatograms (XICs) for all charge states, indicates a common elution time of 25.0 minutes (Figure 7C). The assigned dot-map fragmentation diagram of this digestion product, labeling all matched fragment ions and confirming complete (100%) sequence coverage from elevated-energy MS^E fragmentation, is displayed in Figure 7D.

In addition to the information presented in panels A–D of Figure 7, the MAP Sequence App v2.0 provided more detailed fragment ion data, as shown in the table from Figure 8A. This table lists all assigned fragment ions resulting from the fragmentation of the U31:U44 digestion product, along with their respective mass accuracies. Across all assigned digestion products, the average RMS mass accuracy for assigned fragment ions remains within the sub-ppm range (<1 ppm). Furthermore, individual isotopic distributions for each fragment ion can also

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

14

be visualized, as illustrated in Figure 8B.

The sequence coverage results for the Waters sgRNA, obtained after processing all three digests, are shown in Figure 9. The RapiZyme Cusativin enzyme achieved complete (100%) digested oligonucleotide sequence coverage due to its unique digestion specificity - cleaving at seven distinct dinucleotide sites - and its capacity to generate controlled missed cleavages. These longer digestion fragments, resulting from 2–4 missed cleavages, are more likely to have unique masses, and their sequences were also verified by fragment ion assignments, as illustrated in the figure containing the MAP Sequence Report output (Figure 10). RapiZyme Cusativin specifically cleaves at the 3'-end of cytidine residues, with four major cleavage sites (C_A, C_U, C_G, and U_A) and three minor sites (A_U, G_U, and U_U).[6-11] The enzyme predominantly produces digestion products terminating with a 3' cyclic phosphate (cP).

RapiZyme MC1 also demonstrated excellent performance, achieving 91% digested oligonucleotide sequence coverage. As a T2 ribonuclease, MC1 cleaves at the 5'-end of uridine residues, showing three major cleavage preferences (A_U, C_U, and U_U) and two minor ones (C_A and C_G), [6-11], likewise yielding products with a 3' cP. Due to these enzymatic properties, both Cusativin and MC1 provide greater sequence coverage compared to conventional RNase T1 digestion.[6-11] As shown in Figure 9, the MAP Sequence 2.0 App was able to combine results from multiple enzyme digest maps, ensuring complete and redundant sequence coverage for the Waters sgRNA.

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0
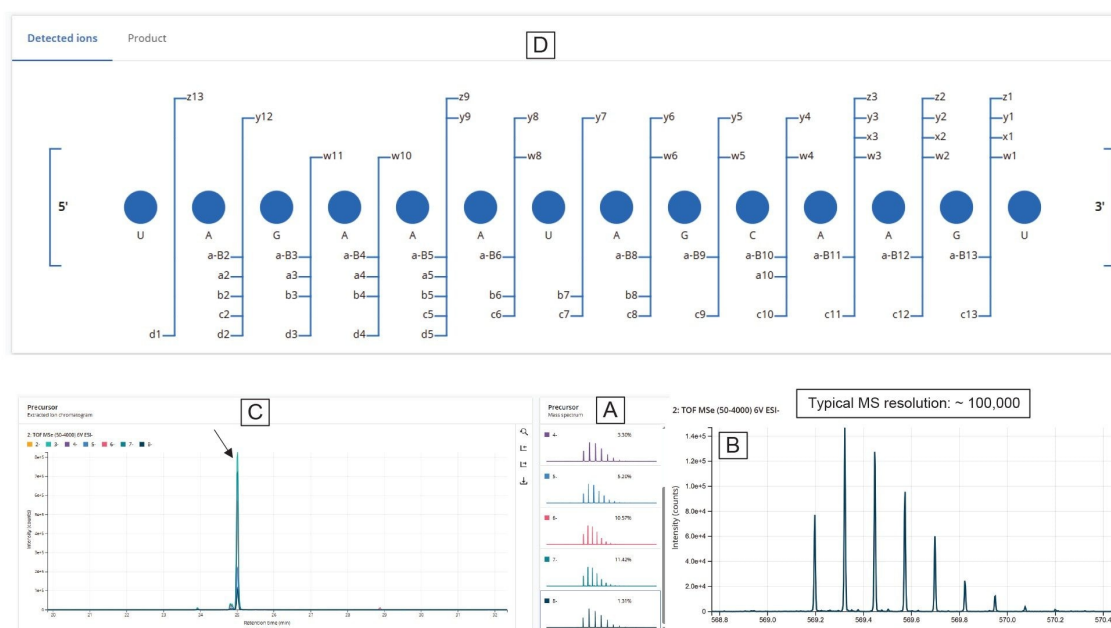
15

Figure 7. Detailed analysis of the U31:U44 MC1 digestion product (14-mer) as provided by the MAP Sequence App (v2.0): (A) Isotopic envelopes for all seven detected charge states (from –2 to –8); (B) zoomed view of the $[M–8H]^{8-}$ charge state; (C) overlaid XICs (extracted ion chromatograms) for all charge states, showing that this digestion product elutes at 25.0 min; (D) dot-map fragmentation diagram displaying all assigned fragment ions, demonstrating that full (100%) sequence coverage was achieved following elevated-energy MSE fragmentation. The Xevo MRT Mass Spectrometer provided a resolution of 100,000, fully resolving the isotopic distribution of the $[M–8H]^{8-}$ charge state.

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0
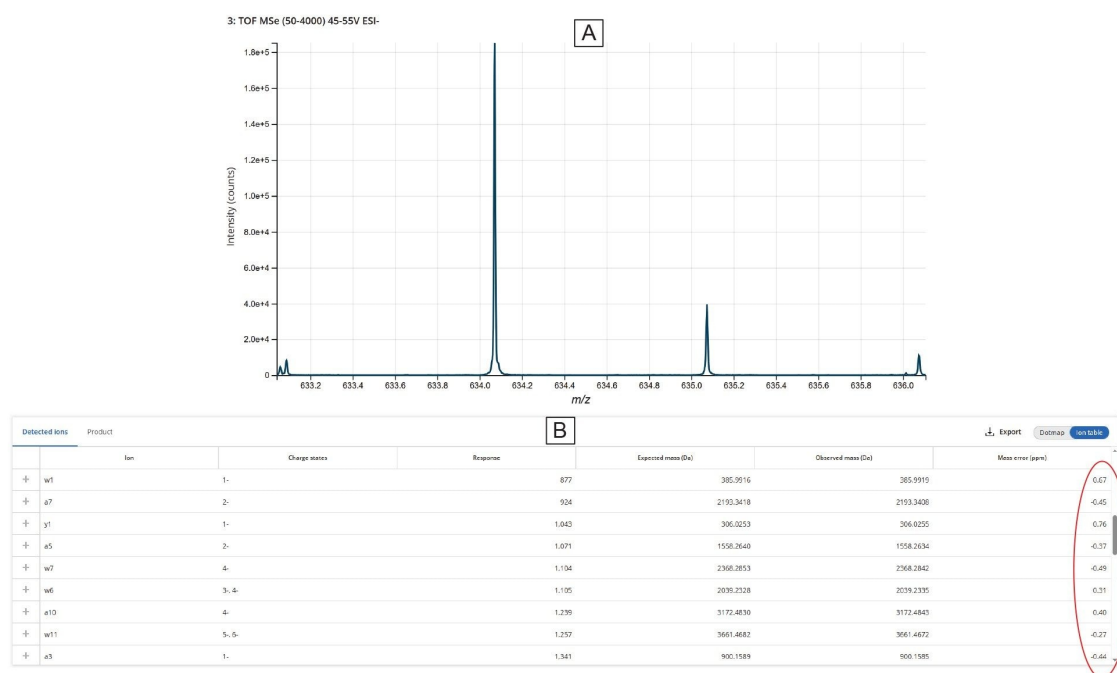
16

Figure 8. Fragment ion assignments for the U31:U44 MC1 digestion product, as generated by the MAP Sequence App (v2.0). Multiple precursor charge states (–2 to –8) underwent simultaneous elevated-energy MS^E fragmentation in the collision cell of the Xevo MRT Mass Spectrometer. (A) Portion of the fragment ion table showing sub-ppm (<1 ppm) mass accuracy for several fragment ions; (B) zoomed view of the isotopic pattern recorded for a singly charged fragment ion. The average mass accuracy for all assigned fragment ions is better than 1 ppm.

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

17

Figure 9. Comparison of sequence coverages obtained using MAP Sequence v2.0 for digestion of the Waters sgRNA with three enzymes— RapiZyme Cusativin, RapiZyme MC1, and RNase T1— assuming up to four missed cleavages. The combined sequence coverage demonstrates that complete sequence characterization of the Waters sgRNA can be achieved through the multi-enzyme digestion approach outlined in the workflow.

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

18

Digestion product	Sequence	Length	Expected mass (Da)	Observed mass (Da)	Response	Mass error (ppm)	Observed RT (min)	Sequence coverage (%)	Dotmap	Fragment count
X1:C4	XYYC	4	1312.1116	1312.1122	652253.25	0.44	8.62	100	●●●●	16
X1:C8	XYYCAACC	8	2580.2992	2580.3016	3098082.2	0.93	18.38	100	●●●●●●●●	37
X1:A9	XYYCAACCA	9	2909.3517	2909.3580	6125619.5	2.16	20.21	77.78	●●●●●●●○○	42
X1:C11	XYYCAACCAUC	11	3520.4183	3520.4245	14867263	1.75	22.25	100	●●●●●●●●●●●	66
A5:C11	AACCAUC	7	2208.3067	2208.3086	705488.6	0.83	13.57	100	●●●●●●●	20
A9:C11	AUC	3	940.1191	940.1202	1528482.6	1.12	1.91	100	●●●	10
U12:C16	UCGAC	5	1590.2078	1590.2103	15499664	1.53	5.48	100	●●●●●	31
U17:C20	UCGC	4	1261.1553	1261.1562	10648742	0.72	2.66	100	●●●●	20
U17:U22	UCGCGU	6	1912.2280	1912.2307	1241434.5	1.37	8.98	100	●●●●●●	22
G19:G21	GCG	3	995.1362	995.1343	10597.752	-1.82	0.94	100	●●●	10
G19:U22	GCGU	4	1301.1615	1301.1623	467344.34	0.62	2.69	100	●●●●	8
G19:U23	GCGUU	5	1607.1868	1607.1888	318810.06	1.29	4.91	100	●●●●●	14
G21:U23	GUU	3	957.0980	957.0992	4352871	1.26	1.74	100	●●●	9
G21:U24	GUUU	4	1263.1233	1263.1241	959135.44	0.64	2.75	100	●●●●	13
U23:C30	UUUAGAGC	8	2571.3171	2571.3193	3771432.5	0.85	14.72	100	●●●●●●●●	41
U24:C30	UUAGAGC	7	2265.2918	2265.2947	13236517	1.26	12.58	100	●●●●●●●	41
U25:C30	UAGAGC	6	1959.2665	1959.2691	18643324	1.32	9.84	100	●●●●●●	35
A26:C30	AGAGC	5	1653.2412	1653.2429	236029.52	1.01	6.64	100	●●●●●	9

Figure 10. Excerpt from the MAP Sequence v2.0 report confirming that all digested product assignments for the Waters sgRNA were extensively verified through corresponding fragment ion data. The UPLC-MS^E results shown here illustrate the oligonucleotide map obtained with RapiZyme Cusativin digestion.

Conclusion

- An informatics workflow, incorporating the waters_connect MAP Sequence App (v2.0) was successfully demonstrated for UPLC-MS^E oligonucleotide sequence mapping of sgRNA digests using the Xevo MRT QTOF Mass Spectrometer.

- Confident peak assessments for analysis of sgRNA digests was achieved with the Xevo MRT QTOF Mass Spectrometer, which delivered the required sensitivity and dynamic range while maintaining an average RMS mass accuracy of less than 1 ppm for both precursors and fragment ions of digestion products.

- Complete sgRNA sequence coverage (100%) was achieved by employing a combination of digestion enzymes.

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

19

- RapiZyme MC1 and RapiZyme Cusativin provided unique cleavage specificity and predictable missed cleavages, features that generated longer, unique, overlapping digestion products. Automated confirmation of modifications and linear sequence was achieved with the combined digests.

References

1. Jinek M, Chylinsky K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity, *Science*, 2012, 337, 816–821.

2. Jiang F, Doudna JA. CRISPR-Cas9 Structures and Mechanisms, *Annu Rev Biophys*, 2017, 46, 505–529.

3. Ganbaatar U, Liu C. CRISPR-Based COVID-19 Testing: Toward Next Generation Point of Care Diagnostics, *Front Cell Infect Microbiol*, 2021, 11, https://doi.org/10.3389/ fcimb.2021.663949 < https://www.frontiersin.org/journals/cellular-and-infection-microbiology/articles/10.3389/fcimb.2021.663949/full>

4. LC-MS Analysis of siRNA, Single Guide RNA and Impurities using the BioAccord System with ACQUITY Premier System and New Automated INTACT Mass Application, 2022, Waters application note. 720007546.

5. Goyon A, Scott B, Kurita K, Crittenden CM, Shaw D, Lin A, Yehl P, Zhang K. Full Sequencing of CRISPR/Cas9 Single Guide RNA (sgRNA) via Parallel Ribonuclease Digestions and Hydrophilic Interaction Liquid Chromatography High-Resolution Mass Spectrometry Analysis, *Anal Chem*, 2022, 93, 14792–14801, doi: 10.1021 /acs.analchem. 1c03533 <https://pubs.acs.org/doi/10.1021/acs.analchem.1c03533> .

6. RNA Digestion Product Mapping Using an Integrated UPLC-MS and Informatics Workflow, 2024, Waters application note. 720008553.

7. Tunable Digestion of RNA Using RapiZyme RNases to Confirm Sequence and Map Modifications, 2024, Waters application note. 720008539.

8. Oligo Mapping of mRNA Digests Using a Novel Informatics Workflow, 2025, Waters application note. 720008677.

9. Analysis of mRNA Cap Impurity Profiles and Capping Efficiency Using RapiZyme MC1 Ribonuclease, 2025, Waters application note. 720008793.

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

20

10. Grunberg S, Wolf EJ, Jin J, Ganatra MD, Becker K, Ruse C, Taron CH, Correa IR, Yigit E. Enhanced Expression and Purification of Nucleotide-specific Ribonucleases MC1 and Cusativin, *Protein Expr Purif Acid Res*, 2022, 190, 105987, doi: 10.1016/j. pep.2021.105987 < https://www.sciencedirect.com/science/article/pii/S1046592821001704> .

11. 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, 23, 7021.

Featured Products

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

Xevo MRT Mass Spectrometer <https://www.waters.com/nextgen/global/products/mass-spectrometry/mass-spectrometry-systems/xevo-mrt.html>

waters_connect Informatics Platform <https://www.waters.com/nextgen/global/products/informatics-and-software/waters_connect/waters-connect-informatics-platform.html>

720009130, November 2025

Sequence Mapping of sgRNA Digests: Leveraging Xevo™ MRT Mass Spectrometer Performance and Streamlining Data Analysis with MAP Sequence 2.0

21