MITSUI KNOWLEDGE INDUSTRY

Application Note

Comprehensive mRNA Sequence Mapping using Multi-Enzyme Digestion, DIA LC-MS/MS, and Automated AQXeNA Software Analysis

Authors

Yuki Matsubara[1], Akari Ito[1], Yasuto Yokoi[1], Nick Pittman[2], Catalin Doneanu[3], Matthew Gorton[2], Scott Berger[3]

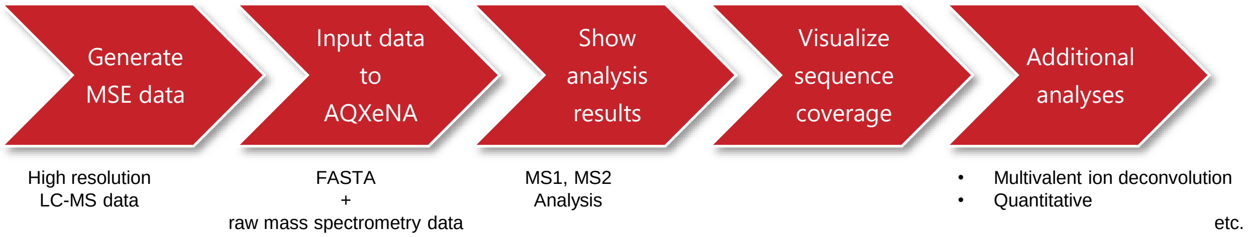
1. Mitsui Knowledge Industry, 2. Waters Corporation, Wilmslow UK, 3. Waters Corporation, Milford MA,

Abstract

In this Application Note, we describe an RNA Sequence Mapping workflow for confident product characterization and analysis by combining (1) multi-ribonuclease digestion using RNase T1 and RapiZyme MC1 to maximize sequence coverage, (2) data-independent acquisition (DIA) to ensure comprehensive data collection and accurate digest product identification, and (3) automated data analysis with the AQXeNA software to streamline processing, deconvolution, and sequence mapping. We demonstrate the application of this workflow to a GFP (jellyfish green fluorescent protein) mRNA, achieving high sequence coverage and confident digest product identification, superior to single-enzyme digestions.

Benefit

・The AQXeNA software, combined with a multi-ribonuclease digestion strategy and DIA, automates data processing, deconvolution, and sequence mapping. This leads to more comprehensive digest product detection, improved sequence coverage, and increased confidence in mRNA sequence confirmation for RNA therapeutic research and development.



Generate MSE data → Input data to AQXeNA → Show analysis results → Visualize sequence coverage → Additional analyses

Generate MSE data: High resolution LC-MS data

Input data to AQXeNA: FASTA + raw mass spectrometry data

Show analysis results: MS1, MS2 Analysis

Additional analyses:
- Multivalent ion deconvolution
- Quantitative

etc.

Figure 1: Workflow diagram illustrating the steps involved data analysis using AQXeNA software.

Introduction

The rapid growth of RNA therapeutics, including mRNA vaccines and gene editing technologies, demands robust analytical methods to ensure product quality, safety, and efficacy. Sequence confirmation is a critical step in mRNA characterization, verifying the correct sequence has been synthesized, and the absence of unintended mutations or errors. Comprehensive mRNA characterization also includes identification of modifications and assessment of degradation

Traditional sequencing methods (Sanger, NGS – Next Generation Sequencing) are cost-effective and high-throughput but can struggle with modified RNA molecules and may not provide the confidence level required for therapeutic mRNA.

LC-MS/MS offers high specificity and sensitivity, enabling precise identification and quantification of RNA modifications, providing a powerful alternative for reliable sequence confirmation. One example of this is provided by Waters, offering an automated sequence confirmation workflow as part of the suite of LC-MS and informatics tools for RNA attribute monitoring.

Traditional LC-MS-based workflows for RNA oligonucleotide sequence analysis have historically faced several challenges:

- Difficulty in Analyzing Long RNA Molecules: Intact mRNA molecules are too large and heterogeneous for intact high-resolution LC-MS/MS analysis.
- Incomplete Sequence Coverage with Single-Enzyme Digests: Digestion with a single ribonuclease often results in incomplete sequence coverage. Some mRNA regions may not be represented by unique, detectable oligonucleotides.
- Limitations of MS1-Only Identification: Using only the oligonucleotide precursor mass (MS1) to characterize oligonucleotide sequences has the potential for ambiguity due to the presence of isomeric digest products.
- Data Acquisition Challenges with Data-Dependent Acquisition (DDA): Traditional DDA methods can miss low-abundance oligonucleotides, further reducing sequence coverage.
- Complex Data Analysis: LC-MS/MS data from RNA digests can contain numerous MS peaks, chemical noise, and composite fragment ion spectra (MS2) containing fragment ions from multiple precursor ions, requiring time consuming data processing.

In order to fully characterize larger RNA constructs early in product development, MS2 analysis provides greater confidence in assignment of digest products. This application note presents an integrated LC-MS/MS workflow that directly addresses these challenges, combining multi-enzyme digestion, DIA, and the AQXeNA software for automated data analysis.

Experimental

Materials and Methods

Dipropylethylamine (DPA, 99 % purity, catalogue number D214752-500ML) and 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP, 99% purity, catalogue number 105228-100G) were purchased from Millipore Sigma (St Louis, MO, USA). Methanol (LC-MS grade, catalogue number 34966-1L) was obtained from Honeywell (Charlotte, NC, USA). HPLC grade Type I deionized (DI) water was purified using a Milli-Q system (Millipore, Bedford, MA, USA). Ultrapure nuclease-free water (catalogue number J71786.AE) for mRNA digestions was purchased from Thermo Fisher Scientific (Waltham, MA, USA).

An mRNA construct based on the jellyfish green fluorescent protein (GFP) sequence was custom-made via IVT (in vitro transcription) synthesis by Biosynthesis (Lewisville, TX, USA). The mRNA molecule was synthesized with a Cap1 structure (with the sequence: 7MeGpppA(2'-OMe) with elemental composition: $C_{32}H_{42}N_{15}O_{26}P_{5}$), followed by 1019 nucleotides and no Poly(A) Tail sequence.

Chromatographically purified, animal free, ribonuclease T1 (catalogue no LS01490, 500kU), isolated from Aspergillus oryzae, was purchased from Worthington Biochemical Corporation (Lakewood, NJ, USA). The lyophilized enzyme was dissolved in 5 mL of 100 mM ammonium bicarbonate (catalogue no 5.33005-50G, Millipore Sigma) to prepare a solution containing 100 units/µL. For mRNA digestion with RNase T1, 5 µL of 5 µM GFP mRNA were mixed with 25 µL of nuclease-free water and 10 µL of RNase T1 enzyme (1000 units) and the digestion was allowed to proceed at 37oC for 15 min. The digest was analyzed immediately by LC-MS using 5 µL injections.

RapiZyme MC1 (Waters P/N 186011190, 10000 units/tube) is a novel RNA digestion enzymes recently introduced by Waters Corporation [1-3]. Before RapiZyme MC1 digestion, the GFP mRNA (10µL, 5 µM solution) was denatured at 90oC for 2 min in a buffer containing 200 mM ammonium acetate pH 8.0. The samples were cooled on ice and microcentrifuged to collect the sample droplets. After adding 50 units of digestion enzyme (1 µL of RapiZyme MC1) and 8 µL of nuclease-free water to obtain a final volume of ~ 20 µL, the mRNA was digested at 37oC for 60 min in an Eppendorf thermomixer. The enzymatic digestion was stopped by heating to 70 ºC for 15 min, to inactivate the enzyme. The digest was analyzed immediately by LC-MS using 5 µL injections.

Oligonucleotide digests were separated on a 2.1 x 150 mm AQUITY PREMIER BEH C18 UPLC column (Waters P/N 186009486) via ion-pair reversed-phase (IP-RP) chromatography. The ACQUITY PREMIER UPLC System (Waters) used a mobile phase containing 10 mM dipropylamine (DPA) and 40 mM hexafluoroisopropanol (HFIP) in DI water as Eluent A and 10 mM DPA and 40 mM HFIP in 50% methanol as Eluent B. All digests were separated using the same gradient conditions (from 0 to 40% Eluent B over 45 min) with a total run time of 1 hour for each injections. High-resolution (>30,000) ESI-MSE (DIA) spectra were acquired in negative ion mode on a Xevo G3 QTof instrument (Waters, Milford, MA, USA).

Data Analysis

All datasets were acquired using the waters_connect UNIFI App version 3.6.0.21 and converted to the .mzML file format. AQXeNA software was used for peak extraction, deconvolution, oligonucleotide mapping, and sequence coverage analysis.

AQXeNA Workflow:

Data Import: waters_connect datasets (*.uep projects) acquired in DIA (MSE) mode were converted to open-source .mzML files for data sharing and were imported into the AQXeNA software.

Sequence Input: The mRNA sequence of the GFP mRNA was entered into the software in FASTA format.

Enzyme Specificity Definition: The cleavage specificities of RNase T1 and RapiZyme MC1 were selected from a comprehensive list of pre-defined enzymes within the software, respectively. The number of missed cleavages allowed was specified for each enzyme (e.g. up to 2 missed cleavages for RNase T1 and up to 4 for RapiZyme MC1).

Automated Data Processing: AQXeNA's deconvolution algorithms were applied to automatically process the DIA (MSE) data.

The deconvolution process in AQXeNA involves several steps:

- Peak Detection: The software identifies potential precursor ions (MS1 peaks) based on their intensity and isotopic patterns.
- Isotopic Pattern Matching: The software compares the observed isotopic distribution of each peak with theoretical isotopic distributions to confirm its elemental composition and determine its charge state.
- Reconstructed MS2 Spectra Generation: For each precursor ion, the software searches for the corresponding fragment ions based on their mass differences and elution profiles. These fragment ions are then used to reconstruct an MS2 spectrum, which represents the fragmentation pattern of the precursor ion.

Reconstructed MS2 spectra are generated by correlating precursor and fragment ions, enabling confident oligonucleotide identification.Sequence Mapping and Coverage Calculation: Identified oligonucleotides were automatically assigned to the target mRNA sequence based on their mass and reconstructed MS2 spectra. Sequence coverage was calculated as the percentage of the mRNA sequence covered by identified oligonucleotides.

Results and Discussion

Multi-Enzyme Digestion

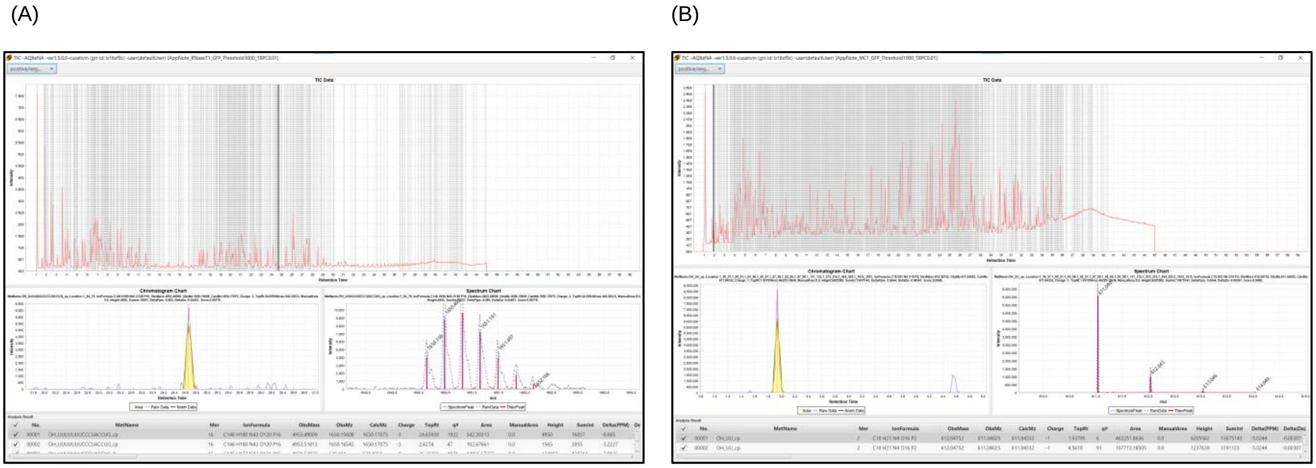
As intact mRNA molecules are too large and heterogeneous for direct, high-resolution LC-MS/MS analysis, enzymatic digestion is necessary to achieve confident sequence information. However, a single ribonuclease often yields incomplete sequence coverage. In order to improve sequence coverage, we employed a dual-enzyme approach.

RNase T1: Cleaves specifically at the 3’ end of guanosine (G) residues.
RapiZyme MC1: Primarily cleaves at the 5’-end of uridine (U) residues (major sites: A_U, C_U, U_U; minor sites: C_A, C_G), and frequently produces missed cleavages.

RNA undergoes separate digestion and analysis by two enzymes in parallel and the resulting data is compared against the expected sequence to overcome areas lacking coverage by a single enzyme approach. The chosen enzyme combination generates a greater number of overlapping oligonucleotide fragments, increasing the likelihood of unique masses and significantly improving sequence coverage compared to single-enzyme digests.

Data-Independent Acquisition (DIA)

Traditional data-dependent acquisition (DDA) can miss low-abundance ions precursors. Furthermore, relying solely on precursor ion mass (MS1) for identification can lead to false positives. To overcome these limitations, we employed DIA (MSE) data acquisition. Unlike data-dependent acquisition (DDA), which selects specific precursors for fragmentation, DIA fragments all precursors, resulting in composite spectra containing fragment ion peaks from multiple precursor ions. This ensures that no digest products, even low-abundance ones, are missed.



(A)

TIC -AQIleNA -ver1.5.0.6-cusativin (git-id: b16cf0c) -user(defaultUser) [AppNote_RNaseT1_GFP_Threshold1000_T8PC0.01]

Analysis Result

✓	No.	MetName	Mer	IonFormula	ObsMass	ObsMz	CalcMz	Charge	TopRt	q#	Area	ManualArea	Height	SumInt	Delta(PPM)	De
✓	00001	OH_UUUUUUUCCCUACCUG_cp	16	C146 H180 N43 O120 P16	4953.49009	1650.15608	1650.17075	-3	24.65938	1922	542.30313	0.0	4950	16057	-8.885	
✓	00002	OH_UUUUUUUCCCUACCUG_cp	16	C146 H180 N43 O120 P16	4953.51813	1650.16540	1650.17075	-3	2.6274	47	102.67661	0.0	1565	3055	-3.2227	

(B)

TIC -AQIleNA -ver1.5.0.6-cusativin (git-id: b16cf0c) -user(defaultUser) [AppNote_MC1_GFP_Threshold1000_T8PC0.01]

Analysis Result

✓	No.	MetName	Mer	IonFormula	ObsMass	ObsMz	CalcMz	Charge	TopRt	q#	Area	ManualArea	Height	SumInt	Delta(PPM)	Delta(Da)
✓	00001	OH_UU_cp	2	C18 H21 N4 O16 P2	612.04752	611.04025	611.04332	-1	1.93795	6	462251.8636	0.0	6205582	13675143	-5.0244	-0.00307
✓	00002	OH_UU_cp	2	C18 H21 N4 O16 P2	612.04752	611.04025	611.04332	-1	4.5618	91	107713.18505	0.0	1237639	3191123	-5.0244	-0.00307

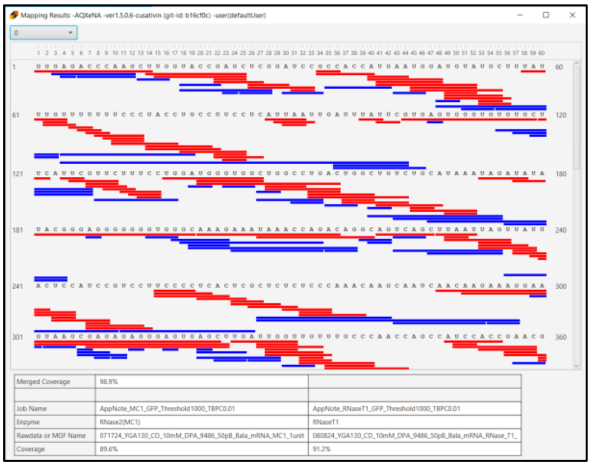
Figure 2: Total ion chromatograms (TICs) of the GFP mRNA digested with (A) RNase T1 and (B) RapiZyme MC1, demonstrating the distinct chromatographic profiles resulting from the different cleavage patterns of the two enzymes.

Automated Processing with AQXeNA

As detailed in the Experimental section, the AQXeNA software automatically processed the DIA data from both RNase T1 and RapiZyme MC1 digests, assigning and identifying oligonucleotides to the GFP mRNA sequence. Figure 1 shows the workflow diagram.

When comparing the total ion chromatograms (TICs) of the GFP mRNA digests (Figure 2) with RNase T1 (A) and RapiZyme MC1 (B), the distinct chromatographic profiles demonstrate the different cleavage patterns of the two enzymes. RNase T1, with its G-specific cleavage, produces a larger number of smaller peaks, while RapiZyme MC1, with its preference for U-rich regions and tendency for missed cleavages, generates longer peaks that tend to elute later.

We can visualize the sequence coverage map obtained by combining the data from both enzymes (Figure 3). Identification based on reconstructed MS2 Data was processed with AQXeNA. The use of fragment ion information (MS2) allows for unambiguous oligonucleotide identification, resolving isomeric species that would be indistinguishable based on MS1 data alone. Oligonucleotides with the same overall mass, same elemental composition, but different sequences will fragment differently when subjected to fragmentation. These different fragmentation patterns, observed in the MS2 spectra, are characteristic of each sequence isomer, enabling their differentiation. This coverage map represents the sequence coverage achieved at the MS2 level, considering only sequences that obtained a confidence score above a defined threshold based on the reconstructed MS2 spectra evaluation with AQXeNA. The combined approach achieved a 97.8% sequence coverage, demonstrating the comprehensive and reliable sequence information obtained with the combined enzyme digestion and automated workflow.



Merged Coverage	98.9%	
Job Name	AppNote_MC1_GFP_Threshold1000_TBPC0.01	AppNote_RNaseT1_GFP_Threshold1000_TBPC0.01
Enzyme	RNase2(MC1)	RNaseT1
Rawdata or MGF Name	071724_YGA130_CD_10mM_DPA_9486_50pB_Bala_mRNA_MC1_1unit	080824_YGA130_CD_10mM_DPA_9486_50pB_Bala_mRNA_RNase_T1_
Coverage	89.6%	91.2%

Merged Coverage	98.9	
Enzyme	RNase T1	MC1
Coverage (Rank first and Above threshold)	91.2	89.6

Figure 3: GFP mRNA sequence coverage map obtained with the combined RNase T1/MC1 digestion, generated at the MS2 level, considering only sequences that obtained a confidence score above a pre-defined confidence threshold based on the reconstructed MS2 spectra evaluation.

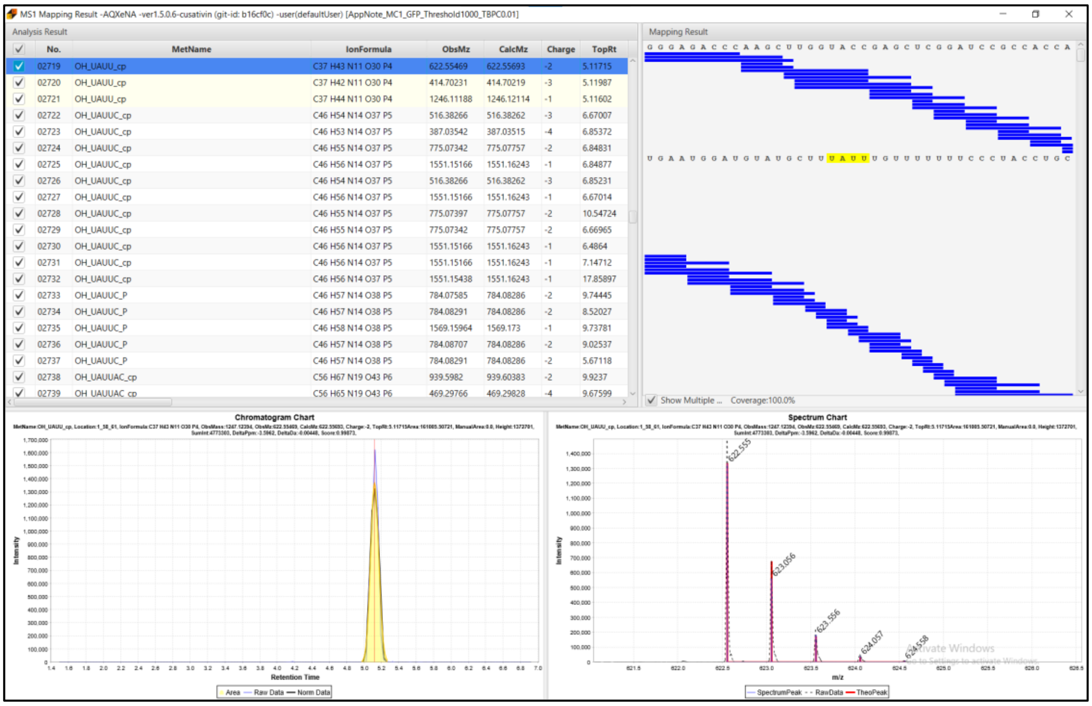
Comparing sequence coverage at the MS2 level for RNase T1, RapiZyme MC1, and their combination (Table 1) shows the combined approach (97.8%) provides higher sequence coverage than either RNase T1 (86.8%) or RapiZyme MC1 (83.5%) digestions alone.AQXeNA also achieved high sequence coverage at the MS1 level, which is valuable for providing a comprehensive overview of the sequence. Figures 4 and 5 illustrate this, showing the sequence coverage map and analysis results as displayed in the AQXeNA software. While this MS1-level coverage offers a useful general picture, it is based solely on precursor ion information (MS1). For greater confidence in accurate sequence assignment, the incorporation of confidence scores from reconstructed MS2 spectra evaluation is essential, as MS1 data alone does not ensure this level of accuracy and can be prone to the detection of ambiguous digest products due to the presence of isomeric oligonucleotide species.

Table 1: Comparison of sequence coverage results obtained at the MS2 level, achieved with RNase T1 digestion alone, RapiZyme MC1 digestion alone, and the combined RNase T1/MC1 digestion. Sequence coverage values at the MS2 level are shown, considering only sequences that obtained a confidence score above a defined threshold based on the reconstructed MS2 spectra evaluation.

Enzyme	MS2 coverage (%) Rank first and Above threshold	MS2 coverage (%) Rank first only
RNase T1	91.2	97.8
RNase 2(MC1)	89.6	100
RNase T1 and MC1	98.9	100

As shown in Table 1, the MS2 coverage was generally lower than the MS1 coverage. This reflects the increased stringency of MS2-based identification, which significantly reduces false positive assignments that may be present in MS1-only analysis. The “Rank first and Above threshold” criterion considers all identified oligonucleotides that meet a certain confidence threshold [4], while the “Rank first only” criterion considers the most likely oligonucleotide assignment for each precursor ion.

These results demonstrate the utility of the combined RNase T1/RapiZyme MC1 and AQXeNA software workflow for comprehensive mRNA sequence confirmation using DIA LC-MS/MS.



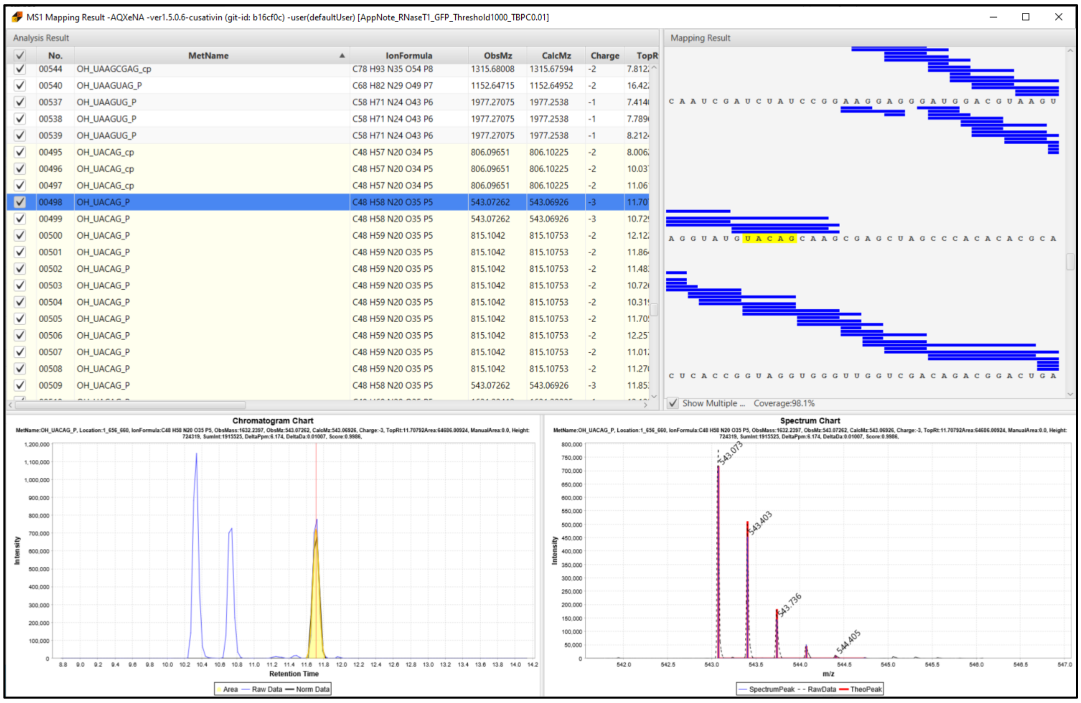
MS1 Mapping Result -AQXeNA -ver1.5.0.6-cusativin (git-id: b16cf0c) -user(defaultUser) [AppNote_MC1_GFP_Threshold1000_TBPC0.01]

Analysis Result

✓	No.	MetName	IonFormula	ObsMz	CalcMz	Charge	TopRt
✓	02719	OH_UAUU_cp	C37 H43 N11 O30 P4	622.55469	622.55693	-2	5.11715
✓	02720	OH_UAUU_cp	C37 H42 N11 O30 P4	414.70231	414.70219	-3	5.11987
✓	02721	OH_UAUU_cp	C37 H44 N11 O30 P4	1246.11188	1246.12114	-1	5.11602
✓	02722	OH_UAUUC_cp	C46 H54 N14 O37 P5	516.38266	516.38262	-3	6.67007
✓	02723	OH_UAUUC_cp	C46 H53 N14 O37 P5	387.03542	387.03515	-4	6.85372
✓	02724	OH_UAUUC_cp	C46 H55 N14 O37 P5	775.07342	775.07757	-2	6.84831
✓	02725	OH_UAUUC_cp	C46 H56 N14 O37 P5	1551.15166	1551.16243	-1	6.84877
✓	02726	OH_UAUUC_cp	C46 H54 N14 O37 P5	516.38266	516.38262	-3	6.85231
✓	02727	OH_UAUUC_cp	C46 H56 N14 O37 P5	1551.15166	1551.16243	-1	6.67014
✓	02728	OH_UAUUC_cp	C46 H55 N14 O37 P5	775.07397	775.07757	-2	10.54724
✓	02729	OH_UAUUC_cp	C46 H55 N14 O37 P5	775.07342	775.07757	-2	6.66965
✓	02730	OH_UAUUC_cp	C46 H56 N14 O37 P5	1551.15166	1551.16243	-1	6.4864
✓	02731	OH_UAUUC_cp	C46 H56 N14 O37 P5	1551.15166	1551.16243	-1	7.14712
✓	02732	OH_UAUUC_cp	C46 H56 N14 O37 P5	1551.15438	1551.16243	-1	17.85897
✓	02733	OH_UAUUC_P	C46 H57 N14 O38 P5	784.07585	784.08286	-2	9.74445
✓	02734	OH_UAUUC_P	C46 H57 N14 O38 P5	784.08291	784.08286	-2	8.52027
✓	02735	OH_UAUUC_P	C46 H58 N14 O38 P5	1569.15964	1569.173	-1	9.73781
✓	02736	OH_UAUUC_P	C46 H57 N14 O38 P5	784.08707	784.08286	-2	9.02537
✓	02737	OH_UAUUC_P	C46 H57 N14 O38 P5	784.08291	784.08286	-2	5.67118
✓	02738	OH_UAUUAC_cp	C56 H67 N19 O43 P6	939.5982	939.60383	-2	9.9237
✓	02739	OH_UAUUAC_cp	C56 H65 N19 O43 P6	469.29766	469.29828	-4	9.67599

Mapping Result
☑ Show Multiple ... Coverage:100.0%
Chromatogram Chart
MetName:OH_UAUU_cp, Location:1_58_61, IonFormula:C37 H43 N11 O30 P4, ObsMass:1247.12394, ObsMz:622.55469, CalcMz:622.55693, Charge:-2, TopRt:5.11715Area:161005.50721, ManualArea:0.0, Height:1372701, SumInt:4773303, DeltaPpm:-3.5962, DeltaDa:-0.00448, Score:0.99873.
Intensity
Retention Time
Area — Raw Data — Norm Data
Spectrum Chart
622.555
623.056
623.556
624.057
624.558
m/z
SpectrumPeak - - RawData — TheoPeak

Figure 4: AQXeNA software screenshot displaying the sequence coverage map and analysis results for the RapiZyme MC1 digest alone. The sequence coverage map represents the coverage achieved at the MS1 level. (A) GFP mRNA sequence with identified oligonucleotides highlighted. (B) List of identified oligonucleotides (partial). (C) Chromatogram of a selected oligonucleotide precursor (UAUU). (D) ESI-MS spectrum of the selected precursor.



MS1 Mapping Result -AQXeNA -ver1.5.0.6-cusativin (git-id: b16cf0c) -user(defaultUser) [AppNote_RNaseT1_GFP_Threshold1000_TBPC0.01]

Analysis Result

✓	No.	MetName	IonFormula	ObsMz	CalcMz	Charge	TopR
✓	00544	OH_UAAGCGAG_cp	C78 H93 N35 O54 P8	1315.68008	1315.67594	-2	7.812
✓	00540	OH_UAAGUAG_P	C68 H82 N29 O49 P7	1152.64715	1152.64952	-2	16.42
✓	00537	OH_UAAGUG_P	C58 H71 N24 O43 P6	1977.27075	1977.2538	-1	7.414
✓	00538	OH_UAAGUG_P	C58 H71 N24 O43 P6	1977.27075	1977.2538	-1	7.789
✓	00539	OH_UAAGUG_P	C58 H71 N24 O43 P6	1977.27075	1977.2538	-1	8.212
✓	00495	OH_UACAG_cp	C48 H57 N20 O34 P5	806.09651	806.10225	-2	8.006
✓	00496	OH_UACAG_cp	C48 H57 N20 O34 P5	806.09651	806.10225	-2	10.03
✓	00497	OH_UACAG_cp	C48 H57 N20 O34 P5	806.09651	806.10225	-2	11.06
✓	00498	OH_UACAG_P	C48 H58 N20 O35 P5	543.07262	543.06926	-3	11.70
✓	00499	OH_UACAG_P	C48 H58 N20 O35 P5	543.07262	543.06926	-3	10.72
✓	00500	OH_UACAG_P	C48 H59 N20 O35 P5	815.1042	815.10753	-2	12.12
✓	00501	OH_UACAG_P	C48 H59 N20 O35 P5	815.1042	815.10753	-2	11.86
✓	00502	OH_UACAG_P	C48 H59 N20 O35 P5	815.1042	815.10753	-2	11.48
✓	00503	OH_UACAG_P	C48 H59 N20 O35 P5	815.1042	815.10753	-2	10.72
✓	00504	OH_UACAG_P	C48 H59 N20 O35 P5	815.1042	815.10753	-2	10.31
✓	00505	OH_UACAG_P	C48 H59 N20 O35 P5	815.1042	815.10753	-2	11.70
✓	00506	OH_UACAG_P	C48 H59 N20 O35 P5	815.1042	815.10753	-2	12.25
✓	00507	OH_UACAG_P	C48 H59 N20 O35 P5	815.1042	815.10753	-2	11.01
✓	00508	OH_UACAG_P	C48 H59 N20 O35 P5	815.1042	815.10753	-2	11.27
✓	00509	OH_UACAG_P	C48 H58 N20 O35 P5	543.07262	543.06926	-3	11.85

Mapping Result
☑ Show Multiple ... Coverage:98.1%
Chromatogram Chart
MetName:OH_UACAG_P, Location:1_656_660, IonFormula:C48 H58 N20 O35 P5, ObsMass:1632.2397, ObsMz:543.07262, CalcMz:543.06926, Charge:-3, TopRt:11.70792Area:64686.00924, ManualArea:0.0, Height:724319, SumInt:1915525, DeltaPpm:6.174, DeltaDa:0.01007, Score:0.9986,
Intensity
Retention Time
Area — Raw Data — Norm Data
Spectrum Chart
543.073
543.403
543.736
544.405
m/z
SpectrumPeak - - RawData — TheoPeak

Figure 5: AQXeNA software screenshot displaying the sequence coverage map and the analysis results for the RNase T1 digest abne. The sequence coverage map represents the coverage achieved at the MS1 level. (A) GFP mRNA sequence with identified oligonucleotides highlighted. (B) List of identified oligonucleotides (partial). (C) Chromatogram of a selected oligonucleotide precursor (UACAG). (D) ESI-MS spectrum of the selected precursor.

Conclusion

This application note demonstrates a streamlined LC-MS/MS workflow for mRNA sequence confirmation that directly addresses key challenges in the field. By combining multi-ribonuclease digestion (RNase T1 and RapiZyme MC1) for comprehensive sequence coverage, data-independent acquisition (DIA) for complete data collection and accurate identification, and the AQXeNA software for automated data processing and deconvolution, this workflow enables reliable and efficient mRNA sequence analysis. This approach supports the development and quality control of RNA therapeutics by providing a robust and confident method for sequence verification.

References

1. Tunable Digestion of RNA Using RapiZyme RNases to Confirm Sequence and Map Modifications, 2024, Waters application note P/N 720008539EN.
2. RNA Digestion Product Mapping Using an Integrated UPLC-MS and Informatics Workflow, 2024, Waters application note P/N 720008553EN.
3. Oligo Mapping of mRNA Digests Using a Novel Informatics Workflow, 2025, Waters application note P/N 720008677EN.
4. Nakayama H. *et al.* “Ariadne: a database search engine for identification and chemical analysis of RNA using tandem mass spectrometry data.” *Nucleic Acids Research* 37, e47 (2009).