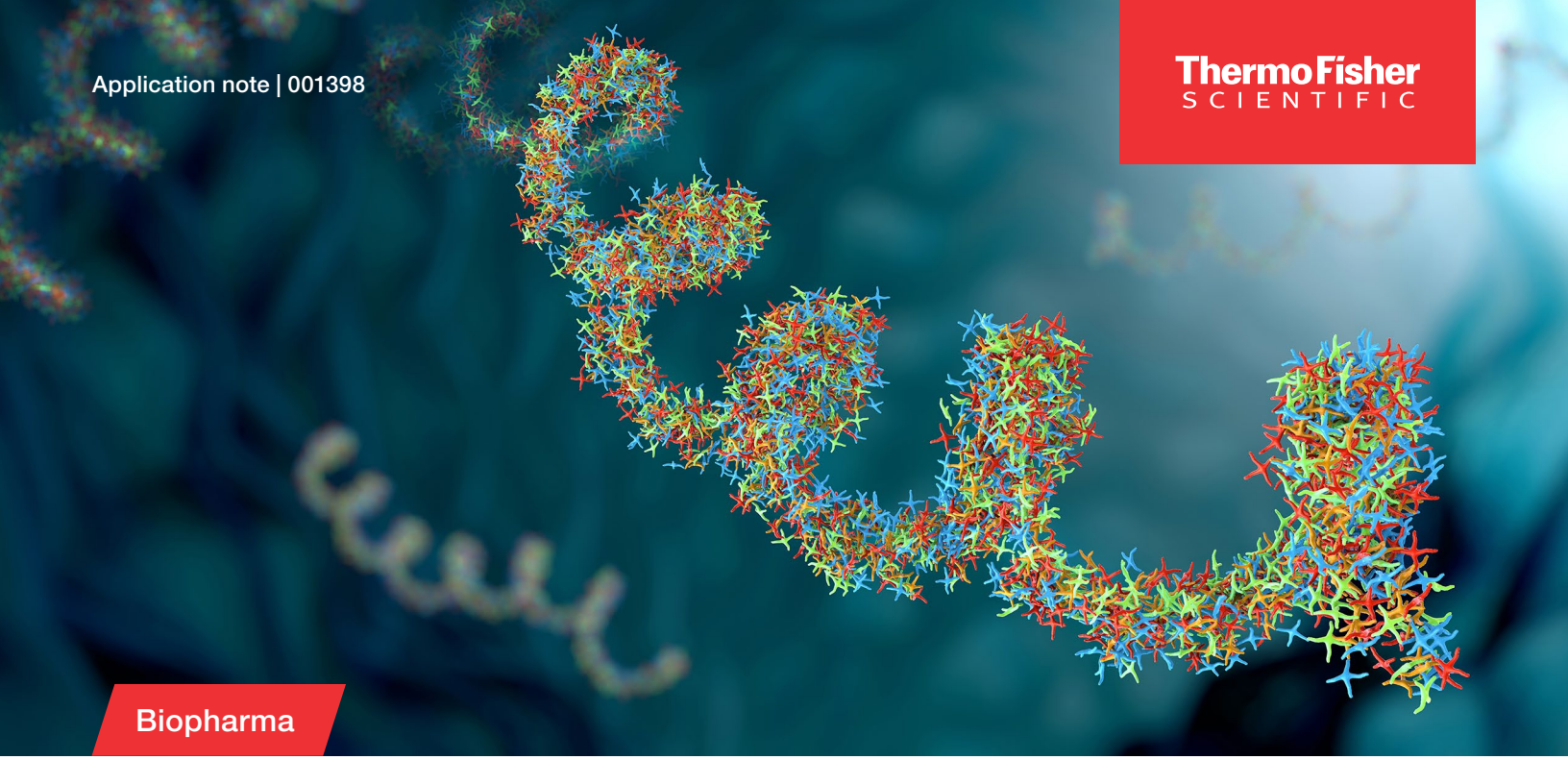


Application note | 001398

Biopharma

Streamlining characterization and monitoring of oligonucleotide impurities using an Orbitrap-based LC-HRAM-MS platform

Authors

Hao Yang[1], Keeley Murphy[2], Yi Zhang[3], Min Du[2]

Thermo Fisher Scientific, [1]San Jose, CA, US, [2]Boston, MA, US, [3]Philadelphia, PA, US

Keywords

Modified ribonucleic acids (RNAs), therapeutic oligonucleotides, full-length product (FLP), liquid chromatography–high-resolution accurate-mass mass spectrometry, ion-pairing reversed-phase liquid chromatography (IPRP-LC), Vanquish Horizon UHPLC, Vanquish Flex UHPLC, Orbitrap Exploris 240 mass spectrometer, Orbitrap Exploris MX mass detector, DNAPac reversed-phase column, BioPharma Finder software, Chromeleon Chromatography Data System (CDS) software, eWorkflow procedure

Application benefits

- Comprehensive characterization of oligonucleotides and their impurities using Thermo Scientific™ BioPharma Finder™ 5.1 software for processing data-dependent MS^2 ($ddMS^2$) data that are collected on the Thermo Scientific™ Orbitrap Exploris™ 240 mass spectrometer
- Confident base-by-base sequence confirmation and localization of modifications for oligonucleotides and their impurities
- Thermo Scientific™ Chromeleon™ eWorkflow™ procedure enables facile deployment of an LC-HRAM-MS method on the Orbitrap Exploris 240 mass spectrometer or the Thermo Scientific™ Orbitrap Exploris™ MX mass detector for oligonucleotide impurity analysis, crucial for obtaining consistent results for relative quantitation of FLP and impurities

Goal

- Demonstrate the use of BioPharma Finder 5.1 software for impurity identification and sequence mapping using $ddMS^2$ data collected on the Orbitrap Exploris 240 mass spectrometer
- Demonstrate the use of Chromeleon CDS 7.3.1 software for relative quantitation of FLP and impurities using full MS data collected on the Orbitrap Exploris MX mass detector
- Demonstrate the use of the Chromeleon eWorkflow procedure for direct method transfer between Orbitrap Exploris-based LC-HRAM-MS instruments for quantitative impuritiy analysis

Introduction

Synthetic therapeutic oligonucleotides, such as antisense and small interfering ribonucleic acid that selectively bind RNA through Watson-Crick base pairing and thereby alter gene expression, have gained significant interest over recent years. At present, 15 oligonucleotide therapeutics have been approved in the United States to treat various rare diseases.[1] Many of these therapeutic oligonucleotides are heavily chemically modified (e.g., 2'-O-methyl, 2'-O-methoxyethyl, 2'-fluoro, phosphorothiolate, and phosphorodiamidate linkages, etc.) to increase activity, enhance stability against nuclease degradation, modulate protein binding, and decrease immunogenicity.[2] The increased structural complexity brings numerous analytical challenges for the product/process development and manufacturing groups to develop highly selective, sensitive, reproducible, and robust analytical methods to characterize and determine the impurity profile to ensure drug safety and quality.

A variety of chromatographic techniques for detecting and quantifying product-related impurities in oligonucleotide therapeutics have been reported.[3,4] Among those techniques, ion-pairing reversed-phase liquid chromatography (IPRP-LC) has been established as the main chromatographic technique for the analysis of oligonucleotide therapeutics, especially when connected to mass spectrometry.[5] The addition of IP agents such as triethylamine and others with fluorinated mobile phase modifier such as 1,1,1,3,3,3-hexafluoro-2-propanol has resulted in significant improvements in resolving product-related impurities for therapeutic oligonucleotides.[6] However, chromatographic approaches alone still suffer from an inherent lack of selectivity. Recently, Rentel et al. reported an IPRP-LC method coupled with nominal-mass mass spectrometry for identification and quantification of co-eluting impurities.[7] This method demonstrated sufficient selectivity for impurities that are at least 4 Da apart from the full-length product (FLP), and the same group also published a separate high-resolution mass spectrometric method for analysis of impurities that are less than 4 Da apart from the FLP.[8] In both methods, impurity identification was based on full MS data and *m/z* value at a specific charge state.

Herein, we report the use of BioPharma Finder 5.1 software to process ddMS^2 data collected on the Orbitrap Exploris 240 mass spectrometer for comprehensive sequence characterization of modified RNA samples with site-specific localization of modifications, which was not possible using full MS data. The sequences were confirmed by matching the experimental with the predicted MS/MS spectra and fragment coverage maps to ensure high confidence in the identified sequences. We then used Chromeleon CDS 7.3.1 software for targeted quantitation of the impurities with confirmed sequences using both full MS and UV data collected on the Orbitrap Exploris 240 mass spectrometer and the Orbitrap Exploris MX detector. The Chromeleon eWorkflow procedure was used for seamless execution of the impurity analysis on one Orbitrap Exploris 240 mass spectrometer and two Orbitrap Exploris MX mass detectors coupled with Vanquish UHPLC systems. We were able to achieve consistent impurity profiles with similar relative quantitation of the impurities with minimal instrument-to-instrument variation. This streamlined approach of using our Orbitrap-based LC-HRAM-MS platforms can be a valuable tool for characterization and monitoring of therapeutic oligonucleotide impurities.

Experimental

Reagents and consumables

- Four RNA samples (Integrated DNA Technologies, Inc.)
- 1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP), 99.9%, Thermo Scientific™ (P/N AC293410500)
- Triethylamine (TEA), 99%, Thermo Scientific™ (P/N 157911000)
- Water, UHPLC-MS grade, Thermo Scientific™ (P/N W8-1)
- Methanol, UHPLC-MS grade, Thermo Scientific™ (P/N A4581)
- Thermo Scientific™ DNAPac™ RP HPLC column, 2.1 × 250 mm, 4 µm (P/N 303324)
- Thermo Scientific™ 9 mm Screw Thread Vials, Polypropylene, 12 × 32 mm, 400 µL (P/N C4000-11)
- Thermo Scientific™ 9 mm Autosampler Vial Screw Thread Caps, Polypropylene (P/N C5000-50)

RNA sample preparation

Four RNA samples were previously desalted and HPLC purified. The sequence information and theoretical monoisotopic mass output from BioPharma Finder 5.1 software are listed in Table 1. The lyophilized samples were resuspended in UHPLC water to make a final concentration of 10 pmol/µL.

Table 1. RNA sample sequences and theoretical mass output from sequence manager in BioPharma Finder 5.1 software

Sample name	Sequence	Theoretical mass (Da)
Unmodified RNA	5'-Ur-pUr-pGr-pAr-pCr-pAr-pCr-pCr-pAr-pGr-pAr-pGr-pCr-pCr-pAr-pAr-pCr-pUr-pGr-pGr-pUr-pAr-pAr-pUr-pGr-3'	7656.0758
2'F RNA	5'-Ur-pUf-pGr-pAr-pCr-pAr-pCr-pCr-pAr-pGr-pAr-pGr-pCr-pCr-pAr-pAf-pCr-pUr-pGr-pGr-pUr-pAr-pAr-pUr-pGr-3'	7660.0671
2'MOE RNA	5'-Ur-pUr-pGe-pAr-pCr-pAr-pCr-pCr-pAr-pGr-pAr-pGr-pCr-pCr-pAr-pAe-pCd-pUr-pGr-pGr-pUr-pAr-pAr-pUr-pGr-3'	7756.1646
3'Biotin-TEG RNA	5'-Ur-pUr-pGr-pAr-pCr-pAr-pCr-pCr-pAr-pGr-pAr-pGr-pCr-pCr-pAr-pAr-pCr-pUr-pGr-pGr-pUr-pAr-pAr-pUr-pGr-3'-x	8225.2930

Sequence annotation:
A/C/G/U: Standard nucleotide bases (adenine/cytosine/guanine/uracil); p: phosphate backbone; r: ribose; d: deoxyribose; f: 2' ribose fluorinated (+1.996 Da); e: 2' ribose 0-methoxyethylated (+58.042 Da); x: 3' biotinylated (+569.228 Da)

Chromatography

The Thermo Scientific™ Vanquish™ Horizon and Thermo Scientific™ Vanquish™ Flex UHPLC systems were used for the applied gradient detailed in Table 2. The modules included in the system are listed in Table 3. Ten microliters, or a total of 100 pmol of RNA sample, were injected onto the DNAPac column for all characterization and monitoring experiments. The systems were set up to collect both UV and MS data in a single run. UV data were acquired at 260 nm with a sampling rate of 20 Hz.

Mass spectrometry

For characterization experiments, a ddMS^2 method with a cycle time of 1 s was developed on the Orbitrap Exploris 240 mass spectrometer. For impurity analysis, a full MS method was developed on the Orbitrap Exploris 240 mass spectrometer and also executed on the Orbitrap Exploris MX mass detector. Detailed instrument methods and source parameters for both MS systems are summarized in Table 4.

Injection sequences

Two injection sequences were created in this study. Sequence #1 consisted of one injection using the ddMS^2 method followed by three replicate injections using the full MS method for each RNA sample. In total, 16 injections for four RNAs were created on the Orbitrap Exploris 240 mass spectrometer. Sequence #2 consisted of just the 12 total injections (i.e., three for each RNA) using the full MS method, and this sequence was used for relative quantitation of FLP and impurities on the Orbitrap Exploris 240 mass spectrometer and Orbitrap Exploris MX mass detector.

Table 2. LC and autosampler conditions

Parameter	Value
UHPLC column	DNAPac RP 2.1 × 250 mm, 4 µm (P/N 303324)
Flow rate	0.3 mL/min
Solvent A	Water with 30 mM TEA and 100 mM HFIP
Solvent B	50% methanol with 30 mM TEA and 100 mM HFIP
Gradient	Time (min) %B 0.0 5 1.0 5 42.0 25 45.0 90 60.0 90 60.5 5 75.0 5
Injection volume	10 µL
Needle wash solution	10% methanol
Seal rinse solution	10% methanol (Vanquish Flex) or 75% isopropanol with 0.1% formic acid (Vanquish Horizon)
Autosampler temperature	6 °C
Thermostatting mode	Still air
Column oven temperature	50 °C

Table 3. Vanquish Horizon and Vanquish Flex UHPLC system modules and part numbers

Modules	Vanquish Flex (P/N)	Vanquish Horizon (P/N)
Vanquish System Base F/H	VF-S01-A-02	VF-S01-A-02
Vanquish Binary Pump F	VF-P10-A-01	VH-P10-A-01
Vanquish Split Sampler FT	VF-A10-A-02	VH-A10-A-02
Vanquish Column Compartment H	VF-C10-A-03	VH-C10-A-03
Vanquish Variable Wavelength Detector F	VF-D40-A	VF-D40-A
Vanquish Variable Wavelength - Semi-Micro Bio Flow Cell, 2.5 µL, 7 mm, 50 bar UV cell	6077.0300	6077.0300

Table 4. Instrument method and source parameters for the Orbitrap Exploris 240 mass spectrometer and Orbitrap Exploris MX mass detector. Unless otherwise indicated, default parameters were used.

Instrument	Orbitrap Exploris 240 mass spectrometer	Orbitrap Exploris MX mass detector
MS source parameters		
Negative ion (V)	2,500	Same as Orbitrap Exploris 240 mass spectrometer
Sheath gas (Arb)	35	
Aux gas (Arb)	10	
Sweep gas (Arb)	0	
Ion transfer tube temperature (ºC)	300	
Vaporizer temperature (ºC)	150	
Full scan parameters		
Expected LC peak width (s)	6	Same as Orbitrap Exploris 240 mass spectrometer
Resolution	120,000	
Scan range (*m/z*)	420–1,600	
Time range (min)	1–50	
AGC target	1E6	
Maximum injection time (ms)	50	
Data-dependent MS^2 scan parameters		
MIPS filter		Not applicable
Monoisotopic peak determination	Peptide	
Intensity filter		
Intensity threshold (counts)	5.0e4	
Charge states filter		
Include charge state(s)	2–50	
Dynamic exclusion filter		
Dynamic exclusion mode	Custom	
Exclusion duration (s)	6	
$ddMS^2$ scan		
Isolation window (*m/z*)	2	
HCD collision energy (%)	17, 20, 23	
Resolution	30,000	
AGC target	3e5	
Maximum injection time (ms)	100	
Microscans	3	

Oligonucleotide characterization using BioPharma Finder 5.1 software

Sequence creation for native RNA and modified RNA

RNA sequences were created in the sequence manager. For the 2'F RNA sequence, a replacement of OH with a fluorine (F) on the 2' ribose was created, labeled as f (Figure 1). With fluorination added to the 2' ribose of the second uridine (U2) and adenosine at site 15 (A15), the 2'F RNA sequence resulted in a theoretical monoisotopic mass of 7660.0671 as shown in Table 1. Similarly, the 2'-O-2-methoxyethyl group (2'MOE) was created (labeled as e) and added to the 2' ribose of the third guanosine (G3) and adenosine at the 15th location (A15). In addition, the 2' ribose of the cytosine at position 16 was converted to a 2' deoxyribose (labeled as d), which resulted in a theoretical monoisotopic mass of 7756.165 for the 2'MOE RNA sequence. For 3'Biotin-TEG RNA, a biotin triethylene glycol (Biotin-TEG) linker was created (labeled as x) with a chemical formula of $C_{22}H_{42}N_3O_{11}PS$ and added to the 3' terminal of the native RNA sequence. This resulted in a theoretical monoisotopic mass of 8225.293 for 3'Biotin-TEG RNA sequence.

Assigning variable modifications for native RNA and modified RNA

For all RNA sequences, dephosphorylation was selected for 5' terminal modification; phosphorylation was selected for 3' terminal modification; and base loss (A/U/G/C), deamination (A/C), oxidation (A/U/G/C), depurination (A/G), depyrimidination (C/U), defluorination (f), RNA triple loss and triple add (pAr, pUr, pGr, pCr) were selected for oligo variable modification. A maximum of one modification is allowed for sequence identification.

Identifying and mapping of oligonucleotide sequences

The identification and mapping of oligonucleotides was performed in the Oligonucleotide Analysis workflow within BioPharma Finder software. The processing method for searching $ddMS^2$ data was based on the "Basic Default Method" with the following changes to the identification parameters: set find for all ions in the run, set S/N threshold to ten for component detection, increased the maximum oligonucleotide mass to 20,000, and set mass accuracy to 10 ppm for identification. Each $ddMS^2$ raw data set was processed using this method against its target sequence created above.

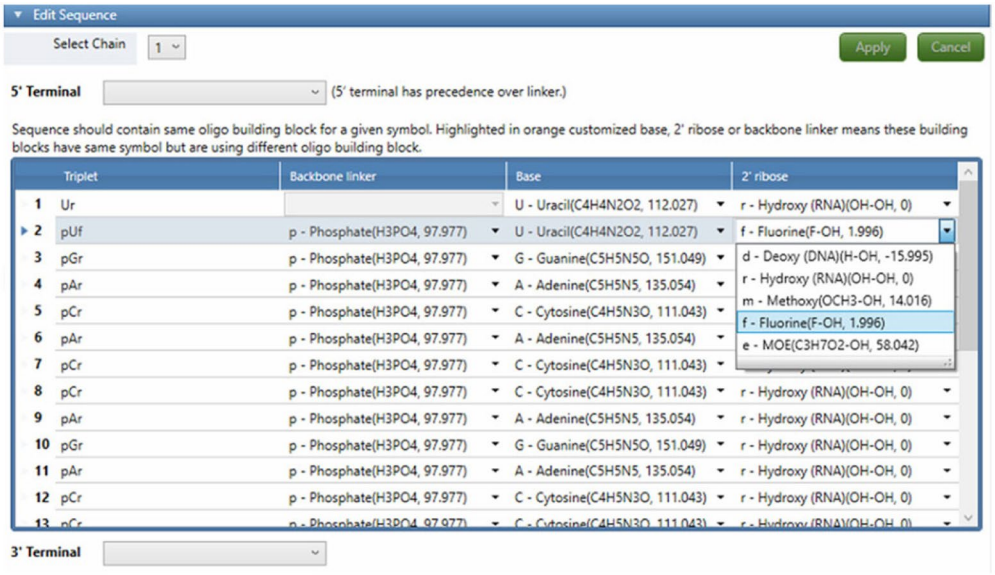
Exporting component results in Chromeleon workbook

Detected charge states were filtered by using the following criteria:

- Delta ppm less than 10 ppm
- Confidence score above 80
- Average structural resolution (ASR) value less than 1.5, but not 0

The filtered results were exported in a Chromeleon workbook with the top five isotopes for each charge states, with one workbook for each RNA sample.

A



B

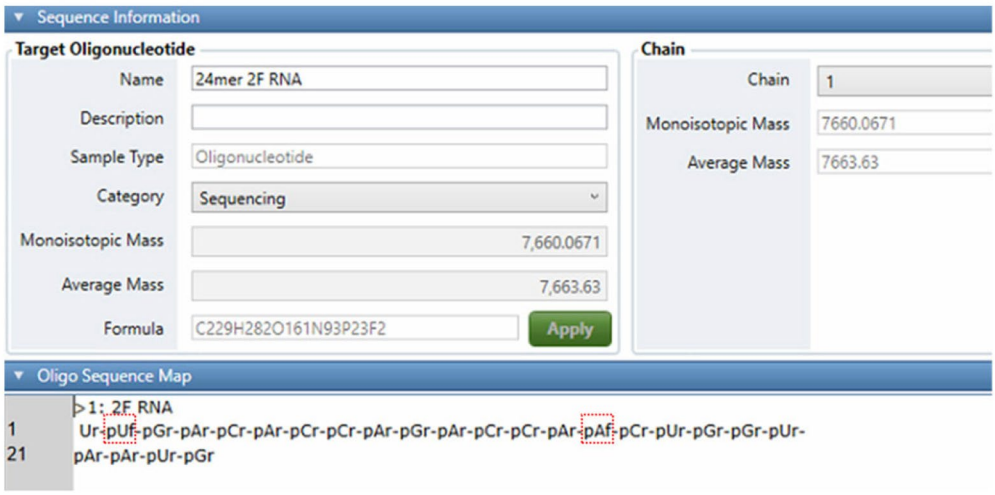


Figure 1. Creating the 2′F RNA sequence in BioPharma Finder 5.1 software. (A) Modifying the 2′ ribose of the second uridine by replacing the 2′OH with 2′F. (B) Determination of theoretical monoisotopic and average mass output based on the input 2′F RNA sequence. The created sequence also highlighted two fluorination sites at U2 and A15.

Oligonucleotide and impurity monitoring using Chromeleon CDS 7.3.1

A Chromeleon MS processing method was created for each RNA sample. The filtered charge states were imported from the workbook to the MS component table of the respective processing method, the RT of the individual charge state was adjusted, and the peak integration parameters were optimized to ensure accurate component detection and consistent peak integration across datasets. For each processing method, the following settings were applied: MS detection algorithm ICIS, manually defined mass tolerance 10 ppm, inhibit integration for TIC channel, Gaussian smoothing five points.

Oligonucleotide impurity report

An oligonucleotide impurity report was created to monitor and quantify the impurities using both UV and full MS data. For UV data, peak areas of detected components were integrated using the default Cobra detection algorithm. Relative quantitation of each impurity was calculated by taking the peak area of that component divided by the peak area sum of all components. For MS data, peak areas for detected charge states were integrated using the ICIS detection algorithm. Relative quantitation of each impurity was calculated by taking the peak area sum of all associated charge states, including the top five isotopes per charge states, and dividing by the peak area sum of all detected impurities.

Chromeleon eWorkflow procedure

A Chromeleon eWorkflow procedure was created for oligonucleotide impurity analysis and executed on an Orbitrap Exploris 240 mass spectrometer and two Orbitrap Exploris MX mass detectors. This eWorkflow procedure bundles the full MS instrument method, processing method, an injection sequence #2, view setting, and impurity report that were previously developed on the Orbitrap Exploris 240 mass spectrometer.

Software

Chromeleon 7.3.1 CDS software was used for all data acquisition, relative quantitation of impurities, and reporting. BioPharma Finder 5.1 software was used for oligonucleotide sequence creation, sequence identification and mapping, and exporting components in a Chromeleon workbook for MS processing.

Results and discussion

Chromatographic separation

Previously, we have shown the combination of the Vanquish Flex system or Horizon UHPLC system and a DNAPac RP column can deliver high-resolution separation of single strand DNAs up to 40mer,[9] and double-stranded siRNA and its diastereomers.[10] In this study, rather than optimizing the gradient for each RNA sample, we developed a general gradient that could separate the impurities from each RNA sample. Excellent reproducibility was observed for the separation of 2′F RNA and 2′MOE RNA as shown in Figure 2. For each RNA sample, three UV traces from replicate injections were overlaid, and the impurities were identified by the delta mass difference between the experimental and theoretical monoisotopic mass, and sequence confirmed by matching the experimental against predicted MS^2 spectrum. Complete characterization of these impurities using $ddMS^2$ data, processed by Biopharma Finder software, will be discussed in the following section.

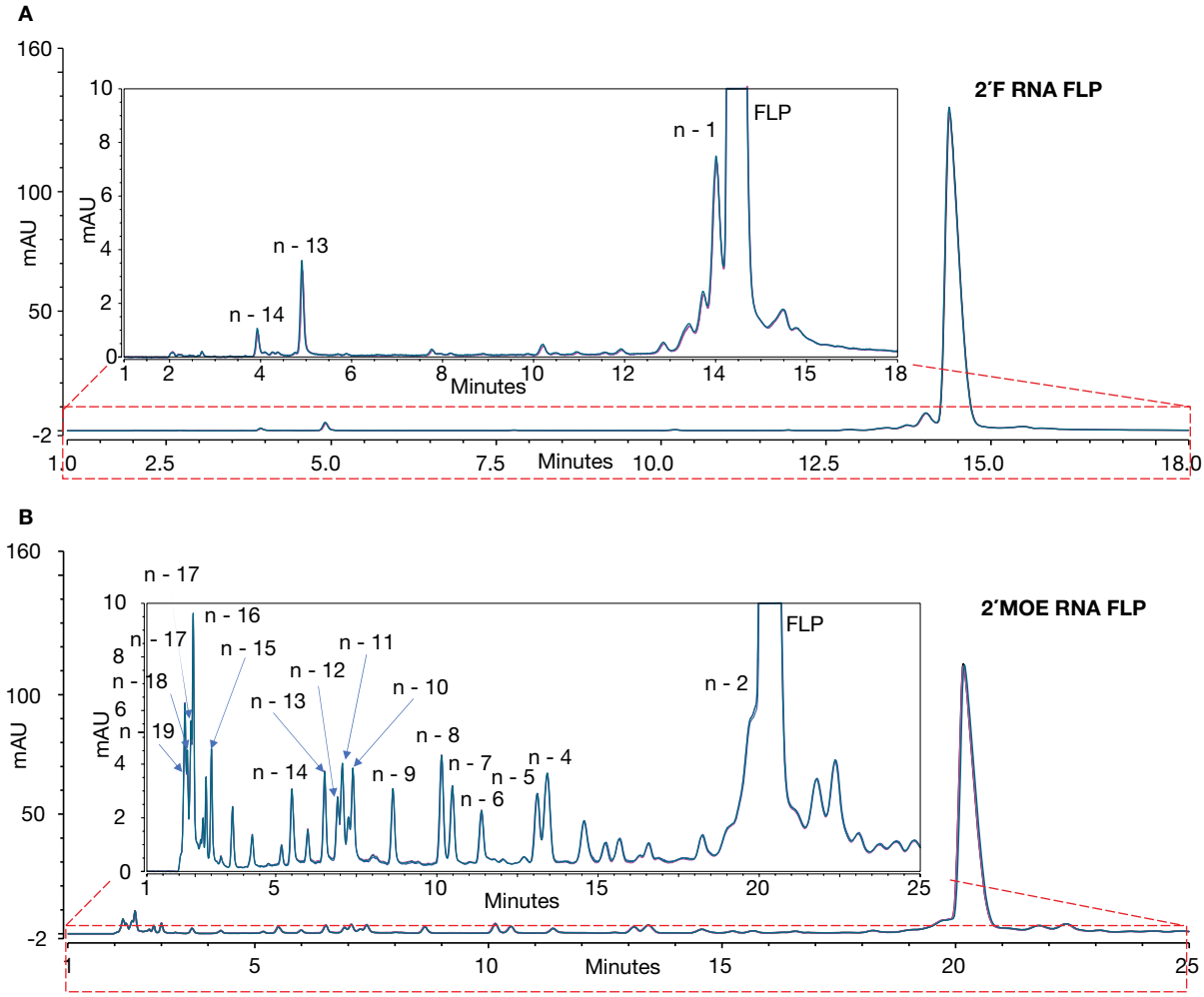


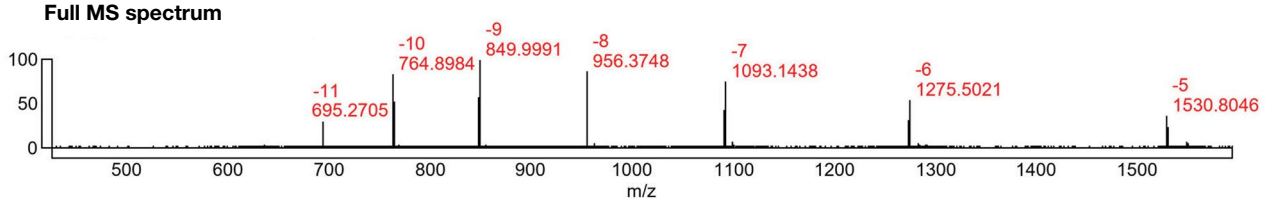
Figure 2. IPRP-LC separation of RNA samples and impurities using a DNAPac RP HPLC column. Three UV traces for the separation of (A) 2′F RNA and (B) 2′MOE RNA were overlaid. Only those identified impurities that had confirmed sequences were labeled in the zoomed-in insets. Impurities were represented as n-x where x is the number of truncated nucleotides from the 5' end.

Oligonucleotide characterization

For characterization of FLP and impurities, ddMS^2 data were collected with stepped normalized collision energy (NCE) for optimizing the fragmentation of both FLP and impurities. Similar to the previously described strategy,[11] we varied the stepped NCE to find a balance between sufficient fragmentation to produce intense a/b/c/w/y/z ion pairs while minimizing the formation of non-unique internal fragments. It was found that stepped NCE of 17, 20, 23 was the optimal NCE setting for fragmenting both FLP and impurities. With this NCE, we were able to achieve a 100% confidence score with less than 4.0 delta ppm between the measured and theoretical monoisotopic mass for all RNA samples (Figure 3). In addition, this optimized NCE generated an average structural resolution (ASR) of 1.0 for all detected charge states (Figure 3) for each FLP. This combination of high confidence score with low delta ppm and low ASR value provided absolute confidence for the identified impurity.

A. FLP characterization for unmodified RNA

Full MS spectrum

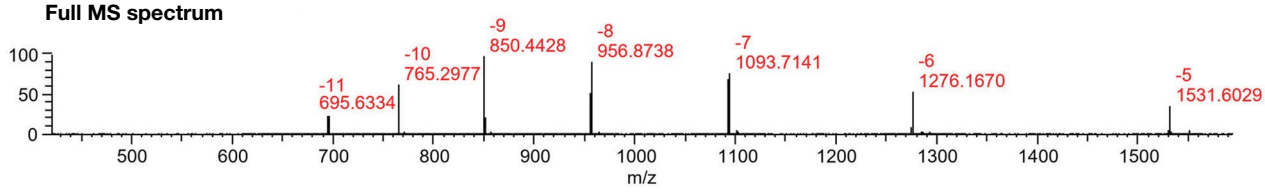


Results

	Identification	Mod	Site	M/Z	Charge St.	Mono Mass Exp.	Avg Mass Exp.	Theor. Mass	Δ ppm	Conf. Score	Best ASR	RT	MS Area
		Aa	Aa no...	Aa	=	=	=	=	≤ 10	≥ 80	=	=	=
1	1:U1-G24 = 7656.076m	None		849.999	-9	7656.0571	7659.43	7656.0758	-2.44	100.0	1.0	11.66	25,640,566.00
2	1:U1-G24 = 7656.076m	None		956.375	-8	7656.0566	7659.40	7656.0758	-2.50	100.0	1.0	11.66	22,595,164.00
3	1:U1-G24 = 7656.076m	None		764.898	-10	7656.0571	7659.41	7656.0758	-2.44	100.0	1.0	11.67	21,241,718.00
4	1:U1-G24 = 7656.076m	None		1093.144	-7	7656.0576	7659.44	7656.0758	-2.37	100.0	1.0	11.65	20,183,250.00
5	1:U1-G24 = 7656.076m	None		1275.502	-6	7656.0518	7659.43	7656.0758	-3.14	100.0	1.0	11.65	14,258,628.00
6	1:U1-G24 = 7656.076m	None		1530.805	-5	7656.0557	7659.43	7656.0758	-2.63	100.0	1.0	11.65	9,576,778.00
7	1:U1-G24 = 7656.076m	None		695.271	-11	7656.0532	7659.42	7656.0758	-2.95	100.0	1.0	11.67	7,483,665.00
8	1:U1-G24 = 7656.076m	None		637.331	-12	7656.0488	7659.70	7656.0758	-3.52	100.0	1.0	11.68	1,215,254.62

B. 2′F RNA

Full MS spectrum



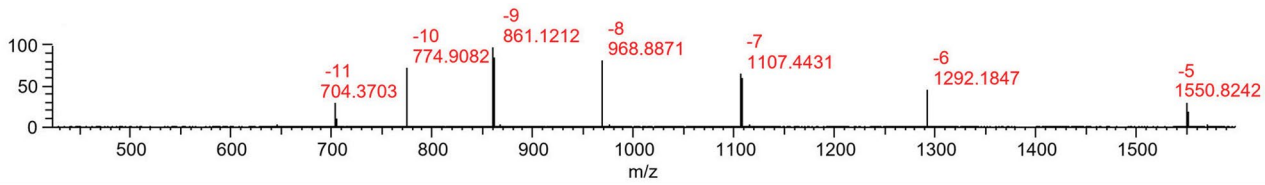
Results

	Identification	Mod	Site	M/Z	Charge St.	Mono Mass Exp.	Avg Mass Exp.	Theor. Mass	Δ ppm	Conf. Score	Best ASR	RT	MS Area
		Aa	Aa	=	=	=	=	=	=	≥ 8.	= (...	=	=
1	1:U1-G24 = 7660.067m	None		850.443	-9	7660.0498	7663.38	7660.0671	-2.26	100.0	1.0	14.38	16,669,580.00
2	1:U1-G24 = 7660.067m	None		956.874	-8	7660.0488	7663.41	7660.0671	-2.39	100.0	1.0	14.38	14,982,251.00
3	1:U1-G24 = 7660.067m	None		1093.714	-7	7660.0488	7663.41	7660.0671	-2.39	100.0	1.0	14.38	12,701,187.00
4	1:U1-G24 = 7660.067m	None		765.298	-10	7660.0459	7663.40	7660.0671	-2.77	100.0	1.0	14.36	11,268,586.00
5	1:U1-G24 = 7660.067m	None		1276.167	-6	7660.0444	7663.39	7660.0671	-2.96	100.0	1.0	14.37	8,748,044.00
6	1:U1-G24 = 7660.067m	None		1531.603	-5	7660.0444	7663.41	7660.0671	-2.96	100.0	1.0	14.37	5,756,438.50
7	1:U1-G24 = 7660.067m	None		695.633	-11	7660.0425	7663.39	7660.0671	-3.22	100.0	1.0	14.36	4,506,887.00
8	1:U1-G24 = 7660.067m	None		637.496	-12	7660.0381	7663.72	7660.0671	-3.79	100.0	1.0	14.36	668,528.25

Figure 3 (panels A and B). Characterization of FLP for all RNA samples. For each RNA sample, the deconvoluted MS spectrum is shown on the top; the source full MS spectrum with all the detected charge states is shown in the middle; and the component detection results containing identification, calculated delta ppm, confidence score, ASR value, retention time, and MS area, for each detected charge state with respective *m/z* value, are shown in the bottom.

C. 2′MOE RNA

Full MS spectrum

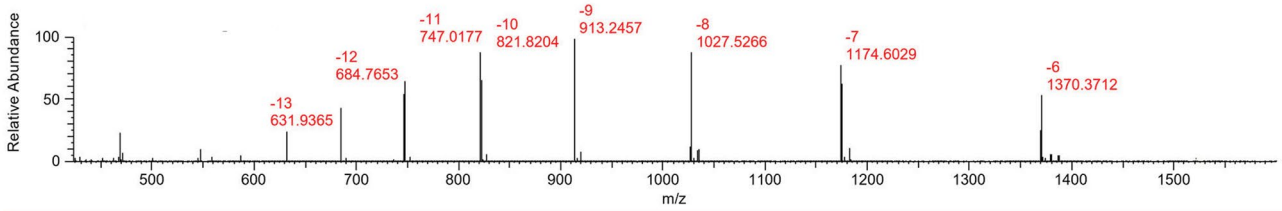


Results

	Identification	Mod	Site	M/Z	Charge St.	Mono Mass Exp.	Avg Mass Exp.	Theor. Mass	Δ ppm	Conf. Score	Best ASR	RT	MS Area
	Aa	Aa no...	Aa	=	=	=	=	=	≤ 10	≥ 80	= (...	=	=
1	1:U1-G24 = 7756.165m	None		861.121	-9	7756.1572	7759.54	7756.1646	-0.95	100.0	1.0	19.33	21,354,794.00
2	1:U1-G24 = 7756.165m	None		968.887	-8	7756.1538	7759.58	7756.1646	-1.39	100.0	1.0	19.34	18,432,166.00
3	1:U1-G24 = 7756.165m	None		774.908	-10	7756.1533	7759.54	7756.1646	-1.45	100.0	1.0	19.33	15,760,586.00
4	1:U1-G24 = 7756.165m	None		1107.443	-7	7756.1528	7759.57	7756.1646	-1.52	100.0	1.0	19.34	14,831,593.00
5	1:U1-G24 = 7756.165m	None		1292.185	-6	7756.1484	7759.57	7756.1646	-2.08	100.0	1.0	19.34	10,155,108.00
6	1:U1-G24 = 7756.165m	None		704.370	-11	7756.1494	7759.56	7756.1646	-1.96	100.0	1.0	19.33	6,686,354.50
7	1:U1-G24 = 7756.165m	None		1550.824	-5	7756.1523	7759.57	7756.1646	-1.58	100.0	1.0	19.33	6,649,672.50
8	1:U1-G24 = 7756.165m	None		645.589	-12	7756.1440	7759.61	7756.1646	-2.65	100.0	1.0	19.32	781,771.56

D. 3′Biotin-TEG RNA

Full MS spectrum



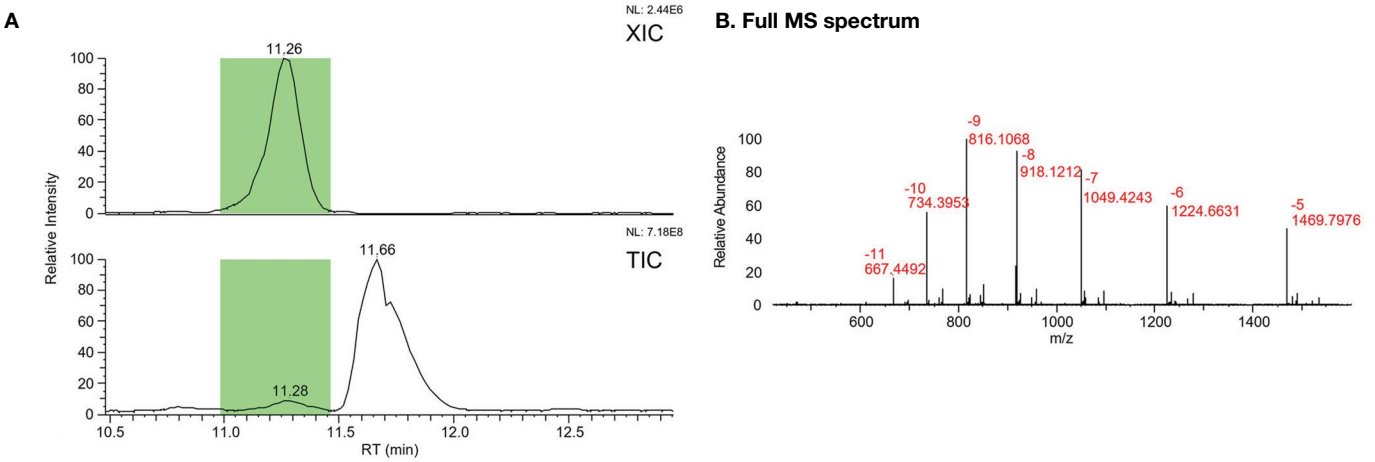
Results

	Identification	Mod	Site	M/Z	Charge St.	Mono Mass Exp.	Avg Mass Exp.	Theor. Mass	Δ ppm	Conf. Score	Best ASR	RT	MS Area
	Aa	Aa	Aa	=	=	=	=	=	≤ 10	≥ 80	=	≥	=
1	1:U1-G24 = 8225.293m	(BioTEG)	(G24)	913.246	-9	8225.2754	8229.14	8225.2932	-2.17	100.0	1.0	34.53	19,835,790.00
2	1:U1-G24 = 8225.293m	(BioTEG)	(G24)	1027.652	-8	8225.2666	8229.15	8225.2932	-3.24	100.0	1.0	34.53	18,006,032.00
3	1:U1-G24 = 8225.293m	(BioTEG)	(G24)	821.820	-10	8225.2734	8229.19	8225.2932	-2.41	100.0	1.0	34.54	17,756,728.00
4	1:U1-G24 = 8225.293m	(BioTEG)	(G24)	1174.460	-7	8225.2656	8229.07	8225.2932	-3.36	100.0	1.0	34.55	16,631,451.00
5	1:U1-G24 = 8225.293m	(BioTEG)	(G24)	747.018	-11	8225.2686	8229.15	8225.2932	-3.00	100.0	1.0	34.53	13,048,383.00
6	1:U1-G24 = 8225.293m	(BioTEG)	(G24)	1370.371	-6	8225.2656	8229.18	8225.2932	-3.36	100.0	1.0	34.55	11,259,548.00
7	1:U1-G24 = 8225.293m	(BioTEG)	(G24)	684.765	-12	8225.2656	8229.17	8225.2932	-3.36	100.0	1.0	34.52	9,166,661.00
8	1:U1-G24 = 8225.293m	(BioTEG)	(G24)	631.937	-13	8225.2617	8229.04	8225.2932	-3.83	100.0	1.0	34.53	4,630,277.50
9	1:U1-G24 = 8225.293m	(BioTEG)	(G24)	586.655	-14	8225.2627	8228.82	8225.2932	-3.71	100.0	1.2	34.54	922,604.00

Figure 3 (panels C and D). Characterization of FLP for all RNA samples. For each RNA sample, the deconvoluted MS spectrum is shown on the top; the source full MS spectrum with all the detected charge states is shown in the middle; and the component detection results containing identification, calculated delta ppm, confidence score, ASR value, retention time, and MS area, for each detected charge state with respective *m/z* value, are shown in the bottom.

While full MS-based methods are the most popular methods for characterization of various oligonucleotide impurities and degradation products,[3,4] these methods only provide intact mass confirmation and do not provide base-by-base sequence information and site-specific modifications. Sequence mapping with direct site localization of these modifications, such as truncation in the middle of a sequence, oxidation, and deamination, are only possible with targeted MS/MS sequencing. In this work, ddMS^2 raw data for each RNA sample was searched against the target sequence with the list of variable modifications assigned above. In addition to common 5′ end truncated impurities (e.g., n – x with x ranging from 1 to 23), we found several other modifications.

For instance, Figure 4 shows the detection of loss of cytosine likely at location 13 (e.g., ~ C13) for native RNA. This impurity was present at ~8% relative to the FLP estimated from the total ion chromatogram of the run (Figure 4A), and we were able to detect six charge states with mass accuracy of less than 2 ppm for deconvoluted monoisotopic mass, confidence score of 99.9, and ASR of 1.0 (Figure 4B, 4C, 4D). We also identified oxidation on uridine at site 17 (U17) for 2′MOE RNA, which was present at ~1% relative to the FLP, and oxidation on uridine at site 20 (U20) for 3'Biotin-TEG, which was present at ~3% relative to the FLP (data not shown). As illustrated in Figure 4D, the ability to perform spectra matching between the experimental and predicted MS/MS fragmentation is invaluable to providing confident identification and sequencing of the impurity. Without MS^2 information, the true identification of these impurities cannot be determined due to the prevalence of many possible identities that have the same accurate mass.



C

Results

	Identification	Mod	Site	M/Z	Charge St.	Mono Mass Exp.	Avg Mass Exp.	Theor. Mass	Δ ppm	Conf. Score	Best ASR	RT	MS Area
51	1:U1-G24 = 7351.0347m(~C13+RNA_Cytosine triple loss)	RNA_Cytosine triple loss	~C13	1469.797	-5	7351.0225	7354.13	7351.0345	-1.64	99.9	1.0	11.27	414,045.41
52	1:U1-G24 = 7351.0347m(~C13+RNA_Cytosine triple loss)	RNA_Cytosine triple loss	~C13	1049.424	-7	0.0000	7354.12	7351.0347	0.00	99.9	1.0	11.27	753,516.56
53	1:U1-G24 = 7351.0347m(~C13+RNA_Cytosine triple loss)	RNA_Cytosine triple loss	~C13	1224.663	-6	7351.0215	7354.13	7351.0345	-1.77	99.9	1.0	11.27	553,913.81
54	1:U1-G24 = 7351.0347m(~C13+RNA_Cytosine triple loss)	RNA_Cytosine triple loss	~C13	918.121	-8	7351.0283	7354.17	7351.0345	-0.84	99.9	1.0	11.27	815,708.62
55	1:U1-G24 = 7351.0347m(~C13+RNA_Cytosine triple loss)	RNA_Cytosine triple loss	~C13	734.395	-10	7351.0225	7354.17	7351.0345	-1.64	99.9	1.0	11.27	457,254.94
56	1:U1-G24 = 7351.0347m(~C13+RNA_Cytosine triple loss)	RNA_Cytosine triple loss	~C13	816.107	-9	7351.0273	7354.12	7351.0345	-0.97	99.9	1.0	11.27	868,029.06
57	1:U1-G24 = 7656.076m(Na+)	Na+		1096.426	-7	0.0000	7681.31	7656.0760	0.00	100.0	1.1	11.64	369,362.44
58	1:U1-G24 = 7656.076m	None		1275.502	-6	7656.0518	7659.43	7656.0758	-3.14	100.0	1.0	11.65	14,258,628.00
59	1:U1-G24 = 7656.076m	None		1093.144	-7	7656.0576	7659.44	7656.0758	-2.37	100.0	1.0	11.65	20,183,250.00
60	1:U1-G24 = 7656.076m	None		1530.805	-5	7656.0557	7659.43	7656.0758	-2.63	100.0	1.0	11.65	9,576,778.00
61	1:U1-G24 = 7656.076m	None		956.375	-8	7656.0566	7659.40	7656.0758	-2.50	100.0	1.0	11.66	22,595,164.00

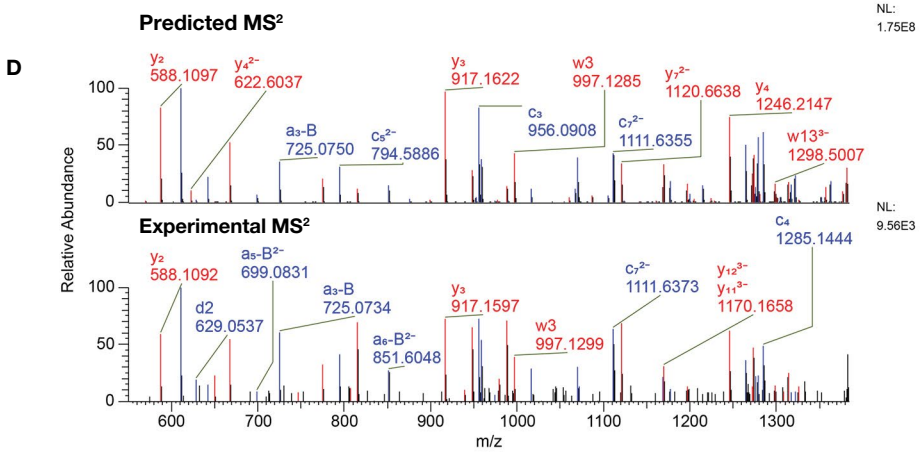


Figure 4. Confident identification of cytosine loss impurity in native RNA by BioPharma Finder 5.1 software. (A) Highlights the detected peak in the extracted ion chromatogram (XIC) and the total ion chromatogram (TIC) for this impurity. (B) Shows the deconvoluted MS spectrum (top) with measured monoisotopic mass with its isotopic distribution of this impurity from the source MS spectrum (bottom). (C) Shows the six detected components that are associated with this impurity; the results also indicated the type of modification with site-specific localization. (D) Shows a comparison of the predicted (top) and measured MS/MS spectra for fragmentation of *m/z* 1469.7996, -5 charge state, with automated peak annotation of matching fragment ions.

Oligonucleotide impurity quantification

Impurities were quantified using both LC-UV data and LC-HRAM-MS data, and the results were compared. When quantifying the impurities using UV signal, this approach often requires full resolution of all impurities from each other and FLP, which is quite challenging considering the number of impurities, and their chemical similarity to each other. On the contrary, quantifying impurities using MS intensities does not require full LC resolution, but rather relies on the ability of HRAM to mass resolve the differences; however, it has been reported that ion suppression could occur for co-eluting impurities, particularly if they are co-eluting with the FLP.[5] Hence, a minimum of near baseline separation between all impurities is required for a fair comparison between the two approaches. As a proof of demonstration, Figure 5 shows the impurity profile for 2′F RNA in both UV and XIC traces. There were three truncated impurities plus the FLP, all of which were identified using BioPharma Finder 5.1 software with high confidence. The impurities were near baseline separated from each other and FLP as shown in Figure 5, and this resulted in similar relative quantitation of the impurities between the two approaches (Table 5).

Moreover, we created an eWorkflow procedure as previously described,[12] containing all the optimized methods, sequence #2, and reports that were developed on the Orbitrap Exploris 240 mass spectrometer for seamless deployment of the impurity analysis on two fit-for-purpose Orbitrap Exploris MX mass detectors. As illustrated in Table 5, using the LC-HRAM-MS approach, excellent precision was achieved across three setups with less than 10% instrument-to-instrument variation. Similar precision was obtained for impurities with relative quantitation above 0.25%, which is near the limit of detection for the LC-UV approach.

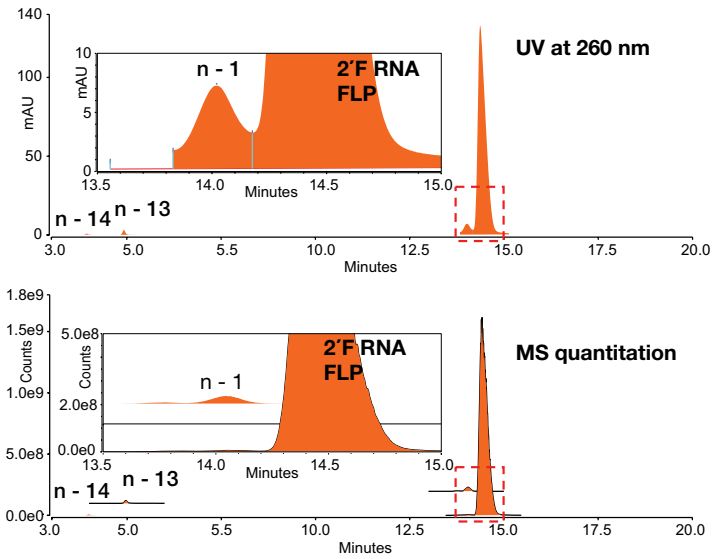


Figure 5. A comparison between LC-UV (top) and LC-HRAM-MS (bottom) profile for the separation of 2′F RNA FLP and impurities

Table 5. A comparison of the relative quantitation of impurities in the 2′F RNA sample across three systems. System #1 is the Vanquish Horizon UHPLC coupled to the Orbitrap Exploris 240 mass spectrometer; System #2 is the Vanquish Horizon UHPLC coupled to the Orbitrap Exploris MX mass detector #1; System #3 is the Vanquish Flex UHPLC coupled to the Orbitrap Exploris MX mass detector #2. Only the average values were reported for % relative quantitation (% relative), and % relative standard deviations (% RSD) were calculated based on three replicate injections.

A. Relative quantitation of impurities in 2′F RNA using LC-UV data at 260 nm wavelength

	System #1		System #2		System #3	
Components	% relative	% RSD	% relative	% RSD	% relative	% RSD
FLP	94.4	0.2	94.8	0.1	94.9	0.1
N-1	4.5	2.0	4.2	1.2	4.2	0.5
N-13	0.83	4.4	0.75	4.4	0.74	3.4
N-14	0.26	26.8	0.20	13.1	0.20	7.0

B. Relative quantitation of impurities in 2′F RNA using LC-HRAM-MS data

	System #1		System #2		System #3	
Components	% relative	% RSD	% relative	% RSD	% relative	% RSD
FLP	97.4	0.1	97.6	0.1	97.7	0.1
N-1	1.7	2.0	1.64	2.7	1.65	4.8
N-13	0.62	8.1	0.54	4.5	0.48	1.1
N-14	0.25	9.1	0.19	5.4	0.16	2.1

Note: For MS data, FLP and truncated impurities were quantified using the peak area sum of detected charge states and isotopes per charge states. For instance, a total of 7 charge states (-5 to -11) were detected for the FLP, the top 5 isotopes per charge states were extracted from TIC, and the sum of the peak area was used for calculating the relative abundance.

Conclusions

We developed a streamlined approach for characterization and monitoring of oligonucleotides. The approach contains the following:

- BioPharma Finder 5.1 software was used to process ddMS^2 data collected on the Orbitrap Exploris 240 mass spectrometer for sequence identification and coverage mapping.
- Identifications of site-specific modifications are enabled via automated MS/MS spectra matching and fragment coverage mapping using ddMS^2 data.
- The instrument methods developed on the Orbitrap Exploris 240 mass spectrometer were included in the Chromeleon eWorkflow procedure and directly transferred to the Orbitrap Exploris MX mass detector without any manual adaptations.
- Full MS data collected on the Orbitrap Exploris 240 mass spectrometer and Orbitrap Exploris MX mass detector provided consistent relative quantitation of impurities.
- Relative quantitation of impurities using full MS data is comparable to the results obtained using UV data.

References

1. Igarashi, J. et al. Research and development of oligonucleotide therapeutics in Japan for rare diseases. *Future Rare Diseases* **2022,** *2*(1). https://doi.org/10.2217/frd-2021-0008
2. Bennett, C.F. Therapeutic antisense oligonucleotides are coming of age. *Ann. Rev. Med.* **2019** *70*, 307–321. https://doi.org/10.1146/annurev-med-041217-010829
3. Sutton, J.M. et al. Current state of oligonucleotide characterization using liquid chromatography-mass spectrometry: insight into critical issues. *J. Am. Soc. Mass Spectrom* **2020,** *31,* 1775.
4. Pourshahian, S. Therapeutic Oligonucleotides, impurities, degradants, and their characterization by mass spectrometry. *Mass Spectrometry Reviews,* **2019.** https://doi.org/10.1002/mas.21615
5. Fountain, K.J.; Gilar, M.; Gebler, J.C. Analysis of native and chemically modified oligonucleotides by tandem ion-pair reversed-phase high-performance liquid chromatography/electrospray ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **2003,** *17,* 646–653.
6. Basiri, B. et al. The role of fluorinated alcohols as mobile phase modifiers for LC-MS analysis of oligonucleotides. *J. Am. Soc. Mass Spectrom.* **2017,** *28*(1), 190–199. https://doi.org/10.1007/s13361-016-1500-3
7. Rentel C. et al. Assay, purity, and impurity profile of phosphorothioate oligonucleotide therapeutics by ion pair–high-performance liquid chromatography–mass spectrometry. *Nucleic Acid Therapeutics* **2022,** *32*(3), 206–220. https://doi.org/10.1089/nat.2021.0056
8. Rentel C. et al. Determination of oligonucleotide deamination by high resolution mass spectrometry. *J. Pharm. Biomed. Anal.* **2019,** *173,* 56–61.
9. Thermo Fisher Scientific, Application Note 21476: Separation of mixed-base oligonucleotides using a high-resolution, reversed-phase chromatography column. https://assets.thermofisher.com/TFS-Assets/CMD/Application-Notes/AN-21476-LC-Mixed-Base-Oligonucleotides-AN21476-EN.pdf
10. Thermo Fisher Scientific, Application Brief 74058: Characterization of a synthetic double stranded siRNA using high-resolution mass spectrometry. https://assets.thermofisher.com/TFS-Assets/CMD/Application-Notes/ab-74058-hram-characterization-double-stranded-srna-ab74058-en.pdf
11. Thermo Fisher Scientific, Application Brief 73789: Oligonucleotide mapping using BioPharma Finder software. http://assets.thermofisher.com/TFS-Assets/CMD/Application-Notes/an-73789-lc-ms-identification-mapping-oligonucleotides-an73789-en.pdf
12. Thermo Fisher Scientific, Application Note 000463: The MAM 2.0 workflow enables seamless transition from research and development to quality control. https://assets.thermofisher.com/TFS-Assets/CMD/Application-Notes/an-000463-lc-ms-mam-workflow-an000463-na-en.pdf

Learn more at **thermofisher.com**