Waters™

응용 자료

A Xevo™ G3-based Workflow for Purity Determination, Intact Mass Measurement, and MS/MS Sequencing of Impurities Detected in Synthetic Oligonucleotides

Catalin E. Doneanu, Jonathan Fox, Brad J. Williams, Chris Knowles, Ying Qing Yu

Waters Corporation

Abstract

This study evaluates how UPLC™-QToF-MS with integrated informatics can facilitate the characterization of synthetic oligonucleotides and their impurities. The combination of oligonucleotide intact mass measurement and MS/MS fragmentation-based sequencing is highly effective at confirming the expected sequence of the Full-Length Product (FLP) and identifying modified variants that can represent potential impurities. The Xevo G3 QTof System is a bench-top High Resolution Mass Spectrometer, which is operated on the compliant-ready informatics platform, waters_connect™, enabling integrated data acquisition and processing. Data independent (MS^E) and targeted MS/MS acquisitions were utilized to characterize the modified oligonucleotides. The holistic analytical and informatics workflow is suitable for both regulated and non-regulated laboratories, enabling both characterization and routine oligonucleotide quality assessments.

Benefits

- The INTACT™ Mass App provides quick and automated mass assignment of the Full-Length Product and

putative identification of oligonucleotide impurities

- The CONFIRM™ Sequence App streamlines the annotation of oligo fragment ions from both DIA (MS^E) and targeted MS/MS data
- Finding sequence omissions and insertions or identifying sites of sequence modification and scrambling
- The Xevo G3 QTof MS paired with the ACQUITY™ Premier UPLC provides a sensitive, robust, and flexible analytical platform for synthetic oligonucleotide analysis capable of being deployed in both regulated and non-regulated laboratories

Introduction

Many oligonucleotide therapeutics (ONTs) are chemically synthesized and exhibit a combination of synthesis artefacts and degradation products broadly classified as product-related impurities. The most popular classes of ONT molecules include antisense oligonucleotides (ASO) and interfering RNA (siRNA oligos) used to modulate gene and protein expressions, and aptamer RNAs that modulate protein and macromolecule functions.[1-2] As synthetic oligonucleotides gain increasing interest as therapeutic drug products, and with a trend to increasing structural complexity, having advanced analytical methodologies for impurity characterization is critical to ensure product safety and efficacy.

Several LC-MS workflows for oligonucleotide analysis have been recently introduced by Waters Corporation.[3-6] These workflows were optimized using a GMP-ready UPLC-Tof MS System that provides a robust and reproducible LC-MS platform, accessible to non-MS experts and supported by streamlined informatics applications that reduce the need for manual data interpretation. Previous studies show that ACQUITY Premier Technologies using Max Peak Technology can improve the chromatographic separation of synthetic oligonucleotides in terms of recovery and reproducibility.[7,8] In addition, the informatics platform used for data processing enables automated peak picking and fast deconvolution of oligonucleotide spectra, providing fast impurity assignments as well as the resulting mass accuracy and quantitative capability to support impurity analysis.[5-6]

The CONFIRM Sequence waters_connect App was introduced for automated processing of both MS/MS and MS^E (data independent) fragmentation spectra for sequence verification.[6] The sequence coverage can be viewed in a "dot-map" diagram to easily assess sequence coverage and to localize sites of modification.

In the studies described here, the Xevo G3 QTof System (Figure 1), along with two waters_connect applications (INTACT Mass and CONFIRM Sequence) were used for analysis of synthetic nucleotide impurities. The ability to detect and confirm low-level impurities was assessed.

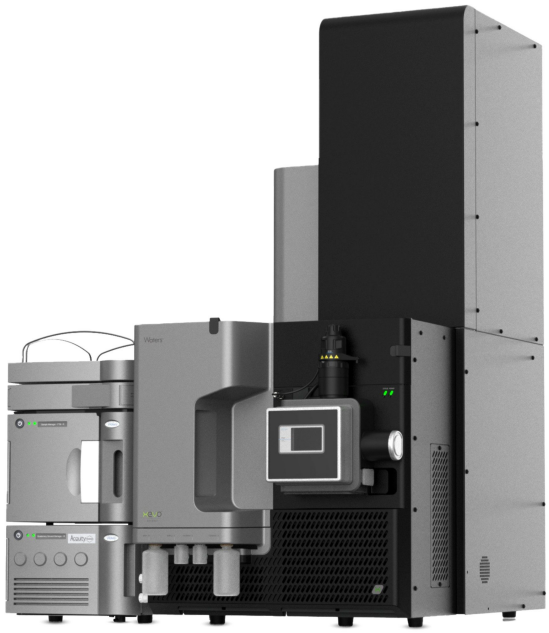

Figure 1. Xevo G3 QTof Mass Spectrometer with the ACQUITY Premier UPLC System and waters_connect Informatics.

Experimental

A 21-mer synthetic siRNA oligonucleotide, containing 13 modified nucleotides, (sequence: GCC UCA GTC TGC TTC GCA CCT), was obtained from ATDBio (Southampton, UK). The oligonucleotide sequence uses colored letters to denote the modified nucleotides, while the regular (unmodified) nucleotides are black. A 2′-O-(2-methoxyethyl) - (2′-MOE) modification was attached to two guanosines (G) and one adenosine (A). Also, in addition to the attachment of the 2′-MOE functional group to one uridine (U) and six cytidines (C), the nucleobases of these two nucleotides were further modified by the attachment of a 5-Methyl (5-Me) group to

produce one 2′-MOE 5-Me uridine (U) and six 2′-MOE 5-Me cytidines (C). Out of the nine cytidine residues present in the 21-mer sequence, three cytidines were modified only at the nucleobase by the addition of a 5-Me functional group (C). Finally, there were eight nucleotides left unmodified: one adenosine (A), three guanosines (G) and five deoxythymidines (T). Stock solutions were prepared in DI water at a concentration of 1 µM (or 2.4 µg/mL), from which a 10 µL volume was injected, which corresponds to loading 10 picomoles of the 21-mer oligonucleotide on-column. Figure 2 displays the structure of the synthetic oligonucleotide (21 mer).

Data acquisition and processing was performed using the waters_connect informatics platform. ESI-MS spectra of intact oligonucleotides were processed automatically using the INTACT Mass App. The oligonucleotide purity assessments were calculated based on UV response as part of the automated processing workflow. Individual MS/MS spectra and MS^E datasets were processed by the CONFIRM Sequence App using the optimum collision energy for fragmentation of each oligonucleotide precursor to sequence the Full-Length Product and product impurities.

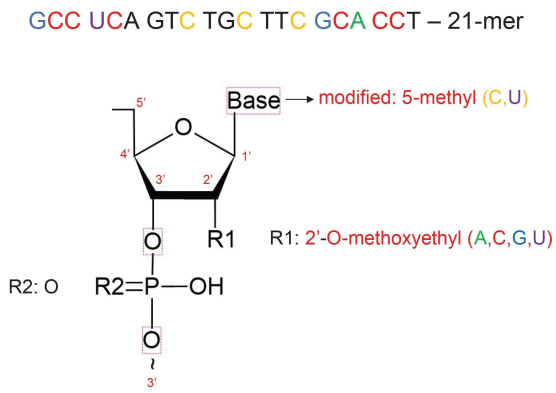

Figure 2. Chemical modifications incorporated in the sequence of the 21-mer siRNA oligonucleotide. Two types of modifications are used for various nucleotides including base modifications for cytidines (C) and uridines (U), as well as 2′-methoxyethyl (2′-O-MOE) for all four nucleotides (A, C, G, U). Eight nucleotides (highlighted in black) out of the twenty-one contain no chemical modifications.

LC Conditions

LC system:	ACQUITY UPLC Premier BSM
Column:	ACQUITY Premier OST 2.1 x 150 mm, packed with 1.7 µm CSH™ C_{18} particles (p/n: 186009486)
Column temperature:	60 °C
Flow rate:	300 µL/min
Mobile phases: Solvent A:	7 mM triethylamine (TEA) and 40 mM hexafluoro-isopropanol (HFIP) in Milli- Q water (pH 8.6)
Solvent B:	3.5 mM TEA and 20 mM HFIP in 50% methanol
Gradient table:	insert Table II
Injection volume:	10 µL
Wash solvents:	Purge solvent: 50% MeOH
Sample manager wash solvent:	50% MeOH

Gradient Table

Time (min)	Flow rate (mL/min)	Solvent A composition (%)	Solvent B composition (%)	Curve profile
0.00	0.3	87	13	Initial
25.00	0.3	77	23	6
30.00	0.3	77	23	6
30.50	0.3	15	85	6
32.50	0.3	15	85	6
33.00	0.3	87	13	6
40.00	0.3	87	13	6

MS Conditions

MS system:	Xevo G3 QTof mass spectrometer
Ionization mode:	ESI(-)
Capillary voltage:	2.8 kV
Cone voltage:	25 V
Source offset:	60 V
Source temperature:	120 °C
Desolvation temperature:	450 °C
Cone gas flow:	50 L/h
Desolvation gas flow:	500 L/hr
TOF mass range:	400-2000 (MS^E acquisition)
Acquisition rate:	0.5 sec
Low energy CE:	6 V
High energy CE ramp:	15 to 45 V
Data acquisition:	waters_connect 3.6.0.21
Data processing:	waters_connect 3.6.0.21
INTACT mass app:	version 1.8.0.10

CONFIRM sequence app:	version 1.4.0.13

Results and Discussion

Data Acquisition using the ACQUITY Premier UPLC with Max Peak Technology System coupled with the Xevo G3 QTof MS System

The siRNA sample (21-mer) used for this study was synthesized using a typical solid phase approach. We previously demonstrated that an ACQUITY Premier UPLC Column, featuring High Performance Surfaces Technology, improves the chromatographic performance for oligonucleotide impurities with reduced need for column conditioning, improved peak shape, and enhanced recovery of low abundance impurities, when compared to a traditional stainless-steel column.[4] Sixteen synthesis-related impurities resolved for this siRNA oligomer are shown in the LC-UV chromatogram displayed in Figure 3. The lowest abundance impurity, according to the UV data, is the first eluting peak with an abundance of 0.05% relative to the peak area of the FLP. The ACQUITY UPLC Premier System used for this impurity assay, is a very inert system which allows for robust separation of trace-level oligonucleotide impurities.

The Xevo G3 QTof MS System has been purposefully developed with capabilities optimized for biopharmaceutical applications: 1) improved ion transmission of low mass labile species; 2) enhanced quantitative capabilities with wide inter and intra-spectrum dynamics range for comprehensive Quan/Qual analysis; 3) new ion optics for improved robustness and reproducibility. When combined with waters_connect informatics, the platform is well suited for product characterization in Development, or for more routine assays in regulated Development, Manufacturing and Quality organizations.

The combined ESI-MS spectrum recorded for the lowest abundance impurity (peak 1 from Figure 3, a 6-mer phosphorylated impurity) is displayed in Figure 4. The corresponding doubly and triply charged precursors are highlighted in this spectrum indicating the ability of the Xevo G3 instrument to detect trace-level impurities.

Sensitive transmission of precursors produced by lower-abundance oligonucleotide impurities ultimately improves the fragmentation through collision-induced-dissociation for both targeted MS/MS or DIA (MS^E) acquisition modes. As shown in Figure 5, a high sequence coverage (83%) was obtained from the MS/MS

fragmentation of the doubly charged precursor (*m/z*=1120.7) of the 5′-phosphorylated 6-mer (identified as peak 1 in Figure 3) when the data was analyzed with the CONFIRM Sequence app. The combination of intact mass measurement and MS/MS sequencing data synergize to provide both semiquantitative data and strong confidence in impurity assignment.

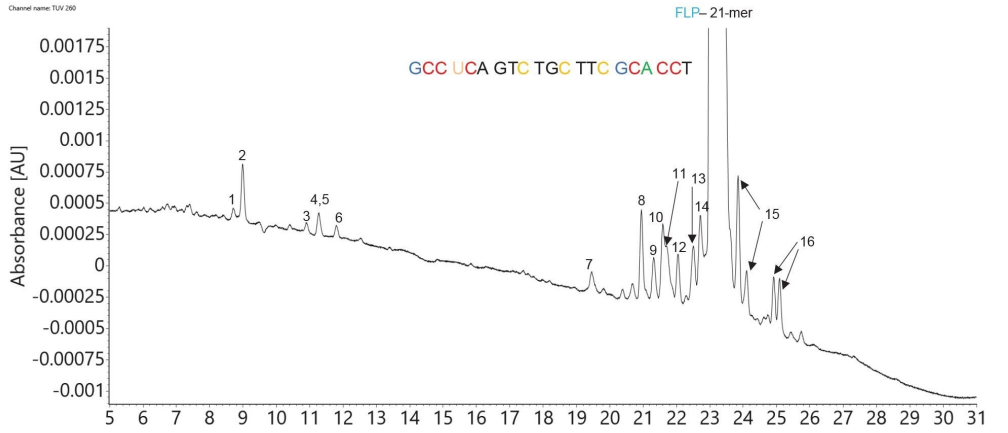

Figure 3. Zoomed-in view of the LC-UV chromatogram recorded on the Acquity UPLC Premier System for the separation of the 21-mer siRNA oligonucleotide and its impurities. Sixteen low-level impurities were detected, the lowest abundance impurity, labeled as peak 1, present at 0.05% relative to the FLP.

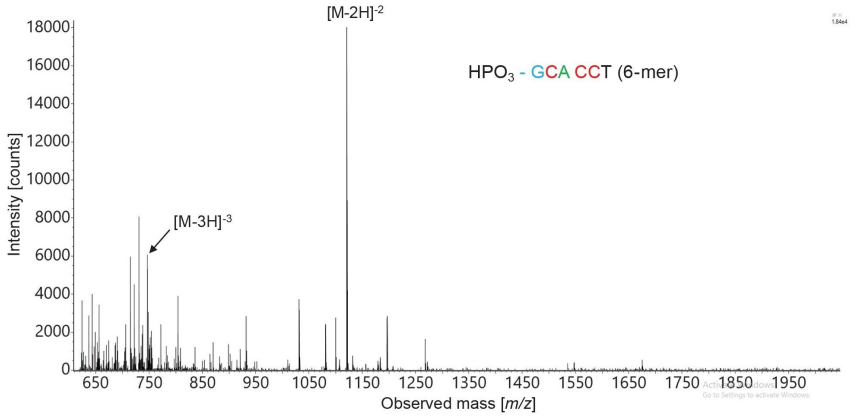

Figure 4. Combined ESI-MS spectrum of a 6-mer 5'-phosphorylated oligonucleotide impurity, which is the least abundant oligonucleotide impurity present in the 21-mer FLP (0.05% relative abundance according to UV peak areas), labeled as peak 1 in the chromatogram shown in Figure 3.

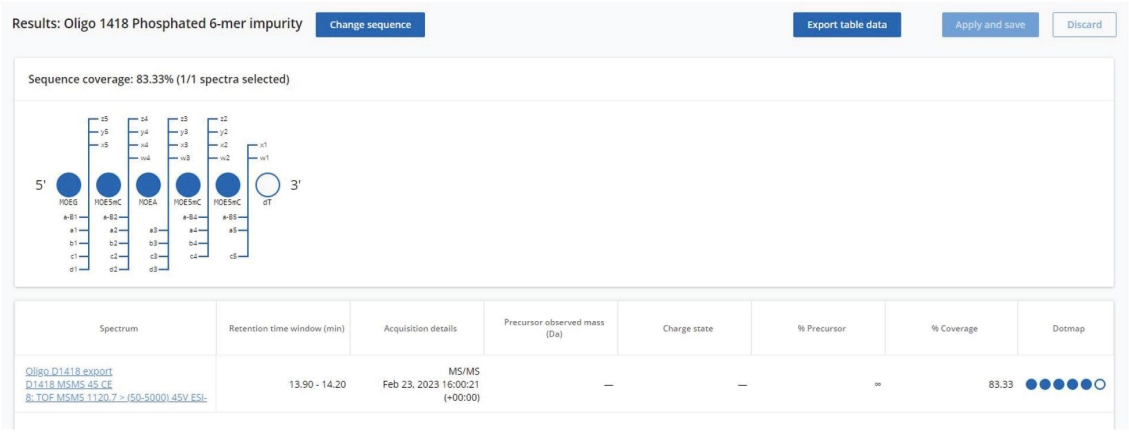

Figure 5. Dot-map sequence coverage (83%) resulting from the MS/MS fragmentation of the [M-2H]$^{-2}$ precursor (m/z=1120.7) of the 5'-phosphorylated 6-mer least abundant oligonucleotide impurity, labeled as peak 1 in Figure 1. The precursor was fragmented with an optimized fixed collision energy (45 V) in the collision cell of the Xevo G3 QTof mass spectrometer.

Automated Mass Assignments with the INTACT Mass app

Calculating the molecular weight and the relative abundance of impurities compared to the FLP peak area (based on the TUV and/or the TIC trace) is performed for each oligo impurity detected above a specified threshold. The INTACT Mass App software was used to process the IP-RP UPLC-UV-MS chromatograms in an untargeted fashion, using automated BayesSpray™ deconvolution to calculate the neutral monoisotopic mass for all components that meet criteria set in the data processing method (RT window and the ion count thresholds).

The untargeted mass analysis can be converted into a targeted analysis simply by updating a target list in the processing method that includes the oligo identification and the expected accurate monoisotopic mass (neutral mass). The INTACT Mass results shown in Figure 6 are presented as three panels, including the deconvolved spectrum (panel A – top), the raw ESI-MS spectrum showing a wide variety of charge states (panel B – middle) and the mock spectrum used for assessing the fidelity of the deconvolution process (panel C – bottom). A portion of the table containing the INTACT Mass processing result is displayed (Figure 7). Out of sixteen impurities (total number of impurities detected in this sample), nine impurities shown in this table along with the FLP, were putatively assigned with mass accuracies better than 10 ppm.

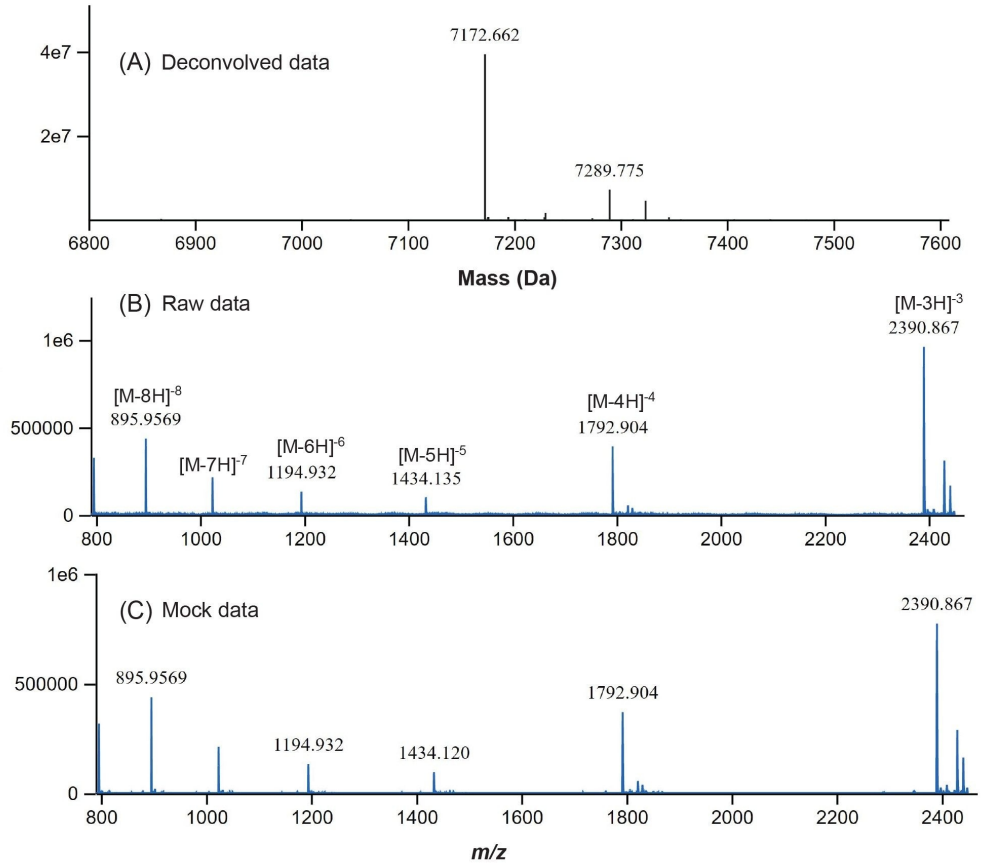

Figure 6. INTACT Mass ESI-MS spectra of the 21-mer heavily modified siRNA oligonucleotide: (A): deconvolved spectrum of the neutral species obtained using the BayesSpray deconvolution algorithm; (B) raw MS spectrum; (C) mock MS spectrum used for assessing the fidelity of the deconvolution process.

Peak no:	Type	Molecule ID	Component	Observed mass (Da)	Expected mass (Da)	Mass error (ppm)	Identity result
7	Impurity	7,172.57	7172.57 2'-MOE-5Me Cytidine,n-dT	6,491.427	6,491.425	0.2	Pass
8	Impurity	7,172.57	7172.57 2'-MOE-5Me Cytidine,5Me-Cytidine	6,492.316	6,492.409	-14.4	Warning
11	Impurity	7,172.57	7172.57 2'-MOE Guanosine	6,769.459	6,769.481	-3.2	Pass
9	Impurity	7,172.57	7172.57 2'-MOE Adenosine	6,785.440	6,785.476	-5.3	Pass
13	Impurity	7,172.57	7172.57 2'-MOE-5Me Uridine	6,794.473	6,794.487	-2.1	Pass
10	Impurity	7,172.57	7172.57 2'-MOE-5Me Cytidine	6,795.455	6,795.471	-2.4	Pass
15	Impurity	7,172.57	7172.57 n-dG	6,843.526	6,843.517	1.3	Pass
12	Impurity	7,172.57	7172.57 n-dT	6,868.558	6,868.524	4.9	Pass
11	Impurity	7,172.57	7172.57 5Me-Cytidine	6,869.450	6,869.508	-8.5	Pass
FLP	Product	7,172.57	7,172.57	7,172.549	7,172.570	-2.9	Pass

Figure 7. Screenshot showing a portion of the processing results table generated by the INTACT Mass App for the 21-mer oligonucleotide and its impurities. The dataset was deconvolved using the BayesSpray charge deconvolution algorithm and 16 oligonucleotide impurities were identified with mass accuracies of under 10 ppm.

Sequence Confirmation of oligonucleotide impurities using the CONFIRM Sequence App

The automated annotation of data independent fragmentation spectra (high energy MS^E) as well as annotation of targeted MS/MS spectra were accomplished using the waters_connect CONFIRM Sequence App. Multiple MS/MS and/or MS^E spectra can be analyzed against one or more oligonucleotide sequences.

The Synthetic Library, which stores the oligonucleotide sequences, was populated with the 21-mer extensively modified siRNA oligonucleotide (Figure 8). The $[M-4H]^{-4}$ precursor was selected for MS/MS fragmentation using the quadrupole and fragmented in the collision cell of a Xevo G3 QTof MS using an ***optimized fixed*** collision voltage (40 V). The result obtained following data processing with the CONFIRM Sequence App - (Figure 9) shows complete (100%) sequence coverage for the FLP. This “dot map” coverage view displays the 21-mer sequence from the 5′ to the 3′ end, with assigned fragment ions displayed above and below the dots (individual nucleotides).

In addition to the FLP, targeted MS/MS was performed on six out of the sixteen oligonucleotide impurities with

TUV peak area at or above 0.05% of the FLP peak area. Data from these analyses are summarized in Table I, including nucleotide sequence, elemental composition, neutral monoisotopic mass, most abundant precursors, optimized collision energies, sequence coverage (from a single MS/MS spectrum as well as from two combined MS/MS spectra) and impurity abundance relative to the FLP (as the TUV area percentage). The lowest abundant impurity, measured at 0.05% was identified as 5′-phosphorylated GCA CCT, according to the ESI-MS and MS/MS spectra shown in Figures 4 and 5. The overall combined sequence coverage resulted from the analysis of MS/MS fragmentation spectra of the doubly and triply charged precursors of this impurity is 89.5%, as indicated in Table I.

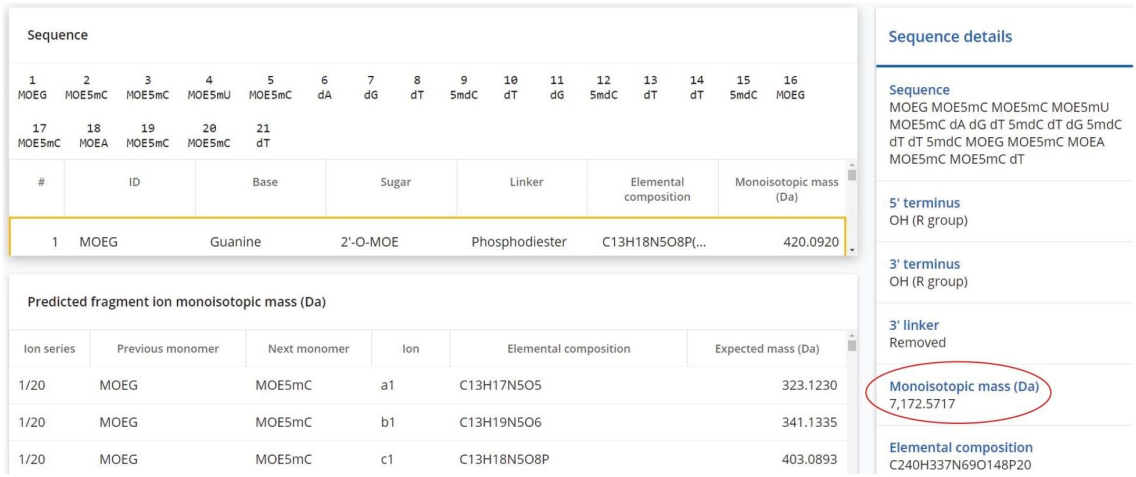

Figure 8. CONFIRM Sequence App screenshot showing the sequence of the 21-mer extensively modified siRNA oligonucleotide as entered in the Synthetic Library.

Figure 9. Complete sequence coverage (100%) resulted from the MS/MS fragmentation of the $[M-4H]^{-4}$ precursor (m/z=1792.14) of the 21-mer siRNA oligonucleotide. The precursor was fragmented with an optimized fixed collision energy (40 V) in the collision cell of a Xevo G3 QTof mass spectrometer.

Peak label	Retention time (min)	Oligonucleotide sequence	Elemental composition	Delta mass (mass difference from FLP)	MW (neutral monoisotopic mass)	Most abundant precursor monoisotopic mass / charge	Second precursor / charge	Optimum CE (V)	MS/MS seq coverage (%)	COMBINED seq coverage (%)	TUV area percentage (%)
1	8.7	5'-phosphorylated GCA CCT	C75 H111 N21 O47 P6	- 4929 Da (loss of GCCUCAGTCTGCTTC)	2243.5367	1120.7611 (-2)	746.8383 (-3)	45 / 26	83.3	89.5	0.05
2	9.0	CA CCT	C_{6}2 H92 N16 O36 P4	- 5412 Da (loss of GCCUCAGTCTGCTTCG)	1760.4811	879.2332 (-2)	1759.4738 (-1)	36 / 70	100.0	100.0	0.27
3	10.9	GCA CCT	C75 H110 N21 O44 P5	- 5009 Da (loss of GCCUCAGTCTGCTTC)	2163.5704	1080.7779 (-2)	720.1828 (-3)	40 / 28	74.2	82.4	0.06
4	11.2	C GCA CCT	C_{5}5 H124 N24 O50 P6	- 4706 Da (loss of GCCUCAGTCTGCTT)	2466.6324	1232.3089 (-2)	821.2053 (-3)	38 / 28	65.1	76.8	0.09
5	11.3	TC GCA CCT	C95 H137 N26 O57 P7	- 4402 Da (loss of GCCUCAGTCTGCT)	2770.6784	1384.3319 (-2)	922.5522 (-3)	38 / 30	65.7	75.2	0.07
6	11.8	TTC GCA CCT	C105 H150 N28 O64 P8	- 4098 Da (loss of GCCUCAGTCTGC)	3074.7245	1536.3496 (-2)	1023.8825 (-3)	38 / 32	68.6	78.3	0.08
7	19.5	GC* UCA G*C TGC TTC GCA CCT	C217 H304 N64 O133 P18	- 681 Da (2'-MOE-5Me C & dT)	6491.4269	2162.8017 (-3)	1621.8495 (-4)	–	–	–	0.14
8	20.9	GC* UCA GTC TGC TT* GCA CCT	C217 H303 N63 O134 P18	- 680 Da (2'-MOE-5Me C & 5MeC)	6492.4109	2163.1297 (-3)	1622.0955 (-4)	–	–	–	0.45
9	21.3	CC UCA GTC TGC TTC GC* CCT	C227 H319 N64 O141 P19	- 387 Da (2'-MOE A)	6785.4774	2260.8119 (-3)	1695.3621 (-4)	–	–	–	0.29
10	21.6	GC* UCA GTC TGC TTC GCA CCT	C229 H321 N64 O140 P19	- 377 Da (2'-MOE-5Me C)	6795.4981	2264.1588 (-3)	1697.8673 (-4)	–	–	–	0.58
11	21.7	*CC UCA GTC TGC TTC GCA CCT	C227 H319 N64 O140 P19	- 403 Da (2'-MOE G)	6769.4824	2255.4869 (-3)	1691.3633 (-4)	–	–	–	0.25
12	22.0	GCC UCA G*C TGC TTC GCA CCT	C230 H324 N67 O141 P19	-304 Da (dT)	6868.5257	2288.5013 (-3)	1716.1242 (-4)	–	–	–	0.32
13	22.5	GCC *CA GTC TGC TTC GCAA CCT	C227 H318 N67 O139 P19	- 378 Da (2'-MOE-5Me U)	6794.4889	2263.8227 (-3)	1697.6149 (-4)	–	–	–	0.45
14	22.7	GCC UCA GTC TGC TT* GCA CCT	C230 H323 N66 O142 P19	-303 Da (5Me C)	6869.5097	2288.8293 (-3)	1716.3702 (-4)	–	–	–	0.65
15	23.1	GCC UCA GTC TGC TTC GCA CCT	C240 H337 N69 O148 P20	–	7172.5717	1792.1357 (-4)	2389.8500 (-3)	40 / 55	100.0	100.0	94.51
16	23.8 / 24.1	GCC UCA *TC TGC TTC GCA CCT	C230 H325 N64 O142 P19	-329 Da (dG)	6843.5192	2280.1658 (-3)	1709.8725 (-4)	–	–	–	1.17
17	24.9 / 25.1	GCC UCA GTC TGC TTC GCA CCT	C240 H342 N69 O151 P20	+53 Da (H5O3)	7225.5956	2407.5246 (-3)	1805.3916 (-4)	–	–	–	0.57
										TOTAL	100.00

Table I. Sixteen oligonucleotide impurities were identified from a 21-mer siRNA modified oligonucleotide. Six impurities and the Full-Length Product (FLP) were sequenced using a Xevo G3 mass spectrometer, individual MS/MS spectra fragmented with optimum collision energies were processed using the CONFIRM Sequence App and high sequence coverage confirming the identity of the FLP (100%) and its impurities were obtained (> 75%). The lowest abundance impurity (0.05% relative abundance) is a 6-mer 5'-phosphorylated oligonucleotide which was sequenced with 83% sequence coverage. The total sequence coverage listed in the table corresponds to the combined sequence obtained from the MS/MS fragmentation of two precursors of each oligonucleotide.

Conclusion

- The ability to identify and confirm low-level impurities is critical to effectively characterize oligonucleotide therapeutics and monitor the quality of product batches.
- The synergy of intact mass analysis and fragmentation-based sequencing, enables holistic impurity profiling and targeted confirmation of impurity assignments, enabling confidence in synthetic production processes for this important class of therapeutics.
- The automated assignment of putative impurities detected in synthesized oligonucleotides with both UV data and the Xevo G3 QTof MS detector was accomplished using the waters_connect INTACT Mass App with typical mass errors below 10 ppm.
- The MS/MS spectra recorded for six oligonucleotide impurities along with the MS/MS spectra recorded for the 21-mer extensively modified FLP were acquired on this same system and processed using the CONFIRM Sequence App to verify their expected sequences.
- The capability of the CONFIRM Sequence App to achieve high sequence coverage (83.3%) of lower abundance impurities (~ 0.05% relative abundance) was confirmed.

References

1. Sharma VK, Watts JK Oligonucleotide therapeutics: chemistry, delivery and clinical progress, *Future Med Chem*, 2015, 7(16), 2221–2242.
2. Roberts TK, Langer R, Wood MJA Advances in oligonucleotide drug delivery, *Nat Reviews*, 2020, 19, 673–694.
3. Catalin E. Doneanu, Jonathan Fox, Emma Harry, Chris Knowles, Ying Qing Yu, Joseph Fredette, Weibin Chen, An Automated Compliance-Ready LC-MS Workflow for Intact Mass Confirmation and Purity Analysis of Oligonucleotides, Waters application note, 720006820, 2020.
4. Catalin E. Doneanu, Chris Knowles, Jonathan Fox, Emma Harry, Ying Qing Yu, Joseph Fredette, Weibin Chen, Analysis of Oligonucleotide Impurities on the BioAccord System with ACQUITY Premier, Waters application note, 720007301, 2021.

5. Catalin E. Doneanu, Patrick Boyce, Henry Shion, Joseph Fredette, Scott J. Berger, Heidi Gastall, Ying Qing Yu, LC-MS Analysis of siRNA, Single Guide RNA and Impurities using the BioAccord System with ACQUITY Premier System and New Automated INTACT Mass Application, Waters application note, 720007546, 2022.

6. Catalin E. Doneanu, Chris Knowles, Matt Gorton, Henry Shion, Joseph Fredette, Ying Qing Yu, CONFIRM Sequence: A waters_connect Application for Sequencing of Synthetic Oligonucleotides and Their Impurities, Waters application note, 720007677, 2022.

7. DeLano M, Walter TH, Lauber MA, Gilar M, Jung MC, Nguyen J, Boissel C, Patel AV, Bates-Harrison A, Wyndham K Using Hybrid Organic-Inorganic Surface Technology to Mitigate Analyte Interactions with Metal Surfaces in UHPLC, *Anal Chem*, 2021, 93, 5773–5781.

8. Gilar M, DeLano M, Gritti F Mitigation of Analyte Loss on Metal Surfaces in liquid chromatography, *J Chrom A*, 2021, 1650, 5773–5781.

Featured Products

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

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

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

720008480, August 2024