AN AUTOMATED WORKFLOW FOR INTACT MASS, PURITY AND SEQUENCE CONFIRMATION OF SYNTHETIC OLIGONUCLEOTIDES AND THEIR IMPURITIES

Catalin Doneanu[1], Chris Knowles[2], Matthew Gorton[2], Joe Fredette[1] and Ying Qing Yu[1]

[1]Waters Corporation, Milford, MA USA; [2]Waters Corporation, Wilmslow, UK

OVERVIEW

Here we are demonstrating an automated, compliance-ready LC-MS workflow for intact mass confirmation , purity analysis and sequence confirmation of synthetic oligonucleotides including small interfering RNA (siRNA) oligos.

INTRODUCTION

- Synthetic oligonucleotides have emerged in recent years as a powerful alternative to small molecule and protein therapeutics [1].
- Manufacturing and quality control of oligonucleotide therapeutics requires highly selective and sensitive LC-MS methods for impurity detection and quantification.
- The most often used mass spectrometry-based method for oligonucleotide analysis has been reversed-phase chromatography employing a variety of ion-pairing reagents and modifiers in negative ESI-MS mode (IP-RP LC-MS).
- An automated workflow for analysis of synthetic oligonucleotides employing the BioAccord™ LC-MS System was recently described [2-5].
- Here we are introducing an automated workflow for oligonucleotide impurity analysis that combines two *waters_connect™* Software applications: INTACT Mass and CONFIRM Sequence. INTACT Mass provides purity information as well as confirmation of intact masses of oligonucleotides and their impurities. CONFIRM Sequence provides sequence coverage information based on the MS/MS fragmentation of oligonucleotides.
- The workflow is supported on TOF and QTof instruments operated under *waters_connect™* Software.

BioAccord™ LC-MS System

METHODS

Sample Preparation

A 21-mer heavily modified oligonucleotide, containing a 2'-OMe modification on 19 of its nucleosides, having the sequence GUA ACC AAG AGU AUU CCA UTT and the elemental composition C229H306N76O143P20 was purchased from ATDBio (Southhampton, UK). Stock solutions were prepared in DI water at a concentration of 1 µM (or 2.34 µg/mL), from which a 10 µL volume was injected, which corresponds to loading 10 picomoles of the 21-mer oligonucleotide on-column.

Experimental Conditions

A BioAccord LC-MS System with an ACQUITY™ Premier UPLC™ System equipped with a 2.1 x 100 mm ACQUITY™ Premier OST Column (P/N 186009485) was used for all oligonucleotide separations. For separation of the 21-mer siRNA oligonucleotide and its impurities, a IP-RP mobile phase containing 7 mM TEA, 40 mM HFIP pH 8.6 was used as Solvent A, while the composition of Solvent B was 3.5 mM TEA, 20 mM HFIP in 50% methanol. Gradient separations were performed from 25% B to 35% B over 25 min. The column flow rate was 0.3 mL/min and the column temperature was 60°C.

RESULTS

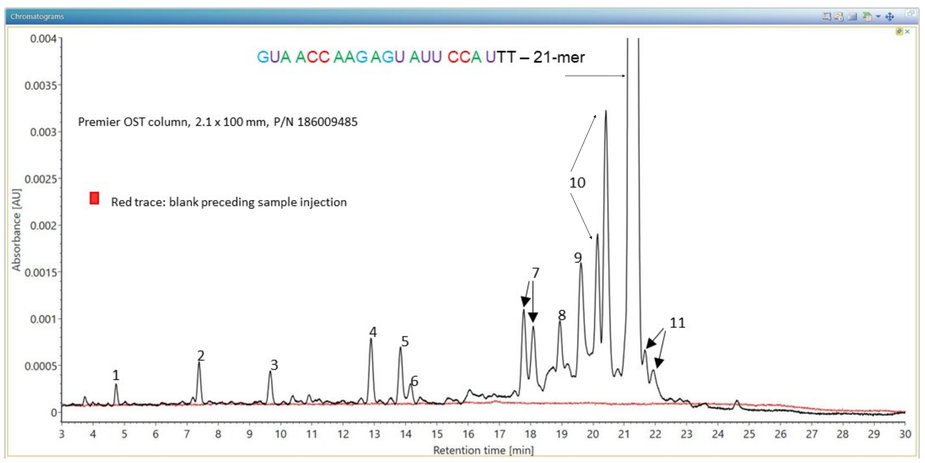

Figure 1. TUV chromatograms showing the separation of oligonucleotide impurities from a 21-mer heavily modified oligonucleotide on an ACQUITY Premier OST Column.

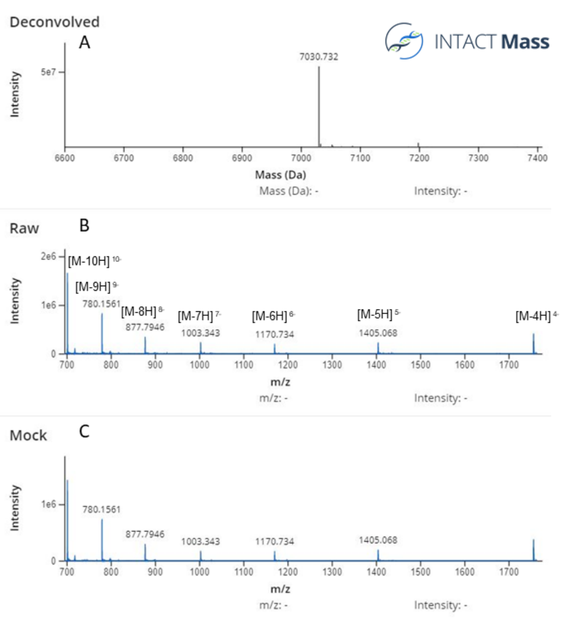

Figure 3. INTACT Mass ESI-MS spectra of the 21-mer heavily modified oligonucleotide: (A): deconvolved spectrum obtained using the BayesSpray deconvolution algorithm; (B) raw MS spectrum; (C) mock spectrum used for assessing the fidelity of the deconvolution process.

Peak no:	Component	Observed mass (Da)	Expected mass (Da)	Mass error (ppm)	Identity result	Observed TIC RT (mins)	Observed UV RT (mins)	LC area	LC amount (%)
1	D1423 n-OMeA[5] & n-OMe5MeC[2] & n-OMeG[2] & n-OMe5MeU	3,593.702	3,593.707	-1.5	Pass	4.74	4.69	1,222	0.2
2	D1423 n-OMeA[4] & n-OMe5MeC[2] & n-OMeG[2] & n-OMe5MeU	3,936.762	3,936.775	-3.4	Pass	7.37	7.34	2,737	0.5
3	D1423 n-OMeA[3] & n-OMe5MeC[2] & n-OMeG & n-OMe5MeU	4,638.916	4,638.907	1.9	Pass	9.66	9.61	2,625	0.5
4	D1423 n-OMeA[2] & n-OMe5MeC[2] & n-OMeG & n-OMe5MeU	4,981.988	4,981.975	2.7	Pass	12.88	12.83	5,276	0.9
5	D1423 n-OMeA[2] & n-OMe5MeC & n-OMeG & n-OMe5MeU	5,315.046	5,315.048	-0.3	Pass	13.80	13.77	4,704	0.8
6	D1423 n-OMeA & n-OMeG & n-OMe5MeU	5,991.181	5,991.188	-1.2	Pass	17.76	17.72	7,942	1.4
7	D1423 n-OMeA	6,684.310	6,684.308	0.3	Pass	19.59	19.55	13,473	2.4
9	D1423 n-OMe5MeU	6,693.332	6,693.320	1.9	Pass	20.13	20.08	13,283	2.3
10	D1423 n-OMe5MeC	6,694.315	6,694.304	1.8	Pass	20.39	20.34	26,001	4.5
8	D1423 unknown[-NH5]	7,008.354	7,008.334	2.9	Pass	20.13	20.08	13,283	2.3
	D1423 MAIN PEAK	7,027.390	7,027.376	2	Pass	21.22	21.19	469,430	82.0
11	D1423 Deamination	7,028.315	7,028.360	-6.5	Pass	20.39	20.34	26,001	4.5

Figure 2. Screenshot with the processing results generated by the INTACT Mass software for the analysis of a 21-mer oligonucleotide and its impurities. The dataset was deconvolved using the BayesSpray charge deconvolution algorithm and eleven oligonucleotide impurities were identified with mass accuracies of under 10 ppm. The first impurity displayed in this table, an 11-mer oligonucleotide, has the lowest detected abundance at 0.2%, according to the UV measurement.

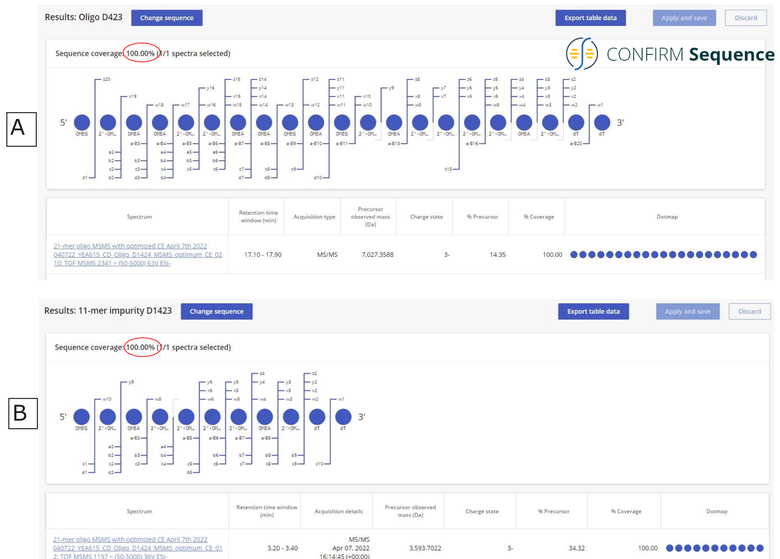

Figure 4. CONFIRM Sequence screenshots showing excellent MS/MS fragmentation coverage (100%) in a dot-map format: (A) for the $[M-3H]^{-3}$ precursor of the 21-mer heavily modified oligonucleotide (m/z = 2342.0), fragmented with an *optimized fixed* collision energy of 63 V); (B) for the $[M-3H]^{-3}$ precursor of the least abundant oligonucleotide impurity (m/z = 1196.9), 11-mer present at 0.2%, fragmented with an *optimized fixed* collision energy of 36 V. Both spectra were recorded on a Xevo™ G2-XS QTof Mass Spectrometer.

Oligo length	Peak label	Retention time (min)	Oligonucleotide sequence	Elemental composition	Most abundant precursor monoisotopic mass / charge state	Second precursor/ charge state	Optimum Collision Energies (V)	MSMS Sequence Coverage (%)	COMBINED Seq Coverage (%)	TUV Area Percentage (%)
11-mer	1	4.73	GU AUU CCA UTT	C119 H161 N33 O77 P10	1196.8952 (-3)	1795.8464 (-2)	36 / 60	100.0	100.0	0.18
12-mer	2	7.39	AGU AUU CCA UTT	C130 H175 N38 O83 P11	1311.2512 (-3)	1967.3805 (-2)	40 / 59	100.0	100.0	0.51
14-mer	3	9.67	AG AGU AUU CCA UTT	C152 H203 N48 O96 P13	1545.295 (-3)	2318.4461 (-2)	46 / 60	100.0	100.0	0.46
15-mer	4	12.89	AAG AGU AUU CCA UTT	C163 H217 N53 O102 P14	1659.651 (-3)	2489.9802 (-2)	49 / 60	81.3	86.7	0.89
16-mer	5	13.83	C AAG AGU AUU CCA UTT	C174 H233 N56 O109 P15	1770.6752 (-3)	1327.7546 (-4)	51 / 40	75.0	83.2	0.83
17-mer	6	14.15	CC AAG AGU AUU CCA UTT	C184 H261 N59 O116 P16	1881.7307 (-3)	1411.0622 (-4)	-	-	-	1.34
18-mer	7	17.8/18.1	CC AAG AGU AUU CCA UTT	C184 H261 N59 O116 P16	1996.0555 (-3)	1496.7898 (-4)	-	-	-	2.38
21-mer	8	18.94	AUA ACC AAG AGU AUU CCA UTT	C229 H306 N76 O142 P20	2336.1198 (-3)	1751.8380 (-4)	-	-	-	2.32
20-mer	9	19.62	GU* ACC AAG AGU AUU CCA UTT	C218 H292 N71 O137 P19	2227.0954 (-3)	1670.0697 (-4)	59 / 50	76.2	76.2	2.32
20-mer	10	20.1/20.4	GUA A*C AAG AGU AUU CCA UTT	C218 H290 N73 O136 P19	2230.4272 (-3)	1672.5686 (-4)	59 / 50	80.0	80.0	6.29
21-mer	FLP	21.28	GUA ACC AAG AGU AUU CCA UTT	C229 H306 N76 O143 P20	2341.4514 (-3)	1755.8367 (-4)	63 / 50	100.0	100.0	82.03
21-mer	11	21.6/21.9	GUA ACC AAG AGU AUU CCA UTT	C229 H307 N76 O143 P20	2341.7874 (-3)	1756.0887 (-4)	-	-	-	0.45

Table I. Eleven oligonucleotide impurities were identified in a 21-mer extensively modified oligonucleotide. Eight impurities and the full length product (FLP) were sequenced using a Xevo G2-XS QTof Mass Spectrometer and the individual MS/MS spectra fragmented with optimum collision energies were processed using the CONFIRM Sequence app. The MS/MS sequence coverage for the FLP and its impurities were above 75%. The lowest abundance impurity, an 11-mer oligonucleotide was sequenced with 95% sequence coverage, while the sequence of the FLP (21-mer oligonucleotide) was confirmed with 100% coverage. The total sequence coverage corresponds to the combined sequence obtained from the MS/MS fragmentation of two precursors of each oligonucleotide.

CONCLUSIONS

- The compliance-ready *waters_connect* INTACT Mass Application performs automated, fast deconvolution of oligonucleotide spectra across the entire chromatographic space, providing fast impurity assignments as well as the required metrics (mass accuracy and abundance) to support impurity analysis.
- The INTACT Mass Application is shown to provide better than 10 ppm mass accuracy for intact mass confirmation of siRNA oligonucleotides and their impurities, down to 0.2% relative abundance levels.
- The CONFIRM Sequence application is used for fast processing of both MS/MS and MS^E (no specific precursor selection) fragmentation spectra for verification of sequence coverage. The software displays the relevant matching information (graphically and in table format) and provides statistical analysis on each matched fragment ion. The sequence coverage can be viewed in a "dot-map" form to easily assess the coverage of a predicted sequence, or to locate an impurity modification, and is capable of high-throughput data analysis of pre-acquired data.

REFERENCES:

1.Sharma VK, Watts JK Oligonucleotide therapeutics: chemistry, delivery and clinical progress, *Future Med Chem*, 2015, 7 (16), 2221-2242
2.An Automated Compliance-Ready LC-MS Workflow for Intact Mass Confirmation and Purity Analysis of Oligonucleotides, 2020, Waters application note, P/N 720006820EN.
3. Analysis of Oligonucleotide Impurities on the BioAccord System with ACQUITY Premier, 2021, Waters application note, P/N 720007301EN.
4. LC-MS Analysis of siRNA, Single Guide RNA and Impurities using the BioAccord System with ACQUITY Premier System and New Automated INTACT Mass Aplication, 2022, Waters application note, P/N 720007546EN.
5. CONFIRM Sequence: a waters_connect Application for Sequencing of Synthetic Oligonucleotide and Their Impurities, 2022, Waters application note, P/N 720007677EN.

TO DOWNLOAD A COPY OF THIS POSTER, VISIT WWW.WATERS.COM/POSTERS