Application Note

Biopharma/Pharma

Agilent

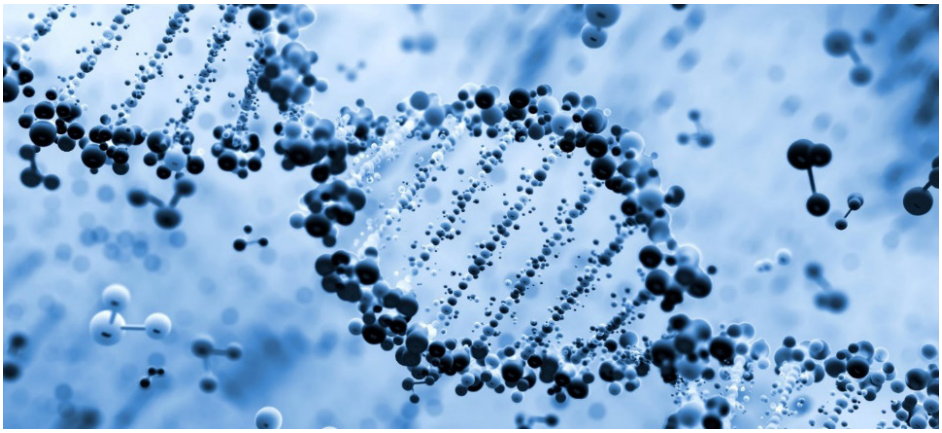
Trusted Answers

Comprehensive and Integrated Workflow for Oligonucleotide Sequence Confirmation by Agilent High-Resolution LC/Q-TOF

Robust and sensitive sequence confirmation of synthetic oligonucleotides and impurities

Authors

David L. Wong and Peter Rye
Agilent Technologies, Inc.



Abstract

This application note demonstrates an oligonucleotide sequence confirmation workflow using Agilent MassHunter BioConfirm software with an Agilent 1290 Infinity II LC and Agilent 6545XT AdvanceBio LC/Q-TOF. The automated data analysis workflow is highly accurate, fast (<1 minute), and powerful – providing complete and high-confidence coverage for short oligonucleotides from just one injection.

Introduction

Oligonucleotides (including small interfering RNA, antisense oligonucleotides, aptamers, and CRISPR guides) have become fast-growing modalities in recent years. With the development of these candidates comes the increased need for robust analytical methods and easy-to-use data analysis workflows to characterize them. Commonly sought-after attributes of oligonucleotides include the mass, purity (and relative quantity of impurities present), and sequence.

The characterization of oligonucleotide mass and related impurities has already been described.[1] This Target Plus Impurities (TPI) workflow used an Agilent 1290 Infinity II LC coupled to the Agilent 6545XT AdvanceBio LC/Q-TOF. This previous work used an MS1-based acquisition method and an automated Agilent MassHunter BioConfirm software 12.0 data analysis method to provide multiple examples of the identification and relative quantification of oligo targets, impurities, modifications, and degradation products.

This application note describes the use of the same analytical system and software, using a targeted MS/MS-based method, to provide fragment confirmation maps on a wide variety of oligonucleotide samples, including single- and double-stranded RNA and DNA. This sequencing workflow uses fragment confirmation at the MS2 level by matching isotope patterns against expectations that are calculated from the oligo sequence. These new software features demonstrate the power of a high-resolution accurate mass system (HRAM) and targeted MS/MS data to structurally characterize oligonucleotides by confirming heavily modified sequences and determining the positions of specific chemical groups.

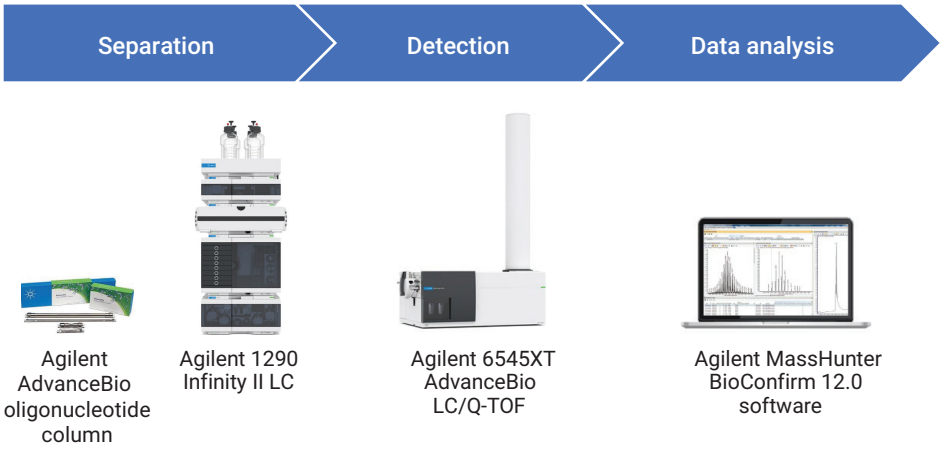


Figure 1. Analytical components of the oligonucleotide analysis—Sequence Confirmation workflow.

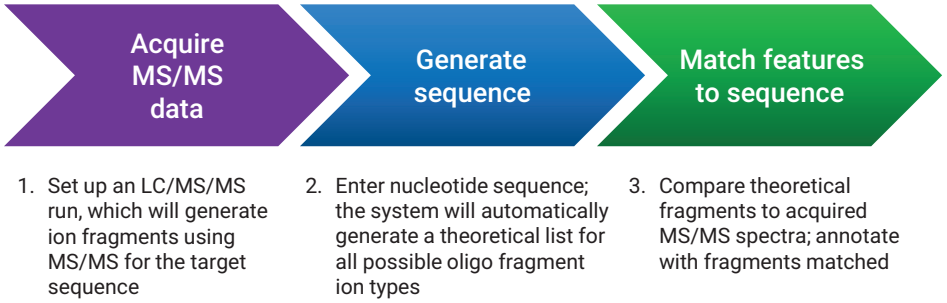


Figure 2. Oligonucleotide Sequence Confirmation data analysis workflow in Agilent MassHunter BioConfirm software, version 12.0.

Experimental

Materials and methods

- Triethylamine (TEA) and 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Methanol (InfinityLab Ultrapure LC/MS grade, part number 5191-4497) was obtained from Agilent Technologies.
- The oligonucleotide (RNA) resolution standard (part number 5190-9028) were from Agilent.
- All other synthetic oligonucleotides were purchased from Integrated DNA Technologies, Inc. (Coralville, IA, USA).

Sample preparation

- The Agilent oligonucleotide (RNA) resolution standards were dissolved with 1 mL de-ionized (DI) water before use. The final concentration was 2 pmol/µL.
- All synthetic oligonucleotide samples were also dissolved with 1 mL of DI water without further purification. Samples were then diluted to 0.50 mg/mL in stock solution.

Instrumentation

- Agilent 1290 Infinity II LC including:
  - Agilent 1290 Infinity II High-Speed Pump (G7120A)
  - Agilent 1290 Infinity II Multisampler (G7167B) with Agilent Infinity II Sample Cooler (Option #100)
  - Agilent 1290 Infinity II Multicolumn Thermostat (G7116B)
- Agilent 6545XT AdvanceBio LC/QTOF

LC/MS analysis

LC/MS analyses were conducted on a 1290 Infinity II LC coupled with a 6545XT AdvanceBio LC/Q-TOF equipped with an Agilent Dual Jet Stream ESI source. Agilent MassHunter Acquisition 11.0 workstation software with compliance features enabled was used. LC separation was obtained with an Agilent AdvanceBio oligonucleotide column (2.1 × 50 mm, 2.7 µm, part number 659750-702).

Tables 1 and 2 list the detailed LC/MS parameters that were used.

Data processing

All LC/MS/MS data files of the oligonucleotide standards and synthetic oligonucleotides samples were processed using Agilent MassHunter BioConfirm software, version 12.0.

Table 1. Liquid chromatography parameters.

Agilent 1290 Infinity II LC	
Column	Agilent AdvancedBio oligonucleotide, 2.1 × 50 mm, 2.7 µm (p/n 659750-702)
Thermostat	4 °C
Solvent A	15 mM TEA and 400 mM HFIP in water
Solvent B	Methanol
Gradient	Time (min) %B 0 to 1 10 1 to 10 10 to 40 10 to 11 40 to 95
Column Temperature	65 °C
Flow Rate	0.5 mL/min
Injection Volume	5.0 µL

Table 2. MS data acquisition parameters.

Agilent 6545XT AdvanceBio LC/Q-TOF	
ESI Source	Dual AJS
Polarity	Negative
Gas Temperature	275 °C
Drying Gas	12 L/min
Nebulizer	35 psig
Sheath Gas Temperature	350 °C
Sheath Gas Flow	12 L/min
VCap	3,500 V
Nozzle Voltage	2,000 V
Fragmentor	175 V
Skimmer	65
Quad AMU	95
Reference Mass	966.000725
Acquisition Mode	Standard mass range, HiRes (4 GHz)
Mass Range	350 to 3,000 *m/z*
Acquisition Rate	1 spectra/sec
Targeted MS/MS Range	100 to 3,000 *m/z*
Min MS/MS Acquisition Rate	2 spectra/sec
Isolation Width	Medium (~4 *m/z*)
Collision Energy (CE)(V)	Optimized individually
Acquisition Time	500 ms/spec

Results and discussion

The liquid chromatography/mass spectrometry (LC/MS) technique has become a powerful analytical tool for synthetic oligonucleotide analysis. However, as larger data files are being generated, a comprehensive characterization of an oligonucleotide sample can be a challenging and time-consuming process. Recently, an integrated workflow has been developed for the analysis of targeted oligonucleotides and their impurities by high resolution mass spectrometry (HRMS).[1] This study demonstrates that obtaining good MS/MS fragmentation data and performing accurate data analysis are two of the most critical attributes in the oligonucleotide Sequence Confirmation workflow.

HPLC separation and HRMS analysis of synthetic oligonucleotides

Excellent chromatographic separation and mass accuracy (sub-ppm) for expected oligonucleotides were achieved in a recent study. This study used the optimized LC/MS-based methodology to analyze the oligonucleotide standards. Figure 3 illustrates the LC/MS analysis of Agilent oligonucleotide (RNA) resolution standards containing four sequences (14-, 17-, 20-, and 21-mer). Baseline separation was achieved in a short HPLC run for oligonucleotides that differ with only a single nucleotide base (20-mer versus 21-mer, Figure 3A).

Figures 3B and 3C demonstrate the excellent MS isotopic resolution for the –3 charge state of the 21-mer RNA standard (Figure 3B, inset) and the good mass accuracy (–1.09 ppm) using the Find-by-Formula (FBF) algorithm. Overall, our optimized LC/MS analysis generated excellent MS results on all four RNA standards with mass errors of 1 ppm (Table 3).

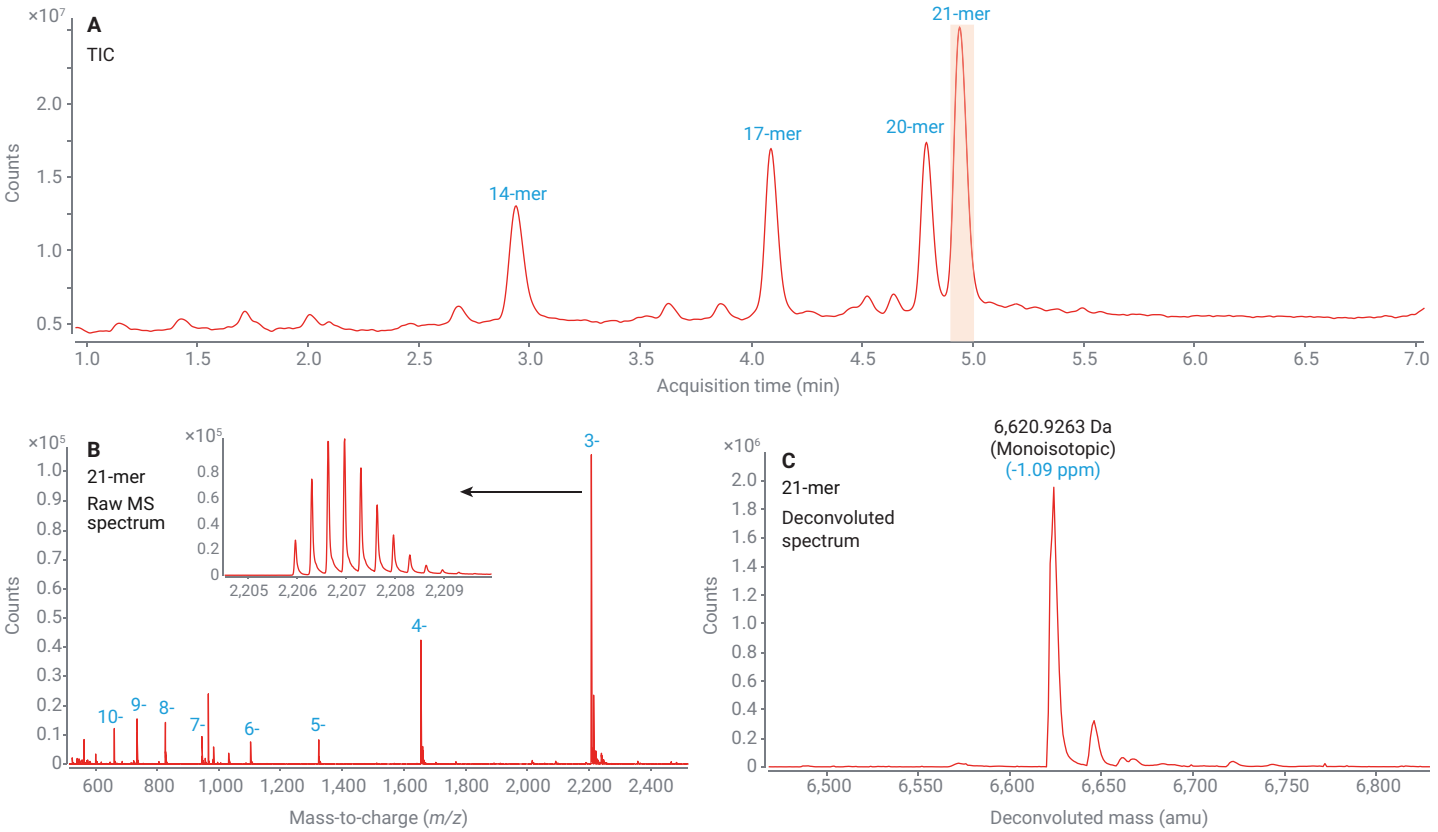


Figure 3. LC/MS analysis of Agilent oligonucleotide (RNA) resolution standards. (A) Total ion chromatography (TIC) of the resolution standards. (B) Raw MS spectrum of the selected 21-mer RNA. (C) Deconvoluted MS spectra of the 21-mer RNA standard.

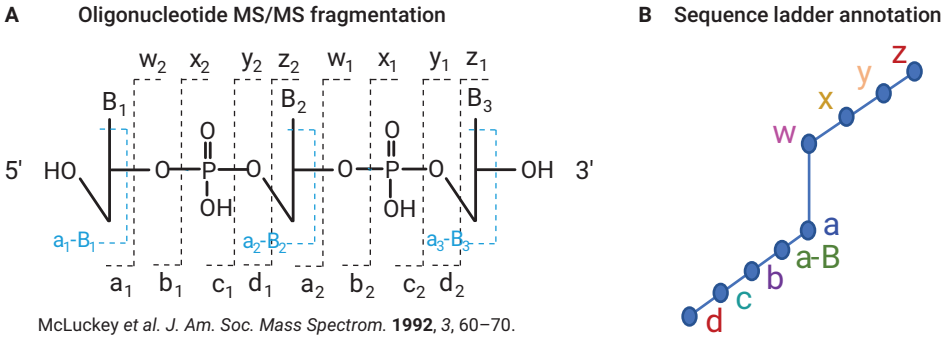
Table 3. LC/MS analysis summary of Agilent oligonucleotide RNA resolution standards. All four oligonucleotides posted approximately 1 ppm of mass accuracy using the Find-by-Formula (FBF) algorithm.

Oligonucleotide	Oligo Length	Sequence	Calculated Monoisotopic Mass (Da)	Measured Mass (Da)	Mass Accuracy (ppm)
Oligonucleotide (RNA) Resolution Standard	14	rCrArCrUrGrArArUrArCrCrArArU	4,395.6479	4,395.6429	– 1.14
	17	rUrCrArCrArCrUrGrArArUrArCrCrArArU	5,335.7670	5,335.7623	– 0.88
	20	rUrCrArUrCrArCrArCrUrGrArArUrArCrCrArArU	6,275.8861	6,275.8800	– 0.97
	21	rGrUrCrArUrCrArCrArCrUrGrArArUrArCrCrArArU	6,620.9335	6,620.9263	– 1.09

Oligonucleotide sequence confirmation data analysis workflow

Oligonucleotide sequence confirmation by the combination of liquid chromatography and electrospray ionization mass spectrometry (LC/ESI-MS) is a well-established technique[2,3] used by the biopharmaceutical industry. However, the lack of an automated workflow in the data processing and result interpretation has been the rate-limiting step for most biopharmaceutical analytical laboratories. Newly developed BioConfirm 12.0 software includes both the Target Plus Impurities (TPI) and Sequence Confirmation workflows. The TPI workflow uses MS1 data to profile (i.e., provide identities and relative quantitation) the target oligonucleotide and related impurities. While the Sequence Confirmation workflow uses MS/MS data to confirm the targeted oligonucleotide sequences as well as any associated modifications.

The sequence confirmation algorithm in BioConfirm 12.0 adopted the standard nomenclature of oligonucleotide fragmentation for data processing. A unique fragment ions annotation for the Fragment Confirmation Ladder was also developed (Figure 4).



McLuckey *et al. J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60–70.

Figure 4. Nomenclature of oligonucleotide fragmentation (A) and annotation for oligonucleotide Fragment Confirmation Ladder (B). The colors of letters represent the different types of fragment ions identified and labeled in the MS/MS spectra.

Figure 5 highlights the user interface of the Sequence Confirmation workflow in BioConfirm 12.0 software, where the user can set up parameters for data processing. The Biomolecule Fragment Spectrum displays either the individual MS/MS data of the identified biomolecule with its theoretical isotopic distribution comparison or the combined MS/MS spectrum of all identified biomolecules (fragment ions). The detailed results, such as identified oligonucleotide fragment ions and their abundances, are shown in the Biomolecules table. The novel Fragment Confirmation Ladder displays all identified fragment ions with the targeted oligonucleotide sequence coverage percentage.

To develop and test the oligonucleotide Sequence Confirmation workflow, many oligonucleotide samples with various sequences, lengths, and modifications have been investigated in this study. The LC/MS/MS methodology has been optimized for each of these oligonucleotides in terms of charge state selection, collision energy (CE), and MS/MS acquisition time to get the best of fragmentation data. It is worthwhile to mention that the BioConfirm 12.0 algorithm allows users to group multiple MS/MS scans for data analysis either by charge state or by collision energy. Therefore, the user can set up a single LC/MS method using various CEs on different charge state ions to obtain the best oligonucleotide sequence coverage.

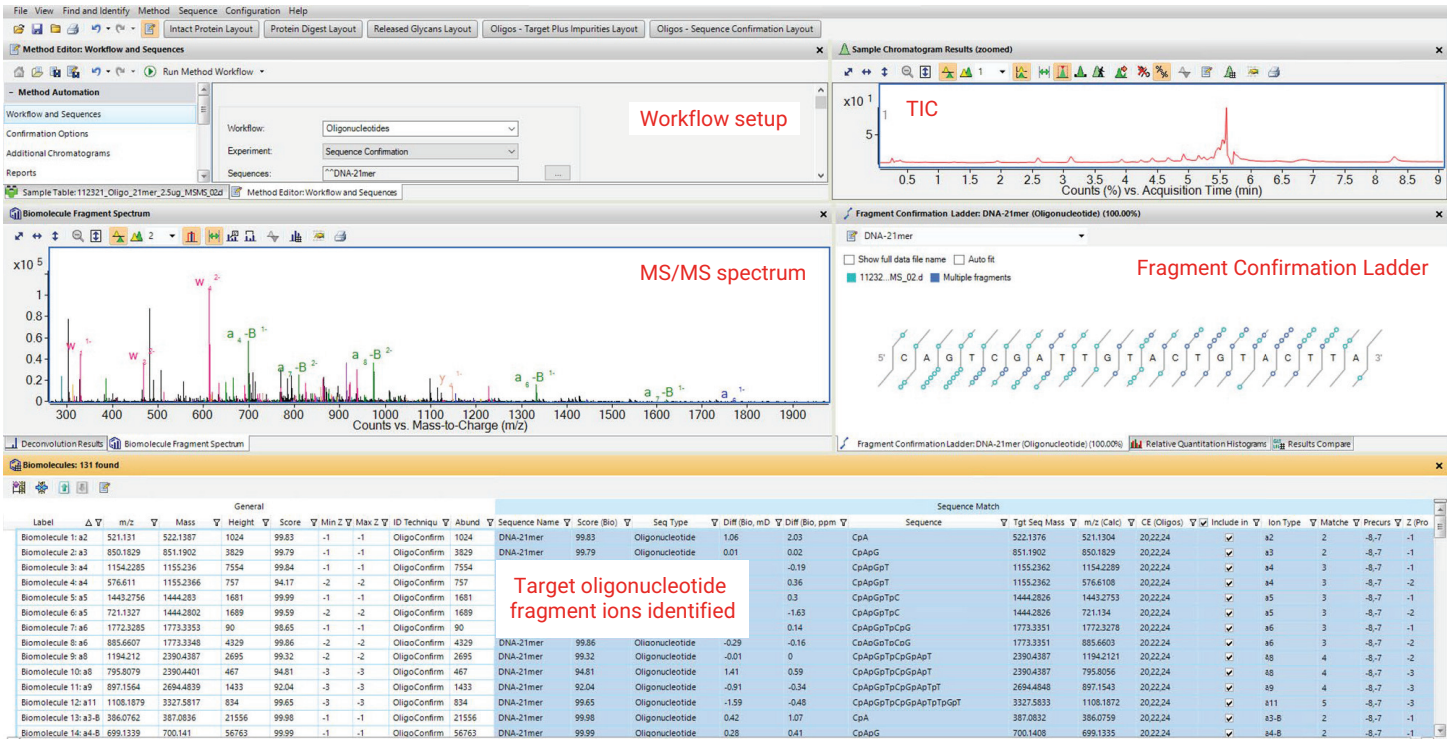


Figure 5. Overview of Agilent BioConfirm software, version 12.0 with Sequence Confirmation workflow.

The following are the experimental details of sequence confirmation on various oligonucleotides:

MS/MS fragmentation of synthetic RNA standards (20-mer versus 21-mer)

Even though accurate masses were obtained on the RNA resolution standards (Table 3) within 1 ppm of mass accuracy, good MS/MS data on the 20-mer and 21-mer RNA standards were generated (Figure 6). Any of the a, a-B, b, c, and d fragment ions of the 21-mer were able to be used to determine the additional rG at its 5' end. Both oligonucleotide sequences were confirmed with 100% sequence coverage from just one injection of each.

Sequence confirmation of isomeric 21-mer oligonucleotides

The BioConfirm 12.0 oligonucleotide Sequence Confirmation workflow allows determination of the correct sequences for isomeric oligonucleotides. A pair of oligonucleotides with exact masses was also tested, but two bases swapped (21-mer versus 21-mer BS). Figure 7 illustrates the LC/MS/MS analysis of both oligonucleotides. As can be seen in the upper panels, both molecules had the same retention time and mass. Therefore, only the MS/MS data can be used to confirm their isomeric characteristic in sequence.

Figure 8 displays the zoomed-in MS/MS spectrum of the a_5 (–1) fragment ions, which had a mass difference distribution (red boxes over peaks) applied to determine the confidence level (Bio score) for the identification. Similarly, base swapping at position 17 was confirmed with any w, x, y, and z ions (≥5).

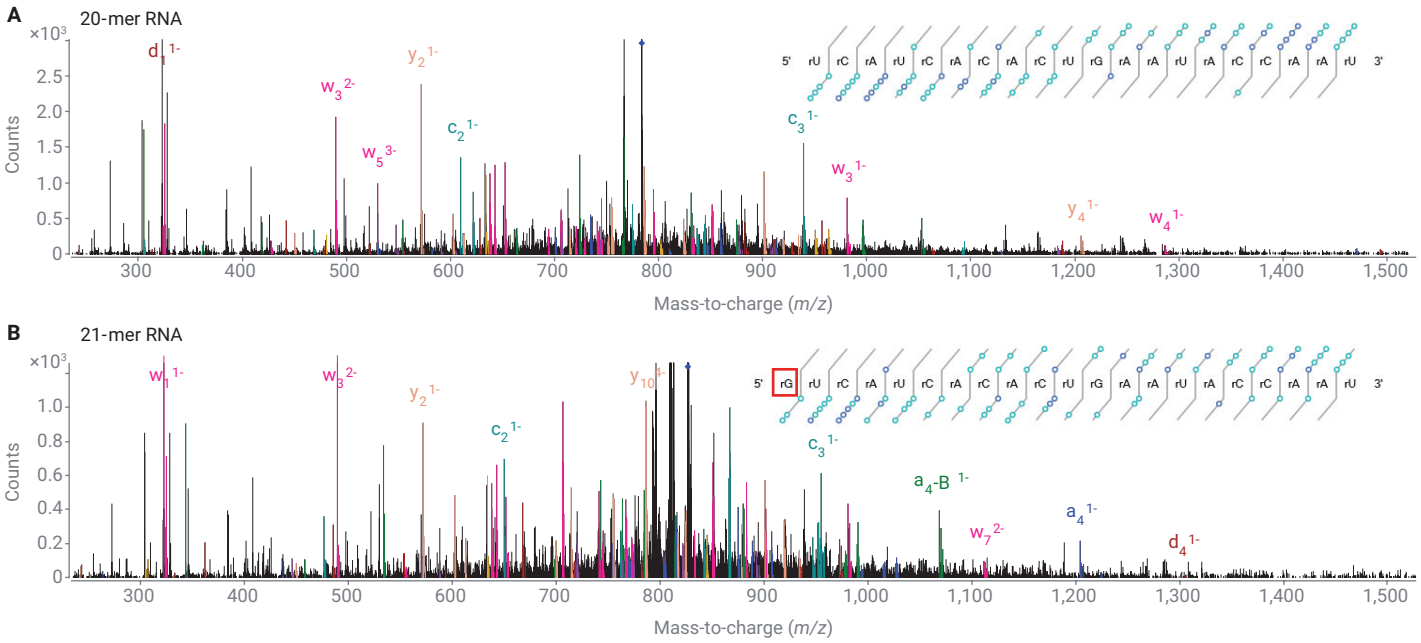


Figure 6. LC/MS/MS analysis on the Agilent oligonucleotide (RNA) resolution standard (20-mer and 21-mer). MS/MS spectrum of 20-mer and 21-mer RNA with Fragment Confirmation Ladder displayed.

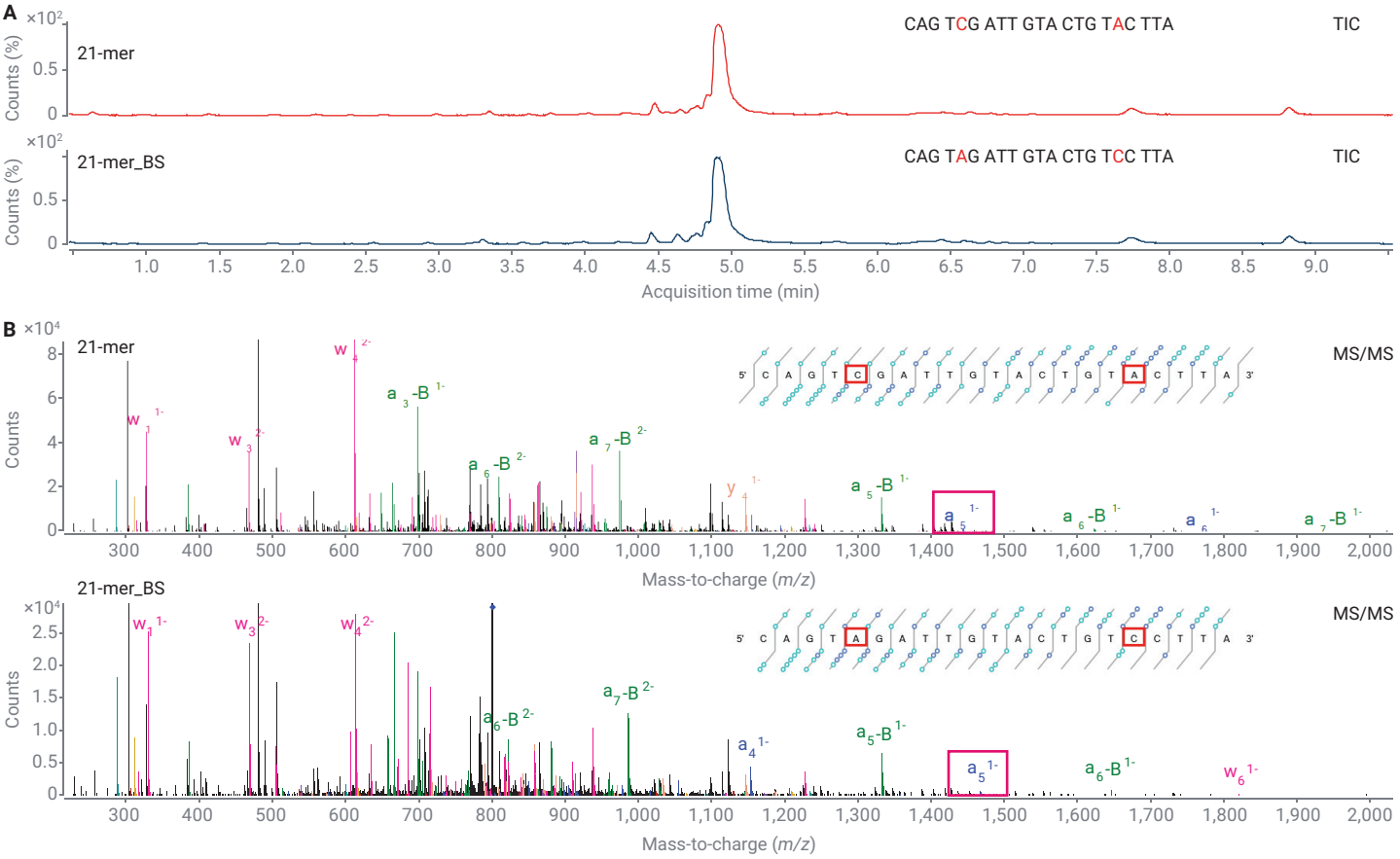


Figure 7. LC/MS/MS analysis on synthetic oligonucleotides (21-mer and 21-mer BS). Top two panels: Total ion chromatography (TIC) of the 21-mer and 21-mer BS (base swap) oligonucleotides. Bottom two panels: MS/MS spectrum of the 21-mer and 21-mer BS oligonucleotides.

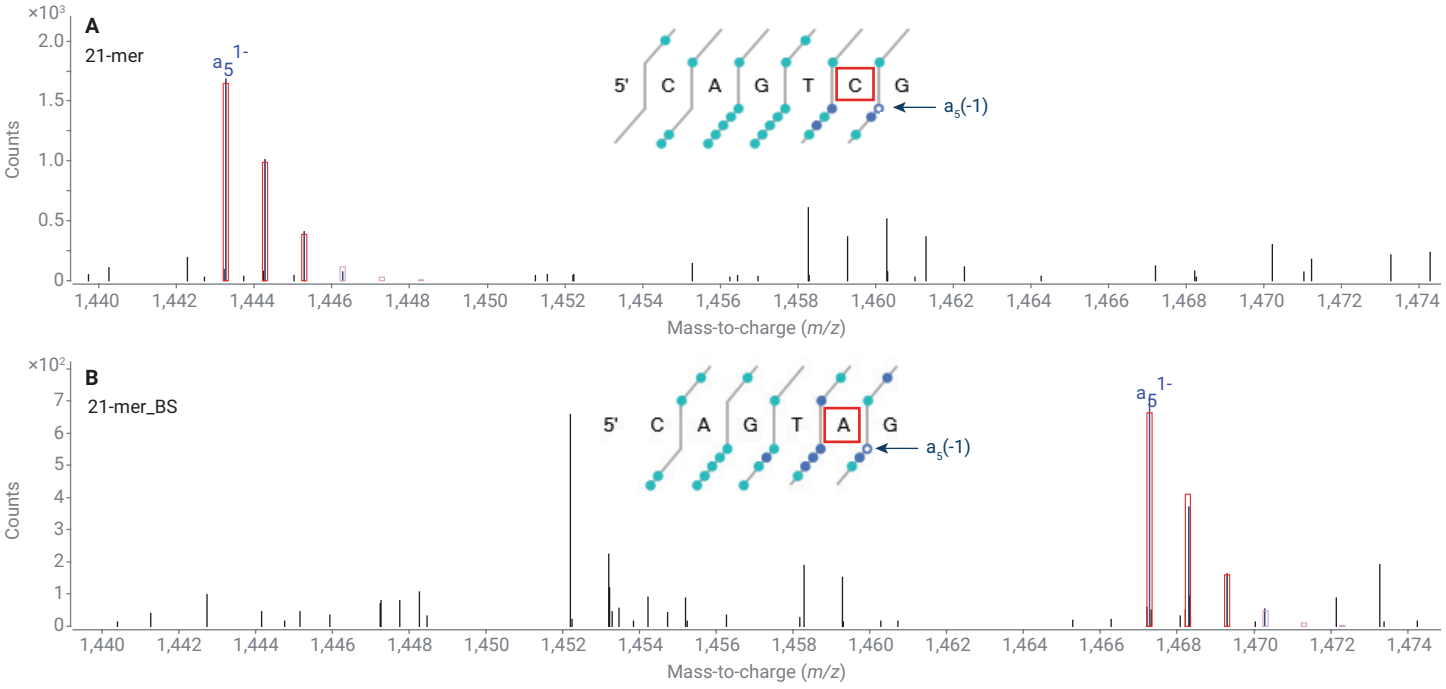


Figure 8. Zoomed MS/MS spectrum of a5 (–1) fragment ions of oligonucleotides (21-mer and 21-mer BS).

Sequence verification of double-stranded oligonucleotide (21-mer DNA and its complementary strand)

The LC/MS-based analysis and sequence confirmation of noncovalent double-stranded oligonucleotides are of great interest to the biopharmaceutical industry. To test the BioConfirm 12.0 Sequence Confirmation workflow on the analysis of double-stranded oligonucleotides, a complementary strand (21-mer complement DNA) of the 21-mer DNA was synthesized and mixed. The noncovalent duplex was then analyzed by LC/MS/MS. Under the traditional LC/MS conditions with organic solvents and ion paring reagents, both DNA strands were coeluted at the same retention time as shown in Figure 9A. However, their identities were confirmed by the Find-by-Formula algorithm in BioConfirm 12.0 Target Plus Impurities (TPI) workflow. The sequences of both DNA strands were further confirmed by the Sequence Confirmation workflow using their MS/MS data. Complete sequence coverages (100%) on both oligonucleotides were achieved with just a single sample injection each.

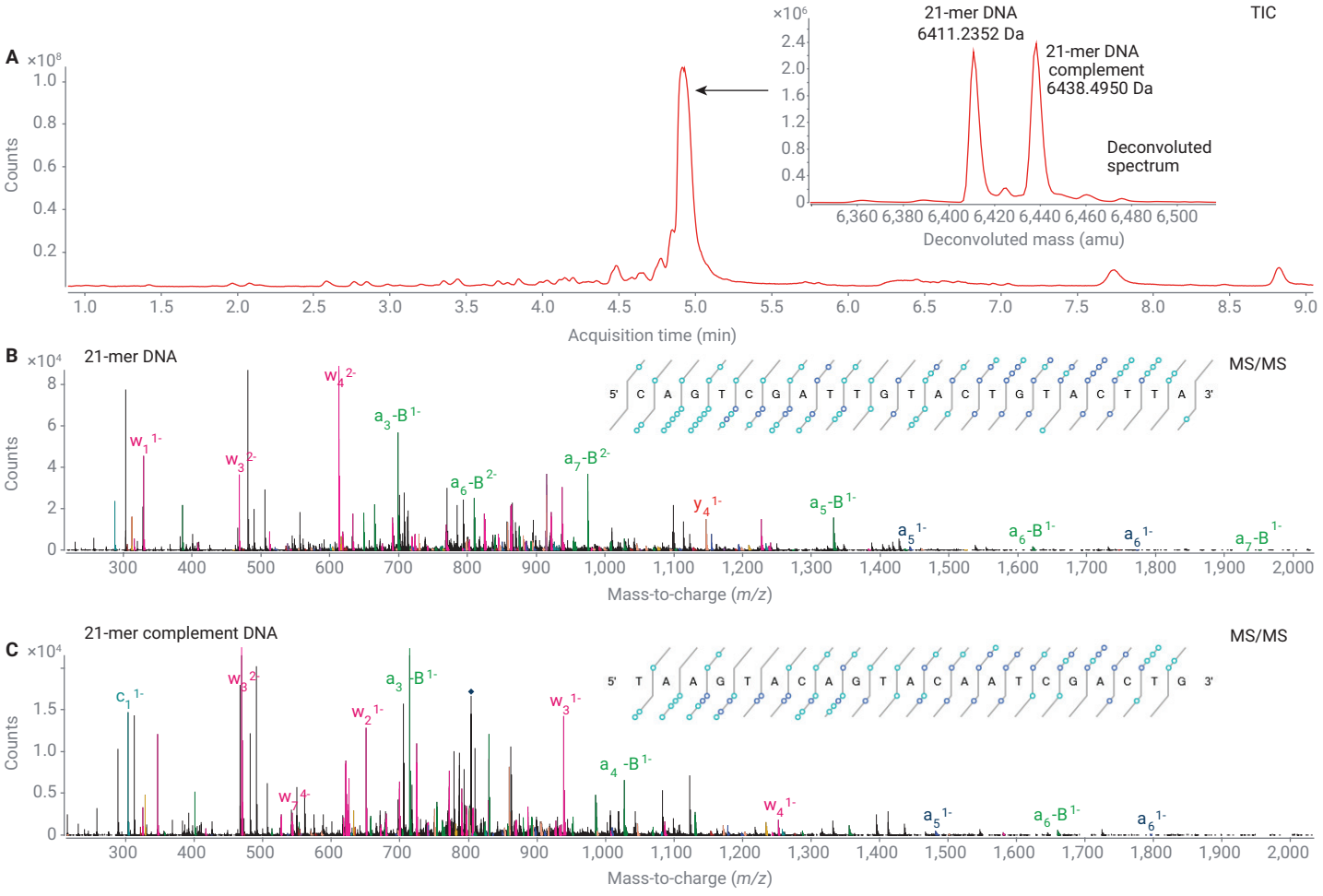


Figure 9. LC/MS/MS analysis of double-stranded DNA (21-mer and 21-mer complement). (A) Total ion chromatography (TIC) of double-stranded DNA (21-mer and 21-mer complement). (B) and (C) MS/MS spectrum of double-stranded DNA (21-mer and 21-mer complement).

MS/MS fragmentation of heavily modified oligonucleotide

Chemically modified oligonucleotides, including antisense oligonucleotides (ASOs), have great biopharma application potential in recent years as they are expected to perform in the target cells to prevent translation and inhibit gene and protein expression.[4]

In this study, a heavily modified 18-mer ASO was used to generate MS/MS data for sequence confirmation testing. A chemical group "2-methoxyethoxy (2'-MOE)" was incorporated into every base of an oligonucleotide at the time of synthesis. Figure 10 shows the LC/MS/MS analysis of this ASO sample. During MS/MS data acquisition, the charge state −8 of the targeted ASO was selected for subsequent fragmentation using various CEs. Excellent MS/MS spectrum of the ASO was generated, and we also achieved 100% sequence coverage by the BioConfirm 12.0 software program (Figure 10C).

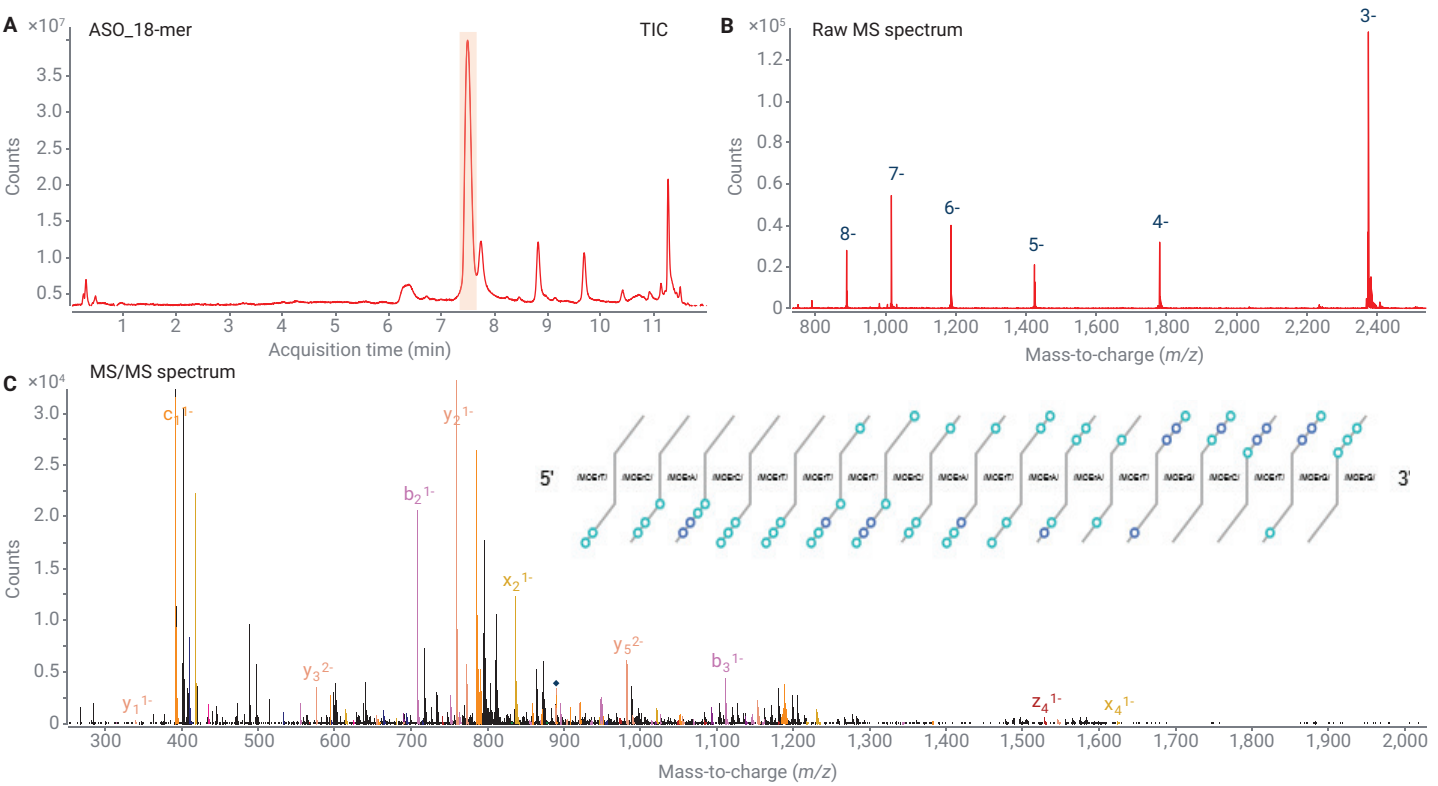


Figure 10. LC/MS/MS analysis on oligonucleotide with specific modifications (ASO_18-mer). (A) Total ion chromatography (TIC) of ASO_18-mer. (B) Raw MS spectrum of ASO_18-mer. (C) MS/MS spectrum of ASO_18-mer with Fragment Confirmation Ladder displayed.

Identification and sequence confirmation of truncated oligonucleotide impurities

In the Target Plus Impurities application note[1], the full target (21-mer), as well as its 5' truncated impurities, were identified, and relative quantitation was accurately determined. However, the sequences of all impurities have yet to be confirmed. In this study, high-quality MS/MS data were acquired for both the N-1 (2-21) and N-2 (3-21) impurities, even though their abundances were low (<5%).

Figure 11 demonstrates the MS/MS spectrum and the Fragment Confirmation Ladder for the full-length target (21-mer) and the N-1 (2-21) and the N-2 (3-21) impurities. All three oligonucleotide sequences were confirmed with 100% coverage.

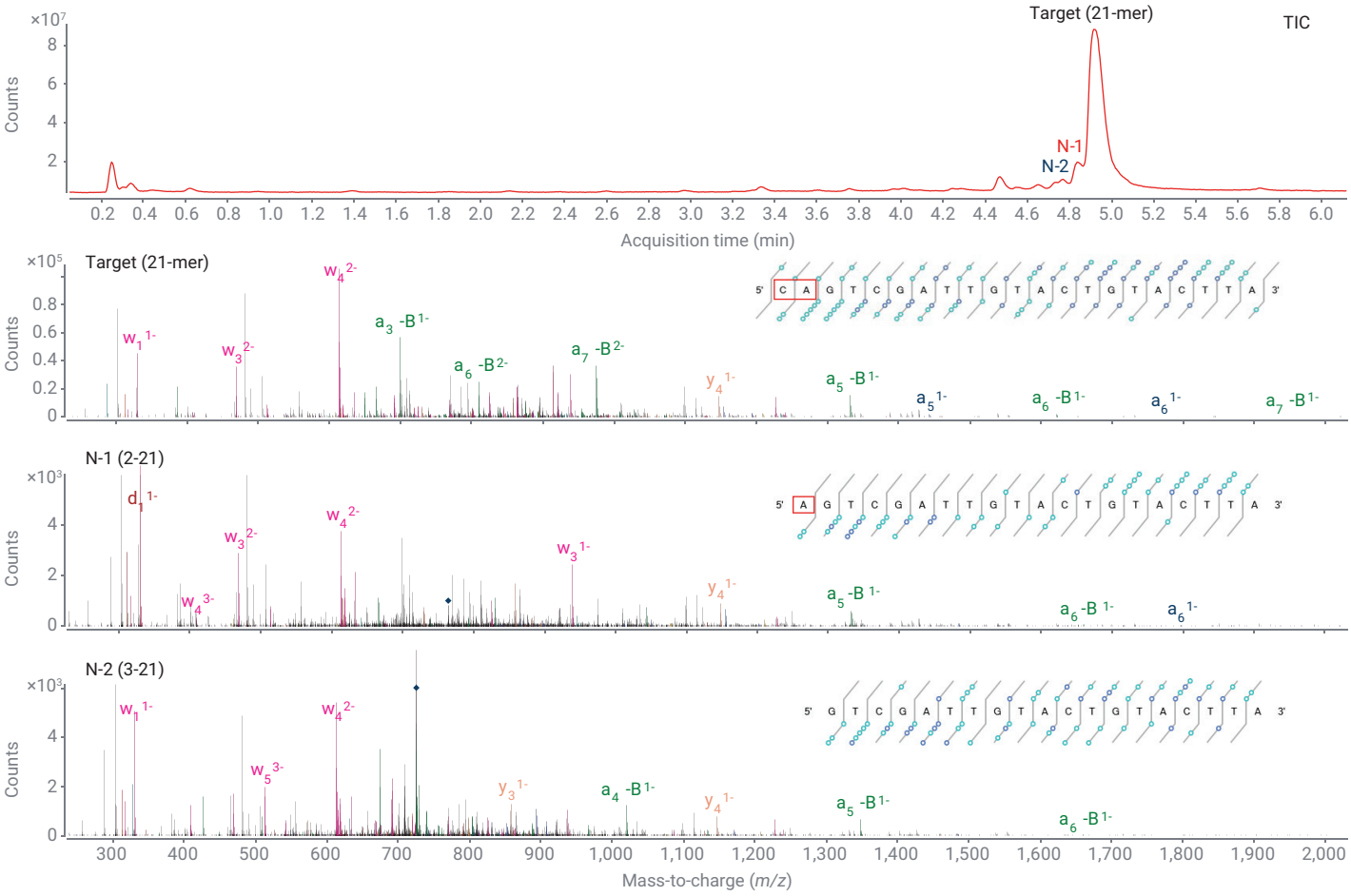


Figure 11. Oligonucleotide sequence confirmation of full-length target (21-mer) and the N-1 (2-21) and the N-2 (3-21) impurities.

Conclusion

An oligonucleotide sequence confirmation workflow that turns the time-consuming data processing into automated and integrated data analysis was developed. This novel Agilent MassHunter BioConfirm software version 12.0 uses the predicted isotope distribution program to confirm the oligonucleotide fragment ions, which increases the identification accuracy and confidence significantly. The Agilent 1290 Infinity II LC coupled with an Agilent 6545XT AdvanceBio LC/Q-TOF also generated superior HRAM MS/MS data that resulted in 100% sequence coverages from just one injection of each for all oligonucleotides analyzed.

References

1. Wong, D.; Rye, P. An Integrated Workflow for the Analysis of Oligonucleotides and Their Impurities by Agilent High-Resolution LC/(Q-) TOF Mass Spectrometry. *Agilent Technologies application note*, publication number 5994-4817EN, **2022**.
2. McLuckey, S. *et al.* Tandem Mass Spectrometry of Small, Multiply Charged Oligonucleotides. *J. Am. Soc. Mass Spectrom*. **1992**, *3(1)*, 60–70.
3. McCloskey, J. *et al.* Interpretation of Oligonucleotide Mass Spectra for Determination of Sequence Using Electrospray Ionization and Tandem Mass Spectrometry. *Anal. Chem.* **1996**, *68(1)*, 1989–1999.
4. Bahal, R. *et al.* Antisense Oligonucleotides: An Emerging Area in Drug Discovery and Development. *J. Clin. Med*. **2020**, *9*, 2004.

www.agilent.com

DE15315662

Printed in the USA, July 13, 2022
5994-5071EN

Agilent
Trusted Answers