Application Note

Biopharma/Pharma

Agilent
Trusted Answers

An Integrated Workflow for the Analysis of Oligonucleotides and Their Impurities by Agilent High-Resolution LC/(Q-)TOF Mass Spectrometry

Separation, characterization, and relative quantitation of target oligonucleotides and their impurities

Authors

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

Introduction

Oligonucleotides (including small interfering RNA, antisense oligonucleotides, aptamers, and CRISPR guides) have become fast-growing modalities in recent years. Along with the development of these candidates comes the increased need for robust analytical methods and easy-to-use data analysis workflows to characterize them. In addition, the characterization of product-related impurities is an important task in the development of new biotherapeutics. Common impurities include conversion of phosphorothioate to phosphodiester, truncations, extensions, and abasic oligonucleotides.[1,2]

Advanced analytical methods, such as LC/MS analysis, are indispensable for the characterization of target oligonucleotides and their impurities, which are often numerous, present at very low abundances, and found in combination with one another. As such, software that supports and automates these profiling efforts can be of great value.

To overcome these obstacles, Agilent has developed the novel, automated Agilent MassHunter BioConfirm software version 12.0 that supports Find-by-Formula (FBF) and Maximum Entropy algorithms for the identification of the target, plus its impurities. Figure 2 describes the details of the Target Plus Impurities (TPI) data analysis workflow.

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. Oligonucleotide (DNA) Ladder Standard (part number 5190-9029), Oligonucleotide (RNA) Resolution Standard (part number 5190-9028), and RNA Standard (100-mer) were all obtained from Agilent.

A 21-mer (CAG TCG ATT GTA CTG TAC TTA) and a 40-mer (CCA CGA CCA AGT GAC AGC AAT GAA TCG AGT CGA GAT CCA T) oligonucleotide were purchased from Integrated DNA Technologies, Inc. (Coralville, IA, USA) with standard desalting purification.

Sample preparation

- Both oligonucleotide (DNA) Ladder Standard and (RNA) Resolution Standard were dissolved with 1 mL deionized (DI) water before use. The final concentration of each was 2 pmol/µL.
- The concentration of 100-mer RNA standard sample was 0.4 mg/mL.
- The 21-mer and 40-mer 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.

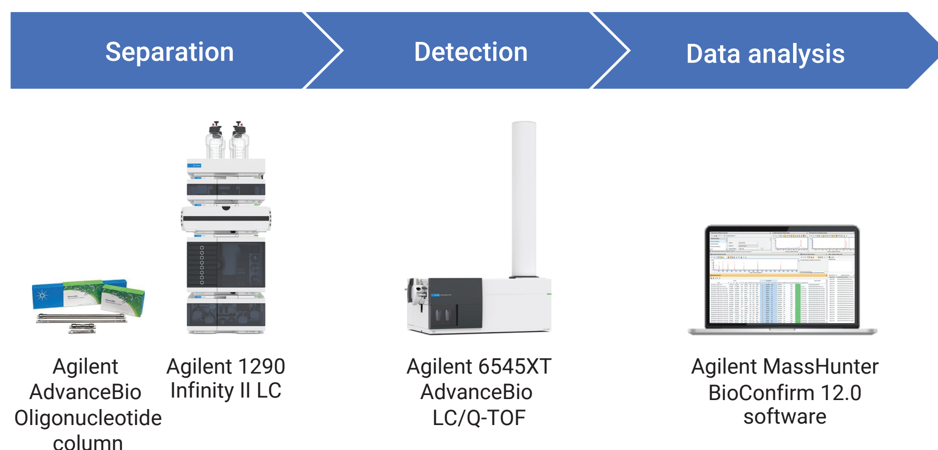
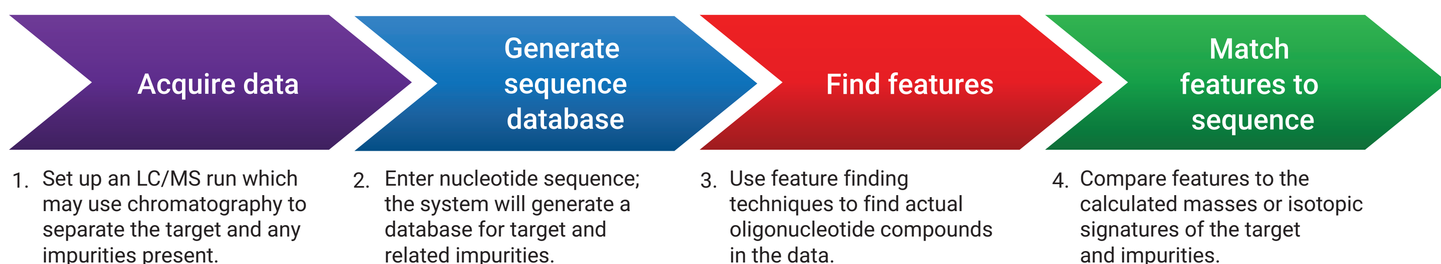


Figure 1. Analytical components of the oligonucleotide analysis - Target Plus Impurities (TPI) workflow.



1. Set up an LC/MS run which may use chromatography to separate the target and any impurities present.
2. Enter nucleotide sequence; the system will generate a database for target and related impurities.
3. Use feature finding techniques to find actual oligonucleotide compounds in the data.
4. Compare features to the calculated masses or isotopic signatures of the target and impurities.

Figure 2. Target Plus Impurities (TPI) data analysis workflow in Agilent MassHunter BioConfirm software, version 12.0.

Instrumentation

- Agilent 1290 Infinity II LC including:
  - Agilent 1290 Infinity II High-Speed Pumps (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/Q-TOF

LC/MS analysis

LC/MS analyses were conducted on a 1290 Infinity II LC coupled with a 6545XT AdvanceBio LC/Q-TOF system equipped with a Dual Agilent Jet Stream source. Agilent MassHunter Acquisition Workstation software (version 11.0) was used, with compliance features enabled. 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 used.

Data processing

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

Table 1. Liquid chromatography parameters.

Agilent 1290 Infinity II LC	
Column	AdvanceBio 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	0 to 1 min, 10% B 1 to 10 min, 10 to 40% B 10 to 11 min, 40 to 95% B
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 System	
Parameter	**Setting**
Source	Dual AJS
Polarity	Negative
Gas Temperature	275 °C
Gas Flow	12 L/min
Nebulizer	35 psi
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 V
Acquisition Mode	HiRes (4 GHz)
Mass Range	300 to 3,200 *m/z*
Acquisition Rate	4 spectra/s

Results and discussion

A comprehensive characterization of an oligonucleotide sample can be a challenging and time-consuming process, as it is not only to profile the target oligonucleotide, but also to obtain the identities and relative quantitation of all related impurities.

HPLC separation of various oligonucleotide standards

To begin the investigation, an LC/MS-based methodology with the excellent chromatographic resolving power and high-resolution accurate-mass (HRAM) detection was optimized. Figure 3 illustrates the LC/MS analysis of two Agilent oligonucleotide standards: DNA Ladder Standard and RNA Resolution Standard. Outstanding chromatographic separation was achieved using ion-pairing reversed-phase chromatography. Separation and detection of the impurities (minor peaks) from the main peaks was also achieved.

LC/MS analysis of synthetic oligonucleotides

Highly sensitive and mass-resolved MS data were produced using the same method. Figure 4 shows the LC/MS results from a synthetic oligonucleotide (40-mer) sample. Approximately 2.5 µg of sample was injected onto the column using an 11-minute gradient with a flow rate of 0.5 mL/min. The charge state distribution of the 40-mer oligonucleotide was in the mass range of *m/z* 600 to 3,000 (–5 to –19). The zoomed view on the insert of Figure 4B demonstrates the excellent MS isotopic resolution for the –13 charge state of the oligonucleotide.

The Q-TOF source conditions were optimized, and excellent quality MS spectra with low mass error (3.6 ppm) was obtained (Table 3). In addition, low-abundance truncation species (~12,000 Da) and the extension species (~12,600 Da) were detected (Figure 4C).

A similar high-quality MS result was obtained on the 100-mer RNA standard sample, as shown in Figure 5. Again, good mass accuracy (9.96 ppm) was observed.

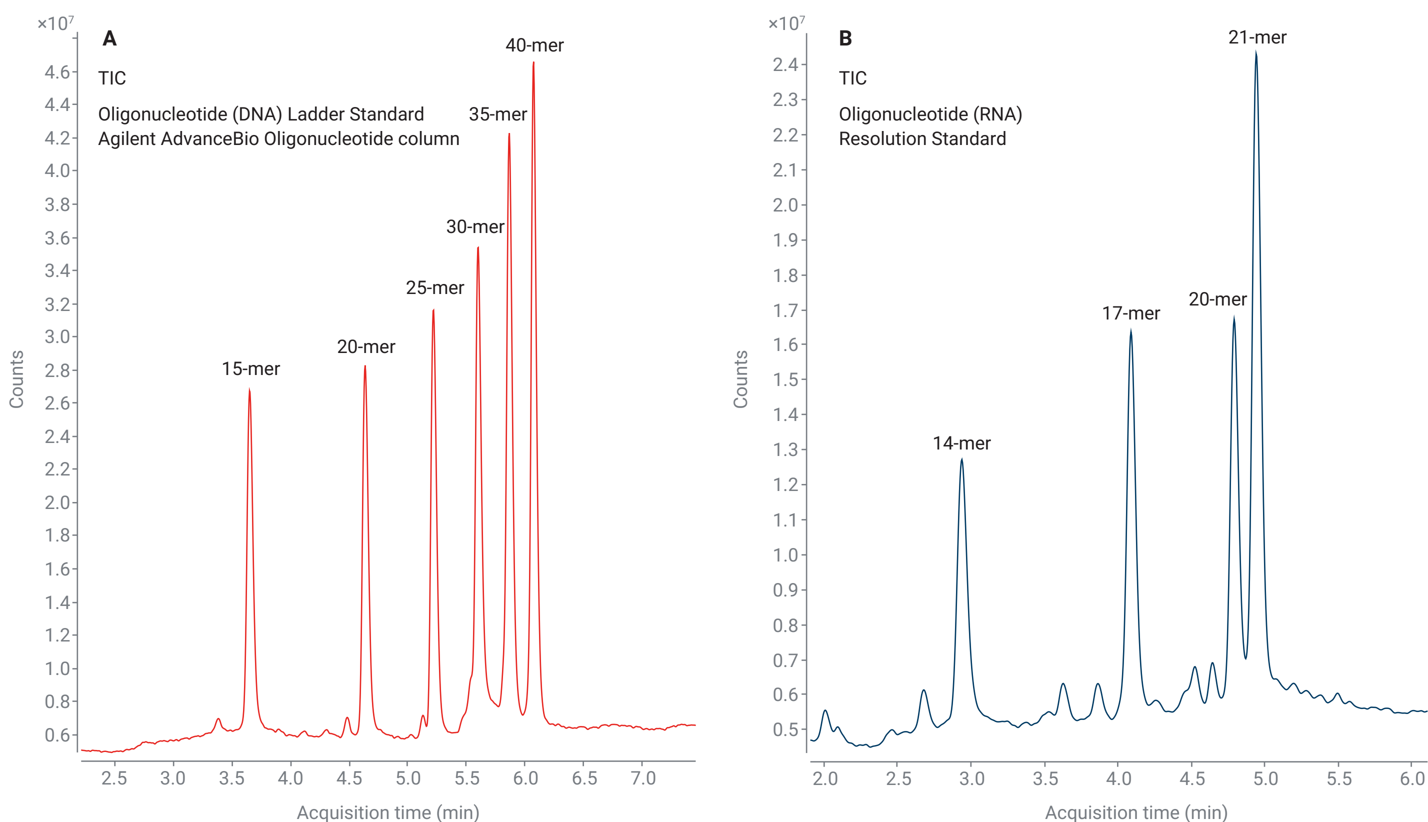


Figure 3. LC/MS Analysis of Agilent Oligonucleotide Ladder Standard (DNA) and Agilent Resolution Standard (RNA).

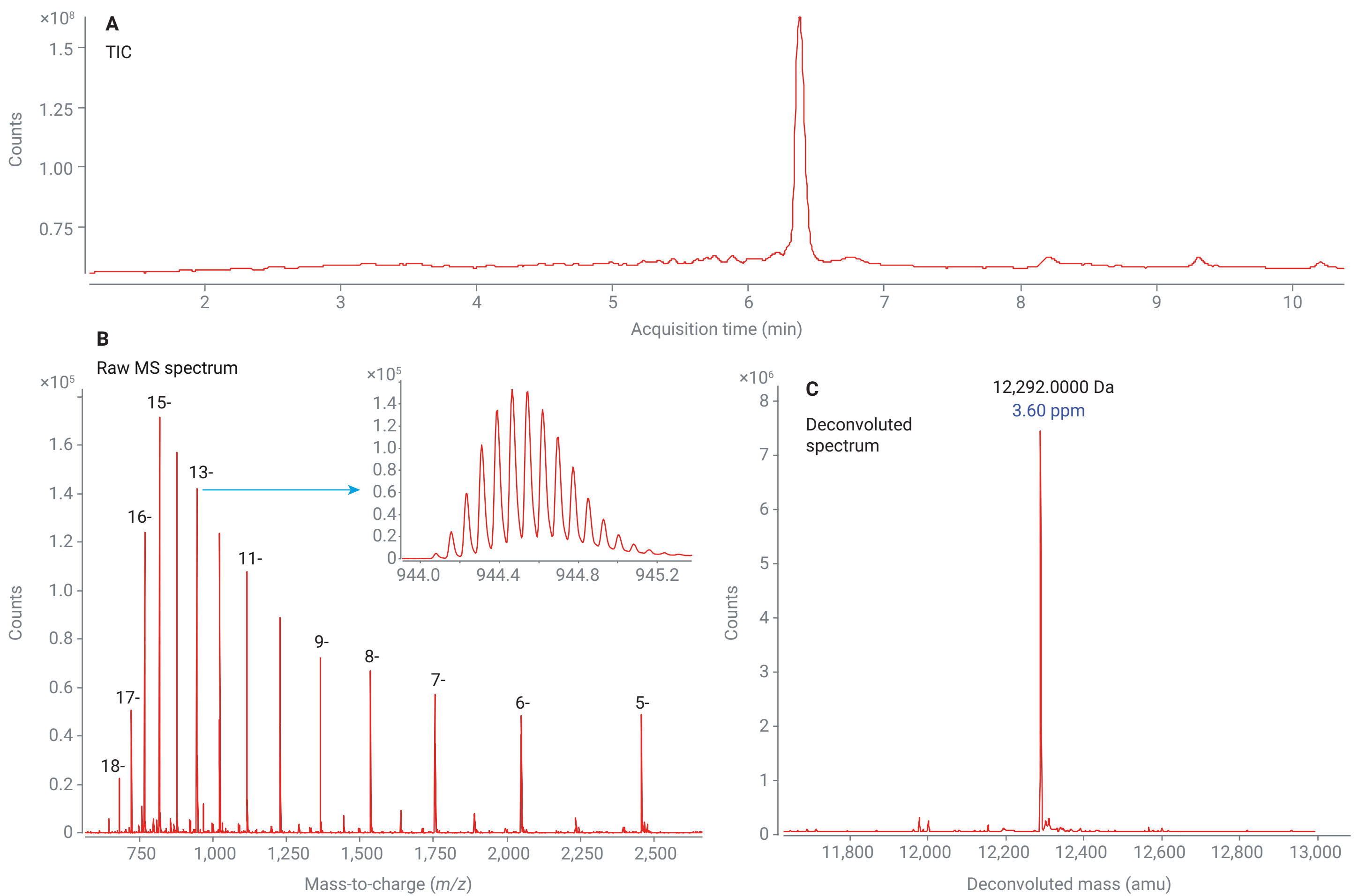


Figure 4. LC/MS analysis of synthetic oligonucleotide (40-mer). (A) Total ion chromatography (TIC) of the 40-mer oligonucleotide. (B) Raw MS spectrum of the 40-mer. (C) Deconvoluted MS spectra of the 40-mer.

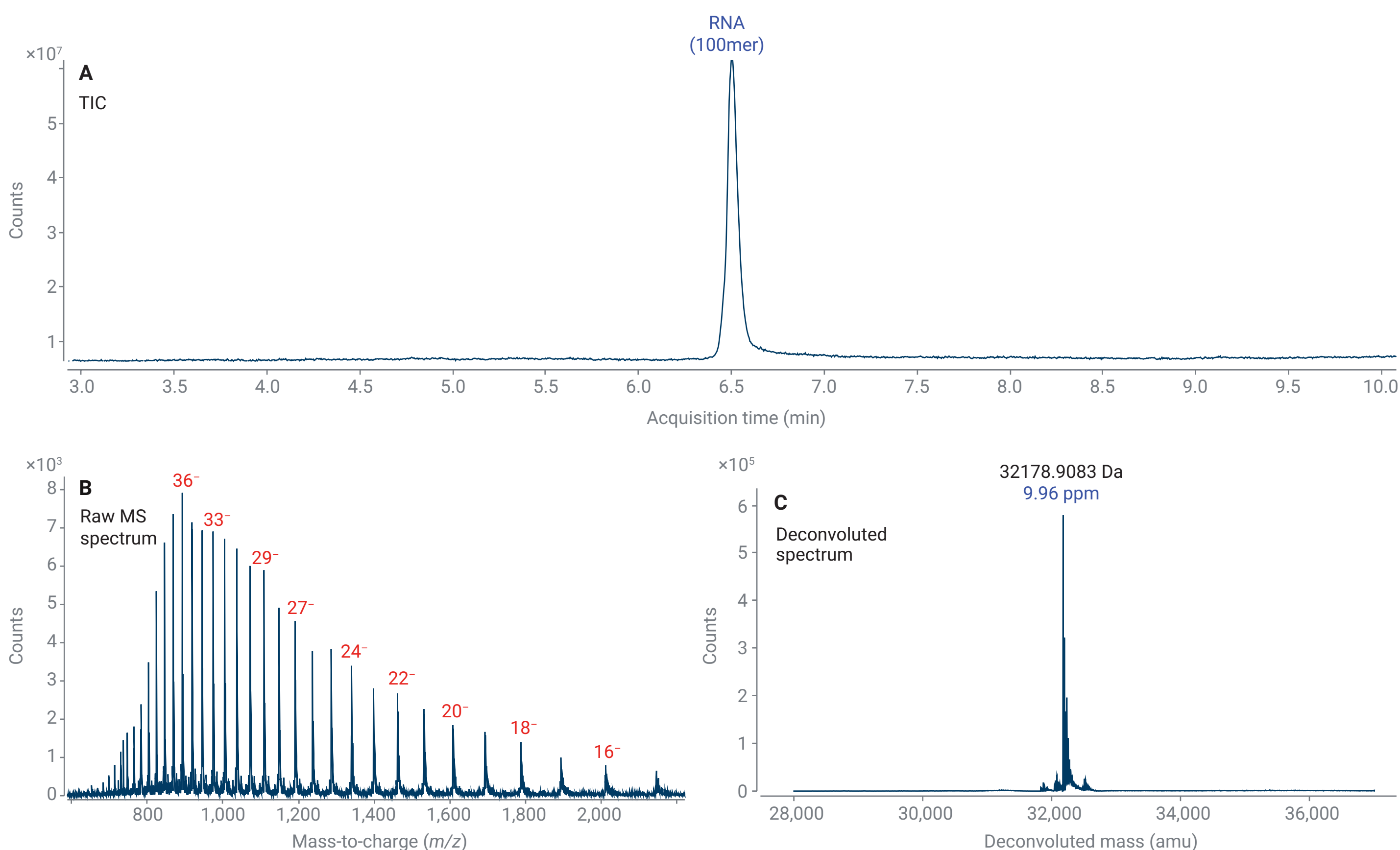


Figure 5. LC/MS analysis of a synthetic oligonucleotide (RNA, 100-mer).

Table 3. List of oligonucleotides analyzed. Calculated masses highlighted in green are monoisotopic masses (matched using FBF) and the numbers highlighted in blue are average masses (matched using Maximum Entropy deconvolution). Overall, excellent mass accuracies were obtained on all oligonucleotide samples analyzed.

Oligonucleotide	Oligo Length	Sequence	Calculated Mass (Da)	Measured Mass (Da)	Mass Accuracy (ppm)
Oligonucleotide (DNA) Ladder Standard	15	TTTTT TTTTT TTTTT	4,498.7348	4,498.7319	-0.64
	20	TTTTT TTTTT TTTTT TTTTT	6,018.9650	6,018.9635	-0.25
	25	TTTTT TTTTT TTTTTTTTTT TTTTT	7,539.1952	7,539.1989	0.50
	30	TTTTT TTTTT TTTTT TTTTT TTTTT TTTTT	9,063.8431	9,063.7988	-4.89
	35	TTTTT TTTTT TTTTT TTTTT TTTTT TTTTT TTTTT	10,584.8111	10,584.8065	-0.43
	40	TTTTT TTTTT TTTTT TTTTT TTTTT TTTTT TTTTT TTTTT	12,105.7790	12,105.8295	4.17
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
DNA-21	21	CAGTCGATTGTACTGTACTTA	6,408.0961	6,408.0952	-0.14
DNA-40	40	CCACGACCAAGTGACAGCAATGAATCGAGTCGAGATCCAT	12,291.9558	12,292.0000	3.60
RNA Standard (Long)	100	AACACCACCAUACAGUGCAGGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCUUU	32,178.5878	32,178.9083	9.96

Target Plus Impurities (TPI) data analysis workflow

While getting good MS results is necessary, it is equally essential to have a powerful software program to interpret the results. Agilent has developed novel and automated MassHunter BioConfirm 12.0 software that supports both Find-by-Formula and Maximum Entropy algorithms for the identification of species present. These Target Plus Impurities (TPI) workflows have been developed in BioConfirm 12.0, alongside a Sequence Confirmation workflow that uses MS/MS data. This application note demonstrates the unique features of the TPI workflow. The sequencing workflow is described in a separate application note.

The TPI workflow in BioConfirm 12.0 uses the oligonucleotide MS data to profile (i.e., identities and relative quantitation) the target oligonucleotide and related impurities. Figure 6 highlights the user interface of the TPI workflow, where the user defines the target oligo sequences, the potential modifications, and the matching rules (5' or 3' Truncation, Deletion, or Split) for data processing. Detailed results, including the identified target and its oligonucleotide impurities, can be displayed in multiple windows/tables format.

Two options for TPI workflow

The TPI workflow can be run using FBF (a targeted approach) or Maximum Entropy Deconvolution (an untargeted approach), where the Workflow Transition Mass function determines which algorithm is used for each sample. FBF will be run if the mass of the sample sequence is smaller than the Workflow Transition Mass. Maximum Entropy deconvolution will be run if the mass of the sample sequence is equal to (or greater than) the Workflow Transition Mass. This user-defined value can be informed by multiple factors, including the resolution of the MS instrument, the masses of the target and impurities, the preference of a targeted versus untargeted analysis, and mass accuracy requirements.

Table 3 shows the LC/MS analysis summary of various intact oligonucleotide samples, with their measured masses and mass accuracies. Sub-ppm mass accuracies were obtained for many of the oligonucleotide sequences shorter than 30-mer, using the FBF algorithm. For larger oligo samples, low ppm in mass error was also achieved by running the Maximum Entropy algorithm.

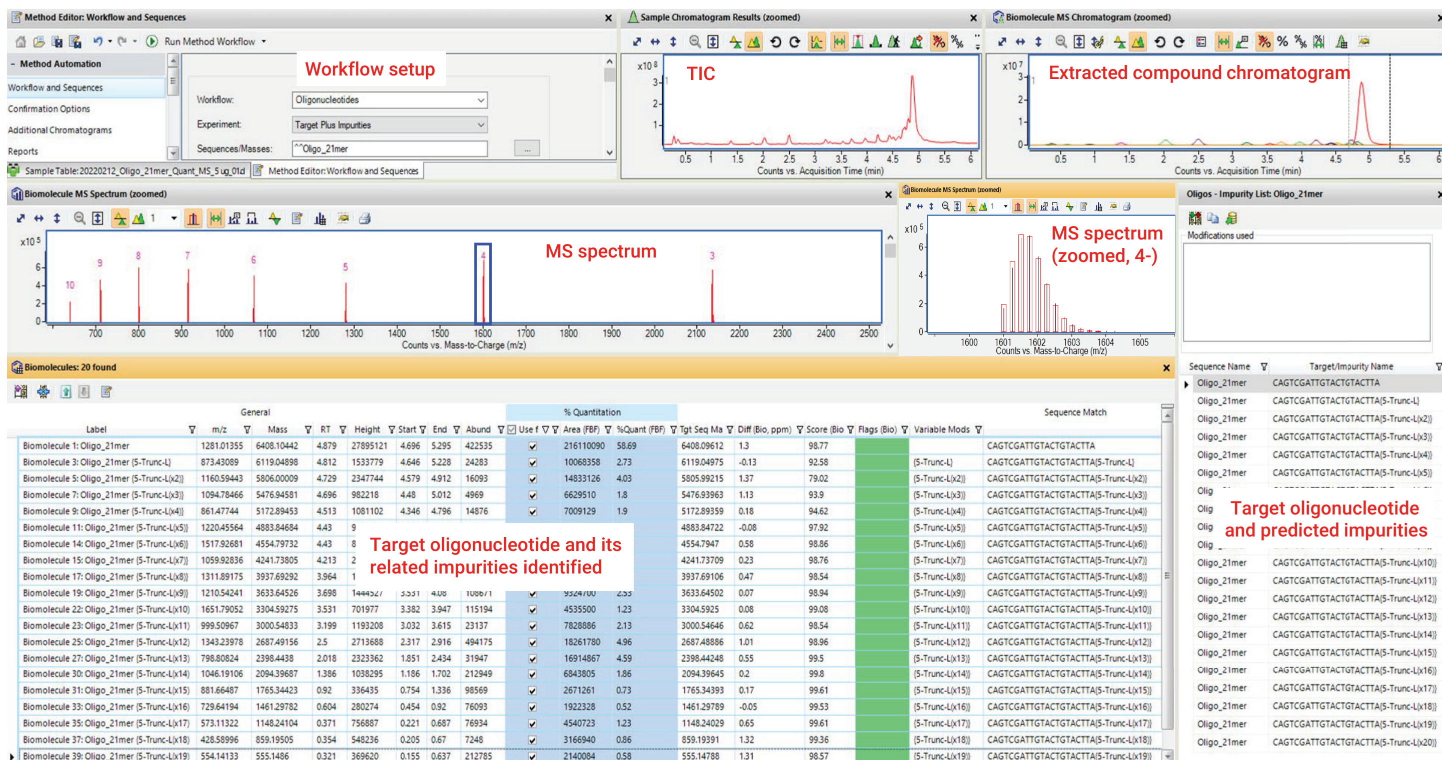


Figure 6. Overview of Agilent BioConfirm software, version 12.0 with Target Plus Impurities (TPI) workflow.

Oligonucleotide impurity analysis

Oligonucleotide impurity analysis is a critical task to characterize product-related impurities of synthetic oligonucleotides. Due to imperfect chemical coupling efficiency in the oligonucleotide synthesis, many types of impurities, such as truncations, additions, and abasic oligonucleotides have been reported.[2] There are analytical challenges to maintaining a comprehensive profile of all product-related impurities. Chromatographic separation and good MS sensitivity are desirable for detecting low level impurities.

This study developed a rapid LC/MS method for characterization of a full-length oligonucleotide target and its impurities. Figure 7 demonstrates the LC/MS profile of a 21-mer synthetic oligonucleotide and a set of related impurities that were determined using FBF. The majority of 5'-truncated impurities of this 21-mer were well separated with a short, 11-minute LC gradient. In fact, these impurities and the target were eluted within 6 minutes (Figure 7A). Data analysis using the BioConfirm 12.0 Find-by-Formula algorithm provided accurate monoisotopic masses, and the relative quantitation results for all targeted impurities, as shown in Figure 7B. As this 21-mer synthetic oligonucleotide sample was not further purified after the initial desalting process, almost all possible 5'-truncated impurities of the target oligonucleotide were identified under 0.5% in relative quantitation.

Table 4 summarizes all 19 oligonucleotide impurities that were targeted for detection with excellent deconvoluted mass accuracies (most at the sub-ppm level) and relative quantitation reproducibility. Note that many other types of target-related impurities, such as 5'-truncates with linker, or 3'-truncates (with or without linker) were also detected in very low abundance levels (data not shown).

To identify a series of 5' truncations on a 40-mer oligonucleotide, Maximum Entropy deconvolution was used. The majority of 5'-truncated impurities of a 40-mer oligonucleotide sample were identified, and the top 10 impurities (as well as the 40-mer target) are shown in Figure 8. The details of the relative quantitation analysis using the Maximum Entropy approach are listed in Table 5.

Similar to the FBF results, high mass accuracies and accurate relative quantitation results were achieved for a larger oligonucleotide molecule (40-mer). The results also demonstrate excellent sensitivity in the detection of low-level oligonucleotide impurities. Among the top 15 impurities of the 40-mer oligonucleotide sample, as low as 0.65% in relative quantitation was obtained (Table 5). As described, many other types of target-related impurities were also detected in very low abundance levels (data not shown).

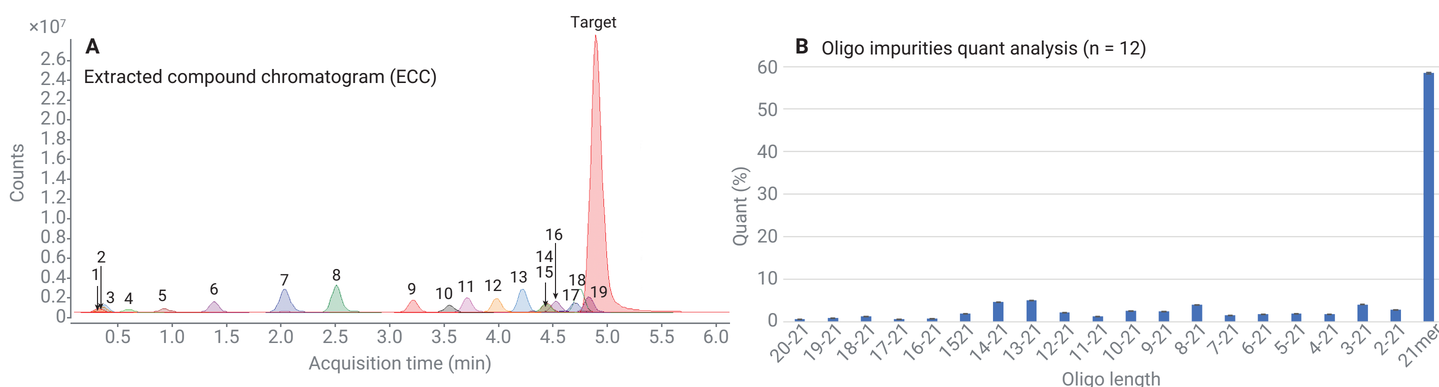


Figure 7. Relative quantification analysis of synthetic oligonucleotide (21-mer) and targeted impurities by the Find-by-Formula algorithm of Agilent BioConfirm software. (A) Extracted compound chromatography of the 21-mer oligonucleotide and its impurities. (B) Relative quantitation analysis results of the 21-mer oligonucleotide and its impurities. Excellent reproducibility with very low RSD (<3%) was achieved over a total of 12 sample injections.

Table 4. Impurity analysis summary on 19 oligonucleotide impurities of a 21-mer synthetic oligonucleotide (n = 12).

Impurity Peak	Oligo Length	RT (min)	Calculated Mono Mass	Measured Mass	Avg Mass Accuracy (ppm)(n = 12)	Avg % Quant (n = 12)	Std Dev	RSD (%)	Sequence
1	20–21	0.321	555.1479	555.1486	1.21	0.57	0.01	2.39	TpA
2	19–21	0.354	859.1939	859.1950	1.09	0.89	0.02	1.76	TpTpA
3	18–21	0.371	1,148.2403	1,148.2410	0.81	1.28	0.02	1.44	CpTpTpA
4	17–21	0.604	1,461.2979	1,461.2978	0.15	0.53	0.01	1.18	ApCpTpTpA
5	16–21	0.920	1,765.3439	1,765.3442	0.57	0.72	0.01	1.72	TpApCpTpTpA
6	15–21	1.386	2,094.3964	2,094.3969	0.65	1.86	0.01	0.66	GpTpApCpTpTpA
7	14–21	2.018	2,398.4425	2,398.4438	0.69	4.61	0.04	0.91	TpGpTpApCpTpTpA
8	13–21	2.500	2,687.4889	2,687.4916	0.73	4.98	0.04	0.72	CpTpGpTpApCpTpTpA
9	12–21	3.199	3,000.5465	3,000.5483	0.33	2.14	0.02	0.75	ApCpTpGpTpApCpTpTpA
10	11–21	3.531	3,304.5925	3,304.5928	0.04	1.23	0.01	1.05	TpApCpTpGpTpApCpTpTpA
11	10–21	3.698	3,633.6450	3,633.6453	0.21	2.55	0.02	0.93	GpTpApCpTpGpTpApCpTpTpA
12	9–21	3.964	3,937.6911	3,937.6929	0.15	2.45	0.02	0.76	TpGpTpApCpTpGpTpApCpTpTpA
13	8–21	4.213	4,241.7371	4,241.7380	0.47	3.97	0.03	0.70	TpTpGpTpApCpTpGpTpApCpTpTpA
14	7–21	4.430	4,554.7947	4,554.7973	0.33	1.49	0.01	0.97	ApTpTpGpTpApCpTpGpTpApCpTpTpA
15	6–21	4.430	4,883.8472	4,883.8468	−0.23	1.74	0.01	0.85	GpApTpTpGpTpApCpTpGpTpApCpTpTpA
16	5–21	4.513	5,172.8936	5,172.8945	0.05	1.91	0.02	0.80	CpGpApTpTpGpTpApCpTpGpTpApCpTpTpA
17	4–21	4.696	5,476.9396	5,476.9458	1.22	1.81	0.01	0.79	TpCpGpApTpTpGpTpApCpTpGpTpApCpTpTpA
18	3–21	4.729	5,805.9921	5,806.0001	1.51	4.04	0.04	0.99	GpTpCpGpApTpTpGpTpApCpTpGpTpApCpTpTpA
19	2–21	4.812	6,119.0498	6,119.0490	−0.32	2.75	0.05	1.92	ApGpTpCpGpApTpTpGpTpApCpTpGpTpApCpTpTpA
Target	21-mer	4.879	6,408.0961	6,408.1044	1.29	58.49	0.20	0.34	CpApGpTpCpGpApTpTpGpTpApCpTpGpTpApCpTpTpA

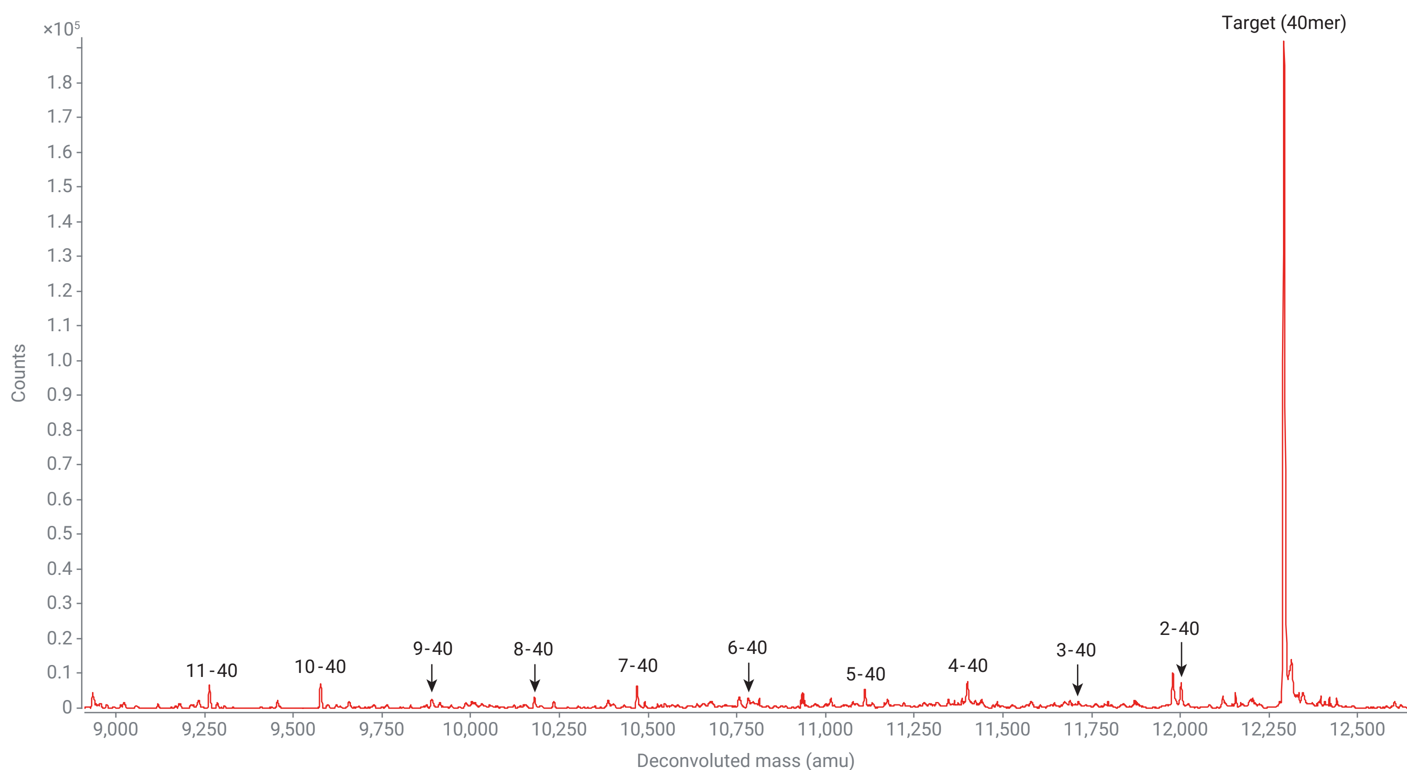


Figure 8. Relative quantification of a 40-mer and its 5' truncations (2-40 to 11-40) using the Maximum Entropy MS Deconvolution approach.

Table 5. Impurity analysis summary on the top 15 oligonucleotide impurities of a 40-mer synthetic oligonucleotide.

Oligo Length	RT (min)	Measured Mass	%Quant	Sequence
16–40	5.861	7,699.0915	2.08	AGCAATGAATCGAGTCGAGATCCAT
15–40	5.878	7,988.2299	2.26	CAGCAATGAATCGAGTCGAGATCCAT
14–40	5.994	8,301.3760	1.24	ACAGCAATGAATCGAGTCGAGATCCAT
13–40	5.990	8,631.0440	1.35	GACAGCAATGAATCGAGTCGAGATCCAT
12–40	6.048	8,935.0245	1.60	TGACAGCAATGAATCGAGTCGAGATCCAT
11–40	6.073	9,263.8039	2.39	GTGACAGCAATGAATCGAGTCGAGATCCAT
10–40	6.148	9,577.3536	2.54	AGTGACAGCAATGAATCGAGTCGAGATCCAT
9–40	6.193	9,891.1191	0.94	AAGTGACAGCAATGAATCGAGTCGAGATCCAT
8–40	6.185	10,179.5068	1.18	CAAGTGACAGCAATGAATCGAGTCGAGATCCAT
7–40	6.214	10,468.9206	2.36	CCAAGTGACAGCAATGAATCGAGTCGAGATCCAT
6–40	6.289	10,782.0017	1.08	ACCAAGTGACAGCAATGAATCGAGTCGAGATCCAT
5–40	6.276	11,111.0793	2.03	GACCAAGTGACAGCAATGAATCGAGTCGAGATCCAT
4–40	6.280	11,400.2121	2.82	CGACCAAGTGACAGCAATGAATCGAGTCGAGATCCAT
3–40	6.334	11,712.5835	0.65	ACGACCAAGTGACAGCAATGAATCGAGTCGAGATCCAT
2–40	6.339	12,002.3602	2.68	CACGACCAAGTGACAGCAATGAATCGAGTCGAGATCCAT
Target	6.384	12,292.1749	72.80	CCACGACCAAGTGACAGCAATGAATCGAGTCGAGATCCAT

Conclusion

Two novel, automated, and integrated oligonucleotide data analysis approaches were developed for the characterization of the target oligonucleotide and its related impurities using HRAM MS data. The analytical results demonstrate that excellent chromatographic separation and mass accuracy (sub-ppm) for expected oligonucleotides were achieved. The LC/MS results also show accurate relative quantification of the observed oligonucleotides and their impurities, with great reproducibility. The newly developed Agilent MassHunter BioConfirm software (version 12.0) enabled automated TPI data processing in a high-throughput manner, which significantly reduced data analysis times.

References

1. Capaldi, D. *et al.* Impurities in Oligonucleotide Drug Substances and Drug Products. *Nucleic Acid Ther.* **2017**, *27*, 309–322.
2. Okafo, G.; Elder, D.; Webb, M. Analysis of Oligonucleotides and Their Related Substances. *ILM Publications* **2012**.

www.agilent.com

DE29613125

This information is subject to change without notice.

© Agilent Technologies, Inc. 2022

Printed in the USA, May 13, 2022
5994-4817EN

Agilent
Trusted Answers