

Poster Reprint

ASMS 2022 Poster number WP386

# Comprehensive, Automated, and Integrated Software for Oligonucleotide Characterization and Sequence Confirmation

David Wong<sup>1</sup>; Peter Rye<sup>2</sup>; Stephen Madden<sup>1</sup>; Gordon Slysz<sup>1</sup> and Crystal Cody<sup>1</sup>

<sup>1</sup>Agilent Technologies, Inc., Santa Clara, CA <sup>2</sup>Agilent Technologies, Inc., Lexington, MA

#### 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 has come the increased need for robust analytical methods and easy-to-use data analysis workflows to characterize them. Critical quality attributes of oligo samples include the determination of the target, confirmation of its sequence, and identification of impurities present – each of which can be time consuming, difficult, and tedious. As such, software that supports and automates these efforts can be of great value. In this work, we present novel, automated, and integrated software to support these workflows using HRAM MS data.

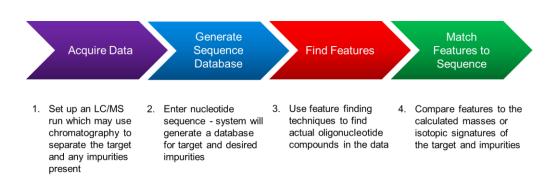


Figure 1. Target Plus Impurities (TPI) data analysis workflow in Agilent MassHunter BioConfirm 12.0 software



#### 1. Set up an LC/MS run 2. Enter nucleotide sequence 3. Compare theoretical

#### Experimental

#### **Oligonucleotide Samples**

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.

21-mers and a 40-mer oligonucleotide were purchased from Integrated DNA Technologies, Inc. (Coralville, IA, USA) with standard desalting purification.

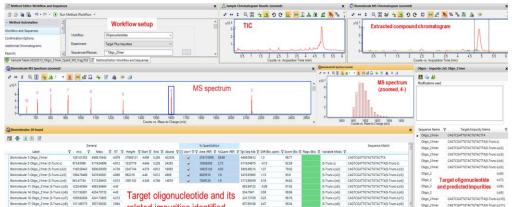
Most oligonucleotide samples were dissolved with DI water to 0.50 mg/mL without further purification.

#### **Oligonucleotide Samples Analysis**

LC/MS analyses of oligonucleotides were conducted on the Agilent 1290 Infinity II LC coupled to the 6545XT AdvanceBio LC/Q-TOF system. The AdvanceBio Oligonucleotides column was used with mobile phases containing traditional ion-pairing additives (triethylamine and hexafluoroisopropanol).



oligonucleotide analysis - Target Plus Impurities (TPI) and Sequence Confirmation workflows.



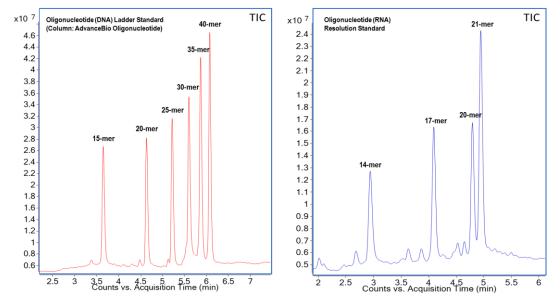
- I. Set up an LC/MS run which will generate ion fragments using MS/MS for the target sequence
- Enter nucleotide sequence - system will automatically generate a theoretical list for all possible oligo fragment ion types
- Compare theoretical fragments to acquired MS/MS spectra; annotate with fragments matched

edutoischis (Uroliĝo"s unis (s-unis-citel)	1201001112	323/ 83232	3,304			1	m · · · ······························		i al a m tit		921/031/0	(UA)	20.34	(b-unic-cital)	CHOICOHILOTHEIGHEITHIS-UDIC-ENGI	Oligo 21mer	CASTCGATTGTACTGTACTTA/5-Trunc-Lix111
Biomolecule 19: Oligo_21mer (5-Trunc-L(x9))	1210.54241	1633.64526	3.690	rela	nec	1 111	punt	les	identif	led	633.64502	0.07	98.94	(5-Trunc-L(x9)	CAGTCGATTGTACTGTACTTA(5-Trunc-L)(8)	Oligo 21mer	CAGTCGATTGTACTGTACTTAG-Trunc-Lix12]
Biomolecule 22: Oligo_21mer (5-Trunc-L(x10)	1651,79052	3304,59275	3.531								304.5925	0.05	99.08	(5-Trunc-L(x10)	CAGTCGATTGTACTGTACTTA(5-Trunc-L(x10)	Oligo_21mer	CAGTCGATTGTACTGTACTTA/5-Trunc-Liv131
Biomolecule 23: Oligo, 21mer (5-Trunc-4)(11)	999.50967	3000.54833	3.199	1193206	3.032	3.615	23137	*	7828886	2.13	3000.54646	0.62	90.54	(5-Trute-L(x11))	CAGTCGATTGTACTGTACTTA(5-Trunc-L(x11))		
Biomolecule 25: Oligo_21mer (5-Trunc-L(x12)	1343,23978	2687,49156	2.5	2713688	2317	2.916	494175		15251780	4.96	2687.45886	1.01	98.96	(5-Trunc-L(x12))	CAGTCGATTGTACTGTACTTA(5-Trunc-L)x12)	Oligo_21mer	CAGTCGATTGTACTGTACTTA(5-Trunc-L(x14))
Biomolecule 27: Oligo_21mer (5-7runc-4)(13)	798.60824	2395.4438	2,018	2323362	1.851	2.434	31947	~	16914867	4.59	2398.44248	0.55	99.5	(5-Trunc-4)(13)	CAGTCGATTGTACTGTACTTA(5-Trunc-L(x13)	Oligo_21mer	CAGTCGATTGTACTGTACTTA(5-Trunc-L)(15)
Biomolecule 30: Oligo_21mer (5-Trunc-L)x14)	1046.19106	2094,39687	1.386	1038295	1.186	1.702	212949		6843805	1.56	2094.39645	0.2	99.5	(5-Trunc-L)(14)	CAGTCGATTGTACTGTACTTAS-Trunc-Lix14()	Oligo_21mer	CAGTCGATTGTACTGTACTTA(5-Trunc-L(x16))
Biomolecule 31: Oligo_21mer (5-7runc-L(x15)	881.66487	1765.34423	0.92	336435	0.754	1.336	98569		2671261	0.73	1765.34393	0.17	99.61	(5-7/unc-L)/15)	CAGTCGATTGTACTGTACTTA(5-Trunc-L(x15)	Oligo_21mer	CAGTCGATTGTACTGTACTTA(5-Trunc-L(x17))
Biomolecule 33: Oligo_21mer (5-Trunc-L)x16)	729.64194	1461.29782	0.604	280274	0.454	0.92	76093		1922328	0.52	1461,29789	-0.05	99.53	(5-Trunc-L(x16))	CAGTCGATTGTACTGTACTTA(5-Trunc-Lix16)	Oligo_21mer	CAGTCGATTGTACTGTACTTA(5-Trunc-L(x18))
Biomolecule 35: Oligo_21mer (5-7runc-L(x17)	573.11322	1145,24104	0.371	756887	0.221	0.687	76934		4540723	1.23	1148,34029	0.65	99.61	(5-Trunc-L(x17))	CAGTCGATTGTACTGTACTTA(5-Trunc-L(x17))	Oligo_21mer	CAGTCGATTGTACTGTACTTA(5-Trunc-L)(19)
Biomolecule 37: Olgo_21mer (5-Trunc-L(x18)	428.58996	859.19505	0.354	545236	0.205	0.67	7248		3166940	0.86	859.19291	1.32	99.36	(5-Trunc-L(x18))	CAGTCGATTGTACTGTACTTA(5-Trunc-L(x18)	Oligo_21mer	CAGTCGATTGTACTGTACTTAS-Trunc-Lix20]
Biomolecule 39: Oligo_21mer (5-Trunc-L)x19)	554.14133	555.1486	0.321	369620	0.155	0.637	212785	*	2140084	0.58	555.14788	1.31	98.57	(5-Trunc-L(x19))	CASTCSATTSTACTSTACTTAS-Trunc-L(x19)		

Figure 2. Oligonucleotide Sequence Confirmation data analysis workflow in Agilent MassHunter BioConfirm 12.0 software Figure 4. Overview of Agilent BioConfirm software, version 12.0 with Target Plus Impurities (TPI) workflow.

2

#### **Results and Discussion**



#### HPLC separation of various oligonucleotide standards

Figure 5. LC/MS Analysis of Agilent Oligonucleotide Ladder Standard (DNA) and Agilent Resolution Standard (RNA).

#### LC/MS analysis of synthetic oligonucleotides

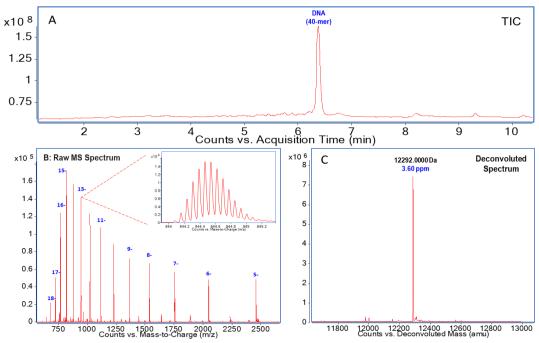
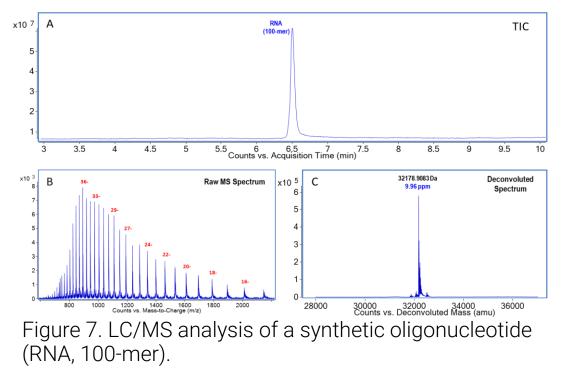


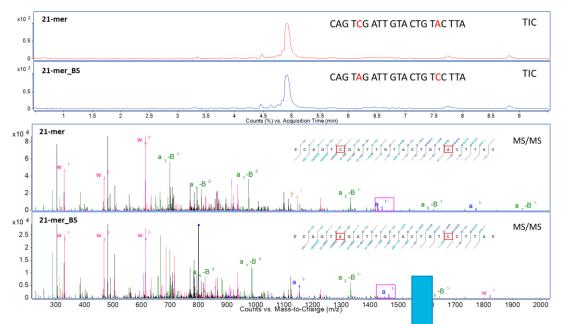
Figure 6. LC/MS analysis of synthetic oligonucleotide (DNA, 40-mer).

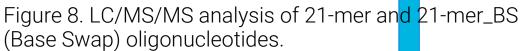


Oligonucleotide	Oligo Length	Sequence	Cal. Mass (Da)	Measured Mass (Da)	Mass Accuracy (ppm)
Oligonucleotide (DNA) Ladder Standard	15 20 25 30 35 40	TTTTT TTTTT TTTTT TTTTT TTTTT TTTTT TTTTT TTTTT TTTTTT TTTTT TTTTT TTTTTTTT	4498.7348 6018.9650 7539.1952 9063.8431 10584.8111 12105.7790	4498.7319 6018.9635 7539.1989 9063.7988 10584.8065 12105.8295	-0.64 -0.25 0.50 -4.89 -0.43 4.17
Oligonucleotide (RNA) Resolution Standard	14 17 20 21	rCrArCrUrGrArArUrArCrCrArArU rUrCrArCrArCrUrGrArArUrArCrCrArArU rUrCrArUrCrArCrArCrUrGrArArUrArCrCrArArU rGrUrCrArUrCrArCrArCrUrGrArArUrArCrCrCrArArU	4395.6479 5335.7670 6275.8861 6620.9335	4395.6429 5335.7623 6275.8800 6620.9263	-1.14 -0.88 -0.97 -1.09
DNA-21	21	CAGTCGATTGTACTGTACTTA	6408.0961	6408.0952	-0.14
DNA-40	40	CCACGACCAAGTGACAGCAATGAATCGAGTCGAGATCCAT	12291.9558	12292.0000	3.60
RNA Standard (Long)	100	AACACCACCAUACAGUGCAGGUUUUAGAGCUAGAAAUAG CAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGCUUUU	32178.5878	32178.9083	9.96

Table 1. 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).

## LC/MSMS analysis and sequence confirmation of synthetic oligonucleotides





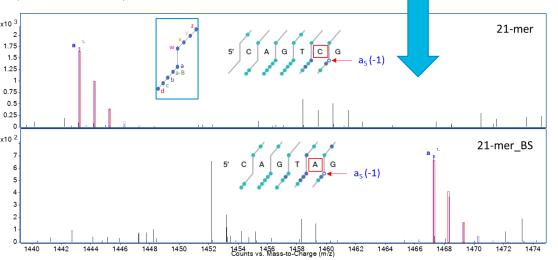


Figure 9. Sequence confirmation of 21-mer and 21mer\_BS (Base Swap) oligonucleotides (zoom-in spectrum).

3

#### **Results and Discussion**

### LC/MSMS analysis and sequence confirmation of a heavily modified oligonucleotide

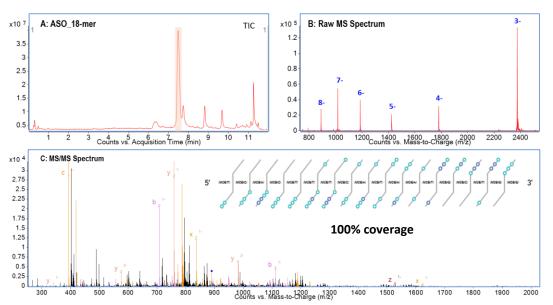
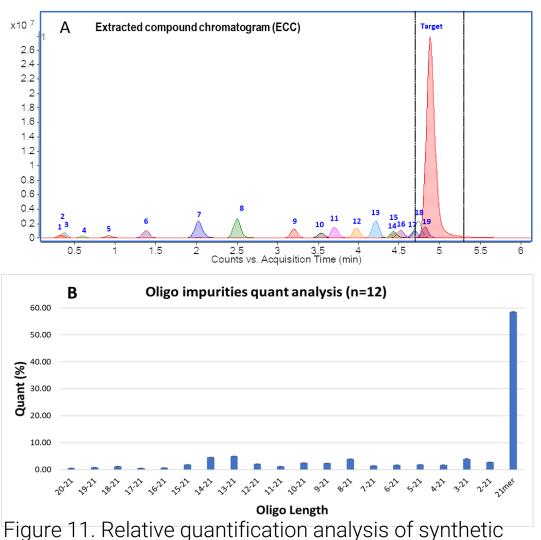


Figure 10. LC/MSMS analysis of a heavily modified oligonucleotide.

### **Oligonucleotides impurity analysis**



oligonucleotide (21-mer) and targeted impurities by the

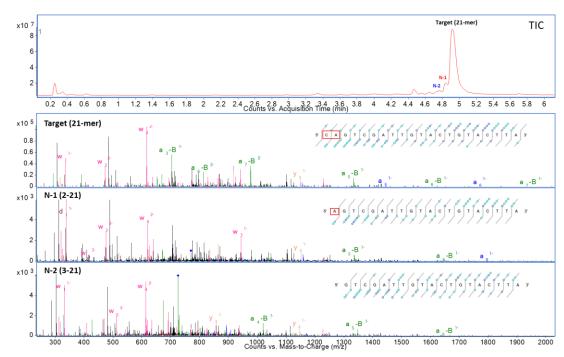


Figure 12. Sequence confirmation of 21-mer and its related impurities (N-1 and N-2).

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	ТрА
2	19-21	0.354	859.1939	859.1950	1.09	0.89	0.02	1.76	ТрТрА
3	18-21	0.371	1148.2403	1148.2410	0.81	1.28	0.02	1.44	СрТрТрА
4	17-21	0.604	1461.2979	1461.2978	0.15	0.53	0.01	1.18	АрСрТрТрА
5	16-21	0.920	1765.3439	1765.3442	0.57	0.72	0.01	1.72	ТрАрСрТрТрА
6	15-21	1.386	2094.3964	2094.3969	0.65	1.86	0.01	0.66	GpTpApCpTpTpA
7	14-21	2.018	2398.4425	2398.4438	0.69	4.61	0.04	0.91	ТрGpTpApCpTpTpA
8	13-21	2.500	2687.4889	2687.4916	0.73	4.98	0.04	0.72	CpTpGpTpApCpTpTpA
9	12-21	3.199	3000.5465	3000.5483	0.33	2.14	0.02	0.75	ApCpTpGpTpApCpTpTpA
10	11-21	3.531	3304.5925	3304.5928	0.04	1.23	0.01	1.05	ТрАрСрТрGрТрАрСрТрТрА
11	10-21	3.698	3633.6450	3633.6453	0.21	2.55	0.02	0.93	GpTpApCpTpGpTpApCpTpTpA
12	9-21	3.964	3937.6911	3937.6929	0.15	2.45	0.02	0.76	TpGpTpApCpTpGpTpApCpTpTpA
13	8-21	4.213	4241.7371	4241.7380	0.47	3.97	0.03	0.70	TpTpGpTpApCpTpGpTpApCpTpTpA
14	7-21	4.430	4554.7947	4554.7973	0.33	1.49	0.01	0.97	ApTpTpGpTpApCpTpGpTpApCpTpTpA
15	6-21	4.430	4883.8472	4883.8468	-0.23	1.74	0.01	0.85	GpApTpTpGpTpApCpTpGpTpApCpTpTpA
16	5-21	4.513	5172.8936	5172.8945	0.05	1.91	0.02	0.80	CpGpApTpTpGpTpApCpTpGpTpApCpTpTpA
17	4-21	4.696	5476.9396	5476.9458	1.22	1.81	0.01	0.79	TpCpGpApTpTpGpTpApCpTpGpTpApCpTpTpA
18	3-21	4.729	5805.9921	5806.0001	1.51	4.04	0.04	0.99	GpTpCpGpApTpTpGpTpApCpTpGpTpApCpTpTpA
19	2-21	4.812	6119.0498	6119.0490	-0.32	2.75	0.05	1.92	ApGpTpCpGpApTpTpGpTpApCpTpGpTpApCpTpTpA
Target	21mer	4.879	6408.0961	6408.1044	1.29	58.49	0.20	0.34	CpApGpTpCpGpApTpTpGpTpApCpTpGpTpApCpTpTpA

Table 2. Impurity analysis summary on 19 oligonucleotide impurities of a 21-mer synthetic oligonucleotide (n = 12).

#### Conclusions

- A highly automated and fully integrated oligonucleotide data (MS1 and MS/MS) analysis program was developed for comprehensive characterization of the targeted oligonucleotide and its related impurities.
- Our analytical results demonstrate that excellent chromatographic separation and mass accuracy (sub-ppm)

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. for expected oligonucleotides were achieved.

- The LC/MS results also show accurate relative quantification of the observed oligonucleotides and their impurities, with very good reproducibility (RSD < 3%).
- We have achieved 100% sequence coverages on all oligonucleotides analyzed for sequence confirmation.

https://explore.agilent.com/asms

This information is subject to change without notice.

DE16130343

© Agilent Technologies, Inc. 2022 Published in USA, May 20, 2022

