Recent Advances in Oligo Purity and Sequence Determinations by LCMS: Maximizing throughput, confidence, and coverage

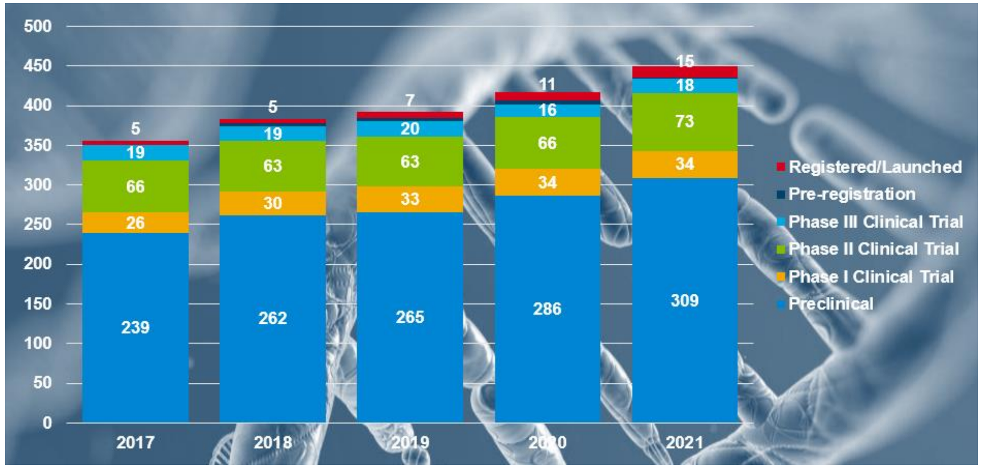
Agilent
Trusted Answers

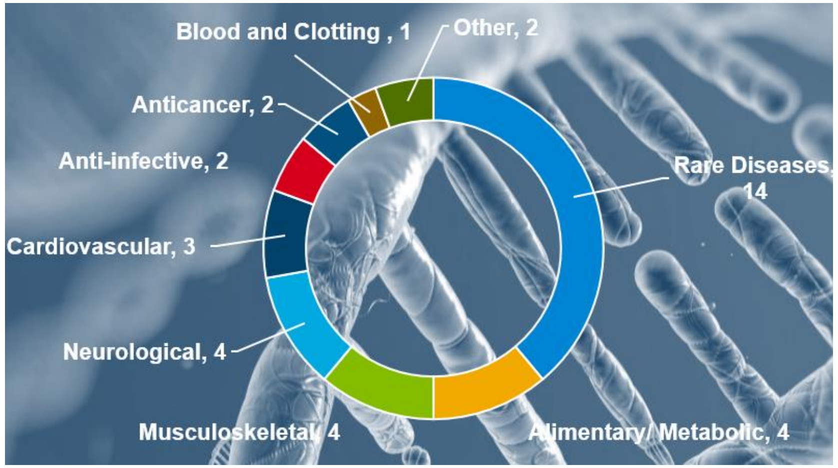
Peter Rye, Ph.D.
(peter.rye@agilent.com)

September 28, 2022

Introduction to Oligonucleotides (Oligos)

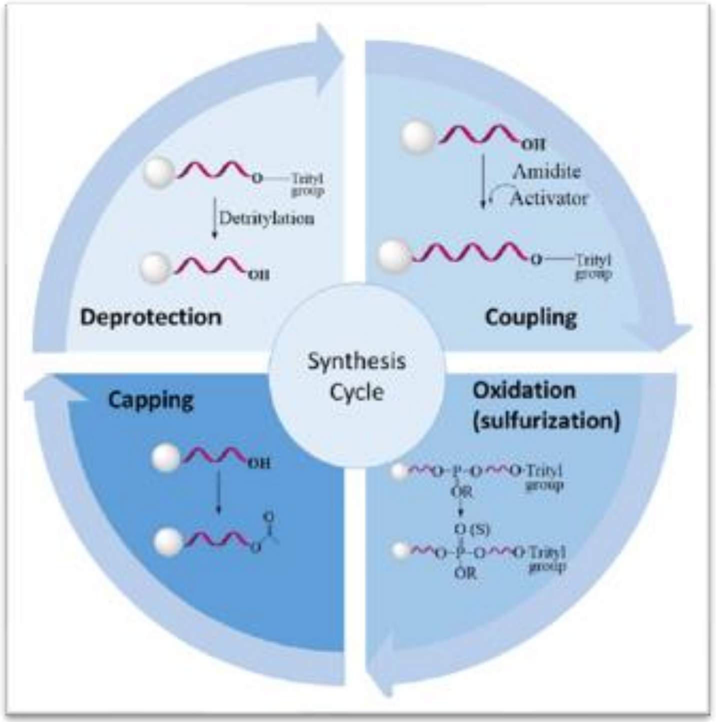
type	characteristics
miRNA	microRNA (~ 22 nt) with partial complementarity to mRNA that inhibits translation
siRNA	20 – 24 bp dsRNA forms part of RISC complex that results in slicing of mRNA (complementary)
ASO	Allele-specific, 15 – 20 nt
Aptamer	20 – 80 nt, ss
CRISPR/Cas9	gRNA ~ 100 nt
asRNA	Short (< 200 nt) or Long (> 200 nt), antisense ss complementary to mRNA
mRNA	~ 800-4500 nt





Data from Informa database

Manufacturing Challenges Associated with Synthetic Oligos



- The manufacturing batch consists of both the target and failed closely related sequences
- Imperfections in the manufacturing process leads to the formation of the impurities.

Shortmers (N-1) – oligonucleotides missing one or more nucleotides

Longmers (N+1) – oligonucleotides that include more than the intended number of nucleotides

Oligos with incomplete deprotections (benzoyl, isobutyryl, cyanoethyl, etc.

Other product related species from events including depurination, depyrimidation, oxidation, PS to PO, and nuclease degradation

Biotechnol J. 2020 Aug;15(8):e1900226. doi: 10.1002/biot.201900226. Epub 2020 May 4.

You are asked to characterize a handful of synthetic oligos for purity and confirm that the nucleotides are in the correct order.

What do you do?

Appropriate Agilent LCMS Hardware for Characterizing Oligos

BioLC

6230 LC/TOF

Oligo impurity analysis

6545XT AdvanceBio LC/QTOF

Oligo impurity and sequencing analysis

- BioLC mitigates bio-oxidation and non-specific binding of oligos throughout flow-path
- TOF/QTOF single-click tuning for all synthetic oligo types
- 5 orders of in-spectra dynamic range making it possible to see very low abundance impurities, even when they coelute
- Constant MS resolution, even when running fast chromatography

Oligo Target Plus Impurities (TPI) Workflow

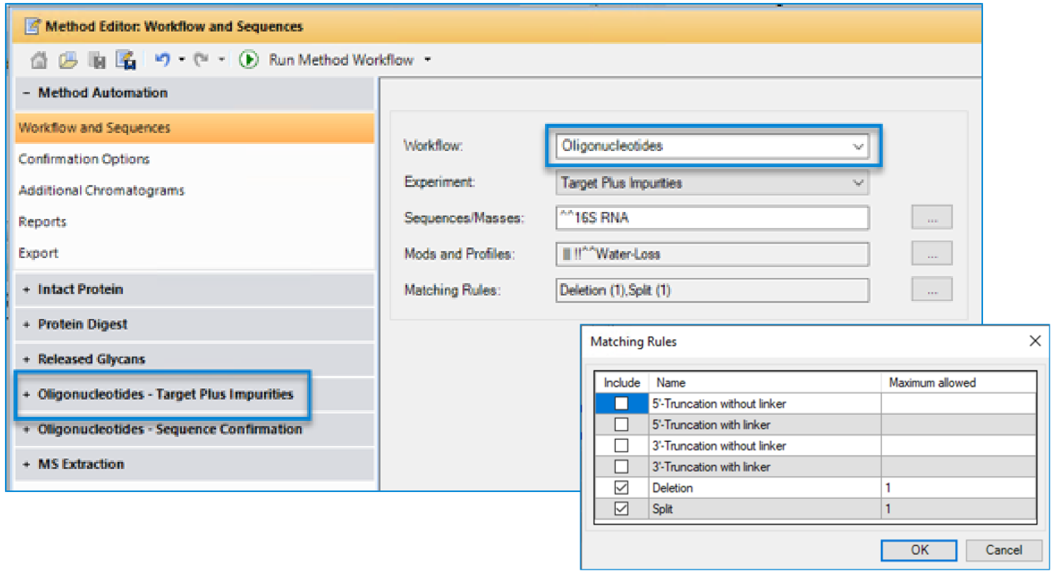
- Acquire MS1 Data
- Define Sequence(s)
- Select Impurities to look for
- Match Features to TPI list

Find by formula (FBF) for <u>targeted</u> impurity analysis looks for defined impurities.

Maximum Entropy Deconvolution for <u>untargeted</u> impurity analysis finds all peaks and matches against impurity list.

Color-coded purity results – by height or area.

Please refer to Agilent application note 5994-4817EN



Workflow: Oligonucleotides

Experiment: Target Plus Impurities

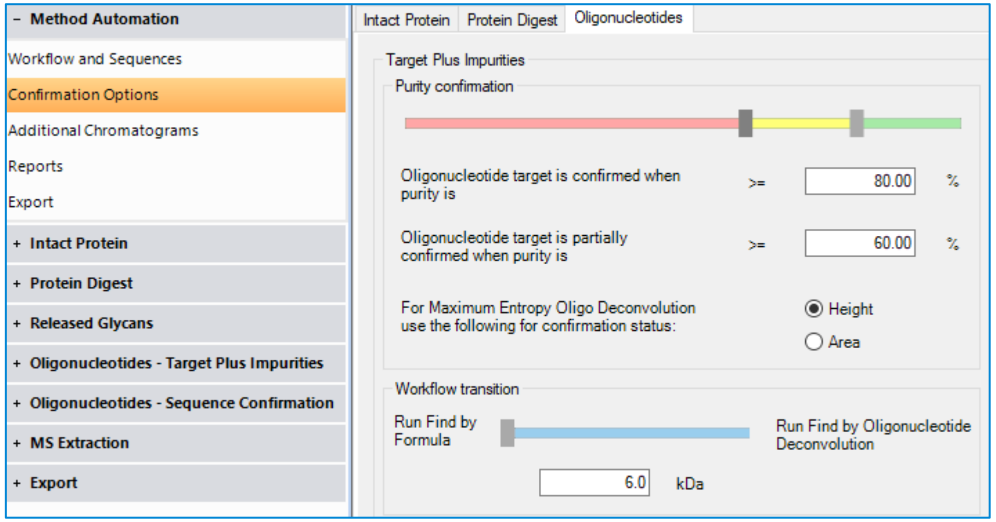
Sequences/Masses: ^^16S RNA

Mods and Profiles: !!^^Water-Loss

Matching Rules: Deletion (1),Split (1)

Matching Rules

Include	Name	Maximum allowed
☐	5'-Truncation without linker	
☐	5'-Truncation with linker	
☐	3'-Truncation without linker	
☐	3'-Truncation with linker	
☑	Deletion	1
☑	Split	1



Target Plus Impurities

Purity confirmation

Oligonucleotide target is confirmed when purity is >= 80.00 %

Oligonucleotide target is partially confirmed when purity is >= 60.00 %

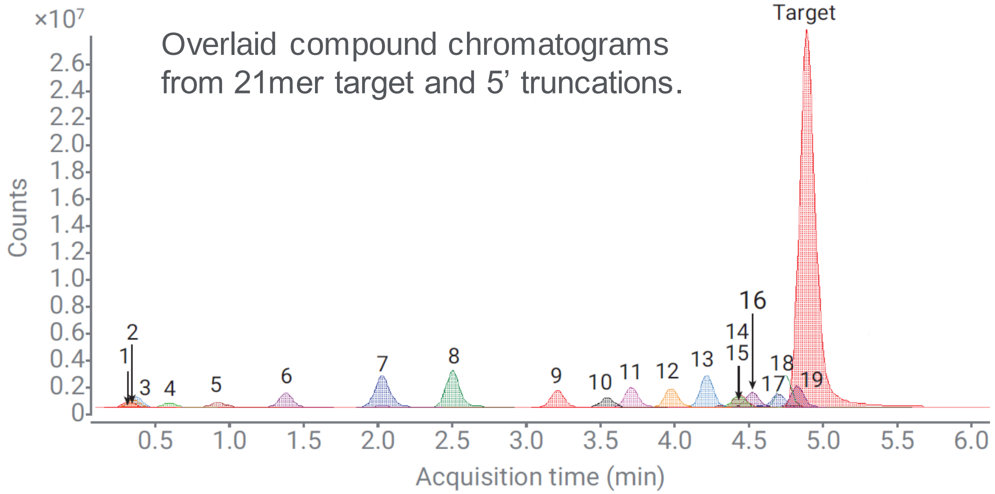
For Maximum Entropy Oligo Deconvolution use the following for confirmation status: ● Height ○ Area

Workflow transition

Run Find by Formula — Run Find by Oligonucleotide Deconvolution

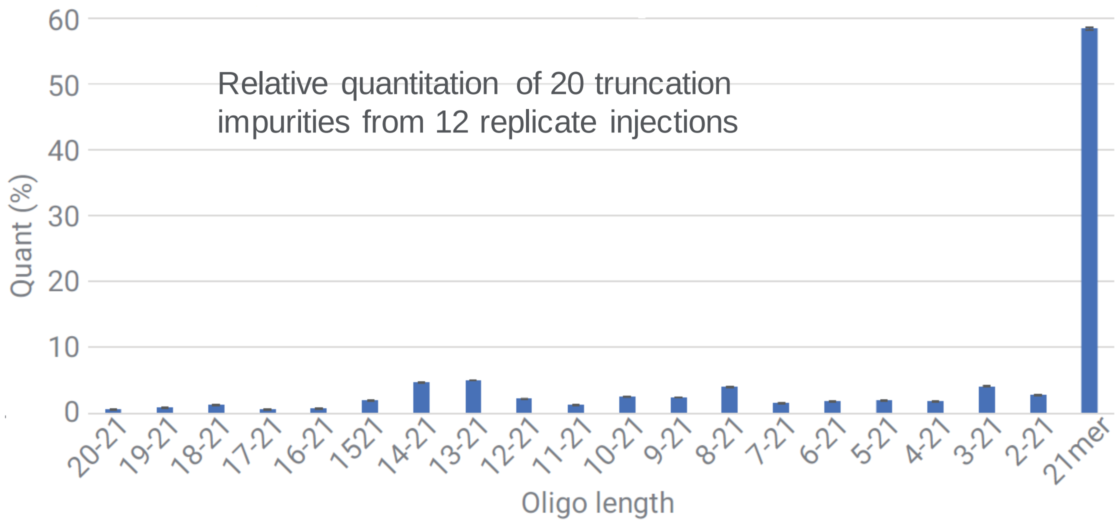
6.0 kDa

Example TPI Experiment Data from Targeted FBF Analysis


Overlaid compound chromatograms from 21mer target and 5' truncations.

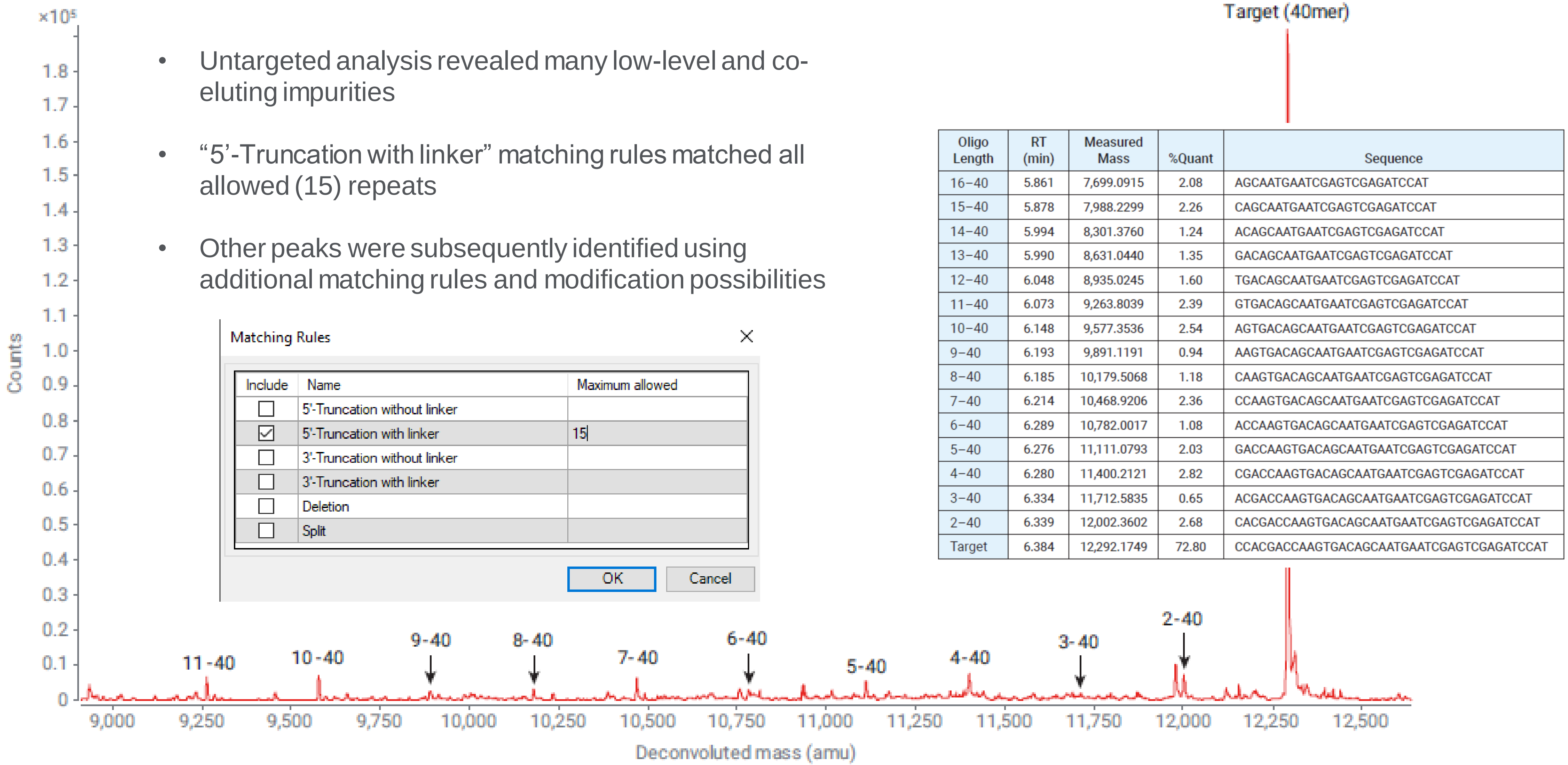
- Targeted analysis found complete set of 5' truncation impurities
- Many low abundance truncations were chromatographically separated and undetectable in TIC.
- Excellent mass accuracy (average < 1 ppm)
- Reproducible relative quant from 12 replicates

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


Relative quantitation of 20 truncation impurities from 12 replicate injections

Example TPI Experiment Data from Untargeted MaxEnt Deconvolution

- Untargeted analysis revealed many low-level and co-eluting impurities
- “5’-Truncation with linker” matching rules matched all allowed (15) repeats
- Other peaks were subsequently identified using additional matching rules and modification possibilities



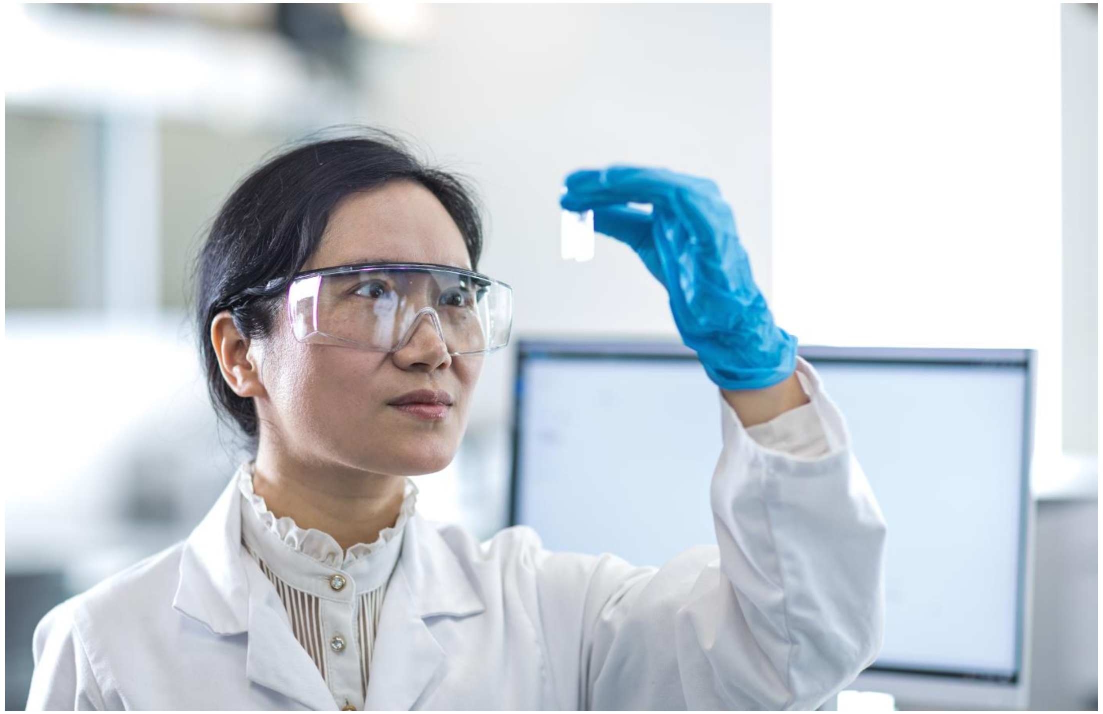
Matching Rules

Include	Name	Maximum allowed
☐	5'-Truncation without linker	
☑	5'-Truncation with linker	15
☐	3'-Truncation without linker	
☐	3'-Truncation with linker	
☐	Deletion	
☐	Split	

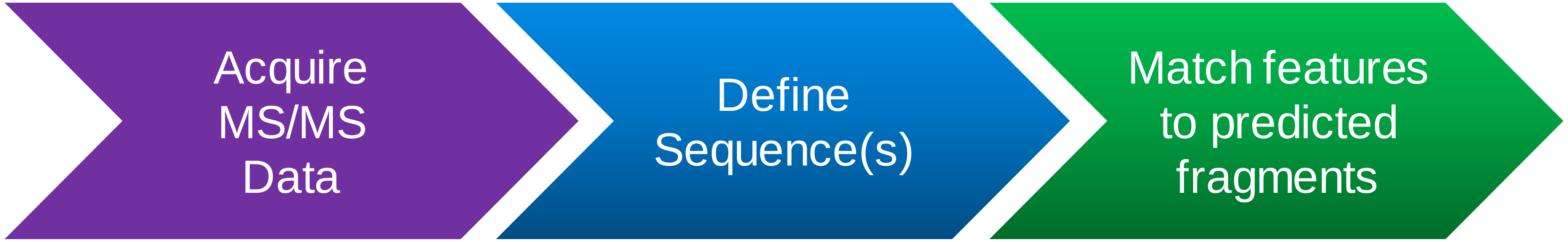
OK Cancel

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

How do you really know the oligo sequences are correct?



Sequencing Workflow in MassHunter BioConfirm 12.0

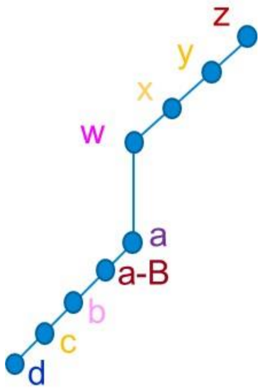


Acquire MS/MS Data

Define Sequence(s)

Match features to predicted fragments

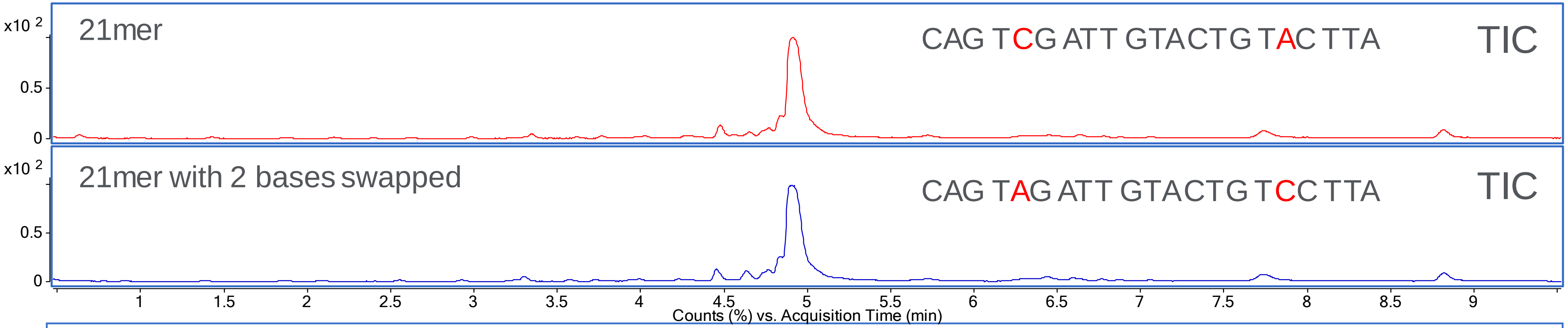
Sequencing results are displayed using a fragment confirmation ladder

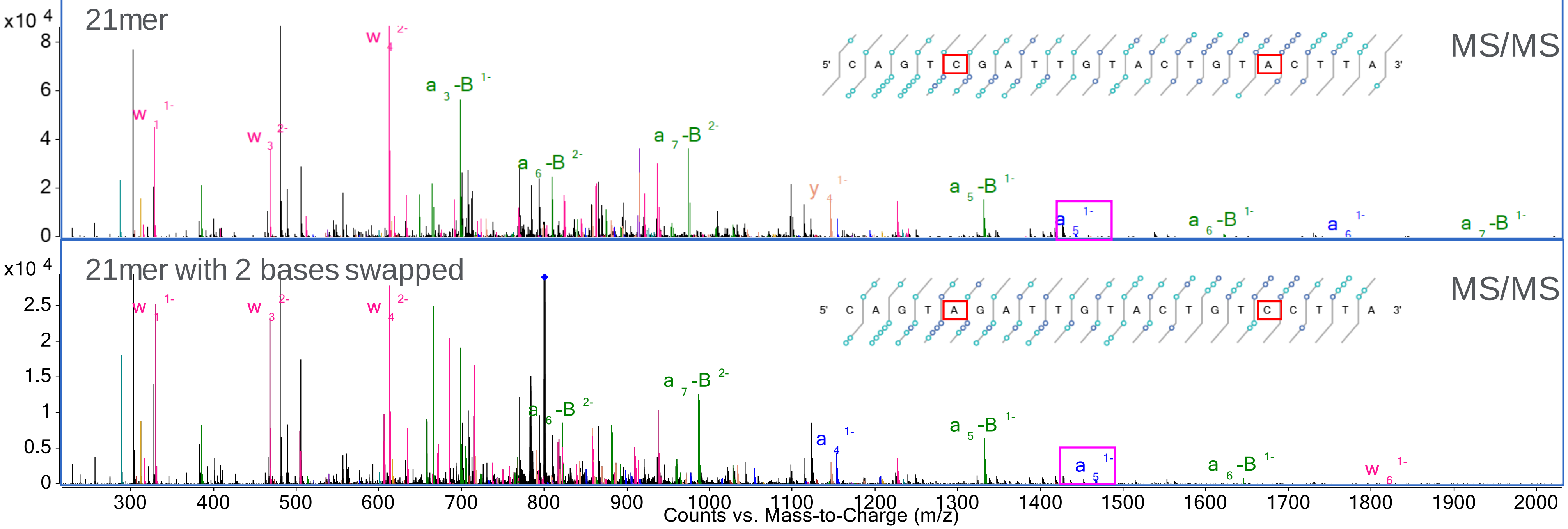


- Dots indicate found fragments
- Dot position indicates fragment type (d, c, b, etc)
- Dot color indicates number of times that fragment was found (e.g. from different files or charge states)
  - Dark blue means found 2+ times
  - Any other color indicates file from which the fragment was found once
- Open dots indicate fragment(s) selected for view in MSMS spectrum

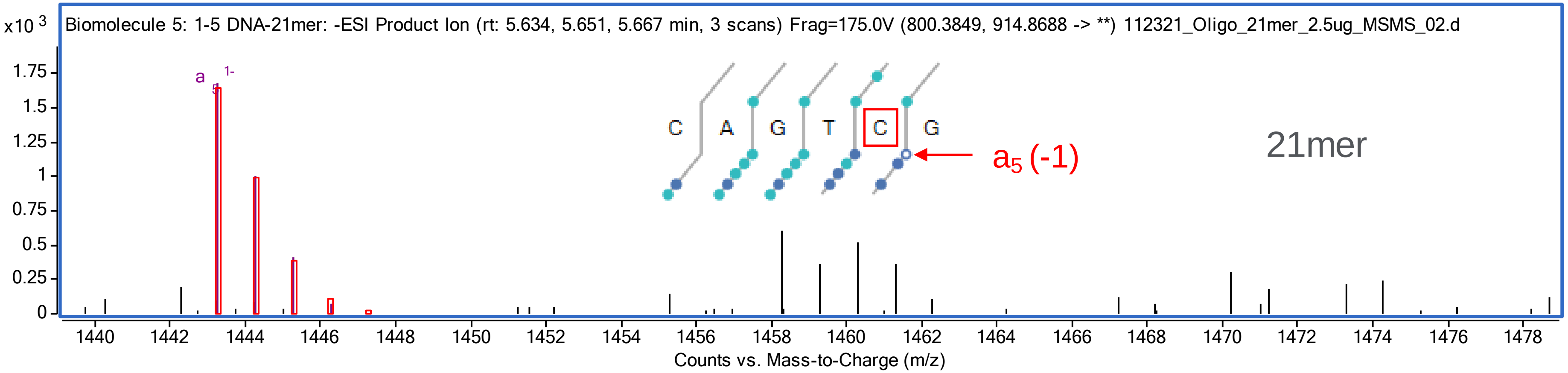
Please refer to Agilent application note 5994-5071EN

Sequence Data for Isomeric Oligos

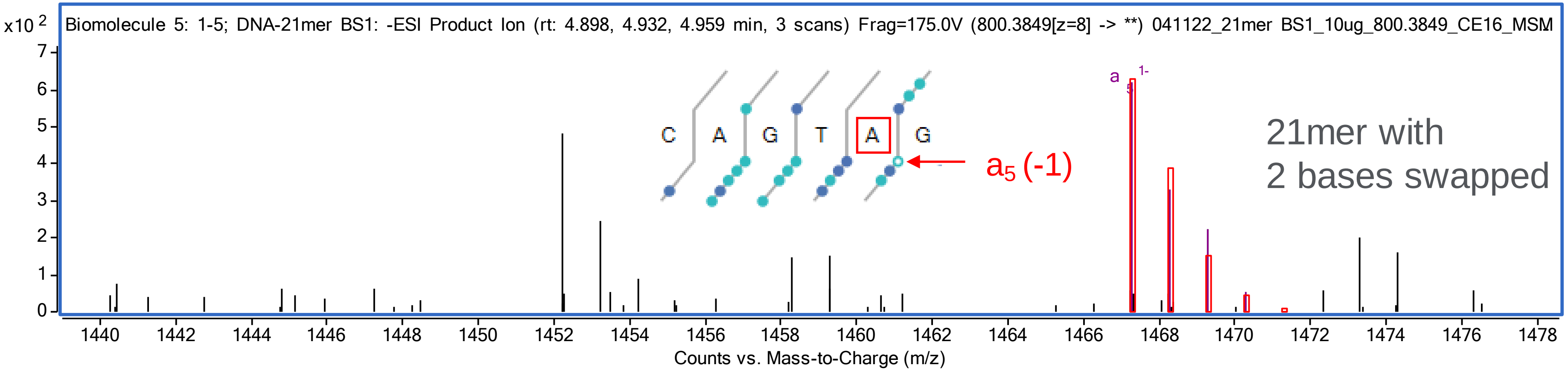




Sequence Data for Isomeric Oligos

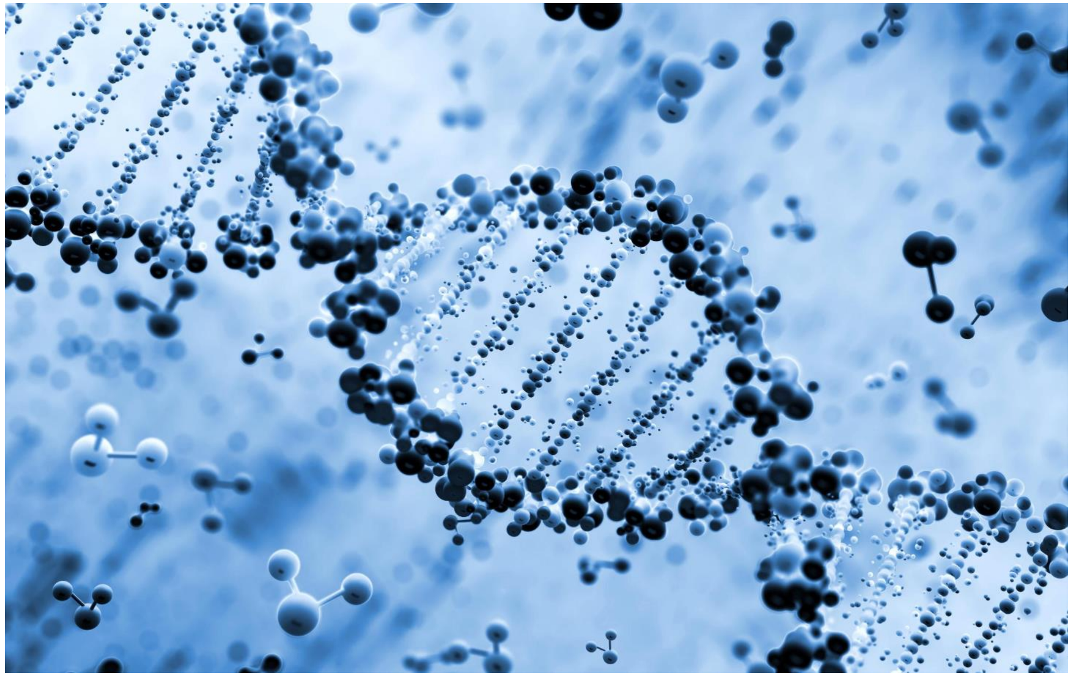


21mer

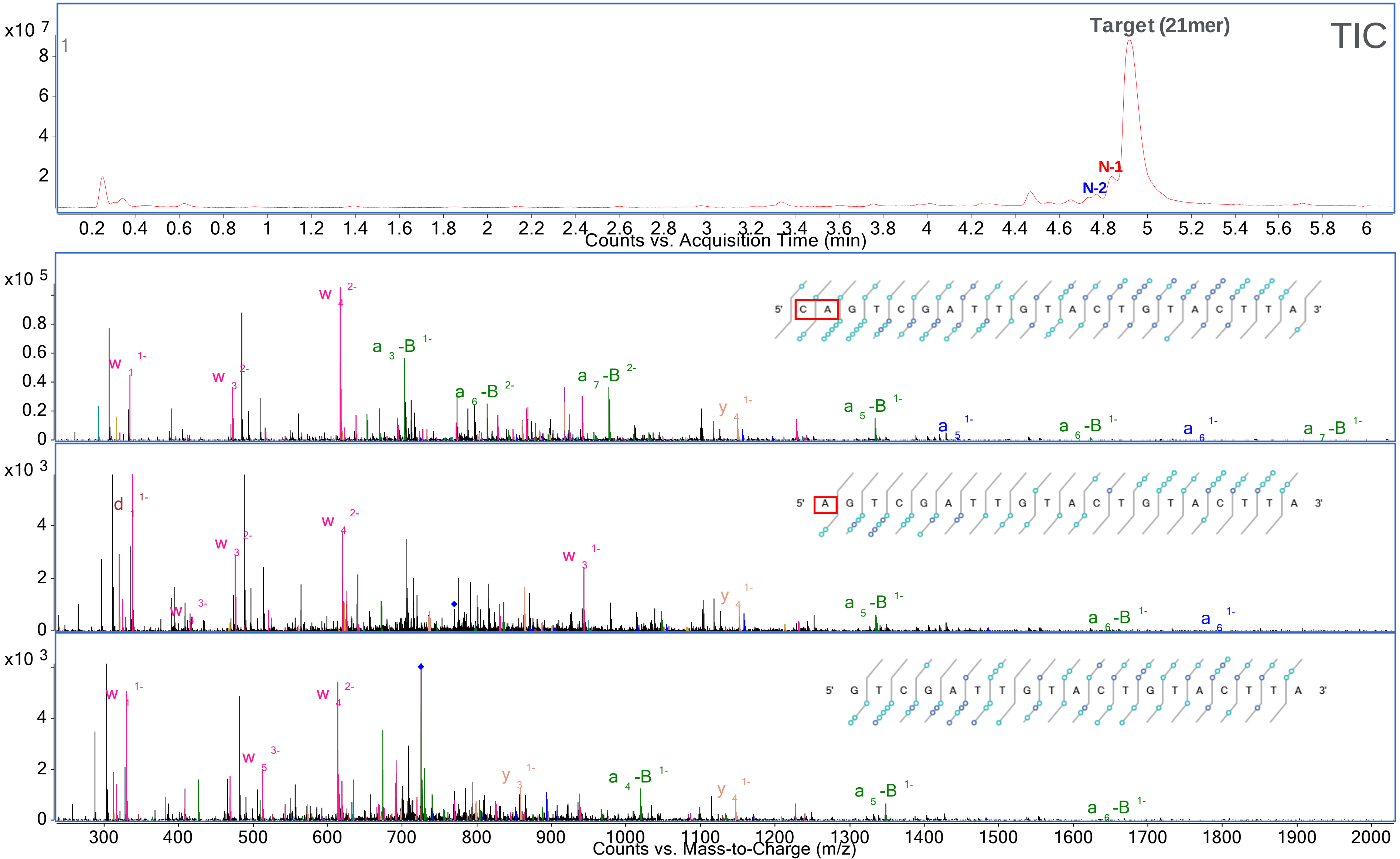


21mer with
2 bases swapped

What if you need to confirm the sequence of an impurity identified from the TPI workflow?



Identification of Truncation Impurities



Target (21mer)

5' N-1 (2-21)

5' N-2 (3-21)

You were successful in profiling the impurities of the oligos and confirming the sequence!

Your success brings more requests – this time on more complex samples.

ASO

/52MOErT//i2MOErC/*/i2MOErA/*/i2MOErC/*/i2MOErT/*/i2MOErT/*/i2MOErT/*/i2MOErC/*/i2MOErA/*/i2MOErT/*/i2MOErA/*/i2MOErA/*/i2MOErT/*/i2MOErG/*/i2MOErC/*/i2MOErT/*/i2MOErG/*/32MOErG/*

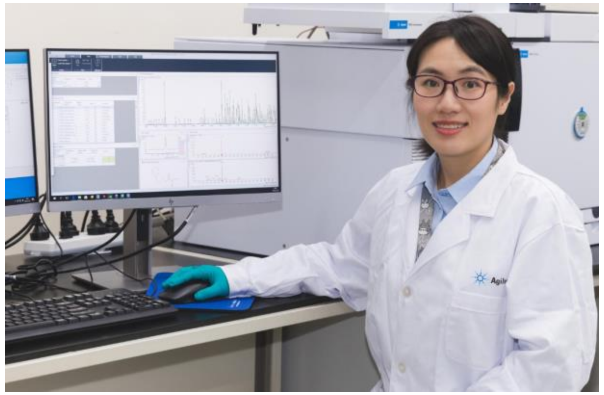
Aptamer

/52FC/mGmGrArA/i2FU//i2FC/mAmG/i2FU/mGmAmA/i2FU/mG/i2FC/

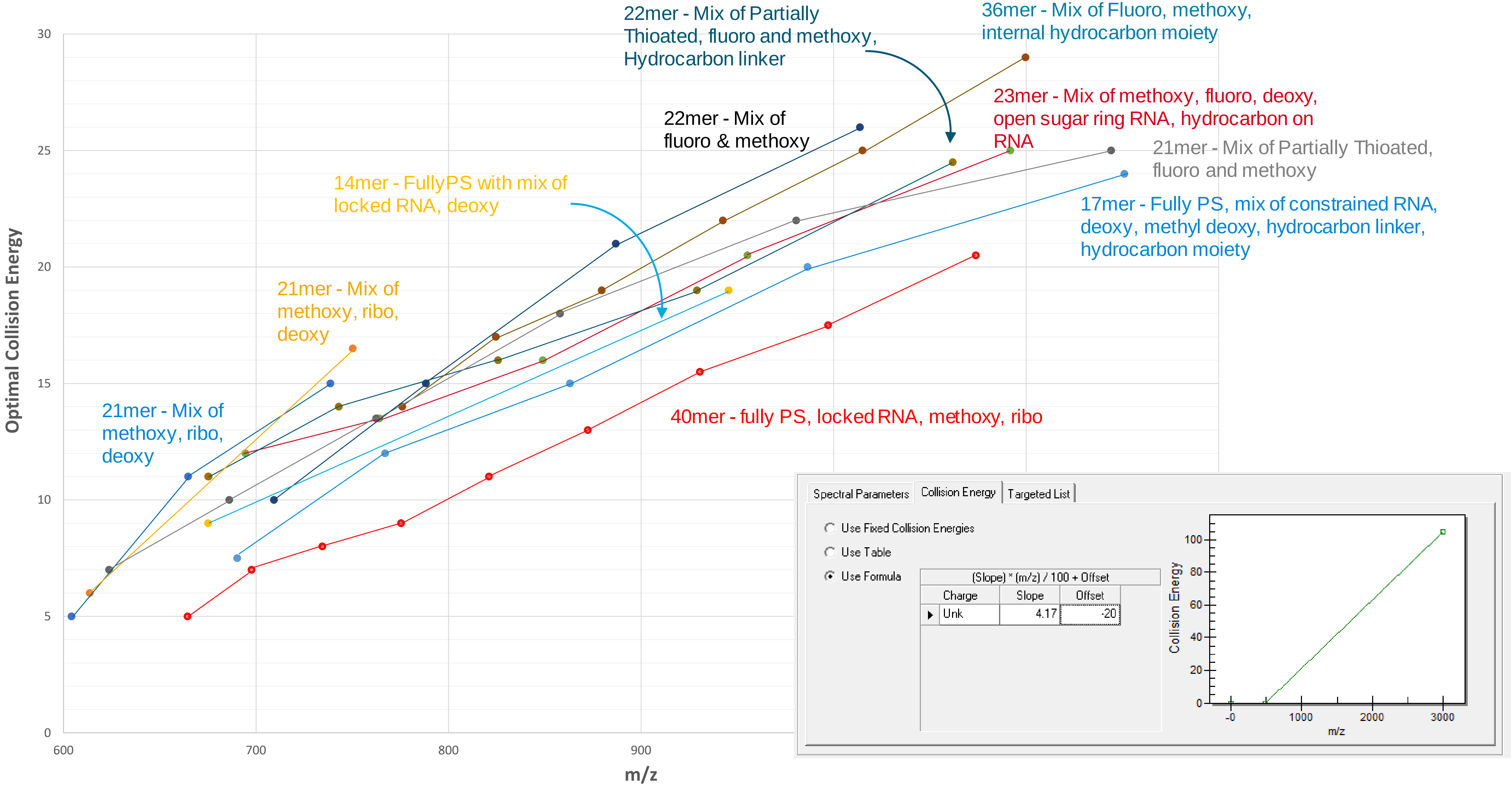
/i2FU//i2FU/mA/i2FU/mA/i2FC/mA/i2FU//i2FC//i2FC/mG/3InvdT/

You're wondering if these samples will also work

and if they will be harder to characterize.



m/z vs. Optimal Collision Energy



22mer - Mix of Partially Thioated, fluoro and methoxy, Hydrocarbon linker

36mer - Mix of Fluoro, methoxy, internal hydrocarbon moiety

23mer - Mix of methoxy, fluoro, deoxy, open sugar ring RNA, hydrocarbon on RNA

22mer - Mix of fluoro & methoxy

21mer - Mix of Partially Thioated, fluoro and methoxy

14mer - FullyPS with mix of locked RNA, deoxy

17mer - Fully PS, mix of constrained RNA, deoxy, methyl deoxy, hydrocarbon linker, hydrocarbon moiety

21mer - Mix of methoxy, ribo, deoxy

21mer - Mix of methoxy, ribo, deoxy

40mer - fully PS, locked RNA, methoxy, ribo

○ Use Fixed Collision Energies

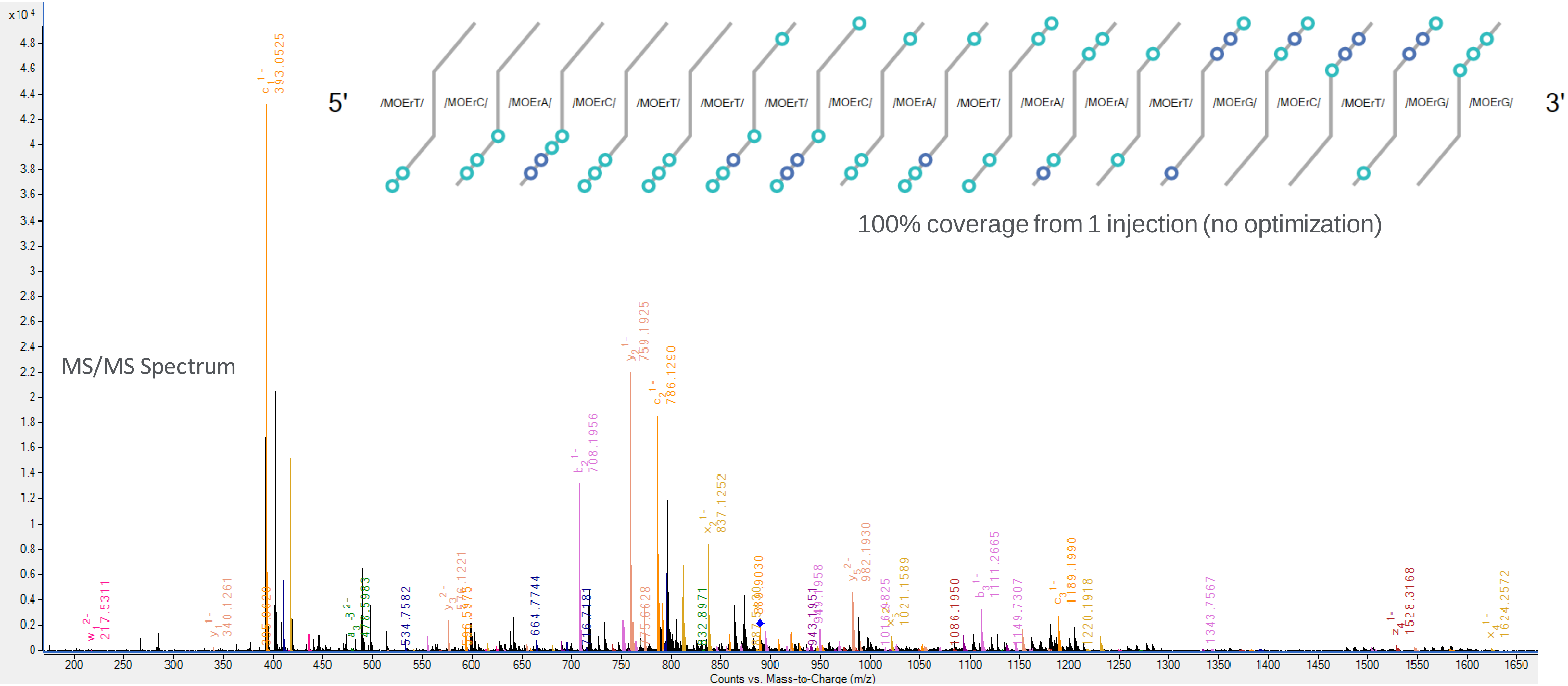
○ Use Table

● Use Formula

(Slope) * (m/z) / 100 + Offset

Charge	Slope	Offset
Unk	4.17	-20

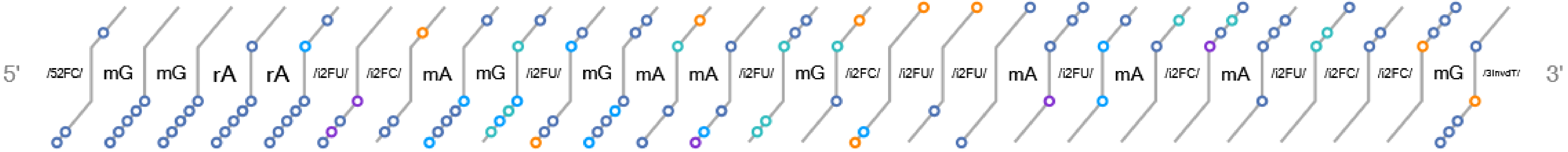
Sequence Data for Heavily Modified ASO run by IPRP LCMS



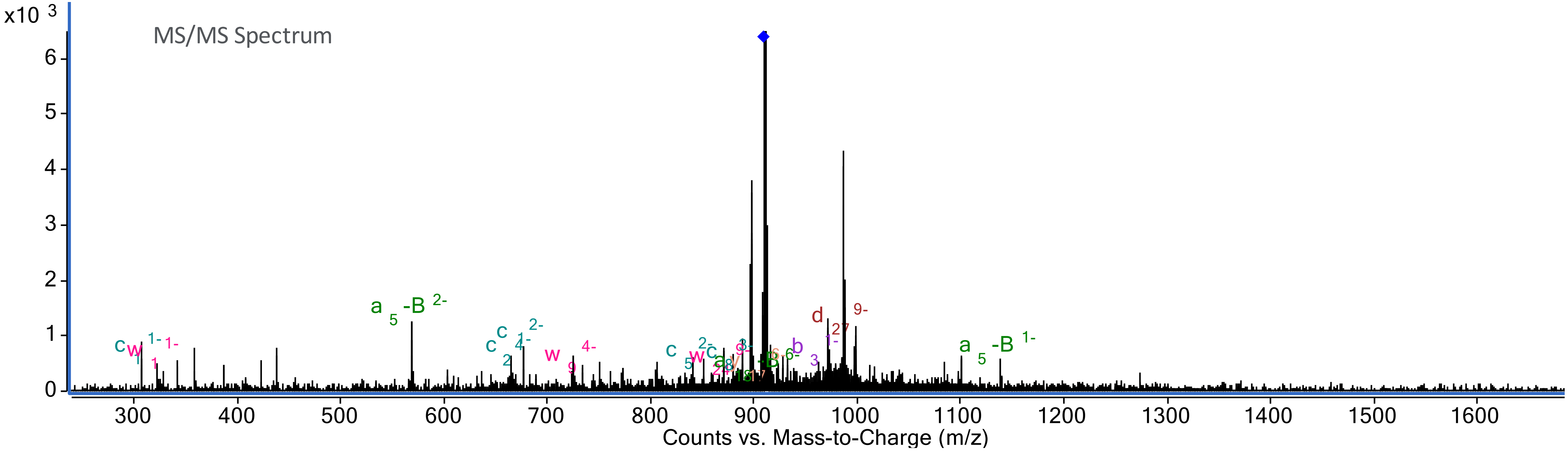
100% coverage from 1 injection (no optimization)

MS/MS Spectrum

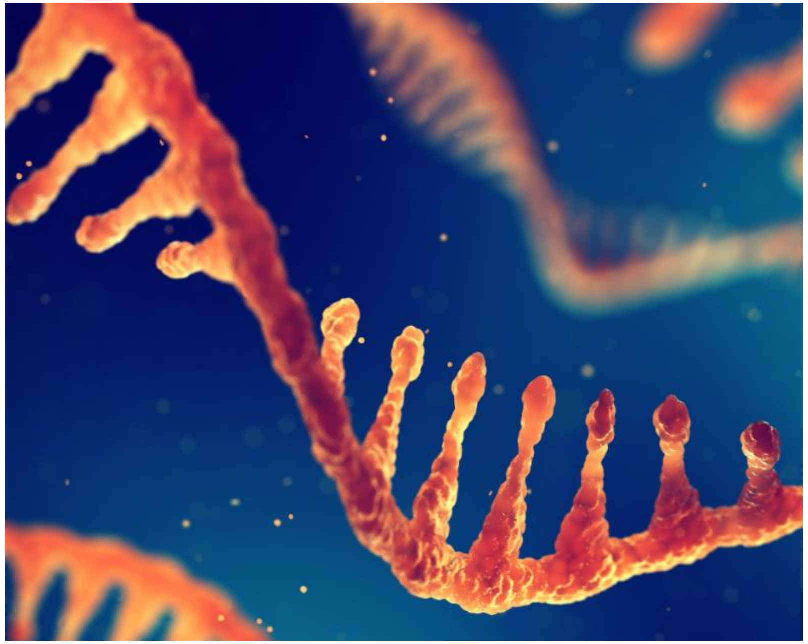
Sequence Data for Modified Aptamer run by HILIC LCMS



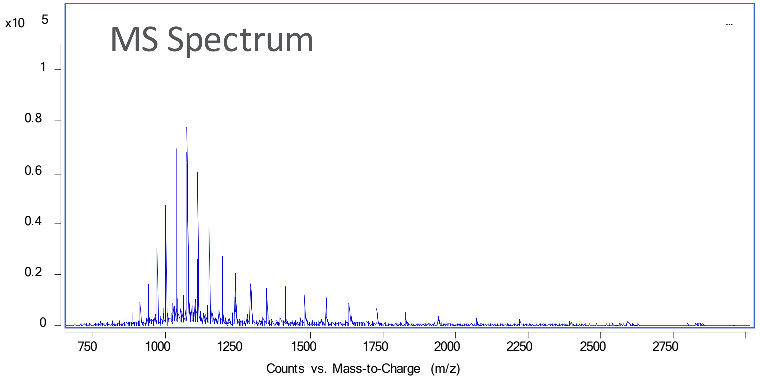
100% coverage from 1 injection (no optimization)

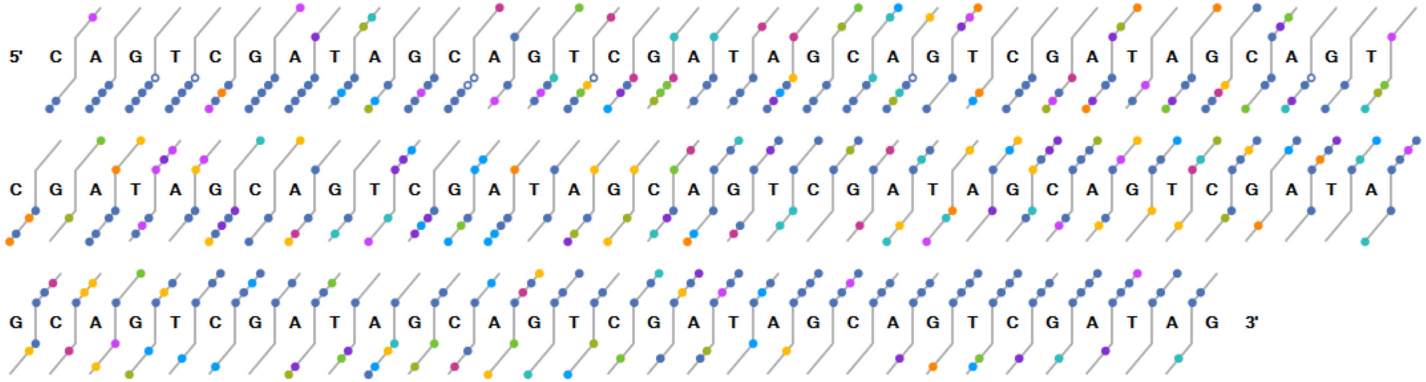


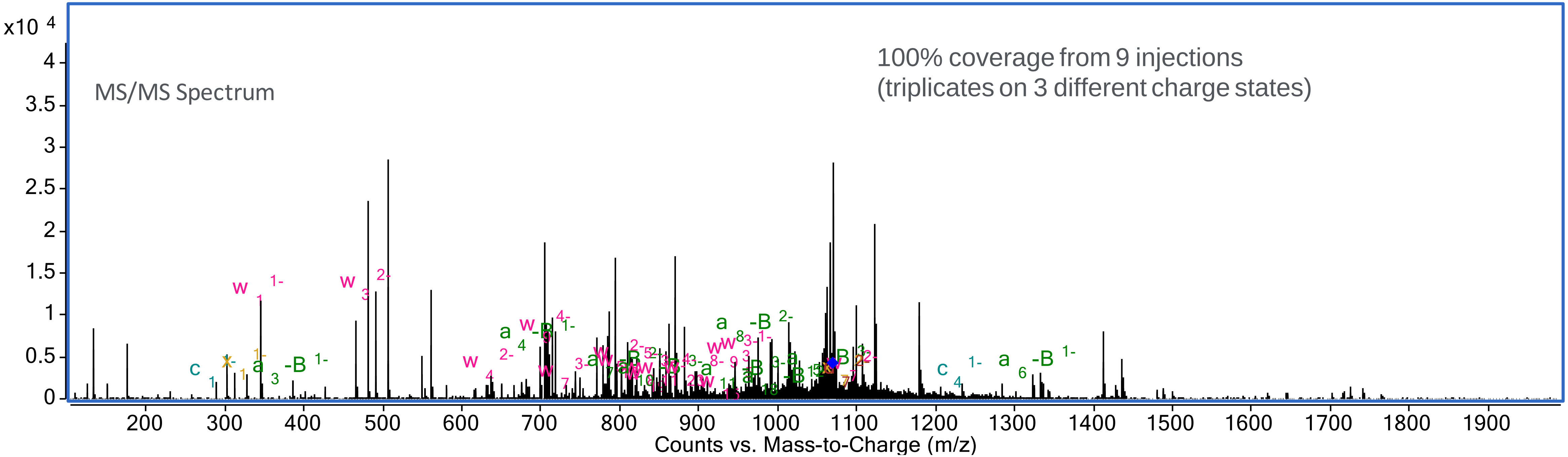
What if you work with longer synthetic oligos and can’t digest them?



Sequence Data for a Long (100mer) Synthetic Oligo





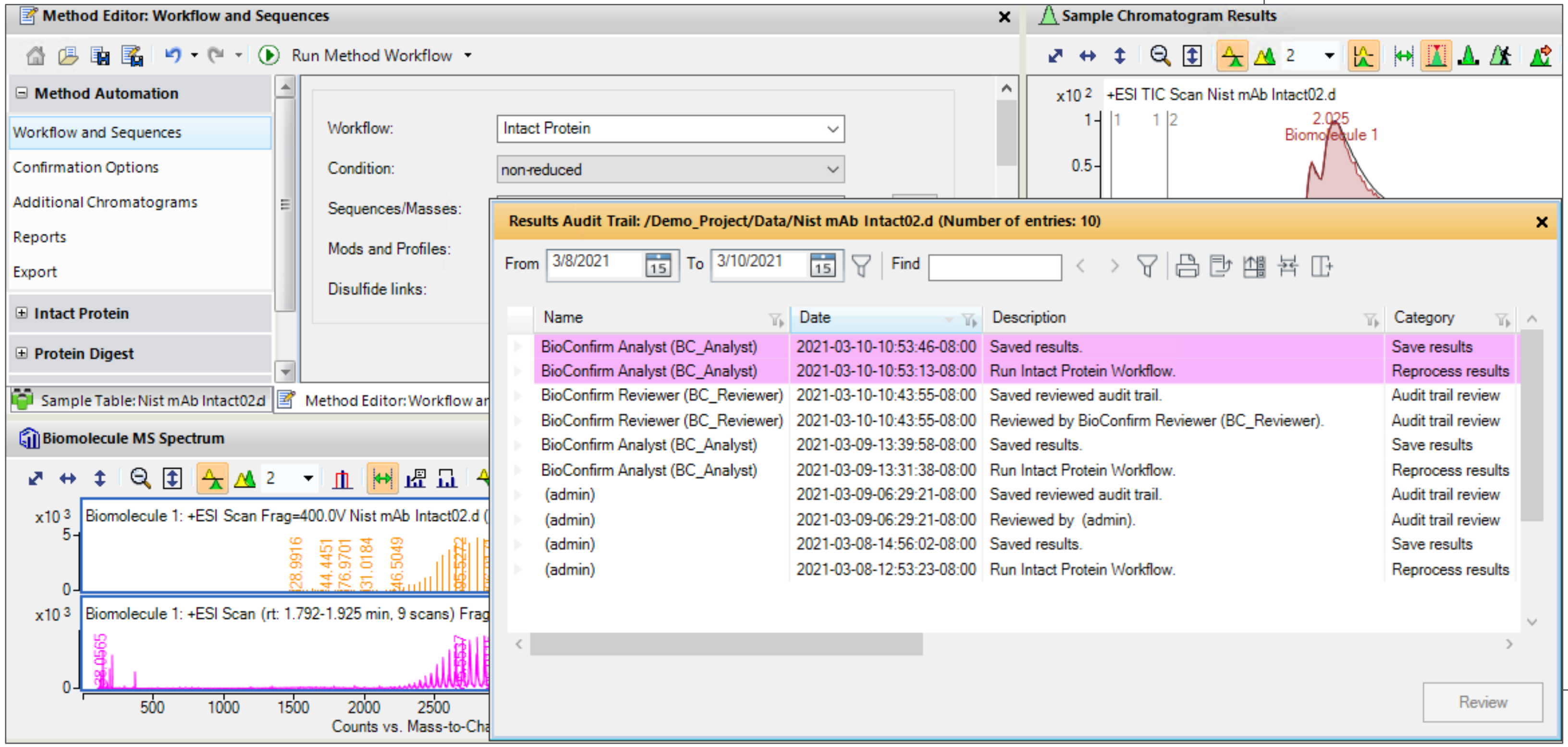


100% coverage from 9 injections
(triplicates on 3 different charge states)

Compliance-Ready LCMS

MassHunter offers technical controls for customers to meet the requirements of 21 CFR Part 11 compliance

Agilent also has offerings for Data integrity, Instrument qualification, Computer system validation, consulting, and Validation starter kits



Workflow: Intact Protein

Condition: non-reduced

Sequences/Masses:

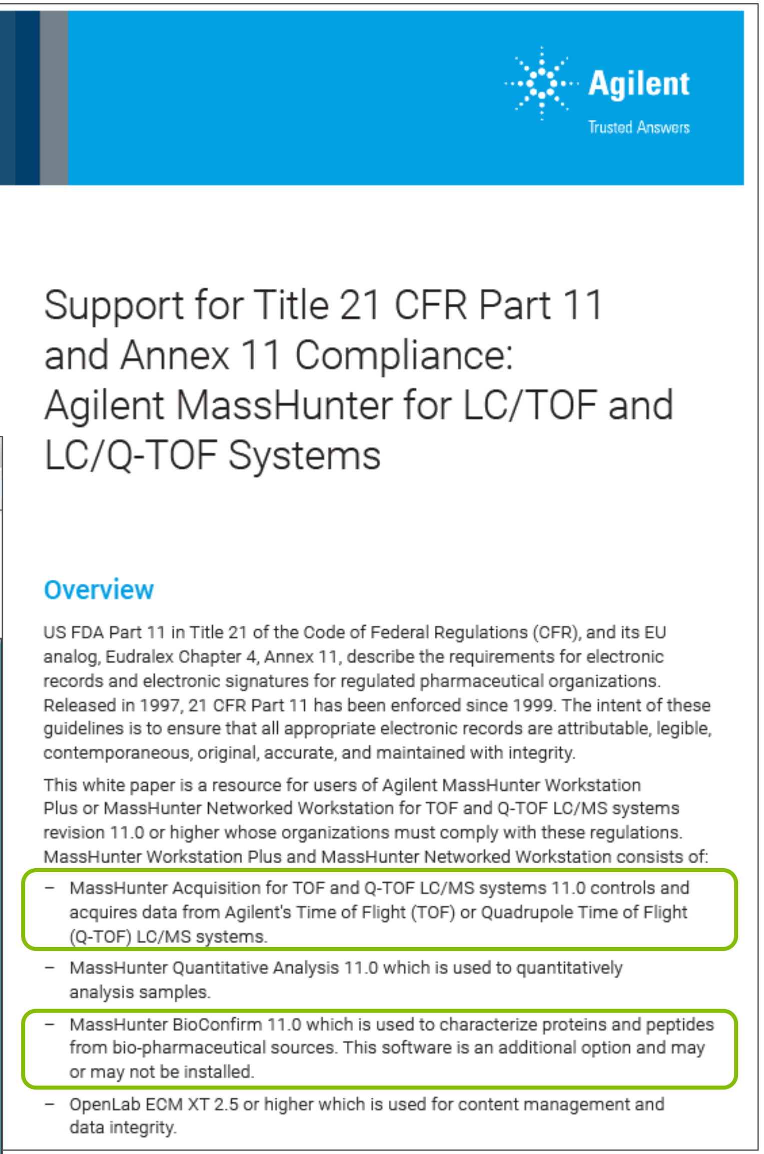
Mods and Profiles:

Disulfide links:

Results Audit Trail: /Demo_Project/Data/Nist mAb Intact02.d (Number of entries: 10)

From 3/8/2021 To 3/10/2021 Find

Name	Date	Description	Category
BioConfirm Analyst (BC_Analyst)	2021-03-10-10:53:46-08:00	Saved results.	Save results
BioConfirm Analyst (BC_Analyst)	2021-03-10-10:53:13-08:00	Run Intact Protein Workflow.	Reprocess results
BioConfirm Reviewer (BC_Reviewer)	2021-03-10-10:43:55-08:00	Saved reviewed audit trail.	Audit trail review
BioConfirm Reviewer (BC_Reviewer)	2021-03-10-10:43:55-08:00	Reviewed by BioConfirm Reviewer (BC_Reviewer).	Audit trail review
BioConfirm Analyst (BC_Analyst)	2021-03-09-13:39:58-08:00	Saved results.	Save results
BioConfirm Analyst (BC_Analyst)	2021-03-09-13:31:38-08:00	Run Intact Protein Workflow.	Reprocess results
(admin)	2021-03-09-06:29:21-08:00	Saved reviewed audit trail.	Audit trail review
(admin)	2021-03-09-06:29:21-08:00	Reviewed by (admin).	Audit trail review
(admin)	2021-03-08-14:56:02-08:00	Saved results.	Save results
(admin)	2021-03-08-12:53:23-08:00	Run Intact Protein Workflow.	Reprocess results



White Paper

Agilent
Trusted Answers

Support for Title 21 CFR Part 11 and Annex 11 Compliance: Agilent MassHunter for LC/TOF and LC/Q-TOF Systems

Overview

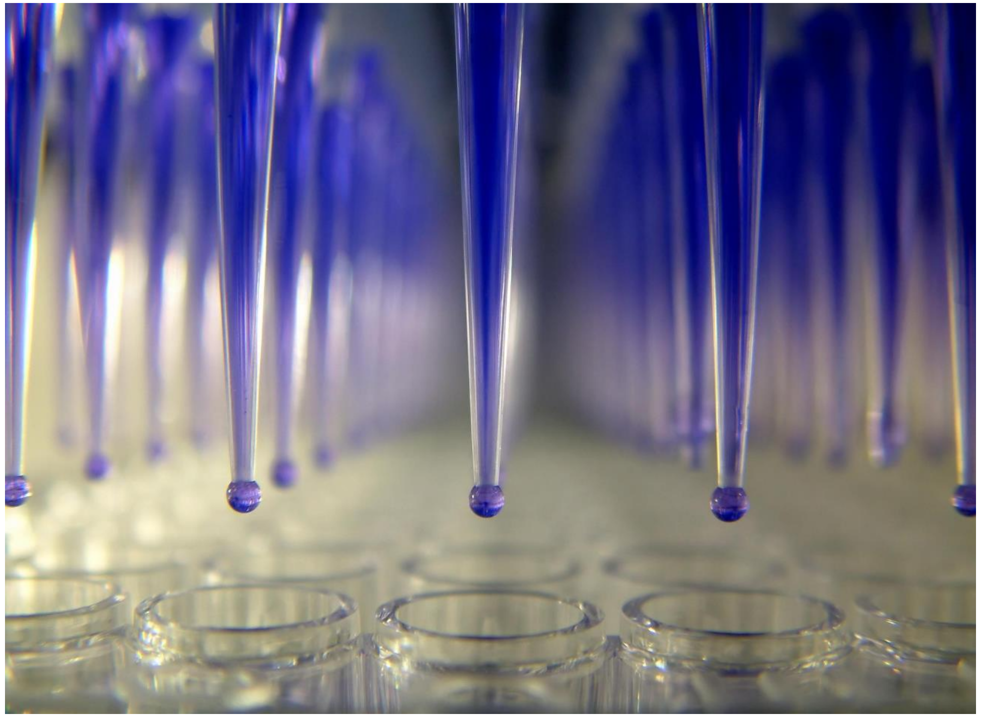
US FDA Part 11 in Title 21 of the Code of Federal Regulations (CFR), and its EU analog, Eudralex Chapter 4, Annex 11, describe the requirements for electronic records and electronic signatures for regulated pharmaceutical organizations. Released in 1997, 21 CFR Part 11 has been enforced since 1999. The intent of these guidelines is to ensure that all appropriate electronic records are attributable, legible, contemporaneous, original, accurate, and maintained with integrity.

This white paper is a resource for users of Agilent MassHunter Workstation Plus or MassHunter Networked Workstation for TOF and Q-TOF LC/MS systems revision 11.0 or higher whose organizations must comply with these regulations. MassHunter Workstation Plus and MassHunter Networked Workstation consists of:

- MassHunter Acquisition for TOF and Q-TOF LC/MS systems 11.0 controls and acquires data from Agilent's Time of Flight (TOF) or Quadrupole Time of Flight (Q-TOF) LC/MS systems.
- MassHunter Quantitative Analysis 11.0 which is used to quantitatively analysis samples.
- MassHunter BioConfirm 11.0 which is used to characterize proteins and peptides from bio-pharmaceutical sources. This software is an additional option and may or may not be installed.
- OpenLab ECM XT 2.5 or higher which is used for content management and data integrity.

Please also refer to Agilent tech note 5994-3546EN

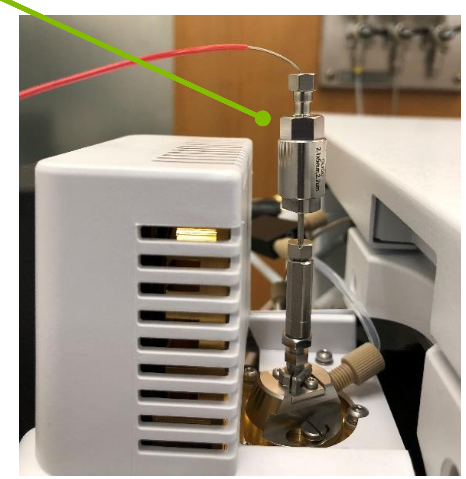
What if you have a large number of oligos to run each week?



Fast LC using Dual Needles with Smart Overlap and a Guard Column

HPLC: Agilent 1290 Infinity II Binary pump, Multi-sampler with Dual Needles

LC Conditions			
Column	AdvanceBio Oligo UHPLC Guard column, 1.7 um, 2.1 x 5mm pn: 821725-921		
Column temperature	room temperature		
Injection volume	1 µL		
Smart Overlap	Enabled, alternating needle		
Autosampler temp	4 ºC		
Needle wash	MeOH:Water 50:50		
Mobile phase	A = 15 mM TEA and 400 mM HFIP in water B = MeOH		
Flow rate	1.00 mL/min		
Gradient program	Time (min)	Time (sec)	B (%)
	0.00	0.00	20
	0.03	1.80	20
	0.24	14.4	50
	0.25	15.0	100
	0.30	18.0	100
	0.31	18.6	20
	0.59	35.0	20
Stop time	0.60 min		
Post time	0.00 min		



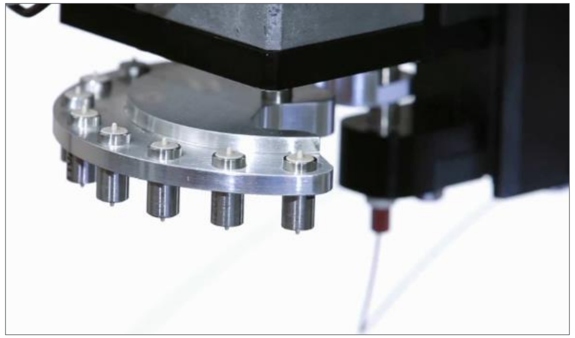
~40 seconds per sample, over 2,000 samples a day

Please refer to Agilent application note 5994-3753

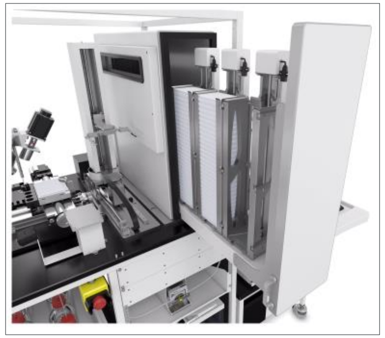
High-throughput MS by RapidFire

Ultrafast autosampler & online solid phase extraction (SPE) system

- No chromatography .. high-throughput desalter
- Couples directly to mass spec (just like an LC)



Automated cartridge changer



Temperature controlled
sample storage

Made to handle large sample sets

- Can run up to 90 x 1536-well plates (=138,240 samples) without intervention
- Has temperature-controlled sample storage unit
- Integrated plate handler and bar code scanner
- Plates can be heat sealed
- SPE cartridges are good for thousands of injections each
- 12 cartridge changer with automatic switching features

Made to go fast

- Multiple pumps and valves to switch instantly from desalt to elute conditions
- MS acquisition is constant through sample set

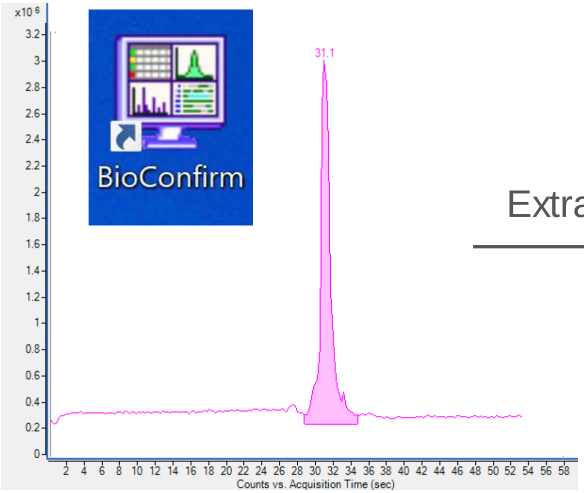
IPRP or HILIC methods available

~15 seconds per sample, over 5,000 samples a day

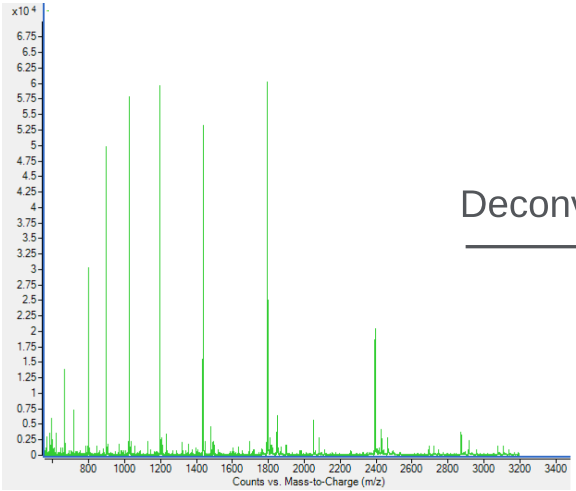
RapidFire
400

Please refer to Agilent application notes 5994-3753 and 5994-4945

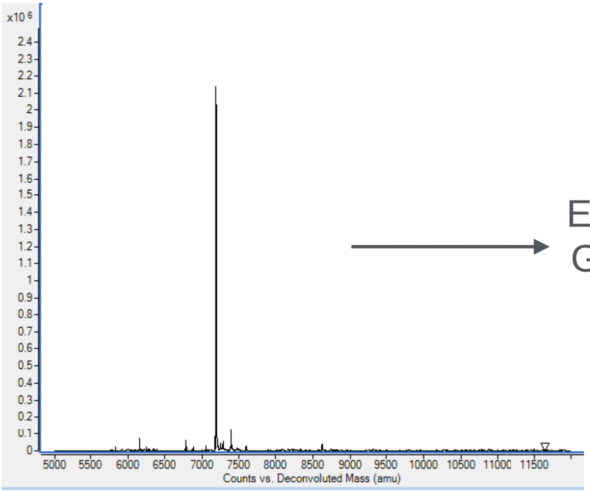
TPI – Automated DA in BioConfirm and Review in BioMS Reviewer



Extract



Deconvolute



Export peak tables
Generate Reports

BioMS
Reviewer

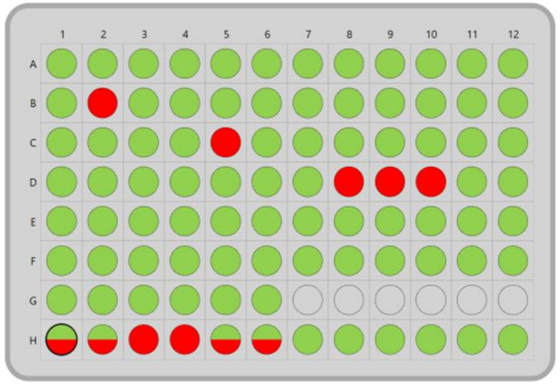
Open results in BioMS Reviewer

Color codes each sample position for "target confirmed" (top color) and "target meets purity threshold" (bottom color)

Displays full sample details/report for position selected

Supports sample pass/fail

exporting/printing of results for whole rack



54, 96, and 384-well formats

Summary

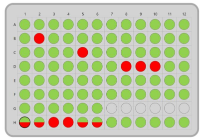
The high-resolution 6230 LC/TOF combined with BioConfirm 12.0 provides rich TPI results using multiple data mining techniques including Find-by-formula (targeted) and Maximum Entropy Deconvolution (untargeted).

The high-resolution 6545XT AdvanceBio LC/Q-TOF combined with BioConfirm 12.0 provides maximal performance and capabilities, supporting both TPI and oligo sequencing workflows.

Agilent supports both ion-pair reverse-phase and HILIC (non-ion-pairing) oligo methods on canonical, heavily modified, and long synthetic samples.

Multiple high-throughput options for TPI are supported including Fast LC and RapidFire HRAM MS.

Data analysis is automated in BioConfirm 12.0 and data review is streamlined by the BioMS Reviewer software.



Acknowledgements

Jim Lau, Guannan Li, and Yannan Yang in the LCMS AE group, AFO

Andrew Coffey and Andrea Tripodi in the Chemistries and Supplies Division

Jason Epps and Todd Kreutzian in the Analytical Development Group at the Nucleic Acid Solutions Division

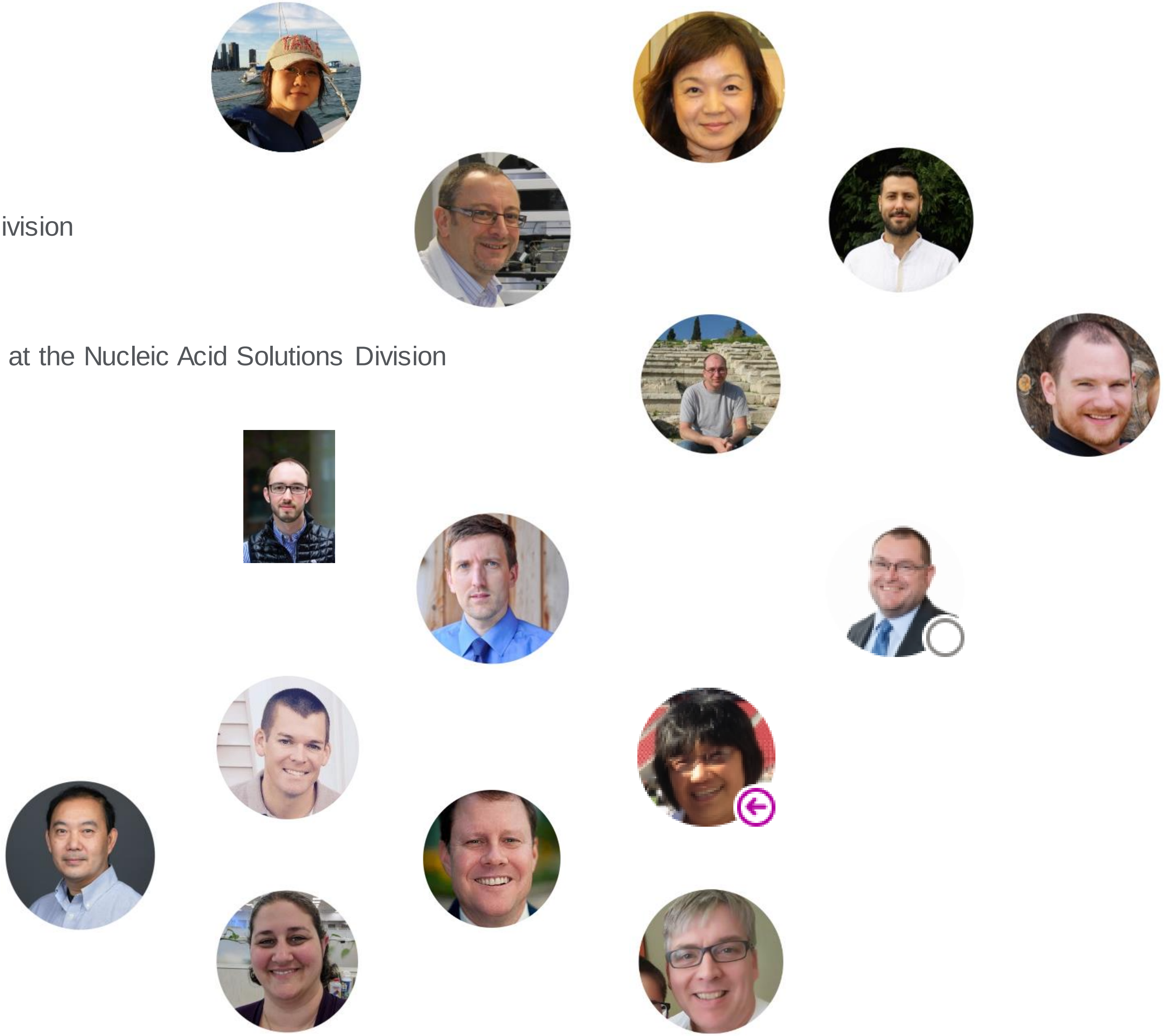
Cody Schwarzer, Senior Scientist at Foghorn Therapeutics

Richard Blankley and Andy Wright in the LCMS AE group, EMEAI

Patrick Cronan and Lisa Zhang in the LC AE group, AFO

David Wong and Steve Madden in Biopharma/Product Marketing

Crystal Cody and Gordon Slysz in R&D



Agilent

Trusted Answers