Automated determination and profiling of therapeutic oligonucleotide impurities and metabolites in LC-HRMS data

Eva Duchoslav; Harini Kaluarachchi; Thanh Ngu; Lyle Burton; Peter Liuni and Jason Causon SCIEX, 71 Four Valley Drive, Concord, ON, L4K 4V8 Canada

ABSTRACT

This presentation highlights several aspects of a typical metabolite and impurity data processing pipeline, that need refinement to support analysis of therapeutic nucleic acids modalities ranging in size from 5 000 to 10000. Entries in a customizable library of oligonucleotides defined in terms of their 5 building blocks, were used to describe the parent molecule and predict its metabolic and impurity products. Isotope pattern fidelity in collected HR MS was instrumental in compound confirmation and the signal from series of isotope peaks was utilized in profiling workflows. Automated data processing of series of samples was enabled through the Molecule Profiler App integrated in the SCIEX OS software.

INTRODUCTION

Nucleic acid modalities, such as siRNA, anti-sense oligonucleotides and aptamers, have become emerging therapeutic candidates in biopharmaceutical development. To be able to succeed in the development and clinical pipelines and meet the regulatory requirements, the biophysical properties and purity of these candidate molecules must be well understood. Recent advancements in the chromatographic separations coupled to high resolution high sensitivity accurate mass spectrometry systems facilitate workflows that are capable of confident identification and relative quantitation of oligonucleotides, as well as their process-related impurities and metabolic or degradation products. To provide timely accurate qualitative and quantitative results, dedicated data reduction tools focused on the mass spectrometry of nucleic acids, need to be deployed in the analytical pipeline.

MATERIALS AND METHODS

Sample Preparation:

representative set of oligonucleotides (11- to 34-mer) a pure oligonucleotide solution spiked with shortmers at low concentrations was investigated.

Instrumental Analysis:

LC-HRMS System:

Agilent 1290 Infinity LC and SCIEX ZenoTOF 7600 system or TripleTOF 6600+ system. LC column: Waters ACQUITY UPLC Oligonucleotide BEH C18, 2.1mmx50mm 1.7u 130 Å.

LC/MS Acquisition Method:

The column was maintained at 60 °C, with a mobile phase consisting of Solvent A: 100 mM hexafluoro-2propanol [HFIP] and 10 mM triethylamine (TEA), and Solvent B: methanol.

After HPLC separation, the samples were analysed with quadrupole-time-of-flight (QTOF) tandem mass spectroscopy (MS/MS) on a TripleTOF 6600 system (SCIEX) equipped with a DuoSpray ion source.

Oligonucleotide structure details in hypothesis-driven data mining

Many expected metabolic or degradation products are results of hydrolysis, cleaving a portion of the molecule, or modifications associated with a specific part of oligonucleotide. For efficient data analysis, studied oligonucleotide molecule is described as a chain of residues. Each residue has up to 5 building blocks that describe the base, sugar, linker (phosphate) core and 2 linker atoms that connect the linker core from 3' and 5' sides to a sugar. Some of the parts can be missing for special cases of residues, such as 3' and 5' termini and some linker cores (such as morpholino). Residue symbols have special characters and abbreviated information on modifications to base, sugar and linker core, for example, "*" indicates phosphorothioate linker.

- ability to predict synthetic impurities, including internal omission and addition impurities, and assess compatibility with a putative degradation reaction

DISCUSSION

Oligonucleotide chemical space

type	# nucleotides	Molar mass g∙mol−1	monoisotopic to base peak	Index of base peak
DNA, phosphorothioate	1*21	6682.35	15.0%	5
2-methoxyethyl DNA mix, phosphorothioate	1*20	7594.76	11.7%	5
2-methoxyethyl DNA mix, phosphorothioate	1*20	7183.08	11.6%	5
Phosphorodiamidate morpholino	1*30	10305.89	4.3%	6
Phosphorodiamidate morpholino	1*25	8647.401	8.9%	5
2-methoxyethyl	1*18	7500.86	12.2%	5
Double stranded RNA with 2'-O-methyl	2 * 22	13424.39	2.2%	7
Double stranded 2' O-methyl and 2' fluoro	21+23	16300	0.5%	9
5' 40kDaPEG single stranded and double stranded mix	27	~50 000	<0.1%	
single stranded and double stranded mix	variable		<0.1%	

Table 1. List of 10 FDA approved drugs to 2020¹, their chemical make-up and characteristics to be considered in LC-MS data analysis.

The last two drugs do not have the isotope pattern resolved in TOF MS at R=40000, and thus they were not considered in our project.



Figure 1. Isotope pattern of 20-mer ssDNA (M=6065.04) CGG CTA CCT TGT TAG CAC AT



Figure 2. Five-part definition of RNA residue and the respective in-silico generated MS/MS fragments.

<u>Five -part oligonucleotide structure representation enables:</u>

- prediction of hydrolytic cleavages
- in-silico generation of accurate MS/MS fragments

Transforming LC/HRMS oligonucleotide data into qualitative workflow results

Challenge

The TOF MS signal of an oligo is distributed To capture the total biomolecule response, series across isotope patterns for range of charge of accurate TOF MS XICs are summed or, the peak states and, optionally, it has different H-alkali areas in accurate XICs are summed. metal exchanges

Isotope patterns have many peaks and Average mass supported with a score for isotope may be immersed in noise mass confirm compound mass. *through Q1 and results in complex MS/MS* pattern provide more confident fragment *spectrum with fragment isotope signal* assignment

monoisotopic peak for low abundant species pattern fit and where possible monoisotopic Single precursor isotope peak not transmitted Fragment charge state and fragment isotope

Impurities and metabolites are not fully | HRMS enables the orthogonal separation of *chromatographically separated* closely eluting peaks *Ionization efficiencies of parent and impurities* Parallel UV detection and processing may not be the same



Figure 3. Role of Isotope Score in metabolite assignment.

As large biomolecules have complex isotope patterns, two putative molecules of similar size are expected to have common peaks in their isotope pattern series. Isotope Score reflects how well each peak in respective isotope series fits the theoretical biomolecule isotope distribution, and thus it is instrumental in resolving competing identification assignments.



Solution

Intensity, cp	600 400 200 0	514.0838	516.4197 516.5825 518.0597 518. 5 516 517 518	4106 518.7437 52 519 520 m/z, Da	0.4357 521.1296 521	522.0890 522.5 522	5899 523 	0822 524.060 1 524)4
Assigned: 106 of 7586 peaks, MSMS Peak Area Assigned: 38.6%, Sequence Coverage: 21 of 21 residues									
Fra	Fragments: 106 of 106 Proposed Formulae 🛛 🕅 🔭 🍞						\bigtriangledown		
	Use	Mass (m/z)	Sequence		Ion	Charge	Error (ppm)	Intensity (cps)	*
31	\checkmark	513.9114	/H2PO3:/ dTdTdAdGdCdAdCdAdTdA		w10	6	-4.0	9.0	
32		515.7521	/H2PO3:/ dAdCdAdTdA		w5	3	1.5	1015.5	
33	\checkmark	518.4106	dCdGdGdCdT /:H2PO3/		d5	3	-0.4	76.3	
34	<	521.5876	dCdGdGdC /:C-4H-6N-3O-2/		a4 - B	2	1.5	1149.2	
35	✓	539.0800	dCdGdGdCdTdAdC /:H2PO3/		d7	4	-4.2	80.3	v

Figure 2 illustrates steps in sequence characterization of an "n+dA" spiked impurity. Terminal MS/MS fragment assignments lead to unambiguous confirmation of sequence that has adenine attached at 3' terminus (B). Isotope patterns of fragments aid in sequence confirmation (C).



Figure 6. Processing steps in a workflow for relative quantification and structural confirmation of metabolites integrated in the Molecule Profiler App.

Fin								
	d Metabolit	es						
	MS S	ample:		Sa	mple/Controls		_	
1	2020	1126_IDA_Ir Spike - IDA	npurity_0-1pc Sp \ Top 3 35eV - XA	iike_01.wiff A1 195 (QOL active) - 2.	.5%/min Grad		X Sample	Olig
2	MS S 2020 0.5%	ample: 1126_IDA_Ir Spike - IDA	mpurity_0-5pc Sp \ Top 3 35eV - X/	oike_01.wiff A1 195 (Q0L active) - 2.	.5%/min Grad		X Sample	Olig
	MS 5 2020	ample: 1126_IDA_Ir	mpurity_0pc Spik	e_01.wiff			X Sample	Olig
3	0% 5	Spike - IDA Te	op 3 35eV - XA1	195 (Q0L active) - 2.59	6/min Grad		X Controls	
4	2020	1126_IDA_In Spike - IDA	npurity_1pc Spik \ Top 3 35eV - XA	e_01.wiff A1 195 (Q0L active) - 2.	.5%/min Grad		X Sample X Controls	Olig
5	MS Sample: 20201192, IDA_Impurity_2pc Spike_01 wiff 2.0% Spike - IDA Top 3 35eV - XA1 195 (QQL active) - 2.5%/min Grad					X Sample	Olig	
6	MS S 2020	ample: 1126_IDA_Ir	mpurity_10pc Sp ⁱ	ike_01.wiff			X Sample	Olig
	10%	арке - пла	100 3 3364 - XX	Tibl (doc active) - 2.	System Grau		Controis	
29	Grou	ps of 1 Group ID	71 Poten	tial Metabol Neutral Mass	ite Peak R.T. (min)	S Count Charge	MS Area 0 %	MS Area 0
1		61	Parent	6065.02	5.81	50 From -3 To -	10 1.24E+06	
2		62	3'(0,1)	5760.97	5.69	15 From 3 To 1	9 1.29E+03	
3	×	G3	3'(n-4)	4845.83	3 5.12	10 From -3 To -	8	
4	~	G4	3'(n+dA)	6378.07	5.93	4 From -3 To -	6	
5		G5	5'(n-1)	5775.97	5.73	7 From -3 To	4 3.42E+03	
6		G6	5'(n-14)	1775.36	6 0.81	6 From -2 To -3	2 2.09E+03	. :
7		G7	5'(n-4)	4828.83	3 5.53	6 From -3 To -3	3 1.09E+03	
8		G8	n-dA@14	5751.97	5.68	3 From -3 To -3	3	
Сс	orrelat	ion De	tails					
- 🌒 3	"(n-1), Neu	tral Mass=5	760.97, RT=5.69	min				
= 3 3 3	I'(n-1), Neu I'(n-4), Neul I'(n+dA), Ne	tral Mass=5 Iral Mass=48 eutral Mass=1	760.97, RT=5.69 45.83, RT=5.12 r 8378.07, RT=5.9	min nin 3 min				
= 3 3 3	l' (n-1), Neu l'(n-4), Neul l'(n+dA), Ne 1.0e5 -	tral Mass=57 Iral Mass=48 eutral Mass=1	760.97, RT=5.69 145.83, RT=5.12 r 8378.07, RT=5.9	min min 3 min				
- 3 3 3	f'(n-1), Neu f'(n-4), Neu f'(n+dA), Ne 1.0e5 9.0e4	tral Mass=57 Iral Mass=48 eutral Mass=4	760.97, RT=5.69 145.83, RT=5.12 r 6378.07, RT=5.90	min nin 3 min				
- 3 3 3	f'(n-1), Neu F(n-4), Neu F(n+dA), Ne 1.0e5 9.0e4 8.0e4	tral Mass=57 Iral Mass=48 utral Mass≕i	760.97, RT=5.69 145.83, RT=5.12 (6378.07, RT=5.90	min min 3 min				
- 3 3	F(n-1), Neur F(n-4), Neur F(n+dA), Ne 1.0e5 9.0e4 8.0e4 7.0e4	tral Mass=5; Iral Mass=48 ⊮utral Mass≕i	760.97, RT=5.69 145.83, RT=5.12 6378.07, RT=5.9	min nin 3 min				
- • 3 • 3 • 3	r'(n-1), Neu ('(n-4), Neu ('(n+dA), Ne 1.0e5 - 9.0e4 - 8.0e4 - 7.0e4 - 6.0e4 -	tral Mass=5 Iral Mass=48 utral Mass≕	760.97, RT=6.69 45.83, RT=5.12 6378.07, RT=5.9	min min 3 min				
Area	r'(n-1), Neu r'(n-4), Neu r'(n+dA), Ne 1.0e5 9.0e4 8.0e4 7.0e4 6.0e4 5.0e4	tral Mass=50 tral Mass=48 eutral Mass=1	760.97, RT=5.69 45.63, RT=5.12 6378.07, RT=5.9	min min 3 min				
Area 3 3 3 3	(n-1), Neu (n-4), Neu (n+dA), Ne 1 0e5 - 9 0e4 - 8 0e4 - 7 0e4 - 6 0e4 - 5 0e4 - 4 0e4 -	tral Mass=5(tral Mass=4 uutral Mass=4	760.97, RT=5.69 45.83, RT=5.12 6378.07, RT=5.9	min min 3 min				
Area 3	F(n-1), Neu F(n-4), Neu (n-40A), Ne 1.0e5 9.0e4 8.0e4 7.0e4 6.0e4 5.0e4 4.0e4 3.0e4	trai Mass=6; frai Mass=48 uutrai Mass=4	760.97, RT-5.69 45939, RT-5.12, 6378.07, RT-5.9	min min 3 min				
Area	r(n-1), Neu (n-4), Neu (n-4), Neu 1.0e5 - 9.0e4 - 8.0e4 - 7.0e4 - 6.0e4 - 5.0e4 - 4.0e4 - 3.0e4 - 2.0e4 -	tral Mass=5 ral Mass=48 uutral Mass=1	760.97, RT5.0 45.83, RT-5.12 16378.07, RT-5.9	min rsn 3 mn	2.00			
Area	r(n-1), Neu r(n-4), Neu r(n+dA), Ne 1.0e5 - 9.0e4 - 8.0e4 - 7.0e4 - 6.0e4 - 5.0e4 - 3.0e4 - 2.0e4 - 1.0e4 -	tral Mass=48 Iral Mass=48 Iral Mass=1	760.97, RT-6.69 45.83, RT-5.12 16378.07, RT-5.9	min min min 100	200			

CONCLUSIONS

- analysis in support of these new modalities.
- oligonucleotide drug design.
- quantitative impurity profiles.

REFERENCES

TRADEMARKS/LICENSING

The SCIEX clinical diagnostic portfolio is For In Vitro Diagnostic Use. Rx Only. Product(s) not available in all countries. For information on availability, please contact your local sales representative or refer to www.sciex.com/diagnostics. All other products are For Research Use Only. Not for use in Diagnostic Procedures.

Frademarks and/or registered trademarks mentioned herein, including associated logos, are the property of AB Sciex Pte. Ltd. or their respective owners in the United States and/or certain other countries (see www.sciex.com/trademarks).

© 2021 DH Tech. Dev. Pte. Ltd. RUO-MKT-10-13959-A



Negative polarity LC HRMS with the information dependent acquisition gave isotopically resolved accurate mass data for series of tested sequences of oligonucleotides having molar masses up to 10000. The highresolution MS/MS data allowed both quantitative and qualitative analysis to be performed in a single run. Expanded LC-HRMS data processing workflow for metabolite and impurity profiling streamlined data

Five-part nucleotide residue description was developed to cover the breadth of emerging chemistries in the

The new tool, integrated into SCIEX OS software as a Molecule Profiler App, was instrumental in detecting, resolving and annotating expected and unexpected cleavage impurities as well as constructing semi-

Roberts, T.C., Langer, R. & Wood, M.J.A. Advances in oligonucleotide drug delivery. *Nat Rev Drug* Discov 19, 673–694 (2020). https://doi.org/10.1038/s41573-020-0075-7 2 Tozaki, T., Karasawa, K., Minamijima, Y., Ishii, H., Kikuchi, M., Kakoi, H., Hirota, K., Kusano, K., Nagata, S. A Marker-based Method to Detect Phosphorothioated Oligonucleotides in Equine Plasma Using Spectroscopic Analysis. Chromatography Today, August / September 2019, 50-54 3 Karasawa, K., Burton, L., Duchoslav, E. Method Development for Metabolite and Impurity Profiling of Oligonucleotide Therapeutics. ASMS Conference 2019, poster WP 631