Characterization and relative quantification of oligonucleotide impurities

Identify and confirm the structure of oligonucleotide impurities using the X500B QTOF system and Molecule Profiler software from SCIEX

Remco van Soest, Kerstin Pohl and Todd Stawicki
SCIEX, USA

This technical note describes the identification, relative quantification and structural confirmation of oligonucleotides and related impurities. Relative quantification of a spiked-in impurity could be achieved at 0.1% (*w*/*w*) and full sequence confirmation was obtained at a level as low as 1% (*w*/*w*).

Oligonucleotide therapeutics and gene therapies are rapidly gaining attention as their potency improves and delivery challenges are addressed. Modalities such as antisense oligonucleotides (ASOs) are becoming more important due to their high specificity and ability to reach formerly undruggable targets. To ensure safe drugs, methods for the identification and characterization of the full-length product (FLP) and impurities are critical. Accurate mass spectrometry can be used for the identification of potential impurities, by comparing the measured accurate masses and isotope patterns with those calculated. However, there is a lack of powerful yet intuitive processing software, limiting cumbersome and time-consuming manual interpretation. Furthermore, structural confirmation leveraging MS/MS adds an additional level of complexity.

This technical note shows the identification and relative quantification of the 5'(*n*-1) impurity of a fully phosphorothioated FLP spiked into an FLP sample at levels between 0.1 and 10% (*w*/*w*), using the Molecule Profiler software to overcome challenges related to data analysis (Figure 1). In addition, unknown impurities present in the FLP were identified as well. Relative quantification of impurities based on time-of-flight mass spectrometry (TOF MS) data was performed within the software and MS/MS fragment ions of the impurities were assigned to confirm their structures, facilitating the characterization of therapeutics in development.

Key features of the Molecule Profiler software for oligonucleotide impurity analysis

- Excellent data quality and high mass accuracies for TOF MS and TOF MS/MS allow for confident assignment of oligonucleotide FLP and impurities using the Molecule Profiler software
- Straightforward relative quantification based on TOF MS peak areas is achieved by grouping charge states
- Sequence coverages of the proposed sequences of impurities are displayed based on user-selectable fragment ion types

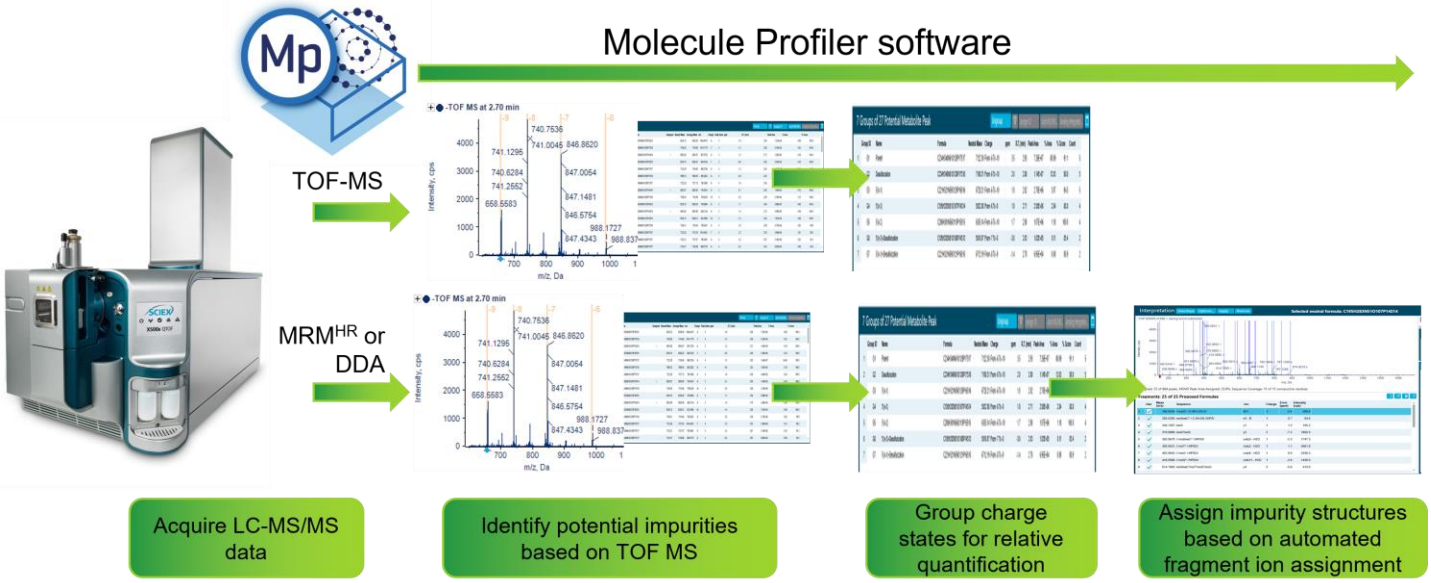


Figure 1. Workflow representation for relative quantification and structural confirmation of impurities using the Molecule Profiler software.

Methods

Samples and reagents: A 18-mer 2'-O-methoxyethyl phosphorothioated oligonucleotide (ASO-18PS2MOE-FLP) with the same sequence as nusinersen, a drug developed for the treatment of spinal muscular atrophy and its 5'(*n*-1) impurity were ordered (desalted) from Integrated DNA Technologies, Inc. The ion-pairing reagents 1,1,3,3,3-hexafluoroisopropanol (HFIP, ≥ 99.8%) and diisopropylethylamine (DIEA, ≥ 99.5%,) were purchased from Sigma Aldrich.

Sample preparation: Samples of FLP in mobile phase A were spiked with the 5'(*n*-1) shortmer at 0.1, 0.3, 1, 3 and 10% (*w*/*w* relative to a total amount of 20 µg/mL) to mimic a process-related impurity. The FLP was used as a control sample.

Chromatography: A Shimadzu LC-20 series HPLC system was used with water containing 15 mM N, N-diisopropylethylamine and 35 mM hexafluoro-isopropanol as mobile phase A and Methanol as mobile phase B. A gradient from 5–50% B in 9 min with a 2 min wash step at 95% B was used at a flow rate of 0.3 mL/min. The column was a Waters ACQUITY PREMIER Oligonucleotide C18 (2.1 × 150 mm, 1.7 µm, 130 Å) at 70ºC and the injection volume was 10 µL, resulting in 0.2 µg of total oligonucleotide on column.

Mass spectrometry: A SCIEX X500B QTOF system was used in negative polarity using an MRM^{HR} method (method details available on request). To determine the precursor masses, the data from a TOF MS scan of the 10% spiked sample was processed using the Molecule Profiler software and the *m*/*z* values for the most abundant charge states for the FLP, the spiked-in 5'(*n*-1) impurity and other abundant identified impurities were extracted. Collision-induced dissociation (CID) was used and collision energies (CE) were selected that ensured the generation of fragment rich MS/MS spectra. The parameters for the final MRM^{HR} method are summarized in Table 1.

Data processing: The molecular weight (MW) of the FLP was confirmed by reconstruction using the Bio Tool Kit in the SCIEX OS software. Data were processed using the SCIEX Molecule Profiler software. Considering the structure of the oligonucleotide, the number of bonds to break in the parent structure and a comprehensive list of possible transformations (77 for the work presented), the software identifies the different charge states of potential impurities. This assignment is based on the accurate mass match (here a 20 ppm tolerance was used) and the match between the measured and theoretical isotope ratios. Additional transformations can be added, while also custom nucleotide residues can be used by defining the 5' and 3' linkers, sugar, base and phosphate groups (Figure 2). The list of found potential impurities can be manually curated and the different charge states of each potential impurity are grouped together for relative quantification based on the TOF MS data.

For confirmation of the structure of each potential impurity, MS/MS spectra can be automatically annotated with *a*, *b, c, d* and *w*, *x* and *y* fragments, including water and base losses (see Figure 2). Sequence coverage is automatically determined by the software, considering user-selectable fragment types, signal-to-noise ratio and the number of bonds to break. A map of the proposed structure, indicating the fragments for which MS/MS evidence was found, facilitates the assessment of sequence coverages of analytes.

Table 1. MS parameters.

Parameter	MS	MS/MS
Scan mode	*TOF MS*	*MRM^{HR}*
Polarity	*negative*	
Gas 1	*60 psi*	
Gas 2	*60 psi*	
Curtain gas	*30 psi*	
Source temperature	*350°C*	
Ion spray voltage	*-4,500 V*	
Declustering potential	*-80 V*	
CAD gas	*7*	
Start mass	*600 m/z*	*150 m/z*
Stop mass	*2,000 m/z*	*3,000 m/z*
Q1 resolution	*N/A*	*Low*
Accumulation time	*0.2 s*	*0.03 s*
Collision energy	*-10 V*	*available upon request*
CE spread	*0 V*	*5 V*
Time bins to sum	*6*	*12*
QJet RF amplitude	*190 V*	*190 V*

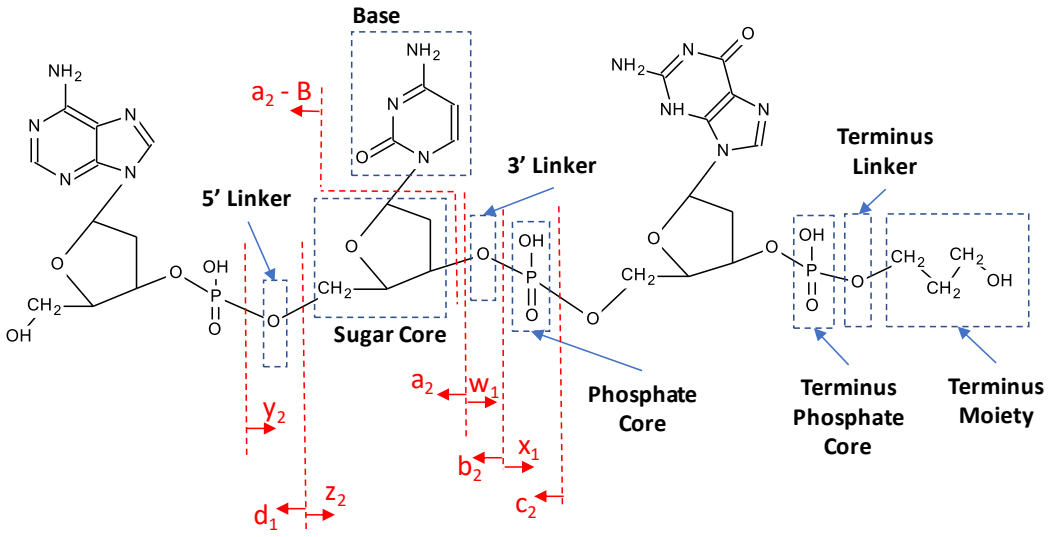


Figure 2. Representation of an oligonucleotide showing the different groups that can be defined for custom nucleotides. Also, the nomenclature of the MS/MS fragments used for sequence confirmation is shown in this figure, taken from the Molecule Profiler software.

Raw data evaluation and MW determination

High data quality in negative ion mode is incumbent for any subsequent analysis of oligonucleotide data. As an example, Figure 3 shows the TOF MS spectrum of the FLP acquired with the X500B QTOF system, a zoom-in of the -8-charge state and the reconstructed masses as calculated by the Bio Tool Kit software within SCIEX OS software. The accurate X500B QTOF system allows for baseline resolution of the isotopes of each charge state. Excellent mass accuracy below 2 ppm was observed upon isotopically resolved reconstruction of the raw data with a maximum entropy algorithm. Reconstruction of raw data can greatly simplify data review as all charge states are collapsed to a 0 charged mass. Therefore, accurate reconstruction is of high importance for judging data quality in a simplified manner.

Considering oligonucleotides can contain a variety of species, which are close in mass or isobaric, MS/MS is a necessity to verify product quality and identify unknowns. Figure 4 shows an example of the raw data quality of an MS/MS spectrum derived from the spike-in of a shortmer. The high data quality at MS/MS level is key to ensure correct matching and understanding sequences as discussed in the next sections.

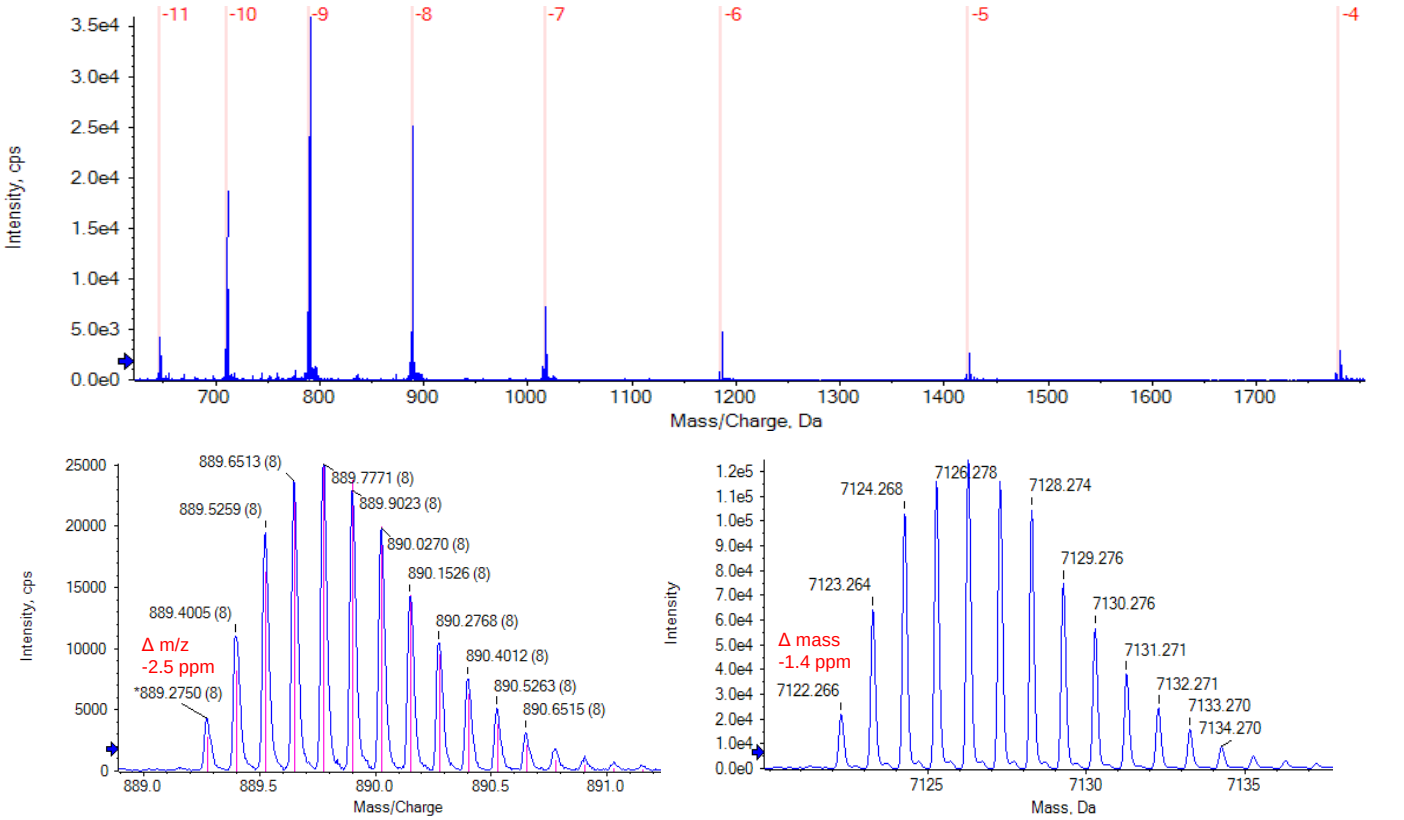


Figure 3. Data from ASO-18PS2MOE-FLP. Top figure shows the full TOF MS spectrum with a charge state envelope from -4 to -11. The lower figure shows a zoom-in of the -8-charge state (left) and the isotopically resolved reconstruction with a step mass of 0.03 Da (right). The mass error was -1.4 ppm for the monoisotopic mass of the reconstructed data.

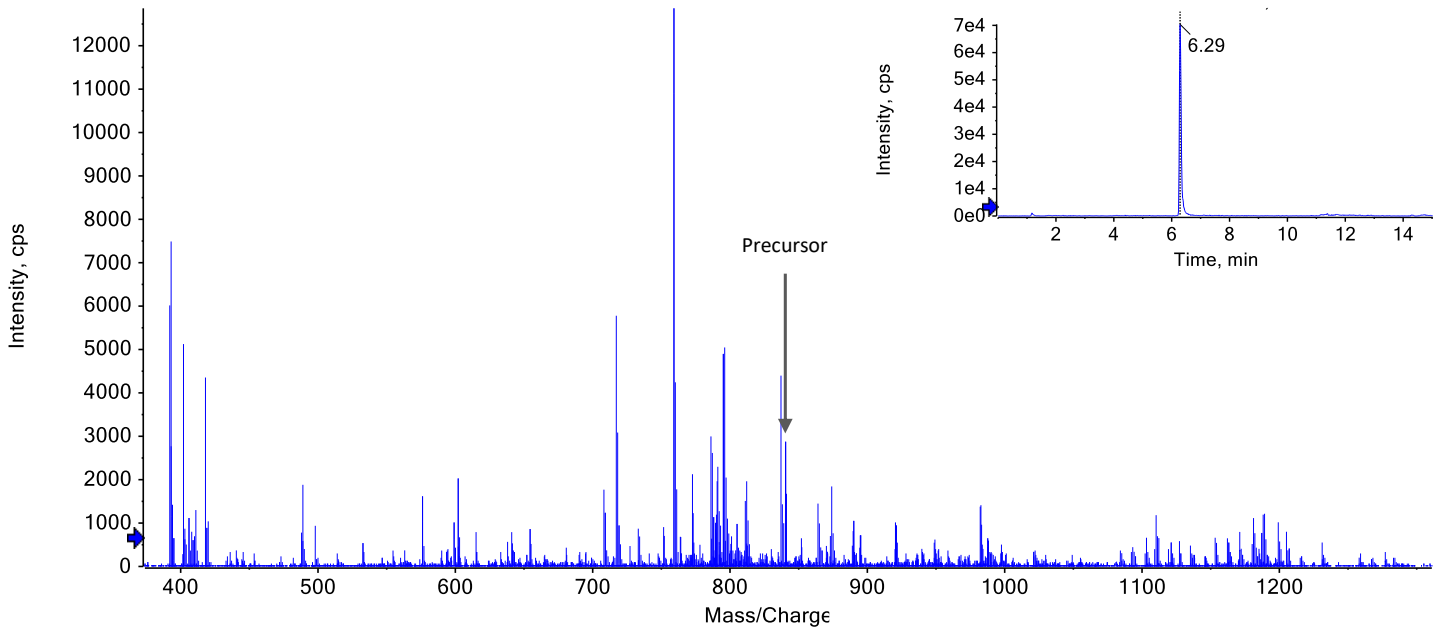


Figure 4. Data of the [M-8H]8- charge state of the 5'(*n*-1) shortmer at the 10% spike-in level. Main figure shows the MS/MS data with precursor indicated. Rich fragmentation was achieved. The inset on the top right shows the XIC of the precursor (*m/z* = 840.5) of the TOF MS data.

Identification of FLP and impurities based on TOF MS data

Identifying oligonucleotides and their impurities can be a cumbersome task. Traditionally, there has been a lack of intuitive and powerful software, allowing for all needed modifications on the phosphate backbone or the bases, which are linked to a variety in proprietary drug modifications. Here, the Molecule Profiler software was used to identify the FLP and related impurities (shortmers) based on the user input of the known sequence of the FLP automatically. As a first step, identification is carried out using the matching of TOF MS data with theoretical information. Ppm error cut offs and isotope ratios are used. Figure 5 shows a zoom-in to the measured and centroided MS spectrum as dark blue sticks and the theoretical isotope patterns as light blue profiles for 2 of the charge states of the FLP.

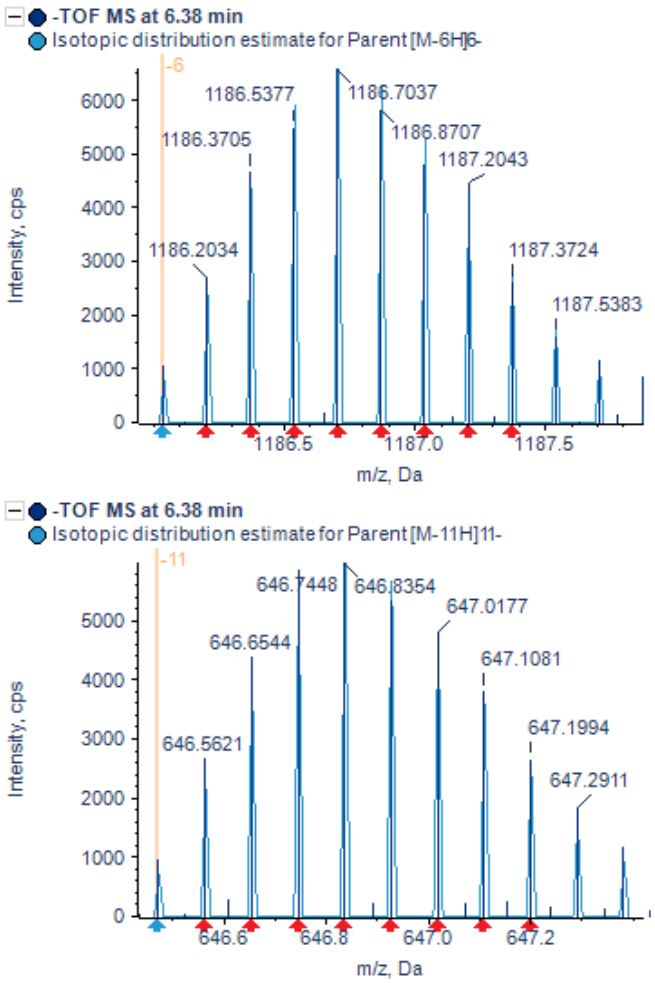


Figure 5. Example charge states for ASO-18PS2MOE-FLP. Zoom-in to the [M-6H]6- (top) and [M-11H]11- (bottom) charge states of ASO-18PS2MOE-FLP shows an excellent match between the theoretical isotope distribution (light blue) and the centroided MS spectrum (dark blue).

The monoisotopic peak is indicated by the blue arrow, whereas the first 8 isotopes are indicated by red arrows. The excellent match of monoisotopic mass, isotopes and isotope ratios allows for confident identification.

Each sample was injected in triplicates and processed individually. The scoring was performed using the matching of isotope ratios, while an *m/z* tolerance of 10 mDa was used for matching. Charge states from -4 to -11 were considered for matching and a low scoring cut-off (40%) was selected to ensure that low abundance impurities were considered. In addition to terminal fragments, the matching of internal fragments was allowed for the search despite those typically being found with an %area below 0.1% (data not shown).

Figure 6 shows the top 15 of 43 identified potential impurities with a mass accuracy < 10 ppm, based on %area after grouping charge states for an injection of the ASO-18PS2MOE-FLP spiked with 3% (of total weight) of its 5'(*n*-1) shortmer. The %area is calculated by summing the areas of all charge states assigned to a particular analyte and dividing it by the total area of all charge states of all assigned analytes.

43 Groups of 89 Potential Metabolite Peaks

Name	Neutral Mass	ppm	% Area	Count
Parent	7122.22	-7.2	69.34	8
Desulfurization	7106.25	-7.3	15.05	8
5'(n-1)	6728.17	-8.3	2.77	7
Isobutyryl (iBz)	7192.28	-5.1	2.57	8
Loss of Sulfur	7090.27	-1.6	1.25	4
Depurination of G	6989.19	-6.8	0.80	2
Phosphorylation + Depyrimidination of U	7108.25	3.4	0.45	1
DMF	7177.30	-4.4	0.67	3
CNEt	7175.27	-4.7	0.36	2
5'(n-3)	5932.03	-7.6	0.42	3
5'(n+moA*)	7525.30	-5.7	0.19	1
3'(n-5)+3':H2PSO2/+Loss of Water	5155.84	-5.2	0.26	2
5'(n+moG*)	7541.35	0.8	0.15	1
DMF	7177.30	-4.1	0.21	2
Phosphorylation + Loss of G	7051.19	-0.7	0.14	1

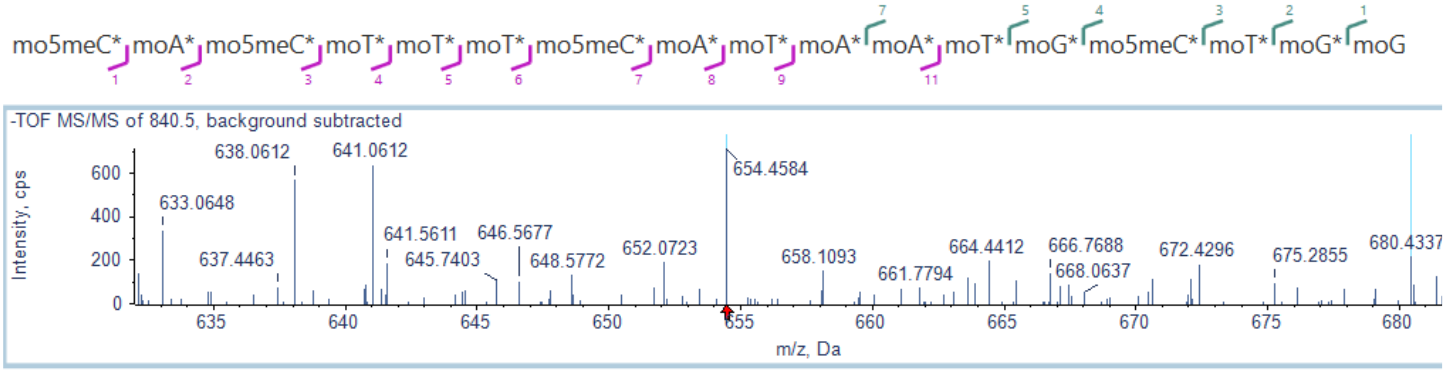
Figure 6. Top 15 of 43 identified potential impurities for spiked ASO-18PS2MOE-FLP. FLP was spiked with 3% of 5'(*n*-1) shortmer and filtered to show only impurities with mass accuracy < 10 ppm. Impurities were sorted by decreasing %area after grouping identified charge states. "Count" shows the number of charge states of the impurities that were found. When multiple potential hits were listed for a specific peak, the lowest scoring hit was removed.

When multiple potential hits were listed for a specific peak, the lowest scoring hit was removed manually. The spiked 5'(*n*-1) impurity was identified with 7 charge states and with a 2.77% area, which matches with the calculated spike-in of 3%. The high percentage of the "Desulfurization" impurity was likely caused during sample storage. Although in-source formation of this product is possible, it could be ruled out as a difference in retention time for the FLP and desulfurization product was observed. The "Isobutyryl" impurity is likely to originate from the use of isobutyryl as a protecting group for 2'-O-methoxyethyl G during synthesis, while "CNET" (cyanoethylthymine) is also used as a protection group. Other identified potential impurities are either byproducts from the synthesis or formed during sample storage. Since the oligonucleotide samples used in this study were not HPLC purified after synthesis, the identification of impurities across all replicates are plausible.

Confirmation of structures using MS/MS data

Therapeutic oligonucleotides are synthesized products, which may contain a variety of impurities derived from each step of the synthesis. The impurities can be isobaric in nature, therefore MS information alone is not enough to understand their identity. The fragmentation of oligonucleotides is quite complex resulting in fragments with highly similar *m*/*z*. Data with high mass accuracy and intuitive software are essential to confirm the identity of analytes.

The Molecule Profiler software used here was to annotate the peaks in the MS/MS spectra and generate a sequence coverage map. For both the FLP and 5'(*n*-1) impurity at the 1% spike-in level the full sequences could be confirmed using terminal fragment ions with low mass errors (< 10 ppm). Figure 7 shows the sequence coverage as displayed in the Molecule Profiler software for the 5'(*n*-1) impurity at the 3% spike-in level. The graphical display of the sequence with matched fragments and numbering of those, facilitates review and verification of the identifications. Likewise, evidence could be seen in the MS/MS spectra of the potential isobutyryl impurity, which was identified with TOF MS at < 3%. Assigning the isobuturyl group to the last guanine in the sequence, multiple 3' terminal fragments were identified in the MS/MS spectrum that matched with the proposed modification providing further evidence of this modification being present (Figure 8).

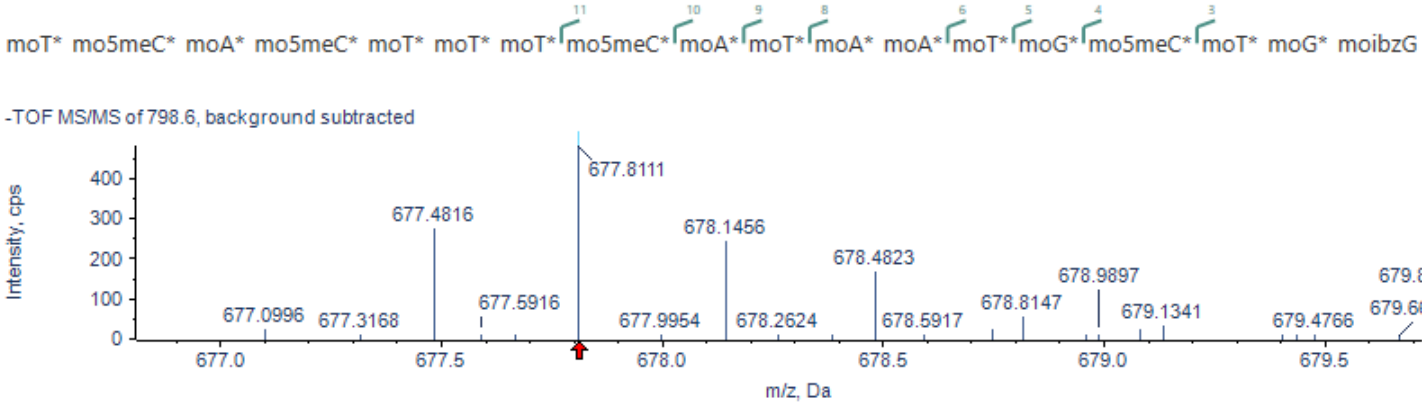


Assigned: 44 of 204 peaks, MSMS Peak Area Assigned: 38.4%, Sequence Coverage: 17 of 17 residues

Fragments: 44 of 44 Proposed Formulae

	Use	Mass (m/z)	Sequence	Ion	Charge	Error (ppm)	Intensity (cps)
1	✓	689.8795	moA*moT*moG*mo5meC*moT*moG*moG	y7	4	4.9	54.4
2	✓	982.1923	moG*mo5meC*moT*moG*moG	y5	2	-1.7	544.7
3	✓	654.4584	moG*mo5meC*moT*moG*moG	y5	3	-2.8	164.7
4	✓	772.6630	mo5meC*moT*moG*moG	y4	2	3.0	522.3
5	✓	1153.2501	moT*moG*moG	y3	1	-2.3	317.8

Figure 7. Confirmation of the sequence of the ASO-18PS2MOE 5'(*n*-1) shortmer spiked-in at the 3% level using MS/MS of the [M-8H]8- charge state. Using *a*, *w*, *c*, *y* and *d* terminal ions and allowing for max. 1 water or base loss, the full sequence could be confirmed considering fragments with a minimal S/N ratio of 50. MS/MS spectrum was centroided and de-isotoped by the Molecule Profiler software to facilitate review of data. Note that only the most intense peaks (44) were assigned, resulting in an assigned peak area of 38.4%.



Assigned: 14 of 668 peaks, MSMS Peak Area Assigned: 2.9%, Sequence Coverage: 11 of 18 residues

Fragments: 14 of 14 Proposed Formulae

	Use	Mass (m/z)	Sequence	Ion	Charge	Error (ppm)	Intensity (cps)
1	✓	677.8111	moG*mo5meC*moT*moG*moibzG	y5	3	5.3	479.6
2	✓	808.1570	moA*moA*moT*moG*mo5meC*moT*moG*moibzG	y8	4	3.0	251.6
3	✓	1017.2219	moG*mo5meC*moT*moG*moibzG	y5	2	7.0	164.8
4	✓	805.7473	moA*moT*moA*moA*moT*moG*mo5meC*moT*moG*moibzG	y10	5	-1.5	142.9
5	✓	704.4609	/H2PSO:/ moG*mo5meC*moT*moG*moibzG	x5	3	5.6	65.6
6	✓	847.6518	/H2PSO:/ mo5meC*moT*moG*moibzG	x4	2	-4.9	54.4

Figure 8. Confirmation of the sequence of the potential ASO-18PS2MOE impurity with 1 isobutyryl protecting group on G. 14 3' terminal fragment ions with S/N >50 in the MS/MS spectrum of the [M-9H]9- charge state were consistent with the presence of 1 isobuturyl group on the last guanine. Note that only the most intense peaks (14) were assigned, resulting in an assigned peak area of 2.9%.

Relative quantification

Identification of the main compound and impurities based on MS and MS/MS data is the first important step. However, as part of the characterization of a sample relative quantification of the different compounds is needed to understand the quality of a batch. Answering questions if a certain purity was achieved or if certain impurities are not exceeding set thresholds require relative quantification. As an example, the Molecule Profiler software was used to determine the relative amount (based on the MS summed intensities of all found charged states including isotopes) for the spiked 5'(*n*-1) shortmer. Great linearity with an R^2 of 0.9992 was found (Figure 9). The reported %areas of the 3 replicate injections for each spike-in level are listed in Table 2. Considering the (low) level of the 5'(*n*-1) shortmer found in the FLP control sample without spike-in (data not shown) and the purity of the 5'(*n*-1) shortmer itself of 76% (data not shown), the correlation of the calculated %areas with the spiked- amounts was excellent and precision was good as well (Table 2).

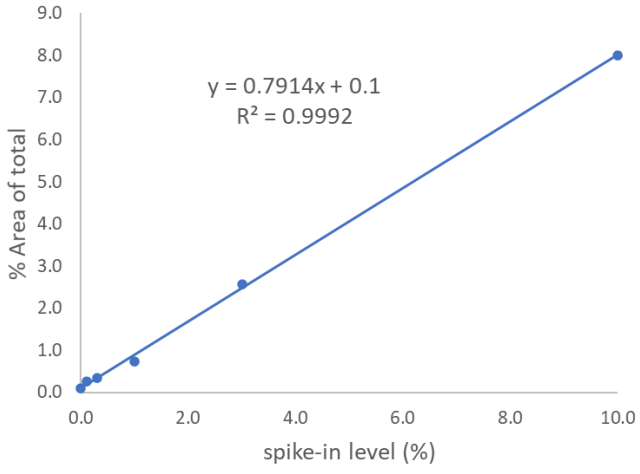


Figure 9. %area of total area for the 5'(*n*-1) shortmer as a function of the spike-in percentage. Great linearity with an R^2 of 0.9992 was observed.

Table 2. Average %areas measured for the different 5'(*n*-1) spike-in levels (*n* = 3). As the spiked-in shortmer was only 76% pure as determined by TO -MS and Molecule Profiler software, the measured values were in very good correlation with the spiked-in amounts.

Spike-in level (%)	Average %area of 5'(*n*-1)	RSD (%)
solvent blank	*0.00*	*0.00*
0.00	*0.10**	*103*
0.10	*0.26*	*3.85*
0.30	*0.34*	*6.73*
1.0	*0.73*	*13.7*
3.0	*2.6*	*4.12*
10	*8.0*	*3.47*

*The FLP was not HPLC purified and was found to contain a small amount of 5'(*n*-1) impurity, which was detected in 2 out of 3 replicates.

Conclusions

- The risk of missing important analytes can be minimized by excellent TOF MS data quality and mass accuracy, which allows for the detection of oligonucleotides and impurities down to the 0.1% level
- Customizable modifications of oligonucleotide backbone (linker, sugar, termini) and bases in Molecule Profiler software as well as a wide range of pre-populated modifications and transformations allow for speeding up the analysis of proprietary oligonucleotides
- Relative quantification showed good correlation for samples spiked with a shortmer impurity from 0.1% to 10% (*w/w* of total) proving the capability of this assay for stream-lined relative quantification in Molecule Profiler software
- The confirmation of the structure of oligonucleotide-based drugs and impurities is greatly accelerated by automatic annotation of TOF MS/MS spectra with the most commonly found fragment ions based on a proposed structures in Molecule Profiler software

Headquarters
500 Old Connecticut Path | Framingham, MA 01701 USA
Phone 508-383-7700
sciex.com

International Sales
For our office locations please call the division headquarters or refer to our website at
sciex.com/offices