Monoisotopic mass confirmation of modified and unmodified oligonucleotides by LC-HRAM MS

Authors
Hao Yang¹, Keeley Murphy², Kristi Akervik³, Min Du²
¹Thermo Fisher Scientific, San Jose, CA, US
²Thermo Fisher Scientific, Boston, MA, US
³Thermo Fisher Scientific, Austin, TX, US

Application benefits
Quick and confident monoisotopic mass confirmation of modified and unmodified oligonucleotides up to 100mer with less than 3 parts per million (ppm) mass accuracy using an LC-HRAM MS method that was developed on a Thermo Scientific™ Orbitrap Exploris™ MX mass detector

Goal
• Demonstrate the use of UHPLC-HRAM MS with the Thermo Scientific™ Chromeloen™ Chromatography Data System (CDS) for accurate monoisotopic mass confirmation of unmodified and modified oligonucleotides
• Demonstrate the use of the Thermo Scientific™ BioPharma Finder™ software oligonucleotide mass calculator for determining the chemical formula, monoisotopic and average mass based on input oligonucleotide sequences

Therapeutic oligonucleotide characterization and analysis is an important component of drug development and quality control. Given the diversity of the types of oligonucleotides, a robust and accurate analytical method is needed to confirm the identity and determine the purity for quality control needs. While conventional intact mass analysis of oligonucleotide products is performed using nominal mass detectors to provide average mass confirmation, high-resolution accurate-mass mass spectrometric techniques, which allow monoisotopic mass confirmation with a high degree of mass accuracy, have gained increasing attention.

Keywords
Oligonucleotide analysis, intact mass determination, monoisotopic mass, liquid chromatography high-resolution accurate-mass mass spectrometry (LC-HRAM-MS), ion-pairing reversed phase liquid chromatography (IPRP-LC), Vanquish Horizon UHPLC, Orbitrap Exploris MX mass detector, DNAPac reversed phase column, Chromeloen Chromatography Data System (CDS), BioPharma Finder oligonucleotide mass calculator
Herein, we describe an intact mass analysis that uses the Xtract algorithm in Chromeleon CDS for monoisotopic mass confirmation of modified and unmodified oligonucleotides with lengths up to 100 nucleotides. Intact mass analysis was executed in four quick steps as shown in Figure 1. First, we obtained the chemical formula by entering the oligonucleotide sequence in the oligonucleotide mass calculator in BioPharma Finder software, version 5.2. Next, the chemical formula and set mass accuracy were entered in the target formula and target tolerance columns, respectively, in the injection sequence prior to data acquisition. After data were acquired, mass deconvolution was performed using the Xtract algorithm with a sequence specific isotope table. Lastly, a report was generated showing the evaluation of measured monoisotopic masses against the theoretical monoisotopic masses that were calculated based on the chemical formula of the input sequences. Using this method, accurate monoisotopic mass confirmation with less than 3 ppm mass accuracy was achieved for all tested oligonucleotides.

Figure 1. Schematics for intact mass analysis of modified and unmodified oligonucleotides. Step 1: input oligonucleotide sequences into the oligonucleotide mass calculator in BioPharma Finder software to generate the chemical formula. This step is optional if the user already has the chemical formula. Step 2: import the chemical formula list to the target formula column and enter the set mass accuracy in the target tolerance column in the injection sequence. Step 3: acquire and process raw data using the Xtract algorithm with a sequence specific isotope table. Step 4: evaluate the measured monoisotopic mass and mass accuracy for all tested samples in the summary report. Both Steps 3 and 4 are automated with auto reporting enabled.
**Experimental**

**Reagents and consumables**

- Oligonucleotide samples, HPLC purified, see Table 1 for details
- 1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP), 99.9% (Thermo Scientific™ Chemicals, P/N 293410500)
- Triethylamine (TEA), 99% (Thermo Scientific™ Chemicals, P/N 157911000)
- Dibutylamine (DBA), 99.5% (Sigma-Aldrich, 471232-100ML)
- Thermo Scientific™ DNAPac™ RP HPLC column, 2.1 × 50 mm, 4 µm (P/N 088924)
- Thermo Scientific™ Water, UHPLC-MS grade (P/N W8-1)
- Thermo Scientific™ Methanol, UHPLC-MS grade (P/N A458-1)
- Thermo Scientific™ Acetonitrile, UHPLC-MS grade (P/N A956-1)
- Thermo Scientific™ 9 mm Screw Thread Vials, Polypropylene, 12 × 32 mm, 400 µL (P/N C4000-11)
- Thermo Scientific™ 9 mm Autosampler Vial Screw Thread Caps, Polypropylene (P/N C5000-50)

**Oligonucleotide sample preparation**

Oligonucleotide samples with lengths ranging from 10 to 55 were obtained from Life Technologies and were prepared at a concentration of 2.5 pmol/µL. The 100 oligomer and phosphorothiolated (PS) oligonucleotides were purchased from Integrated DNA Technologies and were reconstituted in UHPLC grade water at a concentration of 1 pmol/µL. Detailed oligonucleotide sequences, chemical formula, and theoretical monoisotopic mass for each of the samples are listed in Table 1.

---

**Table 1. Oligonucleotide sequences, information, chemical formula, and theoretical monoisotopic mass.** Both chemical formula and theoretical monoisotopic mass were obtained from the oligonucleotide mass calculator in BioPharma Finder software, version 5.2.

<table>
<thead>
<tr>
<th>Oligonucleotide name</th>
<th>Sequence*</th>
<th>Chemical formula</th>
<th>Monoisotopic mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10mer</td>
<td>GAG CGG CTG T</td>
<td>C98H123O59N40P9</td>
<td>3082.5493</td>
</tr>
<tr>
<td>20mer</td>
<td>GAG CGG CTG TGA GCG GCT GT</td>
<td>C196H245O120N80P19</td>
<td>6227.0543</td>
</tr>
<tr>
<td>30mer</td>
<td>GAG CGG CTG TGA GCG GCT GTG AGC GGC TGT</td>
<td>C294H367O121N120P29</td>
<td>9371.5593</td>
</tr>
<tr>
<td>40mer</td>
<td>GAG CGG CTG TGA GCG GCT GTG AGC GGC TGT GAG CGG CTG T</td>
<td>C392H489O242N160P39</td>
<td>12516.0643</td>
</tr>
<tr>
<td>50mer</td>
<td>GAG CGG CTG TGA GCG GCT GTG AGC GGC TGT GAG CGG CTG TGA GCG GCT GTG AGC GGC TGT</td>
<td>C490H611O303N200P49</td>
<td>15660.5693</td>
</tr>
<tr>
<td>55mer</td>
<td>GAG CGG CTG TGA GCG GCT GTG AGC GGC TGT GAG CGG CTG TGA GCG GCT GTG AGC GGC TGT GAG CGG CTG TGA GCG GCT GTG AGC GGC TGT</td>
<td>C539H671O332N223P54</td>
<td>17249.8309</td>
</tr>
<tr>
<td>PS_Full</td>
<td>Ur-sAr-sAr-sCr-sAr-sGr-sAr-sUr-sCr-sGr-sGr-sCr-sCr-sUr-sGr-sGr-sAr-sAr-sUr</td>
<td>C190H237O118N76P19S19</td>
<td>6666.4589</td>
</tr>
<tr>
<td>PS_1</td>
<td>Ur-sAr-sAr-sAr-sGr-sAr-sUr-sAr-sCr-sGr-sGr-sCr-sCr-sGr-sGr-sAr-sGr-sAr-sUr-sAr-pUr</td>
<td>C190H237O119N76P19S18</td>
<td>6650.4817</td>
</tr>
<tr>
<td>PS_2</td>
<td>Ur-pAr-sCr-sAr-sAr-sGr-sAr-sUr-sAr-sCr-sGr-sGr-sCr-sGr-sGr-sAr-sAr-sAr-pUr</td>
<td>C190H237O120N76P19S17</td>
<td>6634.5046</td>
</tr>
<tr>
<td>PS_3</td>
<td>Ur-pAr-sCr-sAr-sAr-sGr-sAr-sUr-sAr-sCr-sGr-sGr-sGr-sGr-sAr-sAr-sAr-pUr</td>
<td>C190H237O121N76P19S16</td>
<td>6618.5274</td>
</tr>
<tr>
<td>PS_4</td>
<td>Ur-pAr-sCr-pAr-sGr-sAr-sUr-sAr-sCr-sGr-sCr-sGr-sGr-sGr-sAr-sAr-pAr-pUr</td>
<td>C190H237O122N76P19S15</td>
<td>6602.5503</td>
</tr>
<tr>
<td>PS_5</td>
<td>Ur-pAr-sCr-pAr-sGr-sAr-sUr-sAr-sGr-sGr-sGr-sGr-sGr-sAr-sAr-pAr-pUr</td>
<td>C190H237O123N76P19S14</td>
<td>6586.5731</td>
</tr>
<tr>
<td>100mer</td>
<td>CAG TCA GTC AGT CAG TCA GTC AGT CAG TCA GTC AGT CAG TCA GTC AGT CAG TCA GTC AGT CAG TCA GTC AGT CAG TCA GTC AGT CAG TCA GTC AGT CAG T</td>
<td>C975H1226O598N375P99</td>
<td>30818.1076</td>
</tr>
</tbody>
</table>

*Sequence annotations
ATCG: DNA bases; U: Uracil; r: ribose; p: phosphodiester bond; s: phosphorothioate bond
Chromatography
The Thermo Scientific™ Vanquish™ Horizon UHPLC system was used for IPRP-LC separation of the oligonucleotide standards. For the 100mer oligo, a gradient of 1–25 %B at a flow rate 0.25 mL/min over 11 minutes with a mobile phase consisting of 15 mM DBA and 25 mM HFIP in water (solvent A) and acetonitrile (solvent B) was used to separate its impurities from the full-length product. For the other oligonucleotides, a flow rate of 0.7 mL/min with a previously reported step gradient consisting of 1 %B for 0.4 minutes, 25 %B for 0.6 minutes, 100 %B for another 0.6 minutes, followed by 1 %B for 2.4 minutes was applied. The column temperature was set to 70 °C in forced air mode. For all sample analyses, 10 µL of sample was injected onto the DNAPac column.

Mass spectrometry
A full scan method with a resolution setting of 120,000 at m/z 200, mass range from m/z 600 to 1600, and 3 microscans was applied using the Orbitrap Exploris MX mass detector. For the 100mer, data were acquired in intact protein mode with a low pressure setting, whereas for the other oligonucleotides, data were acquired in peptide mode with the standard pressure setting. For experiments performed at 0.25 mL/min, the negative ion spray voltage was set at 2,500 V, sheath/aux/sweep gases were set at 50/10/1, respectively, and ion transfer tube and vaporizer temperatures were set at 325 and 350 °C, respectively. For experiments performed at 0.7 mL/min, the negative ion spray voltage was set at 3,000 V, sheath/aux/sweep gases were set at 60/15/2, respectively, and ion transfer tube and vaporizer temperatures were set at 350 °C.

Oligonucleotide mass calculator
The chemical formula is required for mass accuracy calculation and the use of a sequence specific isotope table for mass deconvolution of sulfur-containing oligonucleotides. As shown in Figure 1, the chemical formula for each tested sample was obtained from the input sequence using the oligonucleotide mass calculator in BioPharma Finder software, version 5.2.

Intact mass analysis using compliant-ready Chromeleon CDS 7.3.2
The data processing step for the intact mass analysis was executed automatically using pre-defined processing methods. The methods are detailed here. Full scan data were processed using the Xtract deconvolution algorithm under the integrated intact protein deconvolution feature within Chromeleon CDS version 7.3.2. This algorithm is specifically designed for processing of isotopically resolved accurate mass data. For the 100mer oligo, a single peak was detected within the retention time window between 11.8 to 12.1 minutes; whereas for the other oligonucleotides, a single peak was detected within the retention time window between 0.75 to 0.95 minutes. For each peak, the source spectrum was automatically generated from averaging the data points across the peak. Unless specified, default settings in “average over selected retention time with Xtract” method were used for mass deconvolution. The oligonucleotide output mass range was set between 1,000 and 35,000 Da, and the charge range was set between 3 and 50.

Intact mass analysis summary report
As shown in Figure 2, an intact mass analysis summary report template was built based on the default oligonucleotide report for optimal presentation of the obtained results. In this report, both expected and measured monoisotopic mass, mass accuracy, and target mass tolerance for all tested samples are shown.

**Figure 2. Example intact mass analysis summary report.** The report reflects the assessment of the mass accuracy for each analyzed sample based on the provided target chemical formula, and reflects a “pass” result when the achieved mass accuracy is below the target mass tolerance set to 3 ppm.
Results and discussion

Intact mass analysis of unmodified, single-stranded DNA with length ranging from 10 to 100mer and partial and fully phosphorothiolated RNA (PS_Full) with 20 nucleotide bases was performed using a high-resolution accurate-mass method with a scan resolution of 120,000 at m/z 200. As illustrated in Figure 3 for the analysis of PS_Full, the source spectrum was generated by averaging full scan data across the highlighted peak in the total ion chromatogram (TIC) with a retention time ranging from 0.75 to 0.95 minutes. A total of seven charge states ranging from -5 to -11 were observed, which resulted in a deconvoluted monoisotopic mass of 6666.4719. Similarly, a charge state profile ranging from -11 to -20 was observed for the analysis of 55mer, which resulted in a deconvoluted monoisotopic mass of 17249.8251 as shown in Figure 4; a charge state profile ranging from -20 to -42 was observed for the analysis of 100mer, which resulted in a deconvoluted monoisotopic mass of 30818.1371 as shown in Figure 5. With this method, isotopic resolution of all peaks was achieved, and the measured monoisotopic masses were all within 3 ppm mass accuracy with respect to the expected monoisotopic masses. When a resolution setting of 60,000 was used for the analysis of the 100mer, peaks could not be isotopically resolved, and the deconvolution provided an average mass instead of monoisotopic mass (data not shown). The measured monoisotopic masses are compared to the expected monoisotopic masses calculated from the input chemical formula in the sequence, and the results are shown in a summary report illustrated in Figure 2. A green “pass” is awarded if the absolute value of the achieved mass accuracy is below the target mass tolerance, here set to 3 ppm.
Conclusion

We developed a simple full scan method with mass resolution of 120,000 at m/z 200 on the Orbitrap Exploris MX mass detector for intact MS analysis of unmodified and modified oligonucleotides. This method provides the following:

- A quick way to obtain chemical formula, monoisotopic mass, and average mass for any given oligonucleotide sequences using the oligonucleotide mass calculator within BioPharma Finder software, version 5.2
- Confident monoisotopic mass confirmation with less than 3 ppm mass accuracy for all tested oligonucleotides
- A summary report for quick evaluation of the measured mass accuracy against set target mass accuracy representing pass/fail results for each tested oligonucleotide

Reference