Charge Deconvolution and Automatic Sequence Matching for Oligonucleotides

Wilfred Tang[1], Marshall Bern[1], Rose Lawler[1], James Moore[1], David Garby[2], Nicholas Skizim[2]
[1] Protein Metrics Inc., Cupertino, CA [2]GreenLight Biosciences, Inc., Medford, MA Contact: rlawler@proteinmetrics.com www.proteinmetrics.com

Introduction

Synthetic and modified oligonucleotides are receiving renewed attention due to new chemical modifications that improve their drug properties. This has led to more oligonucleotide therapeutics in the pharmaceutical pipeline.

A practical analytical method for analyzing oligonucleotides involves reverse-phase HPLC with ion pairing agents that complex with the negatively-charged phosphate groups of the oligonucleotide backbone. RP-IP-HPLC utilizing triethylammonium acetate and hexafluoroisopropanol as ion pairing agents, coupled to negative-mode ESI-MS on a single quadrupole, QTOF, or Orbitrap instrument is a common combination.

Data analysis, however, can be challenging

- Many charge deconvolution software programs cannot handle negative-mode mass spectra
- There are a variety of instrument vendors and data file formats
- Spectra may be isotope-resolved, or not, or a combination depending on the charge
- Oligonucleotides are difficult to desalt, and the adduct peaks present an additional complication to peak mass assignment

Here we present an automated processing and reporting platform for the analysis of oligonucleotides using Intact Mass™ software from Protein Metrics Inc. We show data from the Waters MassPREP™ OST (Oligonucleotide Separation Technology) standard and from an experiment assessing mRNA capping.

Methods

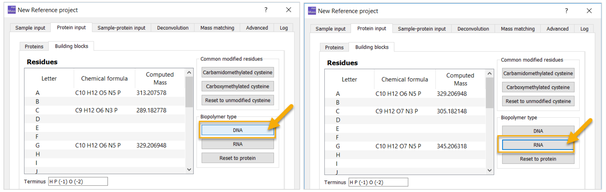
Waters MassPREP OST Standard. This standard contains approximately equimolar amounts of 15, 20, 25, 30, and 35 nucleotide sequences of (deoxy)thymidines with mass ranging from 4,500 to 10,585 Da. Data was acquired with negative mode ionization, resolution mode, on a Waters Xevo® G2 QTOF. The UPLC method utilized the Waters BEH C18 100 x 2.1 mm 1.7 µm column with triethylammonium acetate and hexafluoroisopropanol in methanol and water as ion pairing agents.

mRNA Capping Experiment. To assess the efficiency of 5'-cap addition to GreenLight Biosciences' mRNA products, we used an approach similar to that outlined in Beverly, et al. Briefly, we used a biotinylated oligonucleotide capture probe that would hybridize to and thereby capture the 5'-end of the target mRNA product (~30 nt). We then enzymatically cleaved the mRNA to free the non-hybridized portion and retain solely the bound 5'-end. After dissociating this fragment from the capture probe and desalting, we were able to directly analyze the 5'-end of the mRNA product as an ESI-friendly oligonucleotide. We utilized UPLC and QTOF conditions as described above.

Data Analysis. The data was analyzed using a development build of Protein Metrics Intact Mass software to produce deconvoluted neutral masses, which were automatically matched against a list of target species.

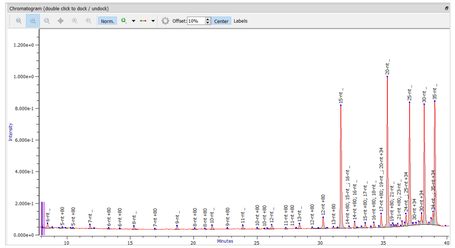
Software Platform

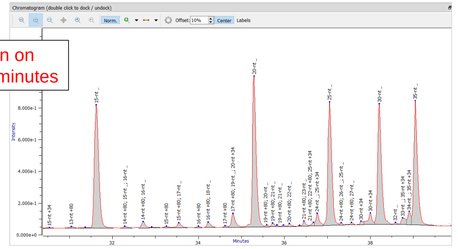
Protein Metrics Intact Mass has been enhanced to better support the analysis of different types of biopolymers. This includes a new set of three buttons to specify the type of molecule being studied (DNA, RNA, or protein). When a button is pressed, the appropriate chemical formula and mass for each one-letter code are automatically filled in.



Waters MassPREP OST Standard

We identified all of the expected species (oligonucleotides of length 15, 20, 25, 30, and 35), along with many less abundant species (nearly all oligonucleotides of length 6 through 35 as well as oligonucleotides with an additional mass of +80 or +34).





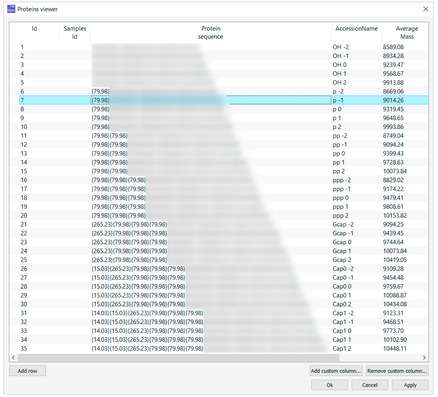
Zoom in on 30-40 minutes

mRNA Capping Experiment

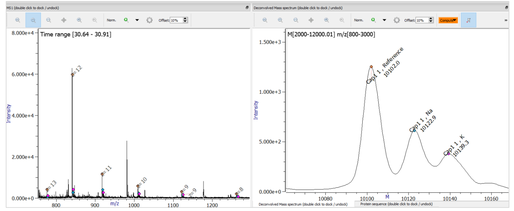
Sequence and/or mass can be used to specify the expected targets.

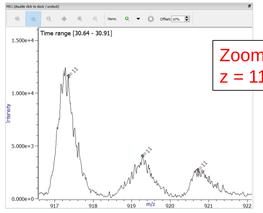
For the mRNA capping experiment, the target list had 35 species

- 7 different chemistries at the 5' end
- 5 different oligonucleotide lengths (due to the uncertainty of exactly where RNase H would cleave, we considered the target sequence as well as shorter by 1, shorter by 2, longer by 1, and longer by 2)



The figures below show the identification of a capped mRNA oligonucleotide. The complex nature of the sample can lead to some misidentifications. A further check is to look for salt adducts (Na+ and K+).



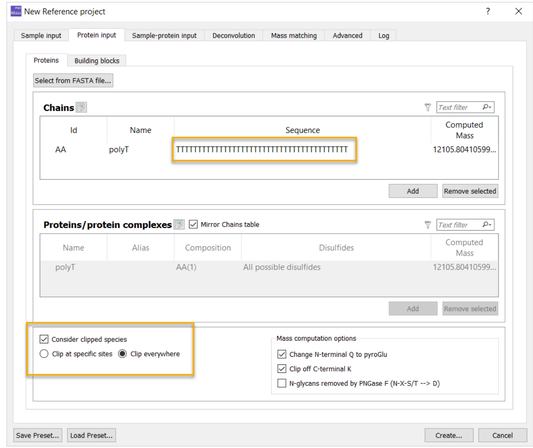


Zoom in on z = 11

Clipping

Intact Mass can also automatically look for clipped species. The software can be directed to clip at specific (user-defined) sites or everywhere.

As an example, when analyzing the Waters MassPREP OST standard, rather than specifying each of the major and minor species individually, an alternative is to specify a single poly-T sequence and automatically look for clips of that single poly-T sequence



Conclusions and Acknowledgments

In summary:

- Intact Mass can be used for the analysis of oligonucleotides in addition to proteins
- Mass assignments can be further confirmed through adduct peaks
- Intact Mass can automatically look for species clipped either at specific motifs or everywhere.

References
Beverly, M., Dell, A., Parmar, P. et al. Anal Bioanal Chem (2016) 408: 5021. https://doi.org/10.1007/s00216-016-9605-x