IMSC 2022, POSTER #115

SEQUENCE VERIFICATION AND SIDE PRODUCT IDENTIFICATION OF SYNTHETIC RNA OLIGONUCLEOTIDES BY LC-ESI-PASEF AND OLIGOQUEST SOFTWARE

[1]Stuart Pengelley, [2]Julia Schneider, [1]Eckhard Belau, [1]Christian Albers, [1]Alexander Bunkowski, [3]Dan Fabris, [1]Peter Sanders, [2]Stephan Seiffert, [2]Ingo Röhl, [1]Detlev Suckau; [1]Bruker Daltonics, Bremen, Germany, [2]Axolabs GmbH, Kulmbach, Germany , [3]University of Connecticut, Department of Chemistry, USA

Introduction

Oligonucleotide characterization by mass spectrometry has gained significant interest recently with the increased use of DNA and RNA as research reagents as well as pharmaceutical drug molecules. Typical DNA primer molecules are in the range of 20mers while single stranded guide RNA (sgRNA) involves the analysis of 100mers at 33 kDa molecular weight.

We developed the dedicated OligoQuest software workflow for the characterization of undigested oligonucleotides in the range 10-100mers using RP-UPLC-UV-ESI-MS/MS which provides the following:

- Sequence confirmation by automatic assignment of MS/MS fragment ions
- LC-UV Quantitation of side products
- MS/MS based sequence proposition of side products

Methods

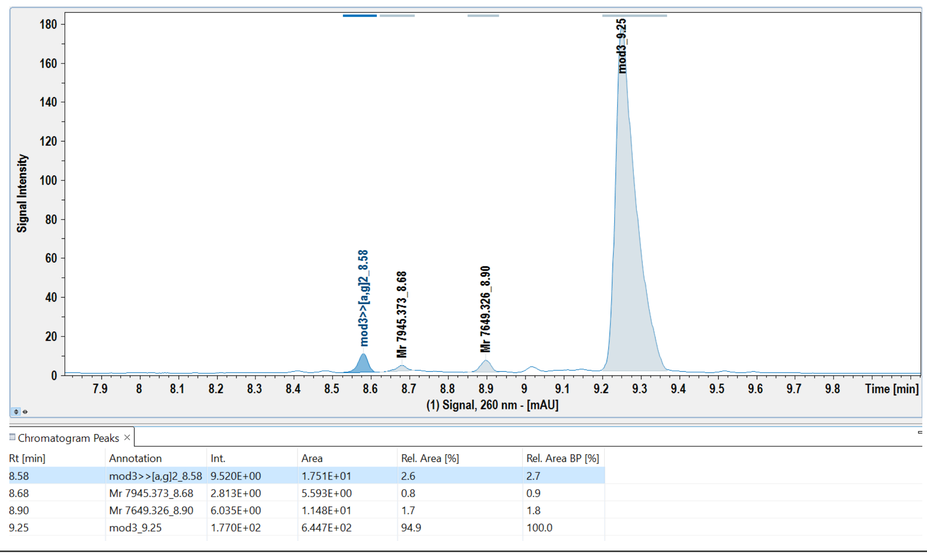
Eight 2´-permethylated RNA 24mers (Axolabs) were UPLC separated (Elute with UV detector at 260 nm) and mass spectra were acquired on a timsTOF Pro 2 with VIP-HESI ion source (all Bruker).

The OligoQuest workflow was implemented in Bruker´s BioPharma Compass software.

In the workflow, multiple MS/MS spectra are accumulated per charge state and the monoisotopic MS/MS peaklist is calculated using the SNAP algorithm.

OligoQuest matches a monoisotopic fragment ion list against the theoretical fragment ions calculated from oligonucleotide sequences - including multiple modifications.

Chromatographic UV peaks are quantified and product purity as well as side product content is determined using both the LC-peak area as well as their MS-peak composition.



Chromatogram Peaks

Rt [min]	Annotation	Int.	Area	Rel. Area [%]	Rel. Area BP [%]
8.58	mod3>>[a,g]2_8.58	9.520E+00	1.751E+01	2.6	2.7
8.68	Mr 7945.373_8.68	2.813E+00	5.593E+00	0.8	0.9
8.90	Mr 7649.326_8.90	6.035E+00	1.148E+01	1.7	1.8
9.25	mod3_9.25	1.770E+02	6.447E+02	94.9	100.0

Fig. 3 Top: Chromatogram of the mod3 analysis. Side product peak at 8.58 min is highlighted, which is suggested to contain a nucleotide exchange variant of a-to-g. Bottom: Chromatogram peaks table which indicates a mod3 content of 94.9% based on peak area. The sequence variant peak at 8.58 min makes up for 2.6 %.

Row	Confirmed	Annotation	Mr Ref	Mr Sample	Δ Mr [ppm]	Rel. Int. [%]	Rt [min]	Intensity Coverage [%]	Sequence Coverage [%]	MSMS Score
1	☑	mod3	7969.3865	7969.3742	-1.54	88.9	9.25	61.0	100.0	61.04
2	☑	mod3>>[a,g]16	7985.3814	7985.3737	-0.97	1.4	8.58	63.7	45.8	29.18
3	☑	mod3>>[a,g]2	7985.3814	7985.3737	-0.97	1.4	8.58	42.3	20.8	8.81

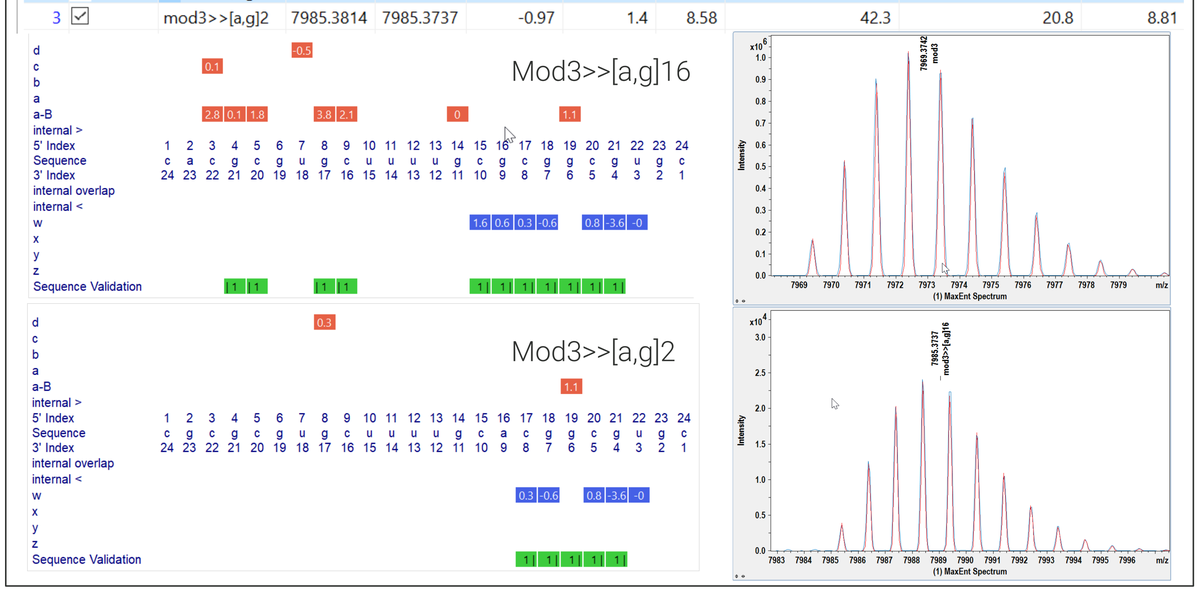


Fig.4. **Top**: Quantitative analysis combining LC and MS data yields 88.9% purity for the target oligo mod3. The [a,g] substitution variant is a 1.4% impurity.
Variant analysis: the MS/MS score for [a,g]16 and [a,g]2 with 29.2 vs. 8.8 indicate the presence of [a,g]16. **Left**: [a,g]16 is clearly the best match based on the fragment ion coverage of 5´-fragments 3-9 and 3´-fragments 22-15. **Right**: confirmation of accurate mass and isotope pattern: blue experimental and red theoretical isotope pattern of mod3 (Top) and mod3 [a,g]16 (bottom).

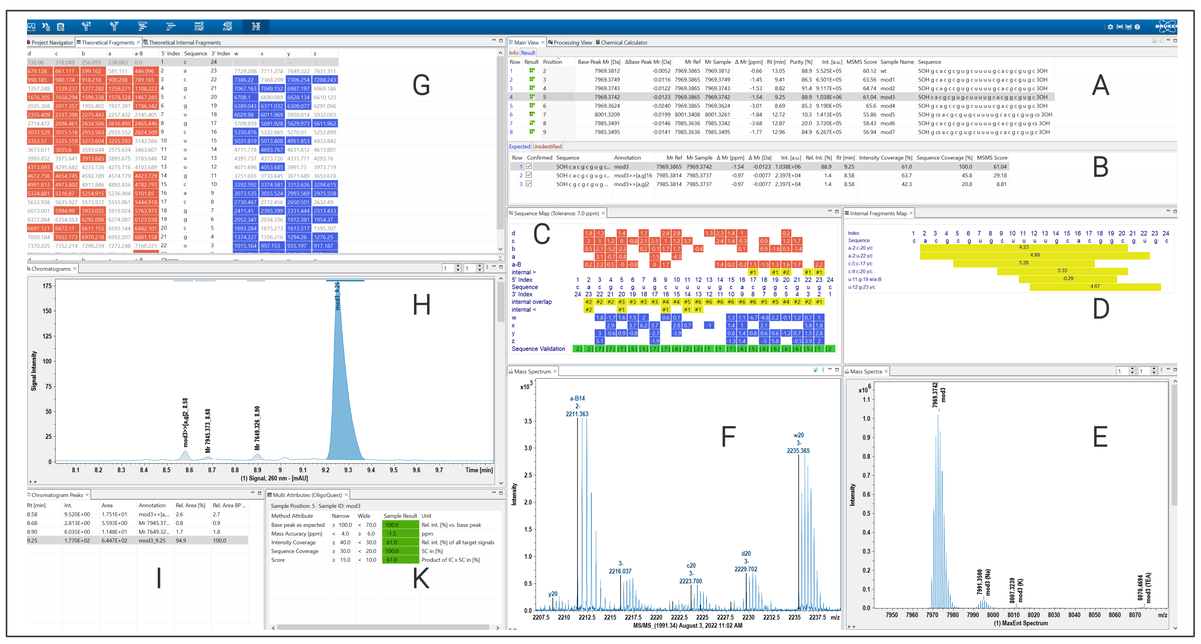


Fig. 1 User interface of the BioPharma Compass result perspective with the analysis overview of 8 oligos of varying sequences:
A Result overview with variant mod3 being selected; B mod3 analysis details incl. sequence suggestions for side products; C Sequence Map with terminal fragment ion matches shown as colored bricks with ppm errors; D Internal fragment ion matches; E intact mass spectrum with annotated adducts; F MS/MS spectrum (detail) with annotated fragment ions; G Theoretical fragment ions with matches highlighted; H UV (260 nm) chromatogram; I Chromatogram peaks quantitation; K Legend for the reporting attribute fields shown in A.

Expected | Unidentified

Row	Confirmed	Annotation	Mr Ref	Mr Sample	Δ Mr [ppm]	Rel. Int. [%]	Rt [min]	Intensity Coverage [%]	Sequence Coverage [%]	MSMS Score
1	☑	mod3	7969.3865	7969.3742	-1.54	88.9	9.25	61.0	100.0	61.04
2	☑	mod3>>[a,g]16	7985.3814	7985.3737	-0.97	1.4	8.58	63.7	45.8	29.18
3	☑	mod3>>[a,g]2	7985.3814	7985.3737	-0.97	1.4	8.58	42.3	20.8	8.81

Sequence Map (Tolerance: 7.0 ppm)

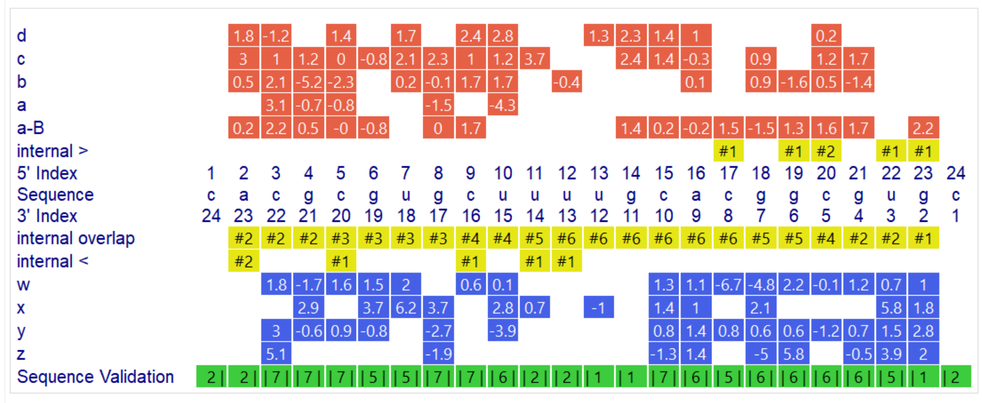


Fig. 2 Top: The target mod3 sequence analysis result incl. MS and MS/MS match quality assessment (ppm Mr error and MSMS Score)
Bottom: Sequence map of matching fragment ions. 100% sequence coverage is achieved highlighted by the uninterrupted row of green bricks

Results

- The sequencing result validation was enabled in a single user interface allowing to review quantitative results as well as the validation of MS/MS spectra which provides a multiple sample analyses overview (*Fig 1*)
- Terminal as well as internal fragment ions were used to validate the mod3 oligo yielding 100% sequence coverage (*Fig. 2*)
- The UV chromatogram was used to determine the purity of the target oligo mod3 with 95% and a side product contribution of 2.6% (*Fig. 3*)
- The side product was identified based on matching fragment ions as mod3-a16g: a nucleotide exchange product, using specific fragment ions and a convincing match of its intact molecular ion isotope pattern providing for monoisotopic mass determination using the SNAP algorithm. The pattern agreed well with the theoretical isotope pattern (*Fig. 4*)

Summary

- BioPharma Compass 2023 with its OligoQuest workflow allows to analyze oligonucleotides using LC-UV-ESI-MS and MS/MS
- Negative mode high isotope fidelity MS and MS/MS spectra are accurately mass analyzed (< 5 ppm) using the SNAP algorithm
- MS/MS spectra can be validated automatically using generated scores and by supporting the operator to directly investigate MS/MS spectra in a sequence driven fashion
- Even low abundant side products can be qualified based on MS and MS/MS data.
- Quantitation is supported using LC-UV data traces alone or by adding the mass composition of LC peaks to account for co-eluting peaks that cannot be considered by LC alone.

Oligonucleotide Biologics