Software enabled oligonucleotide mapping analysis of Erythropoietin mRNA

Alexander Bunkowski[1], Waltraud Evers[1], Eckhard Belau[1], Thomas Meid[1], Lars Vorwerg[1], Stuart Pengelley[1], Yun Yang[2], Guillaume Tremintin[2], Detlev Suckau[1]

[1]Bruker Daltonik GmbH & Co. KG, 28359 Bremen, Germany

[2]Bruker Daltonics Inc., San Jose, CA, USA

Introduction to mRNA digest analysis

- With mRNA vaccines emerging as a novel class of biotherapeutic agents the need for their characterization increased.
- As for protein biotherapeutics, the key method for their LC-MS/MS characterization is the digestion of mRNA into smaller oligonucleotides.
- In this work, we developed the digest conditions, analytical LC-separation, and optimized CID experiments for the characterization of the 615nt EPO mRNA.
- New software was developed to analyze these oligo maps and to eliminate the otherwise overwhelming efforts of manual data interpretation of such datasets.
- The OligoQuest Digest workflow was implemented in the BioPharma Compass for ease-of-use analysis of these complex datasets and to simplify data interpretation.

Methods

Twenty µg of 615mer EPO mRNA (RiboPro, Oss, NL) was digested by RNase T1 (Worthington) in 15 min at 37 °C and 2 µg of the resulting oligonucleotides were subjected to analytical LC and a timsTOF Pro 2 (Bruker) in autoMSMS mode of operation; internal mass calibration in negative ion mode was used.

A 1.5 sec autoMSMS cycle was used in which precursors were fragmented at 2 Hz with CID by applying 40 eV CE at 500 m/z, increasing to 70 eV at 2000 m/z in a linear fashion.

The resulting data were analyzed with prototypic OligoQuest Digest software as part of BioPharma Compass. The specificity of RNase T1 was defined in OligoQuest to cleave at the 3´-end leaving a 3´-phosphate. After chromatographic peak picking, deconvoluted MS spectra were calculated and the MS/MS spectra obtained. Sequences were matched to the MS/MS spectra with an accuracy of typically better than 3 ppm. Mass tolerance was set to 7 ppm.

Results

The sequence coverage map of EPO mRNA was generated and a sequence coverage of 88% was obtained from a single digest.

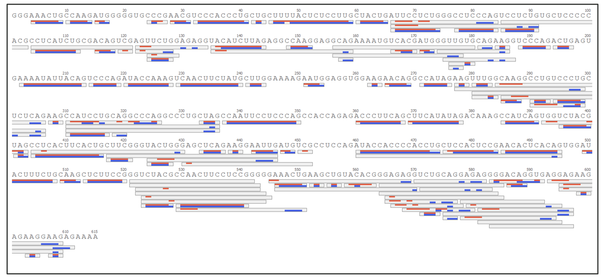


Fig. 1 Sequence coverage map of EPO mRNA. The identified oligonucleotides are shown as grey bars and the 5´- and 3´-fragments in red and blue, respectively

Individual oligos were selected in the sequence coverage map to verify their sequence (Fig. 2) on the level of the MS/MS spectrum and on the level of individual fragment ions (Fig. 3).

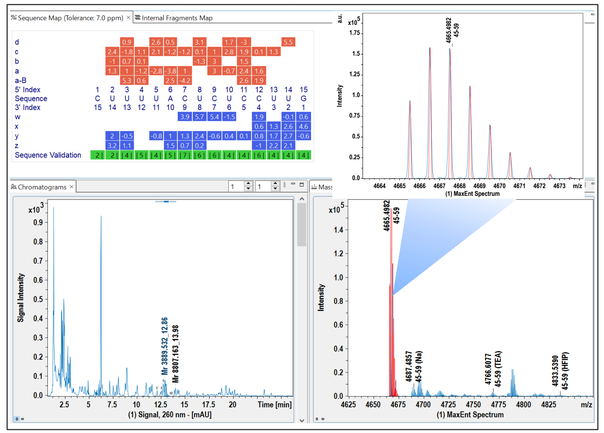


Fig. 2 Fragment ion coverage map and MS spectrum of EPO mRNA 45-59 (top), the chromatogram and the MS spectrum with adducts (bottom). Red and blue tiles show the 5´ and 3´ confirmed fragments, respectively. The green Sequence Validation tiles stringently indicate how many fragments confirm individual nucleotide positions.

The MS/MS spectrum obtained at 12.86 min (Fig. 3) as an example of the obtained data quality in such analyses.

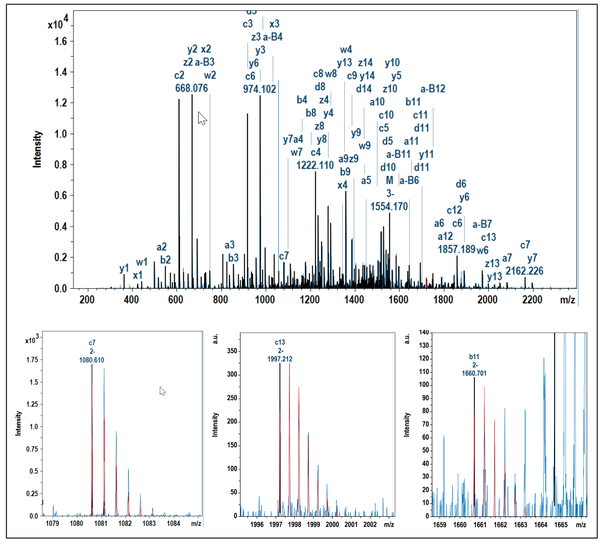


Fig. 3 MS//MS spectrum of EPO mRNA 45-59 (top). Selected, particular low abundant fragment ions were overlayed with their theoretical isotope pattern from the nucleotide fragment annotation (red).

The MS/MS spectra were analyzed and individual fragments were verified across a broad dynamic range. In the MS/MS spectrum of EPO mRNA 45-59 (Fig. 3), fragments are overlayed with their theoretical isotopic pattern to verify their proper identification. This can be applied by the user to any isotopic distribution and can be particularly useful for low abundant peaks picked against a complex background.

For peak picking, the SNAP algorithm was used for MS as well as MS/MS peaklist calculation as part of the OligoQuest workflow. SNAP matches an experimental isotope pattern with an expected pattern based on the molecular composition of the analyte, in this case the oligonucleotide model was employed.

Summary

The OligoQuest Digest workflow in the BioPharma Compass software enables to rapidly obtain a comprehensive overview of oligonucleotide digests. In addition, the software allows to go into full analytical detail and confirm every fragment using highly accurate MS/MS information obtained on Bruker QTOFs.

Short sequences of 2-3mers occurred several times in the sequence (e.g., AG 17x and AAG 9x), these were not included in calculating the sequence map.

The workflow allows the user to:

- Process the LC-MSMS datasets generating monoisotopic peaklists
- Overview the fragment coverage of the digested mRNA
- Select individual digest oligonucleotides to review their fragment ion maps
- Review and explore the MS/MS profile spectra even on the level of low abundant peaks matching the sequence

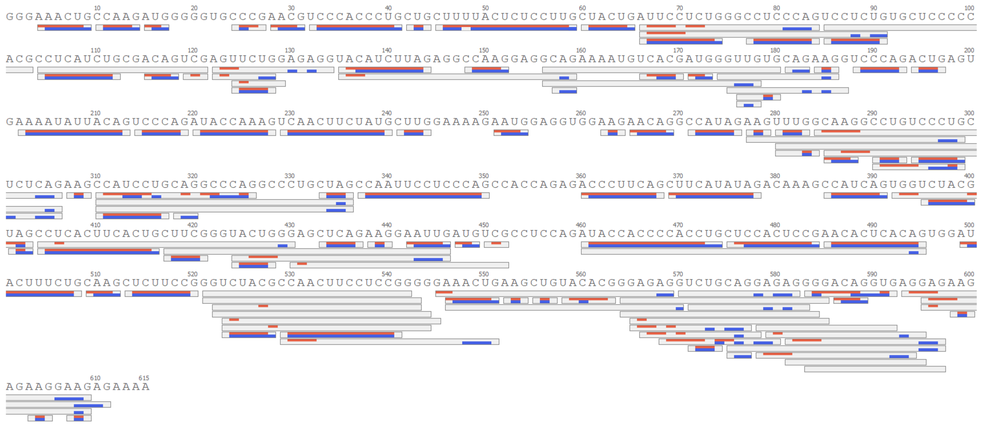
The data quality of the profile data generated using the timsTOF Pro 2 allowed to reliably pick monoisotopic masses in complex spectra, generating high confidence in the reconstructed sequence of oligonucleotide digests.

The OligoQuest Digest workflow provides a comprehensive set of processing and analysis tools that facilitate the analysis of mRNA digests greatly and increases the reliability of the generated results.

Conclusion

- A workflow was developed for the analysis of mRNA digests, it comprised of a 15 min digest step and 25 min LC separation.
- The developed workflow, OligoQuest Digest, allowed to process these data and visualize it for further analysis and detailed data validation.
- In the example, a single analysis provided 88% sequence coverage of the 615nt EPO mRNA.
- High mass accuracy and true isotopic pattern quality allowed to reliably pick even low abundant peaks in complex MS/MS spectra.

mRNA drugs


GGGAAACUGCCAAGAUGGGGGUGCCCGAACGUCCCACCCUGCUGCUUUUACUCUCCUUGCUACUGAUUCCUCUGGGCCUCCCAGUCCUCUGUGCUCCCCC
ACGCCUCAUCUGCGACAGUCGAGUUCUGGAGAGGUACAUCUUAGAGGCCAAGGAGGCAGAAAAUGUCACGAUGGGUUGUGCAGAAGGUCCCAGACUGAGU
GAAAAUAUUACAGUCCCAGAUACCAAAGUCAACUUCUAUGCUUGGAAAAGAAUGGAGGUGGAAGAACAGGCCAUAGAAGUUUGGCAAGGCCUGUCCCUGC
UCUCAGAAGCCAUCCUGCAGGCCCAGGCCCUGCUAGCCAAUUCCUCCCAGCCACCAGAGACCCUUCAGCUUCAUAUAGACAAAGCCAUCAGUGGUCUACG
UAGCCUCACUUCACUGCUUCGGGUACUGGGAGCUCAGAAGGAAUUGAUGUCGCCUCCAGAUACCACCCCACCUGCUCCACUCCGAACACUCACAGUGGAU
ACUUUCUGCAAGCUCUUCCGGGUCUACGCCAACUUCCUCCGGGGGAAACUGAAGCUGUACACGGGAGAGGUCUGCAGGAGAGGGGACAGGUGAGGAGAAG
AGAAGGAAGAGAAAA