Characterization of mRNA Therapeutics Using a Novel LC-MS Based Workflow

Angela Criscuolo[3], Tabiwang Arrey[3], Andrew Williamson[2], Ken Cook[2], Keeley Murphy[1], Jennifer Sutton[1], Alexander Harder[3], Jim Shofstahl[1], Mark Sanders[1]
[1]Thermo Fisher Scientific, San Jose, CA, USA; [2]Thermo Fisher Scientific, Hemel Hempstead, United Kingdom; [3]Thermo Fisher Scientific GmbH, Bremen, Germany

ABSTRACT

Purpose: To develop an RNase digestion strategy to support an LCMS based mRNA mapping workflow through the accurate control of sample digestion time, thereby facilitating the generation of unique digestion fragments and subsequent sequence confirmation.

Methods: mRNA sample was enzymatically digested using RNase T1, immobilized on magnetic beads. Digestion times ranging from 5 to 30 minutes were evaluated. The digested samples were analyzed using a Thermo Scientific™ Orbitrap Exploris™ 240 mass spectrometer coupled with a Thermo Scientific™ Vanquish™ UHPLC system. MS and MS/MS experimental data was analyzed using Thermo Scientific™ Biopharma Finder™ software.

Results: Partial digestion of the mRNA was evaluated with the optimal digestion time observed at 15 minutes. The identified digestion fragment sequences ranged from 5 to 43 bases in length for the mRNA digest. The chromatic results demonstrated an elution order corresponding to digestion fragment length. MSMS matching and annotation was used to identify digestion fragments within the digested mRNA sample. Confidence score values were automatically generated by the data processing software based on upon MSMS matching results. The application of a confidence score filter of greater than 85% resulted in an overall mRNA sequence mapping coverage of 91% for the digested sample.

INTRODUCTION

Messenger RNA (mRNA) based drugs represent a new prospective class of products. In fact, the first vaccine to be approved for emergency use against COVID-19 was based on mRNA technology. The response to the Coronavirus pandemic has demonstrated that mRNA therapies are a growing field and have the potential to protect against other diseases, e.g., Zika virus and the treatment of rare genetic disorders. For this reason, there is currently significant demand for the development of new and improved analytical methods for the characterization of mRNA therapeutics. The aim of this work was to develop a mass spectrometry-based workflow covering sample preparation, LC-MS analysis and data processing in order to characterize mRNA therapeutics.

METHODS

Sample Preparation

An mRNA sample consisting of 3500 nucleotide bases was obtained and sample digestion performed using immobilized RNase T1 on magnetic beads. mRNA samples were diluted to a concentration of 0.5mg/mL in digestion buffer and 2.5 µL of the immobilized T1 beads were added. The sample was incubated at 37°C with continuous agitation and evaluated across an incubation time range of 5 to 20 minutes. The reaction was stopped at the designated time point using a magnet to completely remove the magnetic bead resin, followed by centrifugation. An equal amount of 1% formic acid was added to the remaining solution.

HPLC Conditions

44 uL of the resulting sample solution was injection onto a Thermo Scientific™ DNAPac RP™ column and separated chromatographically using a Vanquish UHPLC system. Mobile phase A was prepared with water containing 0.2% TEA and 1% HFIP. Mobile phase B was prepared with methanol containing 0.2% TEA and 1% HFIP. The DNAPac RP column was maintained at 50°C, with a flow rate of 300 µL/min.

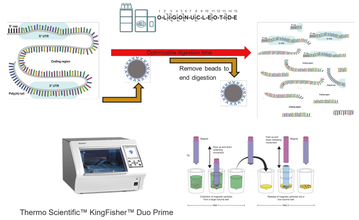
Mass Spec Conditions

All the data was collected using an Orbitrap Exploris 240 mass spectrometer. Sample analysis was performed using data dependent Full Scan and data dependent MS/MS acquisition (Table 1)

Table 1. MS source, global, full scan, and data dependent settings.

MS Source Settings	Value	Scan Settings	Value
Spray voltage (V)	-2500	**Full Scan**	
Vaporizer temp.	350	Scan range (m/z)	450-4000
Capillary temp.	300	Resolution	120000
Sheath gas	50	AGC target value (%)	200
Aux gas	15	Max inject time (ms)	Auto
Global Scan Settings	**Value**	**dd-MS/MS**	
Polarity	Negative	Scan range (m/z)	150-2000
Application mode	Peptide	Resolution	30000
Pressure mode	Standard	Isolation window	3
RF lens (%)	38	AGC target value (%)	50
Expected peak width	Standard	Max inject time (ms)	Auto
Default charge state	5	HCD collision energy (V)	Stepped 25,28,31

Figure 1. mRNA sample digestion was performed using RNase T1 immobilized on magnetic beads to control digestion time and subsequent digestion fragment length.



RESULTS

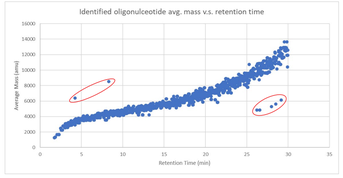
Sample digestion

Digestion optimization was performed on an mRNA sample containing 3500 bases and evaluated to the determine optimal digestion time required to produce digestion fragments ranging between 15 and 40 bases in length. mRNA sample digestion was evaluated at times of 5, 10, 15, 30 minutes. The aim of the partial digestion strategy is to induce missed cleavages throughout the mRNA sequence, thereby producing digestion fragment lengths whose sequence location can be uniquely determined within the larger mRNA sequence, thereby producing the greatest overall sequence mapping coverage. The optimal digestion time was produced with magnetic bead removal at a digestion time of 15 minutes. Magnetic bead removal can be performed manually using a handheld magnet or through the implementation of an automated purification system (Figure 1).

Chromatographic separation

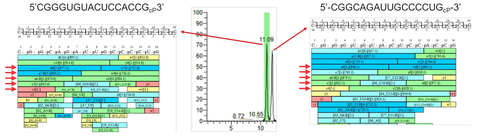
Chromatographic separation of the mRNA digestion fragments was achieved using a 40-minute gradient ranging from 5% to 20% mobile phase B. Oligonucleotide elution order from the column was observed to be dependent on molecular size, with the larger digestion fragments demonstrating increased chromatographic retention. Molecular weight versus chromatographic elution order for each identified digestion fragment was plotted to verify the elution order trend (Figure 2). The elution order plot was evaluated to identify trend outliers and add additional confidence to digestion fragment identification and overall data review.

Figure 2. Oligonucleotide molecular weight versus chromatographic elution time for each identified digestion fragment. Potential outliers are circled in red.



The column and method gradient provided efficient chromatographic separation, demonstrating the resolution of isobaric and isomeric compounds. Chromatographic separation of isomeric oligonucleotides is a critical tool for the accurate identification of mRNA digestion fragments, enabling the collection of diagnostic MSMS fragment ions without interference from overlapping isomers. Chromatographically separated isomers were automatically identified by the data analysis software and confirmed using MSMS fragment identification. Diagnostic MSMS fragments used to confirm oligonucleotide sequence were reviewed using the fragment coverage map tool provided in the data analysis software (Figure 3).

Figure 3. Chromatographically resolved mRNA digestion fragment isomers and the respective fragment coverage map. Diagnostic fragmentation ion identifications are indicated by red arrows for a selection of 5' series fragments.



MSMS fragment identification

MSMS fragment identification is a critical component of LCMS mRNA mapping. In addition to identification by full scan analysis and the mass deconvolution of multiple charge states; fragmentation identification provides the confirmation of sequence order as well as information pertaining to potential modifications and the modification site within the sequence. MSMS fragmentation identification and annotation was automatically generated, and color coded by the data analysis software. In addition to fragment annotation, the observed fragmentation spectrum is automatically compared to a predicted fragmentation model generated for each identified sequence (Figure 4). This comparison is subsequently used to generate a confidence score for each identification. Confidence score values greater than 80% indicate a high confidence sequence identification. MSMS fragmentation identification was also evaluated using the fragment coverage map tool, which provides a visual representation of all assigned MSMS fragmentation and is color coded by intensity for each digestion fragment. Fragment coverage map views were evaluated for each identified charge state and the display customized based on charge state and fragment type. Base peak and selected ion chromatograms for the identified sequence fragments were also evaluated in addition to the customized fragment coverage maps to assist with data review (Figure 5).

Figure 4. Color coded annotated MSMS spectrum is automatically generated for each oligonucleotide identification. Predicted fragmentation for the identified sequence is displayed (top) and compared to experimental results (bottom)

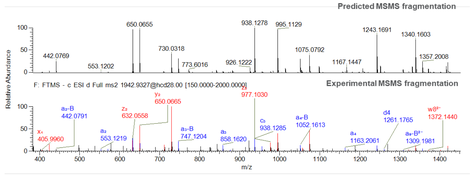
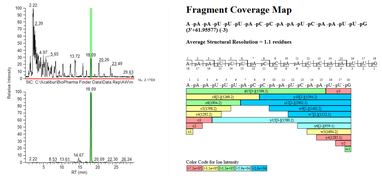


Figure 5. Base peak and selected ion chromatograms for an 18 base digestion fragment (left). Fragment coverage map the 18 base digestion fragment at charge state -3, filtered to remove internal fragmentation annotation in the fragment coverage map (right).



mRNA sequence coverage

Sequence coverage for mRNA mapping by LCMS is greatly improved by increasing the length of the digestion fragments produced. While shorter digestion fragments can be confidently identified, shorter sequence fragments are more likely to be found in multiple sections of the mRNA sequence, these are defined as non-unique identifications. By restricting the digestion time of the mRNA sample and inducing missed cleavages, longer digestion fragments are produced, which can then be assigned with a unique region of the mRNA sequence. Using the results table within the data analysis software, digestion fragment identifications were filtered and sorted to display each identification based on sequence length. Identified sequence fragments were then evaluated and confirmed across multiple charge states and the range of missed cleavages for the sample digestion determined (Figures 6 & 7). The 15-minute mRNA digest produced missed cleavages ranging from 1 to 11 across the identified digestion fragments.

Figure 6. Identification of a 30 base digestion fragment with missed cleavages circled in red. (top) Charges states evaluated in the digestion fragment identification, with confidence score, m/z value, and MS peak area. (bottom)

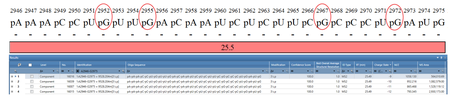
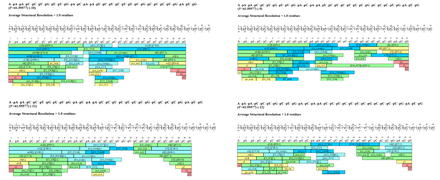
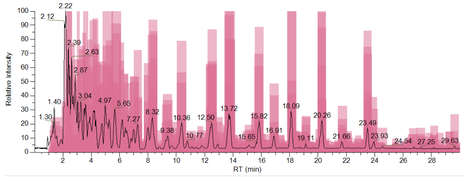


Figure 7. Fragment coverage map for charge states -9, -10, -11, and -12 for 30 base length digestion fragment.



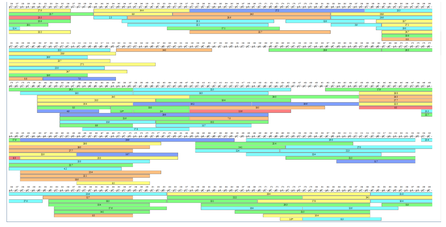
Overall mRNA sequence coverage was evaluated using the oligonucleotide sequence coverage map and chromatogram component identification tools with the data analysis software. The chromatogram component identification tool was used to display all identified digestion fragments in red and pink across the chromatogram (Figure 8) and evaluated to identify any gaps in identification and utilized during LCMS method optimization.

Figure 8. Chromatogram component identification view of the digested mRNA sample. Identified digestion fragments highlighted in pink.



The oligonucleotide sequence coverage map was used to evaluate the overall mRNA sequence mapping coverage. Mapping coverage is determined based on the sum of identified digestion fragments and the respective sequence location assignment in the mRNA sequence (Figure 9). The sequence coverage map provides a percent value for sequence coverage and was filtered to only include results meeting specified criteria. Overall sequence map coverage was determined using digestion fragments assigned with MSMS fragment identification only, with confidence scores greater than 85%, and with identification fragments that only match expected cleavage sites (i.e., no non-specific digestion fragments). These acceptance criteria demonstrated an overall sequence coverage of 89% for unique only identifications and 91% with inclusion of non-unique identifications.

Figure 9. Example selection of the sequence coverage map for the digested mRNA sample. Identified sequence fragments between bases 225 and 600 with a confidence score greater than 85% are displayed.



CONCLUSIONS

- A sample digestion procedure was successfully developed to control mRNA digestion using immobilized RNase on magnetic beads.
- Optimization of the sample digestion time provided an increase in both the overall length of the digestion fragments produced and the percentage of unique sequence identifications.
- The identified digestion fragment sequences ranged from 5 to 43 bases in length in the mRNA digest.
- The chromatographic separation method provided resolution for the identified digestion fragment isomers.
- MSMS fragmentation identification provides increased confidence for sequence identification and confirmation.
- Overall sequence coverage was determined to be 89% for unique only identifications and 91% with inclusion of non-unique identifications.

ACKNOWLEDGEMENTS

The authors would like to thank and acknowledge Dr. Zhongqi Zhang for his work on oligonucleotide identification, fragmentation prediction, and mRNA mapping in support of this project.

TRADEMARKS/LICENSING

PO86111 EN0921S