# ANALYSIS OF mRNA CRITICAL QUALITY ATTRIBUTES (CQAs) USING THE BIOACCORD UPLC-TOF MS SYSTEM AND INTACT Mass SOFTWARE

*Rebecca D'Esposito[1], Catalin Doneanu[1], Heidi Gastall[2], Scott J. Berger[1], and Ying Qing Yu[1]*

*[1]Waters Corporation, Milford MA, USA; [2]Waters Corporation, Wilmslow, UK*

## INTRODUCTION

Recent development and approval of the two COVID mRNA based vaccines has brought RNA-based therapeutics to the forefront of the biopharma industry. As such, development of analytical methods for monitoring of the CQAs of RNA-based therapeutics has become a high priority for ensuring proper control of the manufacturing process.

This study illustrated three CQAs (Critical Quality Attributes) analytical workflow for mRNA using a bench top UPLC-TOF MS instrument and customized data processing software.[1-3]

These CQAs are:

- *5' capping[2]*
- *Poly(A) Tail heterogeneity*
- *Sequence mapping*

## METHODS

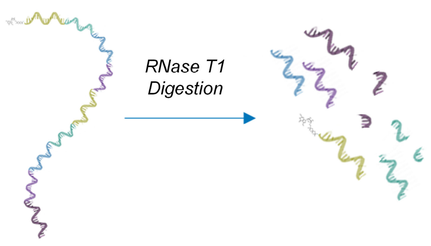
**mRNA CQAs**


**Instrument**

Figure 1: BioAccord™ UPLC-TOF MS System with Max Peak™ High Performance Surfaces (HPS).

UPLC Column: ACQUITY Premier OST Column 1.7 µm, 130 Å, 2.1 x 150 mm

**Sample Preparation**




***Sample used for RNA Mapping:***

- sgRNA (100 mer, IDT) - contains six PTMs, such as 2'O-mehtylguanosine, 2'O-methyluridine and 2'O-methyladenosine and each modification has a phosporothioate group.
- Firefly luciferase (Fluc) mRNA (Amp Tec)

***Sample used for Poly(A) Tail Heterogeneity***

- Firefly luciferase (Fluc) mRNA (Amp Tec)

***Sample used for 5' Capping demonstration***

5'-GUAGAACUUCGUCGAGUACGCUCAA (New England Biolab)

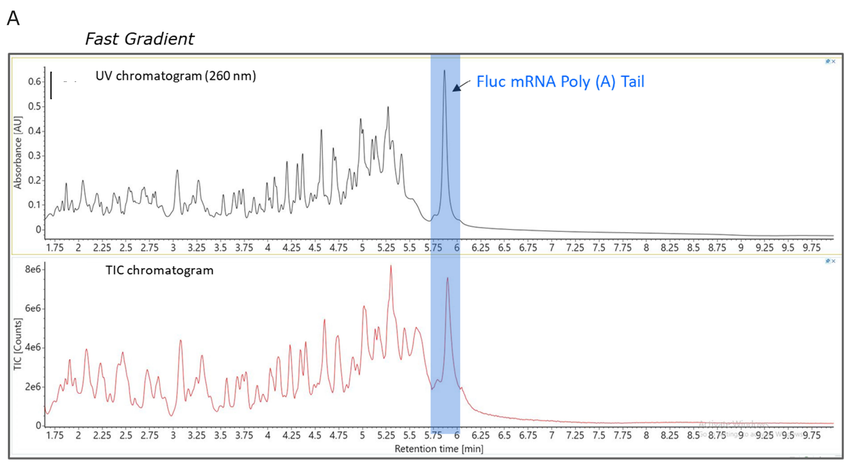
**Software**


INTACT Mass App (software) was the main software used for 5' capping analysis, sequence mapping and Poly(A) Tail hetero-geneity profiling.

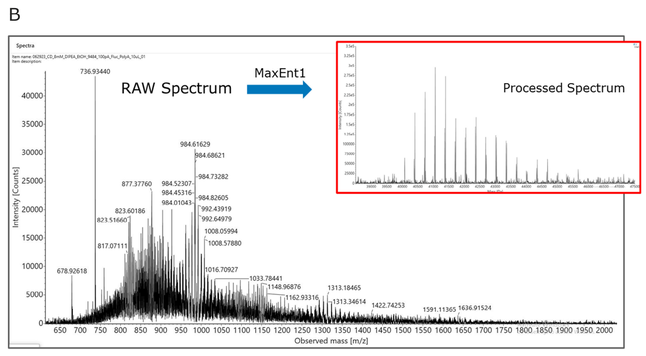
## RESULTS

### *Poly A Tail*

A




B




C

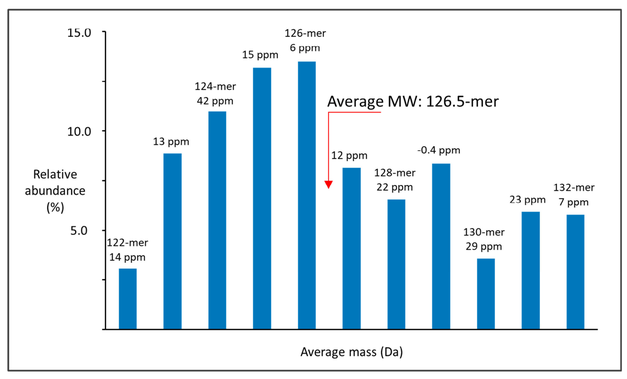
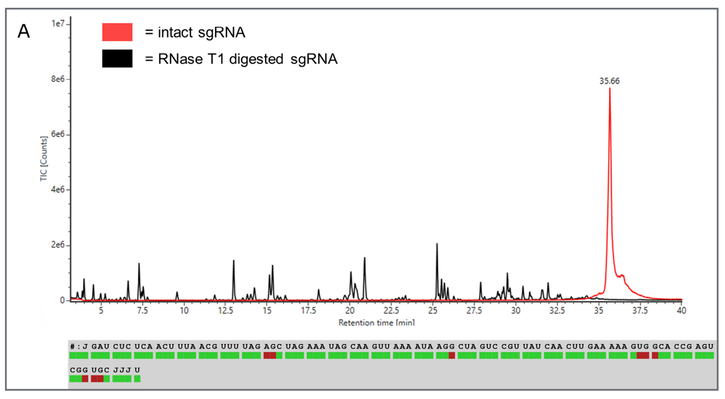



Figure 3. A) A fast IPRP UPLC-MS chromatogram is shown here (UV and TIC). Poly A tail components were focused to a sharp peak using a fast chromatographic separation,[3] the goal was to improve the MS sensitivity for the highly heterogeneous Poly A components from the RNase T1 digestion. B) Summed MS scans showed the level of heterogeneity of the poly A tail. MaxEnt 1 charge deconvolution is used to process the raw MS data to obtain MW series of the Poly A tail. C) Weighted average Poly A tail length was calculated, showing at 126.5 mer long.

### *Sequence Confirmation*







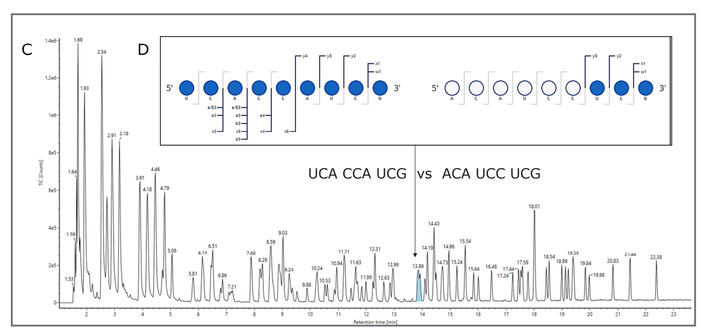



Figure 4. sgRNA and mRNA were digested by RNase T1 and analyzed using the BioAccord LC-MS system. Data analysis is performed using INTACT Mass App with the assistance of mRNA Cleaver App (to generate theoretical mass of the digested fragments) and Sequence Viewer (for coverage% calculation and review of the data). A) Example data were shown from a ~100 mer sgRNA. B) LC-MS chromatogram of RNase T1 digested Firefly luciferase mRNA(~ 4000 mer) and the sequence coverage in a "MAP viewer". We included the ambiguous sequence in the calculation. Data generated from elevated energy fragmentation can be processed in another software, CONFIRM Sequence, to differentiate ambiguous sequence. For example C-D) A better match between two candidate oligo sequences was highlighted based on the fragment ion coverage. These two candidates are *ambiguous sequences* since they have the same MW (isobaric masses), but contain scrambled sequences.

### *5' capping*

A

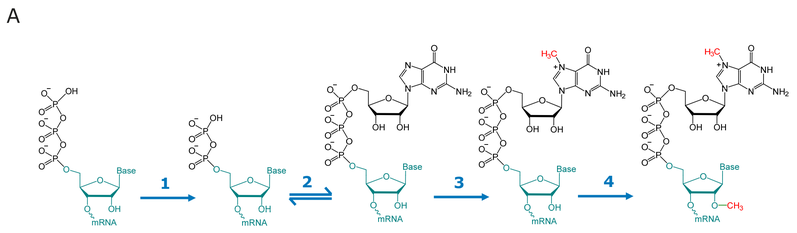
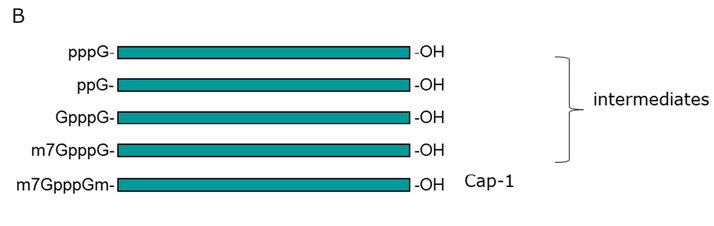



Figure from New England BioLabs Inc.

B

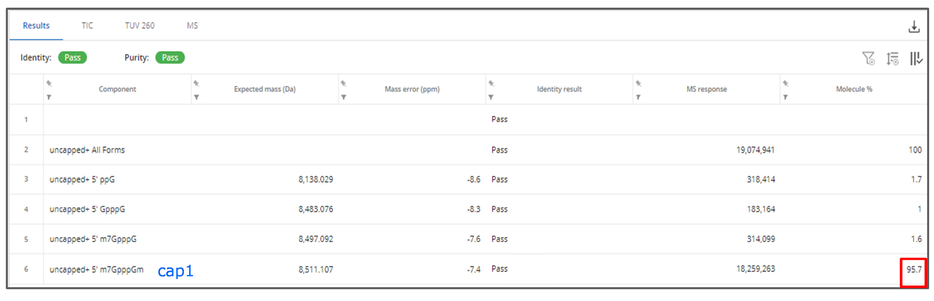



$$\text{Capping efficiency} = \frac{\text{Cap1 MS response}}{\text{(Cap1 + all intermediates) MS response}} \times 100\%$$

Figure 2. (A) Illustration of the Cap-1 vs. Precursor impurities; (B) Synthetically made Cap-1 (25 mer long) was mixed with synthetically made intermediate byproducts in 999:1, 99:1 and 9:1 ratio before LC-MS analysis to quantify the Cap-1 product rel% using MS signals. C) Result table from Cap-1 mixed in at 100:1 ratio with other impurites. LC-MS analysis and data processing using the INTACT Mass App generated the rel% calculation of the Cap-1 species.

C

Cap 1 was mixed with impurity at 99:1 Ratio




| | Component | Expected mass (Da) | Mass error (ppm) | Identity result | MS response | Molecule % |
|---|---|---|---|---|---|---|
| 1 | | | | Pass | | |
| 2 | uncapped+ All Forms | | | Pass | 19,074,941 | 100 |
| 3 | uncapped+ 5' ppG | 8,138.029 | -8.6 | Pass | 318,414 | 1.7 |
| 4 | uncapped+ 5' GpppG | 8,483.076 | -8.3 | Pass | 183,164 | 1 |
| 5 | uncapped+ 5' m7GpppG | 8,497.092 | -7.6 | Pass | 314,099 | 1.6 |
| 6 | uncapped+ 5' m7GpppGm cap1 | 8,511.107 | -7.4 | Pass | 18,259,263 | 95.7 |

## CONCLUSIONS

- Three analytical workflow were developed on the BioAccord UPLC-MS system equipped with the waters_connect informatics platform for three mRNA CQAs analysis. These CQAs are 5' capping efficiency, Poly(A) Tail heterogeneity and sequence mapping analysis.
- Software tools plays a critical role in reducing the time used for data interpretation. The software used in this study is the INTACT Mass App with two newly added capabilities, such as the Enzyme Cleavage Tool and Sequence Mapping Viewer.
- Future work will be focused on improving the sequence coverage by 1) combining MS1 and MS2 results; 2) automation of the entire data processing, with less manual intervention, to reduce the level of ambiguous assignments.

## REFERENCES:

1. Waters Application Note: 720008130 "RNA CQA Analysis using the BioAccord LC-MS System and INTACT Mass waters_connect Application".
2. Waters Application Note: 720007329 " Rapid Analysis of Synthetic mRNA Cap Structure Using Ion-Pairing RPLC with the BioAccord LC-MS System".
3. Waters Application Note: 720007925 "Ion Pairing Reversed Phase LC-MS Analysis of Poly(A) Tail Heterogeneity Using the BioAccord LC-MS System".

***TO DOWNLOAD A COPY OF THIS POSTER, VISIT WWW.WATERS.COM/POSTERS***