Application brief | 003060

Thermo Fisher SCIENTIFIC

Biopharma

Accurate and consistent analysis of poly(A) tails of mRNA therapeutics on a UHPLC-HRAM-MS platform

Authors

Hao Yang[1], Robert Ross[2], Min Du[2]

[1]Thermo Fisher Scientific, San Jose, CA, US

[2]Thermo Fisher Scientific, Lexington, MA, US

Keywords

Polyadenylated (poly(A)) tail analysis, intact mass determination, monoisotopic mass, liquid chromatography high-resolution accurate-mass mass spectrometry, ion-pairing reversed phase liquid chromatography, IPRP-LC, Vanquish Horizon UHPLC system, Orbitrap Exploris 240 mass spectrometer, Orbitrap Exploris MX mass detector, DNAPac reversed phase column, Chromeleon Chromatography Data System (CDS), high-resolution accurate mass, HRAM

Application benefits

- Accurate and consistent measurement of poly(A) tail length and distribution of *in vitro* transcribed (IVT) mRNA
- Automated and compliance-ready analysis and reporting of poly(A) tail lengths for IVT mRNA
- Seamless transfer of the IPRP-LC-HRAM-MS method from one LC-MS system to another via a Thermo Scientific™ Chromeleon™ CDS eWorkflow™ procedure

Goal

- Develop an automated and compliance-ready analytical solution for accurate and consistent analysis of poly(A) tails of IVT mRNA using an IPRP-UHPLC-HRAM-MS method from sample to report
- Execute the developed solution on three different LC-MS systems to demonstrate seamless transfer and data consistency

Introduction

The introduction of messenger RNA (mRNA) therapeutics as effective tools against COVID-19 has accelerated the development of mRNA technologies for combatting infectious disease and cancer.[1-3] An mRNA-based approach offers the potential for faster and more efficient production processes compared to traditional virus-based and protein-based medicines.[3] Typically, mRNAs are produced through *in vitro transcription*

(IVT), and a poly(A) tail is attached to the 3' phosphate of the mRNA transcript to enhance translational efficiency and mRNA stability. This attachment can be done either through encoding the polyadenosine sequence into the template DNA or by post-synthesis addition using a polyadenylase.[4] Either approach will generate some level of sequence variants through non-homologous addition of the adenosine to the 3' end. The ability to accurately measure the tail length of IVT synthesized mRNA is therefore crucial to evaluate its impact on protein expression and stability of the mRNA-based therapeutics. Therefore, accurate and robust analytical methods are needed to characterize and monitor the length and distribution of the poly(A) tails generated through IVT, from development to commercial production. Various analytical methods, such as chromatographic separation, PCR, and next-generation sequencing (NGS) have been reported for measuring the mRNA's tail length. However, these methods either lack the resolution to confidently determine the actual poly(A) tail length or are expensive and require a pool of cDNA fragments and library preparation.

Here, we present a simple LC-MS method for direct and robust measurement of polyadenylated tails in synthetic mRNA transcripts using a high-resolution accurate mass mass spectrometry (HRAM-MS) method developed on the Thermo Scientific™ Vanquish™ Horizon UHPLC system coupled to the Thermo Scientific™ Orbitrap Exploris™ mass spectrometers. To demonstrate reproducibility and consistency of results, a Chromeleon eWorkflow procedure (Figure 1) was executed on three LC-MS systems for the automated analysis of poly(A) tails generated from IVT mRNA digests. Profiling of the poly(A) tails from three individually prepared IVT Cas9 mRNA digest samples was achieved at single nucleotide resolution. The distribution of the tail lengths ranged from 117 to 132, with a median length around 125. The tail length distributions were consistent across all three LC-MS systems, and the measured monoisotopic masses of the poly(A) tails were within 5 ppm of the expected values.

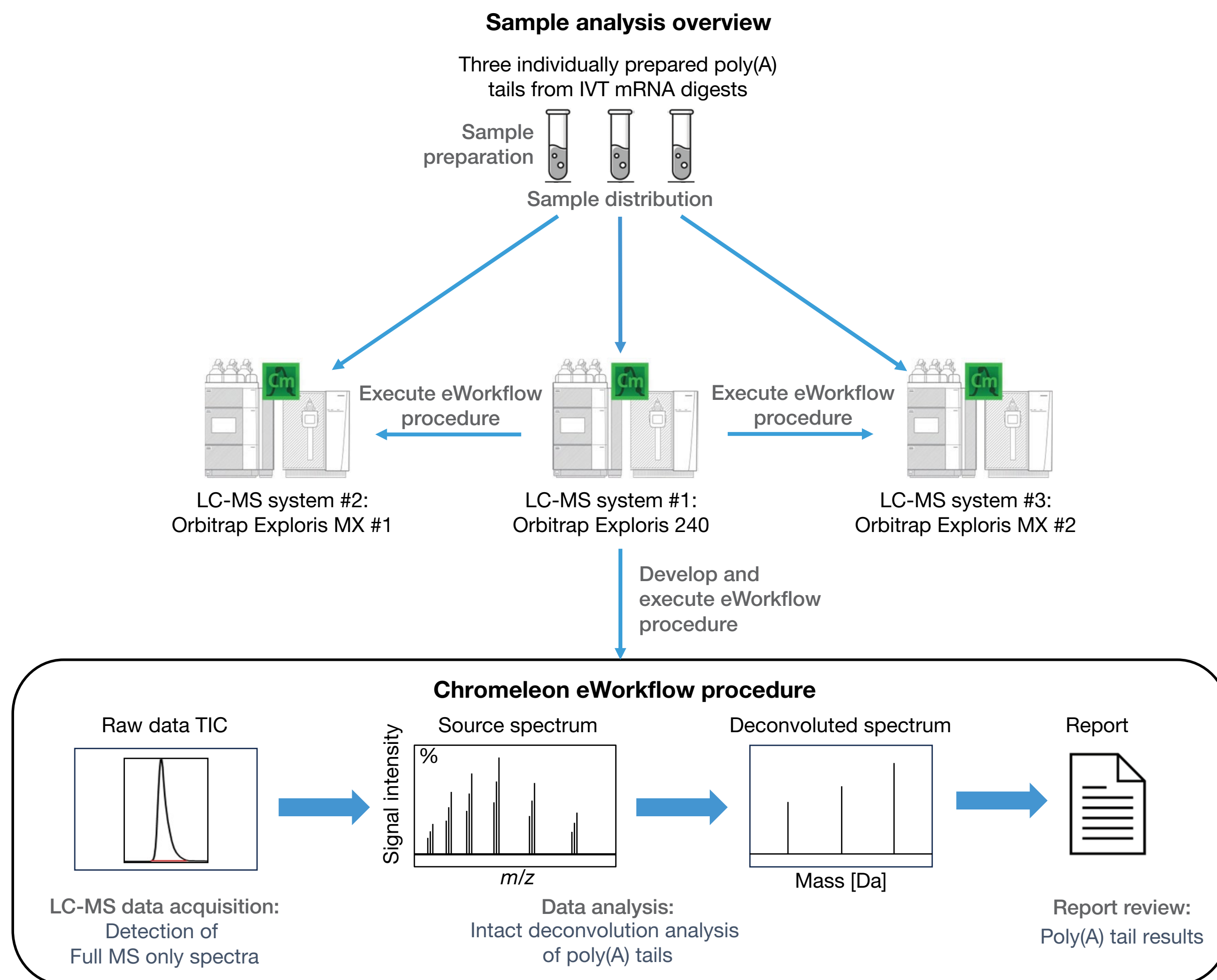

Figure 1. Schematic overview of the study layout for automated analysis and reporting of poly(A) tails from three different IVT Cas9 mRNA digests. The Chromeleon eWorkflow procedure was developed on an Orbitrap Exploris 240 mass spectrometer and then executed on the same system as well as on two Orbitrap Exploris MX mass detectors.

Experimental

Reagents and consumables

- CleanCap™ Cas9 mRNA (TriLink BioTechnologies, L-7606)
- Thermo Scientific™ RNase T1 (1000 U/µL) (P/N EN0542)
- Invitrogen™ Dynabeads™ Oligo(dT)$_{25}$ (P/N 61002)
- Invitrogen™ Nuclease-Free Water (not DEPC-Treated) (P/N 9938)
- Thermo Scientific™ MagJET Separation Rack, 12 × 1.5 mL tube (P/N FERMR02)
- Invitrogen™ Nonstick, RNase-free Microfuge Tubes, 1.5 mL (P/N AM12450)
- Thermo Scientific™ Savant™ SpeedVac™ Medium Capacity Vacuum Concentrators (P/N SPD140DDA-230)
- 1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP), 99.9% (Thermo Fisher Chemicals™, P/N 293410500)
- Dibutylamine (DBA), 99.5% (Sigma-Aldrich, 471232-100ML)
- Thermo Scientific™ DNAPac™ RP HPLC column, 2.1 × 50 mm, 4 µm (P/N 088924)
- Thermo Scientific™ Water, UHPLC-MS grade (P/N W8-1)
- Thermo Scientific™ Acetonitrile, UHPLC-MS grade (P/N A956-1)
- Thermo Scientific™ 9 mm Screw Thread Vials, Polypropylene, 12 × 32 mm, 400 µL (P/N C4000-11)
- Thermo Scientific™ 9 mm Autosampler Vial Screw Thread Caps, Polypropylene (P/N C5000-50)

Sample preparation

Three poly(A) tail samples were extracted using a previously published procedure.[6] Briefly, 20 µg of Cas9 mRNA sample each were digested with 1,000 U of RNase T1 for 1 hr in a thermal mixer set to 37 °C with shaking at 400 rpm. Subsequently, the poly(A) tails were extracted using 100 µL of Dynabeads™ Oligo(dT)$_{25}$ using the recommended procedure as outlined in the manual. Extracted samples were dried down with a SpeedVac vacuum concentrator and reconstituted with 100 µL of 95:5 solvent A:solvent B (v/v) prior to injection.

Chromeleon eWorkflow procedure

A Chromeleon eWorkflow procedure consisting of a pre-defined injection sequence, acquisition method, processing method, and fit-for-purpose report template was created, concluding in an automatic reporting of the poly(A) tail lengths. It was developed on a Vanquish Horizon UHPLC coupled to a Thermo Scientific™ Orbitrap Exploris™ 240 mass spectrometer and subsequently executed on three different LC-MS systems, two of which comprised Thermo Scientific™ Orbitrap Exploris™ MX mass detectors. Below outlines the details for each component in the Chromeleon eWorkflow procedure.

1. Injection sequence

The injection sequence for poly(A) tail analysis consisted of a blank injection (i.e., 10 µL of solvent A) followed by injections of the three poly(A) tail samples. The sample injections were paired with the acquisition method and a processing method. The injection sequence was also setup to enable “Perform Intact Protein Deconvolution auto-processing” under Sequence Properties/Auto Reporting tab, a new feature in Chromeleon CDS 7.3.2, which automatically processes the raw data to determine the poly(A) tail length in the report.

2. Chromatography

A Vanquish Horizon UHPLC system consisting of the following modules was used for IPRP-LC separation of the poly(A) tail analysis:

- Vanquish System Base (P/N VF-S01-A-01)
- Vanquish Binary Pump H (P/N VH-P10-A-01)
- Vanquish Split Sampler FT (P/N VH-A10-A-02)
- Vanquish Column Compartment H (P/N VH-C10-A-03)

Ten microliters of sample were injected onto a DNAPac column and separated using a 15-minute gradient as outlined in Table 1.

Table 1. LC and autosampler conditions

Parameter	Setting	
HPLC column	DNAPac RP 2.1 × 100 mm, 4 µm	
Flow rate	0.4 mL/min	
Solvent A	15 mM DBA and 25 mM HFIP in water	
Solvent B	Acetonitrile	
Gradient	*Time (min)*	*%B*
	0	15
	0.5	15
	10	40
	10.5	90
	12	90
	12.5	15
	15	15
Injection volume	10 µL	
Needle wash	Before draw, 24 µL/s for 10 s with 10% methanol	
Thermostatting mode	Forced Air	
Column oven temperature	70 °C	

3. Mass spectrometry

The Orbitrap Exploris 240 mass spectrometer (P/N BRE725535) and Orbitrap Exploris MX mass detector (P/N BRE725536) equipped with the BioPharma Option (P/N BRE725539) were used for poly(A) tail analysis. A Full scan method operated in intact protein application mode with low pressure setting was applied to obtain a higher signal-to-noise ratio of the poly(A) tails, thereby improving the sensitivity of the method. Table 2 outlines the ion source and Full scan parameter settings. The same method was executed on the Orbitrap Exploris MX mass detector upon import of the Chromeleon eWorkflow procedure for automated poly(A) tail analysis.

Table 2. MS source and scan parameter settings

Parameter	Setting
Instrument	Orbitrap Exploris 240 and Orbitrap Exploris MX
MS source parameters	
Negative ion (V)	3,000
Sheath gas (Arb)	50
Aux gas (Arb)	10
Sweep gas (Arb)	1
Ion transfer tube temperature (°C)	320
Vaporizer temperature (°C)	320
Full scan parameters	
Application mode	Intact protein
Pressure mode	Low pressure
Expected LC peak width (s)	12
Resolution	180,000
Scan range (*m/z*)	1,000–2,000
Time range (min)	0–10 minutes
RF lens (%)	100
AGC target	300%
Maximum injection time (ms)	200
Microscans	3

4. Processing method – intact mass deconvolution

Intact mass deconvolution was used to determine the poly(A) tail lengths. Source spectra were obtained by averaging scans across the set retention time range from 5.6 to 5.9 minutes and were deconvoluted using the Xtract algorithm parameters as outlined in Table 3.

Table 3. Intact mass deconvolution processing parameters

Parameter	Setting
Deconvolution algorithm	Xtract (Isotopically Resolved)
Output mass range	30,000 to 60,000
Output mass	M
S/N threshold	3
Relative abundance threshold (%)	0
Charge range	5 to 50
Min. number of detected charge	5
Isotope table	Nucleotide

5. Summary report – Poly(A) tail lengths

A summary report was built based on the default MS report template version 2.0 with a few adjustments for optimal presentation of the results. The report comprises four sections: 1) a total ion chromatogram (TIC), 2) a source spectrum obtained by averaging the full MS scans across the retention time range from 5.6 to 5.9 minutes, 3) a deconvoluted spectrum using the Xtract algorithm, and 4) detailed deconvolution results providing measured monoisotopic masses, the number of detected charge states, relative abundances, theoretical monoisotopic mass of poly(A) tails, delta ppm between the measured and theoretical monoisotopic mass, matching tail length, and the median tail length.

The deconvolution results table exclusively displays components with matching tail lengths. To identify the poly(A) tail length, an MS scoring criterion was established, comparing the deconvoluted monoisotopic masses to the theoretical monoisotopic masses of the poly(A) tails. The scoring criterion employed the following formula:

$$\frac{(\mathit{component.massToChargeRatio} - \mathit{ipdComponent.mass})}{(\mathit{component.massToChargeRatio})} \cdot 1{,}000{,}000$$

The calculated value had to be within the range from -20 to 20 ppm. In this formula, *"component.massToChargeRatio"* represents the theoretical monoisotopic masses for poly(A) tails, which were imported from the Biopharma Finder software intact workbook containing all previously detected poly(A) tails with lengths ranging from 117 to 132.[5] *"ipdComponent.mass"* are the experimental monoisotopic masses obtained from the

deconvolution results. All analyzed poly(A) tails were measured with a mass accuracy of below 5 ppm (Figure 2). Additionally, a separate “Tail Length” column was created to keep the numeric values of the matching tail lengths. These numeric values were subsequently used to calculate the median tail length for each sample injection.

Software

Chromeleon CDS version 7.3.2 MUb was used for all data acquisition, intact mass deconvolution, targeted MS processing, and reporting. The MS systems were operated under Orbitrap Exploris Tune 4.2 SP1 software.

Results and discussion

In this study, we prepared three separate Cas9 mRNA digests, and the purified extracts were analyzed on three different Orbitrap-based LC-MS systems using the IPRP-LC-HRAM-MS method through the deployment of a Chromeleon eWorkflow procedure. The extraction of the poly(A) tails from a complete RNAse T1 digest has been previously described.[6] Using the Dynabeads Oligo$(dT)_{25}$ magnetic beads, we were able to selectively extract the intact poly(A) tails from the digestion mixture, efficiently purifying the samples, which resulted in a relatively clean TIC. With high resolution mass spectrometric methods, we were able to determine poly(A) tail variants, at single nucleotide base resolution via an automated intact mass deconvolution of the source spectrum. A customizable report was developed for easy and quick data review and reporting. Figure 2 shows an example of the result summary provided in the report for the analysis of the poly(A) tails. Within the reporting, the calculation of the mass difference, referred to as delta parts-per-million (ppm), between the measured and the theoretical monoisotopic mass of a specific tail length is performed. If the calculated mass errors were within the predefined range of ±20 ppm, a positive match to that specific tail length was obtained (Figure 2, panel D). As shown in the result table of Figure 2, accurate mass measurements of the poly(A) tails from 117 to 132 adenosines at single nucleotide resolution were achieved with mass accuracies of less than 5 ppm. Based on the identified tail lengths, a median poly(A) tail length for that specific sample injection was automatically determined and is presented on the bottom right of the deconvolution results table. This allowed quick assessment of the poly(A) tail distribution and median tail length for every single sample injection without any manual intervention.

To demonstrate reproducibility of the assay, three replicate digestions of the Cas9 mRNA were injected per LC-MS system and analyzed using the Chromeleon eWorkflow. As illustrated in Figure 3, the obtained deconvoluted spectrum from the first injection on each LC-MS system was highly comparable. The determined poly(A) tail distribution was not only consistent across three samples, but the obtained results were also reproducible across replicate injections of each sample and consistent across the three LC-MS systems as shown in Table 4. In addition, by enabling “Perform Intact Protein Deconvolution auto-processing” under the Sequence Properties/Auto Reporting tab, a new feature introduced in Chromeleon CDS 7.3.2, reports containing full deconvolution results with matching tail lengths were automatically generated at the end of the acquisition without the need for manual processing. This allowed quick reporting of the median poly(A) tail length based on the obtained distributions, which was centered around 124 to 125 adenosines for IVT Cas9 mRNA in this evaluation. This new feature will be useful for validating this method as a potential quality control assay for mRNA therapeutics.

A

Chromatogram and Results

Injection Details

Injection Name:	**polyA tail**	*Run Time (min):*	**10.02**
Vial Number:	**Y:E3**	*Injection Volume:*	**10.00**
Injection Type:	**Unknown**	*Dilution Factor:*	**1.0000**
Calibration Level:		*Sample Weight:*	**1.0000**
Instrument Method:	**PolyA tail IPD**		
Processing Method:	**PolyA IPD**		
Injection Date/Time:	**28/Apr/23 16:29**		

B

Total Ion Chromatogram

Source Spectrum at 180,000 Orbitrap Resolution

C

Deconvoluted Spectrum

D

Deconvolution Results

Result Component	Measured Monoisotopic Mass	Number of Detected Charge States	Relative Abundance %	Theoretical Monoisotopic Mass	Delta PPM	Matching Tail Length	Tail Length
Result Component 18	38437.3607	5	6.07	38437.1890	4.47	117A tail	117
Result Component 14	38766.3792	7	14.11	38766.2416	3.55	118A tail	118
Result Component 11	39095.3934	11	30.45	39095.2941	2.54	119A tail	119
Result Component 9	39424.4242	12	47.03	39424.3466	1.97	120A tail	120
Result Component 6	39753.4846	15	71.42	39753.3991	2.15	121A tail	121
Result Component 3	40082.5754	15	91.75	40082.4516	3.09	122A tail	122
Result Component 1	40411.6110	15	100.00	40411.5042	2.64	123A tail	123
Result Component 2	40740.6818	16	94.38	40740.5567	3.07	124A tail	124
Result Component 4	41069.7187	15	89.64	41069.6092	2.67	125A tail	125
Result Component 5	41398.7478	14	77.20	41398.6617	2.08	126A tail	126
Result Component 7	41727.8089	13	59.76	41727.7142	2.27	127A tail	127
Result Component 8	42056.8565	13	50.61	42056.7668	2.13	128A tail	128
Result Component 10	42385.8967	13	40.45	42385.8193	1.83	129A tail	129
Result Component 12	42715.0441	10	26.83	42714.8718	4.03	130A tail	130
Result Component 13	43044.0180	9	19.52	43043.9243	2.18	131A tail	131
Result Component 15	43373.1336	6	12.46	43372.9768	3.61	132A tail	132
						Median Length	124.5

Figure 2. Example summary report for poly(A) tail length showing injection details (panel A), the total ion chromatogram and a source spectrum generated by averaging the scan spectra across the peak (panel B), a deconvoluted spectrum (panel C), and a detailed deconvolution result table with matching tail length (panel D). The "Deconvolution Results" table was sorted by the monoisotopic mass column to display from the lowest to the highest value.

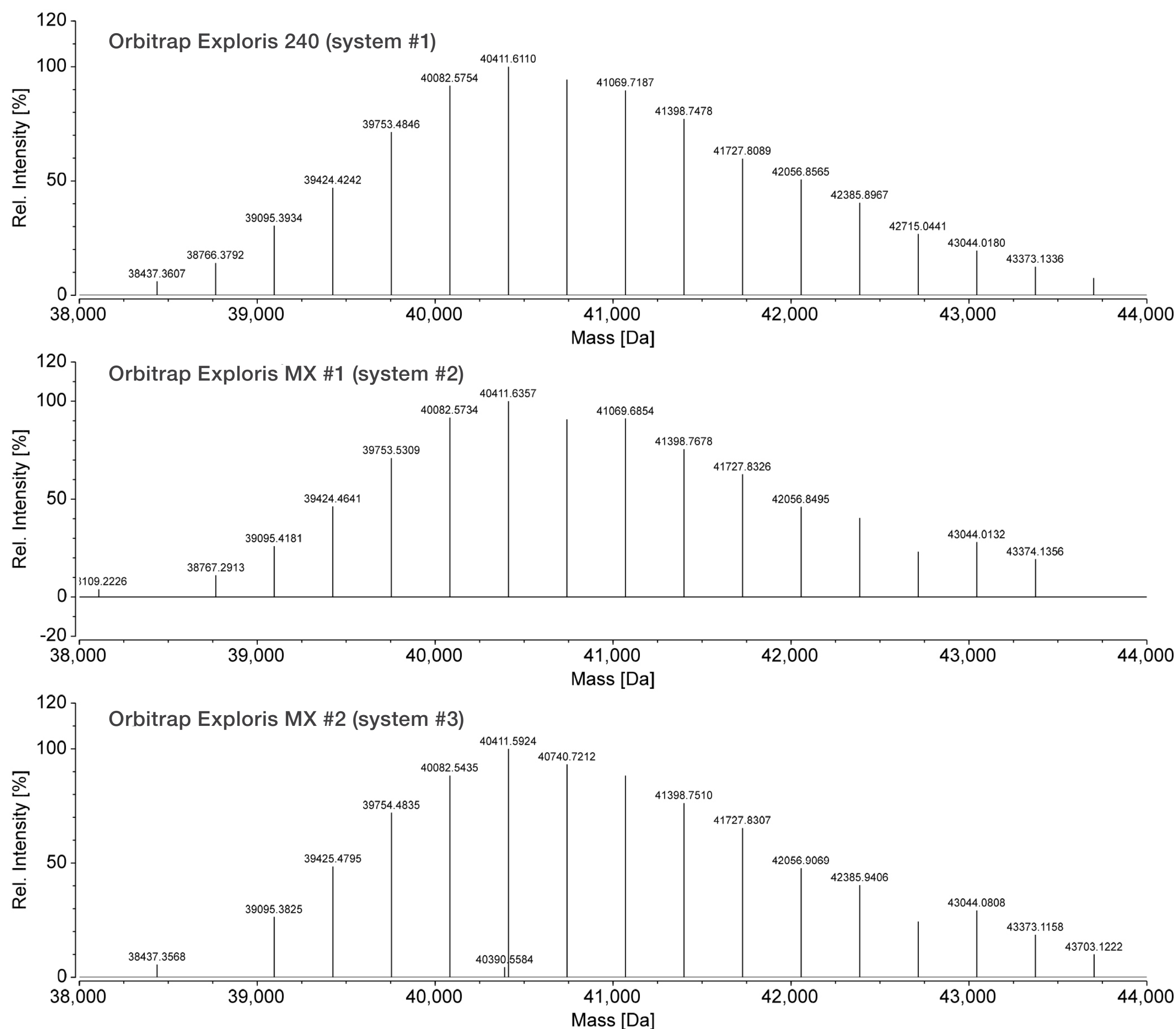

Figure 3. Poly(A) tail deconvolution results obtained on three Orbitrap-based LC-MS systems displaying the first injection each

Table 4. Evaluation of poly(A) tail distribution and median length across three Orbitrap-based LC-MS systems using the Chromeleon eWorkflow procedure. The distribution results from three replicate injections are shown for each LC-MS system, and median tail length was estimated based on the distribution ranging from the lowest to the highest detected tail length across three injections.

Poly(A) tail	LC-MS system #1		LC-MS system #2		LC-MS system #3	
	Range	Medium	Range	Medium	Range	Medium
Injection 1	117–132	124.5	119–131	125	117–132	124.5
Injection 2	118–132	125	118–132	125	119–132	125.5
Injection 3	117–132	124.5	118–132	125	118– 32	125

Thermo Fisher SCIENTIFIC

Conclusions

Here, we demonstrate the development and deployment of an eWorkflow procedure that is an automated and compliance-ready analytical solution for accurate and consistent analysis of poly(A) tails of IVT Cas9 mRNA using a LC-MS method with a resolution setting of 180,000 (at *m/z* 200), operated in intact protein application mode with the low-pressure setting enabled via the BioPharma option.

The obtained results demonstrate:

- Accurate mass measurements of the poly(A) tails were achieved at single nucleotide resolution with mass accuracies of less than 5 ppm.
- Reproducible and consistent results were obtained for three separately prepared IVT Cas9 mRNA digests analyzed on three different LC-MS systems comprising of Orbitrap Exploris 240 and Orbitrap Exploris MX systems upon deployment of an eWorkflow procedure.
- The compliance-ready assessment of the poly(A) tail distributions and median poly(A) tail lengths via an automated fit-for-purpose reporting resulted in median poly(A) tail length of 124.5 and 125 adenosines, respectively.

The Chromeleon eWorkflow procedure allows for easy transfer across instruments and across laboratories, and with that supports significantly reduced implementation effort of the solution in a QC environment.

References

1. Clifford T. H. et al, mRNA-LNP expressing PfCSP and Pfs25 vaccine candidates targeting infection and transmission of Plasmodium falciparum, *NPJ Vaccines* **2022,** *7*(1). DOI: 10.1038/s41541-022-00577-8
2. Neuzil, K. M., An mRNA Influenza Vaccine—Could It Deliver?, *New England Journal of Medicine* **2023.** DOI: 10.1056/NEJMcibr2215281
3. Qin, S. et al, mRNA-based therapeutics: powerful and versatile tools to combat diseases, *Signal Transduction and Targeted Therapy* **2022**, *7*, DOI: 10.1038/s41392-022-01007-w
4. Beverley, M. et al, Poly A tail length analysis of in vitro transcribed mRNA by LC-MS, *Analytical and Bioanalytical Chemistry* **2018,** *410,* 1667–1677.
5. Thermo Fisher Scientific, Application Note 001183: Characterization of in vitro-transcribed (IVT) mRNA poly(A) tail by LC-HRAM-MS and BioPharma Finder 5.0 software. https://assets.thermofisher.com/TFS-Assets/CMD/Application-Notes/an-001183-bt-orbitrap-exploris-240-ms-mrna-hram-lc-ms-an001183-na-en.pdf

Learn more at **thermofisher.com/mrna**

General Laboratory Equipment – Not For Diagnostic Procedures. © 2024 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. CleanCap is a trademark of TriLink BioTechnologies, LLC. This information is presented as an example of the capabilities of Thermo Fisher Scientific products. It is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others. Specifications, terms and pricing are subject to change. Not all products are available in all countries. Please consult your local sales representative for details. **AB003060-EN 0724S**

thermo scientific