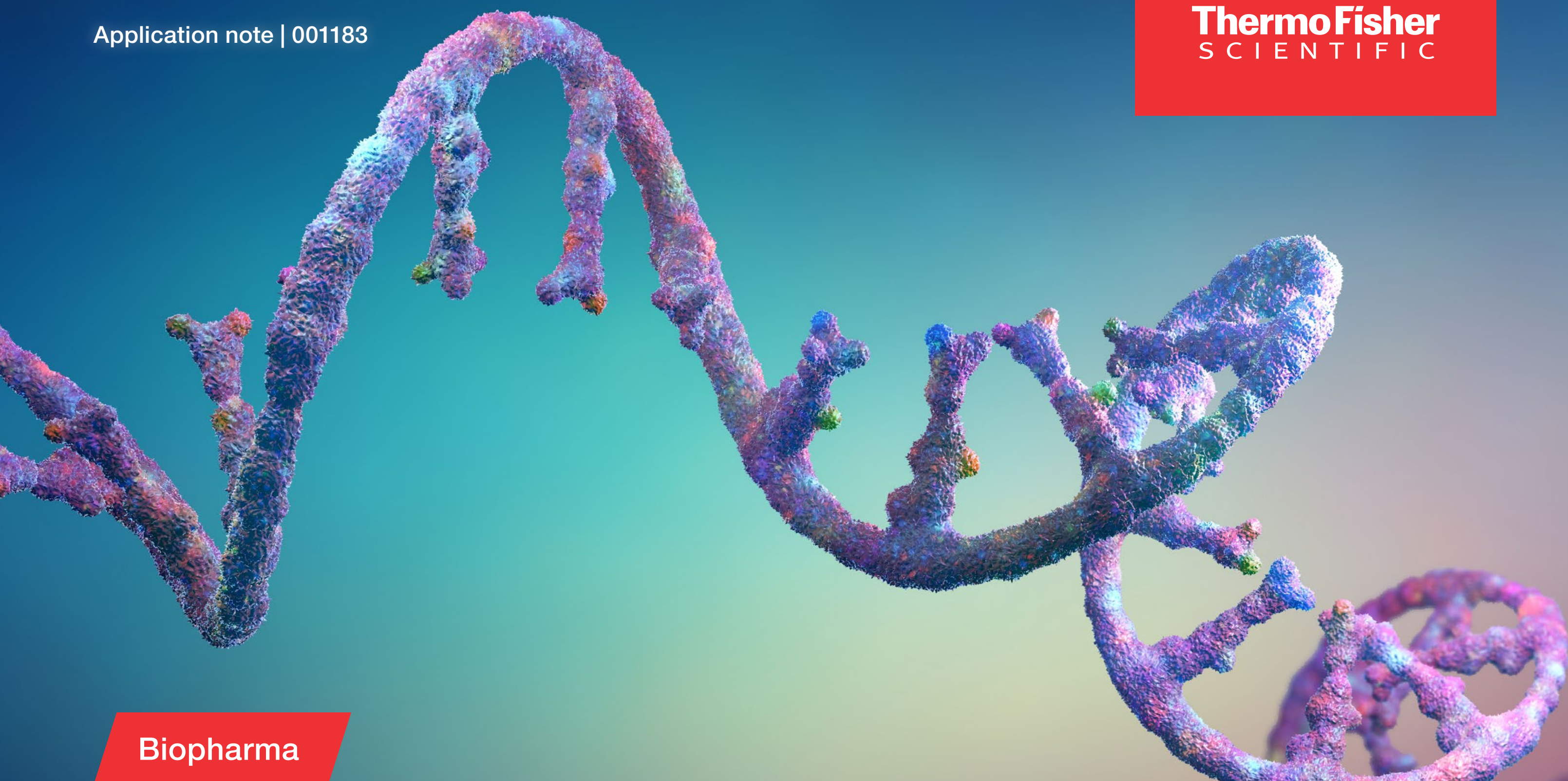


Characterization of *in vitro*-transcribed (IVT) mRNA poly(A) tail by LC-HRAM-MS and BioPharma Finder 5.0 software

Authors

Robert L. Ross[1], Jenny England[2], Rhonda Bassignani[2], Keeley Murphy[3], Yi Zhang[4], Min Du[3]

[1]Thermo Fisher Scientific, Franklin, MA
[2]Thermo Fisher Scientific, Bedford, MA
[3]Thermo Fisher Scientific, Cambridge, MA
[4]Thermo Fisher Scientific, Philadelphia, PA

Keywords

LC-HRAM-MS, Orbitrap Exploris 240 MS, BioPharma Finder software, IVT, Poly(A), mRNA

Application benefits

- Confident identification and sequence confirmation of polyadenylated tails in synthetic mRNA transcripts using high-resolution, accurate mass spectrometry (HRAM)
- Streamlined workflow for the characterization of the polyadenylated tail using Thermo Scientific™ BioPharma Finder™ 5.0 software

Goal

To develop a sensitive and robust LC-MS method for identification and sequence confirmation of 3' polyadenylated tails from IVT mRNA products

Introduction

The central dogma of molecular biology[1] states that, in general, the genetic information stored in our DNA (deoxyribonucleic acid) is transcribed into RNA (ribonucleic acid), which is then translated into protein. During transcription, a "copy" of the expressed gene is made that carries the message encoded in the DNA. This transcript is a strand of RNA, messenger RNA or mRNA, that once processed and exported to the ribosome, is then translated into proteins that keep cellular processes functioning. The mature mRNA is composed of five regions, each with specific biological significance: the 5' cap, the 5' and 3' untranslated regions, the reading frame, and the poly(A) tail.

The poly(A) tail, as its name suggests, is a string of adenosines added to the end of a nascent mRNA transcript.[2] The enzyme responsible for the polyadenylation was discovered in 1960, isolated from bovine thymus tissue, and was the first polymerase identified to use a nucleotide triphosphate in the biochemical addition.[3] The authors went on to show that the poly(A) polymerase, or PAP, was adding the adenosine biopolymer to the 3' end of mRNA.[4] Further investigation revealed the poly(A) tail recruits poly(A) binding proteins.[5] These RNA-protein complexes have various biological roles such as nuclear export post processing,[6] facilitating translation,[7,8] and regulating transcript degradation.[9,10]

For *in vitro*-produced mRNA, characterization of poly(A) tail length is a crucial part of transcript design. Determination of tail length assists in determination of translation efficiency,[11] which has a direct impact on the efficacy of the therapeutic. Analytical measurements of the poly(A) tail have been mostly RNA-Seq based,[12] relying on reverse transcription followed by amplification. These experiments can be costly due to instrumentation requirements, sample workup, and data analysis. We present here an HRAM-LC-MS method for determining the length of polyadenylated tails from *in vitro*-transcribed mRNA.

Experimental

Reagents, consumables, and lab equipment

- Thermo Scientific™ UHPLC-MS Water (P/N W81)
- Thermo Scientific™ UHPLC-MS Acetonitrile (P/N A9554)
- 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)
- Thermo Scientific™ F1-ClipTip™ Variable Volume Single Channel Pipettes (P/N 4641210N)
- Invitrogen™ Nonstick, RNase-free Microfuge Tubes, 1.5 mL (P/N AM12450)
- Invitrogen™ Nonstick, RNase-free Microfuge Tubes, 2 mL (P/N AM12475)
- Thermo Scientific™ Savant™ SpeedVac™ Medium Capacity Vacuum Concentrators (P/N SPD131DDA)
- Thermo Scientific™ NanoDrop™ One/OneC Microvolume UV-Vis Spectrophotometer (P/N ND-ONEC-W)
- 1,1,1,3,3,3-Hexafluoro-2-propanol, 99%, for analysis, Thermo Scientific™ (P/N AAA1274722)
- Di-n-butylamine, 98+%, Thermo Scientific™ (P/N AAA11671AP)

Sample preparation — digestion of IVT mRNA with RNase T1

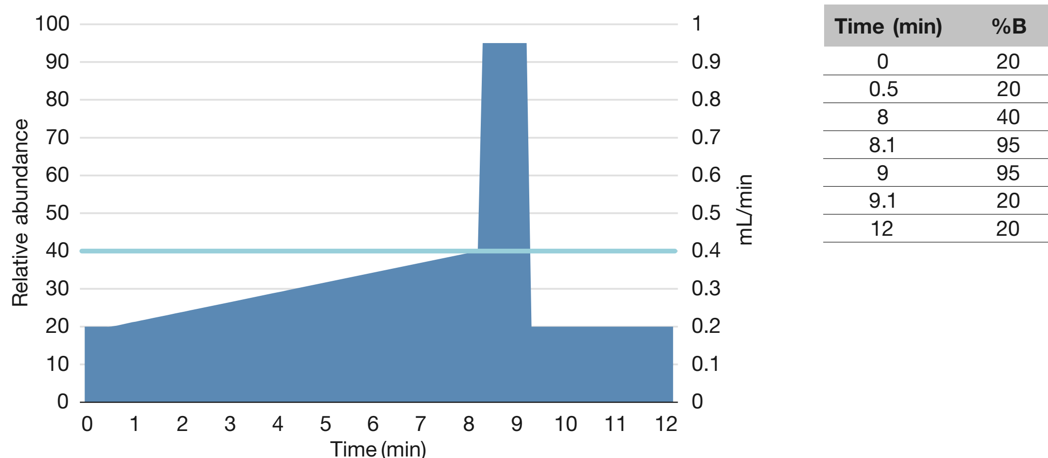
To a low binding Eppendorf tube was added 100 pmol of IVT-produced mRNA, 10 µL of 0.1 mM $ZnCl_2$, and 1,000 U RNase T1. The solution was brought up to 50 µL with Nuclease-Free water and allowed to digest for 1 hr in a thermal mixer set to 37 °C and 400 RPM. At the end of the hour, the tube was removed and the polyadenylated tails were purified from the digest using 100 µL of Dynabeads Oligo$(dT)_{25}$ magnetic beads, following the recommended procedure.

Ion-pair reversed-phase liquid chromatography (IP-RP LC)

For all experiments, the Thermo Scientific™ Vanquish™ Horizon UHPLC system was used, consisting of:

- Thermo Scientific™ Vanquish™ System Base (VF-S01-A-02)
- Thermo Scientific™ Vanquish™ Binary Pump H (P/N VH-P10-A-02)
- Thermo Scientific™ Vanquish™ Split Sampler HT (P/N VH-A10-A-02)
- Thermo Scientific™ Vanquish™ Column Compartment H (P/N VH-C10-A-02)
- Thermo Scientific™ Viper™ MS Connection Kit Vanquish™ LC systems (P/N 6720.0405)
- Mobile phase A: 25 mM HFIP, 15 mM dibutylamine in H_2O
- Mobile phase B: 25 mM HFIP, 15 mM dibutylamine in ACN

Oligonucleotide separations were performed with a Thermo Scientific™ DNAPac™ RP column (4 µm, 2.1 × 100 mm, P/N 088924) using a Vanquish Horizon UHPLC system. The autosampler was held at 5 °C while the column was maintained at 70 °C with the column oven thermostat mode set to Still Air. The LC gradient used in this study is shown in Figure 1.



Time (min)	%B
0	20
0.5	20
8	40
8.1	95
9	95
9.1	20
12	20

Figure 1. Reversed phase gradient used in these experiments. Both mobile phase A (water) and B (acetonitrile) contain the modifiers hexafluoroisopropanol (HFIP) and dibutylamine (DBA).

Mass spectrometry

The Thermo Scientific™ Orbitrap Exploris™ 240 mass spectrometer (P/N BRE725535) was operated with Thermo Scientific™ Chromeleon™ 7.3.1 Chromatography Data System (CDS) software and controlled by Thermo Scientific™ Orbitrap Exploris™ Series 2.0 instrument control software (ICSW). Instrument calibration was performed using Thermo Scientific™ Pierce™ FlexMix™ calibration solution. Data acquisition was performed in negative ion mode. The MS method was built in the method editor using the standard MS template provided with the Orbitrap Exploris instrument control software and then modified accordingly. Table 1 lists the scan parameters used in these acquisitions. Global settings were default for the flow rate (400 µL min^{-1}) used in the experiment.

Table 1. MS parameters used in the analysis

Master scan	
Full scan	**Value**
Orbitrap resolution	240,000
Scan range (*m/z*)	1,000–2,000
RF lens (%)	70
AGC target	Custom
Normalized AGC target (%)	200
Maximum injection time mode	Custom
Maximum injection time (ms)	300
Microscans	2
Data type	Profile
Polarity	Negative

Software

- Chromeleon 7.3.1 CDS software (P/N CHROMELEON7)
- Thermo Scientific™ Freestyle™ 1.8.2 software
- Thermo Scientific™ BioPharma Finder™ 5.0 software (P/N OPTON-30988)

Intact mass analysis in BioPharma Finder software

BioPharma Finder software has two independent deconvolution algorithms in its Intact Mass Analysis workflow: Xtract™, used to deconvolute isotopically resolved mass spectra, and ReSpect™, which deconvolves isotopically unresolved mass spectra. For determination of poly(A) length, we use the Xtract algorithm. As in previous mRNA analysis utilizing BioPharma Finder software, we begin with the Sequence Manager. Accessing the Sequence Manager allows for creation of the poly(A) sequence used to match the acquired data against (Figure 2).

The length of the tail from an IVT construct is known. By creating a poly(A) sequence longer than the expected mass, we can use the created sequence to compare the deconvoluted masses against. For our experiments, we created a sequence 140 adenosines long, ~15 adenosines longer than our theoretical poly(A) tail. This was exploited to determine and annotate the deconvoluted spectra using a terminal truncation search option under the Sequence Manager pane *Assign Variable Modifications* (Figure 3).

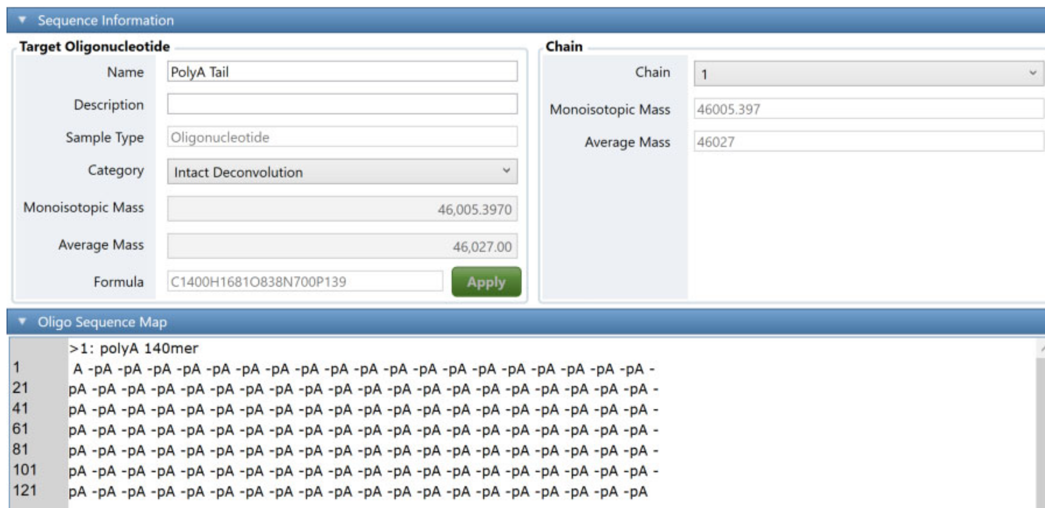


Figure 2. Sequence Manager pane showing a 140mer polyA sequence under the category of Intact Deconvolution

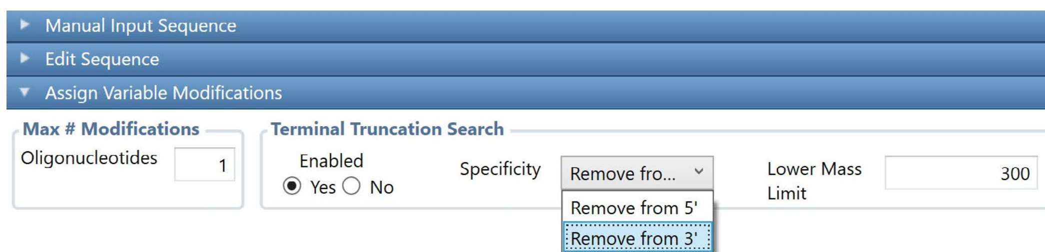


Figure 3. Terminal Truncation Search in BioPharma Finder software, Assign Variable Modifications pane. Using a truncated end search, the software can align deconvoluted masses against the poly(A) sequence to assign annotation.

Results and discussion

UHPLC-HRAM MS analysis of the mRNA poly(A) tail

Characterization of the mRNA 3' tail is accomplished through the application of the enzyme RNase T1. RNase T1 is a naturally occurring endonuclease and shows base specificity, primarily cleaving ssRNA at guanosine residues to generate oligonucleotides with a 3' Gp. We can exploit this base specificity in the characterization of poly(A) tailed mRNA, as the enzyme will digest the intact mRNA at the guanosine residues, generating small oligonucleotides, while leaving the poly(A) tail intact. The population of poly(A) tail is then purified from the digestion using oligo$(dT)_{25}$ magnetic beads as represented in Figure 4A.

Due to the negatively charged phosphate backbone, ion-pairing chromatography is used to resolve the biopolymer. Figure 4B shows the chromatographic peak from an injection of oligo$(dT)_{25}$ purified T1 digest of poly(A) mRNA. It was found that the use of dibutylamine (DBA) resulted in better peak shape and retention than of the commonly used triethylamine (TEA).[13] It should be noted that the use of RNase T1, specifically in these experiments, is due to the last nucleotide in our transcript being a guanosine. Should a different nucleotide be present, the use of RNase A with RNase T1 can be used. RNase A is an endonuclease that shows preference for pyrimidine residues and is active in the same reaction conditions as T1.

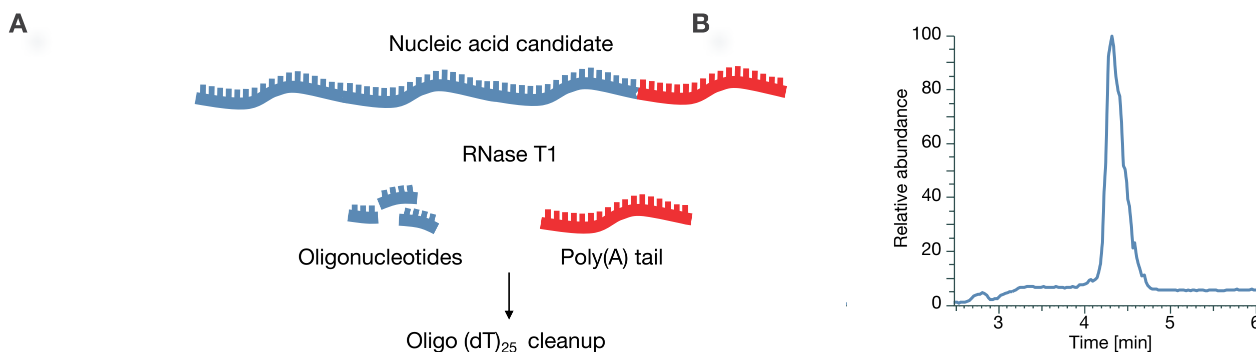


Figure 4. (A) Representation of the digestion workflow; (B) Chromatographic peak of purified poly(A) tails

The Orbitrap Exploris 240 mass spectrometer provides the quantitative precision and accuracy needed for identification of the adenylated strand. The length of poly(A) tails averages around 125 nucleotides, with the main charge-carrying atom in the strand being the hydroxyl in the phosphate backbone. The electronegative backbone can help with desorbing from the electrospray droplet, as coulombic repulsion between the phosphate backbone and the surface charge on the droplet repel, releasing the biopolymer into the gas phase where it can be sampled in the mass spectrometer. Figure 5 shows a mass spectrum of a poly(A) pool purified from a RNase T1 digest of a commercially available mRNA. The presence of overlapping multiply charged ions results in spectra too complicated to manually interpret. The use of BioPharma Finder software's Intact Mass Analysis workflow deconvoluted the multiply charged spectra into its individual monoisotopic masses. Figure 6 shows a zoomed image of a single isotopically resolved charge state from the mass spectrum in Figure 5. The use of the Xtract deconvolution algorithm is recommended when isotopically resolved peaks are present. Xtract uses clusters of isotopically resolved peaks and their spacing to determine an initial estimate of mass. Once the initial estimate is determined, the algorithm uses an Averagine distribution[14] to determine a monoisotopic mass that best fits the estimate. The algorithm then combines the results determined for all the observed charge states for each mass to report the single monoisotopic mass value.

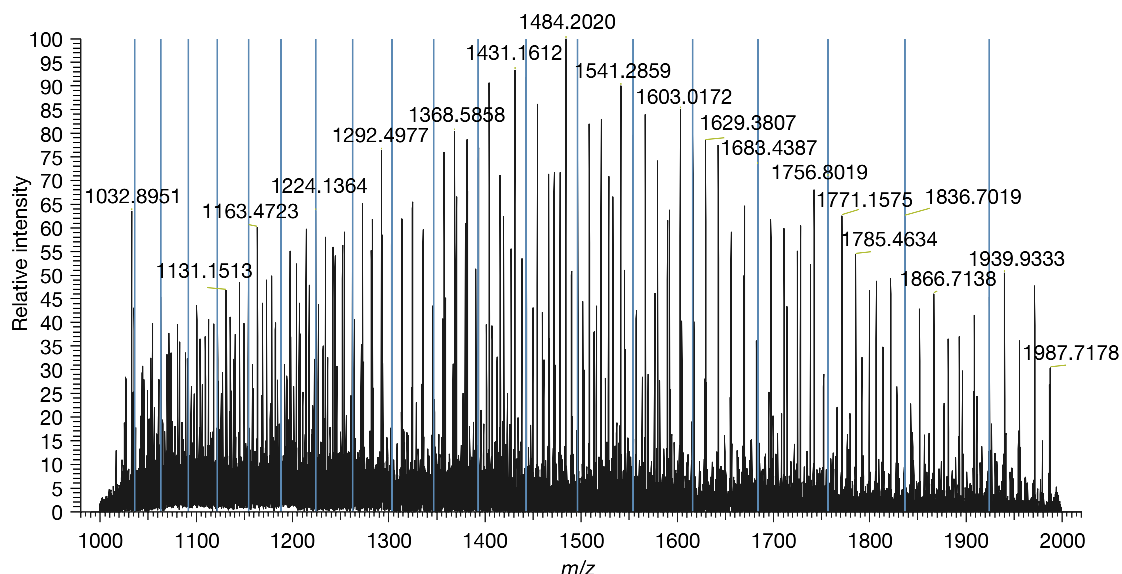


Figure 5. Mass spectrum of a purified poly(A) pool. Blue lines indicate charge states used for Xtract algorithm deconvolution of the mass of a singular length of poly(A) in the sample.

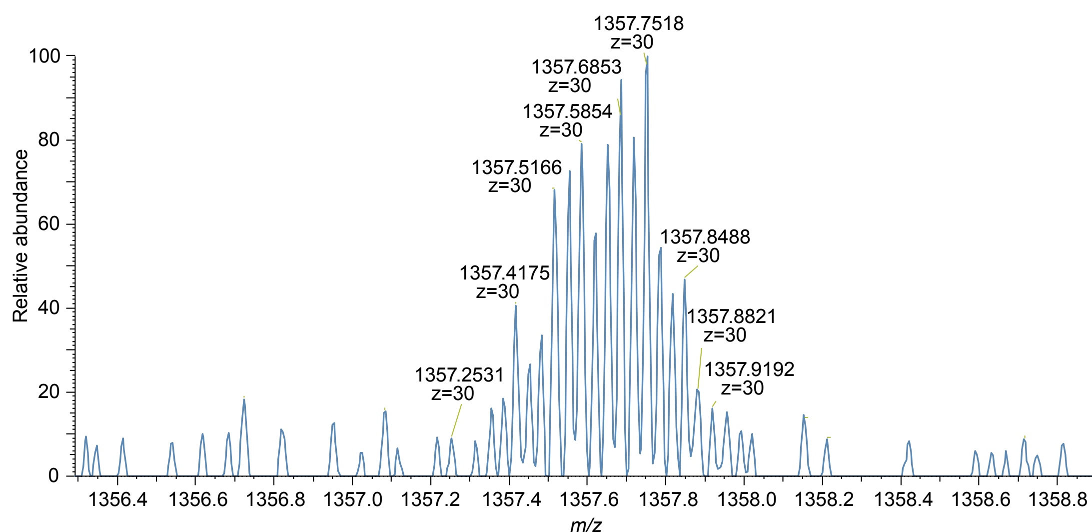


Figure 6. Zoomed image of an isotopically resolved 30- charge state. Isotopic resolved peaks are used with the Xtract deconvolution algorithm for determination of monoisotopic mass.

Deconvoluted results are returned in table format with the option to choose the fields to display. Table 2 lists the results of the poly(A) Intact Mass Analysis listing the annotated Modification, the determined monoisotopic mass, theoretical mass, and the mass difference between the two. The matched mass error for this analysis was between 6 and 3 ppm for the detected features, an extraordinarily accurate measurement when considering the complexity of the sample and mass of the oligonucleotide mixture. Also listed in Table 2 is the relative abundance of the detected feature, the number of detected charge states per deconvoluted mass, and the distribution of charge state between them. It is interesting to note that the charge state distribution for the poly(A) tails falls between 20 and 40. As a rule of thumb, charge state is normally half the length, so a 100mer would theoretically have a charge state distribution of ~50. The lower distribution for poly(A) highlights the challenge posed by the analysis and reinforces the need for proper sample preparation and mobile phase modifier pairing.

The goal of this application was to develop a method for determination of tail length in mRNA vaccine development. By using Biopharma Finder software Intact Mass Analysis with the isotopically resolved oligonucleotide data, we can produce a table of monoisotopic masses and their relative abundances. To determine tail length, we use a theoretical mass from a 140mer poly(A) sequence created in Sequence Manager. The BioPharma Finder software output from a deconvolution experiment annotates the monoisotopic masses determined through deconvolution based on their mass difference from the theoretical 140mer and is included in the results table under the Modification field. Figure 7 shows the resulting output from the Intact Mass Analysis with the monoisotopic masses annotated as truncations from the theoretical sequence. Because we know that the oligonucleotide is made from a polymeric chain of adenosines joined 5' to 3' by a phosphodiester bond, and that each individual mass unit is the mass of an adenosine monophosphate minus a water molecule, we can divide the monoisotopic masses listed in the table by 330 to get the number of adenosines in the poly(A) tail. This is then plotted against the relative abundance to give the distribution of tail lengths in the sample (Figure 8).

Table 2. Intact Mass Analysis Results table listing deconvoluted results

#	Modification	Monoisotopic Mass	Theoretical Mass (Da)	Matched Mass Error (ppm)	Average Mass	Relative Abundance	Number of Charge States	Charge State Distribution
1	9x3'Truncation,1xA	43373.153	43372.977	4.1	43,393.41	3.34	5	34 - 42
2	10x3'Truncation,1xA	43044.155	43043.925	5.3	43,064.26	11.10	10	30 - 41
3	11x3'Truncation,1xA	42715.071	42714.872	4.7	42,735.04	18.11	12	27 - 41
4	12x3'Truncation,1xA	42386.007	42385.820	4.4	42,405.82	30.70	18	24 - 41
5	13x3'Truncation,1xA	42056.934	42056.767	4.0	42,076.59	38.88	18	23 - 41
6	14x3'Truncation,1xA	41727.866	41727.715	3.6	41,747.36	54.37	19	22 - 40
7	15x3'Truncation,1xA	41398.802	41398.662	3.4	41,418.15	72.26	20	21 - 40
8	16x3'Truncation,1xA	41069.766	41069.610	3.8	41,088.96	80.17	19	22 - 40
9	17x3'Truncation,1xA	40740.735	40740.557	4.4	40,759.77	89.78	20	21 - 40
10	18x3'Truncation,1xA	40411.680	40411.505	4.3	40,430.56	100.00	19	21 - 39
11	19x3'Truncation,1xA	40082.647	40082.452	4.9	40,101.38	91.92	19	21 - 39
12	20x3'Truncation,1xA	39753.578	39753.400	4.5	39,772.15	73.92	19	20 - 38
13	21x3'Truncation,1xA	39424.504	39424.347	4.0	39,442.92	35.60	15	22 - 38
14	22x3'Truncation,1xA	39095.466	39095.295	4.4	39,113.74	18.06	13	24 - 38
15	23x3'Truncation,1xA	38766.380	38766.242	3.6	38,784.50	6.20	7	31 - 37

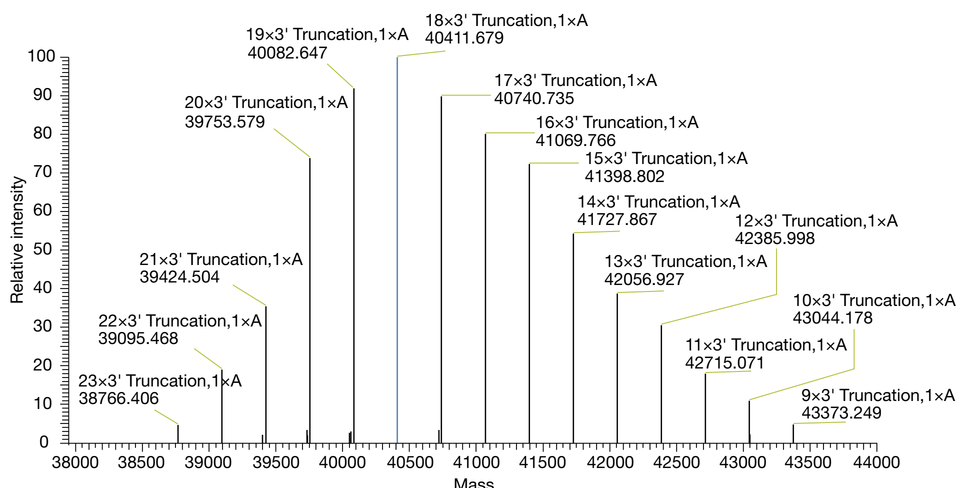


Figure 7. BioPharma Finder software Intact Mass Analysis deconvolution results of a poly(A) pool analysis. Monoisotopic masses are listed based on their relative abundance and annotated based on the difference between the detected mass and the theoretical mass of a 140mer.

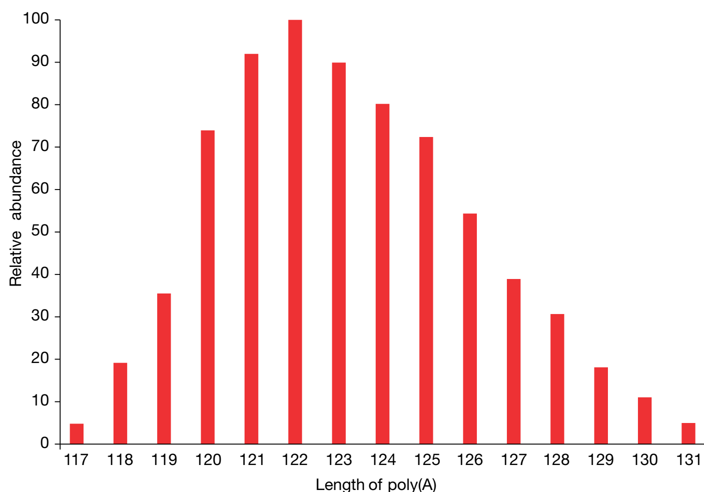


Figure 8. Length of poly(A) tails in a commercial mRNA sample. The figure was produced by dividing the monoisotopic mass by the mass of AMP-H_2O (330 Da) to give the number of adenosines in the individual tails.

Conclusion

The Orbitrap Exploris 240 mass spectrometer, coupled to the Vanquish Horizon UHPLC system, utilizing the DNAPac RP column for separation and BioPharma Finder software for data processing, yields a robust and powerful analytical platform for determination of poly(A) tail length in *in vitro*-produced mRNA.

- The Orbitrap Exploris 240 mass spectrometer returns exceptional mass accuracy for large polyadenylated species.
- BioPharma Finder software Intact Mass Analysis Xtract deconvolution algorithm effortlessly handles complex spectra, yielding precise measurements for poly(A) tail characterization workflows.
- DNAPac RP columns continue to provide excellent chromatographic retention for large oligonucleotides.

References

1. Crick, F. Central Dogma of Molecular Biology. *Nature* **1970**, *227*(5258), 561–563. https://doi.org/10.1038/227561a0
2. Millevoi, S.; Vagner, S. Molecular Mechanisms of Eukaryotic Pre-MRNA 3' End Processing Regulation. *Nucleic Acids Res.* **2010**, *38*(9), 2757–2774. https://doi.org/10.1093/nar/gkp1176
3. Edmonds, M.; Abrams, R. Polynucleotide Biosynthesis: Formation of a Sequence of Adenylate Units from Adenosine Triphosphate by an Enzyme from Thymus *Nuclei. J. Biol. Chem.* **1960**, *235*(4), 1142–1149. https://doi.org/10.1016/S0021-9258(18)69494-3
4. Edmonds, M.; Vaughan, M. H., Jr; Nakazato, H. Polyadenylic Acid Sequences in the Heterogeneous Nuclear RNA and Rapidly-Labeled Polyribosomal RNA of HeLa Cells: Possible Evidence for a Precursor Relationship. *Proc. Natl. Acad. Sci. U. S. A.* **1971**, *68*(6), 1336–1340. https://doi.org/10.1073/pnas.68.6.1336
5. Deo, R. C.; Bonanno, J. B.; Sonenberg, N.; Burley, S. K. Recognition of Polyadenylate RNA by the Poly(A)-Binding Protein. *Cell* **1999**, *98*(6), 835–845. https://doi.org/10.1016/S0092-8674(00)81517-2
6. Natalizio, B. J.; Wente, S. R. Postage for the Messenger: Designating Routes for Nuclear MRNA Export. *Trends Cell Biol.* **2013**, *23*(8), 365–373. https://doi.org/10.1016/j.tcb.2013.03.006
7. Weill, L.; Belloc, E.; Bava, F.-A.; Méndez, R. Translational Control by Changes in Poly(A) Tail Length: Recycling MRNAs. *Nat. Struct. Mol. Biol.* **2012**, *19*(6), 577–585. https://doi.org/10.1038/nsmb.2311
8. Gallie, D. R. The Cap and Poly(A) Tail Function Synergistically to Regulate MRNA Translational Efficiency. *Genes Dev.* **1991**, *5*(11), 2108–2116. https://doi.org/10.1101/gad.5.11.2108
9. Goldstrohm, A. C.; Wickens, M. Multifunctional Deadenylase Complexes Diversify MRNA Control. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*(4), 337–344. https://doi.org/10.1038/nrm2370
10. Garneau, N. L.; Wilusz, J.; Wilusz, C. J. The Highways and Byways of MRNA Decay. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*(2), 113–126. https://doi.org/10.1038/nrm2104
11. Park, J.-E.; Yi, H.; Kim, Y.; Chang, H.; Kim, V. N. Regulation of Poly(A) Tail and Translation during the Somatic Cell Cycle. *Mol. Cell* **2016**, *62*(3), 462–471. https://doi.org/10.1016/j.molcel.2016.04.007.
12. Chang, H.; Lim, J.; Ha, M.; Kim, V. N. TAIL-Seq: Genome-Wide Determination of Poly(A) Tail Length and 3' End Modifications. *Mol. Cell* **2014**, *53*(6), 1044–1052. https://doi.org/10.1016/j.molcel.2014.02.007
13. Roussis, S. G.; Pearce, M.; Rentel, C. Small Alkyl Amines as Ion-Pair Reagents for the Separation of Positional Isomers of Impurities in Phosphate Diester Oligonucleotides. *J. Chromatogr. A* **2019**, *1594*, 105–111. https://doi.org/10.1016/j.chroma.2019.02.026
14. Senko, M. W.; Beu, S. C.; McLaffertycor, F. W. Determination of Monoisotopic Masses and Ion Populations for Large Biomolecules from Resolved Isotopic Distributions. *J. Am. Soc. Mass Spectrom.* **1995**, *6*(4), 229–233. https://doi.org/10.1016/1044-0305(95)00017-8

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

General Laboratory Equipment – Not For Diagnostic Procedures. © 2022 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. 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. **AN001183-na-en 0823**

thermo scientific