Waters™

Nota de aplicación

Ion-Pair Reversed-Phase Liquid Chromatography Method for Analysis of mRNA Poly(A) Tail Heterogeneity

Martin Gilar, Maissa M. Gaye

Waters Corporation

Abstract

Messenger RNA (mRNA) biopolymers are new therapeutic modalities. Development of effective mRNA vaccines against SARS-CoV-2 virus accelerated the research of novel mRNA therapies. Promising mRNA applications include protein replacement therapy and vaccination approaches for cancer treatment. Rapid progress in mRNA technology relies on robust analytical methods for product testing. This application note describes the analysis of 3′ poly(A) tail modifications using efficient Ion-Pair Reversed-Phase Liquid Chromatography (IP RP LC) combined with ultra-violet (UV) detection. The method relies on the digestion of an mRNA molecule with RNase T1 to release the poly(A) tail and an optimized ion-pairing mobile phase system for separation of poly(A) tail length variants. The method is suitable for in-depth analysis of poly(A) tail heterogeneity.

Benefits

- Robust analysis of poly(A) tail heterogeneity
- LC UV method suitable for quality control testing
- n/n-1 resolution of poly(A) tail variants up to 150 long oligonucleotides

Introduction

Therapeutic messenger mRNA molecules are single stranded nucleic acids, typically 2–10 thousand nucleotides long. Therapeutic mRNA have an anatomy that consists of several distinct parts, including a modified 5′ cap, 3′ and 5′ untranslated regions (UTRs), the mRNA coding region or gene transcript sequence, and a poly(A) tail on the 3′ end (see Figure 1). The 5′ end cap minimizes RNA degradation while facilitating the translation process by latching on to ribosomes in the cell. The 3′ and 5′ untranslated regions directly impact translation and are often sequence optimized to achieve maximum RNA translation efficiency. The mRNA coding region encodes for the gene of interest (GOI). This is the part that gets translated into the desired protein, and it often contains modified nucleotides (typically N1-methylpseudouridine) to reduce rates of clearance and increase protein production. Finally, there is the poly(A) tail on the 3′ end, which plays a critical role in minimizing RNA degradation and maintaining *in-vivo* stability of the mRNA. The poly(A) tail portion in mRNA vaccines is typically 100–150 nucleotides (nt) long.

The success of the SARS-CoV-2 vaccines has accelerated research activities on mRNA therapies. Number of mRNA clinical trials are in progress. The research, development and manufacturing of mRNA therapies requires robust analytical methods. Ion-Pair Reversed-Phase Liquid Chromatography (IP RP LC) is a method for choice for analysis of oligonucleotides.[1–5] The n/n-1 resolution is achievable up to approximately 60 nt long species and becomes progressively difficult for longer oligonucleotides and mRNA. The analysis of poly(A) tail heterogeneity of mRNA is a difficult task even for state-of-the-art separation methods and mass spectrometry (MS). The intact mRNA molecules are often digested prior to analysis with selective nucleases such as RNase T1. RNase T1 cleaves the RNA sequence at the 3′ terminus of G nucleotides[6] and produces a mixture of relatively short oligonucleotides that could be analyzed by LC MS/MS.[1,5,7] The poly(A) tail portion of mRNA does not contain G nucleotides to be cleaved. In turn, the 3′ poly(A) tail is liberated from mRNA as an 100–150 nt long oligonucleotide (see Figure 1). Separation of long oligonucleotides differing by a single nucleotide in length presents a challenge for polyacrylamide gel electrophoresis[8], capillary gel electrophoresis[9, 10] or chromatographic methods.[8,9,10,11,13]

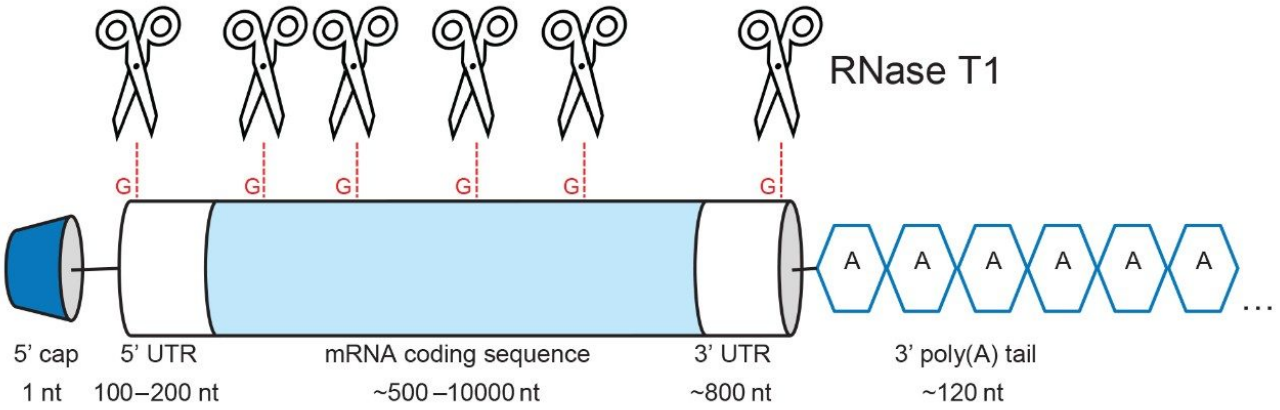

Figure 1. Schematic representation of mRNA structure. Scissors indicate putative cleavage sites of RNase T1. The digest produces short RNA oligonucleotides and liberates the 3'-end poly(A) tail.

In this application note, we describe a high-resolution IP RP LC method for analysis of poly(A) tail heterogeneity. The poly(A) tail liberated with RNase T1 is resolved by IP RP LC into oligoribonucleotide adenosines. The IP RP LC method n/n-1 resolution was achieved up to ~150 nt oligonucleotides. This method is useful for analysis of poly(A) tail heterogeneity, and identification of the dominant oligonucleotide species length in the poly(A) tail. Alternatively, a simple size exclusion chromatography method with ultraviolet detection (SEC UV) can be used to estimate an average poly(A) tail length. The SEC UV method is described in a separate application note 720007853.

Experimental

Sample Preparation

EPO mRNA standard was obtained from TriLink Biotechnologies, part number L-7209. EPO mRNA contains Cap 1 5′ structure, it is polyadenylated, and contains modified 5-methoxyuridine throughout the sequence. Fluc-beta mRNA sample was obtained from AmpTec, part number M1436/1000-C1-A120-NM-P0. The mRNA is polyadenylated, and contains unmodified uridines in the sequence. 1 mg/mL solution of mRNA samples were digested with RNase T1 (Thermo Fisher, part number EN0542, 1000 U/µL) using the following protocol: 50 µL of 1

mg/mL mRNA was mixed with 10 µL of rCutSmart 10x buffer (New England Biolabs, part number B6004S) and 2 µL of RNase T1 was added. The sample was digested for 30 minutes at 37 °C. 1 µL of Quick CIP enzyme (New England Biolabs, part number M0525S, 5000 µ/mL) was added to the sample after the digestion to remove 3′ phosphate groups from the oligonucleotides. Dephosphorylation was performed for 30 minutes at room temperature. The sample was spiked with 20 µL of 1M TEAA (Sigma, part number 90358) and 137 µL of RNase free water (Sigma, part number W4502). The final volume of ~ 200 µL of sample was cleaned up using Oasis™ HLB 96-well µElution Plate (Waters™, part number 186001828BA) using the following protocol: Wash the plate with 50 µL of acetonitrile followed with 100 µL of 0.1 M TEAA solution. Load the sample using ~0.5 mL/minute flow rate. Collect the sample and re-load for the second time. Elute the sample from the plate with 40 µL of 30% acetonitrile solution. Inject 1–5 µL of the eluent on Ion-Pair Reverse-Phase LC column for analysis. 100 nt oligoribonucleotide adenosine was ordered as custom oligonucleotide from IDT DNA (Coralville, IA, USA).

IP RP LC Conditions

LC system:	ACQUITY™ Premier UPLC™ (with SM-FTN and QSM)
Detection:	PDA, Ti 5 µL cell, 260 nm
Column:	ACQUITY Premier Oligonucleotide BEH C_{18} 300 Å, 2.1 x 150 mm, 1.7 µm, (p/n: 186010541)
Column temperature:	60 °C
Sample temperature:	10 °C
Injection volume:	1.0 µL (sample)
Flow rate:	0.3 mL/min
Mobile phase A :	100 mM octylammonium acetate (OAA) in 40% acetonitrile and 1% hexafluoroisopropanol (HFIP), v:v

Mobile phase B:	100 mM OAA in 90% acetonitrile and 1% HFIP
Mobile phase D:	100% acetonitrile
Gradient:	From 46% A, 54% B, and 0% D to 28 % A, 62% B, and 10% D in 40 min
Chromatography software:	Empower™ v 3.0

Mobile phase A preparation: mix 116.13 g of water with 60.86 g of acetonitrile; add 1.145 mL of glacial acetic acid, 3.305 mL of octylamine, and 2 mL of HFIP. Mix well.
Mobile phase B preparation: mix 19.355 g of water with 136.92 g of acetonitrile; add 1.145 mL of glacial acetic acid, 3.305 mL of octylamine, and 2 mL of HFIP. Mix well.

Results and Discussion

IP RP LC Method Development Strategies

IP RP LC separation of oligonucleotides utilizes ion-pairing mobile phase additives.[1,2,14] IP additives, such as alkylamines, are adsorbed on a hydrophobic RP sorbent, and promote ionic interaction of positively charged alkylamines with negatively charged phosphate groups of oligonucleotides. While IP RP LC can easily resolve short 15–30 nt oligonucleotides, the separation task becomes more difficult for long 40–60 nt and very difficult for 60–100 nt oligonucleotides.[13] This is because the 10 nt vs 11 nt oligonucleotides differ 10% in charge (9 or 10 phosphate groups, respectively), while 100 versus 101 nt species differs only in 1%, which reduces the 100/101 nt separation selectivity. The following were the strategies explored for resolution optimization of long oligonucleotides. (i) Use efficient hydrophobic alkylamines (hexylamine, octylamine).[1,14] (ii) Use a high concentration of IP agent in the mobile phase (e.g. 50–100 mM). (iii) Use efficient columns (long columns packed with sub two micron particles). (iv) Use shallow gradients, such as 0.1–0.5% acetonitrile per minute.[2] We combined the use of an ACQUITY Premier Oligonucleotide BEH C_{18} 300 Å, 2.1 x 150 mm, 1.7 µm, Ccolumn with a 100 mM octylammonium acetate IP mobile phase, and 0.25% acetonitrile per minute gradient to achieve an

optimal resolution for poly(A) tail oligonucleotides up to ~150 nt long. Concentrations of OAA higher than 100 mM are not recommended, because the required >80% acetonitrile concentrations used for elution may result in on-column nucleic acid precipitation. Some commercial mRNA samples contain constituents detrimental to IP RP LC column performance. Sample cleanup with Oasis HLB 96-well µElution Plate is recommended prior to IP RP LC analysis.

Application of IP RP LC Method to Poly(A) Tail Analysis

Figure 2 shows the optimized IP RP LC method was applied to analysis of RNase T1 digested EPO mRNA (TriLink) and Fluc Beta mRNA (AmpTech). Short 2–30 nt oligonucleotides produced by the digest eluted before five minutes (data are not shown). Figure 2 highlights the poly(A) species eluting between 12 to 35 minutes. The group of peaks amidst 124 nt length (blue chromatogram) represents analysis of the EPO mRNA poly(A) tail. The red chromatogram reveals that Fluc Beta mRNA poly(A) tail is somewhat longer, grouped around 128 nt peak.

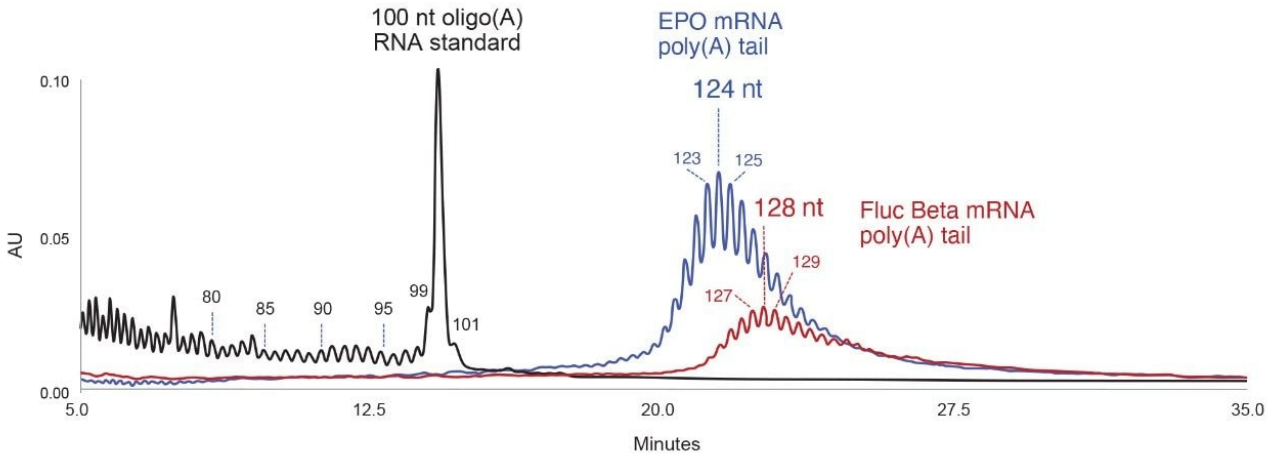

Figure 2. Separation of 100 nt oligo(A) synthetic RNA oligonucleotide standard (black chromatogram), RNase T1 digested EPO mRNA (blue chromatogram) and RNase T1 digested Fluc Beta mRNA (red chromatogram) with an IP RP LC method using an ACQUITY Premier Oligonucleotide BEH 300 Å 1.7 µm Column. The most abundant peak amidst the EPO mRNA poly(A) tail series of peaks corresponds to a 124 nt long species. The most abundant peak in the Fluc Beta mRNA poly(A) tail grouping of peaks corresponds to a 128 nt long species.

The developed IP RP LC method is not strictly compatible with MS detection. The length of the most abundant poly(A) oligonucleotide was estimated from the retention time using 100 nt oligoribonucleotide adenosine

standard (black chromatogram, Figure 2). Synthetic 100 nt oligoribonucleotide(A) contains shorter n-x impurities resolved from the full-length product. We plotted the linear trend L(nt) = $a \times t_r + b$ for 80, 85, 90, 95, 99, 100, and 101 nt species labeled in the Figure 2. The linear calibration plot L(nt) vs retention time t_r is shown in Figure 3. This trend was used to estimate the length L(nt) of the dominant poly(A) tail peaks found in the EPO mRNA and Fluc Beta mRNA digests (compare Figure 2 and Figure 3). Because the RNase T1 digest cleaves after G position, it is possible that 5′ end of liberated poly(A) oligonucleotide contains one or several additional C, A, or U nucleotides. This was the case for Fluc Beta RNA poly(A) tail which contains an additional C nucleotide. The full sequence of EPO mRNA is not provided by TriLink manufacturer. The mRNA sequence must be considered for an accurate length assignment of the poly(A) tail species length. When needed, RNase T1 digestion can also be complemented with RNase A or other specific RNases.

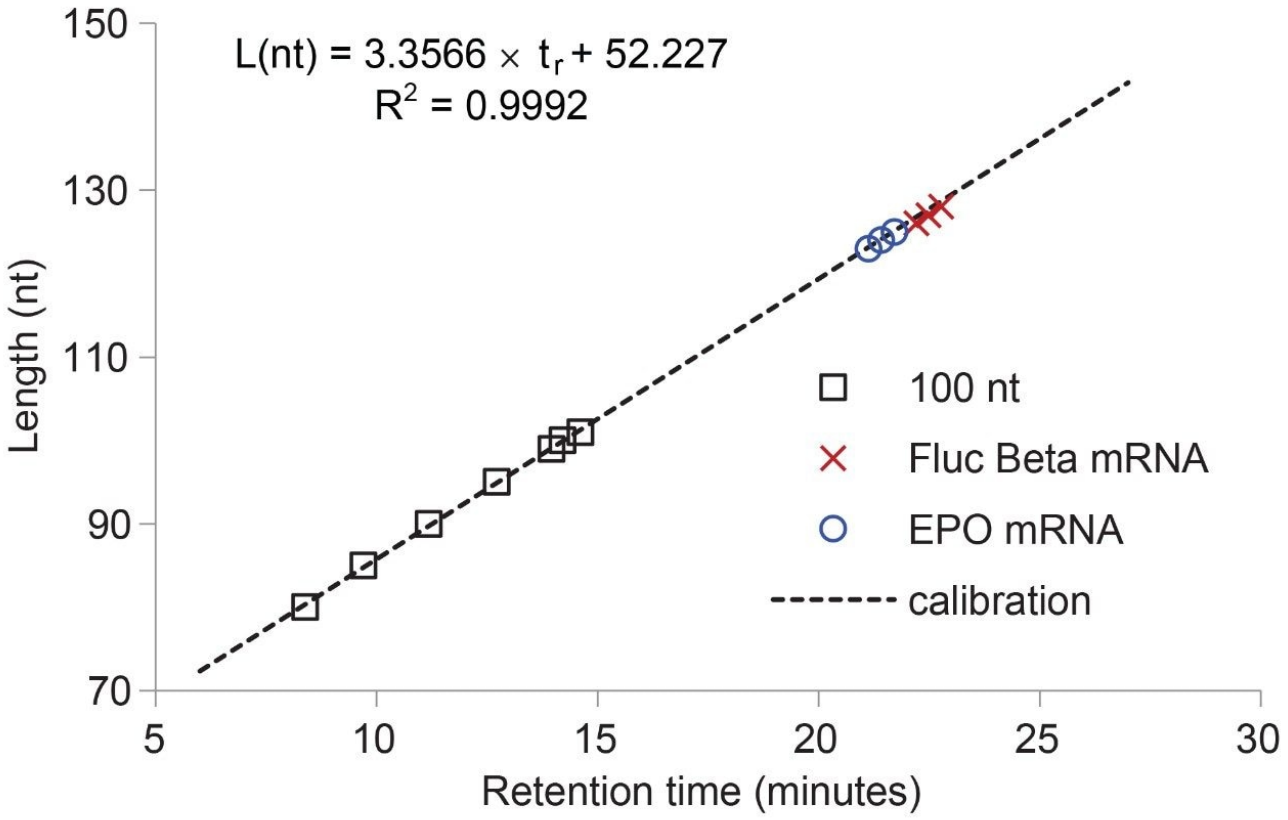

Figure 3. Calibration curve L(nt) = a × t_r + b was constructed using 100 nt synthetic RNA adenosine standard and its truncated synthetic sequences (black chromatogram in Figure 2; 80, 85, 90, 95, 99, 100, and 101 nt peaks). The calibration was used to calculate the length of the most dominant poly(A) tail peaks in mRNA samples. The dominant poly(A) tail peaks were 124 nt and 128 nt for the EPO mRNA and Fluc Beta mRNA digests, respectively. For the clarity, only the three most abundant peaks detected in Figure 2 were plotted in the Figure 3.

To confirm the assignment of the most abundant species in these two polyA tail samples, we separately performed an LC-MS experiment with diisopropylethylamine/HFIP ion-pairing mobile phase or a SEC measurement of poly(A) tail described in the application note 720007853. All three methods (IP RP LC UV, LC-MS, and SEC) provided a consistent length measurement of the most abundant poly(A) tail peak, *i.e.* 124 nt ± 1 nt and 128 nt ± 1 nt for for EPO mRNA and Fluc Beta mRNA digests, respectively.

Conclusion

An IP RP LC-UV method was developed for the analysis of mRNA poly(A) tail heterogeneity. The optimized method uses a wide pore ACQUITY Premier Oligonucleotide BEH C_{18} 300 Å, 150 x 2.1 mm, 1.7 mm Column and an efficient octylammonium acetate ion-pairing buffer. The resolution of long oligonucleotides up to 150 nt is sufficient to study the length and heterogeneity of poly(A) tail samples. The developed method uses simple and robust UV detection that is amenable to use as a quality control assay. The length of the dominant poly(A) tail oligonucleotides was assigned by calibration with synthetic oligoribonucleotide(A) standards. An ACQUITY Premier UPLC Ssystem and a column made with MaxPeak™ High Performance Surfaces were employed to ensure high, reproducible recoveries of the oligonucleotide analytes .[4,15–17]

References

1. M. Donegan, J.M. Nguyen, and M. Gilar, *J. Chromatogr. A*, 1666, 462860 (2022). DOI: 10.1016/j.chroma.2022.462860.
2. M. Gilar, K.J. Fountain, Y. Budman, U.D. Neue, K.R. Yardley, P.D. Rainville, R.J. Russell, 2nd, and J.C. Gebler, *J. Chromatogr. A*, 958, 167–182 (2002).
3. S.G. Roussis, M. Pearce, and C. Rentel, *J. Chromatogr. A*, 1594, 105–111 (2019). DOI: 10.1016/j.chroma.2019.02.026.
4. G.J. Guimaraes, J.M. Sutton, M. Gilar, M. Donegan, and M.G. Bartlett, *J. Pharm. Biomed. Anal.*, 208, 114439 (2022). DOI: 10.1016/j.jpba.2021.114439.
5. A. Goyon, P. Yehl, and K. Zhang, *J. Pharm. Biomed. Anal.*, 182, 113105 (2020). DOI: 10.1016/j.jpba.2020.113105.
6. T. Jiang, N. Yu, J. Kim, J.R. Murgo, M. Kissai, K. Ravichandran, E.J. Miracco, V. Presnyak, and S. Hua, *Anal. Chem.*, 91, 8500–8506 (2019). DOI: 10.1021/acs.analchem.9b01664.
7. V.B. Ivleva, Y.Q. Yu, and M. Gilar, *Rapid Commun Mass Spectrom*, 24, 2631–2640 (2010). DOI: 10.1002/rcm.4683.

8. L.T. França, E. Carrilho, and T.B. Kist, *Q. Rev. Biophys.*, 35, 169–200 (2002). DOI: 10.1017/s0033583502003797.

9. E. Carrilho, M.C. Ruiz-Martinez, J. Berka, I. Smirnov, W. Goetzinger, A.W. Miller, D. Brady, and B.L. Karger, *Anal. Chem.*, 68, 3305–3313. (1996).

10. M. Gilar, A. Belenky, Y. Budman, D.L. Smisek, and A.S. Cohen, *J Chromatogr B Biomed Sci Appl*, 714, 13–20 (1998).

11. J.R. Thayer, *Thermo Scientific Application note 21996*, (2014).

12. C.G. Huber, E. Stimpf, P.J. Oefner, and G.K. Bonn, *LC-GC*, 14, 114–127 (1996).

13. M. Gilar and U.D. Neue, *J. Chromatogr. A*, 1169, 139–150 (2007). DOI: 10.1016/j.chroma.2007.09.005.

14. S.M. McCarthy, M. Gilar, and J. Gebler, *Anal Biochem*, 390, 181-188 (2009). DOI: 10.1016/j.ab.2009.03.042.

15. J.M. Nguyen, M. Gilar, B. Koshel, M. Donegan, J. MacLean, Z. Li, and M.A. Lauber, *Bioanalysis*, (2021). DOI: 10.4155/bio-2021–0115.

16. M. Gilar, M. DeLano, and F. Gritti, *J. Chromatogr. A*, 1650, 462247 (2021). DOI: 10.1016/j.chroma.2021.462247.

17. M. DeLano, T.H. Walter, M.A. Lauber, M. Gilar, M.C. Jung, J.M. Nguyen, C. Boissel, A.V. Patel, A. Bates-Harrison, and K.D. Wyndham, *Anal. Chem.*, 93, 5773–5781 (2021). DOI: 10.1021/acs.analchem.0c05203.

Featured Products

ACQUITY Premier System <https://www.waters.com/waters/nav.htm?cid=135077739>

Empower Chromatography Data System <https://www.waters.com/10190669>

720007873, February 2023