

mRNA疫苗中的poly A分析前处理策略

漆倩

赛默飞世尔科技（中国）有限公司

关键词

mRNA polyA Oligo (dT)25 Orbitrap 高分辨液相色谱质谱

1. 引言

真核生物mRNA通常在其 3' 末端以一段腺苷终止，根据物种的不同，其长度可以从20到200多个碱基不等。3'-多聚腺苷 (poly A) 尾通过与多聚腺苷酸结合蛋白 (PABP)、5' 帽和真核起始因子 (eIF) 相互作用，影响细胞质中的mRNA翻译和稳定性。在设计用于产生免疫刺激性抗原的治疗性或预防性mRNA药物 (mRNA 疫苗) 中，通过将多聚腺苷序列编码到模板 DNA 中或使用聚腺苷酶在体外转录 (IVT) 合成产生的mRNA 3' 端添加poly A尾，可影响蛋白质表达的持续时间和强度，从而影响mRNA疫苗的寿命和免疫原性。由上可知了解mRNA疫苗中的poly A尾长及其与mRNA稳定性和蛋白质表达的关系十分必要，因此，需要建立合适的方法用于准确表征IVT合成mRNA的poly A尾长。

目前有许多技术被用于polyA尾长的分析，例如色谱法 (poly dT或poly U 柱)、基于PCR的分析 (如LM-PAT和ePAT) 以及二代测序方法 (如TAIL-seq和PAL)。其中，二代测序方法 (例如TAIL-seq) 可报告单个碱基，具有较高分辨率。但该方法是一种间接检测的方法，需要通过扩增长链腺苷转化为cDNA后检测，因此检测过程非常复杂^[1]。与上述方法不同，质谱法 (MS) 是一种更加高效且直接的测量技术，能够通过其质量差异来区分单个核苷酸。质谱法已被用于准确测定长度可达 500 个碱基的 PCR 产物的碱基组成。在商业化的诊断平台中也可看到质谱法的应用，例如用质谱法鉴定 120 mer的 PCR 产物，可同时识别多种微生物病原体^[2]。其中高分辨质谱，尤其Orbitrap，以其超高的分辨率及高灵敏度的优势，能够获得远超前于单核苷酸分子量的分辨率，且可同时直接测量多个核酸的

混合物及可能发生的修饰，因此非常适用于poly A 尾长的表征。完整的mRNA疫苗序列通常由几千个碱基组成，质谱难以实现超大分子量物质的检测。通过一定的前处理步骤，获得仅含有3' poly A的短片段可极大提升MS检测mRNA的能力。例如，有文献报道使用 oligo DT磁珠富集纯化polyA，结合质谱检测的方法，能够准确检测poly A尾分布及其修饰^[3]。本文在此方法的基础上进一步优化了前处理策略，并与未经磁珠纯化分析的方法做了对比，为mRNA 疫苗中的poly A尾表征分析提供更多选项。

2. 实验部分

2.1 样品

某公司提供的新冠mRNA疫苗 (约3900mer)

2.2 仪器及试剂

质谱仪器: Q Exactive Focus (赛默飞世尔科技, 美国);

色谱仪器: Vanquish Flex 液相色谱系统 (赛默飞世尔科技, 美国);

色谱柱: DNAPac RP (2.1 x 100 mm, 4 μm, P/N 088923)

试剂: 二次去离子水, 甲醇 (质谱级), DIPEA (质谱级), 六氟异丙醇 (质谱级), EDTA, RNase-Free Distilled Water, Dynabead sTM mRNA 纯化试剂盒 (包括Oligo (dT)25磁珠、binding buffer和washing buffer, P/N 61006), RNase T1(1000 U/μL, P/N EN0541)。

其他: DynaMagTM-2 磁力架

2.3 仪器方法

1) IP-RPLC-MS;

表1 poly A 检测色谱条件

流动相	A相 (%) : 2%HFIP-0.1%DIPEA-H ₂ O	B相 (%) : 0.075%HFIP-0.0375%DIPEA-甲醇
流速 (ml/min)	0.3	
柱温 (°C)	60	
色谱梯度	时间 (min)	B相浓度 (%)
	0	20
	1	20
	15	40
	16.5	90
	18	90
	18.1	20
	20	20

表2 质谱条件

Polarity	Negative
Resolution	17,500
Scan ranges	600-3000 m/z
AGC target	3 e6
Maximum IT	200 ms
Spectrum data type	profile
Spray Voltage (-)	3300
Capillary Temperature	320 °C
Sheath Gas	40
Aux Gas	10
S-Lens RF Level	60
Probe Heater Temp.	350 °C

2.4 数据采集及分析

Thermo Xcalibur 进行数据采集; BioPharma finder 4.1进行数据处理

2.5 前处理

2.5.1 Dynabeads™ mRNA试剂盒纯化

取100 µg mRNA样品, 加入10 µL RNase H buffer和 5 µL RNase T1, 用无酶水稀释至100 µL, 样品置于37度孵育3 hr。取500 µL磁珠置于磁力架30 s, 去除上清; 取200 µL binding buffer重悬磁珠, 去除上清; 再加入100 µL binding buffer; 将酶切样品加入磁珠, 室温温和旋转5 min; 置于磁力架, 吸取上清并遗弃; 分别加入washing buffer 100 µL清洗磁珠2次; 分别加入100 µL 100 mM乙酸铵清洗磁珠2次; 加入100 µL 75%甲醇, 80度水浴1-2 min, 置于磁力架上冷却1-2 min, 取上清, 旋干; 加入50 µL 1 mM EDTA (10 mM Tris HCl pH7-8或水) 复溶, 上样。

2.5.2 酶解后直接进样方法

50 µg mRNA样品中加入10 µL RNase H buffer, 5 µL RNase T1, 加无酶水补足到100 µL, 37度分别孵育15 min、30 min、60 min、3 hr, 取20 µL酶解液 (加入1mM EDTA) 上样。

3. 结果与讨论

3.1 Dynabeads™ mRNA试剂盒纯化方法

mRNA疫苗通常由上千个碱基组成, 由于分子量太大而难以被质谱直接检测。因此对于mRNA加帽或者加尾分析, 通常需要增加前处理

步骤，通过探针纯化及特异性酶切的方法，获取含有5'端或3'端的短链RNA片段后进行质谱检测。对于加尾表征分析实验，通常步骤如下：首先，使用RNase T1将mRNA酶解成短片段（RNase T1酶切位点在G之后，且具有较高的特异性）；然后，加入Oligo DT磁珠探针，通过磁珠与末端poly A碱基互补配对结合捕获poly A片段；最后，在高温条件下用75%甲醇将poly A从磁珠上洗脱。具体实验步骤可参考Michael Beverly文献^[9]。本研究对文献方法进一步优化，基于文献中的磁珠用量很难得到相应的poly A产物峰（图1），本研究尝试将Oligo (dT)25磁珠用量从75 μ L逐渐增加到500 μ L，可以测得较高丰度的polyA产物峰，由于该磁珠回收率较低，建议500 μ L以上的磁珠用量对应100 pmol mRNA样品。

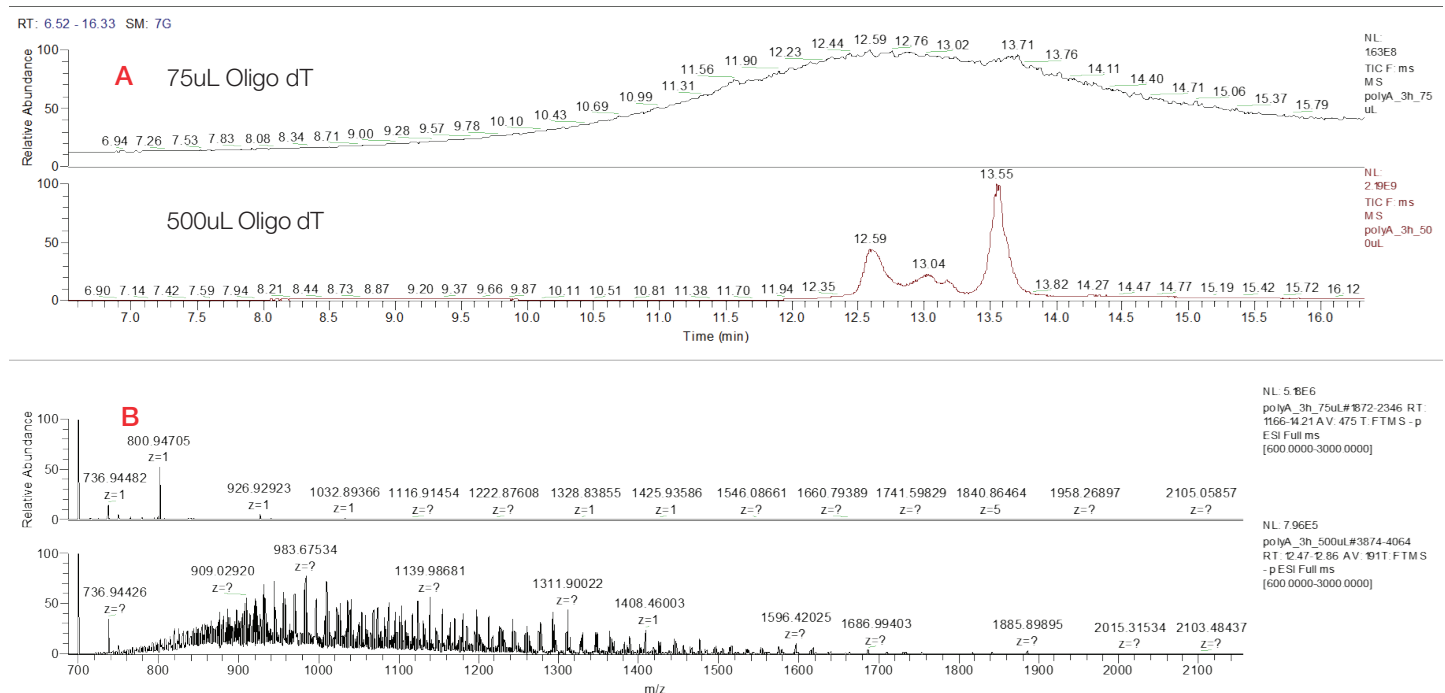


图1: mRNA样品分别经过75 μ L及500 μ L Oligo (dT)25磁珠纯化后, LC/MS检测图谱。A)TIC色谱图; B)质谱图。

图2为经500 μ L Oligo (dT)25磁珠纯化的LC/MS检测结果。通过解卷积TIC图上显示的三个峰（图2），获得相应的分子量。将获得的分子量与poly A理论分子量比对，确定peak1对应polyA样品。peak2和peak3异质性强，推测这两个峰为mRNA较大酶解片段，由于非特异性吸附导致。因此下一步我们增加了洗脱步骤，分别用washing buffer及100 mM乙酸铵清洗2-3次，将非特异性吸附峰降至最低（图3）。

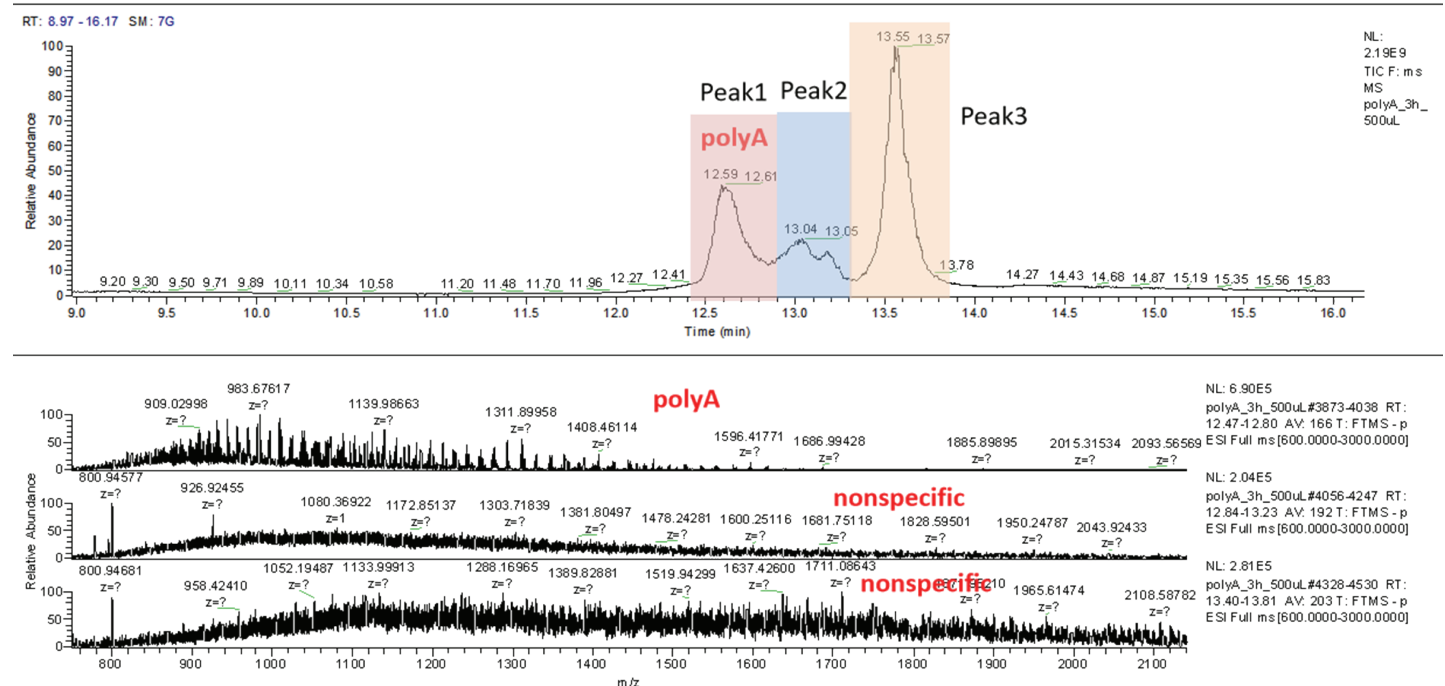


图2: mRNA样品经过500 μ L Oligo (dT)25磁珠纯化后, LC/MS检测图谱。

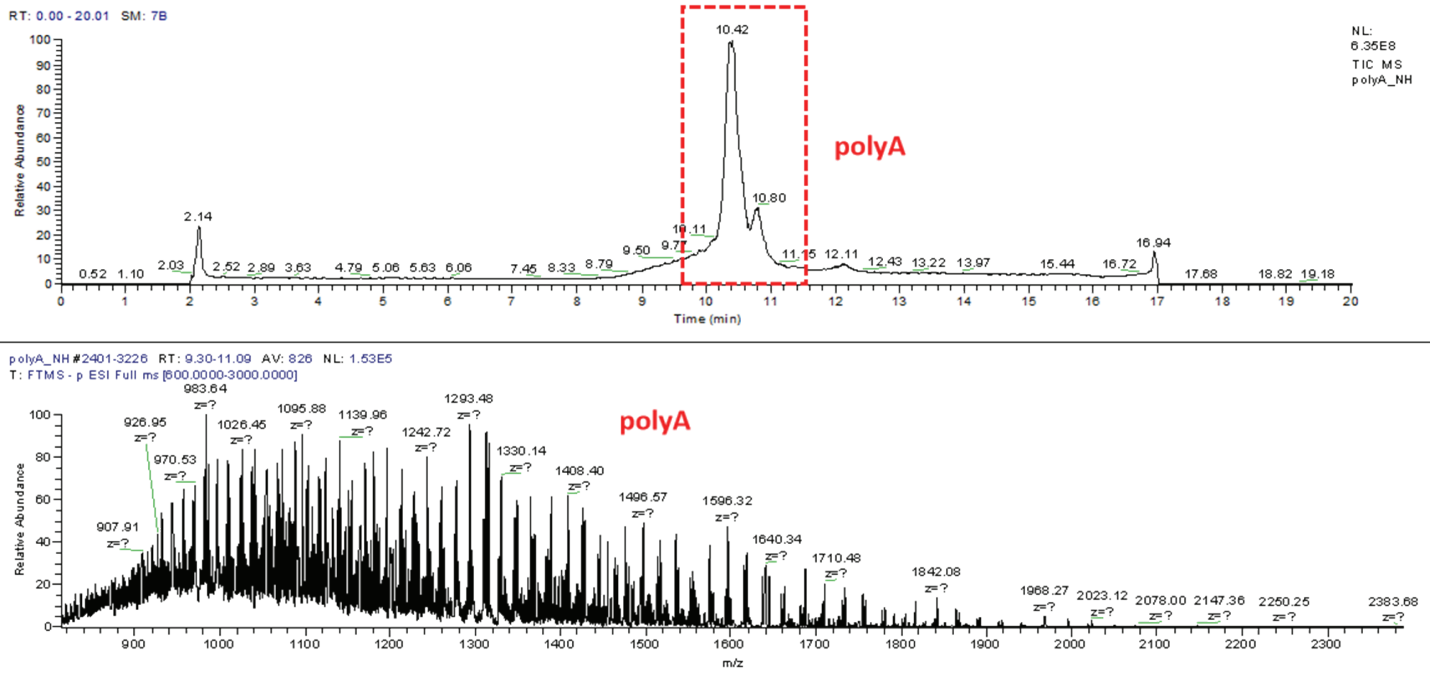


图3: mRNA样品经过500 μ L Oligo (dT)25磁珠纯化, 增加清洗步骤后的LC/MS检测图谱。

注: 更改了洗脱梯度, 因此图2和图3洗脱时间稍有不同

从试剂盒纯化方法步骤中可以看出, 磁珠的用量较大, 且前处理时间较长, 是否有更简便且可行的方法。基于此, 我们尝试直接使用RNase T1溶液酶解mRNA的方法 (无需磁珠和探针), 并将优化后的结果与试剂盒纯化方法对比。首先, 我们对比了不同酶解时间对poly A分布的影响。poly A长度通常在100 mer以上, 根据离子对反相色谱出峰规律, poly A通常最后被洗脱, 如图4红色框显示。随着酶解时间的增加, 从15 min增加到30 min, poly A前面出现了更多的峰, 说明酶解片段随酶解时间的增加而增加; 酶解时间从30 min增加到3 hr, TIC谱图没有明显变化。对四个图谱中保留时间在13 min的poly A峰分别解卷积, 得到poly A分布如图5。15 min酶解时间对应的分子量结果与其他条件不同, 推测由于未充分酶解导致; 30 min到3 hr的结果分子量分布少许不同, 但每个峰对应的分子量基本一致, 因此建议将酶解时间设置30 min以上。需要注意的是, RNase T1酶一直存在于酶解液中, 尚无有效方法快速终止该酶解反应, 因此poly A样品建议现配现做, 或者使用RNase T1 Mag Bulk Kit磁珠试剂盒 (货号: 60120-101) 将RNase T1与酶解液有效分离, 避免过度酶解。

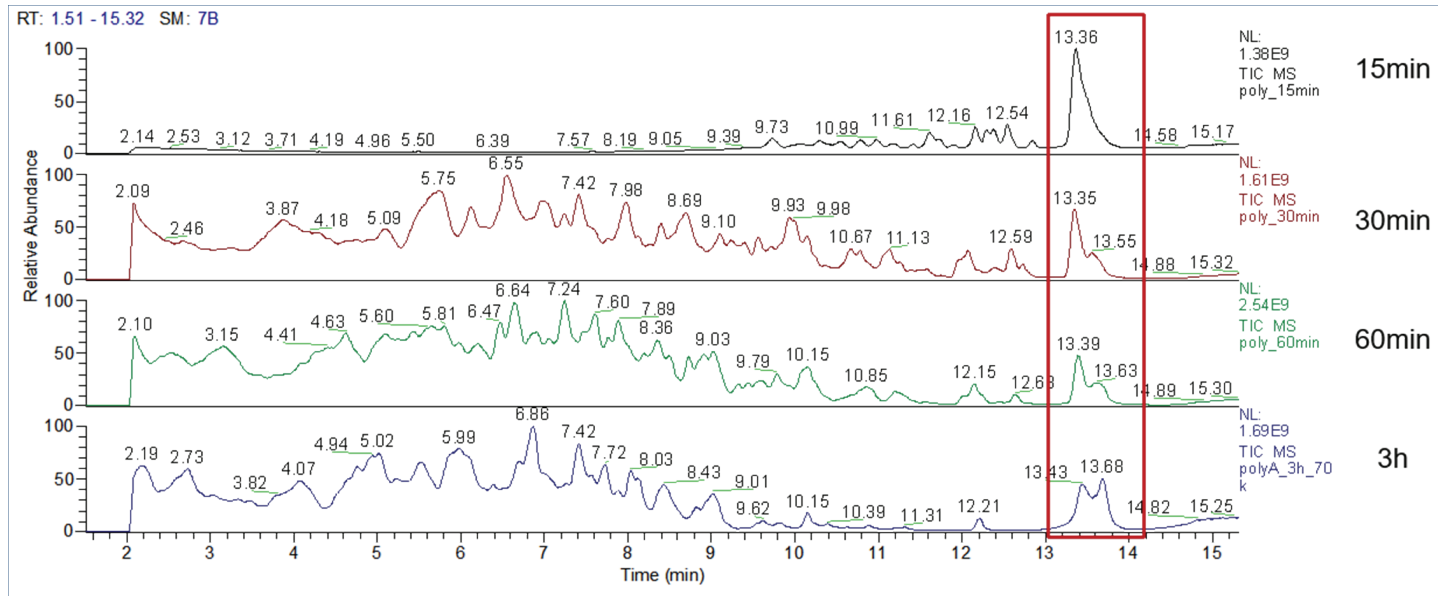


图4: mRNA样品经过RNase T1酶切 (酶解时间分别从15 min到3 hr) 后, 直接质谱检测所得TIC图谱。

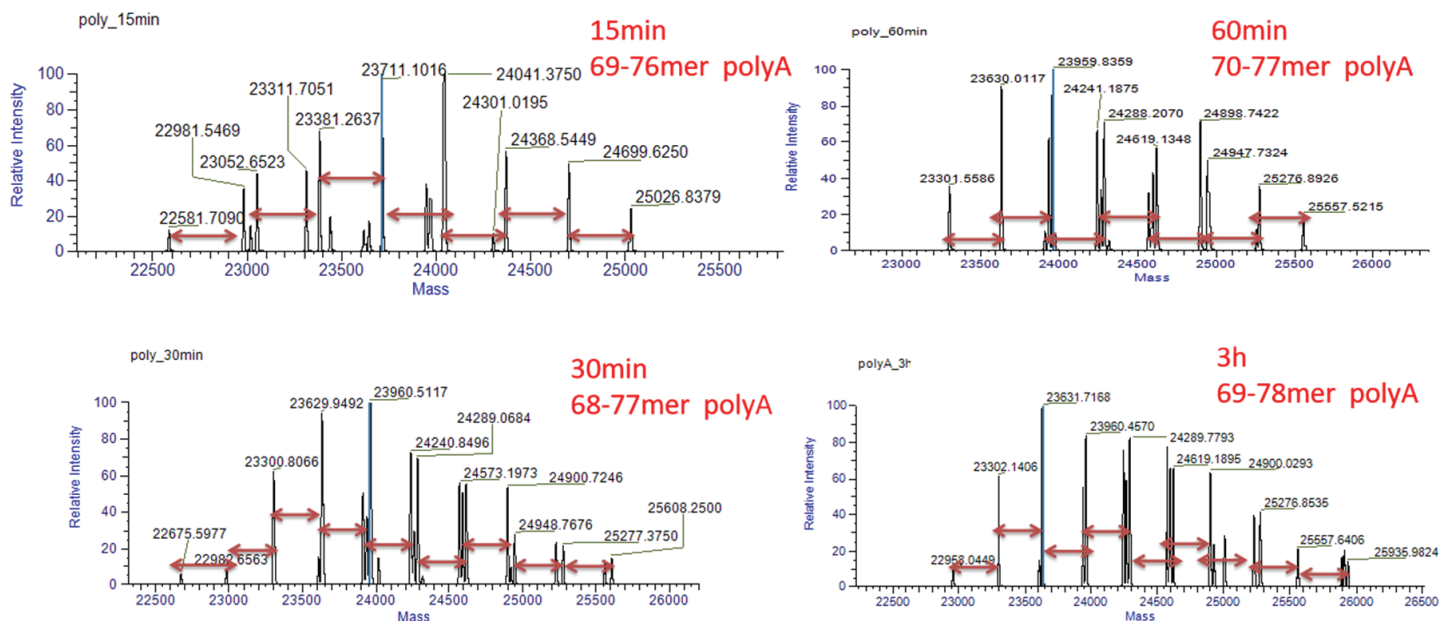


图5: 不同酶解时间对应的polyA峰解卷积结果。

将同一批mRNA样品分别用磁珠法和溶液酶解法处理30 min, 随后进样分析。解卷积分子量分布如图6所示, 磁珠法获得的样品由于纯度高、干扰峰少, 可检测到更多低丰度的poly A峰。因此, 磁珠法测得的poly A分布比溶液酶解法获得的poly A样品分布略宽。但从整体上看, 两种前处理方式得到的主峰分布相近, 且分子量准确度相当, 未经磁珠纯化的溶液酶解前处理方法可作为一种简单可行的选项用于mRNA样品的poly A分析前处理。

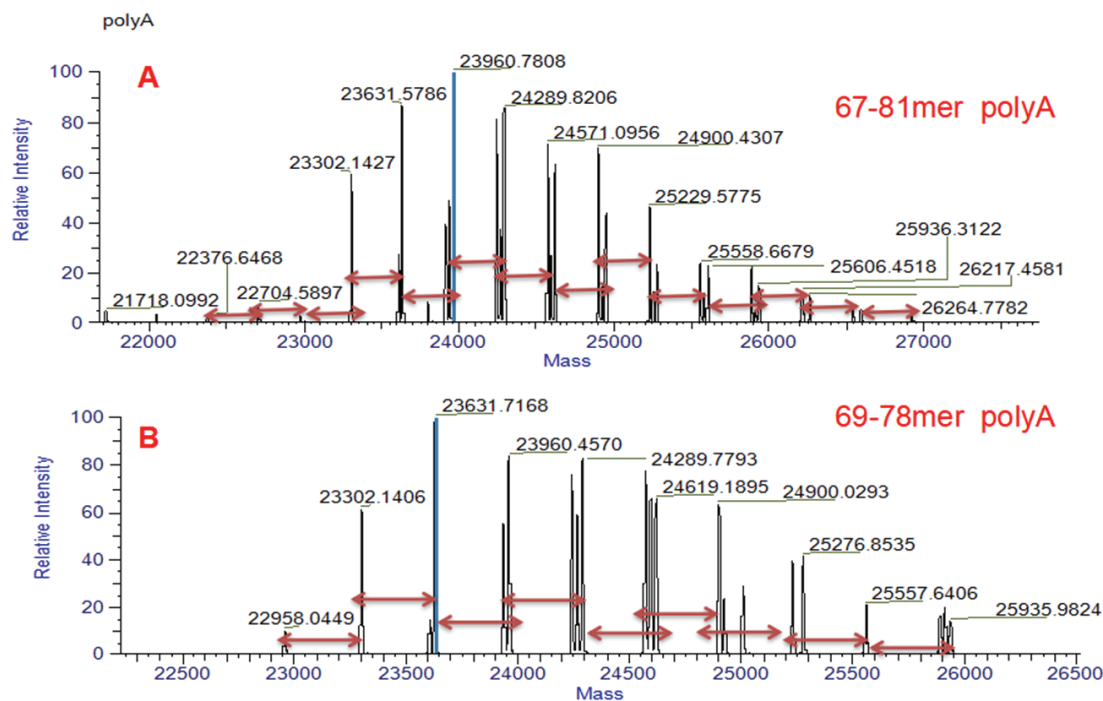


图6: poly A样品经过Biopharma Finder™ 4.1软件解卷积分子量分布图。

A为mRNA样品经过Oligo (dT)25磁珠纯化处理所得poly A分子量分布, B为相同mRNA样品经过溶液酶解30 min所得poly A分子量分布。

为了进一步验证溶液酶解法的可行性和准确性, 我们使用合成的mRNA样品进行测试, 酶解条件参考2.4.2, 酶解时间为1 hr。使用Biopharma Finder™ 4.1软件可轻松建立理论序列的建立简单且可设定修饰, polyA是由不同长度腺嘌呤组成的混合物, 软件中其理论序列的设置有两种方式: 1) 建立多个理论序列, 即将每个长度的polyA分别设置一个序列, 如120A, 121A, 122A等; 2) 设置一个理论序列, 建议取丰度最高组分对应的poly A分子量作为基准, 其他长度poly A通过设置为该组分的可变修饰来表示, 如 +A, +2A, -A, -2A等。可变修饰的添加如图7所示, 只需输入修饰的精确分子量和平均分子量信息及修饰位点, 即可添加新修饰至已有的修饰库中。常见的修饰如磷酸化、硫代磷酸化、甲基化等在软件自带的修饰库中均有, 且可以自定义修饰位点于5' 端、3' 端、碱基、backbone linker等任意位置。

The screenshot displays the Biopharma Finder 4.1 software interface. On the left, the 'Experiment Types' menu has 'Sequence Manager' highlighted. Below it, the 'Oligo Variable Modification' section shows a list of residues (A, A2, A3, A4, A5, A6, -A, -A2, -A3, -A4) and buttons for 'Add', 'Remove', and 'Load Default Mods'. A red arrow labeled 'Input Sequence' points from this area to the 'Sequence Information' panel on the right. This panel shows 'Target Oligonucleotide' with fields for Name (130A), Description, Sample Type (Oligonucleotide), Category (Intact Deconvolution), Monoisotopic Mass (42,714.8718), Average Mass (42,734.94), and Formula (C1300H1561O778N65C). Below this is the 'Oligo Sequence Map' showing a sequence of 121 residues, all labeled as 'pA'. A red arrow labeled 'Add Variable Modification' points from the 'Oligo Variable Modification' section to the 'Building Block and Variable Modification Editor' panel. This panel has two sub-sections: 'Oligo building block' and 'Oligo Variable Modification'. The latter is highlighted with a red box and shows fields for Subunit (Base), Name (A), Site (A), Monoisotopic Mass (329.053), Average Mass (329.21), and Formula.

图7: Biopharma Finder™ 4.1软件新建理论序列及添加新修饰流程图。

图8A展示了Biopharma Finder™ 4.1软件解卷积结果及原始质谱图，解卷积结果呈现出良好的正态分布，每个峰都匹配上了不同长度的poly A。解卷积结果中的每一个归属峰，在原始质谱图中均有多个电荷价态检出。例如123A对应的峰是质谱图中19个价态峰解卷积的结果（蓝线），可信度高。图8B展示了解卷积结果的放大图，图中所有的峰都与理论序列匹配上，呈现正态分布状，每个峰之间相差约329 Da，即一个腺嘌呤核苷分子量。峰之间呈现良好的基线分离，显示了Orbitrap质谱的高分辨及高灵敏度。

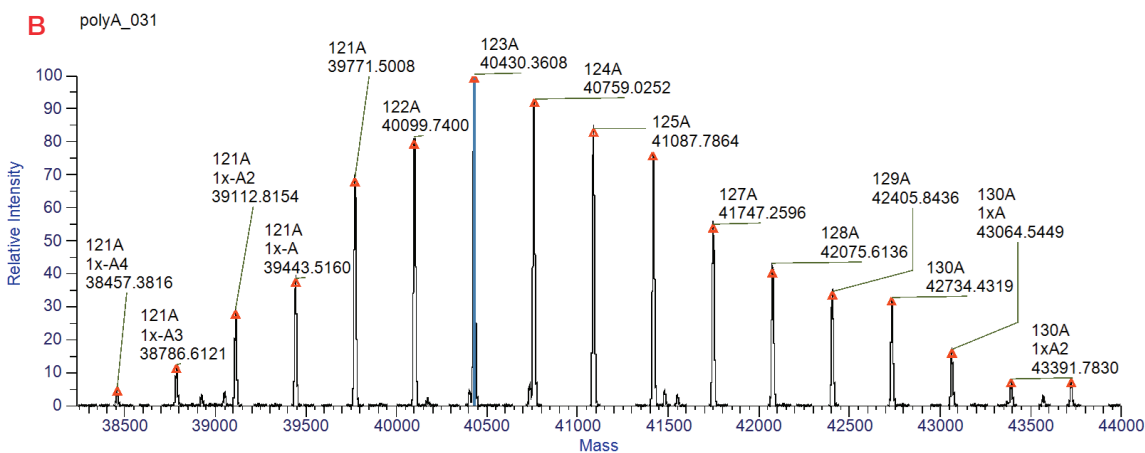
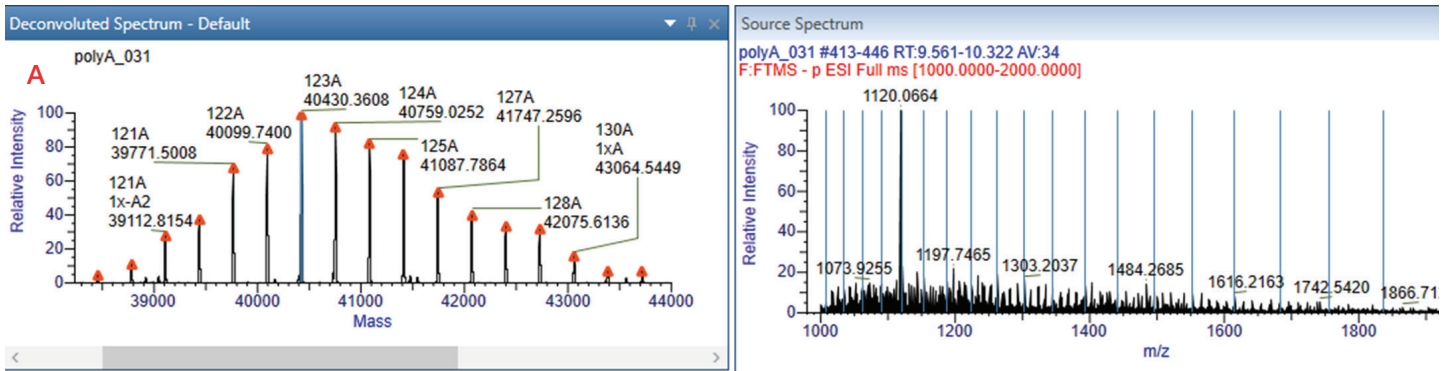


图8: poly A样品经Biopharma Finder™ 4.1软件解卷积的结果展示。

Biopharma Finder™ 4.1软件中解卷积的结果列表详细而全面（图9）。包括匹配的理论序列、修饰、实测分子量、理论分子量、质量偏差、质谱响应强度、每个组分的相对含量、score可信度分值、用于解卷积的价态数目、价态分布、质量偏差、组分的保留时间。结果列表显示匹配的序列分布从117A到133A，每个匹配组分的丰度可参考质谱强度和相对丰度值。质量偏差列表显示，即使分子量超过40,000 Da的组分，质量偏差基本都在1 Da以内，结果准确度高。匹配的电荷价态数量和分布可用作结果可靠性的参考值，所有的匹配组分均有近二十个电荷价态的匹配，显示了结果的高度可信。Apex IT为匹配的每个组分提供精确的保留时间值。

Results														
	Sequence Name	Modification	Average Mass	Theoretical Mass (Da)	Matched Mass Error (Da)	Sum Intensity	Relative Abundance	Fractional Abundance	Score	Number of Charge States	Charge State Distribution	Delta Mass	Apex RT	
	Aa	Aa	=	=	=	=	=	=	=	=	Aa	=	=	
1	123A		40430.3608	40430.4870	0.1262	1.39E+08	100.00	11.84	76.92	19	22 - 40	0.0000	10.12	
2	124A		40759.0252	40759.6939	0.6687	1.28E+08	92.15	10.91	88.79	20	21 - 40	328.6644	10.12	
3	125A		41087.7864	41088.9008	1.1144	1.16E+08	83.18	9.85	91.48	20	22 - 41	657.4256	10.12	
4	122A		40099.7400	40101.2800	1.5400	1.11E+08	79.64	9.43	93.16	20	21 - 40	-330.6208	10.12	
5	126A		41417.6025	41418.1077	0.5053	1.06E+08	76.22	9.02	87.33	21	21 - 41	987.2417	10.16	
6	121A		39771.5008	39772.0731	0.5724	9.51E+07	68.32	8.09	80.74	18	22 - 39	-658.8600	10.09	
7	127A		41747.2596	41747.3147	0.0550	7.57E+07	54.35	6.43	88.86	19	23 - 41	1316.8989	10.16	
8	128A		42075.6136	42076.5216	0.9080	5.67E+07	40.70	4.82	90.54	21	22 - 42	1645.2528	10.16	
9	121A	1x-A	39443.5160	39442.8631	0.6529	5.29E+07	37.98	4.50	77.96	20	20 - 39	-986.8447	10.12	
10	129A		42405.8436	42405.7285	0.1151	4.70E+07	33.72	3.99	76.22	18	25 - 42	1975.4828	10.19	
11	130A		42734.4319	42734.9354	0.5035	4.46E+07	32.02	3.79	81.27	18	25 - 42	2304.0711	10.19	
12	121A	1x-A2	39112.8154	39113.6531	0.8378	3.94E+07	28.31	3.35	83.90	19	21 - 39	-1317.5454	10.12	
13	130A	1xA	43064.5449	43064.1454	0.3995	2.25E+07	16.15	1.91	68.23	17	26 - 42	2634.1841	10.19	
14	121A	1x-A3	38786.6121	38784.4431	2.1690	1.63E+07	11.70	1.39	48.77	14	24 - 37	-1643.7487	10.09	
15	130A	1xA2	43391.7830	43393.3554	1.5724	1.04E+07	7.50	0.89	79.17	16	28 - 43	2961.4222	10.19	
16	130A	1xA3	43723.3891	43722.5654	0.8237	1.04E+07	7.50	0.89	61.90	16	28 - 43	3293.0283	10.23	

图9: poly A样品经Biopharma Finder™ 4.1软件分析的结果列表展示。

结论

综上所述，本文优化了mRNA 表征中的poly A 分析前处理策略，通过对比发现磁珠法和溶液酶解法均是可行的poly A前处理方式。磁珠法获得的样品分析结果相对更精确和完整，而溶液酶解法更经济便捷。此外，分析软件的多功能及操作简便对于poly A尾分析同样重要。Biopharma Finder™ 4.1核酸分析模块可以实现理论序列编辑、加帽加尾自动分析、核酸序列确证、修饰位点及含量分析和mRNA mapping等，是目前商品化软件中对于核酸分析功能最全面的软件。通过合成的mRNA样品验证，得益于Orbitrap的高分辨及高灵敏度等优势，可实现poly A尾完整且精确的表征，解卷积谱图里的每个峰都能实现良好的基线分离，且质量偏差小、结果可信度高。

参考文献

- [1] Subtelny AO, Eichhorn SW, Chen GR, Sive H, Bartel DP. Poly(A)-tail profiling reveals an embryonic switch in translational control. Nature. 2014;508(7494):66–71.
- [2] Wolk DM, Kaleta EJ, Wysocki VH. PCR-electrospray ionization mass spectrometry: the potential to change infectious disease diagnostics in clinical and public health laboratories. J Mol Diagn. 2012;14(4):295–304.
- [3] Beverly M, et al. Poly A tail length analysis of in vitro transcribed mRNA by LC-MS. Analytical and Bioanalytical Chemistry. 2018.



赛默飞
官方微信

热线 800 810 5118
电话 400 650 5118
www.thermofisher.com

Thermo Fisher
SCIENTIFIC