

使用高分辨率 LC/MS 快速分析 mRNA 5' 端加帽

前言

由于 SARS-CoV-2 疫情的持续蔓延,mRNA 的大规模工业化生产受到越来越多的关注。mRNA 疫苗已被证明是对抗病毒最有效的工具之一,其依赖体外转录合成高质量的 mRNA^[1]。mRNA 有效翻译成蛋白的关键取决于名为加帽的 5' 端二核苷酸修饰,该修饰在转录过程中^[2] 或转录之后^[3] 添加到 mRNA 上。重要的是,成功加帽材料的百分比和加帽结构的类型受起始材料的质量、反应条件和 mRNA 序列等因素的影响。因此,mRNA 5' 端加帽是一个关键质量属性,应对其进行全面表征和监测。

本研究使用 Agilent 6545XT AdvanceBio LC/Q-TOF 进行快速 mRNA 加帽分析,样品前处理和分析时间仅为 75 分钟。使用热稳定性酶加速样品前处理,以在更高的反应 温度下释放含有加帽结构的 5' 端寡核苷酸(图 1)。然后使用 AdvanceBio 寡核苷酸 色谱柱,仅通过简单的纯化流程将加帽的寡核苷酸从样品基质中分离出来。

为证明这种方法的实用性,我们优化了一种用于挑战性 mRNA 序列的牛痘病毒酶加 帽反应。我们的结果展示了高分辨率 LC/MS 作为灵敏、高效的方法,在核酸疗法生 产过程优化和质量控制方面的价值。

本研究中使用的缩写

- DNA: 脱氧核糖核酸
- PCR: 聚合酶链式反应
- mRNA: 信使核糖核酸
- ARCA: 抗反向帽类似物
- UTR: 非翻译区
- GTP: 三磷酸鸟苷
- NTP: 三磷酸核苷
- SAM: S-腺苷甲硫氨酸

作者

Brian Liau 安捷伦科技有限公司

实验部分

所有 PCR 引物和 RNA/DNA 嵌合探针均由 Integrated DNA Technologies 公司合成。

体外转录和 mRNA 加帽

编码 T7 启动子下游约 3800 nt 序列模型的质粒购自 Sino Biological 公司。该序列包括 5' UTR 和 3' UTR,以及一个编码区。使用 Herculase II Fusion DNA 聚合酶试剂盒(安捷伦部件号 600677)对 T7 启动子和目标序列进行 PCR 扩增,并使用 StrataPrep PCR 纯化试剂盒(安捷伦部件号 400771)进行纯化。在 Agilent 2100 生物分析仪上使用 DNA 7500 试剂盒(部件号 5067-1506)检测 PCR 反应的特异性和产物浓度。

然后使用 T7 RNA 聚合酶 (New England Biolabs M0251) 按照 制造商推荐的方案将 PCR 产物转录为 mRNA。使用 100 µL 不含核酸酶的水稀释 100 µL 体外转录反应产物,然后使用 DNA 酶 I (New England Biolabs M0303) 消化,以去除 DNA 模板和残留的 PCR 引物。然后加入 70 µL 8 mol/L LiCl 沉淀 mRNA,在 -20 °C 下冷却过夜,然后在 4 °C 下以 12000x g 离心 15 分钟。mRNA 团块在 70% 乙醇中洗涤两次,风干, 然后溶于不含核酸酶的水中,并在 Agilent 2100 生物分析仪上 使用 RNA 6000 Nano 试剂盒 (部件号 5064-1511) 进行定量。 然后将 mRNA 分装为 5 µg 的等分试样,于 -80 °C 下冷冻。

使用两种方法对 mRNA 进行加帽:

- 1) 使用 ARCA 进行共转录加帽
- 2) 使用牛痘病毒加帽酶进行酶促加帽

进行 ARCA 加帽时,用含有 ARCA 与 GTP(比例为 4:1)的 NTP 混合物转录产生 mRNA 样品。进行酶促加帽时,使用 牛痘病毒加帽系统试剂盒 (New England Biolabs M2080)对 5-10 μg 纯化的未加帽 mRNA 进行加帽,然后使用 Monarch RNA 纯化试剂盒 (New England Biolabs T2040) 纯化,并用不 含核酸酶的水洗脱,待分析。

定点 RNA 酶 H 裂解

如 Lapham J. 等人^[4] 所述,进行溶液中 RNA 酶 H 定点裂解。 将 15 nt 的 2'-O-甲基化 RNA/DNA 嵌合探针设计为与 mRNA 样 品互补(图 1),并用于在以下两种条件下指导 RNA 酶 H 裂 解: (1) 1 小时热退火程序,然后在 37 ℃ 下用 RNA 酶 H 裂解 30 分钟,或 (2) 在 50 °C 下用热稳定性 RNA 酶 H 裂解 30 分钟。裂解后,释放含有 5' 加帽结构、长为 40 nt 或 50 nt 的寡 核苷酸,然后使用 Monarch RNA 纯化试剂盒 (New England Biolabs T2040) 纯化消化样品,然后使用小 RNA 试剂盒 (部件 号 5067-1548) 在 Agilent 2100 生物分析仪上进行可视化分析。

5' 端加帽寡核苷酸的 LC-DAD/MS 分析 仪器

- 一 配备二极管阵列检测器的 1290 Infinity II 液相色谱仪(部 件号 G7117B)
- 6545XT AdvanceBio LC/Q-TOF

注意去除流路中的玻璃,减少碱金属加合效应。将安捷伦 Nalgene 瓶(部件号 9301-6460)用作流动相容器,每个溶剂 管线配备一个不锈钢滤芯。使用安捷伦聚丙烯样品瓶(部件 号 5190-2242)。在每天实验开始之前,用 50% MeOH + 0.1% 甲酸将液相色谱系统和色谱柱冲洗 30 分钟,进一步减少碱金 属加合物。

在 AdvanceBio 寡核苷酸色谱柱(2.1 × 50 mm, 2.7 μm, 120 Å, 部件号 659750-702)上分离消化的 mRNA 样品。流动相 和液相色谱梯度如表 1 所示,质谱仪设置如表 2 所示,在 MassHunter BioConfirm 10.0 中进行数据分析,解卷积设置如 表 3 所示。

表 1. 流动相和液相色谱梯度

Agilent 1290 Infinity II 液相色谱系统					
色谱柱	AdvanceBio 寡核苷酸柱,2.7 μm, 2.1 × 50 mm 120 Å				
溶剂 A	15 mmol/L 二丁胺 + 25 mmol/L HFIP 的去离子水溶液				
溶剂 B	15 mmol/L 二丁胺 + 25 mmol/L HFIP 的甲醇溶液				
梯度	时间 (min) % B 0 10 4 30 25 51 25-28 90				
柱温	50 °C				
流速	0.4 mL/min				
进样量	20 µL				
mRNA 每次进样量	2 µg				

表 2. 质谱仪设置

Agilent 6545XT AdvanceBio LC/Q-TOF				
采集模式	负离子,标准 (3200 m/z) 质量数范围,高灵敏度 (2 GHz)			
干燥气温度	350 °C			
气体流速	12 L/min			
雾化器	25 psig			
鞘气温度	275 °C			
鞘气流速	10 L/min			
V _{cap}	4500 V			
喷嘴电压	2000 V			
碎裂电压	250 V			
锥孔电压	65 V			
MS1 范围	400-3200 m/z			
MS1 扫描速率	2 Hz			
参比质量	1033.9881			

表 3. 解卷积设置

MassHunter BioConfirm 10.0 设置				
提取色谱图 (MS)	二极管阵列检测器			
MS1 延迟时间	0.06 s			
解卷积算法	最大熵			
扣除基线	1			
加合物	质子丢失			
质量数范围	10000-30000 Da			
质谱精度	0.5 Da			
使用有限的 m/z 范围	1040-3200			

结果与讨论

对于标准 LC/MS 分析而言全长 mRNA 通常过长,我们首先优 化了 RNA 酶 H 的定点裂解。设计了两个 RNA/DNA 嵌合探针与 5'UTR 内的不同序列互补,引导 RNA 酶 H 裂解^[4] 为核苷酸 40 或 50 (图 1A),以便在成功裂解后释放长度为 40 nt 或 50 nt、 带 5' 端加帽结构的寡核苷酸。这些较小的 5' 端加帽寡核苷酸 适用于 LC/MS 分析。



图 1. mRNA 样品前处理与分析。(A) 与 mRNA 样品互补的嵌合探针设计用于在 40 nt(紫色)和 50 nt(金色)处引导 RNA 酶 H 裂解。划线的探针核 苷酸由 DNA 组成,其余由 2'-O-甲基化修饰的 RNA 组成。(B) 使用非热稳定性 RNA 酶 H 的样品前处理方案(上图方案)或使用热稳定性 RNA 酶 H 的 样品前处理方案(下图方案)。使用下图方案的总分析时间:裂解(30 分钟)+ 纯化(10 分钟)+ LC/MS(25 分钟)+ 数据分析(10 分钟)= 75 分钟 如其他文献^[5] 所述,此类实验的典型工作流程需要在添加非 热稳定性 RNA 酶 H 前进行单独的一小时热退火步骤,然后在 37°C 下进行样品裂解(图 1B,上图方案)。有必要进行单 独的步骤,因为热退火过程中使用的高温(约 95°C)可能 会:(1)使非热稳定性 RNA 酶 H 变性,以及(2)由于反应缓冲 液中存在二价阳离子而促进 mRNA 水解^[6]。

我们假设,如果使用 GC 含量足够高的嵌合探针,则可以在适度升高的温度 (50°C) 下用热稳定的 RNA 酶 H 进行裂解反应 来加速反应(图 1B,下图方案)。原则上,升高的温度有助于解开 mRNA 二级结构,允许高 GC 含量的嵌合探针与其互补的 mRNA 区域结合,并通过热稳定性 RNA 酶 H 直接裂解。

如图 2 所示,使用两种不同的高 GC 含量探针(探针 40 和探 针 50,每种的 GC 含量 60%)和非热稳定性或热稳定性 RNA 酶 H 进行定点裂解。使用 Agilent 2100 生物分析仪自动化电 泳工具和小 RNA 试剂盒(部件号 5067-1548)分离产生的消 化产物,并使用虚拟凝胶图像进行定性评估。

在两种条件下探针 40 和探针 50 均可成功引导 mRNA 样品的 定向裂解。然而,当使用非热稳定性的 RNA 酶 H 时,探针 50 表现出脱靶结合,导致产生意外产物(图 2,红色箭头)。使 用热稳定性 RNA 酶 H 时,探针 50 可引导特异性裂解,从而 在较短时间内获得纯度更高的样品。



图 2. 使用热稳定性和非热稳定性 RNA 酶 H 评估定点裂解。探针 40 和探针 50 分别引导核苷酸 40 和核苷酸 50 处的直接裂解。红色箭头表示脱靶裂 解产物。L:分子量标准品

接下来,通过 IP-RP LC/MS 分析使用探针 50 和热稳定性 RNA 酶 H 消化未加帽 mRNA。图 3 显示了在 Agilent AdvanceBio 寡核苷酸色谱柱(2.1 × 50 mm, 2.7 μm, 120 Å,部件号 659750-702)上从探针 50 和剩余 mRNA 中分离 5'端未加帽 寡核苷酸的结果。使用 MassHunter BioConfirm 10.0 中内置 的自动化数据分析工作流程,从 UV 色谱图中确定的峰提取质 谱图,然后使用表 3 中所示的设置进行自动化解卷积。表 4 显示了每个反应中可能存在的预期的未加帽、加帽中间体和完 全加帽类型。

由于未加帽 mRNA 的 5' 端包含一个不稳定的三磷酸基团, 根据反应条件,可能会观察到一些含二磷酸基团或单磷酸基 团⁽³⁾ 的水解 mRNA。图 4 显示了来自含三磷酸基团峰和含二 磷酸基团峰的提取质谱图和解卷积质量数。优化的热稳定性 RNA 酶 H 裂解方案仅产生少量 (4.2%) 含二磷酸基团的寡核苷酸,且未检测到含单磷酸基团的寡核苷酸,表明反应条件适宜 温和。

未加帽样品中约 15% 含三磷酸基团的寡核苷酸是含额外非模 板 + G 核苷酸的序列变异体(图 4D)。这可能是由于 T7 转 录滑脱引起的,当重复的 G 核苷酸出现在转录序列的起始处 时,通常会出现这种情况^[7]。如图 5 所示,在与 ARCA 共转录 产生的 mRNA 样品中,无论是在 Cap 1 的后续甲基化反应之 前还是之后,滑脱序列变异体都很明显。滑脱序列变异体可被 识别为不同的色谱峰,但未实现基线分离。所有的加帽寡核苷 酸与未加帽寡核苷酸均实现了良好分离。

表 4. 未加帽、加帽中间体和加帽类型。红色字母表示带有"p"标记的磷酸基团的 5' 端核苷酸,可通过酶反应精细加工或替换为加帽结构



图 3. 从嵌合探针和样品基质中分离 5' 端未加帽寡核苷酸的色谱图。峰 1:5' 端含二磷酸基团的寡核苷酸,峰 2:5' 端含三磷酸基团的寡核苷酸



图 4.5' 端未加帽的 mRNA 寡核苷酸质谱图。来自图 3 峰 1 的 (A) 提取质谱图和 (B) 解卷积质谱图。来自图 3 峰 2 的 (C) 提取质谱图和 (D) 解卷积质谱图。红色数字表 示从含二磷酸基团和含三磷酸基团的寡核苷酸提取的质谱图中每个峰的电荷状态。用星号标记的质量峰(16131.16 Da 和 16211.79 Da)与其含二磷酸基团和含三磷 酸基团的寡核苷酸推断特征相匹配(序列插图),质量误差小于 10 ppm(表 5)。在图 D 中,观察到含三磷酸基团 + G 序列变异体(+345.17 Da)与含三磷酸基团的 未加帽寡核苷酸共洗脱



图 5. 加帽寡核苷酸与未加帽寡核苷酸的分离。峰 1 至峰 3: ARCA Cap 0,峰 4 至峰 6: ARCA Cap 1。加帽类型含有 0-2 个非模板 G 核苷酸,为滑脱序列变异体

图 6 显示了从 ARCA Cap 0 和 ARCA Cap 1 峰提取的质谱图和 解卷积质量数。与已发表的文献^[2]一致,我们发现 ARCA Cap 0 共转录加帽效率较高 (91.3 ± 1.76%),但并不完全。在随后甲 基化为 ARCA Cap 1 的样品中,加帽效率与 ARCA Cap 0 没有 显著差异 (90.6 ±1.53%),表明所有 Cap 0 结构基本上完全转 化为 Cap 1。

最后,使用 LC/MS 方法评估牛痘病毒酶加帽的效率。尽管研 究证明这种反应^[3] 比 ARCA 共转录加帽更有效,且与样品序列 无关,但据我们的经验,也可能恰恰相反。使用制造商推荐的 反应条件,牛痘病毒酶加帽产生了未加帽寡核苷酸、加帽中 间体和完全加帽寡核苷酸的混合物(图 7 和表 6)。如图 8 所 示,大多数样品 (90.9 ± 0.6%)为含 5' 端二磷酸基团或三磷酸 基团的未加帽寡核苷酸,6.5 ± 0.7% 为加帽但未甲基化的寡核 苷酸,即 G-Cap。只有 2.6 ± 0.5% 的样品被加帽并甲基化为 Cap 0。 表 5. ARCA 加帽寡核苷酸和加帽寡核苷酸中间体的质量数准确度

ARCA 共转录加帽					
特征	理论质量数	实测质量数	准确度 (ppm)		
ARCA Cap 0	16504.59	16504.62	2.01		
ARCA Cap 0 + G	16849.80	16849.69	6.37		
ARCA Cap 0 + 2G	17195.01	17195.03	1.50		
ARCA Cap 1	16518.59	16518.85	15.77		
ARCA Cap 1 + G	16863.80	16864.11	18.47		
ARCA Cap 1 + 2G	17209.01	17208.73	16.13		
含二磷酸基团	16131.52	16131.37	9.21		
含三磷酸基团	16211.51	16211.50	0.38		
含三磷酸基团 + G	16556.55	-	-		

表 6. 酶促加帽寡核苷酸和加帽寡核苷酸中间体的质量数准确度

酶促加帽						
特征	理论质量数	实测质量数	准确度 (ppm)			
Cap 0	16490.59	16490.65	3.55			
Cap 0 + G	16835.80	16836.00	11.97			
G-Cap	16476.57	16476.65	4.85			
含二磷酸基团	16131.52	16131.49	1.95			
含三磷酸基团	16211.51	16211.61	6.24			
含三磷酸基团 + G	16556.72	16556.97	15.33			



图 6. ARCA Cap 0 mRNA 寡核苷酸的质谱图。(A-C) 图 5 中峰 1-3 的解卷积质谱图。(D-F) 图 5 中峰 4-6 的解卷积质谱图。用星号标记的质量峰(16504.62 Da、 16849.69 Da、17195.36 Da、 16518.85 Da、 16864.11 Da 和 17208.73 Da)与其推断特征相匹配(序列插图),质量误差小于 20 ppm(表 5)。绿色下划线字母表示 可能由于 T7 转录滑脱而产生的非模板核苷酸



图 7. 牛痘病毒酶加帽寡核苷酸和加帽中间体的 LC/MS。(A) 峰 1-2: Cap 0 和 Cap 0 + G 序列变异体,峰 3: G-加帽寡核苷酸,(B-D) 加帽寡核苷酸的解卷积质谱 图。用星号标记的所有质量峰(16490.60、16836.00 和 16476.65 Da)均与其推断特征相匹配(序列插图),且质量误差小于 12 ppm(表 6)。绿色下划线字母表示 可能由于 T7 转录滑脱而产生的非模板核苷酸



图 8. 由 LC/MS 指导的牛痘病毒酶加帽流程的优化

为提高加帽效率,我们将起始 mRNA 的数量分别减少为 0.5 和 0.25,从而提高 SAM、GTP 和牛痘病毒加帽酶相对 mRNA 的摩尔比。该策略相当成功,导致未加帽的寡核苷酸减少 (0.5:73.6±1.3%为未加帽,0.25:58.6±0.8%为未加帽), 加帽的寡核苷酸相应增加 (0.5:17.7±1.8%为 Cap 0, 0.25:32.9±0.4%为 Cap 0)。单独增加单个反应物的浓度不 会提高加帽效率 (数据未显示),表明单个组分没有问题。

结论

本研究开发了一种更快的 LC/MS 方法来定量 mRNA 5' 端加 帽。该工作流程中的一个关键改进是使用热稳定性 RNA 酶 H 和具有足够高 GC 含量的 RNA/DNA 嵌合寡核苷酸探针来加 快 mRNA 裂解,无需单独的热退火步骤。GC 含量过高的探针 可能不合适,因为可能会增加脱靶结合或形成二级结构的可 能性。事实上,在 37 °C 下进行裂解时,我们观察到探针 50 (GC 含量为 67%,图 2)的脱靶裂解。由于较低的反应温度 有利于脱靶结合,在较高温度下进行裂解除可加速反应外,还 产生了纯度更高的样品。在单独的实验中,发现具有低 GC 含 量 (40%)的探针无法成功引导热稳定性 RNA 酶 H 进行定点裂 解(数据未显示)。

我们使用两种不同的加帽方法展示了此工作流程:(1)使用 ARCA 进行共转录加帽,(2)使用牛痘病毒加帽酶进行酶促加 帽。虽然 ARCA 加帽结果与文献一致,但牛痘病毒加帽酶加 帽结果却令人失望。对样品序列的分析突显出了一个可能的 原因:前12个核苷酸的GC含量非常高,超过90%,增加了 二级结构形成的可能性,这可能干扰了牛痘病毒加帽酶的反 应。由于均聚G重复序列也可导致将非模板G核苷酸包含在 内形成序列变异体(图5),这些结果表明重新设计该序列的 5'UTR 可处理的便利性和产品质量。

www.agilent.com

仅供科研使用。 不用于临床诊断用途。

本文中的信息、说明和指标如有变更,恕不另行通知。

© 安捷伦科技(中国)有限公司,2021 2021 年 9 月 21 日,中国出版 5994-3984ZHCN PR7000-3098 我们注意到,M.Beverly 等人^[5] 最近发表了一种类似的分析方 法,其中嵌合探针共价结合到磁珠上,可同时实现 RNA 酶 H 的定点裂解和裂解寡核苷酸的亲和富集。对我们来说,尽管进 行了多次优化,但这种方法未能成功引导样品进行特异性裂解 (数据未显示)。这可能是因为与样品 5' 端互补的嵌合探针具 有非常高的 GC 含量,从而导致广泛的非特异性结合。我们发 现工作流程更加灵活,因为可以通过设计嵌合探针避免样品中 具有挑战性的区域。此外,与基于亲和力的方法相比,使用硅 胶类离心柱进行样品纯化的序列依赖性可能更小,样品回收率 可能更高。

总之,本应用简报重点介绍了使用 Agilent 6545XT AdvanceBio LC/Q-TOF 对 mRNA 5' 端加帽进行快速高分辨率分析。我们预 计这一工作流程可用于药物开发和质量控制实验室。

参考文献

- Hassett, K. J. *et al.* Optimization of Lipid Nanoparticles for Intramuscular Administration of mRNA Vaccines. *Mol. Thera. Nucleic Acids* **2019**, *15*, 1–11
- Stepinski, J. et al. Synthesis and Properties of mRNAs Containing the Novel "Anti-Reverse" Cap Analogues 7-methyl(3'-0-methyl)GpppG and 7-methyl(3'-deoxy) GpppG. *RNA* 2001, 7, 1486–1495
- Fuchs, A. L. et al. A General Method for Rapid and Cost-Efficient Large-Scale Production of 5' Capped RNA. *RNA* 2016, 22, 1454–1466
- Lapham, J. et al. Site-Specific Cleavage of Transcript RNA. Methods in Enzymology 2000, 317, 132–139
- Beverly, M. et al. Label-Free Analysis of mRNA Capping Efficiency using RNase H Probes and LC-MS. *Anal. Bioanal.Chem.* **2016**, 408, 5021–5030
- 6. AbouHaidar, M. G. et al. Non-Enzymatic RNA Hydrolysis Promoted by the Combined Catalytic Activity of Buffers and Magnesium Ions. *Z Naturforsch C J Biosci* **1999**, *54*, 542–548
- T.Conrad et al. Maximizing Transcription of Nucleic Acids with Efficient T7 Promoters. *Nat. Comm. Biol.* 2020, 3, 1-8

