

使用安捷伦高分辨率 LC/(Q-)TOF 质谱分析寡核苷酸及其杂质的一体化 工作流程

目标寡核苷酸及其杂质的分离、表征和相对定量

作者

David L. Wong, Peter Rye

安捷伦科技有限公司



前言

近年来，寡核苷酸（包括小干扰 RNA、反义寡核苷酸、适配体和 CRISPR 向导）已成为快速增长的药物形式。随着此类候选药物的开发，对使用稳定分析方法和简便易用的数据分析工作流程开展表征工作的需求有所增加。此外，产品相关杂质的表征是新生物治疗药物开发中的一项重要任务。常见的杂质包括硫代磷酸酯转化形成的磷酸二酯以及截短、延伸和碱基切除的寡核苷酸^[1,2]。

LC/MS 等先进的分析方法对于表征目标寡核苷酸及其杂质是必不可少的，这些杂质通常数量众多，丰度极低，并且相互结合。因此，支持这些分析工作并能实现自动化的软件可能具有很大的价值。

为克服上述障碍，安捷伦开发了新型自动化 Agilent **MassHunter BioConfirm** 软件 12.0 版，支持使用分子式查找 (FBF, Find-by-Formula) 和最大熵算法来鉴定目标物及其杂质。图 2 描述了目标物与杂质 (TPI, Target Plus Impurities) 数据分析工作流程的详细信息。

实验部分

材料与amp;方法

三乙胺 (TEA) 和 1,1,1,3,3,3-六氟-2-丙醇 (HFIP)，购自 Sigma-Aldrich (St. Louis, MO, USA)。甲醇 (InfinityLab 超纯 LC/MS 级，部件号 5191-4497) 来自安捷伦科技公司。

寡核苷酸 (DNA) 分子量标准品 (部件号 5190-9029)、寡核苷酸 (RNA) 分离度标准品 (部件号 5190-9028) 和 RNA 标准品 (100 mer) 均来自安捷伦。

21 mer (CAG TCG ATT GTA CTG TAC TTA) 和 40 mer (CCA CGA CCA AGT GAC AGC AAT GAA TCG AGT CGA GAT CCA T) 寡核苷酸购自 Integrated DNA Technologies, Inc. (Coralville, IA, USA)，经过标准脱盐纯化。

样品前处理

- 寡核苷酸 (DNA) 分子量标准品和 (RNA) 分离度标准品在使用前均用 1 mL 去离子 (DI) 水溶解。每种标准品最终浓度为 2 pmol/μL
- 100 mer RNA 标准样品的浓度为 0.4 mg/mL
- 21 mer 和 40 mer 寡核苷酸样品也用 1 mL DI 水溶解，无需进一步纯化。然后将样品用储备液稀释至 0.50 mg/mL



图 1. 寡核苷酸分析 — 目标物与杂质 (TPI) 工作流程的分析组件



图 2. Agilent MassHunter BioConfirm 软件 12.0 版中的目标物与杂质 (TPI) 数据分析工作流程

仪器

- Agilent 1290 Infinity II 液相色谱系统包括：
 - Agilent 1290 Infinity II 高速泵 (G7120A)
 - Agilent 1290 Infinity II Multisampler (G7167B), 配备 Agilent Infinity II 样品冷却装置 (选件编号 100)
 - Agilent 1290 Infinity II 大容量柱温箱 (G7116B)
- Agilent 6545XT AdvanceBio LC/Q-TOF

LC/MS 分析

将 1290 Infinity II 液相色谱系统与配备双安捷伦喷射流离子源的 6545XT AdvanceBio LC/Q-TOF 系统联用, 进行 LC/MS 分析。使用 Agilent MassHunter 采集工作站软件 (11.0 版), 启用法规认证功能。使用 Agilent AdvanceBio 寡核苷酸色谱柱 (2.1 × 50 mm, 2.7 μm, 部件号 659750-702) 进行液相色谱分离。

表 1 和表 2 列出了所采用的详细 LC/MS 参数。

数据处理

寡核苷酸标准品和合成寡核苷酸样品的所有 LC/MS 数据文件均使用 Agilent MassHunter BioConfirm 软件 (12.0 版) 进行处理。

表 1. 液相色谱参数

Agilent 1290 Infinity II 液相色谱系统	
色谱柱	AdvanceBio 寡核苷酸色谱柱, 2.1 × 50 mm, 2.7 μm (部件号 659750-702)
恒温箱	4 °C
溶剂 A	15 mmol/L TEA + 400 mmol/L HFIP 水溶液
溶剂 B	甲醇
梯度	0–1 min, 10% B 1–10 min, B 由 10% 升至 40% 10–11 min, B 由 40% 升至 95%
柱温	65 °C
流速	0.5 mL/min
进样量	5.0 μL

表 2. 质谱数据采集参数

Agilent 6545XT AdvanceBio LC/Q-TOF 系统	
参数	设置
离子源	双 AJS
极性	负
气体温度	275 °C
气体流速	12 L/min
雾化器	35 psi
鞘气温度	350 °C
鞘气流速	12 L/min
毛细管电压	3500 V
喷嘴电压	2000 V
碎裂电压	175 V
锥孔电压	65 V
采集模式	HiRes (4 GHz)
质量数范围	300–3200 m/z
采集速率	4 幅谱图/秒

结果与讨论

寡核苷酸样品的全面表征可能是具有挑战性且耗时的过程，因为不仅需要分析目标寡核苷酸，还要鉴定所有相关杂质并相对定量。

各种寡核苷酸标准品的 HPLC 分离

研究开始时，对具有出色色谱分离能力和高分辨率精确质量数 (HRAM) 检测的 LC/MS 方法进行了优化。图 3 展示了两种安捷伦寡核苷酸标准品的 LC/MS 分

析：DNA 分子量标准品和 RNA 分离度标准品。使用离子对反相色谱实现了出色的色谱分离。研究还实现了从主峰中分离并检测杂质（小峰）。

合成寡核苷酸的 LC/MS 分析

使用相同方法获得了高灵敏度和高质量分辨率的质谱数据。图 4 显示了合成寡核苷酸 (40 mer) 样品的 LC/MS 结果。使用 11 min 梯度以 0.5 mL/min 的流速将大约 2.5 μg 样品进样至色谱柱。40 mer 寡核苷酸的多电荷分布图，质量数范围 m/z

600–3000（价态 -5 到 -19）。图 4B 插图的放大视图展示了 -13 价态寡核苷酸的出色质谱同位素分辨率。

研究优化了 Q-TOF 离子源条件，获得了优质的质谱图，具有较低的质量偏差 (3.6 ppm) (表 3)。此外，还检测到低丰度截短 (约 12000 Da) 和延伸组分 (约 12600 Da) (图 4C)。

100 mer RNA 标准样品获得了类似的高质量 MS 结果，如图 5 所示。同样观察到良好的质量数准确度 (9.96 ppm)。

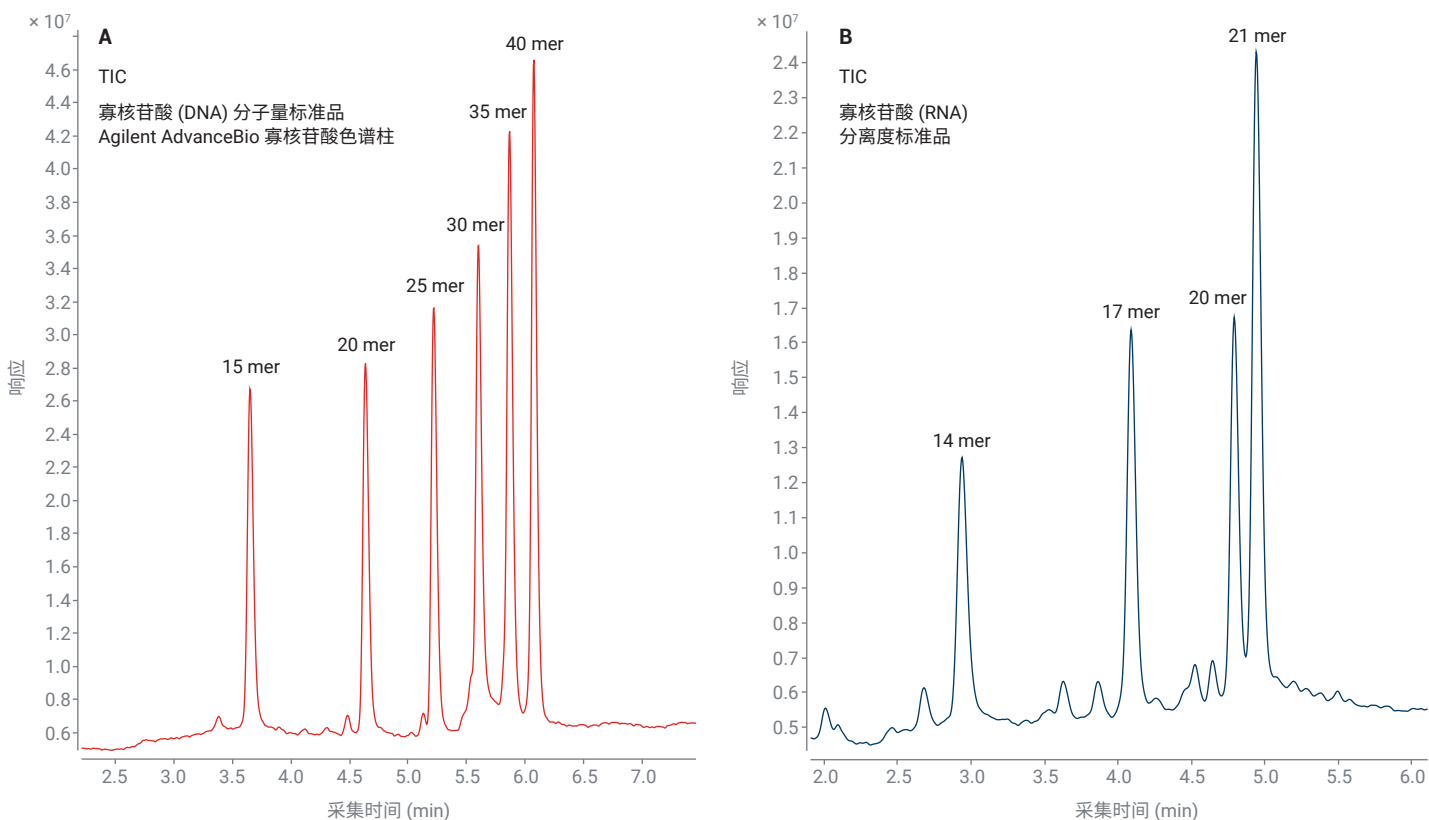


图 3. 安捷伦寡核苷酸分子量标准品 (DNA) 和安捷伦分离度标准品 (RNA) 的 LC/MS 分析

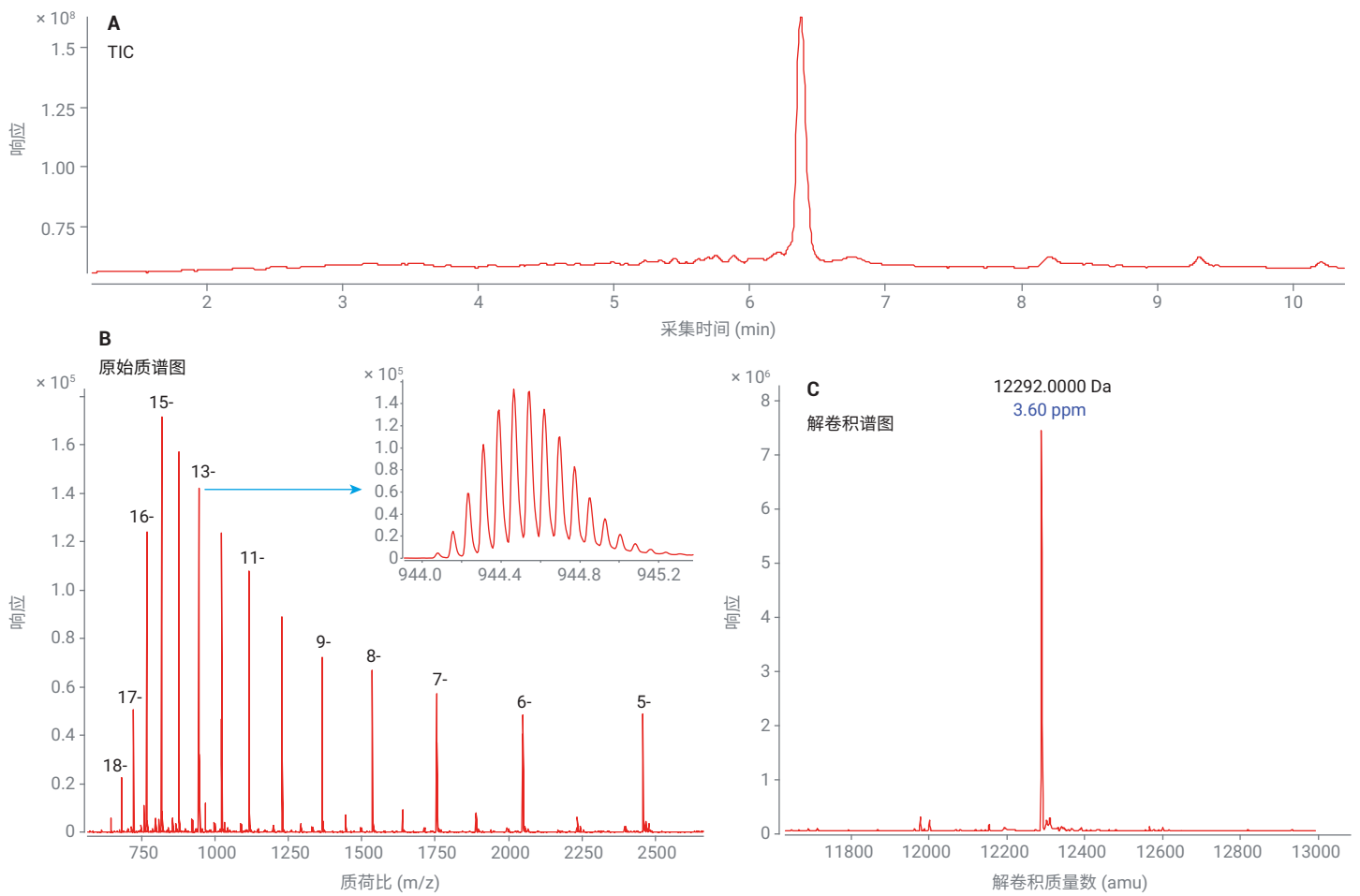


图 4. 合成寡核苷酸 (40 mer) 的 LC/MS 分析。(A) 40 mer 寡核苷酸的总离子色谱 (TIC)。(B) 40 mer 的原始质谱图。(C) 40 mer 的解卷积质谱图

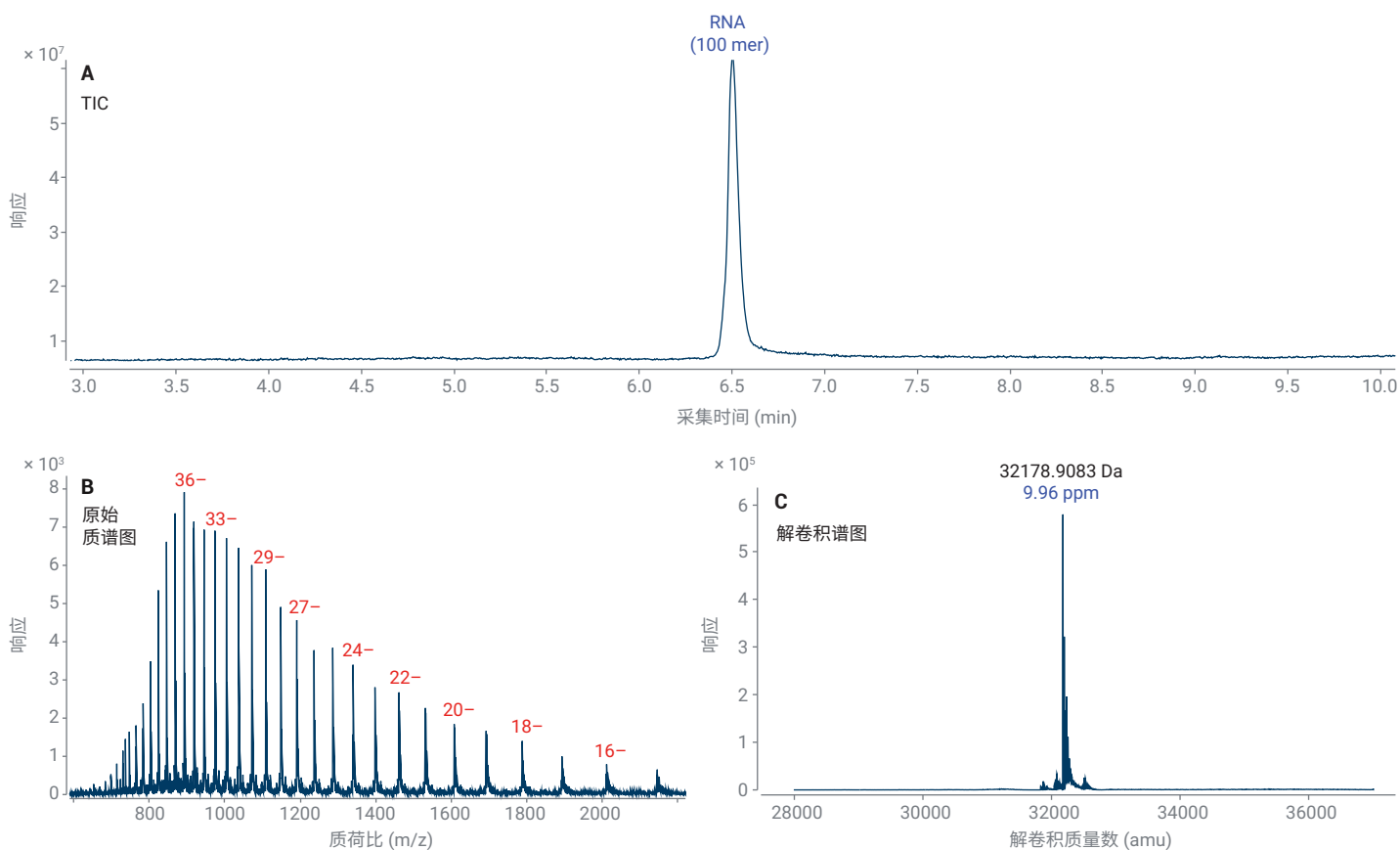


图 5. 合成寡核苷酸 (RNA, 100 mer) 的 LC/MS 分析

表 3. 分析的寡核苷酸列表。以绿色突出显示的计算质量数是单同位素质量数 (使用 FBF 匹配), 以蓝色突出显示的数字是平均质量数 (使用最大熵解卷积匹配)。总体而言, 使用所有分析的寡核苷酸样品均获得了出色的质量数准确度

寡核苷酸	寡核苷酸长度	序列	质量数计算值 (Da)	质量数测定值 (Da)	质量数准确度 (ppm)
寡核苷酸 (DNA) 分子量标准品	15	TTTTT TTTTT TTTTT	4498.7348	4498.7319	-0.64
	20	TTTTT TTTTT TTTTT TTTTT	6018.9650	6018.9635	-0.25
	25	TTTTT TTTTT TTTTTTTTTT TTTTT	7539.1952	7539.1989	0.50
	30	TTTTT TTTTT TTTTT TTTTT TTTTT TTTTT	9063.8431	9063.7988	-4.89
	35	TTTTT TTTTT TTTTT TTTTT TTTTT TTTTT TTTTT	10584.8111	10584.8065	-0.43
	40	TTTTT TTTTT TTTTT TTTTT TTTTT TTTTT TTTTT TTTTT	12105.7790	12105.8295	4.17
寡核苷酸 (RNA) 分离度标准品	14	rCrArCrUrGrArArUrArCrCrArArU	4395.6479	4395.6429	-1.14
	17	rUrCrArCrArCrUrGrArArUrArCrCrArArU	5335.7670	5335.7623	-0.88
	20	rUrCrArUrCrArCrArCrUrGrArArUrArCrCrArArU	6275.8861	6275.8800	-0.97
	21	rGrUrCrArUrCrArCrArCrUrGrArArUrArCrCrArArU	6620.9335	6620.9263	-1.09
DNA-21	21	CAGTCGATTGTACTGTACTTA	6408.0961	6408.0952	-0.14
DNA-40	40	CCACGCCAAGTGACAGCAATGAATCGAGTCGAGATCCAT	12291.9558	12292.0000	3.60
RNA 标准品 (长序列)	100	AACACCACCAUACAGUGCAGGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUCUUU	32178.5878	32178.9083	9.96

目标物与杂质 (TPI) 数据分析工作流程

虽然获得良好的 MS 结果是必要的, 但具备解析结果能力的强大软件程序同样重要。安捷伦开发了创新的自动化 MassHunter BioConfirm 12.0 软件, 支持利用分子式查找和最大熵算法来鉴定所含物质。在 BioConfirm 12.0 中开发了目标物与杂质 (TPI) 分析工作流程, 以及使用 MS/MS 数据的序列确认工作流程。本应用简报展示了 TPI 工作流程的独特功能。测序工作流程在单独的应用简报中进行介绍。

BioConfirm 12.0 中的 TPI 工作流程使用寡核苷酸质谱数据来分析 (即, 鉴定和相对定量) 目标寡核苷酸和相关杂质。图 6 显示了 TPI 工作流程的用户界面, 用户可以在其中定义目标寡核苷酸序列、可能的修饰以及匹配规则 (5' 或 3' 截短、删除或拆分), 以此进行数据处理。详细结果, 包括已鉴定的目标物及其寡核苷酸杂质, 可以以多个窗口/表格格式显示。

TPI 工作流程的两个选项

TPI 工作流程可以使用 FBF (靶向方法) 或最大熵解卷积 (非靶向方法) 运行, 其中工作流程转换质量数 (Workflow Transition Mass) 功能可确定每个样品使用哪种算

法。如果样品序列的质量数小于工作流程转换质量数, 则运行 FBF。如果样品序列的质量数等于 (或大于) 工作流程转换质量数, 则运行最大熵解卷积。此用户定义值可能需要多个因素的信息, 包括 MS 仪器的分辨率、目标物和杂质的质量数、靶向分析与非靶向分析的偏好以及质量数准确度要求。

表 3 显示了各种完整寡核苷酸样品的 LC/MS 分析汇总, 以及其质量数测定值和质量数准确度。使用 FBF 算法, 对多个短于 30 mer 的寡核苷酸序列实现了亚 ppm 级质量数准确度。对于较大的寡核苷酸样品, 通过运行最大熵算法也可以实现低 ppm 级的质量偏差。

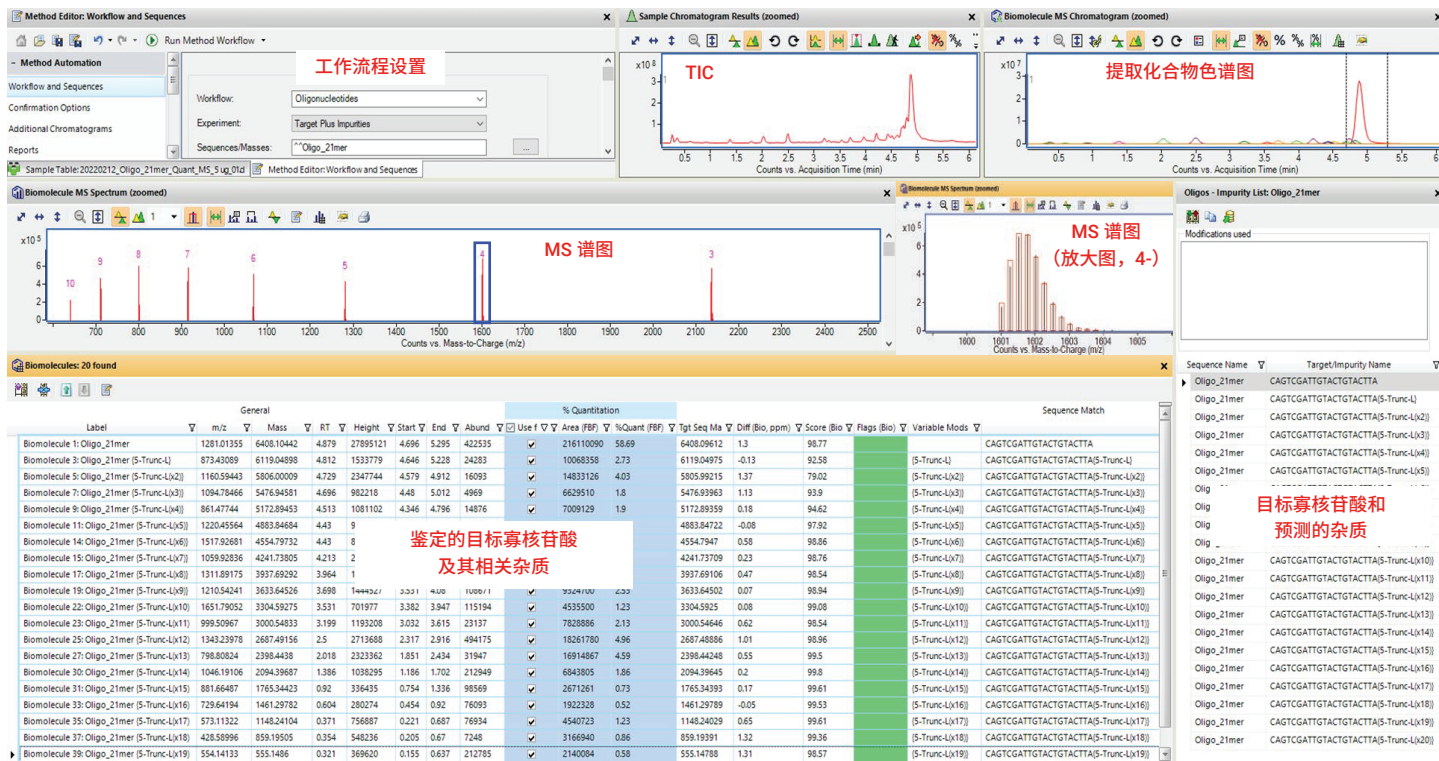


图 6. 采用目标物与杂质 (TPI) 工作流程的 Agilent BioConfirm 软件 (12.0 版) 概览

寡核苷酸杂质分析

寡核苷酸杂质分析是表征合成寡核苷酸产品相关杂质的一项关键性任务。由于寡核苷酸合成中的化学偶联效率仍待提升，有研究报告存在多种杂质，例如截短、添加和碱基切除寡核苷酸^[2]。要保持对所有产品相关杂质进行全面分析仍存在挑战。色谱分离和良好的质谱灵敏度非常利于检测低含量杂质。

本研究开发了一种用于表征全长寡核苷酸目标物及其杂质的快速 LC/MS 方法。图 7 展示了使用 FBF 测定的 21 mer 合成寡核苷酸和一组相关杂质的 LC/MS 图谱。21 mer 的大部分 5' 截短杂质可通过 11 min 的短液相色谱梯度得到良好分离。事实上，杂质和目标物在 6 min 内就

已洗脱（图 7A）。使用 BioConfirm 12.0 分子式查找算法的数据分析提供了准确的单同位素质量数，以及所有目标杂质的相对定量结果，如图 7B 所示。由于该 21 mer 合成寡核苷酸样品在初始脱盐过程后未进一步纯化，因此目标寡核苷酸几乎所有可能的 5' 截短杂质在相对定量中均低于 0.5%。

表 4 汇总了所有 19 种寡核苷酸杂质检测目标物，具有出色的解卷积质量数准确度（大多数在亚 ppm 水平）和相对定量重现性。请注意，许多其他类型的目标物相关杂质，例如带有连接子的 5' 截短体或 3' 截短体（带或不带连接子）也可在极低的丰度水平下被检测到（数据未显示）。

使用了最大熵解卷积鉴定一系列 5' 截短的 40 mer 寡核苷酸。研究鉴定了 40 mer 寡核苷酸样品的大部分 5' 截短杂质，前 10 位杂质（以及 40 mer 目标物）如图 8 所示。表 5 列出了使用最大熵方法进行的相对定量分析的详细信息。

与 FBF 结果类似，对于较大的寡核苷酸分子 (40 mer)，可以获得高质量数准确度和准确的相对定量结果。结果还表明在检测低含量寡核苷酸杂质方面具有出色的灵敏度。在 40 mer 寡核苷酸样品的前 15 种杂质中，获得的相对定量低至 0.65%（表 5）。如前所述，许多其他类型的目标物相关杂质也可在极低丰度水平被检测到（数据未显示）。

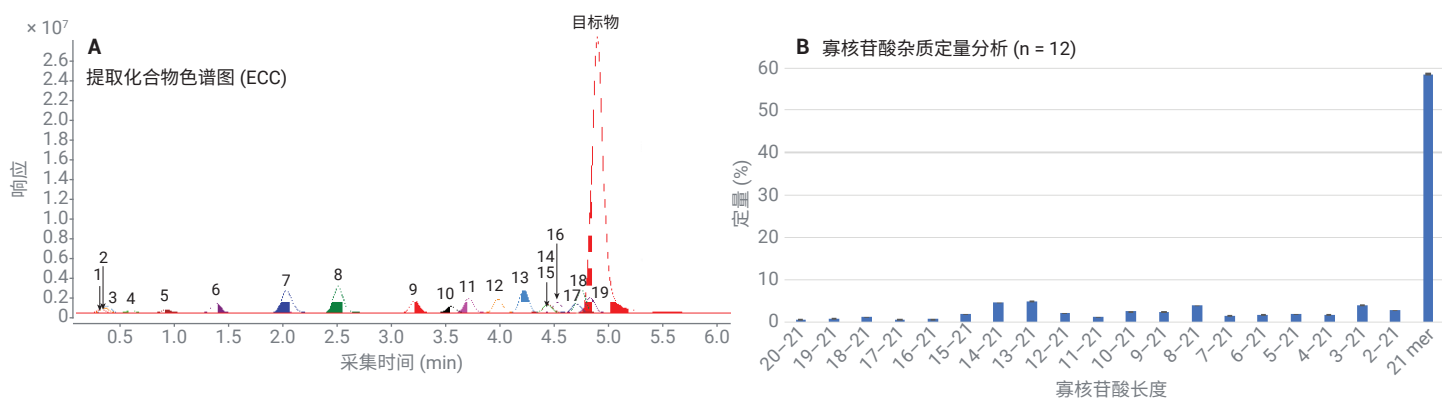


图 7. 通过 Agilent BioConfirm 软件分子式查找算法对合成寡核苷酸 (21 mer) 和目标杂质进行相对定量分析。(A) 21 mer 寡核苷酸及其杂质的提取化合物色谱。(B) 21 mer 寡核苷酸及其杂质的相对定量分析结果。在总共 12 次样品进样中实现了出色的重现性和极低的 RSD (<3%)

表 4. 21 mer 合成寡核苷酸的 19 种寡核苷酸杂质的杂质分析汇总 (n = 12)

杂质峰	寡核苷酸长度	RT (min)	单质量数计算值	质量数测定值	平均质量数准确度 (ppm)(n = 12)	平均定量% (n = 12)	标准偏差	RSD (%)	序列
1	20-21	0.321	555.1479	555.1486	1.21	0.57	0.01	2.39	TpA
2	19-21	0.354	859.1939	859.1950	1.09	0.89	0.02	1.76	TpTpA
3	18-21	0.371	1148.2403	1148.2410	0.81	1.28	0.02	1.44	CpTpTpA
4	17-21	0.604	1461.2979	1461.2978	0.15	0.53	0.01	1.18	ApCpTpTpA
5	16-21	0.920	1765.3439	1765.3442	0.57	0.72	0.01	1.72	TpApCpTpTpA
6	15-21	1.386	2094.3964	2094.3969	0.65	1.86	0.01	0.66	GpTpApCpTpTpA
7	14-21	2.018	2398.4425	2398.4438	0.69	4.61	0.04	0.91	TpGpTpApCpTpTpA
8	13-21	2.500	2687.4889	2687.4916	0.73	4.98	0.04	0.72	CpTpGpTpApCpTpTpA
9	12-21	3.199	3000.5465	3000.5483	0.33	2.14	0.02	0.75	ApCpTpGpTpApCpTpTpA
10	11-21	3.531	3304.5925	3304.5928	0.04	1.23	0.01	1.05	TpApCpTpGpTpApCpTpTpA
11	10-21	3.698	3633.6450	3633.6453	0.21	2.55	0.02	0.93	GpTpApCpTpGpTpApCpTpTpA
12	9-21	3.964	3937.6911	3937.6929	0.15	2.45	0.02	0.76	TpGpTpApCpTpGpTpApCpTpTpA
13	8-21	4.213	4241.7371	4241.7380	0.47	3.97	0.03	0.70	TpTpGpTpApCpTpGpTpApCpTpTpA
14	7-21	4.430	4554.7947	4554.7973	0.33	1.49	0.01	0.97	ApTpTpGpTpApCpTpGpTpApCpTpTpA
15	6-21	4.430	4883.8472	4883.8468	-0.23	1.74	0.01	0.85	GpApTpTpGpTpApCpTpGpTpApCpTpTpA
16	5-21	4.513	5172.8936	5172.8945	0.05	1.91	0.02	0.80	CpGpApTpTpGpTpApCpTpGpTpApCpTpTpA
17	4-21	4.696	5476.9396	5476.9458	1.22	1.81	0.01	0.79	TpCpGpApTpTpGpTpApCpTpGpTpApCpTpTpA
18	3-21	4.729	5805.9921	5806.0001	1.51	4.04	0.04	0.99	GpTpCpGpApTpTpGpTpApCpTpGpTpApCpTpTpA
19	2-21	4.812	6119.0498	6119.0490	-0.32	2.75	0.05	1.92	ApGpTpCpGpApTpTpGpTpApCpTpGpTpApCpTpTpA
目标物	21 mer	4.879	6408.0961	6408.1044	1.29	58.49	0.20	0.34	CpApGpTpCpGpApTpTpGpTpApCpTpGpTpApCpTpTpA

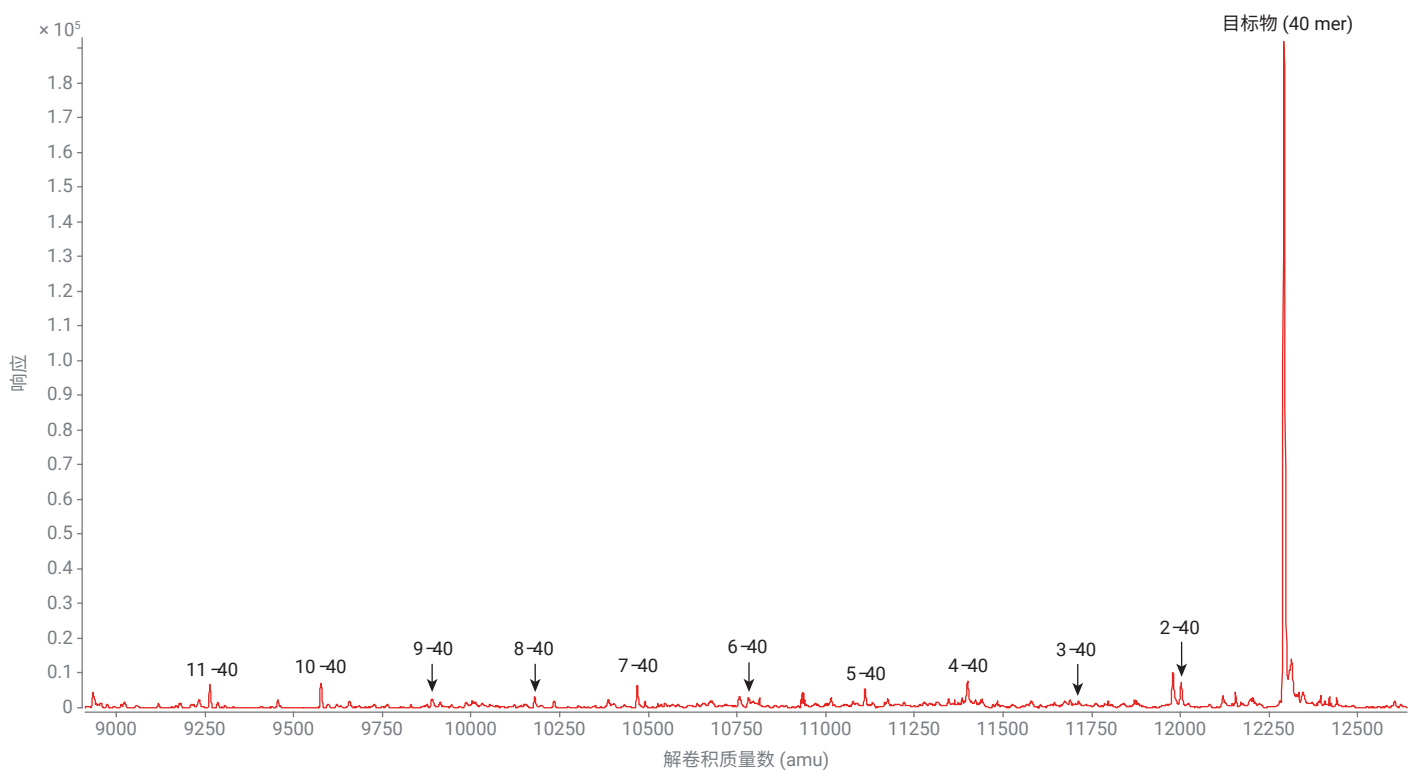


图 8. 使用最大熵质谱解卷积方法对 40 mer 寡核苷酸及其 5' 截短体 (2-40 到 11-40) 进行相对定量

表 5. 40 mer 合成寡核苷酸的前 15 种寡核苷酸杂质的杂质分析汇总

寡核苷酸长度	RT (min)	质量数测定值	定量%	序列
16-40	5.861	7699.0915	2.08	AGCAATGAATCGAGTCGAGATCCAT
15-40	5.878	7988.2299	2.26	CAGCAATGAATCGAGTCGAGATCCAT
14-40	5.994	8301.3760	1.24	ACAGCAATGAATCGAGTCGAGATCCAT
13-40	5.990	8631.0440	1.35	GACAGCAATGAATCGAGTCGAGATCCAT
12-40	6.048	8935.0245	1.60	TGACAGCAATGAATCGAGTCGAGATCCAT
11-40	6.073	9263.8039	2.39	GTGACAGCAATGAATCGAGTCGAGATCCAT
10-40	6.148	9577.3536	2.54	AGTGACAGCAATGAATCGAGTCGAGATCCAT
9-40	6.193	9891.1191	0.94	AAGTGACAGCAATGAATCGAGTCGAGATCCAT
8-40	6.185	10179.5068	1.18	CAAGTGACAGCAATGAATCGAGTCGAGATCCAT
7-40	6.214	10468.9206	2.36	CCAAGTGACAGCAATGAATCGAGTCGAGATCCAT
6-40	6.289	10782.0017	1.08	ACCAAGTGACAGCAATGAATCGAGTCGAGATCCAT
5-40	6.276	11111.0793	2.03	GACCAAGTGACAGCAATGAATCGAGTCGAGATCCAT
4-40	6.280	11400.2121	2.82	CGACCAAGTGACAGCAATGAATCGAGTCGAGATCCAT
3-40	6.334	11712.5835	0.65	ACGACCAAGTGACAGCAATGAATCGAGTCGAGATCCAT
2-40	6.339	12002.3602	2.68	CACGACCAAGTGACAGCAATGAATCGAGTCGAGATCCAT
目标物	6.384	12292.1749	72.80	CCACGACCAAGTGACAGCAATGAATCGAGTCGAGATCCAT

结论

本研究开发了两种创新的自动化和一体化寡核苷酸数据分析方法，用于使用 HRAM 质谱数据表征目标寡核苷酸及其相关杂质。分析结果表明，方法可对预期寡核苷酸实现出色的色谱分离和质量数准确度（亚 ppm）。LC/MS 结果也显示了实测的寡核苷酸及其杂质的准确相对定量，具有高重现性。新开发的 Agilent MassHunter BioConfirm 软件（12.0 版）能够以高通量方式自动处理 TPI 数据，从而显著缩短数据分析时间。

参考文献

1. Capaldi, D. et al. Impurities in Oligonucleotide Drug Substances and Drug Products. *Nucleic Acid Ther.* **2017**, 27, 309–322
2. Okafo, G.; Elder, D.; Webb, M. Analysis of Oligonucleotides and Their Related Substances. *ILM Publications* **2012**

查找当地的安捷伦客户中心：

www.agilent.com/chem/contactus-cn

免费专线：

800-820-3278, 400-820-3278 (手机用户)

联系我们：

LSCA-China_800@agilent.com

在线询价：

www.agilent.com/chem/erfq-cn

www.agilent.com

DE29613125

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，2022
2022 年 5 月 13 日，中国出版
5994-4817ZHCN