

イオン対液体クロマトグラフィーと質量分析を用いた合成 mRNA のオリゴマッピング

Maissa M. Gaye, Jonathan Fox, Johannes P.C. Vissers, Ian Reah, Chris Knowles, Matthew A. Lauber

Waters Corporation

要約

メッセージ RNA は、これを用いた癌治療に対する評価および COVID-19 に対する mRNA ワクチンの FDA 承認に示されているように、ヒトの医療にとって急速に重要な治療法になっています。mRNA ワクチンやその他のクラスの mRNA 医薬品の急速な開発は、分析手法の進歩によって支えられています。このような方法の重要な側面の 1 つは、液体クロマトグラフィー質量分析 (LC-MS) による配列のマッピングにより、治療用 mRNA のアイデンティティ、純度、修飾を確認することです。RNA の質量分析に基づくシーケンシングは、フラグメントの直接分子検出が提供されるため、テンプレート化による RNA シーケンシングより優れており、これを使用して、ヌクレオチド不純物の位置を特定し、重要な構造特性 (5' キャップおよびポリ A テール) を識別することができます。そのため、精密質量マッチから得られる mRNA 成分の注釈付きクロマトグラムをもたらす、mRNA の包括的なボトムアップ LC-UV-MS 特性解析のワークフローを、当社は提案します。

アプリケーションのメリット

- ACQUITY™ Premier Oligonucleotide BEH™ C₁₈ 300 Å カラムと組み合わせたイオン対逆相クロマトグラフィーを使用することによる、高いクロマトグラフィー分離能および MS 感度
- *in silico* mRNA 消化の計算および waters_connect/UNIFI™ サイエンスライブラリーの適用によって促進される精密質量マッチングに基づく、mRNA 消化物の自動注釈

はじめに

SARS-CoV-2 パンデミックは、核酸ベースの医薬品、特に合成 mRNA の急速な開発に刺激を与えました¹。1961 年に Brenner らによって発見されてから 40 年になります² [i]。mRNA は、癌治療のためのヒトでの臨床試験の開始¹⁻³および米国食品医薬品局による 2 種類の COVID-19 mRNA ワクチンの完全承認（それぞれ 2021 年 8 月および 2022 年 1 月）によって示されているように、大きな可能性がある重要な治療法に進化しています。mRNA ワクチンやその他のクラスの mRNA 医薬品の急速な開発は、分析手法の進歩によって支えられています。このような方法の重要な側面の 1 つは、質量分析と結合した液体クロマトグラフィー（LC-MS）によるオリゴマッピングおよびシーケンシングにより、治療用 mRNA のアイデンティティ、純度、修飾を確認することです。サンガー シーケンシングや次世代シーケンシング（NGS）などの核酸シーケンシングテクノロジーは、医薬品開発者に貴重な情報を提供します。ただし、タンデム MS（LC-MS/MS）⁴ や MS^E（低コリジョンエネルギーと高コリジョンエネルギーを交互に使用）⁵ ベースのフラグメンテーションと組み合わせて LC を使用することによって達成できる、レベルを上げた分析もあります。プロテオミクスのボトムアップアプローチと同様に、LC-MS/MS や MS^E ベースの配列決定では、ヌクレオチド不純物の検出と位置確認、および脂質付加核酸塩基⁶、末端キャッピングした残基、ポリ A テール修飾のような重要な構造特性など、RNA フラグメントを直接分子検出する利点があります^{4,7}。大量のデータ解析ソリューションが存在するボトムアッププロテオミクスワークフローとは異なり、RNA マッピングの選択肢は限られています。LC、UV 検出、MS 測定で構成される単一プラットフォーム内で、所定の合成 mRNA を特性解析するボトムアップアプローチに基づく、オリゴマッピングのワークフローを、当社は提案します。消化物成分を、社内で開発した自由に使用できる *in silico* 消化ライブラリーカリキュレーター（mRNACalcondemand）と waters_connect™ を組み合わせて解析し、注釈付きクロマトグラムを生成しました。ここでは、RNase T1 消化ルシフェラーゼ mRNA を使用して、mRNA 配列マッピングの分析アプローチを実証します。

実験方法

サンプル情報

Bijoyita Roy（New England Biolabs、マサチューセッツ州 Ipswich）からご提供いただいた Cypridina ルシフェラーゼ mRNA（キャップなし、ポリ A テール修飾なし）約 90 μg を、3' -グアノシンの特定の RNaseT1（Worthington Biochemical Corporation、ニュージャージー州 Lakewood）を使用して消化しました。このワークフローを、10 μg の TriLink Biotechnologies（CleanCap®FLuc mRNA、カリフォルニア州 San Diego）のホタルルシフェラーゼ mRNA（未翻訳の配列は特許取得済み）を使用して繰り返したところ、同等の結果が得られました。ルシフェラーゼ mRNA は、消化の前に、ヌクレアーゼフリーバッファー（10 mM トリス、0.1 mM EDTA 水溶液（pH 7.5）、Integrated DNA Technologies, Inc、アイオワ州 Coralville）中に調製した 20 μL 尿素（8 M）を用いて、80 °C で 5 分

間変性しました。次に、ヌクレアーゼフリーバッファーに再懸濁した 24 µg (約 10 kU) の RNase T1 (Worthington Biochemical Corporation、ニュージャージー州 Lakewood) を、変性 mRNA に室温で添加し、この混合液を 37 °C で 30 分間培養しました。培養終了時に、ヌクレアーゼフリーバッファー (40 µL) を添加してサンプル合計量を 80 µL にしました。最終アリコートは、ポリプロピレン製 300 µL オートサンプラーバイアル (製品番号: [186002639 < https://www.waters.com/nextgen/global/shop/vials-containers--collection-plates/186002639-polypropylene-12-x-32-mm-screw-neck-vial-with-cap-and-preslit-pt.html>](https://www.waters.com/nextgen/global/shop/vials-containers--collection-plates/186002639-polypropylene-12-x-32-mm-screw-neck-vial-with-cap-and-preslit-pt.html)) に移しました。得られた消化物は、それ以上の操作を行わずにイオン対逆相クロマトグラフィー (IP-RPLC) で処理し、その後 BioAccord™ RDa™ 検出器を使用してネガティブイオン化モードで MS 検出しました。

LC 条件

LC システム:	ACQUITY UPLC™ Premier BSM システム (BioAccord システムの一部として)
検出器:	ACQUITY UPLC TUV 検出器
波長:	260 nm
カラム:	ACQUITY Premier Oligonucleotide BEH C18、2.1 × 150 mm、300 Å、1.7 µm (製品番号: 186010541)
カラム温度:	70 °C
サンプル温度:	4 °C
注入:	5 µL
流速:	0.4 mL/分
移動相 A	0.1% N,N-ジイソプロピルエチルアミン (DIPEA) を IP 試薬とし、1% 1,1,1,3,3,3-ヘキサフルオロイソプロパノール (HFIP) を脱イオン水中に調製
移動相 B	0.0375% DIPEA、0.075% HFIP 含有 65:35 アセトニトリル: 水

グラジエントテーブル

時間 (分)	mL/分	A (%)	B (%)	曲線
初期条件	0.4	97	3	*
60	0.4	70	30	6
60.5	0.4	5	95	6
61	0.4	97	3	6
70	0.4	97	3	6

分析時間 = 70分

MS 条件

MS システム:	BioAccord LC-MS システム
検出器:	ACQUITY RDa 検出器
モード:	フラグメンテーションによるフルスキャン
極性:	ネガティブ
コーン電圧:	40 V
フラグメンテーションコーン電圧:	80 ~ 200 V
質量範囲:	高 (m/z 400 ~ 5,000)
スキャンレート:	2 Hz
キャピラリー電圧:	0.80 kV
脱溶媒温度:	400 °C

結果および考察

C₁₈ 固定相を用いるイオン対逆相クロマトグラフィー (IP-RPLC) は、オリゴヌクレオチドの分析での実証済みのアプローチになっています^{4,7}。移動相には、一般にはアルキルアミンなどのイオン対試薬が含まれます。アルキルアミンは C₁₈ 固定相^{8,9,10} に吸着され、これによって保持機構のようなミックスモードが導入されます⁸⁻¹⁰。このアプリケーションで使用する N,N-ジイソプロピルエチルアミン (DIPEA) /1,1,1,3,3,3-ヘキサフルオロイソプロパノール (HFIP) 移動相システムは、光学 UV 検出およびネガティブイオン化モード質量分析のいずれにも適合します^{4,7,-10}。HFIP は、エレクトロスプレーイオン化を促進するために使用します⁸。

RNase T1 消化ルシフェラーゼ mRNA を、ACQUITY Premier Oligonucleotide BEH C₁₈ (2.1 × 150 mm、300 Å、1.7 μm) カラムに注入し、ACQUITY UPLC TUV 検出器を装着した ACQUITY Premier バイナリー LC を使用してグラジエントを開発しました。この試験で使用した ACQUITY Premier Oligonucleotide BEH C₁₈ カラムは、ACQUITY Premier Oligonucleotide BEH C₁₈ 130 Å カラムと類似していますが、ポアサイズが大きいため、より長いオリゴヌクレオチド分子種の分離が向上しています。BioAccord ベンチトップ LC-MS システムの ACQUITY RDa 検出器で、ネガティブイオン化モード質量分析を使用してデータを 3 回繰り返し取り込みました。さらに、後で LC ピーク同定の確認に使用できる高エネルギーフラグメントイオンスペクトルがスキャン 1 回おきに生成されるように、ACQUITY RDa 検出器による MS^E データ測定をプログラムしました。図 1 に、RNase T1 対照サンプル (上のトレース)、mRNA 対照サンプル (下のトレース)、消化済み mRNA (中央のトレース) のトータルイオンクロマトグラム (TIC) が、示されています。RNase T1 によるルシフェラーゼ mRNA の消化で生じたオリゴヌクレオチドフラグメントは、4 シグマピークキャパシティ 613 で容易に分離されました。全体として、クロマトグラフィーピークはシャープかつ対称で、3 回繰り返し注入にわたる保持時間 (RT) のばらつきは約 0.01 分でした。グラジエント時間 60 分間のメソッドで、消化物の成分は 2 ~ 23 分で溶出し、消化が不完全な一部の mRNA は 29 分付近、インタクト RNaseT1 は 54 分付近で溶出しました。図 1 の上のトレースに示されているように、RNase T1 対照サンプルでは、保持時間 50 分の後にのみシグナルが見られました。これにより、RNase T1 によって mRNA 消化済み成分の保持時間ウィンドウ内に干渉が持ち込まれないことが確認されました (図 1、中央のトレース)。同様に、図 1 の下のトレースにより、インタクトルシフェラーゼ mRNA は約 38 分に溶出することが示されており、消化済みサンプルで 29 分に観察されたピーク (図 1、中央のトレース) は、消化が不完全な mRNA に対応することが確認されます。ここで、5' キャップ構造とポリ A テール構造で構成される合成 mRNA の場合、29 分と 37 分付近のピークが消化後に観察されることがわかります。約 38 分から約 37 分へのわずかなシフトは、消化されなかったポリ A 構造を示す可能性があります (このクロマトグラフィーの動作の検討については、今後のアプリケーションノートで説明します)。

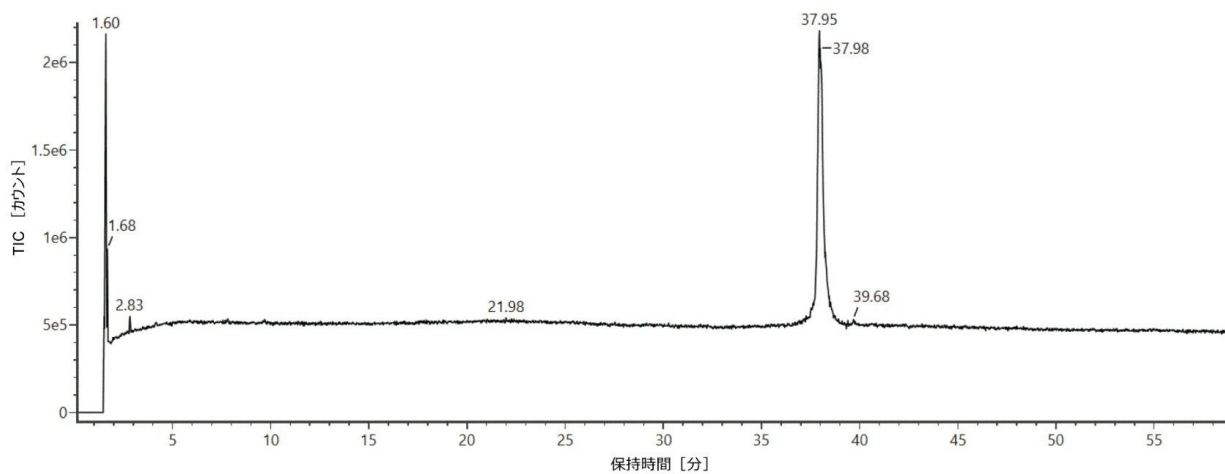
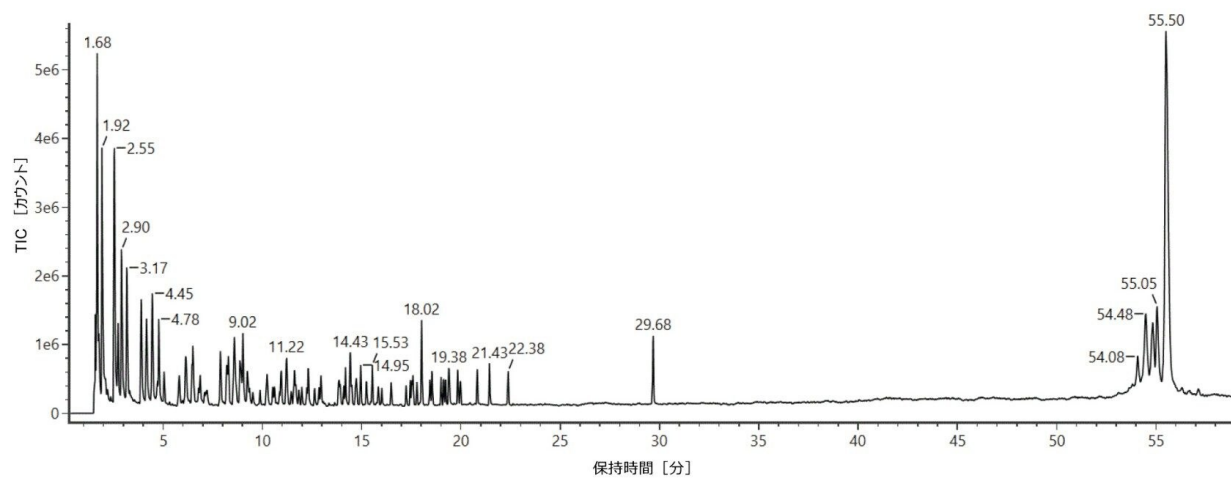
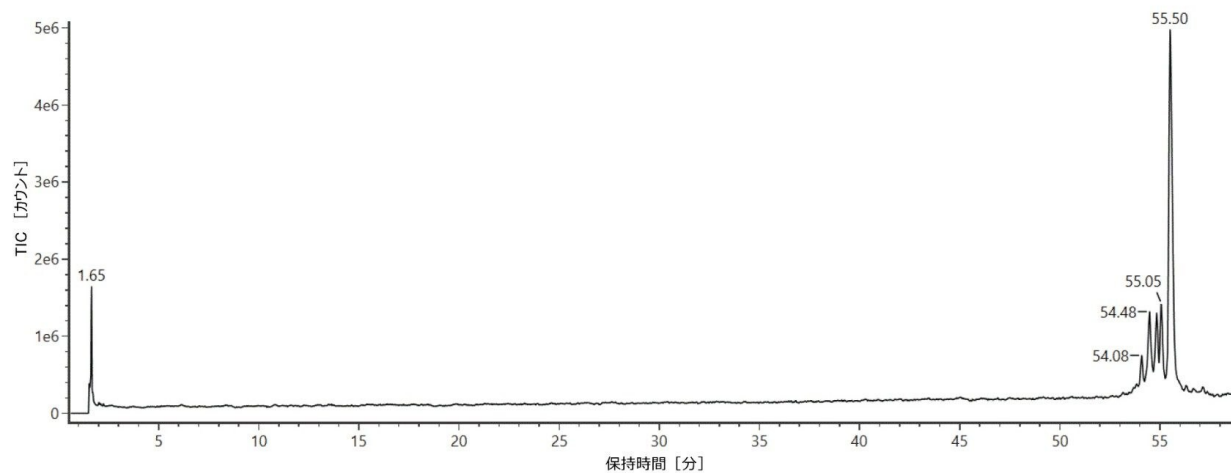


図 1: *RNase T1* 対照サンプル (上のトレース)、*mRNA* 対照サンプル (下のトレース)、および *RNase T1* で消化し、

ACQUITY UPLC I-Class システム (ACQUITY Premier Oligonucleotide BEH C₁₈ カラム、2.1 x 150 mm、300 Å、1.7 μm) およびネガティブイオン化モードの BioAccord ACQUITY RDa 検出器を使用して分析したルシフェラーゼ mRNA のイオン対逆相クロマトグラフィー (IP-RPLC) で得られた消化物 (中間トレース) の TIC クロマトグラム。

in-silico 消化 mRNA カリキュレーター (mRNAlcondemand) のグラフィカルユーザーインターフェース (GUI) がスキーム 1 に示されています。塩基配列の隣に、修飾、酵素、欠落している開裂など、いくつかの消化パラメーターが指定されています。カリキュレーターにより、この情報に基づいてチャージ状態や *m/z* 範囲などの MS 固有の設定、およびモノアイソトピック質量または平均質量に基づいて計算を行う機能が、既定に設定されます。生成された出力は、フラットテキスト csv ファイル形式であり、UNIFI または waters_connect ソフトウェアで活用、または補完するダウンストリーム分析に使用できます。

 digested_sequence....csv ^

スキーム 1: *in silico* mRNA 消化物質量計算用の mRNA カリキュレーター GUI。

次に、ライブラリーを作成し、ここでは消化物成分は個々の分析種と見なされます。これは、カリキュレーターの出力をスプレッドシートとして UNIFI サイエンスライブラリーにインポートすることで達成されました。作成したライブラリーを使用して、HRMS スクリーニング分析メソッドで、取り込み後の消化物化合物をターゲットとし、ユーザー指定の許容範囲を用いることができます。質量許容範囲ベースのライブラリーマッチの使用が、以下に示されています。ライブラリー検索の結果、注釈付きクロマトグラムが自動的に生成されます (図 2)。保持時間ウィンドウ 17 ~ 20 分の拡大表示に対応するクローズアップが、図 2 に示されています。

ロマトグラフィーピークに対応する一意の 2 つの ID が生成されました。異性体 ID（化学組成は同じで配列が異なる分子種）と同重体 ID（化学組成が異なるが [精密質量が異なる] ノミナル質量が似ている分子種）が存在することにも気付きました。最終的に、精密質量マッチに基づいて 90 の一意の成分を同定でき、これらは、異性体のヌクレオチド配列および同重体イオン（表 1 でグレーで強調表示されているセル）を含めて、表 1 に報告されています。

成分名	化学式	検出保持時間 (分)	ニードル質量 (Da)	観測された m/z	電荷	質量誤差 (ppm)	観測された追加のイオン
UGUp_1	C38H48N14O31P4	1.64	1320.15604	659.0689	-2	-4	-2H, -H
CCUUCGCCGp	C92H120N32O73P10	1.65	3150.40375	1049.1302	-3	0.4	-3H
UCUUCUGCGp	C84H107N30O66P9	1.7	2870.35772	1434.1633	-2	-6.3	-2H
CCGGUCAAAGp	C96H121N41O89P10	1.93	3261.45959	1086.1526	-3	3.9	-3H
GAGp	C40H50N20O27P4	2.1	1366.21047	682.0924	-2	-9.2	-2H
CCGp_8	C28H33N11O23P3	2.4	973.14057	485.567	-3	6.6	-2H, -H
GAGACCCUAGp	C115H145N49O82P12	2.54	3895.55339	1297.5089	-3	-3.2	-3H
CUCGp	C47H61N19O37P5	2.55	1624.21331	1623.2078	-1	1	-H
CUg	C28H37N10O23P3	2.55	974.12459	973.1177	-1	0.4	-H, -2H
GAGACUUAUUGp	C115H143N47O84P12	2.55	3897.52142	1298.1735	-3	3.5	-3H
UAUUGAUAGCAGp	C115H142N46O85P12	2.55	3898.50544	1948.2401	-2	-3.1	-2H
ACGp	C29H38N13O21P3	2.73	997.1518	996.1407	-1	-3.8	-H, -2H
CAGp	C29H38N13O21P3	2.73	997.1518	996.1407	-1	-3.8	-H, -2H
AUGp	C29H37N12O22P3	2.92	998.13582	997.1285	-1	0	-H, -2H
AAg_4	C30H38N15O20P3	3.18	1021.16304	1020.153	-1	-2.7	-H, -2H
ACACCCUAGAAAgp	C115H145N49O81P12	3.44	3879.55848	1292.17	-3	-8.8	-3H, -6H, -9H
UGAUUCUUGp	C75H94N26O60P8	3.9	2566.30045	1282.1387	-2	-3.9	-2H
CCCGp	C37H50N14O29P4	3.91	1278.18186	638.0888	-2	6.9	-2H, -3H, -H
CCUg	C37H49N13O30P4	3.91	1279.16587	638.5822	-2	9	-2H, -3H, -H
CUCUGUGp	C75H95N27O60P8	3.91	2581.31135	1289.6535	-2	3.4	-2H
GUUCUUGp	C75H96N28O59P8	3.91	2580.32733	1289.1652	-2	6.2	-2H
AAGCCUGp	C77H97N30O56P8	4.17	2627.36578	1312.675	-2	-1.1	-2H
UAGp_1	C39H49N17O28P4	4.79	1327.18834	1326.1802	-1	-0.7	-H, -2H, -3H
AAAgp	C40H50N20O28P4	5.06	1350.21556	1349.2049	-1	-2.5	-H, -2H, -3H
CUCUGp	C46H60N19O38P5	5.81	1685.19118	791.5962	-2	9	-2H, -3H, -H
UCCCGp	C46H61N19O37P5	5.83	1684.20716	791.1011	-2	5.1	-2H, -3H, -H
CACCGp	C47H62N19O38P5	6.18	1697.23438	802.6148	-2	4.8	-2H, -3H, -H
AUCUGp	C47H60N17O37P5	6.51	1692.20241	803.6021	-2	9.2	-2H, -3H, -H
AUCCGp	C47H61N18O36P5	6.53	1698.18239	803.1076	-2	6.1	-2H, -3H, -H
AAAAACUUGUUGCCGp	C144H179N60O103P15	6.52	4860.66775	1619.2101	-3	-4.7	-3H
CAAAGp	C49H62N23O33P5	6.79	1655.25684	826.625	-2	3.7	-2H, -3H, -H
CCUCUGp	C55H72N18O45P6	7.89	1890.23246	944.1145	-2	5.1	-2H, -3H, -4H, -H
CUCCCGp	C55H73N19O44P6	7.91	1889.24845	943.6212	-2	3.7	-2H, -3H, -4H, -H
CCCAAGp	C56H74N22O42P6	8.29	1912.27566	955.1342	-2	3	-2H, -3H, -4H, -H
AUCCAAGp	C57H73N23O42P6	8.58	1937.27091	967.6315	-2	2.6	-2H, -3H, -4H, -H
CAUUAAGp	C57H72N22O43P6	8.58	1938.25493	1937.2514	-1	1.9	-H
AUCUCUGp	C56H71N19O45P6	8.63	1915.22771	956.614	-2	6.9	-2H, -3H, -4H, -H
AUAACAAGp	C58H73N25O41P6	8.88	1961.28215	979.6375	-2	3	-2H, -3H, -4H, -H
UAACAAGp	C58H73N25O41P6	8.88	1961.28215	979.6375	-2	3	-2H, -3H, -4H, -H
AUCCAAGp	C57H73N23O42P6	8.89	1937.27091	967.6353	-2	6.5	-2H, -3H, -H
AACUCUGp	C57H73N23O42P6	9.03	1937.27091	967.634	-2	5.2	-2H, -3H, -4H, -H
ACACAAGp	C58H74N26O40P6	9.26	1960.29813	979.1457	-2	3.2	-2H, -3H, -4H, -H
ACAAAAGp	C59H74N28O39P6	9.52	1984.30936	991.1508	-2	2.7	-2H, -3H, -4H, -H
UCUUCUGp	C64H82N19O54P7	9.89	2197.24178	1097.6182	-2	3.5	-2H, -3H, -4H
AUCUCUGp	C65H83N22O52P7	10.23	2220.269	1109.1359	-2	7.1	-2H, -3H, -4H
UCUACUGp	C65H83N22O52P7	10.23	2220.269	1109.1359	-2	7.1	-2H, -3H, -4H
UCUACUGp	C65H83N22O52P7	10.61	2220.269	1109.1302	-2	2	-2H, -3H, -4H
AUUCUUGp	C66H84N25O50P7	10.53	2243.29621	1120.643	-2	1.2	-2H, -3H, -4H
CAUUCUGp	C66H84N25O50P7	10.53	2243.29621	1120.643	-2	1.2	-2H, -3H, -4H
CAUUCUGp	C66H84N25O50P7	10.96	2243.29621	1120.6441	-2	2.2	-2H, -3H, -4H
UAUUAAGp	C66H83N24O51P7	10.95	2244.28023	1121.141	-2	6.8	-2H, -3H, -4H, -H
AAACUUGp	C67H85N28O48P7	11.22	2266.32343	1132.1603	-2	4.5	-2H, -3H, -4H, -H
ACUAACAAGp	C67H85N28O48P7	11.22	2266.32343	1132.1603	-2	4.5	-2H, -3H, -4H, -H
CCACAAGp	C67H86N29O47P7	11.25	2265.33942	1131.6655	-2	2	-2H, -3H, -4H
UCUCUUGp	C73H94N22O61P8	11.69	2502.28307	1250.1365	-2	1.1	-2H, -3H, -4H, -5H
AAACAAGp	C69H86N33O45P7	11.84	2313.36188	1155.6762	-2	1.5	-2H, -3H, -4H
UCUUAAGp	C74H95N25O59P8	11.99	2525.31028	1261.6501	-2	1.2	-2H, -3H, -4H, -5H
CCACUUGp	C75H97N29O56P8	12.32	2547.35349	1272.6742	-2	3.1	-2H, -3H, -4H
AAAUUCUUGp	C76H96N30O56P8	12.63	2572.34873	1285.1683	-2	0.3	-2H, -3H, -4H, -5H
ACUAUCUGp	C76H96N30O56P8	12.63	2572.34873	1285.1683	-2	0.3	-2H, -3H, -4H, -5H
ACUAUCUGp	C76H96N30O56P8	12.85	2572.34873	1285.1676	-2	-0.2	-2H, -3H, -4H
ACUAUCUGp	C76H96N30O56P8	12.96	2572.34873	1285.1705	-2	2.1	-2H, -3H, -4H
CUCGCCGp	C56H73N21O44P6	12.64	1929.25459	642.0863	-3	9.8	-3H
GGCGp	C39H50N18O28P4	13.37	1358.19415	1357.1889	-1	1.5	-H
CCUACACGp	C84H109N32O63P9	13.84	2852.39477	1425.1981	-2	5	-2H, -3H, -4H, -5H
ACUAUCUGp	C84H108N31O64P9	13.87	2853.37879	1425.6931	-2	7.1	-2H, -3H, -4H, -5H
UCACUUCGp	C84H108N31O64P9	13.87	2853.37879	1425.6931	-2	7.1	-2H, -3H, -4H, -5H
UCACUUCGp	C84H109N31O64P9	14.1	2853.37879	1425.6993	-2	4.5	-2H, -3H, -4H, -5H
AUUCUUCUGp	C83H108N30O69P9	13.93	2833.30362	1415.653	-2	5.4	-2H, -3H, -4H, -5H
ACUAUCUGp	C85H107N32O64P9	14.19	2878.37404	1438.1666	-2	-9.7	-2H
UAUUAUUGp	C84H105N28O67P9	14.19	2856.33083	1427.1719	-2	9.1	-2H, -3H, -4H
UGAUUGAUUCUUGp	C131H162N44O105P14	14.19	4464.50161	1487.1526	-3	-6.6	-3H
AUAUAUUGp	C85H106N31O65P9	14.43	2879.35805	1438.6866	-2	9.8	-2H, -3H, -5H
ACACCUUAGp	C85H109N34O62P9	14.44	2876.406	1437.2048	-2	5.8	-2H, -3H, -4H, -5H
AACAACUUGp	C86H109N36O61P9	14.72	2900.41724	1449.2093	-2	5	-2H, -3H, -4H, -5H
AAACAACUUGp	C87H109N38O60P9	14.96	2924.42847	1461.2135	-2	3.9	-2H, -3H, -4H, -5H
CAACAACUUGp	C88H110N41O58P9	14.96	2947.45569	1472.7265	-2	3.5	-2H, -3H, -4H, -5H
AAAAACUUGp	C88H109N40O59P9	15.25	2948.4397	1473.2225	-2	6.2	-2H, -3H, -4H, -5H
CUGACCUUAUUCGp	C140H178N51O107P15	15.25	4749.61192	1582.19	-3	-5.8	-3H
UUUACACAGp	C94H119N35O71P10	15.55	3183.41532	1590.7086	-2	4.7	-2H, -3H, -4H, -5H
UAUAUAUUGp	C95H118N36O71P10	15.84	3208.41057	1603.2007	-2	1.2	-2H, -3H, -4H, -5H
AUAUAUAUUGp	C144H178N50O103P15	16.02	4845.65685	1614.2191	-3	3.1	-3H, -6H
CUACAUAAGp	C96H120N40O68P10	16.02	3230.45377	1614.2191	-2	-0.8	-2H, -3H, -4H, -5H
GGCUg	C48H61N20O37P5	16.02	1664.21946	1663.208	-1	-2.5	-H
UCCACUUAUUGp	C102H130N35O80P11	16.49	3465.42939	1731.7095	-2	0.8	-2H, -3H, -4H, -5H, -6H
ACUAUAUUGp	C104H130N39O78P11	17.24	3513.45186	1755.7214	-2	1.1	-2H, -3H, -4H, -5H, -6H
AUCUUAUUGp	C105H130N41O77P11	17.45	3537.46309	1767.7284	-2	1.9	-2H, -3H, -4H, -5H
ACUAUAUUGp	C114H143N45O83P12	18.45	3841.52036	1919.7549	-2	0.6	-2H, -3H, -4H, -5H, -6H
CUCUGGp	C57H72N22O45P6	18.45	1970.24476	1969.2473	-1	5	-H
GUUGp	C39H49N17O28P4	18.55	1355.17917	1358.1705	-1	-0.3	-H
AGGGp	C40H50N20O28P4	19.39	1382.20539	690.0949	-2	-2	-2H, -H
GAGGp_1	C40H50N20O28P4	19.39	1382.20539	690.0949	-2	-2	-2H, -H
UCUUGAUUCUUAUUGp	C188H234N68O144P20	19.39	6366.78303	2121.2642	-3	3.8	-3H
UAUAUAUAUUGp	C124H156N19O88P13	19.43	4169.58887	1388.9648	-3	4.8	-3H, -2H, -4H, -5H, -6H
CAGGp	C39H50N18O28P4	22.38	1342.19924	1341.1865	-1	-4.1	-H, -2H

表1: 精密質量マッチに基づいて、10 ppm の質量誤差範囲内で暫定的に同定およびバリデーションした、ルシフェラー

ゼ mRNA 消化物成分。異性体配列または同重体イオンに対応するセルは、グレーで強調表示されています。

図 3 に、消化成分 UCCACUCUAUGp の同定が得られたデータ例が示されています。この成分は、左上のクロマトグラムに表示されているように 16.49 分に溶出し、*in silico* 消化ライブラリーで、 m/z 576.5732 ($[M-6H]^{6-}$)、692.0876 ($[M-5H]^{5-}$)、865.3603 ($[M-4H]^{4-}$)、1154.1438 ($[M-3H]^{3-}$)、1731.7095 ($[M-2H]^{2-}$) の、2～6 個の負電荷がある 5 つのイオン（図 3、左下のトレース）に基づいて同定されました。 $[M-5H]^{5-}$ イオンおよび $[M-2H]^{2-}$ イオンの同位体分布が示され、チャージ状態の割り当てを裏付けるデータが示されています（図 3、右）。

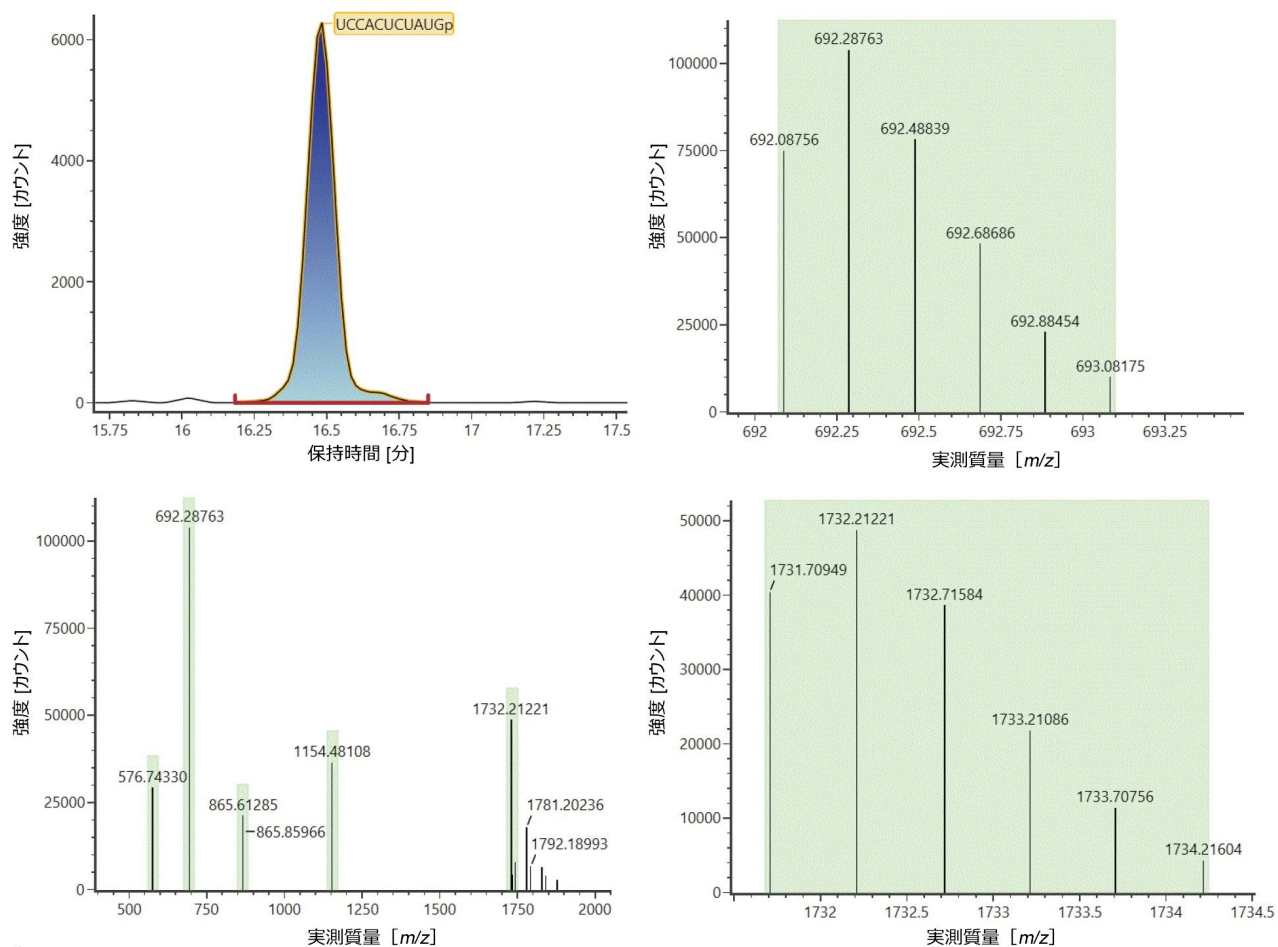


図 3: 16.49 分に溶出する消化物成分 *UCCACUCUAUGp* の同定 (左上のトレース)。この成分は、 m/z 576.5732 ($[M-6H]^{6-}$)、692.0876 ($[M-5H]^{5-}$)、865.3603 ($[M-4H]^{4-}$)、1154.1438 ($[M-3H]^{3-}$)、1731.7095 ($[M-2H]^{2-}$) の、2～6 個の負電荷を帯びた 5 つのイオンに基づいて生成された *in silico* 消化物から、同定されました。実験的に観察された $[M-5H]^{5-}$ イオンと $[M-2H]^{2-}$ イオンの同位体分布が、右上のトレースと下のトレースにそれぞれ示されています。

異性体または同重体イオンの存在による曖昧さは、表 2 に報告されているように、waters_connect CONFIRM Sequence™ アプリケーションを使用して MS^E スペクトルを解釈することで解決されました。位置 623～631 (ACAUCCUCGp) および 551～559 (UCACCAUCGp) の異性体配列は、RNase T1 の *in silico* 消化物から予測され、どちらも同じ保持時間 13.87 分に割り当てられています。インタクト質量分析を使用して正しい割り当てを行うことはできません。同じ注入で得られた MS^E データを用い、waters_connect CONFIRM Sequence™ アプリケーションを使用してこの割り当ての正しい配列を解析しました。そこでは、各配列に対して高エネルギーフラグメントイオンを予測し、カスタマイズされたアルゴリズムで波形解析した生データの同位体クラスターにマッチしました。確認されたフラグ

メントイオンをドットマップ上に示すことで、シーケンスカバー率を速やかに評価できます (図 4、B)。

UCACCAUCGp について、シーケンス完全カバー率が得られ、正しい割り当てとして容易にバリデーションできました (表 2、インデックス番号 63)。

手動でバリデーションした消化成分 (表 1) について得られた中性プリカーサー質量は、973.1406 Da (CCGp、RT 2.54 分) から 6366.7830 Da (UCAUUGAGUUCUCAAACUGp、RT 19.39 分) の 20-mer ヌクレオチドまでにわたっていました。最も早く溶出した成分は UGUUp で、実測中性質量 1320.1560、保持時間 1.64 分と特定されました。手動でバリデーションした一連の ID の中で溶出する最後のルシフェラーゼ mRNA 消化物成分 (表 1) は CAGGp (1342.1992 Da) で、22.38 分に溶出するのが観察されました。このような遅い保持時間での CAGGp の溶出は、密接に関連する配列 UAAGp が約 4 分に溶出したことを考えて、予想外でした。疑陽性の問題に対処するため、waters_connect CONFIRM Sequence アプリケーションを使用して、成分の ID をさらに特性解析しました。MS^E データが含まれている MS データに基づいてバリデーションしました。表 2 に示されているように、90 成分のうち CAGGp を含む 34 成分では、精密質量マッチによって配列割り当てがサポートされましたが、配列割り当てをそれ以上バリデーションするために十分な数の MS^E フラグメントが生成されませんでした。もう 1 つの興味深い別の例は、割り当てられた配列 AAAACAUGUUGCCGp (4860.6678 Da、15-mer、プリン 9 個) と AUACAUUUGACAAAGp (4845.6568 Da、15-mer、プリン 9 個) の同じデータセット内で、同じプリン含有量で 10 分離れた溶出が観察されたことです。MS^E データを使用して、AAAACAUGUUGCCGp を疑陽性として除外し、AUACAUUUGACAAAGp の同定を確認することができました。これにより、精密質量マッチとフラグメンテーションスペクトルの両方を戦略的に用いて、消化済み mRNA 配列から得られた成分を曖昧さなしに同定する補助とすることの、重要性が実証されています。さらに、得られた結果から、消化済み mRNA 成分の保持は、最初に考えたほど予測可能ではない可能性があること、およびオリゴヌクレオチドとクロマトグラフィー固定相の相互作用をさらに詳しく調べてモデル化することが急務であることを認識しました。

最後に、手動でバリデーションしたマッチ済み消化物成分を mRNA 配列と比較することで、シーケンスカバー率を手動で推定しました。401 の初期のマッチのシーケンスカバー率の暫定推定で、カバー率の値約 76% が得られました。厳密にバリデーションしたマッチを mRNA 配列に対して確認したところ、得られたカバー率の値は約 22% でした。観察された消化成分の多くはルシフェラーゼ mRNA 配列の複数の位置にマッピングされ (表 2)、修飾または完全修飾された核酸配列にはわずかに 4 種類の一意の残基のみが存在するため、冗長性が期待されました。

索引	精密質量マッチングを使用したバリデーションしたフラグメント mRNA 相対組成成分 (質量比 10 ppm)	測定保持時間 (分)	mRNA 配列上の位置	mRNA 配列上の追加の位置	MS ^E データを使用したバリデーション (CONFIRM Sequence)			
					カバー率 (%)			
					はい / いいえ	ピーク 1	ピーク 2	ピーク 3
1	UGUGp_1	1.64	217-229	1553-1556	はい	50		
2	CCUUCGCCGp	1.65	896-905	NA	はい	0		
3	UCUACUGCGp	1.7	71-79	NA	はい	0		
4	CCGUCGAAAGp	1.83	1520-1529	NA	はい	10		
5	GAAAGp	2.1	378-381	445-448, 635-638, 712-715	はい	25		
6	CCGp_e	2.54	59-61	246-248, 284-286, 838-840, 855-857, 1073-1075, 1076-1078, 1223-1225, 1520-1522	はい	0		
7	GAGACACCCUAGp	2.54	1667-1678	NA	はい	0		
8	CUCGp	2.55	25-29	NA	はい	0		
9	CUUp	2.55	177-179	422-424, 427-429, 452-454, 512-514, 560-562, 852-854, 906-908, 940-942, 965-967, 1430-1432, 1493-1495	はい	50		
10	GAGACAUCAUUp	2.55	455-466	NA	はい	16.67		
11	UAUUGAUAGCAGp	2.55	165-176	NA	はい	0		
12	ACGp	2.73	779-781	809-811, 820-822, 958-960	はい	0		
13	CAGp	2.73	174-176	273-275, 287-289	はい	50		
14	AUGp	2.92	209-211	291-293, 407-409, 571-573, 709-711, 806-808, 1100-1102, 1193-1195, 1271-1273, 1367-1369, 1415-1417, 1427-1429, 1467-1469, 1581-1583, 1688-1690	はい	100		
15	AAGp_4	3.18	43-45	146-148, 155-157, 379-381, 446-448, 636-638, 713-715, 1105-1107, 1139-1141, 1220-1222, 1304-1306, 1433-1435, 1475-1477, 1490-1492, 1679-1681	はい	100		
16	ACACCCUAGAAGp	3.44	1670-1681	NA	はい	16.67		
17	UGAUUUUp	3.9	1530-1537	NA	はい	50		
18	CCCGp	3.91	902-905	NA	はい	0		
19	CCUGp	3.91	1439-1442	1439-1442, 1564-1567, 1699-1602	はい	25		
20	CUUCUGUUp	3.91	1307-1314	NA	はい	12.5		
21	GUUCCUUp	3.91	1048-1055	NA	はい	12.5		
22	AAGCCUUp	4.17	1220-1227	NA	はい	12.5		
23	UAAGp_1	4.79	644-647	1371-1374	はい	50		
24	AAAGp	5.06	312-315	1147-1150, 1645-1648	はい	100		
25	CUCUGp	5.81	1307-1311	NA	はい	60		
26	UCCCGp	5.83	1232-1236	NA	はい	0		
27	CACCGp	6.18	1470-1474	NA	はい	0		
28	AUCUGp	6.51	639-643	1445-1449	はい	100		
29	AUCCGp	6.53	30-34	NA	はい	20		
30	AAAAACAUGUCCCGp	6.52	234-248	NA	はい	0		
31	CAAAAGp	6.79	822-826	NA	はい	40		
32	CCUCUGp	7.89	687-692	NA	はい	50		
33	CUCCCGp	7.91	968-973	NA	はい	33.33		
34	CCCCAGp	8.29	653-658	NA	はい	16.67		
35	AUCCAGp	8.58	734-739	NA (共置換)	はい	33.33	33.33	
36	CAUAGp	8.58	65-70	NA	はい	100		
37	AUUCUGp	8.63	1204-1209	NA	はい	66.67		
38	AUACAGp	8.88	1155-1160	NA	はい	66.67		
39	UACAAGp	8.88	841-846	NA	はい	100		
40	AUCCAGp	8.89	734-739	NA	はい	33.33	33.33	
41	AACCGp	9.03	110-115	332-337, 400-405, 984-989	はい	100		
42	ACACAGp	9.26	659-664	NA	はい	16.67		
43	ACAAAGp	9.52	1025-1030	NA	はい	33.33		
44	UCUUCUGp	9.89	828-834	NA	はい	71.43		
45	AUCUCUGp	10.23	679-685	NA	はい	14.29		
46	UCUACUGp	10.23	71-77	NA	はい	57.14	28.57	
47	AUUCAGp	10.53	318-324	NA	はい	14.29		
48	CAUACUGp	10.53	815-821	NA	はい	57.14		
49	AUUUCAGp	10.95	1406-1412	NA	はい	71.43		
50	AAACUUGp	11.22	1483-1489	NA	はい	71.43		
51	ACUACAGp	11.22	347-353	1285-1291	はい	71.43		
52	CCACAAGp	11.25	1567-1563	NA	はい	14.29		
53	UCUCUUCGp	11.69	1512-1519	NA	はい	37.5		
54	AAACAGp	11.84	1187-1203	NA	はい	71.43		
55	UCCUACGp	11.99	102-109	NA	はい	37.5		
56	CCACCAUGp	12.32	35-42	NA	はい	25		
57	AAAUCCUGp	12.63	798-805	NA	はい	12.5		
58	ACAUACUGp	12.63	1277-1284	NA	はい	14.29	62.5	40
59	CUGCCGp	12.64	852-857	NA	はい	0		
60	GCGp	13.37	367-370	NA	はい	0		
61	CCUCACACGp	13.84	929-937	NA	はい	22.22		
62	ACAUCCUGp	13.87	623-631	NA	はい	33.33		
63	UCACCAUCGp	13.87	551-559	NA	はい	100	33.3	
64	AUUCUUUUp	13.93	1418-1426	NA	はい	55.56		
65	ACAUCAUUp	14.19	458-466	NA	はい	22.22		
66	AUUCUUUUp	14.19	1083-1071	NA	はい	44.44		
67	UCAUUAUUCUUUUp	14.19	1413-1426	NA	はい	714		
68	AUACAUUUp	14.43	1016-1024	NA	はい	55.56		
69	ACACCUUAGp	14.44	1670-1678	NA	はい	11.11		
70	AACAUCGp	14.72	740-748	NA	はい	22.22		
71	AAACAUCGp	14.96	1182-1170	NA	はい	55.56		
72	CAAAACAAGp	14.96	665-673	NA	はい	77.78		
73	AAAAACAUGp	15.25	234-242	NA	はい	77.78		
74	CUGACCCUACUUCGp	15.25	512-526	NA	はい	0		
75	UUUCACACAGp	15.55	1109-1118	NA	はい	80		
76	UACAUUAAGp	15.84	1346-1355	NA	はい	80		
77	AUACAUUUGACAAAGp	16.02	1016-1030	NA	はい	20		
78	CUACAUAAAAGp	16.02	1653-1662	NA	はい	50		
79	GCGUGp	16.02	420-424	NA	はい	0		
80	UCCACUCUAGp	16.49	786-796	NA	はい	72.73		
81	AACAUUCUAGp	17.24	300-310	NA	はい	63.64		
82	AUCUJAAAAGp	17.45	698-708	NA	はい	63.64		
83	ACAUACUUCAGp	18.45	197-208	NA	はい	50		
84	CUUGCGp	18.45	177-182	NA	はい	0		
85	QUCGp	18.55	506-509	914-917, 1312-1315	はい	0		
86	AGGp	19.39	1603-1606	NA	はい	0		
87	GAGGp_1	19.39	632-635	1214-1217	はい	0		
88	UCAUUGAGUUCUACAAUCGp	19.39	593-612	NA	はい	0		
89	UAACUACAACCAAGp	19.43	1392-1404	NA	はい	69.23		
90	CAGGp	22.38	287-290	NA	はい	0		

表 2: 精密質量マッチ、および *waters_connect CONFIRM Sequence* アプリケーションと収集した MS^E スペクトルを

使用したさらなるバリデーションに基づく、ルシフェラーゼ mRNA 消化成分の同定およびバリデーション。

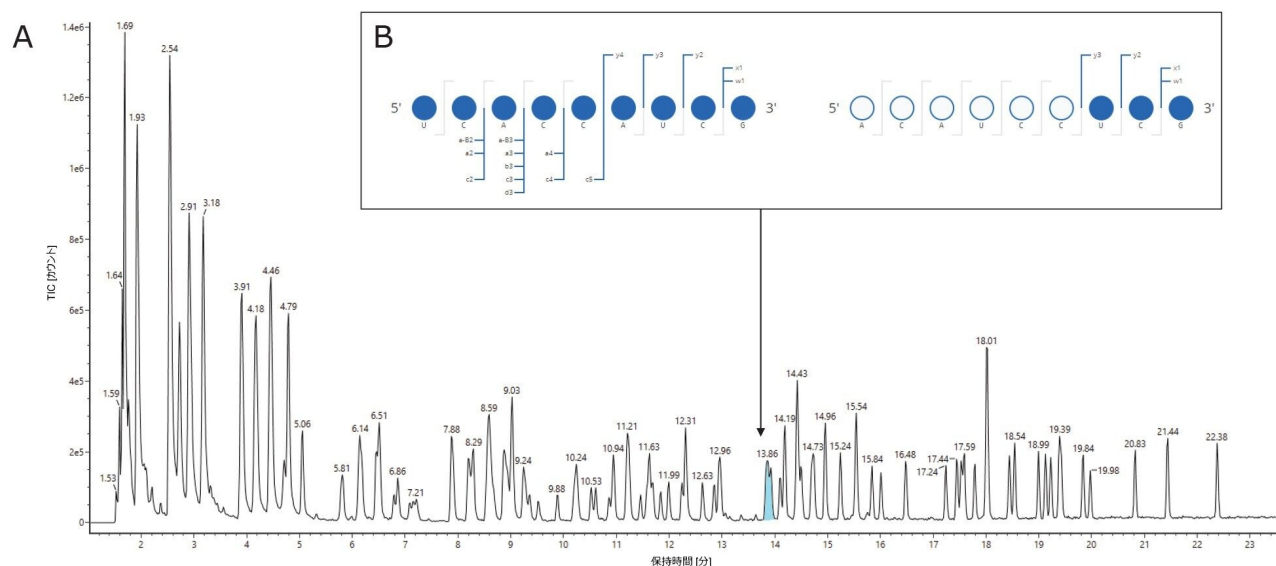


図 4: (A) 位置 623 ~ 631 (ACAUCCUCGp) および 551 ~ 559 (UCACCAUCGp) の消化済みフラグメント成分が RNase T1 消化物から予測され、TIC の同じ RT のピークに割り当てられます。インタクト質量情報を使用して正しい割り当てを決定することはできません。(B) 同じ注入から得られた MS^E データを使用して、この割り当ての正しい配列を解明できます。waters_connect CONFIRM Sequence アプリケーションを用い、McLucky の注釈¹⁾を使用して各配列に対して高エネルギーフラグメントイオンを予測し、カスタマイズされたアルゴリズムで波形解析した生データの同位体クラスターにマッチしました。このソフトウェアにより、確認済みのフラグメントイオンがドットマップに提示され、シーケンスカバー率がすばやく評価されます。

結論

本研究で、IP-RPLC と MS を用いた、合成 mRNA のオリゴマッピング用の頑健な分析ワークフローを確立しました。

- 合成 mRNA を、わずか 10 μg の試料から出発して、RNase T1 を用いて高い再現性で消化し、追加のサンプルクリーンアップなしで ACQUITY Premier Oligonucleotide BEH C₁₈ (2.1 x 150 mm、300 Å、1.7 μm) カラムに注入しました

- ACQUITY Premier LC でイオン対逆相クロマトグラフィーを使用して、高いクロマトグラフィー分離が達成され、これにより、消化物成分を消化が不完全な mRNA および残存する酵素から容易に分離でき、BioAccord ACQUITY RDa 検出器で効率的に検出できました
- *in silico* での mRNA 消化計算および waters_connect/UNIFI サイエンスライブラリーの適用によって促進される精密質量マッチに基づいて、注釈付き mRNA 消化クロマトグラムを生成しました
- 消化済み成分に割り当てた配列は、waters_connect CONFIRM Sequence アプリケーションを用いて、MS^E スペクトルに基づいてさらにバリデーションしました。さらに、ドットマップによる視覚化を使用して、可能性のある割り当てのフラグメントイオンカバー率を速やかに確認しました

本研究の目的は、mRNA 分子のボトムアップ特性解析を容易にするために必要な、クロマトグラフィー、検出、データ解釈のアプローチを確立することでした。RNase T1 消化は、データの収集と分析のワークフローを確立する、最初の例および概念の証明としてのみ、適用しました。とはいえ、(1) 複数のさまざまなヌクレアーゼを用いて補完性のある追加の配列マッピング情報を生成する、および (2) データ取り込みに多重化アプローチを採用することにより、所定の mRNA 構造をより包括的に精査する機会は十分にあります。包括的なシーケンスカバー率の達成を目的とするこれらの側面については、今後の研究で詳しく探査されるでしょう。

参考文献

1. Xu, S.; Yang, K.; Li, R.; Zhang, L., mRNA Vaccine Era—Mechanisms, Drug Platform and Clinical Prospection. *International Journal of Molecular Sciences* 2020, 21 (18), 6582.
 2. Brenner, S.; Jacob, F.; Meselson, M., An Unstable Intermediate Carrying Information From Genes to Ribosomes for Protein Synthesis. *Nature* 1961, 190 (4776), 576–581.
 3. Weide, B.; Pascolo, S.; Scheel, B. Derhovanesian, E.; Pflugfelder, A.; Eigentler, T. K. Pawelec, G.; Hoerr, I.; Rammensee, H. G.; Garbe, C., Direct Injection of Protamine-Protected Mrna: Results of a Phase 1/2 Vaccination Trial in Metastatic Melanoma Patients. *J Immunother* 2009, 32 (5), 498–507.
 4. Jiang, T.; Yu, N.; Kim, J.; Murgo, J.-R.; Kissai, M.; Ravichandran, K.; Miracco, E. J.; Presnyak, V.; Hua, S., Oligonucleotide Sequence Mapping of Large Therapeutic mRNA s via Parallel Ribonuclease Digestions and LC-MS/MS. *Anal.Chem.* 2019, 91 (13), 8500–8506.
 5. Plumb, R. S.; Johnson, K. A.; Rainville, P.; Smith, B. W.; Wilson, I. D.; Castro-Perez, J. M.; Nicholson, J. K., UPLC/MS^E; A New Approach for Generating Molecular Fragment Information for Biomarker Structure Elucidation. *Rapid Communications in Mass Spectrometry* 2006, 20 (13), 1989–1994.
-

6. Packer, M.; Gyawali, D.; Yerabolu, R.; Schariter, J.; White, P., A Novel Mechanism for the Loss of mRNA Activity in Lipid Nanoparticle Delivery Systems.*Nat. Commun.*2021, 12 (1), 6777.
7. Goyon, A.; Scott, B.; Kurita, K.; Maschinot, C.; Meyer, K.; Yehl, P.; Zhang, K., On-line Sequencing of CRISPR Guide RNAs and Their Impurities via the Use of Immobilized Ribonuclease Cartridges Attached to a 2D/3D-LC-MS System.*Anal. Chem.*2021.
8. Guo, L.; Worth, A. J.; Mesaros, C.; Snyder, N. W.; Glickson, J. D.; Blair, I. A., Diisopropylethylamine/Hexafluoroisopropanol-Mediated Ion-Pairing Ultra-High-Performance Liquid Chromatography/Mass Spectrometry for Phosphate and Carboxylate Metabolite Analysis: Utility for Studying Cellular Metabolism.*Rapid Commun Mass Spectrom* 2016, 30 (16), 1835–1845.
9. Birdsall, R. E.; Gilar, M.; Shion, H.; Yu, Y. Q.; Chen, W., Reduction of Metal Adducts in Oligonucleotide Mass Spectra in Ion-Pair Reversed-Phase Chromatography/Mass Spectrometry Analysis.*Rapid Commun Mass Spectrom* 2016, 30 (14), 1667–1679.
10. Fountain, K.; Gilar, M.; Budman, Y.; Gebler, J., Purification of Dye-Labeled Oligonucleotides by Ion-Pair Reversed-Phase High-Performance Liquid Chromatography.*Journal of chromatography.B, Analytical technologies in the biomedical and life sciences* 2003, 783, 61–72.
11. McLuckey, S. A.; Van Berkel, G. J.; Glish, G. L., Tandem Mass Spectrometry of Small, Multiply Charged Oligonucleotides.*J. Am.Soc.Mass Spectrom.*1992, 3 (1), 60–70.

謝辞

本研究に貴重な貢献をしたウォーターズの同僚 Ana-Maria Rotaru、Emanuela Petreanu、Claudia Florea、Dave Jackson、Simon Jones に感謝致します。New England Biolabs での共同研究者である Bijoyita Roy、Siuhong Chan、Ivan R. Corrêa Jr.、Erbay Yigit、G. Brett Robb に、ルシフェラーゼ mRNA および多くの議論を提供して頂いたことについて、感謝致します。

ソリューション提供製品

ACQUITY UPLC I-Class PLUS システム <<https://www.waters.com/134613317>>

バイオ医薬品のための BioAccord LC-MS システム <<https://www.waters.com/waters/nav.htm?cid=135005818>>

ACQUITY UPLC チューナブル UV 検出器 <<https://www.waters.com/514228>>

ACQUITY RDa 検出器 <<https://www.waters.com/waters/nav.htm?cid=135077027>>

UNIFI 科学情報システム <<https://www.waters.com/134801648>>

waters_connect <<https://www.waters.com/waters/nav.htm?cid=135040165>>

720007669JA、2022 年 6 月



© 2024 Waters Corporation. All Rights Reserved.

[利用規約](#) [プライバシー](#) [商標](#) [サイトマップ](#) [キャリア](#) [クッキー](#) [クッキー環境設定](#)