

GCxGC-TOFMSとmsFineAnalysis AIによるHeLa細胞中市販マスペクトルデータベース未登録代謝物の構造推定

関連製品: 質量分析計 (MS)

はじめに

ガスクロマトグラフ質量分析計 (GC-MS) は、豊富なマスペクトルデータベース (DB)、操作性の高さ、測定の再現性に優れる点からメタボローム解析に広く用いられている。GC-MSは揮発性化合物の分析を得意とする一方、トリメチルシリル (TMS) 化などの誘導体化処理を行うことでアミノ酸、有機酸、糖類などの極性が高い水溶性代謝物の測定も可能となる。そのため、血漿や組織抽出物など複雑マトリクス中の代謝物プロファイリングにおいて重要な分析手法となっている。しかし、生物の代謝物は数多くの化合物から構成されており、従来の一次元GC-MS測定では成分の共溶出により解析が困難な場合がある。このようなケースでは、包括的二次元ガスクロマトグラフィー質量分析 (GCxGC-MS) 法が複雑なピーク分離に有効な手法とされる。

GC-MSやGCxGC-MSを用いた代謝物の定性解析は、市販のマスペクトルDBとの比較分析によって化合物推定を行うことが一般的であるが、DBに未登録の化合物が検出されることも多い。我々はこの課題に対し、ガスクロマトグラフ高質量分解能飛行時間質量分析計 (TOFMS) のデータと、深層学習によるマスペクトル予測を組み合わせた自動構造解析手法¹⁾ (以後AI構造解析) を搭載したソフトウェアmsFineAnalysis AIを開発し、2022年にリリースした。さらに、2025年にAI構造解析の更なる精度向上、GCxGCデータ解析機能、AI分子式レコメンド機能を搭載したVer. 3をリリースした。本ソフトウェアには、約2億化合物の構造式と予測EIマスペクトルのDB (以下、AIライブラリー) が収録されており、統合解析²⁾ により決定した分子式情報を活用することで、迅速に未知物質の構造式を推定することが可能である。

本アプリケーションノートでは、HeLa 細胞中の代謝物をGCxGC-TOFMS で測定し、NIST DBには登録されていないものの構造既知である代謝物である*N*-メチルウリジル酸³⁾について、統合解析による分子式決定とAI構造解析を行った結果について報告する。

実験

サンプル準備

サンプルには、ヒューマンメタボロミクスにおけるスタンダードとして広く用いられているHeLa細胞を使用した。1×10⁷ cellsまで培養した細胞試料から、クロロホルム/メタノール/水の抽出混合液で代謝物を抽出し、メタノール/水層を得た。その後、遠心濃縮により溶媒を除去し、水溶性代謝物を得た。抽出の際、シナピン酸を内部標準物質 (I.S.) として添加した。抽出した代謝物は、メトキシアミン塩酸塩ピリジン溶液によるメトキシ化 (MeOX) 誘導体化およびMSTFAによるトリメチルシリル (TMS) 誘導体化を行い、測定に供した。

GCxGC-MS 測定、解析

測定は、ガスクロマトグラフ飛行時間質量分析計 JMS-T2000GC AccuTOF™ GC-Alpha 2.0 (日本電子製、図 1) とサーマルモジュレーター-INSIGHT-Thermal (SepSolve社製) を組み合わせたGCxGC-TOFMSシステムで実施した。イオン化法には、EI法および電界イオン化 (FI) 法を用いた。解析には、未知物質構造解析ソフトウェアmsFineAnalysis AI (日本電子製) を使用した。本ソフトウェアにより、EI法およびFI法で得られたデータを統合して解析した。その他の詳細条件は表 1に示す。



図1 JMS-T2000GC AccuTOF™ GC-Alpha 2.0

表 1 GCxGC-MS 測定条件

GCxGC		MS	
Column	1st: BPX5, 30 m× 0.25 mm, 0.25 μm, 2nd: Rxi17-SilMS, 3.4 m× 0.15 mm, 0.15 μm	Ion Source	EI/FI combination ion source
Inlet	250 °C, Splitless	Ionization	EI+: 70 eV, 300 μA, FI+: -10 kV, 6ms, 12, A
Oven	80 °C (2 min) → 5 °C/min → 325 °C (9 min)	Acq. Speed	EI: 50 Hz, FI: 25 Hz
Carrier flow	He, 1.2 mL/min (Constant Flow)	<i>m/z</i> Range	EI: <i>m/z</i> 33 - 800, FI: <i>m/z</i> 33 - 800
Modulation period	6 sec		

結果と考察

TICCと統合解析結果

図 2にEI法で取得したGCxGC トータルイオンカレントクロマトグラム (GCxGC TICC) および、主要成分のアノテーション結果を示す。GCxGC の高分離能により、800以上のピークが検出された。そのうち、サンプル由来の成分と推定されるものは、287化合物であった。検出された主要な化合物は、アミノ酸や糖、核酸関連物質などの代謝物と考えられる成分であった。

さらに、GCxGC TICC (二次元プロット) の1st RT 43.0-51.0 min, 2nd RT 3.5-5.3 min付近の拡大図を図 3に示す。1st RT 44.5 min, 2nd RT 4.5 minにおいて、今回着目している代謝物のN-メチルウリジル酸³⁾と推定される成分が検出された。なお、二次元プロットでは、N-メチルウリジル酸と近い2nd RT付近に、5'-メチルチオアデノシン (MTA) やアデノシンーリン酸 (AMP) といった核酸関連物質が検出された。今回のGCxGC-TOFMS測定におけるカラムセットでは、1stカラムに微極性、2ndカラムに中極性カラムを使用しており、横軸方向は沸点順、縦軸方向は極性順でそれぞれの成分が溶出する。このように、GCxGC-TOFMS測定により得られた二次元プロットでの溶出位置が構造特性を反映していることが確認できた。また、図 3中のピークBは、N-メチルウリジル酸と同じ1st RTを持つ。つまり、沸点が同じで極性が異なる化合物であり、GCxGC-TOFMS測定により両化合物のマススペクトルが分離されたことで、解析精度の向上に繋がった。

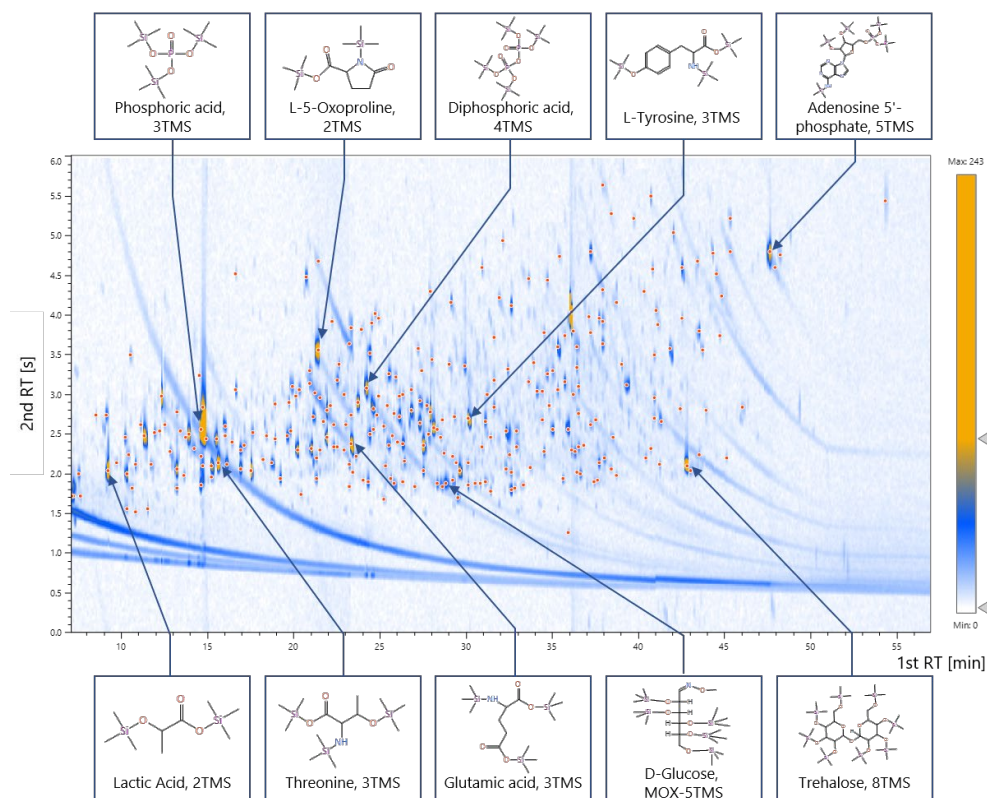


図 2 GCxGC TICC (EI法) と強度上位成分のアノテーション結果

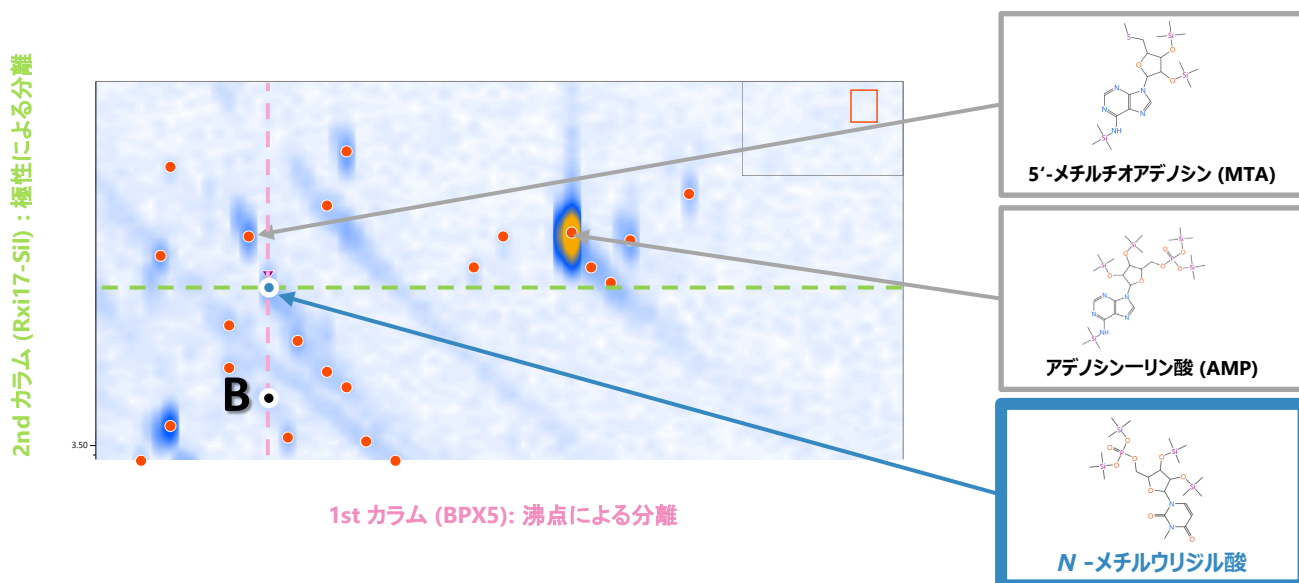


図 3 GCxGC TICC (EI法) の拡大結果

この資料に掲載した商品は、外国為替及び外国貿易法の安全輸出入管理の規制品に該当する場合がありますので、輸出するとき、または日本国外に持ち出すときは当社までお問い合わせください。

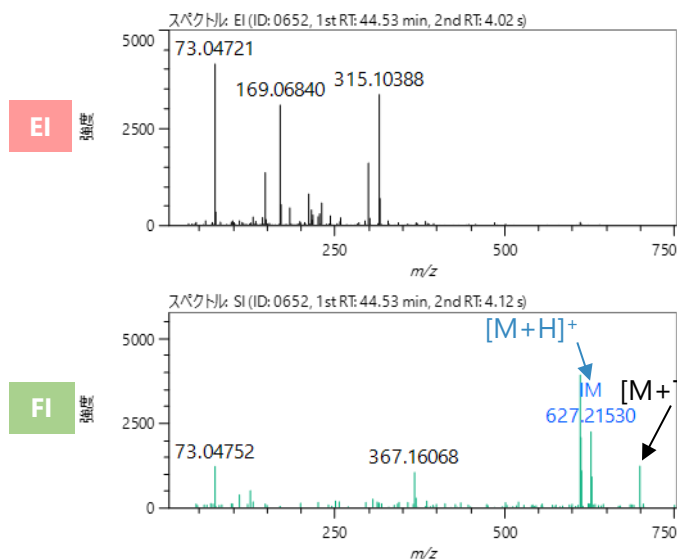
TICCと統合解析結果

次に、*N*-メチルウリジル酸と推定される成分のEI法およびFI法により取得したマスペクトルと統合解析結果を図 4に示す。まず、マスペクトルにおいては、本成分はEI法では分子イオンが検出されず、FIでのみプロトン付加分子 ($[M+H]^+$) と推定されるピーク m/z 627.21530が検出された。本成分についてNIST DB検索を行うと、1位の候補としてUridine-5'-monophosphate, pentakis-TMSがヒットした (図 4 右上)。しかし、この候補の分子式は $C_{24}H_{53}N_2O_9PSi_5$ (計算精密質量は684.23295) であり、 $[M+H]^+$ と推定されるイオン m/z 627.21530と一致しない。NIST DBでヒットしている他の候補についても、 m/z が一致する候補は確認できず、統合解析の結果本成分はNIST DBに未登録であることが確認できた。今回検出された m/z 627.21530は、*N*-メチルウリジル酸にTMSが4つ付加した分子の $[M+H]^+$ である $C_{22}H_{48}N_2O_9PSi_4$ の計算精密質量627.21690と質量誤差 -1.60 mDaで一致するため、本成分は*N*-メチルウリジル酸由来だと考えられた。

また、本成分のEIマスペクトルを用いたAI構造解析結果の上位18候補を図 5に示す。AI構造解析結果では、実測のマスペクトルと構造式から予測したマスペクトルの間の類似度 (AIスコア) が高い順に、構造式候補が羅列している。msFineAnalysis AIのAIライブラリーには、*N*-メチルウリジル酸と同様の分子式をもつ化合物 (異性体) が53成分登録されている。このうち、*N*-メチルウリジル酸の構造式は2位の候補として得られ、AIスコアは819と良好な値を示した。また、実測マスペクトルと予測マスペクトルは、主要なマスピークの m/z 値及び全体的なパターンが良く一致していることが確認できた。また、AI構造解析結果の1位や3位の候補については、*N*-メチルウリジル酸のウラシル環のメチル基の位置のみが異なる異性体である。その他の上位候補の構造式についても、*N*-メチルウリジル酸と類似した構造式が並んでいる。このように、上位候補の多くが*N*-メチルウリジル酸の構造式を反映しており、構造式推定時に有益な情報であった。

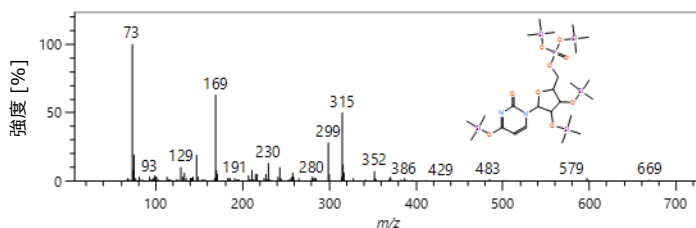
以上の結果から、AI構造解析により*N*-メチルウリジル酸のようにNIST DBに登録されていない未知化合物であっても構造式を推定できることが示された。*N*-メチルウリジル酸は多くのがん細胞にて発現が増加することが知られている新規代謝物で、ウラシルの窒素にメチル基が付加された構造を持ち、RNA修飾由来の代謝物と考えられている化合物である。このように、従来のNIST DB検索のみでは同定が困難であった代謝物に対しても、AI 構造解析を使用することで迅速に構造式推定を行うことができるため、がん研究をはじめとするメタボローム解析での応用が期待される。

● マスペクトル



● NIST DB検索で得られた1位の候補

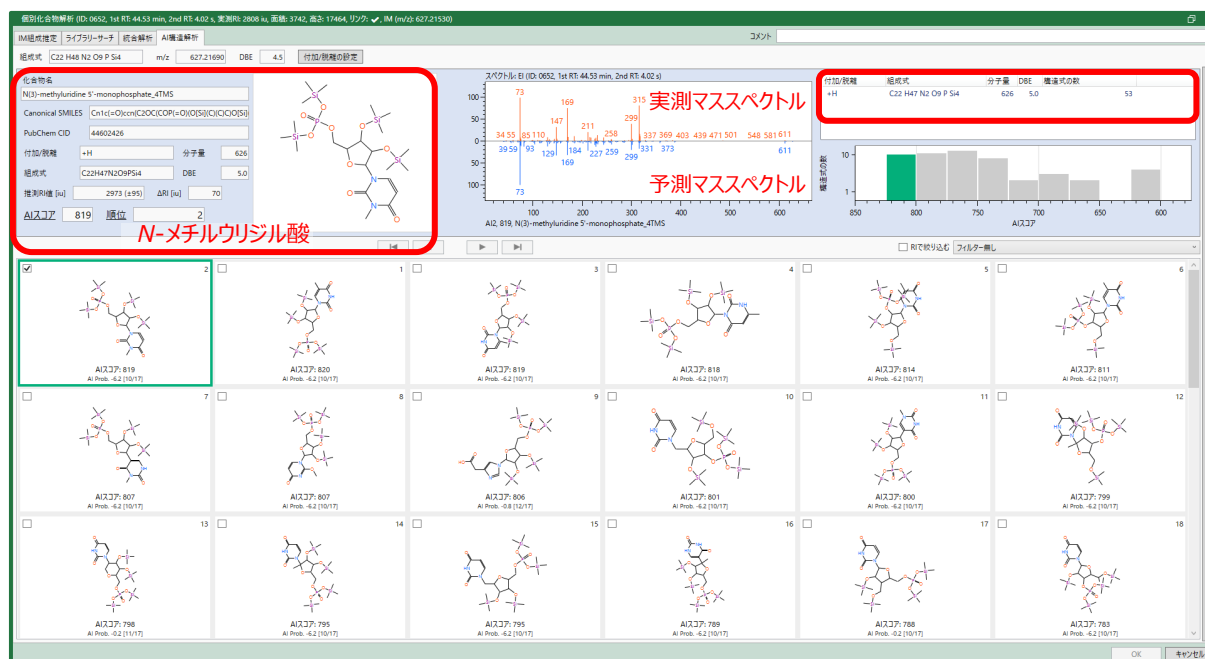
(Uridine-5'-monophosphate, pentakis-TMS, 分子式: $C_{24}H_{53}N_2O_9PSi_5$)



● 統合解析結果

IM組成推定 (m/z : 627.21530)					
組成式	DBE	m/z (計算)	質量誤差 [mDa]	同位体マッチング	カバー率
$C_{22}H_{48}N_2O_9PSi_4$	4.5	627.21690	-1.60	0.70	100

図 4 *N*-メチルウリジル酸のマスペクトルと統合解析結果



構造式候補数
(AIライブラリー
中の構造式数)

構造式の候補 (AIスコア順に羅列)

図 5 AI構造解析結果 (上位18候補)

結論

本MSTipsでは、HeLa細胞中の代謝物をGCxGC-TOFMSで測定し、msFineAnalysis AIを用いて解析した結果について報告した。その結果、アミノ酸や糖、核酸関連物質をはじめとする、多数の代謝物を検出することができた。さらに、構造既知であるもののNIST DBに未登録であるN-メチルウリジル酸についてAI構造解析を行ったところ、有力な候補構造を提示することができ、本手法の有用性が示された。

以上の結果から、GCxGC-TOFMSとmsFineAnalysis AIの組み合わせは、生体試料中の既知代謝物の網羅的な把握に加え、NIST DBをはじめとする市販のマススペクトルライブラリーに未登録の成分について候補構造の取得を可能とすることから、メタボロミクスにおける有用なワークフローであると考えられる。

参考文献

- 1) A. kubo et al, *Mass Spectrom.*, **2023**, *12*, A0120.
- 2) M. Ubukata et al, *Rapid Commun Mass Spectrom.*, 2020, *34*, e8820.
- 3) Z. Lai et al., *Nature Methods*, 2018, *15*, 53-56.

謝辞

HeLa細胞抽出物の提供ならびに測定・解析においてご指導を賜りました東京農工大学の津川裕司教授および倉田明咲様に、深く感謝申し上げます。