

統合定性解析ソフトウェア msFineAnalysis IQ Ver.3におけるAI構造解析機能の紹介

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) を用いた定性解析では、アメリカ国立標準技術研究所 (NIST) が提供する電子イオン化 (EI) マススペクトルライブラリーを用いたNISTデータベース (DB) 検索が広く利用されている。ただし、このDBスペクトルとのマッチファクターのみを指標に定性解析を行うと、複数の有意な候補が得られる場合や、誤った候補が同定結果として選択される場合がある。このような場合、光イオン化 (Photoionization, PI) 法や化学イオン化 (Chemical Ionization, CI) 法をはじめとするソフトイオン化 (SI) 法による分子イオンの確認が有効となる。

弊社では2021年に、ガスクロマトグラフ四重極質量分析計 (GC-QMS) で測定したEI法およびSI法のデータを自動で統合して解析するソフトウェア msFineAnalysis IQ をリリースした。本ソフトウェアでは、EI法で取得したマススペクトルを用いたDB検索と、SI法で取得したマススペクトル中の分子イオンの解析を組み合わせる“統合解析”により確度の高い定性解析結果を得ることができる。ただし、msFineAnalysis IQにおける統合解析は、NIST DB検索を前提としており、DBに未登録の化合物については手動での構造推定が必要であり、未知物質解析には依然として高いハードルがあった。

そこで我々は、GC-MSデータを用いた手動構造推定の困難さの課題解決として、深層学習によるマススペクトル予測を組み込んだ構造推定手法 (以後AI構造解析と称す) を搭載したmsFineAnalysis IQを開発した。AI構造解析は、日本電子製GC-TOFMS専用ソフトウェアmsFineAnalysis AIで開発された技術を基盤としており、この知見を応用することで、GC-QMSによる自動構造推定を実現した。また、AI技術の導入をより明確に示すため、製品名を“msFineAnalysis IQ”として刷新している。

本MSTipsでは、msFineAnalysis IQにおけるAI構造解析機能について紹介する。

統合解析とAI構造解析の関係性

統合解析とAI構造解析の関係性を図1に記載する。上述した通り、統合解析では、EI法で取得したマススペクトルを用いたNIST DB検索と、SI法で取得したマススペクトル中の分子イオンの解析を組み合わせる“統合解析”により確度の高い定性解析結果を得ることができる。

統合解析の結果、その候補がマッチファクター700以下の候補しか得られなかった場合、DB未登録成分と判断され、AI構造解析のフェーズへと進む。

AI構造解析について

AI構造解析機能とは、機械学習の一種である深層学習を活用した未知物質の自動構造推定機能である。

図2に、AI構造解析のワークフローを示す。本解析では、深層学習を用いて構造式からEIマススペクトルを予測するAIモデルを構築し、約2億化合物の構造式に対して予測したEIマススペクトルとRI値をAIライブラリーとしてソフトウェアに内包している (RI値は微極性カラムのみ)。これにより従来のNIST DB検索と同様の操作感で、スペクトルパターンに基づく未知物質の構造推定を可能にした。

なお、AIライブラリーの全2億化合物を対象とした総当たりのDB検索を行うのではなく、統合解析により決定された分子量と一致する化合物のみに候補を絞り込んで検索を行う。これにより、誤同定のリスクを抑えつつ、適切な構造式候補へ迅速に到達できるよう工夫されている。

最終的には、分子量で絞り込んだ候補化合物の予測EIマススペクトルと、実測EIマススペクトルの類似度 (AIスコア) に基づいてスコア計算を行い、構造式候補をAIスコア順に一覧表示する。最終的に得られる結果については次頁に示す。

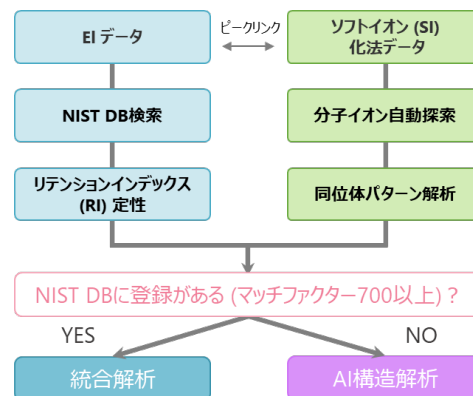


図1 msFineAnalysis IQの解析ワークフロー

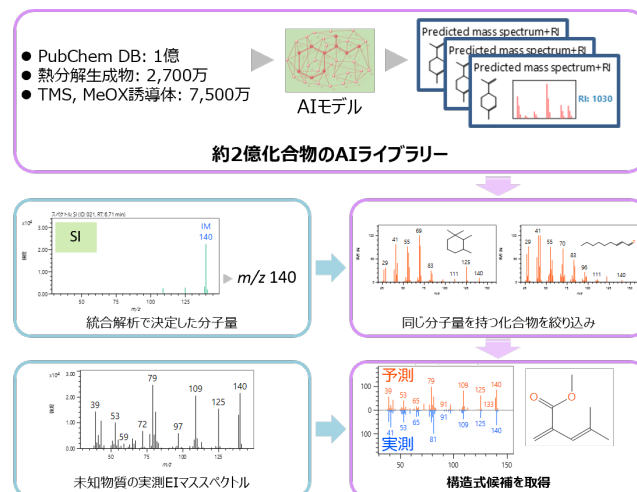


図2 AI構造解析ワークフロー

AI構造解析のGUI

図3に、msFineAnalysis IQによるアクリル樹脂の熱分解生成物に対するAI構造解析結果画面を示す。解析対象成分は、NIST DBに未登録であったダイマー成分である。

本解析結果画面では、画面下部に構造式候補の一覧が表示されており、AI構造解析結果を俯瞰的に確認することができる (1)。各構造式の下には、予測マススペクトルと実測マススペクトルとの類似度を表すAIスコアが表示されており、構造式候補はAIスコアが高い順に羅列している。

また、現在選択されている (緑枠で囲われた候補) の構造式の情報が画面左上に示される (2)。加えて、AIスコアの分布と構造式候補数のヒストグラムにより、選択中の構造式が候補全体の中でどの位置にあるかを確認できる (3)。さらに、実測マススペクトルとAIライブラリー中の予測マススペクトルの比較結果も表示されるため、候補の妥当性を視覚的に評価することが可能である (4)。

本ソフトウェアには、部分構造やモノマー情報、予測RI値によるフィルタリング機能も搭載しており (5, 6)、得られた構造式候補の絞り込みを効率的に行うことができる。構造式の候補数は、AI ライブラリー中の化合物を分子量で絞り込んだ結果、得られた件数を表示する (7)。また、SI法マススペクトルで検出されているイオンがプロトン付加分子 ($[M+H]^+$) である場合などは、その付加/脱離を考慮した候補数も表示できる。

最終的には、得られた構造式候補に加えて、サンプル情報や従来の分析結果から蓄積された知見・ノウハウなどを総合的に併せて判断することで、正しい構造式を選択する。

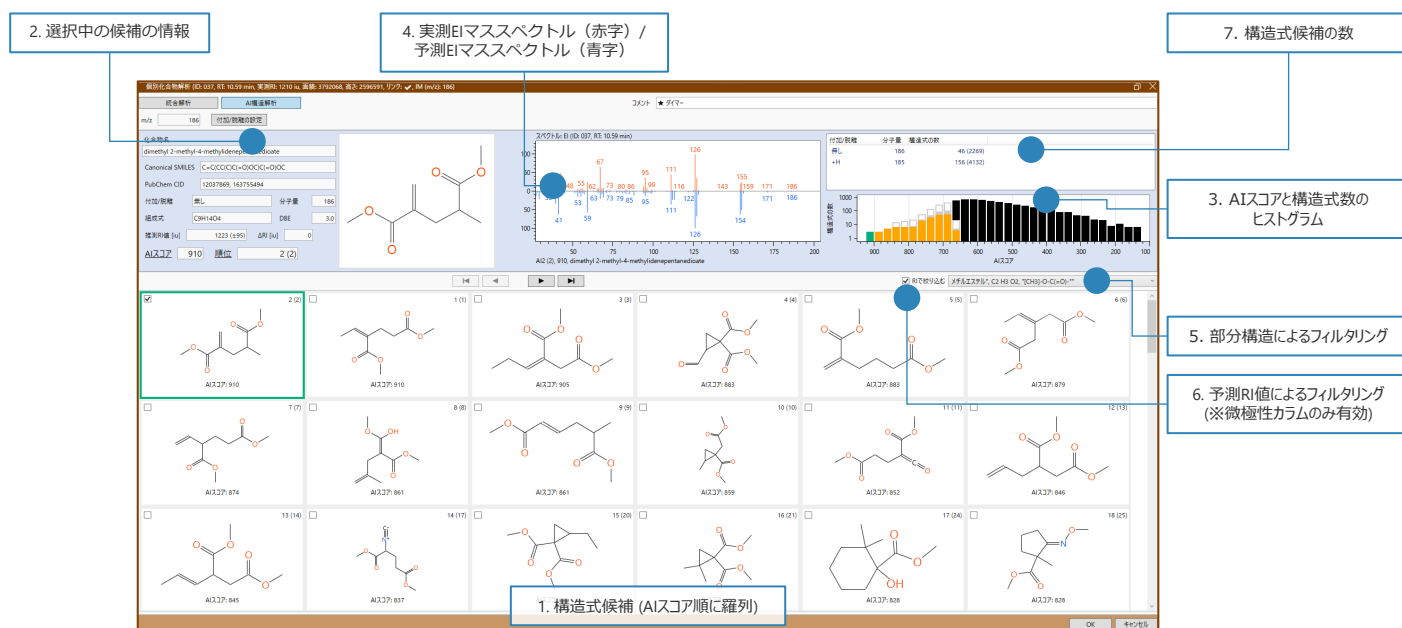


図3 msFineAnalysis IQのGUI

結論

本MSTipsでは、AI構造解析機能を搭載したソフトウェアmsFineAnalysis IQのAI構造解析機能について紹介した。本ソフトウェアは深層学習によるマススペクトル予測を組み込んだ手法により、未知物質の構造推定を自動で行うことができる。従来のNIST DB検索と同様の操作感で解析できるため、複雑なマススペクトルの解釈をユーザーが行う必要はなく、質量分析やAIに関する専門知識がなくても利用できる点が特長である。

msFineAnalysis IQは、GC-MSの定性解析や構造推定を効率化する有用なツールであり、化学・材料・食品・環境など、さまざまな分野における定性解析での活用が期待される。