

未知物質構造解析ソフトウェア msFineAnalysis AI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

ライブラリーデータベース(DB)未登録の化合物の定性分析では、分子イオンやプロトン付加分子を与えやすいソフトなイオン化法と精密質量測定ができる飛行時間型質量分析計(TOFMS)の組み合わせが有用である。電界イオン化(FI)法などのソフトイオン化法にて分子イオンを観測し、更に精密質量に対して組成推定を行うことでDB未登録の未知化合物であってもその分子式を決定できる。弊社ではEI法とソフトイオン化法で得た2つのマススペクトルを用いた定性解析手法を“統合解析”とし、これを搭載したソフトウェアmsFineAnalysisを2018年にリリースした。

TOFMSを用いた統合解析ではフラグメントイオンの組成式も得られるため、部分構造情報を得ることができる。また、部分構造情報と分子式情報を組み合わせることで化合物の構造解析が可能になる。しかし、最終的な構造推定については解析者自身による考察が必要であり、その作業には質量分析や科学に関する知見と多くの時間が求められた。

今回我々は、GC-MSデータを用いた手動構造解析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網羅的な構造解析手法(以後AI構造解析と称す)を搭載したmsFineAnalysis AIを開発した (Figure 1)。AI構造解析手法では上述した統合解析で得た情報と深層学習モデルにより得た予測マススペクトルを用いて構造を推定する。本MSTipsでは、msFineAnalysis AIのAI構造解析機能について紹介する。

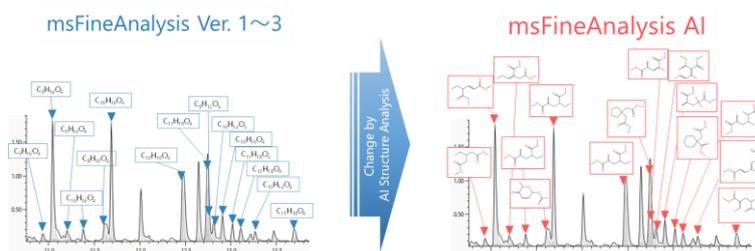


Figure 1 Image of analysis result in msFineAnalysis AI

AI構造解析機能について

AI構造解析機能とは、機械学習と深層学習を相補的に組み合わせた2つのAI(メインAI、サポートAI)を用いて未知物質の構造解析を自動で行う機能である。

Figure 2にメインAIによるAI構造解析のワークフローを示す。メインAI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モデルを構築し、1億個の化合物の構造式から予測されるEIマススペクトルを“AIライブラリー”データベースとしてソフトウェアに内包させ、従来のライブラリーデータベース検索と同様にスペクトルパターンによるデータベース検索機能を実装した。また、メインAIでは1億個すべての予測マススペクトルを用いた総当たりのデータベース検索ではなく、統合解析にて一意に決定した分子式による構造式候補の絞り込みを行っており、より正しい構造式が迅速に得られるよう工夫されている。分子式で絞り込まれた予測EIマススペクトルと、実測のEIマススペクトルを比較してスペクトルパターンによるスコア計算を行い、構造式候補をスコア順に掲載する。得られた構造式候補と、サンプル情報、今までの分析結果の知見・ノウハウなどを合わせ、最終的に正しい構造式を選択する。

Figure 3にサポートAIによる部分構造予測のワークフローを示す。サポートAIは実測マススペクトルを解析し、部分構造の有無を予測する。精密質量から得られたフラグメントイオンやニュートラルロスの組成式を解析し、メインAIで提案された構造情報の解釈を補助することが可能である。

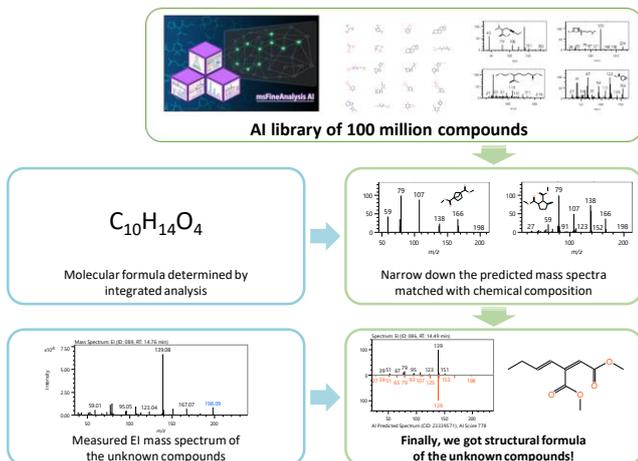


Figure 2 Main AI workflow

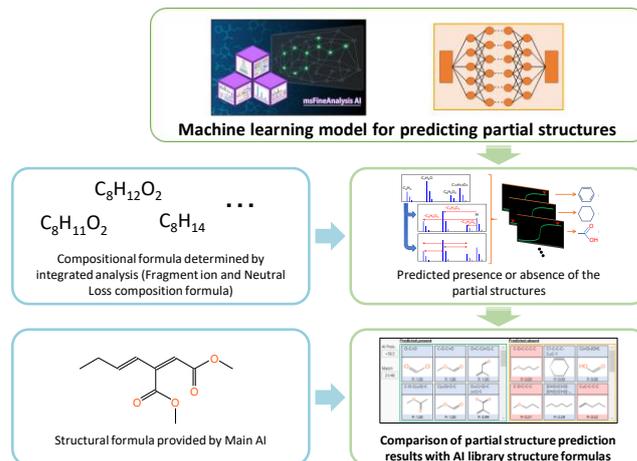


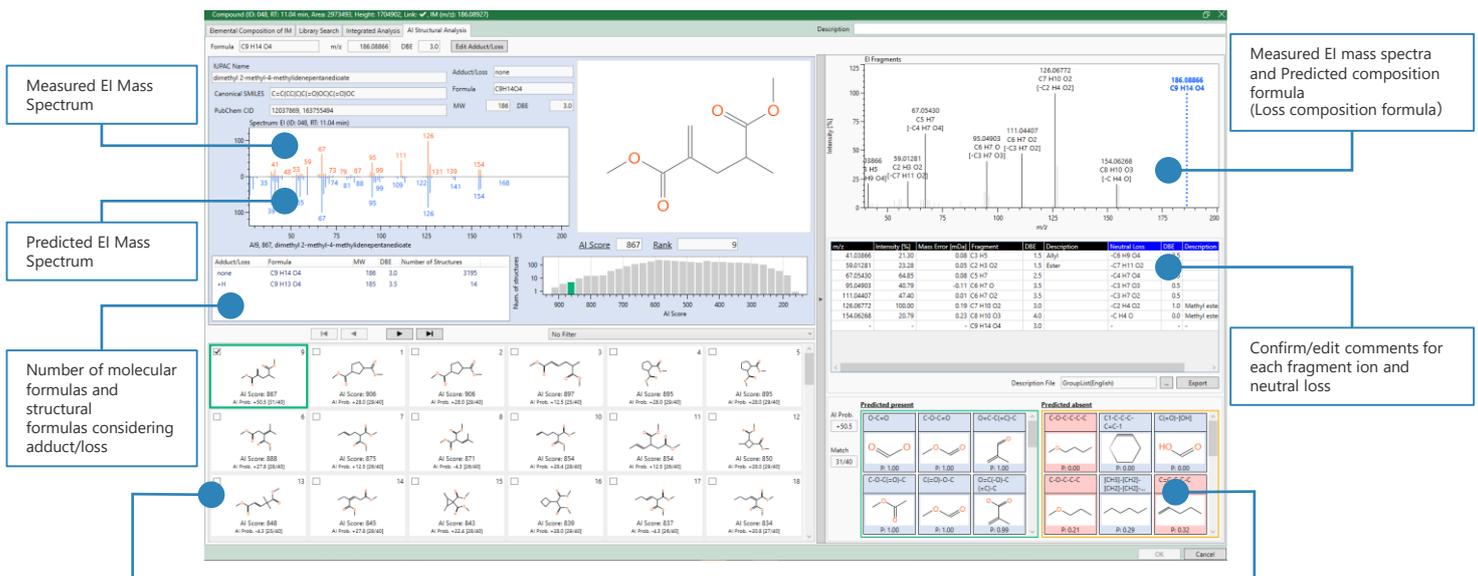
Figure 3 Support AI workflow

AI構造解析のGUI

Figure 4にmsFineAnalysis AIによるアクリル樹脂のAI構造解析結果画面を示す。解析対象成分はNISTライブラリーデータベースに未登録であったダイマー成分である。解析結果画面の左側はメインAIによる解析結果、右側はサポートAIによる解析結果が示されており、データベースに未登録な未知物質においても詳細な構造情報が取得できている。

メインAIの解析結果では、画面下部に予測構造式の一覧が示されており、AI構造解析結果を俯瞰して確認することが可能である。各構造式の下には、AIライブラリーと実測マスペクトルの類似度を表すAIスコアを表示している。さらに選択されている構造式の情報が画面上部に掲載されている。AIスコアのヒストグラムにより選択している候補がどの程度のランクか確認することができるほか、実測マスペクトルとAIライブラリーの予測マスペクトルを比較して表示することが可能である。また、部分構造やモノマーによるフィルタリング機能も搭載しており、後述するサポートAIで予測された部分構造の有無を反映した構造解析結果を得ることができる。

サポートAIの解析結果では、画面下部に部分構造の存在予測結果が示されており、左列に含まれている、右列に含まれていないと予測した部分構造が示されている。背景が青の部分構造はメインAIで選択している構造式と合致するもの、赤の部分構造は合致していないものを表現している。画面上部には実測マスペクトルと各フラグメントイオン・ニュートラルロスの推定組成式が示されている。各推定組成式に対するコメントを確認・編集することも可能である。



Structural Formula Prediction by Main AI

- Displays ranked structural formulas in list.
- Selecting a structural formula updates the information that is displayed.
- Below each structural formula is an AI score that indicates the match percentage between structural formula and mass spectrum.

Partial Structure Prediction by Support AI

- Displays the predicted partial structure information
- Partial structure predicted and present are on the left, predicted and absent are on the right.
- Those with a blue background are the partial structures that match the selected structural formulas. Those with a red background are those that do not match.

Figure 4 GUI of msFineAnalysis AI

結論

本MSTipsでは新たに開発したAI構造解析手法を搭載したソフトウェアmsFineAnalysis AIのAI構造解析機能について紹介した。本ソフトウェアは機械学習と深層学習を相補的に組み合わせた2つのAI（メインAI、サポートAI）を用いて未知物質の構造解析を自動で行う。複雑なマスペクトルの解釈はソフトウェアが自動で行うため、質量分析やAIに関する知識が無くても解析可能である。GC-MS定性分析／構造解析における有益なツールであり、今後様々なGC-MS定性分析シーンにおける活用が期待される。

Copyright © 2022 JEOL Ltd.
このカタログに掲載した商品は、外国為替及び外国貿易法の安全輸出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。

