

Technical Review

高精度MSⁿ測定データを用いた組成推定ソフトウェアの開発

*猪鼻祐介、山口真一、向 紀雄、平野一郎

Development of a unique software tool to predict empirical formulae utilizing accurate mass MSⁿ measurementsYusuke Inohana^{*}, Shinichi Yamaguchi, Norio Mukai and Ichiro Hirano*Analytical Applications Department, Analytical & Measuring Instruments Division, Shimadzu Corporation**1, Nishinokyo-kuwabaracho, Nakagyo-ku, Kyoto 604-8511, Japan***Abstract**

Since a hybrid quadrupole ion trap TOF system generates MS and MSⁿ spectra, the prediction tool set can use the fragment ion data to reduce the list of potential candidates and remove invalid formulae. Although the main variables affecting the returned scores for potential candidates are mass accuracy and isotopic distribution for the MS spectra, results presented will also highlight the importance of fragment ion data. The fragment ion data helps in filtering candidates by allowing for the ability to remove invalid formulae thus helping to identify the correct empirical formula with a higher degree of confidence without the need for using sub-ppm mass accuracy data. This paper will also show the importance of advanced modeling tools for fitting experimental and theoretical mass accuracy and isotope distribution data applied to empirical formula prediction using a diverse chemical library as a test system.

Keywords : quadrupole ion trap, TOF, MSⁿ, mass accuracy, isotopic distribution, empirical formula

緒言

高精度質量分析法は、飛行時間型質量分析計 (TOF)、QqTOFやIT-TOFといったTOFベースのハイブリッド質量分析計およびFTMSといった数々の技術基盤の基で達成可能なものとなってきている。これらの技術基盤は、医薬品分野における代謝物同定や合成化合物の組成式確認などに大きな役割を果たしている[1, 2]。このような中、測定で得られた精密質量データと理論的な同位体分布および化学的ルールを組み合わせ、帰属した実験式の確かさを高めるソフトウェア

ツールの応用がなされてきており、MS²のデータを活用する手法についてもいくつかの報告例がある[3, 4]。本稿では、今回新たに開発した、高精度MS情報だけでなく高精度なMSⁿデータも考慮できる組成式推定ソフトウェアについて述べる。

開発背景・概要

本ソフトウェアの開発にあたって、まず高精度なMSⁿデータを組成推定に応用する有効性を示す一例として、質量精度

〒604 8511 京都府京都市中京区西ノ京桑原町1
株式会社島津製作所 分析計測事業部 応用技術部
TEL : 075-823-1089 FAX : 075-841-9326
E-mail : ino-yu@shimadzu.co.jp

と組成候補数の関係について検討した (Figure 1)。ここでは、以下の計算モデルを用いた。元素は、炭素、水素、窒素、酸素の4元素を用い、Cの数をn、Hの数を $n - 2n + 2$ 、OおよびNをそれぞれ $0 \sim n/5$ とし、 $n = 140$ まで計算し、10、1および0.1 ppmの質量精度を用いて、整数質量間における識別不可能な組成候補がいくつ存在するかを求めた。

実際に計算すると、10 ppmでは200付近から一義的に決定することは難しくなり、1 ppmと0.1 ppmでは、550付近から、その傾向が見られる。1 ppmと0.1 ppmの始点が同じ理由は、はじめに出てくる質量の近いもの同士の差が非常に小さなものだからである。(Ex: $C_{35}H_{41}O_5$: 553.29540, $C_{35}H_{35}N_7$: 553.29539 $0.01 \text{ mu}/500 \text{ u} = 0.02 \text{ ppm}$) 若干、乱暴な考察ではあるが、この結果から質量が倍になると候補の数は10倍程度になり、質量精度が10倍になると約2倍の質量の組合せが決定できるようになると考えられる。また、今回の計算モデルには、硫黄やリンなどの元素を考慮すると、それぞれの質量精度での限界質量は低くなり、曲線の立ち上がりは急になることが予想できる。またSunghwanらも、C, H, N, O, Sの元素の組合せでは500 uの質量を持つ分子の組成を唯一に決定するには、0.1 mu (200 ppb)の精度が必要であると報告している [5]。

Figure 2にAccurate- MS^n コンセプトを示す。組成推定は目的成分の質量が小さく、目的成分の元素が特定でき、かつ測定値の質量精度が高ければ高いほど、その候補数が少なくな

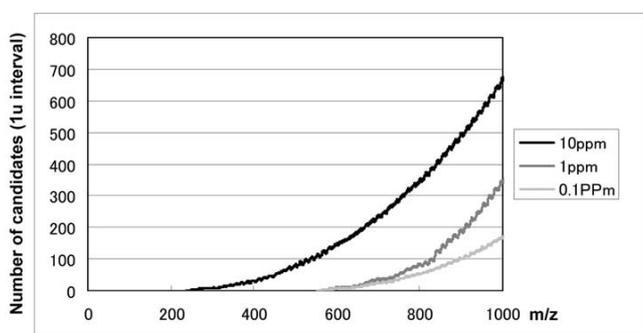


Figure 1. Relationship of mass accuracy and the number of the candidates

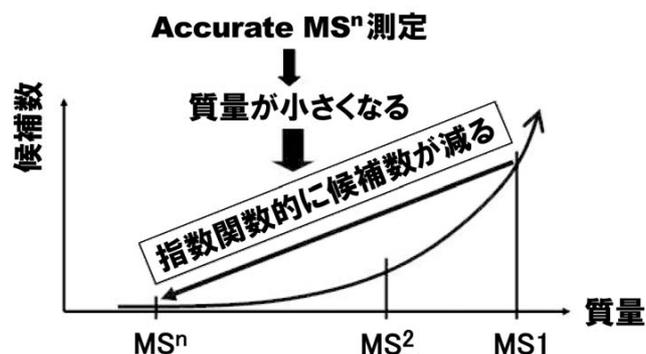


Figure 2. Accurate- MS^n concept

Table 1. Overview of software functionalities and parameters

Stage	概要	パラメータ詳細
1	MSスペクトルデータ	
	分子イオンもしくは付加イオンと質量精度を考慮する。	正しい組成式を同定する可能性を上げるために、不飽和度、HC比、窒素ルール、アダクトイオン情報、質量精度閾値といった汎用的なフィルタ機能を持つ。
2	MS ⁿ スペクトルデータ	
	分子イオンから生成したMS ⁿ プロダクトイオンを考慮し、それぞれのプロダクトイオンにおける理論的なcomplementation (プリカーサイオンとプロダクトイオンの差) を計算する	MS ⁿ で得られるプロダクトイオン情報を用いて、対応するプリカーサイオンの組成候補から、プロダクトイオン情報に一致しない組成候補を除外する。この考え方はそれぞれのステージのMS ⁿ データで繰り返し使用される。例えば、MS ² で得られたプロダクトイオンの組成候補を用いて、MS ³ で選択したプリカーサイオンの組成候補をフィルタリングする。(MS ³ イオンは、MS ² イオンをフィルタリングし、MS ² イオンは、MS ¹ での分子イオンのフィルタリングに用いられる。)
3	同位体分布データ	
	可能性のある組成候補の理論的な同位体分布スペクトルを生成する。	最小二乗法を用いて、測定データの同位体分布と候補組成式の理論上の同位体分布を比較する。尤度スコアリングアルゴリズムは、理論的同位体分布への近似度、質量精度およびMS ⁿ データを用いフィルタした候補リストを考慮している。対数変換した順位スコアにより組成候補を信頼性の順にリストアップする。(スコア値: 0-100)

り、信頼性が高くなる。3次元イオントラップ - 飛行時間型質量分析計LCMS-IT-TOF (島津製作所) が実現した高精度でのMSⁿ測定と高精度なMSⁿデータも考慮できる今回の組成推定ソフトウェアでは、MSⁿで観察される「より質量の小さなプロダクトイオン」および「complementation (プリカーサイオンとプロダクトイオンの差)」をはじめに組成推定し、その結果を親イオンの推定に反映させることにより、効果的に候補数を絞り込むことができる。本ソフトウェアの概要をTable 1に示す。

実験

今回開発した組成推定ソフトウェアの評価には、3次元イオントラップ - 飛行時間型質量分析計LCMS-IT-TOF (島津製作所) によって得られた2つの医薬関連化合物データ (化合物AおよびB) を用いた。実験に用いたLCMS-IT-TOF装置の外観とイオン光学系をFigure 3に示す。

結果と考察

化合物Aの構造とMSⁿスペクトルをFigure 4に示す。MS¹データのみ依存する場合、質量精度と分解能が正しい組成式を同定する決め手となる。組成式候補の信頼性を向上させるには、測定データと理論的な同位体分布パターンを比較することが重要となる。この場合、測定化合物の組成式は「 $C_{28}H_{39}N_5O_4$ 」であり、ソフトウェアは同じ組成式を第一候補として予測している (Figure 5)。

Figure 6にMSⁿデータをもとにした組成推定結果を示した。高質量精度のプロダクトイオン情報を用いることによって、これらと一致しない候補を除外することが可能となり、

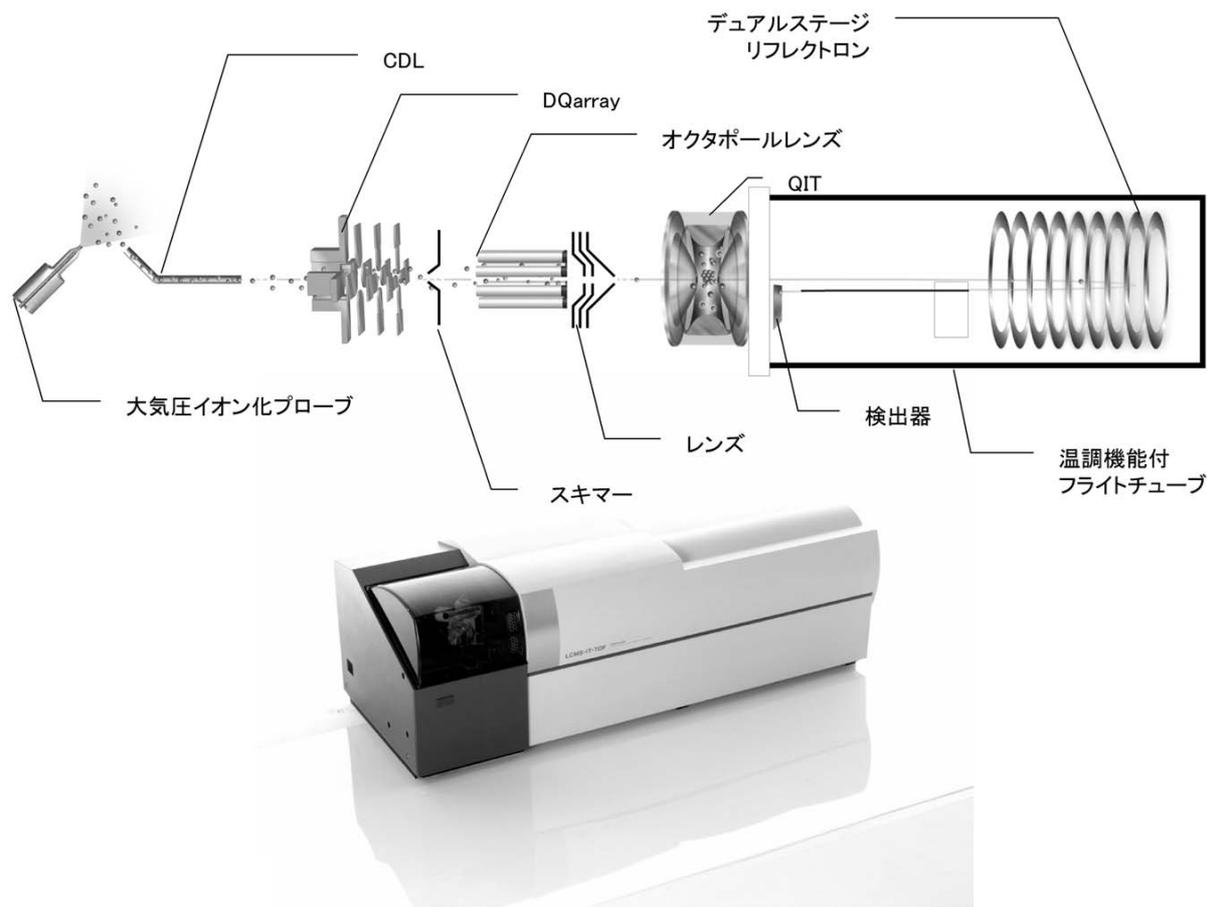


Figure 3. The product overview and ion path of LCMS-IT-TOF

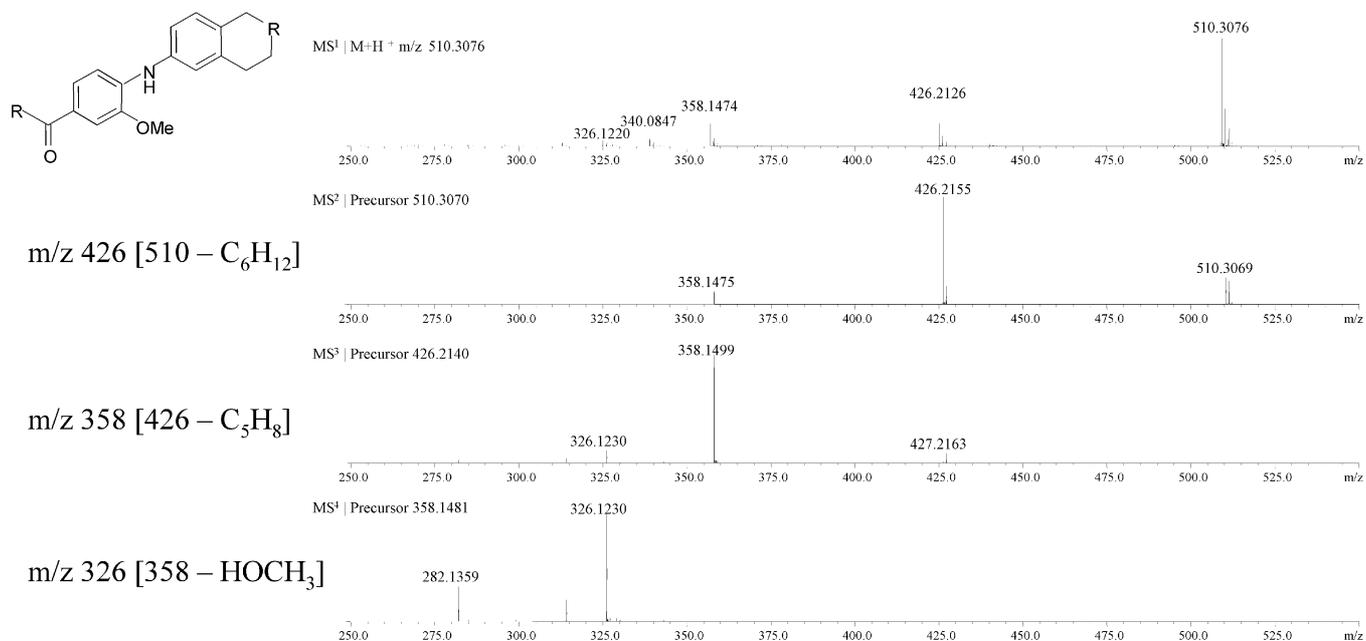
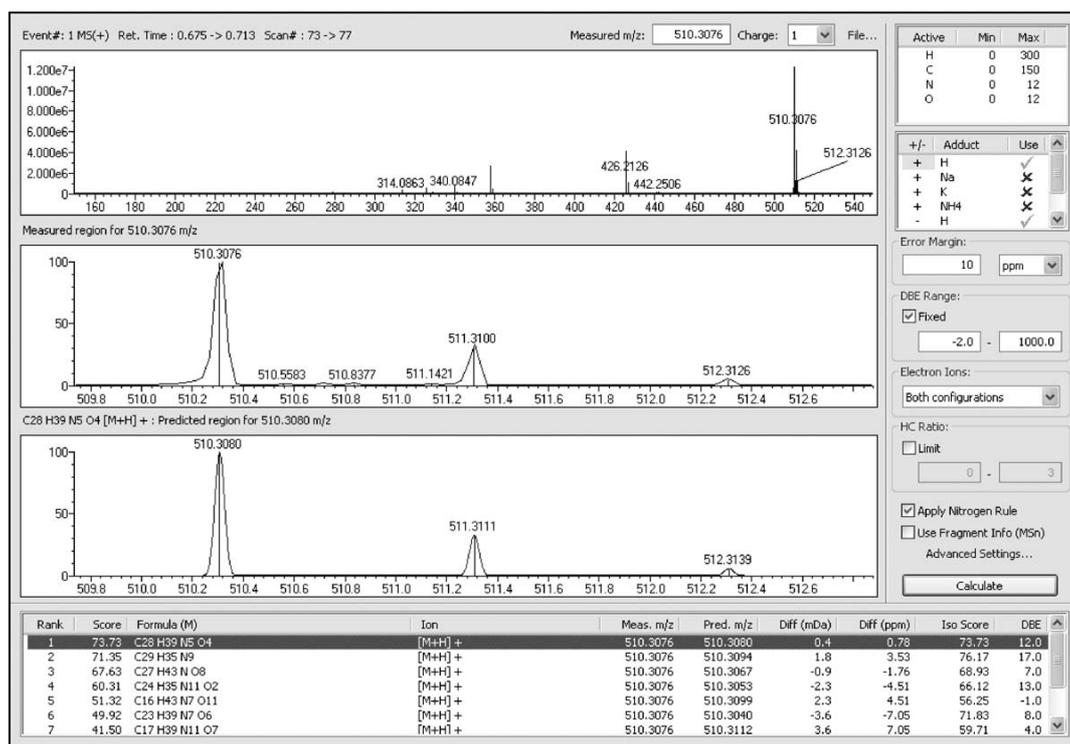
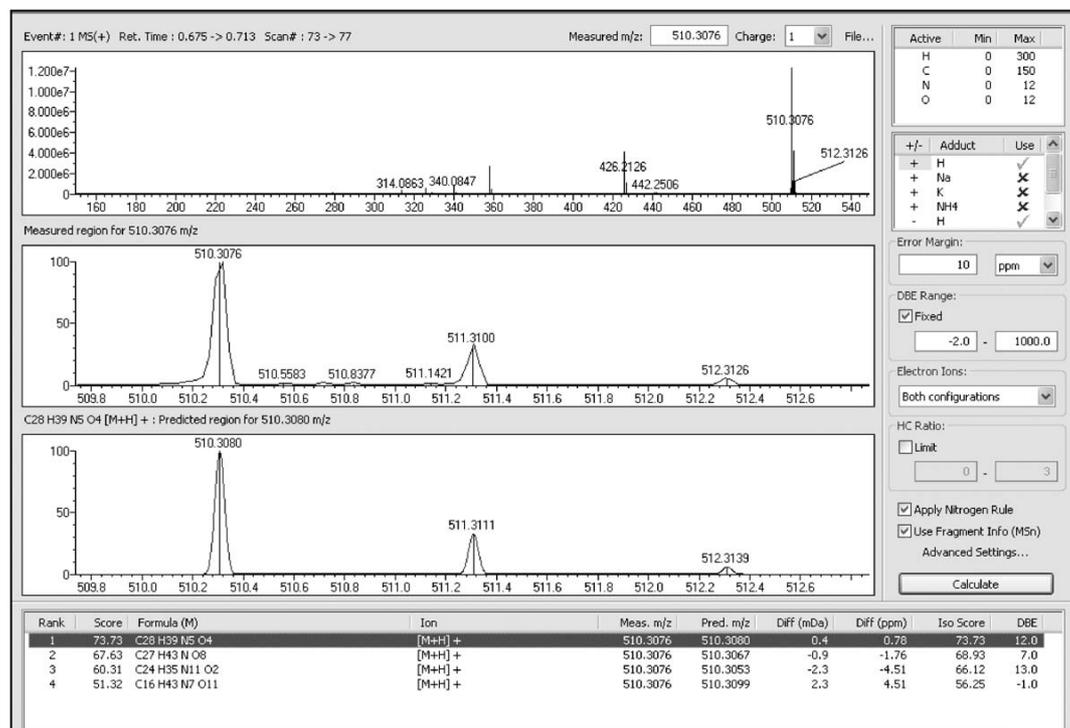


Figure 4. Structure and MSⁿ spectra of compound A

Figure 5. Result of empirical formula prediction (using only MS¹ data)Figure 6. Result of empirical formula prediction (using MSⁿ data)

正しい組成式を決定できる。合わせてcomplement ionの同定結果をFigure 7に示す。

次に、Figure 8に示す構造を持つ化合物BのMSⁿスペクト

ルから、同様に組成推定を行った。MS¹データの場合、正しい組成候補である「C₁₉H₂₁O₃N₃S」は第16位にランキングされているが(Figure 9)、MSⁿデータを用いることによ

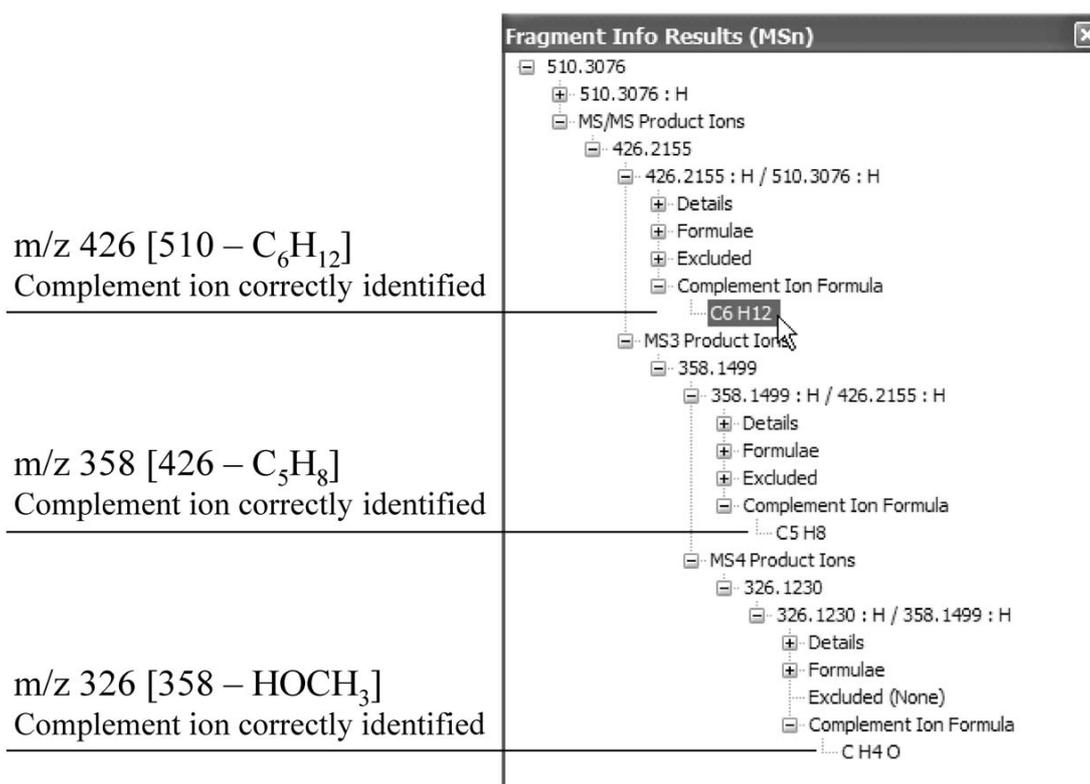
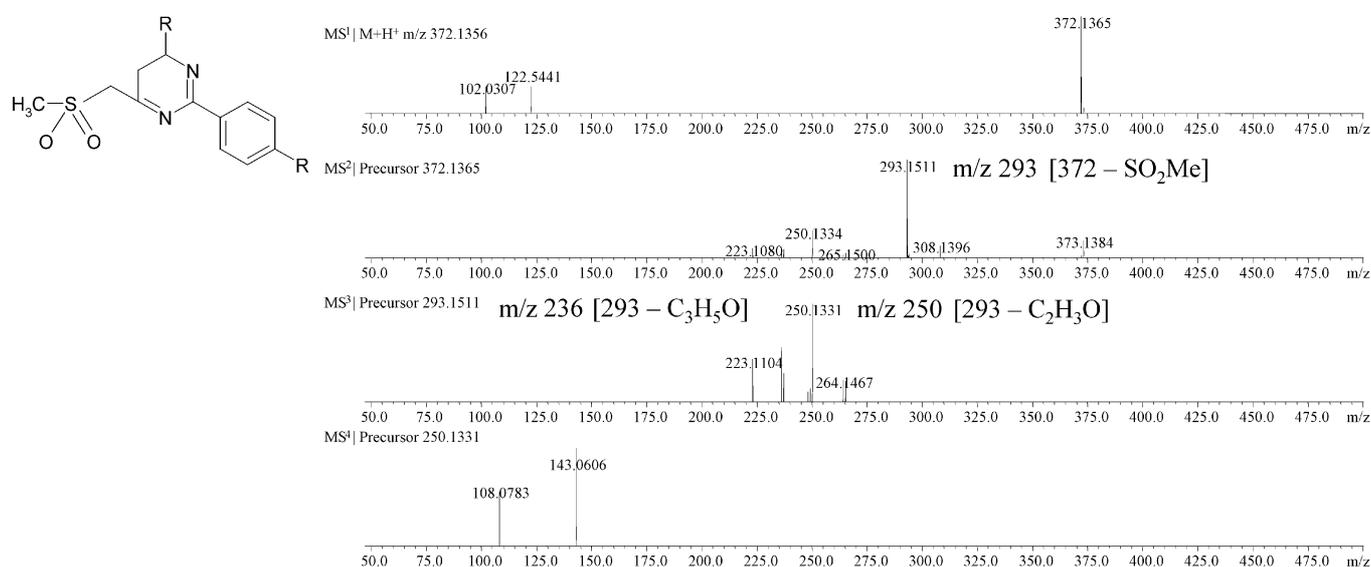


Figure 7. Identification result of complement ions

Figure 8. Structure and MSⁿ spectra of compound B

て、他の組成候補を除外することが可能となり、正しい組成候補である「C₁₉H₂₁O₃N₃S」が一義的に同定できている (Figure 10)。この場合、MSⁿデータによって観測された比較的低い m/z 領域でのプロダクトイオンの情報が第一候補組成式の信頼性を向上させるのに大きな役割を果たしている。

結論

精密質量データから組成式を推定するプログラムは、一般的に化学ルールや同位体パターンフィッティングを考慮しているが、我々はそれらに加え、MSⁿデータの応用が候補数の減少と正しい組成式を推定する可能性を向上させることを紹介した。正しい候補を同定する際の蓋然性 (確実性の度合い) に、明らかに影響する要因はいくつか存在するが、精密

Using only MS¹ data the true candidate was identified as hit number 16

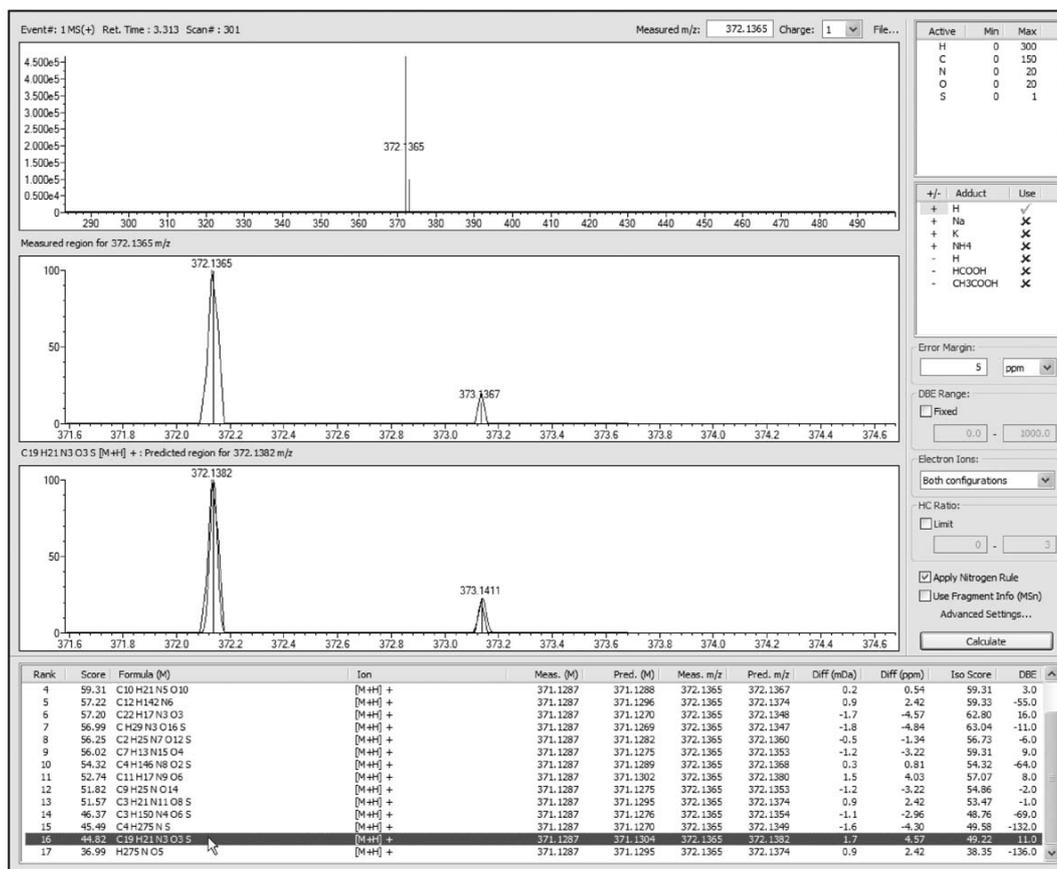


Figure 9. Result of empirical formula prediction (using only MS¹ data)

質量がもっとも重要な要因である。しかしながら、MSⁿによるプロダクトイオン情報と complement ion (プリカーサイオンとプロダクトイオンの差) 情報を擬似分子イオン情報に加えて考慮することを可能とするソフトウェアツールを提供することにより、このアプローチが、ただ単にMS¹データのみを用いる手法に比べ、より高い蓋然性で正しい候補組成式を同定することが可能であることを証明している。

謝辞

本研究にあたり試料のご提供をいただいた、アストラゼネカ (英国) Richard Gallagher博士に深謝いたします。

参考文献

- [1] Michelsen P.; Karlsson A.A, *Rapid Commun. Mass Spectrom.* **1999**; *13*: 2146
- [2] Thurman E.M.; Ferrer I.; Parry R., *J.Chromatogr. A* **2002**; *957*: 3
- [3] Grange A.H.; Osemwengie Li; Brilis G.M.; Sovocool G. W., *Environ. Forensics* **2001**; *2*: 61
- [4] Suzuki S., Ishii T.; Yasuhara A.; Sakai S., *Rapid Commun. Mass Spectrom.* **2005**; *19*: 3500
- [5] Sunghwan K.; Ryan P.R; Alan G.M., *Int. J. Mass Spectrom.* **2006**; *251*: 260

Using MSⁿ data the true candidate was identified as hit number 1

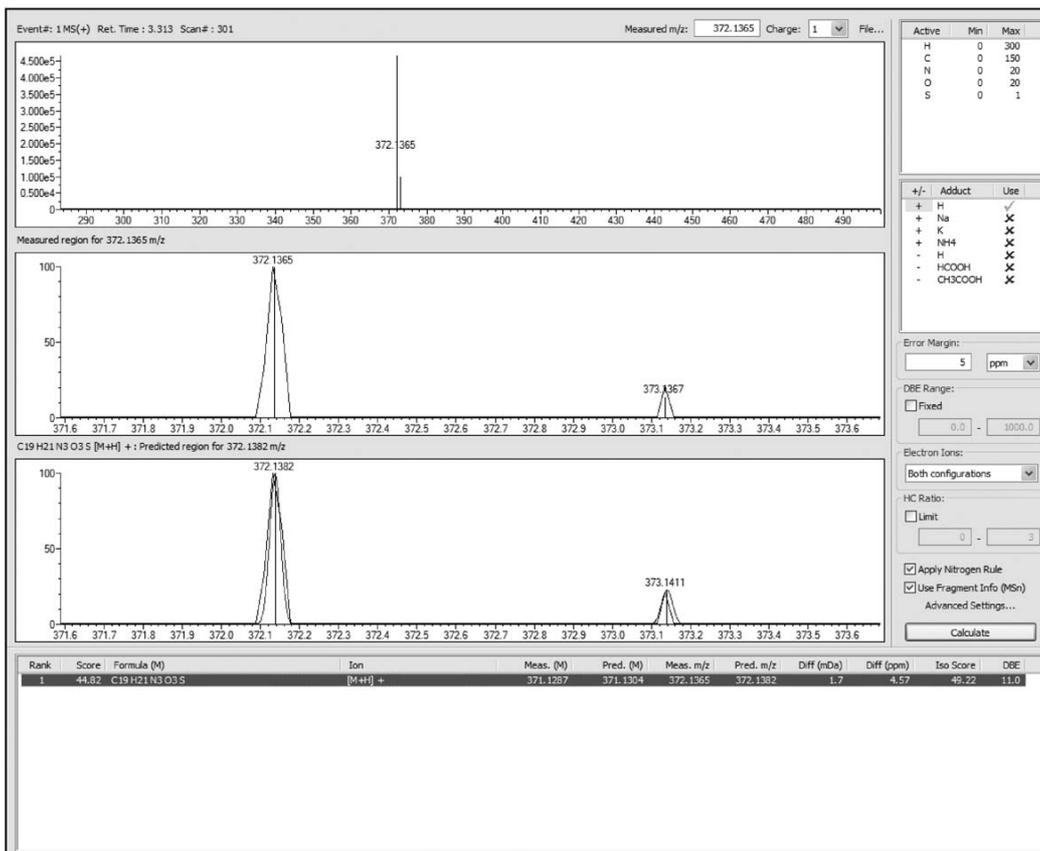


Figure 10. Result of empirical formula prediction (using MSⁿ data)