

DOI:10.7524/j.issn.0254-6108.2023071002

CSTR:32061.14.hjhx.2023071002

杨伦, 吴仕希, 付庆龙. 基于超高分辨率质谱技术的溶解性有机质双电荷离子研究[J]. 环境化学, 2024, 43(12): 4059-4071.

YANG Lun, WU Shixi, FU Qinglong. Elucidation of doubly charged ions in DOM by ultrahigh-resolution mass spectrometry[J]. Environmental Chemistry, 2024, 43 (12): 4059-4071.

基于超高分辨率质谱技术的溶解性有机质双电荷离子研究*

杨 伦 吴仕希 付庆龙**

(中国地质大学(武汉), 环境学院, 武汉, 430078)

摘 要 溶解性有机质(DOM)是各类环境介质中广泛存在的有机混合物, 参与了表生地球系统的诸多生物地球化学和环境化学过程. DOM的傅里叶变换离子回旋共振质谱(FT-ICR MS)谱图仅考虑单电荷离子峰, 但对双电荷离子峰的研究十分有限. 本研究结合FT-ICR MS技术, 提出基于单电荷离子前体的双电荷离子筛选方法, 并将其应用于各类环境介质中DOM的FT-ICR MS谱图中多电荷离子的化学多样性分析. 结果表明, 相比传统仅考虑 ^{13}C 同位素峰的方法, 基于单电荷离子前体的双电荷离子方法能够有效提高DOM的FT-ICR MS谱图中双电荷离子峰的识别(提高44.90%). 相比单电荷离子, 双电荷离子拥有更高的O/C、DBE(最小双键当量)、 AI_{mod} (修正芳香性指数)和NOSC(碳的名义氧化态)值以及更低的H/C值, 表明大部分双电荷DOM是富羧基的脂环分子. 同时, 双电荷离子的数量与样品类型有关, 土壤DOM中的双电荷离子数量更高. 具体来说, 各类样品中双电荷离子的数量和占比如下: 海洋水DOM(540, 0.18%)、天然淡水DOM(3486, 0.81%)、土壤DOM(6334, 4.06%)和废水DOM(41, 0.13%). 这表明单电荷离子峰是各类环境介质DOM的FT-ICR MS谱图中的主要离子峰. 本研究为DOM的化学多样性研究提供了一种新的数据分析方法, 首次系统性地揭示了不同类型DOM中多荷离子的性质, 并证实了双电荷离子对不同类型DOM的超高分辨质谱(UHR-MS)数据分析的影响相对有限.

关键词 双电荷离子, 傅里叶变换离子回旋共振质谱, 天然有机质, 前体离子.

中图分类号 X-1; O6 文献标识码 A

Elucidation of doubly charged ions in DOM by ultrahigh-resolution mass spectrometry

YANG Lun WU Shixi FU Qinglong**

(China University of Geoscience(Wuhan), School of Environment Studies, Wuhan, 430078, China)

Abstract Dissolved organic matter (DOM) is a heterogeneous organic mixture ubiquitous in various environmental media, participating in many biogeochemistry and environmental chemical processes in the supergene earth system. Despite the extensive occurrence of single-charged ion peaks in the mass spectra of DOM by Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS), information regarding their double-charged ion peaks is very limited. In this study, a doubly charged ions algorithm based on their singly charged precursors was developed to characterize the chemodiversity of multiply charged ions in the FT-ICR MS spectra of DOM in various environmental media. Results showed that compared with the traditional method only

2023 年 7 月 10 日收稿(Received: July 10, 2023).

* 国家自然科学基金(42107484)资助.

Supported by National Natural Science Foundation of China (42107484).

** 通信联系人 Corresponding author, E-mail: fuqinglong@cug.edu.cn

considering ^{13}C isotopic peaks, the doubly charged ions method based on singly charged precursors could effectively improve 44.90% number of doubly charged ion peaks in the FT-ICR MS spectra of DOM. The doubly charged ions had higher values of O/C, minimum double-bond equivalence (DBE), modified aromaticity index (AI_{mod}), nominal oxidation state of carbon (NOSC) and lower values of H/C than those for singly charged ions, indicating that most of the doubly charged DOM were carboxyl-rich alicyclic molecules. Moreover, the number of doubly charged ions was dependent on sample type. Soil DOM was most abundant in the doubly charged ions. The number and percentage of doubly charged ions in various samples were in the order of: marine DOM (540, 0.18%), natural fresh water DOM (3486, 0.81%), soil DOM (6334, 4.06%) and waste water DOM (41, 0.13%). These results suggested that the singly charged ion peaks were dominant in the FT-ICR MS spectra of DOM for various environmental matrices. In addition to providing a novel data analysis method for FT-ICR MS data analysis, for the first time, this study reveals the properties of multiply charged ions in different types of DOM and confirms the neglectable effect of doubly charged ions on ultrahigh-resolution mass spectrometry data analysis for different types of DOM.

Keywords doubly charged ions, FT-ICR MS, natural organic matter, precursor ions.

溶解性有机质(Dissolved organic matter, DOM)是指能溶解于水且能透过 $0.45\ \mu\text{m}$ 滤膜的复杂有机混合物^[1-2], 其广泛存在于天然水体、土壤和大气等生态系统中, 对于维持全球碳循环具有十分重要的意义^[3-5]. 近年来, 傅里叶变换离子回旋共振质谱(Fourier transform ion cyclotron resonance mass spectrometry, FT-ICR MS)因其超高的分辨率(分辨率大于 200000)和质量精确度(质量误差小于 10^{-6})而广泛应用于分子水平的 DOM 化学多样性研究^[1, 3, 6]. DOM 的 FT-ICR MS 谱图中通常含有数千乃至上万个离子峰, 各个离子峰的分子式匹配和单双电荷状态的确定是 DOM 超高分辨质谱(Ultrahigh-resolution mass spectrometry, UHR-MS)峰数据分析的难点. 针对 DOM 分子式匹配的难题, 目前有一系列工具可用于分子式自动匹配, 包括仪器软件(如 DataAnalysis 和 Composer)、自动化程序(Formularity^[7]、UltraMassExplorer^[8] 和 TMDS^[9] 等)以及脚本(MFassignR^[10]、CIA^[11]、SQ^[12] 和 TRFu^[13] 等).

DOM 质谱峰电荷状态的确定使得 DOM 的高分辨质谱峰更加复杂. DOM 分子中通常含有大量的含氧官能团, 如羧基和羟基^[14, 15], 这使得 DOM 在电离过程中易形成多电荷离子, 因此确定峰的电荷状态也成为了一大困扰^[16]. 针对 DOM 的 UHR-MS 离子峰电荷状态的确定, 早在 1995 年 Senko 等^[17] 就根据单同位素峰($^{12}\text{C}_n$ 峰)和 ^{13}C 同位素峰($^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰)确定了质谱中的物质都带单个电荷. Brown 等^[18] 根据 $^{12}\text{C}_n$ 峰和 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰确定了单电荷离子, 发现所观察到的离子均为单电荷离子并提出无法排出多电荷离子的存在. Leenheer 等^[19] 在质谱上观察到的单个腐殖质离子被认为是带多个电荷, 而据其它文献报道, 电喷雾离子源(Electrospray ionization, ESI)质谱中 DOM 离子只带一个电荷^[20, 21]. Kujawinski 等^[22, 23] 认为在电喷雾正电荷模式[ESI(+)]质谱中只有极少数腐殖质离子带多个电荷, 而在电喷雾负电荷模式[ESI(-)]质谱中可能包含多个多电荷离子. Stenson 等^[20] 认为低分子量腐殖酸在 ESI(-)质谱中更容易电离形成多电荷离子, 高分子量离子大多带一个电荷. 在处理多电荷离子的问题上, 大多数学者直接将质谱中的离子峰默认为单电荷离子峰^[18, 22, 24, 25], 而部分学者以 $^{12}\text{C}_n$ 峰和 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰之间的质量差来判断多电荷离子的存在^[2, 9, 26]. 这就导致双电荷离子的存在与 ^{13}C 同位素峰存在因果关系, 然而当不存在 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰时, $^{12}\text{C}_n$ 峰也有可能是以双电荷离子峰的形式存在. 此外, 传统方法没有考虑含 N 的双电荷离子的情况. 根据氮规则和 ESI(-)模式下由天然有机质离子峰的经验规则(即天然有机质的离子峰常位于名义质量至名义质量+0.3 范围内)^[27] 可知: 含有奇数个 N 的双电荷离子 $^{12}\text{C}_n$ 峰会出现在质谱图上名义质量+0.5 Da 的位置, 而 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰则会出现在质谱图上名义质量+1 Da 的位置. 因此含有奇数个 N 的双电荷离子将会使双电荷离子的确定更加复杂化, 这也给传统 ^{13}C 同位素方法带来了新的挑战.

基于此, 本研究结合 FT-ICR MS 技术, 提出了一种基于单电荷离子前体的双电荷离子分子式匹配方法, 并进一步探讨了该方法对 DOM 的 UHR-MS 谱图中双电荷离子峰的识别效果以及不同类型

DOM 中多电荷离子的性质, 加深了对 DOM 化学多样性的认识.

1 材料与方法(Materials and methods)

1.1 试剂与样品制备

本研究所使用的玻璃瓶均在稀盐酸溶液(体积比为 5%)中浸泡 24 h 后, 然后用超纯水和甲醇(LC-MS 纯)反复清洗^[28]. 实验步骤: (1)用电子天平(购自波兰 RADWAG, AS 60/220.R2)准确称取 10.0 mg SRNOM 标准品(购自国际腐殖质学会, 2R101N), 将其溶解在 50 mL 超纯水中制备成 200 mg·L⁻¹ 的储备液; (2)取 5 mL SRNOM 储备液, 采用 0.45 μm 的纤维滤膜过滤; (3)将上述 SRNOM 储备液与甲醇溶液(LC-MS 纯, 购自 Fisher Scientific 公司)等体积混合.

此外, 本研究还获取了 127 个样品谱图, 数据来源于文献^[29]. 该 127 个样品由 5 个不同类型的来源组成, 分别为海水 DOM(包含 16 个海洋水和 23 个海岸带海水样品)、自然淡水 DOM(包含 24 个淡水、15 个地下水、5 个农田水、14 个 SRNOM 和 2 个 SRFA 样品)、土壤 DOM(20 个样品)、废水 DOM(6 个样品)以及其它类型 DOM(包含 1 个木质素和 1 个树叶浸提液样品).

1.2 检测设备

本研究利用 FT-ICR MS 仪(型号为 Solarix 2XR 7T, 德国 Bruker Daltonik 公司), 采用手动直接进样的方式, 电荷模式为 ESI(-), 进样流速为 120 μL·h⁻¹, 离子源喷针电压设置为 5.0 kV, 样品采用负离子宽谱扫描模式, 质量(mass-to-charge ratio, m/z)扫描范围为 100—1100, 离子累积时间为 0.025 s, 扫描次数为 1000 次, 数据容量为 4 megaword. 在样品谱图采集前, 先用 100 mg·L⁻¹ 三氟乙酸钠标准液(50% 甲醇水溶液)对 m/z 值进行外标校准. 所有 FT-ICR MS 谱图以淡水天然有机质的 CHO 同系化合物为内标物质, 采用 Data Analysis 软件(Bruker Daltonik, version 5.0)对谱图进行校准^[13].

1.3 分子式匹配方法

本研究采用 TRFu 代码^[29] 进行分子式匹配. TRFu 代码运行流程图见图 1, 首先将质荷比(m/z)、强度(Intensity)和信噪比(Signal-to-noise ratio, S/N)输入到 MATLAB 中(必须确保 m/z 、Intensity 和 S/N 的顺序一致). 计算条件如下: 确认 S/N 大于等于 6 的峰, 测量质量误差小于等于 0.75×10^{-6} ; 仅考虑 C、H、O、N、P 和 S 等 6 种元素, 元素个数及比例设置如下: C(4—50)、N(0—5)、S(0—3)、P(0—1); 氢碳比(H/C)范围: $0.3 \leq H/C \leq 2.25$, 氧碳比(O/C)范围: $0 \leq O/C \leq 1.15$. 双键当量(Double-bond equivalence, DBE)大于等于 0, 电荷值为 -1 和 -2. 进一步采用如下规则对得到的分子式进行过滤: 设定单电荷离子的双键当量减氧^[25](DBE-O)范围为 $-10 \leq \text{DBE-O} \leq 10$; 双电荷离子的 DBE-O 范围进一步扩大为 $-12 \leq \text{DBE-O} \leq 12$ 且同时满足 $N+S+P$ 原子个数总数最小和 N 规则.

由于 DOM 分子中非氧杂原子含量较低^[30], 同一 m/z 值下可能存在的多个分子式, 通过最佳分子式筛选规则进行筛选. 最佳分子式筛选规则主要分为三个部分: 初步筛选、前体筛选和 ¹³C 同位素筛选. 初步筛选规则如下: (1) $N+S+P$ 原子个数之和最小; (2) $S+P$ 原子个数之和最小; (3)最小质量误差. 在初步筛选规则下得到的 $[M-2H]^{-2}$ 型分子式还需要进行前体筛选, 规则如下: (a)当 $[M_1-2H]^{-2}$ 型分子存在 $[M_1-H]^{-1}$ 型分子时, 则保留该 m/z 值(m/z 值为 $[M_1-2H]^{-2}$ 型分子的质荷比, 同下)所匹配分子式; (b)当 $[M_1-2H]^{-2}$ 型分子式不存在 $[M_1-H]^{-1}$ 型分子且含有非氧杂原子时, 则删除该 m/z 值所匹配分子式; (c)当 $[M_1-2H]^{-2}$ 型分子式不存在 $[M_1-H]^{-1}$ 型分子和非氧杂原子时, 则判断该 m/z 值是否匹配了 $[M_2-H]^{-1}$ 型分子式(M_1 、 M_2 为不同的分子式). 判断条件如下: ①如果未匹配到 $[M_2-H]^{-1}$ 型分子式, 则删除该 m/z 值所匹配分子式; ②如果匹配到单个 $[M_2-H]^{-1}$ 型分子时, 则将该 m/z 值匹配为 $[M_2-H]^{-1}$ 型分子式; ③如果匹配到多个 $[M_2-H]^{-1}$ 型分子时, 则将所有匹配到的 $[M_2-H]^{-1}$ 型分子继续按照以下顺序进行筛选: (i) $N+S+P$ 原子个数之和最小; (ii) $S+P$ 原子个数之和最小; (iii)最小质量误差. 最后将上述未匹配的峰、确定的峰以及选定的最佳峰组合后进行 ¹³C 同位素筛选. 规则如下: (1)存在单同位素峰(¹²C_{*n*} 峰); (2)存在 ¹³C 同位素峰(¹³C₁¹²C_{*n-1*} 峰), 且 ¹³C₁¹²C_{*n-1*} 峰与 ¹²C_{*n*} 峰之间的 m/z 值相差 1.0031 Da 至 1.0035 Da; (3)满足 ¹²C_{*n*} 峰的相对丰度(Relative abundance, RA)大于等于 5% 或 ¹³C₁¹²C_{*n-1*} 峰的 RA 值小于 ¹²C_{*n*} 峰且 ¹²C_{*n*} 峰的 RA 值小于 5%.

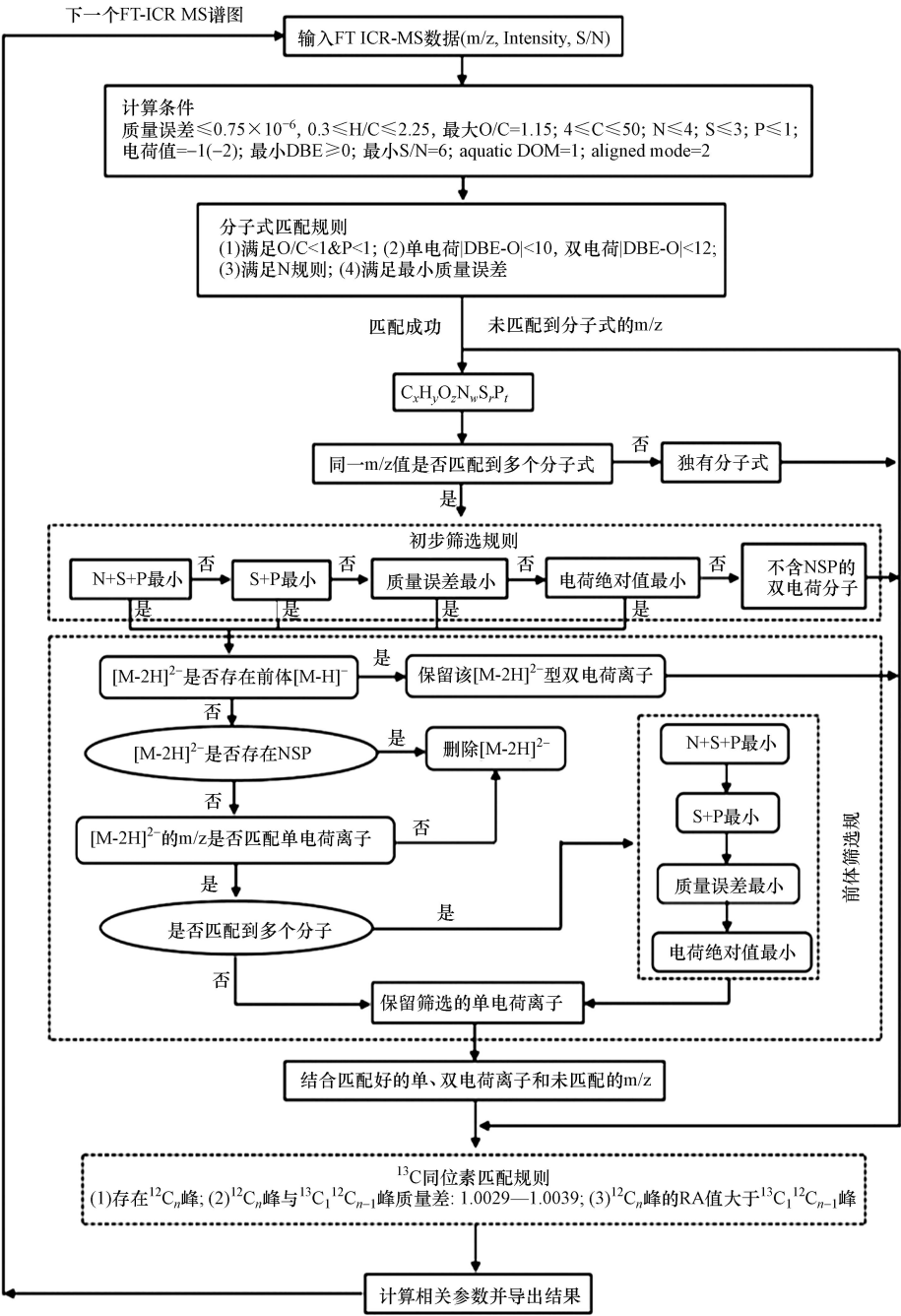


图 1 TRFu 代码运行流程图
Fig.1 Flowchart of the TRFu algorithm

1.4 分子参数计算和统计分析

本研究根据每个离子峰所属的化学式,采用以下方程分别计算 DBE、修正的芳香性指数(Modified aromaticity index, AI_{mod})和碳名义氧化态(Nominal oxidation state of carbon, NOSC). 其中 DBE 是指分子中双键和脂环总数; AI_{mod} 被用来确定 DOM 中是否存在芳香结构(AI_{mod}>0.5)或缩合芳香结构(AI_{mod}≥0.67)^[31-32]; NOSC 指数与有机化合物的生物地球化学反应性和生物利用度有关^[33]. 功能多样性指数 D_F 是以 FT-ICR MS 数据计算为距离函数(Rao 的二次熵)^[34], D_F 可理解为两个分子之间相对于所选 DOM 分子性质的预期差异值. 分子式 C_xH_yO_zN_wS_rP_t 的相关参数(包括 DBE、AI_{mod}、NOSC 和 D_F)计算公式分别如下式(1)—式(4)所示:

$$DBE = 1 + 0.5(2x - y + w + t)$$
 (1)

$$AI_{mod} = \frac{1 + x - 0.5z - r - 0.5y}{x - 0.5z - r - w - t}$$
 (2)

$$\text{NOSC} = 4 - \frac{4x + y - 3w - 2z - 2r + 5t}{x} \quad (3)$$

$$D_F(c) = \sum_{i=1}^{n-1} \sum_{j=i+1}^n p_i \cdot p_j \cdot |c_i - c_j| \quad (4)$$

公式(1)、(2)和(3)中的 x 、 y 、 z 、 w 、 r 和 t 分别指每个分子式中的碳、氢、氧、氮、硫和磷原子的化学计量数. 公式(4)中 n 为样品匹配的分子式总数, p_i 和 p_j 分别是分子式 i 和 j 的 RA 值, c_i 和 c_j 分别是分子式 i 和 j 的特定参数(如 O/C、H/C、DBE、 AI_{mod} 和 NOSC 等).

单个 ^{13}C 同位素的 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰的峰强度偏差的计算公式如下(式(5)和式(6)):

$$^{13}\text{C峰峰强偏差} = \frac{^{13}\text{C峰峰强测量值} - ^{13}\text{C峰峰强理论值}}{^{13}\text{C峰峰强理论值}} \quad (5)$$

$$^{13}\text{C峰峰强理论值} = ^{12}\text{C峰峰强测量值} \times 0.010816 \times n_C \quad (6)$$

式(6)中, 0.010816 为 ^{13}C 与 ^{12}C 同位素的自然丰度比; n_C 是 $^{12}\text{C}_n$ 峰分子式中的碳原子数. 当样本数据通过 Kolmogorov-Smirnov 正态性检验时, 使用单因素方差分析 (One-way analysis of variance, ANOVA) 在 0.05 的显著性水平上检验比较组之间的显著性差异; 否则采用非参数 Kruskal-Wallis 单因素方差分析.

2 结果与讨论(Results and discussion)

2.1 ^{13}C 同位素验证双电荷 DOM 分子

实验中共找到 410 个双电荷 $^{12}\text{C}_n$ 峰(对应 410 个双电荷 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰, 统称为 A 类峰)和 116 个单、双电荷同 m/z 值的 $^{12}\text{C}_n$ 峰(对应 116 个单电荷 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰和 116 个双单电荷 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰, 统称为 B 类峰). 为验证双电荷 DOM 分子($[\text{M}-2\text{H}]^2$ -型分子, $[\text{M}-\text{H}]^2$ -型分子为单电荷 DOM 分子)的准确性, 本实验将从质量误差和峰强度偏差两个角度进行 ^{13}C 同位素验证.

从质量误差的角度来看, A 类 $^{12}\text{C}_n$ 峰与 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰的 m/z 差值为 $(0.5016704 \pm 0.00007) \text{ Da}$, 这与其理论值 0.5016775 高度接近, 说明双电荷 DOM 的 $^{12}\text{C}_n$ 峰和 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰分子式均匹配正确. 假设 A 类 $^{12}\text{C}_n$ 峰和 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰所匹配的分子式是假阳性的, 结合 A 类峰的 m/z 值在质谱图中的位置来看, 所有 $^{12}\text{C}_n$ 峰($^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰)匹配的分子式将包含偶(奇)数个 N 或奇(偶)数个 N 加单个 ^{13}C . 进一步, 在初始分子式匹配条件下(如最小质量误差 0.75×10^{-6} , $0.3 \leq \text{H/C} \leq 2.25$, $0 \leq \text{O/C} \leq 1.15$, $-12 \leq \text{DBE-O} \leq 12$), A 类 $^{12}\text{C}_n$ 峰和 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰的 m/z 值分别匹配到 2144 和 1955 个双电荷分子式. 其中 $^{12}\text{C}_n$ 峰匹配到的双电荷分子式类型包括 X、 $^{13}\text{C}_1\text{N}_1$ 、 N_2 、 $^{13}\text{C}_1\text{N}_3$ 、 N_4 和 $^{13}\text{C}_1\text{N}_5$ 型分子式(X 表示不含 ^{13}C 和 N 元素的 CHO 类分子式, 其他元素未给出); $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰匹配到的双电荷分子式类型包括 $^{13}\text{C}_1$ 、 N_1 、 $^{13}\text{C}_1\text{N}_2$ 、 N_3 、 $^{13}\text{C}_1\text{N}_4$ 和 N_5 型分子式. $^{12}\text{C}_n$ 峰和 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰各自匹配的 5 类分子式的质量误差如图 2a 和图 2b 所示, 结果显示: $^{12}\text{C}_n$ 峰所匹配的分子式中 X 型与其它 4 类分子式均存在显著相关性($P < 0.05$); $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰所匹配的分子式中 $^{13}\text{C}_1$ 型与 N_1 、 $^{13}\text{C}_1\text{N}_2$ 、 N_3 和 $^{13}\text{C}_1\text{N}_4$ 型分子式存在显著性差异($P < 0.05$), 与 N_5 型无显著性差异. 以上结果表明, $^{12}\text{C}_n$ 峰和 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰所匹配的分子式中 CHO 类的质量误差最低, 分别为 $(0.087 \pm 0.084) \times 10^{-6}$ 和 $(0.142 \pm 0.114) \times 10^{-6}$. 进一步从非氧杂原子的角度来看, 由于 DOM 中非氧杂原子含量较低^[30], $^{12}\text{C}_n$ 峰和 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰的分子式匹配为 CHO 类的准确性更高.

从峰强度偏差来看, A 类峰中 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰的测量 RA 值与理论 RA 值(由公式(6)得出)的分布如图 3a 所示. $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰的理论 RA 值与测量 RA 值的比值为 0.94 ± 0.17 , 二者的相关系数(R^2)为 0.744, 这表明由 A 类 $^{12}\text{C}_n$ 峰预测的理论 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰与实际 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰基本符合. 从 $\pm 30\%$ 的偏差阈值来看^[35], 存在 347 个(约占 A 类双电荷离子峰总数的 84.63%)A 类 $^{12}\text{C}_n$ 峰满足 ^{13}C 同位素偏差验证. 而在 $\pm 30\%$ 偏差阈值之外, 通过 ^{13}C 同位素偏差验证时产生了较大的偏差(如图 3b 中 $6 \leq S/N \leq 20$ 的区域). 如 m/z 值为 228.064009 的 $^{12}\text{C}_n$ 峰 $[\text{C}_{20}\text{H}_{24}\text{O}_{12}]^2$, 其 RA 值和 S/N 值分别为 1.15 和 7.2; 以及 m/z 值为 228.565737 的 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰 $^{13}\text{C}_1[\text{C}_{19}\text{H}_{24}\text{O}_{12}]^2$, 其 RA 值和 S/N 值分别为 1.04 和 6.3, 通过公式(5)计算将会产生 3.18 的偏差值. 产生这个结果原因并不是分子式匹配错误, 因为所有分子式赋值都满足严格的匹配条

件,这可能是由于 $[C_{20}H_{24}O_{12}]^{2-}$ 峰的 S/N 值过低(仅为 7.2)从而差生较大的偏差. 因此,部分 A 类峰的 ^{13}C 同位素偏差较大可归结为 S/N 值过低($6 \leq S/N \leq 20$)从而导致检测到的 RA 值偏低.

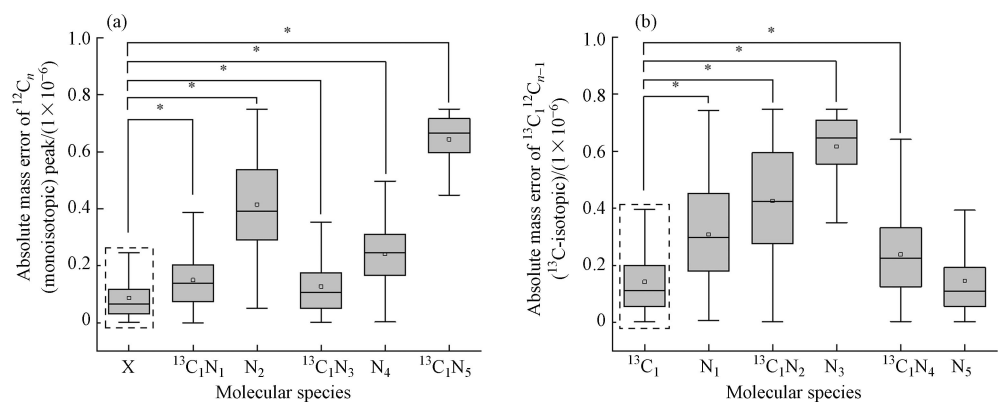


图 2 A 类双电荷 $^{12}C_n$ 峰(a)和 $^{13}C_1^{12}C_{n-1}$ 峰(b)的 m/z 值所赋值的分子式
*表示 $P < 0.05$ 水平的显著性差异

Fig.2 Molecular formula assigned by the m/z value of the class A doubly charged $^{12}C_n$ peak(a) and $^{13}C_1^{12}C_{n-1}$ peak(b)
the significant differences at $P < 0.05$ are indicated by the star symbol (*)

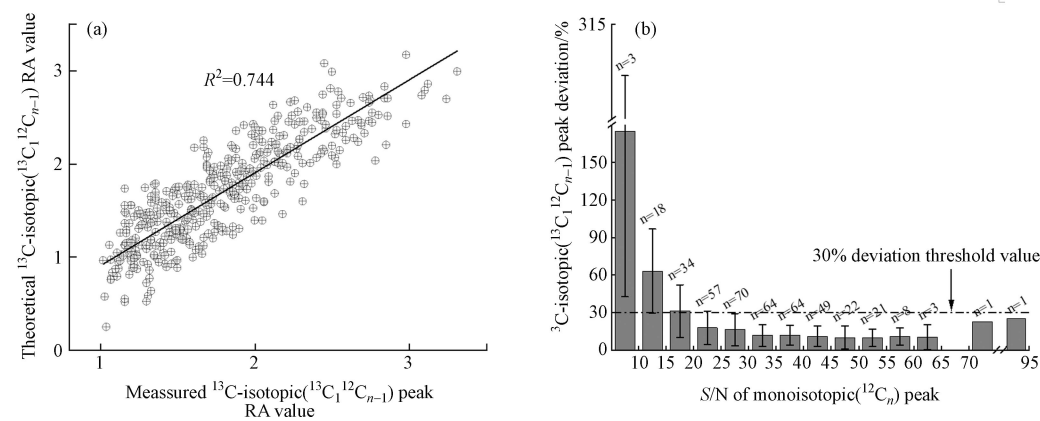


图 3 A 类双电荷 $^{13}C_1^{12}C_{n-1}$ 峰的理论 RA 值与测量 RA 值之间的散点图(a), A 类 $^{12}C_n$ 峰的 S/N 值与 $^{13}C_1^{12}C_{n-1}$ 峰的峰强度偏差的直方图(b)
 n 表示该 S/N 值范围内的 $^{12}C_n$ 峰个数,数值和误差棒代表均值和标准偏差($\pm SD$)

Fig.3 Scatter plot between theoretical RA value and measured RA value of doubly charged $^{13}C_1^{12}C_{n-1}$ peak in class A (a),
histogram of peak intensity deviation between S/N value of $^{12}C_n$ peak and $^{13}C_1^{12}C_{n-1}$ peak in class A (b)
 n is the sample number of each method. The values and error bars represent average and $\pm SD$, respectively

质谱图上相同 m/z 值处的峰信号强度是由多个离子的信号叠加而成,这些离子可能来自多种不同的电荷态. 因此,对于 B 类 $^{12}C_n$ 峰的峰强度偏差验证,这里采用单、双电荷 $^{13}C_1^{12}C_{n-1}$ 峰的 RA 值反向验证 $^{12}C_n$ 峰. 如图 4a—图 4c 所示,在 m/z 值为 377.051388 的位置存在单电荷 $[C_{17}H_{13}O_{10}]^-$ 和双电荷离子 $[C_{34}H_{26}O_{20}]^{2-}$,同时分别在约为 $M + 1$ 和 $M + 0.5$ 的 m/z 差值处检测到 ^{13}C 同位素峰 $[^{13}C_1C_{16}H_{13}O_{10}]^-$ 和 $[^{13}C_1C_{33}H_{26}O_{20}]^{2-}$. 单、双电荷 $^{12}C_n$ 峰的理论 RA 值可由 $^{13}C_1^{12}C_{n-1}$ 峰的 RA 值经过公式(6)分别反向推求出,结果分别为 36.76% 和 6.91%. 因此, m/z 值为 377.051388 的理论 RA 值为 43.67%, 峰强度偏差为 -3.67% . 若验证峰强度偏差时只考虑单电荷(或双电荷) $^{13}C_1^{12}C_{n-1}$ 的 RA 值,实验将会产生更大的偏差. 如图 4d 所示,采用单、双电荷 $^{13}C_1^{12}C_{n-1}$ 峰相结合的验证方式与只采用单电荷(或双电荷) $^{13}C_1^{12}C_{n-1}$ 峰的验证方式之间存在显著性差异($P < 0.05$),采用单、双电荷 $^{13}C_1^{12}C_{n-1}$ 峰相结合的所得到的分子式具有更低的强度偏差值($5.37\% \pm 5.254\%$).

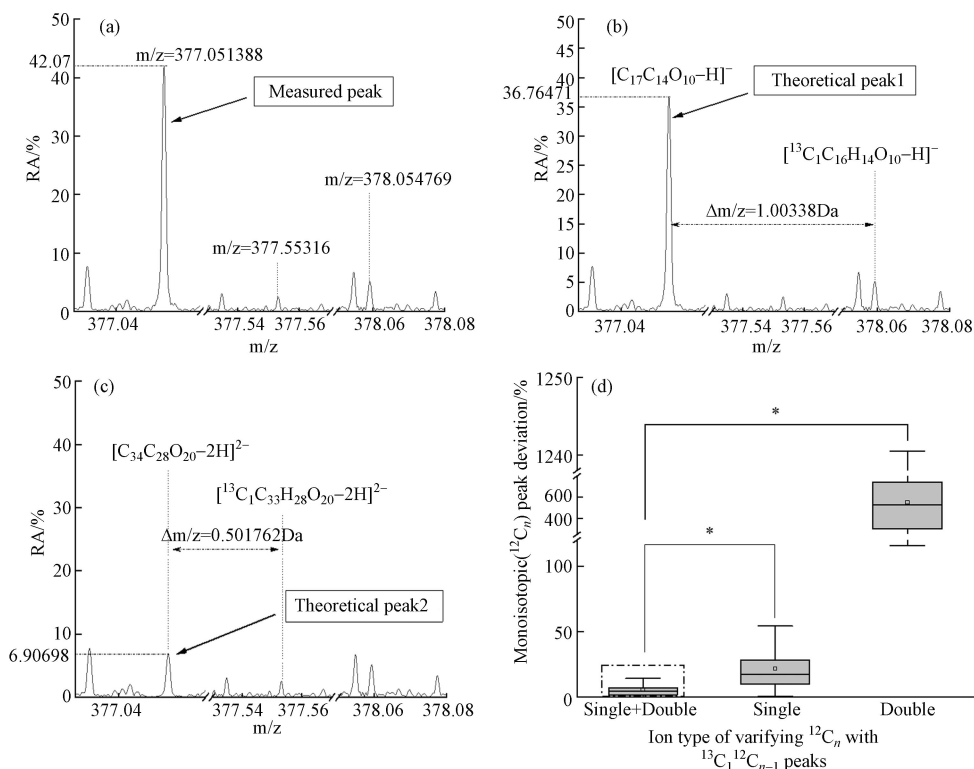


图 4 待验证的单、双电荷 $^{12}\text{C}_n$ 峰(a), 单电荷 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰预测 $^{12}\text{C}_n$ 峰(b), 双电荷 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰预测 $^{12}\text{C}_n$ 峰(c)和 B 类峰的峰强度偏差验证(单电荷+双电荷 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰、单电荷 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰和双电荷 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰验证)(d)

*表示 $P < 0.05$ 水平的显著性差异

Fig.4 The singly and doubly charged $^{12}\text{C}_n$ peak to be verified(a), the singly charged $^{13}\text{C}_1^{12}\text{C}_{n-1}$ peak is predicted to be $^{12}\text{C}_n$ peak(b), the doubly charged $^{13}\text{C}_1^{12}\text{C}_{n-1}$ peak prediction $^{12}\text{C}_n$ peak(c) and class B peak intensity deviation verification(singly charged + doubly charged $^{13}\text{C}_1^{12}\text{C}_{n-1}$ peak, singly charged $^{13}\text{C}_1^{12}\text{C}_{n-1}$ peak and doubly charged $^{13}\text{C}_1^{12}\text{C}_{n-1}$ peak verification)(d) The significant differences at $P < 0.05$ are indicated by the star symbol (*)

以上结果表明, A 类峰中大部分双电荷离子(约占 84.63%)通过了 ^{13}C 同位素偏差验证, 少数双电荷离子存在较大偏差可归结为部分 $^{12}\text{C}_n$ 峰的 S/N 值过低(大约小于 20); 由 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰反向验证 $^{12}\text{C}_n$ 峰的峰强度偏差结果表明: B 类峰的信号强度是由相同 m/z 值的单、双电荷 $^{12}\text{C}_n$ 峰信号叠加而成, 采用单、双电荷 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰相结合的验证方式准确度更高, 这也进一步表明 B 类双电荷离子有效的通过了 ^{13}C 同位素偏差验证。

2.2 基于单电荷离子前体方法与 ^{13}C 同位素方法的差异

对整个质谱数据进行分子式匹配和筛选后, DOM 的 FT-ICR MS 谱图中双电荷离子峰总共 2097 个, 其中以前体离子方法筛选得到的双电荷离子为 1762 个(约占双电荷离子总数的 84.02%)。以图 5a 中名义质量 325 和 651 为例, 1 号到 9 号峰均包含双电荷离子峰, 其中 2 号、6 号和 9 号峰为单、双电荷共存的离子峰($[\text{C}_x\text{H}_y\text{O}_z]^-$ 和 $[\text{C}_{2x}\text{H}_{2y}\text{O}_{2z}]^{2-}$ 具有相同的 m/z), 同时存在强度较高的 1' 号到 9' 号单电荷离子峰前体。从 DOM 分子中羧基的数量和间距来看^[19], 1 号到 9 号双电荷离子峰的出现可能受到羧基的影响。传统 ^{13}C 同位素方法是以分子式之间的质量差($\Delta m/z=^{13}\text{C}_1^{12}\text{C}_{n-1}-^{12}\text{C}_n$)来判断单、双电荷离子, 如图 5b(峰 1 和峰 1' 为单电荷离子峰; 峰 2 和峰 2' 为双电荷离子峰)所示, 该方法得到的双电荷离子数目为 1216 个(约占双电荷离子总数的 57.99%), 这其中还包含 27 个含 N 的双电荷离子。值得一提的是, 传统 ^{13}C 同位素方法确认的双电荷 $^{12}\text{C}_n$ 峰(m/z 值对应名义质量)和 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰(m/z 值对应名义质量+0.5 Da)如图 5b 中的峰 2 和峰 2' 所示, 但该方法寻找的双电荷离子并不全面, 这是因为传统 ^{13}C 同位素方法并没有考虑含 N 的双电荷离子。如图 5b 中的峰 3 和峰 3' 所示, 含奇数个 N 原子的双电荷 $^{12}\text{C}_n$ 峰(m/z 值对应名义质量+0.5 Da)和 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰(m/z 值对应名义质量+1 Da)与传统 ^{13}C 同位素方法确认的双电荷 $^{12}\text{C}_n$ 峰和 $^{13}\text{C}_1^{12}\text{C}_{n-1}$ 峰的位置不同, 这也是传统 ^{13}C 同位素方法所未考虑的。其中峰 3 的 m/z 为 334.530568, 离子化学式为 $[\text{C}_{29}\text{H}_{19}\text{O}_{18}\text{N}_1]^{2-}$; 峰 3' 的 m/z 为 335.032374, 离子化学式为

[¹³C₁C₂₈H₁₉O₁₈N]²⁻(峰 3'的 *S/N* 值仅为 5.5, 这里仅为说明传统¹³C 同位素方法的不足). 此外, 当强度较低或者不存在¹³C 同位素峰时, 传统¹³C 同位素方法则不能判断¹²C_{*n*} 峰(峰 a 到峰 d)以及对应的¹³C₁¹²C_{*n-1*} 峰(峰 a'到峰 d'). 同时存在 6 个双电荷¹²C_{*n*} 峰(峰 f 到峰 k)没有匹配到¹³C₁¹²C_{*n-1*} 峰, 而峰 f 到峰 k 均满足严格的赋值条件^[13], 这说明以传统¹³C 同位素方法来寻找双电荷离子并不可靠. 对比两种方法可知, 传统¹³C 同位素方法并没有考虑单电荷离子前体是否存在^[16, 19, 20, 36], 以名义 *m/z* 323 和 *m/z* 647 为例(图 5d). 在 *m/z* 为 323 和 647 的质量窗口上验证了双电荷¹²C_{*n*} 峰 [C₃₁H₁₈O₁₆]²⁻和双电荷¹³C₁¹²C_{*n-1*} 峰 [¹³C₁C₃₀H₁₈O₁₆]²⁻(分别对应图 5d 中峰 a 和峰 a'), 其中峰 a 存在单电荷离子峰前体 [C₃₁H₁₉O₁₆], 但峰 a'却并未找到单电荷离子峰前体. 在该样品中总共发现 335 个双电荷离子(约占双电荷离子总数的 15.98%)不存在单电荷离子前体, 出现这种结果可能是多个羧基的存在使 DOM 更易电离形成双电荷离子^[19]. 以上结果表明, 相比传统¹³C 同位素方法, 单电荷离子前体方法能够有效识别 DOM 的 FT-ICR MS 谱图中双电荷离子峰, 可以进一步提高(提高 44.90%)双电荷离子的个数.

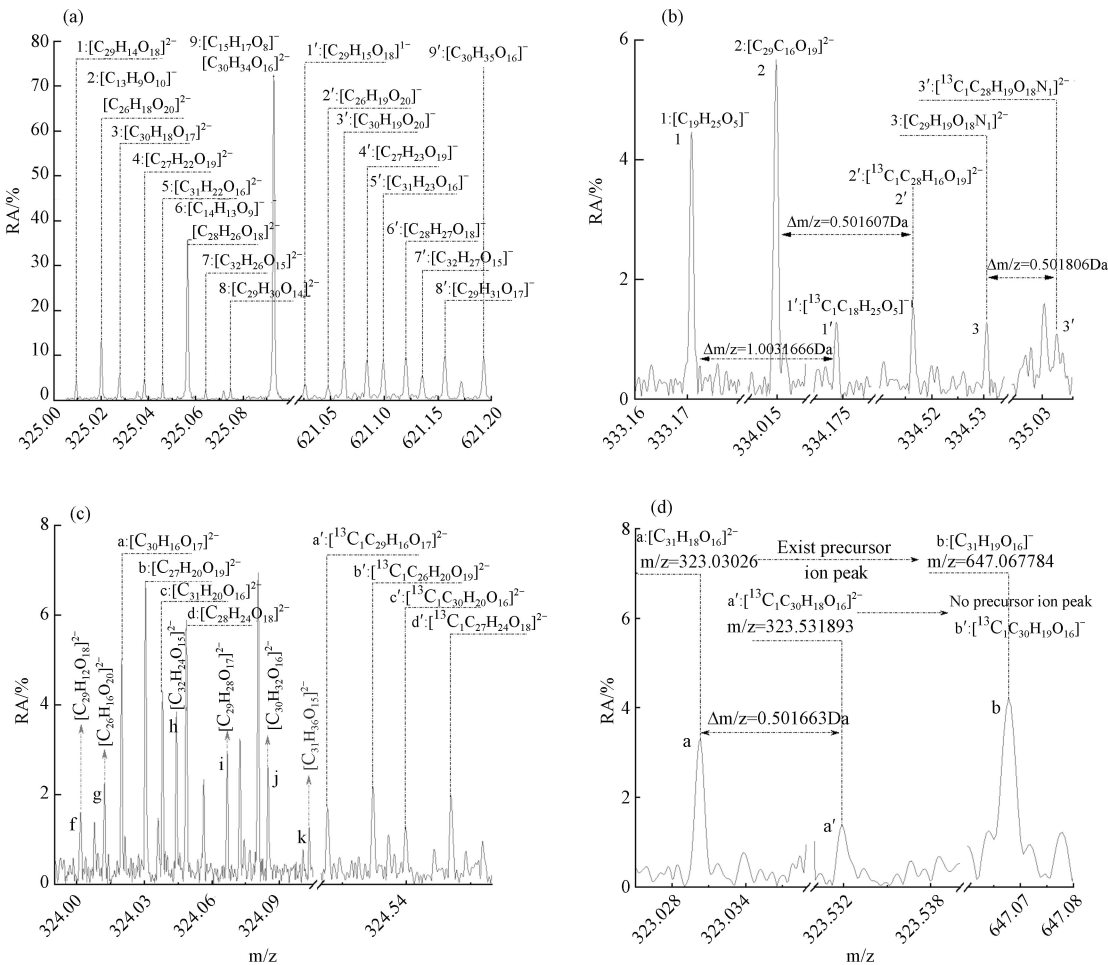


图 5 在名义 *m/z* 325 和 651 质量窗口处利用单电荷离子前体确定双电荷离子(a), 传统¹³C 同位素确定单(双)电荷离子以及含 N 双电荷离子示例(b), 在名义 *m/z* 324 质量窗口利用¹³C 同位素确定的双电荷离子(c)和名义 *m/z* 320 和 641 质量窗口处 [¹³C₁C₂₄H₂₀O₂₀]²⁻无单电荷离子前体(d)

Fig.5 Determination of doubly charged ions by singly charged precursor ions at the nominal *m/z* 325 and 651 mass windows(a), examples of determining singly (doubly) charged ions with ¹³C isotopes and containing N doubly charged ions(b), determination of doubly charged ions by ¹³C isotopes at the nominal *m/z* 324 mass window(c) and [¹³C₁C₂₄H₂₀O₂₀]²⁻ without precursor singly charged ions at the nominal *m/z* 320 and 641 mass windows(d)

2.3 双电荷 DOM 分子的特征及影响

2.3.1 单、双电荷 DOM 分子的参数差异

从单、双电荷 DOM 分子的 van Krevelen^[14] 可知(图 6a), 大部分双电荷 DOM 分子(约占比双电荷

离子总数的 99.80%)具有较高的 O/C 值(0.39—0.84)和较低的 H/C 值(0.45—1.33),这意味着双电荷 DOM 分子具有较高的不饱和度,说明双电荷 DOM 分子经历了更多的氧化过程从而具有较高的氧化态(NOSC)^[37].从单、双电荷 DOM 分子的 5 个参数箱线图(图 6b—图 6f)可以看到,5 个参数值在单、双电荷 DOM 分子之间存在显著性差异($P<0.05$).其中 O/C、DBE、 AI_{mod} 和 NOSC 值在双电荷 DOM 分子中均值更高,而 H/C 值更低.以上结果表明双电荷 DOM 分子具有更高的芳香性、不饱和度和氧化态,这可能与双电荷 DOM 分子中存在较多的羧基有关^[38].

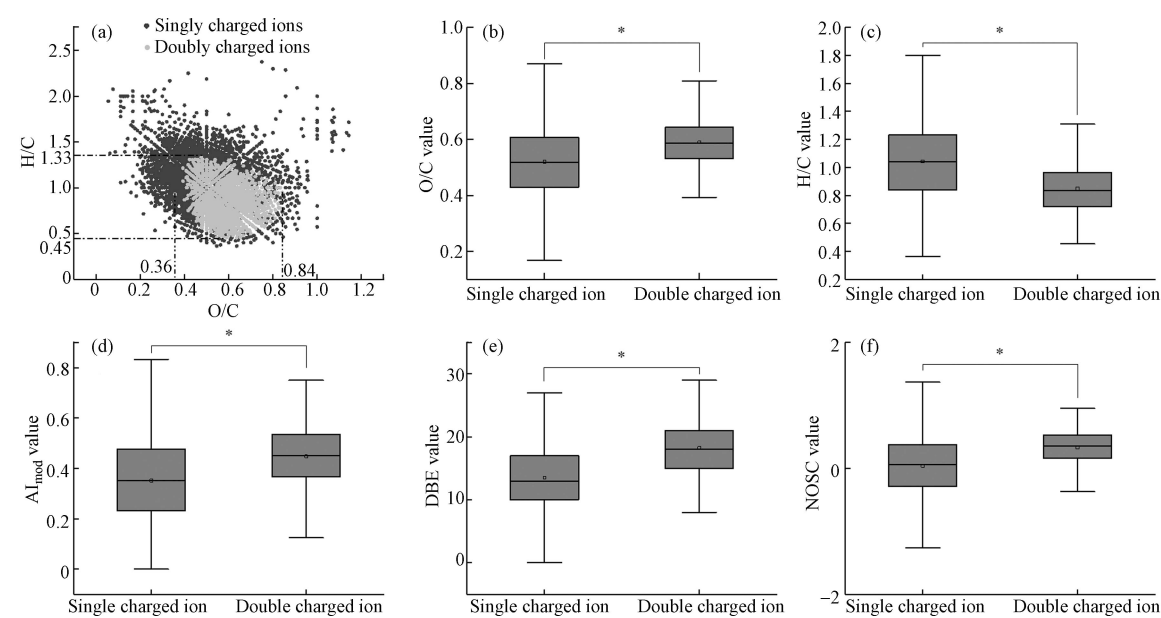
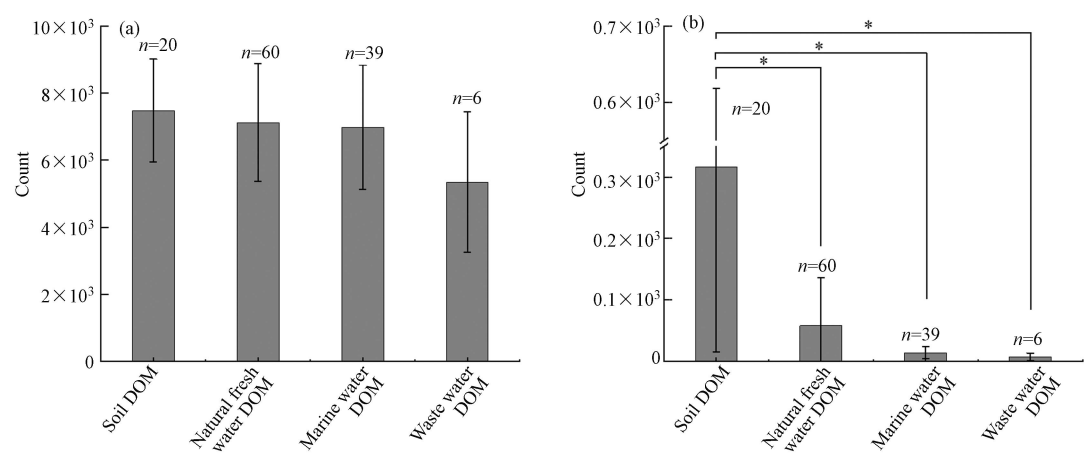


图 6 单、双电荷 DOM 分子的 van Krevelen 图(a)和单、双电荷 DOM 分子的 5 个参数箱线图(b—f)
*表示 $P < 0.05$ 水平的显著性差异

Fig.6 The van Krevelen diagrams of singly and doubly charged DOM molecules(a)and significance analysis of molecular parameters of singly and doubly charged DOM(b—f)
The significant differences at $P < 0.05$ are indicated by the star symbol(*)

2.3.2 不同类型双电荷 DOM 分子差异

对 4 类样品(土壤、自然淡水、海水和废水)进行显著性分析(ANOVA, $P<0.05$)可知(图 7a 和图 7b),单电荷 DOM 分子的数量在不同类型样品之间没有显著性差异.相反,双电荷 DOM 分子的数量在不同类型样品之间存在显著性差异,土壤双电荷 DOM 分子与海水、自然淡水和废水双电荷 DOM 分子之间存在显著性差异($P<0.05$).以上结果表明双电荷 DOM 分子的数量在不同类型样品中存在显著性差异,土壤 DOM 中的双电荷离子更丰富.



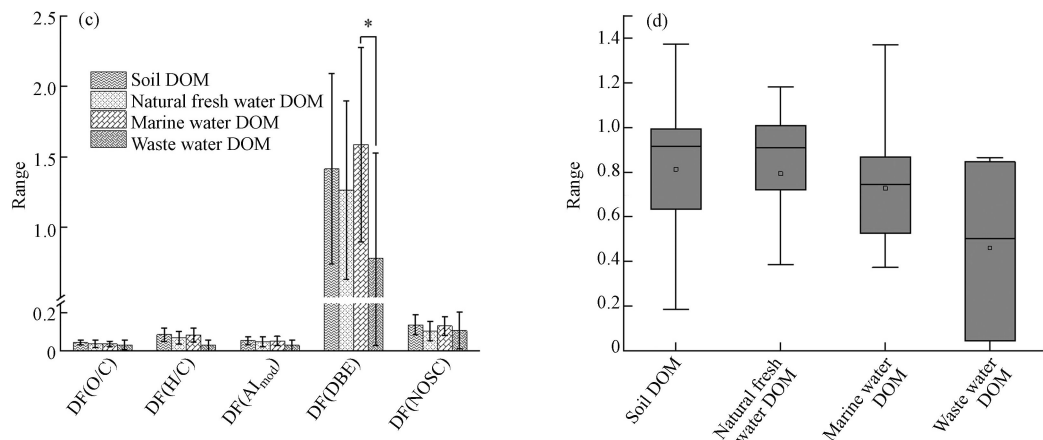


图 7 不同类型样品中单(a)、双(b)电荷 DOM 分子数量的显著性, 不同类型样品间的功能多样性指数(D_F)(c)和综合功能多样性指数分布箱线图(d).
数值和误差条分别表示平均值和标准偏差($\pm$ SD). *表示 $P < 0.05$ 水平的显著性差异.

Fig.7 Significance of the number of singly(a)、doubly(b)charged DOM molecules in different types of samples, significance of different functional diversity indices(D_F)(c)and distribution box diagram of comprehensive function diversity index(d)
The values and error bars represent average and $\pm$ SD, respectively. The significant differences at $P < 0.05$ are indicated by the star symbol(*)

根据分子式中 O/C 和 H/C 值, 可将 DOM 分子的 van Krevelen 图划分为 7 类^[13], 4 类样品的双电荷 DOM 分子的 van Krevelen 分类结果如表 1 所示. 可以看到, 所有样品均以丰富的木质素为主, 其中在土壤、自然淡水、海水和废水的双电荷 DOM 中分别占比 76.71%、79.00%、88.15% 和 85.36%. 此外丹宁在四类样品中也有少量占比, 分别为 4.91%、11.04% 和 3.33% 和 7.32%. 这表明与大多数 DOM 的来源相似, 不同类型的双电荷 DOM 分子可能来自植物源^[39].

表 1 不同类型样品中双电荷 DOM 分子的 van Krevelen 分类结果

Table 1 Results of van Krevelen classification of double-charged DOM molecules in different types of samples

样品类型 Sample types	木质素 Lignin	丹宁 Tannins	稠环芳烃 Condensed aromatic	饱和化合物 Saturated compounds	氨基糖 Aminosugars	不饱和和碳氢化合物 Unsaturated hydrocarbons	碳氢化合物 Carbo-hydrates	占比 ¹⁾ Percentage	总计 ²⁾ Total
土壤	76.71%	4.91%	16.06%	2.00%	0.25%	0.01%	0.05%	4.06%	155947
自然淡水	79.00%	11.04%	9.75%	0.14%	0.06%	0	0	0.81%	430608
海洋水	88.15%	3.33%	2.30%	2.40%	0.74%	0	0	0.18%	304919
废水	85.36%	7.32%	0	7.32%	0	0	0	0.13%	32080

注: 1)代表双电荷离子与离子总数(单电荷离子数目和双电荷离子数目总和)的占比; 2)代表离子总数.
Note: 1) represents the ratio of doubly charged ions to the total number of ions (the summed number of singly and doubly charged ions); 2) represents the total number of ions.

进一步结合四类样品不同性质参数的 D_F 值来看 (图 7c), D_F (O/C)、 D_F (H/C)、 D_F (AI_{mod}) 和 D_F (NOSC) 值在四类样品中均无显著性差异. 土壤、自然淡水和海水 DOM 中双电荷的 D_F (DBE) 值与废水 DOM 存在显著性差异 ($P < 0.05$), 废水双电荷 DOM 的 DBE 值更低, 这可能与废水属于难降解的稳定化合物有关^[40]. 从范围来看, 四类样品的 D_F (O/C) 值相差不大, 这与图 6a 所述的双电荷 DOM 分子的高 O/C 值一致. 海水双电荷 DOM 有着更高的 D_F (DBE) 和 D_F (NOSC) 值, 这说明海水 DOM 在生物地球化学反应中有更高的活性^[37]. 从 D_F (H/C) 和 D_F (AI_{mod}) 可知, 相比废水双电荷 DOM 分子, 土壤、自然淡水和海水双电荷 DOM 的 H/C 和 AI_{mod} 值均更高, 表明它们的生物可利用度更高^[40] 且有着更高的芳香程度. 综合分子式中所有性质参数的期望值来看 (图 7d), 土壤双电荷 DOM 分子的综合功能多样性指数均值更高, 其次是自然淡水、海水和废水双电荷 DOM 分子. 自然淡水和海水双电荷 DOM 分子的综合功能多样性指数分布更集中, 而土壤和废水双电荷 DOM 分子则更分散, 这可能与不同地区的土壤 DOM 分子性质相差较大以及不同来源的废水有关. 以上结果表明不同类型样品的双电荷 DOM 分子的多样性指数受到样品类型的影响, 土壤双电荷 DOM 分子具有更高的芳香程度和不饱和度.

2.3.3 双电荷离子对数据分析的影响

根据“1.3”节前体筛选规则可知: [M-H]⁻型分子和 [M-2H]²⁻型分子最终匹配的分子式是相同的, 因此对于分子式匹配而言, 存在前体单电荷 DOM 分子的双电荷 DOM 分子的 UHR-MS 数据可以不考虑, 并且不存在前体单电荷 DOM 分子的双电荷 DOM 分子的数量较少(约占单、双电荷离子总数的 15.98%)。此外, 由 5 类样品中单、双电荷 DOM 分子的元素组成上可知, 如表 2 所示。结果显示, 单电荷 DOM 分子包含的 8 类化合物中以丰富的 CHO 类化合物为主(51.98%), 其次为 CHON 类化合物(18.72%)和 CHOS 类化合物(12.25%)。而双电荷 DOM 分子仅包含 4 类化合物, 其中 CHO 类化合物最为丰富(1.07%), 其次为 CHON 类化合物(0.06%)、CHOS 类化合物(0.02%)和 CHOP 类化合物(0.01%)。以上结果表明: 相比双电荷 DOM 分子, 单电荷 DOM 分子在元素组成和化合物类别上更复杂, 在数量上也更多。在 FT-ICR MS 谱图上成千上万个离子峰中, 单电荷离子峰是主要的离子峰, 而在双电荷离子中不存在单电荷离子前体的少之又少, 其对谱图中成千上万个离子峰而言微不足道。因此, 在分子式匹配过程中, 往往可以默认为只考虑单电荷离子而忽略双电荷离子, 双电荷离子对质谱数据分析的影响十分有限。

表 2 不同类型样品中所有单、双电荷离子的元素组成及占比

离子类型 Ion types	化合物类型 Type of compound								总计* Total
	CHO	CHON	CHOS	CHOP	CHONS	CHONP	CHOSP	CHONSP	
单电荷离子 (single charged ions)	51.98%	18.72%	12.25%	3.03%	6.56%	2.83%	1.57%	1.88%	890383
双电荷离子 (double charged ions)	1.07%	0.06%	0.02%	0.01%	0	0	0	0	10497

注: *代表离子总数(单电荷离子数目和双电荷离子数目总和)。
Note: * represents the number of ions (the summed number of singly and doubly charged ions).

3 结论(Conclusion)

本研究首次提出了单电荷离子前体确认双电荷离子的方法, 该方法相比传统¹³C 同位素方法能够有效提高 DOM 的 FT-ICR MS 谱图中双电荷离子峰的识别, 双电荷离子比例提高 44.90%。同时该方法能确认不存在¹³C₁¹²C_{n-1} 峰的¹²C_n 峰是否为双电荷离子峰以及含 N 双电荷离子。从单、双电荷离子的角度来看, 相比单电荷离子, 双电荷离子通常具有更高的 O/C、DBE、AI_{mod} 和 NOSC 值以及更低的 H/C 值, 表明大部分双电荷 DOM 是富羧基的脂环分子, 它们有着更高的芳香性、不饱和度和氧化态。从样品类型的角度来看, 不同类型的样品中双电荷 DOM 分子数量显著不同, 土壤 DOM 样品中的双电荷离子数量更丰富。进一步, 结合功能多样性指数来看, 土壤双电荷 DOM 分子的综合多样性指数最高, 说明其具有更高的芳香程度和不饱和度, 这与样品的类型以及样品来源有关。

此外, 鉴于大多数双电荷离子存在单电荷离子前体以及质谱图中成千上万个单电荷离子峰, 在不同类型 DOM 的分子式匹配过程中, 往往可以默认为单电荷离子而忽略双电荷离子, 双电荷离子对质谱数据分析的影响相对有限。该方法为解析质谱图仅考虑单电荷离子提供了依据, 同时也为 UHR-MS 数据分析提供了一个全新的视角。

参考文献 (References)

[1] QI Y L, XIE Q R, WANG J J, et al. Deciphering dissolved organic matter by Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS): From bulk to fractions and individuals[J]. *Carbon Research*, 2022, 1(1): 3.

[2] D'ANDRILLI J, FISCHER S J, ROSARIO-ORTIZ F L. Advancing critical applications of high resolution mass spectrometry for DOM assessments: Re-engaging with mass spectral principles, limitations, and data analysis[J]. *Environmental Science & Technology*, 2020, 54(19): 11654-11656.

[3] McDONOUGH L K, ANDERSEN M S, BEHNKE M I, et al. A new conceptual framework for the transformation of groundwater dissolved organic matter[J]. *Nature Communications*, 2022, 13: 2153.

- [4] van der WAL A, de BOER W. Dinner in the dark: Illuminating drivers of soil organic matter decomposition[J]. *Soil Biology and Biochemistry*, 2017, 105: 45-48.
- [5] HAN R X, LV J T, LUO L, et al. Molecular-scale investigation of soil fulvic acid and water-extractable organic matter by high-resolution mass spectrometry and ¹H NMR spectroscopy[J]. *Environmental Chemistry*, 2019, 16(2): 92.
- [6] SMITH D F, PODGORSKI D C, RODGERS R P, et al. 21 tesla FT-ICR mass spectrometer for ultrahigh-resolution analysis of complex organic mixtures[J]. *Analytical Chemistry*, 2018, 90(3): 2041-2047.
- [7] TOLIĆ N, LIU Y N, LIYU A, et al. Formularity: Software for automated formula assignment of natural and other organic matter from ultrahigh-resolution mass spectra[J]. *Analytical Chemistry*, 2017, 89(23): 12659-12665.
- [8] LEEFMANN T, FRICKENHAUS S, KOCH B P. UltraMassExplorer: A browser-based application for the evaluation of high-resolution mass spectrometric data[J]. *Rapid Communications in Mass Spectrometry*, 2019, 33(2): 193-202.
- [9] KUNENKOV E V, KONONIKHIN A S, PERMINOVA I V, et al. Total mass difference statistics algorithm: A new approach to identification of high-mass building blocks in electrospray ionization Fourier transform ion cyclotron mass spectrometry data of natural organic matter[J]. *Analytical Chemistry*, 2009, 81(24): 10106-10115.
- [10] SCHUM S K, BROWN L E, MAZZOLENI L R. MFAssignR: Molecular formula assignment software for ultrahigh resolution mass spectrometry analysis of environmental complex mixtures[J]. *Environmental Research*, 2020, 191: 110114.
- [11] KUJAWINSKI E B, BEHN M D. Automated analysis of electrospray ionization Fourier transform ion cyclotron resonance mass spectra of natural organic matter[J]. *Analytical Chemistry*, 2006, 78(13): 4363-4373.
- [12] LI P H, TAO J C, LIN J, et al. Stratification of dissolved organic matter in the upper 2000 m water column at the Mariana Trench[J]. *Science of the Total Environment*, 2019, 668: 1222-1231.
- [13] FU Q L, FUJII M, RIEDEL T. Development and comparison of formula assignment algorithms for ultrahigh-resolution mass spectra of natural organic matter[J]. *Analytica Chimica Acta*, 2020, 1125: 247-257.
- [14] 曹冬, 耿方兰, 饶子渔, 等. 傅里叶变换离子回旋共振质谱分子表征 ¹⁵N 同位素标记羟胺衍生化天然有机质 [J]. *环境化学*, 2023, 42(4): 1118-1127.
- CAO D, GENG F L, RAO Z Y, et al. Molecular characterization of ¹⁵N-labelled hydroxylamine-derivatized natural organic matter by FTICR-MS[J]. *Environmental Chemistry*, 2023, 42(4): 1118-1127 (in Chinese).
- [15] STEVENSON F J. Humus chemistry: genesis, composition, reactions[M]. 2nd ed. New York: Wiley, 1994.
- [16] KIM S, RODGERS R P, MARSHALL A G. Truly “exact” mass: Elemental composition can be determined uniquely from molecular mass measurement at ~0.1 mDa accuracy for molecules up to ~500 Da[J]. *International Journal of Mass Spectrometry*, 2006, 251(2/3): 260-265.
- [17] SENKO M W, BEU S C, McLAFFERTY F W. Automated assignment of charge states from resolved isotopic peaks for multiply charged ions[J]. *Journal of the American Society for Mass Spectrometry*, 1995, 6(1): 52-56.
- [18] BROWN T L, RICE J A. Effect of experimental parameters on the ESI FT-ICR mass spectrum of fulvic acid[J]. *Analytical Chemistry*, 2000, 72(2): 384-390.
- [19] LEENHEER J A, ROSTAD C E, GATES P M, et al. Molecular resolution and fragmentation of fulvic acid by electrospray ionization/multistage tandem mass spectrometry[J]. *Analytical Chemistry*, 2001, 73(7): 1461-1471.
- [20] STENSON A C, LANDING W M, MARSHALL A G, et al. Ionization and fragmentation of humic substances in electrospray ionization Fourier transform-ion cyclotron resonance mass spectrometry[J]. *Analytical Chemistry*, 2002, 74(17): 4397-4409.
- [21] LEENHEER J A, FERRER I, FURLONG E T, et al. Charge characteristics and fragmentation of polycarboxylic acids by electrospray ionization—Multistage tandem mass spectrometry[M]//ACS Symposium Series. Washington, DC: American Chemical Society, 2003: 312-324.
- [22] KUJAWINSKI E B, HATCHER P G, FREITAS M A. High-resolution Fourier transform ion cyclotron resonance mass spectrometry of humic and fulvic acids: Improvements and comparisons[J]. *Analytical Chemistry*, 2002, 74(2): 413-419.
- [23] KUJAWINSKI E B, FREITAS M A, ZANG X, et al. The application of electrospray ionization mass spectrometry (ESI MS) to the structural characterization of natural organic matter[J]. *Organic Geochemistry*, 2002, 33(3): 171-180.
- [24] NOVOTNY N R, CAPLEY E N, STENSON A C. Fact or artifact: The representativeness of ESI-MS for complex natural organic mixtures[J]. *Journal of Mass Spectrometry:JMS*, 2014, 49(4): 316-326.
- [25] D’ANDRILLI J, CHANTON J P, GLASER P H, et al. Characterization of dissolved organic matter in northern peatland soil porewaters by ultra high resolution mass spectrometry[J]. *Organic Geochemistry*, 2010, 41(8): 791-799.
- [26] STENSON A C, MARSHALL A G, COOPER W T. Exact masses and chemical formulas of individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectra[J]. *Analytical Chemistry*, 2003, 75(6): 1275-1284.
- [27] FU Q L, FUJII M, WATANABE A, et al. Formula assignment algorithm for deuterium-labeled ultrahigh-resolution mass spectrometry: Implications of the formation mechanism of halogenated disinfection byproducts[J]. *Analytical Chemistry*, 2022, 94(3): 1717-1725.

- [28] ZHANG H F, ZHANG Y H, SHI Q, et al. Study on transformation of natural organic matter in source water during chlorination and its chlorinated products using ultrahigh resolution mass spectrometry[J]. *Environmental Science & Technology*, 2012, 46(8): 4396-4402.
- [29] FU Q L, FUJII M, MA R. Development of a gaussian-based alignment algorithm for the ultrahigh-resolution mass spectra of dissolved organic matter[J]. *Analytical Chemistry*, 2023, 95(5): 2796-2803.
- [30] KOCH B P, DITTMAR T, WITT M, et al. Fundamentals of molecular formula assignment to ultrahigh resolution mass data of natural organic matter[J]. *Analytical Chemistry*, 2007, 79(4): 1758-1763.
- [31] KOCH B P, DITTMAR T. From mass to structure: An aromaticity index for high-resolution mass data of natural organic matter[J]. *Rapid Communications in Mass Spectrometry*, 2006, 20(5): 926-932.
- [32] KELLERMAN A M, DITTMAR T, KOTHAWALA D N, et al. Chemodiversity of dissolved organic matter in lakes driven by climate and hydrology[J]. *Nature Communications*, 2014, 5: 3804.
- [33] LaROWE D E, van CAPPELLEN P. Degradation of natural organic matter: A thermodynamic analysis[J]. *Geochimica et Cosmochimica Acta*, 2011, 75(8): 2030-2042.
- [34] KIDA M, MERDER J, FUJITAKE N, et al. Determinants of microbial-derived dissolved organic matter diversity in Antarctic Lakes[J]. *Environmental Science & Technology*, 2023, 57(13): 5464-5473.
- [35] FU Q L, FUJII M, KWON E. Development and application of a high-precision algorithm for nontarget identification of organohalogens based on ultrahigh-resolution mass spectrometry[J]. *Analytical Chemistry*, 2020, 92(20): 13989-13996.
- [36] GASPAR A, KUNENKOV E V, LOCK R, et al. Combined utilization of ion mobility and ultra-high-resolution mass spectrometry to identify multiply charged constituents in natural organic matter[J]. *Rapid Communications in Mass Spectrometry:RCM*, 2009, 23(5): 683-688.
- [37] CHEN S, XIE Q R, SU S H, et al. Source and formation process impact the chemodiversity of rainwater dissolved organic matter along the Yangtze River Basin in summer[J]. *Water Research*, 2022, 211: 118024.
- [38] BAE E, YEO I J, JEONG B, et al. Study of double bond equivalents and the numbers of carbon and oxygen atom distribution of dissolved organic matter with negative-mode FT-ICR MS[J]. *Analytical Chemistry*, 2011, 83(11): 4193-4199.
- [39] YU K, DUAN Y H, GAN Y Q, et al. Anthropogenic influences on dissolved organic matter transport in high arsenic groundwater: Insights from stable carbon isotope analysis and electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry[J]. *The Science of the Total Environment*, 2020, 708: 135162.
- [40] CHEN X, LIU J H, CHEN J F, et al. Oxygen availability driven trends in DOM molecular composition and reactivity in a seasonally stratified fjord[J]. *Water Research*, 2022, 220: 118690.