Original Article

Identification of Anthropogenic Compounds in Urban Environments and Evaluation of Automated Methods for Reading Fragmentation—A Case of River Water

Atsushi Yamamoto*,1,2,3, Naoko Matsumoto2, Hideya Kawasaki2, and Ryuichi Arakawa2

1 Osaka City Institute of Public Health and Environmental Sciences, 8–34 Tojocho, Tennōji, Osaka 543–0026, Japan
2 Kansai University, 3–3–35 Yamatecho, Suita, Osaka 564–8680, Japan
3 Tottori University of Environmental Studies, 1–1–1 Wakabadaikita, Tottori 689–1111, Japan

A workflow based on liquid chromatography/high-resolution mass spectrometry (LC/HR-MS) was applied for the identification of compounds in urban environments. Substances extracted by solid-phase extraction from river water were wholly analyzed by LC/HR-MS without any purification. Fragmentation in collision-induced dissociation was manually studied for the 20 most intense ions in positive- and negative-ion electrospray ionization with accurate mass determination at a resolution of 100,000. Sixteen anthropogenic compounds in the extract were identified and confirmed using standard reference reagents. These compounds consisted of pharmaceuticals, surfactants, flame retardants, and industrial intermediates. The majority of the compounds are common in our daily life. In the identification process, two automated methods, MAGMa and MetFrag/MetFusion, for reading fragmentation were evaluated for the sixteen compounds. Although automated methods could be used to retrieve the correct molecular structures in most cases, they could not always be promoted to the top rank. Automated methods have yet to be a complete solution for identifying chemical compounds, but will considerably reduce the burden for humans in reading fragmentation.

Copyright © 2016 Atsushi Yamamoto, Naoko Matsumoto, Hideya Kawasaki, and Ryuichi Arakawa. This is an open access article distributed under the terms of Creative Commons Attribution License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Please cite this article as: Mass Spectrom (Tokyo) 2016; 5(1): A0045

Keywords: high-resolution mass spectrometry, urban environment, environmental analysis, fragmentation

(REceived January 18, 2016; Accepted March 8, 2016)

INTRODUCTION

Mass spectrometry (MS) is a key technology for the identification and structural elucidation of small molecules. The molecular formulas of small molecules can be determined from accurate mass values measured by high-resolution (HR) MS because relative atomic masses other than carbon are not strictly integer quantities. Furthermore, there is a close relationship between molecular structure and reactivity of ions in MS, so information on molecular structure can be obtained by interpretation of their fragment ions in the mass spectra. HRMS can make interpretation of mass spectra fairly certain due to its high accuracy and precision. Although a mass spectral library of reference standards is quite helpful in the process of interpreting mass spectra, existing libraries are insufficient for reading all observed data. Of course, attempts to enrich the libraries are continuing, but the situation is not improving rapidly. In addition, ionization and fragmentation are sometimes inherent to instrumental conditions particularly in liquid chromatography/mass spectrometry, which is often applied to polar compounds. Therefore, manual interpretation of the resulting data continues to be time consuming and cumbersome.

During the last five years, several promising automated methods for interpreting accurate mass spectral data have appeared on the scene. One of the methods is the utilization of public molecular structure databases such as KEGG (Kyoto Encyclopedia of Genes and Genomes), PubChem, and ChemSpider, which are more comprehensive than are mass spectral libraries. Insufficiency in mass spectral libraries is considered serious even in the field of environmental research because the number of unregulated emerging micropollutants has increased over the last few decades. The environmental risk caused by unregulated pollutants is of great concern, and environmental researchers are always finding new targets. Many were studied and some of them have been introduced into regulation. However, there is no method by which environmental researchers can evaluate all emerging pollutants and their transformation products.

*Correspondence to: Atsushi Yamamoto, Osaka City Institute of Public Health and Environmental Sciences, 8–34 Tojocho, Tennōji, Osaka 543–0026, Japan, e-mail: ay@pfos.jp
Regulation has been implemented in order of priority level, but potential threats to ecosystems and human health by overlooked compounds remain.

During the last decade, new methodology coupled with HR-MS, referred as to suspect screening and nontarget analysis, has been introduced in environmental analysis. Hug et al. and Schymanski et al. found novel micropollutants and transformation products in their nontarget strategy. In this study, we used two automated methods, MAGMa and MetFrag/MetFusion, to identify compounds in an urban aquatic environment and evaluate the two methods using the results. The methods are available on the Internet and they performed well in a public contest, CASMI2013 (Critical Assessment of Small Molecule Identification).

EXPERIMENTAL

Sample preparation
In February 2015, 500 mL of river water was collected from Neyagawa River, Osaka City, Japan, and passed through a solid phase extraction cartridge, Presep-C Agri Short (Divinylbenezene polymethacrylate copolymer sorbent, 200 mg, Wako Pure Chemical Industries, Ltd.) without pH adjustment. After drying at reduced pressure by a vacuum manifold assembly, substances that had absorbed on the cartridge were eluted by 5 mL of acetonitrile. The eluate was analyzed by LC/MS without further cleanup. A blank sample preparation of an Ultimate 3000 (Dionex, Sunnyvale, CA, USA) and an Exact Orbitrap Mass Spectrometer (Thermo Fisher Scientific Inc., Waltham, MA, USA) was used. Both positive- and negative-ion electrospray ionization (ESI) were applied. Mass spectra were acquired from m/z 100 to 2,000 with a resolution of 100,000 and a mass accuracy of 5 ppm in ESI. Without advance information, we usually cannot know whether these ions indicate which kind of ionic form, such as molecular ions [M]+, sodiated molecules [M]+Na, or protonated molecules [M+H]+, are retrieved from a compound database by querying on monoisotopic mass. Next, fragmentation of structural skeletons (i.e., non-hydrogen atoms) is performed by removing each non-hydrogen atom sequentially and collecting substructures. During the fragmentation, a simple penalty score that depends on the type of bond is taken into account. Then in silico generated substructures are assigned to accurate m/z values in a mass spectrum. Referred molecular structures are retrieved from chemical compound database such as PubChem, Human Metabolite Database (HMDB), and KEGG. An online version, MAGMa, has a user-friendly web interface, http://www.emetabolomics.org/magma/. Use of this method was selected as the best automated tool of the international contest, CASMI2013, which tested the ability to explain blind mass spectral data. In the present study, PubChem was selected as the compound database.

Approaches to explain the mass spectra examined in this study

MAGMa
Ridder et al. introduced an automated method for the interpretation of accurate mass spectra in 2012. The method is based on an algorithm for candidate substructure annotation of multistage accurate mass spectral trees without relying on a spectral library. First, candidate structures are retrieved from a compound database by querying on monoisotopic mass. Next, fragmentation of structural skeletons (i.e., non-hydrogen atoms) is performed by removing each non-hydrogen atom sequentially and collecting substructures. During the fragmentation, a simple penalty score that depends on the type of bond is taken into account. Then in silico generated substructures are assigned to accurate m/z values in a mass spectrum. Referred molecular structures are retrieved from a chemical compound database such as PubChem, Human Metabolite Database (HMDB), and KEGG. An online version, MAGMa, has a user-friendly web interface, http://www.emetabolomics.org/magma/. Use of this method was selected as the best automated tool of the international contest, CASMI2013, which tested the ability to explain blind mass spectral data. In the present study, PubChem was selected as the compound database.

MetFrag/MetFusion
MetFrag was developed in 2010 by Wolf et al. MetFrag also simulates fragmentation to explain ions in a mass spectrum. Before bond disconnection of candidate structures, a small set of rules for molecular rearrangements are applied. Then all bonds to be disconnected are labeled linear or ring. A redundancy check is performed at every disconnection process. The in silico fragments are matched against the query peak list. Referred molecular structures are retrieved from a chemical compound database such as PubChem, KEGG, or ChemSpider. A web application is available at the following URL: http://msbi.ipb-halle.de/MetFrag/. MetFusion is the successor of MetFrag and was released by Gerlich et al. in 2013. MetFusion incorporates the similarity using a mass spectral library in MetFrag scoring. Its spectral library refers to MassBank, HMDB, NIST'11, and METLIN. As with MAGMa, PubChem was selected as the compound database.

Preparation of queries for automated methods
Initially, obvious noise and substances detected in the blank experiment were excluded from the examination for compound identification. The intense ions in the mass spectra were extracted from each LC/MS analysis according to intensity. In ESI, without advance information, we usually cannot know whether these ions indicate which kind of ionic form, such as molecular ions [M]+, protonated molecules [M+H]+, sodiated molecules [M+Na]+, or fragments of unstable precursors. Relevant ions with identical chromatographic peak shapes were extracted in order to assign specific ionic forms to them. The difference in m/z values between the ions offers good criteria for annota-
tion. For example, the m/z differences of [M+Na]⁺ from [M+NH₄]⁺ and [M+H]⁺ are 4.955 and 21.982, respectively. Another LC/MS analysis in opposite polarity also can be used as criteria. Ions with an m/z difference of 2.011 with identical chromatographic peak shapes in both polarities is evidence of the presence of [M+H]⁺ and [M−H]⁻. Ions related to isotopologues and adduct ions were grouped as a single feature. The 20 most intense features were extracted. After the assignment of precursor ions, analyses were carried out in AIF mode. Ions with the same chromatographic peak shapes as precursor ions were extracted to find fragment ions, just as the aforementioned assignment of a feature. Ions with great differences in the first decimal place of the m/z values from that of precursor ions were excluded. A list of accurate masses of precursor and fragment ions was consolidated to obtain a virtual precursor ion and product ion relationship for each feature. This process is quite laborious but necessary because our Exactive is not a multistage mass spectrometer. The list was converted to appropriate formats depending on the requirements of each automated method. All m/z and peak area values used for evaluating automated methods are summarized in the electronic supplementary material (Table S7).

RESULTS AND DISCUSSION

Table 1 is a list of the 20 most intense features in positive- and negative-ion ESI and m/z values of ions with highest intensity in each feature. In positive ion mode, feature P1 with the highest intensity among the features showed an m/z value of 240.1497. P1 included a weak ion that had an m/z of 479.2950, the value of twofold minus 1.005. This relationship could be considered that of [M+H]⁺ and [2M+H]⁺. Therefore, the ion of m/z 240.1497 could be annotated as [M+H]⁺. The m/z values of the examples of the major fragment ions were 106.0654, 108.0811, 133.0763, and 223.1232. Typical chromatograms are shown in Fig. 1. Table 2 is a list of m/z value, peak area, and its isotopologue composition of suspected fragment ions of P1. The isotopologue composition indicated the absence of elements that had a specificity at X+2 such as chlorine, bromine, and sulfur, and the presence of roughly 15 carbon atoms. In addition, the presence of C₆H₄⁺ (P1AIF1 in Table 2) suggested a benzyl or tolyl substructure. The candidate formula of fragment ion P1AIF7 corresponded to the neutral loss of fragment ions P1AIF3 or P1AIF4 in Table 2. The values of ring and double bond equivalents (RDB) of P1AIF3, 4, and 7 were fairly consistent with the inclusion of a tolyl or benzyl substructure. Therefore a highly symmetric structure was suspected. Although the m/z value, 240.1497, has several candidate formulas such as C₅H₅NOP, C₁₅H₁₈N₃, and C₁₇H₂₀O, only C₁₅H₁₈N₃ is consistent with the elemental compositions of the fragment ions of P1 because the precursor ion must include nitrogen atoms. Given that feature P1 is composed of two identical substructures, the consistent substructures were considered to be aminobenzyl, anilinomethyl, benzylamino, or toluidino groups. Each substructure contributes 4 of RDB. Hence, the residual structure must consist of one atom each.

Table 1. Twenty most intense features in the LC/MS analysis of river water extract. Representative accurate masses, retention time, and peak areas in each feature are shown. Related ions such as isotopologues and adducts are consolidated in one feature.

| Feature | Accurate mass | RT/min | Peak area |
|---------|--------------|--------|-----------|
| P1      | 240.1497     | 6.5    | 21795672  |
| P2      | 212.1184     | 4.4    | 11645483  |
| P3      | 399.2509     | 21.3   | 4832006   |
| P4      | 368.4253     | 23.0   | 4224291   |
| P5      | 235.1806     | 4.8    | 3697944   |
| P6      | 298.2226     | 4.1    | 2799989   |
| P7      | 204.1385     | 16.5   | 2521985   |
| P8      | 312.3626     | 22.1   | 2312859   |
| P9      | 342.2489     | 4.1    | 1517764   |
| P10     | 391.2303     | 13.7   | 1315704   |
| P11     | 230.2480     | 17.5   | 1282643   |
| P12     | 515.2446     | 17.0   | 1255808   |
| P13     | 252.2171     | 14.5   | 1071096   |
| P14     | 369.3342     | 22.0   | 1070924   |
| P15     | 326.3738     | 21.5   | 81233     |
| P16     | 248.3275     | 15.0   | 795483    |
| P17     | 294.3550     | 13.9   | 786037    |
| P18     | 748.4850     | 13.7   | 7537304   |
| P19     | 284.3313     | 21.3   | 749116    |
| P20     | 310.2590     | 16.5   | 735019    |

| Feature | Accurate mass | RT/min | Peak area |
|---------|--------------|--------|-----------|
| N1      | 325.1834     | 20.1   | 1503900   |
| N2      | 311.1676     | 18.9   | 14519946  |
| N3      | 339.1990     | 21.6   | 10083557  |
| N4      | 297.1523     | 17.7   | 5756181   |
| N5      | 421.2236     | 20.4   | 1298573   |
| N6      | 180.9721     | 2.5    | 1071585   |
| N7      | 248.9598     | 2.5    | 573028    |
| N8      | 396.2415     | 16.2   | 564249    |
| N9      | 313.1112     | 3.0    | 465045    |
| N10     | 253.2164     | 24.3   | 417956    |
| N11     | 651.3744     | 20.0   | 356371    |
| N12     | 623.3431     | 18.9   | 341778    |
| N13     | 316.9471     | 2.5    | 269727    |
| N14     | 513.2285     | 17.0   | 262597    |
| N15     | 128.9585     | 2.5    | 215132    |
| N16     | 283.1367     | 16.6   | 162474    |
| N17     | 281.0849     | 3.1    | 137739    |
| N18     | 264.9338     | 2.5    | 129963    |
| N19     | 543.0455     | 16.5   | 125226    |
| N20     | 269.0369     | 17.8   | 109648    |

*RT, retention time.

Fig. 1. Extracted ion chromatograms of the most intense ion, m/z 240.1497, under MS mode and its five suspected fragment ions under AIF mode. Data was acquired using a river extract sample by LC/MS running in positive ion ESI. The m/z values extracted are shown in the figure.
of carbon, hydrogen, nitrogen, and one double bond, and connect two substructures at the carbon atom. Features with substructures of the benzylamino and toluidino groups can be considered having a guanidine structure. Figure 2 shows the probable structures for \( P1 \), which correspond to the toluidino (A), benzylamino (B), anilinomethyl (C), and aminobenzyl (D) substructures. The cleavage positions of bonds to generate the fragment ions listed in Table 2 also shown in the figure. The numbers of positions correspond to those in the rightmost row of Table 2.

![Fig. 2. Probable structures of P1: (A) Ditolylguanidine; (B) Dibenzylguanidine; (C) 1,3-Dianilinopropane-2-imine; and (D) 1,3-Bis(aminophenyl)-propane-2-imine. Cleavage positions of bonds to generate the fragment ions listed in Table 2 also shown in the figure. The numbers of positions correspond to those in the rightmost row of Table 2.](image)

| No. | Accurate mass | Peak area | Isotope pattern | Candidate formula | Exact mass | Error (ppm) | RDB | Counterpart to \([M+H]^+\) | Corresponding scission* |
|-----|---------------|-----------|-----------------|-------------------|------------|-------------|-----|----------------------|----------------------|
| P1AIF1 | 91.0546 | 187745 | 100 | 7.1 | \( \text{C}_7\text{H}_7 \) | 91.0542 | 4.4 | 4.5 | \( \text{C}_6\text{H}_7\text{N}^+ \) | iii or ii, vi (+1) |
| P1AIF2 | 93.0576 | 207917 | 100 | 4.3 | \( \text{C}_8\text{H}_{11}\text{N}_3 \) | 93.0573 | 3.2 | 4 | \( \text{C}_7\text{H}_6\text{N}_2 \) | ii, iv (+2) or v |
| P1AIF3 | 106.0866 | 1770545 | 100 | 7.3 | \( \text{C}_9\text{H}_{11}\text{N}_2 \) | 106.0651 | 2.8 | 4.5 | \( \text{C}_8\text{H}_6\text{N}_2 \) | ii |
| P1AIF4 | 108.0811 | 3758787 | 100 | 7.6 | \( \text{C}_{10}\text{H}_{11}\text{N}_2 \) | 108.08088 | 2.8 | 3.5 | \( \text{C}_9\text{H}_8\text{N}_2 \) | ii (+2) |
| P1AIF5 | 116.0497 | 254521 | 100 | — | \( \text{C}_{11}\text{H}_{11}\text{N}^+ \) | 116.0495 | 1.7 | 6.5 | \( \text{C}_{10}\text{H}_9\text{N}_2 \) | i, ii (-1) |
| P1AIF6 | 118.0654 | 75726 | 100 | — | \( \text{C}_{12}\text{H}_{11}\text{N}^+ \) | 118.0651 | 2.5 | 5.5 | \( \text{C}_{13}\text{H}_9\text{N}_2 \) | i, ii (+1) |
| P1AIF7 | 133.0763 | 4445934 | 100 | 8.6 | \( \text{C}_{13}\text{H}_{12}\text{N}_2 \) | 133.0765 | 2.3 | 5.5 | \( \text{C}_{14}\text{H}_8\text{N}_2 \) | ii |
| P1AIF8 | 208.0998 | 79498 | 100 | 10.9 | \( \text{C}_{15}\text{H}_{14}\text{N}_2 \) | 208.09955 | 1.4 | 10 | \( \text{C}_{16}\text{H}_8\text{N}_2 \) | i, iv |
| P1AIF9 | 223.1232 | 1394438 | 100 | 12.2 | \( \text{C}_{17}\text{H}_{16}\text{N}_2 \) | 223.12307 | 0.9 | 9.5 | \( \text{C}_{18}\text{H}_9\text{N}_2 \) | i |
| P1AIF10 | 240.1497 | 3500331 | 100 | 16.2 | \( \text{C}_{19}\text{H}_{18}\text{N}_3 \) | 240.14955 | 0.8 | 8.5 | \( \text{C}_{20}\text{H}_{10}\text{N}_3 \) | |

* The numbers of scission positions in this column are illustrated in Fig. 2. The number in parentheses indicates mass shift in nominal mass.

Table 2. Accurate masses and peak areas of suspected fragment ions of \( P1 \) in all ion fragmentation (AIF) mode. Isotopologue composition and candidate molecular formulas are also shown.
dine could easily be inferred. It was purchased from Tokyo Chemical Industry (Tokyo, Japan) to confirm the identification. After submitting queries of diphenylguanidine to both MAGMa and MetFrag, the correct answer ranked 3rd in 1,442 candidates and 88th in 1,499 candidates, respectively.

In a similar way, fragmentation data were sequentially interpreted to determine compounds. All identification processes were completed by using authentic standard reagents. Identified compounds and their results with MAGMa and MetFrag are listed in Table 3. Detailed interpretation of each feature is provided in the electronic supplementary material. Identification could be completed for more than half of the features. Automated methods were able to retrieve the correct answers but not always promote them to be top rank.

In the case of negative ion mode, the top four features had a constant interval of \( m/z \) 14.015 corresponding to \( \text{CH}_2 \). Their chromatograms are shown in Fig. 3 and indicate the presence of many isomers. Furthermore, in their fragmentation, these substances resembled a principal fragment ion, \( m/z \) 183.0110. Neutral losses of the precursor ions correspond to the values of \( \text{C}_n\text{H}_{2n+2} \). This indicates the presence of an identical substructure and an alkyl chain with a different chain length and different substitution position. The \( m/z \) values of the common fragment ions were 183.0110 and 119.0491. The isotopologue composition of these fragment ions was different in \( X-2 \). The mass defect of these fragment ions indicated a loss of an element with a large negative mass defect such as sulfur. The accurate mass difference 63.9619 between the two fragment ions corresponded exactly to sulfur dioxide. The elemental composition candidates of the fragment ion 119.0491 included \( \text{C}_6\text{H}_5\text{N}_3 \) and \( \text{C}_8\text{H}_7\text{O} \). \( \text{C}_8\text{H}_7\text{O} \) was more likely because the RDB of \( \text{C}_6\text{H}_5\text{N}_3 \) was

| No. | Substance | Structure | Submitted ions | MAGMa | MetFrag |
|-----|-----------|-----------|----------------|-------|---------|
| P1  | 1,3-Di-o-tolyguanidine (\( \text{C}_{15}\text{H}_{17}\text{N}_3 \), CAS RN 97-39-2) | ![Structure](image) | 11 6(14)/2708* | 100(6)/2738 |
| P2  | 1,3-Diphenylguanidine (\( \text{C}_{13}\text{H}_{13}\text{N}_3 \), CAS RN 102-06-7) | ![Structure](image) | 5 3(26)/1442 | 88(10)/1499 |
| P3  | Tris(butoxyethyl)phosphate (\( \text{C}_{18}\text{H}_{39}\text{O}_7\text{P} \), CAS RN 78-51-3) | ![Structure](image) | 5 1(5)/785 | 8(16)/1249 |
| P7  | Crotamiton (\( \text{C}_{14}\text{H}_{37}\text{NO} \), CAS RN 483-63-6) | ![Structure](image) | 12 13(10)/6116 | 13(10)/6409 |
| P9  | Ethylene oxide | ![Structure](image) | | |
| P10 | 2,4,6-Tris(methoxymethyl)aminol-1,3,5-triazine (\( \text{C}_{23}\text{H}_{30}\text{N}_6\text{O}_6 \), CAS RN 3089-11-0) | ![Structure](image) | 8 18(5)/119 | 59(3)/390 |
| P11 | Dimethyldodecylamine oxide (\( \text{C}_{14}\text{H}_{31}\text{NO} \), CAS RN 1643-20-5) | ![Structure](image) | 2 1(1109)/1109 | 1(1137)/1137 |
| P12 | Telmisartan (\( \text{C}_{33}\text{H}_{30}\text{N}_4\text{O}_2 \), CAS RN 144701-48-4) | ![Structure](image) | 3 1(26)/424 | 2(27)/1029 |
| P14 | Stearamidopropyl dimethylamine (\( \text{C}_{23}\text{H}_{48}\text{N}_2\text{O} \), CAS RN 7651-02-7) | ![Structure](image) | 4 1(11)/114 | 2(10)/103 |
| P16 | Benzalkonium chloride (\( \text{C}_{17}\text{H}_{30}\text{N}^+ \), CAS RN 46917-11-7) | ![Structure](image) | 4 1(36)/2821 | 1(10)/113 |
| P18 | Clarithromycin (\( \text{C}_{38}\text{H}_{69}\text{NO}_13 \), CAS RN 81103-11-9) | ![Structure](image) | 5 3(11)/48 | 2(10)/47 |

* The denominators, numerators, and parentheses, respectively, denote the number of retrieved substances, the order of the correct answer, and the number of substances with the same score as the correct answer.
Nontarget Analysis of the Aquatic Environment

Fig. 3. Extracted ion chromatograms of features N1–N4 (C_{10}LAS–C_{13}LAS) in the negative ion ESI analysis. The same sample as used in Fig. 1 was used. The m/z values extracted are shown in the figure.

Table 4. Identified compounds in negative ion mode, number of ions in queries, and results by two automated methods.

| No. | Substance | Structure | Submitted ions | MAGMa | MetFrag |
|-----|-----------|-----------|----------------|-------|---------|
| N1  | C_{12}-LAS (C_{18}H_{29}O_{3}NaS, CAS RN 25155-30-0) | ![Structure](image1) | 3 | 1(184)/184* | 8(89)/515 |
| N2  | C_{11}-LAS (C_{17}H_{27}O_{3}NaS, CAS RN 27636-75-5) | ![Structure](image2) | 3 | 1(170)/170 | 16(52)/579 |
| N3  | C_{13}-LAS (C_{19}H_{31}O_{3}NaS, CAS RN 26248-24-8) | ![Structure](image3) | 4 | 2(153)/154 | 8(32)/124 |
| N4  | C_{10}-LAS (C_{16}H_{25}O_{3}NaS, CAS RN 1322-98-1) | ![Structure](image4) | 3 | 1(219)/219 | 4(442)/446 |
| N5  | Aerosol OT (C_{20}H_{37}O_{7}SNa, CAS RN 577-11-7) | ![Structure](image5) | 5 | 1(42)/4160 | 35(30)/300 |
| N11 | Dimer of C_{12}-LAS | | | | |
| N12 | Dimer of C_{11}-LAS | | | | |
| N14 | Telmisartan (C_{33}H_{30}N_{4}O_{2}, CAS RN 144701-48-4) | ![Structure](image6) | 4 | 364(144)/497 | 4(18)/739 |
| N19 | Bicalutamide** (C_{33}H_{26}N_{4}O_{2}F_{3}S, CAS RN 90357-06-5) | ![Structure](image7) | 5 | 25(6)/1071 | 3(5)/1190 |

* The denominators, numerators, and parentheses, respectively, denote the number of retrieved substances, the order of the correct answer, and the number of substances with the same score as the correct answer.

** The query was submitted under the condition of 5 ppm tolerance and halogen inclusion.
were submitted to both MetFrag and MAGMa. MAGMa and MetFrag retrieved 219 and 446 candidates for C_{10}LAS, respectively. The correct answer ranked 1st in MAGMa and 4th in MetFrag, respectively. However, numerous candidates ranked at the same position as the correct answer. Although the retrieved candidates included LAS, neither tool could explain these fragment ions. In the fragmentation of sulfonate-type surfactants, a product ion, SO_{3}−, is conceivable. The present collision energy, however, was not enough to generate SO_{3}−, which would improve the results. Identified compounds and their results of MAGMa and MetFrag are listed in Table 4. Detailed interpretation of each compound is provided in electronic supplementary material.

Schymanski et al. has proposed a level system to facilitate the communication of compound identification confidence in the field of environmental research.15) According to the level system, the 16 present compounds reached Level 1 confidence (Confirmed structure). These compounds consisted of pharmaceutical compounds, surfactants, a flame retardant, and other industrial materials. Osaka City has a population of about 2.6 million and is situated at the mouth of the Yodo River on Osaka Bay. There are many streams and watercourses, which often become stagnant or even flow backwards. Hence, the aquatic environment is considerably influenced by human activity.16,17) The compounds found in the present study also reflected this situation. Crotamiton, telmisartan, and clarithromycin are common pharmaceutical compounds that are often targeted in the context of environmental research.18,19) In Japan, more than 6 million patients suffer from high blood pressure.20) The unequivocal detection of telmisartan appears quite conceivable. However, several compounds such as bicalutamide and 1,3-di-o-tolyguanidine have rarely been targeted. 1,3-Di-o-tolyguanidine is known not only as a vulcanization accelerator in rubber production but also as a ligand of the σ receptor in the central nervous system21) and designated in Monitoring Chemical Substances found in Japan’s Act on the Evaluation of Chemical Substances and Regulation of Their Manufacture, etc. Its annual manufactured and imported quantity in Japan was over 100 t from FY 2007 to FY 2009.22) Subsequent environmental research on these compounds is expected.

In Schymanski's level system, several features remained at Level 5 (Exact mass). Several of them had a relationship like a constant m/z difference. Figure 4 shows an example of two homologous series ions with an m/z difference of 28.031 corresponding to CH_{3}. Features P4, P8, P15, and P19 were included in these series. Although the m/z differences of 14.015 were found between the series, it seemed the two series had different structures because the retention time was not consecutive. These features were remarkable but it is unusual to keep a record of such features in scientific communication because their exact identification could not be ascertained. Even though their formulas could not be assigned, we believe that reciprocation of such data under assignment Level 5 is still meaningful for a better understanding of chemical substances in the environment.

**CONCLUSION**

In this study, nontarget analysis was conducted to identify anthropogenic compounds in an urban aquatic environment and two automated methods were evaluated using the results. Both MAGMa and MetFrag retrieved correct answers in most queries. They also usually retrieved many incorrect answers with the same score as the correct answer. Increasing the number of submitted fragment ions has not always improved results. Both tools appeared to be equal in their results in the present sample, but MAGMa offered a great advantage in terms of processing time.

To explain the fragment ions in a query, MAGMa and MetFrag use disconnecting bonds of numerous compounds in molecular structure databases. This procedure is impossible to conduct manually, and considerably reduces the burden on humans in reading fragmentation. Of course, that reason is not enough to recommend the procedure and it should be complemented from broader perspectives such as chemical reaction probability, isotope pattern, consistency between precursor structure and product structure, and specific neutral loss on substituent group or element, as described in the electronic supplementary material. In addition to the tools used here, many automated tools that include different perspectives are available. To complement legacy target analysis for environmental risk evaluation of chemical compounds, new methodology is anticipated. Such future development would also be a breakthrough in the field of environmental research.

**Acknowledgements**

This work was financially supported by JSPS KAKENHI Grant Number 26241026 and 26505011.
Electronic supplementary material
Supplementary material to this article can be found online.

REFERENCES

1) F. Hernández, O. J. Pozo, J. V. Sancho, F. J. López, J. M. Martín, M. Ibáñez. Strategies for quantification and confirmation of multi-class polar pesticides and transformation products in water by LC-MS2 using triple quadrupole and hybrid quadrupole time-of-flight analyzers. Trends Anal. Chem. 24: 596–612, 2005.

2) M. Petrovic, D. Barceló. Application of liquid chromatography/mass spectrometry to the environmental analysis. J. Mass Spectrom. 41: 1259–1267, 2006.

3) A. C. Hogenboom, J. A. van Leerdam, P. de Voogt. Accurate mass screening and identification of emerging contaminants in environmental samples by liquid chromatography–hybrid linear ion trap Orbitrap mass spectrometry. J. Chromatogr. A 1216: 510–519, 2009.

4) M. Krauss, H. Singer, J. Hollender. LC-High resolution MS in environmental analysis: From target screening to the identification of unknowns. Anal. Bioanal. Chem. 397: 943–951, 2010.

5) S. Suzuki. Non-target analysis of environmental contaminants by LC/Q-ToFMS/MS. Bunseki Kagaku 62: 379–391, 2013, in Japanese.

6) C. Hug, N. Ulrich, T. Schulze, W. Brack, M. Krauss. Identification of novel micropollutants in wastewater by a combination of suspect and nontarget screening. Environ. Pollut. 184: 25–32, 2014.

7) E. L. Schymanski, H. P. Singer, P. Longrée, M. Loos, M. Ruff, M. A. Stravs, C. R. Vidal, J. Hollender. Strategies to characterize polar organic contamination in wastewater: Exploring the capability of high resolution mass spectrometry. Environ. Sci. Technol. 48: 811–818, 2014.

8) L. Ridder, J. J. van der Hoof, S. Verhoeven, R. C. H. de Vos, R. van Schaik, J. Vervoort. Substructure-based annotation of high-resolution multistage MSn spectral trees. Rapid Commun. Mass Spectrom. 26: 2461–2471, 2012.

9) S. Wolf, S. Schmidt, M. Müller-Hannemann, S. Neumann. In silico fragmentation for computer assisted identification of metabolite mass spectra. BMC Bioinform. 11: 148, 2010.

10) M. Gerlich, S. Neumann. MetFusion: Integration of compound identification strategies. J. Mass Spectrom. 48: 291–298, 2013.

11) T. Nishioka, T. Kasama, T. Kinumi, H. Makabe, F. Matsuda, D. Miura, M. Miyashita, T. Nakamura, K. Tanaka, A. Yamamoto. Winners of CASMI2013: Automated tools and challenge data. Mass Spectrom. (Tokyo) 3: S0039, 2014.

12) P. A. Lyon, W. L. Stebbings, F. W. Crow, K. B. Tomer, D. L. Lippstreu, M. L. Gross. Analysis of anionic surfactants by mass spectrometry/mass spectrometry with fast atom bombardment. Anal. Chem. 56: 8–13, 1984.

13) Japan’s Ministry of Economy, Trade and Industry. Public announcement of the annual quantity of manufacture and import of Chemical Substances. 2015.

14) L. Solar, S. Rubio, D. Pérez-Bendito. Differentiation and quantification of linear alky benzene sulphonate isomers by liquid chromatography-ion-trap mass spectrometry. J. Chromatogr. A 1031: 17–25, 2004.

15) E. L. Schymanski, J. Jeon, R. Gwede, K. Fenner, M. Ruff, H. P. Singer, J. Hollender. Identifying small molecules via high resolution mass spectrometry: Communicating confidence. Environ. Sci. Technol. 48: 2097–2098, 2014.

16) A. Yamamoto, N. Kakutani, K. Yamamoto, T. Kamiura, H. Miyakoda. Steroid hormone profiles of urban and tidal rivers using LC/MS/MS equipped with electrospray ionization and atmospheric pressure photoionization sources. Environ. Sci. Technol. 40: 4132–4137, 2006.

17) F. Adachi, A. Yamamoto, K. Takakura, K. Kawahara. Occurrence of fluoroquinolones and fluoroquinolone-resistance genes in the aquatic environment. Sci. Total Environ. 444: 508–514, 2013.

18) M. Hanamoto, H. Sugishita, N. Yamashita, H. Tanaka, I. Hoawa, C. Koshi. Evaluation of dynamics of PPCPs at Yodo River System. Proc. Environ. Eng. Res. 45: 29–37, 2008, in Japanese.

19) R. Loo, R. Carvalho, Y. C. António, S. Comero, G. Locoro, S. Tavazzi, B. Paracchini, M. Ghiani, T. Lettieri, L. Blaha, B. Jarosova, S. Voorspoels, K. Servaes, P. Haglund, J. Fick, R. H. Lindberg, D. Scheweig, B. M. Gawlik. EU-wide monitoring survey on emerging polar organic contaminants in wastewater treatment plant effluents. Water Res. 47: 6475–6487, 2013.

20) Japan’s Ministry of Health, Labour and Welfare. Patient Survey. 2011.

21) P. M. Beart, R. D. O’Shea, D. T. Manallack. Regulation of ß2-receptors: High- and low-affinity agonist states, GTP shifts, and up-regulation by rimcazole and 1,3-di(2-tolyl)guanidine. J. Neurochem. 53: 779–788, 1989.

22) Japan’s Ministry of Economy, Trade and Industry. Public announcement of the annual quantity of manufacture and import of Type II Monitoring Chemical Substances. 2008, 2009, and 2010.