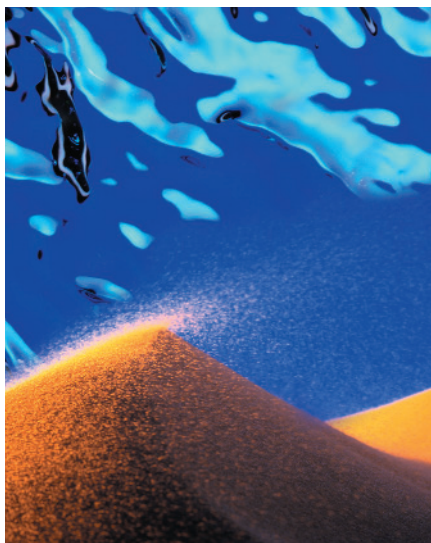


# Identification of Unanticipated Compounds by High-Resolution Mass Spectrometry

Andrew H. Grange and G. Wayne Sovocool

Often, mass spectra of compounds found in water supplies are not found in mass spectral libraries. Ion composition elucidation (ICE) provides the elemental compositions of the ions observed in their mass spectra. This information is used to identify the compounds.



Localized outbreaks of acute illness could result from deliberate addition of unanticipated compounds into water, air, or food. Cancer clusters resulting from long-term exposure to trace levels of compounds are more difficult to detect, but can be revealed by epidemiological studies. To find the causative agents in either case, the compounds in sample extracts must be identified before any contribution they make toward the observed malady can be assessed.

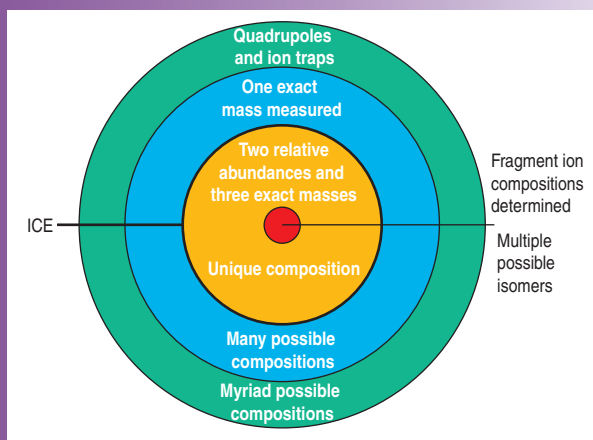
Most trace-level compounds will pose little or no risk. However, all compounds unique to the locality should be identified, quantified, and classified by their toxicity to ensure that no toxic compounds have been overlooked.

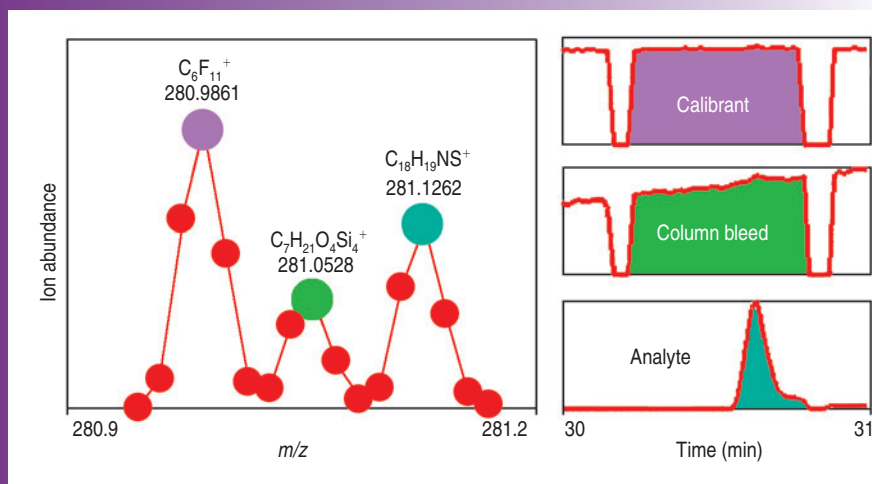
Sample extracts are often complex mixtures. Separation of components in time by gas chromatography followed by mass selection of ions using mass spectrometry (GC-MS) provides distinctive mass spectra for many organic compounds. Mass spectrometer data systems permit automated searches of mass spectral libraries for each compound's mass spectrum. When a single good match is found, a standard can be purchased, and its mass spectrum and retention time on the column can be compared with those of the analyte to confirm its identity. But two problems with this strategy often occur: multiple library matches may provide several candidate compounds, (1) or worse yet, no matches may be found because mass spectra for most compounds are absent from the libraries.

In these cases, determining the numbers of atoms for all elements that compose the molecular ion (a molecule missing one electron) greatly reduces the number of compounds that could

**Andrew H. Grange** is a research chemist and **G. Wayne Sovocool** is a research chemist at Environmental Sciences Division, NERL, U.S. EPA, P.O. Box 93478, Las Vegas, NV 89193-3478. Grange may be contacted at (702) 798-2137, fax: (702) 798-2142, e-mail: grange.andrew@epa.gov. Sovocool may be contacted at (702) 798-2122, fax: (702) 798-2137, e-mail: sovocool.wayne@epa.gov.

**Figure 1.** Target diagram illustrating the dependence of the number of possible compounds on measurements of exact masses and relative abundances. Once the molecular ion composition is known, determination of the fragment ion compositions limits the number of possible isomers.





**Figure 2.** Mass peak profiles plotted from selected ion recording data acquired with 5000 resolving power for three ions with a nominal mass of  $m/z$  281. The three ion chromatograms display (top) a steady signal from a calibrant ion, (middle) a slowly increasing signal from a column bleed ion as the oven temperature increased, and (bottom) a chromatographic peak due to an eluting analyte ion.

produce the mass spectrum. Determination of the elemental compositions of the fragment ions produced by electron ionization further constrains the number of possible compounds, so that searches of the commercial and chemical literature for each compound become feasible. Converging toward a compound's identity based on this knowledge is illustrated by Figure 1. We will discuss ion composition elucidation (ICE) later in this article.

Quadrupole and ion-trap mass spectrometers are most commonly used for GC-MS. These instruments provide insufficient resolving power to separate mass peaks from different combinations of atoms that provide the same nominal (or integer) mass. For example, Figure 2 shows three mass peak profiles with the same nominal mass ( $m/z$  281) that were acquired with a double-focusing mass spectrometer using a resolving power of 5000 (10% valley definition) (2). (Assume two mass peak profiles with equal abundance partially overlap with a 10% valley [10% of the maxima] between them. The resolving power is the average mass of the two profiles divided by the mass difference between the maxima [ $\bar{m}/\Delta m$ ].) The profiles were plotted from selected ion recording data, as described later. They arose from the  $C_6F_{11}^+$  ion of the calibrant (perfluorokerosene), the  $C_7H_{21}O_4Si_4^+$  ion from column bleed,

and the  $C_{18}H_{19}NS^+$  ion produced from a compound in a sample extract. The measured exact masses of the three ions are provided above each profile. The calculated exact mass for the analyte ion was 281.1238 atomic mass units (amu), 2.4 millimass units (mmu) (8.5 ppm of the exact mass) less than the measured value, and 71 mmu (253 ppm) and 138 mmu (490 ppm) greater than the measured masses for the other two ions.

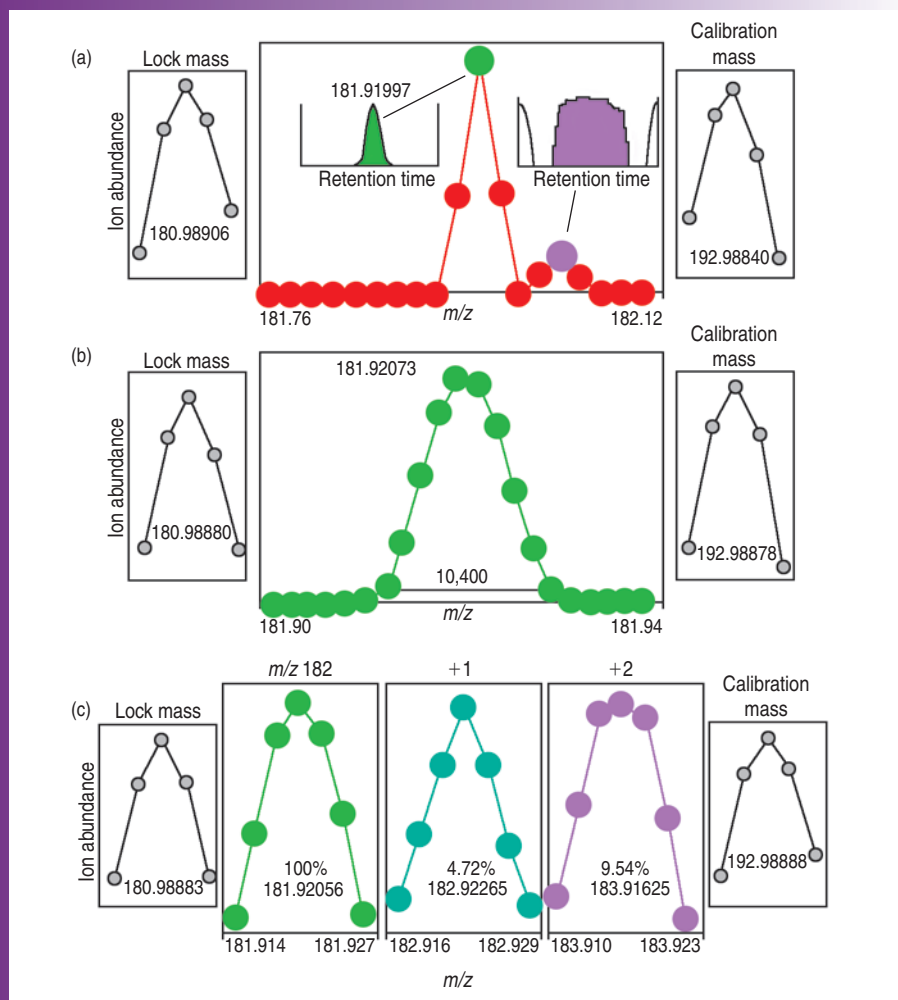
Based on this exact mass measurement, can we conclude  $C_{18}H_{19}NS^+$  is the analyte ion composition? Not yet, because many other combinations of atoms are possible. For an error limit of 3.4 mmu (12 ppm), the error limit for a single exact mass determination made with 5000 resolving power (3), 104 compositions are possible based on combinations of C, H, N, O, F, P, and S atoms. The measured exact mass of the molecular ion reduced the number of possible compositions, but fell far short of providing a unique composition. Additional MS measurements were required before this ion was unequivocally determined to be  $C_{18}H_{19}NS^+$ .

**Higher-mass isotopes of elements.** The presence of atoms of the higher-mass isotopes ( $^{13}C$ ,  $^2H$ ,  $^{15}N$ , and  $^{34}S$ ) provide mass peak profiles +1 and +2 amu heavier than the monoisotopic profile for which all atoms are the most abundant isotopes ( $^{12}C$ ,  $^1H$ ,  $^{14}N$ , and  $^{32}S$ ). Measured exact masses of the +1 and +2 profiles and

their abundances relative to the monoisotopic profile provide four additional values to distinguish among compositions. The number of possible compositions increases exponentially as an ion's mass increases (4), doubles when the error limit doubles (4), and increases each time a new element is added to those deemed possible. Comparison of calculated and measured values of three exact masses and two relative abundances, rather than of a single exact mass, increases fourfold the upper mass limit of ions for which unique ion compositions can be determined (3). As illustrated in Figure 1, knowing the composition of the molecular ion limits a compound's identity to a number of possible isomers. Determining the compositions of fragment ions then reduces this number.

**Instrumental requirements.** To measure all five quantities for analytes that are eluted from a gas chromatograph with a residence time of less than 10 s in the ion source of the mass spectrometer, an instrument must provide at least one scan each second, excellent mass accuracy (because the number of possible compositions doubles when the error limit is doubled), a linear dynamic range of at least  $10^3$  (because the relative abundance of the +2 profile can be 1% or less), and high resolving power to discriminate against mass interferences, if complex mixtures are analyzed. When a recently developed scanning procedure is used, double-focusing mass spectrometers provide the best combination of these instrumental requirements (4).

**Mass peak profiling from selected ion recording data.** Figure 3 shows full and partial mass peak profiles plotted by connecting discrete points. Each point is the area of a chromatographic peak observed in the ion chromatogram acquired at the  $m/z$  of the point. For example, the maximum point in Figure 3a is the area under the chromatographic peak to the left of the point. Another point corresponding to the smaller maximum is the area under the simulated chromatographic peak above it. The compound that produced this ion is ever-present. The baselines in the ion chromatogram were induced by reversing the polarity of a voltage in the ion source for 5 s both before and after the



**Figure 3.** (a) Two full mass peak profiles, each delineated by three points, for ions from an analyte and an ever-present compound. The resolving power was 3000 (10% valley definition) and the mass increment was 100 ppm. (b) The full mass peak profile for the analyte ion delineated by 10 points out of the 21 monitored. The resolving power was 10,000 and the mass increment was 10 ppm. (c) Partial profiles plotted from seven points for the  $m/z$  182 ion and its +1 and +2 profiles acquired using a resolving power of 10,000 and a 10 ppm mass increment. For all three experiments, the partial profiles for the lock mass and calibration mass were each monitored by five  $m/z$  values.

chromatographic peak of the analyte responsible for the larger mass peak profile eluted. Using a Finnigan MAT 900S (Thermo Electron, Waltham, MA) double-focusing mass spectrometer, each of 31  $m/z$  values was monitored for 20 ms during each 1-s selected ion recording cycle. Partial profiles for the lock mass and calibration mass, which always bracket the analyte mass, were each plotted from five  $m/z$  values. The remaining 21  $m/z$  values were used to monitor a mass range, a single profile, or three partial profiles using seven  $m/z$  values to delineate the top portions of each profile.

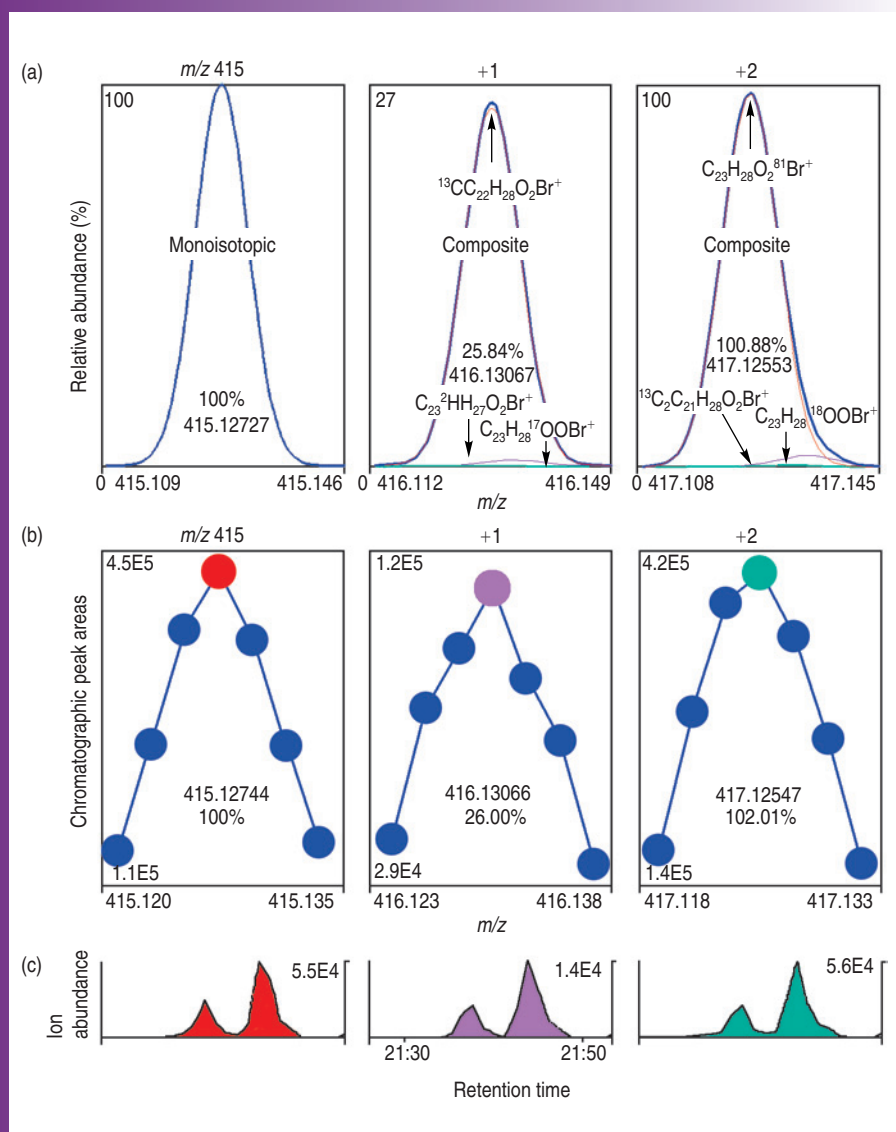
The three exact masses and two relative abundances were measured in the

last experiment of a set of three experiments that employs successive approximation (5). The data displayed in Figure 3a determined how many mass peak profiles were present across a relatively wide mass range (2000 ppm or 364 mmu in this example) centered about the mass observed for an analyte ion in a low-resolution mass spectrum. The data were acquired with a resolving power of 3000 and exact masses were monitored every 100 ppm (18.2 mmu). Two mass peak profiles, each delineated by three points, revealed the presence of an analyte ion (chromatographic peak) and an ever-present calibrant (steady signal), perfluorokerosene. The estimated exact mass of

the analyte profile, calculated as the weighted average of its three points, was used as the center mass in the next experiment. For Figure 3b, exact masses were monitored every 10 ppm (1.82 mmu) to provide a well-defined profile delineated by 10 points. The resolving power was 10,400 as measured from the 5% of maximum levels of the full profile. The exact mass, accurate to within 6 ppm, was then entered into a profile generation model (3), which provided a list of compositions consistent with the measured exact mass within its error limits. The user then chose a hypothetical composition from the list. The center masses of the hypothetical ion and its +1 and +2 partial profiles were then used as the center masses of the three partial profiles monitored in Figure 3c, again with 10 ppm mass increments and a resolving power of 10,000. The exact masses were calculated as the weighted average of the points from each profile and the relative abundances as the ratio of the summed points for the +1 and +2 partial profiles divided by the summed points for the  $m/z$  182 profile multiplied by 100%. The ion investigated was determined unequivocally to be  $C_3H_7AsS_2^+$ . The measured values in Figure 3 were consistent with the calculated values for the three partial profiles: 181.92051, 182.92272, 183.91636, 4.71%, and 8.89%.

Figure 4 shows calculated and measured mass peak profiles for a  $C_{23}H_{28}O_2Br^+$  ion. The theoretical exact masses and relative abundances of the composite +1 and +2 profiles are listed under the profiles in Figure 4a. In Figure 4b, these values were obtained from top portions of the mass peak profiles, which were plotted from areas under selected ion recording chromatograms acquired as the two isomers evident in Figure 4c were eluted into the mass spectrometer. The summed chromatographic peak areas in each panel of Figure 4c provided the maxima of the partial profiles.

**Mass spectral interpretation.** Determination of the compositions of the molecular ion (if observed) and the fragment ions in a mass spectrum using ICE greatly reduces the number of possible compounds that could produce the spectrum. Even then, an expert in mass spectral interpretation is required



**Figure 4.** Plots showing (a) calculated profiles for the  $C_{23}H_{28}O_2Br^+$  ion, the three +1 ions and three most abundant +2 ions, and the composite +1 and +2 profiles; (b) partial  $m/z$  415, +1, and +2 profiles plotted from chromatographic peak areas under ion chromatograms for seven  $m/z$  values across each profile; and (c) ion chromatograms for the  $m/z$  values at the maxima of the partial profiles.

to hypothesize possible isomers based on common fragmentations from ions formed from numerous classes of compounds and to search the chemical and commercial literature for candidates.

**Example 1: a well pollutant.** Table 1 is a partial output from the profile generation model based on the experimental values from Figure 4b. The seven compositions listed are possible based on the measured exact mass of the monoisotopic ion within the error limits of its measurement (3 ppm for one determination at 20,000 resolving power). The mass defects are the non-integer part of the exact masses of the monoisotopic ion and the +1 and +2 profiles. Only the last composition provides calculated exact masses and relative abundances consistent with the five measured values.

In Figure 5, the  $m/z$  415 ion was observed in the raw and background-subtracted mass spectra. Because  $^{79}Br$  and  $^{81}Br$  have nearly equal isotopic abundances, the two mass peaks at  $m/z$  415 and 417 visible above the chemical noise suggested a monobrominated compound might be present in an extract of chlorine-disinfected well water. The  $m/z$  430 and 432 ions became apparent above the chemical noise in the background-subtracted mass spectrum. ICE was also used to determine that the  $m/z$  430 ion was  $C_{24}H_{31}O_2Br^+$ . No compounds with this composition were found in the NIST or Wiley mass spectral libraries.

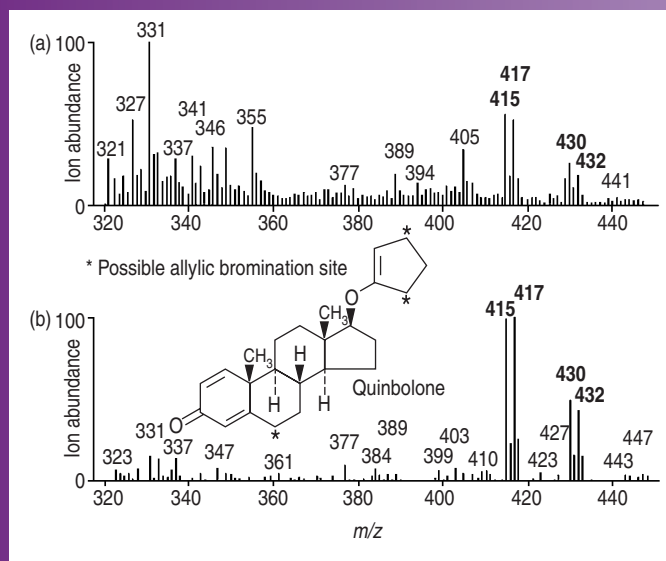
With a single composition to consider for the apparent molecular ion, chemical reasoning and searches of the chemical and commercial literature can lead to

**Table I. Seven of 26 possible compositions for a fragment ion.**

Composition	$m/z = 415.12744 \pm 3 \text{ ppm (1.25 mDa)}$			Relative abundances		
	Mass defects	0	+1	+2	%+1 (%+1 range)	%+2 (%+2 range)
$C_{16}H_{28}N_6SBr$	.12795	.13070	.12591	.12591	20.20 (16.65–23.85) X	103.20 (88.77–117.94)
$C_{17}H_{29}N_4OPBr$	.12623	.12931 X	.12430	.12430	19.65 (16.18–23.14) X	96.96 (83.22–110.57)
$C_{18}H_{31}F_3SBr$	.12819	.13148	.12616	.12616	20.90 (17.97–23.85) X	102.42 (90.13–114.90)
$C_{18}H_{27}N_3O_2FBr$	.12707	.13023	.12515	.12515	20.90 (17.64–24.20) X	98.37 (86.63–110.29)
$C_{19}H_{31}NO_2PBr$	.12758	.13092	.12567	.12567	21.64 (18.61–24.70) X	98.43 (86.59–110.46)
$C_{20}H_{29}O_3FBr$	.12841	.13182	.12652	.12652	22.55 (19.41–25.69) X	98.86 (86.79–111.17)
$C_{23}H_{28}O_2Br$	.12727	.13067	.12541	.12541	25.80 (22.22–29.39)	98.57 (86.49–110.93)
Exp'l values	.12744	.13066	.12547	.12547	26.00	102.01

An X indicates inconsistency between the measured and calculated values.





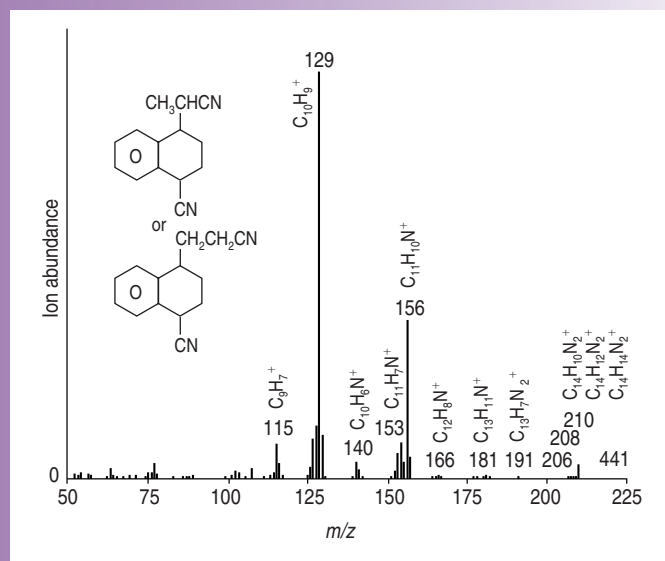
**Figure 5.** (a) Raw mass spectrum of a compound found in a well and (b) its background subtracted mass spectrum.

compound identification. Chlorination of the well water, which contained bromide ions, could brominate organic compounds. The structure of quinbolone, an anabolic steroid, is shown in Figure 5b and has three possible allylic bromination sites — more than enough to account for the two isomers observed in the ion chromatograms in Figure 4c. Substitution of a Br atom for an H atom would provide the observed composition. A feedlot was located near the well and anabolic steroids are often used to stimulate livestock growth. Purchase of quinbolone, its chlorination in the presence of bromide ions, and examination of the mass spectra and retention times of the products would be logical next steps in the identification process for this compound.

#### Example 2: pollutants in Toms River, NJ.

When chemicals are found in a drinking water supply servicing a locality suffering an elevated occurrence of illness, it is important to identify those chemicals to facilitate toxicological studies and to ascertain the source of the chemicals. Such was the case for water from a municipal well near Toms River, NJ, where an increased incidence of childhood cancer was observed. Several isomers in a well-water extract provided similar mass spectra, which were not found in the NIST or Wiley mass spectral libraries. One example is shown in Figure 6.

The compositions of the molecular ion and 10 fragment ions listed in Table II were determined using ICE. The fragment ion chromatograms tracked those



**Figure 6.** A background-subtracted mass spectrum for the most abundant of five isomers found in a municipal well water extract. Compositions were determined for 11 ions.

of the molecular ion and provided the same retention times. The exact mass alone of each fragment ion provided the correct composition, because the strict elemental limits provided by the molecular ion (0–14 C, 0–14 H, and 0–2 N atoms) limited the number of possible compositions for each fragment ion to no more than three. Agreement to within 0.5 mmu between the calculated and measured exact masses identified the correct compositions shown in red in Table II.

Consideration of these fragment ions and the corresponding composite neutral losses by an expert in mass spectral interpretation indicated the compound contained a tetralin core with external  $C_3H_4N$  and CN groups as depicted by

**Table II. Exact masses and possible compositions for 10 fragment ions.**

Experimental mass	0 N	Error (mmu)	1 N	Error (mmu)	2 N	Error (mmu)
208.1001					$C_{14}H_{12}N_2$	+0.1
206.0846					$C_{14}H_{10}N_2$	+0.2
191.0607			$C_{14}H_9N$	-12.8	$C_{13}H_7N_2$	-0.2
181.0887	$C_{14}H_{13}$	-12.6	$C_{13}H_{11}N$	-0.4	$C_{12}H_9N_2$	+12.5
166.0661	$C_{13}H_{10}$	-12.2	$C_{12}H_8N$	+0.4	$C_{11}H_6N_2$	+13.0
156.0816	$C_{12}H_{12}$	-12.3	$C_{11}H_{10}N$	+0.3	$C_{10}H_8N_2$	+12.9
153.0578	$C_{12}H_9$	-12.6	$C_{11}H_7N$	0.0	$C_{10}H_5N_2$	+12.5
140.0501	$C_{11}H_8$	-12.5	$C_{10}H_6N$	+0.1	$C_9H_4N_2$	+12.7
129.0707	$C_{10}H_9$	+0.3	$C_9H_7N$	+12.9	$C_8H_5N_2$	+25.4
115.0548	$C_9H_7$	0.0	$C_8H_5N$	+12.6	$C_7H_3N_2$	+25.2

the structures in Figure 6. Where these external groups were attached to the rings could not be determined from the table. Even so, only a small number of possible isomers remained. A search of the chemical literature soon located 1:2 styrene:acrylonitrile by-products of an industrial polymerization process as the candidate compounds. A standard from an industrial process currently used was

obtained. Three isomers were observed in the standard with the same retention times and mass spectra as three of the five isomers observed in the well water. These confirmations established the industrial source of the pollutants.

**Example 3: Superfund site samples.**

Chemical by-products of production dumped into Superfund sites can yield very complex mixtures of compounds.

For one such site, ion compositions were determined for 51 compounds, most of which contained NS, N<sub>2</sub>S, or N<sub>2</sub>S<sub>2</sub>, including the analyte from Figure 2, C<sub>18</sub>H<sub>19</sub>NS (2). Most of the summed chromatographic peak area was due to alkyl benzothiazoles as indicated by the presence of the characteristic C<sub>8</sub>H<sub>7</sub>NS<sup>+</sup> ion in many of the mass spectra. These compounds are used in the dye and rubber industries. Characterization of these compounds provided confirmation that the dumped material was from a nearby dye plant.

**Needed research.** These examples demonstrated that ICE is a powerful analytical tool when applied to analytes introduced into the mass spectrometer in the gas phase. But most compounds are too polar, ionic, thermolabile, or too high in mass to traverse a GC column or to volatilize from a heated probe. Such compounds require liquid sample introduction using electrospray ionization (ESI) or atmospheric pressure chemical ionization (APCI). These low-energy ionization techniques yield adduct ions (that is, the molecular ion plus Na or K) and few fragment ions. Additional factors that make application of ICE to such compounds difficult are poorer chromatographic resolution, much lower sensitivity than electron ionization, competition between calibrants and analytes for ionization which depletes the analyte signal, and the lack of commercial ESI or APCI mass spectral libraries. Despite these limitations, we will soon investigate the use of ICE to characterize and identify compounds separated by high performance liquid chromatography.

**ICE plus containment.** Although one laboratory in Germany (6) has used a modified form of mass peak profiling from selected ion recording data to identify a naturally occurring, bioaccumulative, organohalogen compound, only our lab now performs ICE as described herein. This lab is not equipped to work with unanticipated compounds that could be extremely toxic. The Environmental Chemistry Branch of the U.S. EPA's Environmental Chemistry Division is ready and willing to transfer ICE technology to containment labs

within secure facilities. If necessary, we will adapt the ICE code for the data systems of other models of double focusing mass spectrometers. Labs expected to identify potentially toxic compounds intentionally added to water, air, or food supplies would then have a power-

ful new analytical tool for doing so. In addition, we hope to transfer this technology to labs with double focusing mass spectrometers that desire to investigate the mystery compounds observed in total ion chromatograms for more common sample extracts.

**Conclusion.** Instances occur where a single chemical or a complex mixture of chemicals not found in mass spectral libraries is added intentionally or unintentionally to water, air, or food. These compounds must be identified before their toxicological properties can be investigated. Determination of the elemental compositions of the ions in mass spectra sufficiently limits the number of possible isomers for each compound to make searches of commercial and chemical literature practical. If a compound were not found in the literature, isolation and concentration of the compound would be necessary to perform other analyses, such as NMR, FT-IR, or x-ray crystallography to identify each compound.

## Acknowledgments

The U.S. Environmental Protection Agency (EPA), through its Office of Research and Development, funded and performed the research described. This manuscript has been subjected to the EPA's peer and administrative review and has been approved for publication.

The authors are grateful to Floyd A. Genicola for providing the As-containing compound for which the data used to construct Figure 1 were obtained. These data helped confirm his identification of the compound as 2-methyl-1,3,2-dithiarsolane.

## References

1. A.H. Grange, L. Osemwengie, G. Brilis, and G.W. Sovocool, *Internat. J. of Environ. Forensics* **2**, 61–74 (2001).
2. A.H. Grange and W.C. Brumley, *LC-GC* **14**(11), 978–986 (1996).
3. A.H. Grange and W.C. Brumley, *J. Amer. Soc. for Mass Spectrom.* **8**, 170–182 (1997).
4. A.H. Grange, F.A. Genicola, and G.W. Sovocool, *Rapid Commun. Mass Spectrom.* **16**, 2356–2369 (2002).
5. A.H. Grange, J.R. Donnelly, and W.C. Brumley, *Anal. Chem.* **68**, 553–560 (1996).
6. W. Vetter, L. Alder, and R. Palavinskas, *Rapid Commun. Mass Spectrom.* **13**, 2118–2124 (1999). ■