APPLICATION NOTE

Deconvolution of Isobaric Interferences in Mass Spectra

Juris Meija and Joseph A. Caruso

Department of Chemistry, University of Cincinnati, Cincinnati, Ohio, USA

The concept of isobar deconvolution using the mass domain and signal intensity based domains is described. The intensity domain-based approach employs the reconstruction of the observed isotope pattern from the isolated patterns of the isobaric species. The quantitative information is adjusted with the use of the least squares algorithm. The mass domain-based approach employs signal deconvolution by forming Gaussian components for which the peak width and position can be predicted a priori. The latter method is applicable to medium resolution instruments, such as TOF-MS, while the pattern reconstruction approach is applicable also to low resolution instrumentation, such as quadrupole-based ICP-MS or GC/MS. An example of CH_zSe ($z = 0-4$) cluster analysis in dimethyl diselenide mass spectra is given to illustrate the concepts underlying both approaches. (J Am Soc Mass Spectrom 2004, 15, 654–658) © 2004 American Society for Mass Spectrometry

Very recently Roussis and Proulx clearly demonstrated the need of mathematical data treatment when interpreting mass spectra [1]. The presence of isobaric interferences is a common issue in mass spectrometry. In plasma mass spectrometry (ICP-MS), this may lead to false identification of the elements and consequently, may affect quantitation of a particular element. Instrumental approaches, through either high resolution mass spectrometers or collision/reaction cell quadrupole instruments, are often used to minimize or eliminate polyatomic interferences [2]. The general mathematical strategy to correct isobaric interferences in atomic mass spectrometry practice is the subtraction of the contribution of interfering isotopes from measurements of non-interfering isotopes. For example, the $^{40}Ar^{35}Cl^+$ contribution on the $m/z = 75$ signal (determination of $^{75}As^+$) can be estimated by monitoring $^{40}Ar^{37}Cl^+$ at $m/z = 77$. From the signal at $m/z = 77$, one can then estimate the abundance of $^{40}Ar^{35}Cl^+$ by knowledge of the natural chlorine isotope ratio. The idea behind this approach is simple; however, it is applicable only for a single polyatomic ion interference.

In molecular mass spectrometry, isobaric interferences lead to a shift in the observed peak mass centroid. In isotope dilution GC/MS, for example, the presence of isobaric interferences affects the apparent isotope ratios and thus they have to be accounted for to achieve accurate results [3]. It is therefore useful to understand the main aspects of isobar deconvolution as a tool to enhance the validity of the data obtained.

In this study we utilize the several approaches to isobaric interference estimation, as recently outlined by Roussis and Proulx [1]. In particular, these approaches were demonstrated to be useful in the interpretation of diselenide and selenosulfenate mass spectra [4] and here we give a more technical report of the various isobar deconvolution approaches. These are isotope cluster reconstruction using isotope pattern superimposition combined with the least square optimization algorithm, signal deconvolution to its components in the mass domain, and mass shift analysis of the overlapping clusters. Although these approaches are known (particularly to those in chemometrics), the majority of mass spectra are interpreted solely based on the single mass measurements and presence of isobars is usually detected qualitatively. In this communication we aim to emphasize the benefits and simplicity of quantitative isobar detection.

Results and Discussion

Isotope Pattern Reconstruction

The presence of isobaric interferences can be easily detected by the distortion of the isotope pattern. One can use the advantage of an isotopic pattern to estimate the amount of overlapping species. Reconstruction of the observed isotope pattern is achieved by the superimposition of the separate isotope profiles of the overlapping species in conjunction with least square optimization of the pattern intensities.

Assuming the mass spectral pattern of interest (Y) can have a contribution (η) from n components, then according to the superposition principle [1], signal

abundance (Y_j) at a certain mass (j) is a linear combination of the isotope patterns of the isobaric components (I):

$$Y_j \propto \sum_{i=1}^{n} \eta_i \cdot I_{ij} \qquad (1)$$

Isotope pattern deconvolution involves selection of contributing species and sequentially finding the relative contributions (η) for each of the components. Isotopic profiles of the individual component can be calculated by the classical polynomial expansion approach [5] or by Fourier transformation methods [6].

Isotope pattern abundances for each of the n individual components can be represented as an $m \times n$ matrix, F (usually non-symmetric), where m is the size of the pattern. I_{ij} represents the abundance of mass j of the i-th component (each of the i-th components in the F matrix has the same mass):

$$F = \begin{pmatrix} I_{11} & I_{21} & \ldots & I_{n1} \\ I_{12} & I_{21} & \ldots & I_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ I_{1m} & I_{2m} & \ldots & I_{nm} \end{pmatrix} \qquad (2)$$

If the instrumentation produces distortion of the theoretical pattern (e.g., due to the differences in ion transmission), appropriate corrections need to be implemented in the abundances I_{ij}. The matrix of the experimental abundances for the cluster of interest is an $m \times 1$ vector Y and solution of the least squares is an $n \times 1$ vector a, from which the relative contributions of the individual components (η_i) then can be calculated:

$$a = (F^T F)^{-1} F^T Y \qquad (3)$$

$$\eta_i = \frac{a_i}{\sum_i a_i} \qquad (4)$$

Error estimate of η_i can be obtained from the diagonal elements of the inverse information matrix $(F^T F)^{-1}$ via the error estimate of a_i:

$$s_{a_i}^2 = s_r^2 \cdot (F^T F)_{ii}^{-1} \qquad (5)$$

Here s_r is the residual standard deviation, obtained from the residual matrix $(Y - Fa)$.

$$\frac{s_{\eta_i}}{\eta_i} = \sqrt{\frac{1}{m-1}(Y - Fa)^T(Y - Fa)}$$

$$\cdot \sqrt{\left(\frac{1}{a_i}(F^T F)_{ii}^{-1}\right)^2 + \sum_{k=1}^{n}\left(\frac{1}{a_k}(F^T F)_{kk}^{-1}\right)^2} \qquad (6)$$

One of the characteristics of this approach is that the relative uncertainty of η_i is very similar across the values of $i = 1–n$. This approach gives better η_i estimate for elements with rich isotopic patterns (such as Te, Sn, Cd, or Hg where each of the elements has more than 6 major isotopes). Eventually, higher numbers of included fitting components need to be justified with the amount of input information (via the size of the isotope pattern). In the case of selenium, which has five major isotopes, five species deconvolution of the isotopic pattern is readily accessible with a precision better than 1%.

Signal Deconvolution in the Mass Domain

Signal peak shape analysis. Because of the different mass defects of the elements, isobaric interferences will differ slightly in their mass. This discrepancy will result in a broadening of the signal in the mass domain. For time-of-flight detection, the observed signal peak width is linearly proportional to the mass and the mass-peak width relationship can be easily obtained using a calibration standard, such as *heptacosa*fluoro-tributylamine. Broadening is more apparent at the base of the peaks, so peak width at 20–30% height (instead of 50%) can be efficiently used for broadening characterization, however higher precision is attained at the 50% level. Accordingly, the reference peak width is compared with the signals of interest and any significant signal width difference indicates the presence of isobaric interferences.

After the presence of isobaric interference is established, the magnitude can be estimated by signal deconvolution using the Gaussian function (or any other more appropriate signal peak function, such as Lorentzian). Under normal operating conditions in time-of-flight mass spectrometry, ions will be normally distributed with respect to their arrival time (represented as mass) and the peak width of two isobar signals can be safely assumed to be the same.

Note that this procedure is not merely an empirical fitting. All the parameters of the Gaussian components have a relationship to the physical properties such as the resolution (peak width), exact mass of the component (peak position) and relative amount of the interfering species (normalized peak area).

Peak centroid mass analysis. Peak centroid mass, which is often used to represent mass spectrometric data, is affected in the presence of isobaric interferences. To illustrate this, assume there are two interfering species A and B with the relative abundance of x and $(1 - x)$ respectively. The analytical signal as a function of mass can be represented with the following expression:

$$I = I(A)_i + I(B)_i \propto \frac{x}{\Delta m} e^{-2((m_i - m_A)/\Delta m)^2}$$

$$+ \frac{1-x}{\Delta m} e^{-2((m_i - m_B)/\Delta m)^2} \qquad (7)$$

The maximum of the observed resulting signal profile (m_{max}) depends on the relative contributions of isobars A and B. Such relationships can be computed from eq 7. For example, when the peak width is larger than the mass difference of the isobars, $\Delta m \geq |m_A - m_B|$, it can be shown that the observed peak centroid mass is the weighted average of the isobar masses:

$$m_{max} \cong x_A \cdot m_A + x_B \cdot m_B \qquad (8)$$

According to this, in the presence of isobaric interferences, the accurate mass measurements will be biased if an uncorrected peak centroid mass is used. Moreover, care should be taken on the actual centroid algorithm itself. If the average mass is used (or 80% top-of-the-peak average etc.) it no longer exactly represents the maximum mass, m_{max}, and additional mass bias thus can be introduced.

Application to Dimethyl Diselenide Mass Spectra

It is known that the electron impact fragmentation of dimethyl diselenide does not render a pure cluster of CH_3Se as the region at $m/z = 88$–98 u clearly shows a distorted isotope pattern of selenium [4] (see also Figure 1). The following example illustrates the deconvolution of this isobaric overlap.

Deconvolution in the intensity domain. The isotope pattern reconstruction algorithm described above was applied to the CH_zSe^+ cluster of dimethyl diselenide EI+ mass spectrum. The contribution of the species with $z = 0$–4 was allowed in the pattern reconstruction and the F matrix (taking into account ^{13}C contribution as well) therefore is constructed as follows:

	CH_0Se	CH_1Se	CH_2Se	CH_3Se	CH_4Se
$m/z = 88$	8.9	–	–	–	–
89	7.6	8.9	–	–	–
90	23.4	7.6	8.9	–	–
91	–	23.4	7.6	8.9	–
92	49.2	–	23.4	7.6	8.9
93	–	49.2	–	23.4	7.6
94	9.1	–	49.2	–	23.4
95	–	9.1	–	49.2	–
96	–	–	9.1	–	49.2
97	–	–	–	9.1	–
98	–	–	–	–	9.1

$F =$ (matrix as above)

From here the least squares optimized isobar abundance contribution matrix (η_i) was calculated according to the eqs 3–6. The results were as follows (with the standard deviation given in the parenthesis): $\eta_0 = 2.0(0.8)\%$ (CH_0Se^+); $\eta_1 = 49.8(0.6)\%$ (CH_1Se^+); $\eta_2 = 19.7(0.6)\%$ ($CH_2Se^{\cdot+}$); $\eta_3 = 22.7(0.6)\%$ (CH_3Se^+) and η_4
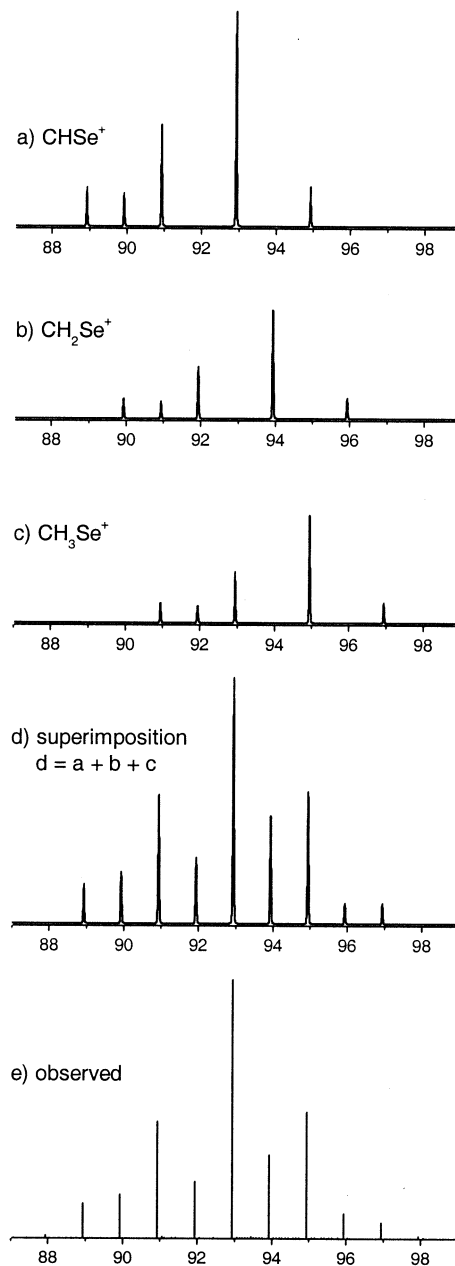

Figure 1. Calculated isotope patterns of the $CHSe^+$ (**a**), CH_2Se^+ (**b**), and CH_3Se^+ (**c**) clusters, and the result of the 2:1:1 superimposition of these species (**d**) along with the experimentally observed CH_zSe^+ cluster in EI+ mass spectra of dimethyl diselenide (**e**).

$= 5.9(0.6)\%$ ($CH_4Se^{\cdot+}$), which is consistent with the 2:1:1 contribution from $CHSe^+$, $CH_2Se^{\cdot+}$ and CH_3Se^+, respectively, and the result obtained is in good agreement with the experimentally observed CH_zSe isotope pattern (Figure 1).

Deconvolution in the mass domain. A single signal from the dimethyl diselenide mass spectrum at $m/z = 92.924$ u was selected for subsequent analysis. The results of the isotope pattern reconstruction procedure suggest that there are two contributions at this m/z value,
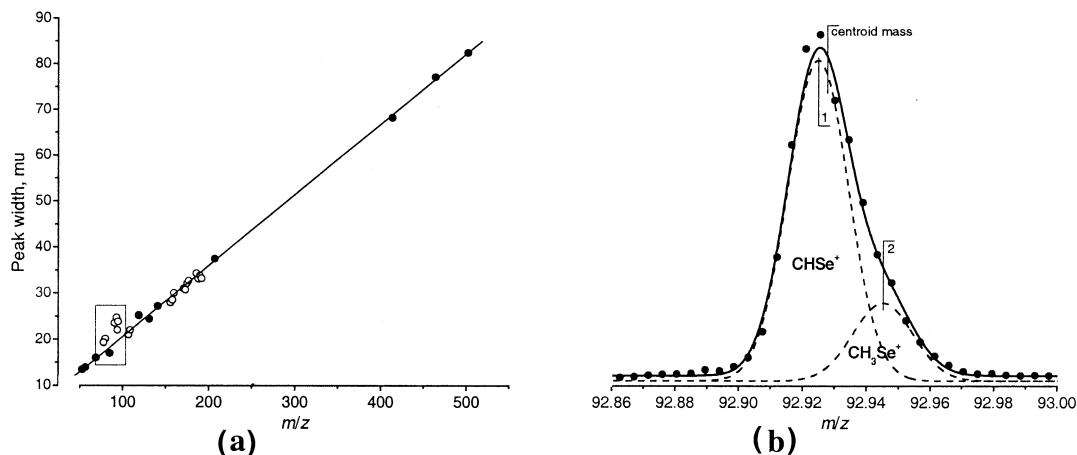
Figure 2. (**a**) Peak width (at half-height) and mass relationship in time-of-flight mass analyzer as obtained from the isobar-free signals in the mass spectrum of *heptacosa*fluoro-tributylamine (filled circle). Squared region shows the isobar interference related peak broadening in mass spectra of dimethyl diselenide (open circle) around $m/z = 93$. (**b**) High resolution mass spectra deconvolution. $m/z = 92.924$ u fragment of the dimethyl diselenide EI+ mass spectra (filled circle) and the reconstructed (fitted) signal peak profile (dash) showing the contribution of $CH^{80}Se$ and $CH_3^{78}Se$ (in the ratio of 4:1). Note that 80% top-of-the-peak centroid mass does not correspond to the mass of either of the components, $CH^{80}Se$ (1) or $CH_3^{78}Se$ (2), nor to the maximum signal.

namely $CHSe^+$ and CH_3Se^+ (refer to the *F* matrix). Accordingly, the observed signal peak at $m/z = 92.924$ u is deconvoluted (fitted) with two Gaussians of equal width:

$$I(m) = \frac{A_1}{\Delta m} e^{-2((m-m_A)/\Delta m)^2} + \frac{A_2}{\Delta m} e^{-2((m-m_B)/\Delta m)^2}$$

The numerical variables in this fitting procedure are peak width (Δm), peak masses (m_A and m_B) and peak abundances (A_1 and A_2). The fitting procedure can be performed using a nonlinear least squares fitting algorithm with simplex and Levenberg-Marquardt minimization algorithms and statistical data weighting ($w = I^{-1}$) for improved peak valley fitting. These methods are implemented using data analysis software, such as Origin.

The fitted Gaussian parameters then are compared with their theoretically (peak position) and experimentally (peak width) predicted values. The single species peak width was estimated from the linear regression of mass *vs.* peak width of the *heptacosa*fluoro-tributylamine mass spectrum (Figure 2a). Gaussian profile fitting results are shown in Figure 2b.

Results of the Gaussian signal fit agree well with the predicted values given in the parentheses: $CH^{80}Se^+$ and $CH_3^{78}Se^+$ peak width 18.0 ± 0.4 mu (19.5 ± 2 mu is the experimentally estimated peak width at $m/z = 93$ u from the data shown in Figure 2a), peak area ratio 4.1 ± 0.5 (a value of 4.21 corresponds to the $CHSe/CH_3Se$ ratio of 2:1), and $CH^{80}Se^+$ and $CH_3^{78}Se^+$ peak separation 20 mu (theoretical value 17 mu). The obtained results clearly show the presence of $CH^{80}Se^+$ and $CH_3^{78}Se^+$ in the ratio of about 4:1. This corresponds to

the $CHSe/CH_3Se$ ratio of 2:1 (referring to the same Se isotopes in both species) which is in agreement with results obtained from the isotope pattern reconstruction approach.

Peak centroid mass shift analysis. In the above example of dimethyl diselenide, the difference between the mass of $CH^{80}Se$ (92.9243 u) and $CH_3^{78}Se$ (92.9408 u) is 17 mu and peak width (Δm) at this mass is estimated to be 20 mu. The observed centroid mass is 92.9273 ± 0.0005 u. Since eq 8 holds true in this case ($\Delta m \geq |m_{CH^{80}Se} - m_{CH_3^{78}Se}|$), the observed centroid mass is in a simple relationship with the isobar abundances:

$$92.9273 \text{ u} \cong x_{CH^{80}Se} \cdot 92.9243 \text{ u} + x_{CH_3Se} \cdot 92.9408 \text{ u}.$$

From here $CH^{80}Se/CH_3^{78}Se = 4.5 \pm 0.8$. Centroid mass shift is an important isobaric interference consequence, as peak centering is common practice in mass spectrometry data analysis.

In conclusion, the $CH^{80}Se/CH_3^{78}Se$ amount ratio in $m/z = 92.924$ signal of dimethyl diselenide EI+ mass spectra obtained by centroid mass shift analysis (4.5 ± 0.8) is comparable to that obtained by isotope pattern reconstruction (4.6 ± 0.4) and peak shape deconvolution approaches (4.1 ± 0.5) thus giving credence to the validity of all methods.

Acknowledgments

References

1. Roussis, S. G.; Proulx, R. Reduction of Chemical Formulas from the Isotopic Peak Distributions of High-Resolution Mass Spectra. *Anal. Chem.* **2003,** *75,* 1470–1482.
2. Tanner, S. D.; Baranov, V. I.; Bandura, D. R. Reaction Cells and Collision Cells for ICP-MS: A Tutorial Review. *Spectrochim. Acta B* **2002,** *57,* 1361–1452.
3. Centineo, G.; Rodriguez-Gonzalez, P.; Blanco-Gonzalez, E.; Garcia-Alonso, I. J.; Sanz-Medel, A. Simultaneous Determination of Mono-, Di- and Tributyltin in Environmental Samples Using Isotope Dilution Gas Chromatography Mass Spectrometry. *J. Mass Spectrom.,* in press.
4. Meija, J.; Beck, T. L.; Caruso, J. A. Interpretation of Alkyldiselenide and Selenosulfenate Mass Spectra, unpublished.
5. Yergey, J. A. A General Approach to Calculating Isotopic Distributions for Mass Spectrometry. *Int. J. Mass Spectrom. Ion Phys.* **1983,** *52,* 337–349.
6. Rockwood, A. L.; Van Orden, S. L.; Smith, R. D. Rapid Calculation of Isotope Distributions. *Anal. Chem.* **1995,** *67,* 2699–2704.