Evaluation of minimal ^{13}C-labelling for stable isotope dilution in organic analysis

Ana González-Antuña,[a] Pablo Rodríguez-González,[a] Giuseppe Centineo[b] and J. Ignacio García Alonso*[a]

Received 23rd November 2009, Accepted 18th February 2010
First published as an Advance Article on the web 11th March 2010
DOI: 10.1039/b924432h

A procedure for Stable Isotope Dilution Analysis in molecular Mass Spectrometry which does not require a methodological calibration graph and can be applied in combination with minimal labelling has been evaluated. This alternative approach is based on the determination of the molar fractions for each pure isotope pattern (natural abundance or labelled) contributing to the isotope pattern observed in the mixture of natural abundance and labelled molecules by multiple linear regression. Two labelled compounds, $^{13}C_1$-labelled or $^{13}C_6$-labelled phenol, were compared to study the influence of the number of ^{13}C atoms in the labelled molecule. The procedure was evaluated by comparing the results obtained for the determination of phenol in NIST 1584 CRM by GC-EI-MS using the classical isotope dilution calibration procedure and the new procedure based on multiple linear regression of isotope patterns without a calibration graph. The results obtained using the proposed procedure agreed well with both the certified values and those obtained using the classical Isotope Dilution Mass Spectrometry (IDMS) calibration procedures. For the evaluated procedure, a full uncertainty budget determination has been developed taking into account all uncertainty sources, including those derived from the uncertainties in the isotope patterns of the natural and labelled compounds. The measurements with the $^{13}C_1$-labelled phenol provided lower propagated uncertainties in comparison to the use of $^{13}C_6$-labelled phenol.

Introduction

Isotope Dilution Mass Spectrometry (IDMS) currently stands out as the analytical technique for trace analysis providing results with the highest metrological quality in comparison to more traditional measurement methods such as external calibration or standard additions. IDMS requires the addition to the sample of an isotopically labelled form of the element or compound to be determined and the measurement of the 'ratio' of labelled to unlabelled analyte by mass spectrometry. Based on this principle, IDMS results are not affected by analyte losses, low recoveries or matrix effects and therefore, IDMS has been considered as a primary method directly traceable to the International System of Units for the quantity 'amount of substance' or mole. Also, it is currently the technique of choice in many intercomparison exercises or certifications of reference materials by National Metrology Institutes[1] and it has been widely applied in different analytical fields: elemental[2] and organic analysis,[3] elemental speciation[4] and bioanalytical chemistry[5] (particularly in the field of quantitative proteomics[6]).

The theory and practice of stable isotope dilution mass spectrometry in both organic and elemental analysis has evolved during the last 50 years following a somewhat diverging trajectory. Recent attempts to apply the concepts of inorganic IDMS to organic molecules have been based on the derivation of equations in which the efficiency of all the chemical steps prior to instrument analysis needs to be included.[7] In elemental analysis no isotopic effects are expected and a simple equation is used to relate the measured isotope ratio in the mixture with the unknown concentration of the analyte. In such cases there is no need to resort to any methodological calibration of the response in the mass spectrometer with the concentration of the analyte.[2,4] However, organic isotope dilution analysis traditionally uses a calibration graph that is prepared using the labelled analyte as internal standard.[3] This calibration graph is linear when there is no mass overlap between the labelled and unlabelled analyte and its slope is normally determined experimentally as it depends on the isotopic composition of both the analyte and its labelled analogue as well as on possible isotopic effects during sample preparation. Therefore, the labelled compound in organic isotope dilution is usually selected to provide no mass overlap with the unlabelled analogue. For this purpose, a significant mass difference of at least 3 mass units is normally recommended for molecules containing between 10 and 20 carbon atoms.[3] However, when multiple labelling of a molecule is performed, differences in its physicochemical properties in comparison with the natural abundance compound have been observed.[8–11] This is particularly true with deuterated compounds, because of slightly different hydrogen bonding capabilities, resulting in slightly different extraction and derivatisation yields from those observed for the natural abundance compounds.[8] Also, undesired changes in retention times both in GC[9] and LC[10] separations have been reported. Most of these isotopic effects are small and can be compensated for when the calibration graph is constructed following the analytical procedure used for the samples.

[a]*Department of Physical and Analytical Chemistry, Faculty of Chemistry, University of Oviedo, Spain. E-mail: jiga@uniovi.es*
[b]*Innovative Solutions in Chemistry S.L., Edificio Científico-Tecnológico Campus de "El Cristo", Oviedo, Spain*

Nevertheless, other effects like the matrix suppression in LC-ESI-MS are not properly corrected by the use of isotopically labelled analytes because of their differences in retention times.[10,11] In these situations, a minimal labelling (*e.g.* a single ^{13}C label in the molecule) might be useful as no isotopic effects or changes in retention time are expected. The problem with this choice is that, due to the spectral overlap, the minimal labelling provides non-linear isotope dilution calibration graphs and this has been traditionally avoided in organic IDMS. Previous works have been devoted to the linearization of the calibration curves by applying complex mathematical linearizing methods.[12,13] Also, Thienpont *et al.*[14] proposed the use of labelled compounds with a mass increment of only 1–2 units using single point-calibration methods. For this purpose, the contribution of the unlabelled analyte at the *m/z* value monitored for the labelled analogue (and *vice versa*) must be determined previously.

The concept of isotope pattern deconvolution (IPD) was first described in the textbook by Lambert *et al.*[15] to calculate the extent of isotope incorporation in molecules and to check for cluster purity and the presence of overlapping $M - H^+$ ions. This procedure was initially based on a stepwise manual calculation and was improved later by the application of direct calculation procedures based on multiple least squares. In the last few years IPD was applied for the evaluation of isobaric interferences in mass spectra of alkylselenium and alkyltin compounds containing several overlapping $(M + zH)^+$ ions ($z = 0$–4)[16,17] and, more recently, the technique has been also applied for elemental analysis[18,19] and trace element speciation[20–23] including metabolism studies.[24–27] Isotope Dilution Analysis, based on the deconvolution of isotope patterns by multiple linear regression, requires the measurement of part of the isotope pattern (mass isotopomer distribution) of the mixture of natural and labelled analyte. The isotopic composition in the mixture is assumed to be a linear combination of two isotope patterns: the isotope pattern of the natural abundance compound and the isotope pattern of the labelled analogue. The individual contribution of each 'isotope pattern' to the overall mass spectrum can be calculated by multiple linear regression and provides the molar fractions of both labelled and unlabelled compound in the sample.

In this paper we evaluate the use of minimal labelling of the isotopically enriched molecules to avoid isotopic effects. As a proof of concept, we have selected the determination of phenol in waters for two main reasons: first, it requires a multi-step sample preparation procedure including derivatisation, liquid–liquid extraction and capillary GC separation (which are prone to isotopic effects) and, second, phenol is one of the few compounds which can be obtained commercially both as single ($^{13}C_1$) and fully labelled ($^{13}C_6$) in ^{13}C. The developed procedure included the calculation of propagated uncertainties that take into account the uncertainties of the natural and labelled mass isotopomer distributions.[28]

Experimental

Instrumentation

Chromatographic analysis was performed with a gas chromatograph model 6890N (Agilent Technologies, Tokyo, Japan) fitted with a split/splitless injector and an HP-5MS capillary column (cross-linked 5% phenyl-methyl siloxane, 30 m × 0.25 mm i.d., 0.25 μm coating). The gas chromatograph was equipped with a mass spectrometric detector model 5975B (Agilent Technologies). An analytical balance model AB204-S (Mettler Toledo, Zurich, Switzerland) was used for the gravimetric preparation of all solutions.

Reagents and materials

Phenyl acetate and phenol were purchased from Sigma-Aldrich (St. Louis, MO, USA). Labelled compounds $^{13}C_1$ phenol (99% nominal enrichment) and $^{13}C_6$ phenol (99% nominal enrichment) were obtained from Cambridge Isotope Laboratories Inc. (Andover, MA, USA) in pure solid form. Stock solutions were prepared by dissolving the corresponding standards in methanol (Merck, Darmstadt, Germany). All stock solutions were stored at 4 °C and employed to prepare daily gravimetrically diluted working standard solutions in methanol. Acetic acid anhydride, sodium chloride and sodium carbonate were supplied by Sigma-Aldrich and hexane was obtained from Merck. All solvents and reagents were of analytical reagent grade. Ultra pure water was obtained from a Milli-Q system (Millipore Co., Bedford, USA). Derivatisation and liquid–liquid extraction was performed in 100 ml glass volumetric flasks with glass stoppers or in 15 ml amber glass vials equipped with screw caps. They were cleaned with chromic acid for 24 h and rinsed with Milli-Q water. The Certified Reference Material NIST 1548 (Priority Pollutants Phenols in methanol) was purchased from the National Institute of Standards and Technology (USA) and was used to validate the different procedures.

Procedures

Derivatisation and extraction of phenols from water samples. A sample of 10 ml of water was placed in a 15 ml amber glass vial with screw cap and mixed with appropriate amounts of the natural abundance phenol standard or NIST 1584 and the corresponding labelled phenol standard dissolved in methanol ($^{13}C_1$ phenol or $^{13}C_6$ phenol). In order to correct for volumetric errors, the amounts of sample and spike added were determined gravimetrically. Then, the pH was adjusted to 9.2 with 0.05 g Na_2CO_3 and the ionic strength was adjusted with 0.35 g of sodium chloride. After total dissolution (*ca.* 10 min), 0.2 g of acetic acid anhydride and 2 ml of hexane were added and the flask was shaken for another 10 min. Finally, most of the organic layer was transferred to a glass vial with the help of a Pasteur pipette and stored at −18 °C until measurement. Initial optimization experiments were performed on 100 ml volumetric glass flasks.

Separation and measurement of mass isotopomer distributions by GC-EI-MS. A 2 ml min^{-1} flow of He was employed as the carrier gas flow. The temperature program consisted in a linear gradient between 50 and 300 °C at 30 °C min^{-1}. The injection was performed at 270 °C using a split/splitless injector using the pulsed splitless mode (1 min purge time and 60 psi of pressure pulse). The electron ionization was performed at an electron energy of 70 eV. The temperature of the transfer line, ion source and quadrupole was set at 280, 230 and 150 °C, respectively. Under these conditions phenol eluted after 4.70 min. Initial

qualitative analyses were performed from *m/z* 40 to 400 in the full-scan mode. Quantitative analyses were performed in selective ion monitoring (SIM) mode. The measurement of mass isotopomer distributions for the derivatised phenol was performed on the M − 43 molecular cluster (loss of the acetyl group, C_2H_3O–) using a 10 ms dwell time per mass. This dwell time assured the acquisition of more than 10 data points along the chromatographic peak even when ten consecutive *m/z* values were monitored. Between two and ten consecutive *m/z* values (starting from $m/z = 93$) were measured for the SIM mode for each natural abundance phenol, $^{13}C_1$ phenol, $^{13}C_6$ phenol or their mixtures. The isotope patterns were determined by dividing the peak area obtained at each mass by the sum of all peak areas measured. A triplicate injection was performed always to evaluate measurement variability.

Results and discussion

Theoretical development

Quantification by multiple linear regression in organic mass spectrometry. In the application of multiple linear regression in organic mass spectrometry the sample, containing a given compound showing natural isotopic composition, is spiked with a known amount of the same compound isotopically labelled in one or several of its constituting elements. Thus, in the spiked sample (or mixture) we have molecules from the same compound showing two different mass isotopomer distributions or isotope patterns: the natural abundance and the isotopically labelled. Assuming that there are not isotope effects on the instrumental responses, the isotope distribution of a given molecular fragment of the molecule in the mixture can be measured by low resolution mass spectrometry as indicated in the procedures. In the proposed approach, part or the whole isotope pattern can be measured. The total amount (moles) of a compound in a given mixture, N_{mix}, can be distributed between the two sources by:

$$N_{mix} = N_{nat} + N_{lab} \quad (1)$$

where N_{mix} is the number of moles of the compound found in the mixture; N_{nat} is the number of moles coming from the compound with natural isotope abundances; and N_{lab} is the number of moles of the labelled compound in the mixture coming from the spike added. Similar amount balances can be also obtained for all mass isotopomers of the compound considered, as shown by eqn (2), illustrated for a nominal mass *i* as an example:

$$N^i_{mix} = N^i_{nat} + N^i_{lab} \quad (2)$$

Note that different isotope compositions of the same compound may coexist at the same nominal mass. For example, phenol at nominal mass 94 is constituted only of $^{12}C_6{}^1H_6{}^{16}O$ (monoisotopic mass) while the same compound at nominal mass 95 will be a mixture of $^{12}C_5{}^{13}C_1{}^1H_6{}^{16}O$, $^{12}C_6{}^1H_5{}^2H_1{}^{16}O$ and $^{12}C_6{}^1H_6{}^{17}O$. For the sake of this demonstration we will assume that only low resolution measurements are performed and that all isotope compositions possessing the same nominal mass are added together in the mass spectrum without differences in ionization efficiencies or isotopic effects. Then, eqn (2) can be expressed as a linear combination of the total amount of each compound (natural abundance or labelled) and their corresponding mass isotopomer distributions:

$$N_{mix} \cdot A^i_{mix} = N_{nat} \cdot A^i_{nat} + N_{lab} \cdot A^i_{lab} \quad (3)$$

where A^i_{mix} is the experimentally measured relative abundance of the compound at nominal mass *i* in the mixture; A^i_{nat} is the theoretical relative abundance of the natural abundance compound at the same nominal mass and A^i_{lab} is the theoretical relative abundance in the isotopically labelled compound. The values of A^i_{mix} are determined experimentally by peak area integration at different consecutive masses in the same mass fragment. The isotope abundance of each isotopomer is then calculated by dividing the peak area at a given mass by the sum of all areas as indicated in the procedures. When we divide eqn (3) by eqn (1), the following expression is obtained:

$$A^i_{mix} = x_{nat} \cdot A^i_{nat} + x_{lab} \cdot A^i_{lab} \quad (4)$$

where

$$x_{nat} = \frac{N_{nat}}{N_{nat} + N_{lab}} \quad (5)$$

and

$$x_{lab} = \frac{N_{lab}}{N_{nat} + N_{lab}} \quad (6)$$

In expressions (5) and (6) the variables x_{nat} and x_{lab} indicate the molar fractions of the compound in the spiked sample arising from the two different sources: natural abundance or isotopically labelled. A typical organic compound may show different isotope compositions at *n* nominal masses so, equations for all these nominal masses can be obtained in a similar way to eqn (4) and they can be expressed in matrix notation as:

$$\begin{bmatrix} A^1_{mix} \\ A^2_{mix} \\ A^3_{mix} \\ \cdots \\ A^{n-1}_{mix} \\ A^n_{mix} \end{bmatrix} = \begin{bmatrix} A^1_{nat} & A^1_{lab} \\ A^2_{nat} & A^2_{lab} \\ A^3_{nat} & A^3_{lab} \\ \cdots & \cdots \\ A^{n-1}_{nat} & A^{n-1}_{lab} \\ A^n_{nat} & A^n_{lab} \end{bmatrix} \cdot \begin{bmatrix} x_{nat} \\ x_{lab} \end{bmatrix} + \begin{bmatrix} e^1 \\ e^2 \\ e^3 \\ \cdots \\ e^{n-1} \\ e^n \end{bmatrix} \quad (7)$$

As we have more parameters (nominal masses) than unknowns (molar fractions) an error vector is included in eqn (7). The best values of x_{nat} and x_{lab} are found by least square minimization of the error vector '***e***'. If we name the vector of the isotope abundances in the mixture as '***y***', the matrix of the 'theoretical' isotope abundances of the pure components as '***A***', and the vector of the unknown molar fractions as '***x***', the least square solution of this over-determined system of equations can be calculated as:

$$\boldsymbol{x} = (\boldsymbol{A}^T \times \boldsymbol{A})^{-1} \times (\boldsymbol{A}^T * \boldsymbol{y}) \quad (8)$$

where $\boldsymbol{A}^T$ indicates the transpose of $\boldsymbol{A}$ and superscript −1 the inverse.

Fig. 1 illustrates the procedure for a mixture of natural abundance phenol and $^{13}C_1$ phenol. The isotope pattern measured for the mixture at the chromatographic peak of phenol is compared with the theoretical isotope patterns of both natural

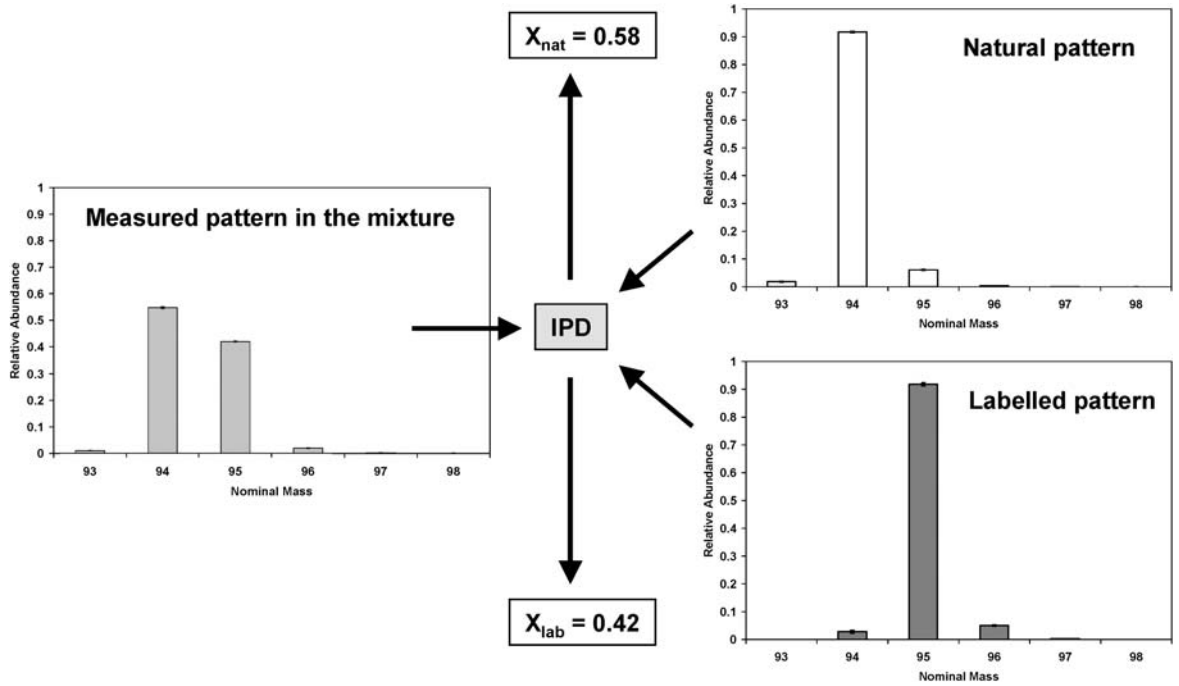

Fig. 1 Illustration of the isotope pattern deconvolution (IPD) process by multiple linear regression for natural abundance phenol and $^{13}C_1$ phenol.

and labelled phenol and the relative contribution of those theoretical patterns to the observed distribution is calculated by multiple least squares. In this particular case the results were $x_{nat} = 0.58$ and $x_{lab} = 0.42$.

When we have more parameters than unknowns (a minimum of three nominal masses) we can also determine the standard uncertainties for those parameters using the variance–covariance matrix, $\boldsymbol{V}(\boldsymbol{x})$. The diagonal elements of this matrix are the variances of the variables $\boldsymbol{x}_{nat}$ and $\boldsymbol{x}_{lab}$. From the multivariate linear regression this matrix can be calculated as:

$$\boldsymbol{V}(\boldsymbol{x}) = (\boldsymbol{A}^{T} \cdot \boldsymbol{A})^{-1} \cdot \boldsymbol{s}_e^2 \quad (9)$$

where $\boldsymbol{s}_e^2$ is the variance of the regression model (the sum of squares of errors divided by the degrees of freedom, $n - 2$). All these matrix calculations can be performed in a single step using the function 'LINEST' in Microsoft Excel. Please note that eqn (7) will provide an exact solution when $n = 2$ and a least squares solution when $n > 2$. Hence, the selection of the number of masses to be used in the calculations will depend on practical considerations such as counting statistics and chromatographic peak profiles.

Once the molar fractions of the compound and their uncertainties are computed by the linear regression the amount of natural abundance compound can be calculated, as N_{lab} is known, using the equation:

$$\frac{N_{nat}}{N_{lab}} = \frac{x_{nat}}{x_{lab}} = R \quad (10)$$

Eqn (10) comes from dividing eqn (5) by eqn (6). Note that eqn (10) takes the form of straight line of slope 1 and intercept 0 when the ratio of moles (N_{nat}/N_{lab}) is plotted against the ratio of molar fractions ($R = x_{nat}/x_{lab}$), regardless of the labelling of the molecule.

Determination of cluster purity. In the previous demonstration we have assumed that the isotope distribution measured in the mixture by mass spectrometry comes from a single ion type, $(M)^+$. However, different ion clusters with one or several mass units of difference can overlap because of different fragmentation pathways.[16,17,29] So, the main ion cluster $(M)^+$ may overlap with $(M - zH)^+$ ions (usually $z = 1$) and this is more the rule than the exception.[29] These spectral interferences may prevent the application of eqn (7) when the isotope compositions, A_{nat} and A_{lab}, are not calculated taking into account this possible spectral overlap.

Spectral overlap due to $(M - zH)^+$ ions has been described in classical textbooks on organic spectroscopy. Lambert *et al.*[15] employed a stepwise isotope pattern deconvolution procedure in which the isotope pattern observed for the natural abundance compound was used to calculate the contributions at M − 2H and M − H. This procedure was used to study isotope incorporation in labelled compounds.[15] In our case, we studied cluster purity using the procedure described by Meija *et al.* based on multiple least squares.[16,17] The fragmentation patterns of organometallic compounds such as dimethyldiselenide[16] or butyltin compounds[17] were determined by GC-EI-MS using multiple least squares. For these calculations matrix $\boldsymbol{A}$ corresponded to the 'theoretical' isotope compositions of the compound taking into account the possible loss or gain of hydrogen atoms. For example, eqn (11) shows the calculation of the experimental fragmentation pattern of natural abundance phenol using three theoretical isotope compositions:[28] phenol $(C_6H_6O)^+$, phenol with the loss of one hydrogen atom $(C_6H_5O)^+$,

and phenol with the gain of one hydrogen atom $(C_6H_7O)^+$, at nominal masses $i = 93$ to 98:

$$\begin{bmatrix} A^{93}_{nat} \\ A^{94}_{nat} \\ A^{95}_{nat} \\ A^{96}_{nat} \\ A^{97}_{nat} \\ A^{98}_{nat} \end{bmatrix} = \begin{bmatrix} 0.00000 & 0.93468 & 0.00000 \\ 0.93457 & 0.06155 & 0.00000 \\ 0.06165 & 0.00362 & 0.93446 \\ 0.00363 & 0.00015 & 0.06175 \\ 0.00015 & 0.00000 & 0.00363 \\ 0.00000 & 0.00000 & 0.00015 \end{bmatrix} \cdot \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} + \begin{bmatrix} e^{93} \\ e^{94} \\ e^{95} \\ e^{96} \\ e^{97} \\ e^{98} \end{bmatrix} \quad (11)$$

The values of α, the fraction of the cluster in the form of $(M)^+$, β, the fraction of the cluster in the form of $(M - H)^+$ and γ, the fraction of the cluster in the form of $(M + H)^+$ are calculated using eqn (11) based on experimentally measured A^i_{nat} (relative abundances at nominal masses 93–98). Once the values of α, β and γ are calculated, then the 'estimated' isotope compositions of the natural abundance and labelled compound, to be used in eqn (7), can be computed based on a linear combination. For the case of the natural abundance compound we have:

$$A^i_{nat} = \alpha \cdot A^i_M + \beta \cdot A^i_{M-H} + \gamma \cdot A^i_{M+H} \quad (12)$$

where A^i_{nat} is the natural 'estimated' isotope composition of the ion cluster for nominal mass i and A^i_M, A^i_{M-H} and A^i_{M+H} are the theoretical isotope compositions of each cluster calculated based on the natural isotope abundances of the corresponding elements. Eqn (11) and eqn (12) could be extended to more ion types (*e.g.* M − 2H and/or M + 2H) if necessary. The determination of the 'estimated' isotope composition of the labelled compound, A^i_{lab}, will be performed with an equation analogue to eqn (12) but using the theoretical isotope compositions of A^i_M, A^i_{M-H} and A^i_{M+H} for the labelled compound and the same values of α, β and γ.

It is important to clarify that the fragmentation characteristics of the compound need to be determined, using either multiple least squares[16] or stepwise calculation procedures,[15] for the direct application of eqn (7) in IDMS. Clearly, the values of the isotope compositions used in matrix ***A*** should take into account the fragmentation pattern of the compound. Thus, the different values of α, β, γ, *etc.* should be independent from: (i) the isotope composition of the molecule (natural abundance or labelled), (ii) the ionization conditions in the ion source and (iii) the concentration of the compound injected in the chromatographic system.

Study of the fragmentation of phenyl acetate and derivatised natural abundance phenol

First, it was observed that the fragmentation pattern of a phenyl acetate standard was exactly the same as that of derivatised natural abundance phenol. The main peak in the mass spectrum $(M - C_2H_3O)^+$ at *m*/*z* around 94 contained only two ion types: $(M)^+$ and $(M - H)^+$ at a ratio of approximately 50:1. The fragmentation factors α, $(M)^+$, and β, $(M - H)^+$, calculated using eqn (11) were constant and independent of the concentration of phenyl acetate, the helium carrier gas flow, the temperature program and the electron energy in the ion source (35, 70 and 140 eV tested). In addition, during the optimization of the derivatisation and extraction conditions of phenol in the water samples the fragmentation pattern did not change during the course of a whole working day. From all these data the fragmentation factors α and β for natural abundance phenol were estimated to be 0.98 and 0.02 respectively.

Fragmentation pattern of the $^{13}C_1$- and $^{13}C_6$-labelled phenols

For the study of the fragmentation patterns for the labelled phenols we prepared aqueous solutions of 100 ml containing *ca.* 40 ng g^{-1} of $^{13}C_1$ phenol and $^{13}C_6$ phenol, respectively. For comparison purposes a solution containing natural abundance phenol was also prepared. Each solution was derivatised as indicated in the procedures but using 1 g of acetic acid anhydride instead of 0.2 g. Each standard was injected in triplicate and the average isotope distributions obtained for each compound were calculated. It was found that the measured isotope distributions for both the pure $^{13}C_1$ phenol and the $^{13}C_6$ phenol showed contamination from natural abundance phenol. Several experiments were performed in order to find out whether the sources of contamination were the labelled compounds themselves or some other reagent. Finally, the acetic acid anhydride used for derivatisation was identified as the only source of contamination with natural abundance phenol. As the contribution of natural abundance phenol in the labelled compounds could not be avoided the determination of their fragmentation patterns had to be performed by extrapolation, particularly for the singly labelled phenol. Thus, an experiment was performed in which the isotope distributions for both labelled standards were measured using increasing amounts of acetic acid anhydride for derivatisation (from *ca.* 0.2 to 2 g). The results obtained for $^{13}C_1$ phenol and $^{13}C_6$ phenol are given in Table 1 using four theoretical isotope patterns to fit the data: nat-$C_6H_6O^+$, nat-$C_6H_5O^+$, lab-$C_6H_6O^+$ and lab-$C_6H_5O^+$. We can observe the increase of the contribution of natural abundance phenol (see the contribution of nat-$C_6H_6O^+$ in Table 1) with the amount of acetic acid anhydride for both isotopically labelled standards. Thus, the determination of the fragmentation pattern of both labelled standards requires the elimination of the contribution of natural abundance phenol to the observed isotopic distribution. For the fully labelled compound the results were satisfactory as the four patterns could be fitted to the model. However, for the singly labelled phenol only three patterns could be fitted because of the almost identical isotopic composition of nat-$C_6H_6O^+$ and lab-$C_6H_5O^+$. So, for the singly labelled phenol the fragmentation factors will have to be obtained by extrapolation at 0 g of acetic acid anhydride to eliminate the contribution of natural abundance phenol in the observed fragmentation factors.

For the $^{13}C_6$ phenol, the plot of the nat-$C_6H_6O^+$ molar fraction *vs.* the amount of acetic acid anhydride added provided a linear graph with an intercept of 0.0046 ± 0.0046 indicating that no impurities of natural abundance phenol in this labelled standard existed. Clearly, the amount of natural abundance phenol detected was due solely to the contamination of the acetic acid anhydride. The extrapolation of the contribution of lab-$C_6H_5O^+$ at 0 g of acetic acid anhydride provided the β value (loss of one hydrogen atom) for the fully labelled compound. Also, for each different amount of acetic acid anhydride added, the ratio of contributions from lab-$C_6H_6O^+$ and lab-$C_6H_5O^+$ in the labelled

Table 1 Experimental fragmentation factors obtained for $^{13}C_1$ phenol and $^{13}C_6$ phenol when changing the amount of acetic acid anhydride used for derivatisation. The uncertainties are shown in brackets for the last two digits

	Amount of acetic acid anhydride added for derivatisation (g)				
Fragmentation factors for $^{13}C_1$	0.2132	0.5491	1.0944	1.6251	2.1727
lab-$C_6H_6O^+$	0.9696 (20)	0.9548 (15)	0.9334 (14)	0.9246 (13)	0.9184 (14)
lab-$C_6H_5O^+$ + nat-$C_6H_6O^+$	0.0317 (19)	0.0456 (15)	0.0665 (14)	0.0749 (13)	0.0810 (14)
nat-$C_6H_5O^+$	0.0017 (19)	0.0019 (15)	0.0021 (14)	0.0024 (13)	0.0024 (13)
	Amount of acetic acid anhydride added for derivatisation (g)				
Fragmentation factors for $^{13}C_6$	0.2330	0.5753	1.0824	1.6369	2.1622
lab-$C_6H_6O^+$	0.9672 (12)	0.9566 (11)	0.9393 (11)	0.9349 (12)	0.9073 (15)
lab-$C_6H_5O^+$	0.0196 (11)	0.0191 (12)	0.0195 (11)	0.0189 (11)	0.0189 (13)
nat-$C_6H_6O^+$	0.0109 (11)	0.0216 (11)	0.0378 (11)	0.0427 (11)	0.0691 (14)
nat-$C_6H_5O^+$	0.0003 (11)	0.0005 (11)	0.0009 (11)	0.0011 (11)	0.0017 (13)

compound should be constant as the contribution of the natural abundance phenol has been computed separately. Both calculation methods provided β values for $^{13}C_6$ phenol of 0.0195 ± 0.0002 and 0.0200 ± 0.0004 respectively. These results were in agreement with those calculated from a triplicate injection of natural abundance phenol ($\beta = 0.0218 \pm 0.0018$).

For the $^{13}C_1$ phenol the fraction of lab-$C_6H_5O^+$ can only be calculated by extrapolation to 0 g of acetic acid anhydride as the contribution of natural abundance phenol (nat-$C_6H_6O^+$) cannot be separated from that of lab-$C_6H_5O^+$. The intercept of this curve was calculated to be $\beta = 0.0206 \pm 0.0020$ by applying a quadratic regression. This β value is in agreement with the fragmentation values calculated for the other phenols.

According to all the previous results, it was decided to use a fragmentation pattern of $\alpha = 0.980$ for the $C_6H_6O^+$ ion and $\beta = 0.020$ for the $C_6H_5O^+$ ion in the calculations of the 'estimated' isotope patterns for all compounds. The standard uncertainties of the fragmentation factors were calculated to be 0.002 for both α and β. The isotope patterns shown in Fig. 1 for the natural and labelled compounds were calculated taking into account these fragmentation factors.

Uncertainty sources

Apart from the usual sources of uncertainty in IDMS (concentration of the spike, weights of sample and spike taken, measurement uncertainty) there are two uncertainty sources unique to this alternative approach that must be studied: the uncertainty from the multiple linear regression calculations and the uncertainty in the theoretical isotope composition of the natural abundance and labelled compounds.[28]

Uncertainty from the multiple linear regression. Eqn (9) shows that the uncertainty of the molar fractions will be a function of the regression variance s_e^2 and the diagonal values of the matrix $(\boldsymbol{A}^T \times \boldsymbol{A})^{-1}$. The regression variance s_e^2 will depend basically on experimental measurement errors while the value of the matrix $(\boldsymbol{A}^T \times \boldsymbol{A})^{-1}$ will depend on the actual values of the theoretical isotope composition both for the natural abundance and the labelled compound and represents what, in elemental isotope dilution analysis, is called the error magnification factor. If measurement errors cannot be optimized we can still minimize the uncertainty of the multiple linear regression by minimizing the values of the $(\boldsymbol{A}^T \times \boldsymbol{A})^{-1}$ matrix.

We have calculated the value of the $(\boldsymbol{A}^T \times \boldsymbol{A})^{-1}$ matrix for organic molecules of general formula C_mH_{2m+2} where m varied from 5 to 200. For that purpose we have first calculated the isotopic composition of these molecules when they are labelled with 0, 1, 2 or 4 carbon-13 atoms enriched at 99% using a calculation procedure described in detail elsewhere.[28] Then, we have computed the diagonal terms of the $(\boldsymbol{A}^T \times \boldsymbol{A})^{-1}$ matrix for the combination of the unlabelled compound ('0' carbon-13 atoms) with the presence of either 1, 2 or 4 carbon-13 atoms in the molecule. The values found for the first diagonal term correspond to the square of the error magnification of the uncertainty of the molar fraction of the natural abundance compound and are plotted in Fig. 2. As expected, the minimum value of the error magnification factor is 1. Then, the value increases as the number of carbon atoms in the molecule, m, grows. The lower error propagation values are always found when 4 carbon-13 atoms are used for labelling and it could be expected that lower values still will be found for larger ^{13}C-labelling. However, for small organic molecules, where

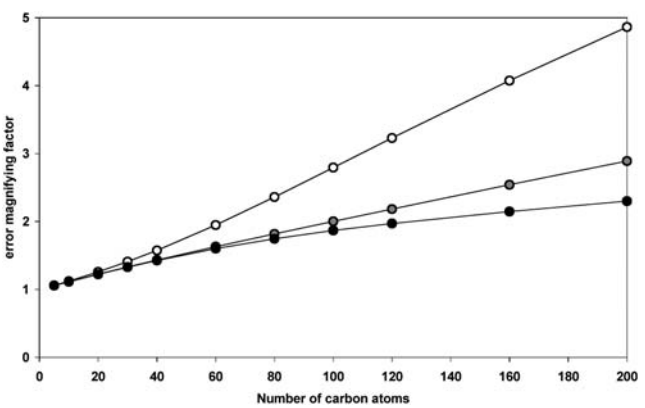

Fig. 2 Square root of the values of the first diagonal term of the $(\boldsymbol{A}^T \times \boldsymbol{A})^{-1}$ matrix for alternative labelling of a C_mH_{2m+2} molecule with (○) 1 carbon-13 atom, (◍) 2 carbon-13 atoms or (●) 4 carbon-13 atoms.

$m < 20$, there is almost no difference in the error magnification factors when 1, 2 or 4 carbon-13 atoms are used for labelling. When the number of carbon atoms in the molecule is between 20 and 60 at least 2 carbon-13 atoms would be required. For molecules containing more than 60 carbon atoms the use of 4 or more carbon-13 atoms would be recommended. Thus, we can conclude that for small organic molecules up to 20 carbon atoms, currently found in environmental analysis, the choice of single or multiple ^{13}C-labelling will not have a drastic influence on the overall uncertainty of the procedure. For larger molecules, multiple labelling would be advisable for lower propagated uncertainties.

Another source of uncertainty in the multiple linear regression is the fact that molar fractions, x_{nat} and x_{lab}, are variables of constant sum and, as indicated by Meija and Mester[30] for elemental isotope abundances, are correlated variables. For a mixture of only two patterns, as it is the present case, the correlation coefficient is -1.[30] So, when calculating the uncertainty of the ratio $R = x_{nat}/x_{lab}$, s_R, for the application of eqn (10), the correlation between both molar fractions will need to be taken into account. Following the general formula of error propagation and taking into account the correlation between the molar fractions, we have obtained the following equation:

$$\frac{s_R}{R} = \sqrt{\left(\frac{s_{x_{nat}}}{x_{nat}}\right)^2 + \left(\frac{s_{x_{lab}}}{x_{lab}}\right)^2 + 2\,\frac{s_{x_{nat}} s_{x_{lab}}}{x_{nat} x_{lab}}} \qquad (13)$$

which allows us to calculate the relative uncertainty in the ratio, s_R/R, from the determined uncertainties in the molar fractions. The values of the uncertainties in x_{nat} and x_{lab} are obtained from the variance–covariance matrix $\boldsymbol{V}(\boldsymbol{x})$ in eqn (9) and depend on the experimental measurement conditions.

Uncertainty from the theoretical isotope composition of the compound. Another possible source of uncertainty in this particular approach for isotope dilution analysis will be the uncertainty in the theoretical isotope composition for both the natural abundance and the labelled compound, which appear in matrix $\boldsymbol{A}$, and are shown graphically in Fig. 1. These uncertainties will have three potential sources: (a) the natural variability of the isotopic composition of the natural abundance compound, due to natural variations in the ^{13}C, ^{2}H, ^{17}O, ^{18}O, *etc.* isotope abundances, (b) the uncertainty in the isotope enrichment of the labelled compound, and (c) the uncertainties of the fragmentation factors α, β, *etc.* determined experimentally for the target compound.

Traditionally, in organic isotope dilution analysis, the natural variability of the isotopic composition of the elements has been ignored because no theoretical isotope compositions are involved in the calculations. However, in this alternative approach the uncertainty in the isotope composition of the target molecule must be evaluated. We have developed[28] spreadsheet software in Microsoft Excel to compute the isotopic composition of molecules and their uncertainties using a Visual Basic macro following the computation algorithm published by Kubinyi.[31] The program was tested by calculating the isotopic composition of bovine insulin ($C_{254}H_{377}N_{65}O_{75}S_6$) with satisfactory results in comparison to those published previously.[31] The uncertainties of the isotopic composition of the molecules were calculated based on Kragten's procedure of uncertainty propagation[32] and taking also into account the correlations between the isotope compositions of the elements as additional uncertainty sources.[28,30] For bi-isotopic elements, such as C, H, Cl, Br, *etc.* the correlation coefficient was taken always as -1. For tri-isotopic elements, such as oxygen, the correlation coefficients were calculated using the mass dependent fractionation law.[28] The detailed demonstration for the calculation of these uncertainties and correlation values for different elements is given elsewhere.[28]

The theoretical isotope composition for phenol and the standard uncertainties calculated based on the natural variability uncertainties are shown in Table 2. As can be observed, for the most abundant nominal mass (mass 94) the abundance is 93.46% with a standard uncertainty of 0.23%. In a similar way we have calculated the isotope composition for the $^{13}C_1$- and the $^{13}C_6$-labelled phenols and those are also given in Table 2 including their standard uncertainties. For the labelled phenols the uncertainty of the isotope enrichment of carbon was taken from the values given by the manufacturer. Both compounds were indicated as 99% enriched in ^{13}C so the uncertainty in the isotope composition was taken as $\pm 1\%/\sqrt{3}$ as recommended by the Eurachem Citac Guide.[33] As can be observed, the uncertainties for the labelled compounds are larger than for the natural abundance phenol and increased with the number of enriched ^{13}C atoms in the molecule. For the $^{13}C_6$-labelled phenol the most abundant nominal mass (mass 100) shows an abundance of 93.9% with standard uncertainty of 3.3%. It is clear from these results that, for highly labelled compounds, a better certificate of its isotope composition is required.

The third source of uncertainty influencing the total uncertainty budget is the fragmentation factors α, β, *etc.* Their experimental uncertainties can be calculated based on eqn (9) using the Kragten procedure as well as the uncertainties in the isotope composition of the molecules and their correlation factors. Based on the previous results obtained for α and β, the 'estimated' isotope compositions of phenol were calculated and are also given in Table 2. As can be observed, there is a small increase in the uncertainties of the isotope composition due to the measurement of the fragmentation pattern. As a general rule it can be expected that, when α tends to 1 and β, γ, *etc.* tend to 0, the uncertainty in the fragmentation pattern will have a smaller influence in the total uncertainty. So, it is advisable to select molecular clusters showing the highest purity of the $(M)^+$ ion. The uncertainties shown in Table 2 are also presented in Fig. 1 for the natural and singly labelled phenol.

Comparison of theoretical and experimental IDMS calibration graphs

Theoretical calibration graphs were computed by calculating the isotopic distribution of the M − 43 cluster in the different blends prepared. For this purpose, we assumed no mass discrimination effects in the mass spectrometer and equal fragmentation and ionization efficiencies for the natural and labelled compound. The theoretical isotopic compositions of the different mixtures both for a pure $C_6H_6O^+$ cluster and for a mixture containing 2% of $C_6H_5O^+$ were calculated following a linear molar mixture model using the calculated isotopic composition of the natural and labelled compounds shown in Table 2:

Table 2 Theoretical mass isotopomer distributions for the natural abundance, $^{13}C_1$- and $^{13}C_6$-labelled phenol, $C_6H_6O^+$, used to calculate the theoretical IDMS calibration graphs and for the multiple linear regression calculations[28]

	Theoretical mass isotopomer distributions					
	100% $C_6H_6O^+$					
Nominal mass	Natural phenol	Uncertainties	$^{13}C_1$ phenol[a]	Uncertainties	$^{13}C_6$ phenol[b]	Uncertainties
93	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
94	0.93457	0.00228	0.00945	0.00545	0.00000	0.00000
95	0.06165	0.00215	0.93575	0.00578	0.00000	0.00000
96	0.00363	0.00014	0.05161	0.00184	0.00000	0.00000
97	0.00015	0.00001	0.00307	0.00011	0.00002	0.00006
98	0.00000	0.00000	0.00012	0.00001	0.00144	0.00214
99	0.00000	0.00000	0.00000	0.00000	0.05688	0.03289
100	0.00000	0.00000	0.00000	0.00000	0.93861	0.03333
101	0.00000	0.00000	0.00000	0.00000	0.00112	0.00021
102	0.00000	0.00000	0.00000	0.00000	0.00193	0.00010
	Theoretical mass isotopomer distributions					
	98% $C_6H_6O^+$ and 2% $C_6H_5O^+$					
Nominal mass	Natural phenol	Uncertainties	$^{13}C_1$ phenol[a]	Uncertainties	$^{13}C_6$ phenol[b]	Uncertainties
93	0.01869	0.00187	0.00019	0.00011	0.00000	0.00000
94	0.91711	0.00284	0.02798	0.00566	0.00000	0.00000
95	0.06049	0.00211	0.91807	0.00593	0.00000	0.00000
96	0.00356	0.00013	0.05064	0.00181	0.00000	0.00000
97	0.00015	0.00001	0.00301	0.00012	0.00005	0.00007
98	0.00000	0.00000	0.00012	0.00001	0.00255	0.00220
99	0.00000	0.00000	0.00000	0.00000	0.07452	0.03228
100	0.00000	0.00000	0.00000	0.00000	0.91986	0.03271
101	0.00000	0.00000	0.00000	0.00000	0.00114	0.00021
102	0.00000	0.00000	0.00000	0.00000	0.00189	0.00010

[a] The isotope enrichment of the enriched $^{13}C_1$ was 99% as given by the manufacturer. The other 5 carbons in the cluster were assumed to be of natural abundance (98.93% ^{12}C). [b] The isotope enrichment of the enriched $^{13}C_6$ was 99% as given by the manufacturer.

$$A^i_{mix} = \frac{(A^i_{nat}N_{nat} + A^i_{lab}N_{lab})}{(N_{nat} + N_{lab})} \quad (14)$$

where N_{nat} and N_{lab} are the number of moles of natural and labelled phenols added to prepare the blends, and A^i_{mix}, A^i_{nat} and A^i_{lab} are the calculated and theoretical abundances (Table 2) of mass i in the cluster for the mixture, the natural abundance phenol and the labelled phenol, respectively. The theoretical calibration graphs were prepared plotting the calculated ratio $A^{94}_{mix}/A^{95}_{mix}$(or $A^{94}_{mix}/A^{100}_{mix}$) *vs.* the N_{nat}/N_{lab} ratio. These theoretical calibration graphs are plotted in Fig. 3A for the singly labelled phenol and Fig. 3B for the fully labelled phenol. Please note that for the singly labelled phenol (Fig. 3A) two theoretical graphs are plotted. The continuous line is the theoretical graph calculated using a 100% of $C_6H_6O^+$ in the cluster while the dashed line was calculated taking into account the presence of 2% of $C_6H_5O^+$ in the cluster.

Increasing concentrations (0–100 ng g^{-1}) of unlabelled phenol were prepared in 10 ml of Milli-Q water and spiked with the necessary amount of the labelled analogue to get a resulting concentration of *ca.* 30 ng g^{-1}. The spiked sample was derivatised and extracted as described in the procedures. The IDMS experimental calibration graphs were prepared by plotting the intensity ratios (peak area ratios of the *m*/*z* 94 and 95 for the $^{13}C_1$ phenol and peak area ratios of the *m*/*z* 94 and 100 for the $^{13}C_6$ phenol) *vs.* the molar ratio between natural and labelled compound in the water sample. The experimental results are shown also in Fig. 3 and were obtained in two independent sets of experiments carried out with more than one year of difference: in January 2008 and in May 2009. Another difference between both sets of experiments was that in 2008 a group of 6 (for $^{13}C_1$) or 10 (for $^{13}C_6$) consecutive masses were measured while in 2009 only two masses were monitored to improve counting statistics (94 and 95 or 94 and 100 respectively).

The results obtained for the singly labelled phenol are shown in Fig. 3A. As can be observed, the theoretical and experimental IDMS calibration graphs are non-linear as expected by the mass overlap. The experimental data of both sets of experiments (2008 and 2009) were in agreement with both theoretical curves but the fit was better when the contribution of $C_6H_5O^+$ was taken into account. We can conclude, from the data shown in Fig. 3A, that eqn (14) can predict accurately the experimental IDMS calibration graphs when there is mass overlap between the natural and labelled compound. Obviously, the isotope compositions used to build the calibration graph should take into account the fragmentation characteristics of the molecule.

The data for the $^{13}C_6$ phenol are shown in Fig. 3B. In this case, the theoretical plot is linear with a theoretical slope of 0.995 and an intercept of 0.000. This is due to the six mass units of difference between natural and labelled phenol which eliminates mass

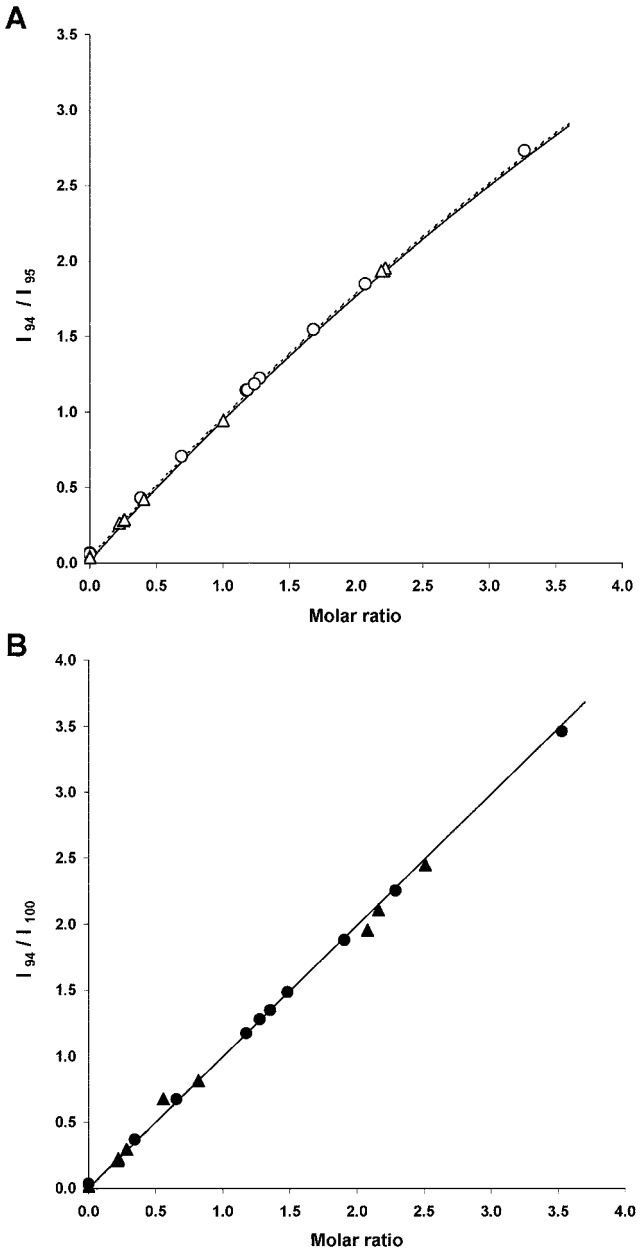

Fig. 3 Comparison between the theoretical calibration graphs, both for the pure $C_6H_6O^+$ cluster (—) and for the 2% of $C_6H_5O^+$ contribution (---), and the experimental IDMS results for: (A) $^{13}C_1$ phenol [data obtained in 2008 (△) and 2009 (○)] and (B) $^{13}C_6$ phenol [data from 2008 (▲) and 2009 (●)].

overlap. Both theoretical lines obtained from the data given in Table 2 (ignoring or taking into account the 2% of hydrogen loss) overlap almost completely so only one theoretical plot is shown. The slope of the experimental IDMS calibration graph obtained in the experiments of 2008 (black triangles) was 0.957 ± 0.015 while the slope obtained in 2009 (black circles) was 0.974 ± 0.002 with improved counting statistics when measuring only two masses. It seems that there is a significant difference between the theoretical slope of 0.995 and the experimental slopes measured here. We think that this effect is related to small differences in the physicochemical properties between natural abundance phenol and $^{13}C_6$-labelled phenol during the processes of derivatisation and liquid–liquid extraction.

Application of multiple linear regression in IDMS

For the multiple linear regression calculations the peak areas measured for masses 94 and 95, for the mixtures of natural abundance phenol and singly labelled phenol, and 94 and 100, for the mixtures of natural abundance phenol and the fully labelled compound, were used. In another set of experiments, masses 93, 94, 95, 96, 97 and 98, for the singly labelled compound, and masses 93, 94, 95, 96, 97, 98, 99, 100, 101 and 102, for the fully labelled compound, were also employed. The experimental isotope abundances at a given mass were calculated by dividing the peak area of that mass in the mixture by the sum of all areas. The molar fractions x_{nat} and x_{lab} were calculated using eqns (7) and (8). Matrix ***A*** corresponded to the values shown in Table 2 and Fig. 1 taking into account the presence of 2% $C_6H_5O^+$ in the fragmentation pattern of phenol. The plot of the ratio of molar fractions *vs.* the ratio of molar concentrations should give a straight line of slope 1 and intercept 0 according to eqn (10). The results obtained using both $^{13}C_1$ and $^{13}C_6$ phenol are shown in Fig. 4 including a reference line of slope 1 and intercept of 0 for comparison. Every point in Fig. 4 is the average of three independent injections. The experimental standard deviations ($n = 3$) are smaller than the size of the points. The white data points (circles, 2 masses, 2009; triangles, 6 masses, 2008; squares, 6 masses, 2009) correspond to the $^{13}C_1$ phenol while the black data points were obtained with $^{13}C_6$ phenol (circles, 2 masses, 2009; triangles, 10 masses, 2008; squares, 10 masses, 2009).

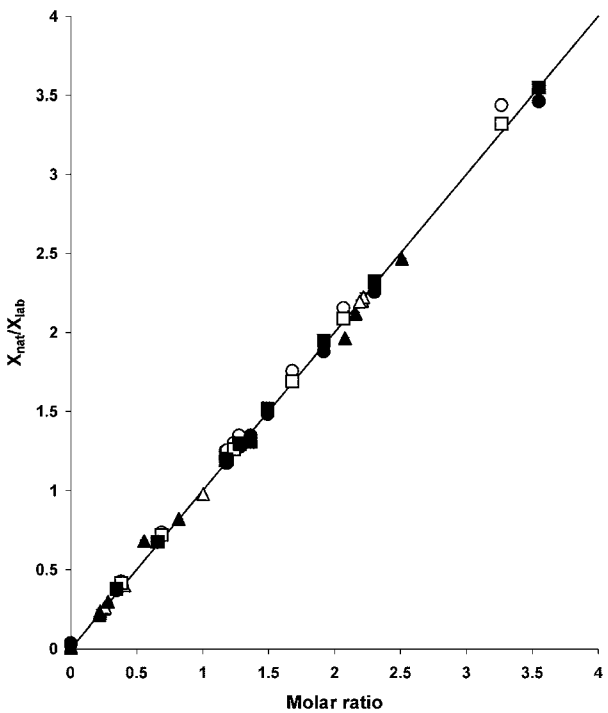

Fig. 4 Application of isotope pattern deconvolution in IDMS for the determination of phenol using $^{13}C_1$ phenol (○, □ and △) and $^{13}C_6$ phenol (●, ■ and ▲). The line corresponds to a slope of 1 and an intercept of 0. For explanation of the symbols see text.

Table 3 Slopes and intercepts of the multiple linear regression calibration functions and their standard uncertainties

Experiment	Slope	Intercept	Slope different from 1?[a]
$^{13}C_1$, 6 masses, 2008	0.996 ± 0.004	0.004 ± 0.004	No
$^{13}C_1$, 6 masses, 2009	1.007 ± 0.004	0.026 ± 0.005	No
$^{13}C_1$, 2 masses, 2009	1.037 ± 0.003	0.027 ± 0.004	Yes
$^{13}C_6$, 10 masses, 2008	0.962 ± 0.015	0.029 ± 0.019	No
$^{13}C_6$, 10 masses, 2009	0.989 ± 0.005	0.033 ± 0.007	No
$^{13}C_6$, 2 masses, 2009	0.9655 ± 0.0008	0.035 ± 0.001	Yes

[a] For 95% confidence.

Several conclusions can be extracted from Fig. 4. First, all data points loosely follow the theoretical line of slope 1 and intercept 0 as expected. That means that eqn (10) can be applied and calibration-free IDMS measurements could be performed by this procedure. Second, the data points obtained in 2008 agree very well with data obtained in 2009 for 6 and 10 masses measured (triangles and squares) indicating that the procedure is reproducible. Third, when comparing the measurement of 2 or 6 and 10 *m*/*z* values for the same mixtures (circles and squares), it seems that better agreement with the theoretical line is observed when more than two masses are selected. Finally, the slopes and intercepts of the IPD calibration functions of the set of experiments carried out in 2008 and 2009 are summarized in Table 3.

As can be observed, for the singly labelled compound the slope is not significantly different from 1. Also, there seems to be no clear isotopic differences between $^{13}C_1$ and $^{13}C_6$ phenol from the data obtained in 2009. However, data obtained in 2008 showed some isotopic effects similar to those found using the IDMS calibration data (Fig. 3). From a purely theoretical point of view, no isotopic effects should be detected using the minimal labelling procedure (the use of $^{13}C_1$ phenol as spike) and standard single collector MS instrumentation. Carbon isotopic fractionation effects during chemical analysis have been described to be, in general, not significant.[34] So, a molecule labelled with a single ^{13}C atom should behave exactly the same as the natural abundance analogue which is the basic assumption for the application of calibration-free isotope dilution analysis procedures.

Analysis of NIST 1584 certified reference material

The proposed multiple linear regression procedure was evaluated by comparing the results obtained for the determination of phenol in the Certified Reference Material NIST 1584 (Priority Pollutants in methanol) by GC-EI-MS using the classical isotope dilution calibration procedure and the new procedure based on multiple linear regression [applying eqn (10)]. Both $^{13}C_1$-labelled and $^{13}C_6$-labelled phenol were employed in both calibration strategies to study the influence of the number of ^{13}C atoms (isotope effects) on the final analyte concentration. Additionally, measurements were performed using only 2 masses, in order to improve counting statistics in the mass spectrometer, or the full 6 or 10 masses to improve the multiple linear regression statistics. For the IDMS calibration graph using minimal labelling, the experimental data were adjusted to a second order polynomial equation. The results obtained by all alternative procedures are shown in Table 4 for a triplicate analysis of the certified reference material. As can be observed, no significant differences between the average values can be detected and all procedures provide average values in agreement with the certified value. The main differences were found in the total combined uncertainties calculated for the different procedures as discussed below.

Calculation of uncertainty budgets and evaluation of uncertainty sources

In the application of a calibration graph in isotope dilution analysis the expected main sources of uncertainty are: the experimentally measured peak area ratios, the calibration graph itself, the concentration of the natural standard used and the uncertainty of the sample preparation procedure (weights taken, dilution factors and so on). However, when applying the alternative multiple linear regression procedure the sources of uncertainty are different. First, no calibration graph is used as

Table 4 Results obtained in the analysis of the certified reference material NIST 1584 applying multiple linear regression using the $^{13}C_1$-labelled or $^{13}C_6$-labelled phenol and the classical isotope dilution calibration strategy using both $^{13}C_1$-labelled or $^{13}C_6$-labelled phenol measuring 2 or 6 masses for $^{13}C_1$ phenol and 2 or 10 masses for $^{13}C_6$ phenol. The uncertainty for the independent replicates corresponds to the combined uncertainty calculated by the Kragten approach.[32] The uncertainty of the average values corresponds to the total combined uncertainty

		Multiple linear regression		Calibration	
	Sample (μg g^{-1})	Masses 94 and 95	Masses 93 to 98	Masses 94 and 95	Masses 93 to 98
Phenol $^{13}C_1$	NIST M1	36.3 ± 0.4	36.7 ± 0.6	36.9 ± 0.2	37.1 ± 0.5
	NIST M2	36.6 ± 0.4	37.0 ± 0.5	37.1 ± 0.2	37.4 ± 0.2
	NIST M3	36.6 ± 0.4	36.8 ± 0.5	37.1 ± 0.2	37.2 ± 0.2
	Average	**36.5 ± 0.4**	**36.8 ± 0.5**	**37.0 ± 0.2**	**37.2 ± 0.4**
		Multiple linear regression		Calibration	
	Sample (μg g^{-1})	Masses 94 and 100	Masses 93 to 102	Masses 94 and 100	Masses 93 to 102
Phenol $^{13}C_6$	NIST M1	36.6 ± 1.7	37.3 ± 1.7	36.9 ± 0.2	36.8 ± 0.2
	NIST M2	36.1 ± 1.7	36.9 ± 1.7	36.5 ± 0.2	36.4 ± 0.6
	NIST M3	36.3 ± 1.7	36.8 ± 1.7	36.6 ± 0.2	36.5 ± 0.4
	Average	**36.3 ± 1.7**	**37.0 ± 1.7**	**36.7 ± 0.3**	**36.6 ± 0.5**
	Certified value (μg g^{-1})	**37.5 ± 1.1**			

every single measurement provides the concentration of the analyte in the sample. However, in this case the uncertainty in the concentration of the labelled compound needs to be taken into account for the calculation of the uncertainty budget. In general, the concentration of the labelled compound will be calculated from reverse isotope dilution experiments and its final uncertainty will show a contribution of the uncertainty in the concentration of the natural abundance standard. The second important source of uncertainty in the multiple linear regression procedure is the isotope composition of both natural abundance and labelled compounds which are used to construct the matrix $\boldsymbol{A}$ in eqns (7) and (8). In our case, the uncertainties in the isotope composition of the natural and labelled compounds should take into account also the uncertainties in the fragmentation factors as shown in Table 2. For the IDMS calibration procedure the isotope composition of the natural and labelled compound is irrelevant. Finally, when more than two *m*/*z* values are used in the multiple linear regression procedure the uncertainty due to the multiple linear regression can be calculated using eqn (9) and propagated to the ratio of molar fractions using eqn (13). In summary, all possible uncertainty sources by both alternative procedures are given in Table 5.

The combined standard uncertainties in each replicate for the determination of phenol in CRM NIST 1584 were calculated by the spreadsheet procedure published by Kragten[32] and recommended by Eurachem.[33] Using this procedure it is possible to estimate the contribution of the uncertainty of each parameter (in %) to the overall total combined uncertainty for each analytical result. Full uncertainty budgets were calculated for all individual replicate analyses of CRM NIST 1584 using the $^{13}C_1$-labelled or $^{13}C_6$-labelled phenol by applying both multiple linear regression and the classical calibration isotope dilution procedure. The final uncertainty results are given in Table 4 while the typical contributions (in %) to the combined uncertainties are shown in Table 5. As can be observed in Table 4, the lowest propagated uncertainties correspond always to the calibration graphs calculated with 2 masses measured regardless of the labelling of the molecule. The measurement of 6 or 10 masses for the calibration graph increased the uncertainty by a factor of 2, on average. For the multiple linear regression procedure the use of the singly labelled molecule provided propagated uncertainties between 3 and 4 times lower than the use of the fully labelled compound which provided the worst propagated uncertainties. For the singly labelled phenol the use of 2 or 6 masses in the multiple linear regression calculations shows similar uncertainty values. Table 5 shows that the uncertainty in the multiple linear regression procedure arises mainly from the isotope composition of the labelled compounds with a significant contribution from the concentration of the isotopically labelled standard. So, for the reduction of these uncertainty sources a better certificate of the labelled standards both in isotope composition and concentration should be given by the manufacturers. This is standard practice in elemental isotope dilution analysis but it has never been required in organic IDMS. As can be observed from Table 5, the concentration and isotope composition of the labelled compounds are irrelevant for the calibration graph procedure. In these cases, the concentration of the natural abundance standards, the calibration graph and the experimental measurement of isotope ratios are the major contributors to the total uncertainty.

Basic steps in the implementation of the multiple linear regression procedure for other organic compounds

The general implementation of the multiple linear regression procedure for the analysis of organic compounds requires a series of initial studies before the procedure can be finally applied for a routine basis:

(1) *The calculation of the theoretical isotope composition of the target molecule (both natural abundance and labelled) and its uncertainty*. The procedure described elsewhere[28] can be used for this purpose. The best labelling of the molecule can be selected

Table 5 Typical contribution to the total combined uncertainty (in % of total uncertainty) from the different uncertainty sources considered in the analysis of CRM NIST 1584

Labelled compound	$^{13}C_1$-labelled phenol				$^{13}C_6$-labelled phenol			
Calculation procedure	Calibration		Multiple linear regression		Calibration		Multiple linear regression	
Number of masses measured	2	6	2	6	2	10	2	10
Sample preparation (weights, dilutions, *etc.*)	9–16	1–7	2–3	1–2	9–11	1–8	<1	<1
Concentration of natural abundance standard	39–64	5–31	—	—	34–44	5–28	—	—
Concentration of isotopically labelled standard	—	—	17–18	8–11	—	—	18	17–18
Experimental measurement of the peak area ratios	2–41	5–83	—	—	30–45	50–92	—	—
Experimental measurement of the isotope abundances	—	—	1–5	1–2	—	—	<1	1–5
Calibration graph	18–16	10–58	—	—	12–16	2–15	—	—
Isotope composition of natural abundance phenol	—	—	18–20	9–12	—	—	<1	<1
Isotope composition of labelled phenol	—	—	57–62	27–39	—	—	81	77–80
Uncertainty of the multiple linear regression	—	—	—	21–35	—	—	—	<1

based on the value of the $(\boldsymbol{A}^{\mathrm{T}} \times \boldsymbol{A})^{-1}$ matrix (Fig. 2). For small organic compounds a minimal labelling could be recommended.

(2) *The measurement of the fragmentation pattern of the standards in the ion source before mixing the natural and the labelled analogues*. Particular attention should be paid to the loss of hydrogen atoms and the correction of possible mass bias effects in the mass spectrometer. Additionally, the stability of the fragmentation pattern should be tested at different concentration levels and different ionization conditions. For this purpose, the measurement of at least 2 *m*/*z* is required.

(3) *The calculation of the 'estimated' isotope composition of the natural abundance and labelled molecule taking into account its fragmentation pattern (matrix* $\boldsymbol{A}$*)*. The uncertainty in the 'estimated' isotope composition should be calculated as well.

(4) *The determination of the concentration of the labelled molecule by reverse isotope dilution analysis using a natural abundance standard as reference*. When minimal labelling is selected and no isotopic effects are expected the same multiple linear regression procedure [eqn (10)] can be used here. For this purpose, the measurement of at least 2 *m*/*z* is required.

(5) *The validation of the procedure by the analysis of a reference material and the calculation of the total combined uncertainty*.

When points 1–5 have been carried out, the practical application of the procedure is very simple as every injection will provide directly the concentration of the analyte in the sample.

Conclusions

It has been demonstrated that minimal labelling in combination with multiple linear regression can be considered as a direct calculation procedure for the determination of organic compounds by IDMS. The procedure does not require a methodological calibration graph as the ratio of molar fractions measured is equal to the ratio of molar concentrations. From a practical point of view, the possibility of avoiding the need of a calibration curve in organic stable isotope dilution analysis reduces drastically the total sample preparation time and facilitates the applicability of IDMS for routine analysis. The only disadvantage of the proposed approach in comparison with the classical calibration strategies is the slightly higher associated uncertainty due to the influence of the uncertainties in the isotope distributions of the natural and labelled compounds. Nevertheless the RSD obtained using the proposed multiple linear regression approach in combination with minimal labelling did not exceed 1.6%. In addition, such uncertainty can be significantly reduced with an accurate certification of the isotopic enrichment of the labelled analogues. Current work in our laboratory is devoted to explore the performance of this methodology for the determination of other compounds of environmental, toxicological and biomedical interest.

Acknowledgements

The authors are grateful for financial support from the Spanish Ministry of Science and Innovation through project ref. CTQ2006-05722. A.G.-A. acknowledges her doctoral grant to FICYT (Asturias) and P.R.-G. acknowledges his research contract to the Spanish Ministry of Science and Innovation through the Juan de la Cierva Program.

References

1 N. G. Dodder, S. C. Tai, L. T. Sniegoski, N. F. Zhang and M. J. Welch, *Clin. Chem.*, 2007, **53**, 1694–1699.
2 K. G. Heumann, *Anal. Bioanal. Chem.*, 2004, **378**, 318–329.
3 M. Rychlik and S. Asam, *Anal. Bioanal. Chem.*, 2008, **390**, 617–628.
4 P. Rodríguez-González, J. M. Marchante-Gayón, J. I. García Alonso and A. Sanz-Medel, *Spectrochim. Acta, Part B*, 2005, **60**, 151–207.
5 A. P. De Leenheer and M. Thienpont, *Mass Spectrom. Rev.*, 1992, **11**, 249–307.
6 S. A. Gerber, J. Rush, O. Stemman, M. W. Kirschner and S. P. Gygi, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 6940–6945.
7 D. G. Burke and L. D. Mackay, *Anal. Chem.*, 2008, **80**, 5071–5078.
8 N. Itoh, M. Numata and T. Yarita, *Anal. Sci.*, 2007, **23**, 1245–1248.
9 N. Itoh, M. Numata, Y. Aoyagi and T. Yarita, *J. Chromatogr., A*, 2007, **1138**, 26–31.
10 J. Wieling, *Chromatographia*, 2002, **55**, S107–S113.
11 E. Stokvis, H. Rosing and J. H. Beijnen, *Rapid Commun. Mass Spectrom.*, 2005, **19**, 401–407.
12 J. A. Jonckheere, P. De Leenheer and H. L. Steyaert, *Anal. Chem.*, 1983, **55**, 153–155.
13 J. F. Sabot and H. Pinatel, *Analyst*, 1993, **118**, 831–834.
14 L. M. Thienpont, B. Van Nieuwenhove, D. Stöckl and A. P. De Leenheer, *J. Mass Spectrom.*, 1996, **31**, 1119–1125.
15 J. B. Lambert, H. F. Shurvell, D. Lightner and R. G. Cooks, in *Introduction to Organic Spectroscopy*, Macmillan Publishing Company, New York, 1987, p. 353.
16 J. Meija and J. A. Caruso, *J. Am. Soc. Mass Spectrom.*, 2004, **15**, 654–658.
17 J. Meija, G. Centineo, J. I. García Alonso, A. Sanz-Medel and J. A. Caruso, *J. Mass Spectrom.*, 2005, **40**, 807–814.
18 J. A. Rodríguez-Castrillón, M. Moldovan, J. R. Encinar and J. I. García Alonso, *J. Anal. At. Spectrom.*, 2008, **23**, 318–324.
19 J. A. Rodríguez-Castrillón, M. Moldovan and J. I. García Alonso, *Anal. Bioanal. Chem.*, 2009, **394**, 351–362.
20 J. Meija, L. Yang, J. A. Caruso and Z. Mester, *J. Anal. At. Spectrom.*, 2006, **21**, 1294–1297.
21 P. Rodriguez-Gonzalez, M. Monperrus, J. I. Garcia Alonso, D. Amouroux and O. F. X. Donard, *J. Anal. At. Spectrom.*, 2007, **22**, 1373–1382.
22 J. Meija and Z. Mester, *Anal. Chim. Acta*, 2008, **607**, 115–125.
23 J. Meija, L. Ouerdane and Z. Mester, *Analyst*, 2009, **134**, 466–471.
24 P. Rodríguez-González, A. Rodríguez-Cea, J. I. García Alonso and A. Sanz-Medel, *Anal. Chem.*, 2005, **77**, 7724–7734.
25 J. A. Rodríguez-Castrillón, M. Moldovan, J. I. García Alonso, J. J. Lucena, M. L. García-Tomé and L. Hernández-Apalaza, *Anal. Bioanal. Chem.*, 2008, **390**, 579–590.
26 H. G. Iglesias, M. L. Fernández, J. I. García Alonso and A. Sanz-Medel, *Anal. Bioanal. Chem.*, 2007, **389**, 707–713.
27 C. L. Rojas, F. J. Romera, E. Alcántara, R. Pérez-Vicente, C. Sariego, J. I. García Alonso, J. Boned and G. Marti, *J. Agric. Food Chem.*, 2008, **56**, 10774–10778.
28 J. I. García Alonso, P. Rodríguez-González, A. González-Gago and A. Gonzalez-Antuña, *Anal. Chim. Acta*, 2010, DOI: 10.1016/j.aca.2010.01.063.
29 M. C. B. Moraes, J. G. A. Brito Neto, V. F. Juliano and C. K. L. do Lago, *Int. J. Mass Spectrom.*, 1998, **178**, 129–141.
30 J. Meija and Z. Mester, *Metrologia*, 2008, **45**, 459–463.
31 H. Kubinyi, *Anal. Chim. Acta*, 1991, **247**, 107–119.
32 J. Kragten, *Analyst*, 1994, **119**, 2161–2165.
33 *Quantifying Uncertainty in Analytical Measurements*, EURACHEM/CITAC Guide, 2nd edn, 2000.
34 T. C. Schmidt, L. Zwank, M. Elsner, M. Berg, R. U. Meckenstock and S. B. Haderlein, *Anal. Bioanal. Chem.*, 2004, **378**, 283–300.