Analysis of labelled compounds by mass spectroscopy in the presence of heavy isotopes at natural abundance: the NAIC program

K. Naveed D. Barakzai,[a] Brendan J. Howlin,[a] William J. S. Lockley,[a*] and Konstantinos Sfyrakis[b]

A program is described, which allows the intensity contributions from naturally occurring heavy isotopes to be eliminated from the mass spectrometric ions of an artificially labelled compound. This allows the accurate measurement of the isotopomer distribution of the compound and hence its percentage isotopic abundance. In the case of radioactively labelled compounds the specific radioactivity can then be derived from the abundance data. The program utilizes a graphical user interface based upon the C^{++} Builder 6.0 environment and VCL library.

Introduction

A number of utilities are available, which enable the calculation of the relative intensities of the peaks in a mass spectrometric molecular ion cluster from the molecular formula of the analyte. Some of these are integrated into dedicated MS software control and analysis packages.[1] However, there are also excellent open-access web-based and Windows®-based programs available.[2a–g] Not all of these packages allow the inclusion of isotopes. Some, however, can be utilized for the theoretical prediction [3a,b] of the mass spectrometric ion intensities of compounds labelled at 100% atom abundance with one or more atoms of a heavy isotope, while one excellent application [2f] has recently been modified to allow for customized isotopic abundance data. Nevertheless, for most isotopic chemists these packages have limited application since synthetic products only rarely contain isotopic atoms at 100% abundance. Indeed, most isotopically labelled products are mixtures of isotopomers containing varying numbers of isotopic atoms, making the analysis of their MS ions a far from trivial task.

This paper describes a program, the Natural Abundance Isotope Correction program (NAIC), designed to address the MS analysis of labelled compounds. An early version of the NAIC program,[4] written in BASIC has been in use since the 1980s. Subsequently, the program was ported to a more efficient programming language, significantly refined, extended, validated and distributed for the use in several isotopic synthesis laboratories. The modified program runs under the Microsoft Windows® operating system and is written in C^{++}. The graphical user interface (GUI) was created using the C^{++} Builder 6.0 environment and the VCL library.[5] Data input and output is via the GUI with an option to export results to the Excel Microsoft Office® spreadsheet.

The deconvolution of isotopically labelled MS ions has been addressed in the past [6,7] though the procedures available at the time were time-consuming and somewhat laborious owing to the limited computing resources available. Recently an alternative spreadsheet-based and open-source program, IsoPat2, was described and utilized for the deconvolution of the MS ions from isotopomer mixtures generated in the microwave deuteration of ketones.[8] This matrix algebra approach to the deconvolution operation, which includes a least-squares error minimization procedure, differs from the approach of NAIC, which utilizes the more conventional algebraic subtraction process. Both approaches, however, are simply alternative ways to derive solutions to a set of linear simultaneous equations arising from an analysis of the contribution from each isotopomer present in a mixture to each MS peak in the observed ion cluster (Figure 1). In our hands, comparisons of the two programs for a wide variety of partially labelled MS ion cluster types have yielded identical results (within rounding errors) for both synthetic and real MS ion data. This has

[a]Division of Chemical Sciences, Faculty of Health and Medical Sciences, University of Surrey, Guildford, GU2 7XH, UK

[b]Blue Brain Project, Ecole Polytechnique Fédérale de Lausanne, Lausanne, CH-1015, Switzerland

*Correspondence to: William J S Lockley, Division of Chemical Sciences, Faculty of Health and Medical Sciences, University of Surrey, Guildford, GU2 7XH, UK. E-mail: w.lockley@surrey.ac.uk
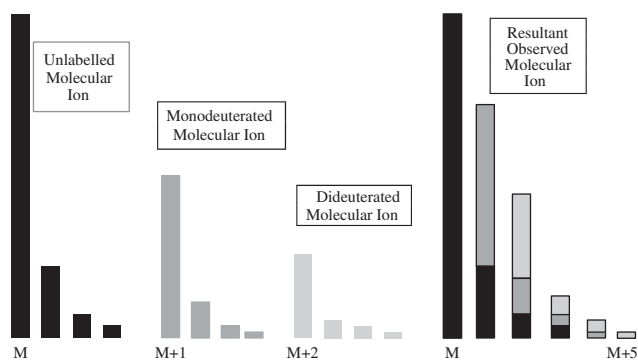
Figure 1. Contribution of isotopomeric forms to the observed molecular ion of a labelled compound consisting of a 4:2:1 mixture of unlabelled, monodeuterated and dideuterated isotopomers.

Figure 2. Simplified operation of the NAIC program.

prompted us to report here the detailed operation of the NAIC program for the first time.

Discussion

The mass spectrum of a mixture of isotopically labelled compounds consists of ions arising from all the isotopomers present. The intensities of the observed peaks in a particular ion cluster arise from a superimposition of ions arising from each isotopomer, paying due regard to the change in mass and also to any differences in volatility, ionization, fragmentation, etc, which might be occasioned by the isotopic substitution. A simplified example of this process is given by the molecular ion of a hypothetical mixture of deuterated isotopomers shown in Figure 1.

In this case, the observed mass spectrometric molecular ion is considered as a simple summation of the appropriate ion intensities. This situation is indeed typical of the molecular ions of most labelled compounds, as in the majority of cases isotopic effects on volatilization, ionization, fragmentation, etc, are rarely of a primary bond-breaking nature and hence may be ignored, at least as a first approximation. In other cases, the experimental MS conditions (e.g. the ionization voltage) or data collection regime (e.g. integration of all scans for GC-MS or HPLC-MS peaks) may need to be adjusted to minimize any isotope effects observed.

Operation of the NAIC program

The program has a number of functionalities as shown in Figure 2. Initially, data on the intensities of the peak cluster constituting the MS ion of interest are either input or calculated. As shown above, these peak intensities arise from contributions from naturally abundance isotopes of the elements present in the ion. Hence, the intensities of the various peaks in the cluster can be calculated from the molecular formula of the ion and the known natural abundances of the appropriate isotopes. NAIC can carry out this calculation for simple molecules containing C, H, O, N, P, I and F, however more comprehensive web-based facilities [2a–f] are freely available, which provide for more complex molecular species. Data from these systems can be easily incorporated during the use of NAIC. However, it should be stressed that direct experimental measurement of the intensities of the ions in the unlabelled material is likely to give more reliable analyses since the only assumption made in this
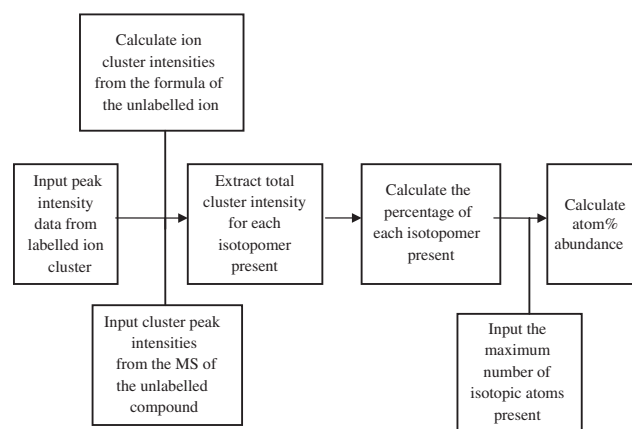
case is of the similar behaviour of the labelled and unlabelled compound under the optimized MS conditions. Once the data from the unlabelled compound and the corresponding data for the labelled compound have been calculated, or input, the program then addresses the determination of the contribution of each isotopomeric component in the labelled compound to the ion cluster intensities.

The analysis of an ion such as the one in Figure 1 may be approached in a number of ways but the simplest method (and the one underlying the procedure used by the NAIC program) is by constructing a series of simultaneous equations, which, taken together, express the contributions of the ion clusters from each isotopomer to the overall intensities of the peaks in the MS ion cluster.

As an example of the process, the intensity, $I_{(M+2)}$ of the peak at $M+2$ in the 'Resultant observed molecular ion' cluster of Figure 1 has contributions from several isotopomers, as shown. The overall intensity of the peak is constituted from the $M+2$ peak of the unlabelled species plus the $M+1$ peak of the monodeutero-species plus the M peak of the dideutero-species. This can be expressed as

$$I_{(M+2)} = a \times M_{D0} + b \times M_{D1} + M_{D2}$$

where M_{D0}, M_{D1} and M_{D2} are the intensities of the nominal mass peaks of the unlabelled, monodeuterated and dideuterated species, and where a and b are the $M+2/M$ and the $M+1/M$ ratios of intensities for the unlabelled compound, respectively. Similar relationships between the observed intensity of a peak and the contributions from the various isotopomers to that peak can be defined for each one of the other peaks in the cluster. The resulting relationships can then be utilized via an iterative incremental subtraction procedure,[6,7] to compute the total intensity, across all masses, arising from each isotopomer. As the quantities of each of the isotopomers present will be proportional to the combined intensities of all the peaks associated with each isotopomer, the ion cluster has now been deconvoluted and the isotopic mixture analysed.

Having extracted and quantified the total peak intensities for each isotopomer present in the labelled compound ion, the calculation of the percentage contribution of each isotopomer to the overall ion cluster merely requires the expression of the result as a percentage of the total ion cluster intensity. The resultant data are then easily converted to the atom% abundance of the isotope by a summation of the isotopic

560

www.jlcr.org

J. Label Compd. Radiopharm **2009**, 52 559–562

contributions to the total, while noting that isotopomers with increasing numbers of isotopic atoms will contribute to a correspondingly increased degree to the overall abundance, such that:

$$\text{Atom\% abundance, } A\% = 100 \times \sum_{i=0}^{i=n}(A_i . i/n)$$

Where A_i is the fractional abundance of each isotopomer, i is the number of isotopic atoms present in the isotopomer and n is the maximum number of possible isotopic replacements in the molecule.

In the case of the molecular ion of a radioactively labelled compound, this abundance value can then be converted to specific radioactivity (S) provided that the molar specific activity (M) of the radioactive isotope at 100% abundance is known, as it is for most common nuclides, via the simple relation: $S = A\% \times M$.

The interface between users and the NAIC program is shown in Figure 3. It is composed of various input and output objects. Initially two choices must be made using the Choices Dialog form (Bottom right). First, since the isotopic pattern of the unlabelled compound may be calculated either from the molecular formula or by entering the experimental intensities from the MS, a decision on these options is required from the user. Second, the type of the isotope in the labelled compound must be specified as either differing from the light isotope by one mass unit, for example, ^{2}H, ^{13}C, ^{15}N, ^{17}O or differing by two mass units, for example, ^{3}H, ^{14}C, ^{18}O.

If the isotopic pattern of the unlabelled compound is to be determined from the molecular formula, then the formula is selected in the Data Entry form (Bottom left) that allows the input of the number of atoms of each element present. The elements supported in the current release are C, H, N, O, P, I, F. If, instead, the isotopic pattern of the unlabelled compound is to be provided from the experimental MS or from the output of a web-based public domain calculator,[2a–f] then the data on the intensities of the peaks in the ion cluster of the unlabelled compound are entered in the Data Entry form under the heading of 'Molecular ion intensities'. The peak intensities from the corresponding ion cluster of the labelled compound are also input in the same Data Entry form, this time under the 'Intensities from labelled spectra' heading. The program will accept the input of floating point numbers for both the 'Molecular ion intensities' and for the 'Intensities from labelled spectra' data. Finally, there is an option to specify the total number of label atoms, which is to be used when calculating the atom% abundance.

When data entry is complete the 'Calculate' button is selected. The results are then output in the Results window (Top). The output comprises; the natural abundance isotope corrected MS ion data, a graphical display of the corrected spectrum, the calculated atom% abundance and the molecular weight (if using the molecular formula option). An option to export the output data to an excel sheet is also provided.
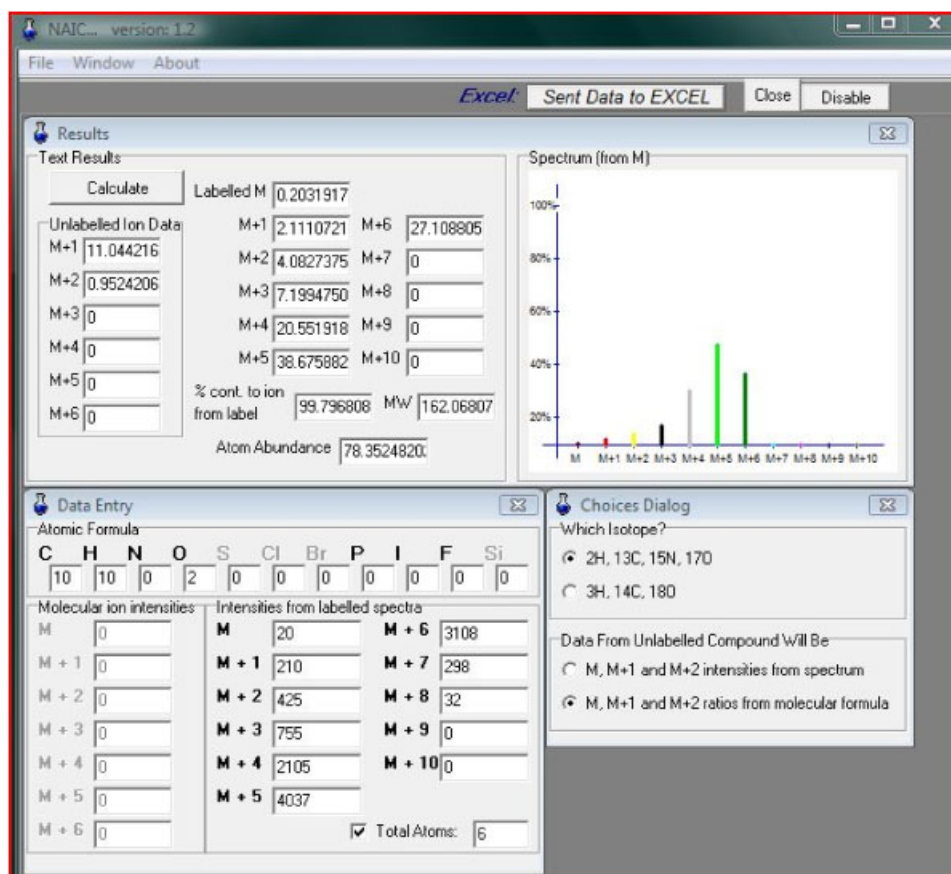

Figure 3. Graphical user interface of NAIC.

J. Label Compd. Radiopharm **2009**, 52 559–562

Copyright © 2009 John Wiley & Sons, Ltd.

www.jlcr.org

561

Further information and program limitations

Provision of extensive programming data and detailed instructions for use of the program would be inappropriate for the Journal. Hence, further technical information on the NAIC program together with the examples of the use of the program have been lodged at an open-access website.[9]

A number of limitations of the program should be noted. First, the unlabelled ion data entry is limited to the values from M to M+6 while the labelled ion data are limited to the values from M to M+10. These limits will enable the analysis of the overwhelming bulk of labelled organic compounds. However, it is worth noting that the IsoPat2 spreadsheet[8] embraces a larger range of values. Second, the algorithm for the calculation of the natural abundance MS ion intensities is only approximate, though quite appropriate for the analysis of small molecules ($<C_{50}$ or so). For high precision calculations relating to larger molecules, a more accurate algorithm such as those available in the web-based public domain calculators [2]a–f should be utilized and the resulting natural abundance isotope pattern then entered into NAIC.

The NAIC program is available free of charge from w.lockley@ surrey.ac.uk. Though efforts will be made to correct any errors or bugs brought to the attention of the authors, the use of the program and any errors occasioned by its use are entirely the responsibility of the user.

Acknowledgement

References

[1] For a typical example see: MassLynx Version 4.0 User's Guide, pp. 457–458 via http://www.waters.com [22 June 2009].

[2] For open-access isotope pattern calculators see: (a) Sheffield ChemPuter isotope pattern calculator, M Winter: http://winter. group.shef.ac.uk/chemputer/ [22 June 2009]; (b) Molecular weight calculator, M. Monroe: http://www.alchemistmatt.com/mwtwin. html. Available at http://ncrr.pnl.gov/software/ [22 June 2009]; (c) IsoPro3, M. Senko: via: http://www.oc.uni-koeln.de/ms-oc/ download/isopro3.zip [22 June 2009]; (d) Molecular Weight and Isotopic Abundance calculator. W. Zheng: http://www.ifa.hawaii. edu/~zhengwj/Applets/IsoPattern.html [27th Jan 2009]; (e) Isotope cluster. http://www.colby.edu/chemistry/NMR/IsoClus.html. archive [22 June 2009]; (f) Isotope Pattern Calculator v4.5, Junhua Yan: http://www.geocities.com/junhuayan/pattern1.htm [22 June 2009]; (g) For a review of web-based and windows-based calculators see: K. Massila, R. D. Stein, S. M. Suhaizan, A. A. Azlianor, *Proc. World Acad. Sci. Eng. Tech.* **2007**, *20*, 146–149. ISSN 1307-6884 (h).

[3] For theoretical approaches see: (a) C. S. Hsu, *Anal. Chem.* **1984**, *56*, 1356–1361 and (b) J. A. Yergey, *Int. J. Mass Spec. Ion Phys.* **1983**, *52*, 337–349.

[4] W. J. S. Lockley, *Appl. Radiat. Isotop.* **1986**, *I-II*, 37.

[5] Codegear at http://www.embarcadero.com/ [22 June 2009].

[6] M. Hesse, H. Meier, B. Zeeh, *Spectroscopic Methods in Organic-Chemistry*, Thieme, New York, **1997**, pp. 256–267.

[7] H. Budzikiewicz, *Massenspektrometrie*, Wiley-VCH, Weinheim, Germany, **1998**, pp. 70–71.

[8] C. C. Gruber, G. Oberdorfer, C. V. Voss, J. M. Kremsner, C. O. Kappe, W. Kroutil, *J. Org. Chem.* **2007**, *72*, 5778–5783.

[9] http://www.surrey.ac.uk/Chemistry/NaicLMSP/ [23 June 2009].