Determination of High-Precision Isotope Ratios from Experimental Isotopic Distributions

Parminder Kaur,1,2,3 Cheng Zhao,2,3 and Peter B. O’Connor1,2,3

1Department of Electrical and Computer Engineering, Boston University
2Cardiovascular Proteomics Center, Boston University School of Medicine
3Mass Spectrometry Resource, Department of Biochemistry, Boston University School of Medicine.

Introduction  Isotope variability due to natural processes provides important information for studying a variety of complex phenomenon like determining the genesis of a given sample, dietary studies of species, nitrification rates in trees etc. These measurements require very high-precision determination of isotope ratios of a particular element involved. Isotope Ratio Mass Spectrometers (IRMS) are widely employed tools for such a high-precision analysis. IRMS instruments accept the sample analyte in the form of only a limited number of gases which must represent the isotopic characteristics of the original sample, which causes lack of flexibility. This work aims at overcoming the limitations inherent to IRMS by estimating the elemental isotopic abundance from the experimental isotope distribution.

Theory  Experimental isotopic distribution is a measure of the isotopic abundances of various elements present in the molecule. It can be represented by the joint convolution of the isotopic abundances of each of the individual atoms. Mathematical techniques have been developed in order to factor out the known isotopic abundance contributions from various elements followed by calculation of the corresponding unknown values for the element of interest. An estimate for the required isotopic abundance is generated from each of the observed isotopic peaks by solving the convolution equations as follows. Let the elemental composition of a given molecule be CnHmNpOpSp. Let

\[
P_C = [1 - p_c \text{ } p_c]
\]

\[
P_H = [1 - p_h \text{ } p_h]
\]

\[
P_N = [1 - p_n \text{ } p_n]
\]

be the isotopic abundances of C, H, N, O and S respectively, where \(P_X(i)\) represents the \(i^{th}\) isotope of element \(X\). Let \(T[n]\) represent the Theoretical Isotopic Distribution (TID)[2]

\[
T[n] = P_C[n] \otimes N \sum_{p}^N P_C[n] \otimes P_H[n] \otimes P_N[n] \otimes P_O[n] \otimes P_O[n] \otimes P_O[n] \otimes P_O[n] \otimes P_S[n] \otimes P_S[n] \otimes P_S[n]
\]

where * denotes convolution operator and \(\otimes\) denotes multiple convolutions defined as follows:

\[
z[n] = x[n] * y[n] = \sum_{j=1}^{\infty} x[j] \text{ } y[n + 1 - j]
\]

\[
x[n] \otimes x[n] = x[n] * x[n] * x[n]...N \text{ times}
\]

Assume that the isotopic abundances of all the elements except, say, C are given and \(p_c\) \((^{13}C\text{ abundance})\) is the unknown value to be determined, then all the terms on the RHS in equation 1 are known except for the \(P_C\) term which depends upon the \(p_c\) value. In the limit of sufficiently high number of ions, the Experimental Isotopic Distribution (EID) approaches the Theoretical Isotopic Distribution (TID)[1], and hence, can be used to substitute \(T[n]\) on the LHS in equation 1. Then the task of estimating \(p_c\) reduces to solving each of the above polynomial equations, with each equation representing one of the isotopes observed in the EID.

Results  Computer generated simulations were carried out in order to generate the experimental isotopic distributions for a given molecule with known elemental composition and isotopic abundances. The results thus obtained were subjected to the developed theoretical framework in order to estimate the isotopic abundance for Carbon from each isotopic peak, with the abundance values for the other elements being taken into consideration in the calculation. These estimated results are shown to be in good agreement with their true values. Increasing the number of ions for generating the experimental isotopic distribution greatly improves the estimate (Fig 1a). This is because in the limit of infinite number of ions, experimental isotopic distribution approaches its theoretical
counterpart and is a true representative of the composition of its constituents. High molecular weight molecules are shown to be particularly advantageous because of the presence of larger number of isotopic peaks in the isotopic distribution leading to a greater amount of information (Fig 1b). Initial results reveal that with sufficiently high number of ions and multiple experiments, it is possible to distinguish between the samples varying very slightly in the Carbon isotopic abundance. For example, the estimate can distinguish whether the sample originated from marine plankton (C13 abundance = 1.09%, Delta C13=-19.5) or meat from an animal feeding on a C4 plant (C13 abundance=1.1%, Delta C13=-12.5). Experiments were carried out with bovine ubiquitin spectra for calculating values. Analysis of 26 spectra from 392 different isotopic peakse revealed the estimated value of $\delta$ being -27.55 ($p_{\delta_{true}}=1.08\%)$, indicating that sample fed primarily on C3 plants (Fig 2).

Fig 1(a) Estimate improves with increase in number of ions used to generate the simulated isotopic distribution, $\delta_{true}=-25.5$, MW=9000, (b) Estimate improves with increase in MW due to higher number of isotopes present in the isotopic distribution, $\delta_{true}=-25.5$, $\delta_{Est}=-25.6802$. ($p_{\delta_{true}}=0.010832, p_{\delta_{Est}}=0.010830$)

Fig 2 (a) Mass Spectrum of bovine ubiquitin (b) EID (blue) matches well with the TID (red) (c) Delta estimate using 26 spectra of bovine ubiquitin, median value=-27.55

This approach eliminates the need to convert the sample into gaseous form. The results are applicable to any type of mass spectrometer providing isotopic resolution, and for any type of sample. The results can also be extended to estimate the isotopic abundance of any unknown element provided the isotopic abundance of the other elements are known a priori. Overlapping isotopic distributions are difficult to handle with this approach. For optimal results, this method requires minimal artifacts in the experimental isotopic distribution.

Acknowledgements This work was supported in part by Federal funds from the National Center for Research Resources under grant No. P41-RR10888 and the National Heart, Lung, and Blood Institute under NO1HV28178. We are thankful to Catherine E. Costello, Amit Juneja, Alan Rockwood, Konstantin Aizikov, Raman Mathur, Jason J Cournoyer, Cheng Lin, and Vera Ivleva for their valuable support.

References

