

SOFTWARE

Open Access

# Software LS-MIDA for efficient mass isotopomer distribution analysis in metabolic modelling

Zeeshan Ahmed<sup>1,7†</sup>, Saman Zeeshan<sup>1,8†</sup>, Claudia Huber<sup>5</sup>, Michael Hensel<sup>2</sup>, Dietmar Schomburg<sup>3</sup>, Richard Münch<sup>4</sup>, Wolfgang Eisenreich<sup>5</sup> and Thomas Dandekar<sup>1,6\*</sup>

## Abstract

**Background:** The knowledge of metabolic pathways and fluxes is important to understand the adaptation of organisms to their biotic and abiotic environment. The specific distribution of stable isotope labelled precursors into metabolic products can be taken as fingerprints of the metabolic events and dynamics through the metabolic networks. An open-source software is required that easily and rapidly calculates from mass spectra of labelled metabolites, derivatives and their fragments global isotope excess and isotopomer distribution.

**Results:** The open-source software “Least Square Mass Isotopomer Analyzer” (LS-MIDA) is presented that processes experimental mass spectrometry (MS) data on the basis of metabolite information such as the number of atoms in the compound, mass to charge ratio ( $m/e$  or  $m/z$ ) values of the compounds and fragments under study, and the experimental relative MS intensities reflecting the enrichments of isotopomers in  $^{13}\text{C}$ - or  $^{15}\text{N}$ -labelled compounds, in comparison to the natural abundances in the unlabelled molecules. The software uses Brauman’s least square method of linear regression. As a result, global isotope enrichments of the metabolite or fragment under study and the molar abundances of each isotopomer are obtained and displayed.

**Conclusions:** The new software provides an open-source platform that easily and rapidly converts experimental MS patterns of labelled metabolites into isotopomer enrichments that are the basis for subsequent observation-driven analysis of pathways and fluxes, as well as for model-driven metabolic flux calculations.

## Background

Metabolism is central for all cellular processes including adaptation of organisms to their respective life style and conditions. Triggered by the presence and activity of metabolic enzymes and the metabolite fluxes through pathways, cellular reactions constitute a highly dynamic network that can be rapidly and efficiently modulated in response to environmental changes. A number of theoretical techniques has been established to predict metabolic fluxes [1-4]. Implementing different mathematical parallel and sequential algorithms, several desktop and web based batch and interactive software applications [5] have been also developed towards quantitative metabolic flux analysis and modeling [6].

In contrast, only few methods allow direct determination of metabolic fluxes, one of which is based on *in vivo* experiments using stable isotope labelled precursors, such as  $^{13}\text{C}$ -glucose or  $^{13}\text{CO}_2$ . The transfer of label to the metabolic network and the specific isotope distribution in metabolic products can then be taken as evidence of metabolic pathways and fluxes during the experimental period. However, robust technology is required to quantitatively determine the isotopomer abundances in multiple metabolites. Specifically, experimental intensities of mass signals (typically of silylated derivatives, metabolites and fragments thereof in GC/MS experiments) have to be converted into relative and molar isotopomer abundances.

**Isotopologues** are species of a compound that differ only in their isotopic composition [7]. The term **isotopomer** is a contraction of ‘isotopic isomer’, grouping isotopologues into those molecules which contain the same number of a specific isotope (e.g.  $^{13}\text{C}$ ) at different positions. As an

\* Correspondence: dandekar@biozentrum.uni-wuerzburg.de

†Equal contributors

<sup>1</sup>Department of Bioinformatics, Biocenter, University of Würzburg, Würzburg, Germany

<sup>6</sup>EMBL, Structural and Computational Biology Unit, Heidelberg, Germany

Full list of author information is available at the end of the article

example, 64 stable carbon isotopologues exist for glucose. Out of these, six species constitute isotopomers with one  $^{13}\text{C}$ -atom at position 1, 2, 3, 4, 5, or 6. In natural compounds, i.e. obtained from the natural environment, the **natural abundance** is the consequence of the natural isotope abundance (i.e. ca. 1.1% for  $^{13}\text{C}$  at a given carbon position) that is diverted through the complete population of isotopologues due to statistical reasons. In contrast, increased isotopologue abundances are observed in labeling experiments where isotope-enriched precursors (e.g.  $^{13}\text{C}$ -labelled) are supplied to the organism under study. This results in the enrichment of specific isotopologues, i.e. on top of the natural abundances, in the metabolic products. The deconvolution of mass intensities yielding **isotopomer enrichment** is the key task of the software described in this manuscript. Notably, mass intensities provide information on the abundances of isotopologues harbouring a specific number of the isotope, i.e. one, two, three, etc.  $^{13}\text{C}$ -atoms, and therefore, the enrichment of isotopomeric groups (**isotopomer distribution**) is obtained. Since metabolic pathways lead to specific isotopomer enrichments and, as a consequence, to specific isotopomer distributions, the latter values can be used to identify and to quantify the relative contributions of metabolic routes from the labelled precursor to the products observed by MS.

So far, three different methods are available for positional isotopomer determination, nuclear magnetic resonance (NMR), mass spectrometric analysis of a sufficient number of useful metabolite fragments [8,9], and multiple reaction monitoring (MRM). These methods can provide orthogonal information and can be combined using our software to improve positional isotopomer determination [10].

In this manuscript, a new freely available software is described that is capable of providing a user friendly graphical interface for the efficient and independent (no third party application is needed) data storage, management and processing towards mass isotopomer distribution analysis [11]. The implemented software enables the user to load data from previously created data files or add data manually into the software application at run time and to process it. Furthermore, it directly parameterizes input experimental data to Brauman's algorithm for accurate estimation of natural and relative abundances. No such application exists, implementing similar mathematics into a user friendly software package.

Currently, only commercial software or user-specific approaches are available for the conversion of mass intensities (provided by the specific software implemented to the mass spectrometer) to the relative and molar isotopomer enrichments, such as tandem mass spectrometric data computing for positional

isotopomer distributions [12], measurements of mass distributions by mass spectrometry [13], isotopomer analysis using GC-MS [10,14], and GC-MS analysis for isotopomer balancing [15].

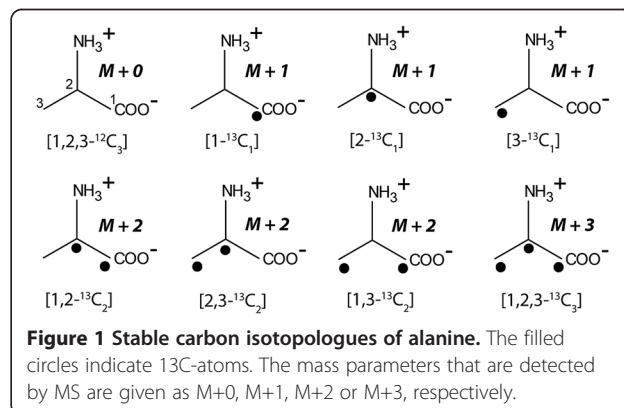
However, for a broad range of users, a open-source software compatible to data exchange with the standard mass software packages is highly desirable. Here, we present a new open-source software using Brauman's least square method for the calculation of isotopomer enrichment that can be used in GC/MS and LC/MS experiments (including tandem-MS/MS) by calculating relative and absolute isotopomer abundances from the mass ratios of signals in experimental MS spectra.

## Implementation

### Algorithm

The software treats experimental raw data from MS. Specifically, MS intensities of metabolic products (typically  $^{13}\text{C}$ - or  $^{15}\text{N}$ -labelled) are analysed on the basis of their  $m/e$  values and the number of C or N atoms, respectively, in the given molecule, derivative or fragment thereof. Overall  $^{13}\text{C}$ - or  $^{15}\text{N}$ -enrichment and the relative and molar contribution of isotopomers are then calculated using Braumann's least squares algorithm [16]. Mass distribution measured by MS display enrichments of isotopomer groups (i.e. isotopologues with a given number of the label (Y or 1 in our notation), for example one, two, three etc.  $^{13}\text{C}$ -atoms). For example, the isotopomer distribution of the  $\text{C}_3$ -compound alanine (see Figure 1) is calculated from the abundances of the unlabelled compound (e.g. 000 for three  $^{12}\text{C}$ -atoms), of the isotopomer group containing one label (00Y, where Y can be at any carbon position), two labels (0YY, the  $^{12}\text{C}$ -atom 0 can be again at any position in alanine), and three labels 111.

The software is composed of two parts: (i) generation of an appropriate set of linear simultaneous equations



and (ii) solution of these equations. The proposed matrix equation to calculate isotopomer enrichments is the least square method using the Moore–Penrose pseudo inverse (the latter is a powerful mathematical method for matrix calculations, nothing specific to isotopomers or metabolism):

$$X = (A^T * A)^{-1} A^T P \quad (1)$$

where  $X$  = calculated relative intensities of the fragments,  $A$  = matrix of relative natural abundance values for all possible isotopomers and  $P$  = set of the experimental relative intensities of the fragments (observed during experimentation). The used mathematics follows Brauman's approach [16] and recommended considerations for its application [14,17]. The number of  $C$  atoms in the fragments contributes to relative natural abundance distributions. It is nevertheless worth while to look for contaminating fragment ions in the mass spectrum, as its impact is substantially corrected through the subtraction of relative natural abundance values. The proposed binomial expression [14] calculating isotopomer fragment distribution taking into account relative natural abundance is

$$A = n! / [(i!) * (n-i)!] * P_0^{(n-i)} * P_1^i \quad (2)$$

where  $A$  = relative natural abundance,  $n$  = number of carbon atoms,  $i$  = index variable to count  $n$  iterations.  $P_0$  and  $P_1$  stand for the abundance of  $^{12}C$  and  $^{13}C$ , respectively.

"The solution of these equations gives the abundance of each organic moiety. Because of the way in which the problem was formulated, the total abundance of the organic moieties must remain constant." This statement by Brauman [16] has now to be put into practical calculations. However, the results of this technique depend upon a number of factors: (i) the analysis is based on the assumption that the fragmentation patterns for all heteroatom isotopes are identical (i.e., no differential isotope effect), (ii) the experimental relative abundance of  $^{12}C$  and  $^{13}C$  isotopes induced through derivatization is known and (iii) the relative natural abundances of the isotopes are either known or measured. The LS-MIDA software package itself is not designed to perform integration of the original MS signals. Thus, another software package must be used first. For our examples, it relied on pre-processing by the software LabSolutions by Shimadzu which is standard software directly supplied with the instrumentation for GC-MS. However, any type of pre-processing software can be used in combination with LS-MIDA.

To predict the relative isotopomer contributions in natural abundance compounds, linear regression analysis is performed by drawing an abundance matrix (eq. 3)

using the known or estimated natural abundance values of isotopes by binomial expression.

$$A = \begin{pmatrix} A_1 & 0 & 0 & 0 \\ A_2 & A_1 & 0 & 0 \\ A_3 & A_2 & A_1 & 0 \\ A_4 & A_3 & A_2 & A_1 \\ \dots & \dots & \dots & \dots \\ A_n & A_{n-1} & A_{n-2} & A_{n-3} \end{pmatrix} \quad (3)$$

Here,  $A_1, A_2, A_3, A_4 \dots A_n$  are the estimated relative natural abundance values of fragments using eq. 2. These values are used in the abundance matrix  $A$  [17,18], based for linear regression analysis. In general, obtaining the coefficients of matrix  $A$  (eq. 3) is the non-trivial part of the method. More specifically, the atoms and their isotope distributions from the derivatisation agent must be taken into account, if present. Overlaps of mass traces by impurities have obvious effects and the mode of ionization (positive or negative) has effects on the  $m/z$  values. Our implementation tackles the processing of the pre-processed data to overcome some of these problems. For example, contributions due to the derivatisation agent are filtered out.

Next, Brauman's least square algorithm (eq. 4) is applied for the estimation of relative intensity values for the fragment spectrum:

$$Ri_{(1-n)} = A^{-1} * P \quad (4)$$

Here,  $Ri_{(1-n)}$  are the string of predicted relative intensity values with respect to the  $m/e$  values. The length of the abundance matrix depends upon the total number of  $m/e$  measurements and experimental relative intensity values. The set of linear equations used to draw the abundance matrix and multiplications for quantitative analysis [14] is given in eq. 5.

$$\begin{aligned} S_0 T_0 &= U_0 \\ S_1 T_0 + S_0 T_1 &= U_1 \\ S_2 T_0 + S_1 T_1 + S_0 T_2 &= U_2 \\ \dots & \\ \dots & \\ S_n T_{m-1} + S_{n-1} T_m &= U_{(n+m-1)} \\ S_n T_m &= U_{(n+m)} \end{aligned} \quad (5)$$

Here, the linear regression analysis (initially used by Brauman [16]) is performed for spectral data analysis, where  $U$  is the mass isotopomer distribution, and  $S$  and  $T$  are the isotope abundances for  $^{12}C$  and  $^{13}C$ , respectively.

To compute isotopomer abundances for each fragment, again linear regression analysis is performed by calculating the abundance matrix, but the input values are now the observed relative intensity values ( $Ri_{(1-n)}$ ) in

the MS-traces. The length of the abundance matrix depends on the number of fragments with the result:

$$Abundance\ Matrix\ (Ri_{(1-n)}) = \begin{pmatrix} Ri_1 & 0 & 0 & 0 \\ Ri_2 & Ri_1 & 0 & 0 \\ Ri_3 & Ri_2 & Ri_1 & 0 \\ Ri_4 & Ri_3 & Ri_2 & Ri_1 \\ \dots & \dots & \dots & \dots \\ Ri_n & Ri_{n-1} & Ri_{n-2} & Ri_{n-3} \end{pmatrix} \quad (6)$$

In eq. 6,  $Ri_1, Ri_2, Ri_3, Ri_4, \dots, Ri_n$  are the observed relative intensity values with respect to the m/e values. Now, eq. 1 is applied to estimate the string of calculated relative abundance values ( $Ra_{(1-n)}$ ) for the fragments, where  $A =$  Abundance Matrix ( $Ri_{(1-n)}$ ).

With eq. 7 absolute  $^{13}C$  enrichments are then calculated:

$$Abs^{13}C = (\sum A_{0...n} * n) / a \quad (7)$$

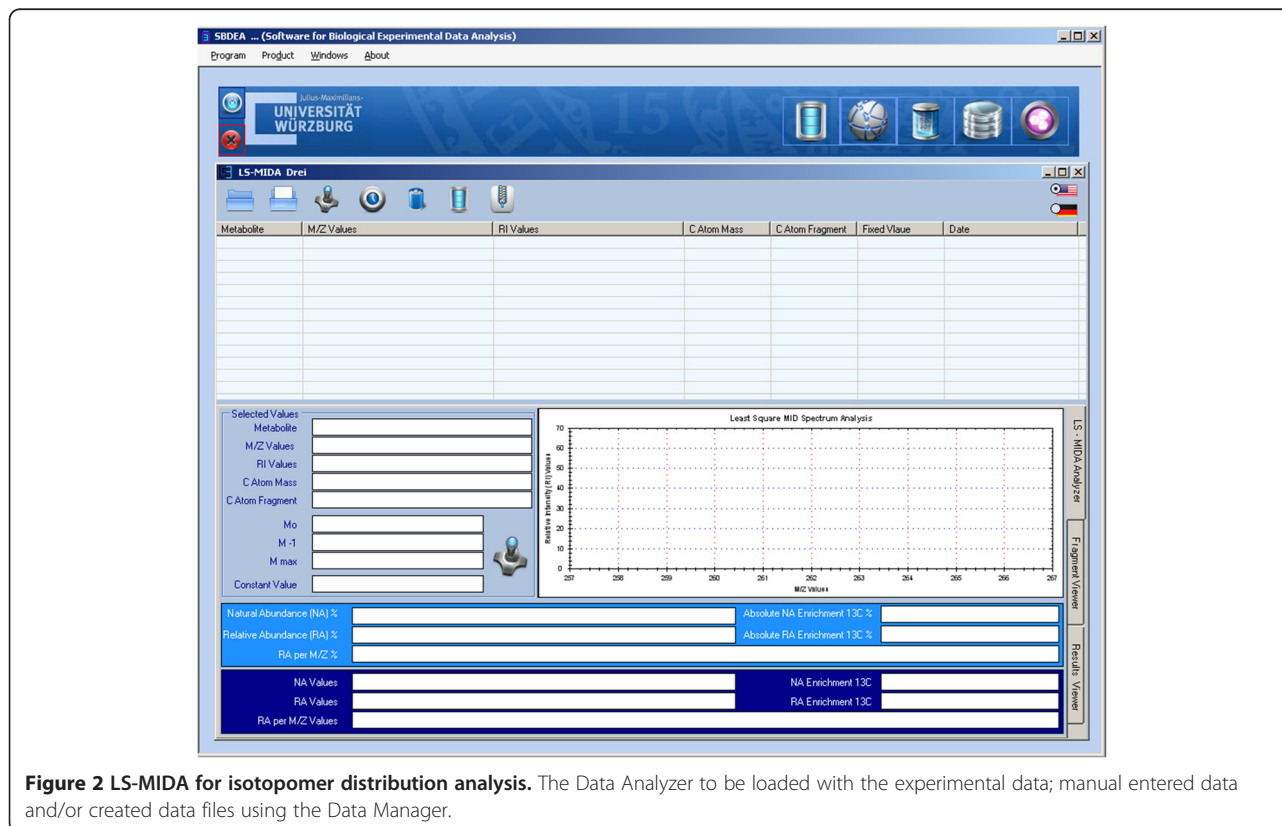
The absolute  $^{13}C$ -enrichment is equal to the sum of all labeled isotopomers multiplied with the respective number of labels (0 to n), divided by the number of carbon atoms in the fragment under study. A indicates the labeled isotopomer, the index 0 till n indicates the number of labeled atoms in the fragment, and a indicates the number of carbon atoms in the fragment. The matrix

calculations are mathematical simple, however, the optimized combination of pre-filtering software and abundance calculations by LS-MIDA takes into account all required steps as well as experimental complications (e.g. filtering out derivatization agent) in a single user-friendly and open-source software package.

## Development

The above described mathematics and calculations were implemented into the software "LS-MIDA" (executable available as Additional file 1; test data in Additional file 2; pre-filtering software considerations see above). We show that the implementation of Brauman's least square method and the inclusion of binomial expression allow accurate calculations of isotopomer enrichments using experimental GC/MS data of  $^{13}C$ -labelled silylated amino acids. LS-MIDA is a UML designed [19] and successfully evaluated third party tool independent reusable desktop application (batch) with user friendly graphical interface, capable of sequentially processing standard input and producing visual output presentation (text and spectrum).

The available and tested version of LS-MIDA provides two main modules, the data analyzer (see Figure 2) and the data manager (see Additional file 3, installation and technical overview). The data analyzer is capable of processing input data (metabolite information, m/e values

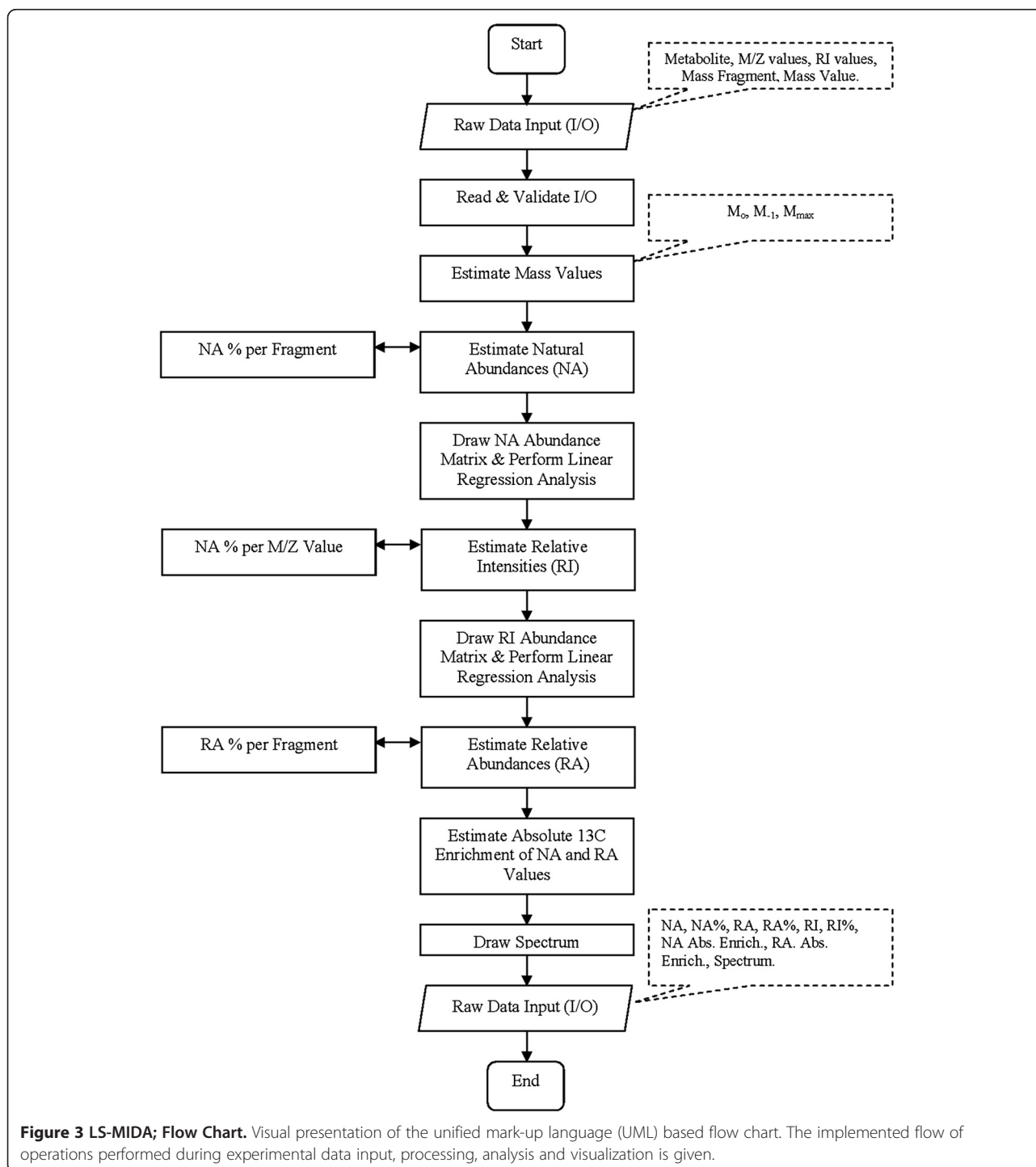


**Figure 2 LS-MIDA for isotopomer distribution analysis.** The Data Analyzer to be loaded with the experimental data; manual entered data and/or created data files using the Data Manager.

and experimental relative MS intensity values). It then estimates mass values ( $M_o$ ,  $M_{-1}$ ,  $M_{max}$ ), predicts relative natural abundance values, and calculates the actual isotopomer abundances from the MS patterns. Finally, it allows drawing the isotopomer distribution of the calculated values (see Figure 3).

### Graphical user interface

LS-MIDA provides an intuitive graphical user interface for file based experimental raw data manipulation and management. It is capable of not only managing user input experimental raw data but also provides options to manage resultant data (output of LS-MIDA). It allows



**Figure 3 LS-MIDA; Flow Chart.** Visual presentation of the unified mark-up language (UML) based flow chart. The implemented flow of operations performed during experimental data input, processing, analysis and visualization is given.

the user to create new data files, manage created data files, merge new or already made data files into one or more new or already created data files and manipulate entries of data files. It is an independent file based data management system that does not require any external or third party database to install and use. LS-MIDA is implemented and tested using Microsoft Windows (version XP and 7) operating system as it is developed using Microsoft Dot net framework and C Sharp (object oriented) programming language. The licensed software is freely available for academic use on request.

#### Database manager

LS-MIDA advantageously provides a file-based data management system for experimental metabolic mass isotopomers based raw data. The data manager is a supporting utility, developed as a user-friendly file-based experimental data management system. It allows the user to create new experimental data files that later can be used for the analysis using data analyzer. The experimental data is organised following a new data format especially proposed (with extension "\*.ls") for LS-MIDA data files. Data manager allows the user to read, add, edit, update, delete and merge data (from other source files of the same extension) into a file.

### Results

#### Calculations

The implemented mathematical procedure in LS-MIDA version 3.0 (see Figure 4) starts with the input (I/O) of metabolite information (e.g., name) and of the raw data from MS spectra, i.e., m/e values, and the experimental relative intensity ( $R_i$ ) values. After I/O validation, the mass values  $M_0$ ,  $M_{-1}$ ,  $M_{max}$  are calculated to adjust the potential mass distribution.

Then, using binomial expansion, relative natural abundances  $N_a$  with percentages per each fragment are estimated (see Figure 5). Next, linear regression analysis is performed and the abundance matrix is drawn with the application of Brauman's least square method using estimated  $N_a$  values to derive the relative intensity values per m/e value for natural abundance compounds. The relative intensity values  $R_i$  are then used to calculate the isotopomer abundances  $R_a$  (including their percentage amounts in each fragment) for the labelled compounds under study.

For this purpose, once again a linear regression analysis is performed drawing the abundance matrix with the implementation of Brauman's least square method. Using the calculated  $N_a$  and  $R_a$  values, absolute  $^{13}\text{C}$  enrichment is then calculated for each fragment (see Figure 6).

The output ( $N_a$  and  $R_a$  values) is presented in numeric format and in special notation format (based on the number of C atoms in the fragments) and the isotopomer distribution is shown graphically (see Figure 7).

#### Application in metabolite measurements

The software was now tested with different data sets.  $^{13}\text{C}$ -Labelled amino acid samples (analyzed as TBDMS-derivatives) were obtained from hydrolysates of *Salmonella enterica* grown in medium containing  $[\text{U-}^{13}\text{C}_6]\text{glucose}$  [20]. We have shown earlier that  $[\text{U-}^{13}\text{C}_6]\text{glucose}$  is efficiently incorporated into most amino acids of *Salmonella enterica* via intermediates of glycolysis. Under these conditions, alanine is mainly composed of the unlabelled isotopomer (derived from unlabelled glucose in the medium) and the fully  $^{13}\text{C}$ -labelled isotopomer due to *de*

Selected Values	
Metabolite	Ala - 260
M/Z Values	259,15#260,1#261,1#262,1#263,1#264,1#265,1#
RI Values	0,07#8,53#2,3#2,39#44,59#8,79#3,88#
C Atom Mass	3#
C Atom Fragment	3#
$M_0$	259 #
$M_{-1}$	258 #
$M_{max}$	278 #
Constant Value	16#

**Figure 4 LS-MIDA; Input Interface.** This figure presents input information (example: alanine; Metabolite name, m/e values,  $R_i$  values, C Atom Values, C Atom fragment values) and estimated mass values ( $M_0$ ,  $M_{-1}$ ,  $M_{max}$ ) for Ala 260. The input and resultant values are also presented in Table 1.

Natural Abundance (NA) %	96.7068262369% 3.2564842893% 0.0365527107% 0.0001367631%	Absolute NA Enrichment 13C %	1.11 %
Relative Abundance (RA) %	70.0008468870352% 0% 0% 29.9991531129648%	Absolute RA Enrichment 13C %	29.9991531129648 %
RA per M/Z %	0.490919298819507% 11.8009464807093% 3.28643028850377% 5.33928863323921% 61.5344908095128% 12.2289817521541% 5.31894273706134%		
NA Values	0.967068262369# 0.032564842893# 0.000365527107# 1.367631E-06#	NA Enrichment 13C	0.0111 #
RA Values	1.01811623727288# 0# 0# 0.436317933953475#	RA Enrichment 13C	0.436317933953475 #
RA per M/Z Values	0.323882817095428# 7.78564582764497# 2.16821441444225# 3.52258272994665# 40.5972311128173# 8.06804106015085# 3.50915957427469#		

**Figure 5 LS-MIDA; Calculation.** Estimated natural abundance, relative abundance, relative intensity values and absolute enrichments for alanine, fragment weight is 260.

*novo* synthesis of alanine from [U-<sup>13</sup>C<sub>3</sub>]pyruvate made from [U-<sup>13</sup>C<sub>6</sub>]glucose via glycolysis.

As an example for the involved numerical steps, Table 1 and Figure 4 show input parameters and experimental MS raw data of three fragments for labelled TBDMS-alanine such as metabolite information, m/e values of the relevant fragment, experimental intensity values, atomic mass values and the number of atoms in the fragment. During input, data file preparation, and management, the data manager structures data into experimental data files which are then used by data analyzer for the calculations. The obtained results are shown in Table 1b, Table 1c, and Figures 4, 5 and 6. Observed results are  $M_o$ ,  $M_{-1}$ ,  $M_{max}$  (the range of mass in which to pick the correct intensities), the predicted relative natural abundances, and the relative abundances of isotopomers in the labelled sample with its absolute <sup>13</sup>C enrichment

As shown in Figure 7, abundances of four different alanine isotopomers are observed. For the isotopomer 000 (i.e. <sup>12</sup>C only), the estimated natural abundance is 96.7%, whereas the same isotopomer accounts for 70.0% in the labelled compound. The isotopomer group with one <sup>13</sup>C-atom has 3.26% or 0% abundance in the unlabelled or labelled compound, respectively. The relative abundances for the isotopomer group comprising two <sup>13</sup>C-atoms are 0.037% or 0% in the unlabelled or labelled alanine sample, respectively, whereas the abundances for the fully labelled isotopomer are 0.00014% or 30.0% in the unlabelled or labelled sample, respectively. On this basis, the observed absolute enrichment value of <sup>13</sup>C in the labelled sample from *S. enterica* results in 30.0%.

This is in line with our expectations and calculations also using other software tools for isotopomer analysis.

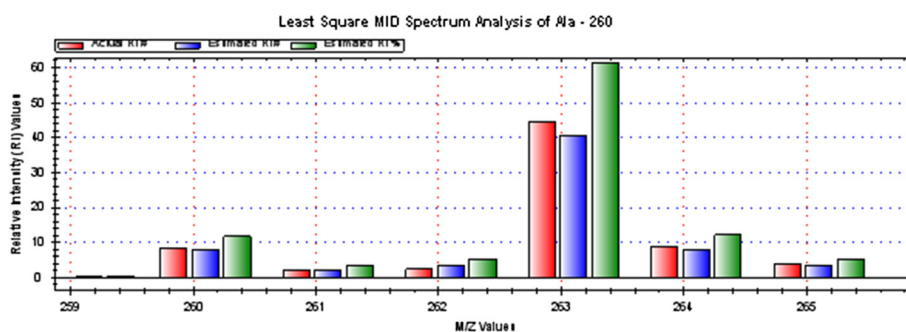
The resulting spectrum is shown in Figure 7. At constant m/e values the peaks of the drawn spectrum may vary according to their molecular composition [16]. The strongest observed relative intensity <sup>13</sup>C isotopomer peak in this example is at 263.1 in the correct range of  $M_o$ ,  $M_{-1}$ ,  $M_{max}$  so the values for 260.1, 263.1 and 264.1 are shown in Figure 7 for the alanine mass spectrum. For more results of labelled TBDMS-amino acids from the same labelling experiment [20,21], please have a look at Additional file 3: Table S2-S3.

For the pathway analyses of Salmonellae mentioned above and in similar studies on other bacteria, we tested the processing of different data sets; input data file preparation and management, experimentation and data analysis. This allowed also an analysis of error rates. Accuracy of the calculation itself is quite high (error less than 1 part per billion). This estimate is based on several hundred test runs of the software and meticulous testing for bugs and unexpected behavior. Regarding errors from the non-trivial coefficients of matrix A involving experimental errors from atoms of the derivatization agent present and overlap of fragment spectra as well as the mode of ionization (positive or negative), we expect error rates of less than 1%.

Another inherent source of error is the analysis of network fluxes for complex biological system that typically result in notably higher deviations when looking at the results from replicates. As these are indirectly inferred from the changes of measured isotopomer concentrations, these data already carry the error due to the

Metabolite	Groups	Natural Abundance Percentage	Relative Abundance Percentage	Natural Abundance Value	Relative Abundance Value
Ala - 260 259 #	[000]	96.7068262369 %	70.0008468870352 %	0.967068262369 #	1.01811623727288 #
	[xxx]1	3.2564842893 %	0 %	0.032564842893 #	0 #
	[xxx]2	0.0365527107 %	0 %	0.000365527107 #	0 #
	[111]	0.0001367631 %	29.9991531129648 %	1.367631E-06 #	0.436317933953475 #

**Figure 6 LS-MIDA; Abundance.** Calculated percentages of abundance values for Alanine in fragment groups.



**Figure 7 LS-MIDA; Spectrum.** Alanine Mass Spectrum ( $R_i$  to  $m/e$ ). Spectrum is drawn and shows estimated relative intensity values, based on the input experimental intensity values with respect to the  $m/e$  values. All data can be processed at once (click the button "Measure all the data").

mathematical procedures. In addition, experimental errors regarding growth conditions must be taken into account (cell number, actual glucose concentration, sample preparation). Indeed, the resulting overall error for flux analyses considering all these effects can be as high as 10% based on our experience. Some limitations arise from the fact that not in all pathways metabolized label is present, but this is an inherent prediction limitation for the approach.

## Discussion

$^{13}\text{C}$  Labelling of metabolites has proven to be a powerful method in quantifying metabolic routes and fluxes, especially, if there are alternative pathways between two metabolites. Isotopomer balancing provides the basis for deducing metabolic pathways and fluxes.

In own studies, we have shown that  $^{13}\text{C}$ -incorporation studies coupled to software-based isotopomer calculation allowed us to identify the pathways of amino acid biosynthesis [22] under relevant non-standardized conditions (e.g. proline biosynthesis in *Listeria monocytogenes* and its modulation by the transcription factor PrfA [21]). Another example concerns nutrient supply for *Salmonella* in the *Salmonella* containing vacuole [20]. Again, the unequivocal determination of nutrient flow across the vacuole to *Salmonella* depends on isotopologue data with the use of suitable processing software.

In general, without transforming the mathematics into applied software doing the calculations, none of these and other insights [6,20,21] regarding flux modifications and usage of different metabolic pathways in different organisms is possible.

Despite this potential power of the methodological approach, it is still difficult to perform metabolic flux analyses due to the lack of user-friendly and open-source software tools. This limitation also demands the development of mathematical modelling of metabolism for each substrate to obtain more detailed and accurate results. Before the implementation of LS-MIDA, we relied on the usage of a lab-specific Excel/Solver-based

software doing the required calculations. However, this approach did not provide user-friendly output formats nor included a database allowing extensive comparative studies. As an alternative to this lab-specific solution, commercial packages may also be used [12,13,15]. However, these software packages are not freely available.

In order to establish tools that can be widely distributed, we have established the LS-MIDA software. Briefly, Brauman's least square algorithm is used and developed in the form of a versatile software application iteratively analyzing the estimated abundance resonances [23] after binomial expansion for the calculation of isotopomer enrichments in labelled metabolites.

Furthermore, LS-MIDA provides a file-based data management system for fast and accurate MS-based isotopomer analyses.

In comparison with other existing approaches [24-26], the combined features in LS-MIDA are not available in standard packages for metabolite modelling such as Metatool [18], Yanasquare [27], Gepasi [28] or FiatFlux [1] (here fluxes are predicted after the isotopologue data have been processed). There are two software solutions available for isotopomer data processing, Envelop [24] and Isotope Pattern Calculator [25], but none uses binomial expression for data extension. The implementation of Brauman's least square method with the inclusion of binomial expression allows rapid and accurate calculation of isotopomer data.

LS-MIDA was compared with in-house software demonstrating its robustness. This showed that LS-MIDA can be used as an open-source platform for many (even non-expert) users in consortia in research programs such as the ongoing priority DFG program dealing with "host adapted metabolism of pathogens" in our example as well as other interested academic groups (German and foreign) for which we did provide and also will provide in future the software free of charge. Training courses and service are offered to support the use of the software within the program and for other users. In consequence, the software is freely available for the worldwide academic community.



**Table 1 Calculation results**

**(A) Experimental raw data of alanine (Ala)**

Metabolite	m/e values	R <sub>i</sub> values	C Atom Metabolite	C Atom Fragment
Alanine (Ala)	259.15#260.1#261.1#262.1# 263.1#264.1#265.1#	0.07#8.53#2.3#2.39# 44.59#8.79#3.88#	3	3

**(B) Alanine experimental raw data analysis<sup>1</sup>**

Metabolite	M <sub>0</sub>	M <sub>-1</sub>	M <sub>max</sub>	N <sub>a</sub>	R <sub>a</sub>	N <sub>a</sub> Abs.	R <sub>a</sub> Abs.	R <sub>i</sub> values
						<b>Enrichment</b>	<b>Enrichment</b>	
Alanine (Ala)	259	258	278	0.967068262369#	1.01811623727288#	0.0111 #	0.436317933953475 #	0.323882817095428#
				0.032564842893#	0#			7.78564582764497#
				0.000365527107#	0#			2.16821441444225#
				1.367631E-06#	0.436317933953475#			3.52258272994665#
								40.5972311128173#
								8.06804106015085#
								3.50915957427469#

**(C) Isotopomer calculation results (example: alanine)**

Isotopomeric group	N <sub>a</sub> %	R <sub>a</sub> %
[000]	96.7068262369%	70.0008468870352%
[XXX] 1	3.2564842893%	0%
[XXX] 2	0.0365527107%	0%
[111]	0.0001367631%	29.9991531129648%

<sup>1</sup> Here and in the following long (many digits) calculation results are shown to illustrate arithmetics.

## Conclusions

LS-MIDA is a versatile, open-source, and user friendly software with rapid calculation, integrated overview on all isotopomers, least square correction and database management system, with good application potential for biology and biotechnology such as studying the core metabolism of organisms, the pathways and fluxes leading to desired products in biotechnology, and complementing methods from genomics, proteomics or metabolomics.

## Availability and requirements

LS-MIDA is free available software for all academic users with open license; a commercial license can be obtained on request.

LS-MIDA is developed using the Microsoft C# (sharp) programming language and Microsoft Dot Net framework 2008. It is compatible (install and use) for all Microsoft Windows operating systems. Moreover, LS-MIDA automatically adopts the language of the installed operating system and presents numerical values accordingly e.g. in case of English language decimal values are '.' (dot) separated and in case of German language decimal values are ',' (comma) separated. Numerical values are separated by '#' (hash) symbol for all languages.

Further details are available in the Additional file 3 Tutorial (installation, evaluation, further data, glossary).

## Additional files

**Additional file 1: LS-MIDA software application.** An executable software file (setup) is included.

**Additional file 2: Example data for different amino acids.** Example data for different amino acids analyzed by LS-MIDA.

**Additional file 3: Tutorial, installation, evaluation, further data, glossary.**

## Competing interests

The authors declare no competing interest (no reimbursements, fees, funding, or salary from an organization that may in any way gain or lose financially from the publication of this manuscript, either now or in the future, stocks or non-financial competing interests (political, personal, religious, ideological, academic, intellectual, commercial or any other).

## Authors' contributions

ZA: software designing, programming and testing. SZ assisted ZA. DS, RM, TD provided software expertise. CH, WE, MH provided experimental expertise. TD and WE lead and guided the study. All authors (ZA, SZ, CH, MH, DS, RM, WE, TD) participated in data analysis and evaluation of the software as well as in writing of the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

We would like to thank Dr. U. Rapp-Galmiche for native speaker corrections and the Deutsche Forschungsgemeinschaft (grant Da 208/10-2, TR34/Z1 and El-384/5-1) for funding. This publication was funded by the German Research Foundation (DFG) and the University of Würzburg in the funding programme Open Access Publishing.

## Author details

<sup>1</sup>Department of Bioinformatics, Biocenter, University of Würzburg, Würzburg, Germany. <sup>2</sup>Department of Microbiology, University of Osnabrück, Osnabrück, Germany. <sup>3</sup>Department of Bioinformatics and Biochemistry, Technical University Braunschweig, Braunschweig, Germany. <sup>4</sup>Institute for Microbiology, Technical University Braunschweig, Braunschweig, Germany. <sup>5</sup>Lehrstuhl für Biochemie, Technische Universität München, München, Germany. <sup>6</sup>EMBL, Structural and Computational Biology Unit, Heidelberg, Germany. <sup>7</sup>Department of Neurobiology and Genetics, Biocenter, University of Würzburg, Würzburg, Germany. <sup>8</sup>Institute of Molecular and Translational Therapeutic Strategies, Hannover Medical School, Hanover, Germany.

Received: 28 May 2012 Accepted: 23 June 2013

Published: 9 July 2013

## References

1. Nicola Z, Eliane F, Uwe S: **FiatFlux – a software for metabolic flux analysis from <sup>13</sup>C-glucose experiments.** *BMC Bioinforma* 2005, **6**:209.
2. Dauner M, Bailey J, Sauer U: **Metabolic flux analysis with a comprehensive isotopomer model in *Bacillus subtilis*.** *Biotechnol Bioeng* 2001, **76**:144–156.
3. Wiechert W: **<sup>13</sup>C metabolic flux analysis.** *Metab Eng* 2001, **3**:195–206.
4. Rantanen A, Rousu J, Jouhten P, Zamboni N, Maaheimo H, Ukkonen E: **An analytic and systematic framework for estimating metabolic flux ratios from <sup>13</sup>C tracer experiments.** *BMC Bioinforma* 2008, **9**:266.
5. Ahmed Z, Majeed S, Dandekar T: **Computational Feature Performance and DSA Evaluation of Applications towards MFA.** *Recent Patents Comput Sci* 2012, **5**:3.
6. Dandekar T, Fieselmann A, Majeed S, Ahmed Z: **Software Applications toward Quantitative Metabolic Flux Analysis and Visualization,** Briefings in Bioinformatics, First online published. ; 2012. November 9.
7. Brenninkmeijer CA, Janssen C, Kaiser J, Röckmann T, Rhee TS, Assonov SS: **Isotope effects in the chemistry of atmospheric trace compounds.** *Chem Rev* 2003, **103**:5125–5161.
8. Rousu JA, Rantanen RA, Ketola C, Juha T, Kokkonen C: **Isotopomer distribution computation from tandem mass spectrometric data with overlapping fragment spectra.** *Spectroscopy* 2005, **19**:53–67.
9. Bequette BJ, Sunny NE, El-Kadi SW, Owens SL: **Application of stable isotopes and mass isotopomer distribution analysis to the study of intermediary metabolism of nutrients.** *J Anim Sci* 2006, **84**:E50–E59.
10. Christensen B, Nielsen J: **Isotopomer analysis using GC-MS.** *Metab Eng* 1999, **1**:E8–E16.
11. Previs SF, Fernandez CA, Yang D, Soloviev MV, France D, Brunengraber H: **Limitations of the Mass Isotopomer Distribution Analysis of Glucose to Study Gluconeogenesis.** *J Biol Chem* 1998, **277**:16853–16859.
12. Rantanen A, Rousu J, Ketola RA, Kokkonen JT, Tarkiainen V: **Computing positional isotopomer distributions from tandem mass spectrometric data.** *Metab Eng* 2002, **4**:285–294.
13. Winden WW, Wittman C, Heinze E, Heijnen J: **Correcting mass isotopomer distributions for naturally occurring isotopes.** *Biotechnol Bioeng* 2002, **80**:477–479.
14. Lee WN, Byerley LO, Bergner EA, Edmond J: **Mass isotopomer analysis: theoretical and practical considerations.** *Biol Mass Spectrom* 1991, **20**:451–458.
15. Dauner M, Sauer U: **GC-MS analysis of amino acids rapidly provides rich information for isotopomer balancing.** *Biotechnol Prog* 2000, **16**:642–649.
16. Brauman JI: **Least Squares Analysis and Simplification of Multi-Isotope Mass Spectra.** *Anal Chem* 1966, **38**:607–610.
17. Korzekwa K, Howald WN, Trager WF: **The use of Brauman's least squares approach for the quantification of deuterated chlorophenols.** *Biomed Environ Mass Spectrom* 1990, **19**:211–217.
18. Pfeiffer T, Sanchez Valdenebro I, Nuno JC, Montero F, Schuster S: **METATOOL: for studying metabolic networks.** *Bioinformatics* 1999, **15**:251–257.
19. Ahmed Z, Majeed S, Dandekar T: **Unified Modeling and HCI Mockup Designing towards MIDA.** *IJES* 2012, **2**:361–382.
20. Dandekar T, Fieselmann A, Popp J, Hensel M: **Salmonella enterica: a surprisingly well-adapted intracellular lifestyle.** *Front Microbiol* 2012, **3**:164.
21. Eisenreich W, Dandekar T, Heesemann J, Goebel W: **Carbon metabolism of intracellular bacterial pathogens and possible links to virulence.** *Nat Rev Microbiol* 2010, **8**:401–412.
22. Papageorgopoulos C, Caldwell K, Shackleton C, Schweingrubber H, Hellerstein MK: **Measuring Protein Synthesis by Mass Isotopomer Distribution Analysis (MIDA).** *Anal Biochem* 1999, **267**:1–16.

23. Baverel G, Conjard A, Chauvin MF, Vercoutere B, Vittorelli A, Dubourg L, Gauthier C, Michoudet C, Durozard D, Martin G: **Carbon 13 NMR spectroscopy: a powerful tool for studying renal metabolism.** *Biochimie* 2003, **85**:863–871.
24. Sykes MT, Williamson JR: **Envelope: interactive software for modeling and fitting complex isotope distributions.** *BMC Bioinforma* 2008, **9**:446.
25. Massila K, Soong HC, Azlianor AA, Muhammad SS: **Reinforcing the concept of calculating isotope pattern using theoretical isotope generator (TIG).** *WSEAS Trans Inf Sci Appl* 2008, **5**:949.
26. Jennings ME, Matthews DE: **Determination of complex isotopomer patterns in isotopically labeled compounds by mass spectrometry.** *Anal Chem* 2005, **77**:6435–6444.
27. Schwarz R, Liang C, Kaleta C, Kuhnel M, Hoffmann E, Kuznetsov S, Hecker M, Griffith G, Schuster S, Dandekar T: **Integrated network reconstruction, visualization and analysis using YANAsquare.** *BMC Bioinforma* 2007, **8**:313. 10 pp.
28. Mendes P: **GEPASI: A software package for modelling the dynamics, steady states and control of biochemical and other systems.** *CABIOS* 1993, **9**:563–571.

doi:10.1186/1471-2105-14-218

**Cite this article as:** Ahmed *et al.*: Software LS-MIDA for efficient mass isotopomer distribution analysis in metabolic modelling. *BMC Bioinformatics* 2013 **14**:218.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

