MetabQ

Report date: 10 October 2014
Issued: Sergey Tumanov & Morgan Han
Template: Morgan Han

1. Introduction:

MetabQ is a package in R software created for automated data processing of raw GC-MS data files performing data extraction and calculation of absolute metabolite values. The package processes NetCDF files employing AMDIS mass spectral libraries and reports, generating data spreadsheet with absolute concentration values of metabolites. Functions of *MetabQ* package provide high flexibility in correcting parameters of analysis and graphical representation of ion chromatograms.

Developed by Yuri Zubenko, Vladimir Obolonkin and Sergey Tumanov.

Maintenance: Yuri Zubenko (8dark8@gmail.com), Sergey Tumanov (stum447@aucklanduni.ac.nz)

Contents

2. Installations

2.1. Software to be installed before using *MetabQ*

1. AMDIS
2. R software and packages (*xcms, tcltk2, scatterplot3d*)
3. ChemStation (Agilent) or
4. Xcalibur (Thermo Fisher Scientific)

2.1.1. How to install AMDIS

a. Download AMDIS software from http://chemdata.nist.gov/mass-spc/amdis/downloads/
b. Click "**AMDIS32_V2.71.exe**".
c. AMDIS will install automatically.

Note: You can download the latest version of AMDIS from http://chemdata.nist.gov/mass-spc/amdis/downloads/

2.1.2. How to install R

a. Download R software from http://cran.r-project.org/bin/windows/base/
b. Click "**R- 3.1.1-win.exe**".
c. R will install automatically.

Note, MetabQ_1.0 was built under R 3.1.0, but you can download the latest version from http://www.r-project.org/

2.2. How to install packages required for *MetabQ* in R

a. Open R software
b. To install *xcms* package type the following code in console

```
source("http://bioconductor.org/biocLite.R")
biocLite("xcms")
```

c. To install *tcltk2* package, on the menu bar, click "**Packages**" → select "**Install package(s)…**" → choose the country from a pop-up window (e.g. New Zealand), click OK and find *tcltk2* package in the list.
d. To install *scatterplot3d* package please follow the procedure as for *tcltk2*
e. To check if you have all the required packages please type the code in console:

```
library (xcms) (press "Enter")

library (tcltk) (press "Enter")

library(scatterplot3d) (press "Enter")
```

If no ERROR messages appear – proceed with *MetabQ* installation

2.3. How to install *MetabQ* in R

a. Install *MetabQ* by typing the following codes in the R console. Please **don't forget** to provide the path to the file

```
install.packages("C:/Users/...../MetabQ_1.0.tar.gz",type="source",report=NULL)
```

b. To check if *MetabQ* packages was installed please type the code in console:

```
library (MetabQ) (press "Enter")
```

3. Preparing documents to run *MetabQ*

3.1. The following documents are required to run *MetabQ*

1) AMDIS batch report(s) (in a text file)
2) CDF files organising in their conditional folders
3) MetabQ.settings.csv file

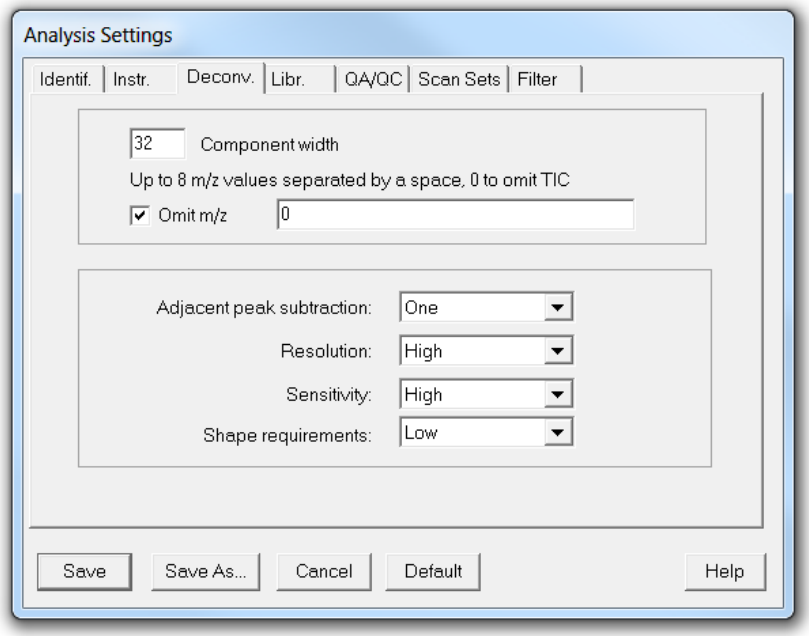
> *Note: All the above documents are located in a single folder*

3.1.1. How to create an AMDIS batch report

3.1.1.1. Setup AMDIS analysis settings

a. On the menu bar, click "**Analyse**" → select "**Settings…**" → under "**Identif.**" tab bar set "**75**" for Minimum match factor and **remove tick** for "**Multiple identifications per compound**".

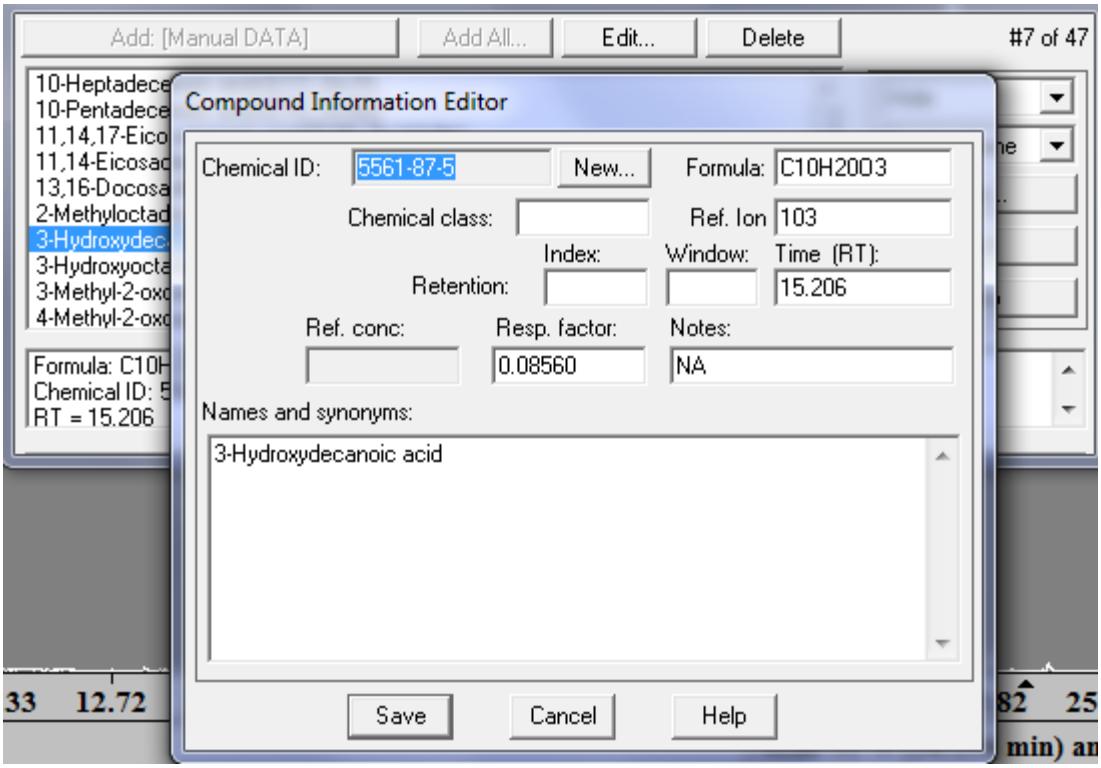
b. → Under "**Deconv.**" tab bar, suggests the following settings:



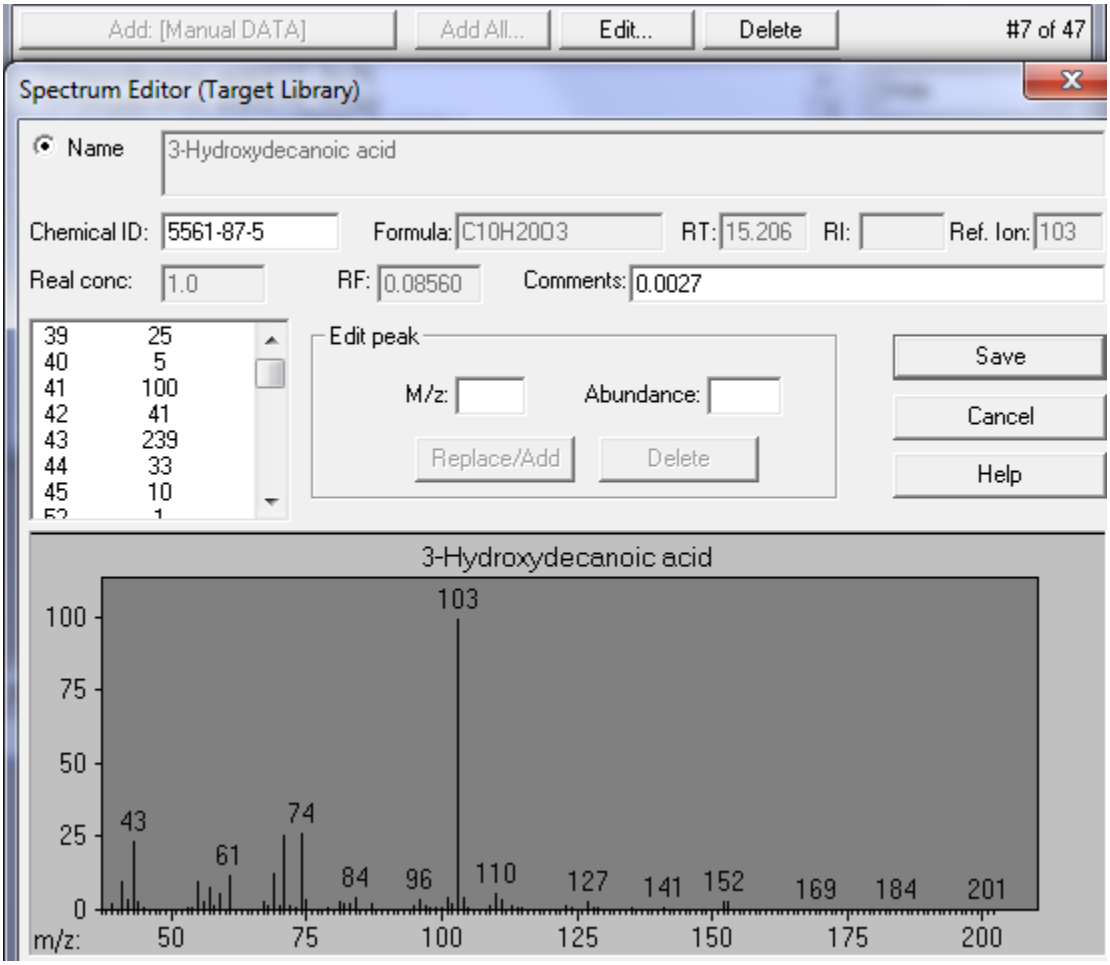
c. → Under "**Libr.**" tab bar, select "**Target Compounds Library**" → click "**Target Compounds Library**" button → select required ***.msl** metabolite library

3.1.1.2. Set or change metabolite response factor and y-intercept

a. To perform absolute metabolite quantitation, AMDIS metabolite library must contain the information about response factor and y-intercept for each metabolite. To set the response factor or change it, on the menu bar, click "**Library**" → select "**Build One Library...**" choose the metabolite and click "**Edit**" → "**Compound**" and insert/change the response factor value. After that click "**Save**":



b. To set the y-intercept or change it, on the menu bar, click "**Library**" → select "**Build One Library...**" choose the metabolite and click "**Edit**" → "**Spectrum**" and insert/change y-intercept value in **Comments** field. After that click "**Save**":

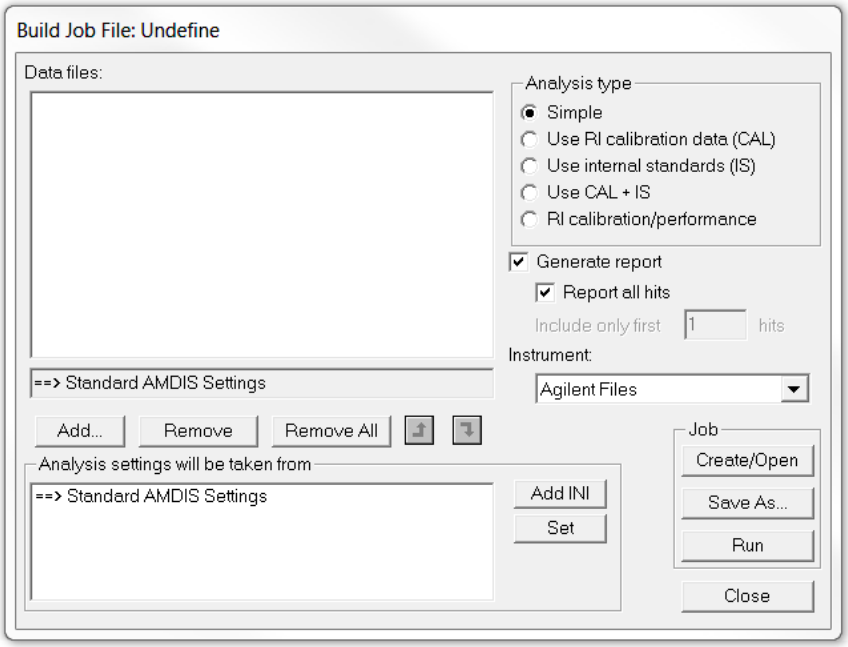


Note: Please make sure that AMDIS library contains metabolite response factors and y-intercepts in order to perform absolute quantitation.

Note: Please verify the RF and y-intercept values as they are not universal for all GC-MS machines

3.1.1.3. To generate AMDIS batch report

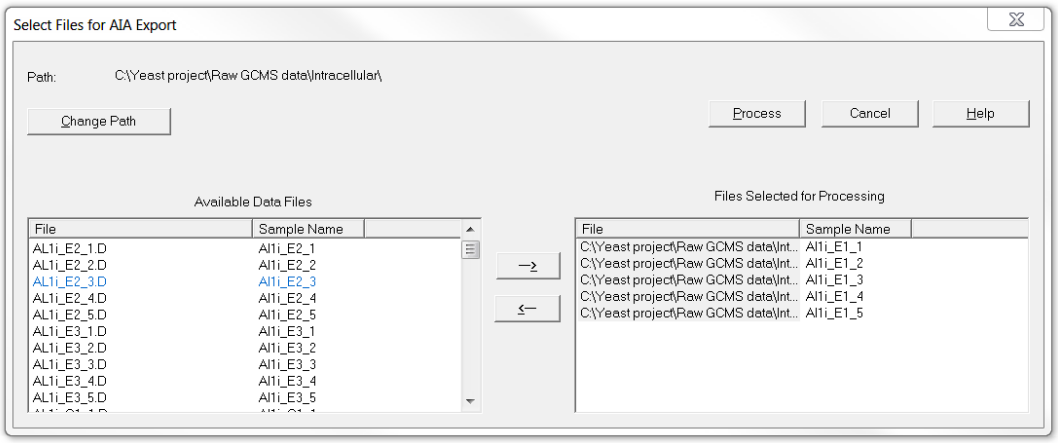
a. On the menu bar, click "**File**" → click "**Batch Job**" → click "**Create and Run Job…**"
b. Make sure "**Simple**" is Analysis Type and tick "**Generate report**" (see figure below).
c. Click "**Save As…"** to choose where to save the report.
d. Click "**Add…**" to input raw GC-MS files.
e. Click "**Run**" to generate an AMDIS batch report.



3.1.2. How to generate CDF files?

3.1.2.1. Generating CDF files from Agilent raw GC-MS data

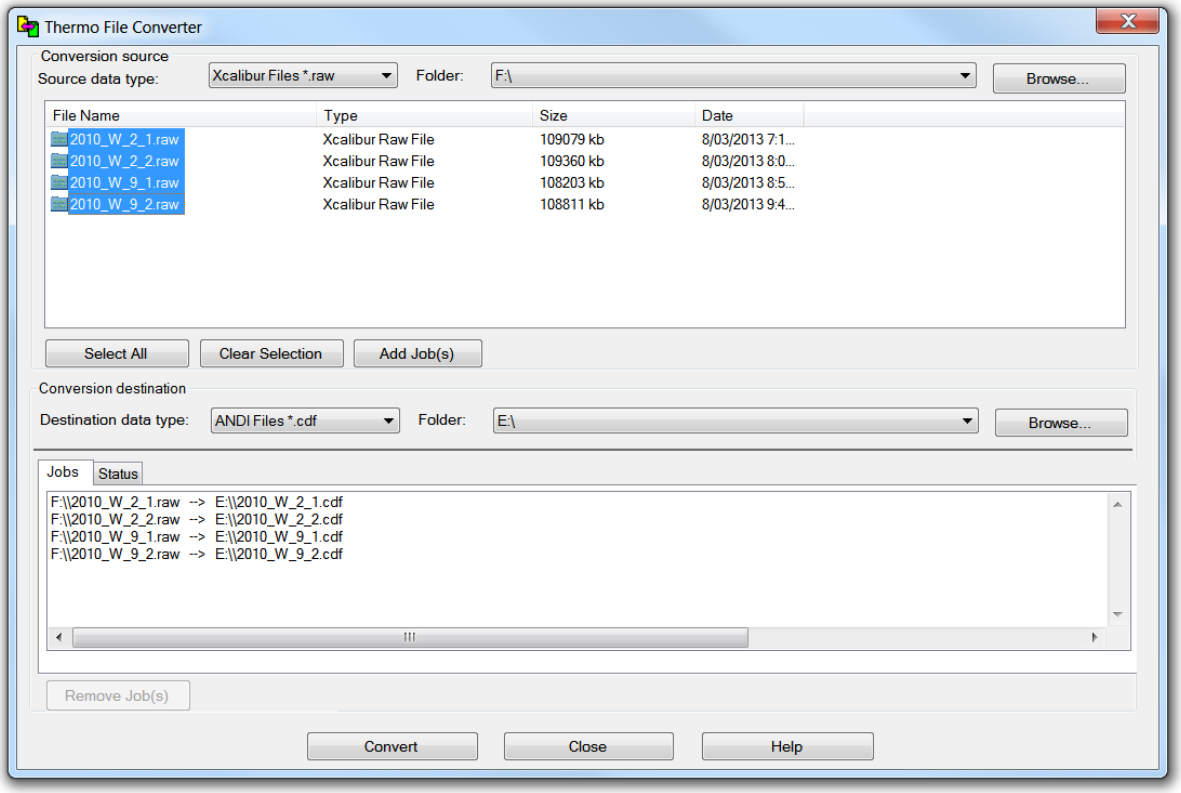
a. Open ChemoStation.
b. On the left window select the folder containing raw GC-MS data.
c. In the menu bar, click "**File**" → choose "**Export data to AIA format…**"→ select "**Create New Directory**" → select a location where all CDF file saved.
d. Afterward, a window will appear → input files by choose files in left window and click right arrow head button (see figure below).



e. Click "Process" button to generate CDF files

3.1.2.2. Generating CDF files from Thermo raw GC-MS data

a. Use search from Window start menu to find "**xconvert**".
b. Choose "**Xcalibur Files*.raw**" as the source data types and browse the location of raw thermo GC-MS data.
c. Choose "**ANDI Files*.cdf**" as the destination data type and browse the folder where cdf files are saved (see figure below).
d. Select raw thermo files on the top window
e. Click "**add job(s)**" button
f. Click "**Convert**" button and raw thermo GC-MS data will convert to cdf files.
g. Change the file extension from ".cdf" to ".CDF", otherwise metab won't work.



3.1.3. MetabQ.settings.csv file

This file will be used by *MetabQ* package and contains the information in order to calculate the absolute metabolite values in the initial sample. Values of only those parameters that are presented in the table below should be changed by user:

Parameter	Value	Description
Extention	CDF	Check the case type of cdf-filesas R is case sensitive (CDF of cdf)
Int standard	d_4-Alanine	Spell the internal standard exactly as it in the AMDIS library
Concentration (d_4-Alanine)	10	Concentration of used internal standard, mM
Concentration (DBP)	1	Concentration of used mass reference internal standard, mM
Sample volume, uL	40	The aliquot of a sample used for analysis
Extract volume, uL	400	The aliquot of chloroform used for extraction of metabolites during sample derivatization process
Sample aliquot, uL	100	The aliquot of MCF derivatized sample taken for analysis
Pooled sample aliquot, uL	50	The aliquot of d-MCF derivatized pooled sample taken for analysis
DBP aliquot, uL	20	The aliquot of mass reference internal standard taken for analysis
a	0.0353	Response factor of d_4-Alanine
b	0.0348	Y-intercept of d_4-Alanine

4. Run *MetabQ*

Make sure all the required documents are placed in a single folder (Read: 3. Prepare documents to run METAB)

a. Open R
b. Load the *MetabQ* package by typing:

```
library(MetabQ)
```

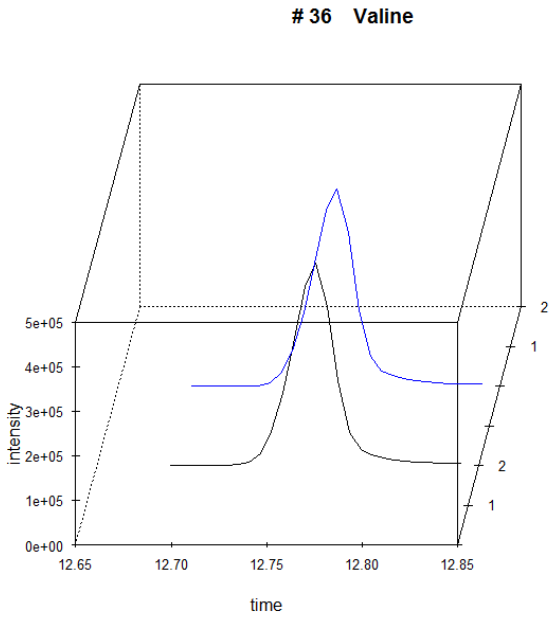
c. Start *MetabQ* by typing:

```
settings()
```

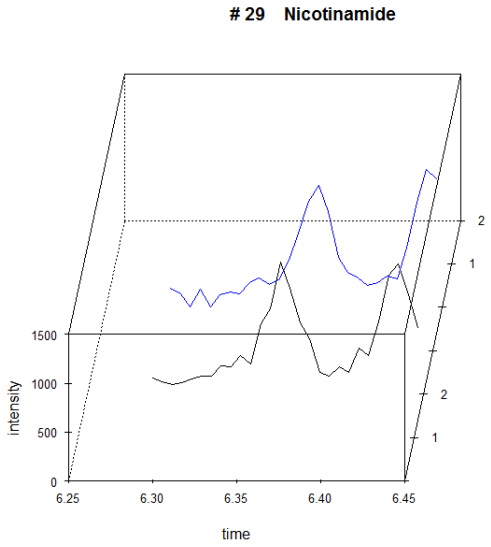
d. A pop-up window will appear for your to browse the folder where all the required documents are located
e. A pop-up window will appear for your to browse the folder with AMDIS libraries (*.msl) and select the AMDIS MSL library
f. This step will generate "folder name"_lib.csv file which contains all the analytical parameters required for further data extraction step.
g. Start the next function by typing :

```
relative()
```

h. The result of this step is "folder name"_auto.csv file containing the extracted abundances of detected metabolites. This list includes both non- and deuterated metabolite derivatives. Also *.png files are generated for each identified metabolite:



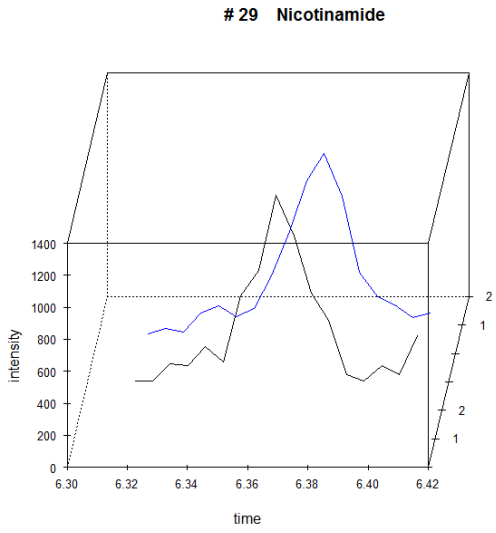
i. Please check all metabolite.png files by means of "One metabolite – one peak" rule (as on figure above for valine). In case there are more than one peak occurred in selected time window (e.g. nicotinamide), the time window should be changed by using *correct()* function.



The retention time for nicotinamide is 6.35 min, however, within set time window 6.25-6.45 min (6.35±0.1 min) two peaks occurred. In order to extract the abundance of correct peak the time window should be changed to 6.30-6.40 min (6.35±0.05 min) by typing the following in console:

```
correct(29,0.05,0.05)
```

This function re-draws the figure according to settings (29-metabolite #, 0.05 and 0.05 – time deviation, min). After this step, "folder name"_correct.csv file was generated with corrected abundances. This step can be run as many times as required and each time the _correct.csv file would be overwritten.



j. After all the corrections have been made, the final function will calculate the absolute values for identified metabolites. To start the function type in console:

```
quant()
```

This function will generate the "folder name"_mg_per_L.csv file containing absolute metabolite concentration values in mg/L.