User Guide

RoWinPro

Rolling window spectral deconvolution of intact protein mass spectrometry datasets
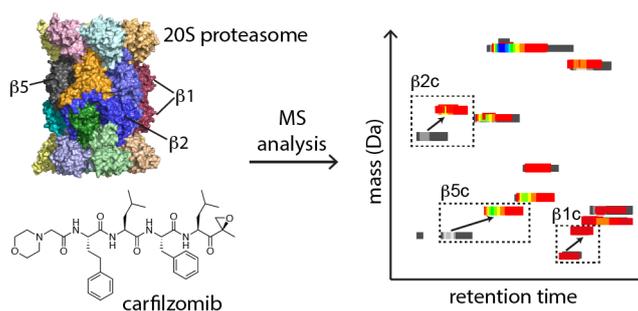
Content

Version 1.4

November 2014

1. Introduction

RoWinPro is intended for an automated rolling windows deconvolution analysis of intact protein mass spectra that were recorded following separation of the sample e.g. on a C4 column. It is particularly useful for the analysis of complex protein mixtures of purified protein complexes or cell extracts. Most commonly, there are more species than peaks in the total ion current chromatogram in these samples which makes the detection of all species by manual selection of spectra challenging.

The program therefore facilitates an automated, comprehensive data analysis that covers all spectra, and provides data for an intuitive visualization of the results in the form of a „sample map". This map plots all detected protein masses over their respective retention times, with data points being color-coded according to the deconvolution intensity. Groups of data points with the same mass can then be clustered to describe the number of species in the sample. Moreover, overlay of sample maps allows the qualitative comparison of these samples and enables the analysis of sample integrity.

The analysis provides insights e.g. into the presence of truncation products, or the stoichiometry of post-translational modifications that are hard to obtain through conventional peptide-based mass spectrometric methods.

For each data file, a range of scans to be analyzed must be given. RoWinPro then carries out the following two steps for all sets of scans („windows") to be analyzed:
1. Extract averaged spectra from a specified Xcalibur raw-file.
2. Run the Promass Znova deconvolution program on these spectra
After completion, all data are compiled in a single table that lists the Promass results and allows for subsequent visualization or further processing (e.g. clustering).

There are three different modes implemented for the selection of sets of scans („windows") to be analyzed. These are schematically depicted in the following figure with an exemplary „window size" of 10 scans.

Sequential window mode
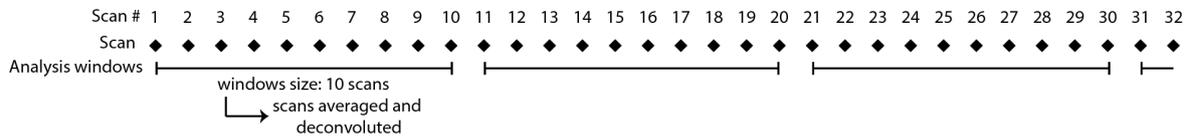
Overlapping window mode
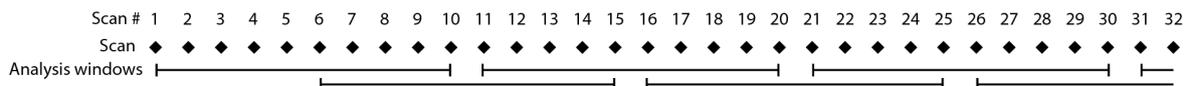
Rolling window mode
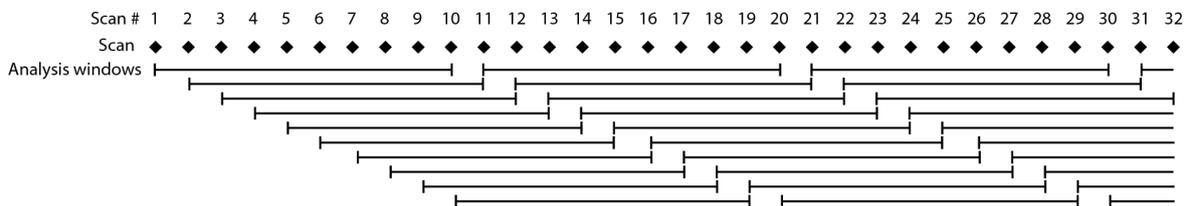
- Sequential window and overlapping window mode allow for a rather fast data analysis, that is particularly suited for the analysis of data from ion trap spectrometers that have a high scan rate and require a higher number of scans (typically >30) to be averaged for analysis.

- The rolling window mode enables the deconvolution of all sequential sets of scans. This mode is particularly suited for the analysis of data generated on FT-ICR spectrometers with low scan rates. It is also recommended when low abundance proteins are to be detected since it avoids any bias that could arise from an unsuitable averaging of scans. However, this mode has the longest analysis times, depending on the CPU performance of the computer.

See the scheme on the following page for a representation of the workflow of a proteasome sample.

Schematic representation of the workflow and follow-up analysis options:

Sample preparation

Subunit separation by reverse phase chromatography

High resolution mass spectrometry

Rolling-window spectral deconvolution data analysis

Selection of scans, Xtract deconvolution and Gauss fitting for precise mass analysis

Automated weighted k-means clustering and subunit assignment

Overlay of sample maps
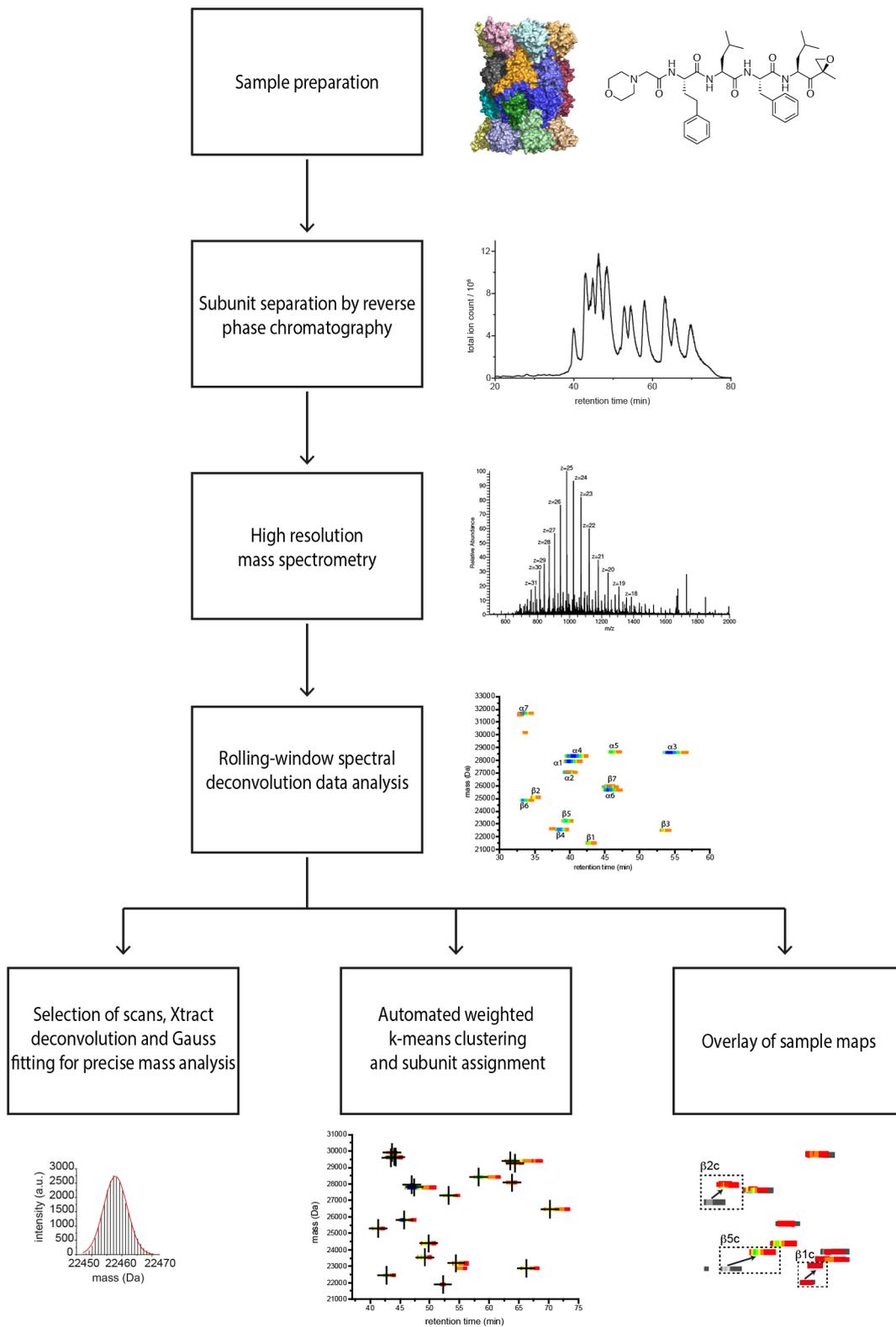
2. System Requirements & Installation

System Requirements:

- RoWinPro is compiled for Windows and has been tested under Windows 7 and Windows 8 (both 32 and 64 bit).

- .NET Framework from Microsoft at Version 4.5.1 or higher is required.

- For data handling, Xcalibur 2.2 or higher including the Xcalibur developer kit (XKD) are required, which can be obtained from Thermo Scientific.

- RoWinPro requires a licensed copy of Promass for the deconvolution step which can be obtained from Enovatia:
  http://www.enovatia.com/products/promass/promass-download/

Raw data

- RoWinPro works with data files of the 'raw' file type as generated by Thermo Scientific mass spectrometers. There are programs available on the internet that convert other data files into the 'raw' file format (e.g. OpenChrome).

- The software extracts consecutive full scans for averaging and further processing. The software can therefore also be used to analyze full scans from data files that contain more than one analysis step (e.g. MS^2 in top down experiments).

Installation

- Visit www.oc2.ch.tum.de, click on "Research" and then on "RoWinPro"

- Download the latest version of "RoWinPro.zip" and unzip it at any location on your hard drive – the exact path is not important

- There is no installation process necessary, simply click on "RoWinPro.exe" to start the program

- If you would like to use RoWinPro in command line mode, use "rowinpro-cmd.exe".

3. RoWinPro Analysis Procedure

This tutorial shows how a .raw file is analyzed by RoWinPro.

- Open the program by double-clicking on RoWinPro.exe
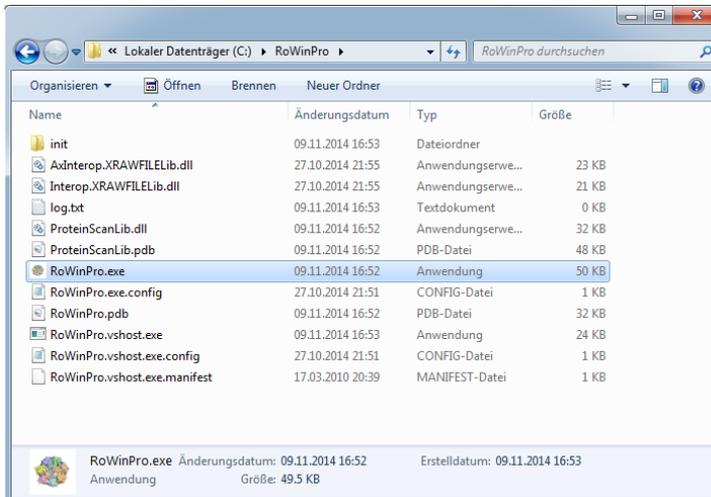
- (1) Click on "Add" at the top right corner to add a new job. (2) Select the .raw-file to be analyzed. This must be at a path were data can be written (e.g. no read-only network drives). (3) Click "Open".
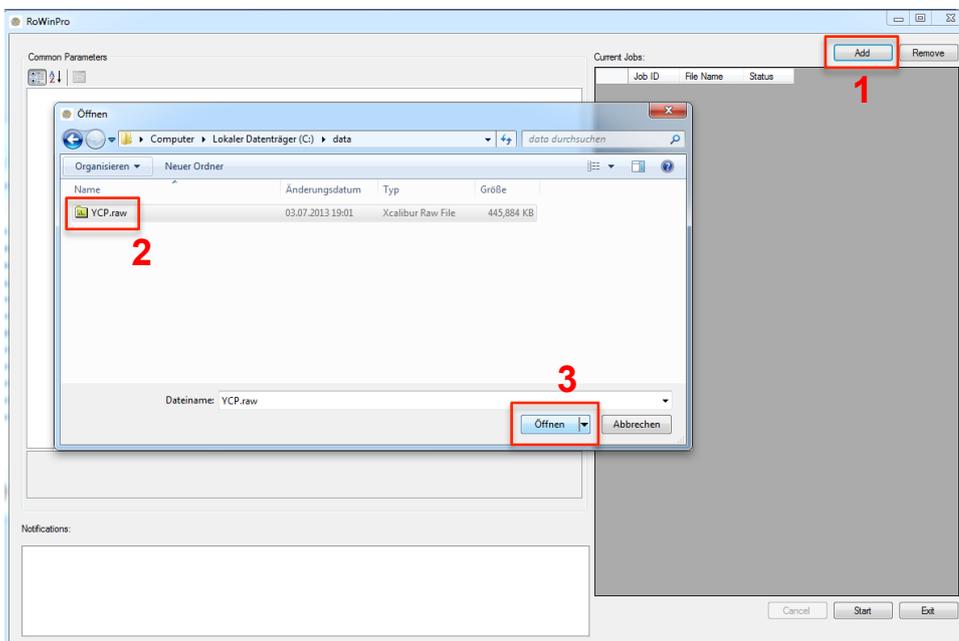
- The RoWinPro GUI consists of three main parts: A list of parameters on the left, a list of jobs on the right and a notifications box at the bottom.
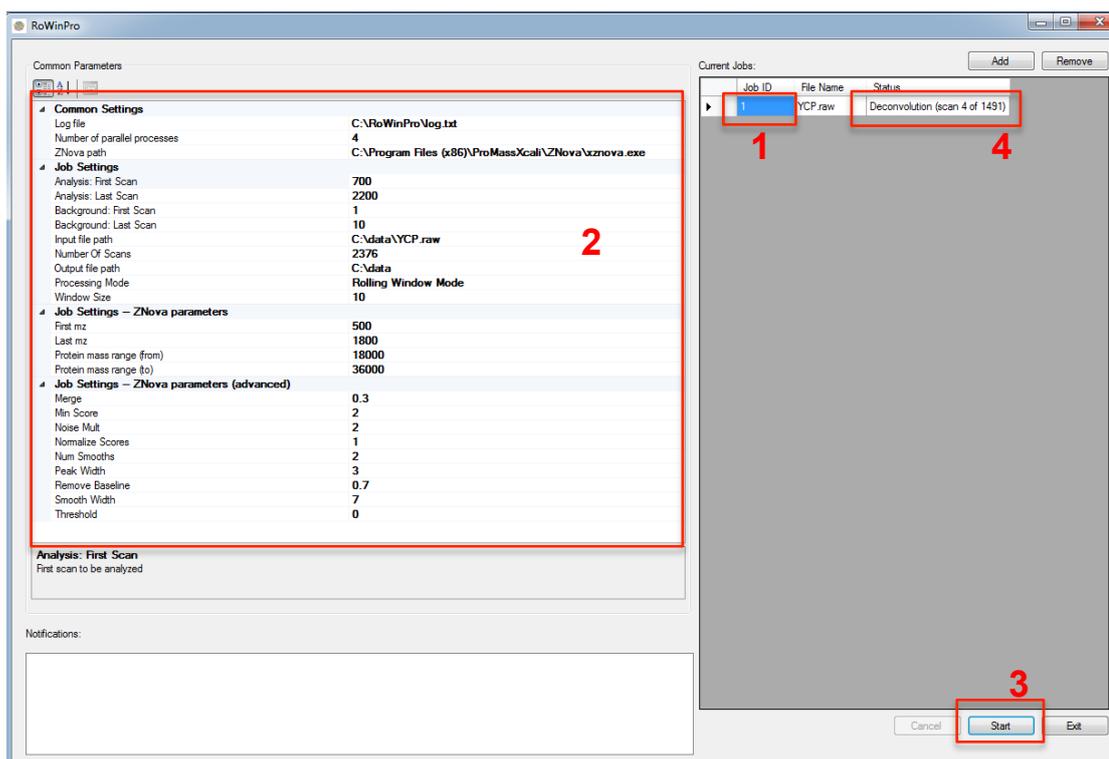
- (1) Select the job that you would like to configure by clicking on the Job ID in the list of jobs on the right which will then turn blue.

- (2) Configure parameters as needed (see below for a detailed list of parameters).

- If you want to analyze multiple samples, or the same sample with different parameters, add jobs to the list with the "Add" button.

- (3) Once all jobs have been configured, start the analysis by clicking "Start" in the bottom right.

- (4) You can track the status of the job in the "Status" column in the Current Jobs list on the right. Once the field shows "Finished", a results file has been generated at the path under "Output file path". Its filename consists of the name of the .raw-file, followed by "_Job_X_results.csv", wherein X is the job ID. Please note that while the analysis is running, no job can be added to the queue.

- Remove old jobs from the queue, before you start new jobs, otherwise they will be executed again (which is intended for an easy correction of mistakes).

4. List of parameters

`Log file`

    Path of a log file which records the processing of all jobs.

`Number of parallel processes`

    This number is detected from the number of CPUs available. Deconvolution will be carried out in parallel if > 1.

`Znova path`

    Path to Znova.exe (which is part of Promass from Enovatia).

`Analysis: First Scan`

    The analysis starts from this scan in the .raw-file. Select a suitable analysis range by looking at the total ion current of the .raw-file in Xcalibur.

`Analysis: Last Scan`

    The analysis ends at this scan in the .raw-file.

`Background: First Scan`

    For spectral averaging, a background signal must be established. Specify a representative range of scans in the raw file where no protein was measured. Default settings for a window size of 10 are scans 1 to 10. Make sure that no protein was detected in these scans; otherwise adjust the background scan range.

`Background: Last Scan`

    See above

`Input file path`

    This field specifies the .raw file to be analyzed.

`Number of Scans`

    This field gives the number of scans that have been detected in the input file.

`Output file path`

    This field specifies where the results file will be generated.

`Processing Mode`

Select between "Rolling Window Mode", "Sequential Window Mode" and "Overlapping Window mode". See the introduction in this document for a visual depiction of these modes.

`Window Size`

This field specifies the window size, i.e. the number of sequential scans that are averaged for analysis.

`First mz`

Specifies from which m/z value onwards the spectra are deconvoluted.

`Last mz`

Specifies until which m/z value the spectra are deconvoluted.

`Protein mass range (from)`

Gives the minimum protein mass in Dalton that can be detected (Promass parameter).

`Protein mass range (to)`

Gives the maximum protein mass in Dalton that can be detected (Promass parameter).

`Znova parameters (advanced)`

See the documentation of Promass available at http://www.enovatia.com/downloads/presentations/ProMass_training_slides.pdf for a detailed description of these parameters.

5. Plotting data

The following tutorial shows how a sample map can be created by plotting the RoWinPro results in Microcal OriginPro 9.1. Any other data plotting software may of course be used.

- Open OriginPro 9.1 and select File -> Import -> Import Wizard
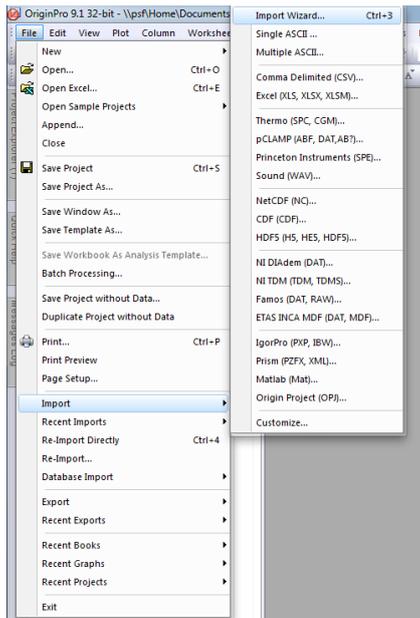
- Select the RoWinPro results file (XXX_Job_X_results.csv) and click through the wizard. Make sure that „tab" is selected as the delimiter separating the columns.
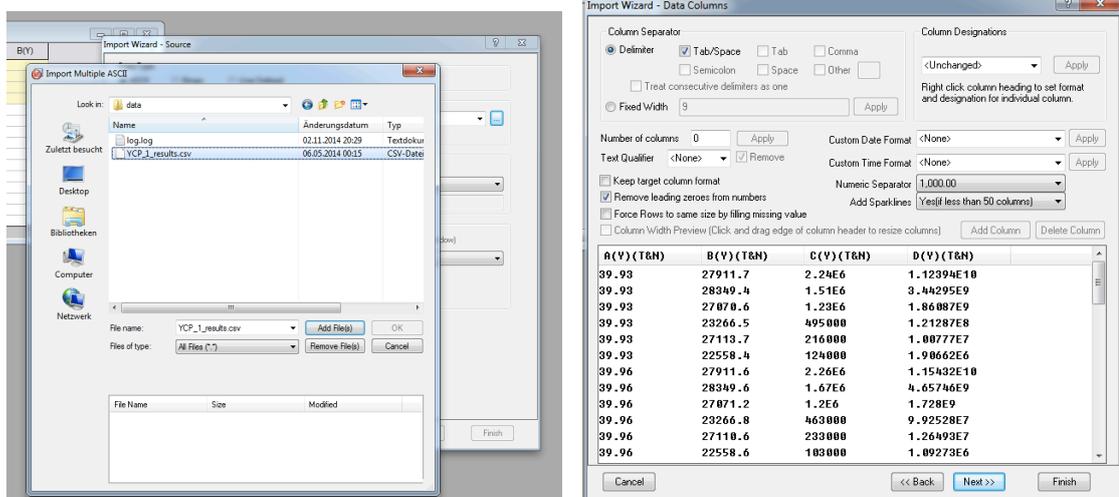
- The workbook created contains four columns:

  Column 1 lists the retention time in min,

  Column 2 lists the protein mass in Da,

  Column 3 lists the deconvolution intensity (a.u.) and

  Column 4 lists a transformed deconvolution intensity which can facilitate the seeding step in k-means clustering.

- Click on Column 3 and enable filtering by clicking on the grey funnel (filtering is implemented in OriginPro 9 and newer versions only). Then select the filtering mode by clicking on the filtering symbol in the column header, select „Top 10" and depending on the dataset start by showing the top 5 to 20 %. This cut-off value must be optimized at the end to ensure that no low-abundance species are omitted and filtering is only used to reduce scattered background and peak tailing.
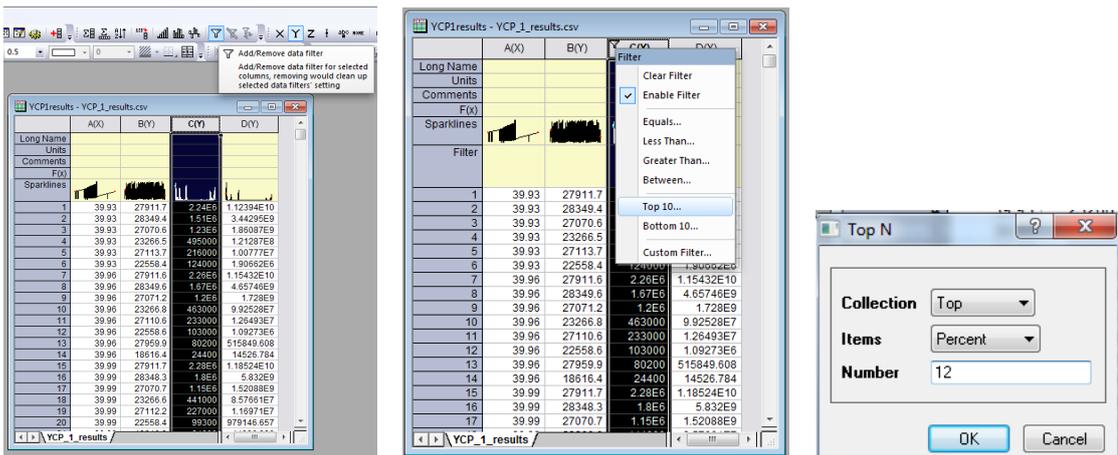
- Select the first three columns and create the map by clicking on Plot -> Symbol -> Color Mapped.
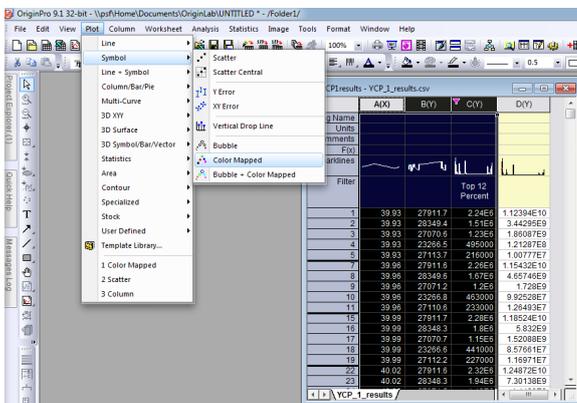
- Disable speed mode to ensure that all data points are displayed in the figure. Click on the layer number (1) in the top left corner of the graph, select Layer Properties and uncheck the options for maximum points per curve and dimension (boxed).
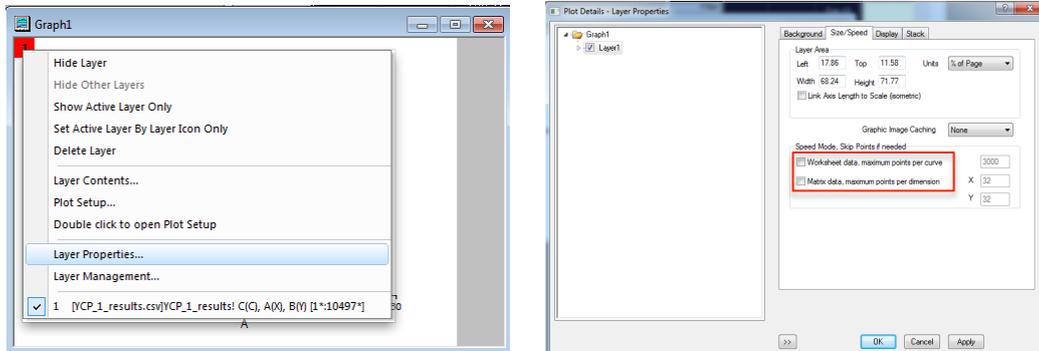
- To change the color scale or the size of the dots, right-click on any data point, select Plot Details and make changes under Symbol and Colormap.
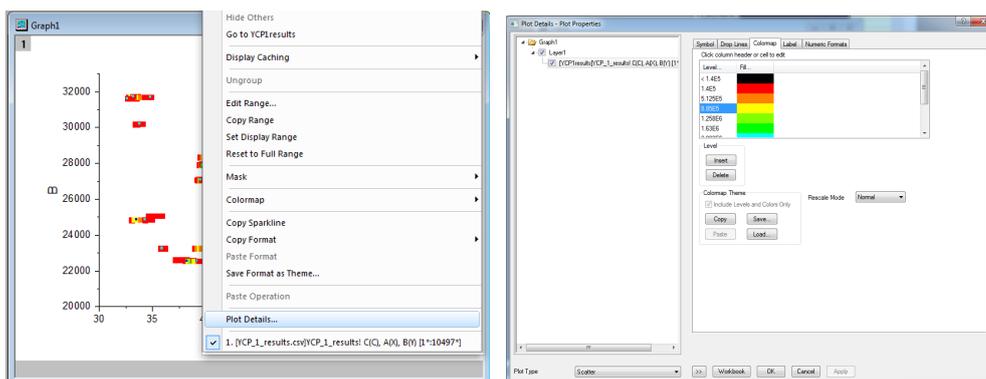
- Make sure to manually adjust the filter level to include all species before exporting the graph through File -> Export -> *.ai or *.eps

- Alternatively, the Plot -> Bubble + Color Mapped option can be chosen in order to scale the size of the dots with the intensity of the detected species.

6. Contact information

RoWinPro was developed at the

Technische Universität München,

Department of Chemistry,

Lichtenbergstr. 4,

D-85748 Garching, Germany.

www.oc2.ch.tum.de

In case of questions, comments and requests, feel free to contact

malte.gersch@mytum.de or stephan.sieber@mytum.de

If you find RoWinPro useful for your research, please cite:

Gersch M, Hackl M, Dubiella C, Dobrinevski A, Groll M, and Sieber SA:
   Intact protein mass spectrometry of proteasomes reveals complex integrity,
   phosphorylation stoichiometry and inhibitor specificity
Submitted, 2014