Proline Tutorial

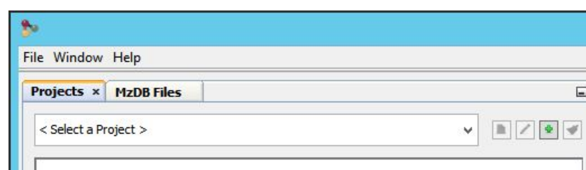
Label-free LC-MS quantitation

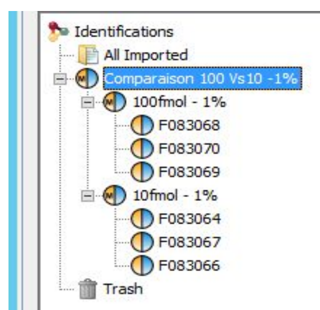
I/ Start Proline

The same project as the one created for previous tutorials will be used. See Tutorial 1 *Proline Basic* for authentication and Tutorial 2 *Quantify by Spectral Counting* for dataset organization.

A. Open project

Select the desired project in the drop-down list.



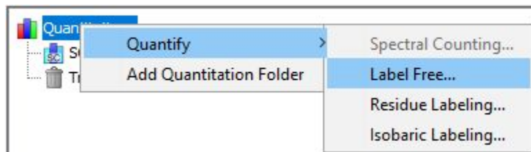
Action	Choose the same project as the one you created in tutorial 2, . The dataset hierarchy should be as shown in the following screenshot. If not, you should create a datasets hierarchy, validate and merge result files as explained in tutorial 2. 

The LC-MS label free quantitative analysis will be done on a sample of 2µg of yeast cells lysate spiked with 10 fmol of UPS1 on one side and with 100 fmol of UPS1 on the other side. 3 replicates of each sample have been analysed on a VELOS ETD spectrometer.

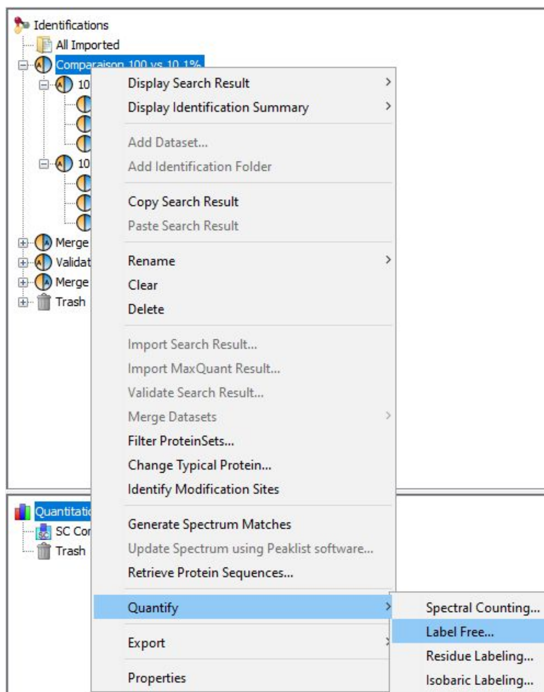
II/ QUANTIFY

A. QUANTITATION VIEW

Quantitation could be launched by using any of these 2 methods :



From the Quantitation tree, at the left bottom of the Proline window. Right click on *Quantitations* node and select « Quantify> Label Free » in the drop down menu.



From *the identification* tree. Select the reference dataset containing the datasets to quantify and from which proteins sets list will be extracted. Right click and select « Quantify > Label Free... »

Step1- Experimental Design dialog.

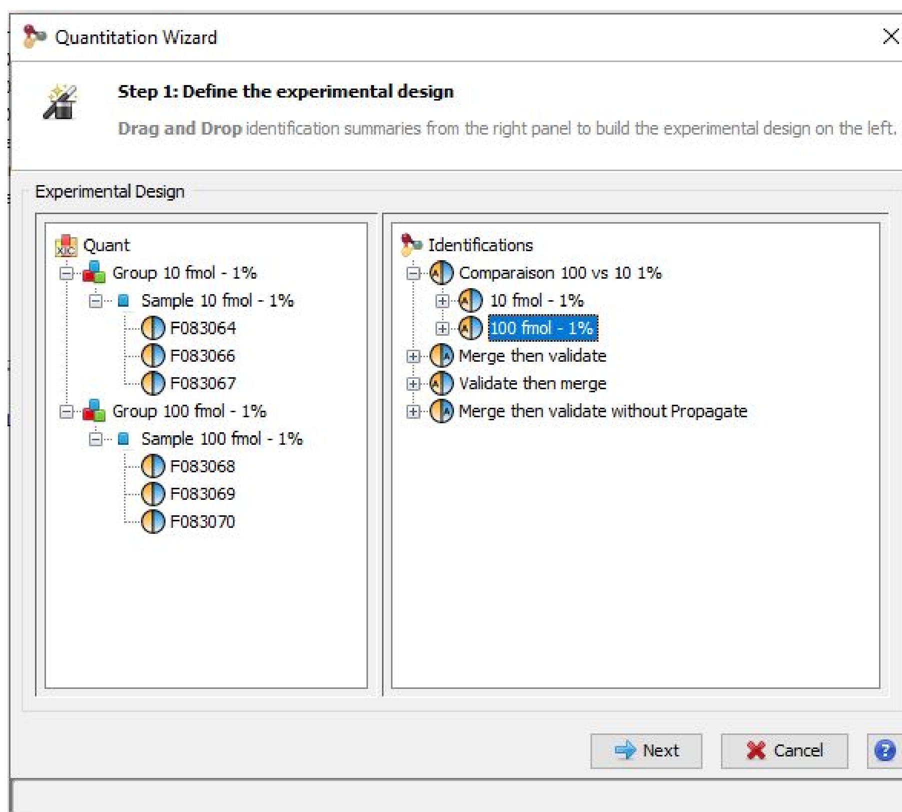
On the left part the tree view shows the experimental design (groups, samples and replicates) that will be built from the identification datasets represented in the panel on the right. Drag and drop identification summaries from le right to the left to build the quantitation experimental design. Depending on the method used to run quantitation, only child datasets of reference dataset or all identification hierarchy will be shown.

The experimental design is composed of :

- groups or biological conditions
- samples or biological replicates
- technical replicates

Note	*If you drag a dataset and drop it on the Quant node, the groups and samples nodes will be automatically generated, depending on the "group-biological replicate-technical replicate" tree.*

Action	Drag and drop the two datasets 100fmol et 10fmol to create 2 groups, a unique sample per group and 3 replicates per samples (see result below).



B. Rename nodes of the experimental design

Nodes of the experimental design could be renamed by one of the following actions:

- Select the node to rename and press the F2 key
- long click on the node to rename it
- Left click the node and choose "Rename" in the popup menu

It is recommended to rename the top level node "Quant" to choose a more appropriate name.

Action	Rename the node to quant_100vs10. Do the same with the samples according to the concentration of spiked UPS1.

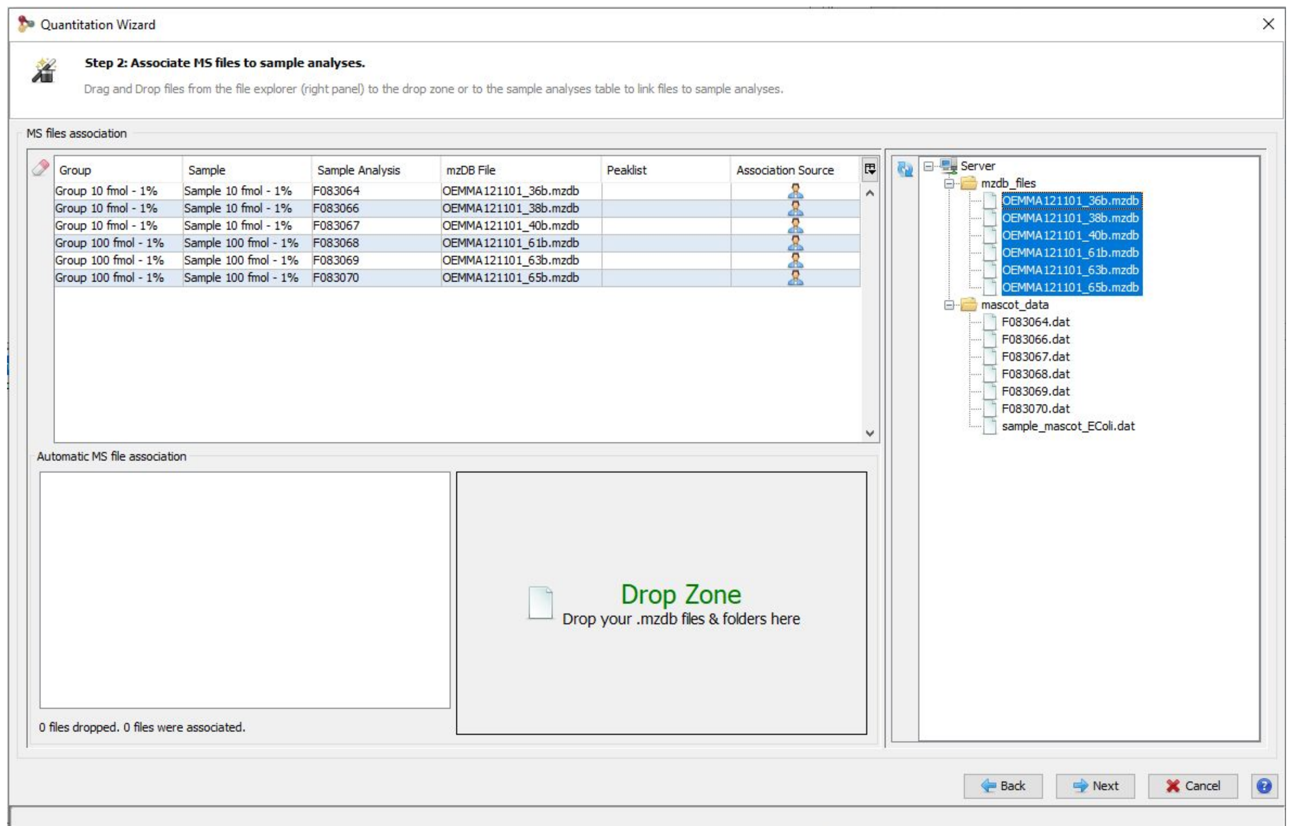
Note	*Nodes can also be deleted from the experimental design by right clicking on the node and choosing “Delete” in the popup menu.*

C. Link datasets and MS files

To perform a quantitation, Proline needs the MS files from which identification results originated. The link between identification results and MS files must be specified by the user in the quantitation dialog (step 2). MS files are required to be supplied in mzDB format and uploaded to the server. Raw files are converted into mzDB format through a converter tool named *raw2mzDB.* This could be done in ProlineStudio via the MS Files tab but batch conversion and upload of raw files could also be automated by using MSAngel or mzdbWizard.

When quantifying an identification result Proline tries to find in the server database an mzDB file already linked to that result. If there is no mzDB file linked, the user must supply a file path by choosing a file from the right part of the panel and drag and drop this file on the corresponding row of the table.

Note	*If values are displayed in the peaklist column, Proline can match peaklist values to file names to link MS files and identification datasets. Drag and drop a batch of files in the Drop Zone area and let Proline associate files with identification results automatically.*



Action	Make sure that each identification dataset is linked to an MS file.

D. Quantification parameters

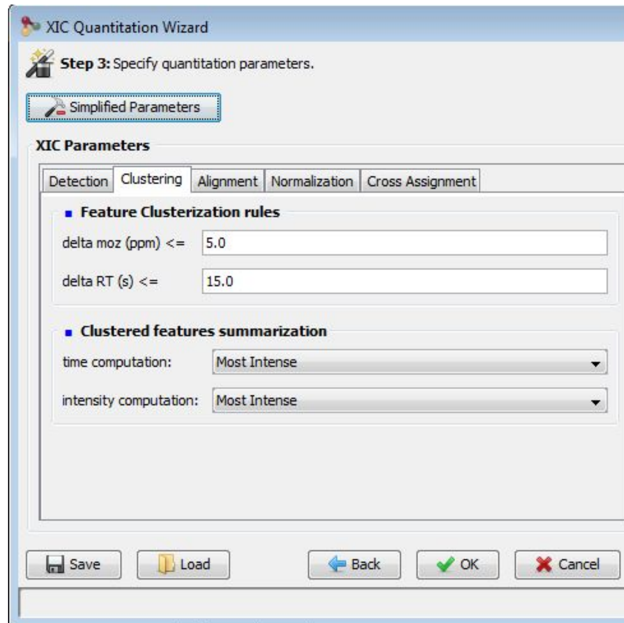
By clicking *Next*, the parameters dialog opens. Not all of the quantitation parameters are displayed at first. To access a complete set of parameter settings, click on *“Advanced Parameters”.*

Extraction moz tolerance: the error tolerance on m/z between the mass at the apex of the peak to extract from the mzDB file and the mass recorded in preceding or following scans.

PSM/Peakel matching moz tolerance: the error tolerance on m/z between the mass of the peakel and the PSMs experimental mass.

Clustering

Extracted elution peaks are clustered as soon as they are matched to the same identified ion and according to the following parameters.



moz tolerance - time tolerance : error tolerance applied to cluster detected features.

time computation : calculation method to be used to determine the elution time of created clusters (elution time of the most abundant feature or median of elution times)

intensity computation : calculation method to determine the intensity value of created clusters (most intense item of the cluster or sum)

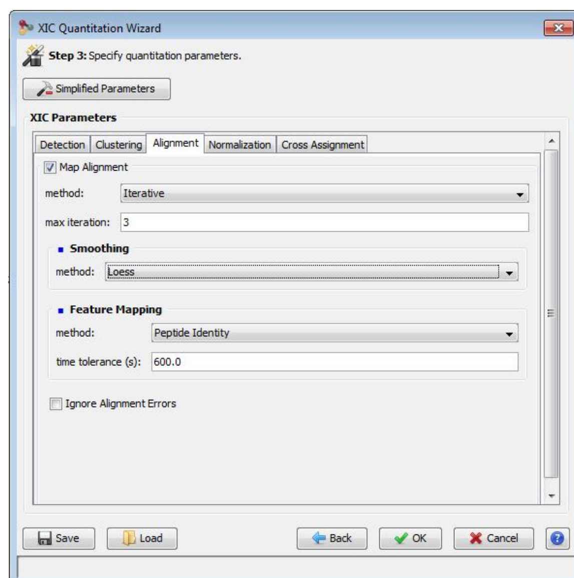
Alignment parameters

Map Alignment :

method : Two alignment methods can be used :

Iterative: for the iterative algorithm, a reference map is first chosen randomly, then every other map is aligned against the reference and the algorithm computes the distance for each pair of maps. The map that has the shortest distance becomes the reference map. The 2 previous steps are re-iterated until either the reference map stays the same between two iterations or the maximum number of iterations is reached.

Exhaustive: the comprehensive algorithm computes the distance between maps for each possible pair of maps and selects the map with the lowest sum of distances to be the reference map. Then all other maps are aligned to this computed reference map and their retention times are corrected.



max iteration : this option is available only for the iterative method. This is a stop condition of this algorithm. When the algorithm has reached its maximum number of iterations, it stops

Smoothing

When the alignment is done, a trend can be extracted with a smoothing method permitting the correction of the aligned map retention time.

Number of landmarks/time interval:

- If the selected smoothing method is *landmark range*, the retention times in the aligned map are adjusted using the median computed on a sliding window whose size is determined by the number of landmarks it must contain. This is set up with the parameter *number of landmarks*. The smoothing method is applied considering the number of landmarks present in the window, and computes the median point for this window.
- If the selected smoothing method is set to *time window*, the retention times in the aligned map are adjusted using the median in a time window. You have to provide the fixed time interval. This time interval corresponds to the window size in which time median will be computed.

Minimum number of landmarks in window: this option is only available for time window smoothing method. This allows you to specify the minimum number of landmarks a window must contain to compute a median on it; it is not significant to compute a median on less landmarks.

Sliding window overlap: overlap is used to compute the step to move the smoothing window forward to calculate a smoothing point for this new smoothing window. This parameter gives the percentage of overlapping between two consecutive windows. For example, if the window size is 200 (seconds or landmarks depending on which smoothing method is selected) and overlap is 20%, the step forward is 200*((100-20)/100) = 160 seconds or landmarks, i.e. the smoothing window is moved forward by a step of 160, so two successive windows overlap each other by a step of 40 seconds or landmarks corresponding to 20% of 200.

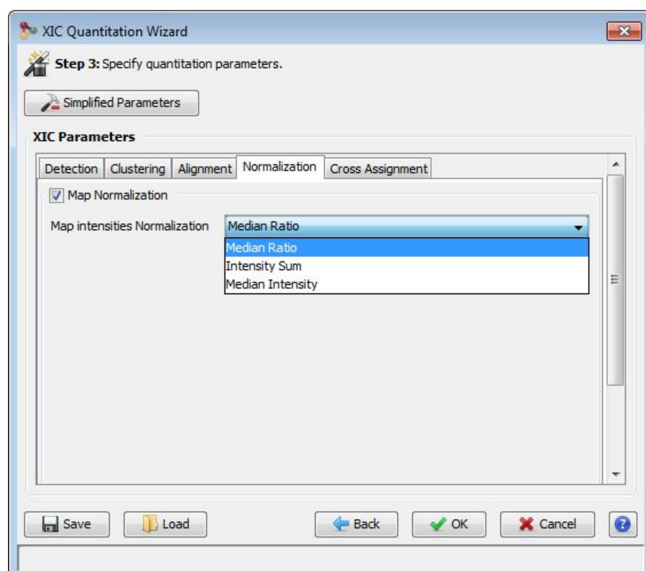
Feature Mapping

method: controls how feature mapping will be performed: by matching feature coordinates (mz, rt) or by using peptide identity associated with identified and quantified features.

Moz tolerance (ppm) -**time tolerance** : mass and time tolerance used to match features coordinates (only used if method is set to feature coordinates, ignored if the method is based on peptide identity).

Normalization

During this step, map intensities could be normalized using three different methods :



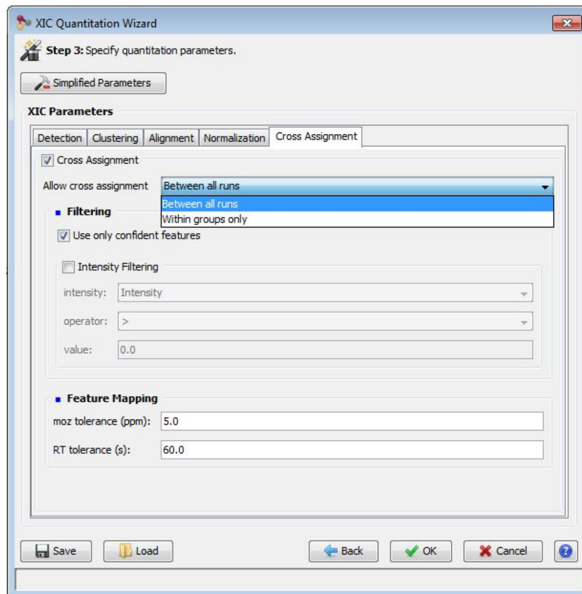
Intensity sum : compute feature intensities sums for each map, set the reference map to the median map, normalization factor for map M = intensities sum of reference map / intensities sum of map M

Median intensity: compute median intensity for each map, set the reference map to median map, normalization factor for map M = reference map median intensity / map M median intensity

Median ratio: compute sum of feature intensities for each map of the map set and sort maps by computed intensities. The map ranking nearest from the median is taken as the reference map. Then for each master map feature, compute ratio as reference map feature intensity / feature intensity

for the considered map. The normalization factor corresponds to the median of the computed ratios

Cross assignment(Master map creation):
This step consists in creating the "master map" (also called consensus map). This map results in the superimposition of all compared maps.



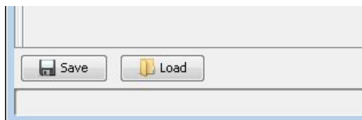
Retain only reliable cross assigned features: if checked, features found by cross assignment will be assessed by checking monoisotopic mass and charge state.

Feature filter : optional parameter to filter detected features. This option can be useful to ignore features of low intensity.

Feature Mapping

moz tolerance (ppm)/ time tolerance : error tolerance used to match features between maps. This procedure is also known as "match between runs" or "cross assignment" in other quantification software.

Astuce	*Parameter settings can be saved and reused to configure a new quantification process.*



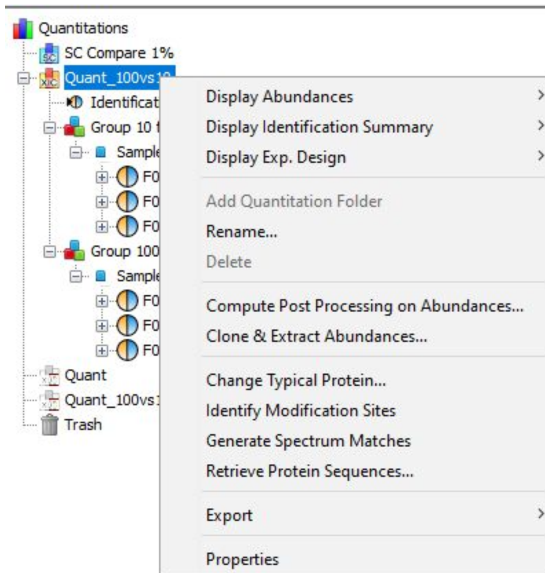
Click *OK* to start the quantification process.

III/ Visualize the quantification results

A. Export quantification results

Export of the quantification includes separate tabs, some providing general information about the analysis, and some giving several levels of information about the results: the « Protein Set » level, the « Protein Match » level, and the peptidic and ionic levels.

Right-click on the quantification to display this menu:



Display Abundances :

Peptide Ion/ Peptides/ Proteins Sets: visualize the ions intensity (with redundancy on the charge), the peptides intensity and the protein abundance.

Modification Sites/Modification Clusters: Identify Modification Sites should have been run before. Visualize modification sites or modification clusters (grouping sites) with corresponding abundance data.

Display Identification summary: same as for the identifications.

Display Exp.Design : visualize the parameters that were used to realize the quantification (experimental design - parameters). Map alignments for all runs or between 2 specific ones could also be visualized.

Rename

Delete

Compute Post Processing on Abundances: post-process the data (performs statistics). Also known as Profilizer, this will be further detailed.

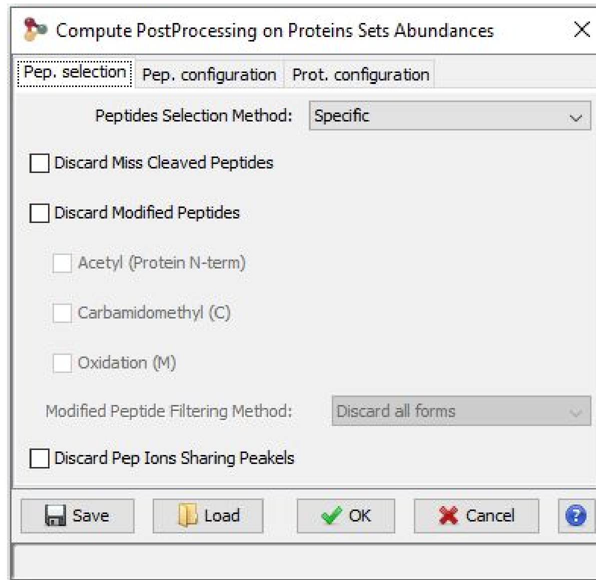
Clone & Extract Abundances: re-launch the quantification on this dataset with updated quantification parameters.

Identify Modification Sites: identify all modification sites in the dataset and cluster these sites according to specified parameters.

Change Typical Protein/ Generate Spectrum Matches/ Retrieve Protein Sequences: same as for the identifications.

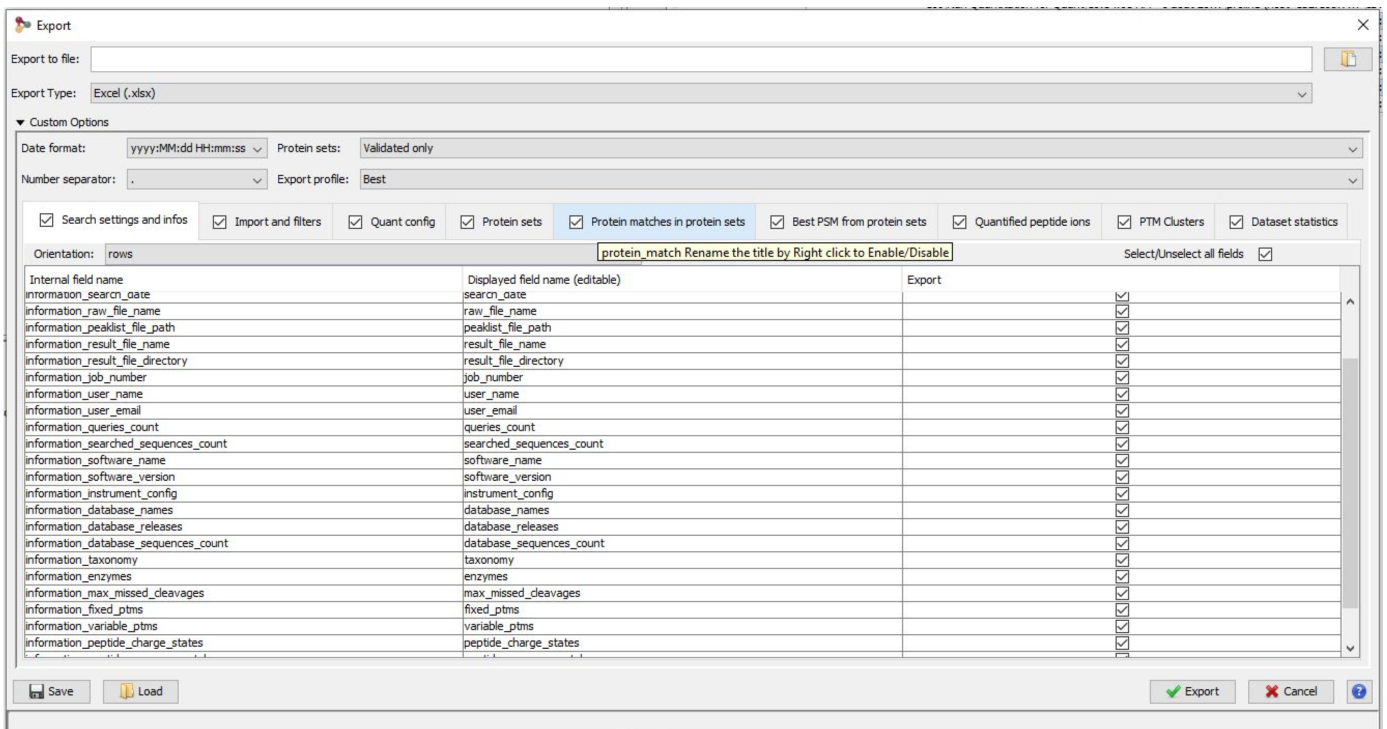
Export : Excel format or list of spectra (list of transitions)

Action	**Compute Post Processing on Abundances** with the following parameters: ● Use specific peptides only (in pep. selection tab) ● Do not *Apply Normalization* for peptides nor for protein sets (in Pep. configuration and Prot. configuration tabs) ● Use *Sum* for *Abundance Summarizer Method* (in Prot. configuration tab).



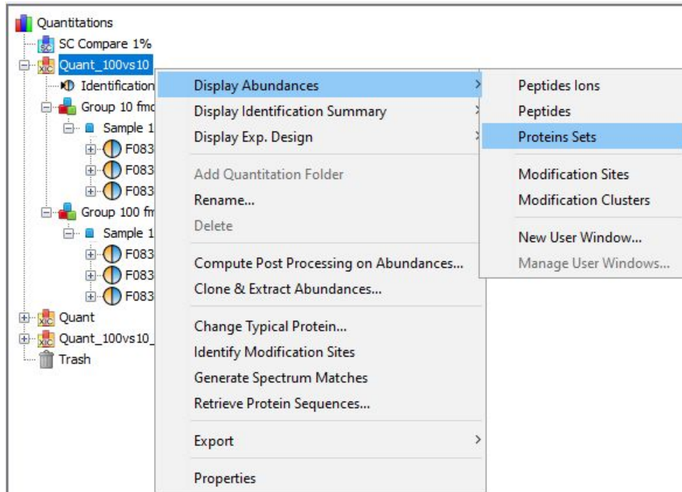
Action	Export the quantification results in an .xls file (version 2003) with the Protein sets tab checked (in addition to the 3 first tabs, checked by default).

Note	*In the export configuration tabs, you will find 2 values for the abundance, depending on the profilizer setup: the raw abundance and the modified abundance (named only 'abundance').*



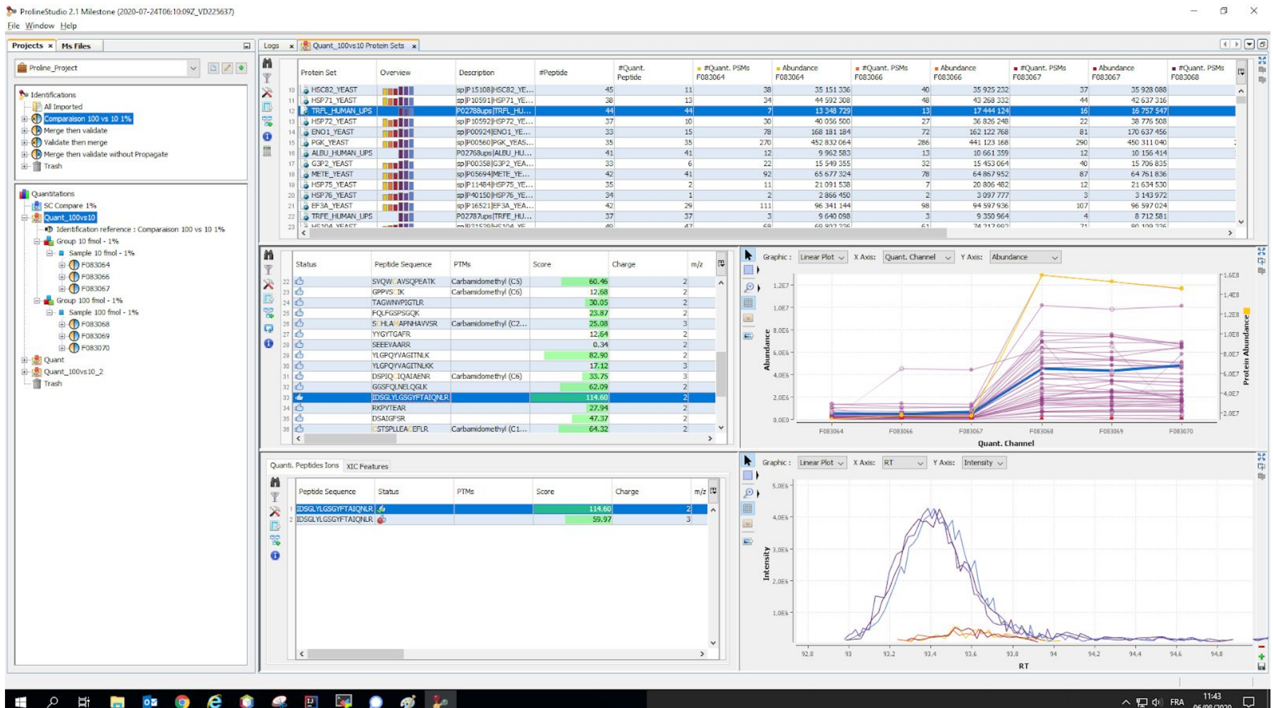
B. Visualize quantified proteins and peptides

Action	Use the *Display Abundances* menu to display the quantified proteins and peptides tables.

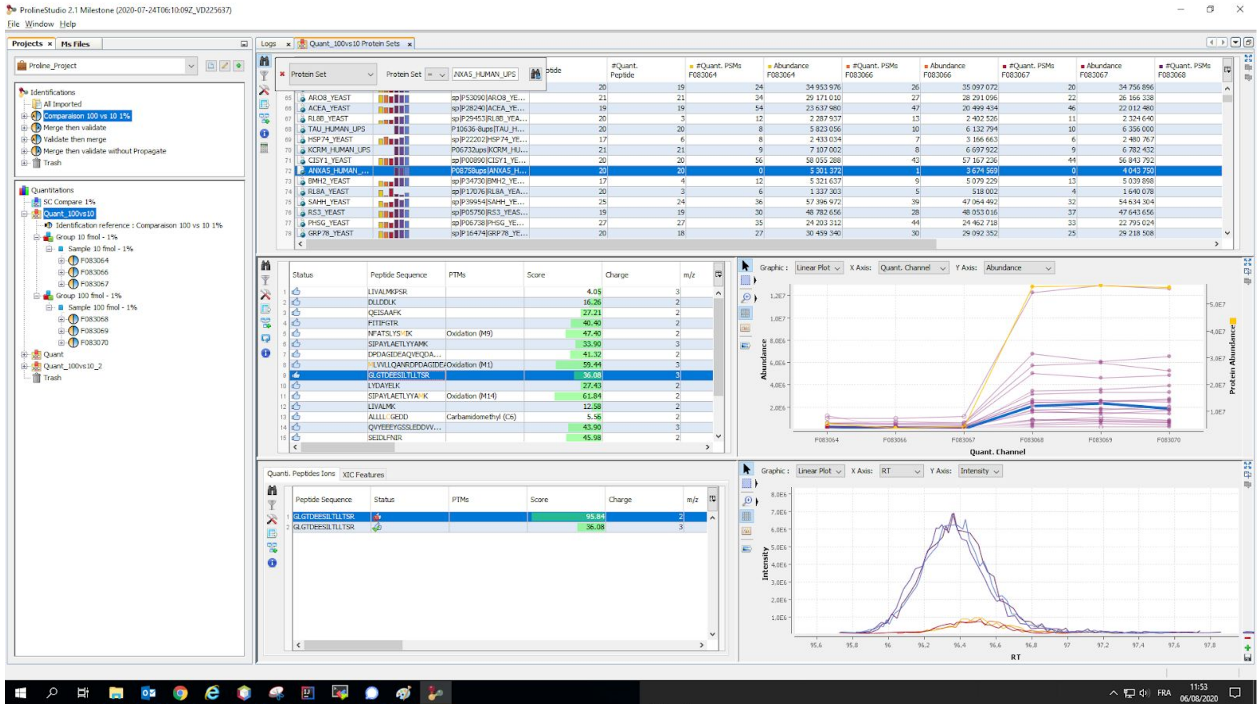


Action	Select the yeast proteins and the UPS1 proteins showing a variation between the two conditions 100 fmol and 10 fmol.

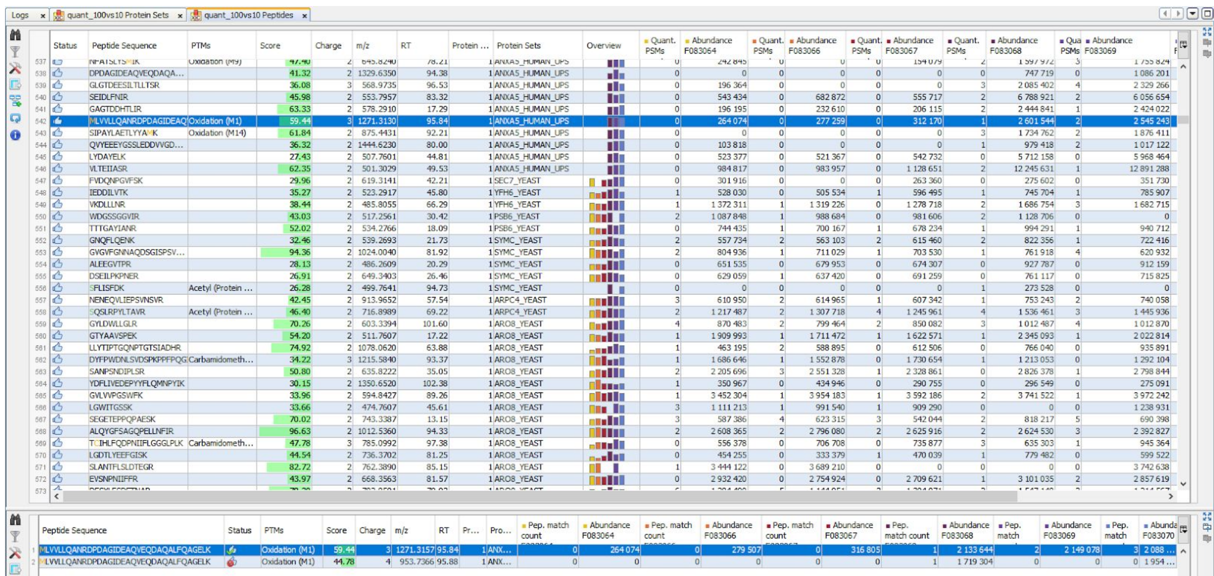
Protein Set: TRFL_HUMAN_UPS



Protein Set: ANXA5_HUMAN_UPS



Example: peptidic view (top) with the table of the peptide ions (bottom) associated with the selected peptide:



Each column can be sorted. The tables can be exported to XLS (Excel) or TSV files.