

***my*ProMS**

# **User's Guide**

**(02/03/18)**

# Table Index

[Introduction](#)

[Connection to ProMS server](#)

[Login](#)

[Server main sections](#)

[Users management](#)

[User classes and access privileges](#)

[Bioinformaticians](#)

[Massists](#)

[Data managers and workgroups](#)

[Biologists](#)

[Account management](#)

[Projects](#)

[Project selection](#)

[Project creation and settings](#)

[Project lifespan](#)

[Accessibility](#)

[Project navigation](#)

[Navigation frame](#)

[Sub-navigation frame](#)

[Option frame](#)

[Results frame](#)

[Project organization](#)

[Experiments](#)

[Samples](#)

[Analyses](#)

[2D gels and spots](#)

[Search results data import: MS Analyses](#)

[Supported search engines](#)

[Collecting search files](#)

[Selecting data files \(Masco, Proteome Discoverer and X!Tandem\)](#)

[Importing analyses](#)

[Import parameters \(Masco, Proteome Discoverer and X!Tandem\)](#)

[Importing MassQuant data](#)

Importing DIA quantification data

- From PeakView

- From OpenSWATH

- Running OpenSWATH quantification

- From Spectrona

Analysis

[Pair-wise peptide-level comparison](#)

[Saving a comparison](#)

[Search for proteins](#)

[Single protein file](#)

[Peptide quantification](#)

[Data import from search results file](#)

[Data export from LC/MS file into MassChroQ](#)

[Managing mXML files](#)

[Running XIC extraction](#)

[Virtual peptides and proteins](#)

[Displaying peptide quantification data](#)

[Protein quantification](#)

[Absolute abundance quantification](#)

[emPAI \(label-free\)](#)

[SIn \(label-free\)](#)

[MassQuant: In-silico LFQ, iBAQ](#)

[Displaying single abundance quantification data](#)

[Relative abundance quantification](#)

[Single-Analyses quantification \(labeled\)](#)

[Design-based quantifications](#)

[Displaying relative abundance quantification data](#)

[Label-free quantifications](#)

[TnPQ](#)

[Comparing multiple protein quantifications](#)

[Exporting multiple quantifications](#)

[Dealing with PTMs](#)

[PTMs relevant to project](#)

[Displaying PTMs distribution](#)

[Comparing modification sites from different projects](#)

[Quantifying modification sites](#)

[Biological samples management](#)

[Properties](#)

[Treatments](#)

[Recording a biological sample](#)

[Linking biological samples to MS Analyses](#)

[Gene Ontology analyses](#)

[GO summarization](#)

[GO enrichment analysis](#)

[Quantitative gene enrichment analysis](#)

[Exploration analysis](#)

[Labeling exploration analysis](#)

[Principal Component Analysis \(PCA\)](#)

[2D-Clustering](#)

[Annotation database management](#)

[Sequence database](#)

[Database](#)

[Listing database](#)

[Adding a new database](#)

[Editing a database](#)

[Deleting a database](#)

[Spectral \(SWATH\) libraries management](#)

[Listing spectral libraries](#)

[Adding a new library](#)

[Merging a library](#)

[Editing a library](#)

[Updating a library](#)

[Resolving the previous version of a library](#)

[Searching for proteins in a library](#)

[Exporting a library](#)

[Deleting a spectral library](#)

[GO files management](#)

[Onlog files](#)

[Annotation files](#)

[Species](#)

[Listing species](#)

[Adding or editing a species](#)

[Deleting a species](#)

[Sequence modifications](#)

[Editing or merging PTMs](#)

[Sequences](#)

[Inserting sequences](#)

[Validation templates management](#)

[Server management](#)

[Log files](#)

Services  
Testing

# Introduction

Proteomic Mass Spectrometer (MS) generates complex data that requires multiple steps of computational and manual processing to be translated into meaningful biological information. To be successful, this process requires the skills of MS specialists, bioinformaticians and biologists. To facilitate such collaboration, we have developed mProMS ([Poillee et al., 2007](#)), a web-based tool that rationalises this data processing workflow while allowing multiple users to interact with the data according to their respective needs.

Typically, output files from MS tools such as Mascot or Proteome Discoverer are imported into mProMS database within a defined experimental context. Spectrometric parameters, protein identifications and variable modification (e.g. phosphorylation) positions can be validated by MS specialists either automatically through dedicated algorithms or manually by visual inspection of each spectrum. Only validated data become accessible to biologists for further investigation. Differential quantification and differential analysis methods are provided through interactive interfaces. Results are displayed as interactive graphics to help users visualise and mine the data. Further biological interpretation is possible using integrated Gene Ontology analyses and enrichment linking to external resources.

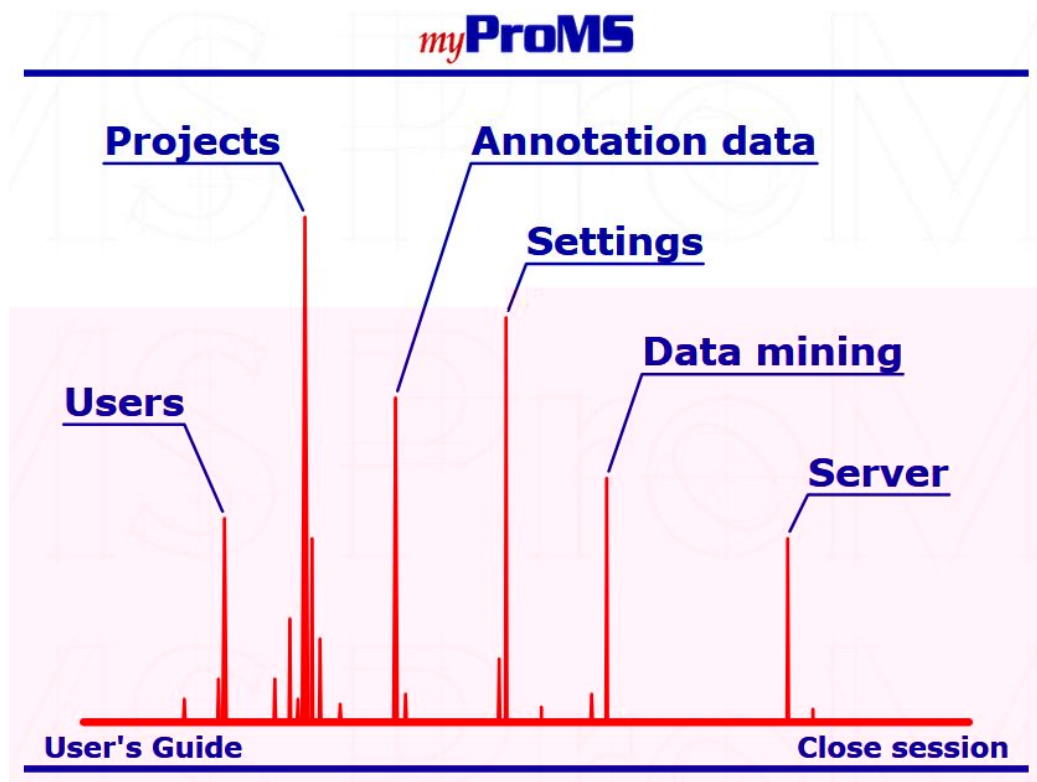
# Connection to myProMS server

## Login

Access to myProMS data requires a login and password. Connecting to local myProMS administrator(s) requires an account. You will then be able to login to the server by clicking on the Start Session button from the home page. During login, you can choose between web- or full screen- display modes.

## Server main sections

Once logged on myProMS, your login, server class (see the **Users management** chapter below for more information) and connection date are displayed at the top of the browser window. Depending on your server class, either **myProMS main window** (Massive and bioinformatician) or the **Project selection window** (biologist) will be displayed as shown below. Users (including biologists) can access the main window at any time by selecting *Main Window* from the Project selection window.





The main window displays links to the 6 areas of m ProMS:

- Users management ,
- Projects access and management ,
- Project annotation data management ,
- Settings management ,
- \*Data mining section,
- Server management ,

*\*Not yet available.*

Each of these sections is described in a dedicated chapter (see below).

Some sections might not be accessible to users depending on the access privileges. Typically, most end users will have access only to their account and projects. See **Users management** chapter below for more information.

# Users management

## User classes and access privileges

Data access in m ProMS is tightly controlled and the user-level permissions are defined in a priori and in a priori. Multiple classes of users are defined based on the specific requirements to perform the different data/ users management and processing tasks available in m ProMS. In addition, in certain classes, users can be granted additional privileges if their specific MS data processing justifies it.

There are 4 classes of users defined in m ProMS ordered by decreasing access privileges: bioinformaticians, massifs, data managers and biologists.

### Bioinformaticians

This class of users is intended for server administration and annotation data management. Although bioinformaticians have full access to all functionalities of m ProMS, they should not be used to perform routine data processing such as MS data validation as they might not have the necessary expertise. We recommend to keep the number of bioinformatician accounts as low as possible and to their extended ability to modify the data,

### Massifs

The massifs class represents MS experts who are in charge of MS data import, validation and reporting. Massifs also manage user accounts and project creation. By default massifs have access to all m ProMS functionalities except those normally dedicated to bioinformaticians.

### Data managers and workgroups

Data managers have the same privileges as massifs but restricted to projects and users of their workgroup. Workgroup usage is optional but is particularly useful for multiple MS-based research labs sharing a common MS facility. In this case, a single instance of m ProMS in a workgroup is attributed to each lab with permissions to access while maintaining management centralisation on the MS facility.

### Biologists

Biologists are end users of m ProMS. They have access to the projects they participate in with various levels of privileges depending on their expertise and involvement in each project. Project access privileges for biologists come in multiple flavors:

#### + Project involvement-based privileges:

- **Guest:** A guest user can only access the project data but cannot modify them.
- **User:** Can access and modify project data.
- **Administrator:** Same as a **user**. In addition, a project Administrator can grant other users access to project.

+ Enterprise-based privileges:

Biologists (users and administrators) can be granted or denied access to MS data and validation of their knowledge of the procedures involved is judged sufficient:

- **Power** (User/Administrator) Can enter validation mode or validate procedures  
identification data only.

- **Super** (User/Administrator).

# Projects

All MS search results and data sequences generated are organized in projects. A project groups sets of data that belong to the same user or group of users and generated in the context of a defined scientific project. End-user (biologists and managers outside their workgroup) accessible to the data is defined at the project level.

## Project selection

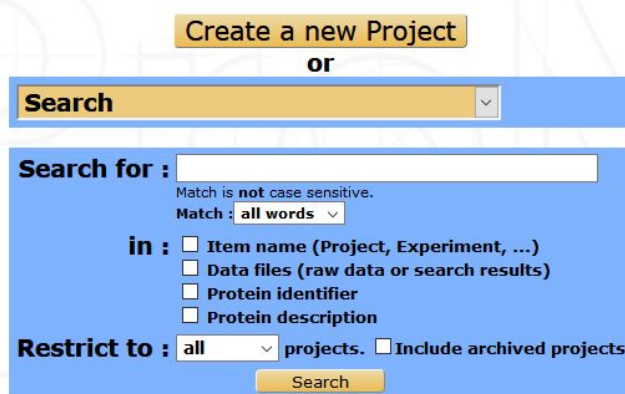
From the ProMS main window, follow the **Projects** link to display the **Project selection** interface.

<Figure Project selection>

For biologists, a search for and list of the projects they have access to will be displayed. For other classes of users, projects can be organized based on the following topics: *Workgroups* (default), *Project owners*, *Active projects* and *Archived projects*. Corresponding projects are listed by ascending name. Their description, owner and/or workgroup together with the access credentials of current user.

If *On-going analyses* is selected, the list generated is composed of Analyses still undergoing validation. Their name, description, data file name, creation date and corresponding project are displayed. This list can be sorted by Importance, Name, MS type, Validation status and Data file name.

Alternatively, a *Search* can be performed using various criteria:



The screenshot shows a web interface for project selection. At the top, there is a yellow button labeled "Create a new Project" followed by the word "or". Below this is a blue box containing a search form. The form has a yellow "Search" button at the top left. Inside the blue box, there is a "Search for :" label followed by a text input field. Below the input field, it says "Match is not case sensitive." and "Match : all words" with a dropdown arrow. Underneath, there is a section labeled "in :" with four checkboxes: "Item name (Project, Experiment, ...)", "Data files (raw data or search results)", "Protein identifier", and "Protein description". At the bottom of the form, there is a "Restrict to :" label followed by a dropdown menu set to "all" and the text "projects." followed by an unchecked checkbox labeled "Include archived projects". A yellow "Search" button is at the bottom right of the form.

Projects are then listed together with the items that were matched during the search. Once listed, click on the *Open* button corresponding to the project wanted to access it.

## Project creation and settings

Only bioinformaticians, mass spectrometrists and data managers can create projects.

From the **Project selection** interface click on the *Create a new Project* button. The following form is then displayed:

**Creating a new Project**

**Name :**

**Description :**

**Protein visibility :**

- ☒ A protein is **Visible** only when **Alias** of a Match Group.
- ☐ A protein is **Visible** everywhere if **Alias** of at least 1 Match Group.
- ☐ A protein is **Visible** everywhere if **Alias** or made **Visible** in at least 1 Match Group.

**Identifier conversion :**

**Relevant PTMs :**

<input type="checkbox"/> Acetyl (A)	<input type="checkbox"/> Biotin (B)	<input type="checkbox"/> Biotine-phenol, (B)
<input type="checkbox"/> ,BG-PEG9-NHS, (BG)	<input type="checkbox"/> Carbamidomethyl (C)	<input type="checkbox"/> Dimethyl (D)
<input type="checkbox"/> Methyl (M)	<input type="checkbox"/> Oxidation (O)	<input type="checkbox"/> Phospho (P)
<input type="checkbox"/> Propionyl (P)	<input type="checkbox"/> Sulfo (S)	<input type="checkbox"/> ,Snap-Tag, (ST)
<input type="checkbox"/> ,Snap-Tag oxyde, (STO)	<input type="checkbox"/> Trimethyl (T)	<input type="checkbox"/> GG (U)

**Project owner :**

**Workgroup :**

**Start date :** 07/02/2018 18:06:01

**Status :** Starting

**Comments :**

- **Name:** Provide a mandatory name for the project.

- **Description:** an optional description for the project.

- **Protein visibility:** Select the visibility of the project - provide protein visibility rules to be used. See **Match Groups and protein visibility** below for detailed information on this concept.

- **PTMs:**

- **Comments:** an optional comments for the project.  
Click on the Save button to create project.  
Projects can be edited at any time to modify any of these settings.

## Project life span

- **On-going:** Once created, a project is set as **active** and **on-going**. This means that it can be populated with new items and data. On-going projects are flagged with a yellow icon in the project selection window.
- **Ended:** If the project is judged completed, it can be edited and **ended** by clicking on the *End* button at the bottom of the edition form. Ending a project will automatically end all partially validated analyses in the reporting (see **Validations** and **Reporting** sections in the **Analysis management** chapter below for more information). Once ended, a project is still active and accessible but can no longer be edited or populated. Ended projects are flagged with a green icon in the project selection window.
- **Archived:** As time passes, some projects might no longer be accessed by analysts. These projects can be archived to save space on the server. All data files stored on the database will be compressed. Archived projects are flagged with a red icon and are no longer accessible for data exploration. They can however be listed in the Project selection window by selecting *List of: Archived projects*.
- **Restoration:** Archived and Ended projects can be fully restored to an active state if necessary by clicking on the appropriate button in the project home page.

## Accessibility

Bioinformaticians and massifs have full access to all projects recorded in m ProMS. Data managers have full access to all projects within their workgroup. Biologists and managers outside their workgroup may be explicitly granted access to projects when needed. The project access management interface is accessible from the project home page by clicking on the *Project Accessibility* button in the **option frame**.

## Accessibility to Project **User1** No Workgroup assigned.

### Users allowed to access this Project

User	Status	Workgroup	Access Right*
No users			

**Allow** --Choose from List-- **to access this Project.**

#### Access rights description:

<b>Guest :</b>	Read access to validated data.
<b>User :</b>	Read/Write access to validated data.
<b>Administrator :</b>	<b>User</b> + Project access management.
<b>Power (User/Administrator) :</b>	<b>User/Admin.</b> with additional read/write access to non-validated protein data.
<b>Super (User/Administrator) :</b>	<b>User/Admin.</b> with full access rights on the current project.
<b>Manager :</b>	Full access rights on all projects of a workgroup.

The interface summarizes the list of users able to access the project together with their credentials. New users can be added one at a time. Once added to the access list, users are given a default. Select the credentials to assign to provided each user with. The access rights available are listed below the user access form. See also **User classes and access privileges** above for more information.

Click on the Save button to validate any changes.

## Project navigation

Navigation frame

Sub-navigation frame

Option frame

Results frame

## Project organization

Data in a project are hierarchically organized as shown in the figure below :

<Figure Project hierarchy>

## Experiments

An **Experiment** item represents an actual biological experiment for which MS data will be collected.

To create a new experiment, select the project element in the top left navigation frame and click on *Add Experiment(s)* in the top right frame.

### Adding new Experiment(s)

**Name :**

**Multiple entries labels :**  Use ' ' between single values and '-' for range (eg. 1,3,5-10).

**Description :**

**Start date :** 07/02/2018 18:15:57

**Preferred species :** - Select -

**Comments :**

Save Clear Cancel

Provide a name and optional description and/or comments. Multiple experiments can be created at once if the field Multiple entries labels is filled in. Labels defined in this field will be sequentially appended to the name each experiment created. Labels can be defined individually using a comma-separated string (eg. A,D,G ) or a range string using a - (eg. 1-5 );

### Samples

A **Sample** item is a loose entity that can represent a single or multiple mixed (e.g. for labelled quantification) biological samples. It can be linked as a sub-experiment or Analysis-containing item. It is possible to define its function depending on the experimental context of the analyses it contains.

To create a new sample, select its parent experiment in the navigation frame and click on *Add Sample(s)* in the top right frame.

### Adding new Sample(s)

**Name :**

**Multiple entries labels :**  Use ' ' between single values and '-' for range (eg. 1,3,5-10).

**Description :**

**Start date :** 07/02/2018 18:16:50

**Comments :**

Save Clear Cancel

Provide a name and optional description and/or comments. Multiple samples can be created as described for experiments (see the **Experiments** paragraph above).



## Analyses

An **Analysis** corresponds to a database imported from a single search engine result file: most of the MS/MS spectra (except for PMF results), the peptide/protein identifications and associated quantities are then present in the file. Analyses data may be imported, validated and reported before end users can access them and further process their results. These procedures are described in the chapter [Analysis data import and validation](#) below.

## 2D gels and spots

**2D-Gel** (Two-dimensional gel electrophoresis) can be recorded and is aliased in mProMS. A picture of the gel (JPEG format only) may be uploaded in mProMS so as to keep a visual record.

To create a 2D-gel, select its parent experiment in the upper navigation frame and click on *Add 2D-Gel* in the option frame. The form shown below will be displayed.

**Adding a new 2D-Gel**



Name :

Gel image file :

(Jpeg only)

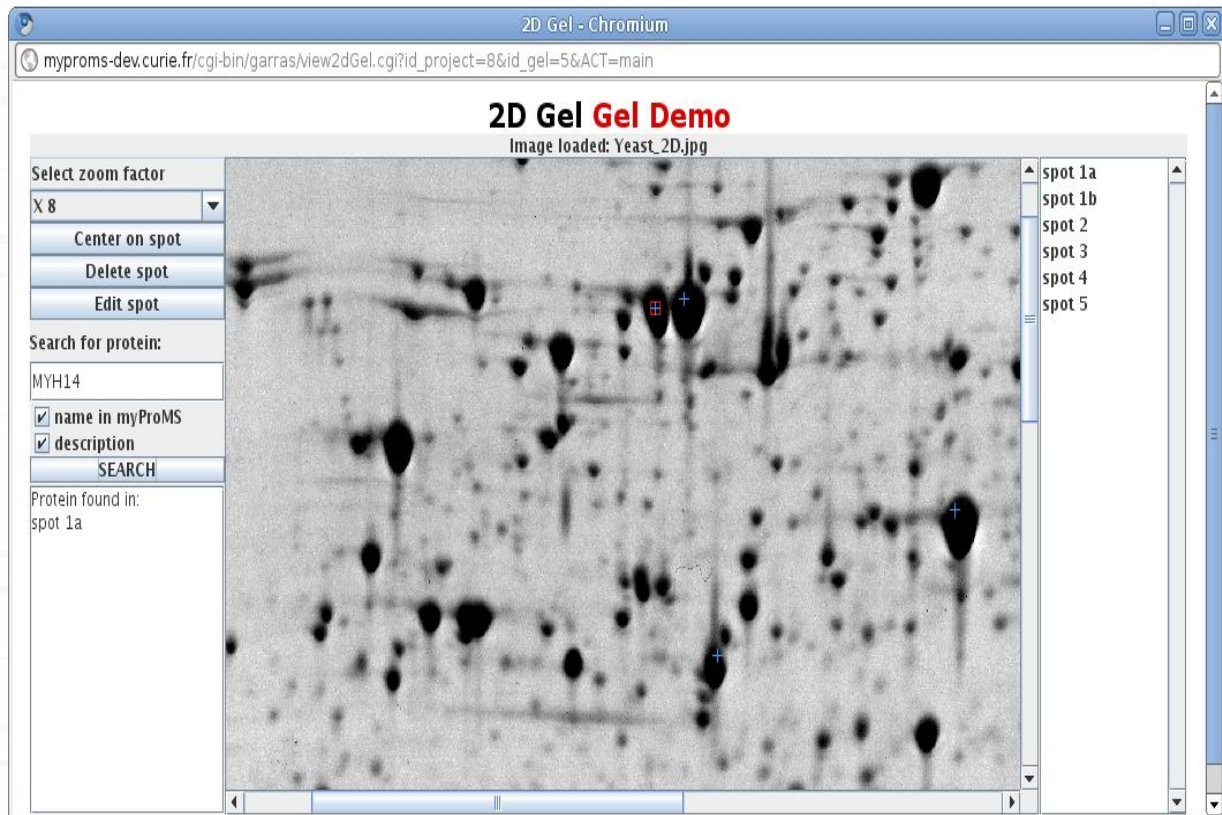
Description :

Start date : 06/08/2013 10:54:40

Comments :

Provide a name, optional description/comments and select the JPEG image file of the gel to be uploaded. Click on **Save** as soon as the new 2D-Gel element in mProMS.

To alias a 2D-gel, select the gel element in the upper left navigation frame and click on *Display 2D Gel* in the option frame. A pop-up window will display the gel image as shown below (JAVA may be enabled in your browser and your machine accepts the security issues).



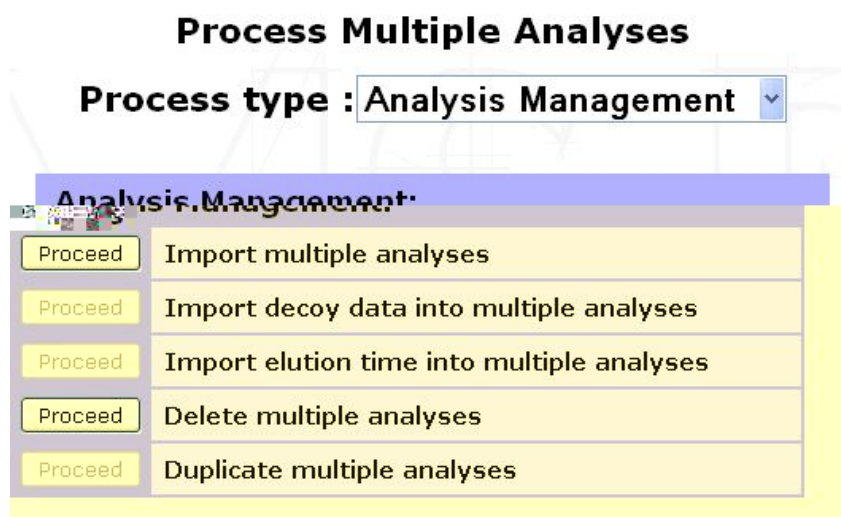
To record a new spot, double click on the gel image where the spot is located. The pop-up box will provide a mandatory name and other optional information (isoelectric point, molecular weight, external identifier if you use another image processing software like ImageMaster, ...). A blue crosshair is then displayed to represent the spot recorded. You can select/deselect a spot by clicking on the corresponding cross. When selected, a spot can be edited or deleted. Protein identification data can be appended to a spot by linking this spot to an existing **Sample**. This sample will no longer be listed in the navigation window and identification data will become accessible through the spot only. When modifying or a spot with linked identification data, the corresponding protein (besides identification in the associated analysis) is displayed in the spot pop-up box. You can also search a protein of interest (in the above example, MYH14) and spot(s) containing this protein will be highlighted with a red square.

In addition to the graphical visualization of the gel, all spots are listed in the lower navigation frame once a 2D-gel is selected. Each spot can be edited/deleted from the same interface.

## Search results data import: MS Analysis

The collection of spectra/peptides/proteins(/quantification) data contained in a search results file are imported into ProMS as an **Analysis**. Only bioinformaticians, massists and managers can import Analyses.

Select the **Experiment** or **Sample** or **2D Gel** in which the Analyses must be imported and click *Process Analyses* in the option frame. From the selection menu displayed, select *Analysis Management* to display the list of available options.



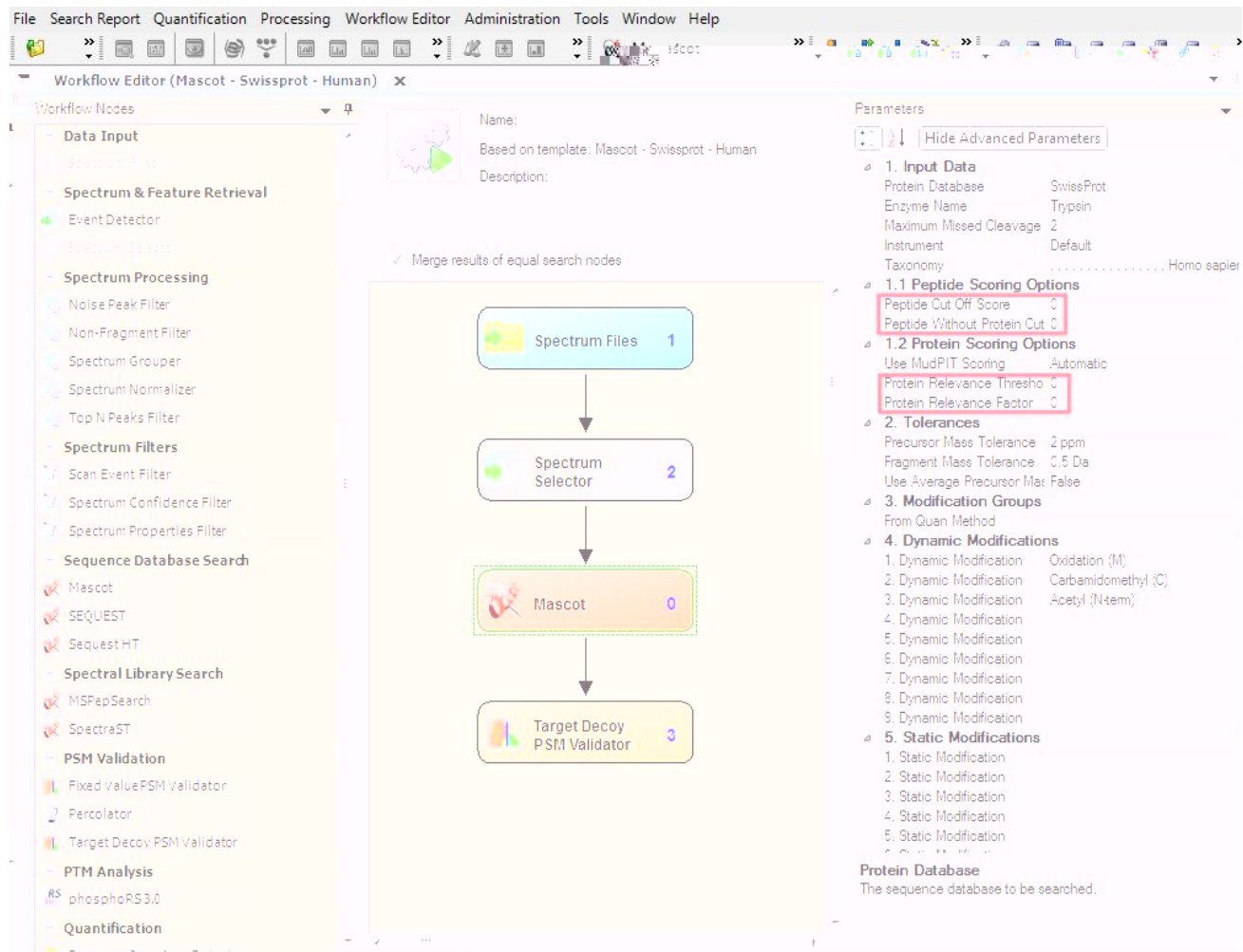
## Supported search engines

ProMS allows to import from various search engines :

- Mascot (DAT file or MSF file from Proteome Discoverer Software by Thermo Scientific).
- Paragon (XML generated from ProteinPilot™ Software by AB SCIEX, group2 molecule).
- Seqes (MSF file from Proteome Discoverer Software by Thermo Scientific).
- Phen (XML file generated through Phen platform by GeneBio). DEPRECATED!
- Andromeda/MaQan (mqpar. ml and 3-4 files are required).
- X!Tandem (XML file from X! Tandem pipeline ([PAPPSO](#)) or from Trans-Proteomic Pipeline ([TPP](#))).
- PeakView (Exported Excel XLXS file for SWATH data).
- OpenSearch (TSV file from [OpenSWATH](#) workflow).
- Specrona (TSV file generated from Spectra no™ by Biognosis).

**Important note:** if to perform Mascot searches with Proteome Discoverer (PD) Software, make sure to disable Protein and Peptide filters. In the **Workflow Editor**, click on Mascot node and then, set the following filters **Peptide Cutoff Score**, **Peptide Without Protein Cut**, **Protein Relevance Threshold** and **Protein Relevance Factor** to 0. If these filters are not turned-off in ProMS import options such as predefined False Discovery Rate (FDR) will not be

accuracy.



## Collecting search files

Multiple **Analyses** can be imported at once as long as the corresponding searches were performed with the same search engine and protein database(s).

## Selecting data files (Mascot, Proteome Discoverer and X!Tandem)

Click on *Proceed* next to the **Import multiple analyses** process as shown below. The following form will be displayed to select the source of the search files to be imported.

## Select a File Source for Import of Multiple Analyses

**Possible sources :**

☐ A user directory on server : ppoullet Clean My Directory

☐ Any directory on server\* :    
 \* for bioinformatician only

☐ Mascot server :

☐ Upload Zip archive :  Parcourir...

☐ Upload multiple files : File # 1 :  Parcourir...

Proceed Cancel

Multiple import sources are available:

- **A user directory on server:** Following upload, files are stored in a user-dedicated directory on server. These files will stay available on the server until he decides to delete them; either just after import or later. In the latter case the user can still access his directory for file management purpose by clicking on the **Clean My Directory** button.
- **Any directory on server:** This option is available to bioinformaticians only. The user can provide an path on the server where ProMS should look for search results files.
- **Mascot server:** If a Mascot server is declared in the ProMS configuration file, it can be accessed, searched for specific search results and files directly loaded in the ProMS server directory.

## Select a File Source for Import of Multiple Analyses

**Possible sources :**

☐ A user directory on server : ppoullet Clean My Directory

☐ Any directory on server\* :    
 \* for bioinformatician only

☒ Mascot server :

Server : mascot02 Log files : searches.log

**Search filters :**

-Date range : from 20131125 to 20131125 (yyyymmdd)

-Job range : from F  .dat to F  .dat

-Title contains :  Search

**File info:**

-Name: F020710.dat  
-Status: Available  
-Search Title:  
-User ID: 0

**Expand** **Collapse**

**mascot02**

- ☒ 20131125
  - ☒ F020710.dat
  - ☒ F020711.dat
  - ☒ F020712.dat
  - ☒ F020713.dat
  - ☐ F020714.dat
  - ☐ F020715.dat
  - ☐ F020716.dat
  - ☐ F020717.dat
  - ☐ F020718.dat
  - ☐ F020719.dat

☐ Upload Zip archive :  Parcourir...

☐ Upload multiple files : File # 1 :  Parcourir...

Proceed Cancel

User can search results files by **date** or **job number** range or keywords in the files **search title**. The listing of files is then displayed and grouped by date of creation. Specific information on a file (name, availability, search title and user ID) can be displayed by clicking on the file name. If access to Mascot is restricted (user access setup), the Mascot serverID must be defined in



m ProMS as well (see Account management section above). In this case, only Mascot files accessible to the server will be displayed.

- **Upload Zip archive:** If a large number of files must be imported, they can be uploaded at once as the zip archive. The archive will be unzipped on the server.

- **Upload multiple files:** Alternately, up to 10 files can be uploaded as separate files.

Once one or more files have been selected, click on *Proceed* to initiate file retrieval from the selected source. This procedure may take a few minutes depending of the number and size of the files. Once the transfer is complete, a file import interface will be displayed.

**Important Note:** Most browsers do not support upload of files with (total) size > **2 Gb**. If files larger than 2 Gb must be uploaded, we recommend to use **Google Chrome**. This limitation does not apply when retrieving files directly from a Mascot server.

## Importing analyses

### Import parameters (Mascot, Proteome Discoverer and X!Tandem)

Files retrieved are listed in alphabetical order together with pertinent information about the search performed: the MS file, search type (MS2, MS1 or mix of both), databank(s) and a comment and search file used. See figure below.

**Select Files to be Imported**  
(From 'ppoullet' directory)

☒ Delete imported files afterwards.

<input type="checkbox"/> Search file	Analysis name	Parents	MS file	Search type	Databank(s)	Taxonomy	Title
<input checked="" type="checkbox"/> F012631.dat	F012631	Sample 1 (new)	F4260FD.RAW	MS/MS Ions Search	SwissProt	Saccharomyces Cerevisiae	F4260FD_Node:2
<input checked="" type="checkbox"/> F175887.dat	F175887	Sample 2 (new)	TV130107_HEK01.raw	MS/MS Ions Search	CPS_human	All entries	TV130107_HEK01_Node:2
<input type="checkbox"/> F175992.dat	F175992	= Select =	TV130114_HEK01.raw	MS/MS Ions Search	CPS_human	All entries	TV130114_HEK01_Node:2

**F7841FD.msf**

☒ Request 2 F7841FD Sample 3 (new) Mascot [Show Search Parameters](#)

**F7873FD.msf**

☐ Request 2 F7873FD = Select = Sequest HT [Show Search Parameters](#)

**Mascot** [Show Search Parameters](#)

UniProt-ID: #2: --Choose from list-- #3: --Choose from list--

Threshold score for: 1% FDR (Use ≤ 0 to disable this option)  
Default is set to: 20 for Mascot, 90 for Paragon, 5 for Phenix, 1 for Sequest

[Proceed](#) [Clear](#) [Cancel](#)

\*Applies only to interpretations based on MS/MS

**F8242DL-01.msf**

☐ Request 2 F8242DL-01 = Select =

Databank(s): #1: SwissProt-Mascot (SwissProt)

Description:

Threshold score\*: •If decoy search, estimate threshold  
•Otherwise use default (default)

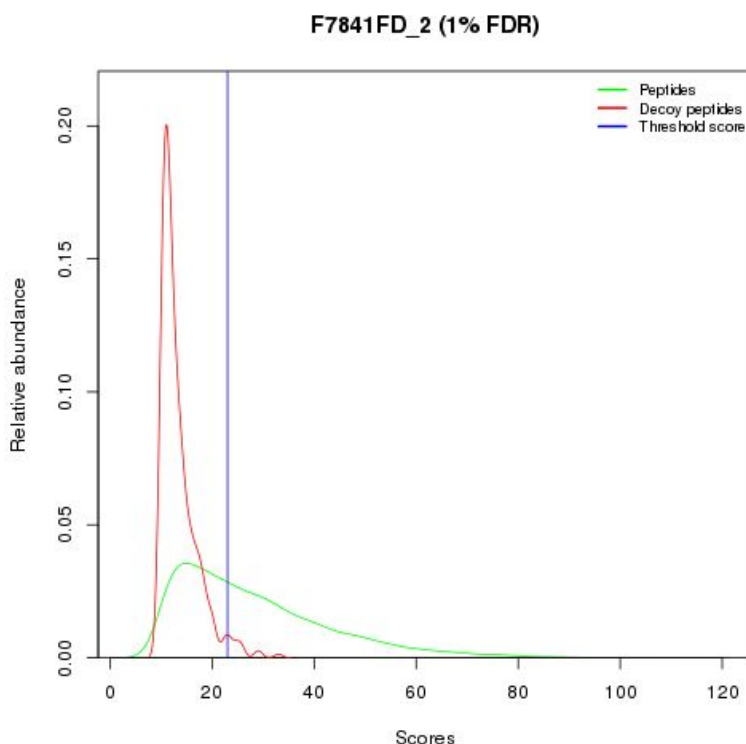
Max. rank\*: 1

Comments:

**Note:** The same Proteome Discoverer msf file can contain multiple searches results (e.g. a search performed with Mascot and another with Sequest). In this case, separate entries will be listed for each search performed together with more details on the parameters used. Each search result can be imported separately and distinct Analyses items will be created.

Proceed as follows to continue data import:

1. Select the files to be imported from the list by checking the boxes on the left-hand side of the files name.
2. Provide a name for the Analysis to be created (**Analysis name** column). This name can be picked by the user or selected to match either the name of the search file or half of the original MS file used for the search.
3. If the parent is an Experiment (or a 2D Gel), each new Analysis must be associated with a specific Sample (or Spot). Pre-existing Samples (or Spots) can be selected from a dropdown menu or they can be created on the fly (Samples only): To create a new Sample, select New from the dropdown menu of **Parents** column. A pop up window will ask you to provide a name for the new Sample.
4. Select the databases to be searched to extract protein annotations (**Databank(s)** field). If multiple databases were searched during the search (possible with Mascot for instance), the corresponding number should also be selected here. All search files to be imported in the same batch should have been performed using the same or equivalent database(s).
5. Define a filtering rule for the data to be imported (**Threshold score**):  
 If a decoy search was performed, data can be filtered based on a user-defined **False Discovery Rate (FDR)** on peptide identification (default is 1%). A threshold score for peptide identification will be determined so that the data imported will (eventually) match the defined FDR value as illustrated in the figure below. Threshold score calculation can either use the **quality** algorithm ([Käll et al. Bioinformatics 2008,25\(7\)](#)), the **Mayu** algorithm or the **DT count** algorithm. In DT count mode, decoy (D) and target (T) peptides are simultaneously sorted in descending score order until the proportion of the 2 populations matches the selected FDR value.



If no decoy search was performed or the FDR value was equal or less than 0, the filtering will be performed according to a minimum (**threshold**) score for peptide identification. A default (search engine-specific) threshold score will be applied unless a different one is provided by the user.

**6.** Select the maximum number of interpeptide ions allowed of the same fragment ion spectrum (**Max. rank**). The default is 1, but up to 10 can be chosen.

**Note:** When performing FDR filtering, it is recommended to set this value to 1 since the FDR calculation is based on 1 interpeptide ion per spectrum.

**7.** Provide optional **description** and **comments**.

**8.** Decide whether the files should be deleted after import or no (**Delete imported files afterwards**). Unless selected for deletion, files will remain on the server for re-import until the user decides to delete them.

**9.** Click on *Proceed* to initiate the data import into the ProMS database.

## Importing MaxQuant data

3 to 5 files are required to import a MaxQuant search/quantification into the ProMS:

1. mqpar.ml (saves all located in the root directory of the MaxQuant search),
2. evidence. (file from the Combined/ directory),
3. Peptides. (idem),
4. proteinGroups. (idem. Optional, only to import protein quantification data),
5. msms. (idem. Optional, only to display peptide fragment ion spectra).



Files 2 to 5 m s be compressed in a common archive before import .

Select an Experiment in which to import the data. From the *Process Analyses* window, select either the **Import multiple Analyses** or **Import quantification** processes and click on *Proceed* next to **Import MaxQuant quantification** to display the form below .

<Image>

Provide the files mentioned above and the protein sequence database(s) used for the search. If contaminants were searched, provide a matching contaminant database. Finally, select whether to search for protein aggregation in the machine groups: in ProMS or MaxQuant. Specify also if to search for protein quantification data (the proteinGroups file must be provided in the archive in this case). Submit the form to start the import. Data import will take a few minutes. Samples, Analyses and an experimental Design, a peptide quantification and 1 or multiple protein quantifications will be added to the selected Experiment according to the information extracted from the files loaded. Peptides and proteins will be automatically validated since they were used in protein quantification.

### Importing DIA quantification data

User can import DIA quantification data from three different software : PeakView , OpenSWATH and Specrona . In the ProMS select an Experiment and from the Process Analyses window, select the **Analysis quantification** process and click on *Proceed* next to **Import PeakView/OpenSWATH/Spectronaut data** to display the associated form.

#### - From PeakView

Two files are required : the Excel worksheet file generated by PeakView and the spectral library . In the PeakView , once the experimental SWATH data analysis is over, you can export the results into an Excel file by clicking on the **Quantitation** tab on the toolbar and selecting **SWATH Processing/Export/All**.

The PeakView search parameters can be filled in the following form to be saved in the ProMS database to ensure retraceability . Then submit the form to launch the import .

#### Select options to import SWATH data

File :  Aucun fichier sélectionné.

Library name :

FDR :  %

**PeakView parameters :**

**Peptide filter :**

Number of peptides per protein :

Number of transitions per peptide :

☐ Exclude modified peptides

**XIC options :**

XIC extraction window (min) :

XIC width :

### From OpenSWATH

The file from OpenSWATH is required, this is a TSV file generated by the last step of the workflow (TRIC). The library has still to be selected to analyse the experimental data is also needed as the associated export parameter file. Then, the user must provide the TRIC method used and the number version of OpenSWATH.

### Select options to import OpenSwath data

Result file :	Parcourir...	Aucun fichier sélectionné.
Library name :	-- Select Library --	
Library export parameter file :	Parcourir...	Aucun fichier sélectionné.
Library export file :	Parcourir...	Aucun fichier sélectionné.
TRIC method used:	LOCAL MST*	*Recommended option
Software version :	ex : 1.2, 2.1.3 ...	
<input type="button" value="Submit"/> <input type="button" value="Clear"/> <input type="button" value="Cancel"/>		

### Running OpenSWATH quantification

The OpenSWATH workflow can be launched directly from mProMS. The process will analyse the experimental files and import the results in a same step. Some parameters are required as the library name, the mXML results files, the iRT file (in TraML format) and the DIA findings file. The user can import his own library considered for OpenSWATH, and choose to merge his results with other existing analysis.

### OpenSwath quantification

Library name :	-- Select Library --		
Library export management :	<input type="radio"/> Use the selected library <input type="radio"/> Import your own library formatted for Openswath		
OpenSwath parameters files :	iRT file : <input type="button" value="Parcourir..."/> Aucun fichier sélectionné.		
	Windows file : <input type="button" value="Parcourir..."/> Aucun fichier sélectionné.		
mzXML files :	<input type="radio"/> Upload multiple files		
	<input type="button" value="Parcourir..."/> Aucun fichier sélectionné.		
	<input type="radio"/> Import from shared data directory		
OpenSwath workflow options:	mz_threshold :	<input type="text" value="0.05"/>	
Pyprophet options:	d_score.cutoff :	<input type="text" value="1"/>	
TRIC method:	LOCAL MST*	*Recommended option	
Merge with other experiment:	<input type="checkbox"/> OpenSwath_all_20transitions_90maxRTdiff		
	<input type="checkbox"/> OpenSwath_all_6transitions_30maxRTdiff		
<input type="button" value="Submit"/> <input type="button" value="Clear"/> <input type="button" value="Cancel"/>			

### From Spectronaut

The import Spectronaut data form is similar to OpenSWATH's form.

## Select options to import Spectronaut data

<b>Result file :</b>	Parcourir...	Aucun fichier sélectionné.
<b>Library name :</b>	== Select Library ==	
<b>Library export parameter file :</b>	Parcourir...	Aucun fichier sélectionné.
<b>Library export file :</b>	Parcourir...	Aucun fichier sélectionné.
<b>Software version :</b>		ex : 1.2, 2.1.3 ...
<input type="button" value="Submit"/> <input type="button" value="Clear"/> <input type="button" value="Cancel"/>		

Once the process is over, samples and analyses will be added to the selected Experiment. Peptides and proteins will be automatically validated. Transition quantification data will be imported but no peptide or protein quantification results as a dedicated pipeline is available in mProMS to perform this task (see the **Protein quantification** chapter below for more information).

### Analysis summary

Files will be imported sequentially and the Analyses (and possible Samples) newly generated will appear in the open navigation window. The Analysis summary is shown below :

#### Analysis **F7228FD**

<b>Name :</b>	F7228FD
<b>Description :</b>	
<b>Position :</b>	1
<b>Start date :</b>	12/12/2012 14:29:05
<b>MS type :</b>	MS/MS Ions Search <input type="button" value="Export Elution Data"/>
<b>Search engine :</b>	Mascot from Proteome Discoverer
<b>Search file :</b>	F7228FD_2.pdm <input type="button" value="Show Search Parameters"/>
<b>MS data file :</b>	F7228FD.RAW
<b>Databank(s) :</b>	•SwissProt-Mascot <b>Type:</b> UniProt - ID
<b>Taxonomy :</b>	Homo sapiens (human)
<b>Labeling :</b>	None
<b>Max. rank :</b>	1
<b>Min. score :</b>	20.52 [FDR-based] <input type="button" value="Show Score Distribution"/>
<b>False discovery :</b>	<b>Targeted FDR:</b> 1% - <b>Observed FDR:</b> [Data not reported: No peptides validated.] <b>Decoy method:</b> Automated decoy search.
<b>Status :</b>	<b>Not validated.</b> <input type="button" value="Show validation history"/>
<b>Comments :</b>	


Diverse information is available on the search as the MS type, search engine, databank(s) (selected in mProMS), protein identifier type, annotation, labeling method if any, threshold score value and strategy (FDR-based or user-defined), validation status,... More search parameters and score distribution for FDR comparison can be displayed on demand.




## Analysis validation

Search results data associated with an Analysis must be validated before being accessible by end-user for further interpretation. Data processing of Analysis data is a multi-step process. Analysis icons are color-coded based on their data processing status to help users to easily determine the validation level of each of them:

: Protein annotation import not completed.

: Data not validated.

: Data partially validated.

: Data validated and reported.

**Note:** Protein annotation import from the database file(s) is not part of the validation process. It is triggered as a background task immediately after search results data import. This process can take several minutes depending on the number of proteins identified. However, validated data cannot be reported (see the Reporting section below for more details) before protein annotation data import has been completed.

## Automated peptide/protein validation

### FDR (False discovery rate) - based validation

FDR-based validation is initiated after the data import step of filter out data below the corresponding threshold score (see **Importing Analyses** above).

### Qualitative validation

### Comparative validation

### Validation templates

## Manual peptide/protein validation

### Peptides selection/exclusion

### Protein exclusion and filtering

## Lower-scoring peptides activation

## Clear peptide/protein selections

## Validation traceability

## Sequence modification validation

### Phosphorylation sites validation with PhosphoRS

**PhosphoRS** ([T. Tassie et al., J. Proteome Res., 2011](#)) is an algorithm used to determine phosphorylation positions on a peptide, based on its sequence and MS2 fragmentation spectrum. This tool is included in our ProMS and can be used to correct imported phosphorylation data.

From **Process analyses** menu, select **Peptide/Protein Selection** process type, and click on **Start PhosphoRS analysis**. The following form is then displayed:

<Figure phosphoRS form>

PhosphoRS parameters can be set in **PhosphoRS Analysis Rules** section:

- :

- **Activation type**: select the fragmentation mechanism of the analysis instrument. PhosphoRS is optimized for each activation type.

- **Mass Deviation**: mass error tolerance when PhosphoRS matches experimental spectra with theoretical spectra.

### Manual validation of modifications

## Validation traceability

## Reporting

# Validated proteins

## Match groups and protein visibility

**Match groups** are a **key feature** in m ProMS. Understanding and setting properly the rules that control match group organization and protein visibility is essential for accurate data analysis and interpretation.

Because the shotgun mass spectrometry technique (used to generate all Analysis data) allows to identify peptides and not proteins, multiple proteins can be matched with the same (set of) peptides. This creates an inherent ambiguity on the identity of the proteins contained in the experimental dataset. m ProMS deals with this problem by organizing the proteins identified in **match groups** representing groups of proteins sharing the same (set of) peptides. To avoid protein inflation, by default only **1 (top) protein** per match group is made **visible** and all others will be **hidden** (not considered as identified in the sample studied). The top protein is also used as **alias** for the match group. Only visible proteins are considered as present in the sample analyzed. Hidden proteins will not be listed (unless the user specifies otherwise) and not used in all subsequent data processing such as protein quantification or Gene Ontology analyses.

## Top protein selection rules

For each match group, m ProMS attempts to select the protein most likely identified by the corresponding set of peptides as top protein using the following rules sequentially until only 1 protein remains:

1. the protein is matched by all peptides in the set.
2. \* the protein is the most often found as top proteins in previously validated analyses.
3. the best scoring protein among those meeting the previous criteria.
4. the protein with best sequence coverage among those meeting the previous criteria.
5. the best annotated protein among those meeting the previous criteria. Annotation quality is estimated as follows:
  - i. SwissProt identifier.
  - ii. rEMBL identifier.
  - iii. none of the following keywords found in the protein description.
    - i. the *hypothetical* keyword is found in the protein description.
    - ii. the *unknown* or *unnamed* keywords are found in the protein description.
  - iv. the protein description is missing.
6. the shortest protein among those meeting the previous criteria.
7. the protein with identifier first in alphabetic order.

\* m ProMS will preferentially select a protein that has been identified often in previous samples.

## Project-wide protein visibility rules

By default, only the top protein of each match group is visible. However, the user can select 1 of 3 predefined project-wide protein visibility rules to alter this default behavior. **Edit** the

corresponding projec . The follo ing sec ion is par of he edi ion form:

<Fig re pro ein isibili r les in projec edi ion form>

These r les are self e plana or and ordered b decreasing s ringenc . Selec ing r le #2 or #3 ill al er m ProMS defa l beha ior and can po en iall lead o m liple isible pro eins per ma ch gro p. Click on **Save** o alida e o r changes. Visibili of all iden ified pro eins ill re-e al a ed based on he r le selec ed e cep if in ol ed in q an ifica ions or GO anal ses. Keep also in mind ha **protein lists** and sa ed **comparisons** (see corresponding chap ers belo ) can be modified b he res l ing changes in pro ein isibili .

### Checking for conflicting match groups

I is possible o check for inconsis enc in ma ch gro ps across m liple anal ses; meaning o de ec pro eins i h inconsis en isibili (**visible s hidden**) across m liple anal ses.

From he *Summary* ie of an projec i em con aining a leas 2 alida ed anal ses, click on he *Scan for conflicts* b on righ of he n mber of isible/ o al pro eins alida ed. A lis of s ch pro eins (if an ) ill be displa ed i h he n mber of anal ses here each pro ein is fo nd isible or hidden as sho n in fig re belo .

<Fig re lis conflic s>

Click on he **[+]** icon o displa he lis of he anal ses in ol ed. From his lis , o can ei her:

- Edi a specific ma ch gro p b selec ing an anal sis (radio lis ) and clicking on he *Edit match group* b on a he op of he able (see he **Manual edition** paragraph belo for help).
- Displa de ailed informa ion on a pro ein in he con e of he anal sis of o r choice b direc l clicking on ha anal sis name.

### Displaying match group composition

**List Proteins** a Anal sis-le el (see he **Project lists** chap er belo for more informa ion) and check *Show Match groups* a he op of he lis . Click on he **[+]** icon ne o he alias pro ein iden ifier o displa he con en of he ma ch gro p. No **[+]** icon indica es ha he pro ein is alone in is gro p. As sho n in he screen cap re belo , isible pro eins are lis ed in bold hile hidden ones appear in ligh fon .

<Fig re ma ch gro p in pro ein lis >



## Manual edition

side from project-wide protein visibility rules, an match group can be manually edited online. If the option is checked, the visibility of an protein in the group. Modification of the protein's name and its associated analysis is not associated with protein qualifications nor GO annotations. In the Match Group box on the bottom of the match group's header editing mode.

<Figure edit Match group par 1)>

In par 1) of the form, the op (alias) protein can be changed and another protein can be made visible meaning that one believes they are indeed present in the biological sample analysed.

<Figure edit Match group par 2) & 3)>

In par 2) of the form, one can propagate the changes made in par 1) upward in the project tree. One can choose to propagate independently the alias, visible and hidden protein selection. Finally - par 3) - one must decide whether or not changes can contradict the current **project-wide protein visibility rule**. If this option is unchecked, any changes made have no effect in the project-wide rule will be ignored.

## Protein list comparison

Full protein-level comparison

Pair-wise protein-level comparison

Pair-wise peptide-level comparison

Saving a comparison

## Search for proteins

## Single protein view

## Peptide quantification

Peptide quantification is a necessary step for peptide-based protein quantification; whether the quantification is based on MS-spectra (SILAC, TnPQ, XIC-based label-free quantification,...) or on MS/MS fragments (iTRAQ, TMT, SWATH...).

## Data import from search results file

Some search results files already contain peptide quantification data. It is also the case for MS/MS fragment-based quantification such as iTRAQ for which the peptide ion intensities are part of the MS/MS spectrum data. Some search results files (Proteome Discoverer MSF or MaxQuant) may also contain peptide XIC data if a quantification was performed after the search process.

When peptide quantification data are contained in the imported search results file(s) in ProMS, you will automatically import these data either during search data import or at the **Validated data Report** step if data validation must be performed. Only quantification data related to validated peptides will be kept (see the [Virtual peptides](#) section below for **important** additional information).

## Data extraction from LC/MS file with MassChroQ

XIC-based peptide quantification can be performed within ProMS whether or not peptide quantification data were already available in the search results file. ProMS uses the tool [MassChroQ](#) (Valo *et al.*, *Proteomics*, 2011) to perform this task. However, the corresponding LC/MS file(s) must be provided in **mzXML** format.

### Managing mzXML files

To manage the list of mzXML files available within a given project, select an **Experiment** or **Sample** from the project navigation window and click on the *Process Analyses* button in the option frame. From the selection menu displayed, select *Analysis Quantification* to display the list of available options. Click on *Proceed* next to the **Manage mzXML files** process as shown below.

**Process Multiple Analyses**

**Process type :** Analysis Quantification ▼

Monitor on-going quantifications

**Peptide Quantification:**

Proceed	Manage mzXML files
Proceed	XIC extraction with MassChroQ

**Protein Label-free Quantification:**

Proceed	Import emPAI data
Proceed	SILAC-based quantification

**Protein Label Quantification:**

Proceed	SILAC-based quantification
Proceed	iTRAQ-based quantification

The following form will be displayed to allow you to either import a new mzXML file or delete already imported ones.

fdf

## Manage mzXML files

**Upload a new file:**

**Files already imported:**

- ☐ G130322\_0174\_c\_ich.mzXML
- ☐ G130322\_0175\_c\_ich.mzXML
- ☐ G130327\_0204\_c\_ich.mzXML
- ☐ G130327\_0205\_c\_ich.mzXML
- ☐ G130402\_0216\_c\_ich\_130402090421.mzXML

**Note:** We perform LC/MS files (RAW & WIFF formats) conversion to mXML in [ProteoWizard](#) tool using default settings. Other format conversion tools are not used. We also recommend not to change the file name (except for the mXML extension) to ease Analysis/mXML file managing in the quantification launch step.

## Running XIC extraction

Go to the *Analysis Quantification* options (as shown above) and click on *Proceed* next to the **XIC extraction with MassChroQ** process to display the form shown below.

### Select Analyses in Experiment **Tg Experiment** for Ext. ion chrom. Quantification

**Name :** Ext. ion chrom. extraction

**Raw-data settings :** Extraction type:  (for mzXML)

**Alignment settings :** Alignment algorithm:  Reference:  Align from 400 to 1200 m/z window

**Peptide selection :** ☐ Extract all charge states of the peptides (even if no MS/MS exists for it)

**Quantification settings :** Type of XIC:

	Analysis	MS type & File	Labeling method	Instrument	Search file & Engine	Databank(s) Taxonomy	Min. score Max. rank	Selected proteins
<input type="checkbox"/>	G130327_0202b_c_ich mzXML file: G130327_0202b_c_ich.mzXML	MS/MS G130327_0202b_c_ich.raw	None	ESI-FTICR	F058447.dat MASCOT	SwisProt-Mascot Rodentia	28.1422 1	585 (1348)
<input type="checkbox"/>	G130327_0203_c_ich_130329121732 mzXML file: G130327_0203_c_ich_130329121732.mzXML	MS/MS G130327_0203_c_ich_130329121732.raw	None	ESI-FTICR	F058448.dat MASCOT	SwisProt-Mascot Rodentia	29.753 1	350 (926)
<input type="checkbox"/>	G130327_0204_c_ich mzXML file: G130327_0204_c_ich.mzXML	MS/MS G130327_0204_c_ich.raw	None	ESI-FTICR	F058449.dat MASCOT	SwisProt-Mascot Rodentia	29.707 1	347 (915)
<input type="checkbox"/>	G130327_0205_c_ich mzXML file: G130327_0205_c_ich.mzXML	MS/MS G130327_0205_c_ich.raw	None	ESI-FTICR	F058450.dat MASCOT	SwisProt-Mascot Rodentia	30.8247 1	345 (911)
<input type="checkbox"/>	G130402_0216_c_ich_130402090421 mzXML file: G130402_0216_c_ich_130402090421.mzXML	MS/MS G130402_0216_c_ich_130402090421.raw	None	ESI-FTICR	F058451.dat MASCOT	SwisProt-Mascot Rodentia	28.584 1	528 (1205)

In the first part of the form, multiple parameters can be set for the extraction:

**-Name** of the quantification: All extraction data collected will be regrouped in a single quantification carrying this name.

**-Extraction type:** Profile or centroid

**-Isotope labeling:** If isotope labeling was performed on your sample, it is possible to use XIC extraction to retrieve it. To do so, you need to choose SILAC. Up to 3 different channels can be retrieved at a time (e.g. heavy, light and medium) and have to be named. For each channel, one or more **quantification label** can be added given the experimental design. Each **quantification label** is linked to a post-translational modification that explains it. Specify the **modification target** on which it occurs (*side chain*, *n-ter* or *c-ter*). If *side chain* is chosen, don't forget to give the residue where the post-translational modification occurs.

Here is provided an example of the filled form in a SILAC experiment where Lysine and arginine residues were designed as heavy isotope. **13C6-15N4** was renamed to **Arg10** for clarity.

**NB :** it is really important to always define a *light* channel if a biological experiment/condition/analysis matches the light version. Otherwise, the *light* version of the peptide will not be retrieved in the end.

**Name :** Ext. ion chrom. extraction

**Raw-data settings :** Extraction type: Profile (for mzXML)

**Isotope labeling :** SILAC

**Channel1 name:** Heavy

**Quantification Label:**

Label Name: 13C6-15N2

Modification target: Side chain

Modification: Label:13C(6)15N(2) / +8.0142 Da on K

**Quantification Label:**

Label Name: Arg10

Modification target: Side chain

Modification: Label:13C(6)15N(4) / +10.0083 Da on R

Add quantification label

Remove quantification label

**Channel2 name:** Light

**Quantification Label:**

Label Name: Light

Modification target: Side chain

Modification: Light / +0.0000 Da on

Add quantification label

**Channel3 name:**

**Quantification Label:**

Label Name:

Modification target: Side chain

Modification: -- Select -- on

Add quantification label

**Alignment settings :** Alignment algorithm: OBI-Warp Reference: -- Select --

Align from 400 to 1200 m/z window

**Peptide selection :** ☐ Extract all charge states of the peptides (even if no MS/MS exists for it)

**Extract XIC traces :** No ☒ Yes

**Quantification settings :** Type of XIC: BasePeakXIC

More settings

**Note :** o r modification is not selectable in the **modification target** option? Check the status of the modification (see **Sequence modification** section below). Maybe the modification or are missing is not added as label. If not, change this and save it by editing the modification.

**-Alignment settings:** Multiple LC/MS runs can be quantified at once. MassChroQ can align all

runs on each feature across different runs. Users must provide an **alignment algorithm** (OBI-Warp or ms2), a **reference** run by selecting the corresponding analysis and an **m/z window** (for OBI-Warp algorithm).

**-Peptide selection:** Whether to extract or not all charge states of a given peptide.

**-Type of XIC** extraction to be performed: basePeak area (most intense peak in the range of masses) or TIC area (summed intensity across the range of masses).

**-More settings** are also available by clicking on the corresponding button.

Finally, click on *Launch Quantification* to start the extraction. A pop-up window will appear to allow you to monitor the quantification progress. XIC extraction is a long process that can last up to an hour or more depending on the number of Analyses to be aligned, the complexity of the LC/MS run and the computer power available. You can continue using mProMS in the meantime and enroll other quantifications. All on-going quantification jobs are displayed in the **monitor quantifications** window (see figure below). As new jobs are launched or old ones completed, they will appear or disappear from the list. Additionally, if an error occurs during quantification, a message will appear for the corresponding job. The user will be able to display the content of the error message and delete the failed job and all associated temporary data.

<Figure Monitor Quantification window>

If this window is closed inadvertently (closing it has no effects on the on-going jobs) or did not appear (pop-up window for mProMS URL must be enabled in your browser), it can be displayed again by clicking on the *Monitor on-going quantifications* button in the *Analysis Quantification* options (Process Analyses > Analysis Quantification).

**Note 1:** Check the [MassChroQ manual](#) for help on setting these parameters properly.

In the second part of the form, you must select the Analyses for an extraction will be performed and associated with the proper mXML file. If the mXML file name matches the **MS data file** recorded for the **Analysis**, mProMS will do the job for you.

**Note 2:** Only runs with reproducible retention-times (e.g. biological or technical replicates) should be selected for alignment. Runs potentially of different sets of features (e.g. sample fractions separated on a gel) should be extracted separately.

## Virtual peptides and proteins

During the quantification process, intensities of parent ions can be calculated even though the corresponding peptide did not end up in the list of peptides validated. For instance, in the case of a SILAC-labeled analysis, the label-free form of a peptide can be validated by itself is labeled correctly; either because the latter falls under the threshold score set or was not identified at all. However, these data are not available for the quantification since both peptide forms are required for ratio calculation. mProMS solves this issue by adding these missing peptides to the list of validated peptides by giving them a special status: **virtual peptides**. This strategy of peptide addition also applies to alternative charge states of a given peptides (e.g. if the 2+ charge state

of a peptide is validated, all other quantified charge states will be added as virtual peptides). Virtual peptides remain hidden unless their presence is required for proper data interpretation. When applied to a label-free quantification where 2 or more analyses are aligned, a peptide validated in the reference Analysis but missing (or not validated) in an aligned one can be rescued as a virtual peptide. If this peptide does not belong to an validated proteins of the aligned Analysis, the protein(s) matching this peptide in the reference will be added to the Analysis aligned as **virtual protein(s)**. Virtual proteins appear in italics in most protein lists.

## Displaying peptide quantification data

Once the peptide quantification data are available (after a **Report** for Search file extraction or an **XIC extraction** within ProMS), they can be displayed for individual **Analysis** by selecting the corresponding Analysis in the Project navigation frame and clicking on the *Internal Quantifications* button in the option frame. From the window displayed in the result frame, select the name of the quantification (in the Peptide quantification block). A window similar to the one below will be displayed showing a summary of the quantification parameters used (if any: no parameters are displayed in case of a direct extraction from a search results file) and a list of proteins with identified peptides and corresponding XICs or fragments area for DIA extraction. In case of labeled quantification, peptide sets (label isoforms) are grouped into a single peptide row. The peptide set sequence, variable modification, position, charge, score(s) and XIC(s) are displayed.



## Single-Analysis Peptide Quantification

Select: **SILAC 2plex (Arg10, Lys6) (Custom) [SILAC]**

Label :	SILAC	
Channel :	1	2
Signal name :	- MMS	+MMS
Isotope(s) :	None	Lys6: Label:13C(6) (K)
Total signal :	4.13e+09	2.93e+09

**TIM44\_YEAST:** RecName: Full=Mitochondrial import inner membrane translocase subunit TIM44; AltName: Full=Inner membrane import site protein 45; Short=ISP45; AltName: Full=Membrane import machinery protein MIM44; AltName: Full=Mitochondrial protein import protein 1  
*Saccharomyces cerevisiae* (431 aa)

#	Peptide sets	Start	Charge	Scores	- MMS	+MMS
1	AQRGSTIVGK	98	2 <sup>+</sup>	34.95/-	3318730	2113160
2	EYSEIDDGESSRYGGFITK	158	2 <sup>+</sup>	50.71/44.31	10081900	6240680
3	EYSEIDDGESSRYGGFITK	158	3 <sup>+</sup>	-/-	3749330	2450730
4	KLDESFEFVRQTK	142	2 <sup>+</sup>	51.11/56.01	32805400	19407800
5	KLDESFEFVRQTK	142	3 <sup>+</sup>	-/-	80268600	48058300
6	KTGETMEHIATK	111	2 <sup>+</sup>	78.48/63.09	3106750	1802680
7	KTGETMEHIATK + Oxidation (M:6)	111	2 <sup>+</sup>	61.28/56.6	1453800	820991
8	KVEDFKEK	221	2 <sup>+</sup>	39.32/-	6068770	2933450
9	LGSEAYKK	82	2 <sup>+</sup>	45.56/51.64	13590500	8835730
10	LWDESENPLIVVMRK + Oxidation (M:13)	242	2 <sup>+</sup>	32.59/-	1222880	851545
11	LWDESENPLIVVMRK + Oxidation (M:13)	242	3 <sup>+</sup>	-/-	6090870	3179090
12	SNEDAGTAVVATNIESK	199	2 <sup>+</sup>	98.55/88.3	89985600	59983200
13	SNEDAGTAVVATNIESKESFGK	199	2 <sup>+</sup>	80.62/64.9	15554300	9468700
14	SQELQENIK	65	2 <sup>+</sup>	44.17/-	64900700	46250700
15	TGETMEHIATK	112	2 <sup>+</sup>	56.32/-	2087380	1623040
16	TGETMEHIATK + Oxidation (M:5)	112	2 <sup>+</sup>	36.41/42.8	2555150	1624440
17	TLQDASGKLGSEAYK	74	2 <sup>+</sup>	98.2/75.46	7019810	4470020
18	TLQDASGKLGSEAYK	74	3 <sup>+</sup>	-/-	10041000	6254880
19	TLQDASGKLGSEAYKK	74	2 <sup>+</sup>	67.4/-	7262500	3228410
20	TLQDASGKLGSEAYKK	74	3 <sup>+</sup>	-/-	20886900	9840560
21	TVVGRSIQSLK	229	2 <sup>+</sup>	32.13/34.17	65403200	44655600
22	VGGFFAETESSRVYSQFK	261	2 <sup>+</sup>	44.78/30.41	1827420	1197130
23	VGGFFAETESSRVYSQFK	261	3 <sup>+</sup>	-/-	4285890	2825770

Legend: Case of a direct extraction of SILAC-labeled peptide XIC from a search result file

Virtual peptides can be easily identified as they do not have a score.

## Multi-Analysis Quantification

Select: **XIC Tg ms2 [Ext. ion chrom.]**

XIC quantification Name :	XIC Tg ms2	<input type="button" value="Export Results"/>
Raw-data settings :	Extraction type: profile (for mzXML)	
Alignment settings :	Alignment algorithm: ms2 Reference: G130322_0175_c_ich Tendency: 10 - Smoothing: 5 (MS/MS) and 3 (MS)	
Charge states :	Validated charge states extracted	
Quantification settings :	Type of XIC: sum	
	<input type="button" value="More settings"/>	

Legend: Case of a MassChroQ extraction with alignment of multiple Analyses

# Protein quantification

## Absolute abundance quantification

### emPAI (label-free)

The **E**xponentially **M**odified **P**rotein **A**bundance **I**ndex (emPAI) is a spectral-count method that estimates the relative quantification of proteins in a complex mixture ([Hishima et al., Mol Cell Proteomics, 2005](#)) based on protein coverage by peptide matches. mProMS uses the built-in implementation of the Mascot server 2.3 software which is a slightly modified version of the original emPAI value (for more details, have a look at [masco help](#)). As this value is retrieved from Mascot web-server, this label-free method can only be applied to Analyses generated from Mascot DAT files directly imported from a connected Mascot server.

...

### SIn (label-free)

The **S**pectral Index **N**ormalized (SIn) is a normalized label-free quantification method which combines three abundance features: peptide and spectral count with fragmentation (MS/MS) intensity (for more details, see [Griffin NM et al., Nat Biotechnol., 2009](#)). This label-free method is currently available only for Analyses generated from Mascot DAT files. Support for other search engines is planned in future versions of mProMS.

### MaxQuant: Intensity, LFQ, iBAQ

### Displaying single abundance quantification data

...

## Relative abundance quantification

### Single-Analysis quantification (labeled)

If a labeled Analysis has to be quantified, labeling parameters and all peptide XIC data should be readily available in the corresponding search results file. Therefore a search for a target protein quantification can be performed as follows: Go to the *Analysis Quantification* options (Process Analyses > Analysis Quantification) and click on *Proceed* next to the **(SILAC/iTRAQ)-based quantification** process to display the quantification form shown below.

## Protein Quantification based on SILAC-labeled Peptides from Analyses in Sample **Detection2**

**Name :** SILAC-based protein ratios

**\*Labeled states :** #1: WT #2: Mutant

**Peptide selection :** Specificity: Proteotypic Missed cleav.: Allowed PTMs: Not allowed Charges: All Sources: All

**Quantification settings :** Bias correction: Scale normalization

☐ Avoid infinite ratios whenever possible (Always true if more than 2 states selected).

**Advanced settings:**

-Variation coefficient threshold between replicates: Auto

(Ignored if no replicates)

-☒ FDR control to 5 % Method: Benjamini-Hochberg

-p-value threshold for outlier detection: 0.05

-Alternative hypothesis for comparison: Two-sided

-Confidence interval on protein abundance: 0.95 (0-1)

\*Each State will be used as reference for all following States

<input type="checkbox"/>	Analysis	MS type & File	Labeling method	Instrument	Search file & Engine	Databank(s) Taxonomy	Min. score Max. rank	Selected proteins
<input checked="" type="checkbox"/>	F4628MT	MS/MS F4628MT.RAW	SILAC 2plex (Arg10, Lys8) (Custom)	ESI-TRAP	F4628MT_2.pdm MASCOT	NCBI-Mascot All entries	20 1	878 (1007)

**Name:** A name for the quantification.

**Labeled states:** Select the different conditions to be compared. Available labeled states are identified based on labeling design extracted from the search results file. Each condition defined will be used as a reference for the following one(s). 1 state is still associated with 1 condition. However, if more than 2 states are identified (e.g. iTRAQ 4/8-plex) an additional option will be displayed for grouping different states as replicates of the same condition. In addition, if more than 2 conditions are defined, all corresponding ratios will be calculated except reverse ratios (cond B/cond A but not cond A/cond B).

**Note:** It is possible to quantify multiple Analyses at once. Make sure they share identical labeling design. If not, they should be quantified separately.

Multiple filter can be applied on **Peptide selection:**

**Specificity:** Whether to restrict quantification to proteotypic peptides or not.

**Missed cleav.:** Include or no miss-cleaved peptides.

**PTMs:** Peptides with sequence modification can be allowed, not allowed or excluded in comparison to corresponding non-modified peptide.

**Charges:** Include all charge states of a peptide sequence or restrict to sequences having the best signal (selecting peptide with highest XIC value).

**Sources:** If the search results files is a merge of multiple LC/MS runs (e.g. Proteome Discoverer), select peptide sequences from all runs or select only the one with best signal.

**Quantification settings:** Additional options are available to control experimental bias, outliers detection and differential analysis.

**Bias correction:** Select whether to correct or not for signal bias between labeled states and which method to apply: If **Scale normalization** is selected, the assumption is made that the total XIC signal between all states should be equal. Alternatively, if **Reference protein(s)** is selected, a pre-recorded **List** of proteins must be provided. When using this option, it is assumed that a subset of proteins (e.g. Housekeeping proteins) is unchanged amongs all states and therefore only the sums of the XICs matching these proteins are set equal. In both cases, a state-specific correction factor is computed and applied to each individual peptide XIC.

**Avoid infinite ratios:** Infinite ratios (log ratios) can occur when XIC values are missing in 1 of the 2 conditions being compared. When a mixture of normal and infinite peptide ratios exists for the same protein, mProMS must either select the most abundant peptide ratios or quantify the protein (e.g. select protein ratio to +/- infinite (log ratios) if more than 50% of matching peptides have infinite log ratios) or only select the normal ratios even if there are less frequent than the infinite ones (**do not avoid infinite ratios whenever possible**). This latter option is automatically selected if more than 2 conditions are compared or pre-enrichment is used.

More **advanced settings** can be used for **outlier** detection, comparison hypothesis (Two-sided/Lesser/Greater), **FDR** control, ...

Finally, **select the Analysis(es)** to be quantified. If multiple peptide quantification datasets are available for an Analysis, one must be selected. Click on the *Launch Quantification* button. Multiple quantifications will be queued and processed as up to 3 parallel jobs. As described above for [Peptide quantification](#), a pop-up window will appear in the list of all jobs launched in their progress status.

## Design-based quantifications

The use of a design for a quantification is highly recommended, even if it requires only single labeled analysis. It is mandatory to create a design for a quantification that requires more than 1 analysis. Designs are automatically generated when importing protein quantification data from MassQuant analyses.

Conditions

Observations

## Displaying relative abundance quantification data

### Label-free quantifications

Label-free quantifications are methods that allow to determine the relative amount of proteins in two or more biological samples in the absence of stable isotope or chemical labeling. It is based on precursor signal intensity or the number of spectra made for each peptide of a protein.

Here is a brief description of several methods available in mProMS that can be found from the panel button **Process Analyses** and then, **Analysis Quantification**.

### TnPQ

Silva et al. showed in their work on a Q-ToF peptide instrument that it is possible to quantify unknown protein samples in a known unified signal response factor in absolute manner ([Silva et al., Mol Cell Proteomics, 2006](#)). Then, the **Top 3 Protein Quantification** (T3PQ, [Grossmann et](#)

[al. J Proteomics, 2010](#)) extended this method of ion ratio comparisons. The method premises have for each protein identified by a set of peptides, the average of the three most efficient ionized and therefore highest MS signals directly correlated with the input amount of the corresponding protein. In mProMS, we extended this definition to all available peptides for a given protein and called it TnPQ.

### Select Analyses in Design Test\_Design for Protein-Ratio Quantification

Steps involved in TnPQ comparison:

**Step 1 :** retrieval of all available XICs (area) of each peptide of the protein for all conditions

**Step 2 :** removal of incomplete peptide information i.e. peptide with no XIC information in at least one of the replicates of a condition will be removed

**Be careful :** when creating a quantification, avoid to add too many conditions because you will lose a lot of peptide information given the fact that all conditions must provide a XIC for a peptide to be considered more further

**Step 3 :** if a bias correction setting was selected (scale or reference protein normalization), a normalization step is introduced by computing biases images on unique peptides (for more information in this step, please, go to [Yang et al. 2002](#), scale normalization part). All XIC are divided by those bias factors

If **None** was chosen, nothing is done on the data

**Step 4 :** removal of extreme XIC values (outliers) based on the coefficient of variation (standard deviation divided by the mean) of all identified peptides along the replicates in the conditions.

**Step 5 :** compute for each protein the geometrical mean of peptide XICs

**Step 6 :** quality control of the data (normalities on the data and variance sameness)

**Step 7 :** compute the ratio between paired conditions and make a test to assess equality of mean depending on the design made before

for 2 conditions : use Student *t-test* comparison (or Welch *t-test* if variances are not the same)

for more than 2 conditions : use Tukey HSD (honest significant difference) test

**Step 7' :** if chosen, adjust p-values to control FDR level

**Comparing multiple protein quantifications**

**Exporting multiple quantifications**

## **Dealing with PTMs**

**PTMs relevant to project**

**Displaying PTMs distribution**

**Comparing modification sites from different project items**

**Quantifying modification sites**

## **Biological samples management**

**Properties**

**Treatments**

**Recording a biological sample**

**Linking biological samples to MS Analyses**



# Gene Ontology analyses

Different types of analyses using **Gene Ontology (GO)** can be performed on validated protein lists. The GO project provides a controlled vocabulary of terms for describing gene products such as proteins. For more details, see the [GO website](#). A GO analysis can regroup proteins into standardized categories of terms belonging to 3 domains: **Biological Process**, **Cellular Component** and **Molecular Function**. In mProMS, all GO analyses need 2 types of GO files that are managed from **GO files management** section (See corresponding chapter below for more information).

- **Ontology file:** the file contains all term descriptions and their relationships between each other
- **Annotation file:** the file maps each protein identifier to the most specific terms that characterize the protein.

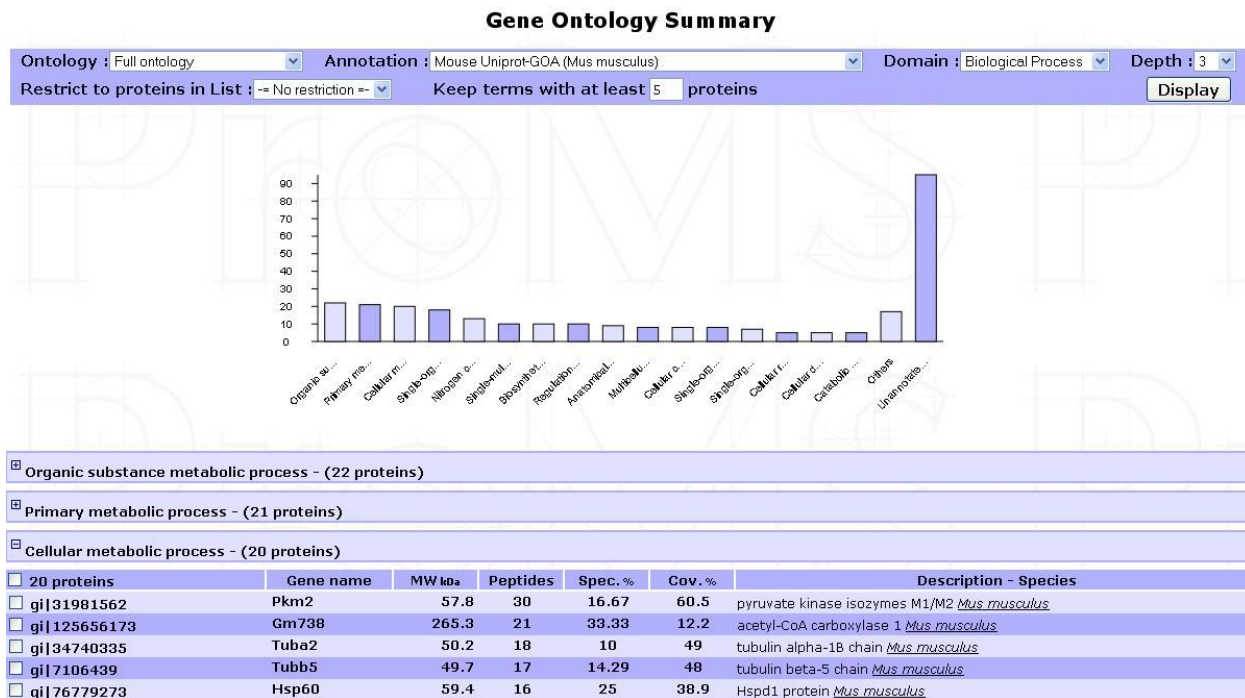
## GO summary

The GO summarization tool can be used to simply regroup proteins sharing common GO terms. This tool can be run from the **option frame** on an **project item**, by clicking on the *Gene Ontology summary* button. The following form is then displayed:

**Gene Ontology Summary**

Ontology : -- Select an ontology file --	Annotation : -- Select an annotation file --	Domain : Biological Process	Depth : 2
Restrict to proteins in List : -- No restriction --	Keep terms with at least 5 proteins	<input type="button" value="Display"/>	

- **Ontology:** the file containing terms that will be used to regroup proteins.
  - **Annotation:** the file containing protein annotations to GO terms.
  - **Domain:** select one of the 3 GO domains the analysis will be focused on.
  - **Depth:** only terms at the specified depth in the GO graph structure will be used. Depth is calculated by counting the distance between a term and the root term of the corresponding ontology domain. If a high depth is selected, a very large number of terms will be displayed and the results may be difficult to read.
  - **Minimal protein per term:** if a selected term contains less proteins than this specified value, this term will be ignored and the matching proteins will be added to the **Other** category. This parameter is optional.
- Click on *Display* to launch the process. After a short calculation time, results are displayed as shown in the example below :



An interactive bar plot shows each term frequency. Click on a bar to display the proteins mapped to the corresponding term. Each protein group can also be viewed by browsing the list of terms displayed below the plot.

## GO enrichment analysis

Enrichment analysis is performed to determine which GO terms are significantly enriched in a **tested set** of proteins when compared with a given **background set** (e.g. the whole proteome of the species studied). All terms will be tested regardless of their depth.

In mProMS, GO enrichment analysis is calculated with the GO::TermFinder package developed for perl ([Boile et al., Bioinformatics, 2004](#)). Briefly, a *P*-value using a hypergeometric distribution is computed to determine whether any GO terms annotate a specified list of proteins at a frequency greater than what would be expected by chance. Multiple hypothesis correction is available with FDR computing.

This tool is accessible by clicking on an experiment and selecting the *Start GO Analysis* button in the option frame. The following form is then displayed:



- **Name:** Provide a name for the enrichment analysis. The analysis is saved and can be retrieved by this name in the **GO analyses** tree displayed in the **sub-navigation** frame.
- **Description:** optional description of the current analysis.
- **Ontology File:** the file containing term relationships.
- **Annotation:** the file containing protein annotations to GO terms.
- **Domain(s):** Select one or more domains of interest.
- **Advanced parameters:**
  - **Estimated number of proteins in organism:** If the background population consists of the whole proteome (more exactly the whole protein set contained in the annotation file), this value can be set to calculate properly the enrichment ratio of GO terms in the tested protein set(s), supposing that the annotation file is incomplete. This option artificially adds unknown proteins to the background.
  - **Background population:** Select the population on which the tested protein set will be compared. A pre-installed **custom list** can be selected, or a local file can be selected instead. This file must contain all protein identifiers that compose the background (1 identifier per row). These identifiers must match the ones contained in the annotation file. If selected background is set to **Unspecified**, the whole protein set contained in the annotation file will be used as background. In this case, be sure that the annotation file contains only proteins from the current species. This can be considered as a whole proteome background if the annotation has a very good coverage of current species proteome. The background population selection strongly affects the significance of terms and must be chosen carefully and coherently with other biological questions.
  - **Statistical settings:** these settings can be set to control the significance cutoff of GO terms. False Discovery Rate (FDR) or p-value criteria can be selected.
  - **Show non-significant terms in graph:** If this option is disabled, non-significant terms will be represented by small dots in graphical view. This can increase significantly the visibility

of the graph if the database contains a large number of significant terms.

- **Include only proteins with at least n peptide(s):** Proteins which contains less peptides than the already specified will be excluded from the result set.

- **Select a protein set:** Select the protein set to be tested. It can be selected from an project item or collection.

Once all parameters have been set, click on *Start Analysis*. The computation may last several minutes depending on the sizes of the protein sets being compared.

The results are directly displayed after the process but can also be accessed later on by selecting the analysis name in the **GO analyses** tree displayed in the **sub-navigation** frame..

The screenshot shows the 'Protein Identification > GO Enrichment analysis' interface. The sidebar on the left contains a 'Demo project' tree with items like 'Protein Identification', 'Protein Labeled Quantification', 'Protein Label-free Quantification', 'Exp. with gels', and 'Peptide Mass Fingerprint'. Below this is a 'Sub-navigation frame' with a 'View: GO Analyses' dropdown and 'Expand'/'Collapse' buttons. The main content area is titled 'Protein Identification > GO Enrichment analysis' and has tabs for 'Summary', 'Biological Process', 'Cellular Component' (selected), and 'Molecular Function'. Under 'Cellular Component', there are 'Table View' and 'Graphical View' buttons. The 'Graphical View' displays a list of enriched terms, including 'Cell', 'Cell part', 'Cortical actin cytoskeleton', 'Cytoplasm', 'Cytoplasmic part', 'Cytoskeletal part', 'Cytosol', 'Cytosolic large ribosomal subunit', 'Cytosolic part', 'Cytosolic ribosome', 'Cytosolic small ribosomal subunit', 'Intermediate filament', 'Intermediate filament cytoskeleton', 'Intracellular', 'Intracellular membrane-bounded organelle', 'Intracellular non-membrane-bounded organelle', 'Intracellular organelle', 'Intracellular organelle lumen', 'Intracellular organelle part', 'Intracellular part', 'Keratin filament', 'Large ribosomal subunit', 'Macromolecular complex', 'Membrane-bounded organelle', 'Membrane-enclosed lumen', 'Non-membrane-bounded organelle', 'Nuclear lumen', 'Nuclear part', 'Nucleolus', 'Nucleoplasm', 'Nucleus', 'Organelle', 'Organelle lumen', 'Organelle part', 'Prefoldin complex', 'Ribonucleoprotein complex', 'Ribosomal subunit', 'Ribosome', 'Signal recognition particle', 'Signal recognition particle, endoplasmic reticulum targeting', and 'Small ribosomal subunit'.

For each domain, results can be displayed in 3 different views accessible at the top of the page:

- **Cloud view:** Highly significant terms (low p-value) are represented with a large font, and less significant terms with a small font. The proteins mapped to a term can be listed by clicking on each term.

- **Table view:** More details can be viewed in table format which contains the p-value and enrichment ratio of each term.

- **Graph view:** Displays a graph of the significant terms as **nodes** with their relationships as **edges**. Each node color is based on the corresponding term's p-value significance. Proteins that are mapped to a term can be viewed by clicking on the corresponding node.

## Quantitative gene enrichment analysis

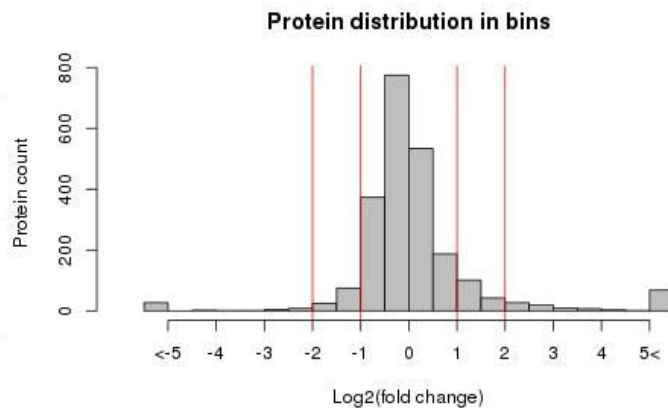
When a quantification is available, a quantitative gene enrichment analysis can be performed as it was originally done for SILAC experiments in an article recently published by [Pan C et al., MCP, 2009](#).

The quantified proteome is divided into five bins corresponding to log2 ratios or bin proportion. Enrichment of GO terms in each bin is then calculated compared to a provided background and a cluster analysis allows to visualize a heatmap of enriched GO-terms in all bins.

Here is how you should proceed to do it.

This option is accessible by clicking on an experiment and selecting the **Start Q. GO Analysis** button in the option frame. After loading a protein set of an **Analysis** or a **Design** related quantification, you need to select the parameters in the following form:

### Gene Ontology Enrichment Analysis On Quantification Data [?]



Bin proportions (%)

Log ratio thresholds

Name:

Description:

Quantification: **SAK vs WT**

Ratio:

Peptides: ☐ Include only proteins containing at least  quantified peptide(s)

Protein-ratio p-value threshold:  (leave empty for no threshold)

Ontology file:  Depth:

Annotation:

Domain: ☒ Biological Process ☐ Cellular Component ☐ Molecular Function

Advanced parameters:

Background population: ☒ All **quantified** proteome  
☐ All **annotated** proteome  
☐ Select a List:   
☐ Upload a local file:  Aucun fichier sélectionné.

Enrichment test statistical settings:

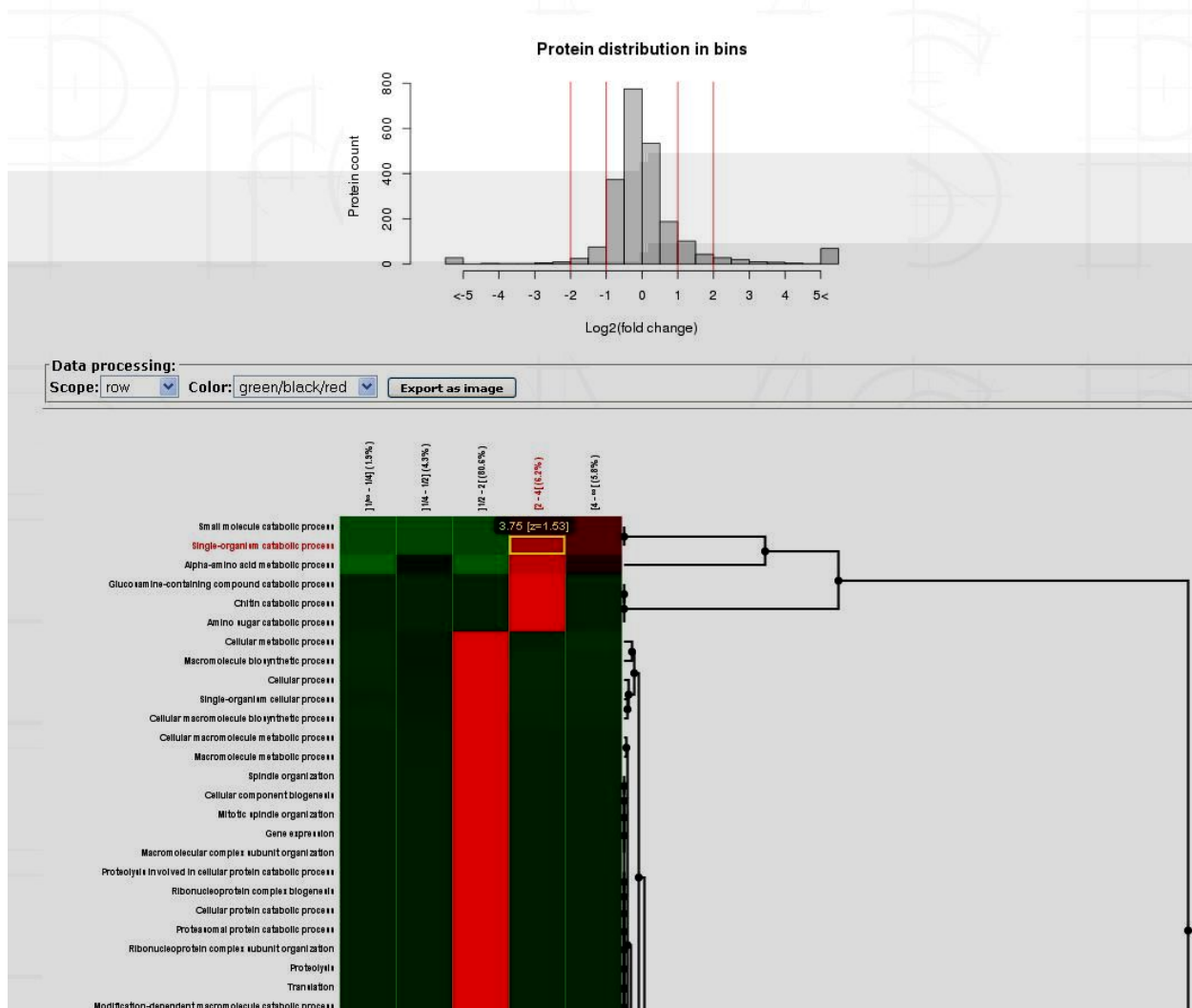
☒ Control FDR at  % with  method  
☐ Use a p-value threshold:  with ☒ Bonferroni correction

- **Name:** Provide a name for the enrichment analysis. The analysis is saved and can be retrieved by this name in the **GO analyses** tree displayed in the **sub-navigation** frame.
- **Description:** Optional description of the current analysis.
- **Ratio:** Choose the ratio considered for the enrichment in the quantification (like heavy/light for SILAC experiments).
- **Peptides:** Make a threshold upon the number of peptides selected to compute the ratio.
- **Protein-ratio p-value threshold:** Make a selection on the associated p-value of the ratio.
- **Ontology file:** The file containing term relationships.
- **Annotation:** The file containing protein annotations to GO terms.
- **Domain:** Select one domain of interest.
- **Advanced parameters:**
  - **Background population:** select the population of which the tested protein set will be compared. See **GO enrichment analysis** section for complete recommendations.
  - **Enrichment test statistical settings:** these settings can be set to control the significance cutoff of GO terms. False Discovery Rate (FDR) or p-value criteria can be selected.

When the enrichment is done, you can get information of the GO-Analysis by clicking on the sub-navigation frame item generated and *Summary*.

Click on the *Heatmap* button to see the output you can get :

## Fold Change-based Gene Ontology Enrichment



Each row represents a GO-Term and each cell is the  $-\log_{10}$  of the  $P$ -value of the enrichment test for the GO-Term in the specific bin (p > 0.1 and then *log-transformed* to 0 if has no enrichment/significance in the bin). Each line is *z-scored*. Then, these *z-scores* are clustered by one-dimensional hierarchical clustering using the function *hclust* in R (the distance function used is theclidean and the agglomeration method used is average).

The heatmap is interactive and can be exported as a jpeg image. Clicking on a cell opens the frame and provides the list of proteins containing the annotated GO-Term in the bin.



# **Exploratory analyses**

**Launching exploratory analyses**

**Principal Component Analysis (PCA)**

**2D-Clustering**

# Annotation data management

## Sequence databanks

The sequence databanks used by the search engines must be referenced in m ProMS so that protein annotations (identifier, description, species and sequence) sometimes not present in search result files (eg. Masco) can be retrieved from the corresponding fasta file during analysis import. A referenced sequence databank is also associated with a specific parser to handle all the m ProMS to properly match and extract the annotation from the fasta file.

### Databank types

Multiple databank types are available in m ProMS depending on the proteomic resource used to download the fasta file:

#### + UniProt:

Typical entry in fasta file:

```
>sp|P15311|EZRI_HUMAN Ezrin OS=Homo sapiens (Human) GN=EZR Ezrin
UniProt - ALL
```

This databank type will extract the entire identifier block `sp|P15311|EZRI_HUMAN` as protein identifier, `Ezrin` as description and `Homo sapiens` as species.

#### UniProt - ACC

Same as above except that the protein identifier used is the **Uniprot accession number** `P15311`. This type is also compatible with the Uniprot isoform naming `ACC#-n`.

#### UniProt - ID

Same as above except that the protein identifier used is the **Uniprot identifier** `EZRI_HUMAN`.

#### SWISSPROT/ rEMBL #1, #2 and #3

These 3 types are equivalent of the 3 UniProt types described above except that they recognize the obsolete fasta entry format:

```
>sp|P15311|EZRI_HUMAN Ezrin (p81) (Cytovillin) (Villin-2) - Homo
sapiens (Human)
```

#### + NCBI:

Typical entry in fasta file:

```
>gi|125987826|sp|P15311|EZRI_HUMAN Ezrin (p81) (Cytovillin)
(Villin-2) [Homo sapiens (Human)]
NCBI - ALL
```

Extracts `gi|125987826|sp|P15311|EZRI_HUMAN` as protein identifier, `Ezrin (p81) (Cytovillin) (Villin-2)` as description and `Homo sapiens` as species.

#### NCBI - GI

Same as above except that only the gi number (`gi|125987826`) is kept as protein identifier.

#### + \*IPI:



Typical entry in fasta file:

```
>IPI:IPI00843975.1|SWISS-PROT:P15311| Tax_Id=9606 Gene_Symbol=EZR  
Ezrin
```

IPI databank

Ezrin IPI00843975 as protein identifier, Tax\_Id=9606 Gene\_Symbol=EZR Ezrin as description and 9606 as species.

\* The IPI resource is no longer maintained. We do not recommend using fasta files from this resource in other MS search engines.

#### + Undefined source:

Protein (User-defined)

This type can be used as a temporary solution for an unknown or cross-compatible entry:

```
>pipe_separated_identifier_block any text
```

Ezrin the identifier block as protein identifier and an following text as description. No species is recorded.

#### + Other types:

New databanks types can be easily added on demand. Please contact your local ProMS administrator or email [proms@cric.fr](mailto:proms@cric.fr) for more information.

### Listing databanks

Only bioinformatician and mass spectrometry managers with granted appropriate privileges can access and manage the protein databanks.

From **myProMS main window**, select *Annotation data* and follow the *Sequence databanks* link. All available databanks are listed in alphabetic order with a short summary of information as shown in the screenshot below. From this window, you can either **add** a new databank, **edit** or **delete** an existing one.

<Figure of databanks list>

### Adding a new databank

From the databank list window, click on the *Add new Databank* button on the top or bottom of the list. The following form will be displayed:

<Figure Add new Databank>

Fill out the form to provide information on the databank you want to add. In particular, you must select the databank type so that the server will know how to extract the protein annotation from

he file. Information on the corresponding parser file is then displayed to help insure the right database is selected. You may also provide a **fasta** file containing the protein data. There are multiple ways to do so:

1. Use a database already referenced by Masco: the ProMS allows you to directly use fasta files stored on the Masco server to avoid data duplication. In this case, the database will be automatically synchronized when updated by Masco.
2. Use a file from a dedicated directory on server (e.g. files previously uploaded by FTP or the directory is shared between local computer and server).
3. Upload a fasta file from your computer.
4. Download the file from the internet: You may provide an HTTP or FTP link to the file.

For the last 3 options, normal and gzip-/zip-compressed files are handled.

If the database contains both large and deco sequences, this may be specified as well as the deco tag used (eg. REV\_).

For the first 2 options (except if a compressed file is used in the 2nd option), it is possible to select the type of annotation rules selected before actually creating the new database: Select a database type, the file to be used and click on the *Test rules* button. Annotations from positions 10 entries from the file will be extracted using the selected rules and displayed. Select another set of rules and rerun if the extraction did not match your expectations.

If the database is species-specific, it is recommended to provide the species scientific name even if already specified in the protein entries of the fasta file.

Click on the *Save* button to submit the database creation form. Once the process is completed, you will be redirected to the database listing window.

## Editing a databank

You can edit all information concerning an existing database except its annotation type, the sequence file used and whether it contains deco sequences.

From the database listing window, click on the *Edit* button on the right side of the database row. A form similar to the one used to add a database will be displayed. Make the desired changes and click on the *Save* button to validate your changes.

You can reset your annotation rules as described above for database addition but regardless of the database file origin.

If your database references a Masco file, it is possible to check if the file has been updated on the masco server by clicking on the *Check for update* button. This can take a few minutes for large database files such as NCBI databases. Checking for file update is nonmandatory since it will be performed automatically once the database is selected during an Analysis import.

## Deleting a databank

We recommend to delete an database that will no longer be used to keep the list displayed as short as possible. Deletion of a database has no effect on the availability of information of analyses using this database. A database can be deleted at any time except during import of analyses using this database. From the database listing window, click on the *Delete* button on the right side of the database row. A prompt will be asked to confirm your decision.

## Spectral (SWATH) libraries management

### Listing spectral libraries

From **myProMS main window**, select *Annotation data* and follow the *SWATH libraries* link. All available libraries are listed in alphabetic order with the following information as shown in the screenshot below. On the left side are listed all existing libraries with the possibility to delete, export, edit or update them. On the upper part of the window, you can either add a new library, merge existing ones, or initiate rerunning processes. You can also search if some desired proteins are existing in one library thanks to the search link and restore the previous version of an updated library.

### List of spectral libraries

**Add new spectral library**

**Merge two spectral libraries**

**Monitor spectral libraries**

#### HFX\_II\_Antoine\_peptFDR2

**Version:** v5  
**Mode:** Unsplit  
**Identifier type:** UNIPROT\_ALL  
**Database(s):** H\_sapiens\_iRT\_DECOY\_20170831  
**RT:** iRT-C18  
**Number of proteins:** 12129  
**Number of unambiguous proteins:** 11680  
**Number of peptides:** 218905  
**Organism:** Homo sapiens  
**Creation date:** 2018-01-30 18:15:53

Delete Edit  
Export Search  
Update Archive  
Restore previous version

### Adding a new library

From the libraries list window, click on the *Add new spectral library* button on the top of the list to display the form below.

## Adding new spectral library to Server

**Task :** ☐ Merge with an other library  
☐ Create new library

**Species :** -- Select species --  
(Update list of reference species if your species is not listed)

**Consensus library options :** ☐ Split ☐ Unsplit

**Files :** ☐ **Import files :**  
 Parcourir... Aucun fichier sélectionné.  
 Parcourir... Aucun fichier sélectionné.  
 Parcourir... Aucun fichier sélectionné.  
 (.dat, .mzXML and .tandem.pep.xml)

☐ **Import from project :**  
 -- Select Project --

☐ **Import archive file :**  
 Parcourir... Aucun fichier sélectionné.  
 (zip or gz archive)

☐ **Select directory from server :**  
  
 (only for bioinformatician)

☐ **Shared data directory**

**Fragmentation type :** -- Select fragmentation type --

**Instrument :** ESI-TRAP

**Databank :** -- Select Databank --

**Mayu options:** FDR estimation with Mayu software.  
 Missed cleavage : 2  
 FDR : 0.01 Type : protFDR

**RT file :** -- Select RT File --

**Description :**

Submit Clear Cancel

You need to select the following parameters in the library creation form :

- **Task :** You can create a new library or merge new data files with an existing library (create a new library from an existing one).
- **Library name :** Provide a name for the library.
- **Species :** Select the species scientific name of the databank list.
- **Consensus library options :** A consensus library is a spectral library in which MS2 spectrum entries with a redundant peptide sequence assignment have been collapsed into a single entry. Two options are provided for consensus library generation: a simple option assumes that all fragmentation spectra are correctly assigned (UNSPLIT) and a more sophisticated option additionally considers retention time when merging

spec ra (SPLIT).

- **Files** : Select the DDA data files used to generate the spectral library. Data from 3 search engines can be selected : Mascot files (.da), X! Tandem files (.ml or .tandem.pep.ml) and Seqes (.ml). For each Mascot, X! Tandem or Seqes file you need to upload the associated mXML file (with the same name as the Mascot, X! Tandem or Seqes file). You can upload your files from your computer, or can import them from an existing project (only for the .da files), upload an archive, or select the files in the shared directory.
- **Instrument** : The mass spectrometer used to acquire the data.
- **Databank** : The fasta file used by the search engines (Mascot, X! Tandem and Seqes).
- **Mayu options** : FDR estimation with MAYU. False Discovery Rate (FDR) and number of missed cleavage can be selected.
- **RT file** : The file containing the list of iRT retention time reference peptides.
- **Description** : Optional description of the current library.

Once the form is filled, click on the *Submit* button to launch the spectral library creation process.

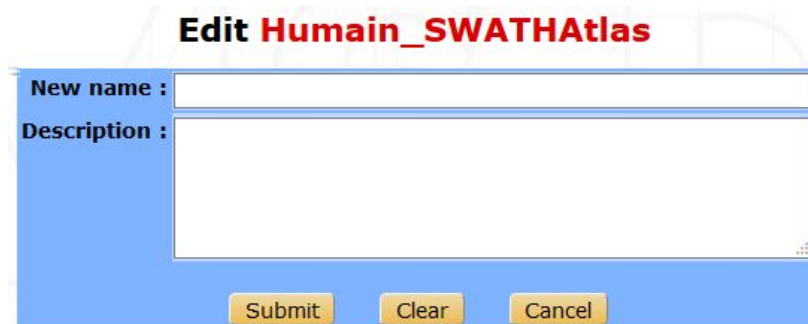
## Merging two library

Two libraries can be merged by clicking on the *Merge two libraries* button on the libraries list window. The displayed form requires the names of each of the 2 libraries, the name of the new library and an optional description. Clicking on *Submit* will fuse the selected libraries to create the new library.

Only two libraries with the same iRT file, databank type and consensus library option (SPLIT or UNSPLIT) can be merged.

## Editing a library

Now you can only the name and the description can be modified. From the libraries list window, click on the *Edit* button on the right side of the library row. The following form will be displayed :



**Edit Human\_SWATHAtlas**

New name :

Description :

Make the desired changes and click on the *Submit* button to save your changes.

## Updating a library

It is also possible to extend a library using another database-search data from the same organism. From the libraries list window, click on the *Update* button on the right of the library row. A form similar to the library creation one will be displayed. Fill in the parameters and click on the *Submit* button to launch the update process.

## Restoring the previous version of a library

An updated library can be downgraded by clicking on the *Restore previous version* button on the right of the library row on the libraries list interface. Every version of a library can be restored by consecutive downgrades.

## Searching for proteins in a library

Another available option is to check whether a protein of interest is present in a library and is aliased by the associated peptides by clicking in the *Search* button of the desired library, on the libraries list window.

Several proteins can be searched at the same time by inserting the accession names, the protein id or the names of the proteins (one per line or separated by either comma or a space character) in the following form.

**Search in Humain\_SWATHAtlas**

Entry :

Species name :

☒ Homo sapiens

Submit

Clear

Cancel

All the selected terms are searched beforehand in Uniprot, and a list of proteins is displayed. Some information such as the protein name, id, accession number, length and corresponding gene names are shown. The number of associated peptides identified is also indicated.

### Results for "histone"

Protein ID (AC)	Gene Names	Protein Names	AA	# Peptides
HDAC1_HUMAN (Q13547)	HDAC1, RPD3L1	Histone deacetylase 1 (HD1) (EC 3.5.1.98)	482	42
P53_HUMAN (P04637)	TP53, P53	Cellular tumor antigen p53 (Antigen NY-CO-13) (Phosphoprotein p53) (Tumor suppressor p53)	393	13
KAT5_HUMAN (Q92993)		Tat-interactive protein) (Tip60) (Histone acetyltransferase HTATIP) (HIV-1 Tat interactive protein) (Lysine acetyltransferase 5) (cPLA(2)-interacting protein)	513	3

The peptide list and the protein's sequence can be displayed by clicking on the number in the #



Peptides col mn.

Peptide list for P04637								
#	Sequence	Modifications	Position	M/Z	Charge	IRT time	Specificity (%)	Found with
1	TYQGSYGFR	-	102-110	539.7513	2+	7.7	100	P04637
2	LGFLHSGTAK	-	111-120	515.7876	2+	-7.2	100	P04637
3	SVTCTYSPALNK	Carbamidomethyl (C:4)	121-132	670.8294	2+	6.4	100	P04637
4	TCPVQLWVDSTPPPGTR	Carbamidomethyl (C:2)	140-156	955.9751	2+	65.8	100	P04637
5	QSQHMTVEVVR	-	165-174	607.8010	2+	-20	100	P04637
6	CSDSDGLAPPQHILIR	Carbamidomethyl (C:1)	182-196	833.4043	2+	24.1	100	P04637
7	CSDSDGLAPPQHILIR	Carbamidomethyl (C:1)	182-196	555.9386	3+	25.6	100	P04637
8	RPILTIITLEDSSGNLLGR	-	249-267	690.0635	3+	95.2	100	P04637
9	RTEEENLR	-	283-290	523.7649	2+	-31.3	100	P04637
10	KGEPHHELPPGSTK	-	292-305	505.2634	3+	-35.2	100	P04637
11	ALPNNTSSSPQPK	-	307-319	670.8439	2+	-18.7	100	P04637
12	KKPLDGEYFTLQIR	-	320-333	569.9858	3+	47.4	100	P04637
13	ELNEALELK	-	343-351	529.7900	2+	28.6	100	P04637

Detailed sequence coverage for P04637								
---------------------------------------	--	--	--	--	--	--	--	--

Peptide coverage : 38.2%

```

1 MEEPQSDPSV EPPLSQETFS DLWKLLPENN VLSPLPSQAM DDLMLSPDDI EQWFTEDPGP
61 DEAPRMPEAA PPVAPAPAAP TPAAPAPAPS WPLSSSVPSQ KTYQGSYGFR LGFLHSGTAK
121 SVTCTYSPAL NKMFCQLAKT CPVQLWVDST PPPGTRVRAM AIYKQSQHMT EVVRRCPHHE
181 RCSDSGLAP PQHLIRVEGN LRVEYLDDRN TFRHSVVVPY EPPEVGSDCT TIHNYMCNS
241 SCMGGMNRRP ILTIITLED SGNLLGRNSF EVRVCACPGR DRRTEEENLR KGEPHHELP
301 PGSTKRALPN NTSSSPQPK KPLDGEYFTL QIRGRERFEM FRELNEALEL KDAQAGKEPG
361 GSRAHSSHLK SKKGQSTSRH KKLMFKTEGP DSD

```

Some information about each peptide such as sequence, modifications, position on the protein, M/Z, charge, IRT time and specificity are shown.

## Exporting a library

You can export a library of sequences in a quantification software. From the libraries list in interface, click on the *Export* button to display the export form.

## Export Humain\_SWATHAtlas

<b>Export format :</b>	-- Select format -- ▾	
<b>Mass range of fragment ions :</b>	Min : 350	Max : 2000
<b>Ion series and charge :</b>	Ions : <input type="text"/> (separated by ',') for example : 'b,y'	
	Charge : 1,2	
<b>Number of ions per peptide :</b>	Min : 3	Max : 20
<b>Files :</b>	Windows SWATH file : <input type="button" value="Parcourir..."/> Aucun fichier sélectionné. File with modifications delta mass : <input type="button" value="Parcourir..."/> Aucun fichier sélectionné. Labelling file : <input type="button" value="Parcourir..."/> Aucun fichier sélectionné. Fasta file : <input type="button" value="Parcourir..."/> Aucun fichier sélectionné.	
<b>Other options :</b>	<input type="checkbox"/> Remove duplicate masses from labelling <input type="checkbox"/> Use theoretical mass Time scale : <input checked="" type="radio"/> seconds <input type="radio"/> minutes UIS order : <input type="text" value="2"/> Maximum permissible error : <input type="text" value="0.05"/> Allowed fragment mass modifications : <input type="text"/>	
<b>Protein list :</b>	<input type="button" value="Parcourir..."/> Aucun fichier sélectionné. (List of desired proteins's accession numbers separated by ',;' or enter/space)	
<input type="button" value="Submit"/> <input type="button" value="Clear"/> <input type="button" value="Cancel"/>		

You have to fill in the following parameters :

- **Export format** : The library can be exported for PeakView or for OpenSWATH, or you can download the final format of the library (sp).
- **Mass range of fragment ions** : Lower and upper mass limits of fragment ions. (min=350 and max=2000 by default).
- **Ion series and charge** : The ion desired type (a, b, c, y, or ...) and charge separated by a comma. (charge=1+ and 2+ by default).
- **Number of ions per peptide** : Minimum and maximum number of ions per peptide. (min=3 and max=20 by default).
- **Files** :
  - **Windows SWATH file** : Upload the file that contains the SWATH ion discovery scheme that has been used for SWATH data acquisition.
  - **File with modifications delta mass** : Optional file containing the modifications not specified by default.
  - **Labelling file** : Optional file containing the amino acid isotopic labelling mass shifts. If this option is selected, heavy transitions will be generated.
  - **Fasta file** : Optional database fasta file used to relate peptides to their proteins.
- **Other options** : You can select another optional options such as the maximum permissible error, the time scale, the UIS order (calculated when using sifting modification; if -1 is set, all transitions for each isoform will be reported; default : 2), or



the list of allowed fragments mass modifications.

- **Protein list** : You can select a file containing a protein list to export these proteins from the library.

Then you can click on the *Submit* button to launch the export process. Once the process is complete, you can download the final file if a download link has appeared.

### Deleting a spectral library

A library can be deleted from the list (by clicking on the *Delete* button of the corresponding library) only if this library has not been used to create another library (merge operation, in this case, a prompt will inform you).

## GO files management

GO analyses require two types of GO files: an ontology file and an annotation file. These files are not project-specific and are thus managed globally in myProMS. Only bioinformaticians and authorized mass spectrometers can manage GO files. From **myProMS main window**, select *Annotation data* and follow the *GO annotations* link to display the list of GO files recorded.

### Ontology files

Ontology files contain the GO terms identifiers, description and relationships between them. To **add** a new ontology file, click on *Add new Gene Ontology file*:

**Add a new Gene Ontology File**

**Name :**

**File :** ☒ **Use a local file:**  
 Aucun fichier sélectionné.

☐ **Use a remote file** (FTP/HTTP URL - .gz accepted):  [Download link](#) - [Info link](#)

**Scope :** ☒ **Complete** ☐ **Slim**

The displayed form requires the following information:

- **Name**: A relevant name for the ontology. This name will be displayed in all GO analyses arising from the ontology selection.

- **File**: The file containing the ontology must be in **OBO** format (not XML nor database dump). Daily updated ontology files can be fetched from [GO ebi.ac.uk](http://go.ebi.ac.uk). The file can be uploaded

directly from your computer or directly retrieved from remote FTP by providing its full URL (e.g. [ftp://ftp.geneontology.org/pub/go/ontology/obo\\_format\\_1\\_2/gene\\_ontology\\_1\\_2.obo](ftp://ftp.geneontology.org/pub/go/ontology/obo_format_1_2/gene_ontology_1_2.obo)).

- **Scope:** Specify if the ontology file contains the **full** gene ontology or a **slim** version. A slim version gives a broad overview of the ontology content in which the detail of the specific fine grained terms. If a slim file is used, make sure to select the *slim* option. In addition, to be able to use a slim ontology for GO analyses, at least one full ontology file must have been also recorded to allow some ProMS to reconstruct missing associations between proteins and the GO terms recorded in the slim file. Running a slim GO analysis in which a corresponding full ontology will cause an error!

Saved ontologies can be **edited**. If the file was retrieved by FTP and a most recent version is available on the disseminator, it can be downloaded again directly by clicking on *Update file*.

## Annotation files

Annotation files contain mapping of protein identifiers to GO terms. They are **species-specific** and must be in **Gene Association File (GAF)** format. A large number of pre-defined annotation files for many species can be fetched from the [Uniprot-GOA database](#). To **add** a new annotation, click on *Add new annotation file*:

**Add a new Annotation File**

<b>Name :</b>	<input type="text"/>
<b>Description :</b>	<input type="text"/>
<b>Species :</b>	<div><div>Drosophila melanogaster (Fruit fly) ▼</div><div>(Only reference species can be selected)</div></div>
<b>File :</b>	<div><div><input type="radio"/> <b>Use a local file:</b> <div><div>Parcourir...</div><div>Aucun fichier sélectionné.</div></div></div><div><input type="radio"/> <b>Use a remote file</b> (FTP/HTTP URL - .gz accepted): <div><input type="text"/></div><div>(e.g. <a href="ftp://ftp.ebi.ac.uk/pub/databases/GO/goa/HUMAN/gene_association.goa_human.gz">ftp://ftp.ebi.ac.uk/pub/databases/GO/goa/HUMAN/gene_association.goa_human.gz</a>)</div></div></div>
<b>Identifier used :</b>	<div>-- Select an identifier type -- ▼</div>
<div><div>Add</div><div>Clear</div><div>Cancel</div></div>	

The displayed form requires the following information:

- **Name:** A relevant name for the annotation, which will be displayed on each GO analysis arising from an annotation selection section.
- **Description:** An optional description for the annotation.
- **Species:** Select the targeted species from the list of available ones (See **Species** below for more information).

- **File:** file can be uploaded from your computer or retrieved remotely from a FTP server (e.g. [ftp://ftp.ebi.ac.uk/pub/databases/GO/goa/HUMAN/gene\\_association.goa\\_human.g](ftp://ftp.ebi.ac.uk/pub/databases/GO/goa/HUMAN/gene_association.goa_human.g) for the human annotation file).

- **Identifier used:** Select the protein identifier has been used in myProMS database for the annotation (e.g. select Uniprot ID or Uniprot AC for Uniprot-GOA files). If Default is selected, the default protein identifier displayed in myProMS will be used. This parameter must be selected to ensure proper GO annotation mapping.

## Species

myProMS automatically records the species associated with an protein validated. Because different strains or variants of the same species are also recorded, it is necessary to maintain all link these entries to the same **reference** species. Furthermore, reference species must be recorded for Gene Ontology analyses. A species management section is provided so that bioinformaticians and authorized mass spectrometrists/managers can maintain record or correct species information. By default, a list of 5 model organisms species database is provided within myProMS as reference.

### Listing species

From **myProMS main window**, select *Annotation data* and follow the *Species* link to access the species management interface.

<Figure Species list>

As shown in the above screen capture, a subset of species can be listed either by **scientific** or **common name** by selecting the appropriate initial letter in one of the 2 alphabets displayed.

### Adding or editing a species

A species can be added or edited by clicking on *Add species* or *Edit buttons* respectively. The following form is then displayed:

<Figure Add/edit species>

The common name, scientific name and accession ID fields are mandatory. A link to the **NCBI Taxonomy** resource is provided to help you find this information if not known. You can either select this species as reference by checking the *Is reference* or link it to a reference one. In addition an optional field allows you to link an species with a reference one by selecting a target species in the drop-down menu.

### **Deleting a species**

A species can be deleted from the list interface (by clicking on the Delete button of the corresponding species) only if this species is no longer associated with

home-named modification of his one binding the non-alid PTM. In the future, m ProMS will automatically applies to his home-named modification the properties of the referenced one. - if his modification has no imported through another name, it should edit the PTM and provide mass and specificity.

Makesure that all PTMs retrieved are valid in order to avoid the other features available in m ProMS to go wrong on top (like fragmentation table of peptides for example).

## Editing or merging PTMs

A PTM can be edited by clicking on *Edit* button.

### Editing modification **Acetyl**

<b>PSI-MS Name :</b>	Acetyl
<b>Interim Name :</b>	Acetyl
<b>Alternative Name(s) :</b>	
<b>Description :</b>	Acetylation
<b>Monoisotopic :</b>	42.0106
<b>Average :</b>	42.0367
<b>Unimod Accession # :</b>	1
<b>Specificity :</b>	Protein N-term, Any N-term, C, H, K, S, T, Y
<b>Hide Specificity Editing</b>	<div> <div>Any N-term</div> <div>Protein N-term</div> <div>A</div> <div>C</div> <div>D</div> <div>E</div> <div>F</div> <div>G</div> <div>H</div> <div>I</div> <div>K</div> <div>L</div> <div>M</div> <div>N</div> <div>P</div> <div>Q</div> <div>R</div> <div>S</div> <div>T</div> <div>V</div> <div>W</div> <div>Y</div> <div>Protein C-term</div> <div>Any C-term</div> </div>
<b>Project display :</b>	-Set code: A -Choose color: 00CC00 <span>Reset color</span>
<b>Is label :</b>	<input type="radio"/> Yes <input checked="" type="radio"/> No
<b>Is substitution :</b>	<input type="radio"/> Yes <input checked="" type="radio"/> No
<b>Merge with :</b>	-- Select --
<div> <div>Save</div> <div>Cancel changes</div> <div>Cancel</div> </div>	

In this mode, you can update the description or the del a-mass of this PTM. A link to UNIMOD is provided by giving the Unimod Accession number. Specificity can be updated given to or to be precise on the PTM and retrieve articles to make have read.

The option *Merge with* gives the opportunity to merge two PTMs in one single entry. This could be useful if you wish to give an alternative name to a modification. Select the modification to merge with the current PTM and click on save. This action will add the name of the current modification to the list of alternative names of the one selected.

For PTMs that you want to make appear in your projects and give special attention to, you need to enter a code (small, a single letter) and a color. Those PTMs will become relevant and will be choosable in every project to manage.

Here is a list of relevant PTMs and their according code-color designation.

<b>Relevant PTMs :</b>	<input type="checkbox"/> Acetyl ( <b>A</b> )	<input type="checkbox"/> Carbamidomethyl ( <b>C</b> )	<input type="checkbox"/> Dimethyl ( <b>D</b> )
	<input type="checkbox"/> Methyl ( <b>M</b> )	<input type="checkbox"/> Oxidation ( <b>O</b> )	<input type="checkbox"/> Propionyl ( <b>P</b> )
	<input type="checkbox"/> Phospho ( <b>P</b> )	<input type="checkbox"/> Label:13C(6) ( <b>Si</b> )	<input type="checkbox"/> Trimethyl ( <b>T</b> )
	<input type="checkbox"/> GlyGly ( <b>U</b> )		

For more information on this topic, please, see the [Project creation and settings](#) section.

## Settings

**Instrument settings**

**Validation templates management**

## Server management

**Log files**

**Server statistics**

**Testing server**