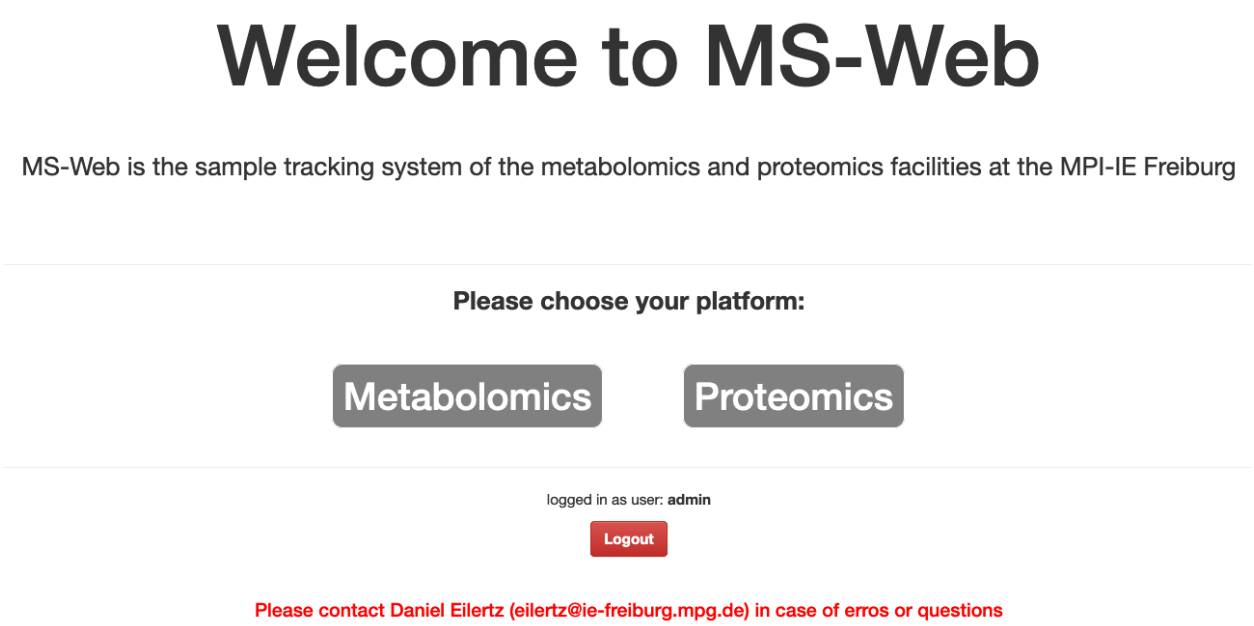
MS-Web

Release 01.06.2019

Oct 16, 2020

Contents



MS-Web is the laboratory information and management system of the metabolomics and proteomics facilities at the MPI-IE Freiburg.

The dockerized, django-based web-service offers sample tracking for metabolomics and proteomics experiments, providing researchers and laboratory staff . . .

MS-Web can be utilized to:

- Upload and track your mass spectrometry experiments
- Download measurement data
- Export related information as Excel spreadsheet
- Start automated software pipelines for data processing

CHAPTER 1

Overview

1.1 Installation

1.1.1 Requirements for MS-Web:

- GWDG-account to access the git repository. GWDG is an IT-service, provided for MPI-empolyees. Registration is for free.
- In order to access and work with MS-Web's repository, please contact Daniel Eilertz (eilertz@ie-freiburg.mpg.de) or Joerg Buescher (buescher@ie-freiburg.mpg.de) for permission.
- You need to have the following tools installed in order to start the MS-Web server:
  - Git
  - Docker (Version: 2.2.0.0 | Engine: 19.03.5)
  - Docker-compose (Version: v1.25.2)

1.1.2 Building and running docker containers

Clone the repository:

```
$ git clone https://gitlab.gwdg.de/daniel.eilertz/ms-web.git
```

Build docker images and start services:

```
$ docker-compose up --build
```

Migrate databases (django models are translated into SQL to create tables):

```
$ docker-compose run web python3 manage.py migrate
```

Create superuser/admin (Please specify name, email and password):

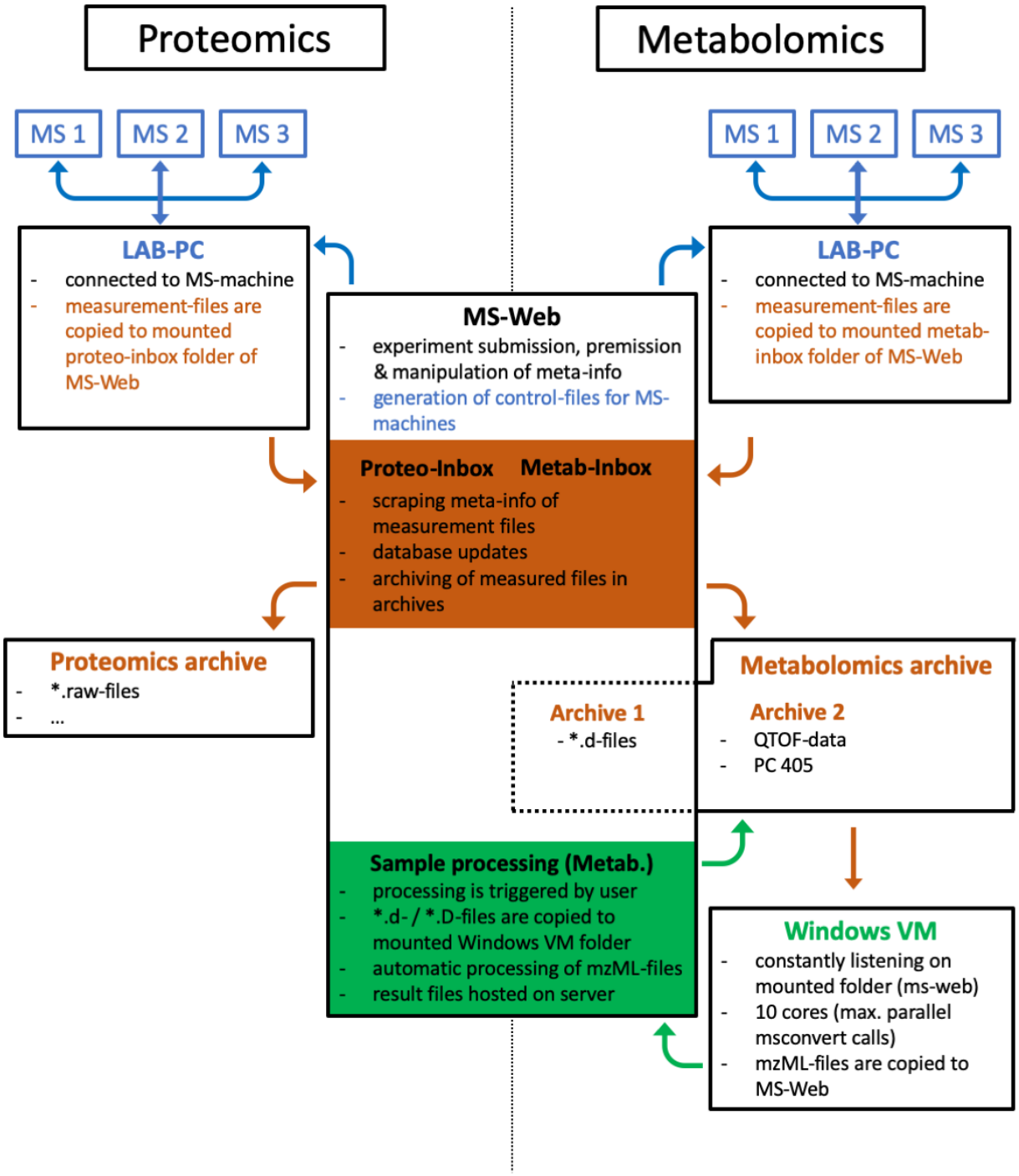
```
$ docker-compose run web python3 manage.py createsuperuser
```

Collect static files:

```
$ docker-compose run web python3 manage.py collectstatic --no-input
```

1.2 MS-Web's Infrastructure

1.2.1 Overview scheme:



Mslims (web server)

Windows VM (file Conversion)

Proteomics archive

Metabolomics archive

1.3 Technologies used by MS-Web

1.3.1 Docker

MS-Web consists of multiple docker (Version: 2.2.0.0 | Engine: 19.03.5) containers (& volumes):

- postgres: postgresql database
- web: django-based webservice
- nginx: webserver (used for production on server)

Click here for more information about MS-Web's containers and an overview about the most frequently used comands.

1.3.2 Django (Python3)

The high-level Python web-framework Django (Version: 2.2.5) is used to run MS-Web's backend. Django extensions

1.3.3 Git & Gitlab

1.3.4 PostgreSQL database

MS-Web uses a PostgreSQL database for its backend data storage. The database is run as the docker service db, which is based on the docker image postgres.

Django's database settings (name, user, password, port) are entirely configured using **/settings.py**.

Host-specific information is stored as and read out from environment variables stored in **/config/.env** (development) and **/config/.env_prod** (production). For reasons of security, these files are not hosted on the respository and must be set before the *the installation*.

Database tables are set up and managed completely as Django models, defined separately for both apps:

- **msweb/proteomics/models.py**
- **msweb/metabolomics/models.py**

To create, modify or delete tables and fields from MS-Web's database, these models have to be translated to SQL-code, which then performs the database operation. Changes to models are therefore saved as **migrations** which allow quick and easy changes to the database via python-code and guarantee traceability via git.

Create a migration file holding information about changes to the database:

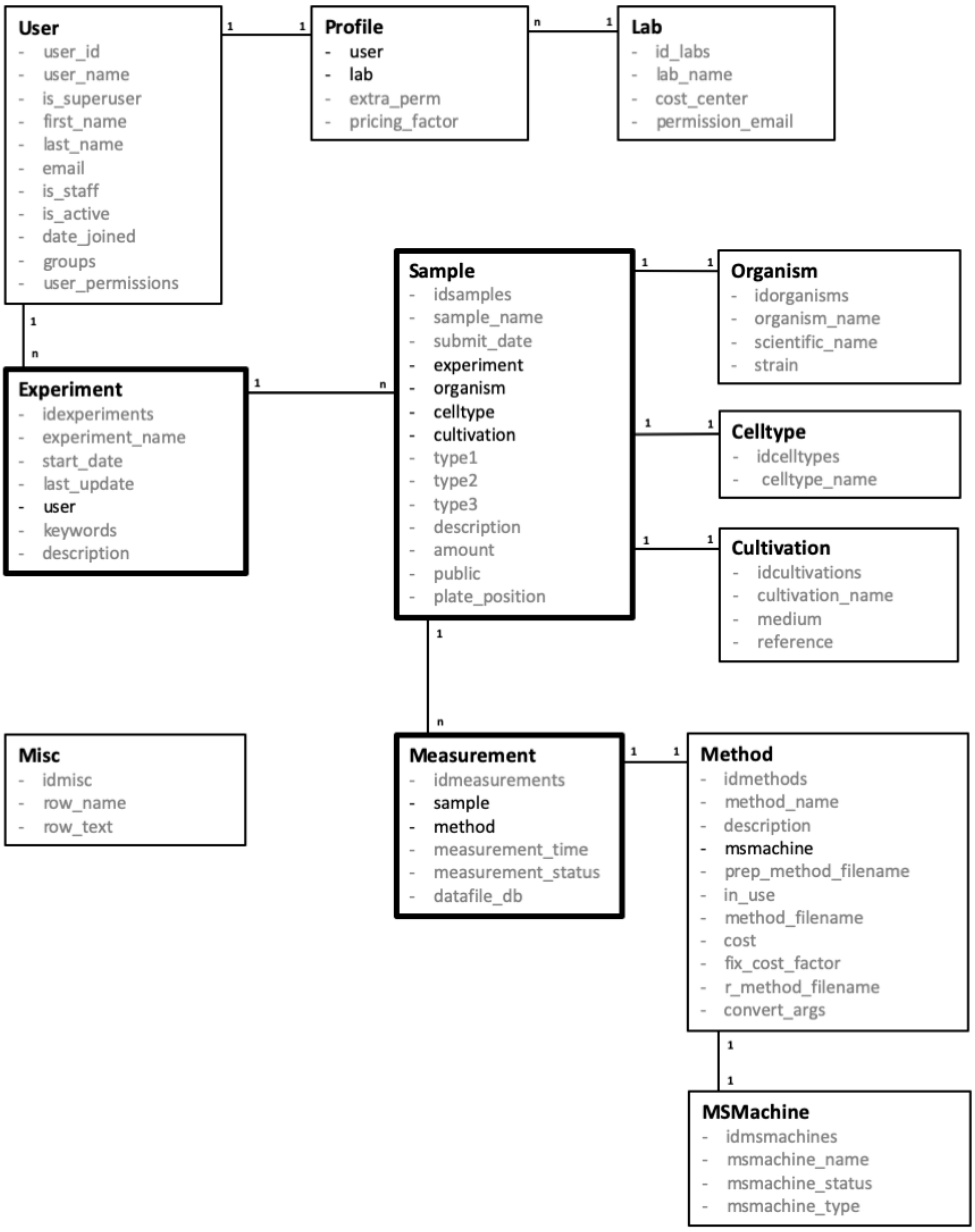
```
$ docker-compose run web python3 manage.py makemigrations
```

Translate a migration file to perform changes to the database:

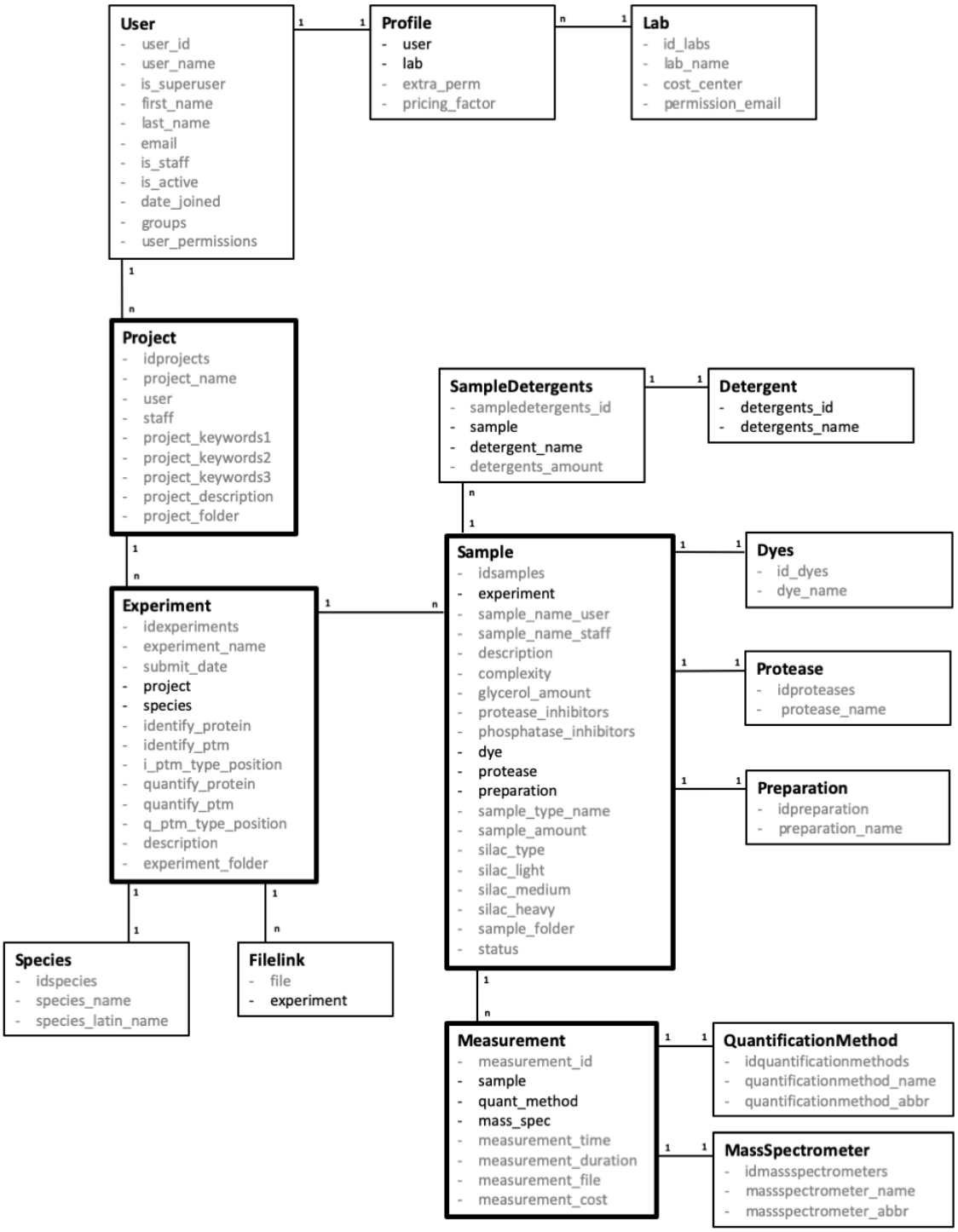
```
$ docker-compose run web python3 manage.py migrate
```

MS-Web's Admin site works as a powerful interface to the database. It lists all database tables and allows admins to create, update and delete most tables' fields using the browser.

Metabolomics database tables:



Proteomics database tables:



1.3.5 R & RStudio

The scripting language R is used for sample processing of the metabolomics app. Packages?

1.3.6 Nginx & Gunicorn

1.4 Metabolomics manual

MS-Web is the sample tracking management system of the proteomics and metabolomics facilities at the Max Planck Institute of Immunobiology and Epigenetics (MPI-IE) in Freiburg. The web based application serves as a tool to facilitate lab users and staff to organize mass spectroscopy measurements and related metadata information. MS-Web was intended to digitalize the facilities' working routines such as sample submission, permission and preparation as well as communication between staff and laboratory users. Besides the physical submission of sample tubes (hand over to facility staff), most work in order to perform and manage mass spec analysis can now be done in silico. This includes raw data storage and backup as well as result and metadata archiving.

All samples submitted to MS-Web are organized hierarchically. One user can have several **experiments**, each consisting of one up to multiple **samples**. In turn, one **sample** can contain at least one up to several **measurements**.

- **Experiments** are e.g. series of altering conditions, series of biological replicates or series of affinity pull-downs (different conditions, tags, cell lines) that are required to finish a project.
- **Samples** refer to tube(s) handed over to the facility = (technical) replicates belonging to one experiment.
- **Measurements** are distinct measured mass spectrometry data files or single LCMS/GCMS runs. Measurements are added to MS-Web automatically as soon as a sample has been measured. Therefore, measurements cannot be uploaded to the system by the users. Nonetheless, users are able to track it online in case sample measurement has been finished. The status of sample(s) has changed to "measured".

When related hierarchical items (experiments, samples and measurement) are mentioned, this tutotial refers to so called child and parent items:

- In this context, a measurement belonging to one specific sample is therefore a child item. The sample in turn is the parent item of this particual measurement.
- The same naming convention applies to samples (child items) of one distinct experiment (parent item).
- By this naming system, users can utilize MS-Web to navigate through their uploaded data in hierarchical way.

The *sample submission page* explains in detail how users can upload their experiments and samples to the facilities in order be measured.

The following topics will guide you through the basic parts and functionalities of MS-Web:

1.4.1 Login

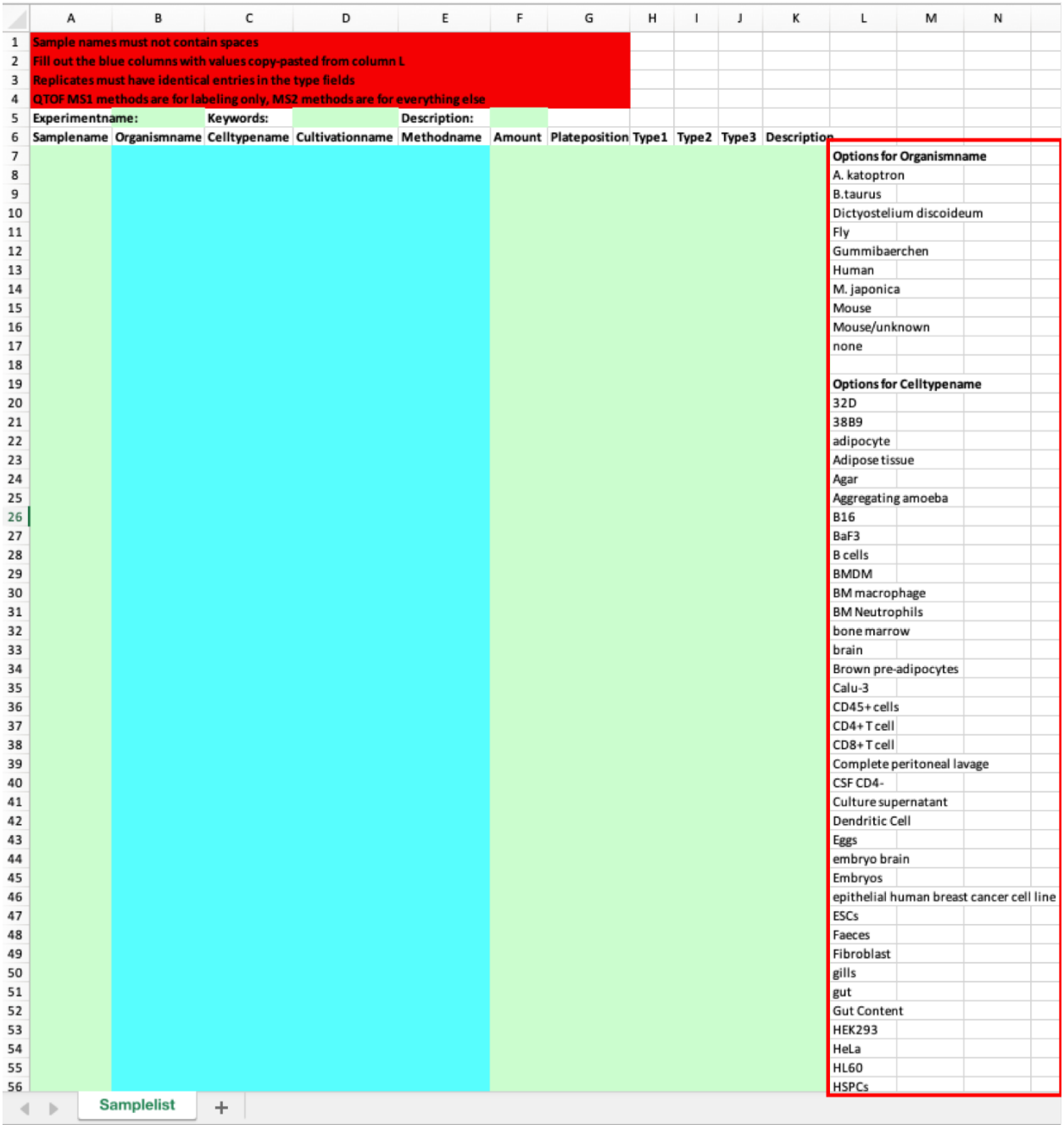
All MPI employees are able to login and use MS-Web once they have access to the institue's intranet. The webservice is also available from outside the insitute's network when using a VPN tunnel (must be installed and registered by the IT department).

Users can use their surname (lower case) as their user name and their (MPI-) email password for authentication.

Please contact Daniel Eilertz (eilertz@ie-freiburg.mpg.de) in case you are not able to login to MS-Web.

1.4.2 Sample submission

Metabolomics **experiments** and **samples** are uploaded to MS-Web via an Excel spreadsheet that can be downloaded from the metabolomics start up page the app or using this link. To submit the table containing your experiment data, please send an email with the attached spreadsheet to the Metabolomics Core Facility (metabolomicscorefacility@ie-freiburg.mpg.de).



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Sample names must not contain spaces													
2	Fill out the blue columns with values copy-pasted from column L													
3	Replicates must have identical entries in the type fields													
4	QTOF MS1 methods are for labeling only, MS2 methods are for everything else													
5	Experimentname:		Keywords:		Description:									
6	Samplename	Organismname	Celltypename	Cultivationname	Methodname	Amount	Plateposition	Type1	Type2	Type3	Description			
7												Options for Organismname		
8												A. katoptron		
9												B.taurus		
10												Dictyostelium discoideum		
11												Fly		
12												Gummibaerchen		
13												Human		
14												M. japonica		
15												Mouse		
16												Mouse/unknown		
17												none		
18														
19												Options for Celltypename		
20												32D		
21												38B9		
22												adipocyte		
23												Adipose tissue		
24												Agar		
25												Aggregating amoeba		
26												B16		
27												BaF3		
28												B cells		
29												BMDM		
30												BM macrophage		
31												BM Neutrophils		
32												bone marrow		
33												brain		
34												Brown pre-adipocytes		
35												Calu-3		
36												CD45+cells		
37												CD4+T cell		
38												CD8+T cell		
39												Complete peritoneal lavage		
40												CSF CD4-		
41												Culture supernatant		
42												Dendritic Cell		
43												Eggs		
44												embryo brain		
45												Embryos		
46												epithelial human breast cancer cell line		
47												ESCs		
48												Faeces		
49												Fibroblast		
50												gills		
51												gut		
52												Gut Content		
53												HEK293		
54												HeLa		
55												HL60		
56												HSPCs		

Samplelist +

Please fill out the Excel spreadsheet to provide the following information about your experiment and samples:

^ Mandatory fields

° Fields with formatting requirements

‘ Fields entries must be choosen from list on the right side of the spreadsheet table. **Please contact the metabolomics facility if your required option is not listed on the spreadsheet.**

Experimentname ^:
Please give your experiment a meaningful name. If the samples listed on the spreadsheet should be attached to an already existing experiment. Please use experiment name of this experiment. Otherwise a new experiment will be created automatically. Measured experiments can be searched for by names in MS-Web. The more precise users name their experiments the more convenient they can be tracked later on.

Keywords:
Uploaded experiments can be tagged with one to serveral keywords as a short summary e.g. about biological conditions. Once uploaded to MS-Web, experiments can be searched for by their given tags.

Description:
A (detailed) description of the experiment and its samples.

Samplename ^°:
Name of each sample tube, that will be handed in to the facilities in order to be measured. Only Letters, numbers, ‘+’ and ‘-’ are allowed as characters for sample names. If the sample name contains spaces, these will be transformed to underscores.

Organismname ‘:
Name of the organism from which the sample was generated.

Celltypename ‘:
Name of the celltype from which the sample was generated.

Cultivationname ‘:
Name of the cultivation condition that was used to generate the sample.

Methodname ‘:
Name of the analytical method to be used for sample measurement.

Amount ^°:
Amount of material in the sample tube (e.g. number of cells, milligram of total protein, microgram of RNA). Only Numbers are allowed in this column or characters that can be interpreted as numbers by Excel (e.g.400000 can be replaced by 4E5 not by 4x10^5).

Plateposition:
Position of the sample tube on the plate. Required only for really polar and soso polar methods.

Type 1 - 3:
Types 1 – 3 can be used to annotate samples for data analysis. Samples with identical entries will be treated as replicates.

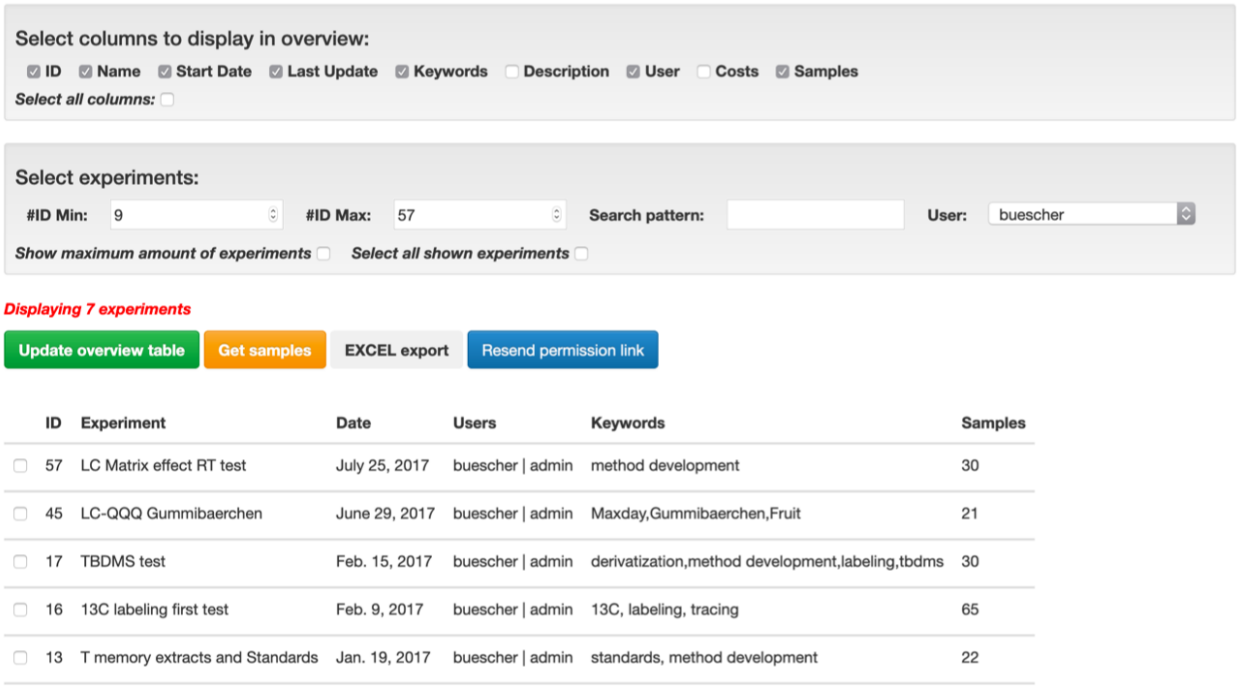
Description:
Free text to describe each sample. Once uploaded to MS-Web, samples can be searched for by their description.

Metabolomics staff will validate your data and eventually contact you to clarify open questions and finally upload the spreadsheet to MS-Web. Once the samples have been transferred, the user's lab head will be automatically notified by email with a detailed description of the samples to be measured and their associated measurement price. The email contains a link which automatically permits the corresponding samples to be measured by the facility. After sample permission, the user will be notified automatically by email.

As soon as all samples belonging to a permitted experiment have been processed, the user will again be informed by an automatic email.

From then on, samples are free to be downloaded or analyzed by the user via the *overview panels*.

1.4.3 Overview panels



Overview panels in MS-Web represent all items (**experiments** and **samples**) a user has uploaded to the system including the **measurements** performed by the metabolomics facilities.

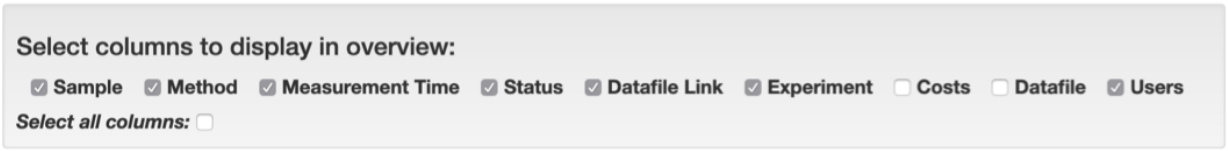
Using these panels users can keep track of their ongoing experiments and follow up the processing state of their samples.

In general, all overview panels consist of four components:

- The *Column section*
- *Search fields*

- *Action buttons*
- The *Overview table*

Column section

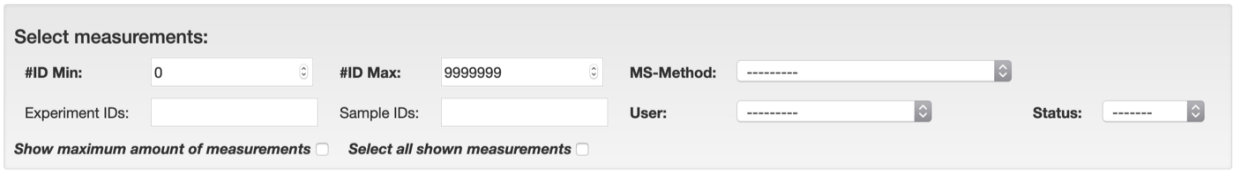


The column section can be used to select the features a user wants to have (actively) shown in the overview tables.

Once a selection has been made using the respective checkboxes the new view can be created using the **update overview table** action button. To reset a column selection in order to list all associated features, the checkbox **select al colums** has to be clicked with a subsequent page update using the **update overview table** action button.

By default, all columns/fields of the respective item overview are shown once the page is loaded initially.

Search fields



The search field bar serves the purpose of querying specific items from all those a user has submitted to MS-Web. Items can be searched by e.g. by their names and/or IDs. The search fields of the **measurement overview panel** also include queries by MS-Method (the method the sample) and the status of the measurement ('pending', 'permitted', 'measured' and failed). The field search pattern queries items based on all information that has been provided by the user (e.g. name, tags and description). Additionally, admins can use the search bar on the user interface to select items by related users.

Once a selection has been made using the form fields, the new view can be created using the **update overview table** action button. In order to list all associated items on the overview table, the checkbox **show all [items]** has to be clicked with a subsequent page update using the **update overview table** action button. The checkbox select all shown [items] can be used to automatically select all items that are currently displayed. Please click the **update overview table** action button in order to perform the automatic selection of items.

Action buttons

Action buttons can be utilized to update the respective overview page and perform actions with listed/selected items such as excel export and item manipulation. Please see the detailed descriptions of actions for each overview page on the respective manual section.

- The **update overview table** action button can be used to refresh the current page according to the customized list of columns (features) and items the user wants to visualize on the respective overview table. Note that all selections and changes will be shown on the page once the **update overview table** button has been clicked.
- As all items in MS-Web (projects, experiments, samples and measurements) are structured hierarchically, users can use the **Get [child item]** action button to navigate through all items that are associated with each other. If a user wants to see all experiments that belong to a specific project that is listed on the project overview page, using the **Get experiments** button MS-Web navigates to the experiment page displaying all relevant experiments. This procedure applies to all overview pages, from project to measurement overview and exclusively works when one single distinct item is select. Please note that one can always navigate back using the internet browser's back button.
- Each overview table page has an **Excel export** action button which can used to download selected items with all respective features as an Excel spreadsheet.
- On the **measurement overview** panel there are two additional buttons to download measurement files and start data processing.

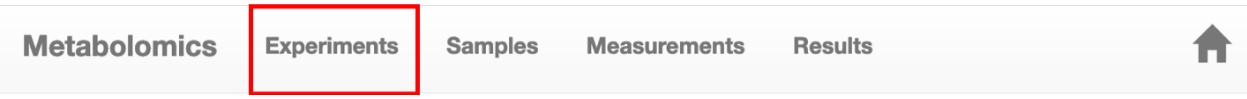
Overview table

	ID	Experiment	Date	Users	Keywords	Samples
☐	57	LC Matrix effect RT test	July 25, 2017	buescher | admin	method development	30
☐	45	LC-QQQ Gummibaerchen	June 29, 2017	buescher | admin	Maxday,Gummibaerchen,Fruit	21
☐	17	TBDMS test	Feb. 15, 2017	buescher | admin	derivatization,method development,labeling,tbdms	30
☐	16	13C labeling first test	Feb. 9, 2017	buescher | admin	13C, labeling, tracing	65
☐	13	T memory extracts and Standards	Jan. 19, 2017	buescher | admin	standards, method development	22

All items in MS-Web (experiments, samples and measurements) are shown in table form according to the information a user has provided during the submission process. Once changes have been made by the user or proteomics staff members, these are automatically shown.

If needed, the tables can be scrolled horizontally and/or vertically by mouse or by using the scrollbars that appear on the top and the bottom of the tables. Please use the **column selection** section in order to limit the number of displayed columns according to your preferences.

Experiment overview



The experiment overview lists all user's experiments with their associated features:

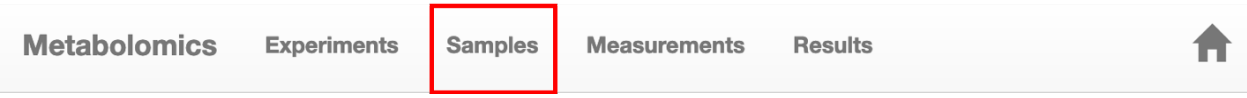
ID	The experiment's ID (can be searched for using the search field)
Name	Name of the experiment
Start Date	Date of experiment upload
Last Update	Time and Date of last update (e.g. adding new samples)
Keywords	Keywords associated with the experiment
Description	Short characterization of the experiment
User	The experiment owner(s)
Costs	Overall costs of the experiment (not selected by default)
Samples	Number of samples associated with the experiment

The experiment overview can be used to download experiment data using the Excel export button. Admins can resend permission links for selected experiments, in case the intital link has been accidently deleted a lab head or the link has expired for any reason.

By clicking the **Get samples** button, users can directly view all samples on the **sample overview** that belong to the selected experiments.

The number of displayed experiments on the overview page is limited to 200 for each user. Please use the search fields to query your view in order to display desired results.

Sample overview



The sample overview lists all user's samples with their associated features:

ID	Sample ID (can be searched for using the search field)
Sample	Name of the sample
Experiment	Name of the associated experiment
Date	Date of sample upload
Organism	Sample organism
Celltype	Sample celltype
Cultivation	Sample cultivation
Type 1 – 3	Sample annotation (for data analysis)
Description	Short description
Amount	Amount of material in sample tube
Public	Is the sample accessible by all users of MS-Web?
Plate Position	Plate position of sample tube
Sample Status	Measurement status of the sample
Costs	Overall costs of the sample (not selected by default)

The sample overview can be used to download sample data using the Excel export button.

By clicking the **Get measurement** button, users can directly view all measurements on the **measurement overview** that belong to the selected samples.

Users can search for samples by the ID of their corresponding experiment in the respective search field.

The number of displayed samples on the overview page is limited to 500 for each user. Please use the search fields to query your view in order to display desired results.

Measurement overview

The measurement overview lists all user's samples with their associated features:

ID	Measurement ID (can be searched for using the search field)
Sample	The associated sample of the measurement
Experiment	The associated experiment of the measurement
User(s)	The measurement owner(s)
Mesaurement time	Duration of the measurement
Method	Analytical method used for measurement
Status	Status of measurement
Datafile Link	Currently not in use
Costs	The cost of the measreuemtn
Datafile	Name of the measurement file

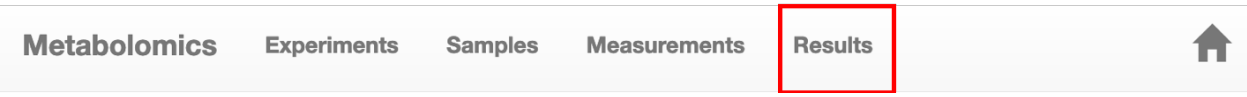
The measurement overview can be used to download mesaurement data using the **Excel export button**.

Users can search for measurements by the ID of their corresponding sample or experiment, the measurement status or the choosen MS-method in the respective search fields. Additionally, admins can query measurement information based on differnet users.

Using the **Start data processing button**, the preselected measurement items are going to be analyzed according to the given MS-method. Please see the section about **data processing** for details (currently under construction). Alternatively, selected measurements can be downloaded as raw files by clicking on the **Dowload raw data** button.

The number of displayed measurements on the overview page is limited to 500 for each user. Please use the search fields to query your view in order to display desired results.

Results overview



The results overview page lists all processing result files a user has generated using the automated **processing pipelines**.

Depending on the respective MS-Method, the the results overview table shows the following columns:

Dataset ID	ID of the processing dataset
Dataset name	Dataset set (given at the start of the processing pipeline)
Owner	Owner of the analysis (results are only visible for the owner)
Excel file	Excel file with signal intensities of detected metabolites
Peakoverview	PDF with chromatograms of detected metabolites for a first visual impression of measured raw data. For best viewing, download the file andt hen open it locally Don't attempt to print it!
Overview heatmap	PDF with an initial statistical analysis of analysed data, provided that you sample submission sheet contained sufficient metadata. For GCMS 13C labeling data this file also contains a mapping of your results on a metabolic map
Processing log	Rscript logfile containing the information that was collected during automatic processing (for debugging)

Selected processing results can be downloaded by clicking on the corresponding link or deleted by using the **Delete selection** button. Please note that deleted results can not be restored again.

For admins, all processing results on the server are displayed, while standard MS-Web users are only able to see their own data.

Please note that results will be removed at irregular intervals without prior warning!

The metabolomics staff while occasionally check the server for occupied storage space and delete older results from the archive. It is strongly suggested to directly download result files as soon as they are avaible on the server.

1.5 Proteomics manual

MS-Web is the sample tracking management system of the proteomics and metabolomics facilities at the Max Planck Institute of Immunobiology and Epigenetics (MPI-IE) in Freiburg. The web based application serves as a tool to facilitate lab users and staff to organize mass spectroscopy measurements and related metadata information. MS-Web was intended to digitalize the facilities' working routines such as sample submission, permission and preparation as well as communication between staff and laboratory users. Besides the physical submission of sample tubes (hand over to facility staff), most work in order to perform and manage mass spec analysis can now be done in silico. This includes raw data storage and backup as well as result and metadata archiving.

All samples submitted to MS-Web are organized hierarchically. One user can have several **projects**, each consisting

of one up to multiple **experiments**. In turn, one experiment can contain at least one up to several **samples**. Finally, the staff of the proteomics core can decide to perform multiple LC-MS measurements of one sample (technical replicate LC-MS injections or measurement of one sample with two different acquisition methods) that are termed **measurements**.

- **Projects** refer to broader scientific problems or questions e.g. all proteomic measurements included in a figure for a publication.
- **Experiments** are e.g. series of altering conditions, series of biological replicates or series of affinity pull-downs (different conditions, tags, cell lines) that are required to finish a project.
- **Samples** refer to tube(s) handed over to the facility = (technical) replicates belonging to one experiment.
- **Measurements** are distinct measured mass spectrometry data files or single LCMS runs. Measurements are added to MS-Web automatically as soon as a sample has been measured by a LCMS machine. Users are able to track it online in case sample **measurement** has been finished. The status of **sample(s)** has changed to "measured".

MS-WEB will automatically notify the head of the proteomics facility (and the user) by email once one or several samples have been submitted electronically via the MS-WEB interface (sample status = "submitted"). Once sample(s) have been approved by the line manager an automatic email is sent to the user (sample status switched to "approved"). Subsequently sample(s) can be physically handed over to the facility.

The following topics will guide you through the basic parts and functionalities of MS-Web:

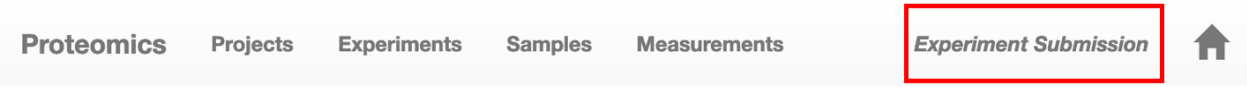
1.5.1 Login

All MPI employees are able to login and use MS-Web once they have access to the institue's intranet. The webservice is also available from outside the insitute's network when using a VPN tunnel (must be installed and registered by the IT department).

Users can use their surname (lower case) as their user name and their (MPI-) email password for authentication.

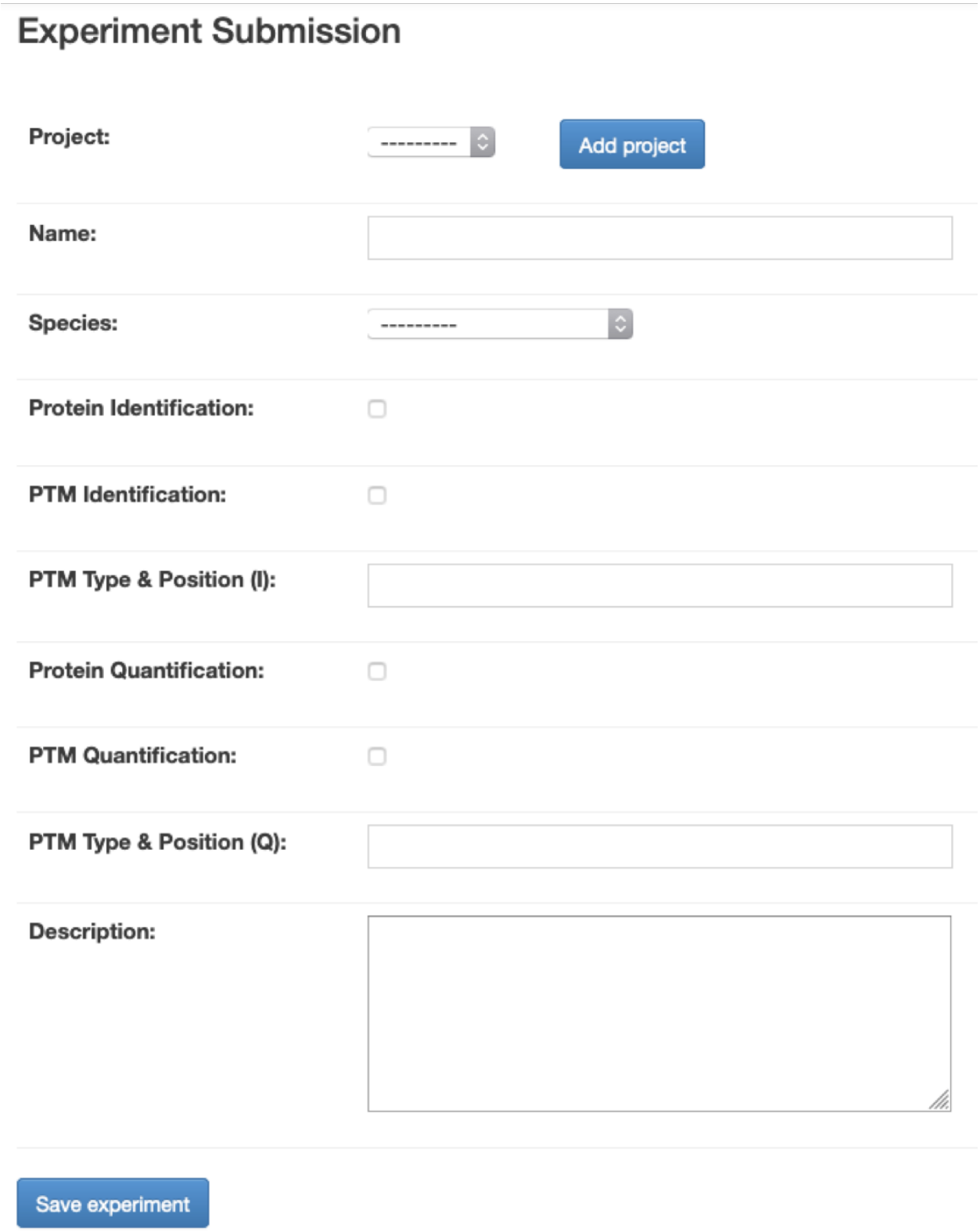
Please contact Daniel Eilertz (eilertz@ie-freiburg.mpg.de) in case you are not able to login to MS-Web.

1.5.2 Experiment submission



In most cases the experiment submission page will be the user's starting point to upload the information related to planned proteomics measurements. **Experiments** in MS-Web refer to a set of physical sample/tubes which are handed in to the proteomics facilities. On the **experiment submission page** one can add experiments to already existing projects or even create new projects to attach experiments to.

Form fields


Experiment Submission
Project:

Add project
Name:
Species:

Protein Identification:
PTM Identification:
PTM Type & Position (I):
Protein Quantification:
PTM Quantification:
PTM Type & Position (Q):
Description:
Save experiment

^ Mandatory fields
° Fields with formatting requirements

Project ^:
The name/title of your scientific project that might be composed of one to several experiments (or a series of experiments). See *project submission* to create a new project to attach the submitted experiment to it.

Name °^:
The name of your experiment that characterizes the biological question or the name of the gene/protein you are working with. If you plan to conduct several experiments it makes sense to number them in ascending order. Only letters, numbers, '-' and '_' are allowed in project name (don't use spaces.).

Species ^:
Biological origin of the sample (species/taxonomy) – in case you do not find the species in the drop down list, please contact the proteomics team.

Protein Identification:
If checked, all identified proteins in the sample will be reported by 1%FDR (default).

PTM Identification:
If checked, you indicate the aim of identifying a posttranslational modification (like phosphorylation, methylation, acetylation etc.). Please note that for a successful experiment you will either have to provide purified protein (1 pmol) or work out an enrichment strategy for the PTM containing peptides together with the proteomics unit.

PTM Type & Postion (I):
Please indicate the chemical identity of the PTM and the modified amino acid (e.g. S, T, Y phosphorylation, K acetylation, K Sumoylation).

Protein Quantification:
If checked, you indicate the aim of quantifying proteins in the sample. Please note that the number of identified proteins is usually slightly higher than the number of proteins that can be quantified; please note that the standard approach is label-free quantification with DDA (shotgun) proteomic data acquisition. In case you plan SILAC labeling or next generation technologies (DIA: data independent acquisition; targeted proteomics with PRM: parallel reaction monitoring) please contact the proteomics team ahead of planning your project and experiments.

PTM Quantification:
If checked, you indicate the aim of quantifying posttranslationally modified peptides in the sample; please note that PTM quantification requires prior **PTM identification** (above) and that the number of quantified PTM harboring peptides is generally smaller than the number of identified PTM containing peptides.

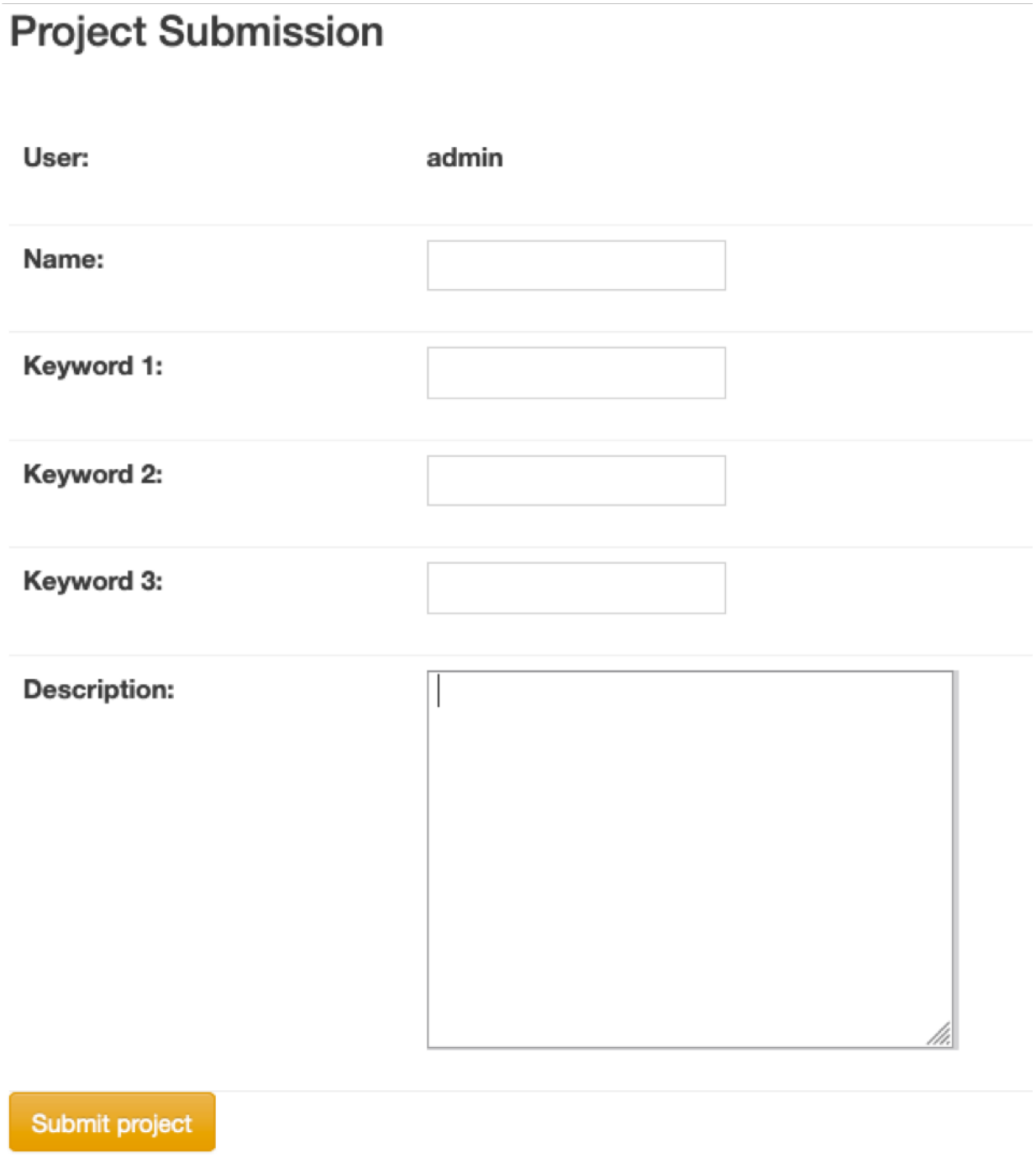
PTM Type & Position (II):
Please indicate the chemical identity of the PTM and the modified amino acid (e.g. S, T, Y phosphorylation, K acetylation) of the peptides (harboring a certain PTM) that should be quantified.

Description:
Detailed description of your experiment. Please note that a global (detailed) description of your biological study (biological question; goal of the project) is given in the project description. You should provide a description and the type of experiment (e.g. biological or technical replicate). You are welcome to indicate the controls and "treatments" (knockout, knockdown, stimulus, bait(s) etc. A short outline of the experimental workflow and the reagents, tools, buffers that are used will be very helpful for us to decide on the optimal sample preparation strategy. Notably, the buffer composition of your submitted samples should be listed in this field.

1.5.3 Project submission

Projects in MS-Web refer to long-term scientific projects and consist of one to several experiments. Projects can be added to MS-Web using the **Add Project Button** on the *experiment submission* page. Alternatively, users can choose an already existing project in the dropdown list on the *experiment submission* page in case they want to extend it with new experiments.

Form fields



^ Mandatory fields
° Fields with formatting requirements

User ^:

Predefined field based on logged in user.

Name ^°:
Name of the project saved in the proteomics database. This name will also appear on MS-Web. Only letters, numbers, '-' and '_' are allowed in project name (don't use spaces.).

Keyword 1 ^:
Keywords might be helpful to organize your uploaded data (e.g. number of planed experiment).

Keyword 2:
Keywords might be helpful to organize your uploaded data (e.g. aim of project).

Keyword 3:
Keywords might be helpful to organize your uploaded data (e.g. biology of project).

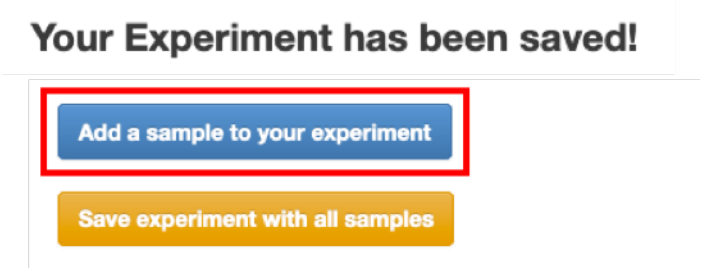
Description ^:
A description might be helpful to organize your uploaded data (like a short outline of the biology behind your project and the scientific question describing your project).

1.5.4 Sample submission

Samples in MS-Web refer to the physical samples/tubes a user hands in to the Proteomics facility in order to subject them to LC-MS analysis (measurement).

Samples can be uploaded in two ways:

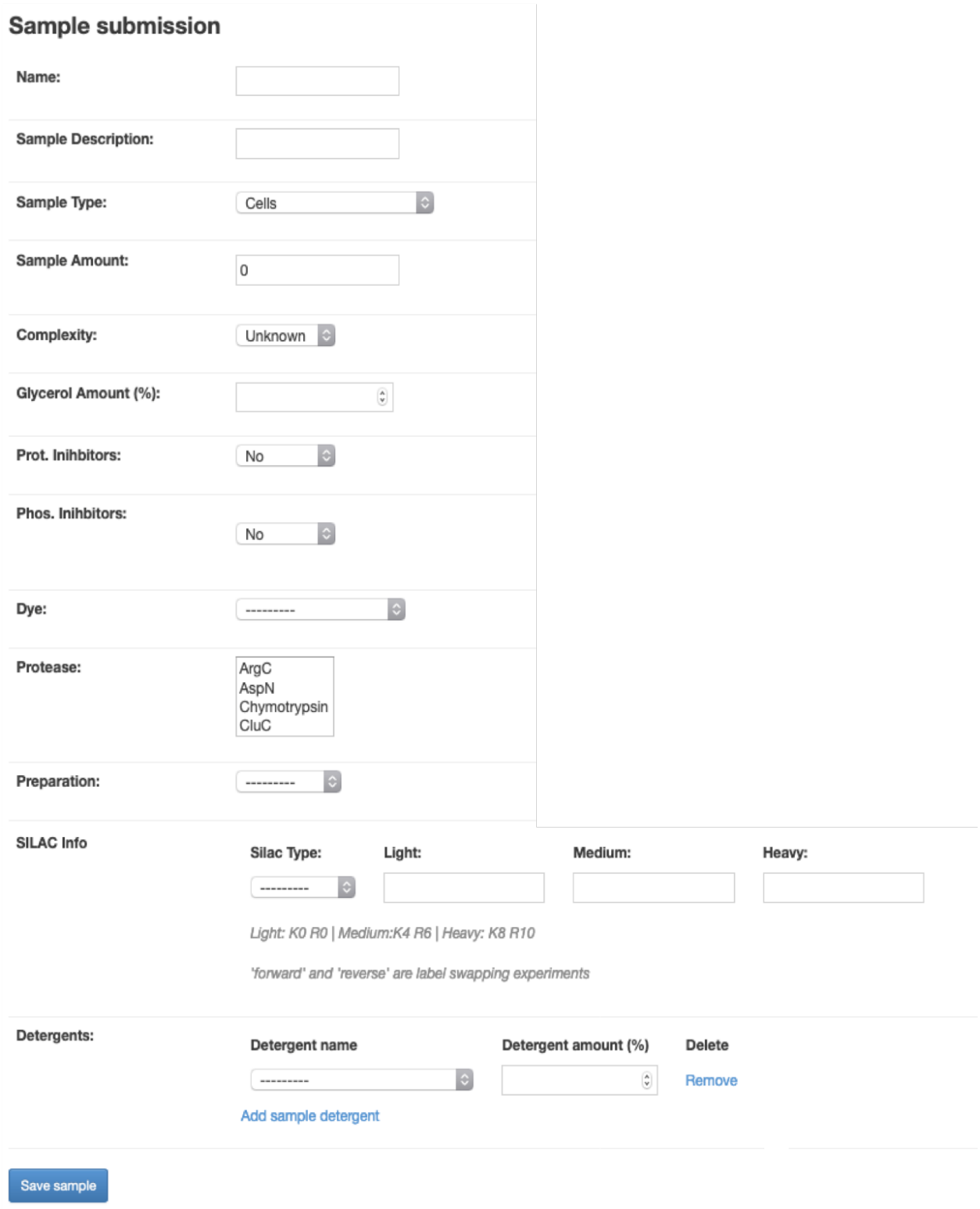
- Firstly, by adding them directly in the course of an ongoing *experiment submission*.



- And secondly, by adding them to an already existing experiment using the **Add samples action button** on the *experiment overview panel*. Either way, uploaded samples can be edited even after they have been initially submitted, using the **Edit sample action button** on the *sample overview panel*. Sample editing by the user is only possible as long as they have not already been approved by the head of the facility. In this way, the user is able to correct mistakes after submission and staff members are still in control to keep all sample meta-information consistent with actual measurement settings once the samples are processed in the lab. Please contact Gerhard Mittler in case you still want to perform changes to your sample information subsequent to approval by the line manager.


Update overview table
Get samples
EXCEL export
Add samples
Edit experiment

Form fields



^ Mandatory fields

° Fields with formatting requirements

Name °^:
Please give your sample a meaningful name. This could be an abbreviation of your gene/protein or cell type. If samples belong to a group of replicates please indicate it in the name by using e.g. _01, _02 or _rep01, _rep02.
Important note: For labeling the tubes that you submit to the facility either use the **sample name** or the unique **sample ID**.

Sample Description:
Describe your sample (e.g. biological condition, treatment). Please indicate if you are submitting biological or technical replicates.

Sample Type:
Specify the type of your sample (e.g. cells) that is physically submitted to the facility

Sample Amount:
Specify the volume of the sample / concentration (e.g. 100 µl of 1 µg/µl total protein) if possible. In case of cells, please indicate the cell number (200,000 or 2 Mio cells).

Complexity:
Select the complexity (e.g. how many thousands of proteins) of your sample in case you are able to judge. Total cell proteomes are highly complex whereas purified proteins obviously exhibit a very low complexity.

Glycerol Amount (%):
Does your sample contain glycerol and if yes, how much?

Protease Inhibitors:
Does your sample contain protease inhibitors and if yes, which kind of (e.g. PMSF, aprotinin, cOmpleteTM EDTA-free, cOmpleteTM plus EDTA)?

Phosphatase Inhibitors:
Does your sample contain phosphatase inhibitors?

Dye:
Only required if you are submitting samples that will be subjected to SDS-PAGE separation performed by the proteomics staff. Please select the gel staining method if possible. Please contact proteomic staff members in case you are not sure about the staining procedure that is most appropriate for your sample(s).

Protease:
Please choose an enzyme for proteolytic digestion (standard proteomic methodology characterizes proteins indirectly by measuring peptides generated through site-specific protease digestion). The selection of two proteases at the same time is possible (double-digestion). Use Ctrl (Strg) key for Windows and Cmd for Mac. The most commonly used

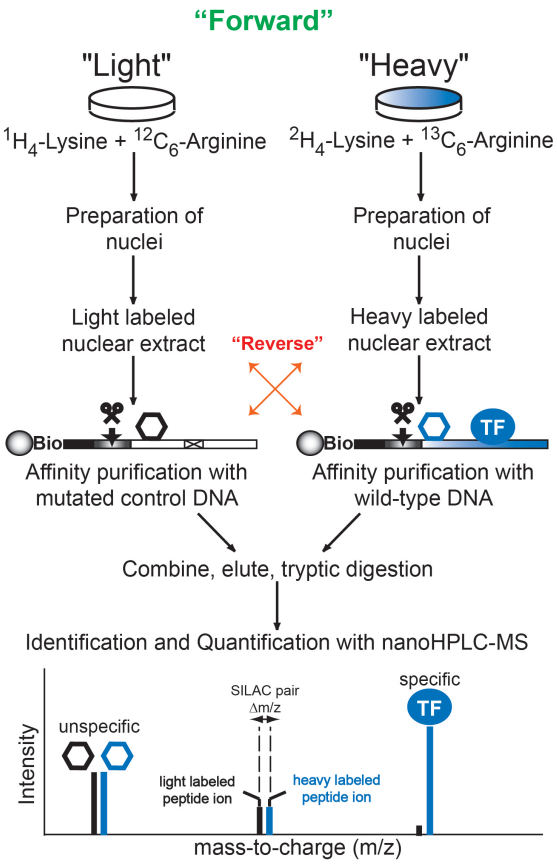
protease is trypsin. For double digestions a combination of LysC and trypsin is commonly employed. Please contact proteomic staff members in case you are not sure about the selection of protease(s).

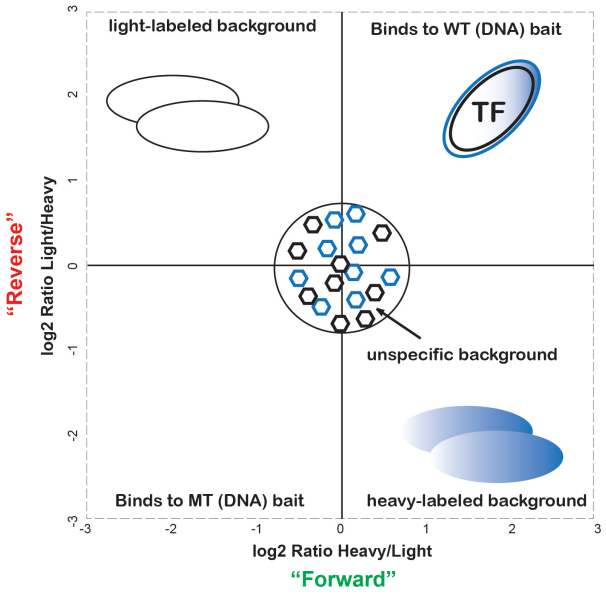
Preparation:
Please choose a preparation method. Please contact proteomic staff members in case you are not sure about the selection of the most appropriate sample preparation method.

SILAC Info:
Accurate and reliable quantitative proteomics in cell culture has been considerably facilitated by the introduction of the stable isotope labeling by amino acids in cell culture (SILAC), combined with high resolution mass spectrometry. SILAC enables the accurate quantification of differentially expressed (between two conditions) or differentially bound (to a wild-type bait vs control bait). There are however several major sources of quantification errors that commonly occur with SILAC techniques, i.e. incomplete incorporation of isotopic amino acids, metabolic arginine-to-proline conversion (if proline concentration in medium is limiting), contamination with exogenous non-labeled proteins, stochastic cell extract-"specific" protein precipitation and experimental errors in final sample mixing. Therefore, "Forward" and "Reverse" SILAC labeling experiments (with label-swapping) are carried out.

- *SILAC Type*: SILAC experiments can have two (standard double SILAC) or up to three (triple SILAC) "channels" = conditions that are mixed into one sample. They are termed "light" (no isotope labeling; cells grown in light non-labeled medium with natural lysine=K0 and arginine=R0), "medium" (cells cultivated in cell culture medium containing lysine-2H4=K4 and arginine-13C6=R6 substituting the natural lysine and arginine) and/or "heavy"(cells grown in medium containing lysine-13C615N2=K8 and arginine-13C615N4=R10 substituting the natural lysine and arginine):
- *Light*: K0 R0
- *Medium*: K4 R6
- *Heavy*: K8 R10

Very often SILAC replicate experiments are carried out in a project. For a "double" SILAC approach at least one "Forward" and one "Reverse" experiment is conducted. They represent SILAC-label swapping experiments that are instrumental in order to sort out false positives (differentially expressed protein candidates) exhibiting irreproducible (often high) ratios. By convention the wild-type or control sample is isotope-labeled (e.g. medium or heavy), whereas the mutant or treatment samples is unlabeled (light) in context of the "Forward" experiment. Likewise the wild-type or control sample is unlabeled (light), whereas the mutant or treatment samples is isotope-labeled (medium or heavy) in context of the "Reverse" experiment.





TF: Transcription Factor (protein binding a specific DNA sequence)

WT: Wild-Type DNA sequence harboring TF binding site

MT: Mutant DNA sequence lacking TF binding site

Label Swapping Experiment:

Forward: WT bait is incubated with "heavy" extract
MT bait is incubated with "light" extract

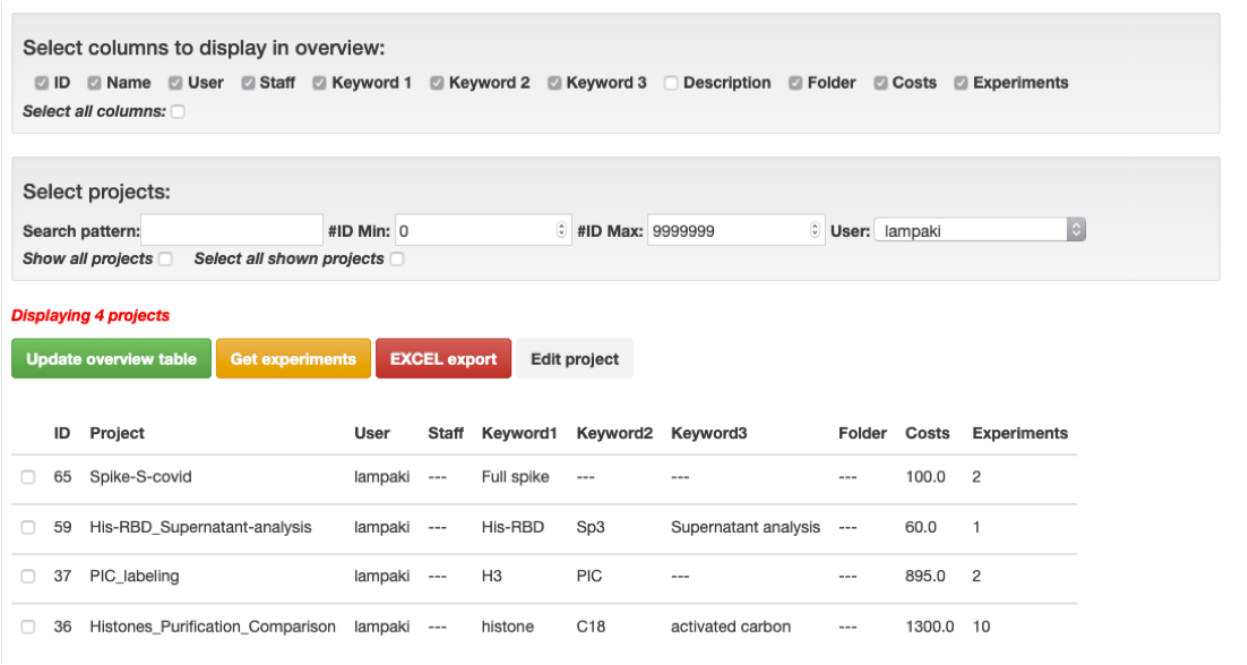
Reverse: WT bait is incubated with "light" extract
MT bait is incubated with "heavy" extract

Detergents:

Detergent information is **extremely crucial** in order to adjust the final sample preparation carried out in the facility. Carry-over of most detergents to the sample that is injected into the LC-MS system will severely hamper the performance of protein ID and quantification.

- *Detergent name*: please choose from list (in case your detergent is not listed, please contact the proteomics team in order to add it to the MS-WEB system)
- *Detergent amount (%)*: please indicate as e.g. 0,1 or 2,0 etc.

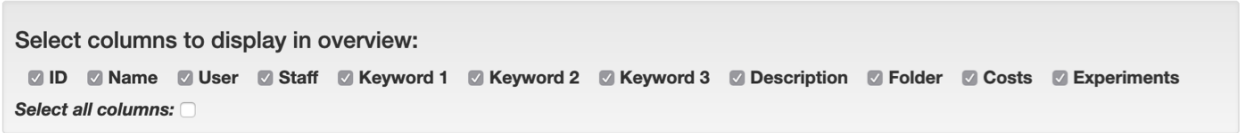
1.5.5 Overview panels



Overview panels in MS-Web represent all items (**projects**, **experiments** and **samples**) a user has uploaded to the system including the **measurements** performed by the proteomics staff. Using these panels users can keep track of their ongoing experiments and follow up the processing state of their samples. In general, all overview panels consist of four components:

- The *Column section*
- *Search fields*
- *Action buttons*
- The *Overview table*

Column section



The column section can be used to select the features a user wants to have (actively) shown in the overview tables. Once a selection has been made using the respective checkboxes the new view can be created using the **update overview table** action button. To reset a column selection in order to list all associated features, the checkbox **select al columns** has to be clicked with a subsequent page update using the **update overview table** action button. By default, all columns/fields of the respective item overview are shown once the page is loaded initially.

Search fields

The search field bar serves the purpose of querying specific items from all those a user has submitted to MS-Web. Items can be searched by e.g. by their names and/or IDs. Additionally, admins can use the search bar on the user interface to select items by related users. Once a selection has been made using the form fields the new view can be created using the **update overview table** action button. In order to list all associated items on the overview table, the checkbox **show all [items]** has to be clicked with a subsequent page update using the **update overview table** action button. The checkbox **select all shown [items]** can be used to automatically select all items that are currently displayed. Please click the **update overview table** action button in order to perform the automatic selection of items.

Action buttons

Action buttons can be utilized to update the respective overview page and perform actions with listed/selected items such as excel export and item manipulation. Please see the detailed descriptions of actions for each overview page on the respective manual section.

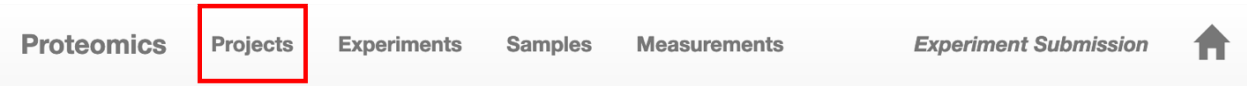
- The **update overview table** action button can be used to refresh the current page according to the customized list of columns (features) and items the user wants to visualize on the respective overview table. Note that all selections and changes will be shown on the page once the **update overview table** button has been clicked.
- As all items in MS-Web (projects, experiments, samples and measurements) are structured hierarchically, users can use the **get [child item]** action button to navigate through all items that are associated with each other. If a user wants to see all experiments that belong to a specific project that is listed on the project overview page, using the **get experiments** button MS-Web navigates to the experiment page displaying all relevant experiments. This procedure applies to all overview pages, from project to measurement overview and exclusively works when one single distinct item is select. Please note that one can always navigate back using the internet browser's back button.
- Each overview table page has an **Excel export** action button which can used to download selected items with all respective features as an Excel spreadsheet.

Overview table

	ID	Project	User	Staff	Keyword1	Keyword2	Keyword3
☐	65	Spike-S-covid	lampaki	---	Full spike	---	---
☐	64	MIMI_RNA	mittler	---	Dominica G. Hilgers lab	CR31451	RAP-MS
☐	63	Amphioxus	schultesasse	---	eggs	sperm	---
☐	62	Spillo_interacion_partners	schultesasse	---	Spillo	Nuclear_soluble	Cytoplasm
☐	61	EggMSandHP1aMS	atinbayeva	---	9 samples	---	---
☐	60	MTaseKO_Chromatin_Proteomics	mittler	---	Montavon	murine MEF chromatin	WT - 5KO cells
☐	59	His-RBD_Supernatant-analysis	lampaki	---	His-RBD	Sp3	Supernatant analysis

All items in MS-Web (projects, experiments, samples and measurements) are shown in table form according to the information a user has provided during the submission process. Once changes have been made by the user or proteomics staff members, these are automatically shown. If needed, the tables can be scrolled horizontally and/or vertically by mouse or by using the scrollbars that appear on the top and the bottom of the tables. Please use the **column selection** section in order to limit the number of displayed columns according to your preferences.

Project overview

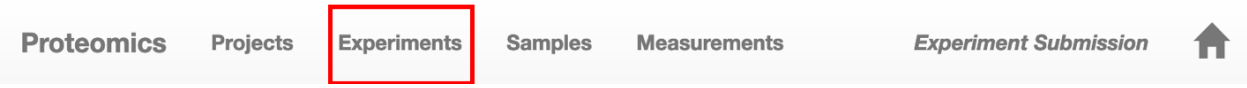


The project overview lists all user's long time projects with their associated features:

ID	The poject's ID (can be searched for using the search field)
Project	Name of the project
User	The project owner
Staff	Associated staff member of the protoemics facility
Keyword1	First keyword option
Keyword2	Second keyword option
Keyword3	Third keyword option
Description	Short characterization of the project
Experiments	The number of experiments that belong to the project
Costs	The overall costs of the project (currently not in use)
Folder	Currently not in use

Listed projects can be edited after submission using the **edit project** action button. Please note, that only one project can be modified at once. This function needs to be handled with extreme care because it solely serve's the purpose of correcting mistakes that occurred during the submission process. Once a project has been approved to be measured by the proteomics facility, it is not possible to edit the project any further. Please contact the facility staff in urgent cases. The number of displayed projects on the overview page is limited to 20 for each user. Please use the search fields to query your view in order to display desired results.

Experiment overview

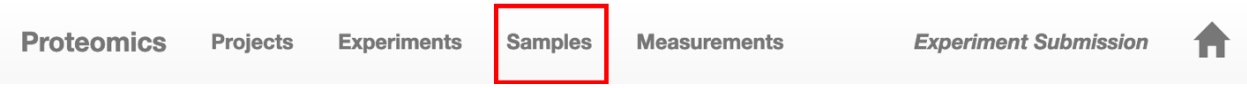


The experiment overview lists all user's experiments with their associated features:

ID	The experiment's ID (can be searched for using the search field)
Name	Name of the experiment
Submission Date	Date of initial submission or last edit by user
Project	Name of the associated project
Species	Biological origin of the experiment's samples
Protein Ident.	Should protein IDs be reported for this experiment?
PTM Ident.	Should PTMs be reported for this experiment?
Type and Position	Chemical identity of PTMs and modified amino acids
Protein Quant.	Should proteins be quantified for this experiment?
PTM Quant.	Should PTMs be quantified for this experiment?
PTM Type & Pos.	Chemical identity of PTMs and modified amino acids
Description	Short characterization of the experiment
Samples	Number of samples associated with an experiment
Costs	Overall costs of the experiment (currently not in use)
Folder	Currently not in use

Listed experiments can be edited after submission using the **edit experiment** action button. Please note, that only one experiment can be modified at once. This function needs to be handled with extreme care because it solely serve's the purpose of correcting mistakes that occurred during the submission process. Once an experiment has been approved to be measured by the proteomics facility, it is not possible to edit it any further. Please contact the facility staff in urgent cases. Using the **add samples** action button, a user can add samples to an already submitted experiment. Samples can only be added to one distinct experiment and as long as it hasn't been approved by proteomics staff members. To list all results associated with selected experiments, users can click on the **get results** action button. The number of displayed experiments on the overview page is limited to 100 for each user. Please use the search fields to query your view in order to display desired results.

Sample overview



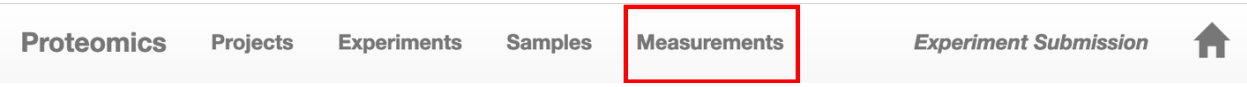
The sample overview lists all user's samples with their associated features:

ID	The sample's ID (can be searched for using the search field)
Name	Name of the sample
Name (Staff)	Names of the sample given by facility staff
Experiment	Name of the associated experiment
Status	Measurement status (submitted, approved, measured)
Complexity	The sample's complexity (unknown, low, medium, high)
Glycerol	Percentage (0-100 %) pf glycerol per sample
Prot. Inhibitors	Does the sample contain protease inhibitors?
Phos. Inhibitors	Does the sample contain phosphatase inhibitors?
Dye	Gel staining method (only for sample separation by SDS-PAGE)
Protease(s)	Enzyme(s) used for proteolytic digestion of proteins
Preparation	The sample preparation method
Sample Type	Sample type (e.g. cells)
Sample Detergents	Type and amount (%) of detergents in the sample
SILAC Type	Forward or Reverse
Light	lysine-0 (K0) and arginine-0 (R0) [no isotope labeling]
Medium	lysine-2H4 (K4) and arginine-13C6 (R6)
Heavy	lysine-13C615N2 (K8) and arginine 13C615N4 (R10)
Amount	The sample's volume / concentration / cell number
Costs	Overall costs of the sample (currently not in use)
Description	Short characterization of the sample
Folder	Currently not in use

Listed samples can be edited after submission using the **edit sample** action button. Please note, that only one sample can be modified at once. This function needs to be handled with extreme care because it solely serve's the purpose of correcting mistakes that occurred during the submission process. Once a sample has been approved to be measured by the proteomics facility, it is not possible to edit it any further. Please contact the facility staff in urgent cases.

The number of displayed samples on the overview page is limited to 500 for each user. Please use the search fields to query your view in order to display desired results.

Measurement overview

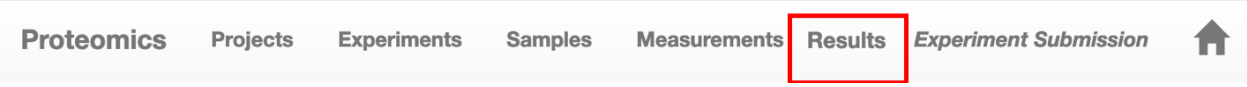


The measurement overview lists all user's measurements with their associated features:

Measurement ID	Measurement's ID (can be searched for using the search field)
Sample	Name of the associated sample
Masurement Time	Time and Date of measurement
Duration	Duration of measurement process
Method	Data acquisition method (e.g. DDA, DIA, PRM)
MS Machine	Mass spec instrument used for measurement
File	Full name of measurement file in proteomics archive
Costs	Measurement costs (currently not in use)

The number of displayed measurements on the overview page is limited to 500 for each user. Please use the search fields to query your view in order to see desired results.

Result overview



The result overview lists all files that have been uploaded to MS-Web by staff members (e.g. graphs, tables). Users can see all their results by clicking on the navbar link item or select specific experiments on the experiment overview panel to obtain respective filelinks only.

1.5.6 Proteomics data management

MS-Web serves as a tool for automatic archiving of measurement data. The file transfer from the mass spectrometer to the final storage is the last part of the proteomics data management process and can be explained according to the succesive steps a sample's lifecycle:

- Project/experiment/sample submission by user.

- Sample approvement by head of facility. Until then, the user is allowed to perform changes to uploaded data.

- XCALIBUR-CSV export of samples to be measured, using the sample overview page. If needed, one can change the ordering of sample in the resulting XCALIBUR-CSV file:
  - Select the XCALIBUR Order column in the column section and update the overview table.
  - Fill out the appearing forms of each sample to export according to your custom ordering.
  - Select all samples to be exported and click the XCALIBUR export button.
  - Your list shoul now be downloaded automatically.

- Physical preparation of sample & loading of XCALIBUR-CSV into software.

- One to many LC-MS measurement(s) of sample(s).

- Copying of measurement files from LS-MS machine to old archive and to mounted folder of MS-Web. Please stick to the guidelines of naming convention:
  - Before copying files to the inbox folder, please make sure to activate the listener using the *file transfer page* of the proteomics which can be accessed here.
  - The filenames must be unique in order to be stored in the archive! If you plan several measurements of the same sample, you will need to add suffixes to the filenames in XCALIBUR.
  - The filenames to be copied must exactly match those of the XCALIBUR-CSV. Otherwise the listener will not accept the files in the mounted folder.
  - Once a measurement file has been rejected by the software, the filename gets an additional flag "_ERROR" and remains in the folder. Therefore it is crucial to always check the inbox folder after the copy process.

- Once all files are copied to the folder, the listener software is activated and scrapes the metainformation from the raw file and updates the database. New measurement(s) are created and attached to respective sample(s) with the following information:
  - Sample
  - Quantification method (Quantification method Abbreviation)
  - Mass spectrometer
  - Measurement time
  - Measurement duration
  - Measurement file

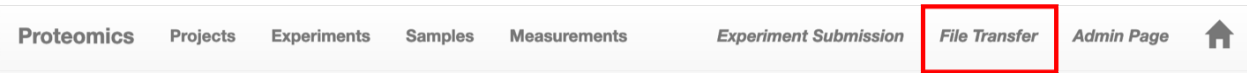
- The measurement file(s) are automatically saved into proteomics archive:
  - Folder Hierarchy:

```
/proteo_archive/<project ID>/<experiment ID>/

E.g. /proteo_archive/00000061/00000145/
```

  - Naming convention:

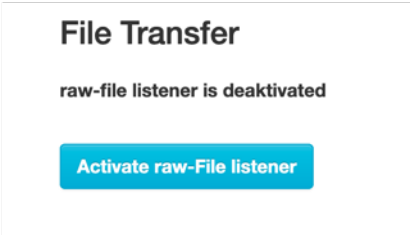
```
<sample ID>___<sample name>_<suffix>.raw

E.g. 00000746___TRIP180rep1_01.raw
```

1.5.7 Proteomics file transfer



The proteomics file transfer page can only be visited by admins and staff members. Its solely purpose is to activate the listener which will detect incoming measurement files that are copied to a mounted folder on the MS-Web server.

If not activated, there will appear a neon blue button for listener activation. As the listener software runs continuously, it only has to activated once after a server restart.



As soon as the listener is activated, the button will disappear and the PID (linux process ID) of the process will be displayed.

For testing purposes or in case of errors, the PID can be used to force the listener to stop.

Log into the web container:

```
(sudo) docker-compose exec web bash
```

Use the PID of the listener to quit the process:

```
(sudo) kill <PID>
```

Alternatively, you can use the following comand to get the listener's PID:

```
(sudo) ps aux
```

1.5.8 MS-Web's proteomics admin panel

1.6 License

1.7 Need help

Please contact daniel.eilertz@ie-freiburg.mpg.de in case of questions