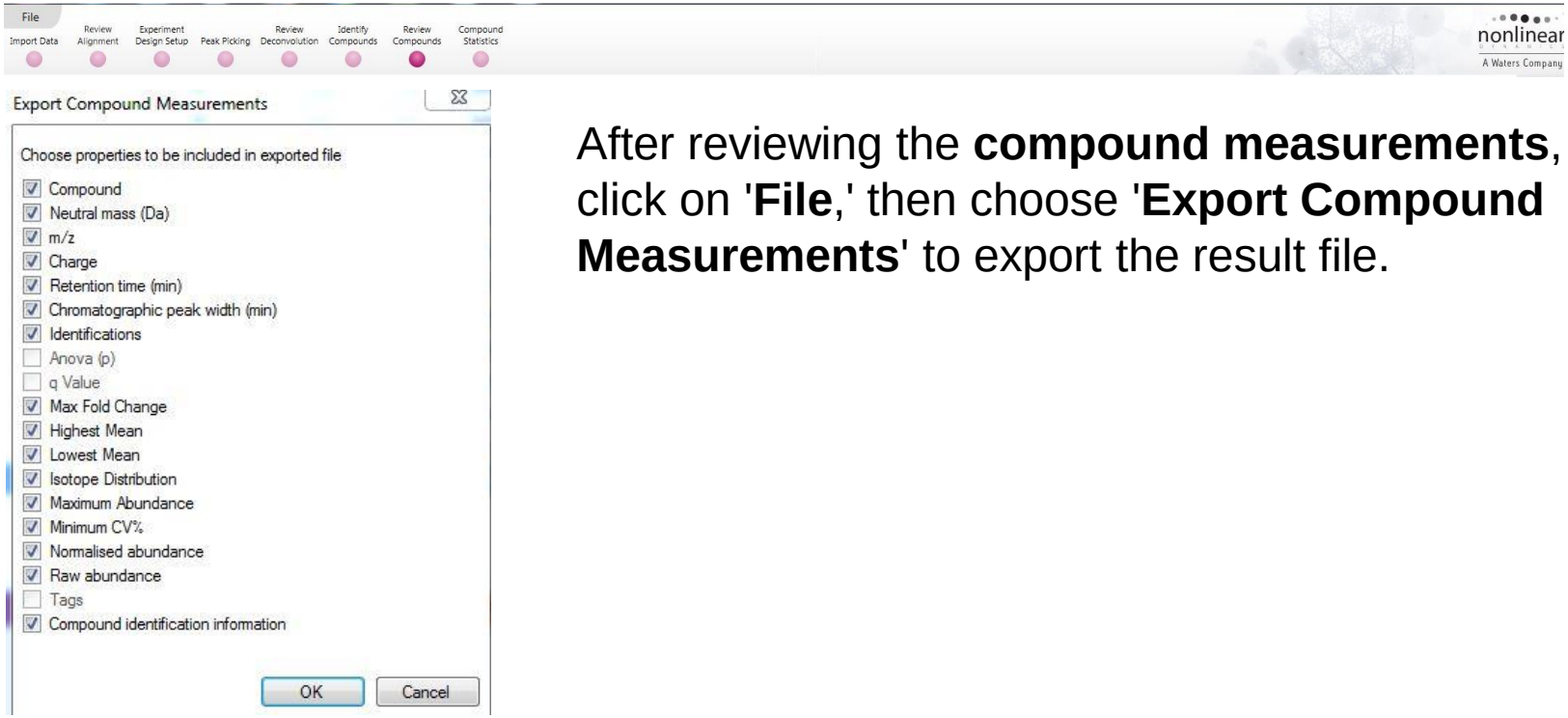
SmallP Manual

Version 2024.12

Contents

1. Prepare files: Progenesis QI

Progenesis QI can deal with massive datasets from multiple instruments for spectral deconvolution in both GC/MS and data-independent MS/MS.
Progenesis QI tutorial:
https://www.waters.com/webassets/cms/promotion/media/designed_page/local_china_common/progenesis_qi/#Progenesis

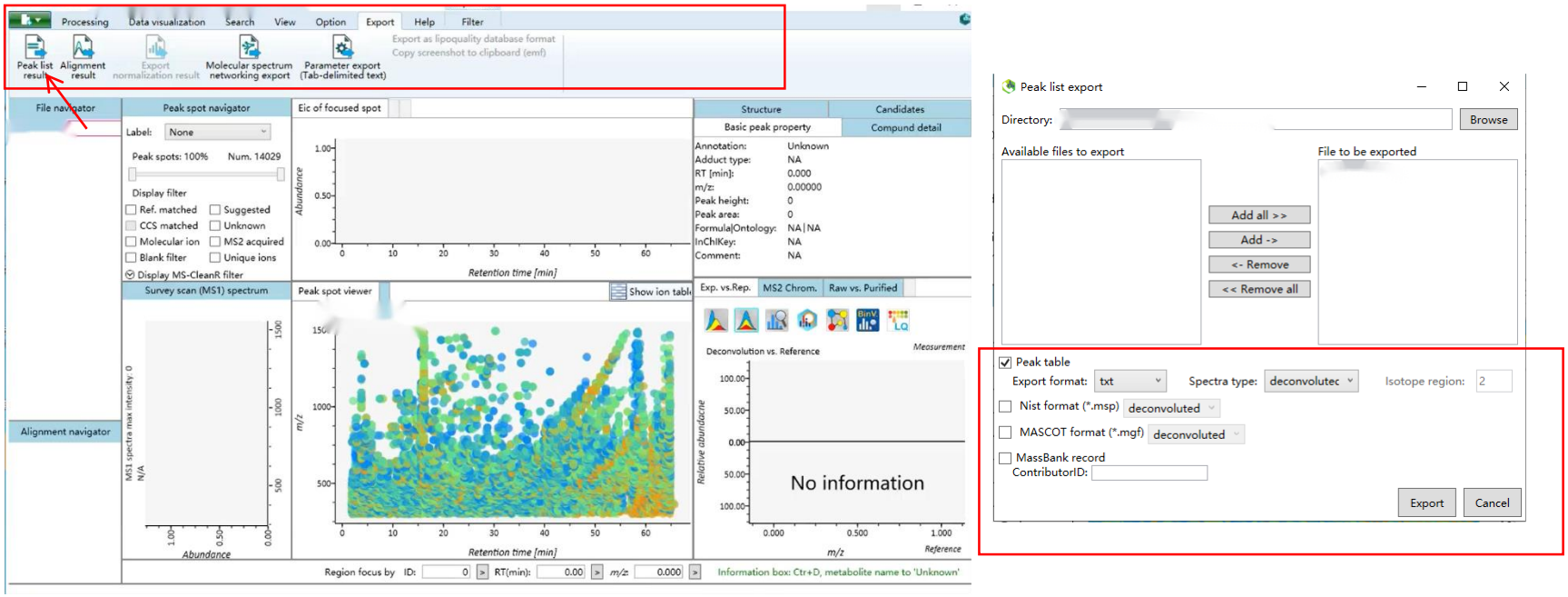


After reviewing the **compound measurements**, click on '**File**,' then choose '**Export Compound Measurements**' to export the result file.

1. Prepare files: MS-DIAL

MS-DIAL

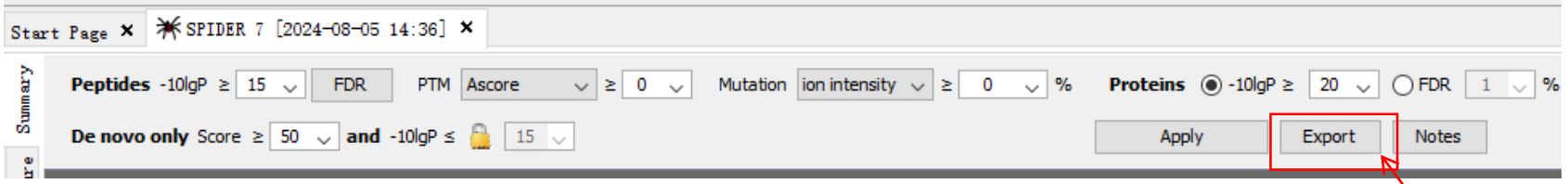
MS-DIAL is a **free** application that supports multiple instruments for spectral deconvolution in both GC/MS and data-independent MS/MS.
MS-DIAL5 tutorial:https://systemsomicslab.github.io/msdial5tutorial/

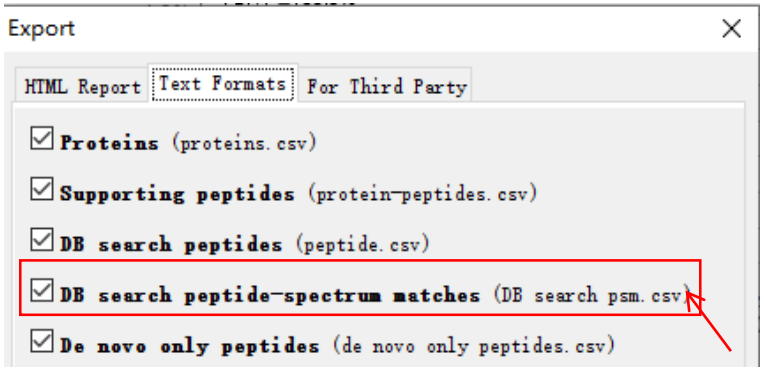


After completing the data import, click on '**Peak List Result**.' Then, export the result as a .txt file.

1. Prepare files: Peaks studio

PEAKS Studio is a software platform with complete solutions for discovery and targeted proteomics.
PEAKS Studio tutorial:https://www.bioinfor.com/peaks-studio/



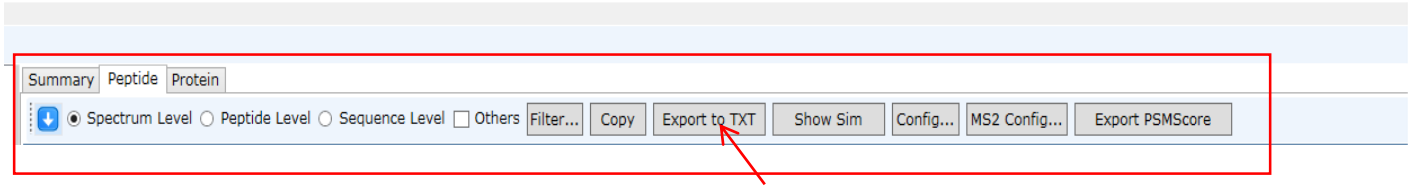


After completing the data processing, click on '**Export**'. Then, Click the **DB search peptide-spectrum matches.**

1. Prepare files: pFind

pFind Studio

pFind Studio is a **free** software platform with complete solutions for discovery and targeted proteomics.
pFind Studio tutorial:http://pfind.org/software/pFind/index.html

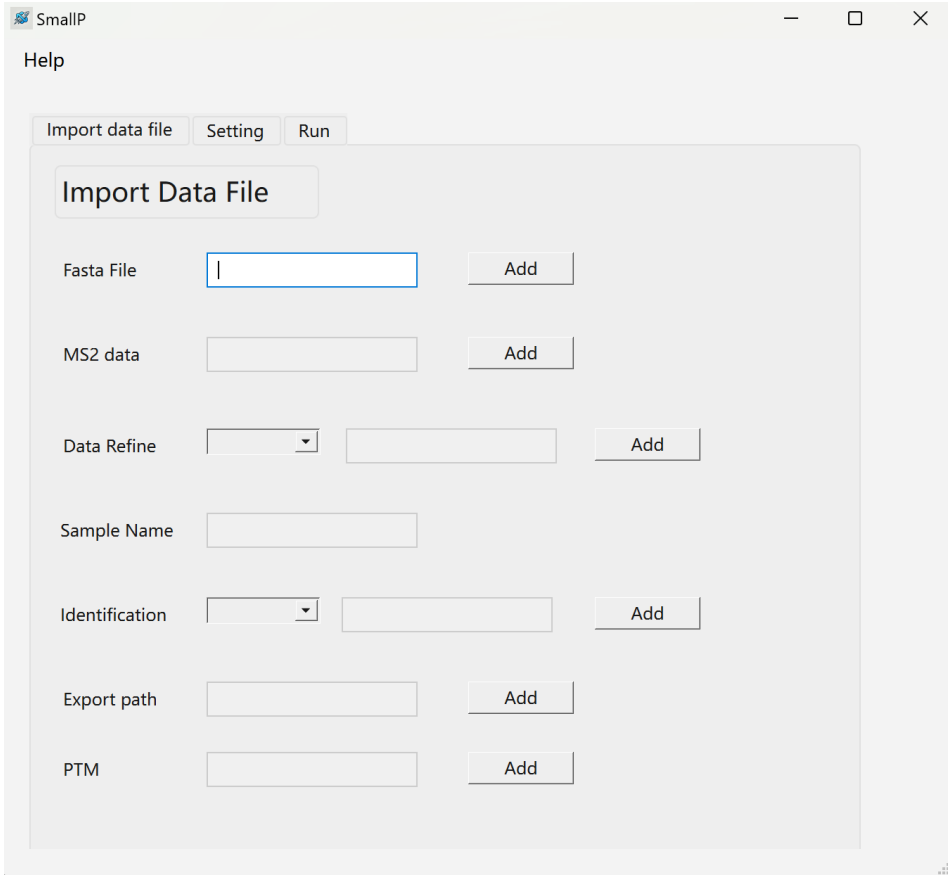


After completing the data processing, it will turn to pbuild program. Click on '**peptide**'. Then, Click the **Export to TXT.**

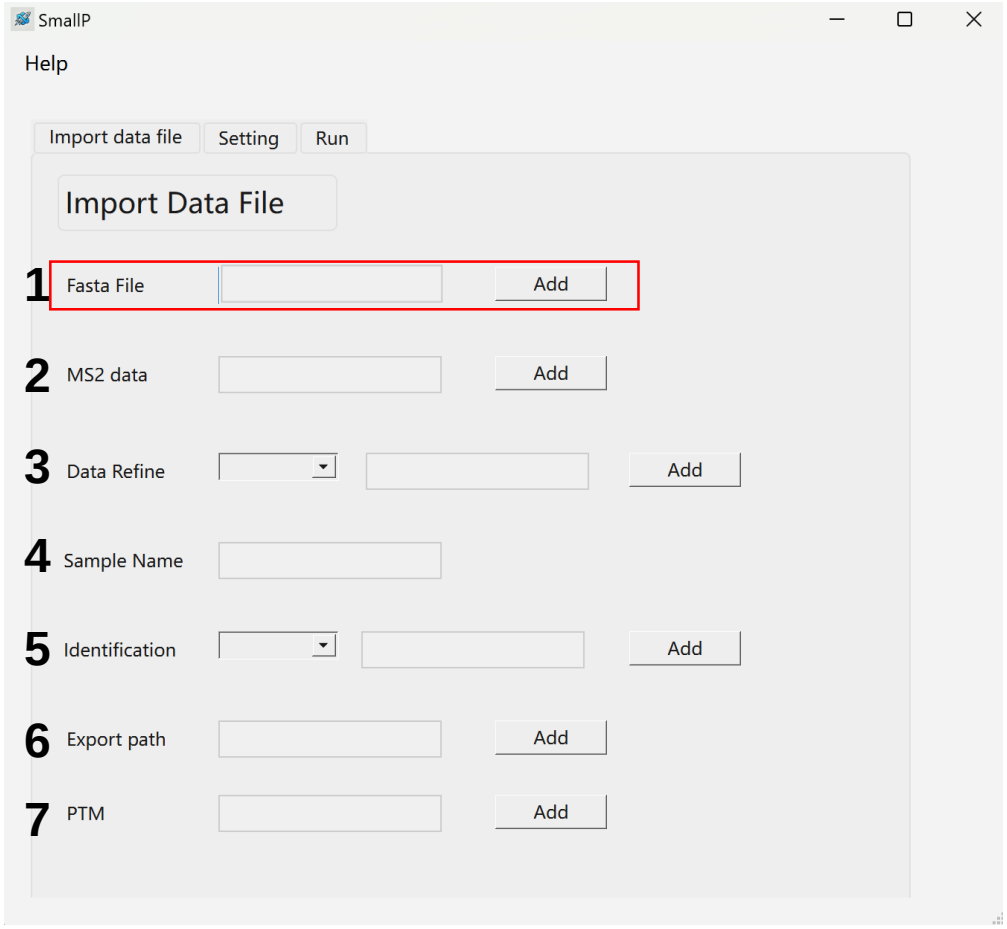
2.Start up

Small P is a free application intended solely for academic use. The program is easy to use for identifying disulfide bonds in peptides. By running Smallp.exe, you can import data, set parameters, and view the results.

For more detailed instructions:



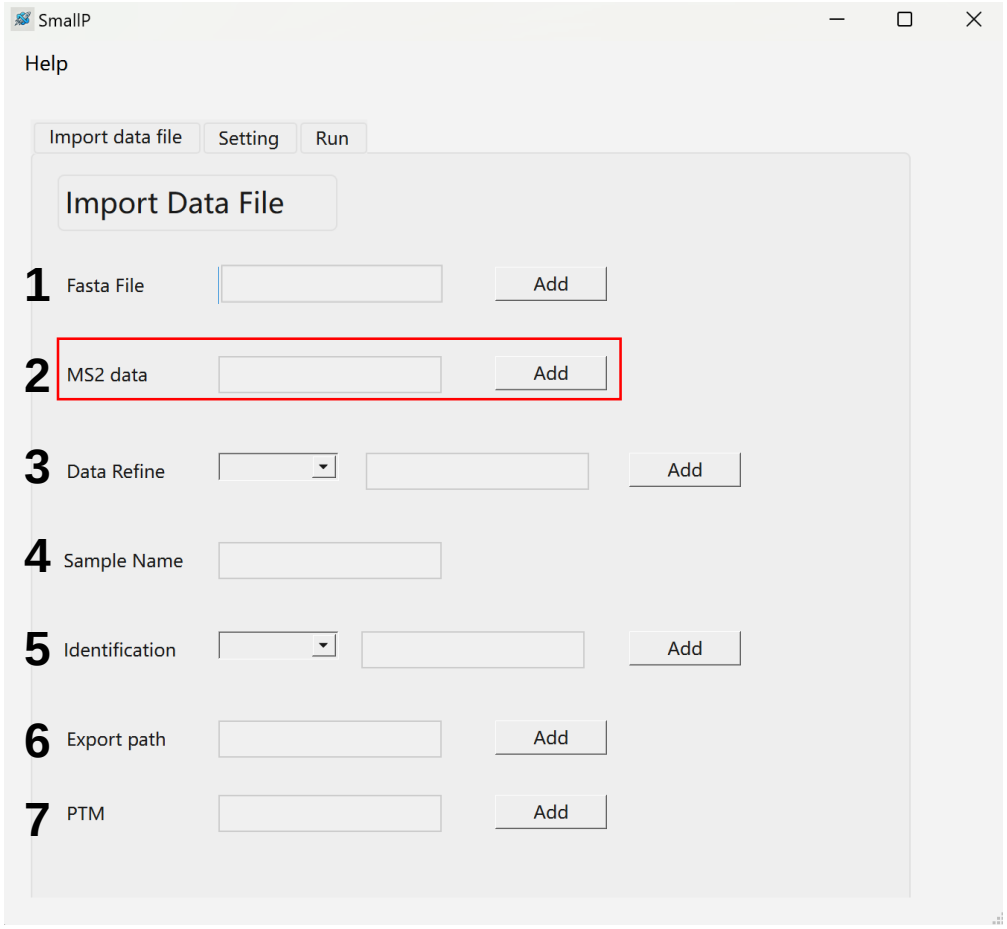
3.Upload datafiles: Database



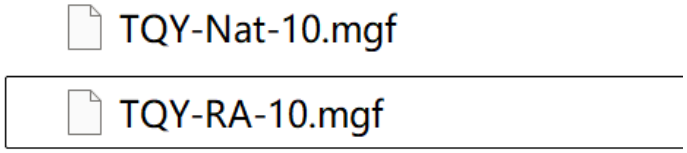
The software requires some profiles to function properly.

Fasta File: Click the “Add” button to select a fasta file (.fasta) as the input file. The fasta file should correspond to the search files you will upload subsequently.

3.Upload datafiles: mgf documents

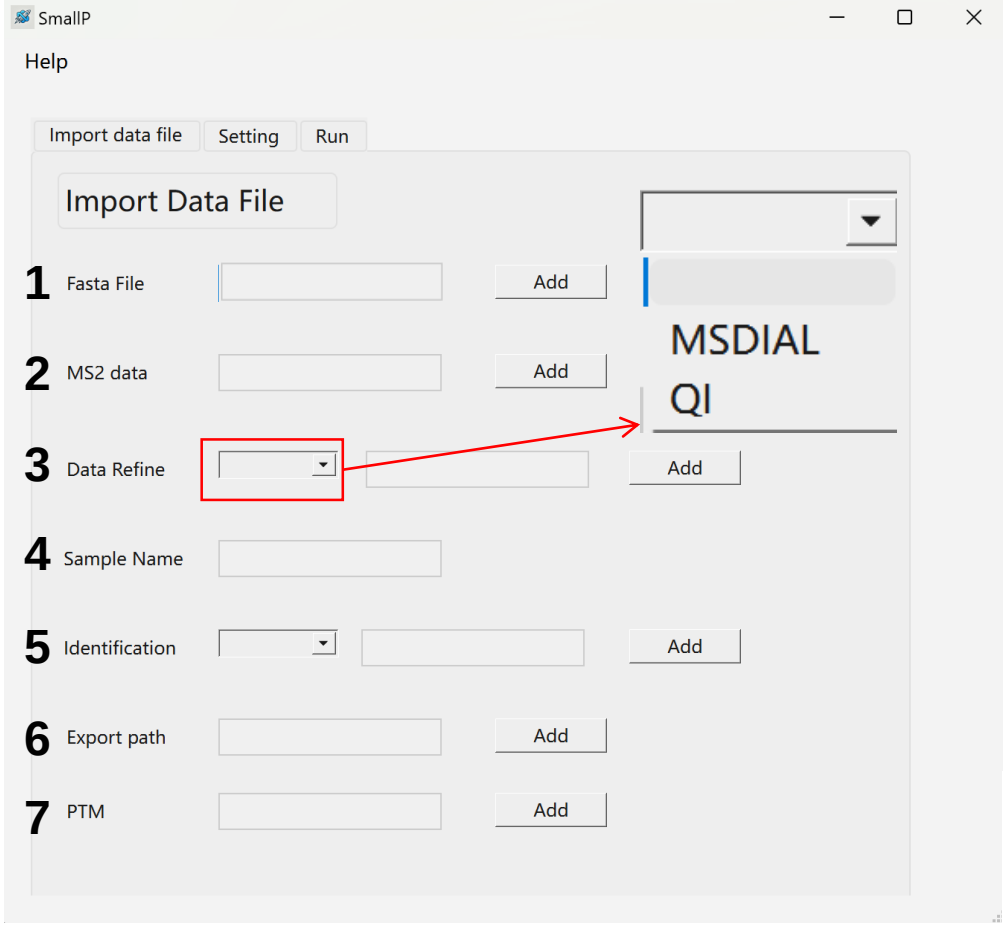


MS2 Data: Click the “Add” button to select the folder path containing the input file (.mgf). The file name should be formatted as xxx-Nat-num, xxx-RA-num, or xxx-RAE-num.

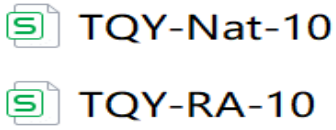


File naming example

3.Upload datafile: Peak detection



Data Refine: Click the “Add” button to select the folder path containing the results files from MSDIAL or QI. The drop-down box allows you to select software. These files could provide the MS information. **See the next section for detailed file export guidelines**.The file name should be formatted as xxx-Nat-num, xxx-RA-num, or xxx-RAE-num.

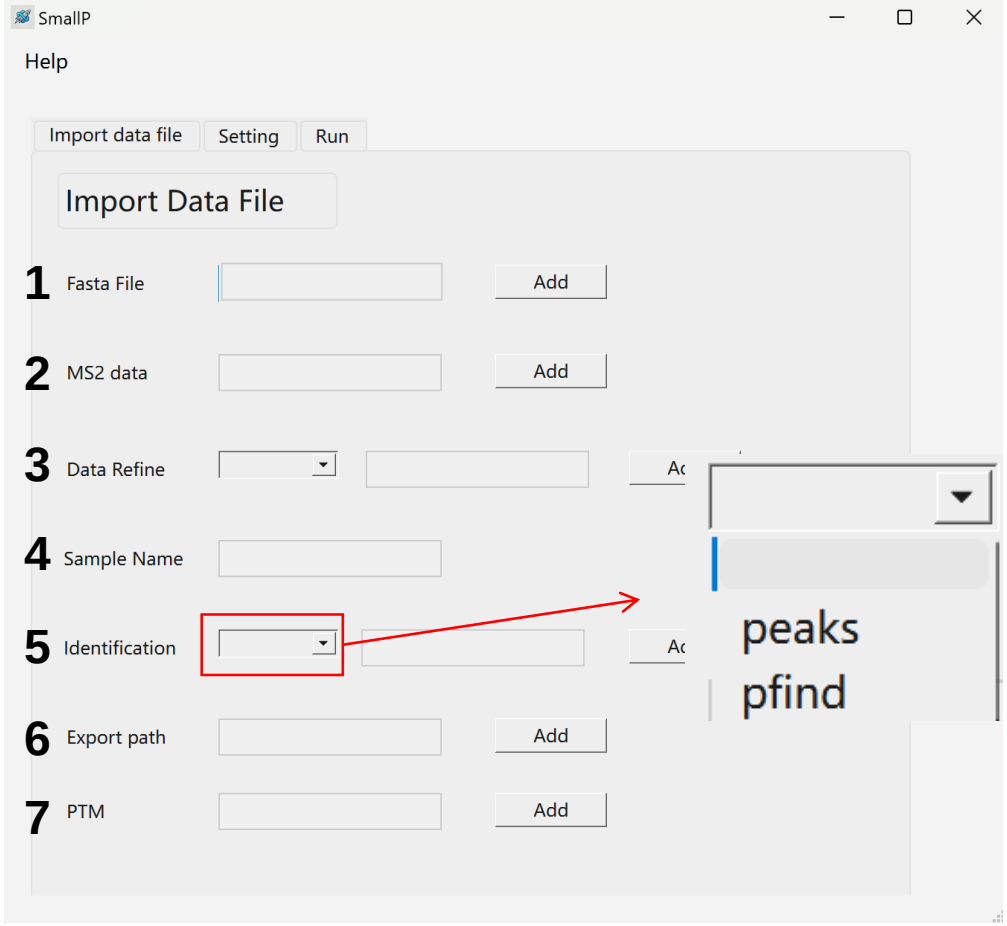


File naming example

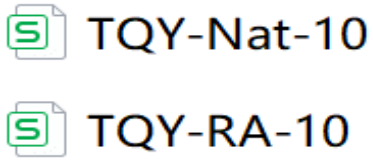
MSDIAL file export

QI file export

3.Upload datafile: Sequence identification



Identification: Click the “Add” button to select the folder path containing the results files from Peaks or pFind. The drop-down box allows you to select software. These files could provide the peptide identification information. **See the next section for detailed file export guidelines**. The file name should be formatted as xxx-Nat-num, xxx-RA-num, or xxx-RAE-num.

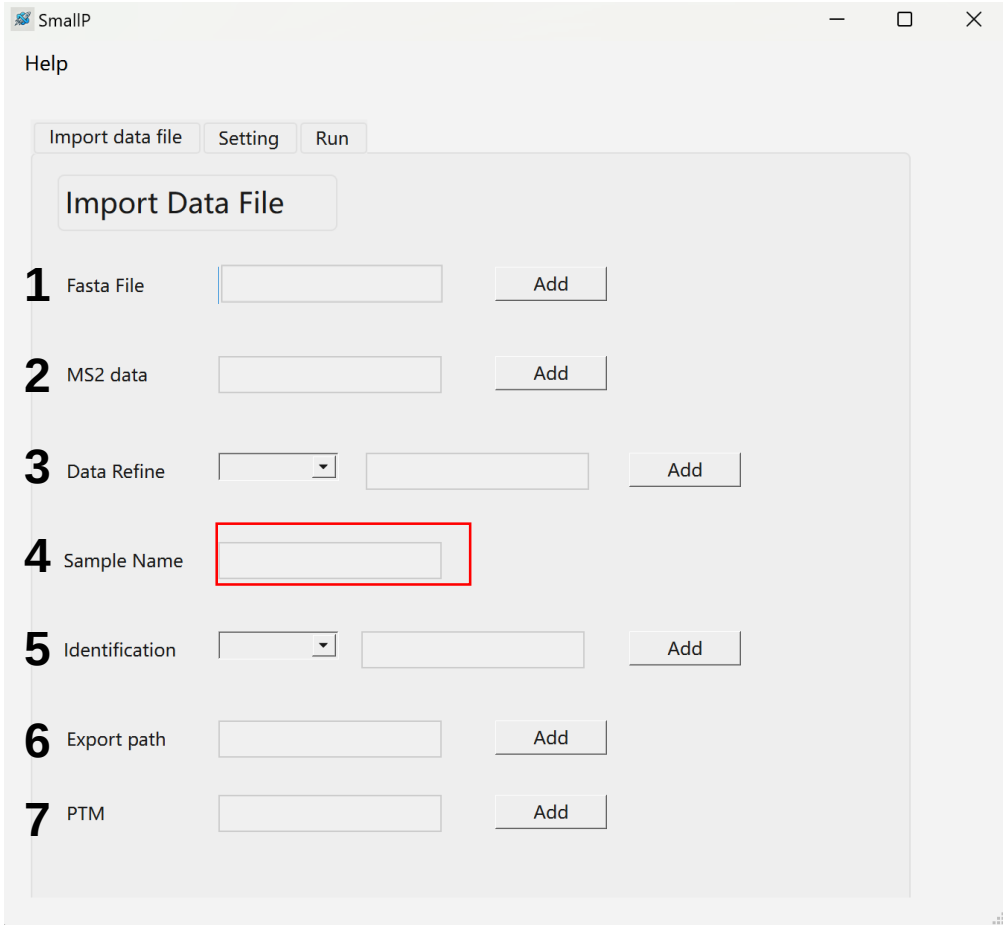


File naming example

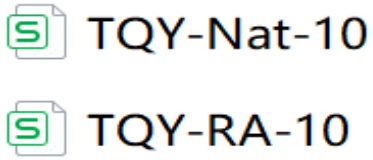
peaks file export

pfind file export

3.Upload datafile: Peak detection

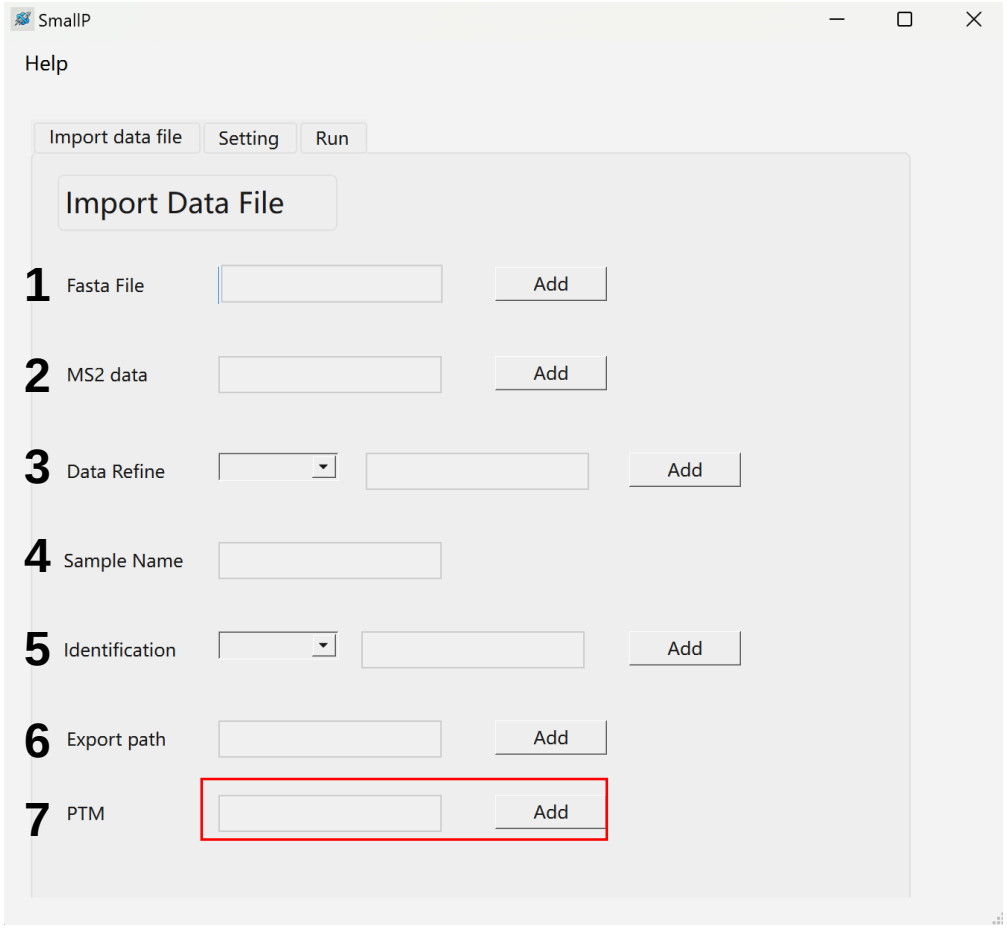


Sample Name: The file name should be formatted as xxx-Nat-num, xxx-RA-num, or xxx-RAE-num. The sample name should be in the format xxx.



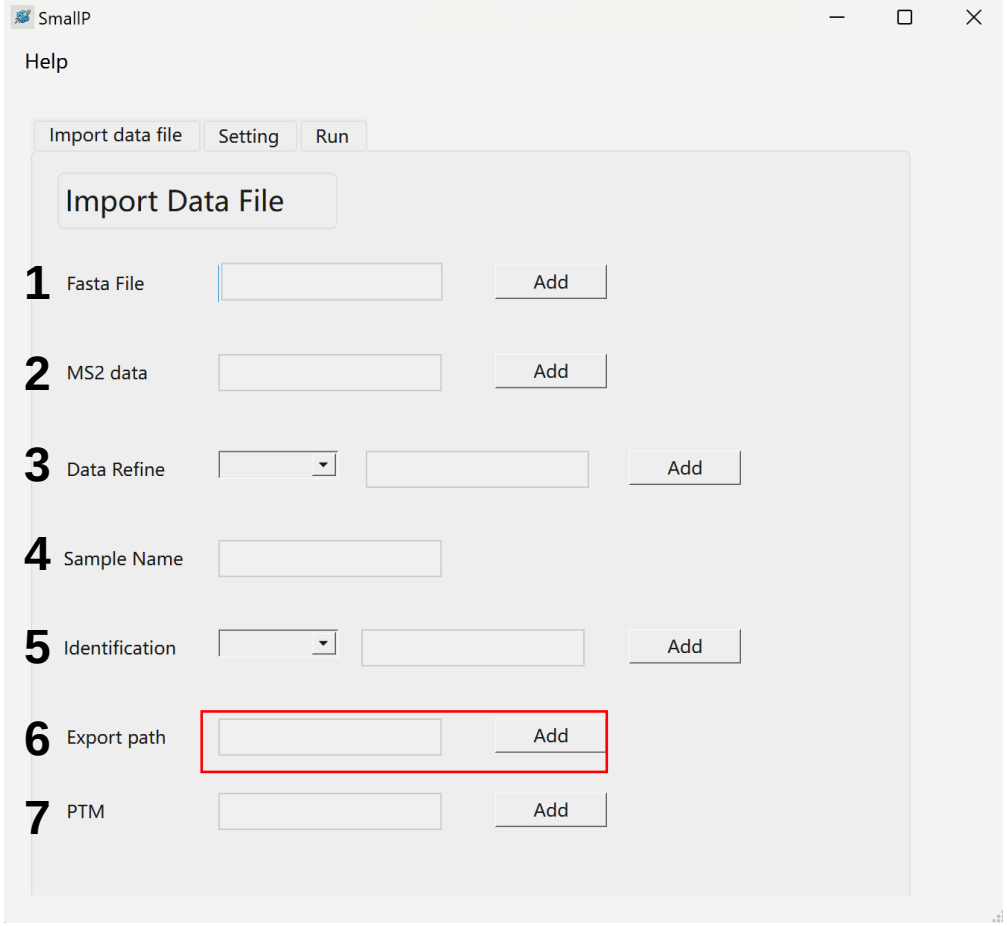
If the files are formatted as shown in the figure, the sample name should be “TQY”.

3.Upload datafile: PTM file



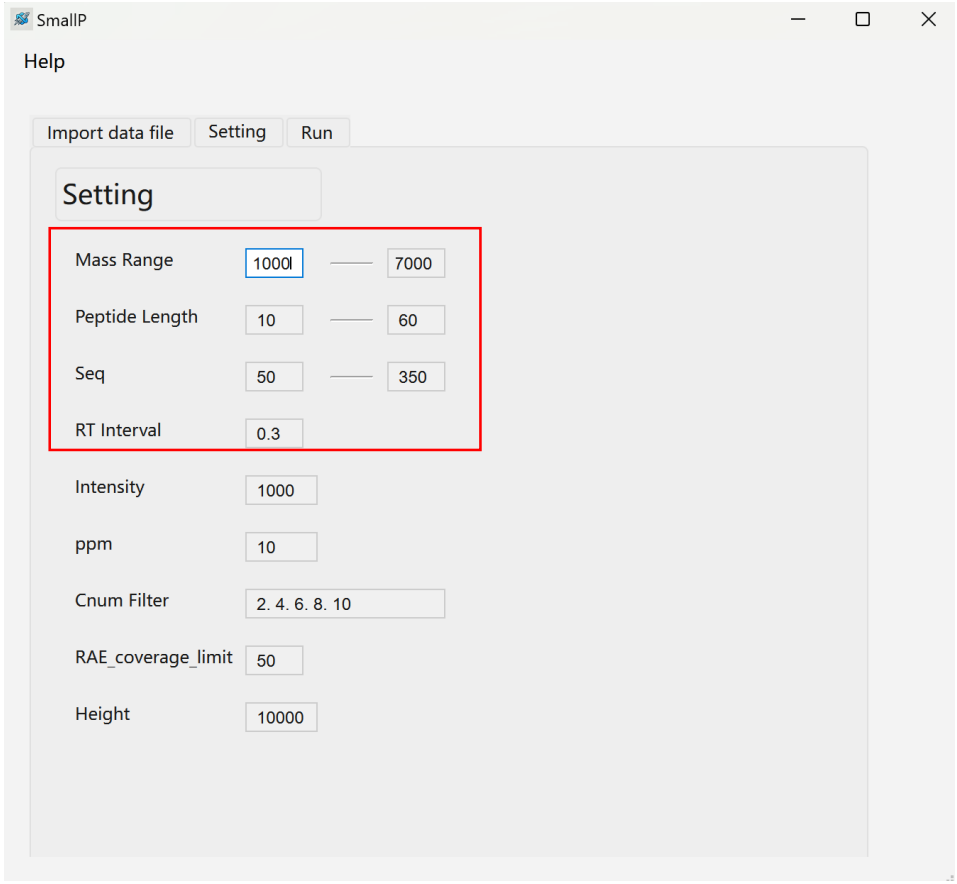
PTM: Click the “Add” button to select the file that contains the post-translational modification information (.xlsx).

4.Setting



Export path: Click the “Add” button to select the folder path where the final result files from the Smallp software will be saved.

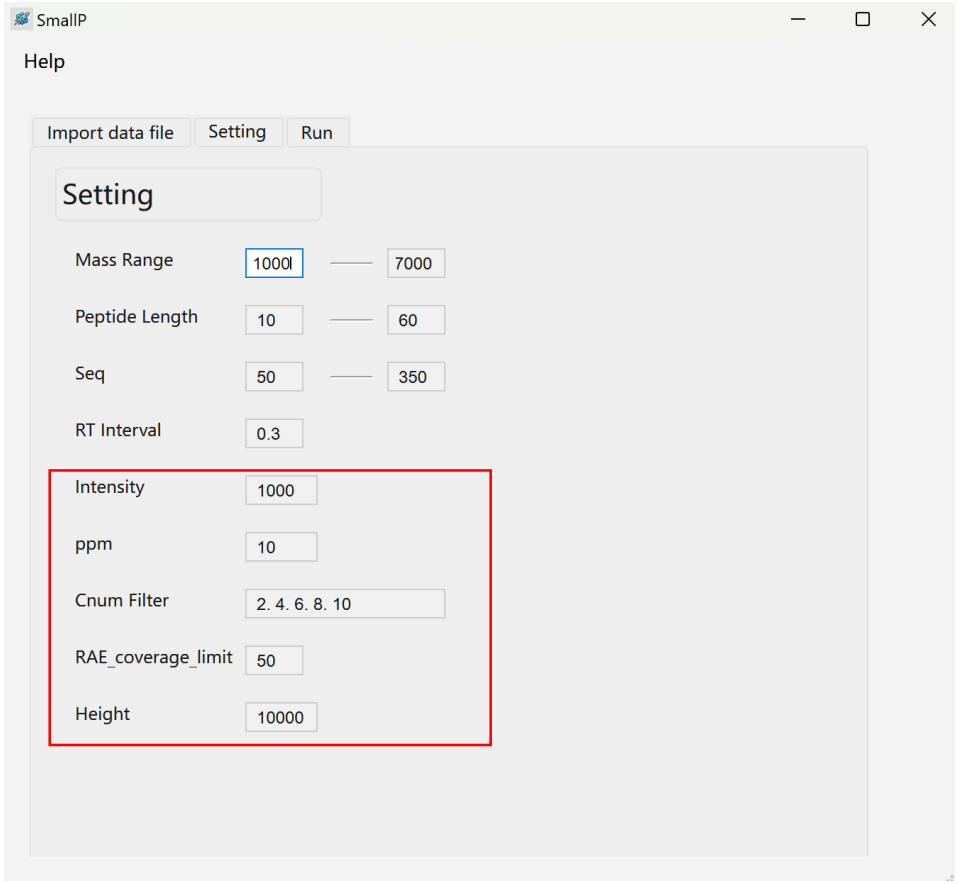
4.Setting



Setting: There are some default parameters you can change depending on your experiment design.

- **Mass range:** The range of mass to be analyzed.
- **Peptide length:** The range of peptide length to be analyzed in the database.
- **Seq:** The range of peptide length to be analyzed in the FASTA file.
- **RT:** The maximum retention time

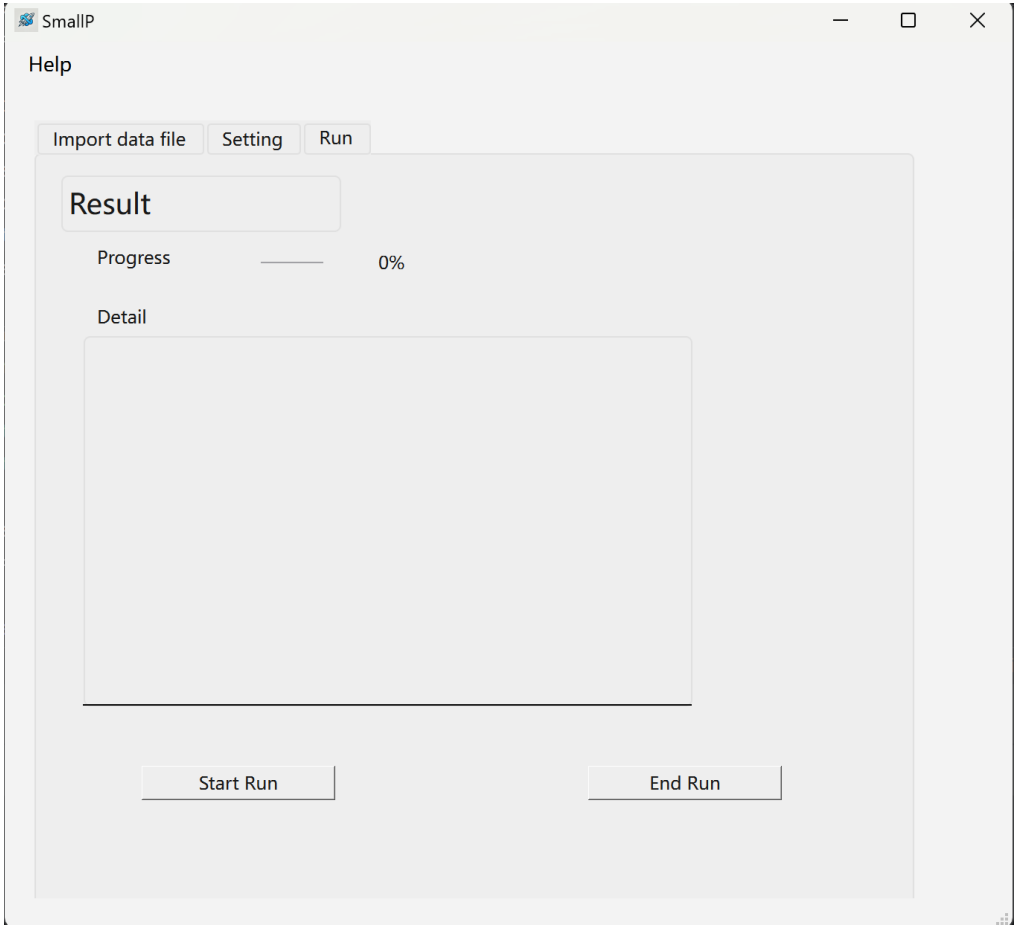
4.Setting



Setting: There are some default parameters you can change depending on your experiment design.

- **Intensity limit:** The minimum mass intensity in the data refine files.
- **Cnum filter:** The number of disulfide bonds you would like to search for. The number should be devided with”,”.
- **RAE coverage limit:** The minimum coverage of RAE.
- **Height:** The minimum ion intensity in the peptide identification files**.**

5.Run



Click the **“Start Run”** button to initiate the program. The progress bar will indicate the progress rate from 0 to 100. The results will be displayed in the detail window.
If an error occurs, you can click the **“End Run”** button to stop the program and rerun it.
There are totally 7 steps during this programs.