Spectra-comp user guide

Release 0.0.1

Table of content

1) Spectra-comp

Spectra-comp is an MS/MS spectra comparison tool. It helps the users to compare between MS/Ms spectra using dot product method.

2) File:

a. Load spectra

The user could load spectra from:

- Proline project:

The user must have an account Proline. The user selects its project and the result summary to set as reference spectra or as test spectra.

Note: You should have Proline software installed on your machine. If you don't have proline installed, you can download it from https://github.com/profiproteomics

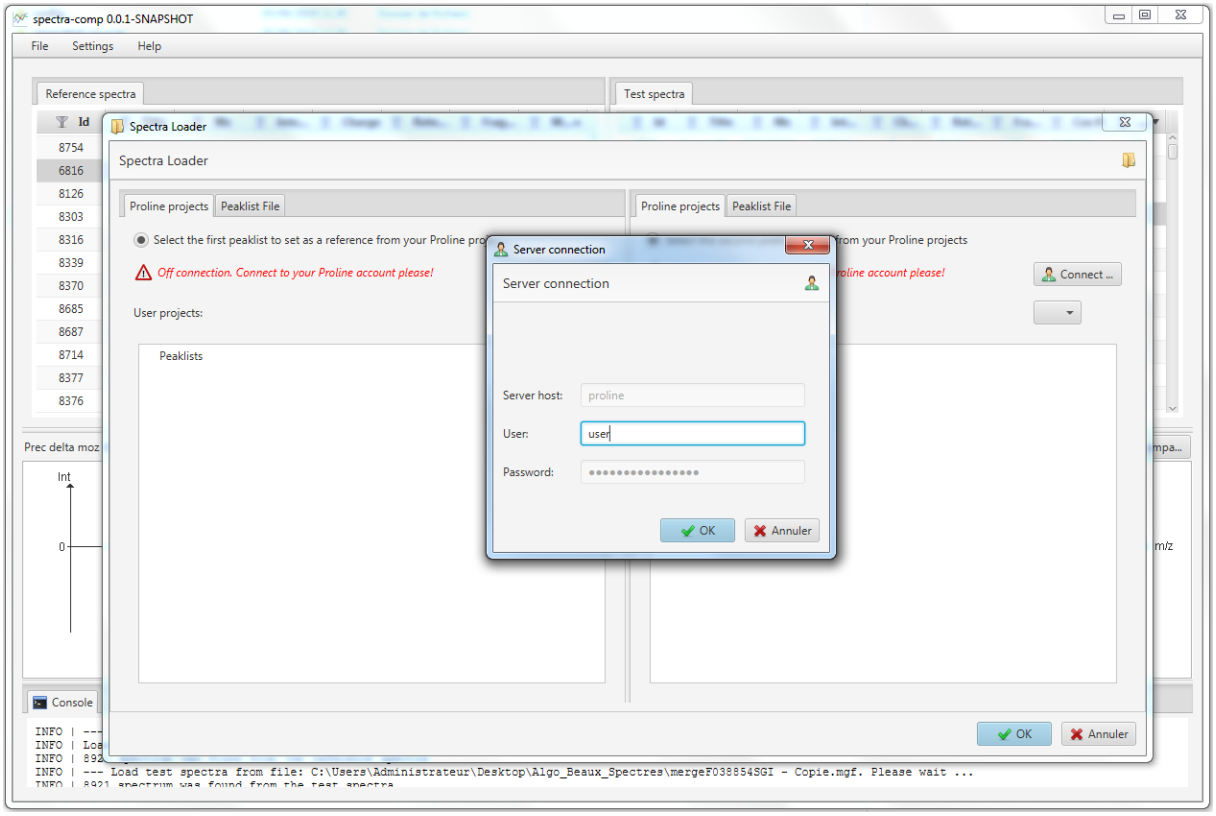
If you don't have an account Proline. You should ask the IT administrator to create you an account. To connect to your proline account, you must enter your Proline login.

Note: you can edit the connection configurations to Proline databases from the "application.conf" file under the folder "../config/application.conf"

```
db-config.driver-type=postgresql

db-config.max-pool-connection=3

//User and Password to connect to databases server

auth-config.user=proline

auth-config.password=Proline_Database

//Databases server Host

host-config.host=proline

host-config.port=5432

db-name=uds_db
```

- Peaklists file:

The user selects a valid peak list file (extension .mgf or .pkl). Spectra-comp will read the peaklist file and extract the spectra. The user chooses the file from where to extract spectra and choose to set it as reference or as test spectra.



In the main view, there are two table views: Reference spectra and test spectra. Each table view contains the information about the loaded spectra.

- Title: The title of the spectrum.
- Precursor (m/z, intensity).
- Fragment number: the fragment number found in the selected spectrum.
- Charge.
- Retention time.
- Fragment number.
- Matched: the number of matched spectra in reference spectra and checked Boolean in test. spectra.
- Cos Θ in test spectra.

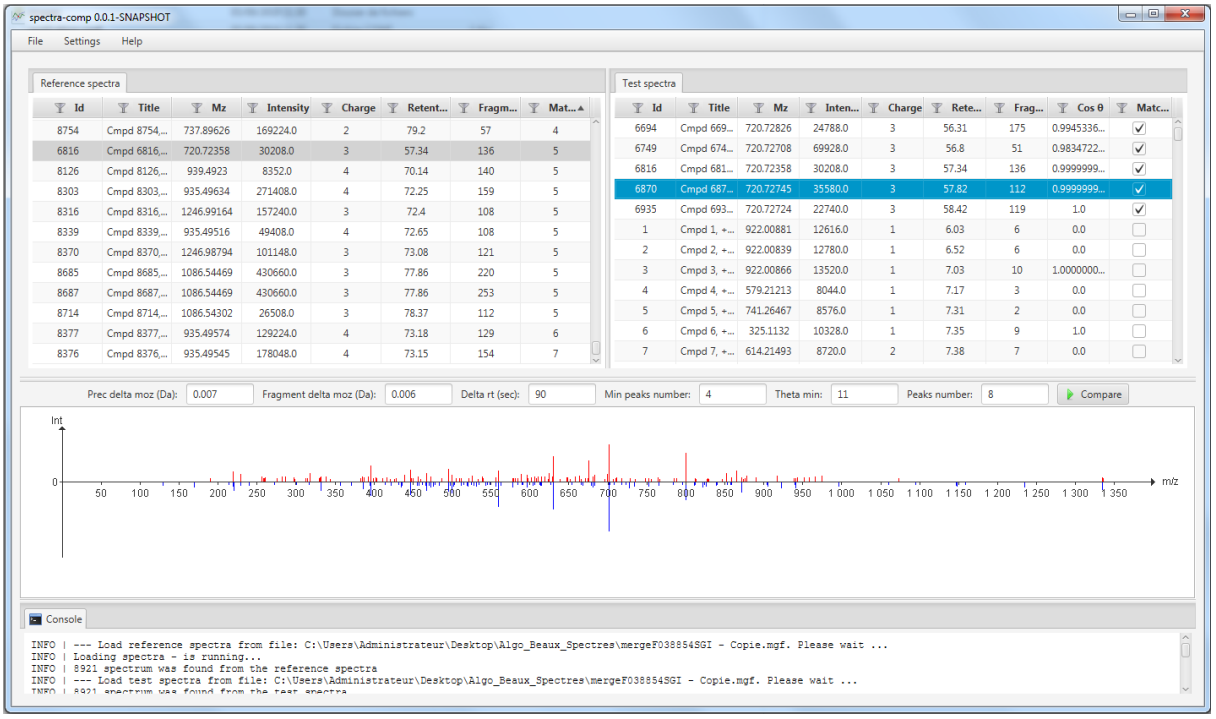
The filters that the user can apply on columns are:

- Boolean: Applied on test spectra column “Matched”.
  The values of the filter could be: True, false or not applied.
- Numeric: Applied on columns: Moz, intensity, charge and retention time.
  The values of filters could be: equals, not equals, greater than, greater than or equals, less than or equals or less than.
  Note: the user can search by a range [GREATER THAN, LESS THAN] or
  ] EQUALS GREATER THAN, LESS THAN OREQUALS] …
- String (Sequence of characters): Applied on columns: for example on spectrum title.
  The values of filters could be:
  -Equals: equals the entered value.
  -Not equals: different from the entered value.
  -Contains: will search on column all the values that contains the entered value.

-Starts with: will search on column all the values that starts with the entered value.
-Ends with: will search on column all the values that ends with the entered value.
Note:
The users may mix the filters. It will be applied as: filter1 AND filter2 AND filterN…
The final result it's Boolean of filter1 AND filter2 AND filterN…

In the main view, the chart represents the selected spectrum from the reference spectra table view and its selected, tested spectrum from the test spectra table view. There are some actions to visualize the graph:

- Zoom: Select an area.
- Cancel zoom: left click.
- Show a sequence of AA: Select a fragment and select the others fragments.
- Cancel the comparison: Ctrl+right click.



b. Export comparison:

This action will export the result of a comparison between the reference spectra and the test spectra.

The result will be exported in pdf file. It contains the source of reference spectra and the source of test spectra (proline project or a peaklist file) as well as, it contains the used parameters in comparison.

c. Exit:

This action will close the resources (database connections when the spectra loaded from Proline projects) and close the spectra-comp software.

3) Settings:

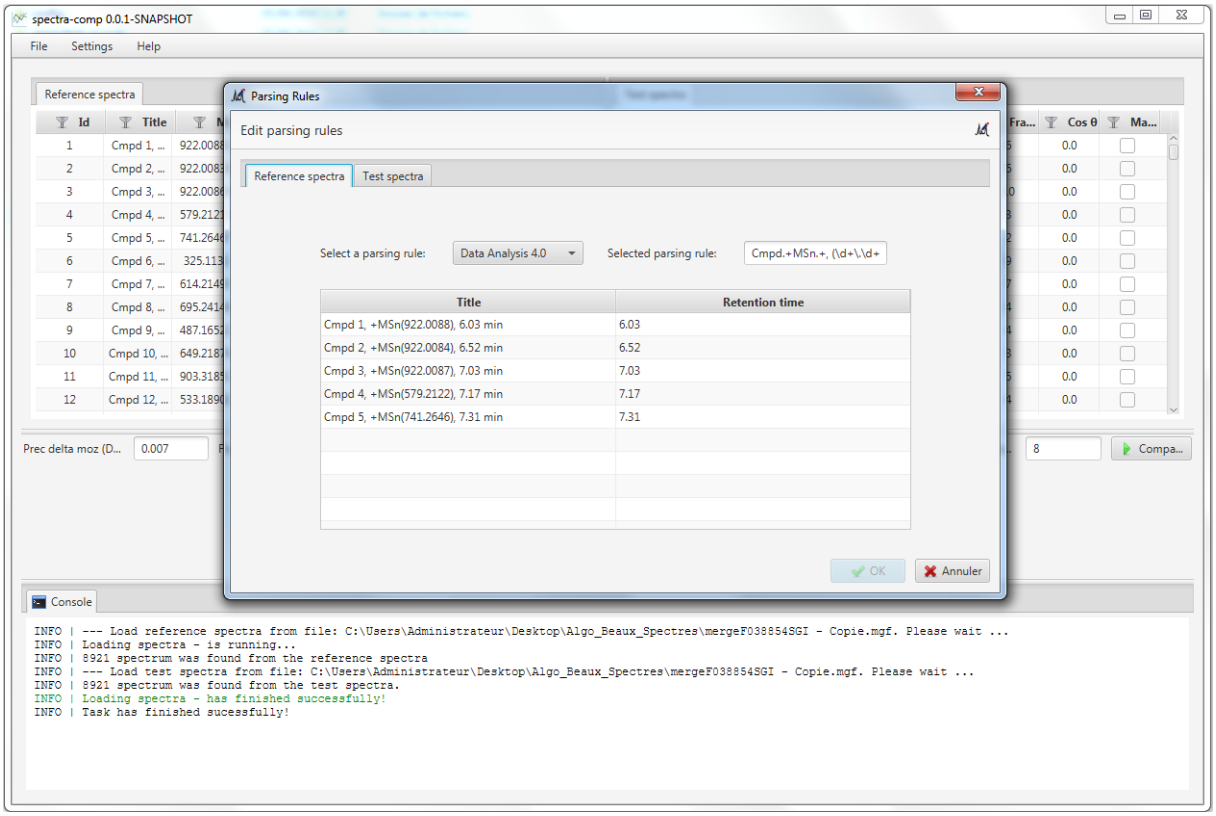
a. Parsing rules

This action lets the user to edit the parsing rules used to retrieve retention time from the title of spectrum. The defined parsing rules are:

- data-analysis*: Cmpd.+MS.+, (\d+\.\d+) min*
- mascot-dll: *Elution: (.+?) to .+? min|Elution: (.+?) min*
- Proline: *first_time:(\d+);*
- msconvert: *RTmin:"(\d+\.\d+)"*
- mascot-distiller: *in range \d+ \(rt=(\d+.\d+)\)|\(rt=(\d+.\d+)\)*
- …

To check the result within the parsing rule, Recover extracts the top five spectrums and shows the result. On apply button the selected parsing will be used, among all the spectra.

Note: The used parsing rule to retrieve the retention time from of titles is applied separately for reference and test spectra.



a. Comparison parameters

This action lets the user define the comparison parameters:

- Precursor delta Moz: the precursor delta Moz (Da).
- Fragment delta Moz: The fragments delta Moz(Da).
- Delta retention time: The delta retention time (seconds).
- Number of minimum peaks: the minimum number of peaks that matched between the reference spectrum and the tested spectrum.

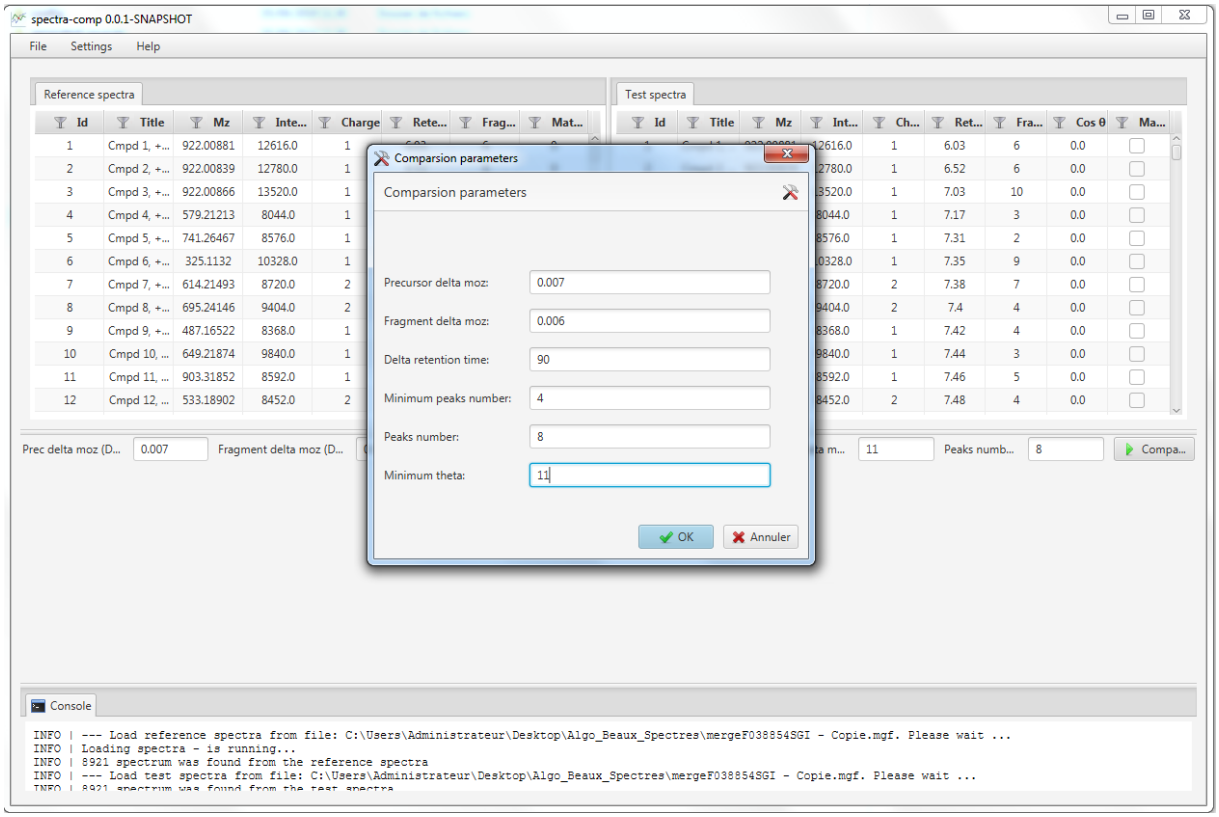
- Number of peaks: the number of peaks that matched: for the tested spectra if a value is present for the same peak of reference spectra, we keep the most intense value of intensity.
- Minimum theta: this value is computed as `Math.cos(Math.toRadians(thetaMin)).`

To compute Cos Θ we use this formula:

```
Cos theta =
∑NB_PEAKS(√RS.peak*√TS.peak)/(√(∑NB_PEAKS(RS.peak))*√(∑NB_PEAKS(TS.peak)))
```

And

```
TS.peak = intensity of peaks TS(tested spectra)
RS.peak = intensity of peaks RS(reference spectra)
```



1) Help

a. User guide

This action will open the user guide file (spectra_comp_user_guide.pdf).

a. About

This action will show the software properties.