MS Ana Spectral Library Search Node for Proteome Discoverer

User Manual

May 2022

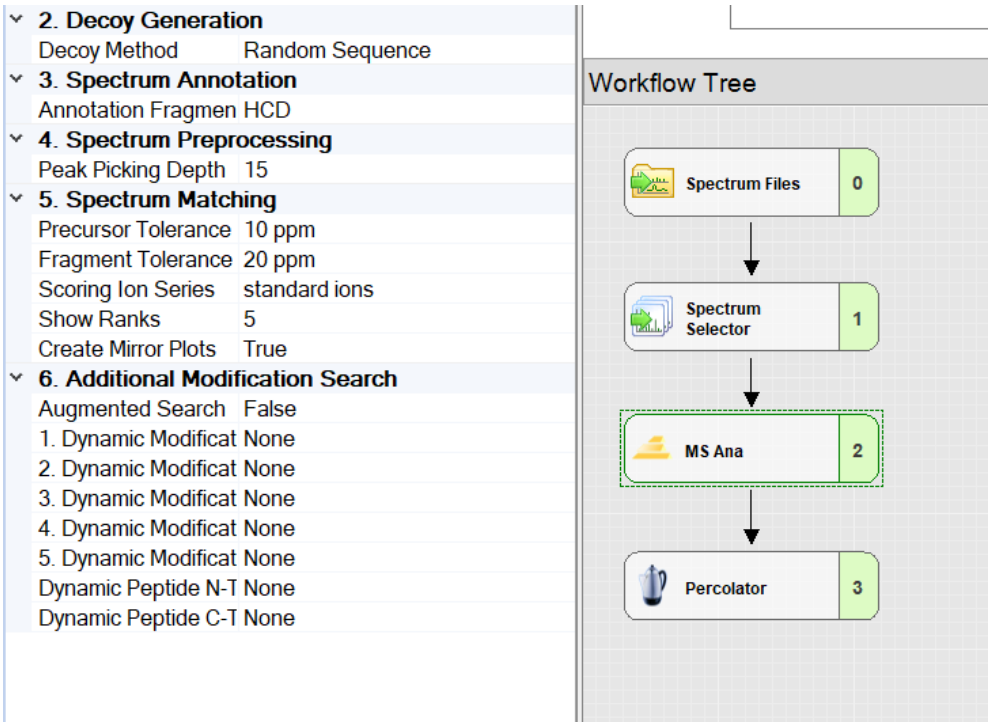
MS Ana is a scoring system to identify peptides in tandem mass spectrometry data using a library of previously identified spectra. Furthermore, MS Ana can create decoy spectral libraries for validation and run searches for additional modifications not found in the spectrum library.

Installing MS Ana

Make sure Proteome Discoverer is closed. Then, run the MS Ana setup file **MS_Ana_PDXX_Node.exe** and accept the license agreement. The setup will copy all relevant files into the Proteome Discoverer directory and register the node. Wait until the console window closes then hit Finish to exit the setup. The MS Ana Spectral Library Search Node is now ready to use! ***Please note: If you are experiencing problems during the installation because of Windows security services you must use the provided certificate to add FHOÖ & IMP as a trusted software source. Double-click the*** **FHOOE_IMP_CERT.cer*****, then choose 'Install Certificate...' and follow the instructions on screen.***

Running MS Ana

Using MS Ana in Proteome Discoverer is similar to using any other identification node like Mascot or SEQUEST. Simply drag & drop the MS Ana node into your workflow and connect it to its predecessor and successor nodes.



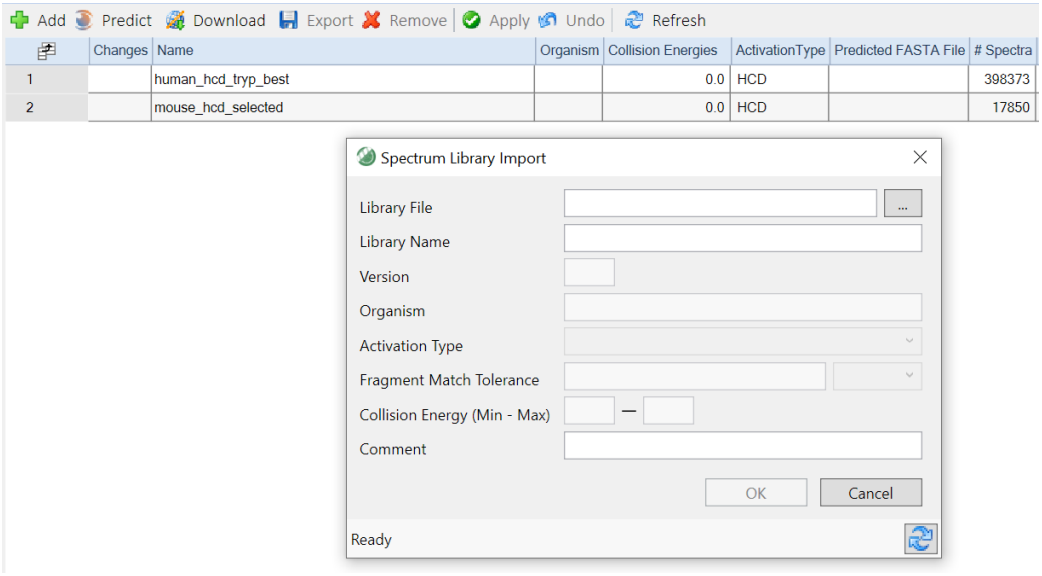
At minimum, you will need to select a spectral library that was previously registered to Proteome Discoverer. Afterwards, please set the parameters (described in a later section) according to your experiment and click the run button.

Please note: The first time you run a search with a spectral library, MS Ana will create the decoy library to allow for result validation using the target-decoy approach. This will take considerably more time than the library search (depending on library size). Subsequent runs will be much faster since the decoy library is stored and used again.

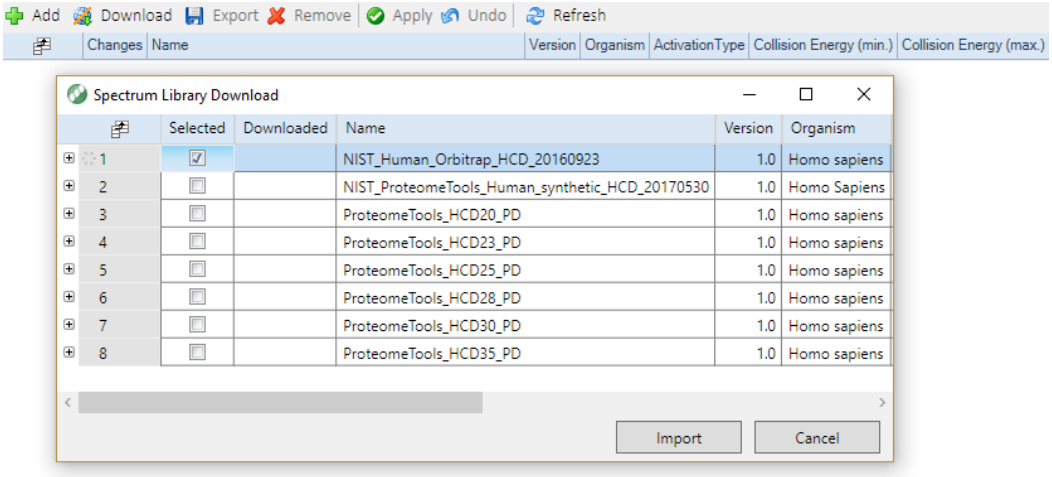
Adding Spectral Libraries to Proteome Discoverer

To run a workflow with MS Ana you must first add a spectral library using the Spectrum Library interface in Proteome Discoverer. Curated spectral libraries can be downloaded from the NIST Mass Spectrometry Data Center or the MassIVE Knowledge Base.

You can import external spectral libraries into Proteome Discoverer. Navigate to Administration: 'Maintain Spectrum Libraries'. This opens a summary view of active spectral libraries. Choose (+) Add to import a new library. Proteome Discoverer can import external libraries in the .MSP, .SPTXT, or NIST binary format.



Alternatively, libraries for different standard organisms can be downloaded directly through Proteome Discoverer. In Proteome Discoverer navigate to Administration: 'Maintain Spectrum Libraries' then select Download to add standard libraries to the Proteome Discoverer database.



Parameters of MS Ana

The following parameters can be set for configuring the MS Ana scoring system:

1. **Input Data**
    a. **Spectral Library**
    The spectral libraries that will be used by MS Ana for peptide identification. Please select one or several libraries that were previously added using the Proteome Discoverer Spectral Library interface.
    Please note: The first time you run a search with a spectral library MS Ana will create the decoy library to allow result validation using the target-decoy approach. This will take considerably more time than the library search (depending on library size). Subsequent runs will be much faster since the decoy library is stored and used again.
    b. **Protein Database**
    The protein sequence database(s) used to map the identified peptides to proteins. Please select any FASTA database(s) previously added using the Proteome Discoverer FASTA File Interface.
2. **Decoy Generation**
    a. **Decoy Method**
    Select the algorithm to be used for the generation of decoy library spectra:
        i. **Random Sequence (Default)**
        Shuffle peptide sequence randomly (except terminals) and move annotated fragment ion peaks accordingly.
        Target S**ESVVYADI**K
        Decoy S**VDIESVYA**K
        ii. **Reverse Sequence**
        Shuffle peptide sequence randomly (except terminals) and move annotated fragment ion peaks accordingly.
        Target S**ESVVYADI**K
        Decoy S**IDAYVVSE**K
        iii. **Shift Sequence**
        Shift peptide sequence by a set number of amino acids (except terminals) and move annotated fragment ion peaks accordingly.
        Target S**ESVVYADI**K
        Decoy S**IESVVYAD**K
3. **Spectrum Annotation**
    a. **Annotation Fragmentation Type**
    Choose fragmentation type that will be used during annotation of spectra. Influences which types of ions will be annotated.
4. **Spectrum Preprocessing**
    a. **Peak Picking Depth**
    Specify the depth for filtering low intensity peaks in the query and library spectra during preprocessing. Only the N highest peaks per 100 m/z window will be retained.

5. **Search Settings**

   a. **Precursor Tolerance**

   The tolerated mass error when matching precursor mass values (MS1 mass error). Possible units are “Da” and “ppm”.

   b. **Fragment Tolerance**

   The tolerated mass error when matching fragment ion mass values (MS2 mass error). Possible units are “Da” and “ppm”. Annotation of fragment ion peaks during decoy spectrum generation will also use the selected Fragment Tolerance.

   c. **Scoring Ion Series**

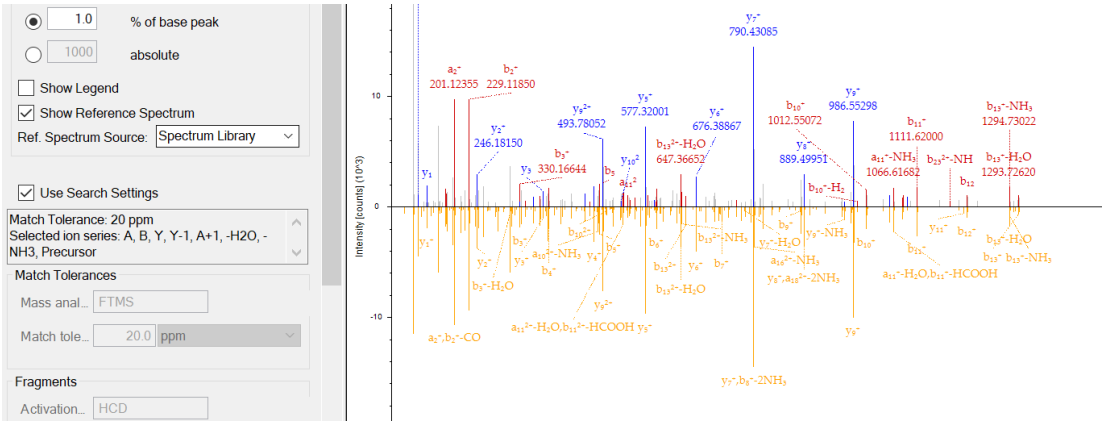
   Pick which fragment ion series will be used in the scoring function. Select ‘standard’ to use a sensible default depending on the fragmentation type. (Standard for HCD fragmentation is y-ions and b-ions)

   d. **Show Ranks**

   Specify how many top hits per spectrum will be reported.

   e. **Create Mirror Plots**

   If set to ‘True’, mirror plots will be created for each identification and saved to the result file. Mirror plots can be viewed in the Peptide Spectrum Match Identification Details of the analysis results. Make sure to check ‘Show Reference Spectrum’ to view mirror plots. Please note: Mirror plot creation can significantly increase processing times and the size of the result files.



Contact

This research project is a collaboration of the Protein Chemistry Group at IMP and the Bioinformatics Research Group at FH Upper Austria, Hagenberg Campus. For any further questions, bug reports, or ideas please contact Sebastian Dorl, Viktoria Dorfer, Stephan Winkler, or Karl Mechtler.