ProSightPD 1.1 Quick Start Guide

The ProSightPD™ application assists you in performing top-down searches in the Thermo Proteome Discoverer™ application. This quick start guide briefly explains the software required to use the ProSightPD application, ProSightPD nodes, ProSightPD workflows, and ProSightPD-specific information contained in the search results.

This guide assumes that you are familiar with both the ProSightPC™ and Proteome Discoverer applications. For detailed information about these applications, refer to their respective user guides: the *ProSightPC User Guide* or the *Proteome Discoverer User Guide*.

Contents

The ProSightPD application uses the following terminology:

- **High/High**—FTMS data with isotopically resolved MS1 and MS2 data.
- **Medium/High**—Short transient low-resolution FT MS1 data and isotopically resolved MS2 data.
- **Low/High**—Ion trap resolution MS1 data and isotopically resolved MS2 data.

System Requirements

The Proteome Discoverer and the ProSightPD applications have configuration requirements that your data system must meet before the software operates correctly.

Hardware Requirements

The *Proteome Discoverer Release Notes* lists the minimum and recommended hardware system requirements.

Table 1. Minimum and recommended hardware requirements for Proteome Discoverer

System	Requirements
Hardware (minimum)	• Data system computer with 2 GHz processor • 2 GB RAM • DVD-ROM drive • Video card and monitor capable of 1280 ×1024 resolution (XGA) • Screen resolution of 96 dpi • 75 GB available on drive C • NTFS format
Hardware (recommended)	• Data system computer with two Intel™ Xeon™ 6-core processors, 2.4 GHz • 24 GB RAM • DVD-ROM drive • Video card and monitor capable of 1280×1024 resolution (XGA) • Screen resolution of 96 dpi • 1 TB available on drive C

Revision A XCALI-97902

thermoscientific

Software Requirements

You must install the following software packages to use ProSightPD version 1.1 in the Proteome Discoverer application.

System	Requirements
Software	• ProSightPC • Proteome Discoverer (x64 version) • ProSight Lite (freeware from prosightlite.northwestern.edu) • Thermo Proteome Discoverer ThirdParty Components installer (also installs the ProSightPD nodes) • Microsoft™ Windows™ 7 x64 Professional (English version) with current service pack

Installing the ProSightPD Nodes

The Thermo Proteome Discoverer ThirdParty Components Installer installs the ProSightPD nodes.

Note ProSightPD nodes that you install with the Thermo Proteome Discoverer ThirdParty Components Installer are automatically licensed for 60 days.

❖ **To install ProSightPD**

Select the **ProSightPD** check box in the Select Features dialog box.

Contact your local Thermo Fisher Scientific sales representative or email info@proteinaceous.net to obtain information on how to license ProSightPD and ProSightPD High Mass.

ProSightPD Nodes

Installing the ProSightPD application adds the following nodes to the Proteome Discoverer application:

- ProSightPD Top-Down High/High cRAWler node—Uses the Xtract algorithm to create experiments from High/High data.
- ProSightPD Absolute Mass Search node—Performs a ProSightPC absolute mass search.
- ProSightPD BioMarker Search node—Performs a ProSightPC biomarker search.
- ProSightPD Gene-Restricted Absolute Mass Search node—Performs a ProSightPC gene-restricted absolute mass search.
- ProSightPD Gene-Restricted BioMarker Search node—Performs a ProSightPC gene-restricted biomarker search.
- ProSightPC Puf Exporter node—Creates PUF files that you can open in the ProSightPC application.

Additional Nodes

Installing the high-mass version of the ProSightPD application adds these two additional nodes:

- ProSightPD Spectrum Selector node—Replaces the default Proteome Discoverer Spectrum Selector node in Low/High and Medium/High searches. It collects the profile data that is required to perform these searches.
- ProSightPD Top-Down Low/High cRAWler node—Uses the ReSpect algorithm to create experiments from Low/High and Medium/High data.
- Parallel Xtract—Transforms spectra so that they contain only singly charged ions.

Searching with ProSightPD

To use the ProSightPD application in a search, follow these topics:

- ProSightPC Search Requirements
- Downloading a ProSightPD Database in ProSightPC
- Creating a StudyDefault ProSightPD Templates
- Setting Up a ProSightPD Database Search
- Adding Files to the Study and Selecting Files

ProSightPC Search Requirements

Every ProSightPD search requires a ProSightPC 4.0 database (PSCW). You can create the ProSightPC database in three ways:

1. Download the database directly from the Proteinaceous web page.
2. Download the ProSightPC database through the ProSightPC application.
3. Create a database using a FASTA file, a UniProtKB XML file, or a UniProtKB flat file with the ProSightPC Database Wizard.

Downloading a ProSightPD Database in ProSightPC

❖ To download a ProSightPC database through the ProSightPC application

1. Start the ProSightPC application.
2. Choose **Databases > Download ProSightPC Databases**.

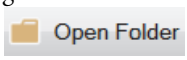
   From proteinaceous.net, you can access top-down and bottom-up databases.
3. Click the date of the database that you are interested in, for example, **July 2016**.
4. Click **Archaebacteria**, **Eukaryotes**, **Prokaryotes**, or **Custom**, as appropriate.

   You must know the taxonomy of the species of the database that you want to download.

   You can choose from one of the following types of databases:

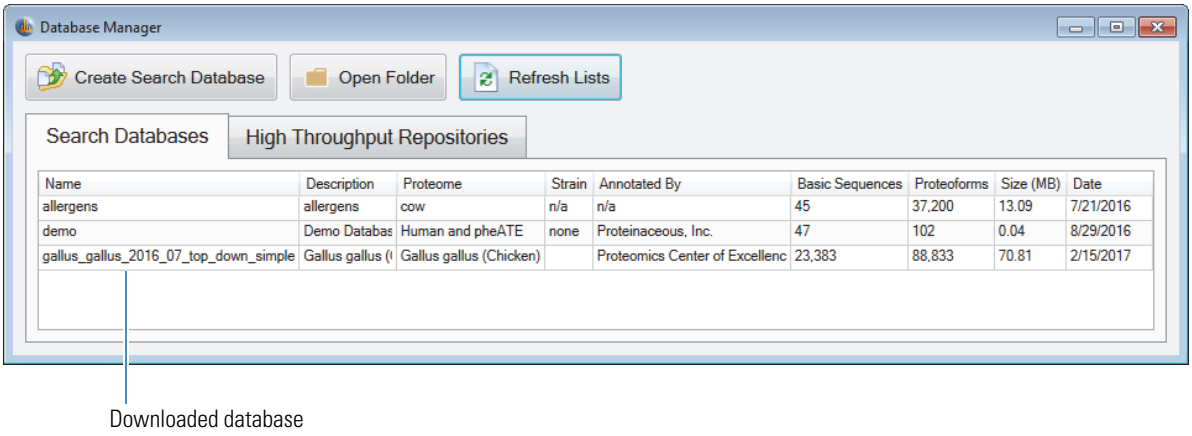
   - TD Complex PSCW—Includes N-terminal acetylation and initial methionine cleavage. This database contains 12 to 15 modifications per entry.
   - TD Simple PSCW—Includes N-terminal acetylation and initial methionine cleavage. This database contains up to three modifications per entry.
   - Bottom Up PSCW—Includes trypsin digestion using two missed cleavages.
   - UniProt XML File—Recreates a PSCW database in the Database Manager.

   The website automatically downloads the database once you select it. The downloaded database (a PSCW or an XML file) appears in the Downloads folder in the following directory:

   C:\Users > *your_name_folder* > Downloads
5. Choose **Databases > Database Manager**, or click the **View Database Info** icon,  , to open the Database Manager window.
6. Click 

in the toolbar of the Database Manager window.
7. Browse to the Downloads folder at the location just given.
8. Copy the downloaded database to the database folder.
9. In the Database Manager, click **Refresh Lists**.

The downloaded database now appears in the list of searchable databases.



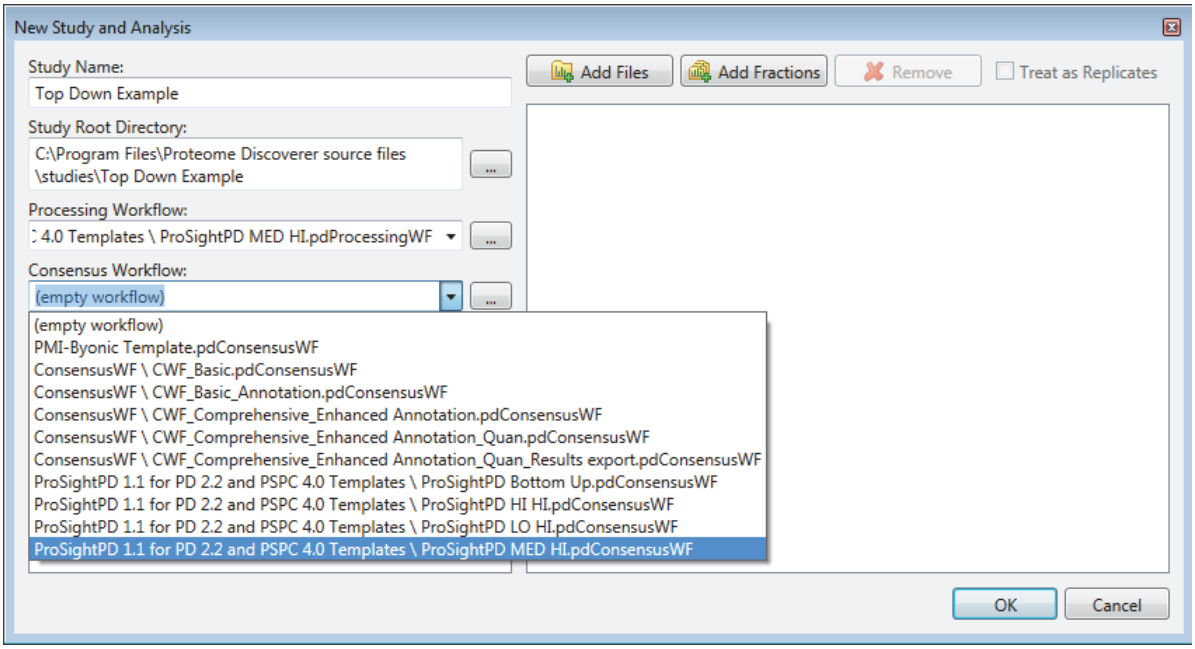
Creating a Study

Start your data analysis by creating a ProSightPD study in the Proteome Discoverer application. For more information on Proteome Discoverer studies, refer to the *Proteome Discoverer User Guide*.

❖ To create a study

1. Start the Proteome Discoverer application.
2. Do one of the following:
   - On the Start Page, click **New Study/Analysis**.
   - Choose **File > New Study/Analysis**.
   - Click the **Create New Study/Analysis** icon,  .

3. In the New Study and Analysis dialog box, do the following:
   a. In the Study Name box type the name of the new study.
   b. In the Study Root Directory box, browse to or type the location of the study root directory.
   c. In the Processing Workflow list, select the processing template to use (see the next topic).
   d. In the Consensus Workflow list, select the consensus template to use (see the next topic).
   e. Click **OK**.



Default ProSightPD Templates

You can find the default ProSightPD templates in Libraries > Documents > Thermo > Proteome Discoverer 2.2 > Common Templates > ProSightPD 1.1 for PD 2.2 and PSPC 4.0 Templates.

These are the default ProSightPD processing workflows:

- ProSighPD Bottom Up Absolute Mass.pdProcessingWF—Performs a peptide analysis.
- ProSightPD HI HI Absolute Mass.pdProcessingWF—Performs a basic top-down proteomic search, using a narrow precursor and product ion mass tolerance.
- ProSightPD HI HI Two Tier.pdProcessingWF—Performs a basic top-down proteomic search, followed by a biomarker search using a narrow precursor and product ion mass tolerance.
- ProSight Hi HI Three Tier Search.pdProcessingWF—Performs a basic two-tier proteomic search, followed by an error-tolerant search as a third step.
- ProSightPD HI HI GRAM.pdProcessingWF—Performs a sequence tag search, followed by a gene-restricted absolute mass search of only those proteins found in the sequence tag search.
- ProSightPD HI HI GRBM.pdProcessingWF—Performs a sequence tag search, followed by a gene-restricted biomarker search of only those proteins found in the sequence tag search.

(For the high-mass version of ProSightPD 1.1 only) These are the available processing workflows:

- ProSightPD LO HI.pdProcessingWF—Performs a search similar to the absolute mass search performed by the ProSightPD HI HI Absolute Mass.pdProcessingWF workflow but uses ReSpect deconvolution for precursor mass determination. It is intended for ion trap MS precursor detection. Because MS/MS is still a high-resolution type of mass spectrometry, a narrow mass tolerance is used for the fragment ions.
- ProSight MED HI.pdProcessingWF—Performs a search similar to absolute mass search performed by the ProSightPD HI HI Absolute Mass.pdProcessingWF workflow but uses ReSpect deconvolution for precursor mass determination. It is intended for low-resolution (for example, 17.5k) Orbitrap™ MS precursor detection. Because MS/MS is still a high-resolution type of mass spectrometry, a narrow mass tolerance is used for the fragment ions.

These are the default ProSightPD consensus workflows:

- ProSightPD Bottom Up.pdConsensusWF—Intended for use with the ProSightPD Bottom Up Absolute Mass.pdProcessingWF processing workflow.
- ProSightPD HI HI.pdConsensusWF—Intended for use with the ProSightPD HI HI Two Tier.pdProcessingWF and ProSight Hi HI Three Tier Search.pdProcessingWF processing workflows.
- ProSightPD LO HI.pdConsensusWF—Intended for use with the ProSightPD LO HI.pdProcessingWF (high-mass version of ProSightPD only) processing workflow.
- ProSightPD MED Hi.pdConsensusWF—Intended for use with the ProSight MED HI.pdProcessingWF (high-mass version of ProSightPD only) processing workflow.

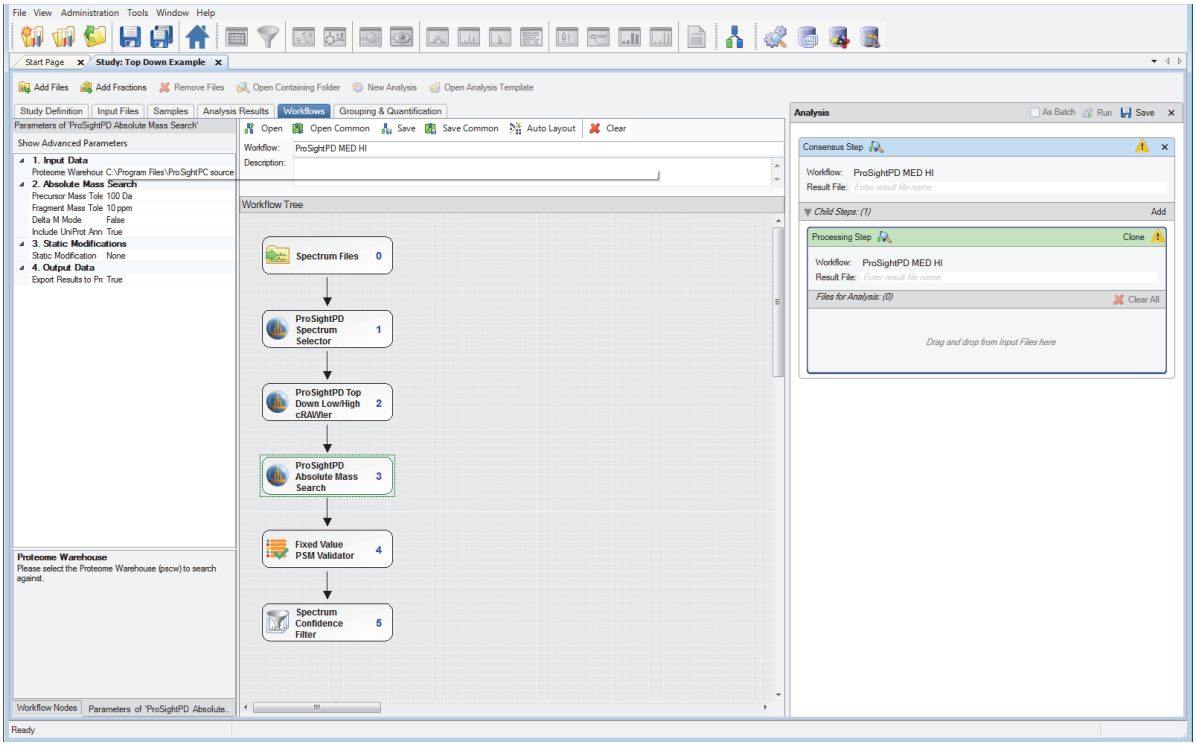
Setting Up a ProSightPD Database Search

If you want to make any changes to the templates to set up a search, follow these instructions. For example, you might want to choose different search parameters.

❖ To set up a ProSightPD database search

1. In the study, click the **Workflows** tab.
2. Click the **Show Workflow** icon,  , on the Processing Step toolbar in the Analysis pane for a processing workflow.

   The processing workflow appears in the Workflow Tree pane.
3. In the workflow, click a search node, for example, the **ProSightPD Absolute Mass Search** node.
4. Select a different setting for a parameter of this node. For example, you might want to select a different database (PSCW file) from the list of databases given by the Proteome Warehouse parameter.
5. For any other search nodes in the workflow, follow step 3 and step 4 for each one.



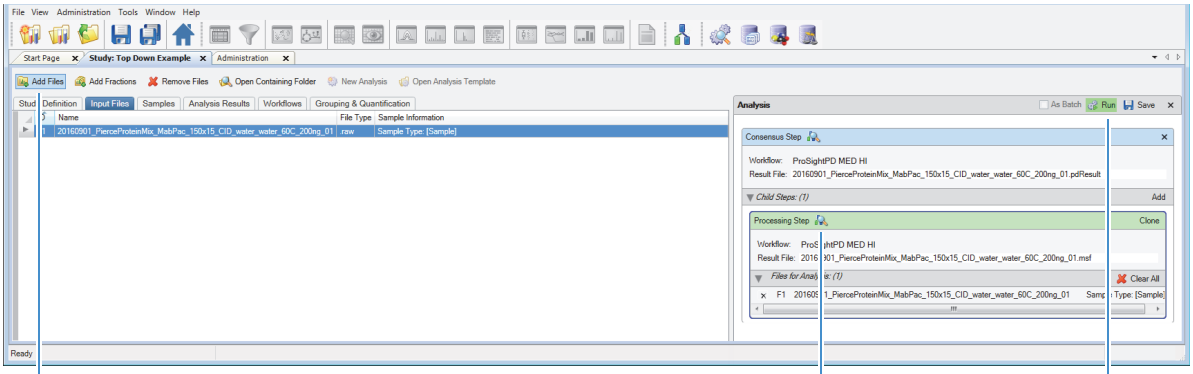
Adding Files to the Study and Selecting Files

If a raw data file is not defined in the creation of a study, follow these steps.

❖ To add files to the study or select files for the search

1. On the Study *Study_name* page of the Proteome Discoverer application, click Add Files.
2. In the Add Files dialog box, select the file and click **Open**.
3. Click the **Input Files** tab.

4. Drag the input file or files to the Processing Step area of the Analysis pane.



Performing the Search

❖ To perform a search

In the Analysis pane, click Run.

Reviewing ProSightPD Search Results

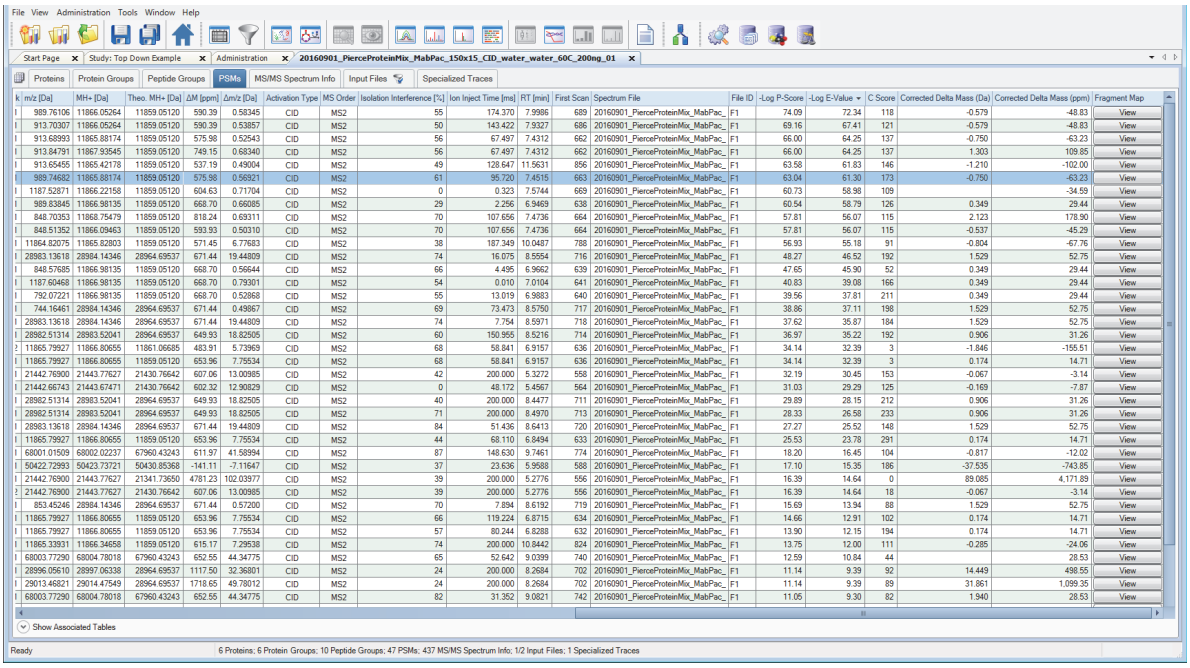
❖ To open the .pdResults report

- In the job queue on the Administration page, select the consensus file, and click Open Results.

–or–

- In the study, click the **Analysis Results** tab, and double-click the name of the .pdResults file.

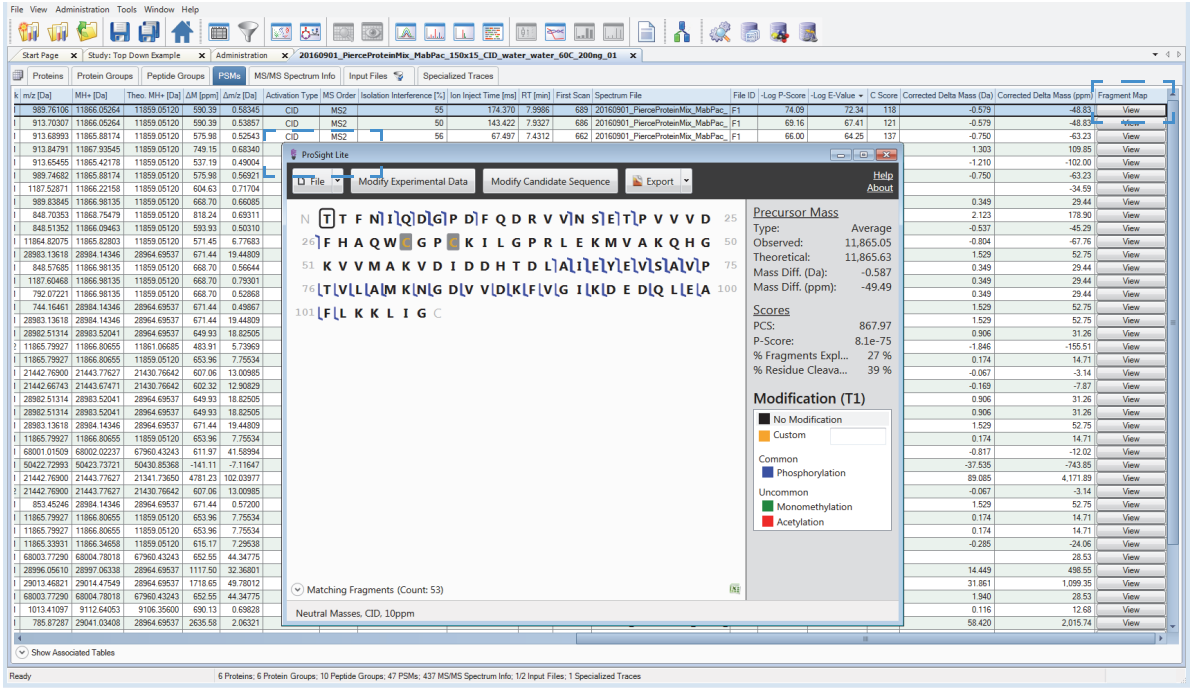
The .pdResults file opens.



The *Proteome Discoverer User Guide* explains in detail how to review search results. Each proteoform identified by the ProSightPD application is stored in the Proteome Discoverer application as a PSM. The results of a ProSightPD search include the following features on the PSMs page:

- Three scoring columns (-Log P-Score, -Log E-Value, and C Score)
- Two Corrected Delta Mass columns (Da and ppm)
- One Fragment Map column

By clicking View in the Fragment Map column, you can view the fragment map associated with the hit in ProSight Lite (see the next figure). If ProSight Lite is not already installed, download and install it from the prosightlite.northwestern.edu website.



References

The following references are useful in understanding the searches in the ProSightPD application. You can direct additional questions to info@proteinaceous.net.

- LeDuc, Richard D.; Kelleher, Neil L. Using ProSight PTM and Related Tools for Targeted Protein Identification and Characterization with High Mass Accuracy Tandem MS Data Current Protocols in Bioinformatics. *Unit 13.6.* **2007**.
- Zamdborg, Leonid; LeDuc, Richard D.; Glowacz, Kevin J.; Kim, Yong-Bin; Viswanathan, Vinayak; Spaulding, Ian T.; Early, Bryan P., Bluhm, Eric J.; Babai, Shannee; Kelleher, Neil L. ProSight PTM 2.0: Improved Protein Identification and Characterization for Top-Down Proteomics. *Nucleic Acids Research 35 (suppl_2),* (Web Server issue). **2007**. W701-6: PMCID: PMC1933126.
- LeDuc, Richard D.; Fellers, Ryan T.; Early, Bryan P.; Greer, Joseph B.; Thomas, Paul M.; Kelleher, Neil L. The C-Score: A Bayesian Framework to Sharply Improve Proteoform Scoring in High-Throughput Top-Down Proteomics. *Journal of Proteome Research*, **2014**, 13 (*7*), 3231-3240. PMCID: PMC4084843.

Trademarks