

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

- [System Requirements](#)
- [Installing the ProSightPD Nodes](#)
- [Searching with ProSightPD](#)
- [Performing the Search](#)
- [Reviewing ProSightPD Search Results](#)
- [References](#)
- [Trademarks](#)

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.

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

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)	<ul style="list-style-type: none">• 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)	<ul style="list-style-type: none">• 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

System Requirements

Hardware Requirements

Software Requirements

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

System	Requirements
Software	<ul style="list-style-type: none">• 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@proteinaeous.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 Proteinaeous 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 proteinaeous.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 **Archaeobacteria**, **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:

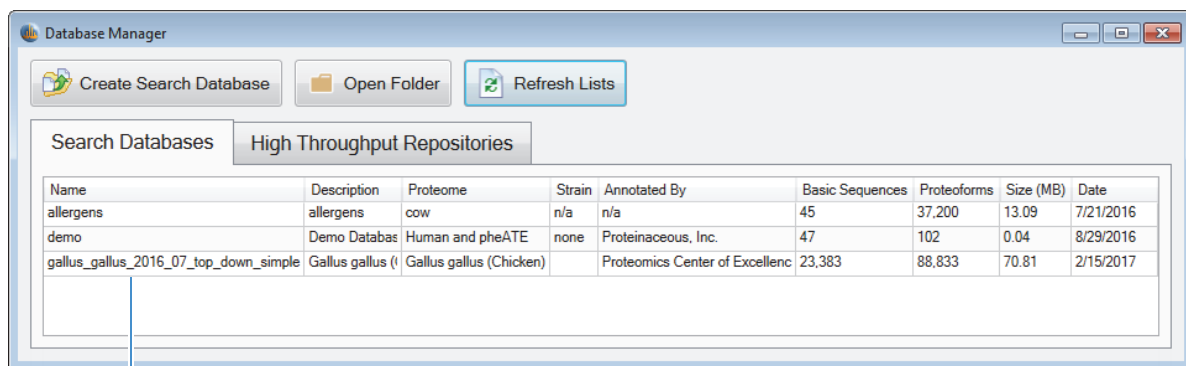
- TD Complex PSCW—Includes N-terminal acetylation and initial methionine cleavage. This database contains 12 to 15 modifications per entry.
- TD Simple PSCW—Includes N-terminal acetylation and initial methionine cleavage. This database contains up to three modifications per entry.
- Bottom Up PSCW—Includes 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  **Open Folder** 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.




Downloaded database

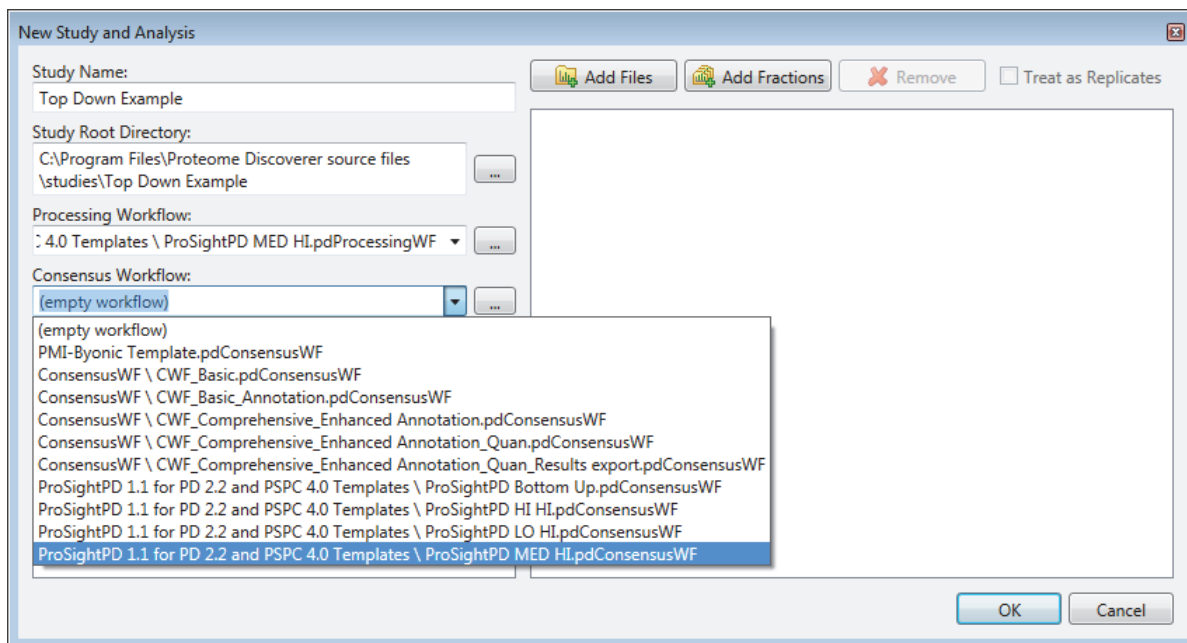
Creating a Study

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 PSC 4.0 Templates.

These are the default ProSightPD processing workflows:

- ProSightPD 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.
- ProSightPD 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 ReSpec 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.
- ProSightPD MED HI.pdProcessingWF—Performs a search similar to absolute mass search performed by the ProSightPD HI HI Absolute Mass.pdProcessingWF workflow but uses ReSpec 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.


Setting Up a ProSightPD Database Search

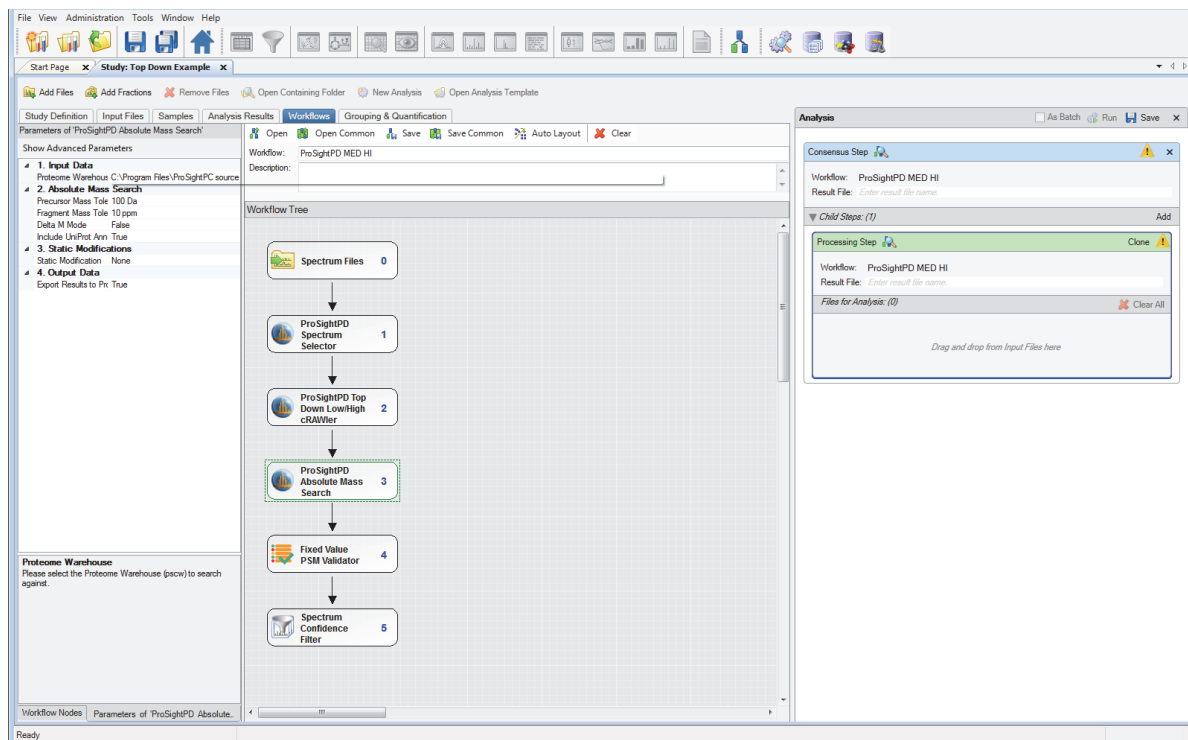
These are the default ProSightPD consensus workflows:

- 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.

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.



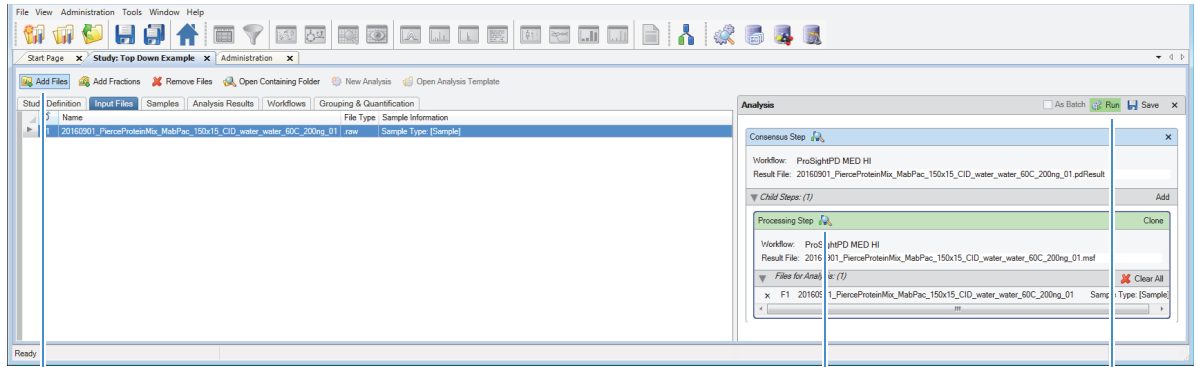
Adding Files to the Study and Selecting Files

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.



Add Files icon

Show Workflow icon

Run icon


Performing the Search

Reviewing ProSightPD Search Results

❖ To perform a search

In the Analysis pane, click .

❖ To open the .pdResults report

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

–or–

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

The .pdResults file opens.

k	m/z [Da]	MH+	[M+H] ⁺	Theor. MH+	[M+H] ⁺	DM (ppm)	[M+H] ⁺	[M+H] ⁺	Activation Type	MS Order	Isolation Interference (%)	Ion Inject Time (min)	RT (min)	First Scan	Spectrum File	File ID	-Log P-Score	-Log E-Value	C Score	Corrected Delta Mass (Da)	Corrected Delta Mass (ppm)	Fragment Map
1	380.76106	11866.05264	11869.05120	690.39	0.5345	CID	MS2	55	174.370	7.9995	639	20160901_PierceProteinMix_MabPac	F1	74.09	72.34	118	0.579				-48.83	View
1	913.70307	11866.05264	11869.05120	590.39	0.53857	CID	MS2	50	143.422	7.9327	696	20160901_PierceProteinMix_MabPac	F1	69.16	67.41	121	0.579				-48.83	View
1	913.68993	11866.05264	11869.05120	575.98	0.52543	CID	MS2	56	67.497	7.4312	662	20160901_PierceProteinMix_MabPac	F1	66.00	64.25	137	0.750				-63.23	View
1	913.84791	11867.93545	11869.05120	749.15	0.68340	CID	MS2	56	67.497	7.4312	662	20160901_PierceProteinMix_MabPac	F1	66.00	64.25	137	1.303				-109.85	View
1	913.65495	11865.42178	11869.05120	537.19	0.49004	CID	MS2	49	128.647	11.5631	856	20160901_PierceProteinMix_MabPac	F1	63.58	61.83	146	-1.210				-102.00	View
1	889.49662	11866.05264	11869.05120	676.96	0.56957	CID	MS2	61	95.720	7.4515	653	20160901_PierceProteinMix_MabPac	F1	63.04	61.30	175	-0.750				-63.23	View
1	1187.82871	11866.22158	11869.05120	694.63	0.71704	CID	MS2	0	0.323	7.5744	689	20160901_PierceProteinMix_MabPac	F1	60.73	58.98	199					-34.59	View
1	989.83845	11866.98135	11869.05120	668.70	0.66095	CID	MS2	29	2.256	6.9469	638	20160901_PierceProteinMix_MabPac	F1	60.54	58.79	126	0.349				29.44	View
1	848.70353	11868.75479	11869.05120	818.24	0.69311	CID	MS2	70	107.656	7.4736	664	20160901_PierceProteinMix_MabPac	F1	57.81	56.07	115	2.123				29.44	View
1	848.51352	11866.09463	11869.05120	593.93	0.50310	CID	MS2	70	107.656	7.4736	664	20160901_PierceProteinMix_MabPac	F1	57.81	56.07	115	-0.537				-45.29	View
1	11864.40275	11865.82903	11869.05120	571.45	6.77683	CID	MS2	38	187.349	10.0487	788	20160901_PierceProteinMix_MabPac	F1	56.93	55.18	91	-0.804				-67.76	View
1	2893.13618	28984.14346	28964.69537	671.44	19.44809	CID	MS2	74	16.075	8.9554	716	20160901_PierceProteinMix_MabPac	F1	48.27	46.52	192	1.529				52.75	View
1	848.57855	11866.98135	11869.05120	668.70	0.56644	CID	MS2	66	4.495	6.9662	639	20160901_PierceProteinMix_MabPac	F1	47.65	45.90	52	0.349				29.44	View
1	1187.60468	11866.98135	11869.05120	668.70	0.79301	CID	MS2	64	0.010	7.0104	641	20160901_PierceProteinMix_MabPac	F1	40.83	39.08	166	0.349				29.44	View
1	792.07221	11866.98135	11869.05120	668.70	0.52688	CID	MS2	65	13.019	6.9883	640	20160901_PierceProteinMix_MabPac	F1	39.56	37.81	211	0.349				29.44	View
1	744.16451	28984.14346	28964.69537	671.44	0.48867	CID	MS2	69	73.473	8.5750	717	20160901_PierceProteinMix_MabPac	F1	38.86	37.11	198	1.529				52.75	View
1	2893.13618	28984.14346	28964.69537	671.44	19.44809	CID	MS2	74	7.754	8.5971	718	20160901_PierceProteinMix_MabPac	F1	37.62	35.87	194	1.529				52.75	View
1	2892.91314	28983.52041	28964.69537	649.93	18.82956	CID	MS2	60	150.955	8.5216	714	20160901_PierceProteinMix_MabPac	F1	36.97	35.22	192	0.906				31.26	View
1	11865.79927	11866.80655	11869.05120	653.96	7.75534	CID	MS2	68	58.841	6.9157	636	20160901_PierceProteinMix_MabPac	F1	34.14	32.39	3	-1.846				-155.51	View
1	11865.79927	11866.80655	11869.05120	653.96	7.75534	CID	MS2	68	58.841	6.9157	636	20160901_PierceProteinMix_MabPac	F1	34.14	32.39	3	0.174				-14.71	View
1	21442.76900	21443.79227	21430.76642	607.06	13.00985	CID	MS2	42	200.000	5.3272	558	20160901_PierceProteinMix_MabPac	F1	32.19	30.45	153	-0.067				-3.14	View
1	21442.66743	21443.67471	21430.76642	602.32	12.90829	CID	MS2	0	48.172	5.4567	564	20160901_PierceProteinMix_MabPac	F1	31.03	29.29	125	-0.169				-7.87	View
1	2892.91314	28983.52041	28964.69537	649.93	18.82956	CID	MS2	40	200.000	8.4477	711	20160901_PierceProteinMix_MabPac	F1	29.89	28.15	212	0.906				31.26	View
1	2892.91314	28983.52041	28964.69537	649.93	18.82956	CID	MS2	71	200.000	8.4970	713	20160901_PierceProteinMix_MabPac	F1	28.33	26.58	233	0.906				31.26	View
1	2893.13618	28984.14346	28964.69537	671.44	19.44809	CID	MS2	84	51.436	8.6413	720	20160901_PierceProteinMix_MabPac	F1	27.27	25.52	148	1.529				52.75	View
1	11865.79927	11866.80655	11869.05120	653.96	7.75534	CID	MS2	44	68.110	6.8494	633	20160901_PierceProteinMix_MabPac	F1	25.53	23.78	291	0.174				14.71	View
1	6801.01509	68002.02237	67960.43243	611.97	41.58994	CID	MS2	67	148.630	9.7461	774	20160901_PierceProteinMix_MabPac	F1	18.20	16.45	104	-0.817				-12.02	View
1	50422.72993	50423.73721	50430.85386	-141.11	-7.19447	CID	MS2	37	23.636	5.9588	568	20160901_PierceProteinMix_MabPac	F1	17.10	15.35	186	-37.535				-743.85	View
1	21442.76900	21443.79227	21431.76500	4781.23	100.03977	CID	MS2	39	200.000	5.2776	556	20160901_PierceProteinMix_MabPac	F1	16.39	14.64	0	89.603				4,171.89	View
1	21442.76900	21443.79227	21430.76642	607.06	13.00985	CID	MS2	39	200.000	5.2776	556	20160901_PierceProteinMix_MabPac	F1	16.39	14.64	18	-0.067				-3.14	View
1	853.45246	28984.14346	28964.69537	671.44	0.57200	CID	MS2	70	7.894	8.6192	719	20160901_PierceProteinMix_MabPac	F1	15.69	13.94	88	1.529				52.75	View
1	11865.79927	11866.80655	11869.05120	653.96	7.75534	CID	MS2	66	119.224	6.8715	634	20160901_PierceProteinMix_MabPac	F1	14.66	12.91	102	0.174				14.71	View
1	11865.79927	11866.80655	11869.05120	653.96	7.75534	CID	MS2	67	80.244	6.8288	632	20160901_PierceProteinMix_MabPac	F1	13.90	12.15	194	0.174				14.71	View
1	11865.39311	11866.34658	11869.05120	615.17	7.29538	CID	MS2	74	200.000	10.3423	824	20160901_PierceProteinMix_MabPac	F1	13.75	12.00	111	-0.285				-24.06	View
1	68003.72290	68004.78018	67960.43243	652.65	44.34775	CID	MS2	65	52.642	9.0399	740	20160901_PierceProteinMix_MabPac	F1	12.59	10.84	44					28.53	View
1	28996.05610	28997.06338	28964.69537	1117.50	32.36801	CID	MS2	24	200.000	8.2684	702	20160901_PierceProteinMix_MabPac	F1	11.14	9.39	92	14.449				498.55	View
1	29013.46821	29014.47549	28964.69537	1718.65	49.78012	CID	MS2	24	200.000	8.2684	702	20160901_PierceProteinMix_MabPac	F1	11.14	9.39	89	31.861				1,099.53	View
1	68003.72290	68004.78018	67960.43243	652.65	44.34775	CID	MS2	82	31.352	9.0821	742	20160901_PierceProteinMix_MabPac	F1	11.05	9.30	82	1.940				28.53	View

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.

The screenshot displays the ProSight Lite software interface. The main window shows a list of protein identifications with columns for k m/z [Da], M+ [Da], Theo. M+ [Da], ΔM [ppm], Am/z [Da], Activation Type, MS Order, Isolation Interference [%], Ion Inject Time [ms], RT [min], First Scan, Spectrum File, File ID, Log P-Score, Log E-Value, C-Score, Corrected Delta Mass (Da), Corrected Delta Mass (ppm), and Fragment Map. A pop-up window titled 'ProSight Lite' is open, showing details for a specific hit. The 'Precursor Mass' section includes Type (Average), Observed (11,865.05), Theoretical (11,865.63), Mass Diff. (Da) (-0.587), and Mass Diff. (ppm) (-49.49). The 'Scores' section lists PCS (867.97), P-Score (8.1e-75), % Fragments Expl. (27%), and % Residue Cleava... (39%). The 'Modification (T1)' section shows a legend for No Modification, Custom, Phosphorylation, Uncommon, Monomethylation, and Acetylation. The protein sequence is shown with modifications highlighted in yellow and red.

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: PMID: 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. PMID: PMC4084843.

Trademarks

Proteome Discoverer is a trademark of Thermo Fisher Scientific Inc. in the United States.

ProSightPC and ProSightPD are registered trademarks of Proteinaceous, Inc. in the United States. UniProt is a registered trademark of the European Molecular Biology Laboratory Incorporated Association.

The following are registered trademarks in the United States and other countries: Microsoft and Windows are registered trademarks of Microsoft Corporation. Intel and Xeon are registered trademarks of Intel Corporation.

All other trademarks are the property of Thermo Fisher Scientific Inc. and its subsidiaries.