Intact Protein Demo Workflows

Biologics Explorer Software Guidelines

Powered by Genedata Expressionist®

SCIEX The power of precision

Genedata EXPRESSIONIST

TPUB-CUST-16031-A

Intact Protein Demo Workflows

CONTENTS OF THIS GUIDE

Part A

Overview of the Intact Protein Demo Workflows

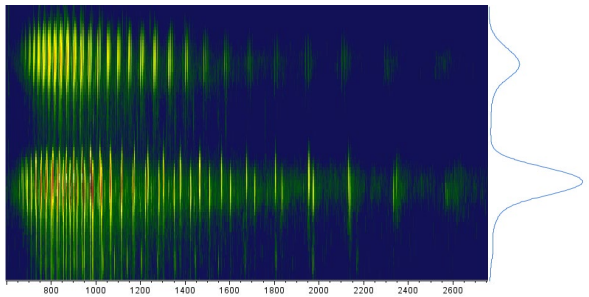
A

Overview of the Applications for Intact Protein Demo Workflows

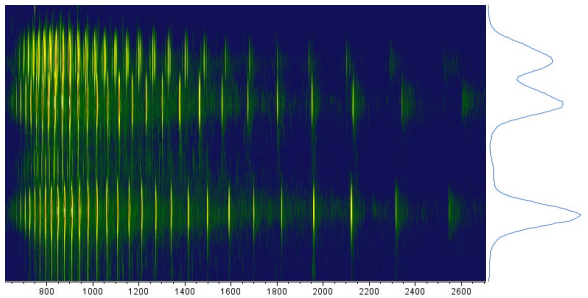
- The Intact Protein demo workflows contain examples of how to analyze intact, reduced, or partially cleaved biotherapeutics.
- **Spectral Deconvolution** is used to analyze species of interest that are chromatographically well resolved.
  - The RT ranges for deconvolution are selected in the TIC or UV data, or can be specified manually.
- **Time-Resolved (TR) Deconvolution** is used to analyze data that contains complex mixtures that are poorly resolved.
  - Each RT scan is deconvoluted to create an ion map of the deconvoluted data.

TR Deconvolution:

Complex Chromatography



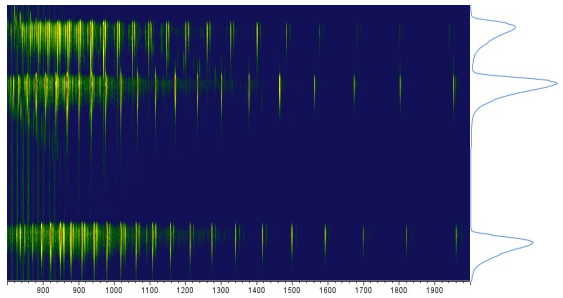
Fragments: unresolved



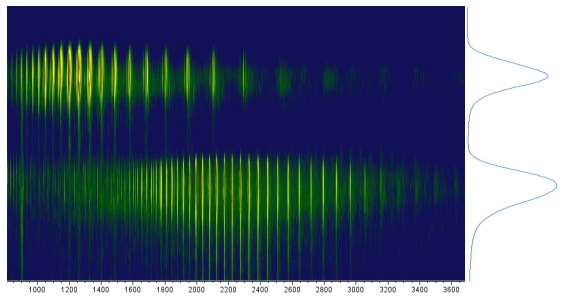
Fragments: overlapped

Spectral Deconvolution:

Simple Chromatography



Fragments: resolved



Subunits: resolved

Overview of the Intact Protein Demo Workflows

Intact_SpectralDeconvolution_MS_Demo:

- An Intact Protein analysis workflow with spectral deconvolution of each RT range that was identified in the TIC.

Intact_SpectralDeconvolution_MS+UV_Demo:

- An Intact Protein analysis workflow with spectral deconvolution of each RT range that was identified in the UV chromatogram.

Intact_TRDeconvolution_MS_Demo:

- An Intact Protein analysis workflow with time-resolved (TR) deconvolution.

Intact_TRDeconvolution_MS+UV_Demo:

- An Intact Protein analysis workflow with time-resolved (TR) deconvolution and UV peak detection.

Pepmap_ReviewSnapshots_Demo:

- A workflow to open or review saved results.

Intact_BatchProcessing_Demo:

- A version of the Intact_SpectralDeconvolution_MS+UV workflow that analyzes multiple data files, one sample at a time.

Part B

Information About Intact Protein Demo Workflows

B

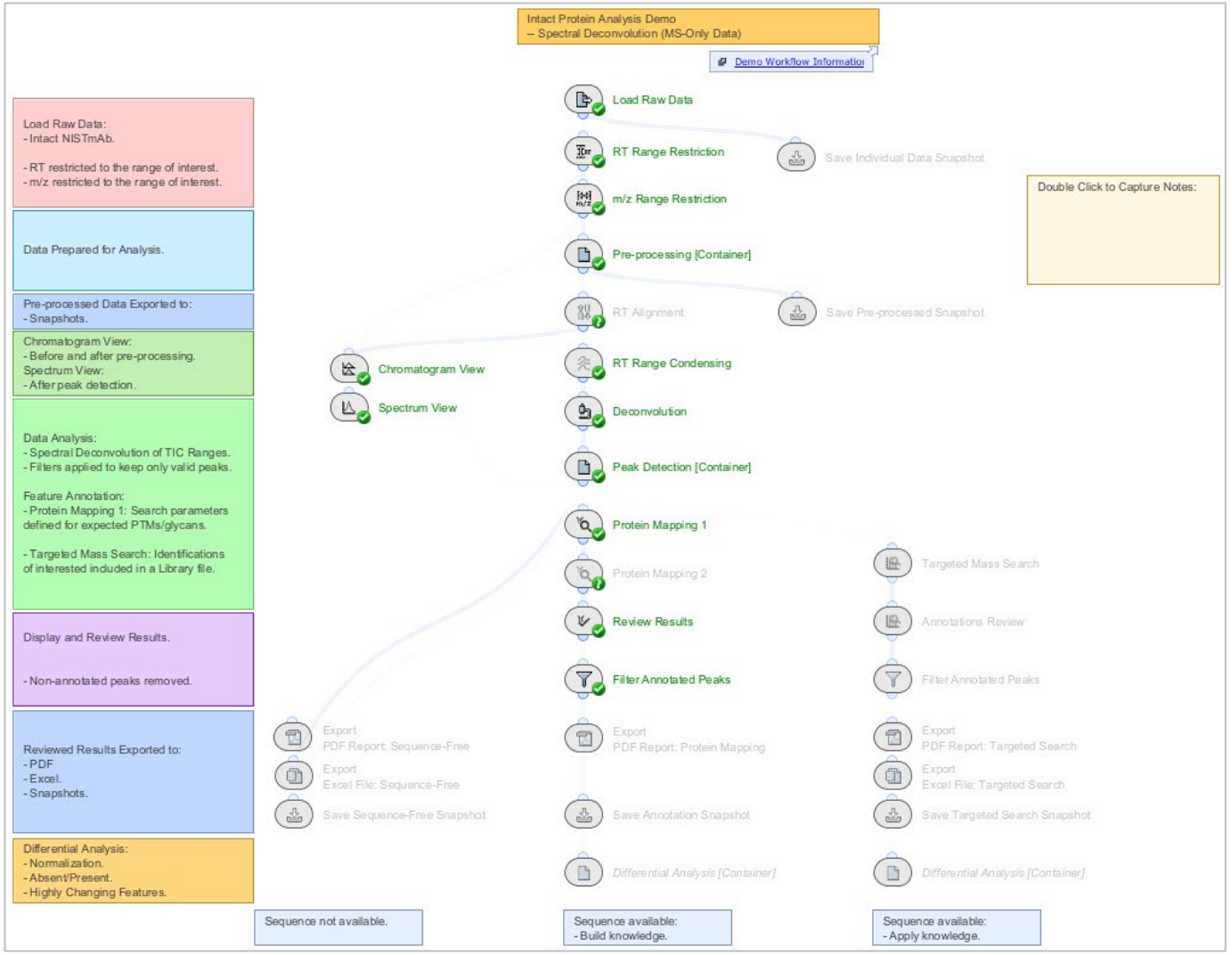
Spectral Deconvolution with MS Data

Demo Workflow Information

B1

Overview and Application: Intact_SpectralDeconvolution_MS_Demo

- This workflow uses data from an intact biotherapeutic molecule.
- This workflow uses the Total Ion Chromatogram (TIC) to select the RT ranges for deconvolution.
- The search parameters in the *Protein Mapping* activity node are optimized to identify the intact (**Fully Connected**) molecule, as well as all other possible combinations with **2 Additional Chains**.
- For more information about how to use Biologics Explorer software, refer to the *Biologics Explorer Quick Guide* and *Intact Protein Template Workflow Guidelines*.



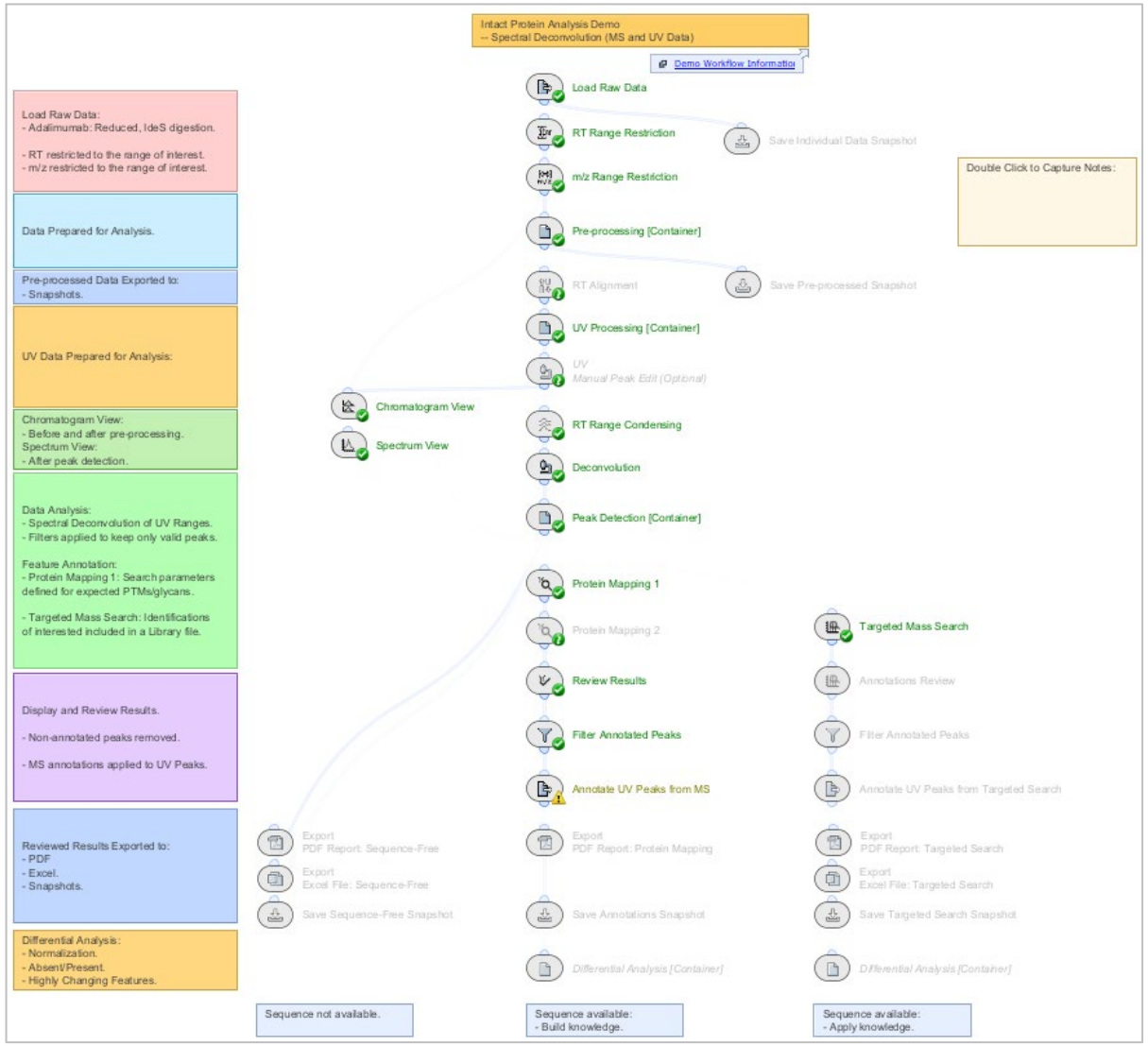
Spectral Deconvolution with MS+UV Data

Demo Workflow Information

B2

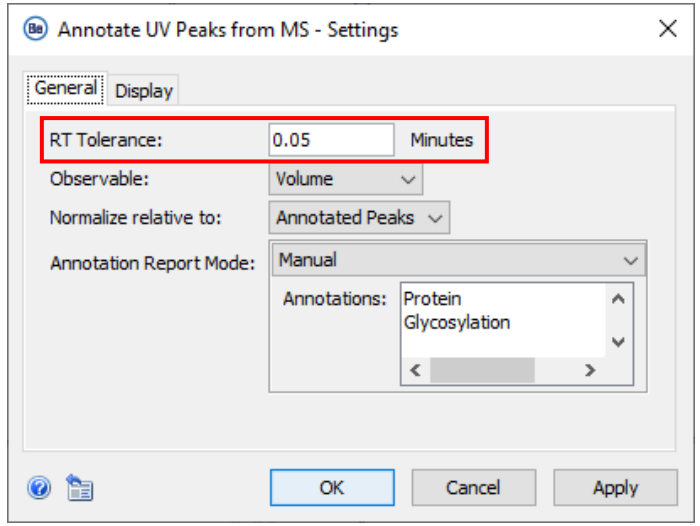
Overview and Application: Intact_SpectralDeconvolution_MS+UV_Demo

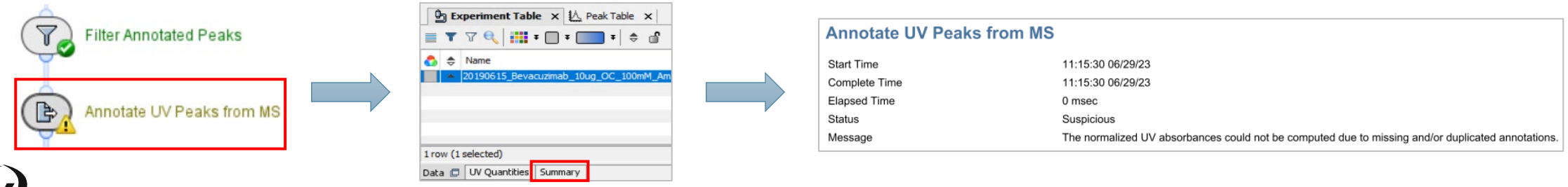
- This workflow uses data from a reduced biotherapeutic molecule after IdeS digestion.
- This workflow uses the UV trace to select the RT ranges for deconvolution.
- The search parameters in the *Protein Mapping* activity node are optimized to identify the **Fully Reduced** subunits (LC, Fd', Fc/2).
- For more information about how to use Biologics Explorer software, refer to the *Biologics Explorer Quick Guide* and *Intact Protein Template Workflow Guidelines*.



Annotate UV Peaks from MS

- This activity node uses MS peak information to annotate the peaks in the **UV Chromatogram**.
  - A related peak must elute in the specified **RT Tolerance**.
- The UV absorbance is used to calculate the relative ratio of the UV peaks.
  - A protein sequence is required to normalize UV absorbance values.
  - If the UV normalization factor cannot be calculated, then the activity node shows a **yellow warning**.
    - UV normalization is not available with *TR Deconvolution*.
    - UV normalization is not applicable for *Targeted Mass Search* because this activity node does not contain the protein sequence.





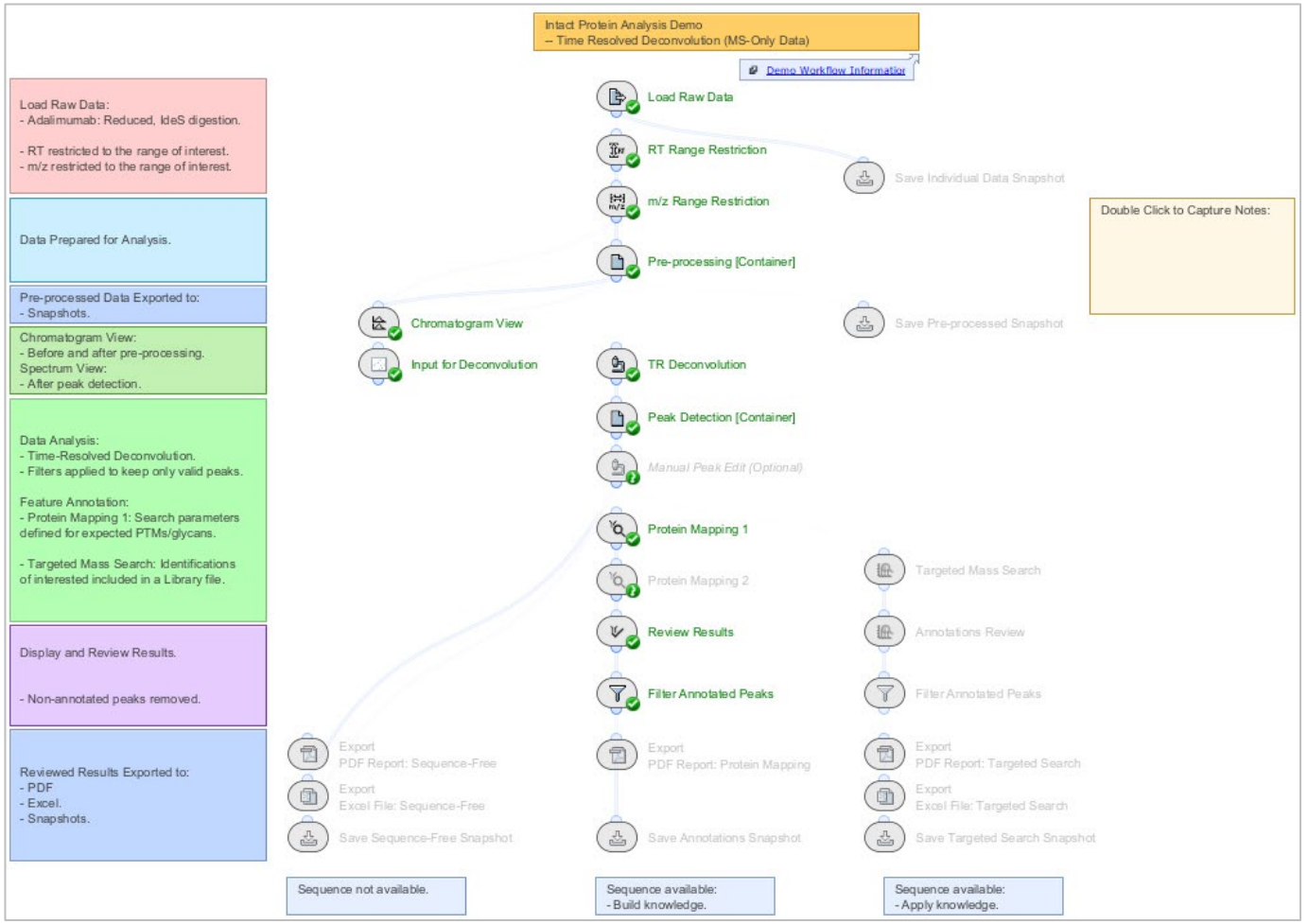
TR Deconvolution with MS-Only Data

Demo Workflow Information

B3

Overview and Application: Intact_TRDeconvolution_MS_Demo

- This workflow uses data from a reduced biotherapeutic molecule after IdeS digestion.
- Each RT scan is deconvoluted create an ion map of the deconvoluted proteins.
- The search parameters in the *Protein Mapping* activity node are optimized to identify the **Fully Reduced** subunits (LC, Fd', Fc/2).
- For more information about how to use Biologics Explorer software, refer to the *Biologics Explorer Quick Guide* and *Intact Protein Template Workflow Guidelines*.



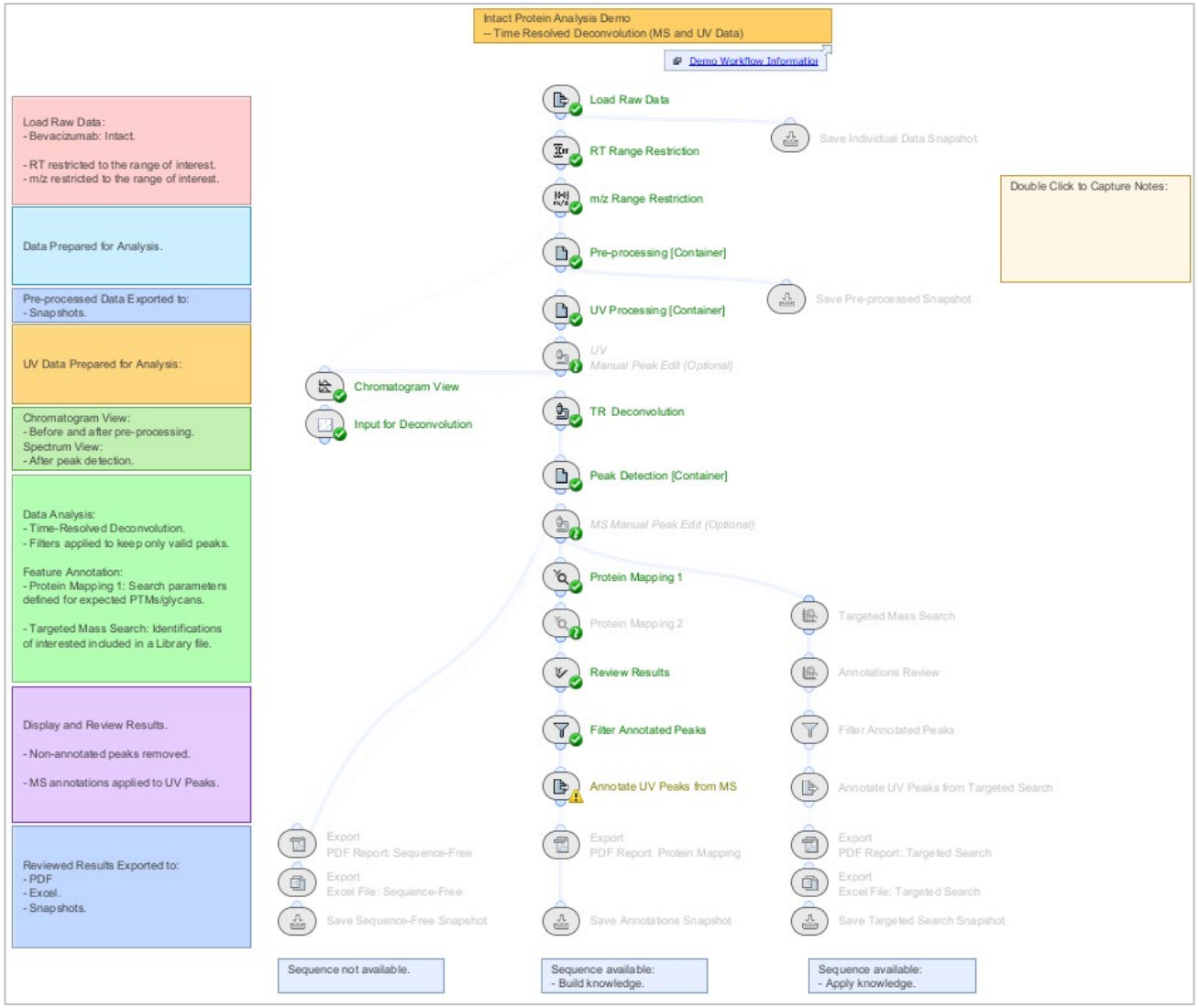
TR Deconvolution with MS+UV Data

Demo Workflow Information

B4

Overview and Application: Intact_TRDeconvolution_MS+UV_Demo

- This workflow uses data from an intact biotherapeutic molecule.
- Each RT scan is deconvoluted create an ion map of the deconvoluted proteins.
- The search parameters in the *Protein Mapping* activity node are optimized to identify the intact (**Fully Connected**) molecule.
- For more information about how to use Biologics Explorer software, refer to the *Biologics Explorer Quick Guide* and *Intact Protein Template Workflow Guidelines*.

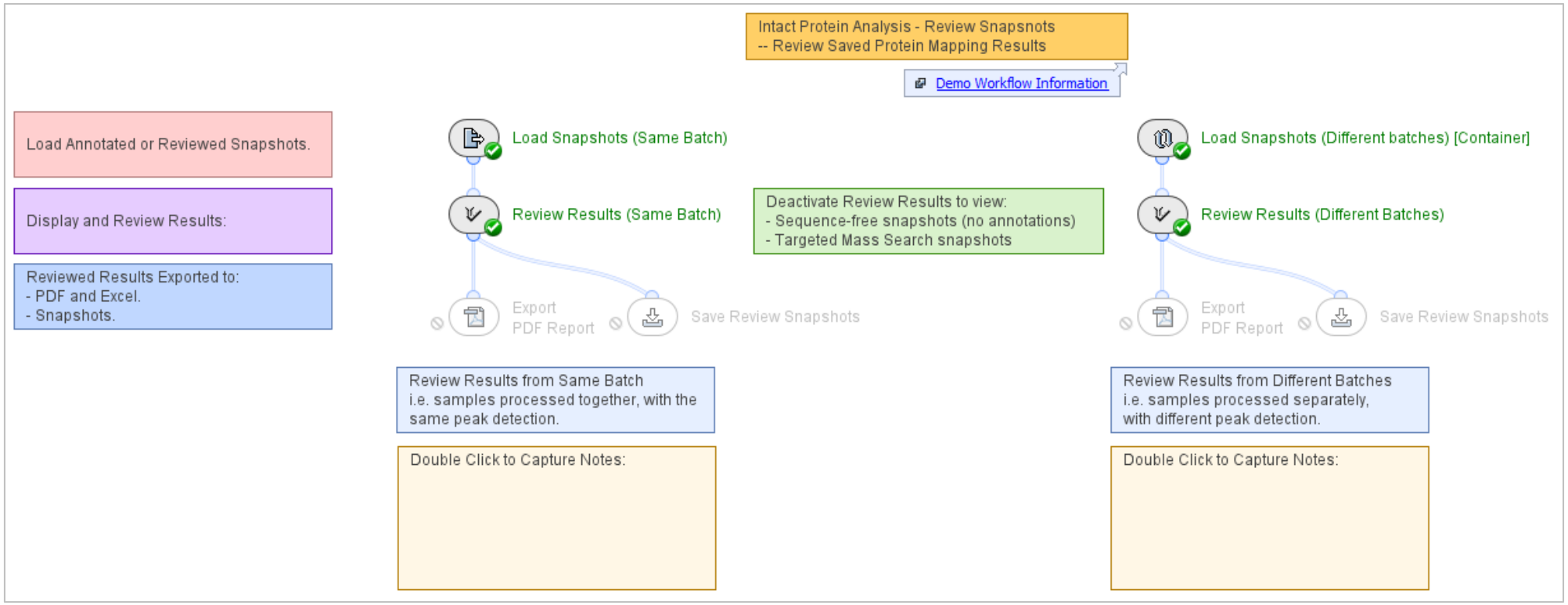


Review Snapshots

Demo Workflow Information

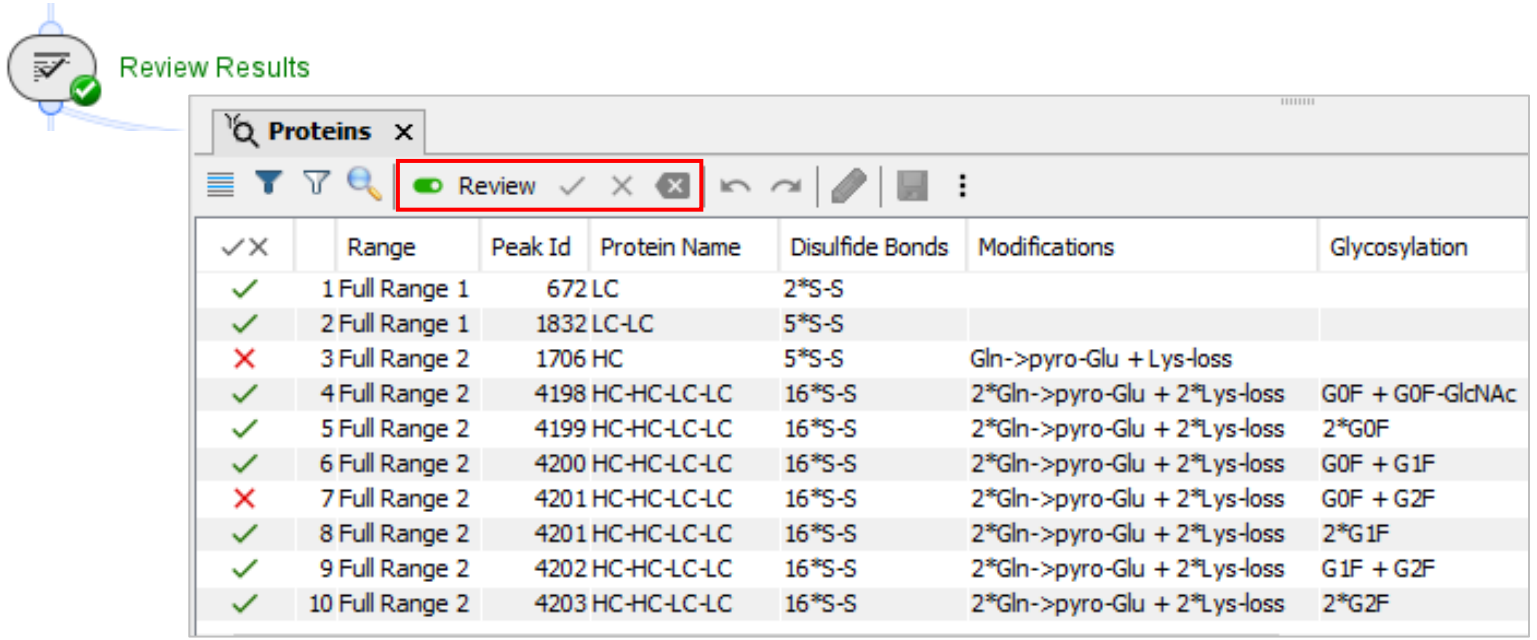
B5

Overview and Application: IntactProtein_ReviewSnapshots_Demo



- This workflow uses saved Snapshots to show results that have protein annotations.
- For more information about how to use Biologics Explorer software, refer to the *Biologics Explorer Quick Guide* and *Intact Protein Template Workflow Guidelines*.

Review Saved Results



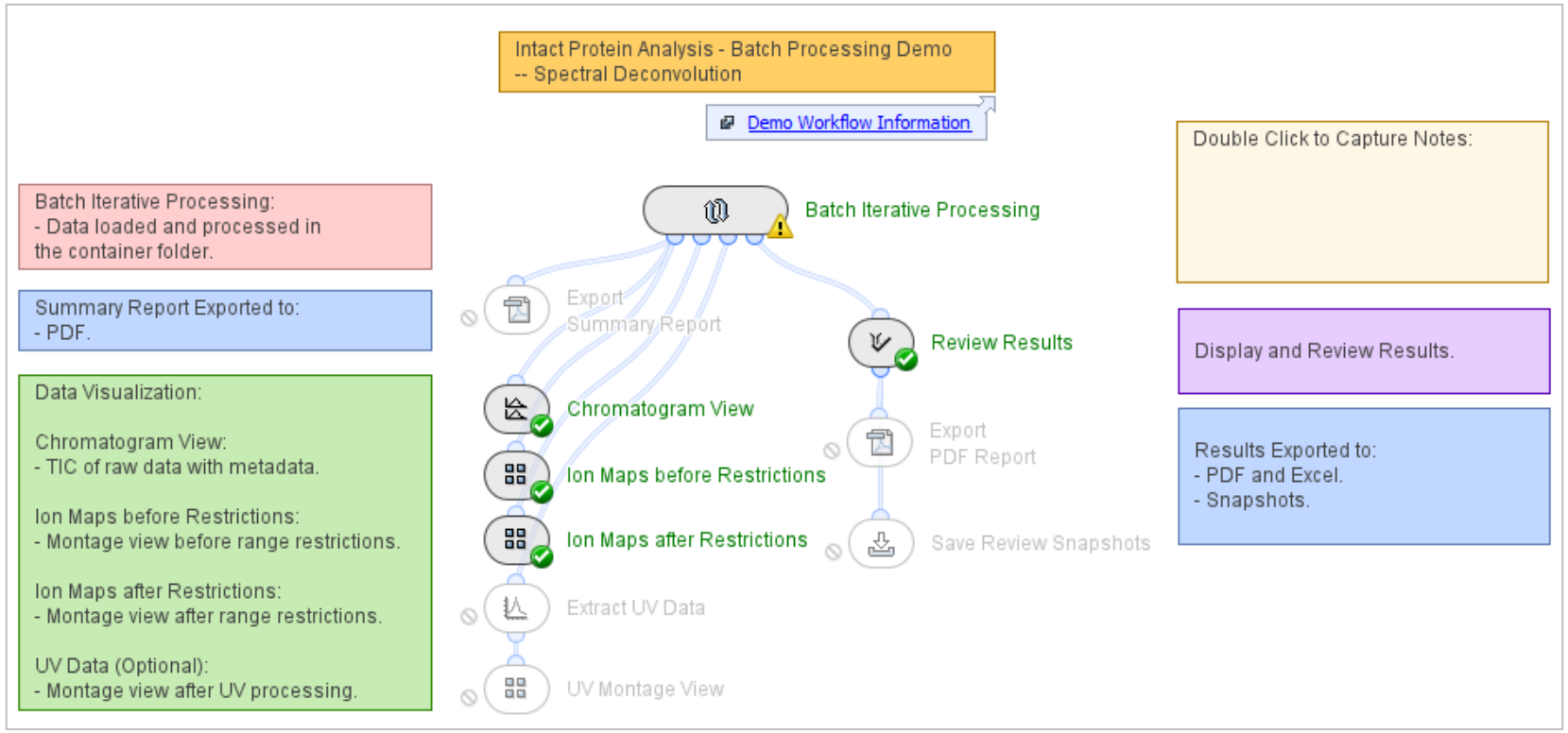
✓✗		Range	Peak Id	Protein Name	Disulfide Bonds	Modifications	Glycosylation
✓	1	Full Range 1	672	LC	2*S-S		
✓	2	Full Range 1	1832	LC-LC	5*S-S		
✗	3	Full Range 2	1706	HC	5*S-S	Gln->pyro-Glu + Lys-loss	
✓	4	Full Range 2	4198	HC-HC-LC-LC	16*S-S	2*Gln->pyro-Glu + 2*Lys-loss	G0F + G0F-GlcNAc
✓	5	Full Range 2	4199	HC-HC-LC-LC	16*S-S	2*Gln->pyro-Glu + 2*Lys-loss	2*G0F
✓	6	Full Range 2	4200	HC-HC-LC-LC	16*S-S	2*Gln->pyro-Glu + 2*Lys-loss	G0F + G1F
✗	7	Full Range 2	4201	HC-HC-LC-LC	16*S-S	2*Gln->pyro-Glu + 2*Lys-loss	G0F + G2F
✓	8	Full Range 2	4201	HC-HC-LC-LC	16*S-S	2*Gln->pyro-Glu + 2*Lys-loss	2*G1F
✓	9	Full Range 2	4202	HC-HC-LC-LC	16*S-S	2*Gln->pyro-Glu + 2*Lys-loss	G1F + G2F
✓	10	Full Range 2	4203	HC-HC-LC-LC	16*S-S	2*Gln->pyro-Glu + 2*Lys-loss	2*G2F

- The *Review Results* activity node opens a copy of the previous analysis.
  - Any previously accepted or rejected proteins have the applicable entry in the **Flags** column.
  - Another stage of review is then possible.
  - The reviewed sbf files and a new report can be saved.

Intact Protein Batch Processing
Demo Workflow Information

B6

Overview and Application: IntactProtein_BatchProcessing_Demo



- This workflow uses sequence information from metadata that was imported into the workflow to analyze multiple samples independently.
- For more information about how to use Biologics Explorer software, refer to the *Biologics Explorer Quick Guide* and *Intact Protein Template Workflow Guidelines*.

How to Use the Batch Processing Workflow

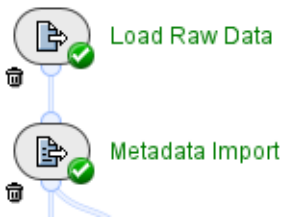
Use this workflow to analyze multiple samples independently.

- The samples do not require consistent chromatography.
- The samples can have different protein sequences.

All samples and their associated metadata are analyzed in the *Batch Iterative Processing* container.

- Intermediate results for each sample are not saved when the *Batch Iterative Processing* container is used for data analysis.
- To optimize workflow parameters, deactivate the **Trash** icon for activity nodes in the *Batch Iterative Processing* container, and use a single representative sample.
- To save memory when large numbers of samples are analyzed, activate the **Trash** icon for activity nodes in the *Batch Iterative Processing* container.

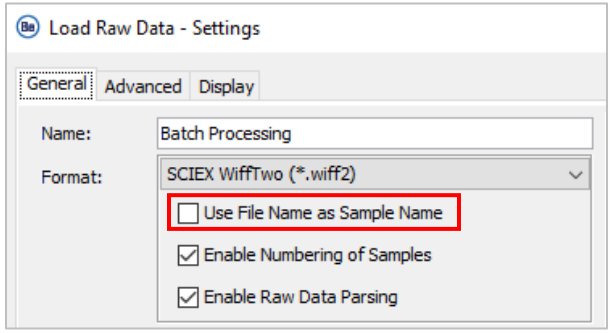
Load Raw Data: Experiment Names and Metadata

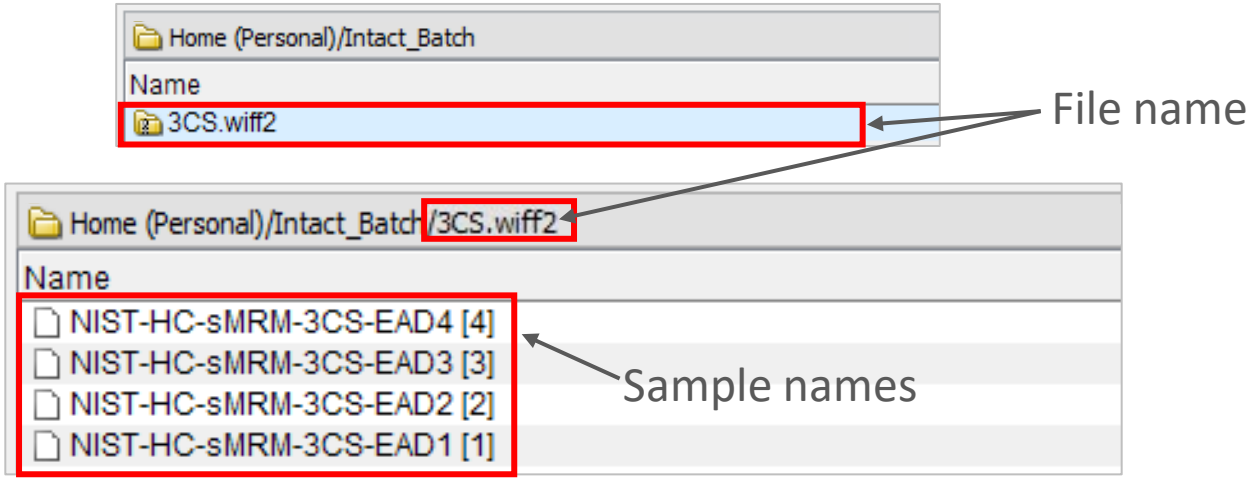


To analyze replicate samples from different acquisition files:

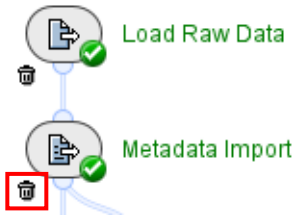
1. Select **Use File Name as Sample Name** in *Load Raw Data*.
2. Use the **File Name** (name of the wiff or wiff2 container file) in the **Experiment** column of the txt file for *Metadata Import*.

- To analyze multiple samples from a single acquisition file:
  1. <u>Do not</u> select **Use File Name as Sample Name** in *Load Raw Data*.
  2. Use the **Sample Name** in the **Experiment** column of the txt file for *Metadata Import*.





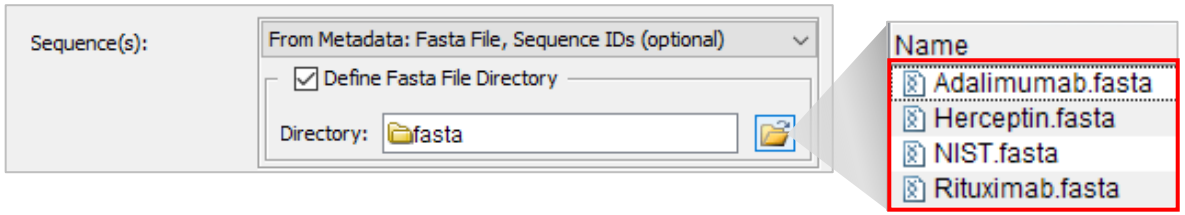
Metadata Import



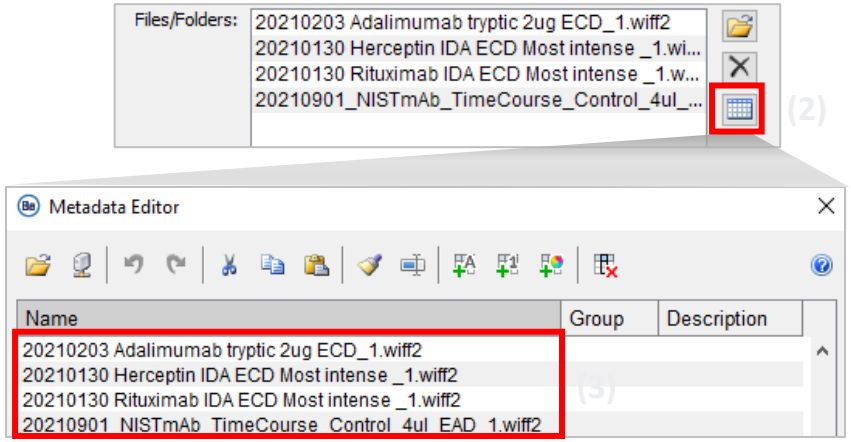
- To analyze multiple samples with the same sequence:
  - Deactivate the **Trash** icon, and then activate the **Bypass** icon for *Metadata Import*.
  - On the **Sequences** tab in *Protein Mapping*, select **From Text** or **From Fasta File**.
- To analyze multiple samples with different sequences:
  - Use *Metadata Import* to select the FASTA file (protein sequence) that will be used for identification in the *Protein Mapping* activity nodes.

- Upload a txt file with *Metadata Import* that links each sample to the correct FASTA file.
  - The name in the **Experiment** column must be the same as in the **Experiment** table in *Load Raw Data*.
  - The name in the **Fasta File** column must be the same as the name of the FASTA file that is in the specified **Fasta File Directory**, including the file extension (fasta or txt).

	A	B
1	Experiment	Fasta File
2	20210203 Adalimumab tryptic 2ug ECD_1	Adalimumab.fasta
3	20210130 Herceptin IDA ECD Most intense _1	Herceptin.fasta
4	20210130 Rituximab IDA ECD Most intense _1	Rituximab.fasta
5	20210901_NISTmAb_TimeCourse_Control_4ul_EAD_1	NIST.fasta



Metadata Import: Create the Metadata File



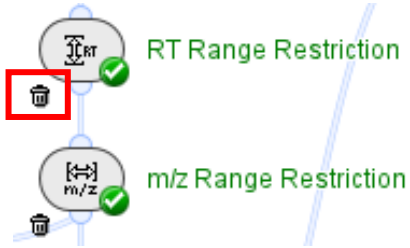
Note: Any metadata added in the *Load Raw Data* **Metadata Editor** table must be completed for all rows (all samples).

Experiment	Fasta File
20210203 Adalimumab tryptic 2ug ECD_1	Adalimumab.fasta
20210130 Herceptin IDA ECD Most intense _1	Herceptin.fasta
20210130 Rituximab IDA ECD Most intense _1	Rituximab.fasta
20210901_NISTmAb_TimeCourse_Control_4ul_EAD_1	NIST.fasta

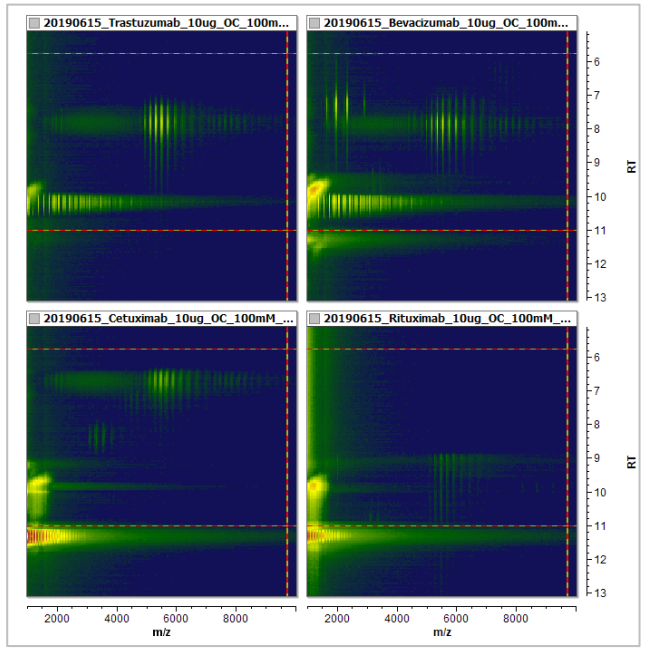
(4)

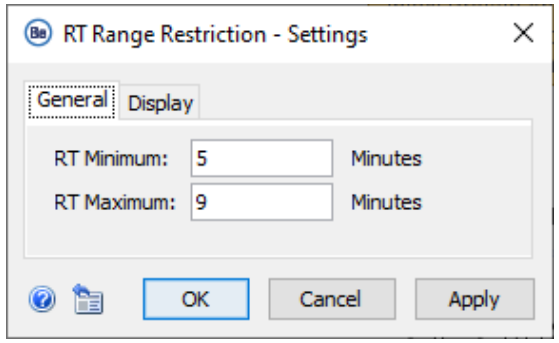
- To create the metadata file in Excel or Notepad:
  1. Select the samples for batch processing in *Load Raw Data*.
  2. Open the **Metadata Editor** table.
  3. Select all of the entries in the **Metadata Editor** table, and then select copy.
  4. Paste the entries into the **Experiment** column of the metadata txt file.
  5. Delete ".wiff" or ".wiff2" from the end of each name. (Tip: Use the Replace command in Excel or Notepad.)
  6. Type the applicable FASTA file name in each row in the **Fasta File** column.
  7. Save the file in txt format, and then upload the file in the *Metadata Import* activity node.

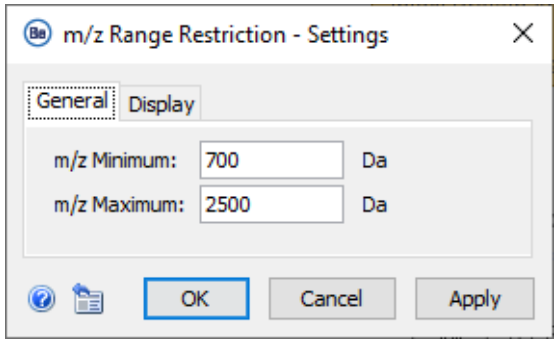
Restriction of RT and *m/z* Ranges



- To optimize settings and select applicable ranges, deactivate the **Trash** icon.
- To identify the RT ranges where there is meaningful data, open (double-click) *Load Raw Data* after the data is loaded.
  - Unless minor components are of interest, limit the ranges to the target protein.
  - Make sure that the selected ranges are wide enough to include all of the samples.
  - If the molecules require very different ranges, then activate the **Bypass** icon on the *Range Restriction* activity nodes

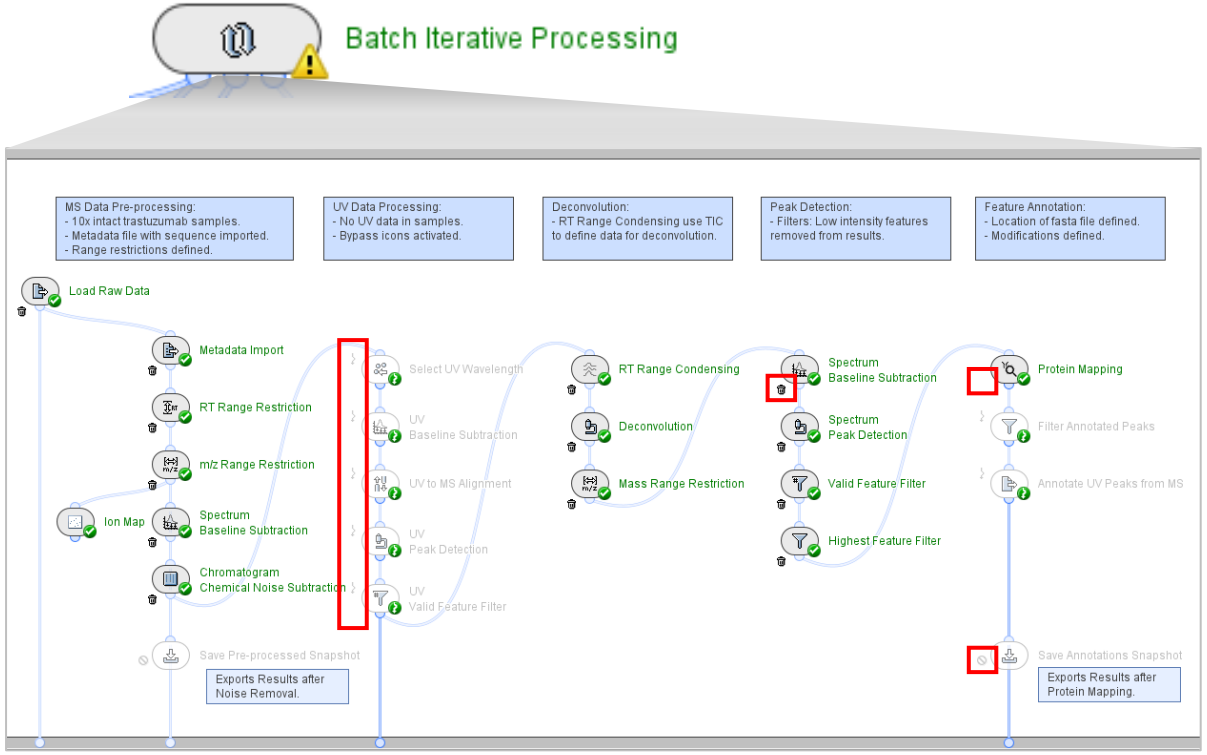






Batch Iterative Processing Container

- The *Batch Iterative Processing* container is not the same as other Biologics Explorer software containers.
  - Only intermediate results from the last sample to be processed can be opened from the activity nodes in the *Batch Iterative Processing* container.



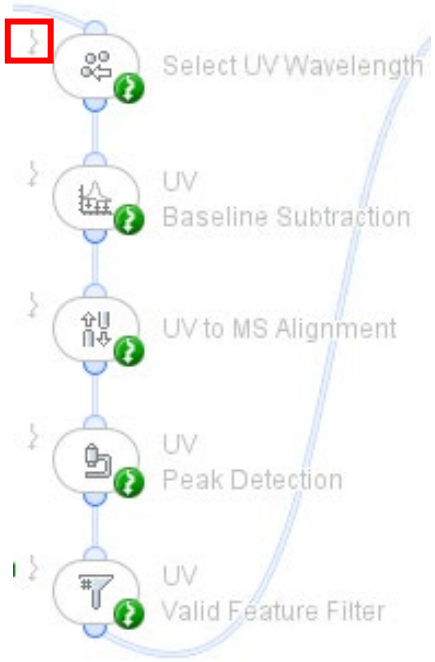
Note: If activity nodes in the container have the **Bypass** icon activated, then the container shows a yellow warning symbol.

- To open the intermediate results of an activity node, deactivate the **Trash** icon before the workflow is started.
- Do not activate the **Trash** icon for activities that are used in the PDF Report.

Activity nodes in the *Batch Iterative Processing* container do not have a **Run** or **Reset** icon.

- Activity nodes in the *Batch Iterative Processing* container cannot be run individually.
  - To use a *Save Snapshot* or *Export* activity node in the *Batch Iterative Processing* container, deactivate the **Block** icon before the workflow is started.

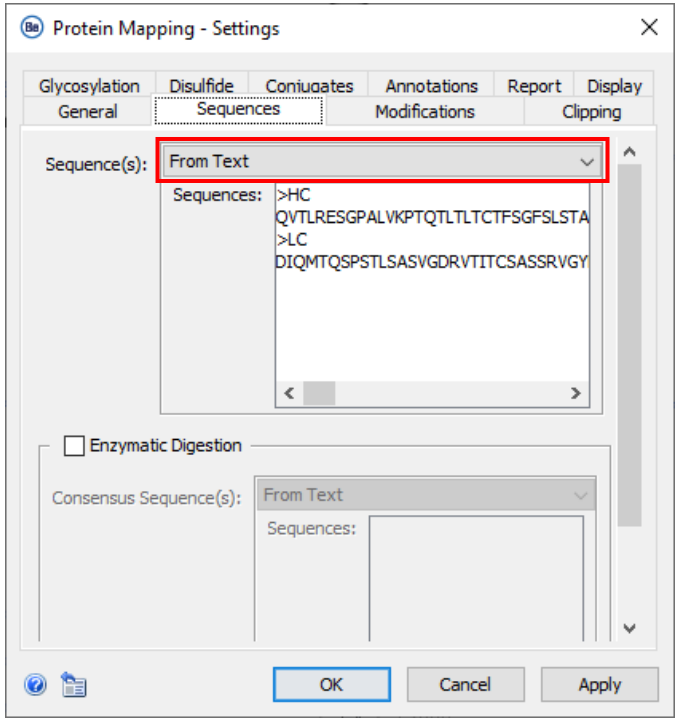
UV Data Processing



- All activity nodes for UV data have the **Bypass** icon activated in the template workflow.
- To use UV peaks in the data to identify the RT ranges for deconvolution:
  1. Deactivate the **Bypass** icon and then activate the **Trash** icon (to save memory) for all of the activity nodes with the prefix UV and *Annotate UV Peaks from MS*.

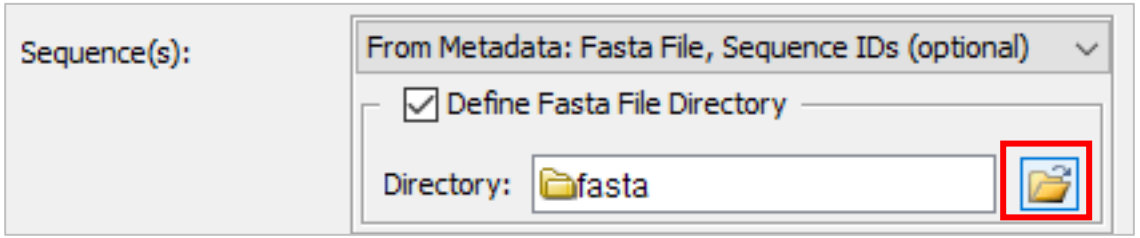
  - *Annotate UV Peaks from MS* uses MS peak identifications to add annotations to the peaks in the UV chromatogram. The peak must elute in the specified **RT Tolerance**.

Protein Mapping: Sequences



Sequences tab:

- **Sequence(s)**:
  - If all samples have the same sequence, then select **From Text** and type the sequence, or **From Fasta File** and select the applicable file.
  - If different samples require different sequences, then select **From Metadata: Fasta File, Sequence IDs (optional)**, and then browse to the location of the folder that contains all of the applicable FASTA files.



- For more information, refer to the next page: *Review Results*: Protein Name in FASTA Files.

Review Results: Protein Name in FASTA Files

- If the protein sequence names <u>are unique</u> across the FASTA files used for identification:

*Bevacizumab.fasta - Notepad

File Edit Format View Help

```
>HC(Bevacizumab)
EVQLVESGGGLVQPGGSLRLSCAASGYTFTNYGMNWVRQAPGKGLEWVGWINTYTGEPTY
>LC(Bevacizumab)
DIQMTQSPSSLSASVGDRVTITCSASQDISNYLNWYQQKPGKAPKVLIYFTSSLHSGVPS
```

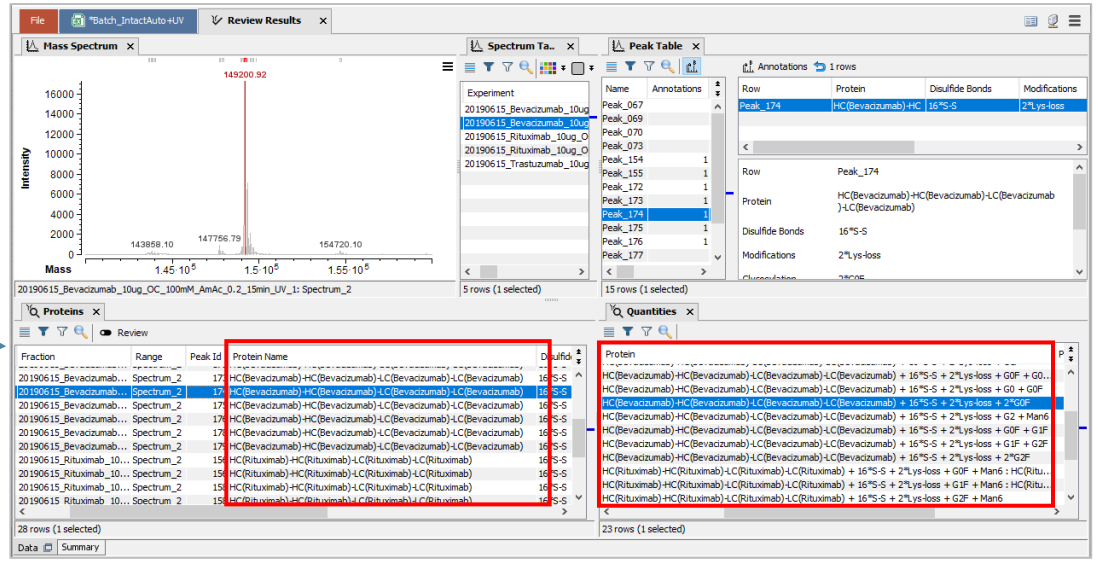
*Rituximab.fasta - Notepad

File Edit Format View Help

```
>HC(Rituximab)
QVQLQQPGAELVKPGASVKMSCKASGYTFTSYNMHWVKQTPGRGLEWIGAIYPGNGDTSY
>LC(Rituximab)
QIVLSQSPAILSASPGEKVTMTCRASSSVSYIHWFQQKPGSSPKPWIYATSNLASGVPVR
```

  - The protein sequence name in *Review Results* is the same as the name in the FASTA file.

Note: If the names are too long, then some table columns in the PDF report might be missing.



- If the protein sequence names <u>are not unique</u> across the FASTA files used for identification:

*Bevacizumab.fasta - Notepad

File Edit Format View Help

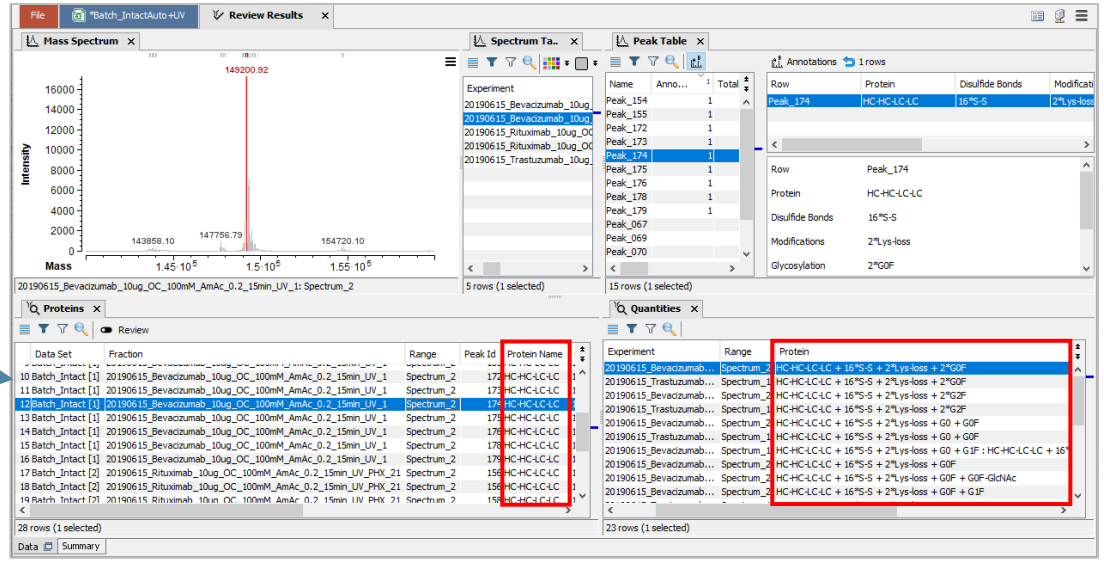
```
>HC
EVQLVESGGGLVQPGGSLRLSCAASGYTFTNYGMNWVRQAPGKGLEWVGWINTYTGEPTY
>LC
DIQMTQSPSSLSASVGDRVTITCSASQDISNYLNWYQQKPGKAPKVLIYFTSSLHSGVPS
```

*Rituximab.fasta - Notepad

File Edit Format View Help

```
>HC
QVQLQQPGAELVKPGASVKMSCKASGYTFTSYNMHWVKQTPGRGLEWIGAIYPGNGDTSY
>LC
QIVLSQSPAILSASPGEKVTMTCRASSSVSYIHWFQQKPGSSPKPWIYATSNLASGVPVR
```

  - The protein sequence name in *Review Results* includes the FASTA file name.



Trademarks/ Licensing

The SCIEX clinical diagnostic portfolio is For In Vitro Diagnostic Use. Rx Only. Product(s) not available in all countries. For information on availability, please contact your local sales representative or refer to www.sciex.com/diagnostics. All other products are For Research Use Only. Not for use in Diagnostic Procedures.

The power of precision

danaher