

# Intact Nucleotide Template Workflows

Biologics Explorer Software Guidelines

Powered by Genedata Expressionist®



TPUB-CUST-16012-B

© 2025 DH Tech. Dev. Pte. Ltd.



# Intact Nucleotide Template Workflow

## CONTENTS OF THIS GUIDE

### A: Overview of the Intact Nucleotide Workflows

1. Overview of Applications for the Intact Nucleotide Template Workflows
2. Intact Nucleotide Template Workflows

### B: Activity Nodes in the Intact Nucleotide Workflows

1. Nucleotide Candidate Generation
2. UV Data Preparation
3. Adduct Grouping
4. Feature Annotation
  - *Mass Mapping*
  - *Annotate UV Peaks from MS*
5. Customized Report Elements

### C: Guidelines for the Intact Nucleotide Workflows

1. Intact Nucleotide with Deconvolution
2. Intact Nucleotide with no Deconvolution

### D: Guidelines for Specific Applications

1. Recommended Settings for Isotopically Nonresolved Data.

---

## Part A

# Overview of the Intact Nucleotide Template Workflow

# A

# Overview of Applications for Intact Nucleotide Workflows

- Use the Intact Nucleotide template workflow with deconvolution to analyze these types of molecules:
  - Large synthetic oligonucleotides with their related impurities and modifications.
  - Large nucleotides with their related impurities and modifications, including Poly(A) Tails and 5' Caps.
- Use the Intact Nucleotide template workflow with no deconvolution to analyze these types of molecules:
  - Smaller oligonucleotides (less than 10 kDa) with their related impurities and modifications.
- The Nucleotide Candidate Generation activity nodes create a list of possible nucleotide forms and impurities to annotate MS features in *Mass Mapping*.
- The RT ranges for deconvolution of MS signals can be identified manually, or by the UV or TIC peaks.

# Intact Nucleotide Template Workflows

## **IntactNucleotide\_Deconvolution:**

- An intact nucleotide analysis workflow with spectral deconvolution for the identification of nucleotides and their impurities from MS1-only data (data without MS/MS fragmentation).

## **IntactNucleotide\_NoDeconvolution:**

- An intact nucleotide analysis workflow with no deconvolution for the identification of nucleotides and their impurities from MS1-only data (data without MS/MS fragmentation).

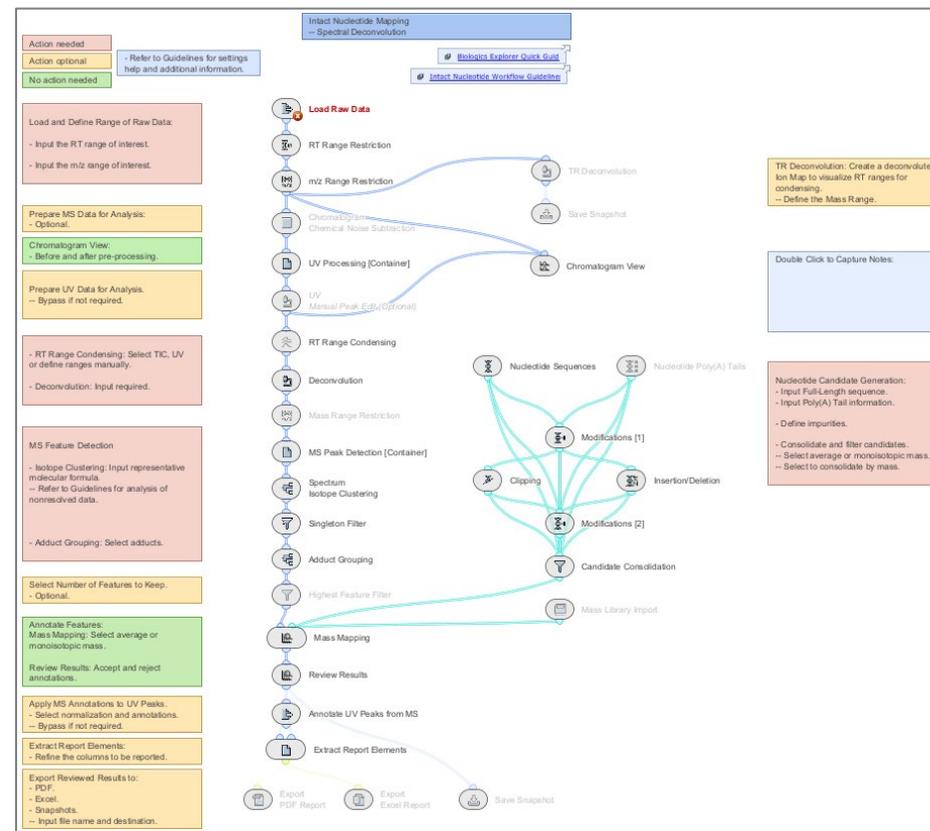
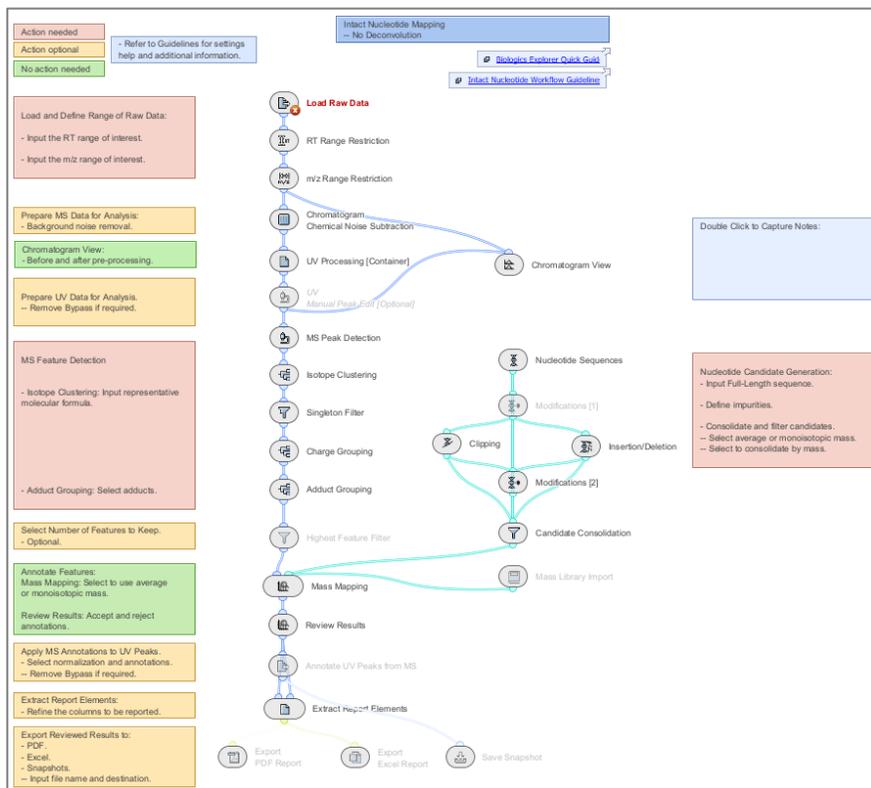
---

## Part B

# Activity Nodes in the Intact Nucleotide Workflows

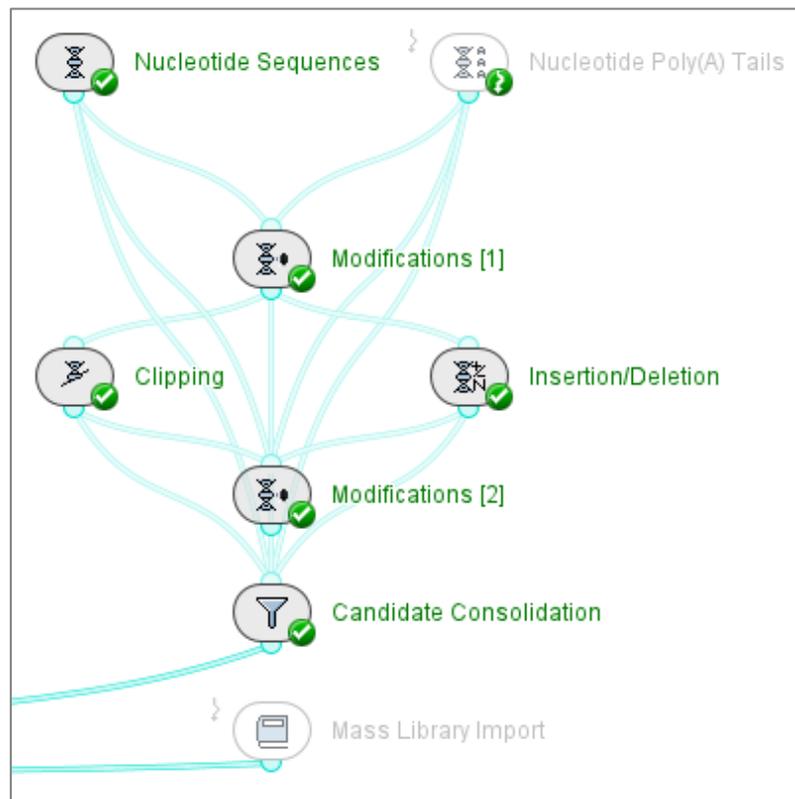
# B

# Intact Nucleotide Template Workflows

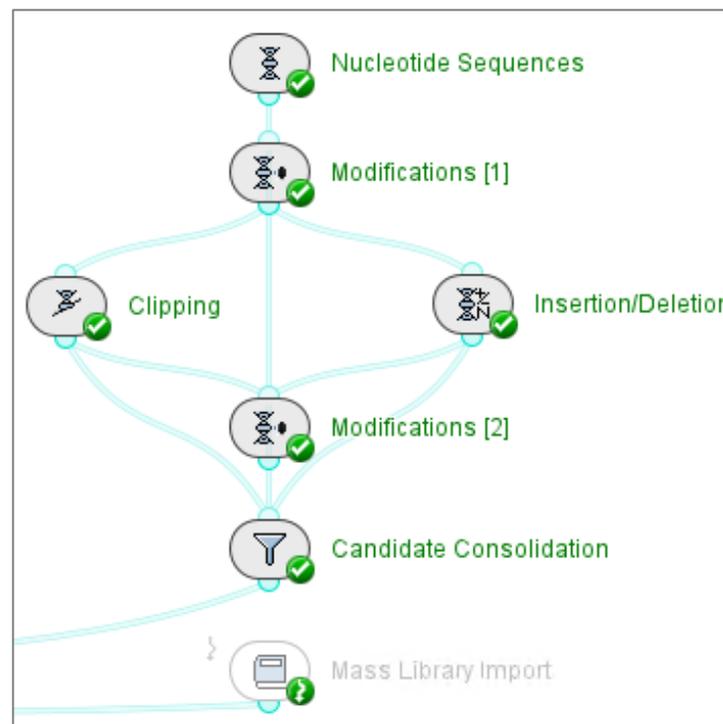


Note: For information about activity nodes that are used in all workflows, for example *Load Raw Data*, *Review Results*, or *Export PDF Report*, refer to the document: [Biologics Explorer Quick Guide](#).

# Nucleotide Candidate Generation



IntactNucleotide\_Deconvolution

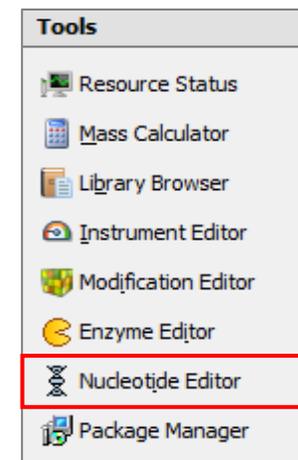


IntactNucleotide\_NoDeconvolution

- Use the Nucleotide Candidate Generation activity nodes to create a list of theoretical candidates of the full-length sequence with the applicable impurities and modifications.

# Nucleotide Candidate Generation: Nucleotide Editor

- To review the Nucleotide Building Blocks or to add custom (**USER**) Building Blocks, browse to **File > Tools > Nucleotide Editor**.
  - The most frequently used nucleotide bases, sugars, and linkers are included as **SYSTEM** building blocks that cannot be edited.
  - Some less frequently used bases and sugars are included as **USER** building blocks that can be edited.



Nucleotide Editor

Enter Filter Text

Type	Abbreviation	Name	Formula	Mass	Source
Sugar	d	Deoxyribose	H <sub>10</sub> C <sub>5</sub> O <sub>4</sub>	134.058	SYSTEM
Sugar	r	Ribose	H <sub>10</sub> C <sub>5</sub> O <sub>5</sub>	150.053	SYSTEM
Linker	o	Phosphate	H <sub>3</sub> O <sub>4</sub> P	97.977	SYSTEM
Linker	s	Phosphorothioate	H <sub>3</sub> O <sub>3</sub> PS	113.954	SYSTEM
Base	A	Adenine	H <sub>5</sub> C <sub>5</sub> N <sub>5</sub>	135.054	SYSTEM
Base	G	Guanine	H <sub>5</sub> C <sub>5</sub> N <sub>5</sub> O	151.049	SYSTEM
Base	C	Cytosine	H <sub>5</sub> C <sub>4</sub> N <sub>3</sub> O	111.043	SYSTEM
Base	T	Thymine	H <sub>5</sub> C <sub>5</sub> N <sub>3</sub> O <sub>2</sub>	126.043	SYSTEM
Base	U	Uracil	H <sub>4</sub> C <sub>4</sub> N <sub>2</sub> O <sub>2</sub>	112.027	SYSTEM
Base	Q	5-Methyl-cytosine	H <sub>7</sub> C <sub>5</sub> N <sub>3</sub> O	125.059	USER
Base	H	Hypoxanthine	H <sub>4</sub> C <sub>4</sub> N <sub>4</sub> O	136.039	USER
Sugar	l	LNA (Locked nucleic acid)	H <sub>10</sub> C <sub>6</sub> O <sub>5</sub>	162.053	USER
Sugar	m	2'-OMe (2'-O-Methylribose)	H <sub>12</sub> C <sub>6</sub> O <sub>5</sub>	164.068	USER
Sugar	f	2'-F (2'-Deoxy-2'-fluororibose)	H <sub>9</sub> C <sub>5</sub> O <sub>4</sub> F	152.048	USER
Sugar	e	2'-MOE (2'-Methoxyethylribose)	H <sub>16</sub> C <sub>9</sub> O <sub>6</sub>	208.095	USER

1 out of 15 items selected

Edit Building Block

Type: Sugar

Abbreviation: m

Name: 2'-OMe (2'-O-Methylribose)

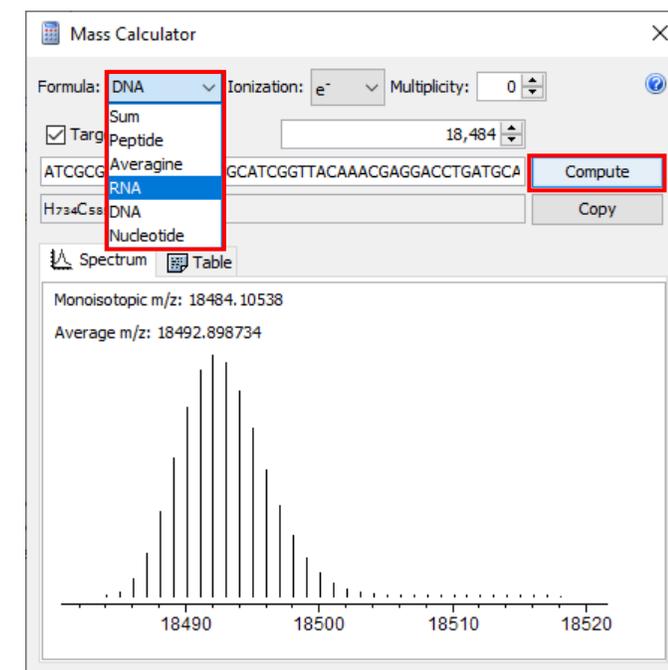
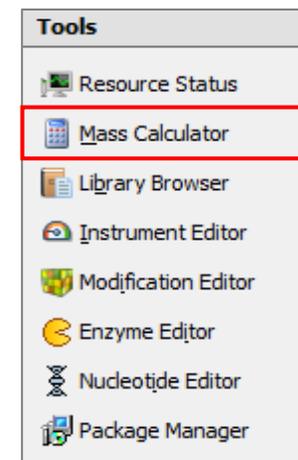
Formula: H12C6O5

- To edit a **USER** building block, select the entry, and then click the  icon.
- To create a new **USER** building block, click the  icon.
  - The abbreviation for a base must be a single, uppercase letter.
  - The abbreviation for a sugar or linker must be a single, lowercase letter.

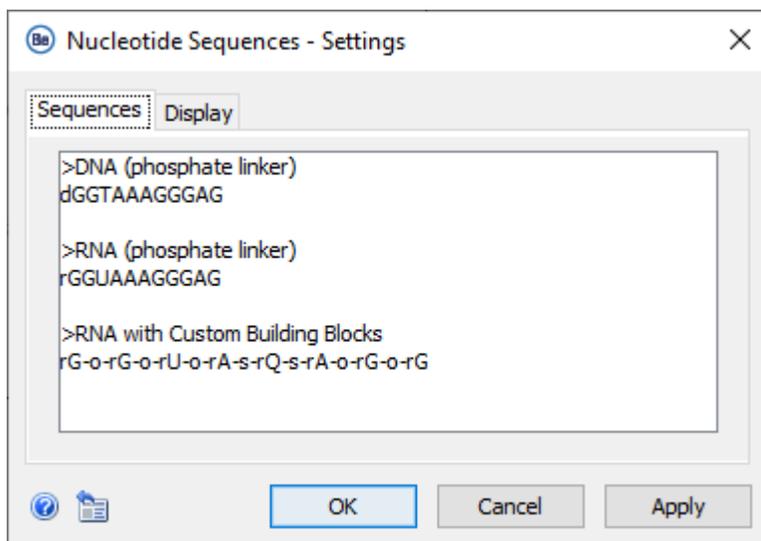
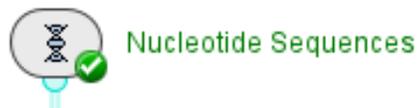
# Nucleotide Candidate Generation: Mass Calculator

- Use the **Mass Calculator** to give the theoretical mass and isotopic distribution of a DNA, RNA, or nucleotide sequence.
  - To open the Mass Calculator, browse to **File > Tools > Mass Calculator**.
  - For a sequence that contains bases with standard (natural) phosphate linkers and ribose (RNA) or deoxyribose (DNA) as the sugar, select **Format: RNA** or **DNA**, and then use single-base notation.
  - For a sequence that contains non-standard (synthetic) **USER** sugars or linkers, select **Format: Nucleotide**, and then use triplet notation:
    - Nucleotides triplets can be given in any order, with or without dashes. For example:
      - fA-o-dA-o-dC-o-dT-o-dA-o-dG-o
      - fAodAodCodTodAodGo
      - lU-oAr-oCr-oUr-oAr-oGr-o
      - lUoAroCroUroAroGro
  - To use the chemical formula in *Spectrum Isotope Clustering*, click the **Copy** button.

Note: For information, refer to the section: [C: Guidelines for the Intact Nucleotide Workflows](#).



# Nucleotide Candidate Generation: *Nucleotide Sequences*



- **Sequences:** Use FASTA format:
  - For the input name, use the prefix **>**.
  - For DNA sequences, use the prefix **d**.
  - For RNA sequences, use no prefix, or the prefix **r**.
  - For sequences with custom (**USER**) Building Blocks, use triplet notation.
- The **Input Sequence** column in the **Nucleotides** Result Table is a reference for the **Backbone Index**.
  - The **Base Index** counts the bases. The **Backbone Index** also counts the linkers:

<b>Input Sequence:</b>	rG	-o-	rG	-o-	rU	-o-	rA	-o-	rA...
<b>Backbone Index:</b>	1	2	3	4	5	6	7	8	9...
<b>Base Index:</b>	1		2		3		4		5...

Identifier	Sequence	Bases	Base Delta	Type	From (Base)	To (Base)	From (Backbone)	To (Backbone)	Calc. Mass	Calc. Avg. Mass	Formula	Input Sequence
DNA (phosphate linker)	dGGTAAAGGGAG	11 N		Full-Length	1	11	1	21	3468.636	3470.295	H <sub>134</sub> C <sub>110</sub> N <sub>52</sub> O <sub>61</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-dG-o-dG-o-dG-o-dA-o-dG
RNA (phosphate linker)	rGGUAAAGGGAG	11 N		Full-Length	1	11	1	21	3630.564	3632.262	H <sub>132</sub> C <sub>109</sub> N <sub>52</sub> O <sub>72</sub> P <sub>10</sub>	rG-o-rG-o-rU-o-rA-o-rA-o-rA-o-rG-o-rG-o-rG-o-rA-o-rG
RNA with Custom Building Blocks	rGGUA-s-rQ-s-rAGG	8 N		Full-Length	1	8	1	15	2633.376	2634.776	H <sub>98</sub> C <sub>79</sub> N <sub>35</sub> O <sub>51</sub> P <sub>7</sub> S <sub>2</sub>	rG-o-rG-o-rU-o-rA-s-rQ-s-rA-o-rG-o-rG

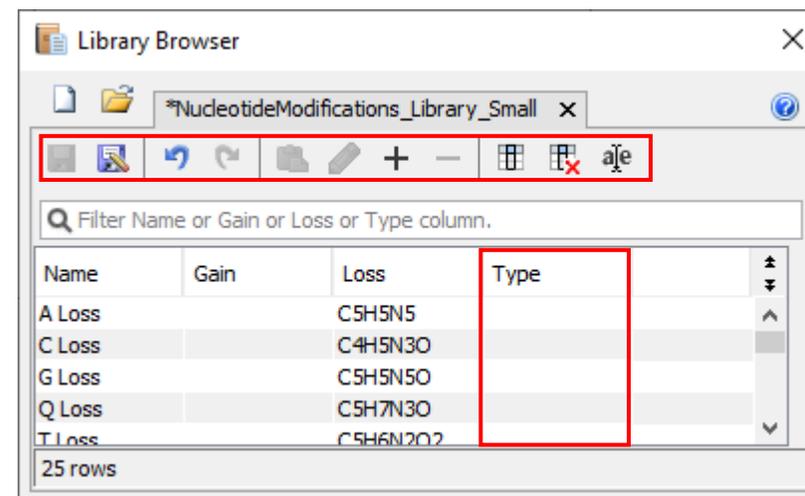
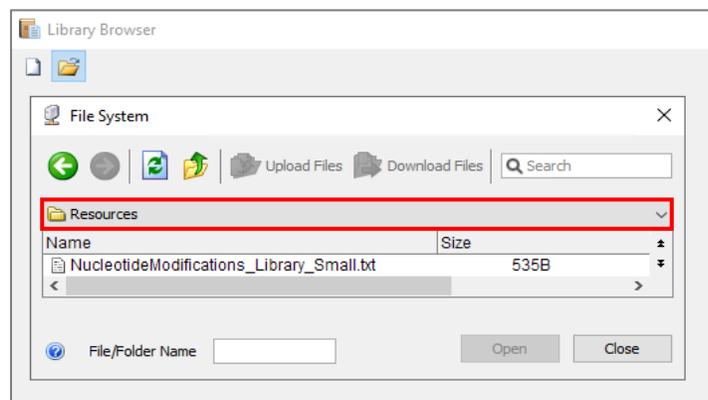
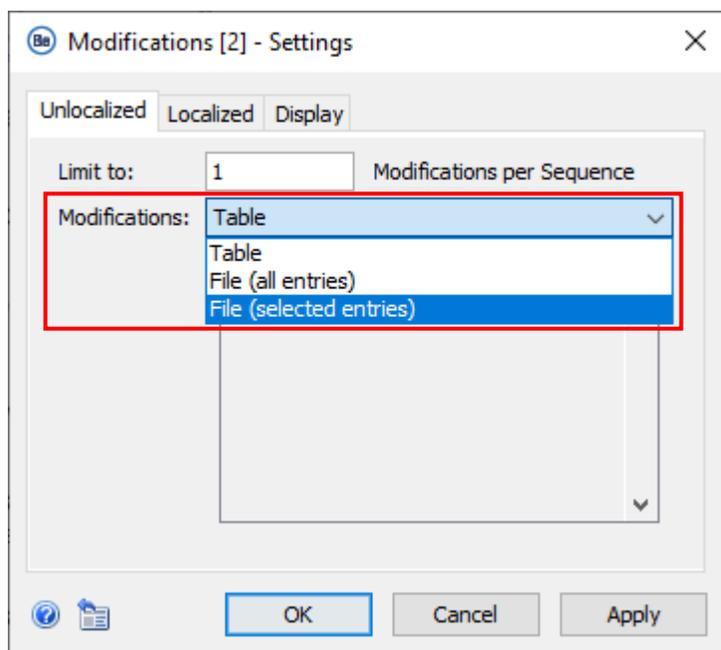


Note: For information about Poy(A) Tail candidate generation, refer to the section: *C1: Guidelines for the Intact Nucleotide with Deconvolution Template Workflow*.

# Nucleotide Candidate Generation: *Modifications* Settings



- Select an option to add the definition of the **Modifications**:
  - **Table**: Type the modifications of interest.
  - **File (all entries)**: Use all of the modifications in a pre-defined txt file.
  - **File (selected entries)**: Use a selection of the modifications from a pre-defined txt file.
- To see a list of example modifications, browse to **Library Browser > Resources > NucleotideModifications\_Library\_Small**.
  - Open this file in the **Library Browser** to customize it for use in *Modifications*.



Note: The number of nucleotide candidates created by the selected settings cannot be more than the threshold (100,000).

# Nucleotide Candidate Generation: *Modifications* Settings



- **Unlocalized** tab: Create modifications on the nucleotide sequence with a **Gain** or **Loss** (by chemical formula or mass in Da).
- **Localized** tab: Create modifications on a specified **Character** or **Location** (backbone index) with a **Gain** or **Loss** (by chemical formula or mass in Da).
  - For analysis of mRNA with a 5' capping species, use **5'** as the **Location** for the 5' cap.

Modifications [1] - Settings

Unlocalized Localized Display

Limit to: 1 Modifications per Sequence

Modifications: Table

Name	Gain	Loss	On Characters	On Locations ...	Type
- pp	H4P2O7	H2O		5'	Uncapped
- ppp	H5P3O10	H2O		5'	Uncapped
- G Cap	C10H16N5O14P3	H2O		5'	Partial
- Cap 0	C11H18N5O14P3	H2O		5'	Capped
- Cap 1	C12H20N5O14P3	H2O		5'	Capped

OK Cancel Apply

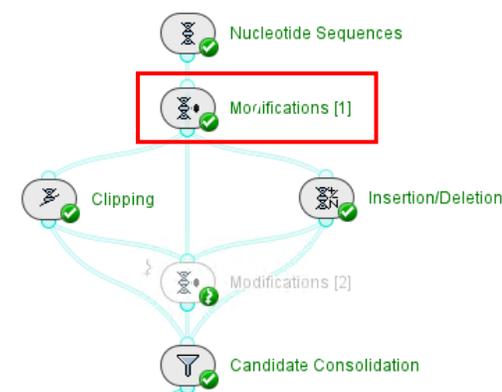
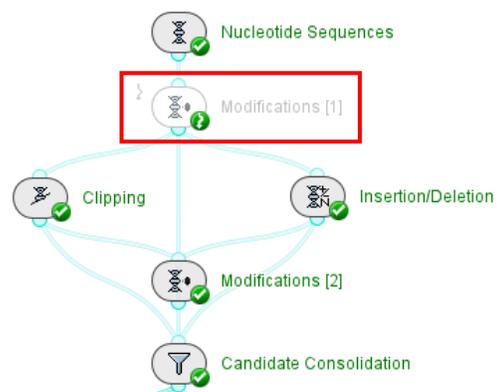
- **Type**: Optionally, create a customized nucleotide Type for each modification.
  - If left blank, then the **Type** will show in the **Nucleotides Results Table** as Modified.

Nucleotides x

Identifier	Sequence	Bases	Base Delta	Type	Mod. Location (Backbone)	Modifications
mRNA Cap 0 [1]	rGGGAGACG...	18 N		Capped	Cap 0 [1]	1
mRNA Cap 1 [1]	rGGGAGACG...	18 N		Capped	Cap 1 [1]	1
mRNA OH	rGGGAGACG...	18 N		Modified	OH	1
mRNA G Cap [1]	rGGGAGACG...	18 N		Partial	G Cap [1]	1
mRNA pp [1]	rGGGAGACG...	18 N		Uncapped	pp [1]	1
mRNA ppp [1]	rGGGAGACG...	18 N		Uncapped	ppp [1]	1
mRNA G Cap [1], OH	rGGGAGACG...	18 N		Modified, Partial	G Cap [1], OH	2

# Nucleotide Candidate Generation: *Modifications* and Impurities

- To create both modified and unmodified versions of candidates with insertions or deletions, activate the **Bypass** icon on *Modifications* [1].
  - If the **Bypass** icon is not activated on *Modifications* [1], then only modified versions of candidates with insertions or deletions are created.



Type	Identifier	Seque...	Bases	Base Delta	Type
3' Clip	100 mer -(dG-o[193])	dATCG...	99 N-1		Deletion (N-1)
3' Clip+Linker	100 mer -(dC-o[195])	dATCG...	99 N-1		Deletion (N-1)
3' Clip+Linker, Modified	100 mer -(dG-o[197])	dATCG...	99 N-1		Deletion (N-1)
3' Clip, Modified	100 mer +(dA-o[1])	dAATC...	101 N+1		Insertion (N+1)
5' Clip	100 mer +(dT-o[3])	dATTC...	101 N+1		Insertion (N+1)
5' Clip+Linker	100 mer +(dC-o[5])	dATCC...	101 N+1		Insertion (N+1)
5' Clip+Linker, Modified	100 mer -(dC-o[195]) G Loss	dATCG...	99 N-1		Deletion (N-1), Modified
5' Clip, Modified	100 mer -(dC-o[195]) T Loss	dAATC...	99 N-1		Deletion (N-1), Modified
Deletion (N-1)	100 mer -(dG-o[197]) A Loss	dATCG...	99 N-1		Deletion (N-1), Modified
Deletion (N-1), Modified	100 mer -(dG-o[197]) C Loss	dATCG...	99 N-1		Deletion (N-1), Modified
Full-Length	100 mer -(dG-o[197]) G Loss	dATCG...	99 N-1		Deletion (N-1), Modified
Insertion (N+1)	100 mer -(dG-o[197]) T Loss	dATCG...	99 N-1		Deletion (N-1), Modified
Insertion (N+1), Modified	100 mer +(dA-o[1]) A Loss	dAATC...	101 N+1		Insertion (N+1), Modified
Modified	100 mer +(dA-o[1]) C Loss	dAATC...	101 N+1		Insertion (N+1), Modified

Type	Identifier	Seque...	Bases	Base Delta	Type
3' Clip	100 mer -(dC-o[195]) G Loss	dATCG...	99 N-1		Deletion (N-1), Modified
3' Clip+Linker	100 mer -(dC-o[195]) T Loss	dATCG...	99 N-1		Deletion (N-1), Modified
3' Clip+Linker, Modified	100 mer -(dG-o[197]) A Loss	dATCG...	99 N-1		Deletion (N-1), Modified
3' Clip, Modified	100 mer -(dG-o[197]) C Loss	dATCG...	99 N-1		Deletion (N-1), Modified
5' Clip	100 mer -(dG-o[197]) G Loss	dATCG...	99 N-1		Deletion (N-1), Modified
5' Clip+Linker	100 mer -(dG-o[197]) T Loss	dATCG...	99 N-1		Deletion (N-1), Modified
5' Clip+Linker, Modified	100 mer +(dA-o[1]) A Loss	dAATC...	101 N+1		Insertion (N+1), Modified
5' Clip, Modified	100 mer +(dA-o[1]) C Loss	dAATC...	101 N+1		Insertion (N+1), Modified
Deletion (N-1), Modified	100 mer +(dA-o[1]) G Loss	dAATC...	101 N+1		Insertion (N+1), Modified
Full-Length	100 mer +(dA-o[1]) T Loss	dAATC...	101 N+1		Insertion (N+1), Modified
Insertion (N+1), Modified	100 mer +(dT-o[3]) A Loss	dATTC...	101 N+1		Insertion (N+1), Modified
Modified	100 mer +(dT-o[3]) C Loss	dATTC...	101 N+1		Insertion (N+1), Modified



# Nucleotide Candidate Generation: *Modifications* Results



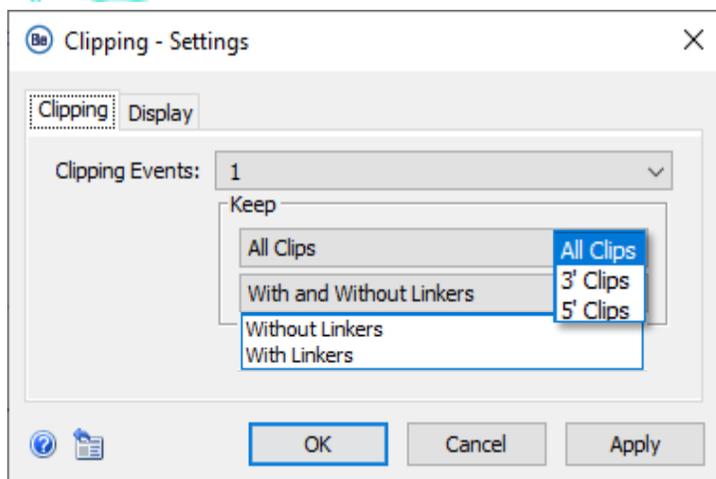
## Nucleotides Result Table:

- The candidate **Identifier** is updated to include the modification.
- The **Nucleotides Results Table** is updated to contain more columns:
  - Type:** Modified or as defined in the *Modifications* settings.
  - Modifications:** The number of modifications on the candidate.
  - Mod. Location (Backbone):** The name and location of the modification.

Identifier	Sequence	Bases	Base Delta	Type	Modifications	Mod. Location (Backbone)	From (Base)	To (Base)	From (Backbone)	To (Backbone)	Calc. Mass	Calc. Avg. Mass	Formula	Input Sequence
DNA 2×A-loss	dGGTAAAGGGAG	11 N		Modified	2 2×A-loss		1	11	1	21	3234.548	3236.072	H <sub>128</sub> C <sub>100</sub> N <sub>42</sub> O <sub>63</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-dG-o-dG-o-dG-o-dA-o-dG
DNA 2×C-loss	dGGTAAAGGGAG	11 N		Modified	2 2×C-loss		1	11	1	21	3282.570	3284.121	H <sub>128</sub> C <sub>102</sub> N <sub>46</sub> O <sub>61</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-dG-o-dG-o-dG-o-dA-o-dG
DNA 2×G-loss	dGGTAAAGGGAG	11 N		Modified	2 2×G-loss		1	11	1	21	3202.558	3204.073	H <sub>128</sub> C <sub>100</sub> N <sub>42</sub> O <sub>61</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-dG-o-dG-o-dG-o-dA-o-dG
DNA 2×T-loss	dGGTAAAGGGAG	11 N		Modified	2 2×T-loss		1	11	1	21	3252.571	3254.098	H <sub>126</sub> C <sub>100</sub> N <sub>48</sub> O <sub>59</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-dG-o-dG-o-dG-o-dA-o-dG
DNA A-loss	dGGTAAAGGGAG	11 N		Modified	1 A-loss		1	11	1	21	3351.592	3353.183	H <sub>131</sub> C <sub>105</sub> N <sub>47</sub> O <sub>62</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-dG-o-dG-o-dG-o-dA-o-dG
DNA A-loss, C-loss	dGGTAAAGGGAG	11 N		Modified	2 A-loss, C-loss		1	11	1	21	3258.559	3260.096	H <sub>128</sub> C <sub>101</sub> N <sub>41</sub> O <sub>62</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-dG-o-dG-o-dG-o-dA-o-dG
DNA A-loss, G-loss	dGGTAAAGGGAG	11 N		Modified	2 A-loss, G-loss		1	11	1	21	3218.553	3220.072	H <sub>128</sub> C <sub>100</sub> N <sub>42</sub> O <sub>62</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-dG-o-dG-o-dG-o-dA-o-dG
DNA A-loss, T-loss	dGGTAAAGGGAG	11 N		Modified	2 A-loss, T-loss		1	11	1	21	3243.560	3245.085	H <sub>127</sub> C <sub>100</sub> N <sub>45</sub> O <sub>61</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-dG-o-dG-o-dG-o-dA-o-dG
DNA C-loss	dGGTAAAGGGAG	11 N		Modified	1 C-loss		1	11	1	21	3375.603	3377.208	H <sub>131</sub> C <sub>106</sub> N <sub>49</sub> O <sub>61</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-dG-o-dG-o-dG-o-dA-o-dG
DNA C-loss, G-loss	dGGTAAAGGGAG	11 N		Modified	2 C-loss, G-loss		1	11	1	21	3242.564	3244.097	H <sub>128</sub> C <sub>101</sub> N <sub>41</sub> O <sub>61</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-dG-o-dG-o-dG-o-dA-o-dG
DNA C-loss, T-loss	dGGTAAAGGGAG	11 N		Modified	2 C-loss, T-loss		1	11	1	21	3267.571	3269.110	H <sub>127</sub> C <sub>101</sub> N <sub>47</sub> O <sub>60</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-dG-o-dG-o-dG-o-dA-o-dG
DNA G-loss	dGGTAAAGGGAG	11 N		Modified	1 G-loss		1	11	1	21	3335.597	3337.184	H <sub>131</sub> C <sub>105</sub> N <sub>47</sub> O <sub>61</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-dG-o-dG-o-dG-o-dA-o-dG
DNA G-loss, T-loss	dGGTAAAGGGAG	11 N		Modified	2 G-loss, T-loss		1	11	1	21	3227.565	3229.086	H <sub>127</sub> C <sub>100</sub> N <sub>45</sub> O <sub>60</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-dG-o-dG-o-dG-o-dA-o-dG
DNA T-loss	dGGTAAAGGGAG	11 N		Modified	1 T-loss		1	11	1	21	3360.603	3362.197	H <sub>130</sub> C <sub>105</sub> N <sub>50</sub> O <sub>60</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-dG-o-dG-o-dG-o-dA-o-dG



# Nucleotide Candidate Generation: *Clipping*



- **Clipping Events:** Select 0, 1, or 2 clips.
- **Keep:** Select the applicable options from the lists:
  - All Clips, 3' Clips, 5' Clips.
  - With and Without Linkers, Without Linkers, With Linkers.

Note: For more information, click the ? icon to open the [Online Help](#).

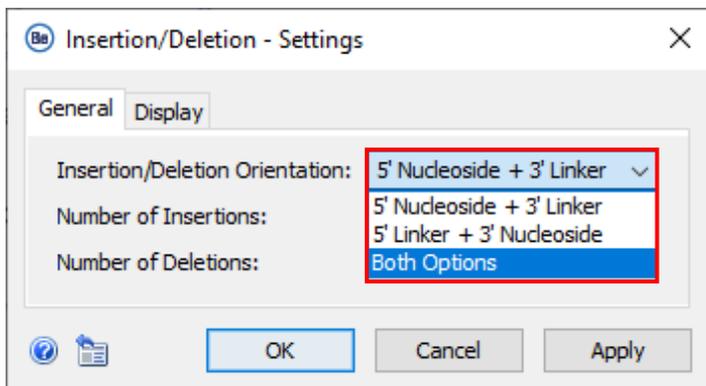
## Nucleotides Result Table:

- The candidate **Identifier** is updated to include the backbone index of the clipped candidate.
- The **Nucleotides Result Table** is updated to contain more columns:
  - **Type:** The position of the clip and if it contains a linker.
  - **Clips:** The number of clips to create the candidate.
  - **Base-Delta:** The number of bases removed from the full-length sequence (N).

Note: The number of nucleotide candidates created by the selected settings cannot be more than the threshold (100,000).

Identifier	Sequence	Bases	Base Delta	Type	Clips
100 mer [1-9]	dATCGC	5	N-95	3' Clip	1
100 mer [1-11]	dATCGCG	6	N-94	3' Clip	1
100 mer [1-10]	dATCGC-o	5	N-95	3' Clip+Linker	1
100 mer [191-199]	dAGCGA	5	N-95	5' Clip	1
100 mer [190-199]	o-dAGCGA	5	N-95	5' Clip+Linker	1
100 mer [3-11]	dTCGCG	5	N-95	3' Clip, 5' Clip	2
100 mer [3-12]	dTCGCG-o	5	N-95	3' Clip+Linker, 5' Clip	2
100 mer [2-11]	o-dTCGCG	5	N-95	3' Clip, 5' Clip+Linker	2

# Nucleotide Candidate Generation: *Insertion/Deletion*



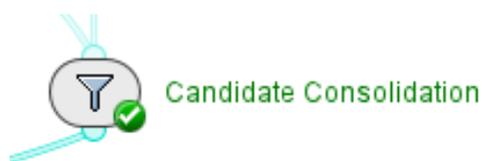
- **Number of Insertions:** Select 0 or 1 repetitions of a nucleotide.
- **Number of Deletions:** Select 0, 1, or 2 removals of a nucleotide.

## Nucleotides Result Table:

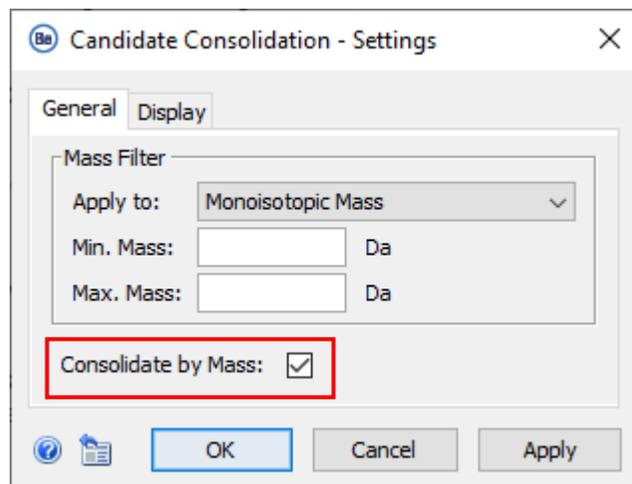
- The **Name** of the candidate is updated to include the inserted or deleted nucleotide and the backbone index location with the following format:
  - Insertion: +(sugarBase-linker[backbone index]).
  - Deletion: -(sugarBase-linker[backbone index]).
- The **Nucleotides** Result Table is updated to contain more columns:
  - **Type:** The change to the full-length sequence (N).
  - **Deletion/Insertion (Base):** The nucleotide with the base index location.
  - **Deletion/Insertions (Backbone):** The nucleotide with the backbone index location.
  - **Bases:** The number of bases in the modified sequence.
  - **Base-Delta:** The change to the number of bases from the full-length sequence (N).

Identifier	Sequence	Bases	Base Delta	Type	Deletion (Base)	Deletion (Backbone)	Insertion (Base)	Insertion (Backbone)
100mer -(dG-o[11-12])	dATCGCGA...	99	N-1	Deletion (N-1)	dG-o[6]	dG-o[11-12]		
100mer -(o-dG[10-11])	dATCGCGA...	99	N-1	Deletion (N-1)	o-dG[6]	o-dG[10-11]		
100mer -(dT-o[3-4], dG-o[11-12])	dACGCGAT...	98	N-2	Deletion (N-2)	dT-o[2], dG-o[6]	dT-o[3-4], dG-o[11-12]		
100mer -(o-dT[2-3], o-dG[10-11])	dACGCGAT...	98	N-2	Deletion (N-2)	o-dT[2], o-dG[6]	o-dT[2-3], o-dG[10-11]		
100mer +(dG-o[11-12])	dATCGCGG...	101	N+1	Insertion (N+1)			dG-o[6]	dG-o[11-12]
100mer +(o-dG[10-11])	dATCGCGG...	101	N+1	Insertion (N+1)			o-dG[6]	o-dG[10-11]

# Nucleotide Candidate Generation: *Candidate Consolidation*



- Select **Consolidate by Mass** to combine candidates from all Nucleotide Candidate Generation activity nodes, and to merge candidates that have the same input name and input sequence from *Nucleotide Sequences*, and the same mass and modifications from the Nucleotide Candidate Generation activity nodes.
  - Use the **Mass Filter** to use only the candidates in a specified mass range.



## Nucleotides Result Table:

- The **Identifier** column is updated to include all candidate information.
- The **Type** column includes the prefix of **Consolidated**.
- For consolidated candidates, columns that would have different information are empty.

Identifier	Sequence	Bases	Base Delta	Type	Clips	Modifications	Mod. Location (...)	From (Base)	To (Base)	From (Backbone)	To (Backbone)	Calc. Mass	Calc. Avg. Mass	Formula	Input Name	Input Sequence
FLP	dATCGCGGA...	100 N		Full-Length	0	0		1	100	1	199	30876.14	30890.81	H <sub>1224</sub> C <sub>976</sub> N <sub>383</sub> O <sub>594</sub> P <sub>99</sub>	FLP	dA-o-dT-o-dC-o-dG-o-d...
FLP [1-15]	dATCGCGGA	8 N-92		3' Clip	1	0		1	8	1	15	2433.46	2434.63	H <sub>98</sub> C <sub>78</sub> N <sub>33</sub> O <sub>45</sub> P <sub>7</sub>	FLP	dA-o-dT-o-dC-o-dG-o-d...
FLP [1-10] A-loss [1], H2O Loss	dATCGCG-o	5 N-95		3' Clip+Linker, Base-loss, Modified	1	2	A-loss [1], H2O Loss	1	5	1	10	1389.19	1389.84	H <sub>56</sub> C <sub>13</sub> N <sub>13</sub> O <sub>20</sub> P <sub>5</sub>	FLP	dA-o-dT-o-dC-o-dG-o-d...
FLP [Oxidation] -(dG-o),-(o-dG)		99 N-1		Consolidated (Deletion (N-1), Modified)	0	1		1	100	1	199	30563.08	30577.61	H <sub>1212</sub> C <sub>966</sub> N <sub>378</sub> O <sub>589</sub> P <sub>98</sub>	FLP	dA-o-dT-o-dC-o-dG-o-d...
FLP [Oxidation] +(dT-o),+(o-dT)		101 N+1		Consolidated (Insertion (N+1), Modified)	0	1		1	100	1	199	31196.18	31211.01	H <sub>1227</sub> C <sub>986</sub> N <sub>383</sub> O <sub>602</sub> P <sub>100</sub>	FLP	dA-o-dT-o-dC-o-dG-o-d...
FLP [A-loss] [1-23]	dATCGCGGA...	12 N-88		Consolidated (3' Clip, Base-loss)	1	1		1	12	1	23	3548.60	3550.30	H <sub>143</sub> C <sub>113</sub> N <sub>12</sub> O <sub>70</sub> P <sub>11</sub>	FLP	dA-o-dT-o-dC-o-dG-o-d...
FLP [H2O Loss] [1-17, 183-199]		9 N-91		Consolidated (3' Clip, 5' Clip, Modified)	1	1						2719.49	2720.80	H <sub>109</sub> C <sub>88</sub> N <sub>35</sub> O <sub>51</sub> P <sub>8</sub>	FLP	dA-o-dT-o-dC-o-dG-o-d...

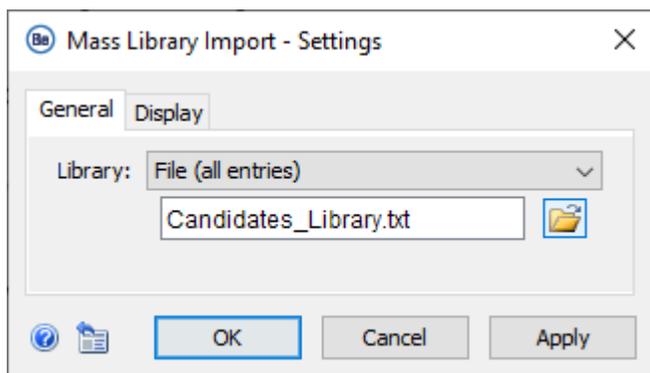


**Note:** Isomeric species with different input names or input sequences are not consolidated

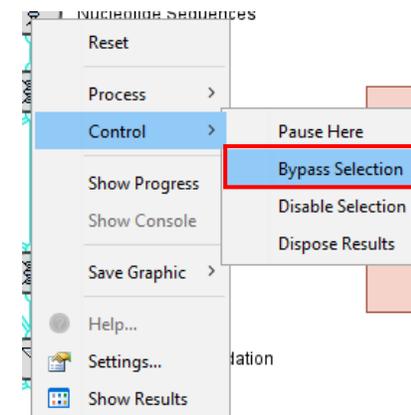
# Nucleotide Candidate Generation: *Mass Library Import*



- Use the *Mass Library Import* activity node to identify features from a list imported as a txt file.



- To **Bypass** the Nucleotide Candidate Generation activity nodes:
  1. Select all of the Nucleotide Candidate Generation activity nodes.
  2. Right-click, and then select **Control > Bypass Selection**.
- To use *Mass Library Import*:
  1. Deactivate the **Bypass** icon.
  2. Browse to a saved output from the other Nucleotide Candidate Generation activity nodes, or to a previously prepared list of candidates.
    - The library txt file must contain a **Calc. Mass** or **Calc. Avg. Mass** column.
      - Select the applicable **Mass Mode** in the *Mass Mapping* activity node.

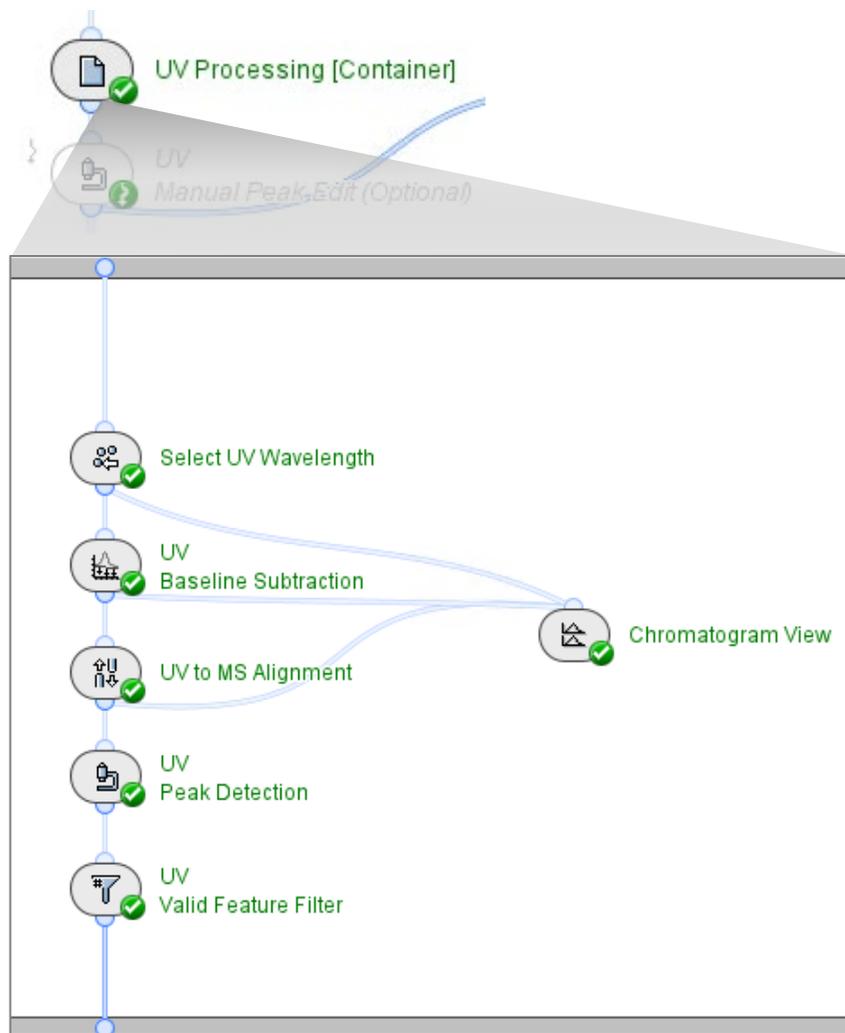


Identifier	Sequence	Bases	Base Delta	Type	Clips	Modifications	Mod. Location (Backbone)	Deletion (Base)	Deletion (Backbone)	Insertion (Base)	Insertion (Backbone)	From (Base)	To (Base)	From (Backbone)	To (Backbone)	Calc. Mass	Calc. Avg. Mass	Formula	Input Sequence
Oligo	dGGTAAAGGGAG	11	N	Library (Full-Length)	0	0						1	11	1	21	3468.636	3470.295	H <sub>131</sub> C <sub>110</sub> N <sub>52</sub> O <sub>41</sub> P <sub>10</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-d
Oligo [1-5]	dGGT	3	N-8	Library (3' Clip)	1	0						1	3	1	5	900.195	900.642	H <sub>38</sub> C <sub>30</sub> N <sub>12</sub> O <sub>17</sub> P <sub>2</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-d
Oligo [1-7]	dGGTA	4	N-7	Library (3' Clip)	1	0						1	4	1	7	1213.253	1213.849	H <sub>50</sub> C <sub>40</sub> N <sub>17</sub> O <sub>21</sub> P <sub>3</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-d
Oligo [1-9]	dGGTAA	5	N-6	Library (3' Clip)	1	0						1	5	1	9	1526.311	1527.056	H <sub>62</sub> C <sub>50</sub> N <sub>22</sub> O <sub>27</sub> P <sub>4</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-d
Oligo [1-17]	dGGTAAAGGG	9	N-2	Library (3' Clip)	1	0						1	9	1	17	2826.526	2827.882	H <sub>110</sub> C <sub>90</sub> N <sub>42</sub> O <sub>50</sub> P <sub>8</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-d
Oligo [1-6]	dGGT-o	3	N-8	Library (3' Clip+Linker)	1	0						1	3	1	6	980.162	980.622	H <sub>39</sub> C <sub>30</sub> N <sub>12</sub> O <sub>20</sub> P <sub>3</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-d
Oligo [1-8]	dGGTA-o	4	N-7	Library (3' Clip+Linker)	1	0						1	4	1	8	1293.219	1293.829	H <sub>51</sub> C <sub>40</sub> N <sub>17</sub> O <sub>21</sub> P <sub>3</sub>	dG-o-dG-o-dT-o-dA-o-dA-o-dA-o-d



- All columns in the library txt file are shown in the output table. The **Type** column includes the prefix of **Library**.

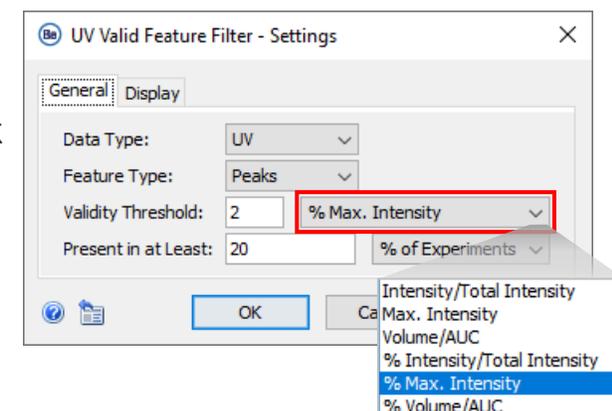
# UV Processing [Container]



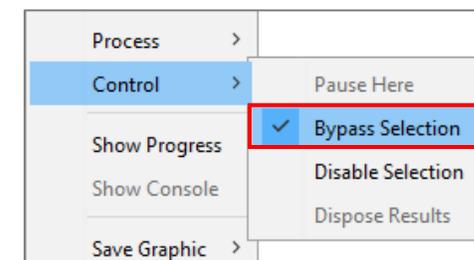
- To use UV data, select the correct value in *Select UV Wavelength*.
- To remove low intensity peaks that are not of interest, set a threshold in *UV Valid Feature Filter*.

- If the **Validity Threshold** is set to a percentage of an observable (**% Intensity/Total Intensity**, **% Max. Intensity** or **% Volume/AUC**), then the largest peak is used to calculate the percentage, not the sum of all peaks.

Note: For more information, click the ? icon to open the [Online Help](#).



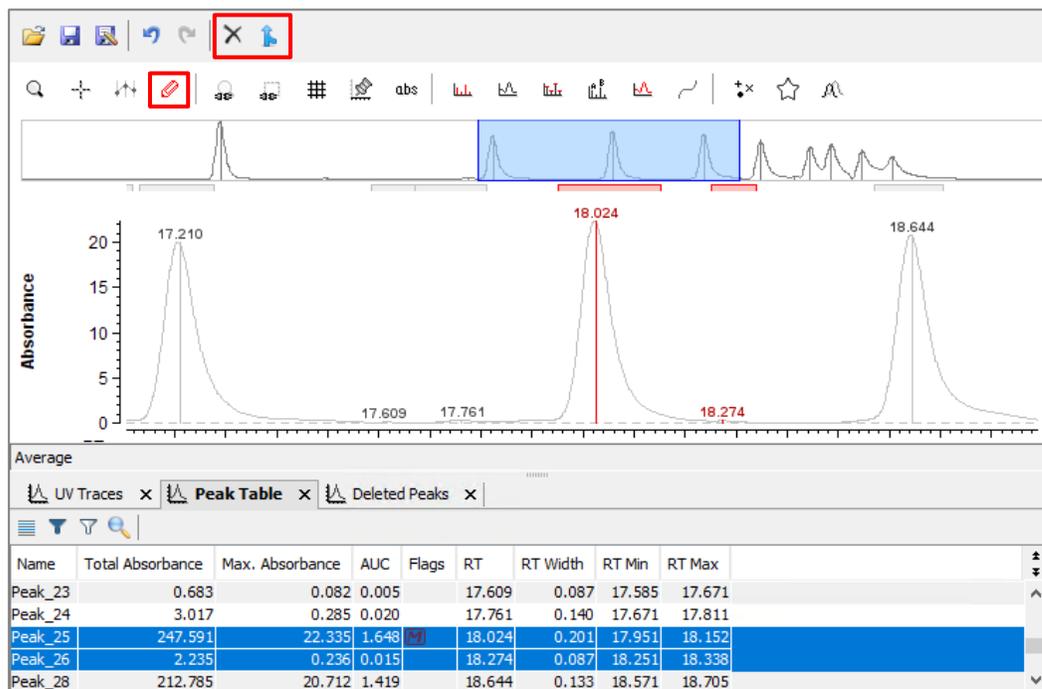
- If there is no UV data:
  1. Select all of the activity nodes in the *UV Processing [Container]*.
  2. Right-click, and then select **Control** > **Bypass Selection**.



# UV Manual Peak Edit

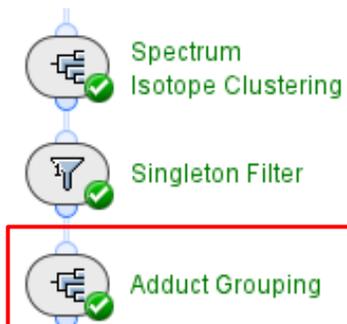


- To manually change the peaks that were detected in the UV chromatogram, use *UV Manual Peak Edit*.



- For complex separations, it is recommended to use *UV Manual Peak Edit* to optimize peak detection, and not to change the *UV Peak Detection* parameters.
- To use *UV Manual Peak Edit*, deactivate the **Bypass** icon.
- Select the **Edit Mode** icon  to:
  - Move the peak boundaries.
  - Merge selected peaks into a single peak.
  - Delete peaks.
  - Draw new peaks.

# Adduct Grouping



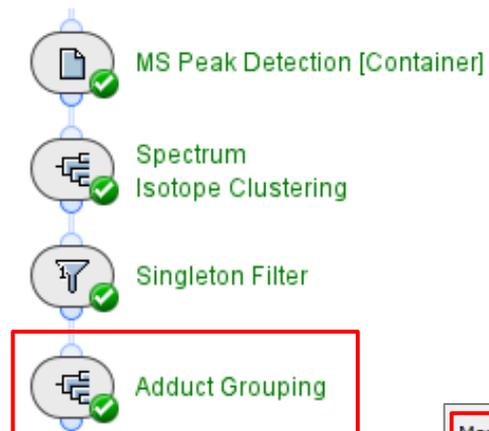
- Use *Adduct Grouping* to group isotopic clusters together that have the same neutral mass and RT.
  - Use the + to select from the list of available adducts.
  - Select the applicable **Mass Mode**:
    - For data with isotopically resolved peaks, or to use **Mass**, select **Monoisotopic**.
    - For data with peaks that are not isotopically resolved, or to use **Avg. Mass**, select **Average**.

The screenshot shows the 'Adduct Grouping - Settings' dialog box. The 'General' tab is selected. The 'RT Tolerance' is set to 0.1 Minutes. The 'mass Tolerance' is set to 20.0 ppm. The 'Allowed Adducts' list contains 'K+' and 'Na+', with a '+' button next to it. The 'Grouping Stringency' is set to 'Relaxed'. The 'Gap Size' is set to 0. The 'Detect Multimers' checkbox is unchecked. The 'Merge Charge and Adduct Groups' checkbox is checked. The 'Mass Mode' section has 'Monoisotopic' selected. The 'OK', 'Cancel', and 'Apply' buttons are at the bottom.

The screenshot shows the 'Select Entries' dialog box. It has a search bar labeled 'Enter Filter Text'. Below the search bar is a list of adducts: TFA, NH4Cl, HFIP, TEA, and -GlcNAc. Each adduct has a checkbox and a star icon to its right. The 'OK' and 'Cancel' buttons are at the bottom.

Note: For information about analysis of data with peaks that are not isotopically resolved, refer to the section: [D1: Guidelines for Specific Applications > Recommended Settings for Isotopically Nonresolved Data](#).

# Adduct Grouping: Merge Charge and Adduct Groups



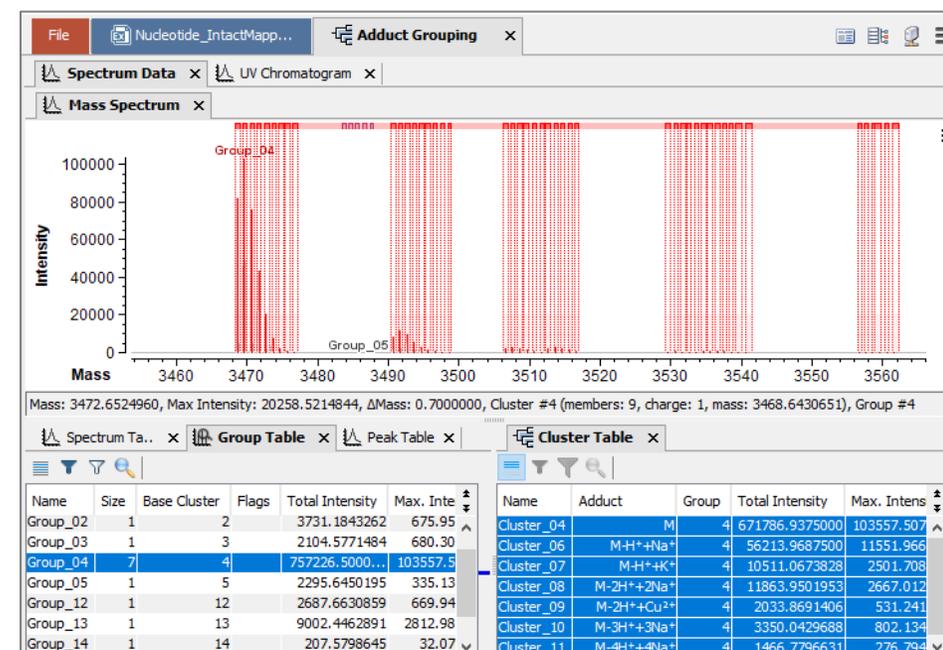
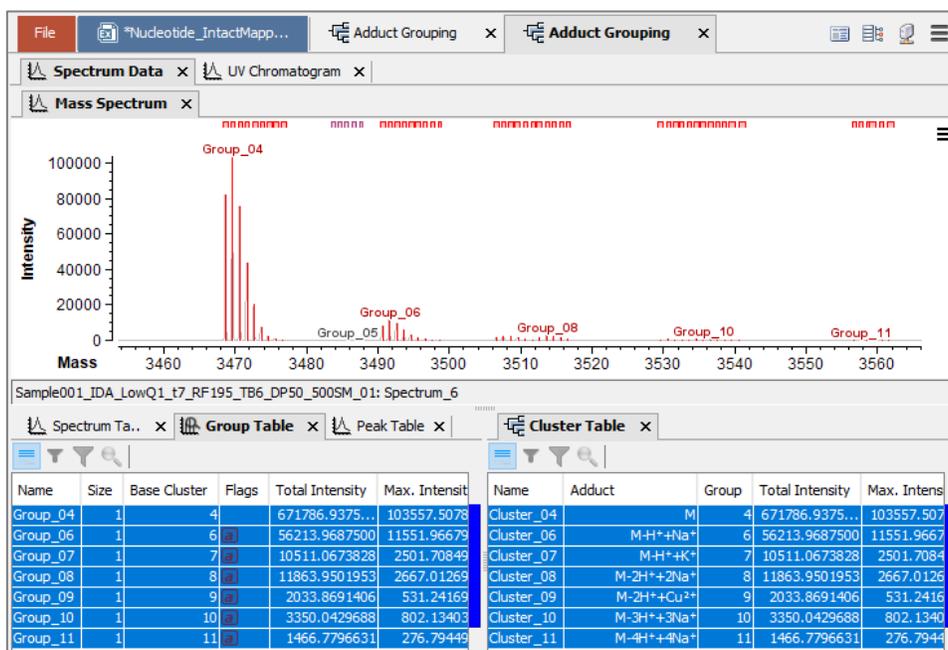
- Select **Merge Charge and Adduct Groups** to group isotopic clusters together into a single group for *Mass Mapping*.
- Do not select **Merge Charge and Adduct Groups** to identify adducts as separate groups for *Mass Mapping*.

Merge Charge and Adduct Groups:

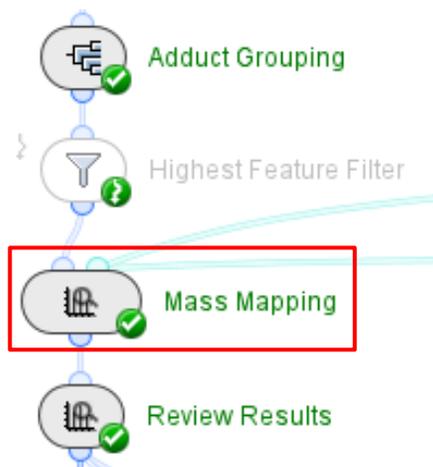
Mass Mode:  Monoisotopic  Average

Merge Charge and Adduct Groups:

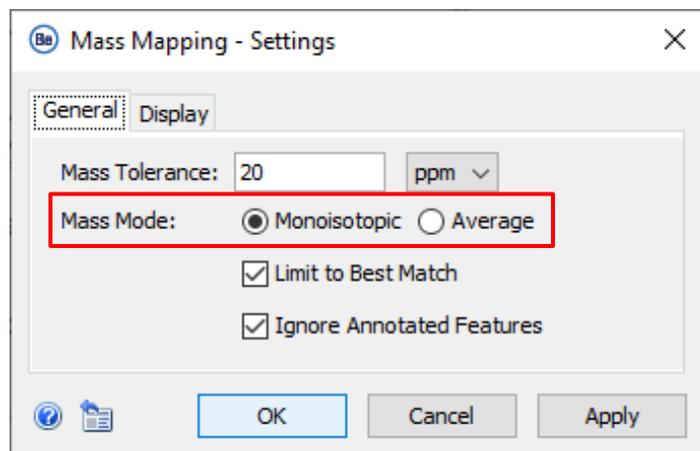
Mass Mode:  Monoisotopic  Average



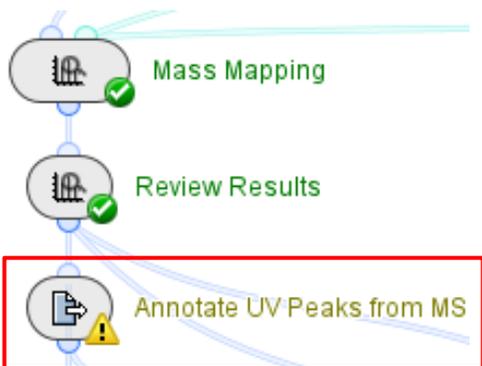
# Mass Mapping



- Use *Mass Mapping* to annotate MS features with candidates created in the Nucleotide Candidate Generation activity nodes, or with candidates from the *Mass Library Import*.
- Select the applicable **Mass Mode**:
  - For data with isotopically resolved peaks, or to use **Mass**, select **Monoisotopic**.
  - For data with peaks that are not isotopically resolved, or to use **Avg. Mass**, select **Average**.
- **Limit to Best Match**: Select to use the nucleotide candidate with the lowest mass delta to the detected mass for annotation.
  - To see all annotations within the specified **Mass Tolerance**, do not select **Limit to Best Match**.



# Annotate UV Peaks from MS



- This activity node uses MS peak information to annotate the related peaks in the **UV Chromatogram**, and then calculate the relative UV and MS abundances.
  - A related peak must elute in the specified **RT Tolerance**.

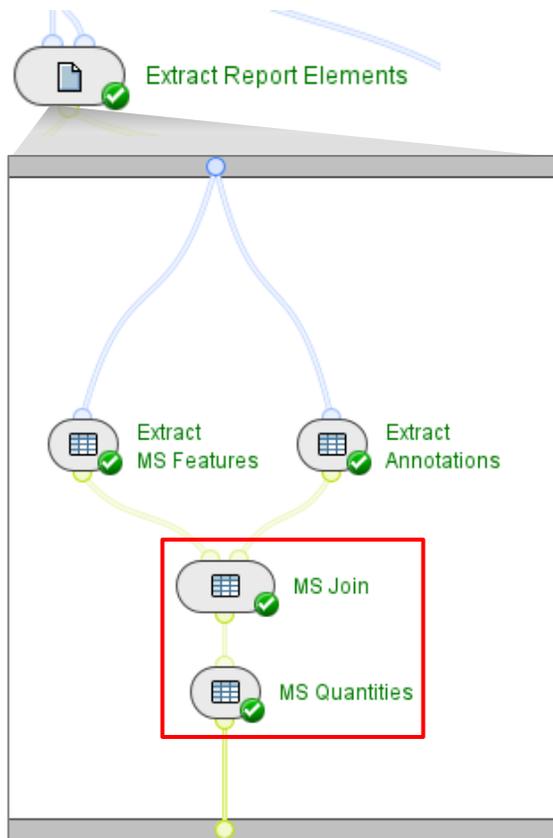
Note: If a UV peak has more than one MS annotation, then the activity node shows a **yellow warning**.

Peak	RT	Annotations	UV Absorbance [AUC]	Relative UV Absorbance
Peak_1	19.15	1	2.07	22.61 %
Peak_2	22.22	3	1.90	20.84 %
Peak_3	23.28	1	1.83	20.0 %
Peak_4	24.15	2	1.76	19.23 %
Peak_5	24.77	1	1.59	17.24 %

Data	UV Quantities	UV Annotations	Summary	UV Peak	RT	Identifier	Type	UV AUC	UV Relative...	Adjusted UV...	MS Relative Intensity...	UV Peak Annotated
				Peak_1	19.15	Impurity	Full-Length	2.07	22.61 %	22.61 %	100.0 %	100.0 %
				Peak_2	22.22	Target [1-46]	3' Clip+Linker	1.90	20.84 %	0.12 %	0.57 %	100.0 %
				Peak_2	22.22	Target [1-59]	3' Clip	1.90	20.84 %	20.13 %	96.58 %	100.0 %
				Peak_2	22.22	Target [60-119]	5' Clip+Linker	1.90	20.84 %	0.59 %	2.85 %	100.0 %
				Peak_3	23.28	Unknown	Full-Length	1.83	20.0 %	20.0 %	100.0 %	100.0 %
				Peak_4	24.15	Unknown [5-79]	5' Clip	1.76	19.23 %	0.06 %	0.32 %	100.0 %
				Peak_4	24.15	Target [1-99]	3' Clip	1.76	19.23 %	19.17 %	99.68 %	100.0 %
				Peak_5	24.77	Target [24-119]	5' Clip+Linker	1.59	17.24 %	0.07 %	0.42 %	100.0 %

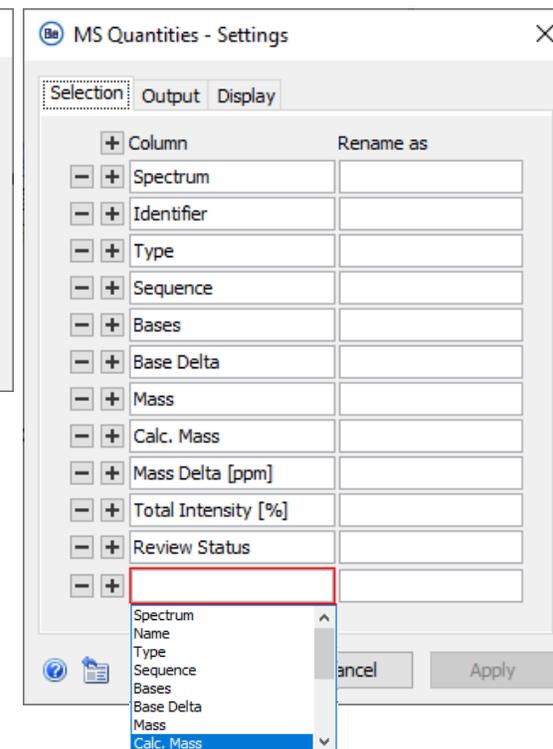
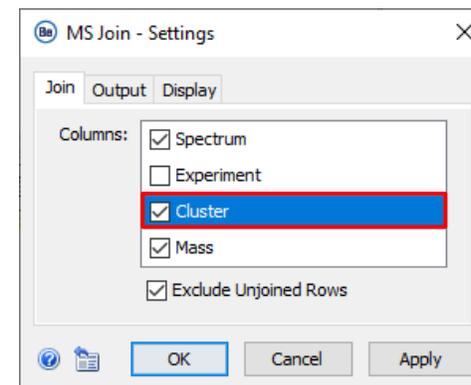
- Normalize relative to:**
  - All Peaks:** Relative UV absorbance is calculated across all detected peaks.
  - Annotated Peaks:** Relative UV absorbance is calculated across annotated peaks.
    - MS Relative Intensity is calculated across annotated peaks.
- Annotation Report Mode: Manual:**
  - Select the information about the annotated features that is included in the result table.

# Extract Report Elements



- Use the *Extract Report Elements* to customize the tables that will be included in the PDF report.
  - To see the columns that are available for selection in *MS Join* or *MS Quantities*, run the activity nodes that are immediately before them.

- *MS Join*: If **Bypass** is activated for *Adduct Grouping*, then select **Cluster**.
- *MS Quantities*: Select the columns of interest for the table in the report.



Note: If a selected column is empty, then the activity node shows a **yellow warning**. For example, if **Review Status** is selected, but there are no accepted identifications.

Elapsed Time	4 msec
Status	Suspicious
Message	The following expected columns are absent or empty: Review Status
Summary	

---

## Part C

# Guidelines for the Intact Nucleotide Workflows

# C

---

Intact Nucleotide with Deconvolution  
Template Workflow Guidelines

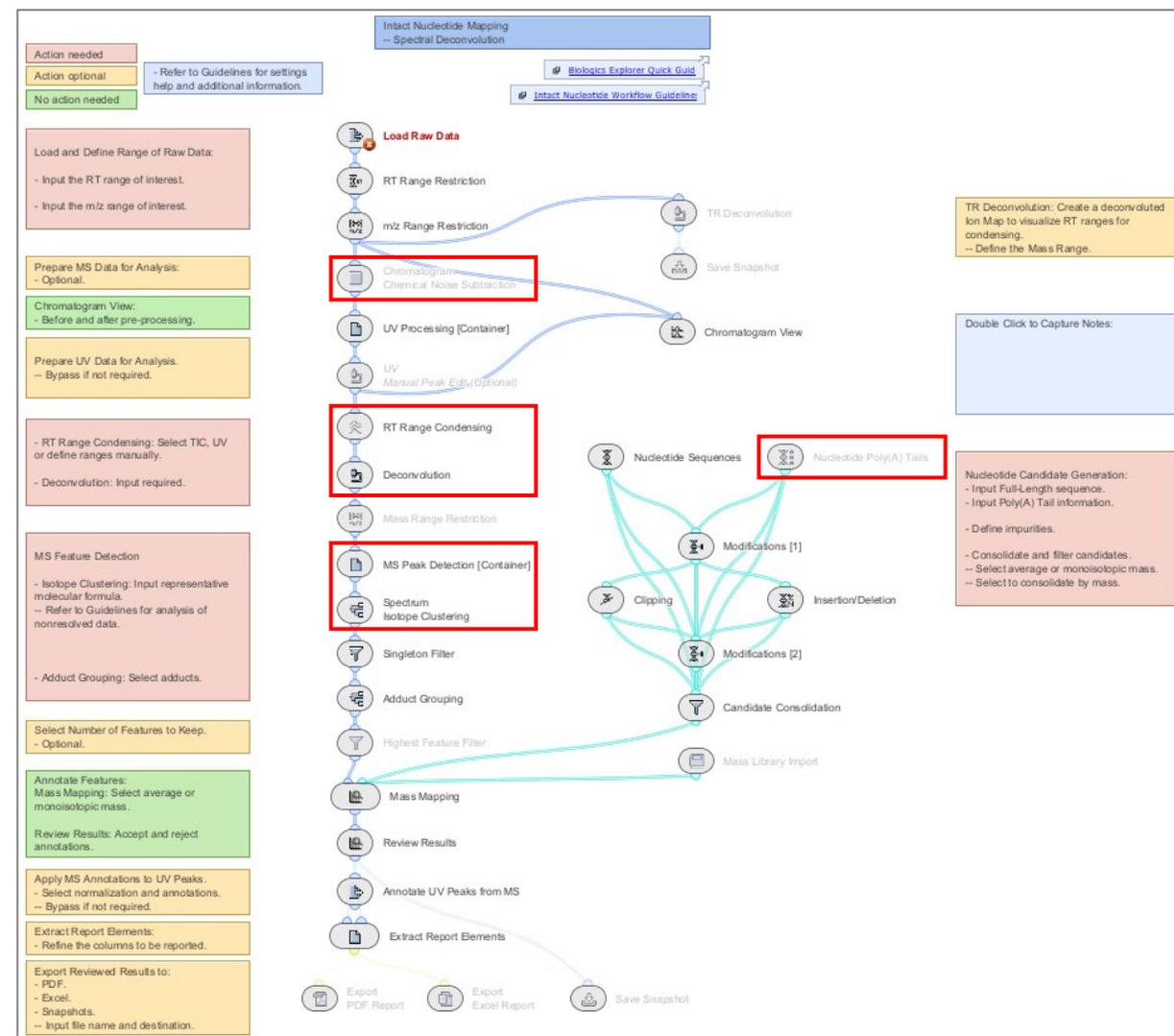
C1

# Intact Nucleotide with Deconvolution Template Workflow

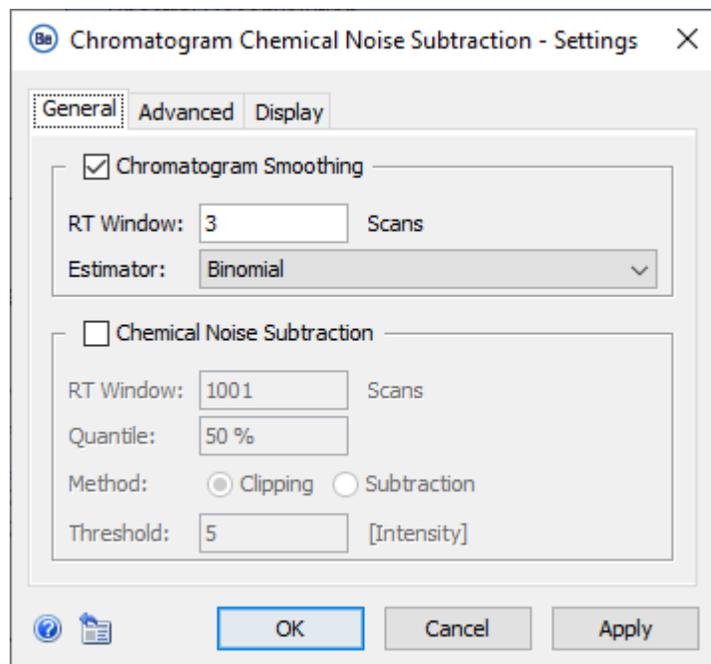
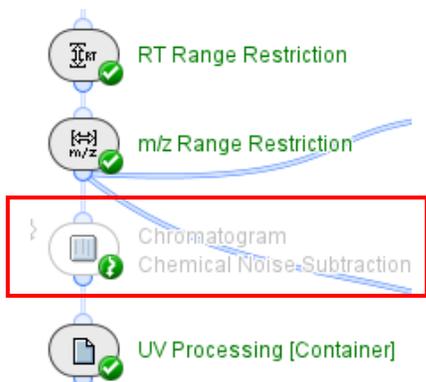
- This section contains information about these activity nodes of interest:

- *Chromatogram Chemical Noise Subtraction*
- *Nucleotide Poly(A) Tails*
- *RT Range Condensing*
- *Deconvolution*
- *Spectrum Peak Detection*
- *Spectrum Isotope Clustering*

Note: For information about activity nodes that are used in all workflows, for example *Load Raw Data*, *Review Results*, or *Export PDF Report*, refer to the document: [Biologics Explorer Quick Guide](#).



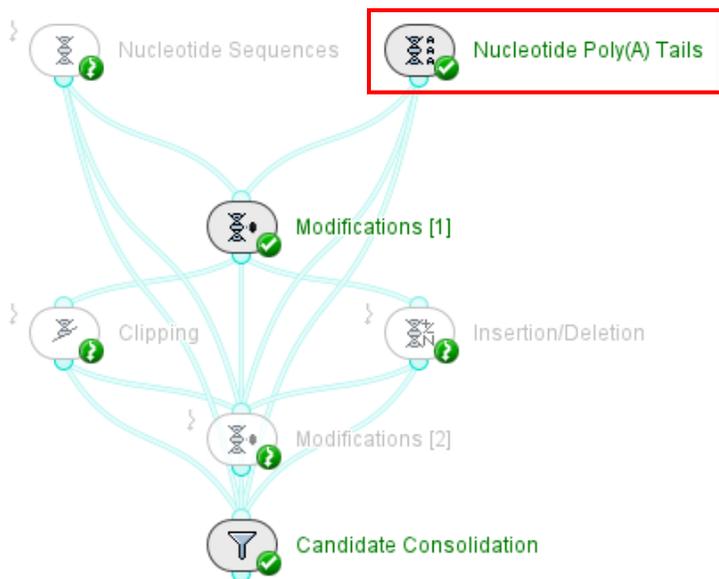
# Chromatogram Chemical Noise Subtraction: Optional



- To use *Chromatogram Chemical Noise Subtraction*, deactivate the **Bypass** icon.
  - **Chromatogram Smoothing** is used to improve the RT profile of peaks for peak detection.
  - **Chemical Noise Subtraction** should be used with care, and only when very high background noise has an effect on the quality of the deconvolution.

Note: For more information, click the ? icon to open the [Online Help](#).

# Nucleotide Candidate Generation: *Poly(A) Tails*



- **Poly(A) Tail Candidates:**

- Use the + icon to add the correct number of rows.
- Add a **5' Sequence** or **3' Sequence**, or leave blank if not required.
- Type the minimum and maximum expected number of adenosines in the Poly(A) Tail under investigation.
- Activate the **Bypass** icon on *Clipping* and *Insertion/Deletion*.

- The **Identifier** and the **Poly(A) Bases** columns in the **Nucleotides** Result Table contain the number of adenosines in that candidate.

Nucleotide Poly(A) Tails - Settings

Poly(A) Tail Candidates Display

Name	5' Sequence	Poly(A) Min.	Poly(A) Max.	3' Sequence
- Segment 1	AAGGAGA...	1	50	
- Segment 2		1	30	GAGG

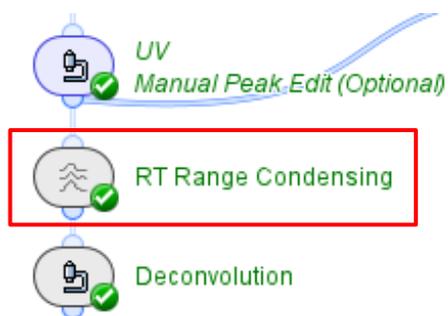
+ OK Cancel Apply

Nucleotides

Identifier	Sequence	Bases	Poly(A) Bases	Type	Input Name
Segment 1 [10*A]	rAAGGAGAAGAGAAGGAAGAGAAAAAAAAA	30	10 Poly(A) Tail		Segment 1
Segment 1 [11*A]	rAAGGAGAAGAGAAGGAAGAGAAAAAAAAA	31	11 Poly(A) Tail		Segment 1
Segment 1 [12*A]	rAAGGAGAAGAGAAGGAAGAGAAAAAAAAA	32	12 Poly(A) Tail		Segment 1
Segment 1 [13*A]	rAAGGAGAAGAGAAGGAAGAGAAAAAAAAA	33	13 Poly(A) Tail		Segment 1
Segment 1 [14*A]	rAAGGAGAAGAGAAGGAAGAGAAAAAAAAA	34	14 Poly(A) Tail		Segment 1
Segment 1 [15*A]	rAAGGAGAAGAGAAGGAAGAGAAAAAAAAA	35	15 Poly(A) Tail		Segment 1
Segment 2 [30*A]	rAAAAAAAAAAAAAAAAAAAAAAAAAAGAGG	34	30 Poly(A) Tail		Segment 2
Segment 2 [29*A]	rAAAAAAAAAAAAAAAAAAAAAAAAAAGAGG	33	29 Poly(A) Tail		Segment 2
Segment 2 [28*A]	rAAAAAAAAAAAAAAAAAAAAAAAAAAGAGG	32	28 Poly(A) Tail		Segment 2
Segment 2 [27*A]	rAAAAAAAAAAAAAAAAAAAAAAAAAAGAGG	31	27 Poly(A) Tail		Segment 2
Segment 2 [26*A]	rAAAAAAAAAAAAAAAAAAAAAAAAAAGAGG	30	26 Poly(A) Tail		Segment 2
Segment 2 [25*A]	rAAAAAAAAAAAAAAAAAAAAAAAAAAGAGG	29	25 Poly(A) Tail		Segment 2

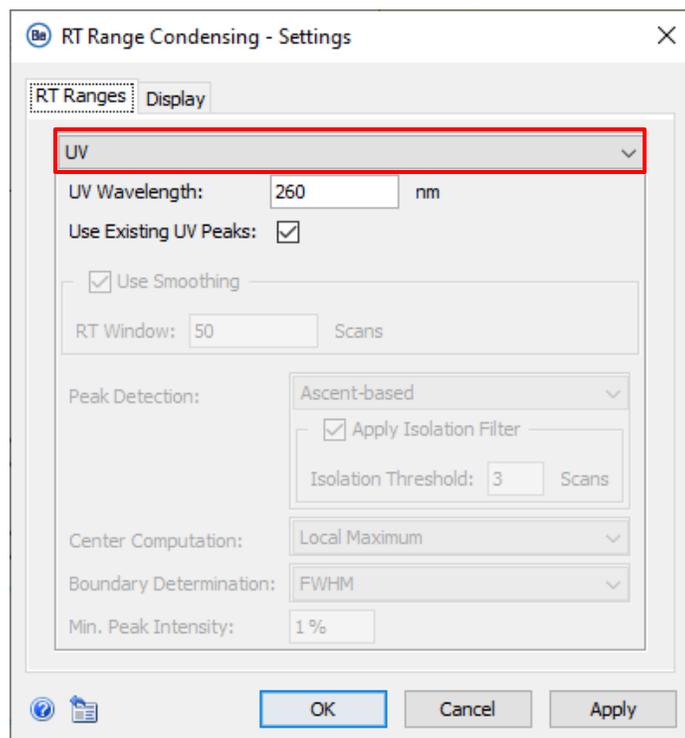


# RT Range Condensing

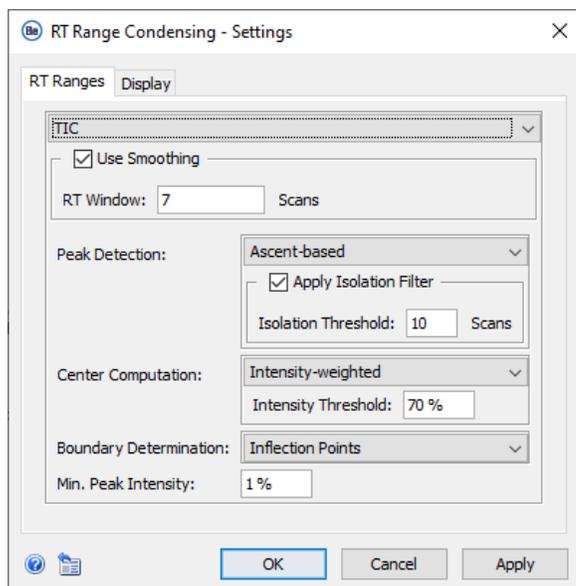
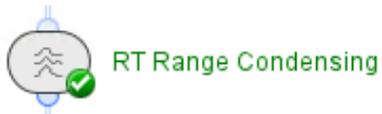


- *RT Range Condensing* detects regions of the ion map that contain signal, and then creates a single spectrum across the RT in these regions.

- Select an option from the list:
  - **TIC:** Uses the peaks in the total ion chromatogram to identify the RT ranges to condense.
  - **UV:** Uses peaks in the UV data to identify the RT ranges to condense.
  - **Manual:** For complex separations, identify the RT ranges to condense manually.



# RT Range Condensing: UV and TIC RT Ranges

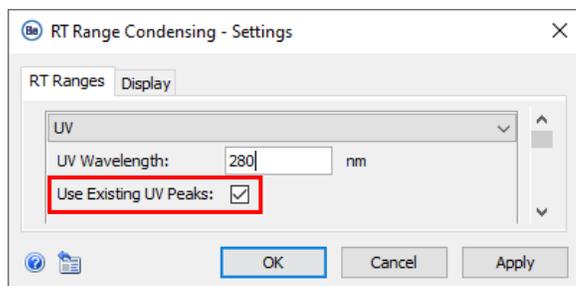


## RT Ranges: TIC

- **Peak Detection:**

- To identify local maxima in the MS signal, select **Ascent-based**.
- To identify changes in the curvature of the MS signal, for example to identify shoulder peaks, select **Curvature-based**.

Note: For more information, click the ? icon to open the [Online Help](#).



## RT Ranges: UV

1. Select the **UV Wavelength**.
  2. Select **Use Existing UV Peaks**.
- If **Use Existing UV Peaks** is selected, then other peak detection settings on this tab are ignored.

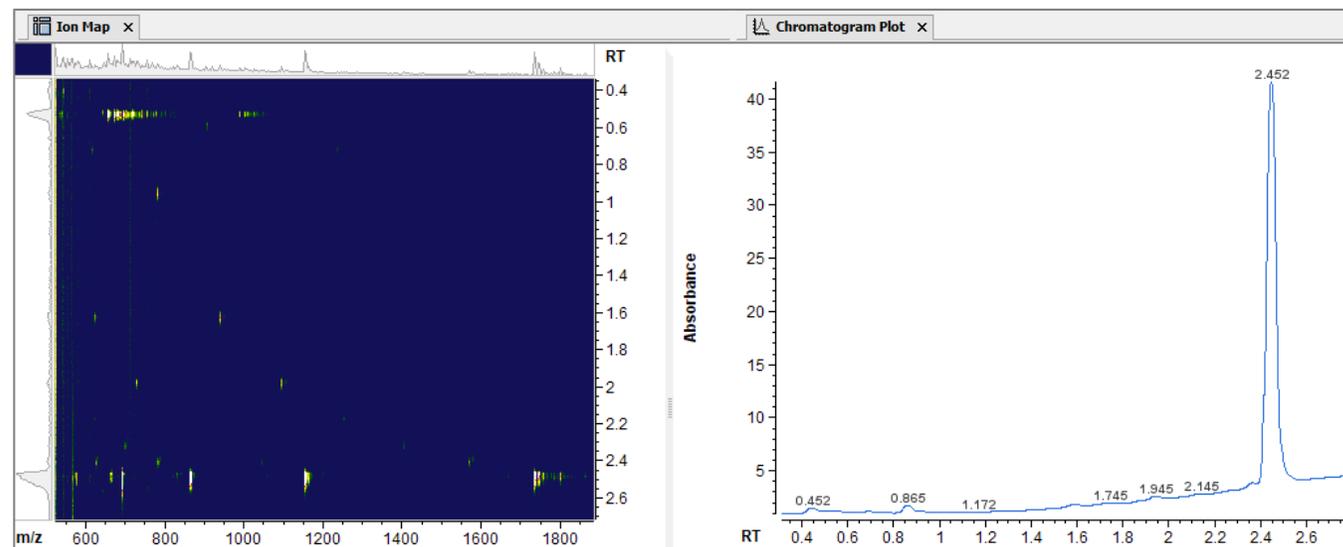
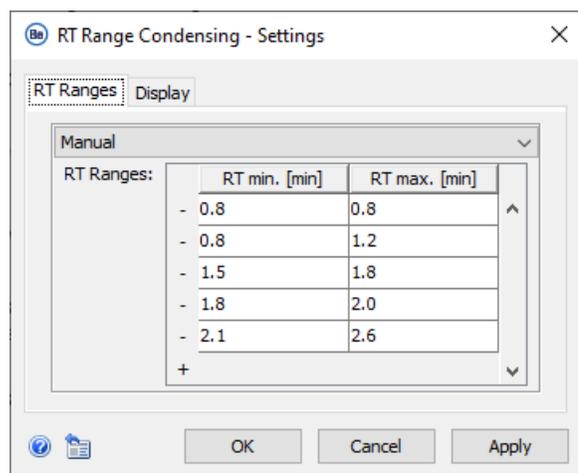
# RT Range Condensing: Manual RT Ranges



RT Range Condensing

## RT Ranges: Manual

- Select RT ranges manually if the components of interest are not chromatographically resolved.
  - For example, if the peaks in the TIC or the UV chromatogram do not show all of the components of interest.

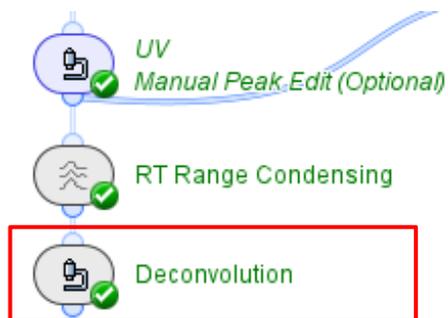


- To use *TR Deconvolution* to help to identify RT ranges of interest:
  1. Deactivate the **Bypass** icon.
  2. Review the results of *TR Deconvolution*.
  3. Type the RT ranges of interest in *RT Range Condensing*.



TR Deconvolution

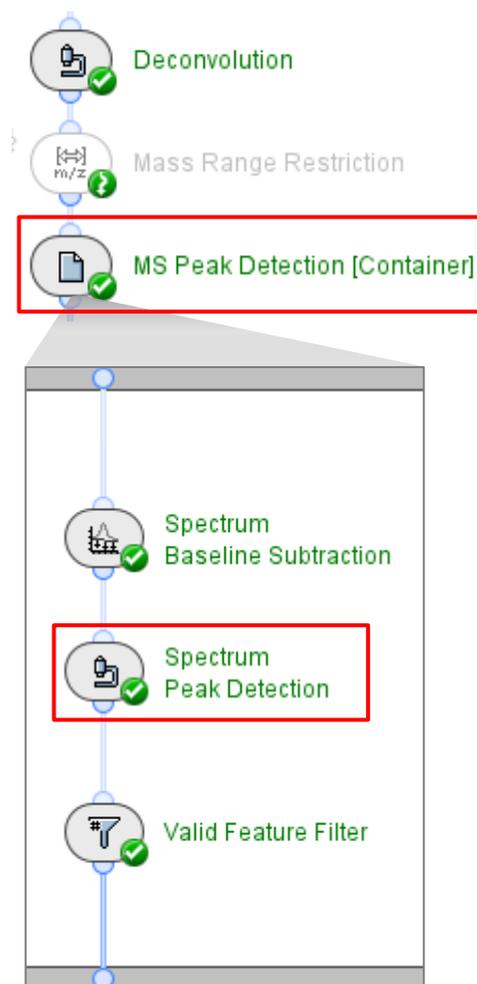
# Deconvolution



- The RT ranges detected in *RT Range Condensing* are deconvoluted.
  - **Deconvolution Quality:**
    - Select **High** for isotopically resolved data.
    - Select **Standard** for lower-resolution data.
  - **Min. Mass and Max. Mass:**
    - Use a wide mass range to decrease the number and intensity of harmonic peaks.
    - It is not recommended to use a **Min. Mass** value that is lower than the maximum  $m/z$  value of the data that will be deconvoluted.
  - **Mass Step:** Set a value that keeps the peak resolution of the data.
    - 0.05 Da to 0.2 Da for isotopically resolved data.
    - 1 Da to 2 Da for lower-resolution data.

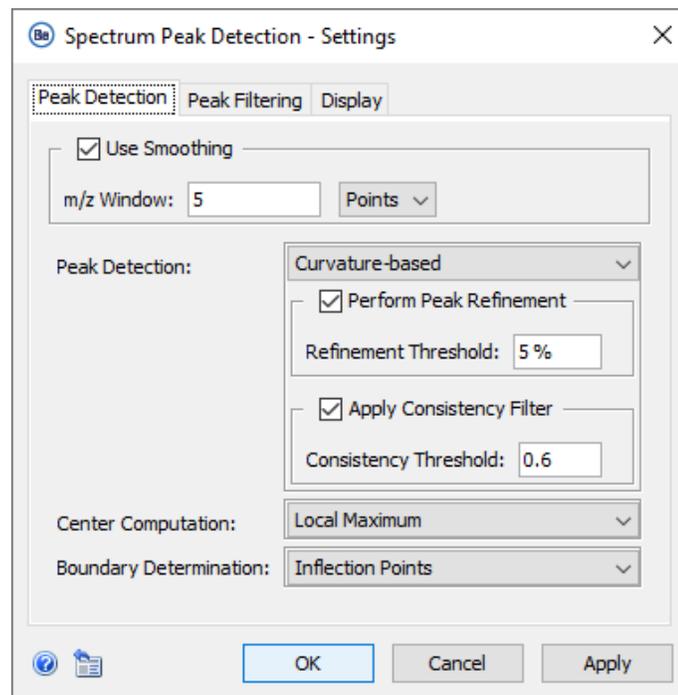
The screenshot shows the 'Deconvolution - Settings' dialog box. It has a 'Deconvolution Options' tab and a 'Display' button. The 'Method' is set to 'Maximum Entropy Deconvolution'. The 'Iterations' field is set to 25, and the 'Deconvolution Quality' is set to 'High'. Under 'Output Mass Spectrum', the 'Min. Mass' is 10 kDa, the 'Max. Mass' is 100 kDa, and the 'Mass Step' is 0.1 Da. The 'Ionization' section has 'Protonation' unselected and 'Deprotonation' selected. At the bottom, there are 'OK', 'Cancel', and 'Apply' buttons.

# MS Peak Detection [Container]: Spectrum Peak Detection



- The default *Spectrum Peak Detection* settings are applicable for most data with isotopically resolved peaks.

Note: For information about analysis of data with peaks that are not isotopically resolved, refer to the section: [D1: Guidelines for Specific Applications > Large Intact Nucleotides \(Nonresolved Data\)](#).

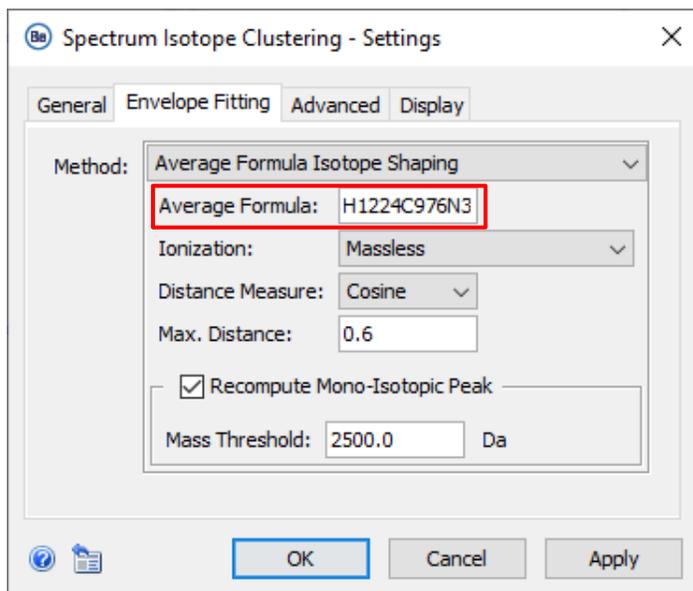
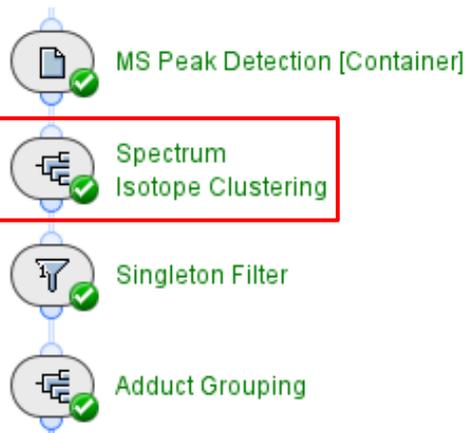


## — Peak Detection:

- To identify local maxima in the MS signal, select **Ascent-based**.
- Increase the **Isolation Threshold** to increase the minimum distance between local maxima for a peak to be detected.
- To identify changes in the curvature of the MS signal, for example to identify shoulder peaks in the  $m/z$  direction, select **Curvature-based**.
- Decrease the **Refinement Threshold** to increase the split sensitivity.
- Increase the **Consistency Threshold** to decrease the split sensitivity.

Note: For more information, click the ? icon to open the [Online Help](#).

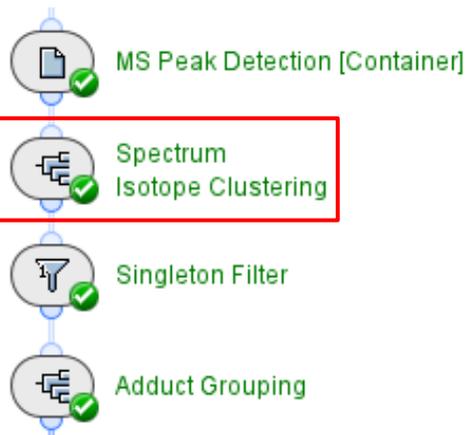
# Spectrum Isotope Clustering



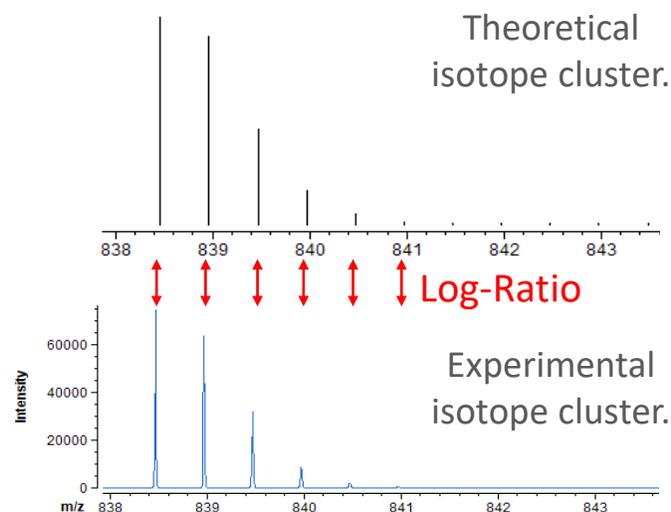
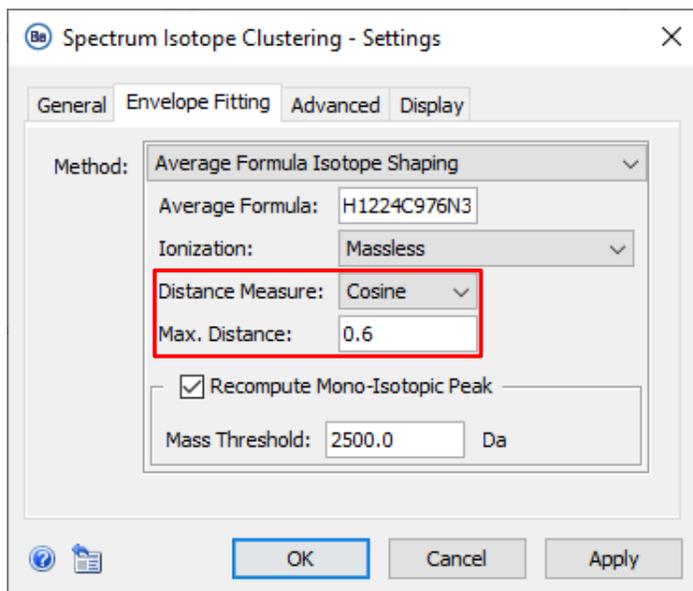
- Use *Spectrum Isotope Clustering* to group peaks together in an isotopic envelope.
  - **Average Formula:** Type a representative molecular formula for the nucleotide under investigation.
    - The chemical formula from the **Mass Calculator** can be copied and pasted here.
  - **Ionization:** Select Massless for deconvoluted data.
- *Spectrum Isotope Clustering* is required for *Adduct Grouping* with isotopically resolved and nonisotopically resolved data.
- Use *Singleton Filter* with isotopically resolved data to remove peaks that are not clustered.

Note: For information about analysis of data with peaks that are not isotopically resolved, refer to the section: [D1: Guidelines for Specific Applications > Recommended Settings for Isotopically Nonresolved Data](#).

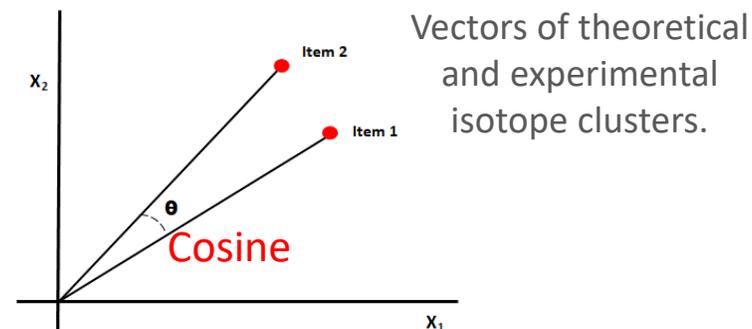
# Spectrum Isotope Clustering: Distance Measure



- The **Distance Measure** compares each experimental peak to the theoretical isotope profile for the representative molecular formula in **Average Formula**.
  - If a peak of interest has not been clustered as required, then compare the results of **Log Ratio** and **Cosine**:



- Log Ratio** treats all peaks in a cluster equally.



- Cosine** converts each cluster into a vector, with the contribution of each peak relative to the abundance. Therefore, smaller peaks have less impact.

---

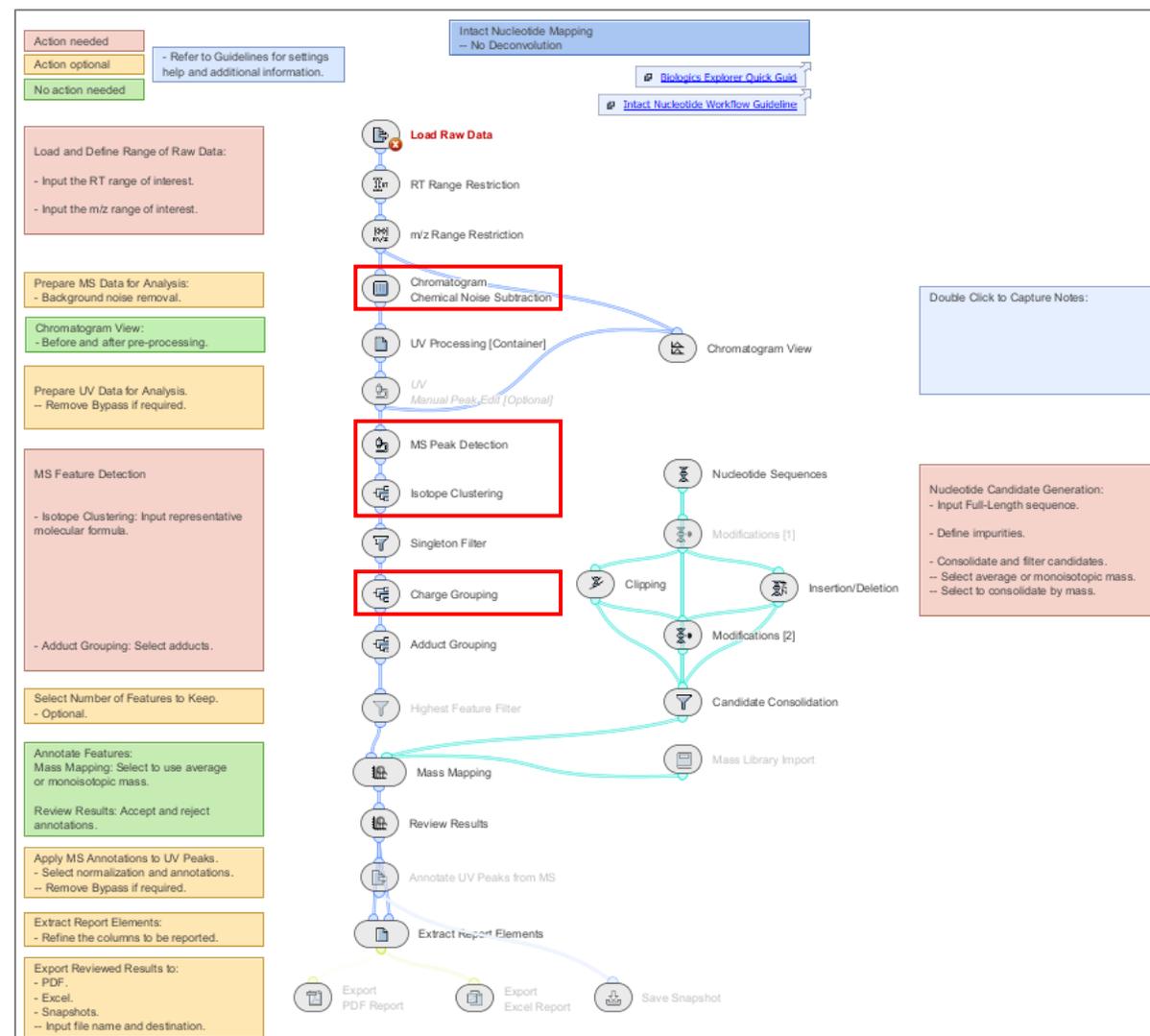
Intact Nucleotide with no Deconvolution  
Template Workflow Guidelines

C2

# Intact Nucleotide with No Deconvolution Template Workflow

- This section contains information about these activity nodes of interest:
  - Chromatogram Chemical Noise Subtraction*
  - MS Peak Detection*
  - Isotope Clustering*
  - Charge Grouping*

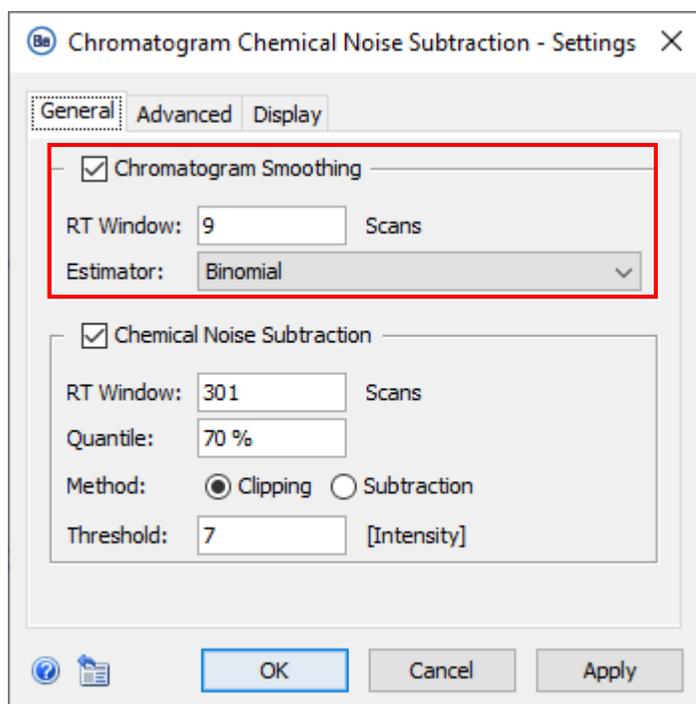
Note: For information about activity nodes that are used in all workflows, for example *Load Raw Data*, *Review Results*, or *Export PDF Report*, refer to the document: [Biologics Explorer Quick Guide](#).



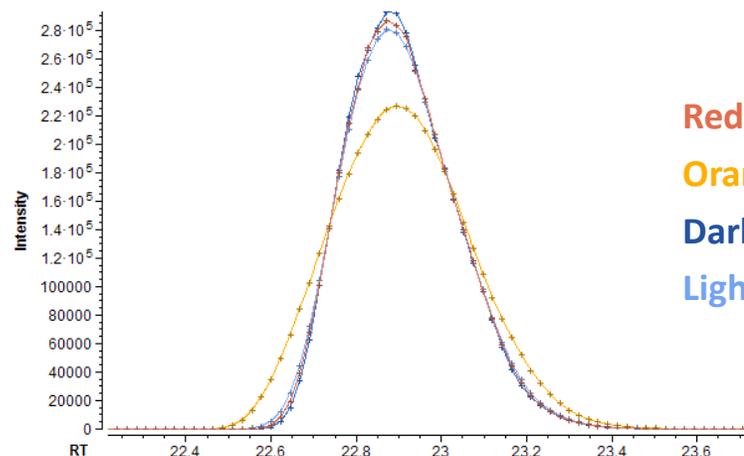
# Chromatogram Chemical Noise Subtraction: Smoothing



**Chromatogram Smoothing** is used to improve the RT profile of peaks for peak detection.



- **Estimator:**
  - **Moving Average** replaces the intensity of each data point with the mean average intensity of the data points in the **RT Window**. High values cause peak widths to increase, but peak volume is not changed.
  - **Binomial** is an iterative form of **Moving Average** that has less effect on peak widths at high scan values.



Red: Moving Average (5 scans)  
 Orange: Moving Average (15 scans)  
 Dark Blue: Binomial (5 scans)  
 Light Blue: Binomial (15 scans)

# Chromatogram Chemical Noise Subtraction: RT Window & Quantile



Chromatogram Chemical Noise Subtraction - Settings

General Advanced Display

Chromatogram Smoothing

RT Window: 9 Scans

Estimator: Binomial

Chemical Noise Subtraction

RT Window: 301 Scans

Quantile: 70 %

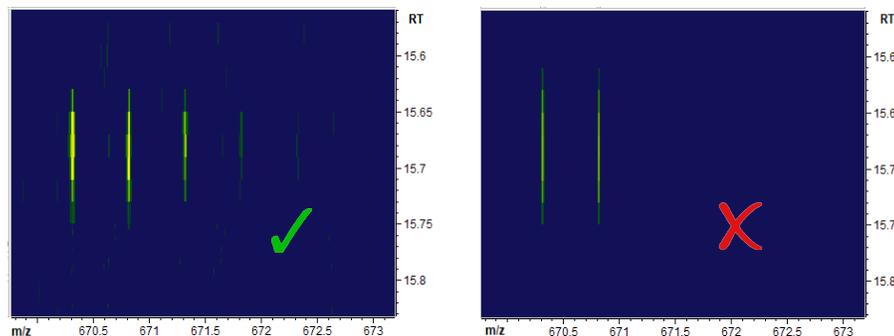
Method:  Clipping  Subtraction

Threshold: 7 [Intensity]

OK Cancel Apply

**Chemical Noise Subtraction** decreases the length of long-tailing peaks.

- Change this setting if the default values remove too much signal.
- If too much signal is removed, it can be identified by:
  - Excessive cutoff of the tails of very wide (extended RT) peaks.
  - Loss of low-intensity isotope peaks from singly (+1) or doubly (+2) charged clusters, or from low-intensity clusters of interest:



- To decrease the amount of noise removal (keep more signal):
  - Decrease the **Quantile**.
  - Increase the **RT Window**.

# Chromatogram Chemical Noise Subtraction: Threshold



Chromatogram Chemical Noise Subtraction - Settings

General Advanced Display

Chromatogram Smoothing

RT Window: 9 Scans

Estimator: Binomial

Chemical Noise Subtraction

RT Window: 301 Scans

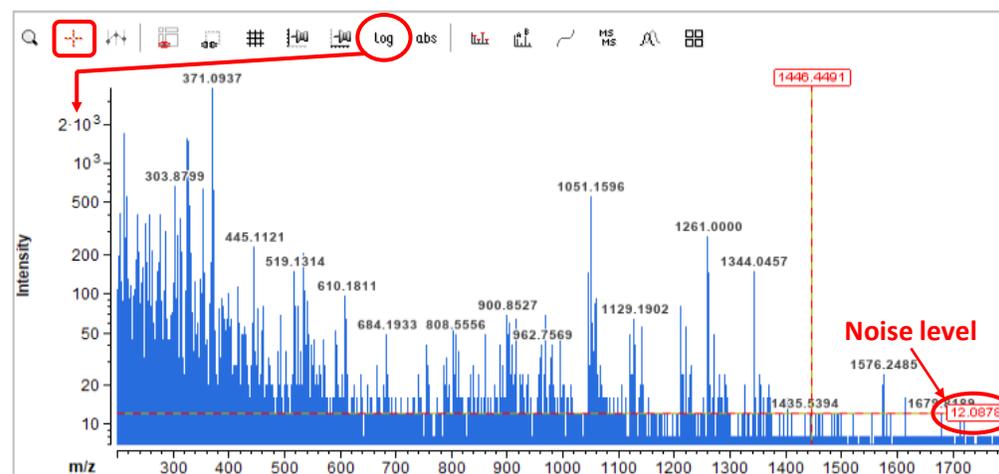
Quantile: 70 %

Method:  Clipping  Subtraction

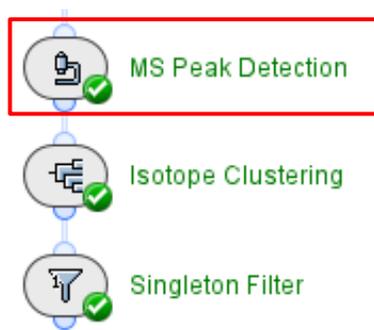
Threshold: 7 [Intensity]

OK Cancel Apply

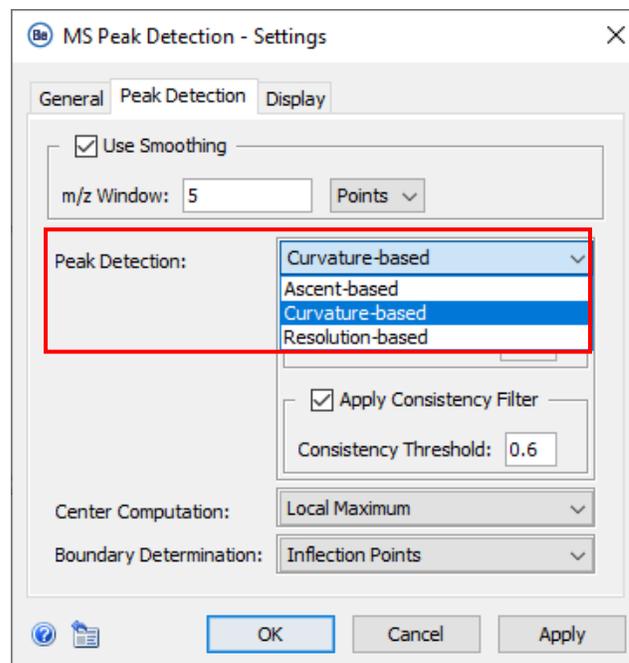
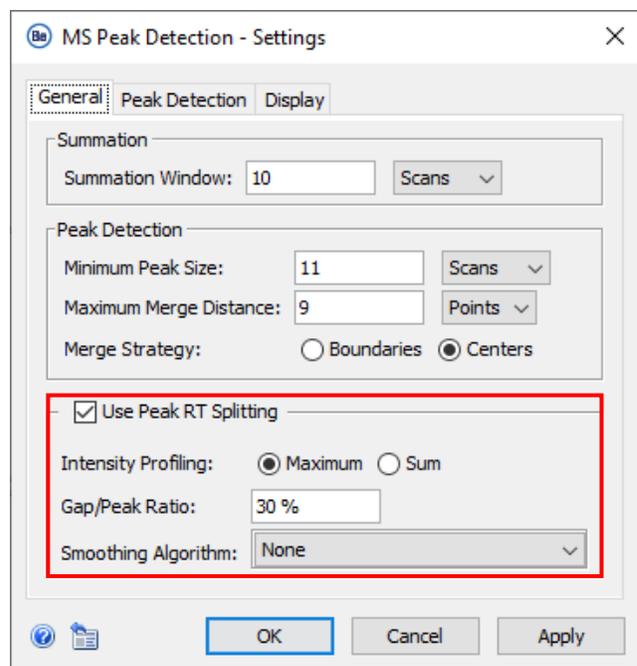
- If the noise level is significantly different from the **Threshold** value in *Chromatogram Chemical Noise Subtraction*, then change this setting.
- To measure the noise level and identify an applicable **Threshold** intensity value:
  - Drag the intensity axis of the mass spectrum until the noise level can be seen, or use the icon in the tool bar to change the axis from the linear to the logarithmic scale.
  - Use the crosshair tool  to measure the intensity of the noise level.



# MS Peak Detection



- The default *MS Peak Detection* settings are applicable for most data that has not been deconvoluted.
  - Use Peak RT Splitting:** To increase the number of peaks detected (increase split sensitivity of shoulder peaks) in the RT direction:
    - Decrease the **Gap/Peak Ratio**.
    - Decrease or remove **Smoothing**.

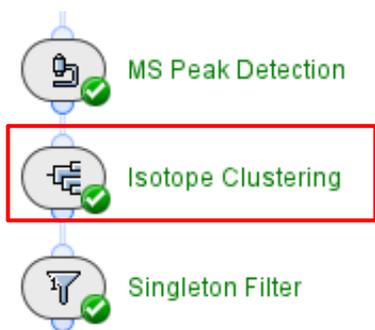


## Peak Detection:

- To identify local maxima in the MS signal, select **Ascent-based**.
- Increase the **Isolation Threshold** to increase the minimum distance between local maxima for a peak to be detected.
- To identify changes in the curvature of the MS signal, for example to identify shoulder peaks in the *m/z* direction, select **Curvature-based**.
  - Decrease the **Refinement Threshold** to increase the split sensitivity.
  - Increase the **Consistency Threshold** to decrease the split sensitivity.

Note: For more information, click the ? icon to open the [Online Help](#).

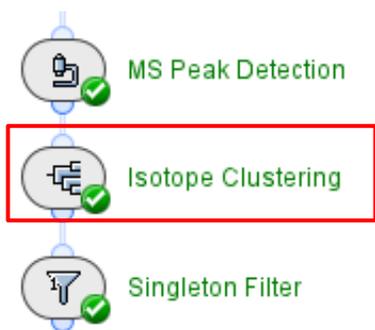
# Isotope Clustering



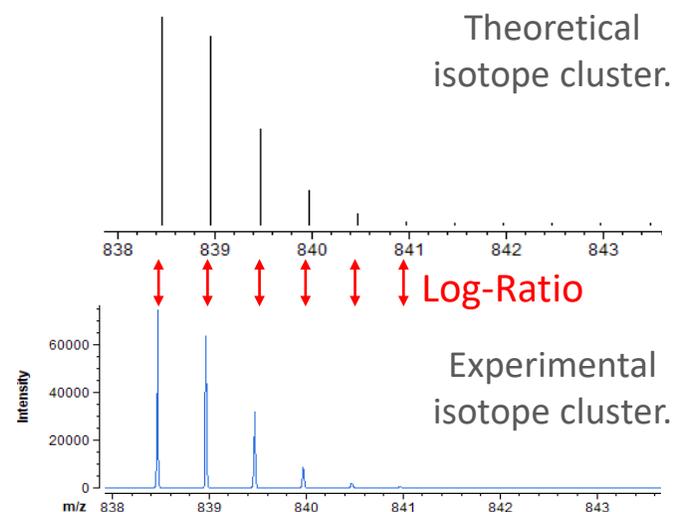
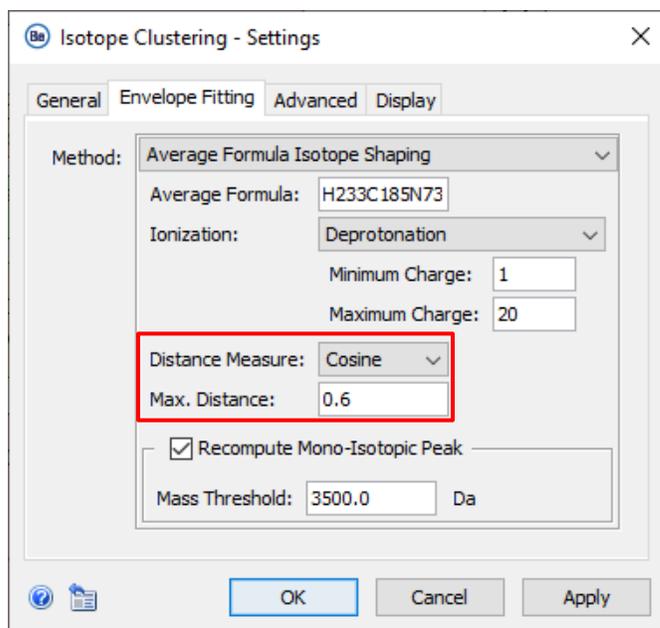
- Use *Spectrum Isotope Clustering* to group peaks together in an isotopic envelope.
- *Isotope Clustering* is required for *Charge Grouping* and *Adduct Grouping*.
  - **Average Formula:** Type a representative molecular formula for the nucleotide under investigation.
    - The chemical formula from the **Mass Calculator** can be copied and pasted here.
  - **Ionization:** Select **Deprotonation** for data that was acquired in negative ion mode.

The screenshot shows the 'Isotope Clustering - Settings' dialog box with the 'General' tab selected. The 'Method' dropdown is set to 'Average Formula Isotope Shaping'. The 'Average Formula' field contains the text 'H233C185N73', which is highlighted with a red box. The 'Ionization' dropdown is set to 'Deprotonation'. Below it, 'Minimum Charge' is set to 1 and 'Maximum Charge' is set to 20. The 'Distance Measure' is set to 'Cosine' and 'Max. Distance' is set to 0.6. The 'Recompute Mono-Isotopic Peak' checkbox is checked. The 'Mass Threshold' is set to 3500.0 Da. At the bottom, there are 'OK', 'Cancel', and 'Apply' buttons.

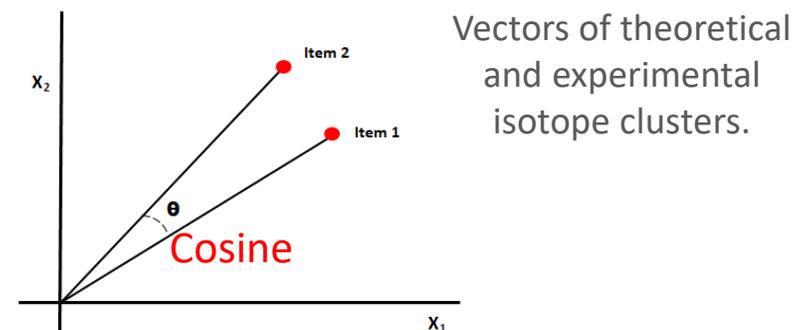
# Isotope Clustering: Distance Measure



- The **Distance Measure** compares each experimental peak to the theoretical isotope profile for the representative molecular formula in **Average Formula**.
  - If a peak of interest has not been clustered as required, then compare the results of **Log Ratio** and **Cosine**:

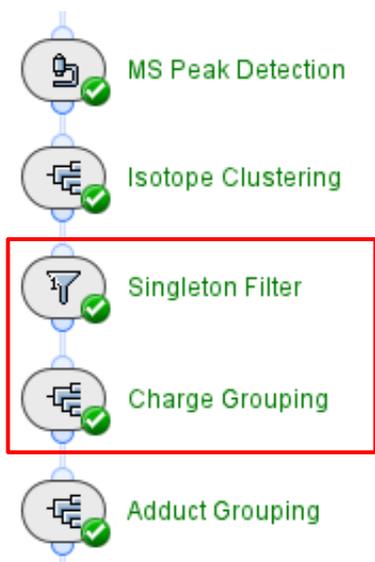


- **Log Ratio** treats all peaks in a cluster equally.

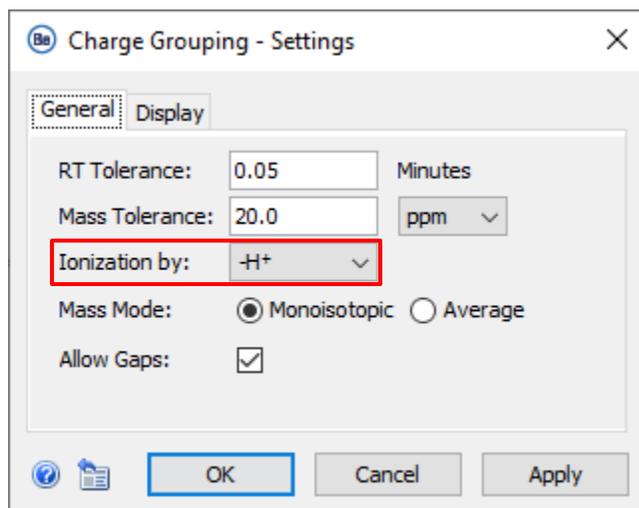


- **Cosine** converts each cluster into a vector, with the contribution of each peak relative to the abundance. Therefore, smaller peaks have less impact.

# Singleton Filter and Charge Grouping



- Use *Singleton Filter* to remove peaks that are not clustered.



- Use *Charge Grouping* to group related clusters with different charge states together.
  - **Ionization by:** Select -H<sup>+</sup> for data that was acquired in negative ion mode.

---

**Part D**  
Guidelines for Specific Applications

**D**

---

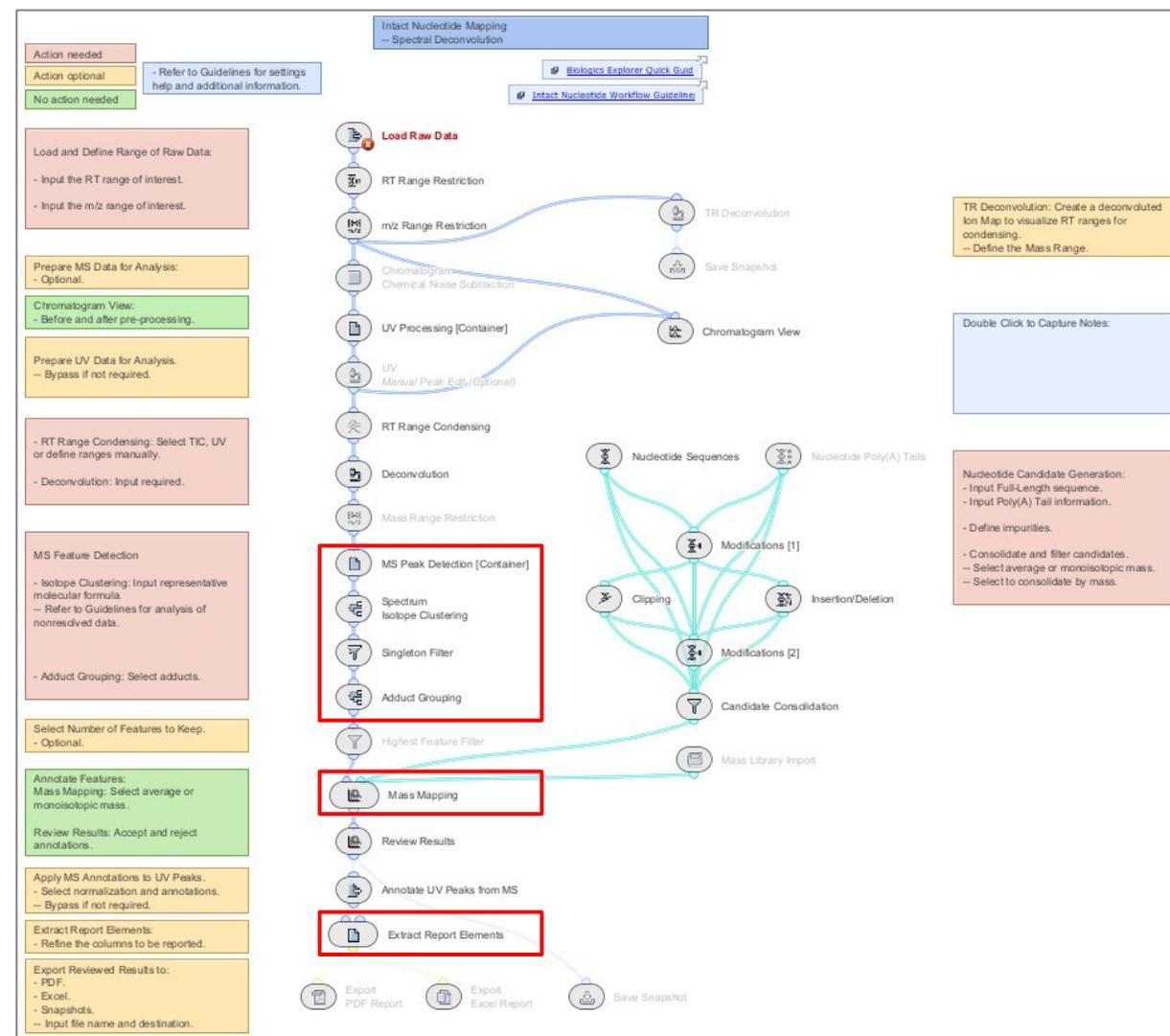
Recommended Settings for Isotopically Nonresolved Data  
Application Specific Information

D1

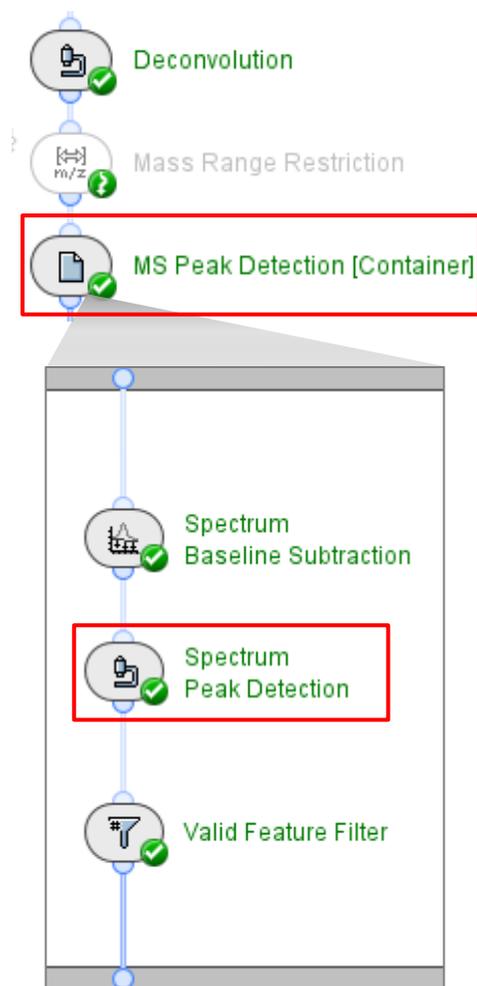
# Nucleotide Deconvolution Workflow for Isotopically Nonresolved Data

- This section contains additional information about these activity nodes:

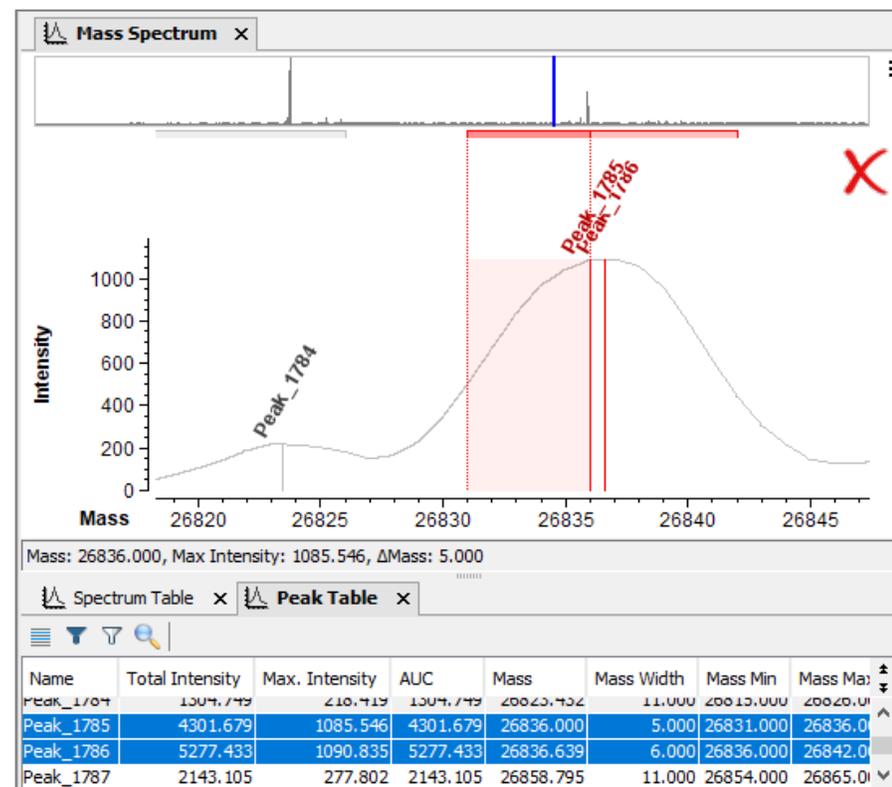
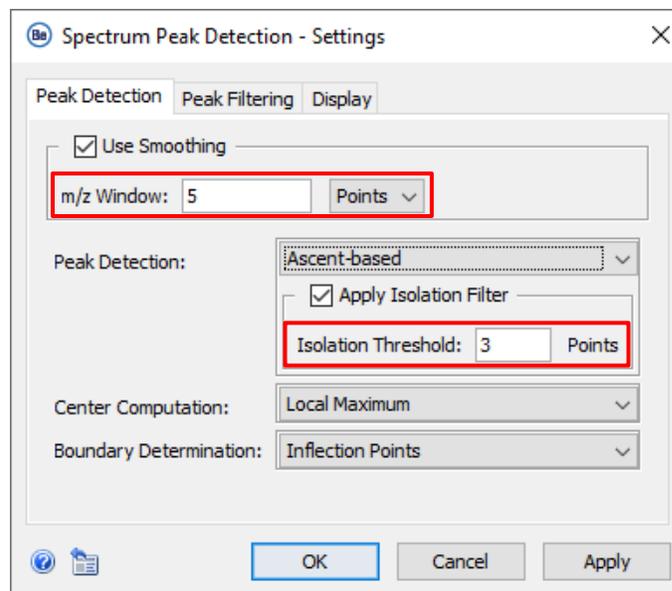
- Spectrum Peak Detection*
- Spectrum Isotope Clustering*
- Singleton Filter*
- Adduct Grouping*
- Mass Mapping*
- Export Report Elements*



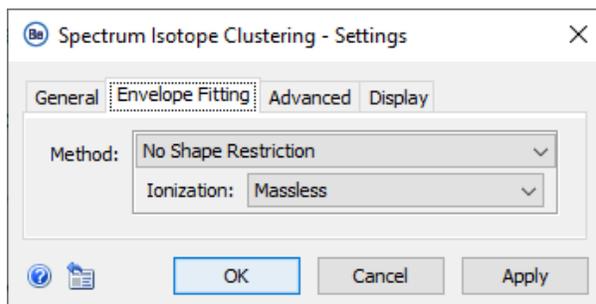
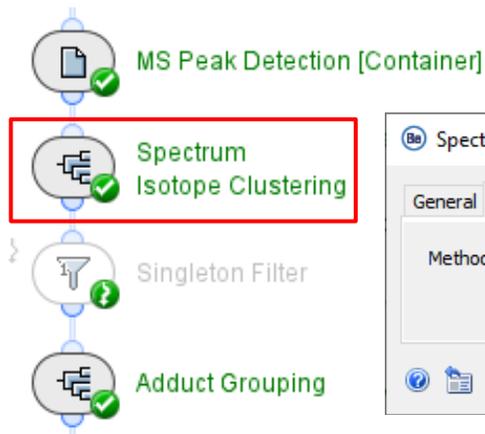
# MS Peak Detection [Container]: Spectrum Peak Detection



- Use **Peak Detection: Ascent-based** for data with peaks that are not isotopically resolved.
  - To remove unwanted detection of peak shoulders, increase the **Smoothing** or **Isolation Threshold**.

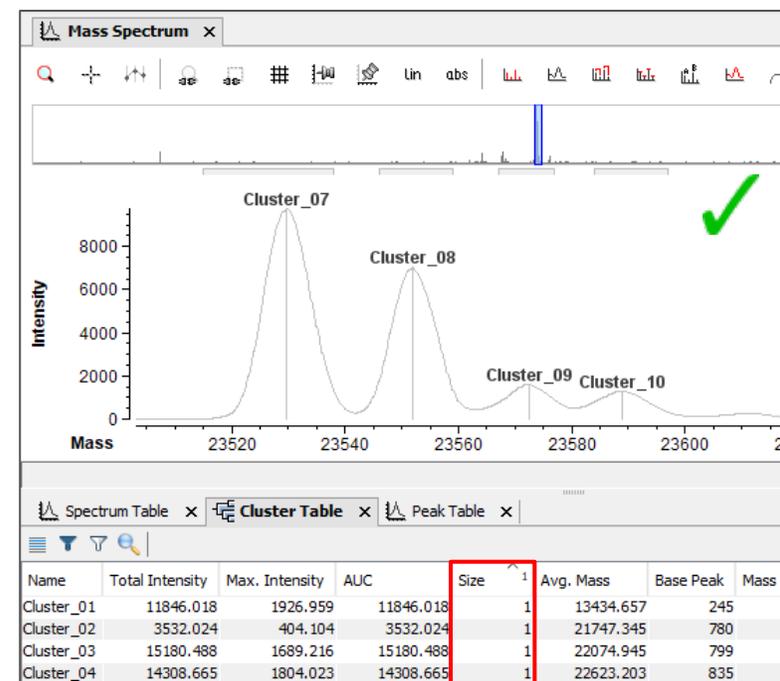
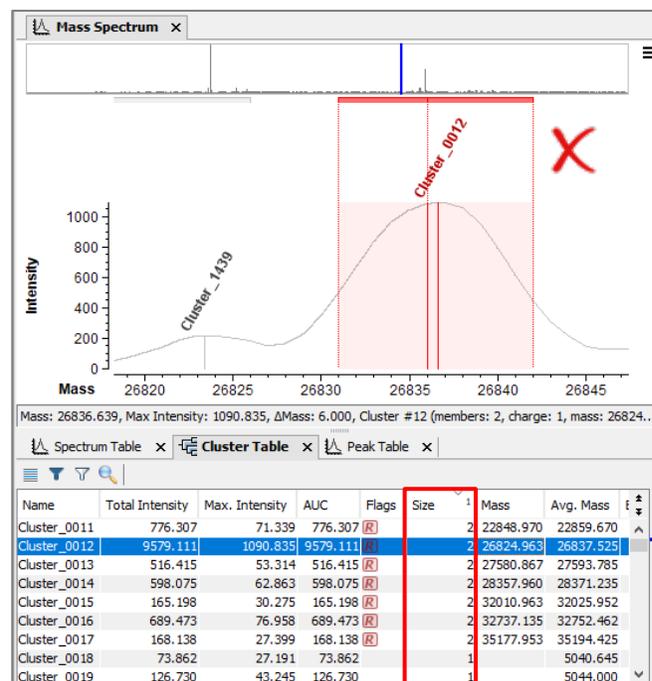


# Spectrum Isotope Clustering

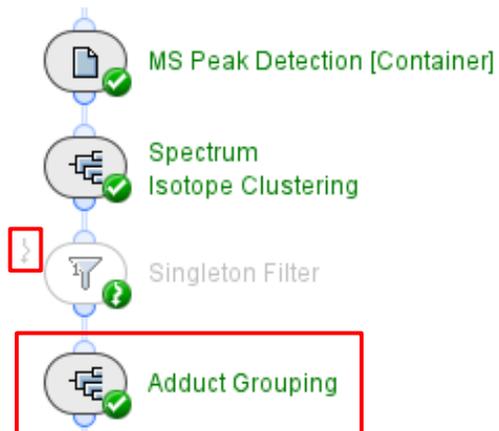


- *Spectrum Isotope Clustering* is required for *Adduct Grouping* with data that is not isotopically resolved.
  - Method: No Shape Restriction
  - Ionization: Massless

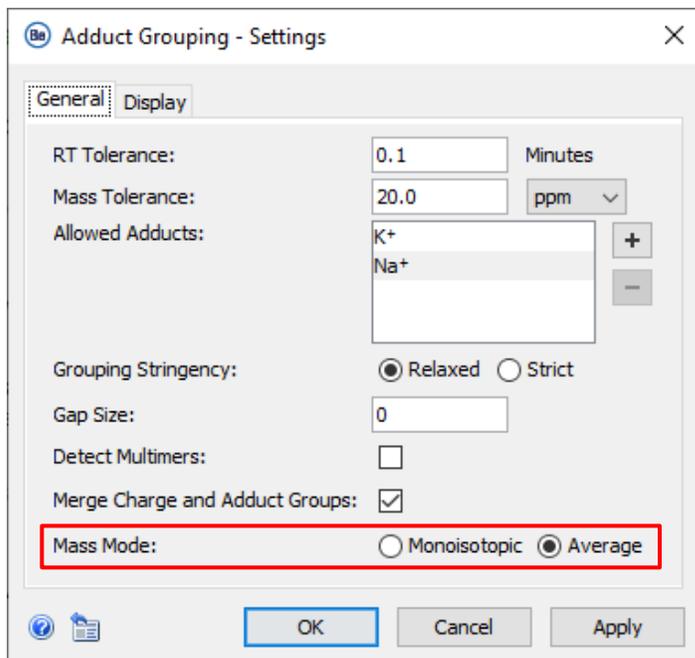
- Make sure that the results of *Spectrum Isotope Clustering* do not contain clusters with a **Size** that is more than 1.
  - To remove unwanted detection of peaks, increase the **Smoothing** or **Isolation Threshold** in *Spectrum Peak Detection*.



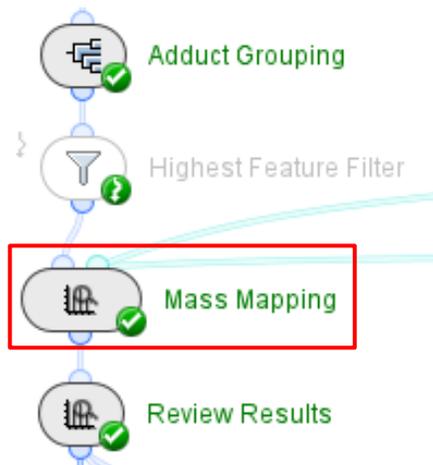
# Singleton Filter and Adduct Grouping



- *Singleton Filter:*
  - Activate the **Bypass** icon.
- *Adduct Grouping:*
  - **Mass Mode: Average.**



# Mass Mapping



- Select **Mass Mode: Average**.

**Mass Mapping - Settings**

General | Display

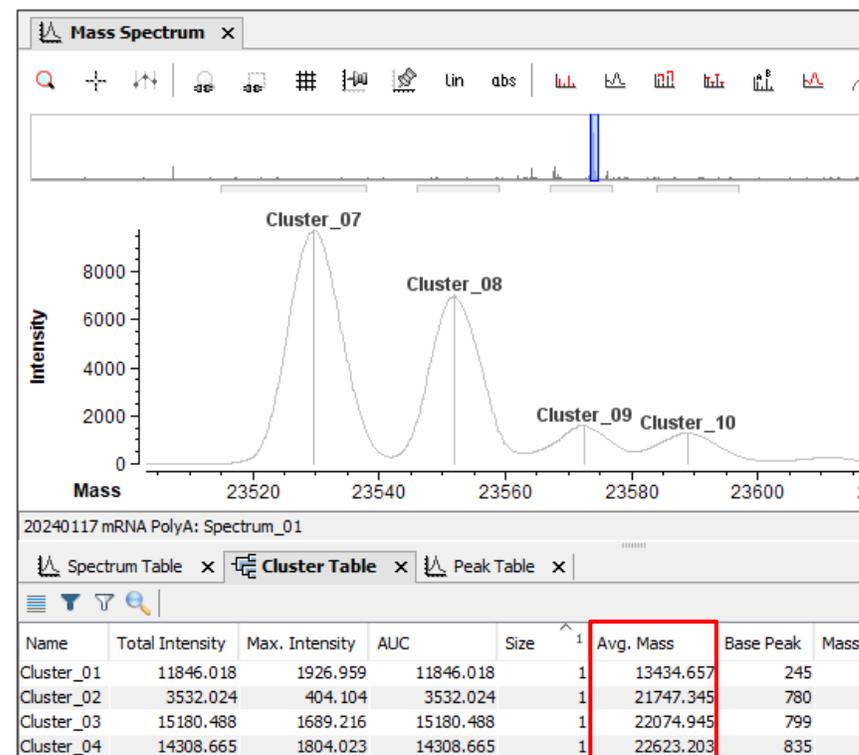
Mass Tolerance: 0.1 Da

Mass Mode:  Monoisotopic  Average

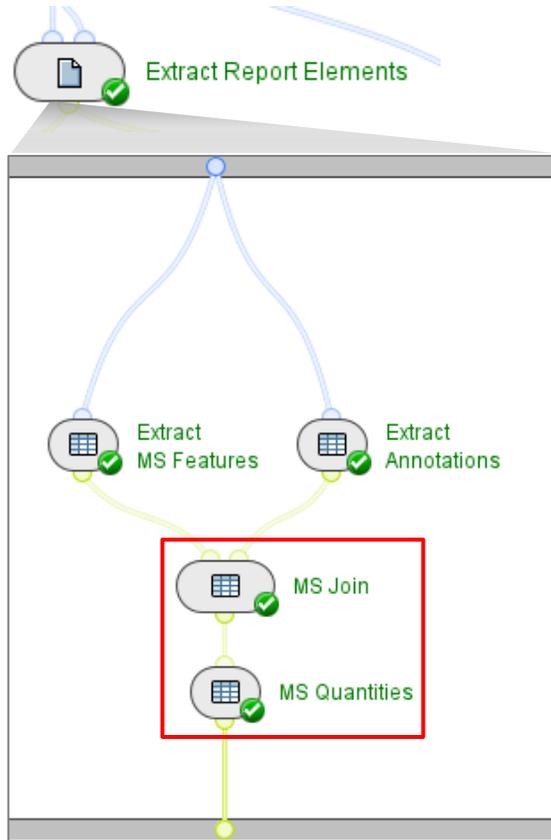
Limit to Best Match

Ignore Annotated Features

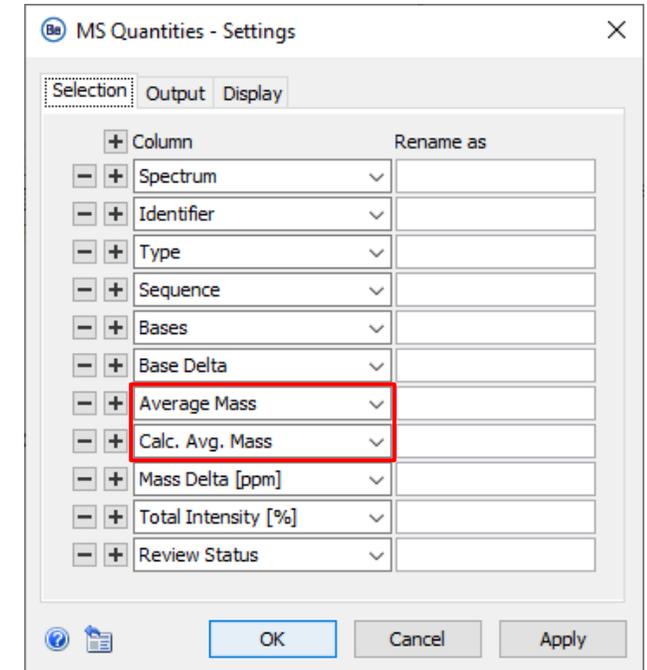
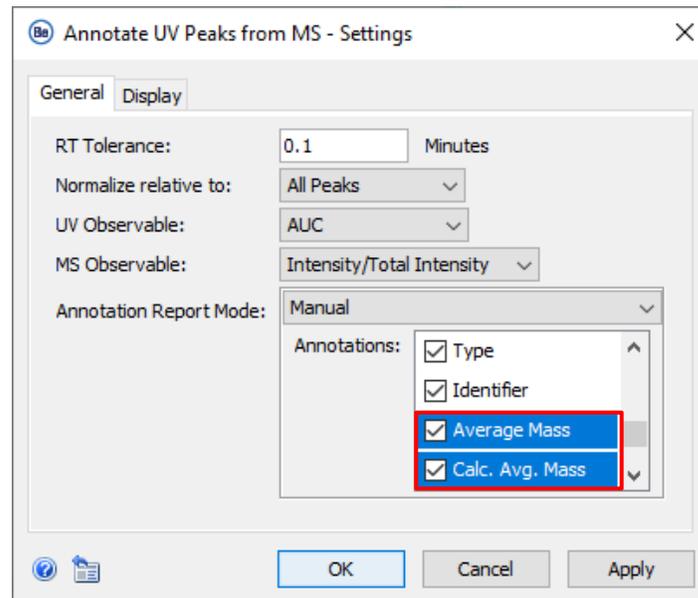
OK Cancel Apply



# Extract Report Elements



- **MS Quantities:**
    - Select **Average Mass** and **Calc. Avg. Mass** from the list.
    - To see the columns that are available for selection, run *MS Join*, and then select from the list.
- Note: If a selected column is empty, then the activity node shows a **yellow warning**. For example, if **Review Status** is selected, but there are no accepted identifications.



- **Average Mass** and **Calc. Avg. Mass** can also be selected in *Annotate UV Peaks from MS*.

# 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](http://www.sciex.com/diagnostics). All other products are For Research Use Only. Not for use in Diagnostic Procedures.

Trademarks and/or registered trademarks mentioned herein, including associated logos, are the property of AB Sciex Pte. Ltd. or their respective owners in the United States and/or certain other countries (see [www.sciex.com/trademarks](http://www.sciex.com/trademarks)).

© 2025 DH Tech. Dev. Pte. Ltd. TPUB-CUST-16012-B





The power of precision

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](http://www.sciex.com/diagnostics). All other products are For Research Use Only. Not for use in Diagnostic Procedures.

Trademarks and/or registered trademarks mentioned herein, including associated logos, are the property of AB Sciex Pte. Ltd. or their respective owners in the United States and/or certain other countries (see [www.sciex.com/trademarks](http://www.sciex.com/trademarks)).

© 2025 DH Tech. Dev. Pte. Ltd. TPUB-CUST-16012-B



**danaHER.**