

# ApQuant

## Manual

### Installation

For Proteome Discoverer Version 2.1.x and earlier:

- Download and install R (available at [CRAN](#))
- Then download and install apQuant with the apQuant installer (available at the [Protein Chemistry Facility website](#))

For Proteome Discoverer 2.3.x and later:

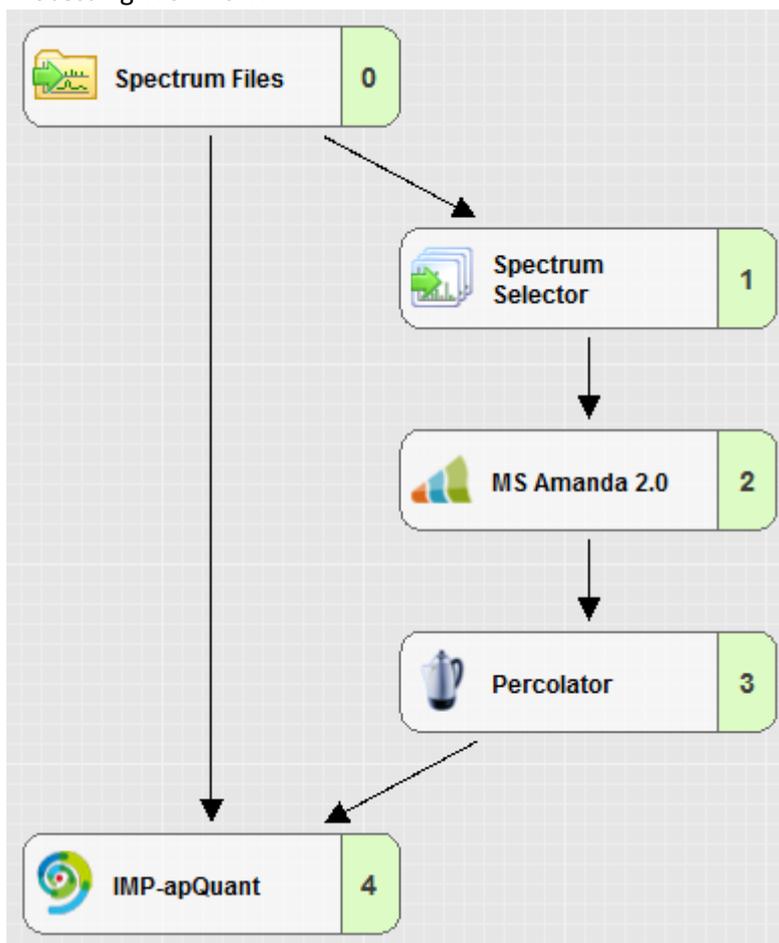
- Download and install apQuant with the apQuant installer (available at the [Protein Chemistry Facility website](#))

### Usage

Usage for Peptides and Proteins:

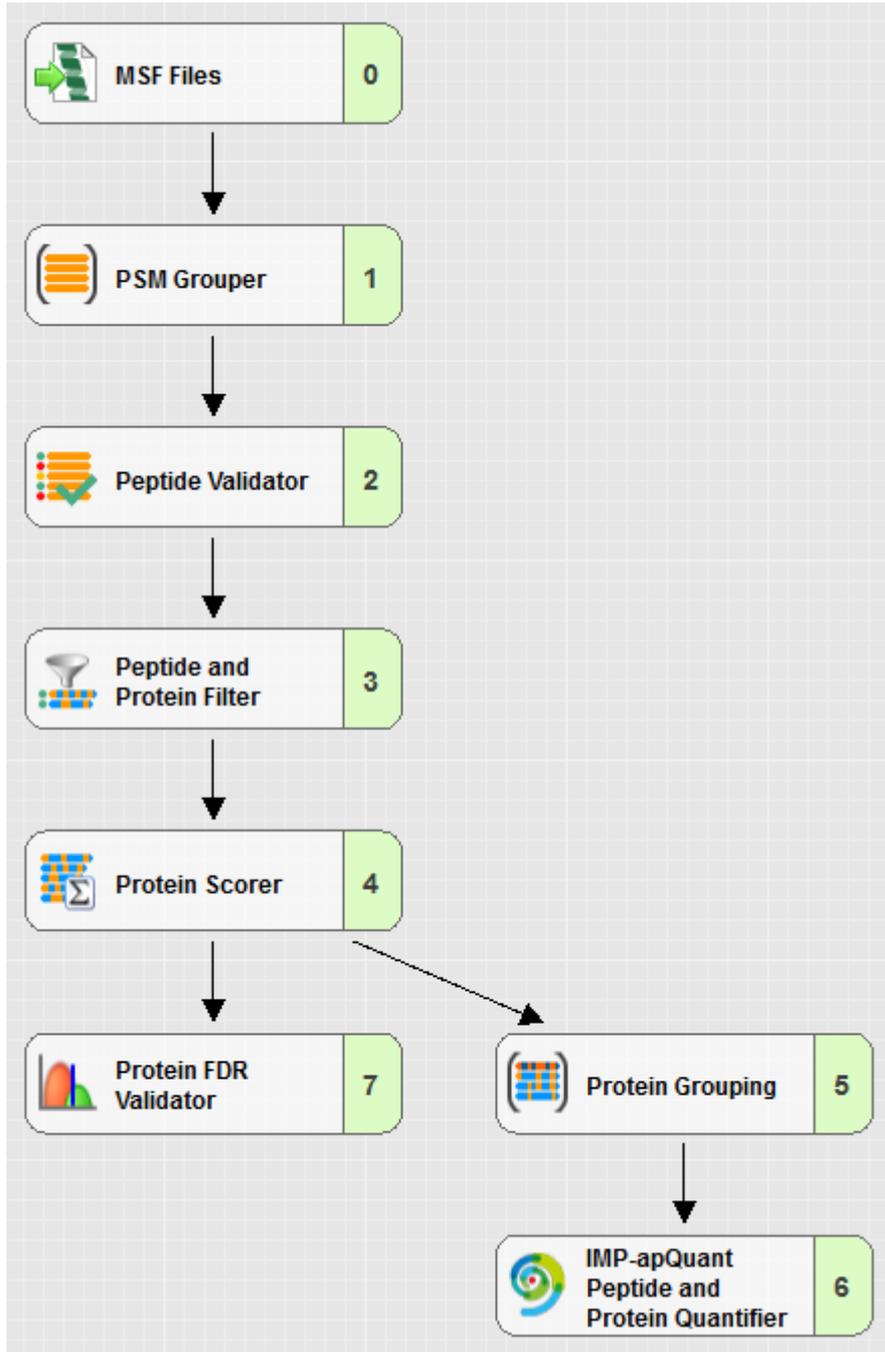
You need to use apQuant in both your Processing and Consensus Workflow

#### 1) Processing Workflow:



Here it needs a connection from the Spectrum Files Node and from your Peptide Validation Node.

2) Consensus Workflow:



The Consensus Node needs a connection from the Protein Grouping Node to function properly.

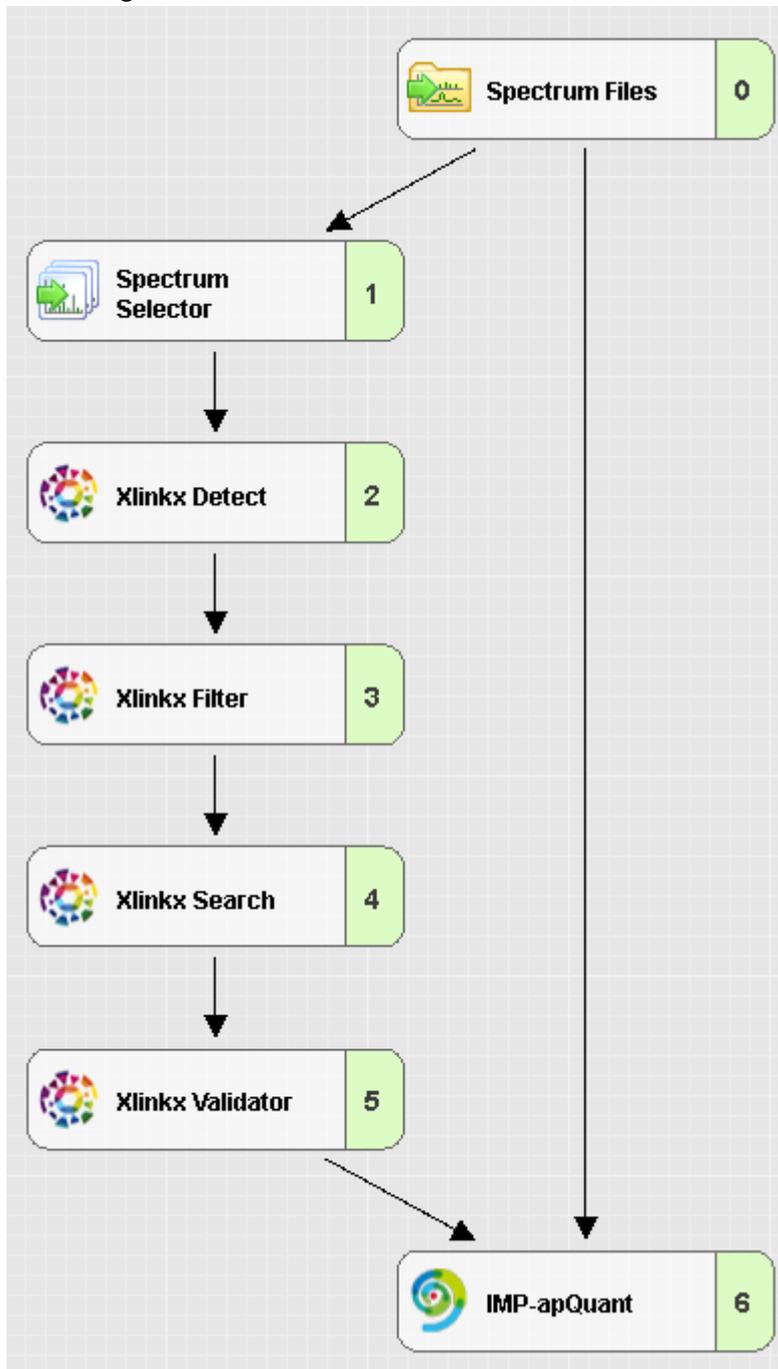
You can re-run the Consensus Node without re-running the whole Processing Workflow, as it will simply reprocess the data generated from the Processing Workflow. But if you want to find additional apQuant Features or QuanResults, the Processing Workflow also needs to be re-run.

Don't forget to set the Parameters correctly (explained in more Details later) and especially don't forget to set the Score Name Parameter and the Score Components to Use Parameter!

Usage for Crosslinks:

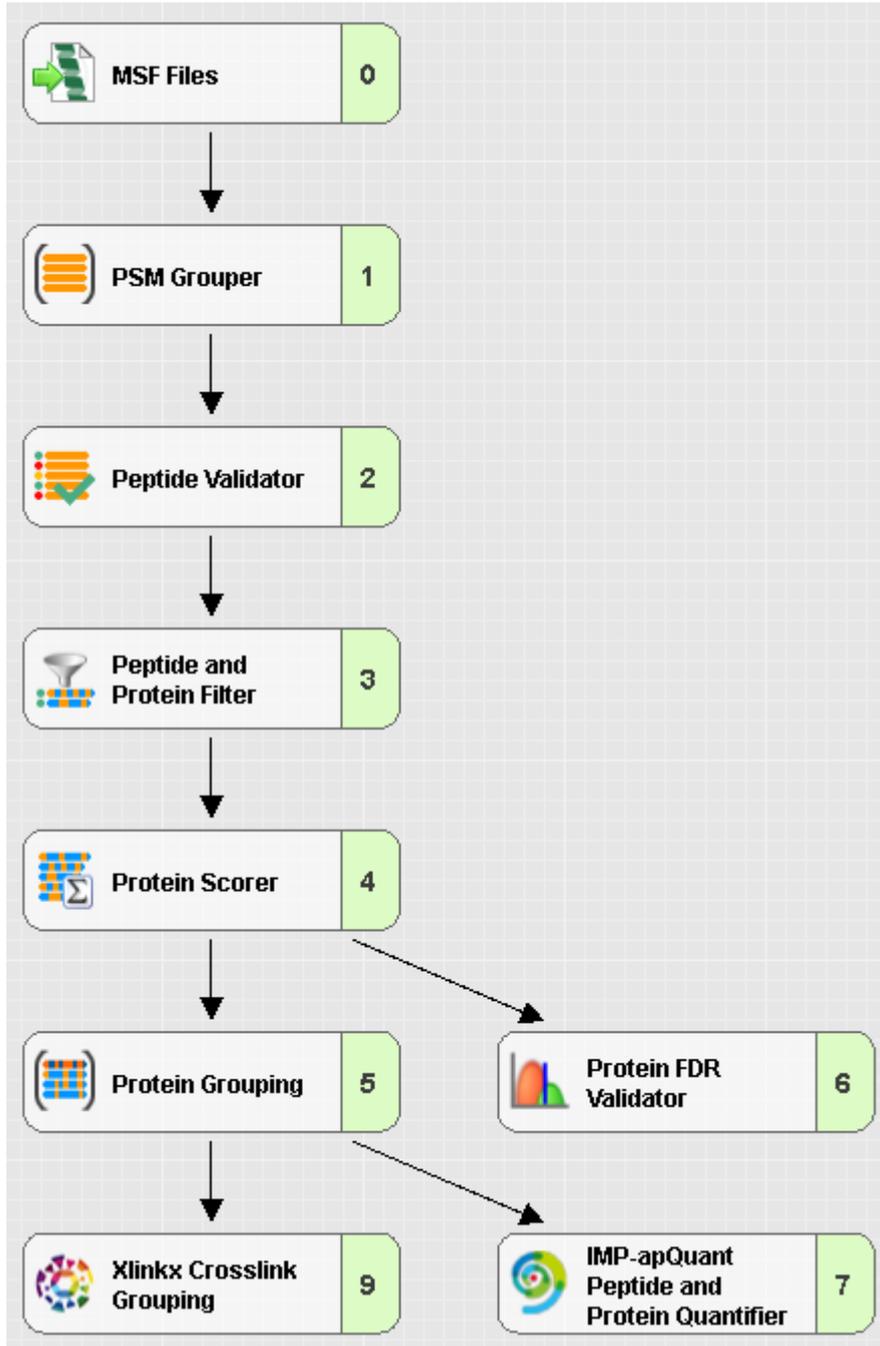
The Workflows for Crosslink data looks almost the same as the Workflow for Peptides and Proteins.

1) Processing Workflow:



The Processing Workflow needs a connection from the Spectrum Files Node and from the Crosslink Validation Node.

2) Consensus Workflow:



The Consensus Workflow has the same connections as the Workflow for Peptides and Proteins, but here it is important, if you use Xlinkx, that the Xlinkx Crosslink Grouping Node is present in the Workflow (it doesn't need a connection to it, it only needs to be present).

## Parameters

ApQuant has a lot of parameters, which will be discussed in more detail below:

Processing Parameters:

### 1. PSM Search Parameters

<i>Parameter</i>	<i>Default Value</i>	<i>Description</i>
<i>PSM Confidence Level</i>	High	The confidence level the PSMs to be considered should have.
<i>Minimum Sequence Length</i>	7	The minimum sequence length the PSMs to be considered should have.
<i>Score Name</i>		The Name of the Node providing the Scores for the PSMs. The text before the colon should exactly match the nodes name! If the nodes name is not in the list, you can add it with Administration → Configuration → MSF Files → PSM scores. Simply add your score value to it.
<i>Minimum Score</i>	70	The minimum score the PSMs to be considered should have.
<i>Search Engine Rank</i>	1	The Search engine rank the PSMs to be considered should have.
<i>Mass Tolerance</i>	5 ppm	The mass tolerance to determine the peaks area.

### 2. Integration Parameters

<i>Parameter</i>	<i>Default Value</i>	<i>Description</i>
<i>Match Between Runs</i>	True	If match between run should be performed. Match between runs tries to find areas in files, although this file does not have a fitting PSM. Only works if 2 or more files are provided, as the PSMs of one file will be tried to be found in the other file.
<i>LOESS Smoothing</i>	True	If Loess Smoothing instead of the Savizky-Golay Smoothing should be used. Smoothing is used for the algorithm to easier detect peaks, and LOESS smoothing is generally recommended.
<i>LOESS window size (in min)</i>	0.2	The window size to be used for the LOESS smoothing.
<i>Use Deisotoping</i>	False	If the spectra should be deisotoped. This improves the quality of the peaks at the cost of runtime.
<i>Tolerance for deisotoping (in Da)</i>	0.005	The tolerance to be used for Deisotoping.
<i>RT Correction</i>	True	If RT correction should be performed. RT correction will not be performed with less than two files.
<i>Retention Time Tolerance</i>	0.5	The tolerance in minutes a peak can be off of its PSM RT.
<i>Missing Peaks</i>	2	How many peaks to one end should be without intensity to see it as the end of the peak.
<i>Minimum Width for Peak (in min)</i>	0.15	How wide a peak should at least be to be counted as a peak.

<i>Noise Level</i>	10000	All signals below this value are treated as non-existent.
<i>Noise Level Percentage</i>	5	For each peak, all signals below the percentage provided here will be cut off, to reduce the trailing of peaks.
<i>FWHM Interpolation</i>	True	If the FWHM should be calculated using interpolation between peaks.

### 3. PhosphoRS/ptmRS Settings

<i>Parameter</i>	<i>Default Value</i>	<i>Description</i>
<i>Probability Threshold</i>	75	Probability Threshold to consider a Modification as fixed.
<i>PhosphoRS Column Name</i>	ptmRS [4]: Best Site Probabilities	Name of the PhosphoRS/ptmRS/anyPTM column where the site probabilities are found in. If the given name is not found, the node searches whether a column with 'Best Site Probabilities' is available, and uses this one.

### 4. Performance Parameters

<i>Parameter</i>	<i>Default Value</i>	<i>Description</i>
<i>Thread Count</i>	0	Number of threads calculating in parallel, 0 to determine automatically (uses all cores available except for one).
<i>Workpackage size</i>	1000	Number of items calculated in parallel. A too high number will result in full RAM, which can lead to a system/program crash.

### 5. Confidence Parameters

<i>Parameter</i>	<i>Default Value</i>	<i>Description</i>
<i>Export peak3d</i>	False	Debug Parameter! Used to export the Peak3D object for manual view.
<i>apQuant Confidence High Cutoff</i>	0.01	False Discovery Rate for highly confident features.
<i>apQuant Confidence Medium Cutoff</i>	0.05	False Discovery Rate for medium confident features.
<i>Score Components to use</i>	[DeltaT1; DeltaT2; DeltaM1; F1; F2; F3]*	The Score Components which should be used for the Scoring. The Default values are currently not set automatically and need to be set manually!
<i>Replacement Value</i>	5	Replacement value for infinity (could not be calculated) score values.
<i>Scoring Method</i>	Percolator	The scoring method to be used. Sum score is more robust if single values are compared, as they're the same every time, but percolator yields better values.
<i>Set Arealess to Low</i>	False	Sets the features with no area to the confidence Low (instead of Unknown).

<i>Use Robust Methods</i>	True	If more robust calculation methods should be used (median instead of mean and IQR instead of SD).
---------------------------	------	---------------------------------------------------------------------------------------------------

\* The Default values are currently not being set automatically and have to be set manually!

#### 6. Crosslink Search Parameters

<i>Parameter</i>	<i>Default Value</i>	<i>Description</i>
<i>CSM Confidence Level</i>	High	The confidence level the CSMs to be considered should have.
<i>Minimum Score</i>		The minimum score the CSMs to be considered should have
<i>Mass Tolerance</i>	5 ppm	The mass tolerance to determine the peaks area.

## Consensus Parameters

### 1. Protein Area

<i>Parameter</i>	<i>Default Value</i>	<i>Description</i>
<i>Peptides to Use</i>	Unique	Specifies which peptides are used for quantification.  Unique: Only peptides that are not shared between different proteins or protein groups are used for the protein quantification. Unique + Razor: Uses all peptides that are not shared between different proteins or protein groups. All shared peptides are used for the protein that has more identified peptides but not for the other proteins they are contained in. All: All peptides are used for the protein quantification.
<i>ID Confidence Level</i>	Medium	Only uses ID apQuant Features with a confidence above or equal to this value for Peptide Group / Protein Quantification
<i>MBR Confidence Level</i>	High	Only uses MBR apQuant Features with a confidence above or equal to this value for Peptide Group / Protein Quantification
<i>Peptides to use for Protein Area:</i>	iBAQ	Method for combining peptide areas to protein areas. The Top X methods take the X most intense peptides. iBAQ uses all quantified peptides and divides the area by the number of theoretical peptides. NiBAQ is the same as iBAQ, but with normalized values.
<i>Grouping Method</i>	Sum	When grouping is enabled, how should the areas from the different files be combined. To use grouping, you have to create a Study Factor called Grouping and assign the raw files accordingly

### 2. Peptide Area

<i>Parameter</i>	<i>Default Value</i>	<i>Description</i>
<i>Confidence to use</i>	mode(lowest)	Which confidence to use for the peptide group areas. If the peptides have a different confidence in different samples, use either the highest or lowest confidence. Mode chooses the most frequent confidence. In case of ties either the highest/lowest is chosen.
<i>Minimum Area</i>	50000	Minimal area to be considered as a real signal. Everything below this value is considered as not existing and will not be used for Peptide Group / Protein Quantification.
<i>Check for overlaps</i>	True	Check if the peaks of different peaks overlap. If there is an overlap, these areas will not be used for Peptide Group / Protein Quantification.
<i>Check for overlaps C13</i>	True	Check if the peaks of a peptide overlaps with the C13 peak of another peptide. If there is an overlap, these areas will not be used for Peptide Group / Protein Quantification.

### 3. Modifications to use

<i>Parameter</i>	<i>Default Value</i>	<i>Description</i>
<i>Use all modifications for Quantification</i>	True	Use all modified peptides for the area calculation. If set to false, only modifications selected below will be used.
<i>Dynamic Modification</i>	None	Select any known or suspected dynamic modification. Will not be used, if "Use all modifications for Quantification" is set to true.
<i>Dynamic Modification</i>	None	Select any known or suspected dynamic modification. Will not be used, if "Use all modifications for Quantification" is set to true.
<i>Dynamic Modification</i>	None	Select any known or suspected dynamic modification. Will not be used, if "Use all modifications for Quantification" is set to true.
<i>Dynamic Modification</i>	None	Select any known or suspected dynamic modification. Will not be used, if "Use all modifications for Quantification" is set to true.
<i>Dynamic Modification</i>	None	Select any known or suspected dynamic modification. Will not be used, if "Use all modifications for Quantification" is set to true.
<i>Static Modification</i>	None	Select any known or suspected static modification. Will not be used, if "Use all modifications for Quantification" is set to true.

### 4. iBAQ Parameters\*

<i>Parameter</i>	<i>Default Value</i>	<i>Description</i>
<i>Enzyme Name</i>	Trypsin (Full)	The reagent used for protein digestion.
<i>Min peptide length</i>	7	The minimum peptide length a peptide should have to be considered.

\*Parameters are only used, if iBAQ Quantification was selected.

### 5. Top X Parameters\*

<i>Parameter</i>	<i>Default Value</i>	<i>Description</i>
<i># peps for protein area</i>	3	How many peptides should be used for the protein area
<i>Usage of peptides</i>	average	How the peptide should be combined for the protein area.

\*Parameters are only used, if one of the Top X methods was selected.