CITRULLIA

Citrullination identification software

Manual version 1.0

Protein Research-group
Institute for Biochemistry and Molecular Biology
SDU

Mads Kierkegaard

Content

Installation

To install Citrullia:

1. Go to https://github.com/Citrullia-SDU/Citrullia.
2. Download the .MSI or the .EXE-install file.
   NB: Users at locations with restrictive security policies on computers such as universities and research facilities may encounter problems with the MSI-installation file. Therefore, the EXE-installation file should be used
3. Follow the installation guide
4. Congratulations. Citrullia has been installed.

To open Citrullia for users mentioned in step 2, it may be necessary to run Citrullia as an administrator.

Requirements

System requirements

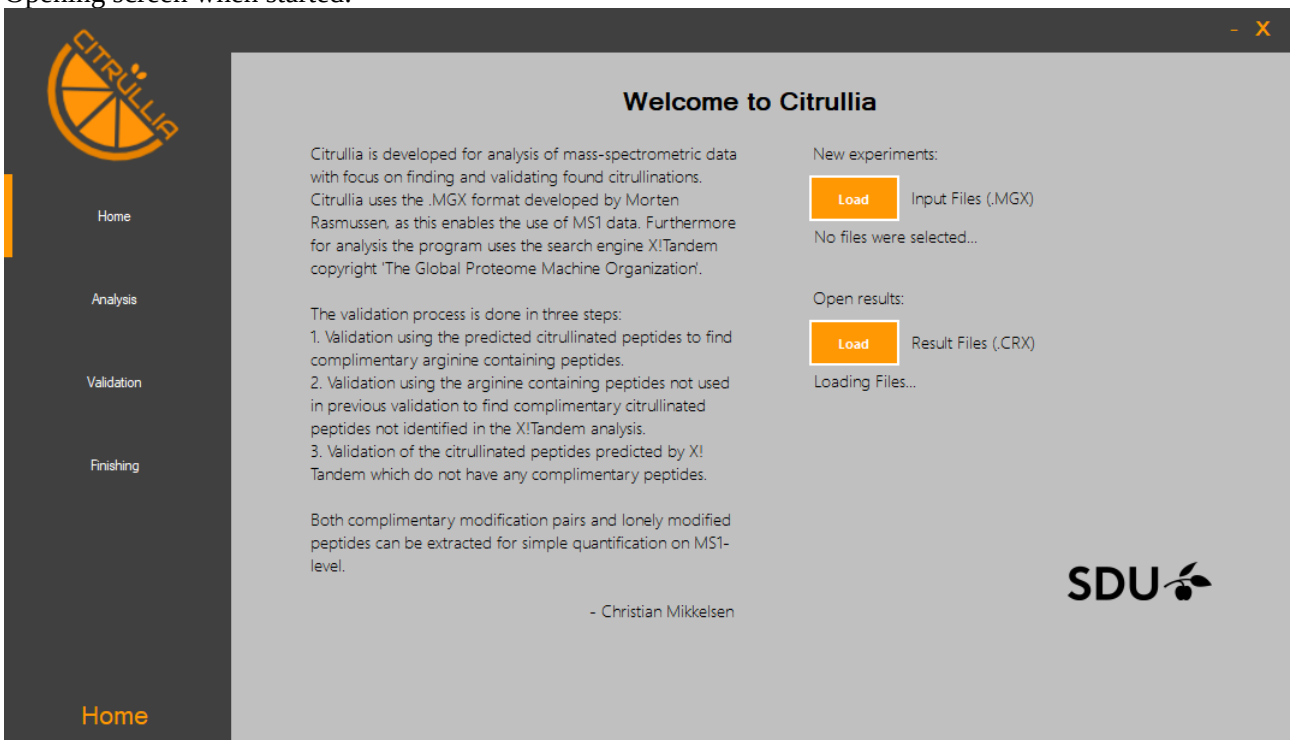
- .NET Framework 4.7.2 (Can be downloaded from https://dotnet.microsoft.com/download/dotnet-framework/net472. Download the Runtime version and not the Dev pack)
- Windows 10 64-bit (Programmed and tested. It might be working on Windows 7, but it has not been tested).
- Screen with minimum size of 1600x1100. (Dual screens are recommended).

File requirements

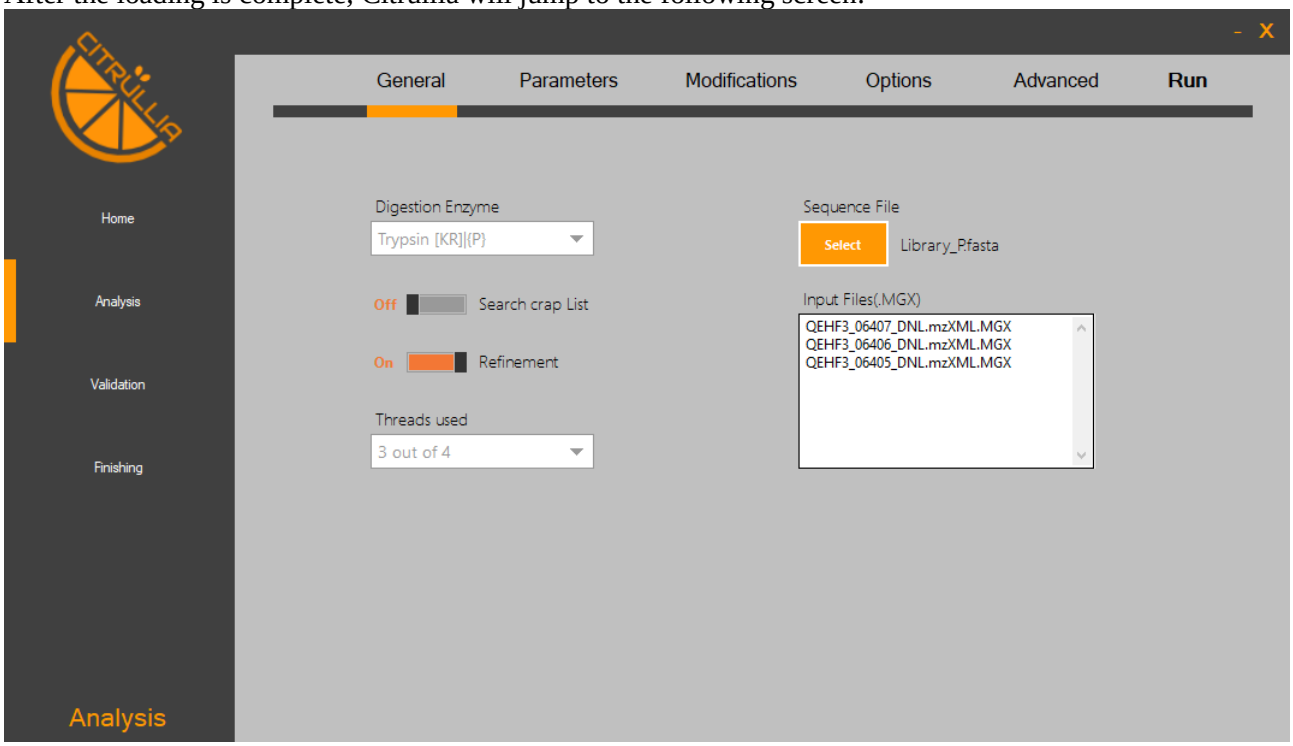
- Input file: The input file must a Mascot Generic Extended (MGX)-file. This is a custom format that contains additional information such as parent mass and MS1 spectra. To create these files MassAI must be used (http://www.massai.dk/).
- Sequence file: A FASTA formatted file containing a list of protein entries with label and protein sequence to be searched . The sequence file can be generated from UniProt etc.

Quick guide (For first time users)

Opening screen when started:

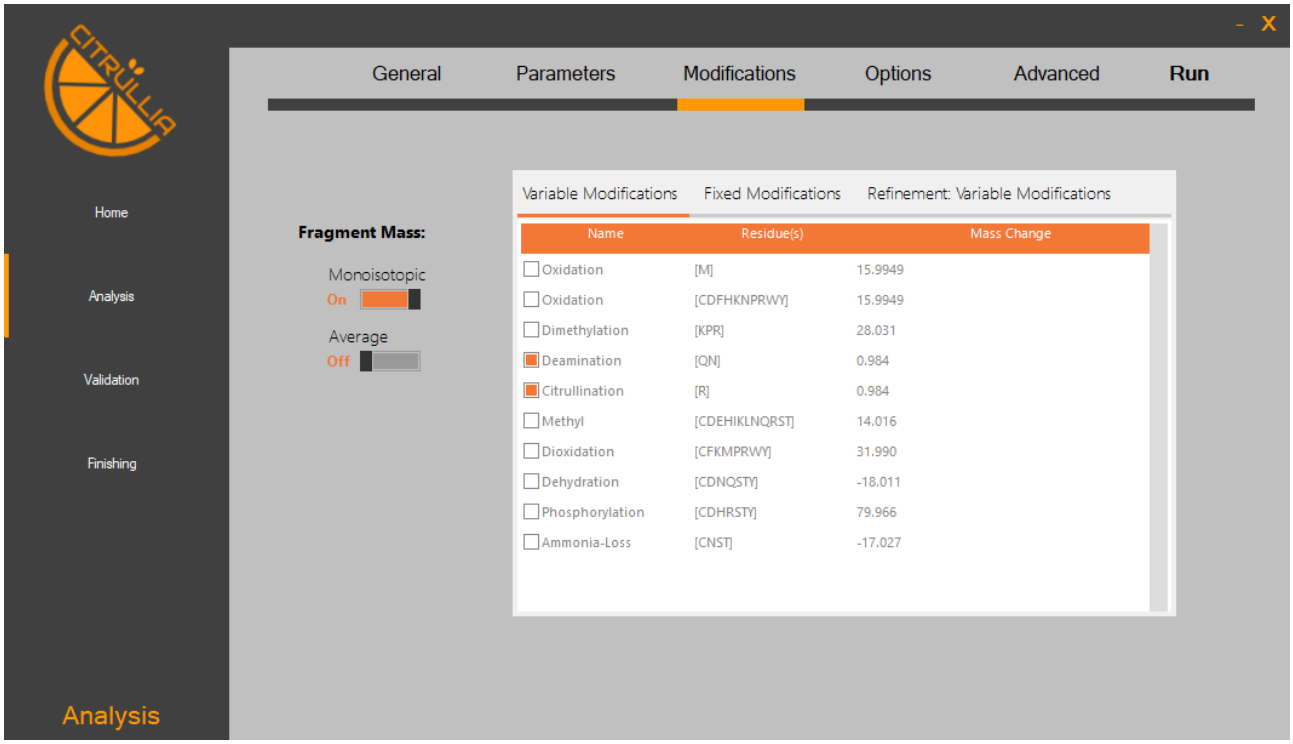


Click on the Load-button under New experiments. A dialog box will pop up prompting for the selection of the files.
Select the MGX-files and click OK. Citrullia will load the files.
After the loading is complete, Citrullia will jump to the following screen:



Select the digestion enzyme used to digest the sample
Select the sequence file

Then click on modifications:

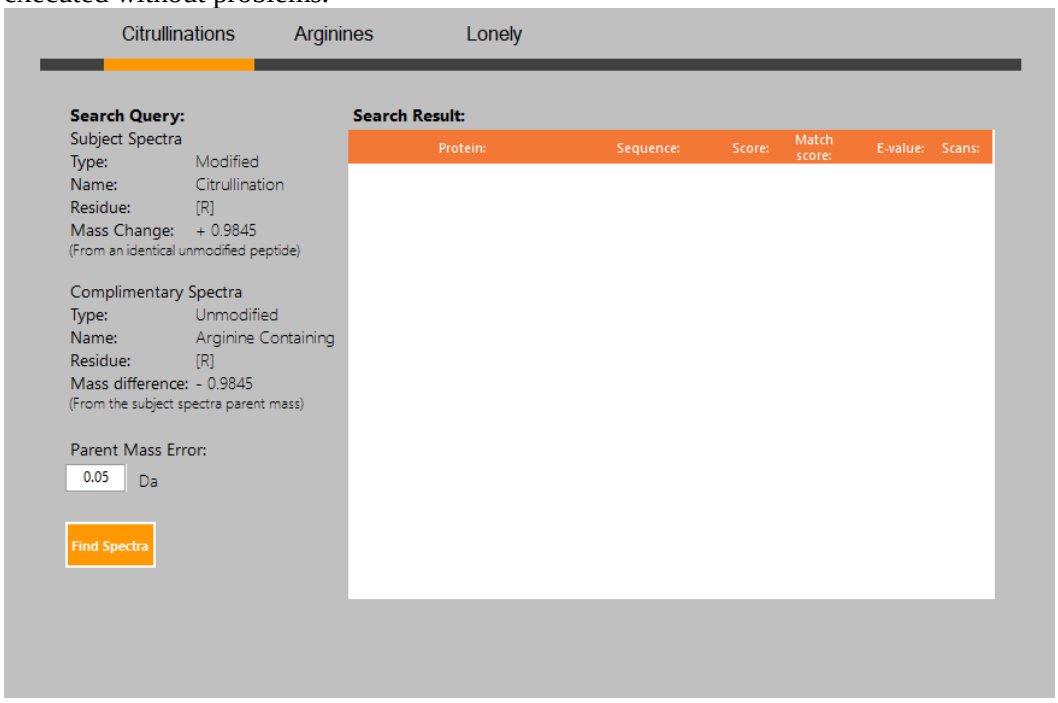


Select Citrullination of Arginine and Deamination of Asparagine and Glutamine by the checking the appropriate box.
The selection of deamination is essential to reduce the number of false positives.
If there are any fixed modifications, click on the Fixed modifications tab and select the modifications.

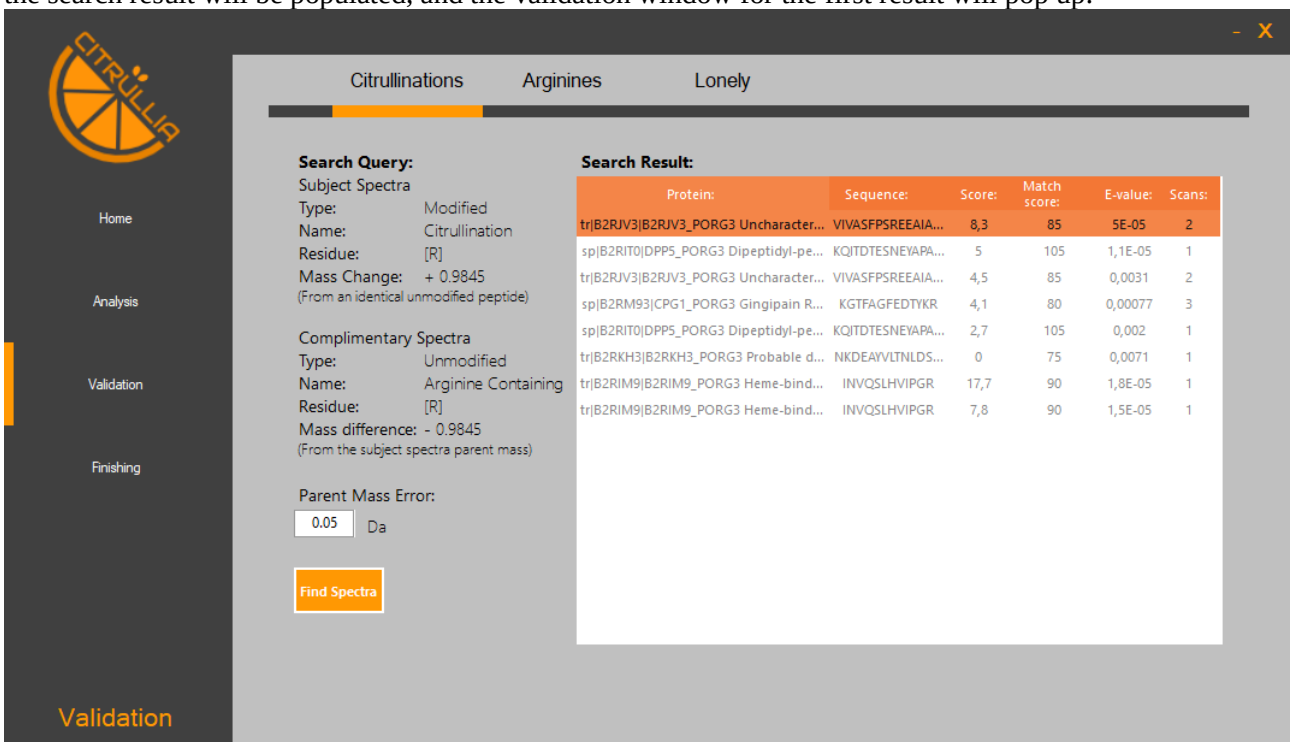
If X!Tandem should account for missed cleavage sites, go to the Options tab and add them.

Search

Afterwards, click Run. The command window will open and show the program of X!Tandem. This will happen for each selected file. If no dialog box pops up and the following window is shown, X!Tandem executed without problems.

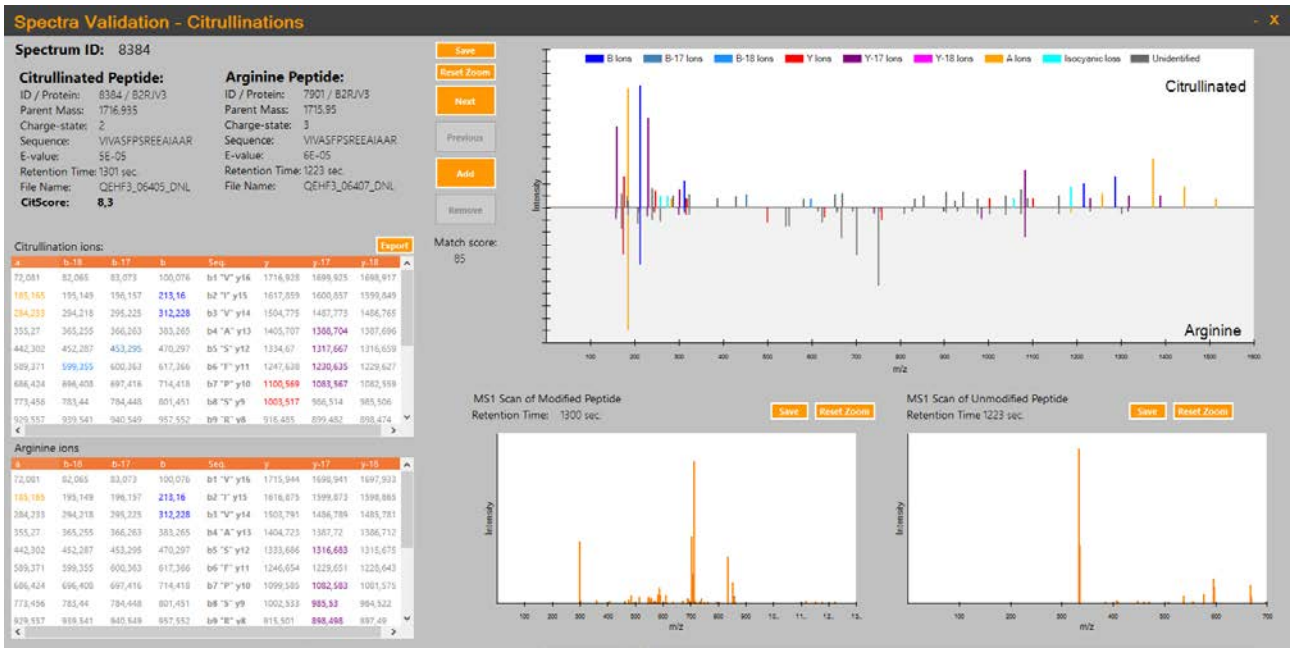


Click on Find Spectra. Citrullia will try to find citrullinations, which will take a little while. When finished the search result will be populated, and the validation window for the first result will pop up:



Validation

The validation window:



In the upper left corner, the information for the citrullination and arginine spectra can be seen.
Bellow it, the ions found are shown for both the citrullinated and the arginine spectra. These can be exported to an Excel spreadsheet by selecting the Export-button.
In the upper right corner, the two MS2 spectra can be seen. The upper spectrum is the citrullinated and the lower is the arginine spectrum. The key to the colours can be seen above the citrullinated spectra.

Below the MS2 spectra, the scan for the MS1 level are shown. A normal peak is orange, while the precursor is blue and possible isocyanic loss (only in the citrullinated MS1 scan) is marked in cyan.
For all three spectra it is possible to zoom in on the chart by marking it with the mouse. This zoom can be reset by clicking the Reset zoom-button. The spectra can be saved to a TIFF-file by selecting the Save-button.
Moreover, it is possible to move between the complementary spectra by clicking the Next and previous button.
To add or remove the validation pair for quantification, click on the Add or Remove button. The possibility of moving between the spectra is disabled by adding the spectra to quantification.

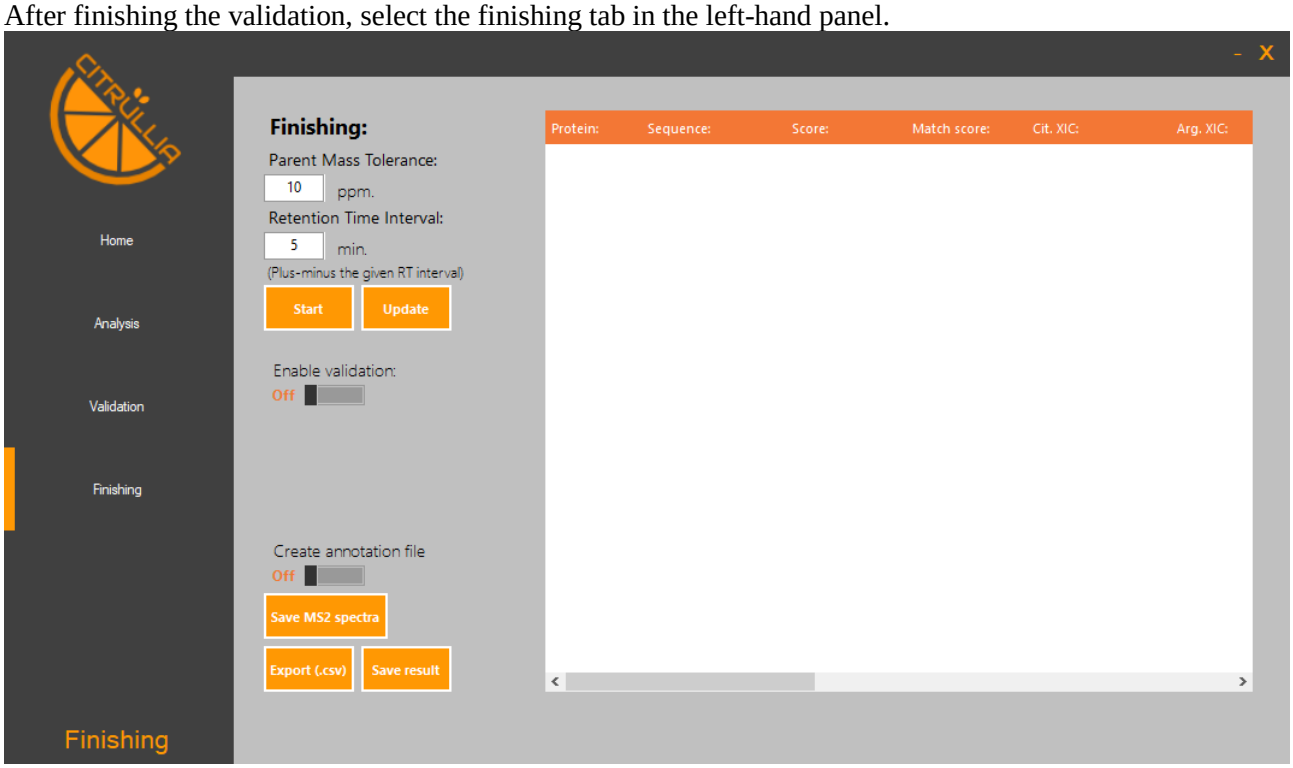
In the cases, where Citrullia cannot pair an arginine spectrum with a citrullinated spectrum, the citrullinated spectrum can be created as an “Lone Citrullination” result.

The principle is the same for Arginine- and Lone Citrullination-validation. However, the validation window for Lone Citrullination looks different, as it only displays a single spectrum.

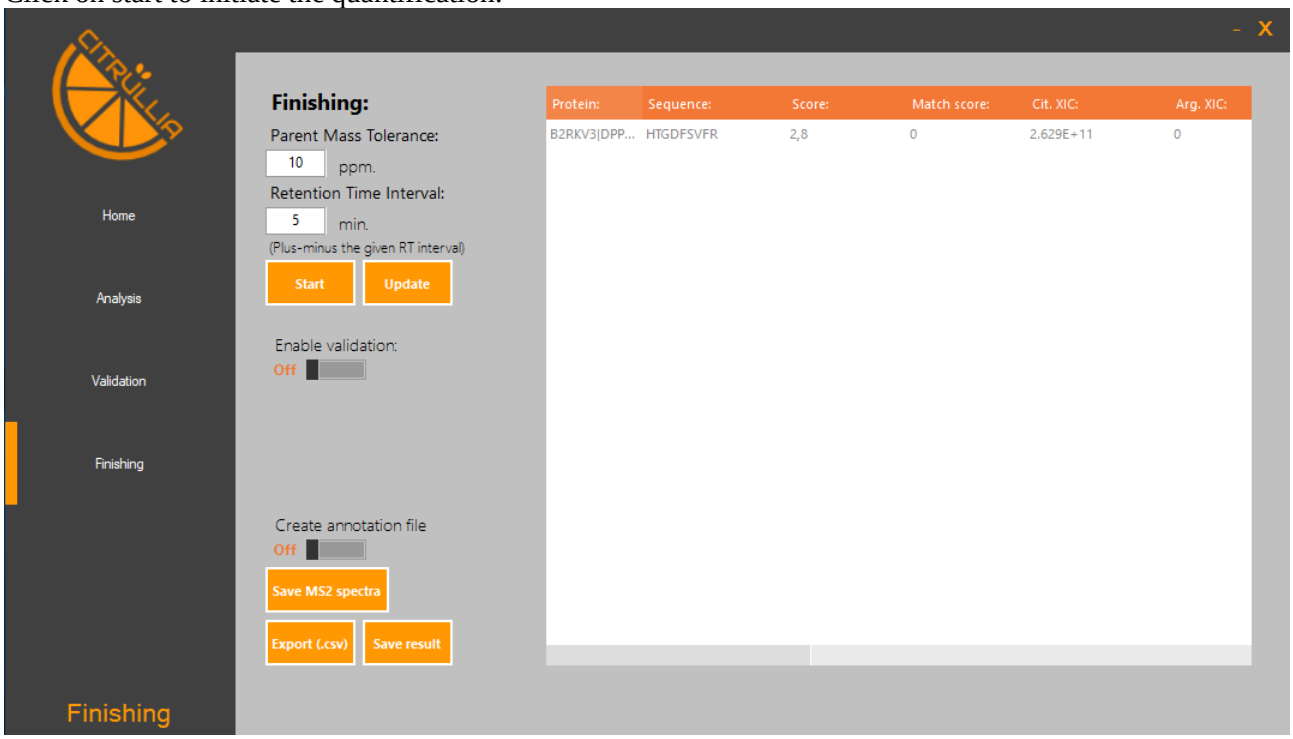
If no results are found for any of the validation types, try to go back and check that the settings are as they should be. It is possible that no results are found in the citrullination validation. However, results should be found in the arginine validation.

Quantification

After finishing the validation, select the finishing tab in the left-hand panel.



Click on start to initiate the quantification:



In the quantification step, Citrullia will calculate the extracted ion chromatogram for the spectra and a citrullination percentage.

From here it is possible to look at the spectra again by selecting the “enable validation”-option. This will open a window like the one in the validation step.

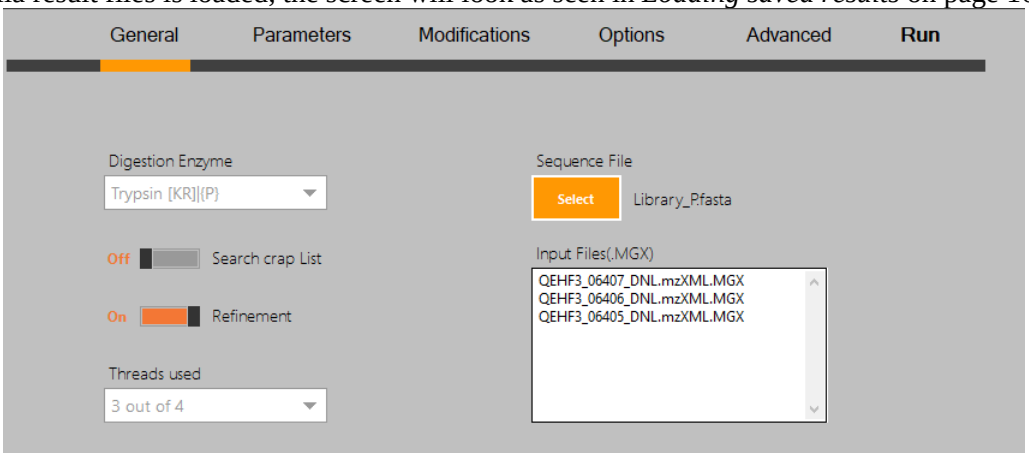
It is also possible to save the quantified annotated spectra to a folder along with a file containing the m/z-values and the ion type for all the spectra (if “Create annotation” is selected).
The quantification data can also be exported to an CSV-file by selecting the Export (.csv)-button.

By selecting the “Save result” you can save the X!Tandem files along the MGX-files. This saves time by bypassing the running of X!Tandem in the future.

Settings

General settings

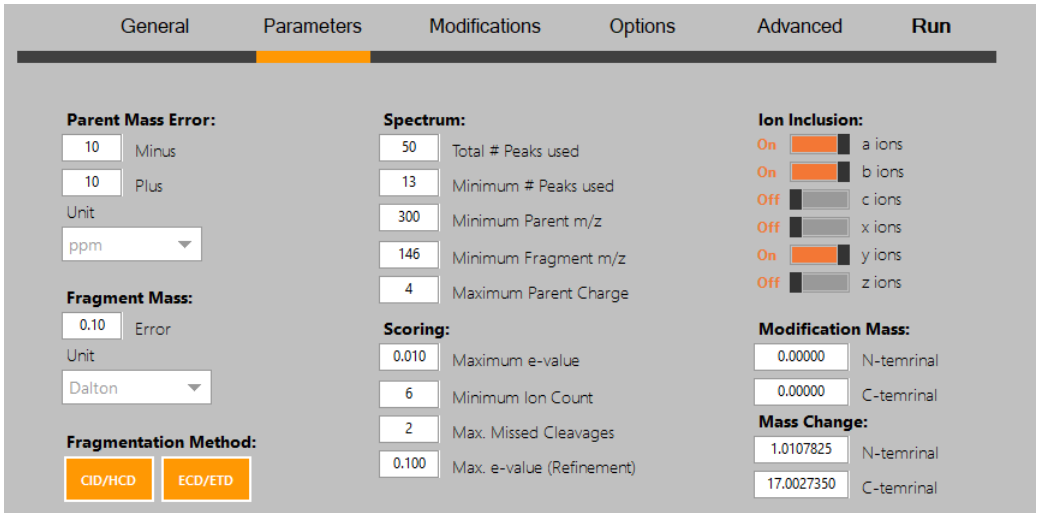
The general settings have two appearances after what type of files loaded. However, this does only change the input files. If MGX-files are loaded, the screen will look as *in Quick guide (For first time users)* on page 3. If Citrullia result files is loaded, the screen will look as seen in *Loading saved results* on page 16.



The select button is used to select a sequence file to use by X!Tandem. This must be a FASTA-file.

Parameter name	**Explanation**	**Default/Choices (In case of drop down)**
Digestion enzyme	The protease used to digest the enzymes prior to running the analysis. This is used by X!Tandem to determine the cleavage points in the sequence.	• Trypsin • Chymotrypsin • Endo Asp-N • Endo Glu-C • Endo Glu-C • Endo Lys-C • Non-Specific
Search crap list	Indicates that X!Tandem should search through a list of proteins, which are commonly used in laboratories, as MS standards or proteins from dust or physical contact with the sample. This provided by The GPM and is included in the installation.	Off
Refinement	Indicates that X!Tandem should try to refine the result and thereby increase speed and accuracy of the protein model	On
Threads used	The number of computer threads used by X!Tandem to perform calculations. Select the highest possible number for fastest computation.	1 to Threads on computer - 1
Sequence file	A FASTA file containing a list of proteins entries with label and protein sequence. There is performed a theoretical digest of these sequences by X!Tandem. These digests are then compared to the data found. The sequence file can be generated from data from UniProt etc.	An FASTA-file.

Parameter settings



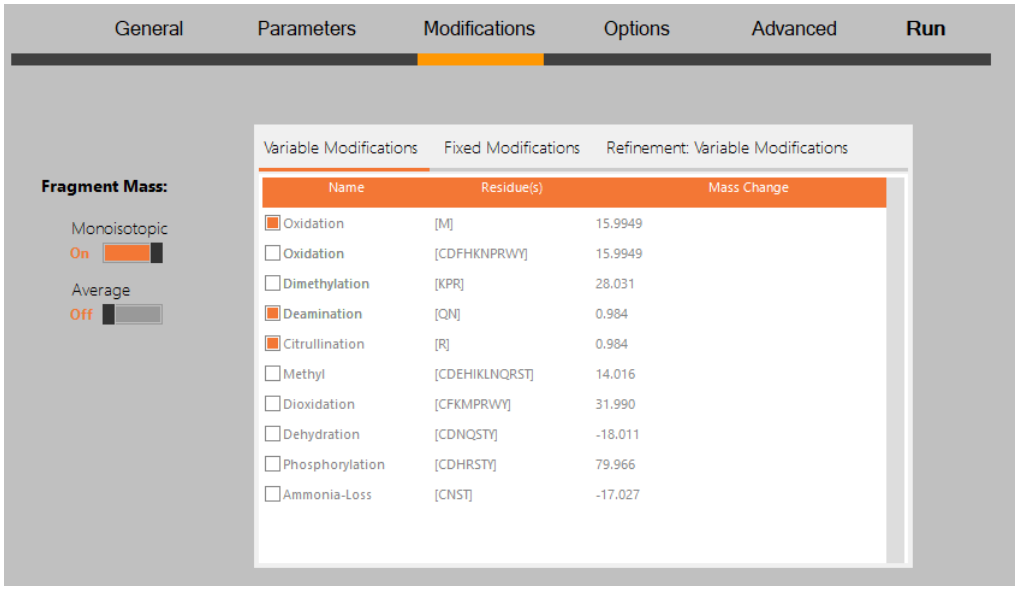
Parameter name	Explanation	Default/Choices (In case of drop down)
Parent Mass Error		
Minus	The lower bound of the parent mass error	10
Plus	The upper bound of the parent mass error	10
Unit	The units of the lower and upper bound.	PPM (Default) Dalton
Fragment Mass		
Error	The fragment mass error	0.1
Unit	The unit of the fragment mass error	Dalton (Default) PPM
Fragmentation method	The fragmentation method used in the mass spectrometer. Used to decide what ion types that should be allowed to contribute to the X!Tandem scoring. Can be either CID/HCD: A-, B- and Y-ions ECD/ETD:	*
Spectrum		
Total number of peaks used	The maximum number of peaks used by X!Tandem to the calculation	50
Minimum number of peaks used	The minimum number of peaks used by X!Tandem for analysis.	13
Minimum parent m/z	The minimum parent m/z-value required for a spectrum to be considered in X!Tandem's calculations.	300
Maximum fragment m/z	The minimum m/z-value required for a fragment to be considered in X!Tandem's calculations.	146

Maximum parent charge	The highest charge for the parent allowed for the spectrum to be considered in X!Tandem's calculations.	4
Scoring		
Maximum e-value	The highest value allowed	0.010
Minimum ion count	The minimum ion count required for X!Tandem to calculate the score.	6
Maximum missed cleavages	The maximum number of missed cleavages.	2
Max e-value (Refinement)	The highest value allowed for refinement.	0.100
Ion inclusion	Which ion type should be allowed to contribute to the X!Tandem scoring.	
A-ions		On
B-ions		On
C-ions		Off
X-ions		Off
Y-ions		On
Z-ions		Off
Modification mass		
N-terminal	Fixed modification mass in Dalton to the terminal residues	0.0
C-terminal		0.0
Mass change		
N-terminal	The mass added to the N-terminal after cleavage	1.0107825
C-terminal	The mass added to the C-terminal after cleavage	17.0027350

Modifications settings

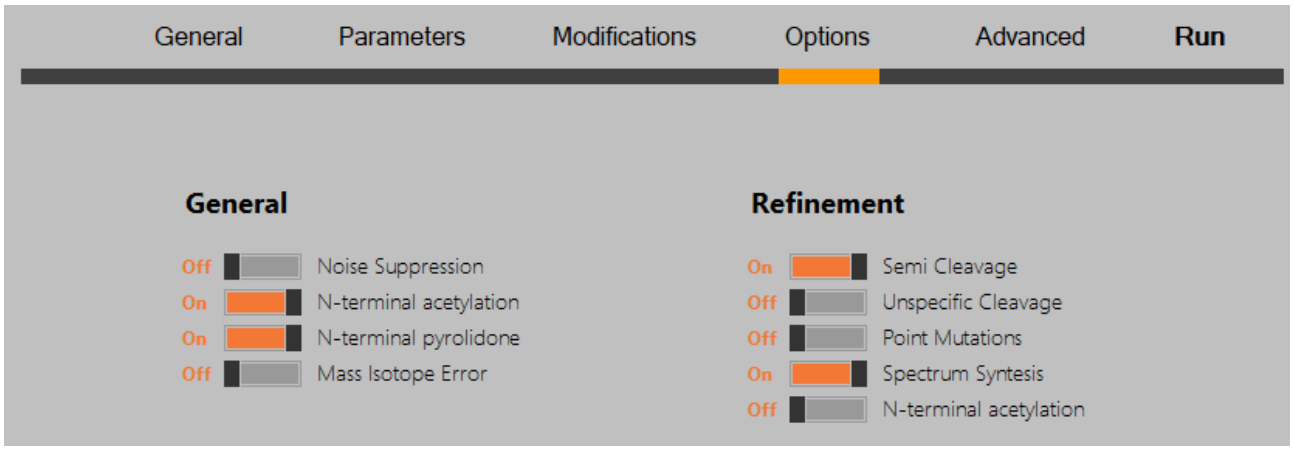
In the modification settings, the post-translational modifications can be set. This can be divided into three groups:
Variable modifications: The modifications that can be present.
Fixed modifications: The modifications that always are present.
Variable modifications (Refinement): The modifications that maybe present. This is only used in the refinement step.

The window shows the modification settings window with the list of variable modifications.
It should be noted that since Citrullia searches for citrullinations, the citrullination PTM should be selected. It would also be beneficial to select the deamination of Q and N, since it will reduce the number of false positive



The two switches choose whether the monoisotopic or the average mass should be used for the PTM masses. As default the monosporic mass is used.

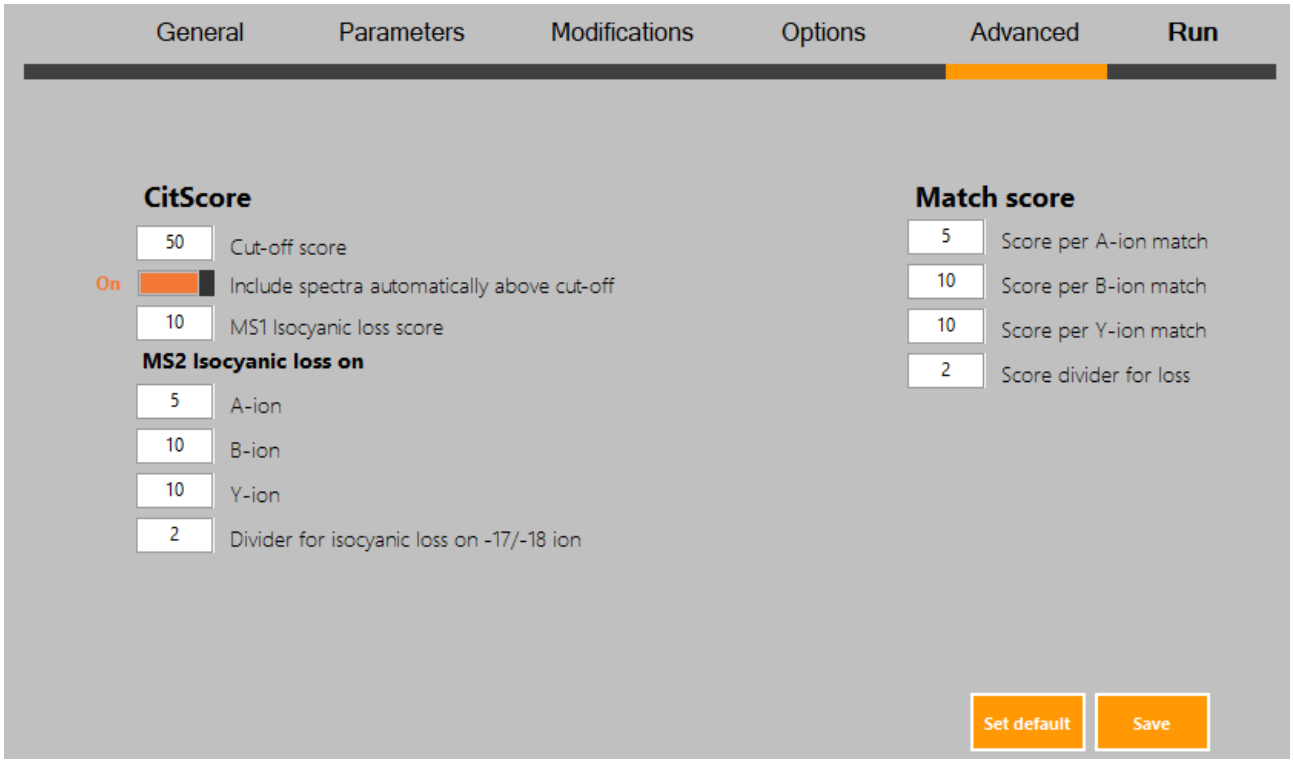
Options settings



Parameter name	Explanation	Default
General		
Noise suppression	Indicate if X!Tandem should perform noise suppression.	Off
N-terminal acetylation	Indicate if X!Tandem should detect N-terminal acetylation	On
N-terminal pyrolidone	Indicate if X!Tandem should detect N-terminal pyrolidone	On
Mass Isotope error	Indicates if X!Tandem should look for multiple tolerance windows centred on carbon-13.	Off
Refinement		
Semi-cleavages	Indicates if semi-cleavages rules should be used.	On

Unspecific cleavage	Indicates if X!Tandem should look for unspecific cleavages.	Off
Point mutations	Indicates if X!Tandem should check for point mutations in the found sequences	Off
Spectrum synthesis	Indicates if X!Tandem should value some of the chemical bonds higher than others.	On
N-terminal acetylation	Indicates to X!Tandem that N-terminal acetylations are a possibility.	Off

Advanced settings



Parameter name	Explanation	Default
CitScore		
Cut-off score	The threshold score for automatically adding the spectra to quantification.	50
Include spectra automatically above cut-off	Indicates that all spectra with a CitScore above the cut-off will be automatically added to quantification	On
MS1 isocyanic loss score	Score for an isocyanic loss on the MS1 scan.	10
MS2 isocyanic loss on		
A-ion	The score for an isocyanic loss from an A-ion	5

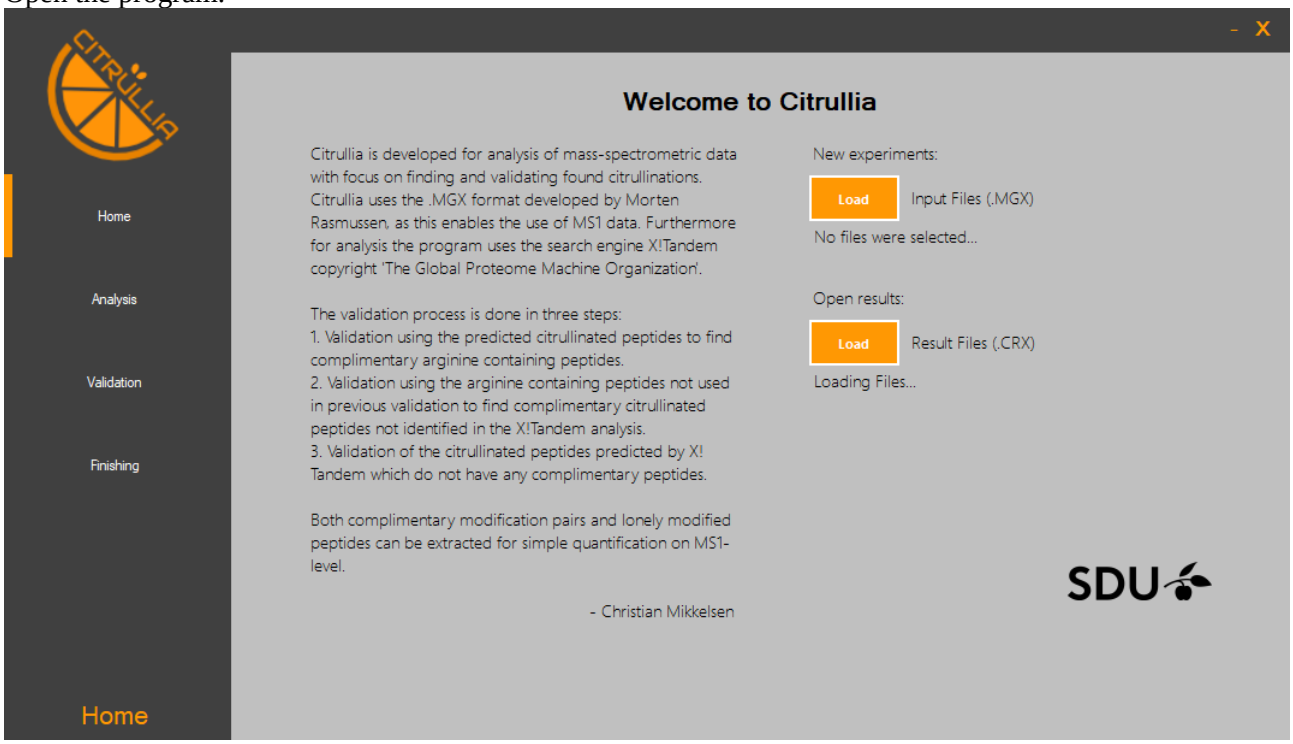
B-ion	The score for an isocyanic loss from an A-ion	10
Y-ion	The score for an isocyanic loss from an A-ion	10
Divider for isocyanic loss on -17/-18 ions	The divider for the score if the isocyanic loss comes from a -17 or -18-ion	2
Match score*		
A-ion	Between A-ions	5
B-ion	Between B-ions	10
Y-ion	Between Y-ions	10
Divider for loss	Divider for the match between a loss ion.	2

*) Score for each match between an ion in the citrullinated and arginine spectra. The match is also determined based on the position of the citrullinated arginine in the citrullinated spectra.

The button “Save” will save the current settings. These will be saved between program sessions.
The button “Set default” will set the score settings according to the values in the table and overwrite previous saved settings.

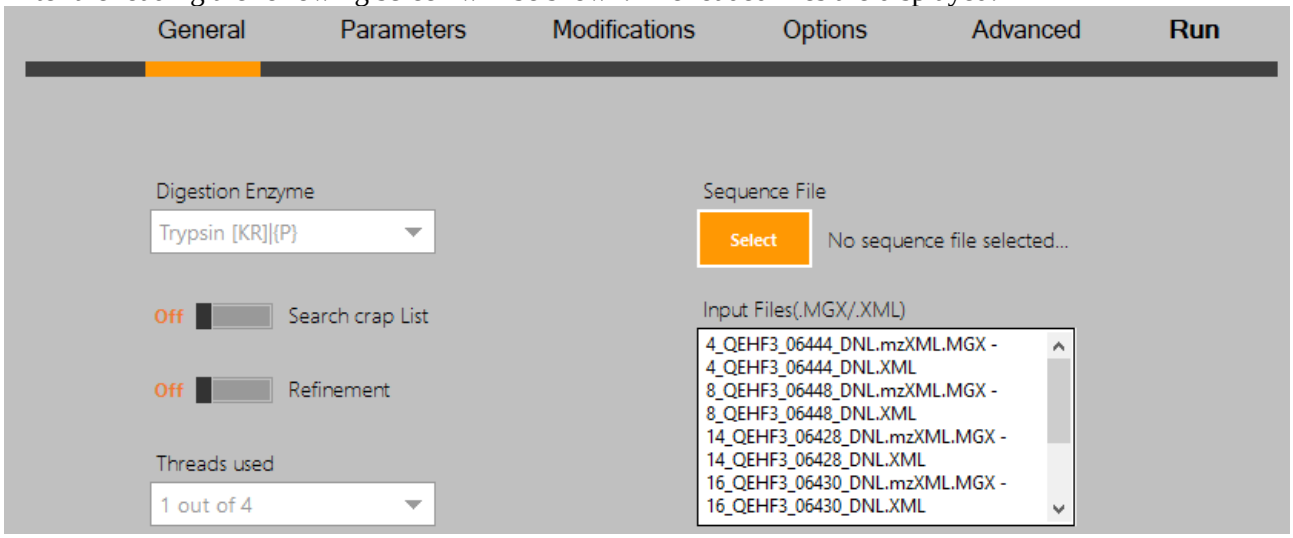
Loading saved results

Open the program:



Select the “Load”-button under Open results. The program will prompt for the loading of a Citrullia result (.CRX)-file. This will load the results.

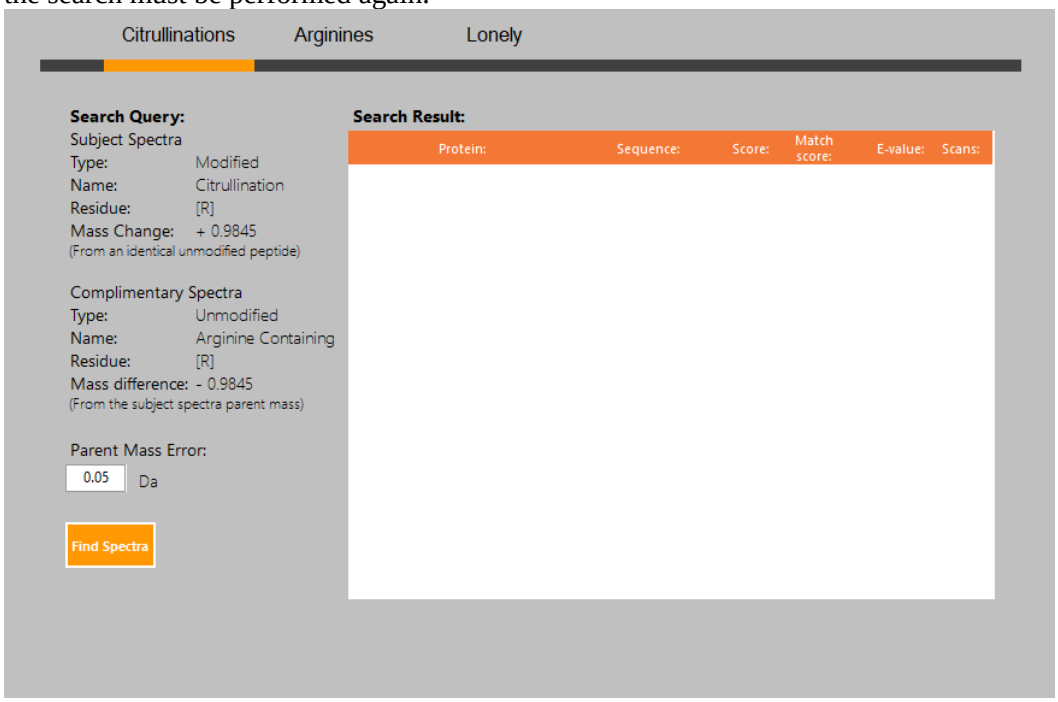
After the loading the following screen will be shown. The loaded files are displayed:



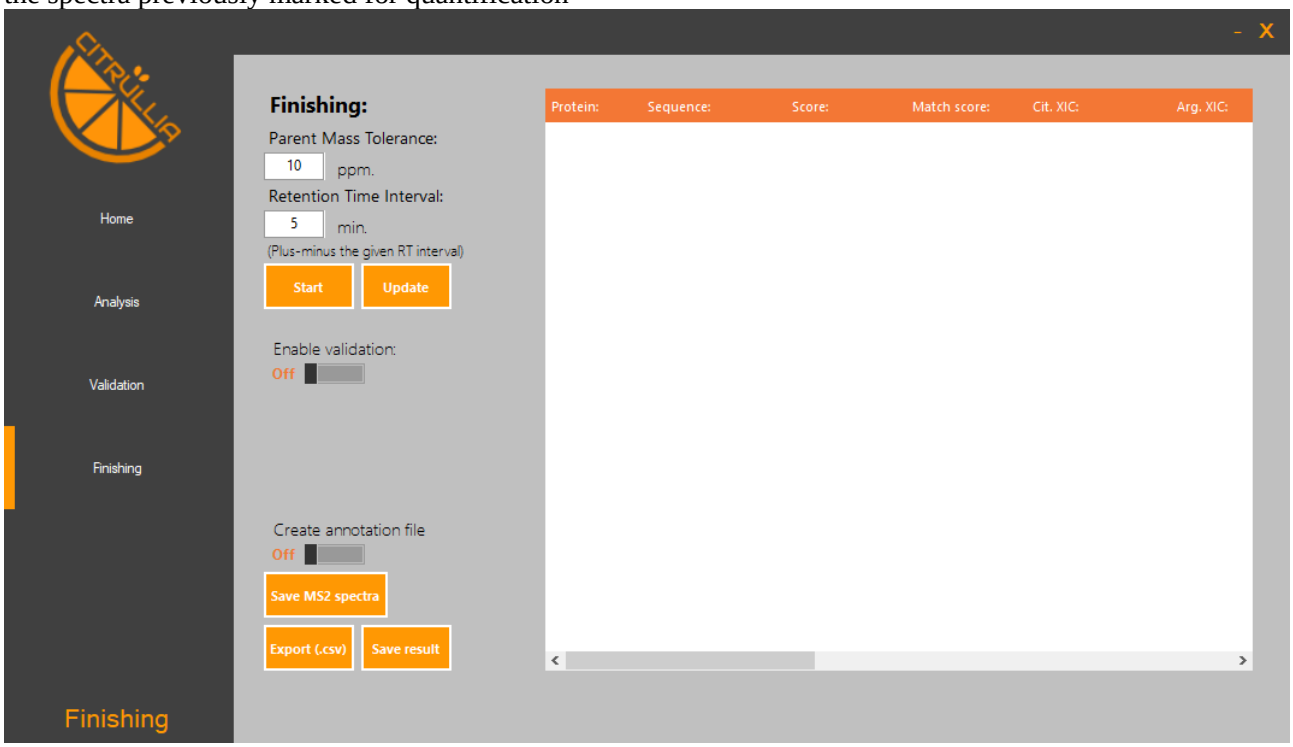
If the score settings should be changed, change them under the section “Advanced”. Se more in paragraph *Advanced settings* on page 14.

If this is done or no settings should be changed, click on Run.

After the run, the search window will show. In order to load the previous spectra marked for quantification, the search must be performed again.



After the search, it is possible to go to quantification-step and start the quantification. This should bring up the spectra previously marked for quantification



Quantification validation

After the quantification it is possible to look at the spectra again. This is done by enabling the “Enable validation”. This will bring up a window as below. However, the title might be different depending on the validation-method used.

Here it is possible to do the same things as in the Validation-step. If the spectrum pair should not be quantified anyway click on “Remove”-button and go to the previous window and click on “Update”. This will update the quantification to the new selection.
If the “Start”-button is clicked it will bring back all the spectra marked for quantification that were initially marked in the validation step.