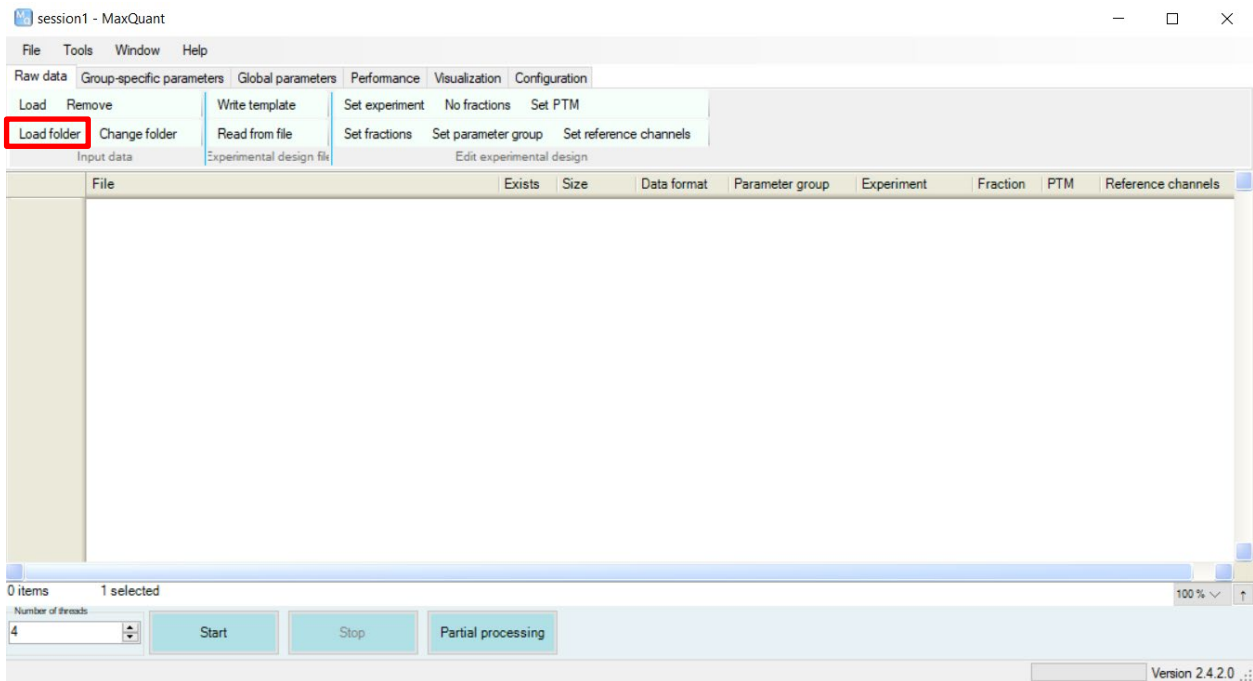# MaxQuant DIA

1. Running MaxDIA with label-free data

Objective: process the data from label-free experiment obtained with Orbitrap DIA.

Data source: **Van Puyvelde, B., Daled, S., Willems, S., Gabriels, R., Gonzalez de Peredo, A., Chaoui, K., Mouton-Barbosa, E., Bouyssié, D., Boonen, K., Hughes, C.J. and Gethings, L.A., 2022. A comprehensive LFQ benchmark dataset on modern day acquisition strategies in proteomics. Scientific Data, 9(1), p.126** (https://doi.org/10.1038/s41597-022-01216-6). For the assignment purposes only two raw files are selected.

1. Launch MaxQuant and click on the "Load folder" button to load mass spectrometry output data into MaxQuant. The data is located in the directory `Data/Assignment_2_MaxQuant_DIA/1_LFQ_Orbitrap/` on your flash drive.

2. Select one row and click "Set experiment" button to set the experiment design of this raw file. Do it for both files. Set the number of threads to be utilized by MaxQuant. You can press Ctrl + Shift + Esc to open Task Manager and select the Performance tab to see how many logical processors your PC has. This determines the maximum number of threads you can set. Consider also how much physical memory (RAM) your PC has. Type "System Information", press Enter and look at the "Installed Physical Memory (RAM)" to find out. For DIA it is recommended to have at least 4 GB of RAM per thread (e.g. 4 threads would need 16 GB of RAM).
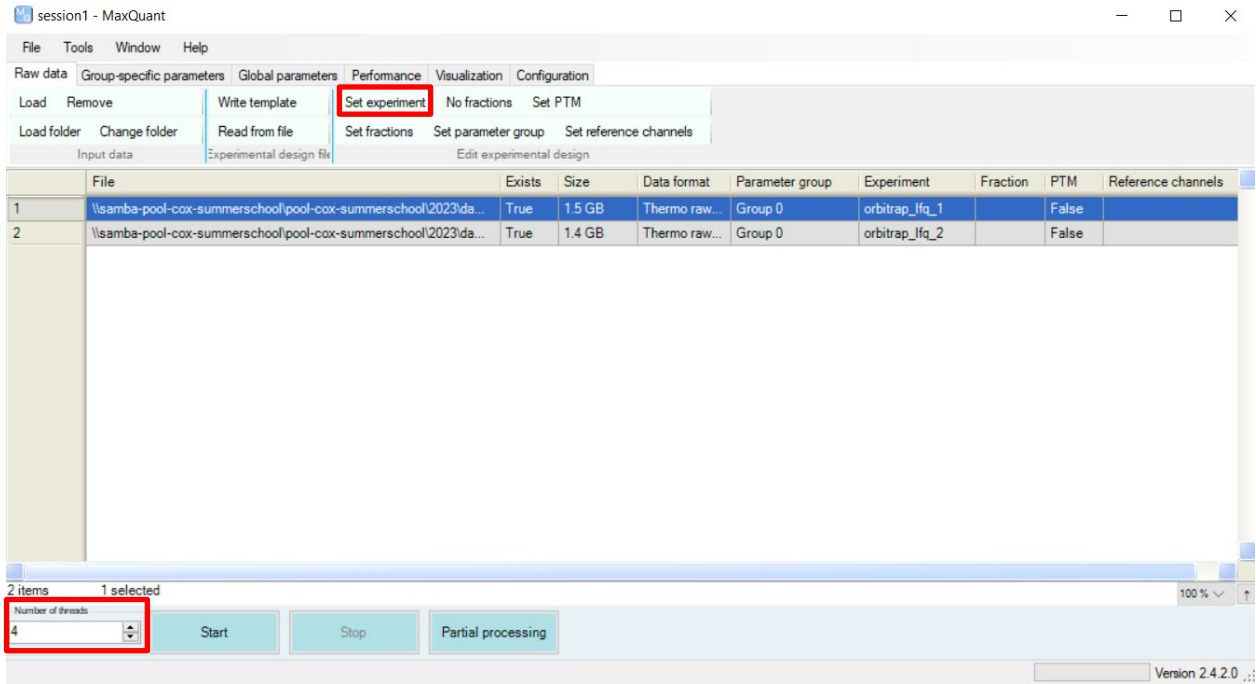
3. Move to the "Group-specific parameters" tab. Here, in the "Type" section, you can select the type of the mass spectrometry run. For this task you need to choose "MaxDIA".
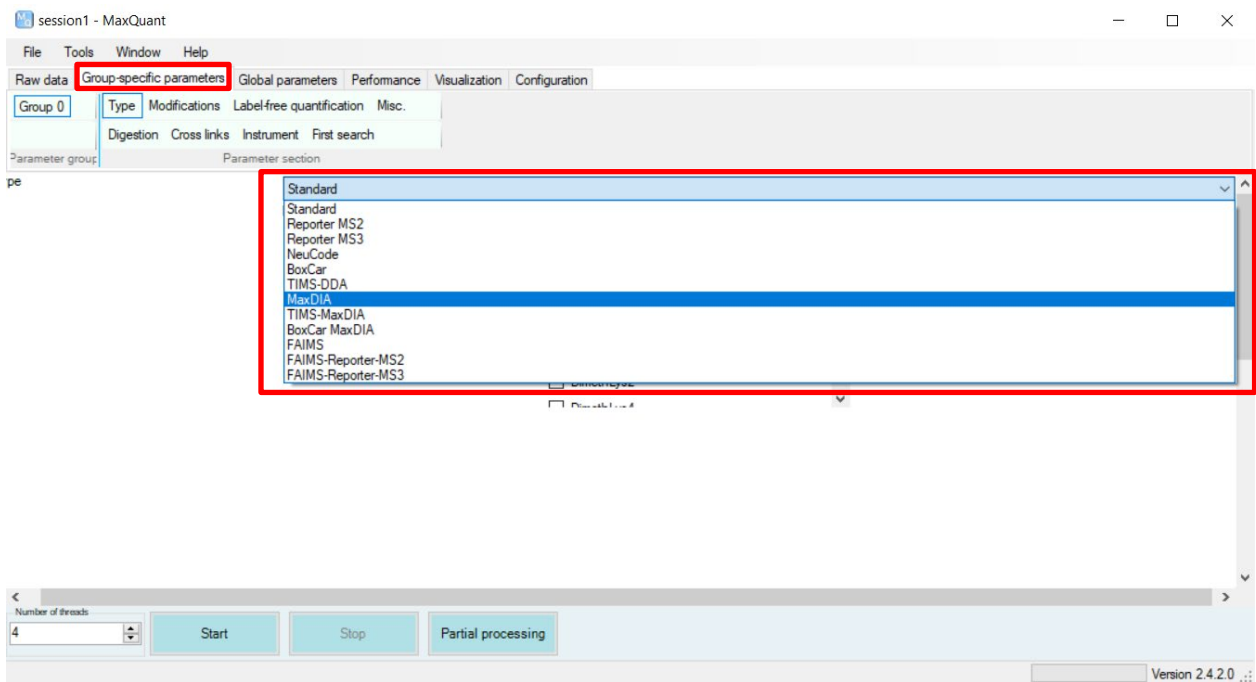

4. Leave the default setting ("MaxQuant") for the "Library type". This way you can use measured library outputted by MaxQuant after the DDA run. You can choose "tsv" option to use libraries from the third-party software that supports a `.tsv` output format but that goes out of the scope of this task.

5. Add `evidence.txt` and `msms.txt` library files into the corresponding boxes. The files are located in the directory `Data/Assignment_2_MaxQuant_DIA/1_LFQ_Orbitrap/measured_library/` on your flash drive.

6. Click on the "Label-free quantification" button to move to the corresponding section of the "Group-specific parameters". Select "LFQ" from the drop-down menu to enable MaxQuant label-free quantification algorithm (MaxLFQ) in this run.
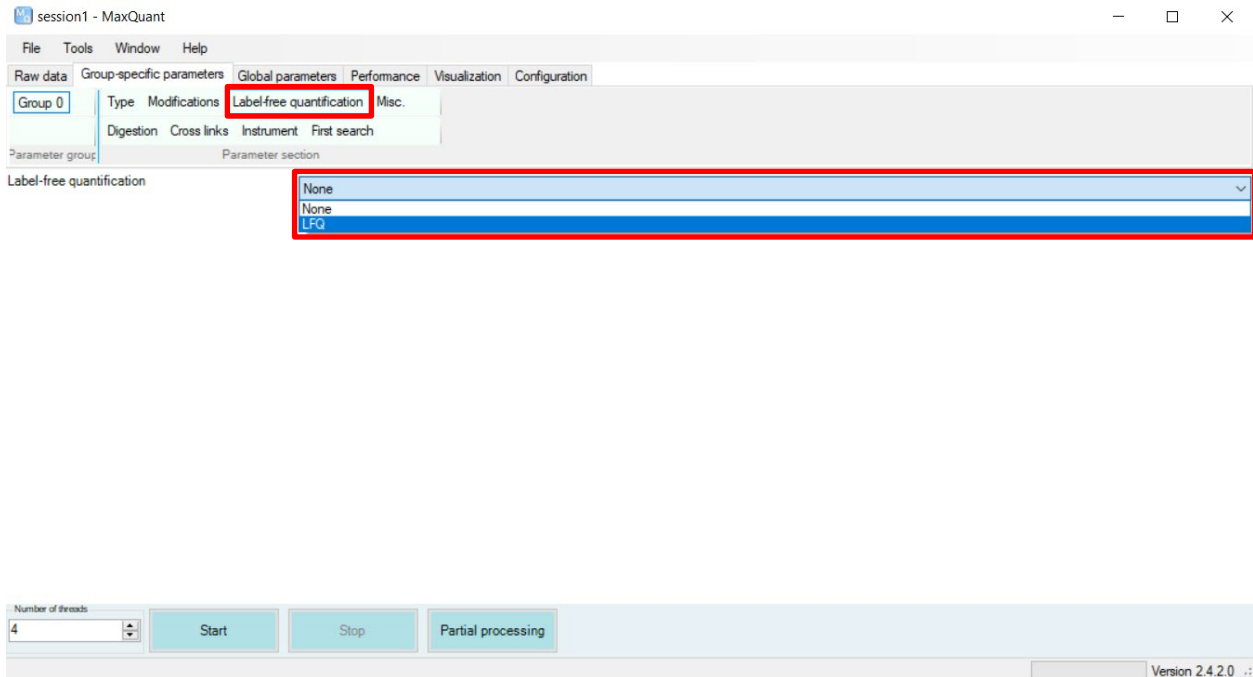
7. Move to the "Global parameters" tab. Here, in the "Sequences" section, you can choose the appropriate FASTA files for the data by clicking "Add" button. You can download FASTA files for different organisms from the UniProt FTP server by going to the
https://ftp.uniprot.org/pub/databases/uniprot/current_release/knowledgebase/reference_proteomes/
webpage. For this task you can use human reference proteome located in the directory
`Data/Assignment_2_MaxQuant_DIA/fasta/` on your flash drive. There are two files here:
`UP000005640_9606.fasta` that contains curated human proteome from SwissProt database and
`UP000005640_9606_additional.fasta` that contains human protein isoforms from TrEMBL database. Use both of them.
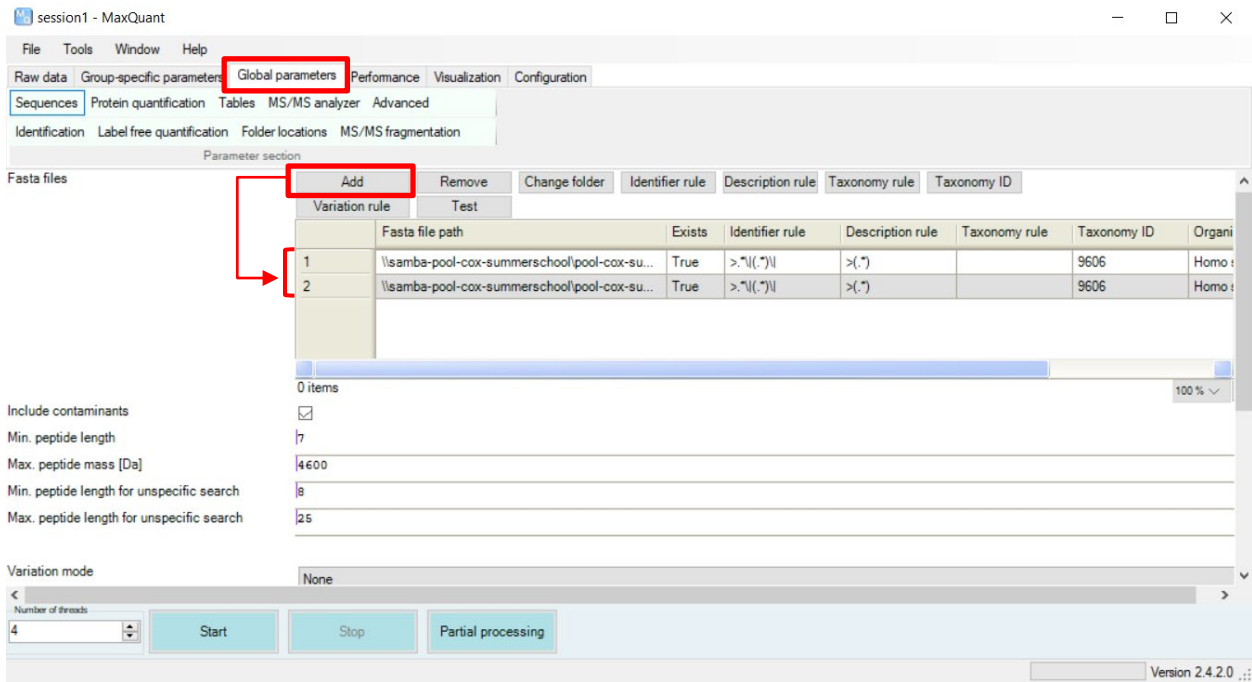
8. Click the "Start" button to start the MaxQuant run.
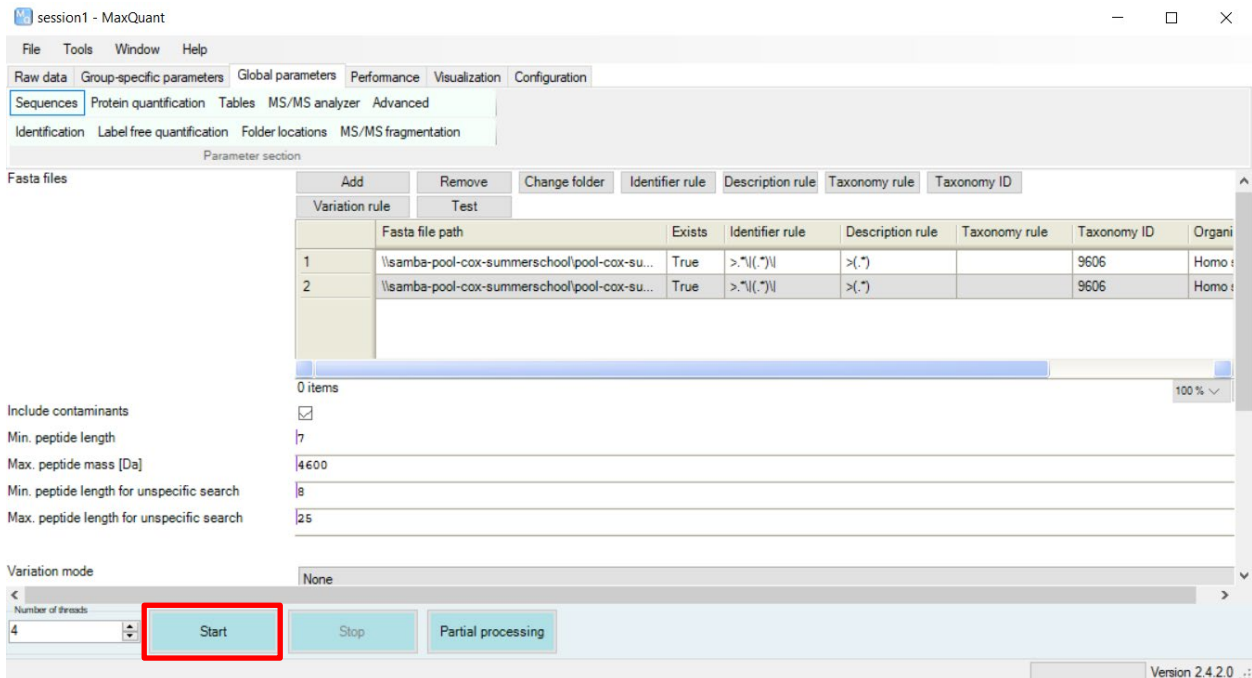
2. Running MaxDIA with MS1-labeled data

Objective: process the data from SILAC experiment obtained with Orbitrap DIA.

Data source: **Pino, L.K., Baeza, J., Lauman, R., Schilling, B. and Garcia, B.A., 2021. Improved SILAC quantification with data-independent acquisition to investigate bortezomib-induced protein degradation. Journal of proteome research, 20(4), pp.1918-1927** (https://doi.org/10.1021/acs.jproteome.0c00938). For the assignment purposes only two raw files are selected.

1. Launch MaxQuant and click on the "Load folder" button to load mass spectrometry output data into MaxQuant. The data is located in the directory `Data/Assignment_2_MaxQuant_DIA/2_MS1_label_Orbitrap/` on your flash drive.
2. Set the experiment design and the number of threads to be utilized by MaxQuant.
3. Move to the "Group-specific parameters" tab and choose "MaxDIA" for the "Type".
4. Add `evidence.txt` and `msms.txt` library files into the corresponding boxes. The files are located in the directory `Data/Assignment_2_MaxQuant_DIA/2_MS1_label_Orbitrap/measured_library/` on your flash drive. Add also contaminant spectral library from `Data/Assignment_2_MaxQuant_DIA/2_MS1_label_Orbitrap/contaminants/`.
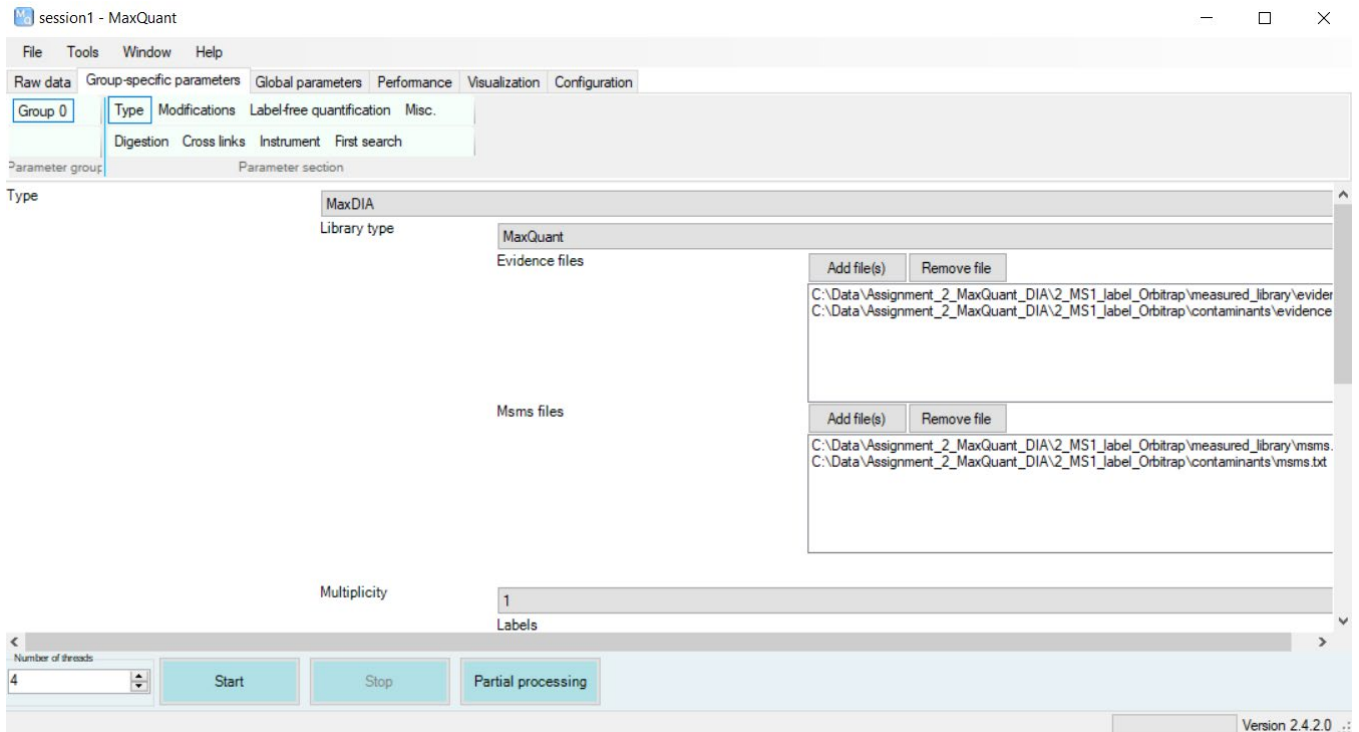
5. Scroll down and select "Multiplicity" 2 from the drop-down menu. Don't tick any boxes in the "Light labels" section, tick "Arg10" and "Lys8" in the "Heavy labels" section.
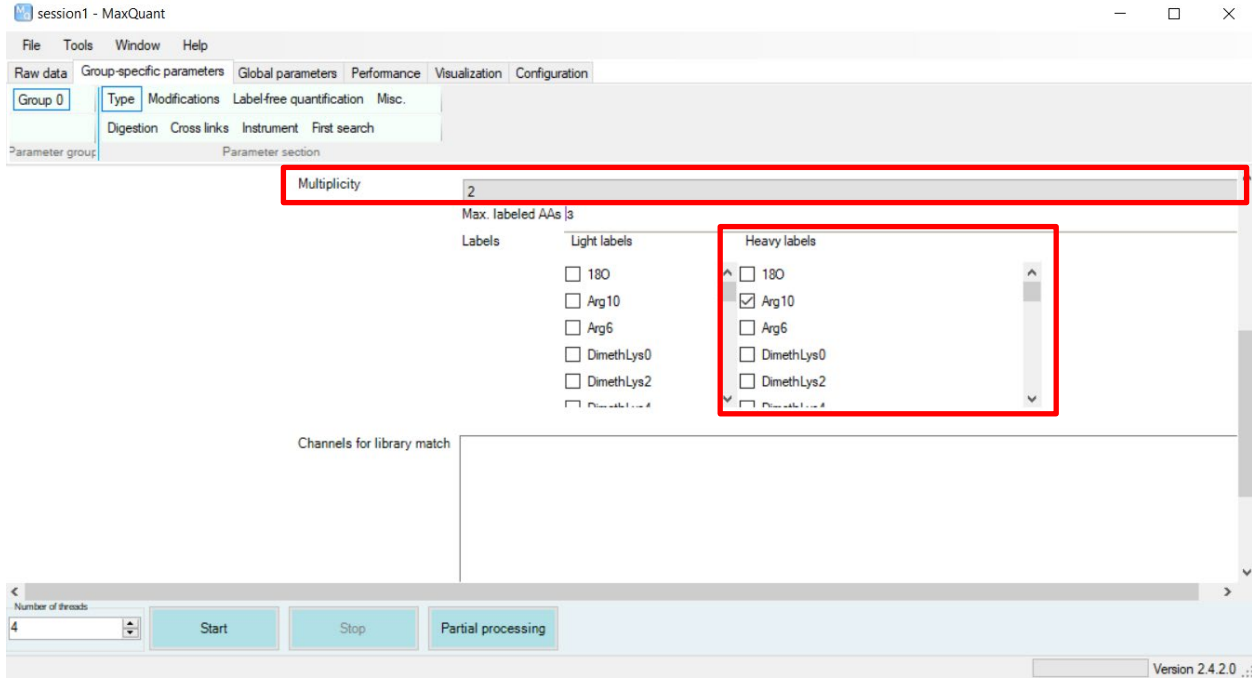
6. Specify channels for library match in the corresponding box. Write 0 (i.e. all channels) in the first row (the first, measured library) and 1 in the second row (the second library, contaminants). This way MaxQuant will search only light channel for contaminants.
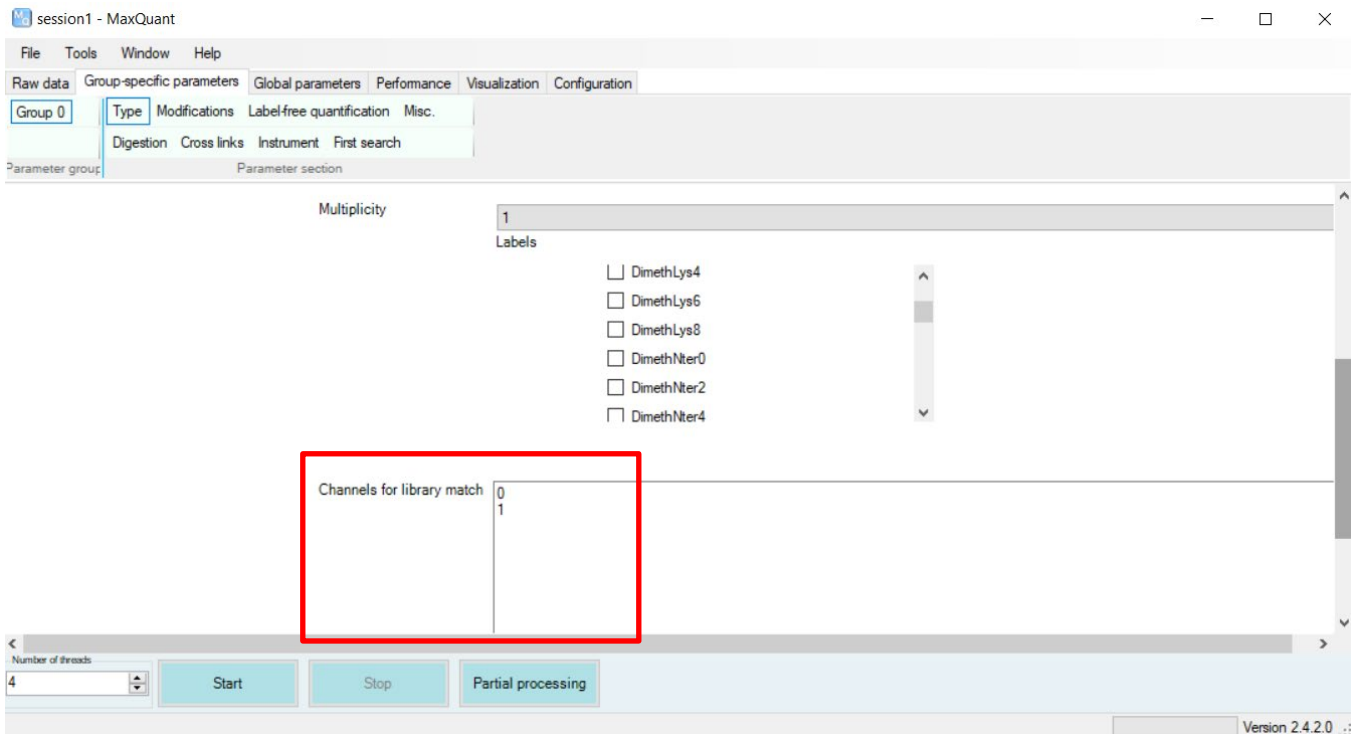
7. Go to the "Label-free quantification" section of the "Group-specific parameters" and select "LFQ" from the drop-down menu. Since SILAC was used in this paper to increase sample multiplicity and not precision, we need MaxLFQ to be able to compare samples. MaxLFQ will account for different raw files as well as multiplexes within each single file.

8. Move to the "Global parameters" tab and add human reference proteome located in the directory `Data/Assignment_2_MaxQuant_DIA/fasta/` on your flash drive. Add also FASTA file with contaminant sequences `contaminants.fasta`.

9. Click the "Start" button to start the MaxQuant run.

3. Running TIMS-MaxDIA with label-free data in discovery mode

Objective: process the data from label-free experiment obtained with timsTOF DIA with predicted libraries.

Data source: **Demichev, V., Szyrwiel, L., Yu, F., Teo, G.C., Rosenberger, G., Niewienda, A., Ludwig, D., Decker, J., Kaspar-Schoenefeld, S., Lilley, K.S. and Mülleder, M., 2022. dia-PASEF data analysis using FragPipe and DIA-NN for deep proteomics of low sample amounts. Nature communications, 13(1), p.3944** (https://doi.org/10.1038/s41467-022-31492-0). For the assignment purposes only two raw files are selected.

1. Discovery mode denotes running MaxDIA with predicted libraries instead of measured libraries. Predicted libraries required for this task have already been downloaded to your flash drive and are located in the directory `Data/Assignment_2_MaxQuant_DIA/predicted_library/`. If you want to download them yourself, you can go to http://annotations.perseus-framework.org/. Navigate to "DiscoveryLibraries" folder then select the organism of your interest.

Note: When you download the predicted library, you will also find FASTA files of the proteome of the selected organism in the `.zip` folder. It is important to use specifically these FASTA files to run DIA in discovery mode, otherwise you will run into conflicts between the library and the reference proteome. In this assignment you should use the same human proteome files as in tasks before, located in the `Data/Assignment_2_MaxQuant_DIA/fasta/`.

2. Launch MaxQuant and click on the "Load folder" button to load mass spectrometry output data into MaxQuant. The data is located in the directory `Data/Assignment_2_MaxQuant_DIA/3_LFQ_TIMS/` on your flash drive.

3. Set the experiment design and the number of threads to be utilized by MaxQuant.

4. Move to the "Group-specific parameters" tab and choose "TIMS-MaxDIA" for the "Type".
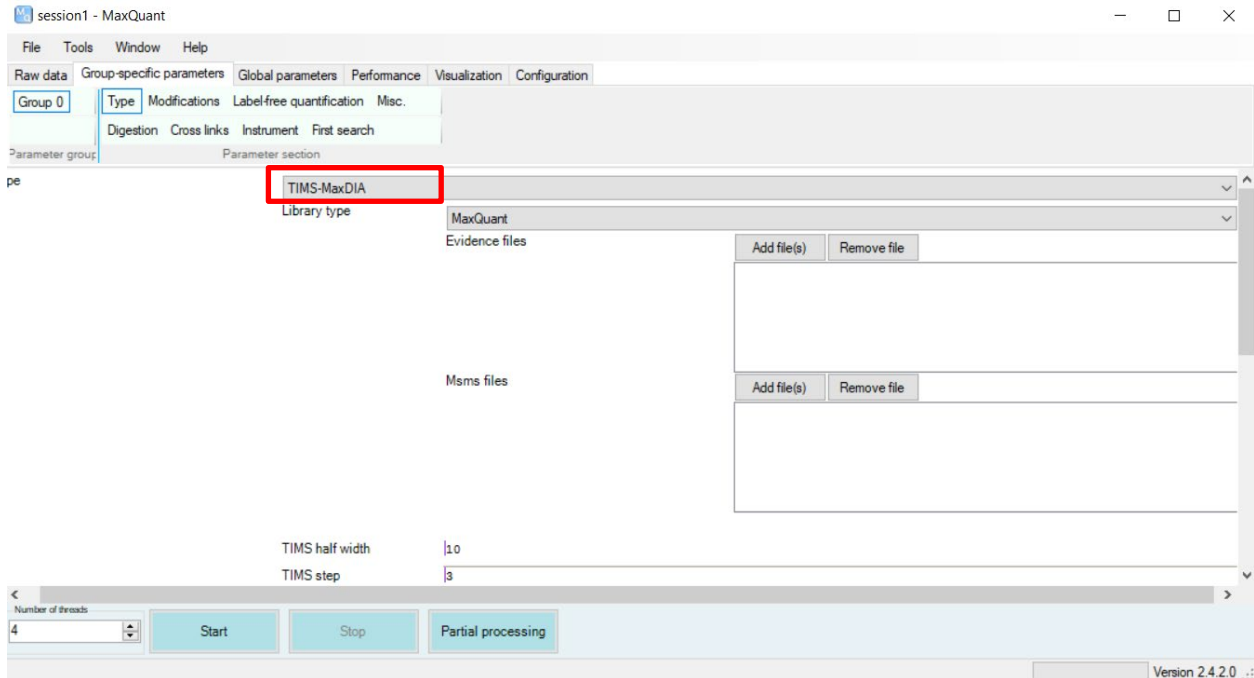
5. Add `evidence.txt` and `msms.txt` library files into the corresponding boxes. The files are located in the directory `Data/Assignment_2_MaxQuant_DIA/predicted_library/` on your flash drive.
6. Go to the "Label-free quantification" section of the "Group-specific parameters" and select "LFQ" from the drop-down menu.
7. Move to the "Global parameters" tab and add human reference proteome located in the directory `Data/Assignment_2_MaxQuant_DIA/fasta/` on your flash drive.
8. Click the "Start" button to start the MaxQuant run.