Please reference the following manuscript:

Rommelfanger, S.R., Zhou, M., Shaghasi, H., Tzeng, S., Evans, B.S., Umen, J.G., Pesavento, J.J.: An Improved Top-Down Mass Spectrometry Characterization of Chlamydomonas reinhardtii Histones and Their Post-translational Modi fi cations. (2021). https://doi.org/10.1021/jasms.1c00029

Before you run the Python script, you must have installed the following Python packages: (1) matplotlib (2) numpy (3) tkinter. You also must know how to use MSCONVERT (Kessner 2008, *Bioinformatics*) to generate an .mzML file from the Thermo .raw file. The .mzML file is then deconvoluted by FASTDeconv (Jeong, 2020 *Cell Syst.*) to generate an MSALIGN text file.

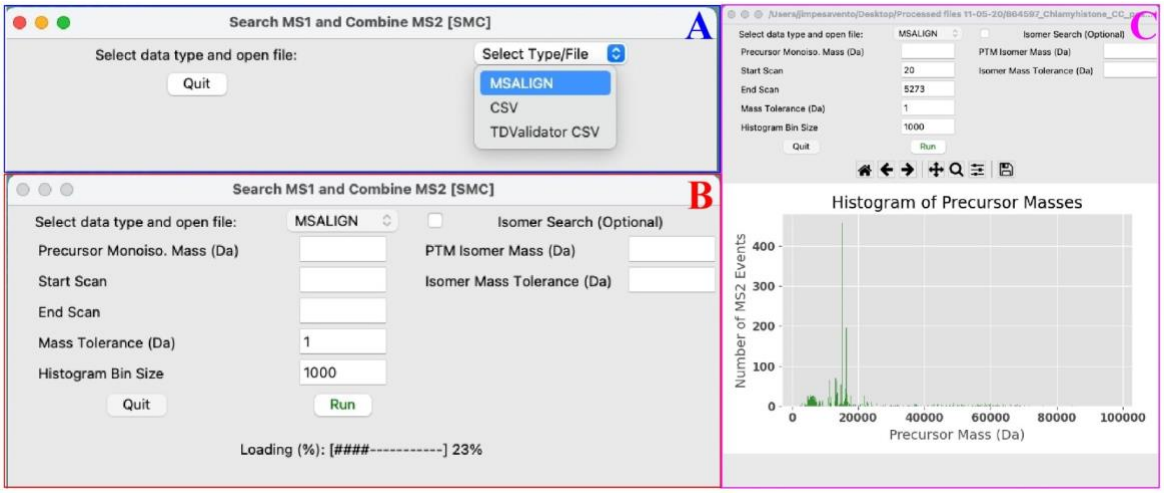

Figure 1. Loading and MSALIGN file into the SMC program. (A) Selection of the MSALIGN file type will open a window for file selection. (B) Depending on the size of the file selected, loading may take several seconds. (C) Once loaded, several search parameter fields are enabled and a histogram of precursor masses selected for MSMS appears below.

Using SMC to find MSMS data for a precursor mass of interest. First, click the 'Select Type/File' drop-down menu and select MSALIGN (Fig. 1A) . A new window should pop up and you should navigate to where the MSALIGN file with MSMS data is located. Once selected, the program will indicate the file is loading (Fig. 1B). Once loaded, the program automatically inputs the first and last scan number, a default setting of 1 Da for the mass tolerance, and a histogram bin size of 1000 (Fig. 1C).

Second, if you know the monoisotopic mass of the targeted precursor ion, you can enter it. Otherwise use the histogram toolbar and adjust the bin size to find a precursor ion. *It is important to note that the navigation toolbar (imported from the matplotlib package) is very buggy and you may need to click an icon several times for it to be selected/enabled.* The default bin size of 1000

pools masses over a ~94.5 Da window, while increasing that value to 10000 reduces it to ~9.4 Da (Fig. 2).

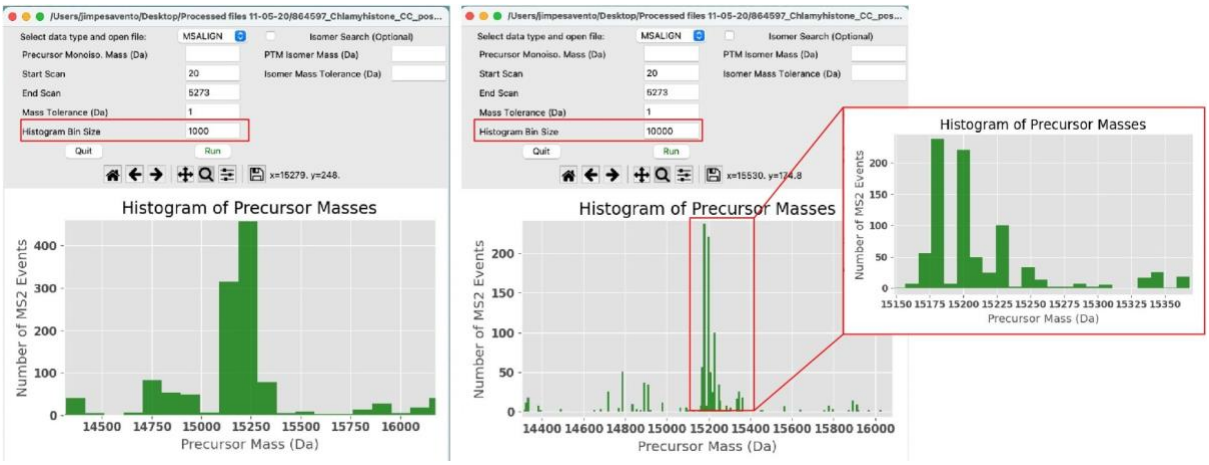

Figure 2. Adjusting the bin size from 1000 to 10000 (red box) provides finer resolution of masses when using the histogram to search for precursor masses to investigate.

Once a precursor mass is selected and the parameters set (Fig. 3A), the program will simultaneously display an output box containing information on the fragmentation information (Fig. 3B) and create a text file in the same directory as the Python script. The output display will have a window designated to the type of activation/fragmentation, in this case both ETD and HCD were performed on masses 11391 +/- 1 Da. For each type of activation, vital information including the precursor mass queried, the scan numbers with MSMS spectra for that precursor mass, the actual precursor mass values, and a list of fragment ions are generated. The text within the output display is selectable and can be copy-pasted into any text editing programs. Shown above this information is the unique filename associated with the output data.

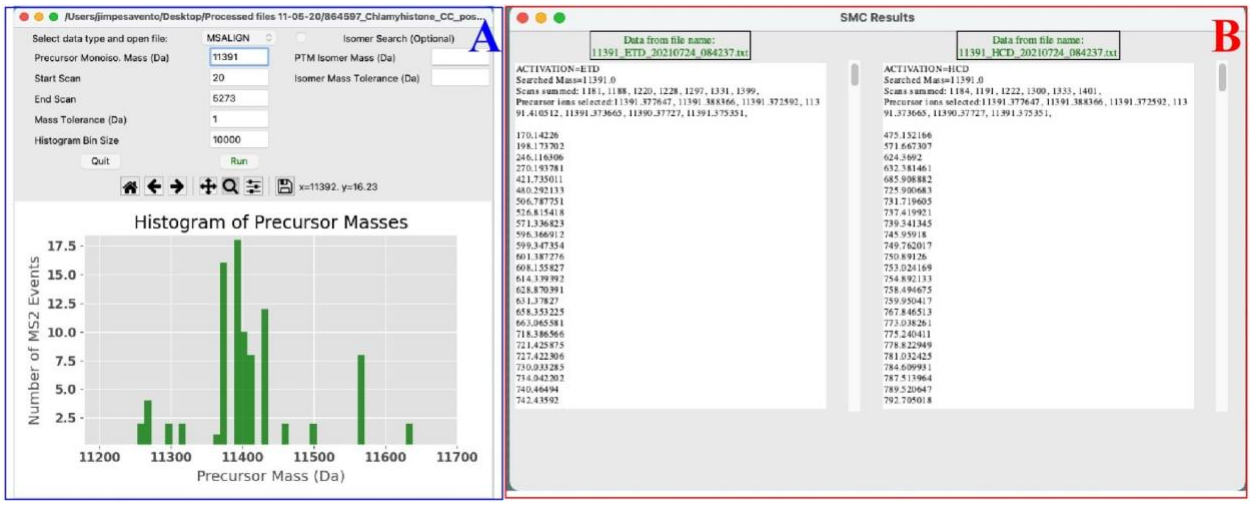

Figure 3. Query of precursor mass 11391 generates MSMS information. (A) Search parameters for 11391 Da included the entire LCMS run length with a mass tolerance of +/- 1 Da. (B) The precursor mass of 11391 +/- 1 Da was selected for both ETD (left panel) and HCD (right panel).

Other considerations. The following are a list of tips and precautions when using the SMC software. (1) It is well known that deconvolution programs, such as FLASHDeconv, are very reliable but do occasionally generate an off-by-one-dalton mass discrepancy. Therefore, increasing the mass tolerance to 2 or 3 Da may reveal additional masses that are from the same precursor mass. (2) You should be familiar with the elution window of your protein/proteoform of interest. Make sure the 'Scans summed' section is in agreement with that elution window. In some cases, different proteins with masses that fall within the search criteria are selected. Usually, this is reflected by large gaps in clustered scan numbers (e.g., 1130, 1150, 1152, 2034, 2039, 2055). Adjusting the 'Start Scan' and 'End Scan' parameter will exclude undesired precursor masses.

Advance feature: 'Isomer Search'. If you believe there to be positional isomers of a known type (e.g., monomethylation), you can search for monoisotopic mass differences (e.g., 14.01 Da) reporting on that type. If successful, the program will create a text file ending in *_isomers.txt in the same directory as the Python script (there will be no popup window containing the data). The *_isomers.txt file is organized as follows: monoisotopic mass of fragment ion 1, monoisotopic mass of fragment ion 2, charge state, mass difference between 1 and 2, ratio of (intensity of fragment ion 1)/(intensity of fragment ion 1 + fragment ion 2). **It is extremely important to manually confirm each fragment ion pair with the MSMS spectra!** The program does NOT differentiate based on fragment ion type (e.g., c ions vs. y ions) and there will be instances of ions of one type pairing with a different type, creating a false positive.

Advance feature: 'TDValidator CSV'. This is a work in progress and will be updated later.