In principle, CHiMA incorporates two major optimizations for database search: 1) Using 50% fragment ion coverage instead of 1% FDR to do peptide-spectrum match (PSM) filtration; 2) Using lysine acetylation, lysine monomethylation, and arginine monomethylation as background variable modifications in addition to your target histone modification. Any database search engine can be used with the CHiMA strategy if these two features are supported.

The following tutorial takes ProLuCID (for peptide-spectrum matching) + DTASelect2.0 (for PSM filtration) as an example. A demo data with related output results can be found in <https://drive.google.com/drive/folders/1qEIvQ81Jpyzr0fv5ALeOGmhiIAVoB1UD?usp=sharing>

Please send comments or questions to Dr. Jinjun Gao by email jinjungao@uchicago.edu.

**1. Requirement:**

**A.** ProLuCID (<http://fields.scripps.edu/yates/wp/?page_id=821>) or ProLuCID GUI (https://github.com/bathyg/Prolucid\_gui)

**B.** DTASelect2.0 (<http://fields.scripps.edu/yates/wp/?page_id=816>)

Both of them were developed by John Yates’s lab and are freely accessible. Please refer to their websites for the installation and configuration of the software.

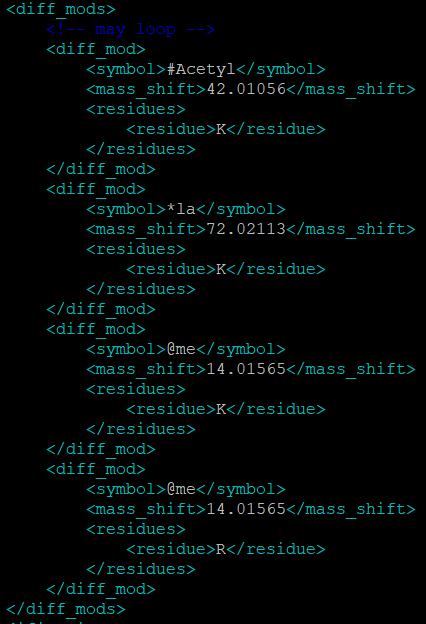
**2. Input:**

For LC-MS/MS data, ProLuCID takes \*.ms2 files as input, which can be easily converted from raw files by tools like RawConverter (<http://fields.scripps.edu/rawconv/>).

**3. Analysis workflow:**

**A.** Make a folder such as “test”, upload LC-MS/MS data in ms2 format.

**B.** Before the database search, you need to edit your parameter file to include Kac, Kme1, and Kme2 as background variable modifications. An example snapshot of the corresponding parameters is shown below.



**C.** Run ProLuCID with the set parameters. ProLuCID will generate all PSMs as \*.sqt file for DTASelect2.0 to do PSM filtration.

**D.** Under the same folder, run DTASelect2.0. Two parameters are needed to specifically set, “-i 0.5” and “--fp 1”, to realize the CHiMA strategy, among which “-i” is to set the lowest proportion of fragment ions observed and “--fp” is to set the false positive rate. Users can also personalize other parameters as they need.

**E.** DTASelect2.0 reads in all PSMs from sqt file, calculates the confident score for each PSM based on all the PSM scores, and summarizes them into a “DTASelect.txt” file. After filtration by the user-specified criterion, a “DTASelect-filter.txt” file should be generated to contain all the reported identifications.