User Manual For plon

Contents

1. Requirements

1) Computing system

pChem search requires a computer with recommended configuration as follows:

- ➢ Microsoft Windows 64-bit
- ➢ Intel Core i7/i9/Xeon Processor
- ➢ 32GB of RAM or more

Note: pIon v1.0 is NOT supported by non-Windows operating systems (incl. MacOS, Linux and so on).

2) MS Data

- ➢ Data dependent acquisition (DDA) with BOTH MS1 and MS/MS spectra recorded in the High-Resolution mode

Note: 1) For automatic performance assessment of chemoproteomic probes, it is recommended to acquire MS data from probe-labeled samples with DMP-tag.

2. Download

1) pIon can be freely downloaded from the website: http://pfind.org/software/pChem/index.html

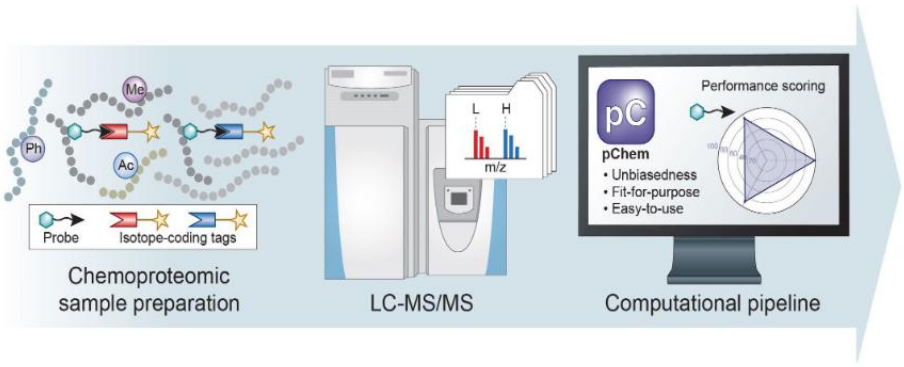
pFind Studio: a computational solution for mass spectrometry-based proteomics

pChem

Introduction - Cite us - Downloads

Introduction

Chemical probe coupled with mass spectrometry (MS)-based proteomics, herein termed chemoproteomics, offers versatile tools to globally profile protein features and to systematically interrogate the mode of action of small molecules in a native biological system. Nonetheless, development of an efficient and selective probe for chemoproteomics can still be challenging. Besides, it is also difficult to unbiasedly assess its chemoselectivity at a proteome-wide scale. Here we present pChem, a modification-centric blind search and summarization tool to provide a pipeline for rapid and unbiased assessing of the performance of ABPP and metabolic labeling probes. This pipeline starts experimentally by isotopic coding of PDMs, which can be automatically recognized, paired, and accurately reported by pChem, further allowing users to score the profiling efficiency, modification-homogeneity and proteome-wide residue selectivity of a chemoproteomic probe.



Cite us

pChem: a modification-centric assessment tool for performance of chemoproteomic probes.
Ji-Xiang He*, Zheng-Cong Fei*, Ling Fu, Cai-Ping Tian, Fu-Chu He, Hao Chi, Jing Yang.
Nat Chem Biol (2022).

https://doi.org/10.1038/s41589-022-01074-8

Downloads

pIon is currently available for free use.

Notice: Sep. 29, 2024 - The functionality of pChem 1.1 is also fully supported in this version. Click to download.

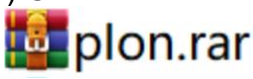
For source code, please refer to github.

For detailed usage, please refer to user guide.

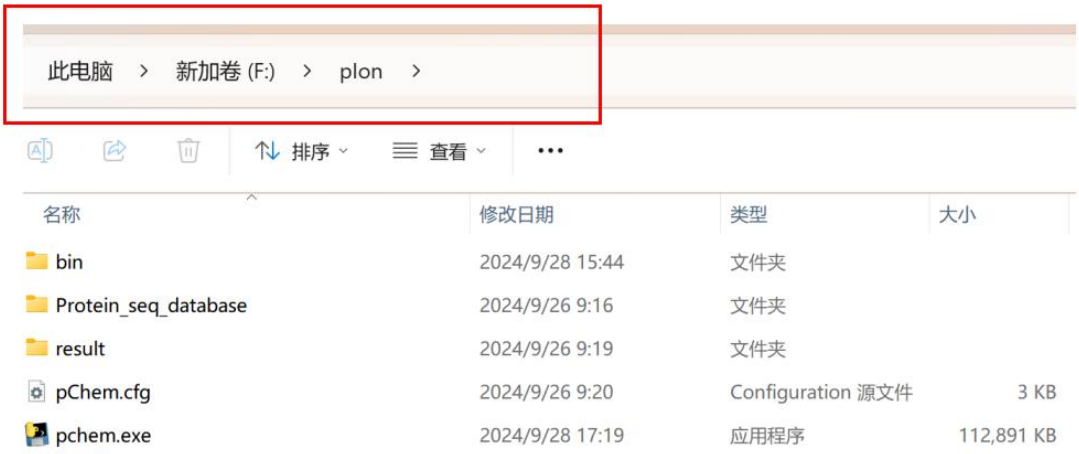
pChem version 1.1 is currently free to use.

Notice: Jan. 10, 2023 - If you are using a version prior to this date, please re-download the pChem software in time. The expiration date is set on Jan. 10, 2026.

2) Click to download, download the RAR compressed file.

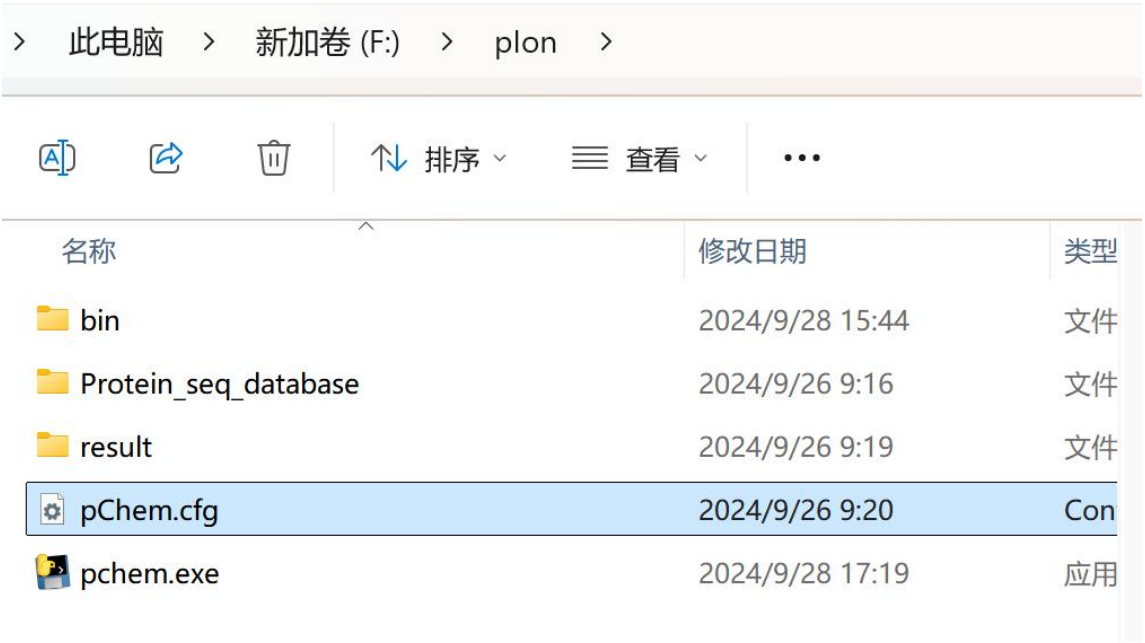


3) Un-zip the “plon.rar” package into a specified file folder (e.g., Local disk F).



3. Configuration

1) Double click *plon* to open the main folder.



2) Open configuration file "*pChem.cfg*" using a text editor, *e.g.*, Microsoft Notepad or Notepad++ (https://notepad-plus.en.softonic.com/).

3) Setting "pChem.cfg".

pChem.cfg

```
################################################################################
# PChem general parameter settings
# Path to the output file
output_path=F:\pIon\result

# Path to the protein sequence database
fasta_path=F:\pIon\Protein_seq_database\Homo_sapiens_uniprot_canonical_20395_entries_20210516.fasta

# Format of MS data, RAW or MZML
msmstype=RAW

# The number and path of MS data
msmsnum=1
msmspath1=G:\pChem_ion-2023-4\pChem-main\20230720\HFX_YangJing_HeJiXiang_IPM_20230701_F1_R1.raw


# Type of MS dissociation method
activation_type=HCD-FTMS

# Usage of open search (True/ False), against Unimod, the common modification can be set if not
open_flag=False
common_modification_number=2
common_modification_list=Carbamidomethyl[C];Oxidation[M];

# Mass range of unknown modification (Da)
min_mass_modification=200
max_mass_modification=1000

# Isotopic pairs of mass shifts with PSMs less than X% of that of overall PDMs were neglected
filter_frequency=5

# If consider the N-side or C-side for amino acid localization (True or False)
side_position=True

# P-value threshold enabling confident amino acid localization
p_value_threshold=0.001

# If report the statistical information (True or False)
report_statistics=True

################################################################################
# If isotope coding is adopted to facilitate the discovery of unknown modifications (True or False)
isotope_labeling=False


# Mass tolerance of the mass shift between light isotope and heavy isotope
mass_of_diff_diff=6.020132


# Isotopic mass difference within empirically defined tolerance(Da)
mass_diff_diff_range=0.005


################################################################################
# If ion labeling is adopted to facilitate the discovery of unknown modifications (True or False)
ion_labeling=True

# One charge mass of ion, it is recommended to keep at least three decimal places
ion_mass=126.128

# In the 0-1 range, the higher the score, the stricter the filtering, and the recommended value is 0.7
ion_filter_ratio=0.7

ion_close_search=False
```

① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩ ⑪ ⑫ ⑬

General Note 1:

For the first-time users, custom settings are required for ①-⑤, ⑧ default settings can be adopted for ⑥, ⑦, ⑨-⑭.

General Note 2:

All parameters (shown in red below) are case sensitive.

General Note 3:

The blank space should be avoided.

1. General Parameters Setting

---

① # Path to the output file

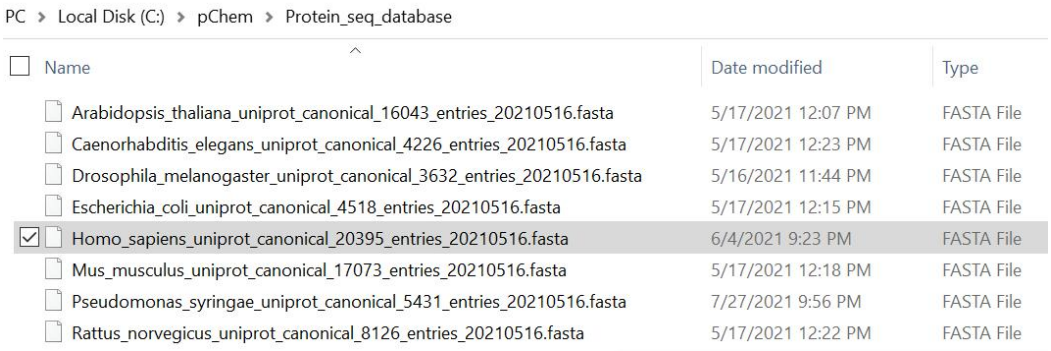
output_path=F:\pIon\result

Note: If the output file folder does not exist, an error will be reported.

② # Path to the protein sequence database

fasta_path=F:\pIon\Protein_seq_database\Homo_sapiens_uniprot_canonical_20395_entries_20210516.fasta

Note: The protein *.fasta database databases of several commonly used species (*e.g.*, *home sapiens*) are included in the subfolder (named as Protein_seq_database) of pChem. Note that the databases of other species can be downloaded from Uniprot as described in **Supporting Protocol 1**.



③ # Format of MS data (RAW or MZML)

msmstype=RAW

Note: Non-Thermo MS data need to be converted into mzML files before pChem search. The users can refer to **Supporting Protocol 2**.

④ # The number and path of MS data

msmsnum=N
msmspath1=X:\XXX\XXX.raw
msmspath2=X:\XXX\XXX.raw
……
msmspathN=X:\XXX\XXX.raw

Note: The suffix of MS data files MUST be input.

Example:

msmsnum=1
msmspath1=G:\data\HFX_YangJing_HeJiXiang_IPM_20230701_F1_R1.raw

⑤ # Type of MS dissociation method

activation_type=HCD-FTMS
illustration: default
Note: 1) pIon and pChem v1.0 can NOT support MS data generated under electron-transfer dissociation ETD, electron-transfer/higher-energy collision dissociation EThcD, and the likes.

⑥ # Usage of open search (True/ False) against Unimod, the common modification can be set if not
open_flag=False
common_modification_number=2
common_modification_list=Carbamidomethyl[C];Oxidation[M];
illustration: default
Note: The names of common modifications should be the same as those appeared in Unimod database. Specifically, you can refer to the modification.ini file in the bin directory.

⑦ # Mass range of unknown modification (Da)

min_mass_modification=200
max_mass_modification=1000
illustration: default
Note: The PDMs generated from the use of bioorthogonal cleavable linkers typically possess masses higher than 200 Da and less than 1000Da.

⑧ # Mass shifts with PSMs less than X% of that of overall PDMs were neglected
filter_frequency=5
illustration: default
Note: This parameter can be set as 0 if one wants to retrieve all PDMs including those with just a few PSMs.

⑨ # If consider the N- or C-termini for amino acid localization (True or False)

side_position=True

illustration: default

⑩ # P-value threshold enabling confident amino acid localization

p_value_ threshold=0.001

illustration: default

⑪ # if report the statistical information (True or False)

report_statistics=False

illustration: default

2. The parameter settings for pChem v1.0: if it is not in isotope mode, you can set *isotope_labeling* to False, and the remaining parameters are the same as in the previous version

---

⑫-1 # If isotope coding is adopted to facilitate the discovery of unknown modifications (True or False)

Isotope_labeling=False

illustration: default

Note: Choose 'False', if pChem is adopted to search endogenous modifications from probe-free and/or label-free protein samples, else, isotope coding is adopted to facilitate the discovery of unknown Modifications.

⑫-2 # Mass tolerance of the mass shift between light isotope and heavy isotope

mass_of_diff_diff=6.020132

Troubleshooting: One needs to confirm this value being correctly input.

⑫-3 # Isotopic mass difference within empirically defined tolerance (Da)

mass_diff_diff_range=0.005

illustration: default

Troubleshooting: If the pChem/pIon search mis-identified the targeted PDMs or even report nothing, one might want to loose the defined mass tolerance (e.g., 0.01Da).

3. The parameter settings for pIon: If ion labeling is adopted to facilitate the discovery of unknown modifications

---

⑬-1 # If ion labeling is adopted to facilitate the discovery of unknown modifications (True or False)

ion_labeling=True

⑬-2 # One charge mass of ion, it is recommended to keep at least three decimal places

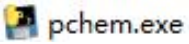
ion_mass=126.128

⑬-3 # In the 0-1 range, a higher score indicates stricter filtering, with a recommended value of 0.7.

ion_filter_ratio=0.7

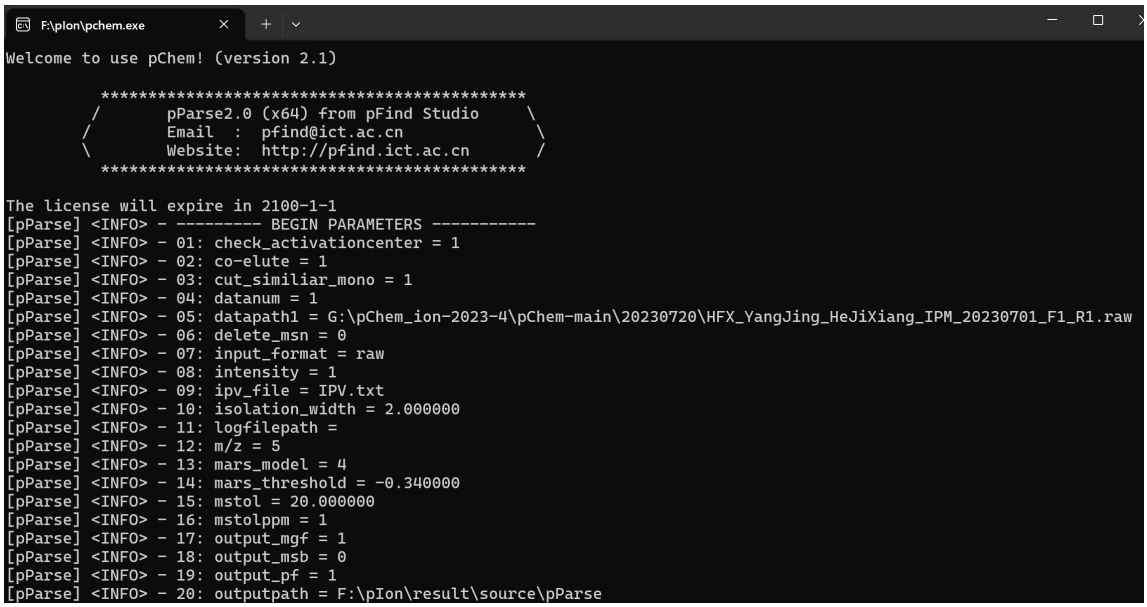
illustration: default

4. Run

Once all parameters have been set, double click “*pChem.exe*” 

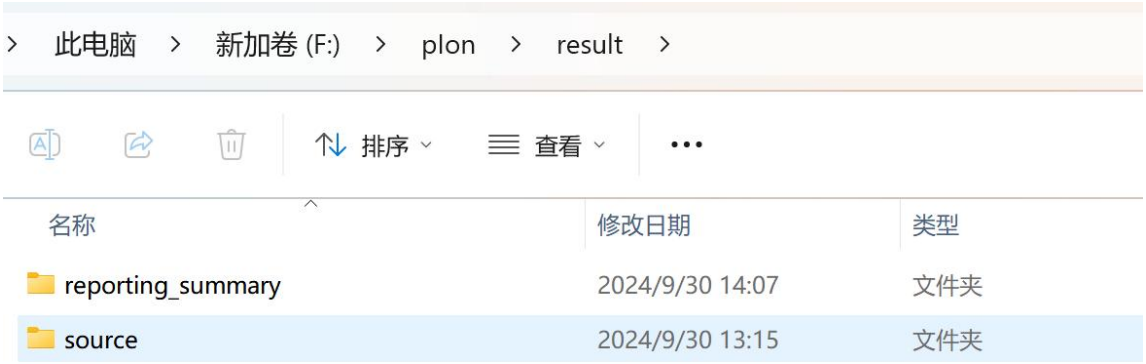
to execute the programming. The message “**Please press any key to continue**” means that program runs to completion.

Note: pChem search will generate several intermediate files in the main folder. do NOT open those files during program running.

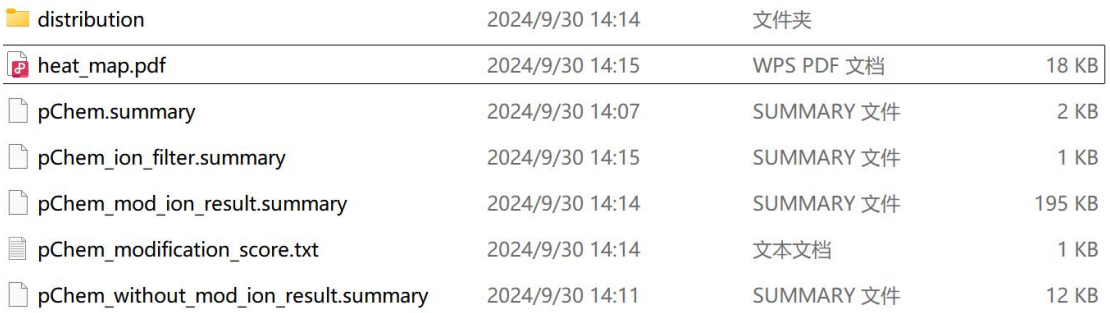


5. Output

1) Double click “*result*” file for searching results.
2) Double click “*reporting summary*”.



3) There are six major output documents.



Note: Users are recommended to copy these output documents and paste into another file. Otherwise, they can be covered by those generated from the next search event.

① pChem.summary

pChem.summary is a tab-delimited text file contains the details of every candidate mass shifts identified by blind search.

pChem.summary

Rank	Modification	Accurate Mass (std, r-squared)	Top1 Site | Probability | p-value	Others	#PSM
1	PFIND_DELTA_334	334.211936 (0.001863, 0.861375)	C|0.873|0.0000	N-SIDE(0.087, 0.0000);	4049
2	PFIND_DELTA_348	348.227660 (0.001815, 0.858382)	C|0.753|0.0000	N-SIDE(0.137, 0.0000); M(0.064, 0.0000);	299
3	PFIND_DELTA_350	350.206903 (0.002114, 0.857968)	C|0.475|0.0000	M(0.25, 0.0000);	160
4	PFIND_DELTA_668	668.424422 (0.00255, 0.80155)	C|0.849|0.0000		106
5	PFIND_DELTA_429	429.249373 (0.002699, 0.85648)	C|0.782|0.0000		78
6	PFIND_DELTA_332	332.196307 (0.002333, 0.861815)	N-SIDE|0.333|0.0001		33
7	PFIND_DELTA_366	366.203270 (0.003142, 0.854716)	C|0.633|0.0000		30
8	PFIND_DELTA_376	376.221914 (0.00201, 0.852762)	C|0.538|0.0000	N-SIDE(0.5, 0.0000);	26
9	PFIND_DELTA_335	335.214659 (0.002066, 0.861158)	C|0.684|0.0000		19
10	PFIND_DELTA_481	481.280325 (0.001364, 0.863835)	C|0.875|0.0000		16
11	PFIND_DELTA_286	286.211184 (0.000441, 0.872325)	M|0.727|0.0000		11
12	PFIND_DELTA_349	349.232975 (0.003745, 0.858172)	C|0.444|0.0001		9
13	PFIND_DELTA_239	239.129133 (0.00394, 0.83333)	N-SIDE|0.778|0.0000		9
14	PFIND_DELTA_333	333.194963 (0.002314, 0.861597)	C|0.625|0.0000		8
15	PFIND_DELTA_317	317.185354 (0.002377, 0.865154)	C|0.714|0.0000		7
16	PFIND_DELTA_362	362.243399 (0.00028, 0.855506)	M|0.667|0.0001		6
17	PFIND_DELTA_351	351.155199 (1.0, 0.848137)	C|0.833|0.0000	N-SIDE(0.667, 0.0002);	6
18	PFIND_DELTA_201	201.083955 (0.004766, 0.912972)	M|0.6|0.0007		5
19	PFIND_DELTA_990	990.504416 (0.001566, 0.910115)	N-SIDE|1.0|0.0007		5
20	PFIND_DELTA_247	247.117956 (0.000734, 0.81509)	C|1.0|0.0003		3

#PSM: The number of PSMs corresponding to modified peptides identified by search engine.

Top1 site | Top1 Probability: The amino acid most likely to be modified with the corresponding localization probability.

Others: Other amino acid sites that may also be labeled by probes and their corresponding localization probability values.

② pChem_ion_filter.summary

pChem_ion_filter.summary is a tab-delimited text file contains the details of every PDM.

pChem_ion_filter.summary

Rank	Modification	Accurate Mass (std, r-squared)	Top1 Site | Probability | p-value	Others	#PSM
1	PFIND_DELTA_334	334.211936 (0.001863, 0.861375)	C|0.873|0.0000	N-SIDE(0.087, 0.0000);	4049

PDM: Probe-derived modifications

#PSM: The number of PSMs corresponding to modified peptides identified by search engine.

Top1 site | Top1 Probability: The amino acid most likely to be modified with the corresponding localization probability.

Others: Other amino acid sites that may also be labeled by probes and their corresponding localization probability values.

③ pChem_mod_ion_result.summary & pChem_without_mod_ion_result.summary

Stores detailed information about all diagnostic ions in both modified and unmodified spectrum categories

pChem_mod_ion_result.summary

Modification info for PFIND_DELTA_334:(PSM = 4049) (filtered PSM = 3603)(neighbor filtered PSM = 2742)

Top 300 characteristic ions in the PFIND_DELTA_334 spectra:

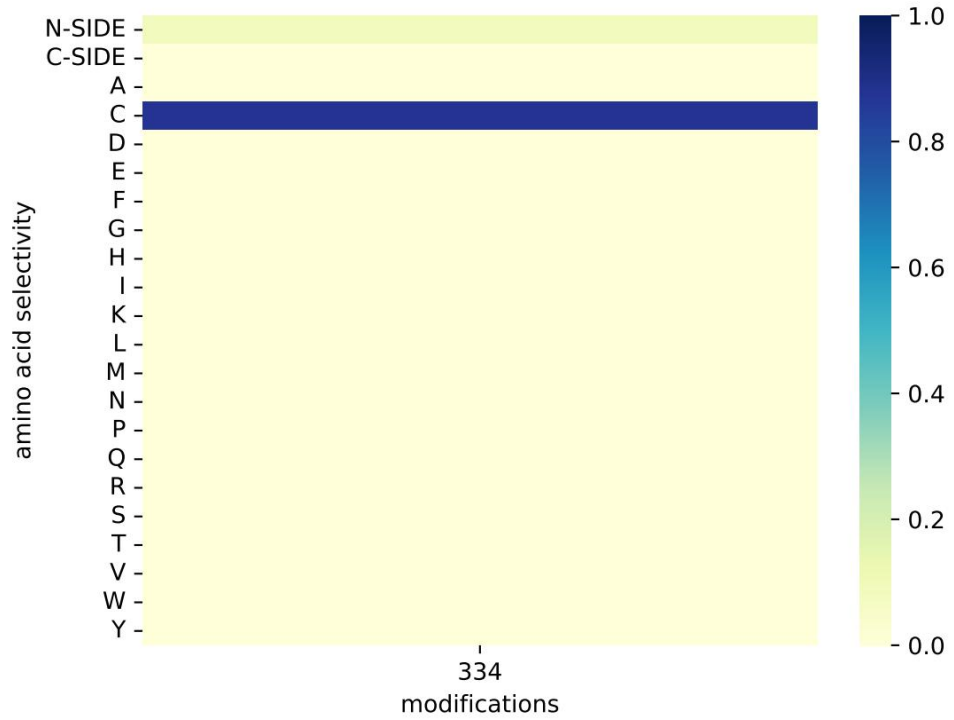
Rank	ion type	ion count	ion accuracy	ion relative peak
1	126.128	3601	126.128188	0.310392
2	120.082	3066	120.081296	0.224466
3	129.102	3434	129.102656	0.203585
4	369.208	2315	369.207955	0.158162
5	136.076	3332	136.076015	0.15778
6	110.072	3359	110.071808	0.106872
7	130.086	3125	130.086604	0.072734
8	201.124	2680	201.123842	0.06852
9	393.208	1322	393.20782	0.063615
10	197.166	2496	197.165395	0.056649
11	159.092	2107	159.091927	0.054452
12	175.12	1675	175.119505	0.052852
13	215.14	2025	215.139708	0.043908
14	143.118	2175	143.118128	0.039902
15	199.18	1354	199.180539	0.039817
16	157.134	1857	157.133865	0.039126
17	173.128	1600	173.128606	0.038781
18	185.166	1585	185.165482	0.036889
19	147.114	1089	147.113209	0.035626

pChem_without_mod_ion_result.summary

Top 300 characteristic ions in the unmodified spectra: (PSM: 76778)

Rank	ion type	ion count	ion accuracy	ion relative peak
1	120.082	70120	120.081292	0.314783
2	129.102	73115	129.10266	0.260202
3	136.076	72398	136.07603	0.200106
4	126.128	73949	126.128164	0.177484
5	110.072	72595	110.071824	0.138366
6	130.086	66916	130.08664	0.091424
7	201.124	60042	201.123831	0.086589
8	175.12	36117	175.119487	0.079598
9	159.092	50733	159.091952	0.062682
10	199.18	33608	199.1806	0.060605
11	185.166	40054	185.16547	0.060537
12	215.14	47334	215.139678	0.058668
13	173.128	37689	173.12864	0.054464
14	143.118	52514	143.118142	0.053285
15	147.114	26163	147.113201	0.052687
16	157.134	45676	157.133858	0.05072
17	183.15	47858	183.149721	0.049035
18	187.144	37573	187.144401	0.04799
19	169.134	49242	169.133869	0.043881
20	233.166	27740	233.165639	0.037374

④ **heat_map.pdf**



Horizontal coordinate: The ∆mass of each PDM

Longitudinal coordinate: The types of amino acids

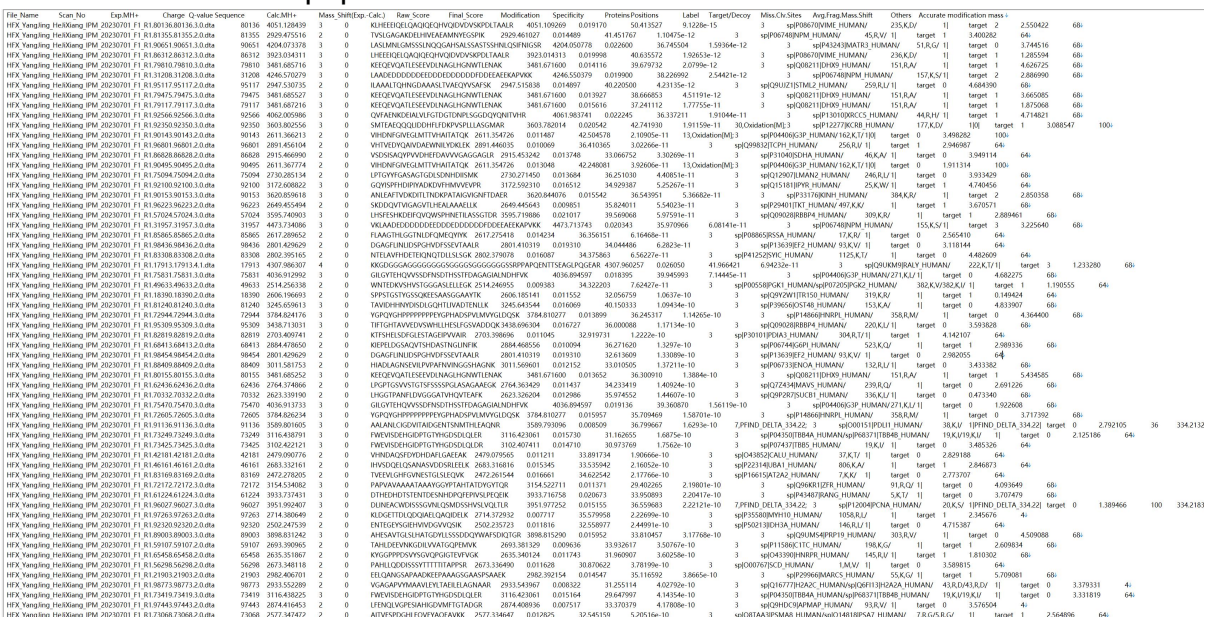
Color gradient: The localization probability that the modification occurs at each potential site.

Note: Those amino acids with p-value higher than p_value_ threshold (0.001 by default) are considered mis-localized sites. As such, their localization probability values are defined to be null. 2) For data generated from non-isotope-labeled or non-ion-labeled samples, heatmap will NOT be provided.

⑤ PSM-level results

F:\plon\result\source\blind\pFind-Filtered.spectra

This file contains the PSM-level information regarding all possible modifications on peptides.



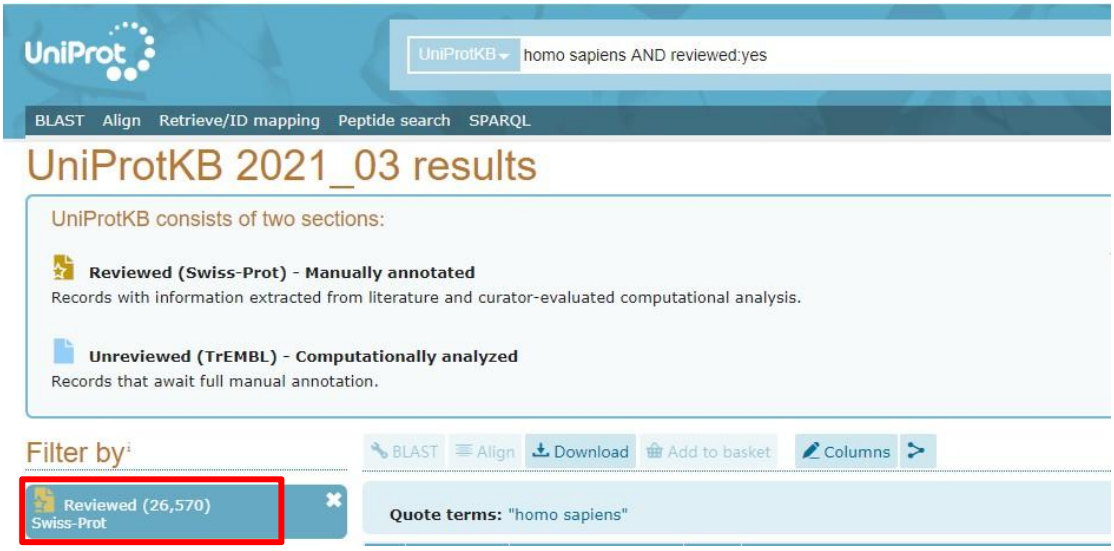
File_Name	Scan_No	Exp.MH+	Charge	Q-value	Sequence	Calc.MH+	Mass_Shift(Exp.-Calc.)	Raw_Score	Final_Score	Modification	Specificity	Proteins	Positions	Label	Target/Decoy	Miss.Cls.Sites	Avg.Frag.Mass.Shift	Others	Accurate modification mass s
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.80136.80136.3.0.dta	80136	4051.128439	3	0	KLHEEEIQELQAQIQEQHVQDVDVSKPDLTAALR	4051.109269	0.019170	50.413527	9.1228e-15	3	sp|P08670|VIME_HUMAN/	235,K,D/	1|	target	2	2.550422	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.81355.81355.2.0.dta	81355	2929.475516	2	0	TVSLGAGAMDELHNVEAEAMNYEGSPIK	2929.461027	0.014489	41.451767	1.10475e-12	3	sp|P06748|NPM_HUMAN/	45,R,V/ 1|	target	1	3.400282	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.90651.90651.3.0.dta	90651	4204.073378	3	0	LASLMNLGMSSSLNQQGAHSALSSASTSSHNLQSIFNIGSR	4204.050778	0.022600	36.745504	1.59364e-12	3	sp|P43243|MATR3_HUMAN/	51,R,G/ 1|	target	0	3.744516	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.86312.86312.3.0.dta	86312	3923.034311	3	0	LHEEEIQELQAQIQEQHVQDVDVSKPDLTAALR	3923.014313	0.019998	40.635572	1.92653e-12	3	sp|P06670|VIME_HUMAN/	236,K,D/	1|	target	1	1.285594	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.79810.79810.3.0.dta	79810	3481.685716	3	0	KEEQEVQATLESEEVDLNAGLHGNWTLENAK	3481.671600	0.014116	39.679732	2.0799e-12	3	sp|Q08211|DHX9_HUMAN/	151,R,A/	1|	target	1	4.626725	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.31208.31208.3.0.dta	31208	4246.570279	3	0	LAADEDDDDDEEEDDDDDDDDDDFDDEEAEEKAPVKK	4246.550379	0.019900	38.226992	2.54421e-12	3	sp|P06748|NPM_HUMAN/	157,K,S/ 1|	target	2	2.886990	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.95117.95117.2.0.dta	95117	2947.530735	2	0	ILAAALTQHNGDAAASLTVAEQVVSAFSK	2947.515838	0.014897	40.220500	4.23135e-12	3	sp|Q9UIZ1|STML2_HUMAN/	259,R,L/ 1|	target	0	4.684390	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.79475.79475.3.0.dta	79475	3481.685527	3	0	KEEQEVQATLESEEVDLNAGLHGNWTLENAK	3481.671600	0.013927	38.666853	4.51191e-12	3	sp|Q08211|DHX9_HUMAN/	151,R,A/	1|	target	1	3.665085	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.79117.79117.3.0.dta	79117	3481.687216	3	0	KEEQEVQATLESEEVDLNAGLHGNWTLENAK	3481.671600	0.015616	37.241112	1.77755e-11	3	sp|Q08211|DHX9_HUMAN/	151,R,A/	1|	target	1	1.875068	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.92566.92566.3.0.dta	92566	4062.005986	3	0	QVFAENKDEIALVLFGTDGTDNPLSGGDLQYQNITVHR	4061.983741	0.022245	36.337211	1.91044e-11	3	sp|P13010|KRCC5_HUMAN/	44,R,H/ 1|	target	1	4.714821	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.92350.92350.3.0.dta	92350	3603.802556	3	0	SMTEAEQQQLIDDHFLFDKPVSPLLLASGMAR	3603.782014	0.020542	42.741930	1.91159e-11	30,Oxidation[M]:3	sp|P12277|KCRB_HUMAN/	177,K,D/ 1|0|	target	1	3.088547	100ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.90143.90143.2.0.dta	90143	2611.366213	2	0	VIHDNFGVEGLMTTVHAITATQK 2611.354726	0.011487	42.504578	2.10905e-11	13,Oxidation[M]:3	sp|P04406|G3P_HUMAN/ 162,K,T/ 1|0|	target	0	3.498282	100ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.96801.96801.2.0.dta	96801	2891.456104	2	0	VHTVEDVQAVVDAVWNLYDKLEK 2891.446035	0.010069	36.410365	3.02266e-11	3	sp|Q99832|TCPH_HUMAN/	256,K,V/ 1|	target	1	2.946987	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.86828.86828.2.0.dta	86828	2915.466990	2	0	VSDSISAQYPVVDHEFDAVVVGAGGAGLR 2915.453242	0.013748	33.066752	3.30269e-11	3	sp|P31040|SDHA_HUMAN/	46,K,A/ 1|	target	0	3.949114	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.90495.90495.2.0.dta	90495	2611.367774	2	0	VIHDNFGVEGLMTTVHAITATQK 2611.354726	0.013048	42.240081	3.92606e-11	13,Oxidation[M]:3	sp|P04406|G3P_HUMAN/ 162,K,T/ 1|0|	target	0	1.911314	100ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.75094.75094.2.0.dta	75094	2730.285134	2	0	LPTGYYFGASAGTGDLSSINHDISNK 2730.271450	0.013684	36.251030	4.40851e-11	3	sp|Q12907|MAN2_HUMAN/	246,R,L/ 1|	target	0	3.933429	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.92100.92100.3.0.dta	92100	3172.608822	3	0	GQYISPFHDIPIYYADKDVFHMVVEVPR 3172.592310	0.016512	34.929387	5.25267e-11	3	sp|Q15181|PYR_HUMAN/	25,K,W/ 1|	target	1	4.740456	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.90153.90153.3.0.dta	90153	3620.859618	3	0	ANLEAFTVDKDITLTNDKPATAIGVIGNFTDAER 3620.844076	0.015542	36.543951	5.36682e-11	3	sp|P33176|KIN1_HUMAN/	384,K,R/	1|	target	2	2.850358	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.96223.96223.2.0.dta	96223	2649.455494	2	0	SKDDQVTVGACAVTLHEALAAAELK 2649.445643	0.009851	35.824011	5.54023e-11	3	sp|P29401|TKT_HUMAN/ 497,K,K/	1|	target	1	3.670571	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.57024.57024.3.0.dta	57024	3595.740903	3	0	LHSFESHKDEIFGVQWSPHNETILASSGTDR 3595.719886	0.021017	39.569068	5.97591e-11	3	sp|Q09028|RBBP4_HUMAN/	309,K,R/	1|	target	1	2.889461	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.83308.83308.2.0.dta	83308	2802.395165	2	0	NTELAVFHDETEKQNQTDLSLSLGK 2802.379078	0.016087	34.375863	6.56227e-11	3	sp|P41252|SYIC_HUMAN/	1125,K,T/	1|	target	0	4.482609	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.17913.17913.4.1.dta	17913	4307.986307	4	0	KKGDGGGAGGGGGGGGGSGGGGSGGGGGGSSRPPAPQENTTSEAGLPQGEAR 4307.960257	0.026050	41.966421	6.94232e-11	3	sp|Q9UKM9|RALY_HUMAN/	222,K,T/ 1|	target	3	1.233280	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.75831.75831.3.0.dta	75831	4036.912992	3	0	GILGYTEHQVVSSDFNSDTHSSTFDAGAGIALNDHFVK 4036.894597	0.018395	39.945993	7.14445e-11	3	sp|P04406|G3P_HUMAN/ 271,K,L/ 1|	target	0	4.682275	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.49633.49633.2.0.dta	49633	2514.256338	2	0	WNTEDKVSHVSTGGGASLELLEGK 2514.246955	0.009383	34.322203	7.62427e-11	3	sp|P00558|PGK1_HUMAN/sp|P07205|PGK2_HUMAN/	382,K,V/382,K,I/ 1|	target	1	1.190555	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.18390.18390.2.0.dta	18390	2606.196693	2	0	SPPSTGSTYGSSQKEESAASGGAAYTK 2606.185141	0.011552	32.056759	1.0637e-10	3	sp|Q9YZW1|TRI50_HUMAN/	319,K,R/	1|	target	1	0.149424	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.81240.81240.3.0.dta	81240	3245.659613	3	0	TAVIDHINVDISDLGQHTIVADTENLLK 3245.643544	0.016069	40.150333	1.09434e-10	3	sp|P39656|OST48_HUMAN/	153,K,A/	1|	target	0	4.833907	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.72944.72944.3.0.dta	72944	3784.824176	3	0	YGPQYGHPPPPPPPPPEYGPHADSPVLMVYGLDQSK 3784.810277	0.013899	36.245317	1.14265e-10	3	sp|P14966|PARP1_HUMAN/	358,R,M/	1|	target	0	4.364400	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.95309.95309.3.0.dta	95309	3438.713031	3	0	TIFTGHTAVVEDVSWHLLHESLFGSVADDQK 3438.696304	0.016727	36.000088	1.17134e-10	3	sp|Q09028|RBBP4_HUMAN/	220,K,L/ 1|	target	0	3.593828	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.82819.82819.2.0.dta	82819	2703.409741	2	0	KTFSHELSDFGLESTAGEIPVVAIR 2703.398696	0.011045	32.919731	1.2222e-10	3	sp|P30101|PDIA3_HUMAN/	304,R,T/ 1|	target	1	4.142107	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.68413.68413.2.0.dta	68413	2884.478650	2	0	KIEPELDGSAQVTSHDASTNGLINFIK 2884.468556	0.010094	36.271620	1.3297e-10	3	sp|P06744|G6PI_HUMAN/	523,K,Q/	1|	target	1	2.989336	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.98454.98454.2.0.dta	98454	2801.429629	2	0	DGAGFLINLIDSPGHVDFSSEVTAALR 2801.410319	0.019310	32.613609	1.33089e-10	3	sp|P13639|EF2_HUMAN/ 93,K,V/	1|	target	0	2.982055	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.88409.88409.2.0.dta	88409	3011.581753	2	0	HIADLAGNSEVILPVPAFNVINGGSHAGNK 3011.569601	0.012152	33.010505	1.37211e-10	3	sp|P06733|ENOA_HUMAN/	132,R,L/ 1|	target	0	3.433382	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.80155.80155.3.0.dta	80155	3481.685252	3	0	KEEQEVQATLESEEVDLNAGLHGNWTLENAK	3481.671600	0.013652	36.300910	1.38844e-10	3	sp|Q08211|DHX9_HUMAN/	151,R,A/	1|	target	1	5.434585	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.62436.62436.2.0.dta	62436	2764.374866	2	0	LPGFTGSVVSTGTSFSSSSFGLASAGAASGK 2764.363429	0.011437	34.233419	1.40924e-10	3	sp|Q7Z434|MAVS_HUMAN/	239,R,Q/	1|	target	0	2.691226	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.70332.70332.2.0.dta	70332	2623.339190	2	0	LHGGTPANFLDVGGGATVHQVTEAFK 2623.326204	0.012986	35.974552	1.44607e-10	3	sp|Q9P287|SUCB1_HUMAN/	336,K,L/ 1|	target	0	0.473340	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.75470.75470.3.0.dta	75470	4036.913733	3	0	GILGYTEHQVVSSDFNSDTHSSTFDAGAGIALNDHFVK	4036.894597	0.019136	39.360870	1.56119e-10	3	sp|P04406|G3P_HUMAN/ 271,K,L/ 1|	target	0	1.922608	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.72605.72605.3.0.dta	72605	3784.826234	3	0	YGPQYGHPPPPPPPPPEYGPHADSPVLMVYGLDQSK 3784.810277	0.015957	35.709469	1.58701e-10	3	sp|P14966|PARP1_HUMAN/	358,R,M/	1|	target	0	3.717392	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.91136.91136.3.0.dta	91136	3589.801605	3	0	AALANLCIGDVITADIGENTSNMTHEAQNR 3589.793096	0.008509	36.799667	1.6293e-10	7,PFIND_DELTA_334.22: 3	sp|O00151|PDLI1_HUMAN/	38,K,V 1|PFIND_DELTA_334.22|	target	0	2.792105	36	334.2132
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.73249.73249.3.0.dta	73249	3116.438791	3	0	FWEVISDEHGIDPTGTYHGDSDLQLER 3116.423061	0.015730	31.162655	1.6875e-10	3	sp|P04150|TBB4A_HUMAN/sp|P68371|TBB4B_HUMAN/	19,K,V/19,K,I/ 1|	target	0	2.125186	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.73425.73425.3.0.dta	73425	3102.422121	3	0	FWEVISDEHGIDPTGTYHGDSDLQLDR 3102.407411	0.014710	30.973769	1.7562e-10	3	sp|P07437|TBB5_HUMAN/	19,K,I/ 1|	target	0	3.485326	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.42181.42181.2.0.dta	42181	2479.090776	2	0	VHNDAQSFDYDHIDAFLGAEEAK 2479.079565	0.011211	33.891734	1.90666e-10	3	sp|O43852|CALU_HUMAN/	37,K,Y/ 1|	target	0	2.829188	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.46161.46161.2.0.dta	46161	2683.332161	2	0	IHVSDQELQSANASVDDSRLEELK 2683.316816	0.015345	33.535942	2.16052e-10	3	sp|P22314|UBA1_HUMAN/	806,K,A/	1|	target	1	2.846873	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.83169.83169.2.0.dta	83169	2472.278205	2	0	TVEEVLGHFGVNESTGLSLEQVK 2472.261544	0.016661	34.622542	2.17766e-10	3	sp|P16615|AT2A2_HUMAN/	7,K,K/ 1|	target	0	2.773707	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.72172.72172.3.0.dta	72172	3154.534082	3	0	PAPVAVAAAATAAAYGGPPTAHTATDYGYTQR 3154.522711	0.011371	29.402265	2.19801e-10	3	sp|Q96KR1|ZFR_HUMAN/	91,R,Q/ 1|	target	0	4.093649	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.61224.61224.3.0.dta	61224	3933.737431	3	0	DTHEDHDTISTENTDESNHDPQFEPVSLPEQEIK 3933.716758	0.020673	33.950893	2.20417e-10	3	sp|P43487|RANG_HUMAN/	5,K,T/ 1|	target	0	3.707479	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.96027.96027.3.0.dta	96027	3951.992407	3	0	DLNEACWDISSSGVNLQSMDSSHVSLVQLTLR 3951.977252	0.015155	36.559683	2.22121e-10	7,PFIND_DELTA_334.22: 3	sp|P12004|PCNA_HUMAN/	20,K,V 1|PFIND_DELTA_334.22|	target	0	1.389466	100	334.2183
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.97263.97263.2.0.dta	97263	2714.380649	2	0	KLDGETTDLQDQIAELQAQIDELK 2714.372932	0.007717	35.579958	2.22699e-10	3	sp|P35580|MYH10_HUMAN/	1058,R,L/ 1|	target	1	2.345676	4ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.92320.92320.2.0.dta	92320	2502.247539	2	0	ENTEGEVSGETHVVDGVVQSIK 2502.235723	0.011816	32.558977	2.44991e-10	3	sp|P50213|IDH3A_HUMAN/	146,R,L/ 1|	target	0	4.715387	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.89003.89003.3.0.dta	89003	3898.831242	3	0	AHESAVTGLSLHATGDYLSSSSDDQYWAFSDIQTGR 3898.815290	0.015952	33.810457	3.17768e-10	3	sp|Q9UMS4|PRP19_HUMAN/	303,R,V/ 1|	target	0	4.509088	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.59107.59107.2.0.dta	59107	2693.390965	2	0	TAHLDEEVNKGDILVVATGQPEMVK 2693.381329	0.009636	33.932617	3.50767e-10	3	sp|P11586|C1TC_HUMAN/	198,K,G/ 1|	target	1	2.609834	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.65458.65458.2.0.dta	65458	2635.351867	2	0	KYGGPPPDSVPSGVQPGIGTEVFVGK 2635.340124	0.011743	31.960907	3.60258e-10	3	sp|O43390|HNRPR_HUMAN/	145,R,V/ 1|	target	1	1.810302	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.56298.56298.2.0.dta	56298	2673.348118	2	0	PAHLLQDDSSSVTTTTITAPPSR 2673.336490	0.011628	30.870622	3.78199e-10	3	sp|O00767|SCD_HUMAN/	1,M,V/ 1|	target	0	3.589815	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.21903.21903.2.0.dta	21903	2982.406701	2	0	EELQANGSAPAADKEEPAAAGSGAASPSAAEK 2982.392154	0.014547	35.116592	3.8665e-10	3	sp|P29966|MARCS_HUMAN/	55,K,G/ 1|	target	1	5.709081	68ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.98773.98773.2.0.dta	98773	2933.552289	2	0	VGAGAPVYMAAVLEYLTAEILELAGNAAR 2933.543967	0.008322	31.255114	4.02792e-10	3	sp|Q16777|H2AC_HUMAN/sp|Q6FI13|H2AZA_HUMAN/	43,K,D/43,R,D/ 1|	target	0	3.379331	4ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.73419.73419.3.0.dta	73419	3116.438225	3	0	FWEVISDEHGIDPTGTYHGDSDLQLER 3116.423061	0.015164	29.647997	4.14354e-10	3	sp|P04150|TBB4A_HUMAN/sp|P68371|TBB4B_HUMAN/	19,K,V/19,K,I/ 1|	target	0	3.331619	64ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.97443.97443.2.0.dta	97443	2874.416453	2	0	LFENQLVGPESAHIGDVMFTGTADGR 2874.408936	0.007517	33.370379	4.17808e-10	3	sp|Q9HDC9|APMAP_HUMAN/	93,R,V/ 1|	target	0	3.576504	4ı
HFX_Yang\lng_HeliXiang_IPM_20230701_F1_R1.73068.73068.2.0.dta	73068	2577.347427	2	0	AITVFSPDGHLFQVEYAQEAVKK 2577.334647	0.012825	32.545159	5.20516e-10	3	sp|O8TAA1|IPSMA8_HUMAN/sp|O14818|PSA7_HUMAN/	7,R,G/5,R,G/ 1|	target	1	2.564896	64ı

6. Supporting protocol 1: Protein sequence database

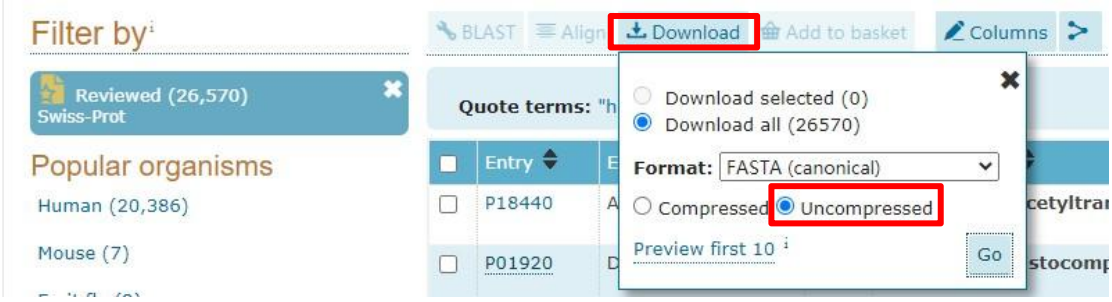
This protocol is used to download protein *.fasta files for database search.

1) Open https://www.uniprot.org/, enter the Latin name of the species (e.g., *home sapiens*), then click search.

2) Click "*Reviewed*" (Swiss-Prot).



3) Select "*Uncompressed*", then Click "*Download*" and "*Go*".

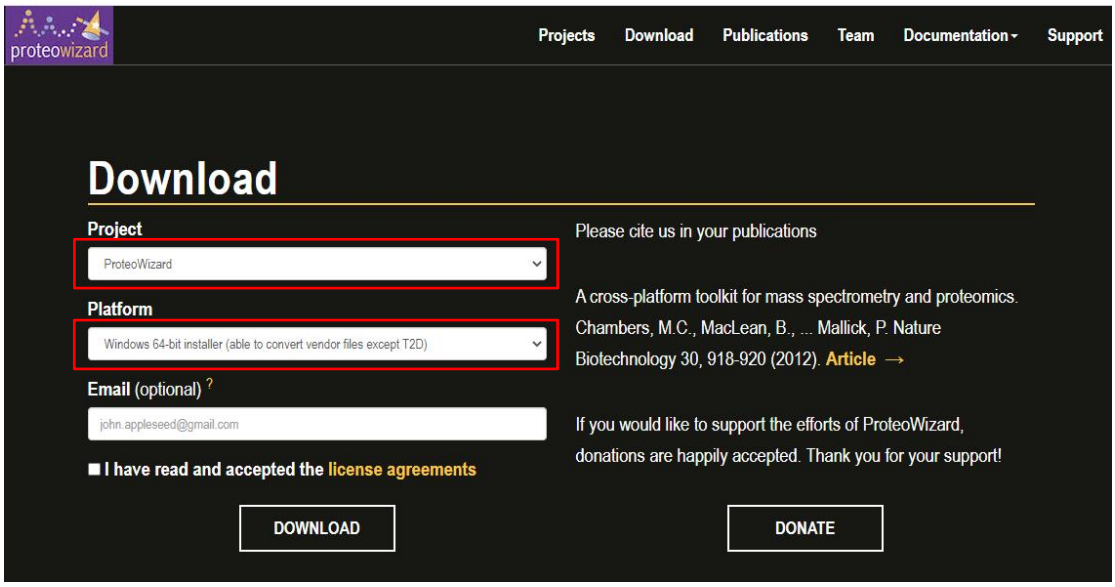


4) Get the *.fasta file.

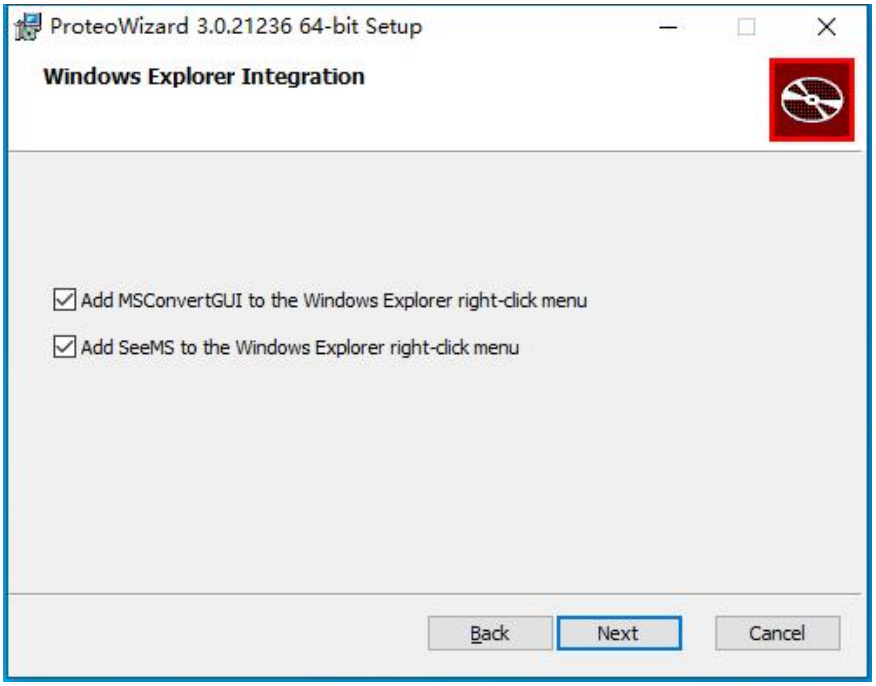
7. Supporting protocol 2: MSconvert

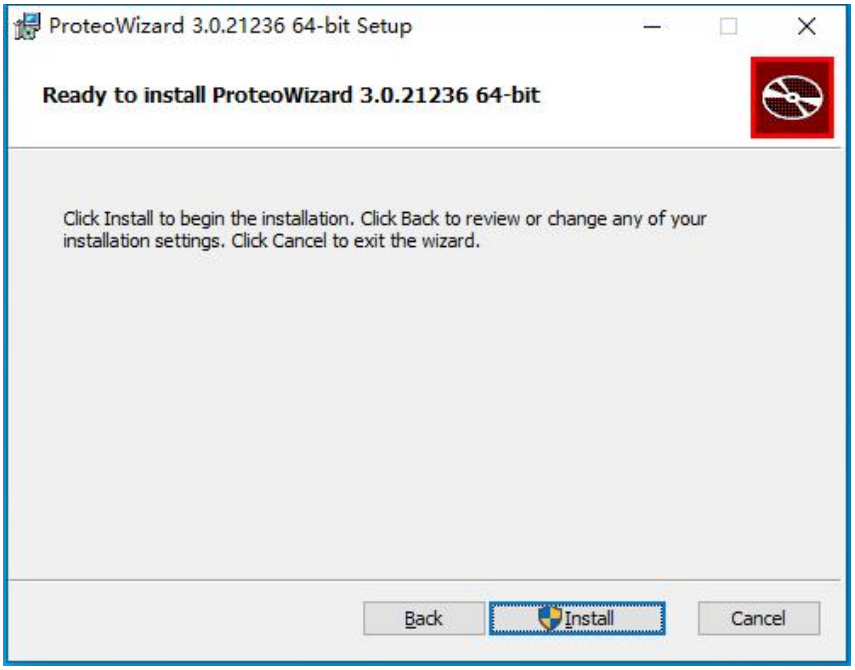
This protocol is used to convert non-Thermo MS data into mzML format files for pChem search.

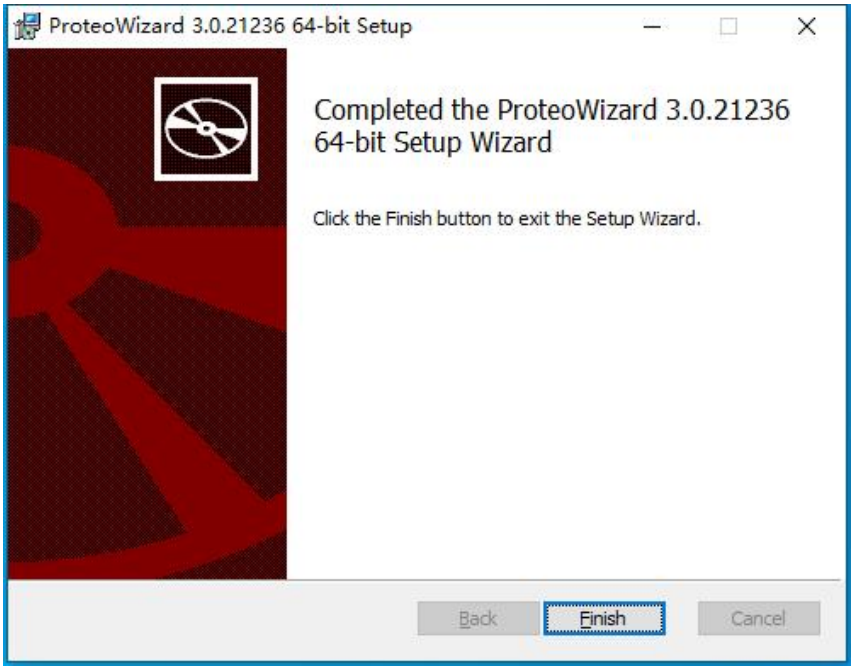
1) Download MSconvertGUI that is embedded in the ProteoWizard platform from: https://proteowizard.sourceforge.io/download.html .



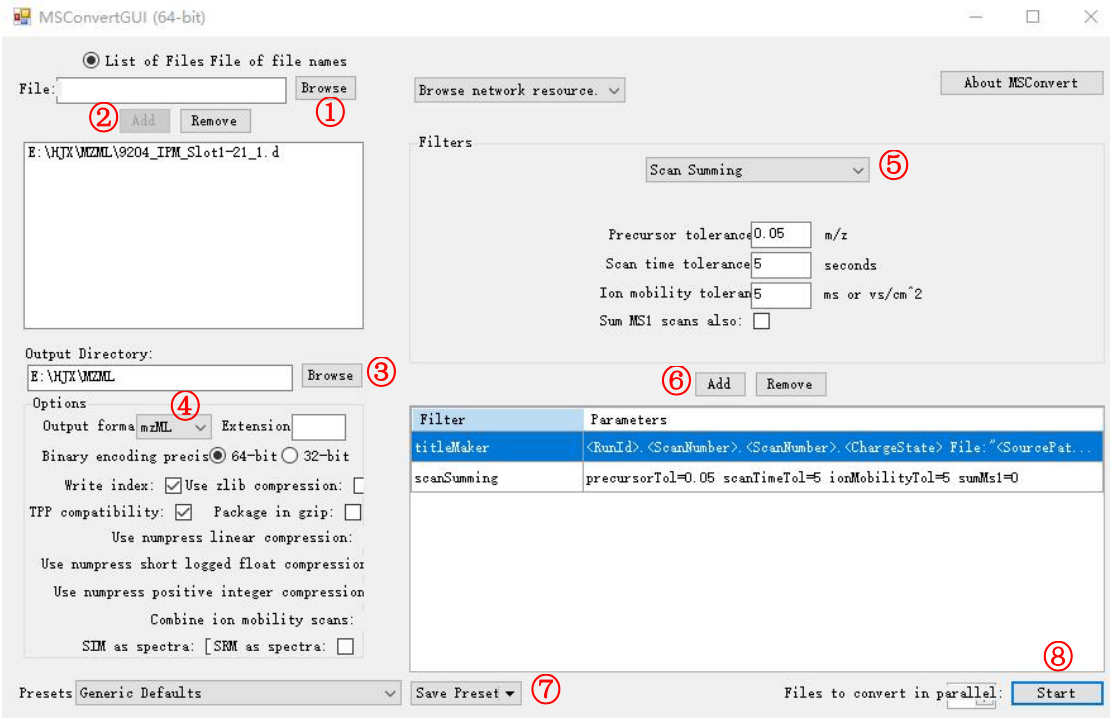
2) Install ProteoWizard according to the following instruction.




ProteoWizard 3.0.21236 64-bit Setup
Ready to install ProteoWizard 3.0.21236 64-bit
Click Install to begin the installation. Click Back to review or change any of your installation settings. Click Cancel to exit the wizard.
Back
Install
Cancel


ProteoWizard 3.0.21236 64-bit Setup
Completed the ProteoWizard 3.0.21236 64-bit Setup Wizard
Click the Finish button to exit the Setup Wizard.
Back
Finish
Cancel

3) Open MSconvertGUI



①-②Browse and add MS data (e.g., *.d, *.WIFF files)

③ Define output route

④ Choose *.mzML as the output data format

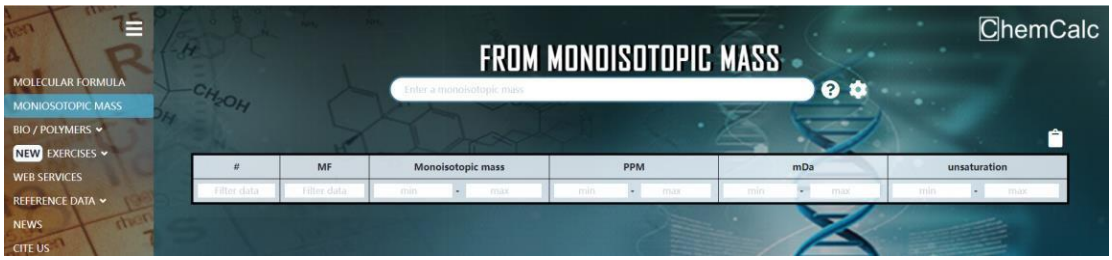
⑤ -⑥ Define parameters for Scan Summing

⑥ -⑧ Save and run

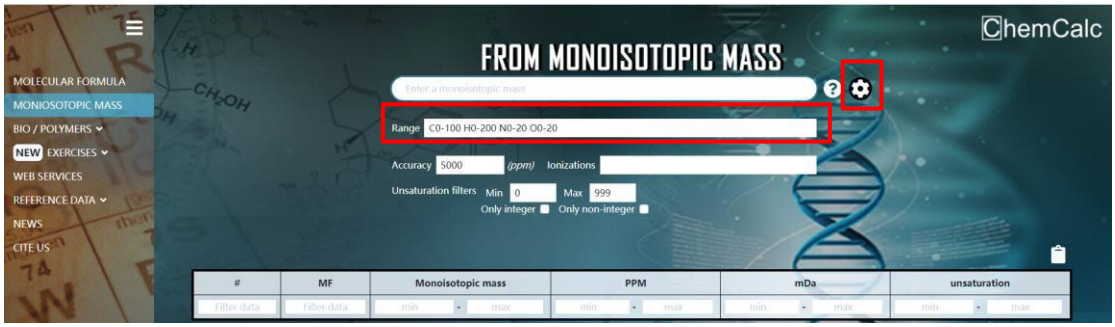
8. Supporting protocol 3: ChemCalc

This protocol is used to estimate candidate molecular formulas from the pChem-determined accurate masses.

1) Open **https://www.chemcalc.org/mf-finder** .



2) Click  , check the element composition.



3) Input the monoisotopic mass of each PDM shown in *pChem.summary* or *pChem_ion_filter.summary* file. The candidate molecular formulas will immediately appear below.

Rank	PDM	Accurate Mass	Top1 Site|Probability	Others	#PSM	#PSM L|H	DFLs
1	PFIND_DELTA_252	252.122339	C|0.988		13876	7368|6508