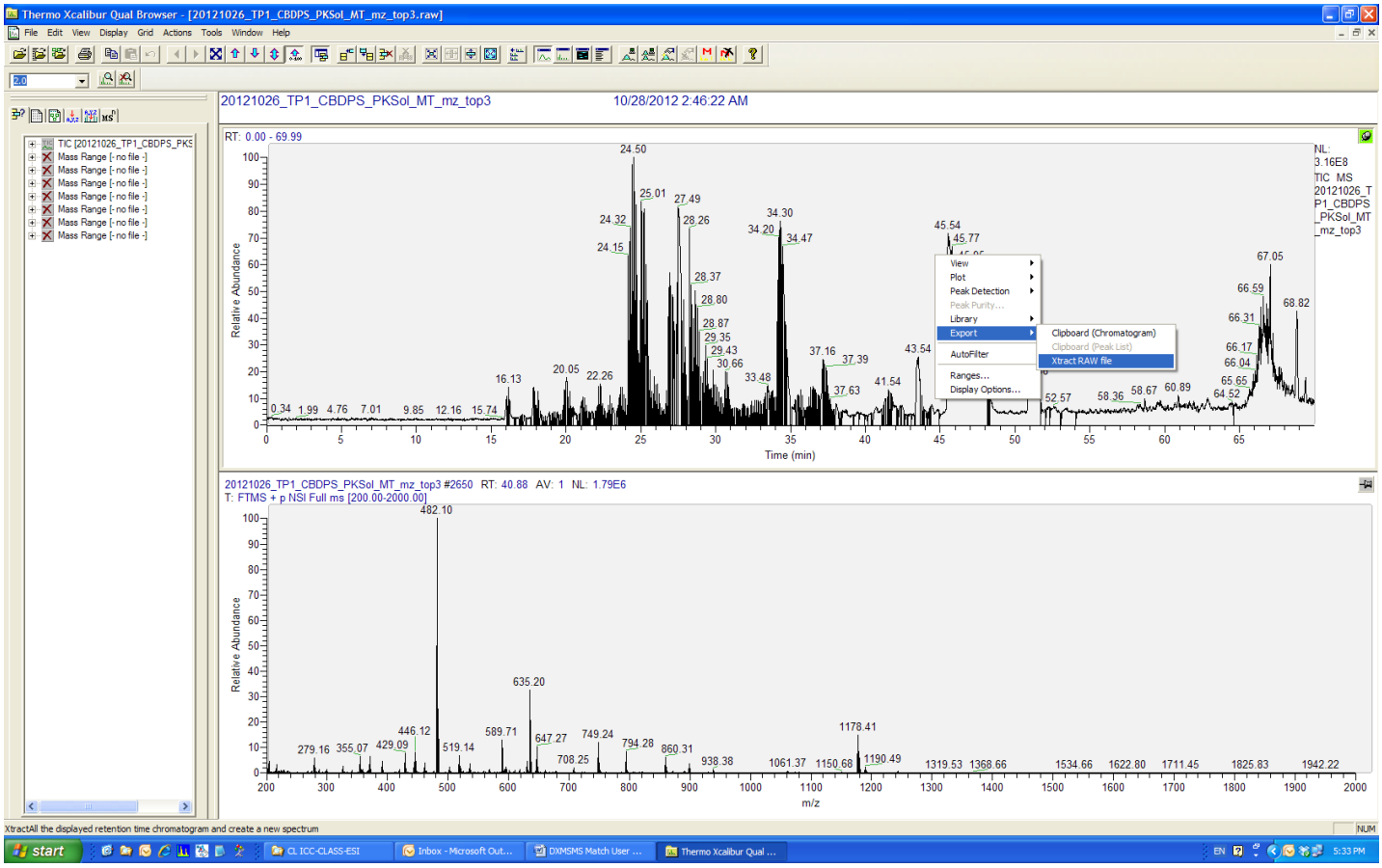
DXMSMS Match User Guide 20141124.

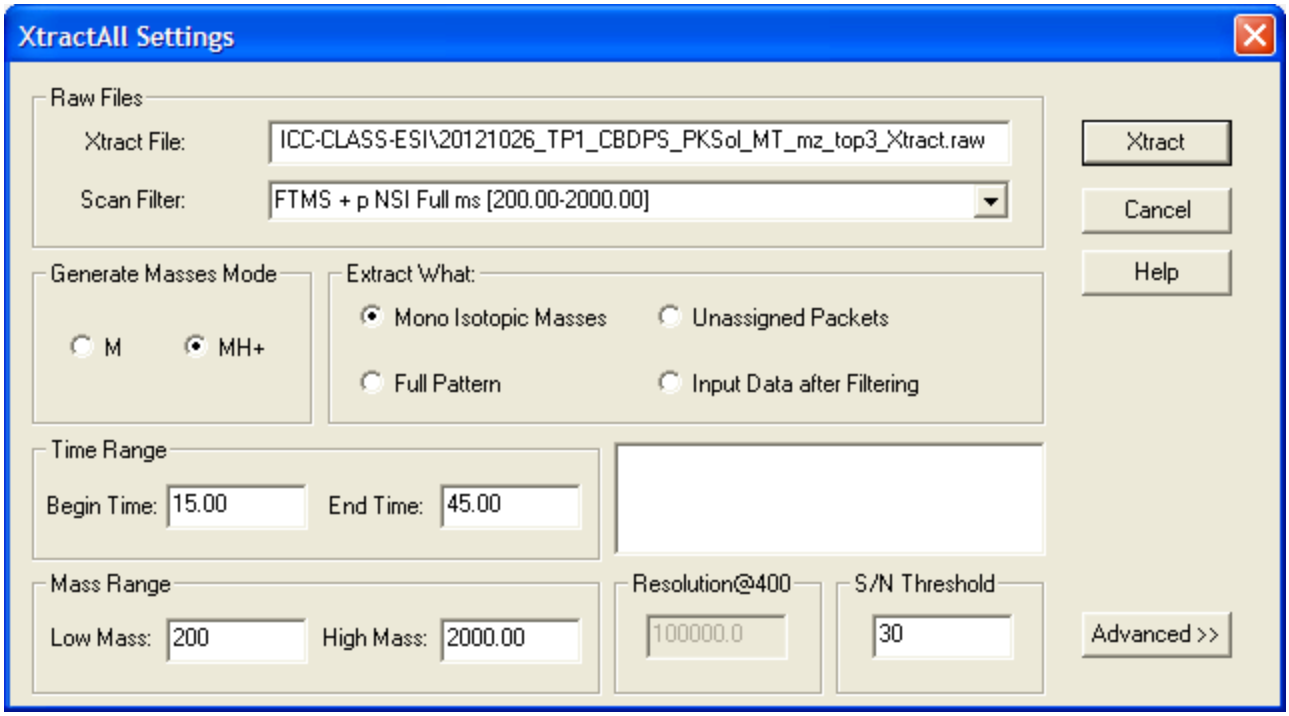
Below is the example how to process mass spectrometric crosslinking data with ICC-CLASS DXMSMS Match.

Example is Orbitrap LC-MS-MS/MS run of test peptide TP1 crosslinked with CBDPS-H8/D8, digested with proteinase K and affinity purified with immobilized avidin. MS data were acquired with Xcalibur (ver. 2.1.0.1140) with Mass Tags and Dynamic Exclusion precursor selection methods enabled in global data dependent settings. For CBDPS-H8/D8 mass difference between light and heavy isotopic forms of 8.05824 Da was used in Mass Tags setting. Mass Tag run used a Top 3 method. MS scans (m/z range from 200 to 2000) and MSMS scans were acquired in the Orbitrap mass analyzer at 60000 and 30000 resolution, respectively. Fragment ions for MSMS acquisition were produced by collision-induced dissociation (CID) at normalized collision energy of 35% for 10 ms and activation q = 0.25.

1. Detecting doublets of signals corresponding to light (H8) and heavy (D8) forms of crosslinks.
When Mass Tags method was used for acquisition, doublets of signals can be obtained in two ways: searching MS spectra (A) or using precursors lists (B). Here we will present both ways.

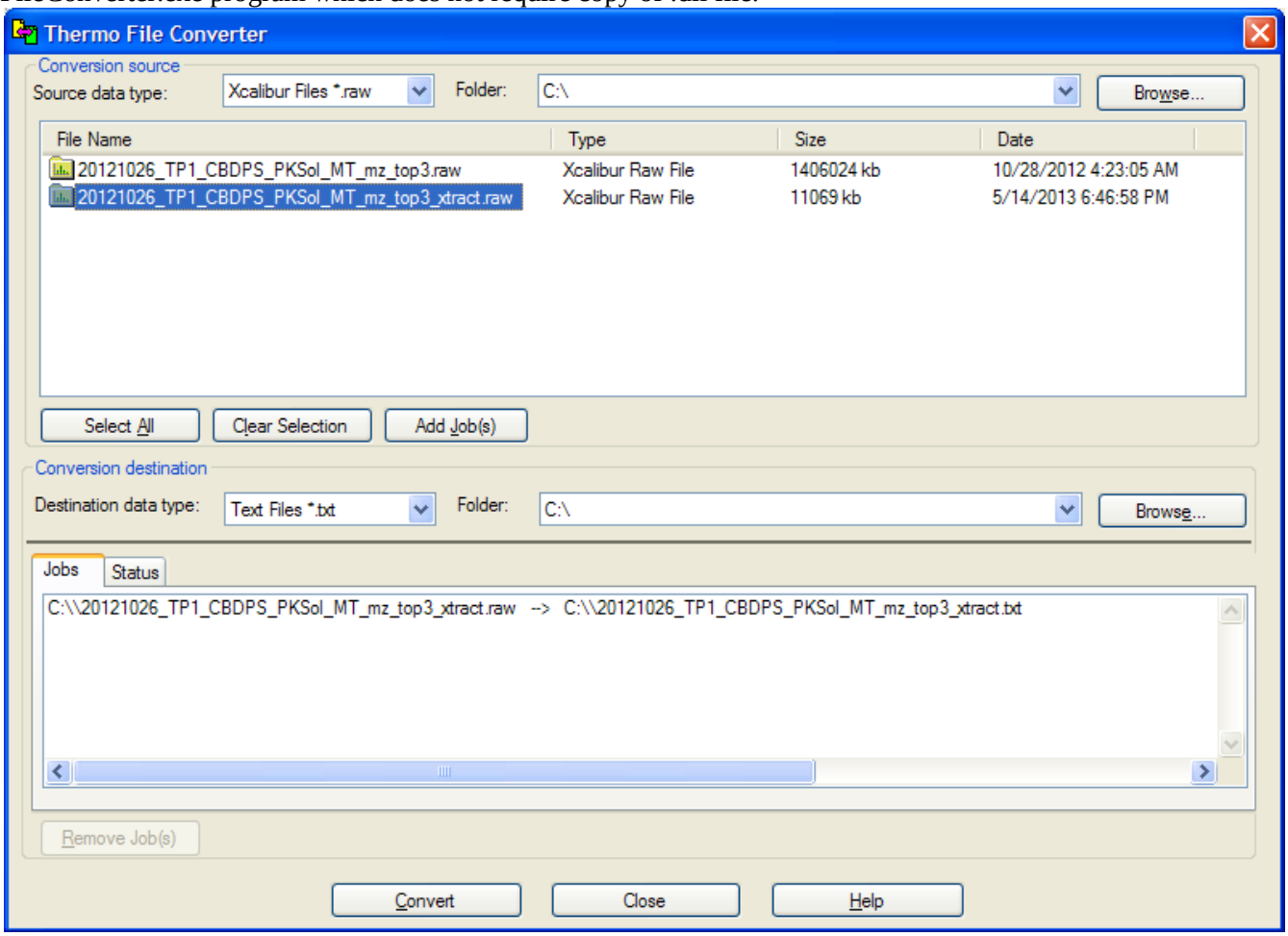
A. Doublets list from MS spectra.
A.1. Extract deconvoluted mass lists using Excalibur's Xtract RAW file utility.





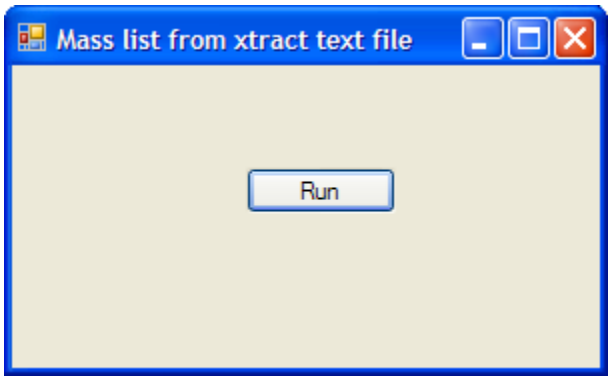
This will create new xtracted .raw file with _xtract addition to the file name.

A.2. Convert this file to the text file using XConvert program of XCalibur package (usually it is located in Xcalibur>system>programs> folder). You can copy it to the data folder together with Thermofisher.Foundation.FConvert.dll file and run it in there. In newer version of XCalibur package this is substituted by FileConverter.exe program which does not require copy of .dll file.



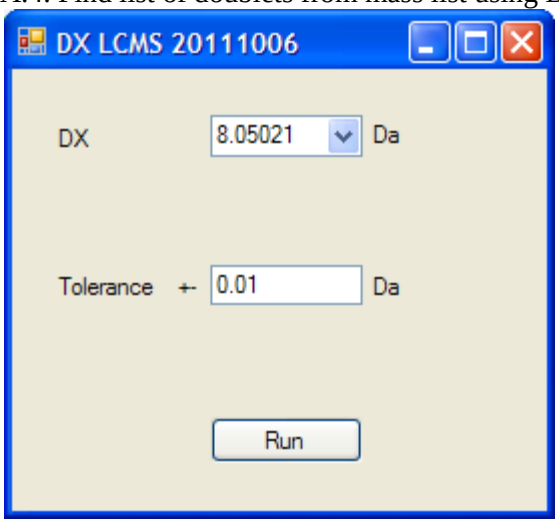
This will create text file containing xtracted MS data _xtract.txt.

A.3. Extract mass list of the run from _xtract.txt file using Mass List From Xtract Text File program.



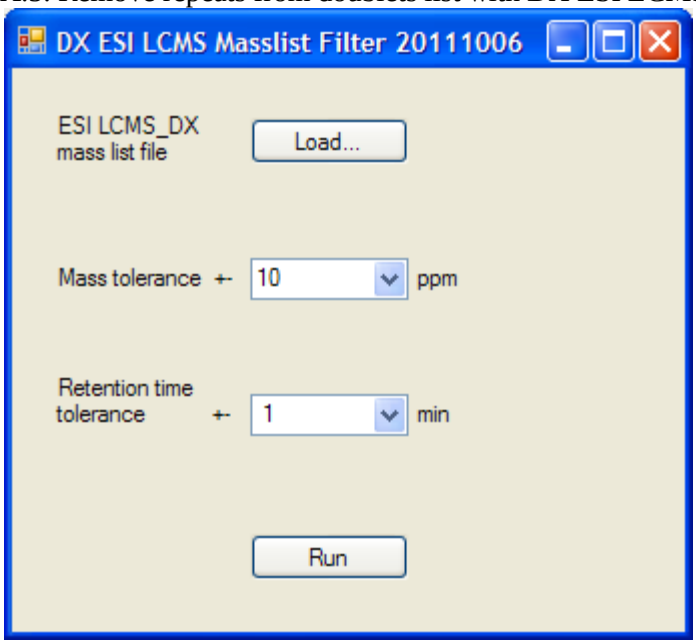
This will create _xtract_MassList.txt file.

A.4. Find list of doublets from mass list using DX ESI LCMS program.



This will create _xtract_MassList_DX.txt file.

A.5. Remove repeats from doublets list with DX ESI LCMS Mass List Filter program.



This will create _xtract_MassList_DX_Filtered.txt file.
This file is ready to be used as input file for DXMSMS Match program.
The file will contain sorted by mass list of light mass, intensity, retention time in minutes, heavy mass, intensity, retention time in minutes and residual difference between heavy and light masses of the doublets minus isotopic label mass difference:

Copy of 20121012_TP1_CBDPS_PKSol_InclList_M1_xtract_MassList_DX_Filtered - Notepad

File Edit Format View Help

402.148163	348798	26.537193	410.192108	799589	26.537193	-.0063
402.148193	663119	38.990860	410.192139	46572	38.990860	-.0063
402.148163	16455	45.375027	410.192169	24858	45.375027	-.0062
424.211914	14206	17.411693	432.256439	66010	17.411693	-.0057
451.113800	405039	33.537858	459.164154	114629	33.537858	.0001
452.188080	363882	26.304358	460.238342	362882	26.304358	.0001
479.177185	1263432	27.176530	487.229889	198017	27.176530	.0025
481.228882	123000	19.081027	489.285492	131703	19.081027	.0064
481.229034	4132908	26.265367	489.285065	120924	26.265367	.0058
481.229309	22578	31.323360	489.285889	38761	31.323360	.0064
495.067322	49767	45.694530	503.107574	65415	45.694530	-.0100
519.138794	166797	15.156860	527.198120	6449	15.156860	.0091
521.135864	37089	54.853192	529.189758	16116	54.853192	.0037
528.115540	519829	41.023860	536.166016	328378	41.023860	.0003
529.189819	4710	54.717192	537.239807	23692	54.717192	-.0002
531.191162	117557	21.501525	539.241455	18312	21.501525	.0001
533.159607	136683	31.647692	541.209778	136342	31.647692	.0000
534.259033	25088	21.252200	542.309143	19100	21.252200	-.0001
535.242798	395640	23.302358	543.292908	279777	23.302358	-.0001
536.165222	72122	17.687192	544.207092	20685	17.687192	-.0083

A.6. Create .mgf file of the run using Thermo's Proteome Discoverer program.
This file is ready to be used as input file for DXMSMS Match program and will look like this:

20121012_TP1_CBDPS_PKSol_InclList_M1 - Notepad

File Edit Format View Help

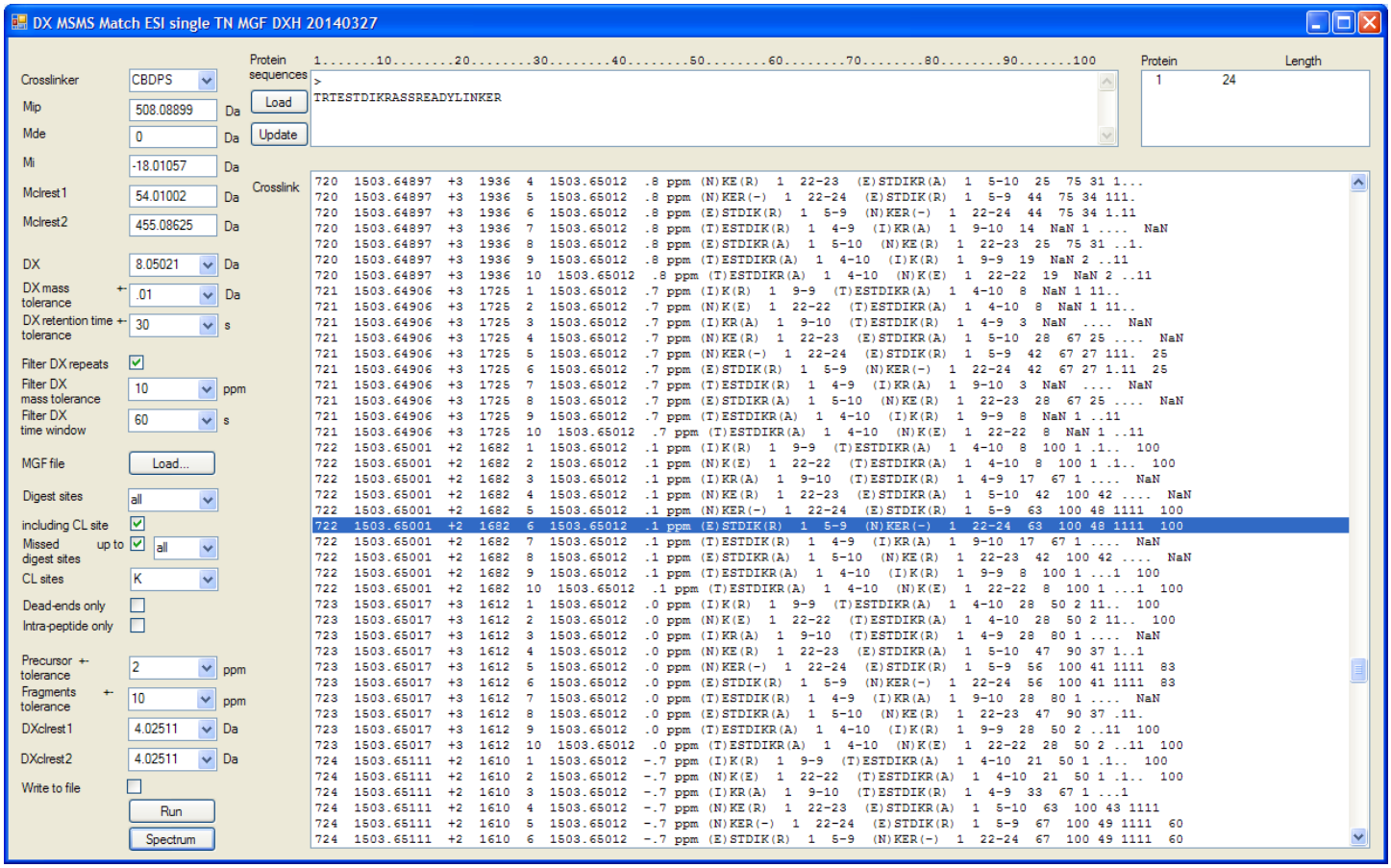
```
MASS=Monoisotopic
BEGIN IONS
TITLE=File4651 Spectrum1 scans: 890
PEPMASS=437.24222 15274.88379
CHARGE=2+
RTINSECONDS=841
SCANS=890
158.09280 32.3536
174.08868 310.584
175.11909 761.718
176.12210 39.779
176.12566 25.3981
191.11519 42.8924
205.14857 53.2761
210.95042 234.749
212.94749 206.815
236.11450 71.974
```

A.7. Perform search for crosslinks masses assignments and verifications using DXMSMS Match ESI DXH program.

B. Doublets list from .mgf file.
When Mass Tags method is used for the precursor selection, Xcalibur acquisition software detects doublets of signals corresponding to isotopic coding mass difference during the run and sequentially acquires MS/MS spectra of light and heavy isotopic forms. Thus, doublets masses can be deducted from the precursors masses stored in .mgf file of the run. For this option there is version of the program DX MSMS Match ESI MGF DXH, which uses only .mgf file and protein sequences as an input. This version can be used also for data acquired using TopN or TopSpeed methods, but if heavy precursors MS/MS spectra were not acquired, isotopic coding information will be lost for those crosslink assignments.

B.1. Create .mgf file of the run using Thermo's Proteome Discoverer program.
This file is ready to be used as input file for DXMSMS Match program.

B.2. Perform search for crosslinks masses assignments and verifications using DXMSMS Match ESI MGF DXH program.



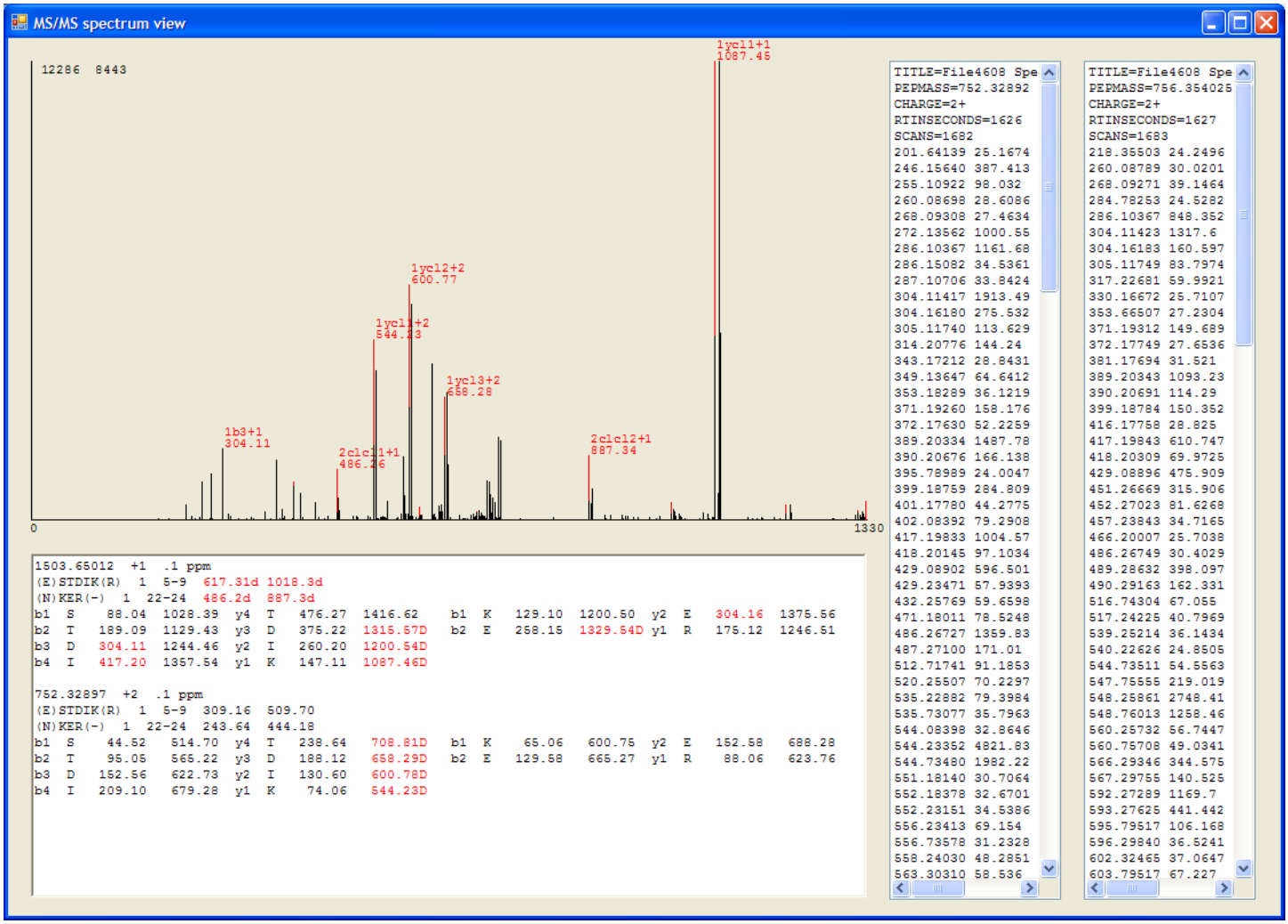
Make necessary selections of the crosslinker, digestion, crosslinking sites and crosslink type restrictions. If user-specific crosslinker is selected, fill in Mip mass value ($[M_{12}+H]^+ = [M_1+H]^+ + [M_2+H]^+ +$ Mip , where M_1, M_2 - masses of free peptides; M_{12} – mass of inter-peptide crosslink), Mclrest1 and Mclrest2 mass values ($[M_{12}+H]^+ = [M_1cl+H]^+ + [M_2cl+H]^+ +$Mcliploss, $[M_1cl+H]^+ = [M_1+H]^+ +$Mclrest, where M_1cl and M_2cl - cleaved peptide products from the cleavage of the inter-peptide crosslink, Mclrest, – mass of the cleaved portion of the crosslinking reagent, attached to the cleaved peptide product of the inter-peptide crosslink).

Load protein sequences using Load button. Sequences can be manually modified in the protein sequences window. Hit Update button to convert all the sequences to uniform text. Number of proteins and their lengths will be displayed in protein window.

Load MGF file. Change, if necessary, precursor mass and retention time tolerance values (precision for finding of MS/MS spectra in .mgf file) and fragments tolerance value (precision for crosslink MS/MS fragments matches to theoretical values).

Hit Run button. Message box with a number of the peptides considered for the search will be displayed. Hit OK button to run the search. Precursors mass list, doublets mass list and filtered doublets mass list (if filtering option is selected) files will be written to the same as program's directory. Found matches will be displayed in the Crosslink text box. If necessary, modify restrictions and run search again. To store output of the search in the file, check Write to file button.

Output will contain peptide and protein data for the found matches, scores and "." or "1" indicators of found crosslinks cleavage products. You can select the line of the match you want to inspect and hit Spectrum button. This will open Spectrum View window, in which you can see the selected MS/MS spectrum and inspect fragment masses matches made.



Found matches for light isotopic form will be highlighted in red. In the window on the right corresponding portion of the light and heavy .mgf files will be displayed. In the window on the bottom theoretical pairs (without and with counterpart crosslinked peptide, which correspondently should manifest in the spectrum as single and doublet signals) of cleavage products, b- and y- fragment masses of the crosslink will be listed. If crosslinker containing fragments were detected as doublets in a merged light and heavy spectrum, they are marked with “D” for peptide backbone fragments and “d” for cleavage products.

Spectrum view can be zoomed in by selecting rectangular range within the spectrum field with left mouse button. Zoomed in mass range can be shifted by selecting direction and length of the shift by left mouse button in the space below x-axis of the plot. Zoom out to the original mass range can be done by single left mouse button click within the spectrum field.

Following the spectrum inspection this window can be closed and next spectrum can be selected for the verification of the assignment from the Crosslink text box of the search DXMSMS Match window.

Output file can be opened in Excel, where it can be sorted, filtered and formatted according to the user preferences. Example of the processed output of the CBDPS-H8/D8 crosslinked test peptide TP1, digested with proteinase K, enriched with monomeric avidin beads and analyzed by LC-ESI-MS and MS/MS on Orbitrap instrument using Mass Tags acquisition method, is presented below. Unique inter-peptide Lys-Lys crosslinks were selected.

Deconvoluted MS Peak (MH+) (L)	MS Intensity (L)	MS Retention Time (min) (L)	MGF Precursor # (L)	Original Precursor (m/z) (L)	z (L)	MS Retention Time (sec) (L)	Scan # (L)	Precursor Prediction #	Theorhetical (MH+) (L)	ppm	Protein (1)	AA# Start (1)	AA# End (1)	Modified AA# (1)	AA -1 (1)	Peptide Sequence (1)	AA +1 (1)	Protein (2)	AA# Start (2)	AA# End (2)	Modified AA# (2)	AA -1 (2)	Peptide Sequence (2)	AA +1 (2)	CID Ions Score	Intensity Share Score	401-1	401-2	401-3	401-4	401 Total
1143.51055	467485	32.59	1151.559984	34087	32.59369	-0.0008		2	1143.5108	0.2	1	7	9	6	T	DIK	R	1	8	9	7	D	IK	R	233	33	1	1	1	1	4
1186.52715	129789	26.20	1194.577318	23574	26.197012	0		4	1186.5278	0.6	1	7	9	6	T	DIK	R	1	9	10	8	I	KR	A	200	35	1	1	1	1	4
1200.54468	6301076	22.26	1208.592847	4321319	22.256845	-0.002		1	1200.5435	-1	1	8	9	7	D	IK	R	1	22	24	21	N	KER	-	267	14	1	1	1	1	4
1257.55364	575784	32.81	1265.603563	51460	32.811512	-0.0003		1	1257.5537	0.1	1	7	9	6	T	DIK	R	1	20	22	19	L	INK	E	163	24	1	1	1	1	4
1258.53826	2195936	35.08	1266.586962	222441	35.076325	-0.0015		1	1258.5377	-0.4	1	7	9	6	T	DIK	R	1	7	9	6	T	DIK	R	100	79	1	1	1	1	4
1314.59763	42559	16.96	1322.64777	13391	16.961503	-0.0001		1	1314.5976	0	1	9	11	8	I	KRA	S	1	22	24	21	N	KER	-	213	32	1	1	1	1	4
1315.56975	193790	26.73	1323.619799	92058	26.728512	-0.0002		2	1315.5704	0.5	1	7	9	6	T	DIK	R	1	22	24	21	N	KER	-	200	35	1	1	1	1	4
1331.59062	50507	31.88	1339.640428	27098	31.881012	-0.0004		2	1331.5905	-0.1	1	5	9	4	E	STDIK	R	1	8	9	7	D	IK	R	170	37	1	1	1	1	4
1332.54973	62909	29.12	1340.600756	6576	29.120012	0.0008		2	1332.5493	-0.3	1	5	9	4	E	STDIK	R	1	21	22	20	I	NK	E	160	40	1	1	1	1	4
1347.54961	360805	28.69	1355.599413	139277	28.69368	-0.0004		4	1347.549	-0.4	1	5	9	4	E	STDIK	R	1	22	23	21	N	KE	R	160	29	1	1	1	1	4
1359.58562	138718	34.53	1367.635301	23170	34.53301	-0.0005		2	1359.5854	-0.2	1	6	9	5	S	TDIK	R	1	7	9	6	T	DIK	R	190	25	1	1	1	1	4
1372.60301	4559167	16.13	1380.650247	2456172	16.132512	-0.003		7	1372.6031	0.1	1	22	24	21	N	KER	-	1	22	24	21	N	KER	-	225	53	1	1	1	1	4
1374.60808	89969	25.61	1382.658129	24960	25.614012	-0.0002		4	1374.6075	-0.4	1	5	9	4	E	STDIK	R	1	9	10	8	I	KR	A	160	37	1	1	1	1	4
1416.61861	2683222	24.76	1424.668456	476935	24.764423	-0.0004		7	1416.6181	-0.4	1	6	9	5	S	TDIK	R	1	22	24	21	N	KER	-	180	35	1	1	1	1	4
1446.61699	218814	34.25	1454.667162	96619	34.253188	0		2	1446.6174	0.3	1	5	9	4	E	STDIK	R	1	7	9	6	T	DIK	R	183	26	1	1	1	1	4
1471.67156	337687	21.51	1479.720859	71367	21.51368	-0.0009		2	1471.6715	0	1	7	10	6	T	DIKR	A	1	22	24	21	N	KER	-	260	26	1	1	1	1	4
1486.64575	91402	16.68	1494.695499	37152	16.677505	-0.0005		2	1486.646	0.2	1	21	24	20	I	NKER	-	1	22	24	21	N	KER	-	260	44	1	1	1	1	4
1503.64972	638881	25.83	1511.699317	182154	25.834845	-0.0006		6	1503.6501	0.3	1	5	9	4	E	STDIK	R	1	22	24	21	N	KER	-	192	26	1	0	1	1	3
1632.69389	3335085	25.61	1640.742561	705904	25.614012	-0.0015		7	1632.6927	-0.7	1	4	9	3	T	ESTDIK	R	1	22	24	21	N	KER	-	157	25	1	1	1	1	4
1659.75107	217941	21.51	1667.800388	47987	21.51368	-0.0009		3	1659.7512	0.1	1	5	10	4	E	STDIKR	A	1	22	24	21	N	KER	-	221	28	1	1	1	1	4
1730.78794	60303	22.70	1738.83801	20399	22.700515	-0.0001		7	1730.7883	0.2	1	5	11	4	E	STDIKRA	S	1	22	24	21	N	KER	-	188	28	1	1	1	1	4