



X! Search Engine Development

X! HUNTER ASL MGF file format (2007.06.01)

[Home](#) | [Statistics](#) | [File formats](#) | [Server](#) | [Source code](#) | [Libraries](#)

The X! Hunter Annotated Spectrum Library (ASL) system normally uses a binary file format to record the spectra and annotation. This format is a beta-test format that will allow the use of text formatted files, using a simple extension of the Mascot Generic File format.

The file has a small header section with three entries:

1. SEARCH=MIS [required]
2. REPTYPE=Peptide [required]
3. LIBSIZE=nnn [required, where nnn = number of spectra in the file]

The annotation and spectra are stored sequentially. Each spectrum in the library begins with BEGIN IONS and ends with END IONS. A typical entry is as follows:

BEGIN IONS	[start of a spectrum]
PEPMASS=353.681	[parent ion m/z value]
CHARGE=2	[parent ion charge]
PEPSEQ=CASLQK	[peptide sequence]
PEPEXP=1.22866e-007	[annotation confidence]
PEPMOD=57.0215@1	[peptide modification and position]
PEPACC=sp ALBU_BOVIN @223	[protein accession number and position]
PEPACC=sp ALBU_HUMAN @224	[protein accession number and position]
GMPp=GMPp0504028113	[peptide accession number]
187.15 3	[ion m/z and intensity]
204.18 22	
215.04 3	
232.07 27	
238.28 3	
257.3 3	
258.26 4	
275.31 9	
301.31 3	
319.13 7	
327.71 5	
387.16 3	
388.36 7	
414.18 10	
475.33 100	
542.34 5	
543.3 8	
546.36 8	
560.25 12	
689.39 4	
END IONS	[end of a spectrum]

NOTES:

1. The spectra are not stored in any particular order: spectra associated with the same protein may be located anywhere within the file.
2. Annotations are based on sequence accession numbers for particular sequence collections, e.g., ENSEMBL, IPI or SWISS-PROT protein accession numbers.
3. X! Hunter ASLs store the twenty (20) most intense peaks for a particular MS/MS spectrum.
4. Parent ion masses are calculated based on the mono-isotopic masses of the peptide residues.