RAW to MGF Converter version 20090904a TITLE line items

Preliminaries

There are 2 distinct ways to determine chromatographic peak maximum and Full Width at Half Maximum (FWHM); they are depicted in Fig. 1 and Fig. 2 below.

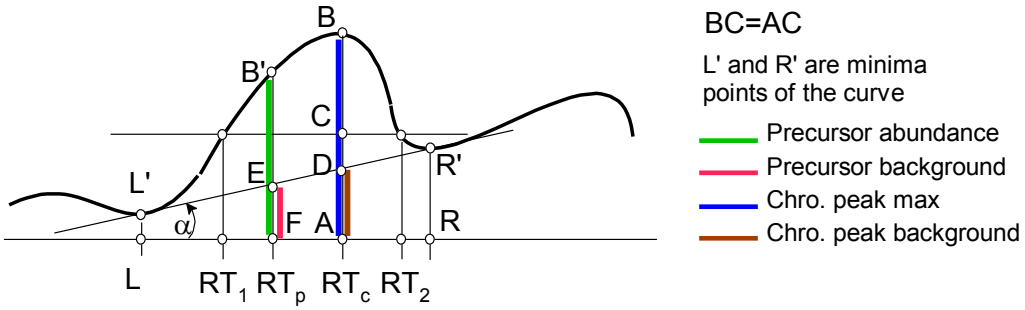

Fig. 1

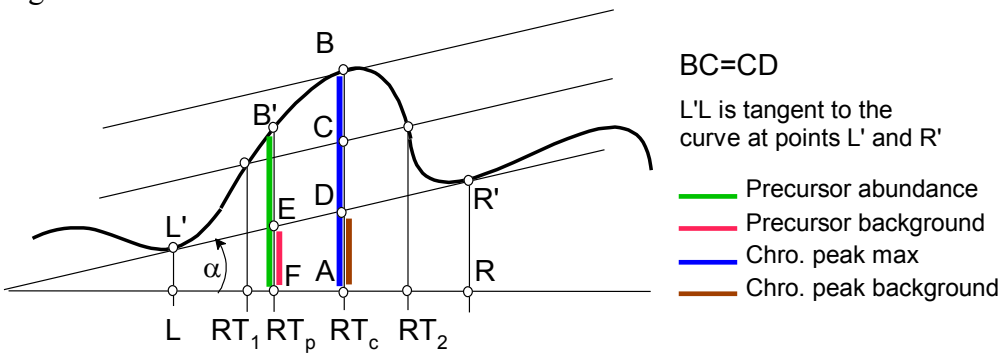

Fig. 2

AB	= max. abundance
RT_2-RT_1	= Full Width at Half Maximum, FWHM
RL	= Width of chromatographic peak foot
AD	= background of the chromatographic peak
B'F	= precursor abundance,
EF/BF	= precursor to background ratio
EF	= precursor background
AD/AB	= chromatographic peak background to the peak max. ratio
(R'R – L'L)/AB	= background line skewness

In the table below, the following are determined from smoothed precursor XIC curve:

Precursor1MaxAb	= **(Chro. Peak max),** in absolute units
Precursor1MaxWidth	= RL (minutes)
Precursor1MaxWid50	= $RT_2 – RT_1$ (minutes)
Precursor1MaxRatio	= **(Precursor abundance)** /**(Chro. Peak max)**
Precursor1MaxBkg	= **(Chro. peak background)** / **(Chro. Peak max)**
Precursor1AbuBkg	= **(Precursor background)** / **(Precursor abundance)**
Precursor1MaxHW	= AC/AB
Precursor1MaxSkew	= (R'R – L'L)/AB

The tokens for method used in Fig2 differ by a single digit, 2 instead of 1, for example Precursor2MaxAb instead of Precursor1MaxAb, etc.

Table of TITLE Items Legend

Footnotes:
[o] – Only for Orbitrap files, only upon "–MonoisoMgfOrbi" option and if available
[p] – search in profile spectrum if available
[2] – only for ms2
[m] – only if chromatographic peak has been found

Abbreviations:
PMR = precursor mass range = ±0.2 (LTQ) or ±0.05 (Orbitrap) centered on Isolation m/z obtained from the ms2 spectrum (if available) or precursor m/z from the Filter
IPMR = isolation precursor mass range, usually ±1.0 around the precursor m/z
ChroMax = precursor ion current chromatographic peak maximum
ChroPeak = precursor ion current chromatographic peak

Table of TITLE Items

	Token	**Value**	**Meaning**
1	Scan:	%ld	Scan number = 1,2,…
2	RT:	%0.3f	Scan retention time, minutes
3	PrecursorScan:	%ld	Precursor scan number as reported by Xcalibur
4[o]	PrecursorMonoisoMZ	%0.4f	Precursor monoisotopic m/z
5[o]	PEPMASS:	Monoiso	PEPMASS in scan below = monoisotopic m/z
6[o]	PEPMASS:Monoiso PrecursorMZ:	%0.4f	Precursor m/z [only if PEPMASS:Monoiso]
7	PrecursorCharge:	%ld	Precursor charge [if available]
8	IBP:	%0.2f	Base peak intensity of the scan
9	ITot:	%0.2f	Total ion current of the spectrum
9a	max2med:	%0.2f	IBP/median of the intensities
10	InjTime:	%0.2f	Scan injection time / ms
11	IsolationMZ::	%0.4f	Isolation m/z (ms2) or m/z from the Filter
12[p]	PrecursorAb:	%0.2f	Max precursor intensity[(p)] within PMR
13	MPY:	%0.2f	1.00 always, for future extensions
14	ms1PrecursorTotAb:	%0.2f	TIC of precursor spectrum abundances
15	ms1PrecursorInjTime:	%0.2f	Precursor scan injection time / ms
16[p2]	ms1PrecursorMZ:	%0.4f	m/z of max precursor intensity[(p)] within IPMR
17[p2]	ms1PrecursorMzAvg:	%0.4f	Avg. m/z weighted with abund. within IPMR
18[p2]	ms1PrecursorMzRMS:	%0.4f	RMS of ms1PrecursorMzAvg
19[p2]	ms1PrecursorIntens:	%0.2f	Max precursor intensity[(p)] within IPMR
	(was ms1PrecursorAb:)		
20	ms1PrecursorRT:	%0.3f	Precursor scan retention time (min)
21	ms2IsolationWidth:	%0.2f	Isolation width of ms2 spectrum, usually 2.0
22	ms1SelMz:	%0.3f-%0.3f	Selected m/z range for accumulated ms1 intensity used for searching the ChroMax
23	ms1SelAvgMZ:	%0.4f	Average m/z within FWHM of the ChroPeak
24	ms1SelRmsMZ:	%0.4f	RMS of ms1SelAvgMZ

	Token	Value	Meaning
25[2]	PrecursorHasMax:	%d	1=> ChroMax found, 2=> only in skewed search (never happened), 0=>not found
25[2]	Precursor1HasMax:	%d	ChroMax found in Fig.1 setting
25a[2]	ms1PrecursorAb:	%0.2f	*Actual Ms1 intensity* interpolated to ms2 RT; same units as ChroMax, e.g. Precursor1MaxAb
25b[2]	ms1PrecursorMax:	%0.2f	Max *actual Ms1 intensity* under the ChroMax peak or zero if ChroMax not found; same units as ChroMax, e.g. Precursor1MaxAb
26[2]	Precursor1HasMax:	%d	ChroMax found in Fig.1 setting
27[2]	Precursor1MaxInjTime:	%0.2f	Precursor scan Inj.Time at Precursor1MaxRT
28[2]	Precursor1MaxTotAb:	%0.2f	TIC of precursor scan at Precursor1MaxRT
29[2]	Precursor1MaxAb:	%0.2f	Height of accumulated within ms1SelMz range and smoothed over time ChroPeak
30[2]	Precursor1MaxRT:	%0.3f	RT of ChroMax / minutes
31[2]	Precursor1MaxWidth:	%0.2f	Width of ChroPeak foot, minutes or 0
32[2]	Precursor1MaxWid50:	%0.2f	FWHM of ChroPeak, minutes or 0
33[2]	Precursor1MaxRatio:	%0.3f	Smoothed precursor abund. / ChroPeak max
34[2]	Precursor1MaxBkg:	%0.2f	ChroPeak background / ChroPeak max
35[2]	Precursor1AbuBkg:	%0.2f	Precursor background / ChroPeak max
36[2]	Precursor1MaxHW:	%0.2f	Intensity at 50% / ChroPeak max
37[2]	Precursor1MaxSkew:	%0.2f	ChroPeak background line skewness
38-49	*Similar to 26-37*		*With 1 replaced by 2 in the token*
50[2]	PrecursorMaxNoise	%0.2f	$\sum\|ab_i-ab_{i-1}\| / \max(ab_i) - 2$ over the ChroMax foot; = 0 if the peak has single local max.
51[2]	PrecursorRTStep:	%0.3f	PrecursorMaxWidth/(num ms1 spectra in it-1)
52	ConvVer:	%s	20080212a – this new version
53	Filter:	%s	Xcalibur Filter string

Note due to a bug, ms1PrecursorMzWidth in previous version was 1.0 instead of 2.0

	PrecursorHasMax: Flags	Flag Meaning
	FoundShiftedPrecursor	ChroMax was discarded (PrecursorHasMax:0) because the max was found beyond the central m/z band, PMR
	NarrowPeakRejected	ChroMax was discarded (PrecursorHasMax:0) as possible noise smoothing artifact
	DblWdth	ChroMax could be found only for smoothing Gaussian is widened by a factor of 2
	NoDblWdth	If the smoothing Gaussian is widened by a factor of 2 the ChroMax cannot be detected; usually occurs in presence of another chromatographic peak nearby
	DblWdth2Norm	ChroMax location for smoothing Gaussian is widened by a factor of 2 was used to find FWHM from the single width Gaussian
	PeakSearchFailed	Although ChoMax was found after 5-point smoothing, it could not be found after Gaussian smoothing
	PeakPresearchFailed	Although ChoMax was not found after 5-point smoothing, it was found after Gaussian smoothing

	PrecursorHasMax: Flags	**Flag Meaning**
	Noisy	Very noisy precursor ion current chromatogram, looks like an uncorrelated noise

Flags example:
PrecursorHasMax:0, FoundShiftedPrecursor, NoDblWdth

	Precursor1HasMax: Precursor1HasMax: Flags	**Flag Meaning**
	Wid50Prb:1	Problem: Could not correctly interpolate RT of the left 50% point of ChroPeak max
	Wid50Prb:2	Problem: Could not correctly interpolate RT of the right 50% point of ChroPeak max
	Wid50Prb:3	Problem: Could not correctly interpolate RT of both 50% points of ChroPeak max

Flags example:
Precursor1HasMax:1,Wid50Prb:1

Newly added MGF TITLE output values

numOCMF:109,109,0,2
-- always present

PrecursorWayMMF:5.54 PrecursorMaxMMF:5.54 mzRmsMax:0.38 mzRmsMs2:0.16 maxDelMz:0.3340,3-0,94,98 ms2DelMz:0.3339,3-0,100,99
-- only for Orbitrap AND when PrecursorMaxMZ:>0 is present

PrecursorMaxMZ:681.2294 PrecursorMaxAb:1194.02 PrecursorMaxRT:16.622
It is output only if PrecursorHasMax: ≠ 0. The values are found from the actual (unsmoothed) intensities; in case of a profile this refers to the area under a single ms peak.

ms1PrecursorMax:13142.38
Max original (unsmoothed) Ms1 intensity found inside the found selected precursor ion chromatographic peak or 0.00 if the peak has not been found; the units are the same as in ms1PrecursorAb (interpolated unsmoothed precursor intensity at the time of ms2 sampling) and Precursor1MaxAb (max. smoothed intensity). The values are obtained by summation of intensities inside the selected m/z range within the isolation m/z range. The may include more than one isotopic peak. The same m/z range is used to obtain the smoothed XIC profile.

Explanation of the output-

numOCMF:109,109,0,2

109=number of ms1 scans found in 3.25 minute RT interval
109=number of points after condensing the ms1 scans
0=number of points in the moving average; 0=>none
2=number of times NNLSQ was called

The following is for Orbitrap, and only if PrecursorMaxMZ>0

PrecursorWayMMF:5.54

5.54=max ms1 mismatch factor for precursor ms1 among ms1 scans from precursor scan to max (above 100 means poor match; 0 may mean precursor ms1 is at the precursor XIC profile max)

PrecursorMaxMMF:5.54

5.54= precursor ms1 mismatch factor for ms1 at the found precursor XIC peak profile max (above 100 means poor match)

mzRmsMax:0.38

0.38=100*weighted average of RMS(mz) over all m/z bands in the XIC peak max
(above 0.3-0.7 usually means impurity)

mzRmsMs2:0.16

0.16=100*weighted average of RMS(mz) over all m/z bands for the precursor ms1
(above 0.3-0.7 usually means an impurity)

maxDelMz:0.3340,3-0,94,98

0.3340=average isotopic delta(m/z) for at least 3 ms1 peaks at the max of XIC peak
3=number of ms1 peaks used to calculate delta(m/z)
0=number of impossible isotopic peaks (below monoisotopic m/z) among the used
94=percentage of ms1 intensity in the isolation width covered by the used ms1 peaks, including impossible isotopic peaks
98=dot product (based on sqrt) of the found isotopic peaks and the expected isotopic pattern
Notes:

- in matching the isotopic pattern, the max. isotopic peak intensity is assumed to be inside the precursor m/z isolation range
- the isolation width is always assumed to be 2;
- all zeroes mean an isotopic sequence of at least 3 peaks has not been found

ms2DelMz:0.3339,3-0,100,99

the same as above; refers to the ms1 scan that immediately precedes the ms2 scan.

PrecursorMaxMZ is the m/z of the ms1 spectrum at the point of maximum intensity inside the range which is output as, for example, ms1SelMZ:680.7091-681.4291. Typically, the ms1 spectrum is a profile. To get intensity in comparable units, the converter does a poor man centroiding: it adds the intensities on the right and on the left of the found maximum until the intensity reaches a minimum or the delta m/z from the max. reaches 0.1 for Orbitrap or 0.4 for others. Currently I am not sure how it would treat LTQ FTMS.

PrecursorMaxAb is the result of the intensity summation.

PrecursorMaxRT is the RT of the ms1 where the m/z and intensity were read. It may differ from the RT of the XIC peak max because no background subtraction was done.

Other changes to the output

Previously, when the found max of the profile XIC peak had little or no contribution from the central band (it was always 0.4 m/z units wide), the output was:

PrecursorHasMax:0,FoundShiftedPrecursor

In the new version, the output is

PrecursorHasMax:1,FoundShiftedPrecursor

Explanation: in the new version, the central band in case of Orbitrap is 0.15 m/z units. This makes the “FoundShiftedPrecursor” flag appear quite often and it does not have as much meaning as before.

Changes to ReAdw4Mascot2 converter starting from version 20090902a – not used in real conversion yet.

3 new items have been added to the TITLE line to characterize Orbitrap precursor ms1 spectrum peaks in the ms2 isolation range. The purpose is to provide information on the purity of the peaks in the isolation range.

Example

PEPMASS:Replaced:1

FilterMzPeakExists(1845):1

PCFD2,2;1,-1,567.3154,1,3,1.0033,0.0003,80,0,3,1.15,0.21,0.03,0.01,0.6,2.0,5.8,994;0,0,568.3494,1,2,1.0038,0,9,0,5,1.18,0.21,0.05,0.02,1.4,2.8,5.0,972

Description of PEPMASS:Replaced:n

Present only if PEPMASS value has been replaced. n=1 means ms1 preceding the ms2 was used; n=2 means ms1 following ms2 was used

Description of FilterMzPeakExists(n):m

m=1 means there is a peak at m/z = Isolation m/z (m/z value in the filter string)
m=0 means the peak doesn't exist
n is ms1 scan number used to calculate PCFD (see below)

Description of PCFD

This item lists best matches of the peaks located in and around the isolation range to the isotopic patterns of a compound that has elemental composition close to "Averagine", $C_{4938}H_{7758}N_{1358}O_{1477}S_{42}$ (Senko et al, J Am Soc MS 1995 pp. 229-233 suggested C 4.9384 H 7.7583 N 1.3577 O 1.4773 S 0.0417) for the tested m/z value and charge.

Three matches of the isotopic pattern were tested for each series of found equidistant peaks from the ms1 spectrum: the one that has the same position of the maximum intensity as the found peaks, and 2 more, each shifted by 1 peak to higher and to lower m/z. The selection of the best match was done based of the highest value of the dot product = $(\sum\sqrt{I_{spec}\times I_{isotopic}})\Big/\sqrt{\sum I_{spec}\times\sum I_{isotopic}}$. The dot product included peaks located beyond the isolation range so that the isotopic pattern is not truncated.

Format of PCFD

PCFD<N>,<M><;match $_1$>…<;match $_N$>
where M ≥ N is the number of found isotopic patterns covering at least 10% of the largest found coverage;
no spaces inside; each match record starts with semicolon; the last match record is followed by a space.

Format of match $_k$

match $_k$ consists of 18 comma-separated numbers in the following order:

(1) [int] 1 or 0: 1 means the filter m/z peak belongs to the found sequence of equidistant peaks
(2) [int] number of peaks missing from the ms1 spectrum at the low mass end of the isotopic pattern; negative value means there have been found extra equidistant peaks.
(3) [%.4f] precursor m/z found for the match
(4) [int][?] charge found for the match; may be followed by a question mark, ?, which means the isotopic envelope search was aborted due to exceeding max. number of comparisons.
(5) [int] number of matched peaks, including peaks beyond the isolation range
(6) [%.4f] average distance between consequent matched peaks (approx 0.5, 0.3333, 0.25, etc.)
(7) [%.4f] RMS of the average distance or 0 if only 2 peaks have been matched
(8) [int] percent of the intensity matched in the isolation range by the found isotopic pattern peaks.
(9) [int] number of missing expected isotopic peaks. Only peaks that are higher than the smallest ms1 peak in the vicinity of the isolation range or above 20% of the largest isotopic peak are counted
(10) [int] percent of missing or wrong intensity of the isotopic envelope is calculated from the normalized matching ms1 peaks and isotopic pattern as $100\times(\sum|I_{spec}-I_{isotopic}|)/\sum I_{isotopic}$ and may exceed 100. Only peaks that are higher than the smallest ms1 peak in the vicinity of the isolation range are counted.
(11) [%.2f] Match quality measure = $\dfrac{\sum \min\left(I_{isotopic,i}, I_{isotopic,i-1}\right)\times\left(\dfrac{I_{spec,i}}{I_{spec,i-1}}+\dfrac{I_{spec,i-1}}{I_{spec,i}}\right)}{\sum \min\left(I_{isotopic,i}, I_{isotopic,i-1}\right)\times\left(\dfrac{I_{isotopic,i}}{I_{isotopic,i-1}}+\dfrac{I_{isotopic,i-1}}{I_{isotopic,i}}\right)}$. Only matching isotopic patterns that produce this measure below 3.2 are output.
(12) [%.2f] Match quality measure is a weighted sum of $\left|\dfrac{I_{spec,i}}{I_{spec,i-1}}-\dfrac{I_{isotopic,i}}{I_{isotopic,i-1}}\right|+\left|\dfrac{I_{spec,i-1}}{I_{spec,i}}-\dfrac{I_{isotopic,i-1}}{I_{isotopic,i}}\right|$ similar to (11)
(13) [int] Percent of missing or wrong intensity of the matched peaks is calculated from the normalized matching ms1 peaks and isotopic pattern as $(\sum|I_{spec}-I_{isotopic}|)/\sum I_{spec}$ and may exceed 100. Only peaks that are higher than the smallest ms1 peak in the vicinity of the isolation range are counted.
(14) [int] The same as above except the summation is over peaks in the isolation range
(15) [%.1f] log10(Max ms1 intensity / max. matched isotopic peak intensity)
(16) [%.1f] log10(Total ms1 intensity / max. matched isotopic peak intensity)
(17) [%.1f] log10(max. matched isotopic peak intensity)
(18) [int] 1000×dot product described above. The dot product includes peaks located beyond the isolation range.

The matches are sorted in order of descending (8), percent of the intensity matched in the isolation range by the found isotopic pattern peaks. Not more than 2 matches are output.