Mascot Distiller

What is Mascot Distiller?

A uniform interface to all the popular MS data file formats
- Interactively, as a data browser
- For applications that need to access "raw" files

A tool for creating high quality peak lists

An interface for submitting Mascot searches and reviewing the results

A tool for calling sequence tags and performing *de novo* sequencing

Support for Quantitation methods that require more than a standard MS/MS peak list.

MASCOT   **Topic:** *Mascot Distiller*   *© 2009-2010 Matrix Science*   **MATRIX SCIENCE**


Most laboratories will have instruments from more than one manufacturer. The instrument data systems are necessarily complex, so there can be a steep learning curve for someone who comes into the lab and just wants to browse their data or generate peak lists. The first benefit of Mascot Distiller is that you can access all of the popular data formats from a single user interface.

Another reason for developing Distiller was to produce high quality peak lists without having to constantly tweak peak detection parameters. Poor quality peak lists translate into poor quality Mascot scores.

Distiller is also a powerful way to review Mascot search results. And, if Mascot fails to get a match, you can perform de novo sequencing and interpret sequence tags for tag searches

Finally, Distiller is used for quantitation methods that require information from the raw data file, either because it is necessary to integrate the elution profile of each precursor peptide or because information is required for precursor peptides that were not used to trigger MS/MS scans, so are missing from the peak list. There is a separate presentation dealing with quantitation.
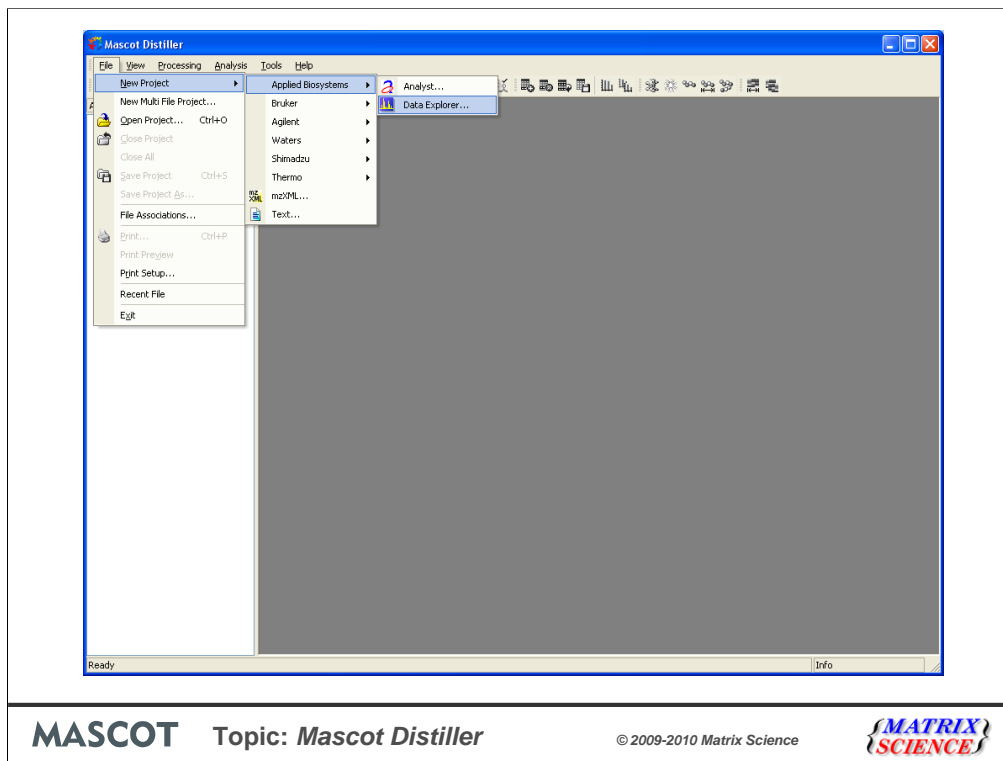
What is Mascot Distiller?
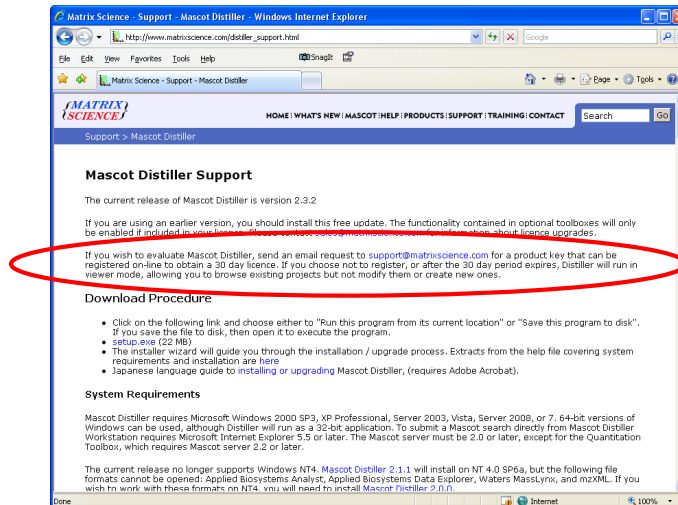

MASCOT Topic: *Mascot Distiller*   *© 2009-2010 Matrix Science*

If you are not already a Distiller user, you can get a free 30 day evaluation licence. Details are on our web site, under Support; Mascot Distiller

Topic: *Mascot Distiller* © 2009-2010 Matrix Science

A binary or "raw" file is initially opened as a new Distiller project. We'll run through the supported formats in a minute

Topic: *Mascot Distiller* © 2009-2010 Matrix Science
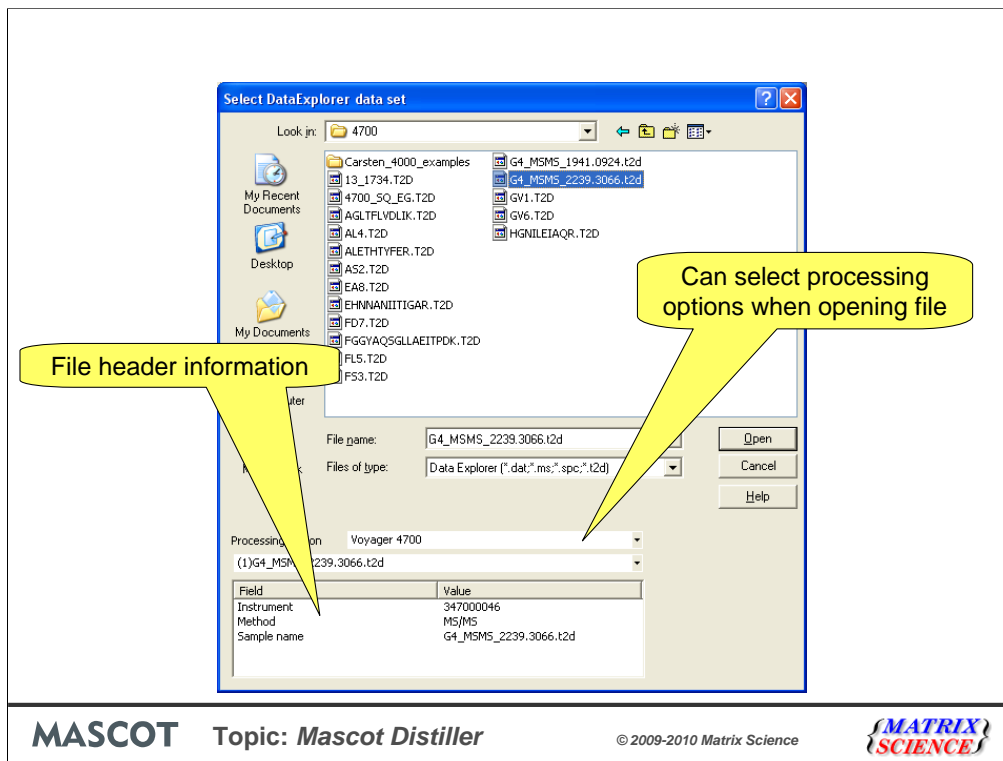
The file browse box displays some header information for the selected file. Also, this is where you can choose the processing options. You can change these later, but choosing the correct set here saves a couple of mouse clicks.

Comprehensive support

Agilent
- DataAnalysis (yep) (LC/MS Trap)
- Analyst (wiff) (LC/MS TOF)*
- MassHunter (Q-TOF)

Applied Biosystems
- Analyst (QStar, QTrap)
- Data Explorer (Voyager*, 4700, 4800)

Bruker
- Data/Flex Analysis (yep) format (Esquire)
- Data/Flex Analysis (baf) format (Apex, MicrOTOF)
- XMASS/XTOF (Reflex, Biflex, etc.)

Shimadzu
- Kompact (Axima)
- LCMSSolution (LCMS-IT-TOF)*

Thermo
- Xcalibur (LCQ, LTQ, Orbitrap)

Waters
- MassLynx (QTof, M@ldi, TofSpec)

mzXML 2.0 and 2.1 (XML interchange format)
Text (ASCII mass and intensity values)

Types of data:
- Single MS spectrum
- Single MS/MS spectrum
- Multiple MS spectra
- Multiple MS/MS spectra (e.g. nano spray)
- 'Triple play' (survey, enhanced/zoom, MS/MS)
- LC-MS/MS (complex arrangements of survey and MS/MS scans)
- Lockspray (Masslynx)

* indicates that original data system library files are required

MASCOT Topic: *Mascot Distiller*   *© 2009-2010 Matrix Science*   *MATRIX SCIENCE*

Mascot Distiller supports all of the mainstream data file formats. In a few cases, Distiller requires library files that are installed as part of the instrument operating system.

Data files can be as simple as a single MS or MS/MS scan, or they can be the most complex mixtures of dependent scans created by IDE-type experiments.

There are some types of data that Mascot Distiller cannot handle. For example, it wouldn't know what to do with SRM (selected reaction monitoring) data.
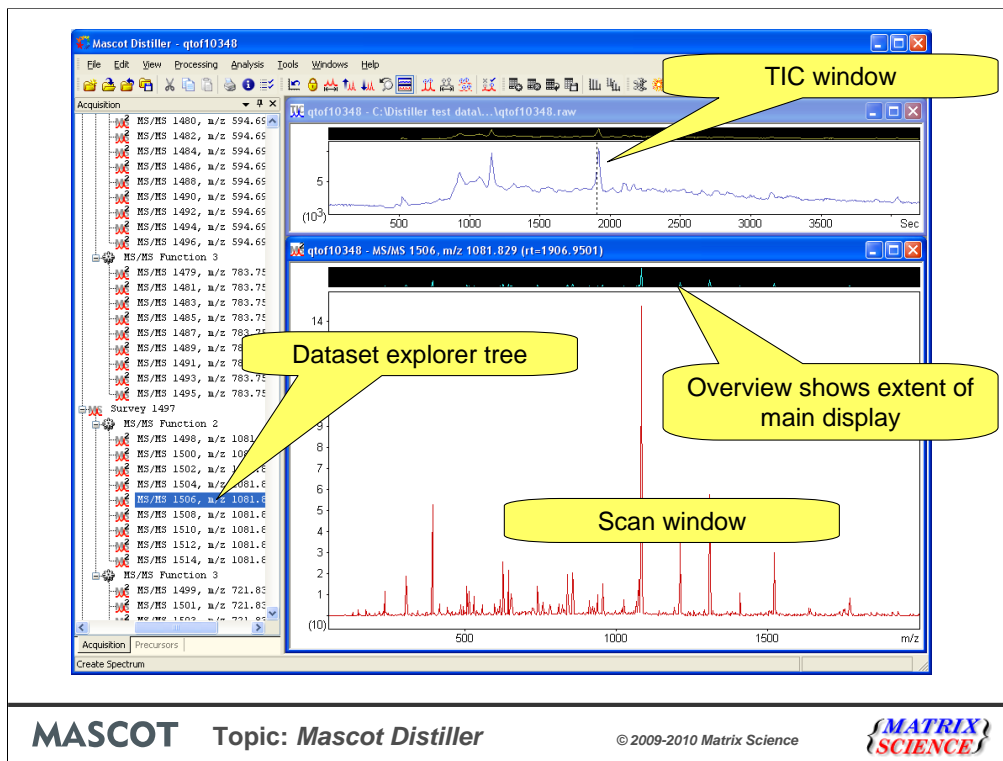
If the raw file contains LC-MS/MS data, this is the general appearance of the Distiller screen.

The Acquisition tab on the explorer tree shows the scan structure. This is Masslynx, so each survey scan is followed by MS/MS scans grouped into functions. If we were looking at (say) Xcalibur triple play data or Analyst data, this structure would look very different.

The total ion chromatogram window is chiefly a navigational aid.

The scan window displays a mass spectrum trace, selected by clicking on the explorer tree or the TIC trace.

At the top of the TIC and scan windows, you can see a representation of the whole trace, called the overview. The black area shows the portion of the trace that is currently displayed. This can be dragged or resized to make zooming and panning around the trace very easy.
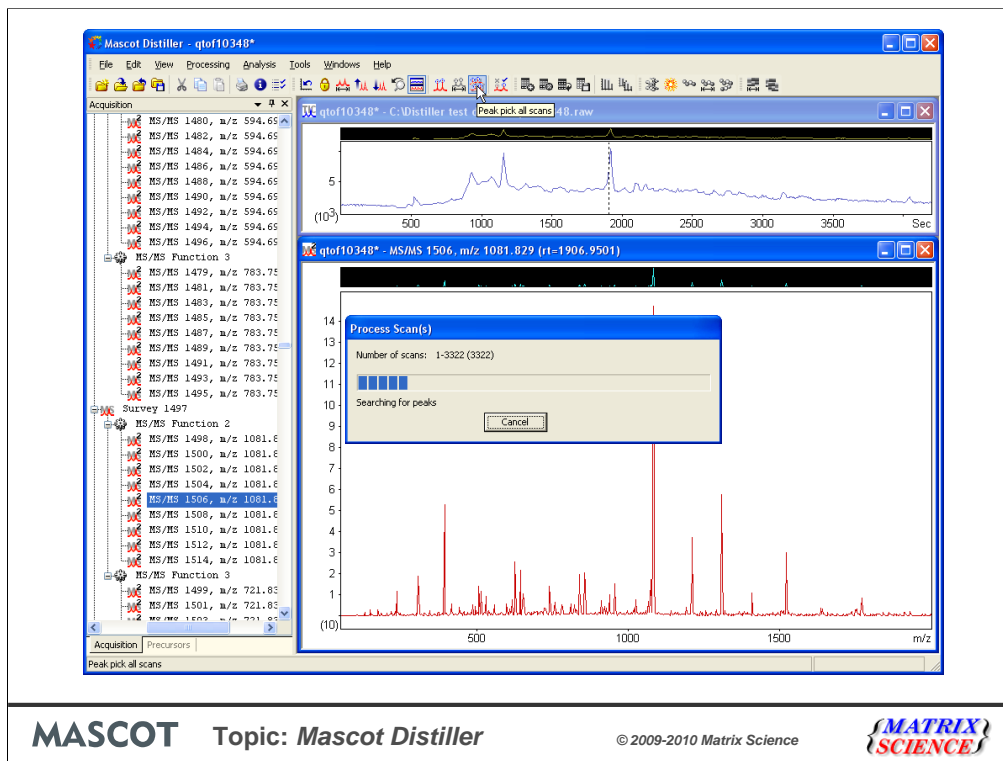
To process the raw data into peak lists suitable for database searching, we just choose process from the Processing menu or toolbar. We can process the current scan, the currently displayed scan range, or all scans.
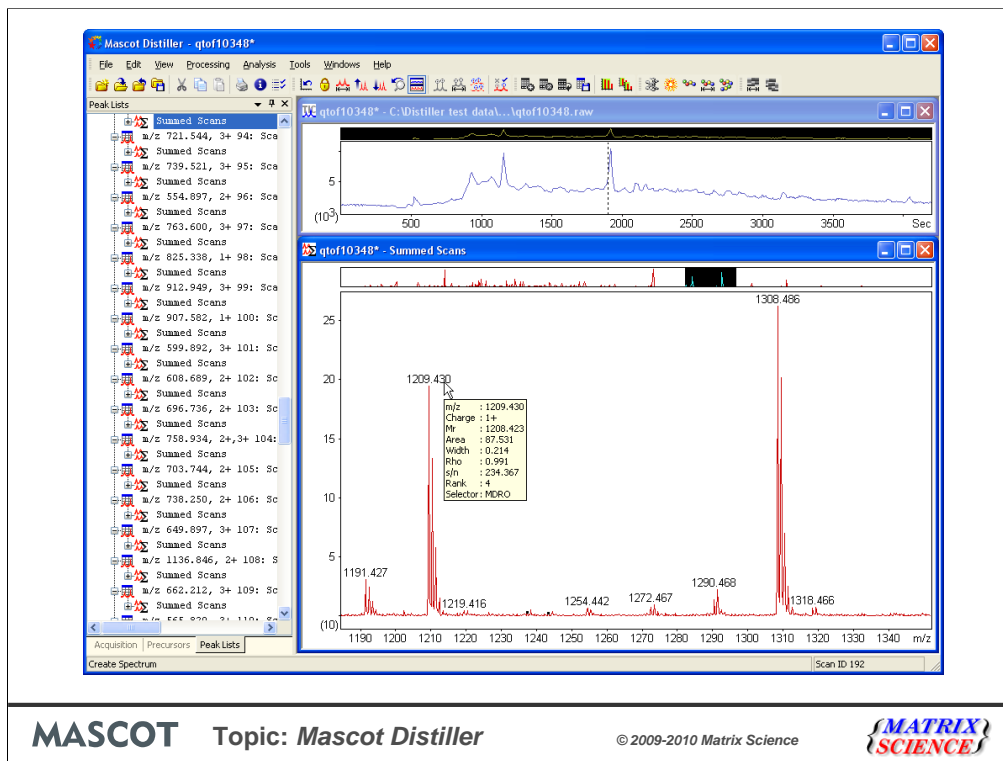
After processing, there is a new tab on the explorer tree: Peak lists. This shows the new data structure, which will usually be different from the original acquisition structure, because scans from the same precursor have been summed together.

When the mouse cursor is over a peak label, we get a tooltip showing complete information about the peak. m/z, charge, Mr, and area are fairly obvious. Width is the full width at half maximum height in Daltons.

Rho is the correlation coefficient, which measures the quality of the peak. Anything over 0.7 is normally a real peak and not a noise spike.

The peak picking procedure also gives us a signal to noise ratio for each peak. Rank is just the order in which the peak was picked and selector shows that this peak was picked by the Distiller library, which we call MDRO (from Mascot Data Reduction Object), as opposed to a manually edited peak

9

Topic: *Mascot Distiller* © 2009-2010 Matrix Science
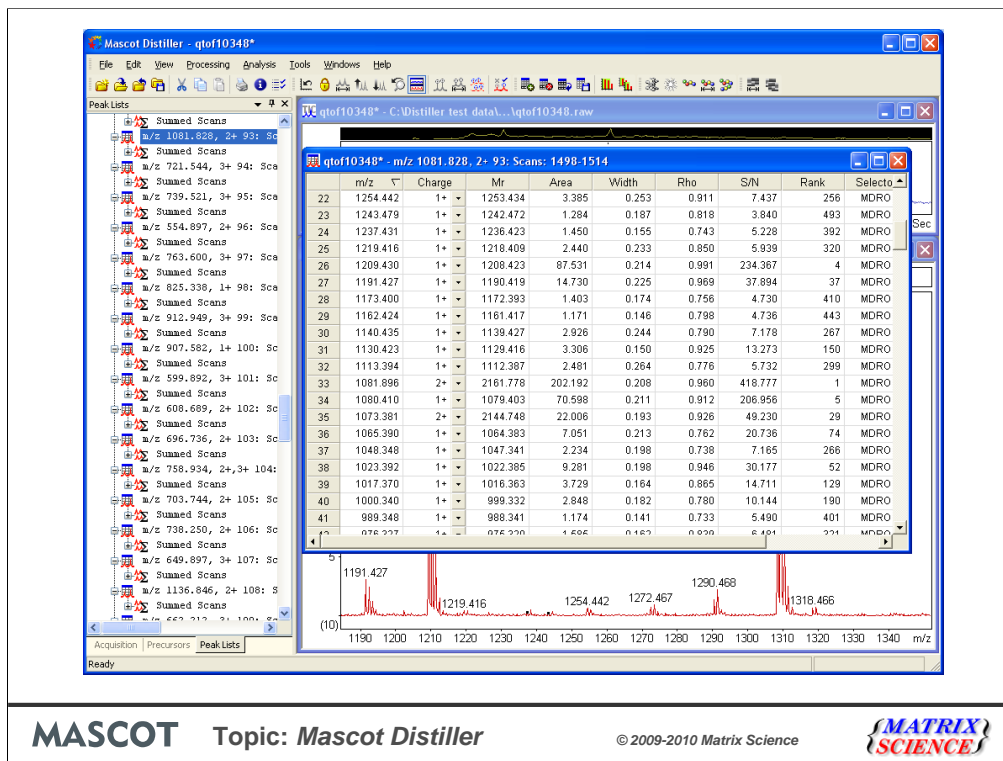
Click on a peak list node on the explorer tree and you get a peak list window containing a grid of these values for all the peaks. This is where you could edit or delete a peak, if you really wanted to. Clicking on the column headers sorts the table on that column
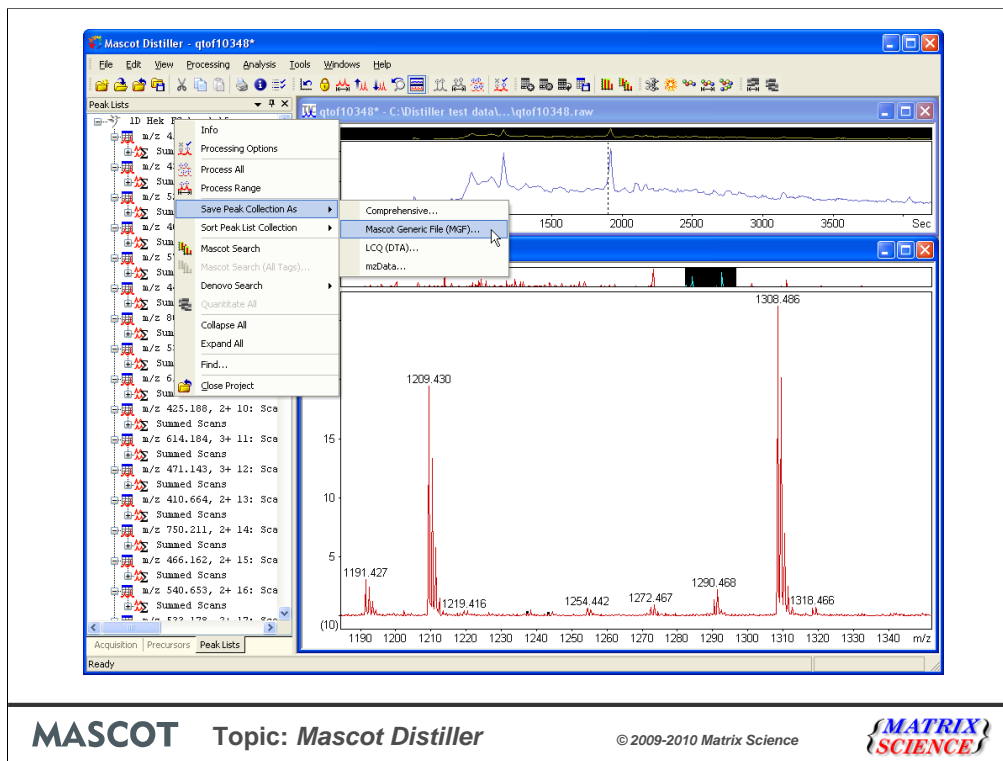
Topic: *Mascot Distiller*   © 2009-2010 Matrix Science

Having created a peak list, you might just want to save it to a file. The supported formats are Mascot Generic fomat, Comprehensive, which includes all of the peak information that we were just looking at, mzData, the XML format from the Proteomics Standards Initiative, and DTA format, which is used by Sequest.

Problems with conventional peak detection

- Failure to pick low intensity peaks
- Picking peaks that are just noise
- Selection of the wrong peak(s) in an isotopic cluster
- Need to continually 'tweak' parameters.

So, what is different about peak picking in Mascot Distiller?

Conventional peak detection works by smoothing the spectrum then looking for a rising gradient, which is the onset of a peak, and a falling gradient at the tail of a peak. Trouble is, this only works well if various parameters are set just right for the particular spectrum. If these parameters are not right, then we see either failure to pick low intensity peaks or picking of peaks that are just baseline noise.

Another common problem is selecting the 13C peak instead of the smaller 12C peak, so that the mass is out by a full Dalton

The need to continually 'tweak' parameters is a big headache if you want to process files automatically, without looking at each spectrum

Peak Picking in Mascot Distiller

1. Choose the most intense feature in the spectrum
2. For each charge state to be considered, calculate the shape of the isotope distribution for an average peptide at that m/z value
3. For each calculated distribution, iteratively adjust the position and peak width to obtain the best fit as measured by the correlation coefficient
4. Select the distribution that gives the best fit, and subtract the fitted area from the spectrum
5. Return to step 1 and repeat until nothing is left but noise

Similar methods have been described by

Peter Berndt et. al., Electrophoresis (1999) 20 3521-3526
Robin Gras et. Al., Electrophoresis (1999) 20 2535-3550.

MASCOT   **Topic:** *Mascot Distiller*   *© 2009-2010 Matrix Science*   *MATRIX SCIENCE*

Mascot Distiller detects peaks by attempting to fit an ideal isotopic distribution to the experimental data. These are the steps in this process.

Advantages of Mascot Distiller peak detection

- The peak list contains just ^{12}C, monoisotopic mass values
- Less likely to get ^{13}C peak by mistake
- Automatically get charge state, total area, and quality statistics for every peak
- Smoothing / filtering not required
- No need to tweak parameters constantly.

And, these are the advantages

Limitations of Mascot Distiller peak detection

- Significantly slower than conventional peak detection
- Can get confused by certain type of isotopic labelling, e.g. 50% ^{18}O
- Accuracy not so good if isotope envelope is distorted, (e.g. badly tuned FTMS)
- Limited benefit if "raw" data already centroided, (e.g. most Xcalibur files).

MASCOT   **Topic:** *Mascot Distiller*   *© 2009-2010 Matrix Science*   *MATRIX SCIENCE*

Of course, nothing is perfect. Here are some of the weaknesses.

Processing options: MS
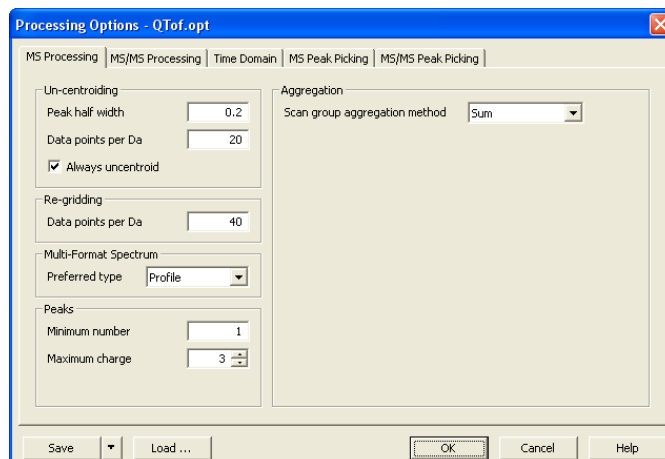
Topic: *Mascot Distiller*   © 2009-2010 Matrix Science

Although you don't have to tweak the peak picking parameters, this doesn't mean there aren't any. The point is that you can set the parameters once, for a particular instrument, and then leave them alone, confident that the peak detection will be consistently acceptable.

Sets of peak picking parameters are stored in XML text files, and can be viewed and modified using the processing options dialog. This has five tabs, and we don't have time today to go into great detail. Just press F1 when this dialog is displayed in Distiller and the on-line help will open up with complete information.

The first tab deals with MS (or survey) scans.

Mascot Distiller works best on profile data, and there is an unavoidable loss of information when data have been badly centroided. However, converting centroided data back to profile data and re-processing may yield some improvement. The Un-centroiding frame specifies how centroided data should be converted back into Gaussian peak profiles.

Some profile data are compressed by dropping runs of zero intensity data points. Before Distiller can process these scans, it is necessary to replace the missing points. This is called re-gridding, and is also used when spectra with non-linear mass scales have to be transformed onto a common m/z axis prior to summing.

For both un-centroiding and re-gridding, it is very important to choose a reasonable value for Data points per Da. There need to be sufficient points to faithfully represent the peak shape. A common rule of thumb is that you need 8 points to accurately define a peak shape. However, the higher this value, the longer peak detection will take. So, don't choose an unnecessarily high value just for the sake of it.

Aggregation is the gouping together equivalent spectra so that they can be summed into a single spectrum. For MS data, the choices are None and Sum. When None is chosen, a separate peak list will be created for each MS scan. When Sum is chosen, all the scans will be summed prior to peak detection.

Maximum charge state should be chosen carefully. Processing time increases in proportion to this number. More importantly, you should only look for high charge states in data with adequate mass resolution. If the instrument resolution means that all charge states above (say) 3 are unresolved, then it is impossible to determine from the gross width of a peak whether a peak is charge state 6 or 7. Baseline resolution is not needed, just enough genuine 'ripples' to indicate that the instrumental broadening is small compared with the gross width of the isotope cluster.

Any peak list that contains less than the Minimum number of peaks will be discarded. If the file contains LC-MS/MS, it is important to set this value to 1, because a survey scan with only one decent peak is still important. Conversely, when processing an MS data file for peptide mass fingerprinting, this value should be set higher, because it is unlikely that you will want to search a spectrum that has less than (say) 10 peaks.

Processing options: MS/MS
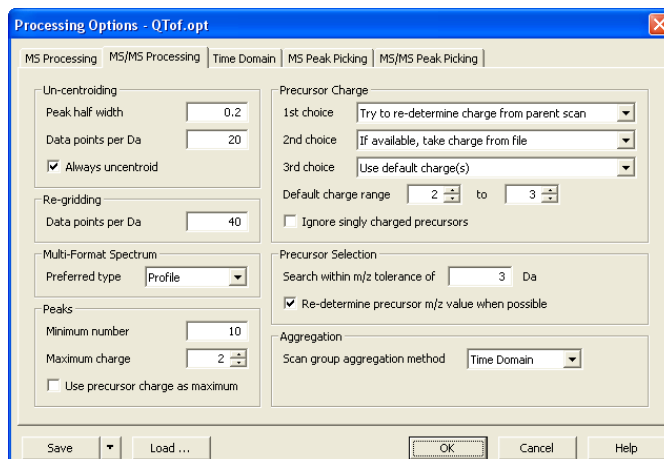
Topic: *Mascot Distiller*

© 2009-2010 Matrix Science

MS/MS Processing parameters apply to all MS/MS scans, whether a single scan, a series of scans, or MS/MS scans within an LC-MS/MS dataset.

Some controls are identical to those on the MS Processing tab. These controls have been duplicated on both tabs because instruments may have different resolution capabilities for MS and MS/MS.

The aggregation method choices for MS/MS are None and Time Domain.

None is only useful for files containing a series of MS/MS scans; it is not applicable to structured LC-MS/MS data. When None is chosen, a separate peak list will be created for each MS/MS scan.

Time domain is invariably the correct choice for LC-MS/MS data. Time domain means that precursor mass and charge information can be derived from survey scans, and that MS/MS scans from a common precursor should be summed together according to the rules on the Time domain tab.

The Precursor charge frame describes the most common decision paths for assigning precursor charge state. Some data systems, such as Xcalibur, seem conservative in determining charge state from a zoom scan. That is, charge state is rarely wrong, but it may be reported as unknown, even though the zoom scan is of good quality. In this situation, these settings would be appropriate.

If there were no zoom scans, then you wouldn't want to try and get the charge state from the low resolution survey scan. You would go straight to using defaults.

If 'Ignore singly charged precursors' is checked, spectra from singly charged precursors will be discarded. This is useful for electrospray analysis of tryptic peptides, where singly charged precursors are often noise or non-peptide contaminants.

The choices for precursor m/z can be simpler, because it is not possible to have a default m/z. The precursor m/z tolerance setting determines the maximum difference allowed between the precursor m/z value in the file and that re-determined by Distiller. The most common problem with precursor m/z is that the instrument data system has taken the 13C peak, so the precursor mass can easily be out by 1 Da

Charge defaults must be specified for those cases where it is not possible to determine the precursor charge state

Maximum Charge state offers a choice between 1+, 2+ and Same as precursor charge. For the purposes of creating a peak list for a Mascot search, the correct setting depends on whether you are outputting MS/MS peaks as m/z or MH+ values, (Peak List Format tab in the Preferences dialog). A conventional peak list contains m/z values, and the maximum charge state that Mascot looks for is 2+. Hence, there is no point in spending time looking for higher charge states. However, if your data definitely includes fragment ions with higher charge states, you should choose to output fragment ions as MH+ values, and select Same as precursor charge.

Processing options: Time domain
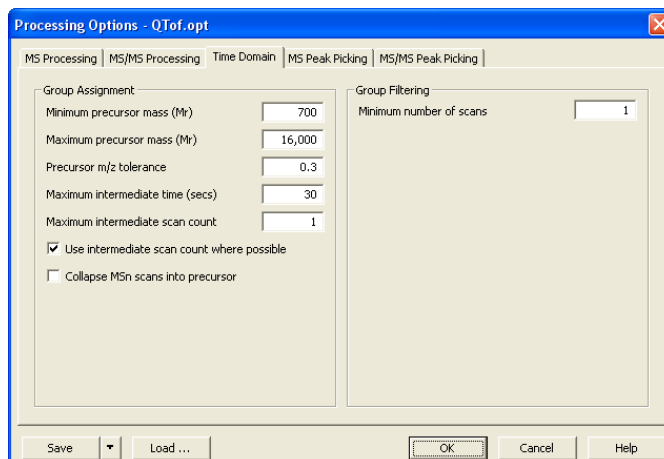
Topic: *Mascot Distiller*   © 2009-2010 Matrix Science

For LC-MS/MS data, the usual setting for Multi-scan data will be to use Time domain processing. The parameters on the Time domain tab then control how MS/MS scans will be summed and filtered. These settings are all fairly conventional.

Spectra from small peptides (< 6 residues) have no value in database searching, because such short sequences can be expected to occur by chance in a large database. A setting of 500 to 700 Da for Minimum precursor mass will generally be appropriate.

The upper limit on a precursor mass in Mascot is 16000 Da, so there is little point in adding larger peptides to the peak list.

Both Precursor m/z tolerance for grouping and Max number of intermediate scans are critical parameters for accurate time domain processing. Precursor m/z tolerance for grouping requires a good estimate of how the mass spectrometer mass precision might drift during a run. Don't set this parameter too wide, because you may end up averaging together spectra from different precursors.

Maximum intermediate time or Maximum intermediate scan count require an estimate of the quality of the chromatography. Might you expect to see the same peptide elute over a period of 10 seconds, or 10 minutes, or what?

Note that scans in this context refers to survey scans. In a MassLynx file, function 1 might be the survey scan while functions 2, 3 and 4 might be MS/MS scans. Only the function 1 scans would count towards the total number of intermediate scans allowed to divide a potential group of MS/MS scans. Some knowledge of how the acquisition method was configured will be helpful in choosing an appropriate value.

If Use intermediate scan count where possible is cleared, the Maximum intermediate scan count is disabled. If there is more than one raw file in the Distiller project, then Maximum intermediate scan count is ignored even if the checkbox is checked, because scan count has no meaning from one file to another

Collapse MSn scans into precursor causes the peaks in (say) an MS3 scan to be added into the peak list of the parent MS2 scan. This seems to be the best way to make use of MSn data in a database search.

Setting minimum number of scans in group requires some knowledge of how the acquisition method was configured. If the method was attempting to acquire 8 scans from each precursor, then having a single scan in the file for a particular precursor could be a good indication that the scan was triggered on a noise spike. If this is not the case, then setting this value to more than 1 could be dangerous, because valid spectra might be discarded.

Processing options: Peak picking
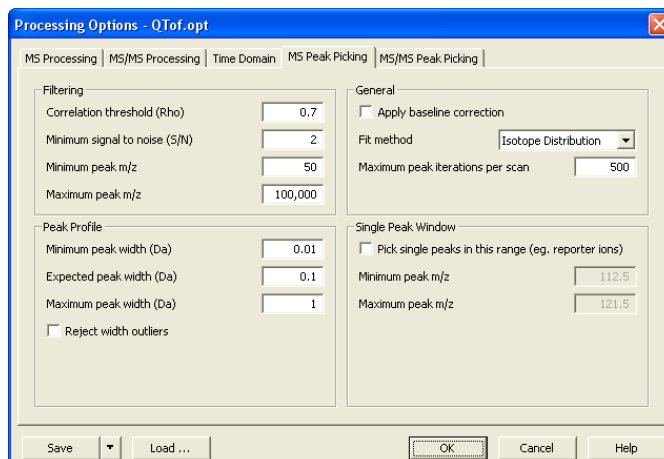
Topic: *Mascot Distiller* © 2009-2010 Matrix Science

There are separate tabs for MS and MS/MS peak picking parameters, with identical controls. This is useful for a hybrid instrument, where the characteristics of the two scan types may be very different. If the same settings can be used for both, there is a checkbox on the MS/MS peak picking tab for 'Same as MS Peak Picking'.

The peak picking code performs a least squares fit between a calculated isotope cluster and each candidate peak, returning the correlation coefficient, which is a measure of the similarity between the peak shapes. Ranges from 1 for perfect correlation to 0 for no correlation. Good fits to strong peaks will normally give correlation coefficients of 0.95 or better. Weak peaks will generally give lower correlation coefficients, and a cut-off of 0.7 seems to work well for all types of data.

S/N stands for signal to noise ratio. Minimum S/N needs to be set empirically for each type of instrument, though not for each dataset. Something between 1 and 10 will usually be appropriate. This is the parameter to tune using a typical spectrum from each instrument type.

Minimum and maximum peak m/z are system limits that should be set well outside the range expected for useful data.

The peak width settings are coarse ones, and do not need tuning. Minimum Peak Width and Maximum Peak Width are safe, conservative limits reflecting the physics of the instrument. Expected Peak Width is a starting point for iteration. The peak width will start here and can potentially go to either limit in the search for the best fit to a peak.

When Reject width outliers is checked, a robust non-parametric routine is applied to detect and remove peak width outliers. It can improve the quality of the peak list in some circumstances.

Baseline correction should be checked if the spectra have a significant baseline 'lift-off'. Generally, this will be true for MALDI-MS, but not for PSD, LCQ, QTOF, etc. If peak detection is performed on a trace with an elevated baseline, the peak list may be full of weak broad peaks, representing the signal 'under the baseline'.

Tip: The 'safe' setting is to have Baseline correction checked. It wastes a little processing time, but avoids the possibility of generating a bad peak list from a trace with an elevated baseline

Sometimes, the peaks in a spectrum cannot be modelled using averagine. In such cases, Distiller can pick single peaks, rather than try to fit complete isotopic distributions by choosing 'Single peak' as the Fit method.

Maximum iterations places an upper limit on the number of detectable peaks. This limit will rarely be reached in practice. The peak picking code iterates until this limit is reached or until no significant peaks remain in the spectrum. Hence, 500 is a safe number. Setting a smaller number may reduce the overall processing time.

Processing options: Peak picking
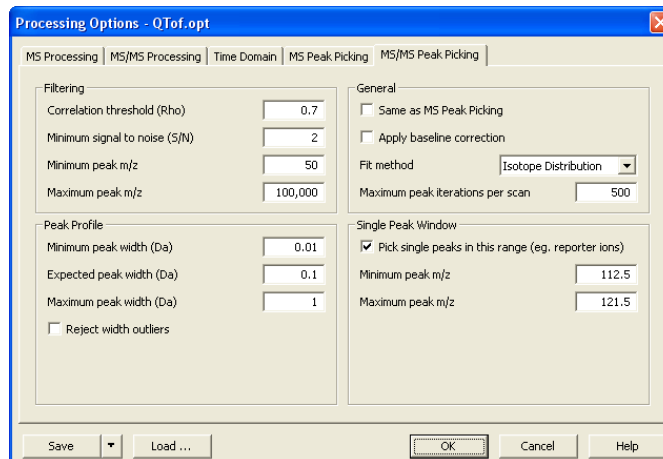
Topic: *Mascot Distiller* © 2009-2010 Matrix Science

For an iTRAQ or TMT experiment, the ideal is to have single peak picking in the reporter ion region and isotope distribution fitting elsewhere. You can achieve this using the 'Single peak window' frame

Since we don't want this behaviour for the MS scans, the checkbox for 'Same as MS Peak Picking' has been cleared
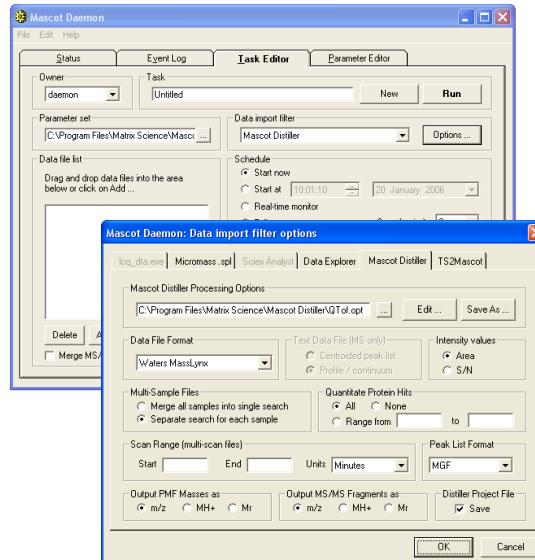
Preferences
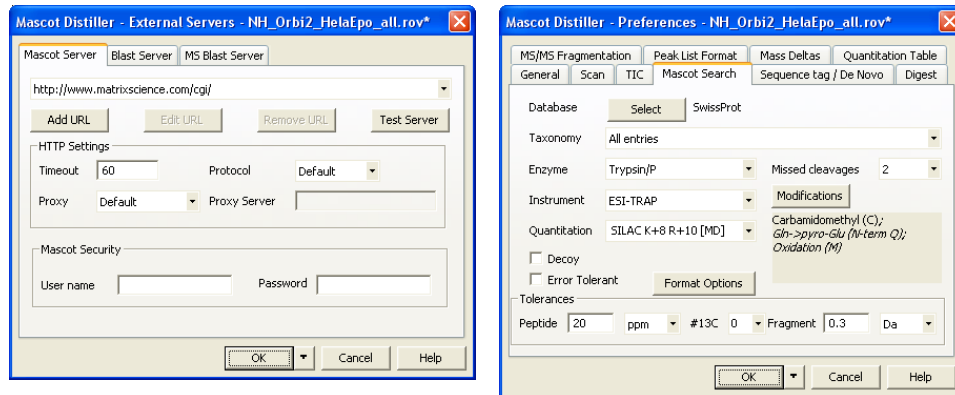
Topic: *Mascot Distiller*

© 2009-2010 Matrix Science

There are many other options dialogs in Distiller. Press F1 at any time to get context sensitive help. Two of the most important are shown here. External Servers is where you select your Mascot Server. Preferences is where you set defaults for many tools, such as Mascot search and de novo.

The MDRO library, which provides the raw file access and peak processing, can be called by other applications, such as Mascot Daemon. We'll go into more detail about how this works in another presentation. As you can see, the Import filter options are relatively simple because all of the processing options are specified by simply selecting an options file.

You can call the Mascot Distiller libraries from your own applications by purchasing a Developer Toolbox licence. Distiller uses COM, so can be called from most Microsoft Windows programming languages. The Developer Toolbox provides a uniform Application Programmer Interface to all of the different raw file formats, greatly reducing development time.

Automation: Developer Toolbox

Here is an example of calling Distiller from VBScript, which is a standard part of Windows. The object oriented interface makes the code very clean and simple.

Search Toolbox

- Import & display Mascot Search results
- Manual *de novo* sequencing
- Automatic *de novo* sequencing
- Automatic sequence tag calling
- Predict fragment ions from a peptide sequence and overlay them on an MS/MS spectrum
- Predict mass values from a protein digest and overlay them on a spectrum.

MASCOT   **Topic:** *Mascot Distiller*   *© 2009-2010 Matrix Science*   *MATRIX SCIENCE*

The Search Toolbox is a collection of tools for protein identification and characterisation
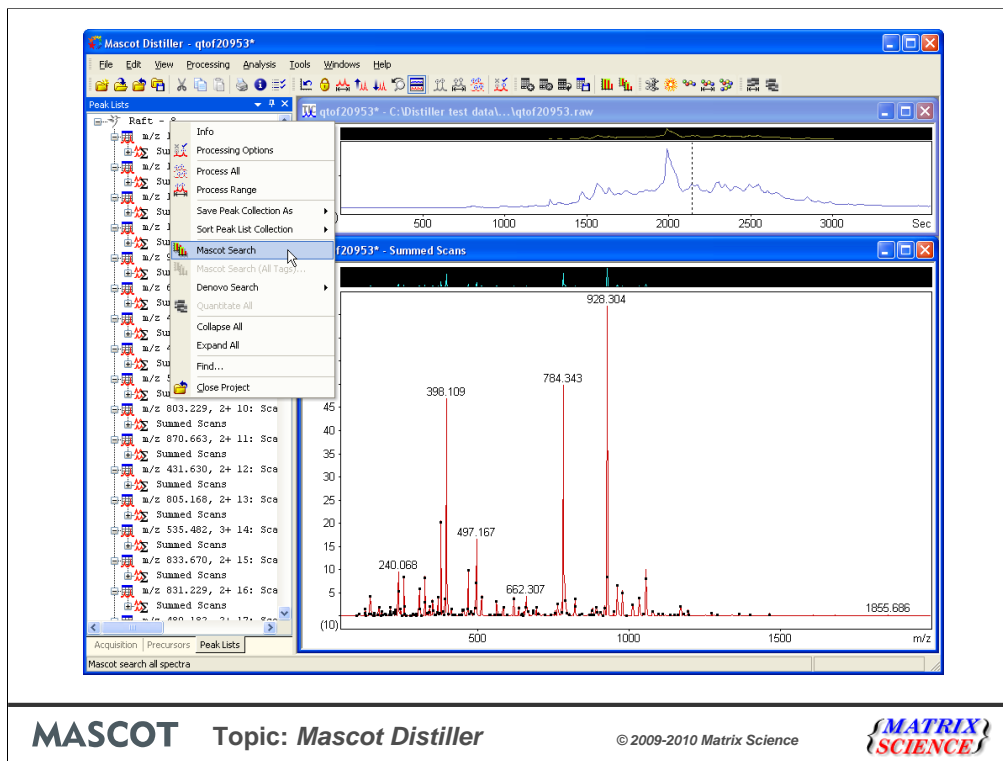
Lets see how to submit a Mascot search from Distiller. You don't have to save the peak list to a file. You just choose Mascot search from the context menu obtained by right clicking the top node of the peak list tree.

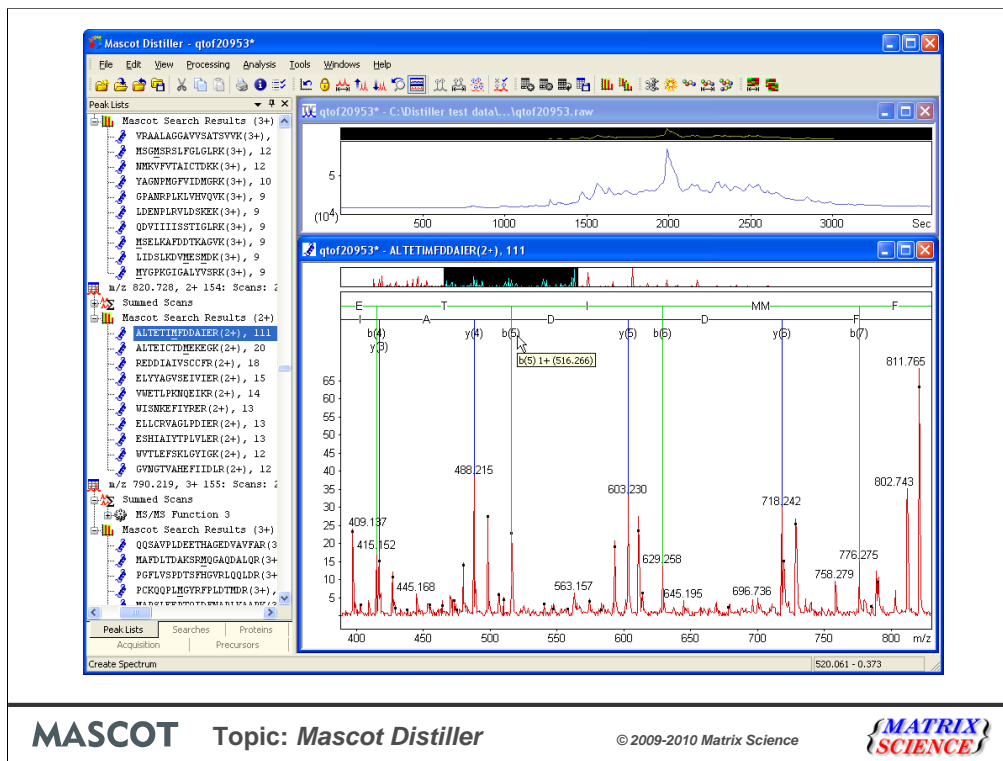MASCOT   Topic: *Mascot Distiller*   *© 2009-2010 Matrix Science*
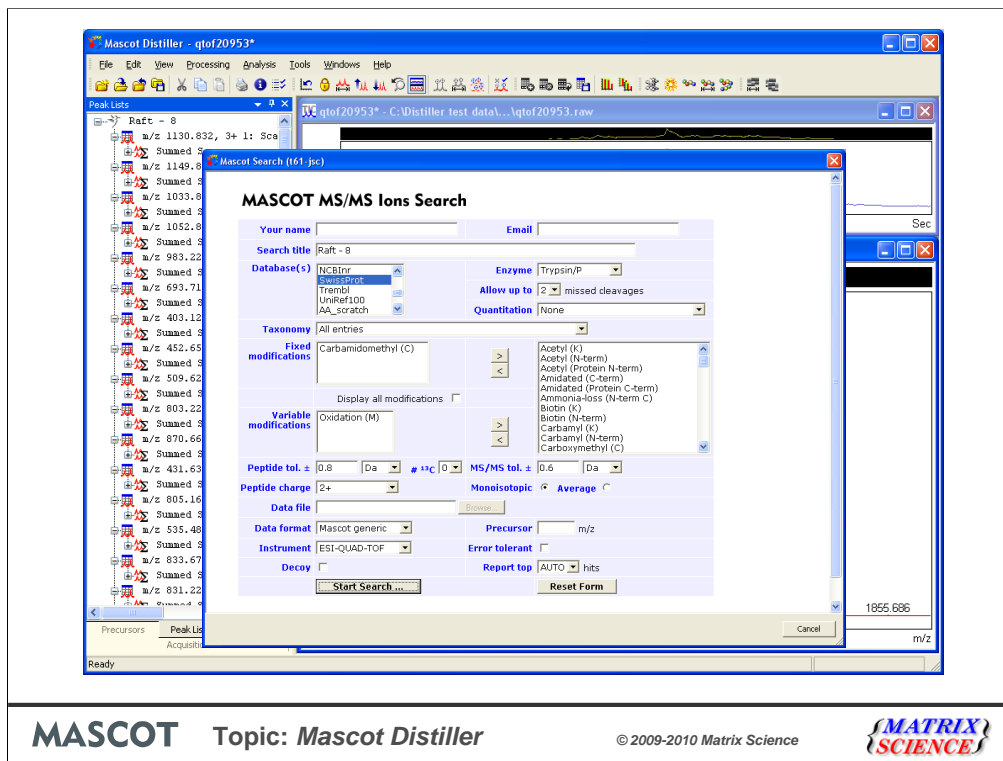
The data are loaded into a search form and the search parameters are set to defaults specified in the Distiller project preferences. The form allows you to make any last minute changes

Topic: *Mascot Distiller* © 2009-2010 Matrix Science
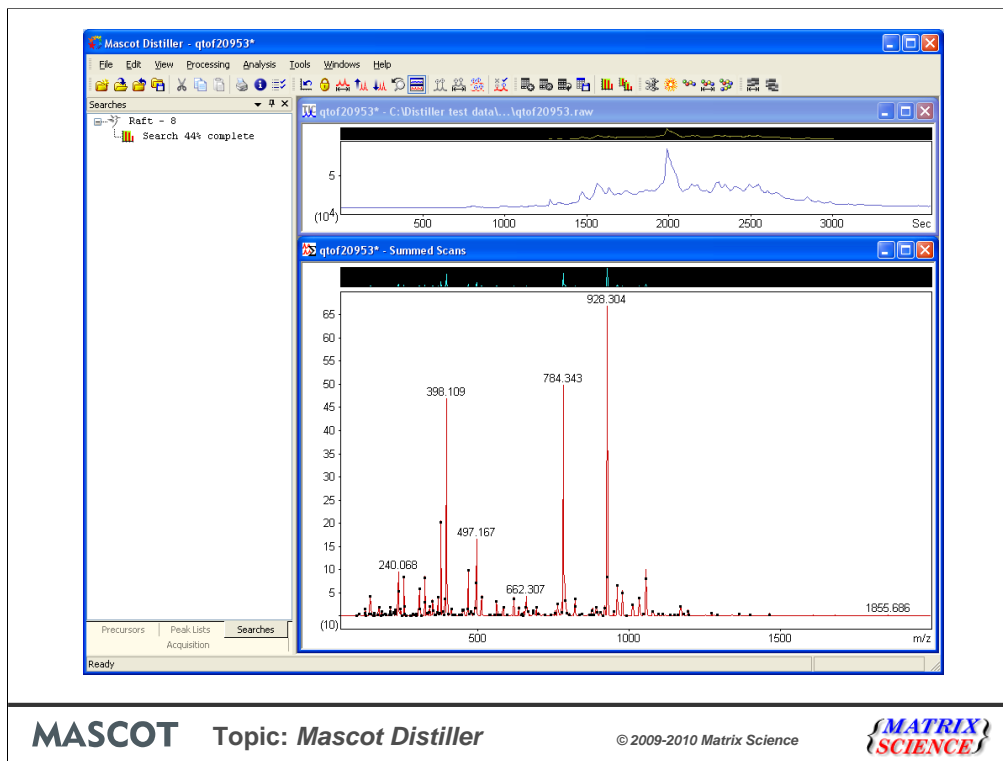
Press submit, the form will vanish, and a new explorer tab will display progress information. The tree can be used to move between multiple sets of search results.

When the search is complete, the behaviour depends on whether you have the search toolbox or not. If you only have the core application,, the results will be displayed in a browser window.
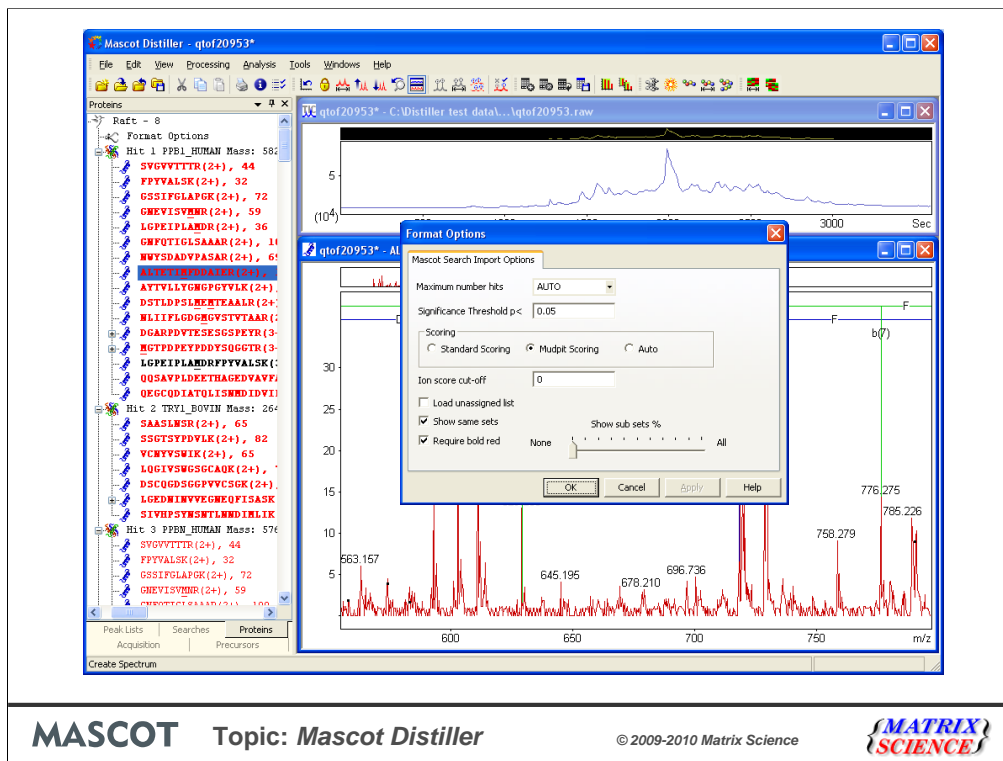
MASCOT   Topic: *Mascot Distiller*   © 2009-2010 Matrix Science

If you have the search toolbox, the results are automatically retrieved from the Mascot server and displayed in DataSet Explorer. In the case of an MS/MS search, each match is labelled with the peptide sequence and Mascot peptide match score. When a peptide match node is selected, the peak assignments are displayed as sequence ladders.

It is much easier to make a judgement about the quality of the matched peaks from this kind of display than from the bitmap in the HTML result report

The Mascot search results are actually displayed in two tabs. One is the peak lists tab, as shown here. This gives an immediate overview of all the matches to each query, and might be described as a spectrum-centric view

The Proteins tab provides a protein-centric view, closely resembling the Mascot Peptide Summary report, with the peptide matches grouped into protein hits. This is a more natural arrangement if you are only interested in the significant matches, or the matches assigned to a particular protein. We added this tab in Distiller 2.2 as part of support for quantitation.

As far as possible, the selected node is synchronised between the two tabs, making it easy to switch between the two views. The Format options node at the top of the proteins tree is equivalent to the Format controls at the top of the Mascot Peptide Summary report

Topic: *Mascot Distiller*   © 2009-2010 Matrix Science
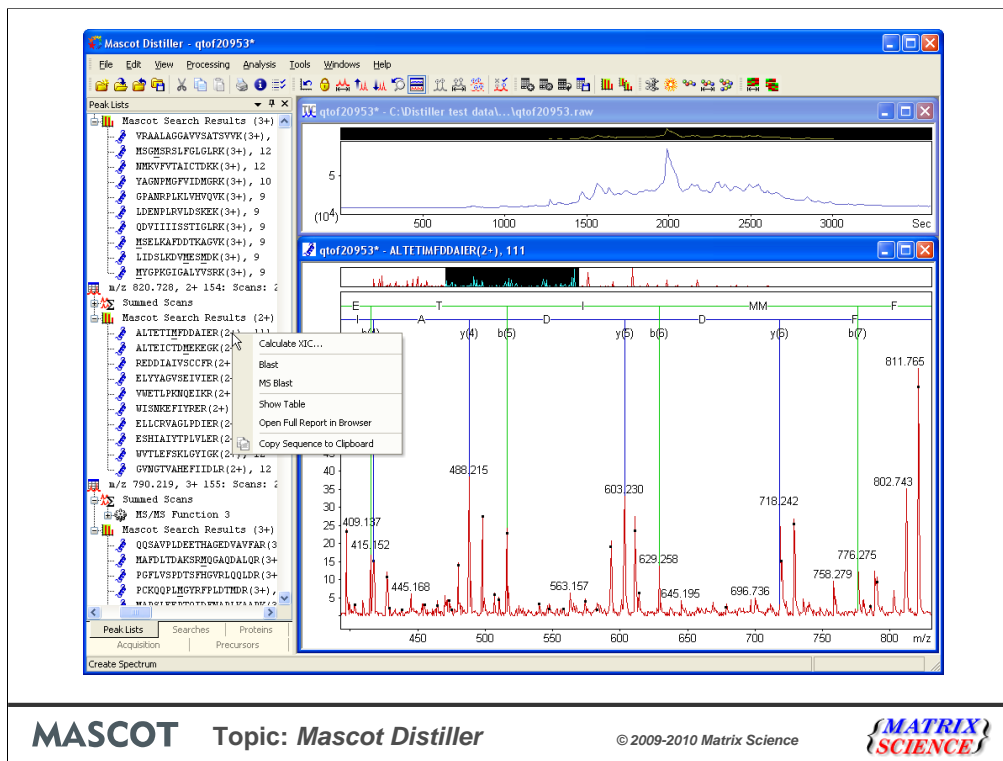
Going back to the peak lists tab, right click any match for a context menu containing relevant options, such as a link out to Blast or MS-Blast

De novo

- New algorithm
- Scores approximate to Mascot scores
- Single spectrum or entire dataset
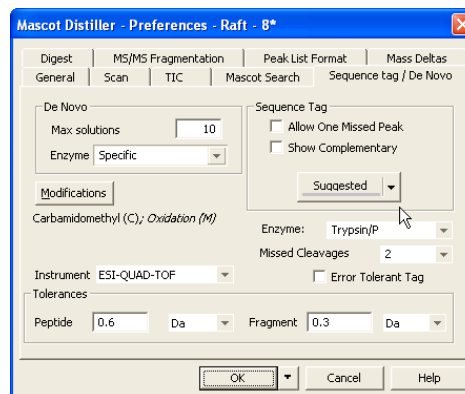
**MASCOT**   Topic: *Mascot Distiller*   © 2009-2010 Matrix Science   *MATRIX SCIENCE*

An important feature of the search toolbox is de novo sequencing.

The score assigned to a de novo solution is a Mascot score. In general, these scores will be higher than you expect to see in a Mascot database search, because the algorithm has selected the best matching sequence from all possible sequences, rather than the limited number of sequences found in any database. So, you should not judge the quality of the match by applying any rule of thumb or significance threshold to the score. However, if you get the same solution by de novo and by database search, using identical parameters, you should find the Mascot scores are very close.

You can de novo sequence a single spectrum or all the MS/MS spectra in the file

Topic: *Mascot Distiller* © 2009-2010 Matrix Science
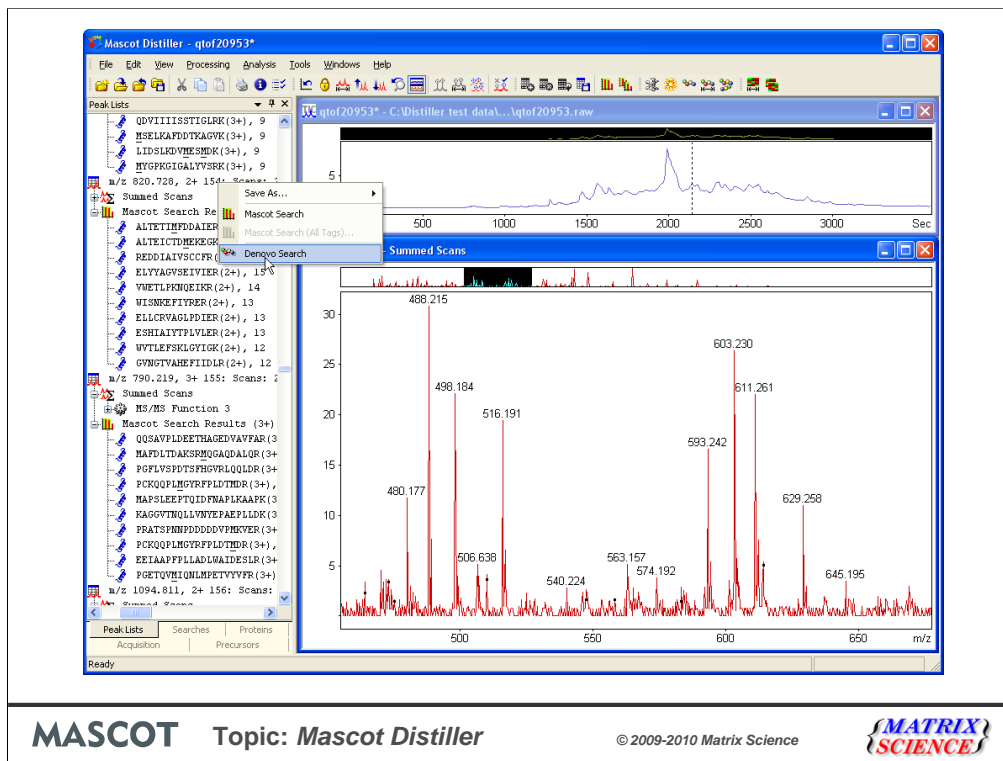
The starting point can be any MS/MS scan that has been processed to create a peak list. Right click the peak list node in Dataset Explorer and choose 'de novo Search', or choose the de novo button from the toolbar when a Summed Scans node is selected.

MASCOT Topic: *Mascot Distiller* © 2009-2010 Matrix Science *MATRIX SCIENCE*
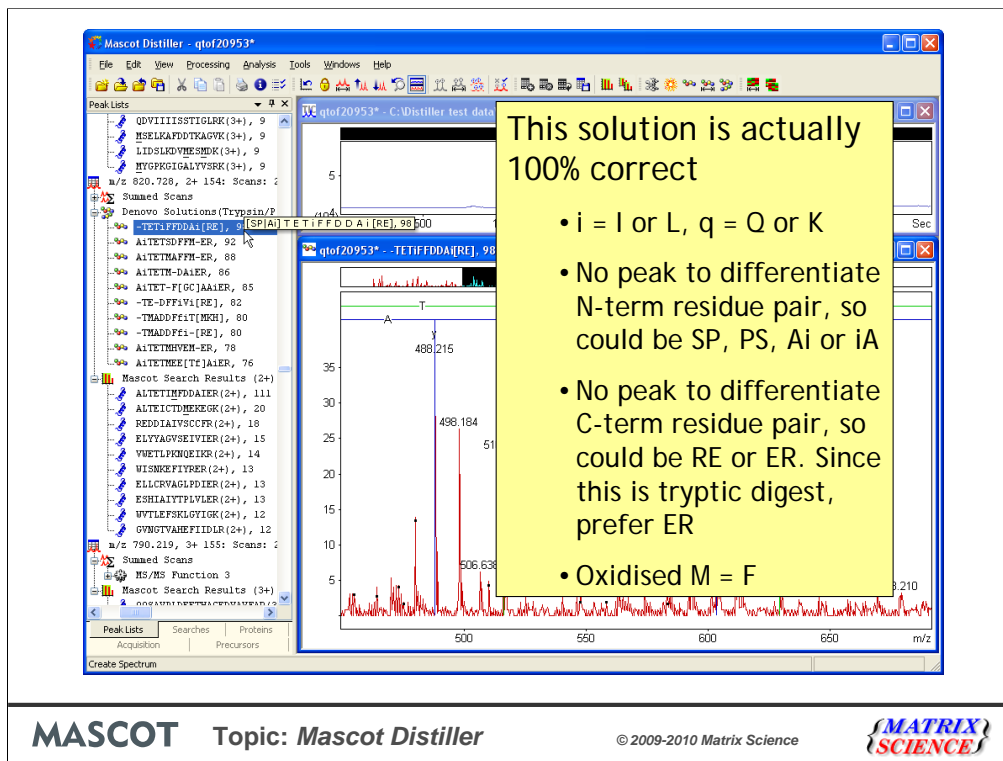
Good signal to noise and good mass accuracy are critical for successful de novo sequencing; much more so than in database searching. GIGO (garbage in - garbage out) is guaranteed.

In a de novo solution, i always represents I or L, and q represent Q or K when the mass tolerance does not allow these residues to be distinguished. However, K is assumed at the C terminus of a peptide when tryptic specificity applies

Ambiguity is indicated by a dash in the sequence. The tooltip shows details of the ambiguity in square brackets, using pipe symbols to separate alternatives. Note that the order of the pairs and triplets is undefined, so that SP could also be PS.

Although the example shown here looks very different to the Mascot database match, they are actually in perfectly agreement. Some uncertainty is unavoidable in de novo, because the search space is so very much larger
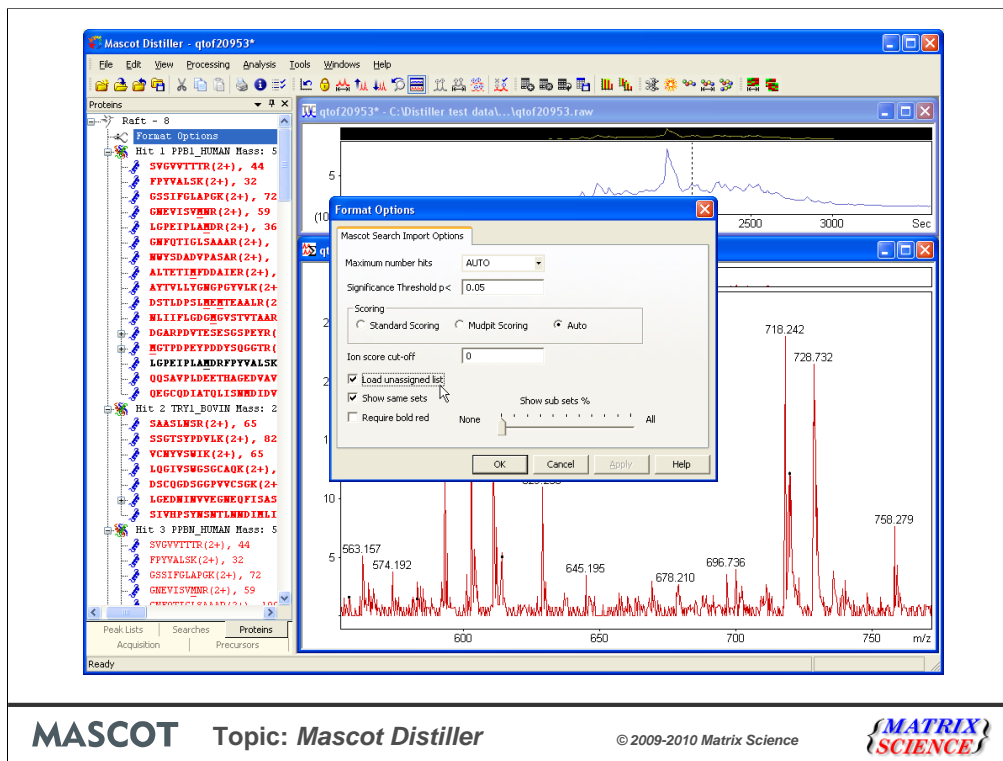
MASCOT   Topic: *Mascot Distiller*   © 2009-2010 Matrix Science

To de novo sequence a complete peak list collection, or the peak lists in the currently displayed TIC range, use the context menu obtained by right-clicking the root (world) node of the peak lists tree.

The most efficient way to de novo only those spectra that failed to give decent matches in the Mascot search is to switch to the proteins tree, click on Format Options, and choose to load the unassigned queries.
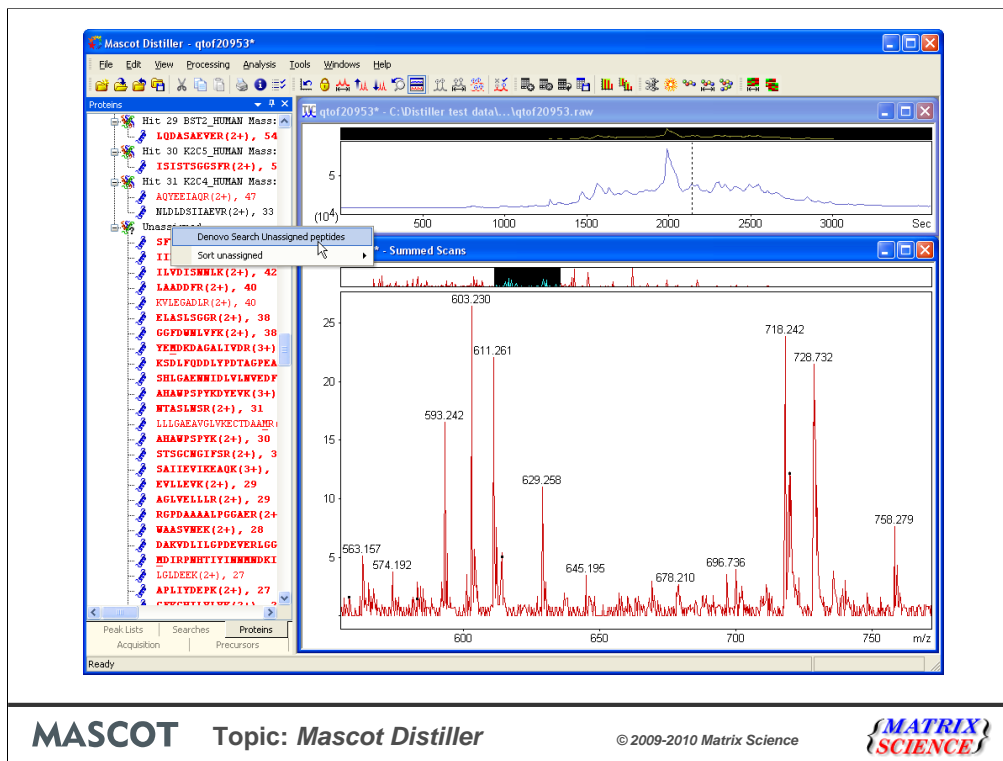
Use the context menu obtained by right-clicking the unassigned node to de novo just the unassigned queries.

MASCOT   **Topic:** *Mascot Distiller*   © 2009-2010 Matrix Science   MATRIX SCIENCE
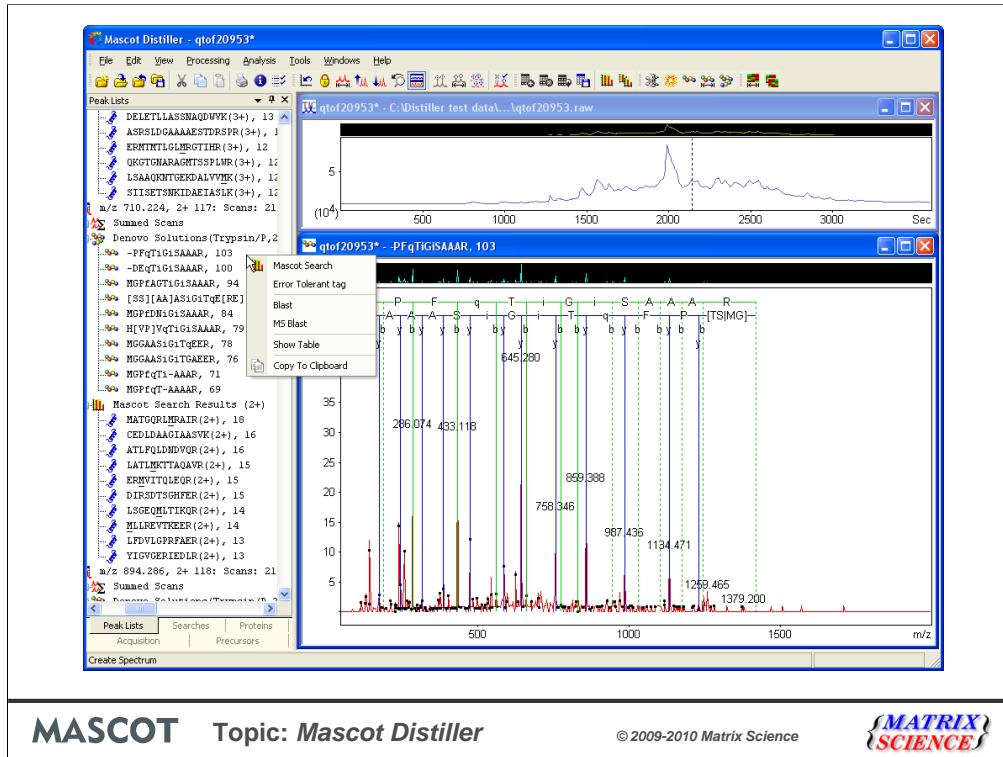
The de novo solutions are added to the peak lists tree, and you can browse down, looking for cases where the database search failed and de novo has a high score.

This looks like a promising case. But, is it right? And, how do we resolve the ambiguity at the N-terminus? One of the most powerful checks is to run an error tolerant sequence tag search, and see whether the match is a modified sequence from a known protein, as we'll discuss this in the next talk.
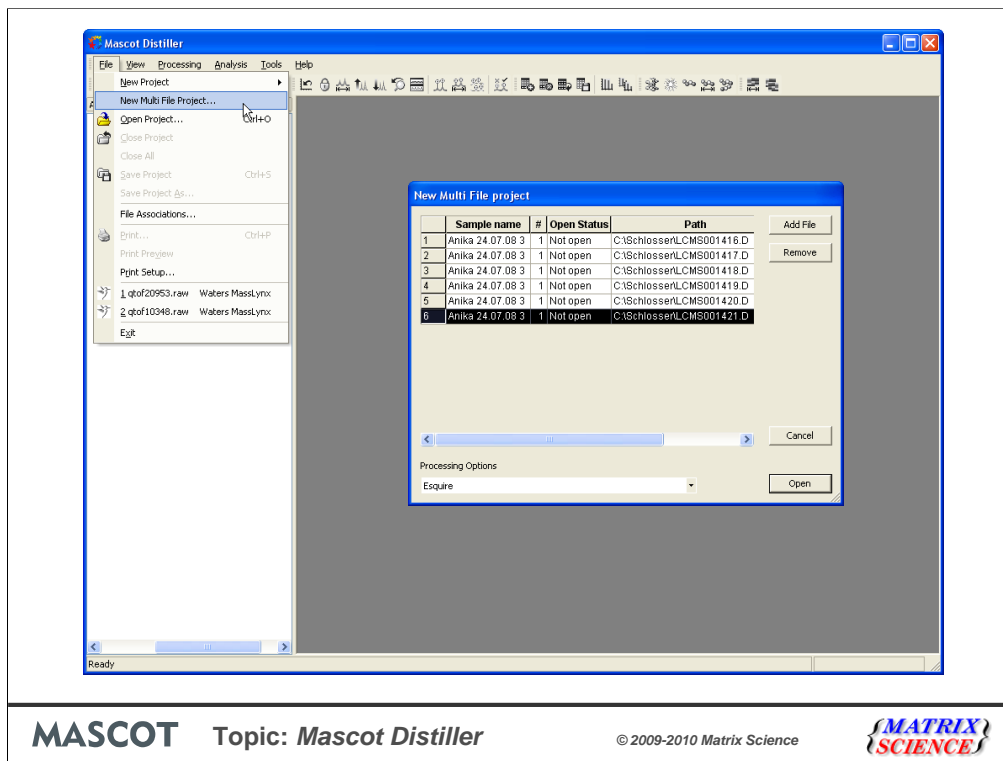
A Distiller project can contain more than one raw file. Choose New multi-file project from the File menu to invoke a file selection box
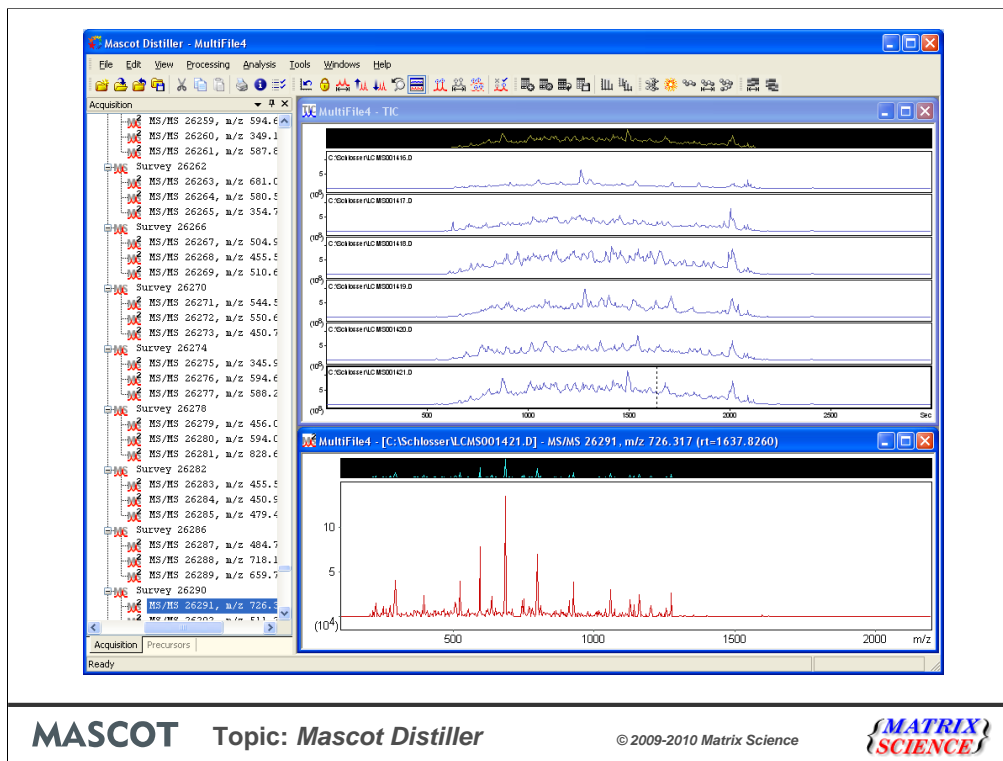
This can be very useful for data from mudpit experiments. By processing the files as a single project, spectra with a similar mass and elution time from each of the files can be summed together. This removes much of the redundancy we see in mudpit search results when the raw files are processed separately.

Projects

- **Can save workspace to a project file (*.rov)**
  - Includes peak lists, search results, tags, etc.
  - Does not include the raw file … just a path reference
- **Projects can also be saved by Mascot Daemon**
- **Project files are zip archives of XML files**
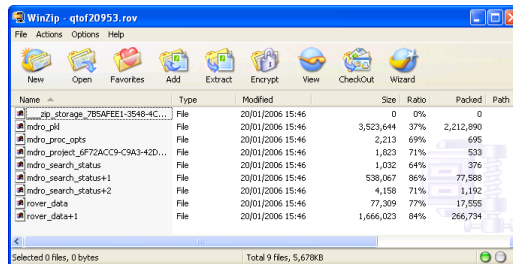
MASCOT **Topic:** *Mascot Distiller*   *© 2009-2010 Matrix Science*   **MATRIX SCIENCE**

Having processed the data and maybe performed some Mascot searches and de novo sequencing, you will often want to be able to save everything in the workspace. You can do exactly that. The only thing that doesn't go into the project file is the raw file itself, for space reasons. If the reference to the raw file is broken, you can easily re-attach the raw file when opening a project.

You can also choose to have Mascot Daemon save Distiller project files.

And, if you are curious, Distiller project files are just zip archives of XML files, so you can take a look at what is in there.

Finally, a few of the most frequent technical support questions:

1. Some of the file access libraries are not fully localised. In Windows regional options, you need to set the decimal separator to a period (full stop)

2. To open Xcalibur 2.1 files, install MSFileReader (for details, see the Xcalibur help page on the Matrix Science web site)

3. De novo settings are in the preferences dialog. Most of these settings will be the same as on the Mascot search tab

4. Agilent MassHunter support is a setup option. This is because it requires Microsoft .Net 2.0 and this is a big install if you don't need it

5. If you don't register and save a licence, Distiller is a read-only project viewer

6. Need to install Adobe SVG plug-in for Internet Explorer to view graphics in quantitation reports. Alternatively, right click in the report window and open in a browser with native SVG support, such as FireFox or Opera