Positive Probability Ltd

Note M4: LCMS Analysis - Peptides

Introduction

Protein identification and the search for biomarkers both make use of the LCMS of peptides from protein digests. Modern instruments operate at high resolution but centroiding is rarely adequate for resolving high m/z, high charge isotope clusters and deconvolution is normally required before deisotoping.

PPL's LCMS analysis program will co-add scans to improve S/N, baseline correct and deconvolve to produce reliable peak tables prior to deisotoping. Unlike other programs, the PPL ***Collapse***™ program will reliably deisotope multi-charge ESI data up to at least Z12 and higher. Background ions, which may be numerous for high S/N data, are identified and removed but genuine elutions that are present on top of background ions are correctly identified and retained. The elutions are then resolved and results presented as tables and maps of m/z (or zero-charge mass), RT and intensity along with all the errors.

Data

The data presented here are 340 scans from the LCMS of a HSA digest. The region covered is from retention time 121-138 minutes (17 minutes of experiment time). Note the intense background ions.

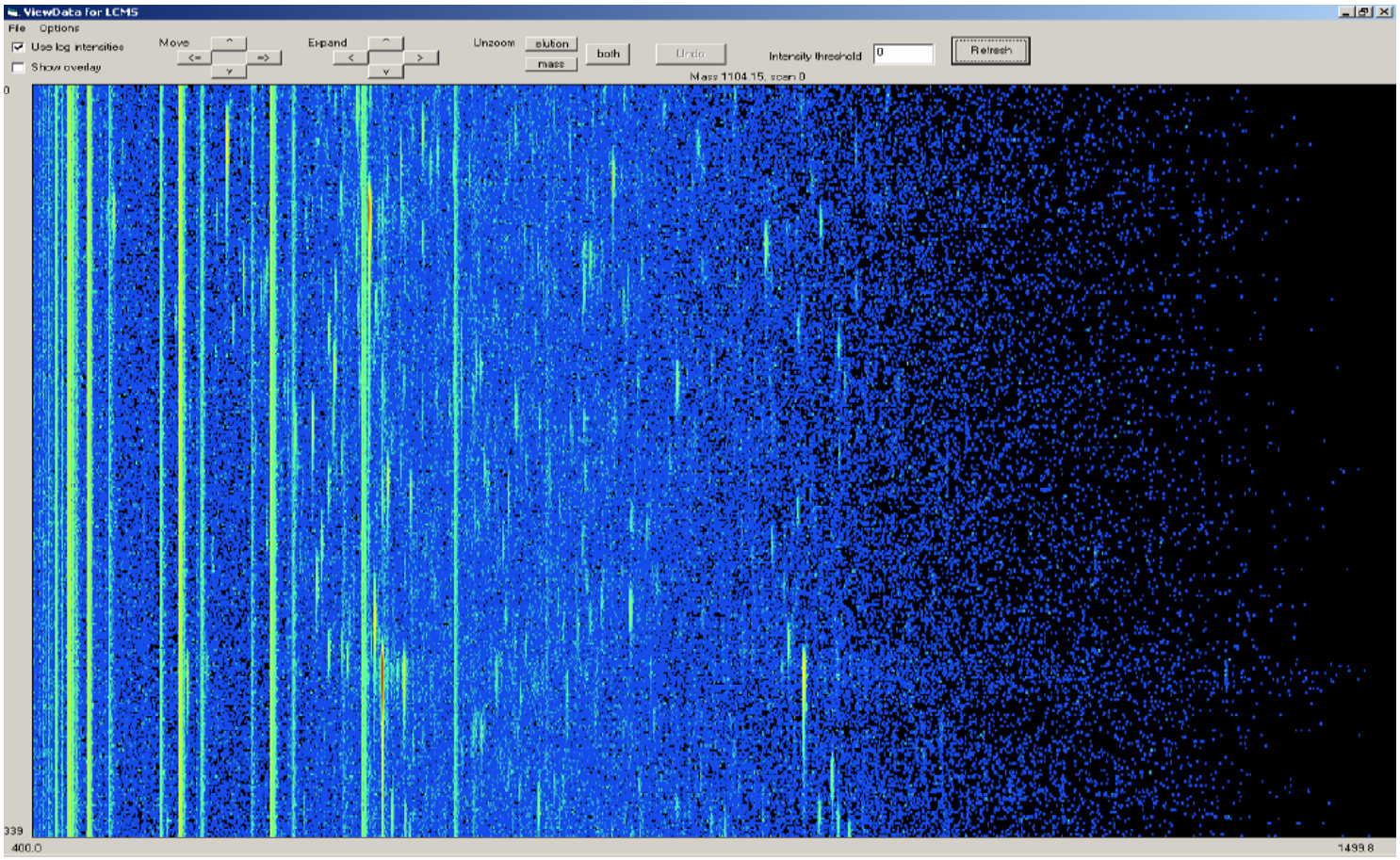

Figure 1. 340 scans from the LCMS of a HSA digest – 17 min. of experiment time.

Evaluating the Chromatography

Scans may be co-added to generate 'blocks' before processing. The number of scans to be co-added may be varied. To improve S/N and retain retention time information, blocks may overlap. A mass must be present in a chosen number of adjacent blocks before it is considered to be genuine and not noise.

Zooming in to the data shown in Figure 1 (Figure 2 below) shows the poor quality of the data and that many peaks are at the noise level.

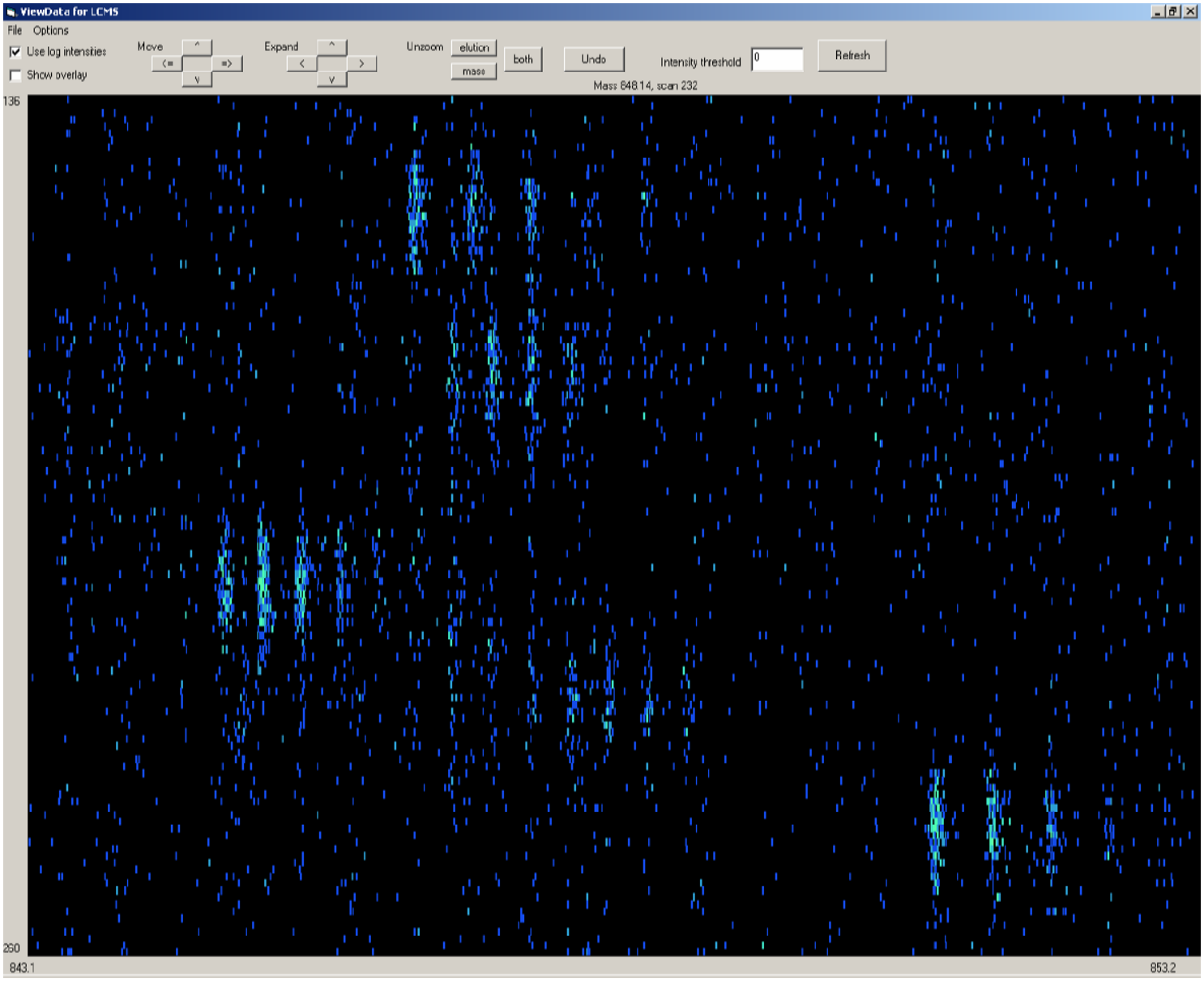

Figure 2. Scans 136-260 and m/z 843-853.

Zooming clearly shows the poor quality of the data. Note that many peaks are indistinguishable from the noise. Elutions for significant peaks (C12 & C12+1) in this region occur over about 15 scans. Because the data are of particularly low S/N, co-adds of 8 scans with an advance of 4 scans and 3 adjacent blocks to confirm an ion would be reasonable.

Data Processing and Results

Preparing the Models

The peak width will be dependent on charge but, depending on the instrument and the acquisition conditions, peaks may also vary with m/z. By co-adding a few suitable scans, models may be generated at different m/z so that any variation in peak width may be established. Any variation will then be taken into account during the deconvolution of each block prior to deisotoping. Note that the models may be saved and used for any experiment run under similar conditions. Each model may be displayed and their widths displayed as a function of m/z so that the list may be edited.

Figure 3 shows the Standards Page (model standards). Models have been generated for the isotope clusters and clusters of peaks throughout the data. The modelling program automatically models the individual isotope peaks and not envelope profiles, even when peaks overlap. It is important to determine the trend in peak width. Because the deconvolution is very tolerant of the models, those that do not conform to any obvious trend should be removed. For these data, Model 4 that happens to be a high charge ion, was significantly narrower than the trend and was therefore rejected. The figure shows Model 6 (red trace) that was obtained by modelling all the peaks in the displayed modelling window.

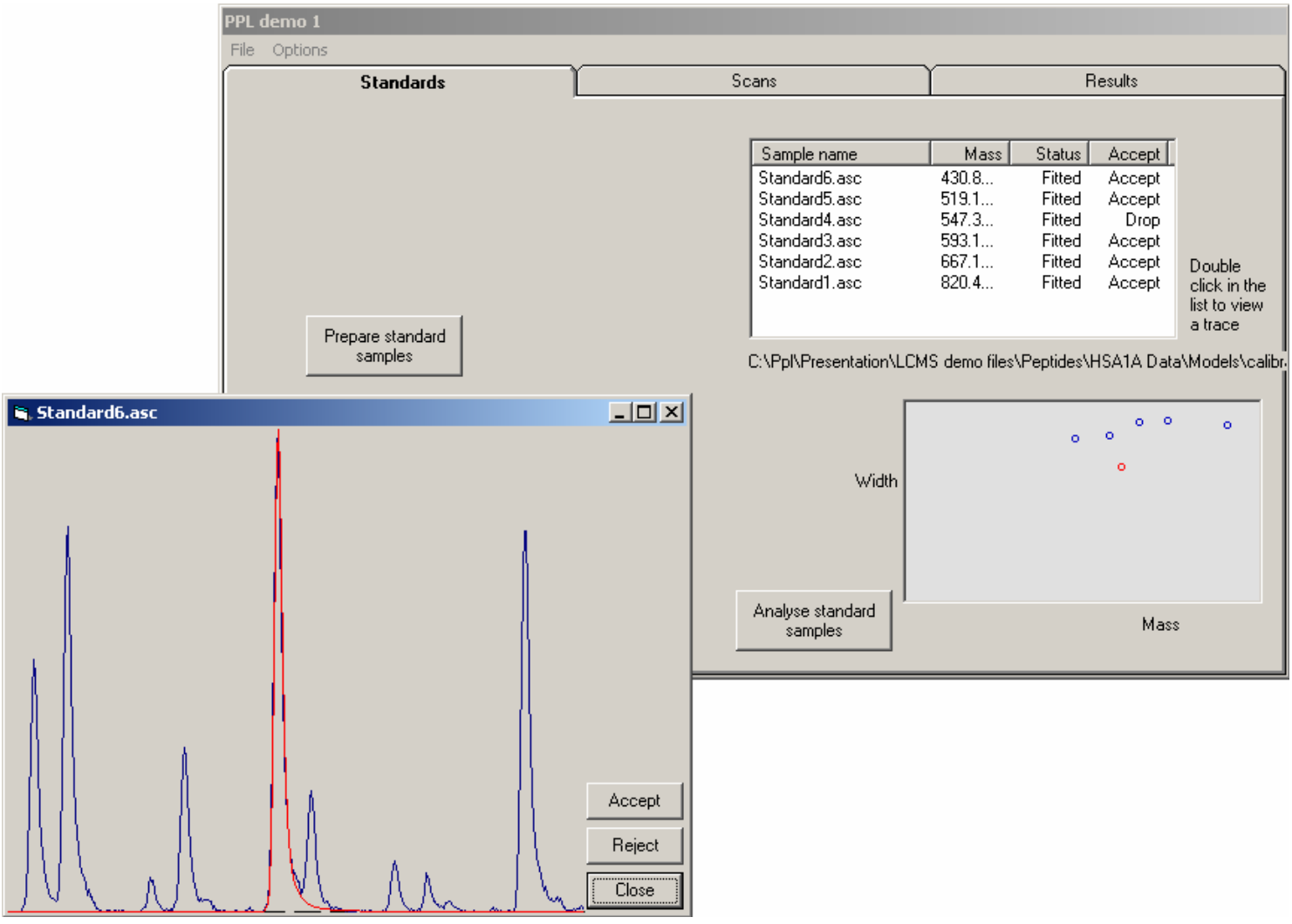

Figure 3. Modelling peaks to determine the way the peak width changes with m/z.

Setting Processing Parameters

The “Scans page” is for setting the input parameters used for processing the data. The chromatography options uses the conditions determined from Figure 2. The deisotoping section allows a suitable unit cell formula (empirical formula) to be input. This determines the way the isotope intensity profiles change with m/z and charge. The minimum and maximum charges that are likely to be present in the data set at the user-expected values and the mass error provides additional freedom for noisy data. C12 positions can be computed from either the reconstructed C12 peak or the first two peaks, the latter being preferred since it reduces the effect of noise. The breakpoints set the number of isotopes required to identify an isotope cluster for different masses. The analysis section allows the deconvolved peak table for each block to be filtered at different confidence levels to remove noise. The panel in the analysis section shows the progress of the computation. Note that that it is unnecessary to reprocess the data from the beginning when determining the method and parameters to be used – standards, deconvolutions and elution profiles may be saved and reloaded for processing with different parameters. Figure 4 shows the processing parameters used for these data.

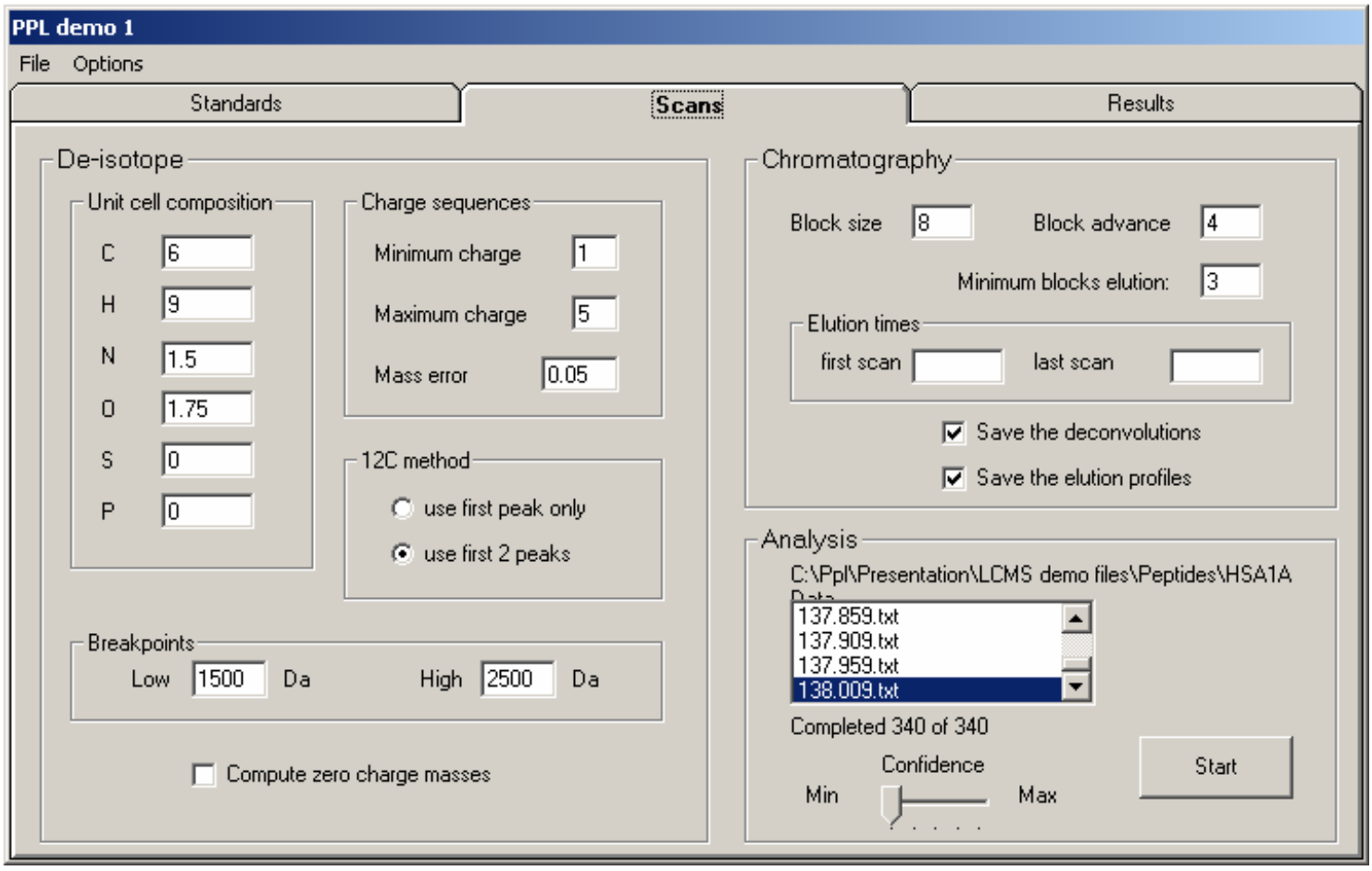

Figure 4. Input parameters for processing the data described here.

At the end of the deconvolution of each block the peak table is filtered according to the set confidence level and deisotoped to C12+1 or zero-charge. Background ions are identified and removed and elution profiles computed along with their retention time and associated errors. Only masses that are present in at least the set minimum number of adjacent blocks are retained. The confidence of each retained mass is known so that the results table may be further filtered to remove obvious noise.

Results

At the end of the computation the results are presented in the “Results Page” as shown in Figure 5.

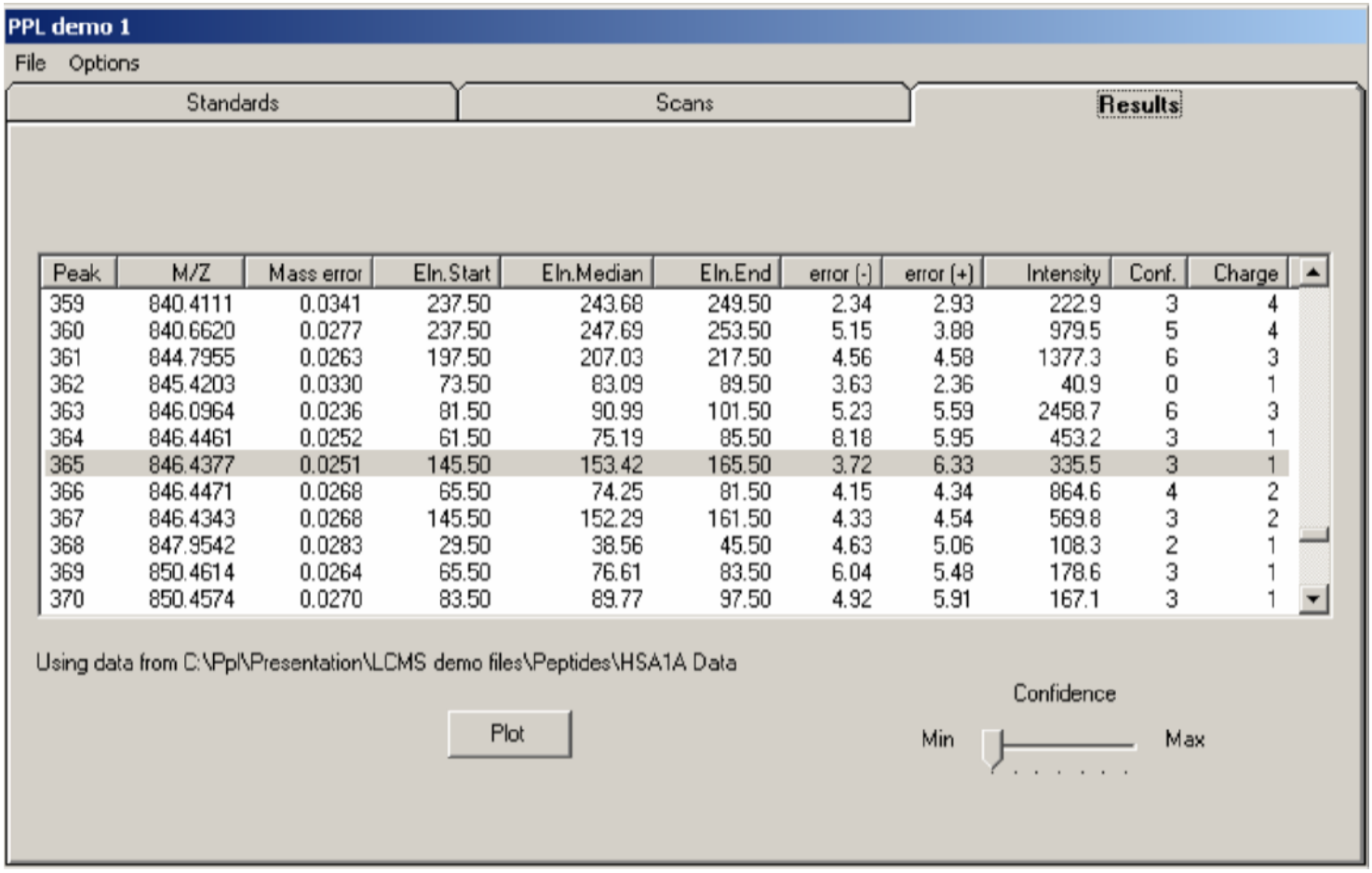

Figure 5. The “Results Page”.

The information contained in the peak table is the reconstructed m/z (or zero-charge mass) and its associated error, the elution median as scan number (or RT), its start and end along with RT errors. These are presented as separate (-) and (+) errors because elutions are generally asymmetric in shape. The reconstructed intensities and confidence levels for each feature are also output.

Note that for simple deisotoping the charge of each ion is shown. For zero-charge results the information about which charges were reconstructed is available. Results peak tables may be loaded into Excel for formatting and sorting according to user requirements. Results may also be presented as a 2-D plot. For simple deisotoping the results may be overlaid with the data.

For these data the total processing time for a 3 GHz single processor P4 computer was 32 sec. for the 17 minutes of experiment time. The processing time is therefore less than 5% of the acquisition time. Given appropriate integration, the processing could be performed on-line so that results were available at the end of the experiment.

The data with the results overlaid are shown in Figure 6 below. The displayed region corresponds to that shown in Figure 2.

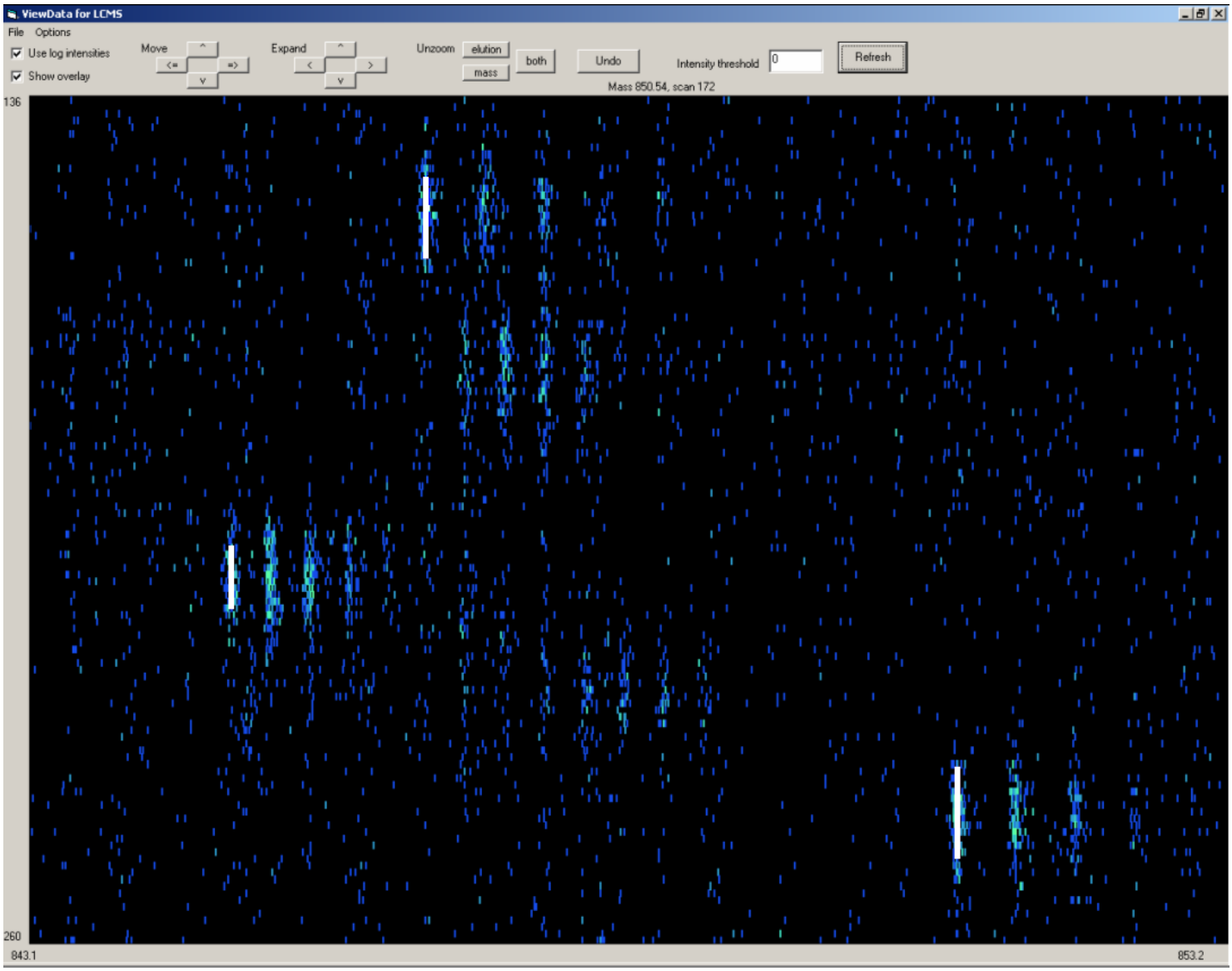

Figure 6. Scans 136-260 and m/z 843-853. Data overlaid with result.

In Figure 6 the white bars represent the 1 SD m/z and RT errors. Note that there are two extremely weak ions that are not reconstructed because they fail to break through the chosen input requirement that an ion must be present on at least three adjacent blocks.

Conclusions

The ***ReSpect***™ data reconstruction methodology has the following benefits over other methods:

1. It is very fast and can therefore operate in real-time.
2. Background ions are easily identified and removed.
3. Results may be presented at any user-chosen confidence level.