



Management of Ultrahigh-Resolution FT-ICR Mass Spectrometry Crude Oil Data

Esha Atolia¹, Yuri E. Corilo², Ryan P. Rodgers^{2,3}, Amy McKenna², Brian M. Ruddy³, Chris L. Hendrickson^{2,3} and Alan G. Marshall^{2,3}

¹Department of Biological Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139

²Ion Cyclotron Resonance Program, National High Magnetic Field Laboratory, 1800 E. Paul Dirac Drive, Tallahassee, FL 32310-4005 ³Department of Chemistry and Biochemistry, Florida State University, 95 Chieftain Way, Tallahassee, FL 32306.

Introduction

Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) uses the frequency of cyclotron motion of the ions in a static magnetic field to determine the mass-to-charge ratio, which is then used to determine the elemental composition of the molecule. The cyclotron equation:

$$\nu = \omega/2\pi = (1.535611 \times 10^8 \text{ B}) / (m/z)$$

allows for this conversion. Note that the ICR frequency for different m/z values is distinct and independent of velocity. Furthermore, FT-ICR MS provides high mass resolving power and mass accuracy required to analyze complex organic molecules such as heavy crude oils.

Heavy crudes are some of the most complex organic mixtures found in nature. As the crude oil industry grows in size and demand for crude oil increases, techniques such as FT-ICR MS that provide high resolving power have become increasingly important. FT-ICR gives a mass resolving power of > 400,000 at m/z = 400 and an average mass error of < 300 ppb, allowing thousands of peaks to be resolved and identified at the level of molecular formula assignment.

FT-ICR MS spectra contain 1,000 to 100,000 peaks and thus a large amount of data must be efficiently processed; data handling and data management are important issues. The presently proposed software will streamline the data analysis process by making it more automated and comprehensive. We also address data processing issues such as unassigned peaks and peak splitting.

Technique: FT-ICR MS

In FT-ICR MS, ions traverse a capillary and then pass through rf ion guides into the analyzer cell that is located at the center of the superconducting magnet. The magnet confines the ions radially whereas the ICR cell confines the ions axially. The analyzer cell consists of four electrodes forming a cylinder. The top and bottom electrodes function as detectors, whereas the main function of the side plates is excitation (Marshall et al., 1998). For detection, the ions are excited by a uniform electric field oscillating at close to the cyclotron frequency of an ion that depends on the ion m/z ratio; the oscillating field is generated by sending radio frequency pulses through an AC circuit. Excitation increases the orbital radius and the kinetic energy of the ion to enable detection.

After excitation, detection is achieved by opposed plates connected to an RC circuit as shown schematically in Figure 1, to generate a time domain signal. The time-domain signal from the detector is transformed into a magnitude vs. frequency by means of a Fourier transform. This mathematical transformation is performed automatically done by NHMFL's PREDATOR software. The program also automatically transforms the frequency domain signal to an m/z domain signal (ion relative abundance vs. m/z) by determining approximate values for B and E (Figure 1).

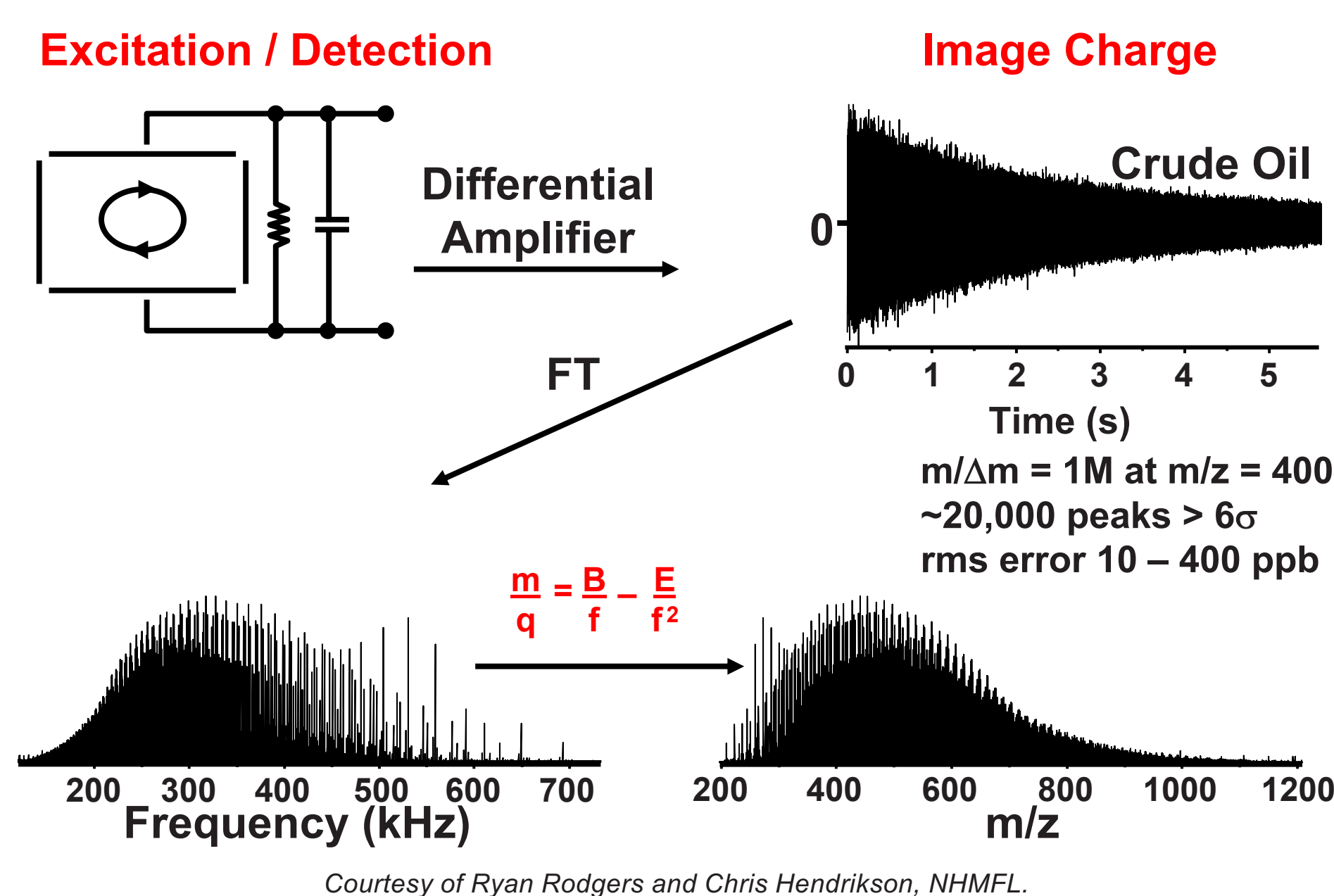


Figure 1. Generation, detection, and processing of a time-domain ICR signal to yield an m/z spectrum. The detection of ions leads to a time domain signal, which is transformed into frequency data by use of the Fourier transformation. ICR frequency is then transformed into mass-to-charge ratio (m/z). As shown, higher frequency corresponds to a lower m/z.

These transformations are necessary for any meaningful use of the data. The resulting m/z contains one peak for each ion elemental composition. Before final analysis, the m/z domain signal needs to be carefully calibrated, i.e., by determining accurate values of the parameters B and E for the particular spectrum.

Discussion

During data analysis/processing two major problems are encountered: unassigned peaks and peak splitting. Our program deals with these issues by making the data processing more automated and sensitive to these issues.

Unassigned Peaks (No Hits). MIDAS is a program designed at Florida State that automatically assigns the peaks in a given spectrum based on user specification of heteroatom and radical content. Thus the program has a landscape of molecules, constrained by user input and by known chemistry of crude oil, that it searches to assign molecular formulas to the peaks in the spectrum. For example, the program exploits homologous series whose members differ by integer multiples of -CH₂ groups (Kendrick series) to assign elemental compositions. However, a small percentage of peaks of low signal-to-noise ratio remain unassigned as shown in Figure 2. Manual reassignment of peaks is slow. The program semi-automates this process by reassigning an entire Kendrick series, given a single assignment in the series.

Negative Ion ESI Ft-ICR MS at 9.4 Tesla

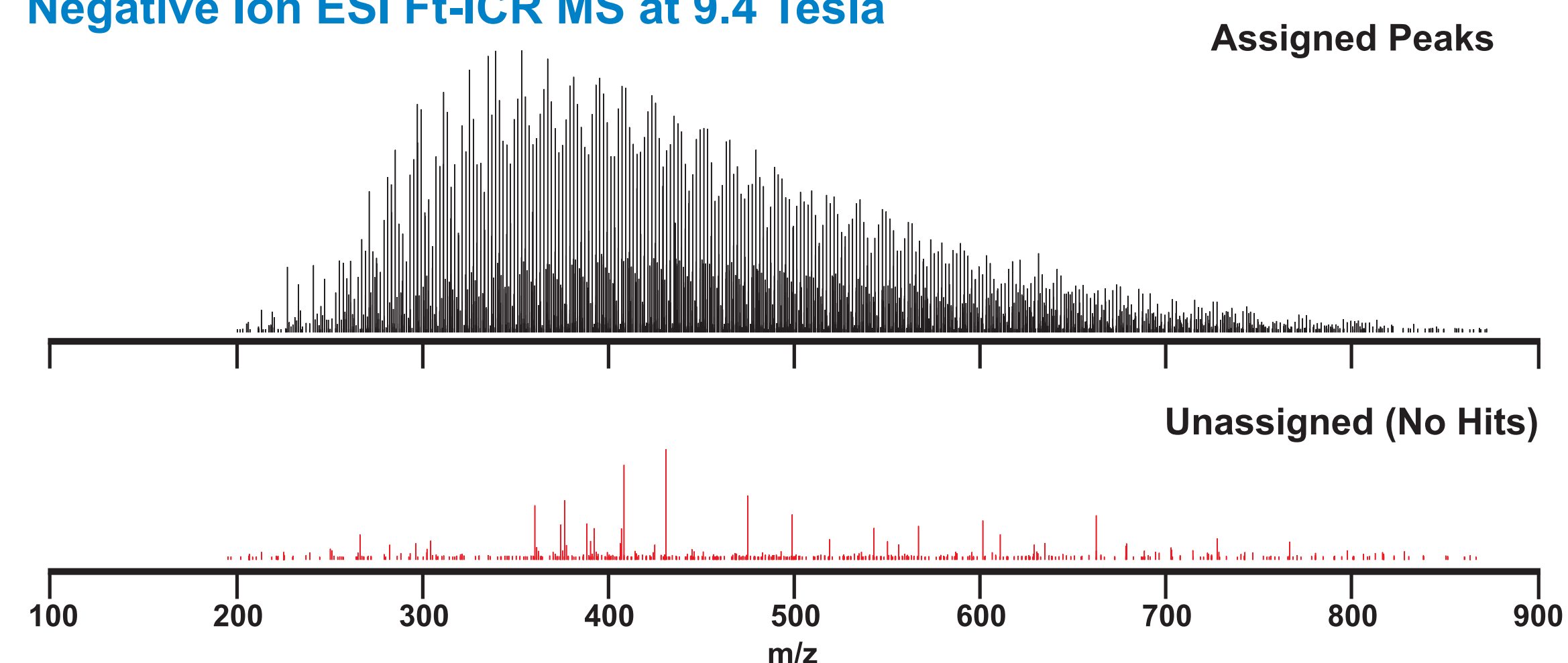


Figure 2. FT-ICR mass spectra separated into assigned and unassigned peaks. Unassigned peaks generally constitute a small percentage and have low signal-to-noise ratio.

Positive Ion ESI Ft-ICR MS at 9.4 Tesla

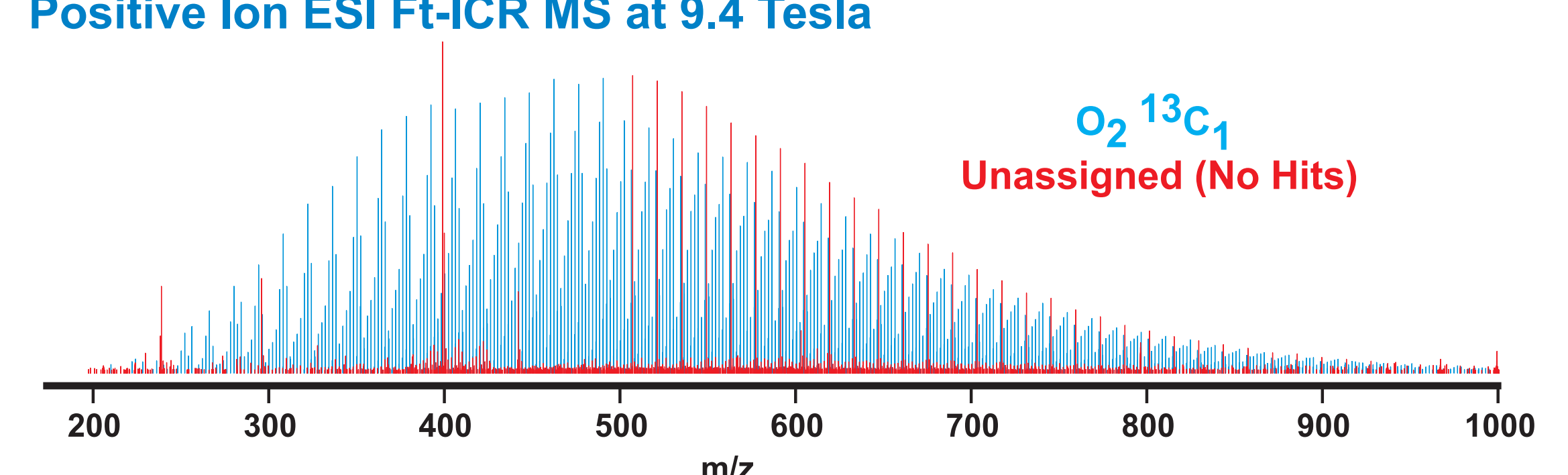


Figure 3. Unassigned peaks, also called the 'No Hits', contain peaks that often belong to an already existing class. A heteroatom class of molecules has a common O, N, and S elemental composition, but different numbers of CH₂ groups. The combination of all homologous series forms a broad m/z distribution. The spectrum clearly shows that the No Hits contain some O₂¹³C₁ peaks because of the clear pattern that some of the No Hits form with the already assigned O₂¹³C₁ peaks.

Peak Splitting. Peak splitting is when a set of ions of one mass leads to two peaks very close in mass instead of one. This phenomenon is caused by a number of factors, but the two major factors are the "comet effect" and inhomogeneous magnetic fields. The comet effect is where a part of the large cloud of ions in the ICR cell trail the large cloud, which leads to two different detected masses very close together. The second reason is that the magnetic field is not constant due to Coulombic shielding effects. Thus all the ions do not feel the same magnetic field and this discrepancy can also lead to peak splitting. This is an important issue in data analysis because currently it is left up to the user to sort through the data and manually hunt down these peaks. The program has automated this process by identifying molecules with the same molecular formula and labeling the lowest abundant peak as unassigned.

Positive Ion APPI FT-ICR MS

at 9.4 Tesla

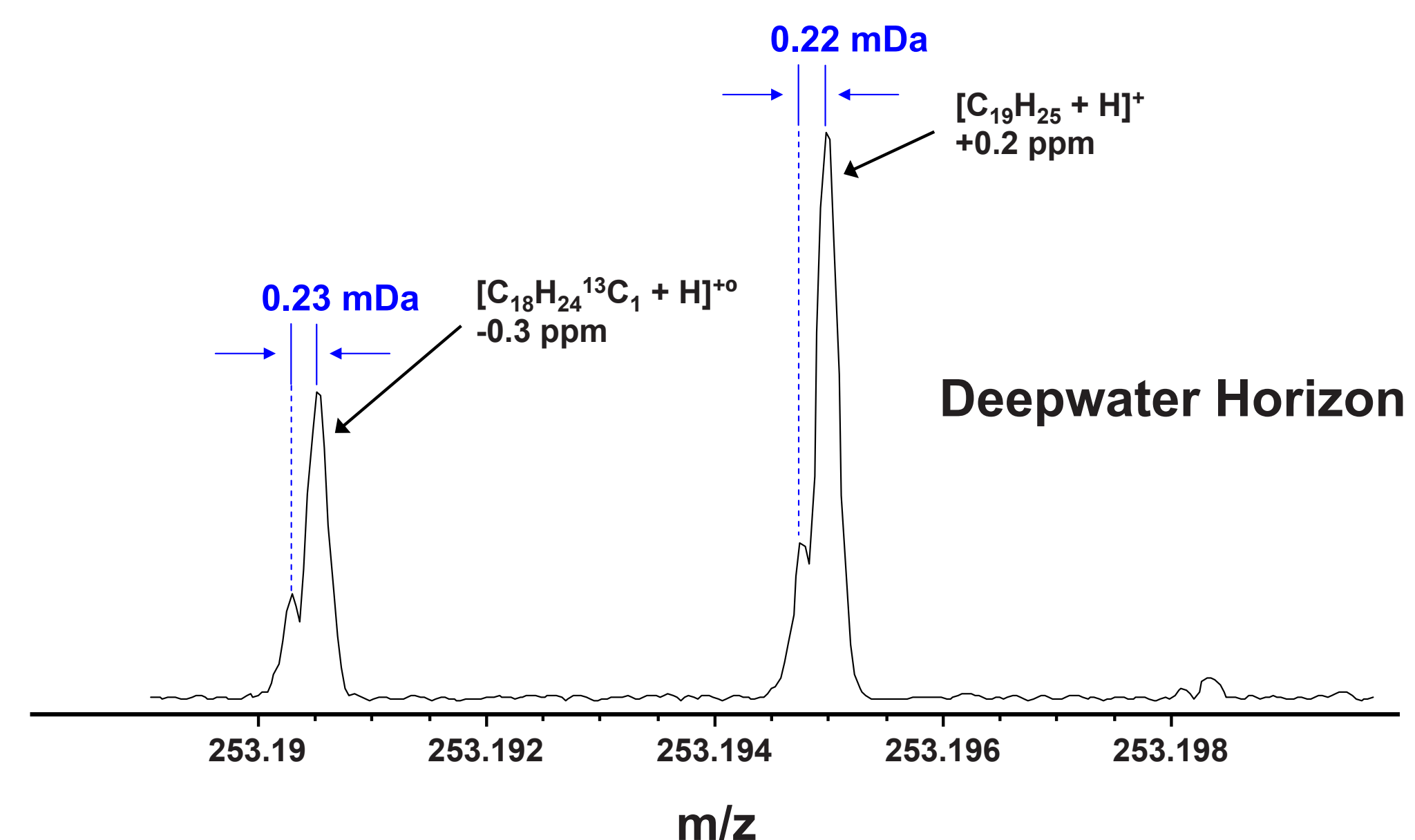


Figure 4. Example of peak splits, showing mass separation, assigned elemental composition, and mass error. In each split, the two peaks are separated by less than the mass of an electron (0.548 mD)!

Positive Ion APPI FT-ICR MS

at 9.4 Tesla

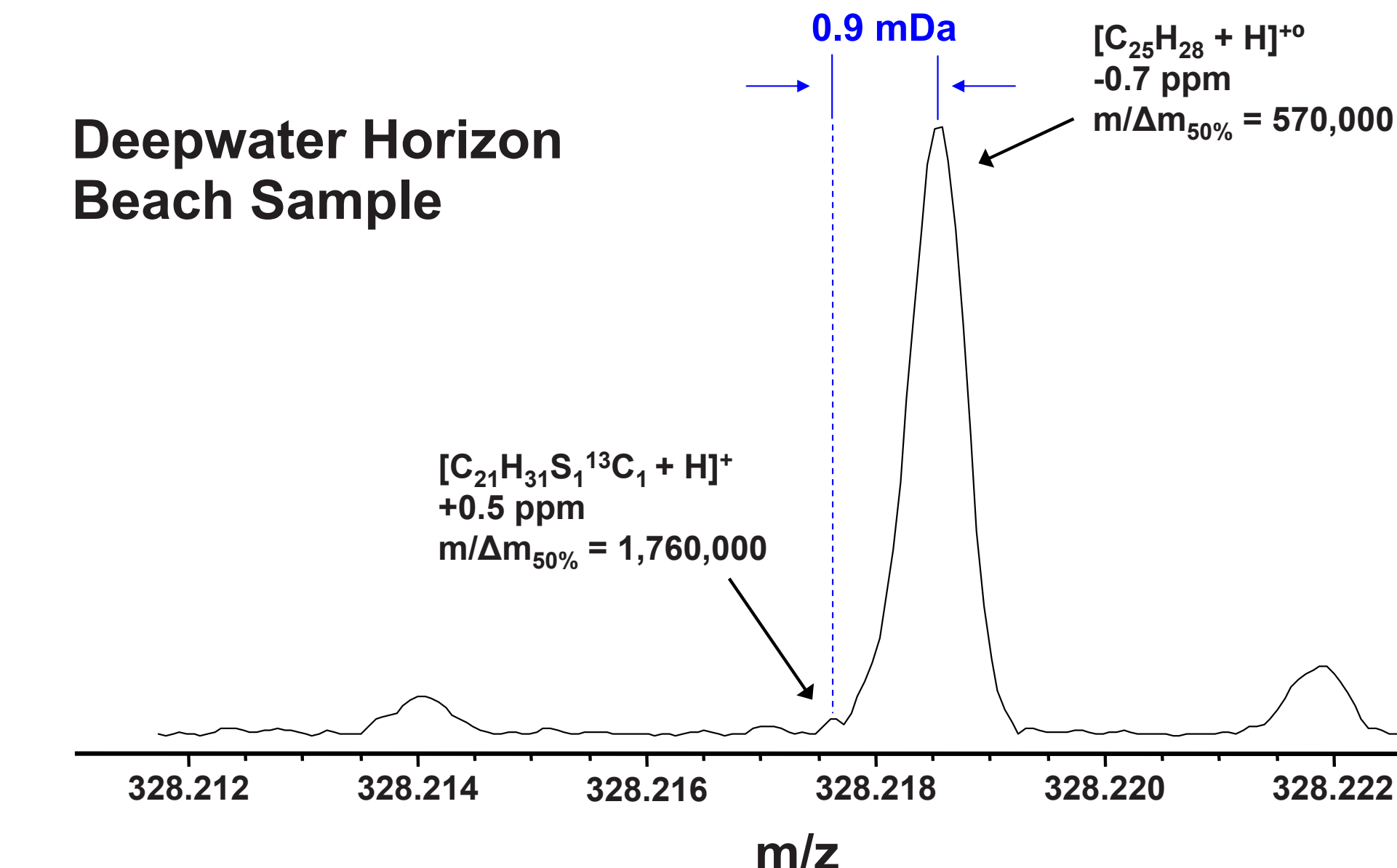


Figure 5. A mass split, annotated as in Figure 4, that may be mistaken by both the user and MIDAS as a split peak. However, these two peaks do indeed correspond to two different molecules, as corroborated by identification of the monoisotopic ion (at ~1 Da lower in mass) corresponding to the lower abundance member of the 328 Da doublet. Thus to account for this issue the program will take the lower-magnitude peak of the pair and label it as unassigned, giving the user the option to reassign or delete that peak as needed.

Conclusion

We present software designed to make data analysis more efficient and sensitive to unassigned peaks and peak splitting. The current version of the software successfully implements these ideas.

References

- (1) Marshall AG, Hendrickson CL, Jackson GS. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer. *Mass Spectrometry Review*, **17**, 1-35, 1998.
- (2) Hughey CA, Hendrickson CL, Rodgers RP, Marshall AG, Qian K. Kendrick Mass Defect Spectrum: A Compact Visual Analysis of Ultrahigh-Resolution Broadband Mass Spectra. *Analytical Chemistry*, **73(19)**, 4676-4681, 2001.