

Cerno Application Note

Extending the Limits of Mass Spectrometry

Spectral Accuracy Analysis of an Ion Series for Structural Elucidation and Elemental Composition Determination

In GC/MS analysis, the rich EI fragments with or without the presence of molecular ion (M^+) are extremely useful for the identification of unknown compounds, especially when the compound has been previously entered into a spectral library such as NIST or Wiley which could then be subsequently searched.¹ If the compound is not contained in a spectral library, accurate mass measurement of the available M^+ and associated fragment ions have been shown to enable the identification of these types of “true unknown” compounds even with a conventional single quadrupole GC/MS system.²

When it comes to ion fragmentation in MS measurements, a wealth of knowledge has been accumulated over the years, as systematically summarized in the authoritative treatise by McLafferty.³ In the context of EI fragments encountered in a typical GC/MS analysis, a table of common fragments have long been established,⁴ also shown in Table 1. The MS Interpreter program that comes with NIST library and search software utilizes such fragment knowledge to complement NIST library search and aid in structural elucidation or understanding.

As shown in Table 1, some common fragments have multiple possibilities, e.g., 28Da may be any one of $-C_2H_4$, $-CO$, and $-N_2$, whereas 29Da may result from $-C_2H_5$ or $-CHO$, etc. While accurate mass measurement may help differentiate these possible fragments, we would like to demonstrate through a real-world example in this application note that these fragments are far better and graphically more intuitively differentiated through spectral differences observed in the profile mode after full spectral calibration, above and beyond just a single point accurate mass measurement alone. Additionally, we will show how different fragments from the same molecular ion (M^+) can all be conceptually and graphically linked together through the concept of spectral accuracy for both structural elucidation and elemental composition determination.

Table 1. Table of Common Fragment Ions

| m/z | Ions | m/z | Ions |
|--------|---|-----|--|
| 15 | CH ₃ | 63 | C ₅ H ₃ ^a |
| 16 | O | 65 | C ₅ H ₅ ^a |
| 17 | OH | 67 | C ₅ H ₇ |
| 18 | H ₂ O | 69 | C ₅ H ₉ |
| 19 | F | 70 | C ₅ H ₁₀ |
| 26 | CN | 71 | C ₅ H ₁₁ , C ₃ H ₇ -C=O |
| 27 | C ₂ H ₃ | 72 | C ₂ H ₅ -CO-CH ₂ +H ^b |
| 28 | C ₂ H ₄ , CO, N ₂ | 73 | C ₃ H ₇ OCH ₂ , C ₂ H ₅ O-C=O C ₃ H ₇ CHOH, C ₂ H ₅ OCHCH ₃ |
| 29 | C ₂ H ₅ , CHO | 74 | CH ₂ -COOCH ₃ +H ^b |
| 30 | CH ₂ NH ₂ , NO | 75 | C ₂ H ₅ O-C=O+2H ^b , C ₂ H ₅ COO+2H ^b |
| 31 | CH ₂ OH, OCH ₃ | 77 | C ₆ H ₅ ^a |
| 32 | O ₂ | 78 | C ₆ H ₅ +H ^{ab} |
| 33 | SH | 79 | C ₆ H ₅ +2H ^{ab} , Br (also 81) |
| 34 | H ₂ S | 81 | C ₆ H ₉ |
| 35, 37 | Cl | 82 | C ₆ H ₁₀ , CCl ₂ (also 84&86) |
| 36,38 | HCl | 83 | C ₆ H ₁₁ , CHCl ₂ (also 85&87) |
| 39 | C ₃ H ₃ ^a | 84 | C ₆ H ₁₂ |
| 40 | CH ₂ CN, Ar | 85 | C ₆ H ₁₃ , C ₄ H ₉ -C=O |
| 41 | C ₃ H ₅ , CH ₂ CN+H ^b | 86 | C ₃ H ₇ -CO-CH ₂ +H ^b |
| 42 | C ₃ H ₆ | 87 | Homologs of 59 & 73 |
| 43 | C ₃ H ₇ , CH ₃ C=O | 88 | CH ₂ -COOC ₂ H ₅ +H ^b |
| 44 | CH ₂ CHO+H ^b , CO ₂ CH ₃ CH-NH ₂ | 89 | C ₃ H ₇ -O-C=O+2H ^b , C ₃ H ₇ COO+2H ^b |
| 45 | CH ₃ CHOH, CH ₂ OCH ₃ , CH ₂ CH ₂ OH, COOH | 90 | C ₆ H ₅ -CH |
| 46 | NO ₂ | 91 | C ₆ H ₅ -CH ₂ , C ₆ H ₅ -CH+H ^b |
| 47 | CH ₂ SH, CH ₃ S | 92 | C ₆ H ₅ -CH ₂ +H ^b |
| 48 | CH ₃ S+H ^b | 93 | C ₇ H ₉ , CH ₂ Br (also 95) |
| 49, 51 | CH ₂ Cl | 94 | C ₆ H ₅ O+H |
| 50 | C ₄ H ₂ ^a | 97 | C ₇ H ₁₃ |
| 51 | C ₄ H ₃ ^a | 99 | Homologs of 71 & 85 |
| 52 | C ₄ H ₄ ^a | 100 | Homolog of 72 & 86 |
| 53 | C ₄ H ₅ | 101 | Homologs of 59, 73, & 87 |
| 54 | CH ₂ CH ₂ CN, CH ₃ CH-CN | 102 | Homolog of 74 & 88 |
| 55 | C ₄ H ₇ | 103 | Homolog of 75 & 89 |
| 56 | C ₄ H ₈ | 105 | C ₆ H ₅ C=O, C ₆ H ₅ -CH ₂ CH ₂ |
| 57 | C ₄ H ₉ , C ₂ H ₅ -C=O | 107 | C ₆ H ₅ -CH ₂ O |
| 58 | CH ₃ -CO-CH ₂ +H ^b | 108 | C ₆ H ₅ -CH ₂ O+H ^b |
| 59 | C ₂ H ₅ OCH ₂ , CH ₃ O-C=O C ₂ H ₅ CHOH, CH ₃ O-CHCH ₃ | 119 | C ₆ H ₅ -C(CH ₃) ₂ |
| 60 | CH ₂ -COOH+H ^b | 127 | I |
| 61 | CH ₃ COO+2H ^b , CH ₃ OCO+2H ^b | | |

^aGood diagnostic for benzene ring compounds

Experimental

Sample information: Calibration tune gas standard PFTBA available from the Agilent GC/MSD instrument and a sample prepared from a natural product (sample).

MS conditions: The PFTBA and the sample were both acquired in profile mode under MassHunter acquisition software at a scan speed 2[^]2 (A/D samples = 4) over a mass range of m/z 29-600 with the ion threshold set to zero. As is typical for GC/MSD, the MS source is heated to 230°C and the quadrupole is heated to 150°C, conditions shown to provide high calibration stability lasting as long as a week without requiring recalibration, based on a previous study.⁵

External PFTBA Calibration and Sample GC/MS Data Acquisition: PFTBA tune gas data were acquired externally through a 5min data run with zero solvent delay and the calibration valve turned on at 2.5min and off at 4.5min, as shown in Figure 1. The GC/MS sample run was acquired with a 3min solvent delay, 30.25min total run time, and two-ramp temperature programming, as shown in Figure 2. Detailed data acquisition setup information can be found elsewhere.⁶

Data acquisition and analysis: Opening the PFTBA run data directly inside the MassWorks™ V6.0 software (Cerno Bioscience, Las Vegas, NV) will automatically trigger a MassWorks calibration by using all available PFTBA-related fragments as standard ions, in a process called AutoCal™.⁷ The calibration file thus created could then be associated with and applied to the GC/MS sample run to achieve both high mass accuracy and spectral accuracy for unknown compounds to be identified. Spectral Accuracy can be computed from the difference between the calibrated and theoretically calculated MS,⁸ either for a given candidate ion (pure)⁸ or a combination of candidate ions (mixtures or ion series) including those with mass spectral overlaps, e.g., from isotopically labelled compounds.⁹ This process of computing Spectral Accuracy is called CLIPS (Calibrated Lineshape Isotope Profile Search).¹⁰

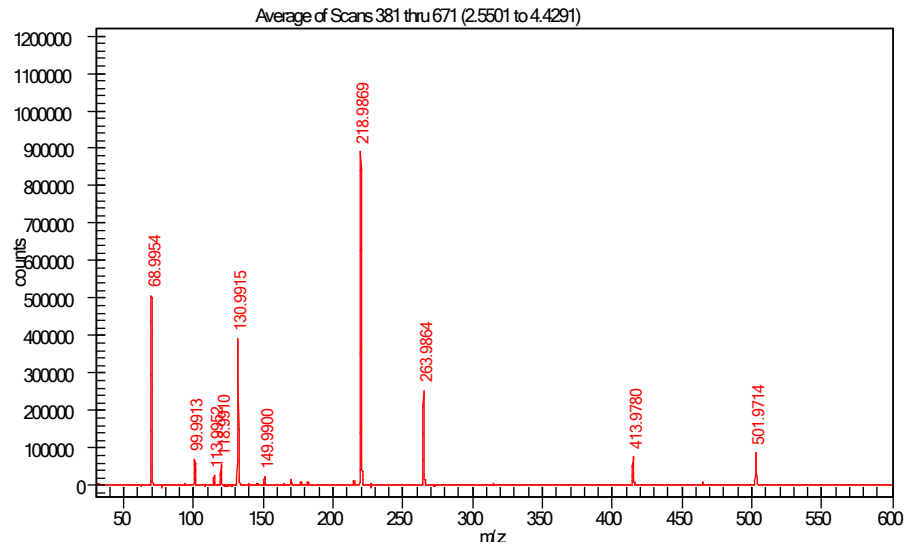
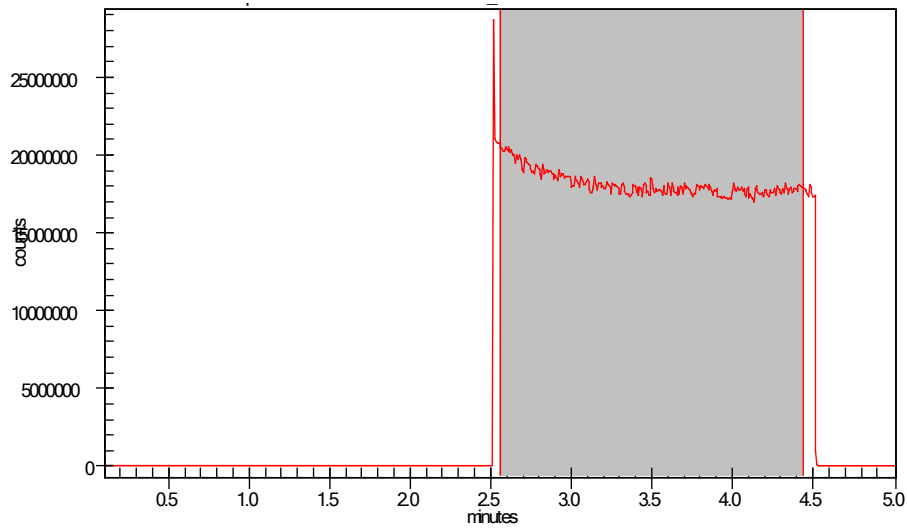


Figure 1. External PFTBA tune gas data acquired for MassWorks AutoCal

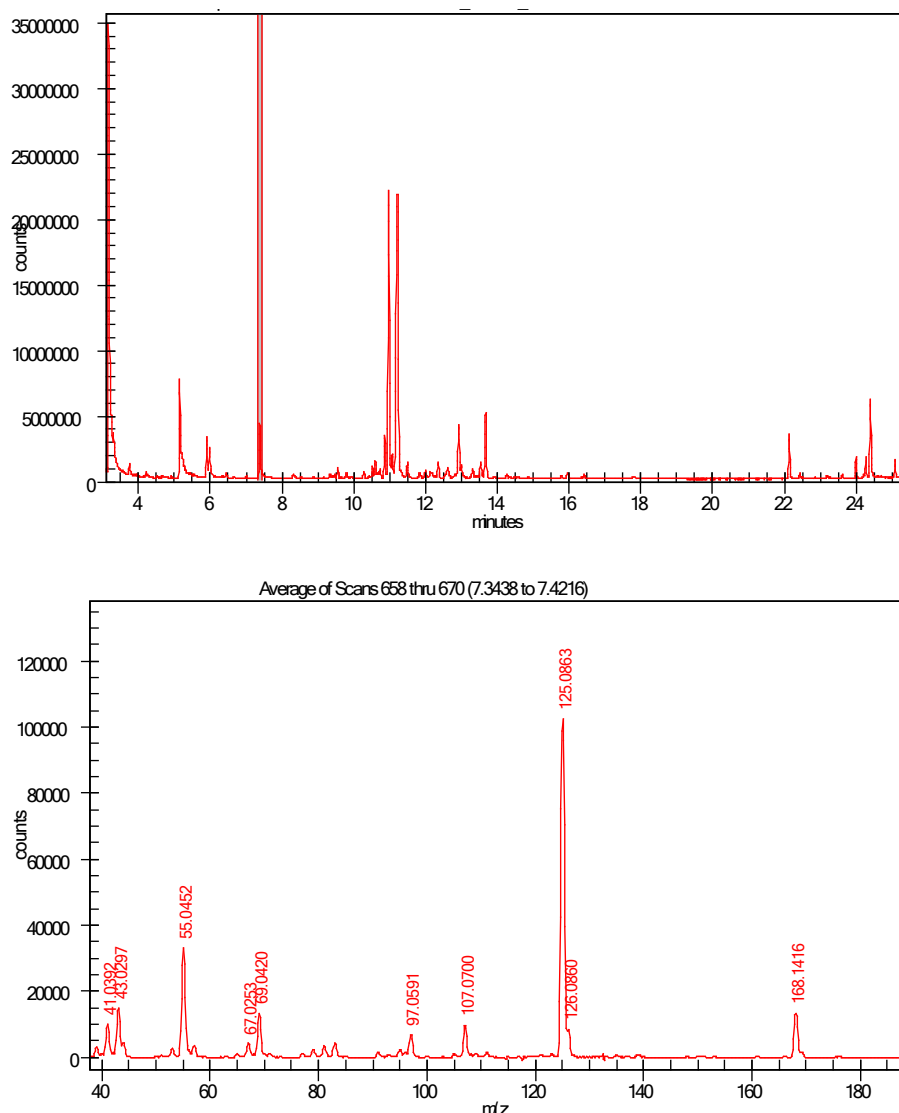


Figure 2. GC/MS analysis of a natural product extract with MassWorks calibration applied

Results

Identification of Possible Molecular (M^+) Ion

For the unknown compound at $R_t=7.4$ Min shown in Figure 2, the ion at m/z 168 with its accurate mass determined as 168.1416Da may possibly be the molecular ion (M^+), which would be used as a working hypothesis for this unknown identification study. A CLIPS search with the generous parameters shown in Figure 3 results in a short list of possible elemental compositions shown in Figure 4. The top ($C_{11}H_{20}O^+$) and the 2nd hit ($C_{10}H_{20}N_2^+$) both have more than 99.0% spectral accuracy with nearly a perfect spectral match between the calibrated and the theoretical MS. The 3rd hit ($C_9H_{16}N_2O^+$) with only 98.5% Spectral Accuracy shows a discernable spectral error around $M+1$ isotopes, more specifically an under-fitting to the calibrated MS around $M+1$, with smaller $M+1$ signal from two less carbons in the corresponding elemental composition. Due to the decrease in one carbon from 1st to 2nd

hit (from 11 carbons in $C_{11}H_{20}O^+$ to 10 carbons in $C_{10}H_{20}N_2^+$), there should have been a discernable decrease in M+1 signal. It turns out that the two nitrogen atoms and the associated increase in M+1 signal from ^{15}N somewhat compensated for the decrease from one less ^{13}C , while the presence of an oxygen in the 1st hit does not contribute to the M+1 signal due to the lack of naturally occurring ^{17}O isotope.

Accurate mass to search: Parameter Set: Actions...

Charge: Analysis Mode: Sequence...

Formula Generation by Mass

Mass Tolerance

mDa

PPM

Electron State

Odd Even Both

Double Bond Equivalent Range

Min: Max:

Empirical Rules

| | Element | Min | Max |
|---|---------|-----|-----|
| 1 | C | 1 | 13 |
| 2 | H | 0 | 26 |
| 3 | N | 0 | 8 |
| 4 | O | 0 | 8 |
| 5 | | | |

Formula Determination / Mixture Analysis by Spectral Accuracy

Profile Mass Range (Da)

Start: End:

Interference Rejection:

Ion Series

| | Repeat Unit | Min | Max |
|---|-------------|-----|-----|
| 1 | | | |

Results

Show All Results Show Top Results

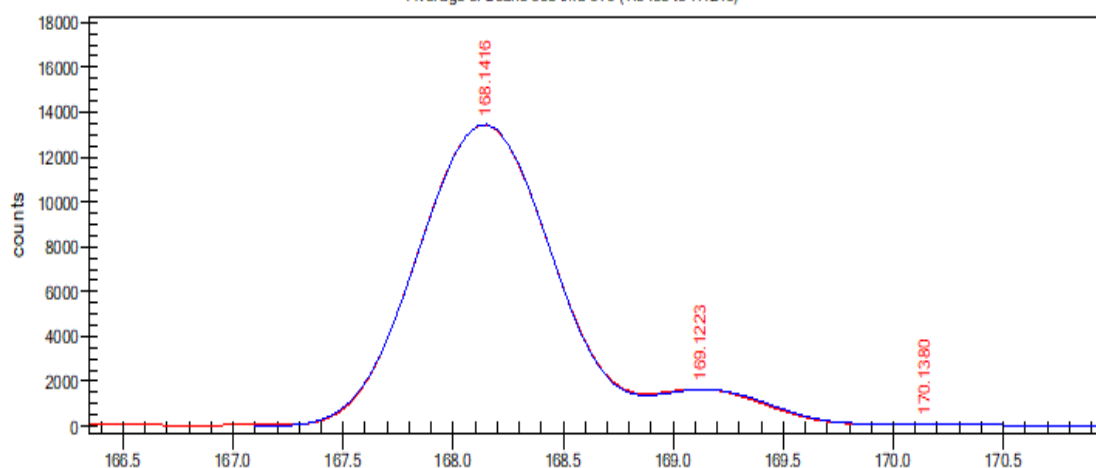
Sort Mass Error By Absolute Value

Show this dialog before each search

Figure 3. CLIPS elemental composition search parameters

| CLIPS Results for average of scans 658 thru 670 | | | | | | | |
|---|--|--------------|------------------|------------------|-------------------|------|------|
| | Formula | Mono Isotope | Mass Error (mDa) | Mass Error (PPM) | Spectral Accuracy | RMSE | DBE |
| 1 | C ₁₁ H ₂₀ O | 168.1509 | -9.2667 | -55.1124 | 99.1179 | 40 | 2.0 |
| 2 | C ₁₀ H ₂₀ N ₂ | 168.1621 | -20.5001 | -121.9215 | 99.0813 | 42 | 2.0 |
| 3 | C ₉ H ₁₆ N ₂ O | 168.1257 | 15.8854 | 94.4765 | 98.4779 | 69 | 3.0 |
| 4 | C ₈ H ₁₆ N ₄ | 168.1369 | 4.6520 | 27.6674 | 98.2393 | 80 | 3.0 |
| 5 | C ₄ H ₁₆ N ₄ O ₃ | 168.1217 | 19.9082 | 118.4013 | 94.6733 | 241 | -1.0 |
| 6 | C ₃ H ₁₆ N ₆ O ₂ | 168.1329 | 8.6748 | 51.5922 | 94.4073 | 253 | -1.0 |

Average of Scans 658 thru 670 (7.3438 to 7.4216)



Average of Scans 658 thru 670 (7.3438 to 7.4216)

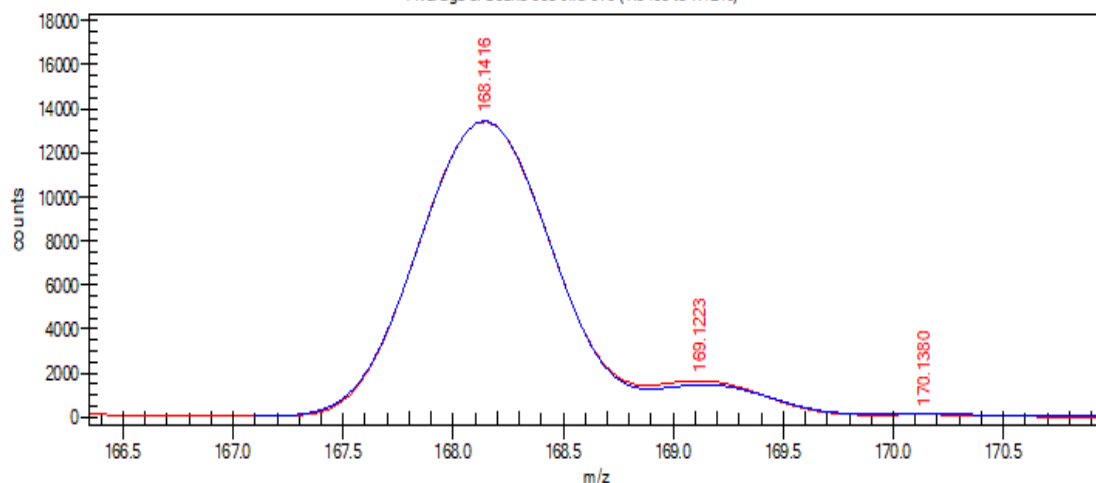


Figure 4. CLIPS search results sorted by Spectral Accuracy with Spectral Overlay (red - calibrated vs blue - theoretical MS) for the top (middle) and 3rd hit (bottom)

Identification of the 43Da Fragment Loss

The above Spectral Accuracy analysis of the M^+ ion including the various higher isotopes narrows down the number of possible elemental compositions to the top 2 candidates. The presence of other EI fragment ions including the most abundant m/z 125 ion will now be utilized to help with both the elemental composition determination and shed light onto its possible structures. Taking advantage of all available EI fragments and the NIST spectral library, one could easily perform the library search with results shown in Figure 5. While the top NIST search hit does have the same elemental composition as the 1st hit from CLIPS search in Figure 4, the search score of less than 800 and the spectral mismatch on key fragments at m/z 125, 69, 55, 41, and 27 indicate that this compound is a possible isomer not contained in the NIST library. Even though the NIST search is inclusive in this case, it is interesting to note that all top 10 hits from the NIST search include at least one oxygen and higher abundance of the m/z 125 fragment ion is associated with either more than one oxygen in the elemental composition or with a single oxygen contained in an OH group in the chemical structure, an important clue on the possible range of structures.

Considering the purported M^+ of m/z 168 and the most abundant fragment at m/z 125 with the mass difference of 43Da and referring to the common fragments listed in Table 1, there seems to be two possibilities for the 43Da fragment, CH_3CO (containing oxygen) or $CH_3CH_2CH_2$ (without oxygen). The ion series analysis available within CLIPS search can be utilized to test and compare these two distinct possibilities. For the unknown molecule to lose a CH_3CO fragment, it would have to contain at least one oxygen to start with, which leaves only the 1st CLIPS hit, $C_{11}H_{20}O^+$, as the possibility. By analyzing both the M^+ ion at m/z 168 and the fragment ion at m/z 125 in one overall spectral fit in the wide spectral profile range as shown in Figure 6, one can test and evaluate the overall Spectral Accuracy both quantitatively and graphically. At the overall Spectral Accuracy of only 97.0% with a significant and systematic spectral shift across all isotopes of the M^+ , the loss of fragment CH_3CO seems quite unlikely. Furthermore, a discernable overfit on the A+1 of the m/z 125 fragment ion likely points to an extra carbon in theoretical MS, indicating that the other 43Da fragment, $CH_3CH_2CH_2$, may hold more promise. With $CH_3CH_2CH_2$ as the 43Da fragment loss and 98.1% Spectral Accuracy as shown in Figure 7, there is neither the spectral shift for M^+ at m/z 168 nor discernable overfit (or underfit) on the A+1 of the m/z 125 fragment ion, clearly indicating $CH_3CH_2CH_2$ is the correct 43Da fragment loss.

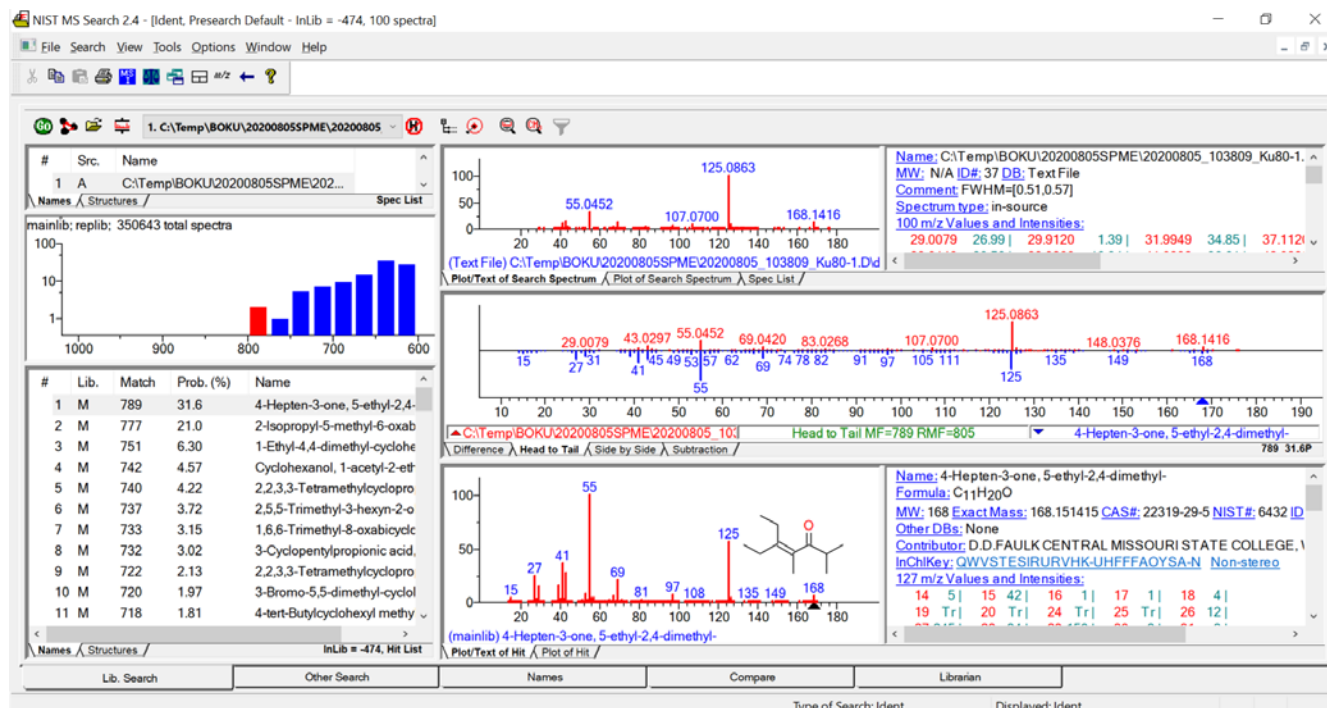


Figure 5. NIST library search of the Rt 7.4 min peak indicates the unknown compound does not exist in the library with top search score of 789

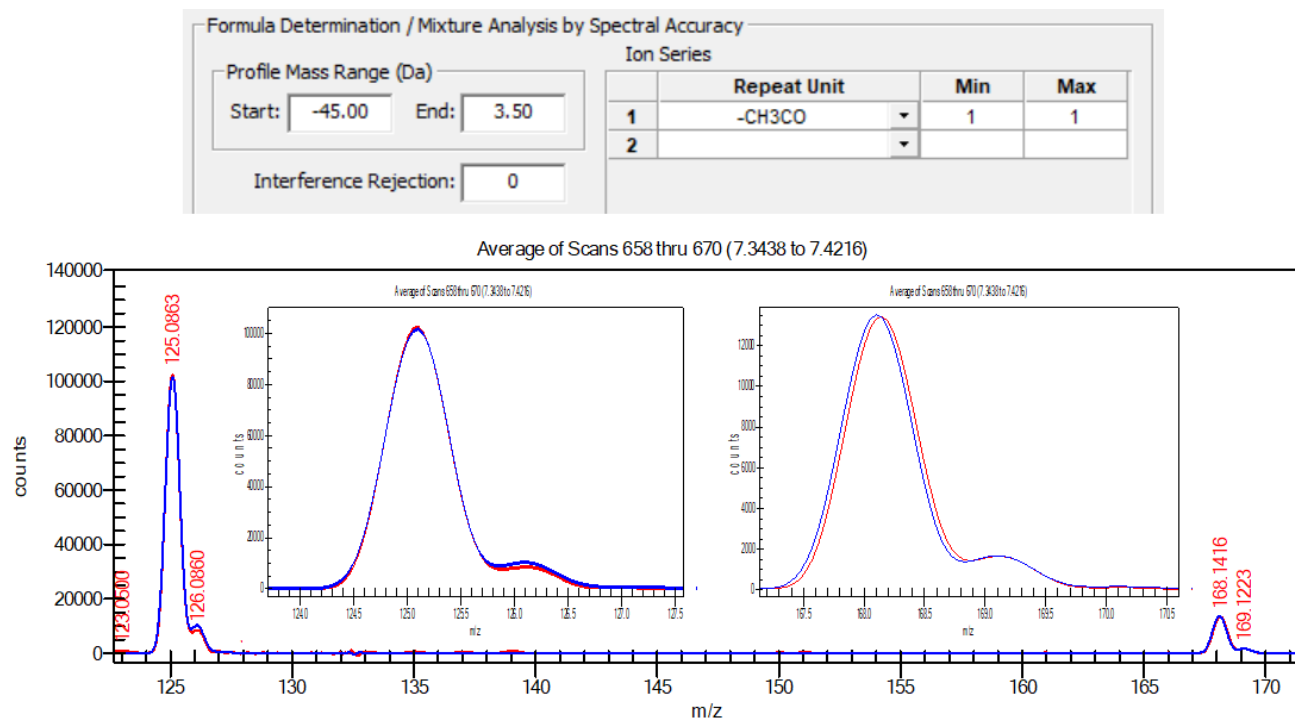


Figure 6. Spectral overlay between calibrated (red) and theoretical (blue) MS with CH₃CO fragment loss and 97.0% Spectral Accuracy

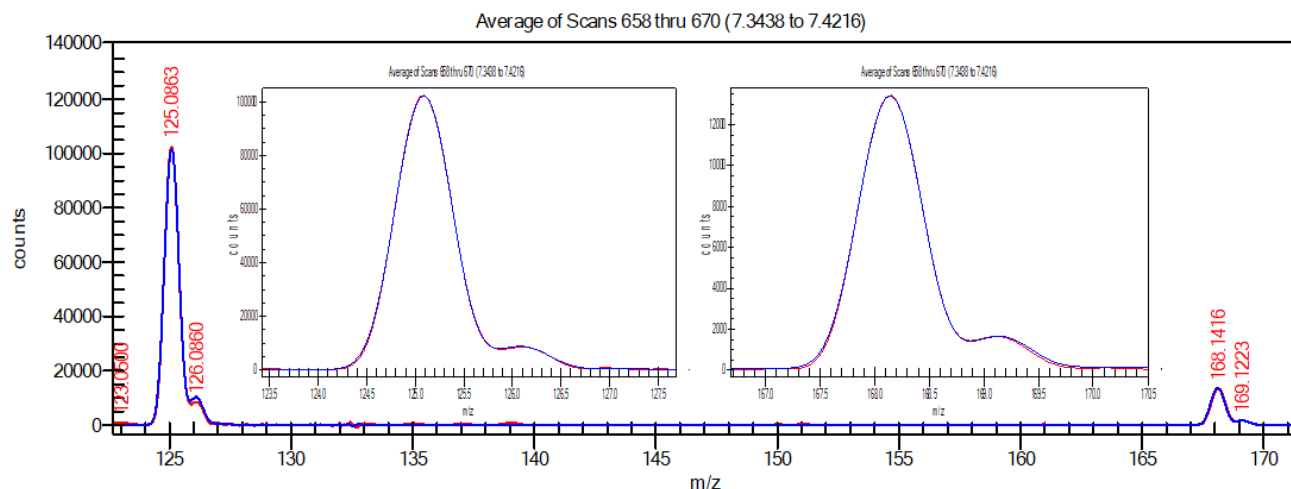
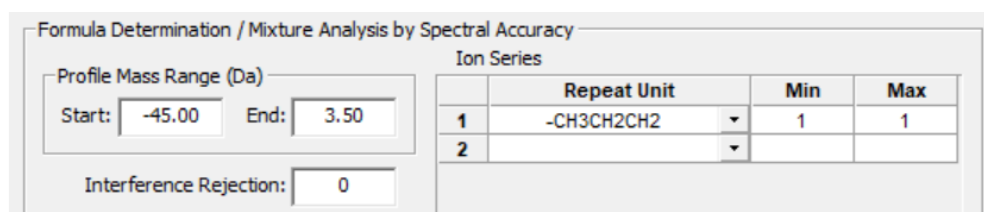
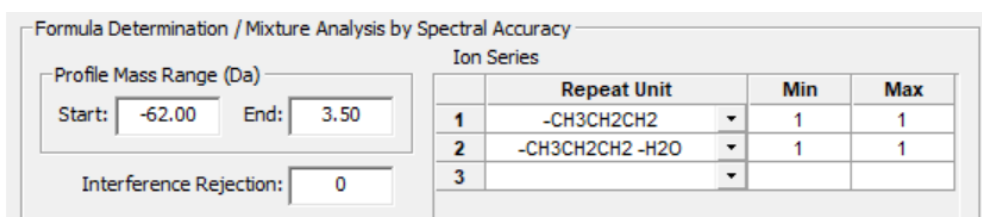


Figure 7. Spectral overlay between calibrated (red) and theoretical (blue) MS with CH₃CH₂CH₂ fragment loss and 98.1% Spectral Accuracy

The Oxygen Question

The confirmation of CH₃CH₂CH₂ as the 43Da fragment loss leaves both the 1st and the 2nd CLIPS hit, C₁₁H₂₀O⁺ and C₁₀H₂₀N₂⁺, in the play as possible elemental compositions. The presence of another smaller fragment at m/z 107Da could potentially help solve the puzzle. This fragment is 18Da away from the most abundant m/z 125 fragment and could be a further fragmentation product. According to the common fragments listed in Table 1, the loss of 18Da has one clear candidate, H₂O. Figure 8 shows the overall spectral fitting with both CH₃CH₂CH₂ and the additional H₂O fragment losses included in the wider spectral profile range covering m/z 106 to 172 with an overall Spectral Accuracy of 97.2%. This somewhat compromised Spectral Accuracy is due mostly to the unmodelled spectral interferences in between the three main ions, e.g., at m/z 109, 111, and 132. The less abundant ions at m/z 109 and 111 may possibly be due to the loss of O and CH₂, respectively, a plausible hypothesis tested as shown in Figure 9, which improves the Spectral Accuracy to 97.6%. This points to the 1st CLIPS hit, C₁₁H₂₀O⁺, as the only viable elemental composition with one oxygen to be lost to form the additional ion fragment observed.



Average of Scans 658 thru 670 (7.3438 to 7.4216)

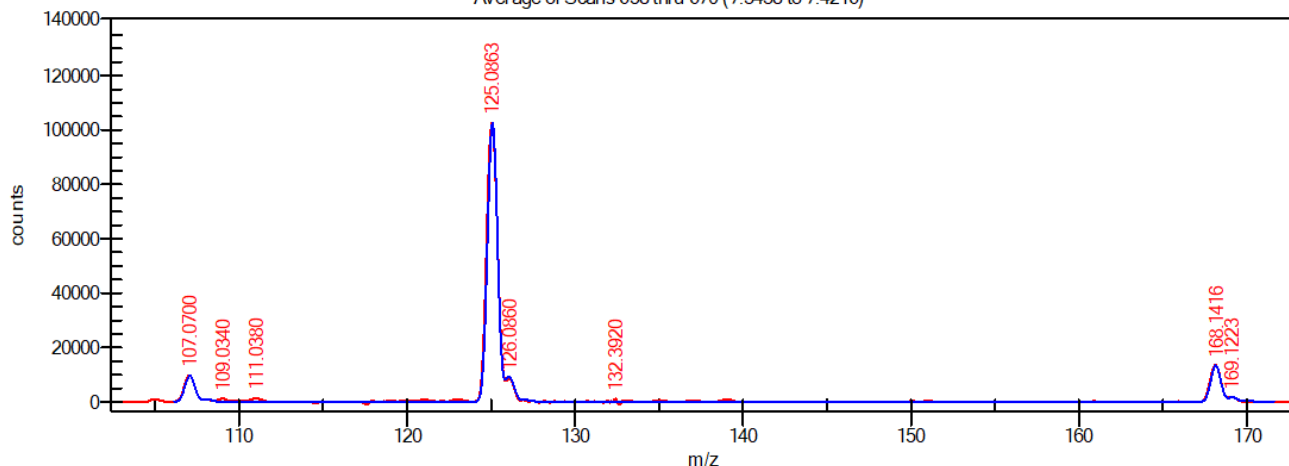
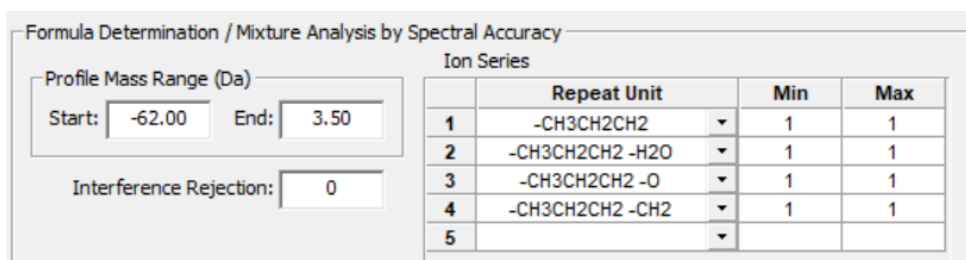


Figure 8. Spectral overlay between calibrated (red) and theoretical (blue) MS with both $\text{CH}_3\text{CH}_2\text{CH}_2$ and an additional H_2O fragment losses (97.2% Spectral Accuracy)



Average of Scans 658 thru 670 (7.3438 to 7.4216)

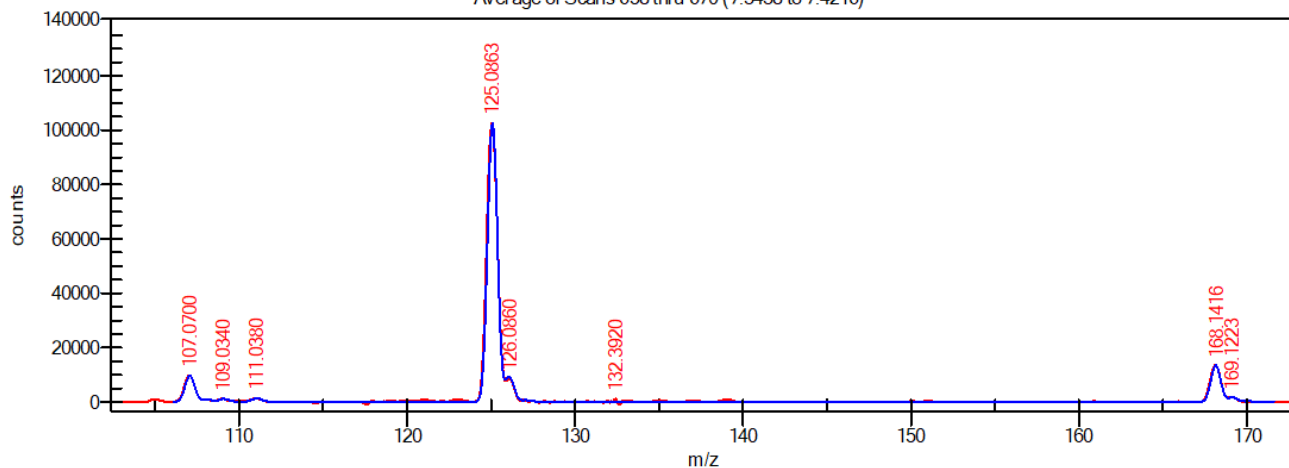


Figure 9. Spectral overlay between calibrated (red) and theoretical (blue) MS with $\text{CH}_3\text{CH}_2\text{CH}_2$, an additional H_2O , and two other fragment losses (97.6% Spectral Accuracy)

Other Smaller Fragments

Using the same methodology, one could analyze other smaller fragments or groups of fragments with the CLIPS ion series feature to identify their respective elemental compositions. Figure 10 shows the results of these additional analysis with the C3 through C8 series of hydrocarbon fragments as indicated. These smaller fragments combined with the larger fragments analyzed previously should help piece together the possible structures of the unknown compound.

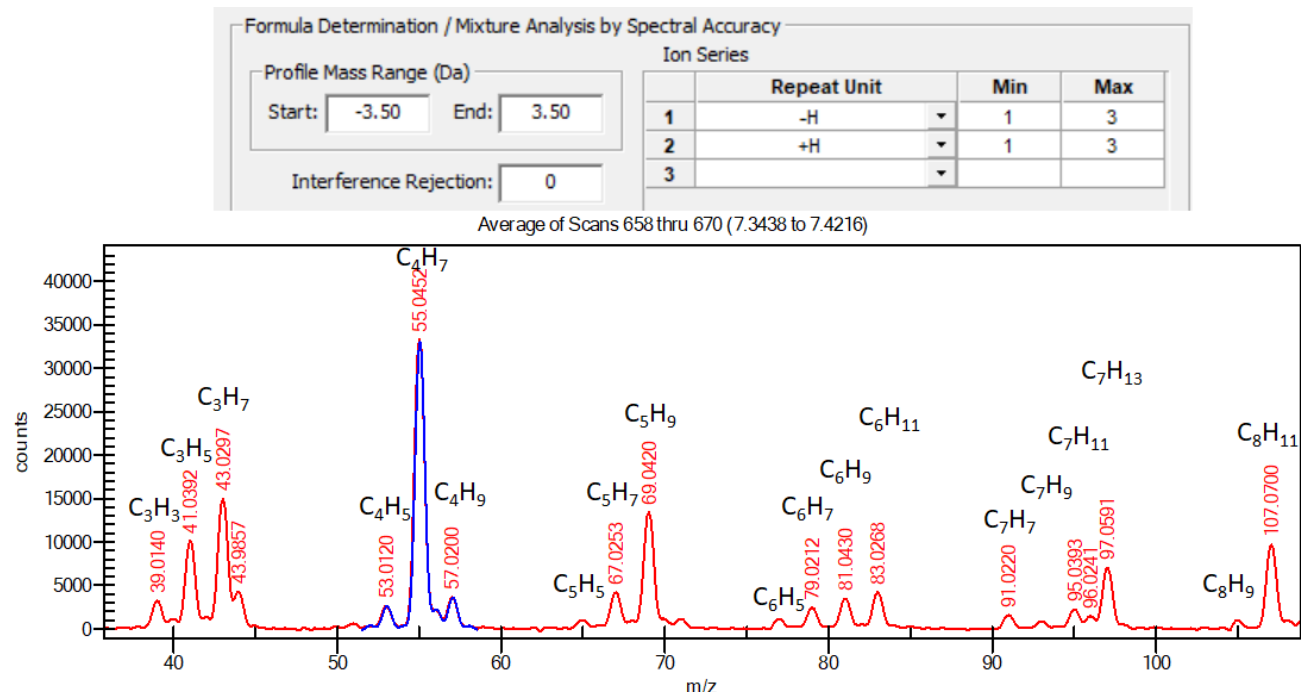
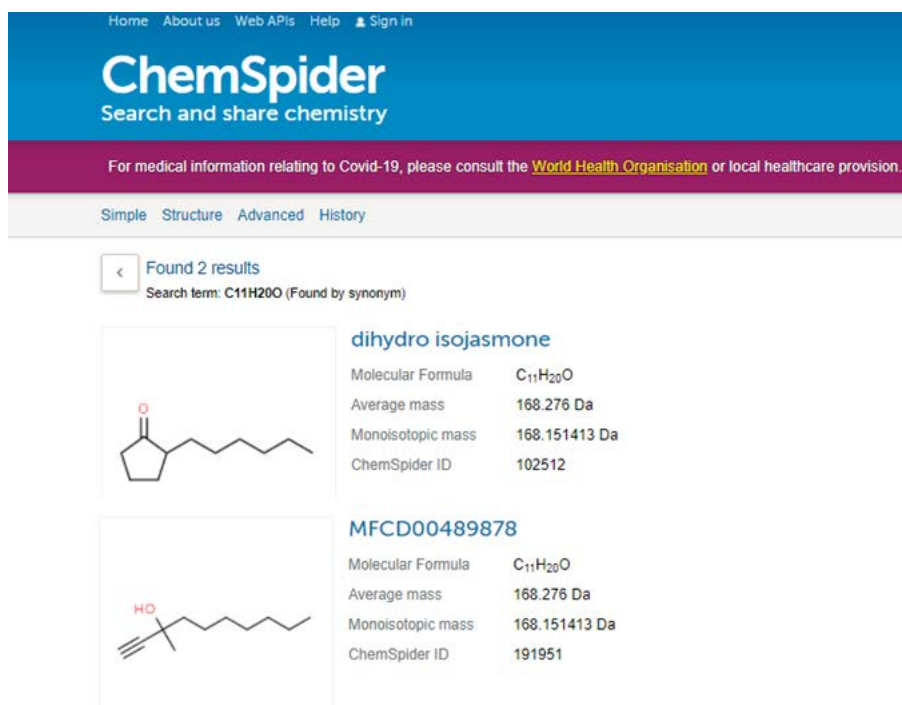


Figure 10. Spectral Accuracy of 98.2% is obtained when CLIPS search is applied to ion series around m/z 55 to reveal possible fragments containing hydrocarbons with other ions series similarly analyzed and fragments proposed

Web Search and Possible Structures

Considering the clues obtained from the unsuccessful NIST search, the CLIPS ion series analysis of both the larger and smaller fragments, and the Web search of possible structures using, e.g., ChemSpider (Figure 11), the possible structure of the compound is likely to be something very similar to the 2nd hit as opposed to the 1st hit from ChemSpider. Similar to the 2nd ChemSpider hit, the correct molecule likely contains a single oxygen through an OH group attached to a carbon, located either in the middle or towards the end of the hydrocarbon chain, with the rest of the hydrocarbon chain containing up to 8 consecutive carbons (largest fragment around m/z 107, shown in Figure 10).



Home About us Web APIs Help Sign in

ChemSpider

Search and share chemistry

For medical information relating to Covid-19, please consult the [World Health Organisation](#) or local healthcare provision.

Simple Structure Advanced History

Found 2 results
Search term: C11H20O (Found by synonym)

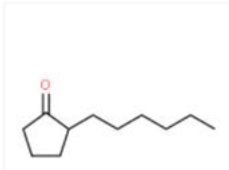
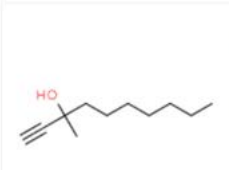
| Chemical Structure | Name | Molecular Formula | Average mass | Monoisotopic mass | ChemSpider ID |
|---|--------------------|-----------------------------------|--------------|-------------------|---------------|
|  | dihydro isojasmone | C ₁₁ H ₂₀ O | 168.276 Da | 168.151413 Da | 102512 |
|  | MFCD00489878 | C ₁₁ H ₂₀ O | 168.276 Da | 168.151413 Da | 191951 |

Figure 11. ChemSpider Web search for possible structures

Conclusions

While accurate mass measurement is quite important for compound identification, it has been demonstrated time and again that the concept of Spectral Accuracy is even more important to not only help narrow down the number of possible elemental compositions but also to help understand and to propose possible structures. When one or more fragments are available, an overall spectral fit in a much wider mass spectral profile range is feasible by including all key fragments into the analysis and thus linking them all together to test and compare various hypothesis. Spectral Accuracy analysis is not only quantitative but also spectrally interpretative, even to the individual isotope level such as ¹³C and ¹⁵N or ¹⁷O (or lack thereof) to help judge which hypothesis is most likely supported by the mass spectral data and help draw proper conclusions about the identity of the compound and its possible structures. Even when a compound is not found in a spectral library, the information from the inconclusive search in the NIST library, when combined with the CLIPS ion series analysis and other Web search such as ChemSpider, could help piece together the possible types of structures for an unknown compound for further study or consideration.

Acknowledgement

We gratefully acknowledge Dr. Stefan Böhmendorfer, Institute of Chemistry of Renewable Resources, Department of Chemistry, University of Natural Resources and Life Sciences (BOKU), Tulln, Austria, for providing the GC/MS data used in this study.

References

- ¹ “Optimization and Testing of Mass Spectral Library Search Algorithms for Compound Identification”, Stein, S. E., J. Am. Soc. Mass Spectrom. 1994, 5, 859–865.
- ² “GC/MSD Formula Determination for Forensics Chemistry”, Wang, Y., Prest, H., Cerno Application Note 106, October, 2009.
- ³ “Interpretation of Mass Spectra”, McLafferty, F. W., Turecek, F., 4th Edition, Print ISBN: 9780935702255 | Online ISBN: 9781938787270, University Science Books, 1993.
- ⁴ “Introduction to Mass Spectrometry: Instrumentation, Applications and Strategies for Data Interpretation”, 4th Edition, Watson, J. T., Sparkman, O. D., Print ISBN: 9780470516348 | Online ISBN: 9780470516898 | DOI:10.1002/9780470516898, John Wiley & Sons, 2007.
- ⁵ “The Robustness of Formula Determination on a Single Quadrupole GC/MS”, J. Mullis, F. Qiu, Y. Wang, Poster, ASMS 2008.
- ⁶ “Cerno Bioscience MassWorks: Acquiring Calibration Data on Agilent GC/MSDs”, Prest, H., Wang, Y., Agilent Application Note 5990-4966EN, November, 2009.
- ⁷ “AutoCal™: Fully Automated Calibration for Accurate Mass and High Spectral Accuracy”, Wang, Y., Cerno Application Note 109, November, 2013.
- ⁸ “The Concept of Spectral Accuracy for MS”, Wang, Y., Gu, M. Anal. Chem. 2010, 82, 7055-7062.
- ⁹ “Direct Quantitation of Complex Mixtures of Isotope Labeled Compounds”, Kuehl, D., Cerno Application Note 113, March, 2017.
- ¹⁰ “Formula Determination for Contaminants, Impurities, and Residuals through Conventional Mass Spectrometry on Real Chromatographic Time Scale”, Wang, Y. and Gu, M. 2007 AOAC INTERNATIONAL Annual Meeting & Exposition.