The Concept of Effective Mass Accuracy (EMA) for Unknown Identification

Yongdong Wang & Don Kuehl

Cerno Bioscience, Las Vegas, NV, USA

Sunday Poster: 6-2-6P

The Holy Grail of Unique Elemental Composition Determination: 50, 5, 1 or 0.1ppm Mass Accuracy?

Mass accuracy has been well established in mass spectrometry (MS) and its importance widely recognized for unknown identification or known compound confirmation. Even with the typically cited mass accuracy of 1-5ppm attainable on most HiRes MS systems, one rarely gets a unique answer, especially for moderately larger compound with m/z>400. While it is theoretically possible to get more confident results through higher mass accuracy, one quickly reaches a point of diminishing economic or practical return at ~1ppm mass accuracy.

MS measurement of a compound typically includes not only its monoisotope but also other relevant higher isotopes and their relative distributions. While the power of combining the mass accuracy and isotope distribution in compound identification has been reported in literature for quite some time (Ref. 1), there does not exist a metric to quantitatively measure how much more powerful such a combination could be. In this presentation, we seek to introduce and demonstrate the concept of Effective Mass Accuracy (EMA) and its power vs mass accuracy by itself, through both computer simulation and experimental data from both high resolution (TOF or Orbitrap) and conventional single or triple quadrupole MS systems.

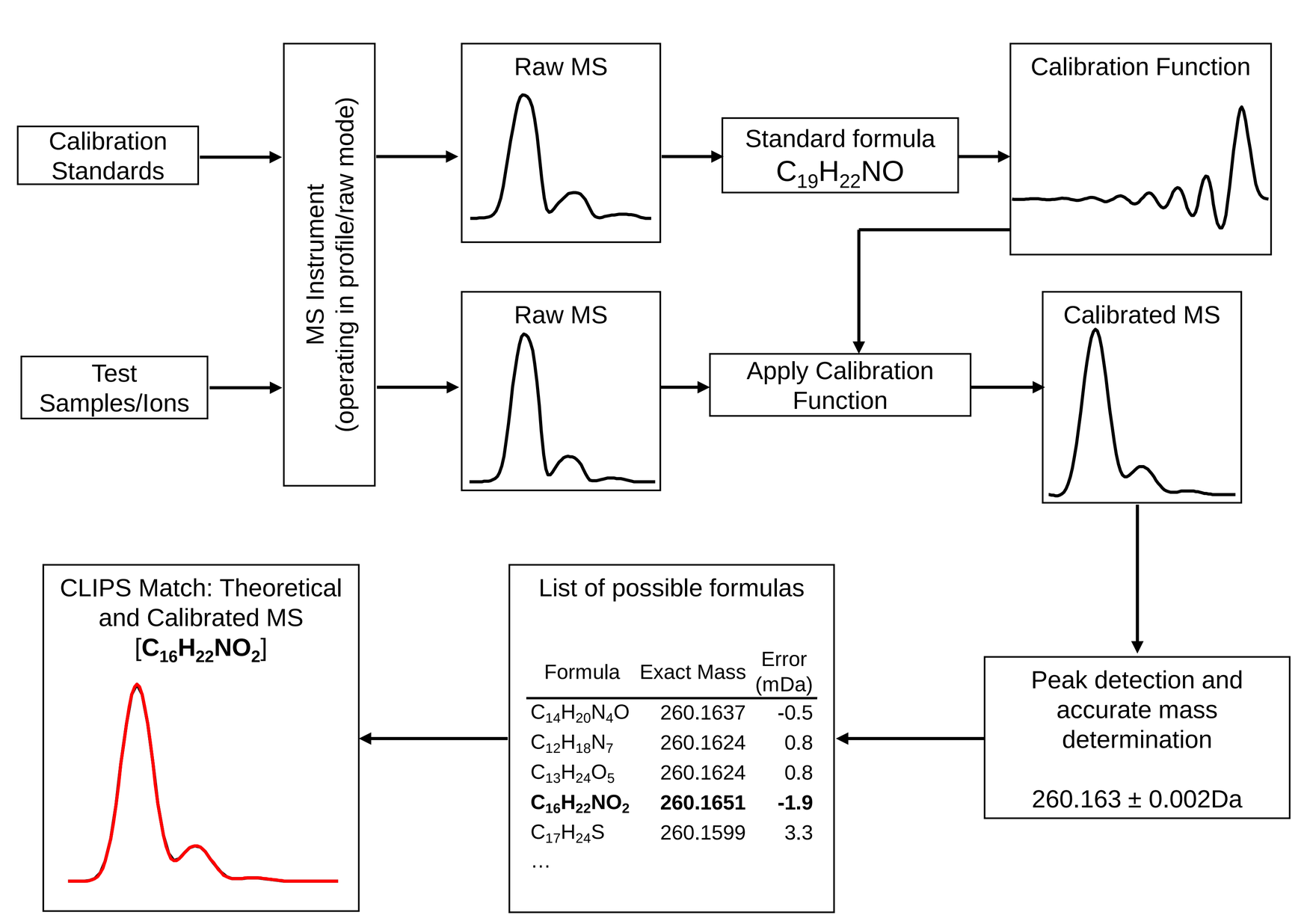
MS TrueCal™: 100x Better Mass Accuracy and High Spectral Accuracy

As outlined in a front cover feature article (Ref 2), acquiring profile mode (raw scan vs centroid) MS data is key to preserving all critical information about elemental compositions and any measured profile mode MS data are composed of two parts:

a. Discrete isotope distributions arising from elemental composition(s)
b. MS instrument- and tune-specific peak shape

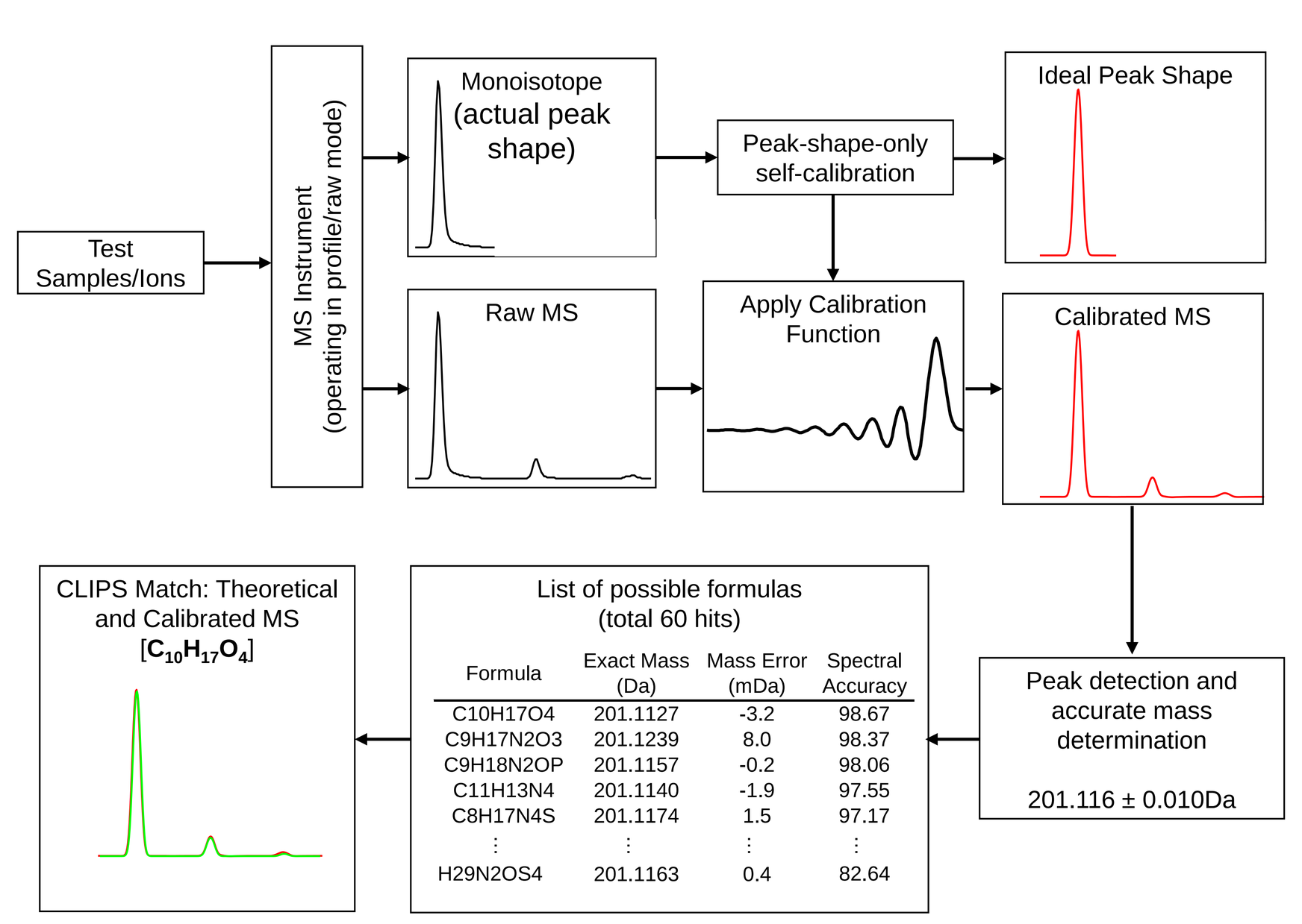
In order to achieve truly accurate MS analysis, it is critically important to perform a new type of MS calibration by involving both the mass position and the MS peak shape (TrueCal). This could be achieved on lower resolution MS system via the use of known standards or on higher resolution system via the use of monoisotopic peak of the test compound itself.

LowRes TrueCal + CLIPS Formula ID



By calibrating both m/z position and MS peak shape through either internal or external standards of known elemental compositions, it is possible to achieve 100x better mass accuracy on a single or triple quad MS system, from 0.x to 0.00x mass accuracy. When combined with the unparalleled spectra accuracy attained during the same calibration process, it becomes practically feasible to determine elemental compositions of unknown compounds under typical GC/MS or LC/MS operating conditions via CLIPS (Calibrated Line-shape Isotope Profile Search).

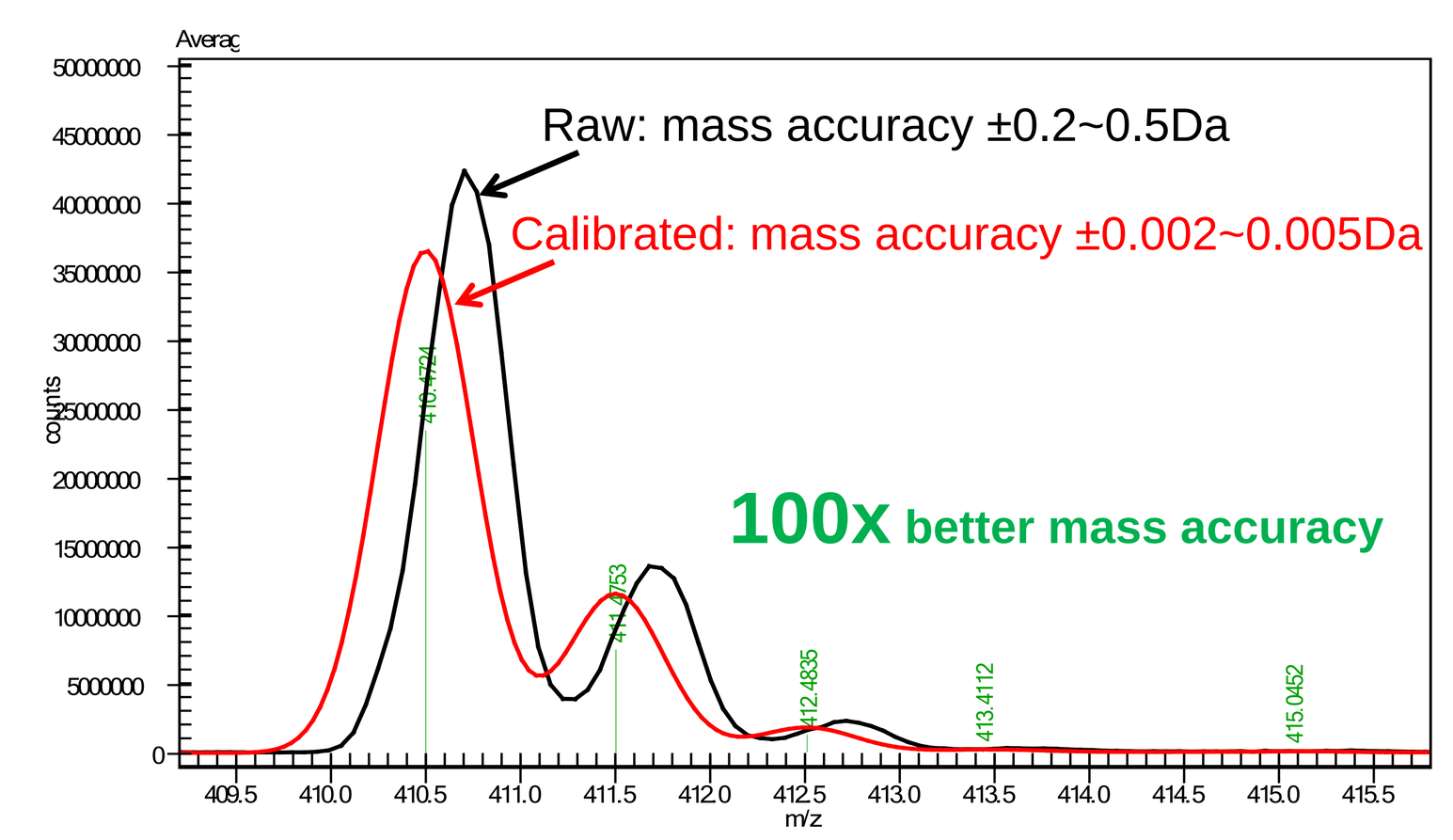
HiRes TrueCal + sCLIPS Formula ID

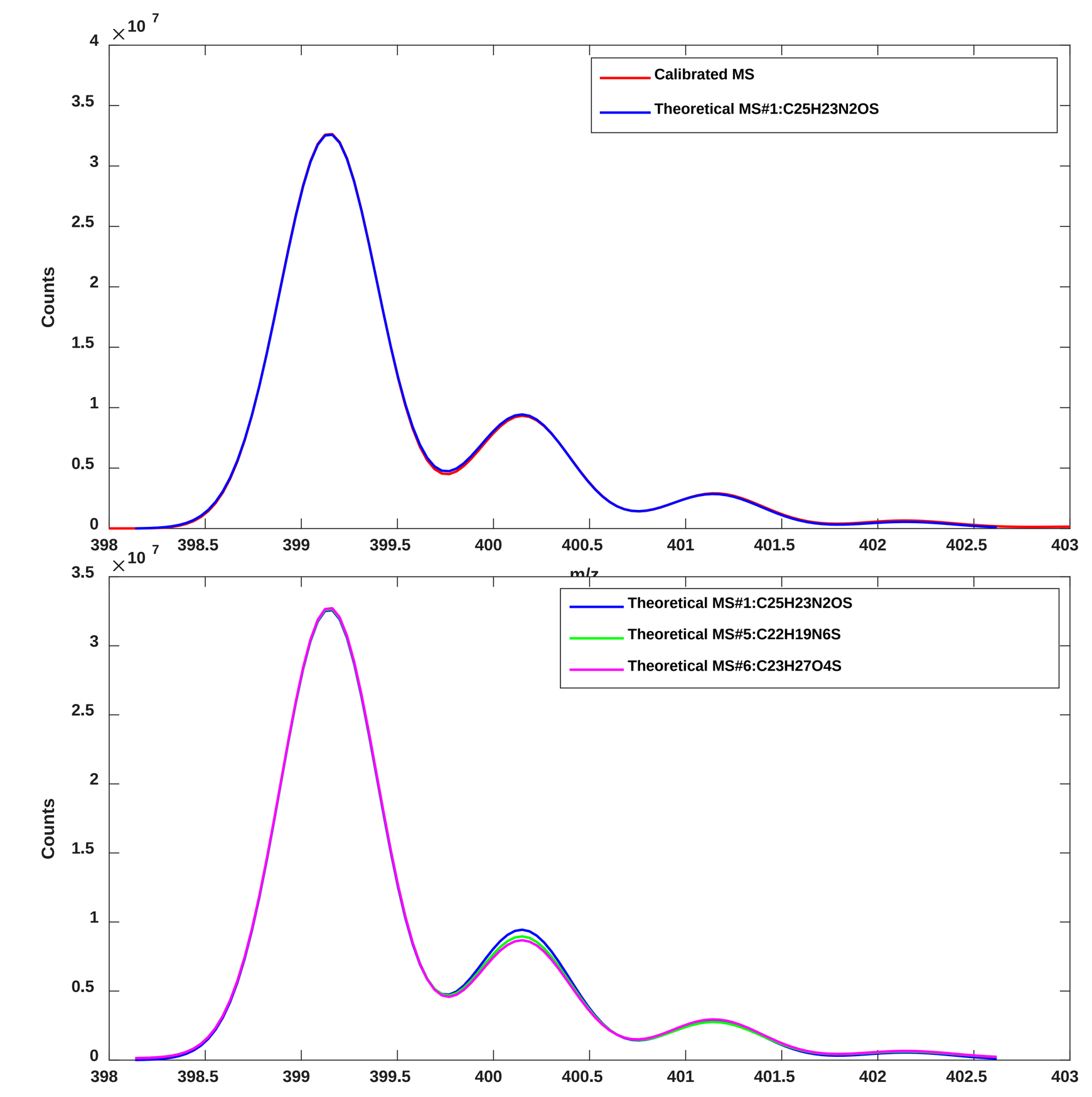


With HiRes TOF or Orbitrap MS data, it is possible to perform a peak-shape-only calibration to transform the actual measured peak shape into a known/perfect mathematical function by using the measured monoisotope peak of the unknown compound itself as the standard. When applied to the whole isotope cluster, it is transformed into a calibrated isotope profile trace conforming to the same known/perfect peak shape, which is then used to calculate the theoretical mass spectrum for any given formula candidate for exact isotope modeling with high spectral accuracy (sCLIPS, self-Calibrated Line-shape Isotope Profile Search).

Results and Discussion

Unit Mass Resolution Single or Triple Quadrupole MS





Accurate Mass 399.1448
Charge 1
Mass Tolerance (mDa) 20
Electron State Even
DBE [Min, Max] [-1, 50]
MS Profile Range (Da) [-1, 3.5]

Empirical Rules Enabled
Empirical Limits Wiley
H/C Ratio Extended
Heteroatom Ratios Extended

Element [Min, Max]
C [1, 32]
H [0, 58]
N [0, 17]
O [0, 17]
S [0, 9]

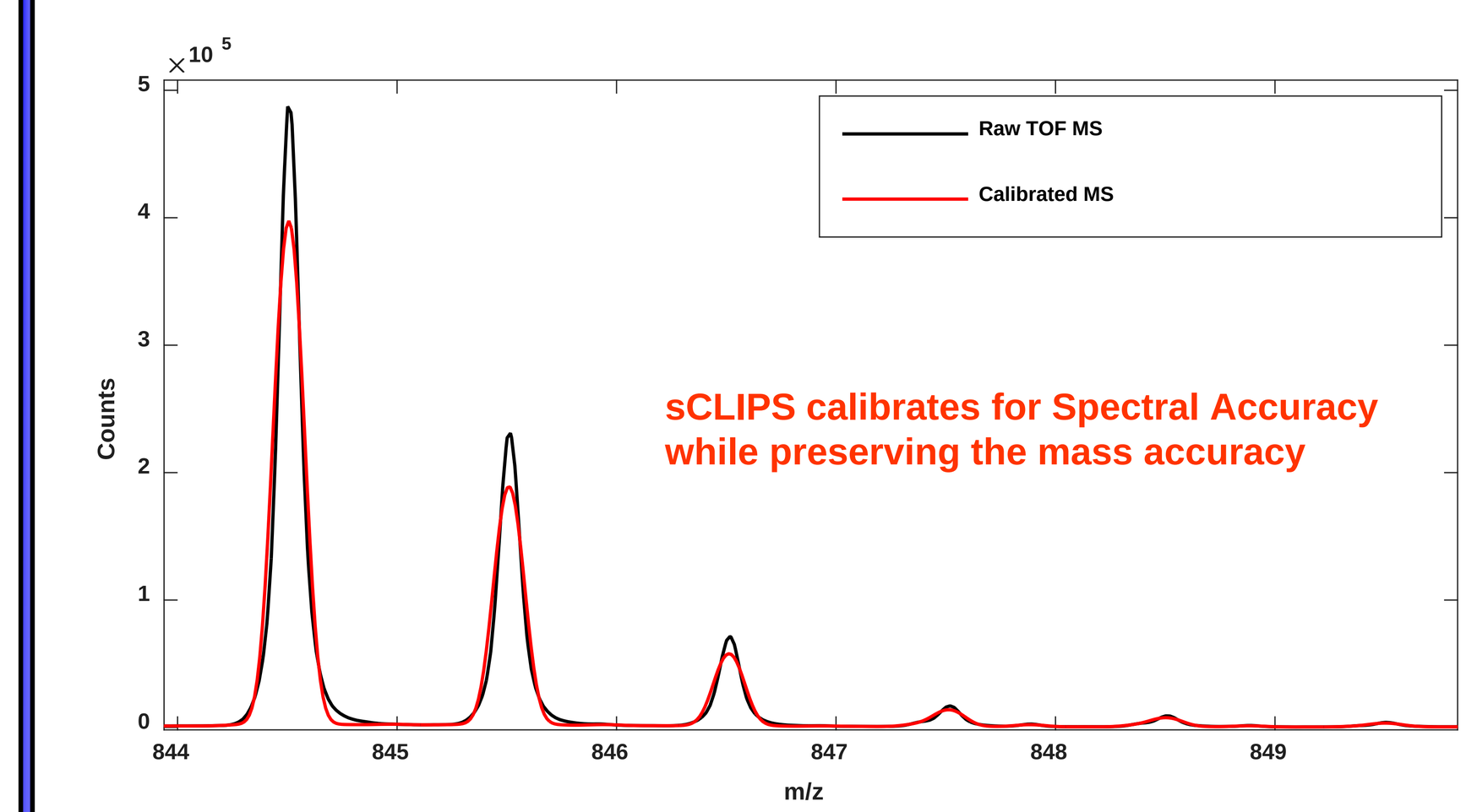
Rank	Formula	Mono Isotope	Mass Error (mDa)	Mass Error (ppm)	Spectral Accuracy (%)	RMSE	DBE
1	C25H23N2OS	399.1526	-7.8	-19	99.10	98,026	15.5
2	C24H23N4S	399.1638	-19.0	-48	99.10	98,063	15.5
3	C26H23O2S	399.1413	3.5	9	98.97	111,214	15.5
4	C23H19N4OS	399.1274	17.4	44	98.86	123,659	16.5
5	C22H19N6S	399.1386	6.2	15	98.58	154,220	16.5
6	C23H27O4S	399.1625	-17.7	-44	98.00	217,331	10.5
7	C22H23O5S	399.1261	18.7	47	97.12	311,894	11.5
8	C21H23N2O4S	399.1373	7.5	19	96.84	342,249	11.5
9	C20H23N4O3S	399.1485	-3.7	-9	96.55	373,689	11.5
10	C26H23O4	399.1591	-14.3	-36	96.52	377,107	15.5
11	C24H19N2O4	399.1339	10.9	27	96.42	388,246	16.5
12	C19H23N6O2S	399.1598	-15.0	-38	96.26	405,937	11.5
13	C22H27N2OS2	399.1559	-11.1	-28	96.20	412,278	10.5
14	C23H19N4O3	399.1452	-0.4	-1	96.19	413,321	16.5
15	C23H27O2S2	399.1447	0.1	0	96.18	413,666	10.5
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
126	C17H35S5	399.1337	11.1	28	84.47	1,683,257	0.5

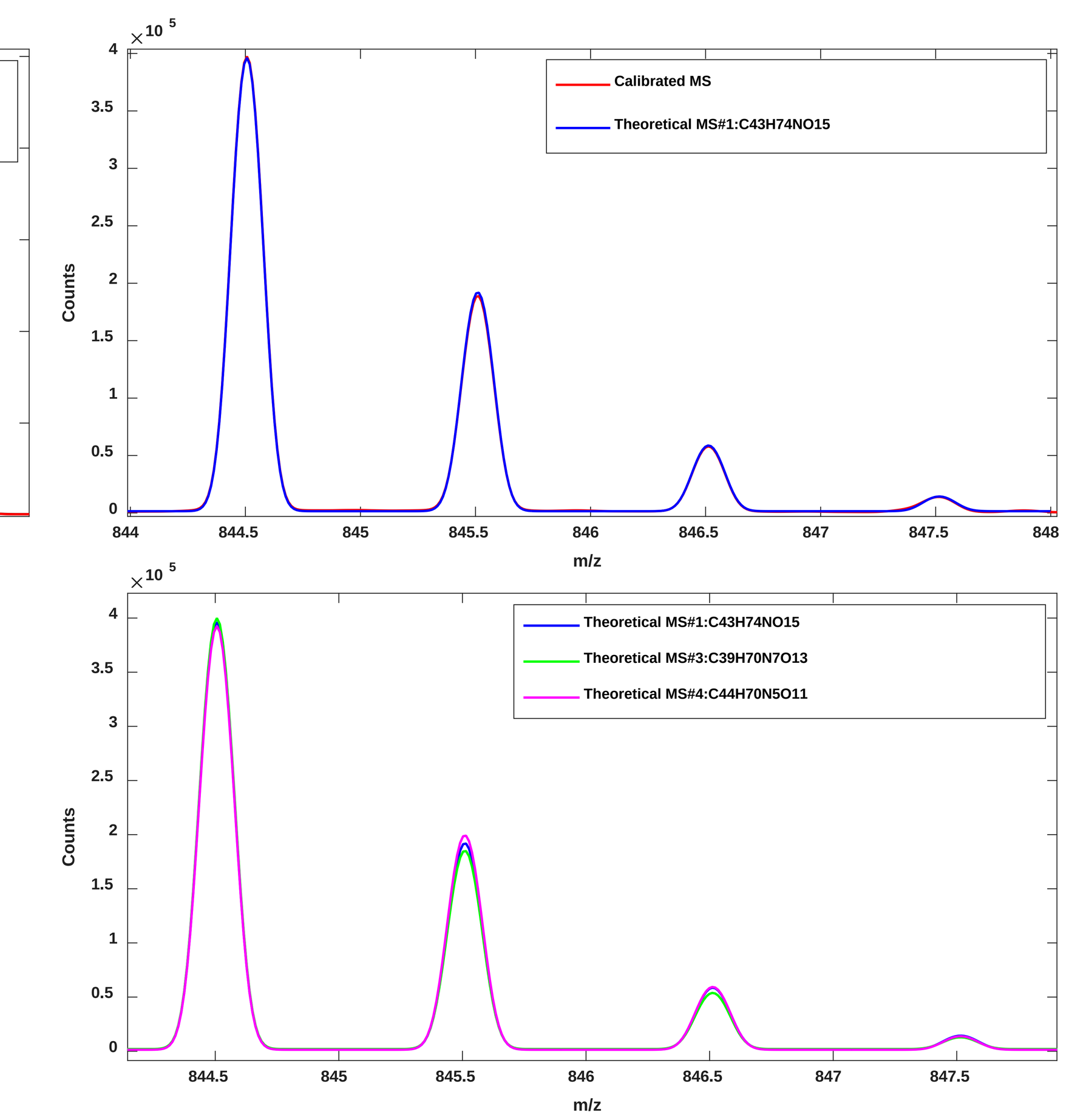
Given that the mass is only accurate to the 0.00x Da level, an unknown CLIPS search with 20mDa mass tolerance is needed, resulting in a long list of 126 possible candidates. When sorted by Spectral Accuracy, however, there are only 5 candidates with ≥ 98.5% Spectral Accuracy.

Rank	Formula	Mono Isotope	Mass Error (mDa)	Mass Error (ppm)	Spectral Accuracy (%)	RMSE	DBE
1	C25H23N2OS	399.1526	-0.1	-0.2	99.10	98,026	15.5
2	C10H27N10OS3	399.1526	-0.1	-0.2	87.94	1,307,659	2.5
3	C17H19N8O4	399.1524	0.1	0.3	93.78	674,171	12.5
4	C10H19N14O2S	399.1531	-0.6	-1.4	90.78	999,851	8.5
5	C17H27N4O3S2	399.1519	0.6	1.5	93.27	729,995	6.5
6	C9H23N10O6S	399.1517	0.8	1.9	88.80	1,213,758	3.5
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
119	C16H31O5S3	399.1328	19.7	49.3	89.35	1,154,517	1.5

When searching around the exact mass of the correct elemental composition and sorting solely on mass accuracy, a mass accuracy of better than 1.5ppm is required to land the correct elemental composition among the top 5 candidates.

Higher Resolution TOF or Orbitrap MS





Accurate Mass 844.5067
Charge 1
Mass Tolerance (ppm) 5
Electron State Even
DBE [Min, Max] [-1, 50]
MS Profile Range (Da) [-1, 3.5]

Empirical Rules Enabled
Empirical Limits Wiley
H/C Ratio Extended
Heteroatom Ratios Extended

Element [Min, Max]
C [1, 66]
H [0, 110]
N [0, 20]
O [0, 25]
S [0, 13]

Rank	Formula	Mono Isotope	Mass Error (mDa)	Mass Error (ppm)	Spectral Accuracy (%)	RMSE	DBE
1	C43H74NO15	844.5053	1.4	1.7	98.97	757	7.5
2	C40H66N11O9	844.5039	2.8	3.3	98.60	1,022	13.5
3	C39H70N7O13	844.5026	4.1	4.8	98.43	1,148	8.5
4	C44H70N5O11	844.5066	0.1	0.1	97.30	1,976	12.5
5	C38H66N15O5S	844.5087	-2.0	-2.3	97.21	2,040	13.5
6	C37H70N11O9S	844.5073	-0.6	-0.7	97.14	2,093	8.5
7	C41H62N15O5	844.5053	1.4	1.7	97.10	2,125	18.5
8	C33H66N17O7S	844.5046	2.1	2.4	96.64	2,461	9.5
9	C41H74N5O11S	844.5100	-3.3	-3.9	96.21	2,771	7.5
10	C39H62N19OS	844.5100	-3.3	-3.9	96.18	2,792	18.5
11	C40H78NO15S	844.5087	-2.0	-2.3	96.11	2,844	2.5
12	C36H74N7O13S	844.5060	0.7	0.8	95.98	2,938	3.5
13	C45H66N9O7	844.5080	-1.3	-1.5	95.50	3,293	17.5
14	C42H58N19O	844.5066	0.1	0.1	95.35	3,399	23.5
15	C32H70N13O11S	844.5033	3.4	4.0	94.84	3,774	4.5
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
63	C46H86NS6	844.5079	-1.2	-1.4	74.70	18,511	4.5

sCLIPS search within the typical 5ppm mass tolerance window results in a long list of 63 candidates! When sorted by Spectral Accuracy, however, there are only 3 or 2 candidates with ≥ 98% or 98.5% Spectral Accuracy, respectively, narrowing the long list down to a much more manageable size.

Rank	Formula	Mono Isotope	Mass Error (mDa)	Mass Error (ppm)	Spectral Accuracy (%)	RMSE	DBE
1	C43H74NO15	844.5053	0.00	0.00	98.96	759	7.5
2	C41H62N15O5	844.5053	0.01	0.02	97.09	2,130	18.5
3	C49H66N9O2S	844.5055	-0.17	-0.20	90.14	7,214	21.5
4	C34H70N17O2S3	844.5055	-0.20	-0.24	90.21	7,163	8.5
5	C34H78N13OS5	844.5050	0.27	0.31	82.25	12,991	2.5
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
139	C39H80N4O9S3	844.5082	-4.19	-4.97	88.65	8,307	2

When searching around the exact mass of the correct elemental composition and sorting solely on mass accuracy, a very high mass accuracy of better than 0.2ppm or 0.02ppm is required in order to land the correct elemental composition among the top 3 or 2 candidates, respectively.

Conclusion

Effective Mass Accuracy (EMA):

- **[20mDa Mass Accuracy, 98.5% Spectral Accuracy] ≡ 1.5ppm: Quadrupole MS**
  20mDa (50ppm @ 400Da) mass error combined with 98.5% spectral accuracy, which can be achieved on most single and triple quad MS systems through TrueCal, has an EMA of 1.5ppm, with the correct elemental composition among the top 5 hits.

- **[5ppm Mass Accuracy, 98.0% Spectral Accuracy] ≡ 0.2ppm: HiRes MS**
  5ppm mass accuracy (@845Da) combined with 98.0% spectral accuracy, which can be achieved on HiRes MS systems with sCLIPS peak shape calibration, has an EMA of 0.2ppm, with the correct elemental composition among the top 3 hits.

- **[5ppm Mass Accuracy, 98.5% Spectral Accuracy] ≡ 0.02ppm: HiRes MS**
  At the same 845Da and 5ppm mass accuracy combined with 98.5% spectral accuracy, which can be achieved on some HiRes MS systems via careful experimentation and sCLIPS peak shape calibration, has an EMA of 0.02ppm, with the correct elemental composition among the top 2 hits, making unique elemental composition determination within reach even for compounds with higher m/z near 1,000Da.

References

1. Kind, T., Fiehn, O., BMC Bioinformatics, 2007, 8:105.
2. Wang, Y., Gu, M., Anal. Chem., **2010**, *82*, 7055.