# ESI-MS Peptide Interpretation Guide

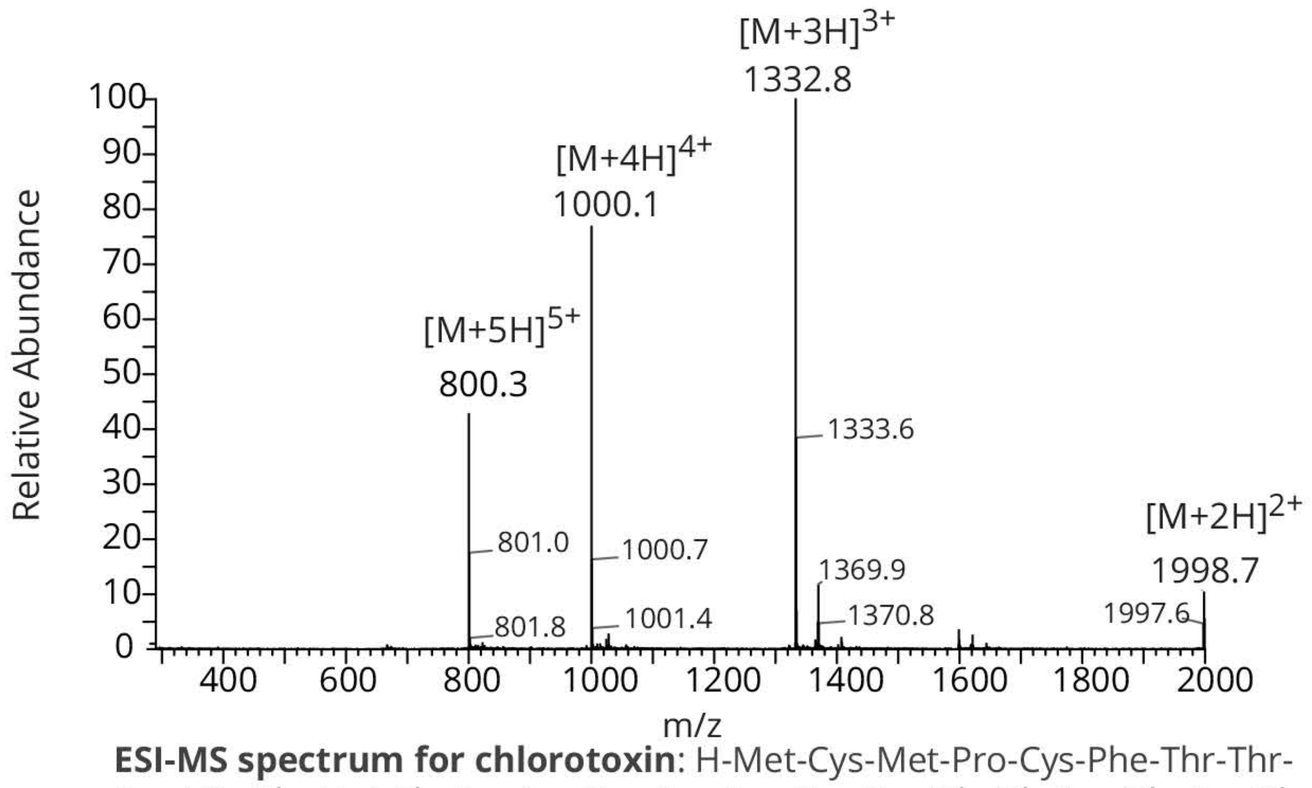
Understanding the Intricacies of Electrospray Ionization Mass Spectrometric Data

JULY 2021

If you're new to electrospray ionization mass spectroscopy (ESI-MS) as a technique for peptide characterization, understanding and interpreting the spectrum can be daunting. Other techniques such as high-performance liquid chromatography (HPLC) show one major peak (i.e., absorbance and purity for RP-HPLC) in the spectra, making the interpretation straightforward. ESI-MS spectra are more complicated due to the multiplicity of the data. While most small molecules (100-500 AMU) behave well in ESI-MS and ionize as a single charged state (i.e., peak $[M + H]^+$), larger biomolecules such as peptides, ionize as multiple charged variants (i.e., $[M + H]^+$, $[M + 2H]^{2+}$, $[M + 3H]^{3+}$, etc.). A major advantage of ESI-MS over other mass spectroscopy techniques is that molecular fragmentation is rare (t-Boc groups and a few other protection groups are an exception) which is why ESI-MS is often referred to as a soft ionization technique.

Generally, an ESI mass spectrometer is composed of three main components: (1) an ion source, (2) the mass analyzer, and (3) a detector. Molecular ions are generated by the ionization chamber where the ion source is kept, and then are transferred to the mass analyzer where the ions are separated according to their mass to charge ratio (m/z value). The ion concentrations and m/z value are then displayed on a chart called a mass spectrum (see spectrum below). The spectrum of ESI-MS contains a Y-axis labeled with the relative intensity (or relative abundance) and an X-axis labeled with m/z. ESI-MS is typically run in positive mode for peptides, because the molecules are easily protonated to form cations ($[M + H]+$, $[M + 2H]^{2+}$, $[M+3H]^{3+}$) or other adducts ($[M + Na]^+$, $[M + K]^+$, $[M + Na + H]^{2+}$, $[M + K + H]^{2+}$, etc.).

In the case of the peptide chlorotoxin (molecular weight = 3995.77 g/mol, molecular formula = $C_{158}H_{249}N_{53}O_{47}S_{11}$), we observe hydrogen ion adducts $[M + 2H]^{2+}$, $[M + 3H]^{3+}$, $[M + 4H]^{4+}$, and $[M + 5H]^{5+}$ in the ESI-MS spectrum. Chlorotoxin contains multiple basic side chains; namely, the guanidino group of arginine and amino group of lysine, any of which may be protonated to form hydrogen ion adducts $(M + zH)+z$. Protonation is common in peptide ESI-MS analysis because the solvents (often acidic water in acetonitrile) are protic and acidic – typically an HPLC buffer (i.e., mobile phase) containing small amounts of trifluoroacetic acid (TFA). In some cases, formic acid is added to the sample to facilitate protonation because the formic acid conjugate base is easier to remove than the conjugate base of TFA.




**ESI-MS spectrum for chlorotoxin**: H-Met-Cys-Met-Pro-Cys-Phe-Thr-Thr-Asp-His-Gln-Met-Ala-Arg-Lys-Cys-Asp-Asp-Cys-Cys-Gly-Gly-Lys-Gly-Arg-Gly-Lys-Cys-Tyr-Gly-Pro-Gln-Cys-Leu-Cys-Arg-NH2
(Disulfide Bridge: Cys1-Cys4, Cys2-Cys6, Cys3-Cys7, Cys5-Cys8)

Molecule Weight: 3995.77 g/mol



CHINESE PEPTIDE

Before confirming the mass of the peptide, it's important to identify the peaks in the spectrum that are associated with M. If we assume that a set of peaks are ion variants of M, then it should be possible to calculate the molecular weight from any of the peaks. If adjacent peaks at 1998.7 and 1332.8 m/z are ion variants of M, then ion 1332.8 m/z should contain an additional hydrogen (i.e., z + 1).

$$\text{Equation for peak 1332.8: } m + 1 = 1332.8(z + 1)$$

$$\text{Equation for peak 1998.7: } m = 1998.7(z)$$

Because the mass of the peptide is the same for each peak, we can set the two equations equal to one another and then solve for charge (z):

$$1332.8z + 1332.8 - 1 = 1998.7z$$

$$1332.8z + 1331.8 = 1998.7z$$

$$1331.8 = 1998.7z - 1332.8z$$

$$1331.8 = 665.9z$$

$$z = 2$$

The results show that the two peaks are separated by one charge unit (+1) and that peak 1998.7 has a charge of +2 (z = 2). Please note that peak $[M + H]^{+1}$ is not visible in the spectrum and the absence of the $[M + H]^{+}$ peak is common for peptides with molecular weights greater than 3000 g/mol.

The remaining peak charges may be calculated in a similar fashion:

| Peak (m/z) | Charge | Peak Label | Calculation | Calculated Mass |
|---|---|---|---|---|
| 1998.7 | 2 | $[M +2H]^{2+}$ | (1998.7 – 1) x 2 = 3995.4 | 3995.4 |
| 1332.8 | 3 | $[M + 3H]^{3+}$ | (1332.8 – 1) x 3 = 3995.4 | 3995.4 |
| 1000.1 | 4 | $[M + 4H]^{4+}$ | (1000.1 – 1) x 4 = 3995.4 | 3995.4 |
| 800.3 | 5 | $[M + 5H]^{5+}$ | (800.3 – 1) x 5 = 3995.5 | 3995.5 |

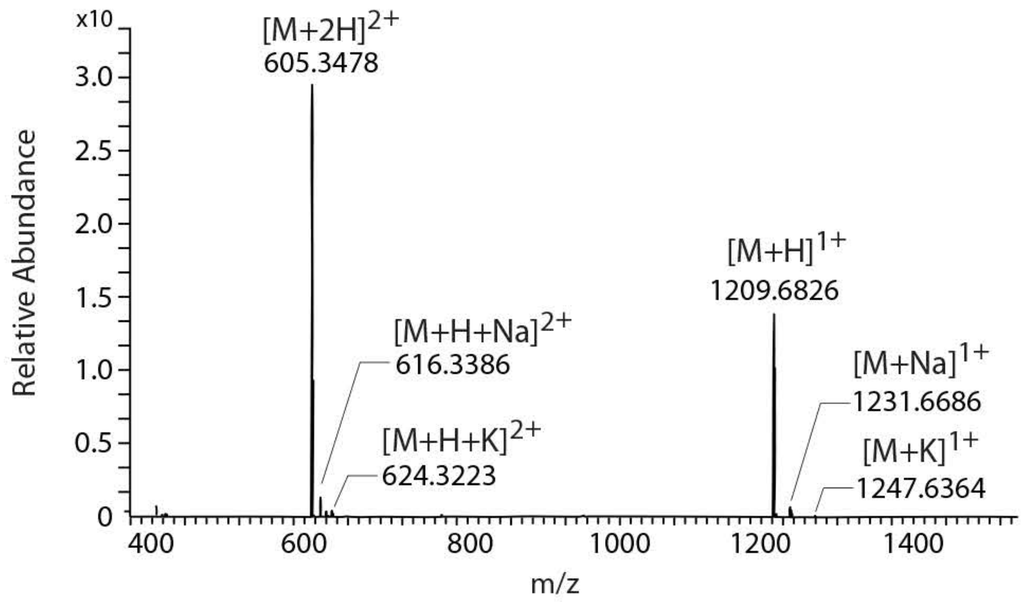
Now that we know the charged state of the peaks, we can calculate the mass of the peptide from each peak. The final molecular weight is derived by calculating the average mass of all the charged peaks:

$$(3995.4 + 3995.4 + 3996.4 + 3995.5) / 4 = 3995.68$$

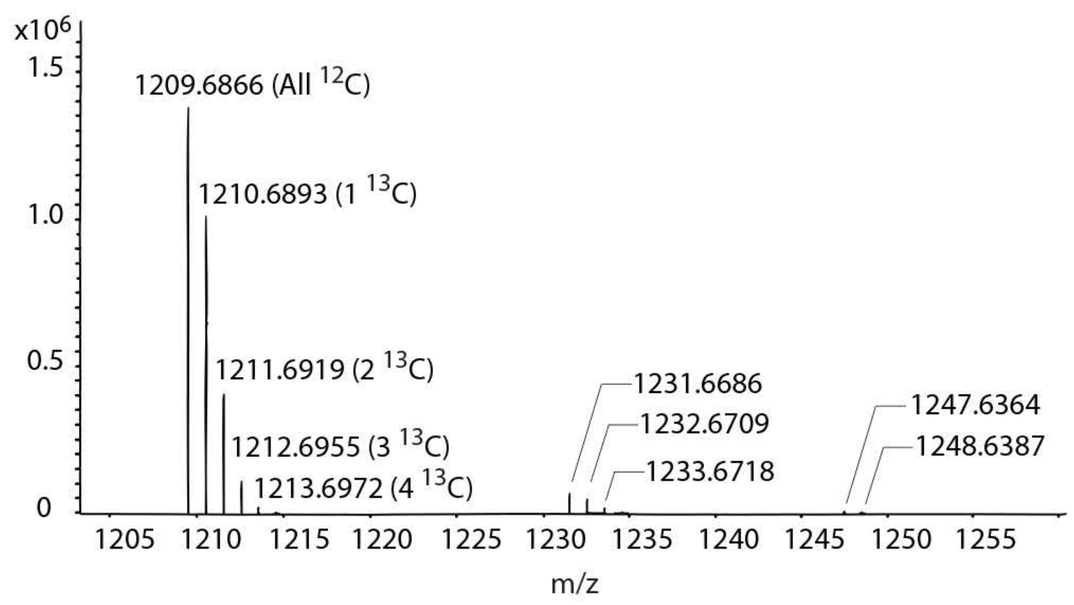
The calculated molecular weight of 3995.68 is within a tenth of a mass unit from the known molecular weight of Chlorotoxin (3995.77 g/mol).

The following leuprolide spectrum was obtained by high-resolution mass spectroscopy (Agilent 6530 Accurate-Mass Q-TOF LC/MS). You may notice by comparison to a low-resolution spectrum, that the high-resolution spectrum reports peaks with significantly higher mass accuracies. Low-resolution spectroscopy instruments typically report masses accurate to one-tenth AMU (e.g., 3995.5), whereas high-resolution MS report masses to the one-ten thousandth of an AMU. This additional accuracy enables the resolution of isotopic mass peaks (see spectrum below). It is the weighted (i.e., abundance) average of the isotopic peaks that accounts for the molecular weight of a compound. Carbon-12 ($^{12}C$) is the common element in peptides; carbon-13 ($^{13}C$) is an isotope of $^{12}C$ and accounts for about 1.2% of the total carbon abundance.

## Leuprolide - High-Resolution ESI-MS




## Isotopic Peak Expansions




Due to the relatively low molecular weight compared to chlorotoxin, the $[M + H]^{1+}$ peak of Leuprolide is clearly visible. In addition to the main ion peak (M +1 or $[M + H]^{+}$), we also observe $[M + Na]^{1+}$, $[M + K]^{1+}$, $[M + 2H]^{2+}$, $[M + H + Na]^{2+}$, and $[M + H + K]^{2+}$. In spectrum expansion, we see the mono-isotopic peaks for $[M + H]^{1+}$ at 1209.6866, 1210.6893, 1211.6919, 1212.6955, and 1213.6972. The molecular weight of leuprolide can be obtained from the weighed average of the four peaks. The peaks are organized by the number of $^{13}C$ present in the peptide with the first peak corresponding to peptide containing only $^{12}C$. The second peak contains exactly one $^{13}C$, and the third peak contains two $^{13}C$. At first glance you may be surprised at the intensities of the second two peaks when you consider $^{13}C$ only accounts for 1.1% of carbon. However, since leuprolide contains 59 carbons in each molecule, the probability of one carbon being a $^{13}C$ isotope is high.

### Peak Report ($1^{+}$ charge)

| Peak | m/z | # $^{13}C$ |
|---|---|---|
| 1209.6866 | $[M + H]^{+1}$ | 0 |
| 1210.6893 | $[M + H]^{+1}$ | 1 |
| 1211.6919 | $[M + H]^{+1}$ | 2 |
| 1212.6955 | $[M + H]^{+1}$ | 3 |
| 1213.6972 | $[M + H]^{+1}$ | 4 |
| 1231.6686 | $[M + Na]^{+1}$ | 0 |
| 1232.6709 | $[M + Na]^{+1}$ | 1 |
| 1233.6718 | $[M + Na]^{+1}$ | 2 |
| 1234.6719 | $[M + Na]^{+1}$ | 3 |
| 1247.6364 | $[M + K]^{+1}$ | 0 |
| 1248.6387 | $[M + K]^{+1}$ | 1 |

### Peak Report ($2^{+}$ charge)

| Peak | m/z | # $^{13}C$ |
|---|---|---|
| 605.3478 | $[M + 2H]^{+2}$ | 0 |
| 605.8487 | $[M + 2H]^{+2}$ | 1 |
| 606.3495 | $[M + 2H]^{+2}$ | 2 |
| 606.8516 | $[M + 2H]^{+2}$ | 3 |
| 616.8397 | $[M + H + Na]^{+2}$ | 0 |
| 617.3397 | $[M + H + Na]^{+2}$ | 1 |
| 617.8391 | $[M + H + Na]^{+2}$ | 2 |
| 624.3223 | $[M + H + K]^{+2}$ | 0 |
| 624.8235 | $[M + H + K]^{+2}$ | 1 |
| 625.3241 | $[M + H + K]^{+2}$ | 2 |

With low-resolution ESI-MS, charge assignment of the peaks requires some assumptions and arithmetic. High-resolution ESI-MS, on the other hand, leaves nothing to chance and requires no estimations. The distance between the monoisotopic peaks determines the charge of the peak cluster. For example, the isotopic peaks at 1209.6866, 1210.6893, 1211.6919, 1212.6955, and 1213.6972 are one AMU apart and therefore correspond to the $[M + H]^{+}$ charged species. The same is true for [M + Na]+ (1231.6686, 1232.6709, 1233.6718, 1234.6719) and $[M + K]^{+}$ – all of which differ by 1.0 m/z. The isotopic peaks for $[M + 2H]^{2+}$, on the other hand, are separated by 0.5 m/z (605.3478, 605.8487, 606.3495, 606.8516 m/z). The isotopic pattern continues with $[M + H + Na]^{2+}$ (616.8397, 617.3397, 617.8391 m/z) and $[M + H + K]^{2+}$ (624.3223, 624.8235, 625.3241 m/z). The final molecular weight is calculated from the weighted averages of the isotopic peaks. The accuracy can often depend on the electrospray software package peak selection algorithm. Interestingly, ESI-MS for lower molecular weight peptides may sometimes incorrectly select the mono-isotopic peak for the mass instead of calculating the average isotopic mass (i.e., molecular weight).



CHINESE PEPTIDE

CHINESE PEPTIDE


Headquarter & Manufacturing Site

US Headquarter

US-based Manufacturing Site

DP Research and Manufacturing Center

**Rocklin, CA, US**
**3880 Atherton Rd, Rocklin, CA 95765**

**San Jose, CA,US**
**160 E Tasman Dr. Suite 200, San Jose, CA 95134**

**Hangzhou,China**
**No.69,12th Street,Qiantang District**

+86 571 86737011

sales@chinesepeptide.com

www.chinesepeptideco.com