Introduction

High resolution mass spectrometers yield to spectra containing numerous signals. Challenging research problems can be solved using those spectra but they are often too complex to be analysed by hand. Therefore it may be required to create a custom analytical tool in order to extract the knowledge hidden in those experimental data. In order to solve quickly those problems, we have developed 3 core functionalities dealing with molecular formula and a scripting language allowing to process mass spectra. Those tools are available online (i.e. www.chemcalc.org).

1. Theoretical isotopic distribution

Isotopic distribution can be calculated using ChemCalc by specifying the FWHM or the resolution of the spectrometer. This will apply a gaussian that will yield to a spectrum that is very close to the experimental results. It is possible not only to specify the charge and to create mixture of molecules / fragments but also to specify the isotopic enrichment of a specific element and to use common group abbreviations.

For example molecular formula may be introduced as:

- HAlaOH i.e. alanine
- HAlaOHH.HAlaOHH i.e. a mixture of 2 decapotide single and doubly protonated
- C12D16O14 i.e. decaene composed of enriched carbon, 98% 13C, 2% 12C

2. Modified peptide fragmentation

Biological activity is observed in many natural or synthetic peptide sequences yielding to an intense research field. Peptide analogs containing non-natural amino acids as well as unconventional side chain modifications have been synthesised. Despite those modification, the mass analysis of such peptides may yield to usual fragmentation (a,b,c and x,y,z) but most of the tools are unable to predict the corresponding fragments masses. Using simple regular expression (i.e. matching a lowercase or parenthesis followed by an uppercase) we are able to virtually split the string representing the sequence in all its fragments. For example "HAlaLys(H++)GlySer(H+Ph)ProGlyOH" describes a peptide sequence with a protonated lysine side chain and the replacement of the alcohol hydrogen of the serine by a phenyl group.

3. MF from monoisotopic mass

The analysis of complex mixtures using high resolution and precision spectrometers requires the development of algorithms that allow to determine possible molecular formula for a specific monoisotopic mass. On ChemCalc it is possible not only to specify the molecular formula range (e.g. C5H6O2 or C5H7O2: i.e. any molecule containing between 0 and 10 carbons and 0 and 22 hydrogens) but also to specify non- cleavable groups of atoms (e.g. (CH)2(CH2)2(CO)2H: allowing 0 to 10 copies of the PEG like chain and between 0 and 10 alanine diadical:). It is even possible to find all possible amino acids compositions of a peptide for a monoisotopic mass up to 1000.

4. Scripting research problems

Based on those 3 tools, script can be done to solve real life problems. As an example, a ruthenium derivative may bind to a peptide in an undetermined position. A script generates all the possible peptide fragmentations and complex with ruthenium derivative in order to help to annotate and assign the experimental spectrum.

Over 5000 molecular formula have been evaluated, i.e.:

1. Generate a theoretical isotopic distribution at a specific resolution
2. Calculate the similarity with the experimental spectrum based on a surface overlap algorithm

The results are sorted in an interactive visualisation tool available from the web browser.

Conclusions

Real life research problems are often not easy to solve using commercial softwares. We show here innovative applications of the core ChemCalc tools to help mass spectrometrists to analyse data directly from a web browser.

Based on a simple scripting language it is possible to automatically match predicted spectra with experimental data and generate interactive reports that can be shared between colleagues. Real case examples involving peptide complexes, isotopic enrichment and GC/MS have been solved.

Don’t hesitate to contact us with your new problems !!!

References

4. https://github.com/npellet/visualizer