Analysis of FTICR-MS Data on Humic Substances and Synthetic Polyelectrolytes Using Different Data Processing Techniques

Erast V. Kunenkov[1], Alexey S. Kononikhin[2], Irina V. Perminova[1], Andrey V. Garmash[1], Eugeniy N. Nikolaev[2], Igor A. Popov[2]

[1] Department of Chemistry, Lomonosov Moscow State University, 119992 Moscow, Russia
[2] Laboratory of Biomacromolecules, Emmanuel Institute of Biochemphysics of RAS, Moscow, Russia

Humic substances (HS) are complex mixtures formed during decay of plant and animal materials. HS are generically classified into fulvic and humic acid fractions according to their solubility properties. The particular features of HS are polydispersity and extreme structural heterogeneity. As a result, the reliable approaches to establishing relationship between their macro-properties and structure are missing. This is mostly due to low resolution ability of the analytical techniques conventionally used for their structural analysis such as elemental analysis, size exclusion chromatography, and others. FTICR mass spectroscopy is a most powerful tool capable to provide ultra high resolution data on structural organization of these substances. However, due to complexity of the humic systems, and, hence, of the FTICR-MS data obtained, development of adequate processing techniques is of particular importance for extraction of the relevant structural information from spectral data.

The goal of this research was to develop the data processing software, which would account for the extreme complexity of humic substances. To reach this goal, Rexx programming language was used. Ultrahigh resolution mass spectra were obtained using 7 Tesla LTQ FT mass spectrometer with electrospray ion source. We used well characterized samples of Suwannee River fulvic and humic acids of the International Humic Substances Society (IHSS). In addition, synthetic polyelectrolyte - polystyrene sulfonate (PSS) with known molecular weight of 2200 Da (size exclusion standards from Polymer Standard Service (Mainz, Germany) was used.

Negative ionization mode was used to obtain FTICR MS spectra of HS samples and PSS. About 3 thousands of different ions were detected in the 400-700 range in each spectrum. To process the data obtained, the software named FIRAN was designed as a set of modules destined for solving the following tasks: 1) ion charge detection based on isotope simulation; 2) data filtering based on the results of charge detection; 3) calculation of differential mass spectra, 4) calculation of Kendrick mass defects; 5) determination of stoichiometric formulas for conventional and differential spectra; 6) evaluation of the total elemental composition of the sample analyzed; and 7) plotting and viewing van Krevelen diagrams.

The principal novelty of the developed data processing software is a conversion of the measured mass spectra into differential mass spectra. A use of differential data processing approach was shown to be very efficient in identification of the monomer fragment in the molecules of regular synthetic polymer like PSS. The most intensive peak was observed for the fragment with molecular weight of 206.001 – which corresponds exactly to the mass of $C_8H_7SO_3Na$-fragment. The obtained differential spectra of HS displayed a lot of small fragments tending to form chains. At the same time, there was no regularly repeated series observed among the large fragments. Nevertheless, this method may help to study ions with high molecular weight, because of differences between masses of such ions are smaller and contain smaller absolute errors as compared to the absolute masses. As a result, there was no Kendrick mass sorting scheme capable to identify most ions in spectra. Implementation of different mass sorting schemes did not detect differences between variations of ion's elemental composition in range 400-700 Da.

According to isotope simulation results, most of HS ions are single-charged, but the number of multiple-charged ions is also significant. In case of not correct determination of ion charge, it may be assigned with a wrong formula, in particular, in case of wide mass window. We can conclude that the careful data filtering is a very important prerequisite for constructing and interpreting van Krevelen diagrams. It should allow for selection of the most intensive peaks and exclude noise from the diagram.