# Multivariate Statistical Analysis of Mass Spectra as a Tool for the Classification of the Main Humic Substances According to Their Structural and Conformational Features

Roberto Baigorri,*,[†] Angel Mari Zamarreño,[†] Marta Fuentes,[†,‡]
Gustavo González-Gaitano,[‡] Jose Maria García-Mina,[†,‡]
Gonzalo Almendros,[§] and Francisco Javier González-Vila[‖]

CIPAV-Roullier Group, Polígono Arazuri-Orkoien, 31160, Orkoien, Spain; Department of Chemistry and Soil Science, University of Navarra, 31080, Pamplona, Spain; Centro de Ciencias Medioambientales, CSIC, Serrano 115B, 28006, Madrid, Spain; and Instituto de Recursos Naturales y Agrobiología, CSIC, PO Box 1052, 41080, Sevilla, Spain

The aim of this work is to explore the suitability of the complementary use of mass spectra and the corresponding statistical analysis (principal components-Pareto analysis (PCA) and discriminant analysis (DA)) of these spectra to differentiate diverse humic samples as a function of their structural and conformational features. To this end, the mass spectra of humic samples belonging to the main humic fraction types (gray humic acid, brown humic acid, and fulvic acid) were obtained by electrospray ionization mass spectrometry (ESI-MS). The results obtained showed that the application of PCA yielded a clear separation between blanks and humic samples. However, a clear differentiation among the humic fraction types was not achieved. The DA of PCA data, however, yielded a clear separation among the humic substances (HS) samples belonging to each HS fraction type considered: gray humic acids, brown humic acids, and fulvic acids. These results showed that the mass spectra of each humic sample include characteristic mass/charge (*m/z*) distribution values that can be considered as a "fingerprint" representative of its specific structural features. Our results also indicate that, although the *m/z* values principally corresponded to single-charged ions, we cannot identify these molecular weight distributions with those of humic samples, since sample molecular fragmentation, as well as partial molecular ionization, cannot be ruled out under our experimental and instrumental conditions.

**KEYWORDS: Humic substances; mass spectrometry; principal components analysis; discriminant analysis; polymers; fragmentation**

## INTRODUCTION

Although many studies indicate that the qualitative and quantitative interpretation of the mass spectra of humic substances (HS) is frequently very difficult (*1–3*), recent studies have demonstrated that the use of sophisticated statistical tools, such as principal components analysis (PCA) or multilayer neural networks, could permit the adequate classification of humic samples as a function of their origin and sampling places (*2, 3*). In principle, these studies indicate that HS mass spectra contain specific groups of *m/z* values that correspond to structural features specific for each family of samples. Thus,

these results suggest that these statistical tools could also be useful to differentiate humic samples with diverse physico-chemical properties and eventually to get an insight into their structural domains and molecular weight.

Recent studies carried out in our laboratory indicate that the main humic fractions that can be obtained as a function of pH and ionic strength (gray humic acids (GHA), brown humic acids (BHA), and fulvic acids (FA)) present different and singular structural features and molecular conformations in solution (*4*). Thus, GHA presented a macromolecular behavior and aliphatic character; BHA, however, presented a supramolecular/macro-molecular behavior and aromatic character, whereas FA presented a supramolecular behavior and aromatic character. It is therefore of interest to explore the suitability of the complementary use of the mass spectrometry and the corresponding statistical data analysis to both discriminate among these fractions and, eventually, obtain new information about the structural features and main size of these humic fractions.

* To whom correspondence should be addressed. Phone: 34 948 324752. Fax: 34 948 324032. E-mail: rbaigorri@timacagro.es.
† CIPAV-Roullier Group.
‡ University of Navarra.
§ Centro de Ciencias Medioambientales.
‖ Instituto de Recursos Naturales y Agrobiología.

Classification of Main Humic Substances by MS

*J. Agric. Food Chem.,* Vol. 56, No. 14, 2008 **5481**

**Table 1.** Elemental and Functional (Potentiometry) Analyses of the Different Humic Fractions

| | g/100 g | | | | | | | acidity in mmol/g HS | | | |
| | C | H | N | *O | H/C | O/C | E4/E6 | strong acidity | COOH | phenolic OH | total acidity |
|---|---|---|---|---|---|---|---|---|---|---|---|
| AGHA | 55.6 | 4.12 | 2.57 | 38.2 | 0.074 | 0.69 | 2.07 | 0.10 | 0.90 | 0.48 | 1.48 |
| ABHA | 45.1 | 2.52 | 0.81 | 51.6 | 0.056 | 1.14 | 6.57 | 0.50 | 2.83 | 1.15 | 4.48 |
| LGHA | 68.9 | 6.84 | 0.66 | 23.6 | 0.099 | 0.34 | 2.33 | 0.14 | 0.30 | 0.25 | 0.69 |
| LBHA | 38.3 | 1.90 | 0.58 | 55.6 | 0.050 | 1.45 | 4.37 | 0.20 | 2.16 | 1.14 | 3.50 |
| ZGHA | 55.2 | 4.01 | 0.51 | 40.3 | 0.073 | 0.73 | 3.06 | 0.12 | 1.42 | 0.71 | 2.25 |
| ZBHA | 43.2 | 2.50 | 0.38 | 53.9 | 0.058 | 1.25 | 4.92 | 0.33 | 2.22 | 0.98 | 3.54 |
| CGHA | 57.6 | 2.67 | 0.66 | 39.1 | 0.046 | 0.68 | 2.52 | 0.30 | 0.63 | 0.65 | 1.58 |
| CBHA | 52.1 | 0.75 | 0.27 | 46.9 | 0.014 | 0.90 | 4.17 | 0.30 | 1.79 | 1.31 | 3.40 |
| BFA | 40.4 | 3.41 | 0.70 | 55.5 | 0.084 | 1.37 | 14.2 | 0.63 | 4.21 | 2.90 | 7.74 |
| ZFA | 31.5 | 2.76 | 0.14 | 65.6 | 0.088 | 2.08 | 16.1 | 1.47 | 2.49 | 0.69 | 4.56 |
| WPFA | 53.6 | 4.24 | 1.07 | 41.8 | 0.079 | 0.78 | 6.80 | 1.18 | 1.72 | 0.86 | 3.76 |
| SRFA | 53.0 | 4.36 | 0.75 | 43.9 | 0.082 | 0.83 | 19.8 | 1.56 | 2.13 | 0.93 | 4.62 |

To this end, we have studied the mass spectra of GHA, BHA, and FA with different origins using a discriminant/multivariate analysis including the Pareto algorithm that has been proposed as especially suitable to the analysis of mass spectra data of organic substances (*5*). In order to facilitate the interpretation of the data obtained under the experimental conditions of our samples and instrumental equipment, we have also carried out complementary studies using well-characterized organic polymers.

## MATERIALS AND METHODS

**Humic Materials.** The humic substances selected for the study were a lignite humic acid obtained from China (CHA), a young brown coal (leonardite) humic acid extracted from a soil sample of the Czech Republic (ZHA) (*6*), a commercial humic acid from Aldrich Chemicals (AHA), and a Leonardite Standard Humic Acid (LHA) purchased from the International Humic Substances Society (IHSS). Two fulvic acids, extracted from a peat of Brazil (BFA) and from a soil sample of the Czech Republic (ZFA) (*6*), and two IHSS standard fulvic acids, Suwannee River fulvic acid (SRFA) and Waskish Peat fulvic acid (WPFA), were included in this work.

**Extraction Procedure.** The different HS were isolated and purified following the IHSS procedure (*7, 8*) and freeze-dried. The humic acids (HA) were fractionated as Swift (*9*) describes on the salting-out procedure by KCl to fractionate HA into gray and brown fractions. In our work, we dissolved 12 g/L of HA in KCl (2 M, pH 7) overnight and subsequently centrifuged. The solution containing BHA and the precipitate containing GHA were purified following the IHSS procedure (*7, 8*) until the ash contents were less than 2%. The samples were named as follows: AGHA (Aldrich gray humic acid), ABHA (Aldrich brown humic acid), LGHA (Leonardite standard gray humic acid), LBHA (Leonardite standard brown humic acid), ZGHA (Czech gray humic acid), ZBHA (Czech brown humic acid), CGHA (Chinese gray humic acid), CBHA (Chinese brown humic acid), BFA (Brazilian fulvic acid), ZFA (Czech fulvic acid), WPFA (Waskish Peat reference fulvic acid), and SRFA (Suwannee River reference fulvic acid). In **Table 1**, we show the main chemical and physical properties of different humic fractions.

**Polymer Standards.** A nominal 1200 Da molecular weight poly-(acrylic acid) (PAA) purchased from Sigma-Aldrich (St. Louis, MO) and a 1100 Da molecular weight poly(ethylene glycol) (PEG) purchased from Polymer Laboratories (Shropshire, U.K.) were used for the study on the possible fragmentation of stable polymers during the ionization process in the mass instrument. All humic materials and polymer standards were dissolved in adequate quantity in mixtures of 50% Milli-Q deionized water and 50% HPLC grade methanol purchased from Merck (Darmstadt, Germany). This solvent was analyzed as a blank before each sample and at the end of the analysis (16 blanks in total).

**Instrumentation and Mass Spectrometry.** The MS/MS experiments were performed in a 3200 Q Trap LC/MS/MS system equipped with an electrospray ionization (ESI) source from Applied Biosystems by direct infusion with a flow rate of 10 mL/min in negative mode. Nitrogen was used as the collision and curtain gas at a pressure of

$10^{-5}$ bar and 20 psi, respectively. The following source voltages were used: declustering potential, $-220$ V; entrance potential, $-10$ V; collision energy, $-5$ V, as low as possible to minimize fragmentation. For each spectrum, 120 scans were summed. All data were recorded and analyzed with Analyst 1.4.2 software.

In order to check if the mass/charge detection range could limit the detection of higher molecular weights, we studied some of the HS samples and polymers using a Time-Of-Flight (TOF) mass spectrometer Micro from Micromass.

**Statistical Software.** The mass spectrometric data were treated by MarkerView 1.1 software supplied by Applied Biosystems, and the multivariate data analyses (discriminant analysis) were performed by SPSS 12.0 software (SPSS Inc., Chicago, IL).

## RESULTS AND DISCUSSION

As can be observed in **Figure 1**, the mass spectra corresponding to the different humic samples appear qualitatively very close to each other, and a discrimination of the different samples according to the shape of the mass spectra and the main *m/z* values seems to be very difficult. These results were in line with those obtained by a number of authors in studies dealing with humic samples with different origins (*2*), different structural features (*10*), or even different fulvic subfractions (*10*). However, Peña-Méndez et al. (*3*) indicated that the statistical analysis of HS mass spectra permits the differentiation of the different HS samples as a function of their origin and place. Thus, the statistical analysis of HS mass spectra yielded a HS classification that demonstrated that these spectra can be used as a "fingerprint" related to HS natural genesis (*3*).

In our case, the application of PCA and discriminant/multivariate analysis permitted us to discriminate among the different humic groups (GHA, BHA, and FA) independently of the origin of the samples. Thus, the statistical study of PCA of the HS spectra (**Figure 2A**) shows that the first two PC explained 89.6% of the cumulative variance. In fact, the scores plot for mass spectra of samples projected on the first two PCs show two main groups: a group involving the blanks and a second group involving the humic samples (**Figure 2B** and **C**). Nevertheless, we cannot differentiate between the different humic fractions. The complexity of direct mass spectra interpretation and the convenience of using complementary statistical tools are reflected in the cases of AGHA and SRFA mass spectrum. Thus, although AGHA mass spectra seem to be rather different from those for the other gray humic fractions (**Figure 1**), PCA analysis shows that AGHA presented the dominant variables that configure the gray humic acid group (**Figure 2B**). On the contrary, although SRFA spectrum seems to be quite similar to those of the other fulvic fractions (**Figure 1**), PCA analysis was not able to include this sample in the fulvic group. However, as is discussed below, the complementary use of DA
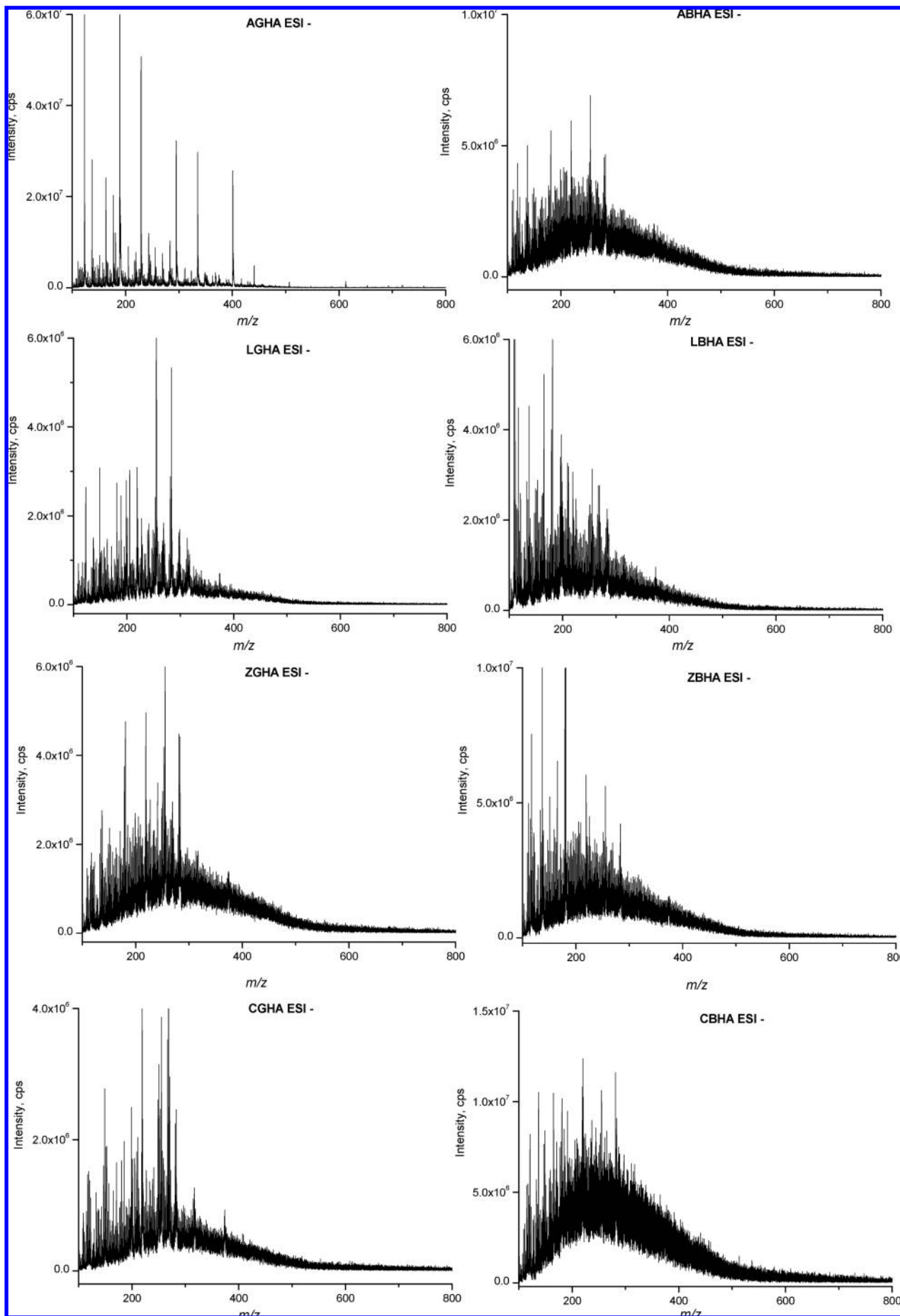
**Figure 1.**   (1 of 2)

Classification of Main Humic Substances by MS

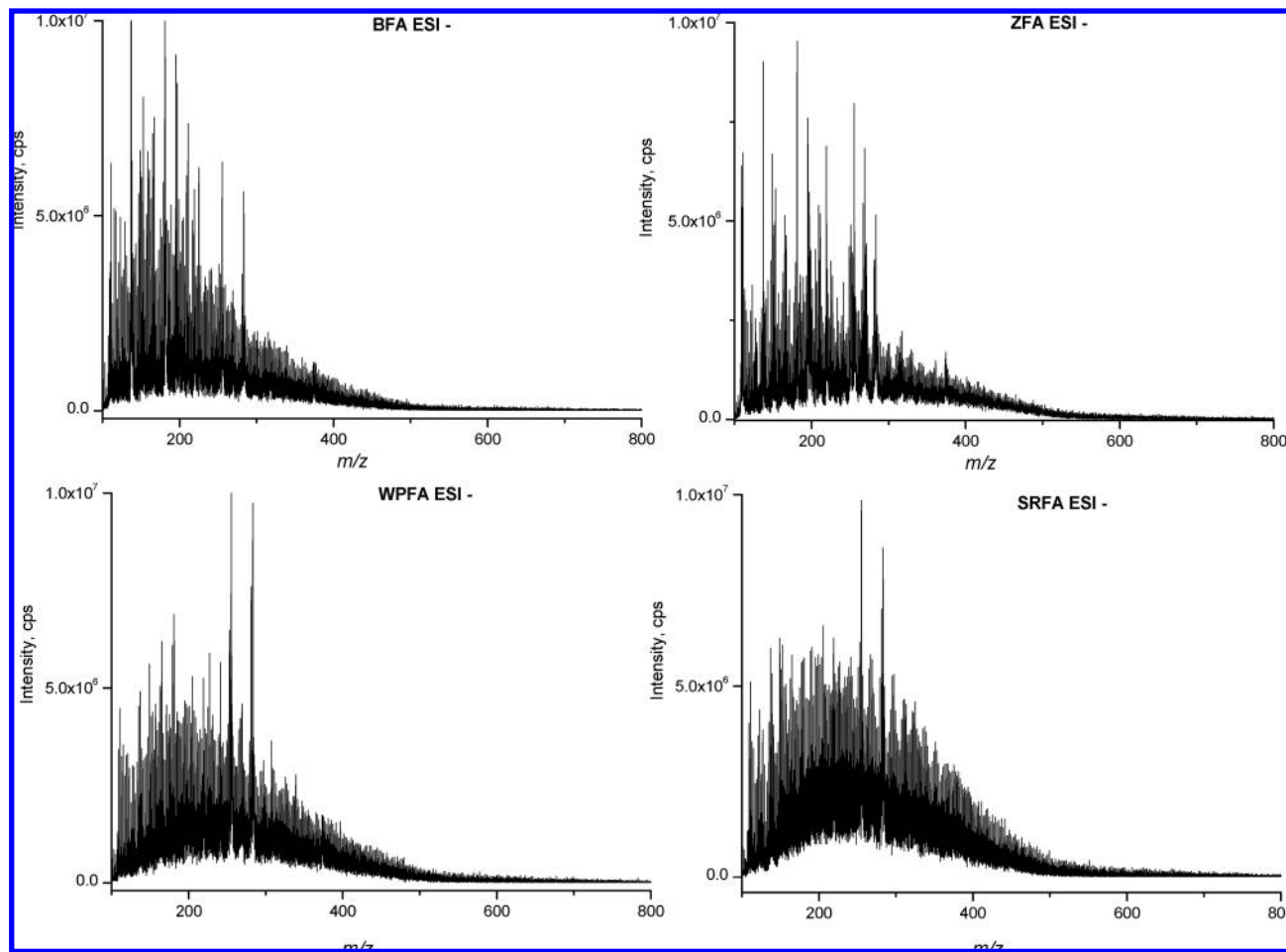*J. Agric. Food Chem.,* Vol. 56, No. 14, 2008 **5483**



**Figure 1.** Mass spectra of different HS fractions in solution. AGHA (Aldrich gray humic acid), ABHA (Aldrich brown humic acid), LGHA (Leonardite standard gray humic acid), LBHA (Leonardite standard brown humic acid), ZGHA (Czech gray humic acid), ZBHA (Czech brown humic acid), CGHA (Chinese gray humic acid), CBHA (Chinese brown humic acid), BFA (Brazilian fulvic acid) and ZFA (Czech fulvic acid), WPFA (Waskish Peat reference fulvic acid) and SRFA (Suwannee River reference fulvic acid).

permitted the adequate classification of all samples according to their main characteristics (gray humic acids, brown humic acids, and fulvic acids).

The discriminant analysis (DA) manages to observe trends in the grouping of the humic fractions, as **Figure 3A** shows. The canonical discriminant functions for the samples explain a very high value of the cumulative variance (99.6% for the two first functions and 100% for the three first functions) and perfectly grouped the samples into four different trends: blanks, GHA, BHA, and FA (**Figure 3A** and **B**).

In principle, these results indicate that the mass spectra of the different family of HS samples contain specific *m/z* values that are valuable structural markers to discriminate among the diverse families of HS samples: GHA, BHA, and FA. If we analyze the *m/z* values (loadings) corresponding to each families (scores) in the graphical representation of the two PCA (**Figure 4**), we observe that, probably, the average molecular weight of these three families follows the same trend (GHA > BHA > FA), in agreement with the results obtained using other techniques (*4*). However, the order of magnitude of these values is rather low (100−1200 Da). Similar results were obtained using a mass instrument with a larger mass range (TOF Micromass, data not shown). The question is whether these *m/z* values correspond to the real value of the molecular weight of the molecules that form each humic system. In principle, three crucial issues are concerned with the answer of this question:

the possible presence of multicharged species, the existence of molecular fragmentation, and the possible partial molecular ionization.

Regarding the possible formation of multicharged species, although a number of authors have reported results indicating the possible presence of multicharge in HS mass spectra (*10–15*), other authors stated that this might be negligible (*16*). We have studied the formation of multicharged species in our mass spectra using the "Enhanced Multiply Charged" (EMC) mode (*17*). The EMC mode is basically an ion processing mode and works on the principle that, once ions have been trapped and cooled for a sufficient time, they have the same kinetic energy. After thermalization, the effective DC trapping barriers depend only on the charge state of the ions and not on their masses. Appropriate settings of the DC voltage and trap emptying time will result in preferential ion release, starting with ions with the lowest charge. We have applied this scan mode to the HS fractions (**Figure 5**). The result is the presence of multicharged ions throughout the entire spectrum. However, if we analyze the EMC/MCA intensity ratios (1.0e5/1.0e7), we can conclude that the formation of the multicharged species is negligible (approximately 1%). These results are in the line with those indicated by Sleighter and Hatcher (*16*).

Concerning the possible molecular fragmentation, the interpretation of the data is much more complicated. A number of authors indicate that a possible fragmentation of humic mol-
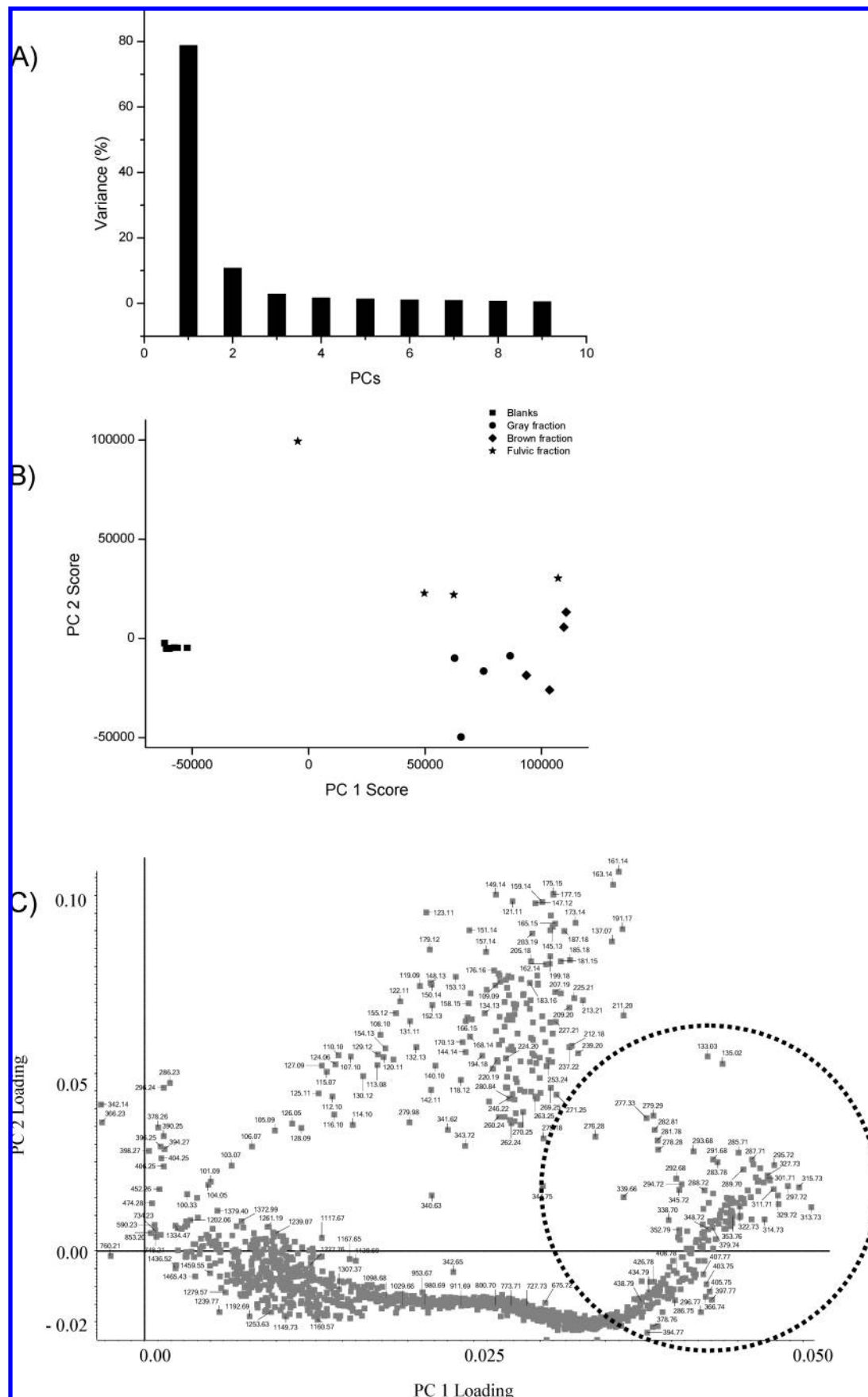
**Figure 2.** (**A**) Distribution of variance (%) on the PCs; (**B**) scores plot for mass spectra of samples (including blanks) projected for the two first principal components; (**C**) loadings plot for mass spectra of samples projected for the two first principal components. Peaks related to all HS fractions in circle.
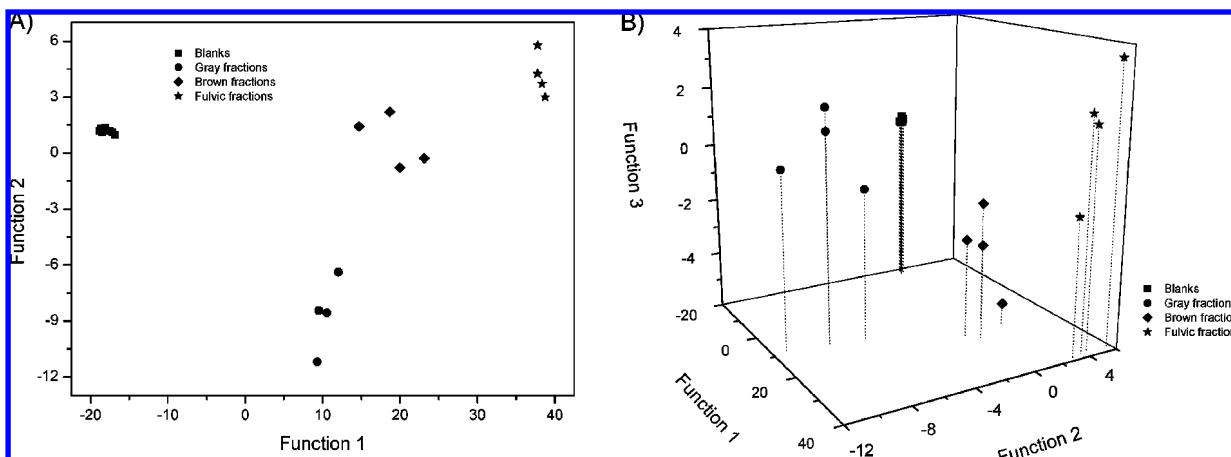
Classification of Main Humic Substances by MS

*J. Agric. Food Chem.,* Vol. 56, No. 14, 2008   **5485**



**Figure 3.** (**A**) Different trends in samples by discriminant analysis (projection on two axes); (**B**) 3D projection of different trends in samples by discriminant analysis.
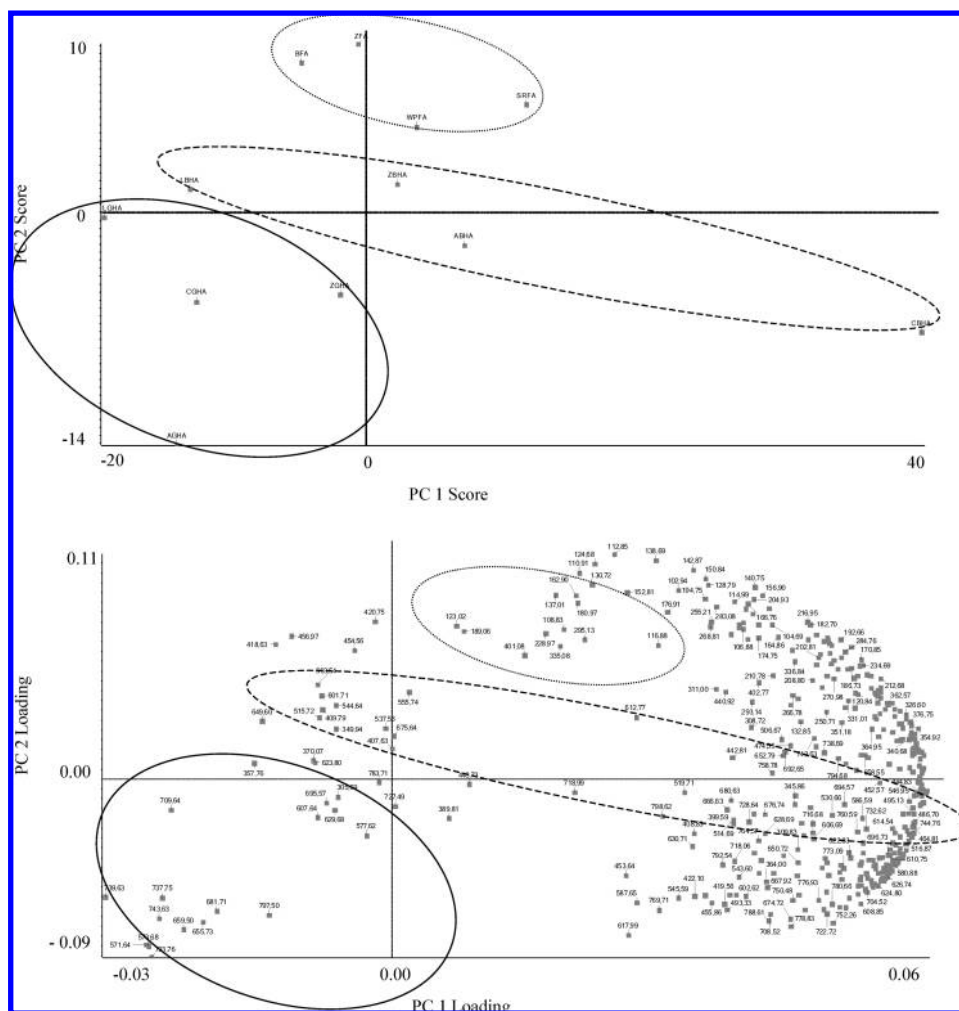


**Figure 4.** PC 1 and PC 2 scores and loadings for the different HS fractions. Related areas (different fractions with different *m/z* peaks) are in circles.

ecules cannot be ruled out (*18, 19*). Thus, studies using stable macromolecules, such as poly(acrylic) acids (PAA), poly(ethylene glycol) (PEG), or polystyrene sulfonic acid (PSSA), reported that these molecules underwent a significant fragmentation under the conditions of ionization used for humic samples (*13, 18, 19*). Other authors, however, stated that fragmentation did not occur when working with humic systems (*16*). This divergence could be associated with the different instrumental conditions used in these studies. In our case, we

have explored the possible presence of molecular fragmentation under the conditions of our instrument by using two stable macromolecules: PAA and PEG. As is shown in **Figure 6A** and **B**, both polymers underwent a significant and characteristic fragmentation associated to 72 amu losses corresponding with the monomer weight for the PAA standard and 44 amu losses corresponding with the monomer weight for the PEG standard. Similar results were obtained using another instrument with a larger mass range (TOF Micromass, data not shown). In
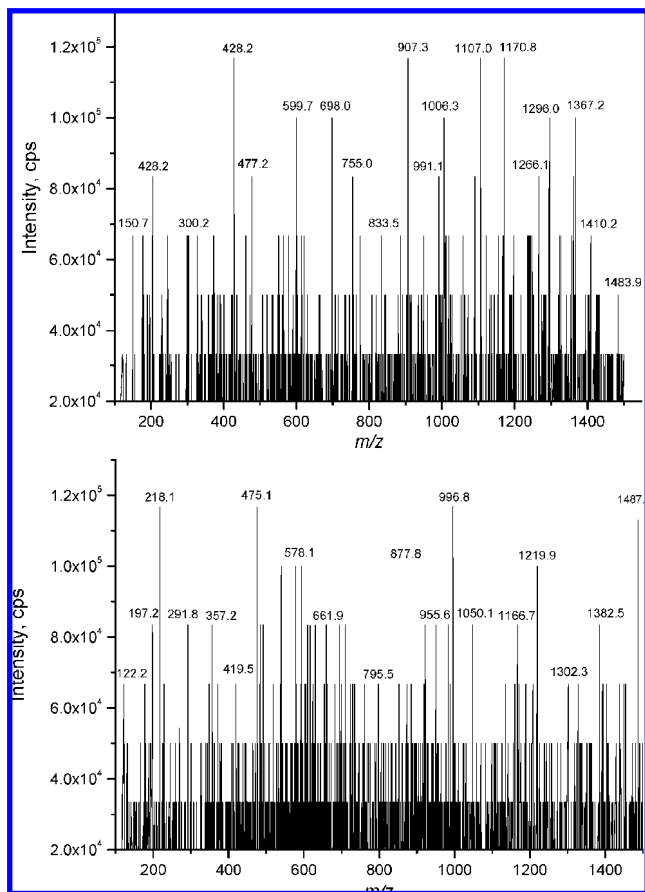
**5486** *J. Agric. Food Chem.,* Vol. 56, No. 14, 2008

Baigorri et al.



**Figure 5.** Multiple charged ions mass spectra of gray humic acid fraction (GHA) on average and brown humic acid fraction (BHA) on average.



**Figure 6.** (**A**) Polyacrylic Acid (PAA) mass spectra and fragmentation peaks series; (**B**) Poly(ethylene glycol) (PEG) mass spectra and fragmentation peaks series. Peaks series in the mass spectra from monomer losses of both polymers.

consequence, although we cannot ensure that humic samples underwent molecular fragmentation, it is clear that this process cannot be ruled out completely under the conditions of our analysis. Therefore, more studies are needed, maybe using alternative ionization methods, in order to elucidate if the distribution of *m/z* values obtained for HS mass spectra corresponds to the real mass distribution in humic systems.

Finally, our data did not permit us to explore whether a possible partial molecular ionization has occurred, although other studies reported that this possibility is also open for humic substances (*20, 21*). In our case, the mass spectra corresponding to AGHA or LGHA, with a strong aliphatic character (close to 80% (*4*)), yielded similar mass spectra to other humic samples with very different chemical structures (*4*). It is possible, therefore, that these mass spectra might be representative of only one fraction of the molecular distribution. In this way, other studies showed that fractions containing paraffin/olefin molecular types or nonpolar compounds could not be accessible by common soft-ionization methods such as ESI or Matrix Assisted Laser Desorption/Ionization (MALDI) (*22, 23*).

In consequence, our study indicates that the complementary use of PCA and DA for the analysis of mass spectra permits an adequate differentiation among the different main groups of humic substances. In principle, these results indicate that the mass spectra contain *m/z* peaks (or *m/z* peaks distributions) that are characteristic and singular of each group of humic substances, giving good information as a "fingerprint" not only according to their origin (from different source) but also to their different structural features.

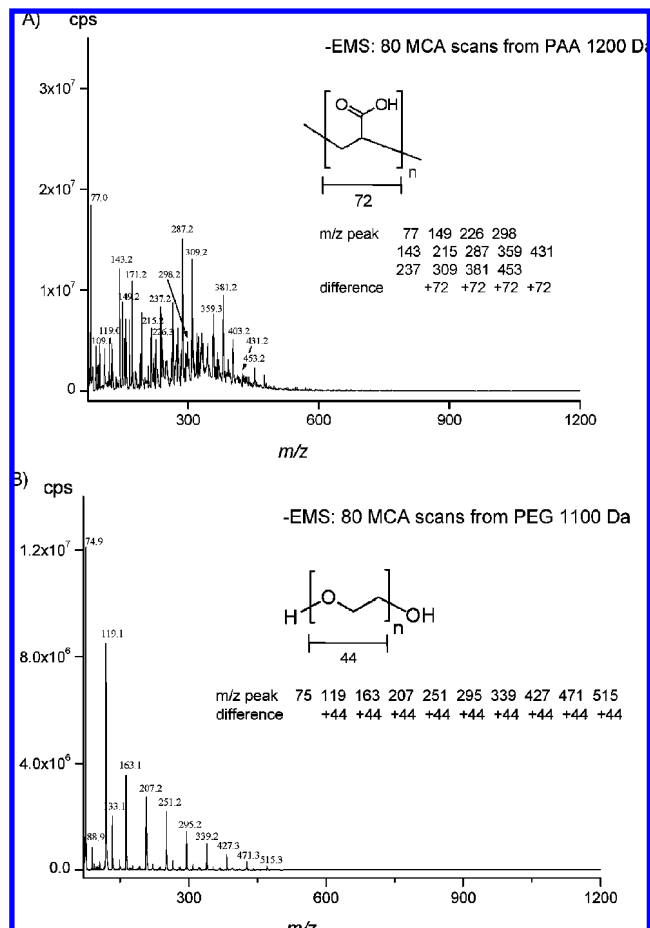Our results also indicate that although the *m/z* values principally corresponded to single-charged species, we cannot ascribe with total confidence these values to the real molecular weight of humic molecules, since sample molecular fragmentation as well as partial molecular ionization cannot be ruled out under the analytical conditions of our study.

## LITERATURE CITED

(1) Schmitt-Kopplin, P.; Kettrup, A. Capillary electrophoresis-electrospray ionization-mass spectrometry for the characterization of natural organic matter: an evaluation with free flow electrophoresis-off-line flow injection electrospray ionization-mass spectrometry. *Electrophoresis* **2003**, *24*, 3057–3066.

(2) Peña-Méndez, E. M.; Gajdošová, D.; Novotná, K.; Prošek, P.; Havel, J. Mass spectrometry of humic substances of different origin including those from Antarctica: A comparative study. *Talanta* **2005**, *67*, 880–890.

(3) Peña-Méndez, E. M.; Novotná, K.; Gajdošová, D.; González, V.; Havel, J. Characterization of humic substances of different origin by means of mass spectrometry and neural networks. *Chemosphere* **2007**, *68*, 2047–2053.

(4) Baigorri, R.; Fuentes, M.; González-Gaitano, G.; García-Mina, J. M. Simultaneous presence of diverse molecular patterns in

Classification of Main Humic Substances by MS

*J. Agric. Food Chem.,* Vol. 56, No. 14, 2008 **5487**

humic substances in solution. *J. Phys. Chem. B* **2007**, *111*, 10577–10582.

(5) MarkerView 1.1 User Manual; Applied Biosystems: Ontario, Canada, 2005.

(6) Novák, J.; Kozler, J.; Janoš, P.; Cežiková, J.; Tokarová, V.; Mandronová, L. Humic acids from coals of the Noth-Bohemian coal field I. Preparation and characterization. *React. Funct. Polym.* **2001**, *4*, 101–109.

(7) Stevenson, F. J. *Humus Chemistry: Genesis, Composition, Reactions*; John Wiley & Sons, Inc.: New York, 1994.

(8) Switf, R. S. Part 3. Chemical Methods. In *Methods of Soil Analysis*; Sparks, D. L., Ed.; Soil Science Society of America: Madison, WI, 1996.

(9) Swift, R. S. In *Humic Substances in Soil, Sediment and Water*; Aiken, G. R., McKnight, D. M., Wershaw, R. L., MacCarthy, P., Eds.; Wiley-Interscience: New York, 1985; pp 392−393.

(10) Fievre, A.; Solouki, T.; Marshall, A. G.; Cooper, W. T. High-resolution Fourier transform ion cyclotron resonance mass spectrometry of humic and fulvic acids by laser desorption/ionization and electrospray ionization. *Energy Fuels* **1997**, *11*, 554–560.

(11) Kujawinski, E. B.; Hatcher, P. G.; Freitas, M. A. High-Resolution Fourier transform ion cyclotron resonance mass spectrometry of humic and fulvic acids: improvements and comparisons. *Anal. Chem.* **2002**, *74*, 413–419.

(12) Stenson, A. C.; Marshall, A. G.; Cooper, W. T. Exact masses and chemical formulas of individual Suwannee river fulvic acids from ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **2003**, *75*, 1275–1284.

(13) Schmitt-Kopplin, P.; Kettrup, A. Capillary electrophoresis-electrospray ionization-mass spectrometry for the characterization of natural organic matter: an evaluation with free flow electrophoresis-off-line flow injection electrospray ionization-mass spectrometry. *Electrophoresis* **2003**, *24*, 3057–3066.

(14) Pfeifer, T.; Klaus, U.; Hoffmann, R.; Spiteller, M. Characterization of humic substances using atmospheric pressure chemical ionization and electrospray ionization mass spectrometry combined with size-exclusion chromatography. *J. Chromatogr., A* **2001**, *926*, 151–159.

(15) Kim, S.; Kramer, R. W.; Hatcher, P. G. Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the Van Krevelen diagram. *Anal. Chem.* **2003**, *75*, 5336–5344.

(16) Sleighter, R. L.; Hatcher, P. G. The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter. *J. Mass Spectrom.* **2007**, *42*, 559–574.

(17) Hopfgartner, G.; Varesio, E.; Tschäppät, V.; Grivet, C.; Bourgogne, E.; Leuthold, L. A. Triple quadrupole linear ion trap mass spectrometer for the analysis of small molecules and macromolecules. *J. Mass Spectrom.* **2004**, *39*, 845–855.

(18) Stenson, A. C.; Landing, W. M.; Marshall, A. G.; Cooper, W. T. Ionization and fragmentation of humic substances in electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **2002**, *74*, 4397–4409.

(19) Leenheer, J. A.; Rostad, C. E.; Gates, P. M.; Furlong, E. T.; Ferrer, I. Molecular resolution and fragmentation of fulvic acid by electrospray ionization/multistage tandem mass spectrometry. *Anal. Chem.* **2001**, *73*, 1461–1471.

(20) Reemtsma, T.; These, A. On-line coupling of size-exclusion chromatography to electrospray ionization tandem mass spectrometry for the analysis of aquatic fulvic and humic acids. *Anal. Chem.* **2003**, *75*, 1500–1507.

(21) These, A.; Reemtsma, T. Limitations of electrospray ionization of fulvic and humic acids as visible from size exclusion chromatography with organic carbon and mass spectrometric detection. *Anal. Chem.* **2003**, *75*, 6275–6281.

(22) Schaub, T. M.; Hendrickson, C. L.; Qian, K.; Quinn, J. P.; Marshall, A. G. High-resolution field desorption/ionization Fourier transform ion cyclotron resonance mass analysis of non-polar molecules. *Anal. Chem.* **2003**, *75*, 2172–2176.

(23) Linden, H. B. Liquid injection field desorption ionization: a new tool for soft ionization of samples including air sensitive catalysts and non-polar hydrocarbons. *Eur. J. Mass Spectrom.* **2004**, *10*, 459–468.