Review

Molecular geochemistry of soil organic matter by pyrolysis gas chromatography/mass spectrometry (GC/MS) technique: A review

Mojtaba Mehrabanian

Department of Soil Science, Faculty of Agricultural Engineering and Technology, Collage of Agriculture and Natural Resources, University of Tehran, Iran.

Accepted 18 March, 2013

Soil organic matter (SOM) is arguably the most complex and least understood component of soils. To analyze SOM structure and reduce SOM large structural heterogeneity, different components of SOM need to be separated into entities that differ in terms of source, composition, and turnover. Pyrolysis-gas chromatography/mass spectrometry (Pyrolysis-GC/MS) is to present the most effectual approach in the study of the pedogenesis of SOM; this uses thermal degradation to cleave bonds in the organic macromolecules and enables a sensitive and rapid characterization of organic constituents. Pyrolysis-GC/MS also aids the molecular characterization of microbial and plant-derived biomass and generates valuable data on the degradation and conservation rates of organic debris. This review provides an overview of commonly measured soil organic chemical constituents identified by Pyrolysis-GC/MS, including (i) aliphatics, fatty acids and sterols, (ii) carbohydrates, (iii) lignin, (iv) aromatic compounds and polycyclic aromatic hydrocarbons (PAHs) and (v) N-containing compounds.

Key words: Soil organic matter (SOM), pyrolysis-gas chromatography/mass spectrometry (GC/MS), molecular geochemistry, organic compounds.

INTRODUCTION

Soil organic matter (SOM) occupies a fundamental role in ecological dynamics. It influences a large group of soil chemical, physical and biological attributes, including soil structure, soil fertility and productivity, soil moisture retention and erodability (Kay et al., 1997). These qualities have a tremendous impact on ecosystem roles, such as nutrient cycling, carbon sequestration and pollutant retention. In fact, SOM composition and its changes during biotic and abiotic transformation are key variables in the processes and the size of soil interactive reactions with the surroundings.

SOM is a combination of all organic materials derived from plants and animals, in different stages of decomposition and degree of association with the mineral matrix (Buurman and Roscoe, 2011). This implies large structural heterogeneity and close linkage to SOM functions. SOM structure analysis helps us to understand chemical processes within soil and to predict changes in the sign and magnitude of terrestrial carbon fluxes in a changing environment. To reduce SOM large structural heterogeneity, different components of SOM need to be separated into entities that differ in terms of source, composition, and turnover. Recent evidence suggests that the classical chemical fractionation into fulvic acids, humic acids, and humin according to solubility characteristics in dilute acid and base (Swift, 1996) is not
useful in this respect. Fractionations that are more promising rely mostly on physical fractionations according to particle size or density (Golchin et al., 1994), and the analysis of dissolved organic matter in the soil solution (Guggenberger et al., 1994), primarily to determine labile fraction of carbon.

In order to expand deep appreciation of the origin and function of organic matter, development of modern fractionation methods is vital. The objectives of this review are to present an account of what has been found about SOM characterization by Pyrolysis-gas chromatography/mass spectrometry (GC/MS) technique. The first part of this review will deal with the strengths of this approach, while the second part will focus on more recent data on the identities of constituent compounds of SOM regarding molecular geochemistry.

**PYROLYSIS-GC/MS\(^1\) TECHNIQUE AND ITS STRENGTHS**

The modern methods used to fractionate SOM into a small number of fractions can be divided into degradative methods involving chemolysis, thermolysis, or thermochemolysis, and non-invasive spectroscopic techniques such as nuclear magnetic resonance (NMR) spectroscopy. Degradative methods provide molecular-level information on specific organic compounds accessible to the degradative step, whereas spectroscopic techniques enlighten about the bulk composition of SOM (Guggenberger, 2006).

In thermolysis, a soil sample is heated according to a heating program, usually with a constant heating rate. A distribution, known as a thermogram, can be produced by measuring the weight loss as a function of temperature. The gases evolved can also be analyzed to determine the amount of carbon or mass lost at different temperatures by different kinds of spectroscopic methods. The thermograms often exhibit different peaks for labile and more resistant SOM (Bruun et al., 2010).

Pyrolysis, coined from the Greek-derived elements pyr "fire" and lysis "separating", along with GC/MS is an effective thermal technique in the study of pedogenetic processes (Abelenda et al., 2011), which uses thermal degradation to cleave bonds in the organic macromolecules. When the heat energy applied to the macromolecule is greater than the energy of specific bonds, those bonds will dissociate in a predictable and reproducible way. The smaller molecules generated by this bond-breaking are identified by the analytical tool and therefore bring an understanding of the original macromolecule and enables a sensitive and rapid characterization of organic constituents.

Pyrolysis-GC/MS is an instrumental technique that enables a reproducible characterization of the intractable and involatile macromolecular complexes found in virtually all materials in the natural ecosystem. It differs from GC/MS in the type of sample analyzed and the method by which it is introduced to the GC/MS system. Instead of the direct injection of a highly refined organic solution, a few amount of the original natural material (for example, soil, sediment etc.) is analyzed directly. The analysis is regularly preceded by an extraction of the sample with an organic solvent to remove any free, unbound components with a low molecular mass that would otherwise obscure analytical data pertaining to the high molecular mass components of interest. Extracted samples are then inserted into a quartz chamber in a pyrolysis unit that is then heated resistively in an oxygen free environment at a preset temperature for a number of seconds. This results in a heat mediated cleavage of chemical bonds within the macromolecular structures of interest producing a suite of low molecular weight chemical moieties of the composition of which is indicative of specific types of macromolecule (for example, lignin, cellulose, chitin etc.). This mixture of compounds is then swept onto the analytical column of the GC and GC/MS proceeds as normal.

In pyrolysis-gas chromatography (PyGC), the fragments generated by pyrolysis are passed through the GC for separation and identification. Frequently, the major peaks in the resulting chromatogram (pyrogram) are easily identifiable and give direct structural information about the material being pyrolyzed. Many polymers, polystyrene and polymethyl methacrylate for example, generate significant quantities of the monomer used in producing the polymer. For other materials, the pyrograms are more complex and serve as "fingerprints" which may be used to distinguish related materials for identification or for quality control.

Most methods used to identify or quantify individual organic compounds require the target chemical be extracted from a solid or liquid matrix. This is often done using a liquid or supercritical fluid extraction. Solvents, particularly basic solutions, can partially oxidize, or otherwise modify the organic matter being studied. In addition, organic molecules can only be identified by GC/MS if they remain in an inert gas stream at 300°C or less. Most organic matrices in the environment are composed of material too large to volatilize at 300°C and cannot be analyzed by GC/MS. Pyrolysis will thermally extract intact molecules or crack large molecules into fragments that can be separated and identified by GC/MS. As such, pyrolysis is an alternative way to "extract" organic matter from complex matrices (White et al., 2004). Pyrolysis is not without its limitations, however. Thermal secondary reaction causes considerable modification of the organic compound, which may bias the interpretation of the pyrolysis products with respect to their mother compounds (Saiz-Jimenez, 1994).

In analytical pyrolysis, low molecular weight compounds and some of the medium molecular weight compounds could be detected at moderate temperatures;

---

\(^1\) Pyrolysis-Gas Chromatography/Mass Spectrometry
however, non-volatile compounds of low thermal stability, in particular organic compounds of high molecular weight, could be analyzed at higher temperatures of above 500 °C (Cabane et al., 2001). Pyrolysis volatile products are chromatographically separated (GC) and identified (MS) based on their ratios of mass to charge (m/z).

The thermal energy required to volatilize a compound indicates bond strength. Thermograms of volatilization, therefore, give useful information on the strengths of chemical bonds within the organic macromolecules, or between organic materials and minerals (Balser, 2005). Shifts in such thermograms to higher pyrolysis temperatures indicate stronger bonds (Schulten and Leinweber, 1999).

MAJOR COMPOUNDS IDENTIFIED BY PYROLYSIS-GC/MS

Analysis of the gaseous mixtures that are produced from the pyrolysis of a given organic compound provides a “Fingerprint” [nature and abundance (GC-MS) of released gases], which can be compared to laboratory data bases. Thus, from this fingerprint, we can identify the complex molecules which constitute the pyrolysed samples. Hence, pyrolysis-GC/MS aids the molecular characterization of microbial and plant-derived biomass, and generate valuable data on the degradation and conservation rates of organic debris. Furthermore, MS of the pyrolysis products can be used to make thermograms of compound classes such as carbohydrates, lignin monomers, lignin dimers, alkyl aromatics, lipids and N-containing compounds, thus, giving rise to a different distribution for each compound class (Bruun et al., 2010).

Based on chemical similarities, the following group of compounds identified by pyrolysis-GC/MS are reviewed in the present paper, including (i) aliphatics, fatty acids and sterols, (ii) carbohydrates, (iii) lignin, (iv) aromatic compounds and polycyclic aromatic hydrocarbons (PAHs) and (v) N-containing compounds.

Aliphatics, fatty acids and sterols

Aliphatics mainly originate from plant waxes and suberin present in tree roots and bark (Lorenz et al., 2007). Aliphatic compounds refer to n-alkanes, n-alkenes, n-methylketones and n-alcohols. The long-chain n-alkanes and n-alkenes are common in fresh litter and in soils with low microbial activity (Abelenda et al., 2011). Short-chain aliphatics can be attributed to microbial products. They are either fragments of microbial lipids, or parts of longer chains, degraded by microorganisms. In addition, long-chain odd C-numbered methylketones probably come from microbially induced oxidation of alkanes (Buurman et al., 2007b).

Fatty acids characterized by pyrolysis are mainly derived from plants, but they can also be of microbial origin (Abelenda et al., 2011). Although C12 fatty acid is probably microbial, fatty acids from plants have a predominance of even over odd chain lengths (Buurman et al., 2007c). The C16 and C18 fatty acids are derived from waxes or lipid biomacromolecules such as cutan and cutin from higher plants. The iso- and anteiso- odd C-numbered fatty acids (mainly C15 and C17), commonly considered to be bacterial biomarkers (Buurman et al., 2007b; Buurman et al., 2007c).

Sterols are part of cell membranes and are generally easily decomposed in aerobic environments. Plant sterols are structures with four rings largely consisting of 27 to 30 C atoms, while microbial sterols (hopanoids) have five rings and up to 35 C atoms. Their identification with pyrolysis-GC/MS is problematic, although quite often a group can be identified (Abelenda et al., 2011).

Carbohydrates

Levoglucosan (anhydroglucosan), cyclopentenone, furan and furaldehyde are some identifiable carbohydrate markers by pyrolysis-GC/MS. Although pyrolysis products of carbohydrates can also have a plant or microbial origin, in most soils polysaccharide products are only major compounds in pyrolysates of fresh plant material. Levoglucosan and cyclopentenone are largely associated with virtually fresh plant-derived matter, especially when found in combination with C3 lignin and small contributions of N-compounds. On the other hand, the occurrence of furans, furaldehydes and anhydrosugars in many soils suggests that these hardly biodegraded polysaccharides originated from residual plant polysaccharides and newly formed microbial polysaccharides, especially when they dominate over levosugars in a given pyrolysate (Buurman et al., 2007b; Abelenda et al., 2011). Kaal et al. (2007) interpreted low furans to levoglucosans ratio as indicating that the majority of carbohydrates is expected to be well preserved.

Lignin

Groups that mainly represent primary SOM are the aliphatics, fatty acids, ketones and lignins. Lignin is commonly derived from wood and also is the second most abundant component of cellular walls after polysaccharides. Within the lignin group, microbial decay preferentially reduces the contribution of syringols. Syringyl-lignin is known to be less stable than guaiacyl-lignin and there is a relative accumulation of guaiacol fragments (Buurman et al., 2007c).
Lignin degradation includes oxidation of the C3 side-chain, which is recognized in pyrolysates. Consequently, the relative proportion of C3-guaiacols to total guaiacols (C3G/GT) and C3-syringols to total syringols (C3S/ST) is commonly utilized as an index of degradation. In pyrolysates, the vinyl-containing phenols and methoxyphenols are indicative of fresh lignin (Abelenda et al., 2011). In addition, the pattern of lignin monomers can be used to identify the primary resource of SOM; in contrast to angiosperm lignin, gymnosperm lignin does not contain syringyl units (Guggenberger, 2006).

Phenols originated largely from lignin, perhaps with a small contribution from carbohydrate-derived phenols (Calvelo Pereira et al., 2011). Pyridines can also be formed by microbial decomposition of plant lignins and other phenolics in the presence of NH₃ (Buurman et al., 2007b).

**Aromatic compounds and polycyclic aromatic compounds**

Aromatic (benzene, toluene, alkylbenzenes, methylbenzenes, dimethylbenzenes, ethylbenzenes and styrenes) and PAH pyrolysis products may be ascribed to a defined origin only in conjunction with other compounds. Thus, aromatic pyrolysates are ascribed to proteins, and specifically to microbial material, when pyridine and toluene are found simultaneously; while a combination of benzene, toluene and PAHs indicates the presence of charred material. PAHs can usually be ascribed to aromatization and condensation reactions upon burning (Abelenda et al., 2011). In other words, the presence of PAHs in SOM pyrolysates is usually considered as testimony of burning (Naafs, 2004; Ross et al., 2005; Rumpel et al., 2007), although small amounts of these may be secondary pyrolysis products of fatty acids (Saiz-Jimenez, 1995).

A significant proportion of the benzenes and PAHs might also originate from charred carbohydrates (Calvelo Pereira et al., 2011). A-type humic acid has a relatively large C-content and large aromatic and carboxyl contents. There is now increasing evidence that much of this aromaticity is due to accumulation of charred plant remains from periodic burning of the vegetation.

Buurman et al. (2007b) conjectured that the naphthalenes Pa4-6 may therefore be pyrolysis products of microbial organic matter. Additionally, series of n-alkylbenzenes (C₃-C₁₇) probably result from incomplete combustion upon burning (Abelenda et al., 2011).

Alkylaromatics are considered a result of the humification of primary plant materials and microbial metabolites to form the skeletons of humic substances (Schulten and Leinweber, 1999). Pyrolysis of fatty acids with one double bond results in traces of toluene and ethylbenzene and such unsaturated fatty acids could be the origin in soil of hydroaromatic and alkylaromatic structures (Buurman et al., 2007b).

Sollins et al. (1996) mentioned similar related behavior in alkyls and aromatics regarding SOM-decomposition. Both fractions are considered relatively recalcitrant to microbial decomposition. Both fractions are considered relatively recalcitrant to microbial biomass origins.

**N-containing compounds**

The study of N-containing compounds comprises the identification of specific protein sources as well as the discrimination of N-compounds of either plants or microbial biomass origins.

Although N-compounds (pyrrole, pyridine, benzonitrile, acetobenzonitrile, indole, methylindole and diketodipyrrole) are, in general, most probably driven from amino acid and amino sugar moieties, there is insufficient knowledge of their more specific sources. Furthermore, whether these moieties are still present as intact polypeptides or as modified entities is as yet unknown. Additionally, the presence of rather unusual isomers of some nitrogen-containing products could imply that these are derived from chemically modified biomolecules formed in the soil (Van Bergen et al., 1998).

A large number of N-compounds are from microbial SOM, but others can be both of microbial and plant origin. Large amounts of N-compounds accompanied by carbohydrate pyrolysis products other than levosugars point toward the presence of large amounts of microbial SOM (Abelenda et al., 2011).

According to Schulten and Schnitzer (1998), heterocyclic N-compounds such as pyrroles, pyrrolidines, pyridines, pyranes, and pyrazoles are significant components of the SOM. The pyridines and pyroles are probably of microbial origin, while the indoles can be predominantly plant derived (Buurman et al., 2007c). Indeed, pyridines can be formed by microbial decomposition of plant lignins and other phenolics in the presence of NH₃. Pyrroles may derive from amino acids such as proline and hydroxyproline. Substituted pyrroles are formed readily upon pyrolysis of porphyrin, which is an essential component of the chlorophyll molecule in terrestrial plants. Indoles are pyrolysis products of tryptophane, an amino acid that occurs in proteins. Furthermore, imidazoles are pyrolysis products of histidine, an amino acid that is essential for tissue growth and repair, but also thermal degradation of grass and soil microorganisms forms imidazoles. Benzonitrile is derived from aromatic amines as well (Buurman et al., 2007b). The work by Buurman and Roscoe (2011) suggests that benzonitrile is presumably mainly microbial.

Acetamides are probably pyrolysis products of chitin. In fact, chitin is a polymer of N-acetylglucosamine that is relatively easy to degrade (Buurman et al., 2007b), and its pyrolysis products are not commonly found in soils. Nonetheless, a large number of chitin pyrolysis products

---

1 An N-containing polysaccharide
were obtained from the allophanic soils of the Azores (Buurman and Nierop., 2007a) and Costa Rica (Buurman et al., 2007b).

This implies that in these soils a significant part of the SOM could derive from fungi and from mesofauna such as arthropods and nematodes. In other words, it appears that decay of plant-derived organics is extremely efficient and that microbially and mesofauna-derived moieties represent a large fraction of the SOM.

Diketodipyrrole may be derived from both plant and microbial sources. Several compounds with nitrile functional groups have been also detected. The origin of these compounds remains unknown, but they may be related to charring (Abelenda et al., 2011).

According to Buurman et al. (2007c), 3-methylpyridine, benzeneacetonitrile and indole were associated with fresh litter, while other N-compounds were more associated with strongly degraded SOM, that is, with microbial activity. Nitrogen compounds showed a clear loss of pyridine compounds, benzonitrile, and diketodipyrrole, while a cluster of N-compounds increased. It is likely that the shift indicates a change from predominantly plant-derived nitrogen compounds towards microbial ones.

Low molecular weight amino acids split into small and volatile compounds that generally elute in the first large bulk of the pyrogram. These substances are poorly identifiable and quantifiable (Kaal et al., 2007).

CONCLUSION

In order to achieve a profound understanding of complex organic matrices, applying Pyr-GC/MS technique as an innovative approach in environmental science and engineering is inevitable and indispensable. The superiority of this technique is its capability to help interpret the chemical composition of SOM in terms of sources, decomposition and stability. Consequently, SOM fingerprinting by means of Pyr-GC/MS technique is indubitably the most effective method to measure SOM chemistry.

The challenge of this technique until recently has been to identify the precursor molecules of the original SOM, which clarifies the limited use of Pyr-GC/MS technique in soil science. However, the fresh rapid progression of the identification of these compounds decipher the genesis of SOM.

Moreover, a sophisticated understanding of SOM chemistry and dynamics will help better modelling of greenhouse gases emissions such as methane and carbon dioxide.

Ultimately, analytical pyrolysis has major limitations that could produce misleading results if disregarded (for example, the effect of the mineral matrix). Nevertheless, within its limitations, analytical pyrolysis has numerous applications and is a useful and valuable tool.

RECOMMENDATIONS

Despite the mentioned uses of Pyr-GC/MS technique, it seems this technique has more relationships with other aspects of soil science. Some recommendations, therefore, are proposed to be the main subject of future studies in this regard.

1) Characterize the type of SOM by pyrolysis-GC/MS that promotes microbial activity and therefore soil denitrification
2) Investigate the effect of SOM on nitrogen cycling especially on N2O emissions
3) Using the derived data in process-based models such as Daycent to simulate N2O fluxes with regard to the nature of SOM

REFERENCES


