Determination of Biomarkers in Petroleum by Multidimensional Gas Chromatography: Fundamentals, Applications, and Future Perspectives

Noroska Gabriela Salazar Mogollón, Paloma Santana Prata, Fabio Augusto.*
Institute of Chemistry, University of Campinas (IQ/UNICAMP).
Instituto Nacional de Ciência e Tecnologia em Bioanalítica (INCTBio).

**Corresponding author: Dr Fabio Augusto,*
Instituto de Química, Universidade Estadual de Campinas, Caixa Postal 6154, 13084-971 Campinas, SP, Brazil
Phone: +55 19 3521-3057 FAX: +55 19 3521-3023 E-mail address: augusto@iqm.unicamp.br

Several improvements were observed in the analytical methods employed for chemical characterisation and determination of biomarkers in petroleum. Emphasis has been placed on sample preparation and instrumental analysis. Comprehensive two-dimensional gas chromatography (GC×GC) has allowed better characterisation of potential biomarkers by dampening co-elution and increasing the signal-to-noise ratio during chromatographic analyses, which has lead to the acquisition of more accurate and reliable mass spectra. As a consequence, reliable biomarkers are now available to ascertain the thermal maturity, extent of biodegradation, evaluation of the oil's migration, and age of the source rock surrounding the petroleum. This review addresses some fundamentals, advances, and future perspectives of gas chromatographic techniques in petroleomics.

1. Introduction

Petroleum or crude oil is primarily a fossil fuel and a non-renewable source of energy. The economic importance of crude oil alongside with the need to perform technological improvements related to the exploration and production of petroleum have stimulated the assessment of the origin of petroleum, which controls its physical properties and chemical composition and, therefore, determines its quality. Crude oils are naturally occurring complex mixtures comprised of more than 20,000 compounds with distinct elemental compositions that exhibit a broad structural and chemical diversity. These oils are predominantly comprised of aliphatic hydrocarbons, aromatics, and many other constituents that contain heteroatoms (N, O, S) with different volatilities [1].

Biological markers (i.e, biomarkers) or chemical fossils are used to ascertain the maturation of crude oils and the extent of their biodegradation [1,2]. The most important biomarkers known today are, but not exclusively, hydrocarbons [1,3]. For instance, biodegraded oils possess little to no aliphatic hydrocarbons in its composition [1]. These compounds are originally found in the organic matter that yields petroleum and are able to withstand the process of biodegradation over the course of its production [1-4]. The resistance to biodegradation of these biomarkers depends largely on its chemical structure. Hence, determination of the biomarkers is used to extract information regarding the organic matter in the source rock that originated the crude oil and the conditions during its deposition [1,4-6]. So, chemical fingerprinting of petroleum is commonly used to ascertain the thermal maturity, extent of biodegradation, evaluation of the oil's migration, and age of the source rock surrounding the petroleum [1-6].

The biomarkers, namely, isoprenoids, terpanes, and steranes are used for the determination of the geochemical parameters mentioned above [1,4-6]. However, chemical characterisation of petroleum remains an extremely challenging task to analytical scientists due to the exceedingly complex nature of the sample: (i) very large number of constituents, (ii) broad structural diversity, (iii) many compounds are isobaric, and (iv) often times important markers are found in trace concentrations. In order to address these shortcomings, chromatographic techniques coupled to mass spectrometry are frequently used to generate accurate analytical profiles for chemical fingerprinting of petroleum. However, the use of conventional chromatographic techniques does not exhibit the required peak capacity (or the number of theoretical plates required) necessary to fully resolve all analytes found in crude oil samples [1,7-8]. Also, even if a highly efficiency GC column was employed for such analyses, it would not display the necessary selectivity/solvation capabilities required to resolve all of the analytes. The latter is crucial in petrochemical studies, where many compounds may have the same boiling point, but have other different physiochemical properties; being an alternative, the introduction of new parameters (selectivities, or polarity index) for the separation of these compounds that suffer from coelution in one-dimensional GC.

2. Conventional Gas Chromatography

The analysis of biomarkers is commonly performed by gas chromatography coupled with mass spectrometry (GC-MS) employing highly efficient and thermally stable open tubular capillary columns (e.g., the ASTM method) [1,9]; however, under practical experimental conditions, GC-MS offers a resolving power far smaller than the required to adequately resolve the analytes that can co-eluting, such as, tri- and pentacyclic terpanes, or that eluete very close, as the

isoprenoids mono- di- and tri aromatics, tri- and pentacyclic terpanes, steranes and aromatic steranes), in these cases, it is often necessary the introduction of an additional procedure during sample preparation. This step consists of the oil's fractionation into saturated hydrocarbons, aromatics, resins, and asphaltenes (i.e., SARA separation) [7,8]. The biomarkers of interest are mostly found in the saturated fraction and, therefore, only the fraction of interest is subjected to GC-MS analysis [1]. During GC-MS analysis, the detection of the analytes is performed under full scan (for untargeted analysis) or single ion monitoring (SIM) mode (for targeted analysis). In either detection mode, under carefully devised experiments, the required qualitative information is obtained and used to provide the necessary information for chemical characterisation of the sample. Nonetheless, the results obtained by these methods are potentially jeopardised by co-elution. For instance, an ion with *m/z* 217 corresponds to a key fragment pertaining to steranes by electron ionisation. However, other molecules may exhibit this ion into its fractionation and co-elution with interfering molecules that exhibit similar fragmentation patterns (i.e., also exhibit *m/z* 217) will lead to erroneous interpretation of the chromatographic data [7,8,10]. If prior knowledge of the biomarkers is available [11], additional selectivity is obtained by exploring gas chromatography coupled to a tandem mass spectrometer (GC-MS^n). For instance, multiple reaction monitoring (MRM) experiments are used exclusively for targeted analysis and/or structural elucidation [1,12,13].

3. Multidimensional Gas Chromatography (MDGC)

Multidimensional separation techniques are powerful methods in which two or more independent separative techniques are linked together for separation [14]. Multidimensional gas chromatographic (MDGC) techniques comprise of two or more independent gas chromatographic separations (i.e, dimensions) coupled in a sequential fashion [15,16]. The paramount requirement to achieve higher peak capacity, (maximum number of peaks separable), is that all GC stationary phases examined must possess distinct/complementary solvation capabilities and, thus, different selectivities [17]. Also, MDGC experiments must be devised to maintain, at least in part, the separation achieved in each dimension so that the resolving power of the composite separation exceeds that of the individual stages [16,17]. While this section is largely confined to the applications of MDGC techniques to the chemical characterisation of crude oils, several excellent reviews describing the fundamentals of MDGC are available for the interested reader [16,18,19-23].

The use of MDGC techniques for the analysis of biomarkers in petrochemical samples has yielded promising results [7,8,11,12]. These techniques can be divided into non-comprehensive and comprehensive methods. The former are heart-cutting MDGC techniques that employ valve-based or flow-controlled microfluidic devices, being the latter collectively known as Deans switch interfaces [19,20]. In these experiments, selected heart-cuts (i.e., fractions of the eluent of a previous dimension) are sampled and transferred to another GC column for additional separation. Therefore, this system allow that only a fraction of the entire sample experiences two or more GC separations, independent of the dimensions of the second column and the sample rate. Although these methods exhibit optimum analytical performance for targeted analysis, the overall increase in peak capacity is modest compared to comprehensive setups [24], which are recommended for untargeted analysis or discovery-based approaches.

Comprehensive MDGC techniques are those in which the whole sample, or a representative fraction of the sample, experience two or more GC separations [24-26]. This is accomplished by means of an interface that continuously samples and periodically transfers the chromatographic band to another dimension for additional separation. These interfaces can be: (a) valve-based, where the focusing of the analytes is effectuated by very rapidly sampling only very narrow fractions from the first column, however, about 90% of the effluent is vented off, so that the far less then 10% of the original injected amount of the sample is transferred to the second column, for this reason, this modulator can does not provide a comprehensive separation, being the type of modulator less used among all [14], (b) flow-controlled, this interface also uses pneumatic valves, here the effluent from the first dimension is accumulated in a loop and then be transferred to the second dimension, however, a loss of considerable percentage of effluent from the first column can take place, in addition this device is subject to providing band broadening, however, new developments are using ultra rapid flow in the second dimension in order to avoid the loss of analyte transferred. [14,23] or (c) thermal modulators, in which accumulated of the analytes in a thick film capillary (modulation capillary) by thermal means (i.e. liquid nitrogen, carbon dioxide or dry ice) and its remobilisation it performed by applicating heat, resulting in narrows peaks and a increase of the peak amplitude and consequently are the type modulator more widely used [14,25]. They operated at sufficiently high sampling frequencies to preserve the separation achieved in the previous dimension.

Is worth noting that, if the peak capacity in the first and second column are, respectively, n_1 and n_2 peaks, the total peak capacity in a GC-GC system will increase arithmetically for n_1+n_2 peaks (remembering that, only a fraction of the sample of the first dimension is transferred to the second dimension), while that in a system GC×GC, where the sample is subjected to separation by two columns of different selectivities, the peak of capacity will increase geometrically for $n_1 \times n_2$, for this reason, the theoretical increase in peak capacity is far greater than those exhibited by heart-cutting techniques [15]. Introduced in 1991 by late Prof. John Phillips [27], comprehensive two-dimensional gas chromatography (GC×GC) is considered by many separation scientists as the biggest milestone in chromatography since the development of GC capillary columns in the late 50's. One of the most important hallmarks of GC×GC is the presence of structured chromatograms where a relationship between the analytes molecular structure and retention coordinates is readily observable. For instance, paraffins, tricyclic terpanes, steranes, and hopanes elute in well defined regions of the chromatogram (see Figure 1) when a saturated fraction of a Brazilian oil sample is analysed by GC×GC coupled to a mass spectrometer with a time

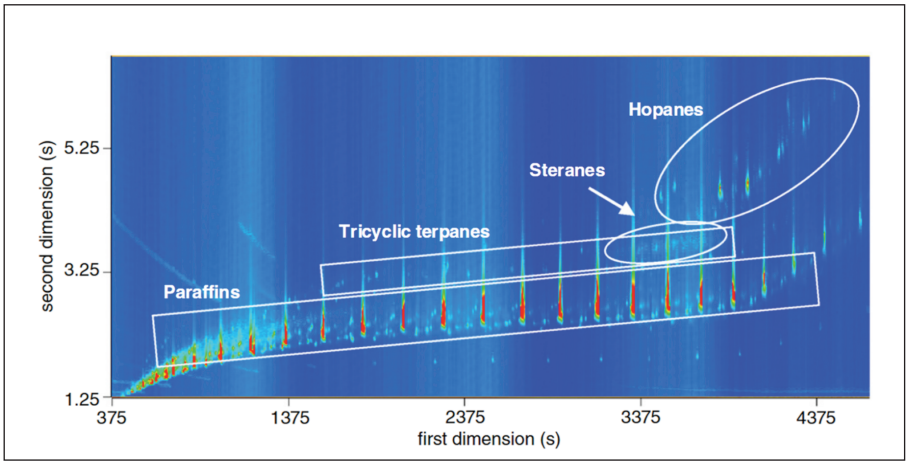

Figure 1. GC×GC-TOFMS chromatogram from the saturated fraction of oil (Camamu-Almada basin, Brazil). ^{1}D column: 30 m × 0.25 mm HP-5MS (polydimethyldiphenylsiloxane with 5% diphenylsiloxane monomer incorporation) (d_f = 0.25 µm). ^{2}D column: 150 cm × 0.10 mm DB-17 (polydimethyldiphenylsiloxane with 50% diphenylsiloxane monomer incorporation) (d_f = 0.10 µm) adapted from Aguiar et al. [10].

of flight mass analyser (GC×GC-TOFMS). Also, the use of MS offers an additional dimension for separation and ascertains qualitative analysis [14,24,28].

3.1. Biomarker Determination by Comprehensive Two-Dimensional Gas Chromatography

In conventional GC-MS analyses of crude oils, the chromatograms exhibit a pronounced "hump" in the detector's baseline that is comprised of an overwhelming number of co-eluting analytes; hence, samples that exhibit this chromatographic profile are known as the unresolved complex mixture (UCM), [12]. The co-elution observed in the UCM is so severe because of the high degree of isomerisation of these compounds that a characterisation by ion monitoring becomes impossible, in this manner, the amount of the information related to geochemical evolution, migration of oils, and extent of biodegradation severely decreases; all of which can provide critical information to understand the factors that regulate the presence of petroleum in surface environments. From the technological standpoint, chemical characterisation of crude oil to determine its composition is used, among other objectives, to ascertain the possibility of generating new cuts in extra heavy gas oil to generate commodities with high revenue (i.e., diesel). Analysis of extra heavy gas oil by GC×GC-TOFMS allowed detection and identification of hopanes, moretanes, paraffins, steranes, and tricyclic terpanes [30]. Frysinger and co-workers examined the chemical composition of crude oil developed in marine environment by combining the qualitative information obtained from both GC×GC-FID and GC-MS analyses [30]. Alkylated aromatics, sulphur-containing aromatics, isoprenoids, steranes, triaromatic steranes, and triterpanes were identified by analysis of the mass spectra and confirmed by their characteristic elution patterns observed in GC×GC, where an overview of oil facilitates the identification of groups of families into the sample, this allowed the calculation of important geochemical parameters, such as biodegradation, thermal maturation and oil migration. As the co-elution between important biomarkers as tri- and pentacyclic terpanes was deleted. [30]. The elution patterns of known biomarkers from crude oil by GC×GC-FID analysis is shown in Figure 2, [31]. Also, the chromatographic resolution of important critical pairs, namely, tricyclic terpanes/pentacyclic terpanes, C30/C30R demethylated homohopane, and hopanes/sterenes from Brazilian crude oils were resolved by GC×GC-TOFMS, which co-eluted by conventional GC-MS analysis [10]. This was also the first report

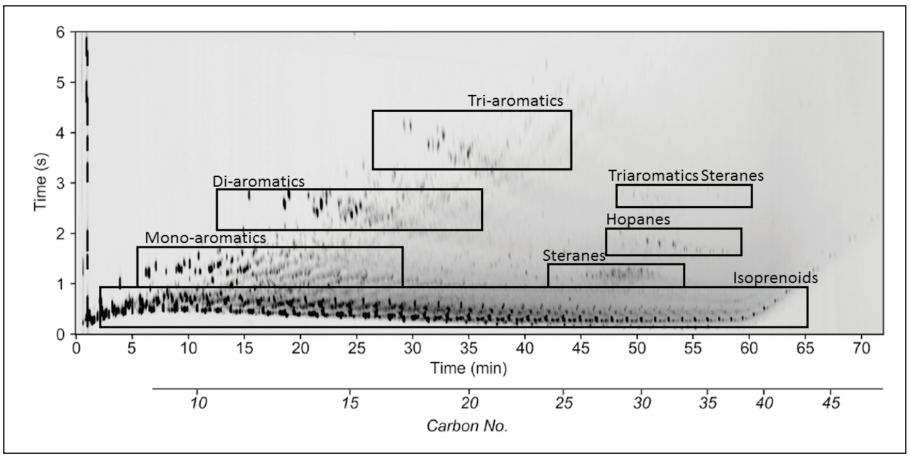

Figure 2. GC×GC-FID chromatogram from Exxon Valdez cargo oil, developed in marine environment. [1]D column: 10 m × 0.10 mm Rtx-1 (polydimethylsiloxane) (d_f = 0.40 µm). [2]D column: 100 cm × 0.10 mm DB-17 (polydimethyldiphenylsiloxane with 50% diphenylsiloxane monomer incorporation) (d_f = 0.10 µm), adapted from Gaines et al. [31]

of demethylated tri- and tetracyclic in Brazilian oils which allowed the study of a new geochemical parameter, these compounds were not developed during diagenesis, indicating that a medium level of biodegradation in the oil is present [10].

Oils of marine origin were analysed by GC×GC-TOFMS and four types of complex mixtures were observed, namely, UCM, UCM I, UCM II, and UCM III. The UCM comprised of C_{36}-C_{40} mono- to tricyclic archaeal isoprenoid diastereoisomers, UCM type I exhibited predominantly mono- to hexacycloalkanes of unknown origin, and UCM type II consisted of C_{35}-C_{40} hydrocarbons that originated from archaeal lipids. Interestingly, the chemical composition of UCM type III was similar to UCM I and II [12].

Several chromatographic techniques, namely, GC-MS, GC×GC-FID, and GC×GC-TOFMS were examined to determine of key petroleum biomarkers to ascertain the cretaceous age and terrestrial organic matter input [32]. Biomarkers of angiosperm origin, namely, lupanoides, olenoides, 18α-oleonane, and 18β-oleonane were identified. GC×GC-TOFMS was used to evaluate the chemical composition of oils from different basins to extract geochemical information from new Brazilian reserves [11]. Several markers including hopanes, steranes, tricyclic terpanes, 8,14-seco-hopanes, onocerane, 3β- and 2α-methylhopanes, and mono- and triaromatics were identified [11]. The 3β-methylhopanes and a series of onoceranes were detected exclusively in samples from depositional lacustrine environments; while 2α-methylhopanes were determined in marine oils [11,33]. It was determined that a ratio between the concentration of C_{31} 3β-methylhopane and C_{30} hopane 100 (3βMHC31/H30) can be used to differentiate lacustrine (3βMHC31/H30 > 1) and marine samples (3βMHC31/H30 < 1) [11]. Also, the identification of biomarkers in oils from the upper Magdalene valley basin (Colombia) by GC×GC-TOFMS was jeopardised by the oil's chemical complexity due to its severe thermal maturity. A chemometric technique, namely multi-way principal component analysis (MPCA), was employed for unsupervised data mining of petroleomic data generated by GC×GC-TOFMS to examine the differences in the oils sampled from several reservoirs [34].

The extent of biodegradation increases the complexity of UCM, which in turn reduces the availability of potential geochemical information, namely, maturity, and origin through peak overlap by GC/MS. Marriott and co-workers were able to resolve several C_1-C_7 alkyl decahydrohopanes from marine, terrestrial, and hybrid oil samples by GC×GC [35]. However, many resolved analytes were not identified possibly due to limited availability of mass spectra in commercial databases.

4. Conclusions and Future Perspectives

Several improvements were observed in the analytical methods employed for chemical characterisation and determination of biomarkers in petroleum. Emphasis has been placed on sample preparation and instrumental analysis. GC×GC has allowed better characterisation of potential biomarkers by reducing the number of compounds co-eluting and increasing the signal-to-noise ratio during chromatographic analyses, which has lead to the acquisition of more accurate and reliable mass spectra. As a consequence, reliable biomarkers are now available to ascertain the thermal maturity, extent of biodegradation, evaluation of the oil's migration, and age of the source rock surrounding the petroleum. It is hoped that improvements in column technology may help expand

the applicability of gas chromatographic techniques in petroleomics by increasing the maximum allowable operating temperature of GC columns stability of stationary phases and manufacturing novel phases with unique selectivities (e.g., ionic liquids and polymeric ionic liquids [14,36]). Also, hyphenation of GC×GC to high-resolution mass spectrometers and/or tandem mass spectrometers will improve structural elucidation of novel biomarkers. Naturally, complex and massive data sets are now being generated by modern analytical systems in petroleomics; therefore, it is expected that updated protocols for data analysis (e.g. chemometrics) will increase in the scientific literature.

5. Acknowledgements

The Brazilian Ministry of Education Agency for Improvement of Graduate Personnel (CAPES), the National Council for Scientific and Technological Development (CNPq), and the São Paulo Research Foundation (FAPESP) are thanked for funding.

6. References

[1] K.E. Peters, C.C Walters, J.M. Moldowan The Biomarker Guide. Biomarkers and Isotopes in Petroleum Systems and Earth History, vols. 1 and 2. Cambridge University Press, USA, 2005, p. 1155.

[2] H.M.E. Van Kaam-Peters, S. Shouten, J.W. de Leeuw, J.S. Sinninghe Damsté, Organic. Geochem. 27 (1997) 399.

[3] S.C. George, H. Volk, A. Dutkiewicz, J. Ridley, R. Buick, Geochim. Cosmochim. Acta 72 (2008) 844.

[4] A. Bechtel, M. Hámor-Vidó, R.F. Sachsenhofer, D. Reischenbacher, R. Gratzer, W. Püttman, Int. J. Coal. Geol. 72 (2007) 33.

[5] P. Farrimond, A. Taylor, N. Telnaes, Org. Geochem. 29 (1998) 1181.

[6] A.E. Pomerantz, G.T. Ventura, A.M. McKenna, J.A. Cañas, J. Auman, K. Koerner, D. Curry, R.K. Nelson, C.M Rodgers, A.G Marshall, K.E. Peters, Mullins, Org. Geochem. 41 (2010) 812.

[7] D.A. Azevedo, J.B. Tamanqueira, J.C.M. Dias, A.P.B Carmo, L. Landau, F.T.T. Gonçalves, Fuel 87 (2008) 2122.

[8] T.F. Silva, D.A. Azevedo, M.D. Rangel, F.R. Aquino Neto, Org. Geochem. 39 (2008) 1249.

[9] ASTM (American Society for Testing and Materials), Standard Test Methods for Comparison of Waterborne Petroleum Oils by Gas Chromatography, D-3328-00, 2000a, W. Conshohocken, PA.

[10] A. Aguiar, A.I. Silva Júnior, D.A. Azevedo, F.R. Aquino Neto, Fuel 89 (2010) 2760.

[11] A.P. Kiepper, A. Casilli, D.A. Azevedo, Org. Geochem. 70 (2014) 62.

[12] C. Eiserbeck, R.K. Grice, J. Curiale, C.M. Reddy, Geochim. Cosmochim. Acta 87 (2012) 299.

[13] E. de Hoffman, V. Stroobant, Mass spectrometry, Wiley, Brussels, 2007, p. 479.

[14] J.C. Giddings, J. Chromatogr. A 703 (1995) 3.

[15] J.B. Phillips, J. Beens, J. Chromatogr. A 856 (1999) 331.

[16] J.V. Seeley, S.K. Seeley, Anal. Chem. 85 (2012) 557.

[17] T.D. Ho, C. Zhang, L.W. Hantao, Anal. Chem. 86 (2014) 262.

[18] P. Marriot, R. Shellie, TrAC, Trends Anal. Chem. 21 (2002) 573.

[19] P.J. Marriott, S. T. Chin, B. Maikhunthod, H.G. Schmarr, S. Bieri, TrAC, Trends Anal. Chem. 34 (2012) 1.

[20] J.V. Seeley, J. Chromatogr. A 1255 (2012) 24.

[21] P.Q. Tranchida, D. Sciarrone, P. Dugo, L. Mondello, Anal. Chim. Acta 716 (2012) 66.

[22] M. Edwards, A. Mostafa, T. Gorécki, Anal. Bioanal. Chem. 401 (2011) 2335.

[23] P.Q. Tranchida, G. Purcaro, P. Dugo, L. Mondello, TrAC, Trends Anal. Chem. 30 (2011) 1437.

[24] J.C. Gidding, Anal. Chem. 56 (1984) 1258.

[25] M. Edwards, A. Mostafa, T. Górecki, Anal. Bioanal. Chem. 401 (2011) 2335.

[26] P.J. Marriott, S.T. Chin, B. Maikhunthod, H.G. Schmarr, S. Bieri, Trends Anal. Chem. 34 (2012) 1.

[27] Z. Liu, J.B. Phillips, J. Chromatogr. Sci. 29 (1991) 227.

[28] J.A. Murray, J. Chomatogr. A 1261 (2012) 58.

[29] B.M.F. Avila, A. Aguiar, A.O. Gomes, D.A. Azevedo, Org. Geochem. 41 (2010) 863.

[30] G.S. Frysinger, R.B. Gaines, J. Sep. Sci. 24 (2001) 87.

[31] R.B. Gaines, G.S. Frysinger, C.M. Reddy, R.K. Nelson, Oil Spill Source Identification by Comprehensive Two-Dimensional Gas Chromatography (GC×GC). Oil Spill Fingerprinting and Source Identification, Chap 5, 2006, p. 169-206.

[32] C. Eiserbeck, R.K. Nelson, K. Grice, J. Curiale, C.M. Reddy, P. Raiteri, J. Chromatogr. A 1218 (2011) 5549.

[33] C.R. Oliveira, A.A. Ferreira, C.J.F. Oliveira, D.A. Azevedo, E.V. Santos Neto, F.R. Alquino Neto, Org. Geochem 46 (2012) 154.

[34] G.T. Ventura, G.J. Hall, R.K. Nelson, G.S. Frysinger, B. Raughuraman, A.E. Pomerantz, O.C. Mullins, C.M. Reddy, J. Chromatogr. A 1218 (2011) 2584.

[35] T. C. Tran, G. A. Logan, E. Grosjean, D. Ryan, P. J. Marriott, Geochim. Cosmochim. Acta 74 (2010) 6468.

[36] L.W. Hantao, A. Najafi, C. Zhang, Anal. Chem. 86 (2014) 3717.