Direct Analysis of Human Sputum for Differentiating Non-small Cell Lung Cancer by Neutral Desorption Extractive Electrospray Ionization Mass Spectrometry

Xiao-Fei Gao,[†] Yipo Xiao, and Yuyou Dai

Jiangxi Key Laboratory for Mass Spectrometry and Instrumentation, East China University of Technology, Nanchang, Jiangxi Province 330013, P. R. China

Human sputum, a typical highly viscous biosample, was directly characterized at the molecular level using neutral desorption extractive electrospray ionization mass spectrometry (ND-EESI-MS) without multi-step sample pretreatment, in an attempt to provide a method for constructing the pattern recognition of rapid diagnosis of lung cancer. Under the optimal experiment conditions, glucose, amino acids, phosphoric lipids and other typical analytes in the sputum sample could be used to conduct qualitative or quantitative (in arginine) analysis. More interestingly, the full scan mass spectra from 50 patients of non-small cell lung cancer, recording the mass spectral fingerprints of sputum samples, were differentiated from the control group (50 healthy individuals) through principal component analysis (PCA). These findings suggest that valuable molecular information concealed in human sputum could be easily revealed and applied for conducting qualitative or quantitative analysis by direct ND-EESI-MS analysis.

Keywords Neutral desorption, extractive electrospray ionization mass spectrometry, human sputum samples, lung cancer, qualitative and quantitative analysis

Introduction

Lung cancer is one of the most common malignancies with high mortality rates worldwide,[1] with non-small cell lung cancer (NSCLC) accounting for 80% of all cases. Most patients with NSCLC are diagnosed in the advanced stage of the disease, with poor survival prospects. It was reported that NSCLC patients have a 55 – 80% chance of surviving 5 years if they are diagnosed at an early stage.[2] Therefore, there is an urgent need for an effective and accurate method for early diagnosis of NSCLS for patients to receive timely prevention. The reported diagnostic methods can be mainly categorized into invasive examinations and non-invasive examinations. The invasive procedures may cause discomfort or pain in patients. Although conventional non-invasive procedures are used in hospitals to diagnose lung cancer, such as human breath gas analysis and sputum cytology,[3] there is still some debatable accuracy of diagnosis of early-stage lung cancer.

Human sputum is a typical mucous mixture secreted by the lungs, bronchi and trachea,[4,5] containing large amounts of mucus, cellular debris and microorganisms under certain pathological conditions.[6] It is usually used in sputum cytology to screen NSCLC in the early stages.[3] However, the sputum samples must be fixed and dyed in cytologic smear analysis, which suffers from several drawbacks, such as complex sample pretreatment, long analysis time and low sensitivity, and thus cannot meet the requirements of rapid clinical diagnosis. In recent years, due to its high sensitivity, specificity and high throughput, advanced mass analyzers (*e.g.*, linear ion trap, Orbitrap) combined with ambient ionization techniques, such as desorption electrospray ionization (DESI),[7] desorption atmospheric pressure chemical ionization (DAPCI),[8] direct analysis in real time (DART),[9] extractive electrospray ionization (EESI),[10] low-temperature plasma probe (LTP),[11] matrix assisted laser desorption/ionization (MALDI),[12] *etc.*, are preferably employed for rapid analysis of complex samples with different forms (solid surfaces,[13] highly viscous samples,[15-20] living objects,[21] complex liquids[22] and gases[23]) in the diagnosis of diseases at the molecular level. In 2007, Chen and Zenobi reported a method using a neutral gas stream to sample the surface of various biological objects for *in vivo* EESI mass spectrometric analysis without sample pretreatment, which is named neutral desorption extractive electrospray ionization mass spectrometry (ND-EESI-MS).[21] The sampling process is totally separated from the ionization process, then non-volatile or volatile compounds can be liberated from virtually any type of highly viscous samples[15,24] (*e.g.*, toothpaste, ionic liquids) for EESI-MS analysis, and the overall result is a very efficient process capable of achieving very low detection limits for complex samples without chemical contamination.

Herein, we tried to use the ND-EESI-MS technique for direct analysis of human sputum samples, to explore the quantitative analysis method for the substance in human sputum and provide a method for constructing the pattern recognition of rapid diagnosis of lung cancer.

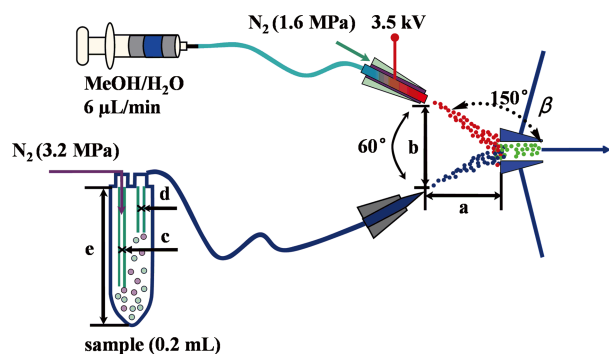† To whom correspondence should be addressed.
E-mail: sandy2005327@126.com

Fig. 1   Schematic illustration of the ND-EESI-MS.   Parametric design: a, 8; b, 2; c, 0.25; d, 2; e, 40 mm.

Experimental

Materials and reagents

This study was approved by the Medical Ethics Committee in the Hospital Institutional Review Board of the Second Affiliated Hospital to Nanchang University, Nanchang, P. R. China. All clinical investigations were conducted according to the principles expressed in the Helsinki declaration of human rights.[25] Patients enrolled in this study were diagnosed with non-small cell lung cancer, confirmed by pathology in the hospital. The patients had an absence of accompanying malignancies and other lung diseases, and had no history of preoperative chemotherapy or radiotherapy. Sputum samples were collected by the hospital from the volunteers after they had fasted for more than 8 h in the morning, and stored at $-80°C$ immediately after collection. The sputum samples were thawed to room temperature before ND-EESI-MS analysis. Methanol (HPLC grade, Fisher Scientific American Company), acetic acid (AR, Shanghai Chemical Reagent Co., Ltd.), and deionized water were used.

Instrument setup

All the experiments were carried out using an LTQ-XL linear ion trap mass spectrometer (US Thermo Electron Corporation with Xcalibur data software) equipped with a homemade ND-EESI source.[26] A KQ3200 ultrasonic cleaning instrument (Kunshan Ultrasonic Instrument Co., Ltd.) was employed for parts cleaning. The setup of ND-EESI is schematically shown in Fig. 1, and is also systematically described elsewhere.[22] Briefly, a gentle N_2 gas flow, saturated with methanol by passing through the pure methanol solvent, was directed to impact the viscous sputum sample, generating aerosols containing matrix and analyte liberated from the sputum. Driven by the nitrogen gas flow passing through a heated sample transfer line ($\sim$250°C), the analyte was introduced into the ionization region of the EESI source, where the neutral molecules were protonated by the primary ionic plume produced by electrospray solution (methanol/water/acetic acid, *v/v/v*, 50/48/2). The ionization region of the EESI source was maintained at 150°C, which facilitated the desolvation and transportation process of the analyte ions, resulting in better sensitivity and improved stability of analyte detection.

For quantitative analysis, the sputum sample was treated using an ultrasonic homogenizer for 10 min to form a homogeneous sputum sample. Then 100 μL arginine (1 ppm) methanol/water solution was added into a 200-μL unified sputum sample as the internal standard. Under the experimental conditions, arginine was detected as protonated molecule (*m/z* 175), which provided

a characteristic fragment of *m/z* 60 in MS/MS spectrum. Thus the operational parameters were optimized by obtaining the highest signal level of the arginine and the instrumental baseline drift was calibrated by using a sputum sample containing constant arginine levels. The LTQ mass instrument was set to work in the positive ion detection mode, the temperature of the heated capillary was 150°C, the lens voltage was 60 V and the N_2 pressure was 1.6 MPa. The spray voltage was +3.5 kV, and the default values for the rest parameters of the LTQ instrument were directly used without further instrument optimization.

Data processing

Principal component analysis (PCA) was used as the statistical tool to picture the difference between the lung cancer samples and the control samples. For PCA, the chemical fingerprint spectrum data of different types of human sputum samples were exported into Excel for further processing using Matlab (Ver. 7.8.0, Math-works, Inc., Natick, MA) software. After normalization of the mass spectrometry data, the "Princomp" function was selected to perform the data analysis. Finally the score and load results of the first three principal components were graphically displayed, showing the differences between the lung cancer samples and the control samples.

Results and Discussion

Optimization of ND-EESI-MS for sputum analysis

The experimental setup is schematically shown in Fig. 1. The operational parameters including the angle β, the distances (a, b, c, d, e), the flow rate of the electrospray solvent for extractive ionization, the ND gas pressure, the ESI spray voltage and the temperature of the heated sample transfer line were optimized using the sputum sample with spiked arginine (10 ppm). After a series of experiments using mass spectrometry (arginine was selected as substrate), the optimal conditions of mass spectrometry were obtained: angle β was 150°; distance a was 8 mm; the distance b was 2 mm; the distance c was 0.25 mm; the distance d was 2 mm; the distance e was 40 mm; the flow rate of the electrospray solution (*i.e.*, methanol/water/acetic acid, 50/48/2, *v/v/v*) was 8 μL/min; the ND gas pressure was 1.6 MPa; the ESI spray voltage was +3.5 kV and the temperature of the heated sample transfer line was maintained at 100°C. The effect of the angle, distance a, flow rate of extraction solvent, N_2 pressure, and ESI spray voltage were examined to optimize conditions of ND-EESI-MS for sputum analysis. Each independent experiment recorded 60 s in MS, and the average values are based on five independent experiments. The error bars are defined as the standard deviation of the five independent experiments (results are not shown here). The signal responding dynamics were similar to those reported in the literatures,[27,28] confirming that the experimental conditions selected for the measurements conducted below were optimal.

Qualitative analysis of sputum by ND-EESI-MS

Sputum samples can be directly analyzed by ND-EESI-MS under either the positive or negative ion detection mode (results not shown). Herein, positive-ion mode was used in this study. Figure 2a shows the mass spectrum recorded from a sputum sample of a healthy volunteer while Fig. 2b shows the spectrum recorded from a sputum sample of a cancer patient under the positive ion detection mode. Amino acids such as glutamine (*m/z* 147), phenylalanine (*m/z* 166), tyrosine (*m/z* 182) and arginine (*m/z* 175), glucose (*m/z* 181) and many other non-volatile chemicals of relatively larger molecular weights were
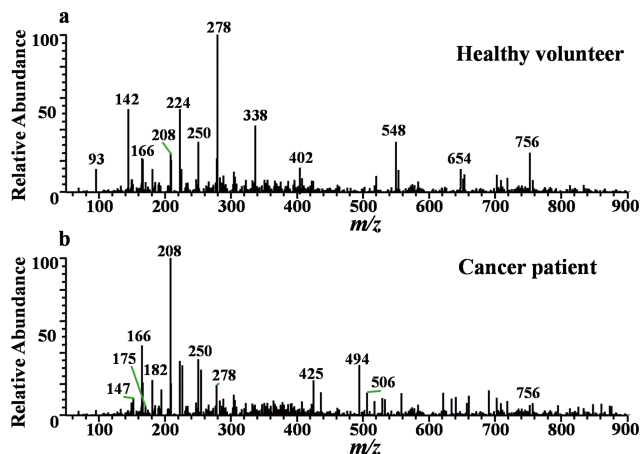
Fig. 2   Typical ND-EESI-MS spectra of sputum samples. a, Healthy volunteer; b, cancer patient.

detected in the mass range of m/z 50 – 900 (Fig. 2a). Figure 2 demonstrates that the ND-EESI was able to liberate the analyte from viscous sputum for mass spectrometry analysis, indicating that the experimental conditions were feasibly selected for further measurement.

Upon collision-induced dissociation (CID) experiment, amino acids such as glutamine (m/z 147), phenylalanine (m/z 166), arginine (m/z 175) and tyrosine (m/z 182) provided characteristic fragments in the CID spectra (see Fig. 3).   For example, protonated glutamine (m/z 147) yielded major fragments of m/z 130, 129 and m/z 101 (Fig. 3a), by the loss of NH_3, H_2O and HCOOH, respectively. The protonated phenylalanine (m/z 166) yielded major fragments of m/z 149 and 120 (Fig. 3b), by the loss of NH_3 and HCOOH, respectively. A similar fragmentation pattern was also observed for protonated tyrosine (m/z 182), which generated m/z 165 and 136 (Fig. 3c) by the loss of NH_3 and HCOOH, respectively. However, the relative abundances of the ionic residue obtained by cleavage of HCOOH varied significantly in the CID spectra, showing that the cleavage of HCOOH was heavily related to the structure of the precursor ions. This finding was further observed in the case of arginine (Fig. 3d). Under similar CID conditions, protonated arginine (m/z 175) cleaved NH_3, H_2O successively to generate fragments at m/z 158 and 157 in the MS/MS spectrum. Alternatively, the precursor ions might lose NH_3 and H_2O simultaneously to generate ionic fragments of m/z 140 in the MS/MS spectrum. The simultaneous loss of ($NH_3 + CO$) from the protonated arginine was also witnessed with the detection of the abundant peak at m/z 130. Under the same conditions, the neutral loss of ($NH=C(NH_2)_2$, MW = 59) was also favored by producing the fragment of m/z 116, which might further lose ($H_2O + CO$) in the spectrum to obtain the highly characteristic ions of m/z 60.

More interestingly, the signal detected at m/z 756 was tentatively assigned as dipalmitoyl phosphatidylcholine sodium salt. In the MS/MS spectrum (Fig. 3e) the major peak at m/z 697 was generated by the loss of trimethyl amine $N(CH_3)_3$ (M_W = 59) from the parent ions. With more collision energy applied, the product ions of m/z 697 yielded major peaks at m/z 573 and 551 by the loss of either a phosphorus acid ethyl $CH_2CH_3PO_4$ (i.e., phosphorylcholine, M_W = 124) or a phosphoryl ethyl + sodium $Na(CH_2)_2PO_4$ (M_W = 146).   The tiny peak detected at m/z 279 was ascribed to sodium palmitate, which was the last ionic residue of the CID experiments (the inset of Fig. 3e). Moreover, the fragmentation pattern was confirmed
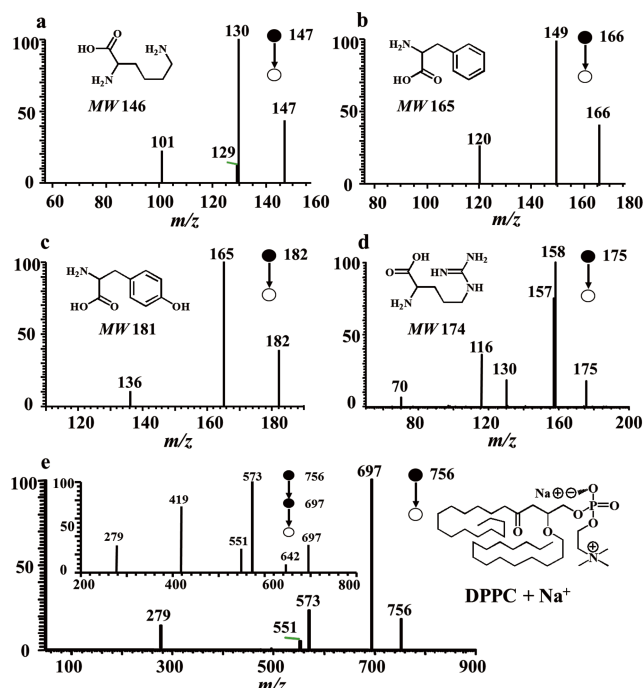

Fig. 3   Tandem mass spectra of analytes detected in sputum samples. a, Glutamine; b, phenylalanine; c, tyrosine; d, arginine; e, dipalmitoyl phosphatidyl choline (DPPC).

using the sodium cluster formed between the authentic dipalmitoyl phosphatidyl choline (DPPC) and sodium ions (DPPC + Na$^+$). Therefore, the signal detected at m/z 756 was attributed to dipalmitoyl phosphatidylcholine sodium salt, a main component of pulmonary surfactant.[29-31]

Quantification of target metabolites in sputum by ND-EESI-MS

Quantitative analysis of target analytes is often required in metabolomics studies. As a demonstration, arginine in sputum was quantitatively detected using ND-EESI-MS/MS. To exclude any potential interference signal, the fragment of m/z 60 generated in MS/MS was selected as the signal to make the calibration curve. Figure 4a shows the linear calibration curve of the arginine signal responded in the range of 1 – 1000 μg/L (It is the average value of five repeated measurements). Quantification is based on the EIC (m/z 60), with a linear equation $y = 12.4x(\mu g/L) - 3.8$ (R^2 = 0.9997), wherein y is the MS signal intensity of arginine in the sputum sample; x is the arginine concentration in the sputum sample. Accordingly, the limit of detection (LOD) of this method for arginine was found to be 0.1 μg/L (arginine concentration in the sputum, S/N = 3, LOD = $c3\sigma/S$), wherein c is the arginine concentration in the sputum sample; σ is the standard deviation of 10 measurements of the signal intensity; S is the signal intensity level averaged for 10 times. It was reported that the amount of L-arginine in sputum of clinically stable patients with cystic fibrosis, in patients with cystic fibrosis before and after treatment for a pulmonary exacerbation, and in healthy control subjects, ranged from 195 to 628 μg/L, detected by liquid chromatography–tandem mass spectrometry.[32] Besides, the linear calibration curve of the arginine signal responded in the range of 1 – 1000 μg/L in our method. This method was considered suitable for arginine quantitative analysis in actual sputum samples.

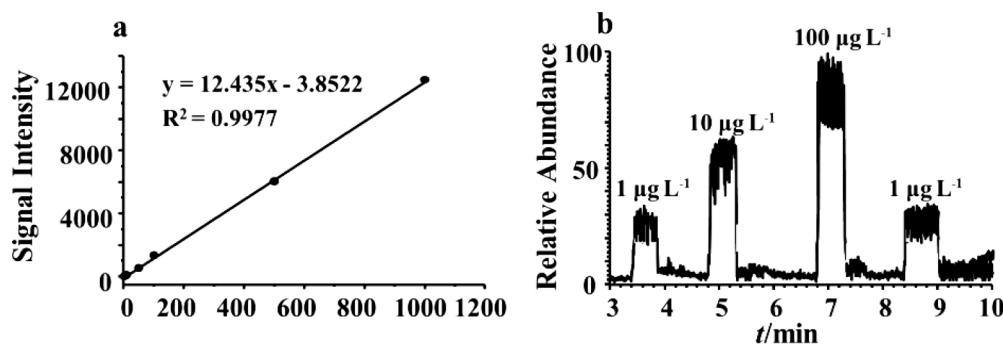The reasonable LOD obtained using MS/MS signal of arginine

Fig. 4   Quantitative analysis of analytes in sputum using ND-EESI-MS. a, Arginine calibration curve; b, selected ion current chromatography of the protonated arginine.

Table 1   Recovery rates and RSD values obtained by ND-EESI-MS/MS for detection of arginine spiked in sputum samples

Added arginine/ $\mu g\ L^{-1}$	Concentration of detected arginine/$\mu g\ L^{-1}$		RSD, %	Recovery, %
	Value/ $\mu g\ L^{-1}$	Mean value/ $\mu g\ L^{-1}$		
5	4.88, 4.83, 4.76, 4.41, 4.66, 4.33, 4.55, 4.46	4.61	4.4	92.2
50	48.9, 47.9, 45.8, 49.4, 45.0, 49.8, 46.3, 48.5	47.7	3.7	95.4
500	477, 495, 470, 49, 462, 489, 460, 496	481	3.3	96.2

suggested that the method established here was of high sensitivity for detection of a given metabolite in sputum samples. Figure 4b shows the selected ion current chromatography of the protonated arginine (m/z 175) recorded using ND-EESI-MS from 4 individual sputum samples spiked with arginine at differential concentrations of 1, 10, 100 and 1 μg/L, respectively.

Arginine was also used to evaluate the recovery rates and reproducibility of the method, wherein the blank sputum sample without signal of arginine, before the addition of spiked arginine, was used as the substrate. Under the optimized experimental conditions, acceptable recoveries between 90.0 – 97.3% and satisfactory relative standard deviation (RSD) values between 3.3 – 4.4% were found for all the sputum samples spiked with different levels of arginine. Table 1 summarizes the analytical results obtained by this method for sputum sample analysis.

ND-EESI-MS differential analysis of sputum samples to differentiate lung cancer

Under the optimized experimental conditions, a total of 100 sputum samples collected from 50 healthy people and 50 lung cancer patients were analyzed by ND-EESI-MS to record the mass spectral fingerprints. All the full scan mass spectra were subjected to principal component analysis, which resulted in a 3D PCA plot by principal component of PC1 (73.5%), PC2 (12.6%) and PC3 (4.1%) where the lung cancer sputum samples were well separated from the non-cancer samples (Fig. 5a). Wherein PC1 – PC3 contains a total of 83.8% of the data variables, two sets of samples were primarily differentiated between principal components PC1 and PC2, indicating the inherent differences that distinguish the two groups of samples. The differential peaks were illustrated in the loading plots (Fig. 5b). For example, phosphatidylcholine (PC) and dipalmitoyl

phosphatidylcholine (DPPC) were responsible for the major differential signals detected as protonated and sodium adducts. These results showed that ND-EESI-MS sputum analysis combined with PCA for data processing may provide the possibility to early diagnosis of non-small cell lung cancer, distinguishing lung cancer patients from the non-cancer group. Samples taken from patients with tuberculosis were also examined by our group, which can also be separated from the the samples from patients with cancer.[33]

Conclusions

Herein, neutral desorption extractive electrospray ionization mass spectrometry (ND-EESI-MS) was successfully adopted for the direct analysis of sputum samples without pretreatment under the optimal mass conditions, with a short time required for each sample analysis. Analytes in the sputum samples can be directly subjected to qualitative (in glucose, amino acids, dipalmitoyl phosphatidyl choline) and quantitative (in arginine) analysis. Moreover, the full scan mass spectra, recording the mass spectral fingerprints of sputum samples collected from 50 healthy people and 50 lung cancer patients, were studied by principal component analysis (PCA), wherein the lung cancer sputum samples were well separated from the healthy samples. These findings suggest that valuable molecular information concealed in human sputum could be rapidly revealed by direct ND-EESI-MS analysis without pretreatment, which may provide an alternative method for allowing non-invasive differentiating of non-small cell lung cancer patients from healthy individuals.

Acknowledgements

References

1. W. Chen, R. Zheng, P. D. Baade, S. Zhang, H. Zeng, F. Bray, A. Jemal, X. Q. Yu, and J. He, *CA-Cancer J. Clin.*,
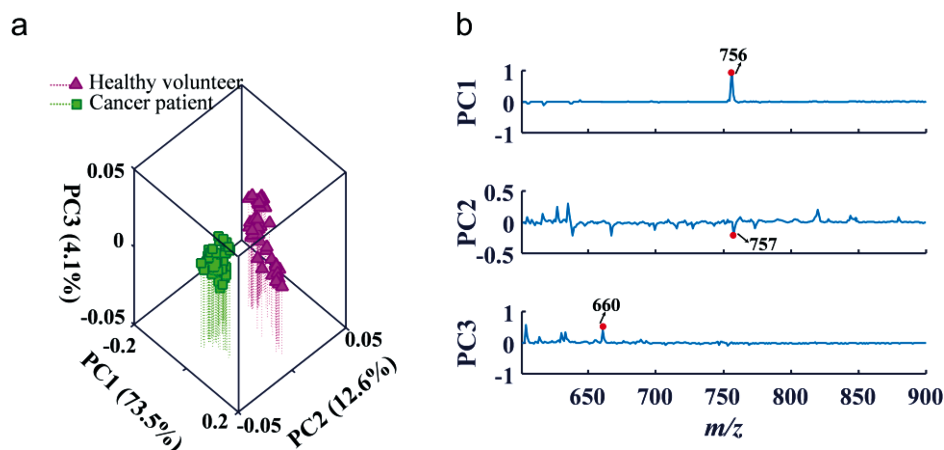
Fig. 5   PCA results of sputum samples analyzed by ND-EESI-MS. a, 3D plot of PCA score results; b, PCA loading results for the PCs. The lipids in each sample have no relevance to any clinical parameter, such as TNM stage, tumor placement in the lung, gender, age or smoking history.

2016, *66*, 115.

2. T. Wang, R. A. Nelson, A. Bogardus, and F. W. Grannis, *Cancer*, **2010**, *116*, 1518.

3. Y. Xie, N. W. Todd, Z. Liu, M. Zhan, H. Fang, H. Peng, M. Alattar, J. Deepak, S. A. Stass, and F. Jiang, *Lung Cancer*, **2010**, *67*, 170.

4. Q.-G. Li, J.-X. Liang, G.-Y. Luan, Y. Zhang, and K. Wang, *Anal. Sci.*, **2000**, *16*, 245.

5. B.-S. Yu, P. Chen, L.-H. Nie, and S.-Z. Yao, *Anal. Sci.*, **2001**, *17*, 495.

6. E. T. Y. Leung, L. Zheng, R. Y. K. Wong, E. W. C. Chan, T. K. Au, R. C. Y. Chan, G. Lui, N. Lee, and M. Ip, *J. Clin. Microbiol.*, **2011**, *49*, 2509.

7. Z. Takáts, J. M. Wiseman, B. Gologan, and R. G. Cooks, *Science*, **2004**, *306*, 471.

8. S. Yang, J. Ding, J. Zheng, B. Hu, J. Li, H. Chen, Z. Zhou, and X. Qiao, *Anal. Chem.*, **2009**, *81*, 2426.

9. J. Y. Yew, R. B. Cody, and E. A. Kravitz, *Proc. Natl. Acad. Sci. U. S. A.*, **2008**, *105*, 7135.

10. H. Chen, A. Venter, and R. G. Cooks, *Chem. Commun.*, **2006**, 2042.

11. Y. Liu, X. Ma, Z. Lin, M. He, G. Han, C. Yang, Z. Xing, S. Zhang, and X. Zhang, *Angew. Chem., Int. Ed.*, **2010**, *49*, 4435.

12. J. Gopal, N. Hasan, M. Manikandan, and H.-F. Wu, *Sci. Rep.*, **2013**, *3*, 1260.

13. R. G. Cooks, Z. Ouyang, Z. Takats, and J. M. Wiseman, *Science*, **2006**, *311*, 1566.

14. X. Li, B. Hu, J. Ding, and H. Chen, *Nat. Protoc.*, **2011**, *6*, 1010.

15. M. Sugimoto, D. T. Wong, A. Hirayama, T. Soga, and M. Tomita, *Metabolomics*, **2010**, *6*, 78.

16. A. Abdel-Rehim and M. Abdel-Rehim, *Biomed. Chromatogr.*, **2014**, *28*, 875.

17. J. J. W. Mikkonen, S. P. Singh, M. Herrala, R. Lappalainen, S. Myllymaa, and A. M. Kullaa, *J. Periodontal Res.*, **2015**, *51*, 431.

18. M. Numako, T. Takayama, I. Noge, Y. Kitagawa, K. Todoroki, H. Mizuno, J. Z. Min, and T. Toyo'oka, *Anal. Chem.*, **2016**, *88*, 635.

19. E. P. Dutkiewicz and P. L. Urban, *Philos. Tr. Roy. Soc., A*, **2016**, *374*, 20150380.

20. T. Takayama, H. Tsutsui, I. Shimizu, T. Toyama, N. Yoshimoto, Y. Endo, K. Inoue, K. Todoroki, J. Z. Min, H. Mizuno, and T. Toyo'oka, *Clin. Chim. Acta*, **2016**, *452*, 18.

21. H. Chen, S. Yang, A. Wortmann, and R. Zenobi, *Angew. Chem., Int. Ed.*, **2007**, *46*, 7591.

22. K. Dettmer, P. A. Aronov, and B. D. Hammock, *Mass Spectrom. Rev.*, **2007**, *26*, 51.

23. H. Chen, A. Wortmann, W. Zhang, and R. Zenobi, *Angew. Chem., Int. Ed.*, **2007**, *46*, 580.

24. H. H. Jeleń, M. Majcher, and M. Dziadas, *Anal. Chim. Acta*, **2012**, *738*, 13.

25. R. V. Carlson, K. M. Boyd, and D. J. Webb, *Brit. J. Clin. Pharmacol.*, **2004**, *57*, 695.

26. H.-W. Chen, J.-H. Lai, Y.-F. Zhou, Y.-F. Huan, J.-Q. Li, X. Zhang, Z.-C. Wang, and M.-B. Luo, *Chin. J. Anal. Chem.*, **2007**, *35*, 1233.

27. Y. Zhou, Z. Wu, C. Li, N. Wang, X. Zhang, H. Chen, and S. Xiao, *Anal. Methods*, **2014**, *6*, 1538.

28. J. Ding, S. Yang, D. Liang, H. Chen, Z. Wu, L. Zhang, and Y. Ren, *Analyst*, **2009**, *134*, 2040.

29. Y. Wei, L. Chen, W. Zhou, K. Chingin, Y. Ouyang, T. Zhu, H. Wen, J. Ding, J. Xu, and H. Chen, *Sci. Rep.*, **2015**, *5*, 10077.

30. M. Kodama, O. Shibata, S. Nakamura, S. Lee, and G. Sugihara, *Colloids Surf., B*, **2004**, *33*, 211.

31. T. W. Jaskolla, K. Onischke, and J. Schiller, *Rapid Commun. Mass Spectrom.*, **2014**, *28*, 1353.

32. H. Grasemann, S. Al-Saleh, J. A. Scott, D. Shehnaz, A. Mehl, R. Amin, M. Rafii, P. Pencharz, J. Belik, and F. Ratjen, *Am. J. Respir. Crit. Care Med.*, **2011**, *183*, 1363.

33. F. Yan, Y. Wei, J. Ding, S. Xiao, B. Jia, and H. Chen, *Chin. J. Exp. Surg.*, **2011**, *18*, 422.