

Chapter 2

Gas Chromatography Mass Spectrometry Coupling Techniques

Zhen Xue, Li-Xin Duan and Xiaoquan Qi

Relative to other metabolomics analysis techniques, gas chromatography mass spectrometry (GC/MS) is one of the earliest applied analysis techniques in metabolomics. The first paper on metabolomics (metabolic profiling) is derived from the application of GC/MS analysis in urine and tissue extracts (Dalglish et al. 1966). With the arrival of omics era and the proposing of metabolomics concept, people began to try using a variety of analytical techniques to obtain metabolomics data. These techniques include chromatography, capillary electrophoresis, mass spectrometry, nuclear magnetic resonance (NMR), infrared spectroscopy, and electrical chemical methods, etc. GC/MS and NMR are the main technologies applied in the early development of metabolomics; in later stage, high-resolution liquid chromatography mass spectrometry (LC-MS) with fast scanning capability is widely used in metabolomics analysis; in recent years, people began trying to integrate a variety of analytical techniques in order to bring into play the advantage of a variety of methods and to make up for the lack of a single analysis. However, there is no one technology can perform quantitative and qualitative analysis for all the endogenous metabolites in the biological sample. GC/MS is the most mature chromatography mass spectrometry coupling technology, suitable for the analysis of metabolites with low polarity, low boiling point, or volatile after being derivatized. GC/MS has been one of the main analytical platforms in plant metabolomics due to the high resolution, high sensitivity, good reproducibility, a large number of standard metabolite spectra libraries, and the relative low cost. This chapter introduces the GC/MS technology principle, plant metabolite extraction and pretreatment technologies, main problems of GC/MS in plant metabolomics analysis, notes, as well as the latest developments.

Z. Xue (✉) · L.-X. Duan · X. Qi
Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China
e-mail: xuezhen@ibcas.ac.cn

L.-X. Duan
e-mail: nlzn@ibcas.ac.cn

X. Qi
e-mail: xqi@ibcas.ac.cn

2.1 GC/MS Principles and Key Technologies

GC can well separate complex mixtures, and MS can detect these compounds. The combination of the two has a more favorite place, for example, both GC and MS can run in the gaseous state; thus, they can be connected directly, and the interface is very simple. Simply speaking, the performance of GC/MS is stable, and the reproducibility is good.

GC plays a role in separation and introduces target substances into MS system by directly injecting analytes into chromatographic column or introducing analytes into chromatographic column after injecting and heating. The chromatographic column is heated thermostatically or program-controlled. Each component is separated by the difference of thermodynamic properties (the difference of boiling points and the difference of selective absorption in the stationary phase) and the different distribution in stationary phase and mobile phase (carrier gas). MS is in fact a detector, mainly including ionization source, mass analyzer, and electron multiplier tubes. Target substances enter into MS through GC and ionized into gaseous ions in the ionization source and then enter into mass analyzer. Ions with different mass-to-charge ratio are sequentially separated and reach the electron multiplier, generating electrical signal, in order to give the 3D information of the target substances, making qualitative analysis more accurate by using ion fragment information. Figure 2.1 shows a schematic diagram of the main parts of GC/MS.

The following are the key technologies of GC/MS (Villas-boas et al. 2007).

2.1.1 Key Technologies of GC in GC/MS

2.1.1.1 Gas Supply System and Mobile Phase

Helium is the commonly used mobile phase in GC/MS, which is provided by compressed gas cylinder, and the flow rate is controlled by pressure and flow

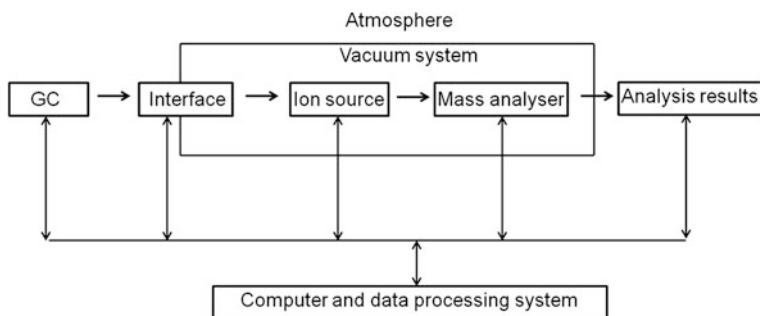


Fig. 2.1 The schematic diagram of the constitution of GC/MS

regulators. GC/MS analysis requires stable airflow, so the gas supply system is a critical component in GC; most of the advanced flow rate control systems are very stable and capable of providing stable and accurate flow rate and pressure control. Gas quality is also an important aspect guaranteeing GC/MS analysis. Gas quality includes gas purity and gas supply. Poor gas quality will cause the emergence of “ghost peaks” in chromatogram. In addition, gas residues, especially oxygen and water, will damage chromatographic column (polar substances in the column are very sensitive to oxygen), and oxygen will reduce the lifetime of lamp filament, and the presence of hydrocarbons will increase signal background. Thus, it is inevitable to use high-purity carrier gas and gas purification apparatus to remove trace oxygen and water in carrier gas. Gas purity should reach 99.9995 %, the purer the better.

2.1.1.2 Chromatographic Columns and Oven

GC columns include packed column and capillary column. The packed column is to coat stationary liquid to carrier particles with uniform size and pack the coated carrier into metal, glass, or plastic tube. Capillary column is to coat stationary phase to the inner wall of the capillary tube, and the capillary tube has no packing. Generally, GC/MS-specific column with MS identification is chose in GC/MS. So far, GC column is classified by the difference of mobile phases: The apolar methyl silicone column is the most used stationary phase; the methyl silicone containing 5 % phenyl groups is the medium polar column; the cyano-propyl methyl silicone stationary phase makes the polarity of the column greater; and the column with the largest polarity is the carbowax-containing stationary phase. These stationary phases are chemically bond into the inner wall of the column, cross-linked to increase its stability. All the columns have their own temperature tolerance ranges; generally speaking, the greater the polarity of the column, the worse the high temperature tolerance.

The diameter of the capillary column and the thickness of the stationary phase determine the β value (the distribution ratio of substance between the gas phase and the stationary phase), that is, the amount of substances distributed in the gas phase and the stationary phase. β value is a core parameter in selecting column. Low β value will increase the retention ability of analytes in the column (the equivalent of more analytes retained in the stationary phase), while reducing the number of plates. Thus, column with thick-film stationary phase (low β value) is typically used for the analysis of volatile compounds, and thin-film column is beneficial for the analysis of less volatile compounds with high boiling point.

Chromatographic peak width is proportional to the square root of the column length, while the column length and the number of theoretical plates have an important relationship. Retention time is proportional to the retaining time of the substance in the stationary phase, that is, the longer the substance retained in the column, the wider the peak width. According to the principle that separation efficiency (theoretical plate number N) is proportional to the square root of column length, if column length increases four times, for example, the resolution will be doubly increased.

Since the distribution efficiency of substances between two phases is strongly depend on the gas temperature, the control of GC temperature is also very important. Suitable temperature program can effectively improve the separation efficiency of substances in the analysis process. At the same time, temperature program can also be used to optimize analysis time.

2.1.1.3 GC Injection System

Injection system is the most critical part of GC, with the aim to convert liquid samples into gaseous state and focus them in the starting end of the column. In metabolomics analysis, imperfections of injection system will cause the incomplete vaporization of sample and the sample cannot timely enter into the column, resulting in peak broadening. Split/splitless injection commonly used injection modes in metabolomics, and also the focus is discussed below.

Split/splitless injection is based on the technique that the sample is rapidly evaporated in a small heated chamber, and the volatile substances are transferred into the column under the action of carrier gas. In the packed column, the flow rate of the carrier gas is high (30–50 ml/min), which is very easy to transfer the sample into the column rapidly and efficiently. However, the introduction of capillary column reduces the flow rate (typically 1–2 ml/min), making the need to change prior techniques to adapt to the characteristics of the capillary column. Initially, it is resolved by venting a part of the sample out of the sample inlet with high flow rate, but this will result in the loss of sample, reducing the sensitivity—this is known as split injection. A later development was to close the split vent during injection, while focusing the analyte in the column rapidly—splitless injection. Figure 2.2 shows the schematic diagram of the split/splitless injection system.

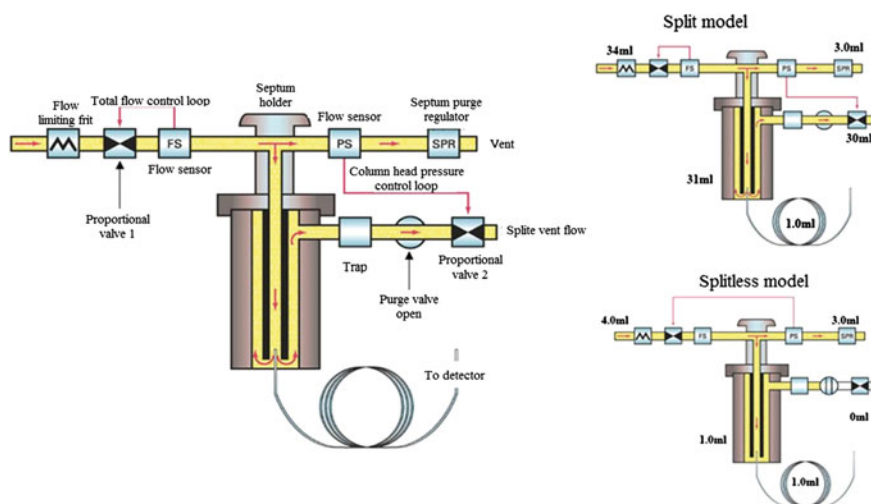


Fig. 2.2 The schematic diagram of the split/splitless injection system

The liner, typically a glass tube, is installed in the heated block of GC inlet. As the evaporation chamber of sample, it has a lot of design patterns, such as with/without packing, various deactivation methods, insertion modes, and sizes. A large volume liner (wide bore) is often used for splitless injection and a smaller liner (narrow bore) for split injection. The liner diameter is usually 2–4 mm, the length is usually 8–10 cm, and the volume of wide-bore liner is about 1 ml.

Injection is to instantly and completely transfer the sample into GC column, which begins from the penetration of syringe into the septum/seal. When the plunger is pushed down, the sample is injected into the hot glass liner, where solvents and analytes are ideally flash-evaporated. Evaporation is a rather complex process. Incomplete evaporation is the main problem it may be encountered. Incomplete evaporation will cause the involatile matrix, droplets, and involatile substances hit and get deposited on the glass of the liner and then slowly released by thermal degradation, seriously affecting the chromatographic separation and even causing the appearing of “ghost peak.” Another problem is that the gas flows through the liner too fast and enters into the inlet of the column before the droplet is completely evaporated, or enters into the inlet of the column by spraying means. Finally, a noteworthy problem is the sample overfilling the injector. 1 μl solvent evaporates to produce 0.5–1 ml gas, thus completely filling a normal wide-bore liner. For split injection, the gas flows through the injector quickly and the evaporated solvent is quickly removed, so it is not prone to overfilling problem, usually by reducing the split ratio to increase the injection volume. But in case of splitless injection where the flow rate through the injector is very low, a larger volume of solvent vapors may overfill the injector. For example, overfilling the liner will lead to cross-contaminations, high variability, and high background.

Split injection is mainly for high-concentration samples. In the split injection, the majority of the carrier gas is split from the bottom of the liner and be vented away. The flow amount into the column is mainly adjusted by split vent. When split injection is carried out, the flow rate of the carrier gas is relatively high, and there is a long distance between injection syringe and the column inlet (injector bottom), allowing more time for the evaporation of the sample. Typically, a narrow bore liner can give effective heat transfer, ensuring the maximum concentration of sample vapors. Although split injection can give sharper peaks, most of the sample will be lost (it will loss 97 % of the whole sample), and nonlinear split will cause quantification distortion.

Splitless injection is used to increase the amount of sample directly into the column by closing the split vent during the injection to make all the gas flow arrive the column through the liner. Since the flow rate of the column is limited in the range of several milliliters per minute, so that it takes quite a while for all the samples entering into the column, typically in the range of 30–90 s, which will cause the expansion of original spectrum very easily. Hence, measures must be taken to focus the sample before it enters into the column, in order to get a better chromatographic separation. In simple terms, the injection time has to be short compared with the peak width in the chromatograms. Solvent recondensation technology can effectively trap sample, concentrating sample at the beginning of the

column. Such concentration technology is done by installing a fused silica column which is deactivated but without stationary phase at the beginning of the column. This silica column is about 2–5 m in length. Solvent will be effectively evaporated when passing through this column, thereby concentrating sample molecules to a narrow band. When all the solvent is evaporated, sample molecules will be focused in a narrow injection band with the carrier gas as a small bandwidth and into the column to be separated.

2.1.1.4 Derivatization Technology for GC

GC requires the sample to be efficiently evaporated in the injector, which is very easy for small molecular compounds with low boiling point (lower than 200–300 °C). However, high boiling point compounds need to be chemically derivatized before evaporation. In metabolomics, interested compounds are usually amino acids, sugars, small molecular organic acids, and other polar metabolites, as well as some apolar metabolites, such as fatty acids and sterols, and the majority of the metabolites are in their nonvolatile state. For compounds containing carboxylic, hydroxylic, or amino groups, derivatization can increase their volatility. The following are the main benefits of derivatization for GC/MS analysis (Wang et al. 2001).

- ① The GC nature of the analytes can be improved. The GC nature of some relatively polar groups, such as hydroxyl and carboxyl, is not good, which has no peak or peak tailing in some common columns. After derivatization, the situation is improved.
- ② The thermal stability of the analytes can be improved. The thermal stability of some analytes is not good. The analyte will decompose or change in the evaporation or chromatographic process. Derivatization can convert it into stable compound under GC/MS detection condition.
- ③ The molecular weight of the analyte can be changed. The molecular weight of most of the derivatized analytes is increased, which is beneficial for the separation between analyte and matrix, reducing the impact of background chemical noise.
- ④ MS behavior of the analyte can be improved. In most cases, derivatized analytes produce more regular and easily be interpreted mass fragments.
- ⑤ The introduction of halogen atom or electron-withdrawing group makes analytes easily be detected by chemical ionization. In most cases, the detection sensitive can be increased, and the molecular weight of the analyte can be detected.
- ⑥ Some particular derivatization methods can be used to split some chiral compounds which are difficult to be separated.

Of course, if the derivatization method is applied improperly, some drawbacks will be brought about. For example, some derivatization reagents need to be removed by nitrogen flow before injection, and improper operation will result in sample loss. Incomplete derivatization reaction will reduce the damage sensitivity of

detection, and the improper use of derivatization reagent will sometimes make the molecular weight of analyte increase too much, approaching or exceeding the mass range of some small MS detectors. Derivatization will produce some artifacts in the sample and will contain some excess reagent, which will seriously interfere with split/splitless injection, and usually, they are not volatile and deposit in the liner.

Derivatization methods suitable for GC/MS analysis include methylation and silylation. There are also many other derivatization methods, and for details, see relevant books (Drozd 1981; Toyooka 1999).

2.1.2 Key Technologies of MS in GC/MS

MS itself can analyze very complex samples. Meanwhile, it is a detector for chromatography, providing very high sensitivity as well as chemical and structural information after connecting chromatography. The development of modern biological MS is more or less driven by the development of metabolomics. MS is a very important analytical means among modern biotechnological analytical methods. Almost all the biotechnological analytical problems can be resolved by MS, from small molecular volatile substances, to complex natural products, as well as proteins and viruses.

The core principle of MS is to determine the mass-to-charge ratio (m/z) of charged compounds. In principle, any charged (or can be charged) substance, which can be transferred into GC at the same time, can be detected by MS. Major development in recent decades is a great expansion of the molecular weight range of MS and the significant improvement of sensitivity. Meanwhile, the mass spectrometer becomes cheaper and easier to be operated. For the instrument constitution of MS, see Fig. 2.3.

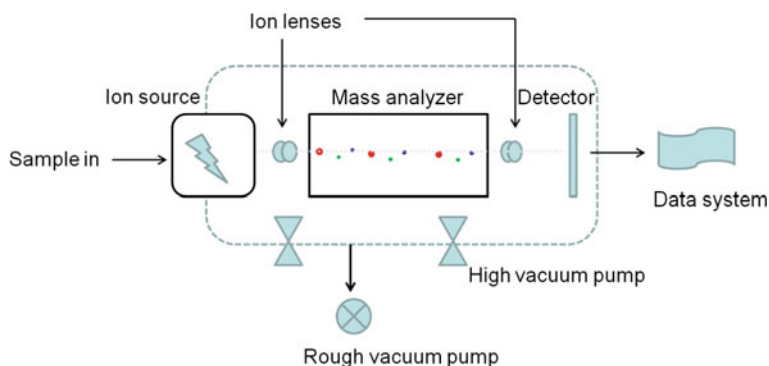


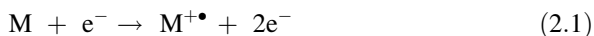
Fig. 2.3 Major constitution of mass spectrometer

2.1.2.1 The Ion Source

GC/MS generally uses open capillary column, the flow rate of carrier gas is low, not to damage the vacuum environment of MS, so that samples can be directly introduced into ion source. The most critical process in ion source is the evaporation, ionization, and transferring of sample into the vacuum system, which depends on the sample type (gas/liquid) and ionization method. These processes can also be carried out in reverse order. For example, the sample is ionized in the solvent, and then, the ions are transferred into gas phase. So far, the most commonly used ionization technique is electron-impact ionization (EI) and electrospray ionization (ESI), the former is mainly coupled with GC, and the later is mainly coupled with LC. EI ion source and GC/MS represent the classic mass spectrometer combination, because the gaseous mobile phase can well flow into MS vacuum. Modern GC/MS system is already very mature technology, is easy to operate, and delivers highly reproducible results.

The lamp filament of EI ion source is usually made of tungsten filament or rhenium filament. Under high vacuum condition, when the current passes through the cathode, the filament temperature can reach up to about 2,000 °C. Hot filament produces electrons, and when the electron energy is higher than the ionization potential of the sample, the sample molecules or atoms are ionized. Ions gain kinetic energy in the electric field and enter into mass analyzer at a certain speed.

In the EI source, the molecules or atoms of analyte lose valence electrons to generate positive ions:



Or capture electrons to generate negative ions:



Generally, the generated positive ions are 10^3 times of negative ions. If not specifically pointed out, conventional MS only studies positive ions. EI energy should be at least equal to the ionization potential of the analyte to make the analyte be ionized to generate positive ions. The ionization potential of elements in the periodic table is 3–25 eV, wherein most of them is less than 15 eV. The ionization potential of organic compounds is 7–15 eV. If EI energy is just equal to the ionization potential of the analyte, all the energy of the electron must be transferred into the analyte to make it be ionized. In fact, the number of molecules or atoms that can obtain all the energy of the electron is limited, and the ionization efficiency is low. Therefore, increasing EI energy is beneficial for increasing the ionization efficiency. To obtain reproducible MS spectra, EI energy is generally 70 eV. But relative high energy will make the remainder energy of the molecular ion greater than the bond energy of some molecules, thereby cracking the molecular ions. Lower ionization voltage can be used to control the number of fragment ions and

increase the intensity of molecular ion peaks. The ionization voltage of common MS instruments can be adjusted in the range of 50–100 V.

A major drawback of EI source is that the solid or liquid sample must be gasified before entering into ion source and therefore not suitable for samples difficult to be volatile or with poor thermal stability. Ion source is a component in MS requiring more attention in operation and maintenance, and many ionization parameters play important roles in obtaining results. Especially when introducing the sample into ion source, the use of solvent is at the core position in the ionization process.

Besides the conventional EI ion source for GC/MS, there are chemical ionization (CI) and field ionization (FI) ion sources. The ionization modes of the latter two are relatively mild, which are soft ionization modes and can obtain molecular weight information of compounds. But the spectral reproducibility is not good as that of EI source, so the use is not so widespread.

2.1.2.2 The Mass Analyzer

Mass analyzer separates charged ions according to mass-to-charge ratio to record the mass and abundance of various ions. According to the types of mass analyzer, MS can be classified into magnetic MS, single quadrupole/MS, triple quadrupole tandem MS, TOF/MS, and ion trap (IT)/MS. Magnetic MS and IT MS are rarely used in GC/MS, and metabolomics research is mainly based on GC-single quadrupole/MS and GC-TOF/MS.

Quadrupole mass spectrometer is currently the most commonly used mass spectrometer. This instrument is characterized by small size, simple structure, low cost, and good performance. Especially for general purposes, it has advantages in cost and performance. The mass analyzer of quadrupole mass spectrometer consists of four parallel electrodes where a radio frequency (RF) voltage supply is connected to adjacent electrodes creating an alternating electric field between the electrodes. Under certain RF voltage and RF frequency, only those ions with certain m/z can successfully pass and reach detector. The amplitude of other ions will constantly increase until they hit the electrode pole and neutralized by electrons to become neutral particles. Quadrupole is also known as mass filter. For principle, see Fig. 2.4.

TOF analyzer is the simplest analyzer, with high scanning speed, very high ion collection efficiency, wide mass range, a resolution of up to 40,000, and dynamic range of up to 10^5 . TOF is very suitable for analyzing very complex metabolomics samples. It can quickly scan the whole spectrum and obtain qualitative information.

The detection principle of TOF technology is that the ion beam is accelerated by high voltage and pushed into the flying tube in pulse, and then “free drift” to reach the detector. Ions with different masses obtain different accelerated speed. Ions with smaller mass have higher speed than ions with larger mass. The arriving time of ion at detector is correlated with ion mass. The principle of TOF detector is relatively simple, and the design art directly affects MS performance. Pusher is used to accelerate particles. In order to get consistent initial kinetic energy and consistent

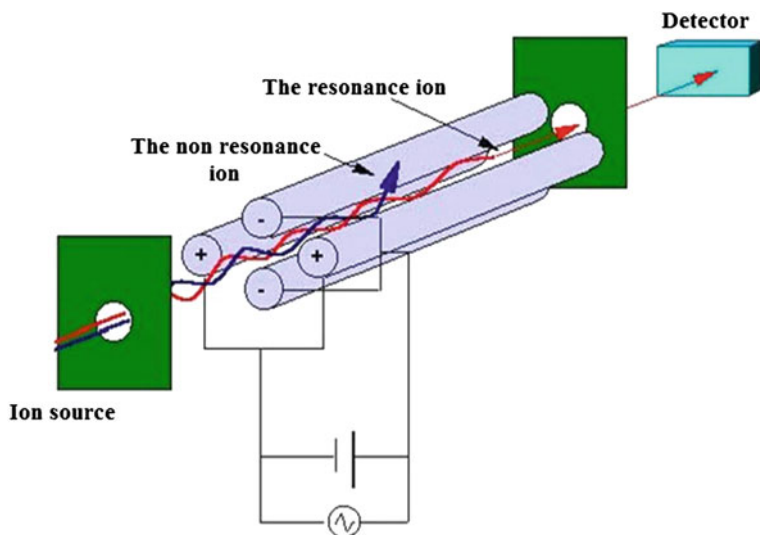


Fig. 2.4 Constitution and analysis principle of quadrupole mass analyzer

initial flight time for charged ions, the ion beam entered into pusher must be very level and narrow to reduce energy diffusion of ions at different directions, which is an important aspect affecting the resolution and accuracy of TOF/MS. The introduction of ion reflector not only increases free flight distance, but also further focuses ions with different kinetic energies. TOF detector requires higher vacuum degree to avoid collisions between ions, and between ions and gas. TOF flight tube requires keeping high stability to reduce subtle changes in mass axis caused by heat expansion and cold contraction, which affects the reproducibility of result and the accuracy of measurement. For typical ion-reflective TOF detector, the flight time of ions with 1,000 Da m/z is not less than $50 \mu\text{s}$ ($50 \times 10^{-6} \text{ s}$). Therefore, the analytic speed of TOF is very fast, requiring the detection speed of ns to ps (10^{-9} to 10^{-11} s). Pusher can reach 20,000 times/s, and the obtained spectrum is the accumulation of multiple ion pushing detection results. To obtain rapid and accurate flight time of ions, the requirement to detector is very high, and TOF-dedicated digital converter (TDC or ADC) is needed. TOF cannot scan ions, cannot store and release ions like ion trap, all the ions are pushed into flight tube one-off, and wait all the ions reach detector. Before next group of ions are pushed into flight tube, the former group of ions cannot be operated. Thus, TOF cannot do selected ion monitoring (SIM) or selected ion recording (SIR), and thus, ion pushing rate affects sensitivity. TOF detector is an ideal match for high-speed GC/MS (advantage in deconvolution) and high-speed HPLC/MS. At present, the quantitative capability of TOF has not yet reached the level of quadrupole, mainly due to the problem previously mentioned for ion pushing in pulse and detector dead time. For principle, see Fig. 2.5.

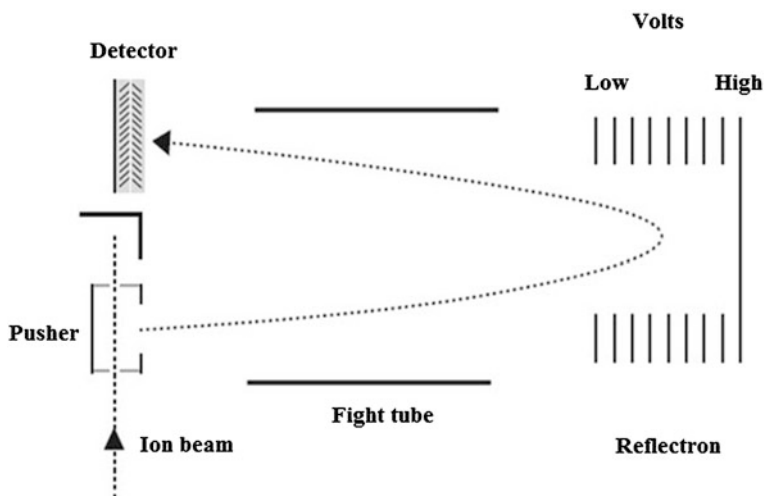


Fig. 2.5 Constitution and analytical principle of TOF mass analyzer

2.1.3 Other Hardware

Besides the above-mentioned components, GC/MS also consists of a pumping system for maintaining required MS vacuum, as well as control electronics and power supplies. The maintenance of high vacuum system is based on two pumping stages to reach a range of 10^{-5} to 10^{-7} hPa, where high-resolution mass analyzer requires higher degree of vacuum. The first stage of pumping system is generally a rotary oil pump backing one or more turbomolecular pumps to achieve minimum pressure for the starting of molecular pump. Generally, these vacuum systems need more care and attention. The second important hardware is the high-voltage power supply system. The stability of MS high-voltage control system is of significance for the resolution, accuracy, and sensitivity of mass spectrometer. Although the high-voltage supply system is very good, they should be maintained over the years, including high-voltage electrical wires and connectors. All modern mass spectrometers are controlled by data analysis system, which not only controls the instrument, but also plays an important role in data analysis. Therefore, data analysis system is known as the fourth leg of mass spectrometer and has equal significance as other accessories.

2.2 Sample Preparation and Analytical Techniques

Small molecular metabolites in the plant will rapidly change when there is a change in external environment. Metabolomics analysis generally requires the sample to be rapidly frozen after collection and stored in the environment below -60°C until the

extraction process begins, to ensure that metabolites are not damaged by enzyme system in the plant. In the extraction process, it should be ensured that metabolites are free from the impact of physical or chemical substances as far as possible. Therefore, it is necessary to keep the ambient temperature constant and avoid using acid or alkali to treat samples. In addition, the use of high-purity solvent for sample pretreatment and the addition of quality control sample in the analytical process are necessary for avoiding the contamination of external samples. The pretreatment of blank sample and target sample in the same batch should be carried out in synchronism. The only difference between the two is that the blank sample does not contain sample materials. In addition, to ensure the reproducibility of biological data, at least six sample replicates are needed.

Prior to sampling, the collection time points should be determined. For example, since the plant leaf is the main part of photosynthesis, the middle time of light cycle should be selected when collecting leaves. For plants in the vegetative stage, collection should be performed before the emergence of the first inflorescence, and the internode at the same part and the same non-aging leaf should be selected. Experience has improved that rapid sampling and rapid quenching are very important. Prior to sample homogenization, all the experimental materials and reagents should be cooled to avoid sample change due to thawing. There are a lot of reports on metabolomics sample pretreatment methods. In addition to referring relevant literatures, experimental methods are also provided by international Web sites maintained by relevant laboratories and institutes. Different laboratories can establish their own methods according to their own needs. Different pretreatment methods may be needed to explore for different experimental materials. Here are just two cases for reference.

Method 1 (Lisec et al. 2006), shown in Fig. 2.6:

Sampling and extraction:

1. 100 mg plant leaf sample is placed in a 2-ml round-bottom tube with screw cap, quickly frozen in liquid nitrogen;
2. Homogenization in a ball mill. Place the sample tube with steel balls in ice-cold ball mill, and grind for 2 min at the vibration frequency of 20 Hz.
Note: The frozen samples can be stored in -80°C refrigerator for no more than 3 months;
3. Add samples into 1,400 μL -20°C precooled methanol, vortex for 10 s;
4. Add 60 μL ribitol (0.2 mg/ml aqueous solution) as an internal standard, vortex for 10 s;
5. Extract using a thermomixer at 70°C , 950 rpm for 10 min;
6. Centrifuge sample at 11,000 g for 10 min;
7. Transfer supernatant to another glass vial;
8. Add 750 μL -20°C precooled chloroform, vortex for 10 s;
9. Add 1,400 μL precooled (-20°C) dH_2O , vortex for 10 s;
10. Centrifuge sample at 2,200 g for 15 min;
11. Transfer 150 μL supernatant (polar phase) into a clean 1.5 ml tube;
12. Transfer another 150 μL supernatant as a backup into a clean 1.5 ml tube;

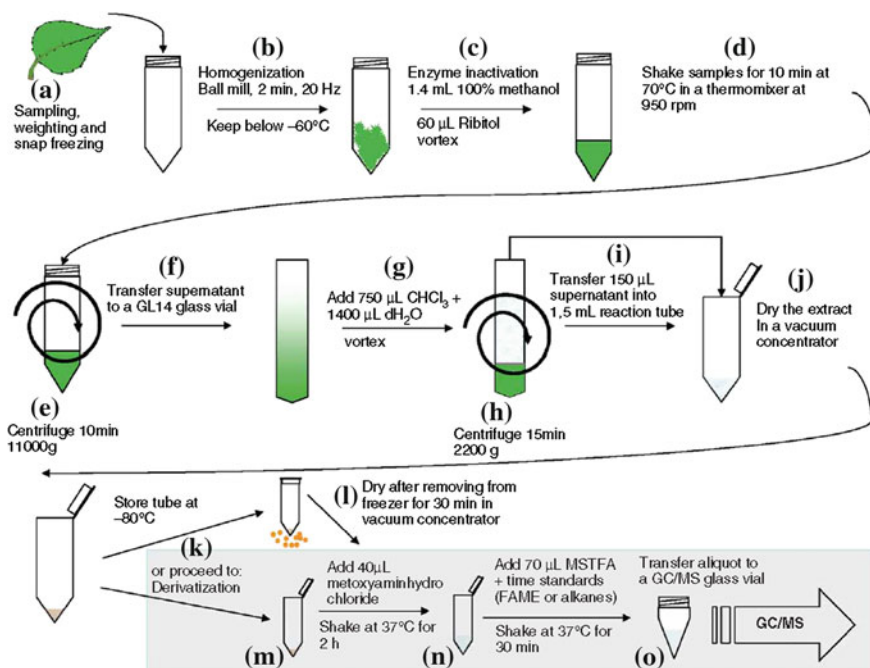


Fig. 2.6 The flowchart of metabolomics sample preparation (Reprinted by permission from Macmillan Publishers Ltd: [Nature Protocols] (Lisec et al. 2006))

13. Drying in a vacuum concentrator at ambient temperature;
14. Fill argon into the sample tube, place all the sample tubes into a sealed plastic bag containing silica-drying agent, and store in -80°C low-temperature refrigerator;
Derivatization:
15. Place the sample tubes which are stored at -80°C in a vacuum concentrator drying for 30 min;
16. Add 40 μL methoxyamination reagent (methoxyamine hydrochloride, 20 mg/ml pyridine solution);
17. Prepare a blank derivatization sample as a control;
18. Shake at 37°C for 2 h;
19. Prepare silylation reagent MSTFA [N-methyl-N-(trimethylsilyl) trifluoroacetamide], add 20 $\mu\text{L}/\text{mL}$ mixture of retention index standard (mixture of saturated fatty acid methyl ester series, FAME ($\text{C}_8\text{--C}_{30}$), and dissolve in chloroform at a concentration of 0.4 mg/mL (liquid standard) or 0.8 mg/mL (solid standard));
20. Add the silylation reagent in step 19 into the sample reaction tube;
21. Shake at 37°C for 30 min;
22. Transfer the derivatized sample to the liner suitable for GC/MS analysis.

Note: The derivatization reagent is extremely toxic, take extra care, wear gloves, and operate under a fume hood.

Key point 1: In the course of derivatization, derivatization reagent is likely to remain in the tube wall or the tube cap of the reaction tube; thus, centrifugation should be performed at each derivatization step;

Key point 2: 15–22 steps are critical. The derivatization reagent in this method is in excess to ensure complete derivatization.

GC/MS instrument model: Agilent G6890 with autosampler, Leco Pegasus IV TOF/MS, MDN-35 capillary column, 30 m length, 0.32 mm inner diameter, 0.25 μm film thickness.

GC injection parameters: injection volume 1 μl , inlet temperature 230 $^{\circ}\text{C}$, splitless injection mode, helium carrier gas, flow rate at 2 ml/min, autosampler injection. For high-concentration samples, split injection mode is suggested, and the split ratio is set to 1: 25.

GC parameters: MDN-35 capillary column (30 m). GC temperature gradient starts at 80 $^{\circ}\text{C}$ and holds constant temperature for 2 min, followed by ramping at 15 $^{\circ}\text{C}/\text{min}$ to 330 $^{\circ}\text{C}$, which is held for 6 min. Transfer line temperature is set to 250 $^{\circ}\text{C}$.

MS parameters: Ion source temperature is set to 250 $^{\circ}\text{C}$, mass scanning range is at m/z 70–600, acquisition rate at 20 spectra/s, MS EI lamp filament starting time at 170 s of chromatographic reagent delay, detector voltage 1,700–1,850 V, MS loss is set to 0, the filament bias current is 70 eV, the instrument is automatically tuned.

Spectral Deconvolution parameters: Deconvolution software ChromaTOF, baseline offset is set to 1 (0.5–1); spectral smoothing for five data points (3–7), peak width 3 s (3–4 s); signal-to-noise ratio (S/N) is 10 (2–15).

Method 2: This method is used for the simultaneously extraction of metabolites, proteins, and RNA (Weckwerth et al. 2004).

Sampling and extraction:

30–100 mg samples of *Arabidopsis* leaves at a developmental stage is harvested.

1. Immediately frozen in liquid nitrogen;
2. Homogenize tissue using the Retsch mill in liquid nitrogen;
3. Add 2 ml precooled (-20°C) extraction solvent methanol: chloroform:water (5:2:2, v/v/v);
4. Mix vigorously at 4 $^{\circ}\text{C}$ for 30 min to precipitate proteins and DNAs/RNAs, and disassociate metabolites from cell membrane and cell wall components;
5. Centrifuge the sample and transfer supernatant;
6. Add 1 ml precooled (-20°C) extraction solvent methanol:chloroform (1:1, v/v) for the second extraction;
7. Centrifuge the sample and transfer supernatant;
8. Combine the organic extracts from the two extraction steps;
9. Add 500 μl water to separate the organic phase and aqueous phase. The methanol–water layer mainly contains sugars, amino acids, and small organic molecules. Chloroform layer mainly contains lipids, chlorophyll, and waxes;
10. Add 1 ml extraction buffer (containing 0.05 M Tris, pH 7.6; 0.5 % SDS; 1 % β -mercaptoethanol) and 1 ml water-saturated phenol to the residue after extraction;

11. Extraction at 37 °C for 1 h;
12. Centrifuge at 14,000 g and transfer supernatant;
13. Separate the phenol phase from supernatant;
14. Precipitate the phenol-phase proteins with ice-cold acetone at -20 °C overnight;
15. Wash three times with ethanol, and dry at ambient temperature;
16. Precipitate proteins remained in the RNA extraction buffer with 200 μ l chloroform;
17. Centrifuge to remove precipitate and separate buffer;
18. Add 40 μ l acetic acid and 1 ml ethanol to precipitate RNA at 37 °C for 30 min;
19. Wash with one volume of 3 M sodium acetate for one time, and wash with one volume of 70 % ethanol for two times;
20. The remaining pellet is dissolved in 100 μ l RNase-free water;
21. Detect the amount and purity of RNA by absorbance at 260 nm and gel electrophoresis in agarose;
Derivatization steps:
22. Evaporate to dry metabolites in the organic phase, and add 50 μ l methoxy-amination reagent (methoxyamine hydrochloride, 20 mg/ml pyridine solution);
23. React at 30 °C for 90 min with continuous shaking;
24. Add 80 μ l MSTFA;
25. React at 30 °C for 30 min;
26. The derivatized samples are stored at ambient temperature for 120 min before injection.

GC-MS instrument model: GC model: HP 5890, Leco Pegasus IV GC-TOF mass spectrometer, 40 m length, inner diameter 0.25 mm, RTX-5 capillary column with a 10 m precolumn.

GC injection parameters: injection volume 1 μ l, inlet temperature 230 °C, splitless injection mode, helium carrier gas, flow rate 1 ml/min, autosampler injection.

GC Parameters: GC temperature gradient starts at 80 °C and holds constant temperature for 2 min, followed by ramping at 15 °C/min to 330 °C, which is held for 6 min.

MS parameters: record 20 spectra per second between m/z 85 and 500. S/N threshold is 20.

2.3 New Technologies and Trends

GC/MS is the most mature chromatography mass spectrometry coupling technology. The one in the spotlight is the comprehensive two-dimensional GC (GC \times GC) developed in 1990s, which is a multi-dimensional chromatography. The separation mechanism is to tandem combine two columns with different stationary phases as a two-dimensional chromatography. A modulator is installed between the two columns, playing a role in trapping and retransmitting. Each fraction separated by the first column must firstly enter into the modulator to be focused in a pulse manner

Table 2.1 The comparison between characteristics of one-dimensional GC/MS and comprehensive two-dimensional GC/MS (Reprinted from Kusano et al. (2007), with permission from Elsevier)

	One-dimensional GC/MS	Comprehensive two-dimensional GC/MS
High throughput	High	Medium
Resolution	High	Very high
Sensitivity	Medium	High
Cost	Low	Medium
Deconvolution	Good	Very good
Data file size (in ASCII and CSV format)	Medium (about 200 MB)	Very large (about 1 GB)

before further separated by the second column. Generally, the second column is relatively short, and the analytical time is short. The comprehensive two-dimensional GC has characteristics such as high resolution and high sensitivity, is one of the currently most powerful separation tools, and is widely applied in the separation of petroleum, tobacco, and pharmaceutical and other complex systems. The key component of comprehensive two-dimensional GC is the modulator between the two chromatographic columns, which is required to completely trap and release the first-dimensional fractions in a very short time period. The modulator is a container connecting the two chromatographic columns, such as the two-stage circular tube. Commonly used method is to trap rapidly at low temperature, such as liquid nitrogen or low-temperature cold trap, and then rapidly warm up to release. Currently, there have been a lot of analytical chemists using comprehensive two-dimensional GC/MS in metabolomics research. Due to the limitation in data analysis methods and the control technique of the modulator, this technology has not yet been widely applied. For the high throughput and reproducibility, usually one-dimensional GC/MS is used to obtain full spectrum, and comprehensive two-dimensional GC/MS is used to obtain more detailed peak information (Table 2.1).

2.4 Common Problems and Cautions

2.4.1 Common Problems in GC/MS-Based Metabolomics Research

1. **GC/MS database.** Metabolite qualification has been a difficult problem in metabolomics. Compared with other chromatography and mass spectrometry coupling technologies, GC/MS has a large number of databases, such as NIST database. However, most of the GC/MS peaks still cannot be resolved by existing commercial MS database. The structures of metabolites in plant are complex, and so metabolomics-targeted database is needed. Many databases, such as human metabolomics database, plant metabolomics database, and

species-specific metabolomics database, are being continuously established and improved. Scholars in Max Planck Institute for Molecular Plant Physiology suggest to accumulate metabolomics qualitative data from everyday metabolomics analysis and to establish database based on retention index and MS data. This method is simple and feasible, and the qualitative effect is more accurate than MS (Wagner et al. 2003; Schauer et al. 2005; Strehmel et al., 2008).

2. **Deconvolution.** The aim of deconvolution is to parse overlapping coeluted peaks and obtain the MS peak of single pure substance. Currently, there are only few good deconvolution softwares, free software such as AMDIS, and commercial software such as ChromaTOF and AnalyzerPro. The deconvolution effects of the three softwares were compared with 36 endogenous standard metabolites mixture. The 36 substances were prepared in standard solutions with five different concentrations. The first group is 500 μM , the second group is 350 μM , the third group is 150 μM , the fourth group is 50 μM , one half of the fifth group is 500 μM , and another half of the fifth group is 50 μM . The 36 standard substances were derivatized and produce 51 metabolites and derivatives. AMDIS performed deconvolution and detected peaks of all the metabolites; ChromaTOF did not detect the 8 metabolites and derivatives in the fourth group; most of metabolites and derivatives were not detected by AnalyzerPro; ChromaTOF and AnalyzerPro produced many false-negative results. However, AMDIS detected the most false-positive peaks, obtaining as much as 522–750, while ChromaTOF obtained 78–173 false-positive peaks and AnalyzerPro obtained fewer false-positive peaks. In addition, the number of deconvoluted peaks and the accuracy of deconvoluted spectra are closely correlated with metabolite concentration, i.e., if the concentration is decreased, the number of deconvoluted metabolites is significantly decreased, and the number of metabolites can correctly match the database is decreased, even some peaks cannot be deconvoluted because the S/N is too low (Lu et al. 2008) (Table 2.2).
3. **Multi-peak phenomena.** In addition to the large number of false-positive peaks caused by the algorithms of different deconvolution software, multiple peaks can

Table 2.2 The comparison of deconvolution effects of different deconvolution softwares

	Deconvolution software	The first group	The second group	The third group	The fourth group	The fifth group
Peak number after deconvolution	ChromaTOF	173	161	121	78	162
	AMDIS	720	620	529	522	720
	AnalyzerPro	67	49	38	14	42
Peak number undetected by deconvolution	ChromaTOF	0	0	0	8	0
	AMDIS	0	0	0	0	0
	AnalyzerPro	2	9	17	38	19
Metabolite number correctly matched with deconvolution	ChromaTOF	37	31	28	14	27
	AMDIS	32	30	20	8	26
	AnalyzerPro	28	24	14	5	18

be produced in sample preparation, extraction, derivatization, and analysis process, especially the variation of derivatization has great impact on metabolic profile. The multi-peak phenomena refer to one metabolite producing multiple peaks, which can be caused by sample degradation, by-products formation, and the introduction of exogenous contaminants. The multi-origination phenomena mean one peak has multiple origins (precursors) (Xu et al. 2010).

Reasons causing the multi-peak phenomena in the derivatization process of GC/MS include:

1. **The multi-peak phenomena are generated in the derivatization process**, including ① forming by-products. The silylation derivatization process is to silylate metabolites, but some functional groups, such as aldehyde, amino, carboxyl, ester, ketone group, and phenolic hydroxyl group, may form a plurality of products. In addition, derivatization reagents, organic solvent impurities, and plastic tube contaminants may also cause by-products. All these non-specific products are referred to as artifacts; ② incomplete derivatization. Many compounds contain a plurality of reactive groups available for derivatization, and when the amount or derivatization time of derivatization reagent is not enough, incomplete derivatization will be produced.
2. **The conversion of metabolites structure.** Geometrical isomers of metabolites may lead to the multi-peak phenomena, for example, linear and cyclic D-glucose can be converted with each other in the solvent. Usually, glucose in the solvent shows at least five different tautomeric forms [such as α -D-glucopyranose (62 %), β -D-glucopyranose (38 %), α -D-glucofuranose (trace), β -D-glucofuranose (trace), and linear-D-glucose (0.01 %)], and all of these tautomeric forms maintain a dynamic balance, their contents are impacted by solvent composition, temperature, and pH, and complex chromatograms will be produced after BSTFA derivatization. Again, inositol alcohol has nine different stereoisomers; arginine can be converted into ornithine with MSTFA derivatization (37 °C, 20 min).
3. **Metabolites will be degraded in the extraction, derivatization, and GC/MS analysis process.** Thermal non-stable compounds prone to be thermally degraded, resulting in multi-peak and multi-origination phenomena. Xu et al. (2001) detected two groups of structural and biological-related substances and found that even the thermally stable compounds, such as phosphorylcholine (PC), 1,2-diacetylglycerol-3-phosphate ester (DAG), and hemolysis phosphocholine (LPC), can generate multi-peak phenomena; structurally similar compounds can produce the same peak (multi-origination phenomena); glycerol, phosphoric acid, fatty acids, and some lipid fragments may be the cracking fragments of these structural similar compounds (including free type and bound type).

2.4.2 Notes on GC/MS-Based Metabolomics Research

- ① Derivatization reagent methoxyamine hydrochloride need temporary preparation, MSTFA need to dry stored in 2–8 °C, avoid absorbing moisture in the air.
- ② Samples must be randomly injected to eliminate systematic errors.
- ③ Set control samples and blank samples (e.g., reagent blank, method blank).
- ④ Low loss injection pad and low loss injection system are important.
- ⑤ Original GC/MS data are transferred to the server, and long-term data preservation should use DVD backup or transfer to a server mirroring system.
- ⑥ System background subtraction: Background peaks of plasticizers, phthalates, and silylation reagent and column bleed as well as the water peak of derivatization reagent should be subtracted.

2.5 Prospects

The data processing in GC/MS-based metabolomics has been a bottleneck hindering the development of the platform, such as the routine applications of data modeling tools in bioinformatics and systems biology analysis.

Challenges in the modernization of GC/MS platform are the parallel control techniques of multiple samples, including ① the automation of sample preparation, sample pretreatment, and the high-throughput and reproducible technology after obtaining data; ② the integration of metabolomics data and other omics data, for example, the analysis of samples by using integrated analysis technique platform and combining proteomic data and transcriptomic data; ③ the spectral recognition of trace compounds or signal molecules among a group of metabolites; ④ the combination of full spectrum analysis and flow analysis; ⑤ reproducible quantitative results, clear system nomenclature, and the comparability of data obtained by different laboratories using different GC/MS platforms; and ⑥ the last but not the least, which may be the most important challenge to all the metabolomics analysis, that is, the structural identification of metabolites in complex spectra of metabolites.

References

- Dalgliesh CE, Horning EC, Horning MG, Knox KL, Yarger K. A gas-liquid-chromatographic procedure for separating a wide range of metabolites occurring in urine or tissue extracts. *Biochem J.* 1966;101:792–810.
- Drozdz J. Chemical derivatization in gas chromatography. *J Chromatogr Libr.* 1981.
- Kusano M, Fukushima A, Kobayashi M, Hayashi N, Jonsson P, Moritz T, Ebana K, Saito K. Application of a metabolomic method combining one-dimensional and two-dimensional gas chromatography-time-of-flight/mass spectrometry to metabolic phenotyping of natural variants in rice. *J Chromatogr B.* 2007;855:71–9.

- Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat Protoc.* 2006;1:387–96.
- Lu H, Dunn WB, Shen H, Kell DB, Liang Y. Comparative evaluation of software for deconvolution of metabolomics data based on GC-TOF-MS. *Trends Anal Chem.* 2008;27:215–27.
- Schauer N, Steinhäuser D, Strelkov S, Schomburg D, Allison G, Moritz T, Lundgren K, Roessner-Tunali U, Forbes MG, Willmitzer L, Fernie AR, Kopka J. GC-MS libraries for the rapid identification of metabolites in complex biological samples. *FEBS Lett.* 2005;579:1332–7.
- Strehmel N, Hummel J, Erban A, Strassburg K, Kopka J. Retention index thresholds for compound matching in GC-MS metabolite profiling. *J Chromatogr B.* 2008;871:182–90.
- Toyo'oka T. Modern derivatization methods for separation science. New Jersey: Wiley; 1999.
- Villas-Boas SG, Roessner U, Hansen MAE, Smedsgaard J, Nielsen J. Metabolomics analysis an introduction. New Jersey: Wiley; 2007.
- Wagner C, Sefkow M, Kopka J. Construction and application of a mass spectral and retention time index database generated from plant GC/EI-TOF-MS metabolite profiles. *Phytochemistry.* 2003;62:887–900.
- Wang ZF, Yang SM, Wu MT, Yue WH. Chromatography coupling techniques. Beijing: Chemical Industry Press; 2001.
- Weckwerth W, Wenzel K, Fiehn O. Process for the integrated extraction, identification and quantification of metabolites, proteins and RNA to reveal their co-regulation in biochemical networks. *Proteomics.* 2004;4:78–83.
- Xu GW, Ye F, Kong HW, Lu X, Zhao XJ. Technique and advance of comprehensive two-dimensional gas chromatography. *Chromatography.* 2001;19:132–6.
- Xu F, Zou L, Ong CN. Experiment-originated variations, and multi-peak and multi-origination phenomena in derivatization-based GC-MS metabolomics. *Trends Anal Chem.* 2010;29:269–80.

Plant Metabolomics

Methods and Applications

Qi, X.; Chen, X.; Wang, Y. (Eds.)

2015, VIII, 319 p. 144 illus., 73 illus. in color., Hardcover

ISBN: 978-94-017-9290-5