18
Gas Chromatography–Mass Spectrometry for Residue Analysis: Some Basic Concepts

H. F. De Brabander, K. De Wasch, and S. Impens
University of Gent. Merebeke, Belgium

R. Schilt
TNO Nutrition and Food Research, Zeist, The Netherlands

M. S. Leloux
State Institute for Quality Control of Agricultural Products, Wageningen, The Netherlands

1. INTRODUCTION

A residue may be defined as a trace of a component, that is present in a matrix after some kind of administration. The matrix may be anything in which a residue may be present, trapped, or concentrated (meat, urine, feces, liver, etc.). There is no general agreement upon the concentration level of a trace. However, the ppb level (μg kg⁻¹) in which residues are present may certainly be considered as a trace or even an ultratrace level. The impact of results in the field of residue analysis in veterinary food inspection is increasing. This is mainly due to the more severe legislation for some residues (class A or forbidden components). Therefore, the highest standard is required from the control methods used. One of the aspects is the specificity of the method. Specificity is defined as the ability of the method to distinguish between the analyte being measured and other substances. The method must have the ability to distinguish between the analyte being measured at trace or ultra trace concentrations and other substances (possibly) present at 10 to 10,000 times higher concentrations. For the control of registered veterinary drugs, next to the determination of the identity of the analyte (qualification), a good quantification is necessary.

441
A residue analysis procedure consists of three distinct steps: First, the analyte has to be extracted from the matrix. Then, the extract is freed from as many interfering products as possible. The third step is the identification and eventually the quantification of the analyte. In modern residue analysis chromatographic techniques are very important. Next to thin-layer chromatography (TLC) and liquid chromatography (LC), gas chromatography–mass spectrometry (GC–MS) is the technique most frequently used. The choice of a very specific detection technique (GC–MS, MS–MS, MS²; high-resolution (HR) GC–MS) could result in a reduction of sample pretreatment, which is, at first sight, an advantage. However, it is not always conceivable for a control laboratory to use all these techniques in routine analyses (e.g., due to high instrument costs). Moreover, as the number of substances of different parts to screen is high and not limited (in the future), one detection technique will never be as specific for all compounds. Another option is high-performance liquid chromatography (HPLC) fractionation. This procedure results in several purified fractions, each containing a limited number of analytes and matrix components. Each HPLC fraction may be analyzed with a specific technique, e.g., GC–MS, and if necessary with different techniques, e.g., TLC, LC–MS.

The choice of the analytical strategy must always be seen in the light of the interpretation of analytical results. Inspection services are interested mainly in a YES/NO answer: Has this animal been illegally treated? Is the concentration of the residue higher than a certain value, e.g., maximum residue limit (MRL), maximum permitted limit (MPL)? In fact, all questions may be subsumed in one major question: Is the law violated? When the answer to that question is YES, actions are taken. Different kinds of actions (rejection of animals, inspection at the farm, etc.) may be performed.

Laboratories, on the other hand, evaluate crude analytical data on the basis of predefined criteria for all relevant parameters; e.g., signal-to-noise ratios (S/N ratios), deviations from observed and target values for reference materials. For residue analysis, such minimum quality criteria are part of EC legislation [1]. Based on these criteria, analytical data are transformed into YES/NO answers. Moreover, in most countries, a system of first and second analysis is used. When the final answer of the first analysis is YES, the legal action is suspended until the results of the second analysis, if any, are known.

2. GAS CHROMATOGRAPHY–MASS SPECTROMETRY APPARATUS IN RELATION TO RESIDUE ANALYSIS

For the routine analysis of residues in meat-producing animals there is an increasing use of coupled techniques. In most cases, low-resolution MS coupled to a chromatographic separation is used. In low-resolution GC–MS, two approaches may be distinguished: (1) The use of instruments, that are able to monitor a whole chromatogram in the full-scan mode, e.g., with 1 spectrum per second, without loosing detection power, e.g., instruments based on ion-trap technology, and (2) linear quadrupole instruments using the full-scan mode for high amounts of analyte, and selective ion monitoring (SIM) for detecting low concentrations (<1 to 10 ng).

Both techniques have their pros and cons, and also their own fans. In full-scan GC–MS, a complete spectrum of each point of the chromatogram as well as all kinds of ion chromatograms can be generated afterwards with a data system. Identification of the analyte by library search may be performed. The result of this search is expressed as a figure, which reflects the fit between the standard and the sample spectrum.

With SIM, a limited number of ions (2 to 4 ions) are monitored during a selected time interval. The presence of the analyte is determined by the presence of these "diagnostic" ions in the correct retention time window and within the correct abundance ratio between certain limits, e.g., ±20% in chemical ionization (CI) and ±10% in electron ionization (EI). In the EU guidelines, the monitoring of 4 ions is mandatory. In practice, the monitoring of 4 ions at low concentrations (<1 ppb) does not always give satisfactory results: some ions disappear at lower concentrations or the ratio between the ions is not reproducible. Moreover, not all analytes in a multiresidue method give enough diagnostic ions with one derivatization method.

Gas chromatography–tandem mass spectrometry (GC–MS–MS) has been available for some time on the larger instruments. However, in most cases, these machines are too expensive for use in field laboratories performing residue analysis. In the mid-1990s, benchtop GC–MS–MS-based ion-trap technology was introduced. In this way, smaller and also less expensive instruments could be constructed. Instead of the classical MS–MS in space, using an instrument consisting of three quadrupoles in series, the MS–MS experiment takes place in one ion trap, in function of time. One ion (the precursor ion, also called parent ion) is isolated and stored in the ion trap. Subsequently, the precursor ion is fragmented by increasing the speed of the ion and by inducing collisions with the helium molecules present in the trap. The product ions (also called daughter ions) are scanned out of the trap resulting in a product (or a daughter) ion mass spectrum. However, it must be noted that the instrument is not limited to only one stage of MS–MS; a target product ion may act as a new precursor ion with formation of secondary product ions (also called granddaughter ions), and so on. Therefore, the nomenclature MS² is preferred to MS–MS throughout this chapter. In practice, the absolute number of ions decreases with each MS–MS step, and therefore MS^n is a practical limit in GC–MS^n.
3. POSSIBLE INTERFERENCES IN GAS CHROMATOGRAPHY—MASS SPECTROMETRY ANALYSIS OF RESIDUES

Gas chromatography—mass spectrometry is often considered as a technique with which no identification mistakes could be made, apart from mistakes due to cross contamination. This may be true or nearly true in major and minor component analysis, but this is certainly not the case in residue analysis.

In extracts of biological material (e.g., urine, meat, feces), a large variety of components in a large variety of concentrations are present. An unknown and variable amount of these matrix components are coextracted with the analyte and introduced into the chromatographic system and the mass spectrometer. Interference between these matrix components, possibly present at relatively high concentrations (ppb range, mg kg⁻¹ or higher), and analytes, present at very low concentrations (ppb range, μg · kg⁻¹), is possible and should be avoided.

Interferences mostly result from coeluting peaks or from background noise. The mass spectrum obtained is a mixture of two mass spectra and false interpretation is possible. Isotope interference is another possibility [2]. This phenomenon may occur with any isotope; ¹³C is used here as an example. Carbon has two natural isotopes: ¹²C and ¹³C with a ratio of 98.9 to 1.1 (the exact figures are rounded for simplicity). In residue analyzes, three important parameters should be taken into account:

1. The very large difference in concentration between the analytes and the matrix components.
2. Analytes and/or interferences are mostly organic molecules containing a relative high amount of carbon atoms.
3. Interferences (most likely) may have analogous structures as the analytes.

In a quadrupole instrument using the SIM mode, many interferences are not observed by the highly selective use of the detector. In an ion-trap instrument, high concentrations of coeluting molecules may influence the ionization time of analytes and so the detection limit. This may cause the following phenomena: false negative and/or positive results, and wrong quantification.

3.1. False Negative Results

In SIM, the diagnostic ions of the analyte must be present in the correct relative intensities (±20% for CI and ±10% for EI) (EC criteria 93/256) [1,3,4]. The higher the number of diagnostic ions monitored, the higher the specificity of the method (fewer false positives) but also the higher the chance of false negative results when the identification criteria are applied (too) strictly. The relative intensity of the ions may be disturbed by background noise and coeluting substances. In Figure 1, the relative abundance of the ion m/z 440 in the sample spectrum is partly due to the analyte (see spectrum standard) and partly due to an interference. The ratio between m/z 440 and 425 is out of range (normal range between 41 and 51). According to the quality criteria, the sample has to be declared negative although the analyte is clearly present. The conclusion "negative" (NO answer) will be the same using SIM or full-scan mode. At full-scan identification of components, the same parameters will influence the (reverse) fit of the spectrum and also disturb the visual comparison. However, in full-scan mode, the presence of interferences (a possible coeluting peak) will be noted more readily.

The interference may be due to a molecule having a molecular or fragment ion equal to one of the diagnostic ions of the analyte. However, interference is also possible with molecules present in a much higher concentration containing ions, that are 1 or 2 amu lower than the diagnostic ions. The latter is of particular interest on shoulder peaks.

3.2. False Positive Results

These may occur when SIM signals are generated by interfering molecules. Nortestosterone (NT), for example, is a well-known anabolic steroid used in cattle fattening. The same component is also endogenous in various animal species [5–7]. For determination of its ditrimethylsilyl (di-TMS) derivative of the β-form, three ions are monitored: m/z 418 (100%), 403 (20%), and 328 (35%). In the urine of pregnant cows, estradiol (E2 in the α-form) is present in concentrations 10⁴ to 10⁵ times higher than the concentration of NT found after illegal application of the drug. It was demonstrated that the di-TMS derivatives of α-E2 and β-NT are not well separated under the chromatographic conditions used in most laboratories. The structures of the di-TMS derivatives of both components are

Figure 1: Formation of false negative results by SIM by interferences that disturb the relative ratios of the diagnostic ions.
4.1. Avoiding False Negatives

False negatives by the loss of the analyte during the cleanup, derivatization, or injection should be monitored by using internal standards. Deuterated analogues of the target components are most suited for this purpose because the behavior of the deuterated component in the extraction procedure, cleanup, and detection is very like that of the analyte. However, their availability is, in number as well as in quantity, is limited. Deuterated components are also very useful for quantification and to prove the absence of a certain analyte in a certain sample (real negatives): the signal of the deuterated (heavy) component is present and that of the light component is not. Moreover, deuterated components may be used for balancing the ion ratios.

False negative results due to not fulfilling the quality criteria, e.g., disturbance of the normal peak ratios of the ions from the analyte by one or more interferences or missing peaks, should be dealt with in another way. The analyst should be aware that the statistical possibility of its occurrence is high. Instead of declaring a sample immediately negative (NO answer) because the ratio of one of the ions is not within the range proposed by the quality criteria, other elements should be added to the analysis. Possible solutions are reinjecting the same derivative on another column, using other derivatization reagents or techniques, using GC–MS, performing a second analysis with a different method, etc.

Although each technique on its own may not completely fulfill the quality criteria (e.g., by disturbed ion ratios) the combination of these techniques may give enough analytical evidence to prove the presence of the analyte "beyond any reasonable doubt." In the new EC criteria, to be published in 2000 [10], a system of identification points is proposed to interpret the analytical data. For the confirmation of substances listed in Groups A and B of Annex I of Council Directive 96/23/EC, minimums of 4 and 3 identification points, respectively, are required. In Table 1, the number of identification points that each of the basic mass spectrometric techniques can earn is given.

However, in order to qualify for the identification points required for confirmation:

1. A minimum of at least one ion ratio must be measured.
2. All measured ion ratios must meet the criteria described above.
3. A maximum of three separate techniques can be combined to achieve the minimum number of identification points.

Of course, each ion may only be counted once.

High-performance LC coupled with full-scan UV diode-array detection (DAD), with fluorescence detection, or to an immunogram, or two-dimensional TLC coupled to spectrometric detection are techniques that may contribute a
Table 1 Number of Identification Points That Each of the Mass Spectrometric Techniques Can Earn

<table>
<thead>
<tr>
<th>MS technique</th>
<th>Identification points earned per ion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-resolution mass spectrometry (LR)</td>
<td>1.0</td>
</tr>
<tr>
<td>LR-MSn precursor ion</td>
<td>1.0</td>
</tr>
<tr>
<td>LR-MSn transition products</td>
<td>1.5</td>
</tr>
<tr>
<td>High-resolution mass spectrometry (HR)</td>
<td>2.0</td>
</tr>
<tr>
<td>HR-MSn precursor ion</td>
<td>2.0</td>
</tr>
<tr>
<td>HR-MSn transition products</td>
<td>2.5</td>
</tr>
</tbody>
</table>

maximum of one identification point (for substances in Group A) provided that the relevant criteria for these techniques are fulfilled.

This approach allows the use of the classical low-resolution GC–MS in the SIM mode (four ions and three correct ratios) for laboratories using older equipment as described in 93/256 [1]. Combinations of modern methods, e.g., GC–MS yielding two ions (and one ratio) in combination with LC–MS yielding two ions (and one ratio), are also allowed. New instruments brought on the market may be included very easily into the system in the future. In some cases, earning four identification points without the power of MS² or HRMS may pose a problem. Therefore, an extra point can be earned with HPLC and HPTLC. Some molecules, e.g., trenbolone and methyltestosterone, yield a very typical fluorescence spectrum with identification power certainly equivalent to an ion in mass spectrometry [11]. On the other hand, the analyte should be restricted to four identification points only. Depending upon the analyte’s analytical experience and skill, the equipment available, and the impact of the analytical result, many more points (evidence) can be gathered (see below).

4.2. Avoiding False Positives

As demonstrated above, false positives may result from the presence of diagnostic ions that do not originate from the analyte but are generated by one or more interferences present at high concentration in the final extract. The fact that the correct ions (with correct ratios) can be produced from the interfering (endogenous) compounds is transparent to the analyst when using the GC–MS in the SIM mode.

With instruments that are not able to take a full-scan run at low concentration, the following strategy can be recommended. In the case of a positive signal, a second full-scan run on the same sample is performed in order to exclude the presence of isotope-generating peaks at the retention time of the analyte. The absence of substantial concentrations of isotope peak generators in the full-scan mass spectrum has to be considered as a quality criterion. The SIM mode could also be used for screening purposes only and suspect samples rechromatographed and fully identified with the other system.

Isotope interferences may be avoided by using instruments capable of operating in the full-scan mode at low concentration levels. The quality criteria (3 to 4 ions) may be extended by using full-scan spectrum matches between the sample spectrum and a (home made) library spectrum. However, in order to obtain a good fit, the sample spectrum should be (reasonably) free of interfering peaks. At lower concentrations (how low: depending upon the cleanup used) the diagnostic ions become less abundant in comparison with the background ions (Fig. 3).

The use of MS–MS is a well-known way to reduce background. If a positive signal is obtained, the sample is rechromatographed (eventually several times) in the MS–MS mode in order to obtain more information. In most cases, there is no need for an extra extraction and cleanup; usually only 1 µl out of 25 µl is injected in the GC and there is enough sample left for the extra injections. Full MS² spectra on each diagnostic ion, e.g., three ions, could be taken during

![Concentration: 10 µg kg⁻¹](image1)

![Concentration: 1 µg kg⁻¹](image2)

![Concentration: 0.1 µg kg⁻¹](image3)

Figure 3 Diagnostic ions in relation to background ions at various concentrations (theoretical example).
Residue Analysis

earned in only one extra GC run using MS$^\circ$. In this technique, a product MS$^\circ$ ion serves as precursor ion of the MS$^{+1}$ run. In Figure 5, the difference between MS$^\circ$ and different MS$^2$ experiments is shown. The latest versions of ion-trap software enable the upgrade from MS$\times$MS to MS$^\circ$ (up to MS$^3$) on the basis of software changes only.

Another approach to obtaining more analytical evidence (and more identification points) is higher mass accuracy. High-resolution GC–MS with magnetic sector instruments has been available for a long time. However, for most laboratories, these instruments are too expensive. Recently, higher mass accuracy may be obtained with more affordable benchtop instruments, based on time-of-flight technology.

5. QUALITATIVE AND/OR QUANTITATIVE ANALYSIS

The classical difference between qualitative and quantitative methods is described below. In quantitative analysis, the result is expressed as a figure (e.g., 75 μg kg$^{-1}$). Quantitative analysis is necessary for the determination of residues of components that may be present in food at maximum allowable concentration, e.g., at the MRL or MPL. The method must be able to establish whether the concentration of the analyte is lower or higher than that limit. The methods must have a limit of quantification (LOQ) that is lower than the MRL. Recently, consensus was reached on the fact that analytical methods to be used for controlling an MRL should have a LOQ of (at least) 0.5 MRL. When the result obtained is higher than the MRL (and action may follow), quality criteria must be used for the qualification of the residue. For values much lower than the MRL, qualitative errors play a less important role unless the (screening) method should miss the analyte completely.

Qualitative methods do not produce figures: the results are expressed as YES/NO answers. These methods could be used for residues of forbidden substances (e.g., products with an estrogenic, androgenic or gestagenic action). However, qualitative methods always have a semi-quantitative character: the minimum amount (a quantitative figure) of analyte to discriminate a signal from the background. In practice, a sort of action limit is applied for the determination of the presence or absence of the analyte.

The difference between qualitative and quantitative methods in residue analysis is not as simple as described above. A method is not necessarily quantitative because a figure is produced, but only when that figure fulfills certain criteria of accuracy (trueness and precision). However, too many people neglect that item and consider a figure produced by any instrument automatically as a quantitative result. An illustration of that (normal) human behavior is the common negligence of the rounding of figures: the number of significant figures must reflect the preci-
In residue analysis, the coefficient of variation increases with decreasing concentration according to the so-called Horwich curve [12]. In most cases, according to the rounding rules, only the magnitude of a result could be given, e.g., 2.10, which is something nobody likes.

For registered veterinary drugs, quantification is only necessary in a small concentration range. The analytical method is validated in the small range close to the MRL, e.g., with an MRL of 50, the range could be from 25 to 75 (Fig. 6). Therefore, the fact that a result (not a method) is qualitative or quantitative does not depend upon the method, but upon the result of the analysis. In Figure 6, the concentration axis is divided into three parts: (1) the quantitative part around the MRL (25 < x < 75; the validation range), (2) a qualitative part in the range of x < 25, and (3) a qualitative part in the range x > 75.

The final result of this quantitative method is not expressed as a figure, but again as a YES/NO answer, i.e., above or below the MRL. In the context of residue analysis, a YES/NO answer is not to be interpreted in terms of “positive” or “negative,” but in more general terms of “violation” or “nonviolation.” The expressions “positive” and “negative” should in principle be avoided: a negative result does not necessarily mean that the residue is absent, because the residue level may be under the MRL or the component may be endogenous in a certain species. The expression “positive” can also be confusing: Recently, a politician declared that the results of the analysis were positive because no residues were found.

In addition, the method of quantification is very important. When MS is used, the quantitative result may be calculated in various ways. When the full signal of the sample versus the signal of the internal standard (calibrated against a series of standards) is used (Figure 7), the quantification is not very reliable because the sample spectrum could be different from the standard spectrum: some ions could be missing and/or some ratios between ions may be disturbed.

Alternatively, the sum of a number of diagnostic ions (including MS ions), the most important ion or an algorithm including the correct ratios of the ions may be used. All these methods of quantification will give different results and can be the cause of contradictions. In Figure 7, for example, peak 2 of the sample is distorted by an interference: taking the blind sum of all peaks will result in false quantification. A correction for the correct peak ratios should be made. The best method of quantification is the use of deuterated internal standards. However, as mentioned before, their availability in number as well as in quantity is limited.

6. CONCLUSIONS

Gas chromatography–mass spectrometry is a very powerful technique for the analysis of residues in veterinary products. During the 1990s, the classical SIM mode was increasingly complemented by other techniques offering full-scan mode at low concentrations, e.g., GC–MS–MS, GC–MS/MS and HR-GC–MS. However, it is still very dangerous to consider GC–MS as an absolute error-less technique. As in any other analytical technique, false positive and false negative results as well as false quantification can be obtained. However, when the analyst is aware of the possible causes of these errors, the application of some simple rules and the investment of a little more time in analyzes may prevent most of these mistakes. A very good strategy is to consider a first “violation” result as just a “suspect” result and to repeat the complete analysis immediately within the laboratory. In that second analysis, deuterated internal standards preferably should be used. Only when the two successive results match, qualitatively as
well as in magnitude of concentration, is the result ready to leave the laboratory. Even then, an open mind for the followup of the results of a second analysis in an independent laboratory is necessary.

REFERENCES


19

Applications of Gas Chromatography–Mass Spectrometry in Residue Analysis of Veterinary Hormonal Substances and Endocrine Disruptors

R. Schilt
TNO Nutrition and Food Research, Zeist, The Netherlands

K. De Wasch, S. Impens, and H. F. De Brabander
University of Gent, Merebeke, Belgium

M. S. Leleux
State Institute for Quality Control of Agricultural Products, Wageningen, The Netherlands

1. INTRODUCTION

1.1. Anabolic Agents and Endocrine Disruptors

In Europe, the word “hormones” has a very bad reputation because of the possible danger for public health of residues of some of these products in foodstuffs of animal origin. Toxicologists have demonstrated that DES (diethylstilbestrol, a synthetic estrogen) is a potential carcinogen [1–3]. In human medicine, analogous experiences with DES were found (the so-called DES-daughters) [4]. Recently, several cases of poisoning have occurred in Spain and France due to the consumption of liver from animals treated with clenbuterol [5,6]. Moreover, some environmentally persistent alkyl-phenolic compounds (such as nonylphenol) and perhaps other chemicals show estrogenic activity [7]. These “environmental” estrogens are brought up in relation to the decreasing quality of human sperm and regarded as an assault on the male.