Mass Spectrometric Analyses in Agriculture and Natural Product Research

Douglas F. Barofsky
Department of Chemistry, Oregon State University, Corvallis, Oregon USA

April, 1999

Mass spectrometry is that branch of analytical science devoted to (1) developing and using instruments to determine the masses of atoms and molecules, (2) deducing the identities or abundances of atoms in physical and biological samples, and (3) elucidating the structural properties, deducing the identities, or determining the concentrations of molecules in physical and biological samples. Mass spectrometry is powerful and versatile because it can be used to qualitatively and (in many cases) quantitatively analyze (1) all elements and essentially all compounds, (2) samples in all states of matter, (3) molecules whose masses range from less than 1 u to greater than 10^6 u, and (4) sample amounts as small as 1 attomole. Ultimately, access to some form of mass spectrometric analysis becomes essential to any branch of science in which study is conducted at the molecular level.

Modern agriculture is increasingly based on chemical, biochemical, and molecular biological knowledge. Hence, agricultural sciences and practices are becoming increasingly dependent on powerful analytical tools, such as mass spectrometry, for investigating and monitoring the characteristics, transport, modification, and actions of a broad range of inorganic and organic chemicals. The field of natural products comprises bioprospecting (i.e., discovery and identification of compounds), ecology (i.e., evaluation of the environmental context in which the compounds are found), and all branches of pharmacology (i.e., pharmacognosy, pharmacodynamics, pharmacogenetics, and pharmacokinetics). These endeavors all involve analytical chemistry in general and mass spectrometry in particular.

This tutorial briefly introduces some basic qualitative and quantitative mass spectrometric methods and illustrates them with examples drawn from research in agriculture and the field of natural products. For a fuller treatment of the fundamentals of mass spectrometry, the interested reader is referred to J.T. Watson's excellent book *Introduction to Mass Spectrometry,* 3rd Edition, Lipincott-Raven, New York, 1997.

The Mass Spectrometer

A mass spectrometer must perform three functions: (1) convert a sample into gas-phase ions, (2) sort (or disperse) the ions according to their respective ratios of mass to ionic charge number (m/z), and (3) produce a record of relative ion abundance versus m/z (i.e., a mass spectrum). The sections of a mass spectrometer responsible for performing these three operations as well as necessary auxiliary components, such as an inlet system for introducing samples into the instrument, are shown schematically in Fig.1. Actual mass spectrometers come in a wide range of shapes and sizes because the mechanisms for producing gas phase ions, dispersing ions, and detecting ions each exist in a variety of forms and there are many ways to combine these different forms in accordance with the general scheme shown in Fig.1.

Examples of mass spectra data produced by three different ionization processes are shown in this tutorial. In order of appearance, these are fast atom bombardment (FAB), electrospray ionization (ESI), and electron impact (chemical ionization?) (EI). FAB is a technique whereby gas phase ions of organic compounds are created in a vacuum chamber by firing 5-10 keV atoms at an analyte dissolved in a viscous liquid like glycerol. In FAB, positive ions are formed by the addition of one or more protons, alkali ions, or other cationic species originating out of the sample matrix, and negative ions are produced by abstraction of one or more protons or other anion. In ESI, an aqueous or organic solution of a sample at atmospheric pressure emerges from the end of an electrostatically charged capillary-needle as a spray of tiny, highly charged droplets. The droplets

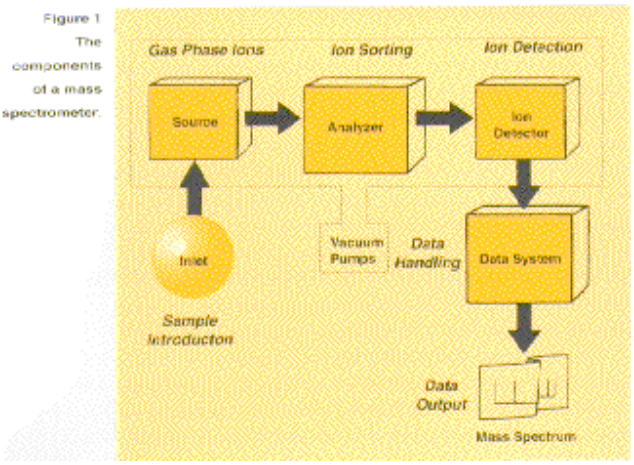

Figure 1.

rapidly evaporate until, through a succession of mechanisms that are not yet understood, only ions of the various analyte-compounds remain; these ions are then conducted into the mass analyzers vacuum chamber. As in FAB, electrospray ions are formed either by addition of one or more protons, alkali ions, or other cation in the case of positive ions and by abstraction of one or more protons or other anion in the case of negative ions. In conventional EI, 70 eV electrons are collided in vacuum with gas molecules of the analyte compounds to produce positively charged radical ions in accordance with the following reaction:

$$M + e^-(70eV) \Rightarrow M^{+\bullet} + e^-(\sim 50eV) + e^-(\text{thermal}).$$

The Mass Spectrum

Mass spectrometers measure the relative masses and relative abundances of (1) intact molecular ions and (2) charged fragments of molecular ions that decomposed after acquiring too much internal energy during either ionization or some other energetic event purposely caused to occur in order to induce dissociation. A mass spectrum is a record of relative ion abundance versus m/z. An example is shown in Fig. 2. The mass of an ion is measured in atomic mass units where

$$1\text{atomic mass unit}(u) = \frac{\text{atomic mass of } {}^{12}\text{C}}{12} = 1.66054 \times 10^{-27}\text{kg}.$$

The mass spectrum in Fig. 2 is in the form of a bar graph. To generate such a spectrum, the signal generated by the mass spectrometer's detector is converted from an analog to a digital form, and then a computer is used to determine the centroid of the digitized signal in m/z units and the height (or the area) of the signal relative to the most intense peak in the spectrum.

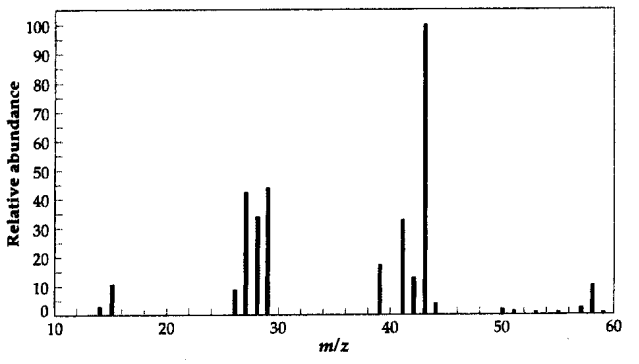

Figure 2

Qualitative Analysis

Since a mass spectrum exhibits signals produced by intact molecular ions and charged fragments, interpreting it requires criteria for recognizing a molecular ions signal and rules for deducing molecular structure from the fragment ions signals. Such criteria and rules are specific to molecular structure, ionization process, and (if required to produce fragments) decomposition-inducing process. Readers interested in learning the fundamentals of mass spectral interpretation are referred to F.W. McLafferty's classical text *Interpretation of Mass Spectra,* 4th Edition, University Science Books, Mill Valley, California, 1993.

Analytical situations that require qualitative mass spectrometric analysis fall generally into three distinct categories: (1) discovering and identifying compounds whose functions and structures are completely

unknown, (2) discovering and identifying compounds whose functions or effects are known but whose structures are completely unknown, (3) monitoring and quantitating compounds whose functions or effects and whose structures are partially or completely known. In the first two cases, mass spectrometry is used in conjunction with other forms of structural analysis, e.g. nuclear magnetic resonance spectrometry and spectrophotometry, to elucidate structure and with biophysical and biochemical tools to elucidate function or effect. In the third case, mass spectrometry is used in conjunction with extraction, filtering, and chromatographic techniques to assay for presence and quantity. Frequently, analyses are performed on compounds that are initially in one of the first two categories so that they can eventually be shifted into the last category.

Fig. 3 shows a positive FAB mass spectrum of a natural product that was isolated from a lipid extract of a blue/green algae *(Lyngbya schizothrix)*. This is an example of a first category compound. Until quite recently, this structure was completely unknown; its function is still under investigation. The structure is drawn in the upper right corner of the Fig. 3 was determined by applying fragmentation rules to the analysis of FAB mass spectra like that reproduced in the figure and by analyzing nuclear magnetic resonance spectra. The compound, which has since been named yanucamide A after the fact that it was discovered in algae found on the Fijian island Yanuca, is a member of a family of cyclic peptides that exhibit both anticarcinogenic and antifungal activities. Hence, this family of natural products is potentially a valuable source not only of pharmaceuticals but also of biotic fungicides for agricultural.

In order to produce a mass spectrum like that shown in Fig. 3, it is first necessary to obtain a pure sample of the compound. This can be accomplished, as it was in the particular case of yanucamide A, by carrying out a series of extractions and chromatographic separations, collecting fractions from the last stage of separation, and performing mass spectrometric analysis on each individual fraction. This tedious procedure usually can not be avoided in first category analyses because of the completely unknown character of the analytes. By contrast, second and third category analyses frequently permit the final stage of chromatographic separation to terminate directly in the ion source of a suitable mass spectrometer. When this is so, the resulting assays generally benefit from gains in speed, sensitivity, qualitative specificity, and quantitative accuracy.

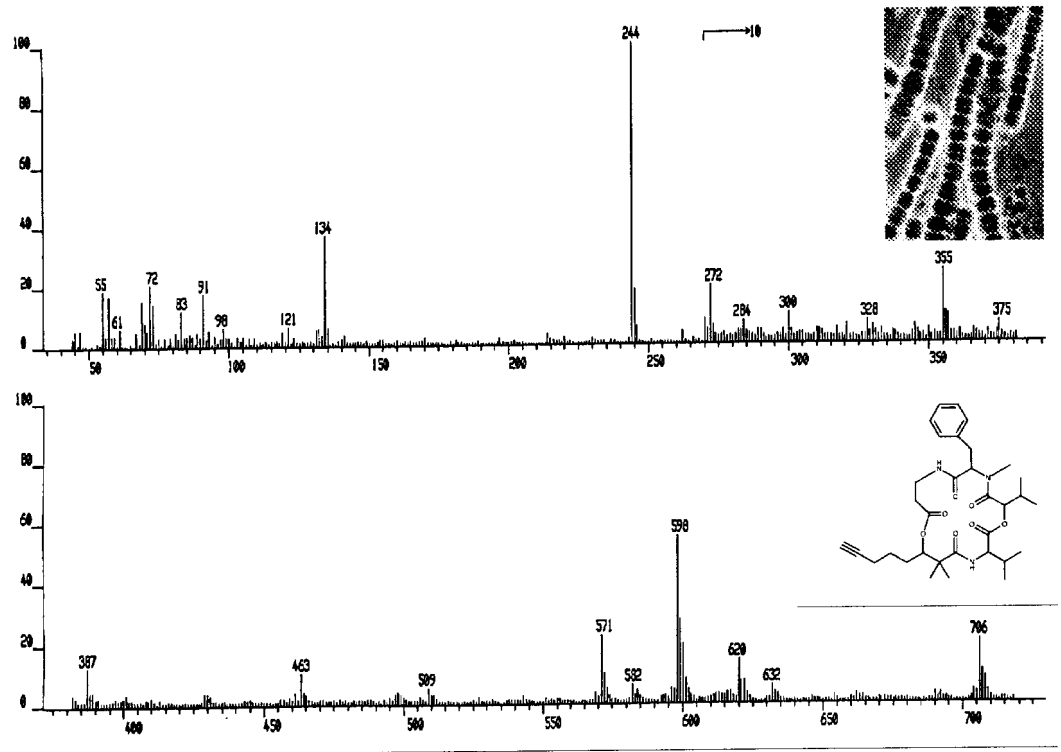

Figure 3

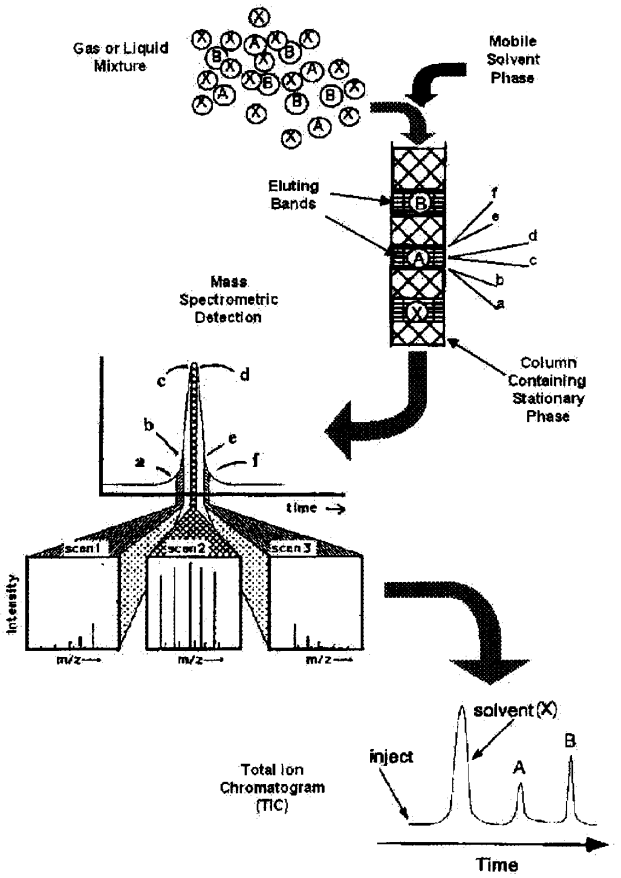

Figure 4

Fig. 4 illustrates the main steps in chromatographic mass spectrometry. A mixture of compounds is injected onto the front end of the column, and simultaneously, the mass spectrometer is triggered to begin repeatedly scanning over some preset range of m/z. The compounds are forced by the continuous introduction of a solvent to move through a column containing a stationary material that imparts a property-specific mobility to each compound, i.e. each compound moves through the column at a speed uniquely determined by how it interacts with the column's stationary phase. Different stationary phase materials select for different molecular properties. Thus, the compounds in the original mixture spread out in bands along the column as they pass through it and arrive at the end of the column at different times. Molecules in the leading edge of a band that is beginning to elute off the end of the column will be almost immediately ionized and recorded in a mass spectrum during the scan cycle that coincides with their arrival in the mass spectrometers ion source. Following a brief dead time during which the mass spectrometer resets itself to the beginning of its scanning range, a mass spectrum of the next front of molecules emerging from the column will be recorded. This periodic sampling of the eluting band continues until it has completely passed of the column, and it begins again immediately when the next band of compounds arrives at the end of the column. Typically, some 20 to 40 mass spectra are recorded in the time it takes a band of molecules to elute of the column. As the data acquisition computer stores each mass spectrum, it plots a number proportional to the total number of ions detected during the scan versus a time characteristic of the scan (e.g. the time it started). This real time plot, which is called a total ion chromatogram (TIC), is a digitized chromatogram of the mixture loaded onto the column.

The recent discovery of paclitaxel as a natural product of fungi found growing in the stems of a cultivar (cultivated variety) of a hazelnut tree commonly known as Gasaway *(Corylus avellana)* [Hoffman et al., *Spectroscopy* **13**, 22-32 (1998)] provides an excellent example of how chromatographic mass spectrometry can be a powerful aid to both second and third category analyses in agricultural research. Hazelnut production in the United States is worth more than $40,000,000 annually. This valuable agricultural enterprise is presently endangered because the trees that produce the best tasting nuts are highly susceptible to Eastern Filbert blight, a highly destructive disease caused by the fungal pathogen *Anisogramma anomala*. Gasaway, which produces a poor tasting nut of no commercial value, is resistant to Eastern Filbert blight.

The study in question began as a search for a statistical correlation between the susceptibility of different clones of hazelnut trees to Eastern Filbert blight and the fragment patterns in the electrospray ionization (ESI) mass spectra of methanol extracts from the stems of these trees. This strategy of exploiting a known biological function, which in this example is the molecular factor that endows the Gasaway cultivar with its resistance to Eastern Filbert blight, to isolate a structurally unknown compound is widely used in second category analyses. In the course of the correlation experiments, a large library of mass spectra of known compounds was searched for any ESI fragmentation patterns resembling those produced from the hazelnut tree- samples. These searches, which are standard practice in third category analyses, soon revealed a correspondence between the mass spectrum of a compound that repeatedly appeared in the methanol extracts of Gasaway and that of paclitaxel, which is an anticancer agent better known by its brand name Taxol (^{TM}Bristol-Myers). Fig. 5 shows a positive ESI mass spectrum of Taxol (M_r = 853 u). In order to create the fragment ion signals seen in the spectrum, the Taxol ions were forced to undergo gas phase collisions with nitrogen molecules at the orifice leading from the ionization chamber to the

mass analyzer. Taxol's structure is drawn in the upper right- hand corner of the spectrum, and dissociation pathways responsible for some of the major fragment ion signals are indicated in on this drawing.

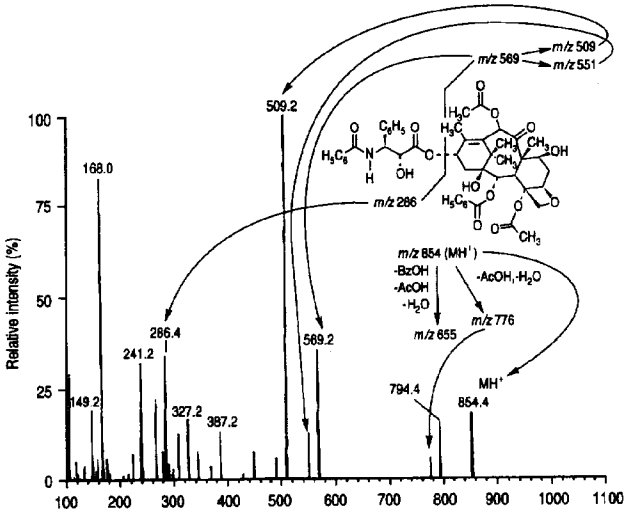

Figure 5 [Hoffman et al., Spectroscopy 13, 22-32 (1998)]

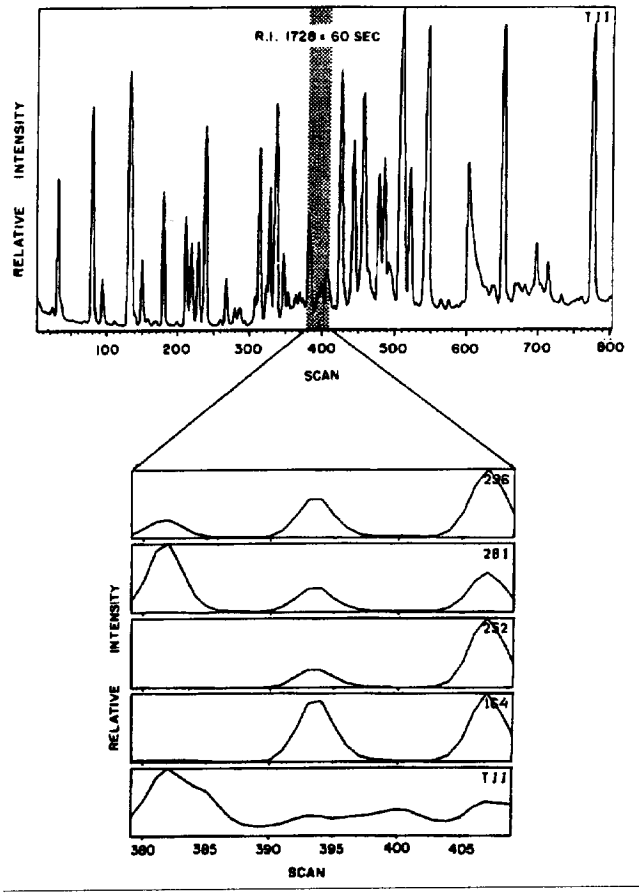

Figure 6 [J.T. Watson, *Introduction to Mass Spectrometry*, 3rd Edition, Lipincott-Raven, New York, 1997]

The uppermost panel in Fig. 6 shows a TIC of a methanol extract of Gasaway stems that putatively contains Taxol and some of Taxols precursors and metabolites. The mass spectra responsible for the chromatographic peak labeled Taxol, closely resemble the spectrum in Fig. 5; however, a strictly visual comparison could be misleading in this instance because the mass spectra responsible for several of the other chromatographic peaks in Fig. 6 also closely resemble the spectrum in Fig. 5. In cases such as this where several closely related structures are involved, the signal to background for a particular compound in the TIC chromatogram can be substantially increased by taking advantage of the underlying mass spectral data to construct a set of unique chromatograms. In a TIC, every mass spectrum recorded during the chroma- tographic separation of a mixture is plotted as a pair of ion-count/scan-time coordinates. It is also possible, however, to construct a mass chromatogram by plotting only the ion-count/scan- time coordinates of mass spectra exhibiting a signal that appears at one particular m/z and exceeds some minimum intensity-threshold are plotted. Any number of mass chromatograms can be constructed from the mass spectra recorded during chromatographic separation of a mixture.

The bottom four panels in Fig. 6 are mass chromatograms for protonated Taxol (m/z 854) and three fragment ions (m/z 569, 509, and 286; see Fig. 5). Although each of the compounds in the mixture produce mass spectra with one, two, or even three of these signals (as would be expected if they are indeed precursors and metabolites of Taxol), only the compound responsible for the chromatographic peak labeled Taxol produces mass spectra that exhibit all four ion signals. Thus because of the high degree of specificity inherent in chromatographic/mass spectrometric data, a second category search for the identity of a compound known to cause a devastating disease in a valuable agricultural resource resulted in the chance discovery of a new biological source of a well characterized compound known to have powerful anticarcinogenic properties. This finding is significant because it suggests that Taxol might be available from a much wider number of plant species than previously thought.

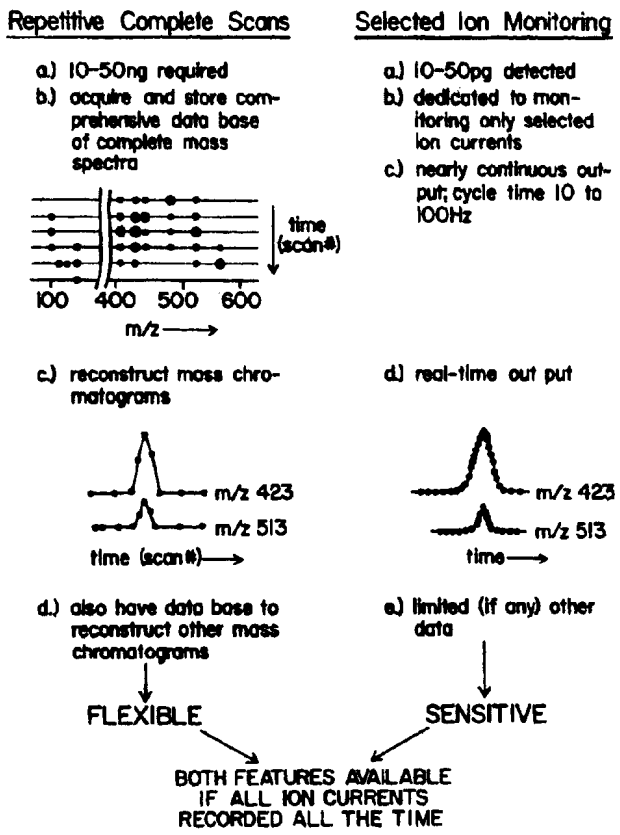

Figure 7 [J.T. Watson, *Introduction to Mass Spectrometry*, 3rd Edition, Lipincott-Raven, New York, 1997]

The identification of Taxol in Gasaway stems is an example of a third category analysis. When only a limited amount of sample is available for an assay of this type, the sensitivity of the mass spectrometric detection can be increased by two or even three orders of magnitude by programming the mass spectrometer to detect ion signals at only a small number of m/z values (e.g., 1 to 5). The fraction of time an instrument devotes to detecting a particular signal is referred to as the duty factor. When a mass spectrometer is operated in a selected ion monitoring mode, the duty factors for the chosen ion signals typically range from 20 to 100% whereas, in a scanning mode, they are typically in the range of 0.1 to 1%. The special case in which an ion signal at only one m/z is monitored (100% duty factor) is called single ion monitoring. Outwardly, selected ion monitoring chromatograms appear the same as mass chromatograms reconstructed from scanned data like those shown in Fig. 6 for Taxol, but in fact, the former provide far superior sensitivity while the latter provide far greater analytical flexibility. The similarities and differences between the two recording modes are summarized in Fig. 7.

When a mass spectrometer is operated in a selected ion monitoring mode and at a high resolving power, exceptional compound specificity and sensitivity are achieved. This is power fully exemplified by the high resolution, electron impact mass spectrum of the steroid 4- androstene-3,17-dione reproduced in Fig. 8 (note that the mass range covered in the mass spectrum is only 0.4 u). The spectrum was recorded at a resolving power of 10,000. The mass accuracy at this resolution is at least $\pm$ 0.001 u. At this accuracy, the measured m/z of 286.193 unambiguously corresponds to the mass of the C_{12}-isotopomer of a compound with the chemical formula $C_{19}H_{26}O_2$, which is the chemical formula for androstenedione. Hence, the structural specificity of the analysis is exceedingly high. The signal to background ratio for the molecular ion of the steroid is effectively zero because the high resolving power of the mass spectrometer completely prevents interference from the ion signals produced by any of the other compounds in the sample. Therefore, high sensitivity (30 pg/μL) is achieved by monitoring only the steriod's signal (the selected ion) through a 0.03 u window centered at m/z 286.193. Steroid analyses such as this find numerous applications in research on the production of agricultural animals.

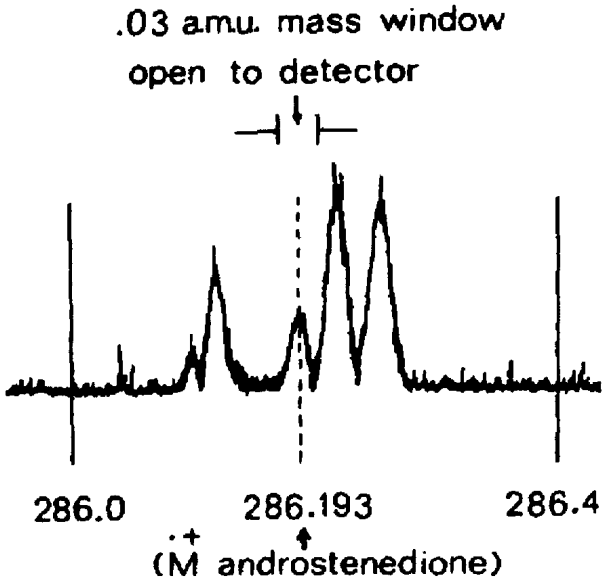

Figure 8 [D.S. Millington in *Applications of Mass Spectrometry to Trace Analysis*, S. Facchetti, ed., Elsevier, Amsterdam, pp. 189-202, 1981]

Quantitative Analysis

A quantitative analysis comprises several steps: (1) obtaining a representative bulk sample from the object of analysis, (2) extracting from the bulk sample a homogeneous laboratory sample, (3) converting the laboratory sample into a form suitable for analysis, which

usually involves dissolving the sample and, if the concentration of analyte is low, concentrating the analyte, (4) removing or masking species that interfere with the chemical analysis, and (5) measuring the concentration of analyte in several aliquots so that the variability (uncertainty) in the analysis can be assessed. Fig. 9 shows the analytical scheme used to determine the concentrations of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) in freshwater fish found in the Willamette River in Oregon, USA [Curtis et al., *Environmental Science & Technology* **27**, 2149-2157 (1993)]. Accumulation of TCDD by fish is a public and regulatory concern because the fish themselves are highly sensitive to the toxicity of TCDD and there is a potential for trophic transfer from contaminated fish to humans and animals that consume them. It is obvious from the figure that that the sampling and sample cleanup steps (i.e., Steps 1-4) are very laborious. In order to assure accuracy in the analysis step (i.e., Step 5), painstaking care must be taken to avoid contamination and losses during each transfer and dilution step. While this sounds easy enough in principle, it can be very difficult to achieve in practice. Patience and painstaking attention to detail are the hallmarks of quantitative analysis. Quantitative mass spectrometric analyses require that signal response be calibrated to some measure of sample amount or concentration. This usually accomplished by introducing varying amounts of a suitably prepared standard of the analyte or analytes into the mass spectrometer (either directly or through a chromatography system) and then plotting the corresponding signal outputs (either as peak height or area) versus sample amount to create a calibration curve. In the study under discussion, this was accomplished by setting the mass spectrometer to transmit only the two most abundant masses of TCDD (319.8965 and 321. 8936) and the most abundant mass of the internal standard, [^{13}C]TCDD (333.9338) at a resolving power of 8,000. The calibration curve was generated from multiple analyses of prepared standards of of natural isotopic abundance TCDD plus 200 pg/μL of [^{13}C]TCDD. Examples of the high resolution selected ion chromatograms of 0.1 pg and 1.0 pg standards are shown in Fig. 10 along with the procedure for calculating the concentration of TCDD from the measured areas of the chromatographic peaks. Note that the two chromatograms have different time scales. The absolute detection sensitivity for TCDD in this study was 10 fg and the detection sensitivity for TCDD residues in fish samples was 0.1-0.2 pg/g. Table 1 lists some of the results obtained for cutthroat trout; analyses of other types of fish and of various river sediments were also made.

The examples presented in this overview are but a small sampling of the ways mass spectrometry is employed in agriculture and natural product research. Nevertheless, they are representative of the most common modes of analysis performed and, hopefully, they will stimulate interested readers to seek mass spectrometric solutions to their analytical problems.

Table 1 - Concentrations (pg/g)[a] of TCDD in Cutthroat Trout taken from Sites on the Willamette River in Oregon, USA.

Location[b]	TCDD	Avg
Mid. Fork/J	0.16, 0.09	0.13
Mid. Fork/O	0.34, 0.25, 0.23	0.27
Harrisburg/J	0.39, 0.35, 0.37	0.37
Harrisburg/O	0.30, 0.57, 0.19	0.35
Halsey/J	0.77, 0.73, 0.92	0.81
Halsey/O	1.82, 1.87, 1.39	1.69

[a] Tissue concentrations of whole fish; [b] J = July, O = October

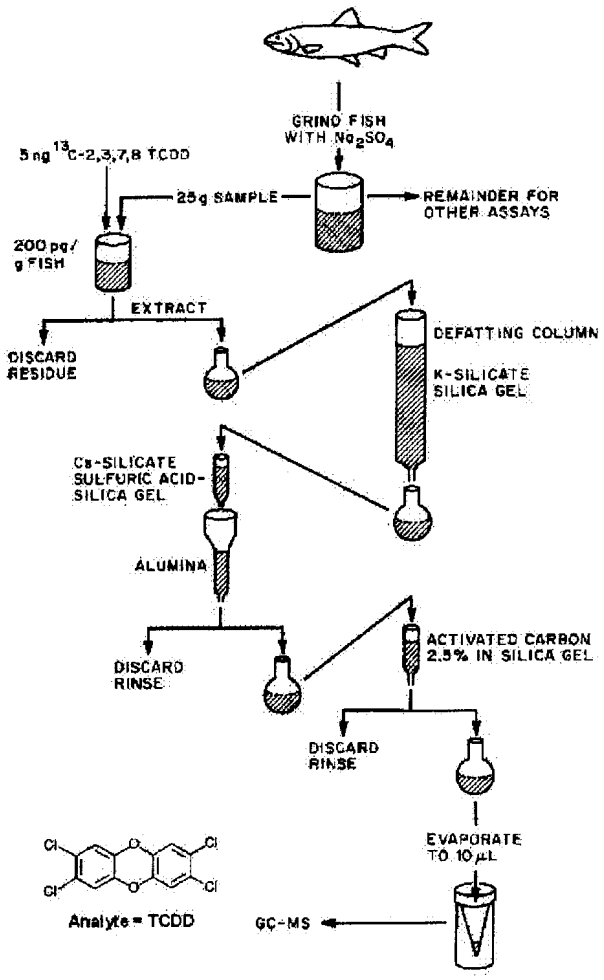

Figure 9

200 pg/g ^{13}C-[TCDD] added to Sample

Assume ^{12}C-[TCDD]/^{13}C-[TCDD] remains constant through workup

Compare GC-MS sample peak area with standard
(1 pg ^{12}C-[TCDD] + 200 pg ^{13}C-[TCDD]

$$[TCDD]_{sample} = \frac{(A_{320} + A_{322}/A_{334})_{sample}}{(A_{320} + A_{322}/A_{334})_{standard}} \text{ (pg/g)}$$

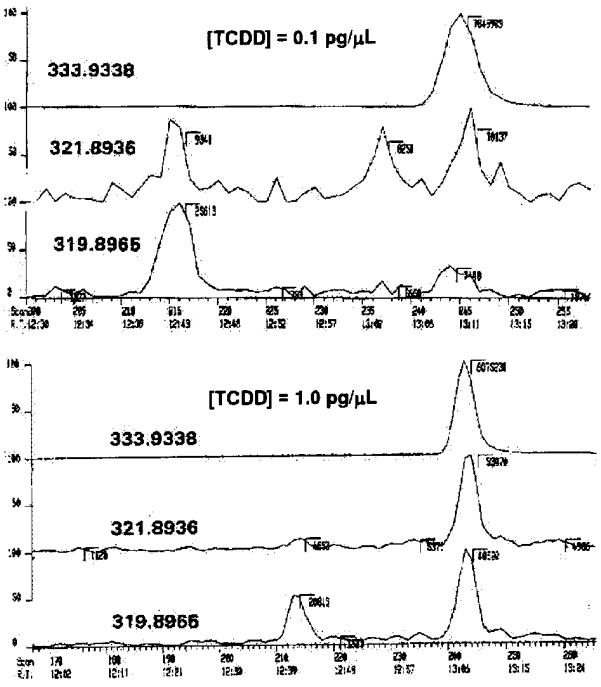

Figure 10