The Application of High Resolution Mass Spectroscopy to Organic Chemistry

By G. W. A. Milne

LABORATORY OF METABOLISM, NATIONAL HEART INSTITUTE, NATIONAL INSTITUTES OF HEALTH, BETHESDA, MARYLAND 20014, U.S.A.

1 Introduction

The fundamental principles of mass spectrometry were established at the turn of the century in a series of classical experiments by, *inter alios*, Wien,[1] and Thomson.[2] These workers demonstrated that a beam of positive ions could be deflected by electrical or magnetic fields and that the angle of deflection of any given ion varied with a number of parameters, one of which is the mass of the ion. Thomson[2,3] designed a mass spectrograph in which he demonstrated that neon gave not one but two lines, since it consists of two stable isotopes, differing in mass. An improved mass spectrograph was subsequently developed by Aston[4] and was used to carry out similar isotope analyses on a variety of elements.[5] At about this time, Dempster[6] published details of his mass spectrometer in which the analysed ions were focused at a point rather than in a plane as in mass spectrographs. In such a system, detection of the ions may be accomplished electrically rather than photographically and this variation, with its inherent greater sensitivity, has come to assume considerable significance in high resolution mass spectrometry. Thus mass spectrometry was effectively dichotomous at birth, and in terms of instrumentation the subsequent development of deflection spectrometers has been in two similar, but rarely intersecting directions. The most sophisticated instruments now available are direct descendants of Aston's mass spectrograph or of Dempster's mass spectrometer, and in the area of high resolution mass spectrometry, to which this Review will be restricted, no fundamental change in either of these designs has enjoyed great success.

The first area of organic chemistry in which mass spectrometry was employed was hydrocarbon analysis,[7] a field in which sample handling problems are minimal, the compounds generally being relatively stable and volatile. From this point, the development of the role of mass spectrometry in organic chemistry was largely arrested, pending the handling of two distinct problems: (i) that of efficiently vaporising and ionising non-volatile or unstable compounds; (ii) that

[1] W. Wien, *Ann. Physik*, 1898, **65**, 440.
[2] J. J. Thomson, *Phil. Mag.*, 1911, **21**, 225.
[3] J. J. Thomson, 'Rays of Positive Electricity', Longmans, Green and Co., London, 1913.
[4] F. W. Aston, *Phil. Mag.*, 1919, **38**, 707.
[5] F. W. Aston, 'Isotopes', Edward Arnold, London, 1922, 1924.
[6] A. J. Dempster, *Phys. Rev.*, 1918, **11**, 316.
[7] H. W. Washburn, H. F. Wiley, and S. M. Rock, *Ind. and Eng. Chem.* (*Anal.*), 1943, **15**, 541.

of recognising and extracting the wealth of information in mass spectra.

Most sample handling systems that are useful with organic compounds have evolved from the premise that the sample should be admitted to the ionising region in the vapour phase with a vapour pressure no greater than about 10^{-5} mm. of mercury. At such pressures the mean free path of a particle is about 5 metres and ion-molecule collisions may be ignored. Early mass spectrometers were designed with gas inlet systems[8] which during the 1940's were modified to accommodate volatile organic liquids.[9] In 1955 the technique of direct introduction of the sample into the ionising region was applied.[10,11] Direct introduction of the sample *via* a vacuum lock[12] is now routine and most commercially available machines have such an inlet system which is probably the most generally useful system for the organic chemist. Using one or another of these various sample handling systems, the organic chemist will be able to obtain mass spectra on the majority of organic compounds. Those which fail to vaporise unchanged however cannot be handled in this way and create a special problem which is discussed later (*v. infra*).

Techniques by which information can be extracted from mass spectra are in their infancy but it is already abundantly clear, thanks to the pioneering efforts of Beynon,[13,14] Biemann,[15] Djerassi,[16] McLafferty[17] and others, that mass spectrometry has enormous potential as an analytical tool in organic chemistry. There is already available a considerable amount of empirical knowledge of fragmentation mechanisms, intelligent application of which has been shown in many instances[18] to be of great assistance in problems of structure determination.

Since the demonstration by Beynon[14] that with instruments of higher resolving power, the mass of an ion may be measured sufficiently accurately to permit the deduction of its atomic composition, a great deal of activity and expenditure has been applied to the field of high resolution mass spectrometry. Mass spectrometers have been improved to the point where such accuracy can be obtained routinely. Computer acquisition and processing of the large quantity of data in any one high resolution spectrum has become fairly common and an exciting

[8] J. H. Beynon, 'Mass Spectrometry and its Applications to Organic Chemistry', Elsevier Publishing Co., London, 1960, p. 147.
[9] Ref. 8, p. 161.
[10] P. Bradt and F. L. Mohler, *Analyt. Chem.*, 1955, **27**, 875.
[11] P. de Mayo and R. I. Reed, *Chem. and Ind.*, 1956, 1481.
[12] J. L. Courtney and J. S. Shannon, *Tetrahedron Letters*, 1963, 13.
[13] J. H. Beynon, 'Advances in Mass Spectrometry', ed. J. Waldron, Pergamon Press, London, 1959, p. 328.
[14] J. H. Beynon, 'Advances in Mass Spectrometry', ed. M. Elliott, Pergamon Press, London, 1961, p. 216. See also ref. 8.
[15] K. Biemann, 'Mass Spectrometry. Organic Chemical Applications', McGraw-Hill, New York, 1962.
[16] H. Budzikiewicz, C. Djerassi, and D. H. Williams, 'Interpretation of Mass Spectra of Organic Compounds', Holden-Day, San Francisco, 1964.
[17] 'Mass Spectrometry of Organic Ions', ed. F. W. McLafferty, Academic Press, New York, 1963.
[18] See, for example, K. Biemann in 'Progress in the Chemistry of Organic Natural Products', ed. L. Zechmeister, XXIV, 1966.

new field, computer-assisted interpretation of mass spectra is now being actively explored.

This Review, which is not intended to be comprehensive, will be concerned with the more important results of these endeavours. A number of important areas, such as the various methods of ionisation, will not be considered in depth and for details of this sort, the reader is referred to the more exhaustive survey by McLafferty and Pinzelik.[19] The purpose of this review is to consider the present position of high resolution mass spectrometry *vis-à-vis* organic chemistry in the light of past results and future prospects.

2 Sample Handling Systems

A large number of gas inlet systems have been designed,[8] most of which use a molecular leak in a pressure reducing system and many of which employ mercury as a cut-off phase. These are therefore of somewhat limited value for organic chemists whose samples are often less volatile than mercury. Replacement[20] of mercury by gallium extends the temperature range of many of these systems to beyond 300° and allows the design of an inlet system that may be used for a large number of non-polar organic molecules.[21] For general use with such compounds, the glass–metal system first described by Beynon[21] is extremely useful. A simpler version of this system has subsequently been designed[22] but is not, however, commercially available as is its forerunner. An enamelled all-metal system has also been developed.[23]

Temperature controlled direct introduction of the sample into the ion source was first reported some twelve years ago[10,11,24] but all these devices were inconvenient because the mass spectrometer could not be kept under vacuum whilst a sample was admitted. This difficulty was overcome by the incorporation of a vacuum lock[12,25,26] and the resulting direct insertion probes, which are now commercially available, can be used with all thermally stable compounds which do not represent extremes in volatility. It has been pointed out moreover[26] that for compounds that are moderately thermally unstable, the direct probe is superior to the older inlet systems in that the sample is, in effect, flash-evaporated and is not heated protractedly. For very volatile samples, a cryogenically cooled probe has been designed.[27] A similar, simpler probe has been developed and used at $-100°$[28] but no commercial models of such systems are available.

[19] F. W. McLafferty and J. Pinzelik, *Analyt. Chem.*, 1966, **38**, 350R. See also ref. 39.
[20] M. J. O'Neil and T. P. Wier, *Analyt. Chem.*, 1951, **23**, 830; M. J. O'Neal, 'Applied Mass Spectrometry', The Institute of Petroleum, London, 1954, p. 27.
[21] Ref. 8, p. 184.
[22] V. J. Caldecourt, *Analyt. Chem.*, 1955, **27**, 1670.
[23] C. Brunee, *Z. Instrumentenk.*, 1965, **73**, 16.
[24] R. I. Reed, *J. Chem. Soc.*, 1958, 3432.
[25] H. C. Hill and R. I. Reed, *J. Sci. Instr.*, 1963, **40**, 259; G. A. Junk and H. J. Svec, *Analyt. Chem.*, 1965, **37**, 1629.
[26] G. L. Kearns, *Analyt. Chem.*, 1964, **36**, 1402; M. Barber, T. O. Merren, and W. Kelly, *Tetrahedron Letters*, 1964, 1063.
[27] H. A. McGee, Jun., NASA, Doc. N6323401, 1963.
[28] W. F. Haddon, E. M. Chait, and F. W. McLafferty, *Analyt. Chem.*, 1966, **38**, 1968.

A simple labelling technique evolves from the fact that so-called 'active hydrogens' may be replaced by deuterium if the sample is admitted into the ion source with D_2O either *via* a probe[29] or an indirect inlet system.[30]

Since Gohlke's first successful attempt[31] to pass the effluent from a gas chromatogram into the source of a mass spectrometer, much work has been done on the dual problems of (i) reducing the recording time of the mass spectrometer to the 1–10 seconds available during the emergence of a gas chromatography peak and (ii) removing the carrier gas used in the chromatography. Electronic improvements, such that magnet scanning at high resolution can be done in under 10 seconds per decade, have been made, and instruments with this capability are available commercially. The efficient removal of carrier gas without loss of solute has proved to be considerably more difficult however and none of the various separators that have been designed is very efficient, although the enormous sensitivity of the mass spectrometer ameliorates this problem to some extent. Three commercially available separators are the Biemann–Watson separator,[32] which is in effect a molecular sieve, the Ryhage molecular jet separator,[33] which removes the molecules of carrier gas as they emerge from a jet with much lower momentum than solute molecules and the Llewellyn semi-permeable membrane separator[34] which permits the passage of organic molecules but not of carrier gas. The use of porous Teflon as a separator has been described,[35] and such a device has been critically compared[36] to the Biemann–Watson separator. Capillary columns may be used without a separator[37] and a general discussion of some of the problems encountered in coupled gas chromatography-mass spectrometry has appeared.[38]

The majority of organic compounds can be handled in one or another of the inlet systems described above, but some highly polar compounds cannot be vaporised without decomposition. The most promising approach to such compounds is the formation of volatile derivatives such as trimethylsilyl ethers and esters.

Compounds which vaporise satisfactorily but fail to give a molecular ion also cause difficulties. The use of lower electron voltages sometimes helps in such cases and the use of field emission rather than electron bombardment holds great promise in spite of its lower sensitivity. But at present, the common approach to this problem is to tolerate it and extract the remaining available information.

[29] See ref. 15, p. 231.
[30] J. S. Shannon, *Austral. J. Chem.*, 1962, **15,** 265.
[31] R. S. Gohlke, *Analyt. Chem.*, 1959, **31,** 535.
[32] J. T. Watson and K. Biemann, *Analyt. Chem.*, 1964, **36,** 1135; 1965, **37,** 844.
[33] R. Ryhage, *Analyt. Chem.*, 1964, **36,** 759; *Arkiv Kemi*, 1966, **26,** 305.
[34] P. M. Llewellyn and D. P. Littlejohn, Pittsburgh Conf. Anal. Chem. App. Spectroscopy, Feb. 1966.
[35] S. R. Lipsky, C. G. Horvath, and W. J. McMurray, *Analyt. Chem.*, 1966, **38,** 1585.
[36] M. A. Grayson and C. J. Wolf, *Analyt. Chem.*, 1967, **39,** 1438.
[37] D. Henneberg, *Analyt. Chem.*, 1966, **38,** 495. See also ref. 47.
[38] W. H. McFaddon and E. A. Day, *Analyt. Chem.*, 1964, **36,** 2362.

3 Instrumentation

Two commercial high resolution mass spectrometers suitable for use with organic compounds are the CEC 21-110B (Consolidated Electrodynamics Corporation, Pasadena, California) and the AEI MS 902 (Associated Electrical Industries, Ltd., Manchester, England). Of the other commercially available high resolution spectrometers, none will be considered here because they are all recent additions to the field and are geometrically similar to either the 21-110B or the MS 902. In addition, relatively little information is available upon the performance of these other machines when coupled with automatic data-handling systems.

Ions are produced in both the 21-110B and the MS 902 by electron bombardment in sources which are basically similar. The ions are then extracted out of the source by an accelerating potential, commonly 8kV, and focused, first in an electrostatic sector and then in a magnetic sector. The resolved ion beams, each one homogeneous with respect to mass to charge ratio,* then enter a detector system.

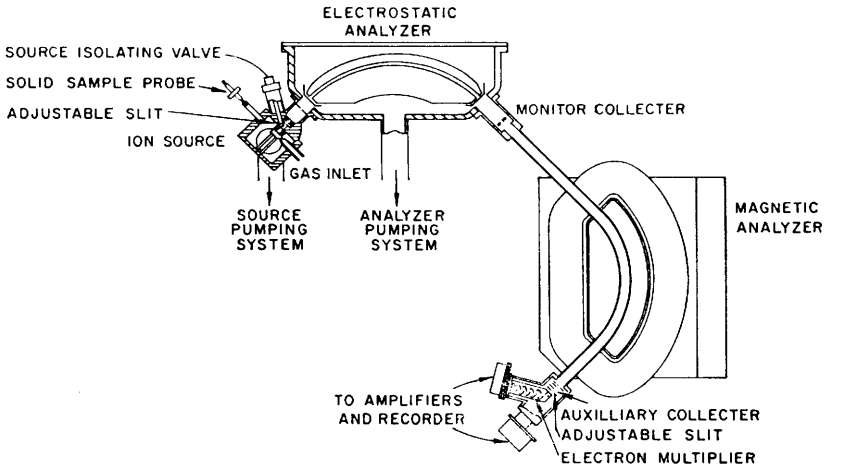


Fig. 1 *Schematic diagram of the* MS 902. *Nier–Johnson geometry.*

The MS 902 is related to Dempster's original spectrometer[6] in that it employs the Nier–Johnson geometry shown in Figure 1. In this system the analysed ion beams are focused in a complex conic surface. At the focal point of this system is placed an exit slit behind which is the first dynode of an electron multiplier. The magnetic field is then scanned, usually from high field (high mass) downwards, and as one mass after another enters the multiplier, the output signal is amplified and recorded in an oscillographic recorder. The mass of a particular ion is related therefore to the time elapsed from the start of a scan when it appears and if the times for standard ions are known accurately, the precise mass of the unknown ion may be calculated.

* The only ions considered in this Review are singly charged. Reference will be made therefore to their 'mass' rather than to their 'mass to charge ratio'.

In contrast to this, the Mattauch–Herzog geometry of the 21-110B, shown in Figure 2, ensures that the analysed ions are focused in a plane, in which, conventionally, a photoplate is placed. This 'spectrometer' is therefore really a 'spectrograph' as was Aston's original design.[4] An image is registered on the plate by each resolved ion in a position which depends upon the square root of the mass. The final density of the line is proportional to the abundance of the ion. No scanning of the magnetic field is required; the precise mass of an ion can be obtained by accurately measuring its position on the photoplate with respect to standard ions. The 21-110B is also equipped with an electrical detector system, based upon an electron multiplier which is placed in the focal plane at the far end of the photoplate (see Figure 2). This represents an alternative to the photoplate and permits the recording of a mass spectrum by scanning the magnetic field, much as with the MS 902.

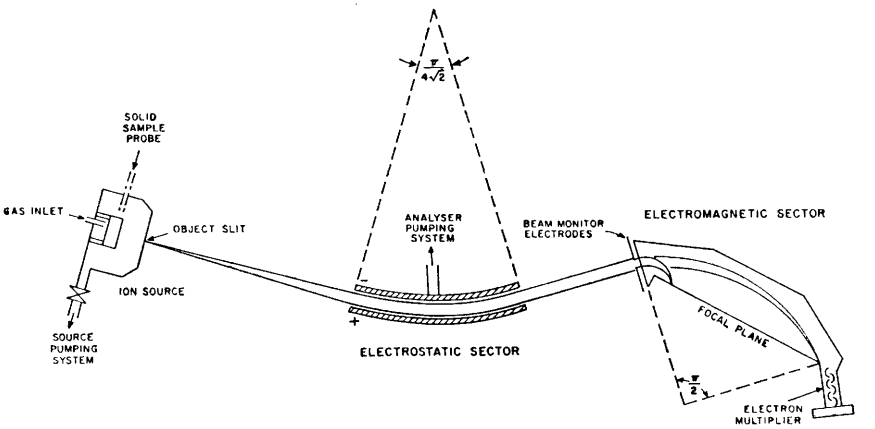


Fig. 2 *Schematic diagram of the* 21-110B. *Mattauch–Herzog geometry.*

A comparison of these two spectrometers is less fruitful than a comparison of the two methods of recording high resolution mass spectra, *i.e.*, the photoplate method and the scanning procedure. The static nature of photoplate recording is very convenient because automatic data acquisition from a photoplate is relatively easy, requiring only a series of accurate linear measurements and photodensitometer readings. Scanning a magnetic field rapidly and reproducibly is difficult and the circuitry required is complex. A further serious technical problem is created by the requirement that the time centroid of each peak be measured with an accuracy of 1 part in 10^5.

The 21-110B is harder to focus[39] and since it cannot avail itself of an amplification system for photoplate recording, it is inherently less sensitive in this mode towards a given number of ions than is a scanning spectrometer, which

[39] R. I. Reed, *Quart. Rev.*, 1966, **20**, 527.

can easily detect a single ion. The photoplate method of recording moreover, does not lend itself to abundance measurements whilst the output from an electrical detection system is effectively a plot of mass *versus* abundance.

A complete spectrum at a resolving power of 1 part in 10^4 can be recorded by either method in about 10 seconds thus permitting the analysis of effluent peaks from a gas chromatogram. Peak matching (*v. infra*) is possible with both machines, but in this respect, the MS 902 has a superior performance. Finally, the MS 902 shows 'metastable' ions considerably more readily than does the 21-110B although this is of no moment in high resolution work when such ions are rarely observed in either spectrometer.

In summary then, whilst it is pointless to argue that one spectrometer is superior to the other, it is fair to say that electrical detection, by virtue of its much higher sensitivity and accuracy, is quite superior to photoplate recording, whose ultimate and serious limitation is the size of the grain in the photoplate emulsion. Handling data from magnetic scans is difficult, but as will be seen, has been accomplished with both machines and will surely prove to be the future method of choice for the recording of high resolution mass spectra.

4 Resolving Power

Two peaks of equal height are said to be just resolved if the valley between them is 10% of the peak height and the resolving power of the system at this point is defined as the difference in the masses of the two ions, *e.g.*, 1 part in 5000 or 200 ppm. This '10% valley definition' of resolving power is not the only definition in use but enjoys wider currency than others such as '1% valley' and will be used here. Instruments with which a resolving power of 1 part in 5000 cannot easily be achieved are generally incapable of mass measurement to the accuracy required for formula assignment to ions from any but the simplest of organic compounds and will not be considered here.

A resolving power of 1 part in 10,000 is often assumed for the purposes of discussion to be a prerequisite for mass measurement of an accuracy sufficient to permit formula assignment but this is a misleading generalisation in two senses. It is perhaps more useful to remember the rule derived from experience, that, depending upon the technique used, the accuracy of mass measurement may be expected to be between one and two orders of magnitude greater than the resolving power being used. Thus a resolving power of 1 part in 10,000 should allow mass measurement with an accuracy of perhaps 1 part in 200,000 or 5 ppm. The other aspect of the problem is that the accuracy required depends only on the problem in hand. If it is not known what elements are present in the sample then the highest feasible accuracy should be sought, but where this information is available, or, as will be discussed later (*v. infra*) can be obtained from the spectrum itself, it is pointless to use a resolving power that is vastly too high for the problem at hand.

As a trivial example, the only two possible ions of mass 140 (ignoring ^{13}C- and ^{2}H-containing ions) in the mass spectrum of a hydrocarbon are $C_{10}H_{20}^{+}$

(m/e 140·1565)* and $C_{11}H_8^+$ (m/e 140·0626). These will be completely separated with a resolving power of 671 ppm (1 part in 1492) and the mass of either could be easily measured accurately enough to permit a choice between the two at a resolving power of 1 part in 1000. In fact this particular problem is so easy that it may be done by manually measuring the position of the peak on the oscillograph output.[40] Hydrocarbons represent the simplest case, and as the number of elements involved increases, the accuracy of mass measurement required to permit an unequivocal assignment of formula increases.

The vast majority of organic compounds however contain no elements other than C, H, O and N. The ions representing simple combinations of these elements that will be closest together (again ignoring minor isotopes) will be separated by 0·0126 mass units which is the difference in mass between CH_2 and N. At mass 500 these should appear as separate peaks at a resolving power of 25 ppm (1 part in 40,000) and should be quite easily mass measured at a resolving power of 1 part in 10,000. Although they are somewhat less common, doublets that are more difficult to resolve do occur, *e.g.*, C_2H_2O and N_3 differ by only 0·00135 mass units. Introduction of sulphur, halogens or carbon-13 into the consideration creates a requirement for even greater accuracy and with currently available instruments, the spectroscopist soon finds himself between the Scylla of high resolution and the Charybdis of adequate sensitivity. In practice, the presence of sulphur, chlorine or bromine can be immediately discerned from their characteristic isotopic distributions whilst iodine is distinguished by its large negative mass defect. Separation of the ^{13}C-^{12}CH and $^1H_2$-^2H doublets is generally considered to be beyond the capabilities of present instruments requiring as it does, a resolving power of 1 part of 25,000 at mass 100. Currently, most instruments are used at a resolving power of about 1 part in 10,000. Although both the 21-110B and the MS 902 are capable of three times that, this figure represents a compromise between sensitivity and resolution.

It is here that the field of high resolution mass spectrometry begins to become distinct from that of low resolution mass spectrometry. Mass spectrometers of low resolving power are of great value in a variety of connections such as trace analysis, reaction kinetics, isotope labelling and structural studies but it is in this latter field that high resolution mass spectrometry has unparalleled potential. Assignment of an unequivocal formula to every ion type in the spectrum and subsequent reassembly of all the fragments to obtain a unique structure for the few micrograms of sample compound is perhaps a breathtaking feat but as will be seen, it has actually been done in a few cases.

5 Peak Matching

In principle, a mass spectrometer could be calibrated and all the variables controlled so that the accurate mass of an ion could be calculated immediately from its position on a photoplate or the time of its appearance during a magnetic

* Based upon the $^{12}C = 12{\cdot}000000$ scale adopted by I.U.P.A.C.

[40] B. H. Johnson and T. Aczel, *Analyt. Chem.*, 1967, **39**, 682.

scan. There are however too many variables to deal with and adequate control of some of these is very difficult. The usual technique therefore, is to calibrate continuously with standard ions and measure the mass of the unknown ion with reference to one or two standard ions. This requires that the sample and the reference compound be admitted simultaneously. Most laboratories use as a standard compound either perfluorokerosene (PFK) or heptacosafluorotributylamine. The latter is a homogeneous compound which is sufficiently volatile to be admitted *via* a gas inlet system whilst the former is normally supplied as a mixture of homologues which requires a heated inlet system. PFK has the great advantage, however, of giving a standard ion at least every 12 mass units whilst in the spectrum of heptacosafluorotributylamine there are some large gaps between abundant ions, and there are no ions above m/e 614.

The only simple method of mass measurement with an accuracy of 10 ppm is the peak matching technique developed by Quisenberry, Scolman and Nier.[41] With the accelerating voltage and the electrostatic analyser voltage kept constant the magnetic field is adjusted so as to bring into focus a standard ion, lower in mass but as close as possible to the unknown ion. A small auxiliary magnet is then swept repetitively over a very small mass range about this ion, *e.g.*, for the standard ion $C_2F_4^+$ (99·99362) from 99·996 to 99·990 and the output displayed on an oscilloscope. The accelerating and electrostatic analyser voltages are then switched back and forth on alternate sweeps to lower values which can be adjusted until the unknown peak is superimposed upon the standard peak. When this condition is fulfilled, the ratio of the higher accelerating potential to the lower accelerating potential can be read directly and is equal to the ratio of the unknown mass to the standard mass. Thus if the unknown ion is superimposed upon the $C_2F_4^+$ standard ion at a ratio of 1·020150, the unknown mass can be immediately calculated to be 102·0085, a number which in practice has six significant figures and is therefore accurate to 1 mmu below m/e 1000. Since the accelerating potential can only be decreased, the magnetic field must be adjusted to focus the lower mass ion and it is convenient to make this the standard ion. The maximum ratio that could be used with Nier's original system was 1·10, *i.e.*, the unknown mass had to be within 10% of the standard mass. A voltage trimming device has subsequently been developed[42] for the MS 902 and extends this range to 2·00 although for other reasons, the useful range is about 1·70. Under these circumstances, heptacosafluorotributylamine proves to be an adequate internal standard and is often used with the peak matching system of the MS 902. With the 21-110B however, the peak matching system has a maximum ratio of 1·10 and since it employs a meter as opposed to an oscilloscope it is quite difficult to use.

Whilst peak matching will probably be overwhelmed by automatic data processing techniques in the fight for survival, its present position is one of great importance. It still is the quickest and simplest way to measure accurately the

[41] K. S. Quisenberry, T. T. Scolman, and A. E. Nier, *Phys. Rev.*, 1956, **102**, 1071.
[42] H. M. Fales, R. Binks, M. Elliott, and R. Freeman, Proceedings of the ASTM, Committee E-14, Dallas, 1966.

mass of a limited number of ions in the spectrum, such as the molecular ion and major fragment ions. With the MS 902, a dexterous and resilient operator can peak match one ion per minute with this technique. This is perhaps 10–100 times slower than the current automatic data processing systems discussed in the following section but is inestimably cheaper. It may be noted parenthetically that Ryhage[43] has adapted the peak matching system to a single focusing mass spectrometer and reported surprisingly accurate mass measurements.

6 Automatic Data Processing

Data processing techniques are inescapable in high resolution mass spectrometry for two reasons. Firstly, the sheer amount of data in just one high resolution mass spectrum is too much to collect manually, by peak matching and secondly, the assignment of formulae to ions of known mass, whilst it can be done in individual cases with the help of tables[44] is, for a large number of ions, a problem tailor-made for a high speed digital computer.

The automatic retrieval of data from a photoplate is relatively easy and was first achieved by Biemann's group in 1964.[45] The equipment necessary is basically a travelling microscope equipped with a photodensitometer. The microscope scans the photoplate and measures the density of any image at intervals of 0·25 microns. This density reading is converted to an electrical signal by a photomultiplier and becomes a y-coordinate which, together with its corresponding x-co-ordinate (the distance from the beginning of the spectrum) is fed to a high speed analogue-to-digital converter. The resulting digital data are discarded unless the y-co-ordinate is greater than a preset 'threshold' value, *i.e.*, a peak is being observed. If this is the case, the data are transferred to magnetic tape. In this way, tape is not wasted recording all the data points between peaks. It now is a relatively simple job for a digital computer to calculate the centroid of each peak, recognise the standard peaks, calculate the exact masses of the unknown peaks, separate the two and discard the former. The results of these operations are accurate mass *versus* intensity data for the unknown ions in the spectrum.

To raise raw data from the MS 902 to this level is somewhat more difficult, but has been accomplished.[46,47] For fast scans, the amplifier system must have a bandpass of at least 10 kc and it is primarily this that distinguishes the MS 902 from its predecessor, the MS 9. The output signal from a scan is recorded

[43] R. Ryhage, Proceedings of the ASTM, Committee E-14, Denver, 1967.

[44] No completely satisfactory tables exist, or could exist. The most popular, J. H. Beynon and A. E. Williams, 'Mass and Abundance Tables for Use in Mass Spectrometry', Elsevier Publishing Co., London, 1963, contain only C, H, O and N combinations, whilst an expensive alternative, D. D. Tunnicliff, P. A. Wadsworth, and D. O. Schissler, '16-Element Mass and Abundance Table with Supplement', 4 vols., Shell Development Co., Emeryville, California, 1965, places an upper limit of 10 on the number of carbon atoms considered.

[45] K. Biemann, P. Bommer, and D. M. Desiderio, *Tetrahedron Letters*, 1964, 1725; K. Biemann, Proceedings of the ASTM, Committee E-14, St. Louis, 1965.

[46] A. J. Campbell, J. S. Halliday, B. N. Green, T. O. Merron, and J. G. Murray, Proceedings of the ASTM, Committee E-14, St. Louis, 1965; C. Merritt, P. Issenberg, M. L. Bazinet, B. N. Green, T. O. Merron, and J. G. Murray, *Analyt. Chem.*, 1965, **37**, 1037.

[47] W. J. McMurray, B. N. Green, and S. R. Lipsky, *Analyt. Chem.*, 1966, **38**, 1194.

directly on an FM tape and this analogue tape is then digitised. Analogue-to-digital conversion in this case is an operation in which the computer measures and records the output voltage at given time intervals—optimally 10–15 micro-seconds in a total scan of 10 seconds which, even with a tape-slowdown factor of 32, requires a high-speed digitising system. The time centroid of each peak is then calculated and the resulting digital information (voltage *versus* time) corresponds to the intensity–distance data derived from a photoplate. In this latter case, however, distance is related simply to the square root of the mass whilst the time–mass relationship in the *quasi*-exponential magnetic scan is more complex. In practice, it has been demonstrated that given three known ions within a mass range of not more than 24 mass units, the mass of any other ion appearing within 12 mass units of either end of this range may be calculated sufficiently accurately. This may not encompass an unknown ion but if PFK is the internal reference, it must include a standard ion which will be recognised as such and so the range of 24 mass units may be shifted 12 mass units to lower mass and the search for an unknown ion resumed. As before a 'threshold' level eliminates the large distance between peaks, standard ions are discarded and the resultant mass-intensity digital data are collected on magnetic tape.

The final stage in the production of an element map is nothing more than a digital computer problem. Each mass must be correlated with a unique elemental formula. So-called ELCOMP programmes have been written which accomplish this in a variety of ways. The simplest and probably most common of these computes the mass of every possible combination of C, H, O and N and any other elements that are requested. If possible, upper and lower limits are placed on the number of atoms of each element present. If these are not known, rough limits can be estimated from the molecular ion. Each computed mass is compared to the observed mass and if the difference between them exceeds a preset error, typically 1–3 mmu, the formula is rejected and the next one is calculated. A self-consistency check may be applied to settle difficult choices such as that between say $C_{20}H_{37}O_5N^+$ and $C_{21}H_{39}O_5^+$ which may be tentatively resolved in favour of the latter if N is absent in all other ions. Various other programmes have been written to cope with this problem[48] and some very ingenious solutions have been proposed. One method[49] is to store all the calculated masses in the computer's memory, or 'core', and simply search for the correct one. A slightly different, and quite promising approach[50] requires the storage of C, O, N combinations in core and the combination of each of these with the appropriate number of hydrogens to arrive at the correct answer. Calculation of the formula from the mass defect of the ion, *i.e.*, the difference between the observed mass and the nearest lower integer has been attempted[51] and an ingenious method based upon a mass scale where $CH_2 = 14{\cdot}000000$ has been reported.[52] The programme of

[48] D. D. Tunnicliff, P. A. Wadsworth, and D. O. Schissler, *Analyt. Chem.*, 1965, **37**, 543, and references cited therein.
[49] A. L. Burlingame, private communication, 1966.
[50] E. Gilbert, unpublished work.
[51] J. Van Katwijk, *Appl. Spectroscopy*, 1964, **18**, 102.
[52] E. Kendrick, *Analyt. Chem.*, 1963, **35**, 2146.

choice depends upon the size and speed of the computer available, the goal being to deal with each mass as rapidly as possible certainly before the next mass on the tape arrives. This is fairly simple as the tape can be slowed or even halted but in the on-line work to be discussed, where the data are processed as they are produced, this refuge is absent and the fastest available systems prove to be barely fast enough.

As has already been mentioned assignment is facilitated considerably by the knowledge of the elements present in the molecule. One way to secure this information is by a careful study of the ions below m/e 100. It appears to be a fairly accurate assumption[53] that all elements of atomic weight less than 100 will appear in some ion below m/e 100 and since the accuracy of mass measurement required for assignment of formulae to these low masses is considerably relaxed, this promises to provide a simple technique for qualitative analysis. Armed with these data, the computer will be far more able to make the difficult choices that arise at higher masses.

The search for a satisfactory internal standard is being continued by many groups in two divergent directions. Improvement of the computational methods should in principle lead to a system in which perhaps ten reference ions each 100 mass units apart would be sufficient for mass measurement. Alternatively, with the rather inefficient software (programmes) presently available, accuracy could be vastly improved if standard ions were present at every mass unit, as would be the case if a straight chain hydrocarbon were to be used as a reference compound. This could, however, lead to the obscuring of hydrocarbon peaks from the sample under investigation and whilst deuterocarbons have been considered, they are relatively inaccessible and in any case fail to solve the problem because of the small separation of the $^{1}H_{2}$-^{2}H doublet.

A major advance that has appeared as fallout from the splutter of programming has been the appearance of bar-graph plotting routines. The input data for these are usually the mass against abundance figures and the output is a neat bar-graph, suitable for filing or publication. This is generally done off-line but has been done on-line[54] in which case, the scan speed of the spectrometer is badly reduced to the slow speed of the plotter.

Venkataraghavan, McLafferty and Amy[55] have reported some computer calculations, based upon photoplate data, in which overlapping bands are mathematically resolved. Such a deconvolution is successful in increasing the effective resolution by probably more than 300%, a result which should promote great activity in this direction.

When the computation techniques are rapid enough to deal with a whole spectrum of perhaps several hundred ions in less than a minute, the real possibility exists that tape recording may be abandoned and the analogue signal from the mass spectrometer fed directly into the computer. Such 'on-line' techniques have been developed to handle the data from a magnetic scan or from a

[53] E. Gilbert, H. M. Fales, and G. W. A. Milne, unpublished work.
[54] R. N. Stillwell, *Analyt. Chem.*, 1966, **38**, 940.
[55] R. Venkataraghavan, F. W. McLafferty, and J. W. Amy, *Analyt. Chem.*, 1967, **39**, 178.

photoplate reader although rather more effort appears to have been devoted to the former.[56] Quite apart from being an impressive *tour de force*, such a system possesses some real advantages. The problem of tape flutter is eliminated; a great deal of valuable computer time is saved; the whole system becomes integral and the interesting vista of feed-back control is opened up. This is clearly one area where the incredible gap between programming ignorance and computer capabilities is being closed a little.

Having arrived at a self-consistent listing of ions in terms of their elemental composition and abundance, the immediate problem is how best to present this. Biemann's element map[45] was the first attempt to solve this problem and in some senses has not subsequently been improved upon. A good feature of this element map, as may be seen from Figure 3, is that all the information is presented concisely yet fairly lucidly. The main criticism that has been levelled at the Biemann element map is that being digital in nature, it is not communicative as is a graphic display, such as a bar graph. In an attempt to cope with this difficulty Venkataraghavan and McLafferty[57] have devised a three-dimensional display, as shown in Figure 4, and met precisely the opposite problem. All the abundances may be seen at a glance, but the elemental compositions cannot. Yet another presentation has been devised[58] in which a complete bar graph is drawn for every combination of heteroatoms, as is shown in Figure 5. This is in some ways the clearest presentation of the three; abundances and compositions are fairly easily arrived at but the transfer of one's attention from one graph to another as a heteroatom is lost in what is, after all, a common enough type of fragmentation, is disconcerting, no less so than is the thought of such a representation for a compound containing six nitrogen atoms and six oxygen atoms.

The conclusion suggested by a study of all these Daedalian efforts is that the information has not perhaps at this stage been processed sufficiently for human digestion and that the computer must be allowed to take the problem further. How this can be done is discussed in the next section.

7 Computer Assisted Interpretation of Spectra

The computer handling of element maps is a branch of mass spectrometry whose age is better measured in months than in years. With the exception of one spectacular and heartening success, the effort so far invested in this problem has failed to yield any substantial dividends and the purpose of this section will be to discuss approaches and methods. Much of this material is unpublished and, as is the case with all unpublished work referred to in this Review, the reader is referred to future issues of Analytical Chemistry where the majority of the papers are expected to appear.

Even a casual glance at an element map will usually yield some structural

[56] W. J. McMurray, S. R. Lipsky, and B. N. Green; C. Merritt, P. Issenberg, and M. L. Bazinet, Proceedings of the ASTM, Committee E-14, Denver, 1967; A. L. Burlingame, D. H. Smith, and R. W. Olsen, *Analyt. Chem.*, 1968, **40**, 13.
[57] R. Venkataraghavan and F. W. McLafferty, *Analyt. Chem.*, 1967, **39**, 278.
[58] A. L. Burlingame and D. H. Smith, private communication.

DEOXYDIHYDRO-N_b-METHYLAJMALINE

	CH	CHO	CHN	CHNO	CHN2	CHN2O
94			6/8 0****			
95	7/11 0**		6/10-0***			
98				5/8 1******		
103	8/7-0***					
106			7/8 0**			
107			7/9-0*			
108			7/10-0***			
110			7/12-0****			
115	9/7-0***					
117			8/7 0**			
118			8/8 0*			
120			8/10 1**			
122			8/12 2*			
123			8/13 2*			
124				7/10 0***		
126				7/12-0**		
127	10/7-0**					
129			9/7 1*			
130			9/8 0***			
131			9/9-0****			
132			9/10-1***			
142			10/8-0*			
143			10/9-0**			
144			10/10-0****			
145			10/11-0***			
152			10/18-0****			
154			10/20-0**			
156			11/10-0*			
157			11/11-0***			
158			11/12 0***	10/8-0**		
159				10/9-0*		
160				10/10-0***		
167			12/9-0*			
168				10/18 0**		
170			12/12 0**			
173				11/11 0**		
181			13/11-0***			
182			13/12-0****	11/20 0******		
183					12/11 2**	
197					13/13 0****	
213					14/17-0***	
269						17/21 1*
326						21/30-0*

Fig. 3 *Element map devised by Biemann* et al. *The first column shows the nominal mass; the second, all the ions containing only* C *and* H*; the third, those containing* C, H *and one* O, *etc. The entry for each ion represents the number of* C *and* H, *the deviation in* mmu *from the theoretical mass and the relative intensity, on a logarithmic scale, represented by* 1—9 *asterisks.* (*Reproduced by permission from* Pure Appl. Chem., 1964, **9**, 104.)

information. For example, it may perhaps be seen at once that the molecular ion loses $\cdot CH_3$ and $\cdot COOH$ and these groups may therefore be assumed to be in the molecule. But this type of information could be extracted, albeit with less certainty, from a low resolution spectrum, and to treat an element map in this way is to ignore most of the information it contains. A serious consideration of every ion in the element map might, on the other hand, be expected to reveal

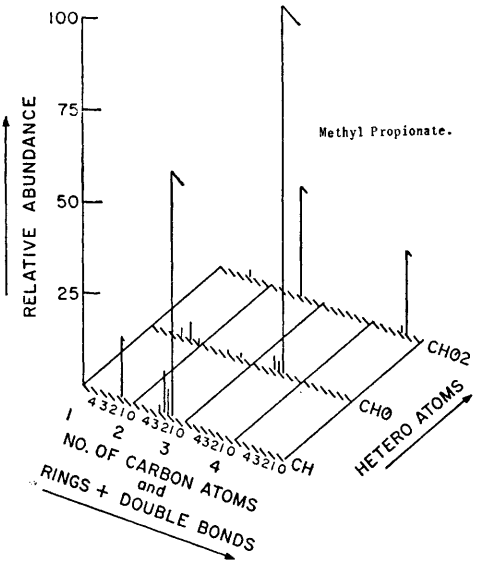


Fig. 4 *Element map devised by Venkataraghavan and McLafferty. Each major division in the X-axis represents the number of carbon atoms and each minor division a whole number of rings and double bonds. The Z-axis represents the heteroatom content and the Y-axis the abundance. (Reproduced by permission of F. W. McLafferty.)*

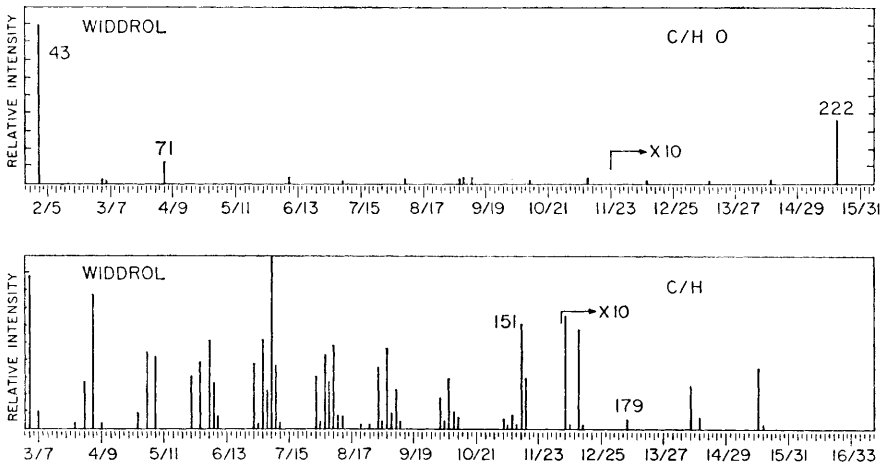
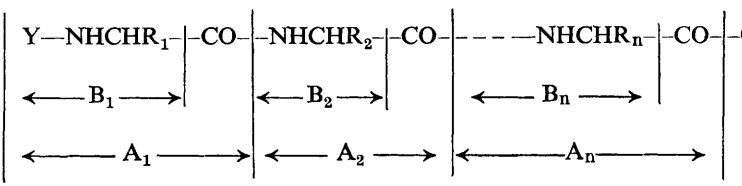


Fig. 5 *Element map devised by Burlingame and Smith. Each plot gives only ions with a specific heteroatom content. Each major division on the X-axis represents a* CH_2 *unit and is divided into* 14 *sub-divisions, each corresponding to one hydrogen or one mass unit. (Reproduced by permission from* J. Amer. Chem. Soc., 1967, **89**, 3232.)

both its source and its fate as will be discussed later. Such information, together with elemental composition data could hopefully be used to arrive at a structure for the molecule. This type of detailed study of element maps is far too time-consuming to be done manually but is extremely easy for even a primitive computer. The essence of the problem is to decide upon the general plan of attack to be adopted. Two approaches will be discussed here. The first relies upon existing knowledge of fragmentation patterns whilst the second assumes no prior information of this sort. The latter course may be intellectually more satisfactory but it is, so far, much less successful than the former approach, and it is very possible that a combination of the two will prove to be better than either one alone.

It has been known for some time[59] that in the mass spectra of peptides, the CO-N bonds are particularly vulnerable to cleavage and it follows therefore that the spectra must contain sequence information. This phenomenon was successfully exploited[60] in the sequence determination of fortuitine, a lipidoacylnonapeptide methyl ester. Much of this work was done manually with low resolution spectra, although the elemental compositions of some crucial ions were confirmed by peak matching.

Two outstanding papers, one from McLafferty's group,[61] the other from Biemann's group[62] represent the first reports of peptide sequencing by computer interpretation of high resolution spectra. The element map of the *N*-acyl peptide methyl ester is produced and used as the input information in the final phase of the computer handling. The peptide chain is assumed to cleave at every position A and also, less commonly, at position B. The spectrum is first searched for an ion of mass corresponding to *N*-acyl-glycyl (*i.e.*, A_1; R_1 = H). If this is absent, a check is made for *N*-acyl-alanyl (A_1; R_1 = Me) and so on until a positive



identification is made. When A_1 has been found, CO (27·9949) is subtracted from its mass to give the mass of B_1 and this confirmatory evidence is sought in the spectrum, although it may be absent if B_1 is correct. This done, a series of fragments A_2 (R_2 = H, Me . . . etc.) are added to A_1 until the continual checking reveals that the correct A_2 has been found. Once again, B_2 is sought and then

[59] K. Heyns and H. F. Grützmacher, *Tetrahedron Letters*, 1963, 1761; *Ann. Chem.*, 1963, **669**, 189; N. S. Wulfson, V. A. Puchkov, B V. Rozinov, A. M. Zyakoon, M. M. Shemyakin, Yu. A. Ovchinnikov, A. A. Kiryushkin, and V. T. Ivanov, *Tetrahedron Letters*, 1965, 2793.
[60] M. Barber, P. Jolles, E. Vilkas, and E. Lederer, *Biochem. Biophys. Res. Comm.*, 1965, **18**, 469.
[61] M. Senn, R. Venkataraghavan, and F. W. McLafferty, *J. Amer. Chem. Soc.*, 1966, **88**, 5593.
[62] K. Biemann, C. Cone, B. R. Webster, and G. P. Arsenault, *J. Amer. Chem. Soc.*, 1966, **88**, 5598.

the whole process is repeated until the molecular ion is found. This approach, which with minor variations is employed by both groups, has a number of corollaries. Side-chain fragmentation independent of the main A-B cleavages can create dilemmas. For example, the rearrangement of and loss of C_3H_6 from valine would result in its being indistinguishable from glycine, but when this happens, it seems never to be an exclusive process and the usual result is that both glycine and valine are indicated as the next amino acid in which case the choice of the latter is elementary. The peptide must be derivatised at both the N-terminal and C-terminal ends in order to be sufficiently volatile to give a mass spectrum. The C-terminal derivative is usually the methyl ester whilst the best of the N-terminal derivatives (Y) that have been studied[62] are acetyl, trideuterioacetyl, trifluoroacetyl and carbobenzoxy. No peptides of more than ten amino acids have been investigated in this way and the results suggest that the upper limit of molecular weight must be in this range, the limitation being volatility.

These methods are of course absolutely specific to peptides, which in terms of fragmentation at least, are remarkably uncomplicated, considering their molecular weight. The general problem of devising the software necessary to handle any molecule, or even any class of molecule other than peptides has not been solved, but work has been begun in this area. One approach[63] has been to cause a computer to go automatically through the interpretive operations that an experienced spectroscopist would apply to a high resolution spectrum, such as ring and double bond calculations, identification of heteroatoms in terms of functional groups and searching for homologous series of ions.

Another very interesting device[64] is that of the computer dialogue technique in which the spectroscopist can communicate with the computer *via* a keyboard, programme it with a set of fragmentation rules for the type of compound in question and then make the computer apply these rules and so use the spectrum to arrive at the structure of the compound in question.

A 'pathfinder' type of programme has been written[65] in which the computer considers each ion in turn and searches back up the element map listing until it finds the 'first possible parent ion' which must contain at least as great a number of atoms of each element as the daughter ion, from which it must differ by a 'permissible' fragment such as H but not a 'non-permissible' fragment such as H_2 or H_3. Thus $C_{10}H_{22}O^+$ could constitute a parent for $C_{10}H_{21}^+$ but not for $C_{10}H_{24}O_2^+$ or for $C_{10}H_{19}O^+$. In this way, an at least partially valid fragmentation map can be assembled and a great deal of structural information is sometimes uncovered.

Programmes have been written[66] which enable a computer to determine the structures of aliphatic hydrocarbons and fatty acids from their mass spectra by

[63] R. Venkataraghavan and F. W. McLafferty, Proceedings of the ASTM, Committee E-14, Denver, 1967.
[64] A. Mandelbaum, P. V. Fennessey, and K. Biemann, Proceedings of the ASTM, Committee E-14, Denver, 1967; K. Biemann and P. V. Fennessey, *Chimia* (*Switz.*), 1967, **21**, 226.
[65] H. M. Fales, E. Gilbert, P. G. Gordon, R. J. Highet and G. W. A. Milne, unpublished work.
[66] B. Pettersson and R. Ryhage, *Arkiv Kemi*, 1967, **26**, 293; *Analyt. Chem.*, 1967, **39**, 790.

the application of simple empirical rules of fragmentation. They appear to give incorrect answers very rarely but give no answer at all in about 20% of the cases studied.

Some very significant work has recently been reported by Lederberg[67] from the Stanford Artificial Intelligence Project. This group has devised a programme which generates all the possible structural isomers corresponding to a molecular formula. It can furthermore avoid generating those isomers which contain chemical absurdities such as –NH–O–O–H and can arrange the remainder in order of plausibility. Thus given a high resolution mass spectrum in which the molecular ion is identified, this list is constructed and 'heuristically' searched (*i.e.*, the search is not comprehensive, as the time required for this would be prohibitive, but is conducted in intelligently selected areas) for a structure consistent with the mass spectrum. This so-called DENDRAL 64 programme is written in LISP, a computer language far better suited to the representation and manipulation of chemical structures than is the more commonly used FORTRAN, and the success attending this departure from convention is educational. The programmes so far completed cannot handle cyclic structures but this problem is said to be now conceptually solved.

A very desirable feature of any data processing system would be access to information that would permit the definitive conclusion that $A^+ \rightarrow B^+ + C$ does occur with the molecule in question. Metastable transitions do provide this information and their automatic recording and interpretation is possible,[68] but they are observable only at low resolving power, thus a second spectrum is required. Moreover, the appearance of a metastable ion is a matter of chance, depending as it does upon a number of variables such as the quite unpredictable rates of fragmentation reactions. Mainly for this last reason, metastable ions constitute only ancillary evidence and an interpretive system built around them would be severely defective.

Of far greater promise is the technique developed by Jennings[69] who has taken advantage of the fact that if the process $M_1^+ \rightarrow M_2^+ + M_3$ occurs in the field-free region between the two sectors of the MS 902, the daughter ion will have been accelerated as M_1^+, but will be magnetically deflected as M_2^+. With a simple change in the circuitry of the instrument, it is possible to study a daughter ion, M_2^+ and identify its parent unequivocally. An interesting outcome of these studies has been the discovery that a daughter ion usually has several parents. This method has been applied recently to a problem in peptide sequencing[70] and the technique is now semi-automatic.

[67] J. Lederberg, NASA Doc. CR-57029, 1964; CR-68898, 1965; CR-68899, 1966. *Proc. Nat. Acad. Sci. U.S.A.*, 1965, **53**, 134. See also, Stanford Artificial Intelligence Project, Memos 49 and 54.

[68] R. E. Rhodes, M. Barber, and R. L. Anderson, *Analyt. Chem.*, 1966, **38**, 48; N. R. Mancuso, S. Tsunakawa, and K. Biemann, *Ibid.*, 1966, **38**, 1775.

[69] K. R. Jennings, *Chem. Comm.*, 1966, 283. See also M. Barber, K. R. Jennings, and R. E. Rhodes, *Z. Naturforsch.*, 1967, **22a**, 15.

[70] M. Barber, W. A. Wolstenholme, and K. R. Jennings, *Nature*, 1967, **214**, 664.

8 Summary

High resolution mass spectrometry is now second only to nuclear magnetic resonance spectroscopy in the organic chemist's armamentarium and it seems fairly clear, its expense notwithstanding, that it is only a matter of time before it assumes first place. Its overwhelming advantage over other techniques is its prodigious sensitivity, rivalled only by that of the scintillation counter. The amount of information per microgram of sample provided by the mass spectrometer permits entire research projects to be carried out at the sub-milligram level and the impact of such a technique in areas of applied organic chemistry, such as biochemistry, is large and obvious.

As a field, high resolution mass spectrometry embraces many disciplines, ranging from mathematics to biochemistry, and perhaps because of this, is an area in which iconoclasm is often attended by success as has been amply demonstrated in the recent past. The resultant absence of dogma is invigorating and invests the whole field with tremendous promise for the future.

I thank Dr J. Daly, Dr H. M. Fales, Dr P. G. Gordon and Dr R. J. Highet for many helpful discussions and suggestions.