Pure & Appl. Chem., Vol.54, No.2, pp.267-279, 1982.
Printed in Great Britain.

0033-4545/82/020267-13$03.00/0
Pergamon Press Ltd

NEW METHODS FOR THE CHARACTERISATION OF BIOPOLYMERS

Howard R. Morris, Anne Dell, A. Tony Etienne, Martin Judkins,
Roy A. McDowell, Mario Panico and Graham W. Taylor

Department of Biochemistry, Imperial College of Science and Technology.
London SW7 2AZ UK.

Abstract - One of the most interesting classes of polymer are those found in living systems, the biopolymers. Within this class, the proteins are perhaps the most fascinating with respect to their versatility of structure and function, size and physico-chemical characteristics such as solubility, elasticity etc. The covalent structure determination of proteins remains a daunting task, and in particular post-translational modifications, now recognised as a common feature in proteins, limit the scope and utility of classical chemical methods. This paper describes recent advances in mass spectrometric methods for the structure elucidation of biopolymers giving examples of total protein structure, and including peptidolipids, glycoproteins and carbohydrates.

INTRODUCTION

Biopolymers represent one of the most interesting areas of modern research in the field of polymer chemistry. There are many lessons to be learned from the study of natural polymers, in particular how their surprising versatility is governed by chemical charge, size, hydrogen bonding etc. This versatility is particularly apparent in the field of protein chemistry, where we find structural proteins - some elastic, some rigid - globular proteins - the enzymes and hormones - and even very small polymers such as the enkephalin pentapeptides, which possess striking biological potency. Structural studies in all these areas are clearly important in order to advance our understanding of the function of biopolymers. The first and often most difficult level of structural analysis is the determination of primary structure - i.e. the sequence of amino acids in the protein. Significant advances have been made in this work over the past ten years, and here we discuss new mass spectrometric methods for the structure elucidation of this class of biopolymer.

PROTEIN/PEPTIDE SEQUENCING

The mass spectrometer is not an obvious choice for sequencing polar, high mass species as, in general, samples must be volatile and of relatively low molecular weight. The success of MS in this field has arisen from developments in methodology in the generation of suitably sized, volatile protein fragments, coupled with the "mixture analysis" capability of the technique (1,2). The protein is digested with a non-specific protease (i.e. elastase) and the fragments converted to volatile N-acetyl N,O permethyl derivatives. In the ion source, the peptide mixture is fractionally volatilised over a temperature range of 100-350°C. By

Thr-Ala-Phe-Leu-Trp-Ala-Gln-Asp-Arg-Asp-Gly-Leu-Ile-Gly-Lys-Asp-Gly-His-Leu-Pro-Trp-His-Leu-Pro-Asp-Asp-Leu-His-Tyr-Phe-Arg-Ala-Gln-Thr-Val-Gly-Lys-Ile-Met-Val-Val-Gly-Arg-Arg-Thr-Tyr-Glu-Ser-Phe-Pro-Lys-Arg-Pro-Leu-Pro-Glu-Arg-Thr-Asn-Val-Val-Leu-Thr-His-Gln-Glu-Asp-Tyr-Gln-Ala-Gln-Gly-Ala-Val-Val-Val-His-Asp-Val-Ala-Ala-Val-Phe-Ala-Tyr-Ala-Lys-Gln-His-Leu-Asp-Gln-Glu-Leu-Val-Ile-Ala-Gly-Gly-Ala-Gln-Ile-Phe-Thr-Ala-Phe-Lys-Asp-Asp-Val-Asp-Thr-Leu-Leu-Val-Thr-Arg-Leu-Ala-Gly-Ser-Phe-Glu-Gly-Asp-Thr-Lys-Met-Ile-Pro-Leu-Asn-Trp-Asp-Asp-Phe-Thr-Lys-Val-Ser-Ser-Arg-Thr-Val-Glu-Asp-Thr-Asn-Pro-Ala-Leu-Thr-His-Thr-Tyr-Glu-Val-Trp-Gln-Lys-Lys-Ala

Fig. 1 The Sequence of Metotrexate-resistent Dihydrofolate Reductase. This was determined solely by mass spectrometric analysis.

comparing the spectra produced at each temperature - observing the concomitant increase in ion intensity as each compound volatilises - the "pure" spectrum of each component of a

mixture of up to five peptides can be determined. Under electron impact conditions the peptide derivative fragments to give "sequence ions" from which the sequence can be deduced (see Ref.1 for a more detailed discussion). Using such procedures both total protein sequences (Fig.1) and innumerable peptide sequences have been determined in this laboratory. Blocked N-terminal peptides can cause major problems in classical sequencing studies. In contrast, MS is fully capable of handling such problems, since the peptide is deliberately 'blocked' during derivatisation prior to analysis. For example, classical sequence analysis of the cytotoxic peptide 'delta haemolysin' (from S. aureus) indicated the N-terminus was blocked. The peptide was deuteroacetyated and permethylated under standard conditions and subjected to EIMS. Signals were observed in the spectrum corresponding to N-formyl-Met (m/e 174) and N-formyl-Met-Ala (m/z 259) defining the blocking group and N-terminal sequence (3). Isotopically labelled reagents (i.e. d_6 acetic anhydride) were used in order to allow identification of any putative natural N-acetyl group in the peptide. Blocked N-terminal sequences are now routinely analysed by EI methods in this laboratory using only 10-50 nmol of material.

NOVEL STRUCTURES

Perhaps the most obvious strength of MS is in the area of new structures where no reference compounds exist. Such a situation arose in sequencing studies on the blood coagulation zymogen Prothrombin. Peptides isolated from the N-terminal region migrated on HVPE with a mobility inconsistent with the classically determined structures. Mass spectrometric analysis confirmed the presence of an anomalous structure in these peptides; full interpretation of the spectra resulted in the structure determination and location in prothrombin of the novel residue γ-carboxyglutamic acid (Gla) (Fig.2). Gla has also been detected and located by this MS method in the N-terminal portions of Factors X_1 and X_2 (1). In a similar manner, the identification and location of other modified amino acids (such as N-methyl lysine) reduces to an almost trivial problem.

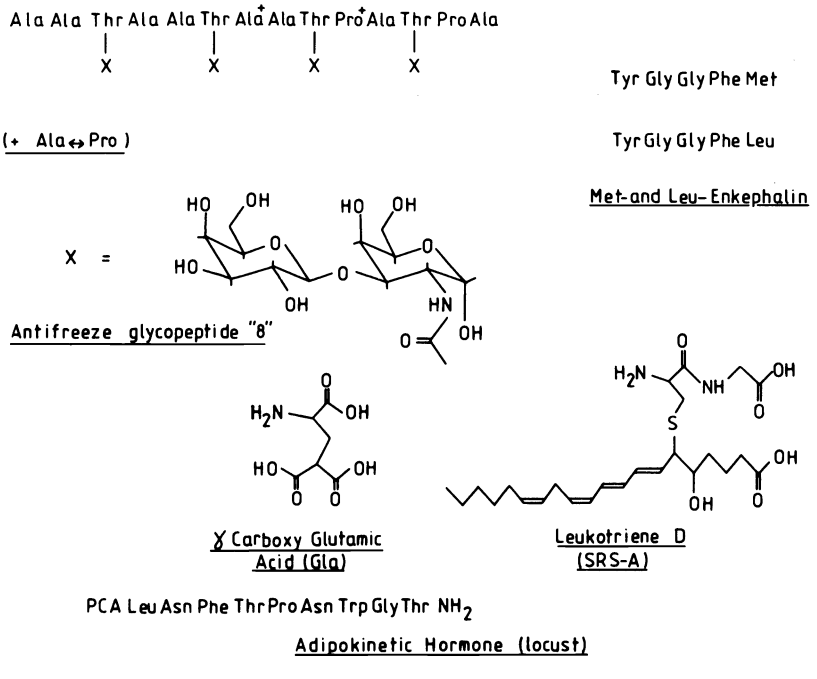


Fig. 2. Important biological substances whose structures were derived by MS.

MS sequencing, unlike classical peptide analysis, is not limited to the "normal" α-amide bond linkages. For example cross linking of the cell surface protein Fibronectin occurs via the formation of a γ-glutamyl lysyl amide bond. The glutamine (Gln) residue specifically involved in the cross-linking has been identified using EIMS. Following competitive labelling studies with a radioactive marker (putrescine) labelled peptides were isolated and subjected to EIMS analysis. From the spectrum obtained (ions at m/z 98, 211, 469 and 497) the partial sequence pGlu-Ala-Gln-putrescine was defined (and confirmed by analysis of the corresponding unlabelled peptide). These data allowed us to determine the cross-linking site at position 3 (Gln) of the fibronectin sequence(4). Because MS is independent of sample class it is particularly well suited to the investigation of compounds of completely unknown structure. A recent example of this was the structure elucidation of slow-reacting substance of anaphylaxis (SRS-A) from lung (SRS-A is a bronchoconstrictor, released in picomole quant-

ities, which is associated with the asthma crisis). Following protein chemical studies on the purified material, a mass spectrum was obtained of the trimethylsilyl acetyl methyl ester derivative on 5µg of material. To enable the interpretation of the SRS-A spectrum (by differentiating significant ions from the background impurities) mixture analysis was used on the 1:1 $CH_3CO:C^2H_3$ CO derivative; ions containing this isotopic label appeared as 1:1 doublets 3 m.u. apart (demonstrating the presence of one free amino group in the natural product) and could be readily located in the spectrum. A detailed analysis of the spectra of labelled SRS-A allowed the structure assignment of the novel peptidolipid 5-hydroxy-6-cysteinylglycinyl-7,9,11,14-eicosatetraenoic acid (Fig.2) (2). This was later termed Leukotriene D.

NEURO CHEMISTRY

The explosion of research into neuroactive peptides has arisen from a realisation of their fundamental role in the control of major body functions; further, the action and inter-relationship of many neuropeptides has led to a belief that an imbalance of these substances may give rise to many mental and physical illnesses including Senile Dementia,Schizophrenia, Parkinsonism etc. MS has already played an important part in this field, for example in structure elucidation of the endogenous opioid peptides - the enkephalins - and, in the first characterisation of a hormone of insect neuroendocrine origin, the locust adipokinetic hormone (1). This blocked decapeptide, PCA-Leu-Asn-Phe-Thr-Pro-Asn-Trp-Gly-Thr-NH_2,could not be sequenced classically but presented few difficulties for MS (Fig. 2). In general however, problems of quantity and purity of neuropeptides in biological fluids preclude MS fromplaying a wider role in neurochemical analysis. To ameliorate these problems, and to aid characterisation of these substances, a high resolution reverse phase HPLC system was developed in this laboratory (2). Using this system 18 neuropeptides of general interest can be separated in one forty minute HPLC run with invariant retention times and in high yields. This system has found wide use in our studies on opioid peptides and precursors, and our investigationsinto the role of neuropeptides in disease states.

A pituitary - derived peptide, with specific aldosterone-stimulating activity on the adrenal zona glomerulosa, was purified to homogeneity in the HPLC system. The peptide was analysed both by protein chemical techniques and by EIMS, and the data was used to determine the structure as a blocked 13 residue peptide: Ac-Ser-Tyr-Ser-Met-Glu-His-Phe-Arg-Trp-Gly-Lys-Pro-Val-NH_2. This is identical to the pro-opiocortin-derived peptide α-melanocyte stimulating hormone (α-MSH), and indicated for the first time that α-MSH may play a role in pituitary regulation of the adrenals (5). Two other peaks of biological activity, eluting just after α-MSH on HPLC, have now been characterised by the nascent technique of Fast Atom Bombardment (FAB) MS (see FD/FAB).

In the analysis of neuropeptides, MS offers an accuracy and specificity not obtainable with other structural probes (e.g. radioimmunoassay). In order to compete most effectively however, picomole sensitivity is required; for this we have turned to Negative Chemical

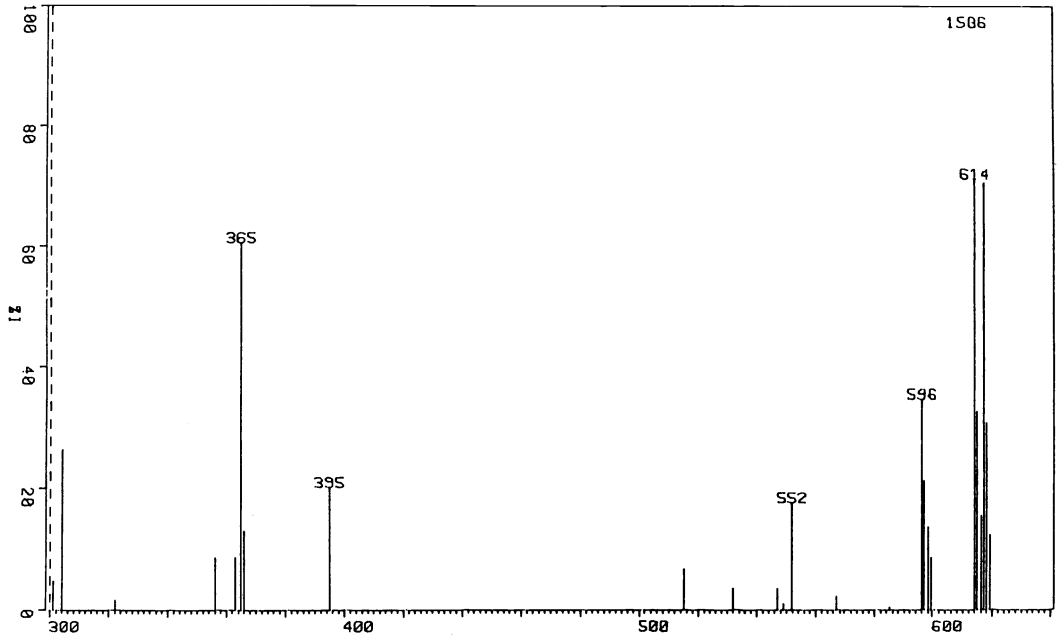


Fig. 3. NCI spectrum of the 1:1 (acetyl: ^{3}H-acetyl) perfluorobenzyl ester of met-enkephalin.

Ionisation (NCI) MS (6). We have already achieved nanomole sensitivity (with full structural information) for perfluoroacyl derivatives of the enkephalins and provided a careful choice of derivative and running procedures is made, the sensitivity can be improved further into the picomole range. This is exemplified in Fig. 3, which shows the NCI spectrum of the N-acetyl (1:1 CH_3CO:C^2H_3CO) perfluorobenzyl ester derivative of Met-enkephalin. Confidence in assignments is aided by the isotope label which results in all ions containing the N-terminus exhibiting a 1:1 doublet separated by 3 m.u. The major high mass signal is the intense doublet at m/z 614/617 corresponding to loss of the perfluorobenzyl group from the molecular ion (resulting in a stabilised carboxyl anion). The intensity of these signals indicated that it should be possible to monitor them for quantitative purposes. In preliminary experiments involving narrow scanning across the mass range spanning this doublet of peaks, we have shown that the derivative can be readily detected at the 50 picomole level. We are now extending these studies by making use of the increased sensitivity resulting from specifically monitoring the ion current at the appropriate mass (selected ion monitoring). In earlier EIMS studies of a Leu-enkephalin derivative we were able to succesfully detect and specifically quantitate, in the presence of other peptides, 25 picomoles of the neuropeptide using selected ion monitoring procedures. Hence we are confident that when this procedure is fully exploited in the NCI mode on derivatives such as those described, picomole sensitivity will be realisable for many neuropeptides.

GLYCOPEPTIDE ANALYSIS

Glycoproteins play a major role in living systems with such diverse functions as disease control, cell surface recognition, blood clotting etc. Despite many years of research the determination of their structure has proven an especially difficult problem requiring the solution of (i) protein sequence (ii) carbohydrate sequence (iii) nature of the protein-carbohydrate linkage and (iv) nature and configuration of the carbohydrate linkages. Mass spectrometry offers great advantages in glycopeptide sequencing because it is not restriced to a particular compound class. This was clearly demonstrated in the structure determination of an antifreeze glycopeptide "AF8" from antarctic fish blood (1). Analysis of the N-acetyl-N, O-permethyl derivative in both electron impact (N-terminal information) and chemical ionisation (C-terminal information) modes gave the (heterogeneous) peptide sequence, defined the sugar as a disaccharide, and specified linkage positions. Analysis of the sugar as a trimethylsilyl derivative showed it to be β-D-galactosyl (1-3) α-N-acetylgalactosamine (Fig. 2). Extension of these studies has resulted in structural data on the carbohydrate portion of prothrombin (7). New data on a series of glycopeptide antibotics and high mass oligosaccharides will be described later (FD/FAB).

High Field Magnet

In many areas of research today there is a growing need for reliable data on molecular weights of species in the 2000-4000 mass range. Problems such as the degree of polymerisaction and the "purity" of polymer produced confront the plastics industry every day.

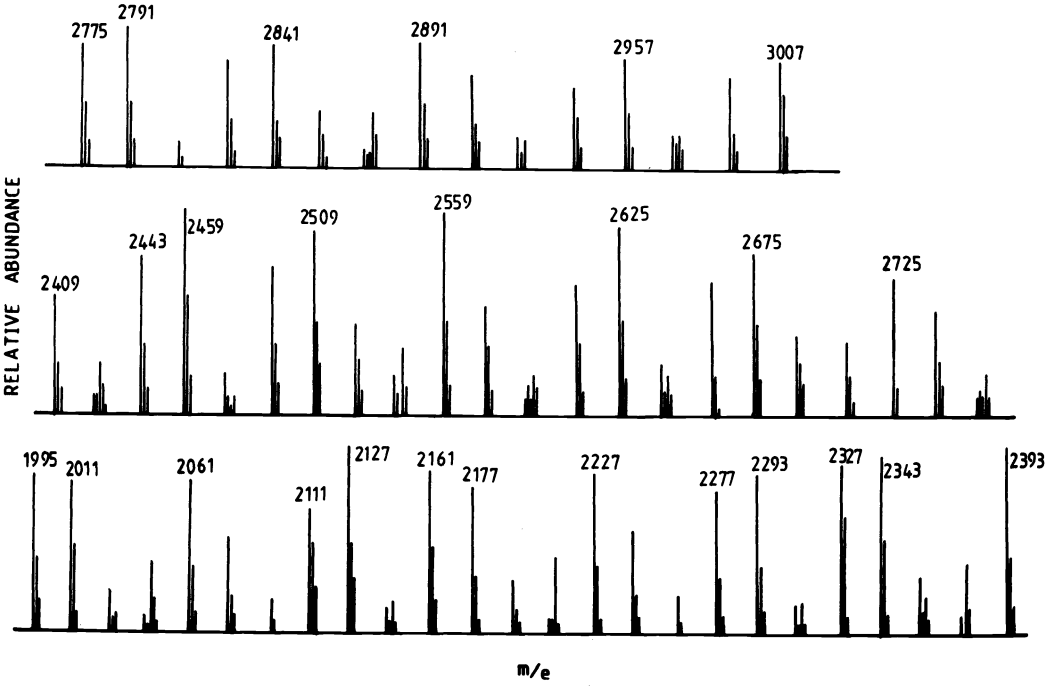


Fig. 4. NCI spectrum between m/z 2000-3000 of Fomblin oil. This material is also used as a high mass reference for FD MS studies.

In our own field, high mass molecular weight information would greatly facilitate the structure elucidation of unknown biological samples. With these problems in mind a High-Field Magnet (HFM) was commissioned for this laboratory enabling an extension of the mass range of conventional mass spectrometers from 1000 to 3000 m.u. at full accelerating voltage and thus full sensitivity (9) (Fig. 4). Apart from the obvious advantages obtained from the HFM (notably increased sensitivity at high mass), the development of the HFM initiated the production of extended range instruments (up to 1800 m.u.) now available from several manufacturers.

The HFM has proved invaluable in our EI and CI studies on peptides and carbohydrates, but perhaps the most successful application of the technology has been in the analysis of high mass, polar and thermally labile substances by Field Desorption (FD) MS and the recently developed technique of Fast Atom Bombardment MS.

FD/FAB

A major limitation of the EI and CIMS techniques described earlier is their requirement for sample volatility. Although this problem has been overcome in many cases by the formation of suitable derivatives (e.g. acetylation/permethylation of peptides), there has always been a need for an MS method which can handle polar, involatile or thermally labile substances, without prior derivatisation (and possible attendent side reactions!). Both field desorption (FD) MS and the recently developed technique of fast atom bombardment (FAB) MS fit the bill. Under FD conditions the sample is ionised and desorbed from an activated tungsten wire by the action of an electric field gradient. This method imparts little internal energy to the sample and results in quasimolecular ions and very little fragmentation. This technique is now well established (9), and has been applied to the solution of many biological problems where classical EIMS has failed.

FD MS has been used in this laboratory to obtain molecular weight data on underivatised antibiotics (Vancomycin, Echinomycin), vitamin B_{12} and its precursors, and, most recently, on high mass oligosaccharides and the bleomycin family of glycopeptide antibiotics. The bleomycins are particularly intractable molecules for MS due to their polarity and thermal instability. In addition, their complexity precludes the use of simple chemical procedures for the preparation of volatile derivatives suitable for EI or CI analysis. Recently we have reported the first successful characterisation by FD MS of bleomycin B_2 and related phleomycins (10). We have now extended this work to other bleomycins and have obtained excellent FD spectra on bleomycin A_2, bleomycin demethyl A_2, tallysomycin A, and Blenoxane (both free and Cu complexed) both as the native materials and as acetylated derivatives (Fig. 5).

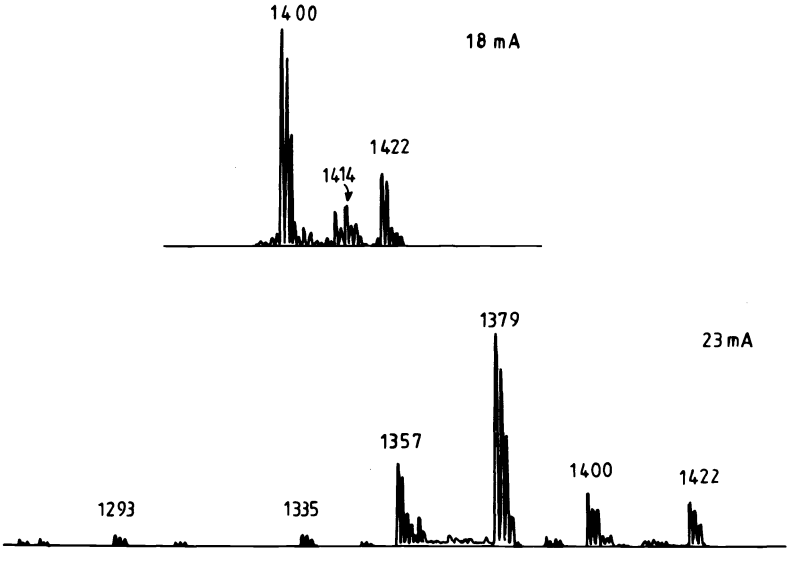


Fig. 5. The FD spectrum of Bleomycin demethyl A_2. The emitter current was increased from 0-30mA. An intense quasimolecular ion ($M+H^+$, m/z 1400) was observed at 18mA with fragmentation occurring at 23mA.

The success of these studies is partly due to the HFM facility - FD data being obtained at full sensitivity - and partly due to the application of sample handling procedures suitable for the FD analysis of thermally fragile and polar molecules. In our hands FD spectra can be obtained routinely; nevertheless considerable skill and experience is required.

In contrast, fast atom bombardment (FAB) MS yields data on polar labile compounds with relative ease. Indeed, in combination with the HFM facilities, this new technique represents a very powerful new tool for the analysis of biopolymers and other high molecular weight species. For FAB MS the sample is loaded in a glycerol matrix on a metallic target and bombarded with a beam of accelerated argon atoms. The sample is ionised by the impinging atoms and long lived ion beams are generated, which are then mass analysed. Both molecular weight and fragment information can be obtained by FAB MS. To complement our FD MS work, we have examined a variety of bleomycins by FAB MS; all have yielded excellent spectra with abundant quasimolecular ions. A typical example is given in Fig. 6a. This shows the FAB spectrum of bleomycin B_2; note the $M+H^+$ species at m/z 1425. Simple chemical modifications can be performed in the glycerol matrix and monitored by FAB MS. Fig. 6b shows the spectrum obtained frum bleomycin B_2 after the addition of $(CH_3CO)_2O/(C^2H_3CO)_2O/CH_3OH$ (1:1:4) to the matrix. The sample is converted mostly to the diacetyl derivative within one minute as shown by the 1:2:1 pattern at m/z 1509,1512,1515 (11).

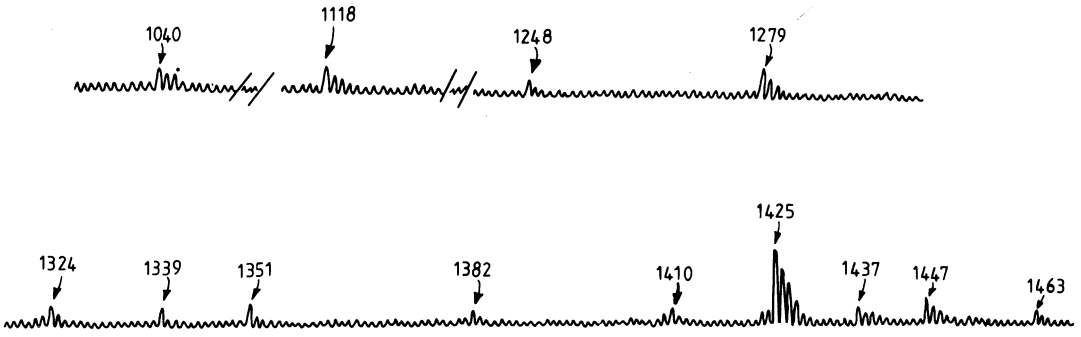


Fig. 6a. Fast Atom Bombardment spectrum of Bleomycin B_2 showing the quasimolecular ion ($M+H^+$, m/z 1425) and fragment ions above m/z 1000.

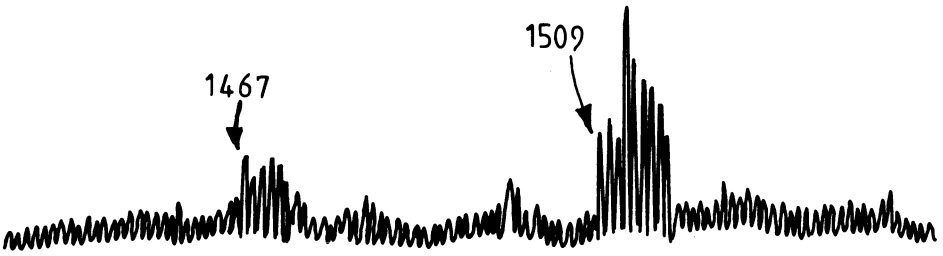


Fig. 6b. Molecular ion region of FAB spectrum obtained after addition of acetic anhydride:^{2}H-acetic anhydride to Bleomycin B_2 in the glycerol matrix.

Both FD and FAB gave molecular weight information on the bleomycins, although, because of greater fragmentation, the FAB spectra are more complex. The sensitivity of the two techniques is comparable; however, unlike FD MS, more intense spectra may be obtained with FAB by increasing sample loadings in the matrix.

We are currently applying FAB MS to our continuing studies in the peptide area. Peptides, of both known and unknown structures, have been examined at the low µg level in both the positive and negative modes. These include both free and blocked peptides in the molecular weight range 400-3000, containing examples of all the protein derived amino acids. All peptides studied give intense quasimolecular ions, together with, in many cases, fragmentation from the N-and C-termini. If fragmentation takes place it occurs via chemical ionisation pathways; no N-C cleavage ions are observed. To facilitate interpretation of the fragmentation patterns the N-(acetyl:^{2}H-acetyl, 1:1) derivative is prepared. N-terminal

ions appear as 1:1 doublets 3 m.u. apart, whereas ions arising from C-terminus are singlets unless they contain Lys or Arg which may also acetylate. We have applied this method to assist colleagues at Cambridge (Dr J. Walker, MRC Molecular Biology Lab.) to complete the sequence of a bovine ATPase derived peptide. Consideration of the FAB spectra of both the free material and its 1:1 acetyl:^{2}H-acetyl derivative allowed the assignment of the partial structure H...Asp-Ala-Thr-Thr-Val...OH. In contrast to the classically derived data the FAB sequence is homologous with the ATPase peptide from E.coli which strongly suggested the FAB derived sequence was correct. Recently, the sequence has been confirmed and extended by classical methods (12). FAB MS has been applied to the solution of many unknown peptide sequences including blocked N-terminals. It is used routinely to obtain molecular data (and sometimes even sequences) at the 1-5 nanomole level. In our current research we have extended this further to the picomole range by a combination of utilising xenon as the bombarding species and using HCl/glycerol for sample loading. This allows observation of quasimolecular ions at the one picomole level for some peptides and, in favourable cases sub-picomole detection. Fig. 7 shows the excellent signal/noise ratio on 15 picomoles of the hexapeptide Lys-Phe-Ile-Gly-Leu-Met NH_2 run under the above conditions.

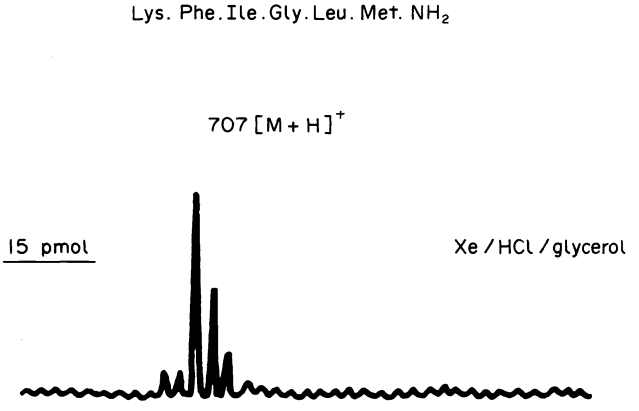
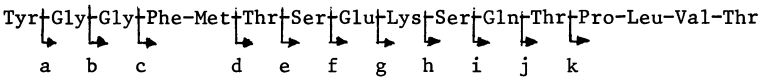


Fig. 7. FAB spectrum of 15 picomoles of the hexapeptide Lys-Phe-Ile-Gly-Leu-Met-NH_2.

The sensitivity of the technique in the neuropeptide field will be of great advantage, intense quasimolecular ions have been produced from Lys-vasopressin ($M+H^+$ 1056; $M+Na^+$, 1078; $M+K^+$, 1094; 2 nmol), α-endorphin ($M+H^+$,1745;300 pmol), and α-MSH ($M+H^+$,1664; 300 pmol). Somewhat higher loadings of α-endorphin resulted in the production of C-terminal ions at m/z 1582,1525,1468,1190,1089,1002,873,745,658,530 and 429 (a-k): Fig. 8).



FAB has already proven invaluable in our pituitary/adrenal work on α-MSH described earlier (neuropeptides section). The two biologically active peptides eluting after α-MSH on HPLC were subjected to FAB MS. The first peptide gave an intense quasimolecular ion at m/z 1706, 42 m.u. higher than that obtained from α-MSH (Fig. 9). Further, this peptide was readily converted to α-MSH on mild base treatment and was indistinguishable from it with respect to amino acid content and EI spectra. It was concluded that the peptide was an O-acetyl derivative of α-MSH; further EI analysis located the acetyl group on the N-terminal acetyl-serine (13). The second peak gave intense quasimolecular ions at m/z 1193 $(M-H)^-$ and 1195 $(M+H)^+$. This, together with EI MS data on the acetyl permethyl derivative was used to deduce the sequence Val-Val-Tyr-Pro-Trp-Thr-Gln-Arg-, a fragment of haemoglobin! We are now investigating the biological role of this peptide.

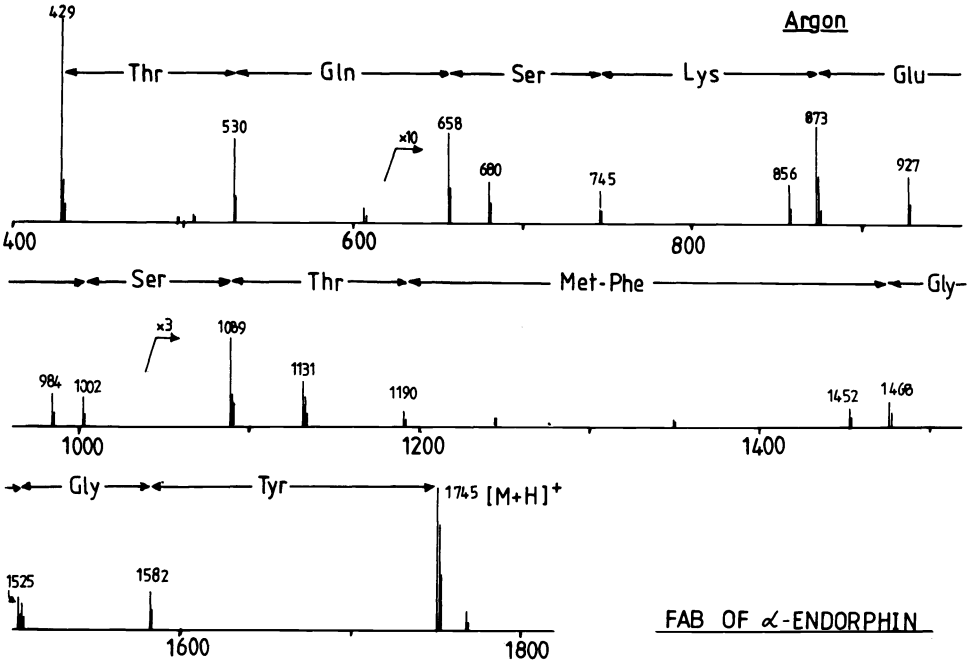


Fig. 8. FAB spectrum (above m/z 400) of α-endorphin.

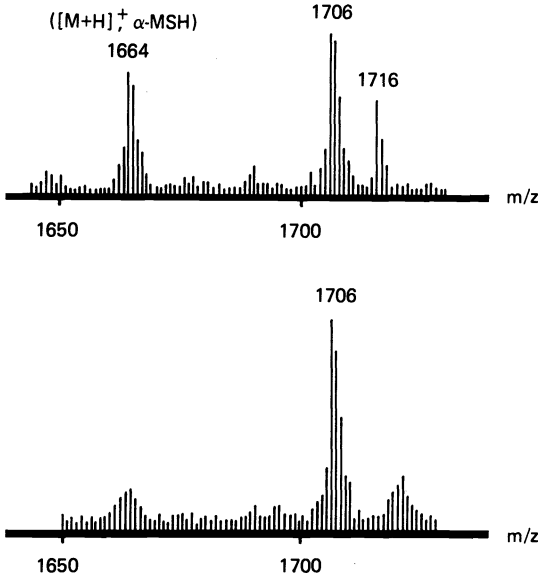


Fig. 9. FAB mass spectra of the two adjacent fractions, upper and lower traces respectively, eluting immediately after α-MSH, obtained from HPLC of a pituitary extract. Signals at m/z 1664 and 1706 are assigned to the pseudomolecular ions of α-MSH and its acetylated derivative respectively.

The mixture analysis capability of MS has already been extensively exploited in EI studies of peptides. Our FAB studies of peptide mixtures have demonstrated that selective ionisation of components can occur. For example FAB of a four component mixture yielded molecular weight and sequence information from Val-Gly-Leu-Ala-Pro-Val-Ala, molecular weight data only from PCA-Tyr-Tyr-Thr-Val-Phe-Asp-Arg and the corresponding Gln peptide, while no data was obtained on the fourth component Ala-Asn-Asn-Lys except under extreme conditions (a weak quasimolecular ion was obtained from the acetyl peptide after total glycerol evaporation) (14). This illustrates the dangers of placing full reliance on any one technique in structural studies.

We are now exploiting (i) the potential of FAB and FD for the ionisation of large involatile substances, and (ii) the ability of our high field magnet facilities to focus high mass ions at good sensitivity, in the study of biological polymers in the mass range 2500-5000 dalton. We have examined a variety of polypeptides up to 3000 M.Wt and all have afforded good FAB spectra, readily countable to the molecular ion. The type of data obtainable is illustrated in Fig. 10. which shows the FAB spectrum of a commercial sample of Melittin, a 26 residue peptide from bee venom. Protonated molecular ions are present at m/z 2845 and 2873 for the free and formyl peptides respectively. Partial structural information can be extracted

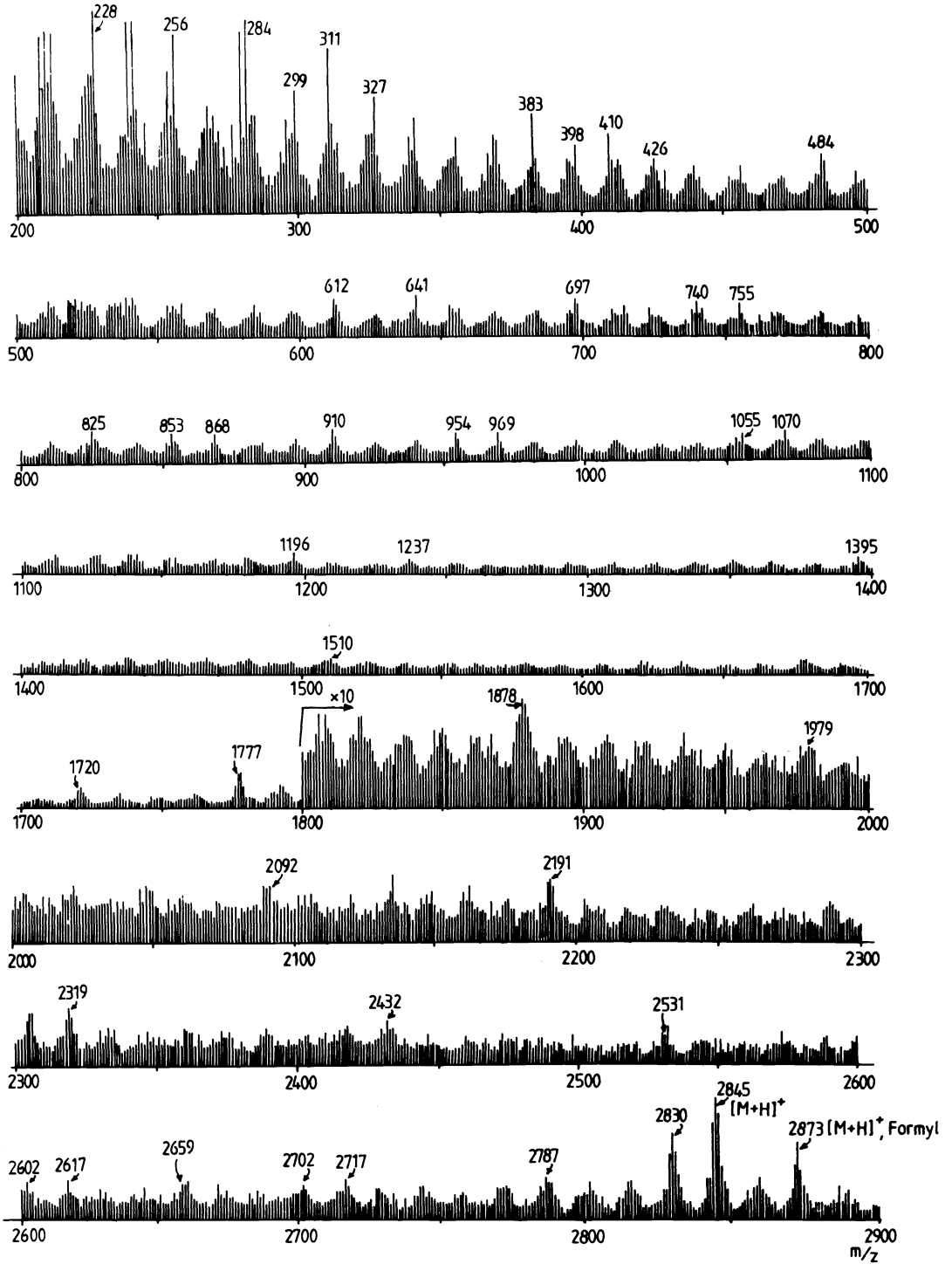


Fig. 10a. The FAB spectrum of Melittin showing pseudomolecular ions for the free and N-formyl species.

from the complex fragmentation pattern, for example C-terminal sequence ions corresponding to the first 12 residues are present above m/z 1700 at m/z 1720, 1735, 1777, 1792, 1878, 1979, 2092, 2191, 2319, 2431, 2432, 2532, 2602, 2659, 2660 and 2787. Interpretation of the Melittin spectrum is, however, considerably facilitated by a knowledge of the sequence, and it should be recognised that the complete characterisation of an unknown peptide of this size requires FAB and/or EI studies of enzyme digests to complement and aid the interpretation of the data obtained from the intact peptide.

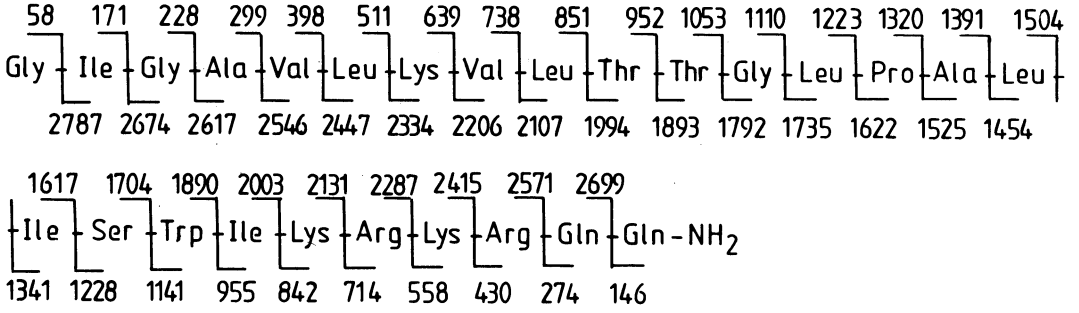


Fig. 10b. Structure of Melittin showing masses of the ions expected from amide bond cleavages.

Peptides above m/z 3000 are also amenable to FAB MS, albeit at lower instrument sensitivity. For example, both glucagon (M.Wt.3480) and the oxidised B chain of insulin (M.Wt.3494) give good molecular ions and typical spectra are shown in Fig. 11. Further, it is even possible to obtain a weak molecular ion from a peptide as large as adrenocorticotropin (ACTH), which has a molecular weight of 4537 dalton. Because of the low abundance of background ions above m/z 3000 manual counting is often not possible and the spectra require mass marking with suitable reference compounds for FAB studies.

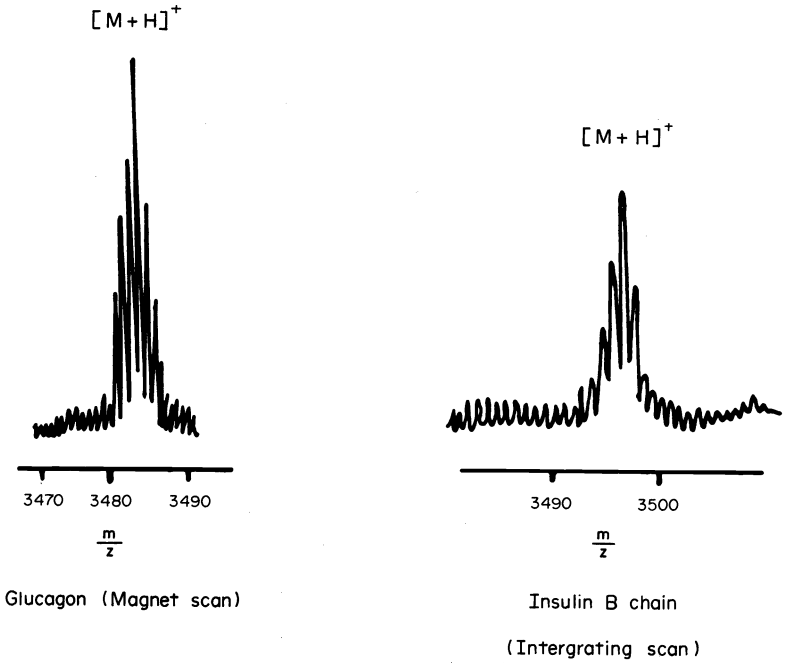


Fig. 11. Molecular ion regions of the FAB spectra of Glucagon and the Insulin B chain. Quasimolecular ions obtained in each case.

The alkali halides are useful references since they yield cluster ions at regular intervals. The insulin and glucagon spectra shown in Fig. 11 were mass marked using the potassium iodide cluster ions at m/z 3356.3 ($K_{21}I_{20}$) and 3522.2 ($K_{22}I_{21}$). The alkali halides can be used to calibrate spectra up to very high mass e.g. we have obtained good spectra of cesium iodide up to m/z 12,344 ($Cs_{48}I_{47}$) using the high field ZAB operated at 2 KV accelerating voltage (12).

Cs_2I	392.71	$Cs_{27}I_{26}$	6887.96
Cs_3I_2	652.52	$Cs_{28}I_{27}$	7147.77
Cs_4I_3	912.33	$Cs_{29}I_{28}$	7407.58
Cs_5I_4	1172.14	$Cs_{30}I_{29}$	7667.39
Cs_6I_5	1431.95	$Cs_{31}I_{30}$	7927.20
Cs_7I_6	1691.76	$Cs_{32}I_{31}$	8187.01
Cs_8I_7	1951.57	$Cs_{33}I_{32}$	8446.82
Cs_9I_8	2211.38	$Cs_{34}I_{33}$	8706.63
$Cs_{10}I_9$	2471.19	$Cs_{35}I_{34}$	8966.44
$Cs_{11}I_{10}$	2731.01	$Cs_{36}I_{35}$	9226.25
$Cs_{12}I_{11}$	2990.81	$Cs_{37}I_{36}$	9486.06
$Cs_{13}I_{12}$	3250.62	$Cs_{38}I_{37}$	9745.87
$Cs_{14}I_{13}$	3510.43	$Cs_{39}I_{38}$	10005.68
$Cs_{15}I_{14}$	3770.24	$Cs_{40}I_{39}$	10265.48
$Cs_{16}I_{15}$	4030.05	$Cs_{41}I_{40}$	10525.29
$Cs_{17}I_{16}$	4289.86	$Cs_{42}I_{41}$	10785.10
$Cs_{18}I_{17}$	4549.67	$Cs_{43}I_{42}$	11044.91
$Cs_{19}I_{18}$	4809.48	$Cs_{44}I_{43}$	11304.72
$Cs_{20}I_{19}$	5069.29	$Cs_{45}I_{44}$	11564.53
$Cs_{21}I_{20}$	5329.10	$Cs_{46}I_{45}$	11824.34
$Cs_{22}I_{21}$	5588.91	$Cs_{47}I_{46}$	12084.15
$Cs_{23}I_{22}$	5848.72	$Cs_{48}I_{47}$	12343.96
$Cs_{24}I_{23}$	6108.53	$Cs_{49}I_{48}$	12603.77
$Cs_{25}I_{24}$	6368.34	$Cs_{50}I_{49}$	12863.58
$Cs_{26}I_{25}$	6628.15	$Cs_{51}I_{50}$	13123.39

Fig. 12. Masses of charged Cesium Iodide clusters $(CsI)_nCs^+$. Discrete ions are observed which are useful as high mass reference markers, independent of resolution.

It is perhaps in the glycopeptide and glycolipid area that our high field magnet MS facilities will continue to have the greatest impact since a large number of biologically important members of these classes of compounds have molecular weights in the 2000-5000 mass range. We have used FD MS to characterise underivatised methylmannose polymers up to 2500 molecular weight (15) and have obtained FD spectra up to m/z 3400 from human milk oligosaccharides analysed as their alditol acetates (16). A typical FD spectrum of the latter class of compound is shown in Fig.13. The sodium cationised molecular ion at m/z 3391 corresponds to a reduced oligosaccharide containing four fucoses and two lactosamines in addition to the basic lacto-N-tetraose unit. The signal at m/z 3349 is assigned to the cationised molecular ion of the mono de-acetylated species.

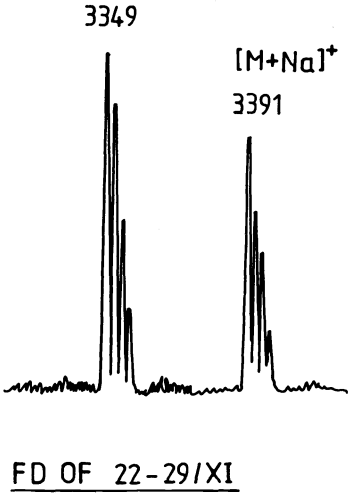


Fig. 13. Molecular ion region of the FD spectrum of a human milk oligosaccharide analysed as the alditol acetate. The molecular weight corresponds to lacto-N-tetraose + lactosamine$_2$+ fucose$_4$.

Our FD and FAB studies have played an invaluable role in the complete structural characterisation of a number of glycosphingolipids. For example, the molecular weights of the I-active ceramide decasaccharide and ceramide pentadecasaccharide from rabbit erythrocyte membranes were established by FD MS of the permethylated derivative. Molecular ions were obtained in the mass ranges m/z 2850 to 2950 and 3950 to 4020 for the deca- and pentadecasaccharides respectively (17,18). FAB MS of permethylated glycosphingolipids of related structures have yielded both molecular weight and structural information, the latter data providing a useful complement to that obtainable from EI analysis (19).

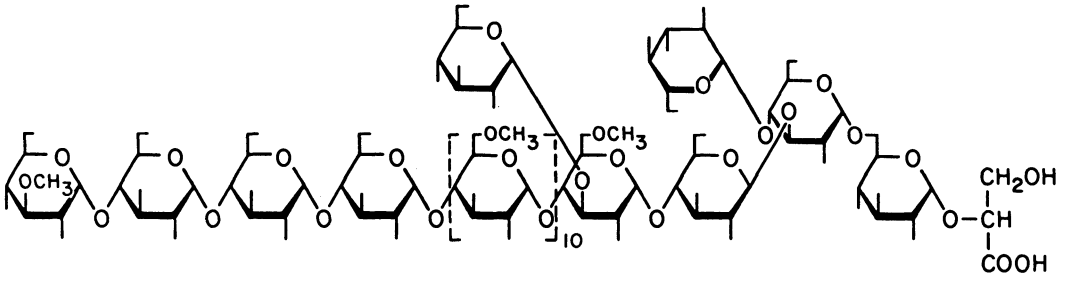


Fig. 14. Structure of a methyl glucose oligosaccharide, one of the many high mass biological polymers under investigation by MS in this laboratory.

Using our high mass facilities we have obtained molecular weight data on a 6-O-methylated glucose polymer (MGP) isolated from a mycobacterial lipopolysaccharide. Our FD and FAB studies have led to the structure revision of this molecule from an 18 sugar unit polymer of molecular weight 3176 to a 20 sugar unit polymer of molecular weight 3514 whose structure is given in Fig. 14. FD MS at 5.5 kV on the permethyl derivative of MGP yielded signals in the region m/z 4200-4250, the complexity resulting from incomplete methylation and cationisation. FAB MS at 6 kV of underivatised MGP afforded the sodium and potassium cationised molecular ions at m/z 3538 and 3554 respectively, unambiguously defining the number of methylated residues present in the polymer (20).

CONCLUSION

In conclusion many of today's most pressing problems in the area of biopolymer mass spectrometry have until recently been intractable because of the difficulties associated with the analysis of microgram or even nanogram quantities of high molecular weight polar substances. In FD analysis the problems are further exacerbated by inherently weak ion beams. The development of the high field magnet (8), together with Fast Atom Bombardment (21) has allowed the application of soft ionisation techniques to the 'next generation' of high mass problems.

Mass spectrometry has already played a major role in the structure elucidation of polymers of biological interest such as neuropeptides, peptidolipids, glycopeptides, antibiotics and proteins. As the technique develops both in terms of mass range and sensitivity MS promises to allow the solution of many important problems only just beyond the scope of present technology. For biological mass spectrometry the next decade should be as exciting and fruitful as the last.

REFERENCES

1. H.R. Morris, Phil. Trans. R. Soc. Lond. A293, 39-51 (1979).
2. H.R. Morris, Nature 286, 447-452 (1980).
3. J.E. Fitton, A. Dell and W.V. Shaw, FEBS Letters 115, 209-212 (1980).
4. R.P. McDonagh, J. McDonagh, T.E. Petersen, H.C. Thorgersen, K. Skorstengaard, L Sottrup-Jensen, S. Magnusson, A. Dell and H.R. Morris, FEBS Letters 127, 174-178 (1981).
5. G.P. Vinson, B.J. Whitehouse, A. Dell, A.T. Etienne and H.R. Morris, Nature 284, 464-467 (1980).
6. D.F. Hunt, G.C. Stafford, F.W. Crow and J.W. Russell, Analyt. Chem. 48, 2098-2105 (1976)
7. G.W. Taylor, H.R. Morris, T.E. Petersen and S. Magnusson, Adv. Mass Spectrom. 8, 1090-1096 (1980).
8. H.R. Morris, A. Dell and R. McDowell, Biomed. Mass Spectrom. 8, 463-473 (1981).
9. H.R. Schulten, J. Mass Spectrom and Ion Physics 32 97-283 (1979).
10. A. Dell, H.R. Morris, S.M. Hecht and M.D. Levin, Biochem. Biophys. Res.Comm. 97, 987-994 (1980).
11. A. Dell, H.R. Morris, M.D. Levin and S.M. Hecht, Biochem. Biophys. Res. Comm. 102, 730-738 (1981).
12. J. Walker, Personal Communication.
13. A. Dell, A.T. Etienne, H.R. Morris, G.P. Vinson, B.J. Whitehouse, M. Barber, R.S. Bordoli, R.D. Sedgwick and A. Tyler, Submitted for publication.
14. H.R. Morris, M. Panico, M. Barber, R.S. Bordoli and A. Tyler, Biochem. Biophys. Res. Comm. 101, 623-631 (1981).
15. M. Linscheid, J.D'Angona, A.L. Burlingame, A. Dell and C.E. Ballou, Proc. Natl. Acad. Sci. U.S.A. 78, 1471-1475 (1981).
16. H. Egge, J. Dabrowski, P. Hanfland, H.V. Nicola, U. Dabrowski and A. Dell, Proc. VIth Int. Symp. Glycoconjugates, Tokyo, pp. 318-319 (1981) Japan Scientific Societies Press.
17. P. Hanfland, H. Egge, U. Dabrowski, S. Kuhn, D. Roelcke and J. Dabrowski, Biochemistry 20, 5310-5319 (1981).
18. P. Hanfland, J, Dabrowski, A. Dell, U. Dabrowski, S. Kuhn and H. Egge, Proc. VIth Int. Symp. Glycoconjugates Tokyo, pp. 98-99 (1981) Japan Scientific Societies Press.
19. A. Dell and H. Egge, Unpublished results.
20. L.S. Forsberg, A. Dell, D.J. Walton and C.E. Ballou, Submitted for publication.
21. M. Barber, R.S. Bordoli, D.R. Sedgwick and A.N. Tyler, Chem. Commun. 325-327 (1981).