Pure & Appl. Chem., Vol. 56, No. 7, pp. 807-819, 1984.
Printed in Great Britain.

0033-4545/84 $3.00+0.00
Pergamon Press Ltd.

STRUCTURAL ANALYSIS OF BLOOD-GROUP ABH, I, i, LEWIS AND RELATED GLYCOSPHINGOLIPIDS. APPLICATION OF FAB MASS SPECTROMETRY AND HIGH RESOLUTION PROTON NMR

Heinz Egge,* Janusz Dabrowski** and Peter Hanfland***

*Institut für Physiologische Chemie der Universität, Nussallee 11, D-5300 Bonn 1, **Max-Planck-Institut für Medizinische Forschnung, Jahnstrasse 29, D-6900 Heidelberg, ***Institut für Experimentelle Hämatologie und Bluttransfusionswesen der Universität, Sigmund-Freud-Strasse, D-5300 Bonn, Federal Republic of Germany

Abstract - The potential of fast atom bombardment (FAB) mass spectrometry and of high resolution proton NMR spectroscopy, in the structure elucidation of complex glycosphingolipids (GSL), was assessed with the aid of unmodified, peracetylated and permethylated derivatives. The results of negative ion FAB MS on native GSL and of positive FAB MS on their permethylated derivatives proved to be highly complementary to those obtained by the use of different NMR techniques. Examples of FAB spectra, of spin-decoupled spectra, of NOE (Nuclear Overhauser effect) difference spectra, of J-resolved two-dimensional spectroscopy, as well as of the COSY (chemical shift correlated) or the SECSY (spin-echo correlated) variant of the two-dimensional spectroscopy, measured at 360 or 500 MHz, demonstrate that with a minimal amount of substance complete information can be gathered on the molecular weight, ceramide moiety, the type and number of sugar constituents, their anomeric configuration, sequence, sites of linkages and branching patterns.

INTRODUCTION

Glycosphingolipids, by definition glycosides of the ceramides, are integral components of the plasma membrane of eucariotic cells. More than 95% of the cellular GSL are anchored asymmetrically in the outer layer of the cell membranes. Although their function is still poorly understood on a molecular level, it is generally accepted that GSL together with other glycoconjugates of the plasma membrane play an important role in the transmembrane communication of neighbouring cells, in the cellular differentiation and malignant degeneration or oncogenesis. They function as cellular antigens, as genetic markers and they can be parts of cellular receptor structures. A real understanding of these normal physiological and pathological processes requires a complete knowledge of the structures involved. Compared to the progress achieved in the sequence determination of other biopolymers, methods used for the sequencing and structure determination of complex glycoconjugates still reside in the dark ages. A great discrepancy exists between the increasing demand for structural information and the vanishingly small amounts of material available, the enormous complexity of the involved structures and the extreme demands of specificity combined with some of the most difficult organic chemistry. During the last years the growing perception of the importance of this field has stimulated the development and the application of new and more powerful methods of isolation, purification and structural elucidation. The complete analysis of GSL implies the determination of

- sugar components and molecular mass
- composition of the ceramide residues
- the anomeric configuration of the sugars
- the conformation of the sugar rings
- the sequence of sugar constituents
- the pattern of branching
- the sites of glycosidic linkages
- secondary structure or spatial orientation.

METHODS

A variety of chemical, biological (immunological), biochemical and physical methods are now available by which different aspects of the structure of glycosphingolipids can be elucidated; hence, there is usually no simple answer to the question of how to reach an unequivocal solution of a given structural problem by investing a possible minimum of means and time. Although physical methods require sophisticated and expensive instrumentation, they are often preferable, being less time-consuming. In the experience of our and several other teams, a combination of mass spectrometry and proton nuclear magnetic resonance proved particularly useful. Both of these methods experienced an unprecedented development during the last decade and are still in a state of rapid instrumental and methodological progress.

Mass spectrometry

Mass spectrometry using electron impact (EI) or chemical ionization (CI) has been successfully employed by several laboratories in the sequence analysis of GSL (Ref. 1,2). As derivatives, permethylated, perme-

thylated and reduced, and in the case of gangliosides, permethylated, reduced and silylated samples were analysed. These results are well documented. Thus, GSL containing up to 12 sugar components could be analysed (Ref. 3). From the mass spectra conclusive evidence can be obtained on the fatty acid and sphingosine of the ceramide residue, on the sugar sequence, on the presence of branching points and also in part on the sites of linkages. Due to the presence of rather intense ions produced by cleavage of the C(2)-C(3) bond of the sphingosine with retention of the fatty acid and the complete carbohydrate chain, it is also possible to analyse mixtures of GSL using selected ion monitoring and programmed increase of probe temperature (Ref. 4). Molecular ions, if any, are normally of very low intensity. The application of these MS techniques is limited by the volatility and thermal stability of the samples. Even with "close probe" techniques and the use of permethylated and reduced derivatives it will be difficult to analyse compounds with molecular weights in excess of 3000 daltons.

Fast atom bombardment. The introduction of FAB has opened new dimensions to the MS analysis of polar compounds (Ref. 5-8). We have recently investigated the potentials of FAB MS in the analysis of permethylated derivatives of GSL and other glycoconjugates. Normally very intense pseudomolecular ions M+H or M+Na are observed. The intensity of these ions can be greatly enhanced by the addition of salts. Since no vaporization process is involved in the formation of these pseudomolecular ions, the mass range is only limited by the specification of the instrument, i.e. strength of the magnetic field and sensitivity of detectors in the high mass range. The most advanced instruments will offer a mass range of 12000 dalton at 10 kV acceleration voltage. In our hands molecular ions were observed with GSL containing up to 25 sugar components (Ref. 9). Besides the molecular ions, a highly specific cleavage occurs at the glycosidic bond between the sugar chain and the ceramide residues, giving rise to intense ions that represent the whole ceramide moiety. From the pseudomolecular ions and the ceramide residue, the overall carbohydrate composition concerning the number of hexoses, N-acetylhexosamines, deoxyhexoses, sialic acids and uronic acids can exactly be calculated. Each of the above components contributes specific mass increments that allow an unequivocal calculation for every possible composition. Furthermore, specific cleavage occurs starting with the terminal carbohydrate components. These can be recognized, as in EI MS, by specific ion pairs such as m/z 219 and 187 for hexose or m/z 260 and 228 for hexosamine, etc. In addition, high intensity sequence ions are produced by the selective cleavage of the glycosidic bonds of the hexosamine residues in the sugar chain. In this way, series of sequence ions are produced, that allow the determination of branching patterns and, e.g., in the lactoseries the location of repetitive lactosamine units. These sequence ions are normally accompanied by daughter ions that are specific for type 1 or type 2 chains, thus offering a means of discrimination between 1-4 and 1-3 linked hexosamines.

Negative ion FAB MS. With the aid of FAB native underivatized GSL are also amenable to the MS analysis. Both positive and negative ions can be recorded. In the presence of cations, however, multiple exchange of protons is observed, thus giving rise to multiple pseudomolecular and fragment ions of relatively low intensity. In contrast, very intense pseudomolecular ions M-1 are observed when recording negative ions. Besides these pseudomolecular ions, fragment ions are formed by cleavage of the glycosidic bonds with the negative charge remaining on the oxygen. Thus, two series of sequence ions are observed starting from both ends of the molecule, that indicate the size of the ceramide residue and the sequence of the sugar components.
In summary, negative ion FAB MS gives conclusive information on the molecular weight, the aglycon, the type, number and sequence of the sugar residues, and the branching pattern of the carbohydrate chain.

Mass spectrometry was performed on a ZAB 1F reversed geometry double focussing instrument equipped with a high field magnet with a mass range of 3500 amu at 8 kV acceleration voltage (VG Analytical, Manchester, U.K.), as described in (Ref. 8).
The isolation and purification of the GSL followed the procedures described elsewhere (Ref. 10, 11).

^{1}H NMR spectroscopy
The outstanding new features of NMR spectroscopy are the extensive use of high-field instruments operating at frequencies of up to 500 MHz, and the utilization of the capabilities of the computer control of experiments. At higher field strength both the spectral dispersion and sensitivity are improved. Computer controlled spectrometers enable the routine application of pulsed Fourier-transform spectroscopy and its important variants, difference spectroscopy and two-dimensional spectroscopy. Thanks to these techniques, the overlapping signals - or at least a part of them - can be assigned so that structure determinations can be based on larger sets of data.

Spin-decoupled difference spectra (SDDS) are obtained by subtracting a control (unperturbed, off-resonance irradiated) spectrum from the spectrum obtained under decoupling conditions, i.e., from one that is on-resonance irradiated during the acquisition time (Ref. 12,13). As seen from the example given in Fig. 1b, the signal of the decoupled proton is neatly isolated from the others in the bulk. Moreover, the whole coupling pattern of the decoupled proton become visible, hence all through-bond connectivities of such protons can be identified and the relevant coupling constants measured with an accuracy sufficient for practical purposes and even for spectrum simulation-iteration, if desired. By applying this procedure consecutively, all proton signals can in principle be assigned. Problems arise when both the irradiated and the observed resonance lie in a crowded spectral region. In such cases two dimensional methods are preferable. On the other hand, none of exactly overlapping multiplet resonances can be seen in J-resolved two-dimensional spectra, whereas each of them can readily be observed by SDDS, if coupled to at least one resonance located elsewhere.

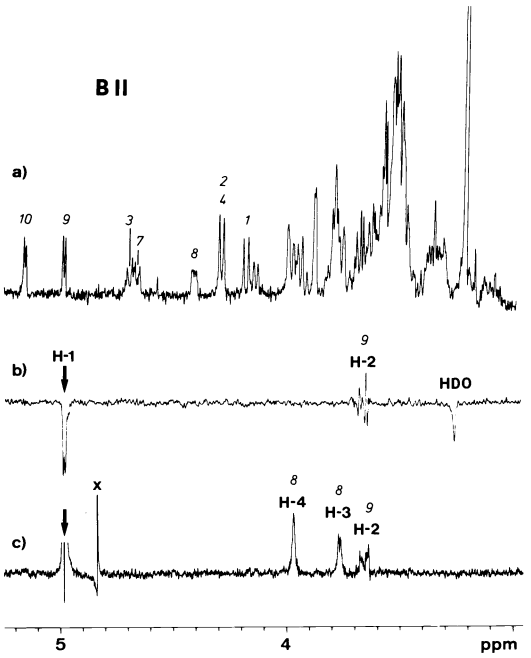

Fig.1 The 3.0-5.2 ppm region of the 360-MHz ^{1}H NMR spectra of the blood-group B-active ceramide octasaccharide, Galα1-3Gal(2← 1αFuc)β1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-1Cer, (B-II) in DMSO-d_6 containing 2% D_2O at 333K. (a) The resolution-enhanced spectrum; (b) SDDS spectrum;Gal-9 H-1 irradiated (indicated by the arrow), H-2 observed; (c) NOE difference spectrum: Gal-9 H-1 irradiated, intraresidue NOE of Gal-9 H-2 and interresidue NOE of Gal-8 H-3 and H-4 observed; x: off-resonance irradiation in the control spectrum.

NOE (nuclear Overhauser effect) difference spectra (Ref. 14-20), which are suitable for ascertaining spatial proximities of nuclei, are produced in a similar manner to SDDS, but irradiation is applied before acquisition. Owing to the low power of this irradiation, NOE experiments are more selective. Some of the through-space connectivities established by NOE will be the same as the through-bond ones obtained by SDDS, viz., those for geminal protons, for vicinal diequatorial or axial/equatorial protons, possibly for H-5/H-6, and for several ceramide protons. Other NOE difference signals will identify new combinations of protons, e.g., synaxial H-1/H-3/H-5 and H-2/H-4, and possibly H-4/H-6. Thus, NOE spectra contribute much to the discovery of hidden resonances and under favourable circumstances may enable one to complete the identification of all spin systems. Even more important than these intraresidue NOEs are the interresidue ones between anomeric and aglyconic protons, because they serve as a basis for the determination of sequences and sites of glycosidic linkage (cf. Fig. 1c).

In the two dimensional (2D) spectra, the effects of chemical shifts and scalar, or dipolar coupling are separated. The contour plot presentation of a homonuclear J-resolved 2D spectrum (Ref. 21, 22) is shown in Fig. 2c. The centres of all resonances are spred along the δ_2 axis (i.e., the chemical shift axis) and can be seen as singlets in the projection onto this axis (Fig. 2b). The cross sections through these centres, yield the multiplets of individual resonances, here in the form of contours (they can, of course, be also depicted in the conventional form). Thus, multiplets that were intermingled in the 1D spectrum (Fig. 2a) are now separated. This technique is only applicable to first-order spectra hence the measurements should be carried out at possibly strong fields. At the same time, higher-order effects are practically meaningless for the J-mediated, shift correlated spectra. The latter indicate scalar coupling between nuclei, without providing precise values of the coupling constants, however. A chemical shift correlated spectrum ("COSY" variant: Ref. 23,24) is presented in Fig. 2d. The contours corresponding to resonances of the 1D spectrum are aligned along the diagonal, in the coordinate system δ_1,δ_2. If two nuclei A and B, resonating at δ^A and δ^B, are coupled to each other, they exhibit off-diagonal contours located in the corners of the square defined by δ^A_1, δ^B_1, δ^A_2 and δ^B_2. Coupled nuclei can therefore be identified by finding such squares on the contour diagram. If any of these nuclei is further coupled to others, additional contours occur at its chemical shift and this simple correlation procedure can be continued until the entire spin system of the given sugar residue is traced.

Thus, a single 2D correlated spectrum can be substituted for a series of consecutive decoupling experiments. The "SECSY" variant of the shift correlated spectroscopy differs slightly in experimental details and in the coordinate system of the presentation, but the connectivity information obtained is the same. Similarly, 2D NOE spectra can be obtained by correlating the chemical shifts of closely located nuclei via the dipolar coupling (i.e., relaxation) between them. Here, too, long series of consecutive 1D NOE experiments can be replaced by one 2D spectrum. Perhaps even a more important advantage of the correlated spectra is that no irradiation by a second radio-frequency field is needed; hence the connectivities between coupled nuclei (scalar or dipolar) can be established even if their signals are heavily overlapped.

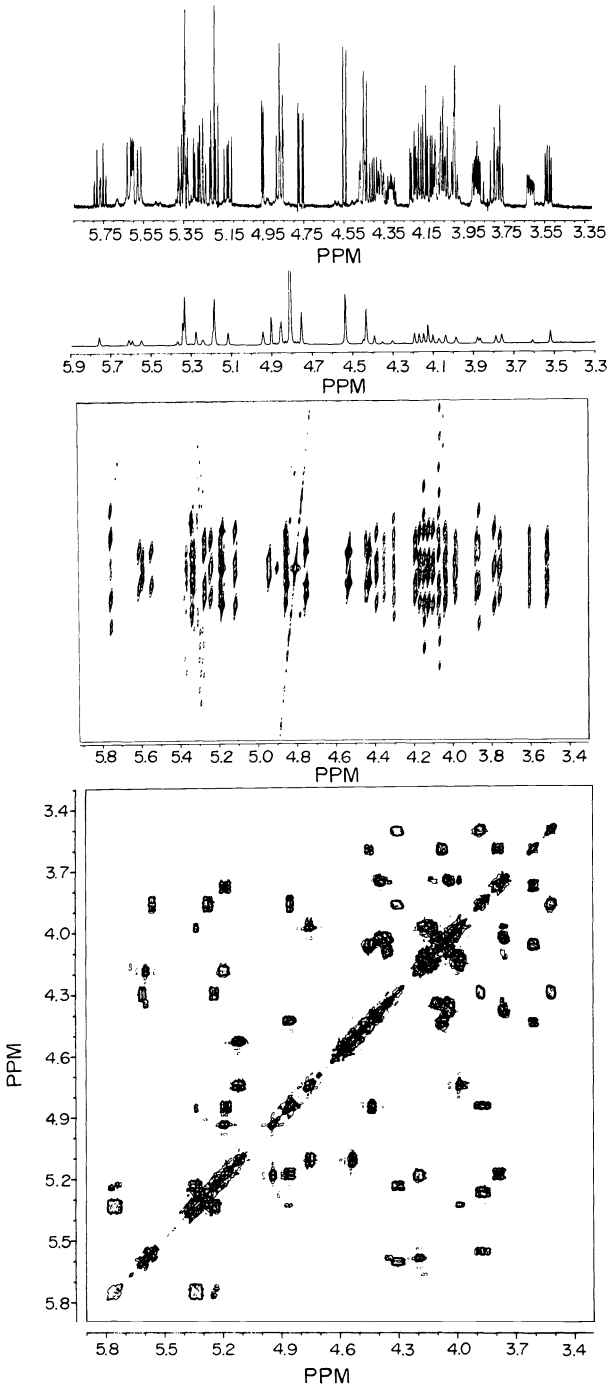

Fig.2 The 3.3-5.9 ppm region of the one- and two-dimensional (2D) 500-MHz ^{1}H NMR spectra of the peracetylated globotetraosylceramide, GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-1Cer, in deuterochloroform at 303 K;

(a) resolution-enhanced 1D spectrum;

(b) projektion of the 2D J-resolved spectrum; all multiplets are quasi-decoupled; the strong peak at 4.81 is an artifact;

(c) contour plot of the J-resolved spectum described in (b); the multiplet structure of most resonances is clearly visible; each multiplet can be plotted individually in the conventional form (not shown here);

(d) contour plot of the J-mediated, chemical shift correlated 2D spectrum (COSY); by convention, the horizontal and the vertical ppm scales are assigned δ_2 and δ_1, respectively (see text).

Elucidation of the elements of the primary structure of GSL by ^{1}H NMR

Ceramide. All signals of the carbon-bound ceramide protons of galactosyl- and glucosyl ceramide were assigned by decoupling, partly in difference mode (Ref. 27,28), and found little changed for several neutral GSL (Ref. 13,20,29-31) and gangliosides (Ref. 29-31).

The number of the CH_2 groups in both chains of the ceramide part can be determined by integration in the 1.0-2.1 ppm area. Similarly, the signals of the cis and trans olefinic protons occurring around 5.5 ppm can be integrated to specify the types and amount of unsaturation. A fatty acid (FA) is characterized by the (-C=O)-CH_2-triplet at ca. 2.03 ppm; the lack of this triplet and a methine (-C=O)-CHOD-signal at ca. 3.82 ppm indicate that sphingosine is acylated by a hydroxy fatty acid (HFA). In neutral GSL, FA can also be distinguished from HFA (Ref. 31) by the sphingosine NH signals at 7.45-7.52 and 7.07-7.17 ppm, respectively (at 25^{0}C).

Type and number of monosaccharide residues, and ring conformation. The anomeric proton signals of the deuterium-exchanged native GSL occur in the otherwise unpopulated region between 4.1 and 4.9 ppm, hence the number of sugar residues can be determined by integrating these signals. The number of amino sugar units is given by the integrals of the methyl singlet signals of the acetamido groups between 1.80 and 1.88 ppm (Ref. 13,20,29-31) and fucose units can be counted by the integral values of their methyl doublets between 1.0 and 1.1 ppm (Ref. 19,32).

The unambiguous ab initio identification of the constituent sugar residues, including their ring conformation, can be performed when all of their ring proton signals and the respective coupling constants have been found. Whether this was done by 1D methods, as in (Ref. 16,19), or 2D ones is meaningless but a combination of J-resolved and J-mediated, shift-correlated 2D spectroscopy proved particularly effective. A synthetic disaccharide derivative (Ref. 33), a percetylated branched hexasaccharide from human milk (Ref. 34), a synthetic branched heptasaccharide related to glycoproteins (Ref. 35) and a branched native ceramide pentadecasaccharide (Ref. 36) have been analysed in this way. Under favourable circumstances the coupling constants can be read with a tolerable accuracy from the shift-correlated spectra (Ref. 29,30) but, due to contour overlap, this can hardly be expected to be extendable substantially beyond the level of a pentasaccharide, which was the largest oligosaccharide unit analyzed in this thorough study of gangliosides (Ref. 30).
It has been pointed out in the work on ceramide pentadecasaccharide (Ref. 36) that, in spite of the enormously increased resolution, furtitious overlap of resonances belonging to different spin systems may still take place, particularly when the molecule contains identical or little differentiated repeating fragments. Multiple overlappings were encountered already on the heptasaccharide level (Ref. 35). It seems, that at the present state of NMR instrumentation the possibility to fully analyze the spectra of oligosaccharides will be limited to molecules of medium size. However, full assignments of NMR parameters are not being made for their own sake and, fortunately, identification of the particular sugar residues can often be unequivocal even though merely based on a partial analysis. This is especially true in the case of GSL, for which the NMR data on the few hexoses known as their constituents (Glc, Gal, GlcNAc, GalNAc, Fuc) were obtained for a great variety of structural combinations (Ref. 13,19,20, 27-32,36,37).

Anomeric configuration. The determination of the anomeric configuration of aldohexopyranoses is based on the extensively studied Karplus-like relationship between the vicinal coupling constant, $^{3}J_{1,2}$, and the dihedral angle φ in the H-C^{1}-C^{2}-H fragment. For the diaxial position of H-1 and H-2 ($\varphi \approx 180^{o}$), this coupling constant is of the order of 6-10 Hz and for their axial/equatorial or diequatorial orientation ($\varphi \approx 60^{o}$), its value is in the range of 1-4 Hz. From here it follows that in the general case of a completely unknown sugar several combinations referring to two possible configurations at C^{2} and two possible chair conformations of the ring would have to be considered. The situation is much simpler, however, if constituents of GSL are concerned. Since the only represented $^{4}C_{1}$ conformations of D-Glc, D-Gal, D-GlcNAc and D-GalNAc, and $^{1}C_{4}$ of L-Fuc imply the axial orientation of H-2 for all these sugar residues, small coupling constants point unmistakably to α anomeric configuration and large ones identify β anomers.

Site of glycosidic linkage and sequence. Glycosylated sites can be identified by glycosylation-induced shifts observed for native or peracetylated oligosaccharides and, at the ab initio level of analysis, by interresidue NOE and interresidue scalar coupling via the glycosidic bond. Substitution at a saccharide ring by another sugar unit induces chemical shift changes ("glycosylation shifts", $\Delta\delta$) for both of them, those for the protons of the glycosylated residue being diagnostic of the glycosylation site. The largest $\Delta\delta$ result for the proton at the linkage site and vicinal to it, other resonances being markedly less changed. Although one or both of the vicinal protons may be even more affected than that at the glycosylation site, the latter can nevertheless be unambiguously determined as corresponding to the signal of the middle proton of these three exhibiting large $\Delta\delta$. Except for special cases of crowding originating with the new sugar residue entering a site vicinal to one already glycosylated (Ref. 16,19,20,30,32), glycosylation shifts regularly follow this pattern and are toward lower field ($\Delta\delta$ is positive). Therefore, glycosidic linkage sites of given sugar residues, of a strain-free GSL, can be determined if chemical shifts for all protons of these residues located terminally in corresponding precursors of this GSL are known.

Glycosylation shifts for peracetylated oligosaccharides are even more characteristic and useful (Ref. 34,38,39). Glycosylating units replace the strongly deshielding acetyl groups hence the shifts are toward higher field, their magnitude being of the order of 1 ppm. Since neighbouring protons are little affected, the discovery of sites of glycosidic linkage is straightforward. Spectra of the peracetylated globotetraosylceramide are shown in Fig. 2.

The NOE signal of the aglyconic proton at the glycosidic linkage site can be obtained upon preirradiation of the anomeric proton of the subsequent sugar residue. This scheme was first used for estimation of the angles between sugar rings (Ref. 15,16) and soon applied to the determination of unknown sequences and sites of glycosidic linkage in native GSL (Ref. 17-20). The identification of the aglyconic proton at the linkage site is unequivocal, except when there is an equatorial proton vicinal to it, as H-4 in 3-glycosylated galactose residues. Such protons exhibit NOEs of approximately the same magnitude, hence they must be located at comparable distance from the anomeric proton across the glycosidic linkage. Although the sequence in this disaccharide fragment is thus unambiguously established, the question of the linkage to site 3 or 4 remains open. The answer to this question can be found by consulting the data on glycosylation shifts: according to what was said above, protons that show practically the same chemical shifts as in the nonglycosylated precursors, cannot be those neighbouring the glycosylation site. Hence, unal-

tered H-2 resonance will indicate glycosylation at C-4 (Ref. 17,20,40).

Another ab initio type approach to resolving sequence and linkage problems puts to use the four-bond scalar coupling between the anomeric and the aglyconic proton. Although this coupling is usually so small as not to be observable even in resolution-enhanced spectra, it can nonetheless be detected (but not measured), being sufficient to cause a magnetization transfer between the protons mentioned, in the so-called delayed COSY experiment (Ref. 24,41). Unfortunately, five-bond interresidue couplings may also produce connectivity contours of comparable intensitiy, hence ambiguities as to the glycosidic linkage site, similar to those described above for NOE, will also have to be taken into account. Although, as outlined in (Ref. 41), the problem can possibly be attacked by determining the relative signs of the coupling constants, this will certainly not be an easy way to go, particularly for large molecules, with the host of long-range inter-and intraresidue couplings expected for them.

All the methods discussed above are, of course, also applicable to the determination of sites of bifurcation of the oligosaccharide chain. However, knowledge of the particular branching points is insufficient for the unambiguous elucidation of the gross branching pattern. This special aspect of sequence will be dealt with in the next section.

The secondary, three-dimensional structure and related problems are beyond the scope of this article.

RESULTS AND DISCUSSION

Recent experiments show that 25% of the ABH blood-group antigens of human erythrocytes are represented by glycosphingolipids. The majority of these, about 4/5, have complex structures that comprise 20-60 sugar components per molecule (Ref. 42,43). A part of these higher GSL could be isolated from the membranes of human B erythrocytes by extraction with butanol/buffer systems followed by various fractionation steps as described elsewhere (Ref. 44). In a final step the peracetylated GSL were purified by preparative h.p.t.l.c. (Ref. 45). The highly purified GSL were characterized by high resolution NMR of the native compounds and by FAB MS of the permethylated derivatives.

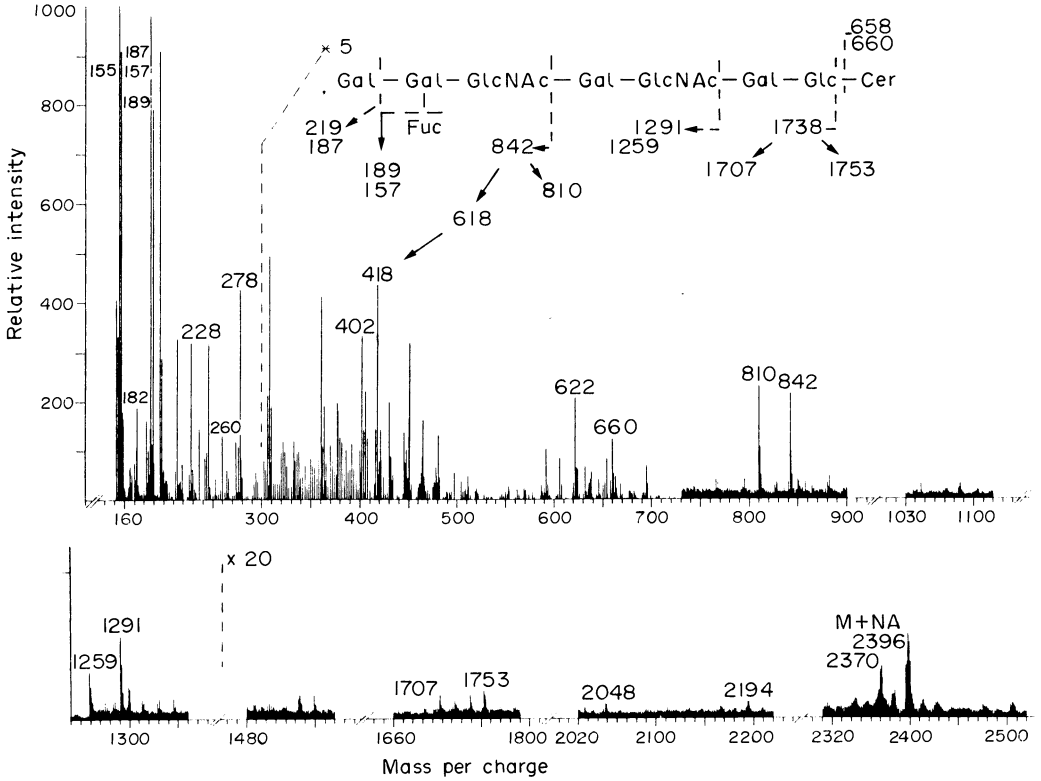

Fig. 3. Positive ion FAB MS spectrum of permethylated blood-group B-active ceramide octasaccharide (B-II).

As an example, the FAB spectrum of the blood-group B active ceramide octasaccharide is shown in Fig. 3 together with the scheme of fragmentation. The complete structural elucidation of this glycolipid, B-II, has already been published (Ref. 10). The major pseudomolecular ion M+Na is m/z 2396 and the ceramide residue is represented by m/z 660 and 658. From these data it can be calculated that the glycolipid contains one deoxyhexose, two hexosamines and five hexoses. The sequence of the sugar residues is represented by several ions. The two ion pairs, m/z 219 and 187, and m/z 189 and 157, clearly indicate the presence of two terminal sugar residues and a branched structure. The terminal tetrasaccharide ion m/z 848 is accompanied by a set of daughter ions that are produced by the alternative or sequential elimina-

tion of methanol, hexose or deoxyhexose, as indicated in the scheme. Especially the elimination of methanol leading to m/z 810 is a strong indication that the position 3 of the GlcNAc residue is not substituted. No pentasaccharide ion is visible but instead a hexasaccharide ion at m/z 1291, which is accompanied by m/z 1259. Both the tetra- and the hexasaccharide ions are accompanied by ions that are 8 amu higher at m/z 856 and 1299, the latter being formed by exchange of one methyl group for one sodium. By addition of potassium ions these ions are shifted to 24 amu higher values. The whole carbohydrate chain, finally, is represented by a series of ions of comparatively low intensity that are produced by elimination of the ceramide residue followed by rearrangement at m/z 1707, 1738 and 1753. A number of ions in the lower mass range, like m/z 609, 503, 451 etc., are derived from the thioglycerol matrix. Thus the GSL B-II can be characterized as a ceramide octasaccharide with a subterminal branching point and two lactosamine units joined to each other.

Further details of the structure of B-II were established by NMR. Anomeric configuration and the number of sugar residues were determined by measuring the coupling constants and integral intensities of the H-1 doublets. The H-1 signal of Gal-8 (Fig. 1) was a higher-order multiplet due to strong coupling between H-2 and H-3 (confirmed by spectrum simulation). The connectivities between the H-1 resonances and those hidden in the unresolved bulk were found by SDDS and/or intraresidue NOE. The sequence Galα1-3(Fucα1-2)Galβ1- was established by interresidue NOE. Thus, irradiation of Fuc H-1 at 5.16 ppm produced two NOE difference signals: the intraunit one of Fuc H-2 at 3.55 ppm, already known from SDDS; and the interunit one of Gal-8 H-2 at 3.77 ppm, whose connectivity with the Gal-8 H-1 resonance at 4.39 ppm was unambiguously established by SDDS. On the other hand, NOEs obtained upon irradiation of H-1 of Gal-9 identified the Galα1-3Gal β linkage, as shown in Fig. 1c and described above in the "site of glycosidic linkage and sequence" section.

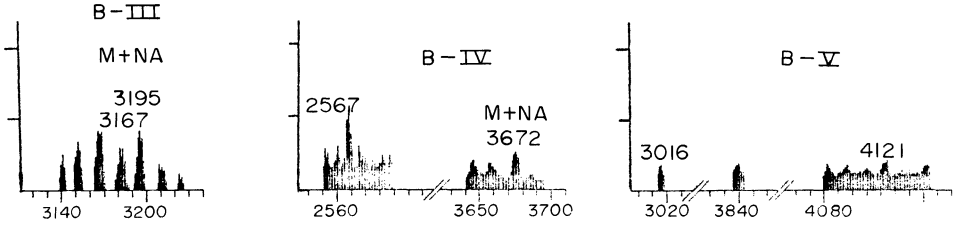

Fig. 4. Molecular ion region observed in the FAB spectra of the permethylated blood-group B-active GSL B-III, B-IV and B-V.

TABLE 1. Anomeric proton chemical shifts (ppm from Me_4Si at 333 K in DMSO-d_6/D_2O, 98:2) of the blood-group B-active GSL, B-III and B-IV, and their nonfucosylated analogue, BI_{rab}-1

BI_{rab}-1[a]

9 8 7 4 3 2 1

Galα1-3Galβ1 — 4GlcNAcβ1 — 3Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-1R
(6)

Galα1-3Galβ1 — 4GlcNAcβ1 — 6Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-1R

4.84 4.30 4.42 4.67 4.30 4.67 4.27 4.17

B-III

9 8 10 7 4 3 2 1

Galα1-3Gal(2←1αFuc)β1 — 4GlcNAcβ1 — 3Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-1R
(6)

Galα1-3Gal(2←1αFuc)β1-4GlcNAcβ1 — 6Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-1R

9' 8' 10' 7'

4.98 4.39 5.16 4.36 4.64 4.30 4.68 4.27 4.17

B-IV

9 8 10 7 6 5 4 3 2 1

Galα1-3Gal(2←1αFuc)β1 — 4GlcNAcβ1-3Galβ1-4GlcNAcβ1 — 3Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-1R
(6)

Galα1-3Gal(2←1αFuc)β1-4GlcNAcβ1 — 6Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-1R

9' 8' 10' 7'

4.98 4.39 5.16 4.37 4.65 4.27 4.68 4.30 4.68 4.27 4.17

[a]Data from (Ref. 17). R denotes a ceramide residue.

In a similar way the higher blood-group B active GSL B-III and B-IV could be characterized by their MS and NMR spectra. From their respective pseudomolecular ions, M+Na 3223 and 3672, they were identified as ceramide dodeca- and ceramide tetradecasaccharides. In addition, a ceramide hexadecasaccharide could be characterized by FAB MS exhibiting M+Na 4121. Their molecular ion regions are shown in Fig. 4. The observed mass increment of 827 amu between B-II and B-III indicates the addition of a second B-determinant branch Gal-(Fuc)-Gal-GlcNAc to the carbohydrate chain in B-III. B-IV and B-V show mass increments of 449 amu, in agreement with the addition of one or two lactosamine units to the molecule. Specific sequence ions indicate a biantennary structure in all three GSL and also the position of the additional lactosamine unit (Ref. 32).

The localization of one of the additional lactosamine units in the 1-3 branch of the biantennary part of B-IV could be inferred from the following NMR data. A comparison of B-III with its nonfucosylated analogue (Table 1) shows that glycosylation induces an upfield shift of the H-1 resonances of the neighbouring GlcNAc-7 and -7' units, the change being from 4.69 to 4.64 ppm for the former and from 4.42 to 4.36 ppm for the latter. In the spectrum of B-IV, there is a signal at 4.37 ppm but none at 4.42 ppm, hence there is no other 1-6 linked GlcNAc unit than that neighbouring the fucosylated antigenic determinant, i.e., the lactosamine unit must be located elsewhere. Taking into account that it was localized by MS in the branched part of the molecule (as indicated by an ion at m/z 1219), it is obviously situated in the 1-3 linked branch.

I and i antigens represent binding sites of red cell surfaces that are recognized by cold-reacting human monoclonal antibodies, which give rise to the cold agglutinin disease (Ref. 46,47), a special form of the autoimmune hemolytic anemia (Ref. 48). The erythrocytes of new born are especially rich in i-determinants whereas those of adult persons show I-activity. Similar changes from i to I structures can be observed with epithelial cell surface markers, which undergo specific structural changes during their cellular differentiation (Ref. 50). On the other hand, reverse changes with the expression of i-antigens have been demonstrated during malignant transformation (Ref. 50-52). The work of several laboratories has shed some light on the structural basis of I/i antigenicities: i-activity is essentially connected with linear structures of repetitive N-acetyllactosamine units that are linked β1-3 to each other whereas I activity is found in branched structures with additional N-acetyllactosamine units linked β 1-6 to the galactose residues. It has been shown that I,i antigens are not restricted in appearance to the human cells. GSL isolated from bovine or rabbit erythrocytes that carry blood-group B-like determinants have been particularly useful for the determination of the structural features of I,i antigenicity. As a result of these investigations, the hexasaccharide structure

$$\begin{array}{l} \text{Gal}\beta1\text{-4GlcNAc}\beta1 \\ \qquad\qquad\qquad\qquad\quad {}^{3}\!\diagdown \\ \qquad\qquad\qquad\qquad\quad\ \ \text{Gal}\beta1\text{-4GlcNAc} \\ \text{Gal}\beta1\text{-4GlcNAc}\beta1 \diagup {}^{6} \end{array}$$

was identified as the region recognized by most anti I antibodies (Ref. 53,54). The basic structure connected with i antigenicity is the linear tetrasaccharide. The relatively small number of GSL tested up to now show, however, only weak I or i activity when compared to the long chain or multibranched structures present on human, bovine or rabbit erythrocytes. Therefore, in order to get a deeper insight into the fine specificity of different I,i antibodies, larger GSL with blood group B-like activity were prepared in high yield from rabbit erythrocytes (Ref. 9). In this series of multibranched GSL, the major problems of structural elucidation were: (a) the arrangement of the side chains either as a series of short branches along one main chain or multibranched like a tree, (b) the sequence of the repetitive N-acetyllactosamine units and (c) the sites of attachment of the different branches either to carbon 3 or 6 of the galactose residues. The GSL BI_{rab}-2 with 15, BI_{rab}-3 with 20 and BI_{rab}-4 with 25 sugar residues were modified by treatment with α-galactosidase and subsequent Smith degradation. These degradation steps were undertaken in order to obtain NMR data on terminal GlcNAc residues, as our previous results on glycosylation shifts (Ref. 13, 20) suggested that this could possibly enable us to resolve the problem of whether the 1-3 or 1-6 linked branches are further bifurcated. The 1-3 linked GlcNAc residues exhibit their H-1 resonance at 4.64 ppm, if in terminal position, and at 4.68 ppm when situated inside of the oligosaccharide chain; for the 1-6 linked one the corresponding numbers are 4.37 and 4.42 ppm. In the spectrum of the ceramide nonasaccharide 3, obtained by degradation from the ceramide pentadecasaccharide 1 via the ceramide dodecasaccha-

TABLE 2. Anomeric proton chemical shifts (ppm from Me_4Si at 333 K in DMSO-$\underline{d}_6$/D_2O, 98:2) of the ceramide pentadecasaccharide 1, ceramide dodecasaccharide 2 and ceramide nonasaccharide 3

	$\underline{9}$ Galα1-3	$\underline{8}$ Galβ1-4	$\underline{7}$ GlcNAcβ1-3	$\underline{7}'$ GlcNAcβ1-6	$\underline{6}$ Galβ1-4	$\underline{5}$ GlcNAcβ1-3	$\underline{4}$ Galβ1-4	$\underline{3}$ GlcNAcβ1-3	$\underline{2}$ Galβ1-4	$\underline{1}$ Glcβ1-1Cer
1[a]	4.84	4.30	4.68	4.42	4.30	4.68	4.30	4.68	4.28	4.17
2		4.22	4.68	4.42	4.30	4.68	4.30	4.68	4.27	4.17
3			4.64	4.37	4.30	4.68	4.30	4.68	4.28	4.17

[a]Data from (Ref. 36).

ride 2 (Table 2), the integral intensity of the 4.64 ppm signals is 1H and the 4.37 ppm signal corresponds to 2H. The branching patterns shown in the heading of Tab. 2 is thus unequivocally established, since the alternative branching, which would be equivalent to the mutual exchange of the 1-3 and 1-6 branches at Gal-4, would require 2H for the 4.64 ppm signal, 1H for the 4.37 ppm one, and 1H for a signal of an internal, 1-6 linked GlcNAc unit at 4.42 ppm that, in fact, is missing.

This analysis can be extended to the higher, polyantennary GSL. For each of them, the lack of any internal, 1-6 linked GlcNAc unit (i.e., the lack of the 4.42 ppm resonance) proves that all of them are branched in the manner shown in the scheme of Fig. 5, with the trisaccharide antennae linked β 1-6 to the galactose residues of the repetitive N-acetyl lactosamine units.

The FAB MS data obtained on permethylated BI_{rab}-4, its α-galactosidase-(BI_{rab}-4α) and Smith degradation products (BI_{rab}-4αSm) are in good agreement with these findings. The UV trace of the molecular ion region of the permethylated BI_{rab}-4, recorded at 4kV acceleration voltage is shown in Fig. 5 together with the scheme of fragmentation. From interpolation between adjacent cesium iodide signals, the center of the main peak was determined as 6185 amu representing the most abundant molecular species of the formula $C_{288}H_{514}N_{10}O_{128}Na$. The major fragment ions observed in the FAB spectrum are indicated.

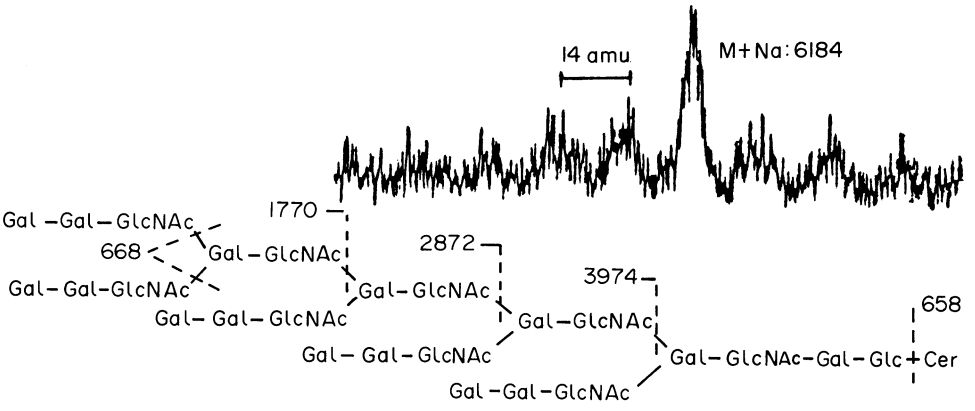

Fig. 5 UV trace of the molecular ion region of permethylated BI_{rab}-4 recorded at 4 kV acceleration voltage on a ZAB HF instrument and the observed fragmentation pattern.

All prominent ions representing the carbohydrate moiety are produced by fission of the glycosidic bonds of the GlcNAc residues. The regular branching pattern is indicated by this series of ions with a mass increment of 1102 amu between each of them. From this pattern of fragmentation, a tree-like branching can clearly be excluded. After treatment of BI_{rab}-4 with α-galactosidase, the molecular ion of the permethylated derivative is shifted to M+Na 5163, thus indicating the loss of five galactoses that were linked α-glycosidically at five branches. The pattern of the sequence ions is also changed in such a way that the difference between them now amounts to 898 amu, equivalent to two N-acetyllactosamine units (Fig. 6a). Due to the presence of five terminal N-acetyllactosamine units, the ion pair m/z 464 and 432 is very intense. A closer inspection of the spectrum shows, that besides the desired product, a small amount of another product is formed where one of the terminal galactoses has not been removed by the enzymatic treatment. This is indicated by the ions m/z 668 and 676. If this residual Gal were evenly distributed over the different branches of the molecule, one should expect that each of the sequence ions were accompanied by an ion of lower intensity at a value 204 amu higher. This is, however, not the case with the ions m/z 1362, 2260 and 3158. A signal significantly above the "chemical noise" of the matrix is only observed at m/z 4168 thus indicating that the Gal attached to the branch closest to the ceramide residue has in part escaped the action of the enzyme. The low proportion of this component did not permit the observance of the corresponding molecular ion.

After Smith degradation, the FAB spectrum of the permethylated product again exhibits a fragmentation pattern that is in line with the tetrabranched structure (Fig. 6b). The most intense ion pair is m/z 260 and 228, produced by the terminal GlcNAc residues.The other expected sequence ions are m/z 954, 1648, 2342 and 3036. Besides m/z 954, all other expected ions are of low intensity or absent. Instead, another series is present at m/z 709, 1403, 2097, 2791, and 204 mass units higher, m/z 2995. From these sequence ions, it can be deduced that during the oxidation with metaperiodate one of the GlcNAc residues was also degraded. It can further be concluded that this happened at the terminal or the subterminal branch. In addition, one of the galactoses linked to the branch closest to the ceramide residue partially escaped the degradation as seen from the ion m/z 2995. Due to the presence of several components in this sample no molecular ions of prominence could be observed.

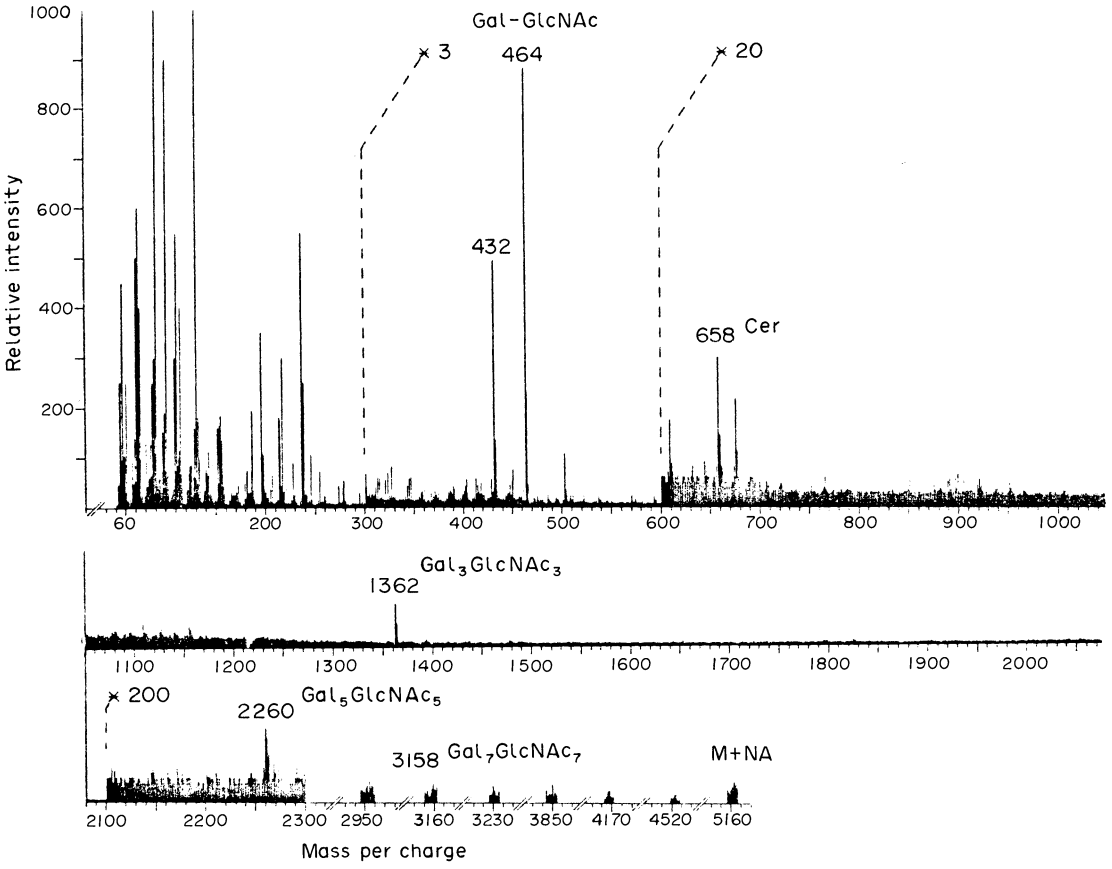

Fig. 6a. FAB MS spectrum of permethylated BI_{rab}-4α.

As outlined above, also native compounds are amenable to FAB MS. Since the ionization process takes place at ambient temperature without any vaporization process, highly polar and also high molecular weight substances can be analysed by this method. Pseudomolecular ions, M-1, of relative high intensity are observed especially in those compounds that carry anionic groups such as COO^- or $-SO_3^-$. As an example the FAB spectrum of the ganglioside GM1 is presented in Fig. 7. As to be seen in the upper trace, the thioglycerol matrix gives rise to intense ions at m/z 107, (M-1), 215 (2M-1), 323 (3M-1) etc. These signals are of high reproducibility. They can therefore be subtracted, thus yielding a "purified" spectrum that is shown in the lower trace of Fig. 7. As indicated in the scheme of fragmentation, sequence ions are formed starting from both ends of the molecule thus giving direct evidence of the sequence of the carbohydrate residues. The sequence ions derived from the ceramide residue always reflect the typical distribution of fatty acids and long chain bases by appropriate intensities as in the M-1 ions. Due to the presence of stearic acid, sphingosine and icosasphingosine as long chain bases, two ion groups appear in this series 28 amu apart starting with m/z 564, 592 for the ceramide residue and m/z 726, for the major glucosylceramide fragment. Due to the substitution of the Gal with the neuraminic acid, no lactosylceramide ions are present but instead the correspondingly higher ions at m/z 1179 and 1207. The further addition of an hexosamine residue to the chain can be deduced from the ion pair m/z 1382 and 1410. Among the ions that are derived from the "nonreducing" terminal of the molecule, m/z 308 and 290 represent the neuraminic acid. The ions produced by the terminal Gal or the Gal-GalNAc (m/z 382) residue are of very low intensity and therefore burried in the noise signals produced by the matrix. Only those ions that carry the NeuAc residue are of higher intensity. These ions can easily be recognized by an accompanying ion two mass units lower, that is formed after elimination of two hydrogen atoms. Such typical ion pairs are observed at m/z 833, 835 and at m/z 995, 997. Thus, two groups of ions give direct evidence on the position of the neuraminic acid in the chain. This is a very important advantage of negative ion FAB MS over more conventional EI or CI MS of the permethylated derivative which give only indirect evidence in most cases. From the analysis of a large number of gangliosides performed in collaboration with the groups of R. Schauer and G. Tettamanti it can be concluded, that negative ion FAB MS is the method of choice for sequence determination. These data in combination with those obtained by high resolution NMR, that have been presented by the group of B. Yu (30) will allow a complete structural elucidation with a minimal amount of substance. This will especially be true for those minor gangliosides that carry O-acetyl groups or lactone rings.

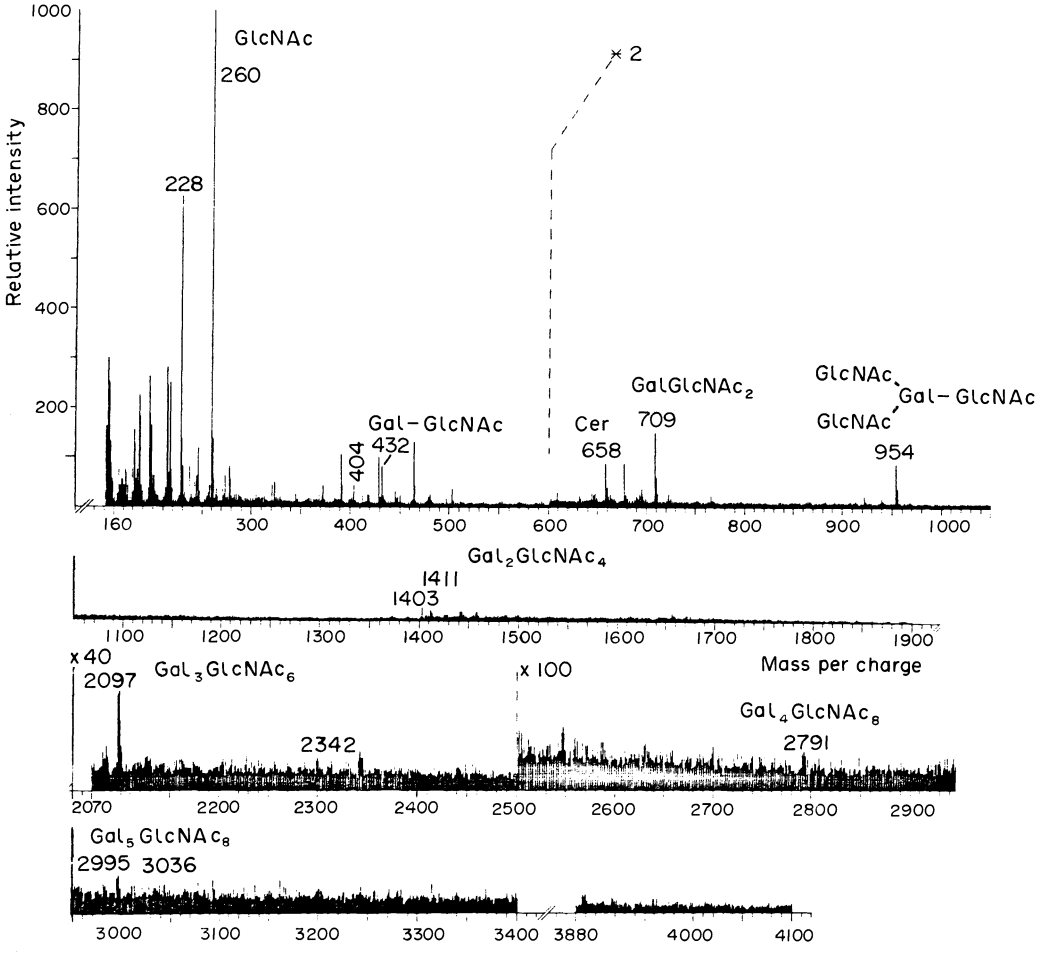

Fig. 6b. FAB MS spectrum of the product isolated after additional Smith degradation and permetylation.

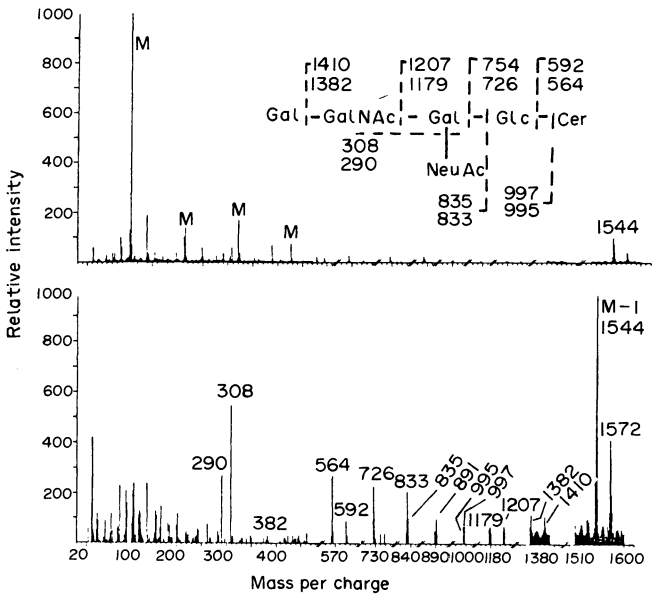

Fig. 7. Negative ion FAB spectrum of native ganglioside GM1 without (top) and with subtraction of the matrix-derived signals (bottom). Matrix-derived signals are indicated by M.

CONCLUSIONS

The presented examples were intended to show, that in most cases a complete elucidation of the primary structure of GSL can be based on the nondestructive NMR measurements and the sensitive methods of mass spectrometry. Referring to the list of parameters that have to be determined during the structural analysis it can be summarized that,

- sugar components and molecular mass are preferably determined by FAB MS using negative ion detection for native or negatively charged GSL or positive ion detection for permethylated (and peracetylated) derivatives, as well as by integration of the anomeric proton signals of the deuterium exchanged GSL, of the H-6 doublets of the fucose residues, and of the methyl singlets of the acetyl groups of hexosamines and sialic acids in the NMR spectra
- composition of the ceramide residue concerning long chain bases and fatty acids is evidenced by specific signals in FAB, EI MS and proton NMR
- the anomeric configuration of the sugars is clearly deducible from the coupling constants of the well separated H-1 doublets of deuterium exchanged GSL
- identification of the component sugar residues and of the conformation of their rings follows from the analysis of the full spin systems of the respective residues
- the sequence of sugar constituents and the pattern of branching is clearly reflected in characteristic sequence ions of the MS
- the sites of glycosidic linkages and the sequence of sugar constituents can be identified by interresidue NOE and scalar coupling across the glycosidic bond,and by specific glycosylation-induced shifts. If applied to sugar residues becoming terminal after α-galactosidase- and Smith degradation treatment, these shifts unambiguously identify the gross branching pattern.

REFERENCES

1. S.-I. Hakomori, Handbook of Lipid Research 3, 62-75, Plenum Press (1983).
2. V.N. Reinhold and S.A. Carr, Mass Spectrom. Rev. 2, 153-221 (1983).
3. M.E. Breimer, G.C. Hansson, K.-A. Karlsson, H. Leffler, W. Pimlott and B.E. Samuelsson, FEBS Letters 124, 299-303 (1981).
4. M.E. Breimer, G.C. Hansson, K.-A. Karlsson, H. Leffler, W. Pimlott and B.E. Samuelsson, Biomed. Mass Spectrom. 6, 231-241 (1979).
5. M. Barber, R.S. Bordoli, R.D. Sedgewick and A.N. Tyler, Nature 293, 270-275 (1981).
6. L.S. Forsberg, A. Dell, D.J. Dalton and C.E. Ballou, J. Biol. Chem. 257, 3555-3563 (1982).
7. M. Arita, M. Iwamori, T. Higuchi and Y. Nagai, J. Biochem. 93, 319-322 (1983).
8. H. Egge, J. Peter-Katalinić and P. Hanfland, in Ganqliosides, Structure, Function and Biomedical Potential (B. Ledeen et al., eds; in print) Plenum Press (1984).
9. P. Hanfland, U. Dabrowski, M. Kordowicz, J. Peter-Katalinić, J. Dabrowski and H. Egge, Proc. 7th Int. Symp. on Glycoconjugates (M.A. Chester et al. eds.) p.p. 413-414 (1983).
10. P. Hanfland, Chem. Phys. Lipids 15, 105-124 (1975).
11. P. Hanfland and H.A. Graham, Arch. Biochem. Biophys. 210, 396-404 (1981).
12. W.A. Gibbons, C.F. Beyer, J. Dadok, R.F. Sprecher and H.R. Wyssbrod, Biochemistry 14, 420-429 (1975).
13. J. Dabrowski, P. Hanfland and H. Egge, Biochemistry 19, 5652-5658 (1980).
14. G. Wagner and K. Wüthrich, J. Magn. Reson. 33, 675-680 (1979).
15. R.U. Lemieux and K. Bock, Jap. J. Antibiotics 32, s. 163-s. 169 (1979).
16. R.U. Lemieux, K. Bock, L.T.J. Delbeare, S. Koto and V.S. Rao, Can. J. Chem. 58, 631-653 (1980).
17. P. Hanfland, H. Egge, U. Dabrowski, S. Kuhn, D. Roelcke and J. Dabrowski, Biochemistry 20, 5310-5319 (1981).
18. J.P. Carver, A.A. Grey, F.M. Winnik, J. Hakimi, J. Ceccarini and P.H. Atkinson, Biochemistry 20, 6600-6606 (1981).
19. J. Dabrowski, P. Hanfland, H. Egge and U. Dabrowski, Arch. Biochem. Biophys. 210, 405-411 (1981).
20. J. Dabrowski, P. Hanfland and H. Egge, in Methods in Enzymology (V. Ginsburg, ed.), Vol. 83, pp. 69-86, Academic Press, New York (1982).
21. K. Nagayama, P. Bachmann, K. Wüthrich and R.R. Ernst, J. Magn. Reson. 31, 133-148 (1978).
22. G. Bodenhausen, R. Freeman, R. Niedermeyer and D.L. Turner, J. Magn. Reson. 26, 133-164 (1977).
23. W.P. Aue, E. Bartholdi and R.R. Ernst, J. Chem. Phys. 64, 2229-2246 (1976).
24. A. Bax and R. Freeman, J. Magn. Reson. 44, 542-561 (1981).
25. K. Nagayama, A. Kumar, K. Wüthrich and R.R. Ernst, J. Magn. Reson. 40, 321-334 (1980).
26. J. Jeener, B.H. Meyer, P. Bachmann and R.R. Ernst, J. Chem. Phys. 71, 4546-4553 (1979).
27. J. Dabrowski, H. Egge and P. Hanfland, Chem. Phys. Lipids 26, 187-196 (1980).
28. A. Yamada, J. Dabrowski, P. Hanfland and H. Egge, Biochim. Biophys. Acta 618, 473-479 (1980).
29. J.H. Prestegard, T.A.W. Koerner, Jr., P.C. Demou and R.K. Yu, J. Am. Chem. Soc. 104, 4993-4995 (1982).
30. T.A.W. Koerner, Jr., J.H. Prestegard, P.C. Demou and R.K.Yu, Biochemistry 22, 2676-2686 (1983).
31. S. Gasa, T. Mitsuyama and A. Makita, J. Lipid Res. 24, 174-182 (1983).
32. P. Hanfland, M. Kordowicz, H. Niermann, H. Egge, U. Dabrowski, J. Peter-Katalinić and J. Dabrowski, Eur. J. Biochem. (submitted).
33. M.A. Bernstein and L.D. Hall, J. Am. chem. Soc. 104, 5553-5555 (1982).
34. J. Dabrowski, H. Egge and U. Dabrowski, Carbohyd. Res. 114, 1-9 (1983).
35. K. Bock, J. Arnarp and J. Lönngren, Eur. J. Biochem. 129, 171-178 (1982).

36. J. Dabrowski and P. Hanfland, FEBS Lett. 142, 138-142 (1982).
37. U. Dabrowski, P. Hanfland, H. Egge, S. Kuhn and J. Dabrowski, J. Biol. Chem. (in print).
38. H. Friebolin, G. Keilich and E. Siefert, Org. Magn. Reson. 2, 457-465 (1970).
39. U. Dabrowski, H. Egge and J. Dabrowski, Arch. Biochem. Biophys. 224, 254-260 (1983).
40. H. van Halbeek, J.F.G. Vliegenthart, H. Winterwerp, W.M. Blanken and D.H. van den Eijnden, Biochem. Biophys. Res. Commun. 110, 124-131 (1983).
41. G. Batta and A. Lipták, J. Am. Chem. Soc. 106, 248-250 (1984).
42. H. Schenkel-Brunner, Eur. J. Biochem. 104, 529-534 (1980).
43. Z. Wilczyńska, H. Miller-Podraza and J. Kościelak, FEBS Letters 112, 277-279 (1980).
44. P. Hanfland and H. Egli, Vox Sang. 28, 438-452 (1975).
45. P. Hanfland, Eur. J. Biochem. 87, 161-170 (1978).
46. T. Feizi, E.A. Kabat, G. Vicar, B. Anderson and W.L. Marsh, J. Immunol. 106, 1578-1592 (1971).
47. S.-I. Hakomori, Semin. Hematol. 18, 39-62 (1981).
48. J.V. Dacie, The Hemolytic Anemias 2nd ed. part II, Grune and Stratton, New York (1962).
49. T. Feizi, Trends Biochem. Sci. 6, 333-335 (1981).
50. A. Kapadia, T. Feizi and H.J. Evans, Exp. Cell Res. 131, 185-195 (1981).
51. J. Picard, D.W. Edward and T. Feizi, J. Clin. Lab. Immunol. 1, 119-128 (1978).
52. A. Kapadia, T. Feizi, D. Jewell, J. Keeling and G. Slavin, J. Clin. Path. 34, 320-337 (1981).
53. T. Feizi, R.A. Childs, K. Watanabe and S.-I. Hakomori, J. Exp. Med. 149, 975-980 (1979).
54. K. Watanabe, S.-I. Hakomori, R.A. Childs and T. Feizi, J. Biol. Chem. 354, 3321-3328 (1979).