Carbon-13 nuclear-magnetic-resonance spectroscopy of Forssman hapten

Theodore A. W. KOERNER, Jr.,* Lewis W. CARY,† Su-Chen LI‡ and Yu-Teh LI‡

*Department of Laboratory Medicine, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510, U.S.A., †NMR Laboratory, Nicolet Technology, 145 E. Dana, Mountain View, CA 94041, U.S.A., and

‡Department of Biochemistry and Delta Regional Primate Research Center, Tulane University, New Orleans, LA 70112, U.S.A.

(Received 10 September 1981/Accepted 21 January 1981)

The ^{13}C n.m.r. spectrum of Forssman hapten was obtained at 25.16 MHz in $[^2\text{H}]\text{chloroform}/[^2\text{H}]\text{methanol}$ (1:1, v/v), using purified glycosphingolipid from canine intestinal mucosa (glycolipid I). All amide, olefin, anomeric, intersaccharide glycosidic ether, amide linkage, methyl and many methylene resonances were resolved and assigned. Analysis of the anomeric region reveals the following pentaglycosylceramide structure as originally proposed [Siddiqui & Hakomori (1971) *J. Biol. Chem.* **246**, 5766–5769]: $\text{GalNAc}(\alpha 1 \rightarrow 3)\text{GalNAc}(\beta 1 \rightarrow 3)\text{Gal}(\alpha 1 \rightarrow 4)\text{Gal}(\beta 1 \rightarrow 4)\text{Glc}(\beta 1 \rightarrow 1)\text{ceramide}$. Analysis of the amide, olefin and methylene regions reveals no α -hydroxy fatty acyl group and $\leq 6\text{ mol}\%$ unsaturated fatty acyl groups are present. Chemical-shift assignments are reported for the anomeric and glycosidic ether carbon atoms of intersaccharide-linked α -galactose and *N*-acetyl- α -galactosamine residues. Two rules are proposed for the assignment of the anomeric form of 1 $\rightarrow$ 3 and 1 $\rightarrow$ 4 linkages of galactose and *N*-acetylgalactosamine residues present in the glycone of glycoconjugates. The present study emphasizes the importance of the anomeric ‘window’ (80–120 p.p.m.) in studies of glycone structure.

Forssman hapten is a heterophilic antigen of sheep erythrocytes that is responsible for their haemolysis in the presence of antiserum and complement. This observation (Forssman, 1911) seventy years ago has served ever since as the basis for most clinical and experimental complement fixation assays (Alving, 1977). Forssman hapten is now known to occur in the cell membrane of erythrocytes and other tissues of many animal species (Alving, 1977; Sweeley & Siddiqui, 1977). Makita *et al.* (1966) were the first to isolate this antigen from horse tissues and concluded that it was a tetraglycosylceramide. However, Siddiqui & Hakomori (1971) using enzymic and rigorous chemical methods concluded Forssman hapten was a pentaglycosylceramide (I) with the structure: $\text{GalNAc}(\alpha 1 \rightarrow 3)\text{GalNAc}(\beta 1 \rightarrow 3)\text{Gal}(\alpha 1 \rightarrow 4)\text{Gal}(\beta 1 \rightarrow 4)\text{Glc}(\beta 1 \rightarrow 1)\text{ceramide}$ (Fig. 1).

Vance *et al.* (1966) first isolated a Forssman hapten from canine gastrointestinal tissue, which Smith *et al.* (1975), Slomiany & Slomiany (1977) and Li (1979) have reported has a structure (Fig. 1) identical with that of the antigen from horse tissues (Siddiqui & Hakomori, 1971). More recently, Slomiany & Slomiany (1978, 1979) have shown that there are three glycolipids from dog gastric mucosa

with Forssman immunoactivity. The Forssman glycolipid used in the present study corresponds to the glycolipid I described by Slomiany & Slomiany (1979).

Recently ^{13}C n.m.r. has been shown to be a powerful method for structural analysis of both neutral glycosphingolipids (Sweeley *et al.*, 1978; Koerner *et al.*, 1979; Li *et al.*, 1980; Dabrowski *et al.*, 1980) and gangliosides (Sillerud *et al.*, 1978; Harris & Thornton, 1978). However, one limitation of this method of analysis has been the absence from the literature of reference chemical shifts for the anomeric carbons of all of the possible intersaccharide glycosidic linkages. An interesting feature of the structure assigned to Forssman hapten is the presence of both α - and β -linked galactose and *N*-acetylgalactosamine intersaccharide residues. This is in contrast with the glycosphingolipids with ganglio-type oligosaccharide chains that have previously been well analysed (Sillerud *et al.*, 1978; Harris & Thornton, 1978), but contain only β -linked galactose and *N*-acetylgalactosamine residues. Thus, we have obtained the ^{13}C n.m.r. spectrum of Forssman hapten to confirm its structure as well as to reveal for the first time the chemical shifts of the anomeric and ether carbon atoms in intersaccharide-

Fig. 1. Structure of Forssman hapten from canine intestine (I). The fatty acyl residue shown is that of tetraosanoic acid ($C_{24:0}$), one of the major long-chain saturated species present.

linked α -galactose and *N*-acetyl- α -galactosamine residues.

Experimental

Glycosphingolipid with Forssman immuno-activity was obtained from dog intestine and isolated by previously published procedures (Vance *et al.*, 1966; McKibbin, 1969; Esselman *et al.*, 1972). The material used in the present study (I) is the major component or glycolipid I in the nomenclature of Slomiany & Slomiany (1979). The structure of compound (I) was confirmed by chemical and enzymic methods (Lindberg, 1972; Li, 1979).

The proton-decoupled natural-abundance ^{13}C n.m.r. spectrum of structure (I) at 40°C was obtained at 25.16 MHz on a Variant XL-100A spectrometer operating in pulsed Fourier-transform mode. A signal-to-noise ratio of 5:1 or more was obtained for most carbon resonances after 100 000 transients (28 h). The sample solution was 100 mg of compound (I) in 2 ml of $[\text{2H}]\text{chloroform}/[\text{2H}]\text{-methanol}$ (1:1, v/v) containing trace tetramethylsilane as internal reference. Chemical shifts are reported in parts per million downfield from tetramethylsilane.

All sugars in the present study are in the pyranose form and have the D-stereochemistry, except for L-rhamnose (Rha).

Results

As previously seen in the spectra of other glycosphingolipids (Koerner *et al.*, 1979; Li *et al.*, 1980; Dabrowski *et al.*, 1980; Sweeley *et al.*, 1978; Sillerud *et al.*, 1978; Harris & Thornton, 1978), the carbon atoms of Forssman hapten (I) resonate in seven distinct regions. Reasons for the assignment of individual carbon atoms (Table 1) within each region are discussed below. Individual carbon atoms are identified by residue and carbon number within each residue (Fig. 1).

Ketone and olefin region

Three resonances are seen in the carbonyl region (170–180 p.p.m.), as expected for the carbonyl carbon atoms of the three amide groups of structure (I). The most downfield resonance (175.0 p.p.m.) is assigned to the amide carbonyl carbon atom of the ceramide moiety (acyl-1), based on its identical chemical shift with the analogous carbon atom of glucosylceramide (Koerner *et al.*, 1979). The remaining amide resonances at 173.9 and 173.5 p.p.m. are thus due to the acetamido carbonyl carbon atoms of the two *N*-acetyl-galactosamine residues. Based on the observed deshielding of the acetamido carbonyl atom of the β -anomer relative to the α -anomer of methyl glycosides of *N*-acetyl-glucosamine (Perkins *et al.*, 1977), 173.9 p.p.m. is

assigned to C-7 of *N*-acetyl- β -galactosamine (see Table 1) and 173.5 p.p.m. to C-7 of *N*-acetyl- α -galactosamine. The two resonances in the olefin region at 134.7 and 129.7 p.p.m. are assigned to C-4 and C-5 of sphingosine respectively, based on their identical chemical shift with resonances from analogous carbon atoms of glucosylceramide (Koerner *et al.*, 1979).

Anomeric region

As anticipated (Koerner *et al.*, 1979; Li *et al.*, 1980), the anomeric 'window' (80–120 p.p.m.) reveals much structural information concerning the oligosaccharide moiety of compound (I). Five distinct resonances are observed in this region, indicating the presence of five saccharide residues and confirming the pentaglycosylceramide structure of Forssman hapten, first proposed by Siddiqui & Hakomori (1971) and later for compound (I) by others (Smith *et al.*, 1975; Slomiany & Slomiany, 1977; Li, 1979). Four of the anomeric resonances (102–105 p.p.m.) are exactly as seen in the spectrum (Sweeley *et al.*, 1978) of globoside [GalNAc(β 1 $\rightarrow$ 3)Gal(α 1 $\rightarrow$ 4)Gal(β 1 $\rightarrow$ 4)Glc(β 1 $\rightarrow$ 1)ceramide], indicating the new resonance at 95.2 p.p.m. must be C-1 of *N*-acetyl- α -galactosamine. Moreover, the location of this anomeric resonance is consistent with its α -linkage, since its is established for monosaccharides (Perlin, 1976), and expected for oligosaccharides, that C-1 of the α -anomer should be shielded relative to C-1 of the β -anomer. In fact, the resonance at 95.2 p.p.m. is quite upfield from the resonances between 102 and 105 p.p.m., one of which must be C-1 of the *N*-acetyl- β -galactosamine residue. Three of the four globoside-like anomeric resonances between 102 and 105 p.p.m. can be assigned for the first time to C-1 of β -glucose, β -galactose and α -galactose, due to their near-identical chemical shifts with resonances of analogous carbon atoms in glucosylceramide (Koerner *et al.*, 1979) and methyl 4'-*O*- α -D-galactopyranosyl β -lactoside (Cox *et al.*, 1978). The remaining resonance at 104.3 p.p.m. is assigned to *N*-acetyl- β -galactosamine by exclusion, as well as its near coincidence with the calculated value for the chemical shift of this carbon (104.2 p.p.m.). The calculated value was obtained by summing the chemical shift of C-1 of *N*-acetyl- β -galactosamine (96.5 p.p.m.) and the glycosidation shift of the anomeric carbon in a β 1 $\rightarrow$ 3 linkage (+7.7 p.p.m.) (both values taken from Colson & King, 1976), the β 1 $\rightarrow$ 3 glycosidation shift having been computed as the difference in the chemical shift of C-1 of *N*-acetylglucosamine in GlcNAc(β 1 $\rightarrow$ 3)L-Rha and in *N*-acetyl- β -glucosamine. To our knowledge included above is the first assignment of chemical shifts for the anomeric carbon atoms of Gal(α 1 $\rightarrow$ 4)Gal and GalNAc(α 1 $\rightarrow$ 3)GalNAc linkages.

Intersaccharide glycosidic ether region

Like the anomeric carbon atoms with which they are 'linkage-mates', these glycosidic ether carbon atoms are well resolved in a class, but are not quite as well resolved individually. Like the resonance assigned to its linkage-mate (C-1 of *N*-acetyl- α -galactosamine), the resonance at 79.6 p.p.m. is assigned to C-3 of *N*-acetyl- β -galactosamine because it is not present in the spectrum of globoside (Sweeley *et al.*, 1978). Of the three resonances in this region that are coincident with the same region of globoside, two of them 80.1 and 79.0 p.p.m. are assigned to C-4 of β -glucose and β -galactose respectively, due to their near-identical chemical shifts with analogous carbon atoms in 4'-*O*- α -D-galactopyranosyl- β -lactoside (Cox *et al.*, 1978). By exclusion, the third resonance at 80.6 p.p.m. is assigned to C-3 of α -galactose.

Ether and secondary alcohol region

Ether carbon atoms not involved in intersaccharide glucosidic linkages (ring ether carbon atoms and C-1 of sphingosine), the secondary-alcohol carbon atoms of the saccharide rings, and C-3 of sphingosine all resonate between 65 and 80 p.p.m. Consequently, 16 closely spaced resonances are present in this region and resolution of individual carbon atoms is not possible at 25.16 MHz. However, by comparison with model compounds (Cox *et al.*, 1978; Koerner *et al.*, 1979), integration and calculation of chemical shifts using known (Colson & King, 1976) Δ^{β} and Δ^{α} substituent effects, the multi-resonance signals in this region can be assigned to groups of individual carbon resonances (Table 1).

Primary-alcohol region

Two multi-resonance signals are observed between 60 and 64 p.p.m., one centred at 61.1 p.p.m. integrating for two carbon atoms and another centred at 62.2 p.p.m., integrating for three carbon atoms. By comparison with model compounds (Perkins *et al.*, 1977; Cox *et al.*, 1978) the upfield signal is assigned to C-6 of β -glucose and β -galactose and the downfield signal to C-6 of α -galactose, *N*-acetyl- β -galactosamine-6 and *N*-acetyl- α -galactosamine-6.

Amide linkage region

Between 45 and 55 p.p.m. are seen three resonances in a region characteristic of alkylamido methine carbon atoms. The resonance at 53.9 p.p.m. is assigned to C-2 of sphingosine since its chemical shift is identical with the analogous carbon atom of glucosylceramide (Koerner *et al.*, 1979). By comparison with the relative chemical shifts of the analogous carbons of the α - and β -methyl glyco-

Table 1. ^{13}C chemical shifts and assignments for Forssman hapten from canine intestine (I)

Carbon number	Residue	$\delta(\text{p.p.m.})^a$							
		...	Sphingosine	<i>N</i> -Acyl ^b	β -Glc	β -Gal	α -Gal	β -GalNAc	α -GalNAc
1			69.6 ^c	175.0	103.6	103.3	102.2	104.3	95.2
2			53.9	36.9	73.9 ^d	72.2 ^e	65.2 ^f	51.6	50.4
3			72.6	26.2	75.3 ^g	76.5	80.6	79.6	68.2
4			134.7	30.0	80.1	79.0	69.6 ^c	69.6 ^c	69.6 ^c
5			129.7	30.0	75.5 ^g	73.9 ^d	71.8 ^e	75.0 ^g	71.5 ^e
6			32.7	30.0	61.1 ^h	61.1 ^h	62.2 ⁱ	62.2 ⁱ	62.2 ⁱ
7			30.0	30.0				173.9	173.5
8			30.0	30.0				22.8 ^j	22.8 ^j
9-14			30.0	30.0					
15			29.7	30.0					
16			32.2	30.0					
17			22.9	30.0					
18			14.2	30.0					
19				30.0					
20				30.0					
21				29.7					
22				32.2					
23				22.9					
24				14.2					

^a The estimated accuracy of the chemical-shift measurements is ± 0.1 p.p.m.

^b Chain length arbitrarily assumed to be C_{24} .

^{c, d, h, i, j} Multi-resonance signals for which the resolution of individual resonances requires a higher resolution spectrum.

^{e, g} Groups of several resonances, the assignments for which may be interchanged.

^f Tentative assignment.

sides of *N*-acetylglucosamine (Perkins *et al.*, 1977), the resonances at 51.6 and 50.4 p.p.m. are assigned to C-2 of the β - and α -anomers of *N*-acetyl galactosamine, respectively.

Methylene and methyl region

All resonances between 14 and 40 p.p.m., except 22.8 p.p.m., are due to the alkyl tails of the sphingosine and acyl residues of structure (I) and are readily assigned (Table 1) since their chemical shifts are identical with the same residues of glucosylceramide (Koerner *et al.*, 1979). The resonance at 22.8 p.p.m. is assigned to C-8 of *N*-acetyl- β -galactosamine and *N*-acetyl- α -galactosamine due to its near-identical chemical shift with that of C-8 of the α - and β -methyl glucosides of *N*-acetylglucosamine (Perkins *et al.*, 1977).

Discussion

The ^{13}C n.m.r. data in Table 1 confirm the structure (Fig. 1) assigned to canine gastrointestinal Forssman hapten (Smith *et al.*, 1975; Slomiany & Slomiany, 1977; Li, 1979), a structure that is identical with that from horse tissues (Siddiqui & Hakomori, 1971). In addition, these data allow a gross fatty acyl analysis to be made for structure (I). Since no resonance was observed at 176.4, 34.9 or 25.6 p.p.m., as was seen in the spectrum of phreno-

sin (Li *et al.*, 1980), it is concluded that no α -hydroxy fatty acyl residue is present. The lack of detectable resonances at 130.2 and 27.4 p.p.m. indicates that 6 mol% or less unsaturated fatty acyl residues are present in structure (I).

Based on the chemical shift assignments several conclusions can be made concerning the resonances of carbon atoms of galactose and *N*-acetylgalactosamine residues involved in 1 $\rightarrow$ 3 and 1 $\rightarrow$ 4 intersaccharide linkages. (1) The anomeric carbon atoms of β 1 $\rightarrow$ 3 or β 1 $\rightarrow$ 4 linkages resonate at lower field than α 1 $\rightarrow$ 3 or α 1 $\rightarrow$ 4 linkages. Thus for 1 $\rightarrow$ 3 and 1 $\rightarrow$ 4 linkages at least, there seems to be no local effect to counter the expectation, based on monosaccharide studies (Perlin, 1976), that the anomeric carbon is an α -intersaccharide linkage is shielded relative to the anomeric carbon in the β -form of the same type of linkage. Tentatively, one may conclude that in chloroform/methanol β -linkages will resonate downfield from, and α -linkages upfield from, 102.5 p.p.m. relative to internal tetramethylsilane. (2) Likewise, the glycosidic ether carbon atoms of β 1 $\rightarrow$ 3 or β 1 $\rightarrow$ 4 linkages resonate at lower field than α 1 $\rightarrow$ 3 or α 1 $\rightarrow$ 4 linkages. Tentatively, one may conclude that in chloroform/methanol β -linkages will resonate downfield from, and α -linkages upfield from, 79.5 p.p.m. relative to internal tetramethylsilane.

The first of the above conclusions may be applied broadly to all possible intersaccharide linkages

between galactose and *N*-acetylgalactosamine residues, since in a test of its ability to predict the anomeric form of the 12 possible Gal→GalNAc or GalNAc→GalNAc and 16 possible Gal→Gal or GalNAc→Gal linkages, this rule was obeyed in every example found in the literature. The validity of the first conclusion may stem from the fact that galactose and *N*-acetylgalactosamine residues should always exist in the 4C_1 conformation for which the anomeric oxygen atom is always oriented axially in an α -linkage and equatorially in a β -linkage. Thus greater relative shielding by axially oriented versus equatorially oriented substituents may be the basis for this generality, as first noted in the mono-saccharide series (Perlin, 1976).

This work was supported by a National Science Foundation grant (no. PCM 79-22466) and National Institutes of Health grants (nos. NS 09626 and RR 00164).

References

- Alving, C. R. (1977) in *The Antigens* (Sela, M., ed.), vol. 4, pp. 1–72, Academic Press, New York
- Colson, P. & King, R. R. (1976) *Carbohydr. Res.* **47**, 1–13
- Cox, D. D., Metzner, E. K., Cary, L. W. & Reist, E. J. (1978) *Carbohydr. Res.* **67**, 23–31
- Dabrowski, J., Egge, H. & Hanfland, P. (1980) *Chem. Phys. Lipids* **26**, 187–196
- Esselman, W. J., Laine, R. A. & Sweeley, C. C. (1972) *Methods Enzymol.* **28**, 140–156
- Forssman, J. (1911) *Biochem. Z.* **37**, 78–115
- Harris, P. L. & Thornton, E. R. (1978) *J. Am. Chem. Soc.* **100**, 6738–6745
- Koerner, T. A. W., Jr., Cary, L. W., Li, S.-C. & Li, Y.-T. (1979) *J. Biol. Chem.* **254**, 2326–2328
- Li, Y.-T. (1979) in *Glycoconjugate Research* (Gregory, J. D. & Jeanloz, R. W., eds.), vol. 1, pp. 1–15, Academic Press, New York
- Li, Y.-T., Koerner, T. A. W., Jr. & Li, S.-C. (1980) *Int. Congr. Pure Appl. Chem.* **27th**, 199–204
- Lindberg, B. (1972) *Methods Enzymol.* **28**, 178–195
- Makita, A., Suzuki, C. & Yosizawa, Z. (1966) *J. Biochem. (Tokyo)* **60**, 502–513
- McKibbin, J. M. (1969) *Biochemistry* **8**, 679–684
- Perkins, S. J., Johnson, L. N., Phillips, D. C. & Dwek, R. A. (1977) *Carbohydr. Res.* **59**, 19–34
- Perlin, A. S. (1976) *Org. Chem., Ser. Two* **7**, 1–34
- Siddiqui, B. & Hakomori, S. (1971) *J. Biol. Chem.* **246**, 5766–5769
- Sillerud, L. O., Prestegard, J. H., Yu, R. K., Schafer, D. E. & Konigsberg, W. H. (1978) *Biochemistry* **17**, 2619–2628
- Slomiany, A. & Slomiany, B. L. (1977) *Eur. J. Biochem.* **76**, 491–498
- Slomiany, A. & Slomiany, B. L. (1979) in *Glycoconjugate Research* (Gregory, J. D. & Jeanloz, R. W., eds.), vol. 1, pp. 291–294, Academic Press, New York
- Slomiany, B. L. & Slomiany, A. (1978) *Eur. J. Biochem.* **83**, 105–111
- Smith, E. L., McKibbin, J. M., Karlsson, K. A., Paschev, I. & Samuelsson, B. E. (1975) *Biochim. Biophys. Acta* **388**, 171–179
- Sweeley, C. C. & Siddiqui, B. (1977) in *The Glycoconjugates* (Horowitz, M. I. & Pigman, W., eds.), vol. 1, pp. 459–540, Academic Press, New York
- Sweeley, C. C., Fung, Y.-K., Macher, B. A., Moskal, J. R. & Nunez, H. A. (1978) in *Glycoproteins and Glycolipids in Disease Processes* (Walborg, E. F., ed.), ACS Symposium Series vol. 80, pp. 47–85, American Chemical Society Press, Washington, DC
- Vance, W. R., Shook, C. P. & McKibbin, J. M. (1966) *Biochemistry* **5**, 435–445