Carbohydrate Research, 97 (1981) 331–336
Elsevier Scientific Publishing Company, Amsterdam – Printed in The Netherlands

Note

Natural-abundance ^{13}C-n.m.r. spectra of some model compounds for uncommon glycopeptide linkages in glycoproteins

ADAM ALLERHAND*, KILIAN DILL†, ELISHA BERMAN,
Department of Chemistry, Indiana University, Bloomington, IN 47405 (U.S.A.)
JEAN M. LACOMBE, AND ANDRÉ A. PAVIA
Laboratoire de Chimie Bioorganique, Faculté des Sciences, 33, rue Louis Pasteur, 84000 Avignon (France)

(Received January 10th, 1981; accepted for publication in revised form, May 8th, 1981)

The carbohydrate–peptide linkages that have thus far been most generally observed in glycoproteins[1] are those of 2-acetamido-1,2-dideoxy-β-D-glucopyranose to N-4 of asparagine[2], 2-acetamido-2-deoxy-α-D-galactopyranose to O-3 of serine or threonine, β-D-xylopyranose to O-3 of serine, and β-D-galactopyranose[2] to O-5 of hydroxylysine[1]. However, there is now strong evidence that linkages of arabinose to hydroxyproline occur in plant glycoproteins[2–6] and proteoglycans[7]; linkages of mannose to serine and/or threonine have also been reported or suggested[8–12], but the evidence was based on the alkaline β-elimination reaction, which is subject to considerable uncertainty[1,13]; linkages of galactose to serine have also been reported[4,5]. The foregoing are only some of the reported "uncommon" carbohydrate–peptide linkages.

We have shown that natural-abundance ^{13}C-n.m.r. spectroscopy can be used to study the structure and dynamic behavior of the carbohydrate residues of glycoproteins[14,15]. Knowledge of ^{13}C chemical shifts of appropriate model compounds may facilitate the use of ^{13}C n.m.r. to study carbohydrate–peptide linkages in intact glycoproteins and in glycopeptides. Use of ^{13}C n.m.r. should also indicate the anomeric configuration of the carbohydrate residue involved in the glycosidic linkage.

In this report, we present the chemical shifts and assignments of the ^{13}C resonances of 3-*O*-(α-D-mannopyranosyl)-L-serine (Man*p*-Ser), 3-*O*-(α-D-mannopyranosyl)-L-threonine[1] (αMan*p*-Thr), 4-*O*-(α-L-arabinofuranosyl)-L-hydroxyproline[1] (αAra*f*-Hyp), and 4-*O*-(β-L-arabinofuranosyl)-L-hydroxyproline[1] (βAra*f*-Hyp). In general, the published reports of Ara-Hyp, Man-Ser, and Man-Thr linkages have

*Address requests for reprints to this author.
†Present address: Department of Chemistry, Clemson University, Clemson, SC 29631, U.S.A.

not included determinations of the anomeric configuration of the arabinosyl and mannosyl groups. However, we have reason to believe that α-D-mannopyranosyl linkages to serine and threonine are present in the glucoamylase from *Aspergillus niger*[16]. The L-arabinosyl group linked to hydroxyproline in potato lectin is the β-furanosyl tautomer[5]. The α-arabinofuranosyl–hydroxyproline linkage has been found in rice-bran proteoglycans[7]. Recently, ^{13}C-n.m.r. spectroscopy was used by Akiyama *et al.*[6] to show that extensin, a cell-wall glycoprotein, contains both α- and β-L-arabinofuranose residues and that the β-furanosyl anomer is linked to hydroxyproline[6].

EXPERIMENTAL

The synthesis of αMan-Ser and αMan-Thr will be reported elsewhere[17,18].

2-N-(Benzyloxycarbonyl)-4-O-(2,3,5-tri-O-benzyl-α-L-arabinofuranosyl)-L-hydroxyproline benzyl ester (**1**) *and its β anomer* (**2**). — Trifluoromethanesulfonic anhydride (0.25 mL, 1.5 mmol) was added in the cold (−15°) to a solution of *N*-(benzyloxycarbonyl)-L-hydroxyproline benzyl ester (Z-Hyp-OBzl, 1.06 g, 3 mmol) in acetonitrile (30 mL). A solution of 2,3,5-tri-*O*-benzyl-L-arabinofuranose (0.42 g, 1 mmol) in dichloromethane (10 mL) was added during 15 min to the foregoing mixture kept at −15°. When t.l.c. showed no starting material (arabinofuranose derivative) remaining (~30 min), the mixture was washed with water (20 mL) and extracted with diethyl ether (2 × 30 mL). The organic layer was washed successively with saturated aqueous sodium hydrogencarbonate and water, and then dried over sodium sulfate and evaporated under vacuum at 50°. Column chromatography on silica gel with 19:1 (v/v) chloroform–diethyl ether as eluent gave **1** as an oil (0.35 g, 45% yield), together with the β anomer (**2**, 0.30 g, 39% yield). Washing of the column with methanol gave the remaining Z-Hyp-OBzl (0.65 g, 90%). Compound **1**, an oil, had $[\alpha]_D^{22}$ −62.4° (*c* 2.1, chloroform); ^{13}C-n.m.r.: δ 172.34 (C=O ester), 155.1 (C=O urethane), 105.35 (C-1 arabinose), 57.0 (C-2 Hyp), 53.34 and 53.02 (C-5 Hyp), and 36.13 and 35.18 (C-3 Hyp). Compound **2** was an oil, $[\alpha]_D^{22}$ +15.5° (*c* 1.8, chloroform); ^{13}C-n.m.r.: δ 172.52 (C=O ester), 154.63 (C=O urethane), 100.55 and 100.15 (C-1 Ara), 58.13 (C-2 Hyp), 51.56 (C-5 Hyp), and 37.5 and 36.8 (C-3 Hyp).

4-O-(α-L-Arabinofuranosyl)-L-hydroxyproline (**3**) *and 4-O-(β-L-arabinofuranosyl)-L-hydroxyproline* (**4**). — Compound **1** (0.37 g, 0.5 mmol), dissolved in a mixture of ethanol (20 mL), acetic acid (15 mL), and water (5 mL) was hydrogenated for 24 h at room temperature in the presence of hydrogen (4 bars) and 10% palladium-on-charcoal (1.3 g). The mixture was filtered through a bed of Celite and the catalyst was thoroughly washed with hot water. The oily residue obtained after evaporation of the aqueous solution under vacuum at 45° was washed with water and evaporated three times. The third time the aqueous solution was treated with Amberlite IRC-50 mesh before evaporation. Compound **3** was obtained crystalline (0.125 g, 95% yield), m.p. 121–123° dec., $[\alpha]_D^{22}$ −115° (*c* 1.0, water). Compound **4** was obtained from

2, by the foregoing procedure, as a crystalline, very hygroscopic product (92% yield); $[\alpha]_D^{22}$ +14.4° (*c* 1.2, water).

Proton-decoupled, natural-abundance, ^{13}C-n.m.r. spectra were recorded at 67.9 MHz (63.4 kG), essentially as described[14], except that a 10-mm probe instead of the 15-mm probe was used for βAra*f*-Hyp. Chemical shifts are reported in p.p.m. downfield from the ^{13}C resonance of tetramethylsilane (Me_4Si), and have an estimated precision of ±0.03 p.p.m. They were measured digitally with respect to a trace of internal 1,4-dioxane, added only when recording spectra for chemical-shift measurements. The chemical shift of 1,4-dioxane was taken as 67.86 p.p.m.

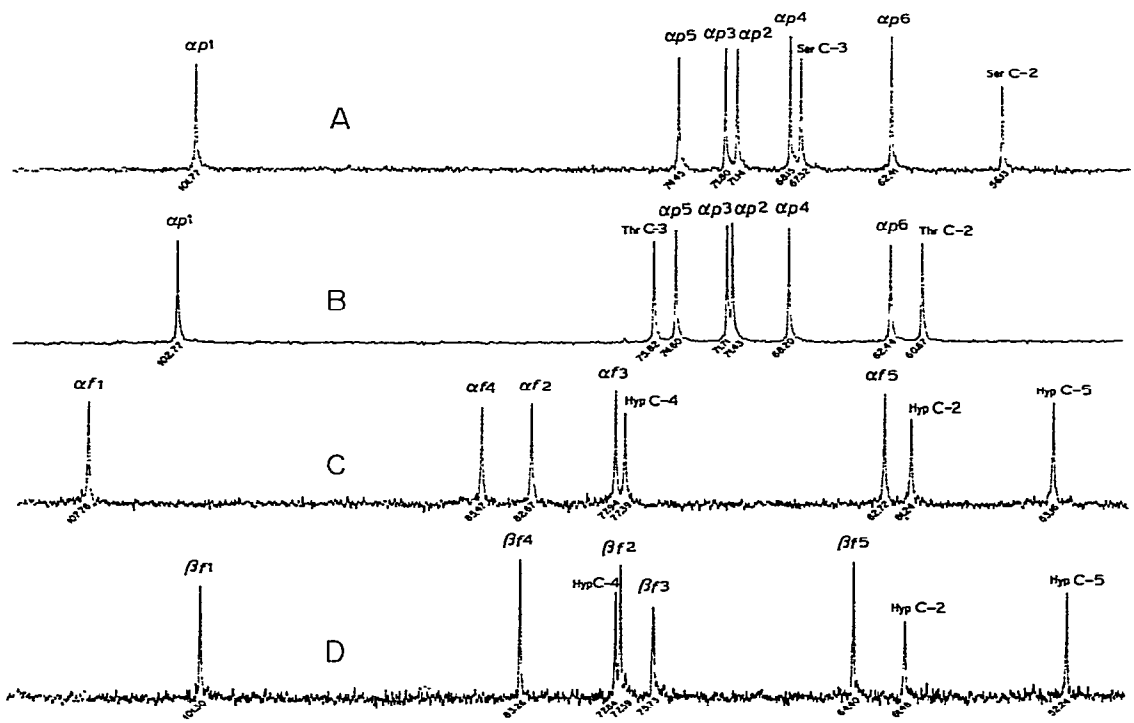

Fig. 1. Regions of aliphatic carbon atoms (except C-4 of threonine and C-3 of hydroxyproline) in the proton-decoupled, natural-abundance, ^{13}C-n.m.r. spectra of some glycosylated amino acids. The number immediately above each peak is the chemical shift. The symbols *αp* and *αf* refer to α-pyranose and α-furanose anomers, respectively. The number following each *αp* and *αf* designation indicates a carbohydrate carbon atom. (A) 55mM αMan*p*-Ser in H_2O (15-mm probe), pH 6.6, ~40°, 4096 accumulations with a recycle time of 5 s. (B) 50mM αMan*p*-Thr, and other conditions as in spectrum A. The C-4 resonance of the threonine residue is at 19.63 p.p.m. (C) 40mM αAra*f*-Hyp, pH 6.5, and other conditions as in spectrum A. The chemical shift of C-3 of the hydroxyproline residue is 36.15 p.p.m. (D) About 0.1M βAra*f*-Hyp (10-mm probe), pH 6.4, 35°, 2800 accumulations, and other conditions as in spectrum A. The chemical shift of C-3 of the hydroxyproline residue is 37.39 p.p.m. Acetate (~0.1M) was present in the sample of βAra*f*-Hyp.

RESULTS AND DISCUSSION

Figs. 1A, 1B, 1C, and 1D show the region of aliphatic carbon atoms (except C-4 of threonine and C-3 of hydroxyproline) in the ^{13}C-n.m.r. spectra of αMan*p*-Ser, αMan*p*-Thr, αAra*f*-Hyp, and βAra*f*-Hyp, respectively, in H_2O at pH 6.5. Chemical shifts and assignments to specific carbon atoms are shown above the peaks. The chemical shifts of C-4 of threonine, and C-3 of hydroxyproline are given in the legend to Fig. 1.

The assignments given in Fig. 1 are based on reported chemical shifts of model compounds (see later). However, because the resonances of C-3 of the serine residue (Fig. 1A), C-2 and C-3 of the threonine residue (Fig. 1B), and C-2 and C-4 of the hydroxyproline residue (Figs. 1C and 1D) are fairly close to carbohydrate-carbon resonances, these resonances of amino acid residues were first identified by means of the effect of pH on chemical shifts: As expected, the chemical shifts of the carbohydrate carbon atoms were relatively unaffected ($\lesssim$0.7 p.p.m. change, see later) when going from the cationic (pH $\lesssim$1.5) to the anionic state (pH $\gtrsim$10) of the amino acid residues, whereas the chemical shifts of C-2 and C-3 of all three amino acid residues and that of C-4 of hydroxyproline changed by more than 2 p.p.m. In the case of βAra*f*-Hyp (Fig. 1D), the resonance of C-2 of the arabinose residue is only 0.25 p.p.m. away from that of Hyp C-4 (at pH $\lesssim$8), and it moves $\sim$0.2 p.p.m. downfield when going from the zwitterionic (pH 6.4) to the anionic state (pH 11.2). Therefore, in order to make one-to-one assignments for these two resonances on the basis of the pH-dependence of chemical shifts, it was necessary to record spectra at pH values in the titration range of the amino group (specifically, pH 9.0 and 9.5).

The specific assignments for C-2 of the serine moiety, C-2 and C-4 of the threonine residue, and C-2, C-3, and C-5 of the hydroxyproline follow readily from the reported chemical shifts of serine[19], threonine[19], and hydroxyproline[20] residues in peptides*. The specific assignments for the amino acid carbon atoms involved in glycosidic linkages (C-3 of serine and threonine, and C-4 of hydroxyproline) follow by elimination. As expected[21], these chemical shifts are $\sim$5–8 p.p.m. downfield from the corresponding reported values for the unglycosylated serine[19], threonine[19], and hydroxyproline residues[20] in peptides.

After all of the amino acid carbon resonances are identified, the specific assignments of the carbohydrate carbon resonances follow readily from the reported ^{13}C chemical shifts[22] of methyl α-D-mannopyranoside, methyl α-L-arabinofuranoside, and methyl β-L-arabinofuranoside.

It is not necessarily true that the chemical shifts of the sugar moieties of αMan*p*-Ser, αMan*p*-Thr, αAra*f*-Hyp, and βAra*f*-Hyp, are similar to the corresponding

*The chemical shifts of Chien and Wise[20] are referenced to CS_2, and must be converted to chemical shifts relative to Me_4Si before comparing with our values. We confirmed our relative assignments of C-2 (a methine carbon atom) and C-5 (a methylene carbon atom) of the hydroxyproline moiety by means of an off-resonance, single-frequency, proton-decoupled, ^{13}C-n.m.r. spectrum of αAra*f*-Hyp.

chemical shifts in glycoproteins because of several reasons. (*i*) It is possible, in principle, that peptide-bond formation affects the ^{13}C chemical shifts of the carbohydrate residues. However, we expect this effect to be small, on the basis of the pH dependence of the chemical shifts: when going from the cationic to the zwitterionic state, the chemical shifts of the carbohydrate moieties of the four compounds under consideration change by 0.2 p.p.m. or less; when going from the zwitterionic to the anionic state, there are a number of significant chemical-shift changes, but all are $\lesssim$0.6 p.p.m. These results strongly suggest that peptide-bond formation will not cause major changes in the chemical shifts of the carbohydrate carbon atoms. (*ii*) It may be necessary to consider the effect of protein folding on the ^{13}C chemical shifts of the carbohydrate moieties. However, we have found for bovine pancreatic ribonuclease B that protein folding does not significantly affect the ^{13}C chemical shifts of even the carbohydrate residue involved in the glycosidic linkage[23]. (*iii*) It is obviously necessary to take into account the effects of participation of a nonanomeric carbon atom in a glycosidic linkage to another carbohydrate residue[21,24].

Fig. 1 has several implications for studies of glycoproteins by ^{13}C-n.m.r.: (*i*) the chemical shift of C-1 of the α-arabinofuranose moiety (Fig. 1C) is downfield of any expected anomeric carbon resonances of the commonly occurring *pyranose* anomers of carbohydrate residues of glycoproteins. Therefore, this resonance may be useful for detecting α-arabinofuranose residues in glycoproteins. However, we must consider possible interference from anomeric carbon atoms of some other carbohydrate residues in the furanose form, such as β-xylofuranose[25,26]; (*ii*) the chemical shift of C-4 of the α-arabinofuranose moiety (Fig. 1C) is downfield of any expected nonanomeric carbon resonances (even those involved in glycosidic linkages) of the commonly occurring pyranose anomers of carbohydrate residues of glycoproteins, so that this resonance may also be useful for detecting α-arabinofuranosyl residues; (*iii*) the use of our chemical-shift values for α-Man*p*-Ser (Fig. 1A) and α-Man*p*-Thr (Fig. 1B) for identifying mannose linkages to serine and threonine residues in glycoproteins will have to take into account the proximity of many other ^{13}C resonances of carbohydrate residues[21]. However, it may be possible to overcome such difficulties by considering differences in linewidths. We find that, for the native glucoamylase I from *Aspergillus niger*, the resonances of C-1 of the α-D-mannopyranosyl residues directly attached to the polypeptide backbone are considerably broader than those of carbohydrate residues not directly bonded to the backbone[16].

ACKNOWLEDGMENT

This work was supported by the United States Public Health Service (Grant GM 22620).

REFERENCES

1 R. D. MARSHALL AND A. NEUBERGER, *Adv. Carbohydr. Chem. Biochem.*, 25 (1970) 407–478.
2 D. T. A. LAMPORT, *Nature*, 216 (1967) 1322–1324.

3 D. T. A. LAMPORT, *Biochemistry*, 8 (1969) 1155–1163.
4 D. T. A. LAMPORT, in F. LOEWUS (Ed.), *Biogenesis of Plant Cell Wall Polysaccharides*, Academic Press, New York and London, 1973, pp. 149–164.
5 A. K. ALLEN, N. N. DESAI, A. NEUBERGER, AND J. M. CREETH, *Biochem. J.*, 171 (1978) 665–674.
6 Y. AKIYAMA, M. MORI, AND K. KATO, *Agric. Biol. Chem.*, 44 (1980) 2487–2489.
7 T. YAMAGISHI, K. MATSUDA, AND T. WATANABE, *Carbohydr. Res.*, 50 (1976) 63–74.
8 D. R. LINEBACK, *Carbohydr. Res.*, 7 (1968) 106–108.
9 J. H. PAZUR, H. R. KNULL, AND D. L. SIMPSON, *Biochem. Biophys. Res. Commun.*, 40 (1970) 110–116.
10 H. GREILING, P. VÖGELE, R. KISTERS, AND H.-D. OHLENBUSCH, *Z. Physiol. Chem.*, 350 (1969) 517–518.
11 A. L. ROSENTHAL AND J. H. NORDIN, *J. Biol. Chem.*, 250 (1975) 5295–5303.
12 M. K. RAIZADA, J. S. SCHUTZBACH, AND H. ANKEL, *J. Biol. Chem.*, 250 (1975) 3310–3315.
13 U. LINDAHL AND L. RODÉN, in A. GOTTSCHALK (Ed.), *Glycoproteins, Their Composition, Structure, and Function*, Elsevier, Amsterdam and New York, 2nd edn., Part A, 1972, p. 510.
14 K. DILL AND A. ALLERHAND, *J. Biol. Chem.*, 254 (1979) 4524–4531.
15 E. BERMAN, A. ALLERHAND, AND A. L. DEVRIES, *J. Biol. Chem.*, 255 (1980) 4407–4410.
16 K. DILL AND A. ALLERHAND, unpublished results.
17 J. M. LACOMBE, A. A. PAVIA, AND J. M. ROCHEVILLE, *Can. J. Chem.*, 59 (1981) 473–481.
18 A. A. PAVIA, S. N. UNG-CHHUN, AND J. M. LACOMBE, *Nouv. J. Chim.*, unpublished results.
19 R. RICHARZ AND K. WÜTHRICH, *Biopolymers*, 17 (1978) 2133–2141.
20 J. C. W. CHIEN AND W. B. WISE, *Biochemistry*, 12 (1973) 3418–3424.
21 S. N. ROSENTHAL AND J. H. FENDLER, *Adv. Phys. Org. Chem.*, 13 (1976) 279–424.
22 P. A. J. GORIN AND M. MAZUREK, *Can. J. Chem.*, 53 (1975) 1212–1223.
23 E. BERMAN, D. E. WALTERS, AND A. ALLERHAND, *J. Biol. Chem.*, 256 (1981) 3853–3857.
24 P. A. J. GORIN, *Carbohydr. Res.*, 39 (1975) 3–10.
25 R. G. S. RITCHIE, N. CYR, B. KORSCH, H. J. KOCH, AND A. S. PERLIN, *Can. J. Chem.*, 53 (1975) 1424–1433.
26 P. A. J. GORIN AND M. MAZUREK, *Carbohydr. Res.*, 48 (1976) 171–186.