Volume 314, number 3, 389–394 FEBS 11906 December 1992

Human urokinase contains GalNAcβ(1-4)[Fucα(1-3)]GlcNAcβ(1-2) as a novel terminal element in *N*-linked carbohydrate chains

Aldert A. Bergwerff[a], Jane E. Thomas-Oates[b], Jan van Oostrum[c], Johannis P. Kamerling[a] and Johannes F.G. Vliegenthart[a]

Bijvoet Center, Department of [a]Bio-Organic Chemistry and [b]Mass Spectrometry, Utrecht University, Utrecht, The Netherlands and *[c]Ciba-Geigy Ltd., Department of Biotechnology, Basel, Switzerland*

Received 28 October 1992; revised version received 4 November 1992

Structural analysis of enzymically released *N*-linked carbohydrate chains of human urokinase (urinary-type plasminogen activator) by ^{1}H NMR spectroscopy and FAB-MS demonstrated that the *N*-linked oligosaccharides on the only *N*-glycosylation site contain diantennary structures with the novel GalNAcβ(1-4)[Fucα(1-3)]GlcNAcβ(1-2) element in the upper or the lower branch.

Urokinase; Plasminogen activator; *N*-Linked carbohydrate chain; Fast atom bombardment mass spectroscopy (FAB-MS); ^{1}H NMR spectroscopy

1. INTRODUCTION

Urinary-type plasminogen activator (u-PA or urokinase) is a serine protease that converts plasminogen into the active fibrinolytic enzyme, plasmin, and is of clinical interest as a thrombolytic agent for treating patients suffering from (acute) vascular occlusions, like myocardial infarction. In addition, u-PA plays a role in extracellular proteolysis during physiological processes, such as gland involution, and in pathological processes, such as tumor growth [1,2]. Human u-PA bears *N*-linked carbohydrate chains only at Asn-302, containing besides Man, Gal, Fuc and Neu5Ac at least four GlcNAc and two GalNAc residues [3,4], and *O*-linked Fuc at Thr-18 [5,6]. Since no detailed structural analysis of the *N*-linked carbohydrate chains of u-PA has been reported, a study was initiated on the carbohydrate part of this glycoprotein. Here we report the occurrence of a novel terminal element, namely GalNAcβ(1-4)[Fucα(1-3)]GlcNAcβ(1-2), present in diantennary oligosaccharides.

Correspondence address: J.P. Kamerling, Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, PO Box 80.075, NL-3508 TB Utrecht, The Netherlands. Fax: (31) (30) 540 980.

Abbreviations: u-PA, urinary type plasminogen activator; t-PA, tissue-type plasminogen activator; EGF, epidermal growth factor; Fuc, fucose; Neu5Ac, *N*-acetylneuraminic acid; CHO, Chinese hamster ovary; PNGase-F, peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F; n.O.e, nuclear Overhauser enhancement; HOHAHA, homonuclear Hartmann-Hahn; MLEV, composite pulse devised by Malcom Levitt; NOESY, nuclear Overhauser enhancement spectroscopy; FAB-MS, fast atom bombardment mass spectrometry; GLC, gas liquid chromatography; FPLC, fast protein liquid chromatography; HPLC, high-pressure liquid chromatography; Le^x, Lewis x.

2. MATERIALS AND METHODS

2.1 *Materials*

Peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F (PNGase-F) from *Flavobacterium meningosepticum* was purchased from Boehringer-Mannheim (Germany). Urokinase (54 kDa), isolated from a human urine pool (Ukidan), was obtained from Laboratoires Serono (Aubonne, Switzerland); each vial contained 500,000 IU urokinase, 1.9 mg NaCl, 20 mg mannitol, 2 mg EDTA and 1.5 mg Na_2HPO_4/NaH_2PO_4.

2.2. *Liberation and isolation of the N-linked carbohydrate chains*

Three batches, each of 10 vials, were used for structural analysis. After removal of additives, the *N*-linked carbohydrate chains were released enzymically from the glycoprotein with PNGase-F, and subsequently fractionated by gel-permeation chromatography, FPLC on Mono Q and HPLC on 5 μm LiChrospher 100-NH_2, essentially as described [7].

2.3. *Analytical methods*

1D 500 MHz ^{1}H NMR spectra were recorded as described [8]. For the recorded 600 MHz 2D HOHAHA spectrum at 300 K [9,10] a MLEV-17 mixing sequence of 120 ms was applied, using a spin-lock field-strength corresponding to a 90° ^{1}H pulse-width of 25.0 μs. A data matrix of 464 × 2,048 points, representing a spectral width of 4,800 Hz in each dimension, was recorded. The 500 MHz 2D NOESY spectrum [11] was acquired at 280 K with a mixing time of 200 ms. A data matrix of 324 × 2,048 points, representing a spectral width of 4,800 Hz in each dimension, was recorded. In the case of the 2D NMR experiments, the ^{1}HO^{2}H signal was presaturated for 1 s during the relaxation delay. Phase-sensitive handling of the date in the ω_1 dimension became possible by the time-proportional phase increment method [12]. 2D NMR data were processed as reported in [13].

For linkage analysis, permethylation was carried out according to [14]. An aliquot of the permethylated oligosaccharide sample was hydrolysed, reduced, and acetylated as described [15]. GLC-MS was carried out on a JEOL JMS-AX505W mass spectrometer with a Hewlett Packard 5980 gas chromatograph, fitted with a CP Sil 5CB column (0.32 mm × 25 m, Chrompack). The sample was injected on-column at 90°C, after 2 min the temperature was increased to 140°C at 30°C/min, and then to 230°C at 4°C/min, and electron impact mass spectra were recorded (at 3 kV).

Positive-ion fast atom bombardment mass spectrometry (FAB-MS) of the permethylated oligosaccharide sample was performed using MS1 of a JEOL JMS-SX/SX102A tandem mass spectrometer, using 8 kV accelerating voltage.

3. RESULTS

The pool of enzymically released *N*-linked carbohydrate chains of urokinase, obtained after separation from the protein, was fractionated on Mono Q, yielding 8 carbohydrate-positive peaks. The neutral FPLC fraction, N1, was subfractionated by HPLC on LiChrospher 100-NH_2, yielding 7 carbohydrate-positive peaks. In this paper, the analysis of HPLC subfraction, N1.4, will be discussed in more detail.

An aliquot of fraction N1.4 was permethylated and analyzed using positive-ion mode FAB-MS (Fig. 1). The most intense pseudomolecular ion was observed at *m/z* 2,477 (corresponding to $[M+H]^+$ for fully methylated deoxyhex$_2$Hex$_3$HexNac$_6$). A^+-type sequence ions [16] were observed at *m/z* 260 (HexNAc$^+$), *m/z* 505 (HexNAc-HexNAc$^+$), *m/z* 679 (deoxyhex$_1$HexNAc$_2^+$) and *m/z* 2,026 (deoxyhex$_1$Hex$_3$HexNAc$_5^+$). These ions, together with the absence of any additional A^+-type ions in the spectrum, indicate that the major component in fraction N1.4 is a diantennary structure with one

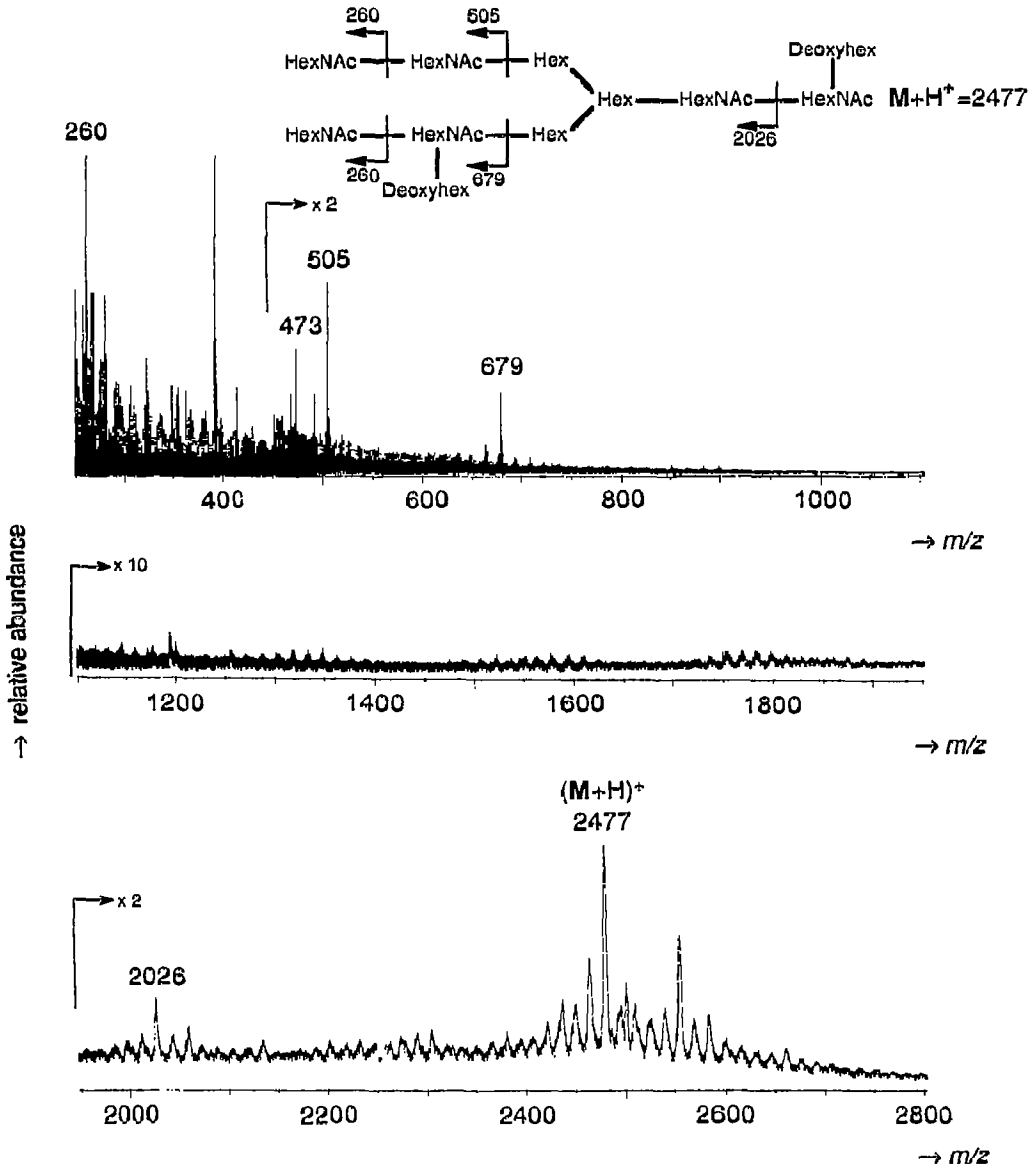


Fig. 1. Fast atom bombardment mass spectrum of permethylated fraction N1.4. The matrix used was thioglycerol, and the bombarding gas was Xe. *m/z* values are quoted as nominal masses. The ion observed at *m/z* 473 arises by β-elimination from *m/z* 679 (see text).

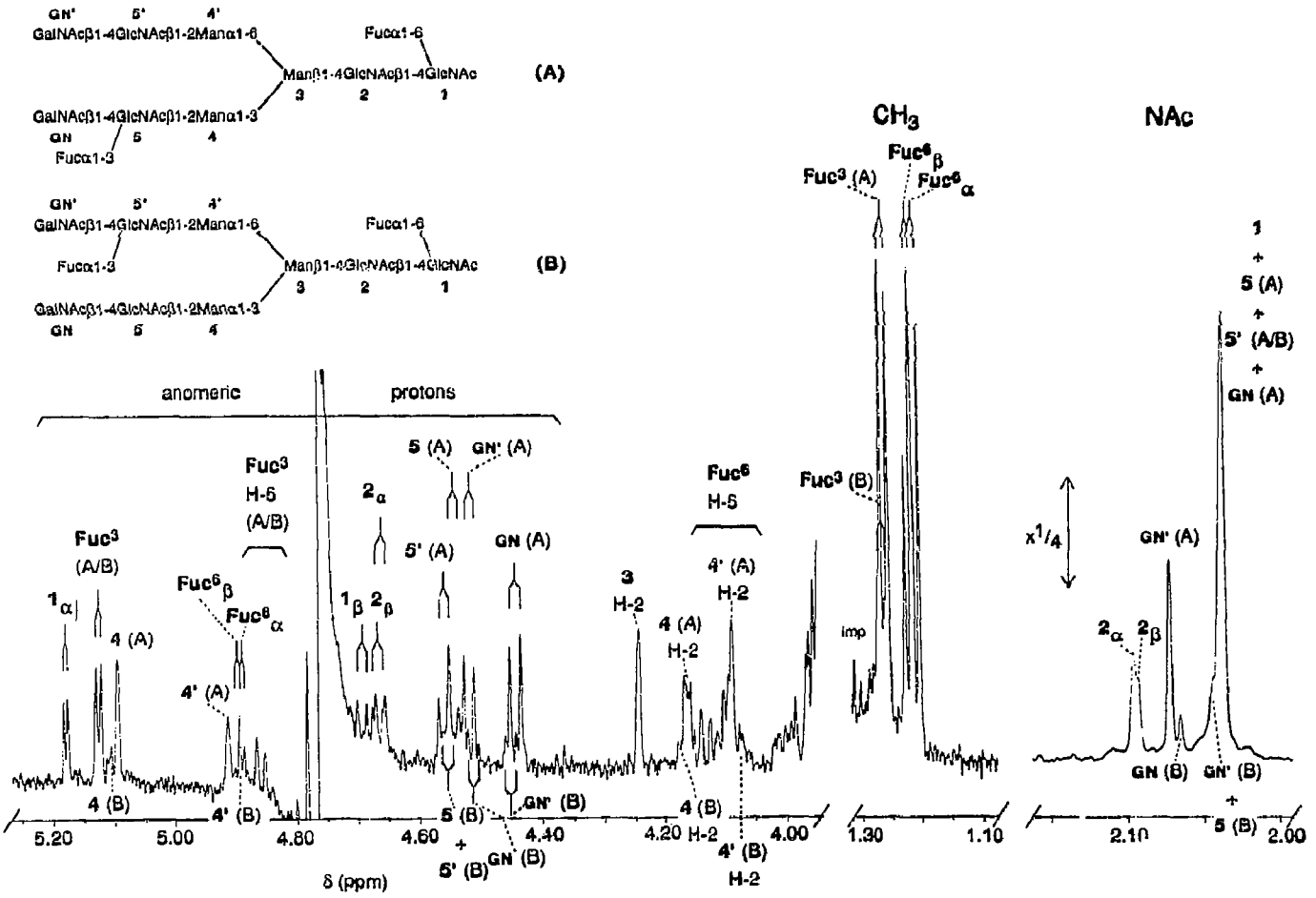
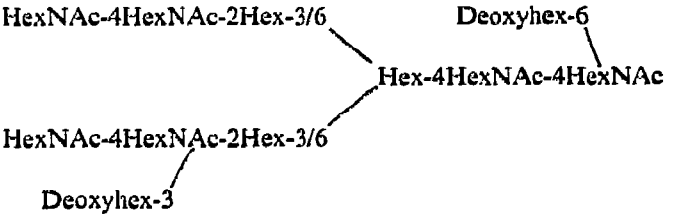


Fig. 2. Structural reporter group signal regions of the resolution-enhanced 500 MHz ^{1}H NMR spectrum at 300 K of fraction N1.4.

branch terminating in a HexNAc-HexNAc-, and the other terminating in a deoxyhex-containing HexNAc-HexNAc- element [17]. The location of this deoxyhexosyl residue is demonstrated by the presence of the additional fragment ion at *m/z* 473 (which arises by β-elimination of deoxyhex from the ion at *m/z* 679) and the absence of an ion at *m/z* 647 (for β-elimination of methanol from *m/z* 679), which indicate that C-3 of the charged hexosamine is substituted with a deoxyhexosyl residue [17]. Moreover, GLC-MS analysis of partially methylated alditol acetates derived from fraction N1.4 showed the presence of a 3,4-substituted HexNAc residue, as well as the derivatives arising from terminal deoxyhex, 2-substituted Hex, 3,6-substituted Hex, terminal HexNAc, 4-substituted HexNAc and 4,6-substituted HexNAc (data not shown). These data indicate the presence of a diantennary oligosaccharide with a conventional core structure, substituted with one antenna which terminates in a HexNAc-4(deoxyhex-3)HexNAc-sequence, and another terminating in HexNAc-4HexNAc-, as follows:



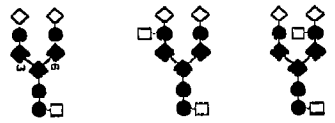
It should be noted that the mass spectrometric data do not allow the site of attachment to the core of the deoxyhex-bearing antenna to be distinguished.

The ^{1}H NMR spectrum of fraction N1.4 (Fig. 2) shows the presence of a major (80%, N1.4A), and a minor (20%, N1.4B) component. The spectral data of N1.4A and N1.4B have similarities with dose of reference compound $GalNAc_2GlcNAc_2Man_3GlcNAc(Fuc)$-GlcNAc (OA; for a detailed structure, see Table I), pointing to the diantennary character of the oligosaccharides, and to the presence of terminal β(1-4)-linked GalNAc. Significant differences exist, which arise from

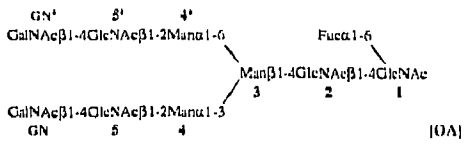
Table I

^{1}H chemical shifts of structural reporter group protons of the constituent monosaccharides of the neutral *N*-linked oligosaccharides OA (reference compound), and N1.4A and N1.4B (derived from human urokinase)

←

Reporter group	Residue	Chemical shift (ppm) in		
				
		OA[a]	N1.4A	N1.4B
H-1	GlcNAc-1α	5.179	5.180	5.180
	GlcNAc-1β	4.694	4.694	4.694
	GlcNAc-$2_α$	4.664	4.664	4.664
	GlcNAc-$2_β$	4.668	4.668	4.668
	Man-4	5.107	5.096	5.109
	Man-4′	4.912	4.914	4.896
	GlcNAc-5	4.553[b]	4.546	4.553
	GlcNAc-5′	4.559	4.561	4.553
	GalNAc-GN	4.514	4.445	4.514
	GalNAc-GN′	4.520	4.521	4.452
	Fuc^3	--	5.127	5.127
	$Fuc_α^6$	4.892	4.892	4.892
	$Fuc_β^6$	4.899	4.900	4.900
H-2	Man-3	4.241	4.245	4.245
	Man-4	4.174	4.164	4.175
	Man-4′	4.090	4.089	4.075
H-5	Fuc^3	--	4.862	4.862
	$Fuc_α^6$	4.097	4.099	4.099
	$Fuc_β^6$	4.130	4.133	4.133
NAc	GlcNAc-1	2.039	2.038	2.038
	GlcNAc-$2_α$	2.096	2.097	2.097
	GlcNAc-$2_β$	2.094	2.094	2.094
	GlcNAc-5	2.044	2.038	2.045
	GlcNAc-5′	2.039	2.038	2.038
	GalNAc-GN	2.066	2.038	2.066
	GalNAc-GN′	2.073	2.073	2.045
CH_3	Fuc^3	--	1.260	1.262
	$Fuc_α^6$	1.208	1.209	1.209
	$Fuc_β^6$	1.220	1.221	1.221

[a] Unpublished data from the authors' laboratory:



[b] GlcNAc-5 H-1 shows an extra virtual coupling.

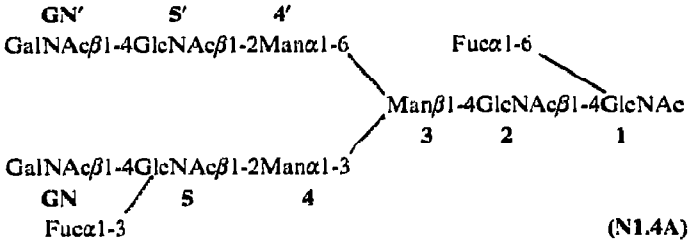
the presence of a Fuc residue located on one of the branches, as concluded from the FAB-MS and linkage analysis data.

Comparison of the 1D ^{1}H NMR data of reference compound, OA, and of component N1.4A demonstrates that the parameters of the Manα(1-6) branch, Man-4′ H-1 and H-2, GlcNAc-5′ H-1 and NAc, and GalNAc-GN′ H-1 and NAc, are identical. The 2D NOESY subspectrum of N1.4A confirms the presence of the GalNAcβ(1-4)GlcNAcβ(1-2)Manα(1-6) branch via inter-residual n.O.e. cross-peaks between GlcNac-5′ H-1 and Man-4′ H-1, and between GlcNAc-5′ H-1 and Man-4′ H-2. Therefore, α(1-3)-linked Fuc (H-1, δ 5.127; $J_{1,2}$ 4.0 Hz) must be present in the Manα(1-3) branch. The structural reporter groups (Table I), characteristic for the GalNAcβ(1-4)[Fucα(1-3)]GlcNAcβ(1-2)-Manα(1-3) branch, were deduced from the 2D NOESY and HOHAHA spectra. The assignment of GlcNAc-5 H-1 at δ 4.546 is based on the observation of inter-residual n.O.e. cross-peaks between this proton and Man-4 H-1 and H-2. The H-1 signal at δ 4.445 ($J_{1,2}$ 84 Hz) belongs to GalNAc-GN, as proven by the typical GalNAc subspectrum on the H-1 track in the 2D HOHAHA spectrum (H-2 and H-3 overlapping at δ 3.72, and H-4 at δ 3.914). The ^{1}H signals of the α(1-3)-linked Fuc residue could be completely traced in the HOHAHA spectrum via cross-peaks of H-4 observed on both the Fuc H-1 and H-6 tracks (H-1 track: H-2, δ 3.669; H-3, δ 3.951; H-4, δ 3.840; H-6 track: H-4, δ 3.840; H-5, δ 4.862). The structural reporter group signals of this residue are similar to those reported for a GalNAcβ(1-4)[Fucα(1-3)]GlcNAcβ(1-3/6) element in the *O*-linked oligosaccharide alditols, N7a and N7b, isolated from sea squirt H-antigen [18]. Finally, the NAc methyl singlet at δ 2.038, having the intensity of 4 methyl signals, comprises not only the NAc methyl signals of GlcNAc-1 and GlcNAc-5′, but also the NAc methyl signals of GlcNAc-5 and GalNAc-GN. It is evident that, going from reference compound OA to N1.4A, the occurrence of the α(1-3)-linked Fuc in the Manα(1-3) branch has considerable upfield shift effects

←

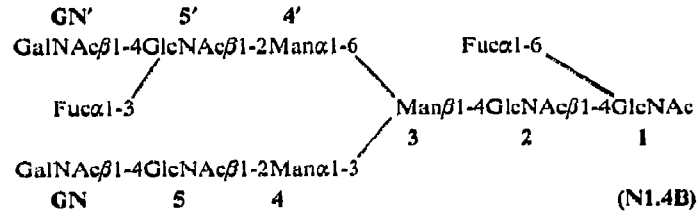
Chemical shifts are given in ppm relative to internal acetone (δ 2.225) in ^2H_2O at 300 K and at p^{2}H 7 [8]. Compounds are represented by short-hand symbolic notation: (□), L-Fuc; (●), D-GlcNAc; (◆), D-Man; (○), D-GalNAc. For numbering of the monosaccharide residues, see text. n.d., not determined. α and β stand for the anomeric configuration of GlcNAc-1. The superscripts 3 and 6 stand for the linkage type of Fuc.

on Man-4 H-1 ($\Delta\delta$ −0.011, Man-4 H-2 ($\Delta\delta$ −0.010), GlcNAc-5 H-1 ($\Delta\delta$ −0.007), GlcNac-5 Nac($\Delta\delta$ −0.006), GalNAc-GN H-1 ($\Delta\delta$ −0.069), and GalNac-GN NAc ($\Delta\delta$ −0.028). In summary, N1.4A has the following structure:



(N1.4A)

Comparison of the 1D ^{1}H NMR data of component N1.4 B with those of OA and N1.4A shows the presence of a GalNacβ(1-4)GlcNAcβ(1-2)Manα(1-3) branch in N1.4B. Evidence for the location of a GalNAcβ(1-4)[Fucα(1-3)]GlcNAc element in the Manα(1-6) branch is based on a similar series of upfield shift effects, as observed for the α(1-3) fucosylation of the Manα(1-3) branch in N1.4A: Man-4' H-1 at δ 4.896 ($\Delta\delta$ −0.016), Man-4' H-2 at δ 4.075 ($\Delta\delta$ −0.015), GlcNAc-5' H-1 at δ 4.553 ($\Delta\delta$ −0.006), GlcNAc-5' NAc at δ 2.038 ($\Delta\delta$ −0.001). GalNAc-GN' H-1 at δ 4.452 ($\Delta\delta$ −0.068), and GalNAc-GN' NAc at δ 2.045 ($\Delta\delta$ −0.028). The α(1-3)-linked Fuc residue is reflected by Fuc H-1 at δ 5.127, H-5 at δ 4.862, and H-6 at δ 1.262. Therefore, N1.4B has the following structure:



(N1.4B)

4. DISCUSSION

The GalNAcβ(1-4)GlcNAcβ sequence has gradually been shown to be a rather common structural element in *N*-linked carbohydrate chains. The presence of Fucα(1-2)Galβ(1-3)GalNAcβ(1-4)GlcNAcβ(1-2) and 3MeGalβ(1-3)GalNAcβ(1-4)GlcNAcβ(1-2) elements were established for *N*-linked carbohydrate chains derived from hemocyanin of the snail, *Lymnaea stagnalis* [19]. The (SO_4^{-4})-GalNacβ(1-4)GlcNAcβ element has been demonstrated to occur in human [20], ovine [21] and bovine [22] lutropins, mouse pro-opiomelanocortin (the precursor of adrenocorticotropin) [23], and human Tamm-Horsfall glycoprotein [13]. The Neu5Acα(2-6)GalNAcβ(1-4)GlcNAcβ(1-2) element is present in the *N*-linked carbohydrate chains of human lutropin [20] and Bowes melanoma t-PA [24], whereas the Neu5Acα(2-3)GalNAcβ(1-4)GlcNAcβ(1-2) sequence has been found in a thrombin-like serine protease, ancrod, from the viper, *Agkistrodon rhodostoma* [25] and in a thrombin-like enzyme, batroxobin, from the snake, *Bothrops atrox moojeni* [26].

In this paper on human urokinase, a novel extension in *N*-linked oligosaccharides is presented, namely GalNAcβ(1-4)[Fucα(1-3)]GlcNAcβ(1-2). This element is present in either the Manα(1-3) branch (80%) or the Manα(1-6) branch (20%) of a diantennary oligosaccharide. Depending on the branch location, the introduction of an α(1-3)-linked Fuc residue gives rise to upfield chemical shift effects for the H-1 and NAc signals of GalNAc-GN or GalNAc-GN', which are larger than those for the H-1 and NAc signals of GlcNAc-5 or GlcNAc-5'. These effects may stem from the closeness in space of the Fuc and GalNAc residues, in an analogous way to that reported for the Galβ(1-4)[Fucα(1-3)]GlcNAc sequence [27], designated Lewis x (Le^x). However, for the terminal GalNAc-containing element no inter-residual n.O.e. contacts could be detected between the Fuc and GalNAc residue like for Le^x.

The Le^x sequence, and its α(2-3)-sialylated form (sialyl-Le^x), are of current interest, since they play a role in cellular adhesion processes, and may be onco-developmentally related compounds [28,29]. In this context, it is noteworthy that u-PA is involved in tissue degradation (e.g. during tumor growth [1]), in which process u-PA receptors are involved.

Since {GalNAcβ(1-4)[Fucα(1-3)]GlcNAcβ(1-3/6)}$_2$-GalNAc-ol and GalNAcβ(1-4)[Fucα(1-3)] GlcNAcβ(1-3)[GalNAcβ(1-4)GlcNAcβ(1-6)] GalNAc-ol from sea squirt H-antigen [18] are allergenic compounds [30], it is interesting to establish whether the GalNAcβ(1-4)-[Fucα(1-3)]GlcNAcβ(1-2)Manα(1-3/6) sequence in urokinase can exhibit a similar feature. In this context it is important to investigate how widely this element found in urokinase is distributed in humans.

Acknowledgements: This investigation was supported by Ciba-Geigy Ltd., by the Netherlands Program for Innovation Oriented Carbohydrate Research (IOP-k), and by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO).

REFERENCES

[1] Danø, K., Andreasen, P.A., Grøndahl-Hansen, J., Kristensen, P., Nielsen, L.S. and Skriver, L. (1985) Adv. Cancer Res. 44, 139–266.
[2] Huber, K., Kirchheimer, J.C., Ermler, D., Bell, C. and Binder, B.R. (1992) Cancer Res. 52, 1717–1720.
[3] Steffens, G.J., Günzler, W.A., Ötting, F., Frankus, E. and Flohé, L. (1982) Hoppe-Seyler's Z. Physiol. Chem. 363, 1043–1058.
[4] McLellan, W.L., Vetterlein, D. and Roblin, R. (1980) FEBS Lett. 115, 181–184.

[5] Kentzer, E.J., Buko, A., Menon, G. and Sarin, K. (1990) Biochem. Biophys. Res. Commun. 171, 401–406.
[6] Buko, A.M., Kentzer, E.J., Petros, A., Menon, G., Zuiderweg, E.R.P. and Sarin, V.K. (1991) Proc. Natl. Acad. Sci. USA 88, 3992–3996.
[7] Damm, J.B.L., Bergwerff, A.A., Hård, K., Kamerling, J.P. and Vliegenthart, J.F.G. (1989) Recl. Trav. Chim. Pays-Bas 108, 351–359.
[8] Vliegenthart, J.F.G., Dorland, L. and Van Halbeek, H. (1983) Adv. Carbohydr. Chem. Biochem. 41, 209–374.
[9] Marion, D., Ikura, M., Tschudin R. and Bax, A. (1989) J. Magn. Reson. 85, 393–399.
[10] Bax, A. and Davis, D.G. (1985) J. Magn. Reson. 65, 355–360.
[11] Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) J. Chem. Phys. 71, 4546–4553.
[12] Marion, D. and Wüthrich, K. (1983) Biochem. Biophys. Res. Commun. 117, 967–974.
[13] Hård, K., Van Zadelhoff, G., Moonen, P., Kamerling, J.P. and Vliegenthart, J.F.G. (1992) Eur. J. Biochem. (in press).
[14] Ciucanu, I. and Kerek, F. (1984) Carbohydr. Res. 131, 209–217.
[15] York, W.S., Darvill, A.G., McNeil, M., Stevenson, T.T. and Albersheim, P. (1986) Methods Enzymol. 118, 3–40.
[16] Dell, A. (1987) Adv. Carbohydr. Chem. Biochem. 45, 19–72.
[17] Fukuda, M., Dell, A., Oates, J. and Fukuda, M. (1984) J. Biol. Chem. 259, 8260–8273.
[18] Ohta, M., Matsuura, F., Kobayashi, Y., Shigeta, S., Ono, K. and Oka, S. (1991) Arch. Biochem. Biophys. 290, 474–483.
[19] Van Kuik, J.A., Sybesma, R.P., Kamerling, J.P., Vliegenthart, J.F.G. and Wood, E.J. (1987) Eur. J. Biochem. 169, 399–411.
[20] Weisshaar, G., Hiyama, J., Renwick, A.G.C. and Nimtz, M. (1991) Eur. J. Biochem. 195, 257–268.
[21] Weisshaar, G., Hiyama, J. and Renwick, A.G.C. (1990) Eur. J. Biochem. 192, 741–751.
[22] Green, E.D., Van Halbeek, H., Boime, I. and Baenziger, J.U. (1985) J. Biol. Chem. 260, 15623–15630.
[23] Skelton, T.P., Kumar, S., Smith, P.L., Beranek, M.C. and Baenziger, J.U. (1992) J. Biol. Chem. 267, 12998–13006.
[24] Chan, A.L., Morris, H.R., Panico, M., Etienne, A.T., Rogers, M.E., Gaffney, P., Creighton-Kempsford, L. and Deli, A. (1991) Glycobiology 1, 173–185.
[25] Pfeiffer, G., Dabrowski, U., Dabrowski, J., Stirm, S., Strube, K.-H. and Geyer, R. (1992) Eur. J. Biochem. 205, 961–978.
[26] Tanaka, N., Nakada, H., Itoh, N., Mizuno, Y., Takanishi, M., Kawasaki, T., Tate, S.-i., Inagaki, F. and Yamashina, I. (1992) J. Biochem. 112, 68–74.
[27] Wormald, M.R., Edge, C.J. and Dwek, R.A. (1991) Biochem. Biophys. Res. Commun. 180, 1214–1221.
[28] Feizi, T. (1991) Trends Biochem. Sci. 16, 84–86.
[29] Springer, A.T. and Lasky, L.A. (1991) Nature 349, 196–197.
[30] Ohta, M., Shigeta, S., Ono, K., Takao, T., Shimonishi, Y. and Oka, S. (1989) Arch. Biochem. Biophys. 275, 151–165.

NOTE ADDED IN PROOF

After the submission of this manuscript, a paper dealing with the GalNAcβ(1-4)[Fucα(1-3)]GlcNAc element in *N*-linked chains of *Schistosoma mansoni* glycoproteins appeared [Scrivatsan, J. et al. (1992) Glycobiology 2, 445–452].