# Isolation and structural characterization of the smaller-size oligosaccharides from desialylated human κ-casein

## Establishment of a novel type of core for a mucin-type carbohydrate chain

H. van Halbeek, J.F.G. Vliegenthart, A.-M. Fiat* and P. Jollès*

*Department of Bio-Organic Chemistry, University of Utrecht, Croesestraat 79, NL-3522 AD Utrecht, The Netherlands* and **Laboratoire des Protéines, Université de Paris V, 45 rue des Saints-Pères, F-75270 Paris Cedex 06, France*

Received 7 May 1985

Alkaline borohydride reductive cleavage (β-elimination) of desialylated human κ-caseinoglycopeptide resulted in the release of a series of oligosaccharides. The smaller-size compounds among them were purified to virtual homogeneity by gel filtration followed by high-performance liquid chromatography. The structures of 9 oligosaccharides were determined by $^1$H-NMR spectroscopy in conjunction with sugar analysis. The tetrasaccharide Galβ(1→3)[Galβ(1→4)GlcNAcβ(1→6)]GalNAc-ol and various partial structures thereof were characterized. Notably, the disaccharide GlcNAcβ(1→6)GalNAc-ol and the trisaccharide Galβ(1→4)GlcNAcβ(1→6)GalNAc-ol were identified; they represent a novel type of core structure for mucin-type carbohydrate chains, namely a peptide-linked GalNAc that is mono-substituted at C-6. In addition, some oligosaccharides ending in GlcNAc-ol could be characterized. Their possible origin is discussed.

*κ-Casein* *Carbohydrate moiety* *HPLC* *$^1$H-NMR spectroscopy* *(Human)*

## 1. INTRODUCTION

The casein complex involved in the milk-clotting process contains among several proteins only one glycoprotein, namely, κ-casein. During the primary phase of this process, para-κ-casein is formed as an insoluble part, while a soluble fraction, κ-caseinoglycopeptide (κ-CGP) is released. The latter contains the complete carbohydrate moiety of κ-casein [1]. Only one prosthetic sugar group was characterized in bovine κ-casein from mature milk whereas two such groups were identified in bovine colostrum [2]. The sugar part was found to be microheterogeneous; the structures of 3 oligosaccharides from bovine mature milk κ-casein have been established in the β-eliminative cleavage product [3,4], namely, NeuAcα(2→3)-Galβ(1 → 3)GalNAc - ol, Galβ(1 → 3)[NeuAcα - (2→6)]GalNAc-ol and NeuAcα(2→3)Galβ(1→3)-[NeuAcα(2→6)]GalNAc-ol. In κ-casein from cow colostrum, 6 oligosaccharides could be characterized possessing GlcNAc as typical extra constituent [5–8].

In human κ-CGP the presence of several (up to 10) prosthetic sugar groups has been demonstrated [9]. All the sugar-peptide linkages are *O*-glycosidic. The carbohydrate content of human κ-CGP is around 55% whereas cow κ-CGP contains only 10% sugar [9]. Not only Gal, GalNAc and NeuAc, but also Fuc and GlcNAc were identified as constituent monosaccharides.

This paper deals with the preparation and the structural characterization of the smaller-size oligosaccharide-alditols isolated from desialylated human κ-CGP after alkaline borohydride reductive cleavage.

*Abbreviations:* κ-CGP, κ-caseinoglycopeptide; HPLC, high-performance liquid chromatography

*Published by Elsevier Science Publishers B.V. (Biomedical Division)*
00145793/85/$3.30 

## 2. MATERIALS AND METHODS

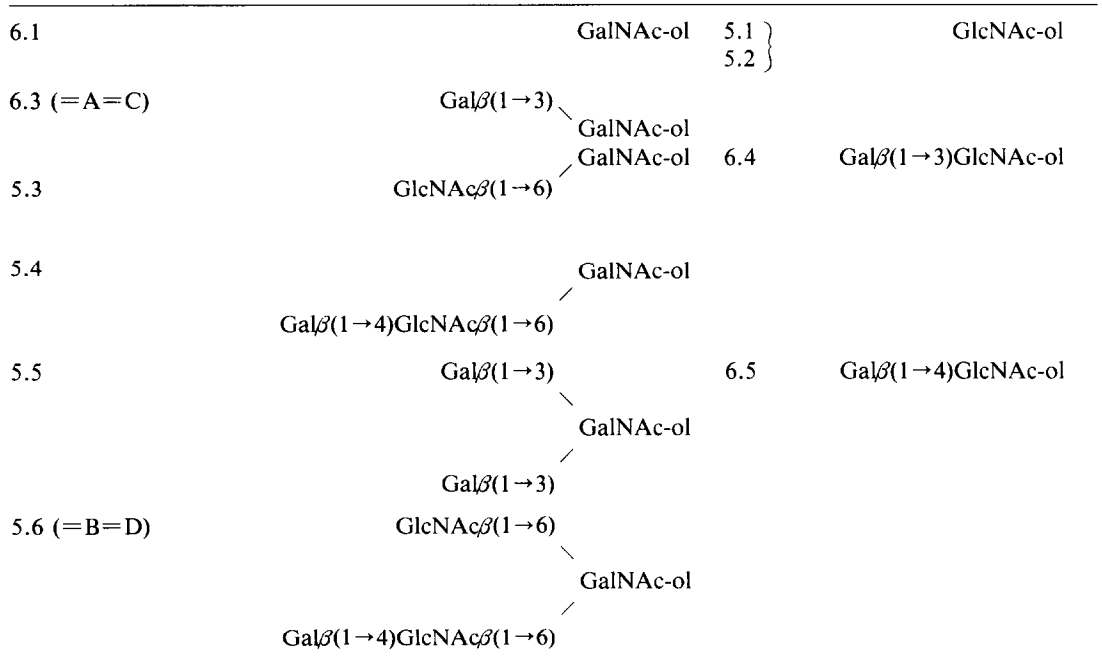
Casein from fresh, pooled human milk was prepared according to Alais and Jollès [10]. Human $\kappa$-CGP soluble in 12% trichloroacetic acid was obtained from the whole casein after chymosin digestion and chromatography on DEAE-Sephadex A-25 according to Fiat et al. [9]. Human $\kappa$-CGP was digested successively by neuraminidase, trypsin and pronase, filtered on Sephadex G-25 and finally chromatographed on DEAE-Sephadex A-25 according to [9].

After shortening of the peptide moieties, the $\kappa$-CGP was treated with alkaline borohydride (0.05 M NaOH and 1.0 M $NaBH_4$) for 16 h at 50°C under nitrogen in the dark [11]. After desalting on Dowex 50 W-X 2 ($H^+$) with 2 mM formic acid as eluent and washing with methanol, the sugar moieties were isolated by filtration on Bio-Gel P-4 (250 × 0.9 cm) with water as eluent. Sugar-positive fractions, stained with the orcinol/sulfuric acid reagent, were further purified either by descending paper chromatography (Whatman no.1) for 18 h in the solvent system ethyl acetate/pyridine/water/acetic acid (5:5:3:1, by vol.) (for the oligosaccharides obtained from short glycopeptides) or by HPLC (for those obtained directly from asialo $\kappa$-CGP).

HPLC of the oligosaccharides was conducted on a Waters chromatograph model ALC/GPC 204 equipped with a MicroPak AX-5 column (30 × 0.4 cm) (Varian). Elution was performed with a linear gradient of acetonitrile-water (85:15–60:40, by vol.) for 90 min at room temperature and a flow rate of 1 ml/min. All solvents were degassed by sonication before application. Sugars were detected at 200 nm.

Sugar analysis of the oligosaccharides was performed by gas-liquid chromatography after methanolysis and trimethylsilylation as in [12]. $^1$H-NMR spectroscopy of $D_2O$ solutions of the oligosaccharides was conducted at 500 MHz, using a Bruker WM-500 spectrometer (SON-NMR facility, Nijmegen, The Netherlands). Further experimental details have been described [5,13].

Scheme 1




Structure of 9 core oligosaccharides that were identified in the $\beta$-elimination product of desialylated human $\kappa$-caseinoglycopeptide

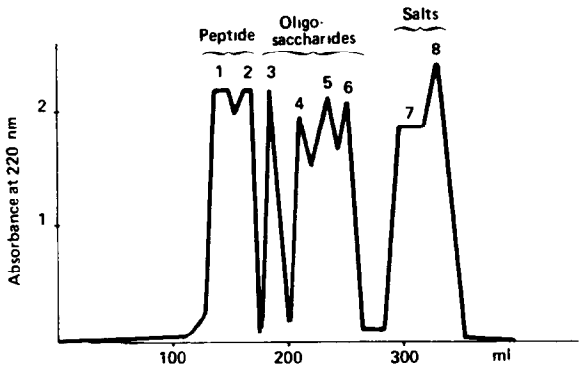



Fig.1. Elution profile of filtration on Bio-Gel P-4 (250 × 0.9 cm) of desialylated human $\varkappa$-caseinoglycopeptide reduced by alkaline borohydride treatment. Water was used as eluent.

## 3. RESULTS

Enzymic digestion of asialo human $\varkappa$-CGP afforded two short glycopeptides designated 1a (amino acid residues 45–65) and 2 (residues 21–37) according to Fiat et al. [9]. These were submitted to alkaline borohydride treatment, filtration on Bio-Gel P-4 and paper chromatography. From glycopeptide 1a, two major oligosaccharides, denoted A ($R_{Gal}$ = 1.0) and B ($R_{Gal}$ = 0.72), were purified. From glycopeptide 2, two oligosaccharides with virtually the same $R_{Gal}$ values as A and B, namely, C ($R_{Gal}$ = 1.0) and D ($R_{Gal}$ = 0.77) were obtained. The structures of A to D were established by 500 MHz $^1$H-NMR spectroscopy (scheme 1).

Alternatively, desialylated human $\varkappa$-CGP (75 mg) was subjected, as such, to alkaline borohydride treatment, and subsequently passed over Bio-Gel P-4. Eight fractions were obtained (fig.1); 1 and 2 contained peptides, 7 and 8 contained just salts, but fractions 3–6 consisted of carbohydrate material. The smaller-size P-4 fractions (5 and 6) were further fractionated by HPLC. The elution profiles were unexpectedly complex. Fraction 5 gave rise to 6 major peaks (denoted 5.1–5.6) (fig.2), and fraction 6 to 5 main peaks (6.1–6.5) (fig.3). The purity of each of these fractions was checked by rechromatography on HPLC. If necessary, a second preparative HPLC run was carried out. In table 1, we report the molar carbohydrate composition of the various HPLC-separated oligosaccharide fractions. Tables 2 and 3 list the $^1$H-NMR characteristics of these fractions, acquired for $D_2O$ solutions at 500 MHz.

Carbohydrate analysis revealed (table 1) that the HPLC-separated subfractions 5.1–5.6 and 6.1–6.5 contained either GalNAc-ol or GlcNAc-ol as the principal reduced monosaccharide. Only in fraction 6.2 were these monosaccharide-alditols both present in considerable amounts (see footnote to table 1). On the basis of the type of reduced monosaccharide, the remaining 10 oligosaccharide-alditols were classified into two groups. The $^1$H-NMR parameters of the compounds ending in GalNAc-ol (6.1, 6.3, 5.3–5.6) have been compiled in table 2, those of the GlcNAc-ol-containing ones (5.1, 5.2, 6.4 and 6.5) in table 3.

Fractions 6.1, 6.3, 5.5 and 5.6 could be readily identified by combination of $^1$H-NMR (table 2)

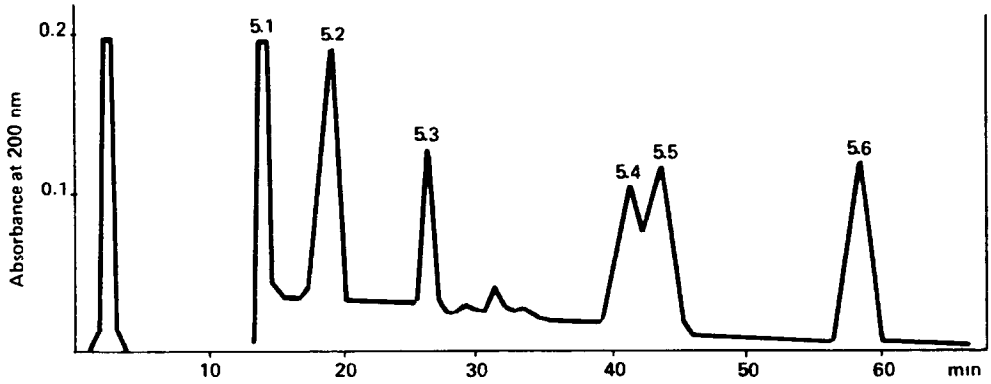



Fig.2. HPLC of Bio-Gel P-4 fraction 5 on a MicroPak AX-5 column (30 × 0.4 cm) (Varian). Elution conditions: a 90-min linear gradient of 85:15–60:40 (v/v) acetonitrile/water. Absorption was detected at 200 nm.

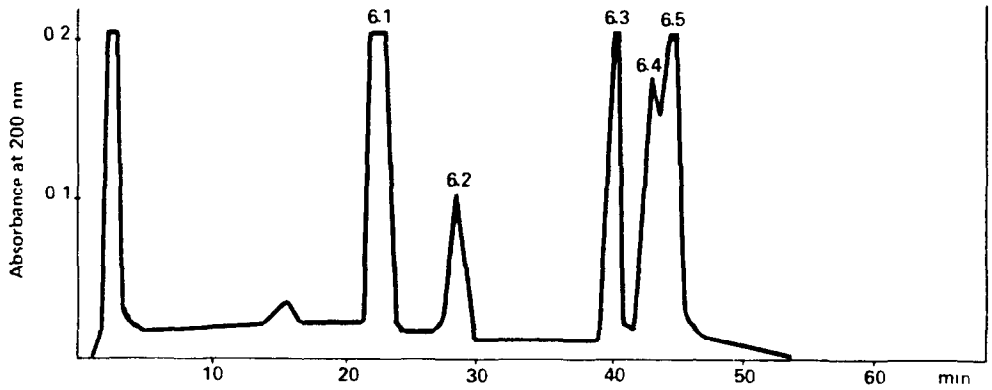



Fig.3. HPLC of Bio-Gel P-4 fraction 6 on a MicroPak AX-5 column (30 × 0.4 cm) (Varian). Elution conditions: a 90-min linear gradient of 85:15–60:40 (v/v) acetonitrile/water. Absorption was detected at 200 nm.

and sugar analysis (table 1) as mono-, di-, tri- and tetrasaccharides with the structures GalNAc-ol, Gal$\beta$(1→3)GalNAc-ol, Gal$\beta$(1→3)[GlcNAc$\beta$(1→6)]GalNAc-ol and Gal$\beta$(1→3)[Gal$\beta$(1→4)GlcNAc$\beta$(1→6)]GalNAc-ol, respectively. The $^1$H-NMR features of each of these fractions match exactly those reported for the same compounds obtained from other sources [4,14,15].

Table 1

Molar carbohydrate composition of the HPLC-separated smaller-size oligosaccharide-alditols, derived from desialylated human $\kappa$-caseinoglycopeptide

| HPLC fraction | Molar ratio[a] of monosaccharides | | | |
|---|---|---|---|---|
| | Gal | GlcNAc | GlcNAc-ol | GalNAc-ol |
| 5.1[b] | + | – | 1.0 | – |
| 5.2 | – | – | 1.0 | – |
| 5.3 | 0.1 | 0.8 | + | 1.0 |
| 5.4 | 1.1 | 0.8 | – | 1.0 |
| 5.5 | 1.1 | 0.9 | – | 1.0 |
| 5.6 | 2.1 | 0.9 | – | 1.0 |
| 6.1 | – | – | – | 1.0 |
| 6.2[bc] | 0.2 | – | 4.0 | 1.0 |
| 6.3 | 1.0 | – | – | 1.0 |
| 6.4[b] | 0.8 | – | 1.0 | + |
| 6.5 | 1.0 | + | 1.0 | + |

[a] Molar ratios were calculated on the basis of one residue of either GalNAc-ol or GlcNAc-ol per molecule

[b] Also contains Gal-ol in detectable amounts

[c] Found to be a mixture of two or three monosaccharide-alditols

+, less than 0.1; –, not detectable at all

The $^1$H-NMR parameters of fractions 5.3 and 5.4 had not been observed before. Carbohydrate analysis (table 1) suggests that a disaccharide and a trisaccharide are concerned, respectively, consisting of GlcNAc and GalNAc-ol (5.3) and of Gal, GlcNAc and GalNAc-ol (5.4). The type of linkage between GlcNAc and GalNAc-ol in 5.3 could be identified by $^1$H-NMR spectroscopy as $\beta$(1→6). The substitution site at GalNAc-ol could be deduced from the set of chemical shifts observed for GalNAc-ol H-2 and H-5 ($\delta$ 4.24 and $\delta$ 4.02, respectively). These are similar to the values reported for the disaccharide NeuAc$\alpha$(2→6)GalNAc-ol [4], and therefore (cf. [16]) proposed to be characteristic of monosubstitution of GalNAc-ol at C-6. The $\beta$-configuration of the linkage is evident from the chemical shift of GlcNAc H-1, in conjunction with its coupling constant ($J_{1,2}$ = 8.2 Hz).

Compound 5.4 can be conceived as an extension of the 5.3 disaccharide by an additional Gal residue. The core type of 5.4 is the same as that of 5.3, as is evident from the set of chemical shifts of GalNAc-ol H-2 and H-5 (see table 2). This implies, that the Gal residue must be attached to GlcNAc. From the chemical shift of Gal H-1 ($\delta$ 4.470, $J_{1,2}$ = 7.8 Hz), as well as from the effect observed on $\delta$ H-1 of GlcNAc ($\Delta\delta$ 0.022) and on NAc of GlcNAc ($\Delta\delta$ 0.002) upon comparison of 5.4 to 5.3, it can be concluded [13] that the linkage between Gal and GlcNAc is $\beta$(1→4).

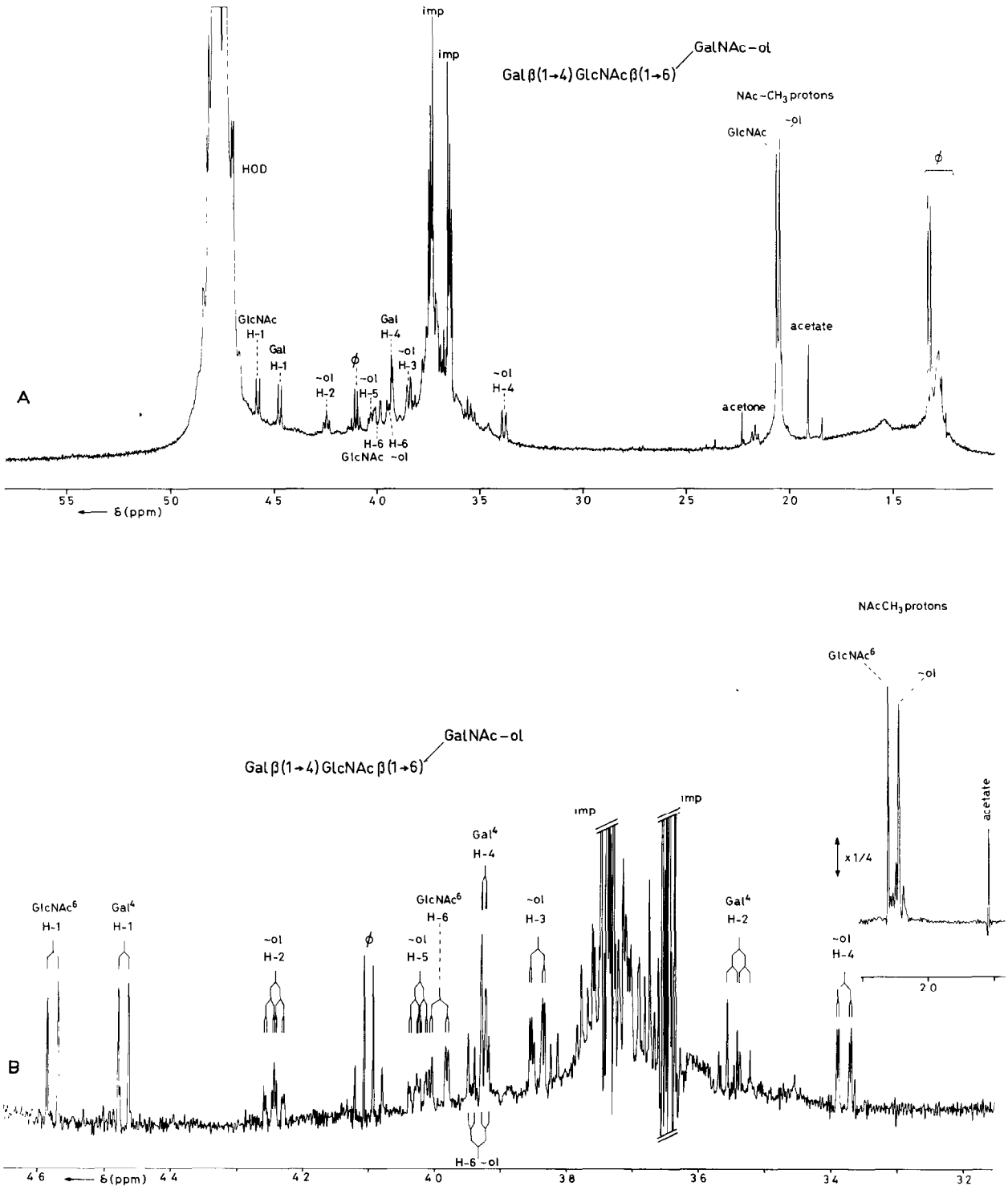



Fig.4. 500 MHz $^1$H-NMR spectrum ($D_2O$, 27°C) of HPLC-fraction 5.4, obtained from human $\kappa$-caseinoglycopeptide, containing the linear trisaccharide-alditol with the indicated, novel structure.

Table 2

Relevant $^1$H-NMR characteristics of constituent monosaccharides for the smaller-size oligosaccharides derived from desialylated human $\varkappa$-caseinoglycopeptide, ending in GalNAc-ol

| Residue[b] | Reporter group | Chemical shift[a] in 6.1 | 6.3 | 5.3 | 5.4 | 5.5 | 5.6 |
|---|---|---|---|---|---|---|---|
| GalNAc-ol | H-2 | 4.253 | 4.392 | 4.243 | 4.243 | 4.392 | 4.392 |
| | H-3 | 3.848 | 4.061 | 3.842 | 3.843 | 4.060 | 4.057 |
| | H-4 | 3.391 | 3.506 | 3.382 | 3.379 | 3.467 | 3.463 |
| | H-5 | 3.929 | 4.194 | 4.022 | 4.025 | 4.278 | 4.281 |
| | H-6 | 3.67 | 3.69 | 3.93 | 3.932 | 3.933 | 3.930 |
| | NAc | 2.057 | 2.051 | 2.047 | 2.045 | 2.065 | 2.067 |
| Gal$^3$ | H-1 | – | 4.476 | – | – | 4.464 | 4.464 |
| | H-2 | – | 3.564 | – | – | 3.542 | 3.538 |
| | H-4 | – | 3.901 | – | – | 3.899 | 3.901 |
| GlcNAc$^6$ | H-1 | – | – | 4.555 | 4.577 | 4.537 | 4.559 |
| | H-6 | – | – | 3.93 | 3.993 | 3.933 | 3.996 |
| | NAc | – | – | 2.059 | 2.061 | 2.065 | 2.064 |
| Gal$^4$ | H-1 | – | – | – | 4.470 | – | 4.469 |
| | H-2 | – | – | – | 3.537 | – | 3.540 |
| | H-4 | – | – | – | 3.925 | – | 3.924 |

[a] Data were acquired at 500 MHz, for $D_2O$ solutions at 27°C. Chemical shifts are in ppm relative to DSS (internal acetone taken at $\delta$ 2.225 ppm)

[b] A superscript at the name of a sugar residue indicates to which position of the adjacent monosaccharide it is linked

(◇) GalNAc-ol, (■—) Gal, (●—) GlcNAc

The carbohydrate analyses (table 1) and $^1$H-NMR features (table 3) of fractions 5.1, 5.2, 6.4 and 6.5 reveal that the former two fractions contain merely GlcNAc-ol, while the latter two contain disaccharides Gal-GlcNAc-ol. From comparison of the $^1$H-NMR data of 6.4 and 6.5 with those published by Hounsell et al. for similar disaccharides [17], it could be readily concluded that 6.4 consists of Gal$\beta$(1→3)GlcNAc-ol while 6.5 is the isomer Gal$\beta$(1→4)GlcNAc-ol. In this respect, it should be noted that, in particular, the chemical shifts of H-2 of GlcNAc-ol ($\delta$ 4.25 vs $\delta$ 4.30) and NAc of GlcNAc-ol ($\delta$ 2.034 vs $\delta$ 2.053) are of discriminative value. The structures of the HPLC-separated oligosaccharides from human $\varkappa$-CGP have been compiled in scheme 1.

## 4. DISCUSSION

The classical approach (cf. [3,4]) of isolation and purification of *O*-glycosidic oligosaccharides, that is, paper chromatography performed on the Bio-Gel P-4 fractionated $\beta$-elimination product of human $\varkappa$-CGP, afforded just two types of core structure, namely, Gal$\beta$(1→3)GalNAc-ol and Gal$\beta$(1→3)[Gal$\beta$(1→4)GlcNAc$\beta$(1→6)]-

Table 3

Relevant $^1$H-NMR characteristics of constituent monosaccharides for the smaller-size oligosaccharides derived from desialylated human $\kappa$-caseinoglycopeptide, ending in GlcNAc-ol

| Residue[b] | Reporter group | Chemical shift[a] in ⊙ 5.1; 5.2 | ⊙(3)— 6.4 | ⊙(4)— 6.5 |
|---|---|---|---|---|
| GlcNAc-ol | H-2 | 4.072 | 4.249 | 4.303 |
| | H-3 | 3.956 | 4.166 | n.d. |
| | H-5 | 3.748 | n.d. | n.d. |
| | H-6 | 3.816 | 3.842 | n.d. |
| | NAc | 2.043 | 2.034 | 2.053 |
| Gal[3] | H-1 | – | 4.498 | – |
| | H-2 | – | 3.565 | – |
| | H-4 | – | 3.902 | – |
| Gal[4] | H-1 | – | – | 4.499 |
| | H-2 | – | – | 3.565 |
| | H-4 | – | – | 3.910 |

[a] Data were acquired at 500 MHz, for $D_2O$ solutions at 27°C. Chemical shifts are expressed in ppm relative to DSS (internal acetone taken at $\delta$ 2.225 ppm)

[b] A superscript at the name of a sugar residue indicates to which position of the adjacent monosaccharide it is glycosidically linked

(⊙) GlcNAc-ol, (■—) Gal; n.d., value could not be determined

GalNAc-ol. It is impressive to notice that combination of HPLC (after Bio-Gel P-4 filtration) and $^1$H-NMR spectroscopy succeeded in elucidation of another 4 (although less abundant) types of core oligosaccharides. The structures resulting from the paper chromatographic working-up procedure were found again (6.3 and 5.6, see scheme 1). Also partial structures thereof, lacking the Gal$\beta$(1→3) and/or Gal$\beta$(1→4) residues, respectively, were demonstrated to occur. In particular, the structures 5.3 and 5.4 containing the GlcNAc$\beta$(1→6)GalNAc-ol element draw the attention, because this is a novel type of *O*-glycosidic core structure (the corresponding trisaccharide has been found to occur in human meconium [18]). Moreover, the occurrence of this structure suggests that in human an alternative pathway for the biosynthesis of mucin-type chains exists; current views are that the GlcNAc$\beta$(1→6) transferase does not recognize GalNAc$\alpha$(1→0)Ser/Thr as a substrate, until GalNAc has been substituted at C-3 by a Gal or a GlcNAc residue [19,20].

In addition to the wanted $\beta$-elimination products ending in GalNAc-ol, a number of smaller-size oligosaccharides were found to end in GlcNAc-ol. Regarding their origin, two possibilities have to be considered. Either, they stem from degradation (peeling reactions [21]) of regular GalNAc-containing chains, or they represent a rather uncommon sugar/peptide linkage, namely, GlcNAc(1→0)Ser/Thr (cf. [22]). HPLC fractionation and $^1$H-NMR analysis of the larger oligosaccharides present in Bio-Gel P-4 fractions 3 and 4, may be helpful to shed more light on this aspect.

## ACKNOWLEDGEMENTS

The authors thank Mr G.J. Gerwig and Mrs J. Chevan for skilful technical assistance with carbohydrate analyses and oligosaccharide preparations, respectively. This investigation was supported by the Netherlands Foundation for Chemical Research (SON/ZWO), the Netherlands Cancer Foundation (KWF) (grant UUKC 83-13), and the CNRS (Unité de Recherches no.102).

## REFERENCES

[1] Jollès, P. (1972) in: Glycoproteins (Gottschalk, A. ed.) BBA Library, vol.5B, pp.782–809, Elsevier, Amsterdam, New York.

[2] Fiat, A.-M., Jollès, J., Loucheux-Lefebvre, M.-H., Alais, C. and Jollès, P. (1981) Z. Physiol. Chem. 362, 1447–1454.

[3] Fournet, B., Fiat, A.-M., Alais, C. and Jollès, P. (1979) Biochim. Biophys. Acta 576, 339–346.

[4] Van Halbeek, H., Dorland, L., Vliegenthart, J.F.G., Fiat, A.-M. and Jollès, P. (1980) Biochim. Biophys. Acta 623, 295–300.

[5] Van Halbeek, H., Dorland, L., Vliegenthart, J.F.G., Fiat, A.-M. and Jollès, P. (1981) FEBS Lett. 133, 45–50.

[6] Saito, T., Itoh, T., Adachi, S., Suzuki, T. and Usui, T. (1981) Biochim. Biophys. Acta 678, 257–267.

[7] Saito, T., Itoh, T. and Adachi, S. (1981) Biochim. Biophys. Acta 673, 487–494.
[8] Saito, T., Itoh, T., Adachi, S., Suzuki, T. and Usui, T. (1982) Biochim. Biophys. Acta 719, 309–317.
[9] Fiat, A.-M., Jollès, J., Aubert, J.-P., Loucheux-Lefebvre, M.-H. and Jollès, P. (1980) Eur. J. Biochem. 111, 333–339.
[10] Alais, C. and Jollès, P. (1969) J. Chromatogr. 44, 573–580.
[11] Carlson, D.M. (1961) J. Biol. Chem. 241, 2984–2986.
[12] Kamerling, J.P. and Vliegenthart, J.F.G. (1982) Cell Biol. Monogr. 10, 95–125.
[13] Vliegenthart, J.F.G., Dorland, L. and Van Halbeek, H. (1983) Adv. Carbohydr. Chem. Biochem. 41, 209–374.
[14] Van Halbeek, H., Dorland, L., Haverkamp, J., Veldink, G.A., Vliegenthart, J.F.G., Fournet, B., Ricart, G., Montreuil, J., Gathmann, W.D. and Aminoff, D. (1981) Eur. J. Biochem. 118, 487–495.
[15] Van Halbeek, H., Dorland, L., Vliegenthart, J.F.G., Hull, W.E., Lamblin, G., Lhermitte, M., Boersma, A. and Roussel, P. (1982) Eur. J. Biochem. 127, 7–20.
[16] Van Halbeek, H. (1984) Biochem. Soc. Trans. 12, 601–605.
[17] Hounsell, E.F., Wright, D.J., Donald, A.S.R. and Feeney, J. (1984) Biochem. J. 223, 129–143.
[18] Hounsell, E.F., Wright, D.J., Donald, A.S.R. and Feeney, J. (1984) in: Abstr. XIIth Int. Carbohydr. Symp., Utrecht (Vliegenthart, J.F.G. et al. eds) p.507, Vonk Publ., Zeist.
[19] Sadler, J.E. (1984) in: Biology of Carbohydrates, vol.II (Ginsburg, V. and Robbins, P.W. eds) pp.199–288, Academic Press, New York.
[20] Schachter, H. and Williams, D. (1982) in: Mucus in Health and Disease, vol.II (Chantler, E.N. et al. eds) pp.3–28, Plenum, New York.
[21] Lloyd, K.O., Kabat, E.A. and Licerio, E. (1968) Biochemistry 7, 2976–2990.
[22] Torres, C.R. and Hart, G.W. (1984) J. Biol. Chem. 259, 3308–3317.