Glycoconjugate J (1985) 2:235—253

^{1}H-NMR Spectroscopic Characterization of Dansyl Glyco-asparagines Derived from Hen Egg White Glycoproteins

HERMAN VAN HALBEEK[1], JOHANNES F G VLIEGENTHART[1]*, HITOO IWASE[2, 3], SU-CHEN LI[2] and YU-TEH LI[2]

[1]*Department of Bio-Organic Chemistry, University of Utrecht, Croesestraat 79, NL-3522 AD Utrecht, The Netherlands*
[2]*Department of Biochemistry, Delta Regional Primate Research Center, Tulane University, New Orleans, Louisiana 70112, USA*
[3]*Present address: Department of Biochemistry, School of Medicine, Kitasato University, 1-15-1 Sagamihara, Kanagawa 228, Japan*

Received July 15, 1985.

Key words: ^{1}H-NMR spectroscopy, intersected hybrid-type glycans, intersected multi-antennary glycans, dansyl derivative, ovalbumin, ovomucoid, ovotransferrin

Dansyl glyco-asparagines were prepared from a partially fractionated mixture of hen egg white glycoproteins. Reverse-phase high performance liquid chromatography (HPLC) on a silica-based octadecyl column yielded ten such derivatives in a virtually pure state. The detailed structures of the glyco-asparagines were identified by 500-MHz ^{1}H nuclear magnetic resonance (NMR) spectroscopy. Two of them were found to be of the oligomannoside *N*-type, four were of the intersected-hybrid *N*-type and another four were of the intersected multi-antennary *N*-type. In monogalactosylated, intersected structures the galactose residue was proved by ^{1}H-NMR to be attached in β(1-4)-linkage to the GlcNAcβ1-4Manα1-3 branch.

Dansyl glyco-asparagines turned out to be suitable derivatives for ^{1}H-NMR spectroscopic analysis. The combination of HPLC and high-resolution NMR spectroscopy of such derivatives proved to be a powerful technique in studying the (micro-)heterogeneity of sugar chains in glycoproteins.

Abbreviations: dns, dansyl (5-dimethylaminonaphthalene-1-sulfonyl); ODS, octadecyl-silica; WEFT, water-eliminating Fourier transform; DSS, sodium 4,4-dimethyl-4-silapentane-1-sulfonate; OVA, ovalbumin; OVM, ovomucoid; OVT, ovotransferrin.

*Author for correspondence

Complete elucidation of the structure of the carbohydrate chain(s) of a glycoprotein is a prerequisite for understanding its biological function. One of the most perplexing features of glycoproteins is the heterogeneity of their carbohydrate chains. Although quite often readily detectable, the detailed characterization of the heterogeneity of sugar chains in glycoproteins is largely restricted by the fractionation and identification techniques available. Hen egg white glycoproteins are known to be extremely heterogeneous in their carbohydrate chains [1, 2]. We now report the fractionation and characterization of their glycans without laborious pre-purification steps (cf. [3-20]). We have used reverse-phase HPLC to fractionate the dansyl derivatives of the glyco-asparagines obtainable from ovalbumin, ovomucoid and ovotransferrin. The separation by HPLC is based on hydrophobic interaction of dansyl groups with silica-bonded octadecyl (ODS) groups [21]. The structures of the dansyl glyco-asparagines isolated by reverse-phase HPLC were subsequently identified by 500-MHz ^{1}H-NMR spectroscopy.

Materials and Methods

Materials

Glycoproteins prepared from pooled hen egg whites according to [22] were partially fractionated as described [23], yielding a mixture of ovalbumin (OVA) and ovomucoid (OVM), and ovotransferrin (OVT). Pronase® was a product from Calbiochem (San Diego, CA, USA), 5-dimethylamino-naphthalene-1-sulfonylchloride (dansyl chloride) from Pierce (Rockford, IL, USA), and acetonitrile from Mallinckrodt (Grossostheim, W. Germany) . All solvents used for HPLC were filtered through a Millipore filter (pore size 0.45μ).

Preparation of Dansyl Glyco-asparagines

Glyco-asparagines from the partially purified OVA/OVM mixture and from OVT were prepared by repeated Pronase digestion. The OVA/OVM compounds were separated into five fractions (A-E) by Dowex-50 anion-exchange chromatography, as described [24]. Dansylation of glyco-asparagines was carried out according to the method of Gray [25].

Fractionation of Dansyl Glyco-asparagines by Reverse-phase HPLC

The HPLC system consisted of an Altex (Rochester, N.Y., USA) Model 110A solvent delivery system, an LKB (Bromma, Sweden) Model 11300 Ultragrad solvent gradient maker, an Altex Ultrasphere ODS reverse-phase column (5 μ; 25 cm × 1 cm i.d.) protected by a Bio-Rad (Richmond, CA, USA) Microguard column (5 cm × 4 mm i.d.) packed with Bio-Rad Bio-Sil ODS reverse-phase hydrocarbon (10 μ), and the following components from Perkin Elmer: a Model-7105 valve sample injector, a Model-650 fluorescence monitor, and a Model-56 recorder. For analytical purposes 2-5 nmol of dansyl glyco-asparagines in 5 μl water, and for preparative purposes, about 500 nmol in 50 μl water were applied to the column. The column was eluted with a linear gradient from water to 7.5% acetonitrile

containing 25 mM sodium borate buffer, pH 7.0 at 3.0 ml/min for 8 h. The effluent was monitored by a fluorescence detector (excitation, 313 nm; emission, 550 nm) and collected when appropriate. Non-dansylated glycopeptides are not retained by the column under the applied conditions.

500-MHz ^{1}H-NMR Spectroscopy

Prior to ^{1}H-NMR spectroscopic analysis, the dansyl glyco-asparagine fractions were repeatedly treated with ^2H_2O at room temperature. After each exchange treatment, the materials were lyophilized. Finally, each sample was redissolved in 400 μl ^2H_2O (99.96 atom% ^{2}H, Aldrich, Milwaukee, WI, USA) for examination in a 5 mm tube (528 PP, Wilmad, Buena, NJ, USA). 500-MHz ^{1}H-NMR spectroscopy was performed on a Bruker WM-500 spectrometer (SON hf-NMR facility, Department of Biophysics, Nijmegen University, The Netherlands), operating in the Fourier transform mode and equipped with a Bruker Aspect-2000 computer [26, 27]. The spectral width used was 5 kHz, which with a data memory of 32K gave a digital resolution of 0.3 Hz. Pulses (90°, 12.5 μs) were used. For solvent-peak suppression, a WEFT-pulse sequence (180°, composite pulse-delay-90°, pulse-acquisition) was applied [28]. Resolution enhancement of the spectra was achieved by the Lorentzian-to-Gaussian transformation [26, 27] routine, implemented in the Bruker software package. The probe temperature was 27°C, and was kept constant within 0.1°. Chemical shifts (δ) are expressed in ppm downfield from internal sodium 4,4-dimethyl-4-silapentane-1-sulfonate (DSS), but were actually measured by reference to internal acetone (δ 2.225 in ^2H_2O at 27°C), with an accuracy of 0.002 ppm.

Results

Fractionation of the OVA/OVM/OVT Dansyl Glyco-asparagines by Reverse-phase HPLC

Glyco-asparagines (120 mg) from hen OVA/OVM were separated into five fractions (A-E) as reported by Huang *et al.* [24]. Fig. 1 shows the fractionation profiles of the dansyl glyco-asparagines A, B and C by reverse-phase HPLC. Glyco-asparagine sample A (1.1 mg) was separated into the subfractions A_1 and A_2 in a ratio of 55:45 (Fig. 1a); sample B (1.2 mg) into the fractions B_1, B_2 and B_3 in a ratio of 30:35:35 (Fig. 1b); and glyco-asparagine fraction C (0.6 mg) yielded one major, C_m (67%), and four minor components (Fig. 1c). Glyco-asparagines D and E were both eluted as single peaks. Glyco-asparagines (8 mg) from hen OVT were separated into two fractions (F_1 and F_2) in the ratio of 1:3. After separation by preparative HPLC, all fractions except the four minor ones of C were subjected to 500-MHz ^{1}H-NMR spectroscopy.

Determination of the ^{1}H-NMR Characteristics of the Dansyl Glyco-asparagines

^{1}H-NMR spectra of ^2H_2O solutions of the glyco-asparagines were recorded at 27°C. As typical examples, the overall spectra of fractions B_3 and C_m are depicted in Fig. 2. In all

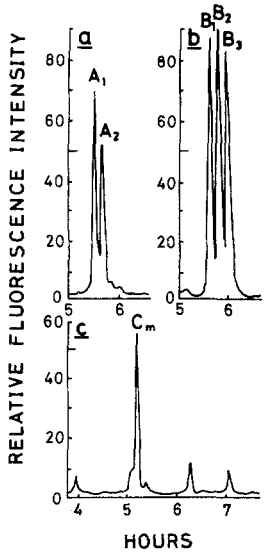


Figure 1. Fractionation of OVA/OVM dansyl glyco-asparagines A, B and C by reverse-phase HPLC. Detailed conditions for HPLC are described in the text.

spectra, the dansyl group can be recognized from the set of aromatic-proton signals in the region between δ 7.6 and δ 8.9, and from the relatively high singlet at δ 2.900, equivalent to six protons, stemming from the dimethylamino function. For interpretation of such ^{1}H-NMR spectra of glycans in terms of primary carbohydrate structures, mainly the so-called structural-reporter group resonances need to be considered [26, 27]. This implies the assignment of the signals from the anomeric protons (δ 4.4-5.4), the mannose H-2 and some H-3 and H-4 atoms (δ 3.9-4.3) and the *N*-acetyl methyl protons (δ 1.95-2.15) (see Fig. 2), with the guidance of reference data of related compounds. The chemical shifts of the structural-reporter groups of the various components of the OVA/OVM/OVT dansyl glyco-asparagine fractions are compiled in Table 1.

The Oligomannoside-type Fractions D and E. The oligosaccharide moieties of the homogeneous dansyl glyco-asparagines D and E appeared to be of the oligomannoside type, containing six and five mannose residues, respectively. The arrangement of the mannose residues in E could be inferred directly from the close similarity of its ^{1}H-NMR spectrum with that of underivatized $Man_5(GlcNAc)_2Asn$, isolated either from the urine of a patient with Gaucher's disease [25] or obtained as ovalbumin fraction V by Kobata and coworkers [2, 3]. A similar reasoning holds for fraction D in comparison to $Man_6(GlcNAc)_2Asn$, being ovalbumin fraction IV [2, 3]; the ^{1}H-NMR spectrum of the authentic fraction IV has been included in the ^{1}H-NMR spectral atlas comprising 72 *N*-glycosidic carbohydrate chains of glycoproteins [27]. The complete structures of D and

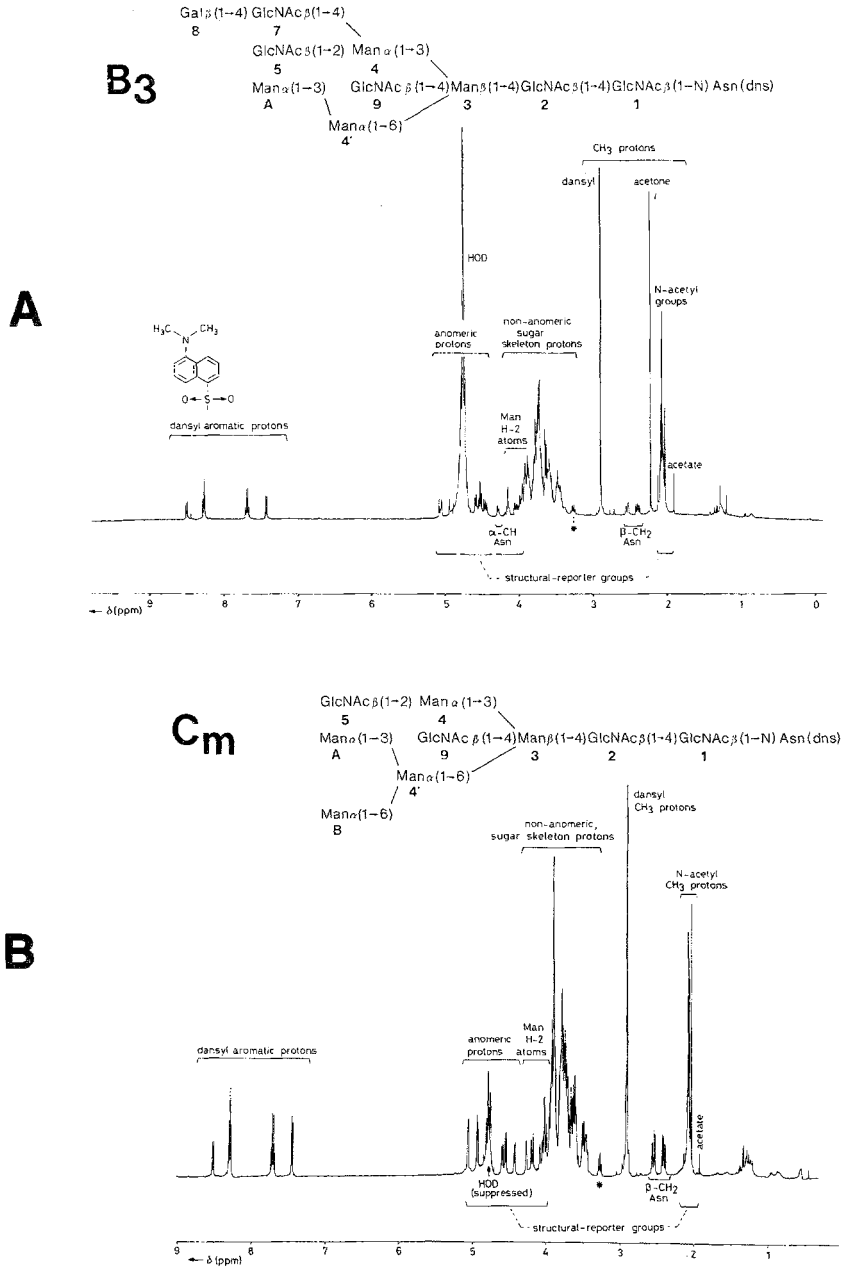
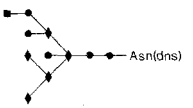
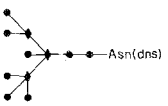


Figure 2. Overall 500-MHz ^{1}H-NMR spectra of OVA/OVM dansyl glyco-asparagine fractions B_3 (A) and C_m (B), in ^2H_2O at 27°C. During the recording of spectrum B, the HO^2H-resonance (HOD) was suppressed by the WEFT pulse sequence (see text). The occurrence of the asterisk-marked doublet-of-doublets at δ 3.26, clearly apart from the bulk of skeleton-proton signals, has been proposed to be correlated with the presence of GlcNAc-**9** in *N*-glycosidic carbohydrates [9, 31]. However, since it cannot as yet be assigned to a certain proton [9, 32], it is not considered here as a structural-reporter group signal.

Table 1. ^{1}H Chemical shifts of structural-reporter groups of constituent monosaccharides for dansyl glyco-asparagines obtained from hen egg-white glycoproteins by revers-phase HPLC.

Reporter group	Residue[a]	Chemical shift[a,b] in			
					
		A_1	A_2	B_1	B_2
H-1	GlcNAc-**1**	4.789[c]	4.790[c]	4.789[c]	4.789[c]
	GlcNAc-**2**	4.582	4.588	4.580	4.586
	Man-**3**	4.730[c]	4.705	4.734[c]	4.703
	Man-**4**	5.048	5.067	5.048	5.068
	Man-**4'**	4.917	4.888	4.916	4.888
	GlcNAc-**5**	4.519	4.539	4.516	4.539
	GlcNAc-**5'**	—	4.558	—	4.555
	GlcNAc-**7**	4.536	4.537	4.516	4.516
	GlcNAc-**7'**	—	4.518[d]	—	4.516[d]
	Gal-**8**	4.471	4.471	—	—
	GlcNAc-**9**	4.410	4.443	4.408	4.443
	GlcNAc-**10'**	—	4.588[d]	—	4.586[d]
	Man-**A**	5.048	—	5.048	—
	Man-**B**	4.926	—	4.928	—
	Man-**C**	—	—	—	—
H-2	Man-**3**	4.161	4.155	4.159	4.151
	Man-**4**	4.291	4.281	4.288	4.278
	Man-**4'**	4.161	4.155	4.158	4.156
	Man-**A**	4.068	—	4.068	—
	Man-**B**	4.005	—	4.005	—
	Man-**C**	—	—	—	—
H-3	Man-**4**	4.039	4.044	4.036	4.042
H-4	Man-**4'**	n.d.	4.192	n.d.	4.191
NAc	GlcNAc-**1**	2.027	2.029	2.027	2.029
	GlcNAc-**2**	2.065	2.070	2.065	2.069
	GlcNAc-**5**	2.051	2.091[e]	2.051	2.091[e]
	GlcNAc-**5'**	—	2.046[e]	—	2.046[e]
	GlcNAc-**7**	2.080	2.079	2.082	2.082
	GlcNAc-**7'**	—	2.095[f]	—	2.095[f]
	GlcNAc-**9**	2.065	2.065	2.065	2.065
	GlcNAc-**10'**	—	2.124[f]	—	2.124[f]

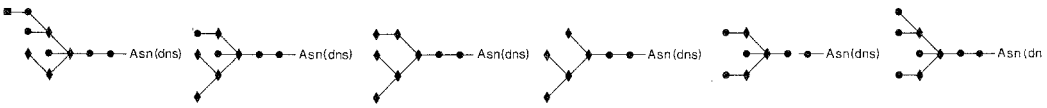
[a] For complete structures and designation of monosaccharide residues, see Fig. 3. In the table-heading, compounds are represented by schematic structures, acording to Ref. 27: ◆, = Man; ●, = GlcNAc; ■, = Gal.

[b] Chemical shifts were acquired at 500 MHz for ^2H_2O solutions at 27°C; they are expressed in ppm downfield from internal DSS.

[c] Partly obscured by HO^2H-line at 27°C (δ = 4.76); visualized by WEFT-technique.

[d,e,f] Assignments may have to be interchanged.

n.d. Value could not be determined by inspection (δ < 4.0).



B_3	C_m	D	E	F_1	F_2
4.789[c]	4.788	4.785	4.787	4.790	4.786
4.584	4.579	4.594	4.594	4.603	4.600
4.736[c]	4.745[c]	4.785	4.785	4.697	4.686
5.052	5.055	5.347	5.097	5.058	5.055
4.941	4.918	4.874	4.873	5.005	5.002
4.518	4.531	—	—	4.553	4.542
—	—	—	—	4.548	4.545
4.535	—	—	—	—	4.515
—	—	—	—	—	—
4.470	—	—	—	—	—
4.441	4.413	—	—	4.465	4.466
—	—	—	—	—	—
5.087	5.046	5.097	5.097	—	—
—	4.928	4.909	4.908	—	—
—	—	5.052	—	—	—
4.158	4.188	4.235	4.253	4.176	4.147
4.296	4.252	4.114	4.077	4.247	4.284
4.158	4.163	4.148	4.148	4.150	4.147
4.068	4.069	4.069	4.069	—	—
—	4.006	3.988	3.978	—	—
—	—	4.069	—	—	—
4.042	n.d.	4.004	n.d.	n.d.	4.046
n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
2.027	2.026	2.029	2.030	2.030	2.030
2.065	2.069	2.068	2.069	2.083	2.083
2.054	2.051	—	—	2.060	2.058
—	—	—	—	2.051	2.051
2.079	—	—	—	—	2.083
—	—	—	—	—	—
2.065	2.064	—	—	2.068	2.064
—	—	—	—	—	—

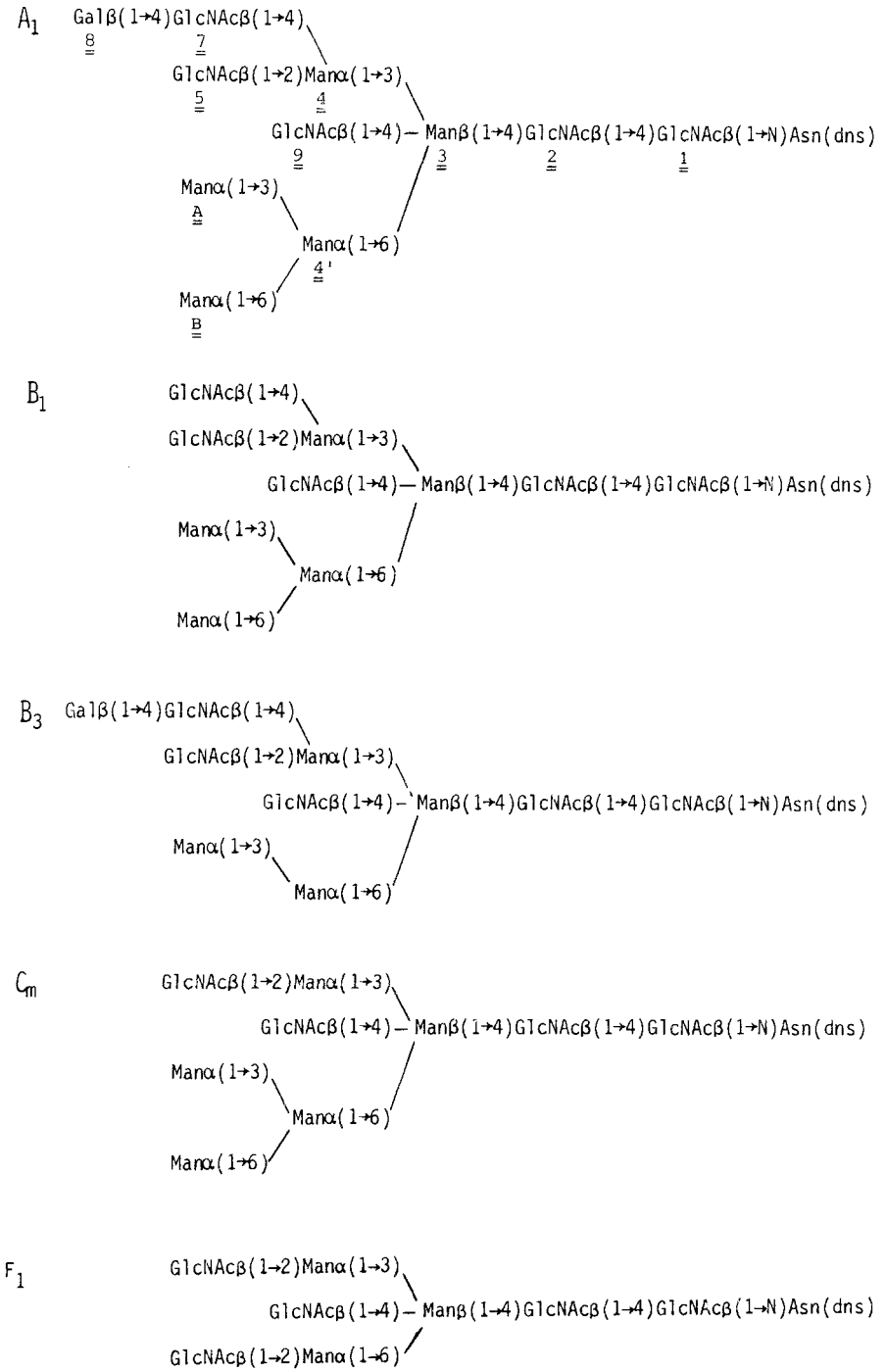
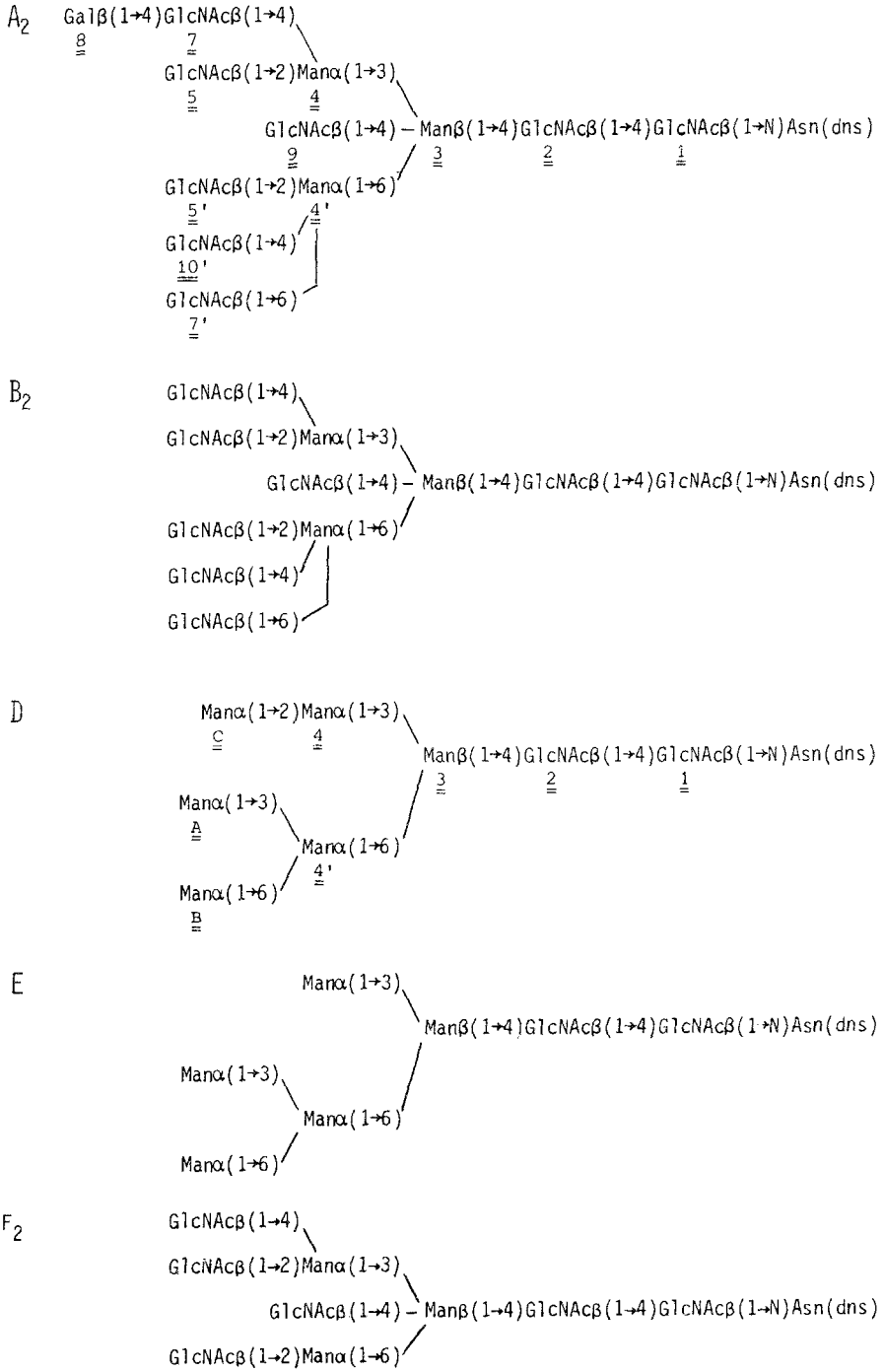


Figure 3. Structures of the main components of the hen egg white glycoprotein dansyl glyco-asparagine fractions separated by reverse-phase HPLC. Designations of monosaccharide residues used for hybrid-type structures are indicated in A_1; those for multi-antennary structures are given in A_2, and those for oligomannoside-type structures in D.


A2
Galβ(1→4)GlcNAcβ(1→4)
8
7
GlcNAcβ(1→2)Manα(1→3)
5
4
GlcNAcβ(1→4)–Manβ(1→4)GlcNAcβ(1→4)GlcNAcβ(1→N)Asn(dns)
9
3
2
1
GlcNAcβ(1→2)Manα(1→6)
5'
4'
GlcNAcβ(1→4)
10'
GlcNAcβ(1→6)
7'
B2
GlcNAcβ(1→4)
GlcNAcβ(1→2)Manα(1→3)
GlcNAcβ(1→4)–Manβ(1→4)GlcNAcβ(1→4)GlcNAcβ(1→N)Asn(dns)
GlcNAcβ(1→2)Manα(1→6)
GlcNAcβ(1→4)
GlcNAcβ(1→6)
D
Manα(1→2)Manα(1→3)
C
4
Manβ(1→4)GlcNAcβ(1→4)GlcNAcβ(1→N)Asn(dns)
3
2
1
Manα(1→3)
A
Manα(1→6)
4'
Manα(1→6)
B
E
Manα(1→3)
Manβ(1→4)GlcNAcβ(1→4)GlcNAcβ(1→N)Asn(dns)
Manα(1→3)
Manα(1→6)
Manα(1→6)
F2
GlcNAcβ(1→4)
GlcNAcβ(1→2)Manα(1→3)
GlcNAcβ(1→4)–Manβ(1→4)GlcNAcβ(1→4)GlcNAcβ(1→N)Asn(dns)
GlcNAcβ(1→2)Manα(1→6)

E, listed in Fig. 3, are in agreement with earlier chemical, enzymic and spectroscopic studies on ovalbumin [3, 4, 6, 9, 10].

Careful comparison of the chemical shift data for D and E with those for the corresponding, underivatized glyco-asparagines [27, 29] demonstrates that the presence of the dansyl group at the *N*-terminus of the amino acid affects the resonance positions of the Asn protons, and of the structural-reporter group protons of the core GlcNAc-residues **1** and **2** in the carbohydrate chain (for the coding of the monosaccharide residues, see Fig. 3). The H-1 signal of the GlcNAc-**1**, characterized by its typical, relatively large $J_{1,2}$ value (9.8 Hz), is shifted upfield from δ 5.07 to δ 4.79. The chemical shift of H-1 of GlcNAc-**2** is less influenced by dansylation, changing from δ 4.605 to δ 4.594. The *N*-acetyl signals of both residues shift downfield upon dansylation of Asn, for GlcNAc-**1** from δ 2.012 to δ 2.029 and for GlcNAc-**2** from δ 2.060 to δ 2.069.

The Intersected Di- and Tri-antennary Fractions F_1 and F_2. Like the OVA/OVM fractions D and E, the dansyl glyco-asparagine fractions F_1 and F_2, obtained from OVT, were found to be homogeneous. The relevant parts of the 500-MHz ^{1}H-NMR spectra of these fractions are presented in Fig. 4. Their oligosaccharide moieties appeared to be of the intersected di- (F_1) and tri-antennary (F_2) type of branching. Taking into account the above mentioned effects of dansylation on the ^{1}H-NMR features of the *N,N'*-diacetylchitobiose unit **1-2**, the structures of F_1 and F_2 could be readily deduced by comparison of the data listed in Table 1 with the ^{1}H chemical shifts reported [27] for the intersected di-antennary heptasaccharide isolated from the urine of a patient with Sandhoff's disease [30], and the intersected tri-antennary glycopeptide isolated from conalbumin (ovotransferrin) [11], respectively. In the latter preparation, being F_2 in the form of the unblocked glycopeptide (Asn-Arg), we detected by ^{1}H-NMR [27] a small amount of the glycopeptide analogue of F_1. However, we were unable to separate these two glycopeptides from each other and hence unable to establish unequivocally the ^{1}H-NMR features of F_1 as a glycopeptide.

The Intersected Hybrid-type and Penta-antennary Fractions A, B and C. The relevant parts of the 500-MHz ^{1}H-NMR spectra of fractions C_m, A_1, B_3 and B_2 are presented in Fig. 5. The signals indicated by dotted lines in the δ 4.7-4.8 spectral regions of Figs. 5B, C and D were obscured by the residual HO^{2}H lines in the spectra recorded at 27°C under standard conditions, but could be visualized by the WEFT technique. Two effects of applying the WEFT pulse sequence, namely, virtually nullifying the HO^{2}H line, by making use of its short T_2, and simultaneously disturbing the signal intensity ratios in the spectra by suppressing narrow signals (e.g. the *N*-acetyl singlets) more effectively than broader (lower intensity) signals (e.g. anomeric doublets), are readily traceable from a comparison of the overall spectra of C_m and B_3 (Fig. 2). The dotted lines in Fig. 5 are assigned to the H-1 atoms of GlcNAc-**1** (δ 4.79; $J_{1,2} \sim 9.7$ Hz) and Man-**3** (δ 4.73-4.75; $J_{1,2} \leqslant 1$ Hz) [27, 29] (see also Table 1 and Fig. 4).

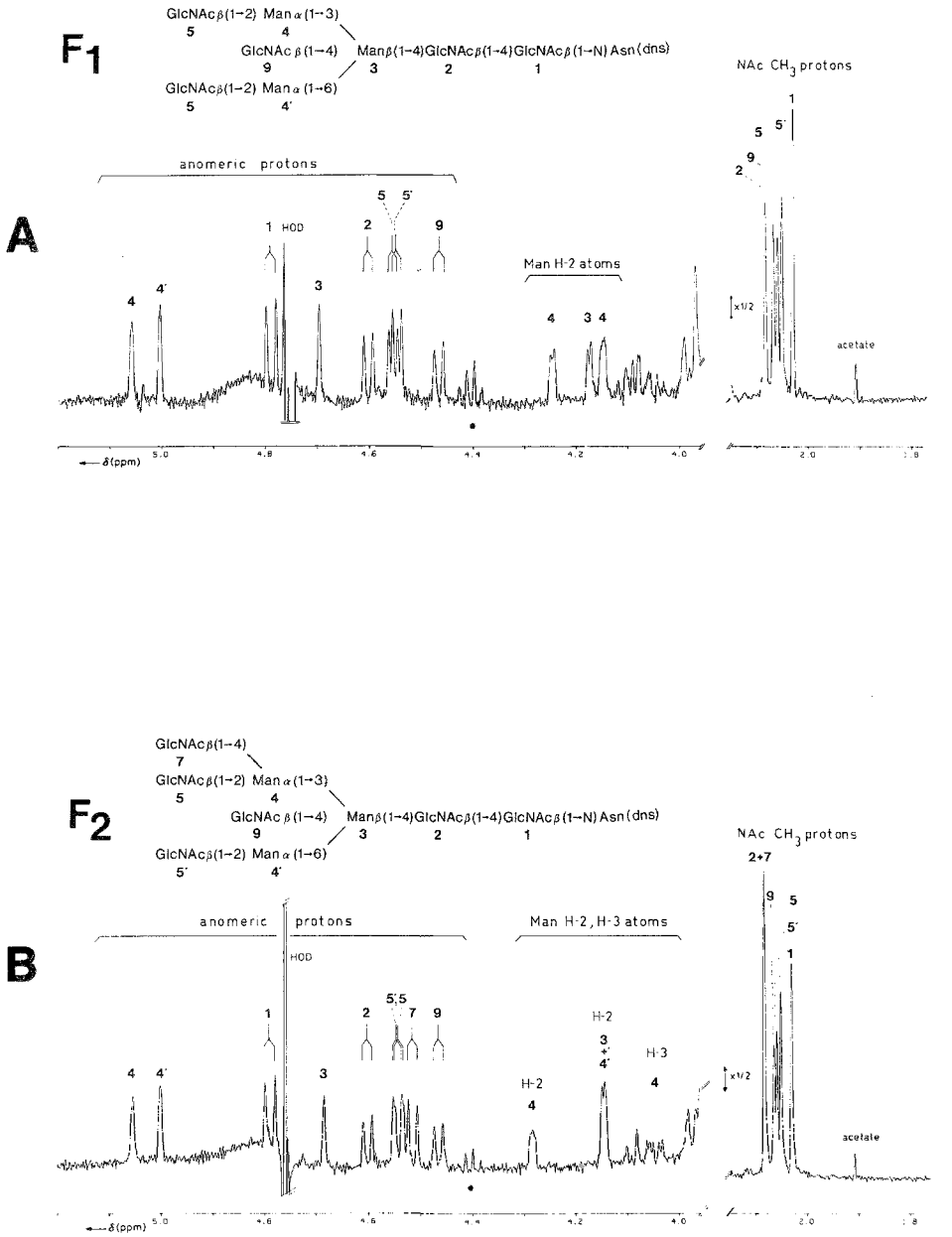


Figure 4. Structural-reporter group regions of the resolution-enhanced 500-MHz ^{1}H-NMR spectra of OVT dansyl glyco-asparagine fractions F_1 (A) and F_2 (B) in ^2H_2O at 27°C. The numbers in the spectra refer to the corresponding residues in the structures. The relative-intensity scales of the *N*-acetyl proton regions differ from those of the other parts of the spectra, as indicated. The asterisk-marked signals at δ 4.4 ppm probably belong to the H-α of dansylated Asn (compare Fig. 2A).

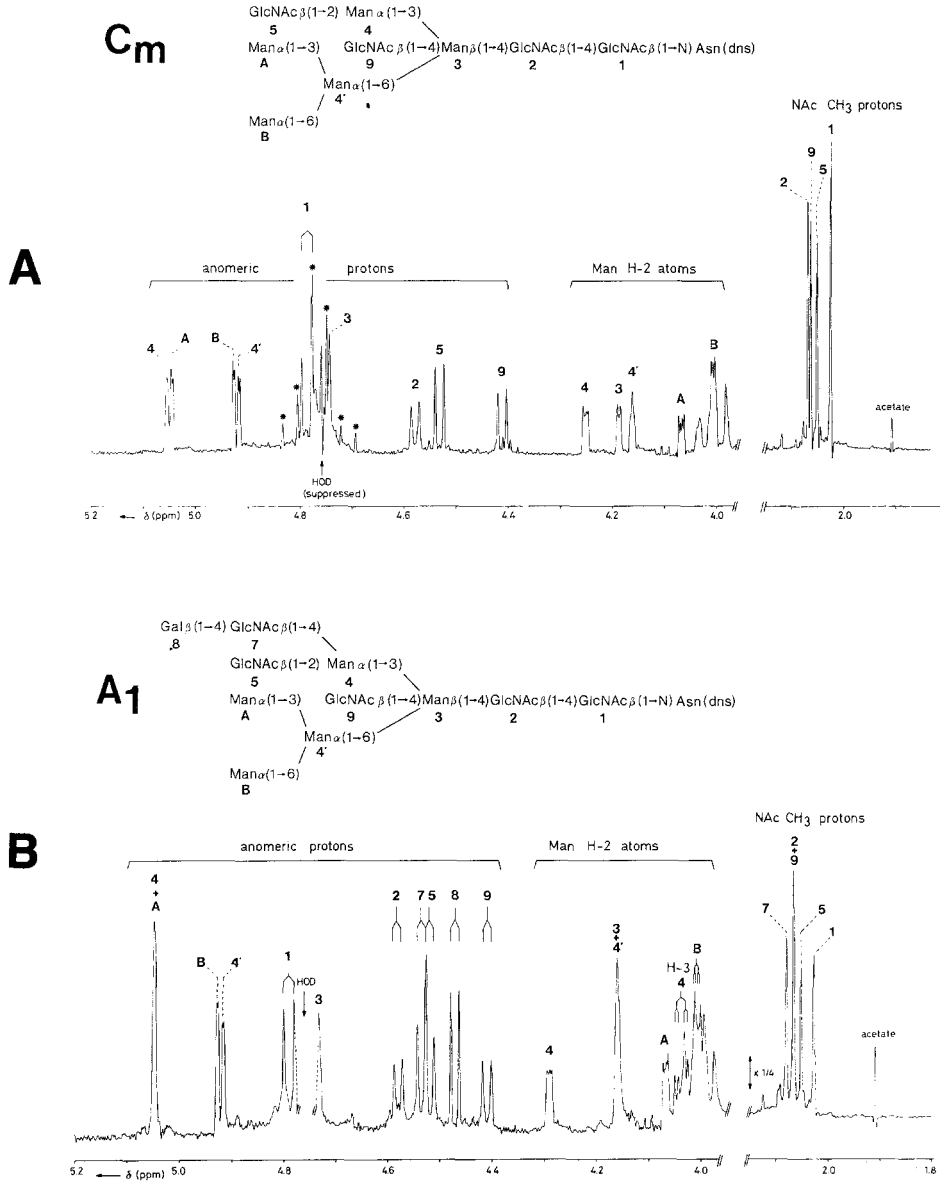
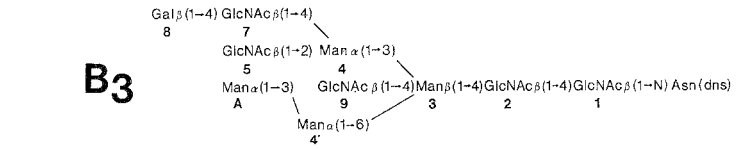
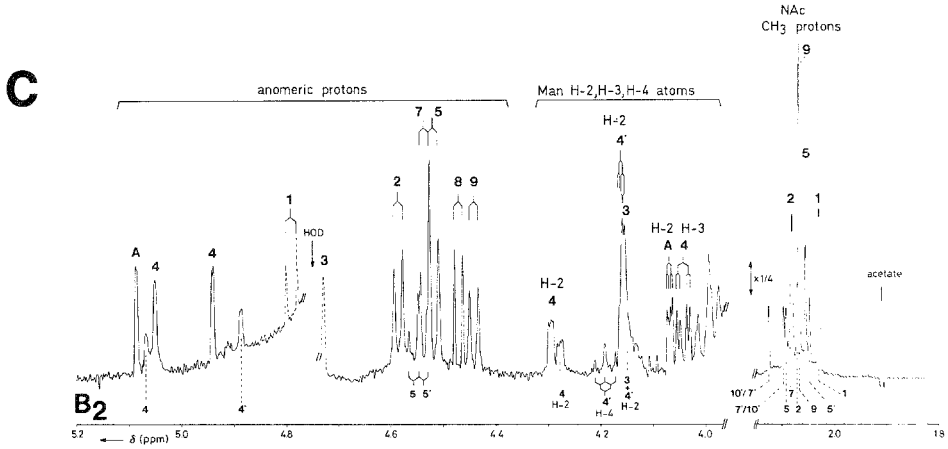
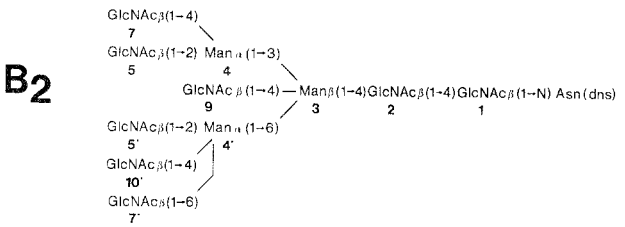
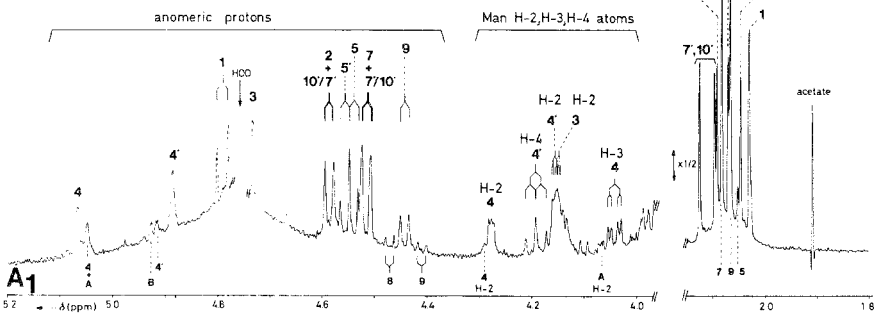


Figure 5. Structural-reporter group regions of the resolution-enhanced 500-MHz ^{1}H-NMR spectra of OVA/OVM dansyl glyco-asparagine fractions C_m (A), A_1 (B), B_3 (C) and B_2 (D), in ^2H_2O at 27°C. The numbers and letters in the spectra refer to the corresponding residues in the structures. The relative-intensity scales of the *N*-acetyl proton regions differ from those of the other parts of the spectra, as indicated. The signals indicated by dotted lines in spectra B-D, at δ 4.7-4.8, were obscured by the residual HO^{2}H(HOD) lines in the spectra recorded at 27°C under standard conditions. However, they could be visualized by the WEFT-technique as illustrated in spectrum A (the asterisks indicate spinning side-bands). It should be mentioned that the WEFT pulse sequence affects also the relative-intensity ratio of NAc signals to anomeric signals. Glyco-asparagines B_3 and B_2 were found to be contaminated with B_2 (25%) and A_1 (10%), respectively, as can be inferred from the lower-intensity signals in spectra C and D. The assignments of the signals attributed to the minor products are indicated below the spectra.


B3
Galβ(1→4) GlcNAcβ(1→4)
8 7
GlcNAcβ(1→2) Manα(1→3)
5 4
Manα(1→3) GlcNAcβ(1→4)Manβ(1→4)GlcNAcβ(1→4)GlcNAcβ(1→N) Asn(dns)
A 9 3 2 1
Manα(1→6)
4'


C
anomeric protons
Man H-2,H-3,H-4 atoms
NAc
CH3 protons
H-2
4'
H-2 H-3
A 4
H-2
4
acetate
x1/4
B2
δ (ppm)


B2
GlcNAcβ(1→4)
7
GlcNAcβ(1→2) Manα(1→3)
5 4
GlcNAcβ(1→4) — Manβ(1→4)GlcNAcβ(1→4)GlcNAcβ(1→N) Asn(dns)
9 3 2 1
GlcNAcβ(1→2) Manα(1→6)
5' 4'
GlcNAcβ(1→4)
10'
GlcNAcβ(1→6)
7'


D
anomeric protons
Man H-2,H-3,H-4 atoms
NAc CH3 protons
H-2 H-2
4' 3
H-4
4'
H-3
4
H-2
4
acetate
x1/2
A1
δ (ppm)

Taking into account that the carbohydrate portion of OVA/OVM consists of only mannose, *N*-acetylglucosamine and galactose [1-20], and that the ^{1}H-NMR parameters of the reporter groups of such residues for *N*-glycosidic carbohydrate chains are fairly well-established [27, 31] the following structural elements of the A, B and C dansyl glycoasparagines can be recognized. The Manβ1-4GlcNAcβ1-4GlcNAcβ1-*N*-Asn(dns) moiety can be verified to be present in (all components of) each fraction based upon the aforementioned chemical shifts of the H-1 atoms and NAc protons of GlcNAc-**1** and -**2**, in conjuction with the occurrence of the typical, broad-lined Man-**3** H-2 doublet [27]. Signals in the δ 4.8-5.3 spectral region, having $J_{1,2} \sim 1.8$ Hz, can be correlated with anomeric protons of α-linked mannose residues. More specifically, those at δ 5.0-5.2 belong to α(1-3)- or α(1-2)-linked mannose, whereas those at δ 4.85-5.0 are attributable to α(1-6)-linked mannose [27]. Each of these signals has a counterpart in the δ 3.9-4.3 spectral region, the mannose H-2 region. In the δ 4.4-4.7 spectral region, a galactose residue at the terminal, non-reducing position, β(1-4)-linked to a *N*-acetylglucosamine, is recognizable from a sharp-lined anomeric doublet at δ 4.47 ($J_{1,2} \sim 7.6$-7.8 Hz) [27]. Other, somewhat broader-lined anomeric doublets in this region can be assigned to β-linked *N*-acetylglucosamine; they are accompanied by a corresponding number of *N*-acetyl signals around δ 2.0.

The Hybrid-type Fraction C_m. The presence of four H-1 signals with $J_{1,2} \sim 1.8$ Hz in the δ 4.8-5.3 region and of three H-1 doublets with $J_{1,2} \sim 8.0$ Hz in the δ 4.4-4.7 region of the spectrum of C_m (Figs. 2B and 5A) points to four α-Man and three β-GlcNAc residues in C_m, in addition to GlcNAc-**1** and Man-**3** showing their anomeric signals at the edges of the HO^{2}H(HOD) line (δ 4.7-4.8). Furthermore, four α-Man H-2 signals (δ 3.9-4.3) and four *N*-acetyl signals (δ 2.0-2.1) support this conclusion. The chemical shift value of the most upfield anomeric doublet, δ 4.413 is characteristic [9, 31] for the intersecting GlcNAc-**9** residue in "hybrid-type" [5] structures. The chemical shift data for the structural-reporter groups of C_m (see Table 1) appeared to be largely identical to those obtained by Carver *et al.* [9] for their ovalbumin glycopeptide C3, at 360 MHz and room temperature. Performance of double-resonance experiments for the four sets of α-Man H-1 and H-2 signals corroborated most of the assignments in [9]. However, the signals at δ 5.055 and δ 5.046 were found to be coupled to the H-2 signals at δ 4.252 and δ 4.069, respectively. Since the latter had been unambiguously assigned to Man-**4** and Man-**A**, respectively, [9, 27, 31], the assignment of the H-1 signals just mentioned has to be interchanged, as compared to [9]. Nevertheless, the structure of C_m is the dansyl analogue of C3 in [9] and of GP-III-A in [2] (see Fig. 3).

The Hybrid-type Fraction B_1. In comparison to the spectrum of C_m, that of fraction B_1 (not shown) reveals the presence of two anomeric doublets coinciding at δ 4.516 ($J_{1,2} \sim$ 8 Hz), while the doublet at δ 4.531 is absent. Moreover, an additional *N*-acetyl singlet is observed at δ 2.082. The newly observed set of H-1 and NAc chemical shifts is consistent with the presence of a terminal GlcNAc-**7** residue that is β(1-4)-linked to Man-**4**. It is in perfect accordance with the data for GlcNAc-**7** in the intersected tri-antennary glycopeptide F_2 derived from OVT (Table 1), as well as for its oligosaccharide-alditol analogue

obtained from hen ovomucoid [12]. This structural feature is supported by the oppositely directed changes in chemical shift ($\Delta\delta \sim 0.03$ ppm) shown by the H-2 signals of Man-**3** and -**4**, going from C_m to B_1 (see Table 1, compare with the step from F_1 to F_2), and by the upfield shift of H-1 of GlcNAc-**5** from δ 4.531 to 4.516. The appearance of the Man-**4** H-3 signal at δ 4.036 is indicative of the 2,4-disubstitution of Man-**4** [12, 26, 27]. The structure of B_1 is included in Fig. 3. It is identical to that of GP-II-B [2, 5], whereas the NMR parameters are in accord with those for agalacto-A3 [9] and for AC-BA [17], recorded at 360 MHz.

The Hybrid-type Fraction A_1. The spectrum of A_1 (Fig. 5B) differs from that of B_1 mainly by the presence of an additional anomeric doublet at δ 4.471 ($J_{1,2}$ 7.6 Hz). This points to the presence of a galactose residue in A_1, β(1-4)-linked to one of the terminal *N*-acetylglucosamine residues in B_1. The precise location of this galactose in the oligosaccharide chain can be deduced from the shift effects introduced in comparison to B_1 (see Table 1). One the the *N*-acetylglucosamine H-1 doublets coinciding at δ 4.516 in the spectrum of B_1 (GlcNAc-**5** or -**7**) has moved downfield to δ 4.536, while the positions of the other anomeric signals are unaffected. Moreover, in the *N*-acetyl region, only the signal which was unambiguously assigned to GlcNAc-**7** shows a small but significant [27] upfield shift, from δ 2.082 to 2.080. Apparently, the shifted H-1 signal belongs also to GlcNAc-**7**. The concerted effects on the H-1 and the *N*-acetyl protons of GlcNAc-**7** prove the galactose residue to be linked to GlcNAc-**7**β (compare [27]). The structure of A_1 is given in Fig. 3. It should be noted that the ^{1}H-NMR parameters for a structure identical to our HPLC-fraction A_1, namely, for ovalbumin glycopeptide A3, or AC-A, have been described in [9] and [17] respectively. Carver *et al.* [9] proved independently the location of galactose to be Gal-**8**, β(1-4)-linked to GlcNAc-**7**, by (^{1}H, ^{1}H) nuclear Overhauser enhancement studies, tracing effects upon presaturation of mannose H-2 resonances. The NMR-deduced structure for A_1 (or A3 [9]) is identical to that elucidated along chemical and enzymic routes for ovalbumin GP-I by Kobata and coworkers [2, 5].

The Hybrid-type Fraction B_3. Considering the signals of relatively high intensity only, the spectrum of dansyl glycopeptide fraction B_3 (Fig. 5C) looks quite similar to that of A_1 (Fig. 5B). A striking difference is that the set of two doublets at $\delta \sim 4.92$ in the spectrum of A_1 is replaced by only one doublet at δ 4.941 ($J_{1,2}$ 1.8 Hz). The latter is coupled to the H-2 signal at δ 4.158. Therefore, B_3 is lacking one of the α(1-6)-linked mannose residues as compared to A_1. Obviously, this missing residue is Man-**B**. The apparent removal of Man-**B**, going from A_1 to B_3, results in shift effects on H-1 of Man-**A** (from δ 5.048 to 5.087) and of GlcNAc-**9** (from δ 4.410 to 4.441), whereas the resonance positions of the other structural-reporter groups remain unaltered. In particular, the H-2 resonance of Man-**4'** is not influenced at all by removal of Man-**B**, corroborating earlier statements that the mannose H-2 resonance position is sensitive to substitution of α-Man at C-3, but not at C-6 [27, 33]. The structure of B_3 (see Fig. 3) has been described for ovalbumin fraction GP-II-A [2, 5]; parts of its ^{1}H-NMR data have been reported for AC-BB in [20]. For data of the agalacto analogue AC-CC, see [18, 19].

The Penta-antennary-type Fraction B_2. The spectrum of fraction B_2 (see Fig. 5D) shows at δ 4.8-5.3 the presence of two anomeric doublets of relatively high intensity. They stem from the α-linked mannose residues **4** and **4'** in the main component of B_2. Around the HO^2H (HOD) resonance, the H-1 signals of GlcNAc-**1** and Man-**3** are observed. The β-anomeric region of the spectrum (δ 4.4-4.7) contains seven relatively intense doublets, all having $J_{1,2} \sim 8$ Hz; one of these belongs to GlcNAc-**2** (δ 4.586). The *N*-acetyl region of the spectrum shows eight singlets that can be ascribed to the main component. The signals at δ 2.2029 and δ 2.069 are attributable to GlcNAc-**1** and -**2**, respectively. The close resemblance of the chemical shifts of the H-1 and H-2 atoms of Man-**3**, -**4** and -**4'** for B_2 (see Table 1) and those described for the oligosaccharide-alditol 11 purified from a hydrazinolysate of hen ovomucoid [12] demonstrates that they have in common the intersected, penta-antennary type of branching. That means that B_2 contains Man-**4**, 2,4-disubstituted by GlcNAc-**5** and -**7'**, respectively; and Man-**4'**, 2, 4, 6-trisubstituted by GlcNAc-**5'**, -**10'** and -**7'**, respectively. These branching patterns are corroborated by the occurrence of a mannose H-3 signal at δ 4.042 (Man-**4**), and a mannose H-4 signal at δ 4.191 (Man-**4'**) (compare [12]). Furthermore, GlcNAc-**9** occurs in β(1-4)-linkage to Man-**3**. The structure of B_2 has been included in Fig. 3. As can be inferred from comparison of the lower intensity signals in Fig. 5D with the spectrum in Fig. 5B, preparation B_2 is contaminated by A_1 ($\sim$ 10%).

The only primary structural difference between compound B_2 and the ovomucoid fraction 11 [12] is the nature of the *N,N'*-diacetylchitobiose unit. In B_2, it occurs in the form of a dansyl glyco-asparagine, whereas in ovomucoid fraction 11 it is present as a reduced oligosaccharide (GlcNAc-**2**, δH-1 4.626; GlcNAc-**1**-OL, δH-2 4.255). The impact of dansylation of Asn on the chemical shifts of the carbohydrate structural-reporter groups has been described to be limited to those of GlcNAc-**1**, GlcNAc-**2** and (H-1 of) Man-**3** (see D, E and F_2). The same holds for reduction of GlcNAc-**1** [12]. However, upon comparison of B_2 and fraction 11 another two significant shift alterations are observed. In the spectrum of B_2, an *N*-acetylglucosamine H-1 doublet is present at δ 4.555, as the counterpart of the doublet at δ 4.545 in the spectrum of fraction 11. Concomitantly, one of the *N*-acetyl singlets at δ 2.055 for fraction 11 is absent from the spectrum of B_2, whereas in the latter a signal occurs at δ 2.091. Going from fraction 11 to B_2, the chemical shifts of the remaining H-1 and NAc signals show deviations not larger than 0.002 or 0.001 ppm, respectively. The unexpected effects on one of the *N*-acetylglucosamine H-1 atoms and on one of the NAc methyl groups must stem from an interaction of the dansyl group with these particular protons. It could be that a ring current shift is responsible for the effects observed. Since our data for H-1 and NAc of GlcNAc-**5** in the hybrid structures B_1 and A_1 show perfect agreement with those described for the corresponding glyco-asparagines without a dansyl group [9, 31] the shifted signals are most likely to be assigned to GlcNAc-**5'**. This is corroborated by molecular (CPK) model building of such a penta-antennary structure. Starting from the fairly well-established conformation of an intersected di-antenna [27, 34], the introduction of GlcNAc-**7**, -**7'** and -**10'** does not necessarily introduce profound changes. A dansyl group attached to Asn can only reach the environment of protons of GlcNAc-**5** and -**5'**, in certain rotamers around the *N*-glycosidic bond. The most probable assignment of the *N*-acetylglucosamine H-1 and NAc signals for B_2 have been listed in Table 1; however, definite assignments require further (conformational) studies.

The Penta-antennary-type Fraction A_2. Fraction A_2 contains a mixture of an intersected penta-antennary product similar to B_2 (60%) and a hybrid type of structure (40%). Upon comparison of the spectrum of A_2 (not shown) with that of Fig. 5C, the minor product appeared to be identical to A_1. The penta-antennary, major product contains an additional galactose residue in β(1-4)-linkage to an *N*-acetylglucosamine as compared to B_2. This is evidenced by the H-1 doublet at δ 4.471. The identity of this galactose residue is proved to be Gal-**8** by the shift of the *N*-acetyl signal of GlcNAc-**7** from δ 2.082 for B_2 to δ 2.079 for A_2, combined with the shift for H-1 of GlcNAc-**7** from δ 4.516 to δ 4.537 (compare [13]). The structure of the main constituent of fraction A_2 is included in Fig. 3.

Discussion

In the present study, we have reported the detection and the detailed characterization of the structural basis of the heterogeneity of hen egg white glycoproteins by employing reverse-phase HPLC to fractionate dansyl derivatives of ovalbumin glyco-asparagines, followed by high-resolution ^{1}H-NMR spectroscopy to determine their structures. The latter have been compiled in Fig. 3.

As to the separation procedure, the following is noteworthy. The introduction of the dansyl group afforded the possibility of further fractionating mixtures of glyco-asparagines that could not be separated completely by Dowex-50 anion-exchange chromatography [24], or other methods (for OVT, see [27]), on the basis of the interaction of this group with silica-based octadecyl chains. Owing to the fluorogenic nature of the dansyl group and to the high resolving power of HPLC, fractionation of dansyl glyco-asparagines by reverse-phase HPLC appeared to be far superior to conventional methods such as gel-permeation chromatography, ion-exchange chromatography and high-voltage electrophoresis [21].

Furthermore, dansyl glyco-asparagines turned out to be suitable derivatives for ^{1}H-NMR spectroscopic studies aimed at primary structural analysis. The signals from the dansyl protons do not interfere with the carbohydrate structural-reporter group resonances (Fig. 2). The effect of dansylation of a glyco-asparagine on the chemical shifts of the relevant carbohydrate signals are largely restricted to those of the core residues GlcNAc-**1** and -**2**. This allows easy comparison with ^{1}H-NMR data for unblocked glycopeptides [9, 27, 31] and oligosaccharide(-alditol)s [12, 13]. The basis of the unexpected influence of dansylation of Asn upon the chemical shifts of the structural-reporter groups of GlcNAc-**5'** in the penta-antennary compounds (A_2 and B_2) requires further investigation.

The galactose residue in A_1, B_3 and also A_2 is attached to GlcNAc-**7**. Galactose could be located in this branch on the basis of the shift effects exerted upon the GlcNAc-**7** reporter groups by its attachment, in comparison to B_1 and B_2. In particular, the reproducibility of ^{1}H-chemical shifts in the *N*-acetyl regions of the spectra permits such structural assignments on the basis of (relative) shift effects that are in the order of magnitude of 0.002 ppm (compare [27]).

Acknowledgements

Thanks are due to Drs. J.H.G.M. Mutsaers and L. Dorland for their help in recording some of the 500-MHz ^{1}H-NMR spectra described. The authors are indebted to Miss C.L.E. de Lint for skilful secretarial assistance. This study was supported by Grant PCM 82-19189 from the National Science Foundation, by the Netherlands Foundation for Chemical Research (SON/ZWO) and by Grant UUKC 83-13 from the Netherlands Cancer Foundation (KWF).

References

1 Montreuil J (1980) Adv Carbohydr Chem Biochem 37:157-223.
2 Kobata A (1984) in Biology of Carbohydrates, vol 2, eds. Ginsburg V, Robbins PW, Wiley, New York, p 87-161.
3 Tai T, Yamashita K, Ogata-Arakawa M, Koide N, Muramatsu T, Iwashita S, Inoue Y, Kobata A (1975) J Biol Chem 250:8569-75.
4 Tai T, Yamashita K, Ito S, Kobata A (1977) J Biol Chem 252:6687-94.
5 Yamashita K, Tachibana Y, Kobata A (1978) J Biol Chem 253:3862-69.
6 Conchie J, Strachan I (1978) Carbohydr Res 63:193-213.
7 Narasimhan S, Harpaz N, Longmore G, Carver JP, Grey AA, Schachter H (1980) J Biol Chem 255:4876-84.
8 Atkinson PH, Grey AA, Carver JP, Hakimi J, Ceccarini C (1981) Biochemistry 20:3979-86.
9 Carver JP, Grey AA, Winnik FM, Hakimi J, Ceccarini C, Atkinson PH (1981) Biochemistry 20:6600-6.
10 Nomoto H, Endo T, Inoue Y (1982) Carbohydr Res 107:91-101.
11 Dorland L, Haverkamp J, Vliegenthart JFG, Spik G, Fournet B, Montreuil J (1979) Eur J Biochem 100:569-74.
12 Paz-Parente J, Wieruszeski JM, Strecker G, Montreuil J, Fournet B, van Halbeek H, Dorland L, Vliegenthart JFG (1982) J Biol Chem 257:13173-76.
13 Paz-Parente J, Strecker G, Leroy Y, Montreuil J, Fournet B, van Halbeek H, Dorland L, Vliegenthart JFG (1983) FEBS Lett 152:145-52.
14 Yamashita K, Kamerling JP, Kobata A (1982) J Biol Chem 257:12809-14.
15 Yamashita K, Kamerling JP, Kobata A (1983) J Biol Chem 258:3099-106.
16 Conchie J, Hay AJ, Lomax JA (1983) Carbohydr Res 112:281-95.
17 Ceccarini C, D'Aniello A, Cacace MG, Atkinson PH (1983) Eur J Biochem 132:469-76.
18 Ceccarini C, Lorenzoni P, Atkinson PH (1983) Biochim Biophys Acta 759:214-21.
19 Nomoto H, Inoue Y (1983) Eur J Biochem 135:243-50.
20 Ceccarini C, Lorenzoni P, Atkinson PH (1984) J Mol Biol 176:161-67.
21 Iwase H, Li S-C, Li Y-T (1983) J Chromatogr 267:238-41.
22 Iwase H, Hotta K (1977) J Biol Chem 252:5437-43.
23 Iwase H, Kato Y, Hotta K (1983) Comp Biochem Physiol 76B:345-48.
24 Huang CC, Mayer Jr HE, Montgomery R (1970) Carbohydr Res 13:127-37.
25 Gray WR (1972) Meth Enzymol 25:121-38.
26 Vliegenthart JFG, van Halbeek H, Dorland L (1981) Pure Appl Chem 53:45-77.

27 Vliegenthart JFG, Dorland L, van Halbeek H (1983) Adv Carbohydr Chem Biochem 41:209-374.
28 van Halbeek H, Vliegenthart JFG, Winterwerp H, Blanken WM, van den Eijnden DH (1983) Biochem Biophys Res Commun 110:124-31.
29 van Halbeek H, Dorland L, Veldink GA, Vliegenthart JFG, Michalski JC, Montreuil J, Strecker G, Hull WE (1980) FEBS Lett 121:65-70.
30 Strecker G, Herlant-Peers MC, Fournet B, Montreuil J, Dorland L, Haverkamp J, Vliegenthart JFG, Farriaux JP (1977) Eur J Biochem 81:165-71.
31 Caver JP, Grey AA (1981) Biochemistry 20:6607-16.
32 Bruch RC, White III HB (1982) Biochemistry 21:5334-41.
33 van Halbeek H, Dorland L, Veldink GA, Vliegenthart JFG, Strecker G, Michalski JC, Montreuil J, Hull WE (1980) FEBS Lett 121:71-77.
34 Carver JP, Brisson JR (1984) in Biology of Carbohydrates, vol 2, eds. Ginsburg V, Robbins PW, Wiley ,New York, p 289-331.