Carbohydrate Research, 235 (1992) 69–81
Elsevier Science Publishers B.V., Amsterdam

A ^{1}H and ^{13}C NMR study of oligosaccharides from human milk. Application of the computer program CASPER

Kerstin Hermansson [a], Per-Erik Jansson [a], Lennart Kenne [a], Göran Widmalm [a] and Frank Lindh [b]

[a] *Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm (Sweden)*
[b] *Reserca AB, S-118 84 Stockholm (Sweden)*

(Received January 16th, 1992; accepted May 1st, 1992)

ABSTRACT

Several oligosaccharides from human milk, containing vicinally branched residues, have been analysed with respect to induced NMR chemical shift changes that originate from the branching. Two types of branching were investigated: (*i*) linear oligosaccharides with a 2-linked residue, which thus becomes vicinally 1,2-disubstituted, and (*ii*) oligosaccharides with either 2,3- or 3,4-branching. It could be concluded that, in ^{13}C NMR spectra of the first type, for which only moderately sized induced changes (< 2 ppm) had been observed previously, large (> 5 ppm) changes are also present. For 2,3- and 3,4-branching, changes similar to those observed earlier were found. In ^{1}H NMR spectra, significant induced shifts for signals from anomeric, aglyconic, and H-5 protons were observed. For most trisaccharides, a unique set of values for the chemical shift differences was found, thus making it suitable to use them for characterisation of substitution patterns in the analysis with the computer program CASPER.

INTRODUCTION

Recently, we described a computer program, CASPER, by which structural analysis of linear and branched oligo- and poly-saccharides can be performed[1,2]. With this method, data from sugar and methylation analysis are used in conjunction with unassigned ^{1}H and/or ^{13}C NMR spectra. The computer program generates, from the data on components and linkages, all possible repeating units or oligosaccharide structures by permutation of the constituent sugar residues and anomeric configurations. If only structures that have compatible $^{1}J_{C,H}$ values for anomeric carbons and $^{3}J_{H,H}$ values for anomeric protons are allowed, compared to those in the experimental spectrum, the number of possible structures is de-

Correspondence to: Dr. P.-E. Jansson, Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm, Sweden.

creased. The NMR spectra for all allowed structures are then calculated using monosaccharide chemical shifts to which disaccharide glycosylation shifts are added. If spectra of oligo- and poly-saccharides with branching at vicinal positions are simulated, correction values are added (see below). The fit between experimental and simulated spectra, normally from a peak-by-peak comparison, is calculated and a ranking list according to fit is generated. The method has been used successfully for oligo- and poly-saccharides of known and unknown structures[3–5].

The reason for having correction values for signals from the three sugar residues involved in the vicinal branching is that changes in the geometries of the glycosidic bonds occur upon branching. Thereby, additional inter-residue interactions often appear. This was early recognized by Lemieux et al. in studies of oligosaccharides related to blood-group determinants[6] and, since then, several other studies of oligosaccharides with vicinally branched regions have been performed[7–13]. These have, inter alia, included studies of chemical shifts and NOEs, as well as correlations of these data to inter-residue atomic distances in computationally derived low-energy conformations.

In order to simulate NMR spectra with CASPER, reference data must be available. There are three databases in CASPER, each with data for both ^{1}H and ^{13}C NMR spectra: (*i*) chemical shifts for monosaccharides, (*ii*) disaccharide glycosylation shifts (i.e., the difference between the chemical shifts of the disaccharide and its constituent monosaccharides), (*iii*) correction values for signals from residues in branch-point regions (the differences between observed and simulated chemical shifts obtained by calculation with only disaccharide glycosylation shifts). These values are implemented for 2,3- and 3,4-branching, and the corrections of the glycosylation shifts are introduced for the branch-point residue and the two sugars substituting it. Reference data have mainly been derived from studies of a number of 3,4-branched trisaccharides[7,8] in which the branch-point residue is methyl α-D-galactopyranoside. Certain approximations to other equatorial–axially substituted sugars (e.g., 3,4-branched L-fucose and 2,3-branched D-mannose or L-rhamnose) are made from these data. The translation to D-mannose and L-rhamnose derivatives may sometimes be unreliable, possibly a result of the vicinity to the anomeric center[13]. Furthermore, the 2-substituted central sugar residue in a linear sequence, sugar1-(1 → 2)-sugar2-(1 → *x*)-sugar3, is also vicinally disubstituted and, for certain anomeric and linkage combinations, steric interactions are expected[14]. However, if $x = 6$, an extra bond, C-5–C-6, is present between the rings; consequently, little interaction is introduced and no correction should be necessary. In order to improve the accuracy of the spectrum simulations, data for new trisaccharides are required. An alternative procedure to synthesis and NMR analysis of the required oligosaccharides has recently been developed[15]. If the desired element is present in a larger oligosaccharide or a polysaccharide and assigned NMR chemical shifts are available, the appropriate glycosylation shifts can be extracted by a simple computational procedure with CASPER and then

entered into the database and used for spectrum simulation of other oligosaccharides.

To obtain glycosylation shift data for branched oligosaccharides and to understand their origin and magnitude, several oligosaccharides isolated from milk, containing branched-trisaccharide elements, have now been investigated. Other NMR studies of some of these oligosaccharides have been published[16,17].

RESULTS AND DISCUSSION

The structures of the oligosaccharides analysed, referred to by their arabic numbers and their abbreviated names, are given in Table I. The ^{13}C and ^{1}H NMR chemical shifts for all but a few signals are given in Tables II and III. For all oligosaccharides, the signals from both anomeric forms of the reducing pyranose residue were assigned, except for **5** and **6**, for which only data for the β-pyranose form are given. Spectra from furanose forms were not investigated.

Chemical shifts from **2–6** offer insight into correction values of six trisaccharide elements (Tables IV and V). For linear trisaccharides having a central 2-linked residue, only *gluco*-trisaccharides with a terminal 3-linked methyl α-D-glucopyranoside residue have been studied earlier[14]. With oligosaccharides **2** and **5**, the influences of differently linked sugars at the reducing end have now been investigated. From compound **4**, trisaccharide elements **V** and **VI**, which have 2,3-diequatorial (*trans*) glycosylation, and from compounds **3** and **6**, trisaccharide elements **VII** and **VIII**, with 3,4-diequatorial (*trans*) glycosylation, the correction values can be extracted. These trisaccharide elements differ substantially from those in the earlier studies, which were all *cis*-disubstituted. A representative from each group, a 1,2-linked trisaccharide, and a 2,3- and a 3,4-substituted trisaccharide, is shown in Fig. 1.

Data for the H-disaccharide, α-L-Fuc*p*-(1 → 2)-D-Gal, were only present in the database approximated from data of the methyl glycoside of α-L-Fuc*p*-(1 → 2)-D-Glc[18] and were thus analysed (Tables II and III). The glycosylation shifts of the disaccharide are given in Tables IV and V. The ^{13}C NMR glycosylation shifts for α-L-Fuc*p*-(1 → 2)-D-Gal were similar to those of α-L-Fuc*p*-(1 → 2)-D-Glc with differences smaller than 0.26 ppm, with the exception of 0.52 ppm for the C-2 signal in the β-D-galactose residue. For ^{1}H NMR glycosylation shifts, all differences but one were smaller than 0.08 ppm, the exception being −0.10 ppm for the H-5 signal in the L-fucosyl residue of compound **1a**.

The NMR spectra of α-L-Fuc*p*-(1 → 2)-D-Gal were complicated due to the occurrence of ~ 25% furanose in addition to α- and β-pyranose forms, giving four different subspectra in the ^{13}C NMR spectrum with 8 signals in the anomeric region. The difference in abundance of the anomeric forms was used to distinguish between them. Furanose forms, ~ 35%, were also found for the B-trisaccharide, with the β-galactofuranose form slightly more abundant than the β-galactopyranose form. The complexity obtained is shown by the anomeric region of the

TABLE I

Structures of investigated oligosaccharides and their abbreviated names

No	Name	Structure
1	H-disaccharide (47:26:9:18) [a]	α-L-Fuc*p*-(1 → 2)-D-Gal
2	2′-Fucosyl-lactose	α-L-Fuc*p*-(1 → 2)-β-D-Gal*p*-(1 → 4)-D-Glc
3	3-Fucosyl-lactose	β-D-Gal*p*-(1 → 4)-D-Glc 3 ↑ 1 α-L-Fuc*p*
4	B-trisaccharide (47:19:9:25) [a]	α-D-Gal*p*-(1 → 3)-D-Gal 2 ↑ 1 α-L-Fuc*p*
5	LNFI	α-L-Fuc*p*-(1 → 2)-β-D-Gal*p*-(1 → 3)-β-D-Glc*p*NAc-(1 → 3)-β-D-Gal*p*-(1 → 4)-D-Glc
6	LNFII	α-L-Fuc*p*-(1 → 4)-β-D-Glc*p*NAc-(1 → 3)-β-D-Gal*p*-(1 → 4)-D-Glc 3 ↑ 1 β-D-Gal*p*
7	LNFIII	β-D-Gal*p*-(1 → 4)-β-D-Glc*p*NAc-(1 → 3)-β-D-Gal*p*-(1 → 4)-D-Glc 3 ↑ 1 α-L-Fuc*p*
8	LNDI	α-L-Fuc*p*-(1 → 4)-β-D-Glc*p*NAc-(1 → 3)-β-D-Gal*p*-(1 → 4)-D-Glc 3 ↑ 1 α-L-Fuc*p*-(1 → 2)-β-D-Gal*p*
9	LNT	β-D-Gal*p*-(1 → 3)-β-D-Glc*p*NAc-(1 → 3)-β-D-Gal*p*-(1 → 4)-D-Glc
10	LNnT	β-D-Gal*p*-(1 → 4)-β-D-Glc*p*NAc-(1 → 3)-β-D-Gal*p*-(1 → 4)-D-Glc

[a] The numbers in parentheses refer to the percentage of α- and β-pyranose and α- and β-furanose, respectively, in D_2O solutions at 70°. The relative amounts are obtained from ^{13}C NMR signal intensities for **1** and from ^{1}H NMR integrals for **4**.

TABLE II

^{13}C NMR chemical shifts in spectra of oligosaccharides **1–6**

Compound	C-1	C-2	C-3	C-4	C-5	C-6	$COCH_3$	C=O
1a. α-L-Fuc*p*-(1 → 2)-α-D-Gal*p*								
Fuc	101.98	69.34	70.52	72.70	68.01	16.12		
Gal	92.87	78.43	69.19	70.38	71.01	62.00		
1b. α-L-Fuc*p*-(1 → 2)-β-D-Gal*p*								
Fuc	100.74	69.46	70.63	72.84	67.73	16.03		
Gal	96.21	80.38	74.29	69.78	75.84	61.81		
2. 2′-Fucosyl-lactose								
Fuc	100.22	69.22	70.66	72.60	67.72	16.08		
Gal	101.28	77.32	74.51	70.01	76.03	61.85		
α-Glc	92.69	72.31	72.13	77.28	71.26	61.06		
β-Glc	96.76	74.95	75.16	77.16	76.14	61.17		
3a. 3-Fucosyl-α-lactose								
Fuc	99.39	69.14	70.33	72.92	67.29	16.11		
Gal	102.69	72.15	73.54	69.27	75.75	62.22		
α-Glc	92.88	73.62	76.01	73.92	71.97	60.85		
3b. 3-Fucosyl-β-lactose								
Fuc	99.22	69.11	70.29	72.92	67.29	16.11		
Gal	102.65	72.15	73.54	69.27	75.75	62.22		
β-Glc	96.73	76.50	78.29	73.92	76.33	60.92		
4a. B-trisaccharide								
Fuc	101.28	68.92	70.47	72.59	67.93	16.05		
Gal	94.50	69.07	70.28	70.32	71.72	62.15 [a]		
α-Gal	92.77	74.66	72.67	65.76	70.78	62.00 [a]		
4b. B-trisaccharide								
Fuc	99.74	68.87	70.66	72.86	67.70	15.96		
Gal	94.43	69.00	70.24	70.38	72.04	62.11 [b]		
β-Gal	96.21	75.89	77.57	64.70	75.42	61.89 [b]		
5. LNFI								
Fuc	100.28	69.09	70.48	72.76	67.29	16.12		
2-Gal	101.11	77.44	74.51	70.03	75.99	61.76 [d]		
GlcNAc	103.88	55.84	78.57	69.46	76.29	61.92	23.17	174.80
3-Gal	103.81	71.04	82.62 [c]	69.32	75.68 [c]	61.63 [d]		
β-Glc	96.66	74.85	75.26	79.56	75.68	61.26		
6. LNFII								
Fuc	98.81	68.83	70.15	72.86	67.59	16.22		
Gal	103.46	71.56	73.46	69.25	75.66 [e]	61.78		
GlcNAc	103.03	56.81	77.00	73.29	76.32 [e]	60.90	23.26	175.39
3-Gal	103.77	70.88	82.86 [f]	69.14	75.71 [f]	62.33		
β-Glc	96.65	74.83	75.28	79.62	75.66	61.25		

[a–f] Pairwise interchangeable.

B-trisaccharide (Fig. 2). The remaining compounds contain a D-glucose residue at the reducing end and no furanose forms were observed.

Using the extraction procedure described previously[15], the correction values for the branched trisaccharide elements **III–VIII** were obtained (Tables IV and V).

Table III

^{1}H NMR chemical shifts [a] of oligosaccharides **1–6**

Compound	H-1	H-2	H-3	H-4	H-5	H-6	H-6′	NAc
1a. α-L-Fuc*p*-(1 → 2)-α-D-Gal*p*								
Fuc	5.09	3.80	3.91	3.84	4.20	1.23		
Gal	5.36	3.80	4.02	4.04	4.11	3.75	3.75	
1b. α-L-Fuc*p*-(1 → 2)-β-D-Gal*p*								
Fuc	5.20	~ 3.81	~ 3.88	~ 3.82	4.30	1.20		
Gal	4.66	3.56	3.83	3.94	3.69	3.76	3.76	
2. 2′-Fucosyl-lactose								
Fuc	5.31	3.81	3.79	3.81	4.22	1.23		
Gal	4.52	3.70	3.87	3.92	3.70	3.77	3.77	
α-Glc	5.23	3.59	3.82	3.70	3.88	3.83	3.87	
β-Glc	4.64	3.31	3.60	3.72	3.48	3.79	3.94	
3a. 3-Fucosyl-α-lactose								
Fuc	5.39	3.81	3.96	3.81	4.78	1.20		
Gal	4.45	3.50	3.66	3.93	3.60	3.76	3.76	
α-Glc	5.20	3.78	3.95	3.88	3.97	3.84	3.89	
3b. 3-Fucosyl-β-lactose								
Fuc	5.44	3.81	3.96	3.81	4.78	1.20		
Gal	4.45	3.50	3.66	3.93	3.60	3.76	3.76	
β-Glc	4.66	3.49	3.78	3.88	3.59	3.84	3.98	
4a. B-trisaccharide								
Fuc	5.12	3.81	3.90	3.82	4.15	1.23		
Gal	5.22	3.88	3.94	4.00	4.33	3.76	3.76	
α-Gal	5.36	4.02	4.17	4.31	4.10	3.77	3.77	
4b. B-trisaccharide								
Fuc	5.26	~ 3.80	~ 3.82	~ 3.81	4.40	1.21		
Gal	5.25	3.89	3.89	4.01	4.22	3.75	3.75	
β-Gal	4.72	3.79	3.96	4.26	3.69	n.a. [b]	n.a.	
5. LNFI								
Fuc	5.21	3.80	3.71	3.78	4.30	1.24		
2-Gal	4.64	3.62	3.84	3.92	3.68	3.78	3.78	
GlcNAc	4.66	3.83	4.01	3.52	3.50	3.77	3.91	2.06
3-Gal	4.44	3.60	3.71	4.13	3.71 [c]	3.78	3.78	
β-Glc	4.66	3.29	3.65	3.64 [c]	3.60 [c]	3.78 [d]	3.92 [d]	
6. LNFII								
Fuc	5.03	3.82	3.89	3.80	4.79	1.19		
Gal	4.52	3.50	3.62	3.92	3.58	3.76	3.76	
GlcNAc	4.79	3.94	4.12	3.77	3.58	3.85	3.97	2.04
3-Gal	4.46	3.63	3.73 [c]	4.15	3.73 [c]	3.79	3.79	
β-Glc	4.67	3.30	3.66	3.65 [c]	3.58 [c]	3.76 [d]	3.92 [d]	

[a] All values for signals from hydroxymethyl groups which have similar chemical shifts are approximate. [b] n.a. = not assigned due to overlapping signals. [c] Assigned from C,H-COSY spectra. [d] Tentative assignments.

TABLE IV

^{13}C NMR glycosylation shifts for disaccharide elements and correction values for vicinally disubstituted trisaccharides, including 1,2-linked residues

Compound		C-1	C-2	C-3	C-4	C-5	C-6
I	(a)	8.86	0.25	0.22	−0.10	0.91	−0.21
(a) (b)	(b)	−0.31	9.08	−0.94	0.10	−0.29	−0.04
α-L-Fuc*p*-(1 → 2)-α-D-Gal*p*							
II	(a)	7.62	0.37	0.33	0.04	0.63	−0.30
(a) (b)	(b)	−1.16	7.42	0.51	0.09	−0.09	−0.03
α-D-Fuc*p*-(1 → 2)-β-D-Gal*p*							
III	(a)	−0.52	−0.24	0.03	−0.24	−0.01	0.05
(a) (b) (c)	(b)	−1.45	−2.00	0.38	0.41	0.05	0.21
α-L-Fuc*p*-(1 → 2)-β-D-Gal*p*-(1 → 4)-β-D-Glc*p*	(c)	0.09	−0.03	−0.14	−2.74	0.52	0.00
IV	(a)	−0.46	−0.37	−0.15	−0.08	−0.44	0.09
(a) (b) (c)	(b)	−2.03	−1.63	0.48	0.49	−0.36	−0.02
α-L-Fuc*p*-(1 → 2)-β-D-Gal*p*-(1 → 3)-β-D-Glc*p*NAc	(c)	0.89	0.16	−4.81	−0.11	0.09	0.30
V	(a)	−1.83	−0.08	−0.06	0.05	−0.21	0.36
(a) (b)	(b)	−0.11	−2.19	−1.08	−1.09	0.00	0.01
α-D-Gal*p*-(1 → 3)-α-D-Gal*p*	(c)	−0.70	−0.42	−0.05	−0.11	−0.08	−0.07
2 ↑ 1 α-L-Fuc*p*							
(c)							
VI	(a)	−1.80	−0.08	0.17	0.23	0.37	0.36
(a) (b)	(b)	−0.01	−2.86	−1.86	−1.49	−0.19	0.11
α-D-Gal*p*-(1 → 3)-β-D-Gal*p*	(c)	−1.00	−0.59	0.03	0.02	−0.03	−0.07
2 ↑ 1 α-L-Fuc*p*							
(c)							
VII	(a)	−1.20	0.25	−0.08	−0.24	−0.32	0.55
(a) (b)	(b)	0.27	1.21	−4.20	−4.60	0.73	−0.37
β-D-Gal*p*-(1 → 4)-β-D-Glc*p*	(c)	−1.13	−0.20	−0.28	0.14	−0.49	0.00
3 ↑ 1 α-L-Fuc*p*							
(c)							
VIII	(a)	−1.42	−0.20	−0.24	0.09	−0.31	0.14
(a) (b)	(b)	0.28	0.98	−5.02	−4.14	1.04	−0.18
α-L-Fuc*p*-(1 → 4)-β-D-Glc*p*NAc	(c)	−0.84	−0.09	−0.06	−0.20	−0.47	−0.03
3 ↑ 1 β-D-Gal*p*							
(c)							

From the studies of 3,4-disubstituted α-D-galactosides[7,8], it was concluded that ^{13}C NMR correction values were between −5 and 2 ppm. The variations in glycosylation shifts for different trisaccharide elements are large and usually a characteristic pattern for each type of substitution is observed. Correlations between glycosylation shifts and calculated inter-residue inter-nuclear distances in low-energy conformations, especially for ^{1}H NMR glycosylation shifts, could be made.

From the study of trisaccharides with a central 2-linked sugar residue[14], the correction values for ^{13}C NMR spectra were found to be fairly small (−0.6 to +1.7 ppm) and with most of the large values positive. For elements **III** and **IV**, seven out of eight correction values for signals from linkage carbons are larger than 0.5 ppm and most of them are larger than 1 ppm. Furthermore, all are negative, in contrast to values found previously, and the two largest are −2.74 and −4.81 ppm, that is, well above the previously observed 1.7 ppm. Thus, it cannot be generalised that a 1,2-disubstitution gives small correction values. Another fairly large value, 0.52 ppm, for the C-5 signal of the reducing sugar residue in element **III** was obtained. The trisaccharide elements **V–VIII** contain di-equatorially 2,3- and 3,4-substituted sugar residues. It cannot be excluded that elements **V** and **VI**, in a larger oligo- or poly-saccharide, would require further corrections due to additional 1,2-substitution. Most signals from linkage carbons have large (> 1 ppm) correction values exclusively with a negative sign, four of them in the range between −4 and −5 ppm. Several correction values for signals from carbons next to the linkage are positive, however, ranging from −1.5 to 1.2 ppm. In principle, no other correction values than for signals from carbons at or next to a linkage are significant.

The ^{1}H NMR correction values for the 3,4-substituted D-galactosides investigated earlier[7,8] were in general fairly small, mostly positive and $< \pm 0.3$ ppm, except for signals for anomeric and H-5 protons in α-linked groups. For elements **V–VIII**, significant correction values for signals from protons at linkage positions are not only confined to those mentioned above but are also found for protons of the branch-point residue. The largest values are found for the H-5 signals in the L-fucosyl group in elements **VII** and **VIII**, namely, 0.45 and 0.46 ppm. The probable reason for this, as well as for other α-linked sugars, is that H-5 is estimated, from energy calculations, to be close to an atom in the aglyconic residue. Depending on the type of atom, the corresponding signal tends to shift either downfield or upfield.

The resemblance between an α-L and a β-D sugar on one hand and an α-D and a β-L sugar on the other hand, mostly reflected in their ^{13}C NMR spectra, has been pointed out earlier[19,20]. In the trisaccharide elements **VII** and **VIII**, the α-L sugar and the β-D sugar are interchanged but still a large degree of similarity in the correction values is observed, supporting the above statement.

The ^{13}C NMR spectra of the oligosaccharides LNT and LNnT (Table I, **9** and **10**) were also recorded. A comparison between the simulated and the experimental ^{13}C NMR spectrum of the β-form of **9** showed a $\Delta\delta$-sum of 1.6 ppm (0.06 ppm/signal). This is an accurate fit, because an average deviation smaller than 0.1

TABLE V

^{1}H NMR glycosylation shifts for disaccharide elements **I** and **II** and extracted chemical shift differences for vicinally disubstituted trisaccharides **III–VIII**

Compound		H-1	H-2	H-3	H-4	H-5	H-6a	H-6b
I	(a)	−0.11	0.03	0.05	0.03	0.00	0.02	
(a) (b)	(b)	0.14	0.02	0.21	0.09	0.08	0.06	0.06
α-L-Fuc*p*-(1 → 2)-α-D-Gal*p*								
II	(a)	0.00	0.04	0.02	0.01	0.10	−0.01	
(a) (b)	(b)	0.13	0.11	0.24	0.05	0.04	0.12	0.04
α-L-Fuc*p*-(1 → 2)-β-D-Gal*p*								
III	(a)	0.11	0.00	−0.09	−0.01	−0.08		
(a) (b) (c)	(b)	−0.03	0.05	0.00	−0.04	−0.04		
α-L-Fuc*p*-(1 → 2)-β-D-Gal*p*-(1 → 4)-β-D-Glc*p*	(b)	−0.03	0.05	0.00	−0.04	−0.04		
IV	(a)	0.01	−0.01	−0.17	−0.04	0.00		
(a) (b) (c)	(b)	0.09	−0.04	−0.04	−0.06	−0.06		
α-L-Fuc*p*-(1 → 2)-β-D-Gal*p*-(1 → 3)-β-D-Glc*p*NAc	(c)	−0.14	−0.03	0.17	−0.04	−0.01		
V	(a)	0.14	0.05	0.06	0.04	0.23		
(a) (b)	(b)	−0.02	0.10	0.11	0.07	0.00		
α-D-Gal*p*-(1 → 3)-α-D-Gal*p*	(c)	0.03	0.01	−0.01	−0.02	−0.05		
2 ↑ 1 α-L-Fuc*p*								
(c)								
VI	(a)	0.15	0.06	0.00	0.04	0.10		
(a) (b)	(b)	0.04	0.09	0.03	0.10	0.01		
α-D-Gal*p*-(1 → 3)-β-D-Gal*p*	(c)	0.06	−0.01	−0.06	−0.01	0.10		
2 ↑ 1 α-L-Fuc*p*								
(c)								
VII	(a)	0.03	−0.04	0.03	0.02	−0.10		
(a) (b)	(b)	−0.01	0.02	0.05	0.17	0.00		
β-D-Gal*p*-(1 → 4)-β-D-Glc*p*	(c)	0.20	0.00	0.08	0.01	0.46		
3 ↑ 1 α-L-Fuc*p*								
(c)								
VIII	(a)	0.05	0.00	0.05	−0.03	0.45		
(a) (b)	(b)	−0.01	0.05	0.19	0.10	−0.05		
α-L-Fuc*p*-(1 → 4)-β-D-Glc*p*NAc	(c)	0.10	−0.05	−0.02	−0.01	−0.12		
3 ↑ 1 β-D-Gal*p*								
(c)								

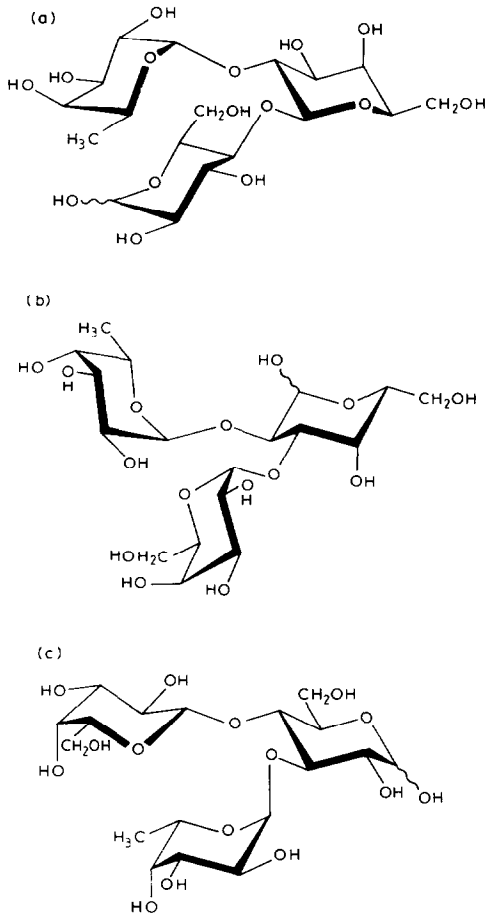

Fig. 1. Schematic representation of some of the investigated branched oligosaccharides: (a) 2′-fucosyl-lactose showing 1,2-disubstitution, (b) the B-trisaccharide showing 2,3-disubstitution, (c) 3-fucosyl-lactose showing 3,4-disubstitution.

ppm per signal is considered good. Likewise, for LNnT, a low $\Delta\delta$-sum of 2.6 ppm (0.10 ppm/signal) was obtained. The small deviations in spectra (i.e., low $\Delta\delta$-sums) for the linear oligosaccharides show that CASPER now handles this type well and that additional effort should be put into collection of data from branched residues.

To test the new correction values, the related compounds LNFIII and LNDI (Table I, **7** and **8**) were chosen. When the ^{13}C NMR spectrum of LNFIII was calculated without correction values, a $\Delta\delta$-sum of 14.4 ppm (0.45 ppm/signal) was obtained. When the correction values from trisaccharide element **VII** (Table IV) were added, the $\Delta\delta$-sum decreased to 4.5 ppm (0.14 ppm/signal). For LNDI, the $\Delta\delta$-sum without and with corrections, using data from fragments **IV** and **VIII**, were 21.6 ppm (0.6 ppm/signal) and 9.9 ppm (0.3 ppm/signal), thus less than half the deviation after use of correction values.

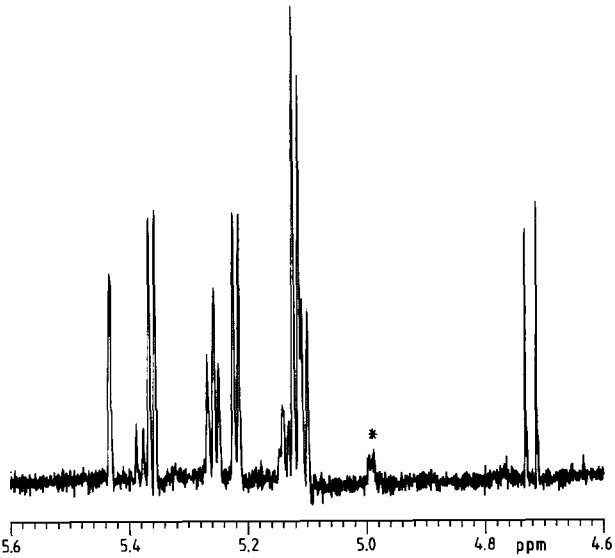

Fig. 2. Anomeric region of the ^{1}H NMR spectrum of the B-trisaccharide, showing complexity due to the different ring forms of the disubstituted galactose; * = unknown.

As another test of the validity of the new correction sets, the ^{13}C NMR data from the *Escherichia coli* O86 O-polysaccharide were used[21]. This polysaccharide has been shown, in ELISA tests, to have blood-group B activity and has a pentasaccharide repeating unit containing a 2,3-disubstituted D-galactose residue (Scheme 1). Several NMR studies of oligosaccharides related to blood-group determinants have been published, inter alia, refs. 22–24. Initial attempts to simulate the ^{13}C NMR spectrum of the *E. coli* O86 O-polysaccharide gave a spectrum that showed deviations from the published chemical shifts for signals that normally would not be expected to have any significant deviations. Re-recording of NMR spectra[25] showed that the indications given by CASPER were correct. The new ^{13}C NMR chemical shifts are therefore given in Scheme 1. The simulated ^{13}C NMR spectrum of the *E. coli* O86 O-polysaccharide was then compared to the assigned experimental chemical shifts. The average deviation is still fairly high, 0.5 ppm/signal, but this derives mainly from other parts of the molecule than the branch-point residue and its substituting sugars. Looking at the ^{13}C NMR signals of the branch-point residue (i.e., from C-1 of residues **a** and **e** and from C-2 and C-3 in residue **b**), only minor differences are observed (< 0.5 ppm). Signals with larger deviation (> 1 ppm) derive from carbons in glycosidic linkages. The ^{13}C NMR spectrum of the polysaccharide was then calculated using branch-point corrections from elements **IV** and **VI**. The element **IV** contains a 2-acetamido-2-deoxy-D-glucose residue instead of a 2-acetamido-2-deoxy-D-galactose residue probably responsible for the still fairly large difference for the signal from C-3 in residue

(b) (c) (d) (e)

→ 2)-β-D-Gal*p*-(1 → 3)-α-D-Gal*p*NAc-(1 → 3)-β-D-Gal*p*NAc-(1 → 4)-α-L-Fuc*p*-(1 →

3
↑
1
α-D-Gal*p*

(a)

Simulated spectrum of the *E. coli* O86 O-polysaccharide.

Residue	C-1	C-2	C-3	C-4	C-5	C-6	$COCH_3$	C=O
(a)	94.4	69.0	70.2	70.4	72.0	62.1		
(b)	101.8	73.4	77.9	65.0	75.0	61.7		
(c)	95.8	50.1	73.4	69.0	71.5	62.0	22.9	175.4
(d)	103.3	52.2	77.2	65.3	75.8	61.9	23.1	175.8
(e)	99.3	69.4	69.8	82.5	67.4	16.1		

Differences between published modified NMR data and the simulated NMR spectrum of the *E. coli* O86 O-polysaccharide

Residue	C-1	C-2	C-3	C-4	C-5	C-6	$COCH_3$	C=O
(a)	−0.1	0.0	0.2	−0.1	0.0	0.1		
(b)	1.4	0.5	0.5	−0.3	0.3	0.1		
(c)	−1.7	−0.5	2.1	0.8	0.3	−0.3	0.5	−0.4
(d)	0.3	−0.3	−1.6	−0.8	0.3	−0.1	−0.1	−1.1
(e)	0.4	−0.1	0.5	0.9	0.4	0.1		

Scheme 1. Simulation of the ^{13}C NMR spectrum of the O-polysaccharide from *Escherichia coli* O86.

c. However, using the correction set, a better fit to the NMR data from residues on the branching region is obtained.

EXPERIMENTAL

Materials.—All oligosaccharides were obtained from Biocarb Biochemicals AB. The oligosaccharides had been purified by several chromatographic procedures and were all more than 95% pure as indicated by HPLC.

NMR spectroscopy.—NMR spectra of solutions in D_2O were recorded at 70°, using a JEOL GSX-270 or a JEOL GX-400 instrument. Chemical shifts are reported in ppm, using sodium 3-trimethylsilylpropanoate-d_4 (δ_H 0.00) and 1,4-dioxane (δ_C 67.40) as internal references. Double quantum filtered COSY, relayed COSY, and C,H-COSY experiments were used to assign signals and were performed according to standard pulse sequences. The assignments were obtained through a combination of the evidence extracted from the two-dimensional spectra, including couplings patterns in the cross-peaks and the relative intensities of signals. The observed couplings indicated that the pyranose rings were in their normal conformation, that is, $^4C_1$ for the D sugars and $^1C_4$ for the L sugars.

CASPER version 2.1[2] was used for the simulations of NMR spectra. The glycosylation shifts from disaccharide fragments which were not available from identical disaccharides were found in the database as approximated values from disaccharides of similar structures. The correction values were also obtained in a similar way.

ACKNOWLEDGMENTS

This work was supported by grants from the Swedish Natural Science Research Council, the Swedish National Board for Technical Development, and Procordias Forskningsstiftelse.

REFERENCES

1 P.-E. Jansson, L. Kenne, and G. Widmalm, *Carbohydr. Res.*, 188 (1989) 169–191.
2 P.-E. Jansson, L. Kenne, and G. Widmalm, *J. Chem. Inf. Comput. Sci.*, 31 (1991) 508–516.
3 P.-E. Jansson, L. Kenne, and G. Widmalm, *Carbohydr. Res.*, 193 (1989) 322–325.
4 H. Baumann, P.-E. Jansson, L. Kenne, and G. Widmalm, *Carbohydr. Res.*, 211 (1991) 183–190.
5 P.-E. Jansson, L. Kenne, and G. Widmalm, *Anal. Biochem.* 199 (1991) 11–17.
6 R.U. Lemieux, K. Bock, L.T.J. Delbaere, S. Koto, and V.S. Rao, *Can. J. Chem.*, 58 (1980) 631–653.
7 H. Baumann, B. Erbing, P.-E. Jansson, and L. Kenne, *J. Chem. Soc., Perkin Trans. 1*, (1989) 2153–2165.
8 H. Baumann, B. Erbing, P.-E. Jansson, and L. Kenne, *J. Chem. Soc., Perkin Trans. 1*, (1989) 2167–2178.
9 G.M. Lipkind, A.S. Shashkov, O.A. Nechaev, V.I. Torgov, V.N. Shibaev, and N.K. Kochetkov, *Carbohydr. Res.*, 195 (1989) 11–25.
10 G.M. Lipkind, A.S. Shashkov, O.A. Nechaev, V.I. Torgov, V.N. Shibaev, and N.K. Kochetkov, *Carbohydr. Res.*, 195 (1989) 27–37.
11 G.M. Lipkind, N.E. Nifant'ev, A.S. Shashkov, and N.K. Kochetkov, *Can. J. Chem.*, 68 (1990) 1238–1250.
12 K. Bock, J.F.-B. Guzman, and R. Norrestam, *Carbohydr. Res.*, 179 (1988) 97–124.
13 N.K. Kochetkov, G.M. Lipkind, A.S. Shashkov, and N.E. Nifant'ev, *Carbohydr. Res.*, 221 (1991) 145–168.
14 A. Adeyeye, P.-E. Jansson, L. Kenne, and G. Widmalm, *J. Chem. Soc., Perkin Trans. 2*, (1991) 963–973.
15 P.-E. Jansson, L. Kenne, and G. Widmalm, *Acta Chem. Scand.*, 45 (1991) 517–522.
16 P. Cagas and C.A. Bush, *Biopolymers*, 30 (1990) 1123–1138.
17 J. Breg, D. Romijn, J.F.G. Vliegenthart, G. Strecker, and J. Montreuil, *Carbohydr. Res.*, 183 (1988) 19–34.
18 P.-E. Jansson, L. Kenne, and E. Schweda, *J. Chem. Soc., Perkin Trans. 1*, (1988) 2729–2736.
19 N.K. Kochetkov, A.S. Shashkov, G.M. Lipkind, and Yu.A. Knirel, *Sov. Sci. Rev., Sect. B*, 13 (1989) 1–73.
20 I. Backman, B. Erbing, P.-E. Jansson, and L. Kenne, *J. Chem. Soc., Perkin Trans. 1*, (1988) 889–898.
21 M. Andersson, N. Carlin, K. Leontein, U. Lindquist, and K. Slettengren, *Carbohydr. Res.*, 185 (1989) 211–223.
22 B.N.N. Rao, V.K. Dua, and C.A. Bush, *Biopolymers*, 24 (1985) 2207–2229.
23 V.K. Dua, B.N.N. Rao, S.-S. Wu, V.E. Dube, and C.A. Bush, *J. Biol. Chem.*, 261 (1986) 1599–1608.
24 B. Bechtel, A.J. Wand, K. Wroblewski, H. Koprowski, and J. Thurin, *J. Biol. Chem.*, 265 (1990) 2028–2037.
25 M. Andersson, personal communication.