Carbohydrate Research, 237 (1992) 11–22
Elsevier Science Publishers B.V., Amsterdam

Computer-assisted analysis of the structure of regular branched polysaccharides containing 2,3-disubstituted rhamnopyranose and mannopyranose residues on the basis of ^{13}C NMR data

Grigory M. Lipkind, Alexander S. Shashkov, Nikolay E. Nifant'ev and Nikolay K. Kochetkov

N.D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Moscow (Russian Federation)

(Received October 30th, 1991; accepted April 15th, 1992)

ABSTRACT

A computer-assisted approach to the analysis of the structure of branched polysaccharides that contain 2,3-di-*O*-glycosylated α-rhamnopyranose and α-mannopyranose residues is based on evaluation of the ^{13}C NMR spectra, using glycosylation effects and their deviations from additivity ($\Delta\Delta$ values) at the branch points. This approach, in combination with monosaccharide and methylation analysis data, has been verified on a series of bacterial polysaccharides of known structure.

INTRODUCTION

A computer-assisted method, based on ^{13}C NMR data and developed[1–4] for the analysis of the structure of regular polysaccharides, has been applied to new linear bacterial polysaccharides[5]. Another approach has been proposed by Jansson et al.[6,7].

The computer-assisted approach involves the following steps[2]: (*1*) generation of all possible structures of a polysaccharide with a given monosaccharide composition, (*2*) calculation of the ^{13}C NMR spectra for each of these structures, and (*3*) a search for the structure whose calculated spectrum is closest to the experimental spectrum.

The ^{13}C NMR spectra of linear polysaccharides can be calculated on the basis of an additive scheme that uses ^{13}C chemical shift data for monosaccharides and the average values of the glycosylation effects. However, for polysaccharides with vicinal branch points, this approach is not valid in general since there are

Correspondence to: Professor N.K. Kochetkov, N.D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Moscow, Russian Federation.

considerable deviations[8–12] (up to 3–4 ppm) between the experimental and calculated ^{13}C chemical shifts.

Therefore, a systematic investigation of the ^{13}C NMR spectra and conformations of branched oligosaccharides with vicinal substitution is being undertaken in order to establish a data base which can be applied to branched polysaccharides. In this context, some 2,3- and 3,4-di-*O*-glycosyl derivatives of methyl β-D-galactopyranoside[13,14] have been studied. Syntheses of the complete series of 2,3-di-*O*-glycosyl derivatives of methyl α-L-rhamnopyranoside with substituents variously with α-D, β-D, α-L, and β-L configurations have been described[15–18], and their conformational and spectral properties[19,20] have been studied as have been some 2,3-di-*O*-glycosyl derivatives of methyl α-D-mannopyranoside[18].

Now we report on the computer-assisted analysis of the structures of branched polysaccharides that contain 2,3-di-*O*-glycosylated α-rhamnose and α-mannose residues. The data[20] on deviations from additivity (ΔΔ values) in the effects of glycosylation in ^{13}C NMR spectra for the corresponding trisaccharides were used in the calculations. The approach was assessed on the basis of known regular polysaccharides, the structures of which have been established by conventional chemical methods and for which ^{13}C NMR data have been reported. Included in this study were the O-specific polysaccharides from *Pseudomonas syringae* pv *tabaci* serogroup VII[21] (**1**), *P. holci* 8300 serogroup I[22] (**2**), *Shigella flexneri* type X[23] (**3**), *P. aeruginosa* V (Verder–Evans)[24] (**4**), *Escherichia coli* 07[25] (**5**), *E. hermannii* strain ATCC 33650[26] (**6**), and *P. holci* 90a serogroup II[27] (**7**).

→ 2)-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 →
2
↑
1
α-D-Fuc*p*3NAc

1

→ 3)-α-L-Rha*p*-(1 → 2)-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 →
3
↑
1
α-D-Fuc*p*3NAc

2

→ 3)-β-D-Glc*p*NAc-(1 → 2)-α-L-Rha*p*-(1 → 2)-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 →
3
↑
1
α-D-Glc*p*

3

→ 4)-α-D-Gal*p*NAcA-(1 → 3)-β-D-Qui*p*NAc-(1 → 2)-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 →
3
↑
1
α-D-Glc*p*

4

→ 3)-α-D-Glc*p*NAc-(1 → 3)-β-D-Qui*p*4NAc-(1 → 2)-α-D-Man*p*-(1 → 4)-β-D-Gal*p*-(1 →
3
↑
1
α-L-Rha*p*

5

→ 4)-β-D-Glc*p*-(1 → 2)-α-D-Rha*p*-(1 → 3)-β-D-Rha*p*-(1 →
3
↑
1
α-DGal*p*

6

→ 3)-α-D-Rha*p*-(1 → 3)-α-D-Rha*p*-(1 → 2)-α-D-Rha*p*-(1 → 2)-α-D-Rha*p*-(1 →
3
↑
1
α-D-Rha*p*

7

RESULTS AND DISCUSSION

The chemical shifts δ (ppm) for the ^{13}C resonances of monosubstituted monosaccharide residues S_i (see Scheme 1) were calculated by equation *1* (ref. 2) where $\delta_0(l)$ are the chemical shifts for the respective nonsubstituted free monosaccharide (l is the number of the carbon), and A(k, l) and B(k′, l) are the effects caused by glycosylation of the unit S_i by the linkages k and k′, respectively.

$$S_{i+1} \xrightarrow{k} S_i \xrightarrow{k'} S_{i-1}$$

Scheme 1.

$$\delta(l) = \delta_0(l) + A(k, l) + B(k', l) \qquad (1)$$

In the same manner, equation *2* for disubstituted residues (unit S_i in Scheme 2) also contains the effect A(k″, l) caused by glycosylation with the second residue of the side chain (the index of linkage k″). In addition, the effects D(k, k″, n) caused by deviations from additivity ($\Delta\Delta$ values) in branched fragments also were taken into account. The $\Delta\Delta$ values * represent the differences between the observed and calculated ^{13}C chemical shifts according to the additive scheme for the corresponding model trisaccharides.

$$\delta(l) = \delta_0(l) + A(k, l) + B(k', l) + A(k'', l) + D(k, k'', l) \qquad (2)$$

$$S_{i+1} \xrightarrow{k} Si \xrightarrow{k'} S_{i-1}$$
$$\Big|_{k''}$$
$$S_{i'}$$

Scheme 2.

* $\Delta\Delta = \delta_{exptl.} - \delta_{calc}$; $\delta_{calc} = \delta_I + \delta_{II} - \delta_{MR}$, where δ_{exptl}, δ_I, δ_{II}, and δ_{MR} are the ^{13}C chemical shifts in the model branched trisaccharide, the respective (1 → 2)- and (1 → 3)-linked disaccharides, and methyl α-L-rhamnopyranoside[19,20].

TABLE I

Deviations from additivity ($\Delta\Delta$, ppm) in the ^{13}C NMR spectra for branch points Sug*p*1-(1 → 2)[Sug*p*2-(1 → 3)]-α-L-Rha*p* / α-D-Man*p*

Absolute configurations of constituents [b]	Type of linkage		D [a]						F [a]	
	k	k″	C-1	C-2	C-3	C-4	C-5	C-6	C-1′	C-1″
LLL/DDD	α-(1 → 2)	α-(1 → 3)	0.6	−0.2	0.2	0.5	0.3	0.6	−0.4	0.3
	α-(1 → 2)	β-(1 → 3)	0.5	−1.0	−0.4	−0.3	0	0	−0.7	−0.2
	β-(1 → 2)	α-(1 → 3)	0.2	0.4	−1.0	0.2	0.2	0.5	0.1	0.2
	β-(1 → 2)	β-(1 → 3)	0.1	−0.8	0.4	−0.6	−0.2	0	−0.3	0.6
LDL/DLD	α-(1 → 2)	α-(1 → 3)	0.4	−0.4	0.2	−0.2	−0.1	0	−0.2	0.5
	α-(1 → 2)	β-(1 → 3)	0.5	−1.7	−0.1	−0.5	0	−0.2	−1.0	0.2
	β-(1 → 2)	α-(1 → 3)	0.4	0	−0.8	−0.6	−0.1	0	0.8	−0.5
	β-(1 → 2)	β-(1 → 3)	0.3	0	−1.1	−0.5	0	−0.2	0.2	0.2
DLL/LDD	α-(1 → 2)	α-(1 → 3)	0.3	0.3	−2.0	1.2	0.3	0.2	−0.5	−0.3
	α-(1 → 2)	β-(1 → 3)	0.6	1.5	0.7	0.3	−0.1	−0.1	0.4	1.1
	β-(1 → 2)	α-(1 → 3)	0.8	−0.2	−0.6	0.6	0.4	0.2	0	0.4
	β-(1 → 2)	β-(1 → 3)	0.4	−0.8	−0.9	−0.4	−0.1	−0.1	−0.5	−0.5
DDL/LLD	α-(1 → 2)	α-(1 → 3)	0.2	0.5	0.2	0.1	0	0.1	0	0.4
	α-(1 → 2)	β-(1 → 3)	0.6	0.5	−1.3	−0.2	0	−0.3	−0.1	0.5
	β-(1 → 2)	α-(1 → 3)	0.3	−0.1	0.2	−0.2	0	0	0	0.4
	β-(1 → 2) [c]	α-(1 → 3)	0.8	−0.9	−1.0	0	0	0	-0.3	−0.9
	β-(1 → 2)	β-(1 → 3)	0.6	−0.5	−0.2	−0.4	−0.1	0	−0.2	0.1

[a] D for C-1/6 in 2,3-di-*O*-glycosylated α-L-Rha*p* and α-D-Man*p* residues; F for C-1′ (Sug*p*1) and C-1″ (Sug*p*2). [b] Related to Sug*p*1, Sug*p*2, and disubstituted residue, respectively. [c] Data[28] for β-D-Glc*p*NAc-(1 → 2)[α-D-Glc*p*-(1 → 3)]-α-L-Rha*p*-OMe.

Vicinal disubstitution affects the chemical shifts of the C-1 resonances in the monosaccharide substituents S_{i+1} and S_i, (Scheme 2). Hence, equation *3* for the calculation of chemical shifts of the C-1 resonances from the S_{i+1} and $S_{i'}$ residues contains the additional term F(k, k″, l).

$$\delta(l) = \delta_0(l) + A(k, l) + B(k', l) + F(k, k'', l) \quad (3)$$

The ^{13}C NMR data for unsubstituted monosaccharides and the glycosylation effects A and B, used in calculations of the ^{13}C NMR spectra of linear polysaccharides, have been reported[2]. The $\Delta\Delta$ values for D and F depend on the absolute and anomeric configurations of the constituents and are given in Table I.

The computer-assisted generation of the possible primary structures of branched polysaccharides took into account all permutations of the constituent monosaccharides of the repeating unit in both the main and side chains (i.e., for a pentasaccharide repeating unit, the side chain may include from 1 to 4 sugar residues) and all possible types of linkage for each disaccharide unit. The ^{13}C NMR spectra were calculated only for structures that accorded with the monosaccharide and methyla-

tion analyses data[21–27]. In addition, the number of α and β linkages, which can be established on the basis of the $J_{C\text{-}1,H\text{-}1}$ values, may also be taken into account.

The S value, calculated[2,3] for each generated structure, is the sum of the squared deviations of the chemical shifts in the observed and calculated spectra normalised to one monosaccharide residue. The possible candidates for the real structures of the polysaccharide are those[2,3] for which the value of S is $\leqslant 1.5$.

The computer-assisted analyses for the polysaccharides **1–7** gave optimal structures (lowest S values) that coincided with the structures reported. The calculated ^{13}C chemical shift data for these optimal structures and the tentative assignment[21–27] of the ^{13}C signals are listed in Table II.

For the *P. syringae* pv. *tabaci* polysaccharide, besides the optimal structure **1** (S 0.9), another structure (**8**) with a relatively low value S (1.1) was found, which also corresponded to the observed ^{13}C NMR spectrum but differed from **1** in the positions of the (1 → 2) and (1 → 3) linkages in the backbone. This coincidence is due to the similarity of the ^{13}C NMR spectra of (1 → 2)- and (1 → 3)-linked rhamnobiosyl fragments[2,3].

→ 3)-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 → 2)-α-L-Rha*p*-(1 →
2
↑
1
α-D-Fuc*p*3NAc

8 (S 1.1)

→ 2)-α-L-Rha*p*-(1 → 2)-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 →
3
↑
1
α-D-Fuc*p*3NAc

9 (S 3.2)

→ 3)-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 →
2
↑
1
α-D-Fuc*p*3NAc-(1 → 2)-α-L-Rha*p*

10 (S 2.2)

→ 3)-α-L-Rha*p*-(1 →
2
↑
1
α-D-Fuc*p*3NAc-(1 → 2)-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*

11 (S 2.3)

The alternative structure (**9**), with the α-D-Fuc*p*3NAc residue as the side chain 3-linked to L-Rha, has a high S value (3.2). This structure differs from **1** in the position of substitution at the branch point. Thus, there is a transition from the

TABLE II

^{13}C chemical shift data [a] calculated (observed) for the O-specific polysaccharides [b] **1–7**

Constituents	Chemical shifts (ppm)					
	C-1	C-2	C-3	C-4	C-5	C-6
Pseudomonas syringae pv. *tabaci* (**1**)						
→ 2)-α-L-Rha*p*-(1 →	101.7	79.8	71.3	73.5	70.0	18.0
	(101.5)	(79.1)	(71.1)	(73.6)	(70.5)	(17.7)
→ 2,3)-α-L-Rha*p*-(1 →	101.0	77.7	76.8	74.1	70.3	18.2
	(100.2)	(77.0)	(76.5)	(73.5)	(70.9)	(17.9)
→ 3)-α-L-Rha*p*-(1 →	103.2	71.4	78.8	72.9	70.0	18.0
	(103.0)	(71.0)	(79.2)	(72.6)	(70.5)	(17.7)
α-D-Fuc*p*NAc-(→	98.1	67.4	52.1	71.8	68.1	16.8
	(98.3)	(67.2)	(52.3)	(71.5)	(68.4)	(16.8)
Pseudomonas holci 8300 (**2**)						
→ 3)-α-L-Rha*p*-(1 →	103.2	71.4	78.8	72.9	70.0	18.0
	(102.8)	(71.2)	(79.0)	(72.8)	(70.4)	(18.0)
→ 2,3)-α-L-Rha*p*-(1 →	102.4	76.1	76.8	71.6	70.0	18.0
	(101.9)	(76.0)	(77.2)	(71.6)	(70.5)	(17.8)
→ 3)-α-L-Rha*p*-(1 →	103.4	71.4	78.8	72.9	70.0	18.0
	(103.2)	(71.2)	(78.8)	(72.5)	(70.4)	(17.8)
→ 3)-α-L-Rha*p*-(1 →	103.4	71.4	78.8	72.9	70.0	18.0
	(103.2)	(71.2)	(78.8)	(72.5)	(70.4)	(17.8)
α-D-Fuc*p*3NAc-(1 →	97.1	67.4	52.1	71.8	68.1	16.8
	(97.3)	(67.0)	(52.4)	(72.2)	(68.1)	(16.5)
Shigella flexneri type X (**3**)						
→ 3)-β-D-Glc*p*NAc-(1 →	103.5	56.5	82.8	69.9	77.2	62.1
	(102.9)	(56.7)	(82.9)	(70.0)	(77.3)	(62.3)
→ 2,3)-α-L-Rha*p*-(1 →	102.8	75.8	75.5	71.8	70.0	18.0
	(102.5)	(75.6)	(75.4)	(72.3)	(70.8)	(17.8)
→ 2)-α-L-Rha*p*-(1 →	102.0	79.8	71.3	73.5	70.0	18.0
	(102.0)	(79.7)	(71.4)	(73.7)	(70.4)	(18.1)
→ 3)-α-L-Rha*p*-(1 →	102.2	71.4	78.8	72.9	70.0	18.0
	(102.5)	(72.0)	(78.7)	(73.0)	(70.4)	(18.0)
α-D-Glc*p*-(1 →	96.4	72.7	74.0	70.9	73.2	61.9
	(96.1)	(72.8)	(74.6)	(71.2)	(72.9)	(61.9)
Pseudomonas aeruginosa V (Verder–Evans) (**4**)						
→ 4)-α-D-Gal*p*NAcA-(1 →	99.9	50.8	68.9	77.5	71.4	73.2
	(99.8)	(50.8)	(69.3)	(77.3)	(71.6)	(73.2)
→ 3)-β-D-Qui*p*NAc-(1 →	103.2	56.2	81.8	76.6	73.2	17.9
	(102.9)	(55.9)	(82.2)	(77.3)	(72.6)	(18.0)
→ 2,3-α-L-Rha*p*-(1 →	103.0	75.8	75.5	71.8	70.0	18.0
	(102.2)	(75.5)	(75.5)	(72.6)	(70.6)	(17.8)
→ 3)-α-L-Rha*p*-(1 →	101.7	71.4	78.8	72.9	70.0	18.0
	(102.1)	(71.2)	(78.8)	(72.6)	(70.5)	(17.8)
α-D-Glc*p*-(1 →	95.8	72.7	74.0	70.9	73.2	61.9
	(96.0)	(72.0)	(74.5)	(70.9)	(73.0)	(61.9)

TABLE II (Continued)

Constituents	Chemical shifts (ppm)					
	C-1	C-2	C-3	C-4	C-5	C-6
Escherichia coli 07 (**5**)						
→ 3)-α-D-Glc*p*NAc-(1 →	98.8	54.1	81.6	70.2	73.3	61.9
	(98.9)	(53.5)	(81.9)	(69.4)	(72.5)	(61.7)
→ 3)-β-D-Qui*p*4NAc-(1 →	104.0	73.6	80.0	57.9	72.5	18.1
	(103.9)	(74.3)	(79.4)	(57.3)	(72.2)	(17.8)
→ 2,3)-α-D-Man*p*-(1 →	100.2	73.8	74.8	65.9	74.1	62.3
	(100.6)	(74.3)	(75.2)	(65.9)	(73.3)	(61.5)
→ 4)-β-D-Gal*p*-(1 →	104.4	71.7	73.6	78.3	76.5	61.5
	(104.7)	(72.0)	(73.3)	(78.0)	(76.2)	(61.4)
α-L-Rha*p*-(1 →	97.0	71.6	71.3	73.5	70.0	18.0
	(97.3)	(71.2)	(71.7)	(73.3)	(69.9)	(17.8)
Escherichia hermannii strain ATCC 33650 (**6**)						
→ 4)-β-D-Glc*p*-(1 →	102.9	74.4	75.6	80.0	75.9	61.3
	(103.2)	(73.9)	(75.4)	(80.4)	(75.9)	(61.3)
→ 2.3-α-D-Rha*p*-(1 →	101.5	79.2	78.2	73.1	70.2	18.5
	(101.4)	(79.1)	(75.8)	(73.4)	(70.8)	(18.0)
→ 3)-β-D-Rha*p*-(1 →	101.0	71.4	81.5	72.5	73.2	18.0
	(101.2)	(71.9)	(81.8)	(72.5)	(73.4)	(18.0)
α-D-Gal*p*-(1 →	102.0	69.6	70.4	70.6	72.2	62.4
	(101.6)	(70.1)	(70.8)	(70.9)	(72.9)	(62.4)
Pseudomonas holci 90a (**7**)						
→ 3)-α-D-Rha*p*-(1 →	103.4	71.4	78.8	72.9	70.0	18.0
	(103.1)	(71.2)	(79.1)	(72.7)	(70.4)	(17.9)
→ 3)-α-D-Rha*p*-(1 →	103.2	71.4	78.8	72.9	70.0	18.0
	(102.8)	(71.4)	(79.1)	(72.8)	(70.5)	(17.9)
→ 2,3)-α-D-Rha*p*-(1 →	102.0	79.9	79.0	73.0	70.0	18.1
	(102.0)	(78.6)	(78.2)	(73.3)	(70.8)	(17.9)
→ 2)-α-D-Rha*p*-(1 →	102.0	79.8	71.3	73.5	70.5	18.0
	(101.8)	(78.9)	(71.4)	(73.7)	(70.7)	(17.9)
α-D-Rha*p*-(1 →	103.8	71.6	71.3	73.5	70.0	18.0
	(103.1)	(71.4)	(71.7)	(73.4)	(70.4)	(18.0)

[a] All ^{13}C signals were assigned in the original papers[21–27]. In some cases the corrections were introduced in order to refer to a common internal standard methanol (50.15 ppm). [b] Data for non-optimal structures are available on request.

substitution mode DLL,αα of the α-L-Rha*p* residue in **1** to LDL,αα in **9** (where the small capitals indicate the absolute configurations of the 2- and 3-substituents and the substituted residue and the Greek letters refer to the anomeric configurations of the 2- and 3-substituents). The ΔΔ values for these modes differ markedly (Table I), namely, −2 ppm for C-3 in DLD,αα and 0.2 ppm for LDL,αα, so that the S values for **1** and **9** differ significantly. Finally, the structures (**10** and **11**) in which the side chain contains more than one monosaccharide residue show high S values (⩾ 2.2).

A choice between the structures **1** (S 0.9) and **8** (S 1.1) generated by computer-assisted analysis may be made on the basis of the NOE data. On pre-irradiation of

H-1 of a 2,3-di-*O*-glycosylated α-L-Rha*p* residue, NOEs were detected[21] on the resonances of H-2 of this residue (20%) and H-3 of the neighbouring residue (40%). Hence, **1** is the most likely structure.

Application of the computer-assisted approach to the *P. holci* 8300 polysaccharide indicates the structure **2** (S 0.6). Other possible structures have S values > 2.6.

For the *Sh. flexneri* type X polysaccharide, besides the optimal structure **3** (S 0.4), two other structures (**12** and **13**) have relatively low S values (0.9 and 1.3, respectively). Structures in which the side chain contains the terminal α-D-Glc*p* residue and another sugar have high S values (> 3.5). This is a consequence of the fact that the effects of glycosylation for the α-L-Rha*p* residues 3-substituted by α-D-Glc*p* and α-L-Rha*p* differ significantly. Thus, for C-2, they are −3.3 and −0.2 ppm, respectively[2]. Calculation revealed unambigously the Glc*p*NAc and Rha*p* residues to be β and α, respectively.

→ 3)-β-D-Glc*p*NAc-(1 → 2)-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 → 2)-α-L-Rha*p*-(1 →
3
↑
1
α-D-Glc*p*

12 (S 0.9)

→ 2)-α-L-Rha*p*-(1 → 2)-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 → 3)-β-D-Glc*p*NAc-(1 →
3
↑
1
αD-Glc*p*

13 (S 1.3)

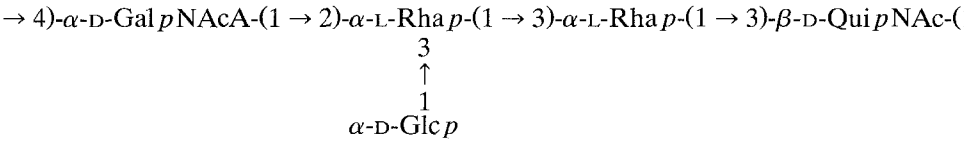
The calculations for the *Sh. flexneri* type X polysaccharide involved ΔΔ values for the glycosylation effects at the branch point β-D-Glc*p*NAc-(1 → 2)[α-D-Glc*p*(1 → 3)-α-L-Rha*p*, which were derived from the ^{13}C NMR spectrum for the model trisaccharide[28] β-D-Glc*p*NAc-(1 → 2)[α-D-Glc*p*(1 → 3)]-α-L-Rha*p*-OMe (see Table I). The use of the ΔΔ values for DDL,βα, determined[20] with the use of the trisaccharide β-D-Glc*p*-(1 → 2)[α-D-Man*p*(1 → 3)]-α-L-Rha*p*-OMe as a standard, afforded slightly worse results although the S value (1.1) for the optimal structure accords with structure **3**.

The alternative structure **13** with an S value of 1.3, corresponds to that of the *Sh. flexneri* type 5a polysaccharide, the ^{13}C NMR spectrum of which is similar to that[23] of the *Sh. flexneri* type X polysaccharide. However, the computation for the type 5a polysaccharide revealed only the correct structure (**14**), for which the S value was smaller (0.7).

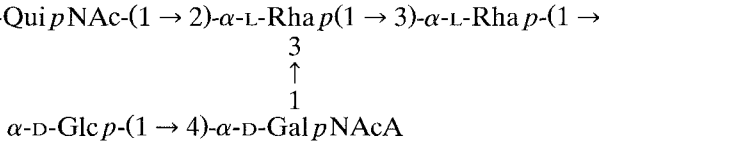
→ 2)-α-L-Rha*p*-(1 → 2)-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 → 3)-β-D-Glc*p*NAc-(1 →
3
↑
1
α-D-Glc*p*

14

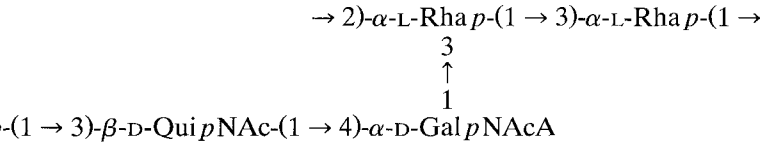
Computer-assisted analysis indicated the structure **4** (S 0.7) for the *P. aeruginosa* V polysaccharide.



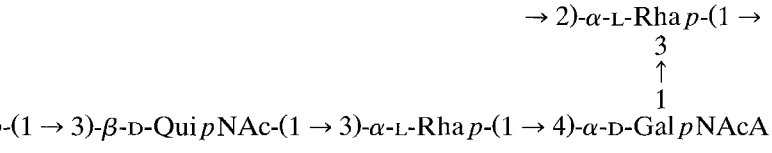
15 (S 1.8)



16 (S 2.5)



17 (S 2.0)



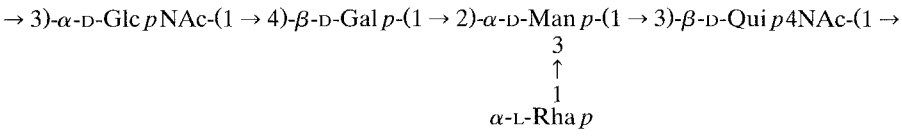
18 (S 2.1)

The alternative structures with one or more sugar residues in the side chain have high S values (> 2.0).

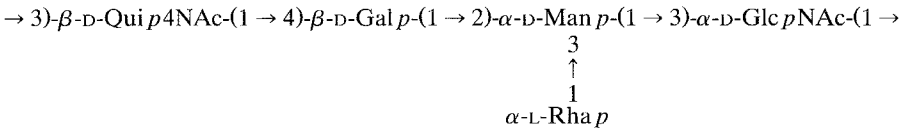
This result shows that the computer-assisted approach simplifies the procedure for the structural determination of branched polysaccharides. Earlier[24], the structure of the polysaccharide **4** was established using various chemical transformations, including the removal of side-chain Glc*p*, selective destruction of the backbone, and analysis of a series of oligosaccharide fragments.

Computer-assisted analysis allows more complicated structures to be established such as that of the *E. coli* 07 polysaccharide (**5**), the repeating unit of which includes five different monosaccharides[25]. This polysaccharide differs from **1–4** in that it contains a disubstituted α-D-Man*p* residue at the branch point instead of an α-L-Rha*p* residue. The computer-assisted analysis revealed only one structure (**5**) with an S value of < 1.5 (0.9). For all other alternative structures (e.g., **19** and **20**), the S values were ≥ 1.9. Structures with side chains that include other residues besides the terminal α-L-Rha*p* have S values of > 4. Previously[25], the

structure **5** was established on the basis of selective transformations and analysis of oligosaccharide fragments.

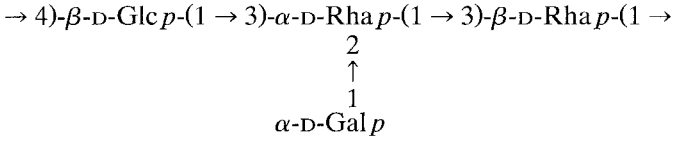


19 (S 1.9)



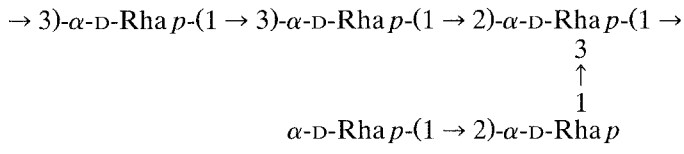
20 (S 2.6)

For the polysaccharides **1–5**, there was good correspondence between the observed and calculated ^{13}C NMR spectra (S < 1.5, see Table II). However, for the *E. hermannii* strain ATCC 33650 polysaccharide (**6**), the best calculated structures had S values of 1.9, 2.8 (**21**), and larger. Nevertheless, the structure with the smallest S value (1.9) was **6**. The rather large S value is due to the relatively high-field position (75.8 ppm) of the C-3 resonance for the disubstituted α-D-Rha*p* residue compared to the calculated value (78.2 ppm, Table II). Nevertheless, the observed ^{13}C signals at 75.8 or 75.9 for the polysaccharide and that calculated at 78.2 ppm may be attributed only to C-3 of the disubstituted α-D-Rha*p*. The chemical shifts of the resonances of the unsubstituted carbons of monosaccharides are in the sequences 81.8, 80.4, 79.1, and 75.8 ppm and 81.5, 80.0, 79.2, and 78.2 ppm in the observed[26] and calculated ^{13}C NMR spectra, respectively, which explains why the optimal calculated variant coincides with the real structure **6**.

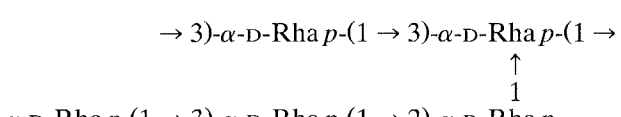


21 (S 2.8)

The computer-assisted analysis of the ^{13}C NMR data[27] for the branched polysaccharide (**7**) of *P. holci* 90a revealed four structures (**7** and **22–24**), each with an S value of 0.9, with the same branch point but differing in the length of the side chain. Thus, for the determination of the structure of the polysaccharide, the ^{13}C NMR and methylation data are not sufficient. Earlier[27], the structure of the polysaccharide was determined by the use of Smith degradation and NOE data.



22 (S 0.9)



23 (S 0.9)

24 (S 0.9)

Thus, the proposed computer-assisted approach indicated the correct structures of the polysaccharides **1–6** with 2,3-di-*O*-glycosylated α-Rha*p* and α-Man*p* residues at the branch points. Future applications of this approach could avoid, or reduce, the need for traditional laborious structural analysis, or at least restrict significantly the number of possible structures and thereby indicate the appropriate chemical work.

EXPERIMENTAL

The program for calculation was developed in the non-dialogue mode, using a BESM-6 computer (Russia) and the language ALCOL-60, the details of which have been described[1,2].

REFERENCES

1 G.M. Lipkind, A.S. Shashkov, and N.K. Kochetkov, *Bioorg. Khim.*, 13 (1987) 633–641.
2 G.M. Lipkind, A.S. Shashkov, Y.A. Knirel, E.V. Vinogradov, and N.K. Kochetkov, *Carbohydr. Res.*, 175 (1988) 59–75.
3 N.K. Kochetkov, A.S. Shashkov, G.M. Lipkind, and Y.A. Knirel, *Sov. Sci. Rev., Sect. B. Chem.*, 13 (1989) 1–73.
4 G.M. Lipkind, A.S. Shashkov, and N.K. Kochetkov, *Carbohydr. Res.*, 198 (1990) 399–402.
5 N.K. Kochetkov, E.V. Vinogradov, Y.A. Knirel, A.S. Shashkov, and G.M. Lipkind, *Bioorg. Khim.*, 18 (1991) 116–125.
6 P.-E. Jansson, L. Kenne, and G. Widmalm, *Carbohydr. Res.*, 168 (1987) 67–77.
7 P.-E. Jansson, L. Kenne, and G. Widmalm, *Carbohydr. Res.*, 188 (1989) 169–191.
8 R.U. Lemieux, K. Bock, L.T.J. Delbaere, S. Koto, and V.S. Rao, *Can. J. Chem.*, 58 (1980) 631–653.

9 G.M. Lipkind, A.S. Shashkov, O.A. Nechaev, V.I. Torgov, V.N. Shibaev, and N.K. Kochetkov, *Bioorg. Khim.*, 15 (1989) 1366–1374.
10 K. Bock, J.F.-B. Guzman, and R. Norrestam, *Carbohydr. Res.*, 179 (1989) 97–124.
11 H. Bauman, B. Erbing, P.-E. Jansson, and L. Kenne, *J. Chem. Soc., Perkin Trans. 1*, (1989) 2153–2165.
12 H. Bauman, B. Erbing, P.-E. Jansson, and L. Kenne, *J. Chem. Soc., Perkin Trans. 1*, (1989) 2167–2178.
13 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.
14 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.
15 N.E. Nifant'ev, L.V. Backinowsky, G.M. Lipkind, A.S. Shashkov, and N.K. Kochetkov, *Bioorg. Khim.*, 17 (1991) 517–530.
16 N.E. Nifant'ev, L.V. Backinowsky, and N.K. Kochetkov, *Bioorg. Khim.*, 16 (1990) 1402–1406.
17 N.E. Nifant'ev, G.M. Lipkind, A.S. Shashkov, and N.K. Kochetkov, *Carbohydr. Res.*, 223 (1992) 109–128.
18 N.E. Nifant'ev, A.S. Shashkov, G.M. Lipkind, and N.K. Kochetkov, *Carbohydr. Res.*, 237 (1992) 95–113.
19 G.M. Lipkind, N.E. Nifant'ev, A.S. Shashkov, and N.K. Kochetkov, *Can. J. Chem.*, 68 (1990) 1238–1250.
20 N.K. Kochetkov, G.M. Lipkind, A.S. Shashkov, and N.E. Nifant'ev, *Carbohydr. Res.*, 221 (1991) 145–168.
21 A.S. Shashkov, G.M. Zdorovenko, E.D. Daeva, L.M. Yakovleva, L.P. Solyanik, R.I. Gvozdyak, Y.A. Knirel, and N.K. Kochetkov, *Bioorg. Khim.*, 16 (1990) 90–97.
22 Y.A. Knirel, G.M. Zdorovenko, A.S. Shashkov, L.M. Yakovleva, N.Y. Gubanova, and R.I. Gvozdyak, *Bioorg. Khim.*, 14 (1988) 172–179.
23 P.-E. Jansson, L. Kenne, and T. Wehler, *Carbohydr. Res.*, 179 (1988) 359–368.
24 Y.A. Knirel, N.A. Kocharova, A.S. Shashkov, N.K. Kochetkov, E.V. Kholodkova, and E.S. Stanislavsky, *Eur. J. Biochem.*, 166 (1987) 189–197.
25 V.L. L'vov, A.S. Shashkov, B.A. Dmitriev, N.K. Kochetkov, B. Jann, and K. Jann, *Carbohydr. Res.*, 126 (1984) 249–259.
26 L.M. Beynon, D.R. Bundle, and M.B. Perry, *Can. J. Chem.*, 68 (1990) 1456–1466.
27 Y.A. Knirel, G.M. Zdorovenko, L.M. Yakovleva, A.S. Shashkov, L.P. Solyanik, and I.V. Zacharova, *Bioorg. Khim.*, 14 (1988) 166–171.
28 L.V. Backinowsky, A.R. Gomtsyan, N.E. Bayramova, and N.K. Kochetkov, *Bioorg. Khim.*, 10 (1984) 79–87.