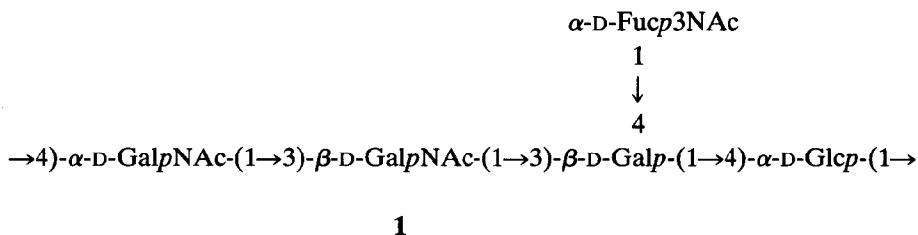## Note

---

Computer-assisted structural analysis of the branched O-specific polysaccharide of *Salmonella arizonae* O63 (*Arizona* O8) on the basis of ^{13}C-n.m.r. data

GRIGORY M. LIPKIND, ALEXANDER S. SHASHKOV, AND NIKOLAY K. KOCHETKOV

N. D. Zelinsky Institute of Organic Chemistry, Academy of Sciences of the U.S.S.R., Moscow (U.S.S.R.)

(Received April 19th, 1989; accepted for publication, July 8th, 1989)

A computer-assisted method, elaborated[1] for the structural determination of regular polysaccharides on the basis of ^{13}C-n.m.r. data, has been modified for branched polysaccharides and we now report its first application, namely to the O-antigen of *S. arizonae* O63 (*Arizona* O8) for which the structure **1** has been proposed[1,2],

<div align="center">

α-D-Fuc*p*3NAc

1

↓

4

→4)-α-D-Gal*p*NAc-(1→3)-β-D-Gal*p*NAc-(1→3)-β-D-Gal*p*-(1→4)-α-D-Glc*p*-(1→

1

</div>

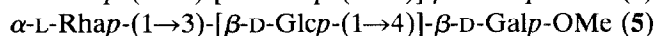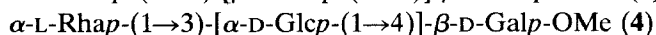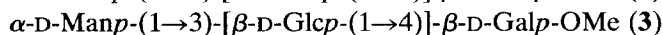where Fuc3NAc is 3-acetamido-3-deoxyfucose.

The following steps are involved in the computer-assisted approach: (*1*) the generation of all possible structures of a polysaccharide, (*2*) the evaluation of the ^{13}C-n.m.r. spectra for each structure starting from the chemical shifts of the resonances of the constituent monosaccharides and glycosylation effects, (*3*) searching for the structure characterised by the calculated spectrum most similar to that observed. In branched polysaccharides, considerable deviations (up to 5 p.p.m.) of experimental ^{13}C chemical shifts from those calculated within the additive scheme[1] are often observed for a monosaccharide at the branching point substituted at vicinal positions. Hence, the formula for the calculation of chemical shifts (δ) of this sugar should be adjusted to

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

where l is the number of the carbon, $A(k,l)$, $A(k'',l)$, and $B(k',l)$ are the effects caused by the glycosylation of the monosaccharide at the branching point (the indexes of linkage types in the main and side chains, k and k'', respectively) and by the formation of the glycosidic linkage (the index of linkage type k'), and $D(k,k'',l)$ is the deviation of the glycosylation effect from the sum of the effects observed in the corresponding disaccharides.

In generating the possible structures of the O63 O-specific polysaccharide, account was taken of all permutations of the constituent monosaccharides in the backbone and the side chain, which may contain 1–4 sugar residues. Those structures were rejected which were inconsistent with the methylation analysis data[2] (Fuc3NAc is the terminal residue of the side chain and 3,4-disubstituted Gal is at the branching point) and the $J_{C-1,H-1}$ values[2] (which showed that, in the repeating unit, three sugar residues are α and two are β).

The effects of glycosylation used for calculation of the ^{13}C-n.m.r. spectra of the polysaccharide have been reported[1]. Deviations (D) for the galactose residue at the branching point are given in Table I as a matrix including four lines which correspond to all possible combinations of glycosidic linkages. The values D were determined from the ^{13}C-n.m.r. data[3] of the model branched trisaccharides **2–5** having the general formula Sug$p$1-(1→3)-[Sug$p$2-(1→4)]-β-D-Gal$p$-OMe.

α-D-Man$p$-(1→3)-[α-D-Glc$p$-(1→4)]-β-D-Gal$p$-OMe (**2**)
α-D-Man$p$-(1→3)-[β-D-Glc$p$-(1→4)]-β-D-Gal$p$-OMe (**3**)
α-L-Rha$p$-(1→3)-[α-D-Glc$p$-(1→4)]-β-D-Gal$p$-OMe (**4**)
α-L-Rha$p$-(1→3)-[β-D-Glc$p$-(1→4)]-β-D-Gal$p$-OMe (**5**)
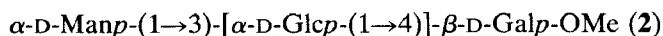
Instead of trisaccharides with the β-D-(1→3) linkage, for which no ^{13}C-n.m.r. data were available, **4** and **5** with α-L-(1→3) linkages were used. Stereochemically, the latter compounds are similar to the former[4] and should have similar D values, as has been exemplified[3]. Corrections (F) for C-1 of Sug1 and Sug2 are also given in Table I.

Values (S) were calculated for each of the generated structures of the poly-saccharide, which are sums of the squared deviations of the chemical shifts in the

TABLE I

DEVIATIONS OF EFFECTS OF GLYCOSYLATION D AND F (p.p.m.) IN TRISACCHARIDES Sug$p$1-(1→3)-[Sug$p$2-(1→4)]-β-D-Gal$p$-OMe

Types of linkage		D Gal residue						F Sug1	Sug2
k'	k''	C-1	C-2	C-3	C-4	C-5	C-6	C-1'	C-1''
α-(1→3)	α-(1→4)	0	−0.4	−0.5	−0.8	0	0	0.3	−0.1
α-(1→3)	β-(1→4)	0	−0.3	0.6	−1.3	0	0	1.0	−0.6
β-(1→3)	α-(1→4)	0	0.7	−2.5	0	0	0	−0.6	0
β-(1→3)	β-(1→4)	0	0.4	−2.4	−2.7	0	0	−0.7	−1.2

TABLE II

CHEMICAL SHIFTS IN THE ^{13}C-N.M.R. SPECTRUM OF THE *S. arizonae* O63 O-SPECIFIC POLYSACCHARIDE (CALCULATED VALUES ARE IN PARENTHESES)

Constituent	Chemical shifts (in D_2O, p.p.m.)					
	C-1	C-2	C-3	C-4	C-5	C-6
→4)-α-D-GalpNAc-(1→	95.3	51.0	68.8	79.2	72.4	61.5
	(94.6)	(51.0)	(68.5)	(79.2)	(72.4)	(61.7)
→3)-β-D-GalpNAc-(1→	103.8	51.8	76.0	65.5	76.0	62.3
	(103.8)	(52.0)	(76.5)	(65.4)	(76.4)	(62.2)
→3,4)-β-D-Galp-(1→	104.1	71.8	80.7	77.1	76.8	61.2
	(104.0)	(71.7)	(81.3)	(78.0)	(76.5)	(61.5)
→4)-α-D-Glcp-(1→	101.3	72.8	72.8	79.8	72.3	61.2
	(101.3)	(72.7)	(72.6)	(79.9)	(71.9)	(61.1)
α-D-Fucp3NAc-(1→	99.7	67.7	52.1	72.2	68.1	16.6
	(101.2)	(67.4)	(52.1)	(71.8)	(68.1)	(16.8)

observed and calculated spectra related to one monosaccharide residue. As has been shown[1], the possible candidates for the real structure of the polysaccharide are those for which the value S does not exceed 1.5 significantly. Such structures of the O63 polysaccharide are shown below in addition to **1** (S 1.0); for all other possible structures, S was >2.5.

α-D-Fucp3NAc-(1→4)-α-D-GalpNAc-(1→4)

$$\search$$

→3)-β-D-GalpNAc-(1→3)-β-D-Galp-(1→4)-α-D-Glcp-(1→   (S 1.5)

α-D-Fucp3NAc-(1→3)-β-D-GalpNAc-(1→4)-α-D-Glcp-(1→4)-α-D-GalpNAc

1

↓

4

→3)-β-D-Galp-(1→

(S 1.7)

Therefore, of 1000 structures considered for the polysaccharide, which are in accord with the methylation data and the number of α and β linkages, only three are consistent with the observed ^{13}C-n.m.r. spectrum, the structure with the least value (1.0) of S being the real structure of the polysaccharide. The calculation also allowed tentative assignment of the signals in the observed ^{13}C-n.m.r. spectrum[1,2] (Table II).

Thus, computer-assisted analysis can be applied in the structural elucidation of branched polysaccharides. As ^{13}C-n.m.r. data for appropriate new branched oligosaccharides accumulate, the accuracy of the method will increase and make possible non-destructive analysis of branched polysaccharides.

EXPERIMENTAL

The ^{13}C-n.m.r. spectrum of the polysaccharide was recorded with a Bruker AM-300 Instrument for a solution in D_2O at 60° (internal MeOH, δ 50.15).

The programme for calculation was developed in the non-dialogue mode, using a BESM-6 computer (U.S.S.R.) and the language ALGOL-60, the details of which have been described[1].

REFERENCES

1 G. M. LIPKIND, A. S. SHASHKOV, Y. A. KNIREL, E. V. VINOGRADOV, AND N. K. KOCHETKOV, *Carbohydr. Res.*, 175 (1988) 59–75.
2 E. V. VINOGRADOV, Y. A. KNIREL, G. M. LIPKIND, A. S. SHASHKOV, AND N. K. KOCHETKOV, *Bioorg. Khim.*, 13 (1987) 1399–1404.
3 G. M. LIPKIND, A. S. SHASHKOV, O. A. NECHAEV, V. I. TORGOV, V. N. SHIBAEV, AND N. K. KOCHETKOV, *Bioorg. Khim.*, 15 (1989) 656–669.
4 A. S. SHASHKOV, G. M. LIPKIND, Y. A. KNIREL, AND N. K. KOCHETKOV, *Magn. Reson. Chem.*, 26 (1988) 735–747.