Chem. Soc. Rev., 1990, **19**, 29–54

HAWORTH MEMORIAL LECTURE*

Microbial Polysaccharides: New Approaches

By N. K. Kochetkov
N. D. ZELINSKY INSTITUTE OF ORGANIC CHEMISTRY,
ACADEMY OF SCIENCES OF THE USSR, MOSCOW, USSR

1 Introduction

The polysaccharides are the biopolymers which constitute the bulk of organic matter of the biosphere and the basis of the structural chemistry of this important class of natural products is forever connected with the life work of Sir Walter Norman Haworth.

Polysaccharides play very important roles in the life of a cell as sources of energy and as construction materials, and are intimately involved in biological processes connected with cell surface specificity, *e.g.* cell-to-cell recognition, interaction of cells with phages and outer media, *etc.*[1] The technological significance of polysaccharides is also very well known and, as an endless self-reproducible source of carbon for the chemical industry, will undoubtably become more important.

Progress in polysaccharide chemistry, as in any branch of the chemistry of natural compounds, depends on the parallel development of methods of structural analysis and synthetic chemistry.

In polysaccharide chemistry one has to distinguish regular polysaccharides from irregular ones. A knowledge of this difference, the result of different pathways in their biosynthesis, is essential in determining the strategy of chemical research. In fact it is in the structural chemistry of regular polysaccharides that rapid progress has been observed. This family comprises, along with simple plant polysaccharides, a great and rapidly growing number of microbial polysaccharides built up of repeating oligosaccharide units. This important class of polysaccharides with highly specific structures attracts much attention because the specific biological activities of its members determine the immunological characteristics of microbial cells and their interrelationships with their environments, and also because of the growth in industrial applications. Naturally, the determination of primary structure is the first step in their investigation.

Application of classical methods of structural analysis, based on the more or less specific depolymerization of the polysaccharide chain and consequent analysis of fragments, can not satisfy modern demands, especially for analysis of small amounts of biological material.

* Delivered at the Spring Meeting of the Royal Society of Chemistry Carbohydrate Group on 30th March 1989 at Cranfield College, Bedford and at the Annual Chemical Congress of the Royal Society of Chemistry on 6th April 1989, at Hull University.

[1] See, for example, 'Receptor in Cellular Recognition', ed. R. M. Gorczynski, Academic Press, 1987, pp. 215–244, 251–256.

The use of mass-spectrometry initiated in our laboratory about 25 years ago[2,3] has proved to be of great service in simplifying methylation analysis[4] and analysis of small oligosaccharide fragments,[5] but unfortunately not very helpful for rapid structural analysis of polysaccharides.

The crucial role in the creation of a new strategy of polysaccharide analysis belongs to ^{13}C-NMR spectroscopy. Unlike ^{1}H-NMR spectra, line broadening and overlapping do not seriously affect ^{13}C-NMR spectra owing to the much larger interval of chemical shifts covered, so ^{13}C-NMR spectroscopy has turned out to be more promising.

Analysis of the ^{13}C-NMR spectrum of a regular polysaccharide, along with easily obtainable data on its monosaccharide composition, allows one to make unambiguous conclusions as to the type of glycosidic linkages (1→2, 1→3, *etc.*), their configuration, and on the nature of the adjacent monosaccharide units. The latter is in fact sequencing of a chain and is of especial importance. Such data amount to a complete structural analysis of regular polysaccharides.[6] A new method for structural elucidation of polysaccharides rests on the computer-assisted analysis of ^{13}C-NMR spectra.

2 Application of ^{13}C-NMR for Structural Elucidation of Regular Polysaccharides

A. Theoretical Background.—To come to reliable conclusions on the primary structure of a polysaccharide from its ^{13}C-NMR spectrum it is obviously necessary to establish clear and explicit relationships between spectral parameters and elements of structure of the polysaccharide. It means that the well known spectral characteristics of individual monosaccharides are to be correlated with those of the same monosaccharide units of a polysaccharide chain. It became evident, several years ago, that chemical shifts in the carbon signal of a monosaccharide unit involved in a chain differ from those of a free monosaccharide.[7] This means that adjacent monosaccharide units as well as other *O*-substituents affect chemical shifts of the monosaccharide under consideration and an understanding of the regularities of this effect could allow one to determine the nature of adjacent monosaccharide units in a polysaccharide chain, and thus solve the fundamental question of the sequence of monosaccharide units in the chain. These regularities were in fact disclosed[8] and employed thereafter in practical work on the structure elucidation of microbial polysaccharides and other regular polysaccharides.

To this end a new spectral parameter—the effect of glycosylation—was

[2] N. K. Kochetkov, N. S. Wulfson, O. S. Chizhov, and B. M. Zolotarev, *Tetrahedron*, 1963, **19**, 2209.
[3] N. K. Kochetkov and O. S. Chizhov, *Adv. Carbohydr. Chem.*, 1966, **21**, 39.
[4] P.-E. Jansson, L. Kenne, H. Liedgren, B. Lindberg, and J. Lonngren, *Chem. Commun. Univ. Stockholm*, 1976, **N8**, 1.
[5] See, for example, A. Dell, *Adv. Carbohydr. Chem. Biochem.*, 1987, **45**, 19.
[6] N. K. Kochetkov, A. S. Shashkov, G. M. Lipkind, and Yu. A. Knirel, *Sov. Sci. Rev., B. Chem.*, 1988, **13** (part 2), 1–72.
[7] See, for example, A. S. Shashkov, *Bioorg. Khim.* (*USSR*), *1983*, **9**, 246.
[8] N. K. Kochetkov, O. S. Chizhov, and A. S. Shashkov, *Carbohydr. Res.*, 1984, **133**, 173.

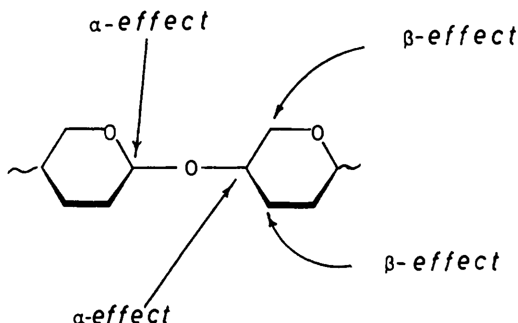


Figure 1

introduced to supplement chemical shift. The effect of glycosylation is the difference between the chemical shift of a given carbon atom in a substituted monosaccharide unit and that in a free monosaccharide. Thus, the chemical shifts are characteristic of the monosaccharide residue, the nature and orientation of *C*-substituents (for example, acetamido-group, deoxy unit, *etc.*), and its ring size. On the contrary, the effect of glycosylation characterizes *O*-substituents of the monosaccharide, that is the nature of the adjacent monosaccharide unit.

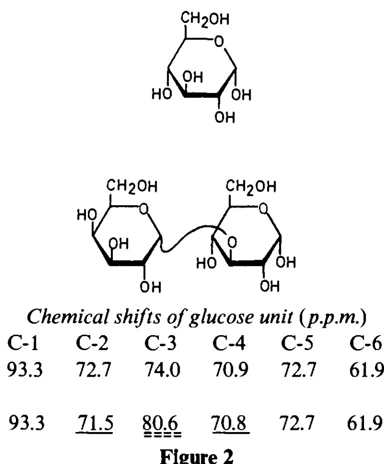
Fundamentally important was that the effects of glycosylation typical of a disaccharide remained practically unaltered when the disaccharide unit constituted a fragment of a more complex oligosaccharide or a polymer. It means that the glycosylation effect is influenced only by the nearest neighbour in the chain whereas more remote monosaccharide units have practically no affect on chemical shifts of carbon signals of a given monosaccharide unit. This allowed us to use the effect of glycosylation revealed from analysis of ^{13}C-NMR spectra of numerous disaccharides which differed in the structure of monosaccharide components and configuration of the glycosidic linkage.

It turned out that the most pronounced effects were usually observed for the carbons in closest proximity to the glycosidic linkage, that is for C-1′ of the glycone, for the aglycone carbon bearing the substituent (the so-called α-effect), and for the neighbouring carbon atoms of aglycone adjacent to the substitution site (β-effects) (Figure 1).

This can be seen for example in the glycosylation effects for glucose upon introduction of a galactose substituent to O-3 (Figure 2)

In brief, the reasons for the appearance of the glycosylation effects and their relationship to the neighbouring glycosidic linkage are as follows. Inductive factors, which strongly affect chemical shifts, are less influential for the glycosylation effects. Much more important is the presence or the absence of inter-unit non-bonded proton–proton interactions.[9] According to Grant and Cheney,[10] such interactions cause the polarization of C–H bonds and hence a change in the paramagnetic component of the shielding of the ^{13}C-NMR nuclei.

[9] G. M. Lipkind and N. K. Kochetkov, *Bioorg. Khim.* (*USSR*), 1983, **9**, 407
[10] D. M. Grant and B. V. Cheney, *J. Am Chem. Soc.*, 1967, **89**, 5315.



Chemical shifts of glucose unit (p.p.m.)

C-1	C-2	C-3	C-4	C-5	C-6
93.3	72.7	74.0	70.9	72.7	61.9
93.3	71.5	80.6	70.8	72.7	61.9

Figure 2

As can be seen, non-bonded interactions are possible between a restricted number of protons. These are H′-proton at the anomeric centre of the glycone, on the one hand, and, on the other, H_n-proton linked to the aglycone carbon atom involved in the glycosidic bond (and this determines the α-effect of glycosylation), and two protons H_{n-1} and H_{n+1} at the adjacent carbon atoms responsible for the β-effects. What is obvious is that the configuration of the glycosidic linkage and relative absolute configurations of both units are critically important for non-bonded proton–proton interactions and hence for the magnitude of the glycosylation effects.

Of especially great importance for monosaccharide sequencing of a polysaccharide chain is the dependence of the glycosylation effects on the spatial structure of an adjacent monosaccharide unit. As is known from theoretical calculations using the 'hard sphere' method,[11] the aglycone portion of a disaccharide in its most favourable conformation is usually *gauche* oriented with respect to the C-1′–O.5′ bond of the glycone, the torsion angle φ being close to $-60°$ for the α- and $+60°$ for the β-configuration of the glycosidic bond.[12] Conformations with a φ-value above $+60°$ and less than $-60°$ are unfavourable.[13]

For disaccharide in these most favourable conformations the distance between H′ of the glycone and H_n of the aglycone is in the range 2.2—3.6 Å that allows rather strong non-bonded interactions. On the other hand, spatial proximity of H-1′ of the glycone and H_{n+1} and H_{n-1} β-atoms of the aglycone can only be realized for their equatorial orientation (Figure 3). Thus an important conclusion is reached in that stereochemical features of the aglycone monosaccharide unit (including those belonging to D- or L-series) and the stereochemistry of the glycosidic linkage affect proton–proton interactions and consequently the mag-

[11] U. Burkett and N. L. Allinger, 'Molecular Mechanics', ACS Monograph, No. 177, American Chemical Society, Washington, 1982.
[12] R. U. Lemieux and S. Koto, *Tetrahedron*, 1974, **30**, 1933.
[13] G. M. Lipkind, V. S. Verovsky, and N. K. Kochetkov, *Carbohydr. Res.*, 1987, **133**, 1.

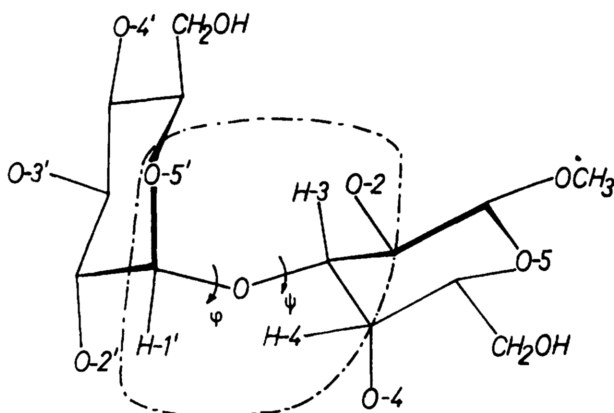
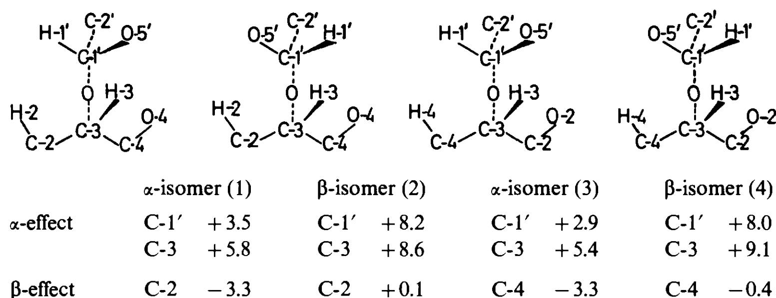


Figure 3



	α-isomer (1)		β-isomer (2)		α-isomer (3)		β-isomer (4)	
α-effect	C-1′	+3.5	C-1′	+8.2	C-1′	+2.9	C-1′	+8.0
	C-3	+5.8	C-3	+8.6	C-3	+5.4	C-3	+9.1
β-effect	C-2	−3.3	C-2	+0.1	C-4	−3.3	C-4	−0.4

Figure 4

nitude of the glycosylation effect. In other words, the effect of glycosylation depends on the stereochemistry of the adjacent monosaccharide unit, that is on its structure.

As an example, two closely related disaccharides containing one equatorial proton at C-2 near the glycosidic linkage region, α-D-Glc-(1→3)-L-Rha (1) and β-D-Glc-(1→3)-L-Rha (2), can be considered.[14–16] The spatial arrangement of atoms near the glycosidic linkage in (1) and (2) (cross-section is made in the plane of C1′–O–C3) is shown in Figure 4. It can easily be seen that in (1) the protons H-1′ and H-3 (α-atoms) are rather remote and the interaction between them is relatively weak. In contrast, H-1′ and H-2 (β-atoms) are *cis* with respect to the plane and interact rather strongly. The opposite situation is with (2). The protons H-1′ and H-3 are closer to each other than in the former case while H-1′ and H-2 are *trans* and more remote. This spatial arrangement, which followed from theoretical conformational analysis, is directly proved by experimental

[14] G. M. Lipkind, S. S. Mamyan, A. S. Shashkov, O. A. Nechaev, V. I. Torgov, V. N. Shibaev, and N. K. Kochetkov, *Bioorg. Khim.* (*USSR*), 1988, **14**, 340.
[15] A. S. Shashkov, G. M. Lipkind, Yu. A. Knirel, and N. K. Kochetkov, *Magn. Res. Chem.*, 1988, **26**, 735.
[16] G. M. Lipkind, A. S. Shashkov, S. S. Mamyan, and N. K. Kochetkov, *Carbohydr. Res.*, 1981, **181**, 1.

determination of the nuclear Overhauser effects (NOE). NOE Experiments have revealed short distances between H-1′ and H-2 in (1) and between H-1′ and H-3 in (2) while the distances between H-1′ and H-3 in the former and between H-1′ and H-2 in the latter are larger. The glycosylation effects are in accord with the above findings. Thus, for the disaccharide (1) rather small α-effects and a relatively large β-effect on C-2 were observed while for (2) the β-effect was much smaller in absolute value along with much larger α-effects. Direct correlation between the effects of glycosylation and the configuration of the glycosidic linkage is obvious.

Now we can change the stereochemistry of the aglycone and compare another pair of related disaccharides, α-D-Glc-(1→3)-D-Gal (3) and β-D-Glc-(1→3)-D-Gal (4),[14–16] the aglycones of which bear equatorial protons on C-4. NOE experiments[14] and theoretical calculations indicate again *cis*-arrangement of H-1′ and H-4 in (3) and their *trans*-arrangement in (4), *i.e.* relative proximity of H-1′ and H-3 in the latter and remoteness in the former. In full accord with this disposition of the protons are the observed effects of glycosylation: α-effects in (3) are smaller than in (4) while much larger (in absolute value) β-effects were typical for (3). The comparison between (1), (2), and (3), (4) demonstrate the influence of stereochemical features of the aglycone monosaccharide on the value of glycosylation effects.

Analogous analysis was applied to other disaccharide systems with common monosaccharide constituents (glucose, galactose, mannose, rhamnose, 2-acetamido-2-deoxy-sugars) linked with 1→4, 1→3, and 1→2 α- or β-glycosidic linkages. It turned out that in all disaccharides the magnitude of the glycosylation effects is clearly associated with configuration of the glycosidic linkage, and depends specifically on the nature, or more correctly, stereochemistry of the adjacent monosaccharide unit.

Most important is that the reverse statement is also valid: the stereochemistry of an aglycone portion of a disaccharide can reliably be established from the glycosylation effects. It is also essential that the glycosylation effects are additive and for a given carbon atom contribute the components which reflect non-bonded interactions between one or another pair of protons.

It means that for practical application of ^{13}C-NMR spectroscopy we need a knowledge of the glycosylation effects in all the common disaccharide systems. The values of these effects were retrieved from literature and estimated from our own spectral data for a number of disaccharides synthesised specially. As an example data for disaccharides with an aglycone of gluco-configuration are listed in Table 1.[17]

With analogous, well specified, data on the effects of glycosylation for other aglycone types[17] and from known data on the chemical shifts of carbon signals for the common aldoses at our disposal[17] we can calculate, with sufficient accuracy, a ^{13}C-NMR spectrum of a regular linear polysaccharide, including microbial ones. The calculation can be carried out according to an additive

[17] G. M. Lipkind, A. S. Shashkov, Yu. A. Knirel, E. V. Vinogradov, N. K. Kochetkov, *Carbohydr. Res.*, 1988, **175**, 59.

Table 1 *Effects of glycosylation on chemical shifts of* [13]*C resonances for Glc-residue in Hex-Glc disaccharides*

Linkage type	*Glc* C-1	C-2	C-3	C-4	C-5	C-6	*Hex* C-1′
α-(1→2)-α	−1.8	+0.4	−1.3	0	0	0	+3.4
α-(1→2)-β	+0.3	+4.7	−1.3	0	0	0	+5.6
α-(1→3)	0	−1.2	+6.8	−0.1	0	0	+6.7
α-(1→4)	0	0	+0.3	+7.4	−1.3	0	+7.6
α-(1→6)	0	0	0	0	−1.8	+5.0	+5.6
β-(1→2)-α	−0.7	+0.9	−1.3	0	0	0	+7.6
β-(1→2)-β	−1.6	+7.5	0	0	0	0	+7.0
β-(1→3)	0	−0.7	+9.5	−1.5	0	0	+6.7
β-(1→4)	0	0	−1.4	+9.0	−1.3	−0.8	+6.3
β-(1→6)	0	0	0	0	−1.0	+7.8	+6.7

scheme for each carbon atom of each monosaccharide constituent of a regular polysaccharide, or, more precisely, of its repeating oligosaccharide unit. The following simple formula (1) is used

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

where l is the number of the carbon atom, $\delta_0(l)$ and $\delta(l)$ are the ^{13}C chemical shifts for a free monosaccharide and for this monosaccharide as a polysaccharide constituent, respectively; $A(k,l)$ is the effect of glycosylation caused by the glycosylation of the unit (the index of linkage type k); $B(k',l)$ is the effect caused by formation of the glycosidic linkage (the index of linkage type k').

The thus calculated spectrum can be compared with the experimental spectrum of the polysaccharide under investigation in order to arrive at a conclusion on its structure. On the basis of this approach a computer-assisted method was developed in our laboratory and used for structural elucidation of linear regular microbial polysaccharides. The only data required are on their monosaccharide composition, demanding a minimum amount of substance, and their ^{13}C-NMR spectral data.

B. Computer-assisted Analysis of ^{13}C-NMR Spectra.—The proposed procedure involves three steps. The first is the generation of all possible structures of a polysaccharide (or, more correctly, its repeating unit) of a given monosaccharide composition. This implies permutations of the monosaccharide constituents in the repeating unit as well as types and configurations of the glycosidic linkages. The number of all possible structures to be considered can generally be very large and for a tetrasaccharide repeating unit, for example, is about 25000. In some cases preliminary consideration of ^{13}C-NMR spectra can be used to rule out *a priori* some structures. Thus, for example, we can judge on the presence of 1→6 linkages from the presence of signals at 65—71 p.p.m. while signals in the region of 60—63 p.p.m. points to the absence of these linkages.

The second step is the evaluation of the ^{13}C-NMR spectra for each of the

Table 2 *Computer-derived structures for a* Salmonella newington *polysaccharide*

		S
(A)	→6)-β-D-Manp-(1→4)-α-L-Rhap-(1→3)-β-D-Galp-(1→	0.3
(B)	→6)-β-D-Manp-(1→3)-α-L-Rhap-(1→3)-β-D-Galp-(1→	1.3
(C)	→6)-β-D-Manp-(1→3)-β-D-Galp-(1→3)-α-L-Rhap-(1→	1.8

structures generated in the first step on the basis of the chemical shift and the glycosylation effect data using the formula and tables indicated before. The evaluated chemical shifts for each monosaccharide constituent are then arranged in descending order for purposes of comparison.

The third step is the computerized search of a structure whose calculated spectrum is closest to the experimental spectrum of the polysaccharide. To this end the squared deviations (Δ^2) are determined for the chemical shifts of the signals with identical numbers, followed by their sum ($\Sigma\Delta^2$), and the normalized value related to one monosaccharide unit ($S = \Sigma\Delta^2/n$), where n is the number of monosaccharide residues in the repeating unit. Our experience has shown that it is the structures with S values not more than 1.5—2 that should generally be taken into consideration.

The programme for calculation was developed using a BESM-6 computer (USSR) and the languages ALGOL-60 and FORTRAN.

C. Application for Structure Elucidation of Microbial Polysaccharides.—The computer-assisted method for the determination of structures was first verified on some O-antigenic polysaccharides of gram-negative bacteria with known structures. An example is given for O-antigenic polysaccharide of *Salmonella newington* with trisaccharide repeating units containing mannose, rhamnose, and galactose. It can be seen from Table 2[18] that the structure A with the lowest S value (0.3) corresponds to the real structure. The closest S values were obtained for structures B (S = 1.3) and C (S = 1.8) which differ from A in only one structural element. This demonstrates the sensitivity of the method.

In the more complex polysaccharide of *Shigella flexneri* variant Y, whose tetrasaccharide repeating unit consists of three rhamnose and one *N*-acetylglucosamine residues, eight structures with S values not exceeding 1.5 were evaluated (Table 3).[17] Again the structure with the smallest S value of 0.3 corresponds to the real structure. It seems that a large number of calculated structures which differ rather insignificantly from the real one may be accounted for by the 'monotonous' character of the repeating unit and by the fact that the ^{13}C resonances for a rhamnose unit glycosylated at position 2, 3, or 4 are similar. At the same time, these facts demonstrate the high reliability of the method.

This new approach has already found wide application in our laboratory for structural analysis of unbranched polysaccharides of microbial origin.

Two more examples of structures determined by means of this computer-

[18] G. M. Lipkind, A. S. Shashkov, and N. K. Kochetkov, *Bioorg. Khim.* (*USSR*). 1987, **13**, 833.

Table 3 *Computer-derived structures for a* Shigella flexneri *variant Y polysaccharide*

Structure	*S*
→2)-α-L-Rhap-(1→2)-α-L-Rhap-(1→3)-α-L-Rhap-(1→3)-β-D-GlcpNAc-(1→	0.3
→2)-α-L-Rhap-(1→3)-α-L-Rhap-(1→3)-α-L-Rhap-(1→3)-β-D-GlcpNAc-(1→	0.7
→2)-α-L-Rhap-(1→3)-α-L-Rhap-(1→2)-α-L-Rhap-(1→3)-β-D-GlcpNAc-(1→	0.8
→3)-α-L-Rhap-(1→2)-α-L-Rhap-(1→3)-α-L-Rhap-(1→3)-β-D-GlcpNAc-(1→	1.0
→2)-α-L-Rhap-(1→2)-α-L-Rhap-(1→2)-α-L-Rhap-(1→3)-β-D-GlcpNAc-(1→	1.1
→4)-α-L-Rhap-(1→2)-α-L-Rhap-(1→3)-α-L-Rhap-(1→4)-β-D-GlcpNAc-(1→	1.1
→2)-α-L-Rhap-(1→2)-α-L-Rhap-(1→4)-α-L-Rhap-(1→3)-β-D-GlcpNAc-(1→	1.2
→3)-α-L-Rhap-(1→2)-α-L-Rhap-(1→2)-α-L-Rhap-(1→3)-β-D-GlcpNAc-(1→	1.5

assisted method can be presented. The repeating unit of the polysaccharide of *Proteus vulgaris* O19 is a tetrasaccharide containing galactose, *N*-acetylglucosamine, *N*-acetylgalactosamine, and *N*-acetyl-L-fucosamine. Of five structures (Table 4)[17] with *S* values of below 2 evaluated by the computer, only one with $S = 0.7$ proved to be consistent with the results of methylation analysis and corresponds to the real structure.

The polysaccharide of *Yersinia frideracsenii* 16,29 was known to be a linear rhamnan with trisaccharide repeating units. It was unknown, however, whether all three rhamnose residues belong to the rare, for this sugar, D-series or whether one of them is the L-isomer. Using data on the chemical shifts and the effects of glycosylation varieties were evaluated. The structures shown in Table 5[19] refer to those with *S* values below 2 for rhamnans both built up of only D-rhamnose and comprising one L-unit. The smallest deviation of 0.5 was found for a structure which was the only one to fit the results of methylation analysis. Thus both structural and configurational questions were answered.

Successful application of this approach might also be exemplified by polysaccharides with tetra- and penta-saccharide repeating units. Data accumulated clearly show broad possibilities of the method so it can be regarded as very useful for elucidation of the structure of linear polysaccharides.

Our efforts are directed now to extend its applicability to branched structures and this is a much more difficult task. The point is that the additivity of the effects of glycosylation is disturbed for monosaccharide units which are in the branch point and are substituted at vicinal 2,3- or 3,4-positions. Additional increments for the glycosylation effects are required in this case. Since theoretical conformational analysis of such branched trisaccharides is only just beginning, experimental evaluation of these increments requires careful spectral analysis of a large number (at least 40) of synthetic, vicinally disubstituted trisaccharides. This work is now in progress in our laboratory and first results show that the approach demonstrated here will be useful for analysis of branched structures as well, as can be illustrated with the first example of such structural analysis. This is the structural elucidation of the *Salmonella arizonae* O63 polysaccharide built up of pentasaccharide repeating units consisting of two *N*-acetylgalactosamine residues and one of each galactose, glucose, and 3-acetamido-3,6-dideoxy-D-

[19] V. A. Zubkov, V. V. Isakov, R. P. Gorshkova, and Yu. S. Ovodov, *Bioorg. Khim* (*USSR*), in press.

Table 4 *Computer-derived structures for a* Proteus vulgaris O19 *polysaccharide*

	S
→3)-α-D-Galp-(1→4)-α-D-GalpNAc-(1→3)-α-L-FucpNAc-(1→3)-β-D-GlcpNAc-(1→	0.7
→2)-α-D-Galp-(1→4)-α-D-GalpNAc-(1→3)-α-L-FucpNAc-(1→3)-β-D-GlcpNAc-(1→	1.0
→4)-α-D-Galp-(1→4)-α-D-GalpNAc-(1→3)-α-L-FucpNAc-(1→3)-β-D-GlcpNAc-(1→	1.1
→4)-α-D-Galp-(1→4)-α-D-GalpNAc-(1→3)-α-L-FucpNAc-(1→4)-β-D-GlcpNAc-(1→	1.6
→3)-α-D-Galp-(1→4)-α-D-GalpNAc-(1→3)-α-L-FucpNAc-(1→4)-β-D-GlcpNAc-(1→	1.9

Table 5 *Computer-derived structures for a* Yersinia friderecsenii 16,29 *polysaccharide*

D,D,D	*S*
→3)-α-D-Rhap-(1→2)-α-D-Rhap-(1→3)-β-D-Rhap-(1→	0.5
→3)-α-D-Rhap-(1→3)-α-D-Rhap-(1→4)-β-D-Rhap-(1→	0.7
→3)-α-D-Rhap-(1→3)-α-D-Rhap-(1→3)-β-D-Rhap-(1→	1.2
→3)-α-D-Rhap-(1→2)-α-D-Rhap-(1→4)-β-D-Rhap-(1→	1.3
→3)-α-D-Rhap-(1→4)-α-D-Rhap-(1→2)-β-D-Rhap-(1→	1.3
→2)-α-D-Rhap-(1→3)-α-D-Rhap-(1→3)-β-D-Rhap-(1→	1.6
D,D,L	
→3)-α-D-Rhap-(1→3)-α-D-Rhap-(1→3)-β-L-Rhap-(1→	0.8
→2)-α-D-Rhap-(1→3)-α-D-Rhap-(1→3)-β-L-Rhap-(1→	1.0
→4)-α-D-Rhap-(1→3)-α-D-Rhap-(1→3)-β-L-Rhap-(1→	0.8

galactose (3-acetamido-fucose). From methylation analysis it is known that the latter sugar is terminal in the branch and galactose is a branching point, but how many monosaccharide units are in the main chain and how many in the branch is unknown. By computer generation of all the possible structures their spectra were calculated, taking into account special increments. These increments were evaluated from spectral data of four synthetic branched trisaccharides and are shown in Table 6.[20] For calculation the modified formula (2) has been used which takes into account the values of these new increments.

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

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

The result of a computerized search of possible structures (a few from several thousands) is shown in Table 7. One can see again that one of these structures has $S = 0.9$ and others have greater S values. It turned out that the structure with the smallest S value again corresponds to the real structure of the polysaccharide of *Salmonella arizonae* O63 which was estimated by classical methods.

[20] G. M. Lipkind, A. S. Shashkov, and N. K. Kochetkov, *Carbohydr. Res.*, 1989 (in press).

Table 6 *Deviation of additivity of the glycosylation effects D and F (p.p.m.) in a branching unit for HexA*(1→3)-[*HexB*(1→4)]-β-D-*Gal-OMe*

Type linkage		*D*						*F**	
k	*k″*	*for Gal resuidue*						*HexA*	*HexB*
		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

* F — corrections for C-1 of HexA and HexB residues

It seems that the computer-assisted method for elucidation of regular polysaccharides is of great practical value for linear structures and has good prospects for branched polymers, although a good deal of work is to be done. Rapid progress in information science, computerization, and calculation methods for complex polymers will possibly enable future solutions of the next most general problem, that is the calculation of spectra of irregular polysaccharide structures.

3 The Synthesis of Regular Polysaccharides

According to the classical academic statement, the structure of a natural substance is considered completely elucidated when its chemical synthesis is carried out. The synthesis of complex natural products influenced greatly the development of general synthetic chemistry and is an essential element of the chemistry and biochemistry of biologically active substances. The chemistry of biopolymers is not an exception. Syntheses of numerous fragments, structural analogues, and, finally, of the very natural biopolymers were of outstanding value as is evident from the development of the chemistry of proteins and nucleic acids. Problems associated with the synthesis of regular polysaccharides including microbial ones have been discussed many times (see, for example, ref. 21). Efficient polymerization or polycondensation procedures, that can be regarded as the only realistic methods for preparation of a polysaccharide chain (for discussion of this problem see[22]), face the main difficulty which is to reach total, or at least a high degree, of stereospecificity of the process. This should open the route to not only regioregular but also to stereoregular polysaccharide chains; the latter demand is much more difficult.

This problem can satisfactorily be solved by the so-called trityl-cyanoethylidene polycondensation which was developed in our laboratory as a rather general method for the synthesis of regular homo- and hetero-polysaccharides with mono- or, respectively, oligo-saccharide units linked through 1,2-*trans*-glycosidic linkages.

[21] N. K. Kochetkov, *Izv. Akad. Nauk SSSR. Ser. Khim.*, 1982, 1543.
[22] N. K. Kochetkov, *ACS Symposium Series*, 1983, **231**, 65.

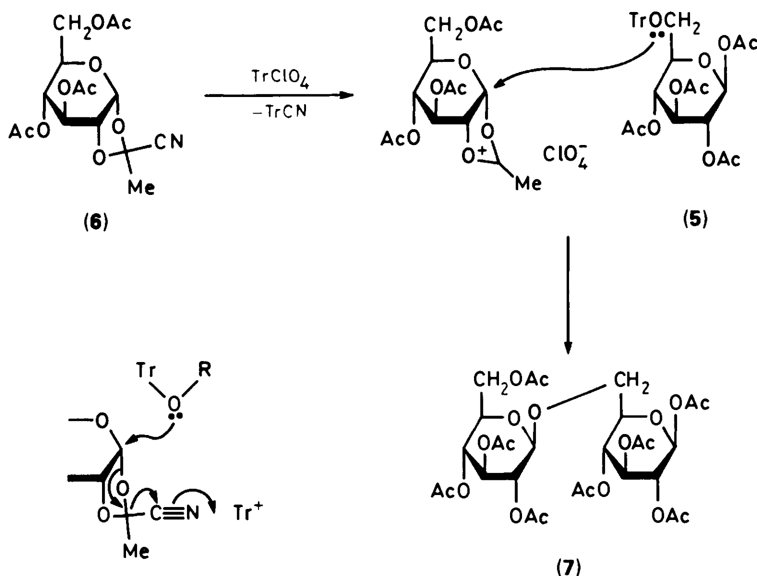
Table 7 *Computer-derived structures for a* Salmonella arizonae O63 *polysaccharide*

```
                                                                          S
→4)-α-GalNAc-(1→3)-β-GalNAc-(1→3)-β-Gal-(1→4)-α-Glc-(1→
                                        4
                                        ↑
                                        1
                                     α-Fuc3NAc                           0.9

→3)β-GalNAc-(1→3)-β-Gal-(1→4)-α-Glc-(1→
                      4
                      ↑
                      1
                   α-GalNAc
                      4
                      ↑
                      1
                   α-Fuc3NAc                                             1.6

→4)-α-GalNAc-(1→4)-β-Gal-(1→
                      3
                      ↑
                      1
                   β-GalNAc
                      3
                      ↑
                      1
                   α-Glc
                      4
                      ↑
                      1
                   α-Fuc3NAc                                             2.5

               →3)-β-Gal-(1→
                      4
                      ↑
                      1
                   β-GalNAc
                      3
                      ↑
                      1
                   α-GalNAc                                              2.4
                      4
                      ↑
                      1
                   α-Glc
                      4
                      ↑
                      1
                   α-Fuc3NAc
```

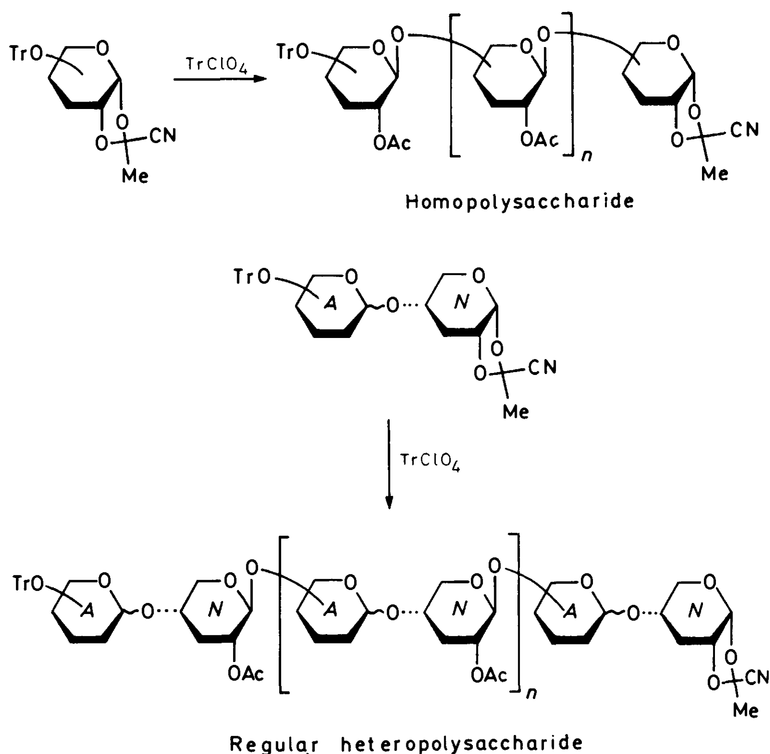


Scheme 1

A. Trityl-Cyanoethylidene Polycondensation.—The new reaction of formation of glycosidic linkage consists in glycosylation of sugar trityl ethers (5) with sugar 1,2-*O*-cyanoethylidene derivatives (6) initiated by triphenylmethylium ion.[23] The formation of the glycosidic linkage may be regarded as a result of nucleophilic attack of the trityl ether oxygen with enhanced nucleophilicity relative to acetoxy oxygens, on the anomeric centre of the cyanoethylidene derivative (Scheme 1). This reaction is believed to follow a concerted push–pull mechanism and results, for steric reasons, in the formation of a disaccharide (7) with a 1,2-*trans*-glycosidic linkage. In the majority of cases the reaction is completely stereospecific, although examples are also known of the limited formation of 1,2-*cis*-glycosidic linkages. Several years ago we demonstrated that this could be corrected by performing glycosylation under high pressure which results in the stereospecific formation of 1,2-*trans*-glycosidic linkage.[24]

The trityl-cyanoethylidene condensation served as the basis for elaboration of a novel general polycondensation reaction.[25] Indeed, if the reaction is carried out with a derivative bearing both the cyanoethylidene and *O*-trityl groups it results in the formation of a chain with monomeric units stereospecifically linked by 1,2-*trans*-glycosidic linkages. Starting from a monosaccharide monomer a

[23] A. F. Bochkov and N. K. Kochetkov, *Carbohydr. Res.*, 1975, **39**, 355.
[24] 24 N. K. Kochetkov, V. M. Zhulin, E. M. Klimov, N. N. Malysheva, Z. G. Makarova, and A. Ya. Ott, *Carbohydr. Res.*, 1987, **164**, 241.
[25] A. F. Bochkov, I. V. Obruchnikov, V. M. Kalinevich, and N. K. Kochetkov, *Tetrahedron Lett.*, 1975, **16**, 3403.



Scheme 2

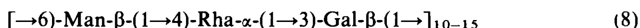
homopolysaccharide is obviously produced. In the case of oligosaccharide as a monomer the product is a regular heteropolysaccharide built up of repeating units (Scheme 2).

The trityl-cyanoethylidene polycondensation made possible the synthesis of a broad range of polysaccharides or higher oligosaccharides of regular structure.[26] This method was also successfully applied to the first chemical syntheses of natural heteropolysaccharides of microbial origin, namely, some O-antigenic polysaccharides and capsular polysaccharides. For the synthesis of these polysaccharides it is necessary to perform polycondensation of a monomer which should correspond to the repeating oligosaccharide unit and be properly functionalized, namely, contain the cyanoethylidene group at the 'reducing' terminus and a tritylated hydroxyl group at the specific position which determines the linkage site between oligosaccharide units. All other hydroxyl groups should be protected. Depending on the location of the O-trityl group the linear or branched polysaccharide chain is produced. All polycondensation reactions were performed

[26] N. K. Kochetkov, *Tetrahedron*, 1987, **43**, 2389.

under standard conditions: at room temperature in dichloromethane in the presence of 10% triphenylmethylium perchlorate. The greatest difficulty in all cases was the synthesis of the complicated oligosaccharide monomers.

B. Synthesis of Microbial Polysaccharides.—Nine years ago the first synthesis of a microbial polysaccharide, namely of O-antigenic polysaccharide *Salmonella newington* was accomplished.[27] This polysaccharide (8) is built up of trisaccharide repeating units composed of mannose, rhamnose, and galactose, the degree of polymerization of the natural biopolymer being about 10 to 15 oligosaccharide units.[28]



$[\rightarrow 6)\text{-Man-}\beta\text{-}(1\rightarrow 4)\text{-Rha-}\alpha\text{-}(1\rightarrow 3)\text{-Gal-}\beta\text{-}(1\rightarrow]_{10-15}$ (8)

$\text{Man-}\beta\text{-}(1\rightarrow 4)\text{-Rha-}\alpha\text{-}(1\rightarrow 3)\text{-Gal}$ (9)

Earlier we synthesized the trisaccharide (9), and at that time it was the first synthesis of a repeating unit of a microbial polysaccharide.[29] Later this synthesis was modified and made more practical[30] when the required trisaccharide peracetate (10) was obtained in a yield of 60%. Conversion of this peracetate into the monomer implied introduction of the 1,2-*O*-cyanoethylidene group and tritylation at O-6 of the mannose residue. To this end (10) was converted into the corresponding glycosyl bromide and then into the deacylated 1,2-O-cyanoethylidene derivative (11). The next step, selective tritylation, was the weak point due to the presence of two primary hydroxyl groups in (11). Although tritylation did produce a mixture of three trityl ethers the necessary isomer (12) has fortunately been isolated, albeit in moderate yield (Scheme 3).

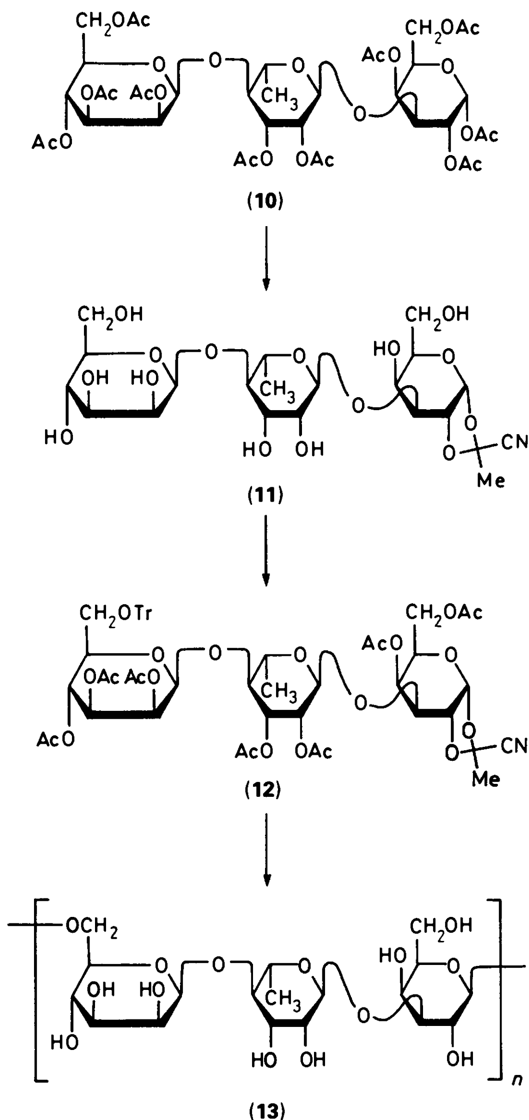
Polycondensation of (12) was accomplished under the standard conditions indicated before, the reaction was quenched with aqueous trifluoroacetic acid, the polymeric product was isolated by chromatography, subjected to Zemplen deacetylation and the synthetic polysaccharide (13), as a mixture of polymer-homologues, was fractionated on Biogel P-4 to give a high-molecular-mass fraction in a yield of 34%.[27] Conventional methods have shown its complete regioregularity and a degree of polymerization of about 10, that is equal to 30 monosaccharide units and a molecular mass of about 4 500. The most critical point was the question of the configuration of the newly formed galactosyl-mannose linkage produced upon polycondensation. That this linkage was β followed from the ^{13}C-NMR spectrum which contained a signal at 104.3 ppm typical for C-1 of β-D-galactopyranosides, while any signal for α-galactose was absent. Final proof of this essential point was achieved by Smith degradation of synthetic (13). This resulted in oxidation of the rhamnose and mannose units

[27] N. K. Kochetkov, V. I. Betaneli, M. V. Ovchinnikov, and L. V. Backinowsky, *Tetrahedron*, 1981, **37**, Suppl. 9, 149.

[28] C. G. Hellerqvist, B. Lindberg, J. Lonngren, and A. A. Lindberg, *Acta Chem. Scand.*, 1971, **25**, 939.

[29] N. K. Kochetkov, B. A. Dmitriev, O. S. Chizhov, E. M. Klimov, N. N. Malysheva, A. Ya. Chernyak, N. E. Byramova, and V. I. Torgov, *Carbohydr. Res.*, 1974, **33**, C5.

[30] V. I. Betaneli, M. V. Ovchinnikov, L. V. Backinowsky, and N. K. Kochetkov, *Carbohydr. Res.*, 1980, **84**, 211.



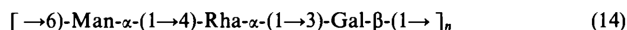
Scheme 3

whilst 3-substituted galactose survived to give galactopyranosyl-glycerol. This was compared with authentic samples of α- and β-galactosyl-glycerol and shown unequivocally to be pure β-anomer. It thus follows that the synthetic polysaccharide contained only β-galactosidic linkage and that the polycondensation was completely stereospecific.

Table 8 *Inhibiting activity of synthetic oligo- and poly-saccharides*

	Minimal inhibitor dose (μg) *in passive haemagglutination reaction*	
	Antiserum	
Inhibitor	3,15	3,19
Man-β-(1→4)-Rha-α-(1→3)-Gal (9)	1.6	1.0
–[→6)-Man-β-(1→4)-Rha-α-(1→3)-Gal-β-(1→]$_n$ (13)	0.2	0.2
–[→6)-Man-α-(1→4)-Rha-α-(1→3)-Gal-β-(1→]$_n$ (14)	100	25

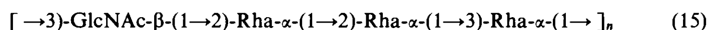
Biological activity of the synthetic polysaccharide and, what is even more essential, its immunochemical specificity, has also been investigated. To this end the synthesis has been undertaken of an analogue (14) of the natural polysaccharide differing only in the configuration of the mannose-rhamnose linkage.[27]



(14)

The results of inhibition of the passive haemagglutination (Table 8) have shown that the synthetic β-mannose-containing polysaccharide (13) is eight times as active as the corresponding trisaccharide repeating unit (9) while the α-analogue (14) is practically inactive. This shows, unequivocally, the high specificity of the former synthetic polysaccharide (13). That the activity of synthetic (13) is lower than that of the natural O-antigenic polysaccharide is certainly accounted for by its lower molecular mass.

The next synthetic target was a more complex O-antigenic polysaccharide of *Shigella flexneri* built up of tetrasaccharide repeating units comprising one *N*-acetylglucosamine and three rhamnose residues (15).[31]



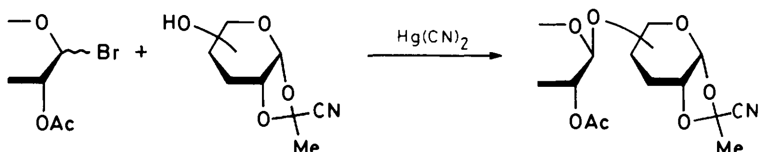
(15)

Here again the main difficulty was associated with the synthesis of the properly functionalized monomer, *viz.*, a tetrasaccharide with a trityl group at O-3 of the glucosamine and cyanoethylidene group in the terminal 3-substituted rhamnose and the approach to the synthesis of a monomer employed earlier was inapplicable here.

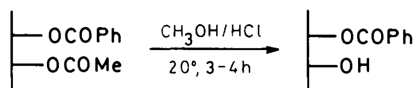
This forced us to develop another, more flexible strategy to the synthesis of complex monomers of this type which involves an assembly of a monomer from the blocks already bearing the required functions and suitable protective groups. This strategy became possible due to two important findings. It turned out that the hydroxyl-containing cyanoethylidene derivatives can be glycosylated themselves under Helferich glycosylation conditions without affecting the cyanoethylidene group (Scheme 4).[32] This allowed us to synthesize oligosaccharide

[31] L. Kenne, B. Lindberg, K. Petterson, E. Katzenellenbogen, E. Romanowska, *Eur. J. Biochem.*, 1978, **91**, 279.

[32] V. I. Betaneli, L. V. Backinowsky, N. E. Byramova, M. V. Ovchinnikov, M. M. Litvak, and N. K. Kochetkov, *Carbohydr. Res.*, 1983, **113**, C1.



Scheme 4



Scheme 5

derivatives from the lower precursors with the cyanoethylidene group preformed.

On the other hand, conditions were found[33] for highly selective removal of an O-acetyl group in the presence of O-benzoyl groups by mild acid-catalysed methanolysis (Scheme 5). This largely facilitated the 'play with protective groups' necessary to ensure regioselective glycosylation and functionalization which are more difficult the larger the oligosaccharide monomer to be obtained. The use of these two simplest and easily introducible acyl protective groups enabled elaboration of a convenient and standardized strategy for the assembly of the complex oligosaccharide monomer.

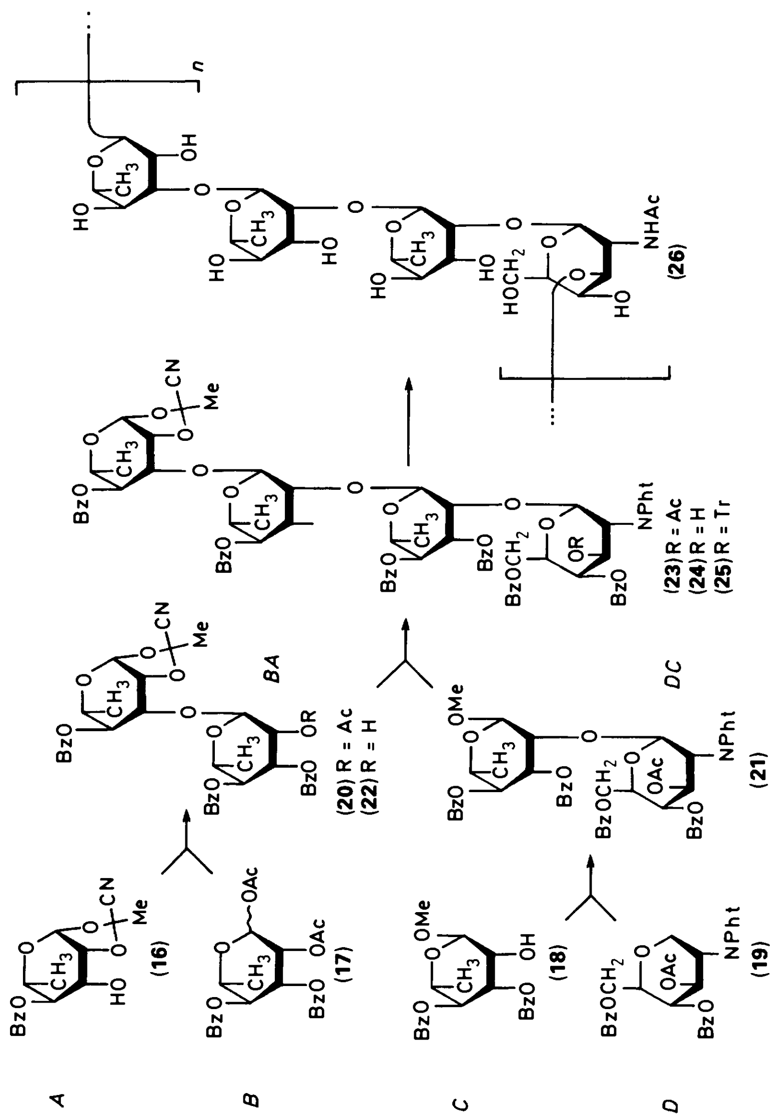
This is well seen from the scheme of synthesis of the *Shigella flexneri* polysaccharide (15) (Scheme 6).[34]

The starting monosaccharide synthons (16)—(19) for the fragments A, B, C, and D of the tetrasaccharide repeating unit, three derivatives of rhamnose and one of glucosamine, have been synthesized by conventional methods and contained a combination of the protective groups which enabled selective liberation of the necessary hydroxyl group at the proper moment.

The assembly of the tetrasaccharide monomer was accomplished by a blockwise (2 + 2) scheme and involved the synthesis of the BA and DC blocks followed by their condensation into the required tetrasaccharide. The synthon A (16) already bearing the cyanoethylidene group and a free hydroxyl at C-3 was glycosylated with a rhamnosyl bromide derived from (17) with benzoyl groups at O-3 and O-4 and an acetyl protection at O-2 to give the BA block derivative (20). For the synthesis of the DC block (18) with only one free hydroxyl group at C-2 was glycosylated with glucosaminyl bromide, obtained from (19), with a phthaloyl-protected amino group, two benzoates at O-6 and O-4, and acetyl protective group at O-3. This afforded a DC fragment (21) which was converted into the corresponding biosyl bromide and used as a glycosyl-donor to glycosylate the BA disaccharide derivative (22) with a free hydroxyl group; (22) was obtained by selective deprotection of (20) making use of the acid-catalysed

[33] N. E. Byramova, M. V. Ovchınnikov, L. V. Backınowsky, and N. K. Kochetkov, *Carbohydr. Res.*, 1983, **124**, C8.

[34] N. K. Kochetkov, N. E. Byramova, Yu. E. Tsvetkov, and L. V. Backınowsky, *Tetrahedron*, 1985, **41**, 3363.



Scheme 6

methanolysis of the only acetyl group (at C-2). The condensation of the bromide derived from (21) with (22) gave rise to derivative (23) which corresponds exactly to the repeating unit of the *Shigella flexneri* polysaccharide and already contains the cyanoethylidene group at the reducing terminus and the only acetyl group at O-3 of the glucosamine unit with all other hydroxyl groups protected as benzoates. After selective deprotection the monohydroxyl derivative (24) was formed. Tritylation of the secondary hydroxyl in (24) with triphenylmethylium perchlorate in the presence of 2,4,6-collidine[35] gave rise to the target monomer (25) which is the properly functionalized repeating unit of the *Shigella flexneri* polysaccharide.

This assembly scheme exemplified the synthesis of a rather complex monomer which proved to be very efficient and was successfully used in other syntheses. The monomer (25) was subjected to polycondensation under standard conditions to give 90% of the protected polysaccharide. After deprotection by hydrazinolysis followed by selective *N*-acetylation a synthetic polysaccharide (26) was obtained which has a completely identical ^{13}C-NMR spectra to the natural one isolated from *Shigella flexneri* variant Y. Spectral data, as well as methylation analysis and gel-chromatography, point to a molecular mass of the synthetic polysaccharide of about 6000, that is to an average degree of polymerization of about 10, close to the natural biopolymer. Immunochemical analysis of the synthetic sample has confirmed its high activity and specificity.[36]

The approach developed in the synthesis of complex heteropolysaccharides was next applied[37] to the preparation of a regular branched polysaccharide, the capsular polysaccharide of *Streptococcus pneumoniae* type 14, built up of repeating tetrasaccharide units with a backbone composed of *N*-acetylglucosamine, galactose, and glucose, the second galactose residue being attached to the *N*-acetylglucosamine as a side chain (27).[38]

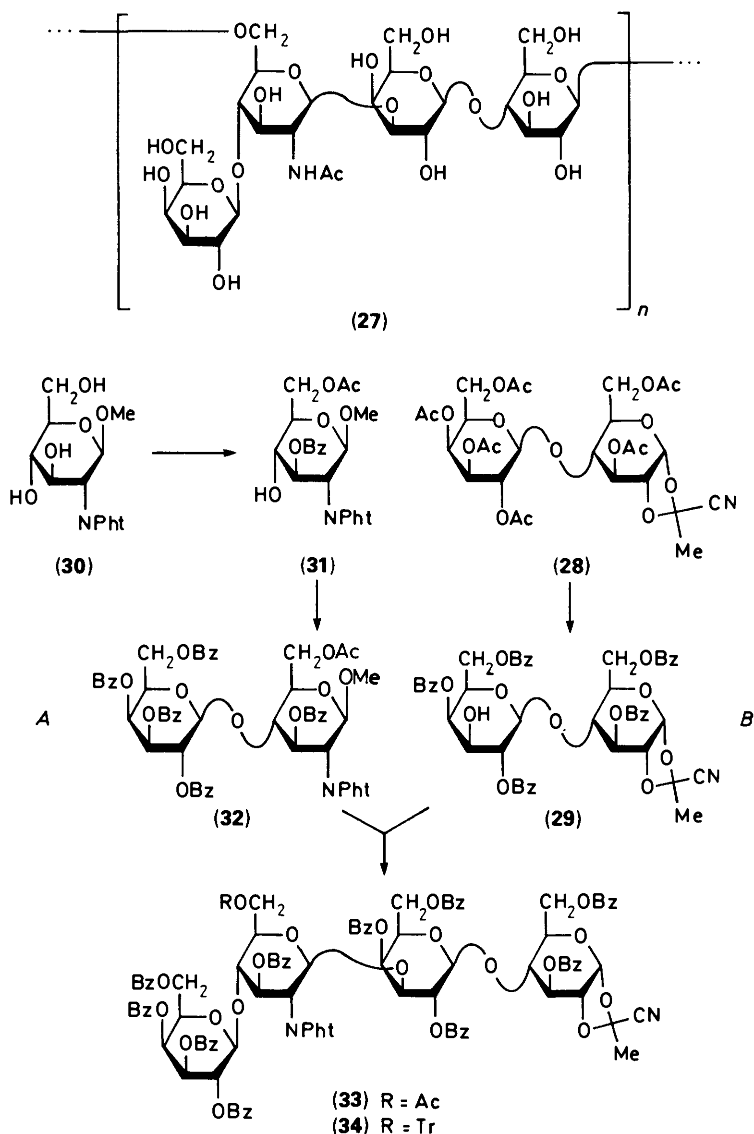
The synthesis of the repeating unit monomer was again the most laborious part. This was accomplished by a blockwise (2 + 2) scheme from the properly protected synthons as in the previous case[34] (Scheme 7). The tetrasaccharide was assembled from two disaccharide blocks, the derivatives of galactosyl-glucosamine (A) and galactosyl-glucose (B). Lactose, the precursor for the B-block, was successively converted into peracetate, to glycosyl bromide, and to the cyanoethylidene derivative (28). After conventional modification it afforded the synthon of the B-block, *viz.*, the disaccharide (29) with the cyanoethylidene group and the only free hydroxyl at C-3 of the galactose moiety. Synthesis of the second disaccharide synthon (block A) was begun from the methyl *N*-phthaloylglucosaminide (30) which, after usual transformations, gave the glucosamine derivative (31) with the only free hydroxyl group at C-4. Glycosylation of this

[35] V. I. Betaneli, M. V. Ovchinnikov, L. V. Backinowsky, and N. K. Kochetkov, *Carbohydr. Res.*, 1979, **76**, 252.

[36] Yu. E. Tsvetkov, N. E. Byramova, L. V. Backinowsky, N. K. Kochetkov, and N. F. Yankina, *Bioorg. Khim. (USSR)*, 1986, **12**, 1213.

[37] N. K. Kochetkov, N. E. Nifant'ev, and L. V. Backinowsky, *Tetrahedron*, 1987, **43**, 3109.

[38] B. Lindberg, J. Lonngren, and D. A. Powell, *Carbohydr. Res.*, 1977, **58**, 177.



Scheme 7

derivative with benzobromogalactose afforded the disaccharide (32) with one *O*-acetyl group at C-6 of the glucosamine and other hydroxyls protected as benzoates. Then the glycosyl bromide obtained from (32) was coupled to (29) to

give a tetrasaccharide cyanoethylidene derivative (33) with all but one hydroxyls protected as benzoates. The only *O*-acetyl group at O-6 of the glucosamine moiety was selectively removed and the hydroxyl group liberated was tritylated giving rise to the monomer (34). Polycondensation of (34) gave a polymer with a backbone composed of glucosamine, galactose, and glucose while the second galactose residue was attached as a side chain to every glucosamine residue at O-4. The structure of the synthetic polysaccharide obtained as (27) was proved by conventional methods including ^{13}C-NMR spectroscopy. Spectral and chromatographic data indicate the molecular mass of the polymer produced to be in the range of 6000, that is the degree of polymerization of the tetrasaccharide unit is about 10. This is of course lower than for the natural capsular polysaccharide which possesses an extremely high molecular mass. This difference does not seriously affect, however, the biological activity of the synthetic polysaccharide. It was half as active as the natural polysaccharide when assayed in the ELISA using the *Streptococcus pneumoniae* 14–anti-14 system.[37]

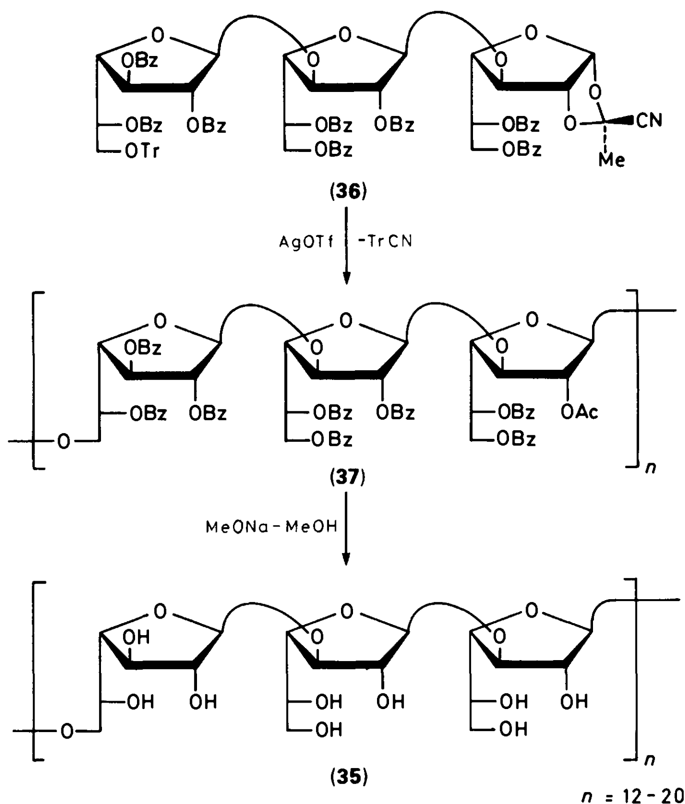
The trityl-cyanoethylidene polycondensation also allowed us to prepare polysaccharides containing the furanose rings. This is a more specific task because the possibilities for synthesis of glycofuranose structures are much more limited. As a successful example the synthesis of the so-called T1 antigen from mutants of some *Salmonella* species can be presented (Scheme 8). This polysaccharide (35) containing trisaccharide repeating units with (1→6)-linkages between them and 'inner' (1→3)-linkages[39] was synthesized[40] by polycondensation of the monomer (36) under usual conditions and after deacetylation of the product (37) afforded the polymer (35) consisting of furanoside repeating units and is identical, according to the literature data, to the natural biopolymer.

These four examples of the syntheses of microbial polysaccharides clearly show that the method is applicable to different objects which contain pyranosidic and furanosidic repeating units with 1,2-*trans*-glycosidic linkages between them and thus opens broad perspectives.

One more modification of this method which is promising a direct route to the synthetic polysaccharide antigens can be mentioned. It is well known that the preparation of artificial antigens is usually connected with the immobilization on a protein or synthetic carrier of a hapten repeating unit or a higher-molecular-weight fragment. The hapten to be immobilized contains a spacer arm at its reducing end with a functional group, necessary for attachment, at some distance from the sugar moiety. A version of the polysaccharide synthesis was elaborated which allowed us to get a polymer bearing at its reducing terminus a functionalized spacer arm. Thus, for example, polycondensation of a mannose bifunctional monomer (38) with both *O*-trityl and cyanoethylidene groups in the presence of a monofunctional compound (39) which contains only the *O*-trityl group and a glycosidically linked spacer arm, results along with the usual 'self-condensation' of the monomer, in a polycondensation whereon the polymeric

[39] M. Berst, O. Luderitz, and O. Westphal, *Eur. J. Biochem.*, 1971, **18**, 361

[40] S. A. Nepogod'ev, L. V. Backinowsky, and N. K. Kochetkov, *Bioorg. Khim.* (*USSR*), 1989, **15**, 1555.



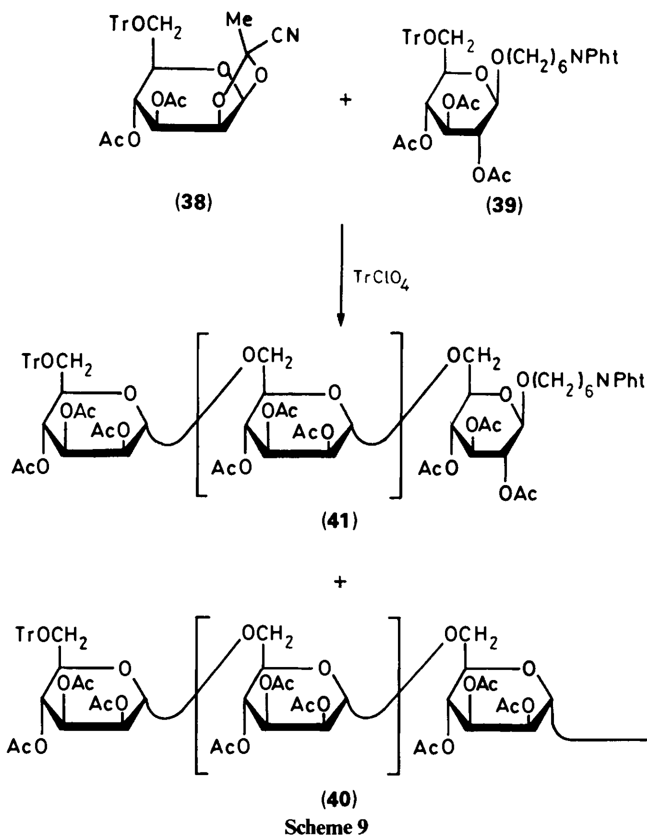
Scheme 8

chain is growing on the monofunctional terminator (Scheme 9).

As a result two polymers are produced, the α-1,6-mannan (40) due to the normal 'self-polycondensation', and the mannan with a 6-phthalimidohexyl glucoside at the reducing terminus (41). Hydrazinolysis of the mixture of (40) and (41) removes *O*-acyl protective groups and liberates free amino group which enables separation of the basic polysaccharide from the neutral one by ion-exchange chromatography.[41] The presence of a free amino group in the basic polysaccharide allows its subsequent immobilization on bovin serum albumin to give an artificial antigen.

This modification of the polysaccharide synthesis was used in preparation of an artificial antigen corresponding to the common polysaccharide antigen of *Pseudomonas aeruginosa*, which is a component of lipopolysaccharides in a majority of *Pseudomonas aeruginosa* species and in some other members of the

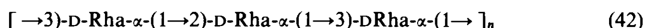
[41] Yu. E. Tsvetkov, A. V. Bukharov, L. V. Backınowsky, and N. K. Kochetkov, *Carbohydr. Res.*, 1988, **175**, C1.



Scheme 9

Pseudomonas family.[42] Synthesis of an artificial antigen can afford a valuable diagnosticum for a large series of diseases caused by *Pseudomonas aeruginosa.*

A common antigen is a linear rhamnan (42) built up of trisaccharide units with D-rhamnose, a sugar rarely occurring in nature.[43]

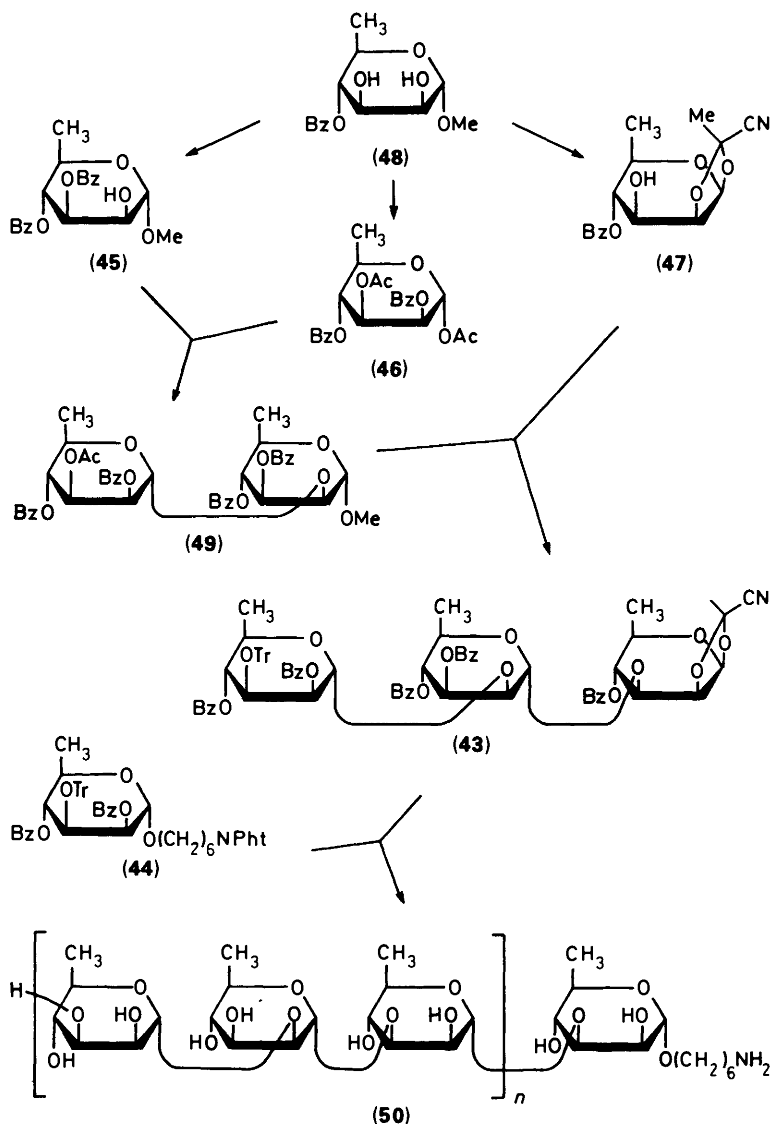


$$[\rightarrow 3)\text{-D-Rha-}\alpha\text{-}(1\rightarrow 2)\text{-D-Rha-}\alpha\text{-}(1\rightarrow 3)\text{-DRha-}\alpha\text{-}(1\rightarrow]_n \qquad (42)$$

The trisaccharide derivative (43) served as a monomer for the synthesis of the polysaccharide. It contained the cyanoethylidene group, and the trityl group at O-3 of the terminal non-reducing rhamnose moiety; as terminator 3-*O*-tritylated rhamnoside with a 6-phthalimidohexyl aglycone (44) was used which could serve later as a spacer arm in subsequent immobilization.

Synthesis of the monomer (43) followed the general strategy outlined earlier

[42] S. Yokota, S. Kaga, S. Sawada, Y. Ataki, and E. Ito, *Eur. J. Biochem.*, 1987, **167**, 203.
[43] N. A. Kocharova, Yu. A. Knirel, N. K. Kochetkov, and E. S. Stanislavsky, *Bioorg. Khim.* (*USSR*), 1988, **14**, 703.



Scheme 10

(Scheme 10).[44] The starting rhamnose synthons (45), (46), and (47) were prepared conventionally from methyl 4-*O*-benzoyl-D-rhamnopyranoside (48) and the trisaccharide assembly was carried out sequentially. First (46) was converted into the corresponding glycosyl bromide and coupled to (45) with the only hydroxyl

[44] Yu. E. Tsvetkov, L. V. Backinowsky, and N. K. Kochetkov, *Carbohydr. Res.*, 1989, **193**, 75.

group unprotected. The bioside (49) was again transformed into the glycosyl bromide and served as a glycosyl-donor in glycosylation of (47) which bears the cyanoethylidene group. The resulting trisaccharide contained the only *O*-acetyl protective group at O-3 of the terminal rhamnose. Selective removal of this acetyl group followed by tritylation gave the required monomer (43). Its polycondensation in the presence of the terminator (44) and triphenylmethylium perchlorate as a catalyst in a ratio of 10:1:1 was performed by slow addition of the monomer (43) solution to a mixture of (44) and the catalyst. From a mixture of two polysaccharides formed after deprotection by hydrazinolysis the polysaccharide with the 6-aminohexyl spacer arm (50) was isolated. Its ^{13}C-NMR spectrum and optical rotation coincided with those for the natural antigen isolated from *Pseudomonas aeruginosa* and *Pseudomonas cerasi.* Coupling of (50) to bovine serum albumin afforded a high molecular-weight glycoconjugate serving as an artificial antigen. This example demonstrates a rational approach for preparation of artificial polysaccharide antigens which can find application in immunochemical studies and for clinical purposes.

The solution of two important general problems of the chemistry of regular polysaccharides, the computer-assisted, time- and material-saving method of structural analysis, and the general approach to their synthesis, seems greatly to favour progress in this complicated field of chemistry of natural compounds.