Pure & Appl. Chem., Vol.53, pp.89-106.
Pergamon Press Ltd. 1981. Printed in Great Britain.

0033-4545/81/0101-0089$02.00/0

THE STRUCTURE OF CARBOHYDRATE CHAINS OF BLOOD-GROUP SUBSTANCES

Varvara A. DEREVITSKAYA

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

ABSTRACT

Treatment of blood group substance H obtained from pig stomach linings with alkaline borohydride results in splitting off ca. 80% of carbohydrate chains. The oligosaccharide fractions obtained after removal of polypeptide were subjected to gel-, paper and anion-exchange chromatography, & 29 individual reduced oligosaccharides, ranging from tri- to undecasaccharide were isolated and their structures were established.
Data obtained revealed gross heterogeneity both in molecular weight and branching degree of carbohydrate chains of blood group substances.
Despite the apparent variety of oligosaccharide structures they have much in common, more complex structures including simpler ones as structural elements.
The possible pathway of carbohydrate chain biosynthesis and general principle of construction of oligosaccharide chains in glycoproteins with O-glycosidic carbohydrate-peptide bonds will be discussed.

INTRODUCTION

Glycoproteins and glycolipids, now called glycoconjugates, are widely distributed in Nature. Practically all serum proteins, some hormones, enzymes, lectins and virus antigens are glycoproteins. The knowledge of the structure of these important biopolymers must be conducive to understanding the mechanism of the complicated life processes.
The role of carbohydrate in the biological function of glycoproteins is not always clear, but in many cases it has been shown that carbohydrate fragments of these biopolymers determine their biological specificity.
The striking example of such glycoproteins are blood group substances (BGS). Soluble BGS are the most complex glycoproteins with molecular weight between 1.10^5 and 1.10^6, which are present in a variety of exocrine gland secretions and serologically identical to erythrocyte antigens of ABO(H) and Lewis blood group systems. They probably play an important role in immunological specificity in human beings and higher animals.
A polypeptide backbone of BGS is linked with numerous carbohydrate chains (1). The proportion of amino acid and carbohydrates in BGS preparations is approximately 1:4. High content of serine, threonine and proline is characteristic of the peptide portion of these glycoproteins; their carbohydrate chains contain L-fucose, D-galactose, 2-acetamido-2-deoxy-D-glucose, 2-acetamido-2-deoxy-D-galactose and in some cases N-acetylneuraminic acid.

The main type of carbohydrate-peptide linkage in BGS is O-glycosidic between an N-acetylgalactosamine residue and serine and threonine residues of the peptide backbone.

POLYPEPTIDE CHAIN OF BGS

There is still only scanty information about the amino acid sequence of the polypeptide chain in BGS. The principal obstacle to the investigation of the structure of polypeptide of BGS is the lack of methods which allow effective removal of the carbohydrate chains without degradation of the polypeptide chain.
In the course of investigation of pig BGS we developed a combination of chemical and enzymatic methods of degradation, which enabled us to isolate twenty glycopeptides and establish their structure (2).
It seems not possible to derive a definite amino acid sequence surrounding the glycosylated hydroxy amino acid which might serve as a recognition signal for incorporation of the first N-acetylgalactosamine unit into a serine or threonine residue of the BGS peptide chain. Watkins isolated a number of peptides from human BGS and came to a similar conclusion (3).
Our experiments with pig BGS allow us to conclude that almost every serine and threonine residue of the polypeptide chain is glycosylated. Indeed, only a trace of serine and threonine were found in peptides formed by fragmentation of polypeptide backbone BGS through alkaline treatment leading to β-elimination of carbohydrate chains followed by bromination of resulting enaminic groups of the peptide chain and repeated alkaline treatment (Fig. 1).

HETEROGENEITY OF CARBOHYDRATE CHAINS IN BGS

It has been mentioned already that the biological specificity of BGS is coded in structure of its carbohydrate fragments, therefore the investigation of these structures is a most intriguing problem.
The investigation of BGS carbohydrate chains was started with the study of the structure of mono- and oligosaccharides determining their biological specificity. Some model compounds and small terminal oligosaccharide fragments isolated from the products of partial degradation of BGS were used for this purpose. These investigations were mostly performed in the laboratories of Watkins and Morgan in London, Kabat in New York and Yosizawa in Sendai. As a result of these very important studies the biological multiple specificity of BGS was established and the structure of determinant oligosaccharide fragments was elucidated (4) but the general structure, dimension and degree of heterogeneity of BGS carbohydrate chains remained unknown.
The solution of this problem necessitates isolation of intact oligosaccharide chains of BGS and determination of their primary structure, but that presents serious difficulties. As these glycoproteins are resistant to proteolytic enzymes, it is not possible to degrade a peptide skeleton and isolate fragments with intact carbohydrate chains.

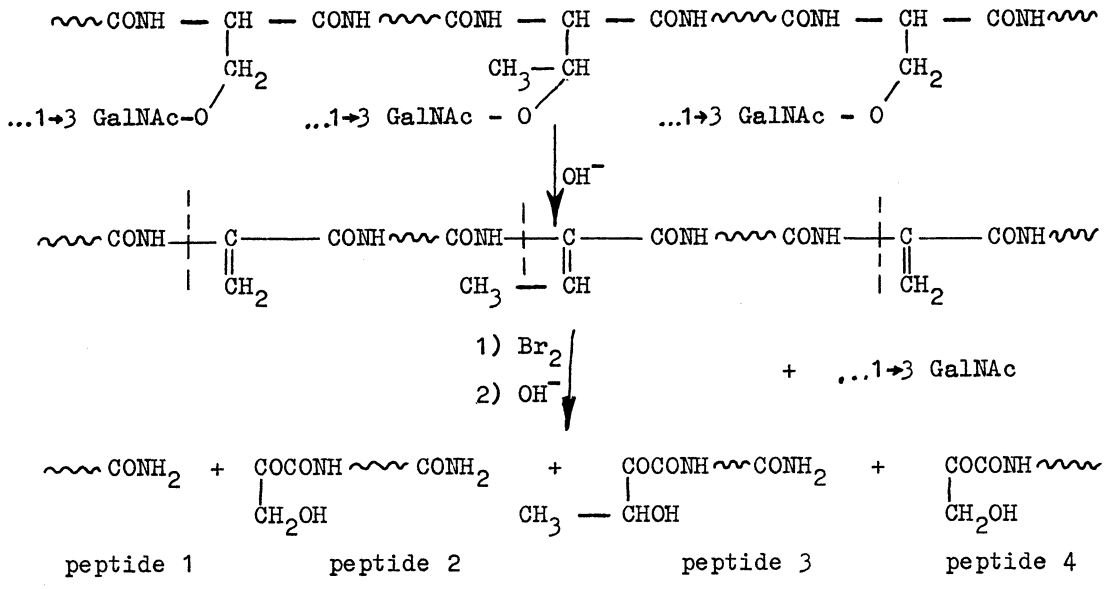


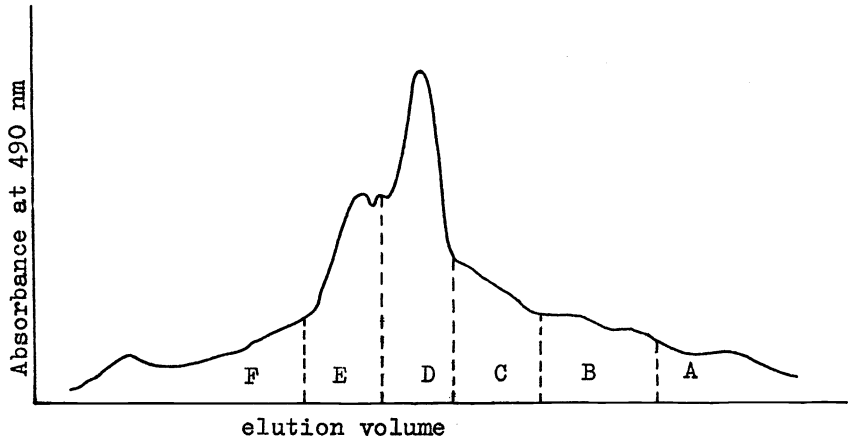
Fig. 1 The fragmentation scheme of BGS peptide chain

The treatment of BGS with bases or alkali results in the cleavage of carbohydrate - protein linkage due to β-elimination with the formation of the corresponding unsaturated amino acid. Resulting oligosaccharides are rapidly degraded from the reducing end because in this region of the carbohydrate chain monosaccharide residues are linked through 1→3 bonds. To prevent alkaline degradation of oligosaccharides they are reduced by sodium borohydride.

We have used this method for the investigation of the structure of carbohydrate chains of BGS H (5,6) from pig stomach linings isolated by the standard procedure modified in our laboratory (7). BGS H was treated under the conditions suggested by Carlson (8) - 0.05 M sodium hydroxide and 1M sodium borohydride at 50°C for 16 hours. These conditions made it possible to isolate reduced oligosaccharides in 85% yield.

The carbohydrate fraction was isolated from the degradation products by ion exchange chromatography. The separation of the carbohydrate fraction on Biogel P-6 (Fig. 2) showed a high degree of heterogeneity of reduced oligosaccharides with molecular weights of fractions ranging from 700 to more than 3000. In this experiment a preparation of BGS from homogenate of several pig stomachs with specificity H was used, but similar heterogeneity was observed when preparations from individual pig and human stomachs were fractionated (Fig. 3). Hence the observed heterogeneity of carbohydrate chains on BGS is not connected with the presence of material from different organisms but seems an inherent property of these glycoproteins.

This result was surprising as previous hypotheses on the structure of BGS were formulated in terms of identical carbohydrate chains in the molecule and suggested the presence of either a large number of short oligosaccharide chains (9) or a small number of an "average" carbohydrate chains containing about twenty monosaccharide units (10, 11). Now these concepts had to be abandoned as carbohydrate chains on BGS are certainly not identical.



Fraction	F	E	D	C	B	A
Yield %	15	20	40	10	12	3
Monosaccharide residues	10	7-10	5-6	4-5	2-3	1

Fig. 2 Chromatography of carbohydrate fraction of Pig BGS H on Biogel P-6, (column 200 x 3.4 cm, eluted with 0.1 acetic acid. Fractions were assayed by absorbance at 490 nm after reactions with phenol and sulfuric acid) and the characterisation of carbohydrate fraction.

ISOLATION AND STRUCTURE ELUCIDATION OF INDIVIDUAL OLIGOSACCHARIDES

All fractions from Biogel P-6 chromatography (Fig. 2) were subjected to additional fractionation except the fraction with a molecular weight of more than 3000, which was obtained in small yield (∿10%) and was not examined further. It is known that the separation of complex mixturesof oligosaccharides of similar composition and without acidic components is a difficult problem. The methods for the separation of the higher hetero-oligosaccharides have not been developed.

Preparative paper chromatography of each fraction from Biogel P-6 leads to a great number of discrete spots but only two of them comprised individual trisaccharides. The others presented complicated mixtures of oligosaccharides.

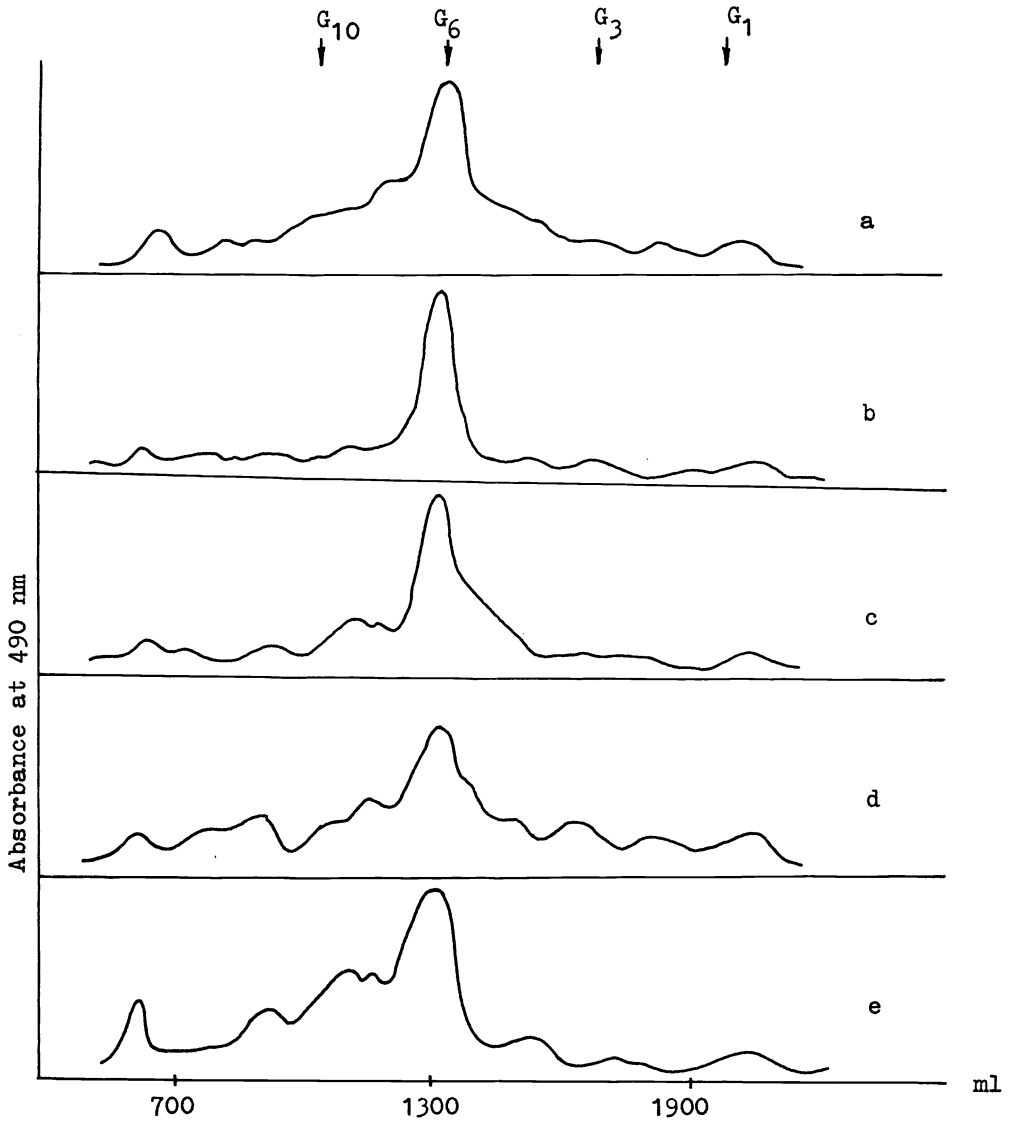


Fig. 3 Gel chromatography of oligosaccharides from the pooled pig stomach linings (a), individual pig stomach linings (b, c, d) and individual human stomach (e). Column -Biogel P-6 (200 x 3.4 cm). Elution with 0.1 M acetic acid. Fractions were assayed after reaction with phenol and sulfuric acid. G_1, G_3, G_6 and G_{10} designate the elution volumes of mono-, di-, hexa- and decasaccharides.

To separate these subfractions we have developed a method of preparative anion-exchange chromatography in borate buffer (Fig. 4).
The method is very effective and convenient for different types of complicated oligosaccharide separations and is widely used in our laboratory.
Thus twenty nine reduced oligosaccharides ranging from trisaccharide to undecasaccharide were isolated and characterised (Table 1).

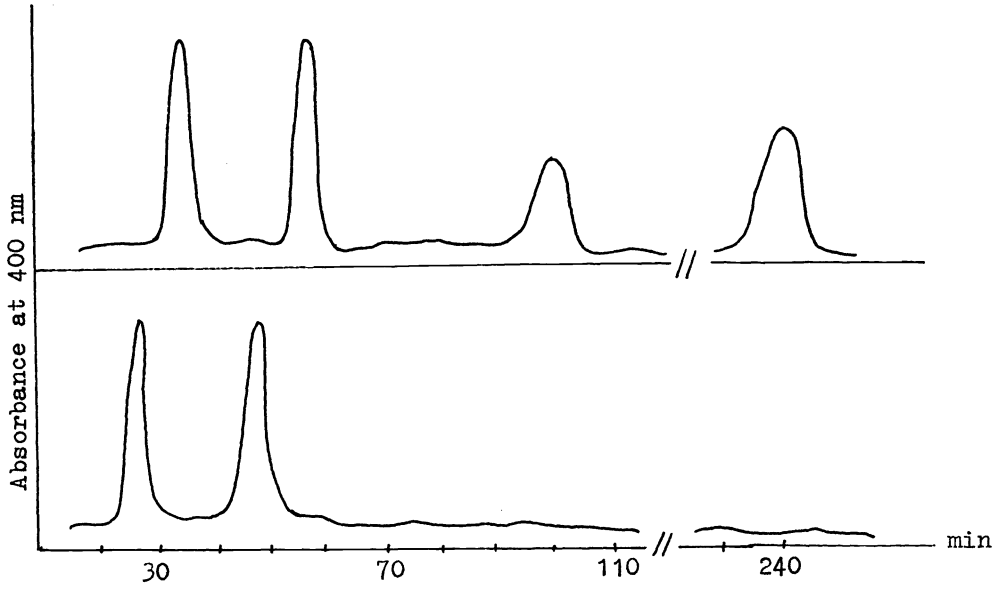


Fig. 4 Chromatography of oligosaccharide fractions (after paper chromatography) on Anionite DAx4, column 60 x 0.8 cm, elution with 0.5 M borate buffer pH 8.5, flow rate 90 ml/h. Fractions were assayed by absorbance at 400 nm after reaction with orcinol and sulfuric acid.

To determine oligosaccharide structure the following methods were used. 1. Total hydrolysis (4 M CF_3COOH, 100°C, 16 h) followed by quantitative ion-exchange chromatography as used for determination of monosaccharide composition. 2. To establish linkage positions in oligosaccharides methylation analysis (methylation by Hakomori method) followed by analysis of methylated alditol acetates with GLC and GLC/mass spectrometry, and periodate oxidation followed by Smith degradation were used. 3. The oligosaccharides were sequenced by partial acidic (0.05 M CF_3COOH, 100°C, 3h) and enzymatic hydrolysis, followed by isolation, anion chromatography, and structural analysis of the resulting fragments. Mass spectrometry of the permethylated derivatives was used for sequencing of oligosaccharides with 3 to 6 monosaccharide units. 4. Configuration of glycosidic linkages was established by the method of Angyal viz., chromium trioxide oxidation of oligosaccharide acetates followed by total hydrolysis and estimation of monosaccharides survived. Hydrolysis of oligosaccharides with specific glycosidases as well as ^{13}C n.m.r. spectrometry were used also for this purpose.

It is well known that accurate estimation of the composition of complex oligosaccharides with glycosidic linkages of different stability to hydrolysis is a difficult task. In particular, this is the case with the oligosaccharides containing a number of hexosamine residues. Therefore the monosaccharide composition given in the Table 1 for some oligosaccharides should be considered as a preliminary one to be checked in the course of structural studies.

The structural elucidation of oligosaccharides is illustrated by two examples - the hexasaccharide 9 and the undecasaccharide 28.

TABLE 1. The elution time (on Anionite DA x 4, column 60 x 0.8 cm) and monosaccharide composition of oligosaccharides.

Number of oligosaccharide	elution time min.	Monosaccharide composition (mole/mole GalolNAc)		
		GlcNAc	Gal	Fuc
1	87	0.0	1.2	0.0
2	158	0.0	1.1	1.0
3	43	0.8	1.1	0.0
4	46	1.7	2.0	0.0
5	86	1.9	1.8	0.8
6	137	1.0	2.3	1.1
7	270	1.1	2.2	2.1
8	33	2.8	1.8	0.0
9	74	2.1	2.2	1.2
10	36	(3.0)	(3.0)	(0.0)
11	92	(2.0)	(3.0)	(1.0)
12	31	(4.0)	(3.0)	(0.0)
13	51	3.3	3.5	1.1
14	150	2.2	2.6	1.8
15	32	(3.0)	(3.0)	(0.0)
16	67	2.2	2.7	0.7
17	28	(4.0)	(3.0)	(0.0)
18	43	2.7	2.5	0.8
19	111	2.9	2.8	2.1
20	34	(4.0)	(3.0)	(0.0)
21	74	2.6	3.2	0.9
22	29	(5.0)	(3.0)	(0.0)
23	45	3.6	2.4	0.7
24	123	3.2	2.8	2.2
25	80	(3.0)	(4.0)	(1.0)
26	25	(6.0)	(4.0)	(0.0)
27	48	3.7	3.9	1.1
28	72	3.5	3.4	1.8
29	135	2.9	3.3	1.9

The scheme of elucidation of hexasaccharide structure is represented in Fig.5. The estimation of monosaccharide composition showed that hexasaccharide consisted of fucose, galactose, N-acetylglucosamine and N-acetylgalactosaminitol residues in the ratio of 1:2:2:1. Periodate oxidation of the oligosaccharide removed essentially all the monosaccharide units except one N-acetylglucosamine residue. Herewith formaldehyde was not formed. 2-Amino-2-deoxythreitol was present among Smith degradation products. This fact is direct evidence of substitution of the N-acetylgalactosaminitol residue at C-3 and C-6.

The attachment of a galactose residue at C-3 of the N-acetylgalactosaminitol residue is shown by the formation in high yield of the reduced disaccharide Gal 1 → 3 GalolN among the products of partial, acidic hydrolysis (1M CF_3COOH, 100°C, 4 h) of the hexasaccharide.

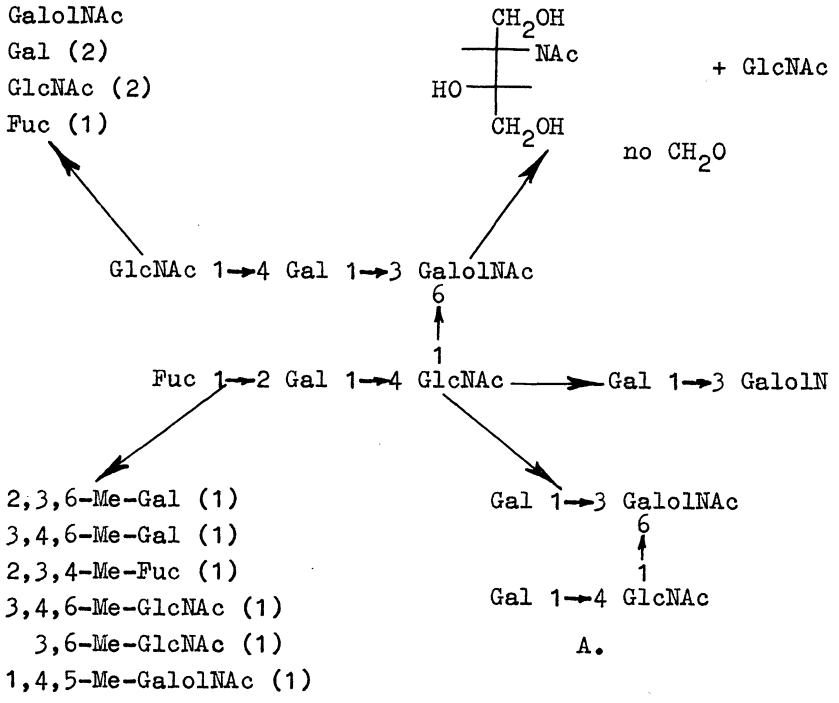


Fig. 5 The scheme for elucidation of hexasaccharide 9 structure.

The methylation analysis showed that one of the galactose residues is linked at C-4, the other at C-2; terminal N-acetylglucosamine and fucose residues are present; the other N-acetylglucosamine residue is substituted at C-4 and the N-acetylgalactosaminitol residue is linked at C-3 and C-6. Tetrasaccharide A was obtained from the hexasaccharide after partial, acidic hydrolysis followed by incubation with α-N-acetylglucosaminidase. It was isolated preparatively and its structure was determined.

The hexasaccharide was sequenced by mass-spectrometry (Fig. 6).

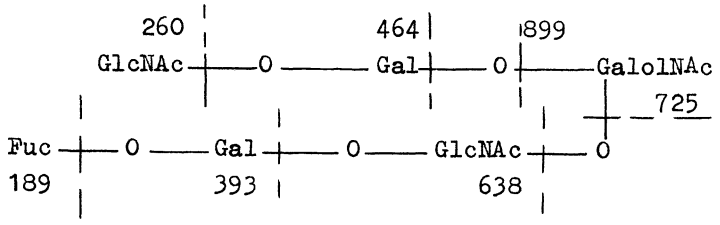


Fig. 6 Mass-spectrometric fragmentation of permethylated oligosaccharide 9.

The configuration of glycosidic linkages followed from results of chromium trioxide oxidation: fucose and one of the N-acetylglucosamine residues survived the oxidation and hence they should be bound by α-linkages. This conclusion was finally supported by ^{13}C n.m.r. (12).

The second example of the structure elucidation is the most complicated oligosaccharide 28, with eleven monosaccharide units. The scheme for the elucidation of this undecasaccharide structure is shown in Fig. 7. The method of mass spectrometry is less effective for sequencing of such oligosaccharides. Consequently the main method of sequencing in this case was partial degradation. Effective separation of the resulting fragments was essential to achieve unequivocal conclusions and the significance of preparative anion exchange chromatography in borate buffer in this respect should be stressed once more.

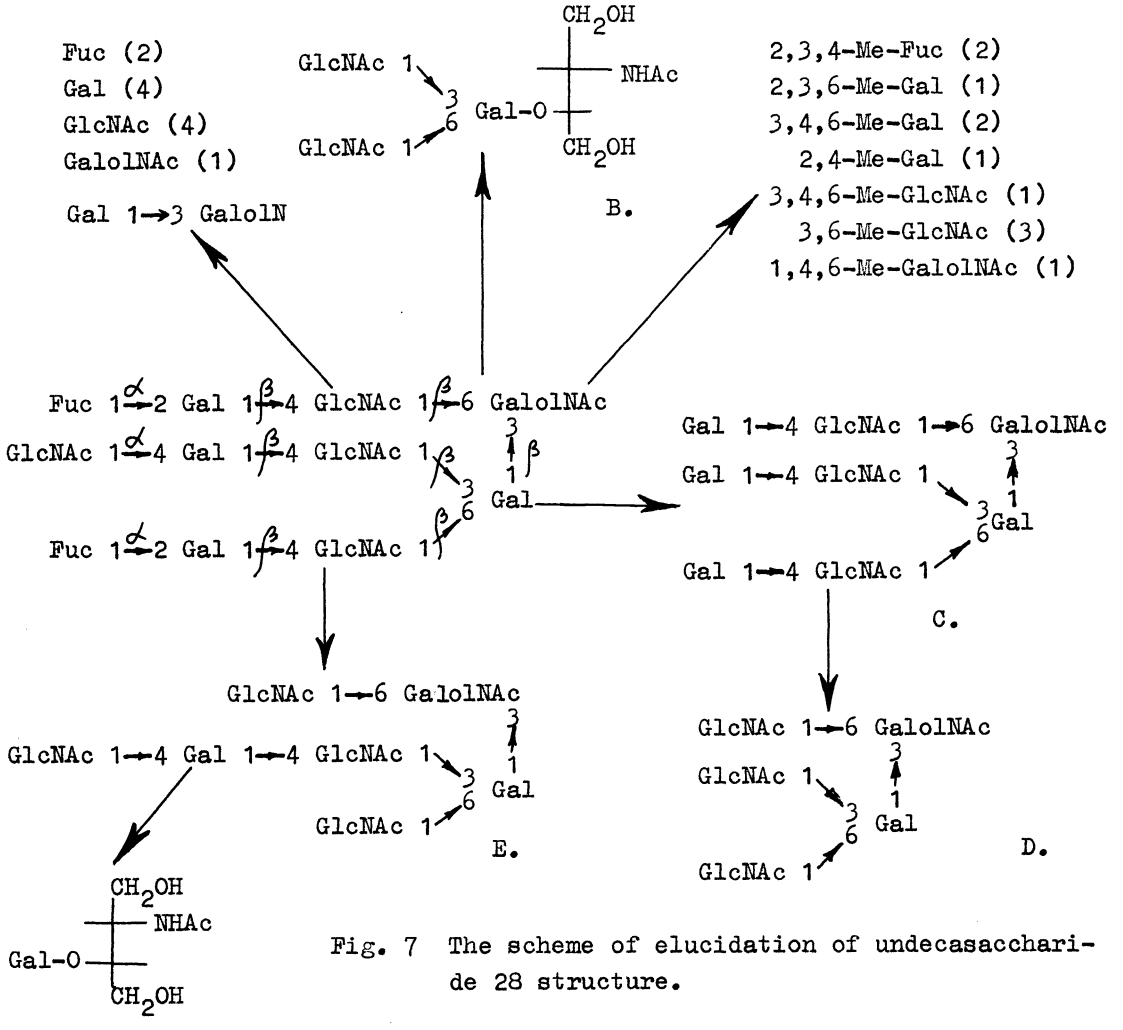


Fig. 7 The scheme of elucidation of undecasaccharide 28 structure.

The undecasaccharide was composed of fucose, galactose, N-acetylglucosamine and N-acetylgalactosaminitol in the ratio of 2:4:4:1. Methylation analysis showed that there are two terminal fucose; one terminal N-acetylglucosamine residue; two galactose units substituted at C-2 and two others, one substituted at C-4, the other at C-3 and C-6; three N-acetylglucosamine residues substituted at C-4, and N-acetylgalactosaminitol linked at C-3 and C-6.

Oligosaccharide B was preparatively isolated after Smith degradation of the undecasaccharide and its structure was confirmed by methylation analysis. Successive mild acidic hydrolysis and incubation with α-N-acetylglucosaminidase (see above) of the undecasaccharide led to oligosaccharide C which is the common fragment of all oligosaccharides of this series. Removal of three galactose residues with β-galactosidase converted oligosaccharide C to the fragment D. The structure of the oligosaccharide D was independently established by periodate oxidation, Smith degradation and methylation analysis. After these procedures only the position of terminal fucose and N-acetylglucosamine units remained obscure. To answer this question the oligosaccharide E was obtained after successive mild acidic hydrolysis and incubation with β-galactosidase of the starting oligosaccharide.

The structure of oligosaccharide E was established by methylation analysis, which revealed the presence of N-acetylgalactosaminitol and galactose residues substituted at C-3 and C-6, three terminal N-acetylglucosamine residues and galactose and N-acetylglucosamine units substituted at C-4. Only galactose was preserved after two successive Smith degradations of the oligosaccharide E. This is direct evidence of the terminal position of N-acetylglucosamine on the chain attached to N-acetylgalactosaminitol at C-3.

The configuration of glycosidic linkages of the studied undecasaccharide followed from the action of specific glycosidases: α-L-fucosidase, β- and α-N-acetylglucosaminidases and β-galactosidase, and the data of high resolution ^{1}H n.m.r. spectrometry (unpublished data of H. van Halbeek).

The structures of all 29 oligosaccharides obtained were elucidated by analogous approaches.

COMMON STRUCTURAL FEATURES OF O-LINKED CARBOHYDRATE CHAINS IN BGS

All oligosaccharides studied can be divided into six series according to their type of structure. The oligosaccharides belong to each series have the common fragment (framed in the following formulas) and differ only in number and position of determinant monosaccharide units of α-fucose and α-N-acetylglucosamine.

Oligosaccharidesof series I

$$\boxed{\text{Gal } 1\xrightarrow{\beta}3 \text{ GalolNAc}}$$

(1) $\quad$ Gal $1\xrightarrow{\beta}3$ GalolNAc

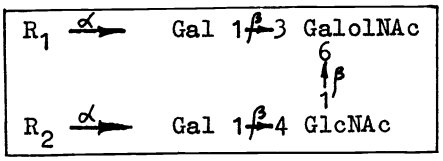
(2) $\quad$ Fuc $1\xrightarrow{\alpha}2$ Gal $1\xrightarrow{\beta}3$ GalolNAc

(3) $\quad$ GlcNAc $1\xrightarrow{\alpha}4$ Gal $1\xrightarrow{\beta}3$ GalolNAc

A disaccharide and two trisaccharides are included in this series. The trisaccharides contain N-acetylgalactosaminitol residues at the reduced end substituted by galactose at C-3 and there are α-fucose and α-N-acetylglucosamine units on the nonreduced end of the chain.

Oligosaccharides of series II

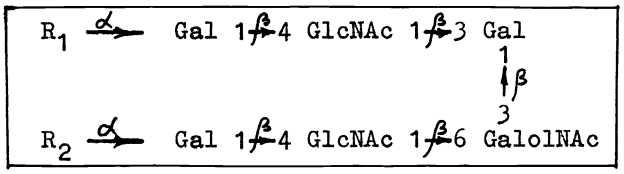
In these oligosaccharides the N-acetylgalactosaminitol residue is the only point of chain branching.



	R_1	R_2
(4)	GlcNAc 1→4	-
(5)	Fuc 1→2	GlcNAc 1→4
(6)	-	Fuc 1→2
(7)	Fuc 1→2	Fuc 1→2
(8)	GlcNAc 1→4	GlcNAc 1→4
(9)	GlcNAc 1→4	Fuc 1→2

These oligosaccharides constitute a major part (∿40%) of the total carbohydrate fraction of BGS.

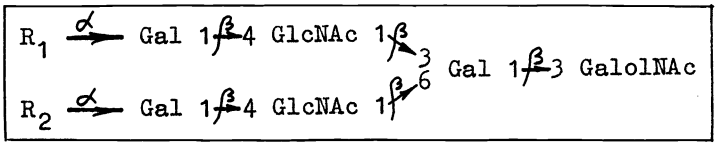
Oligosaccharides of series III



	R_1	R_2
(10)	- - - GlcNAc 1→4	- - -
(11)	- - - Fuc 1→2	- - -
(12)	GlcNAc 1→4	GlcNAc 1→4
(13)	GlcNAc 1→4	Fuc 1→2
(14)	Fuc 1→2	Fuc 1→2

The structure of oligosaccharides of this series is analogous to those of the previous series, but the chain attached to N-acetylgalactosaminitol at C-3 has in addition two monosaccharides.

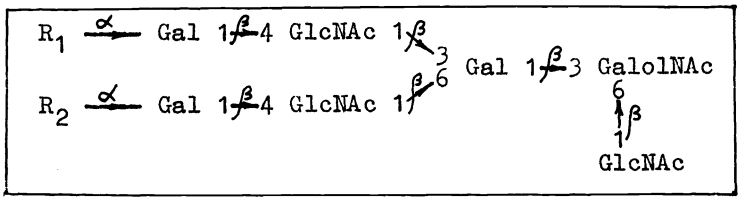
Oligosaccharides of series IV



	R_1	R_2
(15)	- - - GlcNAc 1→4	- - -
(16)	- - - Fuc 1→2	- - -
(17)	GlcNAc 1→4	GlcNAc 1→4
(18)	GlcNAc 1→4	Fuc 1→2
(19)	Fuc 1→2	Fuc 1→2

This series also includes hepta- and octasaccharides, but here galactose is the sole point of branching of the oligosaccharide chain.

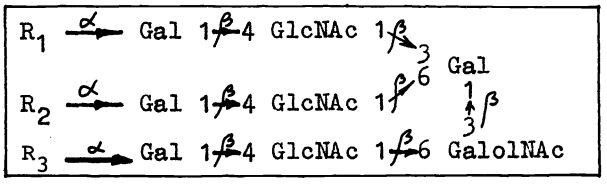
Oligosaccharides of series V



	R_1	R_2
(20)	- - - GlcNAc 1→4	- - -
(21)	- - - Fuc 1→2	- - -
(22)	GlcNAc 1→4	GlcNAc 1→4
(23)	GlcNAc 1→4	Fuc 1→2
(24)	Fuc 1→2	Fuc 1→2

In these oligosaccharides there are two branching points of the chain - N-acetylgalactosaminitol and galactose residues. Only N-acetylglucosamine is attached to the N-acetylgalactosaminitol residue at C-6. We think that the oligosaccharides of this series may be precursors for the following series.

Oligosaccharides of series VI



	R_1	R_2	R_3
(25)	- - - Fuc 1→2	- - -	- - -
(26)	GlcNAc 1→4	GlcNAc 1→4	GlcNAc 1→4
(27)	GlcNAc 1→4	- - - - - Fuc 1→2	- - -
(28)	GlcNAc 1→4	Fuc 1→2	Fuc 1→2
(29)	- - - Fuc 1→2	- - - - - Fuc 1→2	- - -

This series includes the most complicated oligosaccharides, where N-acetylgalactosaminitol residue is substituted with two chains - a linear one at C-6 and a branched one at C-3.

Isolation of a number of reduced oligosaccharides from human BGS was reported by Kabat and co-workers (13, 14). Although those authors failed to obtain individual complex oligosaccharides, some structures reported by them are analogous to those of the oligosaccharides considered here.

Some peculiarities of the oligosaccharides obtained should be noted. On one of the ends a reduced N-acetylgalactosamine residue is always present, which is attached to serine and threonine residues of the peptide chain of BGS. On the other nonreduced end of the oligosaccharides there are α-L-fucose or α-N-acetylglucosamine residues. Presence of characteristic monosaccharide residues on both chain ends of the isolated oligosaccharides is a strong argument for the proposal that these oligosaccharide correspond in structure to native carbohydrate chains of BGS.

Terminal disaccharide fragment Fuc 1→2 Gal determines H-specificity of BGS. In respect of disaccharide α-GlcNAc 1→4 Gal it is known that antibodies

specific to this determinant can be found in human beings and animals after immunization with pig BGS A or H (10).

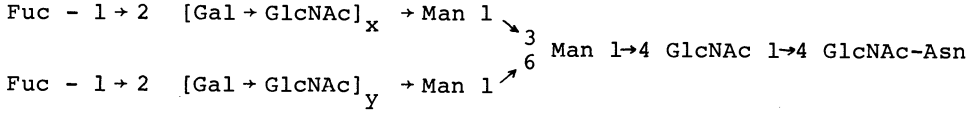
N-LINKED CARBOHYDRATE CHAINS IN BGS

To complete the characterization of carbohydrate chains of BGS it should be remarked that another minor type of carbohydrate-peptide linkage - N-glycosidic, also has been found in these glycoproteins. Glycopeptides with N-glycosidic bonds were directly isolated by us (15) from products of chemical and enzymatic degradation of pig BGS and their structures were established. One of the glycopeptides consisted of N-acetylglucosamine and asparagine residues only; the other contained a sequence of monosaccharide residues typical of the core of glycoproteins with N-glycosidic carbohydrate-peptide linkages:

Man 1 → 3 Man 1 → 4 GlcNAc 1 → 4 GlcNAc → Asn

In order to get these glycopeptides with intact oligosaccharide chains, carbohydrate chains linked by O-glycosidic bonds with serine and threonine residues were split from BGS by alkaline β-elimination, and enaminic groups formed in the peptide chain were destroyed by additional alkali treatment after bromination (Fig. 1). The glycopeptide fraction isolated from degradation products had serological activity of BGS H and consisted of asparagine, mannose, galactose, N-acetylglucosamine and fucose in the ratio of 1:3:25:23:3 (16).

On the basis of the structure of the core and the composition of carbohydrate chains of these glycopeptides, the following hypothetical structure may be suggested:



BIOSYNTHETIC CONSIDERATIONS

It follows from the comparison of the structure of all oligosaccharides discussed that they and consequently the carbohydrate chains of BGS differ from each other in molecular weight, structure and especially in degree of branching.

Considering the structures of oligosaccharide with O-glycosidic linkages we can observe that despite the apparent variety of the structure of these oligosaccharides they have much in common and more complex structures include simpler ones as structural elements (Fig. 8). Disaccharide Gal 1 → 3 GalolNAc is the core of all oligosaccharides. The disaccharide unit of N-acetyllactosamine (Gal 1 → 4 GlcNAc) can then be attached either to an N-acetylgalactosaminitol residue at C-6 or to a galactose residue at C-3 and/or C-6. This process should be followed by transfer of terminal α-fucose or α-N-acetylglucosamine units to galactose residues. The terminal α-fucose unit can be attached only at C-2 whereas α-N-acetylglucosamine only at C-4, after which the chain elongation stops.

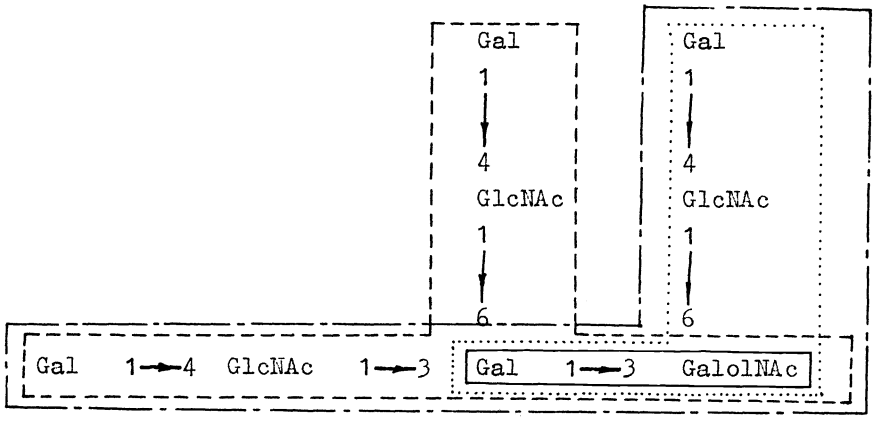


Fig. 8 The structure of common parts of trisaccharides (———), penta- and hexasaccharides (·····), series IV oligosaccharides (-------), series III oligosaccharides (—·——) as elements of complex structure of series VI oligosaccharides.

Thus joining of a fucose residue or α-N-acetylglucosamine residue to the core leads to trisaccharides of series I. The substitution of N-acetylgalactosaminitol at C-6 by an N-acetyllactosamine unit and the subsequent transfer of two terminal monosaccharide units lead to hexasaccharides of series II. The further complication of the structure by substituting galactose units of the core at C-3 and/or C-6 with the N-acetyllactosamine unit, and the subsequent addition of determinant monosaccharide residues lead to more complex oligosaccharides of series III-VI. Thus despite the variety of carbohydrate chains of BGS they are all assembled according to a single biosynthetic programme.

It is known that carbohydrate chains of glycoproteins with O-glycosidic carbohydrate-peptide bonds are synthesized by a "step to step" mechanism which consists of the transfer of sugar residues from sugar nucleotides by glycosyltransferses.

A spectrum of glycosyltransferases is required for the biosynthesis of the complex carbohydrate chains of BGS which are constituted of four different sugars joined in a variety of positional and anomeric linkages.

The specificity of these glycosyltransferases for the biosynthesis of carbohydrate chains of BGS ABO (H) may be deduced in considerable detail from the structures of the considered oligosaccharides.

At least two β-galactosyltransferases and three β-N-acetylglucosaminyltransferases are required for the biosynthesis of these oligosaccharides. One of the galactosyltransferases should be specific to α-N-acetylgalactosaminyl-serine (threonine) residue as an acceptor; the other may use as an acceptor

oligosaccharides with terminal N-acetylglucosamine residues.

TABLE 2. The characterization of glycosyltransferase specificity in biosynthesis of carbohydrate chains of BGS ABO (H)

transferase	acceptor fragment	Position of glycosylation
β-galactosyltransferase	α-N-acetylgalactosamine unit linked to serine	3-OH
β-galactosyltransferase	β-N-acetylglucosamine unit - terminal position	4-OH
β-N-acetylglucosaminyl transferase	α-N-acetylgalactosamine unit substituted at C-3	6-OH
β-N-acetylglucosaminyl-transferase	β-galactose unit substituted at C-3	6-OH
β-N-acetylglucosaminyl-transferase	β-galactose unit - terminal position	3-OH
α-fucosyltransferase	β-galactose unit - terminal position	2-OH
α-galactosyltransferase	β-galactose unit substituted at C-3	3-OH
α-N-acetylgalactosaminyl-transferase	β-galactose unit substituted at C-2	3-OH

BGS carbohydrate chains contained galactose residues substituted at C-3, or C-3 and C-6, but not at C-6. Hence one β-N-acetylglucosaminyltransferase (3-transferase) should be specific to an acceptor with terminal galactose residue and the other (6-transferase) to an acceptor containing a galactose residue substituted at C-3.

The same situation is characteristic of N-acetylglucosaminyltransferase which transfers N-acetylglucosamine to C-6 of N-acetylgalactosamine residues substituted at C-3: there are N-acetylgalactosamine residues substituted at C-3, or at C-3 and C-6 in the oligosaccharides, but one substituted at C-6 is absent.

α-Fucosyltransferase should be specific to terminal galactose residues as well as β-N-acetylglucosaminyltransferase and consequently these two enzymes are competitors in the substitution of terminal galactose residues during biosynthesis.

The attachment of β-N-acetylglucosamine to a terminal galactose residue leads to oligosaccharide chain elongation while the attachment of α-fucose to the same galactose residue leads to the termination of the oligosaccharide chain. The formation of the points of branching at galactose and at N-acetylgalactosamine residues are independent processes but elongation of these branches is determined by the competition of the same glycosyltransferases - β-N-acetylglucosaminyltransferase and α-fucosyltransferase specific to acceptors with a terminal galactose residue. There may be a regulation mechanism which directs the complex processes of BGS carbohydrate chain biosynthesis in one of these directions.

TABLE 3. The hexasaccharides of different BGS isolated from stomach lining

Source and specificity of BGS	Hexasaccharide structure
human BGS H	Fuc 1→2 Gal 1→3 GalolNAc 6 ↑β 1 Fuc 1→2 Gal 1→4 GlcNAc
pig BGS H	Fuc 1→2 Gal 1→3 GalolNAc 6 ↑β 1 GlcNAc 1→4 Gal 1→4 GlcNAc
horse BGS B (17)	Fuc 1→2 Gal 1→3 GalNAc 6 ↑β 1 Fuc 1→2 Gal 1→4 GlcNAc 3 ↑α 1 Gal

The BGS A and B may be considered as derivatives of BGS H. α-Galactosyltransferases are specific to acceptors with β-galactose residues substituted at C-2 by fucose. Glycosyltransferase cannot transfer α-galactose or α-N-acetylgalactosamine residues to galactose before substitution of the galactose residue at C-2 by fucose.

The suggested pathway of BGS carbohydrate chain biosynthesis seems to be common for different organisms. Indeed identical or closely related fragments were isolated from human, pig and horse blood group substances. These include the trisaccharide Fuc 1 $\overset{\alpha}{\downarrow}$ 2 Gal 1 $\overset{\beta}{\downarrow}$ 3 GalolNAc and the hexasaccharide (Table 3) which are the main part of carbohydrate fractions and differ only in terminal determinant monosaccharide units. In pig BGS H an α-N-acetylglucosamine residue substitutes for one of two α-fucose residues characteristic for human BGS H, while in the oligosaccharide from horse BGS B an additional α-galactose residue is present.

CONCLUSION

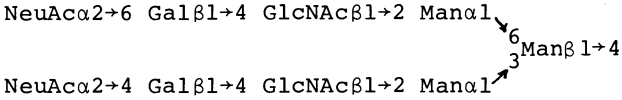
Carbohydrate chains of BGS seem to be among the most complicated structures found in "mucin-type" glycoproteins which contain O-linked carbohydrate units. These units may be represented by the following general formula with the same disaccharide core as in BGS.

$$\begin{array}{ccccc} R_1 & \rightarrow & \text{Gal } 1 & \rightarrow & 3\ \text{GalNAc} \\ & & \downarrow & & \downarrow \\ & & R_2 & & R_3 \end{array}$$

where R_1 and R_2 stand for H, fucose or sialic acid residues and R_3 for H or sialic acid residues (the only exception is ovine submaxillary gland mucin where the disaccharide unit SA 2→6 GalNAc is present).

In carbohydrate chains of BGS the core sequence is extended by addition of

N-acetyllactosamine residues followed by addition of terminal determinant residues, whereas in simple mucin-type glycoproteins the terminal fucose or sialic acid residues are linked directly to the core disaccharide. In other words in BGS carbohydrate chains the N-acetyllactosamine residue serves as a spacer between core and terminal monosaccharide residues.
Similar general structures are characteristic of complex N-linked carbohydrate chains of glycoproteins (18) where N-acetyllactosamine spacers are inserted between the pentasaccharide $Man_3GlcNAc_2$ core sequence and terminal sialic acid or fucose residues as shown for example in the following structure (18)



This similarity of general structure of carbohydrate chains may reflect some similarity in regulation of the biosynthesis and/or biological functions of the different classes of glycoproteins.

Acknowledgements - I am very grateful to all members of our research group whose names are listed in the bibliography, for their contribution to this work and especially to Professor N.K. Kochetkov for support and encouragement and Dr. V.N. Shibaev for helpful discussion.
I am indebted to Dr. H. Halbbek for access to unpublished results on ^1H n.m.r. spectrometry.

REFERENCES

1. N.K.Kochetkov, V.A.Derevitskaya, S.G.Kara-Murza, Carbohyd.Res., 3, 403-415 (1969).
2. N.K. Kochetkov, V.A. Derevitskaya, L.M. Likhosherstov, S.A. Medvedev, Biochem. Biophys. Res. Commun., 56, 311-316 (1974).
3. Shirley D. Goodwin, Winifred M. Watkins, Eur. J. Biochem., 47, 371-382, (1974).
4. W.M. Watkins in "Glycoproteins" ed. A. Gottschalk, Elsevier Amsterdam, London, 1966, p.430-535.
5. Nicolai K. Kochetkov, Varvara A. Derevitskaya, Nicolai P. Arbatsky, Eur. J. Biochem., 67, 129-136 (1976).
6. Varvara A. Derevitskaya, Nicolai P. Arbatsky, Nicolai K. Kochetkov, Eur. J. Biochem., 86, 423-437 (1978).
7. L.M. Likhosherstov, V.A. Derevitskaya, V.I. Fedorova, Biokhimia, 34, 45-50, (1970).
8. I.W. Maye, D.M. Carlson, Carbohyd. Res., 15, 300-303 (1970).
9. W.T.J. Morgan, Bull. Soc. Chem. Biol., 46, 1627 (1964).
10. K.O. Lloyd, E.A. Kabat, E. Licerio, Biochemistry, 7, 2976-2990 (1968).
11. N.K. Kochetkov, V.A. Derevitskaya, L.M. Likhosherstov, M.D. Martynova, S.N. Senchenkova, G.S. Kikot and L.S. Bogdashova, Biochem. Biophys. Res. Commun, 39, 583-588 (1970).
12. Alexander S. Shashkov, Nicolai P. Arbatsky, Varvara A. Derevitskaya and Nicolai K. Kochetkov, Carbohyd. Res., 72, 218-221 (1979).

13. C. Maisonruge-McAuliffe, Elvin A. Kabat, Arch. Biochem. Biophys., 175, 90-113, (1976).
14. L. Rovis, B. Anderson, E.A. Kabat, F. Cruezo, J. Liao, Biochemistry, 12, 5340-5354 (1975).
15. N.K. Kochetkov, V.A. Derevitskaya, L.M. Likhosherstov, S.A. Medveded, Izv. Akad. Nauk SSSR. Ser. Khim. 461-463 (1974).
16. L.M. Likhosherstov, M.D. Martynova, E.A. Kosheleva, V.A. Derevitskaya, N.K. Kochetkov, Proc. Akad. Nauk SSSR, 223, 709-710 (1976).
17. W. Newman, E.A. Kabat, Arch. Biochem. Biophys., 172, 535-550 (1976).
18. A. Kabata, K. Yamashita, S. Takasaki, T. Mizuochi, Y. Tachibana, Glycoconjugates. Proceedings of the fifth international symposium, Kiel, p.6-9, (1979).