HAWORTH MEMORIAL LECTURE*

Synthesis of Complex Oligosaccharide Chains of Glycoproteins

By H. Paulsen
INSTITUT FÜR ORGANISCHE CHEMIE UND BIOCHEMIE,
UNIVERSITÄT HAMBURG, D–2000 HAMBURG 13, WEST GERMANY

1 Introduction

The selective chemical synthesis of complex oligosaccharides has made remarkable progress over the past five years.[1] Biologically interesting oligosaccharides with different building blocks and different types of linkage at the anomeric centres can now be systematically synthesized. Improved methods of reaction selectivity, new catalyst systems, and new separation and analytical techniques have made this possible. In particular, high-field n.m.r. spectroscopy in conjunction with two-dimensional n.m.r. spectroscopy permits excellent monitoring of reaction sequences and a complete analysis of n.m.r. spectra of complex oligosaccharides, in deblocked form also, is now possible.

The synthesis of specific oligosaccharide segments is of most interest in the glycoconjugates class, especially glycoproteins[2] and glycolipids.[3] In glycoproteins and glycolipids, the oligosaccharide section is frequently the determinant that defines the biological function of the substance, and there can be interaction between these carbohydrate structures and other proteins.[4] Examples of such interaction include antigen–antibody reactions or other receptor reactions, as well as interactions that are involved in cell communication, where the molecular structure of the carbohydrate sequence is of great significance as it determines the selectivity of the reaction with the protein or agglutinin.

With synthesized carbohydrate determinants or haptens and modified structures derived from them, information about the type of bonding between the carbohydrate residue and the protein can be obtained. Available data indicate that in this interaction predominantly hydrophobic attractive forces play a greater part than do hydrogen bonds.[5] By linking a synthesized carbohydrate chain to a solid

* Delivered at the Spring Meeting of The Royal Society of Chemistry Carbohydrate Group on 28th March 1983 at Heriot-Watt University, Edinburgh

[1] H. Paulsen, *Angew. Chem.*, 1982, **94**, 184; *Angew. Chem., Int. Ed. Engl.*, 1982, **21**, 155.
[2] J. Montreuil, *Adv. Carbohydr. Chem. Biochem.*, 1980, **37**, 157.
[3] C. C. Sweeley and B. Siddiqui, in 'The Glycoconjugates', ed. M. I. Horowitz and W. Pigman, Academic Press, New York, 1977, Vol. 1, p. 459.
[4] 'Carbohydrate–Protein Interaction', ed. I. J. Goldstein, ACS Symposium Series no. 88, Washington, D.C., 1979.
[5] R. U. Lemieux, in 'IUPAC Frontiers of Chemistry', ed. K. J. Laidler, Pergamon Press, Oxford and New York, 1982, p. 3; O. Hinsgaul, T. Norberg, J. Le Pendu, and R. U. Lemieux, *Carbohydr. Res.*, 1982, **109**, 109.

substrate, an absorber can be obtained on which the protein involved in the interaction can be selectively absorbed, purified, and isolated. This process can be used to obtain agglutinins and enzymes; in addition, the selectivity of lectins can be investigated. If the synthesized carbohydrate determinants are coupled to protein substrates, the resulting synthetic antigens make immunostimulation possible, or even the development of monoclonal antibodies.

There are thus numerous areas of use, of great biological interest, for selectively synthesized complex oligosaccharides. In this article are described the basic methods for the selective synthesis of oligosaccharides that are found in carbohydrate segments, in *O*-glycoproteins as well as in *N*-glycoproteins.

2 Methods of Glycosidic Linkage to Oligosaccharides

For selective glycosidic linkage in an oligosaccharide synthesis, three fundamental procedures, depending on the desired anomeric linkage, have proved to be particularly suitable:[1]

(i) the neighbouring-group-assisted procedure (Scheme 1), by which β-glycosidic linkages in the D-*gluco* and D-*galacto* series and α-glycosidic linkages in the D-*manno* series can be synthesized;

(ii) the *in situ* anomerization procedure (Scheme 2), which permits an α-glycosidic linkage in the D-*gluco* and D-*galacto* series; and

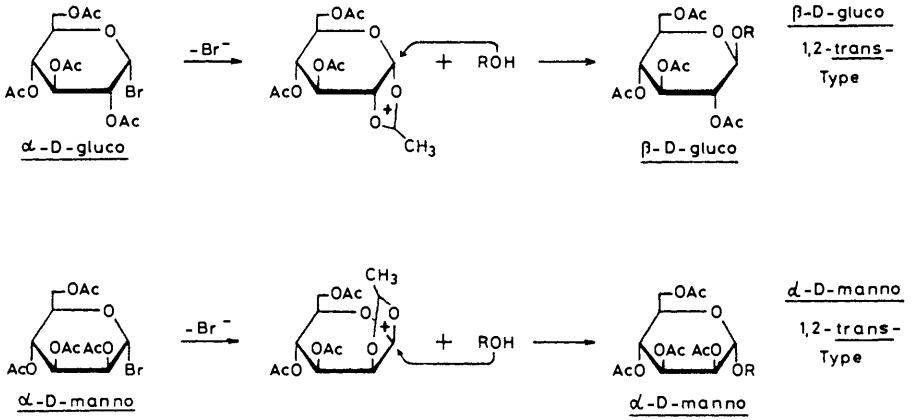
(iii) the heterogeneous catalyst procedure (Scheme 3), by which β-glycosidic linkages in the D-*manno* series can be synthesized.

Overall, the glycosyl halides have proved to be the most successful starting materials for all types of reaction.

In the neighbouring-group-assisted procedure (Scheme 1), it is, as a rule, the more stable α-D-halide with a neighbouring-group-active substituent at C-2, *e.g.* an *O*-acetyl group, that is substituted. There then results, *via* a carboxonium ion, a stabilized cyclic acyloxonium intermediate that can be opened at the anomeric centre by a nucleophile only in a direction *trans* to the substituent at C-2 (1,2-*trans* type). As can be seen in Scheme 1, in the D-*gluco* series this opening results in the β-D-glucopyranoside; analogously, in the D-*galacto* series the reaction results in the β-D-galactopyranoside. Since, however, in the D-*manno* series the substituent at C-2 has an inverted configuration, the result is also an inverted cyclic acyloxonium intermediate. The corresponding *trans* opening (1,2-*trans* type) results in this case, as can be seen in Scheme 1, in the α-D-mannopyranoside, also inverted. The catalysts that are generally used are listed in Scheme 1; as the intermediate is polar, the solvent should be polar or have medium polarity.

It should be mentioned that even in modified form the neighbouring-group-assisted procedure gives very good results if a β-D-1-*O*-acetate group is substituted instead of the halide, whereby the 1-*O*-acetate group with neighbouring-group support functions as a leaving group.[6] The analogous cyclic acyloxonium intermediates thus formed react correspondingly. In this case Lewis acids are used as catalysts. Trimethylsilyl triflate has proved to be particularly effective. This

[6] H. Paulsen, M. Paal, and M. Schultz, *Tetrahedron Lett.*, 1983, **24**, 1759.



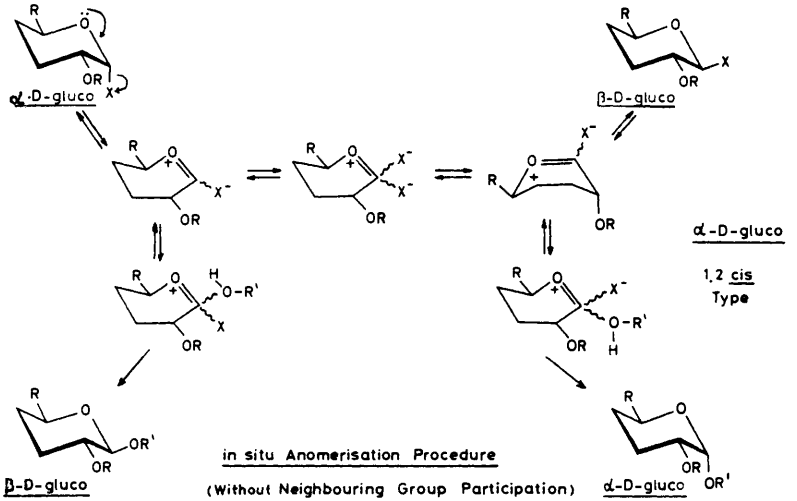
Scheme 1

method, however, requires an especially stable system of protecting groups, especially with acyl residues, as otherwise side-reactions and degredation take place.

In the *in situ* anomerization procedure a neighbouring-group non-active substituent must be present at C-2. Use is made here of the possibility of producing an equilibrium, *via* suitable catalysts, between the α-halide and the β-halide that establishes itself quickly across ion pairs (see Scheme 2).[7] As the β-D-halide (top right) is destabilized by the anomeric effect, there is great excess of α-D-halide (top left) in the thermodynamically controlled equilibrium. If the kinetics of the glycosidation reaction are considered, however, it is found that the reaction of the unstable β-D-halide to give the α-D-glycoside (right path in Scheme 2) is fast in comparison to the conversion of the more stable α-D-halide into the β-D-glycoside (left path in Scheme 2). Under certain reaction conditions this difference in reaction rate can be used such that the reaction proceeds almost exclusively along the fast path from the β-D-halide to the desired α-D-glycoside.[7,8] This also happens if there is a low concentration of the required β-D-halide, but provision is made for it to be re-formed quickly by an effective catalyst across the prescribed equilibrium. In this way selective α-D-glycoside synthesis (1,2-*cis* type) is possible in the D-*gluco* and D-*galacto* series.

Aside from trialkylammonium halides, the most suitable catalysts for starting from α-D-halides in the *in situ* anomerization procedure are mercury salts and silver perchlorate and silver triflate. In this connection modified compounds, too, could

[7] R. U. Lemieux, K. B. Hendriks, R. V. Stick, and K. James, *J. Am. Chem. Soc.*, 1975, **97**, 4056.
[8] H. Paulsen and Č. Kolář, *Chem. Ber.*, 1981, **114**, 306.

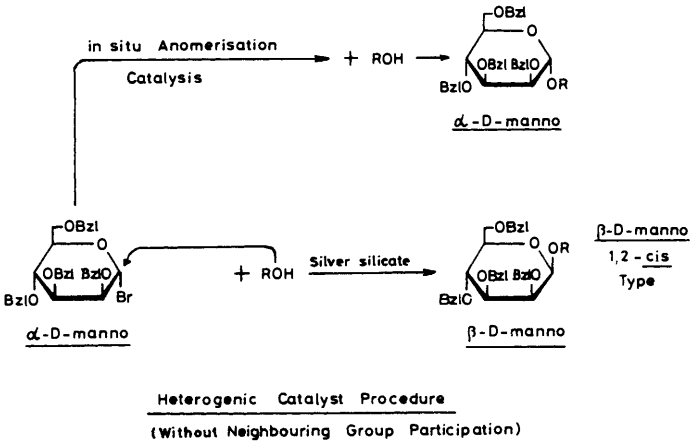


Scheme 2

react as intermediates in place of pyranosyl halides, *e.g.* pyranosyl perchlorates or pyranosyl triflates. For those intermediates, however, all the corresponding stabilities and rates of reaction would be valid so that Scheme 2 is likewise applicable. The scheme was therefore drawn up in a general form. Completely corresponding selectivities regarding the formation of α-D-glycosides are observed with these very much more effective catalysts. Difficulties in selectivity can occur in very reactive hydroxy-groups as the difference between the rate of formation of the α-D-glycosides and of the β-D-glycoside becomes smaller.[9,10] The best selectivities are achieved with hydroxy-groups of medium reactivity.[9]

The *in situ* anomerization procedure would also yield the α-D-mannopyranoside in the D-*manno* series completely analogously to Scheme 2, as the β-D-mannopyranosyl halide is likewise destabilized by the anomeric effect (Scheme 3). In the D-*manno* series the neighbouring-group-assisted procedure and the *in situ* anomerization procedure thus likewise yield the α-D-mannopyranoside and not the β-D-mannopyranoside. In order to achieve this, neighbouring-group reaction and any catalyst of the *in situ* anomerization procedure must therefore be strictly avoided. Only a heterogeneous, insoluble catalyst, which must be applied in the heterogeneous phase under neutral conditions, should be used. Besides other silver catalysts, the very reactive silver silicate catalyst is particularly suitable.[11,12] With this catalyst an α-D-mannopyranosyl halide reacts by inversion to the β-D-mannopyranoside (Scheme 3, 1,2-*cis* type). A very reactive halide and a

[9] H. Paulsen and O. Lockhoff, *Chem. Ber.*, 1981, **114**, 3079.
[10] H. Paulsen and J.-P. Hölck, *Carbohydr. Res.*, 1982, **109**, 89.
[11] H. Paulsen and O. Lockhoff, *Chem. Ber.*, 1981, **114**, 3102.
[12] H. Paulsen, W. Kutschker, and O. Lockhoff, *Chem. Ber.*, 1981, **114**, 3233.



Scheme 3

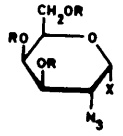
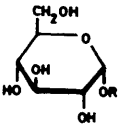
hydroxy-group of good to medium reactivity are necessary for this heterogeneous catalyst procedure; otherwise selectivity decreases.[11]

All three linkage methods are used in the synthesis of the oligosaccharide chains of glycoproteins. For every glycosidic linkage reaction, however, the three parameters given in the Table must be carefully assessed and balanced,[1] namely the reactivity of the halide, that of the catalyst, and that of the hydroxy-component. The reactivity of the pyranosyl halide varies within wide limits with variation of the blocking-group. Ether-substituted compounds are always more reactive than acyl-substituted compounds. The scale of reactivity of the catalysts is also extraordinarily wide. None of the catalysts given in the Table should, of course, be used in the heterogeneous-catalyst procedure. The reactivity of the hydroxy-group is unfortunately the most difficult to assess, even though this very group is often of considerable importance for the course of the reaction. In the linkage step it is always the reactions of polyfunctional compounds that are involved, for which reactivities depend very much on the kind of blocking system. The result of this is that for each glycosidic linkage step optimum conditions should be chosen carefully from the parameters given in the Table in order to find the best compromise between high selectivity and satisfactory yield.

3 *O*-Glycoproteins

The basic structure of the carbohydrate part of the *O*-glycoproteins is shown in Figure 1. It can be seen that the *N*-acetylgalactosamine is linked α-glycosidically to the hydroxy-group of serine or threonine in the peptide chain. As a further saccharide unit, a galactose unit is linked $\beta(1\rightarrow3)$-glycosidically to it. In addition, one or two neuraminic acid residues are linked to this disaccharide unit. This fundamental structure is found in most *O*-glycoproteins.

Table *Reactivity parameters determining selectivity and yield in oligosaccharide syntheses*

Halogen	*Catalyst*	*Alcohol*
	Et_4NBr/molecular sieve (4 Å) $Hg(CN)_2$ $Hg(CN)_2/HgBr_2$ $HgBr_2$/molecular sieve (4 Å) $AgClO_4$, Ag_2CO_3 Ag triflate, Ag_2CO_3	HO—Me > > HO—CH_2R > 6-OH 
R ~ Bzl > Bz ≳ Ac X ~ I > Br > Cl		6-OH > > 3-OH > 2-OH > 4-OH

Glycophorin, the main glycoprotein of the erythrocytes, is an interesting compound. It contains a chain of 131 amino-acids of which the hydrophobic part, amino-acids 71—90, is anchored in the membrane.[13] The part of the peptide chain, amino-acids 1—70, that extends into the exocellular space contains 16 *O*-glycoprotein side-chains and one *N*-glycoprotein side-chain. There are two

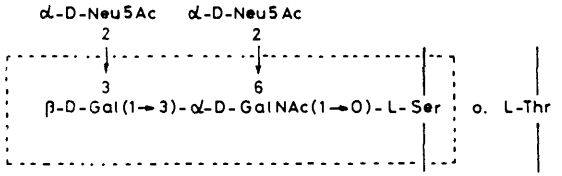


Figure 1 *Fundamental structure of the oligosaccharide chain of* O*-glycoproteins*

different types, whose terminal amino-acids 1—5 are shown in Figure 2. The structure with L-Leu as end group has N-antigen character and the structure with L-Ser as end group has M-antigen character.[13] If glycophorin is treated with neuraminidase, the neuraminic acid residues are cleaved off. T-Antigenicity is ascribed to the disaccharide fundamental structure thus exposed. The synthesis of these kinds of sequences and segments is of great interest.

The production of the α-glycosidic linkage of D-galactosamine to L-serine would indicate the use of the *in situ* anomerization procedure. The two azido-sugars shown in Figure 3 are suitable as neighbouring-group non-active halides.[10] This reaction demonstrates very well the complications of the process that arise when a very reactive hydroxy-component is present, as in the primary hydroxy-group of L-serine. In Figure 3 it can be seen that in the reaction of the α-D-halide with a serine derivative the *in situ* anomerization procedure yields a highly unsatisfactory selectivity of α:β-product, of only 2:1.[10] For better selectivity it would be necessary to reduce the reactivity of the halide, which in this case, however, is not possible

[13] M. Tomita, H. Furthmayer, and V. T. Marchesi, *Biochemistry*, 1978, **17**, 4756.

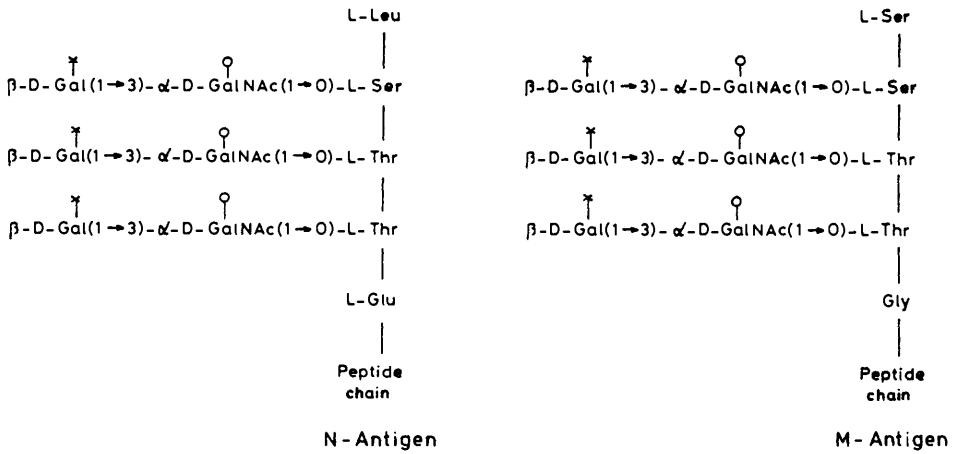


Figure 2 *End groups of glycophorin of the N-antigen and M-antigen types*
[$\overset{*}{|}$ = α-D-Neu5Ac(2→3) −, $\overset{\circ}{|}$ = α-D-Neu5Ac(2→6)-]

since all the hydroxy-groups are already occupied with acyl groups that lower reactivity.

The β-D-chloride, which is directly isolatable in this case, can be substituted for bromide in the reaction. Figure 3 shows that the selectivity of the reaction with β-D-chloride is indeed better but not yet completely satisfactory, with an α:β-product ratio of 7:2 to 9:2.[14] The reason for this is that during the reaction β-D-chloride partially anomerizes to α-D-chloride, which by inversion then yields the undesired β-D-glycoside. By the addition of toluene to the non-polar solvent dichloromethane, which is generally used, it is possible to suppress anomerization.[10] Despite the consequent lowering of the reaction rate, the selectivity of α:β-product (19:1) is now very good (Figure 3).[10]

Using the process it is now possible to synthesize an equivalent *O*-glycopeptide (Scheme 4). The reaction of β-D-halide (1) with the serine derivative (2) results in the α-glycosidically linked product (3), which can be converted *via* compound (4) into the corresponding derivative (5).[10] In a neighbouring-group-assisted procedure α-D-halide (6) of D-galactose can now be linked to the 3-OH group of compounds (5) to form the β-D-glycosidically linked disaccharide (7). By deblocking, the fundamental structure (8), the so-called T-determinant, attached to serine can be obtained from compound (7).[10] A completely analogous reaction sequence can be carried out with threonine.[10]

For the synthesis of larger oligosaccharides, block syntheses in which di- or tri-saccharide halides are applied have in general proved to be very successful.[15] In this way larger saccharide units are obtained more quickly, and with a reduction

[14] B. Ferrari and A. A. Pavia, *Carbohydr. Res.*, 1980, **79**, C1.
[15] H. Paulsen and A. Bünsch, *Carbohydr. Res.*, 1982, **100**, 143.

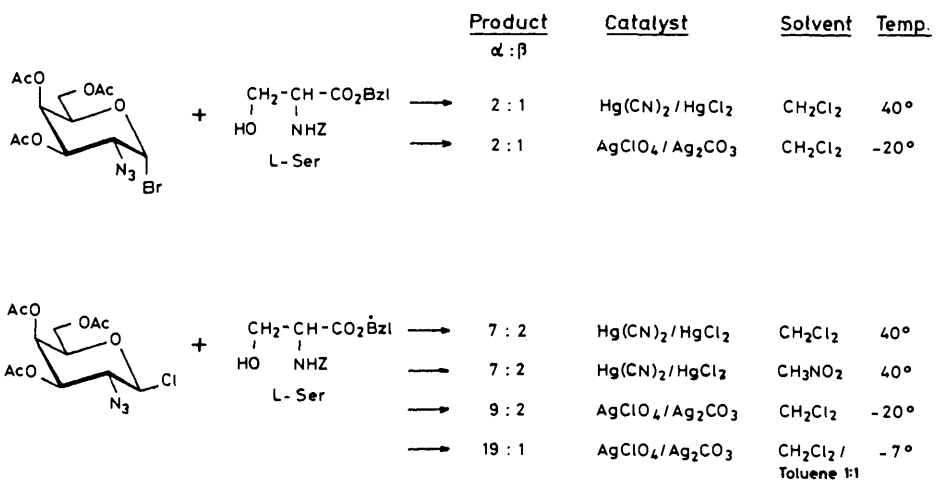


Figure 3 *Reaction of* 3,4,6-*tri*-O-*acetyl*-2-*azido*-2-*desoxy*-D-*galactopyranosyl halides with* N-*benzyloxycarbonyl*-L-*serine benzylester in various conditions*

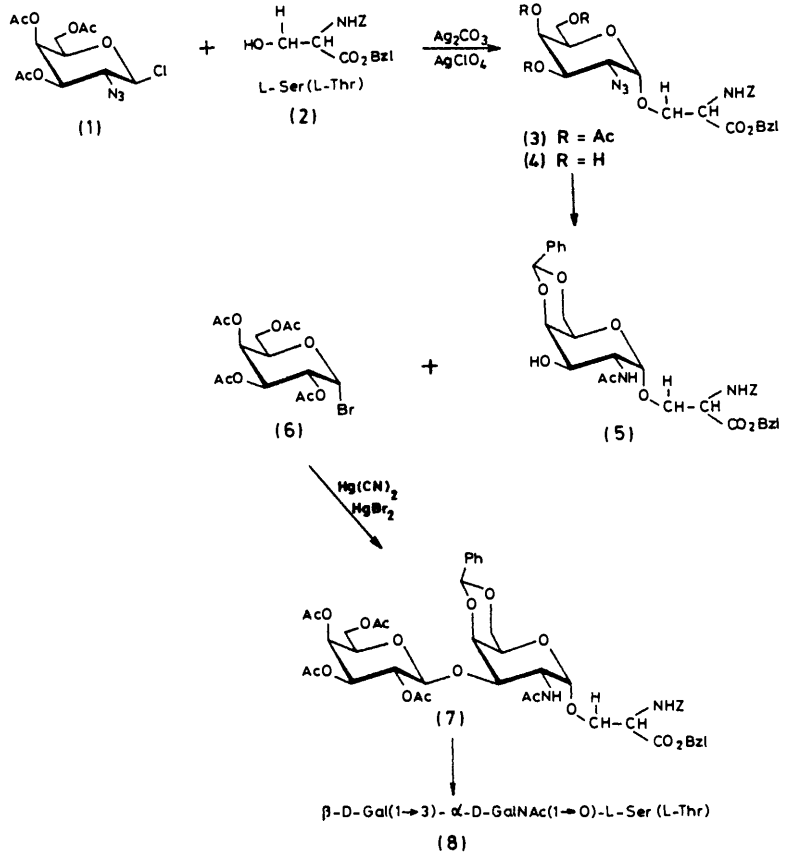
in the number of necessary separation processes. A synthesis block of β-D-Gal(1→3)-D-GalNAc would therefore be very desirable. This synthesis block could then be used for direct linking to hydroxyamino-acids and could possibly be linked to larger peptides as well. In the linking step the stereoselectivity of the reaction with the reactive hydroxy-group of serine or threonine would again be very important.

It became obvious that the production of the desired block could be brought about by reaction of the halide (6) with the easily obtainable derivative (9) (Scheme 5). Unfortunately, however, the expected disaccharide (10) is obtained in low yield only, and with bad selectivity. Surprisingly, the 3-OH group is extraordinarily unreactive in compound (9).[6]

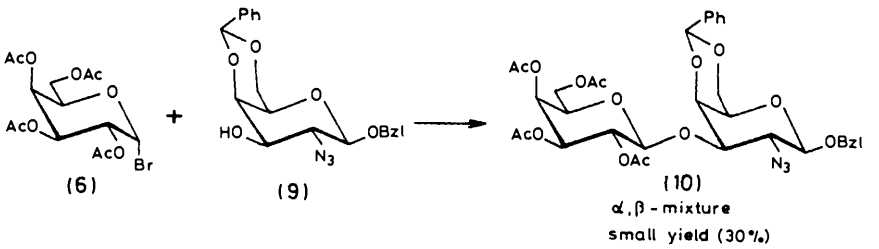
A solution to this important problem was found in a completely new synthesis route (Scheme 6).[6] To begin with, the azidonitrate (11), which can be obtained easily by azidonitration of D-galactal,[16] is converted with sodium methoxide directly into β-D-methylglycoside (12). Selective benzoylation can be carried out with compound (12), resulting in product (13), in which only the 3-OH group has not been substituted. In compound (13) there are now only acyl protecting groups so that glycosidation of (13) with the penta-acetate (14) in the presence of trimethylsilyl triflate *via* the modified neighbouring-group-assisted procedure is successful.[6] In this way the desired disaccharide (15) can be obtained in 85%yield. The methylglycoside (15) can be relatively easily acetolysed to the α-acetate (16), from which by reaction with titanium tetrabromide in anhydrous conditions[15] the newly functionalized synthesis block can be obtained as the α-bromide (17).[6]

Excellent results were obtained for this new synthesis block (17) with the serine derivative (2) or the corresponding threonine derivative.[6] From the reaction

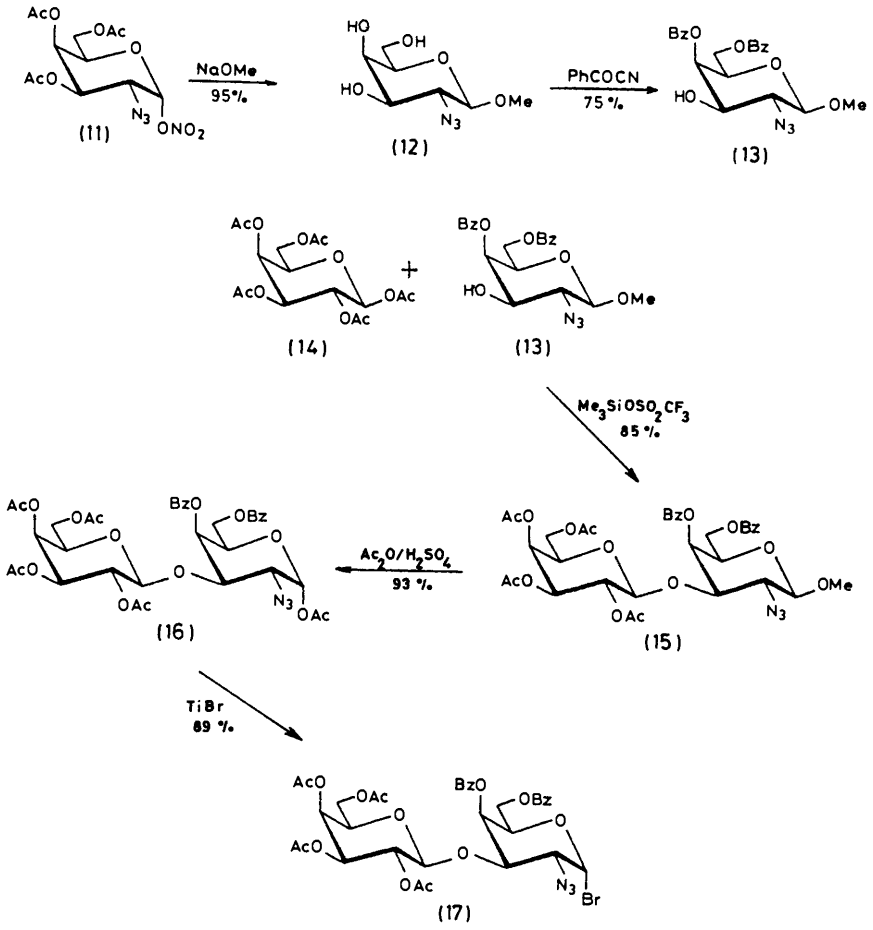
[16] R. U. Lemieux and R. M. Ratcliffe, *Can. J. Chem.*, 1979, **57**, 1244.



Scheme 4



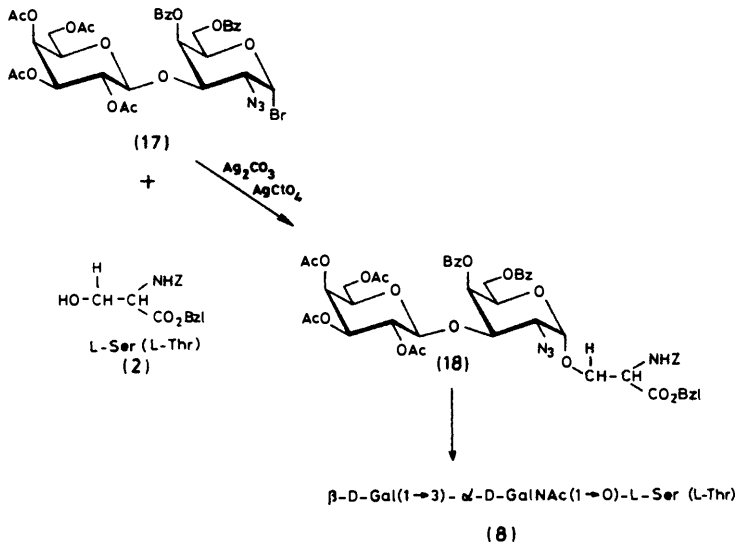
Scheme 5



Scheme 6

(Scheme 7) between compounds (17) and (2) in the presence of silver perchlorate/silver carbonate, the desired glycopeptide (18) is obtained directly from the α-D-bromide *via* the *in situ* anomerization procedure in over 80% yield. The formation of a β-glycosidic product is not observed in this case. The glycopeptide (8) can be obtained by deblocking. With the corresponding threonine derivative the reaction proceeds in exactly the same way, with corresponding selectivity.[6]

It is remarkable that the glycoside synthesis from the α-D-bromide (17) proceeds completely stereoselectively even though the reactive hydroxy-group of serine is again available. The reason for this is that the reactivity of the disaccharide block

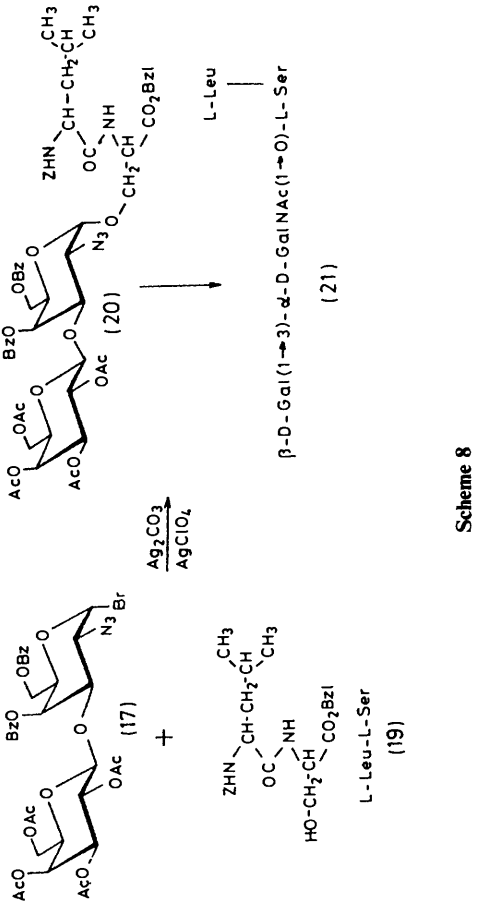


Scheme 7

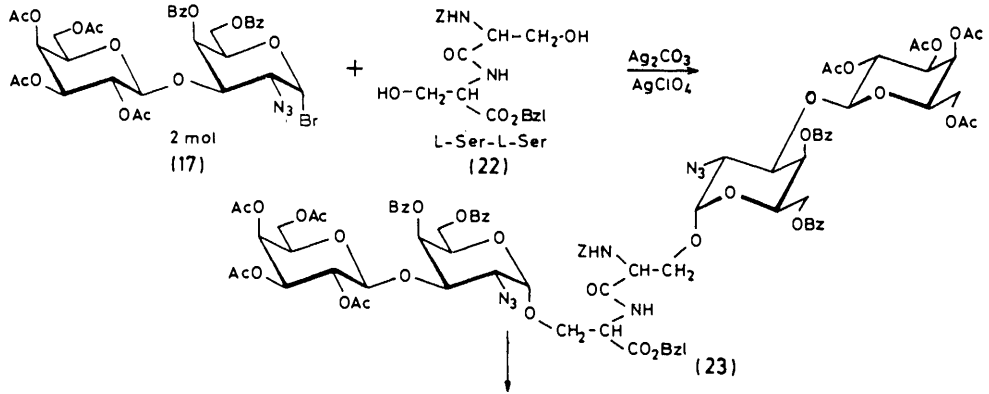
(17) is considerably lower than that of the monosaccharide halides in Figure 3. In this way it becomes feasible to use the *in situ* anomerization procedure since the difference in reaction rate leading to the α- and β-glycosidically linked product is again large enough to achieve good selectivity in favour of the α-glycosidic product. The lower reactivity of the disaccharide halide (17) which is tuned exactly to the reaction, can be attributed on the one hand to the additionally linked galactose residue but on the other hand also to the presence of two benzoyl groups. In compound (17), therefore, an excellent synthesis block is available from which high selectivities in the reaction with hydroxyamino-acids can be expected.

To check the reaction with peptides, the derivative of L-Leu-L-Ser (19) was allowed to react with the synthesis block (17) (Scheme 8).[6] The coupling reaction proceeds similarly with high selectivity and produces exclusively the α-glycosidic glycopeptide (20) in 70% yield. Deblocking (20) results in compound (21), which contains the terminal structural element of glycophorin of the N-antigen type (Figure 2).

It is of the greatest interest to test whether the synthesis block also reacts with several hydroxy-groups of corresponding serine- or threonine-containing peptide chains. In fact it has been possible to allow the derivative of L-Ser-L-Ser (22) to react with 2 moles of the synthesis block (17) in a stereoselectively simultaneous reaction (Scheme 9) to form the glycopeptide (23) with two carbohydrate chains.[6] In this case, too, as proved by careful two-dimensional n.m.r. spectroscopy, the reaction proceeds completely stereoselectively. Components of β-products are not detectable. This demonstrates impressively the efficiency and the high selectivity of


AcO
OAc
AcO
OAc
BzO
OBz
N3
Br
(17)
+
ZHN
CH-CH2-CH
CH3
CH3
OC
NH
HO-CH2-CH
CO2Bzl
L-Leu-L-Ser
(19)
Ag2CO3
AgClO4
(20)
ZHN
CH-CH2-CH
CH3
CH3
OC
NH
CH2-CH
CO2Bzl
L-Leu
β-D-Gal(1→3)-α-D-GalNAc(1→O)-L-Ser
(21)

Scheme 8


2 mol
(17)
L-Ser-L-Ser
(22)
Ag_2CO_3
$AgClO_4$
(23)
β-D-Gal(1→3)-α-D-GalNAc(1→O)-L-Ser
β-D-Gal(1→3)-α-D-GalNAc(1→O)-L-Ser
(24)

Scheme 9

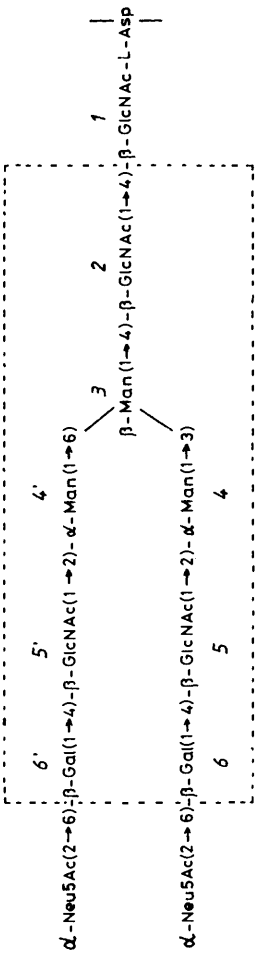


Figure 4 *Fundamental structure of the oligosaccharide chain of* N*-glycoproteins of the lactosamine type*

the synthesis block (17) and it is evident that it can react correspondingly with any serine- or threonine-containing peptide chains. The reaction with longer peptide chains would cause some difficulties concerning their solubility in dichloromethane, which is the solvent of glycoside syntheses. In this case, a step-by-step synthesis of *O*-glycopeptides with mixed substituted amino-acid derivatives, following the methods of the peptide synthesis, is possible to obtain higher *O*-glycopeptides which are part of the glycophorin.

4 *N*-Glycoproteins

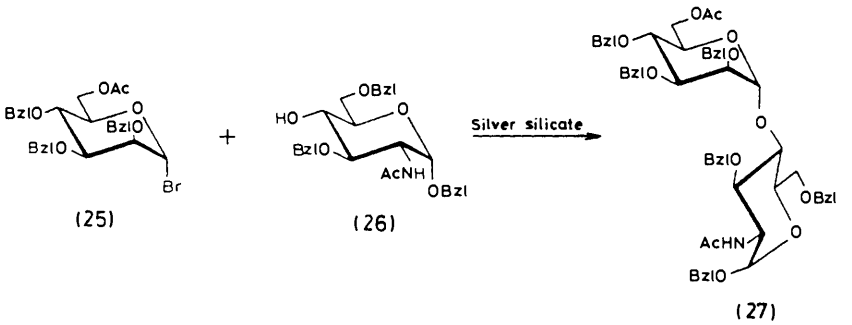
Figure 4 shows the fundamental structure of the oligosaccharide chain as it generally occurs in *N*-glycoproteins of the lactosamine type. The chain is linked β-*N*-glycosidically to the amide group of asparagine, which itself is part of the total peptide chain. The first unit on asparagine is a chitobiose unit, to which the particularly interesting branched middle section consisting of three mannose units is attached. In the lactosamine type, lactosamine antennae, which contain neuraminic acid end-groups, are linked to both mannose residues. In the so-called mannose type of *N*-glycoprotein, the same core structure of chitobiose and branched mannose trisaccharide exists. In this case, though, further mannose groups of different chain lengths are linked to the mannose units. There are numerous variants in which the basic framework is the same but additional branches are present.[2]

The main problem of the synthesis of such an oligosaccharide sequence lies in the production of the β-D-mannosidic linkage in the partial sequence β-D-Man(1→4)-D-GlcNAc (residues 3 and 2 in Figure 4). Until now it has not been possible to prepare this unit by direct linkage, but only indirectly by first of all preparing β-D-Glc(1→4)-D-GlcNAc and then subsequently converting the glucose unit in several steps into a mannose unit.[17] For direct preparation of this key disaccharide the heterogeneous-catalyst procedure should be suitable (Scheme 3) in which, in particular, the reactive silver silicate catalyst could be tested.

The reaction (Scheme 10) in the presence of silver silicate catalyst, however, of the reactive mannose halide (25) with the easily available glucosamine derivative (26) (containing a free 4-OH group) slowly yields a strong preference for the undesired α-D-glycoside (27).[18] It is known that the 4-OH group in derivative (26) is particularly unreactive and therefore should be responsible for the slow reaction. The reaction clearly proceeds to some extent *via in situ* anomerization or *via* the direct formation of a carboxonium ion, both of which should preferentially result in the α-D-glycoside (27).

It was thus necessary to increase the reactivity of the 4-OH group. This is possible since the inverse conformation, as in the 1,6-anhydro-compound (29), can be made to react (Scheme 11). The now free axial 4-OH group has higher reactivity, and in this case the heterogeneous-catalyst procedure in the presence of a silver silicate catalyst does in fact result in satisfactory selectivity, as a β:α-

[17] C. Augé, C. D. Warren, R. W. Jeanloz, M. Kiso, and L. Anderson, *Carbohydr. Res.*, 1980, **82**, 85.
[18] H. Paulsen, R. Lebuhn, and O. Lockhoff, *Carbohydr. Res.*, 1982, **103**, C7.



Scheme 10

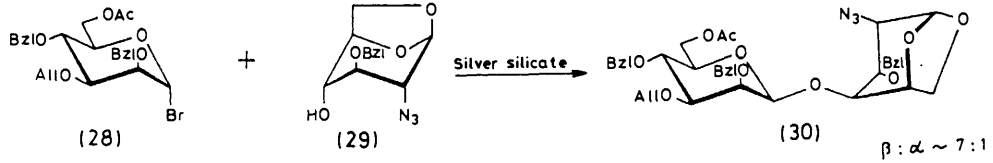
product ratio of 7:1 is found, and the desired β-glycosidic product (30) can be isolated in 65% yield.[18,19] The crucial key disaccharide, which was most suitable for all further synthesis reactions, was thus synthesized.

If the key disaccharide (30) is deacetylated to compound (33), coupling with the mannose halide (35) can be achieved at the liberated 6-OH group in a neighbouring-group-assisted procedure (Scheme 12). The trisaccharide (31) is then obtained by the production of a new α-glycosidic link to mannose. Hence the free trisaccharide (32), which contains the additional mannose residue in an α(1→6)-glycosidic link, can be obtained *via* a series of de-blocking steps.[18,19] If the basic disaccharide (30) is deallylated, primarily to compound (34), which is deblocked at the 3-OH group, a corresponding coupling with the mannose derivative (35) to the trisaccharide (36) is possible. The corresponding deblocking then yields the trisaccharide (37), in which an additional mannose residue is linked α(1→3)-glycosidically to the disaccharide.[18,19]

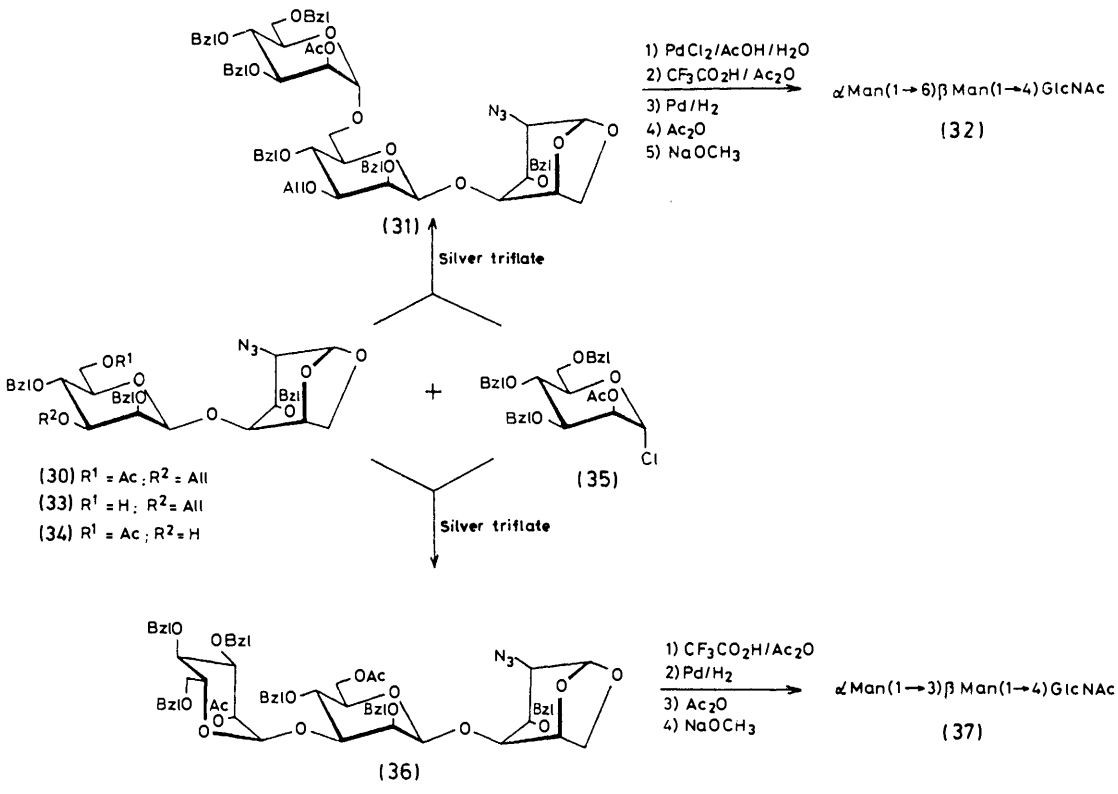
If in the disaccharide (30) both the 6-OH and the 3-OH groups are free, compound (38) is obtained. In a neighbouring-group-assisted procedure there are now two residues of the mannose halide (35) to be linked to compound (38) in a simultaneous reaction (Scheme 13), and the branched tetrasaccharide (39) can be obtained in one step. In a sequence of deblocking steps the desired tetrasaccharide (40), *i.e.* the central branching structure of the core sequence, can be obtained.[18,19] These examples show that, in principle, all sequence fragments can be synthesized.

As the following sequence in the oligosaccharide chain of the lactosamine type of *N*-glycoprotein, in each case a lactosamine unit is linked to the two terminal mannose groups, as shown in Figure 4. It would be very useful if here, too, a block synthesis could be carried out in order to link the lactosamine in one step. From lactosamine an excellent, reactive synthesis block (42) is available.[20] This halide contains a 2-phthalimido-grouping. In a neighbouring-group-assisted reaction this bulky grouping ensures stereoselective glycosidation to a β-D-glycoside. A β-glycosidic linkage with lactosamine with the 2-OH group of the terminal mannose unit is desired. In a trial reaction (Scheme 14), therefore, the trisaccharide (41), which has a free 2-OH group and is easily produced by deacetylation of

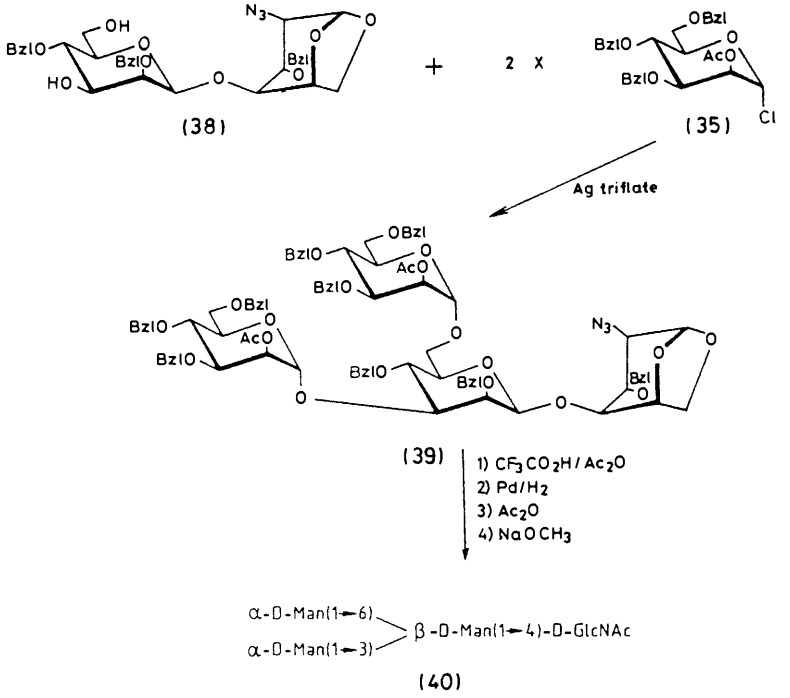
[19] H. Paulsen and R. Lebuhn, *Liebigs Ann. Chem.*, 1983, 1047.
[20] J. Arnap and J. Lönngren, *J. Chem. Soc., Perkin Trans. 1*, 1981, 2070.


OAc
BzlO
O
BzlO
AllO
Br
(28)
+
O
OBzl
O
HO
N3
(29)
Silver silicate
OAc
BzlO
O
BzlO
AllO
O
N3
O
O
Bzl
O
(30)
β : α ~ 7 : 1

Scheme 11


1) PdCl2/AcOH/H2O
2) CF3CO2H/Ac2O
3) Pd/H2
4) Ac2O
5) NaOCH3
αMan(1→6)βMan(1→4)GlcNAc
(32)
(31)
Silver triflate
(30) R1 = Ac ; R2 = All
(33) R1 = H ; R2 = All
(34) R1 = Ac ; R2 = H
(35)
Silver triflate
1) CF3CO2H/Ac2O
2) Pd/H2
3) Ac2O
4) NaOCH3
αMan(1→3)βMan(1→4)GlcNAc
(37)
(36)

Scheme 12

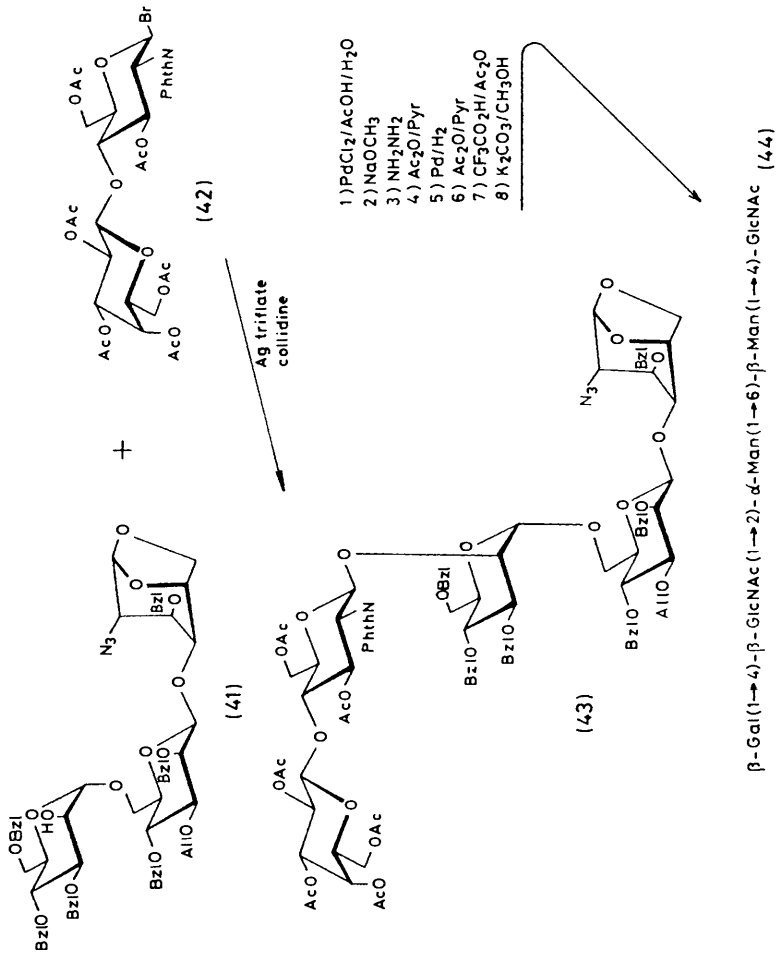


Scheme 13

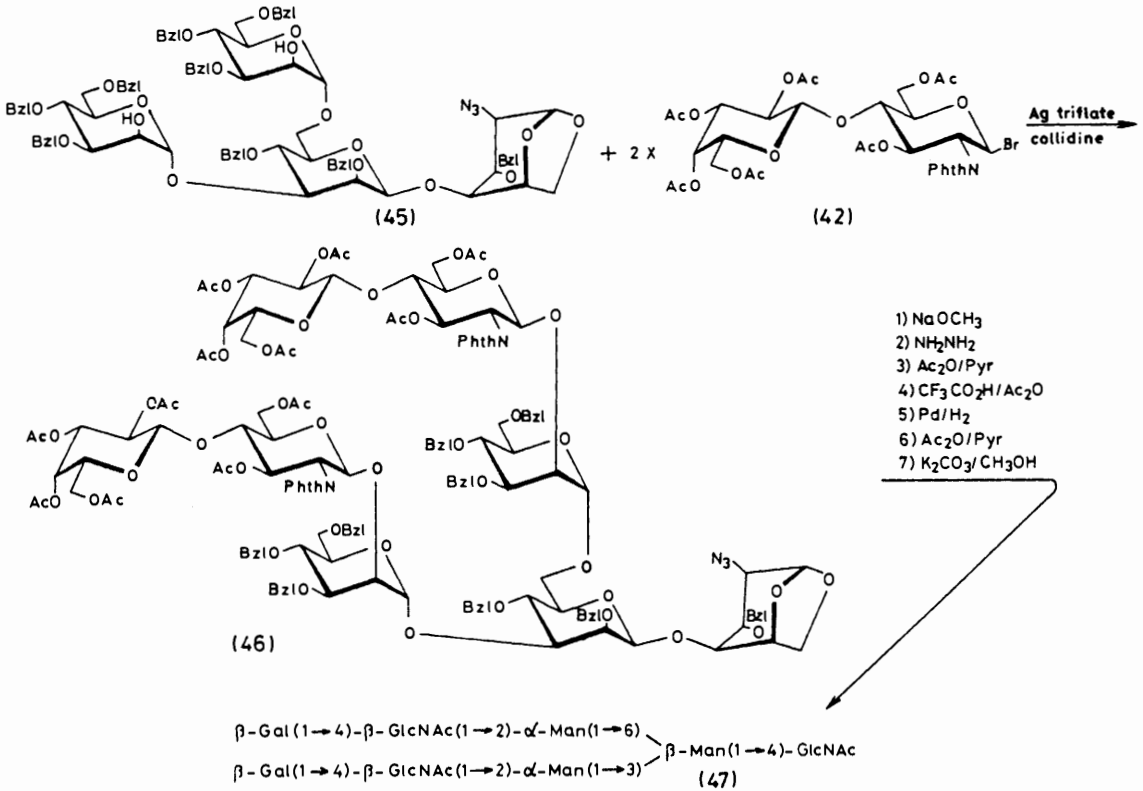
compound (31), was allowed to react with the synthesis block (42) in the presence of silver triflate/collidine. In this way, the desired pentasaccharide (43) is obtained, which contains all the required α- and β-glycosidic linkages. In a series of deblocking reactions the free pentasaccharide (44) can be prepared from compound (43).[21]

An obvious substitution is that of the branched tetrasaccharide (45), which has two free 2-OH groups on the mannose residues and can be easily produced by deacetylation of compound (39), for the reaction with the lactosamine synthesis block (42) (Scheme 15). In fact two moles of (42) can be linked simultaneously to the 2-OH group of (45) under the same conditions in one reaction step, and the desired octasaccharide (46) synthesis is achieved in 70% yield.[21,22] It should be noted that in this case it is not the condensation step that presents the real difficulty. The series of difficult deblocking steps must be undertaken with great care in order not to threaten the newly formed glycosidic links. The cleavage of phthalimido-groups takes place in methanol with hydrazine under very mild conditions. In the

[21] H. Paulsen and R. Lebuhn, *Angew. Chem.*, 1982, **94**, 933; *Angew. Chem., Int. Ed. Engl.*, 1982, **21**, 926.
[22] H. Paulsen and R. Lebuhn, *Carbohydr. Res.*, 1984, **125**, in press.


Ag triflate
collidine
1) PdCl2/AcOH/H2O
2) NaOCH3
3) NH2NH2
4) Ac2O/Pyr
5) Pd/H2
6) Ac2O/Pyr
7) CF3CO2H/Ac2O
8) K2CO3/CH3OH
(41)
(42)
(43)
β-Gal(1→4)-β-GlcNAc(1→2)-α-Man(1→6)-β-Man(1→4)-GlcNAc (44)

Scheme 14


(45)
(42)
Ag triflate
collidine
1) NaOCH3
2) NH2NH2
3) Ac2O/Pyr
4) CF3CO2H/Ac2O
5) Pd/H2
6) Ac2O/Pyr
7) K2CO3/CH3OH
(46)
β-Gal(1→4)-β-GlcNAc(1→2)-α-Man(1→6)
β-Gal(1→4)-β-GlcNAc(1→2)-α-Man(1→3)
β-Man(1→4)-GlcNAc
(47)

Scheme 15

opening of the 1,6-anhydro-ring of the reducing unit with trifluoroacetic acid/acetic anhydride, the acid component must be measured out very accurately so that no further hydrolyses take place. Deblocking proceeds in the prescribed way, and the free octasaccharide (47), which contains the central structure shown in Figure 4, can thus be obtained.[21,22]

Studies were also made to examine the possibility of lengthening the oligosaccharide chain shown in Figure 4 to the asparagine end. For these studies the starting point was again the key disaccharide (30), from which by modification of the protecting groups compound (48) was obtained. After conversion of the azido-group into a phthalimido-group (Scheme 16), the 1,6-anhydro-ring in the resulting compound (49) can be opened by acetolysis with trifluoroacetic acid/acetic anhydride.[23] The acetolysis yields the β-acetate (50), which can be functionalized with titanium tetrachloride to the halide (51). In a neighbouring-group-assisted procedure compound (51) can be coupled with the glucosamine derivative (53) to the trisaccharide (52).[23] This trisaccharide contains the chitobiose unit, which is linked directly to asparagine. The trisaccharide (52) is, moreover, a good starting product for further synthesis reactions. After removal of the two allyl groups, two further mannose residues should be able to link α-glycosidically to the 3-OH and 6-OH groups of the mannose unit.[23] This would be the route to the synthesis of the general saccharide core structure as it occurs both in the *N*-glycoproteins of the lactosamine type and in those of the mannose type.

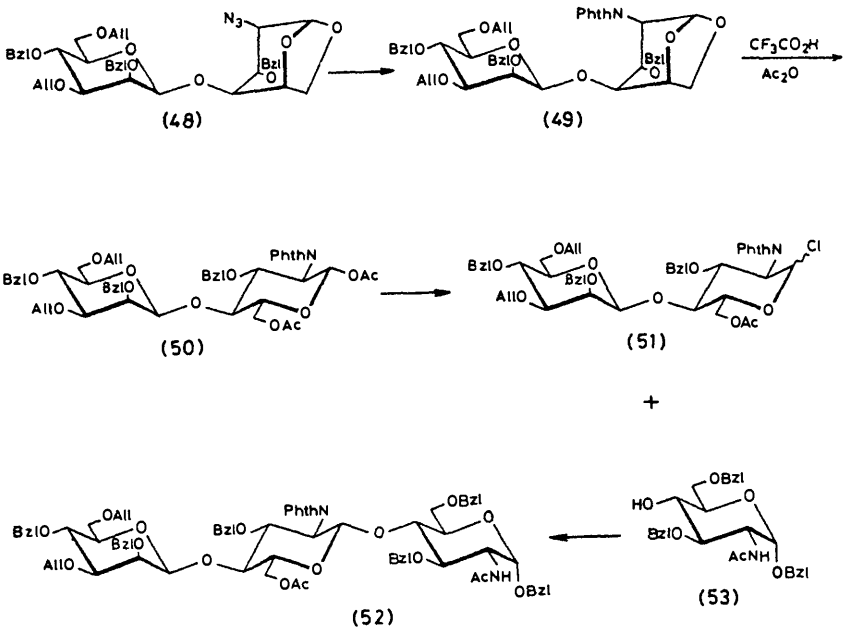
The described reaction sequences indicate that in principle all sequences that occur in *N*-glycoproteins can be chemically synthesized. This is, of course, also the case for all partial sequences that are contained in structures (47) and (52).

5 Neuraminic Acid Oligosaccharides

It would also be highly desirable to discover methods for linking *N*-acetylneuraminic acid to the terminal oligosaccharide chains of glycoproteins, as shown in Figure 4, and of particular interest here is the trisaccharide unit α-D-Neu5Ac(2→6)-β-D-Gal(1→4)-D-GlcNAc. This sequence occurs (see Figure 4) as the terminal segment in the oligosaccharide chain of *N*-glycoproteins of the lactosamine type.

Because of considerable difficulties encountered in synthesizing glycosides by the use of *N*-acetylneuraminic acid, the prospects for a linkage with a reactive 6′-OH group from lactosamine were still the most promising. A corresponding building block was therefore produced from the easily obtained derivative of 2-azido-2-desoxy-lactose (54) (Scheme 17).[24] The halide (54) was first of all converted into the benzylglycoside (55) and then, by variation of the protecting groups, into compound (56). Hydrolysis yielded compound (57), now containing two free hydroxy-groups, 4′-OH and 6′-OH,[25] of which the 6′-OH group should be much more reactive than the 4′-OH group. One would therefore expect preferential

[23] H. Paulsen and R. Lebuhn, *Carbohydr. Res.*, in press.
[24] H. Paulsen and J.-P. Hölck, *Liebigs Ann. Chem.*, 1982, 1121.
[25] H. Paulsen and H. Tietz, *Angew. Chem.*, 1982, **94**, 934; *Angew. Chem., Int. Ed. Engl.*, 1982, **21**, 927.



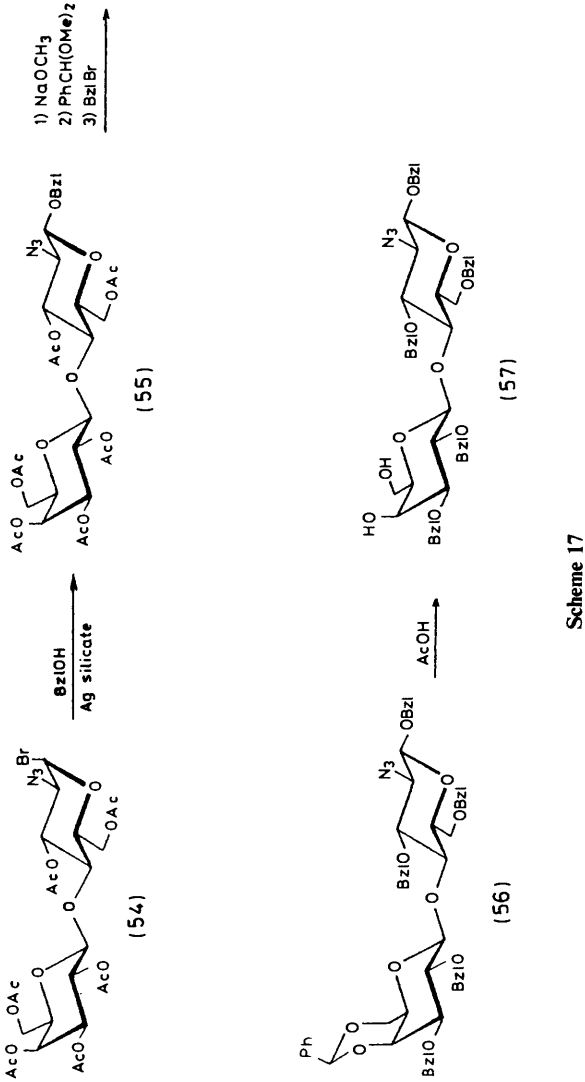
Scheme 16

reaction with the 6'-OH group, and (57) is therefore suitable as substrate for the reaction with an *N*-acetylneuraminic acid halide.

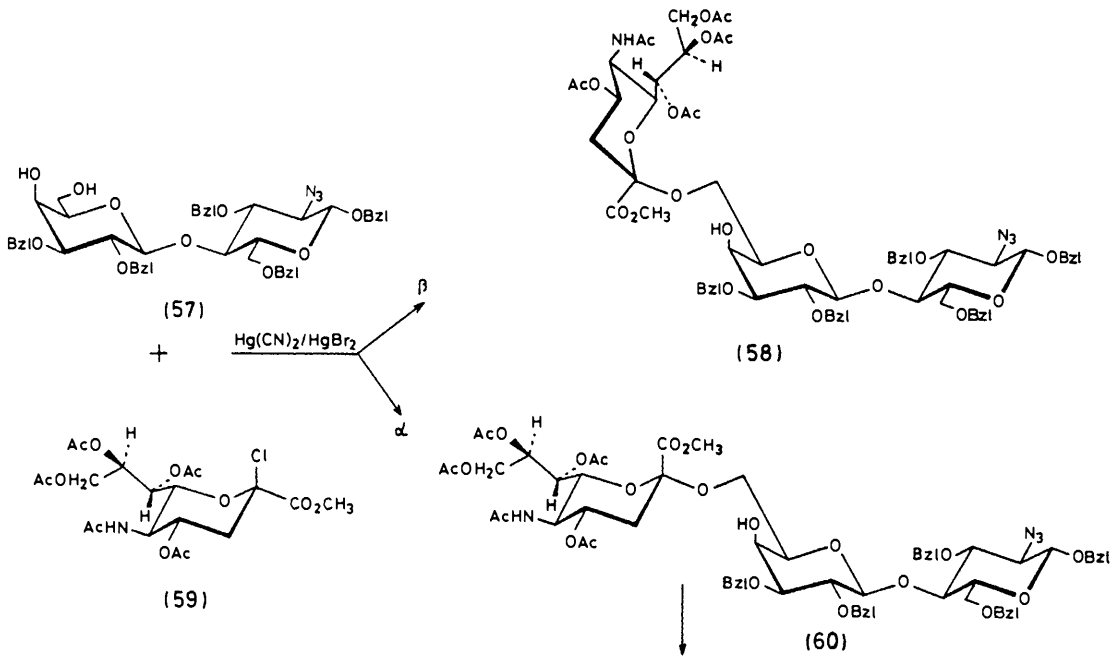
Glycoside syntheses using the halide of the neuraminic acid (59) therefore present considerable difficulties, as the dominating reaction in the presence of virtually all catalysts is elimination with HCl cleavage to give the unsaturated neuraminic acid.[26,27] The yield of neuraminic acid-containing oligosaccharides is therefore extraordinarily low.

It has now been discovered that, with the use of the mercury catalysts (mercury cyanide/mercury bromide) under controlled conditions, the undesired elimination reaction can be considerably suppressed (Scheme 18). Under these conditions, the halide of the neuraminic acid (59) reacted with the lactosamine building block (57) at 50% yield to form neuraminic acid-containing oligosaccharides.[25,26] The reaction does not proceed stereospecifically; the two anomeric compounds (58) and (60) are obtained, in an α:β-product ratio of about 1:1. Both the anomers (58) and (60) can, however, be separated by chromatographic means and can therefore be isolated in this way, both at about 25% yield. Complete deblocking of both

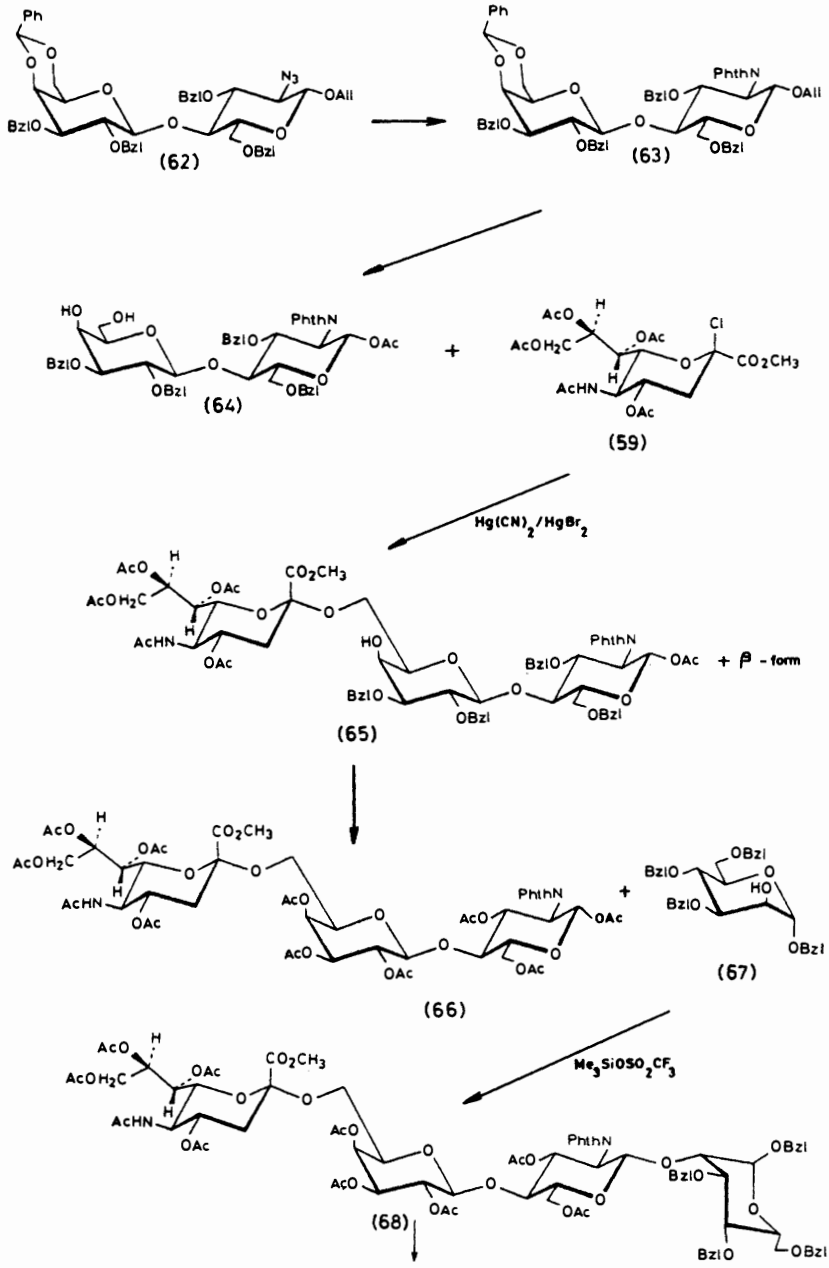
[26] H. Paulsen and H. Tietz, *Carbohydr. Res.*, 1984, **125**, in press.
[27] 'Sialic Acids', ed. R. Schauer, Cell Biology Monographs Vol. 10, Springer Verlag, Wien and New York, 1982.



Scheme 17


HO
OH
BzlO
OBzl
OBzl
BzlO
OBzl
OBzl
N3
(57)
+
Hg(CN)2/HgBr2
β
α
AcO
AcO2C
AcHN
OAc
OAc
Cl
CO2CH3
(59)
CH2OAc
NHAc
OAc
H
AcO
OAc
CO2CH3
(58)
(60)
α-D-Neu5Ac-(2→6)-β-D-Gal(1→4)-D-GlcNAc
(61)

Scheme 18


(62)
(63)
(64)
(59)
$Hg(CN)_2/HgBr_2$
(65)
+ β - form
(66)
(67)
$Me_3SiOSO_2CF_3$
(68)
α-D-Neu5Ac-(2→6)-β-D-Gal(1→4)-β-D-GlcNAc(1→2)-D-Man
(69)

Scheme 19

neuraminic acid trisaccharides was possible.[25,26] From the α-product (60) the free trisaccharide (61) is obtained in this way, and it represents the terminal unit of the saccharide chain of *N*-glycoproteins as shown in Figure 4.

The direct introduction of *N*-acetylneuraminic acid into larger oligosaccharides remains a problem.[27] It would therefore be useful to develop a neuraminic acid-containing synthesis block by means of which the anomeric unit could be converted into a reactive form that would permit linking to any oligosaccharides. The lactosamine residue in (60) could be most suitable in this respect if it were converted into the corresponding phthalimido-compound. With regard to the effectiveness of a glycosidic linkage, a functionalized phthalimido-compound has very good properties.

Therefore, the modified lactosamine derivative (63), which contains a phthalimido-group, was produced (Scheme 19) from compound (62).

The building block (64), which is similarly suitable for reaction with the *N*-acetylneuraminic acid halide (59), is obtained by cleaving the allyl glycoside (63) followed by acetylation and hydrolysis of the benzylidene group. The reaction of (64) with (59) under analogous conditions affords in corresponding yield the two anomeric trisaccharides, which can also be separated chromatographically.[28] The α-D-compound (65) produced in this way now has at the anomeric centre of the reducing unit an acetyl group that is once more functionalizable.

To this end it is necessary first of all to cleave off the benzyl ether groups in compound (65) by hydrogenation and to substitute them with acetyl groups. In this way a more stable compound is obtained that can be used directly in a glycosidation step under the conditions of the trimethylsilyl triflate method (see Scheme 6).[28] A new neuraminic acid-containing synthesis block is available for this.

In a test reaction, the β-acetate compound (66) was allowed to react with the mannose derivative (67), which has a free 2-OH group. In the presence of trimethyl triflate in a neighbouring-group-assisted procedure, the desired tetrasaccharide (68) is obtained by attachment of the new β-glycosidic link to mannose.[28] From (68) the tetrasaccharide chain (69) becomes accessible, representing the terminal sequence of one arm of the lactosamine antennae in the formula shown in Figure 4. This example shows that the block is suitable for further glycoside syntheses. It can thus serve to introduce neuraminic acid into oligosaccharides, with which otherwise practically no reaction would be possible.

6 Conformations of Oligosaccharide Chains

In several cases the oligosaccharide chain of glycoconjugates represents the biological determinant, so therefore the molecular shape (and also conformation) of the chain is of great interest. The molecular shape of the terminal oligosaccharide sequence determines the specificity of the reaction with proteins. Conclusions about the interaction with protein receptors can be drawn from data regarding the conformation of the oligosaccharide chain. For determining the conformation of

[28] H. Paulsen and H. Tietz, to be published.

the oligosaccharide chain, two methods are considered: a theoretical calculation, in which the mutual steric interactions of the saccharide units are taken into account, and the use of the modern methods of high-field n.m.r. spectroscopy, which yield information about the conformation in solution.

In the method of HSEA (hard-sphere exo-anomeric effect)[29] or GESA (geometric saccharide)[30] calculations, the pyranose ring is accepted to be in a rigid chair form, and corresponding data for the chair conformation are inferred from *X*-ray structure analyses. Further, the atomic radii and the exo-anomeric effect are taken into consideration. Rotations to give various conformations are carried out about the different anomeric linkages, and for anomeric linkages with secondary OH groups two angles (ϕ and ψ) can be varied. By varying both angles the energy minimum can be calculated,[29,30] and for glycosidic linkages with secondary hydroxy-groups the minimum was found to be relatively steep so that, in this kind of linking, one conformation is strongly preferred. For anomeric linkages with 6-OH groups, a further rotation about the C-5 – C-6 axis is possible. Taking into account this third variability of rotation, with this kind of linkage, therefore, more minima result, so that variable parts of conformations are present here. The most important minima in this case are a *gg* (*gauche–gauche*) and a *gt* (*gauche–trans*) conformation.[29,30]

For conformation analysis by the n.m.r. method, the ^{1}H n.m.r. spectrum of the substance being analysed must be assigned completely. This is very difficult for deblocked oligosaccharides as there is extensive overlapping of the ring protons of the various saccharide units. An assignment can be made by intensive use of the new methods of two-dimensional n.m.r. spectroscopy. The assignment can be backed up more by the determination of relaxation times and by observation of particular deshielding effects. If a solution to the spectrum has resulted, a difference n.O.e. (nuclear Overhauser effect) spectrum is set up and searched for interglycosidic n.O.e. effects between protons of various saccharide units. Provided that these interglycosidic n.O.e. effects can be observed, statements regarding the distance of the observed protons can be made, from which experimental indications of the conformation occurring in solution can be obtained. The results of the n.m.r. investigations must then be compared with the theoretical calculations.

Both methods described above were applied to the octasaccharide (47), which represents the key element in the oligosaccharide chain of the *N*-glycoprotein shown in Figure 4. A GESA calculation established that within the pentasaccharide section, in which the lactosamine residue is linked at O-3, one conformation is very much preferred.[31] In the lactosamine group, which is linked as a branch at O-6, there is, however, a mixture of *gt* and *gg* conformations. The energy calculation of the minima predicts that in this case the *gt* conformation should predominate; this preferred *gt* conformation is represented in Figure 5. The line drawing (bottom) shows the arrangement of the eight individual pyranose rings; the space-filling atomic model (top) represents the molecule in exactly the

[29] H. Thøgersen, R. U. Lemieux, K. Bock, and B. Meyer, *Can. J. Chem.*, 1980, **58**, 631.
[30] B. Meyer, to be published.
[31] H. Paulsen, T. Peters, V. Sinnwell, and B. Meyer, *Liebigs Ann. Chem.*, in press.

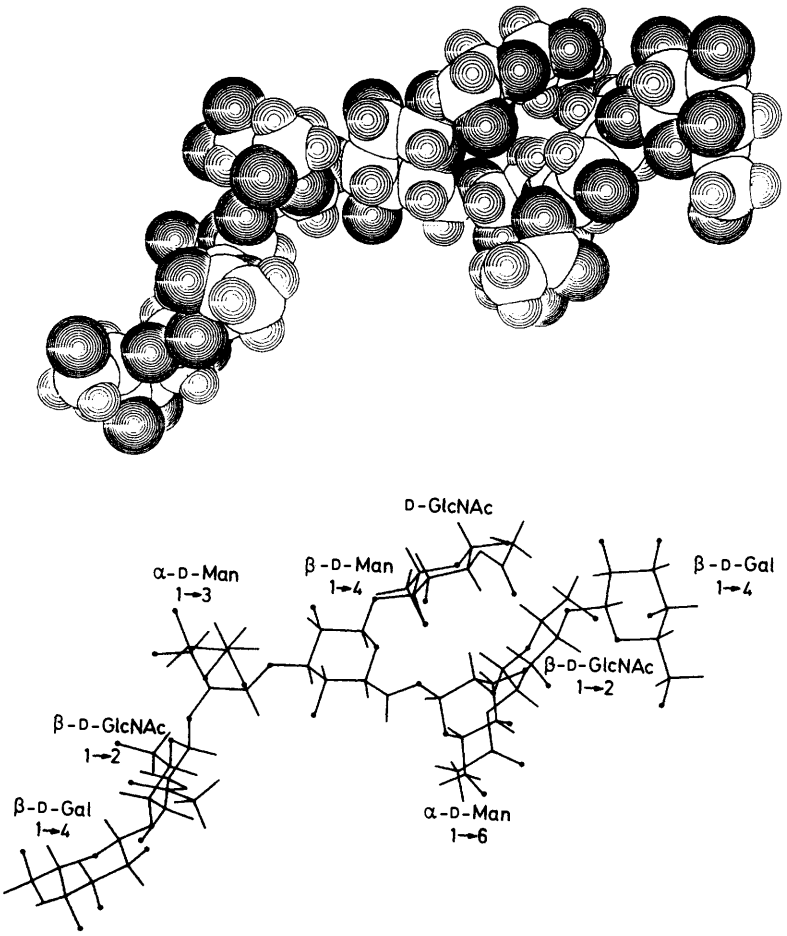


Figure 5 *Conformation of the octasaccharide* (47), *calculated from the GESA program. Bottom: line drawing showing the arrangement of the pyranose rings; top: space-filling molecular model with corresponding conformation*

same position and to the same scale as the line drawing.[31] In both models the two divergent lactosamine antennae can be clearly recognized.

N.m.r. studies of the octasaccharide (47) in D_2O are now possible for the first time, because the amounts of material necessary for such measurements are available by synthesis, leading to results that agree very well with theoretical calculations. The conformations that are preferred according to the calculations should also be preferred in solution, as the interglycosidic n.O.e. effects of the corresponding protons can be reconciled very well with the conformations calculated by

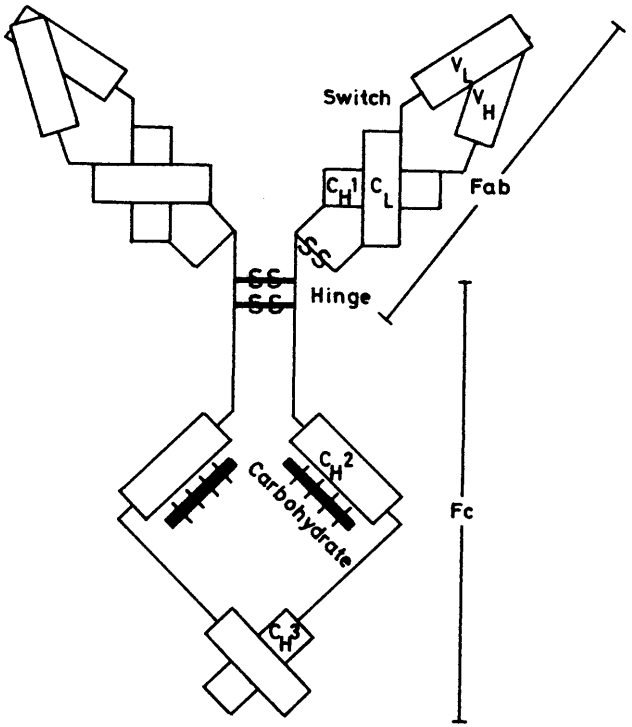


Figure 6 *Schematic diagram of the structure of immunoglobulin IgG1*

means of the GESA program.[31] The n.m.r. studies also indicate that for the lactosamine unit linked to O-6 there is a mixture of *gg* and *gt* conformations. Similarly, the data can best be interpreted in favour of the *gt* conformation.[31]

By very lucky chance, an *X*-ray structure analysis for the octasaccharide (47) is also available. The octasaccharide itself, of course, cannot be crystallized. Huber,[32] however, managed to crystallize immunoglobulin IgG1, isolated pure from human serum, and carried out a complete *X*-ray structure analysis. In addition to several peptide chains in the so-called Fc section, IgG1 contains a small carbohydrate chain. A schematic diagram of the structure of immunoglobulin IgG1 is shown in Figure 6, in which the carbohydrate chain at the C_H2 fragment can be seen. The structure of this carbohydrate chain actually matches to a great extent that of the synthesized octasaccharide (47). The *X*-ray structure analysis gave not only the positions of all the amino-acids but also the positions of the atoms of the carbohydrate chain. Thus there is now also available an *X*-ray structure analysis of the octasaccharide sequence (47). This curious situation arises

[32] R. Huber, *Klin. Wochenschr.*, 1980, **58**, 1217.

because the carbohydrate chain does not crystallize on its own but only in combination with a protein, the crystals thus permitting an *X*-ray structure analysis.

Examination of the conformation in the crystal shows that the octasaccharide is in the *gt* conformation, which from both GESA calculations and n.m.r. analysis in solution should be the conformation preferred in equilibrium. By comparing the arrangement of the individual saccharide residues within the crystal, there can also be seen to be relatively good agreement with the calculated data and with the conformation, as shown in Figure 5. It is quite remarkable that three independent methods for investigating the conformation of the carbohydrate chain of the octasaccharide (47) should lead to similar evidence. These findings are of considerable importance in assessing the conformation of other oligosaccharide chains and their interactions with proteins.