Pure & Appl. Chem., Vol. 69, No. 3, pp. 537–542, 1997.
Printed in Great Britain.
© 1997 IUPAC

Exploring the substrate specificity of sialyltransferases

Johannes A. L. M. van Dorst, Johannis P. Kamerling and Johannes F. G. Vliegenthart

Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, P.O. Box 80.075, NL-3508 TB Utrecht, Netherlands

Abstract: Twelve trisaccharide derivatives designed for detailed exploration of the acceptor specificity of sialyltransferases involved in the biosynthesis of N-glycans have been synthesized. These compounds include β-D-Gal*p*-(1-4)-β-D-Glc*p*NAc-(1-2)-α-D-Man*p*-(1-O)$(CH_2)_7CH_3$ and analogues containing structural variants of D-galactose. All trisaccharides were obtained by condensation of suitably modified glycosyl donors with a single disaccharide acceptor, thus limitating the number of reaction steps required. After deprotection, the compounds were employed to delineate the recognition characteristics of several natural and recombinant sialyltransferases.

INTRODUCTION

In biological systems, proteins and lipids bearing covalently linked carbohydrate chains are found in the cell, on cell membranes and in extra-cellular fluids. Interest in the carbohydrate structures of these glycoconjugates has rapidly increased in recent years, especially since their involvement in a variety of biological phenomena was recognized. For example, the oligosaccharides of cell-surface glycoconjugates have been implicated in cell differentiation, in cell-cell and receptor-ligand interactions, in viral and bacterial infectivity, as well as in tumor progression and metastasis (ref. 1).

Glycoconjugate glycans are synthesized in the cell by a concerted action of many distinct glycosyltransferases, catalysing the addition of monosaccharide units to an existing glycan, a polypeptide, or a lipid acceptor. The biosynthesis of a specific class of protein-bound oligosaccharides (N-glycans) is somewhat more complex, and requires the action of several additional enzymes (ref. 2). Glycosyltransferases are highly specific for their substrates *in vivo*, but a large number of studies have shown that *in vitro* these enzymes tolerate modifications in the acceptor substrate, indicating that some of the hydroxyl groups are of minor importance for effective glycosylation. A thorough analysis of the substrate specificity of glycosyltransferases can be attained by using specifically modified oligosaccharides, probing the contribution of individual hydroxyl groups in recognition and binding. In view of the complexity of carbohydrate organic synthesis, usually the smallest active acceptor substrate and derivatives thereof are synthesized (ref. 3). However, a number of glycosyltransferases have been shown to interact with parts of the acceptor structure remote from the site of glycosylation (ref. 4). This prompted us to undertake a program aimed at a detailed exploration of the recognition characteristics of these enzymes using modified acceptors beyond the minimal substrate, at the interface of a biologically relevant and a synthetically feasible approach.

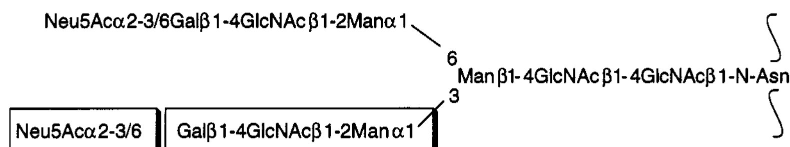
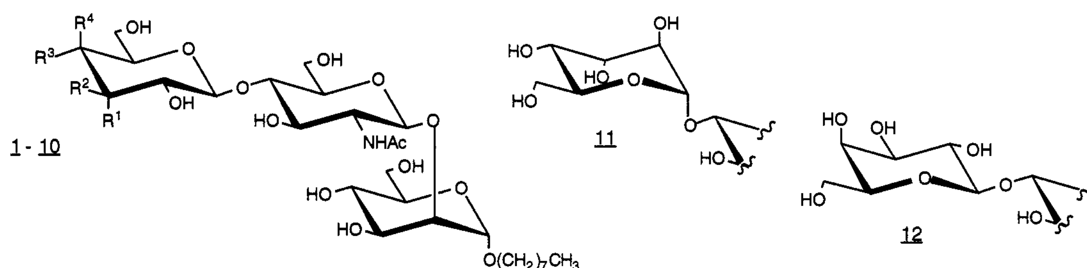


Fig. 1 A diantennary *N*-acetyllactosamine type of N-linked glycan, with Neu5Ac and a trisaccharide branch in shaded boxes.

In a first study we have focussed on sialyltransferases, catalyzing the transfer of sialic acids (a family of derivatives of the monosaccharide neuraminic acid with *N*-acetylneuraminic acid as a major member) to non-reducing termini of distinct glycoconjugate glycans. Sialylated structures are considered to be 'key' determinants in carbohydrate-based recognition (ref. 5); hence the study of the synthesizing enzymes is of special interest. Currently, we have studied the sialyltransferases involved in the α-2,3- or α-2,6-sialylation of terminal galactose units in N-glycoprotein glycans. The trisaccharide β-D-Gal*p*-(1-4)-β-D-Glc*p*NAc-(1-2)-α-D-Man*p*-(1-O)(CH_2)$_7$$CH_3$ and eleven analogues containing structural variants of D-galactose were synthesized (ref. 6, ref. 7) and employed as substrates (ref. 8). These trisaccharides represent a complete branch of a *N*-acetyllactosamine glycan (Fig. 1), 'mimicking' the requirements of the active sites of the enzymes as good as possible. Hydroxyl goups at either C-3 or C-4 of the galactose residue were substituted by hydrogen or fluorine, by amino- or *O*-methyl groups, or were inverted, to determine their involvement in binding and catalytic activity. In addition, trisaccharides containing α-L-altrose (an isomer of β-D-galactose with an inverted hydroxymethyl group at C-5) or β-L-galactose at the non-reducing terminus were constructed as probes.

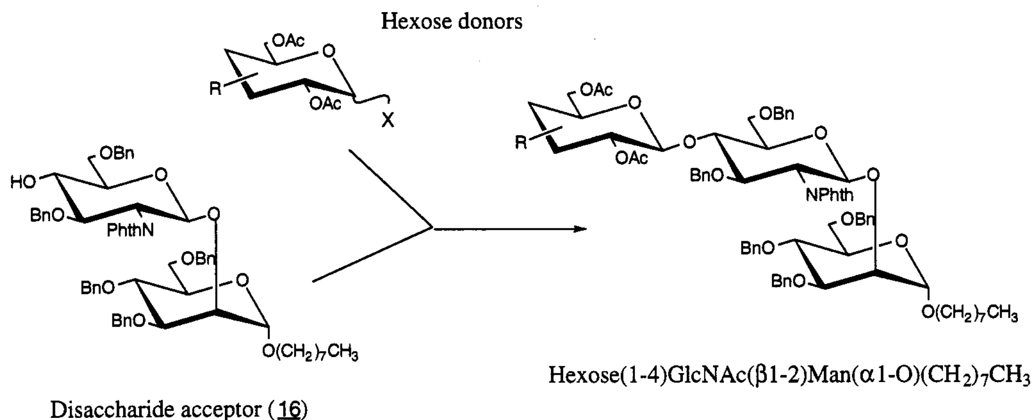
CHEMICAL SYNTHESIS OF OLIGOSACCHARIDES 1-12

Since the target structures 1-12 (depicted in Fig. 2) all contain the element -4)-β-D-Glc*p*NAc-(1-2)-α-D-Man*p*-(1-O)(CH_2)$_7$$CH_3$, a general synthetic strategy was aimed at the preparation of the disaccharide acceptor octyl 2-*O*-(3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-3,4,6-tri-*O*-benzyl-α-D-mannopyranoside (16), followed by elongation at HO-4' using suitably modified glycosyl donors (Scheme 1). This would allow the chemical modification, e.g. deoxygenation, *O*-methylation and fluorination, to be carried out on the level of galactose, to minimize the number of protecting groups that might interfere. In



Nr	R1	R2	R3	R4	Name
1	H	OH	H	OH	β-D-Gal-R
2	H	H	H	OH	3-deoxy-β-D-Gal-R
3	H	F	H	OH	3-deoxy-3-fluoro-β-D-Gal-R
4	OH	H	H	OH	β-D-Gul-R (3-epimer)
5	H	OMe	H	OH	3-*O*-methyl-β-D-Gal-R
6	H	NH_2	H	OH	3-amino-3-deoxy-β-D-Gal-R
7	H	OH	H	H	4-deoxy-β-D-Gal-R
8	H	OH	H	F	4-deoxy-4-fluoro-β-D-Gal-R
9	H	OH	OH	H	β-D-Glc-R (4-epimer)
10	H	OH	H	OMe	4-*O*-methyl-β-D-Gal-R
11	-	-	-	-	α-L-Alt-R (5-epimer)
12	-	-	-	-	β-L-Gal-R (enantiomer)

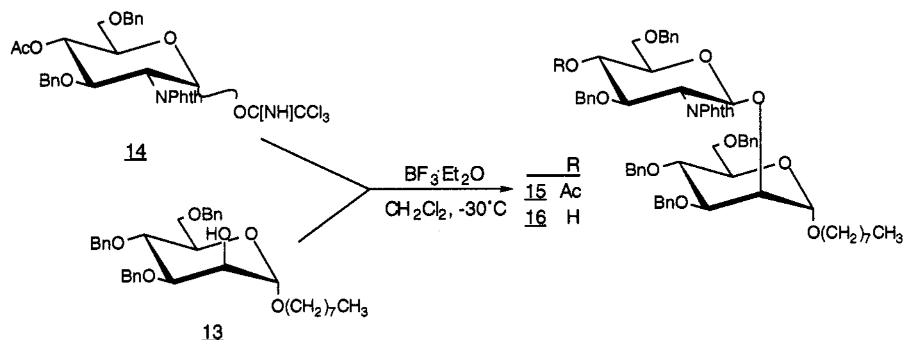
Fig. 2 Structures of synthetic trisaccharides intended for substrate specificity studies of sialyltransferases. For 11 and 12 only the variable part at the non-reducing end of the compounds is shown. The complete synthesis of these trisaccharides is described in ref. 6 and ref. 7. R, (1-4)-β-D-Glc*p* NAc-(1-2)-α-D-Man*p*-(1-O)(CH_2)$_7$$CH_3$.



Scheme 1 General synthetic strategy for the construction of the target trisaccharides 1-12.

addition, the overall number of reaction steps would be limited. Some of the modifications (particularly deoxygenation) turned out to give rise to an increase in reactivity at the anomeric center, necessitating the selection of a glycosyl donor of suitable reactivity.

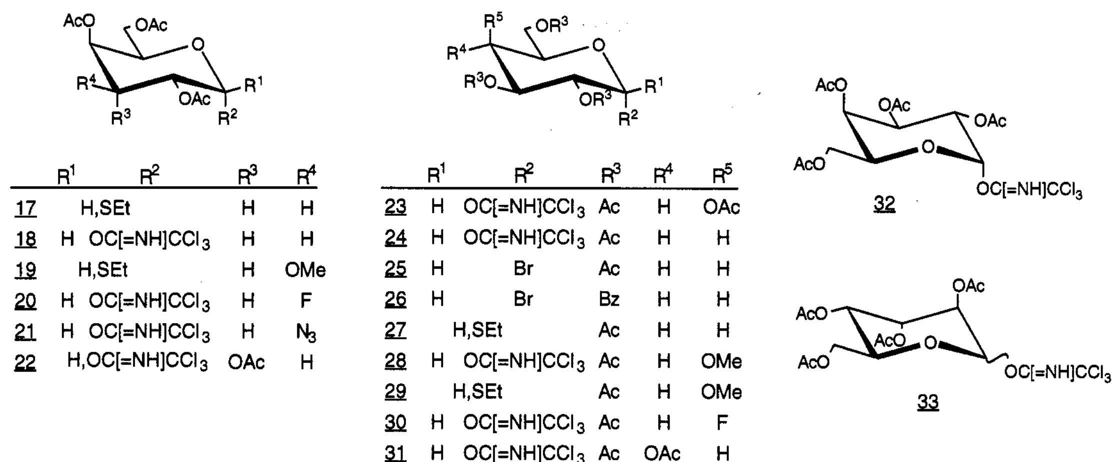
The general disaccharide acceptor 16 was obtained by condensation of octyl 3,4,6-tri-*O*-benzyl-α-D-mannopyranoside (13) (ref. 6) and 4-*O*-acetyl-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-α,β-D-glucopyranosyl trichloroacetimidate (14) (ref. 9), using boron trifluoride etherate as a catalyst (→15, 90%), followed by deacetylation (Scheme 2). Compound 16 was glycosylated, in CH_2Cl_2, with the donor reagents depicted in Scheme 3. The synthesis of 17-33 has been reported in detail elsewhere (ref. 6, ref. 7). Here, the various coupling results will be summarized.



Scheme 2 Synthesis of the key intermediate 16 as a general glycosyl acceptor for the construction of target structures 1-12.

The trichloroacetimidate group proved its general use in the stereoselective coupling of non-activated synthons, such as the unmodified hexoses (D-gulose 22, D-galactose 23, D-glucose 31, L-galactose 32, or L-altrose 33) and the fluorine-containing D-galactoses (20 and 30). In addition, the 3-azido-D-galactosyl imidate 21 was a convenient glycosyl donor. In all glycosylation reactions, trimethylsilyl triflate was used as a catalyst (0.05-0.3 eq) at 0°C, and single trisaccharide products were isolated, by column chromatography, in yields of 63-94%.

Several donors were examined for the preparation of a trisaccharide containing a 4-deoxy-D-galactose residue, since it is known that deoxyglycosyl donors are more reactive than the parent saccharides (ref. 10, ref. 11). The glycosylation behaviour of the bromides 25 and 26, examined towards the corresponding allyl glycoside analogue of 16, revealed that these donors produced mixtures of presumably trisaccharide derivatives



	R^1	R^2	R^3	R^4
17	H,SEt		H	H
18	H	OC[=NH]CCl_3	H	H
19	H,SEt		H	OMe
20	H	OC[=NH]CCl_3	H	F
21	H	OC[=NH]CCl_3	H	N_3
22	H,OC[=NH]CCl_3		OAc	H

	R^1	R^2	R^3	R^4	R^5
23	H	OC[=NH]CCl_3	Ac	H	OAc
24	H	OC[=NH]CCl_3	Ac	H	H
25	H	Br	Ac	H	H
26	H	Br	Bz	H	H
27	H,SEt		Ac	H	H
28	H	OC[=NH]CCl_3	Ac	H	OMe
29	H,SEt		Ac	H	OMe
30	H	OC[=NH]CCl_3	Ac	H	F
31	H	OC[=NH]CCl_3	Ac	OAc	H

Scheme 3 Glycosyl donors used in the glycosylation of disaccharide acceptor 16.

which could not be separated. Therefore, 25 and 26 were not further investigated towards 16. The glycosylation of 16 with imidate 24 at -30°C, using trimethylsilyl triflate as a catalyst, resulted in the formation of a single trisaccharide product (74%). In this reaction, the donor reagent was added dropwise to a solution of the acceptor and the catalyst, requiring a 2.5 fold excess of 24. Alternatively, the use of thioglycosides 27 (1.3 eq) at 0°C, with *N*-iodosuccinimide/triflic acid as a promoter, gave the same trisaccharide in higher yield (81%) without the necessity of a reversed addition of reagents.

Introductory experiments with 4-*O*-methyl-D-galactosyl imidate 28 indicated that the presence of the methyl group enhanced the reactivity at the anomeric center in analogy with the 4-deoxy-D-galactosyl donor. Coupling of 28 with the allyl glycoside analogue of 16 led to loss of stereoselectivity. However, the use of thioglycoside 29 effected the formation of a (β1-4)-linkage only upon glycosylation with 16 (85%). Making use of these observations, the 3-deoxy- and the 3-*O*-methyl-D-galactose residues were introduced by use of the corresponding thioglycosides 17 (81%) and 19 (55%).

From the various glycosylation reactions it was concluded that thioglycosides were most appropriate for stereoselective coupling of activated synthons (carrying deoxy- or *O*-methyl-functionalities), whereas trichloroacetimidates gave high yields with deactivated (e.g. fluorine-containing) synthons. All trisaccharides were deprotected in good overall yields by conventional methods, including dephthaloylation ($H_2NNH_2 \cdot H_2O$ or $H_2N(CH_2)_2NH_2$) and re-*N*,*O*-acetylation, catalytic hydrogenation (H_2, Pd/C), and de-*O*-acetylation (NaOMe, MeOH).

EMPLOYMENT OF TRISACCHARIDES 1-12 AS SYNTHETIC PROBES

The acceptor specificites of rat liver Gal(β1-4)GlcNAc α-2,6-sialyltransferase (RL α-2,6-ST), recombinant full length human liver Gal(β1-4)GlcNAc α-2,6-sialyltransferase (HL α-2,6-rST), and a soluble form of recombinant rat liver Gal(β1-3/4)GlcNAc α-2,3-sialyltransferase (RL α-2,3-rST) were studied using the trisaccharides 1-12 as synthetic probes, as presented in Table 1 (ref. 8).

All substrate analogues, except 12, were accepted to some extent by both α-2,6-sialyltransferases, although a substantial decrease in activity was observed for each acceptor. Previously it has been shown, using a number of modified *N*-acetyllactosamine-type disaccharides, that Gal HO-6 together with the amide group of the penultimate GlcNAc residue are the only groups indispensible for an *effective* transfer of Neu5Ac, provided that they can be properly positioned in the active site (ref. 12). Nevertheless, the enzyme is known to exhibit a high *in vivo* specificity and does not accept Gal(β1-3)GlcNAc nor Gal(β1-6)GlcNAc as a substrate (ref. 13). Therefore, the active site of α-2,6-sialyltransferase must be closely contouring the substrate

TABLE 1. Acceptor specificity of sialyltransferases as determined with trisaccharides 1-12. Relative rates for each acceptor are expressed as a percentage of the incorporation of Neu5Ac into the parent compound 1. R, (1-4)-β-D-Glc*p*NAc-(1-2)-α-D-Man*p*-(1-O)$(CH_2)_7CH_3$.

Nr.	Acceptor	Relative rate of sialylation (%) [a]		
		RL α-2,6-ST	RL α-2,6-rST	RL α-2,3-rST
1	β-D-Gal-R	100	100	100
2	3-deoxy-β-D-Gal-R	7	13	0
3	3-deoxy-3-fluoro-β-D-Gal-R	2	8	0
4	β-D-Gul-R (3-epimer)	1	<1	0
5	3-*O*-methyl-β-D-Gal-R	1	<1	0
6	3-amino-3-deoxy-β-D-Gal-R	2	<1	0
7	4-deoxy-β-D-Gal-R	40	29	2
8	4-deoxy-4-fluoro-β-D-Gal-R	23	17	7
9	β-D-Glc-R (4-epimer)	6	5	0
10	4-*O*-methyl-β-D-Gal-R	2	2	51
11	α-L-Alt-R (5-epimer)	<1	<1	0
12	β-L-Gal-R (enantiomer)	0	0	0

[a] The substrate concentration was 2 mM for all compounds tested.

for an *efficient* transfer of Neu5Ac. Consistent with this, the sterically conservative hydroxyl group substitutions with hydrogen or fluorine were tolerated much better than epimerization or methylation of these hydroxyl functions.

Interestingly, the α-2,6-sialyltransferases were able to accomplish the transfer of Neu5Ac to oligosaccharides with terminal residues lacking a β-D-galacto configuration. The sialylation of 11, containing α-L-altrose instead of β-D-galactose, suggests that this compound exhibits to some extent a proper orientation of Alt HO-6 and the amide group of GlcNAc in the active site. Since 12 was completely ineffective as a substrate, the L-Gal(β1-4)GlcNAc structure in compound 12 apparently lacks the required topography. In order to confirm this assumption, energy maps of the disaccharide derivatives D-Gal(β1-4)GlcNAc(β1-O)Me, L-Alt(α1-4)GlcNAc(β1-O)Me, and L-Gal(β1-4)GlcNAc(β1-O)Me, comprising the variable parts of 1, 11, and 12, respectively, were constructed. The minimum energy conformation for each disaccharide was calculated from its respective contour plot and is depicted in Fig. 3 (ref. 8). It appeared that the hydroxymethyl groups of D-Gal and L-Alt are located in a similar orientation with respect to the NHAc moiety of the GlcNAc residue, namely at one side of the disaccharide. In contrast, these groups are positioned at opposite sides in the L-Gal(β1-4)GlcNAc element.

The α-2,3-sialyltransferase acting on N-glycans is capable of using both Gal(β1-3)GlcNAc and Gal(β1-4)GlcNAc structures as a substrate, suggesting a relatively loose specificity as compared to the α-2,6-sialyltransferase. Using deoxygenated substrate analogues (ref. 12), the enzyme has been shown to require an intact 3,4,6-triol system on the accepting galactose residue, together with some 'unidentified structural features' (ref. 12). The present data are supportive of the requirement for an intact D-*galacto* 3,4-diol system for catalytic action (Table 1), but for binding this topography is not essential. The substrates with an inverted hydroxyl function at either C-3 (4) or C-4 (9) were not sialylated, but nevertheless had inhibitory activities (data not shown). Surprisingly, the 3-deoxy-D-Gal-containing *tri*saccharide 2 proved to be inhibitory towards α-2,3-sialyltransferase, whereas deoxygenation at HO-3 of galactose of the related *di*saccharide D-Gal(β1-4)GlcNAc(β1-O)$(CH_2)_8COOCH_3$ has been found to result in a complete loss of recognition and binding by the enzyme (ref. 14). The mannosyl residue in the trisaccharides studied here may therefore constitute at least some of the above mentioned 'additional structural features' required by α-2,3-sialyltransferase.

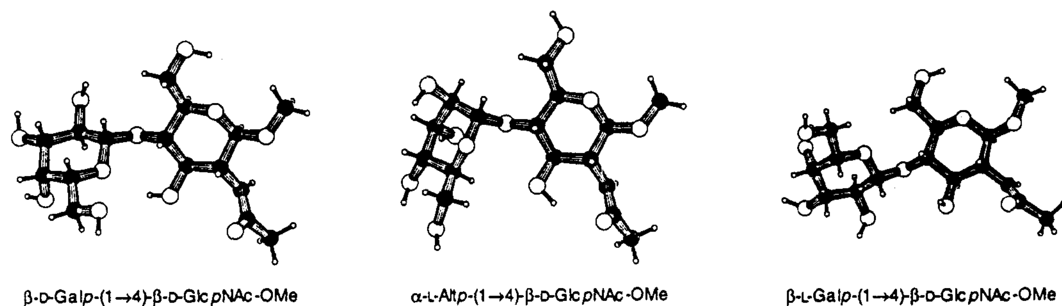


Fig. 3 Minimum energy conformations of D-Gal(β1-4)GlcNAc(β1-O)Me, L-Alt(α1-4)GlcNAc(β1-O)Me, and L-Gal(β1-4)GlcNAc(β1-O)Me.

REFERENCES

1. A. Varki. *Glycobiology* **3**, 97-130 (1993).
2. R. Kornfeld and S. Kornfeld. *Ann. Rev. Biochem.* **54**, 631-664 (1985).
3. see, for example: G. Möller, F. Reck, H. Paulsen, K. J. Kaur, M. Sarkar, H. Schachter and I. Brockhausen. *Glycoconjugate J.* **9**, 180-190 (1992). F. Reck, E. Meinjohanns, M. Springer, R. Wilkens, J. A. L. M. van Dorst, H. Paulsen, G. Möller, I. Brockhausen and H. Schachter. *Glycoconjugate J.* **11**, 210-216 (1994). I. Lindh and O. Hindsgaul. *J. Am. Chem. Soc.* **113**, 216-223 (1991). C. H. Wong, Y. Ichikawa, T. Krach, C. Gautheron-Le Narvor, D. P. Dumas and G. C. Look. *J. Am. Chem. Soc.* **113**, 8137-8145 (1991).
4. see, for example: W. M. Blanken, A. van Vliet and D. H. van den Eijnden. *J. Biol. Chem.* **259**, 15131-15135 (1984). S. C. Ats, J. Lehmann and S. Petry. *Carbohydr. Res.* **252**, 325-332 (1994). D. H. Joziasse, W. E. C. M. Schiphorst, D. H. van den Eijnden, J. A. van Kuik, H. van Halbeek and J. F. G. Vliegenthart. *J. Biol. Chem.* **260**, 714-719 (1985). M. J. Elices and I. J. Goldstein. *J. Biol. Chem.* **264**, 1375-1380 (1989).
5. R. Schauer and J. P. Kamerling. In *Glycoproteins* (J. Montreuil, J. F. G. Vliegenthart and H. Schachter, eds.), New Comprehensive Biochemistry, Vol. 29b. Elsevier, Amsterdam (1996, in press).
6. J. A. L. M. van Dorst, C. J. van Heusden, A. F. Voskamp, J. P. Kamerling and J. F. G. Vliegenthart. *Carbohydr. Res.* **291**, 63-83 (1996).
7. J. A. L. M. van Dorst, C. J. van Heusden, J. M. Tikkanen, J. P. Kamerling and J. F. G. Vliegenthart. *Carbohydr. Res.*, in press (1996).
8. J. A. L. M. van Dorst, J. M. Tikkanen, C. H. Krezdorn, M. B. Streiff, E. G. Berger, J. A. van Kuik, J. P. Kamerling and J. F. G. Vliegenthart. *Eur. J. Biochem.*, submitted for publication (1996).
9. F. Yamazaki, T. Kitajima, T. Nukada, Y. Ito and T. Ogawa. *Carbohydr. Res.* **201**, 15-30 (1990).
10. Z. Zhiyuan and G. Magnusson. *Carbohydr. Res.* **262**, 79-101 (1994).
11. H. Paulsen, V. Rutz and I. Brockhausen. *Liebigs Ann. Chem.*, 747-758 (1992).
12. K. B. Wlasichuk, M. A. Kashem, P. V. Nikrad, P. Bird, C. Jiang and A. P. Venot. *J. Biol. Chem.* **268**, 13971-13977 (1993).
13. J. Weinstein, U. De Souza-e-Silva and J. C. Paulson. *J. Biol. Chem.* **252**, 2363-2371 (1977).
14. O. Hindsgaul, K. J. Kaur, G. Srivastava, M. Blaszcyk-Thurin, S. C. Crawley, L. D. Heerze and M. M. Palcic. *J. Biol. Chem.* **266**, 17858-17862 (1991).