Control of Mucin Synthesis: The Peptide Portion of Synthetic *O*-Glycopeptide Substrates Influences the Activity of *O*-Glycan Core 1 UDPgalactose:*N*-Acetyl-α-galactosaminyl-R β3-Galactosyltransferase†

Inka Brockhausen,*,‡ Gabriele Möller,‡ Gunnar Merz,§ Knut Adermann,§ and Hans Paulsen§

Research Institute, The Hospital for Sick Children, 555 University Avenue, Toronto, M5G 1X8 Ontario, Canada, and Institute for Organic Chemistry, University of Hamburg, Martin-Luther-King-Platz 6, 2 Hamburg 13, Germany

Received May 2, 1990; Revised Manuscript Received July 27, 1990

ABSTRACT: Synthetic *O*-glycopeptides containing one or two GalNAc residues attached to Ser or Thr were used as substrates to investigate the effect of peptide structure on the activity of crude preparations of UDP-Gal:GalNAcα-R β3-Gal-transferase from pig stomach and pig and rat colonic mucosa and of a partially purified enzyme preparation from rat liver. High-performance liquid chromatography used to separate enzyme products revealed that uncharged glycopeptides with an acetyl group at the amino-terminal end and a tertiary butyl or an amide group at the carboxy-terminal end were resistant to proteolysis in crude preparations. The activity of β3-Gal-transferase varied with the sequence and length of the peptide portion of the substrate, the presence of protecting groups, the attachment site of GalNAc, and the number of GalNAc residues in the substrate. The presence and position of Pro had little effect on enzyme activity; ionizing groups near the GalNAc unit interfered with enzyme activity. Since the GalNAc-Thr moieties in many of these *O*-glycopeptides have been shown to assume similar rigid conformations, the variation in enzyme activity indicates that the β3-Gal-transferase recognizes both the peptide and carbohydrate moieties of the substrate. Rat and pig colonic mucosal homogenates contain β3- and β6-GlcNAc-transferases that synthesize respectively *O*-glycan core 3 (GlcNAcβ3GalNAcα-R) and core 4 [GlcNAcβ6(GlcNAcβ3)GalNAcα-R]. These enzymes also showed variations in activity with different peptide structures; these effects did not parallel those observed with β3-Gal-transferase. The results show that the peptide portion in glycoproteins affects the activities of three glycosyltransferases which act on GalNAcα-Ser(Thr)-X. It is predicted that individual O-glycosylation sites carry characteristic *O*-glycan core structures.

The biosynthesis of *O*-glycans requires the sequential transfer of individual sugar residues to the glycoprotein. The glycosyltransferases which carry out these reactions are controlled by their genes, by cofactors, and by the structure and availability of substrates. Some *O*-glycoproteins, such as mucins, are extremely heterogeneous in their oligosaccharide structures; not all the Ser and Thr residues in most *O*-glycoproteins are glycosylated. The number of possible *O*-glycan structures and their positions in the peptide chain are determined in part by the specificity of glycosyltransferases for their substrates. One or more Pro residues are usually found near the glycosylation sites. This may cause increased peptide accessibility for the enzyme that attaches the first GalNAc[1] residue, i.e., UDP-GalNAc:polypeptide α-GalNAc-transferase. As shown by Young et al. (1981) and others, this enzyme is in fact influenced by the length, sequence, and composition of the peptide substrate and especially by Pro residues.

It has been shown for many *N*-glycoproteins that the *N*-glycans at a particular Asn position often differ in a characteristic manner from *N*-glycans at other positions on the same protein. Carver and Cummings (1987) have proposed a mechanism for the role of the peptide in the processing of N-linked carbohydrate at individual glycosylation sites. They suggest that interactions between carbohydrate and peptide may result in different oligosaccharide conformations at different Asn sites. The oligosaccharide conformation may in turn determine which of several competing glycosyltransferases acts at that site. Thus the specificity of glycosyltransferases for certain oligosaccharide conformations may lead to different processing at different glycosylation sites.

Although *O*-glycoproteins usually carry a heterogeneous mixture of oligosaccharides, it is not yet known if *O*-glycan structures differ at individual glycosylation sites or how the peptide portion influences the activities of enzymes that act on GalNAc-Ser(Thr)-X. To study the effect of the peptide moiety on the synthesis of mucin core structures, we have synthesized a series of *O*-glycopeptides containing one or two GalNAc residues attached to peptides of various lengths and compositions (Paulsen et al., 1988; Paulsen & Adermann, 1989). We have used these synthetic *O*-glycopeptides as substrates for a partially purified *O*-glycan core 1 β3-Gal-transferase from rat liver and for core 1 Gal-transferase from pig gastric mucosa and pig and rat colonic mucosal homogenates, as well as core 3 and core 4 GlcNAc-transferases from pig and rat colonic homogenates (Figure 1). Since the crude preparations contain pyrophosphatase and protease activities which destroy the substrates for these enzymes, high-performance liquid chromatography (HPLC) assays were used to separate enzyme products from other assay components. The results suggest that all three glycosyltransferases recognize

† This research was supported by a Canadian Cystic Fibrosis grant to Harry Schachter, Research Institute, The Hospital for Sick Children, Toronto, the Deutsche Forschungsgemeinschaft (to H.P.), and a NATO grant for international collaboration in research (to I.B. and H.P.).

* Address correspondence to this author at the Department of Biochemistry, The Hospital for Sick Children.

‡ The Hospital for Sick Children.

§ University of Hamburg.

[1] Abbreviations: Bn, benzyl; HPLC, high-pressure liquid chromatography; G or Gal, galactose; GA or GalNAc, *N*-acetylgalactosamine; Gn or GlcNAc, *N*-acetylglucosamine; NMR, nuclear magnetic resonance; MES, 2-(*N*-morpholino)ethanesulfonate; PMSF, phenylmethanesulfonyl fluoride; tBu, tertiary butyl ester.

0006-2960/90/0429-10206$02.50/0

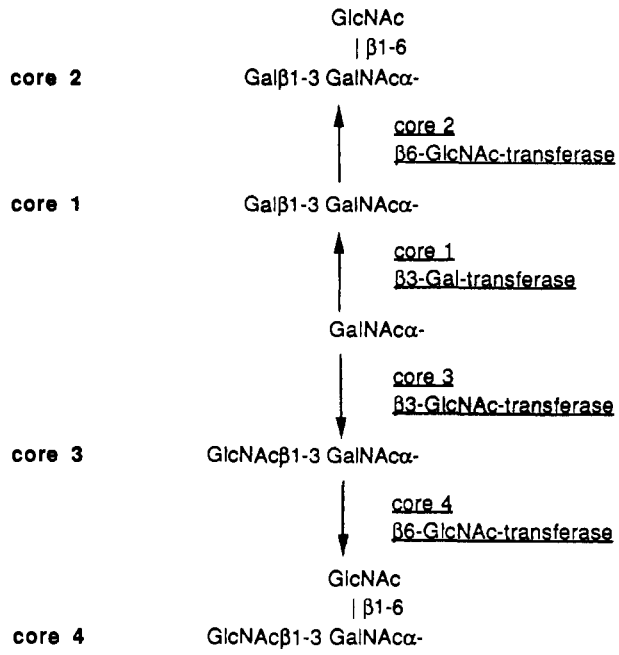

FIGURE 1: Structures and biosynthesis of the four common mucin cores.

the peptide environment near the glycosylated amino acid. We conclude that the peptide structure and the presence of nearby GalNAc residues play a major role in determining the type of *O*-glycan core synthesized at each glycosylation site.

Experimental Procedures

Materials. AG1x8 (100–200 mesh, Cl^- form) was purchased from Bio-Rad. AMP, bovine serum albumin, GalNAcα-benzyl, Triton X-100, and protease inhibitors were purchased from Sigma. UDP-*N*-[1-^{14}C]acetylglucosamine was synthesized as described previously (Brockhausen et al., 1985). UDP-[U-^{14}C]galactose was purchased from Amersham and diluted with UDP-galactose (from Sigma) to yield specific activities of 1088–2220 dpm/nmol. Glycopeptides were synthesized as described previously (Paulsen et al., 1988; Paulsen & Adermann, 1989).

Protein Determination. Protein was determined by the Bio-Rad method with bovine serum albumin as the standard.

Nuclear Magnetic Resonance. Samples were prepared by exchanging two times with 99.8% D_2O (Aldrich) and two times with 99.96% D_2O (Merck Sharpe & Dohme). Samples were dissolved in 99.96% D_2O with acetone as the internal standard (set at 2.225 ppm). Spectra were recorded at the Toronto Carbohydrate Research Centre with a Bruker 500-MHz spectrometer.

Amino Sugar Analysis. The GalNAc content of glycopeptides was determined by amino sugar analysis. Samples containing α-amino-β-guanidinopropionic acid or ε-aminocaproic acid as the internal standard were hydrolyzed in 6 N HCl at 96–100 °C for 1.5 h. Galactosamine analysis of hydrolyzed samples was carried out by the amino acid analysis center in the laboratory of Dr. Max Blum, University of Toronto, with a Beckman Model 121-M amino acid analyzer with an extended short column, the ninhydrin method being used for detection.

Enzyme Preparations. Rat and pig colonic mucosal homogenates were prepared from colonic mucosal scrapings of freshly killed rats and pigs. Mucosa were first cleaned with a glass slide. Scrapings were homogenized with a Potter-Elvehjem homogenizer in 0.25 M sucrose and stored at −70 °C. Pig gastric mucosal microsomes were prepared as described before (Brockhausen et al., 1986). UDP-Gal:GalNAcα-R β3-galactosyltransferase was partially purified from rat liver (I. Brockhausen et al., in preparation). This preparation contained 0.24 mg of protein/mL and was stable for several months at 4 °C in 5 mM $MnCl_2$, 25 mM MES, 20% glycerol, 0.1% Triton X-100, and 0.02% NaN_3. The enzyme contained GlcNAc-transferase II (0.113 μmol min^{-1} mg^{-1}; Bendiak & Schachter, 1987) but was free of pyrophosphatases, proteases, and UDP-Gal:GlcNAc β4-Gal-transferase.

Gal-Transferase Assays. The incubation mixtures for assays with partially purified rat liver Gal-transferase contained the following in a total volume of 25 μl: 1.55 mM UDP-[U-^{14}C]Gal (1088–1225 dpm/nmol), 0.16 M MES, 20 mM $MnCl_2$, 0.2% Triton X-100, 0.08% bovine serum albumin, acceptor substrate (specified under Results), and 5 μL of enzyme (1.2 μg of protein). Mixtures were incubated for 1 h at 37 °C. Reactions were stopped by addition of 0.4 mL of 1 mM EDTA–20 mM borate. Mixtures were passed through Pasteur pipets filled with AG1x8 (100–200 mesh, Cl^- form) and washed with 2.6 mL of water. The eluates were dissolved in 16 mL of scintillation fluid (Readysolve from Beckman) and counted in a liquid scintillation counter. Some of the eluates from the AG1x8 columns were analyzed by HPLC; the eluates were lyophilized and taken up in 200 μL of water, and 100 μL was analyzed by HPLC (see below).

The incubation mixtures for assays with crude enzyme preparations contained the following in a total volume of 40 μL: 0.5–1 mM UDP-[U-^{14}C]Gal (2000–2220 dpm/nmol), 0.125 M MES, 12.5 mM $MnCl_2$, 0.125–0.25% Triton X-100, 0.5 mM PMSF, 2.5 mM AMP (when the rat colon enzyme was used), acceptor substrate (specified under Results), and enzyme (118 μg of protein of rat colonic mucosal homogenate, 135 μg of protein of pig gastric mucosal microsomes, or 220 μg of protein of pig colonic mucosal homogenate). Mixtures were incubated for 2 h at 37 °C. The reactions were stopped with 0.4 mL of 1 mM EDTA–20 mM borate. Mixtures were passed through Pasteur pipets filled with AG1x8 (100–200 mesh, Cl^- form) and washed with 2.6 mL of water. The eluates were lyophilized and taken up in 200 μL of water, and 100 μL was analyzed by HPLC (see below). Radioactive enzyme products were collected and counted. Activity was calculated by subtraction of values obtained from incubations without added substrate.

GlcNAc-transferase Assays. The incubation mixtures for assays with crude enzyme contained the following in a total volume of 40 μL: 1 mM UDP-[1-^{14}C]GlcNAc (4434–5633 dpm/nmol), 0.125 M MES, 0.125 M GlcNAc, 12.5 mM $MnCl_2$, 0.125–0.25% Triton X-100, 0.5 mM PMSF, 2.5 mM AMP (when the rat colon enzyme was used), acceptor substrate (specified under Results), and 20 μL of homogenate (118 μg of protein of rat colonic mucosal homogenate or 220 μg of protein of pig colonic mucosal homogenate). Mixtures were incubated for 2 h at 37 °C. Reactions were stopped by addition of 0.4 mL of 1 mM EDTA–20 mM borate. Mixtures were passed through Pasteur pipets filled with AG1x8 (100–200 mesh, Cl^- form) and washed with 2.6 mL of water. The eluates were lyophilized and taken up in 200 μL of water, and 100 μL was analyzed by HPLC (see below). Activity was calculated by subtraction of values obtained from incubations without added substrate.

Protease Inhibition. Attempts were made in some assays to augment inhibition of proteases by replacing 0.5 mM PMSF in the standard assays with a cocktail of 1.5 mM PMSF, 1.5 mM pepstatin, 1.5 mM antipain, 1.5 mM leupeptin, 1.5 mM

Table I: HPLC Separation of *O*-Glycopeptides on a C18 Column Using Water–Acetonitrile Mixtures as the Mobile Phase

no.	compd[a]	elution times (min)[b] substrate	Gal-T product[c]	Gn-T product[c]	% acetonitrile
4	Ac(GA)T-tBu	24	22	13, 18	10
10	Ac(GA)T-P-tBu	22	16	10, 16	17
11	Ac(GA)T-P-NH_2	27	23	17, 20	2
12	P-(GA)T-V	31			3
13	V-(GA)T-P	24			4
14	AcV-(GA)T-P	38			1
15	AcV-(GA)T-P-tBu	34	21	14, 20	20
16	AcV-(GA)T-P-NH_2	28	22	15, 20	6
17	P-(GA)T-G	9			0
18	A-(GA)T-P	23			0
19	P-(GA)T-P	31			0
20	A-P-(GA)T	12			0
21	A-P-(GA)T-S	26	21	16, 36, 40	0
22	A-P-(GA)T-S-S	18	20, 24	13, 18	0
23	A-P-(GA)T-S-S-S	20	22	16, 28, 40	0
24	A-P-(GA)T-S-S-A	28	25	18, 38	0
25	A-P-(GA)T-S-A-S	40	40	38	0
26	A-P-(GA)T-A-S-S	40	30	38	0
27	A-P-T-(GA)S-S-S	22	18	10, 32	0
28	A-P-(GA)T-S-S-S-T-K	21	22	14	4
30	(GA)T-(GA)T-G-V-A	24	18,[d] 22, 34	26, 36, 48	1
31	A-P-(GA)T-(GA)S-S-S	26	26, 42	16, 27, 42	0
32	A-(GA)T-V-(GA)T-A-G	46			3
33	GlcNAc	3			0
34	GalNAcα-Bn	35	28	18, 27	10

[a] GA, GalNAcα-; tBu, tertiary butyl; Ac, acetyl; NH_2, amide; Bn, benzyl. Amino acids are named by the one-letter code. [b] Compounds were eluted from a C18 column with acetonitrile–water mixtures at a flow rate of 1 mL/min and pressures of 60–80 bar. [c] Products were prepared by incubation with rat colonic mucosal homogenate. Rat colon contains core 1 Gal-transferase and core 3 and core 4 GlcNAc-transferases (Figure 1); two products are therefore expected after incubation of rat colon homogenates with UDP-GlcNAc and GalNAc-R substrates. [d] Elution time of enzyme product with the purified rat liver Gal-transferase.

bestatin, and 0.15% aprotinin.

High-Pressure Liquid Chromatography. HPLC separations were carried out with an LKB system (2152 controller and 2150 pumps). The elution of compounds was monitored at 195 nm with a 2152 variable-wavelength UV monitor. A reverse-phase silica-based C18 column (10 μm, 4.6 × 250 mm, Alltech Applied Science) protected with a C18 guard column was used at a flow rate of 1 mL/min with acetonitrile–water mixtures as the mobile phase and 60–80-bar pressure. Enzyme assays were carried out by recording the absorbance at 195 nm, collecting 2-mL fractions, and counting these in 12 mL of scintillation fluid.

Results

HPLC Fractionation of Synthetic O-Glycopeptides. Table I shows the reverse-phase HPLC separations on a C18 column of the synthetic *O*-glycopeptides used in this work and of the products formed from some of these glycopeptides by the action of core 1 Gal-transferase and core 3 and core 4 GlcNAc-transferases (Figure 1) with rat colonic mucosal homogenate as enzyme source. HPLC conditions were chosen so as to affect the fractionation of radioactive monosaccharides, enzyme products, and other assay components in less than 1 h. Rat colon converts GalNAcα-R to Galβ1–3GalNAcα-R (core 1) and to GlcNAcβ1–3GalNAcα-R (core 3) and converts GlcNAcβ1–3GalNAcα-R to GlcNAcβ1–6(GlcNAcβ1–3)GalNAcα-R (core 4) (Brockhausen et al., 1985). Two elution times are shown for the GlcNAc-transferase products for a particular substrate (Table I); these refer to GlcNAcβ1–3GalNAcα-R and GlcNAcβ1–6(GlcNAcβ1–3)-GalNAcα-R, but the specific retention times of each glycopeptide product have not been determined.

Free Gal or GlcNAc is not retained by the C18 column and elutes after 4 min. Small glycopeptides with one or two amino acids which are either unprotected zwitterions (compounds **1**, **2**, **5**, **7**, **8**, and **9**, Table I) or have a net charge (compounds **3** and **6**, Table I) are not retained by the column (elution in 3–6 min) due to the lack of hydrophobic interaction with the C18 groups; these compounds could not be used for enzyme assays by the HPLC method. When they were used as acceptors for the purified rat liver enzyme, product was counted directly after the AG1x8 step.

The remaining glycopeptides (Table I) are more hydrophobic and are eluted by increasing acetonitrile concentrations in the mobile phase. Elution times depend on the amino acid compositions and sequences of the glycopeptides. Changes in the amino acid sequence (compare compounds **24** and **26** or compounds **18** and **20**, Table I) or in the position of GalNAc attachment (compounds **23** and **27**) can have a significant effect on the interaction with the C18 column, indicating that glycopeptides of the same composition but different sequence may have different hydrophobic surfaces exposed for interaction with the C18 groups. The addition of hydrophilic carbohydrate residues in the formation of enzyme products usually decreases retention times (Table I) but may have no effect or may even increase retention time in some cases.

To ensure that glycopeptide products with a negative charge are not retarded by the AG1x8 anion exchange column used in the standard enzyme assays, borate–EDTA was added to compound **21** [Ala-Pro-(GA)Thr-Ser] followed by passage through AG1x8 under the conditions used for transferase assays, followed by HPLC separation. No retention of glycopeptide on the anion exchange column was detected by measurement of absorbance at 195 nm.

Substrate Specificity of Partially Purified Rat Liver β3-Gal-transferase. Partially purified rat liver β3-Gal-transferase exhibited an activity of 7.1 nmol min^{-1} mg^{-1} toward 2 mM GalNAcα-benzyl (Table II) with a K_M of 1.27 mM and a V_{max} of 12.2 nmol min^{-1} mg^{-1}. The enzyme preparation was free of pyrophosphatase activity since HPLC analysis of the assay mixture showed only a very small amount of free radioactive galactose (<60 dpm/assay) due to the breakdown of UDP-[^{14}C]Gal. HPLC assays carried out with several unprotected glycopeptides (compounds **13**, and **23**–**27**) showed single sharp peaks for both substrates and products, indicating the lack of degrading proteases. The enzyme preparation was free of UDP-Gal:GlcNAc-R β3- and β4-Gal-transferases since it was inactive toward GlcNAc and GlcNAcβ3GalNAcα-benzyl as acceptors.

Unprotected glycopeptide substrates containing one to four amino acid residues were much less effective than GalNAcα-benzyl (Table II). Protection of only the amino-terminal end of GalNAc-Thr, Ala-(GalNAc)Thr, or Val-(GalNAc)Thr-Pro did not improve the effectiveness of these compounds. Protection of both the amino- and carboxy-terminal ends of (GalNAc)Thr-Pro and Val-(GalNAc)Thr-Pro

Table II: β3-Gal-transferase Activities toward GalNAc-glycopeptide Substrates Using the Purified Rat Liver Enzyme[a]

acceptor	dpm of product (1088 dpm/nmol)	enzyme activities (nmol min^{-1} mg^{-1})	apparent K_M (mM)	apparent V_{max} (nmol min^{-1} mg^{-1})	V_{max}/K_M
34, 2 mM GA-Bn	555	7.1	1.27	12.2	9.61
1, 2 mM (GA)S	233	3.0	nd	nd	
2, 2.4 mM (GA)T	61	0.8	nd	nd	
3, 2.9 mM Ac(GA)T	26	0.3	nd	nd	
4, 2.5 mM Ac(GA)T-tBu	62	0.8	nd	nd	
5, 2.1 mM A-(GA)T	31	0.4	nd	nd	
6, 0.76 mM AcA-(GA)T	118	1.5	nd	nd	
7, 1.9 mM (GA)T-A	112	1.4	nd	nd	
8, 2.3 mM (GA)S-A	357	4.5	7.7	17.7	2.31
9, 1.9 mM P-(GA)T	23	0.3	nd	nd	
10, 1.4 mM Ac(GA)T-P-tBu	952	12.2	2.7	34.2	12.7
11, 1.6 mM Ac(GA)T-P-NH_2	523	6.7	3.57	19.1	5.36
12, 2.3 mM P-(GA)T-V	36	0.3	nd	nd	
13, 2.9 mM V-(GA)T-P	18	0.2	nd	nd	
14, 2.5 mM AcV-(GA)T-P	123	1.1	nd	nd	
15, 1.4 mM AcV-(GA)T-P-tBu	773	9.8	2.44	31.0	12.7
16, 1.7 mM AcV-(GA)T-P-NH_2	492	6.3	4.35	22.0	5.06
17, 2.7 mM P-(GA)T-G	31	0.4	nd	nd	
18, 3.0 mM A-(GA)T-P	29	0.4	nd	nd	
19, 2.8 mM P-(GA)T-P	21	0.3	nd	nd	
20, 3.0 mM A-P-(GA)T	34	0.4	nd	nd	
21, 1.5 mM A-P-(GA)T-S	117	1.5	2.94	4.5	1.53
22, 2.0 mM A-P-(GA)T-S-S	703	9.0	2.44	19.5	8.01
23, 1.8 mM A-P-(GA)T-S-S-S	1764	22.5	0.61	30.3	49.7
24, 1.5 mM A-P-(GA)T-S-S-A	2410	30.7	0.55	41.4	75.3
25, 1.5 mM A-P-(GA)T-S-A-S	1699	21.7	1.04	34.2	32.9
26, 1.5 mM A-P-(GA)T-A-S-S	1868	23.8	1.72	38.0	22.1
27, 1.4 mM A-P-T-(GA)S-S-S	395	5.06	0.79	8.3	10.6
28, 2.6 mM A-P-(GA)T-S-S-S-T-K	2034	26.0	0.65	34.2	52.6
29, 2.8 mM A-P-(GA)T-S-S-S-T-K-K-T	2242	28.6	1.35	39.1	29.0
30, 2.3 mM (GA)T-(GA)T-G-V-A	1510	19.3	0.24	22.8	95.1
31, 2.6 mM A-P-(GA)T-(GA)S-S-S	820	10.5	0.38	12.4	32.7
32, 1.2 mM A-(GA)T-V-(GA)T-A-G[b]	1261	16.1	0.53	23.8	44.9

[a] Abbreviations are as in Table I. Gal-transferase assays were carried out as described under Experimental Procedures; product was counted after the AG1×8 step without HPLC separation. The incubation time was 1 h. By use of 1.55 mM UDP-Gal and three to eight different acceptor concentrations in the assays, the apparent K_M and V_{max} were determined by linear double-reciprocal Lineweaver–Burk plots. nd, not determined because the incorporation of Gal into these acceptors was low. [b] Assays were carried out with a more concentrated enzyme preparation; the dpm were calculated relative to 555 dpm for 2 mM GA-Bn as the substrate.

(compounds **10**, **11**, **15**, and **16**) resulted in good acceptor activity. However, protected Ac(GalNAc)Thr-tBu (compound **4**, Table II) was not an effective substrate, indicating that both the absence of a net charge and the absence a peptide chain was required for good substrate activity. The protected amides were about 40% as effective as the corresponding *tert*-butyl esters (compare V_{max}/K_M of compounds **10** and **11** and compounds **15** and **16**, Table II).

Ala-Pro-(GalNAc)Thr-Ser-Ser-Ser (compound **23**) represents the N-terminal sequence of interleukin-2 (Conradt et al., 1985) and is a very effective acceptor (K_M of 0.61 mM and a V_{max} of 30.3 nmol min^{-1} mg^{-1}). Replacement of the carboxy-terminal Ser residue of compound **23** by Ala created the excellent acceptor Ala-Pro-(GalNAc)Thr-Ser-Ser-Ala (compound **24**) with a V_{max}/K_M of 75.3. This hexapeptide was the most effective substrate in the interleukin-2 series (compounds **20**–**29**, Table II); addition of two (compound **28**) or four (compound **29**) extra amino acids did not increase the activity further.

Changing the glycosylation site of Ala-Pro-(GalNAc)-Thr-Ser-Ser-Ser (compound **23**) from Thr to the adjacent Ser (compound **27**) reduced the V_{max} about 3-fold. An additional adjacent GalNAc residue (compound **31**) caused the V_{max} to drop from 30.3 to 12.4 nmol min^{-1} mg^{-1} although the K_M of 0.38 mM is extremely low (about half that of compound **23**). Thus the two GalNAc residues may have a slight inhibitory effect on each other's glycosylation. Glycopeptides **30** and **31** contain two GalNAc residues per molecule, and therefore, three enzyme products are possible for Gal-transferase; only one product for compound **30** (Table I) and two products for compound **31** in a ratio of about 2:1 (Table I) were detected by HPLC. This may be due to the fact that only these products were in fact formed or that not all of the products separated under the given HPLC conditions. However, after prolonged incubation of compound **30** with rat colonic mucosal homogenates (Table III), several products were detected (Figure 2B).

Table III: Core 1 β3-Gal-transferase Activities from Pig Gastric Mucosal Microsomes

acceptor	enzyme activities (nmol h^{-1} mg^{-1})[a]
33, 5 mM GalNAcα-Bn	52.9
20, 3.2 mM A-P-(GA)T	2.5
22, 3.9 mM A-P-(GA)T-S-S	55.5
23, 3.7 mM A-P-(GA)T-S-S-S	80.4

[a] Abbreviations are as in Table I. Compounds were used in Gal-transferase assays without HPLC separation as described under Experimental Procedures. The incubation time was 2 h.

The two enzyme products of compound **31** (eluting at 26 and 42 min) could not be identified due to lack of the appropriate galactosylated standards. Assuming that these two products contain only one Gal residue per molecule, it appears that one GalNAc was preferred over the other. On the basis of the comparison of kinetic data for compounds **23** and **27**, the preferred GalNAc residue would be the one attached to Thr. This is in contrast to the higher activity found for glycopeptides containing (GalNAc)Ser (compounds **1** and **8**) over

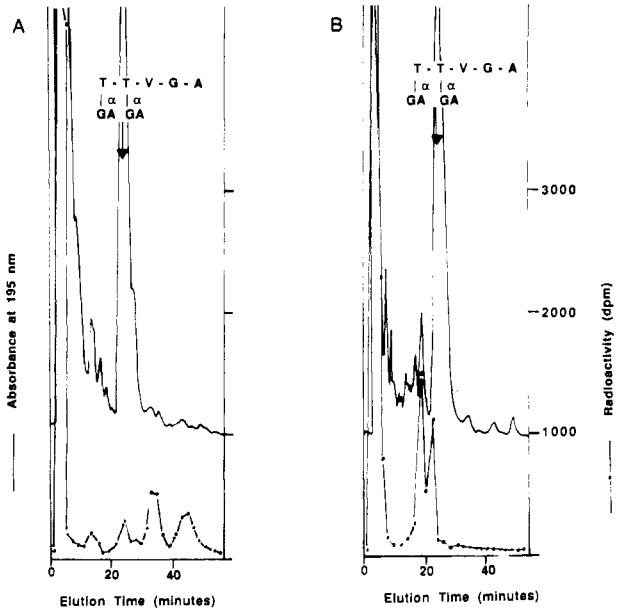

FIGURE 2: Rat colonic mucosal homogenate was incubated with (GA)Thr-(Ga)Thr-Val-Gly-Ala as the substrate in GlcNAc-transferase (A) and Gal-transferase (B) assays as described under Experimental Procedures. The arrows indicate the elution times of the substrate. Separation was carried out on a C18 column with acetonitrile–water = 1/99 as the mobile phase at a flow rate of 1 mL/min.

the corresponding glycopeptides containing (GalNAc)Thr (compounds **2** and **7**).

The presence of Pro or its position in the peptide did not appear to affect Gal-transferase activity. The best acceptor was (GalNAc)Thr-(GalNAc)Thr-Gly-Val-Ala (compound **30**) with a K_M of 0.24 mM and a V_{max}/K_M of 95.1. This glycopeptide lacks Pro but has a 10-fold higher V_{max}/K_M than GalNAcα-benzyl, indicating that the β3-Gal-transferase has a definite preference for a peptide over a hydrophobic benzyl group.

Substrate Specificity of Crude Gal- and GlcNAc-transferase Preparations from Pig Stomach and Colon. Pig gastric β3-Gal-transferase (Table III) was only tested with four substrates. The enzyme appeared to be stimulated by the addition of Ser residues to the interleukin-2 sequence Ala-Pro-(GalNAc)Thr, as was the case for the rat liver enzyme (Table II).

The core 1 β3-Gal-transferase and the core 3 β3-GlcNAc-transferase both compete for the same substrate GalNAcα-R (Figure 1). This competition was studied with pig colonic mucosal homogenate (Table IV), a tissue rich in core 3 synthesis (Brockhausen et al., 1985). Although rat colon makes both core 3 and core 4, pig colon converts GalNAcα-Bn mainly to core 3 with insignificant amounts of core 4. Small glycopeptide substrates [compounds **1–9** (Table II), data not shown, and compounds **12–14** and **17–20** (Table IV)] showed negligible Gal-transferase activity with pig colonic mucosal homogenates. As was found for rat liver Gal-transferase, compounds **23** and **30** proved to be appreciably more effective substrates than GalNAcα-Bn for the pig colon enzyme. Pig colon GlcNAc-transferase showed about the same activity with compounds **23** and **30** as with GalNAcα-Bn.

Substrate Specificity of Crude Gal- and GlcNAc-transferase Preparations from Rat Colon. Protected glycopeptides (compounds **4**, **10**, **11**, **15**, and **16**, Table V) were excellent acceptors for both Gal- and GlcNAc-transferases from rat colon. The amide appears to be a better acceptor than the *tert*-butyl ester for the Gal-transferase, but not for the GlcNAc-transferase. HPLC patterns for the products of compound **11** (data not shown) and compound **10** (Figure 3) show the expected number of enzyme products (i.e., one Gal-transferase product and two GlcNAc-transferase products) and no degradation by proteases. However, some of the unprotected glycopeptides show a reduction in the absorbance peaks of the substrates as well as the occurrence of multiple products (data not shown). A protease inhibitor cocktail could not entirely eliminate this problem.

Despite the protease degradation of substrate and products, the results with rat colon Gal-transferase were similar to those

Table IV: Gal- and GlcNAc-transferase Activities from Pig Colonic Mucosal Homogenate[a]

	nmol h^{-1} mg^{-1}	
acceptor	Gal-T activities	GlcNAc-T activities
33, 2 mM GalNAcα-Bn	3.5	1.9
12, 1.2 mM P-(GA)T-V	<0.1	<0.1
13, 1.5 mM V-(GA)T-P	<0.1	<0.1
14, 1.3 mM AcV-(GA)T-P	<0.1	<0.1
17, 1.3 mM P-(GA)T-G	<0.1	0.4
18, 1.5 mM A-(GA)T-P	0.1	0.3
19, 1.4 mM P-(GA)T-P	<0.1	0.3
20, 1.5 mM A-P-(GA)T	<0.1	0.2
22, 2.0 mM A-P-(GA)T-S-S	1.2	<0.1
23, 1.9 mM A-P-(GA)T-S-S-S	14.6	0.8
30, 1.1 mM (GA)T-(GA)T-G-V-A	10.7	1.4

[a] Abbreviations are as in Table I. Compounds were used in Gal-transferase and GlcNAc-transferase assays with HPLC separation as described under Experimental Procedures. The incubation time was 2 h. GalNAcα-Bn is converted by the pig colonic GlcNAc-transferase to core 3 (GlcNAcβ1–3GalNAcα-Bn) without significant formation of core 4 [GlcNAcβ1–6(GlcNAcβ1–3)GalNAcα-Bn].

Table V: Gal-transferase and GlcNAc-transferase Activities toward GalNAc-glycopeptide Substrates Using Rat Colonic Mucosal Homogenate[a]

	nmol h^{-1} mg^{-1}	
acceptor	total Gal-T activities	total GlcNAc-T activities
33, 2.0 mM GA-Bn	9.5	10.6[b]
4, 1.3 mM Ac(GA)T-tBu	2.4	14.6
10, 1.4 mM Ac(GA)T-P-tBu	6.1	7.1
11, 1.6 mM Ac(GA)T-P-NH_2	22.7	5.1
12, 1.2 mM P-(GA)T-V	<0.1	0.3
13, 1.5 mM V-(GA)T-P	0.3	<0.1
14, 1.3 mM AcV-(GA)T-P	0.2	1.2
15, 1.4 mM AcV-(GA)T-P-tBu	18.4	6.0
16, 1.7 mM AcV-(GA)T-P-NH_2	34.0	
17, 1.3 mM P-(GA)T-G	1.2	0.7
18, 1.5 mM A-(GA)T-P	2.2	<0.1
19, 1.4 mM P-(GA)T-P	<0.1	0.6
20, 1.5 mM A-P-(GA)T	<0.1	<0.1
21, 1.5 mM A-P-(GA)T-S	0.7	2.8
22, 2.0 mM A-P-(GA)T-S-S	3.7	5.9
23, 1.9 mM A-P-(GA)T-S-S-S	6.4	2.8
24, 1.5 mM A-P-(GA)T-S-S-A	2.1	2.6
25, 1.5 mM A-P-(GA)T-S-A-S	11.1	0.4
26, 1.5 mM A-P-(GA)T-A-S-S	1.9	1.8
27, 1.4 mM A-P-T-(GA)S-S-S	4.0	3.4
28, 1.1 mM A-P-(GA)T-S-S-S-T-K	9.1	0.7
30, 1.1 mM (GA)T-(GA)T-G-V-A	18.9	8.1
31, 1.3 mM A-P-(GA)T-(GA)S-S-S	15.0	3.9

[a] Abbreviations are as in Table I. Gal- and GlcNAc-transferase activities were assayed by HPLC as described under Experimental Procedures. The incubation time was 2 h. Rat colon converts GalNAcα-Bn to GlcNAcβ3GalNAcα-Bn due to the core 3 β3-GlcNAc-transferase; this product is further converted to GlcNAcβ6-(GlcNAcβ3)GalNAcα-Bn by the core 4 β6-GlcNAc-transferase. The activities were calculated from the total radioactivity transferred. [b] Core 3 β3-GlcNAc-transferase activity was 5.7 nmol h^{-1} mg^{-1}; core 4 β6 GlcNAc-transferase activity was 4.9 nmol h^{-1} mg^{-1}.

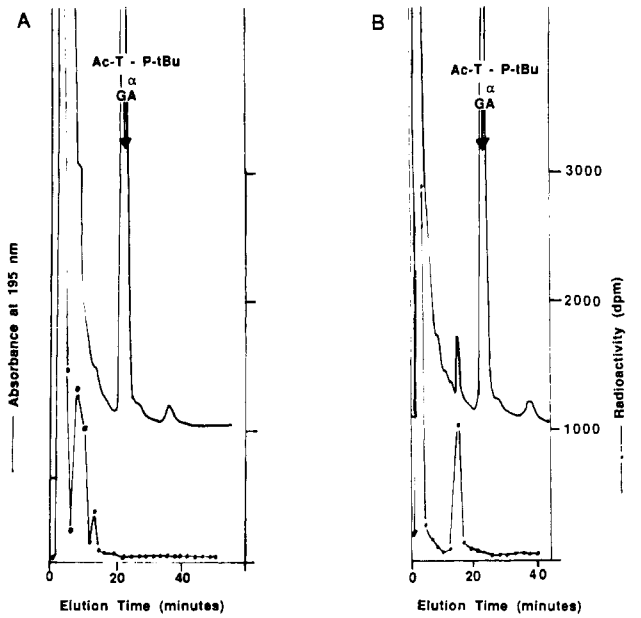

FIGURE 3: Rat colonic mucosal homogenate was incubated with Ac-(Ga)Thr-Pro-tBu as the substrate in GlcNAc-transferase (A) and Gal-transferase (B) assays as described under Experimental Procedures. The arrows indicate the elution times of the substrate. Separation was carried out on a C18 column with acetonitrile–water = 17/83 as the mobile phase at a flow rate of 1 mL/min.

obtained with the purified rat liver enzyme. A few differences were, however, observed between rat colon and liver Gal-transferases. For example, the amides (compounds **11** and **16**) were better substrates than the corresponding *tert*-butyl esters for the rat colon whereas the reverse was the case for rat liver. Rat liver and colon also differ in their relative preferences for glycopeptide **31** with two GalNAc residues and glycopeptide **23** with only one GalNAc residue. These effects may be due to differences between the Gal-transferases in non-mucin-secreting (rat liver) and mucin-secreting (rat colon) tissues or to differences in protease action between the two enzyme preparations. It is likely that the presence of additional GalNAc residues reduces the action of proteases.

GlcNAc-transferases from rat colon (Table V) showed considerable activity toward the protected compounds **4**, **10**, **11**, and **15** as enzyme substrates and showed very little activity toward unprotected glycopeptides with less than four amino acids. The larger glycopeptides with the exception of compounds **25** and **28** were effective substrates. The HPLC patterns showed more than one enzyme product (Figure 3A), indicating that both core 3 and core 4 were synthesized with these glycopeptides, as was previously shown with GalNAcα-Bn (Brockhausen et al., 1985). However, the relative GlcNAc-transferase activities toward these acceptors (especially toward glycopeptides **22–28**, Table V) did not parallel the Gal-transferase activities. These differences suggest that although the peptide moiety affects both Gal- and GlcNAc-transferase activities, the enzymes differ in their responses to the peptide environment.

Discussion

O-Linked oligosaccharides may be found on secreted and membrane-bound glycoproteins. Although their functions are largely unknown, these oligosaccharide chains may play a role in many biological phenomena, such as cellular recognition, in the immune system, in hormonal action, and in fertilization. Carbohydrate has important effects on the physical and chemical properties of proteins. Highly glycosylated GalNAc-Ser/Thr-containing glycoproteins (mucins) synthesized in specialized mucin-secreting cells protect and lubricate underlying epithelial cells. The *O*-glycan chains may contain many antigens and binding sites for microflora. Mucins are also synthesized in many tumor cells and often carry carbohydrate tumor antigens (Magnani et al., 1983; Brockhaus et al., 1985; Feizi et al., 1984; Hanish et al., 1985; Carraway & Spielman, 1986; Lan et al., 1987; Stähli et al., 1988; Yamori et al., 1989; Itzkowitz et al., 1989). To understand the mechanisms and controls of *O*-glycan biosynthesis it is essential to study the detailed specificities of the enzymes that build the core structures which provide the framework for the attachment of antigenic determinants or can represent antigens themselves (Schachter, 1985; Schachter & Brockhausen, 1989).

N-Glycoproteins can carry different oligosaccharide chains on different glycosylation sites. It has been suggested that the role of the peptide in determining the carbohydrate processing may be to provide various accessibilities for different glycosylation sites. However, if competing transferases are subjected to the same degree of accessibility, differences in the final processing products must originate from the specificities of processing enzymes that recognize either the oligosaccharide or the peptide or both. Carver and Cummings (1987) suggested that a mechanism for site-directed processing may be the recognition by processing enzymes of three-dimensional conformations of oligosaccharide substrates. These conformations may be a result of interactions of the glycans with the peptide. The influence of the peptide on *O*-glycan processing has not yet been studied.

In this paper, we have for the first time presented evidence that the synthesis of *O*-glycan cores 1 and 3 is controlled by a direct recognition by glycosyltransferases of the peptide moiety of *O*-glycopeptide substrates. We have studied the activity of the β3-Gal-transferase that synthesizes core 1 (Galβ1–3GalNAcα-) from GalNAcα-R. The enzyme has been described in rat liver (Anderson & Ericsson, 1981) and mucin-secreting tissues (Mendicino et al., 1982; Cheng & Bona, 1982). The partially purified rat liver enzyme (I. Brockhausen et al., in preparation) is highly sensitive to changes in the length, composition, and sequence of the peptide. A length of at least five amino acids is required for efficient synthesis of core 1. The enzyme is also influenced by a charge proximal to the glycosylation site, an adjacent GalNAc residue, and the attachment site of GalNAc. A glycosylation site close to the C-terminal end of a glycoprotein probably does not have a good chance of being elongated due to the charged carboxyl group.

Compounds **20–23**, **28**, and **29** are part of the N-terminal sequence of interleukin-2. This glycoprotein contains only one *O*-glycan chain attached to the Thr of Ala-Pro-Thr-R; this glycan is entirely of sialylated core 1 structure (Conradt et al., 1985). Our results indicate that the peptide would also allow synthesis of core 3 and 4 at this glycosylation site, if the respective enzymes were active in interleukin-2-synthesizing T-cells.

The conformations of several of the *O*-glycopeptides used in the present study (i.e., compounds **5**, **7**, **9**, **20–23**, and **29**) have been analyzed by NMR studies and energy calculations (Pollex-Krüger, 1989; Paulsen et al., 1990). It was determined that the glycopeptides have extended but characteristic conformations, although the amino acid side chains are similar in different compounds. The GalNAc-Thr units assume the same conformations in all glycopeptides, shown by similar coupling constants of Thr protons and nuclear Overhauser effects between Thr and GalNAc protons. Differences in

enzyme activities must therefore reside in differences of the peptide chain adjacent to the glycosylation site. As the β3-Gal-transferase activity was not significantly changed after the peptide chain had been extended past six residues, we suggest that the binding site of this enzyme accommodates the glycosylation site including six amino acids.

Changes in the peptide sequence or overall conformation would result in different binding strengths indicated by changes in the K_M. Differences in activities between GalNAc-Thr and GalNAcα-Bn, or between the protected *tert*-butyl esters and corresponding amides, indicate that one factor determining enzyme kinetics is a hydrophobic effect. It is not clear whether this effect is due to an actual hydrophobic binding site or solubilization properties of substrates favoring binding of hydrophobic substrates. The presence of the hydrophobic amino acid Val in a glycotripeptide was of no advantage, possibly due to interference of a nearby charge, but the best glycopeptide substrate with the highest V_{max}/K_M did contain Val.

Detailed kinetics were carried out with a purified rat liver β3-Gal-transferase. Purified GlcNAc-transferases involved in the synthesis of mucin cores are not yet available. The rat colon provided a tissue containing three enzymes which can glycosylate the GalNAc residue. It was evident that although the rat colon Gal-transferase largely resembled the liver enzyme, the GlcNAc-transferase showed quite different activities toward the same acceptors. To confirm these results, kinetic studies should be carried out upon purification of these GlcNAc-transferases.

Since the mutually exclusive core 1 and core 3 transferases (Figure 1) compete for GalNAc-substrate, the relative activities of these enzymes determine which type of core will be made. It is thus evident that the peptide structure in the substrate can direct mucin core synthesis and each individual O-glycosylation site may carry a characteristic mucin core structure. Structural studies are required to support this theory.

The fact that protected glycopeptides with two or three amino acids proved to be very good acceptors allows future studies of peptide influence on transferases by use of short glycopeptides. The protected amides are recommended for these studies because they rule out a possible artifactual hydrophobic effect and resemble a natural peptide chain.

Acknowledgments

We are grateful to Dr. Harry Schachter for generous support and advice.

References

Anderson, G. N., & Ericsson, L. C. (1981) *J. Biol. Chem. 256*, 9633–9639.

Bendiak, B., & Schachter, H. (1987) *J. Biol. Chem. 262*, 5775–5783.

Briand, J. P., Andrews, S. P., Cahill, E., Conway, N. A., & Young, J. D. (1981) *J. Biol. Chem. 256*, 12205–12207.

Brockhaus, M., Wysocka, M., Magnani, J. L., Steplewski, Z., Koprowski, H., & Ginsburg, V. (1985) *Vox Sang. 48*, 34–38.

Brockhausen, I., Matta, K. L., Orr, J., & Schachter, H. (1985) *Biochemistry 24*, 1866–1874.

Brockhausen, I., Matta, K. L., Orr, J., Schachter, H., Koenderman, A. H. L., & van den Eijnden, D. (1986) *Eur. J. Biochem. 157*, 463–474.

Carraway, K. L., & Spielman, J. (1986) *Mol. Cell. Biochem. 72*, 109–120.

Carver, J. P., & Cumming, D. (1987) *Pure Appl. Chem. 59*, 1465–1476.

Cheng, P.-W., & Bona, S. (1982) *J. Biol. Chem. 257*, 6251–6258.

Conradt, H. S., Geyer, R., Hoppe, J., Grotjahn, L., Plessing, A., & Mohr, H. (1985) *Eur. J. Biochem. 153*, 255–261.

Feizi, T., Gooi, H. C., Childs, R. A., Picard, J. K., Uemura, K., Loomes, L. M., Thorpe, S. J., & Hounsell, E. F. (1984) *Biochem. Soc. Trans. 12*, 591–596.

Hanish, F.-G., Uhlenbruck, G., Dienst, C., Stottrop, M., & Hippau, E. (1985) *Eur. J. Biochem. 149*, 323–330.

Itzkowitz, S. H., Yuan, M., Montgomery, C. K., Kjeldsen, T. Takahashi, H. K., Bigbee, W. L., & Kim, Y. S. (1989) *Cancer Res. 49*, 197–204.

Lan, M. S., Bast, R. C., Colnaghi, M. I., Knapp, R. C., Colcher, D., Schlom, J., & Metzgar, R. S. (1987) *Int. J. Cancer 39*, 68–72.

Mendicino, J., Sivakami, S., Davila, M., & Chandrasekaran, E. V. (1982) *J. Biol. Chem. 257*, 3987–3994.

Paulsen, H., & Adermann, K. (1989) *Liebigs Ann. Chem.*, 751–769, 771–780.

Paulsen, H., Adermann, K., Merz, G., Schultz, M., & Weichert, U. (1988) *Starch 40*, 465–472.

Paulsen, H., Pollex-Krüger, A., & Sinnwell, V. (1990) *Carbohydr. Res.* (in press).

Pollex-Krüger, A. (1989) Ph.D. Thesis, University of Hamburg, FRG.

Schachter, H. (1985) *Biochem. Cell. Biol. 64*, 163–181.

Schachter, H., & Brockhausen, I. (1989) *Symp. Soc. Exp. Biol. 43*.

Stähli, C., Caravatti, M., Aeschbacher, M., Kocyba, C., Takacs, B., & Carmann, H. (1988) *Cancer Res. 48*, 6799–6802.

Yamori, T., Ota, D. M., Cleary, K. R., Hoff, S., Hager, L. G., & Irimura, T. (1989) *Cancer Res. 49*, 887–894.

Young, J. D., Tsuchiya, D., Sandlin, D. E., & Holroyde, M. J. (1979) *Biochemistry 18*, 4444–4448.