THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 273, No. 16, Issue of April 17, pp. 9608–9614, 1998

Printed in U.S.A.

Mutation of the Sialyltransferase S-sialylmotif Alters the Kinetics of the Donor and Acceptor Substrates*

(Received for publication, December 31, 1997)

Arun K. Datta‡, Abhishek Sinha, and James C. Paulson

From Cytel Corporation and the Department of Chemistry and Molecular Biology, Scripps Research Institute, San Diego, California 92121

Protein sequence analysis of the cloned sialyltransferase gene family has revealed the presence of two conserved protein motifs in the middle of the lumenal catalytic domain, termed L-sialylmotif and S-sialylmotif. In our previous study (Datta, A. K., and Paulson, J. C. (1995) *J. Biol. Chem.* 270, 1497–1500) the larger L-sialylmotif of ST6Gal I was analyzed by site-directed mutagenesis, which provided evidence that it participates in the binding of the CMP-NeuAc, a common donor substrate for all the sialyltransferases. However, none of the mutants tested in this motif had any significant effect on their binding affinities toward the acceptor substrate asialo α_1-acid glycoprotein. In this study, we have investigated the role of the S-sialylmotif of the same enzyme ST6Gal I. In total, nine mutants have been constructed by changing the conserved amino acids of this motif to mostly alanine by site-directed mutagenesis. Kinetic analysis for the mutants which retained sialyltransferase activity showed that the mutations in the S-sialylmotif caused a change of K_m values for both the donor and the acceptor substrates. Our results indicated that this motif participates in the binding of both the substrates. A sequence homology search also supported this finding, which showed that the downstream amino acid sequence of the S-sialylmotif is conserved for each subgroup of this enzyme family, indicating its association with the acceptor substrate.

The transfer of sialic acid from its nucleotide donor CMP-Sia[1] to the nonreducing terminus of oligosaccharyl structures of various glycoproteins and glycolipids is mediated by sialyltransferases, a group of about 15 enzymes belonging to the family of glycosyltransferases (1). Cloning of the first three sialyltransferases, namely ST6Gal I (see Ref. 2 for nomenclature), ST3Gal I (3), and ST3Gal III (4), and analysis of their protein sequences deduced from the respective cDNAs revealed that all of these gene products have the type II membrane protein topology common among all the members of the glycosyltransferases examined to date. This topology is characterized by the presence of a short N-terminal cytoplasmic domain, an uncleavable hydrophobic signal-anchor sequence that serves as the membrane-spanning domain, a proteolytically sensitive "stem" region, and a large catalytic domain that projects into the lumen of the Golgi apparatus. In contrast to other glycosyltransferases, however, the sequences derived from the sialyltransferase gene family revealed the presence of two conserved protein motifs in the lumenal catalytic domain, termed as L- and S-sialylmotifs, consisting of about 48 and 23 amino acids, respectively (5–7). This unique feature was exploited to clone new genes of this family by a PCR-based strategy using degenerated primers designed based on the invariant amino acids present at either end of these two motifs. This strategy was highly successful in cloning nine additional sialyltransferases (5, 7–14). The cloning of these new genes by PCR confirmed that the presence of sialylmotifs in any new gene is, in fact, the "cardinal" feature of the sialyltransferase gene family. By now, in total 13 genes have been cloned (2). Analysis of their protein sequences revealed that each of these enzymes shares this common structural feature: the presence of L-sialylmotif and S-sialylmotif. Comparison among the members showed that the larger L-sialylmotif consisting of 48–49 amino acids in the center of each molecule exhibits >70% identity among homologous groups, 40–60% among heterologous groups, and has eight invariant amino acids including one invariant cysteine. The S-sialylmotif, on the other hand, has two invariant amino acids in a stretch of 23 residues. Interestingly, one of these is also cysteine. These two cysteines are predicted to be involved in disulfide linkage formation (6).

Such a unique feature of the presence of these two sialylmotifs in all the cloned sialyltransferases suggests that these motifs might contribute to a structural feature related to the common function of these enzymes. For example, each enzyme transfers sialic acid from the common donor substrate, CMP-NeuAc, to an oligosaccharide acceptor substrate. Thus, these motifs could form part of the binding sites for either the donor or acceptor substrates, or both. Nevertheless, their role in the formation of a specific conformation required for its enzymatic activity may not be ruled out.

Previously, we examined the role of the L-sialylmotif using site-specific mutants of the enzyme ST6Gal I (EC 2.4.99.1) as a model. This enzyme, which forms the Neu5Acα2,6Galβ1,4GlcNAc sequence common to many Asn-linked oligosaccharides, was cloned from rat liver (15). Our results showed that the alanine mutation of some of the invariant amino acids in the L-sialylmotif of this enzyme resulted in an increased K_m toward the donor substrate CMP-NeuAc without significant effect on the acceptor binding affinity, suggesting that this motif contributes to the binding of the common donor substrate (16, 17). Here, we have examined the role of the S-sialylmotif of the same enzyme ST6Gal I and show that the single point mutation of its conserved amino acids affected the K_m values of the mutant enzymes for both the donor and acceptor substrates.

* This work was supported in part by United States Public Health Service Grant GM27904 (to J. C. P.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Chemistry and Molecular Biology, BCC 338, 10550 N. Torrey Pines Rd., San Diego, CA 92037. Tel.: 619-784-2441; Fax: 619-784-2409; E- mail: adatta@scripps.edu.

[1] The abbreviations used are: CMP-Sia, CMP-sialic acid; ASGP, asialo α_1-acid glycoprotein; PCR, polymerase chain reaction; CMP-NeuAc, CMP-neuraminic acid; nt, nucleotide(s).

EXPERIMENTAL PROCEDURES

Construction of Mutants—The single point mutants were constructed following the megaprimer method described earlier (16, 18). The cDNA for the ST6Gal I (GenBank accession no. M18769), previously subcloned in pBluescript KS+ plasmid (19), was used as a template. The nomenclature for the mutant sialyltransferases has been assigned the number that reflects the amino acid residue number for which alanine has been substituted; *e.g.* for P318A, proline 318 is replaced by alanine. The mutagenic antisense oligonucleotides (the substituted nucleotides are underlined) used for construction of the corresponding cDNAs were as follows (5′-3′): P318A, GCCGGAGGATGCGGGATTTGGCTGA (nt 969–939); S319A, CCAGCATGCCGGAGGCTGGGGGATT (nt 970–946); S320A, CCAGCATGCCGGCGGATGGGGGATT (nt 970–946); G321A, GATACCCAGCATGGCGGAGGATGG (nt 975–952); C332A, CCTGGTCAGCCAGCGTCATC (nt 984–1003); V335A, CGTAAATATCTGCCTGGTCACAC (nt 1015–993); V335L, CGTAAATATCTAGCTGGTCACAC (nt 1015–993); E339A, GGGAGGAACGCGTAAATATCTA (nt 1025–1004); F340A, GGATGGGAGGGCCTCGTAAATA (nt 1029–1008). As described earlier (16, 18), in the first step of PCR, the 5′-end oligonucleotide primer, GCTCTAGAATTCCAATCCTCAGTTACCACAG (5′-3′, nt 214–236), and the mutant antisense oligonucleotide (25–50 pmol each) were used to introduce the desired mutation using *Pfu* DNA polymerase (Stratagene); the conditions used were: 94 °C, 30 s; 56 °C, 1 min; 73 °C, 2 min for 20 cycles. The gel analysis showed the generation of a single band for about 780 base pairs depending on the position of the oligomer. This double-stranded DNA product from each reaction was purified using Geneclean II (Bio101, San Diego, CA) and used as a megaprimer in the second step of PCR. Our earlier results showed that the yield of the final desired product is improved by 5 cycles of linear amplification using only one primer, the megaprimer. The reaction mixture for the linear amplification contained 10 ng of template, 80 μM dNTPs in the 1 × *Pfu* DNA polymerase buffer and ~50 ng of megaprimer (estimated from the agarose gel) and started with the addition of 2.5 units of *Pfu* DNA polymerase following these conditions: 94 °C, 1 min; 73 °C, 3 min for 5 cycles. The 3′-end antisense oligonucleotide primer, CCAGGAGAGGATCCATAAAATGAC (5′-3′, nt 1270–1247), was then added, and the reaction was continued as follows: 94 °C, 1 min; 68 °C, 1 min; 73 °C, 3 min for 20 cycles. The products were analyzed by agarose gel electrophoresis, which showed the generation of a major single band of 1.05 kilobases for the specific product. This band was purified by agarose gel electrophoresis followed by Geneclean to separate from the megaprimer. The gel-purified product for the mutants was digested with *Bst*BI (at nt 824) and *Bsp*E1 (at nt 1194). The 370-base pair fragment containing the mutation was purified for each mutant and subcloned into a similarly digested and purified larger fragment of spST-2 (20) following a standard procedure (21). The ligation mixture was used to transform competent cells of *Escherichia coli*. The colonies were isolated, and the plasmids were obtained using Promega's plasmid miniprep kit (Promega). The mutation was confirmed by dideoxy double-stranded sequencing (22) of the entire fragment that has been subcloned, including the restriction sites used.

Expression of the Wild-type and Mutant Sialyltransferases—For the analysis, the wild-type ST6Gal I and its mutants were transiently expressed in COS-1 cells as described earlier (16). Cells (1–2 × 10^6 cells/100-mm dish) were transfected using 2.0 μg of plasmid DNA using LipofectAMINE reagent according to the supplier (Life Technologies, Inc.). Expression of transfected proteins was typically allowed to continue for 36–48 h post-transfection before harvesting the cells. The culture medium was then collected and concentrated 10-fold by ultrafiltration using micro-concentrators (MWCO 10; Amicon Inc., Beverly, MA). The concentrated medium containing the soluble-expressed sialyltransferase was used directly for analysis of the enzymatic activity as described earlier (16). These transfected cells were also used for radiolabeling of the expressed proteins. Transfection was repeated at least three times for each mutant with plasmid DNAs from different preparations.

Pulse-chase Labeling of Transfected COS-1 Cells and Analysis of the Transiently Expressed Proteins—Metabolic labeling of cells using Trans ^{35}S-Express protein label (NEN Life Science Products; 100 μCi/ml) was carried out essentially as described elsewhere (23). The radiolabeled medium from the transfected COS-1 cells was used for immunoprecipitation as follows: the radiolabeled medium (500 μl) was incubated with 10 μl of pre-immune rabbit serum for 20 min at room temperature. 25 μl of protein A-Sepharose (Amersham Pharmacia Biotech) was added and incubated with rotation at room temperature for 30 min. These were centrifuged, and the supernatants were collected. 10 μl of affinity-purified rabbit anti-rat ST6Gal I was added and incubated at 4 °C for overnight. 20 μl of protein A-Sepharose was added and incubated at room temperature for 30 min. The precipitated immune complex obtained by centrifugation was washed three times with immunoprecipitation buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.5% Nonidet P-40, 0.1% SDS) and once with 10 mM Tris-HCl, pH 7.5, 0.1% SDS. Proteins were eluted from the pellet by boiling for 5 min in Laemmli gel sample buffer containing 10% β-mercaptoethanol. Immunoprecipitated proteins were electrophoresed on 10% SDS-polyacrylamide gels according to the method of Laemmli (24). The gel was then soaked for fixing once in 10% acetic acid, 30% ethanol for 40 min and then in Enlightning rapid autoradiography enhancer as instructed by the supplier (NEN Life Science Products) before exposed to Kodak Biomax-MR film at −80 °C.

Other Methods—The Western blot, protein determination, sialyltransferase assay, and the enzyme assay for kinetic analysis were done as described earlier (16).

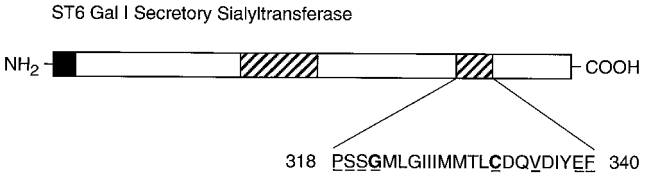

FIG. 1. **spST-2 construct for the soluble form of ST6Gal I and its mutants in relation to the domain structure of sialyltransferase.** The spST-2 was constructed by replacing the first 71 amino acids from the N-terminal of ST6Gal I with the cleavable signal anchor sequence from the dog pancreatic proinsulin (20). The S-sialylmotif in this enzyme spans from amino acids 318 to 340 and contains two invariants, indicated by *boldfaced letters*. The *underlined* amino acid residues were changed to alanine by single point mutagenesis. ■, cleavable signal sequence; ▨, sialylmotifs; □, catalytic domain.

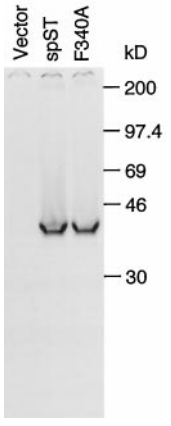

FIG. 2. **Immunoprecipitation of wild-type ST6Gal I (spST-2) and its mutants from ^{35}S-labeled medium of transfected COS-1 cells.** Metabolic labeling of transfected COS-1 cells using Trans ^{35}S-Express protein label (DuPont NEN; 100 μCi/ml) and immunoprecipitation of expressed proteins with rabbit anti-rat ST6Gal I was performed using medium from transfected cells of a 48–60-h post-transfection, essentially as described earlier (16). The fluorogram was shown only for F340A as a representative of mutants. As shown, the mutant sialyltransferases moved similarly with the wild-type enzyme (spST-2) in the SDS-polyacrylamide gel, indicating no gross rearrangement because of the mutation.

RESULTS

Expression of the S-sialylmotif Mutants in COS-1 Cells and Comparison of Their Enzyme Activities—For expression of the wild-type ST6Gal I and its mutants, the mammalian expression vector construct used was based on pSVL (Pharmacia), which contained the soluble form of the cDNA for rat ST6Gal I as mentioned earlier (16). The mutants were derived from this expression vector of the wild-type enzyme spST-2, in which the N-terminal cytoplasmic tail, signal anchor domain, and the stem regions (first 71 amino acids) were replaced with a cleavable signal peptide of dog pancreas insulin (20). Thus, when the cDNAs for the wild-type ST6Gal I enzyme and its mutants

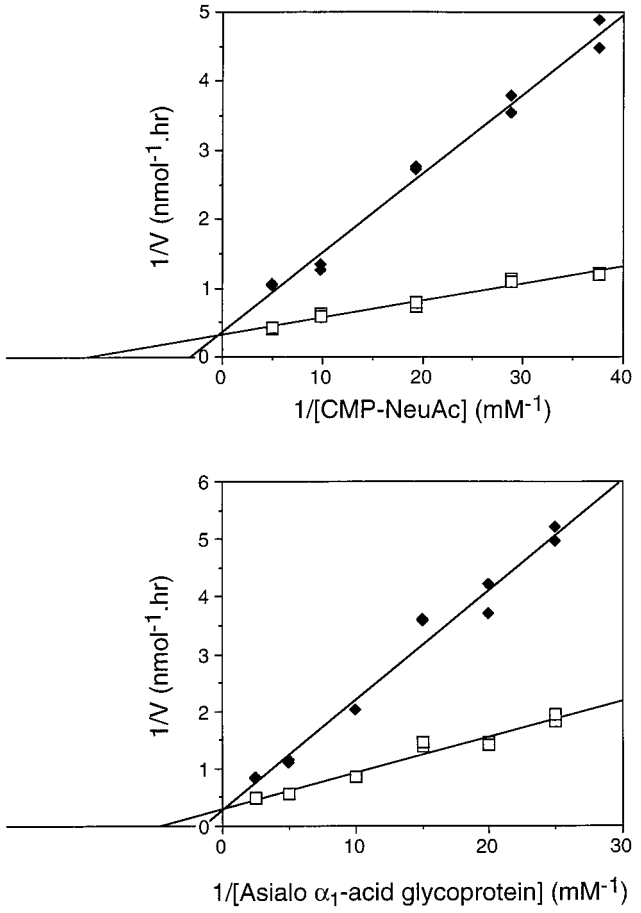

FIG. 3. **Double reciprocal plots of initial rate data with CMP-NeuAc (*top*) or asialo α_1-acid glycoprotein (*bottom*) as the varied substrate.** *Top*, rate data with CMP-NeuAc as the donor (26.5–201 μM) were determined at fixed concentration of the acceptor asialo α_1-acid glycoprotein, 50 μg (equivalent to 20.75 μmol of Gal acceptor unit). *Bottom*, the concentration of the acceptor asialo α_1-acid glycoprotein was varied (0.04–0.4 mM) at a fixed concentration of the donor, 0.15 mM. The plot was shown for E339A (♦) as a representative of the mutants and compared with that of the wild type (spST-2) of sialyltransferase (□). The K_m (apparent) values were determined from the X-intercept ($-1/K_m$ (apparent)) (61) using the Cricket Graph program (Cricket Software, Malvern, PA).

were transfected into COS-1 cells, the proteins were expressed as a soluble form and were detected in the culture medium. For this experiment, we designed the single point mutants, changing the corresponding amino acids to alanine. A total of eight such mutants for the S-sialylmotif were constructed (Fig. 1). In addition, we also constructed mutant V335L by replacing Val^{335} with Leu.

For rapid evaluation of the effect of mutation of the conserved amino acids, we used COS-1 because of its low endogenous sialyltransferase activity. Medium from COS-1 cells mock-transfected using only the vector (pSVL) showed typically less than 2% activity compared with that of expressed soluble sialyltransferases. Freshly grown COS-1 cells were transfected with the cDNAs for both wild-type ST6Gal I and its mutants. The expression of the desired sialyltransferase proteins were checked by immunoprecipitation of the metabolically labeled proteins. Immunoprecipitation of ^{35}S-radiolabeled proteins using the anti-ST6Gal I sialyltransferase antibody, followed by SDS-polyacrylamide gel electrophoresis, indicated that all mutant sialyltransferases were expressed and exhibited similar molecular weight to the wild-type sialyltransferase (Fig. 2). Their similar migration in the SDS-polyacrylamide gel electrophoresis indicated that there may not be a gross rearrangement in their native structure because of the introduced point mutation. To determine the enzymatic properties of the mutant sialyltransferases, media from the transfected cells were concentrated (Amicon Inc., Beverly, MA) and used as a source of crude enzyme. By Western blot, we estimated the relative amount of proteins that is used for sialyltransferase activity. By comparative analysis, it was noted that, although the expression levels were apparently similar to that of the wild-type enzyme for S320A, G321A, V335A, V335L, E339A, and F340A, the enzyme activities obtained were 41, 75, 83, 79, 35, and 15%, respectively. On the other hand, although the protein expression for S319A and C332A seemed to be within 2–5-fold, the sialyltransferase activity measured was less than 5%. For P318A, the expression achieved was, however, very low, if any. Repeated experiments of transfection for the expression of P318A failed to achieve detectable expression levels for this mutant protein. This proline is conserved in most of the sialyltransferases cloned. Proline is known to confer unique conformational constraints on the peptide chain of many biologically important proteins. This conformational restriction seems to be important for the biological functions of these proteins, as has been evidenced by a site-directed mutagenesis study (for a review, see Ref. 25). It is possible that the alanine mutation of this Pro^{318} induces some conformational change of ST6Gal I that affected its expression.

TABLE I
Kinetic constants for the wild-type ST6Gal I and its S-sialylmotif mutants

Kinetic analysis was performed using concentrated medium of COS-1 cells transfected with cDNAs of wild-type secretory sialyltransferase (spST-2) and mutants for S-sialylmotif. Methods have been described under "Experimental Procedures" and also in the text.

ST6Gal I constructs	Apparent K_m values[a]			
	CMP-NeuAc (μM)	-fold	Acceptor (ASGP)[b] (μM)	-fold
Wild type	50	1.0	330	1.0
Mutants:				
S320A	315	6.3	604	1.8
G321A	153	3.0	947	2.8
V335A	89	1.8	71	0.2[c]
V335L	80	1.6	322	1.0
E339A	308	6.1	710	2.1

[a] The apparent K_m values listed was the average of three experiments. In each case, the range of values observed was found to be within $\pm$0.5-fold of the indicated value.

[b] Asialo α_1-acid glycoprotein (ASGP) was prepared by mild acid hydrolysis of α_1-acid glycoprotein (Sigma) following the procedure of Schmid (62).

[c] Mutation of Val^{335} to Ala showed a 5-fold decrease in its K_m value toward the acceptor substrate.

Kinetic Analysis of the Mutant Enzymes—The mutants, which retained ST6Gal I sialyltransferase activity, were used for the comparative kinetic analysis. Analysis of the kinetics of the wild-type and each mutant sialyltransferase expressed in COS-1 cells was performed as described under "Experimental Procedures," except for F340A, which had low activity and was not evaluated. The K_m values for each sialyltransferase were determined for the donor substrate, CMP-NeuAc, and the best glycoprotein acceptor substrate of the wild-type enzyme, asialo α_1-acid glycoprotein (ASGP). This glycoprotein acceptor contains five *N*-linked oligosaccharides that are predominately bi- and tetra-antennary structures, each branch of which is terminated with the acceptor sequence Galβ1,4GlcNAc (26). Data obtained for each of the enzymes produced Lineweaver-Burke plots with correlation coefficients of 0.95 or greater (*e.g.* Fig. 3).

Table I summarizes the K_m values obtained for both the substrates. This shows that the K_m values for the donor substrate CMP-NeuAc were significantly altered, particularly for S320A and E339A with increases in K_m (apparent) of about

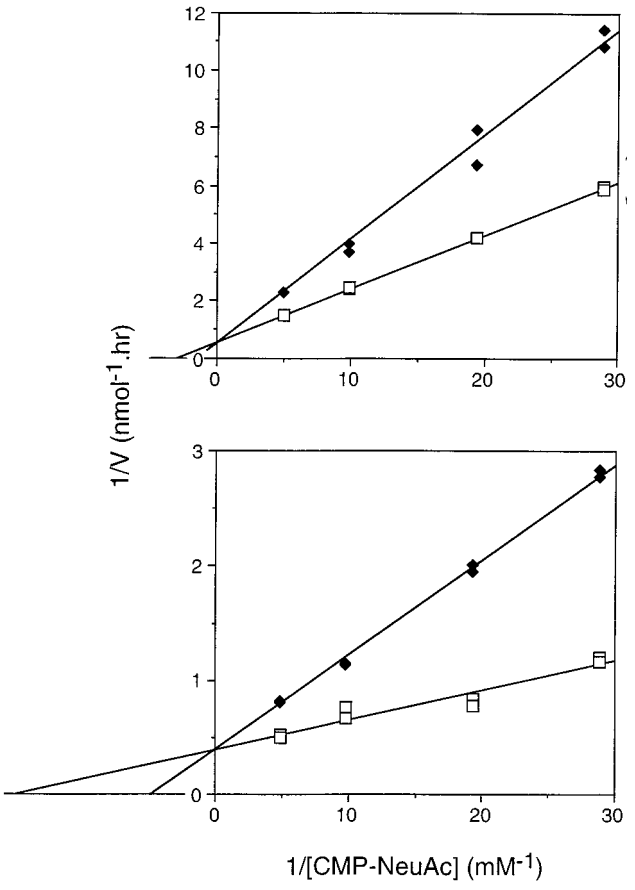

FIG. 4. **Inhibition kinetics of sialyltransferase by CDP.** Initial rate data for the wild-type enzyme (*bottom*) were determined in the presence of varied concentrations of the donor substrate CMP-NeuAc (0.025–0.2 mM) and in the absence (□) and presence (♦) of a fixed concentration of 12.5 μM CDP. Similarly, kinetics for the mutant S320A (*top*) were determined in the presence of the varied concentrations of CMP-NeuAc and in the absence (□) or presence (♦) of a fixed concentration of 16.7 μM CDP. The concentration of the acceptor asialo α_1-acid glycoprotein was kept constant at 50 μg throughout for both enzymes. The K_i values were extracted from the x intercept in the presence of inhibitor (61), where x intercept $= -(1/K_m)(1+[I]/K_i)$, K_m is obtained in the absence of the inhibitor, CDP, and [I] is the concentration of CDP.

6-fold. The K_m values for asialo α_1-acid glycoprotein were also altered by about 2-fold for these two mutants. Interestingly, Val^{335} to Ala showed increased binding affinity toward the acceptor substrate (about 5-fold decreased K_m), which remained unchanged when mutated to Leu. Thus, the interaction of the enzyme with the acceptor substrate is more favored when Ala was introduced. The mutation of Gly^{321} to Ala, on the other hand, increased the K_m values by about 3-fold for both the donor and the acceptor substrates. It is possible that this invariant amino acid is essential for the hydrogen bond formation through the interaction with the OH group of both the donor and the acceptor substrates. Therefore, even a conserved mutation of Gly to Ala affected the binding affinities for both the substrates.

Increase in K_i of CDP for S320A—Cytidine analog CDP is known to be a potent inhibitor of sialyltransferases, exhibiting equilibrium dissociation constants similar to that of CMP-NeuAc (27). Earlier, it was shown that the mutants of the L-sialylmotif, which had increased K_m values toward the CMP-NeuAc, also had increased K_i for CDP compared with that of the wild-type enzyme, as expected (16). To determine if this is also true for the mutants of S-sialylmotif, CDP was tested as a competitive inhibitor of CMP-NeuAc for mutant S320A as a representative (Fig. 4). The K_i for S320A was determined to be

TABLE II
Summary of analysis for the sialylmotif mutants of ST6Gal I

Mutants were constructed by site-directed mutagenesis for the conserved amino acids present in L- and S-sialylmotifs. For S-sialylmotif, methods for analysis have been described under "Experimental Procedures" and in the text. For L-sialylmotif, these were described earlier (16).

	Sialylmotifs	
	L[a] (# (%))	S (# (%))
Number of mutants	11	9
Proteins expressed	11	8
Detectable (>5%) enzyme activity	8	6
K_m values determined	8 (100)	5 (100)
Change of K_m values[b]		
CMP-NeuAc	6 (75)	3 (60)
acceptor	0	3 (60)

[a] Results were taken from Datta and Paulson (16).
[b] Calculated for the mutants with change in K_m values by >2-fold.

50 μM, representing about a 10-fold increase over the wild-type enzyme (4.7 μM). This increase in K_i for CDP is comparable in magnitude with the increase in K_m (which is about 6-fold) for CMP-NeuAc, suggesting that Ser^{320} participates in the binding of the common cytidine moiety in these two ligands.

Identification of a Putative Linkage-specific Sequence—-A protein sequence homology search among the cloned sialyltransferases showed that the downstream amino acid sequence of the S-sialylmotif is conserved for each subgroup of this enzyme family. For example, the sequence LYGFWPF is present in the members of the ST8Sia subfamily only. Similarly, conserved sequence could be identified for the members of the other subfamilies. This finding suggests that such a conserved sequence may confer the linkage specificity for the sialic acid transferred to the glycosyl moiety of the acceptor substrate.

DISCUSSION

The family of glycosyltransferase enzymes transfers respective sugars from its activated nucleotide-sugar donor to various glycolipids and glycoproteins. By now, a considerable number of clones have been obtained for different members of this glycosyltransferase family (see Ref. 29 for a review) that include sialyltransferases, fucosyltransferases, and galactosyltransferases. Analysis of the deduced protein sequence indicated that all of these members have common structural features: the presence of a short N-terminal cytoplasmic protein domain, which is not essential for the catalytic activity, followed by about a 20-amino acid transmembrane domain that determines the retention signal sequence for these Golgi lumenal enzymes. Apart from the stem region, the rest of the protein that resides in the lumen confers the catalytic activity. There is evidence to support such a model. For example, ST6Gal I sialyltransferase, the subject of this study, transfers sialic acid from CMP-NeuAc to the Galβ1,4GlcNAc sequence common to many Asn-linked oligosaccharides (30, 31). The analysis of its protein sequence deduced from the cDNA cloned from rat liver (15) suggests that the primary structure of this enzyme consists of a short N-terminal cytoplasmic domain (amino acids 1–9), a membrane anchor and signal sequence (amino acids 10–27), a stem region (amino acids 28–62), and a large lumenal catalytic domain (amino acids 63–403). This enzyme is found to be active even after truncation of 71 amino acids from its N-terminal that defines the cytosolic, transmembrane, and stem regions. Similarly, Macher and his colleagues (32) have obtained evidence that 20% of the N-terminal amino acid sequence of two human α1,3/4 fucosyltransferases (FucT III and V) is not required for enzyme activity, whereas truncation of the C terminus of these enzymes results in their inactivation. Although these studies showed that the lumenal C-terminal

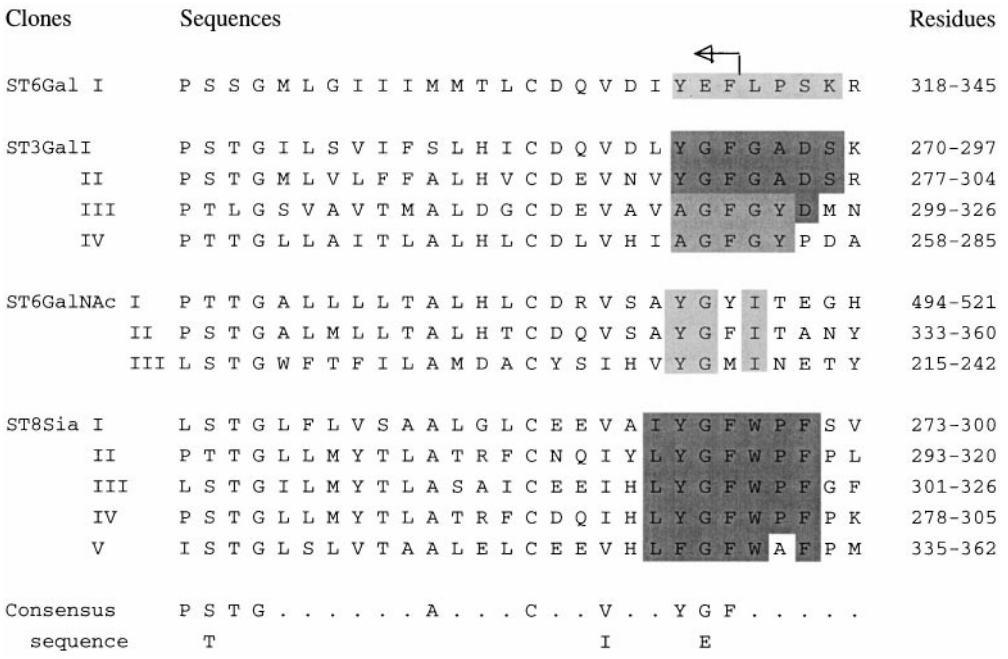

FIG. 5. **Amino acid alignment for the S-sialylmotif of the cloned sialyltransferases.** Members of the sialyltransferase enzyme family are grouped according to the linkage formed by these enzymes (see Ref. 2). The sequences are obtained from rat ST6Gal I (15), porcine ST3Gal I (3), rat ST3Gal II (9), rat ST3Gal III (4), human ST3Gal IV (10), chicken ST6GalNAc I (8), chicken ST6GalNAc II (7), rat ST6GalNAc III (13), mouse ST8Sia I (38), rat ST8Sia II (5), mouse ST8Sia III (11), hamster ST8Sia IV (41), and mouse ST8Sia V (14). The consensus sequence consisting of 10 conserved amino acids are shown. The amino acid residues shown are present in more than 75% of the cloned sialyltransferases. This consensus sequence also includes two invariant amino acids, glycine and cysteine. The restricted variations observed for the other conserved amino acids are also shown. The *arrow* indicates the C-terminal end of the previously described S-sialylmotif (17, 28).

part of these enzymes confers the catalytic activity, nothing has been known about the substrate binding sites until recently. Studies have been initiated by different groups to define the substrate binding sites of various members of this glycosyltransferase family using different techniques, including chemical and site-directed mutagenesis (32–37).

We have initiated the structure-function analysis of sialyltransferases to find out its possible substrate binding sites. This enzyme family has a unique structural feature that is not present in other members of the glycosyltransferases. Each of these enzymes shares the presence of two conserved protein domains termed sialylmotifs (see Ref. 17). Because of this conserved nature, we have studied these sialylmotifs for their possible role in the catalytic activity of these enzymes using ST6Gal I as a model. Mutants have been constructed for this purpose by site-directed mutagenesis of the conserved amino acids present in its two sialylmotifs. Earlier, our kinetic analysis showed (see Table II) that a majority of the alanine mutants for the L-sialylmotif had a significant effect on the binding affinity toward the donor substrate, indicating that the L-sialylmotif predominantly participates in binding of the common donor substrate CMP-NeuAc (16).

In the present study, similar analyses have been performed with the single point mutants for the S-sialylmotif. Out of nine mutants, five mutant enzymes showed enzyme activity sufficient for kinetic analysis. The analysis showed that the alanine mutation affected the K_m of these mutants for both the donor and acceptor substrates (Tables I and II), suggesting that the S-sialylmotif participates in binding with both substrates. This result is supported by the following observation: so far, 13 clones have been obtained for sialyltransferases with distinct substrate specificity. Considering their strict acceptor substrate specificity, these cloned sialyltransferases are subdivided into various groups or subfamilies (2). Protein sequence analysis among the members of this subfamily revealed that each of these share a common sequence near the 3′-end of the S-sialylmotif. For example, members of the ST8Sia subfamily share a conserved amino acid sequence LYGFWPF at the 3′-end of the S-sialylmotif (Fig. 5). The only variation observed in this sequence is by similar amino acid residues: I for L as in ST8Sia I and F for Y as in ST8Sia V. Nevertheless, all the members share this conserved sequence except ST8Sia V, where Pro is substituted by Ala. Each of these members utilizes CMP-NeuAc as a common donor substrate and various glycolipid and glycoprotein acceptors with restricted substrate specificity. Although ST8Sia I transfers sialic acid to the terminal sialic acid of G_{M3} to form G_{D3} (38–40) and G_{T3} (41), the other members, ST8Sia II–IV, transfer it to the terminal sialic acid of the *N*-linked glycan to form dimeric or polymeric sialyl structures (5, 11, 12, 42–46). The recently cloned ST8Sia V, on the other hand, utilizes G_{D1a} and G_{T1b} as acceptors to form G_{T1a} and G_{Q1b}, respectively (14). However, one common feature among all these members is that each of these transfers sialic acid to the C-8 position of the sialic acid of the various oligosaccharide acceptors. This sequence, LYGFWPF, is present only in the ST8Sia subfamily and not in others, indicating that this motif may be serving a common purpose unique only for these members. In the absence of any experimental data, it is difficult at this stage to understand the functional role, if any, of this conserved sequence. Nevertheless, considering the conserved nature of this sequence present only among the members of the ST8Sia subfamily, it may be proposed that this sequence may well determine the linkage specificity, *i.e.* the position C-8 on the sialic acid of the oligosaccharide acceptor, where another sialic acid is transferred from CMP-sialic acid by these enzymes. Thus, it is not surprising that, by using the reverse degenerated oligonucleotide primer designed based on the sequence at the 3′-end of the S-sialylmotif of ST8Sia subfamily, the new clones obtained were for the enzymes with α2,8-sialyltransferase activity (11, 14). Interestingly, the genomic analysis of two recently cloned members of this enzyme subfamily, ST8Sia II and ST8Sia III, showed that this

conserved region originates from a common exon (47, 48).

This characteristic feature of amino acid sequence homology near the C-terminal end of S-sialylmotif is also observed among the members of other subfamilies of sialyltransferases. For example, the sequence YGFGADS is present among the members of the ST3Gal sialyltransferases that transfer sialic acid to the C-3 position of the terminal galactosamine residue of *O*-linked glycoproteins (3, 9, 49–52). The sequence, AGFGYD, on the other hand, is present in ST3Gal III, which transfers sialic acid to the C-3 position of the terminal galactose residue of the *N*-linked glycoproteins (4, 53). It will be interesting to find out whether the acceptor substrate specificity changes from *O*-linked to *N*-linked glycoproteins by changing the Ala residue to Tyr (and *vice versa*) by site-directed mutagenesis. Nevertheless, it should be noted that the enzyme, ST3Gal IV (10, 54), which transfers sialic acid to the C-3 position of the terminal galactose unit of both *O*-linked and *N*-linked glycoproteins, has an AGFGYPD sequence instead. Similarly, the sequence YG*X*I (*X* is any amino acid) is found conserved among the members of ST6GalNAc subfamily (7, 8, 13). These enzymes transfer sialic acid from its nucleotide sugar to C-6 of GalNAc of *O*-linked glycoproteins. Although ST6GalNAc I and -II transfer sialic acid to asialo Galβ1,3GalNAc-Ser/Thr (7, 8), the ST6GalNAc III transfers it to the C-6 of GalNAc of Neu5Acα2,3Galβ1,3GalNAc-Ser/Thr (13). For the ST6Gal subfamily, only one enzyme has been obtained so far, though from different species (15, 55–59). It remains to be seen whether the sequence YEFLPSK, present in ST6Gal I, could also be found conserved among other ST6Gal enzymes (see Ref. 60 for review) once cloned. Thus, it seems that the sequence homology near the C-terminal part of the S-sialylmotif determines the linkage specificity by which the sialic acid moiety is attached to the terminal galactose, galactosamine, or sialic acid of the oligosaccharide acceptor unit of the glycoprotein or glycolipid.

As mentioned above, protein sequence analysis also showed that each of these two sialylmotifs of ST6Gal I contains one cysteine residue. These two residues are invariantly present in all the cloned sialyltransferases (17). In our previous study (16) and in the present experiment, we observed that the alanine mutation of these two conserved cysteine residues, Cys^{181} and Cys^{332}, caused the mutant ST6Gal I enzymes inactive. Our preliminary result (data not shown) also indicated that the wild-type ST6Gal I is sensitive to dithiothreitol and other reducing agents. These findings suggest that these two cysteine residues are important for the enzyme activity and are involved in disulfide linkage formation. From the fact that both the sialylmotifs participate in the binding of the donor substrate CMP-NeuAc, it seems that these two cysteine residues bring the two sialylmotifs closer together by forming an intra-disulfide linkage and, thus, form the conformation required for binding of the donor substrate. Our present result, along with the finding of a probable linkage-specific sequence near the C-terminal end of the S-sialylmotif, also suggests that this smaller motif is most likely associated with the binding site of the acceptor substrate.

Acknowledgments—We thank Dr. Eric Szoberg for critical reading of the manuscript and Dorothy Wharry for secretarial help.

REFERENCES

1. Paulson, J. C., and Colley, J. C. (1989) *J. Biol. Chem.* **264,** 17615–17618
2. Tsuji, S., Datta, A. K., and Paulson, J. C. (1996) *Glycobiology* **6,** v–vii
3. Gillespie, W., Kelm, S., and Paulson, J. C. (1992) *J. Biol. Chem.* **267,** 21004–21010
4. Wen, D. X., Livingston, B. D., Medzihradszky, K. F., Kelm, S., Burlingame, A. L., and Paulson, J. C. (1992) *J. Biol. Chem.* **267,** 21011–21019
5. Livingston, B. D., and Paulson, J. C. (1993) *J. Biol. Chem.* **268,** 11504–11507
6. Drickamer, K. (1993) *Glycobiology* **3,** 2–3
7. Kurosawa, N., Kojima, N., Inoue, M., Hamamoto, T., and Tsuji, S. (1994) *J. Biol. Chem.* **269,** 19048–19053
8. Kurosawa, N., Hamamoto, T., Lee, Y.-C., Nakaoka, T., Kojima, N., and Tsuji, S. (1994) *J. Biol. Chem.* **269,** 1402–1409
9. Lee, Y.-C., Kojima, N., Wada, E., Kurosawa, N., Nakaoka, T., Hamamoto, T., and Tsuji, S. (1994) *J. Biol. Chem.* **269,** 10028–10033
10. Kitagawa, H., and Paulson, J. C. (1994) *J. Biol. Chem.* **269,** 1394–1401
11. Yoshida, Y., Kojima, N., Kurosawa, N., Hamamoto, T., and Tsuji, S. (1995) *J. Biol. Chem.* **270,** 14628–14633
12. Yoshida, Y., Kojima, N., and Tsuji, S. (1995) *J. Biochem.* (*Tokyo*) **118,** 658–664
13. Sjoberg, E. R., Kitagawa, H., Glushka, J., van Halbeek, H., and Paulson, J. C. (1996) *J. Biol Chem.* **271,** 7450–7459
14. Kono, M., Yoshida, Y., Kojima, N., and Tsuji, S. (1996) *J. Biol. Chem.* **271,** 29366–29371
15. Weinstein, J., Lee, E. U., McEntee, K., Lai, P. H., and Paulson, J. C. (1987) *J. Biol. Chem.* **262,** 17735–17743
16. Datta, A. K., and Paulson, J. C. (1995) *J. Biol. Chem.* **270,** 1497–1500
17. Datta, A. K., and Paulson, J. C. (1996) *Indian J. Biochem. Biophys.* **34,** 157–165
18. Datta, A. K. (1995) *Nucleic Acids Res.* **23,** 4530–4531
19. Lee, E. U., Roth, J., and Paulson, J. C. (1989) *J. Biol. Chem.* **264,** 13848–13855
20. Colley, K. J., Lee, E. U., and Paulson, J. C. (1992) *J. Biol. Chem.* **267,** 7784–7793
21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
22. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467
23. Colley, K. J., Lee, E. U., Adler, B., Browne, J. K., and Paulson, J. C. (1989) *J. Biol. Chem.* **264,** 17619–17622
24. Laemmli, U. K. (1970) *Nature* **227,** 680–685
25. Vanhoop, G., Goossens, F., De Meester, I., Hendriks, D., and Scharpe, S. (1995) *FASEB J.* **9,** 736–744
26. Yoshima, H., Matsumoto, A., Mizuochi, T., Kawasaki, T., and Kobata, A. (1981) *J. Biol. Chem.* **256,** 8476–8484
27. Paulson, J. C., Beranek, W. E., and Hill, R. L. (1977) *J. Biol. Chem.* **252,** 2356–2362
28. Tsuji, S. (1996) *J. Biochem.* (*Tokyo*) **120,** 1–13
29. Field, M. C., and Wainwright, L. J. (1995) *Glycobiology* **5,** 463–472
30. Weinstein, J., de Souza-e-Silva, U., and Paulson, J. C. (1982) *J. Biol. Chem.* **257,** 13835–13844
31. Weinstein, J., de Souza-e-Silva, U., and Paulson, J. C. (1982) *J. Biol. Chem.* **257,** 13845–13853
32. Xu, Z., Vo, L., and Macher, B. A. (1996) *J. Biol. Chem.* **271,** 8818–8823
33. Yadav, S., and Brew, K. (1990) *J. Biol. Chem.* **265,** 14163–14169
34. Aoki, D., Appert, H. E., Johnson, D., Wong, S. S., and Fukuda, M. N. (1990) *EMBO J.* **9,** 3171–3178
35. Wang, Y., Wong, S. S., Fukuda, M. N., Zu, H., Lie, Z., Tang, Q., and Appert, H. E. (1994) *Biochem. Biophys. Res. Commun.* **204,** 701–709
36. Holmes, E. H., Xu, Z., Sherwood, A. L., and Macher, B. A. (1995) *J. Biol. Chem.* **270,** 8145–8151
37. Legault, D. J., Kelly, R. J., Natsuka, Y., and Lowe, J. B. (1995) *J. Biol. Chem.* **270,** 20987–20996
38. Sasaki, K., Kurata, K., Kojima, N., Kurosawa, N., Ohta, S., Hanai, N., Tsuji, S., and Nishi, T. (1994) *J. Biol. Chem.* **269,** 15950–15956
39. Nara, K., Watanabe, Y., Maruyama, K., Kasahara, K., Nagai, Y., and Sanai, Y. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 7952–7956
40. Haraguchi, M., Yamashiro, S., Yamamoto, A., Furukawa, K., Takamiya, K., Lloyd, K. O., Shiku, H., and Furukawa, K. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 10455–10459
41. Nakayama, J., Fukuda, M. N., Hirabayashi, Y., Kanamori, A., Sasaki, K., Nishi, T., and Fukuda, M. (1996) *J. Biol. Chem.* **271,** 3684–3691
42. Eckhardt, M., Muhlenhoff, M., Bethe, A., Koopman, J., Frosch, M., and Gerardy-Schahn, R. (1995) *Nature* **373,** 715–718
43. Nakayama, J., Fukuda, M., Fredette, B., Ranscht, B., and Fukuda, M. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 7031–7035
44. Kojima, N., Yoshida, Y., Kurosawa, N., Lee, Y.-C., and Tsuji, S. (1995) *FEBS Lett.* **360,** 1–4
45. Kojima, N., Yoshida, Y., and Tsuji, S. (1995) *FEBS Lett.* **373,** 119–122
46. Scheidegger, E. P., Sternberg, L. R., Roth, J., and Lowe, J. B. (1995) *J. Biol. Chem.* **270,** 22685–22688
47. Yoshida, Y., Kurosawa, N., Kanematsu, T., Kojima, N., and Tsuji, S. (1996) *J. Biol. Chem.* **271,** 30167–30173
48. Yoshida, Y., Kurosawa, N., Kanematsu, T., Taguchi, A., Arita, M., Kojima, N., and Tsuji, S. (1996) *Glycobiology* **6,** 573–580
49. Kitagawa, H., and Paulson, J. C. (1994) *J. Biol. Chem.* **269,** 17872–17878
50. Lee, Y.-C., Kurosawa, N., Hamamoto, T., Nakaoka, T., and Tsuji, S. (1993) *Eur. J. Biochem.* **216,** 377–385
51. Kurosawa, N., Hamamoto, T., Inoue, M., and Tsuji, S. (1995) *Biochim. Biophys. Acta* **1244,** 216–222
52. Kojima, N., Lee, Y.-C., Hamamoto, T., Kurosawa, N., and Tsuji, S. (1994) *Biochemistry* **33,** 5772–5776
53. Kitagawa, H., and Paulson, J. C. (1993) *Biochem. Biophys. Res. Commun.* **194,** 375–382
54. Sasaki, K., Watanabe, E., Kawashima, K., Sekine, S., Dohi, T., Oshima, M., Hanai, N., Nishi, T., and Hasegawa, M. (1993) *J. Biol. Chem.* **268,** 22782–22787
55. Kurosawa, N., Kawasaki, M., Hamamoto, T., Nakaoka, T., Lee, Y.-C., Arita, M., and Tsuji, S. (1994) *Eur. J. Biochem.* **219,** 375–381
56. Grundman, U., Nerlich, C., Rein, T., and Zettlmeissl, G. (1990) *Nucleic Acids Res.* **18,** 667
57. Bast, B. J. E. G., Zhou, L.-J., Freeman, G. J., Colley, K. J., Ernst, T. J., Munro, J. M., and Tedder, T. F. (1992) *J. Cell Biol.* **116,** 423–435

58. Stamenkovic, I., Asheim, H. C., Deggerdal, A., Blomhoff, H. K., Smeland, E., and Funderud, S. (1990) *J. Exp. Med.* **172,** 641–643
59. Hamamoto, T., Kawasaki, M., Kurosawa, N., Nakaoka, T., Lee, Y.-C., and Tsuji, S. (1993) *Bioorg. Med. Chem. Lett.* **1,** 141–145
60. Harduin-Lepers, A., Recchi, M.-A., and Delannoy, P. (1995) *Glycobiology* **5,** 741–758
61. York, J. L. (1986) in *Textbook of Biochemistry* (Devlin, T. M., ed) 2nd Ed., pp. 117–175, John Wiley & Sons, Inc., New York
62. Schmid, K. (1975) in *The Plasma Proteins* (Putnam, F. W., ed) 2nd Ed., Vol. I, pp. 183–228, Academic Press, New York