The HNK-1 carbohydrate epitope is found on many neural cell adhesion molecules. Its structure is characterized by a terminal sulfated glucuronyl acid. The glucuronyltransferases, GlcAT-P and GlcAT-S, are involved in the biosynthesis of the HNK-1 epitope, GlcAT-P as the major enzyme. We overexpressed and purified the recombinant human GlcAT-P from Escherichia coli. Analysis of its enzymatic activity showed that it catalyzed the transfer reaction for N-acetyllactosamine (Galβ1-4GlcNAc) but not lacto-N-biose (Galβ1-3GlcNAc) as an acceptor substrate. Subsequently, we determined the first x-ray crystal structures of human GlcAT-P, in the absence and presence of a donor substrate product UDP, catalytic Mn2⁺, and an acceptor substrate analogue N-acetyllactosamine (Galβ1-4GlcNAc) or an asparagine-linked biantennary nonasaccharide. The asymmetric unit contains two independent molecules. Each molecule is an α/β protein with two regions that constitute the donor and acceptor substrate binding sites. The UDP moiety of donor nucleotide sugar is recognized by conserved amino acid residues including a DXXD motif (Asp195-Asp196-Asp197). Other conserved amino acid residues interact with the terminal galactose moiety of the acceptor substrate. In addition, Val220 and Asn221, which are located on the C-terminal long loop from a neighboring molecule, and Phe224 contribute to the interaction with GlcNAc moiety. These three residues play a key role in establishing the acceptor substrate specificity.

Carbohydrate molecules on the cell surface modulate a variety of cellular functions, including cell-to-cell interactions (1, 2). The HNK-1 carbohydrate epitope, which is recognized by HNK-1 monoclonal antibodies, is found on many neural cell adhesion molecules such as NCAM (3), myelin-associated glycoprotein (4), L1 (3), transiently expressed axonal glycoprotein-1 (5), P0 (6), and also on some glycolipids (7, 8). Expression of the HNK-1 carbohydrate epitope is spatially and temporally regulated during development of the central and peripheral nervous systems (9–11). In addition, the HNK-1 carbohydrate epitope is presumed to be involved in cell-to-cell interactions such as cell adhesion (12), migration (13), and neurite extension (14).

The structure of HNK-1 carbohydrate epitope is known to be HS-O-3-Glcαβ1-3Galβ1-4GlcNAc-R, which is shared by glycolipids and glycoproteins (7, 8, 15). Because the inner structure, Galβ1-4GlcNAc, is found commonly in various glycoproteins and glycolipids, and the terminal sulfo-3-glucuronyl group is essential for both immunoreactivity with HNK-1 monoclonal antibodies (16) and for their functions (17), the terminal structure is unique and important. Therefore, in order to elucidate the functions of the HNK-1 carbohydrate epitope, it is important to characterize the enzymes in the biosynthesis pathway, which are unique to the HNK-1 epitope. The HNK-1 carbohydrate epitope is synthesized in a stepwise manner through the addition of β-1,3-linked glucuronic acid (GlcA)³ by glucuronyltransferase(s) to precursor N-acetyllactosamine followed by the addition of sulfate group by sulfotransferase(s).

It has been reported that there may be two types of glucuronyltransferases associated with the biosynthesis of the HNK-1 carbohydrate epitope in the mammalian brain (18–21). We have recently purified and cloned the HNK-1-associated glucuronyltransferase, GlcAT-P, from rat and human (22–24). More recently, we succeeded in generating mice with targeted deletion of the GlcAT-P gene. The GlcAT-P-deficient mice exhibited reduced long-term potentiation at the Schaffer collateral-CA1 synapses and defects in spatial memory formation (25).

Based on the amino acid sequence information of GlcAT-P cDNA, another HNK-1-associated glucuronyltransferase, GlcAT-S, and a proteoglycan-associated glucuronyltransferase, GlcAT-I, have been cloned and characterized (19, 26–30). These enzymes catalyze the transfer of GlcA from a donor substrate, uridine diphosphoglucuronic acid (UDP-GlcA), to a reducing terminal residue of oligosaccharide chain in the presence of a terminal acceptor substrate. The UDP moiety of donor nucleotide sugar is recognized by conserved amino acid residues including a DXXD motif (Asp195-Asp196-Asp197). Other conserved amino acid residues interact with the terminal galactose moiety of the acceptor substrate. In addition, Val220 and Asn221, which are located on the C-terminal long loop from a neighboring molecule, and Phe224 contribute to the interaction with GlcNAc moiety. These three residues play a key role in establishing the acceptor substrate specificity.
mogeneity of manganese. GlcAT-I transfers GlcA to Galβ1-3Galβ1-4Xylβ1-O-Ser in a biosynthesis pathway of proteoglycan (19, 26). The substrate binding and the reaction mechanisms of GlcAT-I have been discussed at an atomic level based on its crystal structure in complex with donor and acceptor substrate (31–33). On the other hand, GlcAT-P can transfer GlcA to Galβ1-4GlCNac-R but not to lacto-N-biose (Galβ1-3GlCNac) (23, 34). To elucidate acceptor substrate specificity of GlcAT-P, we overexpressed it in Escherichia coli and purified and characterized acceptor substrate specificity in vitro. We then solved crystal structures of GlcAT-P in four different forms: (1) apo-form, (2) with UDP-GlcA and Mn2+, (3) with UDP-GlcA, Mn2+, and an acceptor substrate analogue, N-acetyllactosamine, and (4) with UDP-GlcA, Mn2+, and a natural acceptor substrate, an asparagine-linked biantennary nonasaccharide. Based on the structural and biochemical results, we describe the molecular mechanism of substrate recognition of HNK-1 associated glycoproteins, GlcAT-P.

EXPERIMENTAL PROCEDURES

Materials—UDP-[14C]glucuronic acid (12.1 Bq/mmol) was purchased from ICN Radiochemicals. Asialo-orysocoumid (ASOR) was kindly provided by Dr. Koizumi (Kyowa Hakko Kogyo Co., Ltd.). N-acetylglucosamine was purchased from Sigma-Aldrich. Lacto-N-biose was purchased from Sigma-Aldrich. Lacto-N-neotetraose was kindly provided by Dr. Koizumi (Kyowa Hakko Kogyo Co., Ltd.).

Protein Expression and Purification—The coding region of the catalytic domain of GlcAT-P was amplified from a cloned human GlcAT-P cDNA (24) as a template by polymerase chain reaction (PCR). The amplified DNA fragment was cloned into a bacterial expression vector, pET-28a(+) (Novagen). The expressed protein contained the following sequence: Leuβ1–His6 (the amino acid positions of human GlcAT-P are shown by superscript numbers). An E. coli strain BL21(DE3) pLysS (Stratagene), which was transformed with the plasmid, was cultivated in an LB broth containing 0.4% glucose and 20 μg/ml kanamycin at 30 °C by vigorous aeration, and then induced by the addition of 1 mM isopropyl-β-D-thiogalactopyranoside. The E. coli cells were harvested by centrifugation and suspended in a buffer containing 0.15 M NaCl, 20 mM Tris-HCl, pH 8.0.

Crystallization—Crystallization conditions for the apo-form of GlcAT-P were screened using the hanging drop vapor diffusion method. The apo-form of GlcAT-P was crystallized in 2.0-μl hanging drops over 0.2-ml reservoirs containing 20% (w/v) PEG5000MME, and 0.24 M di-sodium tartrate. The crystals were grown to 0.3 × 0.1 × 0.1 mm in 2–3 days at 289 K.

Table I: Data collection statistics

| Data set | Apo | Complex 1 | Complex 2 | Complex 3 |
|----------|-----|-----------|-----------|-----------|
| Crystallographic data | | | | |
| Space group | P2₁2₁2₁ | P2₁2₁2₁ | P2₁2₁2₁ | P2₁2₁2₁ |
| Unit cell | 61.5/85.7/123.0 | 61.3/85.8/122.8 | 61.7/85.7/122.9 | 61.1/85.8/122.9 |
| α/β/γ (°) | 90.0/90.0/90.0 | 90.0/90.0/90.0 | 90.0/90.0/90.0 | 90.0/90.0/90.0 |
| Cell volume (Å³) | 61.5/85.7/123.0 | 61.3/85.8/122.8 | 61.7/85.7/122.9 | 61.1/85.8/122.9 |
| Cell volume (Å³) | 90.0/90.0/90.0 | 90.0/90.0/90.0 | 90.0/90.0/90.0 | 90.0/90.0/90.0 |
| Resolution (Å) | 40–1.85 (1.92–1.85) | 50–1.82 (1.97–1.97) | 50–1.9 (1.97–1.97) | 50–1.9 (1.97–1.97) |
| Completeness (%) | 96.6 (79.9) | 97.1 (88.4) | 98.2 (87.5) | 99.4 (96.4) |
| Rmerge (%)** | 7.5 (41.3) | 6.0 (31.9) | 4.4 (26.6) | 16.0 (5.8) |

** Complex 1 contains Mn2+, UDP, and N-acetyllactosamine.

** Complex 2 contains Mn2+, UDP.

** Complex 3 contains Mn2+, UDP, and asparagine-linked bi-antennary nonasaccharide.

Values in parentheses are for the highest resolution shell.
For crystal structure analysis, the data sets were collected at 100 K with synchrotron radiation at PF-AR-NW12 and were processed using HKL2000 (36). All crystals belong to the orthorhombic space group $P\bar{2}12121$ with similar unit cell parameters, $a = 61, \ b = 86, \ c = 123 \ \text{Å}$, and 57% solvent content. In all crystal structures, there were two monomers in an asymmetric unit.

Data collection and processing statistics of the data sets are summarized in Table I.

### Table I

| Data set         | Apo                  | Complex 1<sup>a</sup> | Complex 2<sup>b</sup> | Complex 3<sup>c</sup> |
|------------------|----------------------|-----------------------|-----------------------|-----------------------|
| Resolution (Å)   | 40–1.85              | 40–1.82               | 40–1.9                | 40–1.9                |
| No. of reflections | 54,639               | 57,093                | 51,210                | 51,364                |
| $R_{	ext{work}}/R_{	ext{free}}$ | 19.8/22.8            | 20.5/24.4             | 20.1/22.9             | 20.2/22.6             |
| Number of non-hydrogen atoms | 3954               | 3944                  | 3939                  | 3920                  |
| Protein atoms    | 50                   | 50                    | 50                    | 50                    |
| Mn<sup>2+</sup> ions | 2                   | 2                     | 2                     | 2                     |
| Acceptor sugars atoms | 52                  | 52                    | 52                    | 52                    |
| Tartrate atoms   | 10                   | 10                    | 10                    | 10                    |
| Water molecules  | 381                  | 360                   | 414                   | 326                   |
| R.m.s.d. from ideal values |
| Bond length (Å)  | 0.005                | 0.005                 | 0.005                 | 0.005                 |
| Bond angle (°)   | 1.22                 | 1.44                  | 1.47                  | 1.46                  |
| Ramachandran plot (%) |
| Most favored     | 88.3                 | 89.3                  | 87.8                  | 88.0                  |
| Additionally allowed | 10.0                | 9.3                   | 11.0                  | 10.2                  |
| Generously allowed | 1.7                 | 1.4                   | 1.3                   | 1.2                   |
| Disallowed       | 0                    | 0                     | 0                     | 0                     |
| Average B - value (Å<sup>2</sup>) |
| Protein atoms (chain A/chain B) | 22.8/23.6 | 23.4/23.0 | 26.4/25.9 | 27.7/27.0 |
| UDP atoms        | 24.7                 | 28.0                  | 28.0                  | 28.0                  |
| Mn<sup>2+</sup> ions | 18.5                | 20.0                  | 21.0                  | 21.0                  |
| Acceptor sugars atoms | 26.4                | 31.1                  | 31.1                  | 31.1                  |
| Tartrate atoms   | 23.8                 | 23.4                  | 25.2                  | 24.9                  |
| Water molecules  | 30.4                 | 27.4                  | 32.4                  | 31.4                  |

<sup>a</sup> Complex 1 contains Mn<sup>2+</sup>, UDP, and N-acetyllactosamine.

<sup>b</sup> Complex 2 contains Mn<sup>2+</sup> and UDP.

<sup>c</sup> Complex 3 contains Mn<sup>2+</sup>, UDP, and asparagine-linked bi-antennary nonasaccharide.

<sup>d</sup> $R$-factor = $\Sigma ||F_o(h)|| - |F_c(h)||/\Sigma |F_o(h)||$, and $R_{\text{free}}$ was calculated using 5% of data excluded from refinement.

---

**Fig. 1.** The dimer structure of GlcAT-P in complex with Mn<sup>2+</sup>, UDP, and N-acetyllactosamine. The overall structure of the apo-form is almost similar to the quaternary complex (r.m.s.d. of 0.38 Å for the Co atoms). A, stereo diagram of the GlcAT-P quaternary complex (top view). The secondary structures are highlighted (α-helix, blue; β-strand, green in molecule A, α-helix, red; β-strand, yellow in molecule B) and the other regions are colored in gray (molecule A, light gray; molecule B, dark gray). UDP, Mn<sup>2+</sup>, and N-acetyllactosamine molecules are shown in ball-and-stick models. B, stereo diagram of the GlcAT-P quaternary complex (side view). This panel is rotated by 90° with respect to panel A, along a horizontal axis.

---

**X-ray Data Collection and Processing**—For crystal structure analysis, the data sets were collected at 100 K with synchrotron radiation at PF-AR-NW12 and were processed using HKL2000 (36). All crystals belong to the orthorhombic space group $P\bar{2}12121$, with similar unit cell parameters, $a = 61, \ b = 86, \ c = 123 \ \text{Å}$, and 57% solvent content. In all crystal structures, there were two monomers in an asymmetric unit.

Data collection and processing statistics of the data sets are summarized in Table I.

**Structure Determination and Refinement**—The crystal structure of the apo-form of GlcAT-P was solved by the molecular replacement method using the human GlcAT-I (PDB code: 1FGG) as a search model. Molecular replacement calculations were carried out using the program MOLREP (37). The current model with 381 water molecules and one tartrate ion was refined using CNS (38) for the resolution range from 40.0 to 1.85 Å, and has an $R$-factor of 19.8% and an $R_{\text{free}}$ of 22.8. In the Ramachandran plot, 88.3% are in the most favored regions and no
residues are in the disallowed regions. The following residues have not been modeled because of weak or no associated electron density: in molecule A (chain A), residues 157–161 and in molecule B (chain B), residues 151–162. For refinement of the complexes, the apo-form models were refined first against the complex data using CNS. Substrates and co-factors were then manually built into the F_A–F_C electron density maps followed by additional rounds of refinement. The final refinement statistics are shown in Table II. Figures were produced using MOLSCRIPT (39), RASTER-3D (40), CONSCRIPT (41), and GRASP (42).

**Computer-aided Model Building**—The structure of the complete substrate-enzyme complex was calculated by adding a GlcA moiety to the crystal structure of the quaternary complex (UDP-GlcA, Mn2⁺, and N-acetyllactosamine) since the x-ray structure of the quaternary complex did not show electron density for the GlcA moiety of the UDP-GlcA. The UDP moiety in the crystal structure was modified by adding the GlcA residue at the binding cleft. Water molecules, which collide with the UDP-GlcA structure was modified by energy minimization at the binding cleft. The initial conformation of the glucuronyl group was selected at the binding cleft by searching a proper position for the SN2-type reaction with the hydroxyl group of the galactose residue of N-acetyllactosamine. Several water molecules overlapped with the newly introduced glucuronyl group were removed, and only the remaining water molecules were energy-minimized, using Discover 3 (version 98.0, Molecular Simulations Inc. San Diego, CA). After capping solvent water provided from the Assembly module of Insight II (version 2000, Molecular Simulations Inc), the complex structure was energy-minimized until the final root mean square deviation (r.m.s.d.) became less than 0.1 kcal/mol/Å. During this minimization, residues longer than 10 Å away from the substrates, UDP-GlcA and N-acetyllactosamine, all Co atoms, and the Mn2⁺ atom were fixed. All minimizations and molecular dynamics calculations were carried out under the same conditions, unless otherwise mentioned. The entire system was then covered by solvent water (20 Å thick) and these solvent water molecules were fixed in the following molecular dynamics calculations. The energy-minimized complex structure was further optimized with molecular dynamics calculations at 298 K with the cell multipole method, a distance-dependent dielectric constant, and a time step of 1 fs for 100 ps by sampling conformations every 1 ps using Discover 3. One hundred conformations were minimized in this way until the final r.m.s.d. became less than 0.1 kcal/mol/Å, and the lowest energy conformation was selected as an optimized structure.

To generate a model for lacto-N-biose in the acceptor binding site of GlcAT-P, the 1,4-glycosidic bond of N-acetyllactosamine was changed to a 1,3-glycosidic bond maintaining the position of the galactose residue in the substrate-binding cleft. Water molecules, which collide with the N-acetylglucosamine residue, were removed. Following an energy minimization of the remaining water molecules using Discover 3, the cap water molecules were added at the acceptor binding site within 20 Å from the reaction center at C1 of the glucuronic acid moiety, and the complex structure was optimized by the minimization/molecular dynamics procedures described above.

A Lewis X structure was generated by connecting α-L-fucose residue to C-3-OH of N-acetyllactosamine in the substrate-binding cleft. After removal of water molecules that collide with the fucose residue, the complex structure was optimized by the minimization/molecular dynamics procedure within the water shell described above.

**RESULTS AND DISCUSSION**

**Overall Structure of GlcAT-P**—The recombinant protein was designed as an N-terminal truncated form because the GlcAT-P is a type II membrane protein and its N terminus contains a transmembrane and a stem regions. The truncated GlcAT-P (residues 83–334) was expressed in *E. coli* and purified as described under “Experimental Procedures.” The purified pro-
tein showed a glucuronyltransferase activity, transfer of GlcA from UDP-GlcA to a donor substrate (described below). It indicates that the recombinant enzyme purified from E. coli maintains the activity of the natural enzyme (23, 34). Using the recombinant protein, we succeeded in obtaining crystals of GlcAT-P in a variety of conditions: apo, with a donor substrate and manganese, and another with additional acceptor substrate.

We first solved the structure of the apo-form of GlcAT-P by the molecular replacement method using the catalytic domain of GlcAT-I (PDB code: 1FGG) as a search model at 1.85 Å (Tables I and II). The asymmetric unit of the crystal lattice contains one protein dimer as in the GlcAT-I crystal structure (31). The asymmetric unit contains two independent molecules, related by a non-crystallographic 2-fold axis (Fig. 1). One of the two molecules, molecule B, in the asymmetric unit has a longer disordered region (residues 151–162) than the other molecule A (residues 157–161). Therefore, the structure of molecule A is described here.

The overall structure of GlcAT-P is a GT-A fold (33) with a Rossmann-like fold consisting of twelve β-strands and seven α-helices (Fig. 2). It is divided into two regions as in the case of GlcAT-I. These two regions are connected by the DXD motif (Asp195–Asp196–Asp197), which are conserved in many UDP-sugar-dependent glycosyltransferases (33). The N-terminal region (residues 83–197), referred to as a UDP-sugar binding site toward the Golgi lumen for easy access by substrates. The N-terminal transmembrane region was deleted, indicates that it is not important for the dimerization. Terayama et al. (23) reported that native GlcAT-P purified from rat brain exists as a homodimer under non-denaturing conditions. Analytical ultracentrifugation of a recombinant GlcAT-P, in which the N-terminal transmembrane region was deleted, indicates that it forms a homodimer in solution.2 Therefore, the homodimeric structure observed in our crystal is not an artifact caused by crystallographic packing but a true form of the enzyme. As shown in Fig. 1, the N termini of both molecules are located on the opposite side of the substrate binding site. Since GlcAT-P is a type II membrane protein, the N-terminal transmembrane region anchors into the membrane thus presenting the active site toward the Golgi lumen for easy access by substrates.

Complexes between GlcAT-P and Substrates—Next, the structures of GlcAT-P in complexes with UDP and Mn2+ with and without N-acetyllactosamine were solved by the molecular replacement method using the apo-form structure of GlcAT-P as a search model at 1.9 and 1.82 Å, respectively (Tables I and II and Fig. 3). The overall structures of the complexes are very similar with that of the apo-form (data not shown). Despite numerous attempts to soak the apo-form crystals into a solution containing UDP-GlcA, no electron density of GlcA moiety was observed. Two possibilities may account for this. The first possibility is that the UDP-GlcA molecule bound at the catalytic site is hydrolyzed by the enzyme, and the GlcA moiety is released during or after the soaking. The second is that the GlcA moiety is highly flexible and not observed.

We compared the three-dimensional structures of GlcAT-P complexed with UDP, Mn2+, and acceptor substrate analogue (N-acetyllactosamine) and GlcAT-I complexed with UDP, Mn2+, and acceptor substrate analogue (Galβ1–3Galβ1–4Xyl). The overall structure of GlcAT-P reveals significant similarity to that of GlcAT-I (Fig. 4, A and B). In particular, the N terminus of the donor substrate binding region (residues 84–148) shows a strong similarity to that of GlcAT-I with an r.m.s.d. of 0.40 Å for the Ca atoms. But in the other parts of the molecule, there are some differences between GlcAT-P and GlcAT-I structures. In the GlcAT-I structure, there is a long insertion loop between α3 and β4 (Fig. 5). The other part of GlcAT-P (residues 179–322) superimposes with an r.m.s.d. of 1.00 Å. While the overall structures are similar between GlcAT-P and GlcAT-I, their electrostatic surface potential landscapes around the catalytic site are in stark contrast: one rather basic and the other acidic (Fig. 4, C and D). This may reflect the difference in specificity for acceptor molecules, i.e.

---

2 Y. Kobayashi and Y. Nishi, personal communication.
both glycoproteins and glycolipids for GlcAT-P whereas only glycoproteins for GlcAT-I.

Donor Substrate Binding Site—In the complex crystals, the Mn²⁺ and UDP molecules bind to the N-terminal region. Their binding causes side-chain reorientations in the basic residues forming the donor binding site to accommodate UDP and manganese. As shown in Fig. 6A and Table III, the uridine moiety of UDP is recognized by Pro⁹¹, Asp¹²², Lys¹⁵³, and Asp¹⁹⁶ through hydrogen bonds. The side chain of Tyr⁹₃ is involved in the stacking interaction with the uridine base ring about 3.5 Å
The phosphoric acid moieties of UDP are recognized by Arg165, Asp 197, His 311, and Arg 313 through hydrogen bonds. The Mn$^{2+}$ molecule, which interacts with Asp 197 (O$^{1}$ and O$^{2}$), UDP (O2A and O3B), and two water molecules, forms an approximately octahedral coordination (Fig. 3C). In comparison with the apo-form, side chains of the three basic residues (Lys153, Arg165, and Arg313) undergo significant conformational changes upon UDP binding (Fig. 6B).

All these residues involved in the donor substrate binding are conserved among GlcAT-P, GlcAT-S, and GlcAT-I not only on their primary sequences (Fig. 5) but also on their stereochemical dispositions (data not shown). The DXD motifs, which are found in many UDP-sugar-dependent glycosyltransferases, are thought to play an important role in recognition of nucleotide-sugar and Mn$^{2+}$ ions. This is also the case with GlcAT-P: Asp195, Asp196, Asp197. These suggest that GlcAT-P has the molecular scaffold to recognize its donor substrate similar to those of other glycosyltransferases. The crystallographic data is consistent with the previous systematic mutational analysis (43). For example, none of the mutants, Asp$^{122}$, Asp$^{195}$, Asp$^{196}$, Asp$^{197}$, or His$^{311}$ to alanine abolished glucuronyltransferase activity (43). In addition, two amino acid residues (Val$^{220}$ and Asn$^{321}$) on the C-terminal long loop from the neighboring molecule in the homodimer are involved in the interaction.

Catalytic Mechanism—As described above, the overall structures of GlcAT-P and GlcAT-I show strong similarity especially in the N-terminal donor substrate binding region. Although there is no electron density corresponding to GlcA in our quaternary complex, the computer-aided model based on energy minimization and molecular dynamics calculations as described under “Experimental Procedures” shows that the GlcA moiety could fit to the binding cleft without any stereochemical clashes (Fig. 8A). In this model structure, the O2 atom of GlcA interacts with His$^{311}$ (N$^{62}$) of GlcAT-P with a hydrogen bond. The same interaction was observed between GlcA and His$^{306}$ of GlcAT-I, which corresponds to His$^{311}$ of GlcAT-P (32, 33). It has been proposed that the histidine residue play an important role to distinguish GlcA from other sugars (33). Since this histidine residue is found not only in GlcAT-P and GlcAT-I but also in

![Fig. 6. Stereoview of the active site of GlcAT-P. A, residues involved in the interactions between GlcAT-P, acceptor, donor, and Mn$^{2+}$ are shown. Mn$^{2+}$ ion is colored orange. Metal interactions are shown in solid light blue lines. Hydrogen bonds are represented as light blue dashed lines. The superscript N denotes residues of the neighboring molecule B (shown in dark gray). UDP and N-acetyllactosamine molecules are shown in orange red ball-and-stick models colored according to atom types (nitrogen, blue; carbon, black; oxygen, red; phosphorous, purple). B, comparison between the apo-form and the quaternary complex with UDP, Mn$^{2+}$, and N-acetyllactosamine. The apo-form is colored green (molecule A) or light gray (molecule B). The quaternary complex is colored yellow (molecule A) or dark gray (molecule B).](image-url)
GlcAT-S, the architecture for recognition of GlcA might be conserved among the enzymes, which catalyze UDP-GlcA as a donor substrate.

In the GlcAT-I structure in complex with UDP, Mn$^{2+}$ and an acceptor sugar substrate analogue (Galβ1-3Galβ1-4Xyl), Oe2 of Glu$^{281}$, which corresponds to Glu$^{284}$ in GlcAT-P, is located 2.66 Å from the nucleophilic O3 atom of the 3-hydroxyl group of galactose moiety (31). It has been demonstrated that the glutamate residue plays a key role as the catalytic base in the glycosyltransfer reaction in GlcAT-I (31). In the GlcAT-P quaternary complex structure, Oe2 of Glu$^{284}$ is also located 2.66 Å from the nucleophilic O3 atom of the 3-hydroxyl group of galactose moiety. Therefore, Glu$^{284}$ of GlcAT-P might play a key role as the catalytic base in the glycosyltransfer reaction. This possibility is confirmed by the observation that GlcAT-P mutant, E284A, loses its enzymatic activity (43). Thus the catalytic mechanism of glycosyltransfer from the donor sugar-nucleotide to the acceptor might be the same as that proposed in GlcAT-I (31–33). In this scheme, the glutamate residue acts as a catalytic base that deprotonates the 3-hydroxyl of the terminal galactose of the acceptor substrate to form a strong nucleophile. Then the nucleophile attacks the C1 carbon of the GlcA of the UDP-GlcA donor and forms an oxo-carbenium ion-like transition state. Formation of a GlcAβ1-3Gal linkage and a subsequent dissociation of UDP follow.

**Acceptor Substrate Specificity**—In order to study the substrate specificity, the glucuronyltransferase activity of recombinant GlcAT-P toward various acceptor substrates was measured. The specific activities for ASOR (Galβ1-4GlcNAc-R), N-acetyllactosamine (Galβ1-4GlcNAc), and lacto-N-neotetraose (Galβ1-4GlcNAcβ1-3Galβ1-4Glc) were calculated to be 1,680, 180, and 160 nmol/min/mg, respectively. On the other hand, GlcAT-P did not catalyze the transfer reaction for lacto-N-biose (Galβ1-3GlcNAc) as an acceptor substrate. This substrate specificity is consistent with the previous study, in which the glucuronyltransferase activity of the native GlcAT-P was inhibited by N-acetyllactosamine but not by lacto-N-neotetraose (23). We also measured the activity of recombinant GlcAT-S. It shows activities not only for ASOR (2,150 nmol/min/mg), N-acetyllactosamine (270 nmol/min/mg), and lacto-N-neotetraose (260 nmol/min/mg) but also for lacto-N-biose (100 nmol/min/mg).

These observations are confirmed by the other experiments where protein A-fused GlcAT-P shows the activity for N-acetyllactosamine and lacto-N-neotetraose but not for lacto-N-
biose, whereas protein A-fused GlcAT-S shows the activities for all of them. Another glucuronyltransferase, GlcAT-I, transfers GlcA to Galβ1–3Galβ1–4Xylβ1-O-Ser but not to ASOR (44).

A comparison of their amino acid sequences and the crystal structures turns out to be extremely useful to clarify the acceptor substrate specificities of glucuronyltransferases. Many amino acid residues (especially the acidic residues; Glu228, Asp254, and Glu284 in GlcAT-P), which have been shown to interact with acceptor sugar substrate both in GlcAT-P and GlcAT-I crystals, are conserved among GlcAT-P, GlcAT-S, and GlcAT-I (Fig. 5). However, there are some differences among them. First, our crystal structure of the GlcAT-P quaternary complex shows that the C-terminal long loop from the neighbor molecule serves as the key in the recognition of N-acetyllactosamine. Two amino acid residues, Val320 and Asn321 of the C-terminal long loop from the neighboring molecule, interact with the GlcNAc residue through a hydrophobic interaction and a hydrogen bond, respectively (Fig. 6A). Val320 is found only in GlcAT-P, but not in GlcAT-S nor GlcAT-I (Fig. 5). Asn321 is conserved in GlcAT-P and GlcAT-S, but not in GlcAT-I (Fig. 5). We propose that this C-terminal long loop is critical for the acceptor specificity.

Second, Phe245 is important for the interaction with acceptor substrate, N-acetyllactosamine, in our crystal structure. As described above, there is the stacking interaction between the side chain of Phe245 and the GlcNAc moiety of N-acetyllactosamine (Fig. 6A). GlcAT-S and GlcAT-I have tryptophan at the position corresponding to Phe245 of GlcAT-P (Fig. 5). Although a similar aromatic interaction was found between

---

**Fig. 8. Computer-aided substrate-enzyme complex models.** A, the model of GlcAT-P in complex with UDP-GlcA, Mn2+, and N-acetyllactosamine is shown. Since there is no electron density corresponding to GlcA in our quaternary complex, the UDP moiety in the crystal structure was modified by adding the GlcA moiety at the β-phosphate oxygen, and the UDP-GlcA moiety was modified by energy minimization at the binding cleft. The UDP-GlcA and N-acetyllactosamine are shown in orange red except oxygen, phosphorous, and nitrogen atoms (colored in red, purple, and blue, respectively). B, the GlcAT-P complex model with lacto-N-biose as an acceptor substrate is superimposed to the crystal structure of the quaternary complex with N-acetyllactosamine. The N-acetyllactosamine and the lacto-N-biose are shown in orange red and green, respectively. C, the GlcAT-P complex model with Lewis X as an acceptor substrate is superimposed to the crystal structure of the quaternary complex with N-acetyllactosamine. The N-acetyllactosamine and the Lewis X are shown in orange red and green, respectively.

---

**Fig. 9. A**, the structure of bi-antennary nonasaccharide used for cocrystallization is shown. Fmoc is a protecting group for chemical synthesis. B, the omit Fo–Fc electron density map of the bi-antennary nonasaccharide contoured at 1.75 σ, superimposed with a ball-and-stick model colored according to atom types (nitrogen, blue; carbon, black; oxygen, red). The gray chain is the neighboring molecule in the asymmetric unit. The residues that involve interaction with nonasaccharide are shown in ball-and-stick models colored according to atom types.
Trp^{243} of GlcAT-I and the galactose ring of the acceptor substrate, both the aromatic ring of Trp^{243} and the galactose ring are tilted by 30 degrees keeping their parallelism as compared with those in GlcAT-P (Fig. 7). Similar aromatic ring structures are frequently found in other carbohydrate enzymes and lectins (45). The importance of the aromatic interaction between Phe^{245} and acceptor substrates has been shown by the mutational study (43); a F245A mutant lost its activity completely, whereas F245Y retained about 43% of the wild-type activity. These observations suggest that the aromatic ring plays important roles in the acceptor substrate recognition. These differences in the C-terminal residues of the neighboring molecule, stacking between an aromatic residue and galactose ring account for the differences in the specificity of the substrate acceptor recognition.

To elucidate how GlcAT-P has a preference in recognizing β1–4 linkage (N-acetyllactosamine) to β1–3 (lacto-N-biose), computer-aided model building of the substrate-enzyme complexes was carried out. The model structure of the GlcAT-P complex with lacto-N-biose was built by energy minimization and molecular dynamics calculations as described under “Experimental Procedures.” These results show that the aromatic ring of Phe^{245} stacks to the GlcNAc ring of lacto-N-biose much more poorly than to that of N-acetyllactosamine (Fig. 8B). In addition, the signature interaction between the acceptor substrate and Val^{320} or Asn^{321} from the neighboring molecule is lost in the lacto-N-biose-bound model. These differences support the notion that Phe^{245}, Val^{320}, and Asn^{321} play important roles in distinguishing two different glycosidic linkages, β1–4 and β1–3.

In addition, it has been demonstrated that protein A-fused GlcAT-P shows a very low activity for Lewis X (Galβ1–4(Fucα1–3)GlcNAc) as an acceptor, 0.4% of that for N-acetyllactosamine. 8 It is interesting to note that Lewis X has the same reducing terminal structure Galβ1–4GlcNAc as N-acetyllactosamine. Why is Lewis X such a poor acceptor? To shed light on this specificity difference, we first constructed a simple model in which the bound N-acetyllactosamine is replaced by Lewis X, based on the GlcAT-P quaternary complex structure. In this model, the fucose moiety of Lewis X stereochemically clashes with Gly^{223}, Gl^{224}, and the lactosamine moiety of Lewis X itself. To see if this can be alleviated, we then employed the same method as described above to generate a new model structure of the complex between GlcAT-P and Lewis X. The modified complex structure shows several conformational changes of Lewis X (Fig. 8C). First, the stacking interaction between Phe^{245} and GlcNAc ring is lost. Second, the interaction between the acceptor substrate and Val^{320} or Asn^{321} from neighbor molecule is also lost. In short, Phe^{245}, Val^{320}, and Asn^{321} again are responsible for the low glycosyltransferase activity toward Lewis X.

**Recognition of Branched Oligosaccharide Substrate—**The structure of HNK-1 carbohydrate epitope on P0 glycoprotein, which is thought to be synthesized by GlcAT-P, is shown to be a bi-antennary oligosaccharide chain with its terminal of HSO3−3GlcAβ1–3Galβ1–4GlcNAc (15). It has been reported that native GlcAT-P purified from rat brain transfers GlcA to N-acetyllactosamine branches of bi-, tri-, and tetra-antennary oligosaccharide chains with different efficiencies (34). Protein A-fused GlcAT-P overexpressed and purified from E. coli shows a tendency to prefer branched oligosaccharides as an acceptor substrate in comparison with those that are unbranched. 9 To investigate the molecular interaction between GlcAT-P and a branched oligosaccharide from a structural viewpoint, we tried to determine the complex structure of GlcAT-P with UDP-GlcA, Mn^{2+}, and an asparagine-linked bi-antennary nonasaccharide (Fig. 9A). The crystal structure is almost completely the same to the quaternary complex of GlcAT-P, UDP-GlcA, Mn^{2+}, and N-acetyllactosamine. Although additional electron density was observed adjacent to the GlcNAc moiety, the volume was too small to fit a whole asparagine-linked nonasaccharide (Fig. 9B). The electron density is supposed to be the third manose of either of the two reducing termini since the bi-antennary nonasaccharide has the same structure on its branched chains (Fig. 9A). Unfortunately, the lack of electron density for the remaining part of the nonasaccharide prevents an unequivocal assignment of the Gal-GlcNAc-Man density to either of the two. One of the possible reasons for the lack of electron density is that the residues after the GlcNAc moiety is exposed to the solvent and thus fluctuates. The other possibility is that the interactions between GlcAT-P and the remaining part of nonasaccharide is inherently weak and is difficult to detect in the present crystal packing. Such interactions may be further elucidated by future studies of an improved crystal structure and an NMR chemical shift analysis.

**CONCLUSIONS**

The overall crystal structure of HNK-1-associated glucuronyltransferase, human GlcAT-P, is similar to that of GlcAT-I. The donor substrate, UDP, is recognized by a number of amino acid residues, which are conserved among GlcAT-P, GlcAT-S, and GlcAT-I. The active center, Glc^{284}, is also structurally conserved. Thus we conclude that the catalytic mechanism of glycosyltransfer from a donor sugar-nucleotide to an acceptor is the same as that proposed for GlcAT-I.

On the other hand, the acceptor substrate specificity of GlcAT-P is different from the other enzymes. GlcAT-P transfers GlcA to Galβ1–4GlcNAc but not to Galβ1–3GlcNAc whereas GlcAT-S transfers to both acceptor substrates. The crystal structure of the GlcAT-P quaternary complex and computer-aided model building explain this specificity. The unique combination of the aromatic interaction of Phe^{245}, the hydrophobic interaction of Val^{320}, and the hydrogen bond of Asn^{321} are important to determine the acceptor substrate specificity.

**Accession Numbers—**Coordinates have been deposited in the Protein Data Bank (accession codes 1V82, 1V83, and 1V84 for the apo-form of human GlcAT-P, the complex with Mn^{2+}, UDP, and N-acetyllactosamine and the complex with Mn^{2+} and UDP, respectively).

**REFERENCES**

1. Rutishauser, U., Acheson, A., Hall, A. K., Mann, D. M., and Sunshine, J. (1988) Science 240, 53–57
2. Jessell, T. M., Hynes, M. A., and Dodd, J. (1990) Annu. Rev. Neurosci. 13, 227–255
3. Kruse, J., Mailhammer, R., Wernecke, H., Faissner, A., Sommer, I., Goridis, C., and Schachner, M. (1984) Nature 311, 153–155
4. McGarry, R. C., Helfand, S. L., Quarles, R. H., and Roder, J. C. (1983) Nature 306, 576–578
5. Dodd, J., Morton, S. B., Karagogeos, D., Yamamoto, M., and Jessell, T. M. (1988) Neuron 1, 105–116
6. Bollens, E., and Schachner, M. (1987) Neurosci. Lett. 82, 77–82
7. Chou, D. K., Ilyas, A. A., Evans, J. E., Castello, C., Quarles, R. H., and Jungalwala, F. B. (1986) J. Biol. Chem. 261, 1717–1725
8. Ariya, T., Kobiyama, T., Fredro, L., Latov, N., Saito, M., Kon, A., Ando, S., Suzuki, M., Hemling, M. E., Rinehart, K. L., Kusunoki, S., and Yu, R. K. (1987) J. Biol. Chem. 262, 848–853
9. Brenner-Frazer, M. (1986) Dev. Biol. 115, 44–55
10. Schwarting, G. A., Jungalwala, F. B., Chou, D. K., Boyer, A. M., and Yamamoto, M. (1987) Dev. Biol. 120, 65–76
11. Yoshihara, Y., Oka, S., Watanabe, Y., and Mori, K. (1991) J. Cell Biol. 115, 731–744
12. Keilhauer, G., Faissner, A., and Schachner, M. (1986) Nature 326, 728–730
13. Brenner-Frazer, M. (1987) Dev. Biol. 123, 321–331
14. Martin, R., Xin, Y., Schmitz, B., and Schachner, M. (1992) Eur. J. Neurosci. 4, 628–639
15. Voshol, H., van Zuylen, C. W. E. M., Orberger, G., Vliegenthart, J. F. G., and Quarles, R. H. (1988) J. Neurochem. 55, 594–601
16. Schmitz, B., Schachner, M., Ito, Y., Nakano, T., and Ogawa, T. (1994) Glycoconjug. J. 11, 345–352
17. Chou, D. K. H., Flores, S., and Jungalwala, F. B. (1991) J. Biol. Chem. 266,
17941–17947
19. Das, K. K., Basu, M., Basu, S., Chou, D. K. H., and Jungalwala, F. B. (1991) J. Biol. Chem. 266, 5238–5243
20. Kawashima, C., Terayama, K., Ii, M., Oka, S., and Kawasaki, T. (1992) Glycoconjug. J. 9, 307–314
21. Oka, S., Terayama, K., Kawashima, C., and Kawasaki, T. (1992) J. Biol. Chem. 267, 22711–22714
22. Terayama, K., Oka, S., Seiki, T., Miki, Y., Nakamura, A., Kosutsumi, Y., Takio, K., and Kawasaki, T. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 6093–6098
23. Terayama, K., Seiki, T., Nakamura, A., Matsumori, K., Ohta, S., Oka, S., Sugita, M., and Kawasaki, T. (1988) J. Biol. Chem. 263, 30295–30300
24. Mitsumoto, Y., Oka, S., Sakuma, H., Inazawa, J., and Kawasaki, T. (2000) Genomics 15, 166–173
25. Yamamoto, S., Oka, S., Inoue, M., Manabe, T., Takahashi, H., Miyamoto, M., Asano, M., Sakagami, J., Sudo, K., Iwakura, Y., Ono, K., and Kawasaki, T. (2000) J. Biol. Chem. 277, 27227–27231
26. Tone, Y., Kitagawa, H., Imiya, K., Oka, S., Kawasaki, T., and Sugahara, K. (1999) FEBS Lett. 459, 415–420
27. Seiki, T., Oka, S., Terayama, K., Imiya, K., and Kawasaki, T. (1999) Biochem. Biophys. Res. Commun. 255, 182–187
28. Kitagawa, H., Tsuoka, M., Tone, Y., and Sugahara, K. (2001) Biochem. J. 358, 539–546
29. Imiya, K., Ishizaki, T., Seiki, T., Saito, F., Inazawa, J., Oka, S., and Kawasaki, T. (2002) Gene (Amst.) 286, 29–36
30. Marroz, I., Galan, J. J., Berrego, S., and Antinolo, G. (2002) J. Hum. Genet. 47, 677–680
31. Pedersen, L. C., Tsuchida, K, Kitagawa, H, Sugahara, K, Darden, T. A., and Negishi, M. (2000) J. Biol. Chem. 275, 34580–34585
32. Pedersen, L. C., Darden, T. A., and Negishi, M. (2002) J. Biol. Chem. 277, 21869–21873
33. Negishi, M., Dong, J., Darden, T. A., Pedersen, L. G., and Pedersen, L. C. (2003) Biochem. Biophys. Res. Commun. 303, 393–398
34. Oka, S., Terayama, K., Imiya, K., Yamamoto, S., Kendo, A., Kato, I., and Kawasaki, T. (2000) Glycoconj. J. 17, 877–885
35. Kajihara, Y., Suzuki, Y., Sasaki, K., and Juneja, L. R. (2003) Methods Enzymol. 362, 44–64
36. Owczarek, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
37. Vagin, A., and Teplyakov, A. (1997) J. Appl. Crystallogr. 30, 1022–1025
38. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kazueki, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D. 54, 905–921
39. Kraulis, P. J. (1991) J. Appl. Crystallogr. 24, 946–950
40. Merritt, E. A., and Bacon, D. J. (1997) Methods Enzymol. 277, 505–524
41. Lawrence, M. C., and Bourke, P. (2000) J. Appl. Crystallogr. 33, 980–991
42. Nicholls, A., Sharp, K. A., and Honig, B. (1991) Proteins Struct. Funct. Genet. 11, 281–296
43. Ohtsubo, K., Imajo, S., Ishiguro, M., Nakatani, T., Oka, S., and Kawasaki, T. (2000) J. Biochem. 128, 283–291
44. Kitagawa, H., Tone, Y., Tamura, J., Neumann, K. W., Ogawa, T., Oka, S., Kawasaki, T., and Sugahara, R. (1998) J. Biol. Chem. 20, 6615–6618
45. Quiocho, F. A. (1986) Annu. Rev. Biochem. 55, 287–315