J. Biochem. 128, 283–291 (2000)

Studies on the Structure–Function Relationship of the HNK-1 Associated Glucuronyltransferase, GlcAT-P, by Computer Modeling and Site-Directed Mutagenesis[1]

Kazuaki Ohtsubo,[*] Seiichi Imajo,[†] Masamichi Ishiguro,[‡] Takako Nakatani,[*] Shogo Oka,[*] and Toshisuke Kawasaki[*,2]

[*]Department of Biological Chemistry and CREST (Core Research for Educational Science and Technology) Project, Japan Science and Technology Corporation, Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto 606-8501; and Suntory Institutes for [†]Biomedical Research and for [‡]Bioorganic Research, Shimamoto, Osaka 618-0024

Received May 12, 2000; accepted May 30, 2000

All members of a glucuronyltransferase (GlcAT) gene family cloned to date contain four conserved regions (modules I–IV), which are widely located in the catalytic domain. In order to understand the biological significance of these modules, we investigated the structure–function relationship of GlcAT-P by means of the combination of site-directed mutagenesis and computer aided three-dimensional modeling. The wild-type and mutant GlcAT-Ps were expressed in *Escherichia coli* as glutathione-S-transferase (GST)–fused soluble proteins. Most of the mutants in which a polar amino acid within the modules was replaced with alanine lost their transferase activity almost completely, while all of the mutants in which the replacement was outside these modules retained the original catalytic activity. A three-dimensional (3-D) model of GlcAT-P was constructed by computer simulation with the three-dimensional structure of adenylate kinase (1AKE) as a template. This model predicted that the large catalytic domain of GlcAT-P forms a globular shape with a Rossmann-fold motif consisting of five α-helix and β-sheet repeats. The putative catalytic pocket consisting mainly of modules I–III is surrounded by a cluster of polar amino acids, which are essential for the transferase activity and also for the binding to the acceptor substrate (essential amino acids), asialo-orosomucoid. There is the second cluster of essential amino acids almost on the opposite surface of the molecule, in which an aspartic acid repeat (DDD) is located. The biological significance of the second cluster is currently not clear but it may be associated with the interaction of the enzyme with modulation molecules, manganese and membrane phospholipids.

Key words: alanine-scan, glucuronyltransferase, HNK-1 carbohydrate epitope, Rossmann-fold motif, 3D-model.

A prominent feature of the biosynthesis of a carbohydrate chain is template-independent synthesis. In eukaryotic cells, a large variety of sugar chains are expressed on cell surface glycoproteins and glycolipids, and a large number of specific glycosyltransferases and their corresponding genes are involved in the biosynthesis of these sugar chains. Most glycosyltransferases have similar topologies; they are classified as type II transmembrane proteins with a short N-terminal region, which is a putative cytoplasmic tail, a hydrophobic transmembrane domain and a long C-terminal catalytic domain (*1*). However, the molecular structural basis for each specific glycosyltransferase reaction is not well understood.

Currently, a glucuronyltransferase (GlcAT-P), which is a key enzyme in the biosynthesis of the HNK-1 carbohydrate epitope, HSO_3-3GlcAβ1-3Galβ1-4GlcNAc-R, has been isolated and cloned from rat brain (*2–4*), and several homologues of it have been found (*5*, *6*). The epitope is found on neural cell adhesion molecules such as N-CAM, L1, MAG, and P0, as well as on a certain type of glycolipid in the nervous system (*7–10*). The expression of the HNK-1 carbohydrate epitope is spatially and temporally regulated during development of the nervous system (*11*, *12*), and is associated with neural crest cell migration (*13*), neuron and glial cell adhesion (*14*), outgrowth of astrocytic processes, and migration of cell bodies (*15*). GlcAT-P catalyzes the transfer of glucuronic acid (GlcA) from the high-energy donor, uridine diphosphoglucuronic acid (UDP-GlcA), to a terminal

[1] This work was supported in part by a Grant-in-Aid for Scientific Research (A-09307053) and a Grant-in-Aid for Scientific Research (C-11680604) from the Japan Society for the Promotion of Sciences, and a Grant-in-Aid for Scientific Research on Priority Areas (A-10178104) from the Ministry of Education, Science, Sports and Culture of Japan.

[2] To whom correspondence should be addressed. Phone: +81-75-753-4572, Fax: +81-75-753-4605, E-mail: kawasaki@pharm.kyoto-u.ac.jp

Abbreviations: GlcAT, glucuronyltransferase; PCR, polymerase chain reaction; ASOR, asialo-orosomucoid; GST, glutathione-S-transferase.

galactose residue of *N*-acetyllactosamine structure on glycoproteins in the presence of manganese and sphingomyelin, which act as cofactors (*4*). The deduced amino acid sequence of GlcAT-P predicted its type II transmembrane topology as in the cases of many other glycosyltransferases. Sequence alignment of GlcAT-P with its homologues (*5*, *6*), and putative proteins in *C. elegans* (*16*) and *S. mansoni* (*17*) indicated the presence of four highly conserved regions in the catalytic domain, named modules I–IV (*3*), which are presumed to be required for the function of GlcAT-P.

In this study, we investigated the structure-function relationship of GlcAT-P by means of the combination of site-directed mutagensis and computer 3D-modeling. The results indicated that the catalytic domain of GlcAT-P has a globular shape with a Rossmann-fold motif.

MATERIALS AND METHODS

Construction of GlcAT-P Expression Vectors—To construct GST-tagged GlcAT-P expression vectors (pGEX/GlcAT-P), the DNA fragment of GlcAT-P was amplified by PCR using a sense primer, LA-Bam-S, ATAGGATCCGACATCGTGGAGGTGGTGCGCACA (5′-3′; nt 223–246, the nucleotide numbering refers to Fig. 1B in Ref. *3*), and an antisense primer, LA-Eco-A, CCGGAATTCTGTGTAGTTTCAGATCTCCACCGAG (5′-3′; nt 1029–1053), using a full length GlcAT-P in pEF-BOS as a template (*3*). The amplified fragment (nt 223–1053) was digested with *Bam*HI and *Eco*RI, followed by insertion into pGEX-4T-1 (Pharmacia Biotech), which had been digested with the same enzymes.

For the construction of single amino acid substituted mutants of GlcAT-P, sequential PCRs were carried out as described by Ho *et al.* (*18*). In the first PCR, LA-Bam-S and one of the antisense internal mutagenic oligonucleotides listed below were used as the primer pair. The mutagenic antisense oligonucleotides (5′-3′) used (the amino acid numbering refers to Fig. 4) were:

for T121A, GGGCACATGCAATAGCGCGTTGGCCATTCG (nt 349–378);
for E134A, CCTACGTGGAGCATCCGCCACCACCAGCCA (nt 388–417);
for D135A, CGTCCTACGTGGAGCAGCCTCCACCACCAG (nt 391-420);
for R148A, GTTGAGGCCAGTGTCGGCCAGCAGGCGCGC (nt 430–459);
for R175A, GGTGCCACGTGGGATGGCAGGGTCTCGGGC (nt 511–540);
for R196A, CGGCTGAGTGGAGTTCGCTGGGAAGGTCTC (nt 574–603);
for T199A, CACTACACCCGGCTGAGCGGAGTTCCGTGG (nt 583–612);
for D208A, GTACGTGTTGTCGTCAGCTGCGAAGTACAC (nt 610–639);
for D209A, TGTACGTGTTGTCGGCATCTGCGAAGTACA (nt 611–640);
for D210A, CAGACTGTACGTGTTGGCGTCATCTGCGAA (nt 616–645);
for N211A, CTCCAGACTGTACGTGGCGTCGTCATCTGC (nt 619–648);
for E220A, TCTTGTGCTGCGCATCGCTTCAAAGAGCTC (nt 646–675);
for R222A, CACCCTTCTTGTGCTGGCCATCTCTTCAAA (nt 652–681);
for R222K, CCTTCTTGTGCTCTTCATCTCTTCAAAGAG (nt 649–678);
for R226A, AGGCCACACGGACACCGCTCTTGTGCTGCG (nt 664–693);
for S228C, GGCCACAGGCCACACGCACACCCTTCTTGT (nt 670–699);
for R244A, CCCTGCCCCATTCACCGCGGGGGCCTCATA (nt 718–747);
for F258A, TGGCCGGTGGGGGTCGGCGACTGTCTTCCA (nt 760–789);
for F258Y, TGGCCGGTGGGGGTCGTAGACTGTCTTCCA (nt 760–789);
for H261A, TATTGCAAATGGCCGGGCGGGGTCGAAGAC (nt 769–798);
for R262A, GTCTATTGCAAATGGCGCGTGGGGGTCGAA (nt 772–801);
for D267A, ATCCAGCCATGGCTATTGCAAATGGCCGGT (nt 782–811);
for D267E, AAATCCAGCCATCTCTATTGCAAATGGCCG (nt 784–813);
for D267N, AAATCCAGCCATGTTTATTGCAAATGGCCG (nt 784–813);
for K292A, TTCCTGGTAGCCTCCTGCTACCCCACGTAG (nt 862–891);
for E297A, TCGAAGGAGACTGCTTGCCTGGTAGCCTCC (nt 877–906);
for S298A, TTCTCGAAGGAGACTGGCTTCCTGGTAGCC (nt 880-909;
for E311A, GTTTGCTGCCTTCCCTGCTAGATCATTGAG (nt 919-948);
for C317S, AGGATCTTGGTAGAGTTTGCTGCCTTGGGC (nt 933–962);
for K319A, ATGCCAGACCAGGATCGCGGTACAGTTTGC (nt 943–972);
for H323A, TCTGTTCGCGTAGCCCAGACCAGGATCTTG (nt 954–983);
for T324A, TGGCTTCTCTGTTCGCGCATGCCAGACCAG (nt 961–990);
for R325A, CACTGGCTTCTCTGTTGCCGTATGCCAGAC (nt 964–993);
for T326A, CAGCACTGGCTTCTCTGCTCGCGTATGCCA (nt 967–996).

In the second PCR, the respective sense internal mutagenic oligonucleotide (complementary to the antisense internal mutagenic oligonucleotide) and LA-Eco-A were used as the primer pair. pGEX/GlcAT-P was used as the template for both of these reactions. In the final PCR, aliquots of both PCR products from the above reactions were used as the template, LA-Bam-S and LA-Eco-A being used as the primer pair. The final PCR products were digested with *Bam*HI and *Eco*RI restriction enzymes, and then inserted into pGEX-4T-1 previously digested with the same enzymes.

Expression and Purification of the Recombinant GlcATs—The recombinant GST-fusion proteins were expressed in *E. coli* BL21 (TOYOBO, Japan). Growth of the bacteria, induction of expression, homogenization of the cells, binding of the GST-fusion proteins to glutathione-Sepharose (Pharmacia), washing and elution of the bound material were carried out according to the manufacturer's instructions.

Glucuronyltransferase Assay—The transferase activity of the enzymes was assayed as described previously (*4*) with

slight modification. An equivalent amount of each enzyme was incubated at 37°C for 3 h in a reaction mixture, with a final volume of 50 μl, containing 100 mM MES (pH 6.5), 2.5 mM ATP, 0.2% NP-40, 20 mM $MnCl_2$, 20 μg ASOR, 100 μM UDP-[^{14}C]-GlcA (200,000 dpm), and 2 μl of 2% NP-40 extract of rat forebrain, which had been treated at 100°C for 3 min. After incubation, the assay mixture was spotted onto a 2.5-cm Whatman No. 1 disc. The disc was washed with a 10% (w/v) trichloroacetic acid solution three times, followed by with ethanol/ether (2:1, v/v) and then ether. The disc was air-dried and the radioactivity of [^{14}C]GlcA-ASOR on it was counted with a liquid scintillation counter (Beckman LS-6000). The inhibition assay was performed in the presence of 10 mM UDP, *N*-acetyllactosamine and *N*-ethylmaleimide.

ASOR Binding Assay—The binding activity toward acceptor substrate, ASOR, was measured by ELISA. EIA/RIA plates (Costar, USA) were coated with 0.1 μg/well ASOR in PBS overnight at 4°C and then equivalent amounts of the GST-tagged GlcAT-Ps were applied to the wells in the presence of 100 mM MES (pH 6.5) and 20 mM $MnCl_2$. The enzyme bound to the wells was detected *via* the GST-tag using a rabbit anti-GST antibody (Upstate Biotechnology) and HRP-conjugated anti–rabbit IgG (ZYMED, USA), and visualized by means of the HRP chromogenic reaction with *o*-phenylenediamine (WAKO, Osaka). The absorbance at 490 nm was measured after 2 h incubation.

Computer Aided Modeling of GlcAT-P—Homology with the sequence of GlcAT-P was searched for by Web site Network Protein Sequence Analysis (http://pbil.ibcp.fr/NPSA/npsa_server.html) of a sequence database of 3D-structure-known proteins in NRL_3D (sequences of proteins in the Brookhaven protein database) with SSearch. A 3D-structure model of GlcAT-P was built by the homology modeling method, using the adenylate kinase (1AKE) structure in PDB (*19*). The structure model was then optimized with Discover 3 (Molecular Simulations, CA, USA) as described previously (*20, 21*). A UDP-GlcA-*N*-acetyllactosamine-GlcAT-P ternary complex was built by the distance geometry method (DGEOM) (*22*). The complex structure was then optimized by means of the molecular dynamics/minimization procedure at 298 K for 100 ps.

RESULTS

Preparation and Characterization of a GST-GlcAT-P Fusion Protein—In order to investigate the structure-function relationship of GlcAT-P, we constructed a glutathione-*S*-transferase (GST)–tagged GlcAT-P expression vector, which expresses a truncated form of GlcAT-P lacking the NH_2-terminal 74 amino acids, which include the putative cytoplasmic tail and transmembrane domain. The prepared fusion protein (GST-GlcAT-P) was assayed for glucuronyltransferase activity using asialo-orosomucoid (ASOR) as an acceptor substrate. The soluble fusion protein exhibited specific activity of 200 nmol/min/mg, which corresponds to approximately 5% of that of the natural GlcAT-P, which had been isolated from rat brain (4,300 nmol/min/mg). As shown in Fig. 1, the transferase activity of the fusion protein was effectively inhibited by the addition of 10 mM UDP or 10 mM *N*-acetyllactosamine, and was completely abolished by treatment with *N*-ethylmaleimide, which alkylates the thiol residue of Cys-317 (*3*). In addition, GST-GlcAT-P required its specific cofactor, sphingomyelin, for expression of its activity toward a glycoprotein substrate (data not shown). These results are consistent with the characteristics of the native enzyme, indicating that GST-GlcAT-P is a reasonable model for studying the structure–function relationship of natural GlcAT-P.

Effects of Alanine-Substitutions on the Transferase Activity—GlcAT-P has four highly conserved regions in the catalytic domain (modules I–IV) (Fig. 2), which were initially found on sequence alignment with putative proteins in *C. elegans* and *S. mansoni* (*3*). In order to determine how these modules are associated with the function of GlcAT-P, we prepared a series of substitution mutants of the GST-GlcAT-P fusion protein (Fig. 3A), in which amino acids with a polar side chain were substituted with alanine. The amounts of the GST-GlcAT-P and its mutants were estimated by immunoblot analysis with rabbit anti-GST antibodies and HRP-conjugated anti–rabbit IgG, and equivalent amounts of the fusion proteins were used to assay its catalytic activity. As shown in Fig. 3B, most of the substitutions within these modules: T121A, E134A, and D135A (in module I), D208A, D209A, D210A, N211A, E220A, and R222A (in module II), R262A, D267A, E297A, and S298A (in module III), and E311A, K319A, H323A, and T324A (in module IV), dramatically decreased the transferase activity (less than 5% of that of the wild type GST-GlcAT-P), while those in non-conserved regions: R148A, R175A, R196A, and T199A (between module I and module II) and R244A (between module II and module III), did not significantly affect the transferase activity (38 to 100% of that of the wild type GST-GlcAT-P). These findings indicated that the amino acid residues required for catalytic activity were widely distributed in these four modules in the C-terminal catalytic domain and that these modules shared by all GlcATs must be involved in the function of GlcAT-P.

Construction of a 3D-Model of GlcAT-P—There has been no available 3D-protein structure explicitly homologous to that of GlcAT-P. Recently, the crystal structure of a galactosyltransferase has been reported (*23*). However, no significant structural relation found was as to the sequence homology between them. On the other hand, a homology search of the 3D-known-structure database suggested a low

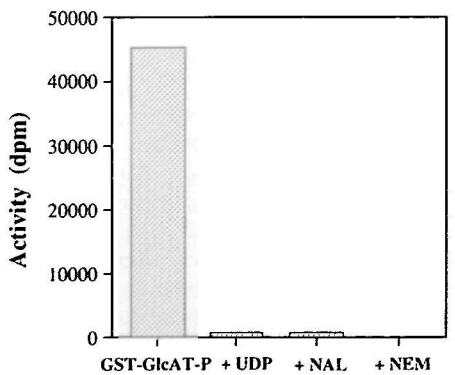

Fig. 1. **The enzymic properties of GST-tagged GlcAT-P.** The assay was carried out as described under "MATERIALS AND METHODS" using 2 μg of GST-GlcAT-P in the presence or the absence of an inhibitor: UDP, NAL (*N*-acetyllactosamine), and NEM (*N*-ethylmaleimide).

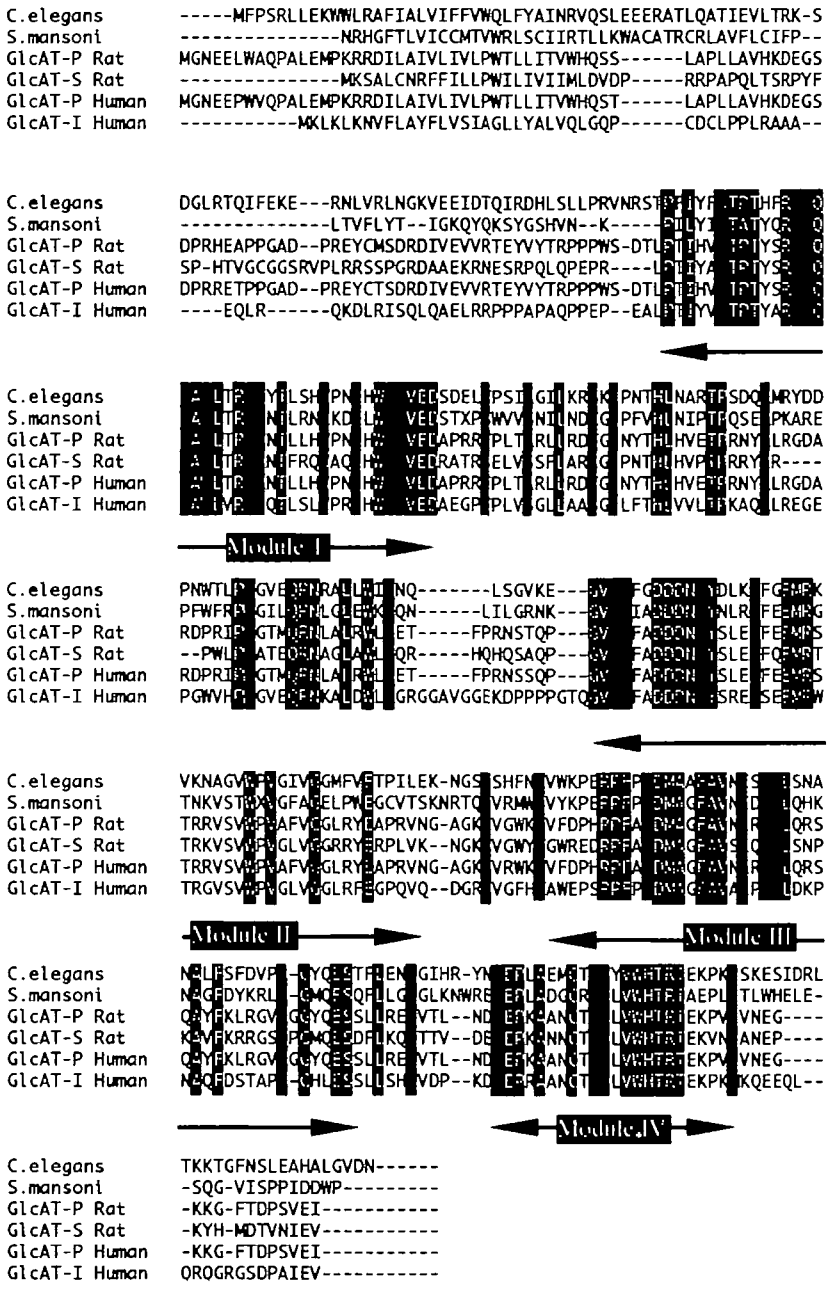

Fig. 2. **Sequence alignment of GlcAT family proteins.** The following sequences are compared: putative proteins of *C. elegans* (*16*), and *S. mansoni* (*17*), rat GlcAT-P (*3*), rat GlcAT-S (*6*), human GlcAT-P (*31*), and human GlcAT-I (*5*). Black and meshed backgrounds indicate completely conserved amino acids and homologous substitution amino acids, respectively. Dashes indicate gaps introduced for maximal alignment. Arrows indicate the locations of modules I–IV.

but significant (about 20% homology) structural relation to adenylate kinases and DNA polymerases, which have nucleotide-binding domains, so-called Rossmann-fold. Of these proteins, adenylate cyclase was the most suitable protein as a template since several residues important for nucleotide binding were conserved in the alignment. Figure 4 illustrates amino acid sequence alignment for GlcAT-P with four adenylate kinase–related proteins.

Adenylate kinase–related proteins show multiple conformations depending on the binding of nucleotides and/or phosphate-accepting substrates (*18*). An adenylate kinase structure, 1AKE, which binds an inhibitor, Ap5A, at the substrate binding site was selected for modeling of the complex structure of GlcAT-P with UDP-glucuronic acid and *N*-acetyllactosamine, since the binding site of the adenylate kinase is occupied by a nucleotide and phosphate-accepting substrate mimic (Ap5A). Provided that UDP occupies the ADP-binding site in 1AKE, and that the 3-hydroxyl group of galactose reacts at C1 of glucuronic acid in an SN2 fashion (inversion at C1) in a mode found in the sulfate transferase (*24*). The resulting model showed that the carboxylate moiety of glucuronic acid bound to Arg226 or Arg196. Asp267 was located close to the 3-hydroxyl group of galactose, suggesting its participation in the activation of the hydroxyl group. Glucosamine was proximal to Phe258 in a manner frequently found in a sugar-aromatic interaction (*25*). The Arg222 residue, which participates in the binding of nucleoside, is well conserved in the ternary complex model. Figure 5 depicts a ribbon diagram of the GlcAT-P model, and shows a Rossmann-fold motif consisting of β1

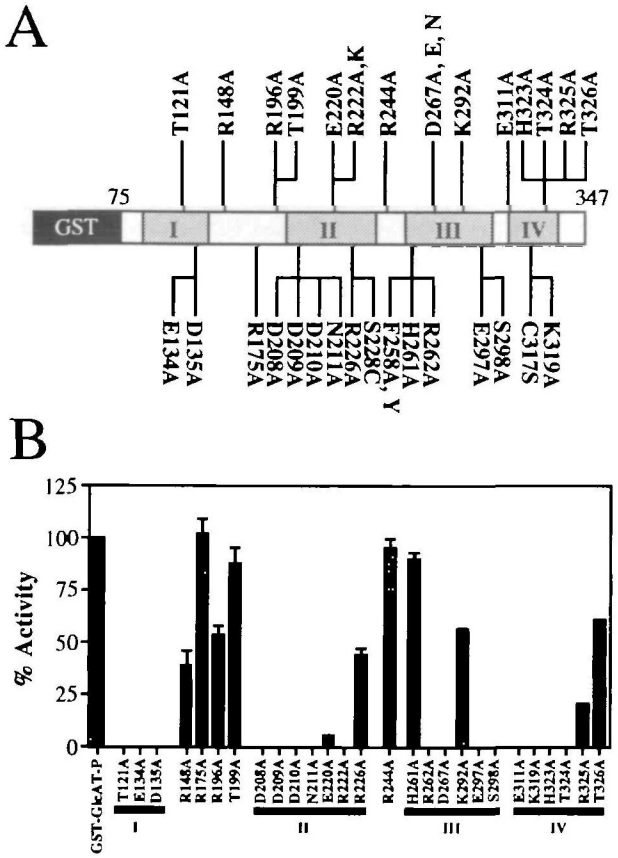

Fig. 3. **(A) Schematic diagram of all the substitution mutants of GST-GlcAT-P.** The chimeric protein of GlcAT-P is fused with GST at Asp75. Modules I–IV are indicated as meshed boxes. **(B) Transferase activity of alanine substitution mutants.** Glucuronyltransferase activity was assayed as described under "MATERIALS AND METHODS" using an equivalent amount of each mutant. Values are the means for at least three experiments, each performed in triplicate, ± SEM.

(113–118), α1 (124–137), β2 (139–141), α4 (172–184), β3 (189–192), α5 (197–205), β4 (210–216), α8 (277–288), β9 (294–300), and α9 (303-316).

Prediction of a Putative Catalytic Pocket by Means of the Combination of Alanine-Scan and the 3D-Modeling—In Fig. 6, each module of GlcAT-P is colored differently, module I in yellow, module II in light blue, module III in magenta, and module IV in green, and the locations of the amino acids essential for the activity are colored in red and those of nonessential amino acids in dark blue. Module I consists of a β-sheet (β1) and an α-helix (α1) in the N-terminal portion, a large proportion of module II consists of α-helices (α5 and 6) and β-sheets (β4, 5, and 6), module III consists mostly of α-helices (α7 and 8) and a β-sheet (β9), and the N-terminal portion of module IV consists of an α-helix (α9). The four modules are involved in the construction of the Rossmann-fold motif and fold together to give a globular shape, on the upper part of which a putative catalytic site is present. This deduced catalytic pocket is surrounded by a cluster of essential amino acids for the transferase activity revealed by the alanine scan: Thr121, Glu134, Asp135, Arg222, Phe258, Arg262, and Asp267. Non-essential amino acids, Arg148, Arg175, Arg196, Arg199, Arg226, Arg244, His261, and Lys292, are localized in the external portion or on the outer surface of the molecule. It is interesting that there is another cluster of essential amino acids at the bottom of the molecule, which consists of Asp208, Asp209, Asp210, Asn211, Glu297, Ser298, and Glu311. The location of this cluster is almost on the opposite side of the globular molecule. The biological significance of this second cluster is discussed below (see "DISCUSSION").

A close-up view of the catalytic pocket is shown in Fig. 7. According to this model, UDP-GlcA is held at this position by the hydrogen-bonding contact between the hydroxyl group of the side chain of Ser228 and the carboxyl oxygen of the glucuronic acid, and by the interaction between the uracyl ring and the side chain of Arg222. The *N*-acetyllactosamine is held in position by the π–π stacking interaction between the side chain of Phe258 and the hexose ring of *N*-acetylglucosamine. The putative catalytic center is Asp267, whose carboxyl oxygen induces deprotonation of the O_3 oxygen atom of the galactose moiety, and then the deprotonated O_3 oxygen attacks the C_1 carbon atom of glucuronic acid.

The deduced model of GlcAT-P predicted that amino acids Arg222, Ser228, Phe258, and Asp267 are directly associated with the catalytic function of the enzyme. To confirm this, we prepared a series of conservative substitution mutants of GlcAT-P. This type of substitution is suitable for the fine diagnosis of suspected amino acids, because the chemical changes in the chain length or modifications of the functional groups are much smaller than those caused by alanine substitutions. As shown in Fig. 8A, the R222K, S228C, D267E, and D267N mutants did not exhibit any glucuronyltransferase activity, while the F258Y mutant retained about 43% of the wild type activity. Additionally, the F258A mutant lost the catalytic activity completely, confirming the important role of the aromatic structure for the catalytic activity.

Effects of Mutations on the Acceptor Substrate Binding—In addition to the transferase activity assay, we designed and performed an acceptor substrate binding assay, which enabled us to directly measure the binding activity of the mutant proteins toward the *N*-acetyllactosamine structure on ASOR (ASOR binding assay). As shown in Fig. 8B, the F258Y mutant retained about 25% of the original ASOR binding activity, this being consistent with the partial retention of the catalytic activity (Fig. 8A). On the other hand, the F258A mutant, which did not show any catalytic activity, lost its binding activity toward ASOR as well. These results suggest that the aromatic ring plays an important role in the specific binding to the acceptor substrate. A similar role of the aromatic ring structure is frequently found in the carbohydrate recognition by enzymes and lectins (*25*). The R222K and S228C mutants retained binding activities toward ASOR (22 and 12%, respectively) comparable to that of the F258Y mutant, but these two mutants had essentially no catalytic activity. These results may suggest that not only the decrease in the binding activity toward ASOR but also the decreases in the binding activity toward UDP-GlcA was due to the complete loss of the trasnferase activity of the enzyme. The substitution mutations in the catalytic pocket, D267E and D267N, abolished the ASOR binding activity. This may suggest that the distance between the carboxyl oxygen of Asp267 and the 3-

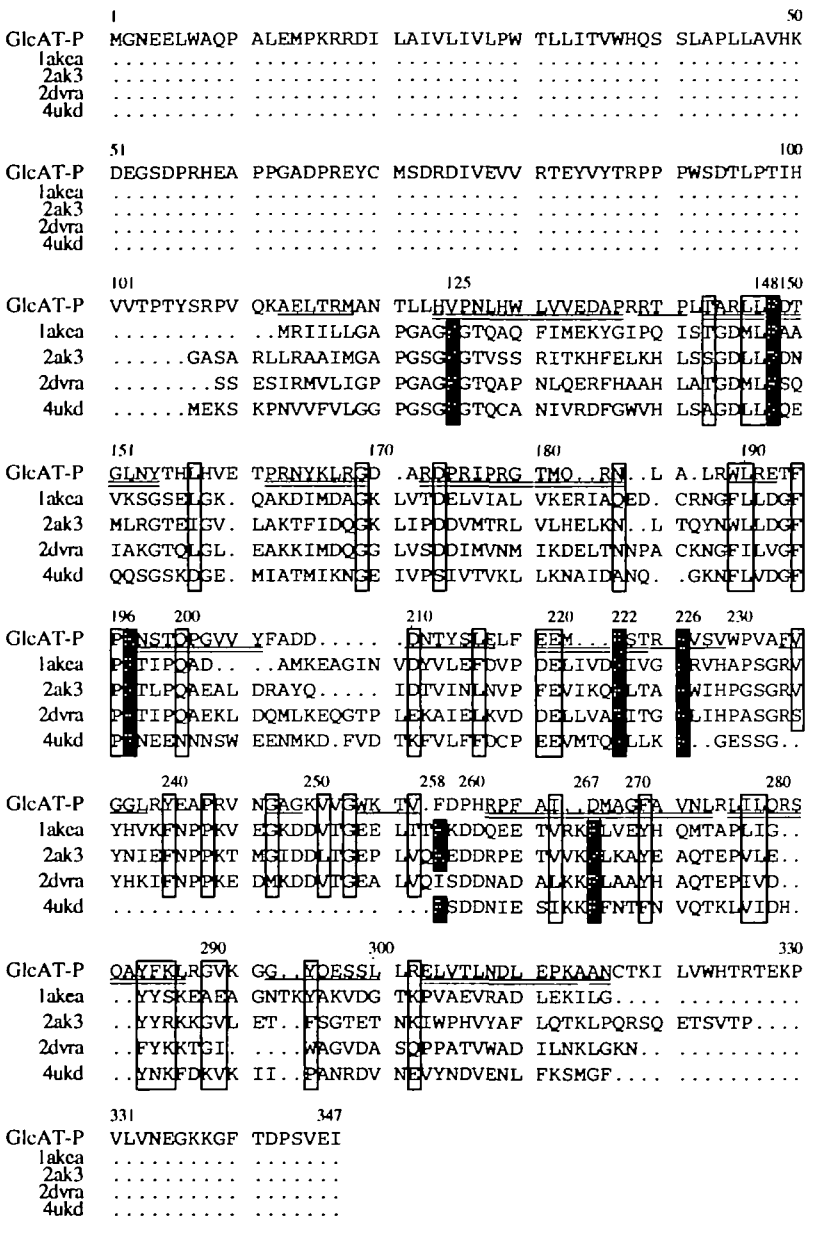

Fig. 4. **Homology alignment between GlcAT-P and four adenylate kinase-related proteins.** GlcAT-P (D88035): HNK-1 associated UDP-glucuronyltransferase (rat); 1akea: Adenylate kinase complex with AP5A (*Escherichia coli*); 2ak3: Adenylate kinase isoenzyme-3 (GTP: AMP phosphotransferase) (bovine); 2dvra: Myokinase (ATP:AMP-phosphotransferase) (mutation: D89V, R165I) (*Saccharomyces cerevisiae*); and 4ukd: UMP/CMP kinase complex with ADP and UDP (*Slime mold*). Black background: basic amino acids directly interacting with Ap5A in 1AKE; enclosures: common residues among the five protein sequences. The single and double underlines indicate predicted β-sheet and α-helix structures, respectively.

hydroxy group of the galactose is critical for the ASOR binding and the catalytic activity. The C317S mutant lost not only the binding activity toward ASOR but also the transferase activity. The structural significance of Cys317 is not clear at the moment, but it might be associated with formation of the catalytic pocket or with the undisclosed function of the second essential amino acid cluster on the opposite side of the molecule. The ASOR binding assay was carried out for all the alanine substitution mutants indicated above, the results being consistent with those of the trasferase activity test as described above (data not shown), suggesting that most of the polar amino acids within the four modules contribute either directly through the interaction with ASOR or indirectly through the proper folding of the catalytic pocket, while those outside these modules do not contribute significantly to the binding of acceptor substrates.

DISCUSSION

The modules (I–IV) were initially found on amino acid sequence alignment between GlcAT-P and homologous putative proteins of *C. elegans* and *S. mansoni* (*3*). No similarity of these modules to other glycosyltransferase families has been found. Based on the premise that these modules are restricted to the GlcAT gene family, several GlcAT genes have been successfully cloned using a hybridization approach or a PCR-based cloning approach (*5*, *6*). These facts suggest that these modules are derived from a common gene and that these modules participate in a structural feature or functions common to all GlcATs.

In this study, we investigated the structure-function relationship of GlcAT-P by means of the combination of site-directed mutagenesis and computer aided 3D-modeling of GlcAT-P. Alanine-scan of highly conserved amino acids in

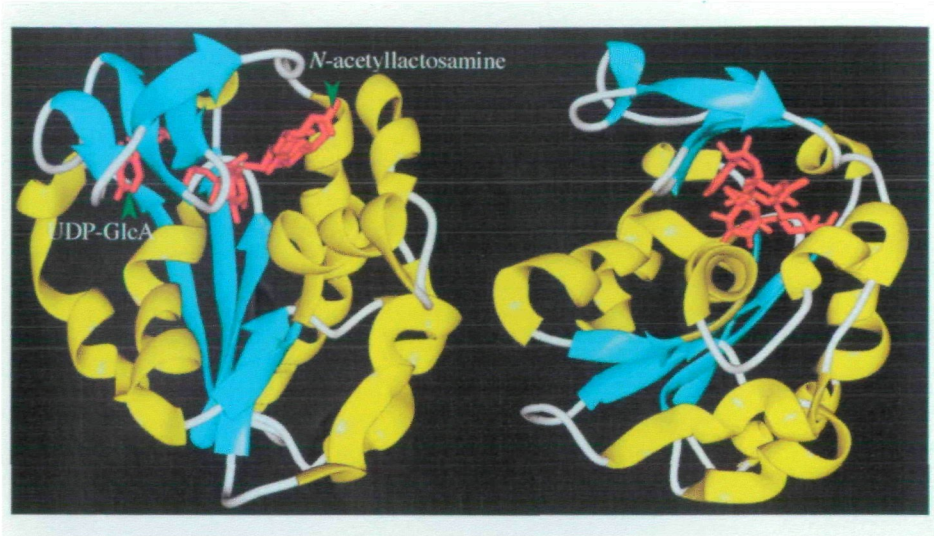

Fig. 5. **Ribbon diagrams of GlcAT-P.** Predicted overall view of the complex of GlcAT-P with its substrates is shown. GlcAT-P has a Rossmann-fold motif. α-Heliices and β-sheets are colored yellow and light blue, respectively. Substrates are colored red.

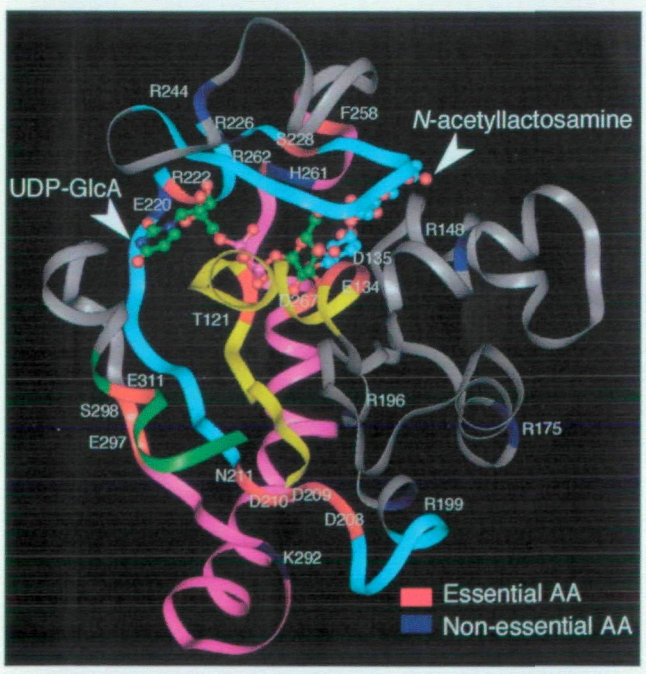

Fig. 6. **Location of essential amino acids and non-essential amino acids in the 3D model.** Modules are color-coded: module I in yellow, module II in light blue, module III in magenta, and module IV in green. A substitution which dramatically reduced the catalytic activity (essential amino acid) is colored red, and one which did not change the catalytic activity significantly (non-essential amino acid) is colored dark blue.

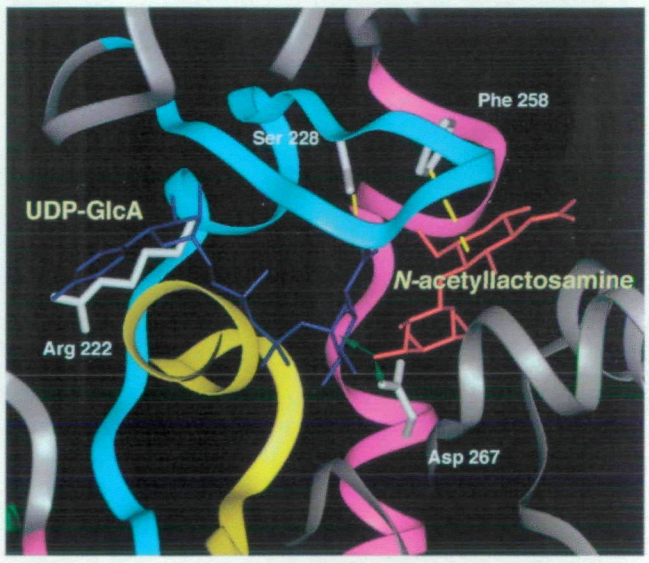

Fig. 7. **A close-up view of the catalytic pocket of GlcAT-P.** Residues participating in the association with substrates are shown. UDP-GlcA and *N*-acetyllactosamine are colored blue and red, respectively. Modules are colored as described in Fig. 6. Some important interactions are indicated by yellow lines. The proposed catalytic mechanism is indicated by green arrows.

these modules indicated that essential amino acids for the catalytic activity were intensively distributed in these modules (Fig. 3). The 3D-modeling of GlcAT-P provided clues as to how these amino acids are positioned and function. The four modules form a putative catalytic pocket, around which the amino acids most probably associated with the catalytic activity, and binding to dono and acceptor substrates are located (Figs. 6 and 7). Similar observations were made for the conserved module in a series of sialyltransferases (*26*, *27*). The sialylmotif, which is a signature module in the sialyltransferase gene family (*28*), is positioned in the center of each molecule and participates in the binding to commonly used donor substrate CMP-NeuAc.

It should be noted that a cluster of essential amino acids for the catalytic activity as well as the binding activity toward ASOR is present at the bottom of the model, which is far from the catalytic pocket (Fig. 6). The functional significance of these amino acids is not clear at the moment, but they somehow participate in maintenance of the active conformation of the enzyme. In this region, an aspartic acid cluster, Asp208-Asp209-Asp210, is present. Many glycosyltransferase have an acidic amino acid cluster(s), DDD or EDD, in common (*29*), and this acidic amino acid cluster has been considered to be involved in the common functional features of glycosyltransferases, such as in the binding to a divalent cation. Since the transferase activity of

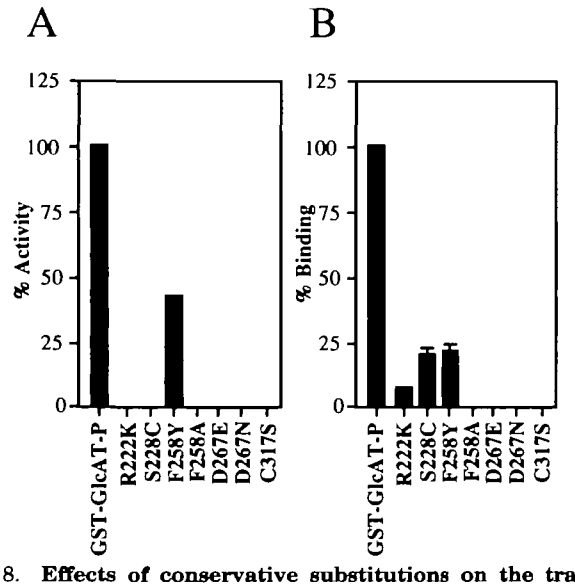

Fig. 8. **Effects of conservative substitutions on the transferase activity and ASOR binding activity of GlcAT-P.** (A) Catalytic activity of conservative substitution mutants. The glucuronyltransferase assay was carried out as described under "MATERIALS AND METHODS" using an equivalent amount of each mutant. (B) Acceptor substrate binding activity of conservative substitution mutants. Values are the means of at least three experiments, each performed in triplicate, ± SEM.

GlcAT-P depends on Mn^{2+} as a divalent cation, the aspartic acid cluster may be involved in the binding to Mn^{2+}. On the other hand, the substrate specificity of GlcAT-P is regulated by membrane phospholipids (*4*, *30*). Sphingomyelin is required for expression of the glucuronyltransferase activity toward a glycoprotein acceptor not only by the native GlcAT-P but also by GSH-GlcAT-P, while phosphatidylinositol is required for the activity toward glycolipid acceptors (Sato, Y., Seki, T., Oka, S., and Kawasaki, T., unpublished results). These results suggest that GlcAT-P contains the binding site(s) for regulatory phospholipids and it is possible to speculate that the aspartic acid cluster might be the site for this interaction. In any case, further study is required to determine the functional significance of the second cluster of essential amino acids clear.

It has previously been shown that Cys317 should play an important role in the transferase reaction (*3*). As shown in Fig. 2, all members of the GlcAT family have conserved Cys at the position equivalent to the Cys of GlcAT-P. In order to determine the function of Cys317, we performed site-directed mutagenesis for Cys317. As shown in Fig. 8, A and B, substitution mutant C317S, which has a conservative substitution of the thiol group with a hydroxyl group at the side chain of Cys317, exhibits neither transferase activity nor binding activity toward ASOR. These results confirm that the thiol group of Cys317 is essential for the binding to ASOR and the catalytic activity of GlcAT-P. In the present 3D-model of GlcAT-P, Cys317 is located just outside of the predictable structure, but its three-dimensional location has not been determined. Cys317 might be associated with the formation of the catalytic pocket or with some undisclosed functions of the second essential amino acid cluster.

Polar amino acids located outside the four modules or on the exterior of the globular molecule do not seem to contribute significantly to the formation of the catalytic pocket of GlcAT-P, since not only the catalytic activity but also the binding activity toward an acceptor substrate was not affected by their substitution with alanine ones. This was further confirmed by constructing chimeric enzymes of GlcAT-P and GlcAT-I. GlcAT-I is involved in biosynthesis of the linkage region of glycosaminoglycans (*5*), but not in that of the HNK-1epitope, but it has the four common modules and the overall structure can be expected to be similar to that of GlcAT-P. However, GlcAT-I is characteristic in containing an extra eight amino acids in the intervening region between module I and module II (see Fig. 3). A chimeric form of GlcAT-P, which contains the GlcAT-I intervening sequence, retained one-third of the original GlcAT-P activity, while chimeric GlcAT-I containing the GlcAT-P intervening sequence did not exhibit any GlcAT-P activity, (Ohtsubo, K., Nakatani, T., Tone, Y., Kitagawa, H., Sugahara, K., Oka, S., and Kawasaki, T., unpublished results), indicating little association of this intervening sequence with the enzyme specificity. These results are consistent with the predicted three-dimensional location of this loop (corresponding to a region covering R148 and R175 in Fig. 6), which is far away from the catalytic pocket, supporting the high validity of the proposed model.

In conclusion, the catalytic domain of GlcAT-P shows a novel topology with a widely opened pocket lined with mostly invariant residues forming the donor and acceptor binding sites in close proximity to the catalytic center. This structure might provide useful structural information for predicting the structures of other related glycosyltransferases.

We would like to thank Ms Y. Nagao for her secretarial assistance.

REFERENCES

1. Paulson, J.C. and Colley, K.J. (1989) Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation. *J. Biol. Chem.* **264**, 17615–17618
2. Oka, S., Terayama, K., Kawashima, C., and Kawasaki, T. (1992) A novel glucuronyltransferase in nervous system presumably associated with the biosynthesis of HNK-1 carbohydrate epitope on glycoproteins. *J. Biol. Chem.* **267**, 22711–22714
3. Terayama, K., Oka, S., Seiki, T., Miki, Y., Nakamura, A., Kozutsumi, Y., Takio, K., and Kawasaki, T. (1997) Cloning and functional expression of a novel glucuronyltransferase involved in the biosynthesis of the carbohydrate epitope, HNK-1. *Proc. Natl. Acad. Sci. USA* **94**, 6093–6098
4. Terayama, K., Seiki, T., Nakamura, A., Matsumori, K., Ohta, S., Oka, S., Sugita, M., and Kawasaki, T. (1998) Purification and characterization of a glucuronyltransferase involved in the biosynthesis of the HNK-1 epitope on glycoproteins from rat brain. *J. Biol. Chem.* **273**, 30295–30300
5. Kitagawa, H., Tone, Y., Tamura, J., Neumann, K.W., Ogawa, T., Oka, S., Kawasaki, T., and Sugahara, K. (1998) Molecular cloning and expression of Glucuronyltransferase I involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. *J. Biol. Chem.* **273**, 30295–30300
6. Seiki, T., Oka, S., Terayama, K., Imiya, K., and Kawasaki, T. (1999) Molecular cloning and expression of a second glucuronyl transferase involved in the biosynthesis of the HNK-1 carbohydrate epitope. *Biochem. Biophys. Res. Commun.* **255**, 182–187
7. McGarry, R.C., Helfand, S.L., Quarles, R.H., and Roder, J.C. (1983) Recognition of myelin-associated glycoprotein by the monoclonal antibody HNK-1. *Nature* **306**, 376–378
8. Kruse, J., Mailhammer, R., Wernecke, H., Faissner, A., Sommer, I., Goridis, C., and Schachner, M. (1984) Neural cell adhesion

molecules and myelin-associated glycoprotein share a common carbohydrate moiety recognized by monoclonal antibodies L2 and HNK-1. *Nature* **311**, 153–155
9. Bollensen, E. and Schachner, M. (1987) The peripheral myelin glycoprotein P0 expresses the L2/HNK-1 and L3 carbohydrate structures shared by neural adhesion molecules. *Neurosci. Lett.* **82**, 77–82
10. Chou, D.K., Ilyas, A.A., Evans, J.E., Costello, C., Quarles, R.H., and Jungalwala, F.B. (1986) Structure of sulfated glucuronyl glycolipids in the nervous system reacting with HNK-1 antibody and some IgM paraproteins in neuropathy. *J. Biol. Chem.* **261**, 11717–11725
11. Schwarting, G.A., Jungalwala, F.B., Chou, D.K., Boyer, A.M., and Yamamoto, M. (1987) Sulfated glucuronic acid-containing glycoconjugates are temporally and spatially regulated antigens in the developing mammalian nervous system. *Dev. Biol.* **120**, 65–76
12. Yoshihara, Y., Oka, S., Watanabe, Y., and Mori, K. (1991) Developmentally and spatially regulated expression of HNK-1 carbohydrate antigen on a novel phosphatidylinositol-anchored glycoprotein in rat brain. *J. Cell Biol.* **115**, 731–744
13. Bronner-Fraser, M. (1987) Perturbation of cranial neural crest migration by the HNK-1 antibody. *Dev. Biol.* **123**, 321–331
14. Keilhauer, G., Faissner, A., and Schachner, M. (1985) Differential inhibition of neurone-neurone, neurone-astrocyte and astrocyte-astrocyte adhesion by L1, L2 and N-CAM antibodies. *Nature* **316**, 728–730
15. Künemund, V., Jungalwala, F.B., Fischer, G., Chou, D.K., Keilhauer, G., and Schachner, M. (1988) The L2/HNK-1 carbohydrate of neural cell adhesion molecules is involved in cell interactions. *J. Cell Biol.* **106**, 213–223
16. Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fraser, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M., Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laisster, N., Latreille, P., Lightning, J., Lloyd, C., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A., Saunders, D., Shownkeen, R., Sims, M., Smaldon, N., Smith, A., Smith, M., Sonnhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Weinstock, L., Wilkinson-Sproat, J., and Wohldman, P. (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. *Nature* **368**, 32–38
17. Davis, R.E., Hardwick, C., Tavernier, P., Hodgson, S., and Singh, H. (1995) RNA trans-splicing in flatworms. Analysis of trans-spliced mRNAs and genes in the human parasite, *Schistosoma mansoni*. *J. Biol. Chem.* **270**, 21813–21819
18. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene (Amst.)* **77**, 51–59
19. Berry, B.B., Meador, B., Bilderback, T., Liang, P., Glaser, M., and Phillips, G.N., Jr. (1994) The closed conformation of a highly flexible protein: The structure of *E. coli* adenylate kinase with bound AMP and AMPPNP. *PROTEINS: Struct. Funct. Genet.* **19**, 183–198
20. Ishiguro, M., Imajo, S., and Hirama, M. (1991) Modeling study of the structure of the macromolecular antitumor antibiotic neocarzinostatin. Origin of the stabilization of the chromophore. *J. Med. Chem.* **34**, 2366–2373
21. Ishiguro, M. (2000) A mechanism of primary photoactivation reactions of rhodopsin: Modeling of the intermediates in the rhodopsin photocycle. *J. Am. Chem. Soc.* **122**, 444–451
22. Blaney, J.M., Crippen, G.M., Dearing, A., Dixon, J.S., and Spellmeyer, D.C. (1995) *DGEOM95-Distance Geometry. QCPE*, No590. Indiana University, Bloomington, IN 47405, USA
23. Gastinel, L.N., Cambillau, C., and Bourne, Y. (1999) Crystal structures of the bovine beta4galactosyltransferase catalytic domain and its complex with uridine diphosphogalactose. *EMBO J.* **18**, 3546–3557
24. Kakuta, Y., Pedersen, L.G., Carter, C.W., Negishi, M., and Pedersen, L.C. (1997) Crystal structure of estrogen sulphotransferase. *Nat. Struct. Biol.* **4**, 904–908
25. Quiocho, F.A. (1986) Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions. *Annu. Rev. Biochem.* **55**, 287–315
26. Datta, A.K. and Paulson, J.C. (1995) The sialyltransferase "sialylmotif" participates in binding the donor substrate CMP-NeuAc. *J. Biol. Chem.* **270**, 1497–1500
27. Wen, D.X., Livingston, B.D., Medzihradszky, K.F., Kelm, S., Burlingame, A.L., and Paulson, J.C. (1992) Primary structure of Gal beta 1,3(4)GlcNAc alpha 2,3-sialyltransferase determined by mass spectrometry sequence analysis and molecular cloning. Evidence for a protein motif in the sialyltransferase gene family. *J. Biol. Chem.* **267**, 21011–21019
28. Livingston, B.D. and Paulson, J.C. (1993) Polymerase chain reaction cloning of a developmentally regulated member of the sialyltransferase gene family. *J. Biol. Chem.* **268**, 11504–11507
29. Yuan, Y.P., Schultz, J., Mlodzik, M., and Bork, P. (1997) Secreted fringe-like signaling molecules may be glycosyltransferases. *Cell* **88**, 9–11
30. Kawashima, C., Terayama, K., Ii, M., Oka, S., and Kawasaki, T. (1992) Characterization of a glucuronyltransferase: neolactotetraosylceramide glucuronyltransferarase from rat brain. *Glycoconj. J.* **9**, 307–314
31. Mitsumoto, Y., Oka, S., Sakuma, H., Inazawa, J., and Kawasaki T. (2000) Cloning and chromosomal mapping of human glucuronyltransferase involved in biosynthesis of the HNK-1 carbohydrate epitope. *Genomics* **65**, 166–173