Rapid Communication J. Biochem. 120, 236-241 (1996)

Human UDP-Galactose Translocator: Molecular Cloning of a Complementary DNA That Complements the Genetic Defect of a Mutant Cell Line Deficient in UDP-Galactose Translocator[1]

Nobuhiko Miura,*[,2] Nobuhiro Ishida,* Masato Hoshino,* Masatake Yamauchi,† Takahiko Hara,‡ Dai Ayusawa,§ and Masao Kawakita*[,3]

Department of Physiological Chemistry, The Tokyo Metropolitan Institute of Medical Science (Rinshoken), 18-22 Honkomagome 3-chome, Bunkyo-ku, Tokyo 113; †Division of Genetics, National Institute of Radiological Sciences, 4-9-1 Anagawa, Inage-ku, Chiba 260; ‡Institute of Molecular and Cellular Biosciences, The University of Tokyo, Bunkyo-ku, Tokyo 113; and §Kihara Institute of Biological Research, Yokohama City University, Maioka-cho, Totsuka-ku, Yokohama 244

Received for publication, June 21, 1996

We have cloned a cDNA that codes for a putative human UDP-galactose translocator (UGT) protein. The cDNA contained an open reading frame of 1,179 base pairs encoding a novel protein of 393 amino acids. Introduction of the open reading frame sequence into a UGT-deficient mouse cell line, Had-1, complemented the genetic defect of the mutant, namely the inability to transport UDP-galactose from the cytosol to the Golgi lumen, as judged from the lectin-sensitivity spectrum of the transformant. To our knowledge, this is the first mammalian nucleotide-sugar translocator whose cDNA sequence has been described.

Key words: CDC91, Golgi apparatus, UDP-galactose, UDP-galactose translocator, ZK370.7.

Golgi membranes are equipped with nucleotide-sugar transporters that carry the substrates of glycosyltransferases from the cytosol, the site of their synthesis, to the Golgi lumen, where orderly addition of sugars to the growing carbohydrate chains takes place (*1*). Both nucleotide-sugar transporters and glycosyltransferases are thus indispensable for the synthesis of glycoconjugates to proceed properly. While substantial progress has been made in the field of glycosyltransferases in recent years, studies on nucleotide-sugar transporters have lagged considerably behind those on glycosyltransferases. Thus, mutant cells defective in nucleotide-sugar transporters were isolated (*2-5*), and a reconstituted proteoliposome system for studying the transport of nucleotide-sugars was also reported (*6*), but the translocator proteins have yet to be characterized. Had-1, a Newcastle disease virus-resistant mutant of mouse FM3A cells, was originally isolated as a hemagglutination-deficient mutant, and was shown to be defective in the galactosylation of cell-surface glycoconjugates (*7-9*). The mutant was also shown to belong to the same complementation group as Lec8, that is defective in the UDP-galactose translocator (*5, 8*). We have attempted to utilize the mutant to clone the human UDP-galactose translocator (UGT) gene by genetic complementation with human DNA, and obtained a segment of DNA that represented a part of the human UGT gene. The isolated gene segment enabled us to map the gene to band Xp11.22-p11.23 on the X chromosome (*10*). Use of this gene segment also enabled us to determine the sequence of a complementary DNA of the UDP-galactose translocator, as described below. When introduced into Had-1 cells, the cDNA was able to complement the genetic defect of the mutant.

We first identified the exon sequences present in the previously isolated 16-kbp long DNA segment that represented a part of the UGT gene (*10*) by exon amplification analysis (*11*). Upon *Eco*RI digestion, the 16-kbp genomic clone generated 5 fragments of 10-, 2.7-, 1.7-, 1-, and 0.5-kbp long, respectively. Of these fragments, the 10- and 2.7-kbp ones were introduced into pSPL3, and then transfected into COS-7 cells using an Exon Trapping System (Life Technologies, Gaithersberg, MD) according to the manufacturer's instructions. RNA molecules with sequences contributing to UGT-exons flanked at both ends by defined sequences derived from the vector were spliced out *in situ* in the transfected COS-7 cells, and then amplified *in vitro* by RT-PCR as described previously (*12*). Fragments that contained 183 bases (E10k exon) and 152 bases (E2.7k exon) of the UGT sequence were obtained from the 10- and 2.7-kbp fragments, respectively (Fig. 1a). Northern blot analysis with the E10k exon fragment as a probe indicated expression of a major mRNA species of about 3-kb long (Fig. 1b). The same result was obtained with the E2.7k exon as a probe (data not shown). A cDNA library derived from human normal diploid fibroblast TIG-1 cells was screened using the E2.7k exon DNA, and a 1.1-kbp DNA clone was isolated. Both the E10k and E2.7k exon se-

[1] This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan, and a grant from Suntory Co., Japan. The nucleotide sequence reported in this paper has been submitted to the DDBJ/GenBank™/EMBL Data Bank under accession number D84454.

[2] Present address: Department of Molecular and Biochemical Toxicology, Faculty of Pharmaceutical Sciences, Tohoku University, Aoba-ku, Sendai 980-77.

[3] To whom correspondence should be addressed. Tel.: +81-3-3823-2101 (Ext. 5285), Fax: +81-3-3823-2965, e-mail: kawakita@rinshoken.or.jp

quences were identified in the clone, but the clone was truncated at both its 5′- and 3′-ends for unknown reasons (Fig. 1a). To determine the entire mRNA sequence, 3′-RACE (*13*) and 5′-RACE (*14*) were performed as described (Fig. 1a). For 5′-RACE, human liver 5′-RACE-Ready cDNA (CLONTECH, Palo Alto, CA) was utilized. Figure 2 shows the deduced mRNA sequence with a predicted open reading frame (ORF) of 1,179 bases flanked by a 5′-untranslated region of 323 bases and a 3′-untranslated region of 1,118 bases in length. As human liver cDNA was used for 5′-RACE, as described above, we checked and confirmed the sequence of 5′-UTR by sequencing the RT-PCR product of the corresponding region obtained from TIG-1 poly(A)$^+$ RNA. The full nucleotide sequence of UGT cDNA in Fig. 2 therefore represents the one from TIG-1 cells. The putative initiation codon, ATG, is accompanied by purines at positions −3 and +4, thus fulfilling the criteria for a consensus initiator methionine (*15*), and is preceded by a termination codon in the reading frame. The deduced protein was composed of 393 amino acids, and its calculated molecular mass was 41 kDa.

Complementary DNA for the open reading frame was isolated by RT-PCR from poly(A)$^+$ RNA of TIG-1 cells, using the PCR primers indicated in Fig. 1a. The ORF cDNA was cleaved with *Eco*RI at the recognition sites introduced in the primers, and then ligated into the *Eco*RI site in the cloning region of an expression vector, pMKIT-neo, a derivative of the SRα vector (*16*) (K. Maruyama, personal communication). Finally, the residual stuffer region in the vector was removed by digestion with *Not*I, followed by self-ligation to complete pMKIT-neo-hUGT. The recombinant plasmid, pMKIT-neo-hUGT, was transfected into Had-1 mutant cells by the procedure described previously (*17*), and then the galactosylation of cell surface glycoconjugates was assessed by determining the lectin-sensitivity of the transformants. As reported previously, the Had-1 mutant is highly sensitive to the cytotoxicity of *Griffonia simplicifolia* lectin GS-II (EY Laboratories, San Mateo, CA), because *N*-acetylglucosamine residues that are recognized by the lectin are exposed at the ends of *N*-linked glycans on the cell surface (*7, 8*). On the other hand, parental FM3A cells are resistant to GS-II and able to grow in the presence of the lectin, while they are sensitive to

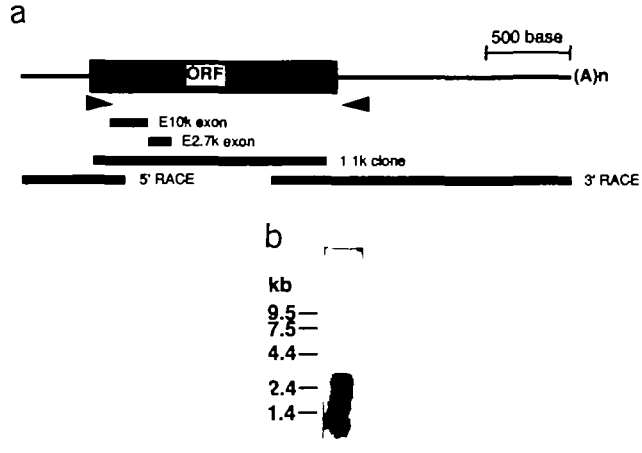

Fig. 1. **Characterization and analysis of UGT cDNA.** a: Cloning strategy for UGT cDNA. The RT-PCR products representing UGT exons of 183 bases (E10k exon) and 152 bases (E2.7k exon) were obtained by exon amplification (*11*) from 10 and 2.7 kbp fragments, respectively. A cDNA clone harboring a 1.1 kbp insert (1.1k clone) was isolated from a human TIG-1 cDNA library with the E2.7k exon as a probe. The sequence information concerning the 5′- and 3′-regions was obtained by 5′-RACE and 3′-RACE, respectively. Arrowheads indicate the positions of the RT-PCR primers used to prepare the UGT ORF fragment. b: Northern blot hybridization showing that the E10k exon hybridized mainly with transcripts of about 3k base long. About 1.5 μg of poly(A)$^+$ RNA from TIG-1 cells prepared using ISOGEN reagent (WAKO, Tokyo) and oligo(dT)-paramagnetic beads (DYNAL, Oslo, Norway) was size-fractionated by formaldehyde agarose gel electrophoresis and then hybridized with the ^{32}P-labeled E10k exon after transfer to a membrane filter. The radioactivity on the blot was visualized with a BAS 2000 image analyzer (Fuji Photofilm, Tokyo).

TABLE I. **Appearance of GS-II-resistant and WGA-sensitive cells among the population of Had-1 cells transformed with the pMKIT-neo-hUGT plasmid.** FM3A cells were transfected with pMKIT-neo, and Had-1 cells with either pMKIT-neo or pMKIT-neo-hUGT as previously described (*17*). Transformants were selected from these pools in ES medium (Nissui Pharmaceutical, Tokyo) supplemented with 2% FCS, that contained 1 mg/ml G418, for 2 weeks, and the survivors were maintained thereafter in the same medium without G418. Equivalent numbers of cells selected from each pool were plated (150 cells for each 60-mmϕ plate) and cultured on 0.35% agarose-ES medium-2% FCS containing G418 (1 mg/ml) alone (Control) or G418 and either GS-II (20 μg/ml) or WGA (20 μg/ml). The numbers of cell-islands formed on the plates were determined at 10 days after inoculation. The mean values obtained for two plates for each set of conditions are presented. Figures in parentheses represent percentages of the numbers of cell islands on control (G418 alone) plates.

Transformants	Control	GS-II	WGA
FM3A-pMKIT-neo	62.5 (100)	76 (122)	0 (0)
Had-1-pMKIT-neo	75.5 (100)	0 (0)	43.5 (58)
Had-1-pMKIT-neo-hUGT	79.5 (100)	66 (83)	24 (30)

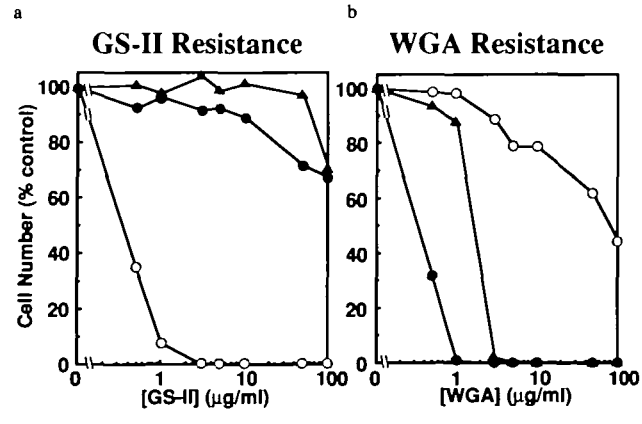

Fig. 3. **Reversal of the lectin-resistance spectrum of Had-1 cells to that of FM3A cells on introduction of human UGT cDNA.** The same number of cells of each clone (10^4 cells in 1 ml of culture medium) was initially inoculated into ES medium supplemented with 2% FCS containing various concentrations of either GS-II (a) or WGA (b), as indicated, and then grown for 3 days in a 24-well culture dish. The number of cells was then determined and presented as a percentage of that of the control cultures grown in the absence of the lectin. ●, Had-1-hUGT-GS-II-3; ○, Had-1-neo-2; ▲, FM3A-neo-2.

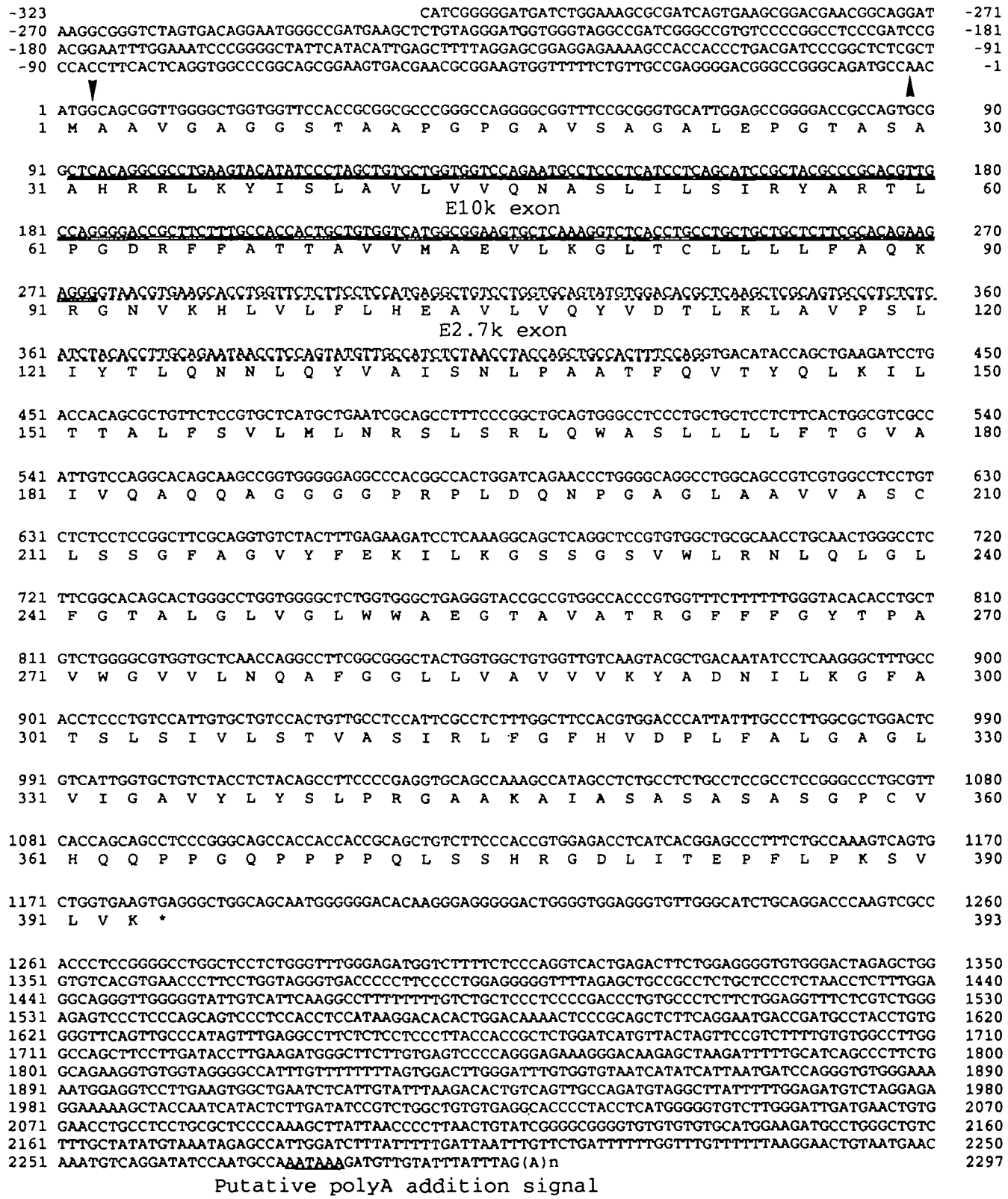

Fig. 2. **Human UGT sequence.** The nucleotide sequence was determined with 373S and 377 DNA sequencers (Perkin-Elmer, Norwalk, CT) using a Taq DyeDeoxy Thermal Cycle Sequencing Kit (Perkin-Elmer). Arrowheads indicate the purines at positions −3 and +4 relative to the putative initiator ATG. A putative poly(A) addition signal located at positions 2274-2279 is underlined. The locations of the E10k exon and the E2.7k exon are indicated by a double underline and a broken underline, respectively.

a

```
hUGT        85 LLFAQKRGNVKHLVLFLHEAVLVQYVDTLKLAVPSLIYTLQNNLQYVAIS 134
ZK370.7      1 MFYKECNFDSRQFSEQVSKYYINAPRELAKMSVPSFAYALQNNLDFVGLS  50
hUGT       135 NLPAATFQVTYQLKILTTALFSVLMLNRSLSRLQWASLLLLFTGVAIVQA 184
ZK370.7     51 NLDAGLYQVTTQLKVVSTAFFMMLFLGRKFSTRRWMAITLLMFGVAFVQM 100
hUGT       185 QQAGGGGP...RPLDQNPGAGLAAVVASCLSSGFAGVYFEKILK.GSSGS 230
ZK370.7    101 NNVSASEANTKRETAENYIVGLSAVLATCVTAGFAGVYFEKMLKDGGSTP 150
hUGT       231 VWLRNLQL...GLFGTALG.LVGLWWAEGTAVATRGFFFGYTPAVWGVVL 276
ZK370.7    151 FWIRNMQMYSCGVISASIACLTDFSR.....ISDKGFFFGYTDKVWAVVI 195
hUGT       277 NQAFGGLLVAVVVKYADNILKGFATSLSIVLSTVASIRLF.GFHVDPLFA 325
ZK370.7    196 LLGVGGLYISLVMRYLDNLYKSMASAVSIILVVVLSMLIFPDIFIGMYFV 245
```

b

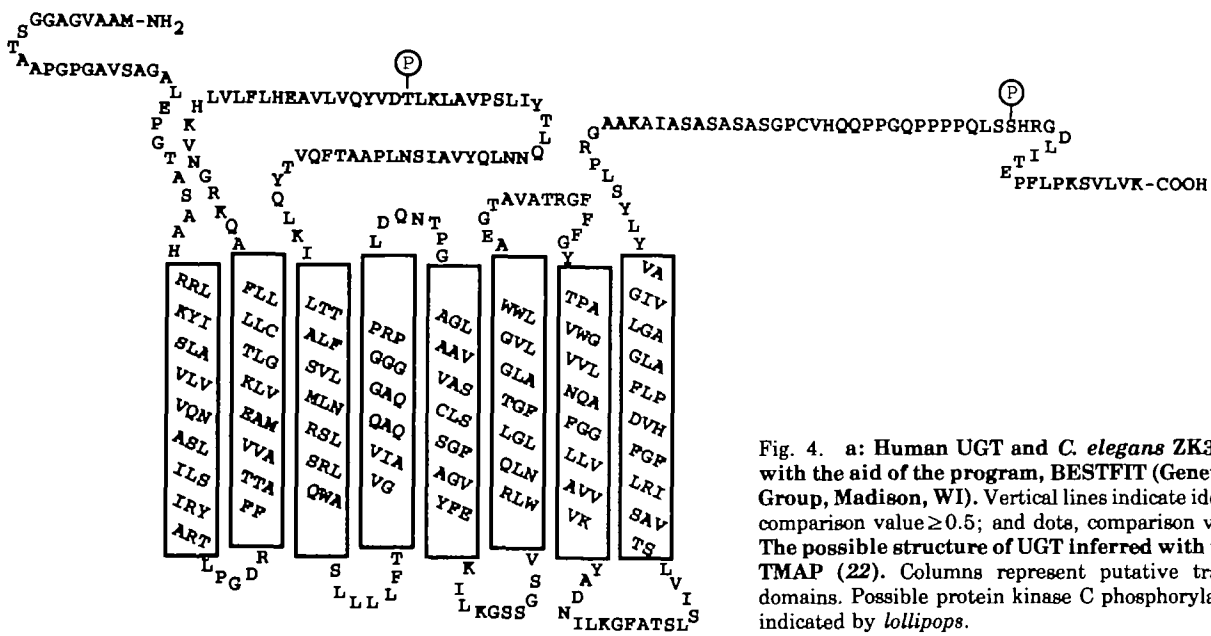

Fig. 4. **a: Human UGT and *C. elegans* ZK370.7 aligned with the aid of the program, BESTFIT (Genetic Computer Group, Madison, WI).** Vertical lines indicate identity; colons, comparison value ≥ 0.5; and dots, comparison value ≥ 0.1. **b: The possible structure of UGT inferred with the program, TMAP (*22*).** Columns represent putative transmembrane domains. Possible protein kinase C phosphorylation sites are indicated by *lollipops*.

wheat germ agglutinin (WGA) (EY Laboratories), to which Had-1 cells are resistant. A significant proportion of the pMKIT-neo-hUGT transformant of Had-1 cells acquired resistance to cytotoxicity of GS-II, while the frequency of WGA-resistant cells among the transformants was definitely reduced (Table I). To determine whether or not the GS-II-resistant clones had recovered a lectin-sensitivity spectrum similar to that of FM3A, as would be expected if the hUGT cDNA in fact complemented the genetic defect of Had-1, several clones were isolated on agarose plates containing G418 and GS-II from the pool of pMKIT-neo-hUGT transformants of Had-1 (Table I). The isolated clones were examined as to their lectin-sensitivity, and representative results are shown in Fig. 3. Had-1-hUGT-GS-II-3 denotes a pMKIT-neo-hUGT transformant clone isolated from an agarose plate containing G418 and GS-II as described above. For reference purposes, Had-1-neo-2 and FM3A-neo-2 were also isolated from the pool of pMKIT-neo-transformants of Had-1 and FM3A, respectively, on agarose plates containing G418, and then their lectin-sensitivity was examined in parallel experiments. The GS-II-resistant clones isolated from the pool of the pMKIT-neo-hUGT transformants of the Had-1 mutant were shown to exhibit a lectin-resistance spectrum that is quite similar to that of the parental FM3A (Fig. 3). This indicates that the defect of the Had-1 mutant, namely the inability to transport UDP-galactose from the cytosol to the Golgi lumen, was complemented by the genetic information carried by the hUGT cDNA.

Reference to a protein database with the BLITZ search (EMBL) (*18*) revealed that UGT shows a significant degree of similarity with the hypothetical 29.4 kDa protein ZK370.7 on chromosome III of *Caenorhabditis elegans* with an unidentified function (*19*), and a quite modest degree of similarity to the PotI protein of *Escherichia coli*, an integral membrane component of the putrescine transporter com-

plex (*20*). Comparison between human UGT and ZK370.7 indicated 66% similarity (38% identity) within the overlapping region (Fig. 4a). Assuming the occurrence of common motifs among nucleotide-sugar transporters, or substantial conservation of the structure of UGT among species, this may imply that ZK370.7 represents a nucleotide-sugar transporter, or may even be a UGT of *C. elegans*. The hydropathy profile and secondary structure prediction of the putative UGT protein (*21, 22*) suggest that the cloned cDNA encodes an intrinsic membrane protein with 8 putative transmembrane domains (Fig. 4b). The predicted structure is consistent with the function of the encoded protein as a transporter. A HPT homology search of the yeast ORFs at the Martinsried Institute for Protein Sequences (http://speedy.mips.biochem.mpg.de/mips/yeast/) suggested the similarity of hUGT with cell division control protein CDC91 (GenBank: U22383) of *Saccharomyces cerevisiae*, which is also a hypothetical protein of 394 amino acid residues whose function is actually unknown. The similarity between human UGT and CDC91 is much lower than that between human UGT and ZK370.7, being 43%, but CDC91 is also predicted to have 8 transmembrane regions like human UGT. As glycoproteins of *S. cerevisiae* lack galactose residues (*23*), the organism may not have a Golgi UDP-galactose translocator. It is intriguing as to whether or not CDC91 represents a transporter of nucleotide-sugar(s) other than UDP-galactose.

The activity of nucleotide-sugar transporters may be detected biochemically in Golgi membrane vesicles (*6*), and mutant cells that lack such transporters have long been available (*2-5, 7, 8*), but their molecular structures, including the amino acid sequences, or even their molecular masses have not been reported yet. Starting from isolation of the Had-1 mutant (*7*), we have eventually determined the primary structure of a putative UGT protein for the first time. To our knowledge, this is the first mammalian sugar-nucleotide transporter whose structure has been elucidated. It is likely that nucleotide-sugar translocators constitute a family of related proteins. The UGT cDNA and its nucleotide sequence information will provide us with important clues for isolating other nucleotide-sugar transporter genes. The expression plasmid encoding the UGT cDNA would also be valuable as a starting point for the production of the recombinant UGT in suitable quantities for biochemical analyses. Moreover, the method of genetic complementation analysis involving the Had-1 mutant and molecularly altered UGT cDNA will prove to be a powerful tool for analyzing the structure-function relationship of the UGT molecule.

We are grateful to Dr. Kazuo Maruyama, Tokyo Medical and Dental University, for providing us with the pMKIT-neo.

REFERENCES

1. Hirschberg, C.B. and Snider, M.D. (1987) Topography of glycosylation in the rough endoplasmic reticulum and Golgi apparatus. *Annu. Rev. Biochem.* **56**, 63-87
2. Briles, E.B., Li, E., and Kornfeld, S. (1977) Isolation of wheat germ agglutinin-resistant clones of Chinese hamster ovary cells deficient in membrane sialic acid and galactose. *J. Biol. Chem.* **252**, 1107-1116
3. Stanley, P. (1985) Membrane mutants of animal cells: Rapid identification of those with a primary defect in glycosylation. *Mol. Cell. Biol.* **5**, 923-929
4. Deutscher, S.L., Nuwayhid, N., Stanley, P., Briles, E.I., and Hirschberg, C.B. (1984) Translocation across Golgi vesicle membranes: A CHO glycosylation mutant deficient in CMP-sialic acid transport. *Cell* **39**, 295-299
5. Deutscher, S.L. and Hirschberg, C.B. (1986) Mechanism of galactosylation in the Golgi apparatus: A Chinese hamster ovary cell mutant deficient in translocation of UDP-galactose across Golgi vesicle membranes. *J. Biol. Chem.* **261**, 96-100
6. Milla, M.E., Clairmont, C.A., and Hirschberg, C.B. (1992) Reconstitution into proteoliposomes and partial purification of the Golgi apparatus membrane UDP-galactose, UDP-xylose, and UDP-glucuronic acid transport activities. *J. Biol. Chem.* **267**, 103-107
7. Hara, T., Hattori, S., and Kawakita, M. (1989) Isolation and characterization of mouse FM3A cell mutants which are devoid of Newcastle disease virus receptors. *J. Virol.* **63**, 182-188
8. Hara, T., Endo, T., Furukawa, K., Kawakita, M., and Kobata, A. (1989) Elucidation of the phenotypic change on the surface of Had-1 cell, a mutant cell line of mouse FM3A carcinoma cells selected by resistance to Newcastle disease virus infection. *J. Biochem.* **106**, 236-247
9. Taki, T., Ogura, K., Rokukawa, C., Hara, T., Kawakita, M., Endo, T., Kobata, A., and Handa, S. (1991) Had-1, a uridine 5′-diphosphogalactose transport-defective mutant of mouse mammary tumor cell FM3A: Composition of glycolipids, cell growth inhibition by lactosylceramide, and loss of tumorigenicity. *Cancer Res.* **51**, 1701-1707
10. Hara, T., Yamauchi, M., Takahashi, E., Hoshino, M., Aoki, K., Ayusawa, D., and Kawakita, M. (1993) The UDP-galactose translocator gene is mapped to band Xp11.23-p11.22 containing the Wiskott-Aldrich syndrome locus. *Somat. Cell Mol. Gen.* **19**, 571-575
11. Buckler, A.J., Chang, D.D., Graw, S.L., Brook, J.D., Haber, D.A., Sharp, P.A., and Housman, D.E. (1991) Exon amplification: A strategy to isolate mammalian genes based on RNA splicing. *Proc. Natl. Acad. Sci. USA* **88**, 4005-4009
12. Ishida, N., Fujita, H., Fukuda, Y., Noguchi, T., Doss, M., Kappas, A., and Sassa, S. (1992) Cloning and expression of the defective genes from a patient with delta-aminolevulinate dehydratase porphyria. *J. Clin. Invest.* **89**, 1431-1437
13. Frohman, M.A. (1993) Rapid amplification of complementary DNA ends for generation of full-length complementary DNAs: Thermal RACE in *Methods in Enzymology* (Wu, R., ed.) Vol. 218, pp. 340-356, Academic Press, New York
14. Edwards, J.B., Delort, J., and Mallet, J. (1991) Oligodeoxyribonucleotide ligation to single-stranded cDNAs: A new tool for cloning 5′ ends of mRNAs and for constructing cDNA libraries by in vitro amplification. *Nucleic Acids Res.* **19**, 5227-5232
15. Kozak, M. (1991) Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.* **266**, 19867-19870
16. Takebe, Y., Seiki, M., Fujisawa, J., Hoy, P., Yokota, K., Arai, K., Yoshida, M., and Arai, N. (1988) SRα promoter: An efficient and versatile mammalian cDNA expression system composed of the simian virus 40 early promoter and the R-U5 segment of human T-cell leukemia virus type 1 long terminal repeat. *Mol. Cell. Biol.* **8**, 466-472
17. Ayusawa, D., Shimizu, K., Koyama, H., Takeishi, K., and Seno, T. (1983) Unusual aspects of human thymidylate synthase in mouse cells introduced by DNA-mediated gene transfer. *J. Biol. Chem.* **258**, 48-53
18. Smith, T.F. and Waterman, M.S. (1981) Identification of common molecular subsequences. *J. Mol. Biol.* **147**, 195-197
19. Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J. *et al.* (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. *Nature* **368**, 32-38
20. Pistocchi, R., Kashiwagi, K., Miyamoto, S., Nukui, E., Sadakata, Y., Kobayashi, H., and Igarashi, K. (1993) Characteristics of the operon for a putrescine transport system that maps at 19 minutes on the *Escherichia coli* chromosome. *J. Biol. Chem.* **268**, 146-152

21. Nakai, K. and Kanehisa, M. (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* **14**, 897-911
22. Persson, B. and Argos, P. (1994) Prediction of transmembrane segments in proteins utilising multiple sequence alignments. *J. Mol. Biol.* **237**, 182-192
23. Herscovics, A. and Orlean, P. (1993) Glycoprotein biosynthesis in yeast. *FASEB J.* **7**, 540-550