Molecular Cloning and Characterization of a Novel α1,2-Fucosyltransferase (CE2FT-1) from Caenorhabditis elegans*

Qinlong Zheng§*, Irma Van Die*, and Richard D. Cummings§‡¶

Received for publication, July 25, 2002
Published, JBC Papers in Press, August 5, 2002, DOI 10.1074/jbc.M207487200

From the §Department of Biochemistry and Molecular Biology and the ¶Oklahoma Center for Medical Glycobiology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73104 and the ¶Department of Molecular Cell Biology, VU University Medical Center, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands

Here we report the discovery of a unique fucosyltransferase (FT) in Caenorhabditis elegans. In studying the activities of FTs in extracts of adult C. elegans, we detected activity toward the unusual disaccharide acceptors Galβ1-4Xyl-R and Galβ1-6GlcNac-R to generate products with the general structure Fucα1-2Galβ1-R. We identified a gene encoding a unique α1,2FT (designated CE2FT-1), which contains an open reading frame encoding a predicted protein of 355 amino acids with the type 2 topology and domain structure typical of other glycosyltransferases. The predicted cDNA for CE2FT-1 has very low identity (5–10%) to the amino acid level to α1,2FT sequences in humans, rabbits, and mice. Recombinant CE2FT-1 expressed in human 293T cells has high α1,2FT activity toward the simple acceptor Galβ1-O-phenyl acceptor to generate Fucα1-2Galβ-R, which in this respect resembles mammalian α1,2FTs. However, CE2FT-1 is otherwise completely different from known α1,2FTs in its acceptor specificity, since it is unable to fucosylate either Galβ1-4Glcβ-R or free lactose and prefers the unusual acceptors Galβ1-4Xylβ-R and Galβ1-6GlcNac-R. Promoter analysis of the CE2FT-1 gene using green fluorescent protein reporter constructs demonstrates that CE2FT-1 is expressed in single cells of early stage embryos and exclusively in the 20 intestinal cells of L1–L4 and adult worms. These and other results suggest that multiple fucosyltransferase genes in C. elegans may encode enzymes with unique activities, expression, and developmental roles.

Fucose-containing glycans are important in a number of biological processes, including their roles as a component of the ligands in mediating leukocyte adhesion to selectins. Mammals possess a number of fucose-containing structures, including those with Fucα1-3GlcNAcβ-R, Fucα1-6GlcNAcβ-R, and Fucα1-2Galβ-R (1, 2). These linkages are generated by a variety of α-fucosyltransferases (FTs) derived from a surprisingly high number of genes. Humans have at least six different α1,3FT genes (III, IV, V, VI, VII, and IX), two α1,2FT genes (FUT1 (H) and FUT2 (Sc) and an α1,2FT pseudogene (SECI)), and at least one α1,6FT gene (3–5). The functions of all of these genes and their cognate fucose-containing glycans are not well understood, but the expression of many of these genes is developmentally regulated and is altered during embryonic development, differentiation, and tumorigenesis. Although genetic approaches in mice have revealed many important insights into the functions of many different genes, particularly in regard to the α1,3FTs and their role in generating cell adhesion ligands (5, 6), genetic approaches to study glycoconjugate roles in development have been frustrated by the complexity of the large multigene families for most of the glycosyltransferases and consequent enzyme redundancy. This situation has led to the consideration of model systems that may accelerate our understanding of the overall biological functions of fucose-containing glycoconjugates.

Caenorhabditis elegans is a highly attractive model system in which to study the potential developmental roles of glycoconjugates, since the worm shares many fundamental biological and biochemical pathways with higher organisms. The potential of using C. elegans to investigate the roles of glycoconjugates in development requires the identification and characterization of specific glycosyltransferases and their cognate structures, which may be important in developmental processes. During the past few years, many different enzymes involved in glycoconjugate biosynthesis have been identified by homology in C. elegans (1). In our previous studies, we identified an α1,3FT and an α1,2FT activity in C. elegans extracts (7). However, further analysis of the acceptor specificity demonstrates that C. elegans has a novel α1,2FT activity toward the unusual acceptors Galβ1-4Xylβ-R and Galβ1-6GlcNac-R to synthesize the Fucα1-2Galβ1-R products; such acceptor specificity would be unique among known α1,2FTs. However, recent studies on the O-glycans of C. elegans glycopolymers document the occurrence of such Fucα1-2Galβ1-R linkages (8). Searches of the C. elegans genome revealed that dozens of putative fucosyltransferase genes are present (3), and the C. elegans genome contains at least 22 different genes encoding putative α1,2FTs with some identity to mammalian α1,2FTs.

In this paper, we report our identification of the C. elegans gene (termed CE2FT-1) encoding the α1,2FT (CE2FT-1) capable of fucosylating the disaccharide acceptors Galβ1-4Xylβ-R and Galβ1-6GlcNac-R but unable to fucosylate lactose. Preliminary studies on other members of this putative α1,2FT family in C. elegans suggest that they differ from CE2FT-1 in acceptor specificity. The expression pattern of CE2FT-1 is also unusual and is limited strictly to the 20 intestinal cells in the adult worm. These unexpected findings suggest that each of the FTs in C. elegans may have a unique enzyme activity and role in development.

* This work was supported by National Institutes of Health Grant RO1 HD037549 (to R. D. C.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Correspondence should be addressed: Dept. of Biochemistry and Molecular Biology, University of Oklahoma Health Sciences Center, 975 N.E. 10th St., BRC417, Oklahoma City, OK 73104. Tel.: 405-271-2481; Fax: 405-271-3910; E-mail: richard-cummings@ouhsc.edu.

§ The abbreviations used are: FT, fucosyltransferase; GFP, green fluorescent protein; pNP, p-nitrophenol; Bzl, benzyl; LNFI, Galβ1–3(Fucα1–4)GlcNacβ1–3Galβ1–4Glc; LNFI, Galβ1–4(Fucα1–3)GlcNacβ1–3Galβ1–4Glc.
min, amplification for 30 cycles at 94 °C for 1 min, 56 °C for 1 min and 72 °C for 2 min using plasmid pcdNA/HisMaxC/CE2FT-1 cDNA as the template. The expected 1420 bp of PCR product was purified on 1.2% TAE-agarose gel and digested by EcoRI and ApoI. The expected 1000-bp DNA fragment was purified and cloned into ApoI (partially digested)/ HindIII-digested pcdNA/HisMaxC/CE2FT-1 cDNA. The subsequent sequence was confirmed. Human 293T cells were transiently transfected with 4.5 μM of plasmid DNA harboring CE2FT-1 cDNA with the HPC4 epitope tag fused to the C terminus (in the plasmid pcdNA/HisMaxC) using FuGene 6 transfection reagent. Seventy-two hours after transfection, cells were harvested and solubilized in 50 mM sodium cacodylate buffer (pH 7.0), and the expressed protein was purified on immobilized HPC4 monoclonal antibody from the cell extracts. Following purification, the epitope-tagged CE2FT-1 was assayed by Western blotting using the HPC4 antibody and was assayed for fusocysferase activity as described below. Fusocysferase Assay—Standard fusocysferase assays were performed in a 50-μl reaction mixture containing 5 μmol of sodium cacodylate (pH 7.2), 1 μmol of MnCl2, 0.2 μmol of ATP, 5 nmol of GDP-[3H]Fuc (3.6 Ci/mmol), 0.2% (v/v) Triton X-100, and 15 μl of cell extract. Aceptor substrate concentrations were as indicated. After incubation for 120 min at 37 °C, the reaction was stopped, and the product was separated from unincorporated label by chromatography on a 1-ml column of Dowex 1-X8 (Cl− form) as described (11, 12); [3H]Fuc-labeled products were isolated by adsorption to Sep-Pak C-18 cartridges (Waters; Milford, MA), washed with water, and eluted with absolute methanol, as described (13). Control assays lacking the acceptor substrate were carried out to correct for incorporation into endogenous acceptors, and assays with mock-transfected cells were conducted to correct for endogenous fusocysferase activity. C. elegans Culture and Preparation of Extracts—The standard laboratory wild type strain N2 worms were grown on NGM plates seeded with OP50 bacteria to obtain lots of mixed stage worms. Worms were washed with M9 buffer and were frozen in liquid nitrogen for 2 min, and the worm pellets were stored at −80 °C. For C. elegans extracts, frozen worms were ground and suspended in 50 mM sodium cacodylate buffer, pH 7.0, containing 1 tablet/10 ml complete protease inhibitor mixture (Roche Molecular Biochemicals), 1% Triton X-100 and then subjected to sonication (three pulses of 10 s each) on a cell disruptor (Branson Sonic Power Co.). The homogenate was incubated on ice for 30 min to allow solubilization of proteins and then centrifuged at 14,000 rpm at 4 °C for 10 min. The supernatant fractions were collected, and they were either used directly or stored as aliquots at −80 °C. Frozen extracts were thawed only once for use as assay material. Characterization of Fusocysferase Assay Products—Products obtained with the acceptor Gal-β-picolinoyl, GDP-[3H]Fuc, and 293T cell extracts expressing recombinant CE2FT-1 were isolated by chromatography on Sep-Pak C-18 cartridges (Waters). The isolated, radiolabeled products were incubated with 0.8 milligrams of α1–2-fucosidase, α1–3,4-fucosidase, and α1,6-fucosidase in 20 μl of reaction buffer 5 (Prozyme, Promega, Madison, WI) for 48 h. Following treatment, the samples were analyzed by descending paper chromatography in a solvent system of ethyl acetate/pyridine/acyclic acid/water (5:5:1:3). Products obtained upon incubation of C. elegans extracts with the acceptors Galβ1–3Galβ1–4Xyl-O-Bzl, Galβ1–4Xyl-O-Bzl, Galβ1–3GalNacO–pNP, and the donor GDP-[3H]Fuc were isolated by chromatography on Sep-Pak C-18 cartridges (Waters). The isolated, radiolabeled products were incubated with 0.8 milligrams of α1–2-fucosidase, α1–3,4-fucosidase, and α1,6-fucosidase, and water (mock treatment) in 50 μl of reaction buffer 5 (Prozyme) at 37 °C for 48 h, respectively. After treatment, the samples were isolated by adsorption to Sep-Pak C-18 cartridges (Waters), washed with water, and eluted with absolute methanol, and the radioactivity in the eluted material was determined by scintillation counting. The percentage of [3H]Fuc released was calculated based on the mock treatment.

Preparation of DNA Constructs for Promoter Analysis—We prepared fusion constructs of the upstream promoter sequences of CE2FT-1 fused to the green fluorescent protein to examine the spatial pattern of CE2FT-1 expression during C. elegans development. A 700-bp fragment of genomic sequence containing the putative promoter region immediately upstream of the fusion site for C. elegans translational fusion was obtained by PCR using C. elegans genomic DNA as template with the primers 5′-AAAAATATCATTATATAATGTTTGGCGG-3′, and 5′-ACTATAAAAAATACCGGGAC-3′. The PCR product was run on an agarose gel, and an ~700-bp DNA band was excised from agarose gel and purified with the Qiagen agarose gel extraction kit. The purified DNA was ligated into pCR3.1 vector and subsequently transformed into Top10F competent cells. Clones containing inserts were selected and
amplified, and minipreps were prepared using the Qiagen kit and then sequenced. The clone containing the right insert was digested with HindIII/XbaI and subcloned into pPD95.67 vector (originally provided by Dr. Andrew Z. Fire (Carnegie Institute of Washington, Baltimore, MD)), creating plasmid pPD95.67/CE2FT-1-prom.

Preparation of Transgenic Worms and Promoter Analysis—DNA injection into the C. elegans germ line was carried out as described (14, 15). Transgenic lines were established from F2 descendants of animals injected with 10 ng/μl CE2FT-1:GFP constructs (pPD95.67/CE2FT-1-prom), and pBX, which carries the PHA-1 gene, which is involved in the development of the pharynx, served as a transformation marker.

Preparing Synchronous Culture of L1 and L2—Adult Worm—N2 worms were grown on NGM plates seeded with OP50 bacteria to obtain lots of gravid adults. These were incubated with 5 ml of alkaline bleach (0.25% KOH, 25% commercial Clorox) at room temperature for 3 min with occasional gentle agitation. This procedure kills all adults and larvae but leaves the eggs alive. The eggs were washed with M9 buffer and centrifuged for 1 min in a clinical centrifuge at 1800 rpm. The supernatant was removed, and the washing was repeated two times. The washed eggs were distributed to one unseeded plate, which was incubated at 20 °C overnight. This allows the eggs to hatch and arrests the animals as starved L1, which produces synchrony of the population. Growth was restarted by distributing these starved L1 onto plates seeded with an adequate amount of food to allow the animals to grow without starving. At times during growth, worms were harvested to generate populations of L1–L4, and young adult worms—500 C. elegans to generate populations of L2 adults and larvae but leaves the eggs alive. The eggs were washed with M9 buffer and centrifuged for 1 min in a clinical centrifuge at 1800 rpm. The supernatant was removed, and the washing was repeated two times. The washed eggs were distributed to one unseeded plate, which was incubated at 20 °C overnight. This allows the eggs to hatch and arrests the animals as starved L1, which produces synchrony of the population. Growth was restarted by distributing these starved L1 onto plates seeded with an adequate amount of food to allow the animals to grow without starving. At times during growth, worms were harvested to generate populations of L1–L4, and young adult worms.

RT-PCR and Quantitative Real Time RT-PCR Analysis of C. elegans CE2FT-1 mRNA during Development—Total RNA was extracted from staged synchronous populations using a total RNA isolation kit (Ambion, Austin, TX). Random primed first strand cDNA was synthesized from 2 μg of total C. elegans RNA using the SuperScript premultiplication system for first strand cDNA synthesis kit (Invitrogen). The first strand cDNA was used as a template in PCRs to amplify cDNAs encoding CE2FT-1 transcripts using the oligonucleotides 5’-GATGAGAAA-GGTGAAAGGACCTTTTACG-3’ and 5’-CATCATTTTTTGAATATTGACC-3’ in the reactions containing 3 μl of first strand cDNA in a volume of 50 μl for 40 cycles. For controls, a C. elegans actin gene-specific primer pair 5’-GACAATGTGAGATCAGTGC-3’ and 5’-ATGAGGTTGTTAGAAAGTCTC-3’ was used to amplify the C. elegans actin gene in the identical conditions described. Following PCR amplification, 20 μl of each PCR product was ran on a 1.5% agarose gel containing 10 μg/ml ethidium bromide. Quantitative RT-PCR was performed using an ABI Prism 7000 sequence detection (PerkinElmer Life Sciences) using the double-stranded DNA binding dye, SYBR Green. Random primed first strand cDNA was synthesized in a 100-μl reaction from 2 μg of the total C. elegans mRNA using TaqMan Reverse Transcription Reagents (PerkinElmer Life Sciences). The first strand cDNA was used as a template in PCRs to amplify cDNAs encoding CE2FT-1 transcripts using gene-specific primers 5’-GATGAGAAA-GGTGAAAGGACCTTTTACG-3’ and 5’-CATCATTTTTTGAATATTGACC-3’, designed by using Primer Express software, in the reactions containing 5 μl of first strand cDNA and 25 μl of 2X SYBR Green Master Mix (PerkinElmer Life Sciences) in a volume of 50 μl. For controls, a C. elegans actin gene-specific primer pair 5’-ATGAGGAGGAAATCTGACG-3’ and 5’-CATCATTTTTTGAATATTGACC-3’ was used to amplify the C. elegans actin gene in the identical conditions described. Cycling parameters were 95 °C for 15 s and 60 °C for 1 min. To confirm the absence of nonspecific amplification, the PCR products were analyzed by a 1.2% agarose gel containing 10 μg/ml ethidium bromide.

RESULTS

Identification of a Fucosyltransferase Gene in C. elegans—Our previous studies indicated that extracts of C. elegans contain an unusual α,1.2FT activity and in addition contained an unidentified α,1.2FT activity (7). To gain a better understanding of the types of fucosyltransferases present in C. elegans, we used a panel of acceptors to assay extracts of adult worms for possible fucosyltransferase activities with radioactive GDP-Fuc as the donor (Table I). The highest activity was found toward the vertebrate type 1 O-glycan structure Galβ1-3GalNAc-O-pNP and Galβ1-4Xylβ-O-benzyl. This disaccharide represents part of the core structure of Galβ1-4Xylβ-O-R found in vertebrate proteoglycans. We considered that the enzyme capable of fucosylating the acceptors Galβ1–3GalNAc-O-pNP and Galβ1–4Xylβ-O-benzyl in C. elegans extracts might be an α,1.2FT acting on the terminal nonreducing Gal residue rather than the penultimate GalNAc or Xyl residues.

To identify C. elegans sequences with possible homology to known α,2FTs, we compared the amino acid sequences of human, rabbit, pig, and mouse α,2FTs to identify potential conserved sequences among them. We then used these identified conserved domains in a BLAST search of the C. elegans genome database. One of the genes identified by this harbored in C. elegans cosmid EGA99 (GenBank™ accession number U80026), derived from chromosome V, encoding a predicted cDNA of 1068 bp encoding the hypothetical protein EGA99.2. A map of the gene structure is shown in Fig. 1A, and the overall gene size is 1814 bp, including 700 bp of the 5’ upstream sequence. The open reading frame is encoded within 10 exons

<table>
<thead>
<tr>
<th>Acceptor (1 μmol)</th>
<th>CE2FT-1a</th>
<th>C. elegans extracts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Galβ1–3GalNAc-O-pNP</td>
<td>14</td>
<td>ND&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Galβ1–4Xylβ-O-benzyl</td>
<td>99</td>
<td>4</td>
</tr>
<tr>
<td>Galβ1–4GlcNAc-Benzyl</td>
<td>ND</td>
<td>4</td>
</tr>
<tr>
<td>Galβ1–6GlcNAc</td>
<td>34</td>
<td>2</td>
</tr>
<tr>
<td>Galβ–O-pNP</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>Galβ–O-phenyl</td>
<td>28</td>
<td>-</td>
</tr>
<tr>
<td>Galα–O-pNP</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Galβ1–3Galβ1–4Xylβ-O-benzyl</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>Galβ1–3Galβ1–4Glc</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>Galβ1–4GlcNAcβ1–2Manp-O-pNP</td>
<td>ND</td>
<td>1</td>
</tr>
<tr>
<td>Galβ1–3GlcNAcβ1–3Galβ1–4Glc</td>
<td>0.5</td>
<td>ND</td>
</tr>
<tr>
<td>GalNAcβ1–4GlcNAcβ3–O-Methyl</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Galβ1–3GlcNAcβ1–3Galβ1–4GlcNAc</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Galβ1–4Glcβ–O-pNP</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>GalNAcβ–S-pNP</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>GalNAcβ–S-p-pNP</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Xylβ–O-Benzyl</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Glcβ–methyl-umbelliferone</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Galβ1–3(Fucα1–4)GlcNAc</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>LNF II/III (mixture)</td>
<td>ND</td>
<td>-</td>
</tr>
<tr>
<td>Galβ1–4GlcNAcβ1</td>
<td>Man</td>
<td>ND</td>
</tr>
<tr>
<td>Galβ1–4GlcNAcβ1</td>
<td>Man</td>
<td>ND</td>
</tr>
<tr>
<td>Galβ1–4GlcNAcβ1</td>
<td>6</td>
<td>ND</td>
</tr>
</tbody>
</table>

<sup>a</sup> Extracts of either transiently transfected human 293 cells expressing recombinant CE2FT-1 or mixed stage C. elegans were assayed toward the indicated acceptors at 1 μmol final concentration using radio-labeled GDP-Fuc, as described under “Experimental Procedures.”

<sup>b</sup> A unit of enzyme is the amount of enzyme capable of catalyzing the transfer of 1 μmol of sugar/min.

<sup>c</sup> ND, not detectable activity.

C. elegans α,1,2-Fucosyltransferase CE2FT-1

39825
predicted to encode a protein with 355 amino acids (Fig. 1B). The exons are small and range in size from 66 to 161 bp (average size of ~107 nucleotides). The introns are also small, ranging in size from 41 to 310 bp, and have an average size of ~101 bp. The small sizes of the exons and introns is consistent with previous observations about expressed genes in *C. elegans*.
(16). The exon/intron junctions at all introns follow the GT/AG rule, although it has been predicted that not all splicing for short introns in C. elegans requires the AG (17). There are three predicted N-glycosylation sites at positions in the predicted luminal domain of the protein at Asn residues 92, 311, and 349 within the N-glycosylation sequon Asn-X-Ser/Thr (where X does not represent Pro). There are no predicted sites for O-glycosylation at Ser and Thr residues, using the NetOGlyc 2.0 Prediction Server (available on the World Wide Web at www.cbs.dtu.dk/services/NetOGlyc/) for O-glycosyla-
tion sites, although this program is predictive for mammalian glycoproteins.

The Kyte-Doolittle hydrophilicity plot (18) of the protein sequence suggests the presence of a 16-amino acid transmembrane domain at the N terminus (Fig. 2). Thus, the encoded protein displays the typical hallmarks of most Golgi-localized glycosyltransferases (i.e. type II membrane orientation with a relatively small cytosolic N terminus and a large, extracellular, C-terminal region) (1). The most conserved region predicted between the CE2FT-1 and other α,2FTs is in the C-terminal domain (Fig. 3), indicating that this region is likely to contain the catalytic domain, as seen for other Golgi glycosyltrans-
ferases. However, overall the predicted amino acid sequence from the CE2FT-1 cDNA displays a relatively low identity (5–10%) to the α,2FT sequences in humans, rabbits, and mice at the amino acid level (Fig. 3).

Cloning of the cDNA from C. elegans Encoding CE2FT-1—Gene-specific primers were designed to amplify the entire coding sequence of the cDNA CE2FT-1 from a C. elegans cDNA λZAP library, as described under “Experimental Procedures.” Following amplification, the PCR product was TA-cloned into the vector pCR3.1 and then subcloned in frame into pcDNA4/HisMax C mammalian expression vector. A clone was isolated, and DNA sequencing confirmed the exon/intron boundaries predicted by the data base.

Expression of C. elegans CE2FT-1 cDNA in the 293T Cell—Transfection of human 293T cells with cDNA encoding CE2FT-1 results in significantly higher enzyme activity over mock-transfected 293T cells by using Gal-β-p-nitrophenol as the acceptor (Table I). The mock-transfected 293T cells have no significant endogenous α,2FT activity. To verify that the ac-
tivity is due exclusively to the expressed cDNA and not to alteration of an endogenous, unknown enzyme, we used West-
ern blot with Anti-Xpress monoclonal antibody to probe the Xpress-tagged CE2FT-1 in cell extracts from transfected 293T cells. Unexpectedly, the results indicated that the N-terminal Xpress tag fused to the N terminus of CE2FT-1 was cleaved, since little antigenic protein was detected. We then constructed an alternative full-length cDNA of CE2FT-1 with the Ca$^{2+}$-de-
pendent HPC4 epitope (9, 10) tag fused to the C terminus. Human 293T cells were transiently transfected with this con-
struct, and the transfected cells were harvested 72 h after transfection. The presence of the HPC4 epitope-tagged CE2FT-1 protein with a molecular mass of ϕ 50 kDa was identified by Western blot using the Ca$^{2+}$-dependent HPC4 monocl-
onal antibody (Fig. 4). The recombinant HPC4 epitope-tagged CE2FT-1 was purified by absorption on immobilized HPC4 (9, 10). Using GDP[3H]Fuc as the donor and Galβ-phenyl as the ac-
ceptor, the immunoabsorbed recombinant HPC4-tagged CE2FT1 generated 2467 cpm of product, whereas immunoab-
sorption from mock-transfected cells generated a background of 58 cpm. These results confirm that the protein encoded by the gene CE2FT-1 is an active α,2FT. To confirm the fucosyl linkages in the reaction product generated by CE2FT-1, the

\[ \text{GDP-Fuc} + \text{Galβ-p-nitrophenol} \rightarrow \text{CE2FT-1} \]

\[ \text{GDP} + \text{Fucα1-2Galβ-p-nitrophenol} \]

Acceptor Specificity of Recombinant CE2FT-1—To further characterize the CE2FT-1 activity, a comprehensive analysis of the acceptor specificity of CE2FT-1 was performed, using a large variety of acceptors, many of which are known to be acceptors for previously described α,2FTs. CE2FT-1 transfers fucose to the monosaccharide acceptor Galβ-p-nitrophenol but not to Galα-p-nitrophenol, indicating absolute specificity for terminal β-linked Gal residues (Table I). However, unexpectedly CE2FT-1 is completely inactive in using Galβ1–3GalNAc-p-NP as the acceptor and is also inactive in using either lactose or Galβ1–4Glcβ-O-pNP as acceptors. In addition, the enzyme is unable to transfer fucose to terminal β1–4-linked galactosyl residues in Man-containing complex-type N-glycans. The enzyme was also inactive with the acceptors containing β1–3/4-linked fucose, such as LNPlI/I and Galβ1–
3/Fucα1–4GlcNAc, indicating that the CE2FT-1 is unable to transfer Fuc in α1,2 linkage to another fucose to generate Fucα1–2Fucα1–R, as seen for at least one other invertebrate α,2FT (19). By contrast, the enzyme demonstrated a clear preference for the unusual acceptor Galβ1–4Xylβ-O-benzyl with about one-third less activity demonstrated toward Galβ1–
6GlcNAc. Importantly, neither of these acceptors has been reported to be an acceptor for previously described α,2FTs.
CE2FT-1 showed some activity toward two of the more complex acceptors with the sequences Gal\(^1\)-H\(\text{925}\)\(\text{2} \rightarrow 3\)Gal\(^1\)-H\(\text{925}\)\(\text{2} \rightarrow 4\)Xyl-O-benzyl and Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 3\)Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 4\)Glc (Table I). These two latter acceptors contain a penultimate Gal residue rather than GlcNAc, as in two of the acceptors that were inactive, the trisaccharide Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 4\)GlcNAc\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 2\)Man-O-pNP and the tetrasaccharide Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 4\)GlcNAc\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 3\)Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 4\)GlcNAc. Overall, CE2FT-1 is inactive toward all acceptors with penultimate GlcNAc or Glc residues (Table I). Thus, it is possible that the enzyme does not recognize acceptors with penultimate GlcNAc residues but prefers those with terminal Gal and a penultimate residue that is neither GlcNAc nor Glc. However, insufficient amounts of product were generated to allow proof at this time as to which of the Gal residues was modified by Fuc in reactions with CE2FT-1. Nevertheless, these results demonstrate that CE2FT-1 is unique compared with known \(\alpha\,1,2\)FTs in its acceptor specificity, since all previously described \(\alpha\,1,2\)FTs are efficient in using lactose or virtually any other acceptor with terminal \(\alpha\,1,4\)- or \(\beta\,1,3\)-linked galactose residues (20, 21). The acceptor specificity of CE2FT-1 suggests that the presence of an unsubstituted C-6 hydroxyl group on the penultimate sugar hinders enzyme recognition of the terminal Gal residue.

CE2FT-1 was compared with other endogenous potential \(\alpha\,1,2\)FTs in extracts of adult \(C.\,elegans\). These extracts contained a fucosyltransferase activity toward the same acceptors identified for CE2FT-1, but in addition, the extracts contained fucosyltransferase activity toward Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 4\)GlcNAc-O-benzyl and the O-glycan type acceptor Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 3\)GalNAc\(^1\)-O-pNP (Table I). It is likely, based on our previous studies (7), that Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 4\)GlcNAc-O-benzyl is fucosylated by the \(\alpha\,1,3\)FT CEFT-1 or related \(\alpha\,1,3\)FTs rather than an endogenous \(\alpha\,1,3\)FT activity. These results suggest that adult worms express an active form of both CE2FT-1 and potentially other fucosyltransferases active toward Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 4\)GlcNAc-O-benzyl and Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 3\)GalNAc\(^1\)-O-pNP. Interestingly, the extracts did not contain significant fucosyltransferase activities toward Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 4\)Glc\(\beta\,1\)-O-pNP. This disaccharide is an excellent acceptor for the known \(\alpha\,1,2\)FTs found in vertebrates, which suggests that CE2FT-1 has a unique acceptor specificity.

The activity of CE2FT-1 was compared with other endogenous potential \(\alpha\,1,2\)FTs in extracts of adult \(C.\,elegans\). These extracts contained a fucosyltransferase activity toward the same acceptors identified for CE2FT-1, but in addition, the extracts contained fucosyltransferase activity toward Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 4\)GlcNAc-O-benzyl and the O-glycan type acceptor Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 3\)GalNAc\(^1\)-O-pNP (Table I). It is likely, based on our previous studies (7), that Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 4\)GlcNAc-O-benzyl is fucosylated by the \(\alpha\,1,3\)FT CEFT-1 or related \(\alpha\,1,3\)FTs rather than an endogenous \(\alpha\,1,3\)FT activity. These results suggest that adult worms express an active form of both CE2FT-1 and potentially other fucosyltransferases active toward Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 4\)GlcNAc-O-benzyl and Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 3\)GalNAc\(^1\)-O-pNP. Interestingly, the extracts did not contain significant fucosyltransferase activities toward Gal\(^1\)-H\(\text{925}\)\(\text{1} \rightarrow 4\)Glc\(\beta\,1\)-O-pNP. This disaccharide is an excellent acceptor for the known \(\alpha\,1,2\)FTs found in vertebrates, which suggests that CE2FT-1 has a unique acceptor specificity.
that of the many potential α1,2FTs encoded by the *C. elegans* genome (3), none of these enzymes efficiently utilize the common acceptor Galβ1-4Glcβ1-O-pNP. To confirm that the fucosyltransferase reaction products using these acceptors were modified by endogenous α1,2FTs in the *C. elegans* extracts, we isolated the products generated by the *C. elegans* extracts using three of the acceptors utilized by the recombinant CE2FT-1. The acceptors we chose were Galβ1,3Galβ1,4Xyl β1-O-Bzl, and Galβ1,3Galβ1,4Xyl-O-Bzl. These α1,2-fucosylated products generated by the endogenous α-fucosyltransferases in *C. elegans* extracts were treated with either α1,2-fucosidase, α1,3/4-fucosidase, or α1,6-fucosidase. α1,2-Fucosidase was quantitatively released from all of the products by α1,2-fucosidase, and only minimal radioactivity was released by the other fucosidases (Fig. 5B). These results indicate that the predominant endogenous α-fucosyltransferase in *C. elegans* extracts is an α1,2FT with the type of acceptor specificity exhibited by recombinant CE2FT-1.

**Expression of CE2FT-1 Promoters at Various Stages in C. elegans Development**—Very little is known about the functions of glycosyltransferases in *C. elegans* (7, 22–26), and no information is available about expression of potential α1,2FT genes. To examine whether CE2FT-1 is expressed during development and its pattern of expression, we first analyzed transcript levels by relative RT-PCR using equivalent amounts of total RNA from different developmental stages (Fig. 6A). The CE2FT-1 gene is expressed at all developmental stages with little relative difference in expression compared with α-actin. Since there are more than 22 putative α1,2FTs in *C. elegans*, and some of them with more than 85% homology to CE2FT-1, we performed quantitative RT-PCR to analyze the mRNA expression level at different stages. The results indicated that there is little relative difference in expression at different stages (see Fig. 6, B and C).

To localize the expression of the gene, we used reporter promoter constructs driving expression of the coelenterate green fluorescent protein (GFP) in the vector pPD95.67/CE2FT-1-prom. The promoter construct consisted of a 5'-700-bp fragment containing all the upstream genomic sequence, which should encompass the putative promoter region immediately upstream of the initiation site for CE2FT-1 translocation (Fig. 1A). In rats, it has been shown that expression of the α1,2FT gene is regulated by functional promoter elements within the 5'-flanking region of that gene (27). Likewise, in *C. elegans*, gene expression is usually regulated by promoter elements in 5'-untranslated regions of genes (14). Expression of the GFP reporter gene in the transgenic worms injected with CE2FT-1 promoter construct pPD95.67/CE2FT-1-prom was examined (Fig. 7, A–H). The results indicate that the promoter for CE2FT-1 is specifically activated in intestinal cells beginning at the 2-fold embryonic stage embryo (Fig. 7, A–D). Promoter function was identified through all later developmental stages with a highly restricted GFP expression in the 20 cells of the adult intestine (Fig. 7, E–H). In these studies, the 20 cells aligned to form the entire intestine were clearly visually identified by transmitted light microscopy compared with the fluorescent images (Fig. 7, E and F). There was a surprising restriction of CE2FT-1 expression in the intestinal cells, and all cells had equivalent expression levels based on visual observations. Virtually no GFP fluorescence was observable in nonintestinal cells in adult worms. In control experiments, we used 5' upstream promoter regions of other potential α1,2FT genes identified in our Blast search of the *C. elegans* genome. GFP promoter analyzers with these other constructs gave no restricted GFP expression in the 20 cells of the intestine, although expression was observed in other cells (data not shown).

**DISCUSSION**

Glycoconjugates with the sequence Fuca1-2Gal-R, as in the ABO(H) blood group antigens, are expressed in a wide variety of tissues and commonly occur in the digestive mucosae of a large number of species ranging from amphibians to mammals (1, 28). The synthesis of the α1,2-fucosylated structures in humans is catalyzed by two α1,2FTs, encoded by *FUT1* and *FUT2* (also known as the Se gene), which differ slightly in acceptor specificity (29–33). In addition, humans possess another allele termed *SEC1*, which is a pseudogene (33). The orthologues for the human *FUT1*, *FUT2*, and *SEC1* genes have been described in rodents and termed *FTA*, *FTB*, and *FTC*, respectively (34). Little is known about the functions of the α1,2-fucosyltransferases and their cognate structure Fuca1-2Gal-R in animals.

In all species so far examined, *FUT2* (Se) is expressed in the
The ileal epithelium of adult mice in normal growth cages is characterized by expression of glycoconjugates with the sequence Fuc1–2Gal-R; however, the intestines of germ-free mice are deficient in such fucose-containing structures (35). Inoculation of the germ-free mice with Bacteroides thetaiotaomicron significantly stimulated expression of Fuc1–2Gal-R structures and transcripts encoding an α1,2FT (35), which could be derived from either SEC1 or FUT2 (36). It was recently shown in germ-free mice that administration of intestinal microbes stimulated expression of MFUT-II, the ortholog of human FUCT2 (37, 38). The bacteria-dependent commensal stimulation of host fucosylation is a complex phenomenon involved in generating metabolic fucose to sustain bacterial flora (39). We do not yet know whether expression of CE2FT-1 in the intestinal cells of C. elegans is induced by bacterial factors, but since the worms are constitutively grown using E. coli as a nutrient source, there may be a role for bacterial products in regulating gene expression.

Other functions of Fuc1–2Gal-R have also been proposed. H-type 1 oligosaccharides with the structure Fuca1–2Galβ1–3GlcNAc-R are differentially expressed during mouse embryogenesis and have been proposed to be involved in early implantation events (40, 41). The α1,2FT activity varies during the estrous cycle and is elevated ~5-fold during estrous (42). However, implantation appears to be normal in mice genetically deficient in both murine homologs of FUT1 and FUT2 (43).

Several other genes are known to be specifically expressed in the intestinal cells of C. elegans, such as the N-acetylglucosaminyltransferase I gene Gly14 (24) and the genes encoding

---

**FIG. 6.** The relative RT-PCR of CE2FT-1 mRNA from C. elegans at different stages. A, total C. elegans RNA from the populations of L1, L2–L4, and adult stage were prepared and used as templates in first strand cDNA synthesis. A 500-bp positive control from the RT-PCR kit was used as a control for the RT-PCR method (Line Positive Control). The arrows indicate the expected size of α-actin and CE2FT-1. Size markers in bp are indicated. B, the quantitative real-time RT-PCR of CE2FT-1 mRNA from C. elegans at different stages. Total C. elegans RNA from the populations of L1, L2–L4, and adult stage were prepared and used as templates in first strand cDNA synthesis. A 106-bp cDNA was amplified. The expression level was indicated by threshold cycles. C, confirmation of a single product of amplification in quantitative real time RT-PCR was performed on 1.2% agarose gel. A 106-bp cDNA fragment was amplified from CE2FT-1, and a 100-bp cDNA fragment was amplified from actin. Size markers in bp are indicated.

**FIG. 7.** Expression of CE2FT-1::GFP in C. elegans. Shown are fluorescence-based microscopic images of transgenic C. elegans expressing the CE2FT-1::GFP reporter construct (A, C, E, and G) and light microscopic images of the corresponding C. elegans profile (B, D, F, and H). Expression of CE2FT-1::GFP reporter in late embryo (A and B), egg (C and D), larva (E and F), and adult (G and H) is shown.
C. elegans α1,2-Fucosyltransferase CE2FT-1

Recent studies have shown that the enzyme α1,2-fucosyltransferase (FucT-α1,2) is involved in the biosynthesis of glycosaminoglycans (GAGs) in mammalian cells. The enzyme catalyzes the transfer of fucose from GDP-fucose to various acceptors, including glycans of the N-linked oligosaccharides. The activity of this enzyme is regulated by the presence of specific transcriptional regulatory elements, such as GATA-like motifs, which are found in the promoters of many fucosyltransferase genes.

In this study, we have identified a novel fucosyltransferase gene, termed CE2FT-1, in C. elegans. This gene is highly conserved in prokaryotic, microbial, and mammalian organisms. The CE2FT-1 gene is not known at present.

We also inspected the 5′-untranslated sequences of some other C. elegans glycosyltransferase genes for the presence of GATA-like elements that might contribute to intestinal expression. The gene encoding CEFT-1, an α1,3FT that synthesizes Fucα1-3Galβ1-4GlcNAc-R linkages, lacks GATA elements within −500 of the translation initiation site. Our recent expression studies on CEFT-1 expression show that it is not expressed in intestinal cells of adult C. elegans but is expressed in specific neural cells. However, an open reading frame within cosmid EGA9,3, which is in the same locus as CEFT-1, also encodes a putative α1,2FT and contains one GATA-like element at −85 upstream of the translation initiation site, suggesting that it may also be specifically expressed in intestinal cells. Interestingly, the Gfy14 gene encoding an N-acetylgalcosaminyltransferase I (CN) was shown to be expressed specifically in intestinal cells. Our inspection of the 5′-untranslated region of the Gfy14 gene reveals that it contains two GATA-like motifs within −300 of the translation initiation site. These comparative results suggest that the GATA-like elements in the 5′-untranslated sequences of some glycosyltransferase genes, such as CE2FT-1, might be important in regulating their intestine-specific expression.

The amino acid sequence of CE2FT-1 displays a low identity (5–10%) to the sequences of previously described vertebrate α1,2FTs. Previous studies have identified three motifs in the C-terminal regions of the vertebrate α1,2FT gene family termed I, II, and III (3) and in prokaryotic (Golgi-resident and constitute the catalytic domain based on associated with its inability to transfer to lactose and other Fucα1-3GlcNAc-R acceptors as do the mammalian α1,2FTs (49).

The roles of these α1,2FT motifs are not known, but their presence in the catalytic domain suggests that they may be involved in interactions with either acceptor glycans or GDP-Fuc. Although purely speculative, the weak conservation of motif II in CE2FT-1 might be associated with its inability to transfer to lactose and other simple Galβ1–4GlcNAc-R acceptors as do the mammalian α1,2FTs.

Little is currently known about the roles of glycoconjugates in C. elegans development. Two of the eight sqv genes regulating vulval epithelial invagination have homologies to glycosyltransferase genes, including sqv-3, homologous to members of the β1,4-galactosyltransferase gene family, which may be involved in elongation of the glycosaminoglycan core linkage to Xyl-Ser, and sqv-8, homologous to two vertebrate β1,3-galactosyltransferases (25, 26, 50–52). A third gene, sqv-7, was recently shown to encode a sugar nucleotide transporter (53).

All of these genes have been shown to be functionally involved in glycosaminoglycan biosynthesis in vivo (54).

The C. elegans genome contains at least 22 homologues of α1,2FT genes, at least four genes encoding α1,3FT, one α1,6FT gene, and probably many more fucosyltransferase genes yet to be defined (3, 7, 55) (also see the Web site of the Consortium for Functional Glycomics at functionalglycomics. mit.edu/cgi-bin.functional_glycomics/glyt/glyt_index.cgi). For example, a cytosolic α1,2FT named Skp1 was recently identified in Dicystostelium (56), which lacks significant homology to any previously identified fucosyltransferases. Skp1 is also significantly different enzymatically and structurally from CE2FT-1 and other known mammalian α1,2FT, in that Skp1 requires divalent cations and reducing conditions for activity, is cytosolic rather than compartmentalized, lacks a membrane anchor domain, and lacks several sequence motifs that are highly conserved in prokaryotic, microbial, and mammalian α1,2FTs (3).

The CE2FT-1 we have identified may be partly responsible for synthesizing the unusual Fucα1-2Galβ1–6Gal-R linkages recently reported in complex-type O-glycans in adult C. elegans glycoproteins (8). The α1,2-linked fucose occurred in complex structures, such as Fucα1–2(Galβ1–6)Galβ1–3(Fucα1–2)Galβ1–3(Glcβ1–6)GalNAcβ1–4GlcNAcβ1–Ser/Thr, in which the Fuc residue was 2-O-methylated. CE2FT-1 acts particularly well on β1–6-branched Gal residues, as in Galβ1–6GlcNAc (Table I), but not Galβ1–4GlcNAc residues (Table I). Interestingly, no Galβ1–4GlcNAc residues were observed in O-glycans from C. elegans (8). In addition, it is possible that the predicted glycan Galβ1–4Xylβ1–Ser generated by enzyme β1,4-galactosyltransferase encoded by the sqv3 gene (54) could be an endogenous acceptor for CE2FT-1, since we found that CE2FT-1 shows the highest acceptor activity toward Galβ1–4Xylβ1–O-benzyl. However, in one study, no α1,2-fucosylated glycosaminoglycan core structures were found in studies on complex-type O-glycans of C. elegans (8), but it is possible that such modifications can be restricted to core structures in only the gut cells of the organism. Further studies will be required to define the exact role of CE2FT-1 and each of the other members of this large family of α1,2FTs in specific glycoconjugate synthesis and worm development.

Acknowledgments—We thank Dr. Janet Duerr, Wietse Schiphorst, Kenneth Hatter, Kenneth Ju, and Dr. Kwame Nyame for help and technical support in these studies. We especially thank Dr. Robert Barstead of the Oklahoma Medical Research Foundation for advice and help.

REFERENCES

Molecular Cloning and Characterization of a Novel α1,2-Fucosyltransferase (CE2FT-1) from Caenorhabditis elegans
Qinlong Zheng, Irma Van Die and Richard D. Cummings

doi: 10.1074/jbc.M207487200 originally published online August 5, 2002

Access the most updated version of this article at doi: 10.1074/jbc.M207487200

Alerts:
  - When this article is cited
  - When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 55 references, 24 of which can be accessed free at http://www.jbc.org/content/277/42/39823.full.html#ref-list-1