Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 10697–10702, October 1996
Biochemistry

Expression cloning of Forssman glycolipid synthetase: A novel member of the histo-blood group ABO gene family

(glycosyltransferase/murine development/ABO blood group)

DAVID B. HASLAM* AND JACQUES U. BAENZIGER†

*Department of Pediatrics, Child Health Research Center of Excellence in Developmental Biology, and †Departments of Pathology and Cell Biology, Washington University School of Medicine, St. Louis, MO 63110

Communicated by Stuart A. Kornfeld, Washington University School of Medicine, St. Louis, MO, July 15, 1996 (received for review April 23, 1996)

ABSTRACT **A phenotypic cloning approach was used to isolate a canine cDNA encoding Forssman glycolipid synthetase (FS; UDP-GalNAc:globoside α-1,3-*N*-acetylgalactosaminyltransferase; EC 2.4.1.88). The deduced amino acid sequence of FS demonstrates extensive identity to three previously cloned glycosyltransferases, including the enzymes responsible for synthesis of histo-blood group A and B antigens. These three enzymes, like FS, catalyze the addition of either *N*-acetylgalactosamine (GalNAc) or galactose (Gal) in α-1,3-linkage to their respective substrates. Despite the high degree of sequence similarity among the transferases, we demonstrate that the FS cDNA encodes an enzyme capable of synthesizing Forssman glycolipid, and demonstrates no GalNAc or Gal transferase activity when closely related substrates are examined. Thus, the FS cDNA is a novel member of the histo-blood group ABO gene family that encodes glycosyltransferases with related but distinct substrate specificity. Cloning of the FS cDNA will allow a detailed dissection of the roles Forssman glycolipid plays in cellular differentiation, development, and malignant transformation.**

Many distinct carbohydrate structures are found at the cell surface conjugated to lipids or proteins. In a number of instances, these structures have been found to be characteristic of cell lineage, differentiation state, or rate of proliferation. Even though glycolipids have been implicated in regulation of cellular differentiation, proliferation, and cell death (1, 2), definitive evidence for their role in these processes awaits the development of tools that will allow precise modulation of glycolipid expression *in vivo*. We have therefore undertaken the cloning of key glycosyltransferases involved in the synthesis of glycolipids that are hypothesized to modulate cellular differentiation and contribute to development.

Forssman glycolipid (FG; globopentosylceramide; IV^{3}αGalNAcGb4; GbO$_5$) is expressed on a variety of cell lineages in a differentiation-specific manner. Synthesis of this glycolipid, which occurs via transfer of *N*-acetylgalactosamine (GalNAc) in α-1,3-linkage to Gb4Cer (Table 1 and ref. 3), is tightly regulated during murine development and has been suggested to be involved in tissue morphogenesis (4). Furthermore, in humans FG is expressed in certain disease states, such as lung and gastrointestinal tumors, where it may serve as a marker for malignant transformation (5, 6).

We report here the cloning of a canine cDNA that encodes Forssman glycolipid synthetase (FS). A phenotypic cloning approach allowed isolation of the FS gene from a cDNA library prepared from canine kidney cells known to produce abundant FG. Interestingly, the FS cDNA clone demonstrates extensive sequence identity to glycosyltransferases that catalyze the final step in synthesis of histo-blood group A and B antigens, raising several implications regarding the evolutionary processes giving rise to these carbohydrate structures. The studies presented herein will form the basis of a detailed dissection of the roles FG plays in cellular and developmental biology.

MATERIALS AND METHODS

Materials, Antibodies and Cell Culture. Glycolipid substrates were purchased from Sigma with the exception of synthetic Histo H glycolipid (Accurate Chemicals). The LacNAc substrate (Galβ1,4GlcNAcβ1,3Galβ1,4Glcβ-MCO) was synthesized from GlcNAcβ1,3Galβ1,4Glcβ-MCO (kindly provided by Ole Hindsgaul, University of Alberta, Edmonton), using bovine β(1, 4)-galactosyltransferase as described (7).

Monoclonal antibody M1/22.21 (rat IgM anti-Forssman antigen) culture supernatant was prepared from the monoclonal hybridoma and was obtained from ATCC; tissue culture supernatants were also prepared from monoclonal hybridoma MC631 provided by the Developmental Studies Hybridoma Bank [maintained by the Department of Biology, University of Iowa (Iowa City) under contract number N01-HD-2-3144 from the National Institute of Child Health and Human Development]; clone 81 FR 2.2 (mouse monoclonal anti-blood group A) was purchased from Dako. Biotinylated lectin BSI-B$_4$ was obtained from Sigma. Cell lines MDCK II and COS-1 were maintained in Dulbecco's modified Eagle's medium (DMEM) that was supplemented with 10% fetal calf serum.

MDCK II cDNA Library Construction. mRNA was isolated from 1×10^8 MDCK II cells using oligo(dT) affinity chromatography. MDCK II mRNA (5 μg) was used in each of two cDNA synthesis reactions using either oligo(dT) or random hexamer primers. Nonpalindromic *Bst*XI adapters were ligated to the cDNA, which was then subjected to size selection (>600 bp) by agarose gel electrophoresis. The cDNA samples were ligated into *Bst*XI-digested vector pcDNAI (Invitrogen) and then electorporated into *Escherichia coli* strain MC1061/P3, resulting in 1.1×10^6 [oligo(dT)] and 1.3×10^6 (random hexamer) independent primary clones.

Isolation of FS cDNA. Plasmid DNA was prepared from pooled oligo(dT) and random hexamer MDCK II cDNA libraries and was used to transfect COS-1 cells. Phenotypic cloning of the FS gene was performed by transfecting 12 dishes (100 mm^3) of COS-1 cells at 70% confluence with 4 μg of pooled cDNA and 50 μl of lipofectamine reagent (GIBCO/BRL) per dish. After 48 h in culture, cells were harvested by trypsinization, pelleted by centrifugation, and resuspended in 12 ml of panning buffer (DMEM plus 10% fetal calf serum and 2% goat serum). Cell suspension (6 ml) was aliquoted into each of two panning dishes that had been coated with goat anti-rat IgM antibodies (Pierce) followed by monoclonal antibody M1/22.25 (8). Following addition of cell suspensions,

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: α-(1,3)GT, α-(1,3)-galactosyltransferase; FS, Forssman glycolipid synthetase; FG, Forssman glycolipid; SSEA-3, stage specific embryonal antigen-3 (Gal-Gb4Cer.

Data deposition: The sequence reported in this paper has been deposited in the GenBank data base (accession no. U66140).

Table 1. Structures of the globoseries glycolipids and the A,B,O histo-blood group antigens

Name	Structure	Reactivity with MC631	Reactivity with M1/22.25
Globoside (Gb4Cer)	GalNAcβ-1,3Galα-1,4Galβ-1,4GlcCer	+	−
Gal-Gb4Cer, (SSEA-3; IV^{3}β-GalGb4)	Galβ-1,3GalNAcβ-1,3Galα-1,4Galβ-1,4GlcCer	+++	−
Forssman antigen (IV^{3}α-GalNAcGb4)	**GalNAcα-1,3**GalNAcβ-1,3Galα-1,4Galβ-1,4GlcCer	−	+++
Histo-blood group A antigen	**GalNAcα-1,3**[Fucα-1,2]Galβ-1,4GlcNAc−	−	−
Histo-blood group B antigen	**Galα1,3**[Fucα-1,2]Galβ-1,4GalNAc	−	−
Gal(α-1,3)Gal epitope	**Galα-1,3**Galβ-1,4GlcNAc−	−	−

The reactivity of the A, B, O histo-blood group with monoclonal antibodies MC631 and M1/22.25 is indicated. +, minimal reactivity; +++, strong reactivity; −, no reactivity.

the dishes were incubated at 4°C for 1 h with occasional rocking. Unbound cells were removed by rinsing six times with panning buffer and gentle aspiration. Plasmid DNA was recovered from adherent cells by Hirt extraction (9) by adding 600 μl of Hirt lysis buffer (10 mM Tris, pH 7.4/10 mM EDTA/0.6% SDS) per dish. Lysed cells were aliquoted into Eppendorf tubes and NaCl was added to a concentration of 1 M. After incubating 6 h at 4°C, debris was removed by centrifugation and the supernatant was subjected to phenol/chloroform and then chloroform extraction. Plasmid DNA was ethanol precipitated following the addition of 5 μg of glycogen. The recovered cDNA was used to electroporate *E. coli* MC1061/P3. Plasmid DNA was prepared from transformed bacteria and 1 μg was used to transfect each of three dishes of COS-1 cells, which were then subjected to panning and Hirt extraction as described above. After a third round of panning and Hirt extraction, bacteria were electroporated and serially diluted before plating. Transformants were pooled from replica plates representing each dilution, and plasmid DNA was prepared from each and was used to transfect COS-1 cells. All of the dilutions, including one composed of only 154 colonies, directed the expression of Forssman antigen in COS-1 cells as detected by indirect immunofluorescence. Pools of 12 individual colonies were prepared from the master plate containing 154 colonies. Plasmid DNA prepared from 3 of 13 pools resulted in reactivity with antibody M1/22.25 when transfected into COS-1 cells. Plasmid DNA prepared from the 12 colonies comprising one positive pool were then tested individually. Two of 12 clones (pFS-7 and pFS-10) resulted in COS-1 reactivity with monoclonal antibody M1/22.25.

Nucleotide Sequence Determination. Plasmid DNA was prepared from clones pFS-7 and pFS-10. Cycle sequencing was performed using Sp6 and T7 or gene-specific primers, fluorescently labeled dideoxynucleotide dye terminators, and *Taq* polymerase under conditions recommended by the manufacturer (Applied Biosystems). The entire 1.97-kb insert from pFS-7 was determined in both directions. Additional sequencing reactions were performed to resolve ambiguous nucleotide designations. Approximately half of the insert contained in plasmid pFS-10 was sequenced and was found to be identical to that of pFS-7 cDNA. Comparison of the FS gene with sequences in the GenBank data base was performed using the program BLAST. Sequence analyses, including translation of open reading frames and hydrophobicity plotting of amino acid sequence, were performed using the GeneWorks software package.

Glycolipid Immunolabeling. Crude lipid extracts were prepared from approximately 2×10^6 MDCK II cells or 2×10^7 COS-1 cells transfected with pFS-7 or vector pcDNAI. Cells were harvested by trypsinization, pelleted by centrifugation, and then resuspended in a volume of methanol equal to the cell pellet. One volume of chloroform was added and, after vortexing vigorously, methanol was added dropwise until organic and aqueous layers resolved into a single phase (approximately one additional volume of methanol). Cellular debris was removed by centrifugation and the organic supernatant was evaporated under a stream of nitrogen. The lipid residue was resuspended in 60 μl of chloroform/methanol/water (65:35:8) and 20 μl of each was spotted onto a TLC plate. After developing in chloroform/methanol/water (65:35:8), the plate was air-dried and blocked with 5% BSA in PBS. Forssman antigen was detected by TLC immunolabeling and detected by ECL as previously described (10). Briefly, the plate was overlaid with antibody M1/22.25 (1:500 dilution of tissue culture supernatant in 5% BSA in PBS). After incubating for 1 h at room temperature, the plate was rinsed three times in PBS, and then overlaid with peroxidase-conjugated goat anti-rat IgM antibodies (1:5000 in 5% BSA in PBS) and incubated for a further 30 min at room temperature. The plate was then washed seven times for 5 min each with 0.05% Tween 20 in PBS, and then immersed for 1 min in chemiluminescence reagent (DuPont/NEN) and immediately exposed to Kodak XAR film.

Fluorescence-Activated Cell Sorter Analysis (FACS). COS-1 cells transfected with pFS-7 or pcDNAI were harvested by trypsinization, and then labeled with antibody M1/22.25 (anti-Forssman), 81 FR 2.2 (anti-A antigen), or lectin BSI-B_4 diluted 1:50 in labeling buffer (5% FBS and 2% goat serum in PBS). After rinsing, secondary antibody (fluorescein isothiocyanate-labeled goat anti-rat IgM or anti-mouse IgG) or phycoerythrin-conjugated streptavidin was added at a 1:100 dilution in labeling buffer and incubated for 30 min at 4°C. After washing three times with labeling buffer, cells were resuspended in 1% paraformaldehyde in PBS and subjected to FACS analysis.

Enzyme Assays. Approximately 5×10^7 COS-1 cells transfected with pFS-7 or pcDNAI were harvested by trypsinization, resuspended in 1 ml ice cold 100 mM Mes (pH 6.7), and disrupted by sonication. Nuclei were removed by sedimentation for 5 min at 6000 × *g*. The supernatants were brought to a final concentration of 20 mM $MnCl_2$ and incubated on ice for 30 min, and the aggregated membranes were pelleted by centrifugation for 5 min at 6000 × *g*. Membrane pellets were resuspended in 500 μl of 1% Triton X-100 in 100 mM Mes and incubated for 1 h at 4°C with gentle rocking. Insoluble material was removed by sedimentation for another 5 min at 6000 × *g*, and protein concentrations were determined by the Bio-Rad *DC* assay. Mes buffer containing 1% Triton X-100 was added as necessary to equalize protein concentrations. Enzyme reactions were performed in a 100 μl volume containing 100 mM Mes (pH 6.7), 10 mM $MnCl_2$, 25 μl of membrane extract (0.2 mg/ml protein), 5 μM ^{3}H-labeled nucleotide sugar (1000 cpm per pmol), and 20 μmol of substrate. After incubating at 37°C for 2 h, reactions were stopped by the addition of 1 ml ice-cold water. Glycolipid products were separated from unincorporated nucleotide sugar by passing over a reverse phase Sep-Pak C_{18} column, washing with 20 ml of water and 10 ml of 20% methanol in water, and then eluting with 3 ml of 100% methanol. One half of the eluted sample was evaporated under reduced pressure, resuspended in 50 μl of chloroform/methanol (2:1), and then spotted onto TLC plates along with a lane of glycolipid standards (5 μg each). After they were developed in chloroform/methanol/water (65:35:8), the plates were air-dried, the lane of glycolipid standards was cut off, and the plates were sprayed with En^{3}Hance and exposed to Kodak XAR film for 12 to 72 h. Glycolipid standards were visualized by orcinol staining.

RESULTS

Isolation of a Canine cDNA Encoding FS. We isolated a cDNA encoding FS using the phenotypic cloning strategy originally devised by Seed and Aruffo (11) and subsequently modified for identification of glycosyltransferase gene products (12–18). MDCK II cells, a subline of MDCK dog kidney cells known to contain FG (19), were used as the source of FS mRNA. COS-1 cells were chosen as the recipient cell line since monkey kidney, from which COS-1 cells were derived, is known to express globoseries glycolipids. Using TLC immunooverlay with monoclonal antibodies MC631 and M1/22.25, we demonstrated that COS-1 cells produce Gb4Cer and Gal-Gb4Cer (SSEA-3; ref. 20) but fail to produce FG (data not shown). Following three rounds of enrichment by transfection with MDCK II cDNA, panning on dishes coated with antibody M1/22.25, and rescue of plasmid DNA by Hirt extraction, a pool of 154 bacterial colonies was identified that resulted in COS-1 reactivity with antibody M1/22.25. These colonies were subdivided until two independent clones were identified that directed the expression of anti-Forssman reactivity when transfected into COS-1 cells.

cDNA Sequence Analysis. Nucleotide sequencing of plasmid pFS-7 revealed the presence of an 1965-bp cDNA insert that encodes a predicted open reading frame consisting of 347 amino acids (Fig. 1*A*). A putative cytoplasmic domain consisting of five amino acids is followed by a hydrophobic sequence consisting of 22 residues that is predicted to serve as a transmembrane domain. This domain is followed by a putative 320 residue catalytic domain (Fig. 1*B*). Thus, the protein encoded by the pFS-7 cDNA is predicted to have a type II transmembrane topology, as do all other Golgi-resident glycosyltransferases cloned to date (21).

FS Is a Member of the Blood Group ABO Gene Family. Comparison of the predicted peptide sequence of FS with the GenBank data base revealed similarity to three previously cloned glycosyltransferases, all of which transfer galactose (Gal) or GalNAc in α-1,3 linkage to their respective substrates (Fig. 2*A*). The highest sequence identity was seen with the histo-blood group A and B transferases (42% amino acid

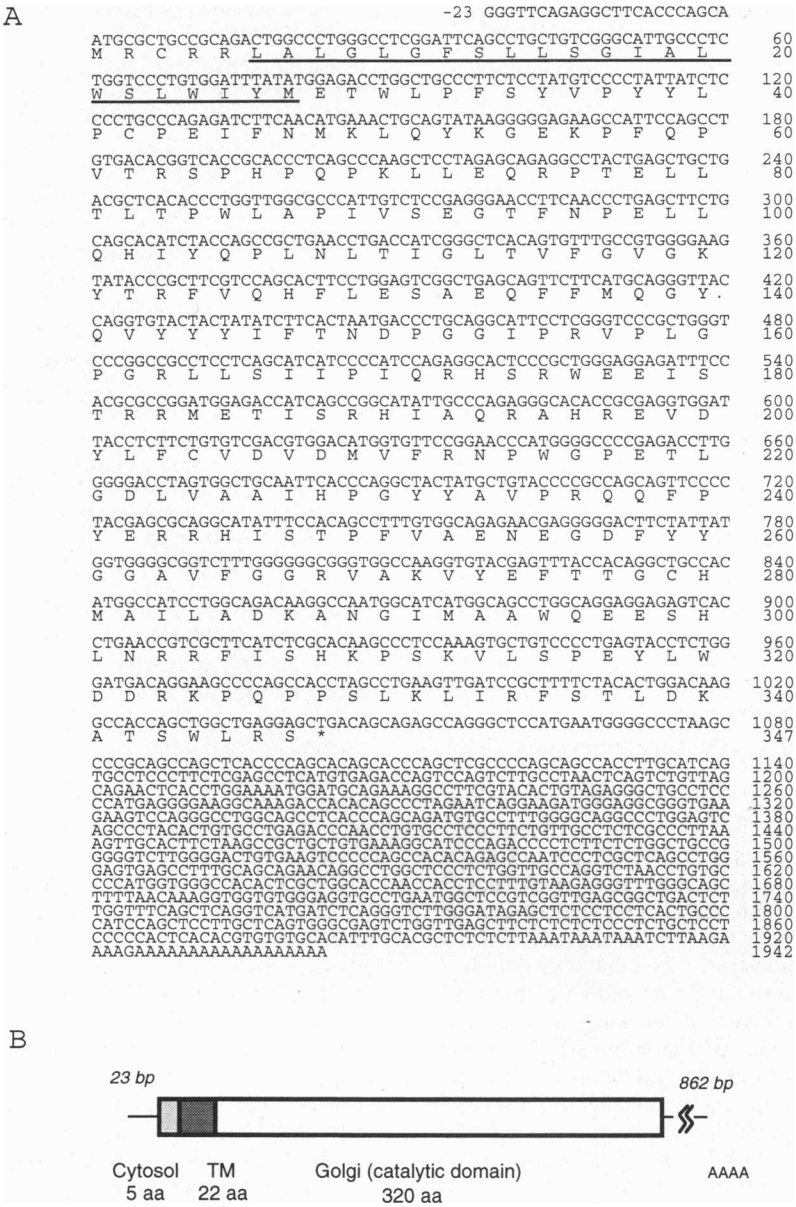

FIG. 1. (*A*) FS nucleotide and deduced amino acid sequence of cDNA contained in plasmid pFS-7. The putative transmembrane domain is underlined. (*B*) Schematic representation of the FS cDNA sequence and relationship to the predicted protein domains of FS. Untranslated nucleotides are indicated by a thin line, and the coding region is indicated by a rectangle. aa, Amino acids; TM, transmembrane domain; AAAA, polyadenylylation tail.

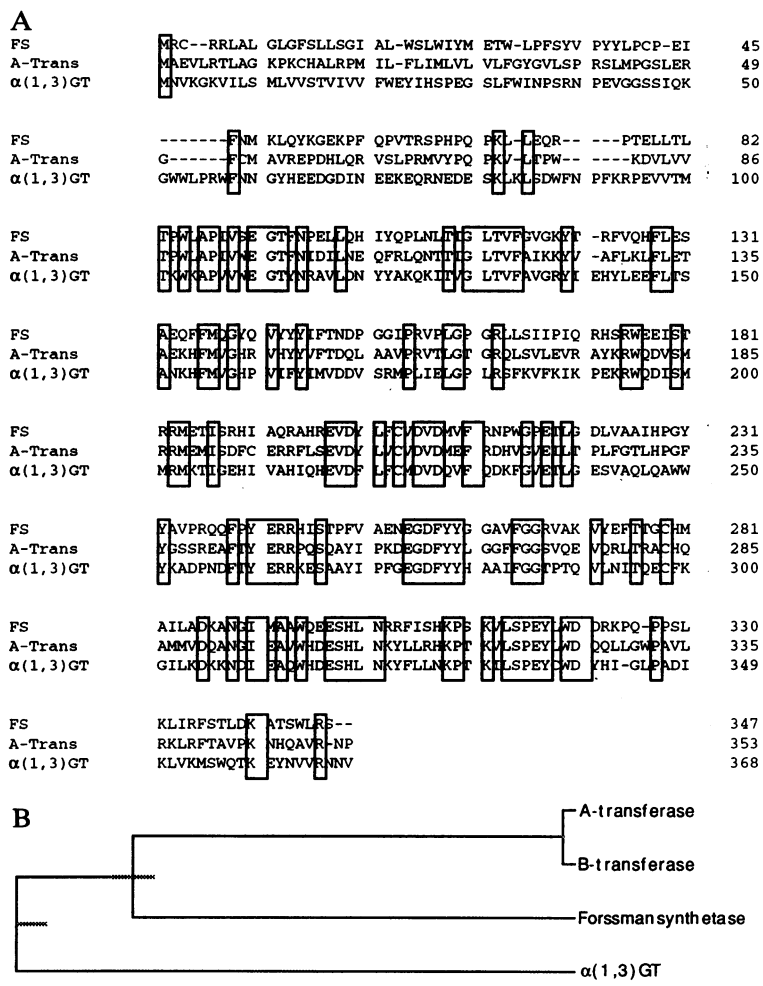

FIG. 2. FS, A-transferase, and α-(1,3)GT are homologous. (*A*) Alignment of the FS, A-transferase, and α-(1,3)GT sequences. Dashes indicate gaps introduced to optimize alignment. Identical residues in all three enzymes are boxed. (*B*) Dendogram representing the degree of sequence similarity between the peptide sequences of FS, A- and B-transferases, and α-(1,3)GT. Horizontal length of lines from their branch point represents the degree of difference between two sequences, but does not necessarily correspond with time elapsed since evolutionary divergence. The unweighted pair group method analysis (UPGMA) was used to construct the tree (GeneWorks software package).

sequence identity to both), which transfer GalNAc and Gal, respectively, to H-type acceptors on either glycolipids or glycoproteins to create the blood group A and B antigens (Fig. 2*B*). Extensive similarity was also demonstrated to α-1,3-galactosyltransferase [α-(1,3)GT] (35% amino acid sequence identity), which adds Gal in α-1,3 linkage to terminal β-4-*N*-acetyllactosamine structures.

Transfection of COS-1 Cells with pFS-7 Results in Anti-Forssman Reactivity. Given the high degree of sequence identity between the canine cDNA encoding FS and the transferases listed above, we undertook a series of experiments to demonstrate that the cloned cDNA encodes an enzyme capable of transferring GalNAc to Gb4Cer to create FG and that lacks the activities demonstrated by related enzymes.

COS-1 cells transiently transfected with the FS cDNA were incubated with antibodies recognizing Forssman antigen (M1/22.25), A-antigen (81 FR 2.2), and the lectin BSI-B_4 (which recognizes terminal α-1,3-linked Gal), and then were analyzed by FACS. Cells transfected with vector alone did not react with antibody M1/22.25, whereas 52% of cells transfected with pFS-7 were intensely labeled with this antibody, indicating *de novo* production of a Forssman-reactive antigen (Fig. 3). In contrast, the intensity and percentage of cells reactive with antibody specific for the A antigen was the same following transfection with pFS-7 or pcDNAI (20.7% and 16.3%, respectively; Fig. 3). This indicated that endogenous A-transferase activity is present in COS-1 cells and is not altered by expression of the FS cDNA. No reactivity with lectin BSI-B_4 was seen with cells transfected with pFS-7 or pcDNAI, indicating the absence of α-(1,3)GT or B-transferase activity, since this lectin recognizes blood group B structures (22) and it has previously been demonstrated that BSI-B_4 reacts with COS-1 cells expressing α-(1,3)GT (12). The cDNA clone encoded by plasmid pFS-7 is therefore capable of directing the expression of a Forssman-reactive antigen in COS-1 cells and does not show evidence of A-transferase or α-(1,3)GT activity.

COS-1 Cells Transfected with pFS-7 Produce FG. To demonstrate that COS-1 cells transfected with pFS-7 produce FG rather than an antigen cross-reactive with antibody M1/22.25, glycolipids were extracted from cells transfected with the FS cDNA or pcDNAI and examined by TLC immuno-overlay using M1/22.25. A strong band of anti-Forssman reactive material that comigrated with authentic FG was detected in glycolipid extracts from cells transfected with pFS-7, but not in cells transfected with pcDNAI. Thus, the Forssman-reactive antigen produced after transfection with pFS-7 has the same properties as authentic FG (Fig. 4).

pFS-7 Encodes a Gb4Cer:*N*-acetylgalactosaminyltransferase. COS-1 cells transfected with pFS-7 or vector alone were examined for the presence of FS activity using Gb4Cer as the acceptor and UDP-[^{3}H]GalNAc as the nucleotide sugar donor. No activity was detected in the media (data not shown), whereas Triton X-100 extracts of COS-1 cells transfected with pFS-7 but not with

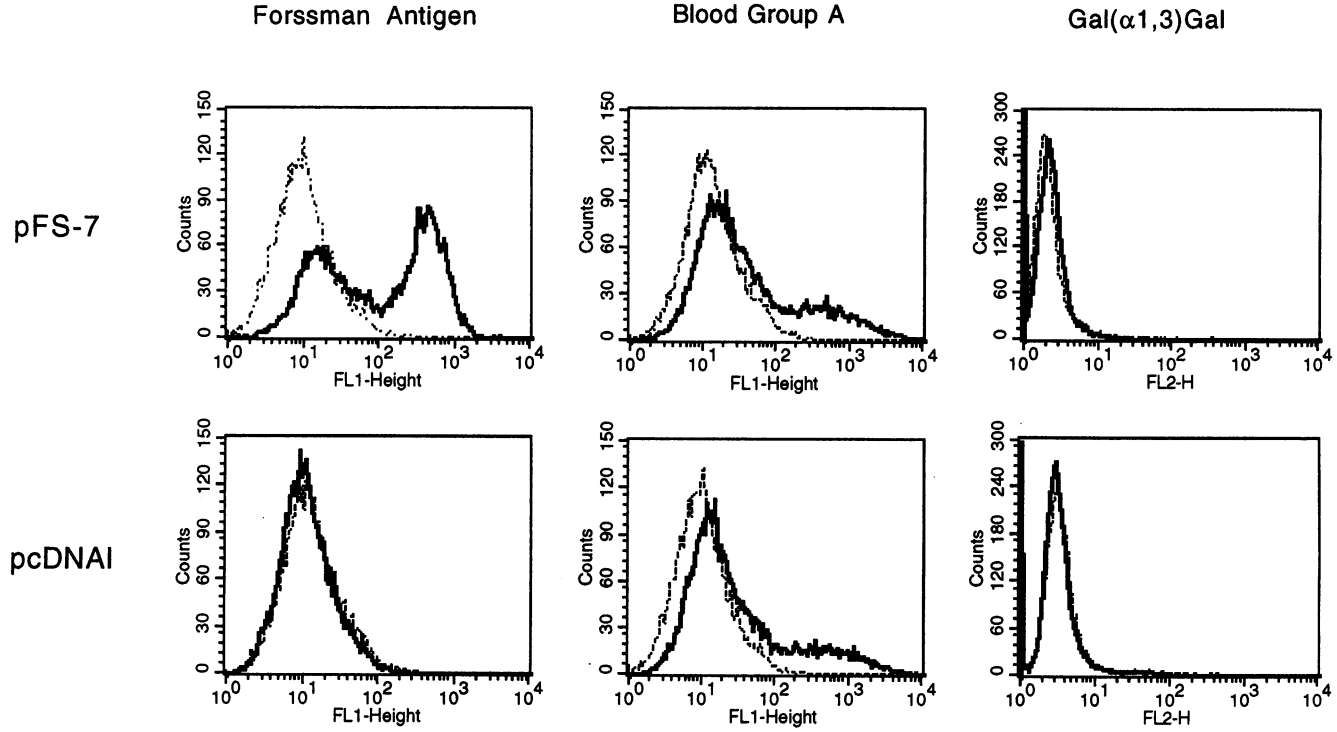

FIG. 3. COS-1 cells transfected with FS cDNA react with monoclonal anti-Forssman antibody. Flow cytometry analysis of COS-1 cells transfected with pFS-7 (*Upper*) or pcDNAI (*Lower*). (*Left*) Cells labeled with monoclonal antibody M1/22.25 (anti-Forssman; thick line) or with secondary antibody alone (thin line). (*Center*) Cells labeled with antibody 81 FR 2.2 (anti -A; thick line) or with secondary antibody alone (thin line). (*Right*) Cells labeled with biotinylated lectin BSI-B_4 (galα-1,3- specific; thick line) or with phycoerythrin-conjugated streptavidin alone (thin line).

pcDNAI were able to mediate transfer of [^{3}H]GalNAc to Gb4Cer (Fig. 5*A*, lanes 1 and 2, respectively). In contrast, there was no difference in the amount of GalNAc transferred to a number of other glycolipid substrates by Triton X-100 extracts prepared from COS-1 cells transfected with either pFS-7 or pcDNAI (Fig. 5*A*, lanes 3 to 14). The low levels of GalNAc transferred to the histo-blood group H acceptor ([Fucα-1,2]Galβ1,4GlcNAc) and GM3 (NeuAcα2,3-Galβ1,4GlcCer) by both pFS-7- and pcDNAI-transfected cells suggest the presence of endogenous A-transferase and GM2 synthetase in COS-1 cells (Fig. 5*A*, lanes 3 and 4 and 11 and 12, respectively).

Since the FS cDNA sequence demonstrated extensive sequence identity to the A- and B-transferases, which catalyze the addition of GalNAc and Gal, respectively, to the histo-H acceptor, we examined the FS gene product for galactosyltransferase activity. As demonstrated above, extracts of COS-1 cells transfected with pFS-7 but not extracts of cells transfected with vector alone mediated transfer of GalNAc to the glycolipid acceptor Gb4Cer producing FG (Fig. 5*B*, lanes 1 and 2). In the absence of

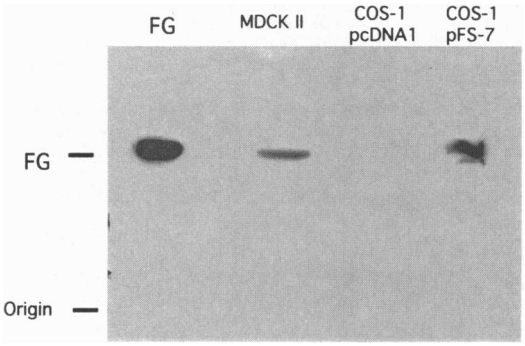

FIG. 4. COS-1 cells transfected with pFS-7 produce FG. Crude lipid extracts or control FG were separated by TLC. FG was detected with antibody M1/22.25. Lane 1, FG standard (1 μg); lane 2, lipid extract from 2 × 10^6 MDCK II cells; lane 3, lipid extract of 2 × 10^7 COS-1 cells transfected with vector pcDNAI; lane 4, lipid extract of 2 × 10^7 COS-1 cells transfected with pFS-7.

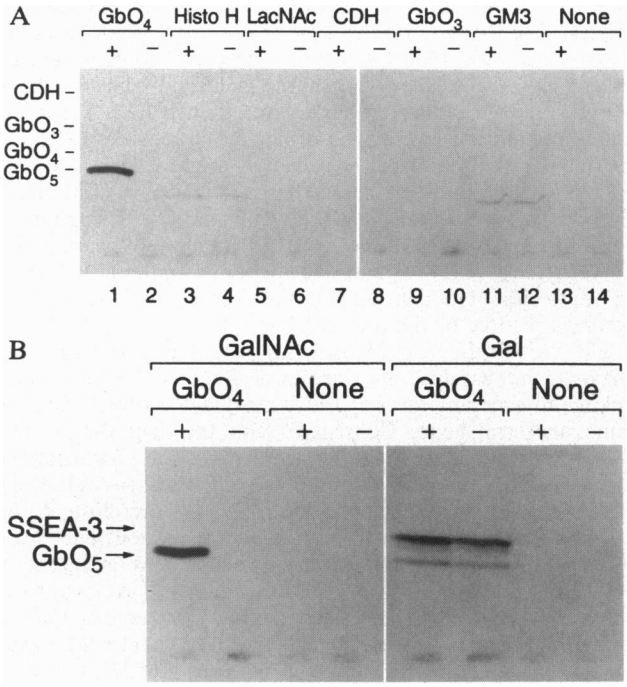

FIG. 5. FS cDNA encodes a UDP-GalNAc:Gb4Cer *N*-acetylgalactosaminyltransferase. Transfer from radiolabeled sugar nucleotide to various substrates was detected by autoradiography after TLC separation. Membrane extracts of COS-1 transfected with pFS-7 (+) and pcDNAI (−) were used as the source of enzyme. (*A*) Transfer of [^{3}H]GalNAc to potential glycolipid acceptors. (*B*) Transfer of [^{3}H]GalNAc and [^{3}H]Gal in the presence or absence of exogenous substrate.

substrate, no labeled product was detected. Extracts of cells transfected with pFS-7 and pcDNAI both transferred [^{3}H]Gal to Gb4Cer, producing Gal-Gb4Cer (SSEA-3; Fig. 5*B*, lanes 5–8; for structure, see Table 1). Detection of SSEA-3 synthetase activity was dependent on the addition of exogenous acceptor and represents endogenous enzyme activity in COS-1 cells as it did not depend on transfection with pFS-7. An additional band migrating below SSEA-3 was synthesized by extracts of cells transfected with both pFS-7 and pcDNAI, indicating the presence of an enzyme capable of modifying SSEA-3, likely by the addition of sialic acid to create sialyl-galactosyl-globoside (SSEA-4). Like SSEA-3, the presence of the slower migrating band is independent of the FS gene product and therefore results from an endogenous enzyme. These experiments demonstrate that transfection of COS-1 cells results in *de novo* production of UDP-GalNAc:Gb4Cer *N*-acetylgalactosaminyltransferase (the FS).

DISCUSSION

We have used a phenotypic cloning approach to isolate a cDNA encoding FS. The cloned cDNA encodes an enzyme with the properties predicted of FS and failed to demonstrate GalNAc or Gal transferase activity when a number of related substrates were examined.

FS displays extensive similarity to three recently cloned glycosyltransferases (Fig. 2); A-transferase (UDP-*N*-acetylgalactosamine:Galβ-1,4[Fucα-1,2]GlcNAc α-1,3 *N*-acetylgalactosaminyltransferase), B-transferase (UDP-galactose:Galβ-1,4[Fucα-1,2] GlcNAc α-1,3-galactosyltransferase), and α-(1,3)GT (UDP-galactose:Galβ1,4GlcNAc α-1,3-galactosyltransferase). Overall, FS displays 42% identity in amino acid sequence to the A- and B-transferases, and 35% identity to α-(1,3)GT. These three previously cloned transferases catalyze the addition of Gal or GalNAc in α-1,3 linkage to their respective acceptor substrates which have closely related structures (12). Despite their extensive sequence identity, each of these glycosyltransferases displays strict substrate and sugar donor specificity. For example, the A- and B-transferases differ only in four amino acids, yet exclusively transfer GalNAc and Gal, respectively, to the histo-H acceptor (23, 24). The B-transferase and α-(1,3)GT are 36% identical, yet the B-transferase requires that the acceptor contain α-(1,2)-linked fucose for transfer, whereas the α-(1,3)GT is inactive in the presence of α-(1,2)-linked fucose. Even though FS is a member of this glycosyltransferase gene family, the structure of its acceptor (Gb4Cer) differs from the acceptors used by the A-, B-, and α-(1,3)GT transferases since it has a terminal β-(1,3)-linked GalNAc rather than terminal β-(1, 4)-linked Gal. Using immunofluorescence of transfected cells, TLC immuno-overlays, and *in vitro* enzyme assays, we found no evidence for transfer of GalNAc or Gal to either blood group H or the Galβ1,4GlcNAc sequence used as substrate by the α-(1,3)GT.

The close sequence identity suggests the possibility that enzymes that catalyze the transfer of Gal or GalNAc in α-1,3 linkage to acceptor carbohydrates may have evolved from the same ancestral gene. Consistent with this hypothesis, it is known that the genes encoding the A-, B-, and O-transferase (a nonfunctional homologue of the A- and B-genes) as well as one of two nonfunctional alleles of the α-(1,3)GT are clustered on human chromosome 9, where it is believed that these genes arose by duplication and subsequent divergence (25). Maintenance of the genes encoding A-, B-, and O- transferases is proposed to be due to a process of balanced selection that has been extant for at least 13 million years, since the A-, B-, and O-alleles are highly conserved among many nonhuman primates (26, 27).

Expression of the carbohydrates produced by each of these glycosyltransferases is known to be highly regulated. Like Forssman antigen synthesis in the mouse, histo-blood group carbohydrates are expressed in a differentiation-specific and developmentally regulated fashion in primates. Moreover, expression of the blood group antigens is know to be altered in a variety of cell types upon malignant transformation, as is production of FG in some tumor cell types. α-(1,3)GT expression, like FS, is tightly regulated during murine development. In all of these cases, synthesis is felt to be controlled at the level of glycosyltransferase gene transcription. Interestingly however, these enzymes differ in their expression patterns; FG is highly expressed in undifferentiated murine teratocarcinoma cells and decreases upon differentiation (28, 29), whereas the level of Gal(α-1,3)Gal expression, the product of the α-(1,3)GT, increases upon differentiation of these same cells (12, 30). Thus, even though the genes for FS and α-(1,3)GT are likely derived from the same ancestral gene, they must be under control of different promoter elements.

We report here the identification and cloning of one of the first genes involved in globoseries glycolipid synthesis. In the past, a number of hypotheses have been raised regarding the roles that these glycosphingolipids play in cellular biology, and the mechanisms by which they exert their effects. Identification of the FS gene will allow us to dissect the function of FG expression during development and cellular differentiation.

Dr. Lora Hooper kindly provided synthetic LacNAc substrate. We are indebted to Drs. Lora Hooper and Stephen Manzella for helpful advice throughout this work, and to Drs. Erika Bach and Linda van Dyk for assistance with FACS. This work was supported by National Institutes of Health Grants R01-CA21923 to J.U.B. and R03-HD33672-01 to D.B.H. D.B.H. was a Fellow of the Pediatric Scientist Development Program when these studies were initiated and is currently a Scholar of the Child Health Research Center of Excellence in Developmental Biology at Washington University School of Medicine (HD33688).

1. Hakomori, S.-I. (1993) *Biochem. Soc. Trans.* **21,** 583–595.
2. Hakomori, S.-I. (1990) *J. Biol. Chem.* **265,** 18713–18716.
3. Siddiqui, B. & Hakomori, S. (1971) *J. Biol. Chem.* **246,** 5766–5769.
4. Willison, K. R. & Stern, P. L. (1978) *Cell* **14,** 785–793.
5. Fredman, P. (1993) *Adv. Lipid Res.* **25,** 213–231.
6. Mori, E., Mori, T., Sanai, Y. & Nagai, Y. (1982) *Biochem. Biophys. Res. Commun.* **108,** 926–932.
7. Hooper, L., Hindsgaul, O. & Baenziger, J. (1995) *J. Biol. Chem.* **270,** 16327–16332.
8. Wysocki, L. & Sato, V. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 2844–2848.
9. Hirt, B. (1967) *J. Mol. Biol.* **26,** 365–369.
10. Haslam, D. B., Borén, T., Falk, P., Ilver, D., Chou, A., Xu, Z., Normark, S. (1994) *Mol. Microbiol.* **14,** 399–410.
11. Seed, B. & Aruffo, A. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 3365–3369.
12. Smith, P. L. & Lowe, J. B. (1994) *J. Biol. Chem.* **269,** 15162–15171.
13. Larsen, R., Rajan, V., Ruff, M., Kukowska-Latallo, J., Cummings, R. & Lowe, J. (1989) *Proc. Natl. Acad. Sci.* **86,** 8227–8231.
14. Kiyomitsu, N., Watanabe, Y., Maruyama, K., Kasahara, K., Nagai, Y. & Sanai, Y. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 7952–7956.
15. Kurosawa, N., Hamamoto, T., Lee, Y. C., Nakaoka, T., Kojima, N. & Tsuji, S. (1994) *J. Biol. Chem.* **269,** 1402–1409.
16. Nagata, Y., Yamashiro, S., Yodoi, J., Lloyd, K. O., Shiku, H. & Furukawa, K. (1992) *J. Biol. Chem.* **267,** 12082–12089.
17. Nara, K., Watanabe, Y., Maruyama, K., Kasahara, K., Nagai, Y. & Sanai, Y. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 7952–7956.
18. Sasaki, K., Kurata, K., Kojima, N., Kurosawa, N., Ohta, S., Hanai, N., Tsuji, S. & Nishi, T. (1994) *J. Biol. Chem.* **269,** 15950–15956.
19. van Genderen, I. L., van Meer, G., Slot, J. W., Geuze, H. J. & Voorhout, W. F. (1991) *J. Cell Biol.* **115,** 1009–1019.
20. Kannagi, R., Levery, S. B., Ishigami, F., Hakomori, S., Shevinsky, L. H., Knowles, B. B. & Solter, D. (1983) *J. Biol. Chem.* **258,** 8934–8942.
21. Field, M. C. & Wainwright, L. J. (1995) *Glycobiology* **5,** 463–472.
22. Murphy, L. A. & Goldstein, I. J. (1977) *J. Biol. Chem.* **252,** 4739–4742.
23. Yamamoto, F.-i., Clausen, H., White, T., Marken, J. & Hakomori, S.-i. (1990) *Nature (London)* **345,** 229–233.
24. Yamamoto, F.-i. & Hakomori, S.-i. (1990) *J. Biol. Chem.* **265,** 19257–19262.
25. Yamamoto, F.-i., McNeill, P. D. & Hakomori, S.-i. (1995) *Glycobiology* **5,** 51–58.
26. Kominato, Y., McNeill, P. D., Yamamoto, M., Russell, M., Hakomori, S.-i. & Yamamoto, F.-i. (1992) *Biochem. Biophys. Res. Commun.* **189,** 154–164.
27. Martinko, J. M., Vincek, V., Klein, D. & Klein, J. (1993) *Immunogenetics* **37,** 274–278.
28. Clark, G. F., Gorbea, C. M., Cummings, R. D., Mattox, S. & Smith, D. F. (1991) *Carbohydr. Res.* **213,** 155–168.
29. Krupnick, J. G., Damjanov, I., Damjanov, A., Zhu, Z. M. & Fenderson, B. A. (1994) *Int. J. Cancer* **59,** 692–698.
30. Cho, S., Yeh, J., Cho, M. & Cummings, R. (1996) *J. Biol. Chem.* **271,** 3238–3246.