Characterization of the Human Forssman Synthetase Gene

AN EVOLVING ASSOCIATION BETWEEN GLYCOLIPID SYNTHESIS AND HOST-MICROBIAL INTERACTIONS*

Huijin Xu, Thomas Storch, Min Yu, Sean P. Elliott, and David B. Haslam‡

From the Departments of Pediatrics and Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri 63110

Differences in glycolipid expression between species contribute to the tropism of many infectious pathogens for their hosts. For example, we demonstrate that cultured human and monkey urinary epithelial cells fail to bind a canine Escherichia coli uropathogenic isolate; however, transfection of these cells with the canine Forssman synthetase (FS) cDNA enables abundant adherence by the same pathogen, indicating that addition of a single sugar residue to a glycolipid receptor has marked effects on microbial attachment. Given the contribution of glycolipids to host-microbial interactions, we sought to determine why human tissues do not express Forssman glycolipid. Query of the GenBank™ data base yielded a human sequence with high identity to the canine FS cDNA. Reverse transcription polymerase chain reaction and Northern blotting demonstrated the presence of FS mRNA in all tissues examined. A human FS cDNA was characterized, revealing identities with the canine FS gene of 86 and 83% at the nucleotide and predicted amino acid sequences, respectively. In contrast to the canine FS cDNA, transfection of COS-1 cells with the human FS cDNA resulted in no detectable FS enzyme activity. These results suggest that variability in glycolipid synthesis between species is an important determinant of microbial tropism. Evolutionary pressure from pathogenic organisms may have contributed to diversity in glycolipid expression among species.

Co-evolution of microbial pathogens with their eucaryotic hosts has resulted in a remarkable degree of specificity in their interaction. As a consequence, pathogenic organisms are typically restricted in their ability to cause disease to a limited number of species. An early stage in the pathogenesis of most infectious diseases is microbial adherence to host cells (1). As a consequence, the capacity for adherence is a major determinant of the host range available to a given pathogen. Interestingly, glycolipids are the initial attachment site for several pathogenic organisms (2, 3). Illness caused by such diverse agents as bacteria, viruses, and toxins is known to be directly affected by the same pathogen, indicating that addition of a single sugar residue to a glycolipid receptor has a marked effect on microbial attachment. Given the contribution of glycolipids to host-microbial interactions, we sought to determine why human tissues do not express Forssman glycolipid. Query of the GenBank™ data base yielded a human sequence with high identity to the canine FS cDNA. Reverse transcription polymerase chain reaction and Northern blotting demonstrated the presence of FS mRNA in all tissues examined. A human FS cDNA was characterized, revealing identities with the canine FS gene of 86 and 83% at the nucleotide and predicted amino acid sequences, respectively. In contrast to the canine FS cDNA, transfection of COS-1 cells with the human FS cDNA resulted in no detectable FS enzyme activity. These results suggest that variability in glycolipid synthesis between species is an important determinant of microbial tropism. Evolutionary pressure from pathogenic organisms may have contributed to diversity in glycolipid expression among species.

* 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 15 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) 163572.

† Supported by the American Heart Association, NIAID Grant AI42817 from the National Institutes of Health, a Basil O’Connor Starter award from the March of Dimes, and Scholarship HD33888 from the Child Health Research Centers of Excellence in Developmental Biology at Washington University School of Medicine. To whom correspondence should be addressed: Washington University School of Medicine, Box 8116, One Children’s Place, St. Louis, MO 63110. Tel. 314-454-6050; E-mail: haslam@kids.wustl.edu.

‡ Supported by the American Heart Association, NIAID Grant AI42817 from the National Institutes of Health, a Basil O’Connor Starter award from the March of Dimes, and Scholarship HD33888 from the Child Health Research Centers of Excellence in Developmental Biology at Washington University School of Medicine. To whom correspondence should be addressed: Washington University School of Medicine, Box 8116, One Children’s Place, St. Louis, MO 63110. Tel. 314-454-6050; E-mail: haslam@kids.wustl.edu.

Viewed from the host perspective it appears that susceptibility to infectious diseases is impacted greatly by factors regulating glycolipid synthesis.

Glycolipids consist of a variable carbohydrate moiety attached to a ceramide backbone and are a component of virtually all eucaryotic cells. Hundreds of distinct glycolipids have thus far been described (9). Variability in glycolipid structure occurs primarily in the carbohydrate moiety due to differences in the number, type, or anomer linkages between sugar residues (10). Each cell type, however, expresses glycolipids containing a limited repertoire of all possible carbohydrate structures. This cell lineage-specific pattern of glycolipid expression is tightly regulated during cellular differentiation and development and varies between species (11–14).

Forssman glycolipid (FG; Forssman antigen) is a member of the globo series glycolipid family, all of which have in common a core galactosyl-α1,3galactose moiety. Unlike many other mammalian species, human cells do not normally produce Forssman glycolipid but produce the precursor glycolipids globotriaosylceramide and globoside (Gb0). In some species, including humans, these precursors serve as an attachment site for bacteria, viruses, and toxins (15–21). In other species it is likely that modification of these glycolipids (such as by the addition of an N-acetylgalactosamine to create Forssman glycolipid) alters adherence of pathogenic organisms, directly affects microbial ecology, and modifies host susceptibility to infectious diseases. It is in this context that we sought to elucidate the mechanisms controlling Forssman glycolipid expression in human cells.

EXPERIMENTAL PROCEDURES

Materials—T24 (human bladder epithelium) and Vero (monkey kidney) cells were obtained from ATCC and passaged in Dulbeco’s modified Eagle’s medium containing 10% fetal calf serum and following stable transfection were maintained in geneticin (Life Technologies, Inc.) at 600 µg/ml (T24) and 200 µg/ml (Vero). Escherichia coli HB101 containing plasmid pRS-1 (encoding the class III P-pilus operon) was kindly provided by Staffan Normark. Anti-Forssman antibody was concentrated from hybridoma M1/22.21 (ATCC) tissue culture supernatant. A human substantia nigra Agt10 cDNA library and a panel of human tissue Agt10 cDNA libraries (CLONTECH, Palo Alto CA) were kindly provided by Jonathan Gitlin. Human multiple tissue RNA blots were purchased from CLONTECH. Double-stranded DNA probes were radiolabeled with [32P]dCTP (3000 Ci/mmol) by random priming using reagents from Roche Molecular Biochemicals. Tag DNA polymerase and dNTPs were from Life Technologies, Inc. Glycolipid substrates and standards were purchased from Sigma. UDP-[3H]galactose and UDP-[3H]N-acetylgalactosamine were purchased from American Radiochemicals.

Stable Transfection of Human and Monkey Uroepithelial Cells with Canine FS cDNA—The canine FS cDNA from pFS-10 (22) was ligated into neomycin-selectable plasmid pCR3.1 to create pFS35. A truncated

1 The abbreviations used are: FG, Forssman glycolipid; FS, Forssman glycolipid synthetase; PCR, polymerase chain reaction; GbO, globoside; HA, hemagglutinin; PBS, phosphate-buffered saline; MES, 4-morpholineethanesulfonic acid; kb, kilobase pair; bp, base pair.
(non-functional) FS cDNA was created by deleting a 400-bp ApoI fragment from pFS-35 to create plasmid pFS-36, used as a transfection control. T24 and Vero cells were transfected with plasmids pFS-35 and pFS-36 using LipofectAMINE reagent as described previously (22). Forty-eight hours after transfection neomycin was added to the media, and resistant clones were allowed to proliferate. After 10 days, the clones were pooled, and the 5% of pFS-35-transfected cells reacting most strongly with anti-Forsman antibody were selected by fluorescence-activated cell sorter and further expanded in the presence of neomycin. As we found that serial passage resulted in a decrease in the percentage of FS-expressing cells over time, the cells were resorted by fluorescence-activated cell sorter for anti-Forsman reactivity prior to performing bacterial adherence assays.

Bacterial Adherence Assays and Immuno-Thin Layer Chromatography—In order to examine the binding of E. coli pRS-1 to monolayers of FS-transfected cells, each cell line was seeded onto slide chambers, allowed to attach overnight, and then were fixed with 4% paraformaldehyde in PBS for 30 min at room temperature and overlaid with a suspension of E. coli:pRS-1 at approximately 1 x 10^9 bacteria/ml. The slides were incubated for 30 min at 37 °C and then washed 5 times with cold PBS. Following another incubation in 4% paraformaldehyde for 15 min at room temperature, cells were stained with Gimsa (Diff. Quick, Roche Molecular Biochemicals) and visualized by light microscopy.

For immunologic detection of Forsman glycolipid and examination of bacterial adherence to cellular glycolipids, a crude lipid extract was prepared from 1 x 10^7 E. coli:pRS-1 at approximately 100 μl per dish and vortexed vigorously. An equal volume of chloroform:methanol (approximately 100 μl) and chloroform:methanol:water and spoted onto TLC plates. Glycolipids were separated in organic solvent (chloroform:methanol:water, 65:35:8); the plates were air-dried and then stained with bovine serum albumin (5% in PBS). Plates were then overlaid either with monoclonal anti-Forsman antibody (MI/M22.21) or bacteria that had been metabolically radiolabeled as described previously (23). The plates were washed and then subjected either to secondary antibody and enhanced chemiluminescence (anti-Forsman glycolipid) or directly to autoradiography (radiolabeled bacteria).

PCR Amplification of Human FS cDNA—Two oligonucleotide primers (MFS-4, 5'-CCCCACCATATAAGAGTCC-3', and MFS-7, 5'-GCCACCATGCTTCGGAGGACC-3') that span intron VI in the human FS gene were used to PCR-amplify from 1 μg of human genomic DNA using primers FS-26 (5'-TTCRCCAGYCMCTTCCGGAGTC-3') and FS-28 (5'-CAGAGGTACAAGGGGACAGC-3') under the conditions described above. The PCR product was subcloned into plasmid pCR3.1 (Invitrogen) to create plasmid pHFS. The plasmid was digested with EcoRI, and the insert was gel-purified, labeled with [32P]dCTP, and used as probe for Northern blots. 50 ng of purified β-actin DNA (CLONTECH) was similarly labeled. Human multiple tissue RNA blots were prehybridized for 30 min at 68 °C in RapidHyb solution (CLONTECH) and then hybridized for 1 h at 68 °C in the same solution containing 1 million cpm per ml of probe. Filters were washed at 68 °C in 0.2x SSC, 0.1% SDS and exposed to Kodak Biomax MS film at 70 °C.

Isolation and Characterization of FS cDNA—A human agt10 substanitia nigra cDNA library was plated on NZCY media, transferred to nitrocellulose, and hybridized overnight with radiolabeled pHFS insert at 65 °C in 5× SSC, 5× Denhardt’s, 1% SDS, and 10 μg/ml herring sperm DNA. Filters were washed in 0.2× SSC, 0.1% SDS at 65 °C and then exposed to autoradiography film. Several hybridizing plaques were identified. Three were purified to homogeneity. One such clone, pHFS-3, contained the entire FS coding region and an unspliced intervening sequence between exons VI and VII.

Nucleotide Sequencing—Two EcoRI restriction fragments obtained from pHFS-3 were subcloned into plasmid pCR3.1 to create plasmids pHFS-1.2 and pHFS-1.8. Cycle sequencing was performed using IC and T7 (Invitrogen) or gene-specific primers, fluorescently labeled dye terminators (BigDye), and Taq polymerase using conditions recommended by the manufacturer (Applied Biosystems). The sequence of both inserts was determined in forward and reverse directions. Additional sequencing reactions were performed to resolve ambiguous nucleotide designations. Nucleotide sequencing of the inserts in the FS expression constructs was performed in a similar manner.

Generation of Human, Chimeric, and Epitope-tagged FS Expression Vectors—To generate a eucaryotic expression vector containing the entire human FS coding region and intron VI, plasmid pHFS-1.2 was digested with XbaI and BamHI and ligated into BglII/XbaI-digested plasmid pHFS-1.8 to create plasmid pHFS. A chimeric human:dog cDNA expression vector was generated by PCR amplifying nucleotides 41 to 357 (encoding amino acid residues 1–119) of the human FS cDNA from plasmid pHFS-1.8 using primers MFS-20 (5'-TTCCTCCGT-GAGGGACAGCC-3') and MFS-8 (5'-CCCCACGCGAAAAC-SGT-GACSCC). Nucleotides 334–1273 (encoding residues 111–347, the stop codon, and 3′-untranslated DNA) from the canine FS cDNA were amplified from plasmid pHFS-10 with primers MFS-31 (5'-GGGCTCAGG-GTTTGGCGTGGG-3') and FS-9 (5'-GCTTACAGCTGAAGGCC-3'). Five microliters of each PCR product were mixed and brought to 100 μl in 1× PCR buffer containing primers MFS-20 and FS-9, 200 μM dNTPs, and 1 unit of Taq DNA polymerase. Twenty five additional cycles of PCR were performed, and the 1.2-kb product was directly subcloned into pCR3.1 to create plasmid pHD. Primer MFS-31 changed residue 113 in the canine FS from leucine to valine to match the human FS at residue 113, and since residues 120–122 are identical between

Human Forssman Synthetase Gene—A 589-bp fragment from exon VII of the human FS gene was PCR-amplified from 1 μg of human genomic DNA using primers FS-26 (5'-TTCRCCAGYCMCTTCCGGAGTC-3') and FS-28 (5'-CAGAGGTACAAGGGGACAGC-3') under the conditions described above. The PCR product was subcloned into plasmid pCR3.1 (Invitrogen) to create plasmid pHFS. The plasmid was digested with EcoRI, and the insert was gel-purified, labeled with [32P]dCTP, and used as probe for Northern blots. 50 ng of purified β-actin DNA (CLONTECH) was similarly labeled. Human multiple tissue RNA blots were prehybridized for 30 min at 68 °C in RapidHyb solution (CLONTECH) and then hybridized for 1 h at 68 °C in the same solution containing 1 million cpm per ml of probe. Filters were washed at 68 °C in 0.2x SSC, 0.1% SDS and exposed to Kodak Biomax MS film at 70 °C.
human and canine FS proteins, this chimera matches the human FS from residues 1 to 122 and matches the canine FS from residues 123 to 347.

A dog/human chimeric cDNA was similarly generated by PCR amplification from the plasmid polylinker to nucleotide 431 of the canine FS cDNA (encoding residues 1–136 of the canine FS except that residue 134 was changed from glutamine to glutamate to match the human FS) using primers T7 and MFS-35 (5′-GAAGAATCTTCTGGAAGATCCAGTCAGCTG-3′). Nucleotides 1298–1966 of the human cDNA (encoding residues 129–347 and the stop codon) were amplified with primers MFS-33 (5′-CCTGGACTGATCCGGGGAGTTCTTC-3′) and MFS-27 (5′-CCGTGTTGAGTCAGCCGAGGAGTTCTTC-3′). The PCR products were mixed and subjected to a further 25 cycles of PCR using primers T7 and MFS-27. The chimeric PCR product was subcloned directly into pCR3.1 to create plasmid pH, which matches the canine FS from residues 1 to 127 and the human FS from residues 128 to 347 (Fig. 8A).

In order to remove exon VI from the human FS expression plasmid, overlapping PCR was performed in a similar manner to that described above. Nucleotides 41 to 359 plus nucleotides 5′ to the splice site of the human FS cDNA were amplified from plasmid pHFS-1.2 using primers MFS-20 and MFS-40 (5′-GGACTGGATGAAATGAGTGTACTTC-3′) and MFS-39 (5′-CACGGTGTTTGCCGTGGGGAAGTACACTCATTTCATC-3′). Nucleotides spanning the splice site of the human FS were amplified from plasmid pHFS-1.2 using primers MFS-39 (5′-CACGGTGTTTGCCGTGGGGAAGTACACTCATTTCATC-3′) and MFS-27 (5′-CCGTGTTGAGTCAGCCGAGGAGTTCTTC-3′) and MFS-27. Primers MFS-39 and MFS-40 were designed to contain complementary and overlapping regions that flank intron VI, such that these ends prime extension in both directions when the two products anneal and delete intron VI. Following five rounds of amplification without flanking primers, the entire coding region was amplified with primers MFS-20 and MFS-27. The resulting product was ligated into plasmid pCR3.1 and its sequence fidelity verified by nucleotide sequence determination.

Addition of an amino-terminal HA epitope tag (consisting of amino acids MGYFPYDVPD) to the human and dog FS cDNAs was accomplished by amplification of either the human or dog FS cDNA, respectively, with primers HA-2 (5′-ATGGGCTATCCTTATGACGTGC-3′) and HA-3 (5′-ATGGGCTATCCTTATGACGTGC-3′) with FS-9 (5′-ATGGGCTATCCTTATGACGTGC-3′) and FS-10 (5′-ATGGGCTATCCTTATGACGTGC-3′) as a control. After developing in chloroform:methanol (2:1), the products were separated from unincorporated nucleotide sugar by passing over a reverse phase C18 column, washing with 5 ml of water, 3 ml of 0.1 M, 1000 cpm per pmol; American Radiochemicals) was substituted for UDP-N-acetylgalactosamine as the sugar nucleotide donor.

**Detection of Epitope-tagged FS Proteins by Western Blotting**—COs cells transfected with epitope-tagged canine or human FS, or untransfected cells, were harvested by scraping into ice-cold PBS. Half of the cells (approximately 4 × 10^6) were used for enzyme assays as described above. The remaining cells were harvested by centrifugation and resuspended directly in 100 μl of SDS-polyacrylamide gel electrophoresis buffer. After heating to 94 °C for 5 min, the samples were subjected to SDS-polyacrylamide gel electrophoresis and then transferred to nitrocellulose. After blocking with 5% skim milk in PBS for 30 min, primary antibody (HA-11, Babco) was added at 1:1000 dilution and incubated for 2 h at room temperature. After washing and incubating in a 1:5000 dilution of secondary antibody (horseradish peroxidase-labeled goat anti-mouse IgG; Roche Molecular Biochemicals), Western blots were developed by enhanced chemiluminescence.

**RESULTS**

**Bacterial Adherence to Cells Expressing Forssman Synthetase**—Canine uropathogenic bacteria commonly produce adhesive fibers terminating with the class III PapG adhesin (24). We examined the ability of E. coli HB101 harboring plasmid pRS-1, encoding the class III adhesin, to adhere to wild-type and FS-transfected uroepithelial cells. Bacteria were overlaid on monolayers of human or monkey kidney cells transfected with pFS-36 (control; WT) or pFS-35 (canine FS; +FS). Adherence was examined by light microscopy. Whereas no bacterial

**Fig. 2. Detection of Forssman glycolipid in extracts of transfected cells by anti-FG antibody (A) and adherence by FG-specific bacteria (B).** Lipid extracts were prepared from T24 (Human) and Vero (Monkey) cells stably transfected with pFS-36 (WT) and pFS-35 (+FS). Lipid extracts and a lane of purified Forssman glycolipid were spotted onto TLC plates and developed in organic solvent. Forssman glycolipid was detected using anti-Forssman glycolipid antibody (A) or radiolabeled E. coli HB101 transfected with plasmid pBS-1 (B) as described in the text. The migration of FG and GbO4 standards on an adjacent lane are indicated.

**Fig. 3. Genomic organization of the human Forssman synthetase gene.** Query of the GenBank*™* data base with the canine FS cDNA yielded high sequence identity to cosmids clones derived from human chromosome 9g34 (GenBank*™* accession numbers AC002319 and AC001643).

*Filled boxes* indicate putative coding exons, and *shaded boxes* indicate 5′ and 3′ non-coding sequence. Subsequent analysis of a human FS cDNA confirmed the assignment of exon boundaries. *Numbers below the boxes* indicate the corresponding nucleotides in the published genomic clone AC002319.
adherence was seen to cells transfected with the truncated FS cDNA (wild-type cells), numerous bacteria bound to cells expressing the full-length FS cDNA (Fig. 1).

Bacterial Adherence to Glycolipids Extracted from Uroepithelial Cells—In order to demonstrate that bacterial adherence was to Forssman glycolipid in transfected cells, extracts were prepared from pFS-36 (WT) and pFS-35 (+FS)-transfected human and monkey uroepithelial cells and examined for the presence of Forssman glycolipid and the ability to mediate bacterial adherence. Immuno-thin layer chromatography demonstrated an intense band of anti-FG reactivity comigrating with authentic FG in lipid extracts of human and monkey cells transfected with plasmid pFS-35, whereas no reactivity is seen to extracts of cells transfected with a truncated FS cDNA (Fig. 2A). Bacterial adherence to glycolipid extracts demonstrated a faint band of adherence to the precursor glycolipid (GbO$_4$, globoside) in extracts from all four cell lines (23, 24). An intense band of bacterial binding to FG was seen in extracts of FS-transfected cells (Fig. 2B).

Identification of the Human FS Genomic Locus—Differences in Forssman glycolipid expression between humans and other species contribute to variable host susceptibility to microbial pathogens. In order to determine whether absent Forssman glycolipid synthesis in humans had a genetic basis, we characterized the human FS gene. GenBank$^\text{TM}$ and dEST data bases were queried with the canine FS nucleotide sequence using the program BLAST, yielding human genomic cosmid sequences that were homologous (99%) to the published human FS cDNA. The human FS gene is separated into at least seven exons spanning more than 8 kb of DNA (Fig. 3). Genomic organization of the human FS gene is very similar to the highly homologous a(1,3)-galactosyltransferase and ABO transferase loci (25, 26). In humans each of these related glycosyltransferase genes are located on chromosome 9q34, and is located in a single large exon in the human cosmid probe (Fig. 3).

Detection of Human FS in cDNA Libraries by PCR—In order to determine whether the FS gene was transcribed in human tissues, PCR was carried out on lysates from cDNA libraries prepared from a panel of human tissues. Primers were designed to amplify specifically a 590-bp product from within exon VII of the FS cDNA. A single band of the expected molecular weight was amplified from cDNA prepared from all tissues tested, suggesting that the FS gene was transcribed in each of these tissues (Fig. 4A). In order to exclude the possibility that the PCR product represented amplification of contaminating genomic DNA, primers spanning intron VI were utilized. Genomic DNA or incompletely processed cDNA was expected to yield a product of approximately 1.5 kb, whereas processed message was expected to yield a product of 528 bp. Brain and kidney cDNA libraries were subjected to PCR with these primers and yielded the expected ~530-bp product, indicating the presence of spliced transcripts in these cDNA libraries (Fig. 4B). Genomic DNA yielded the expected 1.5-kb product and failed to yield the smaller 528-bp product, confirming that the cDNA libraries indeed contained copies of processed FS mRNA. In order to verify that the 528-bp PCR product originated from the FS cDNA rather than a homologous ABO glycosyltransferase, the PCR product was subcloned and characterized. The nucleotide sequence was identical to the published human genomic FS locus except that intron VI had been removed, resulting in a message retaining the expected coding frame (data not shown). Interestingly, brain cDNA yielded a faint 1.5-kb PCR product, either representing genomic DNA contamination or the presence of an incompletely processed pre-mRNA. As described below, isolation of an incompletely processed message from a substantia nigra library indicates that intron VI may be spliced less efficiently from the FS mRNA than are introns I to V.

Northern Blotting of Human RNA Samples—In order to characterize further the pattern of human FS expression, Northern blots of RNA samples from various tissues were hybridized with a probe derived from the putative human FS catalytic domain. Hybridization of the FS probe to a 2.2-kb transcript was found in all tissues examined (Fig. 5A). Highest levels of expression were seen in placenta, ovary, and peripheral blood leukocyte RNA, whereas liver, thymus, and testis...
**FIG. 6.** Sequence analysis of a human Forssman synthetase cDNA. Nucleotide sequence and deduced peptide sequence of the insert in λHFS-3 (A). Bold letters indicate the putative splice site between exons VI and VII. Homology of the derived human FS peptide sequence (HFS) with the canine FS (DFS) (B). Light shading indicates amino acid identities, and dark shading denotes amino acid similarities. Overall, the sequence identity was 83%. Homology of the human FS with the human blood group A transferase (ABO) and the murine α(1,3)-galactosyltransferase (MGT) (C).
RNA demonstrated the weakest hybridization. In addition to the predominant 2.2-kb transcript, hybridizing bands at approximately 3.2 and 8.0 kb were seen in some tissues, possibly representing incompletely or alternately spliced FS mRNA. Notably, no hybridization was seen to the ABO glycosyltransferase transcript that migrates at 5.5 kb on RNA blots (27). The α(1,3)-galactosyltransferase mRNA is not expressed in human tissues and therefore can be excluded as the hybridizing band on RNA blots (28, 29). Hybridization with an actin control probe indicated approximately equal loading of RNA in each lane (Fig. 5B).

Isolation of the Human FS cDNA—A λgt10 human substantia nigra cDNA library was probed with a radiolabeled portion of the human FS gene. Screening of 500,000 plaques identified several positive clones. Three such clones were partially characterized by restriction digestion and polymerase chain reaction. Two did not contain the entire open reading frame and were not fully characterized. The remaining clone (lHFS-3) was characterized in detail. The cDNA insert was released by EcoRI digestion as two fragments of 1.8 and 1.2 kb which were subcloned into plasmid pCR3.1. Nucleotide sequencing revealed 363 bp of 5' non-coding region followed by the putative start codon and a 1041-bp coding region interrupted by a 917-bp intervening sequence corresponding to intron VI (Fig. 6A). The putative coding region of lHFS-3 is 98% identical to previously published genomic sequence. No homology was found in the data base to the first 83 nucleotides of the lHFS-3 insert (nucleotides −363 to −280), indicating that the corresponding genomic sequences have yet to be characterized.

Overall, the coding region of the human FS cDNA demonstrates 86% nucleotide sequence identity and 83% predicted amino acid sequence identity to the canine FS (Fig. 6B). The peptide is predicted to consist of 347 amino acids with a molecular mass of 40.2 kDa and an isoelectric point of 9.0. Like the canine FS, the human peptide is predicted by hydrophobicity calculations to be a type II transmembrane protein wherein the large carboxyl-terminal catalytic domain is located in the Golgi lumen (not shown). Comparison of the human FS peptide sequence with the human blood group A glycosyltransferase and the murine α(1,3)-galactosyltransferase reveals identities of 42 and 38%, respectively (Fig. 6C).

Expression of Human cDNA in COS-1 Cells—Several amino acid substitutions are found in the derived human peptide sequence when compared with the canine FS peptide (Fig. 6B). In order to determine whether the human FS cDNA product possessed Forssman synthetase activity, an expression construct was created by ligating the two FS cDNA fragments in proper orientation into vector pCR3.1 to create plasmid pHFS. COS-1 cells were transiently transfected with plasmid pHFS or the canine FS cDNA (pFS-10). Reverse transcriptase PCR across intron VI was utilized to demonstrate that the human FS cDNA was transcribed and processed in transfected cells (data not shown). Membrane extracts were prepared from transfected cells and used as enzyme source for Forssman synthetase assays (globoside:α(1,3)-galactosaminyltransferase). Whereas extracts of cells transfected with the canine FS cDNA resulted in an intense band of radioactive Forssman glycolipid product, no FS activity was detected in COS cell extracts after transfection with the human FS cDNA, confirming that the
human FS cDNA was incapable of encoding a functional Forssman synthetase (Fig. 7). Control reactions to detect endogenous galactosyltransferase activity in COS cells demonstrated approximately equal amounts of cellular extract in each sample (data not shown).

Expression of Chimeric cDNA Constructs in COS-1 Cells—Most glycosyltransferases are type II transmembrane proteins in which the amino terminus is located in the cytosol, followed by a transmembrane spanning domain, a short stem region, and then the catalytic domain. The greatest homology between FS and related blood group transferases exists in the putative catalytic domain. Interestingly, among all members of the ABO gene family, the putative catalytic domain is transcribed from a single large exon, whereas the amino terminus of the peptide is transcribed from several smaller exons, indicating that the gene structure of this family has been evolutionarily conserved (30). We hypothesized that amino acid substitutions in the catalytic region may account for the lack of enzyme activity encoded by the human FS. In order to test this hypothesis, chimeric constructs were generated. Plasmid pHD contains the amino-terminal (putative non-catalytic) region from the human cDNA fused in-frame to the region encoding the catalytic domain of the canine FS. Conversely, plasmid pHFS consists of the canine amino-terminal coding region fused to the human catalytic domain (Fig. 8A). COS-1 cells were transfected with plasmids pFS-10 (canine), pHFS (human), pHD (human:dog), or pDH (dog:human) FS expression plasmids. Membrane extracts were prepared from transfected cells as described above and used as the source for FS enzyme assays. Forssman synthetase activity was found in extracts of cells transfected with the canine cDNA and the human:dog chimera as indicated by the presence of a band of radiolabeled product comigrating with Forssman glycolipid (Fig. 8B).

Thus, the human:dog chimeric protein functions as a Forssman synthetase, whereas no FS enzyme activity was seen in extracts of cells transfected with the full-length human FS cDNA or the canine:human chimera. These results suggest that nucleotide substitutions encoding the putative catalytic domain render the human protein enzymatically inactive.

Detection of Epitope-tagged FS Proteins and Enzyme Activity in Cell Extracts—Based on the results described above we concluded that the human FS cDNA was unable to encode functional FS enzyme activity. However, we considered two additional possibilities. First, the human FS expression construct used above contained intron VI, whereas the canine and chimeric expression plasmids required no processing to generate a mature mRNA. Additionally, we could not conclusively exclude the possibility that the human FS mRNA was translated less efficiently than the canine FS message and that absent protein synthesis, rather than absent enzymatic function, was responsible for the lack of FS activity in cells transfected with the human FS cDNA. In order to address both concerns a plasmid was constructed in which intron VI was removed from the human FS cDNA. Next, sequences encoding identical HA epitope tags were added to the amino termini of the canine and human FS cDNAs.

After transfection with these HA-tagged cDNA plasmids, COS cells were divided for detection either of epitope-tagged protein or FS enzyme activity. Western blotting of COS cell
glycolipid standards are indicated in A described in the text (B). Protein by Western blot (FS cDNA expression plasmids were examined for the presence of protein transiently transfected with HA epitope-tagged dog (D) or human (H) FS cDNA expression plasmids were examined for the presence of protein by Western blot (A) or assayed for Forssman synthetase activity as described in the text (B). Migration of molecular weight or Forssman glycolipid standards are indicated in A and B, respectively.

extracts demonstrated approximately equal expression of the canine and human HA-tagged FS proteins. The proteins migrate as a doublet of approximately 41 and 43 kDa, presumably reflecting non-glycosylated and glycosylated forms (Fig. 9A). Consistent with the results described thus far, the canine HA-tagged FS resulted in abundant FS activity upon transfection into COS cells, whereas the human HA-tagged FS yielded no detectable FS enzyme activity (Fig. 9B). Control reactions utilizing UDP-galactose as sugar nucleotide donor detected endogenous UDP-galactose:globoside galactosyltransferase activity and demonstrated approximately equal input of cellular extract in each sample (Fig. 9B).

We conclude that in transfected COS-1 cells the canine and human FS proteins were expressed at comparable levels and that the human FS lacks Forssman synthetase activity.

DISCUSSION

Indirect evidence suggests that glycolipids have important roles in normal cellular biology. For example, the expression of distinct glycolipids is tightly regulated during cellular differentiation and malignant transformation (31, 32). Moreover, the tissue-specific regulation of glycolipid expression is highly conserved among individuals within the same species. The mechanisms whereby glycolipids are postulated to affect cellular differentiation and morphogenesis include a direct interaction with membrane-associated signaling molecules or the ability of glycolipids to serve as ligands for cell-cell interactions (33, 34).

Forssman glycolipid is expressed in a tissue-specific and developmentally regulated fashion in such diverse mammals as dog, sheep, horse, guinea pig, and mouse (13, 35–37). Like other glycolipids, FG has been proposed to play important roles in cellular biology. Whereas some human tissues have been reported to react with anti-FG antibody, structural data supporting the presence of Forssman glycolipid in these cells and the presence of Forssman synthetase activity are not available (38–41). Our results suggest either that the anti-FG-reactive antigen was not comprised of the classic Forssman pentasaccharide (GalNAcα1–3GalNAcβ1–3Galα1–4Glcβ1–4Glu) or that some human cells possess an as yet uncharacterized ability to produce Forssman glycolipid. The absence of Forssman glycolipid in normal human tissues indicates either that this molecule is dispensable or that the function of FG has been subsumed by another glycolipid during primate evolution.

We demonstrate here that the inability to produce Forssman glycolipid results from mutations within the putative FS catalytic domain. Yet, widespread tissue expression of the FS mRNA suggests that the human orthologue retains a biologic function, either that it has acquired a novel enzymatic activity (a hypothesis that we have not explored here) or that its remaining biologic function does not require glycosyltransferase activity. In this regard, it is interesting to compare the human FS orthologue with the α(1,3)-galactosyltransferase, another member of the ABO glycosyltransferase gene family. Like Forssman synthetase, the α(1,3)-galactosyltransferase gene is expressed by many mammalian species, apparently losing the ability to encode glycosyltransferase activity recently during primate evolution (42). Interestingly, whereas the human FS retains 83% amino acid sequence identity with the canine orthologue and is expressed widely in human tissues, the α(1,3)-galactosyltransferase gene is not expressed in human cells, and the processed pseudogene contains multiple frameshift mutations (28). Relative conservation of the FS gene during human evolution and widespread tissue expression may imply an as yet unidentified function for the human FS protein.

The ability to express FS or lack of Forssman glycolipid has implications for the host tropism of several infectious diseases. For example, most E. coli isolates causing pyelonephritis demonstrate the capacity for adherence to globoseries glycolipids on uroepithelial cells via a bacterial organelle known as the P-pilus (43). Interestingly, P-pili are subdivided into three classes with slightly differing affinities for members of the globoseries glycolipids, including Forssman glycolipid, due to amino acid substitutions in their adhesive protein (23). This differing affinity confers on the bacteria the ability to infect different species (24). Hence, expression of varying globoseries glycolipids among species is correlated with infection by different bacteria. Other organisms known to bind globoseries glycolipids include parvovirus B19, the causative agent of erythema infectiosum (“fifth disease”) and the shiga toxins responsible for hemolytic uremic syndrome. We recently demonstrated that expression of the FS cDNA in human and monkey cells results in resistance to shiga toxins by the depletion of globotriaosylceramide,2 a precursor to FG that is required for toxin endocytosis (19). Those results suggest that expression of FG protects host cells against some microbial diseases. We demonstrate here that expression of FG in the same human and monkey kidney cells results in their de novo ability to bind bacteria expressing the class III P-pilus adhesin.

Evolution of glycolipids synthesis in eucaryotic species has occurred during constant association with pathogenic organisms. Within the confines of cellular homeostasis, a consequence of this selective pressure appears to be considerable diversity in glycolipid expression among species.

Acknowledgments—Jacques Baenziger, Jonathan Gitlin, Guyu Ho, and Lou Muglia are gratefully acknowledged for advice, support, and critical review of the manuscript.

2 S. P. Elliott and D. B. Haslam, submitted for publication.
