Sequence Properties of the 1,2-Diacylglycerol 3-Glucosyltransferase from Acholeplasma laidlawii Membranes

RECOGNITION OF A LARGE GROUP OF LIPID GLYCOSYLTRANSFERASES IN EU BACTERIA AND AR CHAEA*A

Received for publication, March 22, 2001, and in revised form, April 6, 2001
Published, JBC Papers in Press, April 6, 2001, DOI 10.1074/jbc.M102576200

Stefan Berg‡‡, Maria Edman‡‡, Lu Li†, Malin Wikström‡, and Åke Wieslander††

From the ‡Department of Biochemistry, Umeå University, S-901 87 Umeå, Sweden and the †Department of Biochemistry and Biophysics, Stockholm University, S-106 91 Stockholm, Sweden

Synthesis of the nonbilayer-prone α-monoglucosyl-diacylglycerol (MGlcDAG) is crucial for bilayer packing properties and the lipid surface charge density in the membrane of Acholeplasma laidlawii. The gene for the responsible, membrane-bound glucosyltransferase (aMGS) (EC 2.4.1.157) was sequenced and functionally cloned in Escherichia coli, yielding MGlcDAG in the recombinants. Similar amino acid sequences were encoded in the genomes of several Gram-positive bacteria (especially pathogens), thermophiles, archaea, and a few eukaryotes. All of these contained the typical EX-E catalytic motif of the CAZY family 4 of α-glycosyltransferases. The synthesis of MGlcDAG by a close sequence analog from Streptococcus pneumoniae (spMGS) was verified by polymerase chain reaction cloning, corroborating a connection between sequence and functional similarity for these proteins. However, aMGS and spMGS varied in dependence on anionic phospholipid activators phosphatidylglycerol and cardiolipin, suggesting certain regulatory differences. Fold predictions strongly indicated a similarity for aMGS (and spMGS) with the two-domain structure of the E. coli MurG cell envelope glycosyltransferase and several amphipathic membrane-binding segments in various proteins. On the basis of this structure, the aMGS sequence charge distribution, and anionic phospholipid dependence, a model for the bilayer surface binding and activity is proposed for this regulatory enzyme.

Lipids are the local environment for most integral and peripheral membrane proteins, which often depend on the lipids for optimal function. The large diversity of lipids and the differences in composition and properties between membranes have made it difficult to find out common features of bilayer organization and how lipids and proteins are cooperating in local processes. Lipid-synthesizing pathways have been mapped for the most common types of lipids, and several of the corresponding enzymes catalyzing these reactions have been characterized. However, when it comes to the connection between regulation of bilayer properties and enzyme structure, very little is known (1). So far, only a few lipid-synthesizing enzymes have been crystallized. Which structural properties are involved in the catalytic mechanism of these lipid enzymes, and how are the membrane properties sensed (1)?

In the well-characterized plasma membrane of Acholeplasma laidlawii, the lipid composition is regulated in a manner to maintain (i) lipid phase equilibria, close to a potential bilayer to nonbilayer transition, (ii) a nearly constant radius of spontaneous curvature, and (iii) a certain anionic surface charge density of the lipid bilayer. The synthesis of the major nonbilayer-prone lipid in this membrane, monoglucosyldiacylglycerol (MGlcDAG) (Scheme 1, step I), plays an important role to fulfill the two first points above but also the third, since it is strongly regulated by negatively charged lipids (e.g. the major in vivo lipid phosphatidylglycerol (PG)) (2). MGlcDAG is consecutively processed into diglucosyl diacylglycerol (DGlcDAG) (Scheme 1, step II). Consequently, the formation of this glucolipid, a transfer of Glc from the donor UDP-Glc to the acceptor lipid diacylglycerol (DAG) catalyzed by a glucosyltransferase (EC 2.4.1.157) (3), plays a central part in understanding the total regulation of lipid syntheses in A. laidlawii membranes. Furthermore, glycolipids including nonbilayer-prone ones are major constituents in many cell surface membranes, certain bacterial groups, and most photosynthetic organelles. Fairly little is known about the synthesis and regulation of these. Usually, they are made in a separate pathway (as in A. laidlawii), branching from the conserved one to anionic phospholipids.

In this work, we have cloned the gene for the α-monoglucosyldiacylglycerol synthase from A. laidlawii membranes and a sequence analog from the pathogen Streptococcus pneumoniae and propose these genes, on the basis of sequence similarities, to belong to a new large group of lipid glycosyltransferases that are widely spread in nature. We also present a functional comparison between the two cloned glucosyltransferases and discuss structural properties based on two- and three-dimensional fold predictions from the primary structure. A striking similarity to two new, related structures for an Escherichia coli glycosyltransferase and epimerase, respectively, is indicated.

* This work was supported by the Swedish Natural Science Research Council. 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank®/EBI Databank with accession number(s) AF349769.

‡ These authors contributed equally to this work.

To whom correspondence should be addressed. Tel.: 46-8-16-24-63; Fax: 46-8-15-36-79; E-mail: ake@dbb.se.

© 2001 by The American Society for Biochemistry and Molecular Biology, Inc.

22056 This paper is available on line at http://www.jbc.org

The Journal of Biological Chemistry Vol. 276, No. 25, Issue of June 22, pp. 22056–22063, 2001
This is an Open Access article under the CC BY license.
EXPERIMENTAL PROCEDURES

Urea-PAGE was used to detect a protein that was detected after SDS-PAGE. The protein bands were visualized by silver staining. The molecular masses of the bands were estimated by co-migration with molecular mass markers. The protein bands were excised and subjected to in-gel tryptic digestion. The tryptic peptides were extracted and analyzed by mass spectrometry. The peptide sequences were compared with those of known proteins in the protein databases. The protein was identified as MGlcDAG synthase (ALmgs).

Sequence Features of MGlcDAG Synthase

The DNA sequence of the MGlcDAG synthase gene was determined by direct sequencing of the PCR-amplified DNA. The gene was cloned into the pET15b expression vector and expressed in E. coli BL21 (DE3). The expressed protein was purified and used for analysis.

Enzymatic Assays

The enzymatic activity of MGlcDAG synthase was measured using a radiometric assay. The reaction mixture contained UDP-[14C]glucose and the recombinant protein. The reaction was started by the addition of the enzyme and incubated at 37 °C for 1 h. The reaction was terminated by the addition of an internal standard and quenching reagent. The radioactivity was measured by liquid scintillation counting.

Protein Analysis

The purified protein was analyzed by SDS-PAGE and Western blotting. The protein bands were visualized by Coomassie blue staining and probed with specific antibodies. The molecular mass of the protein was determined by comparison with molecular mass markers.

RESULTS

Gene Cloning and Sequencing

The MGlcDAG synthase gene was isolated from A. laidlawii (3). The gene was cloned into the pET15b expression vector and expressed in E. coli BL21 (DE3). The expressed protein was purified and used for analysis.

Enzymatic Assays

The enzymatic activity of MGlcDAG synthase was measured using a radiometric assay. The reaction mixture contained UDP-[14C]glucose and the recombinant protein. The reaction was started by the addition of the enzyme and incubated at 37 °C for 1 h. The reaction was terminated by the addition of an internal standard and quenching reagent. The radioactivity was measured by liquid scintillation counting.

Protein Analysis

The purified protein was analyzed by SDS-PAGE and Western blotting. The protein bands were visualized by Coomassie blue staining and probed with specific antibodies. The molecular mass of the protein was determined by comparison with molecular mass markers.
Sequence Features of MGlcDAG Synthase

| Position | dsource* | Function/name, species | E value |
|----------|----------|------------------------|---------|
| -2323 to -1051 | gbAF282249 | Arginyl-tRNA synthetase, L. lactis | E-112 |
| -1049 to -4 | gbAAAD5554 | Conserved hypothetical protein, T. maritima | 3E-8 |
| +1 to +1197 | | MGS | |
| +1152 to +1433 | embX61062.1 | tRNA<sup>Leu</sup> A. laidlawii | |

* Database accession number.

cording to SignalP) and given the gene name ALmgs. The amino acid sequence is not related to the ones in the UGT Nomenclature Committee (12). No other potential lipid-synthesizing enzyme genes were present on the contig. However, the two tRNA-amino acid synthases indicate a conserved chromosomal environment. The A. laidlawii tRNA<sup>Leu</sup> has been described earlier in Ref. 13.

The amino acid sequence for ALmgs was used as a query in a PSI-BLAST homology search. A selection of the best hits are presented in Table II; among the top ones putative glycosyltransferases from a large variety of organisms were found. All of these belong to family 4 (retaining GTs) (14) and share the typical residues for α-GTs in this class and belong to family D in the classification by Breton et al. (15). However, the best scores were revealed in a data base for finished and unfinished microbial genomes. Here, the best hit was a sequence coding for a protein in Treponema denticola with 36% amino acid identity to the MGlcDAG synthase (the six best hits all had identities above 30%). The translated sequences from Enterococcus faecalis, Streptococcus pyogenes, S. pneumoniae, and T. denticola (second row) were aligned with the ALmgs sequence (top) (Fig. 1). The conserved residues in all five sequences were mainly focused to three domains: the first 40 residues, residues 90–130, and above all amino acids 280–310, which contains the characteristic motif EX<sub>5</sub>E of family 4 GTs (cf. above).

A search in the protein domain data base (at NCBI) showed that ALmgs and the potential homologs belonged to pfam00534, which is the glycosyltransferase group 1. These indicate glucosylation reaction. The same radiolabeled product was synthesized independent of which of the two substrates, [14C]UDP-Glc or [14C]DAG, was labeled. In order to verify that the lipid product synthesized in vivo (Fig. 2, lane 5) was MGlcDAG, it was extracted from the TLC plate and used in an in vitro assay for DGlcDAG synthesis with purified DGlcDAG synthase from A. laidlawii membranes (7). This enzyme can only use α-MGlcDAG as the lipid substrate and not β-MGlcDAG or Gal variants. The extracted, radiolabeled lipid (Fig. 2, lane 1), could indeed be used as substrate in the DGlcDAG synthesis (Fig. 2). The intensity of the radiolabeled product was increased when 14C-labeled UDP-glucose was used (lane 3). The MGlcDAG TLC spot was also identified as a glycolipid by charring with sulfuric acid/methanol (1:1, v/v) (see “Experimental Procedures”). A typical purple color characteristic for glycolipids was observed (data not shown).

Hence, the two genes ALmgs and SPmgs, with the translated amino acid sequences indicated in Fig. 1, encode analogous enzymes, which both perform the synthesis of the membrane lipid MGlcDAG.

**Lipid Composition in Recombinant Cells**—The lipid composition in the recombinant E. coli strains was analyzed by incorporation of [14C]acetate into the lipids during growth. Four major lipids were recognized on the TLC plates: the glucolipid MGlcDAG, phosphatidylethanolamine (PE), and the negatively charged lipids phosphatidylglycerol (PG) and CL (Fig. 2). The control TOP10<sup>F</sup> strain, containing a pCR-Script vector, had a lipid composition normal for E. coli wild type, with about 72% PE and 28% negatively charged lipids (Table III). The recombinant strain with expressed ALmgs contained a significant fraction of MGlcDAG, about 10%. The fraction of anionic lipids was kept constant, while the nonbilayer lipid PE had decreased to about 63%. However, the homologous glycosyltransferase from S. pneumoniae, although active in vitro, did not affect the lipid composition in vivo, and only traces of MGlcDAG were found (Table III, Fig. 2). The BL21 strain, overexpressing the His-tagged ALmgs, showed a slightly lower synthesis of the glucolipid compared with TOP10<sup>F</sup>, indicating that the N-terminal His extension did not seriously effect enzyme activity. The overexpressed GTs decreased the growth rates compared with controls, and the latter reached the stationary phase faster, which may influence the fractions of CL and PG.

**Lipid Environment and Enzyme Activity**—The native A. laidlawii MGlcDAG synthase is activated in an essential manner by substantial amounts of negatively charged lipids, especially PG (2, 17). E. coli cells overexpressing the GTs were solubilized and diluted with CHAPS/1,2-dioleoyl-sn-glycero-3-phosphocholine micellar dispersions supplemented with various fractions of negative lipids: DOPG and cardiolipin, due to their occurrence in A. laidlawii and S. pneumoniae membranes (18, 19), and phosphatidylserine (1,2-dioleoyl-sn-glycero-3-phospho-L-serine) and PGD because of their PG-like properties.

The results in Fig. 3 show that all four lipids were potent activators for the two GTs, but to different extents. An increased fraction of PG and CL gave sigmoidal-like activation curves for ALmgs, while spMGS responded only to PG. Cardiolipin was able to activate the spMGS but to a very low extent. The lipid-like detergent, PGD, showed only slightly activating effects on spMGS. This was also true for the ALmgs at lower concentrations, but the activity increased significantly at concentrations above 25%. Without supplemented DAG substrate, traces of MGlcDAG product could still be detected. This was...
most probably due to a minor fraction of endogenous DAG present in the added E. coli cell suspension. The total content of E. coli lipids present in an assay was estimated to be less than 50 nmol (cf. the supplemented amounts of 500 nmol). All observed effects on the cloned alMGS (Fig. 3) were in agreement with earlier studies of the native enzyme (2, 17).

**Structure Predictions**—alMGS is firmly anchored in the membrane, and detergents are needed for solubilization (3). Potential hydrophobic transmembrane (TM) segments in the sequences were investigated with a number of prediction methods at the ExPASy Molecular Biology server. One TM (residues 3–22) was proposed according to HMMTOP (20), TMpred (residues 3–24) (21), and TopPred2 (residues 4–24) (22), but not by the SOSUI (23). The orientation of the putative TM was uncertain. This segment also had a substantial amphipathic character according to a hydrophobic moment analysis (24). According to the majority of methods used, spMGS was a soluble protein lacking TM segments, except by HMMTOP (20), predicting one at residues 3–21.

The alMGS sequence showed a low homology to proteins in the structure database (PDB). However, a three-dimensional fold prediction method based on homologous sequence searches (10) listed MurG from E. coli (sequence identity of 14%; 26% similar amino acids), which encodes for a glycosyltransferase catalyzing the last step in the peptidoglycan precursor pathway (25, 26), and a soluble UDP-N-acetylglucosamine 2-epimerase from E. coli (27) (12% identical and 24% similar amino acids). These two have very similar structures (27), but MurG was...
proposed to be attached to the inner membrane (no predicted TM). An alignment of MurG and alMGS by ClustalW was used to find sequence and potential structural similarities (some gaps included) between the two membrane-associated proteins. The aligned sequences were marked with the predicted secondary structure for alMGS (from Jpred (28)) and the established secondary structure and topology for MurG and the epimerase. Besides this, initial circular dichroism studies of a purified His-tagged alMGS suggest both α- and β-structures, in agreement with the α/β open sheet structure determined for the MurG domains.

Furthermore, two regions in this alignment had higher identities (Fig. 4); residues 62–105 and 290–318 (the EX7E motif) in MurG showed similarity to residues 71–114 (25% identity) and 254–278 (37% identity) in MurG. A closer study of residues 74–85 in the alMGS revealed a mixture of hydrophobic and basic amino acids in a predicted amphipathic helix (24). The corresponding sequence in MurG (residues 84–95) is an α-helix proposed to be part of the membrane-binding domain (29). The amphipathic characters are also evident from helical wheel presentations (data not shown). Residues 102–106 in MurG consist of a glycine-rich loop (G loop) localized between a β-strand and an α-helix (29). This motif is related to the one included in the classical Rossmann fold (30, 31). Interestingly, a similar motif is present in the alMGS (putative G loop sequence SXGXXG) (Fig. 4).

Other membrane-binding segments may be present as well. Searching the PDB Intermediate Sequence Library at the SCOP data base (11) with the alMGS sequence as a probe revealed domains in the two related botulinum (PDB structure 3BTA) and tetanus (PDB structure 1A8D) neurotoxins, close to the binding site for the negatively charged neuronal ganglioside lipid (32). Segment sequence 212–260 in alMGS, containing two conserved regions in the potential lipid GTs (Fig. 1), could be modeled on the PDB 1A8D structure template by SwissModel (ExPaSy server (33)), Positions 225–239 had the largest hydrophobic moment (24) for the entire alMGS. Likewise, this segment could also be modeled on the membrane-binding, 62–106 segment in MurG (cf. Fig. 4). Another motif spanning from about residue Ile111 to Tyr127, with a large hydrophobic moment in alMGS, was highly conserved in all the potential lipid GTs (Fig. 1). This amino acid stretch was, despite a low similarity, possible to align with a motif (Pro121, Lys136) conserved among MurG proteins (29). Generally, amphipathic segments (24) were less frequent and of smaller magnitude in MurG than in alMGS.

A theoretical pl was calculated to be around 9 (or higher) for the sequences in Fig. 1, with the exception of spMGS and E. faecalis with a pl of about 5–6. For alMGS, a high number of basic residues were found in the N-terminal half, while the second half was dominated by acidic residues. This polarization of charges along the sequence gave a high pl (~10) for the N-terminal halves and a lower pl (~7) for the C-terminal part, a difference that was analogous but lower for spMGS. Interestingly, a similar pattern of charge distribution is valid for MurG with a pl for the full sequence calculated to be 10.2, while the C-terminal had an acidic pl of 6.2.

Hence, the alMGS lipid glucosyltransferase seems to have several structural features in common with certain membrane-binding proteins of known structure, especially the E. coli GT MurG.
FIG. 4. Alignment of alMGS and the MurG glycosyltransferase. MurG, bound to the inside of the E. coli inner membrane and with a 
recently established structure (29), was proposed to be similar to alMGS by a three-dimensional fold prediction method (10). The two sequences 
were aligned with ClustalW, and the potential secondary structure of alMGS was predicted by Jpred (at the ExPASy server); this is shown along 
the established MurG structure (x-ray). A very good prediction of the determined secondary structures of MurG and the epimerase was also 
achieved by Jpred (data not shown). Light gray, α-helix; dark gray, β-strand. Two interesting regions with high sequence homologies are boxed. 
The first stretch is proposed to be a membrane-associating domain, and the second contains the motif for UDP-sugar binding. The outside 
sequences are represented by a plus sign for positively charged residues (Lys, Arg), minus sign for negative ones (Asp, Glu), and dots for others. 
In the first box, residues in proposed G-loops are marked with boldface italic type (Gly, Ser). Positions 111—127 are strongly conserved in the Fig. 
1 lipid glycosyltransferases. The Glu in boldface type in the second box is conserved for the family 4 of glycosyltransferases. Segment 212—260 in 
alMGS could be modeled on an established membrane-binding toxin structure (PDB 1A8D).

DISCUSSION

Lipid Glycosyltransferase Genes—The genes for the well 
studied MGlcDAG synthase from A. laidlawii strain A-EF22 
was cloned, and the encoded catalytic function was confirmed. 
In the standard assay procedure, rac-1,2-diacylglycerol was 
utilized as the acceptor and UDP-α-glucose as the donor sub-
strate for synthesis of the lipid product MGlcDAG. The stereo-
chemistry of the sugar moiety was characterised indirectly by a 
coupled enzymatic synthesis of the subsequent glucolipid DG-
cDAG, which specifically demands α-MGlcDAG as substrate 
(7). β-MGlcDAG is not present in A. laidlawii (34). Related 
Gram-positive bacteria, like the ones in Fig. 1, all contain 
the lipid α-MGlcDAG in their membranes (35).

In addition, α-MGlcDAG is the structural base of the lipotei-
choic acid in S. pneumoniae, anchoring this cell wall polymer 
into the cytoplasmic membrane (36). The visualization of MGI-
cDAG from the S. pneumoniae gene (clone SPmg3; Fig. 2) 
strongly indicates that the corresponding genes in Fig. 1, being 
more similar to the alMGS sequence than spMGS, all encode 
the MGlcDAG synthesis function. Likewise, treponemas and 
other spirochetes all contain a monoglycosyl-DAG, where the 
hexose is glucose, galactose, or mannose (37). The T. denti-
cola sequence in Fig. 1 most likely encodes the GT needed for this 
synthesis. In an analogous manner to S. pneumoniae, this lipid 
may also be the anchor to the complex outer membrane sheath 
lipid OML521 in T. denticola (38). Additional homologs from 
other pathogens, not shown here, were also found in, for example, 
Streptococcus mutans, Clostridium acetobutylicum, and 
Streptococcus equi.

In the Gram-positive sequences of Fig. 1, the MGlcDAG 
synthase genes are adjacent to another gene (potentially in an 
operon), tentatively identified by us as glycosyltransferases of 
CAZy family 4 (cf. above). No other GT was identified next to 
the A. laidlawii MGS gene (see Table I). Likewise, in the 
related mollicutes Mycoplasma pneumoniae, three potential 
lipid GT genes lay separated on the chromosme. However, 
the alMGS sequence was not related to the M. pneumoniae 
one, in agreement with the different glycolipid structures.

This potential group of lipid GTs seems to be widely spread 
in nature according to the list of selected orthologs in Fig. 1 and 
Table II. Prokaryotes, including Gram-positive and Gram-neg-
ative eubacteria, and archaea are represented, but the list also 
contains sequences from eukaryotes. Interestingly, they were 
found also in the hyperthermophiles Thermotoga maritima and 
Pyrococcus horikoshii, members of eubacteria and archaea, 
respectively. This all indicates that a common ancestor to these 
orthologs was developed very early in the evolution, before the 
separation of the bacterial from the archaeal lineages. All an-
logs (Fig. 1 and Table II) analyzed contained the EX-E motif 
typical for the retaining α-GTs of CAZy family 4 (14). A number 
of the analogs in Table II are involved in the synthesis of 
various lipids or lipid-based molecules. The Borrelia burgdor-
feri enzyme synthesizes monogalactosyl-DAG, in accordance 
with the reported presence of this lipid in B. hermsii (40), and 
the S. pneumoniae homolog encodes the α-MGlcDAG synthase 
(this work). Lactococcus, Deinococcus, Thermotoga, and Phy-
coccus species (Table II) all contain various glycolipids, includ-
ing α-MGlcDAG in the two former ones (35, 41). However, the 
gene from Pseudomonas aeruginosa is not the one synthesizing 
the excreted rhamnolipid (42). The Synecocystis sp., ranked 
as number 11 (Table II), was recently suggested as a lipid 
α-glycosyltransferase catalyzing the synthesis of sulfoquinovo-
syl-diacylglycerol (43).

The Bacillus subtilis gene (Table II) is tuaC involved in 
lipoteichoic acid synthesis (44). Number 17 is Rv0557 from 
Mycobacterium tuberculosis, and it was recently identified as a 
mannosyltransferase (PimB) acting on a phosphatidylinositol 
lipid (45). In CAZy family 4 (14), there are more than a dozen 
open reading frames from Arabidopsis thaliana. One of these 
genes (CAB69850; Table II) showed homology to the MGlcDAG 
synthase from A. laidlawii and an even higher similarity (35% 
identity) to the proposed sulfoquinovosyl-diacylglycerol syn-
thease from Synechocystis (cf. above). Like the β-MGalDAG synthase 
(a GT) from cucumber (46), the two analogs from 
Arabadopsis and Synechocystis (sfr0384) seemed to contain a 
leader and signal peptide of about 103 and 31 amino acids, 
respectively (ChloroP/SignalP prediction). They are probable 
transit peptides required for import through the chloroplast 
envelope and export to a proper Synecocystis compartment.

3 M. L. Rosén and Å. Wieslander, manuscript in preparation.

4 Berg, S., Ostberg, Y., Bergshörm, S., and Wieslander, Å., manuscript in preparation.
Hence, the alMGS enzyme is member of a potentially large and conserved group of lipid glycosyltransferases in nature. Most important, this group is not closely related sequencewise to the corresponding β-MGalDAG synthases in plant chloroplasts (47).

Regulation of Activity—The enzymatic regulation of the MGlCDAG synthesis in A. laidlawii has been extensively characterized (2, 6). Certain lipids activated the alMGS due to their charge properties, with PG as the most potent activator. Here, the two enzymatic activities expressed in E. coli were studied in a mixed micellar system in vitro and with respect to the effects of negatively charged lipids (Fig. 3). The sigmoidal curves shown for PG and CL in the activation of alMGS reached their maximum at a similar fraction of negative charges (two in CL) (Fig. 3B). Similarly, PG also stimulated the activity of spMGS, while a very poor response was given by CL. This difference in regulatory properties is very interesting, since PG and CL are major lipids in S. pneumoniae (19), but only PG has been found in this strain of A. laidlawii (18). The binding of alMGS to lipid bilayers was recently shown to be modulated by electrostatic interactions, with a preference for binding to PG- and CL-enriched membranes. The regulation of spMGS activity may be governed in a different way; the low response to CL and the stimulatory effects by PG indicate a reduced ability to interact with CL. spMGS has substantially different charge properties as illustrated by the lower pi of its N- and C-terminal halves (domains) (see above). This may serve to inhibit synthesis of too much nonbilayer-prone lipid, since both CL and MGlCDAG have such properties. Alternatively, CL is a true inhibitor of the spMGS enzyme. In the A. laidlawii used, lacking CL, this is evidently not the case. In the two AlmgS recombinant clones in vivo, the new nonbilayer-prone glucolipid was synthesized to ~10 mol%. The fraction of PE was lowered to the same extent, while the fraction of negatively charged lipids was kept. This down-regulation of the nonbilayer E. coli lipid may be an enzymatic regulation of the lipid synthesis to keep certain biophysical properties, like the spontaneous curvature (48), intact in the bilayer. However, the nonbilayer-prone β-MGalDAG from cucumber did not cause an analogous reduction of only PE in E. coli (46).

The lack of in vivo MGlCDAG synthesis in the SPmgs clones may depend on (i) the presence of CL in E. coli (cf. above); (ii) a lower density of basic amino acids in spMGS, leading to a weaker binding to an intracellular anionic membrane; or (iii) an inhibitor acting on the enzyme in vivo but not in vitro. Early studies of this enzyme in S. pneumoniae (49) localized the glucosyltransferase activity to a soluble fraction, indicating that the protein was not tightly bound to the membrane. The major nonbilayer-prone lipid PE in E. coli did not act as an inhibitor to this enzymatic activity according to results from in vitro experiments (data not shown). PE has not been found in S. pneumoniae.

Structure Proposal—Up to now only a handful NDP-glycosyltransferases have been structurally determined. The majority of these are using the inverting mechanism and the glycosidic bond formed in the products are in the β-configuration, but one exception is the newly determined structure of LgtC in Neisseria meningitidis (50), which uses a retaining mechanism. The sequence similarity between these structures is low, and they are classified into different glycosyltransferase (CAZy) families (14). However, their three-dimensional structures fall into only two superfamilies (51). The three-dimensional fold prediction (10) for alMGS (see “Results”) is proposed to be similar to one of these, containing the membrane-bound E. coli MurG (29) and soluble UDP-N-acetylglucosamine 2-epimerase (27) but also phage T4 β-GT structures (51). Similar predictions were valid for all of the most closely related sequences in Fig. 1 and for Lactococcus lactis and B. burgdorferi in Table II. Likewise, this was also the case for the cucumber β-MGalDAG GT and its Arabidopsis homolog (data not shown). The latter two and MurG belong to CAZy family 28 (14), strongly indicating structural similarities between the latter and family 4, including alMGS. The LgtC in N. meningitidis, with a retaining mechanism, belongs to family 8. Although they have an analogous catalytic mechanism, the LgtC and alMGS do not show any strong structural homology.

A prediction of alMGS secondary structures and an alignment along the MurG sequence (Fig. 4) revealed several surprising similarities and several regions potentially involved in membrane binding. One of these (positions 212–260; Figs. 1 and 4) could be modeled on an analogous region in two membrane-binding toxins (cf. above), but most typical was the amphipathic character of all of these regions. Such features are described for many proteins binding to lipid bilayer surfaces, and they are analyzed in more detail for a number of established amphipathic helices from the latter recently reviewed by Johnson and Cornell (52). Most similar in this collection of amphipathic helices was a membrane binding segment of DnaA (positions 368–388), initiating chromosome replication in E. coli (53); it aligned with positions 218–240 in the region with the largest hydrophobic moment of the entire alMGS sequence (cf. Figs. 1 and 4). The importance of this DnaA segment for phospholipid interaction is visualized by recent mutant studies (54).

Searching the sequences for several lipid-binding proteins, we found several intriguing similarities. The negatively charged, signal recognition particle receptor FtsY of E. coli, integrating into anionic phospholipids (55), has a positively charged amphipathic α-helix 4 in the structure (56), very similar in sequence to the position 75–91 amphipathic segment in alMGS (Fig. 4). Likewise, but with a slightly lower similarity, was the resemblance of this alMGS segment with the C-terminal amphipathic anchor segments of the LgtC galactosyltransferase (57) and the E. coli phosphatidylemerine syntheses (58). Phosphatidylemerine synthase is the rate-keeping step for the synthesis of the major nonbilayer-prone lipid PE in E. coli (39) and associates to a negatively charged lipid surface (6). The features discussed above and the similarity of this amphipathic, positively charged segment in alMGS with the aligned membrane-binding segment in MurG (first box in Fig. 4) strongly support a similar anchoring function for these two. In the soluble epimerase (above) the corresponding helix sequence segment has fewer positive and more negative charges and is...
now the contact region in the dimer (27), with no membrane attachment.

A model (schematic diagram) for the interaction and anchoring of alMGS with a lipid bilayer surface by a combination of charge-charge and hydrophobic interaction is shown in Fig. 5. It is based on (i) the indicated similarities (see “Results”) between the structurally determined MurG glycosyltransferase and alMGS (and modeled on the former); (ii) the cooperative dependence of alMGS activity on the activator lipid PG (3); (iii) the corresponding dependence of alMGS binding to PG-enriched and (CL)-enriched bilayers; (iv) the ability to release most alMGS from membranes only by detergents and chaotropic agents; and (v) the presence of several potential amphipathic helix segments in the alMGS sequence, typical for many lipid surface-associated proteins (cf. Johnson and Cornell (52)). Here, a close approach of the active site region to the bilayer surface, containing the hydrophobic substrate DAG, may be governed or modulated by the type and amount of amphipathic helix segments in the alMGS sequence, typical for many lipid surface-associated proteins (cf. Johnson and Cornell). Here, a close approach of the active site region to the bilayer surface, containing the hydrophobic substrate DAG, may be governed or modulated by the type and amount of negatively charged activator lipids.

In summary, the enzyme synthesizing the major nonbilayer-prone membrane lipid MGlcDAG in A. laidlawii is related to a large group of lipid glycosyltransferases in nature. It has homologs in related pathogenic bacteria and a structure potentially similar to E. coli MurG, and it is probably attached to the membrane by charge-charge and hydrophobic interactions.

Acknowledgments—We thank Dr. Mari Norgren (Clinical Bacteriology, Umeå University), for supplying the S. pneumoniae strain, and we are grateful to Viola Tegman and Kerstin Hjortberg for technical assistance. We are also grateful to Drs. Olof Karlsson and Dr. Susanne Vikström for valuable discussions and assistance. Dr. Bo Ek (SLU, Uppsala), performed the determination of the N-terminal and internal sequence of the alMGS. We thank Ewa Selstam (Umeå University) for glocolipid identification with the spray assay procedure. Preliminary sequence data were obtained from the Institute for Genomic Research Web site.

REFERENCES

1. Hjelmstad, R., and Bell, R. (1991) Biochemistry 30, 1731–1740
2. Karlsson, O. P., Dahlqvist, A., and Wieslander, Å. (1994) J. Biol. Chem. 269, 23484–23490
3. Karlsson, O. P., Dahlqvist, A., Vikstroem, S., and Wieslander, A. (1997) J. Biol. Chem. 272, 929–936
4. Karlsson, O. P. (1997) Maintenance of Lipid bilayer properties: Key role of two amphipathic helices in proteins. In: Protein Sequence Features of MGlcDAG Synthase. 78, 125–142
5. Johnson, J. E., and Cornell, R. (1999) J. Biol. Chem. 173, 4625–4636
6. Bopp, K., and van Heijenoort, J. (1993) J. Biol. Chem. 175, 1841–1843
7. Campbell, B., Mosimann, T., Tanner, M., and Strynadka, N. (2000) Biochemistry 39, 14993–5001
8. Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M., and Barton, G. J. (1998) Bioinformatics 14, 892–893
9. Ha, S., Walker, D. Y., and Walker, S. (2000) Protein Sci. 9, 1045–1052
10. Branden, C., and Tooze, J. (1988) Introduction to Protein Structure, pp. 47–64, Garland Publishing, Inc., New York
11. Baker, P., Britton, K., Rice, D. R., and Stillman, T. J. (1992) J. Biol. Chem. 266, 662–671
12. Umland, T., Wang, J., Swaminathan, S., Furey, W., Schmidt, J., and Saz, M. (1999) Nat. Struct. Biol. 10, 768–792
13. Lux, R., Karlsson, O. P., and Wieslander, A. (1997) J. Biol. Chem. 272, 9602–9606