Lipase maturation factor 1 (LMF1) is predicted to be a polytopic protein localized to the endoplasmic reticulum (ER) membrane. It functions in the post-translational attainment of enzyme activity for both lipoprotein lipase and hepatic lipase. By using transmembrane prediction methods in mouse and human orthologs, models of LMF1 topology were constructed and tested experimentally. Employing a tagging strategy that used insertion of ectopic glycan attachment sites and terminal fusions of green fluorescent protein, we established a five-transmembrane model, thus dividing LMF1 into the cytoplasmic and terminal domains (the amniotic enzyme). This representative model shows that the carboxyl-terminal domain is essential to lipase maturation. DUF1222 comprises four of the six domains. Three domains were found to face the cytoplasm and terminal domains to face the ER lumen. We showed for the first time, using several naturally occurring variants featuring DUF1222 truncations, that Lmf1 interacts physically with the two largest ones facing the ER lumen. We discussed the role of DUF1222 domain structure.

This article has been withdrawn by Mark H. Doolittle, Saskia B. Neher, Osnat Ben-Zeev, Ciara M. Gallagher, Maryam Hosseini, Peter Walter and Miklós Péterfy. Jo Ling-liao, Fen Yin and Howard Wong have not responded. After a review of the data, one of the corresponding authors (M. P.) learned that there was duplication and inappropriate manipulation of images in Fig. 7, raising concerns about the validity of the conclusions drawn from the results. The Walter laboratory (S. B. N., C. M. G., and P. W.) was not involved in the generation, presentation, or interpretation of the data in Fig. 7.
LMF1 Topology and Lipase Interaction

Pass) membrane protein located to the ER (9), the site of lipase folding and assembly (12,13). Dependent on the position and number of transmembrane (TM) domains, LMF1 is a polytopic protein divided into multiple domains that are oriented to both the cytoplasm and the ER lumen. Such a topological arrangement would place domains within the ER lumen in proximity to lipase nascent chains as they emerge from the translocon during protein synthesis.

LMF1 is also distinguished by an evolutionarily conserved domain of unknown function referred to as DUF1222 in the Pfam database (14); this domain is found in more than 50 proteins covering a wide taxonomic range. The cld mutation causes a large carboxy-terminal deletion of DUF1222 in mouse Lmf1, as does a nonsense mutation identified in the human homolog (Y439X). Like cld in mouse, Y439X causes combined lipase deficiency and hypertriglycerideremia in humans (9). Thus, both cld and Y439X represent loss-of-function mutations and implicate DUF1222 in lipase maturation.

In this study, we show that mouse and human LMF1 contain five TM segments that divide the protein into six separate domains; the cytoplasmic or ER luminal orientation of each domain has been determined experimentally. Notably, DUF1222 comprises four of the six domains, with the two largest oriented to the ER lumen. We provide evidence, for the first time, of the physical interaction of Lmf1 with LPL and HL. Using Lmf1 variants partially or completely lacking DUF1222, we narrowed the lipase interaction site to a single domain of unknown function referred to as DUF1222 in the Pfam database. Lmf1 lacking DUF1222, truncating Lmf1 at position 171 (Lmf1ADUF) with the added peptide sequence GKPATQQLPTPPAP at its carboxyl terminus (9); Lmf1 lacking DUF1222, truncating Lmf1 at position 171 (Lmf1ADUF) with the added peptide sequence AGQAITSYQSRTERE at its carboxyl terminus; Lmf1N430A and Lmf1N430A, removing N-linked glycan attachment sites at positions 396 and 430, respectively; Lmf1InsNA243, with an insertion of NA at position 243 of Lmf1N430A to complete an ectopic 243INAT245 glycan attachment site. Additional ectopic glycan attachment sites were inserted at the following positions of mouse Lmf1N430A: 28, 94, 101, 175, 188, and 350. The following human LMF1 sequences were used: LMF1N430A, encoding the full-length protein (567 amino acids); and LMF1N428A, removing the N-linked glycan attachment site at position 428.

For GFP fusion constructs, human LMF1 was PCR-amplified from human cDNA clone 100062174 (Open Biosystems) and inserted into pcDNA6.1/CT-GFP-TOPO (Invitrogen) containing a carboxy-terminal V5 epitope tag as described (12,17). For experiments using lipase affinity purification, a tandem affinity purification (TAP) tag was synthesized for in-frame integration into an AgeI site occurring just after the V5 epitope tag of pcDNA6 (18). After transfection, the resulting expressed LPL, HL, and PL proteins contained a carboxy-terminal V5-TAG tag consisting of the V5 epitope followed by a single calmodulin-binding peptide domain from the tobacco etch virus (TEV) protease site and endogenous IgG-binding domains derived from pcDNA6 (18).

All mouse and human RNA sequences, except for growth hormone (15), were subcloned into the pcDNA6 expression vector (Invitrogen) containing an EcoRI linker. The following mouse LMF1 cDNA sequences were cloned: LMF1N430A, encoding the full-length LMF1 protein (567 amino acids); LMF1N430A, encoding the full-length LMF1 protein (567 amino acids); LMF1N396A, encoding the full-length LMF1 protein (567 amino acids); LMF1InsNA243, with an insertion of NA at position 243 of Lmf1N430A to complete an ectopic 243INAT245 glycan attachment site. Additional ectopic glycan attachment sites were inserted at the following positions of mouse Lmf1N430A: 28, 94, 101, 175, 188, and 350. The following human LMF1 sequences were used: LMF1N430A, encoding the full-length protein (567 amino acids); and LMF1N428A, removing the N-linked glycan attachment site at position 428.

For GFP fusion constructs, human LMF1 was PCR-amplified from human cDNA clone 100062174 (Open Biosystems) and inserted into pcDNA3.1/CT-GFP-TOPO (Invitrogen) and pcDNA3.1/NT-GFP-TOPO (Invitrogen) according to manufacturer’s instructions. Additionally, human LMF1 cDNA was truncated at positions in the protein corresponding to amino acids 95, 120, 250, and 350 and then inserted into pcDNA3.1/CT-GFP-TOPO (Invitrogen). Control plasmids expressing fluorescent proteins located on the luminal or cytoplasmic side of the ER membrane (ECFP-CD36 and CD36-YFP) were a kind gift from J. Lippincott-Schwartz.

To normalize for transfection efficiency in the LMF1 functional assay, a secreted human placental alkaline-phosphatase (SEAP) reporter gene in the pM1 expression vector was used (X-extremeGENE®, Roche Applied Science). All expression vector plasmids were prepared using the EndoFree maxi kit (Qiagen) according to the manufacturer’s instructions. Diluted
plasmid solutions were quantitated using a NanoDrop 2000 spectrophotometer (Thermo Scientific).

**LMF1 Functional Assay**—Details of the Lmf1 functional assay are reported elsewhere (19). The cld mutant cell line was co-transfected with an LMF1 test sequence along with a mixture of LPL and SEAP. At 24 h post-transfection, a sample of medium was taken for SEAP activity measurements, and heparin was added to a final concentration of 10 units/ml. Two h later, cells were washed with phosphate-buffered saline and lysed in detergent-containing buffer (0.2% sodium deoxycholate, 10 units/ml heparin, 10 mM Tris-HCl, pH 7.5). After sonication and centrifugation, supernatants were used for the measurement of LPL activity and detection of LMF1 protein levels by Western blot analysis (see below).

SEAP activity was measured using the SEAP reporter assay kit (InvivoGen) according to manufacturer’s instructions. LPL activity was measured using a lecithin-stabilized radiolabeled triolein (glycerol tri[9,10(n)-3H]oleate) substrate as described (20). Levels of LPL activity and LMF1 protein were normalized to SEAP activity to account for assay-to-assay variations in transfection efficiency.

**Lipase Affinity Purification**—The details of tandem affinity purification are described elsewhere (18, 21). In experiments using lipase-TAP constructs (Fig. 7), the calmodulin affinity step was used alone. Briefly, HEK293 cells transfected with lipase-TAP constructs were lysed in 2% CHAPS, 0.15M NaCl, 10 mM Tris-HCl, pH 8.0; then applied to calmodulin affinity resin (Stratagene) and incubated overnight with constant mixing. Bound lipase-TAP proteins were eluted by Ca2+ chelation using 0.5% CHAPS, 1 mM captoethanol, 1 mM EGTA, 1 mM Mg acetate, and 10 mM Tris-HCl, pH 8.0; eluates at final concentrations of 2% CHAPS, 10% glycerol, captoethanol, 1 mM Mg acetate, 10 mM Tris-HCl, pH 8.0, and the mixture was placed in a boiling water bath for 2 min. The samples were subjected to SDS-PAGE and Western blotting as described above.

ER localization experiments of Lmf1 protein variants used confocal laser scanning fluorescence microscopy performed on HEK293 cells transfected using the FuGENE 6 transfection reagent (Roche Applied Science) as described (9, 16). Lmf1-Myc was detected with anti-Myc-TRITC antibody (1:100, Santa Cruz Biotechnology), and endogenous calnexin was visualized using rabbit anti-calnexin primary (1:50, Stressgen) and goat anti-rabbit IgG-fluorescein isothiocyanate secondary (1:100, Santa Cruz Biotechnology) antibodies. Images were processed using Adobe Photoshop software (Adobe Systems, Inc.). For the FPP assay, transfected HeLa cells on 35-mm poly-d-lysine-coated glass-bottomed plates (MatTek) were used for visualization of all GFP-LMF1 fusion constructs and of eCFP-CD38 and CD38-YFP in live cells. The FPP assay was performed as described (23) with the following exceptions. Cells were treated with 30 μM geldanamycin (CalBiochem) for 80 s before treatment with 4 μM brefeldin A (Sigma) was mimicked on a Zeiss LSM510 microscope (www.tmb.org/services.info.nih.gov/ij/) was used for confocal laser scanning fluorescence microscopy using individual cells across the entire focal plane. FPP signal intensity per nonfat milk. The Lmf1-Myc tag was similarly detected using a biotinylated rabbit anti-Myc polyclonal antibody (GenTex) at 1:2,500 followed by horseradish peroxidase-conjugated streptavidin (Invitrogen) at 1:50,000. After washing, bound horseradish peroxidase was visualized using a chemiluminescent substrate (ECL+, GE Healthcare), and images were captured on Hyperfilm ECL® (GE Healthcare). Quantitation of protein bands was carried out by densitometric scanning of Western blot films using NIH Image 1.63 software.

For endoglycosidase (endo H) experiments, 20 μg of cell lysates was denatured in 0.5% SDS, 50 mM sodium phosphate buffer, pH 5.9; then 2 μl of N-endoglycosidase H (Roche Applied Science) or water (−endo H) was added to the denatured lysates. After overnight incubation at 37 °C, SDS sample buffer was added to final concentrations of 2% SDS, 2% 2-mercaptoethanol, 10% glycerol, 0.004% bromphenol blue, 10 mM Tris-HCl, pH 8.0, and the mixture was placed in a boiling water bath for 2 min. The samples were subjected to SDS-PAGE and Western blotting as described above.

ER localization experiments of Lmf1 protein variants used confocal laser scanning fluorescence microscopy performed on HEK293 cells transfected using the FuGENE 6 transfection reagent (Roche Applied Science) as described (9, 16). Lmf1-Myc was detected with anti-Myc-TRITC antibody (1:100, Santa Cruz Biotechnology), and endogenous calnexin was visualized using rabbit anti-calnexin primary (1:50, Stressgen) and goat anti-rabbit IgG-fluorescein isothiocyanate secondary (1:100, Santa Cruz Biotechnology) antibodies. Images were processed using Adobe Photoshop software (Adobe Systems, Inc.). For the FPP assay, transfected HeLa cells on 35-mm poly-d-lysine-coated glass-bottomed plates (MatTek) were used for visualization of all GFP-LMF1 fusion constructs and of eCFP-CD38 and CD38-YFP in live cells. The FPP assay was performed as described (23) with the following exceptions. Cells were treated with 30 μM geldanamycin (CalBiochem) for 80 s before treatment with 4 μM brefeldin A (Sigma) was mimicked on a Zeiss LSM510 microscope (www.tmb.org/services.info.nih.gov/ij/) was used for confocal laser scanning fluorescence microscopy using individual cells across the entire focal plane. FPP signal intensity per

RESULTS

**Membrane Topology of LMF1**—Mouse and human LMF1 are predicted to be polytopic membrane proteins for which the topology in the ER is determined by the position and number of TM domains. The LOCATE data base (24) uses MemO, a high throughput, automated pipeline, to orient a protein with respect to the membrane. α-Helical TM prediction was performed by a MemO consensus method utilizing five publicly available predictors (listed under “Experimental Procedures”). The MemO consensus assigned a total of seven TM domains to both mouse and human LMF1 in the same locations for both orthologs (supplemental Table S1). We added three publicly available prediction methods in addition to the five used by MemO (listed under “Experimental Procedures”) and scored a particular TM domain with regard to how often it was predicted in all eight methods for each LMF1 ortholog. Three of the seven TM domains assigned by MemO scored 100%, indicating that for both mouse and human LMF1 proteins, all eight methods were in complete agreement (supplemental Table S1). Of the remaining four TM domains assigned by MemO, three were in agreement at least 60% of the time (88, 75, and 63%), and the last domain was predicted with a score of 44%.
FIGURE 1. The amino and carboxyl termini of LMF1 are oriented to opposite sides of the ER membrane. A, GFP-LMF1 and LMF1-GFP fusions were constructed with green fluorescent protein added to the amino and carboxyl terminus of human LMF1, respectively. HeLa cells transfected with each Lmf1 construct were then subjected to an FPP assay. When trypsin was added at 80 s, access was restricted to cytoplasmic proteins because, unlike ER membranes, the plasma membrane is selectively permeabilized by digitonin added in the assay at t = 0. In the case of GFP-LMF1, the rapid loss of fluorescent signal after the addition of trypsin indicates the location of the amino terminus of LMF1 as facing the cytoplasm; in contrast, the persistence of a signal from LMF1-GFP indicates that the carboxyl terminus of LMF1 is oriented to the ER lumen. B, illustration of the validity of the FPP assay in determining the correct orientation of the amino and carboxyl termini of a known membrane protein. CD3δ is a type 1 membrane protein with its carboxyl and amino termini facing the cytoplasm and ER lumen, respectively. Transfection of HeLa cells with CD3δ-YFP and eCFP-CD3δ confirms that, unlike the cytoplasm-facing carboxyl terminus of CD3δ, the ER lumen-facing amino terminus is protected from trypsin digestion. C, graphs of fluorescent signal intensity versus assay time are shown for all GFP constructs. Each graph shows the percentage of signal remaining relative to t = 0, and each point is the average of 4–6 cells from at least three independent experiments. Error bars show S.D.
Figure 2. Lmf1 has five TM domains. A, MemO-predicated TM domains are positioned with respect to a hydropathy plot of mouse Lmf1. The scores refer to the percentage of times that a given TM domain is predicted in both mouse and human LMF1 orthologs across all eight prediction methods (see supplemental Table S1). Hydropathy values \( \geq 1.6 \) are consistent with TM domains. B, models of Lmf1 membrane topology are illustrated using three, five, or seven TM domains. The three-TM model is built with domains having a top score of 100%; the five-TM model includes two additional domains having the next highest scores (88 and 75%); and the seven-TM model includes all MemO-predicted domains (supplemental Table S1). The position of amino acid 243 is given as a reference in all three models. The cylinders represent the TM domains having an \( \alpha \)-helical structure; DUF1222 is indicated in black. C, the dipeptide NA sequence was inserted after position 243 of mouse Lmf1 to complete an ectopic NAT glycan attachment site (Lmf1\textsuperscript{InsNA243}); the single, naturally occurring glycosylation site at position 430 was removed (N430A; see supplemental Fig. S2). After transfection of HEK293 cells, the cell lysates were treated with (+) or without (−) endo H, which removes N-linked glycans of the high mannose type. Lmf1\textsuperscript{InsNA243} was visualized after Western blot analysis of the treated lysates using the c-Myc epitope tag. The mobility of Lmf1\textsuperscript{InsNA243} treated with endo H was equivalent to unglycosylated Lmf1 (Lmf1\textsuperscript{N430A}; see supplemental Fig. S2).
To limit possible models of LMF1 topology based on MemO prediction results, the cytoplasmic and/or ER lumenal orientation of the amino and carboxyl termini of LMF1 was determined. Besides identifying the location of LMF1 termini, this analysis constrains models of LMF1 topology to having either an even or an odd number of TM domains. To determine orientation, a FPP assay was employed (25), which measures in real time the trypsin proteolysis of fluorescent tags fused to the terminus of a membrane protein of interest. Digitonin, which is added to cells transfected with such a fusion construct, selectively permeabilizes the plasma membrane, and thus fluorescent tags oriented to the cytoplasm by fusion to the end of a membrane protein are accessible to trypsin digestion. In contrast, tags oriented to the ER lumen are protected from trypsin, as the cholesterol-poor ER membrane is relatively resistant to digitonin (26).

Constructs were prepared with the GFP fused to either the amino or carboxyl terminus of human LMF1 (GFP-LMF1 and LMF1-GFP, respectively). The fusion constructs exhibited LMF1-specific activities comparable with a wild type control (supplemental Fig. S1), indicating a native-like topology. When the FPP assay was applied to both constructs (Fig. 1, A and C), GFP at the amino terminus of LMF1 was clearly accessible to trypsin proteolysis, whereas it remained protected when fused to the LMF1 carboxyl terminus. Thus, the amino and carboxyl termini of LMF1 were oriented oppositely to the cytoplasm and the ER lumen, respectively, indicating that LMF1 must have an odd number of TM domains. The ability of the FPP assay to identify the orientation accurately was validated using a known type 1 membrane protein, CD3ε/H9254 (Fig. 1, B and C).

The FPP assay and the MemO prediction analyses were used together to model LMF1 membrane topology based on an odd number (3, 5, or 7) of TM domains. Fig. 2A shows the seven MemO-predicted TM domains superimposed over a hydropathy plot of mouse Lmf1 (27). As expected, the predicted TM domains were localized over peaks with hydropathy scores of >1.6 (window = 19), indicative of membrane-spanning regions (27). Fig. 2B shows a three-TM model of Lmf1 constructed using the three predicted TM domains scoring 100% and a five- and seven-TM model including predicted TM domains having successively lower scores (supplemental Table S1); α-helical TM domains are represented as cylinders. For reference, amino acid position 243 is highlighted in Fig. 2B, showing that this residue is within a loop directed to the cytoplasm in the three- and seven-TM models of Lmf1; in contrast, it is positioned on a loop facing the ER lumen in the five-TM model. The domain of unknown function (DUF) 1222 is indicated as a black line, whereas TM domains within DUF1222 are solid black cylinders (Fig. 2B).
To differentiate between the three possible models, an ectopic glycan attachment site was introduced at position 243 of mouse Lmf1 (Fig. 2C, Lmf1 InsNA243), which has its single, naturally occurring glycosylation site at position 430 removed (supplemental Fig. S2, Lmf1 N430A). Because glycosylation can occur only within the ER lumen, Lmf1 InsNA243 was transfected into HEK293 cells and tested for the presence of attached glycan chains of the high mannose type, indicative of ER-based glycosylation (28). Western blot analysis of the expressed Lmf1 InsNA243 protein showed a band that had increased mobility when treated with endo H (Fig. 2C), which selectively removes high mannose chains from glycoproteins. Thus, site 243 is glycosylated, indicating its exposure to the oligosaccharyltransferase complex in the ER lumen (28). The model consistent with this result is the one based on five TM domains (Fig. 2B); in this model only, position 243 resides on a loop oriented to the ER lumen.

A similar strategy, using additional inserted ectopic glycan attachment sites, was used to verify mouse Lmf1 topology based on the five-TM model (Fig. 3A). Placement of ectopic sites on the model are shown in Fig. 3A, including the single, naturally occurring site at position 430 (supplemental Fig. S2). Notably, N-linked glycosylation at site 430 places the C-terminal domain of mouse Lmf1 within the ER lumen, consistent with the ER localization of the carboxyl terminus of human LMF1 (Fig. 1A). Like position 430, ectopic sites placed on loops A and C, also predicated to be oriented lumenally (Fig. 3A), were similarly glycosylated (Fig. 3B). In contrast, site 28, predicated to be within the N-terminal domain of mouse Lmf1, was not utilized for glycosylation, consistent with the localization of the amino
terminus of human LMF1 to the cytoplasm (Fig. 1A). Like site 28, all additional ectopic sites predicted to be within loops facing the cytoplasm were not glycosylated (Fig. 3, A and B). Notably, functional analysis of all ectopic glycan mutants showed activity as high as the wild type control (data not shown), indicating that native topology was retained. Separate transfection experiments confirmed the glycosylation status of all ectopic sites (supplemental Fig. S3).

Finally, the five-TM model was confirmed by using human LMF1-GFP fusion constructs in the FPP assay (Fig. 4). Carboxyl-terminal GFP fusions were made with human LMF1 truncated at sites within loops predicted by the five-TM model; these included ER-facing loop A (95-GFP), cytoplasm-facing loop B (200-GFP), ER-facing loop C (250-GFP), and cytoplasm-facing loop D (350-GFP). As shown in Fig. 4, the predicted location of each loop was confirmed by the FPP assay, with cytoplasm- and ER-facing loops showing accessibility or resistance to trypsin proteolysis, respectively. Thus, all results using ectopic glycan tags and FPP analysis on both mouse and human LMF1 constructs are consistent with a five-TM model of topology, subdividing LMF1 into three domains facing the cytoplasm (the amino-terminal domain and loops B and D) and three oriented to the ER lumen (loops A and C and the carboxyl-terminal domain). The smallest-to-largest domains of LMF1 are the amino-terminal domain (~28 aa), loop D (~46 aa), loops A and B (~56 aa), loop C (~71 aa), and the large carboxyl-terminal domain (~186 aa). The evolutionarily conserved DUF1222 begins in loop B and extends through most of the carboxyl-terminal domain. Indeed, as shown below, those regions comprising DUF122 are the most likely to have functional roles in lipase maturation.

Coding Variants of Lmf1 Affecting DUF1222—The mouse Lmf1 gene has 11 exons encoding a protein that is 574 amino acids in length (9). However, we determined the existence of several naturally occurring Lmf1 splice variants that eliminate exons 4–11 encoding DUF1222 (Fig. 5A). Two of these variants (1–2–3–3A and 1–2–3*) are present in the dbEST database, and all three have been detected by RT-PCR (Fig. 5A) and confirmed by sequencing. All splice variants comprise exons 1–3, and two include an alternative exon 3A ending in a polyadenylation signal. One of these latter variants (1–2–3*–3A) uses a cryptic splice donor (Fig. 5A, SD) site that results in the inclusion of some intronic sequence downstream of exon 3. The intronic sequence is also present in the variant without exon 3A (1–2–3*), as it uses an alternative downstream polyadenylation site (Fig. 5A, pA). All encode Lmf1 proteins lacking the DUF1222 domain (ΔDUF), and splice variant 1–2–3–3A has the largest coding potential (188 amino acids).

Fig. 5B illustrates the predicted topology of wild type (full-length) Lmf1 and the 1–2–3–3A splice variant, referred to hereafter as ΔDUF, based on the five-TM model described above. DUF1222 begins in loop B and extends throughout loops C and D and most of the carboxyl-terminal domain (Fig. 5B, black line). In contrast, the ΔDUF variant includes the amino-terminal domain, loop A, and the part of loop B preceding DUF1222 (Fig. 5B, gray line). The truncation of Lmf1 caused by the cld
Lmf1 without DUF1222 (Lmf1cld) was enough to cause loss of function (Fig. 6), indicating that Lmf1 interacts best with the two lipases dependent on Lmf1 for their maturation. In contrast, the DUF1222 truncation (Lmf1ΔDUF) failed to co-isolate with any lipase protein, verifying the placement of the lipase interaction site within loop C of DUF1222. When loop C was reinstated as the only ER-facing domain of DUF1222 (Lmf1cld), the resulting Lmf1 protein was shown to co-isolate with HL and LPL but only very poorly with PL (Fig. 7, lanes 9–11). Thus, loop C is critical for lipase interaction, but cannot function independently of the large carboxyl-terminal domain in carrying out the function of lipase maturation (Fig. 6A).

Although TM prediction is limited by the many diverse factors affecting TM usage and insertion during polytopic protein folding. Indeed, some polytopic proteins, such as the cystic fibrosis transmembrane conductance regulator, can assume multiple topological forms (30–33). In this study, we used a number of predictors to locate putative TM domains within LMF1; these predictions were based on Kyte-Doolittle hydrophy plots (27), the dense alignment surface (DAS) method (34), and machine-learning algorithms (hidden Markov models such as TMHMM and HMMTOP), which are trained by analyzing the residues that tend to occupy defined regions in integral membrane proteins (31). However, the training set used by such programs is limited because of sparse information regarding eukaryotic membrane proteins; indeed, hydrophobic cores within soluble regions can be misidentified as TM domains, whereas short TM domains may be overlooked entirely (31). Because of these limitations, we used a number of TM prediction methods to interrogate two mammalian LMF1 orthologs (mouse and human). We assumed that for a given TM domain, a high degree of agreement among different methods in two homologs would provide a more accurate outcome than relying on a single method alone. We started with the predicators employed by the LOCATE data base (24) and used the resulting MemO-derived consensus of seven TM domains as our basis for model building. We compared these MemO predictors along with three others to calculate an alignment score for the seven predicted TM domains (supplemental Table S1) and then used the scores in a hierarchical manner to build additional models based on these three and five TM domains. In all cases, detect Lmf1-lipase interaction, lipase constructs fused to a TAP tag were used to achieve lipase affinity purification under conditions favoring the retention of protein-protein interactions (18). Three lipase-TAP constructs were used to evaluate Lmf1-lipase interaction: LPL, HL, and PL. Although all three are members of the lipase gene family (29), only maturation of LPL and HL is dependent on Lmf1 (5, 6, 9).

Lmf1wt, Lmf1cld, and Lmf1ΔDUF expression constructs were co-transfected in HEK293 cells along with each lipase-TAP construct and a vector only control. Fig. 7 shows that the various lipase-TAP proteins (panel A) and Lmf1 protein variants (panel B) were co-expressed in transfected HEK293 cells lysates; these lysates were then used for affinity purification of lipase proteins. The resulting lipase-TAP isolates were subjected to Western blot analysis and visualized using the Lmf1-specific c-Myc epitope tag (Fig. 7C). Notably, the Lmf1wt protein was shown to co-isolate with HL and LPL but only very poorly with PL (Fig. 7, lanes 1–3), indicating that Lmf1 interacts best with the two lipases dependent on Lmf1 for their maturation.

To verify the importance of DUF1222 in lipase maturation, Lmf1wt, Lmf1ΔDUF, and Lmf1cld protein variants were examined for their ability to restore LPL activity in mutant cld cells. Although LPL transfected into mutant cld cells is inactive and is retained within the ER, co-transfection of cells with fully functional Lmf1 can restore LPL activity at both global and type levels (9, 19). Such an assay was used to determine the ability of Lmf1ΔDUF and Lmf1cld variants to restore LPL activity in mutant cld cells. Lmf1 without DUF1222 (Lmf1ΔDUF) failed to restore LPL activity (supplemental Fig. S4), confirming the importance of this evolutionarily conserved domain in lipase maturation. In fact, removal of the carboxyl-terminal domain of DUF1222 by the cld truncation (Lmf1cld) was enough to cause loss of function (supplemental Fig. S4).

Importantly, the loss of function caused by the DUF1222 deletion or cld truncation was not due to displacement of Lmf1 from the ER, the site of lipase maturation. Fig. 6 shows that Lmf1ΔDUF and Lmf1cld, along with the wild type control (Lmf1wt), co-localized with calnexin, a type 1 membrane protein containing a classical ER retention signal. Thus, although ER localization of Lmf1 remains unimpaired when DUF1222 is absent, this evolutionarily conserved domain is essential in carrying out the function of lipase maturation.

**Lipase Interaction with Lmf1**—Although the mechanisms underlying Lmf1-induced lipase maturation are unknown, a first step in identifying the process is to determine whether Lmf1 physically interacts with lipase proteins in the ER. To determine whether these variants express stable proteins.

Calnexin is a type 1 ER membrane protein with a classical ER retention signal. The downstream carboxyl-terminal sequence (Fig. 5B) expressed Lmf1 c-Myc epitope tag was detected only in transfected cells. Calnexin labeled antibodies, endogenous calnexin was visualized in all cells; the expression construct for full-length Lmf1wt, Lmf1ΔDUF, and the DUF1222-truncated splice variant Lmf1ΔDUF were transfected into HEK293 cells. Using fluorescently labeled antibodies, endogenous calnexin was visualized in all cells; the expressed Lmf1 c-Myc epitope tag was detected only in transfected cells. Calnexin is a type 1 ER membrane protein with a classical ER retention signal.
we found the high degree of homology between mouse and human LMF1 (88% amino acid similarity) reflected in the conservation of the predicted TM domains, which were essentially identical between the two orthologs (supplemental Table S1). Because of this conservation, we tested our models experimentally using mouse and human LMF1 in an interchangeable manner. Our strategy of employing tagging methods at LMF1 termini and within predicted loops is a proven means of testing topological models of polytopic proteins located in the plasma membrane (35, 36), mitochondria (37), and ER (38–40). The agreement in results between LMF1 orthologs using two different techniques (FPP analysis and ectopic glycan tagging) provided confidence in the five-TM model as a reliable representation of mammalian LMF1 topology. This representative model provides a structural basis for determining the functions associated with the six domains identified in our study, half of which faced the cytoplasm (the amino-terminal domain and loops A and D) or the ER lumen (loops A and C and the carboxyl-terminal domain). This model also provides a first look at the arrangement of the DUF1222, an evolutionarily conserved domain located toward the carboxyl-terminal end of many hypothetical integral membrane proteins in bacteria and eukaryotes (14). Indeed, LMF1 is the first example of this family with a known function and, because of this study, an experimentally tested topology. It should be noted, however, that the boundary limits given for each of the TM domains in LMF1 are estimates. This is because the accuracy of TM boundaries assigned by prediction methods has been difficult to assess, as determining TM boundaries by biochemical and structural approaches has proven difficult (41). As shown in supplemental Table S1, for a given TM domain, different prediction methods give boundary limits that can vary by as much as 10–13 amino acids. We used the boundaries provided by the mouse MemO consensus to define the limits of the TM domains in our model of LMF1 (Fig. 3A).
The importance of DUF1222 in lipase maturation has been implied by loss-of-function mutations that truncate this domain at its carboxyl-terminal end. The cld mutation represents the most severe truncation because of intragenesis of a polyadenylation site occurring in the long term repeat of an inserted murine retrovirus in intron 7 of the mouse Lmf1 gene (9). It removes all of the downstream sequence starting from position 361, eliminating the fifth TM domain (aa 368–388) and the entire carboxyl-terminal domain (aa 389–574). However, despite this large deletion, the Lmf1cld protein can be expressed (Fig. 5C) and remains localized to the ER membrane (Fig. 6B). Moreover, the carboxyl-terminal GFP fusion of human LMF1 truncated at position 350 (near position 361) showed a normal topology based on the expected cytoplasmic orientation of the truncation site within loop D (Fig. 4). This result suggests that cld does not affect Lmf1 topology upstream of the truncation site. Notably, Lmf1cld still retains loop C and its associated lipase interaction site, although it is incapable of functioning as a lipase maturation factor. In fact, Lmf1cld appears to bind more LPL and HL than does the wild type protein (see legend for Fig. 7). This may reflect lack of an adjacent carboxyl-terminal domain in Lmf1cld and suggests that although lipases may initially bind to loop C, subsequent interaction with the large carboxyl-terminal domain absent in Lmf1cld may in fact trigger their eventual release. Indeed, it is intriguing to speculate that the juxtaposition of loop C and the carboxyl-terminal domain may promote cooperation between these domains to facilitate the related functions of lipase binding and maturation.

In conclusion, our model of LMF1 topology provides a rational basis for pursuing structure/function studies to identify the specific roles of individual domains of LMF1. This model will also be useful in identifying potential sites for post-translational modifications that may in turn regulate LMF1 maturation.

Acknowledgment—We greatly appreciate the gift of eCFP-CD3 and CD8-YFP constructs from Dr. J. Lippincott-Schwartz.

REFERENCES
1. Hasham, S. N., and Pillarisetti, S. (2006) Clin. Chim. Acta 372, 179–183
2. Mead, J. R., Irvine, S. A., and Ranji, D. P. (2002) J. Mol. Med. 80, 753–769
3. Perret, B., Mballe, M., Martinez, L., Tercé, F., Barbaras, R., and Collet, X. (2002) J. Lipid Res. 43, 1163–1169
4. Santamarina-fojo, S., González-Navarro, H., Freeman, L., Wagner, E., and Nong. Z. (2004) Arterioscler. Thromb. Vasc. Biol. 24, 1750–1754
5. Scow, R. O., Schultz, C. J., Park, J. W., and Blanchette-Mackie, E. J. (1998) Chem. Phys. Lipids 93, 149–155
6. Briquet-Laugier, V., Ben-Zeev, O., White, A., and Doolittle, M. H. (1999) J. Lipid Res. 40, 2044–2058
7. Olivecrona, T., Chernick, S. S., Bengtsson-Olivecrona, G., Paterniti, J. R., Jr., Brown, W. V., and Scow, R. O. (1985) J. Biol. Chem. 260, 2552–2557
8. Davis, R. C., Ben-Zeev, O., Martin, D., and Doolittle, M. H. (1990) J. Biol. Chem. 265, 17960–17966
9. Péterfy, M., Ben-Zeev, O., Mao, H. Z., Weissglas-Volkov, D., Aouizerat, B. E., Pullinger, C. R., Frost, P. H., Kane, J. P., Malloy, M. I., Reue, K., Pajukanta, P., and Doolittle, M. H. (2007) Nat. Genet. 39, 1483–1487
10. Reue, K., and Doolittle, M. H. (1996) J. Lipid Res. 37, 1387–1405
11. Coleman, T., Seip, R. L., Gimble, J. M., Lee, D., Maeda, N., and Semenkovich, C. F. (1995) J. Biol. Chem. 270, 12518–12525
12. Ben-Zeev, O., and Doolittle, M. H. (2004) J. Biol. Chem. 279, 6171–6181
13. Ben-Zeev, O., Mao, H. Z., and Doolittle, M. H. (2002) J. Biol. Chem. 277, 10727–10738
14. Sonnhammer, E. L., Eddy, S. R., and Durbin, R. (1997) Proteins 28, 405–420
15. Boedecker, J. C., Doolittle, M. H., and White, A. L. (2001) J. Lipid Res. 42, 1858–1864
16. Péterfy, M., Phan, J., and Reue, K. (2005) J. Biol. Chem. 280, 32883–32889
17. Ben-Zeev, O., Doolittle, M. H., Davis, R. C., Elovson, J., and Schotz, M. C. (1992) J. Biol. Chem. 267, 6219–6227
18. Doolittle, M. H., Ben-Zeev, O., Basillais, S., Whitelegge, J. P., Péterfy, M., and Wong, H. (2009) J. Lipid Res. 50, 1173–1184
19. Yin, F., Doolittle, M. H., and Péterfy, M. (2009) J. Lipid Res., in press
20. Briquet-Laugier, V.; Péterfy, M., and Doolittle, M. H. (1999) Methods Mol. Biol. 109, 87–97
21. Palin, N. J., and Stokoe, S. (2004) J. Biol. Chem. 279, 218–229
22. Briquet-Laugier, V. (1998) J. Lipid Res. 39, 2101–2109
23. Lippincott-Schwartz, J. (2006) J. Membr. Biol. 213, D213–D217
24. Davis, R. C., Canfield, V., Levenson, R., and Gros, P. (1995) Biochemistry 34, 4402–4411
25. Kast, C., Canfield, V., Levenson, R., and Gros, P. (1996) J. Biol. Chem. 271, 9240–9248
26. Gonzalez-Baro, M. R., Granger, D. A., and Coleman, R. A. (2001) J. Biol. Chem. 276, 43182–43188
27. Joyce, C. W., Shelness, G. S., Davis, M. A., Lee, R. G., Skinner, K., Anderson, R. A., and Rudel, L. L. (2000) Mol. Biol. Cell 11, 3675–3687
28. Lin, S., Cheng, D., Liu, M. S., Chen, J., and Chang, T. Y. (1999) J. Biol. Chem. 274, 23276–23285
29. Schenkel, J. N., and Lippincott-Schwartz, J. (2006) Nat. Meth. 3, 523–527
30. Dowhan, W., and Bogdanov, M. (2009) J. Lipid Res. 50, 993–999
31. Dowhan, W., and Bogdanov, M. (2009) Annu. Rev. Biochem. 73, 1019–1049
32. Ovalle, H. V., and Orrego, V. (2001) J. Biol. Chem. 276, 27155–27161
33. Wong, H., and Schotz, M. C. (2002) J. Lipid Res. 43, 993–999
34. Sirois, M., and Tardif, C. (2006) J. Biol. Chem. 281, 1333–1335
35. Turner, R. J. (2003) J. Membr. Biol. 192, 149–157
36. Byun, J. K., Hwang, J. Y., and Pajukanta, P., and Doolittle, M. H. (2003) J. Bioenerg. Biomembr. 36, 17960–17966
37. Gonzalez-Baro, M. R., Granger, D. A., and Coleman, R. A. (2001) J. Biol. Chem. 276, 43182–43188
38. Joyce, C. W., Shelness, G. S., Davis, M. A., Lee, R. G., Skinner, K., Anderson, R. A., and Rudel, L. L. (2000) Mol. Biol. Cell 11, 3675–3687
39. Lin, S., Cheng, D., Liu, M. S., Chen, J., and Chang, T. Y. (1999) J. Biol. Chem. 274, 23276–23285
40. Man, W. C., Miyazaki, M., Chu, K., and Ntambi, J. M. (2006) J. Biol. Chem. 281, 1251–1260
41. Deber, C. M., Wang, C., Liu, L. P., Prior, A. S., Agrawal, S., Muskat, B. L., and Cuticchia, A. J. (2001) Protein Sci. 10, 212–219