THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 274, No. 20, Issue of May 14, pp. 14122–14129, 1999

Printed in U.S.A.

Targeting Proteins to the Lumen of Endoplasmic Reticulum Using N-terminal Domains of 11β-Hydroxysteroid Dehydrogenase and the 50-kDa Esterase*

(Received for publication, December 31, 1998, and in revised form, February 17, 1999)

Hassan Mziaut‡, George Korza‡, Arthur R. Hand§, Craig Gerard¶, and Juris Ozols‡‖

From the ‡Department of Biochemistry and §Electron Microscopy, University of Connecticut Health Center, Farmington, Connecticut 06030-3305 and the ¶Department of Pediatrics, Children's Hospital, Harvard Medical School, Boston, Massachusetts 02115

Previous studies identified two intrinsic endoplasmic reticulum (ER) proteins, 11β-hydroxysteroid dehydrogenase, isozyme 1 (11β-HSD) and the 50-kDa esterase (E3), sharing some amino acid sequence motifs in their N-terminal transmembrane (TM) domains. Both are type II membrane proteins with the C terminus projecting into the lumen of the ER. This finding implied that the N-terminal TM domains of 11β-HSD and E3 may constitute a lumenal targeting signal (LTS). To investigate this hypothesis we created chimeric fusions using the putative targeting sequences and the reporter gene, *Aequorea victoria* green fluorescent protein. Transfected COS cells expressing LTS-green fluorescent protein chimeras were examined by fluorescent microscopy and electron microscopic immunogold labeling. The orientation of expressed chimeras was established by immunocytofluorescent staining of selectively permeabilized COS cells. In addition, protease protection assays of membranes in the presence and absence of detergents was used to confirm lumenal or the cytosolic orientation of the constructed chimeras. To investigate the general applicability of the proposed LTS, we fused the N terminus of E3 to the N terminus of the NADH-cytochrome *b*5 reductase lacking the myristoyl group and N-terminal 30-residue membrane anchor. The orientation of the cytochrome *b*5 reductase was reversed, from cytosolic to lumenal projection of the active domain. These observations establish that an amino acid sequence consisting of short basic or neutral residues at the N terminus, followed by a specific array of hydrophobic residues terminating with acidic residues, is sufficient for lumenal targeting of single-pass proteins that are structurally and functionally unrelated.

A number of proteins are resident and catalytically active in the lumen of the endoplasmic reticulum (ER).[1] The best characterized motif targeting this compartment is for soluble proteins bearing C-terminal sequence KDEL (1, 2). The molecular signals responsible for the insertion and retention of membrane proteins in the ER with a lumenal orientation of their functional domains are poorly understood. Most ER proteins are targeted for membrane insertion by a hydrophobic signal peptide at the N terminus (3). During translation, the signal peptide is recognized by a cytosolic RNA/protein complex termed the signal recognition particle (SRP) (4). Upon binding a signal peptide, the SRP arrests the protein synthesis until it has facilitated transfer of the nascent polypeptide-ribosome complex to the translocation channel (5). The signals controlling folding and eventual topology of ER lumenal proteins after the nascent chain is transferred to the translocation channel are unknown. One early critical event in the multistep folding process is whether or not the signal peptide is removed. If the signal peptide is removed, then the newly formed N terminus is generally located in the lumen of the ER. In cases where the signal peptide is not removed, the resulting N-terminal hydrophobic segment may function as a membrane anchor, directing the transmembrane insertion of the nascent polypeptide. Therefore, membrane proteins with uncleaved signal peptide may have either cytosolic or lumenal orientation. Mutagenesis experiments on a number of membrane proteins have revealed that the charge distribution, the length of the membrane segment, as well as the charge of the segments following the membrane segment are some of the factors that determine the specific topology of such proteins (6). Detailed analysis of these factors have been complicated by use of multispanning model proteins with a lack of agreement on the topology. Moreover, additional ambiguities are encountered in trying to overexpress polytopic eukaryotic membrane proteins in prokaryotic hosts. Previously, we identified the membrane binding anchor of cytochrome b5 (7), cytochrome *b*5 reductase (8), cytochromes P-450 (9–11), epoxide hydrolase (12), Δ9 stearyl coenzyme A desaturase (13), and three forms of flavin-containing monooxygenases (14). All of these ER proteins appeared to be oriented toward the cytosolic side of the ER membrane by single or multispanning membrane segments. In further studies to define the proteins found or oriented in the lumenal compartment of the ER, we identified two isoforms of 60-kDa carboxylesterases (15) and the lumenal NADP glucose-6-phosphate dehydrogenase (16). The esterases and the glucose-6-phosphate dehydrogenase were devoid of any transmembrane segments. The lumenal esterases carried a C-terminal segment HIEL- and HTEL-related KDEL, the ER lumen retention motif (1). The primary structure of the glucose-6-phosphate dehydrogenase failed to display sequence segments or a motif that would be responsible for its lumenal orientation or retrieval. We then extended our studies to oligosaccharyltransferase complex (OT)

* This work was supported by Grant R01 GM-26351 from the National Institutes of Health (to J. O.). 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.

‖ To whom correspondence should be addressed: Dept. of Biochemistry, School of Medicine of the University of Connecticut Health Center, Farmington, CT 06030-3305. Tel.: 860-679-2211; E-mail: ozols@sun.uchc.edu.

[1] The abbreviations used are: ER, endoplasmic reticulum; 11β-HSD, 11β-hydroxysteroid dehydrogenase; E3, liver microsomal 50-kDa esterase/*N*-deacetylase; GFP, green fluorescent protein; LTS, lumenal targeting signal; PAGE; polyacrylamide gel electrophoresis; PBS, phosphate-buffered saline; SLO, streptolysin O; OT, oligosaccharyltransferase; AEBSF, 4-(2-aminoethyl)-benzenesulfonyl fluoride; PCR, polymerase chain reaction; BSA, bovine serum albumin.

(18). The OT from mammalian tissues was found to be a trimolecular complex consisting of ribophorins I and II and a 50-kDa protein. The predicted membrane orientation of OT indicated that both ribophorins and the 50-kDa proteins have large N-terminal lumenal domains and a short C-terminal cytoplasmic domain (18). Studies on the orientation of the 11β-HSD (17) and the 50-kDa esterase/N-deacetylase (E3) (19) revealed that both proteins shared a short N-terminal cytoplasmic domain and a large C-terminal lumenal domain.

In the present study, we investigated whether the N-terminal membrane binding segment of 11β-HSD and E3 may act as a targeting sequence for the lumenal orientation of proteins in the ER. Here we show that a transmembrane segment consisting of basic residues at the N terminus followed by an array of specific 17 hydrophobic residues terminating with acidic residues constitute a lumenal targeting signal for a set of single-membrane-spanning proteins that are otherwise structurally and functionally unrelated.

EXPERIMENTAL PROCEDURES

Materials—Chemical products were purchased from Sigma unless otherwise noted. COS-7 cells were from American Type Culture Collection (ATCC). Cell reagents, restriction, and modification enzymes were from Life Technologies, Inc. The GFPS65T mutant protein in pEGFP-N1 vector and an anti-GFP mouse monoclonal antibody were from CLONTECH. Anti-stearyl CoA-desaturase polyclonal antibody was raised in a rabbit by injecting keyhole limpet hemocyanin-conjugated 20-residue synthetic peptide corresponding to residues 338–358 of stearyl CoA-desaturase antigen (20). Avian anti-*b*5 and anti-reductase were raised against the polar moieties of rat cytochrome *b*5 and cytochrome *b*5 reductase (19). Anti-calreticulin rabbit polyclonal antibody was from Affinity Bioreagents. Alkaline phosphatase-conjugated anti-mouse and anti-rabbit IgG were from Sigma. Biotin-coupled goat anti-mouse IgG was from Molecular Probes. Iodocarbocyanine (Cy3)-streptavidin and Cy3-conjugated goat anti-rabbit IgG were from Jackson Immunoresearch Laboratories. Proteinase K and 4-(2-aminoethyl)-benzenesulfonyl fluoride (AEBSF) were from Roche Molecular Biochemicals. Gold-labeled secondary antibody was from Amersham Pharmacia Biotech.

Construction of Plasmids—All expression vectors were driven by the cytomegalovirus promoter-enhancer contained in the pEGFP-N1 vector. The N tag-GFP chimeras were constructed by inserting the coding region of GFP at the C terminus of the first 23 and 34 amino acids corresponding to lumen targeting signal (LTS) of 11-βHSD and esterase 3, respectively and variants thereof, as follows (see Fig. 1). First, cDNA encoding residues 1–34 of human esterase 3 preceded by Kozak sequence and *Eco*RI site at 5′ end and extended by *Bcl*I site at 3′ was synthesized using two long overlapping oligonucleotides (5′-CTTCGAATTCCCACCATGGGAAGAAAATCGCTGTACCTTCTGATTGTGGGGATCCTCAATTT-3′ as sense long oligo and 5′-TGGATTGATCATTCTCCATGGCTCCTCAACGTTATCTGGGAGAGGCGTATAAATATAATG-3′ as antisense long oligo) and two short oligonucleotides (5′-CTTCGAATTCCCACCATG-3′ as sense short oligo and 5′-TGGATTGATCATTCTCCA-3′ as antisense short oligo) and the chain polymerase reaction (PCR). Second, the coding sequence of GFP was amplified from pEGFP-N1 by PCR using a 5′-oligonucleotide, which contains a *Bcl*I restriction site (5′-ATGGTGATCAAGGGCGAGGAGCTG-3′) and 3′-oligonucleotide containing *Not*I site (5′-CCTCTACAAATGTGGTATGGC-3′). This was done in order to remove the Kozak sequence preceding the first ATG codon of GFP in pEGFP-N1 vector so that the initiation codon introduced in the N-tag sequences will be used by N-tag-GFP chimeras.

The resulting PCR products were purified and digested with *Eco*RI-*Bcl*I and *Bcl*I-*Not*I for N-tags and GFP, respectively and inserted into *Eco*RI-*Not*I-digested pEGFP-N1 vector in one-step ligation. All the N-tag-GFP chimeras and their mutated or deleted variants were constructed using the experimental procedure described above.

The expression vector pEN1.rab E3 N-b5 red.(31–300) encodes a fusion protein consisting of rabbit esterase 3 LTS, followed by three novel amino acids (PPV) encoded by the restriction site for *Age*I, by amino acids 31–300 of bovine liver NADH-cytochrome *b*5 reductase, and FLAG epitope TAG (LARIKRTGDGSHKSS). This fusion protein was expressed in pEN1 vector obtained by removing the coding sequence of GFP from pEGFP-N1. The rab E3 N-b5 red.(31–300) was constructed as follows. First, the sequence corresponding to amino acids 31–300 of reductase was amplified by PCR using as a template the 940-bp DNA fragment coding for entire bovine liver microsomal reductase inserted in pGEM3zf(+) (21) with a pair of primers, 5′-GATCCACCGGTCCCGGCCATCACGCTCGAGAAT-3′ matching with the amino acids 31–37 of beef reductase preceded by *Age*I site, and 5′-AGAGTCGCGGCCGCTTTAGCTACTCTTGTGGCTCCCATCTCCAGTTCTCTTAATCCTGGCTAAGAAGGCGAAGCAGCGTTCTTT-3′ matching with amino acids 294–300. This oligonucleotide included the desaturase epitope, stop codon, and *Not*I site. The PCR fragment was purified, digested with *Age*I-*Not*I and inserted along with the *Eco*RI-*Age*I fragment corresponding to LTS of rabbit esterase3 into *Eco*RI-*Not*I-digested pEN1 vector. The integrity of the constructs were confirmed by DNA sequencing.

Cell Culture and Transient Transfection—COS-7 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal bovine serum, 100 units/ml penicillin, and 100 mg/ml streptomycin. The cells were seeded at approximately 30% confluence onto 35-mm dishes at 37 °C under 5% CO_2. After 24 h, when the cells became 60–70% confluent, they were transfected using the LipofectAMINE reagent according to the manufacturer's instructions. The cell staining was observed 48–72 h after transfection under Zeiss LSM 410 confocal microscope.

Immunoblot Analysis—Particulate components and the soluble fraction of the cells were separated as described previously (22). Insoluble material was removed by centrifugation, samples resolved by SDS-PAGE, and Western blot analysis was performed using anti-mouse monoclonal antibody against GFP. Alkaline phosphatase-conjugated antibodies were used as secondary antibodies for protein visualization.

Isolation and Protease Digestion of Microsomes from COS-7 Cells—Microsomal membranes were prepared as described by Ref. 23. The pellet fraction was enriched in cytochrome *b*5 by adding 10 μl of 0.5 mM pure detergent free cytochrome *b*5 (24). The mixture was resuspended in a solution containing 250 mM sucrose, 10 mM potassium-Hepes, pH 7.2, 50 mM potassium chloride, 2 mM magnesium chloride, and 2 mM calcium chloride to a final volume of 0.2 ml. Protease digestions were performed at 4 °C for 30 min with increasing concentrations of proteinase K in presence or absence of 0.5% Triton X-100. AEBSF (4 mM) was used to terminate the protease digestions. After 15 min of incubation on ice, the membranes were solubilized with SDS gel loading buffer, and the proteins subjected to SDS-PAGE and immunoblotted as described above.

Selective Permeabilization and Immunofluorescence Analysis—Permeabilization of cells using digitonin, streptolysin O, and saponin was performed 48 h after transfection, as described previously (25). Immunocytofluorescent staining was conducted by incubation of the cells with a 1:100 dilution of anti-GFP mouse monoclonal antibody in the permeabilization buffers containing 0.1% BSA. The cells were then rinsed three times and then incubated with a 1:100 dilution of biotin-coupled goat anti-mouse immunoglobulin. After the wash, cells were incubated with streptavidin coupled-iodocarbocyanine (Cy3) diluted 1:100 in the permabilization buffers containing 3% BSA. The cells were rinsed several times before mounting in Mowiol 4.88 and then viewed under Zeiss LSM 410 confocal microscope.

Electron Microscopic Immunogold Labeling—Transfected COS-7 cells with rab E3 N-GFP were fixed with 3% paraformaldehyde plus 0.15% glutaraldehyde in 0.1 M sodium cacodylate buffer, pH 7.4. The cells were allowed to fix for about 1 h at room temperature, the fixative was removed, and the cells stored at 4 °C in 1% paraformaldehyde in 0.1 M cacodylate, pH 7.4. The cells were removed from the flask by scraping and then pelleted in 0.2% BSA in PBS. The cells were resuspended in 0.1% BSA/PBS, transferred to a tube containing 3% low-gelling agarose and pelleted again. The agarose was chilled on ice, and the tip of the tube was cut off. The embedded cell pellet was rinsed in 0.1 M cacodylate buffer, pH 7.4, dehydrated in methanol, infiltrated, and embedded at −20 °C in Lowicryl K4M resin. After polymerization with ultraviolet light (365 nm), thin sections of the embedded cells were cut with a diamond knife and collected on Formvar-coated, 200-mesh nickel grids.

Immunogold labeling was done essentially as described elsewhere (26). Nonspecific binding was blocked with 1% BSA plus 1% instant milk in PBS. The sections were labeled overnight at 4 °C with anti-GFP mouse monoclonal antibody diluted 1:50 in 1% BSA and 5% normal goat serum in PBS. After rinsing with PBS, the bound immunoglobulins were visualized by incubating with 10-nm diameter gold-labeled secondary antibody diluted 1:15 in PBS for 1 h at room temperature. After thorough rinsing with PBS and distilled water, the sections were stained with uranyl acetate and lead citrate and observed in Philips CM10 TEM.

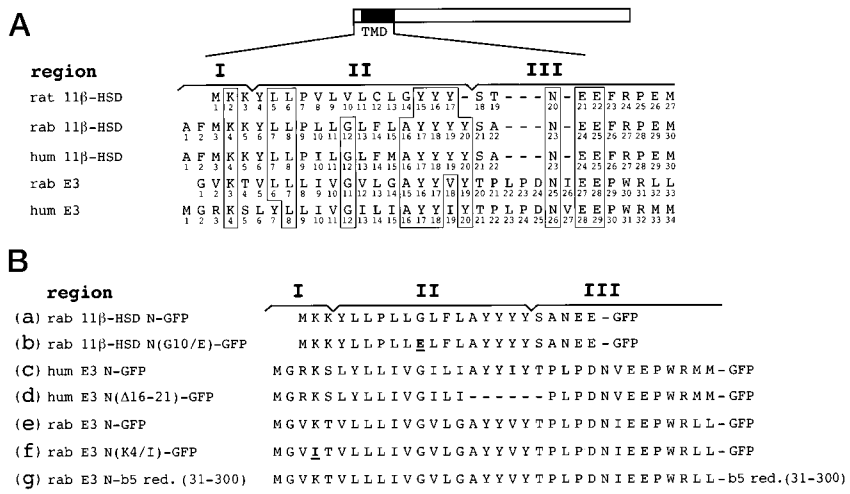

FIG. 1. **N-terminal sequences of native proteins and GFP constructs.** *A*, N-terminal sequences of rat, rabbit, and human 11β-HSD are aligned with those of rabbit and human E3. The positions of N-terminal amino acid residues are *numbered* below the sequence. Conserved amino acids are enclosed in *boxes*. *Dashes* were inserted to maintain alignment of conserved residues. The LTS is denoted by regions I, II and III as indicated above. The Protein Identification Resource accession numbers for the rat, rabbit, and human 11β-HSDs are A34430, A55573, and A41173, respectively. Rabbit and human E3 accession numbers are A58922 and A53856. *B*, amino acid sequences of LTS of native proteins and the derivatives thereof with altered LTS fused to GFPs, which were prepared for the experiments described here. Amino acid residues or segments that were altered are indicated by *letters* in *boldface* or by *dashed line*.

RESULTS

The N Termini of 11β-HSD and Esterase 3 (E3) Determines Their Subcellular Localization—The proposed LTS of rat, rabbit, and human 11β-HSD aligned with sequences of E3 of rabbit and human are shown in Fig. 1*A*. The two protein families share a short positively charged N termini (region I), followed by an array of some 17 hydrophobic residues containing an aromatic cluster Ala-Tyr-Tyr-*X*-Tyr or Gly-Tyr-Tyr-Tyr cluster (region II), terminating with di-glutamyl residues (region III). Sequence similarity beyond the N termini could not be found in these two protein families (19). To ascertain whether N terminus of 11β-HSD is a determinant of its subcellular localization in mammalian cells, constructs were generated encoding N terminus of 11β-HSD fused to GFP (11β-HSD N-GFP) (Fig. 1*B*). The *A. victoria* GFP is widely recognized as a powerful tool in cell biology serving as a reporter for monitoring localization and dynamics of intracellular proteins and organelles (27). When rab 11β-HSD N-GFP (Construct a) was transiently expressed in COS-7 cells and examined by fluorescent microscopy, a staining pattern characteristic of ER localization was seen (Fig. 2*A*). Expression of native GFP showed green fluorescent signals in both the cytoplasm and nucleus (Fig. 2*B*) consistent with the previous studies (28). To explore whether the N-terminal domain of E3 also imparts GFP a membrane localization, GFP constructs hum and rab E3N-GFP (Constructs c and e, respectively) were generated and expressed in COS-7 cells. Cells expressing Constructs c and e showed fluorescence pattern of ER localization, which was similar to that seen with Construct a (Fig. 2, *A*, *C*, and *D*). Analysis of the membrane fractions of cells expressing the Construct e by SDS-PAGE, followed by immunoblotting with GFP antibody showed the localization of GFP construct in the insoluble fraction. Upon SDS-PAGE, a decreased mobility of rab E3 N-GFP as compared with untagged GFP was observed, implying that the signal sequence like N terminus of the construct was not removed by the signal peptidase (Fig. 2*E*). As anticipated, untagged GFP expressed in cells was found only in the soluble fraction.

The N Termini Targets ER Localization—To confirm that the fluorescence distribution along the reticular network observed in Fig. 2 is characteristic of ER localization, we compared the fluorescence pattern of cells expressing rab E3N-GFP (Construct e) with a known ER marker, calreticulin. The fluorescence pattern of Construct e (Fig. 2*F*) is almost identical to the calreticulin staining (Fig. 2*G*). Indeed, the combined fluorescence revealed a yellow staining indicative of the colocalization of the rab E3N-GFP-associated green fluorescence and the calreticulin-associated red fluorescence as evidenced by Fig. 2*H*.

Electron Microscopic Immunogold Labeling—To confirm these results at the ultrastructural level, ultrathin cryosections were analyzed in double-labeling experiments with the immunogold technique, using anti-GFP antibodies to identify the ER. Electron microscopic examination of thin sections of rab E3 N-GFP (Construct e) transfected cells labeled with anti-GFP antibodies revealed the presence of gold particles over the nuclear envelope and the ER (Fig. 3, *A* and *B*). Clusters of heavily labeled ER membranes were found in the perinuclear region. The gold particles were clearly associated with periphery of the ER cisternae. Although the labeling was most frequently seen in association with ribosome-studded membranes, cisternae of apparently smooth ER were also labeled (Fig. 3*C*). The labeling obtained with the polyclonal and monoclonal anti-GFP preparations was approximately equivalent. Nonspecific background labeling was relatively low; very few particles were found over the nuclear chromatin, mitochondria, or other structures.

Transmembrane Topology of Rab E3N-GFP, Construct e—To determine the orientation of Construct e in the membrane, we performed immunofluorescent staining of selectively permeabilized COS-7 cells by saponin, digitonin, and streptolysin O (SLO). As previously reported, digitonin (20 μg/ml) and SLO (200 units/ml) selectively permeabilized plasma membrane leaving internal membranes including ER intact, whereas saponin permeabilized all the membranes (29). For control experiments, in cells expressing untagged GFP treated with digitonin (Fig. 4*A*), or saponin (Fig. 4*C*), the antibodies stained the nucleus as well as the cytoplasm as expected. Cells treated with SLO (Fig. 4*B*) showed a similar staining except for the nuclear membrane, which was not permeabilized by SLO. Cells transfected with cDNA for rab E3N-GFP (Construct e) and permeabilized with either digitonin (Fig. 4*D*) or SLO (Fig. 4*E*) failed to stain with the antibodies to GFP even though these cells expressed the GFP construct as indicated by the green fluores-

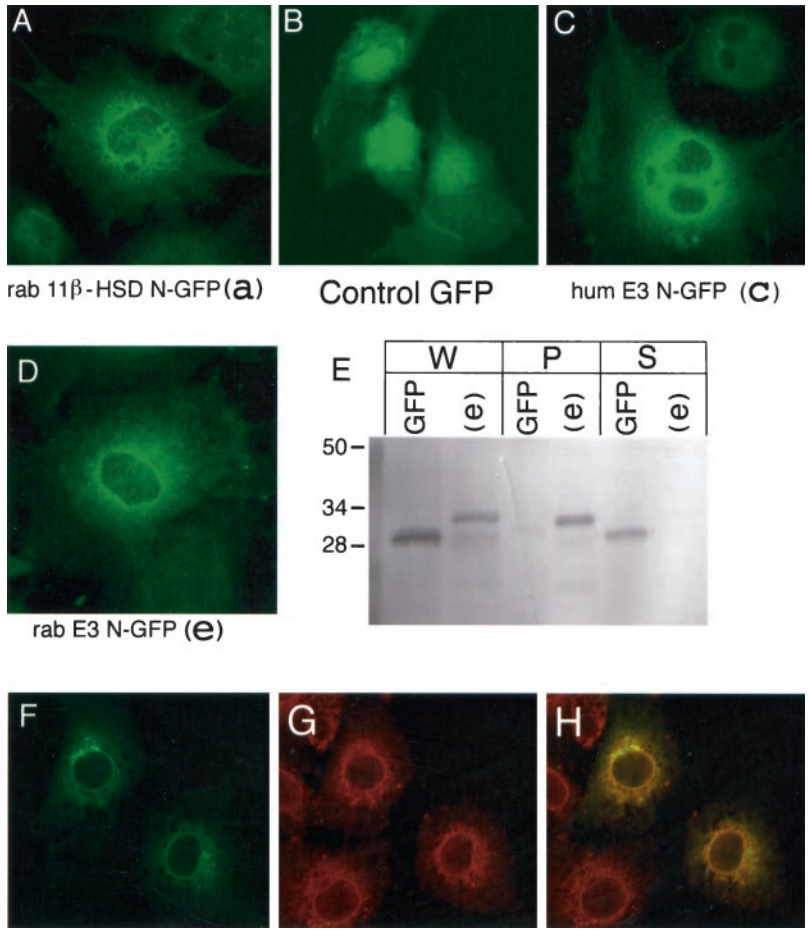

FIG. 2. **Expression of rab11β-HSD N-GFP, hum E3 N-GFP, rab E3 N-GFP, and native GFP in COS-7 cells.** Evidence for the colocalization with the ER marker calreticulin. Representative fluorescence images of COS 7 cells transfected with GFP and the indicated fusion proteins are shown. *A*, Rab 11β-HSD N-GFP (Construct a). *B*, GFP control. *C*, Hum E3 N-GFP (Construct c). *D*, Rab E3 N-GFP (Construct e). The cells were cultured for 48 h after transfection and observed under Zeiss LSM 410 confocal microscope (magnification, ×40). *E*, immunoblot analysis of rab E3 N-GFP (Construct e) and native GFP in cellular subfractions. Cells transfected with cDNA for GFP and rab E3 N-GFP were harvested 48 h after transfection, disrupted, and subjected to centrifugation. The whole cell extract (*W*), soluble fraction (*S*), and pellet fraction (*P*) (50 μg of protein) were subjected to SDS-PAGE and immunoblotted with 5 μg/ml of an anti-GFP mouse monoclonal antibody. *F*, *G*, and *H*, colocalization of rab E3 N-GFP (*green*) with the ER marker calreticulin (*red*) in COS-7 transfected cells permeabilized with 0.2% saponin. *F*, fluorescence image of COS-7 cells expressing rab E3 N-GFP (*green*). *G*, the same cell as in *F* was costained by incubation with a rabbit polyclonal antibody against the ER-resident protein calreticulin followed by incubation with Cy3-conjugated goat anti-rabbit IgG. The cells were then observed under confocal microscope to visualize the staining pattern of the fluorescent dye Cy3 (*red*). *H*, merged images of *F* and *G* reveal considerable colocalization of brightest signal for rab E3 N-GFP and calreticulin in *yellow* (magnification ×40).

cent emission. However, after permeabilization with saponin (Fig. 4*F*), all the cells expressing GFP chimeras became reactive to the GFP antibody. The fact that the GFP fluorescence pattern uniformly overlapped with the GFP antibody staining pattern only when all the cell membranes were completely permeabilized with saponin indicates that the GFP tagged with E3 LTS is oriented toward the lumenal side of ER (Fig. 4*F*).

Evidence for Lumenal Localization by Proteinase K Assay—To confirm the lumenal topography of rab E3N-GFP (Construct e), membranes from cells expressing Construct e were subjected to proteinase K digestion in the presence and absence of detergent. As seen in Fig. 4*G*, in the intact membranes Construct e was not susceptible to proteinase K digestion. Proteolysis of Construct e could be demonstrated when detergent was added to intact membranes prior to incubation with the protease. Taken together, these results are consistent with the lumenal orientation of the expressed chimera. The proteolysis of cytochrome *b*5 was used as a positive control in this proteolysis protection assay. Cytochrome *b*5 is an integral membrane protein oriented toward the cytosolic side of ER, and is readily released from the membrane upon proteolysis (19, 24). Cytochrome *b*5 added to a membrane preparation is spontaneously incorporated into the membrane. As shown in Fig. 4*G*, proteolysis of membrane fractions in the presence or absence of detergent results in the conversion of cytochrome *b*5 to a species with faster mobility upon SDS-PAGE. The latter species lacks the C-terminal 40 residue membrane anchor, but retains its catalytic activity (7). Cleavage of cytochrome *b*5 was observed in intact microsomes of cells containing Construct e at protease concentration where no cleavage of the expressed protein occurred (Fig. 4*G*).

Mutagenesis of the Lumenal Targeting Sequence—Sequences of 11β-HSD from different mammalian species contain two Lys residues in region I. The region I of human E3 contains an Arg-Lys sequence. In the rabbit E3 sequence, the Arg residue is replaced by Val, implying that in region I a single lysyl residue rather than two basic residues do not alter the orientation of E3 in the membrane. Replacement of the single Lys residue by Ile as in rab E3N(K4/I)-GFP (Construct f) led to ER localization of the resulting chimera as evidenced by immunofluorescence (Fig. 5*A*). The fluorescence image of cells expressing Construct f is similar to that seen for rab E3N-GFP (Construct e). The membrane localization of Construct f was confirmed by subcellular fractionation and immunoblot analysis. As seen in Fig.

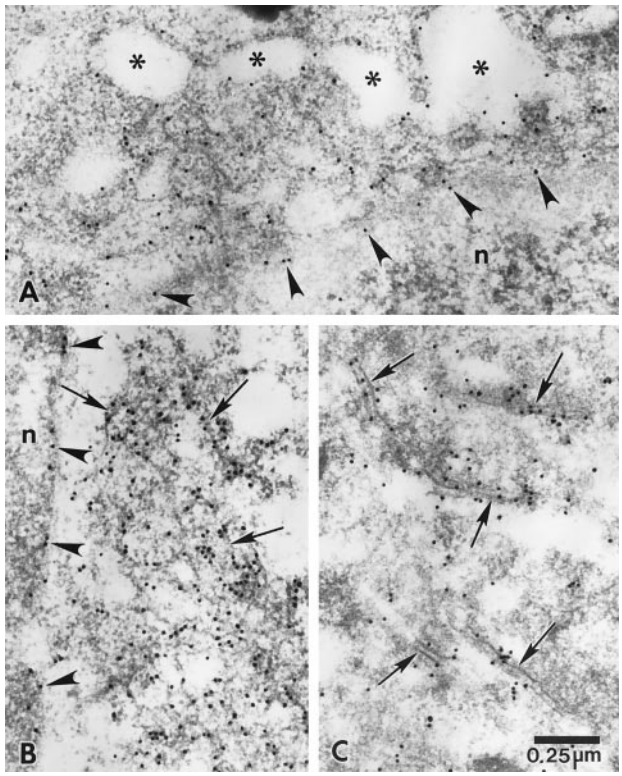

FIG. 3. **Electron microscopic immunogold labeling of rab E3 N-GFP transfected cells with anti-GFP antibodies.** *A* and *B*, illustrate clusters of ER in the perinuclear region. In *A*, the *asterisks* (*) indicate dilated regions of ER with membrane-associated gold particles. In *B,* the *arrows* indicate regions where labeling of the ER is particularly prominent. In both panels, the *arrowheads* indicate gold particles associated with the nuclear envelope. Nucleus is denoted by *n*. In *C*, several smooth-surfaced membranes are labeled (*arrows*). *Scale bar* = 0.25 μm.

5*B*, the expressed protein was present in pellet and absent in the soluble fraction. To establish the orientation of the expressed protein, cells were selectively permeabilized with saponin, digitonin, SLO, and analyzed by immunofluorescence microscopy (Figs 5, *C–E*). Again, the staining pattern of such permeabilized cells was identical to that obtained with Construct e (Fig. 4, *D–F*). This result indicates that positively charged residues are not an essential topogenic signal for the lumenal orientation.

When the single Gly residue in region II was replaced by Glu, and the resulting chimera (Construct b) expressed in cells, such mutation led to the expression of the protein with an apparent cytosolic and nuclear localization (Fig. 6*A*). Deletion of six residues in region II as in the Construct d also led to a cytosolic and nuclear localization of the protein (Fig. 6*B*).

Reorientation of NADH Cytochrome b5 Reductase Conferred by the LTS—To test the utility of LTS as an *in vivo* trafficking signal for proteins other than GFP, we made a fusion protein between the polar segment of microsomal NADH cytochrome *b*5 reductase and the N terminus of rab E3. The catalytically active polar segment of the reductase consists of some 270 residues, and lacks the hydrophobic 28-residue *N*-myristoylated membrane-anchoring domain at the N terminus present in the intact protein (8). The native reductase has a cytosolic orientation. To assess whether the LTS can redirect the reductase derivative to the lumen of ER, we fused the LTS of rabbit E3 to the N terminus of the polar segment of the reductase. In addition, we added to the C terminus of the reductase the highly antigenic C-terminal sequence of stearyl CoA desaturase as an epitope tag. We termed the chimera rab E3-N-b5 red (31–300) or Construct g. The orientation of Construct g in the ER membrane was examined by proteinase K digestion. As shown in Fig. 7, digestion of intact microsomes (*lane 1–3* and *5*) revealed a protease-resistant fragment approximately of 35–40 kDa in mass, indicating that the construct was intralumenal. When microsomal membranes were disrupted with 0.5% Triton X-100, the protein became sensitive to proteinase K digestion and was completely degraded (*lane 4*). The proteinase K protection assay results confirm our previous results of the GFP chimera orientation and show that soluble as well as membranous proteins with a cytoplasmic orientation wearing 11β or E3 LTS are oriented toward the lumenal side of ER.

DISCUSSION

11β-HSD and E3 are two unrelated ER lumenal proteins with type II orientation (17, 19). Analysis of their covalent structures implied that the LTS for 11β-HSD and E3 is encoded in the N-terminal segment. The N-terminal amino acid sequence of 34 residues of this group of proteins from several mammalian species is shown in Fig. 1*A*. For discussion purposes, the LTS sequence is divided into regions I, II, and III. In addition to the positive charge at the N-terminal residue, features of region I that are shared by several species include a single lysyl residue. A stretch of aliphatic and aromatic hydrophobic residues (region II) follows region I, and continues into a segment of negatively charged residues (region III). A striking feature of region I is the variability of the first three residues and the conservation of a single lysyl residue. Region II is a hydrophobic segment containing a cluster of tyrosyl residues. In region III, there is a conserved Asn-*X*-Glu-Glu segment. The remaining sequence of the C-terminal ectodomain shared no obvious homology between the dehydrogenases and the esterases. Despite that the bulk of the polypeptide chain is translocated across the membrane, the N-terminal leader peptide like sequence in this group of proteins escapes signal peptidase cleavage and serves as the membrane anchor.

Fig. 1*B* displays the mutations and deletions of the chimeras constructed by fusion of N-terminal domain of 11β-HSD or E3 to the N terminus of GFP. These constructs were expressed in COS 7 cells, and analysis of their cellular targeting was accomplished by combination of fluorescence microscopy and electron microscopic immunogold labeling. The topology of expressed proteins was established by immunocytofluorescent staining of selectivity permeabilized COS cells. Protease protection assay of membranes in the presence and absence of detergents was also used to distinguish the membrane sidedness of the GFP chimeras.

The subcellular localization and orientation of the constructed GFP chimeras is summarized in Fig. 8 (*A* and *B*). The most important finding of this study is that fusion of the N-terminal sequences from both 11β-HSD and E3 proteins to GFP resulted in the lumenal localization of the chimera in the ER membrane. This finding implied that the information (LTS) for the lumenal orientation of 11β-HSD and E3 is encoded in the N-terminal segment of some 20 to 25 residues.

As displayed in the amino acid sequences of the native proteins from different species, the presence of two neutral or two basic residues in region I of LTS are not essential for the lumenal targeting. This is based on the assumption that the proteins from different animal species share an identical orientation. The proposed essential di-arginine sequences (30) at the cytoplasmic N terminus of type II membrane proteins are not indispensable for the ER localization for the group of proteins described here. Deletion of the single lysyl residue in region I (Construct f) led to ER targeting of the GFP chimera with a lumenal ER orientation (Fig. 5). Therefore, the absence of basic residues in region I does not alter the lumenal target-

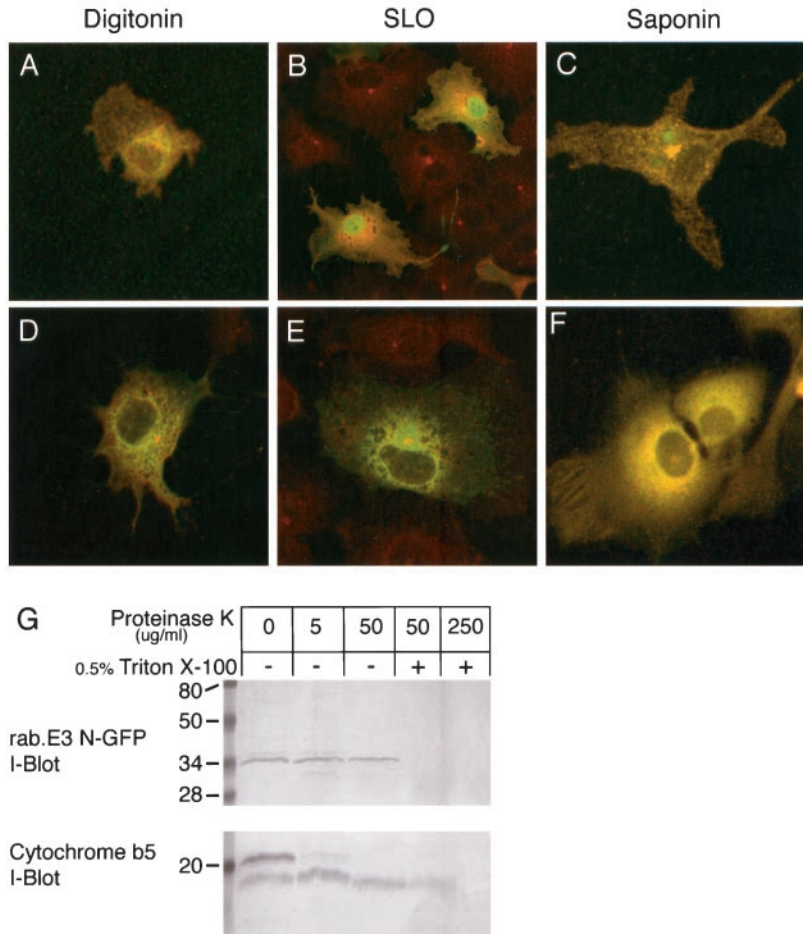

FIG. 4. **Immunocytofluorescent staining of selectively permeabilized COS-7 cells and the protease sensitivity of the expressed chimeras.** COS-7 cells transfected with cDNA coding for GFP (*A–C*), and rab E3 N-GFP (*D–F*) were permeabilized either with digitonin (20 μg/ml), streptolysin O (200 units/ml), or saponin (0.2%) before incubation with anti-GFP mouse monoclonal antibody. Bound antibody was visualized by incubation with biotin-coupled goat anti-mouse and then with Cy3-streptavidin as described under "Experimental Procedures" (magnification, ×40). *G*, protease sensitivity of rab E3 N-GFP and cytochrome *b*5 in membranes. The microsomal membranes of COS-7 cells transfected with cDNA coding for rab E3 N-GFP fusion protein and enriched in cytochrome *b*5 were incubated with the indicated concentration of proteinase K for 30 min on ice in the absence (*lanes 1–3*) or presence (*lane 4* and *5*) of 0.5% Triton X-100. The mobility of rab E3 N-GFP and cytochrome *b*5 was analyzed by SDS-PAGE and immunoblot with 5 μg/ml of an anti-GFP mouse monoclonal antibody or with 1:1,000 dilution of an anti-cytochrome b5 chicken polyclonal antibody.

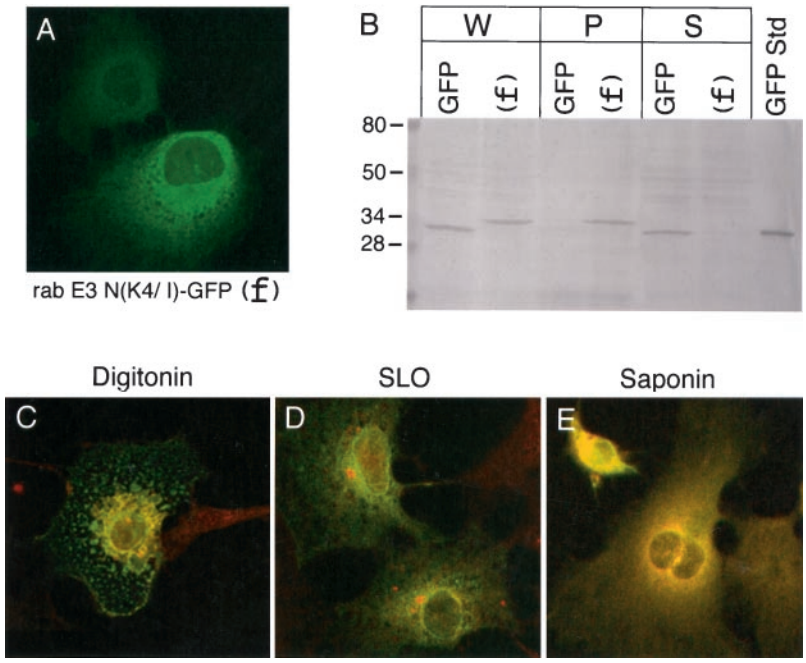

FIG. 5. **Expression of rab E3 N(K4/I)-GFP in COS-7 cells to analyze the importance of positively charged residue of region I in the targeting to the ER.** *A*, subcellular localization of rab E3 N(K4/I)-GFP (Construct f) determined by fluorescence microscopy (magnification, ×40). The *(K4/I)* symbol indicates the position of positively charged residue within E3 region I in which the single lysine was mutated to an isoleucine residue. *B*, immunoblot analysis of COS-7 cells expressing Construct f and fractionated as described in Fig. 2. The fluorescence image and the immunoblot show the same intracellular localization as the constructs E3 N-GFP in Fig. 2 (*C* and *D*). Selective permeabilization with digitonin (*C*), SLO (*D*), and saponin (*E*) of cells expressing Construct f was done as described in Fig. 4.

ing of the protein. Clearly, the positively charged residues are not required for the ER localization, but the extreme negative charge of region III would conform to the N > C charge role for the lumenal orientation.

In order to further characterize the essential residues of the LTS, several chimeras of LTS fused to GFP were constructed and their subcellular localization analyzed. The Construct b bearing a single negative charge in the center of region II led to the expression of protein with fluorescence pattern similar to the control GFP (Fig. 6*A*). Immunoblot of the subcellular fractions of cells expressing Construct b indicated its electrophoretic mobility upon SDS-PAGE identical to native GFP, implying that a truncation of the N terminus of the chimera had occurred (Fig. 6*C*), but the protein failed to enter the

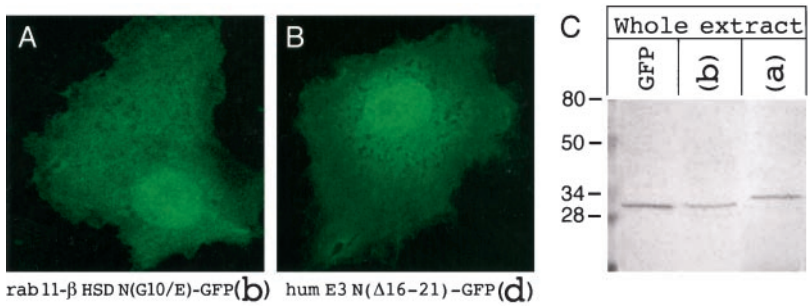

FIG. 6. **Effect of mutations in the hydrophobic region of 11β-HSD and E3 LTS on the localization of GFP chimera proteins.** *A*, fluorescence microscopy of COS-7 cells transfected with rab 11β-HSD N(G10/E)-GFP (Construct b) in which the glycine residue in position 10 was mutated to glutamic acid residue. *B*, fluorescence image of COS-7 cells transfected with hum E3 N(Δ16–21)-GFP (Construct d) in which the hydrophobic segment was shortened by deletion of the six residues AYYIYT. A strong labeling pattern of nucleus, ER, and cytoplasm is seen for the mutant chimera proteins as compared with the ER labeling pattern seen for the wild chimera proteins in Fig. 2 (*C* and *D*) (magnification, ×40). *C*, immunoblot analysis of cells expressing GFP, Construct b, and Construct a.

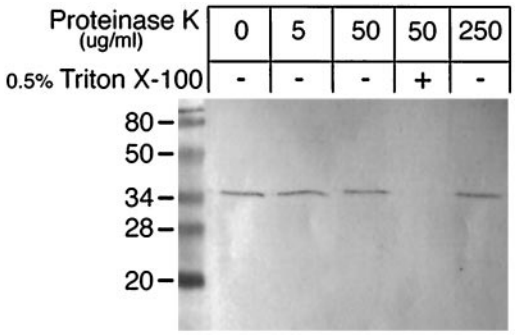

FIG. 7. **Evidence for the lumenal orientation of rab E3 N-b5 red.(31–300) in microsomal membranes of transfected cells.** Microsomes containing the chimeric protein (Construct g) were incubated with the indicated concentrations of proteinase K in the presence (*lane 4*) or absence (*lanes 1–3* and *5*) of 0.5% Triton X-100. Samples of the proteolysis reaction mixture were subjected to SDS-PAGE analysis, followed by immunoblot with 1:2,000 dilution of anti-FLAG polyclonal antibody.

A

	Localization				Orientation in the ER	
	N	ER	C	S	L	Cyt
(a) rab 11β-HSD N-GFP	–	+	–	–	+	–
(b) rab 11β-HSD N(G10/E)-GFP	+	–	+	–		
(c) hum E3 N-GFP	–	+	–	–	+	–
(d) hum E3 N(Δ16-21)-GFP	+	–	+	–		
(e) rab E3 N-GFP	–	+	–	–	+	–
(f) rab E3 N(K4/I)-GFP	–	+	–	–	+	–
(g) rab E3 N-b5 red. (31-300)	–	+	–	–	+	–

B

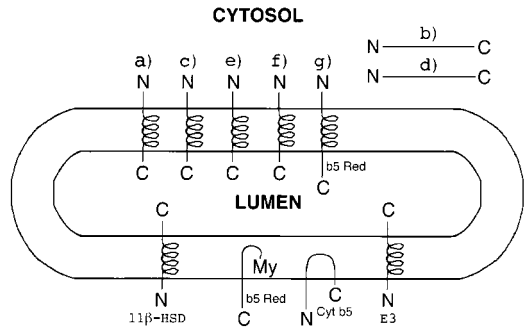

FIG. 8. **Targeting proteins to the lumen of ER with LTS.** *A*, summary of subcellular localization of all the constructs expressed. The constructs are abbreviated *a-g* as shown in Fig. 1*B*. The abbreviations for the subcellular localizations are *N, ER, C, S, L,* and *Cyt*, nucleus, endoplasmic reticulum, cytosolic, secreted, lumen, and cytoplasmic side, respectively, proposed membrane topology of the LTS-GFP constructs. *N* and *C* represent the N and C termini of the native and chimera proteins. The chimeras are displayed in the upper part, and the native proteins in the lower part of the model. *My* at the N terminus of native cytochrome *b*5 reductase denotes a myristoyl residue.

secretory pathway as anticipated. This implied that the signal peptidase was not responsible for the cleavage of the N-terminal signal. What may have compensated for the translocation deficiency of Construct b? It is most likely that the introduction of a glutamyl residue in region II of the LTS resulted in a conformational change interpreted by the cytosolic quality control system as a misfolded protein. Proteolytic elimination of malfolded proteins, including uncomplexed subunits of protein assemblies, is a common defense mechanism of the cell. Mutant forms of cystic fibrosis transmembrane conductance regulator expressed in mammalian cells, or carboxypeptidase Y mutants that fail to translocate across the membrane and accumulate on the surface of ER are readily degraded by the cytoplasmic proteosome pathway (31, 32). One reason we observed the GFP tag in cells expressed with Construct b could be the resistance of the GFP molecule toward proteolysis. Expression of the N terminus of Construct b fused to a polypeptide that is more sensitive to proteolysis than the GFP molecule most likely would result in the degradation of the entire construct.

A striking feature of region II of the parent proteins is the presence of a cluster of tyrosyl residues (AYYXY or GYYY). Repeats of tyrosyl residues (AYPYYA) are found in the transmembrane domain of the 48-kDa subunit of microsomal OT. OT has a short C terminus oriented toward the cytosol, and a single transmembrane domain with the bulk of the protein molecule positioned in the lumen of ER (18). The T-cell receptor delta chain displays a GYYYYV sequence in its membranous segment (33). That a tyrosine-containing motif mediates ER retention of CD3-ϵ chain has also been reported (34). To determine the topogenic importance of the tyrosyl residue cluster GFP chimera, Construct d lacking the tyrosyl residues in region II was generated. Construct d upon its expression in COS cells displayed fluorescence in the nucleus and cytosol (Fig. 6B). This fluorescence pattern of intracellular structures was similar to that observed by the untagged GFP. These findings suggest that a cluster of tyrosyl residues is essential for the correct folding and the lumenal targeting. Further analysis will be needed to establish whether replacement of these residues with phenylalanines or glycines provides an appropriate context for the function of the conserved tyrosyl residues. Deletion of five to six residues in region III of the 11β-HSD-GFP constructs does not affect the lumenal orientation of the constructs (Fig. 2*A*).

Presently, the 11β-HSD and the E3 protein are the only type II, static lumenal ER proteins with a single N-terminal transmembrane segment that have been characterized. The asialo-

glycoprotein receptor (35) and the influenza virus neuramidinase (36) are type II plasma membrane proteins. The glycoprotein, paramyxovirus hemagglutinin-neuroaminidase (HN) is a type II membrane protein localized to the Golgi cisternae (37). The predicted amino acid sequence of the simian virus HN protein includes a 17-residue cytoplasmic tail, a 19-residue membranous segment, and a large 523-residue C-terminal ectodomain (38). Newly synthesized HN oligomerizes into tetramers before transport from ER to Golgi, and alterations of the C-terminal ectodomain can prevent ER to Golgi transport (39). The large family of glycosidases and glycosyltransferases, although type II transmembrane proteins are all located in the Golgi cisternae. Their common features include an N-terminal cytoplasmic tail, a transmembrane domain, and a large catalytic domain oriented toward the lumenal side of the membrane (40). Two mutually complementary models have been proposed to explain the mechanism of Golgi retention of the glycosyltransferases mediated by their transmembrane domains. One model postulates the retention through oligomerization, which prevents proteins from entering transport vesicles (41, 42). The other model suggests that their retention depends on the length of a membrane-spanning domain, which can be accommodated by the specific lipid composition of Golgi complex membrane (43, 44). It has been pointed out that neither the oligomerization nor the membrane lipid composition alone can explain the sorting of Golgi proteins.

Having established that the N-terminal domain of 11β-HSD or E3 can mediate lumenal localization of the following downstream GFP molecule, we sought to test the general applicability of this finding to targeting of ER proteins that display a cytosolic orientation in their native state. The flavoprotein NADH cytochrome *b*5 reductase has an amphiphilic structure in which the hydrophilic, catalytic domain of some 270 residues is linked to a membrane-anchoring, hydrophobic domain that serves to orient the catalytic site of the reductase at the membrane-aqueous interface to permit a rapid electron transfer to cytochrome *b*5 (8). The hydrophobic domain at the N terminus of the reductase consists of some 28 residues as well as *N*-myristoylation of the N-terminal glycine residue (45) (Fig. 8*B*). Fusion of the LTS of rabbit E3 to the N terminus of the polar segment of the reductase (Construct g) resulted in the lumenal targeting of the fusion protein (Figs. 7 and 8*B*).

Whether the topogenic signal described here is position-independent as regards to lumenal targeting of proteins, in addition to identifying the structural features that prevent these constructs transiting to the Golgi, are the topics for further studies. In conclusion, a static lumenal targeting signal of proteins reported here should add to the repertoire of techniques for studies on the structure, organization, and processes of the lumen of ER.

Acknowledgment—We thank Dr. C. S. Ramaro (University of Alabama at Birmingham) for constructing and expressing Constructs a and b.

REFERENCES

1. Munro, S., and Pelham, H. R. B. (1987) *Cell* **48,** 899–907
2. Scheel, A. A., and Pelham, H. R. (1996) *Biochemistry* **35,** 10203–10209
3. Blobel, G. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 1496–1500
4. Schatz, G., and Dobberstein, B. (1996) *Science* **271,** 1519–1526
5. Hegde, R. S., Voigt, S., Rapoport, T. A., and Lingappa, V. R. (1998) *Cell* **92,** 621–631
6. Gafvelin, G., Sakaguchi, M., Andersson, H., and von Heijne, G. (1997) *J. Biol. Chem.* **272,** 6119–6127
7. Ozols, J., and Gerard, C. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 3725–3729
8. Ozols, J., Korza, G., Heinemann, F. S., Hediger, M. A., and Strittmatter, P. (1985) *J. Biol. Chem.* **260,** 11953–11961
9. Ozols, J., Heinemann, F. S., and Johnson, E. F. (1985) *J. Biol. Chem.* **260,** 5427–5434
10. Ozols, J. (1986) *J. Biol. Chem.* **261,** 3965–3979
11. Heinemann, F. S., and Ozols, J. (1982) *J. Biol. Chem.* **257,** 14988–14999
12. Heinemann, F. S., and Ozols, J. (1984) *J. Biol. Chem.* **259,** 797–804
13. Thiede, M. A., Ozols, J., and Strittmatter, P. (1986) *J. Biol. Chem.* **261,** 13230–13235
14. Ozols, J. (1994) *Biochemistry* **33,** 3751–3757
15. Ozols, J. (1989) *J. Biol. Chem.* **264,** 12533–12545
16. Ozols, J. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 5302–5306
17. Ozols, J. (1995) *J. Biol. Chem.* **270,** 2305–2312
18. Kumar, V., Heinemann, F. S., and Ozols, J. (1994) *J. Biol. Chem.* **269,** 13451–13457
19. Ozols, J. (1998) *Biochemistry* **37,** 10336–10344
20. Heinemann, F. S., and Ozols, J. (1998) *Mol. Biol. Cell* **9,** 3445–3453
21. Strittmatter, P., Kittler, J. M., Coghill, J. E., and Ozols, J. (1992) *J. Biol. Chem.* **267,** 2519–2523
22. Mori, M., Morita, T., Ikeda, F., Amaya, Y., Tatibana, M., and Cohen, P. P. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 6056–6060
23. Haynes, R. L., Zheng, T., and Nicchitta, C. V. (1997) *J. Biol. Chem.* **272,** 17126–17133
24. Ozols, J. (1989) *Biochim. Biophys. Acta* **997,** 121–130
25. Otto, J. C., and Smith, W. L. (1994) *J. Biol. Chem.* **269,** 19868–19875
26. Hand, A. R. (1995) in *Introduction to Biophysical Methods for Protein and Nucleic Acid Research* (Glasel, J. A., and Deutscher, M., eds) pp. 205–260, Academic Press, New York
27. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994) *Science* **263,** 802–805
28. Yano, M., Kanazawa, M., Terada, K., Namchai, C., Yamaizumi, M., Hanson, B., Hoogenraad, N., and Mori, M. (1997) *J. Biol. Chem.* **272,** 8459–8465
29. Schulz, I. (1990) *Methods Enzymol.* **192,** 280–300
30. Schutze, M. P., Peterson, P. A., and Jackson, M. R. (1994) *EMBO J.* **13,** 1696–1705
31. Kopito, R. R. (1997) *Cell* **88,** 427–430
32. Hiller, M. M., Finger, A., Schweiger, M., and Wolf, D. H. (1996) *Science* **273,** 1725–1728
33. Hein, W. R., and Dudler, L. (1993) *EMBO J.* **12,** 715–724
34. Mallabiabarrena, A., Jimenez, M. A., Rico, M., and Alarcon, B. (1995) *EMBO J.* **14,** 2257–2268
35. Spiess, M., and Lodish, H. F. (1986) *Cell* **44,** 177–185
36. Bos, T. J., Davis, A. R., and Nayak, D. P. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81,** 2327–2331
37. Parks, G. D. (1996) *J. Biol. Chem.* **271,** 7187–7195
38. Hiebert, S. W., Paterson, R. G., and Lamb, R. A. (1985) *J. Virol.* **54,** 1–6
39. Ng, D. T., Hiebert, S. W., and Lamb, R. A. (1990) *Mol. Cell. Biol.* **10,** 1989–2001
40. Paulson, J. C., and Colley, K. J. (1989) *J. Biol. Chem.* **264,** 17615–17628
41. Colley, K. J. (1997) *Glycobiology* **7,** 1–13
42. Dahdal, R. Y., and Colley, K. J. (1993) *J. Biol. Chem.* **268,** 26310–26319
43. Pelham, H. R. B., and Munro, S. (1993) *Cell* **75,** 603–605
44. Munro, S. (1995) *Biochem. Soc. Trans.* **23,** 527–530
45. Ozols, J., Carr, S. A., and Strittmatter, P. (1984) *J. Biol. Chem.* **259,** 13349–13354

CELL BIOLOGY AND METABOLISM:

Targeting Proteins to the Lumen of Endoplasmic Reticulum Using N-terminal Domains of 11β-Hydroxysteroid Dehydrogenase and the 50-kDa Esterase

Hassan Mziaut, George Korza, Arthur R. Hand, Craig Gerard and Juris Ozols

J. Biol. Chem. 1999, 274:14122-14129.
doi: 10.1074/jbc.274.20.14122