Characterization of the TGN exit signal of the human mannose 6-phosphate uncovering enzyme

Prashant Nair[1,*], Beat E. Schaub[1,*], Kai Huang[2], Xiang Chen[2], Robert F. Murphy[2,3], Janice M. Griffith[4], Hans J. Geuze[4] and Jack Rohrer[1,‡]

[1]Institute of Physiology, University of Zurich, Zurich, 8057, Switzerland
[2]Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[3]Department of Biomedical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[4]Department of Cell Biology and Institute of Biomembranes, Utrecht University School of Medicine, 3584 CX Utrecht, The Netherlands
*These authors contributed equally to this work
‡Author for correspondence (e-mail: rohrer@access.unizh.ch)

Accepted 12 April 2005
Journal of Cell Science 118, 2949-2956 Published by The Company of Biologists 2005
doi:10.1242/jcs.02434

Summary

The human mannose 6-phosphate uncovering enzyme participates in the uncovering of the mannose 6-phosphate recognition tag on lysosomal enzymes, a process that facilitates recognition of those enzymes by mannose 6-phosphate receptors to ensure delivery to lysosomes. Uncovering enzyme has been identified on the trans-Golgi network at steady state. It has been shown to traffic to the plasma membrane from where it is rapidly internalized *via* endosomal structures, the process being mediated by a tyrosine-based internalization motif, Y^{488}HPL, in its cytoplasmic tail. Using immunogold electron microscopy a GFP-uncovering enzyme fusion construct was found to be colocalized with the cation-dependent mannose 6-phosohate receptor in regions of the trans-Golgi network, suggesting that uncovering enzyme might follow a similar pathway of exit from the trans-Golgi network as that of the cation-dependent mannose 6-phosohate receptor. In this study, we identified the signal sequence in the cytoplasmic tail of uncovering enzyme responsible for its exit from the trans-Golgi network. Using GFP fusion constructs of the transmembrane and cytoplasmic domains of uncovering enzyme, we could show, by automated analysis of confocal immunofluorescence images, that residues Q^{492}EMN in the cytoplasmic tail of uncovering enzyme are involved in its exit from the trans-Golgi network. Detailed characterization of the exit signal revealed that residue Q^{492} is the most important to the exit function while M^{494} and N^{495} also contribute. The cytoplasmic tail of the uncovering enzyme does not possess any of the known canonical signal sequences for interaction with Golgi-associated gamma ear-containing adaptor proteins. The identification of a trans-Golgi network exit signal in its cytoplasmic tail elucidates the trafficking pathway of uncovering enzyme, a crucial player in the process of lysosomal biogenesis.

Keywords: Lysosomal biogenesis, Uncovering enzyme, trans-Golgi network, TGN exit, Subcellular location features, Protein distribution comparison

Introduction

Lysosomes are one of the major degradative compartments of mammalian cells and their biogenesis depends on the concerted transport of acid hydrolases from their site of origin to their site of action through a series of intermediate compartments. Lysosomal enzymes are recognized by one of the two known mannose 6-phosphate receptors (MPRs) in the trans-Golgi network (TGN) for transport onward to lysosomes. This crucial recognition event directly depends on the acquisition of mannose 6-phosphate (M6P) tags by the lysosomal enzymes. The tag is generated in a two-step enzymatic reaction which occurs in the Golgi. The first step, the addition of an *N*-acetyl glucosamine phosphate to selected mannose residues of the lysosomal enzyme, is catalyzed by the UDP-*N*-acetylglucosamine lysosomal enzyme, *N*-acetylglucosamine phosphotransferase. The second step, mediated by *N*-acetylglucosamine-1-phosphodiester α-*N*-acetylglucosaminidase, leads to the actual uncovering of the M6P tag by cleaving the GlcNAc residue. The α-*N*-acetylglucosaminidase is therefore termed uncovering enzyme (UCE).

UCE is a type I membrane glycoprotein of 515 amino acids. It has a single transmembrane domain of 27 residues and a 41 residue cytoplasmic tail, which contains several established and potential signal sequences ensuring its correct intracellular sorting (Kornfeld et al., 1999; Lee et al., 2002). UCE has been shown to reside in the TGN at steady state and its constitutive recycling between that organelle and the plasma membrane has been established using a combination of steady state measurements of enzyme activity and localization of green fluorescent protein (GFP) fusion constructs of the enzyme cytoplasmic tail (Rohrer and Kornfeld, 2001). It was shown that the classical endocytosis motif Y^{488}HPL and the C-terminal N^{511}PFKD^{515} motif are involved in the trafficking of UCE (Rohrer and Kornfeld, 2001). However, the dissection of the trafficking itinerary of UCE to elucidate the individual transport steps involved has not yet been reported. Three different types of signals in the cytoplasmic tail of UCE have been described. First, the transmembrane domain and the first 11 residues of the cytoplasmic tail of UCE were shown to be involved in TGN retention of the enzyme, thereby indicating

that the TGN exit signal is excluded from this region of the enzyme. Secondly, the Y^{488} residue has been shown to be functional in internalization from the surface (Lee et al., 2002). Third, the Y^{486} residue was implicated in the return of internalized UCE to the TGN (Lee et al., 2002). The signal(s) leading to TGN exit have been studied in this work.

Signals involved in localization of membrane proteins at specific intracellular locations may be of the retention or the recycling type. A number of transmembrane proteins and receptors are known to cycle between the TGN and the plasma membrane, of which, TGN38, furin, the envelope glycoprotein gp1 of the Varicella Zoster virus and the human Menkes disease protein are notable. The SDYQRL sequence in the cytoplasmic tail of TGN38 has been shown to be involved in its TGN localization (Wong and Hong, 1993). An acidic TGN retention signal and a YKGL TGN retrieval motif have been identified in furin (Schafer et al., 1995). A tyrosine-based motif and a casein kinase II phosphorylation site were shown to be responsible for the TGN localization of the Varicella Zoster envelope glycoprotein gp1 (Alconada et al., 1996). A C-terminal di-leucine motif has been shown to be involved in the correct localization of the human Menkes disease protein in the TGN (Petris et al., 1998). A FHRL sequence in the cytoplasmic tail of carboxypeptidase D has been shown to be important for the TGN retention and retrieval of the protein from the surface (Eng et al., 1999). Although several instances of signals involved in TGN localization exist in the literature, there have been few reports of signal sequences involved in the exit from the TGN of proteins cycling between the TGN and the plasma membrane. Anterograde transport of vesicles from the TGN has been established to be a process mediated by clathrin-coated vesicles and Golgi-associated adaptor proteins. Cargo binding to such adaptors for selective inclusion in transport vesicles depends on crucial signal sequences in its cytoplasmic tail. For instance, it has been shown that the cytoplasmic signals $YKGL^{765}$, LI^{760}, F^{790} and clustered acidic amino acids in furin are crucial for interaction with adaptor protein-1 (AP1) (Teuchert et al., 1999).

Using immunoelectron microscopy we previously found that a GFP-UCE fusion construct was localized to the TGN in close proximity to the cation-dependent MPR (CD-MPR), suggesting that the two proteins might have a similar pathway of exit from the TGN. However, unlike the CD-MPR, which has been shown to interact with AP-1 and Golgi-localized, gamma-ear-containing, adenosine diphosphate ribosylation factor-binding protein (GGAs), UCE does not contain any of the known motifs for interaction with GGAs. The localization of the GFP fusion construct of the transmembrane and cytosolic domains of the human UCE has been previously shown to be identical to that of the endogenous enzyme in HeLa cells (Rohrer and Kornfeld, 2001). Furthermore, it was shown that replacement of the Y^{488} residue with an alanine led to a cell surface accumulation of the GFP fusion constructs thus demonstrating the role of the residue in internalization of the enzyme (Rohrer and Kornfeld, 2001). Additionally, it was reported that a construct truncated at position Y^{486} was mostly trapped in an intracellular compartment. This result was surprising in view of the finding that only 16% of the molecules were found on the cell surface despite the lack of an internalization signal (Lee et al., 2002). This indicated that the construct lacks a TGN exit signal causing it to leave the TGN very slowly.

In this study, we used a series of GFP fusion constructs, confocal immunofluorescence microscopy and automated image interpretation to identify the signal sequence in the cytoplasmic tail of UCE involved in its exit from the TGN. The identification of the TGN exit signal in the cytoplasmic tail of UCE is an important step in elucidating its trafficking itinerary. It is as yet unknown if UCE is packaged into the same or different vesicles from that of the MPR at the TGN. Given that UCE is crucial to the process of lysosomal biogenesis, the potential interaction of the TGN exit signal with specific Golgi-associated adaptor proteins might explain its traffic out of the TGN.

Materials and Methods

Materials

Enzymes used in molecular cloning were obtained from Boehringer-Mannheim, New England Biolabs, Sigma or Promega; α-minimal essential medium (α-MEM) and fetal calf serum were from Gibco-BRL (Basel, Switzerland). Polyethylenimine (PEI; linear, M_r 25,000) was from Polysciences Inc. (Warrington, PA, USA) and cell culture dishes from Falcon. DAPI (4′,6-diamidino-2-phenylindole) was from Boehringer-Mannheim. Lissamine rhodamine sulfonyl chloride (LRSC) was from Molecular Probes (Eugene, OR, USA). Oligonucleotides were synthesized by Microsynth GmBH (Galbach, Switzerland).

Recombinant DNA

All basic DNA procedures were as described previously (Sambrook et al., 1998). The GFP-UCE transmembrane domain-cytoplasmic tail construct was prepared using a PCR procedure (Ho et al., 1989) as described (Rohrer and Kornfeld, 2001). The GFP-UCE 502 stop construct, the double mutant, the single and combined mutant constructs were all made using the pBill-neo GFP-UCE Y488A construct as template for a two-step PCR using appropriate partial complementary pairs of oligonucleotides in which the desired amino acid replacements had been incorporated. The PCR products were put together with fragments from the vector in a triple ligation following digestion with *Bgl*II, *Mlu*I and *Eco*RI. The final PCR products were subcloned into pSFFV-neo as described (Rohrer et al., 1995). All coding sequences created by PCR were verified by sequencing.

Cell culture and transfection

HeLa cells were grown in α-MEM to 30% confluency before transfection with 2.2 μg of DNA using 7.6 μg PEI according to the protocol from Polyplus-transfection SAS, Illkirch, France. Selection for resistance to neomycin (G418) was carried out using 1 mg/ml G418 as the final concentration. Resistant colonies were picked individually and screened for the expression of GFP-UCE by immunofluorescence analysis. Two or three different clones expressing similar or varying amounts of each of the fusion constructs were expanded for further study and maintained in selective medium.

Immunoelectron microscopy

For immunoelectron microscopy, HeLa cells expressing GFP-UCE were fixed with 4% paraformaldehyde in 0.1 M sodium phosphate buffer at pH 7.4, washed in buffer, and embedded in 10% gelatin. The gelatin blocks were infused with 2.3 M sucrose and frozen in liquid nitrogen. Cryosectioning and immunogold labeling have been described before (Geuze et al., 1981; Raposo et al., 1997). Double-

immunogold labeling of GFP and CD-MPR was done with 15 nm and 10 nm protein A-gold particles, respectively.

High resolution confocal immunofluorescence microscopy

Confocal immunofluorescence microscopy was carried out using a variation on a three-color scheme previously developed to facilitate automated image interpretation (Velliste and Murphy, 2002) as follows. Three percent paraformaldehyde-fixed and saponin-permeabilized (final concentration of 1 mg/ml) cells were incubated with 1 μg/ml DAPI (to label total DNA) and a 1:40 dilution of a 1 mg/ml stock solution of LRSC (to label total protein) for 10 minutes before being washed four times with PBS-glycine-saponin. The coverslips were mounted on glass slides with ProLong Antifade (Molecular Probes) for viewing with a Leica TCS SP2 (AOBS) confocal laser-scanning microscope. Serial sections in the z-axis through entire cells (23-110 sections per image stack) were taken with a step size of 0.1628 μm and a pixel size of 0.0977 μm in the x and the y dimensions (1024×1024 pixels per section). At least 80 cells were analyzed for every clone of the mutant constructs and two or three clones were analyzed for each construct. The resulting stacks of images were exported as Tiff files and visually analyzed with the Imaris program (Bitplane AG, Zurich, Switzerland).

Automated image analysis

Automated segmentation of the GFP images using the DNA (DAPI) and total protein (LRSC) channels were carried out using a seeded watershed approach (Velliste and Murphy, 2002). The position of each nucleus was identified by downsampling and thresholding the DNA image, finding each connected object, and eliminating objects whose integrated DNA fluorescence was outside the range observed for normal nuclei. The centers of the nuclei along with the total protein image were supplied to the mmcwatershed function from the SDC Morphology Toolbox for MATLAB (SDC Information Systems, Naperville, IL, USA). This function implements the seeded watershed algorithm described previously (Lotufo and Falcao, 2000); it returns an image divided into regions corresponding to each of the seeds (nuclei). These regions were used as masks to isolate each cell in the GFP and DNA images (partial cells touching the edge of the image were eliminated).

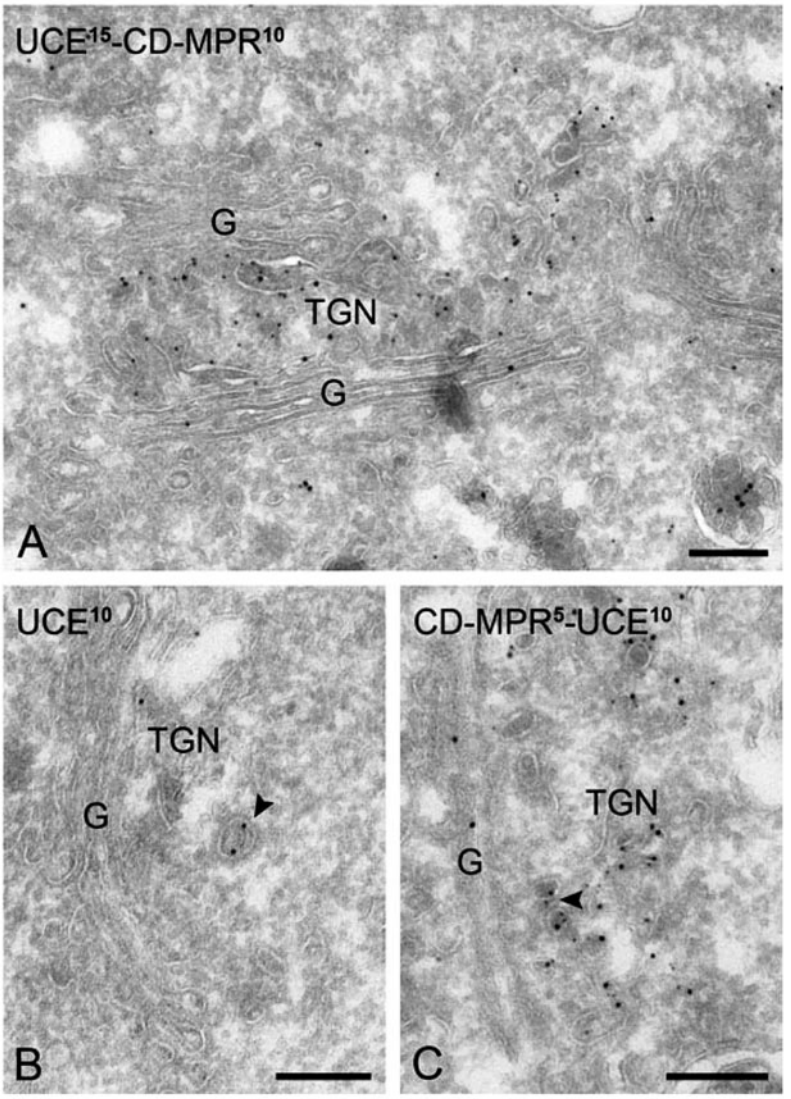

Fig. 1. Single and double immunogold labeling of UCE (B) and of UCE and CD-MPR (A,C) with gold particle sizes (in nm) as indicated in superscript on the figures. A and C show that most of the UCE and CD-MPR localize to the TGN with only a few gold particles present over cisternae of the Golgi stacks (G). In B a clathrin-coated vesicle is labeled for UCE (arrowhead) in a low-expressing cell, while in C a few coated vesicles are labeled for UCE and CD-MPR together (arrowhead). Bars, 100 nm.

The GFP and DNA image for each cell were then used to calculate subcellular location features (SLFs), numerical features describing the subcellular pattern. The features describe various aspects of the pattern, such as the size of objects and the 'evenness' of the pattern (for a review, see Huang and Murphy, 2004). A new feature set, SLF19, was created by combining set SLF11 (Chen and Murphy, 2004) with features SLF9.9-9.14 and SLF9.21-9.28 that capture the relationship of the GFP fluorescence to the parallel DNA image (Velliste and Murphy, 2002). This feature set was calculated for all images of the various UCE constructs, and a set of 52 features (defined as SLF20) was selected from it by stepwise discriminant analysis (SDA) in order to choose only those features that are useful for distinguishing the constructs (Huang and Murphy, 2004). The average feature values were calculated for each construct, and the Mahalanobis distances between each pair of constructs were calculated and used to construct a dendrogram as described previously (Chen et al., 2003). Representative images for each construct were selected using a variation on the TypIC approach described previously (Markey et al., 1999). The cells for each construct were ranked in order of their typicality (defined as the inverse of the Mahalanobis distance from the SLF features for that cell to the mean SLF values for that construct). Each multi-cell image was ranked in order of the average typicality of the cells in the image.

Results

CD-MPR colocalizes with GFP-UCE in HeLa cells

Using immunogold electron microscopy, we observed that the GFP-UCE fusion construct colocalizes with the CD-MPR in tubulovesicular regions of the TGN (Fig. 1A,C). When cell sections were labeled for UCE alone, occasional clathrin-coated buds and vesicles at the TGN were positive (Fig. 1B).

In addition, double-labeling of UCE and CD-MPR showed that the two proteins colocalized in a small fraction of clathrin-coated TGN vesicles and buds (Fig. 1C, arrowhead), suggesting that the proteins have the same exit pathway out of the TGN. However, UCE does not contain any known motifs for interaction with GGA proteins. Therefore, we sought to identify the TGN exit signal of UCE.

Intracellular location of GFP-UCE wild-type versus mutant forms

In order to determine the TGN exit signal in the cytoplasmic tail of the human UCE, GFP fusion constructs of the cytoplasmic tail of the UCE were made in which the lumenal domain of UCE was replaced by GFP. The localization of this fusion construct has been shown to be the same as that of the wild-type UCE in HeLa cells (Rohrer and Kornfeld, 2001). The cytoplasmic tail of UCE has been shown to contain a tyrosine-based internalization signal, residue Y^{488}HPL, which plays a role in rapid internalization from the plasma membrane (Rohrer and Kornfeld, 2001). Previous studies revealed that truncating the cytoplasmic tail at residue 486 resulted in a mutant enzyme trapped in an intracellular compartment with only 16% of the molecules being on the cell surface (Lee et al., 2002). This construct lacks an internalization signal but its low surface levels indicate that delivery to the surface is strongly reduced, possibly due to a slow exit out of the TGN. By contrast, the construct Y^{488}-A/M^{494} Stop, which also lacks a functional internalization signal, was found to be 50% at the cell surface. This suggests that it leaves the TGN faster than the Y^{486} Stop construct because it contains a TGN exit signal (Lee et al., 2002). In order to characterize a TGN exit signal we made an alanine scan of the residues Q^{492}EMNGEPLAA^{501} two at a time and analyzed the effect of the mutations on the localization of the GFP fusion constructs. All the constructs were made in the background of the Y^{488}-A mutation to prevent internalization if they would reach the plasma membrane. If a TGN exit signal was located in the scanned region, these constructs would be localized to the TGN and not at the surface, despite the Y488A internalization defect, as they would be expected to have a reduced rate of exit from the TGN. The rationale for screening the cytoplasmic tail up to residue 502 was based on the fact that signal sequences are usually about six to eight amino acids long, and furthermore, it was important to exclude redundant internalization signals such as the N^{511}PFKD^{515} signal in order to study the effect of the mutations in the background of an internalization defective UCE (Fig. 2A). Conventional approaches to interpret changes in subcellular distribution depend on the investigator's ability to differentiate localization patterns by visual examination of confocal immunofluorescence images. The inherent disadvantage of such a method is the lack of objectivity and inability to obtain quantitative data on the localization pattern. In order to obtain precise and quantitative data on the localization pattern, we used image interpretation methods that have been demonstrated to be capable of discriminating all major subcellular location patterns in 3D images with an accuracy greater than 98% (Chen and Murphy, 2004). These methods use sets of well-characterized numerical features (SLFs) that capture the essence of subcellular patterns while minimizing sensitivity to the position, rotation, size and shape of each cell. Most importantly for our purposes, these features have been demonstrated to be able to distinguish subtle differences in Golgi protein patterns that cannot be distinguished by visual examination (Murphy et al., 2003).

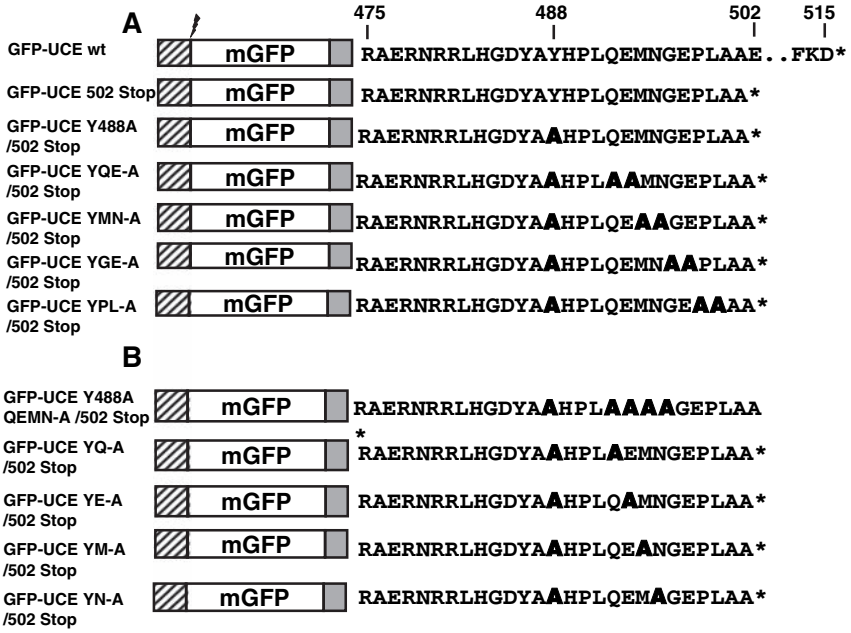

Fig. 2. (A) Schematic illustration of the wild-type and mutant GFP-UCE constructs. The lumenal domain of UCE was replaced with monomeric GFP fused to an N-terminal cleavable preprolactin leader sequence represented by hatched bars. The transmembrane domain of UCE is indicated by shaded bars and the amino acids in the cytoplasmic tail are shown in single letter code. Mutations of the Y^{488} residue and the potential exit signal residues are shown in bold letters. The position of a stop codon is indicated by an asterisk. (B) Schematic illustration of the single and combined mutant GFP-UCE constructs. The constructs have the same organization as those in Fig. 1A. In addition to the Y^{488} and the 502 Stop mutation, residues Q^{492}, E^{493}, M^{494} and N^{494} are either singly mutated or mutated simultaneously and the mutant residues are indicated by bold letters.

Using these features, we studied the relationship among UCE constructs by dendrogram analysis, in which wild-type and mutant constructs were grouped according to their relative similarity. The 'distance' (or inverse of similarity) between the images of each pair of constructs was calculated using feature set SLF20 as described in the Materials and Methods. The distance was expressed as a Mahalanobis distance, which measures the distance weighted by the statistical variation in each feature and any correlation between features. A clear pattern that emerged from the analysis was the similarity in localization of the wild-type construct and that of the 502 Stop construct, clearly indicating that any information for TGN exit is contained before residue 502 of the cytoplasmic tail. All immunofluorescence

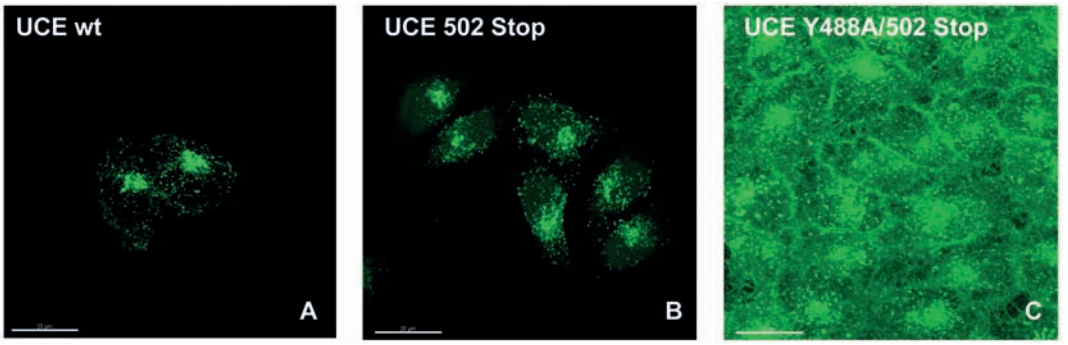

Fig. 3. High-resolution confocal immunofluorescence images of HeLa cells expressing wild-type or mutant GFP-UCE constructs. HeLa cells were stably transfected with wild-type GFP-UCE (A), GFP-UCE 502 Stop mutant (B) or GFP-UCE Y^{488}A/502 Stop mutant (C). The localization of the constructs was analyzed using high resolution three-dimensional confocal immunofluorescence microscopy followed by projection of z-stacks using the Imaris 4.0 software. Bars, 20 μm.

images were additionally ranked in order of their typicality (see Materials and Methods) and the most typical picture for each construct is shown (Fig. 3). Both constructs were localized to the TGN and could be co-labeled with a known TGN marker protein TGN46 (data not shown). Furthermore, as was clearly discernible from a visual examination of the images, the Y^{488}-A/502 Stop construct was to a large extent localized to the cell surface (Fig. 3). The image clustering program separated the different constructs analyzed into three groups – one comprising the wild-type UCE cytosolic tail and the 502 Stop construct, one comprising the Y^{488}; Q^{492}E-A/502 Stop and the Y^{488}; M^{494}N-A/502 Stop constructs, and the last comprising the Y^{488}-A/502 Stop ,Y^{488}; G^{496}E-A/502 Stop and Y^{488}; P^{498}L-A/502 Stop constructs (Fig. 4A). The surface localization of the Y^{488}; G^{496}E-A/502 Stop and Y^{488}; P^{498}L-A/502 Stop constructs (similar to Y^{488}-A/502 Stop construct) suggests that those mutations do not interfere with the TGN exit signal (allowing the UCE to exit the TGN and then become trapped at the cell surface). The intermediate phenotype of the Y^{488}; Q^{492}E-A/502 Stop and Y^{488}; M^{494}N-A/502 Stop constructs indicates that those mutations have compromised the TGN exit signal (Fig. 5).

Analysis of single and combined mutations of the Q^{492}EMN sequence

In order to further characterize the signal and to determine the contribution of the individual residues we made single and combined alanine substitution mutants of the Q^{492}, E^{493}, M^{494} and N^{495} residues (all in the Y^{488}-A/502 Stop background) and analyzed the localization of the constructs by dendrogram analysis as had been done for the double mutants. The dendrogram (Fig. 4B) contained the previous three groups plus a new group consisting of Y^{488}; M^{494}-A/502 Stop, Y^{488}; N^{493}-A/502 Stop, and Y^{488}; Q^{492}EMN-A/502 Stop. This group has a pattern resembling endosomes, suggesting that those constructs can exit the TGN but then may be altered so that

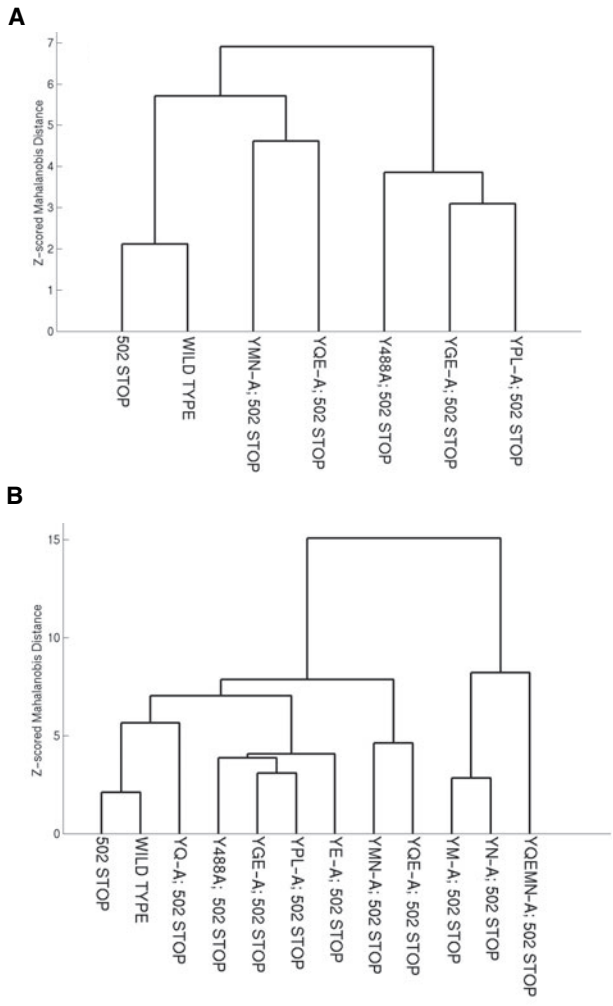

Fig. 4. Dendrograms showing similarity in localization between (A) the GFP-UCE wild type, GFP-UCE 502 Stop, GFP-UCE Y^{488}A/502 Stop and double mutant constructs of the Q^{492}EMNGEPL sequence in the cytoplasmic tail of UCE, and (B) the wild-type and all 11 mutant constructs. The constructs are clustered based on Mahalonobis distance calculated on the SLF20 set of subcellular localization features. The similarity in pattern between any two constructs is inversely proportional to the vertical distance along the path connecting them. For example, wild-type and 502 Stop are the most similar pair of constructs. These dendrograms are available at http://murphylab.web.cmu.edu/services/PSLID/HeLa_UCE/ with a browser interface that allows the underlying images for any branch to be displayed interactively.

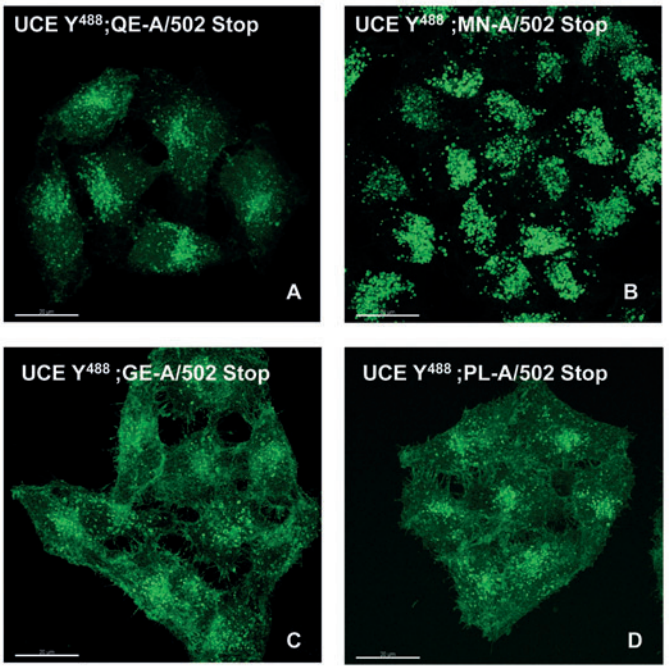

Fig. 5. High resolution confocal immunofluorescence images of HeLa cells expressing double mutant GFP-UCE constructs of the potential TGN exit signal. HeLa cells were stably transfected with GFP-UCE Y^{488}; Q^{492}E-A/502 Stop (A), GFP-UCE Y^{488}; M^{494}N-A/502 Stop (B), GFP-UCE Y^{488}; G^{496}E-A/502 Stop (C) or GFP-UCE Y^{488}; P^{498}L-A/502 Stop (D). Bars, 20 μm.

they are trapped elsewhere. The most important finding is the placement of Y^{488}; Q^{492}-A/502 Stop within the wild-type group, showing that this residue is the most important part of the TGN exit signal. The grouping of Y^{488}; E^{493}-A/502 Stop with the Y^{488}-A/502 Stop group indicates that E^{493} is not required for the exit signal. Combining results from the single and double mutants leads to the conclusion that the exit signal resides in the Q^{492}EMN residues with Q^{492} being most important but the M^{494}N residues being able to confer a partial exit signal even when Q^{492} is mutated. Our own visual observation of the localization of the constructs confirmed the image analysis results shown by the representative images of the single and combined mutants (Fig. 6).

Discussion

The results presented in this study demonstrate that the residues Q^{492}EMN in the cytoplasmic tail of UCE direct its exit from the TGN. Furthermore, the single alanine substitutions revealed that the residue Q^{492} is more important to the TGN exit function than residues M^{494} and N^{495}, which play a contributory role in the exit of UCE from the TGN. UCE has been shown to reside in the TGN at steady state and it constitutively cycles between the TGN and the plasma membrane (Rohrer and Kornfeld, 2001). UCE is synthesized as an inactive proenzyme and is activated by furin in the TGN (Do et al., 2002). It has been previously shown that UCE is internalized rapidly from the plasma membrane with a half-time of 0.65 minutes, using a tyrosine based internalization signal, Y^{488}HPL, by means of clathrin-mediated endocytosis (Lee et al., 2002). The trafficking itinerary of proteins that cycle between the TGN and the plasma membrane is either direct or involves an intermediate endosomal step (Lemmon and Traub, 2000). Proteins that reach the recycling or sorting endosomes are retrieved by transport to the TGN before reaching the surface or go directly to the plasma membrane. The steady state localization of UCE at the TGN thus mainly depends on the rate at which the enzyme exits the TGN, given the fact that it is rapidly internalized from the surface.

The localization and cycling pathway of UCE is important for the correct transport of lysosomal enzymes to the lysosomes in view of its function in the generation of the M6P recognition tag, a process that is completed in the TGN. The TGN exit step of its trafficking pathway is thus crucial to the process of lysosomal biogenesis. Using immunogold electron microscopy we found that a GFP-UCE fusion construct colocalizes with the CD-MPR at the TGN. This raised the question of whether the two proteins might have similar pathways of exit from the TGN. While the exit of the CD-MPR from the TGN has been extensively characterized, little is known about the TGN exit step in the trafficking itinerary of UCE. Using high resolution confocal immunofluorescence microscopy and automated image analysis we could show that the residues

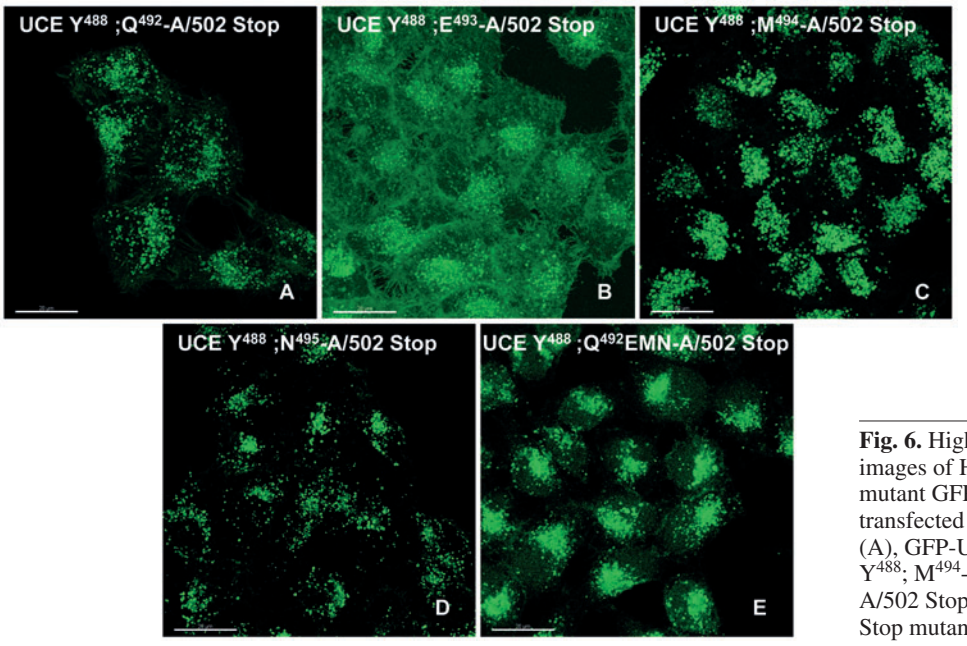

Fig. 6. High-resolution confocal immunofluorescence images of HeLa cells expressing combined and single mutant GFP-UCE constructs. HeLa cells were stably transfected with GFP-UCE Y^{488}; Q^{492}-A/502 Stop (A), GFP-UCE Y^{488}; E^{493}-A/502 Stop (B), GFP-UCE Y^{488}; M^{494}-A/502 Stop (C), GFP-UCE Y^{488}; N^{495}-A/502 Stop (D) or GFP-UCE Y^{488}; Q^{492}EMN-A/502 Stop mutant (E). Bars, 20 μm.

Q^{492}EMN are involved in the exit of UCE from the TGN. The GFP-UCE transmembrane domain-cytoplasmic tail fusion construct has been previously identified at the TGN (Rohrer and Kornfeld, 2001). Truncating the construct at position 502 did not affect the TGN localization of the enzyme, however, an additional Y^{488}-A mutation led to an accumulation of the construct at the surface. Therefore, any construct with the Y^{488}-A mutation which can exit the TGN (and is not otherwise perturbed) must accumulate at the surface. The images made at high resolution were analyzed using feature-based machine learning techniques. The advantages of using these approaches to analyze the images are that (i) it is possible to obtain quantitative data on the localization pattern and (ii) the resolution achieved by the program far outweighs that obtained by visual examination of images as previously demonstrated (Murphy et al., 2003). The Y^{488}; Q^{492}E-A/502 Stop and Y^{488}; M^{494}N-A/502 Stop constructs were mainly localized to the TGN as opposed to the Y^{488}; G^{496}E-A/502 Stop and Y^{488}; P^{498}L-A/502 Stop constructs, which were mainly localized to the plasma membrane. This indicates the role of the residues Q^{492}EMN in exit from the TGN. The localization of the Y488A/502 Stop construct was interesting in that although most of the construct was localized at the surface, a significant amount was still found in the trans-Golgi/TGN area, indicating that part of the Y^{488}HPL motif might also be involved in the TGN exit step of the enzyme trafficking pathway. Thus, in combination with the residues Q^{492}EMN, the Y^{488}HPL sequence functions in ensuring the proper cycling of UCE between the TGN and the plasma membrane.

To determine the relative contribution of each individual residue of the Q^{492}EMN TGN exit signal, single and combined mutants were made and their localization was analyzed using the image clustering program. The closeness of the mutant Y^{488}; Q^{492}-A/502 Stop to the wild-type suggested that Q^{492} plays an important role in exit from the TGN. A single mutation of Q^{492} is sufficient to block the exit function. By contrast, the mutant Y^{488}; E^{493}A/502 Stop showed a similar pattern to the mutant Y488A/502 Stop, suggesting that mutation of E^{493} does not block exit from the TGN. Surprisingly, the Y^{488}; M^{494}-A/502 Stop and Y^{488}; N^{495}-A/502 Stop mutants have a localization pattern that is discrete from that of the wild-type and Y488A/502 Stop mutant. Inspection of the images of these mutants revealed that they may form an endosome pattern, which suggests that the protein exits the TGN but does not enter the plasma membrane properly. It suggests that the protein could enter the plasma membrane properly when both M^{494} and N^{495} residues are intact, while either one of the two residues is sufficient in assisting Q^{492} in the exit from the TGN. The combined mutant had a mixture of Golgi and endosome patterns, which suggests that both the exit from the TGN and plasma membrane delivery are impaired.

Membrane proteins such as the MPRs and Lamp1 exit the TGN for onward transport to the late endosomes or the plasma membrane either by default or using a signal sequence that interacts specifically with adaptor proteins. The rate of exit of UCE from TGN and therefore the rate of its surface delivery, is reduced in the absence of its TGN exit signal. This leads to the conclusion that UCE requires the TGN exit signal for optimal trafficking out of the TGN. The nature of its signal sequence, however, does not resemble the canonical motifs found in other proteins such as the acidic cluster di-leucine motif found in MPRs or the tyrosine-based motifs present in Lamp-1, leading to the speculation that a novel interacting protein might act in the packaging of UCE at the TGN.

This work was supported in part by a grant from the Swiss National Science Foundation (no. 24.7184) to J.R. and by a grant from the United States National Institutes of Health (GM068845) to R.F.M.. We would like to thank Eric G. Berger for the critical reading of the manuscript and for his continuous support.

References

Alconada, A., Bauer, U. and Hoflack, B. (1996). A tyrosine-based motif and a casein kinase II phosphorylation site regulate the intracellular trafficking of the varicella-zoster virus glycoprotein I, a protein localized in the trans-Golgi network. *EMBO J.* **15**, 6096-6110.

Chen, X. and Murphy, R. F. (2004). Robust Classification of Subcellular Location Patterns in High Resolution 3D Fluorescence Microscope Images. In *Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1632-1635. San Francisco, CA.

Chen, X., Velliste, M., Weinstein, S., Jarvik, J. W. and Murphy, R. F. (2003). Location proteomics – Building subcellular location trees from high resolution 3D fluorescence microscope images of randomly-tagged proteins. *Proc. SPIE The International Society for Optical Engineering* **4962**, 298-306.

Do, H., Lee, W. S., Ghosh, P., Hollowell, T., Canfield, W. and Kornfeld, S. (2002). Human mannose 6-phosphate-uncovering enzyme is synthesized as a proenzyme that is activated by the endoprotease furin. *J. Biol. Chem.* **277**, 29737-29744.

Eng, F. J., Varlamov, O. and Fricker, L. D. (1999). Sequences within the cytoplasmic domain of gp180/carboxypeptidase D mediate localization to the trans-Golgi network. *Mol. Biol. Cell* **10**, 35-46.

Geuze, H. J., Slot, J. W., van der Ley, P. A. and Scheffer, R. C. T. (1981). Use of colloidal gold particles in double labeling immunoelectron microscopy on ultrathin frozen tissue sections. *J. Cell Biol.* **89**, 635-665.

Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. and Pease, L. R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51-59.

Huang, K. and Murphy, R. F. (2004). From quantitative microscopy to automated image understanding. *J. Biomed. Opt.* **9**, 893-912.

Kornfeld, R., Bao, M., Brewer, K., Noll, C. and Canfield, W. (1999). Molecular cloning and functional expression of two splice forms of human N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosaminidase. *J. Biol. Chem.* **274**, 32778-32785.

Lee, W. S., Rohrer, J., Kornfeld, R. and Kornfeld, S. (2002). Multiple signals regulate trafficking of the mannose 6-phosphate-uncovering enzyme. *J. Biol. Chem.* **277**, 3544-3551.

Lemmon, S. K. and Traub, L. M. (2000). Sorting in the endosomal system in yeast and animal cells. *Curr. Opin. Cell Biol.* **12**, 457-466.

Lotufo, R. and Falcao, A. (2000). The ordered queue and the optimality of the watershed approaches. In *Mathematical Morphology and its Application to Image and Signal Processing*, vol. 12 (ed. J. Goutsias, L. Vincent and D. S. Bloomberg). Boston, MA: Kluwer Academic Publishers.

Markey, M. K., Boland, M. V. and Murphy, R. F. (1999). Toward objective selection of representative microscope images. *Biophys. J.* **76**, 2230-2237.

Murphy, R. F., Velliste, M. and Porreca, G. (2003). Robust numerical features for description and classification of subcellular location patterns in fluorescent microscope images. *J. VLSI Signal Proc.* **35**, 311-321.

Petris, M. J., Camakaris, J., Greenough, M., LaFontaine, S. and Mercer, J. F. (1998). A C-terminal di-leucine is required for localization of the Menkes protein in the trans-Golgi network. *Hum. Mol. Genet.* **7**, 2063-2071.

Raposo, G., Kleijmeer, M. J., Posthuma, G., Slot, J. W. and Geuze, H. J. (1997). Immunogold labeling of ultrathin cryosections: Application in immunology. In *Handbook of Experimental Immunology*. Vol. 4, 5th edn (ed. L. A. Herzenberg, D. M. Weir and C. Blackwell), pp. 1-11. Malden, UDA, Blackwell Science.

Rohrer, J. and Kornfeld, R. (2001). Lysosomal hydrolase mannose 6-phosphate uncovering enzyme resides in the trans-Golgi network. *Mol. Biol. Cell* **12**, 1623-1631.

Rohrer, J., Schweizer, A., Johnson, K. F. and Kornfeld, S. (1995). A

determinant in the cytoplasmic tail of the cation-dependent mannose 6-phosphate receptor prevents trafficking to lysosomes. *J. Cell Biol.* **130**, 1297-1306.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1998). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Schafer, W., Stroh, A., Berghofer, S., Seiler, J., Vey, M., Kruse, M. L., Kern, H. F., Klenk, H. D. and Garten, W. (1995). Two independent targeting signals in the cytoplasmic domain determine trans-Golgi network localization and endosomal trafficking of the proprotein convertase furin. *EMBO J.* **14**, 2424-2435.

Teuchert, M., Schafer, W., Berghofer, S., Hoflack, B., Klenk, H. D. and Garten, W. (1999). Sorting of furin at the trans-Golgi network. Interaction of the cytoplasmic tail sorting signals with AP-1 Golgi-specific assembly proteins. *J. Biol. Chem.* **274**, 8199-8207.

Velliste, M. and Murphy, R. F. (2002). Automated determination of protein subcellular locations from 3D fluorescence microscope images. In *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 867-870. Bethesda, MD, USA.

Wong, S. H. and Hong, W. (1993). The SXYQRL sequence in the cytoplasmic domain of TGN38 plays a major role in trans-Golgi network localization. *J. Biol. Chem.* **268**, 22853-22862.